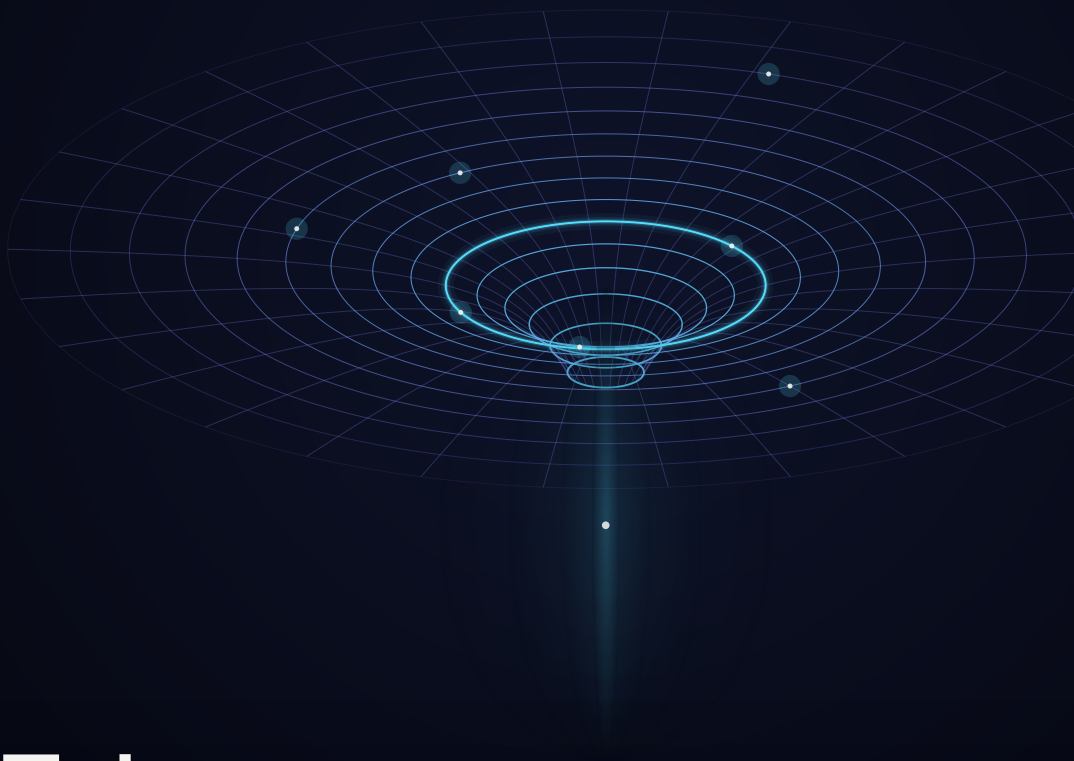
DE LA REALIDAD A LA ABSTRACCIÓN



ÁNGEL ZELEDÓN

Curso completo de matemática universitaria
Del cálculo a la teoría de categorías

2026

Índice general

Índice de figuras **xxxvii**

Cómo usar esta guía **xl**

I Fundamentos: lógica, números y estructuras discretas **1**

1. Conjuntos, lógica y cómo se demuestra algo **2**

1.1. Conjuntos: el material de construcción . . . 3
1.2. Las cinco conectivas y la negación de enunciados . . . 4
1.3. Lógica: proposiciones y conectivos . . . 11
1.4. Tautologías, argumentos válidos y falacias . . . 12
1.5. Cuantificadores . . . 14
1.6. Relaciones: equivalencia, particiones y orden . . . 15
1.7. Los métodos de demostración . . . 18
1.8. Cómo se lee y se escribe una demostración . . . 20
1.9. Contar lo infinito: cardinalidad y el argumento diagonal . . . 20
1.10. Errores típicos de lógica . . . 23
1.11. Funciones, familias e infinitos . . . 23
1.12. Los tres principios de inducción . . . 25
1.13. Diagramas de Hasse: dibujar un orden sin dibujarlo entero . . . 26
1.14. Ejercicios resueltos . . . 28
1.15. Ejercicios propuestos . . . 29

2. Lógica matemática y teoría de conjuntos **30**

2.1. Sintaxis, semántica y completitud . . . 30
2.2. Del alfabeto al teorema: qué es demostrar . . . 31
2.3. Incompletitud . . . 36
2.4. Computabilidad: lo que ninguna máquina puede decidir . . . 38
2.5. Computabilidad: qué se puede calcular y qué no . . . 42
2.6. Teoría de conjuntos: ZFC, ordinales y elección . . . 43
2.7. Complejidad: P, NP y la NP-completitud . . . 45
2.8. Tarski, el segundo teorema y modelos no estándar . . . 47

2.9. Estructuras, satisfacción y conectivos completos . . . 49
2.10. Ejercicios resueltos . . . 53
2.11. Ejercicios propuestos . . . 54

3. Teoría de números: de la divisibilidad a la criptografía 56
3.1. Divisibilidad y el algoritmo de Euclides . . . 56
3.1.1. El máximo común divisor . . . 58
3.1.2. La identidad de Bézout: el regalo escondido del algoritmo . . . 60
3.2. Cimientos: divisibilidad, congruencias y clases de restos . . . 62
3.2.1. La relación de divisibilidad . . . 63
3.2.2. Parte entera y división euclídea . . . 65
3.2.3. Primos y compuestos . . . 67
3.2.4. Máximo común divisor y mínimo común múltiplo . . . 68
3.2.5. Factorización única y valuación p-ádica . . . 72
3.2.6. Congruencias . . . 73
3.2.7. Clases de restos, el conjunto $\mathbb{Z}/n\mathbb{Z}$ y las unidades . . . 76
3.2.8. La ecuación diofántica lineal . . . 78
3.2.9. Los criterios de divisibilidad, y de dónde salen . . . 80
3.3. Los números primos y el teorema . . . 82
3.3.1. Hay infinitos primos . . . 82
3.3.2. El lema de Euclides: la propiedad que hace especiales a los primos 83
3.3.3. El teorema fundamental de la aritmética . . . 84
3.4. Aritmética modular: las matemáticas del reloj . . . 89
3.4.1. Dividir es más delicado: inversos módulo m . . . 91
3.4.2. Las tres reglas de oro . . . 94
3.5. Los criterios de divisibilidad, por fin demostrados . . . 96
3.5.1. La prueba del nueve . . . 97
3.6. Calendarios, ciclos y relojes de todos los tamaños . . . 97
3.7. Dígitos de control . . . 98
3.8. Ecuaciones de congruencia . . . 99
3.8.1. El teorema chino del resto, en versión práctica . . . 100
3.9. El teorema chino del resto . . . 100
3.9.1. Resolviéndolo a mano . . . 101
3.10. Los teoremas de Fermat y Euler . . . 106
3.10.1. El pequeño teorema de Fermat . . . 106
3.10.2. Euler generaliza a cualquier módulo . . . 108
3.10.3. La función φ de Euler y el teorema que generaliza a Fermat . . . 115
3.11. Orden, raíces primitivas y el intercambio . . . 117
3.12. Residuos cuadráticos y el símbolo de Legendre . . . 119
3.13. Criptografía clásica: sustituir letras . . . 121
3.13.1. El cifrado de César . . . 121
3.13.2. El cifrado afín . . . 122

3.13.3. Vigenère: una clave que se mueve . . . 122
3.14. RSA: cómo mandar un secreto, paso a paso . . . 123
3.14.1. La idea . . . 123
3.14.2. Preparando las llaves (esto lo hace quien recibe) . . . 123
3.14.3. Cifrando (esto lo hace quien envía) . . . 124
3.14.4. Descifrando (solo puedo yo) . . . 124
3.14.5. ¿Por qué funciona? Aquí está la magia, y es pura teoría del capítulo 124
3.15. Una joya para cerrar: ¿qué primos son suma de dos cuadrados? . . . 125
3.16. Errores típicos de aritmética modular . . . 127
3.17. Divisores, Wilson y fracciones continuas . . . 127
3.18. Ejercicios resueltos . . . 134
3.19. Ejercicios propuestos . . . 135

4. Números complejos desde el principio 138
4.1. Por qué hicieron falta: las cúbicas de Cardano . . . 139
4.2. Los complejos desde cero: aritmética, plano, forma polar y exponencial . 139
4.3. Definición y forma binómica . . . 147
4.4. Aritmética: sumar, multiplicar, dividir . . . 148
4.5. El plano de Argand . . . 150
4.6. Forma polar: módulo y argumento . . . 150
4.7. Lugares geométricos en el plano complejo . . . 154
4.8. De Moivre y las raíces n-ésimas . . . 156
4.9. Cuadráticas con discriminante negativo . . . 159
4.10. Anticipo: la fórmula de Euler . . . 161
4.11. El logaritmo complejo y las potencias . . . 163
4.12. Transformaciones del plano y la esfera de Riemann . . . 165
4.13. Errores típicos con complejos . . . 167
4.14. Trigonométricas e hiperbólicas complejas . . . 168
4.15. Series complejas y geometría con complejos . . . 169
4.16. Fasores e impedancia . . . 171
4.17. La raíz cuadrada de un complejo y la cuadrática general . . . 172
4.18. El módulo y el conjugado, demostrados . . . 176
4.19. Ejercicios resueltos . . . 179
4.20. Ejercicios propuestos . . . 180

5. Combinatoria y probabilidad 181
5.1. Los dos principios fundamentales . . . 181
5.2. Cómo decidir qué clase de conteo pide un problema . . . 183
5.3. Permutaciones: cuando el orden importa . . . 191
5.4. Combinaciones: cuando el orden no importa . . . 192
5.5. El binomio de Newton y el triángulo de Pascal . . . 194
5.6. Inclusión–exclusión y el principio del palomar . . . 195
5.7. Probabilidad: el modelo de Laplace . . . 196

5.8. Probabilidad condicionada, independencia y Bayes . . . 198
5.9. Ensayos repetidos: la distribución binomial . . . 199
5.10. Errores típicos de combinatoria y probabilidad . . . 200
5.11. Funciones generatrices: contar con series . . . 201
5.12. Combinaciones con repetición e inclusión–exclusión . . . 203
5.13. Varianza y la ley de los grandes números . . . 205
5.14. Variables aleatorias e independencia . . . 206
5.15. Ejercicios resueltos . . . 208
5.16. Ejercicios propuestos . . . 209

II Cálculo y álgebra lineal 211

6. Cálculo en una variable 212
6.1. Los números reales y la completitud . . . 212
6.2. Límites laterales y clasificación de las discontinuidades . . . 214
6.3. La derivada de la función inversa, demostrada . . . 216
6.4. Primitiva, integral indefinida y las dos técnicas demostradas . . . 218
6.5. Las dos partes del teorema fundamental del cálculo . . . 221
6.6. Las hipótesis de la regla de L'Hôpital . . . 223
6.7. Sucesiones . . . 225
6.8. Límites de funciones y continuidad . . . 227
6.9. La derivada . . . 232
6.9.1. Las reglas, con demostración . . . 234
6.9.2. Los teoremas del valor medio . . . 235
6.9.3. Convexidad y trazado de gráficas . . . 238
6.9.4. El teorema de Taylor . . . 239
6.10. La integral de Riemann . . . 240
6.10.1. Técnicas de integración . . . 243
6.10.2. Integrales impropias . . . 243
6.11. Series numéricas . . . 244
6.12. Series de potencias y funciones analíticas . . . 247
6.13. Fundamentos: la construcción de los números reales . . . 248
6.14. Demostraciones completas de los teoremas de compacidad . . . 250
6.15. Sucesiones y series de funciones: convergencia uniforme . . . 251
6.16. Complementos: hiperbólicas, complejos y técnicas finas . . . 255
6.16.1. Funciones hiperbólicas . . . 255
6.16.2. Números complejos y la fórmula de Euler, demostrada . . . 255
6.16.3. Derivación bajo el signo integral . . . 255
6.16.4. Criterio de Dirichlet para integrales oscilantes . . . 256
6.16.5. La irracionalidad de e . . . 256
6.17. Cierre del temario I: aplicaciones de la derivada . . . 256
6.17.1. Asíntotas y trazado completo de curvas . . . 256

6.17.2. Tasas relacionadas . . . 256
6.17.3. Diferenciales, aproximación lineal y método de Newton . . . 257
6.18. Cierre del temario II: aplicaciones de la integral . . . 257
6.18.1. Áreas y volúmenes . . . 257
6.18.2. Longitud de arco y superficies de revolución . . . 257
6.18.3. Valor medio y trabajo . . . 258
6.18.4. Potencias trigonométricas y estrategia general . . . 258
6.19. Profundización final: herramientas finas del análisis . . . 258
6.19.1. Límite superior e inferior . . . 258
6.19.2. Condensación de Cauchy . . . 258
6.19.3. Criterios de Abel y Dirichlet para series . . . 259
6.19.4. Fracciones continuas . . . 259
6.19.5. Euler–Maclaurin . . . 259
6.20. Última capa: los teoremas que faltaban . . . 259
6.20.1. El teorema de Abel del borde . . . 259
6.20.2. Criterios de Raabe y Gauss . . . 260
6.20.3. Productos infinitos . . . 260
6.20.4. Sumabilidad de series divergentes . . . 260
6.20.5. La exponencial construida desde cero . . . 260
6.20.6. El teorema de aproximación de Weierstrass . . . 260
6.21. Ronda de temarios: cierre curricular de Cálculo . . . 261
6.21.1. Las tres formas del resto de Taylor . . . 261
6.21.2. La irracionalidad de π (Niven, 1947) . . . 262
6.21.3. Trapecio y Simpson con cotas de error demostradas . . . 263
6.22. Cierre definitivo: Beta y Stirling demostrado . . . 264
6.22.1. La función Beta . . . 264
6.22.2. La fórmula de Stirling, demostrada . . . 265
6.23. Segunda ronda: más teoría y ejemplos a fondo . . . 266
6.23.1. Integrales de Frullani . . . 266
6.23.2. La media aritmético-geométrica de Gauss . . . 266
6.23.3. Ejemplos finos . . . 266
6.24. Coda: Stieltjes y la variación acotada . . . 267
6.25. Temas extra y ejemplos desarrollados . . . 267
6.25.1. Sumas de Riemann como límites . . . 267
6.25.2. Trapecio y Simpson . . . 267
6.25.3. La ley de Snell por Fermat . . . 267
6.25.4. Producto de Cauchy . . . 267
6.25.5. Wallis y Stirling . . . 267
6.26. Seno, coseno y π desde la serie . . . 268
6.27. Bonnet, medida nula y Stieltjes . . . 269
6.28. Los criterios de convergencia, demostrados . . . 271
6.29. Radio de convergencia y los extremos . . . 275

6.30. Ejercicios resueltos . . . 278
6.31. Ejercicios adicionales resueltos . . . 280
6.32. Taller de ejemplos explicados: ¿converge o no una serie? . . . 281
6.32.1. Series geométricas: las más fáciles . . . 281
6.32.2. Series-p: la regla del exponente . . . 281
6.32.3. El criterio del cociente (el más útil) . . . 282
6.33. Ampliación del temario . . . 282
6.33.1. Regla de L'Hôpital y las formas indeterminadas . . . 282
6.33.2. Técnicas de integración, en detalle . . . 283
6.33.3. Curvas paramétricas . . . 285
6.33.4. Coordenadas polares . . . 285
6.34. Taller de integrales impropias . . . 286
6.35. Taller de indeterminaciones . . . 288
6.36. Centroides y los teoremas de Pappus . . . 291
6.37. Ejercicios propuestos . . . 295

7. Álgebra lineal 296
7.1. Sistemas lineales y eliminación de Gauss . . . 296
7.1.1. Factorización LU y coste computacional . . . 298
7.1.2. Matrices y sus operaciones . . . 298
7.2. Matrices: qué son y cómo se operan . . . 301
7.2.1. Qué es exactamente una matriz . . . 301
7.2.2. Un catálogo de matrices con nombre . . . 303
7.2.3. Sumar matrices y multiplicarlas por un número . . . 304
7.2.4. La traspuesta . . . 305
7.2.5. Por qué el producto de matrices es como es . . . 307
7.2.6. Potencias de una matriz cuadrada . . . 310
7.2.7. Un sistema de ecuaciones escrito como $A\mathbf{x} = \mathbf{b}$ . . . 311
7.2.8. Operaciones de fila, forma escalonada, pivotes y rango . . . 313
7.3. Espacios vectoriales . . . 317
7.3.1. Independencia, bases, dimensión . . . 318
7.3.2. Coordenadas y cambio de base . . . 318
7.4. Transformaciones lineales . . . 321
7.5. Determinantes . . . 326
7.6. Valores y vectores propios . . . 332
7.6.1. Cuando no se puede diagonalizar: Jordan . . . 335
7.7. Producto interno y ortogonalidad . . . 338
7.7.1. Gram–Schmidt, proyecciones y mínimos cuadrados . . . 339
7.7.2. Matrices ortogonales y el teorema espectral . . . 342
7.8. Formas cuadráticas . . . 343
7.9. Descomposición en valores singulares (SVD) . . . 346
7.10. Fundamentos: las demostraciones estructurales . . . 349

7.10.1. Existencia de bases y el lema de intercambio 349
7.10.2. El determinante: existencia y unicidad 351
7.10.3. Polinomio mínimo y descomposición primaria 351
7.10.4. Triangularización de Schur y demostración de Cayley–Hamilton 353
7.10.5. Existencia de la SVD, con demostración 355
7.10.6. El espacio dual 356
7.11. Complementos: bloques, Cholesky, normas y pseudoinversa 357
7.11.1. Matrices por bloques 357
7.11.2. Cholesky 357
7.11.3. Normas matriciales y condicionamiento 358
7.11.4. Pseudoinversa de Moore–Penrose 359
7.12. Cierre del temario: los subtemas restantes 359
7.12.1. Matrices elementales 359
7.12.2. Los cuatro subespacios fundamentales 359
7.12.3. Espacios complejos: hermitianas, unitarias, normales 359
7.12.4. Formas bilineales y congruencia 360
7.12.5. Matrices y grafos 360
7.13. Profundización final: teoría espectral fina 360
7.13.1. Espacios cociente 360
7.13.2. Desigualdades de rango 360
7.13.3. Funciones de matrices 360
7.13.4. Courant–Fischer y entrelazado 360
7.13.5. Perron–Frobenius e iteración de potencias 361
7.13.6. Definidas positivas: las siete caracterizaciones 361
7.14. Última capa: formas canónicas y transformadas 361
7.14.1. La forma de Jordan, demostrada 361
7.14.2. Forma canónica racional 361
7.14.3. Descomposición polar 361
7.14.4. Cauchy–Binet 362
7.14.5. Circulantes y la transformada discreta de Fourier 362
7.15. Ronda de temarios: cociente, bidual y anuladores 362
7.15.1. El espacio cociente y el primer teorema de isomorfismo 362
7.15.2. El bidual, o por qué “canónico” es una palabra técnica 364
7.15.3. Anuladores: la correspondencia exacta 365
7.16. Cierre definitivo: diagonalización simultánea 365
7.17. Segunda ronda 366
7.17.1. Reflexiones de Householder 366
7.17.2. Tridiagonales y el laplaciano discreto 366
7.17.3. La matriz de Hilbert 367
7.18. Coda: perturbación de autovalores 367
7.19. Temas extra y ejemplos desarrollados 367
7.19.1. Cadenas de Markov y PageRank 367

7.19.2. Vandermonde e interpolación . . . 367
7.19.3. Círculos de Gershgorin . . . 368
7.19.4. La matriz de proyección . . . 368
7.19.5. Rotaciones en 3D . . . 368
7.20. Eckart–Young y componentes principales . . . 368
7.21. Complexificación y forma real . . . 369
7.22. Kronecker, vec y ecuaciones matriciales . . . 370
7.23. Proyecciones, complementos y bases infinitas . . . 372
7.24. Las demostraciones pendientes: espectral, rangos y Jordan . . . 373
7.25. Condicionamiento, pseudoinversa y Gram–Schmidt, explícitos . . . 375
7.26. Ejercicios resueltos . . . 377
7.27. Ejercicios adicionales resueltos . . . 380
7.28. Taller de ejemplos explicados: sistemas y valores propios . . . 380
7.28.1. Resolver un sistema por eliminación de Gauss . . . 381
7.28.2. Valores y vectores propios de una matriz 2×2 . . . 381
7.29. Ampliación del temario . . . 382
7.29.1. Geometría analítica en $\mathbb{R}^2$ y $\mathbb{R}^3$ . . . 382
7.29.2. Espacios afines: puntos frente a vectores . . . 383
7.29.3. Optimización y programación lineal: el método Simplex . . . 383
7.29.4. Producto tensorial y álgebra multilineal . . . 384
7.30. Ejercicios propuestos . . . 384

8. Grafos: la matemática de las conexiones 386
8.1. Qué es un grafo . . . 386
8.2. El vocabulario básico de los grafos . . . 387
8.3. Grado y el lema del apretón de manos . . . 396
8.4. Caminos, conexión y los puentes de Königsberg . . . 398
8.5. Recorrer un grafo: anchura y profundidad . . . 401
8.6. Ciclos hamiltonianos . . . 404
8.7. Árboles . . . 405
8.8. Grafos planos y la fórmula de Euler . . . 407
8.9. Grafos bipartitos y emparejamientos . . . 411
8.10. Coloración y el teorema de los cuatro colores . . . 415
8.11. Caminos mínimos: el algoritmo de Dijkstra . . . 418
8.12. Flujo máximo y corte mínimo . . . 421
8.13. Grafos dirigidos: orden topológico y camino crítico . . . 424
8.14. Redes reales: centralidad, PageRank y el mundo pequeño . . . 428
8.15. Conjuntos independientes, cliques y recubrimientos . . . 430
8.16. Errores típicos con grafos . . . 433
8.17. Isomorfismo de grafos . . . 434
8.18. Ejercicios resueltos . . . 436
8.19. Ejercicios propuestos . . . 437

9. Cálculo multivariable **439**
9.1. Topología de $\mathbb{R}^n$ . . . 439
9.1.1. Límites: el peligro de las trayectorias . . . 440
9.2. La derivada parcial: definición, notaciones y el aviso central . . . 442
9.3. Diferenciabilidad . . . 450
9.3.1. Taylor en varias variables . . . 452
9.4. Optimización . . . 455
9.4.1. Extremos libres . . . 455
9.4.2. Extremos condicionados: multiplicadores de Lagrange . . . 457
9.5. Optimización numérica: descenso por gradiente y Newton . . . 462
9.6. Integrales múltiples . . . 467
9.6.1. Aplicaciones físicas . . . 468
9.7. El área de una superficie parametrizada . . . 472
9.8. Derivación bajo el signo integral . . . 476
9.9. Funciones inversas e implícitas . . . 480
9.10. Fundamentos: las demostraciones del cálculo multivariable . . . 481
9.10.1. Igualdad de las parciales cruzadas . . . 481
9.10.2. Parciales continuas implican diferenciabilidad . . . 483
9.10.3. Taylor multivariable, con demostración . . . 483
9.10.4. El teorema de la función inversa, demostrado . . . 483
9.10.5. Lagrange, demostrado con la función implícita . . . 484
9.10.6. Fubini y el cambio de variables: por qué son ciertos . . . 484
9.10.7. Convexidad en varias variables . . . 485
9.11. Complementos: parámetros, cuádricas e impropias múltiples . . . 485
9.11.1. Integrales con parámetro y la integral de Dirichlet . . . 485
9.11.2. Cuádricas clasificadas . . . 485
9.11.3. Impropias múltiples . . . 486
9.11.4. Jensen . . . 486
9.12. Cierre del temario: los subtemas restantes . . . 486
9.12.1. Cinemática: componentes de la aceleración . . . 486
9.12.2. Límites iterados vs. doble . . . 486
9.12.3. Taylor con multi-índices . . . 486
9.12.4. Probabilidad con integrales dobles . . . 486
9.12.5. Sistemas implícitos . . . 487
9.13. Profundización final: teoremas selectos . . . 487
9.13.1. Funciones homogéneas y el teorema de Euler . . . 487
9.13.2. Taylor con resto integral . . . 487
9.13.3. Condiciones KKT . . . 487
9.13.4. Transporte de Reynolds . . . 487
9.14. Última capa: hessiana orlada, Morse y rango . . . 488
9.14.1. La hessiana orlada: el segundo orden de Lagrange . . . 488
9.14.2. El lema de Morse . . . 488

9.14.3. El teorema del rango . . . 488
9.15. Ronda de temarios: envolventes de familias de curvas . . . 488
9.16. Cierre definitivo . . . 490
9.17. Segunda ronda: gaussianas n-dimensionales y la bola . . . 491
9.17.1. La gaussiana multivariable . . . 491
9.17.2. El volumen de la bola n-dimensional . . . 492
9.17.3. Newton multivariable . . . 492
9.18. Coda: dualidad de Lagrange . . . 492
9.19. Temas extra y ejemplos desarrollados . . . 492
9.19.1. D'Alembert por cambio de variables . . . 492
9.19.2. La derivada material . . . 492
9.19.3. Descenso por gradiente en cuadráticas . . . 493
9.19.4. El teorema de la envolvente . . . 493
9.19.5. Curvatura de curvas implícitas . . . 493
9.20. Conjuntos medibles-Jordan . . . 493
9.21. El cambio de variables, demostrado . . . 494
9.22. Las tres demostraciones que sostienen el capítulo . . . 496
9.23. Dominio, conjuntos de nivel y continuidad de un campo escalar . . . 499
9.24. Weierstrass demostrado y los extremos absolutos sobre un compacto . . . 504
9.25. Ejercicios resueltos . . . 508
9.26. Ejercicios adicionales resueltos . . . 510
9.27. Taller de ejemplos explicados . . . 510
9.27.1. Derivadas parciales . . . 510
9.27.2. El gradiente . . . 511
9.27.3. Multiplicadores de Lagrange: un problema completo . . . 511
9.28. Ampliación del temario . . . 511
9.28.1. Integrales múltiples en coordenadas cilíndricas y esféricas . . . 511
9.28.2. Aplicaciones físicas de la integral múltiple . . . 512
9.29. Ejercicios propuestos . . . 513

10. Coordenadas polares y ecuaciones paramétricas 514
10.1. Dónde está un punto: dos respuestas posibles . . . 514
10.2. Polares, curvas polares y curvas paramétricas desde cero . . . 515
10.2.1. El nombre de un punto no es único . . . 516
10.2.2. Qué es una curva polar y qué es una simetría . . . 517
10.2.3. Curvas paramétricas: la pendiente y la concavidad . . . 520
10.2.4. La cicloide, deducida en lugar de anunciada . . . 522
10.2.5. Medir una curva paramétrica: longitud y superficie . . . 523
10.3. Ecuaciones sencillas en polares . . . 524
10.4. Curvas que solo existen cómodamente en polares . . . 527
10.4.1. La cardioide . . . 527
10.4.2. Las rosas . . . 528

10.4.3. La espiral de Arquímedes . . . 528
10.4.4. Las cónicas vistas desde el foco . . . 528
10.5. Dónde se cortan dos curvas polares . . . 531
10.6. Taller: exprimir la ecuación focal de una cónica . . . 534
10.7. Ecuaciones paramétricas: cuando hace falta el tiempo . . . 537
10.8. Tiro parabólico: la parametrización que sí importa . . . 538
10.9. La cicloide y sus dos milagros . . . 538
10.10. Velocidad, aceleración y curvatura . . . 542
10.11. La familia de las ruletas . . . 545
10.12. Áreas, longitudes y superficies de una curva paramétrica . . . 547
10.13. Polares en el espacio . . . 551
10.14. Errores típicos . . . 553
10.15. Tangentes en polares . . . 554
10.16. Frenet, torsión y la segunda ley de Kepler . . . 555
10.17. Cálculo de variaciones: la braquistócrona y la tautócrona, demostradas . . . 557
10.18. Curvatura y torsión: las fórmulas operativas, deducidas . . . 560
10.19. La ecuación focal de las cónicas, deducida . . . 562
10.20. Las fórmulas polares, demostradas, y la curvatura en polares . . . 563
10.21. Curvas cerradas, simples y regulares, con reparametrización por el arco . . . 567
10.22. Ejercicios resueltos . . . 571
10.23. Ejercicios propuestos . . . 571

11. Cálculo vectorial 573

11.1. Curvas y su geometría . . . 573
11.2. Cimientos: campos, curvas, superficies e integrales . . . 576
11.2.1. Campos escalares y campos vectoriales . . . 576
11.2.2. Curvas parametrizadas, vector tangente y elemento de longitud 578
11.2.3. Las dos integrales de línea, que son cosas distintas . . . 579
11.2.4. Independencia del camino, campos conservativos y función potencial 582
11.2.5. Superficies parametrizadas, vector normal y las dos integrales de superficie . . . 585
11.2.6. Los cuatro operadores diferenciales . . . 588
11.3. Campos vectoriales y operadores diferenciales . . . 592
11.4. Integrales de línea . . . 595
11.4.1. Campos conservativos . . . 598
11.5. Integrales de superficie . . . 605
11.6. Los tres grandes teoremas . . . 607
11.6.1. Aplicaciones físicas . . . 613
11.7. Formas diferenciales: el lenguaje unificado . . . 616
11.8. Fundamentos: la teoría detrás de los teoremas integrales . . . 622
11.8.1. Las fórmulas de Frenet, deducidas . . . 622
11.8.2. Existencia del potencial: demostración constructiva . . . 623

11.8.3. Demostración del teorema de Stokes . . . 624
11.8.4. Formas diferenciales con rigor . . . 624
11.8.5. Aplicación: unicidad del problema de Dirichlet . . . 626
11.9. Complementos: coordenadas curvilíneas, Helmholtz y Green . . . 627
11.9.1. Operadores en coordenadas curvilíneas ortogonales . . . 627
11.9.2. Helmholtz . . . 629
11.9.3. Identidades de Green y unicidad . . . 629
11.9.4. Ángulo sólido . . . 630
11.10. Cierre del temario: los subtemas restantes . . . 630
11.10.1. Forma normal de Green . . . 630
11.10.2. Deformación de contornos . . . 630
11.10.3. Parametrizaciones estándar . . . 630
11.10.4. Independencia de la parametrización . . . 630
11.11. Profundización final . . . 631
11.11.1. Potencial vector explícito . . . 631
11.11.2. El problema de Kepler completo . . . 631
11.11.3. Poincaré–Hopf . . . 631
11.11.4. Maxwell rematado: ondas y Poynting . . . 632
11.12. Última capa: Gauss demostrado, Biot–Savart y cohomología . . . 632
11.12.1. Demostración del teorema de la divergencia . . . 632
11.12.2. Biot–Savart . . . 632
11.12.3. El primer grupo de cohomología, calculado . . . 633
11.13. Ronda de temarios: Reynolds y el ángulo sólido . . . 634
11.13.1. Derivada material y teorema del transporte de Reynolds . . . 634
11.13.2. El ángulo sólido . . . 635
11.14. Cierre definitivo: las identidades de Green . . . 636
11.15. Segunda ronda: el planímetro, el dipolo y el generador . . . 639
11.15.1. El planímetro: Green hecho latón . . . 639
11.15.2. El dipolo, trabajado . . . 639
11.15.3. Faraday cuantitativo . . . 639
11.16. Coda: potenciales retardados . . . 639
11.17. Temas extra y ejemplos desarrollados . . . 640
11.17.1. Trabajo–energía y conservación . . . 640
11.17.2. Kepler II para toda fuerza central . . . 640
11.17.3. Deducción de la ecuación del calor . . . 640
11.17.4. Pappus–Guldin . . . 640
11.17.5. Vorticidad y Kelvin . . . 640
11.18. El flujo de un campo . . . 640
11.19. Invertir el laplaciano y Helmholtz . . . 641
11.20. Homotopía y campos conservativos: la implicación que faltaba . . . 643
11.21. Orientación, solenoidal y simplemente conexo . . . 645
11.21.1. Superficies orientables y orientación inducida en el borde . . . 645

11.21.2. Campos solenoidales . . . 646
11.21.3. Dominios simplemente conexos . . . 647
11.22. Ejercicios resueltos . . . 648
11.23. Ejercicios adicionales resueltos . . . 649
11.24. Taller de ejemplos explicados . . . 650
11.24.1. Una integral de línea . . . 650
11.24.2. El teorema de Green (convertir un contorno en un área) . . . 651
11.25. Ampliación del temario: introducción al cálculo tensorial . . . 651
11.26. Ejercicios propuestos . . . 652

12. Ecuaciones diferenciales 654
12.1. Ecuaciones de primer orden . . . 654
12.1.1. Separables . . . 655
12.1.2. Lineales: el factor integrante . . . 655
12.1.3. Exactas y factores integrantes . . . 656
12.1.4. Sustituciones clásicas . . . 656
12.1.5. Existencia, unicidad y el método de Picard . . . 658
12.2. Qué es una ecuación diferencial . . . 662
12.2.1. La incógnita es una función . . . 662
12.2.2. Qué significa ser solución . . . 663
12.2.3. Solución general, particular y singular . . . 665
12.2.4. Problema de valor inicial y problema de contorno . . . 666
12.2.5. El vocabulario que clasifica: lineal, homogénea, coeficientes constantes . . . 667
12.2.6. Ecuaciones de variables separables . . . 671
12.2.7. Ecuaciones lineales de primer orden: de dónde sale el factor integrante . . . 673
12.2.8. Ecuaciones exactas: la definición que faltaba . . . 675
12.2.9. Existencia y unicidad: qué garantiza y qué no . . . 676
12.2.10. Campo de direcciones e isoclinas: ver la ecuación antes de resolverla 678
12.3. Ecuaciones lineales de segundo orden . . . 680
12.3.1. Coeficientes constantes . . . 680
12.3.2. Búsqueda de soluciones particulares . . . 681
12.3.3. Euler–Cauchy . . . 683
12.4. Sistemas lineales: el álgebra lineal en acción . . . 683
12.4.1. El caso diagonalizable . . . 684
12.4.2. La exponencial de matriz . . . 684
12.4.3. Retratos de fase y estabilidad . . . 685
12.5. La transformada de Laplace . . . 690
12.6. Soluciones por series de potencias . . . 694
12.6.1. Puntos ordinarios . . . 695
12.6.2. Puntos singulares regulares: Frobenius . . . 695
12.7. Series de Fourier . . . 698

12.8. Ecuaciones en derivadas parciales . . . 699
12.8.1. La ecuación del calor . . . 699
12.8.2. La ecuación de onda . . . 699
12.8.3. La ecuación de Laplace . . . 700
12.9. Fundamentos: la teoría que sostiene las EDO . . . 702
12.9.1. Picard–Lindelöf, demostrado de verdad . . . 702
12.9.2. Teoría de sistemas lineales: la matriz fundamental . . . 703
12.9.3. La transformada de Laplace, con fundamento . . . 704
12.9.4. Series de Fourier: la teoría de la convergencia . . . 705
12.9.5. Sturm–Liouville: el teorema espectral de las EDP . . . 706
12.9.6. Métodos de energía: unicidad para el calor y la onda . . . 706
12.10. Complementos: análisis cualitativo, Lyapunov y numéricos . . . 707
12.10.1. Ecuaciones autónomas y línea de fases . . . 707
12.10.2. Resonancia amortiguada y factor Q . . . 707
12.10.3. Lyapunov . . . 707
12.10.4. Euler y Runge–Kutta . . . 707
12.11. Cierre del temario: los subtemas restantes . . . 710
12.11.1. Orden n . . . 710
12.11.2. Fourier complejo . . . 710
12.11.3. Clasificación de EDP . . . 710
12.11.4. Dirichlet, Neumann, Robin . . . 710
12.11.5. Legendre y armónicos esféricos . . . 711
12.11.6. El tambor circular . . . 711
12.12. Profundización final: Green, el péndulo exacto y Hamilton . . . 711
12.12.1. Funciones de Green . . . 711
12.12.2. El péndulo exacto . . . 711
12.12.3. Estabilidad numérica . . . 711
12.12.4. Sistemas hamiltonianos . . . 712
12.13. Última capa . . . 712
12.13.1. El método de las características . . . 712
12.13.2. Poincaré–Bendixson y ciclos límite . . . 712
12.13.3. Teoría de Floquet y resonancia paramétrica . . . 713
12.13.4. Bifurcación de Hopf . . . 713
12.13.5. Caos: Lorenz y el mapa logístico . . . 713
12.13.6. Comparación de Sturm . . . 713
12.14. Ronda de temarios: Clairaut y las soluciones singulares . . . 714
12.14.1. La ecuación de Clairaut . . . 714
12.15. Cierre definitivo . . . 715
12.15.1. Existencia sin unicidad: el teorema de Peano . . . 715
12.15.2. El intervalo maximal y la alternativa de escape . . . 716
12.15.3. Existencia global: el criterio de crecimiento lineal . . . 716
12.16. Segunda ronda . . . 717

12.16.1. Lindstedt–Poincaré . . . 717
12.16.2. Ecuaciones en diferencias . . . 717
12.16.3. Resonancia modal completa . . . 717
12.17. Coda: el oscilador cuántico, resuelto . . . 717
12.18. Temas extra y ejemplos desarrollados . . . 718
12.18.1. Diccionario mecánica–electricidad . . . 718
12.18.2. Modos normales . . . 718
12.18.3. De la serie a la transformada de Fourier . . . 718
12.18.4. El pozo cuántico . . . 718
12.18.5. Laplace en el rectángulo . . . 718
12.19. Isoclinas y trayectorias ortogonales . . . 718
12.20. Descartar ciclos, Laplace periódico y Lagrange . . . 720
12.21. Oscilación y comparación: los teoremas de Sturm . . . 722
12.22. Poincaré–Bendixson: por qué el caos necesita tres dimensiones . . . 724
12.23. Definiciones que faltaban . . . 727
12.23.1. Ecuaciones autónomas, equilibrios y estabilidad . . . 727
12.23.2. La línea de fases y el criterio de la derivada . . . 728
12.23.3. El wronskiano, con su definición y su ley de evolución . . . 729
12.23.4. Dirichlet, Neumann y Robin: la tercera condición de contorno . . . 730
12.23.5. Soluciones singulares . . . 731
12.23.6. Clasificar equilibrios en dimensión dos . . . 732
12.24. Ejercicios resueltos . . . 734
12.25. Ejercicios adicionales resueltos . . . 736
12.26. Taller de ejemplos explicados . . . 737
12.26.1. Tipo 1: variables separables . . . 737
12.26.2. Tipo 2: lineal de primer orden (factor integrante) . . . 738
12.26.3. Tipo 3: lineal de segundo orden con coeficientes constantes . . . 738
12.27. Ampliación del temario . . . 739
12.27.1. La ecuación de Bessel y las funciones cilíndricas . . . 739
12.27.2. La transformada Z: Laplace para sistemas discretos . . . 739
12.27.3. Métodos numéricos para EDP: diferencias finitas y elementos finitos . . . 740
12.28. Ejercicios propuestos . . . 741

13. Ecuaciones en derivadas parciales 742
13.1. Qué es una EDP y por qué hay tres tipos . . . 742
13.2. El alfabeto de una EDP: notación, orden, linealidad y solución . . . 744
13.3. Transporte y el método de las características . . . 753
13.4. La ecuación de ondas y la fórmula de d'Alembert . . . 754
13.5. La ecuación del calor: difusión e irreversibilidad . . . 755
13.6. La ecuación de Laplace: el equilibrio . . . 756
13.7. Separación de variables y series de Fourier . . . 757

13.8. Problemas bien planteados . . . 759
13.9. Errores típicos con ecuaciones en derivadas parciales . . . 760
13.10. Soluciones débiles y Lax–Milgram . . . 761
13.11. Solución fundamental, Green y el núcleo de Poisson . . . 763
13.12. Choques, Rankine–Hugoniot y entropía . . . 765
13.13. Sturm–Liouville y funciones especiales . . . 767
13.14. Clasificación general de segundo orden y formas canónicas . . . 768
13.15. Las demostraciones que faltaban: d'Alembert, energía y unicidad . . . 770
13.16. El núcleo del calor, deducido con Fourier . . . 772
13.17. Principios del máximo: elíptico y parabólico, demostrados . . . 773
13.18. El método de las imágenes: problemas en la semirrecta . . . 776
13.19. Ondas en dimensión dos y tres: Kirchhoff, Poisson y Huygens . . . 777
13.20. Lax–Milgram y Poincaré, demostrados . . . 779
13.21. Diferencias finitas: consistencia, estabilidad de von Neumann y la condición CFL . . . 781
13.22. Curvas características, dominio de dependencia y dominio de influencia . . . 784
13.23. Ejercicios resueltos . . . 787
13.24. Ejercicios propuestos . . . 788

14. Métodos numéricos: cuando no hay fórmula 790
14.1. Por qué hacen falta . . . 790
14.2. Error absoluto, error relativo y cifras significativas . . . 791
14.3. El orden de convergencia de un método . . . 794
14.4. Radio espectral y convergencia de los métodos iterativos . . . 797
14.5. Bisección: el método que nunca falla . . . 800
14.6. Newton: seguir la tangente . . . 802
14.7. La secante: Newton sin derivada . . . 804
14.8. Iteración de punto fijo . . . 804
14.9. Interpolación . . . 807
14.10. Ajuste por mínimos cuadrados . . . 810
14.11. Integración numérica . . . 812
14.12. El error de verdad: redondeo y cancelación . . . 813
14.13. Sistemas lineales . . . 816
14.14. Métodos iterativos . . . 820
14.15. Ecuaciones diferenciales: de Euler a Runge–Kutta . . . 822
14.16. Derivación numérica . . . 826
14.17. Monte Carlo . . . 829
14.18. Errores típicos . . . 832
14.19. Romberg, Gauss y el punto flotante . . . 833
14.20. Splines y Newton para sistemas . . . 835
14.21. Los teoremas de error: Newton, interpolación y cuadratura . . . 837
14.22. Cómo se propaga el error de redondeo operación tras operación . . . 840

14.23. El número de condición de un problema numérico . . . 845
14.24. Ejercicios resueltos . . . 849
14.25. Ejercicios propuestos . . . 850

15. Optimización y análisis convexo 852
15.1. Convexidad y condiciones de optimalidad . . . 852
15.2. Los cimientos de la convexidad . . . 854
15.3. Algoritmos de descenso y desigualdades . . . 861
15.4. Dualidad de Lagrange y las condiciones KKT . . . 862
15.5. Funciones no suaves: subgradientes y métodos proximales . . . 865
15.6. Conos, programación semidefinida y relajaciones . . . 866
15.7. Conjugada de Fenchel y dualidad . . . 867
15.8. Separación, Farkas y puntos de silla . . . 870
15.9. Velocidades de convergencia y aceleración . . . 872
15.10. Cómo reconocer que algo es convexo sin derivar dos veces . . . 873
15.11. Ejercicios resueltos . . . 875
15.12. Ejercicios propuestos . . . 876

III Probabilidad y estadística 877

16. Variables aleatorias, distribuciones e inferencia 878
16.1. Espacio muestral, sucesos y la regla de Laplace . . . 879
16.2. Los axiomas de Kolmogórov y el álgebra de sucesos . . . 881
16.3. Probabilidad condicionada, árboles y el teorema de Bayes . . . 889
16.4. Variables aleatorias discretas . . . 892
16.4.1. La distribución binomial, revisitada . . . 894
16.5. La distribución de Poisson: contar sucesos raros . . . 895
16.6. Juegos de azar, esperanza de ganancia . . . 896
16.7. Variables aleatorias continuas . . . 899
16.8. La distribución exponencial y la falta de memoria . . . 900
16.9. La distribución normal . . . 902
16.9.1. Aproximación normal a la binomial . . . 903
16.10. Muestreo: la distribución de la media muestral . . . 906
16.11. Intervalos de confianza . . . 907
16.12. Contraste de hipótesis . . . 910
16.13. Errores típicos de probabilidad e inferencia . . . 914
16.14. La t de Student y el contraste χ^2 . . . 914
16.15. Vectores aleatorios, linealidad de la esperanza y Jensen . . . 916
16.16. Covarianza, correlación y potencia . . . 924
16.17. Desigualdades que valen siempre: Markov, Chebyshev y Chernoff . . . 925
16.18. Funciones generadoras de momentos y funciones características . . . 928
16.19. Modos de convergencia y las leyes de los grandes números . . . 930

16.20. El teorema central del límite, demostrado . . . 932
16.21. Transformar y sumar variables aleatorias . . . 934
16.22. El catálogo completo de las distribuciones discretas . . . 936
16.23. La función de distribución y los cuantiles . . . 941
16.24. Ejercicios resueltos . . . 946
16.25. Ejercicios propuestos . . . 947

17. Estadística descriptiva y análisis exploratorio 948
17.1. Medidas de centro y de dispersión: la cuestión de la robustez . . . 948
17.2. Variables, frecuencias, tipificación y las dos rectas . . . 950
17.3. Dos variables: covarianza, correlación y regresión descriptiva . . . 957
17.4. Cuando el resumen miente: agregación y confusión . . . 958
17.5. La forma de una distribución: histogramas, cuantiles y colas . . . 959
17.6. Errores típicos de estadística descriptiva . . . 963
17.7. Dos variables a la vez: correlación y la paradoja de Simpson . . . 963
17.8. Rangos y variables categóricas . . . 965
17.9. Concentración: Lorenz y Gini . . . 967
17.10. Robustez y atípicos multivariantes . . . 969
17.11. Transformaciones y datos faltantes . . . 971
17.12. Por qué la recta de mínimos cuadrados es esa, y por qué $R^2 = r^2$ . . . 973
17.13. Los conceptos que faltaban por escrito . . . 976
17.14. Ejercicios resueltos . . . 977
17.15. Ejercicios propuestos . . . 978

18. Inferencia estadística 979
18.1. Estimadores: sesgo, varianza y error cuadrático medio . . . 979
18.2. Población, muestra y distribución en el muestreo . . . 980
18.3. Máxima verosimilitud . . . 988
18.4. Intervalos de confianza . . . 990
18.5. Contrastes de hipótesis . . . 991
18.6. Potencia, tamaño muestral y comparaciones múltiples . . . 992
18.7. Errores típicos de inferencia . . . 995
18.8. Máxima verosimilitud y la información de Fisher . . . 996
18.9. Suficiencia y estimadores óptimos . . . 998
18.10. Razón de verosimilitudes, Wald y puntuación . . . 1000
18.11. Inferencia con dos muestras . . . 1002
18.12. Neyman–Pearson e intervalos por inversión . . . 1003
18.13. Las distribuciones del muestreo: χ^2, t y F, y el teorema de Fisher . . . 1006
18.14. Suficiencia, Rao–Blackwell y Lehmann–Scheffé, demostrados . . . 1008
18.15. Información de Fisher y la cota de Cramér–Rao . . . 1010
18.16. Máxima verosimilitud: propiedades asintóticas y el teorema de Wilks . . . 1011
18.17. Cobertura, y el método delta . . . 1013
18.18. Ejercicios resueltos . . . 1014

18.19. Ejercicios propuestos 1015

19. Modelos lineales, regresión múltiple y análisis de la varianza 1016
19.1. El modelo lineal en forma matricial 1016
19.2. Regresión simple, hipótesis del modelo y residuos 1018
19.3. Análisis de la varianza 1025
19.4. Diagnóstico del modelo: qué mirar antes de creerse el resultado 1027
19.5. El análisis de la varianza y los modelos lineales generalizados 1030
19.6. Gauss–Markov y la distribución de los estimadores 1032
19.7. Comparaciones múltiples tras el ANOVA 1033
19.8. Confianza y predicción 1035
19.9. ANOVA de dos factores e interacción 1036
19.10. Cuando fallan los supuestos 1038
19.11. Diseño de experimentos 1039
19.12. La distribución exacta de los estimadores y el teorema de Cochran 1041
19.13. Diagnóstico y selección: las definiciones explícitas 1043
19.14. Ejercicios resueltos 1045
19.15. Ejercicios propuestos 1045

20. Estadística bayesiana 1047
20.1. El teorema de Bayes como maquinaria de aprendizaje 1047
20.2. Verosimilitud, priori y posteriori, con una actualización hecha entera . . . 1049
20.3. Otras conjugaciones y la elección de la priori 1056
20.4. Cuando no hay conjugada: muestrear la posteriori 1057
20.5. Comparar modelos: evidencia, factores de Bayes y la paradoja de Lindley 1059
20.6. La distribución predictiva 1061
20.7. Resumir la posteriori: decisión, pérdida y HPD 1062
20.8. Modelos jerárquicos y Bayes empírico 1064
20.9. Prioris objetivas y cálculo de la evidencia 1065
20.10. Por qué funciona Metropolis–Hastings, y de dónde sale la conjugación normal 1067
20.11. Gibbs, prioris impropias y diagnóstico de convergencia 1070
20.12. Ejercicios resueltos 1072
20.13. Ejercicios propuestos 1072

21. Métodos no paramétricos y remuestreo 1074
21.1. Contrastes basados en rangos 1074
21.2. Qué significa no paramétrico, y las tres piezas básicas 1075
21.3. Bootstrap y remuestreo 1082
21.4. Bondad de ajuste y tablas de contingencia 1084
21.5. Regresión sin modelo: suavizado y splines 1087
21.6. Estimar una densidad sin suponer su forma 1089
21.7. Contrastes de permutación 1090

21.8. Rangos para k grupos: Kruskal–Wallis y Friedman . . . 1092
21.9. Variantes del bootstrap y validación cruzada . . . 1093
21.10. Intervalos exactos y Hodges–Lehmann . . . 1095
21.11. De dónde salen los p-valores: la distribución nula, construida . . . 1096
21.12. Qué es un rango, cómo se tratan los empates y la correlación de Spearman 1098
21.13. El contraste de los rangos con signo y su distribución exacta . . . 1102
21.14. Ejercicios resueltos . . . 1105
21.15. Ejercicios propuestos . . . 1106

22. Procesos estocásticos 1108
22.1. Cadenas de Markov . . . 1108
22.2. Qué es un proceso estocástico y una cadena calculada entera . . . 1110
22.3. Movimiento browniano . . . 1116
22.4. El proceso de Poisson y las colas de espera . . . 1117
22.5. Tiempos de parada, martingalas y el muestreo opcional . . . 1120
22.6. Cálculo de Itô y la ecuación de Black–Scholes . . . 1121
22.7. Clasificación de estados y MCMC . . . 1123
22.8. Máximo del browniano y tiempos de llegada . . . 1125
22.9. Procesos de ramificación . . . 1126
22.10. Tiempo continuo y ecuaciones estocásticas . . . 1128
22.11. Chapman–Kolmogórov y el teorema ergódico . . . 1130
22.12. Filtraciones, estados absorbentes y cómo analizar una cadena . . . 1132
22.13. Ejercicios resueltos . . . 1134
22.14. Ejercicios propuestos . . . 1134

IV Álgebra abstracta 1136

23. Teoría de grupos desde cero: la simetría explicada con las manos 1137
23.1. El experimento del triángulo . . . 1138
23.1.1. Componer movimientos . . . 1138
23.1.2. La tabla completa . . . 1139
23.2. Cimientos: operación, axiomas y primeros ejemplos . . . 1140
23.2.1. La pieza más pequeña: la operación . . . 1140
23.2.2. La definición de grupo, axioma por axioma . . . 1142
23.2.3. Las primeras demostraciones: lo que sale gratis de los axiomas . . . 1143
23.2.4. Un catálogo mínimo de grupos, todos verificados . . . 1145
23.2.5. Orden de un grupo y orden de un elemento . . . 1147
23.2.6. Subgrupos, subgrupo generado y grupos cíclicos . . . 1148
23.2.7. Permutaciones: dos notaciones y un convenio . . . 1150
23.2.8. Homomorfismos, núcleo e imagen . . . 1152
23.2.9. Clases laterales, índice y el teorema de Lagrange . . . 1154
23.2.10. Subgrupos normales, cociente y primer teorema de isomorfía . . . 1155

23.2.11. Productos directos, y qué queda hecho . . . 1157
23.3. Ahora sí: la definición, regla por regla . . . 1157
23.4. Un catálogo de grupos, verificado uno por uno . . . 1159
23.4.1. Los enteros con la suma . . . 1159
23.4.2. El reloj: los enteros módulo n . . . 1159
23.4.3. El grupo de Klein . . . 1159
23.4.4. Ejemplos que *no* son grupos . . . 1160
23.5. El orden de un elemento . . . 1160
23.6. Subgrupos: mundos dentro del mundo . . . 1161
23.7. El grupo del cuadrado: D_4 con todo detalle . . . 1162
23.8. Permutaciones con las manos . . . 1163
23.9. Lagrange con las cuentas hechas . . . 1165
23.10. Cocientes: mirar con lentes borrosos . . . 1166
23.11. Isomorfismos: cuándo dos grupos son el mismo . . . 1167
23.12. Acciones: los grupos puestos a trabajar . . . 1169
23.13. ¿Cuántos grupos hay de cada tamaño? . . . 1171
23.14. Productos directos: fabricar grupos nuevos . . . 1172
23.15. Dónde aparecen los grupos de verdad . . . 1173
23.16. Taller de ejemplos resueltos . . . 1174
23.17. Cierre: qué llevarte de este capítulo . . . 1176
23.18. Grupos cíclicos, homomorfismos y conjugación . . . 1176
23.19. Los teoremas de Sylow, enunciados y demostrados . . . 1179
23.20. Acción, simple y resoluble: los términos que los ejercicios usan . . . 1181
23.21. Ejercicios resueltos . . . 1182
23.22. Ejercicios propuestos . . . 1184

24. Anillos, cuerpos y polinomios desde cero 1186
24.1. El experimento: dos operaciones en vez de una . . . 1187
24.2. El vocabulario estructural completo . . . 1187
24.2.1. Subanillos . . . 1188
24.2.2. Ideales generados y operaciones con ideales . . . 1189
24.2.3. El teorema de correspondencia y los otros dos de isomorfía . . . 1191
24.2.4. Anillos locales: la palabra que faltaba . . . 1191
24.2.5. Localizar: añadir inversos a la carta . . . 1192
24.2.6. Módulos y elementos enteros, en dos definiciones . . . 1193
24.3. Ahora sí: la definición, regla por regla . . . 1194
24.4. Un catálogo de anillos, verificado uno por uno . . . 1195
24.5. Unidades y divisores de cero: los dos defectos de $\mathbb{Z}_6$ . . . 1196
24.6. Dominios de integridad y cuerpos . . . 1197
24.7. Los anillos $\mathbb{Z}_n$ con todo detalle . . . 1198
24.8. Homomorfismos: traducir entre anillos . . . 1198
24.9. Ideales: los subconjuntos que sirven para dividir . . . 1199

24.10. Anillos cociente: fabricar $\mathbb{C}$ a partir de $\mathbb{R}$ . . . 1200
24.11. Polinomios: la aritmética paralela a la de $\mathbb{Z}$ . . . 1201
24.12. Raíces, factores y el teorema del resto . . . 1203
24.13. Irreducibilidad: los criterios que de verdad se usan . . . 1204
24.14. Construir cuerpos finitos con las manos . . . 1205
24.15. Cuando la factorización única se rompe . . . 1206
24.16. La jerarquía completa . . . 1206
24.17. Dónde aparecen los anillos de verdad . . . 1207
24.18. Taller de ejemplos resueltos . . . 1207
24.19. Cierre: qué llevarte de este capítulo . . . 1209
24.20. Característica, isomorfía y fracciones . . . 1210
24.21. Lema de Gauss y varias variables . . . 1211
24.22. Cadenas ascendentes: anillos noetherianos y la base de Hilbert . . . 1213
24.23. Cuerpos finitos: existencia y unicidad . . . 1214
24.24. El teorema de Gauss en su versión fuerte . . . 1215
24.25. Divisibilidad en un dominio: asociados, irreducibles, primos y máximo común divisor . . . 1216
24.26. Dominios euclídeos y la cadena euclídeo, DIP, DFU, demostrada . . . 1219
24.27. Ejercicios resueltos . . . 1222
24.28. Ejercicios propuestos . . . 1225

25. Estructuras algebraicas: grupos, anillos y cuerpos 1226
25.1. Grupos: definición y ejemplos . . . 1226
25.1.1. Permutaciones en detalle . . . 1227
25.1.2. Subgrupos y grupos cíclicos . . . 1230
25.2. Anillos desde cero . . . 1232
25.2.1. Los axiomas, y lo que se deduce de ellos . . . 1233
25.2.2. Subanillos: la palabra que el capítulo usa sin decirla . . . 1234
25.2.3. Ideales: uno, dos o tres tipos según conmute o no . . . 1235
25.2.4. Homomorfismos, cocientes y los teoremas de isomorfía . . . 1238
25.2.5. El anillo de polinomios, definido de verdad . . . 1241
25.3. Lagrange, clases laterales y subgrupos normales . . . 1243
25.4. Homomorfismos y teoremas de isomorfía . . . 1246
25.5. Presentaciones . . . 1248
25.6. Acciones de grupos . . . 1252
25.7. El producto semidirecto y la clasificación . . . 1255
25.8. Tablas de caracteres . . . 1259
25.9. Grupos solubles y grupos simples . . . 1263
25.10. Anillos . . . 1268
25.10.1. La jerarquía aritmética . . . 1269
25.11. Polinomios . . . 1274
25.11.1. Irreducibilidad . . . 1275

25.12. Cuerpos 1278
25.13. Fundamentos: las demostraciones estructurales del álgebra 1282
25.13.1. La ecuación de clases y sus frutos 1282
25.13.2. La simplicidad de A_5, con las cuentas completas 1284
25.13.3. Aritmética de anillos: los teoremas que faltaban 1285
25.13.4. El grupo multiplicativo de un cuerpo finito es cíclico 1288
25.14. Complementos 1289
25.14.1. Productos directos y semidirectos 1289
25.14.2. Automorfismos 1289
25.14.3. Polinomios simétricos 1289
25.15. Cierre del temario: los subtemas restantes 1289
25.15.1. Segundo y tercer teoremas de isomorfía 1289
25.15.2. Jordan–Hölder 1291
25.15.3. Presentaciones 1291
25.15.4. El cuerpo de fracciones 1291
25.15.5. Factorización efectiva 1291
25.16. Profundización final 1291
25.16.1. Wilson 1291
25.16.2. Suma de dos cuadrados 1291
25.16.3. Resultantes 1292
25.16.4. Representaciones y caracteres 1292
25.16.5. La base de Hilbert 1292
25.17. Última capa 1292
25.17.1. Sylow, con demostración 1292
25.17.2. Módulos sobre un DIP: el teorema que unifica 1293
25.17.3. Símbolo de Jacobi 1294
25.17.4. Enteros algebraicos y dominios de Dedekind 1294
25.18. Ronda de temarios: nilpotencia y automorfismos 1295
25.18.1. Grupos nilpotentes 1295
25.18.2. Automorfismos interiores y exteriores 1298
25.19. Cierre definitivo: noetherianidad, Zorn y localización 1299
25.19.1. Anillos noetherianos y el teorema de la base de Hilbert 1299
25.19.2. Zorn y los ideales maximales 1301
25.19.3. Localización: mirar con lupa 1301
25.20. Segunda ronda: Pólya, idempotentes y $\mathbb{F}_{16}$ 1302
25.20.1. El cubo coloreado 1302
25.20.2. Idempotentes: el CRT hecho elementos 1302
25.20.3. $\mathbb{F}_{16}$ completo 1302
25.21. Coda: orden 12 y los teoremas de solubilidad masiva 1302
25.22. Temas extra y ejemplos desarrollados 1302
25.22.1. Burnside 1302
25.22.2. RSA completo 1303

25.22.3. El retículo de $\mathbb{Z}_{12}$ . . . 1303
25.22.4. Los grupos simples finitos . . . 1303
25.23. Módulos y la forma normal de Smith . . . 1303
25.24. Jordan–Hölder y álgebras . . . 1305
25.25. Tres demostraciones que se habían dejado como idea . . . 1307
25.26. Descomposición, normalidad y módulos libres . . . 1309
25.27. Ejercicios resueltos . . . 1313
25.28. Ejercicios adicionales resueltos . . . 1314
25.29. Taller de ejemplos explicados . . . 1315
25.29.1. El grupo $\mathbb{Z}_6$ y el orden de un elemento . . . 1315
25.29.2. Clases laterales y el teorema de Lagrange . . . 1315
25.29.3. Comprobar que algo es un grupo . . . 1316
25.30. Ejercicios propuestos . . . 1316

26. Teoría de Galois 1317
26.1. Extensiones de cuerpos . . . 1317
26.1.1. Elementos algebraicos y polinomio mínimo . . . 1318
26.1.2. Cuerpos de descomposición y normalidad . . . 1322
26.1.3. Separabilidad . . . 1322
26.2. Cimientos: cuerpos, extensiones y grados . . . 1322
26.2.1. Cuerpo, subcuerpo y característica . . . 1323
26.2.2. Extensiones de cuerpos y el grado . . . 1325
26.2.3. Torres de extensiones: el teorema del grado . . . 1326
26.2.4. Elementos algebraicos, trascendentes y el polinomio mínimo . . . 1328
26.2.5. $K[\alpha]$ frente a $K(\alpha)$: el pequeño milagro . . . 1331
26.2.6. Automorfismos: las simetrías de una extensión . . . 1333
26.2.7. Cuerpo de descomposición y cuerpo fijo . . . 1336
26.2.8. El vocabulario que queda, en una página . . . 1337
26.3. El grupo de Galois . . . 1338
26.4. La correspondencia de Galois . . . 1342
26.5. Las fórmulas clásicas: grados 2, 3 y 4 . . . 1348
26.6. El método de Dedekind . . . 1351
26.7. Sumas de Gauss y reciprocidad cuadrática . . . 1355
26.8. Solubilidad por radicales . . . 1360
26.9. Construcciones con regla y compás . . . 1367
26.10. Los períodos de Gauss y la construcción del heptadecágono . . . 1370
26.11. Fundamentos: la demostración de la correspondencia . . . 1373
26.11.1. Independencia de caracteres . . . 1373
26.11.2. El teorema de Artin . . . 1374
26.11.3. Ensamblaje del teorema fundamental . . . 1375
26.11.4. El teorema del elemento primitivo, demostrado . . . 1376
26.11.5. El teorema fundamental del álgebra, vía Galois . . . 1376

26.11.6. El discriminante y el cálculo de grupos de Galois 1378
26.11.7. Extensiones ciclotómicas y de Kummer: los peldaños de las torres 1379
26.12. Complementos 1381
26.12.1. Norma y traza 1381
26.12.2. Cuerpos finitos: teoría completa 1381
26.12.3. La ecuación general 1381
26.12.4. Cardano numérico completo 1381
26.12.5. Panorama 1381
26.13. Cierre del temario: los subtemas restantes 1382
26.13.1. Clausura algebraica 1382
26.13.2. Liouville y los primeros trascendentes 1382
26.13.3. Ciclotómicos completos 1382
26.13.4. La cuártica por su resolvente 1382
26.13.5. Base normal e independencia de radicales 1382
26.14. Profundización final 1382
26.14.1. Hilbert 90 1382
26.14.2. Sumas de Gauss y reciprocidad cuadrática 1383
26.14.3. $x^6 - 2$: el diédrico D_6 1383
26.14.4. El grupo absoluto y Krull 1383
26.14.5. Especialización e irreducibilidad de Hilbert 1383
26.15. Última capa: solubilidad en detalle, p-ádicos y Lüroth 1384
26.15.1. El criterio de solubilidad, con las dos direcciones 1384
26.15.2. Números p-ádicos 1384
26.15.3. El teorema de Lüroth 1385
26.15.4. Cohomología de Galois, en una cápsula 1386
26.16. Ronda de temarios 1386
26.16.1. El teorema de la base normal 1386
26.16.2. Inseparabilidad: lo que la característica p rompe 1386
26.16.3. Bases de trascendencia 1388
26.17. Cierre definitivo: Galois infinita y teoría de Kummer 1388
26.17.1. Teoría de Galois infinita: la topología de Krull en serio 1388
26.17.2. Teoría de Kummer: las abelianas, clasificadas 1389
26.18. Segunda ronda 1390
26.18.1. El teorema de extensión 1390
26.18.2. El Frobenius, visto factorizar 1390
26.18.3. Un mínimo con estrategia: $\sqrt{2} + i$ 1391
26.19. Coda: la quíntica soluble y Chebotarev cuantitativo 1391
26.20. Temas extra y ejemplos desarrollados 1391
26.20.1. El pentágono con regla y compás 1391
26.20.2. Cuerpos cuadráticos y Kronecker–Weber 1392
26.20.3. Un reducible instructivo 1392
26.20.4. Galois computacional 1392

26.20.5. AES y los cuerpos de Galois . . . 1392
26.21. Ciclotómicos, casus irreducibilis y trascendencia . . . 1392
26.22. Gauss–Wantzel: la implicación que faltaba, y el teorema 90 de Hilbert . . . 1394
26.23. Separabilidad, clausura algebraica y el criterio S_p . . . 1397
26.24. Ejercicios resueltos . . . 1402
26.25. Ejercicios adicionales resueltos . . . 1403
26.26. Taller de ejemplos explicados: grupos de Galois concretos . . . 1404
26.26.1. Un caso de grado 2 . . . 1404
26.26.2. Un caso de grado 3 (más rico) . . . 1404
26.27. Ejercicios propuestos . . . 1405

27. Teoría de la representación de grupos 1406
27.1. Representaciones, caracteres y ortogonalidad . . . 1406
27.2. Qué es una representación y la tabla de S_3 construida desde cero . . . 1408
27.3. La representación regular y las aplicaciones . . . 1413
27.4. Los grupos simétricos y los grupos continuos . . . 1414
27.5. Inducción, restricción y el teorema de Burnside . . . 1416
27.6. Schur, Maschke y de dónde salen las tablas . . . 1417
27.7. Operaciones con representaciones . . . 1419
27.8. Grupos compactos y Peter–Weyl . . . 1421
27.9. Restringir, inducir y Frobenius . . . 1422
27.10. Las relaciones de ortogonalidad, demostradas . . . 1423
27.11. El teorema de Burnside, demostrado . . . 1426
27.12. Funciones de clase y el producto tensorial de representaciones . . . 1428
27.13. Ejercicios resueltos . . . 1429
27.14. Ejercicios propuestos . . . 1430

V Análisis 1432

28. Topología: la geometría de lo que no se rompe 1433
28.1. La idea: geometría sin distancias . . . 1433
28.2. Subespacios, subbases y puntos de acumulación . . . 1434
28.3. Espacios métricos . . . 1440
28.4. Conjuntos abiertos y cerrados . . . 1441
28.5. Espacios topológicos: quedarse solo con lo esencial . . . 1445
28.6. Continuidad, redefinida . . . 1446
28.7. Taller: comprobar la continuidad en la práctica . . . 1448
28.8. Compacidad . . . 1450
28.9. Conexión . . . 1454
28.10. Homeomorfismos: cuándo dos espacios son “el mismo” . . . 1457
28.11. Taller: fabricar un homeomorfismo con todas sus letras . . . 1460
28.12. Bases y subbases: describir una topología sin listarla . . . 1462

28.13. Topología producto y topología cociente . . . 1463
28.14. Axiomas de numerabilidad y compactificación por un punto . . . 1464
28.15. Axiomas de separación . . . 1468
28.16. Hausdorff: el axioma que se usa a diario . . . 1470
28.17. Compacidad: las tres caras . . . 1473
28.18. Conexión por caminos . . . 1473
28.19. Una mirada al grupo fundamental . . . 1473
28.20. Espacios recubridores: desenrollar un espacio . . . 1477
28.21. Característica de Euler y superficies . . . 1482
28.22. Taller de ejemplos explicados . . . 1486
28.22.1. Demostrar que un conjunto es abierto . . . 1486
28.22.2. Demostrar que $(0,1)$ no es compacto . . . 1486
28.22.3. Un homeomorfismo explícito . . . 1486
28.22.4. Distinguir dos espacios con un invariante . . . 1486
28.23. Entornos, redes y homotopía: los términos que faltaban . . . 1487
28.24. Ejercicios resueltos . . . 1489
28.25. Ejercicios propuestos . . . 1491

29. Análisis real: los cimientos y las patologías 1492
29.1. Qué falta todavía . . . 1492
29.2. Los cimientos elementales, definidos desde cero . . . 1493
29.3. Contar lo infinito: cardinalidad . . . 1503
29.4. El conjunto de Cantor y la escalera del diablo . . . 1510
29.5. Oscilación . . . 1513
29.6. El teorema de Baire . . . 1517
29.7. Espacios completos y el principio de contracción . . . 1521
29.8. Una función continua sin derivada en ningún punto . . . 1524
29.9. Familias de funciones: Dini y Arzelà–Ascoli . . . 1526
29.10. Aproximación: Bernstein, Weierstrass y Stone . . . 1532
29.11. Reordenar series: el teorema de Riemann . . . 1533
29.12. Variación acotada y continuidad absoluta . . . 1539
29.13. Taller de ejercicios resueltos . . . 1547
29.14. Cierre: qué llevarte de este capítulo . . . 1549
29.15. Medida cero, sin esperar al capítulo de Lebesgue . . . 1549
29.16. Ejercicios resueltos . . . 1550
29.17. Ejercicios propuestos . . . 1553

30. Variable compleja: el puente entre los mundos 1555
30.1. Funciones holomorfas y las ecuaciones de Cauchy–Riemann . . . 1555
30.2. Los cimientos, uno a uno . . . 1558
30.2.1. Funciones de variable compleja . . . 1559
30.2.2. Límite y continuidad en el plano complejo . . . 1560
30.2.3. La derivada compleja . . . 1561

30.2.4. Las ecuaciones de Cauchy–Riemann, deducidas . . . 1562
30.2.5. Holomorfa, entera, analítica: tres palabras y su relación . . . 1564
30.2.6. Funciones armónicas y armónicas conjugadas . . . 1565
30.2.7. Las funciones elementales complejas . . . 1567
30.2.8. Caminos e integral de contorno . . . 1569
30.3. El teorema de Cauchy y la fórmula integral . . . 1574
30.4. Series de Laurent, singularidades y residuos . . . 1579
30.5. Mapas conformes y el teorema de Riemann . . . 1588
30.6. El panorama desde arriba . . . 1593
30.7. Ronda de temarios . . . 1593
30.7.1. El lema de Schwarz y los automorfismos del disco . . . 1594
30.7.2. El núcleo de Poisson y el problema de Dirichlet en el disco . . . 1595
30.7.3. Prolongación analítica: el principio de reflexión . . . 1596
30.7.4. La función Gamma y la fórmula de reflexión . . . 1597
30.8. Cierre definitivo . . . 1598
30.8.1. El teorema del mapeo de Riemann, demostrado . . . 1598
30.8.2. Los teoremas de Picard . . . 1599
30.8.3. Funciones elípticas: $\wp$ de Weierstrass . . . 1600
30.9. Segunda ronda: Rouché y el principio del argumento . . . 1601
30.9.1. El principio del argumento . . . 1601
30.9.2. La inversión de Laplace: Bromwich . . . 1603
30.10. El índice de una curva y el teorema de los residuos, demostrado . . . 1603
30.11. El teorema de Laurent, y el cálculo sistemático de residuos . . . 1606
30.12. Integrales reales por residuos: lema de Jordan, valor principal y el semicírculo pequeño . . . 1608
30.13. Entera, meromorfa y familia normal: las tres palabras que faltaban . . . 1611
30.14. Ejercicios resueltos . . . 1612
30.15. Taller de ejemplos explicados . . . 1613
30.15.1. Aritmética compleja . . . 1613
30.15.2. Una integral real resuelta por residuos . . . 1614
30.16. Ejercicios propuestos . . . 1614

31. La integral de Lebesgue y la transformada de Fourier 1616
31.1. Por qué Riemann no basta . . . 1616
31.2. Qué es medir: sigma-álgebras, medidas y funciones medibles . . . 1617
31.3. La medida de Lebesgue . . . 1626
31.4. La integral de Lebesgue . . . 1633
31.5. El cambio de variable y las coordenadas polares . . . 1639
31.6. Los espacios L^p y la completitud . . . 1641
31.7. La transformada de Fourier, con demostraciones . . . 1646
31.8. Cierre definitivo . . . 1654
31.8.1. Medidas producto y Fubini–Tonelli . . . 1654

31.8.2. Radon–Nikodym: toda medida razonable tiene densidad 1657
31.8.3. El teorema de diferenciación de Lebesgue 1658
31.9. Variación acotada y continuidad absoluta 1659
31.10. Medidas de Lebesgue–Stieltjes y medidas signadas 1662
31.11. El broche: la probabilidad es teoría de la medida 1665
31.12. L^∞, el supremo esencial, y las desigualdades demostradas 1668
31.13. Egorov y Lusin: casi uniforme, casi continua 1670
31.14. El lema de Vitali y la derivación de funciones monótonas 1671
31.15. σ-finitud y medida producto: las hipótesis que las reglas mandan comprobar 1674
31.16. Ejercicios resueltos 1674
31.17. Taller de ejemplos explicados: convergencia y medida cero 1676
31.17.1. Un límite de integrales con el teorema de convergencia dominada 1676
31.17.2. Un conjunto de medida cero 1676
31.18. Ejercicios propuestos 1677

32. Análisis funcional: álgebra lineal en dimensión infinita 1678
32.1. La idea: cuando los vectores son funciones 1678
32.2. Aplicaciones lineales, funcionales y operadores acotados 1679
32.3. Espacios normados y espacios de Banach 1686
32.4. Espacios de Hilbert: la geometría regresa 1692
32.5. Bases ortonormales, Bessel y Parseval 1694
32.6. Operadores acotados 1699
32.7. Los cuatro teoremas grandes 1704
32.8. El principio de contracción de Banach 1711
32.9. Dualidad y el teorema de representación de Riesz 1715
32.10. Operadores compactos y el teorema espectral 1716
32.11. Ecuaciones integrales: el caso del núcleo degenerado 1719
32.12. Convergencia débil y compacidad recuperada 1722
32.13. Formas bilineales, Lax–Milgram y el método de Galerkin 1723
32.14. Taller de ejercicios resueltos 1726
32.15. Cierre: qué llevarte de este capítulo 1728
32.16. Espectro y resolvente: las definiciones y la fórmula de Gelfand 1729
32.17. Espacios de Sobolev, derivada débil y la topología débil-$*$ 1731
32.18. El adjunto y las clases de operadores, definidos 1733
32.19. Ejercicios resueltos 1734
32.20. Ejercicios propuestos 1737

33. Análisis armónico 1739
33.1. Series de Fourier 1739
33.2. Periodicidad, paridad y ortogonalidad de funciones 1740
33.3. La transformada de Fourier 1752
33.4. Convolución, núcleos y el teorema de muestreo 1753
33.5. Ondículas: mirar la frecuencia sin perder el tiempo 1756

33.6. Inversión, Plancherel y distribuciones . . . 1758
33.7. La transformada discreta y el algoritmo rápido . . . 1760
33.8. Fejér, Poisson y suavidad frente a decaimiento . . . 1761
33.9. Integrales singulares y la función maximal . . . 1763
33.10. Sumación de Poisson y varias variables . . . 1765
33.11. El núcleo de Dirichlet y la constante de Lebesgue . . . 1766
33.12. Convergencia puntual: Riemann–Lebesgue, Dini y el teorema de Dirichlet 1769
33.13. Fejér, demostrado, y sus consecuencias . . . 1770
33.14. Calcular una serie de Fourier y decidir cómo converge . . . 1772
33.15. Ejercicios resueltos . . . 1774
33.16. Ejercicios propuestos . . . 1775

34. Teoría analítica de números 1776
34.1. La función zeta y la distribución de los primos . . . 1776
34.2. Series de Dirichlet, productos de Euler y prolongación analítica . . . 1778
34.3. Funciones aritméticas y primos en progresiones . . . 1784
34.4. El teorema de los números primos y la hipótesis de Riemann . . . 1785
34.5. Cribas y la conjetura de los primos gemelos . . . 1787
34.6. Caracteres, funciones L y primos en progresiones . . . 1788
34.7. Continuación analítica y ecuación funcional . . . 1790
34.8. Las funciones de Chebyshev y los órdenes medios . . . 1792
34.9. Sumas exponenciales y el método del círculo . . . 1793
34.10. La criba de Selberg, con los pesos calculados . . . 1795
34.11. La notación asintótica y las funciones aritméticas, por escrito . . . 1797
34.12. Ejercicios resueltos . . . 1799
34.13. Ejercicios propuestos . . . 1800

VI Geometría, topología y abstracción 1801

35. Geometría algebraica 1802
35.1. Variedades e ideales: el diccionario . . . 1802
35.2. El diccionario por escrito: conjuntos algebraicos y división . . . 1803
35.3. Curvas planas: singularidades, género y Bézout . . . 1812
35.4. El plano proyectivo: por qué las cuentas salen bien en el infinito . . . 1813
35.5. Del anillo al espacio: el espectro y los esquemas . . . 1816
35.6. Anillo de coordenadas y morfismos . . . 1817
35.7. Bases de Gröbner y eliminación . . . 1818
35.8. Divisores y Riemann–Roch . . . 1820
35.9. Haces y el espectro . . . 1821
35.10. Multiplicidad de intersección, divisor canónico y el algoritmo de Buchberger 1823
35.11. Ejercicios resueltos . . . 1826
35.12. Ejercicios propuestos . . . 1826

36. Cálculo tensorial **1828**
36.1. Convenio de Einstein . . . 1828
36.2. Índices arriba y abajo y el convenio de Einstein . . . 1829
36.2.1. Índices y componentes . . . 1829
36.2.2. El convenio de sumación . . . 1830
36.2.3. La delta de Kronecker . . . 1832
36.2.4. Arriba, abajo y el cambio de coordenadas . . . 1833
36.2.5. Dos operaciones que ya se pueden definir con índices . . . 1836
36.3. Vectores, covectores y el espacio dual . . . 1837
36.3.1. La ley de transformación . . . 1838
36.4. Tensores y sus operaciones . . . 1840
36.5. La métrica . . . 1846
36.5.1. Subir y bajar índices . . . 1847
36.6. Derivación covariante . . . 1849
36.6.1. Transporte paralelo y geodésicas . . . 1854
36.7. Curvatura . . . 1857
36.8. Aplicaciones físicas . . . 1865
36.8.1. Electromagnetismo covariante . . . 1865
36.8.2. Relatividad general . . . 1865
36.8.3. Medios continuos . . . 1866
36.9. Fundamentos: la teoría detrás del formalismo . . . 1868
36.9.1. El producto tensorial, construido formalmente . . . 1868
36.9.2. La conexión de Levi-Civita: existencia y unicidad . . . 1870
36.9.3. Geodésicas por cálculo de variaciones . . . 1871
36.9.4. Las simetrías del tensor de Riemann, demostradas . . . 1873
36.9.5. El Teorema Egregium y qué significa "intrínseco" . . . 1875
36.10. Complementos: Levi-Civita, Killing y el plano hiperbólico . . . 1876
36.10.1. El símbolo de Levi-Civita . . . 1876
36.10.2. Derivada de Lie y campos de Killing . . . 1876
36.10.3. El plano hiperbólico . . . 1877
36.11. Cierre del temario: los subtemas restantes . . . 1877
36.11.1. Variedades formalmente . . . 1877
36.11.2. Vectores tangentes como derivaciones . . . 1878
36.11.3. Conos de luz y geodésicas nulas . . . 1878
36.11.4. Christoffel esféricos en $\mathbb{R}^3$ . . . 1878
36.11.5. Isometrías y espacios homogéneos . . . 1878
36.12. Profundización final . . . 1879
36.12.1. Qué mide cada curvatura . . . 1879
36.12.2. Segunda forma fundamental . . . 1879
36.12.3. Ondas gravitacionales . . . 1879
36.12.4. El mapa exponencial . . . 1880
36.13. Última capa: Lie, Cartan y Friedmann . . . 1880

36.13.1. Grupos y álgebras de Lie . . . 1880
36.13.2. El método de las formas de Cartan . . . 1881
36.13.3. Las ecuaciones de Friedmann, deducidas . . . 1881
36.13.4. Kerr y la termodinámica de agujeros negros . . . 1881
36.14. Ronda de temarios: desviación geodésica . . . 1881
36.14.1. La ecuación de Jacobi . . . 1882
36.15. Cierre definitivo: Hopf–Rinow, Frobenius y Gauss–Bonnet . . . 1883
36.15.1. Hopf–Rinow: la completitud es una sola . . . 1883
36.15.2. El teorema de Frobenius: cuándo los planos se pegan en hojas . 1883
36.15.3. Gauss–Bonnet, demostrado . . . 1884
36.16. Segunda ronda: Bianchi contraída, polares y el GPS . . . 1885
36.16.1. La identidad que hace consistente a Einstein . . . 1885
36.16.2. La aceleración en polares, desde Christoffel . . . 1885
36.16.3. El GPS con números . . . 1886
36.17. Coda: espinores y la doble cubierta . . . 1886
36.18. Temas extra y ejemplos desarrollados . . . 1886
36.18.1. Relatividad especial como geometría . . . 1886
36.18.2. Cuadrivelocidad y $E = mc^2$ . . . 1886
36.18.3. El cono: curvatura concentrada . . . 1886
36.18.4. El tensor energía–momento . . . 1887
36.18.5. Triángulos que delatan la curvatura . . . 1887
36.19. Conexión y tensor de Weyl: los objetos que faltaban por definir . . . 1887
36.20. Ejercicios resueltos . . . 1888
36.21. Ejercicios adicionales resueltos . . . 1890
36.22. Taller de ejemplos explicados: midiendo sobre la esfera . . . 1891
36.22.1. Longitud de un paralelo . . . 1891
36.22.2. Un símbolo de Christoffel de la esfera . . . 1891
36.23. Ejercicios propuestos . . . 1892

37. Topología algebraica 1893
37.1. Invariantes: la característica de Euler . . . 1893
37.2. Caminos, lazos y por qué el grupo fundamental es un grupo . . . 1894
37.3. Homología: contar agujeros de cada dimensión . . . 1900
37.4. El grupo fundamental: lazos que no se pueden deshacer . . . 1902
37.5. Recubrimientos y el grado de una aplicación . . . 1905
37.6. Cohomología, producto y dualidad de Poincaré . . . 1906
37.7. Complejos de cadenas: qué es la homología . . . 1907
37.8. Funtorialidad, Mayer–Vietoris y Künneth . . . 1909
37.9. Recubrimientos y grupos de homotopía superiores . . . 1911
37.10. Van Kampen y los complejos CW, enunciados . . . 1912
37.11. Equivalencia homotópica, orientabilidad y homología celular . . . 1914
37.12. Ejercicios resueltos . . . 1919

37.13. Ejercicios propuestos . . . 1920

38. Teoría de categorías 1921
38.1. Categorías, funtores y transformaciones naturales . . . 1921
38.2. Los axiomas de categoría, seis ejemplos y los isomorfismos . . . 1923
38.3. Límites, colimites y adjunciones . . . 1928
38.4. El lema de Yoneda: un objeto es lo que hace . . . 1929
38.5. Categorías abelianas y sucesiones exactas . . . 1931
38.6. Mónadas: estructura que se compone consigo misma . . . 1933
38.7. Mono, epi y equivalencia de categorías . . . 1934
38.8. Funtores derivados, Ext y Tor . . . 1936
38.9. Prehaces, Yoneda y mónadas . . . 1938
38.10. Tamaño, diagramas y conos: los cimientos que faltaban . . . 1939
38.11. Ejercicios resueltos . . . 1941
38.12. Ejercicios propuestos . . . 1942

Epílogo: el mismo número, visto seis veces 1944

A. Mapa de cierre del temario 1948

B. Formulario 1967
B.1. Álgebra . . . 1967
B.2. Geometría analítica . . . 1968
B.3. Trigonometría . . . 1968
B.4. Exponenciales y logaritmos . . . 1969
B.5. Límites . . . 1969
B.6. Derivadas . . . 1969
B.7. Integrales . . . 1969
B.8. Series . . . 1970
B.9. Álgebra lineal . . . 1970
B.10. Cálculo vectorial . . . 1970
B.11. Ecuaciones diferenciales . . . 1971
B.12. Geometría del plano y del espacio . . . 1971
B.13. Números complejos . . . 1971
B.14. Coordenadas polares y curvas paramétricas . . . 1971
B.15. Sucesiones, progresiones y matemática financiera . . . 1972
B.16. Combinatoria . . . 1972
B.17. Probabilidad y distribuciones . . . 1972
B.18. Estadística e inferencia . . . 1973
B.19. Aritmética modular y criptografía . . . 1973
B.20. Sistemas de numeración . . . 1974
B.21. Métodos numéricos . . . 1974
B.22. Grafos . . . 1974

B.23. Álgebra abstracta . . . 1974
B.24. Desigualdades y estimaciones útiles . . . 1975
B.25. Constantes . . . 1975

C. Tablas estadísticas 1976
C.1. Normal estándar: función de distribución $\Phi(z)$ . . . 1976
C.2. t de Student: cuantiles $t_{\nu,p}$ . . . 1978
C.3. Chi cuadrado: cuantiles $\chi^2_{\nu,p}$ . . . 1979
C.4. F de Fisher: cuantiles al 95 % y al 99 % . . . 1980
C.5. Rango studentizado $q_{\alpha;k,\nu}$ para el método de Tukey . . . 1981
C.6. Valores críticos de Kolmogórov–Smirnov y de Wilcoxon . . . 1981

D. Glosario de símbolos 1983
D.1. Conjuntos y lógica . . . 1983
D.2. Números . . . 1983
D.3. Análisis . . . 1984
D.4. Álgebra . . . 1984
D.5. Geometría y topología . . . 1985
D.6. Notación escolar de uso diario . . . 1985
D.7. Trigonometría, exponenciales y logaritmos . . . 1986
D.8. Combinatoria, probabilidad y estadística . . . 1987
D.9. Sucesiones, matrices y estructuras discretas . . . 1987
D.10. El alfabeto griego . . . 1988
D.11. Abreviaturas y expresiones latinas . . . 1989
D.12. Convenios de este libro . . . 1989

E. Soluciones a los ejercicios propuestos 1990

F. Bibliografía comentada y siguientes pasos 2013

Índice alfabético 2021

Índice de figuras

1.1. Diagrama de Venn de tres conjuntos 3

2.1. La brecha de Gödel 37
2.2. El mapa de la complejidad 46

3.1. La aritmética modular como un reloj 90

4.1. El plano de Argand 150
4.2. Las raíces quintas de la unidad 157

5.1. El triángulo de Pascal 194
5.2. La biyección de barras y estrellas 203

6.1. Los impostores de Taylor de $\sin x$ 239
6.2. Convergencia uniforme frente a puntual 252

7.1. Un paso de Gram–Schmidt 339
7.2. La SVD: del círculo a la elipse 355

8.1. El grafo de trabajo del vocabulario 391
8.2. Los puentes de Königsberg 399
8.3. Un grafo de seis vértices 400

9.1. Multiplicadores de Lagrange: la tangencia 458
9.2. Curvas de nivel: mínimo y silla 463
9.3. La astroide como envolvente 490

10.1. Rosa, cardioide y espiral 525
10.2. La cicloide gana a la recta 558

11.1. Divergencia y rotacional, vistos 574
11.2. Stokes como cancelación interna 616

12.1. Campo de direcciones de la logística 655
12.2. El plano de fase del péndulo 691
12.3. Fourier de la onda cuadrada y el fenómeno de Gibbs 698

12.4. El ciclo límite de Van der Pol 725

13.1. Características de la ecuación de ondas 754
13.2. Galerkin con funciones sombrero 761
13.3. Choque y rarefacción en la ecuación de Burgers 765
13.4. La frontera parabólica excluye la tapa 774
13.5. Huygens: cáscara en 3D, disco en 2D 778
13.6. El factor de amplificación y el umbral $r = 1/2$ 781

14.1. El método de Newton, tres iteraciones 803
14.2. Truncamiento frente a redondeo 833

15.1. Hiperplano soporte de un convexo 852
15.2. La conjugada como familia de tangentes 868
15.3. Proyección y separación de convexos 870
15.4. Aceleración de Nesterov con $\kappa = 100$ 872

16.1. La regla empírica de la normal 902
16.2. Error de tipo II y potencia 924
16.3. Las cotas de concentración comparadas 926
16.4. La convolución fabrica la campana 934

17.1. Diagrama de caja y valores atípicos 949
17.2. La curva de Lorenz 968
17.3. Atípicos que ninguna variable delata 969
17.4. Regresión es proyección ortogonal 974

18.1. Los dos errores de un contraste 991
18.2. Los tres contrastes asintóticos 1000
18.3. Datos apareados frente a independientes 1003
18.4. Las colas de la t frente a la normal 1006

19.1. Ajuste por mínimos cuadrados y residuos 1017
19.2. Bandas de confianza y de predicción 1035

20.1. Priori, verosimilitud y posteriori 1048
20.2. Predictiva frente a estimación puntual 1061
20.3. Intervalo HPD frente a colas iguales 1063

21.1. Distribución empírica frente a teórica 1084
21.2. La distribución de permutación 1090

22.1. Una cadena de Markov de tres estados 1108
22.2. El principio de reflexión 1125

23.1. Las seis simetrías del triángulo 1138

23.2. El retículo de subgrupos de $\mathbb{Z}_{12}$ 1174

24.1. La jerarquía de dominios de integridad 1197

25.1. La correspondencia de Galois de $\mathbb{Q}(\sqrt{2},\sqrt{3})$ 1227

26.1. La correspondencia de Galois en un retículo 1343

27.1. Los cinco diagramas de Young de S_4 1414

28.1. Compacidad por recubrimientos 1451
28.2. La condición de Hausdorff 1472

29.1. Construcción del conjunto de Cantor 1510

30.1. Contorno semicircular y polos 1579
30.2. El contorno semicircular 1579
30.3. El índice cuenta vueltas con signo 1604
30.4. El contorno indentado para $\int \sin x/x$ 1609

31.1. Riemann rebana el dominio, Lebesgue el rango 1627

32.1. Bolas unidad en tres normas 1686

33.1. Sumas de Fourier y fenómeno de Gibbs 1740
33.2. Sumas parciales de la serie de x^2 1747
33.3. Sumas parciales de la media rampa 1748
33.4. Núcleo de Dirichlet y núcleo de Fejér 1762
33.5. Dirichlet oscila, Fejér no 1767

34.1. La escalera de $\pi(x)$ frente a $\mathrm{li}(x)$ 1777
34.2. La banda crítica de la función zeta 1791

35.1. La ley de grupo de una curva elíptica 1812

36.1. Transporte paralelo sobre la esfera 1829
36.2. Transporte paralelo en la esfera 1849

37.1. Toro y esfera: π_1 los separa 1902
37.2. Mayer–Vietoris aplicada a la esfera 1909

38.1. Propiedad universal del producto 1922

Cómo usar esta guía

Esta guía empieza en cero y termina en la teoría de Galois. No supone nada más que saber sumar, restar, multiplicar y dividir números: la Parte I construye desde ahí la lógica y los conjuntos, la teoría de números, la aritmética modular, los números complejos y la combinatoria, la estadística, las sucesiones y los límites, con el detalle y la paciencia de un libro de colegio. A partir de esa base, el libro sube por el cálculo, el álgebra lineal, el análisis y el álgebra abstracta hasta tres cumbres: las **ecuaciones diferenciales**, la **teoría de Galois** y el **cálculo tensorial**.

El camino es largo, pero está señalizado: cada capítulo contiene la teoría con sus demostraciones, cajas de *Idea clave* que dicen en palabras qué significa cada cosa antes de que aparezca el simbolismo, cajas de *Cuidado* con los errores típicos, ejemplos resueltos paso a paso y ejercicios con solución completa.

Cómo está ordenado el libro. Los capítulos están dispuestos en el orden en que se deben leer, no en el orden en que se escribieron. El libro se divide en cinco partes, y cada parte solo usa herramientas de las anteriores. Si ya sabes cálculo, empieza directamente en la Parte II; si vienes de cero, empieza en la página uno de la Parte I y no te saltes nada.

La ruta de estudio

El orden de los capítulos no es arbitrario: cada uno usa herramientas del anterior.

Parte I — Fundamentos: lógica, números y estructuras discretas. Todo lo que hace falta para poder empezar, y nada de secundaria. **Conjuntos, lógica y cómo se demuestra algo** (capítulo 1) es el idioma en el que está escrito el resto: qué es un conjunto, cómo funciona una implicación y su contrarrecíproco, cómo se niega un enunciado con cuantificadores y cuáles son los métodos de demostración. Es corto y conviene no saltárselo, porque a partir del siguiente capítulo el libro lo da por sabido. **Lógica matemática y teoría de conjuntos** (capítulo 2) da un paso más y convierte la lógica misma en objeto de estudio: sintaxis, semántica, completitud, los teoremas de incompletitud de Gödel y la indecidibilidad del problema de la parada. **Teoría de números: de la divisibilidad a la criptografía** (capítulo 3) es la puerta de entrada al pensamiento algebraico y el capítulo que sostiene medio libro: arranca en la división con resto y el algoritmo de Euclides, pasa por los primos, la factorización única, las congruencias, el teorema chino del resto y los teoremas de Fermat y Euler, y termina en los dígitos de control del código de barras, el intercambio de claves de Diffie y Hellman y RSA explicado paso a paso. **Números complejos** (capítulo 4) cuenta qué es i de verdad, con el plano de Argand, la forma polar y De Moivre, desde la historia real de las cúbicas de Cardano

y no desde la fábula de $x^2 + 1 = 0$. Y **combinatoria y probabilidad** (capítulo 5) enseña a contar sin listar, con el binomio de Newton y la probabilidad elemental hasta Bayes.

Parte II — Cálculo y álgebra lineal. La maquinaria de trabajo. **Cálculo en una variable** (capítulo 6) construye $\mathbb{R}$ desde el axioma del supremo y levanta sobre él sucesiones, continuidad, derivada, integral de Riemann, Taylor y series, con todas las demostraciones. **Álgebra lineal** (capítulo 7) es el lenguaje de casi toda la matemática moderna: sin espacios vectoriales no hay cálculo multivariable serio, ni sistemas de ecuaciones diferenciales, ni tensores. **Grafos** (capítulo 8) viene justo después porque su herramienta central es la matriz de adyacencia; trae los puentes de Königsberg, los árboles, la fórmula de Euler, los cuatro colores y el algoritmo con que el teléfono calcula rutas. **Cálculo multivariable** (capítulo 9) extiende la derivada a varias dimensiones con gradiente, hessiana, jacobiano y multiplicadores de Lagrange; **coordenadas polares y paramétricas** (capítulo 10) llega aquí y no antes porque necesita la integral y el jacobiano. **Cálculo vectorial** (capítulo 11) culmina en Green, Stokes y Gauss. **Ecuaciones diferenciales** (capítulo 12) es la primera cumbre, y **ecuaciones en derivadas parciales** (capítulo 13) la continúa con ondas, calor y Laplace. Cierran la parte **métodos numéricos** (capítulo 14), lo que de verdad hace un ordenador cuando se le pide resolver algo, y **optimización y análisis convexo** (capítulo 15), la teoría que explica cuándo un mínimo local es global.

Parte III — Probabilidad y estadística. La parte que más se usa fuera de un aula y la que peor se interpreta en la prensa. **Variables aleatorias, distribuciones e inferencia** (capítulo 16) va de la regla de Laplace al teorema central del límite. **Estadística descriptiva** (capítulo 17) se ocupa de los datos antes del modelo: robustez, forma, atípicos y cuándo un resumen miente. **Inferencia estadística** (capítulo 18) formaliza estimadores, máxima verosimilitud, intervalos, contrastes, potencia y comparaciones múltiples. **Modelos lineales** (capítulo 19) es regresión y análisis de la varianza en forma matricial, con diagnóstico. **Estadística bayesiana** (capítulo 20) cambia la pregunta y con ella toda la maquinaria. **Métodos no paramétricos** (capítulo 21) es lo que queda cuando no se puede suponer normalidad. Y **procesos estocásticos** (capítulo 22) mete el tiempo dentro del modelo: cadenas de Markov, movimiento browniano y teoría de colas.

Parte IV — Álgebra abstracta. Aquí el libro cambia de naturaleza: deja de calcular con objetos y empieza a estudiar estructuras. **Teoría de grupos desde cero** (capítulo 23) enseña la simetría con un triángulo de cartón antes de escribir un solo axioma. **Anillos, cuerpos y polinomios desde cero** (capítulo 24) hace con dos operaciones lo que el anterior hizo con una: empieza con la tabla de multiplicar de $\mathbb{Z}_6$ y termina construyendo $\mathbb{C}$ y los cuerpos finitos con las manos. **Estructuras algebraicas** (capítulo 25) lo sistematiza todo y añade acciones de grupo, Sylow, módulos y extensiones. **Teoría de Galois** (capítulo 26) es la segunda cumbre: por qué la ecuación de quinto grado no se resuelve con radicales. Y **teoría de la representación** (capítulo 27) convierte los grupos en matrices, que es como se estudian de verdad. Esta parte va *antes* del análisis a propósito: la topología necesita saber qué es un grupo para hablar del grupo fundamental.

Parte V — Análisis. Lo que ocurre cuando se toma en serio la pregunta "¿qué es de verdad un límite, y en qué espacios vive?". **Topología** (capítulo 28) aísla la noción de cercanía sin usar distancias. **Análisis real** (capítulo 29) vuelve sobre los cimientos con lupa y es el capítulo de las patologías: el conjunto de Cantor, la escalera del diablo, una función continua sin derivada en ningún punto y el teorema de Baire. **Variable compleja** (capítulo 30) muestra que derivar una vez sobre $\mathbb{C}$ obliga a ser infinitamente derivable, y de ahí salen los residuos y los mapas conformes. **La integral de Lebesgue** (capítulo 31) reconstruye la integral desde la medida. **Análisis funcional** (capítulo 32) hace álgebra lineal en dimensión infinita, con Banach, Hilbert y el teorema espectral. **Análisis armónico** (capítulo 33) profundiza en Fourier, la convolución y el teorema de muestreo. Y **teoría analítica de números** (capítulo 34) cierra la parte usando el análisis complejo para contar primos, que es una de las ideas más bellas que existen.

Parte VI — Geometría, topología y abstracción. **Geometría algebraica** (capítulo 35) traduce sistemas de ecuaciones polinómicas en objetos geométricos y viceversa. **Cálculo tensorial** (capítulo 36) es la tercera cumbre: generaliza vectores y matrices a objetos que viven en espacios curvos, y es el lenguaje de la relatividad general. **Topología algebraica** (capítulo 37) cuenta agujeros con grupos. Y **teoría de categorías** (capítulo 38) cierra el libro donde tiene que cerrarlo: es el capítulo que mira hacia atrás y descubre que grupos, anillos, espacios vectoriales y espacios topológicos eran ejemplos de una misma cosa. Leído antes sería una colección de cajas y flechas vacías; leído al final, es el resumen de todo.

Consejos prácticos

Haz los ejercicios antes de mirar las soluciones. La matemática se aprende con las manos. Las soluciones están completas precisamente para que puedas comparar tu razonamiento, no para sustituirlo.

No memorices demostraciones: memoriza ideas. Cada demostración incluida aquí tiene una o dos ideas centrales; el resto es rutina. Las cajas de *Idea clave* señalan lo que de verdad hay que retener.

Dosifica el material. Un ritmo realista es: la Parte I en cuatro o seis semanas; la Parte II en tres o cuatro meses, que es la más larga y la más exigente; la Parte III en seis semanas; la Parte IV en dos meses; la Parte V en tres o cuatro; y la Parte VI al ritmo que pida cada capítulo. Nadie lee este libro de corrido, y no está pensado para eso: cada capítulo abre con una caja que dice exactamente qué hace falta traer y de qué capítulo viene, de modo que se puede entrar por donde interese y retroceder sólo lo justo. **Complementa con los libros del apéndice.** Esta guía es autocontenida en lo esencial, pero ningún documento sustituye a los grandes textos; el apéndice final recomienda, tema por tema, dónde profundizar.

Parte I

Fundamentos: lógica, números y estructuras discretas

Capítulo 1

Conjuntos, lógica y cómo se demuestra algo

Lo que necesitas traer

Nada. Éste es el punto de partida del libro: sólo hace falta saber contar y estar dispuesto a leer despacio. Todo lo que aparece aquí se define desde cero, y todos los capítulos posteriores dan por sabido lo que se cuenta en éste: el lenguaje de conjuntos, los cuantificadores, las funciones inyectivas y sobreyectivas, las relaciones de equivalencia y los cuatro métodos de demostración (directa, contrarrecíproca, reducción al absurdo e inducción).

Este es el capítulo cero del libro: el que explica el idioma en el que están escritos todos los demás. En la escuela se enseña a *calcular*; en la matemática seria se enseña a *demostrar*, y el salto entre las dos cosas es el que hace que mucha gente capaz abandone. No es un salto de dificultad técnica, es un salto de género literario. Aquí aprenderás a leer un enunciado que empieza con "sea", a distinguir un teorema de su recíproco, a negar una frase con dos cuantificadores sin equivocarte y a reconocer cuándo un argumento está completo. Todo lo que viene después —el álgebra del álgebra elemental, los límites, y mucho más adelante la teoría de Galois del capítulo 26— usa este idioma sin volver a explicarlo.

Idea clave

Intuición primero. Una demostración no es una ceremonia ni un requisito burocrático: es la única manera que tenemos de estar *seguros*. En física uno cree algo porque el experimento salió bien mil veces; en matemática eso no basta, porque hay afirmaciones que se cumplen para los primeros cuarenta números y fallan en el cuarenta y uno. Una demostración es una cadena de pasos en la que cada eslabón es indiscutible, de modo que si aceptas el punto de partida estás obligado a aceptar el final. Y "obligado" es la palabra: la matemática es el único lugar donde una discusión termina de verdad.

La segunda idea del capítulo es igual de importante: casi todos los objetos de este libro —números, funciones, vectores, grupos, espacios topológicos— son *conjuntos con estructura*. Aprender el lenguaje de conjuntos no es un preámbulo, es aprender de qué

están hechas todas las cosas que vienen.

1.1. Conjuntos: el material de construcción

Definición 1.1. Un *conjunto* es una colección de objetos, llamados sus *elementos*. Escribimos $x \in A$ para decir que x es elemento de A, y $x \notin A$ en caso contrario. Dos conjuntos son iguales cuando tienen exactamente los mismos elementos.

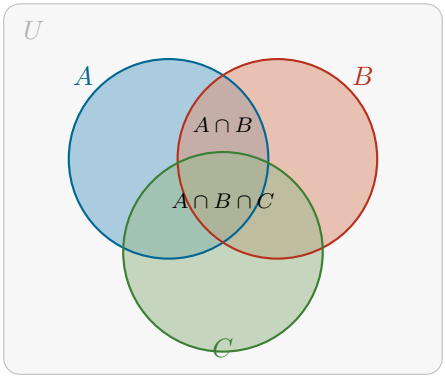


Figura 1.1: Tres conjuntos dividen el universo en ocho regiones, una por cada respuesta posible a las tres preguntas. Ese ocho es 2^3, y es la misma cuenta que da el número de filas de una tabla de verdad con tres proposiciones.

Hay dos maneras de describir un conjunto: por *extensión*, listando sus elementos, como $A = \{1, 2, 3\}$; y por *comprensión*, dando la propiedad que los caracteriza, como $A = \{n \in \mathbb{N} : 1 \leq n \leq 3\}$, que se lee "el conjunto de los n naturales tales que $1 \leq n \leq 3$". Los dos puntos (a veces se escribe una barra $|$) se leen siempre "tal que".

Cuidado

El orden y la repetición no significan nada dentro de las llaves: $\{1, 2, 3\} = \{3, 1, 2\} = \{1, 1, 2, 3\}$. Un conjunto solo sabe *quién* está, no cuántas veces ni en qué posición. Cuando el orden sí importa se usan paréntesis, y entonces hablamos de un *par ordenado* o de una *tupla*: $(1, 2) \neq (2, 1)$. Confundir $\{a, b\}$ con (a, b) es un error que arruina definiciones enteras más adelante.

Definición 1.2. A es *subconjunto* de B, escrito $A \subseteq B$, si todo elemento de A está en B. El *conjunto vacío* $\varnothing$ no tiene elementos y es subconjunto de todos. El *conjunto potencia* $\mathcal{P}(A)$ es el conjunto de *todos* los subconjuntos de A.

Ejemplo 1.3. Si $A = \{a, b\}$, entonces $\mathcal{P}(A) = \{\varnothing, \{a\}, \{b\}, \{a, b\}\}$: cuatro elementos. En general, si A tiene n elementos, $\mathcal{P}(A)$ tiene 2^n, porque construir un subconjunto es decidir, para cada elemento, si entra o no entra: n decisiones binarias independientes. Es el principio multiplicativo del capítulo 5 en su forma más limpia.

Cuidado

$\in$ y $\subseteq$ no son lo mismo y no se pueden intercambiar. Con $A = \{1, 2\}$: es verdad que $1 \in A$ y falso que $1 \subseteq A$; es verdad que $\{1\} \subseteq A$ y falso que $\{1\} \in A$. La pertenencia relaciona un *elemento* con un conjunto; la inclusión relaciona dos *conjuntos*. En el conjunto potencia la cosa se invierte: $\{1\} \in \mathcal{P}(A)$, porque los elementos de $\mathcal{P}(A)$ son conjuntos.

Regla: operaciones con conjuntos

1. *Unión:* $A \cup B = \{x : \ x \in A \text{ o } x \in B\}$ (el "o" es inclusivo: vale que esté en los dos).
2. *Intersección:* $A \cap B = \{x : \ x \in A \text{ y } x \in B\}$.
3. *Diferencia:* $A \setminus B = \{x : \ x \in A \text{ y } x \notin B\}$.
4. *Complemento* (dentro de un universo U): $A^c = U \setminus A$.
5. *Producto cartesiano:* $A \times B = \{(a, b) : \ a \in A, \ b \in B\}$, el conjunto de pares ordenados. El plano es $\mathbb{R} \times \mathbb{R} = \mathbb{R}^2$.

Teorema 1.4 (leyes de De Morgan)**.**

$$(A \cup B)^c = A^c \cap B^c, \qquad (A \cap B)^c = A^c \cup B^c.$$

Idea clave

De Morgan se recuerda con una frase: *negar reparte y voltea*. La negación entra en el paréntesis, se aplica a cada trozo y cambia el "o" por "y" (y viceversa). En castellano: lo contrario de "estudia matemática o física" es "ni estudia matemática ni estudia física" —hay que negar las dos—, mientras que lo contrario de "estudia matemática y física" es "le falta al menos una de las dos". Estas leyes aparecerán idénticas en lógica (negando "o" e "y"), en probabilidad (complementos de uniones, capítulo 5) y en topología (complementos de abiertos, capítulo 28). Es la misma ley tres veces.

Ejemplo 1.5. Sean $U = \{1, \ldots, 10\}$, $A = \{2, 4, 6, 8, 10\}$ (pares) y $B = \{3, 6, 9\}$ (múltiplos de 3).

$A \cup B = \{2, 3, 4, 6, 8, 9, 10\}$; $A \cap B = \{6\}$; $A \setminus B = \{2, 4, 8, 10\}$; $A^c = \{1, 3, 5, 7, 9\}$.

Comprobación de De Morgan: $(A \cap B)^c = \{6\}^c = U \setminus \{6\}$, que tiene nueve elementos; y $A^c \cup B^c = \{1, 3, 5, 7, 9\} \cup \{1, 2, 4, 5, 7, 8, 10\}$, que es también todo U menos el 6.

1.2. Las cinco conectivas con sus tablas completas, la negación paso a paso y el vocabulario de funciones

Un idioma no se aprende con una lista de palabras sueltas, y este capítulo maneja unas cuantas que hasta ahora se han usado sin haber sido presentadas del todo. La tabla de la sección anterior recoge cuatro conectivas y deja fuera la quinta, el "si y solo si", que sin

embargo aparece en casi todos los enunciados del libro. La negación de un "para todo" se ha enunciado como regla, pero negar de verdad una frase con tres cuantificadores encadenados es una operación mecánica que conviene practicar despacio. Y el símbolo $\circ$ de la composición, o el f^{-1} de la inversa, se han escrito sin decir nunca qué significan exactamente. Esta sección cierra esos huecos: no introduce ninguna idea nueva, le da forma exacta a lo que ya se estaba usando.

Las cinco conectivas, completas

Definición 1.6 (Las cinco conectivas y su tabla)**.** Dadas dos proposiciones p y q, las cinco conectivas básicas son la *negación* $\neg p$ ("no p"), la *conjunción* $p \wedge q$ ("p y q"), la *disyunción* $p \vee q$ ("p o q", inclusivo), la *implicación* $p \Rightarrow q$ ("si p, entonces q") y el *bicondicional* $p \Leftrightarrow q$ ("p si y solo si q"). Cada una queda definida por completo por la columna que le corresponde en esta tabla, y por nada más:

p	q	$\neg p$	$p \wedge q$	$p \vee q$	$p \Rightarrow q$	$p \Leftrightarrow q$
V	V	F	V	V	V	V
V	F	F	F	V	F	F
F	V	V	F	V	V	F
F	F	V	F	F	V	V

Las cuatro filas son los cuatro casos posibles, y la tabla no deja ninguno sin respuesta: por eso una conectiva es un objeto matemático y no una opinión sobre el castellano.

Observación 1.7. Las cuatro primeras columnas ya aparecían en la tabla de la sección anterior; lo único nuevo es la quinta. El bicondicional vale V exactamente cuando p y q tienen el *mismo* valor de verdad, sea cual sea ese valor. Por eso "si y solo si" se lee también como "p y q son equivalentes": no dice que ambas sean verdaderas, dice que van juntas.

Ejemplo 1.8 (Una fórmula evaluada fila por fila)**.** Comprobemos que $(p \Rightarrow q) \Leftrightarrow (\neg q \Rightarrow \neg p)$ es una tautología, es decir, que la implicación y su contrarrecíproco dicen lo mismo. Se rellenan las columnas auxiliares y se compara:

p	q	$\neg p$	$\neg q$	$p \Rightarrow q$	$\neg q \Rightarrow \neg p$	bicondicional
V	V	F	F	V	V	V
V	F	F	V	F	F	V
F	V	V	F	V	V	V
F	F	V	V	V	V	V

Las columnas quinta y sexta son idénticas, luego la última da V en las cuatro filas. Esa identidad de columnas *es* la equivalencia lógica, y es exactamente lo que autoriza el método de demostración por contrarrecíproco.

En la misma tabla se ve, de paso, que el recíproco $q \Rightarrow p$ y el contrario $\neg p \Rightarrow \neg q$ tienen ambos la columna V, V, F, V, distinta de la de $p \Rightarrow q$, que es V, F, V, V. Difieren precisamente en las filas segunda y tercera: ahí está, con nombres y apellidos, la razón de que confundir un teorema con su recíproco sea un error y no un matiz.

Cuidado

[Por qué una implicación con antecedente falso es verdadera] Ésta es la casilla que a todo el mundo le rechina, y merece la pena entender por qué el convenio no es arbitrario, sino el *único* razonable. El punto de partida no se discute: si p es verdadera, la implicación debe valer lo mismo que q, porque "si p entonces q" promete que de p se pasa a q. Eso fija las dos primeras filas, V y F. La pregunta es qué poner en las dos filas con p falsa, y solo hay cuatro maneras de rellenarlas. Veámoslas todas:

- Poner F y F da la columna V, F, F, F, que es la de $p \wedge q$: tendríamos una conectiva nueva que resulta ser la conjunción disfrazada, y "si llueve me mojo" significaría "llueve y me mojo", lo que convertiría en falso el enunciado los días de sol.
- Poner F y V da la columna V, F, F, V, que es la del bicondicional: "si llueve me mojo" significaría además que si no llueve no me mojo, y eso no lo dice nadie.
- Poner V y F da la columna V, F, V, F, que es sencillamente la columna de q: la implicación no dependería de p en absoluto, y el "si" sobraría.
- Poner V y V da V, F, V, V, la única que no coincide con ninguna conectiva ya existente, y es la implicación material que se usa.

El argumento decisivo, sin embargo, es otro y es matemático. Considérese el enunciado, obviamente verdadero, "todo múltiplo de 4 es par", que en símbolos es $\forall n\,(4 \mid n \Rightarrow 2 \mid n)$. Un "para todo" es verdadero cuando lo son todos sus casos, y entre esos casos está $n = 6$, donde el antecedente $4 \mid 6$ es falso. Si decidiéramos que una implicación con antecedente falso es falsa, ese caso sería falso y con él todo el enunciado, de modo que "todo múltiplo de 4 es par" resultaría ser mentira. El convenio no es un capricho: es lo que hace que las afirmaciones universales de la matemática entera signifiquen lo que queremos que signifiquen. Dicho de otro modo, $p \Rightarrow q$ solo promete que *no ocurrirá* el caso "p verdadera y q falsa"; si p no se cumple, la promesa sigue intacta. Los enunciados así, verdaderos por no tener ningún caso que los pueda romper, se llaman *vacuamente ciertos*, y son la razón de que "todo elemento del conjunto vacío es rojo" sea verdad.

Negar un enunciado con cuantificadores, paso a paso

Regla: negar un enunciado cuantificado

1. **Escribe el enunciado en símbolos**, con todos los cuantificadores explícitos y en el orden en que aparecen. Si la frase dice "para todo $\varepsilon > 0$", recuerda que eso abrevia $\forall \varepsilon\,(\varepsilon > 0 \Rightarrow \cdots)$, y si dice "existe $\delta > 0$", abrevia $\exists \delta\,(\delta > 0 \wedge \cdots)$.
2. **Pon un $\neg$ delante de todo** y no toques nada más todavía.
3. **Empuja el $\neg$ hacia dentro**, un cuantificador cada vez y de izquierda a derecha, con la regla ya conocida: $\neg\forall$ pasa a $\exists\neg$ y $\neg\exists$ pasa a $\forall\neg$. Los cuantificadores conservan su orden; solo cambian de tipo.
4. **Al llegar al núcleo sin cuantificadores, aplica De Morgan** y la negación de la

implicación: $\neg(A \wedge B) \equiv \neg A \vee \neg B$, $\neg(A \vee B) \equiv \neg A \wedge \neg B$ y $\neg(A \Rightarrow B) \equiv A \wedge \neg B$. Las desigualdades se niegan sin miedo: lo contrario de $<$ es $\geq$, y lo contrario de $=$ es $\neq$.

5. **Deja intactas las restricciones de rango.** Un "para todo $x > 0$" negado da "existe $x > 0$ tal que no..."; la condición $x > 0$ no se niega, porque estaba dentro del antecedente de una implicación y De Morgan la deja como está. Éste es el paso donde más gente se equivoca.
6. **Traduce el resultado al castellano** y compruébalo en voz alta: cada $\exists$ pide exhibir un caso, cada $\forall$ habla de todos.
7. **Control final:** el enunciado y su negación no pueden ser los dos verdaderos ni los dos falsos. Prueba un ejemplo concreto y verifica que exactamente uno de los dos se cumple.

Ejemplo 1.9 (Cinco negaciones resueltas enteras)**. (1) "Todo número primo es impar".** En símbolos, con $P(n) =$ "n es primo" e $I(n) =$ "n es impar":

$$\forall n\,(P(n) \Rightarrow I(n)) \;\xrightarrow{\neg}\; \exists n\,(P(n) \wedge \neg I(n)).$$

El $\forall$ se ha vuelto $\exists$ y la implicación se ha vuelto conjunción. En castellano: "existe un número primo que es par". Control: la negación es verdadera, porque $n = 2$ lo cumple; luego el enunciado original es falso, como debe ser.

(2) "Todo alumno aprobó algún examen". Con $A(a, e) =$ "el alumno a aprobó el examen e":

$$\forall a\,\exists e\; A(a, e) \;\xrightarrow{\neg}\; \exists a\,\forall e\; \neg A(a, e).$$

Dos cuantificadores, dos cambios, y el orden se respeta. En castellano: "hay un alumno que suspendió todos los exámenes". Nótese lo que *no* dice la negación: no dice que todos los alumnos suspendieran algo, ni que hubiera un examen que suspendieran todos. Basta un alumno, y ese alumno tiene que fallarlos todos.

(3) "Existe un real x tal que $xy = y$ para todo y".

$$\exists x\,\forall y\; (xy = y) \;\xrightarrow{\neg}\; \forall x\,\exists y\; (xy \neq y).$$

En castellano: "sea cual sea x, hay algún y al que multiplicar por x lo cambia". Control: el original es verdadero, con $x = 1$, porque $1 \cdot y = y$ para todo y; luego la negación es falsa, y en efecto para $x = 1$ no existe ningún y con $1 \cdot y \neq y$.

(4) "f es inyectiva". La definición es

$$\forall x\,\forall y\; (f(x) = f(y) \Rightarrow x = y) \;\xrightarrow{\neg}\; \exists x\,\exists y\; (f(x) = f(y) \wedge x \neq y).$$

Se han cambiado los dos cuantificadores y se ha deshecho la implicación. En castellano: "hay dos puntos distintos con la misma imagen". Esto es exactamente lo que uno hace para probar que una función *no* es inyectiva: exhibir la pareja. Con $f(x) = x^2$ sobre $\mathbb{R}$, la pareja es $x = 2$, $y = -2$, con $f(2) = f(-2) = 4$ y $2 \neq -2$.

(5) "f es uniformemente continua". Aquí hay tres cuantificadores y dos restricciones de rango:

$$\forall \varepsilon > 0\ \exists \delta > 0\ \forall x\, \forall y\ (|x-y| < \delta \Rightarrow |f(x) - f(y)| < \varepsilon).$$

Negando de izquierda a derecha, y usando en el último paso que $\neg(A \Rightarrow B) \equiv A \wedge \neg B$:

$$\exists \varepsilon > 0\ \forall \delta > 0\ \exists x\, \exists y\ (|x-y| < \delta\ \wedge\ |f(x) - f(y)| \geq \varepsilon).$$

Obsérvese el paso 5 de la regla en acción: "$\varepsilon > 0$" y "$\delta > 0$" siguen ahí sin negar, porque son restricciones de rango y no parte del núcleo. En castellano: hay un margen ε imposible de garantizar, porque por pequeño que se elija δ siempre aparecen dos puntos a distancia menor que δ cuyas imágenes se separan al menos ε. Ésa es, literalmente, la manera de demostrar que una función no es uniformemente continua. La negación de la definición de límite, que tiene la misma estructura, se hizo ya en la sección de cuantificadores.

Las operaciones de conjuntos como definición, y la doble inclusión

Definición 1.10 (Operaciones y igualdad de conjuntos)**.** Las cinco operaciones presentadas operativamente en la sección de conjuntos se recogen aquí con estatuto de definición, para poder citarlas. Dados A y B dentro de un universo U:

$$A \cup B = \{x : x \in A \text{ o } x \in B\}, \qquad A \cap B = \{x : x \in A \text{ y } x \in B\},$$

$$A \setminus B = \{x : x \in A \text{ y } x \notin B\}, \qquad A^c = U \setminus A, \qquad A \times B = \{(a,b) : a \in A,\ b \in B\}.$$

Además, $A = B$ significa que A y B tienen los mismos elementos, lo cual equivale a que se cumplan a la vez $A \subseteq B$ y $B \subseteq A$. Dos conjuntos son *disjuntos* si $A \cap B = \varnothing$.

Regla: demostrar que dos conjuntos son iguales

1. **Parte la igualdad en dos inclusiones:** $A \subseteq B$ y $B \subseteq A$. Casi nunca se demuestran a la vez, y querer hacerlo es la causa habitual de los argumentos confusos.
2. **Para probar $A \subseteq B$, empieza siempre igual:** "sea $x \in A$". Ese x es arbitrario, así que lo que se concluya valdrá para todos.
3. **Traduce la pertenencia a lógica** usando la definición de las operaciones: $x \in A \cup B$ se convierte en "$x \in A$ o $x \in B$", y a partir de ahí se trabaja con conectivas.
4. **Termina en "luego $x \in B$"**, y solo entonces empieza la otra inclusión.
5. **Si los pasos son todos reversibles** —cadenas de "si y solo si"— puedes hacer las dos inclusiones de golpe, pero dilo explícitamente; si alguno no lo es, no vale.
6. **Para refutar una igualdad basta un elemento** que esté en uno y no en el otro.

Ejemplo 1.11 (Una igualdad demostrada por doble inclusión)**.** Probemos que $A \setminus (B \cap C) = (A \setminus B) \cup (A \setminus C)$.

$\subseteq$**.** Sea $x \in A \setminus (B \cap C)$. Entonces $x \in A$ y $x \notin B \cap C$. Que x no esté en $B \cap C$ significa que es falso que "$x \in B$ y $x \in C$", y por De Morgan eso equivale a "$x \notin B$ o $x \notin C$". En el

primer caso $x \in A$ y $x \notin B$, luego $x \in A \setminus B$; en el segundo, $x \in A \setminus C$. En cualquiera de los dos casos x pertenece a la unión.

$\supseteq$. Sea $x \in (A \setminus B) \cup (A \setminus C)$. Entonces $x \in A \setminus B$ o $x \in A \setminus C$. En ambos casos $x \in A$. Y en el primero $x \notin B$, luego no puede estar en $B \cap C$; en el segundo $x \notin C$, y tampoco. Así que $x \in A$ y $x \notin B \cap C$, es decir, $x \in A \setminus (B \cap C)$.

Como valen las dos inclusiones, los conjuntos son iguales. ■ Comprobación numérica con $A = \{1, 2, 3, 4\}$, $B = \{2, 3\}$, $C = \{3, 4\}$: $B \cap C = \{3\}$, luego el lado izquierdo es $\{1, 2, 4\}$; y $A \setminus B = \{1, 4\}$, $A \setminus C = \{1, 2\}$, cuya unión es $\{1, 2, 4\}$. Coinciden.

Imagen, composición, inversa

Definición 1.12 (Imagen, preimagen, identidad, composición e inversa)**.** Sea $f : A \to B$ una función.

- La *imagen* de un subconjunto $S \subseteq A$ es $f(S) = \{f(x) : x \in S\}$; la *imagen de* f es $f(A)$, y f es sobreyectiva justamente cuando $f(A) = B$.
- La *preimagen* de $T \subseteq B$ es $f^{-1}(T) = \{x \in A : f(x) \in T\}$. Ojo: esta notación tiene sentido siempre, exista o no la función inversa.
- La *identidad* de A es la función $\mathrm{id}_A : A \to A$ dada por $\mathrm{id}_A(x) = x$.
- Si $f : A \to B$ y $g : B \to C$, la *composición* $g \circ f : A \to C$ se define por

$$(g \circ f)(x) = g(f(x)).$$

  Se lee "g compuesta con f" y se aplica *de derecha a izquierda*: primero f, después g. Para que la composición tenga sentido, el codominio de f debe ser el dominio de g.
- f es *invertible* si existe $g : B \to A$ con $g \circ f = \mathrm{id}_A$ y $f \circ g = \mathrm{id}_B$; esa g, cuando existe, es única y se llama la *inversa* de f y se escribe f^{-1}.

Ejemplo 1.13 (Composición e inversa con números)**.** Sean $f : \mathbb{R} \to \mathbb{R}$ con $f(x) = 2x + 3$ y $g : \mathbb{R} \to \mathbb{R}$ con $g(x) = x^2$.

(a) Las dos composiciones. $(g \circ f)(x) = g(2x+3) = (2x+3)^2 = 4x^2 + 12x + 9$, mientras que $(f \circ g)(x) = f(x^2) = 2x^2 + 3$. En $x = 1$ la primera vale 25 y la segunda 5: la composición **no es conmutativa**, y el orden en que se escribe importa tanto como en "ponerse los calcetines y luego los zapatos".

(b) La inversa de f. Despejando $y = 2x + 3$ se obtiene $x = (y - 3)/2$, luego $f^{-1}(y) = (y - 3)/2$. Comprobación en las dos direcciones, que es lo que exige la definición: $f^{-1}(f(x)) = ((2x + 3) - 3)/2 = x$ y $f(f^{-1}(y)) = 2 \cdot \frac{y-3}{2} + 3 = y$. Las dos salen la identidad, luego f es invertible y f^{-1} es la inversa.

(c) Una que no tiene inversa. $g(x) = x^2$ sobre $\mathbb{R}$ no es inyectiva ni sobreyectiva, así que no hay ninguna h con $h \circ g = \mathrm{id}_{\mathbb{R}}$: si la hubiera, de $g(2) = g(-2)$ se seguiría $2 = h(g(2)) = h(g(-2)) = -2$. Sin embargo la preimagen sigue existiendo como conjunto: $g^{-1}(\{4\}) = \{2, -2\}$ y $g^{-1}(\{-1\}) = \varnothing$. Compárese con el ejemplo de las cuatro funciones x^2 de la sección de funciones: restringiendo dominio y codominio a $[0, \infty)$ sí hay inversa, la raíz cuadrada.

Proposición 1.14. *Si $f : A \to B$ y $g : B \to C$ son inyectivas, $g \circ f$ es inyectiva. Si ambas son sobreyectivas, $g \circ f$ es sobreyectiva. En consecuencia, la composición de biyecciones es una biyección.*

Demostración. Inyectividad: supongamos $(g \circ f)(x) = (g \circ f)(x')$, es decir $g(f(x)) = g(f(x'))$. Como g es inyectiva, $f(x) = f(x')$; y como f lo es, $x = x'$. Sobreyectividad: sea $c \in C$. Como g es sobreyectiva, existe $b \in B$ con $g(b) = c$; y como f lo es, existe $a \in A$ con $f(a) = b$. Entonces $(g \circ f)(a) = g(b) = c$. □

Observación 1.15. Merece la pena retener el recíproco parcial, que se usa a menudo: si $g \circ f$ es inyectiva, entonces f lo es (aunque g no tiene por qué serlo), y si $g \circ f$ es sobreyectiva, entonces g lo es. La primera sale igual de rápido: si $f(x) = f(x')$, aplicando g se obtiene $(g \circ f)(x) = (g \circ f)(x')$ y por hipótesis $x = x'$.

Clases, cociente y órdenes: los nombres exactos

Definición 1.16 (Clase, cociente, proyección y orden total)**.** Sea $\sim$ una relación de equivalencia sobre A.

- La *clase de equivalencia* de $a \in A$ es $[a] = \{x \in A : x \sim a\}$.
- El *conjunto cociente* $A/\sim$ es el conjunto cuyos elementos son las clases: $A/\sim = \{[a] : a \in A\}$. Sus elementos son *conjuntos*, no elementos de A.
- La *proyección canónica* es la función $\pi : A \to A/\sim$ dada por $\pi(a) = [a]$; es siempre sobreyectiva.

Un orden parcial $\leq$ sobre A es un *orden total* si además dos elementos cualesquiera son comparables: para todos $a, b \in A$ se cumple $a \leq b$ o $b \leq a$. Cuando existen elementos incomparables el orden es estrictamente parcial.

Ejemplo 1.17 (El cociente $\mathbb{Z}/5\mathbb{Z}$, entero)**.** Sobre $\mathbb{Z}$, sea $a \sim b$ cuando $a - b$ es múltiplo de 5. Ya sabemos por la sección de relaciones que esto es de equivalencia. Las clases son cinco:

$$[0] = \{\ldots, -5, 0, 5, 10, \ldots\}, \quad [1] = \{\ldots, -4, 1, 6, \ldots\}, \quad [2], \quad [3], \quad [4],$$

y $\mathbb{Z}/5\mathbb{Z} = \{[0], [1], [2], [3], [4]\}$, un conjunto con exactamente cinco elementos, cada uno de ellos infinito. La proyección envía $\pi(17) = [2]$, porque $17 - 2 = 15$ es múltiplo de 5.

La suma está bien definida. Se quiere definir $[a] + [b] = [a + b]$, y hay que comprobar que el resultado no depende de los representantes elegidos. Supongamos $[a] = [a']$ y $[b] = [b']$, es decir $a - a' = 5k$ y $b - b' = 5m$. Entonces

$$(a + b) - (a' + b') = (a - a') + (b - b') = 5(k + m),$$

luego $[a + b] = [a' + b']$. Comprobación con cifras: $[3] + [4] = [7] = [2]$, y eligiendo otros representantes, $[8] + [9] = [17] = [2]$. Sale lo mismo, como tenía que salir. Ésta es la comprobación de "estar bien definido" que el capítulo 23 exigirá en cada construcción cociente, y no es una formalidad: si la operación dependiera del representante, sencillamente no existiría.

Cuidado

Tres confusiones de notación que cuestan caras. La primera: $f^{-1}(T)$ con T un conjunto es siempre la preimagen y existe siempre; $f^{-1}(y)$ con y un punto solo tiene sentido si f es biyectiva. La segunda: en $g \circ f$ se aplica primero la función de la derecha, aunque se escriba la segunda. La tercera: $[a]$ es un subconjunto de A y un elemento de $A/\sim$ al mismo tiempo, igual que ocurría con $\{1\} \in \mathcal{P}(A)$; no hay contradicción, solo dos niveles.

1.3. Lógica: proposiciones y conectivos

Definición 1.18. Una *proposición* es un enunciado que es verdadero o falso, pero no ambas cosas. "7 es primo" lo es; "$x > 3$" no lo es mientras no se diga quién es x (es un *predicado*); "¿qué hora es?" no lo es nunca.

Los conectivos se definen por su tabla de verdad, que dice qué pasa en cada uno de los casos posibles:

p	q	$\neg p$	$p \wedge q$	$p \vee q$	$p \Rightarrow q$
V	V	F	V	V	V
V	F	F	F	V	F
F	V	V	F	V	V
F	F	V	F	F	V

Cuidado

La cuarta columna es la que a todo el mundo le cuesta: *una implicación con hipótesis falsa es verdadera.* "Si $2 + 2 = 5$, entonces yo soy el rey de Francia" es un enunciado *verdadero*. No es un truco: la implicación $p \Rightarrow q$ solo promete que no ocurrirá el caso "p verdadera y q falsa"; si p no se cumple, la promesa no se ha roto. Los matemáticos llaman a esto *vacuamente cierto*, y es lo que hace que "todo elemento del conjunto vacío es rojo" sea verdad.

Regla: las cuatro formas de una implicación

Dado $p \Rightarrow q$ ("si p entonces q"):

1. El *recíproco* es $q \Rightarrow p$. **No es equivalente.**
2. El *contrario* es $\neg p \Rightarrow \neg q$. **Tampoco es equivalente.**
3. El *contrarrecíproco* es $\neg q \Rightarrow \neg p$. **Sí es equivalente**, siempre.
4. $p \Leftrightarrow q$ ("si y solo si") significa las dos: $p \Rightarrow q$ *y* $q \Rightarrow p$.

Ejemplo 1.19. Sea la afirmación verdadera: "si llueve, el suelo se moja".
Recíproco: "si el suelo está mojado, llovió". Falso: pudo pasar el camión del riego.
Contrario: "si no llueve, el suelo no se moja". Falso, por lo mismo.

Contrarrecíproco: "si el suelo no está mojado, no llovió". Verdadero, y dice *exactamente* lo mismo que el original.

Idea clave

La equivalencia entre una implicación y su contrarrecíproco es una herramienta de demostración de primera categoría, no una curiosidad. Cuando "si p entonces q" es difícil de atacar de frente, muchas veces "si no q entonces no p" es casi trivial. La usaremos, por ejemplo, para probar que si n^2 es par entonces n es par: de frente hay que factorizar n^2; por contrarrecíproco basta escribir $n = 2k + 1$ y elevar al cuadrado.

Cuidado

Confundir un teorema con su recíproco es el error lógico más caro que existe, y no solo en matemática. "Todo cuadrado es un rectángulo" no dice que todo rectángulo sea un cuadrado. "Si una función es derivable, es continua" no dice que toda función continua sea derivable —de hecho es falso, y la función de Weierstrass del capítulo 29 es continua en todas partes y derivable en ninguna—. Cuando un enunciado sea de doble sentido, el libro dirá explícitamente *si y solo si*.

1.4. Tautologías, argumentos válidos y falacias

Una tabla de verdad no sirve solo para definir los conectivos: sirve para decidir, sin discutir con nadie, si un razonamiento es correcto. Esta sección convierte la lógica en una herramienta de control.

Definición 1.20. Una fórmula es una *tautología* si su tabla de verdad da V en todas las filas, sea cual sea el valor de las proposiciones que la componen; es una *contradicción* si da F en todas; y es *contingente* si a veces da V y a veces F. Dos fórmulas son *lógicamente equivalentes* cuando su bicondicional es una tautología, es decir, cuando tienen exactamente la misma columna final.

Idea clave

Un argumento no es más que una lista de premisas y una conclusión. Decir que el argumento es *válido* es decir que la fórmula "(premisas) $\Rightarrow$ (conclusión)" es una tautología: no existe ni una sola fila donde las premisas sean todas verdaderas y la conclusión falsa. Ojo con la palabra: válido no significa cierto. Un argumento válido con premisas falsas puede llevar a una conclusión falsa sin dejar de ser válido. La validez es una propiedad de la *forma*, no del contenido.

Regla: comprobar si un argumento es válido

1. **Extrae las proposiciones atómicas** y ponles nombre: p, q, r. Todo lo demás son conectivos disfrazados de lenguaje corriente.
2. **Traduce cada premisa y la conclusión** a una fórmula. "Solo si" introduce el consecuente; "a menos que" es un $\vee$; "ni... ni" es $\neg p \wedge \neg q$.
3. **Escribe la fórmula del argumento:** $(P_1 \wedge P_2 \wedge \cdots \wedge P_n) \Rightarrow C$.
4. **Construye la tabla completa,** con 2^k filas si hay k proposiciones atómicas: 4 filas con dos letras, 8 con tres, 16 con cuatro.
5. **Busca solo las filas críticas:** aquellas en las que todas las premisas valen V. Si en todas ellas la conclusión también vale V, el argumento es válido.
6. **Si aparece una fila con premisas V y conclusión F,** esa fila *es* el contraejemplo: el argumento no es válido y ya sabes en qué situación falla.
7. **Como atajo, reconoce la forma.** Las cuatro que hay que saber de memoria van en el ejemplo siguiente: dos son válidas siempre y dos son falacias siempre.

Ejemplo 1.21. Las cuatro formas que se construyen con una implicación $p \Rightarrow q$ y un dato adicional.

1. *Modus ponens:* de $p \Rightarrow q$ y p se deduce q. Válido. "Si llueve el suelo se moja; llueve; luego el suelo se moja". La tabla de $((p \Rightarrow q) \wedge p) \Rightarrow q$ da V en las cuatro filas.
2. *Modus tollens:* de $p \Rightarrow q$ y $\neg q$ se deduce $\neg p$. Válido, y es el contrarrecíproco en acción: "el suelo no está mojado, luego no ha llovido". La tabla de $((p \Rightarrow q) \wedge \neg q) \Rightarrow \neg p$ también da V en las cuatro filas.
3. *Falacia de afirmar el consecuente:* de $p \Rightarrow q$ y q concluir p. **No es válido.** La fila $p = F$, $q = V$ tiene las dos premisas verdaderas y la conclusión falsa: el suelo está mojado porque pasó el camión del riego.
4. *Falacia de negar el antecedente:* de $p \Rightarrow q$ y $\neg p$ concluir $\neg q$. **Tampoco.** La misma fila $p = F$, $q = V$ la tumba. Estas dos falacias son, respectivamente, confundir la implicación con su recíproco y con su contrario.

Ejemplo 1.22. Cuatro equivalencias que conviene tener memorizadas, todas comprobables con una tabla de cuatro u ocho filas.

1. $p \Rightarrow q \;\equiv\; \neg p \vee q$. Esta es la que explica por qué una implicación con hipótesis falsa es verdadera: si p es falsa, $\neg p$ ya es verdadera y la disyunción se cumple sola.
2. Leyes de De Morgan: $\neg(p \wedge q) \equiv \neg p \vee \neg q$ y $\neg(p \vee q) \equiv \neg p \wedge \neg q$. Negar una conjunción da una disyunción, y viceversa. Negar "es alto y rubio" es "o no es alto o no es rubio", no "no es alto y no es rubio".

3. Negación de la implicación: $\neg(p \Rightarrow q) \equiv p \wedge \neg q$. La negación de una implicación *no* es otra implicación: es una conjunción. Para refutar "si es primo mayor que 2, es impar" hay que exhibir un primo mayor que 2 que sea par.

4. Distributivas: $p \wedge (q \vee r) \equiv (p \wedge q) \vee (p \wedge r)$ y $p \vee (q \wedge r) \equiv (p \vee q) \wedge (p \vee r)$, con sus ocho filas cada una. Nótese que aquí $\wedge$ y $\vee$ se distribuyen *el uno sobre el otro*, cosa que en la aritmética no ocurre: el producto se distribuye sobre la suma, pero la suma no sobre el producto.

Cuidado

Validez y verdad son cosas distintas y hay que separarlas con cuidado. El argumento "todos los pájaros vuelan; el pingüino es un pájaro; luego el pingüino vuela" es perfectamente *válido*: la forma es impecable. Lo que falla es la primera premisa, que es falsa. Al revés, "la Tierra es redonda; luego $2 + 2 = 4$" tiene premisa y conclusión verdaderas y es un argumento pésimo, porque la conclusión no se sigue de la premisa. Cuando revises una demostración ajena, pregúntate las dos cosas por separado: ¿son ciertas las premisas?, ¿se sigue la conclusión?

Pausa para pensar

Toda tautología con k letras se comprueba mirando 2^k filas, así que en principio la lógica proposicional está *resuelta*: hay un algoritmo que decide cualquier fórmula. Ahora piensa en el coste: con $k = 10$ son 1024 filas, con $k = 30$ ya son más de mil millones, y con $k = 100$ el número de filas supera al de átomos de una ciudad. Decidir si una fórmula puede ser verdadera en alguna fila es el famoso problema SAT, y nadie conoce un método esencialmente mejor que probar. Que exista un algoritmo no significa que sea utilizable; esa distinción entre *computable* y *computable en un tiempo razonable* es la frontera que separa la lógica de la complejidad, y reaparece al hablar de criptografía en el capítulo 3.

1.5. Cuantificadores

Definición 1.23. $\forall$ se lee "para todo" (cuantificador universal) y $\exists$ se lee "existe" o "existe al menos un" (cuantificador existencial). Se escribe $\exists!$ para "existe un único".

Regla: negación de cuantificadores

$$\neg\big(\forall x\, P(x)\big) \equiv \exists x\, \neg P(x), \qquad \neg\big(\exists x\, P(x)\big) \equiv \forall x\, \neg P(x).$$

Es decir: *la negación intercambia los cuantificadores y niega el final.* Para negar "todos los cisnes son blancos" no hace falta decir que ninguno lo es: basta exhibir un cisne negro.

Ejemplo 1.24. Neguemos la definición de límite, que aparecerá formalmente en el capítulo 6:

$$\forall \varepsilon > 0\ \exists \delta > 0\ \forall x\ \big(0 < |x - a| < \delta \Rightarrow |f(x) - L| < \varepsilon\big).$$

Aplicando la regla de derecha a izquierda, y usando que $\neg(p \Rightarrow q) \equiv p \wedge \neg q$:

$$\exists \varepsilon > 0\ \forall \delta > 0\ \exists x\ \big(0 < |x - a| < \delta\ \wedge\ |f(x) - L| \geq \varepsilon\big).$$

En castellano: hay un margen de error ε que no se puede conseguir por más que se apriete δ, porque siempre queda algún x cerca de a cuya imagen se sale. Esa frase es literalmente la demostración de que un límite no existe, y sale sola de la regla.

Cuidado

El orden de los cuantificadores cambia el significado. Compara:

- $\forall x\ \exists y:\ y > x$. Verdadero en $\mathbb{R}$: para cada número hay otro mayor (el y puede depender del x).
- $\exists y\ \forall x:\ y > x$. Falso: afirmaría que hay un número mayor que todos.

Las mismas palabras en distinto orden pasan de verdad a falsedad. Toda la diferencia entre continuidad y continuidad *uniforme* (capítulo 29) es exactamente esto: si el δ puede depender del punto o no. Cuando leas una definición con dos cuantificadores, pregúntate siempre *quién depende de quién*.

1.6. Relaciones: equivalencia, particiones y orden

Ya sabemos formar el producto cartesiano $A \times B$. Un subconjunto cualquiera de $A \times A$ es lo que se llama una *relación* sobre A, y esa idea tan simple es la que sostiene tres construcciones que aparecen en todo el libro: los conjuntos cociente, las fracciones y los órdenes.

Definición 1.25. Una *relación binaria* sobre A es un subconjunto $R \subseteq A \times A$; se escribe $a \sim b$ en lugar de $(a, b) \in R$. La relación puede tener estas propiedades:

- *reflexiva*: $a \sim a$ para todo a;
- *simétrica*: si $a \sim b$ entonces $b \sim a$;
- *antisimétrica*: si $a \sim b$ y $b \sim a$ entonces $a = b$;
- *transitiva*: si $a \sim b$ y $b \sim c$ entonces $a \sim c$.

Una relación reflexiva, simétrica y transitiva es una *relación de equivalencia*. Una relación reflexiva, antisimétrica y transitiva es un *orden parcial*.

Idea clave

Una relación de equivalencia es la manera precisa de decir "para lo que me interesa ahora, estos objetos son el mismo". Al clasificar los enteros por su resto al dividir entre 3 no estamos diciendo que 4 y 7 sean iguales, sino que son intercambiables para todo lo que dependa del resto. El resultado de esa clasificación es una *partición*: el conjunto queda troceado en bloques que no se solapan y que entre todos lo cubren. La equivalencia y la partición son dos vistas del mismo objeto.

Regla: reconocer y usar una relación de equivalencia

1. **Comprueba las tres propiedades por separado,** en este orden: reflexiva, simétrica, transitiva. Si falla una sola, no es de equivalencia y no hay clases que valgan.
2. **Para refutar, basta un contraejemplo concreto;** para confirmar, hay que argumentar con elementos genéricos a, b, c.
3. **Define la clase de** a como $[a] = \{x \in A : x \sim a\}$, el conjunto de todo lo que es equivalente a a.
4. **Usa el hecho clave:** dos clases o son idénticas o son disjuntas, nunca se solapan a medias. Y $[a] = [b]$ si y solo si $a \sim b$.
5. **Elige un representante** de cada clase para trabajar con ella; el resultado no puede depender del representante elegido, y comprobar eso se llama ver que la operación *está bien definida*.
6. **Forma el conjunto cociente** $A/\sim$, cuyos elementos son las clases. Contar sus elementos es contar los bloques de la partición.
7. **Si en vez de simétrica la relación es antisimétrica,** no clasifiques: ordena. Dibuja el diagrama de Hasse poniendo abajo lo pequeño y uniendo con un trazo solo los pares que no se deducen por transitividad.

Ejemplo 1.26. Cuatro relaciones, y qué es cada una.

1. En $\mathbb{Z}$, $a \sim b$ cuando $a-b$ es múltiplo de 3. Es de equivalencia: $a-a=0$ es múltiplo de 3; si $a-b=3k$ entonces $b-a=3(-k)$; y si $a-b=3k$, $b-c=3m$, entonces $a-c=3(k+m)$. Sobre $\{0,1,\ldots,9\}$ las clases salen $[0]=\{0,3,6,9\}$, $[1]=\{1,4,7\}$, $[2]=\{2,5,8\}$: tres bloques disjuntos que cubren los diez números. El cociente $\mathbb{Z}/3\mathbb{Z}$ tiene exactamente tres elementos, y de él se ocupa el capítulo 3.
2. En el conjunto de las rectas del plano, "ser paralela a" (admitiendo que toda recta es paralela a sí misma) es de equivalencia, y cada clase es lo que en geometría se llama una *dirección*. En cambio "ser perpendicular a" no lo es: no es reflexiva ni transitiva, porque dos rectas perpendiculares a una tercera son paralelas entre sí, no perpendiculares.
3. En $\mathbb{Z} \times (\mathbb{Z} \setminus \{0\})$, el par (a,b) se relaciona con (c,d) cuando $ad=bc$. Es de equivalencia, y su conjunto cociente *es* $\mathbb{Q}$: la fracción $\frac{2}{3}$ no es un par de números, es la clase entera

$\{(2,3),(4,6),(-2,-3),\dots\}$. Esto es lo que justifica que simplificar una fracción no la cambie.

4. En $D=\{1,2,3,4,6,12\}$, la relación "a divide a b" es reflexiva, antisimétrica y transitiva: es un orden *parcial*, con 18 pares relacionados de los 36 posibles. No es un orden total, porque 2 y 3 son incomparables: ninguno divide al otro. En el diagrama de Hasse, 1 va abajo del todo y 12 arriba del todo.

Ejemplo 1.27. Cuántas maneras hay de clasificar, y por qué la cuenta importa.

1. Sobre un conjunto de 3 elementos hay $2^9=512$ relaciones posibles, porque hay $3\times 3=9$ pares y cada uno se incluye o no. De esas 512, muy pocas son de equivalencia.
2. El número de particiones de un conjunto de n elementos son los *números de Bell*: $1,2,5,15,52,203,877$ para $n=1,\dots,7$. Con tres elementos hay 5: todo junto; los tres separados; y tres maneras de juntar dos y dejar uno suelto. Como equivalencias y particiones son lo mismo, ese 5 cuenta también las relaciones de equivalencia sobre un conjunto de tres elementos.
3. La partición más fina es la de las clases unitarias, $[a]=\{a\}$, que corresponde a la igualdad estricta; la más gruesa es la de un solo bloque, que corresponde a la relación que lo relaciona todo con todo. Cualquier equivalencia está entre esas dos.
4. Un aviso de conteo: "ser primo entre sí" sobre $\{1,\dots,10\}$ es simétrica pero no transitiva (3 y 4 son primos entre sí, 4 y 9 también, pero 3 y 9 no lo son), y "$a<b$" es transitiva pero ni reflexiva ni simétrica: es un orden *estricto*. Faltando una propiedad, no hay clases ni cociente.

Cuidado

Hay un error clásico que consiste en creer que la simetría y la transitividad implican la reflexividad: "si $a\sim b$, por simetría $b\sim a$, y por transitividad $a\sim a$". El razonamiento tiene una grieta: solo funciona para los a que están relacionados con *alguien*. La relación vacía sobre un conjunto no vacío es simétrica y transitiva por vacuidad, y no es reflexiva. Por eso las tres condiciones se piden por separado y ninguna sobra. El otro error frecuente es confundir simétrica con antisimétrica: no son contrarias. La igualdad es las dos a la vez, y hay relaciones que no son ni una ni otra.

Pausa para pensar

Si toda equivalencia da una partición y toda partición da una equivalencia, ¿qué se gana con tener dos nombres? La respuesta es que cada punto de vista sugiere preguntas distintas. Como relación, uno pregunta si está bien definida una operación sobre las clases, y eso lleva directamente a los grupos cociente del capítulo 23 y a los anillos cociente del capítulo 24. Como partición, uno pregunta cuántos bloques hay y de qué tamaño, y eso lleva al conteo del capítulo 5. Es la misma matemática mirada desde dos ventanas: merece la pena acostumbrarse a saltar de una a otra.

1.7. Los métodos de demostración

Regla: demostración directa

Para probar $p \Rightarrow q$: supón p, deduce consecuencias, llega a q.

Ejemplo 1.28. *Si n es par, entonces n^2 es par.*

Supongamos n par: por definición $n = 2k$ con k entero. Entonces $n^2 = 4k^2 = 2(2k^2)$, que es dos veces un entero, luego es par. ■

Regla: demostración por contrarrecíproco

Para probar $p \Rightarrow q$, prueba $\neg q \Rightarrow \neg p$. Es la misma afirmación.

Ejemplo 1.29. *Si n^2 es par, entonces n es par.*

Probamos el contrarrecíproco: si n es impar, n^2 es impar. En efecto, $n = 2k + 1$ da

$$n^2 = 4k^2 + 4k + 1 = 2(2k^2 + 2k) + 1,$$

que es impar. ■

Regla: demostración por contradicción (reducción al absurdo)

Para probar p: supón $\neg p$ y deduce algo imposible. Como la falsedad no puede seguirse de la verdad, $\neg p$ era falsa.

Teorema 1.30. *$\sqrt{2}$ es irracional.*

Demostración. Supongamos que no: que $\sqrt{2} = p/q$ con p, q enteros sin factores comunes (siempre se puede simplificar hasta llegar ahí). Elevando al cuadrado, $2q^2 = p^2$, luego p^2 es par y, por el ejemplo anterior, p es par: $p = 2m$. Sustituyendo, $2q^2 = 4m^2$, es decir $q^2 = 2m^2$, de donde q también es par. Pero entonces p y q tienen el factor común 2, contra lo supuesto. La contradicción viene de haber supuesto que $\sqrt{2}$ era racional. □

Idea clave

Esta demostración tiene unos veinticinco siglos y sigue siendo el mejor ejemplo de lo que la matemática puede hacer y ninguna otra disciplina puede: establecer un hecho *negativo* con certeza absoluta. Ninguna cantidad de cálculo numérico probaría jamás que $\sqrt{2}$ no es una fracción —solo diría que no es ninguna de las que probamos—. Fue, además, un escándalo: los pitagóricos creían que todo era razón de enteros, y este argumento demolió su programa. La misma técnica, refinada, prueba que e y π son irracionales y, mucho más difícil, trascendentes (capítulo 25).

Regla: refutación por contraejemplo

Para probar que "$\forall x\, P(x)$" es *falsa* basta un único x con $P(x)$ falsa. Para probar que es *verdadera* no basta ninguna cantidad de ejemplos.

Ejemplo 1.31. El polinomio n^2+n+41 da un número primo para $n = 0, 1, 2, \ldots, 39$: cuarenta aciertos seguidos. En $n = 40$ falla, porque

$$40^2 + 40 + 41 = 1681 = 41^2.$$

Cuarenta comprobaciones no son una demostración. Un contraejemplo sí es una refutación.

Regla: inducción matemática

Para probar que $P(n)$ vale para todo $n \geq n_0$:

1. *Base:* comprobar $P(n_0)$.
2. *Paso inductivo:* suponer $P(k)$ para un $k \geq n_0$ arbitrario (*hipótesis de inducción*) y deducir $P(k+1)$.

Entonces $P(n)$ vale para todo $n \geq n_0$.

Idea clave

La imagen del dominó es correcta y conviene tomársela en serio: la base dice que la primera ficha cae; el paso inductivo dice que cada ficha, al caer, tira a la siguiente. Con las dos cosas caen todas, aunque sean infinitas. Lo que *no* se hace en el paso inductivo es suponer lo que se quiere probar: se supone $P(k)$ y se demuestra $P(k+1)$, que son cosas distintas. Sin la base, el argumento se cae entero —se puede "demostrar" inductivamente que todos los números son iguales si uno se salta la comprobación inicial—.

Ejemplo 1.32. *Para todo* $n \geq 1$*:* $1 + 2 + \cdots + n = \dfrac{n(n+1)}{2}$.

Base. Con $n = 1$: el lado izquierdo es 1 y el derecho $\frac{1\cdot 2}{2} = 1$. Correcto.

Paso. Supongamos que vale para k, es decir $1 + \cdots + k = \frac{k(k+1)}{2}$. Entonces

$$1 + \cdots + k + (k+1) = \frac{k(k+1)}{2} + (k+1) = (k+1)\left(\frac{k}{2} + 1\right) = \frac{(k+1)(k+2)}{2},$$

que es la fórmula con $k+1$ en lugar de k. ■

Ejemplo 1.33. *Para todo* $n \geq 1$*,* $n^3 - n$ *es múltiplo de* 6.

Base. $n = 1$: $1 - 1 = 0$, múltiplo de 6.

Paso. Supongamos $k^3 - k = 6m$. Entonces

$$(k+1)^3 - (k+1) = k^3 + 3k^2 + 3k + 1 - k - 1 = (k^3 - k) + 3k(k+1) = 6m + 3k(k+1).$$

Y $k(k+1)$ es producto de dos enteros consecutivos, luego es par: $3k(k+1)$ es múltiplo de 6. La suma de dos múltiplos de 6 lo es. ■

Observación 1.34. Existe una variante, la *inducción fuerte*, en la que para probar $P(k+1)$ se permite suponer $P(n_0), \dots, P(k)$ todos a la vez. Es lógicamente equivalente a la ordinaria y resulta imprescindible cuando el paso no salta de uno en uno: por ejemplo, para demostrar que todo entero mayor que 1 es producto de primos (capítulo 3), donde $n = ab$ obliga a invocar la hipótesis para a y b, que pueden estar muy por debajo de n.

1.8. Cómo se lee y se escribe una demostración

Regla: el vocabulario que hay que traducir

1. **"Sea x..."** anuncia que x es arbitrario: lo que se pruebe valdrá para todos. Es un $\forall$ disfrazado.
2. **"Existe"** obliga a exhibir uno o a demostrar que tiene que haberlo (por ejemplo con el principio del palomar del capítulo 5).
3. **"Sin pérdida de generalidad"** significa que los casos restantes son simétricos y se omiten. Si no lo son, la frase es una trampa.
4. **"Basta ver que..."** reduce el objetivo a otro más simple; el lector debe poder reconstruir por qué basta.
5. **"Recíprocamente"** avisa de que empieza la segunda mitad de un "si y solo si".
6. ■ **o** □ marca el final de la demostración.

Pausa para pensar

Un consejo práctico que vale por todo el capítulo: cuando leas una demostración, no la leas como un texto. Léela con lápiz, y en cada línea pregúntate "¿de dónde sale esto?". Si no lo sabes, la línea está mal escrita o te falta algo, y en ambos casos hay que parar ahí. La matemática es la única lectura en la que ir despacio es ir rápido: cinco páginas entendidas valen más que cincuenta hojeadas, y el resto del libro está escrito suponiendo que lees así.

1.9. Contar lo infinito: cardinalidad y el argumento diagonal

Hasta ahora "cuántos elementos tiene" era una pregunta con respuesta cuando el conjunto era finito. Cantor se dio cuenta de que la pregunta sigue teniendo sentido para conjuntos infinitos, con tal de renunciar a contar de uno en uno y quedarse con lo esencial: emparejar.

Definición 1.35. Dos conjuntos A y B tienen la *misma cardinalidad*, y se escribe $|A| = |B|$, cuando existe una biyección $f : A \to B$, es decir, una correspondencia que asigna a cada elemento de A uno de B sin repetir ninguno (*inyectiva*) y sin dejarse ninguno (*sobreyectiva*).

Un conjunto es *numerable* si tiene la misma cardinalidad que $\mathbb{N}$; es decir, si sus elementos se pueden poner en una lista infinita $a_1, a_2, a_3, \ldots$ en la que aparezcan todos.

Idea clave

La definición renuncia a contar y se queda con emparejar, y eso trae una consecuencia que al principio escandaliza: un conjunto infinito puede tener la misma cardinalidad que una parte propia suya. Hay tantos números pares como números naturales, porque $n \mapsto 2n$ los empareja uno a uno, aunque los pares sean "la mitad". Esto no es una paradoja sino la definición misma de infinito: un conjunto es infinito precisamente cuando admite una biyección con un subconjunto propio.

Regla: decidir si un conjunto es numerable

1. **Intenta primero fabricar la lista.** Si consigues describir un procedimiento que, en un número finito de pasos, llega a cualquier elemento dado, el conjunto es numerable.
2. **Ordena por tamaño y luego alfabéticamente** cuando los elementos tengan un "tamaño" natural: por eso las palabras finitas de un alfabeto finito son numerables aunque haya infinitas.
3. **Usa el truco de las diagonales** para productos: para listar $\mathbb{N} \times \mathbb{N}$ se recorren las diagonales $i + j = 2$, luego $i + j = 3$, y así sucesivamente; cada diagonal es finita, de modo que todo par acaba saliendo.
4. **Aplica los tres resultados de fábrica:** un subconjunto de un numerable es numerable o finito; la unión numerable de conjuntos numerables es numerable; el producto de dos numerables es numerable.
5. **Si sospechas que *no* es numerable, no busques la lista: destrúyela.** Supón dada una lista cualquiera y construye a partir de ella un elemento que no puede estar en ella.
6. **Para construirlo, usa la diagonal:** haz que el elemento nuevo difiera del k-ésimo de la lista precisamente en la posición k. Así no coincide con ninguno, y la lista no era completa.
7. **Concluye por reducción al absurdo:** como la lista era arbitraria, ninguna lista sirve, y el conjunto no es numerable.

Ejemplo 1.36. Cuatro conjuntos que sí son numerables, con su lista explícita.

1. Los pares: $0, 2, 4, 6, \ldots$ mediante $n \mapsto 2n$. Los cuadrados perfectos: $0, 1, 4, 9, 16, \ldots$ mediante $n \mapsto n^2$, y esta fue la observación de Galileo tres siglos antes que Cantor.
2. Los enteros $\mathbb{Z}$, en apariencia "el doble" que $\mathbb{N}$: basta listarlos en zigzag $0, 1, -1, 2, -2, 3, -3, \ldots$, lo que corresponde a la fórmula $n \mapsto n/2$ si n es par y $n \mapsto -(n+1)/2$ si es impar. Todo entero aparece, y en un lugar concreto.

3. Los racionales positivos, recorriendo las diagonales $i+j = 2, 3, 4, \ldots$ del cuadro de fracciones i/j y saltándose las ya simplificadas: $\frac{1}{1}, \frac{1}{2}, \frac{2}{1}, \frac{1}{3}, \frac{3}{1}, \frac{1}{4}, \frac{2}{3}, \frac{3}{2}, \frac{4}{1}, \frac{1}{5}, \frac{5}{1}, \frac{1}{6}, \ldots$ Cada diagonal tiene un número finito de fracciones, así que toda fracción tiene su turno. Los racionales, que parecían "muchísimos más" que los enteros porque hay infinitos entre dos cualesquiera, son exactamente igual de numerosos.

4. Los polinomios de coeficientes enteros son numerables (ordénalos por grado y por suma de valores absolutos de los coeficientes), y como cada uno tiene a lo sumo n raíces, los números *algebraicos* —los que son raíz de algún polinomio entero, entre ellos $\sqrt{2}$ y todas las raíces del capítulo 26— forman también un conjunto numerable.

Ejemplo 1.37. El argumento diagonal, paso a paso, y lo que se sigue de él.

1. Supongamos que los reales del intervalo $[0, 1)$ se pudieran listar. Escribimos los cuatro primeros de una lista imaginaria en forma decimal:

$$x_1 = 0{,}\mathbf{1}234567\ldots, \quad x_2 = 0{,}2\mathbf{3}45678\ldots, \quad x_3 = 0{,}34\mathbf{5}6789\ldots, \quad x_4 = 0{,}456\mathbf{7}890\ldots$$

2. Tomamos los dígitos de la diagonal, en negrita: $1, 3, 5, 7, \ldots$, y fabricamos un número nuevo cambiando cada uno, por ejemplo sumándole 1: obtenemos $y = 0{,}2468\ldots$

3. Ese y no puede ser x_1, porque difieren en el primer decimal; ni x_2, porque difieren en el segundo; ni x_k para ningún k, porque difieren en el decimal k-ésimo. Luego y no está en la lista, y la lista no era completa. Como la lista era arbitraria, $[0, 1)$ no es numerable: $|\mathbb{R}| > |\mathbb{N}|$.

4. El mismo argumento, aplicado a los subconjuntos de $\mathbb{N}$ en vez de a los decimales, da el teorema general de Cantor: para todo conjunto A se cumple $|A| < |\mathcal{P}(A)|$. Para conjuntos finitos esto es la desigualdad $n < 2^n$ —con $n = 1, 2, 3, 4, 5$ da $2, 4, 8, 16, 32$—, pero el argumento vale igual para infinitos, y produce una escalera infinita de infinitos cada vez mayores.

Cuidado

Al construir el número diagonal hay un detalle técnico que no se puede saltar: los desarrollos decimales no son únicos, porque $0{,}4999\ldots = 0{,}5$. Si al cambiar los dígitos uno se topara con una cola de nueves o de ceros, el número construido podría coincidir con uno de la lista escrito de la otra manera. Se arregla eligiendo los dígitos nuevos, por ejemplo, siempre entre 2 y 7: así el resultado no tiene cola de nueves ni de ceros y la comparación decimal es fiable. Este tipo de remiendo, que parece pedante, es lo que separa un argumento convincente de una demostración.

Pausa para pensar

Junta las dos conclusiones: los algebraicos son numerables y los reales no. La resta es inmediata y demoledora: tienen que existir números reales que no son raíz de ningún polinomio de coeficientes enteros —los llamados *trascendentes*, como π y e—, y de

hecho son la inmensa mayoría. Fíjate en lo que acaba de ocurrir: hemos demostrado que existen infinitos objetos de cierto tipo sin exhibir ni uno solo. Se llama *demostración de existencia no constructiva*, y es de las cosas más elegantes y más incómodas que hace la matemática. La misma tensión entre "existe" y "sé cuál es" reaparece en el capítulo 29 con el axioma del supremo y en el capítulo 31 con los conjuntos no medibles.

1.10. Errores típicos de lógica

Cuidado

1. **Afirmar el consecuente.** De "$p \Rightarrow q$" y "q" no se sigue "p". Es usar el recíproco sin haberlo probado.
2. **Negar el antecedente.** De "$p \Rightarrow q$" y "no p" no se sigue "no q".
3. **Confundir $\in$ con $\subseteq$**, y $\{a\}$ con a.
4. **Negar mal un "para todo".** Lo contrario de "todos" es "al menos uno no", no "ninguno".
5. **Intercambiar cuantificadores.** $\forall\exists$ y $\exists\forall$ dicen cosas distintas; casi siempre una es cierta y la otra falsa.
6. **Demostrar con ejemplos.** Verificar cien casos no prueba nada universal; un solo caso sí refuta.
7. **Razonar hacia atrás sin decirlo.** Partir de lo que se quiere probar y llegar a algo cierto solo vale si cada paso es reversible; si no, hay que reescribir el argumento al revés.
8. **Saltarse la base de la inducción.** Sin la primera ficha no cae ninguna, por bien engranadas que estén.

1.11. Funciones, familias indexadas y la escalera de infinitos

Definición 1.38 (Función)**.** Una función $f : A \to B$ es un subconjunto $F \subseteq A \times B$ tal que para cada $a \in A$ existe un único b con $(a, b) \in F$. Es **inyectiva** si $f(a) = f(a') \Rightarrow a = a'$; **sobreyectiva** si su imagen es todo B; **biyectiva** si ambas.

Teorema 1.39. *f es biyectiva si y solo si existe $g : B \to A$ con $g \circ f = \mathrm{id}_A$ y $f \circ g = \mathrm{id}_B$. Además f es inyectiva si y solo si tiene inversa por la izquierda, y sobreyectiva si y solo si la tiene por la derecha (esto último usa el axioma de elección cuando A es infinito).*

Ejemplo 1.40 (La misma fórmula, cuatro funciones distintas)**.** Con $f(x) = x^2$:

$\mathbb{R} \to \mathbb{R}$	ni inyectiva ni sobreyectiva	$(f(-2) = f(2);\ -1$ no se alcanza)
$[0, \infty) \to \mathbb{R}$	inyectiva, no sobreyectiva	
$\mathbb{R} \to [0, \infty)$	sobreyectiva, no inyectiva	
$[0, \infty) \to [0, \infty)$	biyectiva	inversa: $\sqrt{}$

La moraleja es que una función *no es una fórmula*: es una fórmula junto con un dominio y un codominio. Cambiar cualquiera de los tres cambia el objeto.

De paso, el recuento: hay $|B|^{|A|}$ funciones de A en B. Con $|A| = 3$ y $|B| = 2$ son $2^3 = 8$; inyectivas, ninguna (no caben tres elementos en dos); sobreyectivas, $8 - 2 = 6$ (todas menos las dos constantes).

Definición 1.41 (Familias indexadas)**.** Para una familia $\{A_i\}_{i \in I}$ con I arbitrario,

$$\bigcup_{i \in I} A_i = \{x : \exists i,\ x \in A_i\}, \qquad \bigcap_{i \in I} A_i = \{x : \forall i,\ x \in A_i\},$$

y De Morgan se generaliza: el complementario de la unión es la intersección de los complementarios y viceversa.

Ejemplo 1.42 (Dos avisos que se cobrarán más adelante)**. (a)** $A_n = (-1/n,\ 1/n)$ para $n \geq 1$: cada uno es abierto, y

$$\bigcap_{n \geq 1} A_n = \{0\},$$

que no lo es. La intersección *infinita* de abiertos puede no ser abierta.

(b) $B_n = [1/n,\ 1]$: cada uno es cerrado, y

$$\bigcup_{n \geq 1} B_n = (0, 1],$$

que no lo es. Estos dos ejemplos son exactamente la razón de que la definición de topología pida uniones arbitrarias pero solo intersecciones *finitas*, y de que en la teoría de la medida todo se haga con familias *numerables*.

Idea clave

[La paradoja de Russell] Hemos dicho que un conjunto es "una colección de objetos". Eso, tomado literalmente, es contradictorio. Fórmese

$$R = \{x : x \notin x\}.$$

Si $R \in R$, entonces por definición $R \notin R$; y si $R \notin R$, entonces cumple la condición y $R \in R$. Cualquiera de las dos ramas se contradice.

La reparación es no permitir formar conjuntos por comprensión libre, sino solo *dentro* de un conjunto ya dado: se admite $\{x \in A : P(x)\}$ y se prohíbe $\{x : P(x)\}$. Con esa restricción, R simplemente no se puede construir, y la consecuencia es que *no existe el conjunto de todos los conjuntos*. La versión de barbería —el barbero que afeita

exactamente a quienes no se afeitan a sí mismos— tiene la misma estructura, y su moraleja es la misma: la conclusión no es una paradoja del mundo, es que ese barbero no existe.

Teorema 1.43 (Cantor–Schröder–Bernstein)**.** *Si existen inyecciones $f : A \to B$ y $g : B \to A$, entonces existe una biyección entre A y B.*

La demostración parte A en tres trozos según el origen de cada elemento al ir tirando hacia atrás con g^{-1} y f^{-1}: los que tienen una cadena de antecesores infinita, los que terminan en A y los que terminan en B. En los dos primeros se usa f; en el tercero, g^{-1}. No hace falta el axioma de elección.

Ejemplo 1.44 (Dos cardinalidades imposibles de construir a mano)**. (a)** $|[0,1]| = |(0,1)|$**.** Las inclusiones $(0,1) \hookrightarrow [0,1]$ y $x \mapsto \frac{x+1}{3}$ de $[0,1]$ en $(0,1)$ son inyectivas. El teorema concluye. Construir la biyección explícitamente exige el truco de Hilbert (mover una sucesión infinita de puntos); el teorema evita tener que pensarlo.

(b) $|[0,1]| = |[0,1]^2|$**.** La inyección obvia es $x \mapsto (x, 0)$. En sentido contrario, entrelácense los desarrollos decimales: $(0{,}a_1a_2\ldots,\ 0{,}b_1b_2\ldots) \mapsto 0{,}a_1b_1a_2b_2\ldots$ Esa aplicación no es sobreyectiva —por los problemas con los nueves periódicos—, pero sí es inyectiva, que es todo lo que hace falta. El cuadrado tiene tantos puntos como el segmento, resultado que dejó a Cantor escribiendo a Dedekind "lo veo, pero no lo creo".

Observación 1.45. Y por encima de todo eso está la escalera: por el argumento diagonal, $|A| < |\mathcal{P}(A)|$ *siempre*. Partiendo de $\mathbb{N}$ se obtiene una sucesión estrictamente creciente de infinitos, $\aleph_0 < \mathfrak{c} = 2^{\aleph_0} < 2^{\mathfrak{c}} < \cdots$, que no termina nunca. Si hay algún cardinal *entre* $\aleph_0$ y $\mathfrak{c}$ es la hipótesis del continuo, y se sabe desde 1963 que los axiomas usuales no lo deciden.

1.12. Inducción ordinaria, inducción fuerte y buena ordenación

Regla: los tres enunciados

Sobre $\mathbb{N}$, los siguientes principios son *equivalentes*:

1. **Inducción ordinaria.** Si $P(0)$ y $P(n) \Rightarrow P(n+1)$, entonces $P(n)$ para todo n.
2. **Inducción fuerte.** Si para todo n se cumple que $[P(k)$ para todo $k < n] \Rightarrow P(n)$, entonces $P(n)$ para todo n. (Nótese que no hace falta caso base: con $n = 0$ la hipótesis es vacía.)
3. **Buena ordenación.** Todo subconjunto no vacío de $\mathbb{N}$ tiene un mínimo.

Esquema de las equivalencias. (1) $\Rightarrow$ (3): si $S \neq \varnothing$ no tuviera mínimo, se probaría por inducción que ningún n está en S, luego $S = \varnothing$. (3) $\Rightarrow$ (2): si el conjunto de contraejemplos fuera no vacío tendría un mínimo n, y entonces $P(k)$ valdría para todo $k < n$, lo que por hipótesis da $P(n)$: contradicción. (2) $\Rightarrow$ (1) es inmediato. $\square$

Ejemplo 1.46 (Un caso donde la inducción ordinaria no sirve)**. Afirmación.** Todo $n \geq 12$ se puede pagar con monedas de 4 y de 5.

Con inducción ordinaria uno se atasca: saber que n se paga no dice nada útil sobre $n+1$, porque no hay moneda de 1. Con inducción fuerte, en cambio, sale sola.

Base. $12 = 4+4+4$, $13 = 4+4+5$, $14 = 4+5+5$, $15 = 5+5+5$. Hacen falta *cuatro* casos base, tantos como el valor de la moneda mayor, y esa es justamente la señal de que el problema pide inducción fuerte.

Paso. Sea $n \geq 16$ y supóngase la afirmación cierta para todo k con $12 \leq k < n$. Como $n-4 \geq 12$, la hipótesis se aplica a $n-4$: se paga con monedas de 4 y 5. Añádase una moneda de 4 y queda pagado n.

Observación 1.47. El 12 no es arbitrario. El mayor valor *no* pagable con monedas de a y b coprimas es el número de Frobenius $ab-a-b$; con $a=4$ y $b=5$ da $20-9=11$, y en efecto 11 no se puede pagar y todo lo que sigue sí. El problema general con tres o más monedas no tiene fórmula cerrada conocida.

Ejemplo 1.48 (La buena ordenación, escondida en una demostración clásica)**.** En la prueba de que $\sqrt{2}$ es irracional se dice "escríbase p/q en forma irreducible". Esa frase *es* el principio de buena ordenación: entre todas las fracciones que representan al número, se toma la de denominador mínimo, y ese mínimo existe porque los denominadores son naturales.

La versión sin fracciones, llamada *descenso infinito* de Fermat, hace el mismo trabajo explícitamente: si $p^2 = 2q^2$, entonces p es par, $p = 2r$, luego $q^2 = 2r^2$ y (q, r) es una solución *más pequeña*. Repitiendo se obtendría una sucesión estrictamente decreciente de naturales, que no existe. La contradicción no es sobre paridades: es sobre la imposibilidad de descender para siempre dentro de $\mathbb{N}$.

Cuidado

Los dos errores clásicos de la inducción. El primero es olvidar el caso base: sin él se "demuestra" que todos los naturales son iguales, porque el paso inductivo puede ser impecable y no sostener nada. El segundo es más sutil: aplicar el paso a un caso en el que no vale. El famoso argumento de que "todos los caballos son del mismo color" —quitando un caballo del grupo y luego otro— falla exactamente al pasar de $n=1$ a $n=2$, porque ahí los dos subgrupos que se solapan no tienen ningún caballo en común. La demostración parece válida para todo n y solo se rompe en un sitio; encontrar ese sitio es el ejercicio.

1.13. Diagramas de Hasse: dibujar un orden sin dibujarlo entero

Un orden parcial sobre un conjunto de seis elementos puede tener casi cuarenta pares relacionados, y dibujarlos todos produce una maraña ilegible en la que no se ve nada. La observación de Hasse es que la mayoría de esos pares son redundantes: los reflexivos están

siempre, por definición, y los que se obtienen encadenando otros dos se pueden recuperar. Lo único que hay que dibujar es el esqueleto.

Definición 1.49 (Relación de cubrimiento y diagrama de Hasse)**.** Sea $(P, \leq)$ un conjunto parcialmente ordenado. Se dice que b *cubre* a a, y se escribe $a \lessdot b$, cuando

$$a < b \qquad \text{y no existe } c \in P \text{ con } a < c < b.$$

Es decir, b está inmediatamente por encima de a, sin escalones intermedios.

El *diagrama de Hasse* de $(P, \leq)$ es el dibujo que coloca cada elemento de P en un punto del plano de modo que si $a < b$ entonces a queda estrictamente más abajo que b, y que traza un segmento entre a y b *exactamente* cuando $a \lessdot b$. No se dibujan bucles —los pares reflexivos $a \leq a$— ni segmentos entre pares que se deducen por transitividad, ni flechas: la dirección la lleva la altura.

Ejemplo 1.50 (Dos diagramas descritos en palabras)**. (a) Los divisores de** 12 **ordenados por divisibilidad.** El conjunto es $D = \{1, 2, 3, 4, 6, 12\}$ y la relación es "a divide a b". Hay 18 pares relacionados, de los cuales 6 son reflexivos y quedan 12 estrictos. De esos doce, solo *siete* son cubrimientos:

$$1 \lessdot 2, \quad 1 \lessdot 3, \quad 2 \lessdot 4, \quad 2 \lessdot 6, \quad 3 \lessdot 6, \quad 4 \lessdot 12, \quad 6 \lessdot 12.$$

El dibujo tiene cuatro niveles. Abajo del todo, solo el 1. En el segundo nivel, 2 y 3, uno al lado del otro, cada uno unido al 1 por un segmento. En el tercero, 4 y 6: el 4 se une solo al 2, y el 6 se une tanto al 2 como al 3, de modo que del 6 bajan dos segmentos que forman una uve invertida. Arriba del todo, el 12, unido al 4 y al 6. La figura resultante recuerda a un rombo alargado con una rama lateral. Los cinco pares estrictos que no se dibujan —$1 < 4$, $1 < 6$, $1 < 12$, $2 < 12$, $3 < 12$— se leen siguiendo caminos ascendentes: hay camino de 1 a 12 pasando por 2 y 4, luego $1 < 12$. Y hay tres pares incomparables, que son $\{2, 3\}$, $\{3, 4\}$ y $\{4, 6\}$: entre ellos no hay ningún camino ascendente, y en el dibujo se reconocen porque están al mismo nivel o en ramas que no se comunican.

(b) Los subconjuntos de $\{a, b, c\}$ **ordenados por inclusión.** Hay 8 elementos y 27 pares relacionados, 19 de ellos estrictos, y los cubrimientos son exactamente 12: los pares en que un conjunto tiene un elemento más que el otro. El diagrama tiene cuatro niveles, por cardinal: el vacío abajo; los tres unitarios; los tres de dos elementos; y el total arriba. Cada unitario se une al vacío y a los dos pares que lo contienen; cada par se une a sus dos unitarios y al total. El resultado es, literalmente, un cubo dibujado en perspectiva y apoyado sobre un vértice: doce aristas, ocho vértices y las cuatro capas del cubo mirado por su diagonal principal. Que el retículo de subconjuntos de un conjunto de n elementos sea el n-cubo no es una metáfora; es la razón por la que los problemas de conteo sobre subconjuntos y los problemas de recorrido sobre cubos son el mismo problema.

Cuidado

Qué se pierde y qué se gana. Se gana legibilidad y se gana *unicidad*: el conjunto de cubrimientos determina el orden por completo, porque en un conjunto finito todo par

estricto $a < b$ se descompone en una cadena de cubrimientos $a \lessdot c_1 \lessdot \cdots \lessdot b$, y por tanto recuperar $\leq$ es tomar el cierre reflexivo y transitivo de lo dibujado. No se pierde, pues, ni un solo dato. Lo que sí se pierde es la lectura *inmediata* de la comparabilidad: para saber si $a \leq b$ hay que buscar un camino ascendente, y con muchos niveles eso deja de ser un vistazo.

Dos avisos más. El primero, que la posición horizontal no significa nada: dos diagramas que parecen distintos pueden ser el mismo orden, exactamente como pasaba con los dibujos de un grafo. Solo cuentan qué está unido con qué y quién está por encima de quién. El segundo, y más serio, que la construcción exige que los cubrimientos existan. En $(\mathbb{Q}, \leq)$ ningún racional cubre a otro, porque entre dos cualesquiera siempre cabe un tercero: el diagrama de Hasse de los racionales no tiene *ningún* segmento y no recupera absolutamente nada del orden. Los diagramas de Hasse son una herramienta para órdenes finitos, o a lo sumo para órdenes en que toda cadena entre dos elementos es finita; aplicarlos a un orden denso no es difícil, es imposible.

1.14. Ejercicios resueltos

Ejercicio 1.1. Con $A = \{1, 2, 3, 4\}$ y $B = \{3, 4, 5\}$, calcula $A \cup B$, $A \cap B$, $A \setminus B$ y $|\mathcal{P}(A)|$.

Solución. $A \cup B = \{1, 2, 3, 4, 5\}$; $A \cap B = \{3, 4\}$; $A \setminus B = \{1, 2\}$; y $|\mathcal{P}(A)| = 2^4 = 16$. ■

Ejercicio 1.2. Escribe el contrarrecíproco de "si x^2 es irracional, entonces x es irracional" y decide si el enunciado original es cierto.

Solución. Contrarrecíproco: "si x es racional, entonces x^2 es racional". Este es evidentemente cierto (producto de racionales), luego el original también lo es. Nótese que el recíproco es falso: $x = \sqrt{2}$ es irracional y $x^2 = 2$ no. ■

Ejercicio 1.3. Niega: "para todo $\varepsilon > 0$ existe N tal que para todo $n > N$ se tiene $|a_n - L| < \varepsilon$".

Solución. "Existe $\varepsilon > 0$ tal que para todo N existe $n > N$ con $|a_n - L| \geq \varepsilon$". Es decir, la sucesión se sale del margen ε infinitas veces. ■

Ejercicio 1.4. Demuestra por inducción que $1 + 3 + 5 + \cdots + (2n - 1) = n^2$.

Solución. Base: $n = 1$ da $1 = 1^2$. Paso: si $1 + \cdots + (2k - 1) = k^2$, sumando el siguiente impar $2k + 1$ queda $k^2 + 2k + 1 = (k + 1)^2$. ■ (Geométricamente: cada impar es la "escuadra" que agranda un cuadrado de lado k a uno de lado $k + 1$.) ■

Ejercicio 1.5. Demuestra que $2^n > n^2$ para todo $n \geq 5$.

Solución. Base: $n = 5$ da $32 > 25$. Paso: supongamos $2^k > k^2$ con $k \geq 5$. Entonces $2^{k+1} = 2 \cdot 2^k > 2k^2$, y basta ver que $2k^2 \geq (k + 1)^2$, es decir $k^2 - 2k - 1 \geq 0$, cierto para $k \geq 3$. ■ ■

Ejercicio 1.6. Demuestra que no existe un racional cuyo cuadrado sea 3.

Solución. Si $\sqrt{3} = p/q$ irreducible, entonces $3q^2 = p^2$, luego $3 \mid p^2$ y, como 3 es primo, $3 \mid p$ (esto último es el lema de Euclides del capítulo 3). Escribiendo $p = 3m$: $3q^2 = 9m^2$, o sea $q^2 = 3m^2$, y por lo mismo $3 \mid q$. Contradicción con la irreducibilidad. ■ ■

1.15. Ejercicios propuestos

Sin solución desarrollada: solo enunciado y, al final, la respuesta. Los últimos piden decidir *verdadero o falso* y justificar; ahí la respuesta corta es el veredicto, y el trabajo es tuyo.

1.1. Con $U = \{1, \dots, 8\}$, $A = \{1, 3, 5, 7\}$, $B = \{2, 3, 5, 7\}$: calcula $A \cap B$, $A \cup B$, $A \setminus B$ y A^c.

1.2. ¿Cuántos elementos tiene $\mathcal{P}(\mathcal{P}(\{a, b\}))$?

1.3. Escribe por comprensión el conjunto $\{1, 4, 9, 16, 25\}$.

1.4. Da el recíproco y el contrarrecíproco de "si un número es divisible por 6, es divisible por 3", y di cuáles son ciertos.

1.5. Construye la tabla de verdad de $(p \Rightarrow q) \wedge (q \Rightarrow p)$ y di a qué conectivo equivale.

1.6. Niega: "existe un estudiante que aprobó todos los exámenes".

1.7. ¿Es cierto que $\forall x \in \mathbb{R}\ \exists y \in \mathbb{R} :\ xy = 1$?

1.8. Demuestra por inducción que $1^2 + 2^2 + \cdots + n^2 = \dfrac{n(n+1)(2n+1)}{6}$.

1.9. Demuestra que la suma de un racional y un irracional es irracional.

1.10. Verdadero o falso: si $A \subseteq B$ y $B \subseteq C$ entonces $A \subseteq C$.

1.11. Verdadero o falso: $A \times B = B \times A$ para todos A, B.

1.12. Verdadero o falso: $\varnothing \in \mathcal{P}(A)$ para todo A, y también $\varnothing \subseteq A$.

Respuestas

1. $A \cap B = \{3, 5, 7\}$; $A \cup B = \{1, 2, 3, 5, 7\}$; $A \setminus B = \{1\}$; $A^c = \{2, 4, 6, 8\}$. **2.** $\mathcal{P}(\{a, b\})$ tiene 4 elementos, luego su potencia tiene $2^4 = 16$. **3.** $\{n^2 :\ n \in \mathbb{N},\ 1 \le n \le 5\}$. **4.** Recíproco: "si es divisible por 3, lo es por 6" (falso: el 9). Contrarrecíproco: "si no es divisible por 3, no lo es por 6" (cierto). **5.** Equivale a $p \Leftrightarrow q$: verdadera solo cuando p y q tienen el mismo valor. **6.** "Todo estudiante reprobó al menos un examen". **7.** Falso: $x = 0$ no tiene inverso. Sería cierto en $\mathbb{R} \setminus \{0\}$. **8.** Base $n = 1$: $1 = 6/6$. Paso: sumar $(k+1)^2$ y factorizar $(k+1)$ da $\frac{(k+1)(k+2)(2k+3)}{6}$. **9.** Por contradicción: si $q + i = r$ fuera racional, entonces $i = r - q$ sería racional. **10.** Verdadero (transitividad de la inclusión). **11.** Falso salvo si $A = B$ o alguno es vacío: $(a, b) \neq (b, a)$. **12.** Verdadero las dos: el vacío es subconjunto de todo, y por eso es elemento de todo conjunto potencia.

Capítulo 2

Lógica matemática y teoría de conjuntos

Lo que necesitas traer

Del capítulo 1: conjuntos, cuantificadores, funciones y el argumento diagonal de Cantor. Nada más. Los sistemas formales, la aritmética de Peano y las máquinas de Turing se construyen aquí desde el principio; conviene, eso sí, haber digerido bien qué significa *demostrar* algo antes de preguntarse qué *no* se puede demostrar.

Todo el libro ha usado demostraciones sin preguntarse qué es exactamente una demostración. La lógica matemática convierte esa pregunta en objeto de estudio: formaliza qué es un enunciado, qué es una deducción y qué es un modelo, y con ello puede demostrar teoremas *sobre* la matemática misma. Los resultados son de los más sorprendentes de la disciplina, porque establecen límites intrínsecos a lo que cualquier sistema formal puede lograr.

2.1. Sintaxis, semántica y completitud

Definición 2.1. En lógica de primer orden se distingue entre $\Gamma \vdash \varphi$ (la fórmula φ se *deduce* de Γ mediante reglas puramente sintácticas) y $\Gamma \models \varphi$ (φ es *verdadera* en todo modelo de Γ). El *teorema de completitud* de Gödel afirma que ambas nociones coinciden, y el *teorema de compacidad* que si todo subconjunto finito de Γ tiene modelo, entonces Γ lo tiene.

Idea clave

Qué dice esto en cristiano. Hay dos maneras de entender "se sigue de": una mecánica, manipulando símbolos con reglas, y otra semántica, sobre significados y modelos. El teorema de completitud dice que coinciden, lo que es una noticia excelente: todo lo verdadero en todos los modelos es demostrable, y viceversa. La compacidad es su consecuencia más productiva y también la más contraintuitiva: si cada trozo finito de una teoría es satisfacible, la teoría entera lo es, por infinita que sea. De ahí salen modelos con propiedades imposibles de imaginar de antemano.

Ejemplo 2.2 (Compacidad construyendo lo imposible)**. (a) Infinitesimales.** Tómese la teoría de los reales y añádase una constante c con los axiomas $c > 0$, $c < 1$, $c < \frac{1}{2}$, $c < \frac{1}{3}\ldots$ Cada subconjunto *finito* es satisfacible (basta un real pequeño), luego por compacidad el conjunto entero tiene modelo: existe un cuerpo ordenado que extiende a $\mathbb{R}$ con un elemento positivo menor que todo racional positivo. Son los hiperreales, y con ellos se puede hacer análisis con infinitesimales de verdad.

(b) Ningún axioma distingue lo finito. No existe ninguna teoría de primer orden cuyos modelos sean exactamente los conjuntos finitos: añadiendo axiomas "hay al menos n elementos" para todo n, cada trozo finito es satisfacible y por compacidad hay un modelo infinito. La finitud no es expresable en primer orden.

(c) Löwenheim–Skolem. Toda teoría con un modelo infinito tiene modelos de todos los cardinales infinitos. En particular la teoría de conjuntos, si es consistente, tiene un modelo *numerable*, cosa que parece contradecir el teorema de Cantor sobre la no numerabilidad de $\mathbb{R}$. Es la paradoja de Skolem, y se resuelve observando que "numerable" significa dentro del modelo, y la biyección que faltaría vive fuera de él.

(d) Lo que no se puede definir. La compacidad implica que la noción de "ser alcanzable desde el cero sumando uno repetidamente" no es de primer orden, y por eso existen modelos no estándar de la aritmética con elementos infinitos. Ninguna cantidad de axiomas de primer orden consigue fijar $\mathbb{N}$ salvo isomorfismo.

2.2. Del alfabeto al teorema: fórmulas bien formadas, sistemas deductivos, consistencia, corrección y completitud

El capítulo se abrió distinguiendo $\Gamma \vdash \varphi$ de $\Gamma \models \varphi$ y desde entonces ha usado esos dos símbolos en cada página. La sección sobre estructuras y satisfacción explicó con detalle el segundo: qué es un modelo y qué significa que una fórmula sea verdadera en él. Del primero, en cambio, no se ha dicho nunca qué es: no se ha dado ni un solo axioma, ni una sola regla de inferencia, ni se ha escrito jamás una demostración formal completa. Tampoco se ha definido qué es una fórmula —se han manejado φ, ψ, $\mathrm{Dem}_T(x)$ como si el lector supiera qué cadenas de símbolos están permitidas—, ni qué significa exactamente que una teoría sea *consistente*, palabra que aparece en la hipótesis de los dos teoremas de incompletitud. Esta sección construye todo eso desde cero. Es la mitad sintáctica de la historia, y sin ella la mitad semántica no se sostiene.

Qué cadenas de símbolos están permitidas

Definición 2.3 (Alfabeto, términos y fórmulas bien formadas)**.** Fijado un lenguaje de primer orden $\mathcal{L}$ —con sus símbolos de constante, de función y de relación, tal como se definió al hablar de estructuras—, el *alfabeto* consta de esos símbolos más un repertorio infinito de variables $x_0, x_1, x_2, \ldots$, las conectivas $\neg, \wedge, \vee, \to, \leftrightarrow$, los cuantificadores $\forall, \exists$, el símbolo de igualdad $=$, los paréntesis y la coma.

Los *términos* se definen por recursión: toda variable es un término; toda constante es un término; y si g es un símbolo de función de aridad n y $t_1, \ldots, t_n$ son términos, entonces $g(t_1, \ldots, t_n)$ es un término. Nada más es un término.

Las *fórmulas bien formadas* (abreviado FBF) se definen también por recursión:

1. si R es un símbolo de relación de aridad n y $t_1, \ldots, t_n$ son términos, entonces $R(t_1, \ldots, t_n)$ es una FBF, llamada *atómica*; también lo es $t_1 = t_2$;
2. si φ es una FBF, $\neg\varphi$ lo es;
3. si φ y ψ lo son, también $(\varphi \wedge \psi)$, $(\varphi \vee \psi)$, $(\varphi \to \psi)$ y $(\varphi \leftrightarrow \psi)$;
4. si φ lo es y x es una variable, también $\forall x\, \varphi$ y $\exists x\, \varphi$;
5. nada más es una FBF.

La última cláusula no es un adorno: es la que permite demostrar propiedades de todas las fórmulas por inducción sobre su construcción.

Ejemplo 2.4 (Cuatro cadenas, dos buenas y dos malas)**.** En el lenguaje $\mathcal{L} = \{0, 1, +, \cdot, <\}$ de la aritmética:

(a) $\forall x\, \exists y\, (x < y)$ es una FBF. Se construye así: x e y son términos por ser variables; $x < y$ es atómica por la cláusula 1; $\exists y\, (x < y)$ por la 4; y $\forall x\, \exists y\, (x < y)$ otra vez por la 4. Ese registro de pasos es el *árbol de formación*, y toda FBF tiene uno y solo uno.

(b) $\forall x\, (x + 1 < x \cdot x \to \neg(x = 0))$ también lo es: $x + 1$ y $x \cdot x$ son términos por la regla de las funciones, la desigualdad y la igualdad son atómicas, el $\neg$ por la cláusula 2, la flecha por la 3 y el cuantificador por la 4.

(c) $\forall x\, (x + 1)$ *no* es una FBF: $x + 1$ es un término, no una fórmula, y los cuantificadores solo se aplican a fórmulas. Confundir término con fórmula es el error de sintaxis más común, y equivale a escribir "para todo x, siete".

(d) $\forall <\, (x < y)$ tampoco lo es: solo se puede cuantificar sobre variables, nunca sobre símbolos de relación. Precisamente esa prohibición es lo que hace que la lógica sea de *primer* orden; levantarla da la lógica de segundo orden, que caracteriza $\mathbb{N}$ salvo isomorfismo pero pierde el teorema de completitud, como se explicó al hablar de modelos no estándar.

Definición 2.5 (Variables libres y ligadas, sentencia)**.** Una aparición de la variable x en una fórmula está *ligada* si cae dentro del alcance de un cuantificador $\forall x$ o $\exists x$, y está *libre* en caso contrario. Una FBF sin variables libres se llama *sentencia*. Solo las sentencias tienen un valor de verdad absoluto en una estructura; una fórmula con variables libres expresa una propiedad, y su verdad depende de qué elementos se asignen a esas variables, que es exactamente el papel de la asignación s en la definición de satisfacción.

Ejemplo 2.6. En $\exists y\, (x < y)$ la variable y está ligada y x está libre: no es una sentencia, sino una propiedad de x que en $\mathbb{N}$ cumple todo número. En $\forall x\, \exists y\, (x < y)$ no hay variables libres: es una sentencia, y en $\mathbb{N}$ es verdadera mientras que en el conjunto $\{0, 1, 2\}$ con el orden usual es falsa. En $(\exists y\, (x < y)) \wedge (y = 0)$ la misma letra y aparece ligada en el primer trozo y libre en el segundo: es legal, pero es la clase de escritura que conviene evitar renombrando la variable ligada.

Qué es una demostración formal

Definición 2.7 (Sistema deductivo de Hilbert, demostración y $\vdash$)**.** Un *sistema deductivo* consta de un conjunto de *axiomas lógicos* —fórmulas que se admiten sin justificación— y un conjunto de *reglas de inferencia* —procedimientos que producen fórmulas a partir de otras—. Un sistema de Hilbert clásico para la parte proposicional toma como axiomas todas las instancias de los tres esquemas

$$\text{(A1) } \varphi \to (\psi \to \varphi),$$

$$\text{(A2) } (\varphi \to (\psi \to \chi)) \to ((\varphi \to \psi) \to (\varphi \to \chi)),$$

$$\text{(A3) } (\neg\psi \to \neg\varphi) \to (\varphi \to \psi),$$

y como única regla el *modus ponens*: de φ y $\varphi \to \psi$ se obtiene ψ. Para la parte cuantificacional se añaden los esquemas de instanciación universal y la regla de generalización.

Una *demostración formal* de φ a partir de un conjunto Γ de premisas es una *lista finita* de fórmulas $\varphi_1, \varphi_2, \ldots, \varphi_n$ tal que $\varphi_n = \varphi$ y cada φ_i es o bien un axioma lógico, o bien un elemento de Γ, o bien el resultado de aplicar una regla de inferencia a fórmulas anteriores de la lista. Se escribe $\Gamma \vdash \varphi$ cuando tal lista existe, y se dice que φ es *deducible* de Γ. Si Γ es vacío se escribe $\vdash \varphi$ y se dice que φ es un *teorema* del sistema.

Idea clave

Lo importante de esta definición es lo que *no* contiene: no aparece por ninguna parte la palabra “verdad”, ni “significado”, ni “modelo”. Una demostración formal es un objeto combinatorio, una lista de cadenas de símbolos que satisface una condición comprobable renglón a renglón por alguien —o algo— que no entienda nada de lo que está escrito. Ésa es precisamente la gracia: como comprobar una demostración es una tarea mecánica, un ordenador puede hacerlo, y como las demostraciones son objetos finitos, se pueden numerar. De ahí sale la codificación de Gödel y con ella todo el capítulo. La contrapartida es que *encontrar* la lista no es mecánico en absoluto.

Ejemplo 2.8 (Una demostración formal entera: $\vdash \varphi \to \varphi$)**.** Parece el teorema más tonto posible y no hay ningún axioma que lo dé directamente; hacen falta cinco líneas. Escribimos θ como abreviatura de $\varphi \to \varphi$.

1. $(\varphi \to ((\varphi \to \varphi) \to \varphi)) \to ((\varphi \to (\varphi \to \varphi)) \to (\varphi \to \varphi))$. Instancia de (A2) con $\psi := \varphi \to \varphi$ y $\chi := \varphi$.
2. $\varphi \to ((\varphi \to \varphi) \to \varphi)$. Instancia de (A1) con $\psi := \varphi \to \varphi$.
3. $(\varphi \to (\varphi \to \varphi)) \to (\varphi \to \varphi)$. Modus ponens sobre las líneas 2 y 1.
4. $\varphi \to (\varphi \to \varphi)$. Instancia de (A1) con $\psi := \varphi$.
5. $\varphi \to \varphi$. Modus ponens sobre las líneas 4 y 3.

Cada renglón está justificado por su etiqueta y por nada más; no hay ni una palabra de castellano en el argumento, y el último renglón es lo que se quería. Que semejante trivialidad

cueste cinco líneas explica por qué nadie trabaja así en la práctica: los sistemas de Hilbert están diseñados para *razonar sobre* demostraciones —pocos esquemas, una regla, todo fácil de codificar—, no para producirlas. Para producirlas se usan la deducción natural o el cálculo de secuentes, que son equivalentes y mucho más cómodos, y el *teorema de la deducción* —si $\Gamma \cup \{\varphi\} \vdash \psi$ entonces $\Gamma \vdash \varphi \to \psi$— reconcilia ambos mundos y hace innecesario volver a escribir listas como ésta.

Consistencia, satisfacibilidad y las dos palabras "completo"

Definición 2.9 (Consistencia, satisfacibilidad, validez)**.** Sea Γ un conjunto de sentencias.

- Γ es *consistente* si no existe ninguna fórmula φ tal que $\Gamma \vdash \varphi$ y $\Gamma \vdash \neg\varphi$. Equivalentemente, si hay al menos una fórmula que Γ *no* demuestra: de un conjunto inconsistente se deduce absolutamente todo, hecho conocido como *principio de explosión*. Es una noción puramente sintáctica.
- Γ es *satisfacible* si existe alguna estructura $\mathcal{M}$ con $\mathcal{M} \models \Gamma$, es decir, si tiene al menos un modelo. Es una noción puramente semántica.
- Una sentencia φ es *válida*, escrito $\models \varphi$, si toda estructura la satisface; en el caso proposicional eso es exactamente ser una tautología.
- Una teoría T consistente es *completa* si para toda sentencia φ del lenguaje se cumple $T \vdash \varphi$ o $T \vdash \neg\varphi$: la teoría se pronuncia sobre todo.
- Una teoría T es *decidible* si existe un algoritmo que, dada cualquier sentencia, termina siempre y responde correctamente si es o no un teorema de T.

Cuidado

[La palabra "completo" significa dos cosas distintas en este capítulo] Ésta es la confusión más frecuente de toda la lógica matemática, y el capítulo usa las dos acepciones a pocas páginas de distancia.

- El **teorema de completitud de Gödel** habla de la completitud del *cálculo deductivo*: todo lo que es consecuencia semántica se puede deducir sintácticamente, es decir, $\Gamma \models \varphi$ implica $\Gamma \vdash \varphi$. Es una buena noticia sobre el aparato de demostración, y vale para la lógica de primer orden entera.
- Los **teoremas de incompletitud de Gödel** hablan de la completitud de una *teoría* concreta: que la aritmética no decide todas las sentencias de su lenguaje. Es una mala noticia sobre las teorías, no sobre el cálculo.

No hay contradicción alguna entre ambos, y el mismo autor demostró los dos con un año de diferencia. Se pueden tener a la vez porque hablan de cosas distintas: el cálculo demuestra todo lo que es verdadero *en todos* los modelos de PA, y la sentencia de Gödel simplemente no lo es —es verdadera en $\mathbb{N}$ y falsa en algún modelo no estándar—. Dicho en una frase: la incompletitud de la aritmética no es un fallo de la lógica, es que los axiomas no consiguen fijar un único modelo. Hay una tercera acepción, la *completitud*

funcional de un conjunto de conectivas, que ya se trató aparte y no tiene nada que ver con estas dos.

Teorema 2.10 (Corrección del cálculo)**.** *Si* $\Gamma \vdash \varphi$ *entonces* $\Gamma \models \varphi$*. En particular, todo teorema del cálculo es válido, y todo conjunto satisfacible es consistente.*

Demostración. Por inducción sobre la longitud de la demostración. Sea $\mathcal{M}$ un modelo cualquiera de Γ y sea $\varphi_1, \ldots, \varphi_n$ una demostración de φ desde Γ. Se prueba que $\mathcal{M} \models \varphi_i$ para todo i. Si φ_i pertenece a Γ, se cumple por hipótesis. Si es un axioma lógico, se cumple porque los axiomas son válidos: por ejemplo (A1) es la tautología $p \to (q \to p)$, cuya tabla da V en las cuatro filas —cuando p es falsa el antecedente falla y cuando p es verdadera el consecuente $q \to p$ es verdadero—, y lo mismo se comprueba con (A2) en ocho filas y con (A3) en cuatro. Si φ_i se obtuvo por modus ponens de φ_j y $\varphi_k = \varphi_j \to \varphi_i$ con $j, k < i$, por hipótesis de inducción $\mathcal{M}$ satisface ambas, y por la cláusula de la implicación en la definición de satisfacción, satisface φ_i. Como $\varphi_n = \varphi$, queda $\mathcal{M} \models \varphi$.

La segunda afirmación es el contrarrecíproco de la primera: si Γ fuese inconsistente demostraría φ y $\neg\varphi$, luego cualquier modelo suyo tendría que satisfacer las dos, lo cual es imposible; por tanto no tendría modelos y no sería satisfacible. □

Observación 2.11. Corrección y completitud son las dos mitades de la misma moneda y conviene no confundirlas. Corrección dice que el cálculo *no demuestra basura*: todo lo deducible es verdad en todos los modelos. Completitud dice que el cálculo *no se deja nada*: todo lo verdadero en todos los modelos es deducible. La primera es fácil, se demuestra en media página como acaba de verse, y sin ella el sistema no serviría para nada. La segunda es profunda, es el teorema con que abrió el capítulo, y su recíproco combinado da la identidad $\vdash = \models$. De ahí sale, como corolario inmediato, que consistente y satisfacible son lo mismo en primer orden, y de esa equivalencia sale a su vez el teorema de compacidad.

Regla: decidir si un conjunto de fórmulas es consistente

1. **Para probar que *sí* lo es, construye un modelo.** Por corrección, todo conjunto con modelo es consistente, y exhibir una estructura concreta es mucho más fácil que razonar sobre todas las demostraciones posibles.
2. **Elige el modelo más pequeño que puedas:** un universo de uno o dos elementos suele bastar, y las relaciones se dan por tablas.
3. **Para probar que *no* lo es, deriva una contradicción explícita,** es decir, exhibe la demostración de alguna φ y la de $\neg\varphi$. Basta una.
4. **En el caso proposicional, usa la tabla de verdad:** el conjunto es satisfacible si y solo si alguna fila hace verdaderas todas sus fórmulas a la vez. Con k letras hay 2^k filas que revisar.
5. **Si el conjunto es infinito, aplica compacidad:** basta comprobar que cada subconjunto finito tiene modelo, y ésa es la técnica con la que se construyeron los

infinitesimales y los modelos no estándar.

6. **No confundas inconsistente con falso.** Un conjunto puede ser perfectamente consistente y ser falso en el modelo que a uno le interesaba; consistencia solo pide que exista *algún* modelo.

Ejemplo 2.12 (Consistencia, satisfacibilidad y validez con casos concretos)**. (a) Consistente, con modelo diminuto.** Sea Γ el conjunto formado por $\forall x\,\forall y\,(x = y)$ y $\exists x\,(x = x)$. Parece contradictorio y no lo es: la estructura de universo $M = \{a\}$, con un solo elemento, satisface las dos. Luego Γ es satisfacible y, por corrección, consistente.

(b) Inconsistente. El conjunto $\{p \vee q,\ \neg p,\ \neg q\}$ no tiene modelo: repasando las cuatro filas de la tabla de verdad, ninguna hace verdaderas las tres fórmulas a la vez —para que valgan $\neg p$ y $\neg q$ hacen falta p y q falsas, y entonces $p \vee q$ es falsa—. Por completitud del cálculo proposicional, de ahí se deduce formalmente una contradicción.

(c) Satisfacible pero no válida. La sentencia $\exists x\,\exists y\,\neg(x = y)$ es verdadera en cualquier estructura con al menos dos elementos y falsa en la del apartado (a). Es decir, tiene modelos pero no todos: satisfacible y no válida. La diferencia entre "tiene algún modelo" y "los tiene todos" es la diferencia entre $\neg \models \neg\varphi$ y $\models \varphi$, y no distinguirlas es como confundir "existe" con "para todo".

(d) Válida. $\forall x\,(x = x)$ lo es, porque la igualdad se interpreta siempre como la identidad del universo, sea cual sea la estructura. También lo es cualquier instancia de (A1). Por completitud, todas ellas son teoremas del cálculo: $\vdash \varphi$.

(e) Completa pero no decidible. La aritmética verdadera $\mathrm{Th}(\mathbb{N})$ es consistente y completa —para cada φ, o φ o $\neg\varphi$ es verdadera en $\mathbb{N}$— y sin embargo no es decidible ni recursivamente axiomatizable, según se explicó al hablar de axiomatizabilidad. Las cinco propiedades de la definición son independientes entre sí, y este ejemplo lo demuestra.

Pausa para pensar

Merece la pena mirar hacia atrás con estas definiciones en la mano. $\vdash$ es un hecho sobre listas finitas de cadenas de símbolos; $\models$ es una cuantificación sobre todas las estructuras, incluidas las de tamaño inconcebible. La corrección los conecta en una dirección con una inducción de media página. La completitud los conecta en la otra, y ésa fue la tesis doctoral de Gödel en 1929. Un año después demostró que ninguna teoría recursivamente axiomatizable y consistente con aritmética suficiente es completa. Con el vocabulario ya fijado, se ve que los dos resultados hablan de sujetos distintos —el cálculo y las teorías— y que el segundo no debilita al primero ni un milímetro.

2.3. Incompletitud

Teorema 2.13 (Gödel, 1931)**.** *Todo sistema formal consistente, con axiomas efectivamente enumerables y suficientemente potente para expresar la aritmética elemental, contiene enunciados verdaderos que no puede demostrar. Y ningún sistema de ese tipo puede demostrar su*

propia consistencia.

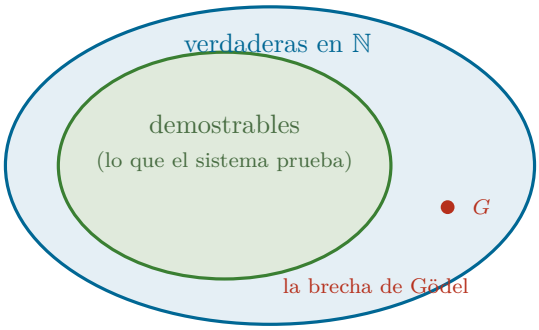


Figura 2.1: El primer teorema de incompletitud dibujado. En todo sistema formal consistente, efectivo y lo bastante potente para la aritmética, lo demostrable queda estrictamente dentro de lo verdadero, y la sentencia G vive en la diferencia. Añadir G como axioma no cierra la brecha: el nuevo sistema tiene su propia G.

Idea clave

Qué dice esto en cristiano. La demostración construye, dentro del propio sistema, un enunciado que dice "yo no soy demostrable". Si el sistema lo demostrara sería inconsistente, y si lo refutara también; luego ni lo demuestra ni lo refuta, y sin embargo es verdadero, porque en efecto no es demostrable. El truco técnico es la *codificación de Gödel*, que asigna un número a cada fórmula y convierte afirmaciones sobre demostraciones en afirmaciones aritméticas. Es el mismo mecanismo de autorreferencia de la paradoja del mentiroso, hecho riguroso.

Ejemplo 2.14 (Qué se cae y qué no)**. (a) Lo que muere es el programa de Hilbert.** Hilbert había propuesto en 1920 fundamentar toda la matemática en un sistema formal del que se demostrara, con métodos finitistas, que era completo y consistente. El primer teorema mata la completitud y el segundo la posibilidad de esa demostración de consistencia. En once años se pasó del optimismo total a la imposibilidad demostrada.

(b) Enunciados concretos indecidibles. No son solo rarezas autorreferenciales: el teorema de Goodstein, sobre unas sucesiones de enteros que sorprendentemente siempre acaban en cero, es verdadero y no demostrable en la aritmética de Peano. La hipótesis del continuo es independiente de los axiomas usuales de la teoría de conjuntos, como demostraron Gödel en 1940 y Cohen en 1963.

(c) Lo que no dice. No dice que haya verdades incognoscibles: el enunciado de Gödel de un sistema *sí* se demuestra en un sistema más fuerte. Tampoco dice que la matemática sea inconsistente, ni que haya cosas que las máquinas no puedan hacer y los humanos sí; ese argumento, popular en divulgación, no se sostiene, porque el humano tampoco puede demostrar la consistencia de su propio sistema.

(d) El primo hermano computacional. El problema de la parada es indecidible: no existe ningún programa que, dado otro programa, decida siempre si terminará. Turing lo demostró en 1936 con el mismo esquema de autorreferencia, y de ahí sale que tampoco hay

algoritmo para decidir si una fórmula de primer orden es demostrable. Lógica y computación resultaron ser la misma cosa vista desde dos lados.

Cuidado

[Las cuatro hipótesis del teorema son todas necesarias] El sistema debe ser *consistente* (de uno contradictorio se demuestra todo), tener axiomas *efectivamente enumerables* (con la lista completa de todos los enunciados verdaderos de la aritmética no hay incompletitud, pero esa lista no es reconocible por ninguna máquina), y ser *suficientemente potente*: la aritmética solo con la suma, sin producto, es completa y decidible, resultado de Presburger de 1929. Segundo aviso: usar el teorema fuera de la matemática para argumentar sobre la mente, la conciencia o los límites de la ciencia es abuso; el enunciado habla de sistemas formales con aritmética dentro, y nada más.

Pausa para pensar

Gödel presentó sus resultados en 1930 en Königsberg, en la misma conferencia donde Hilbert pronunció su célebre discurso terminado con la frase que está grabada en su tumba: debemos saber, sabremos. Hilbert no estaba en la sala cuando Gödel habló, y la trascendencia del anuncio la captó casi solo Von Neumann, que aquella tarde dedujo por su cuenta el segundo teorema y escribió a Gödel para descubrir que este ya lo tenía. Gödel tenía veinticuatro años. Es probablemente el caso más claro de un resultado que cambió por completo la idea que una disciplina tenía de sí misma.

2.4. Computabilidad: lo que ninguna máquina puede decidir

Los teoremas de incompletitud dicen que hay verdades que ningún sistema formal razonable demuestra. La teoría de la computabilidad dice algo emparentado y a la vez independiente: hay preguntas perfectamente bien planteadas que ningún algoritmo responde. Las dos familias de resultados se demostraron casi a la vez, hacia 1936, y hoy sabemos que son dos caras del mismo fenómeno diagonal.

Definición 2.15. Una *máquina de Turing* es una tupla $(Q, \Sigma, \Gamma, \delta, q_0, q_{\text{ac}}, q_{\text{re}})$ formada por un conjunto finito de estados Q, un alfabeto de entrada Σ, un alfabeto de cinta $\Gamma \supseteq \Sigma$ con un símbolo blanco, y una función de transición

$$\delta \colon (Q \setminus \{q_{\text{ac}}, q_{\text{re}}\}) \times \Gamma \longrightarrow Q \times \Gamma \times \{I, D\}$$

que, según el estado actual y el símbolo leído, indica el estado siguiente, el símbolo que se escribe y hacia dónde se mueve el cabezal. Un conjunto $A \subseteq \Sigma^*$ es *decidible* si alguna máquina se detiene siempre y acepta exactamente los elementos de A; es *semidecidible* o *recursivamente enumerable* si alguna máquina acepta los elementos de A y no se detiene (o rechaza) en los demás.

Ejemplo 2.16 (Máquinas de Turing pequeñas)**. (a) Sumar uno en binario.** Con la entrada escrita en binario, la máquina se coloca al final, y mientras lea un 1 lo cambia por 0 y retrocede;

al leer el primer 0 (o el blanco) escribe 1 y para. Bastan tres estados. Con la entrada 1011 el resultado es 1100.

(b) Reconocer palíndromos. La máquina compara el primer símbolo con el último, los borra, y repite. Reconoce el lenguaje de los palíndromos, que ningún autómata finito puede reconocer: la cinta hace de memoria ilimitada, y eso es exactamente lo que separa a las máquinas de Turing de los autómatas.

(c) Lo que no cambia nada. Añadir más cintas, permitir un alfabeto mayor, dejar que el cabezal se quede quieto o hacer la máquina no determinista no aumenta la clase de lenguajes reconocibles: todas esas variantes se simulan entre sí. Sí cambia la *eficiencia*, y ahí es donde vive la pregunta de si P es igual a NP.

(d) La máquina universal. Existe una única máquina U que, dada la descripción de cualquier máquina M y una entrada w, simula a M sobre w. Ésa es la idea del ordenador de programa almacenado: el mismo aparato ejecuta cualquier programa porque el programa es un dato más. Turing la describió en 1936, diez años antes de que se construyera ninguna.

Idea clave

Qué dice esto en cristiano. Una máquina de Turing es la formalización más austera posible de lo que significa "seguir un procedimiento mecánico": una cinta infinita cuadriculada, un cabezal que lee y escribe un símbolo cada vez, y una tabla finita de reglas del tipo "si estoy en el estado 7 y leo un 1, escribo 0, me muevo a la derecha y paso al estado 3". Nada más. La distinción entre decidible y semidecidible es la distinción entre *poder contestar sí y no* y *poder contestar sólo sí*: en el segundo caso, si la respuesta es no, la máquina se queda calculando para siempre y el que espera nunca sabe si es que la respuesta es no o si es que falta un rato más.

Teorema 2.17 (Indecidibilidad del problema de la parada)**.** *El conjunto*

$$K = \{\langle M, w\rangle : \textit{la máquina } M \textit{ se detiene con entrada } w\}$$

es semidecidible pero no decidible.

Demostración. Es semidecidible: basta simular M sobre w y aceptar si la simulación termina. Supongamos que fuera decidible por una máquina H que siempre se detiene y responde correctamente. Construimos D que, con entrada $\langle M\rangle$, ejecuta $H(\langle M, \langle M\rangle\rangle)$ y hace lo contrario: si H dice "se detiene", D entra en un bucle infinito; si H dice "no se detiene", D se detiene. Preguntemos ahora qué hace D con entrada $\langle D\rangle$. Si D se detiene, es porque H dijo que no se detenía, contradicción; si no se detiene, es porque H dijo que sí se detenía, contradicción también. Luego H no existe. □

Regla: Demostrar que un problema es indecidible por reducción

1. **Elegir un problema A ya conocido como indecidible**, típicamente K o su complemento.
2. **Fijar la dirección**: para probar que B es indecidible se reduce A a B ($A \leq_m B$),

nunca al revés. Reducir B a A no demuestra nada sobre B.

3. **Construir una función computable total** f que transforme instancias de A en instancias de B.
4. **Comprobar la equivalencia** $x \in A \iff f(x) \in B$, en las dos direcciones.
5. **Verificar que** f **es efectivamente calculable** y que *siempre* termina: una reducción que no termina no es una reducción.
6. **Concluir**: si B fuese decidible, componiendo el decisor de B con f se decidiría A, que es imposible.
7. **Para semidecidibilidad**, recordar que A es decidible si y sólo si A y su complemento son ambos semidecidibles; ésta es la vía rápida para separar niveles.

Ejemplo 2.18 (Cuatro problemas y su estatus)**. (a) ¿Acepta** M **la palabra vacía?** Indecidible. Dada una instancia $\langle M, w\rangle$ de la parada, se construye M_w que ignora su entrada, escribe w en la cinta y ejecuta M. Entonces M_w acepta la palabra vacía exactamente cuando M acepta w. La construcción de M_w a partir de M y w es puramente sintáctica y siempre termina, luego es una reducción válida.

(b) ¿Es $L(M)$ **vacío?** Indecidible, y ni siquiera semidecidible. Su complemento ("$L(M)$ es no vacío") sí lo es: se simulan en paralelo todas las entradas, avanzando un paso de cada una por turnos, y se acepta en cuanto alguna acepte. Si $L(M) = \emptyset$ fuese semidecidible, ambos lo serían y el problema sería decidible, lo que contradice la reducción desde K.

(c) ¿Es M **equivalente a** M'**?** Indecidible: contiene el caso (b) tomando M' una máquina que rechaza todo. Aquí ni el problema ni su complemento son semidecidibles, lo que lo sitúa estrictamente por encima de la parada en la jerarquía.

(d) ¿Tiene M **más de siete estados?** Decidible, y de hecho trivial: se cuentan los estados en la descripción de M. Esto marca la frontera exacta: preguntas sobre la *descripción sintáctica* de la máquina son decidibles; preguntas sobre el *lenguaje que reconoce* casi nunca lo son.

Teorema 2.19 (Rice)**.** *Sea P una propiedad de lenguajes recursivamente enumerables que no es ni siempre verdadera ni siempre falsa. Entonces el problema de decidir, dada $\langle M\rangle$, si $L(M)$ tiene la propiedad P es indecidible.*

Idea clave

Qué dice esto en cristiano. El teorema de Rice es una máquina de fabricar indecidibilidades: ninguna propiedad interesante del *comportamiento* de un programa se puede decidir automáticamente. "¿Este programa termina siempre?", "¿calcula la función identidad?", "¿acepta algún número primo?", "¿es equivalente a este otro?": todas indecidibles, sin necesidad de construir una reducción para cada una. La única escapatoria son las propiedades *triviales* (las que cumplen todos los lenguajes o ninguno) y las propiedades del código fuente en lugar del comportamiento. Por eso los verificadores

de programas reales son necesariamente incompletos: o rechazan programas correctos, o aceptan incorrectos, o no siempre terminan. No es un defecto de ingeniería, es un teorema.

Cuidado

[Contraejemplo: Rice no dice que todo sea indecidible] Es tentador leer Rice como "nada se puede analizar automáticamente", y es falso por dos motivos. Primero, la propiedad debe ser *del lenguaje*: "M tiene un bucle sintáctico" o "M escribe alguna vez el símbolo 0 en los primeros cien pasos" son decidibles, porque hablan de la máquina y no de lo que reconoce. Segundo, restringiendo el modelo de cómputo todo cambia: para autómatas finitos, la vacuidad y la equivalencia son perfectamente decidibles, y los algoritmos son eficientes. La indecidibilidad es un precio que se paga por la potencia de cálculo, no una maldición universal.

Ejemplo 2.20 (La conexión con Gödel)**. (a) De la parada a la incompletitud.** Si un sistema formal T consistente, recursivamente axiomatizado y suficientemente expresivo demostrase, para cada M y w, o bien "M para con w" o bien su negación, entonces enumerando demostraciones se decidiría la parada: basta ir generando todas las pruebas hasta encontrar una de las dos. Como la parada es indecidible, T deja alguna de esas afirmaciones sin decidir. Eso es el primer teorema de incompletitud, obtenido sin construir ninguna sentencia autorreferente.

(b) La diagonal común. Tanto Gödel como Turing usan el mismo truco: alimentar a un objeto con su propia descripción. Gödel lo hace con la aritmetización de la sintaxis y la sentencia "yo no soy demostrable"; Turing lo hace con $D(\langle D\rangle)$. Cantor lo había hecho en 1891 con los decimales.

(c) La función del castor afanoso. $\Sigma(n)$, el máximo número de unos que puede escribir una máquina de n estados que se detenga, es una función bien definida y no computable; crece más deprisa que cualquier función computable. Se conocen $\Sigma(1) = 1$, $\Sigma(2) = 4$, $\Sigma(3) = 6$, $\Sigma(4) = 13$, y a partir de ahí sólo cotas inferiores gigantescas. Conocer $\Sigma(n)$ resolvería la parada para máquinas de n estados.

(d) Lo que no se sigue. Nada de esto implica que la mente humana supere a las máquinas. El argumento habitual ("yo veo que la sentencia de Gödel es verdadera, luego no soy un algoritmo") olvida que esa verdad es *condicional* a la consistencia del sistema, y que nadie ha comprobado la consistencia del sistema que uno mismo sería.

Pausa para pensar

En 1936, Church publicó su solución del *Entscheidungsproblem* usando el cálculo lambda unos meses antes de que Turing publicara la suya con máquinas. Turing añadió un apéndice demostrando que ambos formalismos definen la misma clase de funciones, y esa coincidencia, repetida después con las funciones recursivas de Gödel–Herbrand, los sistemas de Post y todos los lenguajes de programación inventados desde entonces, es lo que sostiene la *tesis de Church–Turing*: "computable" significa lo mismo en todos

los modelos razonables. No es un teorema, porque "razonable" no está definido; es una regularidad empírica tan firme que nadie espera verla caer.

2.5. Computabilidad: qué se puede calcular y qué no

La incompletitud y la incomputabilidad son la misma cosa vista desde dos ángulos. Turing lo demostró en 1936, con un argumento que además fundó la informática y que es más corto y más transparente que el de Gödel.

Definición 2.21. Una *máquina de Turing* es una cinta infinita, un alfabeto finito, un conjunto finito de estados y una tabla que, dado el estado y el símbolo leído, indica qué escribir, hacia dónde moverse y a qué estado pasar. Una función es *computable* si alguna máquina de Turing la calcula. La *tesis de Church–Turing* afirma que esa noción coincide con la idea intuitiva de "calculable mediante un procedimiento mecánico".

Idea clave

Qué dice esto en cristiano. La definición parece pobre —una cinta y una tabla— y ésa es la gracia: cualquier cosa que se le añada no aumenta lo que puede calcular. Se han propuesto docenas de modelos alternativos, del cálculo lambda de Church a las funciones recursivas de Gödel y Kleene, pasando por cualquier lenguaje de programación real, y todos definen exactamente la misma clase de funciones. Esa coincidencia masiva es la evidencia a favor de la tesis, que no es un teorema —relaciona un concepto formal con uno intuitivo— sino una afirmación empírica que lleva noventa años sin contraejemplos.

Teorema 2.22 (Indecidibilidad del problema de la parada)**.** *No existe ninguna máquina de Turing que, dada la descripción de una máquina M y una entrada w, decida siempre correctamente si M acaba parando sobre w.*

Idea clave

La demostración, en cuatro frases. Supóngase que existe $H(M, w)$ que responde sí o no. Constrúyase $D(M)$ que ejecute $H(M, M)$ y haga lo contrario: si H dice que M para sobre sí misma, D entra en un bucle infinito; si dice que no para, D se detiene. Ahora pregúntese qué hace $D(D)$. Si para, es porque H dijo que no paraba; si no para, es porque H dijo que paraba. Contradicción en ambos casos, luego H no existe. Es el argumento diagonal de Cantor aplicado a programas, y la misma idea que la sentencia de Gödel, sin necesidad de codificar aritmética.

Ejemplo 2.23 (Problemas que no se pueden decidir)**. (a) El teorema de Rice.** Toda propiedad no trivial de la *función* que computa un programa es indecidible. No sólo la parada: tampoco se puede decidir si dos programas calculan lo mismo, si un programa devuelve alguna vez cero, o si acepta algún dato. Sólo son decidibles las propiedades sintácticas, del texto del programa, no las semánticas.

(b) El décimo problema de Hilbert. Decidir si una ecuación diofántica tiene solución entera es indecidible. Lo demostraron Davis, Putnam, Robinson y Matiyasevich entre 1950 y 1970. Un problema formulado en 1900 como "encuéntrese el algoritmo" resultó no tener ninguno.

(c) El problema de la palabra. Dada una presentación de un grupo por generadores y relaciones, decidir si dos palabras representan el mismo elemento es indecidible en general —Novikov, 1955—. Por eso la teoría combinatoria de grupos trabaja con clases restringidas.

(d) El castor afanoso. La función que da el máximo número de pasos que puede dar una máquina de n estados antes de parar crece más deprisa que cualquier función computable. Se conocen sus valores hasta $n = 5$: para $n = 5$ vale $47\,176\,870$, resultado establecido por verificación formal en 2024. Para $n = 6$ se sabe que supera $10^{10^{10^7}}$, y probablemente nunca se sepa cuánto vale exactamente.

Cuidado

[Contraejemplo: indecidible no significa imposible en la práctica] Que el problema de la parada sea indecidible no impide que un analizador estático demuestre la terminación de la inmensa mayoría de los programas reales. Lo que el teorema prohíbe es un método que acierte *siempre* y con *todos* los programas; permite perfectamente un método que responda "sí", "no" o "no lo sé", y que el tercer caso sea raro en la práctica. Confundir ambas cosas lleva a conclusiones filosóficas exageradas sobre los límites de la máquina, y también a las igualmente exageradas sobre la mente humana, que tampoco decide el problema de la parada.

Pausa para pensar

Turing tenía veintitrés años cuando escribió el artículo, en 1936, y su objetivo no era fundar la informática sino resolver el *Entscheidungsproblem* de Hilbert. La máquina fue un medio, no un fin: la inventó para tener una definición precisa de "procedimiento mecánico" que le permitiera demostrar que *no* existía uno. El aparato conceptual más influyente del siglo XX en tecnología nació como herramienta auxiliar de una demostración de imposibilidad.

2.6. Teoría de conjuntos: ZFC, ordinales, cardinales y el axioma de elección

Idea clave

La teoría ingenua de conjuntos —"un conjunto es una colección cualquiera de objetos"— es *contradictoria*. Considérese

$$R = \{x : x \notin x\},$$

el conjunto de los conjuntos que no se pertenecen a sí mismos. ¿Se pertenece R a sí mismo? Si sí, por definición no debería; si no, por definición debería. La paradoja de

Russell (1901) no es un juego de palabras: destruyó el programa de Frege cuando el segundo volumen ya estaba en imprenta.

La reparación de Zermelo consiste en no permitir formar conjuntos por comprensión libre, sino solo *dentro* de un conjunto ya existente.

Regla: los axiomas de Zermelo–Fraenkel

- **Extensionalidad.** Dos conjuntos con los mismos elementos son iguales.
- **Par, unión, partes.** Existen $\{a, b\}$, $\bigcup A$ y $\mathcal{P}(A)$.
- **Separación.** Para cada propiedad P y cada conjunto A existe $\{x \in A : P(x)\}$. Aquí es donde muere la paradoja: R no se puede formar, solo $R \cap A$ para un A dado, y ese conjunto simplemente no pertenece a A.
- **Reemplazo.** La imagen de un conjunto por una función definible es un conjunto.
- **Infinito.** Existe un conjunto inductivo, y de ahí $\mathbb{N}$.
- **Regularidad.** Todo conjunto no vacío tiene un elemento $\in$-minimal; en particular ningún conjunto se pertenece a sí mismo.
- **Elección (AC).** Todo producto de conjuntos no vacíos es no vacío.

Observación 2.24. No existe "el conjunto de todos los conjuntos": si existiera, la separación permitiría formar R dentro de él. Las colecciones demasiado grandes son *clases propias*, y viven fuera de la teoría como abreviaturas del lenguaje. Por eso las categorías necesitan un cuidado especial: **Set** no es un objeto de **Set**.

Definición 2.25 (Ordinales de von Neumann)**.** Un ordinal es un conjunto transitivo y bien ordenado por $\in$. Se construyen desde abajo:

$$0 = \varnothing, \quad 1 = \{0\}, \quad 2 = \{0, 1\}, \quad \dots \quad \omega = \{0, 1, 2, \dots\}, \quad \omega + 1 = \omega \cup \{\omega\}, \ \dots$$

Cada ordinal *es* el conjunto de los ordinales menores que él.

Ejemplo 2.26 (La aritmética ordinal no es conmutativa)**.** Léase $\alpha + \beta$ como "pon una copia de α y detrás una de β".

$$1 + \omega = \omega, \qquad \omega + 1 \neq \omega.$$

En efecto, poner un elemento delante de $\mathbb{N}$ da un orden isomorfo a $\mathbb{N}$ (basta renumerar); ponerlo detrás produce un orden con un último elemento, y $\mathbb{N}$ no lo tiene. Igualmente $2 \cdot \omega = \omega$ y $\omega \cdot 2 = \omega + \omega$, que es el orden de $\mathbb{N}$ seguido de otra copia de $\mathbb{N}$: un buen orden en el que hay un elemento con infinitos predecesores.

Se puede realizar sobre $\mathbb{N}$ mismo: ordénese $0, 2, 4, 6, \dots, 1, 3, 5, 7, \dots$ (primero los pares, luego los impares). Es un buen orden de tipo $\omega + \omega$ sobre un conjunto numerable.

Definición 2.27 (Cardinales)**.** $|A| = |B|$ si hay una biyección. Un cardinal es un ordinal que no es equipotente a ninguno menor. Los cardinales infinitos se enumeran $\aleph_0, \aleph_1, \aleph_2, \ldots$

Teorema 2.28 (Aritmética cardinal)**.** *Para cardinales infinitos,* $\kappa + \lambda = \kappa \cdot \lambda = \max(\kappa, \lambda)$. *En particular* $\aleph_0 \cdot \aleph_0 = \aleph_0$. *La exponenciación, en cambio, sí crece: por el argumento diagonal de Cantor,*

$$2^\kappa > \kappa \qquad \textit{siempre},$$

y en particular $\mathfrak{c} = 2^{\aleph_0} > \aleph_0$.

La **hipótesis del continuo** —que $\mathfrak{c} = \aleph_1$, es decir, que no hay cardinales intermedios— es indecidible en ZFC: Gödel probó en 1938 que no se puede refutar y Cohen en 1963 que no se puede demostrar. Es el ejemplo más limpio de una pregunta matemática perfectamente formulada cuya respuesta los axiomas no determinan.

Teorema 2.29 (Las tres caras del axioma de elección)**.** *Sobre ZF, son equivalentes:*

1. *Todo producto de conjuntos no vacíos es no vacío (AC).*
2. *Todo conjunto parcialmente ordenado en el que toda cadena tiene cota superior posee un elemento maximal (lema de Zorn).*
3. *Todo conjunto admite un buen orden (Zermelo).*

Observación 2.30. Estas tres versiones tienen grados de plausibilidad muy distintos —Jerry Bona bromeaba con que AC es obviamente cierto, el principio del buen orden obviamente falso, y sobre Zorn quién sabe— y sin embargo son el mismo enunciado. Zorn es la forma en que se usa en la práctica: de él salen la existencia de base en todo espacio vectorial, la de ideal maximal en todo anillo, la de clausura algebraica, el teorema de Hahn–Banach y el de Tichonov. Y también consecuencias incómodas: la partición de Banach–Tarski de la esfera y la existencia de conjuntos no medibles de Vitali. Aceptar el axioma es aceptar el paquete completo.

2.7. Lo que se puede calcular pero no deprisa: P, NP y la NP-completitud

Los dos apartados anteriores tratan de lo *indecidible*. Falta la otra mitad de la teoría de la computación, que trata de lo decidible pero caro, y que es donde vive el problema abierto más famoso de la informática.

Definición 2.31 (Las dos clases)**.** P es la clase de los problemas de decisión resolubles por una máquina de Turing determinista en tiempo polinómico en la longitud de la entrada. NP es la clase de los resolubles por una máquina *no determinista* en tiempo polinómico o, de manera equivalente y más útil, la clase de los problemas cuyas respuestas afirmativas admiten un **certificado** verificable en tiempo polinómico.

fuera del rectángulo: la parada, Rice, el décimo de Hilbert

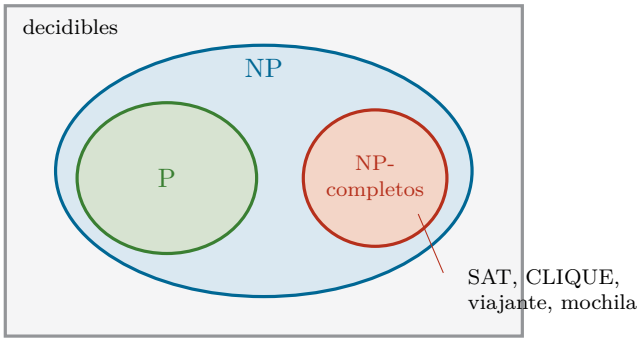


Figura 2.2: Indecidible e intratable no son lo mismo. Fuera del rectángulo no hay algoritmo ninguno; dentro lo hay, y la pregunta es si es rápido. Se sabe que $\mathrm{P} \subseteq \mathrm{NP}$ y que los NP-completos son los más difíciles de NP; si alguno cayera en P, las tres regiones interiores colapsarían en una sola.

La caracterización por certificados es la que hay que retener. Encontrar una asignación que satisfaga una fórmula puede costar una eternidad; *comprobar* que una asignación dada la satisface cuesta leer la fórmula. Factorizar un número de mil cifras es carísimo; verificar una factorización es una multiplicación. $\mathrm{P} = \mathrm{NP}$ preguntaría si encontrar es tan fácil como verificar, es decir, si la creatividad se puede automatizar.

Definición 2.32 (Reducción y NP-completitud)**.** A se **reduce** a B (en el sentido de Karp) si hay una función computable en tiempo polinómico f con $x \in A \iff f(x) \in B$. Un problema es **NP-completo** si está en NP y todo problema de NP se reduce a él.

Teorema 2.33 (Cook–Levin, 1971)**.** *SAT —decidir si una fórmula booleana es satisfacible— es NP-completo.*

La demostración codifica el funcionamiento completo de una máquina no determinista con cota polinómica como una fórmula gigantesca cuya satisfacibilidad equivale a la existencia de un cómputo aceptante. A partir de ahí, para probar que un problema nuevo es NP-completo basta reducirle uno ya conocido, y Karp exhibió veintiún problemas en 1972. Hoy son miles.

Ejemplo 2.34 (De 3-SAT a CLIQUE, con una fórmula concreta)**.** Sea

$$\varphi = (x \vee y \vee \bar{z}) \wedge (\bar{x} \vee y \vee z) \wedge (x \vee \bar{y} \vee z).$$

(a) El grafo. Se dibujan tres columnas de tres vértices: un vértice por cada literal de cada cláusula, nueve en total. Se une un vértice con otro si y solo si están en *cláusulas distintas* y no son *contradictorios* (es decir, no son x y $\bar{x}$).

(b) La equivalencia. φ es satisfacible si y solo si el grafo tiene un clique de tamaño 3. En efecto, un clique de tamaño 3 toma exactamente un literal por cláusula (los de la misma columna nunca están unidos) y esos tres literales son compatibles entre sí, luego se pueden poner todos a verdadero. Y al revés: una asignación satisfactoria elige un literal verdadero por cláusula, y esos tres forman clique.

(c) La comprobación. Con $x = \mathrm{V}$, $y = \mathrm{V}$, $z = \mathrm{V}$ se satisfacen las tres cláusulas, y el clique correspondiente es $\{x$ de C1, y de C2, z de C3$\}$: los tres están en cláusulas distintas y ninguno es la negación de otro. ✓

(d) El coste. La construcción del grafo es cuadrática en el tamaño de la fórmula, luego polinómica. Eso es lo que significa reducir: si alguien encontrara un algoritmo rápido para CLIQUE, tendríamos uno rápido para 3-SAT, y por Cook–Levin para todo NP.

Cuidado

Indecidible e intratable son cosas distintas y se confunden constantemente. El problema de la parada es **indecidible**: no hay algoritmo, ni lento ni rápido, ni con toda la eternidad por delante. El problema del viajante es **decidible**: hay un algoritmo obvio —probar todas las rutas— que termina siempre; lo que no se conoce es uno eficiente. La frontera de Turing es absoluta; la de Cook es, hasta donde sabemos, solo práctica. Y podría no serlo: si P = NP, desaparecería.

Observación 2.35. Las consecuencias de P = NP serían devastadoras y hermosas a la vez. Toda la criptografía de clave pública se derrumbaría, porque su seguridad descansa en que factorizar y calcular logaritmos discretos son caros. Pero también sería automática la búsqueda de demostraciones de longitud acotada, el diseño óptimo de circuitos, el plegamiento de proteínas y buena parte del aprendizaje automático. La opinión mayoritaria, sin demostración, es que P $\neq$ NP; y se sabe que las técnicas conocidas —diagonalización, circuitos naturales, pruebas algebrizantes— son demostrablemente insuficientes para zanjarlo. Es uno de los siete problemas del milenio, y hay quien argumenta que, de resolverse afirmativamente, arrastraría consigo a los demás: buscar demostraciones de longitud acotada pasaría a ser tratable.

2.8. Indefinibilidad de la verdad, el segundo teorema y los modelos no estándar

El primer teorema de incompletitud se ha enunciado y contado. Faltan sus dos hermanos, y ambos salen del mismo lema.

Teorema 2.36 (Lema diagonal)**.** *Para cualquier fórmula $\varphi(x)$ con una variable libre existe una sentencia G tal que*

$$T \vdash G \leftrightarrow \varphi(\ulcorner G \urcorner),$$

donde $\ulcorner G \urcorner$ es el número de Gödel de G.

En cristiano: toda propiedad expresable admite una sentencia que *afirma de sí misma* que la cumple. La autorreferencia, que parecía un juego de palabras, es un teorema. De ahí sale todo lo demás por sustitución.

Regla: tres diagonalizaciones y tres teoremas

- Con $\varphi(x) = \neg\mathrm{Dem}_T(x)$ sale G = "yo no soy demostrable": primer teorema de incompletitud.
- Con $\varphi(x) = \mathrm{Dem}_T(x)$ sale la sentencia de Henkin, H = "yo sí soy demostrable". Y aquí la sorpresa: H *es* demostrable, por el teorema de Löb.
- Con $\varphi(x) = \neg\mathrm{Verdad}(x)$ sale "yo soy falsa", la paradoja del mentiroso. Como una contradicción es imposible, lo que falla es la hipótesis: el predicado Verdad no existe.

Teorema 2.37 (Indefinibilidad de la verdad, Tarski 1936)**.** *No existe ninguna fórmula $V(x)$ del lenguaje de la aritmética tal que, para toda sentencia σ,*

$$\mathbb{N} \models V(\ulcorner\sigma\urcorner) \iff \mathbb{N} \models \sigma.$$

Compárese con la demostrabilidad, que *sí* es expresable: Dem_T existe como fórmula. Ahí está la asimetría fundamental. Un sistema formal puede hablar de sus propias demostraciones pero no de su propia verdad, y la incompletitud es exactamente el hueco entre ambas cosas.

Teorema 2.38 (Segundo teorema de incompletitud)**.** *Si T es consistente, recursivamente axiomatizable y contiene la aritmética suficiente, entonces*

$$T \nvdash \mathrm{Con}(T),$$

donde $\mathrm{Con}(T)$ *abrevia* $\neg\mathrm{Dem}_T(\ulcorner 0 = 1\urcorner)$.

Esquema. El primer teorema dice, informalmente, "si T es consistente entonces G no es demostrable". Ese razonamiento se puede *formalizar dentro de* T, y lo que se obtiene es $T \vdash \mathrm{Con}(T) \to G$. Si además T demostrase $\mathrm{Con}(T)$, entonces demostraría G, contradiciendo el primer teorema. La formalización necesita las tres condiciones de derivabilidad de Hilbert–Bernays sobre Dem_T, que es la parte laboriosa. □

Observación 2.39. Lo que *no* dice el teorema: no dice que la consistencia de la aritmética sea dudosa. Gentzen la demostró en 1936 por inducción transfinita hasta el ordinal ε_0, un argumento perfectamente convincente que simplemente usa un principio que la propia aritmética no puede formalizar. El segundo teorema no cierra la pregunta; dice dónde hay que ir a buscar la respuesta.

Teoría de modelos: por qué $\mathbb{N}$ no se puede fijar

Teorema 2.40 (Löwenheim–Skolem)**.** *Si una teoría de primer orden en un lenguaje numerable tiene un modelo infinito, entonces tiene modelos de* todo *cardinal infinito.*

Ejemplo 2.41 (Un modelo no estándar de la aritmética, construido)**.** Tómese la teoría completa de $\mathbb{N}$ —*todas* las sentencias verdaderas en los naturales, no solo los axiomas de Peano— y añádase una constante nueva c junto con el esquema infinito

$$c > 0, \quad c > 1, \quad c > 2, \quad c > 3, \ \ldots$$

Cualquier subconjunto *finito* de esas sentencias es satisfacible en $\mathbb{N}$ mismo: basta interpretar c como un número mayor que todos los mencionados. Por compacidad, el conjunto entero tiene un modelo M.

En M hay un elemento c mayor que todos los naturales estándar. Y sin embargo M satisface exactamente las mismas sentencias de primer orden que $\mathbb{N}$: es *elementalmente equivalente* y no isomorfo. Su orden es $\mathbb{N}$ seguido de una colección de copias de $\mathbb{Z}$ ordenadas densamente y sin extremos.

Ahí está la respuesta a la afirmación de que ninguna cantidad de axiomas de primer orden fija $\mathbb{N}$ salvo isomorfismo. El culpable es el propio primer orden: el axioma de inducción *de segundo orden* sí caracteriza a $\mathbb{N}$, pero su lógica no tiene un sistema deductivo completo. Se puede tener caracterización o se puede tener completitud, no las dos.

Observación 2.42. La misma construcción, aplicada a $\mathbb{R}$ con una constante infinitesimal, produce los **hiperreales** y con ellos el análisis no estándar de Robinson: una fundamentación rigurosa de los infinitésimos de Leibniz, tres siglos después de que Berkeley los llamara "fantasmas de cantidades difuntas".

Y en la otra dirección, el criterio de Łoś–Vaught da un método sorprendentemente barato para demostrar completitud: si una teoría sin modelos finitos es *categórica* en algún cardinal infinito —todos sus modelos de ese tamaño son isomorfos—, entonces es completa. Así se prueba de un plumazo que la teoría de los cuerpos algebraicamente cerrados de característica fija es completa y decidible.

2.9. Estructuras, satisfacción y conectivos completos: la semántica que el capítulo daba por sabida

El capítulo arrancó distinguiendo $\Gamma \vdash \varphi$ de $\Gamma \models \varphi$ y enunció el teorema de completitud, pero nunca dijo qué es exactamente un modelo ni qué significa el símbolo $\models$. Tampoco definió "recursivamente axiomatizable", pese a que esa hipótesis aparece literalmente en el enunciado del segundo teorema de incompletitud, ni "conjunto completo de conectivos", que un ejercicio propuesto pide decidir. Las cuatro definiciones son elementales y caben en una sección; sin ellas, media docena de enunciados del capítulo son eslóganes.

Qué es un modelo

Definición 2.43 (Lenguaje y estructura)**.** Un *lenguaje de primer orden* $\mathcal{L}$ consta de un conjunto de símbolos de constante, un conjunto de símbolos de función y un conjunto de símbolos de relación, cada uno de estos dos últimos con una *aridad* asociada (el número de argumentos que toma). Una *estructura* o *modelo* para $\mathcal{L}$ es un par $\mathcal{M} = (M, \ \cdot^{\mathcal{M}})$ formado por:

- un conjunto no vacío M, el *universo* o dominio de la estructura;
- para cada símbolo de constante c, un elemento $c^{\mathcal{M}} \in M$;
- para cada símbolo de función g de aridad n, una función $g^{\mathcal{M}} : M^n \to M$ (total: definida en todas partes);
- para cada símbolo de relación R de aridad n, un subconjunto $R^{\mathcal{M}} \subseteq M^n$.

La igualdad no se interpreta: se toma siempre como la identidad de M.

Ejemplo 2.44 (Un lenguaje, tres estructuras distintas)**.** Sea $\mathcal{L} = \{0, 1, +, \cdot, <\}$ con $0, 1$ constantes, $+$ y $\cdot$ funciones binarias y $<$ una relación binaria. Las siguientes son tres estructuras para el mismo lenguaje, y las tres son legítimas:

- $\mathbb{N}$ con su aritmética y su orden usuales;
- $\mathbb{Q}$ con lo mismo;
- el universo $M = \{a, b\}$ con $0^{\mathcal{M}} = a$, $1^{\mathcal{M}} = b$, la suma dada por la tabla del o exclusivo, el producto por la del y lógico, y $<^{\mathcal{M}} = \{(a, b)\}$.

La sentencia $\forall x \, \forall y \, (x < y \to \exists z \, (x < z \wedge z < y))$ —la densidad del orden— es falsa en la primera, verdadera en la segunda y falsa en la tercera. El mismo texto, tres veredictos: eso es exactamente lo que la semántica añade a la sintaxis.

Definición 2.45 (Satisfacción, modelo de una teoría, validez y consecuencia)**.** Sea $\mathcal{M}$ una estructura y s una *asignación*, esto es, una función de las variables en M. El *valor* $\bar{s}(t)$ de un término se define por recursión: $\bar{s}(x) = s(x)$, $\bar{s}(c) = c^{\mathcal{M}}$ y $\bar{s}(g(t_1, \ldots, t_n)) = g^{\mathcal{M}}(\bar{s}(t_1), \ldots, \bar{s}(t_n))$. La relación de *satisfacción* $\mathcal{M} \models \varphi \, [s]$ se define entonces por recursión sobre la fórmula:

$$
\begin{aligned}
\mathcal{M} \models R(t_1, \ldots, t_n) \, [s] &\iff (\bar{s}(t_1), \ldots, \bar{s}(t_n)) \in R^{\mathcal{M}};\\
\mathcal{M} \models (t_1 = t_2) \, [s] &\iff \bar{s}(t_1) = \bar{s}(t_2);\\
\mathcal{M} \models \neg\varphi \, [s] &\iff \text{no } \mathcal{M} \models \varphi \, [s];\\
\mathcal{M} \models (\varphi \wedge \psi) \, [s] &\iff \mathcal{M} \models \varphi \, [s] \text{ y } \mathcal{M} \models \psi \, [s];\\
\mathcal{M} \models (\varphi \vee \psi) \, [s] &\iff \mathcal{M} \models \varphi \, [s] \text{ o } \mathcal{M} \models \psi \, [s];\\
\mathcal{M} \models \forall x \, \varphi \, [s] &\iff \mathcal{M} \models \varphi \, [s(x \mapsto a)] \text{ para todo } a \in M;\\
\mathcal{M} \models \exists x \, \varphi \, [s] &\iff \mathcal{M} \models \varphi \, [s(x \mapsto a)] \text{ para algún } a \in M,
\end{aligned}
$$

donde $s(x \mapsto a)$ es la asignación que coincide con s salvo que envía x a a. Si φ es una *sentencia* (sin variables libres), el valor de verdad no depende de s y se escribe $\mathcal{M} \models \varphi$.

Dada una teoría T (un conjunto de sentencias), $\mathcal{M}$ es *modelo de* T, escrito $\mathcal{M} \models T$, si satisface todas sus sentencias. Una sentencia φ es *válida*, escrito $\models \varphi$, si toda estructura la satisface; y es *consecuencia lógica* de Γ, escrito $\Gamma \models \varphi$, si todo modelo de Γ es modelo de φ.

Observación 2.46 (Por qué esta definición no es una trivialidad)*.* Parece que las cláusulas se limitan a repetir en castellano lo que los símbolos ya decían, y esa impresión es exactamente lo que hay que vencer. Tres puntos.

Primero: la recursión es sobre la *complejidad* de la fórmula, y por eso está bien fundada; no se está definiendo "verdadero" en términos de "verdadero", sino la verdad de una fórmula en términos de la verdad de fórmulas estrictamente más cortas. Esa observación de Tarski, de 1933, es la que convierte la noción de verdad en un objeto matemático manipulable y separa el lenguaje del que se habla del lenguaje en el que se habla.

Segundo: la definición se hace *fuera* de la estructura, en la metateoría. Ese detalle es la razón de que el teorema de indefinibilidad de Tarski no sea una contradicción: la verdad en $\mathbb{N}$ se define perfectamente en teoría de conjuntos, y lo que el teorema prohíbe es definirla mediante una fórmula del propio lenguaje de la aritmética.

Tercero: con $\models$ ya definido, la diferencia entre $\vdash$ y $\models$ deja de ser una intuición. La primera es un hecho combinatorio sobre cadenas finitas de símbolos —existe una lista de fórmulas con ciertas propiedades—, la segunda es una cuantificación sobre *todas* las estructuras, incluidas las de tamaño arbitrariamente grande. Que ambas coincidan, que es el teorema de completitud, no tenía por qué pasar, y de hecho en lógica de segundo orden no pasa.

Teorías recursivamente axiomatizables

Definición 2.47 (Teoría recursivamente axiomatizable)**.** Una teoría T es *recursivamente axiomatizable* (o *efectivamente axiomatizable*) si existe un conjunto A de sentencias tal que T y A tienen exactamente las mismas consecuencias, y además existe un *algoritmo* que, dada una fórmula cualquiera, se detiene siempre y responde correctamente a la pregunta "¿es esta fórmula un elemento de A?". Dicho de otro modo: el conjunto de axiomas es decidible, aunque sea infinito.

Ejemplo 2.48 (Infinito no significa inefectivo)**.** La aritmética de Peano de primer orden tiene infinitos axiomas, porque la inducción no es un axioma sino un *esquema*: uno por cada fórmula $\varphi(x)$. Y sin embargo es recursivamente axiomatizable, porque reconocer si una cadena de símbolos tiene la forma

$$\big(\varphi(0) \wedge \forall x\,(\varphi(x) \to \varphi(x+1))\big) \to \forall x\,\varphi(x)$$

es una comprobación puramente sintáctica que cualquier programa hace en tiempo finito. Lo mismo vale para ZFC, con sus esquemas de separación y reemplazo. La hipótesis no es "pocos axiomas" sino "axiomas reconocibles".

Observación 2.49 (Es la hipótesis exacta del segundo teorema, y sin ella no hay teorema)*.* El segundo teorema de incompletitud afirma que una teoría consistente, recursivamente axiomatizable y con aritmética suficiente no demuestra su propia consistencia. La hipótesis de axiomatizabilidad recursiva es la que permite *escribir* el predicado $\mathrm{Dem}_T(x)$ dentro del lenguaje de la aritmética: para decir "x codifica una demostración" hay que poder decir "cada línea es un axioma o se sigue de las anteriores", y la primera mitad de esa frase solo es expresable si ser axioma es decidible. Sin ella, $\mathrm{Con}(T)$ ni siquiera es una sentencia del lenguaje, y el enunciado del teorema se queda sin sujeto.

Qué pasa si se quita. Considérese la *aritmética verdadera*

$$\mathrm{Th}(\mathbb{N}) = \{\varphi :\ \mathbb{N} \models \varphi\},$$

el conjunto de todas las sentencias verdaderas en los naturales. Es consistente, tiene toda la aritmética que se quiera y es **completa** por construcción: para cada φ, o bien φ o bien $\neg\varphi$ es verdadera en $\mathbb{N}$, y por tanto es uno de sus teoremas. Así que no es incompleta y no contradice a Gödel. Lo que ocurre es que *no* es recursivamente axiomatizable, y no lo es por un argumento breve: si lo fuera, al ser además completa y consistente sería decidible —enumérense las demostraciones hasta encontrar la de φ o la de $\neg\varphi$, cosa que el ejercicio resuelto de este mismo capítulo hace en detalle—, y la aritmética es indecidible por el teorema de Church. Alternativamente, el teorema de Tarski lo da directamente: un conjunto de axiomas decidible haría definible la verdad aritmética.

Esa es la moraleja completa, y es más fina que el eslogan habitual. No es que ninguna teoría completa de la aritmética exista: existe, y es la lista de todas las verdades. Es que esa lista no es reconocible por ninguna máquina, así que no sirve como sistema formal. Lo que Gödel demostró es que *o* se tiene una lista de axiomas manejable *o* se tiene completitud, pero no las dos.

Conjuntos completos de conectivos

Definición 2.50 (Completitud funcional)**.** Una *función de verdad* n-aria es una aplicación $f : \{0,1\}^n \to \{0,1\}$; hay exactamente 2^{2^n} de ellas, es decir 4 unarias, 16 binarias y 256 ternarias. Un conjunto C de conectivos es *completo* (o *funcionalmente completo*) si toda función de verdad, de cualquier aridad, se puede expresar mediante una fórmula que use únicamente conectivos de C y variables proposicionales.

Teorema 2.51 ($\{\neg,\wedge\}$ es completo)**.** *Toda función de verdad se expresa con negación y conjunción solamente.*

Por forma normal disyuntiva. Sea $f : \{0,1\}^n \to \{0,1\}$. Si f es constantemente 0, la fórmula $p_1 \wedge \neg p_1$ la representa, y hemos terminado. En caso contrario, para cada vector $a = (a_1, \ldots, a_n)$ con $f(a) = 1$ constrúyase la conjunción de literales

$$K_a \;=\; \ell_1^{a_1} \wedge \ell_2^{a_2} \wedge \cdots \wedge \ell_n^{a_n}, \qquad \text{donde } \; \ell_i^1 = p_i \;\text{ y }\; \ell_i^0 = \neg p_i.$$

Por construcción K_a vale 1 en la asignación a y 0 en cualquier otra, porque cualquier discrepancia en una coordenada falsifica el literal correspondiente. Entonces la disyunción $\bigvee_{a:\, f(a)=1} K_a$ toma el valor 1 exactamente en los vectores donde f vale 1: es decir, representa f. Esa expresión es la *forma normal disyuntiva* de f, y hasta aquí se han usado $\neg$, $\wedge$ y $\vee$.

Falta eliminar la disyunción, y para eso basta una ley de De Morgan:

$$A \vee B \;\equiv\; \neg(\neg A \wedge \neg B),$$

igualdad que se comprueba con una tabla de cuatro filas. Sustituyendo cada $\vee$ de la forma normal por esa expresión se obtiene una fórmula equivalente escrita solo con $\neg$ y $\wedge$. Como f era arbitraria, $\{\neg,\wedge\}$ es completo. □

Ejemplo 2.52 (La lista que hay que saberse, con las cuentas hechas)**. Completos:** $\{\neg,\wedge\}$, $\{\neg,\vee\}$ y $\{\neg,\to\}$ —en el último, $A \vee B \equiv \neg A \to B$ y luego De Morgan—; también $\{\neg,\wedge,\vee\}$ y

cualquier superconjunto de un completo. Y, con la constante 1 disponible, $\{\oplus, \wedge, 1\}$, base de la representación de Zhegalkin como polinomio sobre $\mathbb{F}_2$: $\neg A \equiv A \oplus 1$.

Completos con un solo conectivo, y son los dos únicos binarios que lo consiguen:

- *NAND*, la barra de Sheffer $A \uparrow B \equiv \neg(A \wedge B)$. En efecto $\neg A \equiv A \uparrow A$ y $A \wedge B \equiv (A \uparrow B) \uparrow (A \uparrow B)$; con eso ya se tiene $\{\neg, \wedge\}$ y se aplica el teorema.
- *NOR*, la flecha de Peirce $A \downarrow B \equiv \neg(A \vee B)$. Análogamente $\neg A \equiv A \downarrow A$ y $A \vee B \equiv (A \downarrow B) \downarrow (A \downarrow B)$.

No completos, con el motivo:

- $\{\wedge, \vee\}$: toda fórmula construida con ellos es *monótona* —cambiar una variable de 0 a 1 nunca hace bajar el resultado—, y $\neg$ no lo es. De las 16 funciones binarias solo se alcanzan 4.
- $\{\rightarrow\}$ sola: toda fórmula que solo use $\rightarrow$ vale 1 cuando todas las variables valen 1, así que la constante 0 y la negación quedan fuera de alcance. Se alcanzan 6 de las 16.
- $\{\wedge, \vee, \rightarrow, \leftrightarrow\}$, es decir, todo menos la negación: las cuatro preservan el 1, luego ninguna combinación suya puede valer 0 con todas las variables a 1. Se alcanzan 8 de las 16.
- $\{\neg, \leftrightarrow\}$ y $\{\neg, \oplus\}$: todas las funciones obtenidas son *afines*, de la forma $c_0 \oplus c_1 p_1 \oplus \cdots \oplus c_n p_n$ sobre $\mathbb{F}_2$, y la conjunción no lo es. Se alcanzan 8 de las 16.
- $\{\neg\}$ sola, o $\{\wedge\}$ sola: obvio por aridad, no se puede combinar información de dos variables en el primer caso ni negar en el segundo.

Observación 2.53 (El criterio general, y por qué esto se construye de verdad). Los cuatro motivos de la lista anterior no son ad hoc: son cuatro de las cinco *clases maximales* del teorema de Post, de 1941, que da el criterio definitivo. Un conjunto de conectivos es completo si y solo si, para cada una de las cinco propiedades siguientes, contiene al menos un conectivo que *no* la tiene: preservar el 0, preservar el 1, ser monótono, ser afín y ser autodual. Comprobar la completitud de un conjunto se reduce entonces a rellenar una tabla de cinco columnas. NAND, por ejemplo, falla las cinco a la vez, y por eso basta él solo.

Y esto tiene consecuencias físicas, no solo lógicas. Que NAND sea completo significa que cualquier circuito digital —un sumador, una memoria, un procesador entero— se puede fabricar con un único tipo de puerta. La industria lo aprovecha literalmente: las familias lógicas se diseñan en torno a NAND y NOR porque en tecnología CMOS son las puertas más baratas en número de transistores, y toda la complejidad se obtiene replicando una sola pieza. La completitud funcional es la razón por la que eso es posible.

2.10. Ejercicios resueltos

Tres problemas que obligan a distinguir lo que un sistema formal *dice* de lo que un sistema formal *demuestra*. Es la distinción entera del capítulo.

Ejercicio 2.1. Demuestre que el conjunto de fórmulas de primer orden en un lenguaje numerable es numerable, y deduzca que hay estructuras que ninguna teoría de primer orden caracteriza salvo isomorfismo.

Solución. Una fórmula es una sucesión finita de símbolos tomados de un alfabeto numerable. El conjunto de sucesiones finitas sobre un alfabeto numerable es unión numerable de conjuntos numerables —uno por cada longitud— y por tanto numerable. Como las fórmulas son un subconjunto de esas sucesiones, también lo son.

De ahí sale la consecuencia. Por Löwenheim–Skolem, toda teoría de primer orden con un modelo infinito tiene modelos de todo cardinal infinito, luego ninguna teoría de primer orden puede fijar el cardinal de sus modelos. La teoría de los reales tiene modelos numerables, y la de conjuntos también: es la paradoja de Skolem, que no es paradoja sino el aviso de que "no numerable" es una afirmación relativa al modelo. ■

Ejercicio 2.2. Explique por qué el segundo teorema de incompletitud no queda refutado por el hecho de que la consistencia de la aritmética de Peano sí se demuestre en la teoría de conjuntos.

Solución. El segundo teorema dice que PA no demuestra Con(PA), no que nadie pueda demostrarlo. Gentzen probó la consistencia de PA en 1936 usando inducción transfinita hasta ε_0, y ZFC también la demuestra. Ninguna de las dos cosas contradice a Gödel, porque ambas usan recursos que PA no tiene.

Lo que sí se sigue es lo incómodo: ZFC tampoco demuestra su propia consistencia, y para demostrarla haría falta una teoría más fuerte, cuya consistencia a su vez exigiría otra. No hay suelo. La pregunta "¿son consistentes las matemáticas?" no tiene respuesta dentro de las matemáticas, y ésa es exactamente la razón de que el programa de Hilbert, tal como fue formulado, sea irrealizable. ■

Ejercicio 2.3. Sea T una teoría consistente, recursivamente axiomatizable y completa. Demuestre que T es decidible, y explique por qué esto, junto con Gödel, obliga a que la aritmética sea incompleta.

Solución. Enumere sistemáticamente todas las demostraciones de T, cosa posible porque los axiomas son reconocibles por un algoritmo. Dada una sentencia σ, vaya generando demostraciones y compare cada conclusión con σ y con $\neg\sigma$. Por completitud, una de las dos aparecerá tarde o temprano; por consistencia, no aparecerán ambas. El procedimiento siempre termina y responde correctamente, luego T es decidible.

Ahora júntelo con Church: la aritmética de Peano es indecidible. Si fuese completa sería decidible por el argumento anterior, luego no es completa. Esta demostración de la incompletitud no construye ninguna sentencia autorreferente y no se parece en nada a la de Gödel; llega al mismo sitio por computabilidad. Que dos caminos tan distintos coincidan es señal de que el fenómeno es robusto y no un artificio de la codificación. ■

2.11. Ejercicios propuestos

Sin solución desarrollada: solo enunciado y, en el apéndice, la respuesta.

2.1. ¿Cuántas fórmulas proposicionales no equivalentes hay con dos variables p y q?

2.2. Escriba la negación de $\forall x\,\exists y\,(x < y \wedge P(y))$ empujando el negador hasta los átomos.

2.3. ¿Es $\{\neg, \wedge\}$ un conjunto completo de conectivos? ¿Y $\{\Rightarrow\}$ solo?

2.4. Enuncie la diferencia entre $\Gamma \vdash \varphi$ y $\Gamma \models \varphi$, y diga qué teorema las relaciona.

2.5. ¿Qué afirma el teorema de compacidad? Úselo para demostrar que un grafo infinito cuyos subgrafos finitos son todos 3-coloreables es 3-coloreable.

2.6. ¿Puede una teoría de primer orden caracterizar a $(\mathbb{N}, +, \cdot, 0, 1)$ salvo isomorfismo?

2.7. Explique en qué sentido la paradoja de Skolem no es una contradicción.

2.8. ¿Cuál es el cardinal de $\mathcal{P}(\mathbb{N})$? ¿Y el de $\mathbb{R}$? ¿Qué afirma la hipótesis del continuo?

2.9. ¿Es demostrable la hipótesis del continuo en ZFC? ¿Y refutable? Cite a quién se debe cada mitad.

2.10. Dé un ejemplo de conjunto bien ordenado que no sea isomorfo a ningún ordinal finito ni a ω.

2.11. ¿Qué es un conjunto recursivamente enumerable pero no recursivo? Dé el ejemplo canónico.

2.12. Verdadero o falso: "Gödel demostró que hay verdades matemáticas que nunca podremos conocer". Justifique la respuesta con precisión.

Capítulo 3

Teoría de números: de la divisibilidad a la criptografía

Lo que necesitas traer

Del capítulo efcap:logica: inducción y demostración por reducción al absurdo. Nada de cálculo, nada de álgebra lineal: este capítulo se lee con lápiz y papel desde la primera línea. Es largo porque va de la división con resto hasta RSA sin saltarse un escalón, y es el que sostiene la teoría analítica de números del capítulo efcap:analitica, la teoría de grupos del capítulo efcap:grupos y los cuerpos finitos del capítulo efcap:anillos.

Una confesión antes de empezar. Los capítulos anteriores están escritos como un mapa para alguien que ya recorrió el territorio: densos, veloces, llenos de atajos y de guiños entre temas. Sirven para *recordar* y para *ver conexiones*, pero no siempre para *aprender por primera vez*. Este capítulo está escrito al revés. Vamos despacio, con números concretos antes que con letras, sin saltarnos ningún paso de las demostraciones, y explicando no solo *qué* es verdad sino *por qué* a alguien se le ocurrió preguntárselo. Casi no necesitas nada de los capítulos anteriores: solo saber sumar, restar y multiplicar enteros. La teoría de números es el lugar perfecto para esto, porque sus preguntas se entienden con la primaria y sus respuestas ocupan la vida entera de los matemáticos.

Mi apuesta es que al terminar este capítulo sientas algo que quizá los otros no te dieron: la sensación de haber *entendido*, no de haber presenciado. Si funciona, el resto del libro se puede releer con esta misma lentitud.

3.1. Divisibilidad y el algoritmo de Euclides

Idea clave

Intuición primero. La teoría de números estudia los enteros, y su rasgo característico es que en los enteros *no se puede dividir*. Esa carencia, que parece una limitación, es justamente lo que hace rica a la teoría: en vez de un cociente exacto hay un cociente y un resto, y todo lo demás sale de organizar bien esa idea.

El algoritmo de Euclides es probablemente el algoritmo más antiguo que se sigue usando tal cual, y su idea cabe en una línea: cualquier divisor común de a y b también divide al resto de dividir a entre b, así que se puede sustituir el par por otro más pequeño sin perder información sobre sus divisores comunes. Repite y los números se desploman hasta que uno divide al otro. Lo que lo convierte en teoría, y no en un truco de cálculo, es la identidad de Bézout: al recorrer el algoritmo hacia atrás se descubre que el máximo común divisor siempre se puede escribir como $ax + by$ con enteros x e y. Ese hecho, aparentemente modesto, es el que sostiene el lema de Euclides, la factorización única, los inversos modulares y, al final del capítulo, el RSA.

Empecemos por lo más elemental. Decimos que un entero a *divide* a otro entero b (y escribimos $a \mid b$) si b es un múltiplo exacto de a, es decir, si existe un entero k con $b = ak$. Así $3 \mid 12$ porque $12 = 3 \cdot 4$, pero $3 \nmid 13$ porque 13 no es múltiplo de 3.

Cuando a *no* divide a b, lo que hacemos en la primaria es dividir con resto: repartir lo más posible y anotar lo que sobra. Esto merece un nombre formal, porque es el cimiento de todo lo que sigue.

Teorema 3.1 (Algoritmo de la división)**.** *Dados enteros a y b con $b > 0$, existen enteros únicos q (el* cociente*) y r (el* resto*) tales que*

$$a = bq + r, \qquad 0 \leq r < b.$$

Demostración. La idea es tan simple que casi da pena escribirla, pero vale la pena hacerlo una vez. Mira todos los múltiplos de b:

$$\ldots,\ -2b,\ -b,\ 0,\ b,\ 2b,\ 3b,\ \ldots$$

Están espaciados uniformemente cada b pasos, así que a cae en algún hueco entre dos múltiplos consecutivos: hay un múltiplo bq que es el más grande que no pasa de a. Es decir, $bq \leq a$ pero $b(q+1) > a$. Define $r = a - bq$. De $bq \leq a$ sale $r \geq 0$; de $a < b(q+1) = bq + b$ sale $r < b$. Ya está la existencia.

Para la unicidad, supón que hubiera dos representaciones: $a = bq + r = bq' + r'$ con ambos restos en $[0, b)$. Restando, $b(q - q') = r' - r$. El lado izquierdo es un múltiplo de b; el lado derecho es la diferencia de dos números que viven en $[0, b)$, así que su valor absoluto es menor que b. El único múltiplo de b con valor absoluto menor que b es el cero. Luego $r = r'$ y, de ahí, $q = q'$. □

Pausa para pensar

Antes de seguir, asegúrate de que la frase "el único múltiplo de b menor que b en valor absoluto es el cero" te resulte obvia. Los múltiplos de b son $0, \pm b, \pm 2b, \ldots$; el más cercano al cero, aparte del cero mismo, ya está a distancia b. Este tipo de argumento —"es un múltiplo de b y a la vez es demasiado pequeño para serlo, salvo que sea cero"— va a reaparecer una y otra vez. Vale la pena tenerlo en el bolsillo.

Regla: dividir con resto, también cuando el dividendo es negativo

Para hallar q y r en $a = bq + r$ con $0 \leq r < b$:

1. Busca el mayor múltiplo de b que *no pase* de a. Ese múltiplo es bq.
2. El resto es lo que falta: $r = a - bq$.
3. Comprueba siempre que $0 \leq r < b$. Si te sale un resto negativo, has cogido un múltiplo demasiado grande; baja uno.

La palabra clave del paso 1 es *no pase*: con a negativo el múltiplo correcto queda por debajo de a, no por encima.

Ejemplo 3.2. Dividamos $a = 47$ entre $b = 7$. Los múltiplos de 7 son $\ldots, 42, 49, \ldots$; el mayor que no pasa de 47 es $42 = 7 \cdot 6$. Luego $q = 6$ y $r = 47 - 42 = 5$, y en efecto $0 \leq 5 < 7$:

$$47 = 7 \cdot 6 + 5.$$

Ahora $a = -47$, el mismo $b = 7$. Aquí la trampa: el mayor múltiplo de 7 que no pasa de -47 no es -42 (que es *mayor* que -47), sino $-49 = 7 \cdot (-7)$. Luego $q = -7$ y $r = -47 - (-49) = 2$:

$$-47 = 7 \cdot (-7) + 2, \qquad 0 \leq 2 < 7.$$

Quien escriba $-47 = 7 \cdot (-6) - 5$ ha hecho una cuenta correcta pero *no* es la división euclídea, porque -5 no está en $[0, 7)$. El teorema afirma que la pareja buena existe y es única, y aquí es $(-7, 2)$.

3.1.1. El máximo común divisor

Dados dos enteros, digamos 48 y 18, sus divisores comunes son los números que dividen a ambos. Los de 48 son $1, 2, 3, 4, 6, 8, 12, 16, 24, 48$; los de 18 son $1, 2, 3, 6, 9, 18$; los comunes son $1, 2, 3, 6$, y el mayor de ellos es 6. A ese número lo llamamos el *máximo común divisor*, $\gcd(48, 18) = 6$.

Listar todos los divisores funciona con números chicos, pero es desesperante con números grandes. ¿Cómo calcularía $\gcd(48, 18)$ alguien que no quiere factorizar nada? Aquí entra una de las ideas más antiguas y más bellas de las matemáticas, con más de dos mil años: el *algoritmo de Euclides*. Se apoya en una observación pequeñita.

Lema 3.3. *Si $a = bq + r$, entonces $\gcd(a, b) = \gcd(b, r)$.*

Demostración. Voy a convencerte de que *el par (a, b) y el par (b, r) tienen exactamente los mismos divisores comunes*; si tienen los mismos divisores comunes, tienen el mismo máximo. Sea d un divisor común de a y b. Como $r = a - bq$, y d divide a a y a b, también divide a cualquier combinación como $a - bq$: divide a r. Así d es divisor común de b y r. Recíprocamente, si d divide a b y a r, entonces divide a $bq + r = a$, así que es divisor común de a y b. Los dos conjuntos de divisores comunes coinciden. □

Este lema convierte un problema en uno *más pequeño*: para conocer $\gcd(a, b)$, basta conocer $\gcd(b, r)$, y r es menor que b. Repite, y los números encogen sin parar hasta que uno se vuelve cero, momento en que la respuesta salta a la vista (pues $\gcd(n, 0) = n$: todo divide al cero, así que el mayor divisor común de n y 0 es el mayor divisor de n, que es n).

Hagámoslo con 48 y 18, paso a paso, sin prisa:

$$\begin{aligned} 48 &= 2 \cdot 18 + 12 && \text{(dividimos 48 entre 18: cociente 2, resto 12)} \\ 18 &= 1 \cdot 12 + 6 && \text{(ahora dividimos 18 entre 12: resto 6)} \\ 12 &= 2 \cdot 6 + 0 && \text{(dividimos 12 entre 6: resto 0, paramos).} \end{aligned}$$

El último resto no nulo es 6: ese es el máximo común divisor. Fíjate en la elegancia: nunca factorizamos nada, solo dividimos tres veces. Con números de cientos de dígitos el algoritmo sigue siendo rapidísimo, y por eso vive dentro de tu teléfono cada vez que abre una conexión segura.

Regla: algoritmo de Euclides, en columna

1. Escribe el par (a, b) con $a \geq b > 0$.
2. Divide a entre b y quédate con el resto r.
3. Sustituye el par por (b, r) y vuelve al paso 2.
4. Cuando el resto sea 0, el *último resto no nulo* es $\gcd(a, b)$.

Nunca hay que factorizar, y el número de pasos es pequeñísimo: nunca más de unas cinco veces el número de cifras del menor de los dos.

Ejemplo 3.4. Calculemos $\gcd(1071, 462)$:

$$\begin{aligned} 1071 &= 2 \cdot 462 + 147 \\ 462 &= 3 \cdot 147 + 21 \\ 147 &= 7 \cdot 21 + 0. \end{aligned}$$

Último resto no nulo: $\gcd(1071, 462) = 21$. Comprobación por factorización, que aquí sí es fácil: $1071 = 3^2 \cdot 7 \cdot 17$ y $462 = 2 \cdot 3 \cdot 7 \cdot 11$; los factores comunes son 3 y 7, y $3 \cdot 7 = 21$ ✓.

Otro más, con un paso extra: $\gcd(252, 198)$.

$$\begin{aligned} 252 &= 1 \cdot 198 + 54 \\ 198 &= 3 \cdot 54 + 36 \\ 54 &= 1 \cdot 36 + 18 \\ 36 &= 2 \cdot 18 + 0, \end{aligned}$$

así que $\gcd(252, 198) = 18$. Obsérvese cómo los números caen: $252, 198, 54, 36, 18$. No hace falta ninguna intuición sobre qué primos intervienen.

3.1.2. La identidad de Bézout: el regalo escondido del algoritmo

Ahora viene lo verdaderamente útil, y quiero que lo descubras en vez de que te lo anuncie. Volvamos a subir por las cuentas de Euclides, pero al revés, despejando el resto en cada línea:

$$\begin{aligned} 6 &= 18 - 1 \cdot 12 && \text{(de la segunda línea)} \\ &= 18 - 1 \cdot (48 - 2 \cdot 18) && \text{(sustituimos } 12 = 48 - 2 \cdot 18\text{, de la primera línea)} \\ &= 3 \cdot 18 - 1 \cdot 48. \end{aligned}$$

¡Mira lo que acaba de pasar! Escribimos el máximo común divisor 6 como una combinación $3 \cdot 18 + (-1) \cdot 48$ de los números originales. Esto no fue suerte; siempre se puede hacer, deshaciendo el algoritmo. Enunciémoslo.

Teorema 3.5 (Identidad de Bézout)**.** *Para cualesquiera enteros a, b (no ambos cero), existen enteros x, y tales que*

$$ax + by = \gcd(a, b).$$

Demostración. Aquí hay una demostración limpia que además explica *por qué* el máximo común divisor aparece y no otro número. Considera el conjunto de todos los valores positivos que puede tomar la expresión $ax + by$ cuando x, y recorren los enteros:

$$S = \{\, ax + by : x, y \in \mathbb{Z},\ ax + by > 0 \,\}.$$

Este conjunto no es vacío (tomando x, y adecuados podemos hacer positiva la combinación), y todo conjunto no vacío de enteros positivos tiene un elemento mínimo. Llamemos d a ese mínimo, con $d = ax_0 + by_0$. Afirmo que $d = \gcd(a, b)$, y lo probaré en dos mitades.

Primero, d divide a a y a b. Dividamos a entre d: $a = dq + r$ con $0 \le r < d$. Entonces

$$r = a - dq = a - q(ax_0 + by_0) = a(1 - qx_0) + b(-qy_0),$$

que también es de la forma $ax + by$. Si r fuera positivo, sería un elemento de S menor que d, contradiciendo que d es el mínimo. Así que $r = 0$, es decir, $d \mid a$. Lo mismo con b. Entonces d es un divisor común.

Segundo, d es el mayor. Sea c cualquier divisor común de a y b. Como $d = ax_0 + by_0$ y c divide a a y a b, también divide a esa combinación: $c \mid d$. Y un divisor de d (positivo) no puede ser mayor que d. Luego d es el máximo divisor común. □

Idea clave

Qué dice esto en cristiano. Bézout dice que el máximo común divisor de dos números siempre se puede escribir combinándolos con coeficientes enteros. Es un resultado de existencia con un algoritmo detrás: el de Euclides extendido, que va guardando los coeficientes mientras calcula el máximo común divisor. De él sale todo lo demás en este capítulo: el lema de Euclides, la factorización única y la manera de calcular inversos módulo m, que es lo que hace funcionar la criptografía RSA.

Ejemplo 3.6 (Euclides extendido, paso a paso)**. (a) El caso base.** Para $a = 1071$ y $b = 462$: $1071 = 2 \cdot 462 + 147$, luego $462 = 3 \cdot 147 + 21$, luego $147 = 7 \cdot 21 + 0$. El máximo común divisor es 21.

(b) Subir los coeficientes. Deshaciendo el camino: $21 = 462 - 3 \cdot 147$, y como $147 = 1071 - 2 \cdot 462$, queda $21 = 462 - 3(1071 - 2 \cdot 462) = 7 \cdot 462 - 3 \cdot 1071$. Comprobación: $7 \cdot 462 = 3234$ y $3 \cdot 1071 = 3213$, cuya diferencia es 21.

(c) Para qué sirve: inversos modulares. Para invertir 7 módulo 26, se busca $7x + 26y = 1$. Euclides da $7 \cdot 15 - 26 \cdot 4 = 105 - 104 = 1$, luego $7^{-1} \equiv 15 \pmod{26}$. Este cálculo es exactamente el que descifra un cifrado afín y el que genera la clave privada de RSA.

(d) La consecuencia teórica. Si $\gcd(p, a) = 1$ con p primo, existen x, y con $px + ay = 1$; multiplicando por b queda $pbx + aby = b$, y si $p \mid ab$ entonces p divide a los dos sumandos, luego a b. Ese es el lema de Euclides, y de él sale la unicidad de la factorización en primos. Toda la aritmética elemental cuelga de esta identidad.

Cuidado

[Bézout da existencia, no unicidad, y solo funciona con el máximo común divisor] Los coeficientes no son únicos: a $7 \cdot 462 - 3 \cdot 1071 = 21$ se le puede sumar cualquier múltiplo de $\frac{1071}{21} = 51$ al primero restando el correspondiente del segundo. Segundo aviso: la ecuación $ax + by = c$ tiene solución entera si y solo si $\gcd(a, b)$ divide a c; para c menor que el máximo común divisor no hay ninguna, y $6x + 4y = 3$ es insoluble porque $2 \nmid 3$. Y tercero: la factorización única no es un hecho universal de todos los anillos, como muestra $6 = 2 \cdot 3 = (1 + \sqrt{-5})(1 - \sqrt{-5})$ en $\mathbb{Z}[\sqrt{-5}]$; en $\mathbb{Z}$ es cierta precisamente porque hay división con resto y por tanto Bézout.

Pausa para pensar

Esta demostración es más sutil que la del algoritmo, así que respira. La idea central es: *el máximo común divisor es el número positivo más pequeño que puedes fabricar combinando a y b con coeficientes enteros.* Esa caracterización —"el mínimo positivo alcanzable"— es sorprendente la primera vez. El truco de dividir a entre d y observar que el resto también es alcanzable (y por tanto debe ser cero) es el corazón del asunto; es el mismo tipo de argumento del algoritmo de la división, reutilizado.

Bézout parece un tecnicismo, pero es la llave que abre casi todas las puertas de este capítulo. Solo un ejemplo de su poder inmediato: si $\gcd(a, b) = 1$ (decimos que a y b son *coprimos*, no comparten factores), entonces existen x, y con $ax + by = 1$. Vamos a usar esto sin parar.

Regla: algoritmo de Euclides extendido (hallar x e y de Bézout)

1. Haz el algoritmo de Euclides normal y *guarda todas las líneas.*
2. Despeja el resto en cada línea: $r = a - bq$.
3. Empieza por la última línea con resto no nulo y ve sustituyendo hacia arriba, línea por línea.

4. En cada sustitución *no multipliques todavía*: agrupa dejando los dos números originales a la vista. Ese es el único paso donde se pierde la gente.
5. Al final queda $\gcd(a, b) = ax + by$. Comprueba la igualdad numéricamente antes de seguir.

Ejemplo 3.7. Retomemos $\gcd(1071, 462) = 21$ y busquemos los coeficientes. Las líneas del algoritmo, con el resto despejado, eran

$$\begin{aligned} 147 &= 1071 - 2 \cdot 462, \\ 21 &= 462 - 3 \cdot 147. \end{aligned}$$

Sustituimos 147 en la segunda usando la primera, sin multiplicar aún:

$$\begin{aligned} 21 &= 462 - 3\,(1071 - 2 \cdot 462) \\ &= 462 - 3 \cdot 1071 + 6 \cdot 462 \\ &= 7 \cdot 462 - 3 \cdot 1071. \end{aligned}$$

Es decir, $1071 \cdot (-3) + 462 \cdot 7 = 21$. Comprobación: $-3213 + 3234 = 21$ ✓.

Ejemplo 3.8. Y con 252 y 198, cuyo máximo común divisor era 18. Restos despejados:

$$54 = 252 - 198, \qquad 36 = 198 - 3 \cdot 54, \qquad 18 = 54 - 36.$$

Hacia arriba:

$$\begin{aligned} 18 &= 54 - 36 = 54 - (198 - 3 \cdot 54) = 4 \cdot 54 - 198 \\ &= 4\,(252 - 198) - 198 = 4 \cdot 252 - 5 \cdot 198. \end{aligned}$$

Comprobación: $1008 - 990 = 18$ ✓. Los coeficientes son $x = 4$, $y = -5$; obsérvese que uno de los dos siempre sale negativo (salvo casos triviales), porque $ax + by$ tiene que *encogerse* hasta el máximo común divisor.

3.2. Cimientos: la divisibilidad, las congruencias y las clases de restos, definidas desde cero

Idea clave

Por qué esta sección existe. La sección anterior ha contado una historia: dividir con resto, encoger los pares con Euclides, descubrir Bézout. Contar la historia es imprescindible, pero no basta. Un capítulo que va a llegar hasta el RSA necesita que cada palabra que use tenga un significado fijado por escrito, y que cada propiedad que emplee esté demostrada en alguna parte y no solo dada por evidente.
Eso es lo que hacemos aquí. No vamos a repetir las explicaciones de la sección anterior: vamos a *fijar* las definiciones (qué significa exactamente $a \mid b$, qué es un primo, qué

es el máximo común divisor, qué es una congruencia, qué es $\mathbb{Z}/n\mathbb{Z}$), a demostrar las propiedades pequeñas que luego se usarán mil veces sin avisar, y a dejar por escrito las tres o cuatro advertencias que separan a quien maneja esto con soltura de quien se equivoca siempre en el mismo sitio. Es la sección menos vistosa del capítulo y la que sostiene todas las demás; si en algún momento más adelante una demostración da un salto que no ves, es muy probable que el peldaño que falta esté aquí.

3.2.1. La relación de divisibilidad

Empecemos por el ladrillo. En los enteros no siempre se puede dividir, y la pregunta "¿cabe exactamente?" es la pregunta central de todo el capítulo. Merece una definición sin ambigüedad.

Definición 3.9 (Divisibilidad)**.** Sean $a, b \in \mathbb{Z}$. Decimos que a *divide* a b, y escribimos

$$a \mid b,$$

si existe algún entero k tal que $b = ak$. En ese caso también se dice que a es un *divisor* o un *factor* de b, y que b es un *múltiplo* de a. Cuando no existe tal k escribimos $a \nmid b$.

Obsérvese lo que la definición *no* dice: no habla de fracciones, ni de restos, ni de calculadoras. Solo afirma que existe un entero que multiplicado por a da b. Toda la teoría se construye sobre esa frase.

Cuidado

[Tres confusiones que hay que cortar de raíz] **Primera: $a \mid b$ no es a/b.** El símbolo $a \mid b$ es una *afirmación*, algo que es verdadero o falso; a/b es un *número*. Escribir "$3 \mid 12 = 4$" es como escribir "$2 < 5 = 3$": una mezcla sin sentido. Lo correcto es "$3 \mid 12$, y el cociente es 4". Y ojo al orden, que es el error más frecuente de todos: en $a \mid b$ el pequeño va delante. Se lee "a divide a b", nunca "a dividido entre b".
Segunda: el 0 se deja dividir por todos. Como $0 = a \cdot 0$ para cualquier a, se tiene $a \mid 0$ *siempre*, incluido $0 \mid 0$. Lo que no ocurre nunca es lo contrario: si $0 \mid b$ entonces $b = 0 \cdot k = 0$, de modo que el único múltiplo de 0 es el propio 0. Esto no contradice la prohibición de dividir entre cero, porque la definición 3.9 no divide nada: solo multiplica.
Tercera: los signos. La divisibilidad ignora los signos: $3 \mid -12$ (pues $-12 = 3 \cdot (-4)$) y $-3 \mid 12$. Por eso, cuando hablemos del "mayor" divisor común, habrá que decir explícitamente que lo buscamos entre los positivos.

Las propiedades que siguen son minúsculas y se usarán constantemente, casi siempre sin mencionarlas. La segunda de la lista, la de las combinaciones, es probablemente el enunciado más utilizado de todo el capítulo: aparece en el algoritmo de Euclides, en la identidad de Bézout, en el lema de Euclides, en los criterios de divisibilidad y en la demostración de la infinitud de los primos.

Proposición 3.10 (Reglas básicas de la divisibilidad)**.** *Sean a, b, c, d, x, y enteros. Entonces:*

1. (Reflexividad) $a \mid a$, *y también* $1 \mid a$ *y* $a \mid 0$.
2. (Transitividad) *Si* $d \mid a$ *y* $a \mid b$, *entonces* $d \mid b$.
3. (Combinaciones) *Si* $d \mid a$ *y* $d \mid b$, *entonces* $d \mid ax + by$ *para cualesquiera enteros* x, y. *En particular* $d \mid a + b$ *y* $d \mid a - b$.
4. (Producto) *Si* $d \mid a$, *entonces* $d \mid ab$ *y* $dc \mid ac$.
5. (Tamaño) *Si* $d \mid a$ *y* $a \neq 0$, *entonces* $|d| \leq |a|$.
6. (Antisimetría) *Si* $a \mid b$ *y* $b \mid a$, *entonces* $a = \pm b$; *si además ambos son positivos,* $a = b$.

Demostración. Todas salen de escribir la definición y no hacer nada más. (1) $a = a \cdot 1$, $a = 1 \cdot a$ y $0 = a \cdot 0$.

(2) Por hipótesis $a = dk$ y $b = a\ell$ para ciertos enteros k, ℓ. Sustituyendo, $b = (dk)\ell = d(k\ell)$, y $k\ell$ es entero: luego $d \mid b$. (3) Por hipótesis $a = dk$ y $b = d\ell$. Entonces

$$ax + by = dkx + d\ell y = d\,(kx + \ell y),$$

y $kx + \ell y$ es un entero. Luego $d \mid ax + by$. Los casos $a + b$ y $a - b$ se obtienen con $(x, y) = (1, 1)$ y $(x, y) = (1, -1)$.

(4) Si $a = dk$, entonces $ab = d(kb)$ y $ac = (dc)k$.

(5) Si $a = dk$ con $a \neq 0$, entonces $k \neq 0$, luego $|k| \geq 1$ y $|a| = |d|\,|k| \geq |d|$. (6) Si a y b son ambos cero, no hay nada que probar; si no, por (5) aplicado en los dos sentidos tenemos $|a| \leq |b|$ y $|b| \leq |a|$, luego $|a| = |b|$, es decir $a = \pm b$. □

Ejemplo 3.11 (Las reglas en funcionamiento)**. (a) Transitividad.** Como $3 \mid 21$ (pues $21 = 3 \cdot 7$) y $21 \mid 861$ (pues $861 = 21 \cdot 41$), la propiedad (2) garantiza $3 \mid 861$ sin necesidad de dividir: en efecto $861 = 3 \cdot 287$.

(b) Combinaciones, el caballo de batalla. Sabemos que $7 \mid 91$ y $7 \mid 63$. Entonces 7 divide a *cualquier* combinación entera suya. Por ejemplo, con $x = 4$ e $y = -5$:

$$91 \cdot 4 + 63 \cdot (-5) = 364 - 315 = 49 = 7 \cdot 7 \quad \checkmark$$

No hemos comprobado la divisibilidad de 49 por separado: la proposición nos la regaló antes de hacer la cuenta. Este es exactamente el gesto que se repetirá en el algoritmo de Euclides, donde el resto $r = a - bq$ es una combinación de a y b y por eso hereda todos sus divisores comunes.

(c) Tamaño. Los divisores positivos de 20 están todos entre 1 y 20: son $1, 2, 4, 5, 10, 20$. La propiedad (5) es lo que garantiza que la búsqueda es finita, y por tanto lo que permite hablar del "mayor" divisor común sin que la frase quede en el aire.

(d) Un uso negativo. ¿Puede 6 dividir a $91 \cdot 4 + 63 \cdot (-5) = 49$? Si 6 dividiera a 49, por (5) habría un entero k con $49 = 6k$, imposible porque $6 \cdot 8 = 48$ y $6 \cdot 9 = 54$. Así que $6 \nmid 49$. Fíjate en la forma del argumento: para negar una divisibilidad basta encajonar el número entre dos múltiplos consecutivos.

3.2.2. Parte entera y división euclídea

La sección anterior enunció el teorema de la división y lo justificó mirando la recta de los múltiplos de b: la idea era que a cae en algún hueco entre dos múltiplos consecutivos. Esa imagen es correcta y es la que hay que retener, pero esconde un supuesto que conviene sacar a la luz, porque es el mismo que sostendrá la demostración de Bézout: *que ese hueco existe*, es decir, que entre los múltiplos de b que no pasan de a hay uno máximo. Vamos a demostrarlo con la herramienta adecuada, y de paso a fijar dos notaciones que el capítulo usa sin haberlas presentado.

Definición 3.12 (Parte entera)**.** Para un número real x, su *parte entera* o *suelo* $\lfloor x \rfloor$ es el mayor entero que no supera a x:

$$\lfloor x \rfloor = \max\{\, n \in \mathbb{Z} : n \le x \,\},$$

lo que equivale a decir que $\lfloor x \rfloor$ es el único entero n con $n \le x < n+1$.

Con números positivos coincide con "quitar los decimales": $\lfloor 3{,}7 \rfloor = 3$. Con negativos *no*, y ahí está la trampa: $\lfloor -3{,}7 \rfloor = -4$, porque -4 es el mayor entero que no supera a $-3{,}7$; el -3 se pasa. Retén ese detalle, porque es exactamente el que hace que la división euclídea de un número negativo salga como sale.

El principio que vamos a usar es el *principio de buena ordenación*: todo conjunto no vacío de enteros no negativos tiene un elemento mínimo. Es una propiedad de $\mathbb{N}$ tan básica que suele darse por descontada, y sin embargo es equivalente al principio de inducción y es la que convierte las ideas geométricas de "hueco" en demostraciones.

Teorema 3.13 (División euclídea)**.** *Sean $a, b \in \mathbb{Z}$ con $b > 0$. Existe* exactamente un *par de enteros (q, r) tal que*

$$a = bq + r \qquad y \qquad 0 \le r < b.$$

A q se le llama cociente *y a r,* resto*. Además $q = \lfloor a/b \rfloor$, y se escribe $r = a \bmod b$.*

Demostración. Existencia. Consideremos el conjunto de todos los "restos candidatos" no negativos:

$$S = \{\, a - bt \;:\; t \in \mathbb{Z},\ a - bt \ge 0 \,\}.$$

Primero, S no es vacío. Si $a \ge 0$ basta tomar $t = 0$, que da $a \in S$. Si $a < 0$, tomemos $t = a$: como $b \ge 1$, se tiene $a - ba = a(1-b) \ge 0$, porque a es negativo o cero y $1 - b$ es negativo o cero, y el producto de dos números de signo ≤ 0 es ≥ 0. Así que S contiene algún elemento.

Al ser S un conjunto no vacío de enteros no negativos, el principio de buena ordenación nos da un elemento mínimo; llamémoslo r, alcanzado con cierto $t = q$, de modo que $r = a - bq \ge 0$, es decir, $a = bq + r$.

Falta ver que $r < b$. Supongamos que no, es decir, $r \ge b$. Entonces

$$r - b = a - bq - b = a - b(q+1) \ge 0,$$

así que $r - b$ también pertenece a S; pero $r - b < r$ porque $b > 0$, y eso contradice que r fuera el mínimo de S. Luego $r < b$, como queríamos.

Unicidad. Supongamos dos representaciones válidas,

$$a = bq + r = bq' + r', \qquad 0 \le r < b, \quad 0 \le r' < b.$$

Restando, $b(q - q') = r' - r$. El lado izquierdo es un múltiplo de b. El derecho es la diferencia de dos números del intervalo $[0, b)$, luego $|r' - r| < b$. Ahora bien, el único múltiplo de b de valor absoluto menor que b es el cero: si $b(q - q') \neq 0$ entonces $|q - q'| \ge 1$ y por tanto $|b(q - q')| \ge b$, contradicción. Así que $r' = r$, y de $b(q - q') = 0$ con $b > 0$ sale $q = q'$.

La fórmula del cociente. De $a = bq + r$ con $0 \le r < b$ se sigue, dividiendo entre $b > 0$,

$$q \le \frac{a}{b} = q + \frac{r}{b} < q + 1,$$

y por la definición 3.12 eso significa exactamente $q = \lfloor a/b \rfloor$. □

Observación 3.14 (El convenio del resto no negativo)*.* El teorema no dice que haya *una* manera de escribir $a = bq + r$: hay infinitas, una por cada valor de q. Lo que dice es que hay exactamente una con el resto en el rango $[0, b)$, y ese es el convenio que adoptamos siempre en este libro. La consecuencia práctica es que **el resto nunca es negativo, tampoco cuando el dividendo lo es**. Muchos lenguajes de programación no siguen este convenio y devuelven restos negativos para dividendos negativos; conviene saberlo antes de confiar en una máquina.

Ejemplo 3.15 (División euclídea con dividendo negativo)**. (a) El caso fácil.** Dividamos 100 entre 7. Como $\lfloor 100/7 \rfloor = \lfloor 14{,}28\ldots \rfloor = 14$, el cociente es $q = 14$ y el resto $r = 100 - 7 \cdot 14 = 100 - 98 = 2$:

$$100 = 7 \cdot 14 + 2, \qquad 0 \le 2 < 7 \quad \checkmark$$

(b) El caso que hay que aprender: $a = -17$, $b = 7$. La tentación es escribir $-17 = 7 \cdot (-2) - 3$, y la cuenta es cierta, pero *no es la división euclídea*, porque -3 no está en $[0, 7)$. Hagámoslo bien. El cociente es

$$q = \left\lfloor \frac{-17}{7} \right\rfloor = \lfloor -2{,}428\ldots \rfloor = -3,$$

donde el suelo baja a -3 y no sube a -2. Entonces

$$r = -17 - 7 \cdot (-3) = -17 + 21 = 4,$$

y en efecto

$$-17 = (-3) \cdot 7 + 4, \qquad 0 \le 4 < 7 \quad \checkmark$$

El par correcto es $(q, r) = (-3, 4)$, y por el teorema es el único con resto en el rango.

(c) Otro más, para fijarlo. Con $a = -100$ y $b = 7$: $\lfloor -100/7 \rfloor = \lfloor -14{,}28\ldots \rfloor = -15$, luego $r = -100 - 7 \cdot (-15) = -100 + 105 = 5$ y

$$-100 = 7 \cdot (-15) + 5 \quad \checkmark$$

Regla mnemotécnica: con dividendo negativo, el cociente "se pasa de largo hacia abajo" y el resto sube. Si te sale un resto negativo, resta uno al cociente y súmale b al resto.

(d) La notación. Con lo anterior, -17 mód $7 = 4$ y -100 mód $7 = 5$. No son -3 ni -2.

3.2.3. Primos y compuestos

Definición 3.16 (Número primo y número compuesto)**.** Un entero $p > 1$ es *primo* si sus únicos divisores positivos son 1 y p. Un entero $n > 1$ que no es primo se llama *compuesto*; equivalentemente, n es compuesto si se puede escribir $n = ab$ con $1 < a < n$ y $1 < b < n$. Los enteros 0 y 1 no son ni primos ni compuestos.

Cuidado

[Por qué el 1 no es primo] No es un capricho ni una convención arbitraria heredada: es una decisión forzada por el teorema fundamental de la aritmética. Si admitiéramos que 1 es primo, la factorización dejaría de ser única de inmediato, porque

$$12 = 2^2 \cdot 3 = 1 \cdot 2^2 \cdot 3 = 1^2 \cdot 2^2 \cdot 3 = 1^{57} \cdot 2^2 \cdot 3 = \cdots$$

serían infinitas factorizaciones distintas del mismo número. Excluir el 1 es lo que permite enunciar la unicidad sin coletillas. La razón profunda, que se ve en el capítulo de anillos, es que el 1 pertenece a otra categoría de objetos: es una *unidad*, un elemento invertible, y las unidades nunca cuentan como factores. Segundo aviso: tampoco los números negativos se clasifican como primos en este capítulo; hablamos siempre de enteros mayores que 1.

Antes de decir nada sobre los primos hay que asegurarse de que un entero cualquiera tiene alguno. Parece obvio y lo es, pero la demostración usa buena ordenación otra vez, y el enunciado se utilizará sin avisar en cuanto empiece la factorización.

Lema 3.17 (Todo entero mayor que uno tiene un divisor primo)**.** *Si $n > 1$ es un entero, existe un primo p con $p \mid n$.*

Demostración. Sea $D = \{\, d \in \mathbb{Z} : d > 1 \text{ y } d \mid n \,\}$, el conjunto de divisores de n mayores que 1. No es vacío, porque $n \in D$ (todo número se divide a sí mismo y $n > 1$). Por buena ordenación, D tiene un elemento mínimo; llamémoslo p.

Afirmo que p es primo. Si no lo fuera, por la definición 3.16 tendría un divisor c con $1 < c < p$. Pero entonces $c \mid p$ y $p \mid n$, luego por transitividad (proposición 3.10(2)) se tendría $c \mid n$; y como $c > 1$, resultaría $c \in D$ con $c < p$, contradiciendo la minimalidad de p. Por lo tanto p es primo y divide a n. □

Teorema 3.18 (Euclides)**.** *Hay infinitos números primos.*

Demostración. Sea $p_1, \ldots, p_n$ cualquier lista finita de primos, y consideremos $N = p_1 p_2 \cdots p_n + 1$. Como $N > 1$, el lema 3.17 garantiza que N tiene algún divisor primo p. Si p estuviera en la lista, dividiría al producto $p_1 \cdots p_n$ y también a N, luego por la propiedad de las combinaciones (proposición 3.10(3)) dividiría a la diferencia $N - p_1 \cdots p_n = 1$, lo cual es imposible para un primo. Así que p es un primo fuera de la lista. Ninguna lista finita los contiene a todos. □

La sección siguiente vuelve sobre esta demostración con calma, con ejemplos numéricos y con la advertencia importante sobre lo que el argumento *no* afirma; aquí solo la hemos

registrado en su forma mínima, apoyada en el lema que le faltaba.

3.2.4. Máximo común divisor y mínimo común múltiplo

Aquí hay que tener cuidado con las palabras. La sección anterior presentó el máximo común divisor como "el mayor de los divisores comunes", que es la descripción intuitiva correcta y la que hay que tener en la cabeza. Pero como *definición* de trabajo es floja: no dice nada sobre el caso $\gcd(0,0)$, y sobre todo no sirve tal cual en contextos donde no hay un orden con el que comparar. La definición que se usa de verdad, y que además hace que todas las demostraciones salgan solas, es esta.

Definición 3.19 (Máximo común divisor)**.** Sean a, b enteros no ambos nulos. El *máximo común divisor* de a y b, denotado $\gcd(a,b)$, es el único entero d que cumple las dos condiciones siguientes:

1. $d > 0$, $d \mid a$ y $d \mid b$ (*es divisor común*);
2. si $c \mid a$ y $c \mid b$, entonces $c \mid d$ (*y todo divisor común lo divide*).

Cuando $\gcd(a,b) = 1$ se dice que a y b son *coprimos* o *primos entre sí*: no comparten ningún factor mayor que 1. Por convenio se toma $\gcd(0,0) = 0$.

Observación 3.20 (Las dos definiciones coinciden, y por qué preferimos esta). La condición (2) dice más que "d es el mayor": dice que los divisores comunes no solo son menores que d, sino que *lo dividen*. Las dos versiones son equivalentes para enteros: si d cumple (1) y (2), entonces cualquier divisor común c cumple $c \mid d$ y por tanto $|c| \le d$ por la proposición 3.10(5), de modo que d es efectivamente el mayor; y recíprocamente, la identidad de Bézout demostrará que el mayor divisor común cumple (2).

La ventaja de la caracterización por divisibilidad es doble. Primero, es la propiedad que de hecho se usa en las demostraciones: cuando queramos probar que cierto número *es* el máximo común divisor, comprobar (1) y (2) es mecánico, mientras que "comprobar que es el mayor" obliga a razonar sobre todos los demás. Y segundo, es la versión que sobrevive al pasar a otros anillos, en los que no hay ningún "mayor" que elegir; el lector que llegue al capítulo de anillos reconocerá aquí la definición general de máximo común divisor.

Regla: el algoritmo de Euclides, paso a paso

Para calcular $\gcd(a,b)$ con $a \ge b \ge 0$ y $a > 0$:

1. Si $b = 0$, la respuesta es a y se termina. (En efecto, $\gcd(a,0) = a$: todo entero divide a 0, así que los divisores comunes de a y 0 son los divisores de a.)
2. Si $b > 0$, haz la división euclídea $a = bq + r$ con $0 \le r < b$.
3. Sustituye el par (a,b) por (b,r) y vuelve al paso 1.
4. El proceso termina siempre, porque los restos $b > r > r' > \cdots \ge 0$ forman una sucesión estrictamente decreciente de enteros no negativos y no puede ser infinita.
5. El resultado es el *último resto no nulo*, que es $\gcd(a,b)$.

La razón de que funcione es el lema $\gcd(a,b) = \gcd(b,r)$, ya demostrado en la sección anterior: a y b por un lado, y b y r por otro, tienen exactamente los mismos divisores comunes, porque $r = a - bq$ y $a = bq + r$ son combinaciones enteras la una de la otra (proposición 3.10(3)).

El algoritmo, además de calcular $\gcd(a,b)$, produce los coeficientes de Bézout si se lleva la contabilidad adecuada. La sección anterior lo hizo "subiendo" por las líneas y sustituyendo hacia atrás, que es el método correcto para entender lo que pasa. Aquí damos la versión en tabla, que es la que conviene usar en cuanto haya más de tres divisiones: baja en lugar de subir, no hay que guardar nada y el riesgo de equivocarse es mucho menor.

Regla: Euclides extendido en tabla

Se construye una tabla con cuatro columnas: q, r, s, t, de modo que en toda fila se cumpla la invariante

$$r = s\,a + t\,b.$$

1. **Dos filas de arranque,** sin cociente: la primera es $r = a$, $s = 1$, $t = 0$ (pues $a = 1 \cdot a + 0 \cdot b$); la segunda es $r = b$, $s = 0$, $t = 1$.
2. **Cociente.** En cada fila nueva, q es el cociente de dividir el resto *dos filas arriba* entre el resto de la fila inmediatamente anterior.
3. **Nueva fila.** Cada entrada se calcula igual, restando q veces la fila anterior a la fila de dos más arriba:

$$r_{\text{nueva}} = r_{-2} - q\,r_{-1}, \quad s_{\text{nueva}} = s_{-2} - q\,s_{-1}, \quad t_{\text{nueva}} = t_{-2} - q\,t_{-1}.$$

4. **Parada.** Se para cuando aparece $r = 0$. La *fila anterior* contiene $r = \gcd(a,b)$ y sus coeficientes s, t con $\gcd(a,b) = sa + tb$.
5. **Comprobación obligatoria.** Sustituye s y t y verifica numéricamente la igualdad. Cuesta diez segundos y detecta cualquier error de signo, que es el error que siempre se comete.

Ejemplo 3.21 (La tabla completa para $\gcd(2024, 748)$)**.** Tomemos $a = 2024$ y $b = 748$. Construimos la tabla fila a fila.

q	r	s	t
—	2024	1	0
—	748	0	1
2	528	1	−2
1	220	−1	3
2	88	3	−8
2	44	−7	19
2	0	17	−46

Veamos de dónde sale cada fila. El primer cociente es el de 2024 entre 748: como $2 \cdot 748 = 1496$ y $3 \cdot 748 = 2244 > 2024$, sale $q = 2$ y $r = 2024 - 1496 = 528$; los coeficientes son $s = 1 - 2 \cdot 0 = 1$ y $t = 0 - 2 \cdot 1 = -2$. La segunda: $748 = 1 \cdot 528 + 220$, con $s = 0 - 1 \cdot 1 = -1$ y $t = 1 - 1 \cdot (-2) = 3$. La tercera: $528 = 2 \cdot 220 + 88$, con $s = 1 - 2 \cdot (-1) = 3$ y $t = -2 - 2 \cdot 3 = -8$. La cuarta: $220 = 2 \cdot 88 + 44$, con $s = -1 - 2 \cdot 3 = -7$ y $t = 3 - 2 \cdot (-8) = 19$. Y la última: $88 = 2 \cdot 44 + 0$, y ahí paramos.

El último resto no nulo es 44, luego $\gcd(2024, 748) = 44$, y su fila da los coeficientes de Bézout:

$$2024 \cdot (-7) + 748 \cdot 19 = -14\,168 + 14\,212 = 44 \quad \checkmark$$

La comprobación del paso 5 es esa suma, y sale. Nótese de paso que la invariante $r = sa + tb$ se cumple en *todas* las filas, no solo en la buena: por ejemplo la del 220 dice $2024 \cdot (-1) + 748 \cdot 3 = -2024 + 2244 = 220$ ✓. Comprobar una fila intermedia es la mejor manera de cazar un error a mitad de camino.

Contraste con la factorización. Aquí es fácil verificar el resultado por el otro método: $2024 = 2^3 \cdot 11 \cdot 23$ y $748 = 2^2 \cdot 11 \cdot 17$, de modo que los factores comunes son 2^2 y 11, con producto $4 \cdot 11 = 44$ ✓. Pero obsérvese el trabajo: Euclides ha necesitado cinco divisiones y ninguna idea; factorizar ha exigido descubrir que 2024 esconde un 23 y 748 un 17. Con números de cien cifras el segundo método es sencillamente imposible y el primero sigue tardando un suspiro.

La identidad de Bézout quedó demostrada en la sección anterior. Lo que sí conviene registrar aquí es su forma fuerte, porque es la que se usará al resolver ecuaciones: no solo el máximo común divisor es alcanzable como combinación, sino que *los alcanzables son exactamente sus múltiplos.*

Teorema 3.22 (El conjunto de las combinaciones enteras)**.** *Sean a, b enteros no ambos nulos y $d = \gcd(a, b)$. Entonces*

$$\{\, ax + by : x, y \in \mathbb{Z} \,\} = \{\, \textit{múltiplos de } d \,\}.$$

En particular, existe una combinación $ax + by$ igual a c si y solo si $d \mid c$.

Demostración. Contenido en los múltiplos. Como $d \mid a$ y $d \mid b$, la propiedad de las combinaciones (proposición 3.10(3)) da $d \mid ax + by$ para cualesquiera x, y. Luego todo elemento del conjunto de la izquierda es múltiplo de d.

Contiene a los múltiplos. Por la identidad de Bézout existen x_0, y_0 con $ax_0 + by_0 = d$. Dado un múltiplo cualquiera kd, basta multiplicar por k:

$$a(kx_0) + b(ky_0) = kd,$$

que es una combinación entera de a y b. Los dos conjuntos coinciden. □

Corolario 3.23 (Caracterización de la coprimalidad)**.** *Dos enteros a y b son coprimos si y solo si existen enteros x, y con $ax + by = 1$.*

Demostración. Si $\gcd(a,b)=1$, el teorema 3.22 con $c=1$ da la combinación. Recíprocamente, si $ax+by=1$, entonces $d=\gcd(a,b)$ divide al lado izquierdo, luego $d \mid 1$, y como $d>0$ se tiene $d=1$. □

Nos falta la pareja del máximo común divisor, que el capítulo usa desde el principio sin haberla definido nunca.

Definición 3.24 (Mínimo común múltiplo)**.** Sean a, b enteros no nulos. El *mínimo común múltiplo* $\operatorname{mcm}(a,b)$ es el único entero m que cumple:

1. $m>0$, $a \mid m$ y $b \mid m$ (*es múltiplo común*);
2. si $a \mid c$ y $b \mid c$, entonces $m \mid c$ (*y divide a todo múltiplo común*).

De nuevo la condición (2) es más fuerte que "es el menor", y de nuevo las dos versiones coinciden en $\mathbb{Z}$: un múltiplo común positivo que divide a todos los demás es necesariamente el menor de ellos, por la propiedad del tamaño.

Teorema 3.25 (Relación entre máximo común divisor y mínimo común múltiplo)**.** *Para cualesquiera enteros positivos a, b,*

$$\gcd(a,b)\cdot \operatorname{mcm}(a,b) = ab.$$

Demostración. Sea $d=\gcd(a,b)$ y definamos $m=ab/d$, que es un entero porque $d \mid a$. Vamos a comprobar que m cumple las dos condiciones de la definición 3.24; por unicidad, será el mínimo común múltiplo, y la fórmula quedará probada.

Es múltiplo común. Escribiendo $a=d\alpha$ y $b=d\beta$ con α, β enteros, tenemos

$$m=\frac{ab}{d}=\frac{(d\alpha)b}{d}=\alpha b = a\beta,$$

lo que exhibe a m a la vez como múltiplo de b y como múltiplo de a.

Divide a todo múltiplo común. Sea c con $a \mid c$ y $b \mid c$; escribamos $c=au$. Por Bézout hay enteros x, y con $ax+by=d$. Dividamos c entre m mirando el cociente:

$$\frac{c}{m}=\frac{cd}{ab}=\frac{c(ax+by)}{ab}=\frac{cx}{b}+\frac{cy}{a}.$$

Ahora bien, c/b es entero porque $b \mid c$, y c/a es entero porque $a \mid c$. Luego c/m es suma de dos enteros, es decir, un entero: $m \mid c$.

Con las dos condiciones verificadas, $m=\operatorname{mcm}(a,b)$ y por tanto $d\cdot m = d\cdot(ab/d)=ab$. □

Ejemplo 3.26 (Mínimo común múltiplo sin factorizar)**.** Sigamos con $a=2024$ y $b=748$, cuyo máximo común divisor ya calculamos: 44. Entonces

$$\operatorname{mcm}(2024,748)=\frac{2024\cdot 748}{44}=2024\cdot\frac{748}{44}=2024\cdot 17=34\,408.$$

Fíjate en el orden de las operaciones: hemos simplificado *antes* de multiplicar, $748/44 = 17$, para no manejar el producto completo. Verificación de la fórmula del teorema 3.25:

$$44 \cdot 34\,408 = 1\,513\,952 \qquad \text{y} \qquad 2024 \cdot 748 = 1\,513\,952 \quad \checkmark$$

Y verificación de que es múltiplo común: $34\,408 = 2024 \cdot 17$ y $34\,408 = 748 \cdot 46$, pues $748 \cdot 46 = 34\,408$ ✓. Lo importante es el método: hemos obtenido el mínimo común múltiplo de dos números de cuatro cifras con una sola división extra sobre el algoritmo de Euclides, sin factorizar nada.

3.2.5. Factorización única y valuación p-ádica

El lema de Euclides y el teorema fundamental de la aritmética se demuestran en la sección siguiente, y no vamos a repetir aquí esas demostraciones. Sí necesitamos dos cosas que allí no aparecen: la versión general del lema (la que no exige que el número sea primo, y que es la que se usará al resolver congruencias) y una notación para leer exponentes en una factorización.

Lema 3.27 (Lema de Euclides generalizado, o lema de Gauss)**.** *Si $c \mid ab$ y $\gcd(c, a) = 1$, entonces $c \mid b$.*

Demostración. Por el corolario 3.23 existen enteros x, y con $cx + ay = 1$. Multipliquemos por b:

$$cbx + aby = b.$$

El primer sumando es múltiplo de c a la vista. El segundo también, porque $c \mid ab$ por hipótesis. Por la propiedad de las combinaciones, c divide a la suma, que es b. □

El lema de Euclides habitual es el caso particular $c = p$ primo: si $p \nmid a$, entonces $\gcd(p, a) = 1$ porque los únicos divisores positivos de p son 1 y p, y se aplica lo anterior. La versión general es estrictamente más útil: por ejemplo, de $6 \mid 4 \cdot 9$ no se deduce nada porque $\gcd(6, 4) = 2 \neq 1$, pero de $7 \mid 4k$ sí se deduce $7 \mid k$.

Teorema 3.28 (Teorema fundamental de la aritmética)**.** *Todo entero $n > 1$ se escribe como producto de primos, y esa escritura es única salvo el orden de los factores. Agrupando los primos repetidos, existe una única expresión*

$$n = p_1^{a_1} p_2^{a_2} \cdots p_k^{a_k}, \qquad p_1 < p_2 < \cdots < p_k \text{ primos}, \ a_i \geq 1.$$

Existencia. Por inducción fuerte sobre n. Si n es primo, ya es un producto de un solo factor. Si es compuesto, se escribe $n = ab$ con $1 < a < n$ y $1 < b < n$; por hipótesis de inducción a y b son productos de primos, y concatenando las dos listas se obtiene una factorización de n. La unicidad, que es la parte con contenido real, se demuestra en la sección siguiente a partir del lema de Euclides. □

Definición 3.29 (Valuación p-ádica)**.** Sea p un primo y n un entero no nulo. La *valuación p-ádica* de n, escrita $v_p(n)$, es el exponente con que p aparece en la factorización de n;

equivalentemente, el mayor entero $e \geq 0$ tal que $p^e \mid n$. Con esta notación, todo entero $n > 1$ se escribe como

$$n = \prod_p p^{\,v_p(n)},$$

producto extendido a todos los primos, en el que solo un número finito de exponentes es no nulo.

Proposición 3.30 (Divisibilidad, máximo común divisor y mínimo común múltiplo por exponentes)**.** *Sean a, b enteros positivos. Entonces:*

1. $v_p(ab) = v_p(a) + v_p(b)$ *para todo primo* p.
2. $a \mid b$ *si y solo si* $v_p(a) \leq v_p(b)$ *para todo primo* p.
3. $v_p(\gcd(a,b)) = \min\{v_p(a), v_p(b)\}$ *y* $v_p(\operatorname{mcm}(a,b)) = \max\{v_p(a), v_p(b)\}$.

Demostración. (1) Multiplicando las factorizaciones de a y de b y agrupando primos iguales, el exponente de p en el producto es la suma de los exponentes; la unicidad del teorema 3.28 garantiza que esa es *la* factorización de ab.

(2) Si $a \mid b$, digamos $b = ak$, entonces por (1) se tiene $v_p(b) = v_p(a) + v_p(k) \geq v_p(a)$. Recíprocamente, si $v_p(a) \leq v_p(b)$ para todo p, el número $k = \prod_p p^{\,v_p(b) - v_p(a)}$ es un entero (todos los exponentes son ≥ 0) y cumple $ak = b$.

(3) Sea $d = \prod_p p^{\min\{v_p(a), v_p(b)\}}$. Por (2), d divide a a y a b, porque su exponente en cada primo no supera al de ninguno de los dos. Y si $c \mid a$ y $c \mid b$, entonces $v_p(c) \leq v_p(a)$ y $v_p(c) \leq v_p(b)$ para todo p, luego $v_p(c) \leq \min\{v_p(a), v_p(b)\}$ y por (2) otra vez $c \mid d$. Las dos condiciones de la definición 3.19 se cumplen, así que $d = \gcd(a,b)$. El caso del mínimo común múltiplo es idéntico invirtiendo las desigualdades. Nótese que de aquí sale una segunda demostración del teorema 3.25, porque $\min\{u,v\} + \max\{u,v\} = u + v$ para cualquier par de números. □

Ejemplo 3.31 (Leer exponentes)**.** Con $2024 = 2^3 \cdot 11 \cdot 23$ tenemos $v_2(2024) = 3$, $v_{11}(2024) = 1$, $v_{23}(2024) = 1$ y $v_p(2024) = 0$ para todos los demás primos, empezando por $v_3(2024) = 0$ y $v_5(2024) = 0$. Con $748 = 2^2 \cdot 11 \cdot 17$: $v_2 = 2$, $v_{11} = 1$, $v_{17} = 1$.

Aplicando la proposición 3.30(3) primo a primo:

$$\gcd = 2^{\min(3,2)} \cdot 11^{\min(1,1)} \cdot 17^{\min(0,1)} \cdot 23^{\min(1,0)} = 2^2 \cdot 11 = 44,$$

$$\operatorname{mcm} = 2^{\max(3,2)} \cdot 11^{\max(1,1)} \cdot 17^{\max(0,1)} \cdot 23^{\max(1,0)} = 2^3 \cdot 11 \cdot 17 \cdot 23 = 34\,408,$$

que coinciden con lo calculado por Euclides ✓. Presta atención a los exponentes cero: el 17 no está en 2024 y el 23 no está en 748, y por eso ninguno de los dos entra en el máximo común divisor, pero *ambos* entran en el mínimo común múltiplo. Ese es el error clásico de este cálculo.

3.2.6. Congruencias

Definición 3.32 (Congruencia módulo n)**.** Sea $n \geq 1$ un entero, llamado *módulo*. Dados $a, b \in \mathbb{Z}$, decimos que a es *congruente* con b *módulo* n, y escribimos

$$a \equiv b \pmod{n},$$

si $n \mid a-b$. Cuando no lo es, se escribe $a \not\equiv b \pmod{n}$.

Proposición 3.33 (Tres maneras equivalentes de decir lo mismo)**.** *Para $n \geq 1$ y $a, b \in \mathbb{Z}$, las siguientes afirmaciones son equivalentes:*

1. *$n \mid a-b$;*
2. *a y b dejan el mismo resto al dividirlos entre n;*
3. *existe $k \in \mathbb{Z}$ con $a = b + kn$.*

Demostración. (1)⇒(2): dividamos ambos según el teorema 3.13, $a = nq_1 + r_1$ y $b = nq_2 + r_2$ con $0 \leq r_1, r_2 < n$. Entonces $a - b = n(q_1 - q_2) + (r_1 - r_2)$, y como $n \mid a-b$ y $n \mid n(q_1-q_2)$, la propiedad de las combinaciones da $n \mid r_1 - r_2$. Pero $|r_1 - r_2| < n$, y el único múltiplo de n menor que n en valor absoluto es el cero: $r_1 = r_2$.

(2)⇒(3): con la misma notación, si $r_1 = r_2$ entonces $a - b = n(q_1 - q_2)$, luego basta tomar $k = q_1 - q_2$. (3)⇒(1): si $a = b + kn$ entonces $a - b = kn$, que es múltiplo de n. □

La caracterización (2) es la que justifica la imagen del reloj, y la (3) es la que se usa al resolver sistemas por sustitución. Ahora la propiedad estructural: las congruencias se comportan como una igualdad.

Proposición 3.34 (La congruencia es una relación de equivalencia)**.** *Para todo $n \geq 1$ y todos $a, b, c \in \mathbb{Z}$:*

1. (Reflexiva) $a \equiv a \pmod{n}$.
2. (Simétrica) *Si* $a \equiv b \pmod{n}$, *entonces* $b \equiv a \pmod{n}$.
3. (Transitiva) *Si* $a \equiv b$ *y* $b \equiv c \pmod{n}$, *entonces* $a \equiv c \pmod{n}$.

Demostración. (1) $a-a = 0$ y $n \mid 0$. (2) Si $n \mid a-b$, digamos $a-b = kn$, entonces $b-a = (-k)n$, luego $n \mid b-a$. (3) Si $n \mid a-b$ y $n \mid b-c$, la propiedad de las combinaciones (proposición 3.10(3)) da que n divide a la suma $(a-b)+(b-c) = a-c$. □

Esto no es una formalidad: es precisamente lo que permite escribir cadenas $a \equiv b \equiv c \equiv \cdots$ y manipular congruencias como si fueran ecuaciones. Y es lo que, dentro de un momento, hará que tenga sentido hablar de "clases".

Proposición 3.35 (Compatibilidad con las operaciones)**.** *Si $a \equiv b \pmod{n}$ y $c \equiv d \pmod{n}$, entonces*

$$a + c \equiv b + d, \qquad a - c \equiv b - d, \qquad ac \equiv bd \pmod{n},$$

y en consecuencia $a^k \equiv b^k \pmod{n}$ para todo entero $k \geq 1$.

Demostración. Sea $a - b = kn$ y $c - d = \ell n$. Para la suma y la resta, $(a \pm c) - (b \pm d) = (a-b) \pm (c-d) = (k \pm \ell)n$. Para el producto, sumamos y restamos el mismo término:

$$ac - bd = ac - bc + bc - bd = c(a-b) + b(c-d) = (ck + b\ell)n.$$

Las potencias salen aplicando el producto repetidamente, formalmente por inducción sobre k. □

Corolario 3.36 (Los polinomios respetan las congruencias)**.** *Si P es un polinomio con coeficientes enteros y $a \equiv b \pmod{n}$, entonces $P(a) \equiv P(b) \pmod{n}$.*

Demostración. Cada monomio $c_j x^j$ cumple $c_j a^j \equiv c_j b^j$ por la proposición 3.35, y sumando todos los monomios se conserva la congruencia. □

Este corolario, que parece técnico, es el motor oculto de los criterios de divisibilidad: un número escrito en base diez *es* el valor de un polinomio en sus cifras evaluado en 10, y todo se reduce a sustituir 10 por su resto.

Cuidado

[La división no está en la lista] La proposición 3.35 habla de sumar, restar y multiplicar. **Dividir no aparece, y no es un olvido.** De $ac \equiv bc \pmod{n}$ *no* se sigue en general $a \equiv b$. El contraejemplo mínimo es este: módulo 15,

$$6 \cdot 4 = 24 \equiv 9 \qquad \text{y} \qquad 6 \cdot 9 = 54 \equiv 9,$$

así que $6 \cdot 4 \equiv 6 \cdot 9 \pmod{15}$; y sin embargo $4 \not\equiv 9 \pmod{15}$, porque $9 - 4 = 5$ no es múltiplo de 15. Cancelar el 6 habría producido una falsedad. Lo que sí se puede hacer, y con exactitud, lo dice el teorema siguiente.

Teorema 3.37 (Ley de cancelación)**.** *Sean $a, b, c \in \mathbb{Z}$ y $n \geq 1$, y sea $g = \gcd(c, n)$. Entonces*

$$ac \equiv bc \pmod{n} \qquad \Longleftrightarrow \qquad a \equiv b \pmod{\tfrac{n}{g}}.$$

En particular, si $\gcd(c, n) = 1$ se puede cancelar c sin tocar el módulo.

Demostración. Escribamos $n = g\,n'$ y $c = g\,c'$, de modo que $n' = n/g$ y, por la proposición 3.30(3) o simplemente por la definición de máximo común divisor, $\gcd(c', n') = 1$: si un $e > 1$ dividiera a los dos, ge sería un divisor común de c y n mayor que g.

($\Rightarrow$) Si $n \mid (a-b)c$, escribamos $(a-b)c = nk$, es decir $(a-b)gc' = gn'k$. Cancelando el factor g, que es un entero no nulo, queda $(a-b)c' = n'k$, luego $n' \mid (a-b)c'$. Como $\gcd(n', c') = 1$, el lema de Gauss (lema 3.27) permite quitar el c': $n' \mid a - b$, es decir, $a \equiv b \pmod{n'}$.

($\Leftarrow$) Si $n' \mid a-b$, digamos $a-b = n'k$, entonces $(a-b)c = n'k\,gc' = n\,(kc')$, luego $n \mid (a-b)c$, es decir $ac \equiv bc \pmod{n}$. □

Ejemplo 3.38 (La cancelación correcta, con el contraejemplo de antes)**.** Retomemos $6 \cdot 4 \equiv 6 \cdot 9 \pmod{15}$. Aquí $c = 6$, $n = 15$ y $g = \gcd(6, 15) = 3$, luego el teorema autoriza a cancelar el 6 *a condición de dividir el módulo entre* 3:

$$4 \equiv 9 \pmod{5}.$$

¿Es cierto? $9-4=5$, que sí es múltiplo de 5 ✓. La conclusión falsa era $4 \equiv 9 \pmod{15}$; la verdadera es $4 \equiv 9 \pmod{5}$. El teorema no prohíbe cancelar: obliga a pagar el precio en el módulo.

Y el caso cómodo: módulo 15, de $7x \equiv 7y$ sí se sigue $x \equiv y \pmod{15}$, porque $\gcd(7,15)=1$ y entonces $n/g = 15$. Antes de cancelar, siempre la misma pregunta: ¿es el factor coprimo con el módulo?

3.2.7. Clases de restos, el conjunto $\mathbb{Z}/n\mathbb{Z}$ y las unidades

El capítulo escribe $\mathbb{Z}/n\mathbb{Z}$ y $(\mathbb{Z}/n\mathbb{Z})^\times$ en varios sitios. Definamos esos símbolos.

Definición 3.39 (Clase de restos y el conjunto $\mathbb{Z}/n\mathbb{Z}$)**.** Sea $n \geq 1$. Para cada entero a, su *clase de restos módulo n* es el conjunto de todos los enteros congruentes con él:

$$[a]_n = \{\, x \in \mathbb{Z} : x \equiv a \pmod{n} \,\} = \{\, a + kn : k \in \mathbb{Z} \,\}.$$

El conjunto de todas las clases se denota

$$\mathbb{Z}/n\mathbb{Z} = \{[0]_n, [1]_n, \ldots, [n-1]_n\},$$

y tiene exactamente n elementos. Un *sistema completo de restos* módulo n es cualquier elección de n enteros, uno de cada clase. Un *sistema reducido de restos* es una elección de un representante por cada clase $[a]_n$ con $\gcd(a,n)=1$; tiene $\varphi(n)$ elementos.

Que $\mathbb{Z}/n\mathbb{Z}$ tenga exactamente n elementos es consecuencia directa de la división euclídea: todo entero a se escribe $a = nq + r$ con $0 \leq r < n$, luego $[a]_n = [r]_n$ y las clases listadas agotan todas las posibilidades; y son distintas dos a dos, porque dos restos distintos del rango $[0, n)$ difieren en menos de n y no pueden ser congruentes.

Ejemplo 3.40 (Sistemas de restos)**. (a) Módulo** 6**.** El sistema completo estándar es $\{0,1,2,3,4,5\}$, pero cualquier otra elección de un representante por clase vale igual. Por ejemplo $\{6, 13, 20, 27, 34, 41\}$ también es un sistema completo, porque sus restos son

$$6 \equiv 0, \quad 13 \equiv 1, \quad 20 \equiv 2, \quad 27 \equiv 3, \quad 34 \equiv 4, \quad 41 \equiv 5 \pmod{6},$$

uno de cada. Y también lo es $\{-2,-1,0,1,2,3\}$, que a veces resulta más cómodo porque los números son pequeños en valor absoluto.

(b) Un sistema reducido módulo 12**.** Los enteros de 1 a 12 coprimos con 12 son $1, 5, 7, 11$: son cuatro, y en efecto $\varphi(12) = 4$. Los demás comparten con 12 el factor 2 o el factor 3. Nótese que $\{13, 17, 19, 23\}$ es otro sistema reducido módulo 12, porque $13 \equiv 1$, $17 \equiv 5$, $19 \equiv 7$ y $23 \equiv 11$.

Para poder *operar* con clases hay que definir la suma y el producto, y ahí surge la pregunta delicada: si defino $[a] + [b] = [a+b]$, el resultado se ha calculado usando los representantes a y b concretos. ¿Qué pasa si alguien elige otros? Si la respuesta cambiara, la operación no existiría. Que no cambie es lo que se llama estar *bien definida*, y hay que demostrarlo.

Proposición 3.41 (Las operaciones de $\mathbb{Z}/n\mathbb{Z}$ están bien definidas)**.** *Las reglas*

$$[a]_n + [b]_n = [a+b]_n, \qquad [a]_n \cdot [b]_n = [ab]_n$$

no dependen de los representantes elegidos: si $[a]_n = [a']_n$ *y* $[b]_n = [b']_n$*, entonces* $[a+b]_n = [a'+b']_n$ *y* $[ab]_n = [a'b']_n$*.*

Demostración. Que $[a]_n = [a']_n$ significa exactamente $a \equiv a' \pmod{n}$, y análogamente $b \equiv b'$. Por la proposición 3.35, $a+b \equiv a'+b'$ y $ab \equiv a'b'$ módulo n; y dos enteros congruentes tienen la misma clase. Luego $[a+b]_n = [a'+b']_n$ y $[ab]_n = [a'b']_n$. □

Idea clave

Merece la pena parar en lo que acaba de ocurrir, porque es un patrón que se repetirá en todo el libro. Hemos tomado un conjunto infinito ($\mathbb{Z}$), lo hemos partido en n paquetes, y hemos comprobado que las operaciones sobreviven al aplastamiento. La proposición 3.35 era un enunciado sobre *números*; la proposición 3.41 lo convierte en un enunciado sobre *estructuras*. Es la misma proposición dicha dos veces, y sin embargo la segunda versión es la que permite decir que $\mathbb{Z}/n\mathbb{Z}$ es un anillo con n elementos y conectar todo este capítulo con el álgebra abstracta.

Definición 3.42 (Unidades e inverso modular)**.** Una clase $[a]_n$ es una *unidad* de $\mathbb{Z}/n\mathbb{Z}$ si existe $[x]_n$ con $[a]_n[x]_n = [1]_n$, es decir, si existe un entero x con

$$ax \equiv 1 \pmod{n}.$$

A ese x se le llama *inverso de a módulo n* y se escribe a^{-1}; es único módulo n. El conjunto de las unidades se denota $(\mathbb{Z}/n\mathbb{Z})^\times$ y tiene $\varphi(n)$ elementos.

Proposición 3.43 (Criterio de invertibilidad y unicidad)**.** *El entero a tiene inverso módulo n si y solo si* $\gcd(a,n) = 1$*. Cuando existe, es único módulo n.*

Demostración. Si $\gcd(a,n) = 1$, el corolario 3.23 da enteros x, y con $ax + ny = 1$; leyendo esa igualdad módulo n, el término ny desaparece y queda $ax \equiv 1$. Recíprocamente, si $ax \equiv 1 \pmod{n}$, entonces $ax - 1 = n(-y)$ para algún y, o sea $ax + ny = 1$, y el corolario 3.23 da $\gcd(a,n) = 1$.

Unicidad: si $ax \equiv 1$ y $ax' \equiv 1$, entonces $ax \equiv ax'$, y como $\gcd(a,n) = 1$ la ley de cancelación (teorema 3.37) con $g = 1$ permite cancelar a sin cambiar el módulo: $x \equiv x' \pmod{n}$. □

Regla: calcular el inverso de a módulo n con la tabla

1. **Comprueba primero** $\gcd(a,n) = 1$**.** Si no lo es, no hay inverso y no se trata de buscar mejor: la proposición 3.43 lo prohíbe.
2. **Monta la tabla de Euclides extendido** con (n, a) en ese orden, es decir, el módulo arriba.
3. **Lee la fila del último resto no nulo,** que será 1: da coeficientes con $1 = s\,n + t\,a$.

4. **El inverso es** t, el coeficiente que acompaña a a. Redúcelo al rango $[0, n)$ sumándole n tantas veces como haga falta.
5. **Verifica multiplicando:** $a \cdot a^{-1}$ debe dejar resto 1 al dividirlo entre n. Este paso no es opcional; cierra casi todos los errores de signo.

Ejemplo 3.44 (El inverso de 35 módulo 88)**.** Primero el paso 1: $\gcd(35, 88) = 1$, porque $88 = 2^3 \cdot 11$ y $35 = 5 \cdot 7$ no comparten factores. Existe, pues, el inverso. Montamos la tabla con $a = 88$ y $b = 35$:

q	r	s	t
—	88	1	0
—	35	0	1
2	18	1	-2
1	17	-1	3
1	1	2	-5
17	0	—	

Las divisiones son $88 = 2 \cdot 35 + 18$, $35 = 1 \cdot 18 + 17$, $18 = 1 \cdot 17 + 1$ y $17 = 17 \cdot 1 + 0$. La fila del 1 dice

$$1 = 2 \cdot 88 + (-5) \cdot 35 = 176 - 175 \quad \checkmark$$

Leyendo módulo 88, el término $2 \cdot 88$ desaparece y queda $(-5) \cdot 35 \equiv 1 \pmod{88}$. El inverso es $t = -5$, que reducido al rango pedido es

$$35^{-1} \equiv -5 + 88 = 83 \pmod{88}.$$

Verificación del paso 5: $35 \cdot 83 = 2905$, y $2905 = 33 \cdot 88 + 1 = 2904 + 1$, resto 1 ✓.

Un caso sin inverso, para contrastar. Busquemos el inverso de 34 módulo 88. Aquí $\gcd(34, 88) = 2$, así que no existe, y la razón se ve directamente: $34x$ es siempre par, $88k$ es siempre par, luego $34x - 88k$ es par y nunca puede valer 1. No es que no lo hayamos encontrado: es que no está.

3.2.8. La ecuación diofántica lineal

Con Bézout en la mano podemos resolver por completo la ecuación entera más sencilla que existe. Una ecuación se llama *diofántica* cuando solo se aceptan soluciones enteras, y la condición de ser entero lo cambia todo: $2x + 4y = 3$ tiene infinitas soluciones racionales y ninguna entera.

Teorema 3.45 (Ecuación diofántica lineal)**.** *Sean a, b enteros no ambos nulos, c un entero y $d = \gcd(a, b)$. La ecuación*

$$ax + by = c$$

tiene solución entera si y solo si $d \mid c$. Si la tiene y (x_0, y_0) es una solución particular, entonces todas *las soluciones son*

$$x = x_0 + \frac{b}{d}\,t, \qquad y = y_0 - \frac{a}{d}\,t, \qquad t \in \mathbb{Z},$$

y valores distintos de t dan pares distintos.

Demostración. El criterio. Es exactamente el teorema 3.22: los valores que puede tomar $ax + by$ son precisamente los múltiplos de d.

Las soluciones propuestas valen. Sustituyendo,

$$a\Big(x_0 + \frac{b}{d}t\Big) + b\Big(y_0 - \frac{a}{d}t\Big) = ax_0 + by_0 + \frac{ab}{d}t - \frac{ab}{d}t = c,$$

y los dos términos con t se cancelan. Además b/d y a/d son enteros, así que x e y lo son.

No hay más. Sea (x, y) otra solución cualquiera. Restando las dos ecuaciones,

$$a(x - x_0) + b(y - y_0) = 0 \quad \Longrightarrow \quad a(x - x_0) = -b(y - y_0).$$

Dividamos todo entre d, escribiendo $a = d\alpha$ y $b = d\beta$:

$$\alpha(x - x_0) = -\beta(y - y_0),$$

donde ahora $\gcd(\alpha, \beta) = 1$ (si tuvieran un factor común $e > 1$, entonces de sería un divisor común de a y b mayor que d). De la igualdad se sigue que β divide al producto $\alpha(x - x_0)$; como β es coprimo con α, el lema de Gauss (lema 3.27) da $\beta \mid x - x_0$, es decir, $x - x_0 = \beta t = \frac{b}{d}t$ para algún entero t. Sustituyendo arriba, $\alpha\beta t = -\beta(y - y_0)$, y cancelando $\beta \neq 0$ queda $y - y_0 = -\alpha t = -\frac{a}{d}t$. Es la forma anunciada.

Por último, si dos valores $t \neq t'$ dieran el mismo x, sería $\frac{b}{d}(t - t') = 0$ con $b \neq 0$, imposible; y si $b = 0$ se razona con la coordenada y. □

Ejemplo 3.46 (Una ecuación diofántica resuelta entera: $42x + 30y = 18$)**. Paso 1: ¿hay solución?** Calculamos $\gcd(42, 30)$ con Euclides:

$$42 = 1 \cdot 30 + 12, \qquad 30 = 2 \cdot 12 + 6, \qquad 12 = 2 \cdot 6 + 0,$$

luego $d = 6$. Como $6 \mid 18$ (pues $18 = 6 \cdot 3$), el teorema 3.45 garantiza que hay soluciones, y de hecho infinitas.

Paso 2: una solución particular. Deshacemos las divisiones:

$$6 = 30 - 2 \cdot 12 = 30 - 2\,(42 - 30) = 3 \cdot 30 - 2 \cdot 42,$$

es decir, $42 \cdot (-2) + 30 \cdot 3 = -84 + 90 = 6$ ✓. Como necesitamos $18 = 3 \cdot 6$, multiplicamos los coeficientes por 3:

$$x_0 = 3 \cdot (-2) = -6, \qquad y_0 = 3 \cdot 3 = 9.$$

Comprobación: $42 \cdot (-6) + 30 \cdot 9 = -252 + 270 = 18$ ✓.

Paso 3: la solución general. Con $b/d = 30/6 = 5$ y $a/d = 42/6 = 7$:

$$x = -6 + 5t, \qquad y = 9 - 7t, \qquad t \in \mathbb{Z}.$$

Paso 4: comprobar varios valores. Con $t = 1$ sale $(x, y) = (-1, 2)$ y $-42 + 60 = 18$ ✓. Con $t = 2$ sale $(4, -5)$ y $168 - 150 = 18$ ✓. Con $t = 3$ sale $(9, -12)$ y $378 - 360 = 18$ ✓. Con $t = 0$ recuperamos la particular. Todo encaja.

Paso 5: ¿y si se piden soluciones positivas? Habría que imponer a la vez $-6+5t>0$, es decir $t>6/5=1{,}2$, y $9-7t>0$, es decir $t<9/7=1{,}285\ldots$ No hay ningún entero en ese intervalo, así que *esta* ecuación no tiene soluciones con x e y ambos positivos, aunque tenga infinitas soluciones enteras. Cambiando el término independiente la situación cambia: para $42x+30y=246$ los coeficientes particulares se multiplican por $246/6=41$, dando $x_0=-82$, $y_0=123$, y la solución general $x=-82+5t$, $y=123-7t$; ahora $t=17$ produce $x=3$, $y=4$, y en efecto $42\cdot 3+30\cdot 4=126+120=246$ ✓.

La moraleja del paso 5 conviene retenerla: el teorema resuelve la existencia de soluciones *enteras*, y la pregunta por las soluciones *no negativas* es un problema distinto y bastante más difícil, del que se ocupa el problema de las monedas de Frobenius.

3.2.9. Los criterios de divisibilidad, y de dónde salen

Todos los criterios que se aprenden de memoria en la escuela son el mismo teorema aplicado nueve veces. Escribamos un número en base diez separando sus cifras $c_k,\ldots,c_1,c_0$:

$$N=c_k10^k+\cdots+c_1\cdot 10+c_0.$$

Esto es un polinomio con coeficientes enteros evaluado en 10. Por el corolario 3.36, para conocer N módulo m basta conocer 10 módulo m, o más finamente, los restos de las potencias 10^j módulo m. Todo lo demás es aritmética de una línea.

Regla: criterios de divisibilidad, con su justificación

1. **Por 2:** la última cifra es par. *Porque* $10\equiv 0 \pmod 2$, luego toda potencia 10^j con $j\geq 1$ es $\equiv 0$ y $N\equiv c_0 \pmod 2$.
2. **Por 5:** la última cifra es 0 o 5. *Porque* $10\equiv 0 \pmod 5$ y, igual que antes, $N\equiv c_0 \pmod 5$.
3. **Por 10:** la última cifra es 0. *Porque* $10\equiv 0 \pmod{10}$ y $N\equiv c_0 \pmod{10}$.
4. **Por 4:** el número formado por las dos últimas cifras es múltiplo de 4. *Porque* $100=4\cdot 25\equiv 0 \pmod 4$, luego $10^j\equiv 0$ para $j\geq 2$ y solo sobrevive $N\equiv 10c_1+c_0 \pmod 4$.
5. **Por 8:** el número formado por las tres últimas cifras es múltiplo de 8. *Porque* $1000=8\cdot 125\equiv 0 \pmod 8$, luego $10^j\equiv 0$ para $j\geq 3$ y $N\equiv 100c_2+10c_1+c_0 \pmod 8$.
6. **Por 3:** la suma de las cifras es múltiplo de 3. *Porque* $10\equiv 1 \pmod 3$, luego $10^j\equiv 1^j=1$ para todo j y $N\equiv c_k+\cdots+c_1+c_0 \pmod 3$.
7. **Por 9:** la suma de las cifras es múltiplo de 9. *Porque* $10\equiv 1 \pmod 9$, exactamente el mismo cálculo que para el 3.
8. **Por 11:** la suma alternada de las cifras, empezando por las unidades con signo más, es múltiplo de 11. *Porque* $10\equiv -1 \pmod{11}$, luego $10^j\equiv(-1)^j$ y $N\equiv c_0-c_1+c_2-c_3+\cdots \pmod{11}$.

9. **Por 6:** ser divisible por 2 y por 3 a la vez. *Porque* $6 = 2 \cdot 3$ con $\gcd(2,3) = 1$: si $2 \mid N$ y $3 \mid N$, el lema de Gauss (lema 3.27) aplicado a $N = 2u$ y $3 \mid 2u$ con $\gcd(3,2) = 1$ da $3 \mid u$, luego $6 \mid N$. La coprimalidad es imprescindible: ser divisible por 2 y por 4 no implica serlo por 8.

Lo importante no son las nueve recetas, sino la observación única de la que salen: los criterios de "últimas cifras" aparecen cuando alguna potencia de diez es $\equiv 0$, y los de "suma de cifras" cuando $10 \equiv \pm 1$. Y hay un regalo: los criterios no solo dicen si divide, dan *el resto exacto*.

Ejemplo 3.47 (Un solo número contra los nueve criterios)**.** Tomemos $N = 3\,517\,284$, cuyas cifras desde las unidades son $4, 8, 2, 7, 1, 5, 3$.

Por 2, 5 y 10. La última cifra es 4: par, luego $2 \mid N$; no es 0 ni 5, luego $5 \nmid N$; no es 0, luego $10 \nmid N$. De hecho $N \equiv 4 \pmod{10}$.

Por 4 y por 8. Las dos últimas cifras forman $84 = 4 \cdot 21$, luego $4 \mid N$; en efecto $N = 4 \cdot 879\,321$. Las tres últimas forman 284, y $284 = 8 \cdot 35 + 4$, así que $8 \nmid N$ y además $N \equiv 4 \pmod{8}$.

Por 3 y por 9. La suma de cifras es $3 + 5 + 1 + 7 + 2 + 8 + 4 = 30$. Como $3 \mid 30$, se tiene $3 \mid N$. Como $30 = 9 \cdot 3 + 3$, resulta $N \equiv 3 \pmod{9}$ y por tanto $9 \nmid N$. Obsérvese la precisión del método: no solo sabemos que 9 no divide, sabemos que sobra 3.

Por 6. Es divisible por 2 y por 3, luego por 6: en efecto $N = 6 \cdot 586\,214$ ✓.

Por 11. Suma alternada desde las unidades:

$$4 - 8 + 2 - 7 + 1 - 5 + 3 = -10.$$

Como $-10 \equiv 1 \pmod{11}$, tenemos $N \equiv 1 \pmod{11}$ y por tanto $11 \nmid N$. Cuidado con el punto de partida: hay que empezar por las unidades con signo +; empezar por la izquierda sin vigilar la paridad del número de cifras cambia el signo global y puede dar un resultado aparentemente distinto.

Resumen. N es divisible por 2, 3, 4 y 6, y no lo es por 5, 8, 9, 10 ni 11. Todo ello sin haber hecho una sola división de siete cifras.

Pausa para pensar

Repasa hacia atrás lo que hemos construido, porque el orden no es casual y conviene verlo como una cadena. La *buena ordenación* dio la división euclídea. La división euclídea dio el algoritmo de Euclides y, con él, la identidad de Bézout. Bézout dio el lema de Gauss, la ley de cancelación, el criterio de invertibilidad y la ecuación diofántica lineal. El lema de Gauss, en su caso primo, dará la unicidad de la factorización. Y la compatibilidad de las congruencias con las operaciones dio a la vez los criterios de divisibilidad y la existencia misma de $\mathbb{Z}/n\mathbb{Z}$.
No hay ningún eslabón suelto: todo el capítulo, incluidos el teorema de Euler y el RSA que lo cierran, cuelga de ese primer principio sobre los enteros no negativos. Si tuvieras

que quedarte con una sola frase de esta sección, que sea esta: *en teoría de números casi todo se demuestra dividiendo con resto y mirando qué es demasiado pequeño para ser un múltiplo, salvo que sea cero.*

3.3. Los números primos y el teorema fundamental de la aritmética

Idea clave

Intuición primero. Los primos son los átomos de la multiplicación: los números que no se pueden partir. El teorema fundamental de la aritmética dice que todo entero se construye con ellos de *una sola manera*, y merece la pena resistir la tentación de darlo por obvio, porque no lo es. Que exista alguna factorización es fácil (parte y sigue partiendo). Que sea única es un teorema con contenido real, y la prueba de que tiene contenido es que en otros sistemas numéricos perfectamente razonables la unicidad *falla*: en los números de la forma $a + b\sqrt{-5}$ ocurre que $6 = 2 \cdot 3 = (1 + \sqrt{-5})(1 - \sqrt{-5})$, con los cuatro factores irreducibles. El descubrimiento de ese fracaso, en el siglo XIX, es lo que dio origen a la teoría de ideales y a buena parte del álgebra moderna.
Sobre la distribución de los primos, la situación es curiosa: sabemos desde Euclides que son infinitos, con una demostración que cabe en tres líneas y es de una elegancia insuperable; y sin embargo las preguntas más simples sobre cómo se reparten —si hay infinitos primos gemelos, por ejemplo— siguen abiertas después de dos siglos.

Un *número primo* es un entero mayor que 1 cuyos únicos divisores positivos son 1 y él mismo: $2, 3, 5, 7, 11, 13, \ldots$ Los demás enteros mayores que 1 se llaman *compuestos* porque se parten en factores más pequeños ($6 = 2 \cdot 3$, $12 = 2 \cdot 2 \cdot 3$). Los primos son los ladrillos con que se construyen todos los enteros por multiplicación, y este capítulo entero gira, en el fondo, alrededor de ellos.

Quiero probar dos cosas sobre los primos, y quiero que veas por qué ninguna de las dos es tan obvia como parece.

3.3.1. Hay infinitos primos

La lista $2, 3, 5, 7, 11, \ldots$ ¿se acaba en algún momento, o sigue para siempre? Euclides respondió hace más de dos milenios, y su demostración sigue siendo de una belleza intacta.

Teorema 3.48 (Euclides)**.** *Hay infinitos números primos.*

Demostración. Vamos a demostrar que ninguna lista finita de primos puede contenerlos a todos. Supón que alguien te presenta una lista finita de primos, cualquiera: $p_1, p_2, \ldots, p_n$. Voy a fabricar un primo que no está en su lista, y con eso quedará claro que la lista estaba incompleta.

Considera el número

$$N = p_1 \cdot p_2 \cdots p_n + 1,$$

es decir, multiplica todos los primos de la lista y súmale 1. Este N es mayor que 1, así que tiene al menos un divisor primo; llamémoslo p. Afirmo que p no está en la lista. ¿Por qué? Porque si p fuera uno de los p_i, entonces p dividiría al producto $p_1 \cdots p_n$, y también divide a N por elección; pero entonces p dividiría a la diferencia $N - p_1 \cdots p_n = 1$, lo cual es imposible: ningún primo divide a 1. Contradicción. Así que p es un primo nuevo, fuera de la lista.

Como *toda* lista finita deja fuera algún primo, no existe una lista finita completa: hay infinitos. □

Ejemplo 3.49 (La demostración de Euclides y sus variantes)**. (a) La construcción.** Con los primos $2, 3, 5$ se forma $2 \cdot 3 \cdot 5 + 1 = 31$, que es primo y no está en la lista. Con $2, 3, 5, 7, 11, 13$ sale $30031 = 59 \cdot 509$: *no* es primo, pero sus factores tampoco están en la lista. Esto último es lo que la demostración necesita, y conviene no contarla mal: Euclides no afirma que $p_1 \cdots p_n + 1$ sea primo.

(b) Cuántos hay. El teorema de los números primos dice que hay aproximadamente $x/\ln x$ primos por debajo de x. Para $x = 10^6$ eso da 72 382 frente a los 78 498 reales, un error del ocho por ciento que se reduce usando el logaritmo integral.

(c) La demostración de Euler. $\sum 1/p$ diverge, luego hay infinitos primos, y además diverge como $\ln \ln x$: crecimiento tan lento que sumando los inversos de todos los primos menores que un billón no se llega a 4. Es más fuerte que Euclides, porque también dice que los primos no son "demasiado escasos".

(d) Huecos arbitrariamente grandes. Los n números consecutivos $(n+1)! + 2, \ldots, (n+1)! + n + 1$ son todos compuestos. Con $n = 5$: los cinco números del 722 al 726 son compuestos. Hay infinitos primos y sin embargo hay desiertos tan largos como se quiera.

Pausa para pensar

Un error común es pensar que la demostración dice que $N = p_1 \cdots p_n + 1$ "siempre es primo". No dice eso, y de hecho es falso: por ejemplo $2 \cdot 3 \cdot 5 \cdot 7 \cdot 11 \cdot 13 + 1 = 30031 = 59 \cdot 509$ es compuesto. Lo que la demostración usa es más fino: N tiene *algún* factor primo, y ese factor, sea quien sea, no puede estar en la lista original. Distinguir "N es primo" de "N tiene un factor primo nuevo" es exactamente el tipo de precisión que hace que una demostración sea correcta. Vuelve a leer el argumento con esto en mente.

3.3.2. El lema de Euclides: la propiedad que hace especiales a los primos

Ahora una propiedad que parece trivial y no lo es. Si un primo p divide a un producto ab, ¿tiene que dividir a a o a b? Con $p = 5$ y $ab = 5 \cdot 6 = 30$, sí: 5 divide a 30 y en efecto divide al factor 5. Pero, ¿es esto *siempre* verdad? Fíjate en que para números *no* primos falla: 6 divide a $4 \cdot 9 = 36$, pero 6 no divide ni a 4 ni a 9. Los primos tienen algo especial, y ese algo es lo que vamos a probar. Es aquí donde Bézout cobra vida.

Lema 3.50 (Lema de Euclides)**.** *Si p es primo y $p \mid ab$, entonces $p \mid a$ o $p \mid b$.*

Demostración. Supón que $p \mid ab$ pero que p *no* divide a a; voy a demostrar que entonces forzosamente $p \mid b$. (Con esto basta: o divide a a, o —si no— divide a b.)

Como p es primo, sus únicos divisores son 1 y p. El máximo común divisor $\gcd(p, a)$ tiene que ser uno de esos dos. No puede ser p, porque eso significaría $p \mid a$, y estamos suponiendo que no. Así que $\gcd(p, a) = 1$: son coprimos. Por Bézout, existen enteros x, y con

$$px + ay = 1.$$

Multiplica toda la ecuación por b:

$$pbx + aby = b.$$

Mira los dos sumandos de la izquierda. El primero, pbx, es obviamente múltiplo de p. El segundo, aby, también es múltiplo de p, porque $p \mid ab$ por hipótesis. La suma de dos múltiplos de p es múltiplo de p; pero esa suma es b. Por lo tanto $p \mid b$. □

Ejemplo 3.51 (El lema de Euclides y dónde falla)**. (a) Uso directo.** Como $7 \mid 91 = 13 \cdot 7$, el lema garantiza que 7 divide a 13 o a 7; efectivamente divide al segundo. Y de $3 \mid 6 \cdot 10 = 60$ se sigue que 3 divide a 6, que en efecto es el caso.

(b) Sin primalidad, falla. $6 \mid 4 \cdot 9 = 36$ y sin embargo 6 no divide a 4 ni a 9. El lema es una propiedad exclusiva de los primos, y de hecho *caracteriza* a los primos entre los enteros mayores que 1.

(c) De dónde sale. Por Bézout, si $p \nmid a$ entonces $\gcd(p, a) = 1$ y existen x, y con $px + ay = 1$. Multiplicando por b: $pbx + aby = b$, y como p divide a los dos sumandos de la izquierda, divide a b. Toda la aritmética de la divisibilidad descansa en Bézout.

(d) En otros anillos. En $\mathbb{Z}[\sqrt{-5}]$ el elemento 2 es irreducible y no cumple el lema: divide a $6 = (1+\sqrt{-5})(1-\sqrt{-5})$ y no divide a ninguno de los dos factores. Por eso allí la factorización no es única, y por eso los algebristas distinguen "irreducible" de "primo", dos palabras que en $\mathbb{Z}$ significan lo mismo.

Pausa para pensar

Detente en el gesto clave: multiplicar la identidad de Bézout $px + ay = 1$ por b. Ese "multiplicar por lo que quiero que aparezca" es una técnica que vas a ver muchísimo. Partimos de un 1 construido con p y a, y al multiplicarlo por b conseguimos que b herede la divisibilidad por p. La coprimalidad de p y a —consecuencia directa de que p es primo y no divide a a— es lo que permitió tener ese 1 para empezar. Sin primalidad, no hay coprimalidad garantizada, y por eso el lema falla para el 6.

3.3.3. El teorema fundamental de la aritmética

Todos "sabemos" que cada entero se factoriza en primos de una única manera: $12 = 2^2 \cdot 3$ y no hay otra forma. Pero, ¿por qué es única? Que *exista* una factorización es fácil (sigue partiendo los factores compuestos hasta que todos sean primos). Que sea *única* es un teorema de verdad, y el lema de Euclides es justo lo que necesitamos.

Teorema 3.52 (Factorización única)**.** *Todo entero $n > 1$ se escribe como producto de primos, y esa escritura es única salvo el orden de los factores.*

Demostración. Existencia. Si n es primo, ya está (es un producto de un solo primo). Si es compuesto, se parte como $n = ab$ con a, b menores que n y mayores que 1; por el mismo razonamiento aplicado a a y a b (que son más pequeños), cada uno se factoriza en primos, y juntando las dos listas obtenemos una factorización de n. Este "seguir bajando" termina porque los enteros positivos no pueden decrecer para siempre.

Unicidad. Supón que un mismo número tuviera dos factorizaciones:

$$p_1 p_2 \cdots p_r = q_1 q_2 \cdots q_s,$$

con todos los p_i y q_j primos. Quiero mostrar que son la misma lista reordenada. Mira el primo p_1: divide al lado izquierdo, luego divide al lado derecho $q_1 q_2 \cdots q_s$. Por el lema de Euclides (aplicado repetidamente: si p_1 divide a un producto, divide a alguno de los factores), p_1 divide a *alguno* de los q_j. Pero q_j es primo, así que sus únicos divisores son 1 y él mismo; como $p_1 \neq 1$, forzosamente $p_1 = q_j$. Cancela ese par p_1 y q_j de ambos lados y repite el argumento con p_2, luego p_3, y así sucesivamente. Al final, cada primo de la izquierda se emparejó con uno idéntico de la derecha: las dos listas eran la misma. □

Ejemplo 3.53 (La factorización única y lo que permite)**. (a) Un ejemplo.** $360 = 2^3 \cdot 3^2 \cdot 5$. De ahí se lee todo: tiene $(3{+}1)(2{+}1)(1{+}1) = 24$ divisores, su suma es $(1{+}2{+}4{+}8)(1{+}3{+}9)(1{+}5) = 15 \cdot 13 \cdot 6 = 1170$, y es divisible por 8 pero no por 16.

(b) Máximo común divisor. Con $360 = 2^3 3^2 5$ y $84 = 2^2 \cdot 3 \cdot 7$, el máximo común divisor es $2^2 \cdot 3 = 12$ y el mínimo común múltiplo es $2^3 3^2 5 \cdot 7 = 2520$. Su producto es $12 \cdot 2520 = 30240 = 360 \cdot 84$, como garantiza la identidad general.

(c) Irracionalidad. Si $\sqrt{2} = p/q$ en forma irreducible, entonces $p^2 = 2q^2$, y contando el exponente del 2 en ambos lados sale un número par a la izquierda y uno impar a la derecha. La factorización única es lo que convierte ese conteo en una contradicción.

(d) Lo que no se sigue. La unicidad de la factorización no dice nada sobre la *dificultad* de encontrarla. Multiplicar dos primos de trescientas cifras es instantáneo; recuperar los factores del producto es, hasta donde se sabe, inviable. Toda la criptografía RSA vive en esa asimetría.

Observación 3.54. Vale la pena saborear que la unicidad *no es un regalo gratis*: descansa por completo en el lema de Euclides, que a su vez descansa en Bézout, que descansa en el algoritmo de la división. Hay sistemas numéricos parecidos a $\mathbb{Z}$ donde la factorización *no* es única, precisamente porque falla el análogo del lema de Euclides; el más famoso es el conjunto de números de la forma $a + b\sqrt{-5}$, donde $6 = 2 \cdot 3 = (1 + \sqrt{-5})(1 - \sqrt{-5})$ son dos factorizaciones genuinamente distintas. Que en los enteros ordinarios la factorización sea única es una propiedad que se *gana*, no que se supone; toda la teoría algebraica de números del siglo XIX nació de intentar recuperarla donde se pierde.

Regla: leer una factorización: divisores, mcd y mcm

Si $n = p_1^{a_1} p_2^{a_2} \cdots p_k^{a_k}$, la factorización contiene ya todas las respuestas:

1. **Número de divisores:** $d(n) = (a_1 + 1)(a_2 + 1) \cdots (a_k + 1)$, porque cada divisor se fabrica eligiendo, para cada primo, un exponente entre 0 y a_i.
2. **Máximo común divisor:** se toma cada primo con el exponente *menor* de los dos.
3. **Mínimo común múltiplo:** cada primo con el exponente *mayor*.
4. **Consecuencia inmediata:** $\gcd(a, b) \cdot \operatorname{mcm}(a, b) = ab$, porque menor más mayor es la suma de los dos exponentes.

Con números grandes esto es inútil (factorizar es caro y Euclides es barato); con números de examen es lo más rápido y, sobre todo, es lo que *explica* por qué las respuestas son las que son.

Ejemplo 3.55. Sea $n = 360$. Factorizamos bajando por primos: $360 = 2 \cdot 180 = 2^2 \cdot 90 = 2^3 \cdot 45 = 2^3 \cdot 3^2 \cdot 5$. Entonces

$$d(360) = (3 + 1)(2 + 1)(1 + 1) = 4 \cdot 3 \cdot 2 = 24$$

divisores. Y en efecto son $1, 2, 3, 4, 5, 6, 8, 9, 10, 12, 15, 18, 20, 24, 30, 36, 40, 45, 60, 72, 90, 120, 180, 360$: veinticuatro ✓.

Ahora $\gcd(84, 360)$ y $\operatorname{mcm}(84, 360)$ con $84 = 2^2 \cdot 3 \cdot 7$ y $360 = 2^3 \cdot 3^2 \cdot 5$:

$$\gcd = 2^{\min(2,3)} 3^{\min(1,2)} 5^{\min(0,1)} 7^{\min(1,0)} = 2^2 \cdot 3 = 12,$$

$$\operatorname{mcm} = 2^3 3^2 5^1 7^1 = 2520.$$

Comprobación con la fórmula del producto: $12 \cdot 2520 = 30240$ y $84 \cdot 360 = 30240$ ✓.

Cuidado

El exponente 0 cuenta. Al calcular el máximo común divisor hay que recordar que 84 tiene 5^0 y 360 tiene 7^0, y $\min(0, 1) = 0$: esos primos *no* entran. El error habitual es incluir en el mcd un primo que solo aparece en uno de los dos números.

Regla: decidir si un número es primo: de la división de prueba a Miller–Rabin

Preguntar "¿es primo?" es fácil; contestarlo deprisa, no tanto. Hay una escalera de métodos, cada uno con su rango de utilidad, y conviene saber en qué peldaño se está. Estos son los siete pasos.

1. **Descarta lo trivial.** Si $n \leq 1$ no es primo; si $n = 2$ o $n = 3$ sí lo es; si n es par y mayor que 2, o múltiplo de 3 y mayor que 3, tampoco. Esto elimina de golpe dos tercios de los candidatos.
2. **División de prueba hasta $\sqrt{n}$.** Basta probar divisores $d \leq \sqrt{n}$, porque si $n = ab$

con $a \le b$ entonces $a \le \sqrt{n}$. Y basta probar los *primos* hasta $\sqrt{n}$. Para números de hasta seis o siete cifras es perfectamente razonable a mano o con calculadora.

3. **Si necesitas muchos primos a la vez, usa la criba de Eratóstenes.** Escribe $2, 3, \ldots, N$, tacha los múltiplos de cada primo empezando por su cuadrado, y lo que queda son los primos. Encontrar todos los primos hasta N cuesta del orden de $N \log\log N$ operaciones, muchísimo menos que probar uno por uno.

4. **Para números grandes, pasa al test de Fermat.** Si n es primo y $\gcd(a, n) = 1$, entonces $a^{n-1} \equiv 1 \pmod{n}$. Elige una base a, calcula la potencia por exponenciación binaria y mira el resultado: *si no da* 1*, n es compuesto, con certeza absoluta.* Es una prueba de composición, no de primalidad.

5. **Desconfía si da** 1**: existen los pseudoprimos.** Que $a^{n-1} \equiv 1$ no garantiza nada. Peor aún, los *números de Carmichael* pasan el test para *toda* base coprima con ellos, así que ninguna cantidad de bases los delata.

6. **Usa Miller–Rabin, que sí distingue.** Escribe $n - 1 = 2^s d$ con d impar. Calcula $a^d \bmod n$ y elévalo al cuadrado s veces, anotando la sucesión. Si n es primo, la sucesión debe empezar en 1 o pasar por -1 antes de llegar a 1. Si llega a 1 desde un valor distinto de ± 1, has encontrado una *raíz cuadrada no trivial de* 1, lo que es imposible módulo un primo: n es compuesto.

7. **Repite con varias bases y estima la confianza.** Cada base falsa testigo tiene probabilidad $\le 1/4$ de dejar pasar un compuesto, así que con k bases independientes el error es $\le 4^{-k}$. Con $k = 20$ eso es menos de 10^{-12}, y para números de hasta 3.3×10^{14} basta comprobar las bases $2, 3, 5, 7, 11, 13$ para tener certeza *demostrada.*

Ejemplo 3.56 (Cinco números interrogados, y dos impostores desenmascarados)**. (a) División de prueba: ¿es primo** 1009**?** Como $31^2 = 961$ y $32^2 = 1024$, tenemos $\sqrt{1009} \approx 31.8$, así que basta probar los primos hasta 31: son $2, 3, 5, 7, 11, 13, 17, 19, 23, 29, 31$, once divisiones. Ninguna da resto cero (1009 es impar, sus cifras suman 10 y no es múltiplo de 3, no acaba en 0 ni en 5, $1009 = 7 \cdot 144 + 1$, $= 11 \cdot 91 + 8$, $= 13 \cdot 77 + 8$, $= 17 \cdot 59 + 6$, $= 19 \cdot 53 + 2$, $= 23 \cdot 43 + 20$, $= 29 \cdot 34 + 23$, $= 31 \cdot 32 + 17$). Por tanto 1009 es primo. Once divisiones, no mil.

(b) La criba hasta 50**.** Tachando múltiplos de 2 desde 4, de 3 desde 9, de 5 desde 25 y de 7 desde 49 —y ahí paramos, porque $11^2 = 121 > 50$— sobreviven

$$2,\ 3,\ 5,\ 7,\ 11,\ 13,\ 17,\ 19,\ 23,\ 29,\ 31,\ 37,\ 41,\ 43,\ 47,$$

quince primos. Obsérvese el detalle del paso 3: los múltiplos de p menores que p^2 ya fueron tachados por un primo menor, por eso se empieza en p^2.

(c) El pseudoprimo 341**.** Calculemos $2^{340} \bmod 341$ por exponenciación binaria. El resultado es

$$2^{340} \equiv 1 \pmod{341},$$

así que 341 *pasa* el test de Fermat en base 2. Y sin embargo $341 = 11 \cdot 31$ es compuesto: es el

menor pseudoprimo de Fermat en base 2. Basta cambiar de base para descubrirlo:

$$3^{340} \equiv 56 \not\equiv 1 \pmod{341},$$

y ese 56 es un certificado irrefutable de que 341 no es primo, aunque no diga cuáles son sus factores.

(d) El número de Carmichael 561. Aquí cambiar de base no sirve. El criterio de Korselt caracteriza a estos números: n es de Carmichael si y solo si es libre de cuadrados y $p-1 \mid n-1$ para cada primo $p \mid n$. Comprobémoslo con $561 = 3 \cdot 11 \cdot 17$, que es libre de cuadrados, y con $560 = 2^4 \cdot 5 \cdot 7$:

$$3-1 = 2 \mid 560, \qquad 11-1 = 10 \mid 560 \ (560 = 10 \cdot 56), \qquad 17-1 = 16 \mid 560 \ (560 = 16 \cdot 35).$$

Con las tres divisibilidades, el teorema chino del resto garantiza

$$a^{560} \equiv 1 \pmod{561} \quad \text{para toda } a \text{ con } \gcd(a, 561) = 1.$$

Comprobado con $a = 2, 4, 5, 7, 13$: las cinco potencias dan exactamente 1. El test de Fermat es ciego ante 561 para siempre.

(e) Miller–Rabin desenmascara a 561 en cuatro cuadrados. Escribimos $560 = 2^4 \cdot 35$, es decir $s = 4$ y $d = 35$. Con la base $a = 2$:

$$2^{35} \equiv 263, \qquad 263^2 \equiv 166, \qquad 166^2 \equiv 67, \qquad 67^2 \equiv 1 \pmod{561}.$$

Verifiquemos el último paso a mano: $67^2 = 4489$ y $561 \cdot 8 = 4488$, luego $4489 \equiv 1$. Aquí está la contradicción: hemos llegado a 1 desde 67, y $67 \not\equiv 1$ y $67 \not\equiv 560 \equiv -1$. Módulo un primo la ecuación $x^2 \equiv 1$ solo tiene las soluciones ± 1, porque $p \mid (x-1)(x+1)$ obliga a $p \mid x-1$ o $p \mid x+1$. Como 561 admite una tercera raíz cuadrada de 1, no puede ser primo. Y de propina, $\gcd(67-1, 561) = \gcd(66, 561) = 33$ da un factor: $561 = 33 \cdot 17$.

(f) Cuántos primos hay, aproximadamente. El teorema de los números primos dice que $\pi(x) \sim x/\ln x$. Para $x = 1000$ el valor exacto es $\pi(1000) = 168$, mientras que

$$\frac{1000}{\ln 1000} = \frac{1000}{6.907755} = 144.76, \qquad \mathrm{Li}(1000) = 177.61.$$

La primera aproximación se queda corta en un 14 % y la segunda se pasa en un 6 %; ambas mejoran en términos relativos al crecer x. La lectura práctica: entre los números de d cifras, aproximadamente uno de cada $2.3\,d$ es primo, así que buscar un primo de 300 cifras probando candidatos impares al azar requiere unos 350 intentos. Por eso la criptografía puede generar primos gigantes sin dificultad.

Cuidado

Distingue con cuidado los dos tipos de conclusión. Un test fallido (Fermat o Miller–Rabin) demuestra que el número es compuesto, sin ninguna probabilidad de error. Un test superado *no* demuestra que sea primo: solo acota la probabilidad de equivocarse. Por eso a los números que pasan muchas rondas se les llama “primos probables” y no “primos”. Existen tests deterministas de primalidad en tiempo polinómico (el algoritmo AKS, de

2002), pero en la práctica son mucho más lentos que veinte rondas de Miller–Rabin.

Pausa para pensar

Hay una asimetría profunda escondida aquí: decidir si un número es primo resulta ser *fácil*, y factorizarlo, *difícil*. Miller–Rabin te dice en milisegundos que un número de 600 cifras es compuesto, y a la vez nadie sabe encontrar sus factores en menos tiempo que la edad del universo. Toda la criptografía de clave pública vive en esa grieta. Y nótese que el apartado (e) nos regaló un factor por accidente: cuando Miller–Rabin encuentra una raíz cuadrada no trivial de 1, el máximo común divisor la convierte en una factorización. Es la idea que, llevada al límite, produce el algoritmo de Shor y explica por qué un ordenador cuántico rompería RSA.

3.4. Aritmética modular: las matemáticas del reloj

Idea clave

Intuición primero. La aritmética modular es la aritmética que ya usas sin darte cuenta cada vez que miras un reloj: si son las 10 y pasan 5 horas, no son las 15, son las 3. Trabajar módulo n es quedarse solo con el resto al dividir entre n y declarar iguales a todos los números que dejan el mismo resto. Lo notable —y lo que hace que esto sea una teoría y no un juego— es que sumar, restar y multiplicar *siguen funcionando* después de ese aplastamiento: puedes reducir antes o después de operar y da lo mismo.
Eso convierte problemas sobre números astronómicos en cuentas pequeñas: para saber la última cifra de 7^{100} no hace falta calcular 7^{100}, basta con trabajar módulo 10. Pero hay una operación que *no* sobrevive intacta, y ahí está toda la sutileza: la división. Módulo 12, el número 3 no tiene inverso, porque 3 y 12 comparten un factor. La condición exacta para poder dividir entre a es que $\gcd(a, n) = 1$, y el inverso se calcula, precisamente, con la identidad de Bézout de la primera sección. Cuando n es primo todos los no nulos tienen inverso, y entonces $\mathbb{Z}_n$ es un cuerpo, que es la conexión directa con el capítulo 25.

Aquí empieza la parte más divertida y la más útil. La idea nace de algo que ya sabes hacer: leer un reloj. Si son las 10 y pasan 5 horas, no son las 15: son las 3, porque el reloj "da la vuelta" al llegar a 12. Acabas de hacer aritmética modular. Vamos a formalizar ese "dar la vuelta".

Definición 3.57. Fijamos un entero $m > 0$, al que llamamos el *módulo*. Decimos que dos enteros a y b son *congruentes módulo* m, y escribimos

$$a \equiv b \pmod{m},$$

si m divide a su diferencia $a - b$. Equivalentemente: a y b dejan el mismo resto al dividirse entre m.

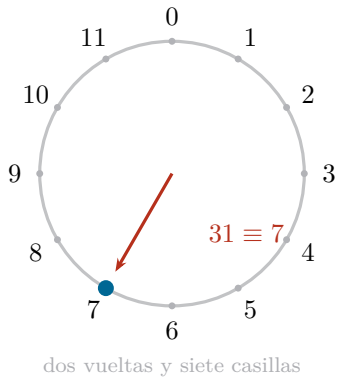


Figura 3.1: Trabajar módulo 12 es contar en un reloj: 31 y 7 ocupan la misma casilla porque difieren en dos vueltas completas.

Ejemplo 3.58 (Congruencias en la vida diaria)**. (a) El reloj.** Trabajar módulo 12 es lo que hace que "cinco horas después de las diez" sean las tres: $10+5=15\equiv 3 \pmod{12}$. Módulo 7 se calculan los días de la semana, y módulo 2 la paridad.

(b) Criterios de divisibilidad. Como $10\equiv 1 \pmod 9$, toda potencia de diez es $\equiv 1$, y por tanto un número es congruente con la suma de sus cifras módulo 9. Para 4728: $4+7+2+8=21\equiv 3$, luego $4728\equiv 3 \pmod 9$ y no es divisible por 9. Como $10\equiv -1 \pmod{11}$, el criterio del 11 alterna signos.

(c) Últimas cifras. La última cifra de 7^{100} se obtiene módulo 10: las potencias de 7 ciclan $7,9,3,1$ con período 4, y como $100\equiv 0 \pmod 4$, la respuesta es 1. Sin congruencias habría que calcular un número de ochenta y cinco cifras.

(d) El dígito de control. El DNI español añade una letra determinada por el número módulo 23; el ISBN y el IBAN usan esquemas parecidos módulo 11 y 97. Detectan errores de una cifra y transposiciones, precisamente porque la congruencia cambia al alterar cualquier dígito.

Por ejemplo, $17\equiv 5 \pmod{12}$ (ambos dejan resto 5 al dividir entre 12; su diferencia 12 es múltiplo de 12), y $10+5=15\equiv 3 \pmod{12}$, que es justo lo que dice el reloj. Trabajar módulo m es agrupar todos los enteros en m "cajones" según su resto: módulo 12 hay doce cajones, el de los que dejan resto 0, el de resto 1, etc.

Lo maravilloso es que las congruencias se comportan casi como igualdades: se pueden sumar, restar y multiplicar libremente.

Proposición 3.59. *Si* $a\equiv b \pmod m$ *y* $c\equiv d \pmod m$*, entonces*

$$a+c\equiv b+d \pmod m \qquad y \qquad ac\equiv bd \pmod m.$$

Demostración. Que $a\equiv b$ significa $m\mid (a-b)$, digamos $a-b=mk$; análogamente $c-d=m\ell$. Para la suma: $(a+c)-(b+d)=(a-b)+(c-d)=mk+m\ell=m(k+\ell)$, múltiplo de m. Para el producto, un pequeño truco de sumar y restar lo mismo:

$$ac-bd=ac-bc+bc-bd=c(a-b)+b(c-d)=c\,mk+b\,m\ell=m(ck+b\ell),$$

múltiplo de m. Listo. □

Ejemplo 3.60 (Lo que se puede y lo que no se puede hacer con congruencias)**. (a) Sumar y multiplicar, sí.** De $17 \equiv 2$ y $23 \equiv 3 \pmod 5$ se sigue $17+23 = 40 \equiv 0$ y $17 \cdot 23 = 391 \equiv 6 \equiv 1 \pmod 5$. En efecto, $391 = 78 \cdot 5 + 1$.

(b) Elevar a potencias, sí. De $a \equiv b$ se sigue $a^k \equiv b^k$ para todo $k \geq 1$, por inducción a partir del producto. Así, $3^{20} \equiv (3^4)^5 \equiv 1^5 = 1 \pmod 5$.

(c) Dividir, no siempre. De $6 \equiv 0 \pmod 6$ y $2 \cdot 3 \equiv 2 \cdot 0 \pmod 6$ no se puede cancelar el 2: daría $3 \equiv 0 \pmod 6$, que es falso. La cancelación de c sólo es válida si $\gcd(c, m) = 1$, o pasando al módulo $m/\gcd(c, m)$.

(d) Cambiar el exponente, tampoco. De $4 \equiv 1 \pmod 3$ *no* se sigue $2^4 \equiv 2^1 \pmod 3$ por ese motivo; los exponentes viven módulo el orden del grupo, no módulo m. Es el error más frecuente al usar el pequeño teorema de Fermat: la base va módulo p y el exponente módulo $p - 1$.

Esto es lo que nos deja calcular con restos como si fueran números pequeños. Por ejemplo, ¿cuál es el último dígito de $7 \cdot 7 \cdot 7 = 343$? El "último dígito" es el resto módulo 10. Trabajando módulo 10: $7 \cdot 7 = 49 \equiv 9$, y $9 \cdot 7 = 63 \equiv 3$. El último dígito es 3, y no necesitamos calcular 343 para saberlo. Con potencias gigantescas esto será una salvación.

3.4.1. Dividir es más delicado: inversos módulo m

Sumar, restar y multiplicar módulo m funcionan sin sorpresas. Dividir es otra historia, y entenderlo bien es la clave de todo lo que viene. "Dividir entre a" significa multiplicar por un *inverso* de a: un número a^{-1} tal que $a \cdot a^{-1} \equiv 1 \pmod m$. ¿Cuándo existe?

Probemos a mano. Módulo 7, busquemos el inverso de 3: queremos $3x \equiv 1 \pmod 7$. Probando, $3 \cdot 5 = 15 = 14 + 1 \equiv 1$: el inverso de 3 es 5. Ahora módulo 6, busquemos el inverso de 3: queremos $3x \equiv 1 \pmod 6$. Pero $3x$ siempre es múltiplo de 3, y los múltiplos de 3 módulo 6 son solo 0 y 3: nunca dan 1. ¡El 3 no tiene inverso módulo 6! La diferencia entre los dos casos es que $\gcd(3, 7) = 1$ pero $\gcd(3, 6) = 3$. Esto no es casualidad, y ya tienes la herramienta para probarlo.

Proposición 3.61. *El entero a tiene inverso módulo m si y solo si* $\gcd(a, m) = 1$.

Demostración. Supón primero que $\gcd(a, m) = 1$. Por Bézout —nuestra vieja llave— existen enteros x, y con $ax + my = 1$. Leamos esta ecuación módulo m: el término my es múltiplo de m, así que es $\equiv 0$, y nos queda

$$ax \equiv 1 \pmod m.$$

Ese x es el inverso de a. Recíprocamente, si a tiene inverso, digamos $ax \equiv 1 \pmod m$, entonces $ax - 1$ es múltiplo de m, digamos $ax - 1 = m(-y)$, o sea $ax + my = 1$; cualquier divisor común de a y m divide al lado izquierdo, luego divide a 1, luego vale 1: son coprimos. □

Pausa para pensar

Fíjate en que la demostración no solo dice *que* el inverso existe: dice *cómo encontrarlo*. Es exactamente el x de la identidad de Bézout, que a su vez sale de deshacer el algoritmo de Euclides. Así que "dividir módulo m" es, operativamente, correr el algoritmo de Euclides. Cuando tu computadora calcula un inverso modular para cifrar un mensaje, está haciendo justo lo que hicimos con 48 y 18 al principio del capítulo. Todo está conectado, pero esta vez la conexión la construimos nosotros paso a paso, no te la anuncié.

Esto también explica cuándo se puede resolver una ecuación lineal $ax \equiv b \pmod m$: si a tiene inverso (es decir, $\gcd(a, m) = 1$), multiplica ambos lados por a^{-1} y listo, $x \equiv a^{-1}b$. Y cuando $\gcd(a, m) = d > 1$, la ecuación tiene solución solo si $d \mid b$, y en ese caso hay exactamente d soluciones distintas módulo m (un fenómeno que vale la pena ver en un ejemplo de los ejercicios).

Regla: calcular a^{-1} módulo m

1. Comprueba $\gcd(a, m) = 1$. Si no lo es, *no hay inverso*, y no se trata de buscar mejor.
2. Corre el algoritmo de Euclides con (m, a) y desándalo hasta obtener $ax + my = 1$.
3. El inverso es x. Redúcelo al rango $[0, m)$ sumando o restando múltiplos de m.
4. Verifica: multiplica $a{\cdot}x$ y comprueba que el resto módulo m es 1. Es una comprobación de dos segundos que cierra el 90 % de los errores.

Ejemplo 3.62. Busquemos 7^{-1} módulo 26. Primero, $\gcd(7, 26) = 1$, así que existe. Euclides:

$$\begin{aligned} 26 &= 3 \cdot 7 + 5 \\ 7 &= 1 \cdot 5 + 2 \\ 5 &= 2 \cdot 2 + 1 \\ 2 &= 2 \cdot 1 + 0. \end{aligned}$$

Desandando desde la penúltima línea:

$$\begin{aligned} 1 &= 5 - 2 \cdot 2 = 5 - 2\,(7 - 5) = 3 \cdot 5 - 2 \cdot 7 \\ &= 3\,(26 - 3 \cdot 7) - 2 \cdot 7 = 3 \cdot 26 - 11 \cdot 7. \end{aligned}$$

Luego $7{\cdot}(-11) \equiv 1 \pmod{26}$, y $-11 \equiv 15$. El inverso es $\boxed{15}$. Verificación: $7{\cdot}15 = 105 = 4{\cdot}26{+}1$ ✓.

Ejemplo 3.63. Resolvamos ahora la ecuación $7x \equiv 3 \pmod{26}$. Con el inverso ya calculado,

$$x \equiv 15 \cdot 3 = 45 \equiv 45 - 26 = 19 \pmod{26}.$$

Verificación: $7 \cdot 19 = 133 = 5 \cdot 26 + 3$ ✓. Hay una única solución módulo 26, como debe ser cuando la base es invertible.

Compárese con $8x \equiv 12 \pmod{20}$, donde $d = \gcd(8, 20) = 4$. Como $4 \mid 12$, hay solución, y el teorema anuncia *cuatro*. Dividiendo todo entre 4 (incluido el módulo) queda $2x \equiv 3 \pmod 5$, y como $2^{-1} \equiv 3$ módulo 5, sale $x \equiv 9 \equiv 4 \pmod 5$. Las cuatro soluciones módulo 20 son entonces

$$x = 4,\ 9,\ 14,\ 19.$$

Comprobación de una de ellas: $8 \cdot 14 = 112 = 5 \cdot 20 + 12$ ✓. Nótese que al dividir hubo que dividir *también el módulo*: ese es el paso que casi todo el mundo olvida.

Regla: resolver una congruencia lineal $ax \equiv b \pmod m$

Esta es la ecuación de primer grado de la aritmética modular, y su respuesta puede ser ninguna solución, una, o varias. Todo lo decide el máximo común divisor.

1. **Calcular** $d = \gcd(a, m)$ con el algoritmo de Euclides. Este número manda sobre todo lo demás.
2. **Si $d \nmid b$, no hay solución.** Y la razón es transparente: $ax - mk = b$ tiene el lado izquierdo siempre múltiplo de d, así que si b no lo es, la igualdad es imposible. Se responde "sin solución" y se termina.
3. **Si $d \mid b$, hay exactamente d soluciones módulo m.** No una: exactamente d, y conviene decirlo así porque el error más común es dar solo una de ellas.
4. **Dividir todo entre d:** la congruencia $ax \equiv b \pmod m$ se convierte en $\frac{a}{d}x \equiv \frac{b}{d} \pmod{\frac{m}{d}}$, donde ahora $\gcd\left(\frac{a}{d}, \frac{m}{d}\right) = 1$. Ojo: *el módulo también se divide*. Olvidarlo es el error clásico del tema.
5. **Invertir el coeficiente** con el algoritmo de Euclides extendido, y despejar: $x \equiv \left(\frac{a}{d}\right)^{-1} \frac{b}{d} \pmod{\frac{m}{d}}$. Esa es *una* solución.
6. **Desplegar las d soluciones módulo m** sumando repetidamente $\frac{m}{d}$: si x_0 es la solución módulo $\frac{m}{d}$, las soluciones módulo m son
$$x_0, \quad x_0 + \frac{m}{d}, \quad x_0 + 2\frac{m}{d}, \quad \ldots, \quad x_0 + (d-1)\frac{m}{d}.$$
7. **Comprobar sustituyendo.** Cuesta diez segundos y detecta cualquier error de inverso.

Ejemplo 3.64 (Los tres desenlaces posibles)**. (a) Solución única:** $7x \equiv 3 \pmod{26}$**.** Como $\gcd(7, 26) = 1$, hay exactamente una solución. Buscamos 7^{-1} módulo 26: por tanteo o por Euclides extendido, $7 \cdot 15 = 105 = 4 \cdot 26 + 1$, luego $7^{-1} \equiv 15$. Entonces

$$x \equiv 15 \cdot 3 = 45 \equiv 19 \pmod{26}.$$

Comprobación: $7 \cdot 19 = 133 = 5 \cdot 26 + 3$. Correcto. (Esta es exactamente la cuenta que hay que hacer para descifrar un cifrado afín en un alfabeto de 26 letras.)

(b) Tres soluciones: $12x \equiv 9 \pmod{15}$**.** Aquí $d = \gcd(12, 15) = 3$, y $3 \mid 9$: habrá tres soluciones. Dividiendo todo entre 3, *módulo incluido*:

$$4x \equiv 3 \pmod{5}.$$

Como $4 \cdot 4 = 16 \equiv 1$, el inverso de 4 es 4, y $x \equiv 4 \cdot 3 = 12 \equiv 2 \pmod{5}$. Desplegando módulo 15 con saltos de $\frac{15}{3} = 5$:

$$x \equiv 2,\ 7,\ 12 \pmod{15}.$$

Comprobación de las tres: $12 \cdot 2 = 24 \equiv 9$, $12 \cdot 7 = 84 = 75 + 9 \equiv 9$, $12 \cdot 12 = 144 = 135 + 9 \equiv 9$. Las tres valen. Quedarse solo con $x = 2$ habría sido responder un tercio del ejercicio.

(c) Sin solución: $6x \equiv 4 \pmod{9}$**.** Ahora $d = \gcd(6, 9) = 3$ y $3 \nmid 4$: no hay solución. Se puede ver a mano recorriendo $x = 0, \ldots, 8$: los valores de $6x$ módulo 9 son solo $0, 6, 3$, y el 4 no aparece nunca. El conjunto de valores alcanzables es siempre el de los múltiplos de d, ni uno más.

Cuidado

Módulo un número compuesto se pierde una regla que en los enteros parece intocable: de $ac \equiv bc$ *no* se sigue $a \equiv b$. Por ejemplo $2 \cdot 4 \equiv 2 \cdot 1 \pmod{6}$ (ambos dan 2) y sin embargo $4 \not\equiv 1 \pmod{6}$. Solo se puede cancelar c cuando $\gcd(c, m) = 1$; si no, hay que cancelar también en el módulo: de $ac \equiv bc \pmod{m}$ se deduce $a \equiv b \pmod{m/\gcd(c, m)}$. Los divisores de cero son la raíz del problema: módulo 6, el 2 y el 3 se multiplican y dan 0 sin ser ninguno cero.

3.4.2. Las tres reglas de oro

Todo lo que sigue en el capítulo se apoya en un hecho: las congruencias se llevan bien con la suma, la resta y el producto. Es decir, se puede reducir *antes* de operar, no solo después. Eso es lo que convierte a la aritmética modular en una herramienta y no en una curiosidad.

Teorema 3.65 (Compatibilidad de las operaciones)**.** *Si* $a \equiv b \pmod{m}$ *y* $c \equiv d \pmod{m}$*, entonces*

$$a + c \equiv b + d \pmod{m}, \qquad a - c \equiv b - d \pmod{m}, \qquad a\,c \equiv b\,d \pmod{m}.$$

En consecuencia, $a^n \equiv b^n \pmod{m}$ *para todo* $n \in \mathbb{N}$*.*

Demostración. Por hipótesis $a - b = km$ y $c - d = \ell m$ para ciertos $k, \ell \in \mathbb{Z}$. Entonces $(a + c) - (b + d) = (k + \ell)m$, lo que prueba el primer caso, y el segundo es idéntico con $k - \ell$. Para el producto,

$$ac - bd = ac - bc + bc - bd = c(a - b) + b(c - d) = (ck + b\ell)m,$$

que es múltiplo de m. La afirmación sobre potencias se obtiene aplicando el producto $n - 1$ veces (formalmente, por inducción sobre n). □

Idea clave

La demostración cabe en tres líneas, pero el contenido es enorme: significa que *podemos tirar la información que no interesa en cualquier momento del cálculo*. Para saber en qué acaba 7^{100} no hace falta calcular 7^{100} (un número de 85 cifras): basta arrastrar el resto paso a paso. Es la misma filosofía que la notación científica, que tira las cifras pequeñas, pero al revés: aquí tiramos las grandes.

Regla: potencias grandes módulo m: búsqueda del ciclo

Para calcular a^n mód m:

1. Calcule $a^1, a^2, a^3, \ldots$ módulo m, reduciendo en cada paso.
2. Los restos se repiten tarde o temprano: hay un ciclo de cierta longitud L.
3. Reduzca el exponente: si $a^L \equiv 1$, entonces $a^n \equiv a^{\,n \bmod L}$.
4. Ojo: si n mód $L = 0$, el resultado es $a^L \equiv 1$, no a^0 mal interpretado.

Ejemplo 3.66 (¿En qué cifra termina 7^{100}?)**.** Terminar en cierta cifra es lo mismo que ser congruente con ella módulo 10. Calculamos:

$$7^1 \equiv 7, \quad 7^2 \equiv 49 \equiv 9, \quad 7^3 \equiv 63 \equiv 3, \quad 7^4 \equiv 21 \equiv 1 \pmod{10}.$$

El ciclo tiene longitud 4. Como $100 = 4 \cdot 25$, es decir $100 \equiv 0 \pmod 4$, resulta $7^{100} \equiv (7^4)^{25} \equiv 1^{25} = 1 \pmod{10}$. 7^{100} **termina en 1.**

Ejemplo 3.67 (3^{2026} mód 7)**.** $3^1 \equiv 3$, $3^2 \equiv 2$, $3^3 \equiv 6$, $3^4 \equiv 4$, $3^5 \equiv 5$, $3^6 \equiv 1 \pmod 7$: ciclo de longitud 6. Ahora $2026 = 6 \cdot 337 + 4$, o sea $2026 \equiv 4 \pmod 6$, luego

$$3^{2026} \equiv 3^4 \equiv 4 \pmod 7.$$

Fíjese en el detalle fino: el exponente se reduce módulo 6, no módulo 7. Los módulos de la base y del exponente son distintos, y confundirlos es el error número uno de este tema.

Ejemplo 3.68 (Las tres últimas cifras de 2^{1000})**.** Trabajamos módulo 1000. Note que $2^{10} = 1024 \equiv 24$. Entonces

$$2^{1000} = (2^{10})^{100} \equiv 24^{100} \pmod{1000},$$

y elevando por pasos ($24^2 = 576$, $24^4 \equiv 576^2 = 331776 \equiv 776, \ldots$) se llega a $2^{1000} \equiv 376$ (mód 1000). Las tres últimas cifras de 2^{1000} son **376**. El número completo tiene 302 cifras y ninguna calculadora de bolsillo lo muestra entero; la aritmética modular lo resuelve con lápiz.

Cuidado

[La división *no* está en la lista] Las tres reglas de oro son suma, resta y producto. **Dividir no vale.** Por ejemplo,

$$2 \cdot 3 = 6 \equiv 16 = 2 \cdot 8 \pmod{10},$$

y sin embargo $3 \not\equiv 8 \pmod{10}$: no se puede "cancelar el 2". La razón es que 2 y 10 comparten el factor 2. Cancelar solo es lícito cuando el factor que se cancela es primo con el módulo, y a eso le dedicamos la sección 3.8.

3.5. Los criterios de divisibilidad, por fin demostrados

En la escuela primaria se aprenden de memoria: "un número es divisible entre 9 si la suma de sus cifras lo es". Nadie explica por qué. La razón cabe en una línea de aritmética modular, y es esta: $10 \equiv 1 \pmod{9}$.

Teorema 3.69 (Criterio del 9 y del 3)**.** *Un entero es divisible entre 9 (resp. entre 3) si y solo si la suma de sus cifras es divisible entre 9 (resp. entre 3). Más aún, un número y la suma de sus cifras son congruentes módulo 9.*

Demostración. Escriba $N = c_k 10^k + \cdots + c_1 10 + c_0$. Como $10 \equiv 1 \pmod{9}$, el teorema 3.65 da $10^j \equiv 1^j = 1 \pmod{9}$ para todo j. Por tanto

$$N \equiv c_k + \cdots + c_1 + c_0 \pmod{9}.$$

Un número es divisible entre 9 exactamente cuando es congruente con 0, de donde se sigue el enunciado. Para el 3 se razona igual, porque también $10 \equiv 1 \pmod{3}$. □

Teorema 3.70 (Criterio del 11)**.** *Un entero es divisible entre 11 si y solo si la suma alternada de sus cifras (empezando por la de las unidades con signo +) es divisible entre 11.*

Demostración. Ahora $10 \equiv -1 \pmod{11}$, luego $10^j \equiv (-1)^j$, y así $N \equiv c_0 - c_1 + c_2 - c_3 + \cdots \pmod{11}$. □

Ejemplo 3.71. $N = 918\,082$. Suma alternada desde las unidades: $2 - 8 + 0 - 8 + 1 - 9 = -22$, que es múltiplo de 11. Y en efecto $918\,082 = 11 \cdot 83\,462$.

Divisor	Criterio	Por qué
2	última cifra par	$10 \equiv 0 \pmod{2}$
3	suma de cifras divisible entre 3	$10 \equiv 1 \pmod{3}$
4	las dos últimas cifras forman múltiplo de 4	$100 \equiv 0 \pmod{4}$
5	última cifra 0 o 5	$10 \equiv 0 \pmod{5}$
6	divisible entre 2 y entre 3	$6 = 2 \cdot 3$, primos entre sí
8	las tres últimas cifras forman múltiplo de 8	$1000 \equiv 0 \pmod{8}$
9	suma de cifras divisible entre 9	$10 \equiv 1 \pmod{9}$
10	última cifra 0	$10 \equiv 0 \pmod{10}$
11	suma alternada divisible entre 11	$10 \equiv -1 \pmod{11}$

Observación 3.72. La tabla revela la lógica oculta: los criterios "de las últimas cifras" (2, 4, 5, 8, 10) salen de que alguna potencia de 10 sea $\equiv 0$, y los criterios "de suma de cifras" (3, 9, 11)

salen de que 10 sea $\equiv \pm 1$. No hay un criterio cómodo para el 7 porque $10 \equiv 3 \pmod 7$ y las potencias de 3 módulo 7 recorren un ciclo largo: $3, 2, 6, 4, 5, 1$. Se puede fabricar un criterio con esos pesos, pero es más trabajo que dividir.

3.5.1. La prueba del nueve

Regla: la prueba del nueve

Para comprobar una multiplicación $a \cdot b = c$:

1. Reduzca a, b y c módulo 9 sumando sus cifras (repita hasta obtener una cifra).
2. Multiplique los dos primeros restos y redúzcalo módulo 9.
3. Si el resultado no coincide con el resto de c, **la multiplicación está mal**.

Ejemplo 3.73. $4321 \times 1234 = 5\,332\,114$. Restos módulo 9: $4+3+2+1 = 10 \to 1$; $1+2+3+4 = 10 \to 1$; producto $1 \cdot 1 = 1$. Y del resultado: $5+3+3+2+1+1+4 = 19 \to 1$. Coinciden: la operación supera la prueba. Si alguien hubiera escrito $5\,331\,414$, la suma daría $5+3+3+1+4+1+4 = 21 \to 3 \neq 1$ y el error quedaría detectado.

Cuidado

La prueba del nueve detecta errores, pero **no garantiza que el resultado sea correcto**. Si usted intercambia dos cifras del resultado, la suma no cambia y la prueba pasa tan campante. Tampoco detecta un error de exactamente 9 unidades. Es un filtro barato, no un certificado.

3.6. Calendarios, ciclos y relojes de todos los tamaños

Ejemplo 3.74. Hoy es lunes. ¿Qué día será dentro de 1000 días? Como $1000 = 142 \cdot 7 + 6$, hay que avanzar 6 días: domingo.

Pausa para pensar

Los años bisiestos siguen una regla que es pura aritmética modular: un año es bisiesto si es múltiplo de 4, *salvo* que sea múltiplo de 100, *salvo* que sea múltiplo de 400. Así, 1900 no fue bisiesto y 2000 sí lo fue. ¿Cuántos días tiene un ciclo completo de 400 años? Son $400 \cdot 365 + 97 = 146\,097$ días. Y ahora la sorpresa: $146\,097 = 7 \cdot 20\,871$ es múltiplo exacto de 7. Por eso el calendario gregoriano se repite idéntico, día de la semana incluido, cada 400 años. Su cumpleaños del año 2426 caerá en el mismo día de la semana que el de 2026.

3.7. Dígitos de control: la aritmética modular que usted lleva en la cartera

Cada vez que se teclea el número de una tarjeta, el código de un libro o el código de barras de un producto, un ordenador comprueba en microsegundos si el número "tiene sentido". No consulta ninguna base de datos: hace una cuenta módulo algo.

Regla: ISBN-10

Un ISBN de diez cifras $c_1c_2\dots c_{10}$ es válido si

$$10c_1 + 9c_2 + 8c_3 + \cdots + 2c_9 + 1\,c_{10} \equiv 0 \pmod{11}.$$

El último dígito c_{10} es el de control y se elige para que la cuenta salga; si tiene que valer 10, se escribe **X**.

Ejemplo 3.75. ISBN 0-306-40615-2. La suma ponderada es

$$10{\cdot}0+9{\cdot}3+8{\cdot}0+7{\cdot}6+6{\cdot}4+5{\cdot}0+4{\cdot}6+3{\cdot}1+2{\cdot}5+1{\cdot}2 = 0+27+0+42+24+0+24+3+10+2 = 132,$$

y $132 = 12 \cdot 11$, luego el ISBN es válido.

Regla: EAN-13 (el código de barras de cualquier producto)

Las trece cifras $c_1 \dots c_{13}$ cumplen

$$c_1 + 3c_2 + c_3 + 3c_4 + \cdots + 3c_{12} + c_{13} \equiv 0 \pmod{10},$$

es decir, se alternan los pesos 1 y 3 empezando por 1.

Ejemplo 3.76. 9780306406157: la suma es $9+21+8+0+3+0+6+12+0+18+1+15+7 = 100 \equiv 0 \pmod{10}$. Válido.

Observación 3.77 (Por qué 11 es mejor que 10). Supongamos que al teclear intercambiamos dos cifras contiguas c_i y c_{i+1}. En el EAN-13 la suma cambia en $2(c_i - c_{i+1})$, que es $\equiv 0 \pmod{10}$ siempre que las dos cifras difieran en 5. Por eso 9780306401657 (hemos intercambiado el 6 y el 1) *pasa* el control siendo un código equivocado. En el ISBN-10, en cambio, el cambio es $(c_i - c_{i+1})$ multiplicado por la diferencia de pesos, que nunca es múltiplo de 11 salvo que las cifras sean iguales: **el módulo primo detecta todas las transposiciones**. Ese es el motivo profundo por el que los sistemas serios eligen módulos primos, y una primera pista de que los primos no son una curiosidad sino una herramienta de ingeniería.

Cuidado

La X del ISBN no es una letra decorativa ni un "desconocido": es el dígito 10 escrito con la cifra romana, precisamente porque el módulo 11 admite once restos posibles y el sistema decimal solo tiene diez símbolos. Si un formulario le rechaza una X, el formulario

está mal programado, no su libro.

3.8. Ecuaciones de congruencia

Resolver $ax \equiv b \pmod{m}$ es el análogo modular de despejar una x. La diferencia con el álgebra escolar es que aquí puede haber *ninguna* solución o *varias*, y el número exacto lo decide el máximo común divisor.

Teorema 3.78 (Congruencia lineal)**.** *Sea $d = \gcd(a, m)$. La congruencia $ax \equiv b \pmod{m}$ tiene solución si y solo si $d \mid b$. Cuando la tiene, hay exactamente d soluciones distintas módulo m.*

Idea de la demostración. $ax \equiv b \pmod{m}$ equivale a que existan x, k con $ax - km = b$; es decir, a que b sea combinación entera de a y m. Por la identidad de Bézout (capítulo 3) el conjunto de esas combinaciones es exactamente el de los múltiplos de d. El recuento de soluciones sale de que, fijada una x_0, todas las demás son $x_0 + t\,(m/d)$ con $t = 0, 1, \ldots, d-1$. □

Ejemplo 3.79.

- $6x \equiv 4 \pmod{10}$: aquí $d = \gcd(6, 10) = 2$ y $2 \mid 4$, luego hay 2 soluciones. Probando: $x = 4$ y $x = 9$.
- $3x \equiv 6 \pmod{9}$: $d = 3$ y $3 \mid 6$, luego 3 soluciones: $x = 2, 5, 8$.
- $3x \equiv 5 \pmod{9}$: $d = 3$ pero $3 \nmid 5$. **No hay solución**, por mucho que se busque.
- $7x \equiv 3 \pmod{11}$: $d = 1$, solución única. Es $x = 2$, porque $14 \equiv 3$.

Corolario 3.80 (Inversos)**.** *El número a tiene inverso módulo m (esto es, existe a^{-1} con $a\,a^{-1} \equiv 1$) si y solo si $\gcd(a, m) = 1$. En ese caso el inverso es único módulo m.*

Regla: calcular un inverso con el algoritmo de Euclides al revés

Para hallar $a^{-1} \bmod m$:

1. Aplique el algoritmo de Euclides a a y m anotando cada división.
2. Si el último resto no nulo no es 1, no hay inverso: pare.
3. Recorra las divisiones hacia atrás despejando los restos, hasta escribir $1 = \alpha a + \beta m$.
4. Entonces $a^{-1} \equiv \alpha \pmod{m}$ (súmele m si sale negativo).

Ejemplo 3.81 (17^{-1} módulo 43)**.** Euclides: $43 = 2 \cdot 17 + 9$, $17 = 1 \cdot 9 + 8$, $9 = 1 \cdot 8 + 1$, $8 = 8 \cdot 1 + 0$. Hacia atrás:

$$1 = 9 - 8 = 9 - (17 - 9) = 2 \cdot 9 - 17 = 2(43 - 2 \cdot 17) - 17 = 2 \cdot 43 - 5 \cdot 17.$$

Luego $-5 \cdot 17 \equiv 1 \pmod{43}$, es decir $17^{-1} \equiv -5 \equiv 38$. Comprobación: $17 \cdot 38 = 646 = 15 \cdot 43 + 1$. ✓

3.8.1. El teorema chino del resto, en versión práctica

Teorema 3.82. *Si m y n son primos entre sí, el sistema*

$$x \equiv a \pmod{m}, \qquad x \equiv b \pmod{n}$$

tiene solución, y esta es única módulo mn.

Ejemplo 3.83. $x \equiv 3 \pmod 5$ y $x \equiv 4 \pmod 7$. De la primera, $x = 3 + 5t$. Sustituyendo en la segunda: $3 + 5t \equiv 4 \pmod 7$, o sea $5t \equiv 1 \pmod 7$. Como $5 \cdot 3 = 15 \equiv 1$, se tiene $5^{-1} \equiv 3$ y por tanto $t \equiv 3 \pmod 7$. Con $t = 3$: $x = 18$. La solución general es $x \equiv 18 \pmod{35}$.

Ejemplo 3.84 (El problema clásico, del siglo III)**.** "Tengo un montón de objetos. Si los agrupo de tres en tres, sobran 2; de cinco en cinco, sobran 3; de siete en siete, sobran 2. ¿Cuántos hay?" El sistema es $x \equiv 2\ (3)$, $x \equiv 3\ (5)$, $x \equiv 2\ (7)$, y la respuesta es $x \equiv 23 \pmod{105}$. La menor cantidad posible es 23; después vienen 128, 233, etc. El enunciado aparece en el *Sunzi Suanjing* chino, y de ahí el nombre del teorema.

Observación 3.85. El teorema chino del resto es, mirado con más altura, la afirmación de que $\mathbb{Z}/mn\mathbb{Z} \cong \mathbb{Z}/m\mathbb{Z} \times \mathbb{Z}/n\mathbb{Z}$ como anillos cuando $\gcd(m, n) = 1$. Esa frase no significa nada todavía, pero significará mucho en el capítulo 24. Guárdela.

3.9. El teorema chino del resto

Idea clave

Intuición primero. El teorema chino del resto resuelve el problema del general que quiere contar sus soldados sin contarlos: los forma de tres en tres y sobran dos, de cinco en cinco y sobran tres, de siete en siete y sobran dos; ¿cuántos hay? La respuesta es que, si los módulos no comparten factores, esos restos determinan el número de forma única dentro de un rango del tamaño del producto de los módulos.
La lectura estructural es más valiosa que el problema: el teorema dice que trabajar módulo $n = pq$ (con p y q sin factores comunes) es *exactamente lo mismo* que trabajar módulo p y módulo q en paralelo, sin perder ni ganar nada. Es una descomposición en piezas independientes, igual que descomponer un vector en coordenadas, y en el lenguaje del capítulo 25 es un isomorfismo de anillos. En la práctica es la razón de que muchos cálculos gigantes se hagan troceados por primos, incluida una de las optimizaciones habituales del RSA.

Aquí hay un acertijo con más de mil quinientos años, planteado por el matemático chino Sun Zi: *hay un número de cosas cuya cantidad desconocemos; al contarlas de tres en tres sobran 2, de cinco en cinco sobran 3, y de siete en siete sobran 2. ¿Cuántas cosas hay?* En nuestro lenguaje, buscamos un x con

$$x \equiv 2 \pmod 3, \qquad x \equiv 3 \pmod 5, \qquad x \equiv 2 \pmod 7.$$

¿Existe tal número? ¿Es único? Vamos a resolverlo despacio y a la vez descubrir el teorema general.

3.9.1. Resolviéndolo a mano

Construyamos la solución en etapas, imponiendo una condición a la vez.

La primera condición, $x \equiv 2 \pmod 3$, dice que x es uno de $2, 5, 8, 11, 14, 17, 20, 23, \ldots$

Ahora impongamos la segunda, $x \equiv 3 \pmod 5$. En vez de probar la lista entera, escribamos $x = 3 + 5t$ (que garantiza la segunda condición) y forcemos la primera: $3 + 5t \equiv 2 \pmod 3$. Como $5 \equiv 2 \pmod 3$ y $3 \equiv 0$, esto es $2t \equiv 2 \pmod 3$, o sea $t \equiv 1 \pmod 3$ (multiplicando por el inverso de 2 módulo 3, que es 2, porque $2 \cdot 2 = 4 \equiv 1$). Así $t = 1 + 3s$, y entonces

$$x = 3 + 5(1 + 3s) = 8 + 15s.$$

Comprobemos: 8 deja resto 2 al dividir entre 3 (bien) y resto 3 al dividir entre 5 (bien). Ya satisfacemos las dos primeras condiciones, y todas las soluciones son $8, 23, 38, \ldots$ (paso $15 = 3 \cdot 5$).

Falta la tercera, $x \equiv 2 \pmod 7$. Escribimos $x = 8 + 15s$ y forzamos: $8 + 15s \equiv 2 \pmod 7$. Como $8 \equiv 1$ y $15 \equiv 1 \pmod 7$, queda $1 + s \equiv 2$, es decir $s \equiv 1 \pmod 7$. Tomando $s = 1$ obtenemos

$$x = 8 + 15 \cdot 1 = 23.$$

Verifiquemos las tres de un tirón: $23 = 3 \cdot 7 + 2$ (resto 2 módulo 3 ✓), $23 = 5 \cdot 4 + 3$ (resto 3 módulo 5 ✓), $23 = 7 \cdot 3 + 2$ (resto 2 módulo 7 ✓). La respuesta al acertijo de Sun Zi es 23 cosas, y todas las demás respuestas válidas son $23,\ 23+105,\ 23+210, \ldots$ (paso $105 = 3 \cdot 5 \cdot 7$).

Observa lo que pasó: la solución existe y es única *módulo el producto* $3 \cdot 5 \cdot 7 = 105$. Esto es el teorema general.

Teorema 3.86 (Teorema chino del resto)**.** *Sean $m_1, m_2, \ldots, m_k$ enteros positivos* coprimos dos a dos *(cada par no comparte factores). Entonces, para cualesquiera restos deseados $a_1, \ldots, a_k$, el sistema*

$$x \equiv a_1 \pmod{m_1}, \quad \ldots, \quad x \equiv a_k \pmod{m_k}$$

tiene solución, y esa solución es única módulo el producto $M = m_1 m_2 \cdots m_k$.

Demostración. Doy una demostración constructiva —te dice cómo armar la solución—, que es más iluminadora que una de mera existencia. Para cada índice i, sea $M_i = M/m_i$ el producto de *todos los módulos menos el i-ésimo*. La clave: M_i es coprimo con m_i (no contiene el factor m_i, y los módulos son coprimos dos a dos), así que M_i tiene un inverso módulo m_i; llamémoslo N_i, de modo que $M_i N_i \equiv 1 \pmod{m_i}$. Ahora define

$$x = a_1 M_1 N_1 + a_2 M_2 N_2 + \cdots + a_k M_k N_k.$$

Verifiquemos que este x funciona, digamos para el módulo m_1 (los demás son idénticos). Cada término $a_i M_i N_i$ con $i \neq 1$ contiene el factor M_i, que a su vez contiene a m_1 (pues M_i es el producto de todos los módulos salvo m_i, e incluye a m_1 porque $i \neq 1$); así que esos términos son $\equiv 0 \pmod{m_1}$ y desaparecen. Solo sobrevive el primer término, y ahí $M_1 N_1 \equiv 1 \pmod{m_1}$ por construcción, dejando $x \equiv a_1 \cdot 1 = a_1 \pmod{m_1}$. Perfecto.

Unicidad módulo M. Si dos números x y x' resuelven el sistema, su diferencia $x - x'$ es divisible por cada m_i (deja resto 0 en cada módulo). Como los m_i son coprimos dos a dos, un número divisible por todos ellos es divisible por su producto M (esto usa el lema de Euclides:

si primos distintos dividen a un número, su producto también). Luego $x \equiv x'$ (mód M): hay una sola solución módulo M. □

Ejemplo 3.87 (El teorema chino del resto, resuelto)**. (a) El problema clásico.** Buscar x con $x \equiv 2$ (mód 3), $x \equiv 3$ (mód 5), $x \equiv 2$ (mód 7). Como $3, 5, 7$ son coprimos dos a dos, hay solución única módulo $M = 105$, y es $x = 23$: en efecto $23 = 7 \cdot 3 + 2$, $23 = 4 \cdot 5 + 3$, $23 = 3 \cdot 7 + 2$. Todas las soluciones son $23 + 105k$.

(b) El método constructivo. Se toman $M_i = M/m_i$, es decir $35, 21, 15$, y se resuelve $M_i y_i \equiv 1$ (mód m_i): sale $y_1 = 2$, $y_2 = 1$, $y_3 = 1$. Entonces $x = 2 \cdot 35 \cdot 2 + 3 \cdot 21 \cdot 1 + 2 \cdot 15 \cdot 1 = 140 + 63 + 30 = 233 \equiv 23$ (mód 105).

(c) Sin coprimalidad, puede no haber solución. El sistema $x \equiv 1$ (mód 4), $x \equiv 2$ (mód 6) no tiene ninguna: la primera obliga a x impar y la segunda a x par. Con módulos no coprimos hay solución si y sólo si las condiciones coinciden en cada máximo común divisor.

(d) Para qué sirve hoy. Para acelerar RSA calculando módulo p y módulo q por separado (unas cuatro veces más rápido), para la aritmética de enteros grandes por residuos y para el reparto de secretos de Asmuth–Bloom. En lenguaje moderno el enunciado es el isomorfismo de anillos $\mathbb{Z}/M \cong \mathbb{Z}/m_1 \times \cdots \times \mathbb{Z}/m_k$.

Observación 3.88. El teorema chino del resto es más que un truco para acertijos. Dice que conocer un número módulo 105 es *exactamente lo mismo* que conocerlo módulo 3, módulo 5 y módulo 7 por separado: la información se parte sin pérdida en piezas independientes, una por cada primo. Esta idea de "mirar un problema un primo a la vez y luego recombinar" es una de las estrategias más poderosas de toda la matemática, y es la razón por la que en criptografía se puede acelerar enormemente la aritmética con números enormes: se trabaja módulo cada primo pequeño y se pega el resultado al final.

Regla: resolver un sistema de congruencias por sustitución

1. Comprueba que los módulos son coprimos dos a dos. Si no, el sistema puede ser incompatible.
2. Toma la congruencia *de módulo mayor* y escribe $x = a_1 + m_1 t$. Empezar por el módulo mayor reduce el número de casos.
3. Sustituye en la siguiente congruencia y despeja t resolviendo una congruencia lineal en t.
4. Sustituye hacia atrás y repite con la siguiente. Tras k pasos, el módulo acumulado es el producto.
5. Verifica la solución en *todas* las congruencias originales, no solo en la última.

Ejemplo 3.89. Resolvamos

$$x \equiv 3 \pmod 7, \qquad x \equiv 4 \pmod 9, \qquad x \equiv 2 \pmod{11}.$$

Los módulos $7, 9, 11$ son coprimos dos a dos, así que habrá una única solución módulo $7 \cdot 9 \cdot 11 = 693$.

Paso 1. De la primera, $x = 3 + 7t$.

Paso 2. Metemos en la segunda: $3 + 7t \equiv 4 \pmod 9$, es decir $7t \equiv 1 \pmod 9$. El inverso de 7 módulo 9 es 4, porque $7 \cdot 4 = 28 = 27 + 1$. Luego $t \equiv 4 \pmod 9$, o sea $t = 4 + 9s$, y

$$x = 3 + 7(4 + 9s) = 31 + 63s.$$

Paso 3. Metemos en la tercera: $31 + 63s \equiv 2 \pmod{11}$. Reducimos: $31 \equiv 9$ y $63 \equiv 8$, luego $9 + 8s \equiv 2$, es decir $8s \equiv -7 \equiv 4 \pmod{11}$. El inverso de 8 módulo 11 es 7 (porque $8 \cdot 7 = 56 = 55 + 1$), así que $s \equiv 7 \cdot 4 = 28 \equiv 6 \pmod{11}$. Tomando $s = 6$:

$$x = 31 + 63 \cdot 6 = 31 + 378 = 409.$$

Verificación. $409 = 7 \cdot 58 + 3$ ✓; $409 = 9 \cdot 45 + 4$ ✓; $409 = 11 \cdot 37 + 2$ ✓. La solución general es $x \equiv 409 \pmod{693}$.

Cuidado

Si los módulos *no* son coprimos, el teorema no se aplica y el sistema puede no tener solución. Por ejemplo $x \equiv 1 \pmod 4$ y $x \equiv 2 \pmod 6$ es imposible: la primera obliga a que x sea impar y la segunda a que sea par. En cambio $x \equiv 1 \pmod 4$ y $x \equiv 3 \pmod 6$ sí tiene solución ($x = 9$, y en general $x \equiv 9 \pmod{12}$), porque los dos restos coinciden módulo el máximo común divisor 2. Esa es exactamente la condición de compatibilidad en el caso general: $a_i \equiv a_j \pmod{\gcd(m_i, m_j)}$.

Regla: decidir si a es un cuadrado módulo p: símbolo de Legendre y reciprocidad

La ecuación $x^2 \equiv a \pmod p$ tiene dos soluciones o ninguna (salvo el caso $a \equiv 0$). Decidir cuál de las dos cosas ocurre sin probar los p candidatos es lo que hace el símbolo de Legendre, y la ley de reciprocidad cuadrática lo convierte en un cálculo tan rápido como el algoritmo de Euclides. Estos son los pasos.

1. **Define el símbolo.** Para p primo impar y $p \nmid a$,

$$\left(\frac{a}{p}\right) = \begin{cases} +1 & \text{si } a \text{ es resto cuadrático módulo } p, \\ -1 & \text{si no lo es.} \end{cases}$$

   Exactamente la mitad de los $p - 1$ residuos no nulos son cuadrados: los $(p-1)/2$ valores de k^2 con $1 \le k \le (p-1)/2$.

2. **Reduce a módulo p y factorízalo.** El símbolo es completamente multiplicativo en el numerador:

$$\left(\frac{ab}{p}\right) = \left(\frac{a}{p}\right)\left(\frac{b}{p}\right),$$

   así que basta saber calcularlo para -1, para 2 y para cada primo impar que aparezca.

3. **Aplica las dos leyes suplementarias.**

$$\left(\frac{-1}{p}\right) = (-1)^{\frac{p-1}{2}} = \begin{cases} +1 & p \equiv 1 \ (4) \\ -1 & p \equiv 3 \ (4) \end{cases}, \qquad \left(\frac{2}{p}\right) = (-1)^{\frac{p^2-1}{8}} = \begin{cases} +1 & p \equiv \pm 1 \ (8) \\ -1 & p \equiv \pm 3 \ (8) \end{cases}.$$

La primera dice cuándo -1 es un cuadrado; la segunda, cuándo lo es 2. Conviene memorizarlas por las clases módulo 4 y módulo 8.

4. **Para primos impares $q \neq p$, usa la reciprocidad cuadrática.**

$$\left(\frac{q}{p}\right)\left(\frac{p}{q}\right) = (-1)^{\frac{p-1}{2}\cdot\frac{q-1}{2}}.$$

En palabras: los dos símbolos son iguales *salvo* que p y q sean ambos $\equiv 3 \pmod 4$, en cuyo caso son opuestos. Esto permite dar la vuelta al símbolo y reducir de nuevo, exactamente como en Euclides.

5. **Itera hasta llegar a símbolos triviales.** Alterna "reducir el numerador módulo el denominador" con "dar la vuelta". Como los números decrecen, el proceso termina, y en un número de pasos comparable al del algoritmo de Euclides.

6. **Contrasta con el criterio de Euler.** Se cumple

$$\left(\frac{a}{p}\right) \equiv a^{\frac{p-1}{2}} \pmod p,$$

y el lado derecho solo puede valer ± 1. Es más lento que la reciprocidad pero completamente mecánico, y sirve de verificación independiente.

7. **Si el símbolo es $+1$, aún falta hallar la raíz.** Saber que existe no es saber cuál es. Si $p \equiv 3 \pmod 4$ hay fórmula directa: $x \equiv a^{(p+1)/4}$. En el caso general se usa el algoritmo de Tonelli–Shanks. *No confundas la existencia con el cálculo*: es la misma distinción que en la teoría de Galois entre saber que hay raíces y saber escribirlas.

Ejemplo 3.90 (Cuatro símbolos calculados, y para qué sirven)**. (a) Los cuadrados módulo 7, a mano.** Elevamos al cuadrado $1, 2, 3$ (basta la mitad, porque k y $7-k$ dan lo mismo):

$$1^2 = 1, \qquad 2^2 = 4, \qquad 3^2 = 2 \text{ (pues } 9 - 7 = 2).$$

Luego los restos cuadráticos módulo 7 son $\{1, 2, 4\}$ y los no restos $\{3, 5, 6\}$: tres y tres, como anuncia el paso 1. Comprobación con las leyes: $7 \equiv 7 \pmod 8$, así que $\left(\frac{2}{7}\right) = +1$, y en efecto $2 = 3^2$. Y $7 \equiv 3 \pmod 4$, así que $\left(\frac{-1}{7}\right) = -1$: en efecto $-1 \equiv 6$ no está en la lista.

(b) $\left(\frac{3}{7}\right)$ por reciprocidad. Ambos primos son $\equiv 3 \pmod 4$, luego los símbolos son opuestos:

$$\left(\frac{3}{7}\right) = -\left(\frac{7}{3}\right) = -\left(\frac{1}{3}\right) = -1.$$

Coincide con la lista de (a): 3 no es un cuadrado módulo 7. Verificación por el criterio de Euler: $3^3 = 27 \equiv 6 \equiv -1 \pmod 7$. Los tres métodos concuerdan.

(c) Un cálculo de verdad: $\left(\frac{30}{103}\right)$. El primo 103 es demasiado grande para listar cuadrados, así que factorizamos el numerador: $30 = 2 \cdot 3 \cdot 5$ y

$$\left(\frac{30}{103}\right) = \left(\frac{2}{103}\right)\left(\frac{3}{103}\right)\left(\frac{5}{103}\right).$$

Primer factor. $103 = 12 \cdot 8 + 7$, luego $103 \equiv 7 \equiv -1 \pmod 8$ y $\left(\frac{2}{103}\right) = +1$.

Segundo factor. Como $103 \equiv 3 \pmod 4$ y $3 \equiv 3 \pmod 4$, hay cambio de signo:

$$\left(\frac{3}{103}\right) = -\left(\frac{103}{3}\right) = -\left(\frac{1}{3}\right) = -1.$$

Tercer factor. Como $5 \equiv 1 \pmod 4$, no hay cambio de signo:

$$\left(\frac{5}{103}\right) = \left(\frac{103}{5}\right) = \left(\frac{3}{5}\right) = -1,$$

porque los cuadrados módulo 5 son $\{1, 4\}$ y 3 no está.

Producto. $(+1)\cdot(-1)\cdot(-1) = +1$, así que 30 *sí* es un cuadrado módulo 103. Y en efecto lo es: las dos soluciones de $x^2 \equiv 30 \pmod{103}$ son

$$x = 37 \quad \text{y} \quad x = 66 = 103 - 37,$$

ya que $37^2 = 1369 = 13 \cdot 103 + 30 = 1339 + 30$. Nótese lo que hemos hecho: hemos decidido la existencia con cuatro líneas de aritmética con números de dos cifras, sin tocar 103 candidatos.

(d) Cuándo -1 es un cuadrado, y el teorema de los dos cuadrados. La primera ley suplementaria dice que $x^2 \equiv -1 \pmod p$ tiene solución exactamente cuando $p \equiv 1 \pmod 4$. Ejemplos: para $p = 13$, $5^2 = 25 = 26 - 1 \equiv -1$; para $p = 17$, $4^2 = 16 \equiv -1$; para $p = 5$, $2^2 = 4 \equiv -1$. En cambio para $p = 7, 11, 19$ no hay solución. Este hecho aparentemente modesto es la puerta del teorema de Fermat sobre sumas de dos cuadrados: un primo impar p es suma de dos cuadrados si y solo si $p \equiv 1 \pmod 4$. Compruébese:

$$5 = 1 + 4, \quad 13 = 4 + 9, \quad 17 = 1 + 16, \quad 29 = 4 + 25, \quad 37 = 1 + 36,$$

mientras que 3, 7, 11, 19 y 23 no admiten ninguna descomposición. La demostración usa precisamente la raíz de -1 que garantiza el símbolo.

Cuidado

La ley de reciprocidad se enuncia para primos impares *distintos*, y el símbolo de Legendre exige que el denominador sea primo. Si el denominador es compuesto hay que usar el símbolo de Jacobi, que se calcula con las mismas reglas pero cuya interpretación cambia: un símbolo de Jacobi igual a -1 sí garantiza que a no es un cuadrado, pero un $+1$ *no* garantiza que lo sea. El ejemplo mínimo es $\left(\frac{2}{15}\right)_J = \left(\frac{2}{3}\right)\left(\frac{2}{5}\right) = (-1)(-1) = +1$, y sin embargo 2 no es un cuadrado módulo 15, porque los cuadrados módulo 15 son $\{1, 4, 6, 9, 10\}$.

Idea clave

La ley de reciprocidad cuadrática fue la joya de Gauss, que dio ocho demostraciones distintas y la llamó *theorema aureum*. Lo asombroso de ella es que relaciona dos preguntas que no tienen ninguna razón aparente para estar conectadas: "¿es q un cuadrado módulo p?" y "¿es p un cuadrado módulo q?" viven en anillos distintos y hablan de objetos distintos. La explicación moderna es que ambas son afirmaciones sobre cómo se descompone el primo en la extensión ciclotómica $\mathbb{Q}(\zeta_p)$, y que la ley es el primer caso de la *teoría de cuerpos de clases*: la generalización de la teoría de Galois del capítulo 26 a las extensiones abelianas de cuerpos de números. Toda la teoría algebraica de números moderna es el intento de contestar "¿cuál es la reciprocidad para grados mayores que dos?".

3.10. Los teoremas de Fermat y Euler

Idea clave

Intuición primero. Estos dos teoremas dicen que elevar a potencias módulo n es una operación *periódica*: tarde o temprano vuelves al punto de partida. El pequeño teorema de Fermat da el periodo cuando el módulo es primo ($a^{p-1} \equiv 1$), y Euler lo generaliza a cualquier módulo sustituyendo $p-1$ por $\varphi(n)$, la cantidad de números menores que n que no comparten factores con él.
La consecuencia operativa es enorme y conviene verla con claridad: si el exponente vive módulo $\varphi(n)$, entonces para calcular a^k basta reducir k módulo $\varphi(n)$ primero. Un exponente de mil cifras se convierte en uno pequeño. Y esa es, literalmente, la puerta trasera del RSA: quien conoce $\varphi(n)$ puede deshacer la exponenciación; quien no lo conoce tendría que factorizar n para averiguarlo. Todo el criptosistema de la sección siguiente es este teorema más la observación de que factorizar un número de seiscientas cifras es, hoy por hoy, inviable.

Ahora llegamos a dos joyas que hacen posible calcular potencias astronómicas con lápiz y papel, y que son el motor secreto de la criptografía moderna.

3.10.1. El pequeño teorema de Fermat

Empecemos experimentando, que es como se descubren estas cosas. Tomemos el primo $p = 5$ y elevemos varios números a la cuarta potencia, mirando el resto módulo 5:

$$1^4 = 1, \quad 2^4 = 16 \equiv 1, \quad 3^4 = 81 \equiv 1, \quad 4^4 = 256 \equiv 1 \pmod 5.$$

¡Todos dan 1! No es casualidad. Fermat observó este patrón, y dice lo siguiente.

Teorema 3.91 (Pequeño teorema de Fermat)**.** *Si p es primo y a no es divisible por p, entonces*

$$a^{p-1} \equiv 1 \pmod p.$$

Demostración. Esta demostración es preciosa porque el truco parece un juego de manos y resulta ser exacto. Considera los primeros $p-1$ múltiplos de a:

$$a,\ 2a,\ 3a,\ \ldots,\ (p-1)a.$$

Afirmo que, al reducirlos módulo p, estos números son *una reordenación* de $1, 2, \ldots, p-1$. Para verlo, dos observaciones. Primero, ninguno es $\equiv 0$: si $p \mid ja$ con $1 \le j \le p-1$, por el lema de Euclides p dividiría a j o a a, pero j es demasiado chico y a no es divisible por p por hipótesis. Segundo, son todos distintos entre sí módulo p: si $ja \equiv ka$, entonces $(j-k)a \equiv 0$, y como a es invertible módulo p (es coprimo con el primo p), podemos cancelarlo para obtener $j \equiv k$. Así que tenemos $p-1$ números, todos distintos, ninguno cero, viviendo en los cajones $1, 2, \ldots, p-1$: forzosamente los ocupan todos, en algún orden.

Ahora multipliquemos todos esos números. Por un lado,

$$a \cdot 2a \cdot 3a \cdots (p-1)a = a^{p-1} \cdot \big(1 \cdot 2 \cdots (p-1)\big) = a^{p-1}(p-1)!.$$

Por otro lado, como módulo p son una reordenación de $1, 2, \ldots, p-1$, su producto es congruente con

$$1 \cdot 2 \cdots (p-1) = (p-1)!.$$

Igualando las dos cuentas módulo p:

$$a^{p-1}(p-1)! \equiv (p-1)! \pmod{p}.$$

Finalmente, $(p-1)!$ es un producto de números todos coprimos con p, así que es invertible módulo p: lo cancelamos de ambos lados y queda $a^{p-1} \equiv 1 \pmod{p}$. □

Idea clave

Qué dice esto en cristiano. El pequeño teorema de Fermat dice que elevar a $p-1$ módulo un primo devuelve siempre 1. No es magia: es el teorema de Lagrange aplicado al grupo multiplicativo $\mathbb{F}_p^*$, que tiene $p-1$ elementos, de modo que el orden de cualquier elemento divide a $p-1$. Su utilidad práctica es doble: permite reducir exponentes gigantescos a exponentes manejables, y da un test de primalidad rapidísimo, aunque no infalible.

Ejemplo 3.92 (Reducir exponentes y detectar primos)**. (a) La comprobación directa.** Con $p = 7$: $2^6 = 64 = 63 + 1 \equiv 1$, $3^6 = 729 = 728 + 1 \equiv 1$ y $5^6 = 15\,625 = 2232 \cdot 7 + 1 \equiv 1$. Todos los elementos no nulos cumplen lo mismo, como manda el teorema.

(b) Reducir un exponente enorme. Para calcular 3^{1000} mód 7: como $3^6 \equiv 1$ y $1000 = 6 \cdot 166 + 4$, queda $3^{1000} \equiv 3^4 = 81 \equiv 4 \pmod{7}$. Un exponente de mil se ha reducido a uno de cuatro sin calculadora.

(c) La generalización de Euler. Cuando el módulo no es primo se usa $a^{\varphi(m)} \equiv 1$ si $\gcd(a, m) = 1$. Para $m = 10$, $\varphi(10) = 4$, y en efecto $3^4 = 81 \equiv 1 \pmod{10}$. De aquí sale que las últimas cifras de las potencias se repiten con período 4: $3, 9, 7, 1, 3, 9, 7, 1, \ldots$

(d) El test de primalidad. Si $a^{n-1} \not\equiv 1 \pmod{n}$ para algún a, entonces n *no* es primo, y se ha demostrado sin factorizar nada. Con $n = 15$ y $a = 2$: $2^{14} = 16\,384 = 1092 \cdot 15 + 4 \equiv 4 \neq 1$, luego 15 es compuesto. Es la base del test de Miller–Rabin, que se usa a diario para generar las claves de RSA.

Cuidado

[El recíproco es falso: existen los números de Carmichael] Que $a^{n-1} \equiv 1$ para algún a, o incluso para todos los a coprimos con n, no demuestra que n sea primo. El menor contraejemplo es $561 = 3 \cdot 11 \cdot 17$, que cumple $a^{560} \equiv 1 \pmod{561}$ para todo a coprimo con él y sin embargo es compuesto. Se llaman números de Carmichael y hay infinitos, lo que se demostró en 1994. Segundo aviso: la hipótesis $p \nmid a$ es imprescindible; con $a = 7$ y $p = 7$ sale $7^6 \equiv 0$, no 1. Y tercero: para módulos compuestos hay que usar $\varphi(m)$ y no $m - 1$; confundirlos rompe todos los cálculos de RSA.

Pausa para pensar

La jugada central —"los múltiplos $a, 2a, \ldots, (p-1)a$ son los mismos residuos $1, 2, \ldots, p-1$ barajados"— merece que la pruebes tú con un caso. Toma $p = 5$, $a = 2$: los múltiplos son $2, 4, 6, 8$, que módulo 5 son $2, 4, 1, 3$: en efecto, $1, 2, 3, 4$ revueltos. Multiplicar y cancelar el factorial es el remate. Este patrón de "multiplicar por algo invertible solo baraja los residuos" reaparecerá; es la versión aritmética de que multiplicar por una matriz invertible no pierde información.

El uso práctico es espectacular. ¿Cuánto vale 2^{1000} módulo 13? Sin Fermat, es impensable a mano. Con Fermat: como 13 es primo y 2 no es múltiplo de 13, sabemos $2^{12} \equiv 1$. Dividimos el exponente entre 12: $1000 = 12 \cdot 83 + 4$. Entonces

$$2^{1000} = \left(2^{12}\right)^{83} \cdot 2^4 \equiv 1^{83} \cdot 16 \equiv 16 \equiv 3 \pmod{13}.$$

Una potencia con más de trescientos dígitos, resuelta en dos líneas. Este es exactamente el tipo de cálculo que hace tu teléfono millones de veces al día.

3.10.2. Euler generaliza a cualquier módulo

Fermat necesita que el módulo sea primo. Euler descubrió cómo hacerlo funcionar con *cualquier* módulo m, y para eso introdujo una función que cuenta cuántos números por debajo de m son coprimos con m.

Definición 3.93. La *función φ de Euler* cuenta, para cada $m \geq 1$, la cantidad de enteros en $\{1, 2, \ldots, m\}$ que son coprimos con m:

$$\varphi(m) = \#\{\, a : 1 \leq a \leq m,\ \gcd(a, m) = 1 \,\}.$$

Por ejemplo $\varphi(10) = 4$, porque de 1 a 10 los coprimos con 10 son $1, 3, 7, 9$. Si p es primo, *todos* los números de 1 a $p - 1$ son coprimos con él, así que $\varphi(p) = p - 1$: la función de Euler generaliza el "$p - 1$" de Fermat. Y para un producto de dos primos distintos p, q hay una fórmula preciosa, $\varphi(pq) = (p - 1)(q - 1)$, que probaremos porque la vamos a necesitar para RSA.

Teorema 3.94 (Euler)**.** *Si* $\gcd(a, m) = 1$, *entonces* $a^{\varphi(m)} \equiv 1 \pmod{m}$.

Demostración. La demostración es *la misma idea de Fermat*, ahora con los residuos coprimos en el papel de $1, \ldots, p-1$. Sean $r_1, r_2, \ldots, r_{\varphi(m)}$ los enteros entre 1 y m coprimos con m. Como a es coprimo con m, al multiplicar cada r_i por a y reducir módulo m obtenemos de nuevo una reordenación de la misma lista $r_1, \ldots, r_{\varphi(m)}$ (mismo argumento que antes: los ar_i siguen siendo coprimos con m, y son distintos entre sí porque a se puede cancelar). Multiplicando todos:

$$(ar_1)(ar_2)\cdots(ar_{\varphi(m)}) = a^{\varphi(m)}\, r_1 r_2 \cdots r_{\varphi(m)} \equiv r_1 r_2 \cdots r_{\varphi(m)} \pmod{m}.$$

El producto $r_1 \cdots r_{\varphi(m)}$ es coprimo con m (cada factor lo es), así que es invertible y lo cancelamos, dejando $a^{\varphi(m)} \equiv 1 \pmod{m}$. □

Un calentamiento antes del gran final. ¿Cuál es el último dígito de 7^{222}? Es 7^{222} módulo 10. Como $\gcd(7, 10) = 1$ y $\varphi(10) = 4$, Euler da $7^4 \equiv 1 \pmod{10}$. Dividimos $222 = 4 \cdot 55 + 2$, así que

$$7^{222} = \left(7^4\right)^{55} \cdot 7^2 \equiv 1 \cdot 49 \equiv 9 \pmod{10}.$$

El último dígito es 9.

Regla: reducir una potencia grande módulo m

Para calcular a^n mód m sin calcular a^n:

1. Reduce primero la base: sustituye a por a mód m.
2. Ve elevando al cuadrado y reduciendo en cada paso: a^2, $a^4 = (a^2)^2$, $a^8 = (a^4)^2$, …, sin dejar nunca que los números crezcan.
3. Escribe n en binario y multiplica solo las potencias que hagan falta.
4. Si $\gcd(a, m) = 1$, atajo: por el teorema de Euler basta reducir el *exponente* módulo $\varphi(m)$, no módulo m.

Ejemplo 3.95. Calculemos 2^{1000} mód 11. Como 11 es primo y $11 \nmid 2$, el pequeño teorema de Fermat da $2^{10} \equiv 1$. Dividimos el exponente: $1000 = 10 \cdot 100$, luego

$$2^{1000} = (2^{10})^{100} \equiv 1^{100} = 1 \pmod{11}.$$

Ahora 7^{222} mód 100. Aquí 100 no es primo, así que usamos Euler: $\varphi(100) = 100\left(1 - \frac{1}{2}\right)\left(1 - \frac{1}{5}\right) = 40$, y $\gcd(7, 100) = 1$. Reducimos el exponente módulo 40: $222 = 5 \cdot 40 + 22$, luego $7^{222} \equiv 7^{22}$. Y ahora por cuadrados sucesivos:

$$7^2 = 49, \quad 7^4 \equiv 49^2 = 2401 \equiv 1, \quad \text{luego} \quad 7^{22} = 7^{4\cdot 5+2} \equiv 1^5 \cdot 49 = 49.$$

Así que 7^{222} termina en 49.

Cuidado

El exponente se reduce módulo $\varphi(m)$, *nunca* módulo m. Véase lo que pasa si uno se confunde: para 2^{100} mód 7 lo correcto es $\varphi(7) = 6$ y $100 = 16 \cdot 6 + 4$, luego $2^{100} \equiv 2^4 = 16 \equiv 2$. Quien reduzca el exponente módulo 7 escribirá $100 \equiv 2$ y concluirá $2^{100} \equiv 2^2 = 4$, que es *falso*. Y hay una segunda condición que se olvida aún más: el teorema de Euler exige $\gcd(a, m) = 1$. Con $a = 2$ y $m = 4$ no hay nada que reducir, porque $2^n \equiv 0$ (mód 4) para todo $n \geq 2$ y la base no es invertible.

Regla: calcular el orden de un elemento y encontrar una raíz primitiva

El *orden* de a módulo m (con $\gcd(a, m) = 1$) es el menor $k \geq 1$ con $a^k \equiv 1$ (mód m). Es el dato que gobierna todos los cálculos con potencias.

1. **Comprobar que** $\gcd(a, m) = 1$**.** Si no, a no tiene orden: ninguna potencia suya será 1, porque todas comparten factor con m.
2. **Recordar la restricción clave: el orden divide a** $\varphi(m)$**.** Es el teorema de Euler más el teorema de Lagrange en el grupo $(\mathbb{Z}/m)^\times$. Así que *nunca* hay que probar $k = 1, 2, 3, \ldots$ uno por uno: basta probar los divisores de $\varphi(m)$.
3. **Listar los divisores de** $\varphi(m)$ **de menor a mayor** y calcular a^k para cada uno hasta encontrar el primero que dé 1. Con m primo, $\varphi(m) = m - 1$.
4. **Atajo para verificar si** a **es raíz primitiva** (es decir, si su orden es exactamente $\varphi(m)$): basta comprobar que $a^{\varphi(m)/p} \not\equiv 1$ para cada primo p que divida a $\varphi(m)$. Son solo dos o tres cuentas, no $\varphi(m)$.
5. **Contar cuántas raíces primitivas hay:** si existen, son exactamente $\varphi(\varphi(m))$. Y existen solo para $m = 1, 2, 4, p^k, 2p^k$ con p primo impar; módulo 8, por ejemplo, no hay ninguna.
6. **Generar el resto una vez hallada una.** Si g es raíz primitiva, las demás son g^j con $\gcd(j, \varphi(m)) = 1$, y el orden de g^j en general es $\frac{\varphi(m)}{\gcd(j, \varphi(m))}$. Con una raíz primitiva se conoce el orden de todos los elementos sin hacer más potencias.
7. **Usarlo para reducir exponentes.** Si a tiene orden k, entonces $a^n \equiv a^{n \bmod k}$. Reducir con el orden es más eficiente que reducir con $\varphi(m)$, porque el orden suele ser bastante menor.

Ejemplo 3.96 (Todo el grupo multiplicativo módulo 13)**.** Como 13 es primo, $\varphi(13) = 12$ y los órdenes posibles son los divisores de 12: $1, 2, 3, 4, 6, 12$.

(a) ¿Es 2 raíz primitiva? Los primos que dividen a 12 son 2 y 3, así que basta comprobar dos cosas: $2^{12/2} = 2^6 = 64 = 52 + 12 \equiv 12 \not\equiv 1$ y $2^{12/3} = 2^4 = 16 \equiv 3 \not\equiv 1$. Ninguna da 1: el orden de 2 es 12, es raíz primitiva. Y en efecto sus potencias recorren todo:

$$2^1, \ldots, 2^{12} \equiv 2, 4, 8, 3, 6, 12, 11, 9, 5, 10, 7, 1.$$

Ahí están los doce elementos no nulos, cada uno exactamente una vez.

(b) Los órdenes de todos los demás, sin más potencias. Escribiendo cada elemento como 2^j y aplicando $\operatorname{ord}(2^j) = 12/\gcd(j, 12)$, la tabla sale sola:

a	1	2	3	4	5	6	7	8	9	10	11	12
j con $2^j = a$	12	1	4	2	9	5	11	3	8	10	7	6
orden	1	12	3	6	4	12	12	4	3	6	12	2

Por ejemplo $3 = 2^4$ y $\gcd(4, 12) = 4$, luego su orden es $12/4 = 3$; comprobación directa: $3^2 = 9$, $3^3 = 27 \equiv 1$. Correcto.

(c) Cuántas raíces primitivas. La fórmula predice $\varphi(\varphi(13)) = \varphi(12) = 4$, y la tabla las muestra: son 2, 6, 7 y 11, exactamente los 2^j con $j \in \{1, 5, 7, 11\}$, es decir con $\gcd(j, 12) = 1$. Predicción y recuento coinciden.

(d) Para qué sirve todo esto. Calculemos 5^{1000} mód 13. Con el teorema de Euler habría que reducir 1000 módulo 12, lo que da $1000 = 83 \cdot 12 + 4$ y 5^4. Pero el orden de 5 es 4, mucho menor, y 1000 es múltiplo de 4:

$$5^{1000} = \left(5^4\right)^{250} \equiv 1^{250} = 1 \pmod{13}.$$

Sin necesidad de calcular nada más. Conocer el orden convierte una exponenciación en una división con resto.

Ejemplo 3.97 (Los criterios de divisibilidad, deducidos en vez de memorizados)**.** Todos los criterios de la escuela salen de una sola observación: para saber el resto de un número al dividirlo entre m, basta conocer los restos de las potencias 10^k módulo m.

(a) El 9 y el 3. Como $10 \equiv 1 \pmod 9$, se tiene $10^k \equiv 1$ para todo k, luego un número es congruente con *la suma de sus cifras*. Para 123456789: la suma es 45, y $45 \equiv 0 \pmod 9$, así que el número es múltiplo de 9 (y de 3, porque $3 \mid 9$). Nótese que el criterio no solo dice si divide: da el resto exacto.

(b) El 11. Ahora $10 \equiv -1 \pmod{11}$, luego $10^k \equiv (-1)^k$: el número es congruente con la suma *alternada* de sus cifras, empezando por la de las unidades con signo $+$. Para 123456789:

$$9 - 8 + 7 - 6 + 5 - 4 + 3 - 2 + 1 = 5,$$

así que $123456789 \equiv 5 \pmod{11}$, y no es múltiplo de 11.

(c) El 7, que "no tiene criterio". Sí lo tiene; lo que pasa es que el ciclo es más largo. Los restos de 10^k módulo 7 son

$$1,\ 3,\ 2,\ 6,\ 4,\ 5,\ 1,\ 3,\ 2, \ldots$$

con período 6. Ponderando las cifras de 123456789 desde las unidades con esos pesos:

$$9 \cdot 1 + 8 \cdot 3 + 7 \cdot 2 + 6 \cdot 6 + 5 \cdot 4 + 4 \cdot 5 + 3 \cdot 1 + 2 \cdot 3 + 1 \cdot 2 = 134,$$

y $134 = 19 \cdot 7 + 1$, luego $123456789 \equiv 1 \pmod 7$. La razón de que el criterio del 9 sea popular y el del 7 no es puramente decimal: $10 \equiv 1$ módulo 9 es un accidente de la base, no una propiedad del 9. En base 8, el criterio fácil sería el del 7.

Regla: resolver ecuaciones diofánticas: lineales, pitagóricas y de Pell

Una ecuación diofántica es una ecuación de la que solo se aceptan soluciones enteras. Tres familias cubren casi todo lo que se ve en un curso, y cada una tiene su método propio. Estos son los siete pasos del procedimiento general.

1. **Decide primero si hay alguna solución.** Para $ax + by = c$ el criterio es exacto: hay solución entera si y solo si $\gcd(a, b) \mid c$. Para ecuaciones no lineales, el arma habitual es mirar la ecuación módulo un número pequeño: si no hay solución módulo 4, módulo 8 o módulo 9, no la hay en los enteros.

2. **Caso lineal: halla una solución particular con Euclides extendido.** Calcula x_0, y_0 con $ax_0 + by_0 = \gcd(a, b)$ y multiplica por $c/\gcd(a, b)$.

3. **Escribe la solución general del caso lineal.** Con $d = \gcd(a, b)$, todas las soluciones son
$$x = x_0 + \frac{b}{d}\,t, \qquad y = y_0 - \frac{a}{d}\,t, \qquad t \in \mathbb{Z}.$$
Comprueba el signo: al sustituir, los dos términos añadidos deben cancelarse.

4. **Caso pitagórico: usa la parametrización, no la fuerza bruta.** Todas las ternas primitivas de $x^2 + y^2 = z^2$ (es decir, con $\gcd(x, y, z) = 1$) son
$$x = m^2 - n^2, \qquad y = 2mn, \qquad z = m^2 + n^2,$$
con $m > n \geq 1$ coprimos y de distinta paridad. Las no primitivas son múltiplos de estas.

5. **Caso de Pell: $x^2 - Ny^2 = 1$ con N no cuadrado.** Halla la *solución fundamental* (x_1, y_1), la de menor y positivo, probando $y = 1, 2, 3, \ldots$ y mirando si $Ny^2 + 1$ es un cuadrado. Si N es un cuadrado perfecto, la ecuación se factoriza y solo hay soluciones triviales.

6. **Genera todas las soluciones de Pell a partir de la fundamental.** Se obtienen de
$$x_k + y_k\sqrt{N} = \left(x_1 + y_1\sqrt{N}\right)^k,$$
o equivalentemente por la recurrencia $x_{k+1} = x_1x_k + Ny_1y_k$, $y_{k+1} = x_1y_k + y_1x_k$. Hay infinitas soluciones, siempre.

7. **Verifica sustituyendo, y comprueba el tamaño.** Las soluciones de Pell crecen geométricamente y pueden ser enormes desde el principio; si tu "solución fundamental" tiene diez cifras, probablemente no te has equivocado. Y para el caso lineal, comprueba que dos valores consecutivos de t den soluciones válidas.

Ejemplo 3.98 (Cinco ecuaciones diofánticas resueltas por completo)**. (a) Lineal:** $17x+13y = 5$**.** Primero el paso 1: $\gcd(17, 13) = 1$, que divide a 5, luego hay solución. Euclides extendido:
$$17 = 1 \cdot 13 + 4, \qquad 13 = 3 \cdot 4 + 1, \qquad 4 = 4 \cdot 1 + 0.$$

Remontamos: $1 = 13 - 3\cdot 4 = 13 - 3(17-13) = 4\cdot 13 - 3\cdot 17$. Es decir, $17\cdot(-3) + 13\cdot 4 = 1$. Multiplicando por 5:

$$x_0 = -15, \qquad y_0 = 20,$$

y en efecto $17(-15) + 13(20) = -255 + 260 = 5$. La solución general es

$$x = -15 + 13t, \qquad y = 20 - 17t, \qquad t \in \mathbb{Z}.$$

Con $t = 1$: $(x, y) = (-2, 3)$ y $17(-2) + 13(3) = -34 + 39 = 5$. Con $t = 2$: $(11, -14)$ y $187 - 182 = 5$. Correcto. Si además se pidieran soluciones *positivas*, habría que imponer $-15 + 13t > 0$ y $20 - 17t > 0$, es decir $t > 15/13$ y $t < 20/17$, condiciones incompatibles: no hay ninguna.

(b) Una lineal sin solución, y el problema de las monedas. La ecuación $6x + 15y = 7$ no tiene soluciones enteras, porque $\gcd(6, 15) = 3$ no divide a 7: el lado izquierdo siempre es múltiplo de 3. En cambio, si se piden soluciones con $x, y \geq 0$ el problema cambia por completo. Con monedas de 5 y de 7, la mayor cantidad *no* pagable es el número de Frobenius $5\cdot 7 - 5 - 7 = 23$; en efecto, $23 - 7 = 16$, $23 - 14 = 9$ y $23 - 21 = 2$ no son múltiplos de 5, y 23 tampoco lo es. A partir de 24 todo se puede pagar: $24 = 2\cdot 5 + 2\cdot 7$, $25 = 5\cdot 5$, $26 = 5 + 3\cdot 7$, $27 = 4\cdot 5 + 7$, $28 = 4\cdot 7$, y de ahí en adelante se suma 5.

(c) Ternas pitagóricas. La parametrización del paso 4 con los primeros pares (m, n) coprimos y de paridad distinta:

m	n	$m^2 - n^2$	$2mn$	$m^2 + n^2$
2	1	3	4	5
3	2	5	12	13
4	1	15	8	17
4	3	7	24	25
5	2	21	20	29
5	4	9	40	41

Todas cumplen la ecuación: por ejemplo $9^2 + 40^2 = 81 + 1600 = 1681 = 41^2$. Y la identidad general es una línea de álgebra:

$$(m^2 - n^2)^2 + (2mn)^2 = m^4 - 2m^2n^2 + n^4 + 4m^2n^2 = (m^2 + n^2)^2.$$

Lo difícil no es comprobar la fórmula, sino demostrar que *no hay más*; ahí interviene que x, y, z primitivos obligan a $z - x$ y $z + x$ a ser dos cuadrados por un argumento de factorización única.

(d) Pell con $N = 2$. Buscamos $x^2 - 2y^2 = 1$. Probando: $y = 1$ da 3, no cuadrado; $y = 2$ da $2\cdot 4 + 1 = 9 = 3^2$. Solución fundamental $(3, 2)$. La recurrencia del paso 6 es $x_{k+1} = 3x_k + 4y_k$, $y_{k+1} = 2x_k + 3y_k$, y genera

$$(3, 2) \to (17, 12) \to (99, 70) \to (577, 408) \to \cdots$$

Verificaciones: $17^2 - 2\cdot 12^2 = 289 - 288 = 1$; $99^2 - 2\cdot 70^2 = 9801 - 9800 = 1$; $577^2 - 2\cdot 408^2 = 332929 - 332928 = 1$. Y hay una sorpresa: los cocientes

$$\frac{3}{2} = 1.5, \quad \frac{17}{12} = 1.416667, \quad \frac{99}{70} = 1.4142857, \quad \frac{577}{408} = 1.41421569$$

convergen a $\sqrt{2} = 1.41421356$, ganando unas dos cifras por paso. No es casualidad: las soluciones de Pell *son* los convergentes del desarrollo en fracción continua de $\sqrt{N}$, que para $\sqrt{2}$ es $[1;\overline{2}]$. Con $N = 3$ la fundamental es $(2, 1)$ y la sucesión es $(2, 1) \to (7, 4) \to (26, 15) \to (97, 56)$, con $97^2 - 3 \cdot 56^2 = 9409 - 9408 = 1$.

(e) Pell con $N = 61$: la trampa de Fermat. Fermat propuso este caso como reto a los matemáticos ingleses, y con razón. La solución fundamental de $x^2 - 61y^2 = 1$ es

$$x = 1\,766\,319\,049, \qquad y = 226\,153\,980.$$

Nueve y diez cifras: buscar por fuerza bruta probando $y = 1, 2, 3, \ldots$ exigiría doscientos millones de comprobaciones. El método correcto pasa por la fracción continua $\sqrt{61} = [7;\overline{1, 4, 3, 1, 2, 2, 1, 3, 4, 1, 14}]$, cuyo periodo tiene longitud 11, un número impar; eso obliga a tomar el convergente número 22 en vez del número 11, y de ahí el tamaño monstruoso. Es el ejemplo clásico de que el tamaño de la solución no guarda ninguna relación con el tamaño de la ecuación.

(f) Cuándo N es un cuadrado. Si $N = 4$, la ecuación $x^2 - 4y^2 = 1$ se factoriza:

$$(x - 2y)(x + 2y) = 1.$$

Como ambos factores son enteros, deben ser 1 y 1 o -1 y -1, lo que fuerza $y = 0$ y $x = \pm 1$. La diferencia con el caso $N = 2$ es abismal —de infinitas soluciones a dos— y el motivo es puramente algebraico: $\sqrt{N}$ es racional, el anillo $\mathbb{Z}[\sqrt{N}]$ se rompe y no hay unidades interesantes que generar.

Cuidado

Comprobar que hay solución módulo todos los números pequeños *no* demuestra que haya solución entera. La ecuación $3x^2 + 4y^2 = z^2$ con soluciones no triviales, o el famoso contraejemplo de Selmer $3x^3 + 4y^3 + 5z^3 = 0$, tienen soluciones módulo todo primo y ninguna en los enteros. A esa propiedad se le llama fallo del *principio local–global* de Hasse–Minkowski, principio que sí es cierto para formas cuadráticas y falla ya en grado tres. Las congruencias son un filtro excelente para descartar, pero casi nunca bastan para construir.

Idea clave

Las tres familias de esta regla esconden la misma estructura algebraica. Las soluciones de $x^2 - Ny^2 = 1$ son exactamente las *unidades de norma uno* del anillo $\mathbb{Z}[\sqrt{N}]$, y el hecho de que haya infinitas y que todas sean potencias de una sola es el teorema de las unidades de Dirichlet. Las ternas pitagóricas son los puntos racionales de la circunferencia $u^2 + v^2 = 1$, y la parametrización del paso 4 no es más que la proyección estereográfica desde $(-1, 0)$. Y las ecuaciones lineales son el caso de rango uno del mismo cuento. Cuando el grado sube a tres, la curva deja de admitir parametrización racional, aparecen las curvas elípticas, y el conjunto de soluciones se convierte en un grupo finitamente generado: el teorema de Mordell–Weil. La demostración de Wiles del último teorema de Fermat vive justo ahí.

3.10.3. La función φ de Euler y el teorema que generaliza a Fermat

En la sección de las potencias buscábamos el ciclo a mano, probando hasta que aparecía un 1. Euler descubrió que la longitud de ese ciclo se puede predecir sin calcular nada, con solo mirar cómo se factoriza el módulo.

Definición 3.99. Para $n \geq 1$, la *función φ de Euler* cuenta cuántos números de 1 a n son primos con n:

$$\varphi(n) = \#\{a : 1 \leq a \leq n,\ \mathrm{mcd}(a, n) = 1\}.$$

Esos son exactamente los que tienen inverso módulo n; por eso se les llama las *unidades* módulo n.

Teorema 3.100 (Euler)**.** *Si* $\mathrm{mcd}(a, n) = 1$*, entonces* $a^{\varphi(n)} \equiv 1 \pmod{n}$*. En particular, si p es primo y p no divide a a, entonces* $a^{p-1} \equiv 1 \pmod{p}$*, que es el pequeño teorema de Fermat.*

Idea clave

El teorema dice que el ciclo de potencias de cualquier unidad se cierra a más tardar en el paso $\varphi(n)$. Eso convierte el cálculo de a^k mód n en un problema de dos líneas: se reduce el exponente k módulo $\varphi(n)$ y se calcula la potencia pequeña que queda. Fíjese en la diferencia de escalas: el exponente vive módulo $\varphi(n)$, la base vive módulo n. Confundir esos dos módulos es el error número uno de todo este capítulo.

Regla: calcular $\varphi(n)$ y reducir un exponente

1. **Factoriza** n en potencias de primos: $n = p_1^{k_1} \cdots p_r^{k_r}$. Sin la factorización no hay fórmula.
2. **Para un primo,** $\varphi(p) = p - 1$: todos los números de 1 a $p - 1$ son primos con p.
3. **Para una potencia de primo,** $\varphi(p^k) = p^k - p^{k-1} = p^{k-1}(p-1)$: de los p^k números se quitan los múltiplos de p, que son p^{k-1}.
4. **Multiplica:** φ es multiplicativa para factores primos entre sí, luego $\varphi(n) = \varphi(p_1^{k_1}) \cdots \varphi(p_r^{k_r})$. Equivalentemente, $\varphi(n) = n \prod_{p|n} \left(1 - \frac{1}{p}\right)$.
5. **Comprueba que la base es unidad:** el teorema exige $\mathrm{mcd}(a, n) = 1$. Si no lo es, no se aplica y hay que ir por el teorema chino del resto.
6. **Reduce el exponente módulo** $\varphi(n)$, no módulo n: si $k = q\varphi(n) + s$, entonces $a^k \equiv a^s \pmod{n}$.
7. **Calcula la potencia pequeña** que queda, si hace falta por cuadrados sucesivos, y no olvides el caso $s = 0$, que da 1.

Ejemplo 3.101 (Cuatro valores de φ, calculados de dos maneras)**.** 1. $\varphi(10)$. Contando: de 1 a 10 son primos con 10 los números $1, 3, 7, 9$, o sea 4. Con la fórmula: $10 = 2 \cdot 5$,

luego $\varphi(10) = \varphi(2)\varphi(5) = 1 \cdot 4 = 4$. ✓

2. $\varphi(36)$. Contar 36 números a mano es tedioso; la fórmula lo hace de cabeza: $36 = 2^2 3^2$, así que $\varphi(36) = \varphi(4)\varphi(9) = 2 \cdot 6 = 12$. Con la otra versión, $36\left(1 - \frac{1}{2}\right)\left(1 - \frac{1}{3}\right) = 36 \cdot \frac{1}{2} \cdot \frac{2}{3} = 12$. ✓

3. $\varphi(100) = \varphi(4)\varphi(25) = 2 \cdot 20 = 40$, y $\varphi(360) = \varphi(8)\varphi(9)\varphi(5) = 4 \cdot 6 \cdot 4 = 96$. Obsérvese cómo un módulo con muchos factores primos pequeños tiene un φ proporcionalmente pequeño.

4. $\varphi(101) = 100$, porque 101 es primo. Y $\varphi(55) = \varphi(5)\varphi(11) = 4 \cdot 10 = 40$: este es exactamente el número que aparecía en la construcción de RSA de la sección siguiente. Ahora ya sabemos de dónde salía $(p-1)(q-1)$.

Ejemplo 3.102 (Cuatro potencias monstruosas, en una línea)**.** 1. 7^{100} mód 10. Como $\mathrm{mcd}(7, 10) = 1$ y $\varphi(10) = 4$, reducimos $100 \equiv 0 \pmod 4$, luego $7^{100} \equiv 7^0 = 1$. La última cifra de 7^{100} es un 1.

2. $2^{1000000}$ mód 55. Aquí $\varphi(55) = 40$ y $1000000 = 25000 \cdot 40$, así que el exponente se reduce a 0 y $2^{1000000} \equiv 1 \pmod{55}$. Un millón de multiplicaciones evitadas con una división.

3. Inversos sin Euclides: como $a^{\varphi(n)} \equiv 1$, resulta que $a^{-1} \equiv a^{\varphi(n)-1}$. Módulo 10, el inverso de 3 es $3^3 = 27 \equiv 7$, y en efecto $3 \cdot 7 = 21 \equiv 1$. Módulo 40, el inverso de 7 es $7^{\varphi(40)-1} = 7^{15} \equiv 23$, y $7 \cdot 23 = 161 = 4 \cdot 40 + 1$. ✓

4. Exponenciación por cuadrados sucesivos, para cuando el exponente reducido sigue siendo grande. Para 3^{13} mód 17 se escribe $13 = 1101_2 = 8 + 4 + 1$ y se elevan al cuadrado sucesivamente: $3^1 \equiv 3$, $3^2 \equiv 9$, $3^4 \equiv 13$, $3^8 \equiv 16$. Entonces $3^{13} \equiv 16 \cdot 13 \cdot 3 = 624 \equiv 12 \pmod{17}$. Se han hecho cinco multiplicaciones en vez de doce, y con exponentes de 300 cifras la diferencia es entre un instante y la edad del universo.

Cuidado

El teorema de Euler exige $\mathrm{mcd}(a, n) = 1$ y esa hipótesis no es decorativa. Módulo 10, la base 2 no es unidad, y en efecto $2^4 = 16 \equiv 6$, no 1; las potencias de 2 nunca vuelven al 1, se quedan girando entre $2, 4, 8, 6$. Antes de reducir un exponente hay que comprobar siempre que la base es prima con el módulo. Y una segunda trampa: $\varphi(n)$ da un exponente que *funciona*, pero no necesariamente el más pequeño. Módulo 7, $\varphi(7) = 6$ y sin embargo $2^3 = 8 \equiv 1$: el ciclo real de 2 tiene longitud 3, no 6.

Pausa para pensar

El teorema de Euler tiene una lectura que lo vuelve casi obvio: los enteros primos con n, con el producto módulo n, forman un grupo de $\varphi(n)$ elementos, y en cualquier grupo finito el orden de un elemento divide al orden del grupo. Esa es la razón profunda, y se demuestra en el capítulo 23 con el teorema de Lagrange, sin hablar de números en ningún momento. Vale la pena parar a admirar la economía: un enunciado sobre restos de potencias resulta ser un caso particular de un teorema sobre simetrías.

3.11. Orden, raíces primitivas y el intercambio de claves de Diffie y Hellman

El teorema de Euler acota la longitud del ciclo, pero no la da. Medirla con exactitud abre la puerta a un truco criptográfico distinto del de RSA, y en cierto sentido más elegante: dos personas que nunca se han visto acuerdan una clave secreta hablando en voz alta delante de todo el mundo.

Definición 3.103. Si $\mcd(a, n) = 1$, el *orden* de a módulo n es el menor exponente $k \geq 1$ con $a^k \equiv 1 \pmod{n}$. Cuando ese orden vale exactamente $\varphi(n)$, se dice que a es una *raíz primitiva* módulo n: sus potencias recorren *todas* las unidades, sin repetir ninguna hasta cerrar el ciclo.

Idea clave

Una raíz primitiva es un generador: da una manera de numerar todas las unidades módulo p como potencias suyas. Eso convierte los productos en sumas de exponentes, igual que los logaritmos ordinarios, y por eso al exponente se le llama *logaritmo discreto*. La gracia es que aquí la analogía se rompe donde interesa: elevar es fácil, pero *deshacer* la exponenciación —encontrar el exponente— no se sabe hacer rápido. Toda la criptografía de Diffie y Hellman vive en esa asimetría.

Regla: hallar el orden, una raíz primitiva y montar un intercambio de claves

1. **Calcula** $\varphi(n)$; el orden de cualquier unidad tiene que ser un divisor suyo, así que solo hay que probar entre los divisores.
2. **Lista los divisores de** $\varphi(n)$ **en orden creciente** y calcula a^d mód n para cada uno; el primer d que dé 1 es el orden.
3. **Para saber si** a **es raíz primitiva módulo un primo** p**,** basta comprobar que $a^{(p-1)/q} \not\equiv 1$ para cada primo q que divida a $p - 1$. No hace falta probar todos los divisores.
4. **Recuerda dónde existen:** hay raíces primitivas módulo p y p^k con p primo impar, y módulo 2 y 4, pero no módulo 8, ni 12, ni en general para un módulo con dos factores primos impares distintos.
5. **Para el intercambio de claves,** publica un primo grande p y una raíz primitiva g: son datos abiertos, no secretos.
6. **Cada participante elige un exponente secreto** —a para Alicia, b para Bruno— y publica solo $A = g^a$ mód p y $B = g^b$ mód p.
7. **Cada uno calcula la clave común** elevando lo que recibió a su propio secreto: Alicia obtiene B^a y Bruno obtiene A^b, y ambos valen g^{ab} mód p. Quien escucha ve p, g, A y B, y para llegar a g^{ab} necesitaría el logaritmo discreto.

Ejemplo 3.104 (La tabla de órdenes módulo 7)**.** Como $\varphi(7) = 6$, los órdenes posibles son los divisores de 6: $1, 2, 3, 6$. Calculamos las potencias sucesivas de cada unidad.

1. 1 da siempre 1: orden 1. Y $6 \equiv -1$ da $6, 1, 6, 1, \ldots$: orden 2, como toda base congruente con -1.
2. 2 da $2, 4, 1, 2, 4, 1$: orden 3, la mitad de lo que permitía Euler. Lo mismo le pasa a 4, que da $4, 2, 1, 4, 2, 1$.
3. 3 da $3, 2, 6, 4, 5, 1$: aparecen las seis unidades $\{1, 2, 3, 4, 5, 6\}$ antes de cerrar. Orden 6, luego 3 *es* raíz primitiva módulo 7. También lo es 5, que da $5, 4, 6, 2, 3, 1$.
4. Recuento: los órdenes son $1, 3, 6, 3, 6, 2$, todos divisores de 6, y hay exactamente dos raíces primitivas. En general, módulo un primo p hay $\varphi(p-1)$ raíces primitivas: aquí $\varphi(6) = 2$. ✓

Ejemplo 3.105 (Alicia y Bruno acuerdan una clave delante de todos)**.** Se publica $p = 23$ y $g = 5$, que es raíz primitiva módulo 23.

1. Alicia elige en secreto $a = 6$ y publica $A = 5^6$ mód 23. Por cuadrados sucesivos: $5^2 = 25 \equiv 2$, $5^4 \equiv 4$, luego $5^6 = 5^4 \cdot 5^2 \equiv 4 \cdot 2 = 8$. Publica $A = 8$.
2. Bruno elige en secreto $b = 15$ y publica $B = 5^{15}$ mód 23. Como $5^8 \equiv 16$, se tiene $5^{15} = 5^8 \cdot 5^4 \cdot 5^2 \cdot 5 \equiv 16 \cdot 4 \cdot 2 \cdot 5 = 640 \equiv 19$. Publica $B = 19$.
3. Alicia calcula $B^a = 19^6$ mód $23 = 2$; Bruno calcula $A^b = 8^{15}$ mód $23 = 2$. Los dos han obtenido la misma clave 2, que es $g^{ab} = 5^{90}$ mód 23, sin haberse dicho nunca sus exponentes.
4. Un espía conoce $p = 23$, $g = 5$, $A = 8$ y $B = 19$. Para pasar de ahí a la clave necesita resolver $5^x \equiv 8 \pmod{23}$, y con números tan pequeños lo hace por fuerza bruta: probando $x = 1, 2, 3, \ldots$ encuentra $x = 6$. Con un primo de 600 cifras esa búsqueda deja de ser practicable, y ahí está toda la seguridad del protocolo.

Cuidado

El intercambio de claves protege contra quien *escucha*, pero no contra quien *interviene*. Si alguien se coloca en medio y se hace pasar por Bruno ante Alicia y por Alicia ante Bruno, acuerda una clave con cada uno y lee todo el tráfico sin romper ninguna matemática. Es el ataque del intermediario, y se resuelve con firmas digitales, no con más aritmética modular. La lección general es que un protocolo criptográfico correcto necesita, además de un problema difícil, una manera de saber con quién se está hablando.

Pausa para pensar

El logaritmo discreto y la factorización son problemas distintos, pero comparten un rasgo inquietante: nadie ha demostrado que sean difíciles. Solo se sabe que nadie ha encontrado un método rápido, lo cual no es un teorema sino un estado de la cuestión. Si mañana apareciera un algoritmo eficiente para uno de los dos, buena parte de la criptografía

actual dejaría de servir de un día para otro. De hecho, para un ordenador cuántico ese algoritmo ya existe sobre el papel, y esa es la razón de que se busquen sistemas basados en otros problemas. Merece la pena retener la idea: la seguridad práctica descansa hoy sobre una ignorancia compartida, no sobre una demostración.

3.12. Residuos cuadráticos y el símbolo de Legendre

Sabemos resolver $ax \equiv b \pmod m$. La pregunta siguiente es la que llevó a Gauss a escribir su primer gran libro: ¿cuándo tiene solución $x^2 \equiv a \pmod p$? Es decir, ¿qué números son cuadrados en el reloj?

Definición 3.106. Sea p un primo impar y a no divisible por p. Se dice que a es *residuo cuadrático* módulo p si la congruencia $x^2 \equiv a \pmod p$ tiene solución. El *símbolo de Legendre* se define como

$$\left(\frac{a}{p}\right) = \begin{cases} +1 & \text{si } a \text{ es residuo cuadrático módulo } p, \\ -1 & \text{si no lo es.} \end{cases}$$

Idea clave

En el reloj de p horas, elevar al cuadrado empareja: x y $-x$ dan el mismo resultado, y nada más coincide. Por eso de las $p-1$ unidades solo la mitad son cuadrados, y la otra mitad no lo es nunca, por mucho que se busque. El símbolo de Legendre no es más que una forma cómoda de escribir ese sí o no, y su virtud es que se comporta como un producto: $\left(\frac{ab}{p}\right) = \left(\frac{a}{p}\right)\left(\frac{b}{p}\right)$. Eso permite descomponer una pregunta difícil en preguntas fáciles.

Regla: decidir si a es un cuadrado módulo p

1. **Reduce a módulo p** primero; el símbolo solo depende de la clase de a.
2. **Factoriza** a y aplica la multiplicatividad: el problema se reduce a calcular el símbolo de -1, de 2 y de cada primo impar que aparezca.
3. **Usa la primera ley complementaria:** $\left(\frac{-1}{p}\right) = +1$ si $p \equiv 1 \pmod 4$, y -1 si $p \equiv 3 \pmod 4$.
4. **Usa la segunda ley complementaria:** $\left(\frac{2}{p}\right) = +1$ si $p \equiv \pm 1 \pmod 8$, y -1 si $p \equiv \pm 3 \pmod 8$.
5. **Usa la reciprocidad cuadrática** para dar la vuelta a los primos impares $q \neq p$: $\left(\frac{q}{p}\right)\left(\frac{p}{q}\right) = (-1)^{\frac{p-1}{2}\cdot\frac{q-1}{2}}$. En palabras: los dos símbolos coinciden salvo que p y q sean ambos congruentes con 3 módulo 4, en cuyo caso son opuestos.
6. **Al dar la vuelta, reduce y repite;** el proceso baja de tamaño en cada paso, igual que el algoritmo de Euclides, y termina siempre.

7. **Si prefieres calcular en vez de razonar,** usa el criterio de Euler: $\left(\frac{a}{p}\right) \equiv a^{(p-1)/2} \pmod p$, cuyo resultado es siempre 1 o $p-1$.

Ejemplo 3.107 (Los cuadrados de tres relojes pequeños)**.** 1. Módulo 7: elevando al cuadrado $1, \ldots, 6$ salen $1, 4, 2, 2, 4, 1$, de modo que los residuos cuadráticos son $\{1, 2, 4\}$ y los no residuos $\{3, 5, 6\}$. Tres y tres, como anunciaba la teoría.

2. Módulo 11 los cuadrados son $\{1, 3, 4, 5, 9\}$, cinco de diez; módulo 13 son $\{1, 3, 4, 9, 10, 12\}$, seis de doce. En los tres casos la lista tiene exactamente $(p-1)/2$ elementos.

3. Resolver $x^2 \equiv 2 \pmod 7$: como 2 está en la lista, hay solución, y probando se ve que $x \equiv 3$ y $x \equiv 4$, que son ± 3. Resolver $x^2 \equiv 3 \pmod 7$: como 3 no está, no hay ninguna solución, y no hace falta buscarla.

4. Comprobación de la multiplicatividad módulo 11: $\left(\frac{3}{11}\right) = +1$ y $\left(\frac{5}{11}\right) = +1$, luego $\left(\frac{15}{11}\right)$ debe valer $+1$; y en efecto $15 \equiv 4$, que es un cuadrado. ✓

Ejemplo 3.108 (El criterio de Euler y la reciprocidad en acción)**.** 1. Criterio de Euler módulo 7, con exponente $(7-1)/2 = 3$: los cubos de $1, 2, 3, 4, 5, 6$ son $1, 1, 6, 1, 6, 6$. Donde sale 1 hay residuo cuadrático, donde sale $6 \equiv -1$ no lo hay, y la lista coincide exactamente con $\{1, 2, 4\}$ frente a $\{3, 5, 6\}$. ✓

2. Primera ley complementaria: -1 es cuadrado módulo $5, 13, 17$ (todos $\equiv 1$ mód 4) y no lo es módulo $7, 11, 19$ (todos $\equiv 3$ mód 4). Esta es la razón última de que un primo se escriba como suma de dos cuadrados si y solo si es $\equiv 1 \pmod 4$.

3. Reciprocidad con números pequeños: $\left(\frac{3}{7}\right)$ y $\left(\frac{7}{3}\right)$. Como $3 \equiv 3$ y $7 \equiv 3$ módulo 4, los símbolos son opuestos. Y en efecto $\left(\frac{7}{3}\right) = \left(\frac{1}{3}\right) = +1$, luego $\left(\frac{3}{7}\right) = -1$, coherente con que 3 no fuera cuadrado módulo 7.

4. Un cálculo grande sin probar nada: $\left(\frac{219}{383}\right)$, con 383 primo. Factorizamos $219 = 3 \cdot 73$. Para el 3: como $383 \equiv 3$ y $3 \equiv 3$ módulo 4, hay cambio de signo, y $\left(\frac{383}{3}\right) = \left(\frac{2}{3}\right) = -1$, luego $\left(\frac{3}{383}\right) = +1$. Para el 73: como $73 \equiv 1 \pmod 4$ no hay cambio, y $\left(\frac{383}{73}\right) = \left(\frac{18}{73}\right) = \left(\frac{2}{73}\right)$, que vale $+1$ porque $73 \equiv 1 \pmod 8$; luego $\left(\frac{73}{383}\right) = +1$. El producto da $\left(\frac{219}{383}\right) = +1$: hay solución. Buscándola resultan ser $x \equiv 169$ y $x \equiv 214$, pero eso ya ha costado trabajo, mientras que *saber* que existía ha costado cuatro líneas.

Cuidado

La reciprocidad cuadrática dice si hay solución, no cuál es. Son dos problemas de dificultad muy distinta: decidir la existencia es rápido, y encontrar la raíz cuadrada módulo p ya requiere un algoritmo aparte. Segunda precaución: el símbolo de Legendre solo está definido para p primo impar. Existe una extensión, el símbolo de Jacobi, que admite módulos compuestos y se calcula igual de rápido, pero pierde la propiedad esencial: un símbolo de Jacobi igual a $+1$ *no* garantiza que haya solución. Es una trampa

clásica, y hay que tenerla presente antes de aplicar la fórmula fuera de su terreno.

Pausa para pensar

Gauss llamó a este resultado el *teorema áureo* y publicó seis demostraciones distintas a lo largo de su vida; hoy se conocen más de doscientas. Vale la pena preguntarse por qué tanto empeño en un enunciado que, a primera vista, es una curiosidad sobre restos. La respuesta es que la reciprocidad relaciona dos primos que no tienen nada que ver el uno con el otro, y esa clase de relación inesperada resultó ser la punta de un iceberg: la teoría de cuerpos de clases y el programa de Langlands son, en cierto sentido, la continuación de esta única fórmula. Los cuerpos finitos del capítulo 24 y las extensiones del capítulo 26 son el lenguaje en que esa continuación se escribe.

3.13. Criptografía clásica: sustituir letras

Numeremos las letras $A = 0$, $B = 1$, …, $Z = 25$ (usamos el alfabeto de 26 letras, sin ñ, que es el convenio internacional). Cifrar es aplicar una función a cada número; descifrar es aplicar la inversa.

3.13.1. El cifrado de César

Regla: César

Con clave k, cada letra de valor p se convierte en

$$c \equiv p + k \pmod{26},$$

y se descifra con $p \equiv c - k \pmod{26}$.

Ejemplo 3.109. Con $k = 3$, ATACAR AL AMANECER se convierte en `DWDFDU DO DPDQHFHU`. Compruebe una letra: la R vale 17, y $17 + 3 = 20$, que es la U. Para descifrar se resta 3 y se recupera el mensaje.

Cuidado

El cifrado de César tiene **25 claves útiles**. Cualquiera puede probarlas todas en un minuto con lápiz y papel; un ordenador tarda microsegundos. Peor aún: ni siquiera hace falta probarlas, porque el cifrado conserva la frecuencia de las letras. En español la E y la A aparecen alrededor del 12–13 % de las veces cada una; si en el texto cifrado la letra más frecuente es la H, casi seguro que la clave es 3. Este ataque, el **análisis de frecuencias**, lo describió al-Kindi en Bagdad en el siglo IX, y liquidó de un plumazo toda la criptografía de la Antigüedad.

3.13.2. El cifrado afín

Generalicemos: en lugar de sumar, apliquemos una función lineal.

Regla: cifrado afín

$c \equiv a\,p + b \pmod{26}$, con $\mathrm{mcd}(a, 26) = 1$. Se descifra con $p \equiv a^{-1}(c - b) \pmod{26}$.

Teorema 3.110. *La aplicación* $p \mapsto ap + b$ mód 26 *es biyectiva si y solo si* $\mathrm{mcd}(a, 26) = 1$.

Demostración. Si $\mathrm{mcd}(a, 26) = 1$, existe a^{-1} por el corolario anterior y la aplicación $c \mapsto a^{-1}(c - b)$ es su inversa. Si $d = \mathrm{mcd}(a, 26) > 1$, entonces p y $p + 26/d$ tienen la misma imagen, así que la aplicación no es inyectiva. □

Ejemplo 3.111. Con $a = 5$, $b = 8$: la A(0) $\to$ 8 =I, la B(1) $\to$ 13 =N, la C(2) $\to$ 18 =S, la D(3) $\to$ 23 =X, la E(4) $\to$ 2 =C. Para descifrar hace falta 5^{-1} mód 26, que es 21 (porque $5 \cdot 21 = 105 = 4 \cdot 26 + 1$).

Observación 3.112. Los valores admisibles de a son los primos con $26 = 2 \cdot 13$, es decir los impares distintos de 13: hay 12. Con 26 valores posibles de b, salen $12 \cdot 26 = 312$ claves. Mejor que 25, pero sigue siendo ridículo: un ordenador las prueba todas antes de que usted termine de leer esta frase.

3.13.3. Vigenère: una clave que se mueve

La debilidad de César es que la misma letra siempre se cifra igual. Vigenère lo arregla usando una palabra clave que se repite: la j-ésima letra del mensaje se desplaza según la j-ésima letra de la clave.

Ejemplo 3.113. Mensaje ATACAR, clave CLAVE (valores $2, 11, 0, 21, 4$):

texto	A	T	A	C	A	R
clave	C	L	A	V	E	C
suma	0+2	19+11	0+0	2+21	0+4	17+2
cifra	C	E	A	X	E	T

El resultado es `CEAXET`. Observe que las tres A del mensaje se han convertido en C, A y E: el análisis de frecuencias directo ya no sirve.

Cuidado

Vigenère se llamó durante tres siglos *le chiffre indéchiffrable*, y no lo era. Si la clave tiene longitud L, las letras en posiciones $1, 1 + L, 1 + 2L, \ldots$ están todas cifradas con el mismo desplazamiento: son un César. Basta averiguar L (Kasiski, 1863, lo dedujo de las repeticiones en el texto cifrado) y aplicar L análisis de frecuencias por separado. La lección es incómoda: **la complicación no es seguridad**.

Idea clave

Todos los cifrados de esta sección comparten un defecto estructural: quien puede cifrar puede descifrar, porque la clave es la misma. Eso obliga a que emisor y receptor se hayan puesto de acuerdo antes, en persona o por un canal seguro; y si ya tienen un canal seguro, ¿para qué quieren el cifrado? Este círculo vicioso, el **problema del intercambio de claves**, bloqueó la criptografía hasta 1976. La salida no vino de complicar más las sustituciones: vino de la aritmética modular con números primos enormes.

3.14. RSA: cómo mandar un secreto, paso a paso

Idea clave

Intuición primero. El problema que resuelve el RSA parece imposible cuando se enuncia: dos personas que nunca se han visto, comunicándose por un canal que todo el mundo puede escuchar, quieren intercambiar un secreto. Sin un encuentro previo no hay clave compartida, y sin clave compartida no hay cifrado clásico posible.
La salida es una idea genuinamente nueva: una operación que cualquiera pueda *hacer* pero que solo el destinatario pueda *deshacer*. Multiplicar dos primos grandes es instantáneo; recuperar los primos a partir del producto no lo es, y esa asimetría es toda la seguridad del sistema. El candado se publica abierto, la llave se guarda. Lo que hace que el mecanismo funcione matemáticamente es el teorema de Euler de la sección anterior: cifrar es elevar a e, descifrar es elevar a d, y se eligen de modo que $ed \equiv 1 \pmod{\varphi(n)}$, con lo cual la composición devuelve el mensaje original. Conviene notar que la seguridad no está *demostrada*: descansa en que nadie ha encontrado una manera rápida de factorizar. Es una apuesta, aunque una apuesta que lleva cincuenta años ganándose.

Ahora vamos a juntar todo lo del capítulo —Euclides, Bézout, inversos, Euler— para construir el sistema de cifrado que protege prácticamente toda la comunicación segura del mundo. Y no lo vamos a describir en abstracto: lo vamos a *ejecutar* con números pequeños, cifrando y descifrando un mensaje de verdad, para que veas cada engranaje girar.

3.14.1. La idea

El problema que RSA resuelve suena imposible: quiero que *cualquiera* pueda enviarme un mensaje secreto, pero que solo *yo* pueda leerlo, aunque el método de cifrado sea público. La solución se apoya en una asimetría: multiplicar dos primos grandes es fácil, pero recuperarlos de su producto (factorizar) es, hasta donde sabemos, descomunalmente difícil. Yo publico el producto; el secreto son los factores.

3.14.2. Preparando las llaves (esto lo hace quien recibe)

Sigamos el procedimiento con números de juguete para poder verlo todo.

Paso 1: elegir dos primos. Tomo $p = 5$ y $q = 11$. (En la vida real tendrían cientos de dígitos; aquí los queremos diminutos para calcular a mano.)

Paso 2: multiplicarlos. $n = pq = 5 \cdot 11 = 55$. Este n será público.

Paso 3: calcular $\varphi(n)$. Como n es producto de dos primos distintos, $\varphi(n) = (p-1)(q-1) = 4 \cdot 10 = 40$. Este número es secreto (calcularlo requiere conocer p y q).

Paso 4: elegir un exponente público e. Necesito un e coprimo con $\varphi(n) = 40$. Elijo $e = 3$; en efecto $\gcd(3, 40) = 1$. La pareja $(n, e) = (55, 3)$ es mi *llave pública*: la anuncio al mundo.

Paso 5: calcular el exponente privado d. Aquí uso los inversos modulares del capítulo. Necesito d con

$$e\,d \equiv 1 \pmod{\varphi(n)}, \qquad \text{es decir} \qquad 3d \equiv 1 \pmod{40}.$$

Corriendo el algoritmo de Euclides (o probando), $d = 27$ funciona: $3 \cdot 27 = 81 = 80 + 1 \equiv 1$ (mód 40). Este $d = 27$ es mi *llave privada*: no se la digo a nadie.

3.14.3. Cifrando (esto lo hace quien envía)

Supón que quieres mandarme el mensaje $M = 7$ (un mensaje es solo un número menor que $n = 55$; los textos reales se convierten en números por bloques). Usas mi llave pública $(55, 3)$ y calculas

$$C = M^e \bmod n = 7^3 \bmod 55.$$

Ahora $7^3 = 343$, y $343 = 6 \cdot 55 + 13$, así que $C = 13$. Me envías el número 13 por un canal completamente público; cualquiera puede verlo.

3.14.4. Descifrando (solo puedo yo)

Recibo $C = 13$ y aplico mi llave privada $d = 27$:

$$M = C^d \bmod n = 13^{27} \bmod 55.$$

Esto parece un cálculo monstruoso, pero se hace por partes reduciendo módulo 55 en cada paso (así los números nunca crecen). Vamos:

$$13^2 = 169 \equiv 4, \quad 13^4 \equiv 4^2 = 16, \quad 13^8 \equiv 16^2 = 256 \equiv 36, \quad 13^{16} \equiv 36^2 = 1296 \equiv 31 \pmod{55}.$$

Como $27 = 16 + 8 + 2 + 1$, multiplicamos las piezas correspondientes:

$$13^{27} = 13^{16} \cdot 13^8 \cdot 13^2 \cdot 13^1 \equiv 31 \cdot 36 \cdot 4 \cdot 13 \pmod{55}.$$

Paso a paso: $31 \cdot 36 = 1116 \equiv 6$ (pues $1116 = 20 \cdot 55 + 16$... verifiquemos con cuidado: $20 \cdot 55 = 1100$, $1116 - 1100 = 16$, así que $\equiv 16$); sigamos, $16 \cdot 4 = 64 \equiv 9$; y $9 \cdot 13 = 117 \equiv 7$ (pues $117 = 2 \cdot 55 + 7$). El resultado es $\boxed{7}$: ¡recuperé exactamente tu mensaje original!

3.14.5. ¿Por qué funciona? Aquí está la magia, y es pura teoría del capítulo

No fue suerte que descifrar deshiciera el cifrado. Ciframos elevando a e y desciframos elevando a d; en total, el mensaje se eleva a ed. Y elegimos d precisamente para que $ed \equiv 1$ (mód $\varphi(n)$), es decir, $ed = 1 + k\varphi(n)$ para algún entero k. Entonces

$$C^d = (M^e)^d = M^{ed} = M^{1+k\varphi(n)} = M \cdot \left(M^{\varphi(n)}\right)^k.$$

Por el teorema de Euler, si M es coprimo con n, entonces $M^{\varphi(n)} \equiv 1 \pmod{n}$, así que $\left(M^{\varphi(n)}\right)^k \equiv 1$ y todo colapsa a

$$C^d \equiv M \cdot 1 = M \pmod{n}.$$

El teorema de Euler es, literalmente, la razón por la que RSA descifra correctamente. (El caso en que M comparte un factor con n se arregla con el teorema chino del resto, mirando módulo p y módulo q por separado; te lo dejo como uno de los ejercicios guiados.)

Observación 3.114. Detente a apreciar lo que acabas de hacer: usaste el algoritmo de Euclides (para el inverso d), la función de Euler y su fórmula $(p-1)(q-1)$, y el teorema de Euler (para que descifrar funcione). Cada pieza del capítulo tuvo su papel, y ninguna fue decorativa. La seguridad descansa en que un atacante que solo ve $(n, e) = (55, 3)$ tendría que hallar $\varphi(n) = 40$ para deducir d, y hallar $\varphi(n)$ equivale a factorizar n en $5 \cdot 11$. Con n de juguete es trivial; con n de seiscientos dígitos, factorizar tomaría más tiempo que la edad del universo con la tecnología actual. Toda la seguridad del comercio electrónico cuelga de esa dificultad.

3.15. Una joya para cerrar: ¿qué primos son suma de dos cuadrados?

Quiero terminar con un resultado que no tiene ninguna aplicación práctica y es, por eso mismo, una de las cosas más hermosas que conozco. La pregunta es de niño: ¿qué números primos se pueden escribir como suma de dos cuadrados perfectos?

Experimentemos, que ya sabes que así se empieza:

$$2 = 1^2 + 1^2, \quad 5 = 1^2 + 2^2, \quad 13 = 2^2 + 3^2, \quad 17 = 1^2 + 4^2, \quad 29 = 2^2 + 5^2.$$

En cambio, por más que lo intentes, $3, 7, 11, 19, 23$ *no* se dejan. ¿Ves el patrón? Míralos módulo 4. Los que sí se pueden: $5, 13, 17, 29$ dejan resto 1 al dividir entre 4. Los que no: $3, 7, 11, 19, 23$ dejan resto 3. (Y el 2 es un caso aparte.) La conjetura se dibuja sola.

Teorema 3.115 (Fermat, sobre sumas de dos cuadrados)**.** *Un primo impar p es suma de dos cuadrados si y solo si $p \equiv 1 \pmod{4}$.*

Idea clave

Qué dice esto en cristiano. El teorema de los dos cuadrados es de los resultados más elegantes de la teoría de números: contesta con exactitud qué primos se escriben como suma de dos cuadrados, y la respuesta depende solo del resto al dividir por 4. Los primos 1 módulo 4 sí, los primos 3 módulo 4 no, y el 2 es el caso aparte. La razón profunda es que en los enteros de Gauss $\mathbb{Z}[\mathrm{i}]$, donde $a^2 + b^2 = (a + b\mathrm{i})(a - b\mathrm{i})$, esos primos dejan de ser primos y se parten, mientras que los otros siguen siendo irreducibles.

Ejemplo 3.116 (Comprobar el criterio)**. (a) Los que sí.** $5 = 1 + 4 = 1^2 + 2^2$; $13 = 4 + 9 = 2^2 + 3^2$; $17 = 1 + 16 = 1^2 + 4^2$; $29 = 4 + 25 = 2^2 + 5^2$; $37 = 1 + 36 = 1^2 + 6^2$. Los cinco son $\equiv 1 \pmod{4}$, pues valen $5 = 4 + 1$, $13 = 12 + 1$, $17 = 16 + 1$, $29 = 28 + 1$ y $37 = 36 + 1$.

(b) Los que no. 3, 7, 11, 19 y 23 son todos $\equiv 3 \pmod 4$, y ninguno es suma de dos cuadrados. Se comprueba a mano: los cuadrados menores que 23 son $0, 1, 4, 9, 16$, y ninguna pareja suma 23. La razón general es que módulo 4 los cuadrados solo valen 0 o 1, así que una suma de dos cuadrados nunca es 3 módulo 4.

(c) La descomposición es única. Para 13 solo hay una manera, $2^2 + 3^2$, salvo orden y signos. En cambio $65 = 5 \cdot 13$ admite dos: $1^2 + 8^2$ y $4^2 + 7^2$. La multiplicidad viene de la identidad de Brahmagupta–Fibonacci, que multiplica dos sumas de cuadrados dando otra, y de que hay dos maneras de emparejar los factores gaussianos.

(d) El criterio para números compuestos. Un entero n es suma de dos cuadrados si y solo si en su factorización todo primo $\equiv 3 \pmod 4$ aparece con exponente *par*. Así $45 = 3^2 \cdot 5$ sí lo es ($45 = 36 + 9 = 6^2 + 3^2$), porque el 3 está al cuadrado; mientras que $21 = 3 \cdot 7$ no lo es, porque tanto el 3 como el 7 aparecen con exponente impar.

Cuidado

[Con más cuadrados el problema cambia por completo] Con *cuatro* cuadrados no hay ninguna restricción: todo entero positivo es suma de cuatro cuadrados, y eso lo demostró Lagrange en 1770. Por ejemplo $7 = 4+1+1+1$ y $23 = 9+9+4+1$. Con tres cuadrados sí hay restricción, y es más fina: falla exactamente para los números de la forma $4^a(8b+7)$, como el propio 7 y el 23. Segundo aviso: el enunciado del teorema es solo para primos impares; el $2 = 1^2 + 1^2$ se trata aparte y es el único primo que se ramifica en $\mathbb{Z}[\mathrm{i}]$, pues $2 = -\mathrm{i}(1+\mathrm{i})^2$.

No voy a dar la demostración completa —requiere una idea genuinamente nueva—, pero sí quiero mostrarte la mitad fácil y contarte de dónde sale la difícil, porque ahí se esconde una conexión que ata este capítulo con el de números complejos.

La mitad fácil: si $p \equiv 3 \pmod 4$, entonces p no es suma de dos cuadrados. Mira los cuadrados módulo 4. Un número par al cuadrado es $\equiv 0$; uno impar al cuadrado es $\equiv 1$ (pues $(2k+1)^2 = 4k^2 + 4k + 1$). Así que *cualquier* cuadrado es $\equiv 0$ o $\equiv 1$ módulo 4, y una suma de dos cuadrados solo puede ser $\equiv 0, 1$ o 2 módulo 4: *nunca* $\equiv 3$. Por lo tanto un primo $\equiv 3 \pmod 4$ está descartado de entrada. Fíjate qué típico de la aritmética modular: sin saber nada de los cuadrados concretos, un simple conteo de restos prohíbe una familia entera.

De dónde sale la mitad difícil. Que todo primo $p \equiv 1 \pmod 4$ *sí* se pueda escribir así es lo profundo. La demostración más luminosa pasa por los *enteros de Gauss*, los números de la forma $a + bi$ con a, b enteros e $i^2 = -1$. En ese mundo, escribir $p = a^2 + b^2$ es lo mismo que factorizar $p = (a + bi)(a - bi)$: preguntar qué primos son suma de dos cuadrados es preguntar qué primos *dejan de ser primos* al mudarse a los enteros de Gauss. Y resulta que el resto de p módulo 4 controla exactamente eso. Es un ejemplo perfecto de una lección que atraviesa toda la matemática: a veces la manera de resolver un problema es *ampliar el universo* en el que vive, hasta que la respuesta se vuelva visible.

Observación 3.117. Si este final te dejó con ganas de más, esa es justo la sensación correcta. La teoría de números es así: preguntas de una sencillez desarmante que abren, al tirar del hilo, mundos enteros —los enteros de Gauss, los cuerpos que construimos en el capítulo de Galois, la función zeta del capítulo de variable compleja—. Este capítulo fue una puerta de entrada

escrita despacio; cada uno de sus teoremas tiene detrás una teoría que podría llenar su propio libro. Pero ahora tienes los cimientos, construidos ladrillo a ladrillo, y puestos por ti mismo.

3.16. Errores típicos de aritmética modular

1. Escribir $=$ donde va $\equiv$. No es lo mismo $15 = 3$ (falso) que $15 \equiv 3 \pmod{12}$ (cierto).
2. Reducir el exponente con el módulo de la base: en $3^{2026} \bmod 7$, el exponente se reduce módulo 6 (la longitud del ciclo), *no* módulo 7.
3. Dar restos negativos: $-7 \bmod 5$ es 3, no -2.
4. Cancelar factores sin comprobar que son primos con el módulo.
5. Creer que $ax \equiv b$ siempre tiene solución única. Tiene $\gcd(a, m)$ soluciones, o ninguna.
6. Buscar el inverso de un número que no lo tiene: 6 no tiene inverso módulo 10.
7. Aplicar el teorema chino del resto con módulos no coprimos.
8. Sumar las cifras para el criterio del 11, en lugar de alternar los signos.
9. Creer que la prueba del nueve certifica que una cuenta está bien.
10. Empezar la suma alternada del criterio del 11 por la izquierda sin cuidar el signo (hay que empezar por las unidades, o comprobar la paridad del número de cifras).
11. En el ISBN, olvidar que la X vale 10.
12. Numerar las letras desde $A = 1$ en un cifrado que supone $A = 0$ (o al revés) y obtener un texto desplazado una posición.
13. Elegir a par o $a = 13$ en un cifrado afín: la función deja de ser invertible.
14. Pensar que un cifrado es seguro porque el atacante no conoce el método. La seguridad debe residir solo en la clave (principio de Kerckhoffs, 1883).
15. En RSA, confundir el módulo de los mensajes (n) con el módulo de los exponentes (φ).

3.17. Tres clásicos que faltaban: divisores, el teorema de Wilson y las fracciones continuas

La factorización única no solo sirve para calcular máximos comunes divisores: permite contar y sumar los divisores de un número sin listarlos, y de ahí sale la pregunta más antigua de la aritmética, la de los números perfectos. Después, un criterio de primalidad exacto que nadie usa y todo el mundo debería conocer, y por último la manera correcta de escribir un número irracional con enteros.

Las funciones divisor y los números perfectos

Definición 3.118 (Funciones divisor y número perfecto)**.** Para $n \geq 1$ se definen

$$\sigma_0(n) = \#\{d \geq 1 : d \mid n\}, \qquad \sigma_1(n) = \sum_{d|n} d,$$

el número de divisores positivos y su suma. Si $n = p_1^{a_1} p_2^{a_2} \cdots p_r^{a_r}$ es la factorización en primos, entonces

$$\sigma_0(n) = \prod_{i=1}^{r} (a_i + 1), \qquad \sigma_1(n) = \prod_{i=1}^{r} \frac{p_i^{a_i+1} - 1}{p_i - 1} = \prod_{i=1}^{r} (1 + p_i + p_i^2 + \cdots + p_i^{a_i}).$$

La razón de ambas fórmulas es la misma: todo divisor de n se escribe de manera única como $p_1^{b_1} \cdots p_r^{b_r}$ con $0 \leq b_i \leq a_i$, luego elegir un divisor es elegir cada exponente por separado, con $a_i + 1$ opciones. Al desarrollar el producto de las sumas geométricas, cada sumando es exactamente un divisor y aparece una sola vez.

Un entero $n \geq 2$ es *perfecto* si es igual a la suma de sus divisores propios, es decir, si

$$\sigma_1(n) = 2n.$$

Ejemplo 3.119 (Las fórmulas en funcionamiento)**. (a)** $n = 12 = 2^2 \cdot 3$**.** Se tiene $\sigma_0(12) = (2+1)(1+1) = 6$, y en efecto los divisores son $1, 2, 3, 4, 6, 12$. Y

$$\sigma_1(12) = \frac{2^3 - 1}{2 - 1} \cdot \frac{3^2 - 1}{3 - 1} = 7 \cdot 4 = 28,$$

que coincide con $1 + 2 + 3 + 4 + 6 + 12 = 28$. Como $28 > 24 = 2 \cdot 12$, el 12 es *abundante*.

(b) $n = 360 = 2^3 \cdot 3^2 \cdot 5$**.** Sin listar nada: $\sigma_0(360) = 4 \cdot 3 \cdot 2 = 24$ divisores y

$$\sigma_1(360) = \frac{2^4 - 1}{1} \cdot \frac{3^3 - 1}{2} \cdot \frac{5^2 - 1}{4} = 15 \cdot 13 \cdot 6 = 1170.$$

Listar los 24 divisores y sumarlos llevaría un buen rato; la factorización lo resuelve con tres multiplicaciones.

(c) Los cuatro primeros perfectos. $\sigma_1(6) = 1 + 2 + 3 + 6 = 12 = 2 \cdot 6$; $\sigma_1(28) = 1 + 2 + 4 + 7 + 14 + 28 = 56 = 2 \cdot 28$; $\sigma_1(496) = 992 = 2 \cdot 496$ y $\sigma_1(8128) = 16256 = 2 \cdot 8128$. Los cuatro eran conocidos en la Antigüedad, y el quinto, 33 550 336, no se documentó hasta el siglo XV.

Teorema 3.120 (Euclides–Euler)**.** *Un número par n es perfecto si y solo si*

$$n = 2^{p-1}(2^p - 1) \quad \textit{con } 2^p - 1 \textit{ primo},$$

lo que obliga a que p sea primo. Los primos de la forma $2^p - 1$ se llaman de Mersenne.

Demostración. Solo la implicación fácil, que es la de Euclides. Sea $q = 2^p - 1$ primo y $n = 2^{p-1}q$. Como q es impar, la factorización de n es $2^{p-1} \cdot q$ con los dos primos distintos, y la fórmula multiplicativa da

$$\sigma_1(n) = (2^p - 1)(1 + q) = q \cdot 2^p = 2 \cdot 2^{p-1}q = 2n.$$

La otra implicación, que *todo* par perfecto tiene esa forma, es de Euler y llegó dos mil años después. □

Observación 3.121. Los cuatro perfectos del ejemplo son $6 = 2^1(2^2-1) = 2\cdot 3$, $28 = 2^2(2^3-1) = 4\cdot 7$, $496 = 2^4(2^5-1) = 16\cdot 31$ y $8128 = 2^6(2^7-1) = 64\cdot 127$, con $p = 2, 3, 5, 7$ y los primos de Mersenne $3, 7, 31, 127$. La correspondencia es exacta: buscar pares perfectos y buscar primos de Mersenne son literalmente el mismo problema, y por eso los récords de primos grandes se anuncian siempre en esa forma.

Y aquí está el hueco más viejo de las matemáticas: nadie sabe si existe algún número perfecto *impar*. Se ha comprobado que no lo hay por debajo de 10^{1500} y se han demostrado docenas de condiciones que tendría que cumplir —al menos nueve factores primos distintos, un factor primo mayor que 10^8—, pero ni una demostración de que no existe ni un ejemplo. La pregunta lleva abierta desde Euclides, unos veintitrés siglos, y sigue siendo elemental de enunciar y aparentemente inaccesible.

El teorema de Wilson

Teorema 3.122 (Wilson)**.** *Si p es primo, entonces*

$$(p-1)! \;\equiv\; -1 \pmod{p}.$$

Recíprocamente, si $n > 1$ cumple $(n-1)! \equiv -1 \pmod{n}$, entonces n es primo. El criterio es, por tanto, una caracterización exacta de la primalidad.

Demostración. Los casos $p = 2$ y $p = 3$ se comprueban a mano: $1! = 1 \equiv -1 \pmod 2$ y $2! = 2 \equiv -1 \pmod 3$. Sea $p \geq 5$.

Cada elemento tiene inverso. Los números $1, 2, \ldots, p-1$ son todos coprimos con p, luego cada uno tiene un inverso módulo p, que además está en esa misma lista. La aplicación $a \mapsto a^{-1}$ es, pues, una permutación de $\{1, \ldots, p-1\}$, y es su propia inversa.

Quiénes son sus propios inversos. Un elemento cumple $a \equiv a^{-1}$ exactamente cuando $a^2 \equiv 1 \pmod p$, es decir, cuando $p \mid a^2 - 1 = (a-1)(a+1)$. Como p es primo, el lema de Euclides obliga a que $p \mid a-1$ o $p \mid a+1$, o sea $a \equiv 1$ o $a \equiv -1 \equiv p-1$. Y estos dos son distintos porque $p \geq 3$. De modo que exactamente dos elementos son sus propios inversos, y los otros $p-3$ se agrupan en $(p-3)/2$ parejas $\{a, a^{-1}\}$ disjuntas, con $a \neq a^{-1}$.

Se multiplica todo. En el producto $(p-1)! = 1 \cdot 2 \cdots (p-1)$, cada una de esas parejas aporta $a \cdot a^{-1} \equiv 1$. Sobreviven solo los dos elementos sueltos:

$$(p-1)! \;\equiv\; 1 \cdot (p-1) \;\equiv\; -1 \pmod{p}.$$

El recíproco. Supóngase n compuesto, con un divisor d tal que $1 < d < n$. Entonces d es uno de los factores del producto $(n-1)! = 1 \cdot 2 \cdots (n-1)$, luego $d \mid (n-1)!$. Si además se tuviera $(n-1)! \equiv -1 \pmod n$, entonces $n \mid (n-1)! + 1$ y con más razón $d \mid (n-1)! + 1$; restando, $d \mid 1$, lo que contradice $d > 1$. Por tanto ningún compuesto satisface la congruencia. □

Ejemplo 3.123 (Comprobaciones y el caso $n = 4$)**.** Con $p = 7$: $6! = 720 = 102 \cdot 7 + 6$, luego $6! \equiv 6 \equiv -1 \pmod 7$. Con $p = 11$: $10! = 3\,628\,800$ y su resto módulo 11 es $10 \equiv -1$. Con $p = 13$ sale 12, otra vez -1.

Para los compuestos el producto se hunde a cero, porque los factores de n ya están dentro del factorial: $(n-1)! \equiv 0 \pmod{n}$ para $n = 6, 8, 9, 10, 12$. La única excepción es $n = 4$, donde $3! = 6 \equiv 2 \pmod{4}$: el argumento del recíproco toma $d = 2$ y concluye $2 \mid 1$, absurdo, y en efecto el resto no es $-1 \equiv 3$ sino 2. El caso $n = 4$ es especial porque su único divisor propio aparece una sola vez en $3! = 1 \cdot 2 \cdot 3$, mientras que en todo otro compuesto $n = ab$ con $a \neq b$ ambos factores aparecen por separado.

Cuidado

Wilson es un criterio de primalidad *exacto*, sin excepciones y sin probabilidad de por medio, a diferencia de Fermat, que tiene los números de Carmichael, o de Miller–Rabin, que es probabilístico. Y aun así no se usa jamás: calcular $(n-1)!$ módulo n exige $n-2$ multiplicaciones, mientras que la simple división de prueba necesita solo $\sqrt{n}$ divisiones y Miller–Rabin unas pocas exponenciaciones. Es el ejemplo perfecto de que ser correcto no basta: en algoritmia, la caracterización más limpia puede ser la más inútil.

Fracciones continuas

Definición 3.124 (Fracción continua)**.** Una *fracción continua simple* es una expresión

$$a_0 + \cfrac{1}{a_1 + \cfrac{1}{a_2 + \cfrac{1}{a_3 + \cdots}}}$$

con $a_0 \in \mathbb{Z}$ y $a_i \geq 1$ enteros para $i \geq 1$. Se abrevia

$$[a_0; a_1, a_2, a_3, \dots],$$

con punto y coma tras la parte entera y comas después. Los truncamientos $[a_0; a_1, \dots, a_k]$ se llaman *convergentes* y son las fracciones h_k/ℓ_k que mejor aproximan al número con denominador acotado.

El desarrollo es finito si y solo si el número es racional. Es *periódico* si y solo si el número es un irracional cuadrático, es decir, raíz de un polinomio de grado dos con coeficientes enteros, y ese es el teorema de Lagrange. El periodo se marca con una barra: $\sqrt{2} = [1; \overline{2}]$ significa $[1; 2, 2, 2, \dots]$, y el número áureo vale $(1+\sqrt{5})/2 = [1; \overline{1}]$.

Regla: desarrollar un número en fracción continua

1. **Tómese la parte entera.** Si x_0 es el número de partida, hágase $a_0 = \lfloor x_0 \rfloor$.
2. **Réstese y compruébese.** Si $x_0 - a_0 = 0$, el desarrollo ha terminado y el número era racional.
3. **Inviértase la parte fraccionaria:** $x_1 = 1/(x_0 - a_0)$, que siempre es mayor que 1 y por tanto tiene parte entera $a_1 \geq 1$.
4. **Repítase** con x_1, luego con x_2, y así sucesivamente.

5. **Para una raíz cuadrada, racionalícese en cada paso** multiplicando numerador y denominador por el conjugado, y **deténgase en cuanto reaparezca un x_i ya visto**: a partir de ahí todo se repite y ese tramo es el periodo.
6. **Compruébese el resultado** calculando los convergentes con la recurrencia $h_k = a_k h_{k-1} + h_{k-2}$ y $\ell_k = a_k \ell_{k-1} + \ell_{k-2}$, arrancando de $h_{-1} = 1$, $h_{-2} = 0$, $\ell_{-1} = 0$, $\ell_{-2} = 1$, y viendo que se acercan al valor buscado alternando por defecto y por exceso.

Ejemplo 3.125 (El desarrollo de $\sqrt{7}$, con la comprobación)**. Paso a paso.** Como $2^2 = 4 < 7 < 9 = 3^2$, se tiene $a_0 = 2$ y $x_0 = \sqrt{7} = 2{,}645\,751\ldots$

$$x_1 = \frac{1}{\sqrt{7}-2} = \frac{\sqrt{7}+2}{3} = 1{,}548\,58\ldots, \qquad a_1 = 1;$$

$$x_2 = \frac{1}{x_1 - 1} = \frac{3}{\sqrt{7}-1} = \frac{\sqrt{7}+1}{2} = 1{,}822\,88\ldots, \qquad a_2 = 1;$$

$$x_3 = \frac{1}{x_2 - 1} = \frac{2}{\sqrt{7}-1} = \frac{\sqrt{7}+1}{3} = 1{,}215\,25\ldots, \qquad a_3 = 1;$$

$$x_4 = \frac{1}{x_3 - 1} = \frac{3}{\sqrt{7}-2} = \sqrt{7}+2 = 4{,}645\,75\ldots, \qquad a_4 = 4;$$

$$x_5 = \frac{1}{x_4 - 4} = \frac{1}{\sqrt{7}-2} = x_1.$$

Ha reaparecido x_1, luego el bloque $1, 1, 1, 4$ se repite para siempre y

$$\sqrt{7} = [2; \overline{1, 1, 1, 4}].$$

En cada paso se ha racionalizado: por ejemplo $3/(\sqrt{7}-1) = 3(\sqrt{7}+1)/6 = (\sqrt{7}+1)/2$, usando que $(\sqrt{7}-1)(\sqrt{7}+1) = 6$.

La verificación numérica. Los convergentes, con la recurrencia del paso 6, son

$$2, \quad 3, \quad \frac{5}{2}, \quad \frac{8}{3}, \quad \frac{37}{14}, \quad \frac{45}{17}, \quad \frac{82}{31}, \quad \frac{127}{48},$$

cuyos valores decimales son 2, 3, 2,5, 2,666 7, 2,642 857, 2,647 059, 2,645 161 y 2,645 833, frente a $\sqrt{7} = 2{,}645\,751$. Se ve la alternancia anunciada —por defecto, por exceso, por defecto…— y la convergencia: el último se equivoca en menos de 10^{-4} con un denominador de dos cifras.

Y una comprobación algebraica exacta. Los convergentes que preceden al final de cada periodo resuelven la ecuación de Pell $x^2 - 7y^2 = 1$:

$$8^2 - 7 \cdot 3^2 = 64 - 63 = 1, \qquad 127^2 - 7 \cdot 48^2 = 16\,129 - 16\,128 = 1.$$

Es decir, $(8, 3)$ es la solución fundamental de Pell para $N = 7$ y $(127, 48)$ la siguiente, lo que cierra el círculo con las ecuaciones diofánticas vistas antes: las soluciones de Pell *son* los convergentes de la fracción continua del radical.

RSA, condensado en una regla

La sección anterior explicó de dónde sale el RSA y por qué funciona, con los primos de juguete $p = 5$ y $q = 11$. Lo que sigue no repite la justificación: la comprime en un procedimiento que se pueda seguir sin releer nada, con las comprobaciones que hay que hacer y en qué orden, y la aplica al ejemplo canónico de la literatura, con números lo bastante grandes como para que se vea por qué hacen falta la exponenciación binaria y el teorema chino del resto.

Regla: RSA paso a paso

Generación de claves, que hace quien va a recibir el mensaje.

1. Elíjanse dos primos distintos y grandes, p y q, y calcúlese el *módulo* $n = pq$.
2. Calcúlese $\varphi(n) = (p-1)(q-1)$. Este número es secreto: obtenerlo sin conocer p y q equivale a factorizar n.
3. Elíjase el *exponente público* e con $1 < e < \varphi(n)$ y $\gcd(e, \varphi(n)) = 1$.
4. Calcúlese el *exponente privado* d como el inverso de e módulo $\varphi(n)$, con el algoritmo de Euclides extendido: $ed \equiv 1 \pmod{\varphi(n)}$.
5. Publíquese la clave pública (n, e) y guárdense d, p y q. Destrúyanse p y q si no se van a usar para acelerar el descifrado.

Cifrado. Codifíquese el mensaje como un entero m con $0 \le m < n$ y calcúlese $c = m^e \bmod n$ por exponenciación binaria.
Descifrado. Calcúlese $m = c^d \bmod n$, de nuevo por exponenciación binaria.
Verificación, que no debe saltarse nunca al montar el sistema:

6. Compruébese que $\gcd(e, \varphi(n)) = 1$; si no, el paso 4 no tiene solución.
7. Compruébese que $ed \bmod \varphi(n)$ vale exactamente 1.
8. Cífrese y descífrese un mensaje de prueba y véase que se recupera el original. Si falla, el error está casi siempre en d.

Ejemplo 3.126 (El ejemplo canónico, resuelto entero)**. Claves.** Tómense $p = 61$ y $q = 53$. Entonces

$$n = 61 \cdot 53 = 3233, \qquad \varphi(n) = 60 \cdot 52 = 3120.$$

Se elige $e = 17$, que es válido porque $\gcd(17, 3120) = 1$. Para hallar d se corre Euclides:

$$3120 = 183 \cdot 17 + 9, \qquad 17 = 1 \cdot 9 + 8, \qquad 9 = 1 \cdot 8 + 1, \qquad 8 = 8 \cdot 1 + 0.$$

Y se deshace hacia atrás:

$$1 = 9 - 8 = 9 - (17 - 9) = 2 \cdot 9 - 17 = 2(3120 - 183 \cdot 17) - 17 = 2 \cdot 3120 - 367 \cdot 17.$$

Luego $-367 \cdot 17 \equiv 1 \pmod{3120}$ y $d = 3120 - 367 = 2753$. Verificación del paso 7: $17 \cdot 2753 = 46\,801 = 15 \cdot 3120 + 1$, y en efecto el resto es 1. La clave pública es $(3233, 17)$ y la privada, 2753.

Cifrado de $m = 65$. Por exponenciación binaria, con $17 = 16 + 1$ y reduciendo módulo 3233 en cada paso:

$$65^2 = 4225 = 1 \cdot 3233 + 992, \qquad 992^2 = 984\,064 = 304 \cdot 3233 + 1232,$$

$$1232^2 = 1\,517\,824 = 469 \cdot 3233 + 1547, \qquad 1547^2 = 2\,393\,209 = 740 \cdot 3233 + 789.$$

Es decir, $65^2 \equiv 992$, $65^4 \equiv 1232$, $65^8 \equiv 1547$ y $65^{16} \equiv 789$. Finalmente

$$c = 65^{17} = 65^{16} \cdot 65 \equiv 789 \cdot 65 = 51\,285 = 15 \cdot 3233 + 2790,$$

de modo que $c = 2790$. Cinco multiplicaciones en lugar de dieciséis, y ningún número por encima de siete cifras.

Descifrado. Habría que calcular 2790^{2753} mód 3233, y con la exponenciación binaria son unas dieciocho multiplicaciones. Pero quien posee p y q tiene un atajo mucho mejor, el teorema chino del resto: se reduce todo módulo 61 y módulo 53 por separado. Como $2790 = 45 \cdot 61 + 45$ y $2753 = 45 \cdot 60 + 53$,

$$m \equiv 45^{53} \equiv 4 \pmod{61},$$

y como $2790 = 52 \cdot 53 + 34$ y $2753 = 52 \cdot 52 + 49$,

$$m \equiv 34^{49} \equiv 12 \pmod{53}.$$

Queda recomponer. El inverso de 53 módulo 61 es 38, porque $53 \cdot 38 = 2014 = 33 \cdot 61 + 1$; entonces $h = 38 \cdot (4 - 12)$ mód $61 = 38 \cdot (-8)$ mód $61 = 1$ y

$$m = 12 + 1 \cdot 53 = 65.$$

Se recupera el mensaje original, que era 65. La comprobación directa también funciona: $2790^{2753} \equiv 65 \pmod{3233}$. El atajo por el teorema chino es unas cuatro veces más rápido, y es exactamente lo que hace cualquier implementación real cuando conserva p y q.

Un detalle que suele omitirse. El paso 4 admite usar el mínimo común múltiplo $\lambda(n) = \operatorname{mcm}(p-1, q-1)$ en lugar de $\varphi(n)$. Aquí se tiene $\lambda(3233) = \operatorname{mcm}(60, 52) = 780$, y el inverso de 17 módulo 780 es 413, mucho menor que 2753. Y también descifra: $2790^{413} \equiv 65 \pmod{3233}$. Todo exponente privado válido lo es módulo $\lambda(n)$, y usar el más pequeño ahorra trabajo sin perder nada; es lo que recogen las normas modernas.

Cuidado

Los números de este ejemplo son de juguete: 3233 se factoriza mentalmente. Un módulo real tiene hoy 2048 bits, unos 617 dígitos decimales, y aun así el esquema tal cual está descrito es inseguro. Le faltan dos cosas. La primera, un *relleno* aleatorio: cifrar m sin más es determinista, de modo que un atacante que sospeche el mensaje puede cifrarlo él mismo y comparar, y mensajes pequeños con e pequeño ni siquiera desbordan el módulo y se recuperan con una raíz e-ésima ordinaria. La segunda, que p y q deben generarse al azar, ser de tamaños parecidos pero no demasiado próximos —si $|p - q|$ es pequeño, el método de Fermat factoriza n en un suspiro— y no reutilizarse jamás entre claves distintas.

3.18. Ejercicios resueltos

Ejercicio 3.1. Calcula $\gcd(1071, 462)$ con el algoritmo de Euclides y exprésalo como combinación entera $1071x + 462y$.

Solución. Dividimos repetidamente, anotando cada línea:

$$1071 = 2 \cdot 462 + 147$$
$$462 = 3 \cdot 147 + 21$$
$$147 = 7 \cdot 21 + 0.$$

El último resto no nulo es 21, así que $\gcd(1071, 462) = 21$. Ahora subimos deshaciendo, para hallar los coeficientes:

$$21 = 462 - 3 \cdot 147 = 462 - 3 \cdot (1071 - 2 \cdot 462) = 7 \cdot 462 - 3 \cdot 1071.$$

Es decir, $1071 \cdot (-3) + 462 \cdot 7 = 21$. Comprobación rápida: $-3213 + 3234 = 21$. ✓ Observa que el método no adivinó los coeficientes: salieron solos de reescribir las divisiones al revés. ■

Ejercicio 3.2. Resuelve la congruencia $6x \equiv 4 \pmod 8$. ¿Cuántas soluciones hay módulo 8?

Solución. Primero miramos $\gcd(6, 8) = 2$. Como 2 divide a 4, hay solución (si no dividiera, no habría ninguna). La teoría anticipa que habrá exactamente 2 soluciones módulo 8. Para hallarlas, dividimos toda la congruencia entre 2, incluyendo el módulo:

$$3x \equiv 2 \pmod 4.$$

Ahora $\gcd(3, 4) = 1$, así que 3 tiene inverso módulo 4; como $3 \cdot 3 = 9 \equiv 1$, el inverso es 3. Multiplicando, $x \equiv 3 \cdot 2 = 6 \equiv 2 \pmod 4$. Es decir, $x \equiv 2 \pmod 4$, lo que módulo 8 significa $x \equiv 2$ o $x \equiv 6$. Verificamos ambas: $6 \cdot 2 = 12 \equiv 4$ ✓ y $6 \cdot 6 = 36 \equiv 4$ ✓. Hay dos soluciones, tal como predijo el máximo común divisor. ■

Ejercicio 3.3. Encuentra el resto de 3^{100} al dividir entre 7.

Solución. Como 7 es primo y 3 no es múltiplo de 7, el pequeño teorema de Fermat da $3^6 \equiv 1 \pmod 7$. Dividimos el exponente entre 6: $100 = 6 \cdot 16 + 4$. Entonces

$$3^{100} = \left(3^6\right)^{16} \cdot 3^4 \equiv 1 \cdot 3^4 \pmod 7.$$

Falta $3^4 = 81$; y $81 = 7 \cdot 11 + 4$, así que $3^4 \equiv 4$. El resto es 4. (Comprobación de cordura: no calculamos 3^{100}, que tiene 48 dígitos, y aun así conocemos su resto exacto. Ese es el regalo de Fermat.) ■

Ejercicio 3.4. Un pirata reparte monedas de oro. En grupos de 7 sobran 6; en grupos de 9 sobran 4. ¿Cuál es la menor cantidad de monedas posible?

Solución. Buscamos el menor $x > 0$ con $x \equiv 6 \pmod 7$ y $x \equiv 4 \pmod 9$. Como 7 y 9 son coprimos, el teorema chino del resto garantiza solución única módulo 63. La construimos: escribimos $x = 4 + 9t$ (asegura la segunda condición) y forzamos la primera, $4 + 9t \equiv 6 \pmod 7$. Como $9 \equiv 2 \pmod 7$, esto es $4 + 2t \equiv 6$, o sea $2t \equiv 2$, luego $t \equiv 1 \pmod 7$ (el inverso de 2 módulo 7 es 4, y $4 \cdot 2 = 8 \equiv 1$; aplicándolo, $t \equiv 4 \cdot 2 = 8 \equiv 1$). Tomando $t = 1$: $x = 4 + 9 = 13$. Verificamos: $13 = 7 + 6$ (resto 6 módulo 7 ✓) y $13 = 9 + 4$ (resto 4 módulo 9 ✓). La menor cantidad es 13 monedas. ■

Ejercicio 3.5. En el RSA de juguete del capítulo ($p = 5$, $q = 11$, $n = 55$, $e = 3$, $d = 27$), cifra el mensaje $M = 9$ y luego descífralo, comprobando que recuperas el 9.

Solución. *Cifrado:* $C = 9^3$ mód 55. Reducimos en cada paso para no arrastrar números grandes: $9^2 = 81 \equiv 26$ (mód 55), y entonces $9^3 \equiv 9 \cdot 26 = 234 \equiv 14$ (mód 55) (pues $234 = 4 \cdot 55 + 14$). Así $C = 14$.

Descifrado: $M = 14^{27}$ mód 55. Reducimos por cuadrados sucesivos: $14^2 = 196 \equiv 31$ (pues $196 - 3 \cdot 55 = 196 - 165 = 31$); $14^4 \equiv 31^2 = 961 \equiv 26$ (pues $961 - 17 \cdot 55 = 961 - 935 = 26$); $14^8 \equiv 26^2 = 676 \equiv 16$ (pues $676 - 12 \cdot 55 = 676 - 660 = 16$); $14^{16} \equiv 16^2 = 256 \equiv 36$ (pues $256 - 4 \cdot 55 = 36$). Como $27 = 16 + 8 + 2 + 1$:

$$14^{27} \equiv 36 \cdot 16 \cdot 31 \cdot 14 \pmod{55}.$$

Multiplicamos por pasos: $36 \cdot 16 = 576 \equiv 26$ ($576 - 10 \cdot 55 = 26$); $26 \cdot 31 = 806 \equiv 36$ ($806 - 14 \cdot 55 = 806 - 770 = 36$); $36 \cdot 14 = 504 \equiv 9$ ($504 - 9 \cdot 55 = 504 - 495 = 9$). Recuperamos $M = 9$. ✓ El teorema de Euler, en acción, garantizaba de antemano que esto tenía que salir. ■

Ejercicio 3.6. Calcule 2^{100} mód 7.

Solución. $2^1 \equiv 2$, $2^2 \equiv 4$, $2^3 \equiv 1$ (mód 7): el ciclo tiene longitud 3. Como $100 = 3 \cdot 33 + 1$, es decir $100 \equiv 1$ (mód 3), resulta $2^{100} \equiv 2^1 = 2$ (mód 7). ■

Ejercicio 3.7. Halle la última cifra de 3^{2026}.

Solución. Módulo 10: $3, 9, 7, 1$ y vuelta a empezar; ciclo de longitud 4. Como $2026 = 4 \cdot 506 + 2$, se tiene $3^{2026} \equiv 3^2 = 9$ (mód 10). Termina en 9. ■

Ejercicio 3.8. Resuelva $8x \equiv 12$ (mód 20).

Solución. $d = \mathrm{mcd}(8, 20) = 4$ y $4 \mid 12$, luego hay 4 soluciones. Dividiendo todo entre 4 (módulo incluido): $2x \equiv 3$ (mód 5). Como $2^{-1} \equiv 3$ (mód 5), sale $x \equiv 9 \equiv 4$ (mód 5). Las soluciones módulo 20 son $x = 4$, 9, 14, 19. Comprobación con $x = 9$: $72 = 3 \cdot 20 + 12$. ✓ ■

Ejercicio 3.9. Un número de tres cifras es divisible entre 9 y entre 11, y su cifra de las centenas es 5. ¿Cuál es?

Solución. Ser divisible entre 9 y entre 11 con $\mathrm{mcd}(9, 11) = 1$ equivale a serlo entre 99. Los múltiplos de 99 de tres cifras son $198, 297, 396, 495, 594, \ldots$ El que empieza por 5 es **594**. En efecto, $5 + 9 + 4 = 18$ es múltiplo de 9 y $4 - 9 + 5 = 0$ es múltiplo de 11. ■

Ejercicio 3.10. Descifre `KROD`, sabiendo que se ha usado un César con clave 3.

Solución. Se resta 3 a cada letra: K(10) → 7 =H, R(17) → 14 =O, O(14) → 11 =L, D(3) → 0 =A. El mensaje es HOLA. ■

Ejercicio 3.11. Con la clave pública $(n, e) = (143, 7)$, cifre el mensaje $M = 9$ y compruebe que se recupera con la clave privada $d = 103$.

Solución. $143 = 11 \cdot 13$, luego $\varphi = 10 \cdot 12 = 120$; y en efecto $7 \cdot 103 = 721 = 6 \cdot 120 + 1$, así que $d = 103$ es correcta. Cifrado: $9^2 = 81$, $9^4 \equiv 81^2 = 6561 = 45 \cdot 143 + 126 \equiv 126$, y $9^7 = 9^4 \cdot 9^2 \cdot 9 \equiv 126 \cdot 81 \cdot 9$. Reduciendo por pasos, $126 \cdot 81 = 10\,206 = 71 \cdot 143 + 53 \equiv 53$ y $53 \cdot 9 = 477 = 3 \cdot 143 + 48 \equiv 48$. Luego $C = 48$. Descifrando, $48^{103} \equiv 9$ (mód 143), como debía ser. ■

3.19. Ejercicios propuestos

Sin solución desarrollada: solo enunciado y, al final, la respuesta. Aquí muchas respuestas son un nombre (“el grupo de Klein”, “sí, por Eisenstein”) más que un número: comprueba que coincida el fondo del argumento, no la redacción.

3.1. Calcula $\gcd(1071, 462)$ con el algoritmo de Euclides y expresa el resultado como $1071u + 462v$.

3.2. Factoriza 2024 en primos.

3.3. Calcula 3^{100} mód 7 usando el pequeño teorema de Fermat.

3.4. Halla el inverso de 7 módulo 26.

3.5. Resuelve el sistema $x \equiv 2\ (3)$, $x \equiv 3\ (5)$, $x \equiv 2\ (7)$.

3.6. Calcula $\varphi(36)$ y $\varphi(3120)$.

3.7. ¿Cuál es el orden de 3 en $(\mathbb{Z}/7\mathbb{Z})^\times$? ¿Es 3 raíz primitiva módulo 7?

3.8. Comprueba el teorema de Wilson con $p = 7$: ¿cuánto vale $6!$ mód 7?

3.9. Calcula el símbolo de Legendre $\left(\frac{5}{7}\right)$ y di si 5 es residuo cuadrático módulo 7.

3.10. Halla $\sigma_0(28)$ y $\sigma_1(28)$. ¿Es 28 perfecto?

3.11. Da el desarrollo en fracción continua de $\sqrt{7}$.

3.12. En RSA con $p = 61$, $q = 53$ y $e = 17$: calcula n, $\varphi(n)$ y la clave privada d.

Respuestas

1. 21; por ejemplo $21 = 1071 \cdot (-3) + 462 \cdot 7$. **2.** $2024 = 2^3 \cdot 11 \cdot 23$. **3.** 4. ($3^6 \equiv 1$, y $100 = 6 \cdot 16 + 4$, luego $3^{100} \equiv 3^4 = 81 \equiv 4$.) **4.** 15 (pues $7 \cdot 15 = 105 = 4 \cdot 26 + 1$). **5.** $x \equiv 23$ (mód 105). **6.** $\varphi(36) = 12$ y $\varphi(3120) = 768$. **7.** Orden 6; sí, es raíz primitiva (el grupo tiene 6 elementos). **8.** $6! \equiv 6 \equiv -1$ (mód 7), como predice Wilson. **9.** $\left(\frac{5}{7}\right) = -1$: no es residuo cuadrático. **10.** $\sigma_0(28) = 6$, $\sigma_1(28) = 56 = 2 \cdot 28$; sí, es perfecto. **11.** $\sqrt{7} = [2; \overline{1, 1, 1, 4}]$. **12.** $n = 3233$, $\varphi(n) = 60 \cdot 52 = 3120$, $d = 2753$.

3.1. Calcule (-23) mód 7 y (-23) mód 11.

3.2. ¿En qué cifra termina 9^{99}?

3.3. Calcule 5^{200} mód 13.

3.4. ¿Es 1 234 567 890 divisible entre 9? ¿Y entre 11?

3.5. Halle la cifra que falta para que 47_3 sea múltiplo de 9.

3.6. Compruebe con la prueba del nueve si $567 \times 432 = 245\,344$ puede ser correcto.

3.7. Resuelva $9x \equiv 6\ (\text{mód } 15)$.

3.8. Halle 11^{-1} módulo 26.

3.9. Resuelva el sistema $x \equiv 1\ (\text{mód } 4)$, $x \equiv 2\ (\text{mód } 9)$.

3.10. Cifre MATEMATICAS con un César de clave 5.

3.11. ¿Cuántas claves distintas tiene un cifrado afín módulo 27 (alfabeto con ñ)?

3.12. Si hoy es jueves, ¿qué día de la semana será dentro de 10^6 días?

Respuestas

1. $-23 \equiv 5 \pmod 7$ y $-23 \equiv 10 \pmod{11}$. **2.** Ciclo de 9 módulo 10: $9, 1$; como 99 es impar, termina en **9**. **3.** $5^4 \equiv 1 \pmod{13}$ y $200 \equiv 0 \pmod 4$, luego $5^{200} \equiv 1$. **4.** Suma de cifras $= 45$, múltiplo de 9: sí. Suma alternada $= -5$: no es múltiplo de 11. **5.** La cifra 4 ($4+7+4+3 = 18$). **6.** $567 \to 0$, $432 \to 0$, producto $\to 0$; pero $245\,344 \to 2+4+5+3+4+4 = 22 \to 4 \neq 0$: **está mal** (lo correcto es 244 944). **7.** $d = 3$, tres soluciones: $x \equiv 4, 9, 14 \pmod{15}$. **8.** $11^{-1} \equiv 19 \pmod{26}$ (porque $11 \cdot 19 = 209 = 8 \cdot 26 + 1$). **9.** $x \equiv 29 \pmod{36}$. **10.** RFYJRFYNHFX. **11.** Los primos con $27 = 3^3$ son los no múltiplos de 3: 18 valores de a, por 27 de b, dan **486** claves. **12.** $10^6 \equiv 1 \pmod 7$ (porque $10 \equiv 3$ y $3^6 \equiv 1$), luego **viernes**.

Capítulo 4

Números complejos desde el principio

Lo que necesitas traer

Del capítulo 1: el lenguaje de conjuntos y funciones. Conviene tener a mano las razones trigonométricas del seno y el coseno, aunque las identidades que hagan falta se deducen aquí mismo con la fórmula de Euler. No se usa cálculo: la exponencial compleja se introduce de forma algebraica y geométrica.

Este capítulo cuenta cómo la matemática se vio obligada a inventar un número cuyo cuadrado es -1, y qué ocurrió cuando lo hizo. La versión que se cuenta en la escuela —"se inventaron para resolver $x^2 + 1 = 0$"— es falsa y además es mala pedagogía, porque nadie necesita resolver esa ecuación: si no tiene solución, se dice que no la tiene y no pasa nada. La historia verdadera es mucho mejor y la contamos abajo. Todo lo que hay aquí es escolar y solo requiere el álgebra del álgebra elemental, la trigonometría de la trigonometría elemental y las coordenadas de la geometría analítica. El tratamiento serio —funciones holomorfas, integrales de contorno, residuos— es el capítulo 30; la razón estructural profunda de por qué $\mathbb{C}$ existe y es único aparece en los capítulos 24 y 26.

Idea clave

Intuición primero. Un número complejo no es un número "imaginario" en el sentido de irreal. Es un *punto del plano* al que se le ha enseñado a multiplicarse. Los números reales son los puntos de una recta y saben sumarse y multiplicarse; los complejos son los puntos de un plano y también saben, y la multiplicación resulta significar algo geométricamente precioso: **girar y estirar**.

Multiplicar por 2 estira al doble. Multiplicar por -1 gira media vuelta. Y entonces, ¿qué gira un cuarto de vuelta? Un número que, aplicado dos veces, gira media vuelta: es decir, un número cuyo cuadrado es -1. Ese número no es un truco de notación, es el giro de $90°$. Se llama i, y toda la aritmética compleja se vuelve transparente en cuanto uno lo ve así.

4.1. Por qué hicieron falta: las cúbicas de Cardano

En el siglo XVI, Cardano y Tartaglia encontraron una fórmula para resolver la ecuación de tercer grado, análoga a la de la cuadrática. Aplicada a

$$x^3 = 15x + 4,$$

la fórmula produce

$$x = \sqrt[3]{2 + \sqrt{-121}} + \sqrt[3]{2 - \sqrt{-121}}.$$

Aquí viene lo interesante: esa ecuación **sí tiene** solución real, y muy sencilla. Comprobemos: $4^3 = 64$ y $15 \cdot 4 + 4 = 64$ ✓. La solución es $x = 4$ (las otras dos son $-2 \pm \sqrt{3}$, también reales). Es decir: la fórmula da la respuesta correcta, pero el *camino* para llegar a ella pasa obligatoriamente por raíces de números negativos que después se cancelan solas.

Si uno acepta trabajar con $\sqrt{-121} = 11i$ como si fuera un número normal, y observa que

$$(2 + i)^3 = 2 + 11i \quad \text{y} \quad (2 - i)^3 = 2 - 11i,$$

entonces la fórmula da $x = (2 + i) + (2 - i) = 4$. *Exacto.*

Idea clave

Esta es la lección histórica y conviene subrayarla: los complejos no se aceptaron para resolver ecuaciones sin solución, sino porque eran *inevitables como paso intermedio* para resolver ecuaciones cuyas soluciones eran reales todo el tiempo. Bombelli, hacia 1572, fue el primero en tragarse el sapo y calcular con ellos sistemáticamente. Los llamó "cantidades sofísticas" porque él tampoco creía en ellos; solo constató que funcionaban. Doscientos cincuenta años después, Gauss y Argand explicaron por qué: eran puntos del plano desde el principio.

Cuidado

El adjetivo "imaginario" es un accidente histórico desafortunado, acuñado por Descartes con intención despectiva. Los complejos son exactamente tan reales (o tan poco) como los negativos, que también costaron siglos de aceptación, o como los irracionales, que a los pitagóricos les parecieron un escándalo. Todos son construcciones; lo que las legitima es que sirven y que son consistentes.

4.2. Qué es exactamente un número complejo, cómo se opera con él, cómo se dibuja y cómo se escribe en forma polar y exponencial

La sección anterior ha contado por qué hicieron falta. Esta monta la maquinaria entera desde cero y sin dar nada por sabido: qué objeto es un número complejo, por qué sus operaciones son las que son y no otras, cómo se suman, se restan, se multiplican y —sobre todo— se dividen dos de ellos, cómo se dibujan en el plano, y cómo se escriben en las dos formas alternativas, la

polar y la exponencial, que son las que convierten las potencias y las raíces en algo trivial. Las secciones siguientes retoman cada uno de estos objetos con más ejemplos, más aplicaciones y las demostraciones que aquí se dejan indicadas; lo que aquí se hace es fijar cada definición y resolver un ejemplo completo de cada operación.

Qué es un número complejo

Definición 4.1 (Los complejos como puntos del plano con una multiplicación)**.** Un **número complejo** es un par ordenado (a, b) de números reales, y el conjunto de todos ellos se denota $\mathbb{C}$. En $\mathbb{C}$ se definen dos operaciones:

$$(a, b) + (c, d) = (a + c,\ b + d), \qquad (a, b) \cdot (c, d) = (ac - bd,\ ad + bc).$$

Los pares de la forma $(a, 0)$ se suman y se multiplican exactamente como los reales a, de modo que se identifican con ellos y se escribe a en lugar de $(a, 0)$; en particular $0 = (0, 0)$ y $1 = (1, 0)$. Al par $(0, 1)$ se le llama **unidad imaginaria** y se le escribe i. Con la regla del producto,

$$i^2 = (0, 1) \cdot (0, 1) = (0 \cdot 0 - 1 \cdot 1,\ 0 \cdot 1 + 1 \cdot 0) = (-1, 0) = -1.$$

Como $(a, b) = (a, 0) + (b, 0) \cdot (0, 1)$, todo complejo se escribe $z = a + bi$: es la *forma binómica*, que es la notación que se usará siempre y que la sección siguiente recoge como definición de trabajo, junto con los nombres de a y b.

Idea clave

Qué se ha ganado y qué ha costado. Se ha ganado un número cuyo cuadrado es -1, y no por decreto: la fórmula del producto, aplicada al par $(0, 1)$, lo produce. Y no se ha perdido nada de lo que uno espera de un sistema numérico —conmutatividad, asociatividad, distributiva, existencia de inversos— salvo una cosa: el *orden*. En $\mathbb{C}$ no hay ningún "menor que" compatible con las operaciones, y por eso preguntar si $2 + i$ es mayor que $1 + 3i$ no significa nada. Se comparan los módulos, que son reales, y nada más.

Proposición 4.2 (La regla del producto no es una elección)**.** *Si se quiere que en $\mathbb{C}$ valgan las propiedades habituales del álgebra —la distributiva, la conmutatividad y la asociatividad del producto— y que además $i^2 = -1$, entonces el producto de $a + bi$ por $c + di$ tiene que ser forzosamente $(ac - bd) + (ad + bc)i$.*

Demostración. Aplicando la distributiva dos veces, $(a + bi)(c + di) = ac + adi + bci + bd\,i^2$. Sustituyendo $i^2 = -1$ en el último sumando y agrupando la parte con i y la parte sin ella,

$$(a + bi)(c + di) = (ac - bd) + (ad + bc)i.$$

No ha habido ninguna decisión: cada paso estaba obligado. Por eso no conviene memorizar esa fórmula final, sino hacer la distributiva y sustituir i^2 al terminar. □

Las cuatro operaciones en forma binómica

Regla: sumar, restar, multiplicar y dividir en forma binómica

Sean $z = a + bi$ y $w = c + di$, con a, b, c, d reales. Se opera tratando i como si fuera una letra, con la única novedad de que $i^2 = -1$.

1. **Suma y resta: parte con parte.**

$$z + w = (a + c) + (b + d)i, \qquad z - w = (a - c) + (b - d)i.$$

2. **Producto: distributiva y luego $i^2 = -1$.** Multiplica los cuatro pares de términos, cambia i^2 por -1 y agrupa:

$$zw = (ac - bd) + (ad + bc)i.$$

3. **Para dividir, primero el conjugado.** El *conjugado* de $w = c + di$ es $\bar{w} = c - di$, y la clave de todo es que

$$w\,\bar{w} = (c + di)(c - di) = c^2 - d^2 i^2 = c^2 + d^2,$$

un número *real* y no negativo, que además es $|w|^2$ con $|w| = \sqrt{c^2 + d^2}$. Es la diferencia de cuadrados de siempre, con el signo que $i^2 = -1$ regala.

4. **Multiplica numerador y denominador por el conjugado del denominador.** Es lo mismo que racionalizar un denominador con raíces:

$$\frac{z}{w} = \frac{z\,\bar{w}}{w\,\bar{w}} = \frac{(a + bi)(c - di)}{c^2 + d^2}.$$

5. **Desarrolla el numerador** como un producto normal y **reparte el denominador**, que ya es un número real, entre la parte real y la imaginaria:

$$\frac{z}{w} = \frac{ac + bd}{c^2 + d^2} + \frac{bc - ad}{c^2 + d^2}\,i.$$

6. **Nunca dejes una i en el denominador.** Una respuesta con i abajo no está terminada, exactamente igual que $1/\sqrt{2}$ no está racionalizado.

7. **Comprueba multiplicando hacia atrás**: el cociente por el divisor debe devolver el dividendo. Y como control rápido, $|z/w|$ tiene que valer $|z|/|w|$.

Ejemplo 4.3 (Las cuatro operaciones con los mismos dos números)**.** Sean $z = 4 + 3i$ y $w = 1 - 2i$.

(a) Suma y resta. $z + w = (4 + 1) + (3 - 2)i = 5 + i$ y $z - w = (4 - 1) + (3 + 2)i = 3 + 5i$. Ojo al signo: restar $-2i$ es sumar $2i$.

(b) Producto. Distributiva: $(4 + 3i)(1 - 2i) = 4 - 8i + 3i - 6i^2$. Como $-6i^2 = +6$,

agrupando queda

$$zw = (4+6) + (-8+3)i = 10 - 5i.$$

(c) Cociente. El conjugado de w es $\bar{w} = 1 + 2i$, y $w\bar{w} = 1^2 + 2^2 = 5$. Entonces

$$\frac{z}{w} = \frac{(4+3i)(1+2i)}{5} = \frac{4+8i+3i+6i^2}{5} = \frac{4+11i-6}{5} = \frac{-2+11i}{5} = -\frac{2}{5} + \frac{11}{5}i,$$

es decir $-0{,}4 + 2{,}2\,i$.

(d) Las dos comprobaciones. Multiplicando hacia atrás,

$$\left(-\tfrac{2}{5} + \tfrac{11}{5}i\right)(1-2i) = -\tfrac{2}{5} + \tfrac{4}{5}i + \tfrac{11}{5}i - \tfrac{22}{5}i^2 = \left(-\tfrac{2}{5} + \tfrac{22}{5}\right) + \left(\tfrac{4}{5} + \tfrac{11}{5}\right)i = 4 + 3i \ \checkmark$$

Y con los módulos: $|z| = \sqrt{16+9} = 5$, $|w| = \sqrt{1+4} = \sqrt{5} = 2{,}236068$, luego $|z|/|w| = \sqrt{5}$; y en efecto $\left|-\tfrac{2}{5} + \tfrac{11}{5}i\right| = \sqrt{\tfrac{4}{25} + \tfrac{121}{25}} = \sqrt{5} = 2{,}236068$ ✓.

Ejemplo 4.4 (Dos cocientes más, uno de ellos con denominador imaginario puro)**. (a)** $\dfrac{2+3i}{4-i}$. El conjugado del denominador es $4+i$ y $(4-i)(4+i) = 16+1 = 17$, luego

$$\frac{(2+3i)(4+i)}{17} = \frac{8+2i+12i+3i^2}{17} = \frac{5+14i}{17} = \frac{5}{17} + \frac{14}{17}i = 0{,}294118 + 0{,}823529\,i.$$

Comprobación: $(5+14i)(4-i)/17 = (20-5i+56i+14)/17 = (34+51i)/17 = 2+3i$ ✓.

(b) $\dfrac{3-i}{2i}$. El conjugado de $2i$ es $-2i$, y $(2i)(-2i) = -4i^2 = 4$:

$$\frac{(3-i)(-2i)}{4} = \frac{-6i+2i^2}{4} = \frac{-2-6i}{4} = -\frac{1}{2} - \frac{3}{2}i.$$

Comprobación: $\left(-\tfrac{1}{2} - \tfrac{3}{2}i\right)(2i) = -i - 3i^2 = 3 - i$ ✓. Caso particular útil: $\dfrac{1}{i} = \dfrac{-i}{-i \cdot i} = -i$, porque $i \cdot (-i) = 1$.

El plano de Argand

Definición 4.5 (Plano de Argand)**.** Se llama **plano de Argand**, o plano complejo, al plano cartesiano en el que cada número $z = a + bi$ se representa por el punto de coordenadas (a, b), o por el vector que va del origen a ese punto. El eje horizontal, formado por los complejos con $b = 0$, se llama *eje real*, y el vertical, formado por los que tienen $a = 0$ —los *imaginarios puros*—, se llama *eje imaginario*. Con esta identificación:

- sumar complejos es sumar vectores, con la regla del paralelogramo;
- $|z| = \sqrt{a^2+b^2}$ es la distancia de z al origen, por el teorema de Pitágoras;
- $|z - w|$ es la distancia entre los puntos z y w;
- conjugar es reflejar respecto del eje real, y cambiar de signo es girar media vuelta.

Ejemplo 4.6 (Leer el plano)**.** Con $z = 4 + 3i$ y $w = 1 - 2i$: el punto z está en el primer cuadrante y w en el cuarto. La distancia entre ambos es

$$|z - w| = |3 + 5i| = \sqrt{9 + 25} = \sqrt{34} = 5{,}830952.$$

El conjugado $\bar{z} = 4 - 3i$ es el reflejado de z en el eje real, y $-z = -4 - 3i$ es su simétrico respecto del origen. La suma $z + w = 5 + i$ es la diagonal del paralelogramo que forman los dos vectores. Y una condición como $|z - 4 - 3i| = 2$ describe los puntos que distan 2 de z: la circunferencia de centro $(4, 3)$ y radio 2, escrita en una línea y sin coordenadas.

Módulo, argumento y forma polar

Definición 4.7 (Módulo, argumento y argumento principal)**.** El **módulo** de $z = a + bi$ es $|z| = \sqrt{a^2 + b^2}$, su distancia al origen. Si $z \neq 0$, un **argumento** de z es cualquier ángulo θ medido desde el semieje real positivo hasta el vector z, es decir, cualquier θ con

$$a = |z| \cos\theta, \qquad b = |z| \sin\theta.$$

Hay infinitos, y se diferencian en múltiplos de 2π. Se llama **argumento principal**, y se escribe $\operatorname{Arg} z$, al único que cae en el intervalo $(-\pi, \pi]$. Con $r = |z|$ y θ cualquier argumento, la **forma polar** de z es

$$z = r(\cos\theta + i \sin\theta) = r \operatorname{cis}\theta.$$

Explícitamente, y esto es lo que hay que usar para no equivocarse de cuadrante,

$$\operatorname{Arg}(a + bi) = \begin{cases} \arctan \dfrac{b}{a} & \text{si } a > 0, \\ \arctan \dfrac{b}{a} + \pi & \text{si } a < 0 \text{ y } b \geq 0, \\ \arctan \dfrac{b}{a} - \pi & \text{si } a < 0 \text{ y } b < 0, \\ \pi/2 & \text{si } a = 0 \text{ y } b > 0, \\ -\pi/2 & \text{si } a = 0 \text{ y } b < 0. \end{cases}$$

El número 0 tiene módulo 0 y *no* tiene argumento: no apunta a ninguna parte.

Cuidado

[La arcotangente no sabe en qué cuadrante estás] La función arctan devuelve siempre un ángulo entre $-\pi/2$ y $\pi/2$, es decir, del primer o del cuarto cuadrante. Si z está a la izquierda del eje imaginario, el cociente b/a es el mismo que el del punto diametralmente opuesto y la calculadora devuelve ese otro: por eso las dos ramas con $a < 0$ llevan el $\pm\pi$. Y si $a = 0$ no hay cociente que valga, porque habría que dividir entre cero. Dibujar el punto antes de escribir el ángulo evita todos estos errores; en un ordenador, la función que hace este trabajo correctamente se llama `atan2`.

Ejemplo 4.8 (Seis argumentos principales, uno por cada situación)**.** (a) $z = 3 + 4i$: $r = \sqrt{9 + 16} = 5$ y, como $a > 0$, $\operatorname{Arg} z = \arctan \frac{4}{3} = 0{,}927295$ radianes, o sea $53{,}130102°$. Primer cuadrante.

(b) $z = -1 + i$: $r = \sqrt{2} = 1{,}414214$; aquí $a < 0$ y $b > 0$, luego $\operatorname{Arg} z = \arctan(-1) + \pi = -\frac{\pi}{4} + \pi = \frac{3\pi}{4} = 135°$. La calculadora, a secas, habría dado $-45°$: el punto opuesto.

(c) $z = -2 - 2\sqrt{3}\,i$: $r = \sqrt{4 + 12} = 4$; ahora $a < 0$ y $b < 0$, luego $\operatorname{Arg} z = \arctan\sqrt{3} - \pi = \frac{\pi}{3} - \pi = -\frac{2\pi}{3} = -120°$. Tercer cuadrante.

(d) $z = 5 - 5i$: $r = 5\sqrt{2} = 7{,}071068$ y, como $a > 0$, $\operatorname{Arg} z = \arctan(-1) = -\frac{\pi}{4} = -45°$. Cuarto cuadrante.

(e) $z = -3i$: $a = 0$ y $b < 0$, luego $r = 3$ y $\operatorname{Arg} z = -\frac{\pi}{2} = -90°$, sin arcotangente ninguna.

(f) $z = -7$: está sobre el eje real negativo, con $r = 7$ y $\operatorname{Arg} z = \pi = 180°$. Escribir $\arctan\frac{0}{-7} = 0$ daría $+7$, el número contrario.

La exponencial compleja y la forma exponencial

Definición 4.9 (Exponencial compleja y forma exponencial)**.** Para θ real se define

$$e^{i\theta} = \cos\theta + i\sin\theta,$$

que es la *fórmula de Euler*, y para $z = x + iy$ cualquiera,

$$e^z = e^{x+iy} = e^x\big(\cos y + i\sin y\big),$$

donde e^x es la exponencial real de siempre. En consecuencia, todo complejo no nulo de módulo r y argumento θ se escribe en **forma exponencial** como

$$z = r\,e^{i\theta},$$

que es la forma polar con otra notación. Más adelante, la sección dedicada a la fórmula de Euler explica de dónde sale esta definición mirando las series de potencias; aquí basta con tomarla como una abreviatura y comprobar que se comporta como una exponencial, que es lo que dice la proposición siguiente.

Proposición 4.10 (La exponencial compleja se comporta como una exponencial)**.** *Para todos $z, w \in \mathbb{C}$ y todo θ real:*

$$e^z e^w = e^{z+w}, \qquad \big|e^{x+iy}\big| = e^x, \qquad \big|e^{i\theta}\big| = 1, \qquad e^z \neq 0, \qquad e^{z+2\pi i} = e^z.$$

Demostración. Basta con el caso imaginario puro, porque la parte real se separa sola. Multiplicando y usando las fórmulas de suma de ángulos de la trigonometría elemental,

$$e^{i\alpha}e^{i\beta} = (\cos\alpha + i\sin\alpha)(\cos\beta + i\sin\beta) = (\cos\alpha\cos\beta - \sin\alpha\sin\beta) + i(\sin\alpha\cos\beta + \cos\alpha\sin\beta),$$

que es $\cos(\alpha+\beta) + i\sin(\alpha+\beta) = e^{i(\alpha+\beta)}$. Para el caso general se combina esto con $e^{x_1}e^{x_2} = e^{x_1+x_2}$ de los reales. El módulo sale de $|\cos y + i\sin y| = \sqrt{\cos^2 y + \sin^2 y} = 1$, de donde $|e^{x+iy}| = e^x \cdot 1 = e^x$, que nunca es cero porque la exponencial real nunca lo es. Y la periodicidad es la de cos y sin: sumar 2π a y no cambia nada. □

Idea clave

Aquí está el corazón del asunto. La igualdad $e^{i\alpha}e^{i\beta} = e^{i(\alpha+\beta)}$ dice que multiplicar complejos *suma argumentos*, y como los módulos se multiplican, multiplicar por z es la operación geométrica "estirar por $|z|$ y girar $\arg z$". Compruébelo con $i = e^{i\pi/2}$: multiplicar por i es girar un cuarto de vuelta, y hacerlo dos veces es girar media, que es multiplicar por -1. Eso es $i^2 = -1$, ya no como definición arbitraria sino como hecho geométrico obligatorio. Y de paso: las fórmulas de suma de ángulos de la trigonometría, tan feas de memorizar, son la distributiva del producto complejo disfrazada.

Regla: pasar de binómica a polar y exponencial, y volver

1. **De binómica a polar: dibuja el punto** $z = a + bi$, aunque sea a mano alzada. Todo lo demás depende de saber en qué cuadrante está.
2. **Calcula el módulo** $r = \sqrt{a^2 + b^2}$, que siempre es positivo y nunca da problemas.
3. **Calcula el argumento con la fórmula por casos** de la definición 4.7, o bien halla el ángulo de referencia $\alpha = \arctan\left|\frac{b}{a}\right| \in [0, \pi/2]$ y colócalo en su cuadrante: primero $\theta = \alpha$; segundo $\theta = \pi - \alpha$; tercero $\theta = \alpha - \pi$; cuarto $\theta = -\alpha$. Si $a = 0$ o $b = 0$, léelo directamente sobre los ejes.
4. **Escribe las dos formas**, que son la misma cosa: $z = r\operatorname{cis}\theta = r\,e^{i\theta}$.
5. **De polar o exponencial a binómica no hay trampa**: $a = r\cos\theta$ y $b = r\sin\theta$, y ya está.
6. **Verifica los signos.** Los de a y b deben coincidir con los del coseno y el seno del ángulo obtenido. Si tu ángulo tiene coseno positivo y tu parte real es negativa, la respuesta está media vuelta desplazada.
7. **Recuerda para qué sirve cada forma**: la binómica para sumar y restar, la polar o la exponencial para multiplicar, dividir, elevar y extraer raíces. Cambiar de forma en el momento adecuado es la mitad del oficio.

Ejemplo 4.11 (Cuatro conversiones en los dos sentidos)**. (a)** $z = -3 + 4i$ **a exponencial.** Módulo $r = \sqrt{9+16} = 5$. Como $a < 0$ y $b > 0$, el argumento principal es $\arctan\left(-\frac{4}{3}\right) + \pi = \pi - 0{,}927295 = 2{,}214297$ radianes, o sea $126{,}869898°$. Luego $z = 5\,e^{2{,}214297\,i}$. Comprobación: $5\cos(2{,}214297) = -3$ y $5\sin(2{,}214297) = 4$ ✓.

(b) $z = 2 + 2i$ **a polar y exponencial.** Módulo $2\sqrt{2} = 2{,}828427$ y argumento $\frac{\pi}{4}$, por estar en el primer cuadrante con $a = b$. Luego $z = 2\sqrt{2}\operatorname{cis}\frac{\pi}{4} = 2\sqrt{2}\,e^{i\pi/4}$.

(c) $z = 3e^{2\pi i/3}$ **a binómica.** Como $\cos 120° = -\frac{1}{2}$ y $\sin 120° = \frac{\sqrt{3}}{2}$,

$$z = 3\left(-\tfrac{1}{2} + \tfrac{\sqrt{3}}{2}i\right) = -\tfrac{3}{2} + \tfrac{3\sqrt{3}}{2}i = -1{,}5 + 2{,}598076\,i.$$

Los signos $(-,+)$ corresponden al segundo cuadrante, que es donde cae $120°$ ✓.

(d) Un caso completo con la exponencial general. $e^{2-i} = e^2(\cos(-1) + i\sin(-1)) = e^2\cos 1 - i\,e^2\sin 1 = 3{,}992324 - 6{,}217676\,i$. Su módulo es $e^2 = 7{,}389056$, como anuncia la

proposición 4.10, y su argumento es -1 radián. Y con $\theta = \pi$ sale $e^{i\pi} = \cos\pi + i\sin\pi = -1$, la identidad famosa, cuyo significado geométrico es sencillo: girar media vuelta es multiplicar por -1.

Potencias y raíces

Teorema 4.12 (Fórmula de De Moivre)**.** *Para todo entero n y todo $z = r\,e^{i\theta}$ con $r > 0$,*

$$z^n = r^n e^{in\theta}, \qquad \textit{es decir} \qquad \big(r\operatorname{cis}\theta\big)^n = r^n \operatorname{cis}(n\theta).$$

Demostración. Para $n \geq 1$, por inducción. El caso $n = 1$ es trivial. Si se cumple para n, entonces, usando la proposición 4.10,

$$z^{n+1} = z^n \cdot z = r^n e^{in\theta} \cdot r\,e^{i\theta} = r^{n+1} e^{i(n\theta+\theta)} = r^{n+1} e^{i(n+1)\theta}.$$

Para $n = 0$ ambos miembros valen 1. Y para $n < 0$, escríbase $n = -m$ con $m > 0$: como $e^{im\theta}e^{-im\theta} = e^0 = 1$, el inverso de $r^m e^{im\theta}$ es $r^{-m}e^{-im\theta}$, que es justamente $r^n e^{in\theta}$. □

Ejemplo 4.13 (Tres potencias que a mano serían inhumanas)**. (a)** $(1+i)^{12}$**.** Módulo $\sqrt{2}$ y argumento 45°, luego el resultado tiene módulo $(\sqrt{2})^{12} = 2^6 = 64$ y argumento $12 \cdot 45° = 540°$, que reducido es 180°. Por tanto $(1+i)^{12} = 64\operatorname{cis}180° = -64$. Compárese con desarrollar un binomio de trece términos.

(b) $(\sqrt{3} - i)^7$**.** Módulo $\sqrt{3+1} = 2$ y argumento $-30°$, luego el resultado tiene módulo $2^7 = 128$ y argumento $-210°$, que equivale a 150°:

$$(\sqrt{3} - i)^7 = 128\left(-\tfrac{\sqrt{3}}{2} + \tfrac{1}{2}i\right) = -64\sqrt{3} + 64i = -110{,}851252 + 64\,i.$$

(c) Exponente negativo. $(1+i)^{-4}$ tiene módulo $(\sqrt{2})^{-4} = \frac{1}{4}$ y argumento $-180°$, luego vale $-\frac{1}{4}$. Y en efecto $(1+i)^4 = -4$.

Teorema 4.14 (Raíces n-ésimas)**.** *Sea $w = R\,e^{i\varphi}$ con $R > 0$ y sea $n \geq 1$ entero. La ecuación $z^n = w$ tiene* exactamente n *soluciones distintas, que son*

$$z_k = \sqrt[n]{R}\; e^{i(\varphi+2\pi k)/n}, \qquad k = 0, 1, \ldots, n-1,$$

donde $\sqrt[n]{R}$ es la raíz n-ésima real y positiva de R. Todas tienen el mismo módulo, de modo que están sobre una misma circunferencia, y sus argumentos se separan $2\pi/n$: son los vértices de un polígono regular de n lados centrado en el origen.

Demostración. Búsquese $z = r\,e^{i\theta}$ con $r > 0$. Por De Moivre, $z^n = r^n e^{in\theta}$, y esto vale $R\,e^{i\varphi}$ si y sólo si los módulos coinciden y los argumentos se diferencian en un múltiplo entero de 2π:

$$r^n = R \qquad \text{y} \qquad n\theta = \varphi + 2\pi k \ \text{ para algún entero } k.$$

La primera ecuación, entre números reales positivos, tiene la única solución $r = \sqrt[n]{R}$. La segunda da $\theta = (\varphi + 2\pi k)/n$. Falta ver cuántos valores distintos produce: dos índices k y k' dan el mismo número si y sólo si sus ángulos difieren en un múltiplo de 2π, es decir si $2\pi(k-k')/n$ lo es, o sea si n divide a $k - k'$. Por tanto hay exactamente n soluciones distintas, las de $k = 0, 1, \ldots, n-1$, y al llegar a $k = n$ se vuelve a la primera. □

Ejemplo 4.15 (Dos extracciones de raíces, con comprobación)**. (a) Las raíces cuartas de** -4**.** Aquí $R = 4$ y $\varphi = 180°$, luego el módulo común es $\sqrt[4]{4} = \sqrt{2} = 1{,}414214$ y los argumentos son 45°, 135°, 225° y 315°, separados 90°. Como $\sqrt{2}\cos 45° = \sqrt{2}\sin 45° = 1$, las cuatro raíces son

$$1+i, \qquad -1+i, \qquad -1-i, \qquad 1-i,$$

los vértices de un cuadrado. Comprobación de la primera: $(1+i)^2 = 2i$ y $(2i)^2 = -4$ ✓. Su suma es 0, como debe ocurrir siempre con $n \geq 2$, y vienen en parejas conjugadas porque el número de partida era real.

(b) Las raíces cúbicas de $2+2i$**.** El módulo de $2+2i$ es $2\sqrt{2} = 2{,}828427$ y su argumento 45°, luego el módulo común de las raíces es $\sqrt[3]{2\sqrt{2}} = \sqrt{2} = 1{,}414214$ y los argumentos son 15°, 135° y 255°. Con $\cos 15° = \frac{\sqrt{6}+\sqrt{2}}{4}$ y $\sin 15° = \frac{\sqrt{6}-\sqrt{2}}{4}$ salen valores exactos:

$$z_0 = \frac{1+\sqrt{3}}{2} + \frac{\sqrt{3}-1}{2}i = 1{,}366025 + 0{,}366025\,i, \qquad z_1 = -1+i, \qquad z_2 = -\frac{\sqrt{3}-1}{2} - \frac{1+\sqrt{3}}{2}i.$$

Comprobación de la segunda, que es la más cómoda: $(-1+i)^2 = -2i$ y $(-2i)(-1+i) = 2i - 2i^2 = 2 + 2i$ ✓. Y la suma de las tres es 0 ✓. Obsérvese que aquí las raíces *no* van en parejas conjugadas: el número de partida no era real.

Cuidado

[En $\mathbb{C}$ no existe "la" raíz] El símbolo $\sqrt{\ }$ deja de designar un número único en cuanto se sale de los reales positivos, y con él se cae la regla $\sqrt{a}\sqrt{b} = \sqrt{ab}$: el contraejemplo clásico es $-1 = i \cdot i = \sqrt{-1}\sqrt{-1} = \sqrt{(-1)(-1)} = \sqrt{1} = 1$, una demostración de que $-1 = 1$ cuyo único fallo es haber usado esa regla fuera de su terreno. Por eso se dice siempre "las n raíces n-ésimas de w", en plural, y dar una sola raíz cúbica es dar un tercio de la respuesta. Segundo aviso: la igualdad $\arg(zw) = \arg z + \arg w$ es cierta *salvo múltiplos de* 2π, y al trabajar con argumentos principales hay que reducir al final; con los módulos, en cambio, $|zw| = |z||w|$ es exacta y sin condiciones.

Pausa para pensar

Merece la pena mirar hacia atrás y ver qué ha pasado. Se ha partido de un par de números reales y una regla de multiplicar; de ahí ha salido, sin ninguna hipótesis extra, que multiplicar es girar y estirar, que las potencias se calculan multiplicando ángulos y que un número tiene exactamente n raíces n-ésimas repartidas en un polígono regular. Nada de eso es visible en la forma binómica, y todo es evidente en la exponencial. Las secciones que siguen explotan esa doble descripción: la binómica para lo aditivo, la polar para lo multiplicativo, y el paso de una a otra —con el cuadrante bien puesto— como la maniobra que hay que tener automatizada.

4.3. Definición y forma binómica

Definición 4.16. Se define la **unidad imaginaria** i como un símbolo que cumple

$$i^2 = -1.$$

Un **número complejo** es una expresión $z = a + bi$ con $a, b \in \mathbb{R}$. El número a es la **parte real** $\mathrm{Re}(z)$ y b la **parte imaginaria** $\mathrm{Im}(z)$. El conjunto de todos ellos se denota $\mathbb{C}$.

Cuidado

La parte imaginaria es b, **no** bi. En $z = 3 - 7i$, $\mathrm{Im}(z) = -7$, un número real. Es una convención universal y confundirla estropea todas las fórmulas siguientes.

Regla: igualdad de complejos

$a + bi = c + di$ si y solo si $a = c$ *y* $b = d$. Una sola ecuación compleja equivale siempre a *dos* ecuaciones reales. Este desdoblamiento es la técnica más útil del capítulo.

Un real es el caso $b = 0$; un complejo con $a = 0$ se llama *imaginario puro*. Así $\mathbb{R} \subset \mathbb{C}$, igual que $\mathbb{Q} \subset \mathbb{R}$.

Potencias de i

Regla: el ciclo de cuatro

$$i^1 = i, \quad i^2 = -1, \quad i^3 = i^2 \cdot i = -i, \quad i^4 = 1,$$

y a partir de ahí se repite. Para calcular i^n, divida n entre 4 y quédese con el resto:

$$i^{2026}: \quad 2026 = 4 \cdot 506 + 2 \;\Rightarrow\; i^{2026} = i^2 = -1.$$

Ejemplo 4.17. i^{35}: como $35 = 4 \cdot 8 + 3$, se tiene $i^{35} = i^3 = -i$.

4.4. Aritmética: sumar, multiplicar, dividir

Regla: suma y resta

Se suman parte con parte, como si i fuera una letra:

$$(a + bi) + (c + di) = (a + c) + (b + d)i.$$

Regla: producto

Se multiplica como binomios y se usa $i^2 = -1$:

$$(a + bi)(c + di) = ac + adi + bci + bdi^2 = (ac - bd) + (ad + bc)i.$$

No memorice la fórmula final: haga la distributiva y sustituya i^2 al terminar. Es más rápido y no se equivoca de signo.

Ejemplo 4.18. $(3 + 2i)(1 - 4i)$.

Paso 1. Distributiva: $3 - 12i + 2i - 8i^2$.
Paso 2. $-8i^2 = +8$.
Paso 3. Agrupar: $(3 + 8) + (-12 + 2)i = 11 - 10i$.

Ejemplo 4.19. $(3 - 2i)^2 = 9 - 12i + 4i^2 = 9 - 12i - 4 = 5 - 12i$. Los productos notables del álgebra elemental funcionan igual; lo único nuevo es el i^2 del final.

Definición 4.20. El **conjugado** de $z = a + bi$ es $\bar{z} = a - bi$. El **módulo** es $|z| = \sqrt{a^2 + b^2}$.

Regla: la identidad que hace posible dividir

$$z\,\bar{z} = (a + bi)(a - bi) = a^2 - b^2 i^2 = a^2 + b^2 = |z|^2.$$

Es decir: *un complejo por su conjugado siempre da un número real no negativo*. Es la diferencia de cuadrados de siempre, con el signo que $i^2 = -1$ le regala.

De ahí sale la división: para dividir entre $c + di$, multiplique arriba y abajo por el conjugado $c - di$. Es exactamente "racionalizar el denominador" del álgebra elemental, y por eso a veces se le llama *realizar* el denominador.

Ejemplo 4.21. Calcula $\dfrac{2 + 3i}{1 - i}$.

Paso 1. Multiplicamos por $\dfrac{1 + i}{1 + i}$:

$$\frac{(2 + 3i)(1 + i)}{(1 - i)(1 + i)}.$$

Paso 2. Denominador: $1^2 + 1^2 = 2$.
Paso 3. Numerador: $2 + 2i + 3i + 3i^2 = 2 + 5i - 3 = -1 + 5i$.
Paso 4. Resultado: $\dfrac{-1 + 5i}{2} = -\dfrac{1}{2} + \dfrac{5}{2}i$.

Ejemplo 4.22. $\dfrac{1 + 2i}{3 - i} = \dfrac{(1 + 2i)(3 + i)}{9 + 1} = \dfrac{3 + i + 6i - 2}{10} = \dfrac{1 + 7i}{10} = \dfrac{1}{10} + \dfrac{7}{10}i.$

Regla: propiedades del conjugado y del módulo

$$\overline{z + w} = \bar{z} + \bar{w}, \qquad \overline{zw} = \bar{z}\,\bar{w}, \qquad |zw| = |z||w|, \qquad \left|\frac{z}{w}\right| = \frac{|z|}{|w|}.$$

Además $\mathrm{Re}(z) = \frac{z + \bar{z}}{2}$ y $\mathrm{Im}(z) = \frac{z - \bar{z}}{2i}$, y z es real si y solo si $z = \bar{z}$.

Cuidado

$|z + w| \neq |z| + |w|$ en general: solo vale $\leq$ (desigualdad triangular). Con $z = 1$ y $w = -1$: la izquierda es 0 y la derecha 2. Y muy importante: en $\mathbb{C}$ **no hay orden**. Escribir $z < w$

para complejos no significa nada; ninguna forma de ordenarlos es compatible con las operaciones. Solo los módulos, que son reales, se comparan.

4.5. El plano de Argand

Identifiquemos $z = a + bi$ con el punto (a, b) del plano cartesiano de la geometría analítica. El eje horizontal se llama *eje real* y el vertical *eje imaginario*. Esta representación se llama **plano de Argand** o plano complejo.

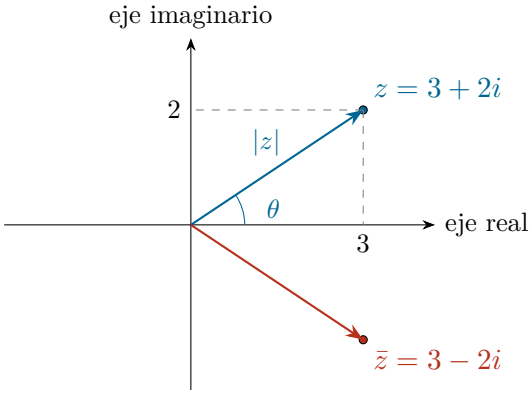


Figura 4.1: El plano de Argand: un número complejo es un punto, su módulo es la distancia al origen y su argumento el ángulo con el eje real. Sumar se convierte en sumar vectores; multiplicar, en girar y estirar.

Con esta imagen, tres cosas que eran fórmulas se vuelven evidentes:

- **Sumar** es sumar vectores: la regla del paralelogramo. Por eso la desigualdad triangular es literalmente la desigualdad triangular de la geometría.
- **El módulo** $|z| = \sqrt{a^2 + b^2}$ es la distancia al origen: es el teorema de Pitágoras y nada más.
- **Conjugar** es reflejar respecto del eje real.

Ejemplo 4.23. $|z - 3| = 2$ describe los puntos que distan 2 del punto 3: una circunferencia de centro $(3, 0)$ y radio 2. Toda la geometría analítica de circunferencias y rectas se reescribe en lenguaje complejo, a menudo con ecuaciones más cortas.

4.6. Forma polar: módulo y argumento

Un punto del plano queda determinado por sus coordenadas (a, b), pero también por su distancia al origen y el ángulo que forma con el eje positivo. Esta segunda descripción es la que hace milagros.

Definición 4.24. Si $z \neq 0$, su **argumento** $\theta = \arg z$ es el ángulo desde el eje real positivo hasta el vector z. Con $r = |z|$ se tiene $a = r\cos\theta$, $b = r\sin\theta$, y por tanto la **forma polar** o trigonométrica

$$z = r(\cos\theta + i\sin\theta),$$

que se abrevia $z = r\operatorname{cis}\theta$.

Cuidado

El argumento no es único: θ y $\theta + 2\pi k$ dan el mismo punto. Se llama *argumento principal* al que está en $(-\pi, \pi]$. Y cuidado con $\theta = \arctan(b/a)$: la arcotangente devuelve siempre un ángulo del primer o cuarto cuadrante, así que si z está a la izquierda del eje imaginario hay que sumar π. Dibujar el punto antes de escribir el ángulo evita el 100 % de estos errores.

Ejemplo 4.25. $z = 1 + i$: $r = \sqrt{1+1} = \sqrt{2}$ y $\theta = 45^\circ = \pi/4$ (primer cuadrante). Luego $1 + i = \sqrt{2}\operatorname{cis}\frac{\pi}{4}$.

Ejemplo 4.26. $z = -1 + i$: $r = \sqrt{2}$, pero $\arctan(1/(-1)) = -45^\circ$, que apunta al cuarto cuadrante. Como z está en el segundo, el argumento correcto es $-45^\circ + 180^\circ = 135^\circ = 3\pi/4$.

Teorema 4.27 (Producto en forma polar)**.**

$$r_1\operatorname{cis}\theta_1 \cdot r_2\operatorname{cis}\theta_2 = r_1 r_2\operatorname{cis}(\theta_1 + \theta_2).$$

Demostración. Multiplicando y usando las fórmulas de suma de ángulos de la trigonometría elemental:

$$\begin{aligned}(\cos\theta_1 + i\sin\theta_1)(\cos\theta_2 + i\sin\theta_2) &= (\cos\theta_1\cos\theta_2 - \sin\theta_1\sin\theta_2) \\ &\quad + i(\sin\theta_1\cos\theta_2 + \cos\theta_1\sin\theta_2) \\ &= \cos(\theta_1+\theta_2) + i\sin(\theta_1+\theta_2).\end{aligned}$$

□

Ejemplo 4.28 (Multiplicar es girar y estirar)**. (a) Un producto sencillo.** $(2\operatorname{cis}30^\circ)(3\operatorname{cis}45^\circ) = 6\operatorname{cis}75^\circ$: se multiplican los módulos y se suman los argumentos. En forma binómica habría que desarrollar cuatro productos y usar identidades trigonométricas.

(b) Multiplicar por i**.** Como $\mathrm{i} = 1\operatorname{cis}90^\circ$, multiplicar por i es girar noventa grados sin cambiar el tamaño. Por eso $\mathrm{i}^2 = -1$: dos giros de noventa grados dan media vuelta.

(c) Potencias. La fórmula de De Moivre $(r\operatorname{cis}\theta)^n = r^n\operatorname{cis}(n\theta)$ sale de aplicar el teorema n veces. Con $(1+\mathrm{i})^8$: el módulo es $\sqrt{2}$ y el argumento 45°, luego el resultado es $(\sqrt{2})^8\operatorname{cis}360^\circ = 16$.

(d) El argumento vive módulo 360 **grados.** Al multiplicar $\operatorname{cis}200^\circ$ por $\operatorname{cis}300^\circ$ sale $\operatorname{cis}500^\circ = \operatorname{cis}140^\circ$. Olvidarse de reducir es el error clásico, y es también la razón de que la función argumento no sea continua en todo el plano: hay que cortar por algún sitio.

Idea clave

Aquí está el corazón del capítulo. Multiplicar dos complejos es *multiplicar los módulos y sumar los argumentos*. Es decir: multiplicar por z es la operación geométrica "estirar por $|z|$ y girar $\arg z$".
Compruébelo con i: tiene módulo 1 y argumento 90°, luego multiplicar por i es *girar un cuarto de vuelta sin estirar*. Y multiplicar por i dos veces gira media vuelta, que es multiplicar por -1. Eso es exactamente $i^2 = -1$, y ahora ya no es una definición arbitraria sino un hecho geométrico obligatorio.
Además explica de golpe por qué las fórmulas de suma de ángulos de la trigonometría son tan feas de memorizar y tan naturales aquí: son la ley distributiva del producto complejo, disfrazada.

Regla: pasar de forma binómica a polar sin equivocarse de cuadrante

Es el cálculo que más suspensos causa del capítulo, y no por difícil sino por confiado: la calculadora sabe hacer arcotangentes, pero no sabe en qué cuadrante está usted.

1. **Dibuje el punto** $z = a + bi$ en el plano de Argand, aunque sea a mano alzada. Todo lo demás depende de este paso.
2. **Calcule el módulo** $r = \sqrt{a^2 + b^2}$. Siempre es positivo, y nunca da problemas.
3. **Si** $a = 0$, no use la arcotangente: el argumento es $\pi/2$ si $b > 0$ y $-\pi/2$ si $b < 0$. Si $b = 0$, es 0 si $a > 0$ y π si $a < 0$.
4. **En otro caso, calcule el ángulo de referencia** $\alpha = \arctan\left|\frac{b}{a}\right|$, que siempre está entre 0 y $\pi/2$.
5. **Coloque el ángulo en el cuadrante que le corresponde** según los signos de a y b: primer cuadrante $\theta = \alpha$; segundo $\theta = \pi - \alpha$; tercero $\theta = \alpha - \pi$; cuarto $\theta = -\alpha$.
6. **Para el camino inverso**, de polar a binómica, no hay trampa ninguna: $a = r\cos\theta$ y $b = r\sin\theta$.
7. **Verifique** comprobando que los signos de $r\cos\theta$ y $r\sin\theta$ coinciden con los de a y b. Si no coinciden, el argumento está media vuelta desplazado.

Ejemplo 4.29 (Siete conversiones, una por cada situación posible)**.** (a) $z = 1 + i\sqrt{3}$. Módulo $r = \sqrt{1+3} = 2$. Ángulo de referencia $\alpha = \arctan\sqrt{3} = 60°$. El punto está en el primer cuadrante, luego $\theta = 60° = \pi/3$ y

$$z = 2\,\mathrm{cis}\,\tfrac{\pi}{3}.$$

(b) $z = -1 + i\sqrt{3}$. El módulo vuelve a ser 2 y $\alpha = 60°$, pero ahora $a < 0$ y $b > 0$: segundo cuadrante, $\theta = 180° - 60° = 120° = 2\pi/3$. La calculadora, si se le pide $\arctan(\sqrt{3}/(-1))$, devuelve $-60°$, que apunta al cuarto cuadrante: exactamente el punto opuesto. Es el error clásico.

(c) $z = -\sqrt{3} - i$. Módulo 2, ángulo de referencia $\arctan(1/\sqrt{3}) = 30°$, tercer cuadrante, luego $\theta = 30° - 180° = -150° = -5\pi/6$. Como argumento principal se prefiere $-5\pi/6$ a $210°$, aunque los dos describen el mismo punto.

(d) $z = 3 - 3i$. Módulo $\sqrt{9+9} = 3\sqrt{2}$, ángulo de referencia $45°$, cuarto cuadrante, $\theta = -45° = -\pi/4$. Y en efecto $3\sqrt{2}\cos(-45°) = 3\sqrt{2} \cdot \frac{\sqrt{2}}{2} = 3$ ✓, con parte imaginaria -3 ✓.

(e) $z = -5$. Aquí $b = 0$ y no hay arcotangente que valga: el número está sobre el eje real negativo, luego $r = 5$ y $\theta = \pi$. Escribir $\arctan(0/(-5)) = 0$ y concluir $\theta = 0$ daría $+5$, que es el número contrario.

(f) $z = -2i$. Ahora $a = 0$: el punto está sobre el eje imaginario, hacia abajo. Módulo 2 y argumento $-\pi/2$. La arcotangente ni siquiera está definida, porque habría que dividir entre cero.

(g) *Camino inverso.* Sea $z = 4\,\mathrm{cis}\,\frac{5\pi}{6}$. Entonces

$$a = 4\cos 150° = 4 \cdot \left(-\tfrac{\sqrt{3}}{2}\right) = -2\sqrt{3}, \qquad b = 4\sin 150° = 4 \cdot \tfrac{1}{2} = 2,$$

luego $z = -2\sqrt{3} + 2i$. Comprobación del módulo: $\sqrt{12+4} = 4$ ✓, y los signos $(-,+)$ corresponden al segundo cuadrante, que es donde cae $150°$ ✓.

Cuidado

Hay una segunda trampa, más sutil, en la fórmula $\arg(z_1 z_2) = \arg z_1 + \arg z_2$: es cierta *salvo múltiplos de* 2π. Multiplicando $z_1 = z_2 = -1 + i$, cuyos argumentos principales valen $3\pi/4$, se obtiene $3\pi/4 + 3\pi/4 = 3\pi/2$, que no está en $(-\pi, \pi]$; el argumento principal del producto es $3\pi/2 - 2\pi = -\pi/2$, y en efecto $(-1+i)^2 = -2i$ ✓. Con módulos la igualdad $|z_1 z_2| = |z_1||z_2|$ sí es exacta y sin condiciones. La función argumento es la primera función multivaluada que aparece en el libro, y de ahí nacen los cortes de rama del capítulo 30.

Pausa para pensar

Muchos lenguajes de programación tienen dos arcotangentes: `atan(y/x)`, que solo recibe el cociente y por tanto no puede saber el cuadrante, y `atan2(y,x)`, que recibe las dos coordenadas por separado y devuelve el argumento correcto en $(-\pi, \pi]$. Esa segunda función existe únicamente por el problema de este apartado. Si al programar cualquier cosa con ángulos escribe usted `atan` en vez de `atan2`, tarde o temprano la mitad del plano le saldrá al revés.

Idea clave

Una forma de no equivocarse nunca: el módulo dice *cuánto* y el argumento dice *hacia dónde*. La coordenada a es $r\cos\theta$ y la b es $r\sin\theta$, así que los signos de a y b son los signos del coseno y del seno del ángulo. Si su respuesta tiene un ángulo cuyo coseno es positivo y una parte real negativa, la respuesta está mal, sin necesidad de revisar

ninguna cuenta.

4.7. Lugares geométricos en el plano complejo

Una vez que z es un punto del plano, cualquier condición sobre z describe un conjunto de puntos, es decir, una figura. La gracia de escribir esas condiciones con módulos y argumentos en lugar de con x e y es que quedan mucho más cortas y, sobre todo, mucho más parlantes: $|z-a|$ *es* la distancia de z al punto a, y ya está.

Idea clave

Traduce siempre el símbolo antes de calcular. $|z-a|$ es la distancia de z a a; $\arg(z-a)$ es la dirección en la que se ve z desde a; $\operatorname{Re} z$ y $\operatorname{Im} z$ son la abscisa y la ordenada. Con ese diccionario, condiciones como "la distancia a a es constante" o "equidista de a y de b" se reconocen sin desarrollar nada: son la circunferencia y la mediatriz de toda la vida.

Regla: identificar el lugar geométrico que define una condición

1. **Lee la condición en lenguaje de distancias.** Sustituye mentalmente $|z-a|$ por "distancia de z al punto a" antes de tocar el álgebra.
2. **Compara con el catálogo.** Distancia constante a un punto: circunferencia. Distancias iguales a dos puntos: mediatriz. Suma constante de distancias a dos puntos: elipse. Diferencia constante: hipérbola. Argumento constante: semirrecta.
3. **Si no reconoces la figura, escribe** $z = x + iy$ y desarrolla, recordando que $|z|^2 = z\bar{z} = x^2 + y^2$. Es preferible elevar al cuadrado los módulos que arrastrar raíces.
4. **Ordena la ecuación resultante** agrupando $x^2 + y^2$ y completando cuadrados en x y en y por separado.
5. **Lee el centro y el radio** de la forma $(x-h)^2 + (y-k)^2 = r^2$; si el segundo miembro sale negativo, el lugar es vacío, y si sale cero, es un único punto.
6. **Comprueba con dos o tres puntos concretos** sustituyéndolos en la condición original.
7. **Dibuja.** Un boceto a mano detecta en dos segundos los errores de signo que el álgebra esconde durante media página.

Ejemplo 4.30 (Cuatro condiciones y sus figuras)**.** 1. $|z-(1+2i)| = 3$. Es directamente la circunferencia de centro $(1, 2)$ y radio 3. Si lo desarrollas obtienes $x^2-2x+y^2-4y+5 = 9$, que completando cuadrados es $(x-1)^2 + (y-2)^2 = 9$: lo mismo, con más trabajo. La desigualdad $|z-(1+2i)| < 3$ es el disco abierto interior.

2. $|z-1| = |z-i|$. Los puntos que equidistan de 1 y de i forman la mediatriz del segmento

que los une. Desarrollando, $(x-1)^2+y^2 = x^2+(y-1)^2$ se reduce a $-2x+2y=0$, es decir la recta $y=x$: en efecto, la mediatriz de $(1,0)$ y $(0,1)$.

3. $|z-2|+|z+2|=6$. Suma constante de distancias a $(2,0)$ y $(-2,0)$: una elipse de focos esos dos puntos, con $a=3$ (la mitad de 6) y $c=2$, luego $b^2=9-4=5$ y su ecuación es $\dfrac{x^2}{9}+\dfrac{y^2}{5}=1$. Comprobación: para $z=3$ las distancias son 1 y 5; para $z=i\sqrt{5}$ ambas valen 3; en los dos casos suman 6.

4. $|z-2| = 2|z+1|$. Aquí las distancias no son iguales sino proporcionales, y la figura ya no es una recta. Elevando al cuadrado, $(x-2)^2+y^2 = 4\left[(x+1)^2+y^2\right]$, de donde $3x^2+12x+3y^2=0$ y, dividiendo por 3, $x^2+y^2+4x=0$, es decir $(x+2)^2+y^2=4$: la circunferencia de centro $(-2,0)$ y radio 2. Se la llama *circunferencia de Apolonio* del par de puntos, y aparece siempre que la razón de distancias es constante y distinta de 1.

Ejemplo 4.31 (Regiones, argumentos y una hipérbola)**.** 1. $\operatorname{Re} z = 1$ es la recta vertical $x=1$, y $\operatorname{Re} z > 1$ el semiplano abierto a su derecha. Curiosamente $|z| = |z-2|$ describe la misma recta, porque es la mediatriz de 0 y 2: dos condiciones de aspecto muy distinto pueden dar la misma figura.

2. $\arg z = \frac{\pi}{4}$ es la semirrecta $y=x$ con $x>0$, *sin* el origen (donde el argumento no está definido) y *sin* la parte del tercer cuadrante, cuyo argumento es $-\frac{3\pi}{4}$. La condición $0 < \arg z < \frac{\pi}{2}$ es el primer cuadrante abierto.

3. $1 \le |z-i| \le 2$ es una corona circular centrada en i: todo lo que queda entre las circunferencias de radios 1 y 2, incluidas ambas. Intersecándola con $\operatorname{Im} z > 1$ te quedas con la mitad superior de esa corona.

4. $\big||z-3|-|z+3|\big| = 4$ es una hipérbola de focos $(3,0)$ y $(-3,0)$, con $a=2$ y $c=3$, luego $b^2=9-4=5$ y su ecuación cartesiana es $\dfrac{x^2}{4}-\dfrac{y^2}{5}=1$. Sin el valor absoluto exterior obtendrías una sola de las dos ramas.

Cuidado

Cuidado con las condiciones sobre argumentos: no son simétricas ni continuas como las de módulo. Escribir $\arg z = \frac{\pi}{4}$ y esperar la recta completa $y=x$ es el error más frecuente, porque el argumento distingue la semirrecta de ida de la de vuelta. Además, al pasar por el eje real negativo el argumento principal salta de π a $-\pi$, de manera que una región descrita como $\alpha < \arg z < \beta$ puede estar rota en dos trozos si el intervalo cruza ese corte. Cuando dudes, comprueba con un punto de cada zona.

Pausa para pensar

Fíjate en que las cónicas del capítulo la geometría analítica han reaparecido aquí definidas por una sola igualdad con módulos, sin ejes ni coeficientes. Esa economía no es casual: el plano complejo es el lenguaje natural de la geometría plana, y muchas demostraciones euclídeas largas se vuelven cálculos de dos líneas al escribirlas con números complejos.

La sección siguiente da un paso más y estudia qué le hacen a estas figuras las funciones que transforman el plano.

4.8. De Moivre y las raíces n-ésimas

Teorema 4.32 (Fórmula de De Moivre)**.** *Para todo entero n,*

$$\big(r\operatorname{cis}\theta\big)^n = r^n\operatorname{cis}(n\theta).$$

Es el teorema anterior aplicado n veces: elevar a la n eleva el módulo a la n y multiplica el ángulo por n.

Ejemplo 4.33. Calcula $(1+i)^{10}$.

Paso 1. Forma polar: $1+i=\sqrt{2}\operatorname{cis}45^\circ$.

Paso 2. De Moivre: $(\sqrt{2})^{10}\operatorname{cis}(450^\circ)=32\operatorname{cis}90^\circ$.

Paso 3. $\operatorname{cis}90^\circ=i$, luego $(1+i)^{10}=32i$.

Compare con desarrollar el binomio de Newton con once términos y once potencias de i. Esa comparación es todo el argumento a favor de la forma polar.

Ejemplo 4.34. $(1+i\sqrt{3})^6$: módulo 2, argumento 60°; luego $2^6\operatorname{cis}360^\circ=64$. Un número real, obtenido de uno que no lo era.

Raíces: la sorpresa de que sean varias

En $\mathbb{R}$, un número positivo tiene una raíz cuadrada positiva y un número tiene una única raíz cúbica real. En $\mathbb{C}$ la situación es mucho más limpia y mucho más rica:

Teorema 4.35 (Raíces n-ésimas)**.** *Todo $z=r\operatorname{cis}\theta$ distinto de cero tiene **exactamente** n raíces n-ésimas, dadas por*

$$w_k=\sqrt[n]{r}\ \operatorname{cis}\!\left(\frac{\theta+360^\circ k}{n}\right),\qquad k=0,1,\dots,n-1.$$

Geométricamente: todas tienen el mismo módulo $\sqrt[n]{r}$, luego están sobre una misma circunferencia, y sus argumentos se separan $360^\circ/n$. *Forman los vértices de un polígono regular de n lados.*

Ejemplo 4.36. Las tres raíces cúbicas de 8.

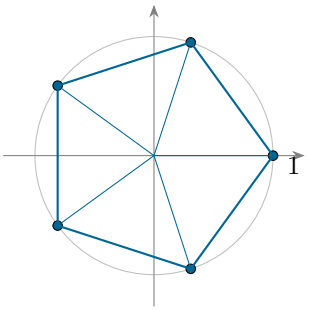
Paso 1. $8=8\operatorname{cis}0^\circ$, luego $\sqrt[3]{8}=2$ y los argumentos son $0^\circ, 120^\circ, 240^\circ$.

Paso 2. Las raíces son

$$2,\qquad 2\operatorname{cis}120^\circ=-1+i\sqrt{3},\qquad 2\operatorname{cis}240^\circ=-1-i\sqrt{3}.$$

Paso 3. Comprobación: $(-1+i\sqrt{3})^3$. Primero $(-1+i\sqrt{3})^2=1-2i\sqrt{3}-3=-2-2i\sqrt{3}$; multiplicando por $-1+i\sqrt{3}$ sale $2-2i\sqrt{3}+2i\sqrt{3}+6=8$ ✓.

Un triángulo equilátero inscrito en la circunferencia de radio 2.



Las cinco raíces quintas de 1: un pentágono regular.

Figura 4.2: Las cinco raíces quintas de la unidad reparten la circunferencia en ángulos de $72°$. Las raíces n-ésimas de cualquier número forman siempre un polígono regular de n lados centrado en el origen.

Ejemplo 4.37. Las raíces cuartas de -16. Como $-16 = 16\,\mathrm{cis}\,180°$, el módulo es $\sqrt[4]{16} = 2$ y los argumentos son $45°, 135°, 225°, 315°$:

$$\sqrt{2}+i\sqrt{2}, \quad -\sqrt{2}+i\sqrt{2}, \quad -\sqrt{2}-i\sqrt{2}, \quad \sqrt{2}-i\sqrt{2}.$$

Un cuadrado. Obsérvese que vienen en pares conjugados, como debe ocurrir siempre que el número de partida sea real.

Observación 4.38. Las raíces n-ésimas de 1 forman un grupo cíclico de orden n con la multiplicación: es el ejemplo más visual de grupo cíclico del capítulo 23, y el hecho de que el pentágono regular sea construible con regla y compás mientras el heptágono no lo sea es un teorema de la teoría de Galois del capítulo 26. Casi todo el álgebra abstracta de la Parte IV se puede motivar mirando este dibujo.

Regla: calcular las n raíces n-ésimas de un número complejo

En $\mathbb{R}$ un número tiene una raíz cúbica y a lo sumo dos raíces cuadradas. En $\mathbb{C}$ todo número no nulo tiene *exactamente* n raíces n-ésimas, ni una más ni una menos, y salen todas de la misma receta:

1. **Escriba el número en forma polar**, $w = R\,\mathrm{cis}\,\varphi$, con el argumento en el cuadrante correcto.
2. **El módulo de todas las raíces es el mismo**: $r = \sqrt[n]{R}$, la raíz n-ésima *real y positiva* de R. Aquí no hay ambigüedad ninguna.
3. **El argumento de la primera raíz** es $\theta_0 = \varphi/n$.
4. **Las demás se obtienen sumando** $2\pi/n$ sucesivamente:

$$\theta_k = \frac{\varphi + 2\pi k}{n}, \qquad k = 0, 1, \ldots, n-1.$$

5. **Pare en** $k = n-1$: al llegar a $k = n$ el ángulo se ha incrementado en 2π y se vuelve a la primera raíz.

6. **Dibújelas**: las n raíces son los vértices de un polígono regular de n lados inscrito en la circunferencia de radio r. Ese dibujo es la comprobación más rápida que existe.
7. **Verifique** que la suma de las n raíces es 0 (si $n \geq 2$) y que elevar una cualquiera de ellas a la n devuelve w.

Ejemplo 4.39 (Seis extracciones de raíces y los polígonos que dibujan)**.** (a) *Raíces cúbicas de* 1. Aquí $R = 1$ y $\varphi = 0$, luego $r = 1$ y $\theta_k = 2\pi k/3$ para $k = 0, 1, 2$:

$$1, \qquad \operatorname{cis} 120^\circ = -\tfrac{1}{2} + \tfrac{\sqrt{3}}{2}i, \qquad \operatorname{cis} 240^\circ = -\tfrac{1}{2} - \tfrac{\sqrt{3}}{2}i.$$

Forman un triángulo equilátero inscrito en la circunferencia unidad. Su suma es $1 - \frac{1}{2} - \frac{1}{2} + i(\frac{\sqrt{3}}{2} - \frac{\sqrt{3}}{2}) = 0$ ✓.

(b) *Raíces cúbicas de* $8i$. Ahora $8i = 8 \operatorname{cis} 90^\circ$, luego $r = \sqrt[3]{8} = 2$ y los argumentos son 30°, $30^\circ + 120^\circ = 150^\circ$ y 270°:

$$2 \operatorname{cis} 30^\circ = \sqrt{3} + i, \qquad 2 \operatorname{cis} 150^\circ = -\sqrt{3} + i, \qquad 2 \operatorname{cis} 270^\circ = -2i.$$

Comprobación de la primera: $(\sqrt{3} + i)^3$ tiene módulo $2^3 = 8$ y argumento $3 \cdot 30^\circ = 90^\circ$, o sea $8i$ ✓. Y la suma es $\sqrt{3} - \sqrt{3} + i(1 + 1 - 2) = 0$ ✓.

(c) *Raíces cuartas de* -16. Como $-16 = 16 \operatorname{cis} 180^\circ$, se tiene $r = \sqrt[4]{16} = 2$ y argumentos 45°, 135°, 225°, 315°:

$$\pm\sqrt{2} \pm \sqrt{2}\, i$$

en las cuatro combinaciones de signos, que forman un cuadrado. Obsérvese que *ninguna* es real: era de esperar, porque ningún número real elevado a la cuarta da un negativo.

(d) *Raíces sextas de* 1. Son $\operatorname{cis}(60^\circ k)$ para $k = 0, \ldots, 5$, es decir

$$1, \ \tfrac{1}{2} + \tfrac{\sqrt{3}}{2}i, \ -\tfrac{1}{2} + \tfrac{\sqrt{3}}{2}i, \ -1, \ -\tfrac{1}{2} - \tfrac{\sqrt{3}}{2}i, \ \tfrac{1}{2} - \tfrac{\sqrt{3}}{2}i,$$

los vértices de un hexágono regular. Dos de ellas son reales, 1 y -1, y las otras cuatro van en parejas conjugadas: eso pasa siempre que el número de partida es real.

(e) *Raíces quintas de* 32. Módulo $\sqrt[5]{32} = 2$ y argumentos $72^\circ k$. Solo una es real, el 2; las otras cuatro forman dos parejas conjugadas. La suma de las cinco vuelve a ser cero, y separando la parte real de esa identidad se obtiene

$$1 + 2\cos 72^\circ + 2\cos 144^\circ = 0.$$

Con $\cos 36^\circ = \frac{1+\sqrt{5}}{4}$ y $\cos 72^\circ = \frac{\sqrt{5}-1}{4}$ la identidad se comprueba a mano, y de paso aparece el número áureo: no es casualidad que el pentágono regular y la razón áurea vayan siempre juntos.

(f) *El producto de todas las raíces.* Para las raíces n-ésimas de la unidad, el producto es $(-1)^{n+1}$: vale 1 si n es impar y -1 si n es par. Con $n = 5$ da 1 y con $n = 6$ da -1, como se comprueba emparejando cada raíz con su conjugada —cuyo producto es 1— y quedándose con las reales sueltas.

Cuidado

El símbolo $\sqrt{\ }$ deja de tener un significado único en cuanto se sale de los reales positivos, y con él se cae la regla $\sqrt{a}\sqrt{b} = \sqrt{ab}$. El contraejemplo canónico es

$$-1 = i^2 = \sqrt{-1}\sqrt{-1} \overset{?}{=} \sqrt{(-1)(-1)} = \sqrt{1} = 1,$$

una demostración de que $-1 = 1$ cuyo único fallo es haber usado esa regla fuera de su terreno. Por eso en $\mathbb{C}$ no se escribe $\sqrt[n]{w}$ como si designara un número: se dice "las n raíces n-ésimas de w", en plural.

Pausa para pensar

Las n raíces n-ésimas de la unidad forman un grupo con la multiplicación: el producto de dos de ellas vuelve a ser una de ellas, y cada una tiene inversa —su conjugada—. Es un grupo cíclico de orden n, generado por $\operatorname{cis}(2\pi/n)$, y es exactamente el mismo objeto que el $\mathbb{Z}_n$ del capítulo 3: el reloj de n horas, dibujado sobre la circunferencia. Esa identificación es una de las puertas de entrada al capítulo 23, y las raíces de la unidad reaparecerán en el capítulo 26 como la pieza que decide qué polígonos regulares se pueden construir con regla y compás.

Idea clave

Todo el método cabe en una frase: *extraer raíces es repartir el ángulo*. Como multiplicar suma argumentos, elevar a la n multiplica el argumento por n; luego extraer la raíz n-ésima lo divide entre n. Y como el argumento de partida está definido salvo vueltas completas, al dividir entre n esas vueltas dejan de ser equivalentes y se abren en n respuestas distintas. Las n raíces no son un capricho: son las n maneras de repartir las vueltas.

4.9. Cuadráticas con discriminante negativo

Volvamos al álgebra elemental. Allí, ante $\Delta = b^2 - 4ac < 0$, dijimos "no hay soluciones reales". Ahora podemos decir la verdad completa: hay dos soluciones complejas conjugadas.

Ejemplo 4.40. $x^2 - 4x + 13 = 0$.

Paso 1. $\Delta = 16 - 52 = -36$.

Paso 2. $\sqrt{-36} = 6i$.

Paso 3. $x = \dfrac{4 \pm 6i}{2} = 2 \pm 3i$.

Paso 4. Comprobación con $x = 2 + 3i$: $(2+3i)^2 = 4 + 12i - 9 = -5 + 12i$; entonces $-5 + 12i - 4(2+3i) + 13 = -5 + 12i - 8 - 12i + 13 = 0$ ✓.

Teorema 4.41 (Raíces conjugadas)**.** *Si un polinomio tiene todos sus coeficientes* reales *y z es raíz suya, entonces $\bar{z}$ también lo es.*

Demostración. Si $p(z) = a_n z^n + \cdots + a_0 = 0$, conjugue la igualdad entera. Como conjugar respeta sumas y productos y deja fijos los reales, sale $a_n \bar{z}^n + \cdots + a_0 = \overline{p(z)} = \bar{0} = 0$, es decir $p(\bar{z}) = 0$. □

Ejemplo 4.42 (Raíces conjugadas y factorización real)**. (a) Un caso concreto.** $x^2 - 2x + 5$ tiene raíces $1 \pm 2\mathrm{i}$: aparecen en pareja, como debe ser. Su suma es 2 y su producto 5, ambos reales, coherente con los coeficientes.

(b) Grado impar. Un polinomio real de grado impar tiene al menos una raíz real, porque las complejas se emparejan y no pueden agotar un número impar de raíces. Para $x^3 - x - 1$ la raíz real es aproximadamente 1,324718 y las otras dos son conjugadas.

(c) Factorización sobre $\mathbb{R}$. Todo polinomio real se factoriza en factores lineales y cuadráticos irreducibles, juntando cada pareja conjugada: $x^4+1 = (x^2+\sqrt{2}x+1)(x^2-\sqrt{2}x+1)$. Ése es el teorema que hace posible la descomposición en fracciones simples con coeficientes reales.

(d) Con coeficientes complejos, falla. El polinomio $x^2 - \mathrm{i}$ tiene raíces $\pm\frac{\sqrt{2}}{2}(1+\mathrm{i})$, que no son conjugadas entre sí. La hipótesis de coeficientes reales no es un detalle: es todo el contenido del teorema.

Idea clave

De ahí sale un corolario que en la escuela se enuncia sin explicación: *todo polinomio real de grado impar tiene al menos una raíz real.* Motivo: las raíces no reales vienen en parejas, luego son un número par de ellas; si el grado es impar, alguna raíz tiene que quedarse sin pareja, y esa es real. Compare con la demostración analítica por el teorema del valor intermedio (los límites elementales): son dos argumentos completamente distintos para el mismo hecho, uno algebraico y uno topológico.

Teorema 4.43 (Fundamental del álgebra)**.** *Todo polinomio no constante con coeficientes complejos tiene al menos una raíz en $\mathbb{C}$; y por tanto todo polinomio de grado n se factoriza en exactamente n factores lineales, contando multiplicidades.*

Ejemplo 4.44 (El teorema fundamental del álgebra, con ejemplos)**. (a) Contar con multiplicidad.** $x^3 - 3x + 2 = (x-1)^2(x+2)$ tiene tres raíces contando multiplicidades: el 1 dos veces y el -2 una. Sin contar multiplicidades sólo hay dos valores distintos, y por eso el enunciado especifica "con multiplicidad".

(b) Raíces de la unidad. $x^n = 1$ tiene exactamente n soluciones, $\mathrm{cis}(360°k/n)$ para $k = 0, \ldots, n-1$, que forman un polígono regular inscrito en la circunferencia unidad. Con $n = 6$ son los vértices de un hexágono, y su suma vale 0.

(c) Por qué hace falta $\mathbb{C}$. Sobre $\mathbb{R}$ el polinomio $x^2 + 1$ no tiene ninguna raíz; sobre $\mathbb{Q}$ tampoco la tiene $x^2 - 2$. Sólo al llegar a $\mathbb{C}$ el proceso se detiene: $\mathbb{C}$ es *algebraicamente cerrado*, y ampliarlo más no aporta raíces nuevas.

(d) No es un teorema algebraico. Pese al nombre, ninguna demostración es puramente algebraica: todas usan alguna propiedad topológica o analítica de $\mathbb{R}$, ya sea el teorema de

Liouville, el grado de una aplicación o la compacidad. La razón profunda es que la definición misma de $\mathbb{R}$ es analítica.

Idea clave

Esto es lo que hace que $\mathbb{C}$ sea el final del camino y no una parada más. Al pasar de $\mathbb{N}$ a $\mathbb{Z}$ ganamos la resta; de $\mathbb{Z}$ a $\mathbb{Q}$, la división; de $\mathbb{Q}$ a $\mathbb{R}$, los límites; de $\mathbb{R}$ a $\mathbb{C}$, las raíces de polinomios. Y ahí se acaba: **no hay que ampliar más**, porque $\mathbb{C}$ ya contiene las raíces de todos sus propios polinomios. Se dice que $\mathbb{C}$ es *algebraicamente cerrado* (capítulo 24). Curiosamente, ninguna demostración de este teorema es puramente algebraica: todas usan análisis o topología. La más corta es de una línea con el teorema de Liouville del capítulo 30.

4.10. Anticipo: la fórmula de Euler

Si se desarrolla en serie de potencias (capítulo 6) la exponencial y se le permite un argumento imaginario, aparece

$$e^{i\theta} = \cos\theta + i\sin\theta,$$

la **fórmula de Euler**. Con ella la forma polar se escribe $z = re^{i\theta}$, la regla del producto se vuelve la ley de exponentes $e^{i\theta_1}e^{i\theta_2} = e^{i(\theta_1+\theta_2)}$, y De Moivre se vuelve $(e^{i\theta})^n = e^{in\theta}$: es decir, los tres teoremas de este capítulo colapsan en una sola propiedad conocida desde los logaritmos.

Poniendo $\theta = \pi$ se obtiene la identidad más citada de la matemática:

$$e^{i\pi} + 1 = 0,$$

que reúne en cinco símbolos el 0 de la aritmética, el 1 de la multiplicación, el π de la geometría, el e del análisis (las sucesiones y series elementales) y el i del álgebra. Su significado geométrico, con lo que ya sabemos, es sencillo: girar media vuelta es multiplicar por -1.

Regla: deducir identidades trigonométricas con la fórmula de Euler

Nadie debería memorizar la lista de fórmulas del ángulo doble, triple y siguientes: se deducen todas en dos líneas a partir de $e^{i\theta} = \cos\theta + i\sin\theta$. El procedimiento tiene dos direcciones.

1. **Retenga las dos fórmulas inversas de Euler**, que salen de sumar y restar $e^{i\theta}$ y $e^{-i\theta}$:
$$\cos\theta = \frac{e^{i\theta} + e^{-i\theta}}{2}, \qquad \sin\theta = \frac{e^{i\theta} - e^{-i\theta}}{2i}.$$
2. **Para pasar de ángulo múltiple a potencias** (por ejemplo $\cos 3\theta$ en función de $\cos\theta$), parta de De Moivre: desarrolle $(\cos\theta + i\sin\theta)^n$ con el binomio de Newton.
3. **Separe la parte real y la imaginaria**: la real da $\cos n\theta$ y la imaginaria da $\sin n\theta$.
4. **Elimine el seno o el coseno** con $\sin^2\theta = 1 - \cos^2\theta$ para dejar la expresión en una sola función.

5. **Para el camino contrario**, de potencias a ángulos múltiples (*linealizar*), sustituya $\cos\theta$ o $\sin\theta$ por su fórmula de Euler, eleve a la potencia con el binomio y agrupe los términos $e^{ik\theta}$ con $e^{-ik\theta}$.
6. **Reconozca cada pareja** $e^{ik\theta}+e^{-ik\theta}=2\cos k\theta$ y $e^{ik\theta}-e^{-ik\theta}=2i\sin k\theta$ para volver a la trigonometría.
7. **Verifique** siempre en $\theta=0$ y en $\theta=\pi/2$, donde todo vale 0 o ±1: dos sustituciones detectan casi cualquier error de signo.

Ejemplo 4.45 (Seis identidades deducidas, ninguna memorizada)**.** (a) *Ángulo doble, de un plumazo.* $e^{2i\theta}=(e^{i\theta})^2=(\cos\theta+i\sin\theta)^2=\cos^2\theta-\sin^2\theta+2i\sin\theta\cos\theta$. Igualando con $\cos 2\theta+i\sin 2\theta$ y separando partes:

$$\cos 2\theta=\cos^2\theta-\sin^2\theta,\qquad \sin 2\theta=2\sin\theta\cos\theta.$$

Las dos fórmulas de un tirón, sin dibujar ningún triángulo.

(b) *Ángulo triple.* Ahora $(\cos\theta+i\sin\theta)^3=\cos^3\theta+3i\cos^2\theta\sin\theta-3\cos\theta\sin^2\theta-i\sin^3\theta$. La parte real es $\cos^3\theta-3\cos\theta\sin^2\theta$, y sustituyendo $\sin^2\theta=1-\cos^2\theta$ queda

$$\cos 3\theta=4\cos^3\theta-3\cos\theta.$$

La parte imaginaria da, del mismo modo, $\sin 3\theta=3\sin\theta-4\sin^3\theta$. Control en $\theta=0$: $4-3=1=\cos 0$ ✓. Control en $\theta=\pi/3$: $4\cdot\frac18-3\cdot\frac12=-1=\cos\pi$ ✓.

(c) *Ángulo quíntuple, sin más esfuerzo que paciencia.* Repitiendo con $n=5$ y limpiando los senos,

$$\cos 5\theta=16\cos^5\theta-20\cos^3\theta+5\cos\theta.$$

Aparece un patrón: $\cos n\theta$ es siempre un polinomio de grado n en $\cos\theta$, con coeficiente principal 2^{n-1}. Ese polinomio tiene nombre —polinomio de Chebyshev T_n— y reaparece en el capítulo 14 como la elección de nodos que evita el desastre de Runge. Es el mismo objeto visto desde dos sitios muy distintos.

(d) *Linealizar* $\cos^2\theta$*.* En la dirección contraria:

$$\cos^2\theta=\Big(\frac{e^{i\theta}+e^{-i\theta}}{2}\Big)^2=\frac{e^{2i\theta}+2+e^{-2i\theta}}{4}=\frac{2\cos 2\theta+2}{4}=\frac{1+\cos 2\theta}{2}.$$

Esta es la identidad que convierte $\int\cos^2 x\,\mathrm{d}x$ de un problema en una cuenta de diez segundos.

(e) *Linealizar* $\cos^4\theta$ *y* $\sin^3\theta$*.* Con el mismo método, elevando a la cuarta y agrupando,

$$\cos^4\theta=\frac{3+4\cos 2\theta+\cos 4\theta}{8},\qquad \sin^3\theta=\frac{3\sin\theta-\sin 3\theta}{4}.$$

Control de la primera en $\theta=0$: $(3+4+1)/8=1$ ✓. Control de la segunda en $\theta=\pi/2$: $(3-\sin\frac{3\pi}{2})/4=(3+1)/4=1$ ✓. Cualquier potencia de senos y cosenos se integra así, sin recurrir a ninguna tabla.

(f) *Una suma que parece imposible.* ¿Cuánto vale $1 + \cos\theta + \cos 2\theta + \cdots + \cos(n-1)\theta$? Sumando en vez las exponenciales, es una progresión geométrica de razón $e^{i\theta}$:

$$\sum_{k=0}^{n-1} e^{ik\theta} = \frac{e^{in\theta} - 1}{e^{i\theta} - 1},$$

y tomando parte real, tras simplificar,

$$\sum_{k=0}^{n-1} \cos k\theta = \frac{\sin\frac{n\theta}{2}}{\sin\frac{\theta}{2}} \cos\frac{(n-1)\theta}{2}.$$

Con $n = 5$ y $\theta = 0{,}37$ ambos miembros valen 3,2061 ✓. Sin complejos, esta suma se resiste; con ellos es la fórmula de la progresión geométrica de toda la vida.

Cuidado

Al usar $e^{i\theta}$ conviene no olvidar que la exponencial compleja *no* es inyectiva: $e^{i\theta} = e^{i(\theta+2\pi)}$, de modo que la exponencial es periódica de periodo $2\pi i$. En consecuencia el logaritmo complejo tampoco es una función bien definida sin elegir antes un corte, y expresiones como i^i tienen infinitos valores —todos reales, curiosamente, el principal $e^{-\pi/2} \approx 0{,}2079$—. Manipular e^z como si fuera la exponencial real es la fuente de la mitad de los errores del capítulo 30.

Pausa para pensar

Merece la pena detenerse en lo que acaba de ocurrir. La trigonometría del bachillerato es una colección de identidades sin patrón visible, que se memorizan y se olvidan. Al meter i en el exponente, todas ellas resultan ser consecuencias de la única ley $e^a e^b = e^{a+b}$. No es que los complejos "sirvan también" para la trigonometría: es que la trigonometría *es* el estudio de la exponencial en la circunferencia unidad, y sin i se está viendo de perfil.

Idea clave

Hay un puente más que conviene dejar señalado. Combinando las fórmulas inversas de Euler con la exponencial real aparecen las funciones hiperbólicas: $\cos(i\theta) = \cosh\theta$ y $\sin(i\theta) = i\sinh\theta$. Es decir, seno y coseno hiperbólicos *son* el seno y el coseno de siempre, evaluados en el eje imaginario. Todas las identidades hiperbólicas se obtienen de las trigonométricas cambiando algunos signos, y la regla mnemotécnica que las gobierna —la llamada regla de Osborn— no es más que la traducción de ese cambio de eje.

4.11. El logaritmo complejo y las potencias

La fórmula de Euler nos da e^z para cualquier z. La pregunta inversa —dado w, ¿qué z cumple $e^z = w$?— es más delicada de lo que parece, y su respuesta explica de una vez por qué expresiones como i^i tienen sentido y por qué tienen infinitos valores.

Idea clave

La exponencial compleja es periódica: $e^{z+2\pi i} = e^z$. Una función periódica no puede tener inversa única, igual que el seno no la tiene sin restringir antes su dominio. Así que el logaritmo complejo no es una función, sino una *familia* de valores que se diferencian en múltiplos de $2\pi i$; elegir uno de ellos es elegir una rama.

Si $w \neq 0$ se escribe en forma polar como $w = r\,e^{i\theta}$ con $r = |w| > 0$, entonces $e^z = w$ para $z = \ln r + i(\theta + 2k\pi)$, con k entero cualquiera. Se llama **logaritmo principal** al que se obtiene tomando el argumento principal $\theta \in (-\pi, \pi]$, y se escribe

$$\operatorname{Log} w = \ln|w| + i \operatorname{Arg} w.$$

La parte real es un logaritmo real de toda la vida; toda la novedad está en la parte imaginaria, que no es más que el ángulo.

Regla: calcular un logaritmo o una potencia compleja

1. **Comprueba que la base no es cero.** El logaritmo de 0 no existe, ni en $\mathbb{R}$ ni en $\mathbb{C}$.
2. **Pasa la base a forma polar**, obteniendo el módulo r y el argumento principal $\theta \in (-\pi, \pi]$; el cuadrante debe elegirse con el mismo cuidado de la sección de forma polar.
3. **Escribe el logaritmo principal** como $\operatorname{Log} w = \ln r + i\theta$, y la familia completa como $\ln r + i(\theta + 2k\pi)$ si el enunciado pide todos los valores.
4. **Para una potencia, usa la definición** $w^s = e^{s \operatorname{Log} w}$, que es la única forma sensata de dar sentido a un exponente no entero.
5. **Desarrolla el exponente** $s \operatorname{Log} w$ separando su parte real y su parte imaginaria: la real controla el módulo del resultado y la imaginaria su argumento.
6. **Aplica la fórmula de Euler** al resultado para escribirlo en forma binómica si hace falta.
7. **Indica si has tomado la rama principal.** Cuando el exponente no es entero hay infinitos valores, y omitir esa aclaración es una respuesta incompleta, no un detalle pedante.

Ejemplo 4.46 (Cuatro logaritmos).

1. $\operatorname{Log}(-1)$. El módulo es 1 y el argumento principal π, luego $\operatorname{Log}(-1) = \ln 1 + i\pi = i\pi$. Todos los valores son $i(\pi + 2k\pi)$. Aquí está, por fin, el sentido de la célebre igualdad $e^{i\pi} = -1$: los negativos sí tienen logaritmo, sólo que imaginario.
2. $\operatorname{Log}(i)$. Módulo 1, argumento $\frac{\pi}{2}$, así que $\operatorname{Log}(i) = i\frac{\pi}{2} \approx 1{,}5708\,i$.
3. $\operatorname{Log}(1+i)$. El módulo es $\sqrt{2}$ y el argumento $\frac{\pi}{4}$, de modo que $\operatorname{Log}(1+i) = \ln\sqrt{2} + i\frac{\pi}{4} \approx 0{,}34657 + 0{,}78540\,i$.

4. $\operatorname{Log}(-3)$. Módulo 3, argumento π: $\ln 3 + i\pi \approx 1{,}09861 + 3{,}14159\,i$. Compáralo con $\operatorname{Log}(2i) = \ln 2 + i\frac{\pi}{2} \approx 0{,}69315 + 1{,}57080\,i$.

Ejemplo 4.47 (Potencias imposibles que sí se pueden)**.** 1. i^i. Por definición vale $e^{i\operatorname{Log} i} = e^{i\cdot i\pi/2} = e^{-\pi/2} \approx 0{,}20788$. Es un número *real*, lo cual desconcierta a todo el mundo la primera vez. Los demás valores son $e^{-\pi/2-2k\pi}$: para $k=1$ sale 0,00039 y para $k=-1$ sale 111,318, todos reales y positivos.

2. $(-1)^i = e^{i\operatorname{Log}(-1)} = e^{i\cdot i\pi} = e^{-\pi} \approx 0{,}043214$, otra vez real.

3. $2^i = e^{i\ln 2} = \cos(\ln 2) + i\sin(\ln 2) \approx 0{,}76924 + 0{,}63896\,i$, un punto de la circunferencia unidad, como corresponde a un exponente imaginario puro sobre una base real positiva.

4. $(1+i)^{1-i}$. Con $\operatorname{Log}(1+i) = \ln\sqrt{2} + i\frac{\pi}{4}$, el exponente es $(1-i)\left(\ln\sqrt{2} + i\frac{\pi}{4}\right) = \left(\ln\sqrt{2} + \frac{\pi}{4}\right) + i\left(\frac{\pi}{4} - \ln\sqrt{2}\right)$, y el resultado principal es $\approx 2{,}80788 + 1{,}31787\,i$.

Cuidado

La propiedad $\log(ab) = \log a + \log b$ **no** sobrevive intacta al paso a los complejos. Tomando $a = b = -1$, el primer miembro es $\operatorname{Log}(1) = 0$ y el segundo $i\pi + i\pi = 2\pi i$: difieren en un múltiplo de $2\pi i$, exactamente el margen de indeterminación del argumento. La igualdad correcta es "módulo $2\pi i$", o bien es exacta si se trabaja con la familia completa de valores en lugar de con la rama principal. Lo mismo le ocurre a $(w^s)^t = w^{st}$, que sólo es fiable cuando los exponentes son enteros.

Pausa para pensar

Elegir el argumento en $(-\pi, \pi]$ es cómodo pero arbitrario, y tiene un precio: la función Log es discontinua a lo largo del semieje real negativo, el llamado *corte de rama*. Si en lugar de cortar el plano permitimos que el argumento crezca libremente al dar vueltas alrededor del origen, se obtiene una superficie en espiral —la superficie de Riemann del logaritmo— sobre la cual el logaritmo sí es una función continua y biyectiva. Esa idea, que reconcilia lo local con lo global, es uno de los hilos conductores del capítulo 30.

4.12. Transformaciones del plano y la esfera de Riemann

Cada función de variable compleja es una manera de mover el plano entero. Las más simples —sumar una constante, multiplicar por una constante, invertir— ya bastan para generar una familia sorprendentemente rica, que aparece en geometría, en relatividad y en el diseño de mapas.

Idea clave

Sumar b es trasladar. Multiplicar por $a = \rho\, e^{i\alpha}$ es girar un ángulo α y escalar un factor ρ a la vez: la multiplicación compleja *es* una semejanza directa. Y $z \mapsto 1/z$ es invertir respecto de la circunferencia unidad y reflejar en el eje real. Componiendo estas tres

operaciones se obtienen todas las transformaciones $z \mapsto \dfrac{az+b}{cz+d}$, llamadas de Möbius.

Regla: estudiar una transformación de Möbius

1. **Comprueba que $ad-bc \neq 0$.** Si el determinante se anula la función es constante y no transforma nada.
2. **Localiza el polo y el valor en el infinito**: la imagen de $z=-d/c$ es ∞ y la de ∞ es a/c. Con esos dos convenios la transformación es una biyección del plano ampliado $\mathbb{C}\cup\{\infty\}$ en sí mismo.
3. **Descomponla en pasos elementales** cuando $c\neq 0$, dividiendo: $\dfrac{az+b}{cz+d} = \dfrac{a}{c} + \dfrac{bc-ad}{c^2}\cdot\dfrac{1}{z+d/c}$, que es una traslación, una inversión, una semejanza y otra traslación.
4. **Recuerda el principio circular**: la imagen de una recta o una circunferencia es siempre una recta o una circunferencia, aunque el tipo pueda cambiar.
5. **Para hallar la imagen de una figura, transforma tres de sus puntos.** Tres puntos determinan una circunferencia o una recta, así que no hace falta más.
6. **Calcula la inversa** despejando z; la inversa de una transformación de Möbius es de nuevo de Möbius, con la matriz inversa de coeficientes.
7. **Busca los puntos fijos** resolviendo $\dfrac{az+b}{cz+d}=z$, que es una ecuación de segundo grado: hay a lo sumo dos, salvo la identidad.

Ejemplo 4.48 (La inversión $z\mapsto 1/z$ en acción)**.** 1. *La circunferencia unidad se queda en su sitio*, pero recorrida al revés: $1/i=-i$ y $1/(-1)=-1$. En general, si $|z|=1$ entonces $1/z=\bar{z}$, que es la reflexión en el eje real. Y $|z|=2$ va a parar a la circunferencia $|w|=\frac{1}{2}$: los puntos lejanos se acercan y los cercanos se alejan.

2. *Una recta que no pasa por el origen se convierte en circunferencia.* La recta $x=1$ tiene por imagen los w con $\operatorname{Re}(1/w)=1$, es decir $\dfrac{X}{X^2+Y^2}=1$, o sea $X^2+Y^2-X=0$: la circunferencia de centro $\left(\frac{1}{2},0\right)$ y radio $\frac{1}{2}$, que pasa por el origen. Comprobación: $1\mapsto 1$, $1+i\mapsto \frac{1}{2}-\frac{1}{2}i$ y $1+2i\mapsto \frac{1}{5}-\frac{2}{5}i$, y los tres distan $\frac{1}{2}$ del centro.

3. *Una recta por el origen sigue siendo una recta*, porque el origen va al infinito. La recta $y=x$ se transforma en $y=-x$: basta ver que $\operatorname{Im}(1/w)=-\operatorname{Re}(1/w)$ equivale a $X+Y=0$.

4. *Puntos fijos.* Los de $z\mapsto 1/z$ son las soluciones de $z^2=1$, o sea ± 1. Los de $z\mapsto \dfrac{z+1}{z-1}$ son $1\pm\sqrt{2}$. Y $z\mapsto z+3$ no tiene ninguno en el plano: su único punto fijo es ∞, lo cual explica geométricamente por qué una traslación no tiene centro.

Ejemplo 4.49 (Del disco al semiplano, y del plano a la esfera)**.** 1. *Una transformación célebre.* $f(z) = \dfrac{z-i}{z+i}$ manda el eje real entero a la circunferencia unidad: $f(0) = -1$, $f(1) = -i$, $f(-1) = i$, y en los tres casos el módulo es 1. Además $f(i) = 0$ y $f(-i) = \infty$, de modo que el semiplano superior se convierte en el disco unidad. Su inversa es $z = i\dfrac{1+w}{1-w}$.

2. *Componer es multiplicar matrices.* A la transformación $\dfrac{az+b}{cz+d}$ se le asocia la matriz $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$, y componer dos transformaciones corresponde a multiplicar sus matrices. Con $\begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}$ y $\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$ el producto es $\begin{pmatrix} 2 & 1 \\ 4 & 3 \end{pmatrix}$, y los determinantes -2 y -1 dan 2, como debe ser. Las transformaciones de Möbius forman así un grupo, tema del capítulo 23.

3. *El punto del infinito, en serio.* Añadir un único ∞ al plano parece un truco contable hasta que se ve la construcción de Riemann: se coloca una esfera de radio 1 apoyada en el origen y se proyecta desde su polo norte. El punto $z = x + iy$ va a parar a $\left(\dfrac{2x}{1+|z|^2}, \dfrac{2y}{1+|z|^2}, \dfrac{|z|^2-1}{1+|z|^2}\right)$, que está siempre sobre la esfera.

4. *Algunos valores de esa proyección.* El origen va al polo sur $(0,0,-1)$; el punto 1 va a $(1,0,0)$; el punto i va a $(0,1,0)$; y $z = 2$ va a $\left(\frac{4}{5}, 0, \frac{3}{5}\right)$. Cuanto mayor es $|z|$, más cerca queda su imagen del polo norte, que es precisamente el punto reservado para ∞.

Cuidado

El principio circular dice que rectas y circunferencias se transforman en rectas o circunferencias, pero no que cada tipo se conserve: la recta $x = 1$ se volvió una circunferencia y la circunferencia unidad podría volverse una recta. La clave está en si la figura pasa o no por el polo de la transformación. Otro error frecuente es creer que el centro de una circunferencia va al centro de su imagen; no ocurre casi nunca, porque estas aplicaciones conservan ángulos pero no proporciones radiales.

Pausa para pensar

En la esfera de Riemann desaparece la distinción entre rectas y circunferencias: unas y otras son circunferencias sobre la esfera, y las rectas son simplemente las que pasan por el polo norte. Ese cambio de punto de vista —compactificar añadiendo un solo punto— es una de las primeras construcciones serias de topología que encontrarás en el capítulo 28, y explica por qué el plano complejo ampliado se comporta mucho mejor que el plano a secas.

4.13. Errores típicos con complejos

1. Escribir $\sqrt{-4}\cdot\sqrt{-9} = \sqrt{36} = 6$. **Falso:** la regla $\sqrt{a}\sqrt{b} = \sqrt{ab}$ solo vale para $a, b \geq 0$. Lo correcto es $2i \cdot 3i = 6i^2 = -6$. Conviértase siempre a la forma bi *antes* de multiplicar.

2. Decir que $\operatorname{Im}(3-7i) = -7i$. Es -7.
3. Olvidar multiplicar por el conjugado al dividir y dejar i en el denominador.
4. Escribir $|a+bi| = a+b$ o $= a^2+b^2$. Es $\sqrt{a^2+b^2}$.
5. Dar el argumento con la arcotangente sin mirar el cuadrante.
6. Comparar complejos: $2+i > 1+3i$ no significa nada.
7. Creer que $z^2 \geq 0$ siempre. Con $z = i$ sale -1.
8. Dar una sola raíz cúbica. En $\mathbb{C}$ hay siempre exactamente tres.
9. Al aplicar De Moivre, olvidar elevar también el módulo: $(1+i)^{10} \neq \operatorname{cis} 450°$.
10. Pensar que $\bar{z} = -z$. Eso solo ocurre si z es imaginario puro.

4.14. Seno y coseno de un número complejo, y su parentesco con las hiperbólicas

Hemos llegado al logaritmo complejo y a las potencias z^w. Falta el paso natural siguiente, y trae una sorpresa.

Definición 4.50. De la fórmula de Euler, despejando:

$$\cos z = \frac{e^{iz}+e^{-iz}}{2}, \qquad \sin z = \frac{e^{iz}-e^{-iz}}{2i},$$
$$\cosh z = \frac{e^{z}+e^{-z}}{2}, \qquad \sinh z = \frac{e^{z}-e^{-z}}{2}.$$

Para z real coinciden con las de siempre; para z complejo son la única extensión holomorfa posible.

Teorema 4.51 (El parentesco)**.**

$$\cos(iy) = \cosh y, \qquad \sin(iy) = i\sinh y, \qquad \cosh(iy) = \cos y, \qquad \sinh(iy) = i\sin y.$$

Y separando en parte real e imaginaria:

$$\cos(x+iy) = \cos x\cosh y - i\,\sin x\sinh y, \qquad \sin(x+iy) = \sin x\cosh y + i\,\cos x\sinh y.$$

Además $\cos^2 z + \sin^2 z = 1$ *sigue valiendo para todo* z *complejo, y las fórmulas de adición también.*

Idea clave

Las funciones circulares y las hiperbólicas no son dos familias parecidas por casualidad: son *la misma función* mirada en dos direcciones perpendiculares del plano complejo. El coseno recorre $[-1,1]$ sobre el eje real y crece sin límite sobre el imaginario. Toda

identidad hiperbólica es una identidad trigonométrica con y sustituido por iy, y de ahí la regla de Osborn que en los libros de bachillerato aparece como truco sin explicación: cambia cada producto de dos senos de signo.

Ejemplo 4.52 (El seno y el coseno no están acotados)**.** De $\cos(iy) = \cosh y$ se sigue

$$\cos(i) = \cosh 1 = 1{,}543\,081 > 1,$$

y $\cos(iy) \to \infty$ cuando $y \to \infty$. La cota $|\cos x| \le 1$, que en el eje real parece una propiedad esencial, es un accidente de mirar solo una recta.

De hecho eso era inevitable: por el teorema de Liouville —toda función entera y acotada es constante— ninguna función entera no constante puede estar acotada en todo $\mathbb{C}$. El seno y el coseno son enteros y no constantes, luego *tienen* que explotar en alguna dirección.

Ejemplo 4.53 (Resolver $\cos z = 2$)**.** Poniendo $w = e^{iz}$, la ecuación $\frac{w+w^{-1}}{2} = 2$ se convierte en

$$w^2 - 4w + 1 = 0 \qquad \Longrightarrow \qquad w = 2 \pm \sqrt{3}.$$

Ambas raíces son reales y positivas, luego $iz = \log w$ es real y

$$z = -i \log(2 \pm \sqrt{3}) = \mp i \log(2 + \sqrt{3}) = \mp 1{,}316\,958\, i,$$

más los múltiplos de 2π que quiera añadirse, porque $2 - \sqrt{3} = (2 + \sqrt{3})^{-1}$.

En $\mathbb{R}$ la ecuación $\cos x = 2$ no tiene solución; en $\mathbb{C}$ tiene infinitas, todas sobre el eje imaginario. Es el mismo salto conceptual que llevó de $x^2 = -1$ a los complejos, ocurriendo otra vez un piso más arriba.

4.15. Series geométricas, sumas trigonométricas y el diccionario geométrico

Series

Teorema 4.54. *Si $|z| < 1$, entonces $\sum_{k\ge 0} z^k = \dfrac{1}{1-z}$, y la convergencia es absoluta. Y para todo z,*

$$e^z = \sum_{k\ge 0} \frac{z^k}{k!}, \qquad e^{z+w} = e^z e^w,$$

esta última por producto de Cauchy y el binomio de Newton. Con eso la fórmula de Euler deja de ser un cálculo informal y pasa a ser un teorema.

También vale la suma finita, para todo $z \neq 1$:

$$\sum_{k=0}^{n-1} z^k = \frac{1 - z^n}{1 - z}.$$

Ejemplo 4.55 (Sumar cosenos sin una sola identidad trigonométrica)**.** Se quiere calcular $\sum_{k=0}^{n-1} \cos k\theta$. En lugar de encadenar fórmulas de adición, se toma la parte real de una geométrica compleja:

$$\sum_{k=0}^{n-1} e^{ik\theta} = \frac{1-e^{in\theta}}{1-e^{i\theta}}.$$

Sacando factor común $e^{in\theta/2}$ arriba y $e^{i\theta/2}$ abajo, y usando $e^{i\alpha}-e^{-i\alpha}=2i\sin\alpha$:

$$\frac{1-e^{in\theta}}{1-e^{i\theta}} = e^{i(n-1)\theta/2}\,\frac{\sin(n\theta/2)}{\sin(\theta/2)}.$$

Tomando partes real e imaginaria salen *las dos* sumas de un golpe:

$$\sum_{k=0}^{n-1}\cos k\theta = \frac{\sin(n\theta/2)}{\sin(\theta/2)}\cos\frac{(n-1)\theta}{2}, \qquad \sum_{k=0}^{n-1}\sin k\theta = \frac{\sin(n\theta/2)}{\sin(\theta/2)}\sin\frac{(n-1)\theta}{2}.$$

Comprobación con $n=3$ y $\theta=2\pi/3$: el factor $\sin(3\pi/3)=\sin\pi=0$, luego ambas sumas valen 0. Y en efecto $\cos 0+\cos\frac{2\pi}{3}+\cos\frac{4\pi}{3}=1-\frac12-\frac12=0$. ✓

El factor $\frac{\sin(n\theta/2)}{\sin(\theta/2)}$ que ha aparecido solo es el *núcleo de Dirichlet*, el mismo que gobierna la convergencia de las series de Fourier. Aquí ha salido de una serie geométrica de cuatro líneas.

El diccionario métrico

Regla: traducir geometría a complejos

Para $z, w \in \mathbb{C}$ vistos como vectores del plano:

producto escalar	$\langle z, w\rangle = \operatorname{Re}(\bar z w)$
doble del área orientada	$\det(z,w) = \operatorname{Im}(\bar z w)$
perpendicularidad	$\bar z w$ imaginario puro
a, b, c alineados	$\dfrac{c-a}{b-a} \in \mathbb{R}$
giro de α en torno a p	$z \mapsto p + e^{i\alpha}(z-p)$

Ejemplo 4.56 (Producto escalar y área de un tirón)**.** Sean $z=3+i$ y $w=1+2i$. Entonces

$$\bar z w = (3-i)(1+2i) = 3+6i-i+2 = 5+5i,$$

de donde $\langle z,w\rangle = 5$ y el área del triángulo de vértices 0, z, w es $|\operatorname{Im}|/2 = 5/2 = 2{,}5$. Un solo producto de complejos da a la vez el producto escalar y el determinante, que en $\mathbb{R}^2$ exigirían dos cuentas separadas.

Ejemplo 4.57 (Triángulo equilátero, caracterizado sin trigonometría)**.** Los vértices a, b, c forman un triángulo equilátero orientado positivamente si y solo si

$$a + \omega b + \omega^2 c = 0, \qquad \omega = e^{2\pi i/3}.$$

Comprobación con $a=0$, $b=1$, $c=\frac{1}{2}+\frac{\sqrt{3}}{2}i$: como $\omega=-\frac{1}{2}+\frac{\sqrt{3}}{2}i$ y $\omega^2=-\frac{1}{2}-\frac{\sqrt{3}}{2}i$,

$$\omega b=-\tfrac{1}{2}+\tfrac{\sqrt{3}}{2}i, \qquad \omega^2 c=\left(-\tfrac{1}{2}-\tfrac{\sqrt{3}}{2}i\right)\left(\tfrac{1}{2}+\tfrac{\sqrt{3}}{2}i\right)=\tfrac{1}{2}-\tfrac{\sqrt{3}}{2}i,$$

y la suma es 0. ✓

La razón de fondo es que multiplicar por ω *es* girar 120°, así que la condición dice literalmente que los tres vértices se llevan unos a otros por giros de un tercio de vuelta. Toda una familia de resultados de geometría elemental —el teorema de Napoleón, el de Van Aubel— se demuestran en unas pocas líneas con este lenguaje, mientras que en coordenadas cartesianas exigen páginas.

4.16. Dónde se usan de verdad: fasores, impedancia y oscilaciones

Todos los capítulos "desde cero" de este libro terminan mostrando para qué sirve lo aprendido. Este es el turno de los complejos, y su aplicación es tan directa que llegó a reorganizar una ingeniería entera.

Idea clave

El problema práctico es este: en un circuito de corriente alterna todo oscila a la misma frecuencia, y lo único que distingue a una señal de otra es su *amplitud* y su *fase*. Sumar dos sinusoides desfasadas con identidades trigonométricas es tedioso y propenso a errores. Pero amplitud y fase son exactamente módulo y argumento: cada sinusoide se codifica en un solo número complejo, y sumar señales se convierte en sumar vectores.

Definición 4.58 (Fasor)**.** A la señal $v(t)=A\cos(\omega t+\phi)$ se le asocia el **fasor** $\mathbf{V}=Ae^{i\phi}$, de modo que $v(t)=\operatorname{Re}(\mathbf{V}e^{i\omega t})$. Entonces

$$\text{derivar} \;\longleftrightarrow\; \text{multiplicar por } i\omega, \qquad \text{integrar} \;\longleftrightarrow\; \text{dividir por } i\omega.$$

Esa es toda la magia: el cálculo diferencial se convierte en aritmética. Una ecuación diferencial lineal con forzamiento sinusoidal pasa a ser una ecuación algebraica en $\mathbb{C}$.

Definición 4.59 (Impedancia)**.** Aplicando la regla a las tres leyes de los componentes:

$$Z_R=R, \qquad Z_L=i\omega L, \qquad Z_C=\frac{1}{i\omega C}=-\frac{i}{\omega C},$$

y la ley de Ohm generalizada es $\mathbf{V}=Z\mathbf{I}$, con las impedancias combinándose en serie y en paralelo igual que las resistencias.

Ejemplo 4.60 (Un circuito RLC en serie)**.** Sea $R=3\,\Omega$, con reactancias $X_L=\omega L=7\,\Omega$ y $X_C=1/(\omega C)=3\,\Omega$. La impedancia total es

$$Z=R+i(X_L-X_C)=3+4i,$$

luego

$$|Z| = \sqrt{9+16} = 5\ \Omega, \qquad \arg Z = \arctan\frac{4}{3} = 53{,}13^\circ.$$

Con una tensión eficaz de 100 V, la corriente es de 20 A y va *retrasada* 53,13° respecto de la tensión, porque el circuito es inductivo ($X_L > X_C$). El factor de potencia es $\cos 53{,}13^\circ = 0{,}6$: solo el 60 por ciento del producto de tensión y corriente se convierte en potencia útil.

Y la resonancia se lee de un vistazo: ocurre cuando $X_L = X_C$, es decir $\omega = 1/\sqrt{LC}$, porque entonces $Z = R$ es puramente real, mínima en módulo, y la corriente es máxima y está en fase. Toda la teoría de la sintonía de una radio cabe en esa frase.

Ejemplo 4.61 (Sumar señales es sumar vectores)**.** Sean $v_1 = 3\cos\omega t$ y $v_2 = 4\cos(\omega t + 90^\circ)$. Con fasores, $\mathbf{V}_1 = 3$ y $\mathbf{V}_2 = 4i$, luego

$$\mathbf{V}_1 + \mathbf{V}_2 = 3 + 4i \qquad \Longrightarrow \qquad v_1 + v_2 = 5\cos(\omega t + 53{,}13^\circ).$$

Amplitud 5 y fase 53,13°, obtenidas con una suma de dos números y sin escribir ni una identidad trigonométrica. Hacerlo con $\cos(a+b)$ ocuparía media página.

Observación 4.62 (Más allá de los circuitos)*.* La misma codificación aparece en cuanto hay oscilaciones. Un exponente complejo con parte real negativa,

$$e^{(-\alpha+i\omega)t} = e^{-\alpha t}\big(\cos\omega t + i\sin\omega t\big),$$

describe una oscilación amortiguada: la parte real gobierna el decaimiento y la imaginaria la frecuencia. Ese único objeto unifica el muelle con rozamiento, el circuito RLC en régimen transitorio y la desintegración con oscilación.

En óptica y en acústica, sumar ondas es sumar fasores, y la interferencia constructiva o destructiva es simplemente que los vectores apunten en el mismo sentido o en sentidos opuestos. Y en mecánica cuántica, la unidad imaginaria está en el corazón de la ecuación de Schrödinger,

$$i\hbar\,\partial_t\psi = H\psi,$$

donde el factor i es lo que convierte una ecuación de difusión —que aplasta— en una ecuación de ondas —que oscila y conserva la norma—. Los "números imaginarios" que Cardano usó a regañadientes como rodeo de cálculo resultaron ser el lenguaje en que está escrita la física fundamental.

4.17. La raíz cuadrada de un número complejo en forma binómica, y la ecuación de segundo grado con coeficientes cualesquiera

La sección anterior resolvió las ecuaciones $ax^2 + bx + c = 0$ con coeficientes *reales* y discriminante negativo, y la de las raíces n-ésimas dio un método polar para extraer raíces. Falta algo intermedio que se usa constantemente y que no se ha dicho: cómo calcular $\sqrt{a+bi}$ *en forma binómica, sin pasar por ángulos*, y qué ocurre con la ecuación de segundo grado cuando los coeficientes son complejos y el discriminante ya no es un número real cuyo signo se pueda mirar.

Idea clave

El método polar de la sección anterior es impecable pero tiene un inconveniente práctico: obliga a calcular un arcotangente, dividirlo entre dos y volver a evaluar un seno y un coseno, con lo que una raíz exacta como $\sqrt{-5+12i} = 2+3i$ sale disfrazada de decimales. El método binómico ataca el problema al revés: se *postula* que la raíz es $x+yi$, se eleva al cuadrado, se identifican partes real e imaginaria y se resuelve el sistema. Sale un sistema de dos ecuaciones que, con un truco de una línea, se desacopla por completo. Cuando la respuesta es exacta, este método la encuentra exacta.

Definición 4.63 (Raíces cuadradas de un complejo)**.** Dado $w \in \mathbb{C}$, se llama *raíz cuadrada* de w a todo $z \in \mathbb{C}$ con $z^2 = w$. Si $w \neq 0$ hay exactamente dos, y son opuestas: si z_0 es una, la otra es $-z_0$. El símbolo $\sqrt{w}$ *no* designa a ninguna de las dos en particular mientras no se fije una rama, de modo que en este libro se escribirá siempre "las raíces cuadradas de w son $\pm z_0$".

Teorema 4.64 (Fórmula binómica de la raíz cuadrada)**.** *Sea $w = a+bi$ con $b \neq 0$ y sea $\rho = |w| = \sqrt{a^2+b^2}$. Entonces las dos raíces cuadradas de w son $\pm(x+yi)$ con*

$$x = \sqrt{\frac{\rho + a}{2}}, \qquad y = \operatorname{sgn}(b)\sqrt{\frac{\rho - a}{2}},$$

donde $\operatorname{sgn}(b)$ vale $+1$ si $b>0$ y -1 si $b<0$. Si $b=0$ y $a>0$ las raíces son $\pm\sqrt{a}$; si $b=0$ y $a<0$ son $\pm i\sqrt{|a|}$.

Demostración. Impónganse $z = x+yi$ y $z^2 = w$. Desarrollando, $z^2 = (x^2-y^2) + 2xy\,i$, y la igualdad de complejos equivale a las dos ecuaciones reales

$$x^2 - y^2 = a, \qquad 2xy = b.$$

Añádase una tercera, que no es independiente pero desacopla el sistema: tomando módulos en $z^2 = w$ se obtiene $|z|^2 = |w|$, es decir

$$x^2 + y^2 = \rho.$$

Sumando y restando esta con la primera se despeja de golpe $x^2 = \frac{\rho+a}{2}$ y $y^2 = \frac{\rho-a}{2}$, ambos no negativos porque $\rho \geq |a|$. Quedan cuatro combinaciones de signos, y la ecuación $2xy = b$ elimina dos: el producto xy debe tener el signo de b. Eligiendo $x > 0$, el signo de y queda forzado a ser el de b; la otra raíz es la opuesta. □

Regla: extraer la raíz cuadrada de un complejo sin usar ángulos

1. **Escribe el número como** $a+bi$ y calcula el módulo $\rho = \sqrt{a^2+b^2}$. Si ρ sale entero, casi seguro la raíz también lo es y merece la pena buscarla exacta.
2. **Calcula** $x^2 = \dfrac{\rho+a}{2}$ **y** $y^2 = \dfrac{\rho-a}{2}$**.** Los dos son automáticamente ≥ 0; si alguno te sale negativo, has equivocado el módulo.
3. **Toma las raíces positivas** x e y de esos dos números.

4. **Ponle a y el signo de b**, dejando x positivo. Este es el único paso donde se puede perder el resultado, y es también el que sustituye a toda la discusión de cuadrantes del método polar.
5. **Escribe las dos respuestas**, $\pm(x+yi)$. Nunca una sola.
6. **Comprueba elevando al cuadrado.** Es una multiplicación de diez segundos y cierra el problema sin margen de duda.
7. **Control de módulos.** Debe cumplirse $|z| = \sqrt{\rho}$. Si el número de partida tenía módulo 25, la raíz debe tener módulo 5.

Ejemplo 4.65 (Cinco raíces cuadradas extraídas y comprobadas)**.** (a) $w = 3+4i$. Módulo $\rho = \sqrt{9+16} = 5$. Entonces $x^2 = \frac{5+3}{2} = 4$ y $y^2 = \frac{5-3}{2} = 1$, luego $x = 2$, $y = 1$ (positivo, porque $b = 4 > 0$). Las raíces son $\pm(2+i)$. Comprobación: $(2+i)^2 = 4+4i-1 = 3+4i$ ✓, y $|2+i| = \sqrt{5} = \sqrt{\rho}$ ✓.

(b) $w = -5+12i$. Módulo $\sqrt{25+144} = 13$. Ahora $x^2 = \frac{13-5}{2} = 4$ y $y^2 = \frac{13+5}{2} = 9$: $x = 2$, $y = 3$. Raíces $\pm(2+3i)$, y en efecto $(2+3i)^2 = 4+12i-9 = -5+12i$ ✓. Obsérvese cómo la parte real negativa hace que la raíz sea "más imaginaria que real".

(c) $w = -7-24i$. Módulo $\sqrt{49+576} = 25$. Sale $x^2 = \frac{25-7}{2} = 9$ y $y^2 = \frac{25+7}{2} = 16$, luego $x = 3$ y, como $b = -24 < 0$, $y = -4$. Raíces $\pm(3-4i)$: $(3-4i)^2 = 9-24i-16 = -7-24i$ ✓. Aquí se ve para qué sirve el paso 4: con $y = +4$ habría salido $-7+24i$, el conjugado, que no es el número pedido.

(d) $w = i$. Módulo 1, $a = 0$, luego $x^2 = y^2 = \frac{1}{2}$ y las raíces son $\pm\frac{\sqrt{2}}{2}(1+i)$. Comprobación: $\left(\frac{\sqrt{2}}{2} + \frac{\sqrt{2}}{2}i\right)^2 = \frac{1}{2} + 2\cdot\frac{1}{2}i - \frac{1}{2} = i$ ✓. Numéricamente, $\frac{\sqrt{2}}{2} = 0{,}707107$.

(e) $w = 2i$. Módulo 2, $x^2 = y^2 = 1$: las raíces son $\pm(1+i)$, y $(1+i)^2 = 2i$ ✓. Este caso conviene memorizarlo, porque aparece constantemente.

Teorema 4.66 (La fórmula cuadrática vale sobre $\mathbb{C}$)**.** *Sean $A, B, C \in \mathbb{C}$ con $A \neq 0$. Las soluciones de $Az^2 + Bz + C = 0$ son*

$$z = \frac{-B \pm \delta}{2A},$$

donde δ es cualquiera *de las dos raíces cuadradas de $\Delta = B^2 - 4AC$. En particular la ecuación tiene siempre exactamente dos soluciones (una doble si $\Delta = 0$), y se cumplen las relaciones de Cardano–Vieta $z_1 + z_2 = -B/A$ y $z_1 z_2 = C/A$.*

Demostración. Es el mismo completar cuadrados de siempre, y en $\mathbb{C}$ funciona sin ninguna salvedad porque la división entre $2A \neq 0$ es legítima y las raíces cuadradas existen siempre. De $Az^2 + Bz + C = 0$, dividiendo entre A y sumando y restando $B^2/(4A^2)$:

$$\left(z + \frac{B}{2A}\right)^2 = \frac{B^2 - 4AC}{4A^2} = \frac{\Delta}{4A^2}.$$

Si $\delta^2 = \Delta$, el miembro derecho es $\left(\frac{\delta}{2A}\right)^2$, luego $z + \frac{B}{2A} = \pm\frac{\delta}{2A}$, que es la fórmula. Escribir $\pm$ delante de δ hace irrelevante cuál de las dos raíces se haya elegido. Sumando y multiplicando las dos soluciones se obtienen las relaciones de Vieta. □

Cuidado

La diferencia con el caso real es que **el signo de Δ ya no se puede mirar**, porque Δ no es un número real y los complejos no se ordenan. La pregunta "¿el discriminante es positivo o negativo?" deja de tener sentido; la única pregunta que queda es "¿es cero o no?". Y las raíces *ya no vienen en pareja conjugada*: el teorema de las raíces conjugadas exigía coeficientes reales, y aquí no los hay. En el ejemplo siguiente, las soluciones $2+3i$ y $1-i$ no son conjugadas la una de la otra ni tienen por qué serlo.

Ejemplo 4.67 (Tres cuadráticas con coeficientes complejos, resueltas enteras)**. (a) Coeficiente principal 1.** Resuélvase $z^2 - (3+2i)z + (5+i) = 0$.

Discriminante. $\Delta = (3+2i)^2 - 4(5+i) = (9+12i-4) - (20+4i) = (5+12i) - (20+4i) = -15+8i$.

Su raíz cuadrada. Módulo $\rho = \sqrt{225+64} = \sqrt{289} = 17$. Entonces $x^2 = \frac{17-15}{2} = 1$ y $y^2 = \frac{17+15}{2} = 16$, y como $b = 8 > 0$ sale $\delta = 1+4i$. Comprobación: $(1+4i)^2 = 1+8i-16 = -15+8i$ ✓.

Las soluciones.

$$z = \frac{(3+2i) \pm (1+4i)}{2} = \frac{4+6i}{2} = 2+3i \qquad \text{o} \qquad \frac{2-2i}{2} = 1-i.$$

Comprobación por Vieta. La suma es $(2+3i)+(1-i) = 3+2i$, que es $-B/A$ ✓. El producto es $(2+3i)(1-i) = 2-2i+3i+3 = 5+i$, que es C/A ✓. Las dos relaciones cuadran, luego las dos soluciones son correctas.

(b) Otro, con una raíz imaginaria pura. Resuélvase $z^2 + (1-i)z + (2+i) = 0$.

Discriminante. $\Delta = (1-i)^2 - 4(2+i) = (1-2i-1) - (8+4i) = -2i-8-4i = -8-6i$.

Su raíz. Módulo $\sqrt{64+36} = 10$; $x^2 = \frac{10-8}{2} = 1$, $y^2 = \frac{10+8}{2} = 9$, y como $b = -6 < 0$, $\delta = 1-3i$. En efecto $(1-3i)^2 = 1-6i-9 = -8-6i$ ✓.

Las soluciones.

$$z = \frac{-(1-i) \pm (1-3i)}{2},$$

que con el signo + da $\frac{-1+i+1-3i}{2} = \frac{-2i}{2} = -i$, y con el signo − da $\frac{-1+i-1+3i}{2} = \frac{-2+4i}{2} = -1+2i$.

Comprobación. Suma: $-i + (-1+2i) = -1+i = -(1-i)$ ✓. Producto: $(-i)(-1+2i) = i - 2i^2 = i+2 = 2+i$ ✓. Y una comprobación directa sobre la primera raíz: $(-i)^2 + (1-i)(-i) + (2+i) = -1 + (-i+i^2) + 2 + i = -1-i-1+2+i = 0$ ✓.

(c) Con coeficiente principal complejo. Resuélvase $i\,z^2 - 3i\,z + (-1+3i) = 0$, donde $A = i$, $B = -3i$ y $C = -1+3i$.

Discriminante. $\Delta = (-3i)^2 - 4 \cdot i \cdot (-1+3i) = -9 - 4i(-1+3i) = -9+4i-12i^2 = -9+4i+12 = 3+4i$.

Su raíz. Ya calculada en el ejemplo anterior: $\delta = 2+i$.

Las soluciones. Con $2A = 2i$,

$$z = \frac{3i \pm (2+i)}{2i}.$$

Para el signo $+$: $\dfrac{2+4i}{2i} = \dfrac{1+2i}{i} = \dfrac{(1+2i)(-i)}{-i\cdot i} = -i+2 = 2-i$. Para el signo $-$: $\dfrac{-2+2i}{2i} = \dfrac{-1+i}{i} = (-1+i)(-i) = i+1 = 1+i$.

Comprobación por Vieta. Suma: $(2-i)+(1+i) = 3$, y $-B/A = 3i/i = 3$ ✓. Producto: $(2-i)(1+i) = 2+2i-i+1 = 3+i$, y $C/A = \frac{-1+3i}{i} = \frac{(-1+3i)(-i)}{1} = i+3 = 3+i$ ✓. El coeficiente principal complejo no ha añadido ninguna dificultad conceptual: solo una división más.

Pausa para pensar

Con este teorema y el fundamental del álgebra, la situación queda perfectamente ordenada: *toda* ecuación de segundo grado con coeficientes complejos se resuelve con una fórmula cerrada, igual que en $\mathbb{R}$, y sin excepciones ni casos. Lo mismo ocurre con el tercer y el cuarto grado, mediante las fórmulas de Cardano y de Ferrari —de las que este capítulo empezó hablando—. A partir del quinto grado ya no hay fórmula con radicales, y la razón no es que nadie la haya encontrado sino que se demuestra que no existe: es el teorema de Abel–Ruffini, y su explicación completa es la teoría de Galois del capítulo 26. Que la frontera esté justo entre el grado cuatro y el cinco es uno de los hechos más sorprendentes de toda la matemática.

4.18. Las propiedades del módulo y del conjugado, demostradas, y la desigualdad triangular

La sección de aritmética recogió en una regla las propiedades $\overline{z+w} = \bar z + \bar w$, $\overline{zw} = \bar z\bar w$ y $|zw| = |z||w|$, y mencionó la desigualdad triangular en una caja de aviso. Ninguna de las cuatro se demostró. Se demuestran aquí, porque son las que convierten a $\mathbb{C}$ en un espacio donde se puede medir, y porque la desigualdad triangular es la propiedad de la que cuelgan después todos los criterios de convergencia de series complejas.

Proposición 4.68 (Propiedades del conjugado)**.** *Para todos $z, w \in \mathbb{C}$ se cumple*

$$\overline{z+w} = \bar z + \bar w, \qquad \overline{zw} = \bar z\,\bar w, \qquad \overline{\bar z} = z, \qquad z\bar z = |z|^2,$$

y $z \in \mathbb{R}$ si y solo si $z = \bar z$, mientras que z es imaginario puro si y solo si $z = -\bar z$.

Demostración. Escríbanse $z = a+bi$ y $w = c+di$. Para la suma, $\overline{(a+c)+(b+d)i} = (a+c)-(b+d)i = (a-bi)+(c-di)$. Para el producto, por un lado

$$\overline{zw} = \overline{(ac-bd)+(ad+bc)i} = (ac-bd)-(ad+bc)i,$$

y por otro $\bar z\bar w = (a-bi)(c-di) = ac-adi-bci+bd\,i^2 = (ac-bd)-(ad+bc)i$: los mismos dos números. La involución $\overline{\bar z} = z$ es cambiar el signo dos veces, y $z\bar z = a^2+b^2 = |z|^2$ es la diferencia de cuadrados ya vista. Finalmente $z = \bar z$ equivale a $b = -b$, es decir $b = 0$, y $z = -\bar z$ a $a = -a$, es decir $a = 0$. □

Teorema 4.69 (El módulo es multiplicativo)**.** *Para todos $z, w \in \mathbb{C}$: $|zw| = |z|\,|w|$, y si $w \neq 0$ también $\left|\frac{z}{w}\right| = \frac{|z|}{|w|}$.*

Demostración. Usando la proposición 4.68 dos veces y la conmutatividad,

$$|zw|^2 = (zw)\overline{(zw)} = z\,w\,\bar{z}\,\bar{w} = (z\bar{z})(w\bar{w}) = |z|^2|w|^2.$$

Como los módulos son reales no negativos, tomar raíces cuadradas conserva la igualdad y da $|zw| = |z||w|$. Para el cociente, aplíquese lo anterior a $z = \frac{z}{w} \cdot w$ y despéjese. Nótese que la demostración no ha necesitado escribir ni una sola coordenada: es más corta y más limpia que desarrollar $(ac-bd)^2+(ad+bc)^2$ y comprobar que coincide con $(a^2+b^2)(c^2+d^2)$ —identidad de Diofanto—, aunque esa cuenta también funciona y es la que demuestra que un producto de dos sumas de dos cuadrados es una suma de dos cuadrados. □

Definición 4.70 (La distancia en el plano complejo)**.** Se define la *distancia* entre z y w como $d(z,w) = |z-w|$. Coincide exactamente con la distancia euclídea entre los puntos correspondientes del plano de Argand, y cumple las tres propiedades de una distancia: es no negativa y solo nula si $z = w$; es simétrica, porque $|z-w| = |-(w-z)| = |w-z|$; y satisface la desigualdad triangular $d(z,u) \leq d(z,w) + d(w,u)$, que es el teorema siguiente aplicado a $z-w$ y $w-u$.

Teorema 4.71 (Desigualdad triangular)**.** *Para todos $z, w \in \mathbb{C}$,*

$$|z+w| \leq |z| + |w|,$$

con igualdad si y solo si uno de los dos es cero o el cociente z/w es un número real positivo *—es decir, si y solo si z y w tienen el mismo argumento—.*

Demostración. El punto de partida es que para cualquier $u \in \mathbb{C}$ se cumple $\operatorname{Re}(u) \leq |u|$, pues si $u = p + qi$ entonces $p \leq \sqrt{p^2+q^2}$. Ahora se desarrolla el cuadrado del miembro izquierdo:

$$|z+w|^2 = (z+w)\overline{(z+w)} = z\bar{z} + z\bar{w} + w\bar{z} + w\bar{w} = |z|^2 + |w|^2 + \left(z\bar{w} + \overline{z\bar{w}}\right),$$

donde se ha usado que $w\bar{z} = \overline{z\bar{w}}$. El paréntesis es $2\operatorname{Re}(z\bar{w})$, de modo que

$$|z+w|^2 = |z|^2 + |w|^2 + 2\operatorname{Re}(z\bar{w}) \;\leq\; |z|^2 + |w|^2 + 2|z\bar{w}| = |z|^2 + |w|^2 + 2|z||w| = \left(|z| + |w|\right)^2,$$

usando el teorema 4.69 y $|\bar{w}| = |w|$. Tomando raíces cuadradas —lícito, porque ambos lados son no negativos— se obtiene la desigualdad.

La igualdad se da exactamente cuando $\operatorname{Re}(z\bar{w}) = |z\bar{w}|$, es decir cuando $z\bar{w}$ es un número real no negativo. Si $w \neq 0$, dividir entre $|w|^2$ muestra que eso equivale a que z/w sea real y ≥ 0, o sea a que z y w apunten en la misma dirección desde el origen. □

Corolario 4.72 (Desigualdad triangular inversa)**.** *Para todos $z, w \in \mathbb{C}$: $\big||z| - |w|\big| \leq |z-w|$.*

Demostración. De $z = (z-w) + w$ y del teorema anterior sale $|z| \leq |z-w| + |w|$, es decir $|z| - |w| \leq |z-w|$. Intercambiando los papeles de z y w se obtiene $|w| - |z| \leq |w-z| = |z-w|$. Las dos desigualdades juntas son exactamente el enunciado con el valor absoluto. □

Proposición 4.73 (Identidad del paralelogramo)**.** *Para todos $z, w \in \mathbb{C}$:*

$$|z+w|^2 + |z-w|^2 = 2(|z|^2 + |w|^2).$$

Demostración. Basta sumar las dos identidades obtenidas en la demostración anterior: $|z+w|^2 = |z|^2 + |w|^2 + 2\,\mathrm{Re}(z\bar{w})$ y, cambiando w por $-w$, $|z-w|^2 = |z|^2 + |w|^2 - 2\,\mathrm{Re}(z\bar{w})$. Los términos cruzados se cancelan. □

Ejemplo 4.74 (Las cuatro desigualdades, con números)**.** Tómense $z = 3+i$ y $w = 1+2i$, con $|z| = \sqrt{10} \approx 3{,}162278$ y $|w| = \sqrt{5} \approx 2{,}236068$.

(a) Triangular. $z+w = 4+3i$, luego $|z+w| = 5$, mientras que $|z|+|w| \approx 5{,}398346$. En efecto $5 < 5{,}398346$: desigualdad estricta, coherente con que z y w no tengan el mismo argumento.

(b) Triangular inversa. $z-w = 2-i$, con $|z-w| = \sqrt{5} \approx 2{,}236068$, y $\big||z|-|w|\big| \approx |3{,}162278 - 2{,}236068| = 0{,}926210$. Se cumple $0{,}926210 \le 2{,}236068$ ✓.

(c) Paralelogramo. $|z+w|^2 + |z-w|^2 = 25+5 = 30$, y $2(|z|^2+|w|^2) = 2(10+5) = 30$ ✓. La identidad se cumple exactamente, sin margen de error, como corresponde a una igualdad y no a una desigualdad.

(d) El caso de igualdad. Tómense ahora $z = 1+2i$ y $w = 2+4i = 2z$, que tienen el mismo argumento. Entonces $z+w = 3+6i$ y $|z+w| = 3\sqrt{5} \approx 6{,}708204$, mientras que $|z|+|w| = \sqrt{5} + 2\sqrt{5} = 3\sqrt{5}$: *iguales*, exactamente como predice el teorema 4.71. Geométricamente, el triángulo se ha degenerado en un segmento.

Teorema 4.75 (Suma y producto de las raíces n-ésimas de la unidad)**.** *Sea $n \ge 2$ y sean $1, \zeta, \zeta^2, \dots, \zeta^{n-1}$ las n raíces n-ésimas de la unidad, con $\zeta = \mathrm{cis}(2\pi/n)$. Entonces*

$$\sum_{k=0}^{n-1} \zeta^k = 0, \qquad \prod_{k=0}^{n-1} \zeta^k = (-1)^{n+1}.$$

Demostración. La primera es la suma de una progresión geométrica de razón $\zeta \neq 1$:

$$\sum_{k=0}^{n-1} \zeta^k = \frac{\zeta^n - 1}{\zeta - 1} = \frac{1-1}{\zeta - 1} = 0,$$

porque $\zeta^n = 1$ por construcción y el denominador no se anula al ser $n \ge 2$. (Segunda demostración, sin fórmulas: las n raíces son los vértices de un polígono regular centrado en el origen, y la suma de los vectores que van del centro a los vértices de una figura con simetría de rotación de orden n debe ser invariante al girar $2\pi/n$; el único vector con esa propiedad es el nulo.)

Para el producto, obsérvese que las n raíces son exactamente las raíces del polinomio $z^n - 1$, cuyo término independiente es -1 y cuyo coeficiente principal es 1; por las relaciones de Cardano–Vieta, el producto de las raíces es $(-1)^n \cdot \frac{-1}{1} = (-1)^{n+1}$. Con $n = 5$ da 1 y con $n = 6$ da -1, que es lo que se comprobó a mano en la sección de las raíces. □

Cuidado

Dos avisos sobre lo anterior. Primero: la desigualdad triangular se extiende a cualquier número finito de sumandos por inducción, $|z_1 + \cdots + z_m| \le |z_1| + \cdots + |z_m|$, y esa es la versión que se usará al estudiar series. Segundo: *no* existe ninguna desigualdad análoga con $|z|$ sustituido por Re o por arg. Es cierto que $\operatorname{Re}(z+w) = \operatorname{Re} z + \operatorname{Re} w$ —eso es una igualdad, no una desigualdad—, pero $\arg(z+w)$ no guarda ninguna relación sencilla con $\arg z$ y $\arg w$: el argumento se lleva bien con el producto y fatal con la suma, mientras que el módulo hace justamente lo contrario. Esa dualidad entre las dos operaciones es la que obliga a cambiar de forma —binómica o polar— según lo que se vaya a hacer.

Pausa para pensar

Todo lo demostrado en esta sección se apoya en un solo hecho, $z\bar{z} = |z|^2$, y en que $\operatorname{Re}(u) \le |u|$. Nada más. Ahora bien, si se relee la demostración de la desigualdad triangular sustituyendo $z\bar{w}$ por un "producto escalar" abstracto, se obtiene palabra por palabra la demostración de la desigualdad triangular en cualquier espacio con producto escalar, real o complejo: la de $\mathbb{R}^n$ del capítulo 7, la de los espacios de funciones del capítulo 32. El paso $\operatorname{Re}(z\bar{w}) \le |z||w|$ es, en ese contexto general, la desigualdad de Cauchy–Schwarz. Los complejos son el caso más pequeño donde todo eso ya ocurre, y por eso son el mejor sitio para verlo por primera vez.

4.19. Ejercicios resueltos

Ejercicio 4.1. Halla x, y reales tales que $(x+yi)(2-i) = 5+5i$.

Solución. Desarrollando: $2x - xi + 2yi + y = (2x+y) + (2y-x)i$. Igualando parte a parte: $2x + y = 5$ y $-x + 2y = 5$. Resolviendo el sistema (el álgebra elemental): $x = 1$, $y = 3$. Comprobación: $(1+3i)(2-i) = 2 - i + 6i + 3 = 5 + 5i$ ✓. ■

Ejercicio 4.2. Calcula $\dfrac{3+4i}{1-2i}$ y su módulo.

Solución. $\dfrac{(3+4i)(1+2i)}{1+4} = \dfrac{3+6i+4i-8}{5} = \dfrac{-5+10i}{5} = -1+2i$. Su módulo es $\sqrt{5}$, que coincide con $|3+4i|/|1-2i| = 5/\sqrt{5} = \sqrt{5}$ ✓. ■

Ejercicio 4.3. Resuelve $z^2 + z + 1 = 0$ e interpreta el resultado geométricamente.

Solución. $\Delta = 1 - 4 = -3$, luego $z = \dfrac{-1 \pm i\sqrt{3}}{2}$. Ambas tienen módulo $\sqrt{\frac{1}{4} + \frac{3}{4}} = 1$ y argumentos $\pm 120°$: junto con $z = 1$ forman las tres raíces cúbicas de la unidad. Coherente, porque $z^3 - 1 = (z-1)(z^2+z+1)$. ■

Ejercicio 4.4. Usando De Moivre, expresa $\cos 3\theta$ en función de $\cos\theta$.

Solución. $(\cos\theta + i\sin\theta)^3 = \cos 3\theta + i\sin 3\theta$. Desarrollando el cubo del binomio, la parte real es $\cos^3\theta - 3\cos\theta\sin^2\theta$. Sustituyendo $\sin^2\theta = 1 - \cos^2\theta$:

$$\cos 3\theta = 4\cos^3\theta - 3\cos\theta.$$

Comprobación en $\theta = 0$: $4 - 3 = 1 = \cos 0$ ✓. Esta identidad, imposible de adivinar con trigonometría sola, sale en tres líneas con complejos; y es la que Cardano necesitaba para el caso irreducible de la cúbica. ■

Ejercicio 4.5. Demuestra que si $|z| = 1$ entonces $\dfrac{1}{z} = \bar{z}$.

Solución. De $z\bar{z} = |z|^2 = 1$ se sigue directamente $\bar{z} = 1/z$. Geométricamente: en la circunferencia unidad, invertir es reflejar. ■

Ejercicio 4.6. Halla todas las soluciones de $z^4 = -16$ y dibuja su posición.

Solución. Ya calculadas arriba: $\pm\sqrt{2} \pm i\sqrt{2}$, los cuatro vértices de un cuadrado inscrito en la circunferencia de radio 2, con los ejes como diagonales del cuadrado girado 45°. ■

4.20. Ejercicios propuestos

4.1. Calcula $(2 - 5i) + (3 + 8i)$ y $(2 - 5i)(3 + 8i)$.

4.2. Simplifica i^{2027}.

4.3. Escribe $\dfrac{4 - i}{2 + 3i}$ en forma binómica.

4.4. Halla el módulo y el argumento principal de $-3 + 3i$.

4.5. Expresa $-2 - 2i\sqrt{3}$ en forma polar.

4.6. Calcula $(\sqrt{3} + i)^6$ usando De Moivre.

4.7. Resuelve $z^2 - 6z + 34 = 0$.

4.8. Halla las cuatro raíces cuartas de 1 y comprueba que suman 0.

4.9. Halla x, y reales con $(x + yi)(3 - 2i) = 12 + 5i$.

4.10. Demuestra que $z + \bar{z}$ y $z\bar{z}$ son siempre reales.

4.11. ¿Qué conjunto del plano describe $|z - i| = |z + i|$?

4.12. Sabiendo que $2 + i$ es raíz de $x^3 - 6x^2 + 13x - 10$, halla las otras dos.

Respuestas

1. $5 + 3i$; y $6 + 16i - 15i + 40 = 46 + i$. **2.** $2027 = 4 \cdot 506 + 3$, luego $i^{2027} = -i$. **3.** $\frac{(4-i)(2-3i)}{13} = \frac{8-12i-2i-3}{13} = \frac{5-14i}{13}$. **4.** $r = 3\sqrt{2}$, $\theta = 135° = 3\pi/4$ (segundo cuadrante). **5.** $r = 4$, $\theta = 240°$: $4\,\mathrm{cis}\,240°$. **6.** $r = 2$, $\theta = 30°$: $2^6\,\mathrm{cis}\,180° = -64$. **7.** $\Delta = 36 - 136 = -100$, $z = 3 \pm 5i$. **8.** $1, i, -1, -i$; su suma es 0 (y en general las n raíces n-ésimas de la unidad suman 0 para $n \geq 2$, por simetría o por la fórmula de la suma geométrica de las sucesiones y series elementales). **9.** $(x + yi)(3 - 2i) = (3x + 2y) + (3y - 2x)i$; de $3x + 2y = 12$ y $3y - 2x = 5$ sale $x = 2$, $y = 3$. **10.** $z + \bar{z} = 2a$ y $z\bar{z} = a^2 + b^2$, ambos reales. **11.** Los puntos equidistantes de i y $-i$: el eje real. **12.** Por el teorema de raíces conjugadas, $2 - i$ también es raíz; su producto es 5, y como el término independiente es -10 con signo $(-1)^3$, la tercera raíz es 2. En efecto $x^3 - 6x^2 + 13x - 10 = (x - 2)(x^2 - 4x + 5)$.

Capítulo 5

Combinatoria y probabilidad

Lo que necesitas traer

Del capítulo 1: conjuntos, funciones inyectivas y biyectivas, e inducción. Del capítulo 3: factoriales y divisibilidad elemental. La probabilidad que aparece aquí es la versión finita de la regla de Laplace; la teoría completa espera en el capítulo 16.

Contar parece la más elemental de las actividades matemáticas: es lo primero que aprende un niño. Pero contar *sin listar* —saber cuántas cosas hay sin escribirlas una por una— es un arte con reglas propias, y es la puerta de entrada a la probabilidad. Este capítulo es autocontenido: no necesita cálculo ni geometría analítica, solo el álgebra del álgebra elemental. Lo ponemos aquí, justo antes de los límites, porque es el último tema realmente escolar que queda pendiente y porque el coeficiente binomial que aparece en él reaparecerá sin avisar en el binomio de Newton, en las series de potencias del capítulo 6, en el conteo de subgrupos del capítulo 23 y hasta en las componentes independientes del tensor de Riemann del capítulo 36.

Idea clave

Intuición primero. Toda la combinatoria elemental sale de una sola pregunta hecha con cuidado: *¿el orden importa?* y *¿se puede repetir?* Las cuatro respuestas posibles dan las cuatro fórmulas que existen, y no hay una quinta. El error clásico del estudiante no es equivocarse en la aritmética, sino elegir mal la fórmula porque no se detuvo a contestar esas dos preguntas. Si aprendes a contestarlas antes de escribir nada, este capítulo se vuelve trivial; si no, se vuelve un catálogo de fórmulas indistinguibles.
La probabilidad, por su parte, no es una teoría nueva: en el caso finito y equiprobable es *exactamente* combinatoria dividida entre combinatoria. "Probabilidad de sacar par de ases" significa "cuente cuántas manos tienen par de ases, cuente cuántas manos hay, divida". Por eso los dos temas van en el mismo capítulo.

5.1. Los dos principios fundamentales

Todo lo demás se deduce de estos dos. Merecen leerse despacio.

Regla: principio multiplicativo (o de la multiplicación)

Si una tarea se realiza en dos etapas sucesivas, la primera de m maneras y la segunda de n maneras *para cada* elección de la primera, entonces la tarea completa se realiza de $m \cdot n$ maneras.
Se extiende a cualquier número de etapas: $m_1 \cdot m_2 \cdots m_k$.

Regla: principio aditivo (o de la suma)

Si una tarea se realiza *o bien* de una de m maneras *o bien* de una de n maneras, y ninguna manera pertenece a los dos grupos a la vez, entonces hay $m + n$ maneras en total.

Idea clave

La palabra que decide cuál usar es la conjunción: "**y**" (primero esto *y* luego aquello) se traduce por *multiplicar*; "**o**" (esto *o* aquello, excluyentes) se traduce por *sumar*. Es literalmente una traducción del castellano a la aritmética, y es fiable.

Ejemplo 5.1. Un restaurante ofrece 3 entradas, 5 platos fuertes y 2 postres. ¿Cuántos menús completos hay?

Paso 1. Es "entrada *y* plato *y* postre": multiplicamos.

Paso 2. $3 \cdot 5 \cdot 2 = 30$ menús.

Ejemplo 5.2. ¿Cuántas placas de auto hay con el formato "tres letras seguidas de tres dígitos", si se usan las 26 letras del alfabeto inglés y los 10 dígitos, permitiendo repeticiones?

Seis etapas independientes: $26 \cdot 26 \cdot 26 \cdot 10 \cdot 10 \cdot 10 = 17\,576 \cdot 1000 = 17\,576\,000$.

Ejemplo 5.3. De una ciudad A a una ciudad B hay 3 vuelos y 4 buses. ¿De cuántas maneras se puede viajar de A a B?

Es "vuelo *o* bus", y ningún viaje es las dos cosas: $3 + 4 = 7$.

Cuidado

El principio multiplicativo exige que el número de opciones de la segunda etapa *no dependa* de cuál se eligió en la primera. Si un restaurante tiene 3 entradas y 5 platos, pero con la entrada de mariscos solo se permiten 2 platos, la cuenta ya no es $3 \cdot 5$: hay que partirla en casos y sumar ($2 \cdot 5 + 1 \cdot 2 = 12$). Esto no es una excepción a la regla, es la regla leída con cuidado.

5.2. El mapa del conteo: permutaciones, variaciones y combinaciones, y cómo saber cuál pide el problema

Los dos principios de la sección anterior bastan, en teoría, para contar cualquier cosa. En la práctica casi todos los problemas de conteo elemental caen en uno de cinco moldes que se repiten una y otra vez, y el trabajo del que resuelve no es inventar nada, sino *reconocer el molde*. Esta sección hace tres cosas. Primera, dar nombre y definición precisa a los cinco moldes, porque la mitad de los errores vienen de llamar "combinación" a lo que es una variación. Segunda, dar un procedimiento de cuatro preguntas que decide el molde sin margen de duda, aplicado a seis problemas distintos resueltos uno por uno. Y tercera, cerrar con dos herramientas que aparecen en cuanto los objetos se reparten en cajas o el problema se define en función de sí mismo: el modelo de bolas y urnas y las recurrencias lineales.

Las secciones siguientes deducen cada fórmula con su justificación; aquí sólo se nombran, se distinguen y se aprende a elegirlas. Léase esta sección como el mapa y las siguientes como el viaje.

Definición 5.4 (Ordenaciones y selecciones)**.** Sea X un conjunto de n elementos distintos y sea k un entero con $0 \leq k$.

(I) Una **permutación** de X es una ordenación de *todos* sus elementos en fila; equivalentemente, una aplicación biyectiva de X en sí mismo. Su número es $n!$, con el factorial que se define en la sección siguiente.

(II) Una **variación** (u ordenación parcial) **sin repetición** de n elementos tomados de k en k es una lista ordenada de k elementos *distintos* de X. Su número es $V(n,k) = P(n,k) = \dfrac{n!}{(n-k)!}$, y exige $k \leq n$.

(III) Una **variación con repetición** es una lista ordenada de k elementos de X en la que se permite repetir. Su número es n^k, y aquí k puede superar a n.

(IV) Una **combinación sin repetición** de n elementos tomados de k en k es un *subconjunto* de k elementos de X: importa quiénes son, no en qué orden se escriben. Su número es $\binom{n}{k}$.

(V) Una **combinación con repetición** es un *multiconjunto* de tamaño k formado con elementos de X, es decir, una elección de k objetos en la que se puede repetir tipo y no importa el orden. Su número es $\binom{n+k-1}{k}$.

(VI) Una **permutación con repetición** es una ordenación de n objetos entre los que hay grupos de objetos indistinguibles, de tamaños $n_1, \ldots, n_r$ con $n_1 + \cdots + n_r = n$. Su número es $\dfrac{n!}{n_1!\, n_2! \cdots n_r!}$.

Observación 5.5. Las cuatro primeras fórmulas son exactamente las cuatro casillas de la tabla que aparece más adelante en el capítulo, bajo el título de los cuatro casilleros; allí se demuestra la única que no es inmediata, la de las combinaciones con repetición, mediante la biyección de

las barras y las estrellas. Las secciones inmediatamente siguientes deducen las otras tres y la de permutaciones con repetición. Aquí interesa sólo distinguirlas.

Sobre la notación: la tradición escolar española escribe $P_n = n!$ para las permutaciones, $V_{n,k}$ para las variaciones sin repetición, $VR_{n,k} = n^k$ para las variaciones con repetición, $C_{n,k} = \binom{n}{k}$ para las combinaciones y $CR_{n,k}$ para las combinaciones con repetición. La tradición anglosajona no tiene la palabra "variación" y llama a $V(n,k)$ *k-permutación*, escribiéndola $P(n,k)$. Este libro usa $P(n,k)$ y $\binom{n}{k}$, que son las que menos se prestan a confusión, pero conviene reconocer las otras porque aparecen en cualquier libro de texto de bachillerato.

Regla: las cuatro preguntas para elegir la fórmula

Antes de escribir un solo número, contesta por escrito y *en este orden*:

1. **¿Cuál es el conjunto base y cuántos se toman?** Escribe n (de cuántos tipos distintos dispones) y k (cuántos coges). Casi todos los errores empiezan aquí, confundiendo el papel de los dos números.
2. **¿Importa el orden?** La prueba infalible: intercambia dos de los elegidos y pregúntate si el resultado es *otro* caso distinto o el *mismo*. Si es otro, importa el orden y estás en una variación o una permutación; si es el mismo, no importa y estás en una combinación. "Oro, plata y bronce" sí; "comité de tres" no.
3. **¿Se pueden repetir?** ¿Puede salir dos veces el mismo elemento? Depende de si el objeto "se gasta" al usarlo. Sacar bolas de una urna *sin* reposición no repite; con reposición, sí. Las cifras de un PIN repiten; los corredores de un podio no.
4. **¿Se usan todos?** Si $k = n$ y hay orden, la variación se convierte en una permutación y la fórmula se simplifica a $n!$. Si además hay elementos indistinguibles entre sí, es una permutación con repetición y hay que dividir por los factoriales de las multiplicidades.
5. **Escribe la fórmula del casillero** al que has llegado:

	importa el orden	no importa el orden
sin repetir	$\dfrac{n!}{(n-k)!}$	$\dbinom{n}{k}$
con repetición	n^k	$\dbinom{n+k-1}{k}$

6. **Comprueba con un caso diminuto.** Baja los datos a $n = 3$, $k = 2$ y enumera a mano las nueve, seis, tres o seis posibilidades según el casillero. Si tu fórmula no reproduce lo que has enumerado, has elegido mal el casillero, no te has equivocado en la aritmética.
7. **Comprueba el orden de magnitud.** Con orden siempre salen más casos que sin orden, y exactamente $k!$ veces más cuando no hay repetición. Si tu respuesta "sin orden" es mayor que la "con orden", hay un error seguro.

Ejemplo 5.6 (Seis problemas resueltos con las cuatro preguntas)**.** (a) *El podio.* En una carrera de 8 corredores se reparten oro, plata y bronce. ¿De cuántas maneras?

P1: el conjunto base son los 8 corredores, $n = 8$; se toman $k = 3$. **P2:** intercambiar el oro y la plata da otro podio, luego *sí* importa el orden. **P3:** nadie gana dos medallas, luego *no* hay repetición. **P4:** $k = 3 < 8 = n$, no se usan todos. Casillero: variación sin repetición.

$$P(8,3) = \frac{8!}{5!} = 8 \cdot 7 \cdot 6 = 336.$$

(b) *El comité.* De esas mismas 8 personas se elige un comité de 3 sin cargos.

P1: $n = 8$, $k = 3$, igual que antes. **P2:** el comité formado por Ana, Beto y Carla es el mismo se escriba en el orden que se escriba: *no* importa el orden. **P3:** nadie está dos veces: sin repetición. Casillero: combinación sin repetición.

$$\binom{8}{3} = \frac{8 \cdot 7 \cdot 6}{3 \cdot 2 \cdot 1} = 56.$$

Y la comprobación del paso 7 cuadra a la perfección: $336 = 56 \cdot 3! = 56 \cdot 6$. Los mismos ocho y los mismos tres, y seis veces menos casos, porque cada comité da lugar a 6 podios distintos.

(c) *El PIN.* ¿Cuántos códigos de 4 cifras hay, del 0000 al 9999?

P1: $n = 10$ dígitos, $k = 4$ posiciones. **P2:** 1234 y 4321 son códigos distintos: *sí* importa el orden. **P3:** 1131 es un código perfectamente válido: *sí* hay repetición. Casillero: variación con repetición.

$$10^4 = 10\,000.$$

Obsérvese que aquí k podría ser mayor que n sin ningún problema: hay 10^6 códigos de seis cifras.

(d) *La fila.* ¿De cuántas maneras se sientan 6 personas en una fila de 6 sillas?

P1: $n = 6$, y se colocan las 6. **P2:** sí importa el orden, es lo único que hay. **P3:** nadie se sienta dos veces. **P4:** $k = n$, se usan todos, y todos son distinguibles. Casillero: permutación.

$$6! = 720.$$

(e) *Los anagramas.* ¿Cuántas palabras distintas, con o sin sentido, se forman con las letras de PATATA?

P1: son 6 letras y se usan las 6. **P2:** sí importa el orden. **P4:** se usan todas, pero *no* son distinguibles entre sí: hay tres A, dos T y una P. Casillero: permutación con repetición.

$$\frac{6!}{3!\,2!\,1!} = \frac{720}{6 \cdot 2 \cdot 1} = 60.$$

La comprobación mental: si las tres A llevaran subíndices y las dos T también, habría 720 palabras; al borrar los subíndices, cada palabra real se ha contado $3! \cdot 2! = 12$ veces, y $720/12 = 60$.

(f) *Los pasteles.* Una pastelería tiene 5 sabores y se compran 4 pasteles, pudiendo repetir sabor. ¿Cuántas compras distintas hay?

P1: $n = 5$ sabores, $k = 4$ pasteles. **P2:** la caja con dos de chocolate y dos de fresa es la misma se metan en el orden que se metan: *no* importa el orden. **P3:** se puede repetir sabor: *sí* hay repetición. Casillero: combinación con repetición.

$$\binom{5+4-1}{4} = \binom{8}{4} = 70.$$

Este es el casillero que casi nadie reconoce a la primera; la señal que lo delata es la pareja "no importa el orden" y "se puede repetir", que en el mundo real suele venir disfrazada de reparto de objetos idénticos.

La tabla resumida de los seis, para ver de un vistazo cómo cada respuesta empuja hacia un casillero distinto:

problema	n	k	orden	repite	resultado
podio	8	3	sí	no	336
comité	8	3	no	no	56
PIN	10	4	sí	sí	10 000
fila	6	6	sí	no	720
anagramas	6	6	sí	indistinguibles	60
pasteles	5	4	no	sí	70

Los seis han sido comprobados enumerando exhaustivamente los casos con un ordenador, y los seis números coinciden con la fórmula.

Cuidado

La pregunta 2 no se contesta mirando el enunciado, sino *el objeto que se cuenta*. "Elegir tres de ocho" no dice nada por sí solo: si lo elegido es un comité, no hay orden; si es un podio, lo hay; si es una contraseña, lo hay y además puede repetir. Cuando el enunciado sea ambiguo, decide tú y escribe explícitamente qué has supuesto: en un examen eso vale casi tantos puntos como el número.
Y un segundo aviso sobre la pregunta 3: "con repetición" se refiere a repetir *tipos*, no objetos. Al comprar cuatro pasteles de cinco sabores no hay cinco objetos de los que se tomen cuatro; hay cinco *tipos* inagotables. Confundir ambas cosas lleva a escribir $\binom{5}{4}$, que además da un número menor que el correcto y debería hacer saltar la alarma del paso 7.

Regla: repartir objetos en cajas: el modelo de bolas y urnas

Un problema de reparto se traduce siempre al mismo modelo: hay k *bolas* que se distribuyen en n *urnas*. Lo único que hay que decidir es si las bolas se distinguen entre sí y si las urnas se distinguen entre sí, y eso da cuatro casos y nada más.

1. **Traduce el enunciado.** Objetos que se reparten ⇒ bolas; destinatarios o grupos

$\Rightarrow$ urnas. Pregúntate si dos objetos son intercambiables sin que cambie el reparto (bolas indistinguibles) y si dos destinatarios lo son (urnas indistinguibles). Personas con nombre y niños con nombre se distinguen; caramelos idénticos y "grupos" sin nombre, no.

2. **Bolas distinguibles y urnas distinguibles:** cada bola elige urna independientemente, así que hay n^k repartos. Es una variación con repetición disfrazada.
3. **Bolas indistinguibles y urnas distinguibles:** sólo importa cuántas caen en cada urna, es decir, las soluciones enteras no negativas de $x_1 + \cdots + x_n = k$. Hay $\binom{n+k-1}{k}$, por barras y estrellas. Es una combinación con repetición disfrazada.
4. **Bolas distinguibles y urnas indistinguibles:** lo que se cuenta es en qué grupos quedan partidas las bolas, sin nombrar los grupos. El número de particiones de un conjunto de k elementos en exactamente j bloques no vacíos se llama *número de Stirling de segunda especie* $S(k, j)$, y la respuesta es $\sum_{j=1}^{n} S(k, j)$.
5. **Bolas indistinguibles y urnas indistinguibles:** sólo queda el reparto de tamaños, sin decir de quién: es el número de maneras de escribir k como suma de a lo sumo n enteros positivos, o sea, el número de *particiones del entero* k en a lo sumo n partes. No hay fórmula cerrada elemental.
6. **Si se exige que ninguna urna quede vacía**, los cuatro números cambian a: las aplicaciones sobreyectivas $n!\,S(k, n)$; la cuenta $\binom{k-1}{n-1}$, que sale de regalar una bola a cada urna y repartir el resto libremente; el propio $S(k, n)$; y las particiones de k en exactamente n partes.
7. **Verifica siempre con un caso pequeño.** Con $k = 2$ y $n = 2$ los cuatro números son 4, 3, 2 y 2: enumerarlos a mano lleva medio minuto y desmonta cualquier duda sobre qué caso es el tuyo.

Ejemplo 5.7 (Cuatro bolas en tres urnas, los cuatro casos enumerados enteros)**.** Se toman $k = 4$ y $n = 3$ y se calculan los cuatro números; todos han sido comprobados generando uno por uno los repartos con el ordenador.

(a) *Bolas distinguibles, urnas distinguibles.* Cuatro regalos distintos para tres niños con nombre: cada regalo elige niño, luego

$$3^4 = 81.$$

Sin urnas vacías, es decir, con los tres niños recibiendo algo: $3!\,S(4, 3) = 6 \cdot 6 = 36$. También se obtiene por inclusión–exclusión, $3^4 - 3 \cdot 2^4 + 3 \cdot 1^4 = 81 - 48 + 3 = 36$.

(b) *Bolas indistinguibles, urnas distinguibles.* Cuatro caramelos idénticos para tres niños con nombre: son las soluciones de $x_1 + x_2 + x_3 = 4$,

$$\binom{3+4-1}{4} = \binom{6}{4} = 15.$$

Y son fáciles de listar: $(4,0,0)$ y sus 3 permutaciones, $(3,1,0)$ con 6, $(2,2,0)$ con 3, $(2,1,1)$ con 3; total $3+6+3+3=15$ ✓. Con todos recibiendo al menos uno: $\binom{4-1}{3-1}=\binom{3}{2}=3$, que son $(2,1,1)$, $(1,2,1)$ y $(1,1,2)$.

(c) *Bolas distinguibles, urnas indistinguibles.* Cuatro personas repartidas en hasta tres grupos sin nombre:

$$S(4,1)+S(4,2)+S(4,3)=1+7+6=14.$$

El 1 es dejarlas todas juntas; los 7 son las maneras de partirlas en dos grupos —cuatro del tipo "uno contra tres" y tres del tipo "dos contra dos"—; los 6 son las de partirlas en tres grupos, que necesariamente son "dos juntas y dos solas", y hay $\binom{4}{2}=6$ maneras de elegir la pareja. Nótese que el total de particiones de un conjunto de cuatro elementos es 15 —el número de Bell—, y sale 14 porque queda excluida la partición en cuatro grupos, que no cabe en tres urnas.

(d) *Bolas indistinguibles, urnas indistinguibles.* Cuatro caramelos idénticos en tres montones sin nombre: sólo cuentan los tamaños,

$$4; \qquad 3+1; \qquad 2+2; \qquad 2+1+1,$$

es decir, 4 repartos. La partición $1+1+1+1$ queda fuera porque necesitaría cuatro montones. Con los tres montones no vacíos queda una sola posibilidad, $2+1+1$.

Los cuatro números para el mismo enunciado —81, 15, 14 y 4— difieren en un factor veinte, y lo único que los separa es qué se considera distinguible. Ahí está la moraleja: en un problema de repartos, decidir qué es distinguible *es* el problema; la fórmula viene después y es la parte fácil.

Definición 5.8 (Recurrencia lineal homogénea con coeficientes constantes)**.** Una sucesión (a_n) satisface una *recurrencia lineal homogénea de orden r con coeficientes constantes* si existen números $c_1,\dots,c_r$, con $c_r\neq 0$, tales que

$$a_n = c_1a_{n-1}+c_2a_{n-2}+\cdots+c_ra_{n-r} \qquad \text{para todo } n\geq r.$$

Se llama *lineal* porque cada término anterior aparece elevado a la primera potencia y sin productos entre ellos, y *homogénea* porque no hay ningún sumando independiente. Los r primeros valores $a_0,\dots,a_{r-1}$ son las *condiciones iniciales*, y determinan la sucesión por completo. Su *polinomio característico* es

$$x^r - c_1x^{r-1} - c_2x^{r-2} - \cdots - c_r.$$

Idea clave

Una recurrencia aparece en cuanto un problema de conteo se puede resolver "mirando el último paso". Si para contar los objetos de tamaño n basta con clasificarlos según cómo terminan, y cada clase se corresponde con objetos de tamaño menor, los dos principios de la suma y del producto producen automáticamente una recurrencia. No es una técnica nueva: es el principio aditivo aplicado a una partición inteligente de los casos.

Regla: resolver una recurrencia lineal con la ecuación característica

1. **Escribe la recurrencia en forma estándar**, con todos los términos a un lado: $a_n - c_1 a_{n-1} - \cdots - c_r a_{n-r} = 0$. Anota aparte las condiciones iniciales.
2. **Escribe la ecuación característica** sustituyendo a_{n-j} por x^{r-j}; para orden 2 queda $x^2 - c_1 x - c_2 = 0$.
3. **Halla sus raíces.** Si son reales y distintas, $\lambda_1 \neq \lambda_2$, la solución general es $a_n = A\lambda_1^n + B\lambda_2^n$.
4. **Si hay una raíz doble** λ, la solución general es $a_n = (A + Bn)\lambda^n$: el factor n es imprescindible, porque con dos constantes hacen falta dos soluciones independientes y λ^n sólo da una.
5. **Impón las condiciones iniciales** sustituyendo $n = 0$ y $n = 1$; sale un sistema lineal de dos ecuaciones con incógnitas A y B.
6. **Despeja A y B** y escribe la fórmula cerrada.
7. **Verifica con el término siguiente:** calcula a_2 con la recurrencia y con la fórmula cerrada, y comprueba que coinciden. Si no, el error está casi siempre en el paso 5.

Ejemplo 5.9 (Tres recurrencias resueltas hasta el final)**.** (a) *El pasillo de dominós.* ¿De cuántas maneras a_n se puede cubrir un rectángulo de $2 \times n$ con fichas de dominó de 2×1, colocadas en horizontal o en vertical?

Planteamiento. Mírese la columna de la izquierda. O bien está cubierta por una ficha vertical, y el resto es un rectángulo $2 \times (n-1)$; o bien está cubierta por dos fichas horizontales, que ocupan también la segunda columna, y el resto es un $2 \times (n-2)$. Son casos excluyentes y agotan todas las posibilidades, luego por el principio aditivo

$$a_n = a_{n-1} + a_{n-2}, \qquad a_1 = 1, \quad a_2 = 2.$$

Cálculo. Los primeros valores son

$$1,\ 2,\ 3,\ 5,\ 8,\ 13,\ 21,\ 34,\ 55,\ 89,$$

de modo que un pasillo de 2×10 se cubre de 89 maneras. Se ha comprobado generando todos los recubrimientos uno por uno hasta $n = 8$, y coinciden.

Lectura. Es la sucesión de Fibonacci desplazada: $a_n = F_{n+1}$. Su fórmula cerrada —la de Binet— sale de la ecuación característica $x^2 = x + 1$, cuyas raíces son $\frac{1 \pm \sqrt{5}}{2}$; la sección sobre funciones generatrices, más adelante en el capítulo, la obtiene por el otro camino, y merece la pena comparar los dos.

(b) *Raíces distintas.* Resuélvase $a_n = 5a_{n-1} - 6a_{n-2}$ con $a_0 = 1$ y $a_1 = 4$.

Paso 2. Ecuación característica: $x^2 - 5x + 6 = 0$. **Paso 3.** Raíces $x = 2$ y $x = 3$, distintas, luego $a_n = A \cdot 2^n + B \cdot 3^n$. **Paso 5.** Con $n = 0$: $A + B = 1$. Con $n = 1$: $2A + 3B = 4$. **Paso 6.** Restando el doble de la primera de la segunda, $B = 2$, y entonces $A = -1$:

$$a_n = 2 \cdot 3^n - 2^n.$$

Paso 7. Verificación: la recurrencia da $a_2 = 5\cdot 4 - 6\cdot 1 = 14$, y la fórmula da $2\cdot 9 - 4 = 14$ ✓. Los ocho primeros términos son $1, 4, 14, 46, 146, 454, 1394, 4246$, y la fórmula los reproduce uno a uno.

(c) *Raíz doble.* Resuélvase $a_n = 4a_{n-1} - 4a_{n-2}$ con $a_0 = 1$ y $a_1 = 3$.

Paso 2. Ecuación característica: $x^2 - 4x + 4 = (x-2)^2 = 0$. **Paso 4.** Raíz doble $\lambda = 2$, luego $a_n = (A + Bn)2^n$. **Paso 5.** Con $n = 0$: $A = 1$. Con $n = 1$: $(1 + B)\cdot 2 = 3$, de donde $B = \frac{1}{2}$. **Paso 6.** Por tanto

$$a_n = \left(1 + \tfrac{n}{2}\right)2^n = 2^n + n\,2^{n-1}.$$

Paso 7. La recurrencia da $a_2 = 4\cdot 3 - 4\cdot 1 = 8$ y la fórmula $4 + 2\cdot 2 = 8$ ✓; a_3 vale $4\cdot 8 - 4\cdot 3 = 20$ por un lado y $8 + 3\cdot 4 = 20$ por el otro ✓. Los primeros términos son $1, 3, 8, 20, 48, 112, 256, 576$. Si se hubiera olvidado el factor n del paso 4 y se hubiera escrito $a_n = A\cdot 2^n$, el sistema del paso 5 sería incompatible —$A = 1$ y $2A = 3$— y ese es justamente el síntoma de que la raíz era doble.

(d) *La misma recurrencia, otro problema.* ¿Cuántas cadenas de n ceros y unos no contienen dos unos consecutivos? Clasificando por el último símbolo: si acaba en 0, lo anterior es cualquier cadena válida de longitud $n-1$; si acaba en 1, el símbolo anterior ha de ser 0 y lo de más atrás es cualquier cadena válida de longitud $n-2$. Otra vez $b_n = b_{n-1} + b_{n-2}$, ahora con $b_1 = 2$ y $b_2 = 3$, y salen $2, 3, 5, 8, 13, 21, 34, 55$: comprobado enumerando las 2^n cadenas hasta $n = 8$. Dos problemas que no se parecen en nada resultan ser el mismo, y esa es la razón de fondo por la que merece la pena resolver la recurrencia en abstracto y no el problema concreto.

Cuidado

Tres errores clásicos con recurrencias. El primero, olvidar el factor n en el caso de raíz doble; el sistema de condiciones iniciales sale incompatible y hay quien concluye que el problema está mal planteado. El segundo, contar mal las condiciones iniciales: una recurrencia de orden 2 necesita exactamente dos, y darle una sola deja infinitas soluciones mientras que darle tres suele producir una contradicción. El tercero, aplicar el método a recurrencias que no son lineales homogéneas con coeficientes constantes: $a_n = a_{n-1} + n$ tiene un término independiente y $a_n = a_{n-1}a_{n-2}$ no es lineal, y en ninguno de los dos casos sirve la ecuación característica tal cual.

Observación 5.10. La ecuación característica no es un truco aislado. Es exactamente el mismo cálculo que resuelve las ecuaciones diferenciales lineales con coeficientes constantes, donde $a_n = \lambda^n$ se sustituye por $y = \mathrm{e}^{\lambda t}$ y la raíz doble obliga igualmente a introducir un factor extra, allí $t\,\mathrm{e}^{\lambda t}$. Y es también un problema de valores propios: la recurrencia $a_n = c_1 a_{n-1} + c_2 a_{n-2}$ se escribe en forma matricial como

$$\begin{pmatrix} a_n \\ a_{n-1} \end{pmatrix} = \begin{pmatrix} c_1 & c_2 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} a_{n-1} \\ a_{n-2} \end{pmatrix},$$

y el polinomio característico de esa matriz es, salvo el signo, el de la definición. Las tres teorías —recurrencias, ecuaciones diferenciales lineales y diagonalización— son la misma idea contada tres veces, y quien entiende una ha entendido las otras dos.

Pausa para pensar

La palabra "variación" es una herencia de la tradición continental y no existe en inglés, donde todo son *permutations* y *combinations*; en cambio la distinción que de verdad importa —orden sí o no, repetición sí o no— es universal y no depende del idioma. La recurrencia del pasillo de dominós es la de los conejos que Leonardo de Pisa, llamado Fibonacci, planteó en su *Liber Abaci* de 1202 como un ejercicio menor entre cientos; ocho siglos después es probablemente la sucesión más famosa de la matemática, y aparece en el conteo de recubrimientos, en las cadenas binarias sin unos consecutivos, en el algoritmo de Euclides y en el número de hojas de una piña. Que un problema de conejos y un problema de dominós den la misma sucesión no es una casualidad pintoresca: es la señal de que ambos tienen la misma estructura recursiva, y reconocer estructuras compartidas es, al final, de lo que trata toda la combinatoria.

5.3. Permutaciones: cuando el orden importa

Definición 5.11. El *factorial* de $n \in \mathbb{N}$ es $n! = n \cdot (n-1) \cdots 2 \cdot 1$, con el convenio $0! = 1$.

Los primeros valores crecen con una violencia que conviene tener presente:

$$1! = 1, \quad 2! = 2, \quad 3! = 6, \quad 4! = 24, \quad 5! = 120, \quad 10! = 3\,628\,800, \quad 20! \approx 2.43 \times 10^{18}.$$

Pausa para pensar

¿Por qué $0! = 1$ y no 0? No es un capricho. La única manera de ordenar el conjunto vacío es no hacer nada, y "no hacer nada" es *una* manera, no ninguna. Además, el convenio es el único que hace que la identidad $n! = n \cdot (n-1)!$ siga valiendo en $n = 1$, y el único que hace que $\binom{n}{0} = 1$ salga de la fórmula general. En matemática, los convenios que se conservan son los que hacen que las fórmulas no tengan excepciones.

Regla: permutaciones de n objetos

El número de maneras de ordenar n objetos distintos en fila es

$$n!.$$

Por qué: para el primer lugar hay n candidatos; una vez ocupado, para el segundo quedan $n-1$; y así. Principio multiplicativo.

Ejemplo 5.12. ¿De cuántas maneras se sientan 5 personas en una fila de 5 sillas? $5! = 120$.

Regla: variaciones (permutaciones parciales)

El número de maneras de elegir k objetos de entre n distintos *y ordenarlos* es

$$P(n,k) = \frac{n!}{(n-k)!} = n(n-1)\cdots(n-k+1) \quad (k \text{ factores}).$$

Ejemplo 5.13. En una carrera de 10 corredores, ¿de cuántas maneras pueden repartirse oro, plata y bronce?

Importa el orden (no es lo mismo oro para Ana que bronce para Ana) y no hay repetición (nadie gana dos medallas):

$$P(10,3) = 10 \cdot 9 \cdot 8 = 720.$$

Regla: permutaciones con elementos repetidos

Si entre n objetos hay n_1 iguales entre sí de un tipo, n_2 de otro, …, n_r del último (con $n_1 + \cdots + n_r = n$), el número de ordenaciones *distinguibles* es

$$\frac{n!}{n_1!\, n_2! \cdots n_r!}.$$

Por qué: si pintáramos los repetidos de colores distintos habría $n!$ órdenes; al despintarlos, cada ordenación real se ha contado $n_1! \cdots n_r!$ veces, una por cada permutación interna de los iguales. Dividimos por lo que sobra.

Ejemplo 5.14. ¿Cuántas "palabras" distintas (con o sin sentido) se forman con las letras de MATEMATICA?

Paso 1. Contamos letras: son 10 en total, con A tres veces, M dos veces, T dos veces, y E, I, C una vez cada una.

Paso 2. Aplicamos la fórmula:

$$\frac{10!}{3!\,2!\,2!} = \frac{3\,628\,800}{6 \cdot 2 \cdot 2} = \frac{3\,628\,800}{24} = 151\,200.$$

Ejemplo 5.15. Con las letras de BANANA (6 letras: A tres veces, N dos veces, B una):

$$\frac{6!}{3!\,2!\,1!} = \frac{720}{12} = 60.$$

5.4. Combinaciones: cuando el orden no importa

Regla: combinaciones

El número de maneras de elegir k objetos de entre n distintos *sin importar el orden* es el *coeficiente binomial*

$$\binom{n}{k} = \frac{n!}{k!\,(n-k)!}.$$

Por qué: hay $P(n,k) = n!/(n-k)!$ maneras de elegirlos y ordenarlos; cada selección desordenada aparece ahí $k!$ veces (una por cada orden interno). Dividimos por $k!$.

Idea clave

La relación entre las dos fórmulas es la clave de todo el capítulo y se resume en una frase: *ordenar es elegir y luego permutar*. En símbolos, $P(n,k) = \binom{n}{k} \cdot k!$. Si te sabes una, te sabes la otra. Y si dudas de cuál usar, pregúntate si intercambiar dos de los elegidos produce un resultado distinto: si sí, es P; si no, es $\binom{n}{k}$.

Ejemplo 5.16. De un grupo de 10 personas se elige un comité de 3 (sin cargos). ¿Cuántos comités hay?

$$\binom{10}{3} = \frac{10 \cdot 9 \cdot 8}{3 \cdot 2 \cdot 1} = \frac{720}{6} = 120.$$

Compárese con el ejemplo de las medallas: mismos 10 y mismos 3, pero allí había 720 y aquí 120, exactamente $3! = 6$ veces menos, porque un comité {Ana, Beto, Carla} es uno solo, mientras que las seis órdenes de esas mismas tres personas eran seis podios distintos.

Ejemplo 5.17. ¿Cuántas manos de póquer (5 cartas de una baraja de 52) hay?

$$\binom{52}{5} = \frac{52 \cdot 51 \cdot 50 \cdot 49 \cdot 48}{120} = 2\,598\,960.$$

Este número aparecerá varias veces más adelante como denominador de probabilidades.

Ejemplo 5.18. De 7 hombres y 5 mujeres se elige un comité de 4 personas con al menos una mujer. ¿Cuántos comités hay?

Paso 1. Contar directamente obligaría a sumar los casos "una mujer", "dos", "tres", "cuatro". Es más corto contar el complemento.

Paso 2. Comités totales: $\binom{12}{4} = 495$. Comités sin ninguna mujer (solo hombres): $\binom{7}{4} = 35$.

Paso 3. Restamos: $495 - 35 = 460$.

Idea clave

La técnica del ejemplo anterior —*contar el complemento*— vale su peso en oro. Siempre que el enunciado diga "al menos uno", el camino corto casi siempre es contar el total y restar el caso "ninguno". La razón es estructural: "al menos uno" se descompone en muchos casos, pero su negación, "ninguno", es un caso solo.

Regla: propiedades del coeficiente binomial

1. *Simetría:* $\binom{n}{k} = \binom{n}{n-k}$. Elegir a los k que entran es lo mismo que elegir a los $n-k$ que se quedan fuera.

2. *Regla de Pascal:* $\binom{n}{k} = \binom{n-1}{k-1} + \binom{n-1}{k}$. Fija un elemento: los grupos que lo contienen y los que no.
3. *Suma total:* $\sum_{k=0}^{n} \binom{n}{k} = 2^n$. Es el número de subconjuntos de un conjunto de n elementos, clasificados por tamaño.
4. *Bordes:* $\binom{n}{0} = \binom{n}{n} = 1$ y $\binom{n}{1} = n$.

5.5. El binomio de Newton y el triángulo de Pascal

Teorema 5.19 (binomio de Newton)**.** *Para todo $n \in \mathbb{N}$,*

$$(a+b)^n = \sum_{k=0}^{n} \binom{n}{k} a^{n-k} b^k.$$

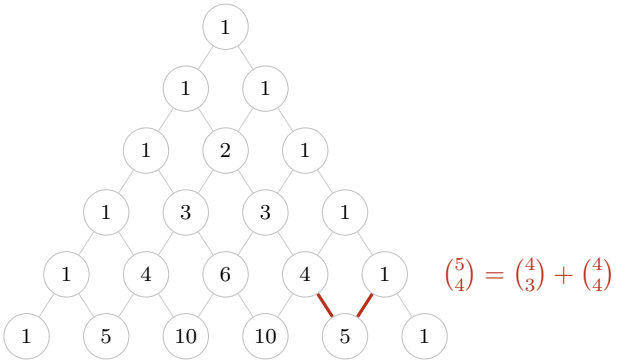


Figura 5.1: El triángulo de Pascal. Cada número es la suma de los dos que tiene encima, y esa relación no es un truco aritmético: contar los subconjuntos de k elementos separando los que contienen al último elemento de los que no lo contienen da exactamente esa identidad.

Idea clave

Por qué es cierto, sin inducción. Escribe $(a+b)^n$ como el producto de n paréntesis idénticos. Al desarrollarlo, cada término del resultado se obtiene eligiendo, en cada paréntesis, o bien la a o bien la b, y multiplicando. Un término con k letras b (y por tanto $n-k$ letras a) aparece tantas veces como maneras hay de decidir *en cuáles k paréntesis* elegimos la b: es decir, $\binom{n}{k}$ veces. El coeficiente binomial no se "mete" en la fórmula: *es* el conteo.

Los coeficientes forman el triángulo de Pascal, donde cada número es la suma de los dos de

arriba (que es exactamente la regla de Pascal):

$$\begin{array}{ccccccccccc} & & & & & 1 & & & & & \\ & & & & 1 & & 1 & & & & \\ & & & 1 & & 2 & & 1 & & & \\ & & 1 & & 3 & & 3 & & 1 & & \\ & 1 & & 4 & & 6 & & 4 & & 1 & \\ 1 & & 5 & & 10 & & 10 & & 5 & & 1 \end{array}$$

Ejemplo 5.20. Desarrollar $(2x-3)^4$.

Paso 1. Aquí $a=2x$, $b=-3$, $n=4$. Los coeficientes de la fila 4 son $1,4,6,4,1$.

Paso 2. Sumamos término a término, con cuidado con los signos (las potencias impares de -3 son negativas):

$$\begin{aligned}(2x-3)^4 &= \binom{4}{0}(2x)^4 + \binom{4}{1}(2x)^3(-3) + \binom{4}{2}(2x)^2(-3)^2 + \binom{4}{3}(2x)(-3)^3 + \binom{4}{4}(-3)^4 \\ &= 16x^4 + 4\cdot 8x^3(-3) + 6\cdot 4x^2\cdot 9 + 4\cdot 2x\cdot(-27) + 81 \\ &= 16x^4 - 96x^3 + 216x^2 - 216x + 81.\end{aligned}$$

Ejemplo 5.21. ¿Cuál es el coeficiente de x^3 en $(1+x)^7$? Directamente $\binom{7}{3} = 35$, sin desarrollar nada.

Cuidado

El error más frecuente en todo el álgebra escolar es escribir $(a+b)^2 = a^2+b^2$, o su versión general $(a+b)^n = a^n+b^n$. El binomio de Newton dice exactamente por qué está mal y cuánto falta: faltan todos los términos cruzados. Solo en un anillo de característica p y con $n=p$ resulta cierto que $(a+b)^p = a^p+b^p$ —el llamado "sueño del estudiante"—, y eso ocurre porque ahí los coeficientes binomiales intermedios son múltiplos de p; lo veremos en serio con el endomorfismo de Frobenius del capítulo 24.

5.6. Inclusión–exclusión y el principio del palomar

Regla: principio de inclusión–exclusión

Para dos conjuntos finitos,

$$|A\cup B| = |A| + |B| - |A\cap B|.$$

Para tres,

$$|A\cup B\cup C| = |A|+|B|+|C| - |A\cap B| - |A\cap C| - |B\cap C| + |A\cap B\cap C|.$$

El patrón general alterna signos: se suman las intersecciones de un número impar de conjuntos y se restan las de número par.

Ejemplo 5.22. ¿Cuántos enteros del 1 al 100 son múltiplos de 3 o de 5?

Paso 1. Múltiplos de 3: $\lfloor 100/3 \rfloor = 33$. Múltiplos de 5: $\lfloor 100/5 \rfloor = 20$.

Paso 2. Los que son de ambos son múltiplos de 15: $\lfloor 100/15 \rfloor = 6$. Esos los contamos dos veces.

Paso 3. $33 + 20 - 6 = 47$.

Regla: principio del palomar (o del casillero)

Si se reparten n objetos en m cajas y $n > m$, alguna caja recibe al menos dos objetos. Más en general, alguna caja recibe al menos $\lceil n/m \rceil$ objetos.

Ejemplo 5.23. En un grupo de 13 personas, al menos dos cumplen años el mismo mes: hay 12 meses (cajas) y 13 personas (objetos).

Ejemplo 5.24. En cualquier grupo de 367 personas hay dos con la misma fecha de cumpleaños, contando el 29 de febrero. Nótese lo débil que es la conclusión comparada con lo que veremos abajo: la *certeza* necesita 367 personas, pero la *probabilidad del* 50 % se alcanza ya con 23.

Idea clave

El principio del palomar parece una tontería y es una de las herramientas más potentes de la matemática discreta. Su fuerza está en que demuestra *existencia* sin construir nada: garantiza que hay dos objetos en la misma caja sin decir cuáles. Ese mismo estilo de argumento —"no puedo exhibirlo, pero tiene que haberlo"— reaparece en el teorema de Bolzano de los límites elementales, en el argumento diagonal de Cantor del capítulo 29 y en el teorema de Lagrange del capítulo 23.

5.7. Probabilidad: el modelo de Laplace

Definición 5.25. Un *experimento aleatorio* es un proceso cuyo resultado no se puede predecir con certeza. El *espacio muestral* Ω es el conjunto de todos los resultados posibles, y un *suceso* es un subconjunto $A \subseteq \Omega$.

Regla: regla de Laplace

Si Ω es finito y *todos sus resultados son igualmente probables*,

$$P(A) = \frac{\text{número de casos favorables}}{\text{número de casos posibles}} = \frac{|A|}{|\Omega|}.$$

Cuidado

La hipótesis de equiprobabilidad no es decorativa: es toda la regla. Al lanzar dos dados y sumar, el espacio "suma $= 2, 3, \ldots, 12$" tiene once resultados, pero *no* son equiprobables,

y $P(\text{suma} = 7)$ no es $1/11$. Hay que trabajar en el espacio de los 36 pares ordenados, que sí lo son: $P(\text{suma} = 7) = 6/36 = 1/6$. Elegir mal el espacio muestral es la fuente número uno de errores en probabilidad elemental.

Regla: propiedades básicas

1. $0 \le P(A) \le 1$, $P(\varnothing) = 0$, $P(\Omega) = 1$.
2. *Complemento:* $P(A^c) = 1 - P(A)$.
3. *Unión:* $P(A \cup B) = P(A) + P(B) - P(A \cap B)$ (inclusión–exclusión otra vez).
4. Si A y B son *incompatibles* ($A \cap B = \varnothing$), $P(A \cup B) = P(A) + P(B)$.

Ejemplo 5.26. Se lanzan cuatro monedas. ¿Cuál es la probabilidad de obtener al menos una cara?

Paso 1. "Al menos una" pide el complemento: el suceso contrario es "ninguna cara", es decir, cuatro cruces.

Paso 2. $P(\text{ninguna cara}) = \dfrac{1}{2^4} = \dfrac{1}{16}$.

Paso 3. $P(\text{al menos una}) = 1 - \dfrac{1}{16} = \dfrac{15}{16} = 0{,}9375$.

Ejemplo 5.27 (el problema del cumpleaños)**.** En un grupo de 23 personas, ¿cuál es la probabilidad de que al menos dos compartan cumpleaños (con 365 días, ignorando bisiestos)?

Paso 1. Complemento: que *todas* sean distintas.

Paso 2. La primera persona puede cumplir cualquier día (365/365); la segunda, cualquiera salvo el de la primera (364/365); y así:

$$P(\text{todas distintas}) = \frac{365}{365} \cdot \frac{364}{365} \cdots \frac{343}{365} = \frac{P(365, 23)}{365^{23}} \approx 0{,}4927.$$

Paso 3. $P(\text{alguna coincidencia}) \approx 1 - 0{,}4927 = 0{,}5073$, algo más del 50 %.

Pausa para pensar

El resultado del cumpleaños desconcierta a casi todo el mundo, y vale la pena entender por qué. La intuición compara 23 con 365 y concluye "poquísimo". Pero el suceso no es "alguien coincide *conmigo*", sino "alguien coincide con alguien": lo que hay que comparar no es el número de personas con el de días, sino el número de *parejas*, que es $\binom{23}{2} = 253$. Con 253 oportunidades de coincidir, cada una de probabilidad $1/365$, la sorpresa se disuelve. La moraleja general es que en probabilidad hay que contar las parejas, no los individuos, siempre que el suceso involucre relaciones.

Ejemplo 5.28. De una baraja de 52 cartas se extraen dos sin reposición. ¿Probabilidad de que sean del mismo palo?

Casos favorables: elegir palo (4 maneras) y dos cartas de ese palo ($\binom{13}{2} = 78$): $4 \cdot 78 = 312$. Casos posibles: $\binom{52}{2} = 1326$. Luego $P = 312/1326 = 4/17 \approx 0{,}2353$.

Ejemplo 5.29 (una mano de póquer)**.** ¿Probabilidad de un *full* (tres de un valor y dos de otro)?

Paso 1. Elegimos el valor del trío (13), sus tres palos ($\binom{4}{3} = 4$), el valor de la pareja (12, distinto del anterior) y sus dos palos ($\binom{4}{2} = 6$).

Paso 2. Favorables: $13 \cdot 4 \cdot 12 \cdot 6 = 3744$.

Paso 3. $P = \dfrac{3744}{2\,598\,960} = \dfrac{6}{4165} \approx 0{,}00144$, es decir, algo menos de una mano de cada 700.

5.8. Probabilidad condicionada, independencia y Bayes

Definición 5.30. La *probabilidad condicionada* de A dado B (con $P(B) > 0$) es

$$P(A \mid B) = \frac{P(A \cap B)}{P(B)}.$$

Idea clave

Condicionar es *cambiar de espacio muestral*. Cuando te dicen "sabiendo que salió par", el mundo ya no tiene seis resultados sino tres, y todas las probabilidades se recalculan dentro de ese mundo más pequeño. La división por $P(B)$ es exactamente la renormalización que hace que el nuevo mundo vuelva a sumar 1. Si entiendes eso, no necesitas memorizar la fórmula: la reconstruyes.

Definición 5.31. A y B son *independientes* si $P(A \cap B) = P(A)P(B)$; equivalentemente, si $P(A \mid B) = P(A)$: saber que ocurrió B no cambia la apuesta sobre A.

Cuidado

Independiente no es lo mismo que incompatible; de hecho son casi opuestos. Dos sucesos incompatibles con probabilidad positiva son *fuertemente dependientes*: si ocurre uno, el otro es imposible. La confusión viene de que ambas palabras suenan a "no tienen que ver", pero una habla de conjuntos disjuntos y la otra de una igualdad numérica.

Teorema 5.32 (Bayes)**.** *Si* $P(A), P(B) > 0$,

$$P(A \mid B) = \frac{P(B \mid A)\,P(A)}{P(B)}, \qquad \text{y si } \Omega = A \sqcup A^c, \quad P(B) = P(B \mid A)P(A) + P(B \mid A^c)P(A^c).$$

Ejemplo 5.33 (la prueba médica)**.** Una enfermedad afecta al 1 % de la población. Una prueba detecta al 99 % de los enfermos (sensibilidad) y da falso positivo en el 5 % de los sanos. Si a alguien le sale positivo, ¿cuál es la probabilidad de que esté enfermo?

Paso 1. Datos: $P(E) = 0{,}01$, $P(+ \mid E) = 0{,}99$, $P(+ \mid E^c) = 0{,}05$, $P(E^c) = 0{,}99$.

Paso 2. Probabilidad total de dar positivo:

$$P(+) = 0{,}99 \cdot 0{,}01 + 0{,}05 \cdot 0{,}99 = 0{,}0099 + 0{,}0495 = 0{,}0594.$$

Paso 3. Bayes:

$$P(E \mid +) = \frac{0{,}0099}{0{,}0594} = \frac{1}{6} \approx 16{,}7\,\%.$$

Pausa para pensar

El resultado anterior es de los que cambian la manera de leer el periódico. Una prueba que "acierta el 99 % de las veces" deja al paciente positivo con solo un 17 % de probabilidad de estar enfermo. La razón es puramente aritmética: en una población de 10 000 personas hay 100 enfermos (de los que 99 dan positivo) y 9900 sanos (de los que 495 dan positivo). Los falsos positivos son cinco veces más numerosos que los verdaderos, simplemente porque el grupo del que salen es cien veces más grande. Cuando la enfermedad es rara, la *prevalencia* domina a la *precisión*. Esta es la falacia de la tasa base, y no se corrige con intuición: se corrige contando.

Ejemplo 5.34 (Monty Hall)**.** Hay tres puertas; detrás de una hay un auto y detrás de las otras dos, cabras. Eliges una puerta. El presentador, que sabe dónde está el auto, abre *otra* puerta con una cabra y te ofrece cambiar. ¿Conviene?

Paso 1. Tu puerta inicial acierta con probabilidad 1/3; el auto está en las otras dos con probabilidad 2/3.

Paso 2. Al abrir una cabra, el presentador no reparte esa probabilidad 2/3 entre las dos puertas restantes: la concentra entera en la que queda cerrada, porque él nunca abriría la del auto.

Paso 3. Cambiar gana con probabilidad 2/3; quedarse, con 1/3. Conviene cambiar.

La clave —y lo que hace el problema difícil— es que el presentador *tiene información* y la usa. Si abriera una puerta al azar y por casualidad saliera cabra, las dos puertas quedarían a 1/2 y daría igual.

5.9. Ensayos repetidos: la distribución binomial

Regla: distribución binomial

Si se repite n veces un experimento independiente que tiene éxito con probabilidad p, la probabilidad de obtener exactamente k éxitos es

$$P(X = k) = \binom{n}{k} p^k (1-p)^{n-k}, \qquad k = 0, 1, \ldots, n.$$

El factor $p^k(1-p)^{n-k}$ es la probabilidad de *una* secuencia concreta con k éxitos; el $\binom{n}{k}$ cuenta cuántas secuencias así hay.

Ejemplo 5.35. Se lanza una moneda justa 10 veces. ¿Probabilidad de exactamente 6 caras?

$$P = \binom{10}{6}\left(\tfrac{1}{2}\right)^6\left(\tfrac{1}{2}\right)^4 = \frac{210}{1024} = \frac{105}{512} \approx 0{,}2051.$$

Ejemplo 5.36. Se lanzan cinco dados. ¿Probabilidad de exactamente tres seises?

$$P = \binom{5}{3}\left(\tfrac{1}{6}\right)^3\left(\tfrac{5}{6}\right)^2 = 10 \cdot \frac{1}{216} \cdot \frac{25}{36} = \frac{125}{3888} \approx 0{,}0322.$$

Definición 5.37. La *esperanza* (o valor esperado) de una variable aleatoria X que toma valores $x_1, \dots, x_m$ con probabilidades $p_1, \dots, p_m$ es

$$\mathbb{E}[X] = \sum_{i=1}^{m} x_i\, p_i.$$

Ejemplo 5.38. Un dado justo: $\mathbb{E}[X] = \frac{1}{6}(1+2+3+4+5+6) = \frac{21}{6} = 3{,}5$. Nótese que la esperanza no tiene por qué ser un valor posible: es un promedio a largo plazo, no una predicción.

Observación 5.39. Para la binomial, $\mathbb{E}[X] = np$. En diez lanzamientos de moneda se esperan cinco caras; en 600 lanzamientos de dado, 100 seises. La demostración limpia usa que la esperanza es *lineal* —$\mathbb{E}[X+Y] = \mathbb{E}[X] + \mathbb{E}[Y]$ incluso si X e Y no son independientes—, escribiendo X como suma de n indicadores de valor esperado p cada uno. Esa linealidad es la propiedad más útil de toda la probabilidad elemental.

Idea clave

La probabilidad de este capítulo es finita y elemental, pero el edificio completo requiere la teoría de la medida: cuando Ω es infinito no numerable (por ejemplo, "elegir un número al azar en $[0,1]$"), "número de casos favorables entre casos posibles" pierde sentido y hay que sustituirlo por *medida* de favorables entre medida del total. Esa es exactamente la construcción de Lebesgue del capítulo 31, y la axiomática moderna de Kolmogorov consiste en decir que una probabilidad *es* una medida con masa total 1. Todo lo que hemos hecho aquí sobrevive intacto: la regla de Laplace es el caso de la medida de conteo normalizada.

5.10. Errores típicos de combinatoria y probabilidad

Cuidado

1. **Confundir combinaciones con variaciones.** Antes de escribir nada, pregunta: "si intercambio dos de los elegidos, ¿es otro resultado?". Sí $\Rightarrow P(n,k)$; no $\Rightarrow \binom{n}{k}$.
2. **Sumar cuando hay que multiplicar.** "Y" multiplica, "o" suma. Un menú de entrada *y* postre no tiene $3+2$ opciones.

3. **Contar dos veces.** En problemas con "al menos" o con casos solapados, casi siempre hay repeticiones; inclusión–exclusión o el complemento las eliminan.
4. **Suponer equiprobabilidad donde no la hay.** La suma de dos dados no es uniforme; el número de hijos varones tampoco. Laplace solo vale si el espacio elegido es simétrico.
5. **Confundir $P(A \mid B)$ con $P(B \mid A)$.** La probabilidad de dar positivo estando enfermo (99 %) no es la de estar enfermo dando positivo (17 %). Es el error de Bayes, y es responsable de errores judiciales reales.
6. **Creer que independiente significa incompatible.** Son cosas distintas y casi contrarias.
7. **La falacia del jugador.** Después de cinco caras seguidas, la moneda no "debe" una cruz: sigue siendo $1/2$. La moneda no tiene memoria; el promedio a largo plazo se restablece por *dilución* de las rachas, no por compensación.
8. **Olvidar dividir por los repetidos.** Las ordenaciones de BANANA no son $6!$, sino $6!/(3!\,2!)$.

5.11. Funciones generatrices: contar con series

Contar a mano funciona hasta que el problema tiene una recursión. A partir de ahí conviene guardar toda la sucesión dentro de una sola función y dejar que el álgebra haga el trabajo. La técnica la inventó Euler y sigue siendo la herramienta más potente de la combinatoria.

Definición 5.40. La *función generatriz ordinaria* de una sucesión (a_n) es la serie formal $A(x) = \sum_{n\geq 0} a_n x^n$. La *exponencial* es $\hat{A}(x) = \sum_{n\geq 0} a_n \frac{x^n}{n!}$. No hace falta que converjan: se manipulan como objetos algebraicos, y el coeficiente de x^n se denota $[x^n]A(x)$.

Idea clave

Qué dice esto en cristiano. Una sucesión infinita es difícil de manejar; una función es fácil. La función generatriz es un colgador donde se cuelgan todos los términos a la vez, cada uno marcado con su potencia de x, y entonces las operaciones combinatorias se convierten en operaciones algebraicas: elegir entre dos opciones disjuntas es sumar funciones, combinar dos estructuras independientes es multiplicarlas, y una recursión lineal se convierte en una ecuación que se despeja. La convergencia da igual porque nunca se evalúa x; sólo se leen coeficientes.

Regla: Resolver una recursión con funciones generatrices

1. **Escribir** $A(x) = \sum a_n x^n$ y multiplicar la recursión por x^n.
2. **Sumar sobre todos los** n donde la recursión es válida.

3. **Reconocer las sumas desplazadas**: $\sum a_{n-1}x^n = xA(x)$ y $\sum a_{n-2}x^n = x^2A(x)$.
4. **Ajustar los términos iniciales**, que suelen quedar sueltos y son la fuente habitual de errores.
5. **Despejar** $A(x)$, que resulta ser una función racional si la recursión era lineal con coeficientes constantes.
6. **Descomponer en fracciones simples** y desarrollar cada una como serie geométrica.
7. **Leer el coeficiente** de x^n, que es la fórmula cerrada buscada.

Ejemplo 5.41 (Cuatro sucesiones, cuatro funciones)**. (a) Fibonacci.** De $a_n = a_{n-1} + a_{n-2}$ sale $A(x) = \dfrac{x}{1 - x - x^2}$, cuyo desarrollo es $x + x^2 + 2x^3 + 3x^4 + 5x^5 + 8x^6 + 13x^7 + 21x^8 + \cdots$ Descomponiendo en fracciones simples con las raíces $\frac{1\pm\sqrt{5}}{2}$ se obtiene la fórmula de Binet sin adivinar nada.

(b) Números de Catalan. Cuentan, entre otras cien cosas, las formas de emparejar n paréntesis y las triangulaciones de un polígono. La recursión $C_{n+1} = \sum C_i C_{n-i}$ es exactamente el cuadrado de la función generatriz, luego $C(x) = 1 + xC(x)^2$, y despejando, $C(x) = \frac{1-\sqrt{1-4x}}{2x} = 1 + x + 2x^2 + 5x^3 + 14x^4 + 42x^5 + \cdots$ De ahí sale $C_n = \frac{1}{n+1}\binom{2n}{n}$.

(c) Desarreglos. Las permutaciones sin puntos fijos tienen función generatriz exponencial $\frac{e^{-x}}{1-x}$, de donde $D_n = n!\sum_{k=0}^{n} \frac{(-1)^k}{k!}$. Los primeros valores son $0, 1, 2, 9, 44, 265$, y el cociente $D_n/n!$ tiende a $1/e = 0{,}368$: la probabilidad de que ningún regalo llegue a su dueño no depende del número de personas.

(d) Particiones. El número de formas de escribir n como suma de enteros positivos tiene función generatriz $\prod_{k\geq 1} \frac{1}{1-x^k}$, la identidad de Euler. Para $n = 6$ da $p(6) = 11$. No existe fórmula cerrada elemental, y la asintótica de Hardy y Ramanujan, $p(n) \sim \frac{1}{4n\sqrt{3}} e^{\pi\sqrt{2n/3}}$, es uno de los resultados más espectaculares de la combinatoria analítica.

Cuidado

[Contraejemplo: el producto no siempre es lo que uno cree] Multiplicar funciones generatrices ordinarias cuenta pares de estructuras *etiquetadas por posición*, y multiplicar las exponenciales cuenta pares de estructuras *sobre conjuntos disjuntos de etiquetas*. Usar la ordinaria donde tocaba la exponencial da resultados que parecen razonables y son falsos. La regla es sencilla: objetos indistinguibles, ordinaria; objetos etiquetados, exponencial. Equivocarse de familia es el error clásico del principiante y no produce ningún síntoma visible.

Pausa para pensar

Euler introdujo las funciones generatrices en 1748 para contar particiones, y su identidad del producto infinito sigue siendo el punto de partida del tema. Pólya las llamó "una máquina de hacer galletas", y Herbert Wilf abrió su libro *generatingfunctionology* con la frase que mejor las resume: una función generatriz es un tendedero donde colgamos una sucesión de números para exhibirla. Lo notable es que, una vez colgada, el tendedero se puede sumar, multiplicar, derivar e integrar, y la sucesión obedece.

5.12. El cuarto casillero: combinaciones con repetición, e inclusión–exclusión en general

Regla: la tabla completa de los cuatro casilleros

Elegir k objetos de entre n tipos:

	importa el orden	no importa
sin repetir	$\dfrac{n!}{(n-k)!}$	$\binom{n}{k}$
con repetición	n^k	$\binom{n+k-1}{k}$

El único casillero que no hemos justificado es el de abajo a la derecha, y su demostración es una de las biyecciones más bonitas de la combinatoria.

Teorema 5.42 (Barras y estrellas)**.** *El número de multiconjuntos de tamaño k tomados de n tipos —equivalentemente, el número de soluciones enteras no negativas de $x_1 + \cdots + x_n = k$— es* $\binom{n+k-1}{k}$.



Figura 5.2: La solución $(3, 0, 5, 2)$ de $x_1 + x_2 + x_3 + x_4 = 10$ codificada como una fila de diez estrellas separadas por tres barras. Elegir la solución es elegir dónde van las barras entre las trece posiciones, y de ahí el $\binom{n+k-1}{k}$ del cuarto casillero de la tabla.

Demostración. Represéntese cada solución como una fila de k estrellas separadas en n grupos por $n-1$ barras. Por ejemplo, con $n = 4$ y $k = 10$, la solución $(3, 0, 5, 2)$ es

$$\star\star\star\,|\,|\,\star\star\star\star\star\,|\,\star\star.$$

Toda fila de k estrellas y $n-1$ barras corresponde a exactamente una solución y viceversa. Hay $k+n-1$ posiciones y hay que elegir cuáles son barras: $\binom{n+k-1}{n-1} = \binom{n+k-1}{k}$. □

Ejemplo 5.43 (Caramelos, con y sin condición)**. (a)** Repartir 10 caramelos idénticos entre 4 niños:

$$\binom{10+4-1}{10} = \binom{13}{3} = 286.$$

(b) Con al menos uno cada uno. Se da un caramelo a cada niño de entrada y se reparten los 6 restantes libremente. Con el cambio $y_i = x_i - 1$:

$$\binom{6+4-1}{6} = \binom{9}{3} = 84.$$

El truco de "regalar el mínimo y volver al caso libre" resuelve todas las variantes con cotas inferiores.

Teorema 5.44 (Inclusión–exclusión, forma general)**.**

$$\Big|\bigcup_{i=1}^{n} A_i\Big| = \sum_{\varnothing \neq S \subseteq \{1,\dots,n\}} (-1)^{|S|+1} \Big|\bigcap_{i\in S} A_i\Big|.$$

Demostración. Cuéntese cuántas veces aparece un elemento que pertenece exactamente a m de los conjuntos. Aparece en $\binom{m}{j}$ intersecciones de tamaño j, con signo $(-1)^{j+1}$, luego su aportación total es

$$\sum_{j=1}^{m} (-1)^{j+1}\binom{m}{j} = 1 - \sum_{j=0}^{m}(-1)^j\binom{m}{j} = 1 - (1-1)^m = 1.$$

Cada elemento se cuenta exactamente una vez. □

Ejemplo 5.45 (Un filtro con tres primos)**.** ¿Cuántos enteros de 1 a 1000 no son divisibles ni por 2, ni por 3, ni por 5? Los múltiplos son 500, 333 y 200; los de dos en dos, 166 (por 6), 100 (por 10) y 66 (por 15); y 33 por 30. Luego

$$1000 - (500 + 333 + 200) + (166 + 100 + 66) - 33 = 266.$$

Comprobado exhaustivamente: son exactamente 266. De paso, $266/1000 = 0{,}266$ se aproxima al producto $(1-\tfrac{1}{2})(1-\tfrac{1}{3})(1-\tfrac{1}{5}) = \tfrac{4}{15} = 0{,}2667$, que es la forma en que la criba de Eratóstenes se convierte en la fórmula de Euler para φ.

Ejemplo 5.46 (Desarreglos, deducidos y no regalados)**.** ¿Cuántas permutaciones de n elementos no dejan ningún punto fijo? Sea A_i el conjunto de las que fijan el elemento i. Entonces $|A_{i_1} \cap \cdots \cap A_{i_j}| = (n-j)!$, y hay $\binom{n}{j}$ maneras de elegir cuáles. Inclusión–exclusión da

$$D_n = n! - \sum_{j=1}^{n}(-1)^{j+1}\binom{n}{j}(n-j)! = n!\sum_{j=0}^{n}\frac{(-1)^j}{j!}.$$

Con $n=4$: $D_4 = 24(1-1+\frac{1}{2}-\frac{1}{6}+\frac{1}{24}) = 24\cdot\frac{3}{8} = 9$, que se puede contar a mano. Con $n=5$, $D_5 = 44$.

Y como $\sum(-1)^j/j! \to e^{-1}$, la proporción de desarreglos tiende a $1/e = 0{,}367\,879$ con una velocidad asombrosa: ya con $n=7$ el error es menor que 10^{-4}. Si diez personas dejan su abrigo en el guardarropa y se los devuelven al azar, la probabilidad de que *nadie* recupere el suyo es 0,367 879, y esa cifra apenas cambia con cien personas o con un millón.

5.13. Varianza, desviación típica y qué significa "a largo plazo"

Hemos definido la esperanza y la hemos descrito como "un promedio a largo plazo, no una predicción". Faltan las dos piezas que hacen operativa esa frase: una medida de cuánto se separa X de su media, y el teorema que garantiza la convergencia.

Definición 5.47 (Varianza)**.**

$$\operatorname{Var}(X) = E\big[(X-\mu)^2\big] = E[X^2] - \mu^2, \qquad \sigma = \sqrt{\operatorname{Var}(X)}.$$

Ante transformaciones afines, $\operatorname{Var}(aX+b) = a^2\operatorname{Var}(X)$: sumar una constante no dispersa, multiplicar dispersa al cuadrado. Y si X e Y son independientes, $\operatorname{Var}(X+Y) = \operatorname{Var}(X) + \operatorname{Var}(Y)$.

Ejemplo 5.48 (Dado y binomial)**. (a) Un dado.** $E[X] = 3{,}5$ y $E[X^2] = \frac{1+4+9+16+25+36}{6} = \frac{91}{6}$, luego

$$\operatorname{Var}(X) = \frac{91}{6} - \frac{49}{4} = \frac{35}{12} = 2{,}916\,667, \qquad \sigma = 1{,}707\,825.$$

(b) Binomial. Escríbase $X = I_1 + \cdots + I_n$ como suma de indicadores independientes con $E[I] = p$ y $\operatorname{Var}(I) = p(1-p)$. La misma descomposición que da $E[X] = np$ da ahora

$$\operatorname{Var}(X) = np(1-p).$$

Para $B(10;0{,}5)$: $\mu = 5$, $\sigma^2 = 2{,}5$, $\sigma = 1{,}581\,139$. La regla empírica dice que unas dos terceras partes de las tiradas darán entre 3,4 y 6,6 caras, y en efecto $P(4 \le X \le 6) = 0{,}656\,25$.

Teorema 5.49 (Desigualdad de Chebyshev)**.** *Para cualquier variable con varianza finita y todo $k > 0$,*

$$P\big(|X-\mu| \ge k\sigma\big) \le \frac{1}{k^2}.$$

Es una cota grosera —para la normal, la probabilidad real de salirse de 2σ es 0,046 y Chebyshev solo garantiza 0,25— pero es *universal*: no supone nada sobre la forma de la distribución. Y de ella sale, en tres líneas, el teorema que da sentido a toda la sección.

Teorema 5.50 (Ley débil de los grandes números)**.** *Si $X_1, X_2, \ldots$ son independientes con la misma media μ y varianza σ^2 finita, entonces para todo $\varepsilon > 0$*

$$P\big(|\bar{X}_n - \mu| \ge \varepsilon\big) \;\le\; \frac{\sigma^2}{n\varepsilon^2} \;\longrightarrow\; 0.$$

Ejemplo 5.51 (Lo que converge y lo que diverge)**.** Lánzese una moneda n veces y sea S_n el número de caras.

n	σ de la *proporción*	σ del *número*
100	0,05	5
10 000	0,005	50
1 000 000	0,0005	500

La proporción de caras se concentra cada vez más alrededor de 0,5: eso es la ley de los grandes números. Pero la diferencia *absoluta* entre caras y cruces se dispersa cada vez más: crece como $\sqrt{n}$.

Y ahí está la respuesta a la falacia del jugador. Tras diez caras seguidas, la moneda no "debe" ninguna cruz, y de hecho la diferencia acumulada de diez no tiende a compensarse nunca: lo que ocurre es que diez se vuelve despreciable frente a n. El promedio no se corrige, se *diluye*. La ley de los grandes números no promete compensación, promete irrelevancia.

Observación 5.52. La ley *fuerte* afirma algo más: $\bar{X}_n \to \mu$ con probabilidad 1, es decir, casi toda trayectoria individual converge, no solo la probabilidad de desviarse. Requiere solo que la media exista, sin condición sobre la varianza (Kolmogórov, 1930). Y la Cauchy, que no tiene media, es el contraejemplo obligado: el promedio de n Cauchy es otra Cauchy idéntica, así que promediar un millón de mediciones no mejora absolutamente nada.

5.14. Variables aleatorias: qué son y qué significa que sean independientes

Se ha hablado de la esperanza de X, de la varianza de X y de que "X e Y independientes" hace aditiva la varianza, sin haber dicho nunca qué es X ni qué significa esa independencia, que no es la de sucesos definida antes. Conviene cerrar ese hueco, porque la diferencia entre las dos nociones es exactamente lo que separa un enunciado correcto de uno falso.

Definición 5.53 (Variable aleatoria)**.** Sea Ω un espacio muestral. Una *variable aleatoria* es una función

$$X : \Omega \longrightarrow \mathbb{R}$$

que asigna un número a cada resultado posible del experimento. No es una variable en el sentido del álgebra ni es aleatoria en ningún sentido misterioso: es una función perfectamente determinada, y lo único aleatorio es el punto $\omega \in \Omega$ en el que se la evalúa. Para $A \subseteq \mathbb{R}$ se escribe

$$P(X \in A) \;:=\; P\big(\{\omega \in \Omega : X(\omega) \in A\}\big),$$

de modo que toda pregunta sobre X se traduce en una pregunta sobre un suceso de Ω. Al lanzar dos dados, Ω tiene 36 elementos y "la suma" es la función $X(i,j) = i + j$; "el máximo" es otra variable sobre el mismo Ω, y "vale 1 si sale doble y 0 si no" es una tercera, llamada *indicador*.

Definición 5.54 (Independencia de variables aleatorias)**.** Dos variables aleatorias X e Y definidas sobre el mismo Ω son *independientes* si

$$P(X \in A,\ Y \in B) \;=\; P(X \in A)\,P(Y \in B) \qquad \text{para todos los conjuntos } A, B \subseteq \mathbb{R}.$$

Si X e Y toman un número finito de valores, la condición equivale a la versión con valores sueltos: $P(X = x,\, Y = y) = P(X = x)\,P(Y = y)$ para todo par (x, y). Una familia $X_1, \ldots, X_n$ es independiente si la probabilidad conjunta se factoriza para toda elección de conjuntos, no solo dos a dos.

Cuidado

La independencia de *sucesos* es una sola igualdad, $P(A \cap B) = P(A)P(B)$, referida a dos conjuntos concretos. La independencia de *variables* es una familia entera de igualdades, una por cada par (A, B): dicho de otro modo, exige que los sucesos $\{X \in A\}$ y $\{Y \in B\}$ sean independientes *simultáneamente para todos* los A y todos los B. Que una de esas igualdades se cumpla no dice nada de las demás.

El único caso en que ambas nociones coinciden es el de los indicadores: si X e Y solo toman los valores 0 y 1, hay esencialmente un suceso por variable y la independencia de esos dos sucesos ya arrastra la de las variables. Esa coincidencia es la culpable de la confusión, porque casi todos los ejemplos introductorios son indicadores.

Ejemplo 5.55 (Un par de sucesos independiente dentro de un par de variables que no lo es)**.** Sean X e Y con valores en $\{0, 1, 2\}$ y ley conjunta dada por la tabla, con denominador 8:

$P(X = x, Y = y)$	$y = 0$	$y = 1$	$y = 2$
$x = 0$	2/8	1/8	1/8
$x = 1$	1/8	1/8	0
$x = 2$	1/8	0	1/8

Las marginales salen $P(X = 0) = \frac{1}{2}$, $P(X = 1) = P(X = 2) = \frac{1}{4}$, y lo mismo para Y por simetría.

(a) Un par de sucesos independiente. Tómense $A = \{X \geq 1\}$ y $B = \{Y \geq 1\}$. Entonces $P(A) = P(B) = \frac{1}{2}$ y, sumando las cuatro casillas del cuadrante inferior derecho, $P(A \cap B) = \frac{1}{8} + 0 + 0 + \frac{1}{8} = \frac{1}{4} = P(A)P(B)$. Los sucesos son independientes.

(b) Las variables no lo son. Basta bajar de nivel: $P(X = 1, Y = 1) = \frac{1}{8}$ mientras que $P(X = 1)P(Y = 1) = \frac{1}{4} \cdot \frac{1}{4} = \frac{1}{16}$. La igualdad falla, luego X e Y no son independientes por más que aquel par de sucesos lo fuera.

(c) Y la varianza no es aditiva. Se calcula $\mathbb{E}[X] = \mathbb{E}[Y] = \frac{3}{4}$, $\mathbb{E}[X^2] = \frac{5}{4}$ y por tanto $\operatorname{Var}(X) = \operatorname{Var}(Y) = \frac{5}{4} - \frac{9}{16} = \frac{11}{16}$. Además $\mathbb{E}[XY] = 1 \cdot \frac{1}{8} + 4 \cdot \frac{1}{8} = \frac{5}{8}$, de donde

$$\operatorname{Cov}(X, Y) = \mathbb{E}[XY] - \mathbb{E}[X]\mathbb{E}[Y] = \tfrac{5}{8} - \tfrac{9}{16} = \tfrac{1}{16} \neq 0.$$

En consecuencia $\operatorname{Var}(X + Y) = \frac{11}{16} + \frac{11}{16} + 2 \cdot \frac{1}{16} = \frac{3}{2}$, mientras que $\operatorname{Var}(X) + \operatorname{Var}(Y) = \frac{11}{8}$. La fórmula aditiva daría 1,375 en lugar del valor correcto 1,5, un error del orden del ocho por ciento en un ejemplo de nueve casillas.

Observación 5.56. Por qué la varianza y la ley de los grandes números necesitan la versión de variables. Todo se reduce a una cuenta. La igualdad

$$\mathrm{Var}(X+Y) = \mathrm{Var}(X) + \mathrm{Var}(Y) + 2\big(\mathbb{E}[XY] - \mathbb{E}[X]\mathbb{E}[Y]\big)$$

es puro álgebra y siempre es cierta; el término correctivo se anula si y solo si $\mathbb{E}[XY] = \mathbb{E}[X]\mathbb{E}[Y]$. Y ese producto de esperanzas se obtiene desarrollando

$$\mathbb{E}[XY] = \sum_x \sum_y xy\, P(X=x,\, Y=y),$$

suma que se factoriza como $\big(\sum_x xP(X=x)\big)\big(\sum_y yP(Y=y)\big)$ únicamente si *cada* probabilidad conjunta se parte en el producto de las marginales. Es decir: hacen falta tantas igualdades como casillas tenga la tabla. La independencia de un suceso con otro suceso proporciona una sola de ellas, y una sola no permite factorizar una suma doble. El ejemplo anterior lo muestra con números: allí una de las igualdades se cumplía y el término correctivo valía $\frac{1}{8}$, nada despreciable.

La ley débil de los grandes números hereda el problema porque su demostración consiste en aplicar Chebyshev a $\bar{X}_n$, y para eso hay que saber que $\mathrm{Var}(\bar{X}_n) = \sigma^2/n$, lo que exige que la varianza de la suma de las n variables sea $n\sigma^2$: la aditividad, otra vez, y ahora repetida $\binom{n}{2}$ veces, una por cada par. Si la hipótesis fuera "el suceso $\{X_1 > 0\}$ es independiente del suceso $\{X_2 > 0\}$", la cota $\sigma^2/(n\varepsilon^2)$ sencillamente no se podría escribir, y la ley se vendría abajo por falta de la única desigualdad que la sostiene.

Dos matices honestos para terminar. El primero, que lo estrictamente necesario para ambos resultados es algo más débil que la independencia: basta con que las variables estén *incorreladas* dos a dos, es decir, con que $\mathbb{E}[X_iX_j] = \mathbb{E}[X_i]\mathbb{E}[X_j]$. El segundo, que aun así se enuncian con independencia porque esta es estable frente a transformaciones —si X e Y son independientes, también lo son $f(X)$ y $g(Y)$ para cualesquiera f y g—, mientras que la incorrelación no lo es en absoluto: la variable X uniforme sobre $\{-1,0,1\}$ y su cuadrado $Y = X^2$ están incorreladas, porque $\mathbb{E}[X^3] = 0 = \mathbb{E}[X]\mathbb{E}[X^2]$, y sin embargo Y está determinada por X. Incorreladas no es independientes, y la distancia entre ambas cosas aparece en cuanto uno pasa de sumas a cualquier otra operación.

5.15. Ejercicios resueltos

Ejercicio 5.1. ¿Cuántos números de tres cifras distintas se pueden formar con los dígitos $1, 2, 3, 4, 5$?

Solución. Importa el orden y no hay repetición: $P(5,3) = 5 \cdot 4 \cdot 3 = 60$. ■

Ejercicio 5.2. ¿De cuántas maneras se pueden repartir 12 personas en dos equipos de 6, si los equipos se distinguen (local y visitante)?

Solución. Basta elegir quiénes van al equipo local; el resto queda determinado: $\binom{12}{6} = 924$. (Si los equipos *no* se distinguieran, habría que dividir entre 2: 462.) ■

Ejercicio 5.3. Calcula el término independiente de x en el desarrollo de $\left(x^2 + \dfrac{1}{x}\right)^6$.

Solución. El término general es $\binom{6}{k}(x^2)^{6-k}x^{-k} = \binom{6}{k}x^{12-3k}$. El exponente se anula cuando $12-3k=0$, es decir $k=4$. El término es $\binom{6}{4}=15$. ■

Ejercicio 5.4. En una urna hay 5 bolas rojas y 3 azules. Se extraen dos sin reposición. ¿Probabilidad de que ambas sean rojas?

Solución. Por conteo: $\binom{5}{2}/\binom{8}{2} = 10/28 = 5/14$. Por condicionada: $\frac{5}{8}\cdot\frac{4}{7} = \frac{20}{56} = \frac{5}{14} \approx 0{,}357$. Los dos caminos coinciden, como debe ser. ■

Ejercicio 5.5. Se lanzan dos dados. ¿Probabilidad de que la suma sea 8, sabiendo que el primero mostró 3?

Solución. Condicionar reduce el espacio a los seis pares $(3, y)$. Solo $(3, 5)$ suma 8, luego $P = 1/6$. (Sin condicionar, $P(\text{suma}=8) = 5/36$.) ■

Ejercicio 5.6. Un examen tiene 8 preguntas de opción múltiple con 4 opciones cada una. Si se responde todo al azar, ¿probabilidad de acertar exactamente 3?

Solución. Binomial con $n=8$, $p=1/4$:

$$\binom{8}{3}\left(\tfrac{1}{4}\right)^3\left(\tfrac{3}{4}\right)^5 = 56\cdot\frac{1}{64}\cdot\frac{243}{1024} = \frac{13\,608}{65\,536} \approx 0{,}2076.$$

■

Ejercicio 5.7. Demuestra que en cualquier conjunto de $n+1$ enteros elegidos entre 1 y $2n$ hay dos que son primos entre sí.

Solución. Agrupa los $2n$ números en las n cajas $\{1,2\},\{3,4\},\ldots,\{2n-1,2n\}$. Con $n+1$ elegidos y n cajas, el principio del palomar garantiza que dos caen en la misma caja; esos dos son consecutivos, y dos enteros consecutivos siempre tienen máximo común divisor 1. ■

5.16. Ejercicios propuestos

Sin solución desarrollada: solo enunciado y, al final, la respuesta. Empiezan mecánicos (aplicar una fórmula) y terminan conceptuales (elegir cuál). Si tu respuesta difiere, casi siempre es porque contestaste mal a "¿importa el orden?".

5.1. Calcula $\binom{9}{4}$ y $P(9,4)$.

5.2. ¿Cuántas palabras distintas se forman con las letras de PARALELO?

5.3. ¿Cuántos comités de 5 personas hay en un grupo de 12?

5.4. De 6 hombres y 4 mujeres se elige un comité de 4 con exactamente 2 mujeres. ¿Cuántos hay?

5.5. ¿Cuántos enteros del 1 al 200 son múltiplos de 4 o de 6?

5.6. Halla el coeficiente de x^5 en $(1+2x)^8$.

5.7. Se lanzan tres monedas. ¿Probabilidad de exactamente dos caras?

5.8. De una baraja de 52 se saca una carta. ¿Probabilidad de que sea figura (J, Q, K) o de corazones?

5.9. Se lanzan dos dados. ¿Probabilidad de que el producto sea par?

5.10. En una fábrica, la máquina A produce el 60 % de las piezas con un 2 % de defectuosas, y la B el 40 % con un 5 %. Se toma una pieza defectuosa: ¿probabilidad de que venga de A?

5.11. Se lanza una moneda justa 7 veces. ¿Probabilidad de al menos una cruz?

5.12. ¿Cuántos apretones de manos hay si 10 personas se saludan todas con todas exactamente una vez, y qué tiene esto que ver con $\binom{10}{2}$?

Respuestas

1. $\binom{9}{4} = 126$ y $P(9,4) = 3024$; el cociente es $4! = 24$. **2.** PARALELO tiene 8 letras con A dos veces y L dos veces: $8!/(2!\,2!) = 10\,080$. **3.** $\binom{12}{5} = 792$. **4.** $\binom{4}{2}\binom{6}{2} = 6 \cdot 15 = 90$. **5.** $50 + 33 - 16 = 67$ (múltiplos de 12 para la intersección). **6.** $\binom{8}{5}2^5 = 56 \cdot 32 = 1792$. **7.** $\binom{3}{2}/2^3 = 3/8$. **8.** $12/52 + 13/52 - 3/52 = 22/52 = 11/26$. **9.** El complemento es "ambos impares": $1 - \frac{9}{36} = \frac{27}{36} = \frac{3}{4}$. **10.** $P(\text{def}) = 0{,}6 \cdot 0{,}02 + 0{,}4 \cdot 0{,}05 = 0{,}032$; $P(A \mid \text{def}) = 0{,}012/0{,}032 = 3/8 = 37{,}5\,\%$. **11.** $1 - 1/2^7 = 127/128$. **12.** 45; un apretón es un par no ordenado de personas, luego es exactamente $\binom{10}{2} = 45$.

Parte II

Cálculo y álgebra lineal

Capítulo 6

Cálculo en una variable

Lo que necesitas traer

Del capítulo 1: cuantificadores (la definición de límite es un $\forall\varepsilon\,\exists\delta$ y no se entiende sin ellos) e inducción. Del capítulo 3: nada imprescindible. Conviene manejar con soltura potencias, raíces, logaritmos y las funciones trigonométricas a nivel operativo; todas sus propiedades se demuestran aquí a partir de la definición.

Este capítulo desarrolla con detalle los cimientos: la estructura de $\mathbb{R}$, sucesiones, límites, continuidad, la derivada, la integral y las series. Aunque ya sepas operar, aquí está el *porqué* de cada regla; los teoremas de este capítulo sostienen todo lo que viene después.

6.1. Los números reales y la completitud

Los racionales $\mathbb{Q}$ tienen agujeros: no existe ningún racional cuyo cuadrado sea 2 (si $p^2 = 2q^2$ con p/q irreducible, p sería par, $p = 2m$, y entonces $q^2 = 2m^2$ haría a q también par: contradicción). Lo que distingue a $\mathbb{R}$ de $\mathbb{Q}$ es un único axioma adicional:

Definición 6.1 (Supremo)**.** Sea $A \subseteq \mathbb{R}$ no vacío. Una *cota superior* de A es un M con $a \le M$ para todo $a \in A$. El *supremo* $\sup A$ es la *menor* de las cotas superiores. Análogamente se define el *ínfimo*.

Teorema 6.2 (Axioma de completitud)**.** *Todo subconjunto no vacío de $\mathbb{R}$ acotado superiormente tiene supremo en $\mathbb{R}$.*

Idea clave

Qué dice esto en cristiano. El supremo es "el techo más bajo posible". Cualquier cota superior sirve de techo, pero el supremo es el que se apoya justo encima del conjunto sin dejar hueco. La diferencia con el máximo es sutil y decisiva: el máximo tiene que *pertenecer* al conjunto, el supremo no. Y el axioma de completitud dice que en $\mathbb{R}$ ese techo mínimo siempre existe, cosa que en $\mathbb{Q}$ es falsa. Ese único enunciado es lo que

separa a los racionales de los reales, y de él salen absolutamente todos los teoremas grandes del cálculo: Bolzano, Weierstrass, el valor medio y el teorema fundamental.

Ejemplo 6.3 (Supremos que se alcanzan y supremos que no)**. (a) Con máximo y sin máximo.** En $A = [0, 1]$ el supremo es 1 y además es máximo, porque $1 \in A$. En $B = (0, 1)$ el supremo sigue siendo 1, pero no hay máximo: ningún elemento de B es el mayor, ya que dado cualquier $b < 1$ siempre existe $\frac{1+b}{2}$, que es mayor y sigue estando en B.

(b) Una sucesión que se acerca sin llegar. El conjunto $\left\{1 - \frac{1}{n}\right\}$ tiene los elementos 0; 0,5; 0,6667; 0,75; 0,8;… Su supremo es 1: es cota superior porque $1 - \frac{1}{n} < 1$ siempre, y es la menor porque cualquier $M < 1$ es superado en cuanto $n > \frac{1}{1-M}$. No hay máximo.

(c) Uno donde sí se alcanzan los dos. Para $C = \left\{(-1)^n \left(1 + \frac{1}{n}\right)\right\}$ los valores son -2; 1,5; $-1{,}3333$; 1,25; $-1{,}2$;… El supremo es 1,5, alcanzado en $n = 2$, y el ínfimo es -2, alcanzado en $n = 1$. Aquí ambos son máximo y mínimo.

(d) El que rompe $\mathbb{Q}$. Sea $A = \{q \in \mathbb{Q} : q^2 < 2\}$. Está acotado (por 2, por ejemplo) y no es vacío. Sus aproximaciones por decimales son 1,4; 1,41; 1,414; 1,4142; 1,41421; 1,414213…, con cuadrados 1,96; 1,9881; 1,999396; 1,99996164…, cada vez más cerca de 2 por abajo. En $\mathbb{R}$ el supremo es $\sqrt{2} = 1{,}4142136$, pero *ese número no está en* $\mathbb{Q}$: dentro de los racionales, este conjunto acotado no tiene supremo. Ahí es donde $\mathbb{Q}$ falla y $\mathbb{R}$ no.

Cuidado

[Supremo no es máximo, ni es límite] Confundir supremo con máximo es el error número uno del tema, y el apartado (b) lo desactiva: 1 es el supremo y no es máximo. Segundo aviso: el supremo tampoco es lo mismo que el límite, aunque a veces coincidan. En el apartado (c) la sucesión no converge y sin embargo tiene supremo perfectamente definido. Y tercero: la completitud es un *axioma*, no un teorema; no se demuestra a partir de las propiedades algebraicas y de orden, porque $\mathbb{Q}$ las cumple todas y no es completo. Es la propiedad que se elige para definir $\mathbb{R}$.

En $\mathbb{Q}$ esto falla: $\{x \in \mathbb{Q} : x^2 < 2\}$ está acotado pero no tiene supremo racional. De la completitud se deduce la *propiedad arquimediana* ($\mathbb{N}$ no está acotado en $\mathbb{R}$) y la densidad de $\mathbb{Q}$ en $\mathbb{R}$: entre dos reales siempre hay un racional.

Idea clave

Prácticamente todos los teoremas “de existencia” del cálculo (valor intermedio, Weierstrass, convergencia de sucesiones monótonas, criterio de Cauchy) son reformulaciones de la completitud. Es *el* axioma del análisis.

6.2. Los límites laterales y la clasificación exacta de las discontinuidades

Este bloque y los cuatro siguientes reúnen, con definición formal y ejemplo resuelto, las piezas que el resto del capítulo utiliza sin descanso. Se pueden leer ahora o consultar cuando aparezcan; lo que no se debe es darlas por sabidas, porque casi todos los errores del cálculo de una variable consisten en usar una de ellas sin haber comprobado sus hipótesis.

Empezamos por la más elemental. Decir «$f(x)$ tiende a L cuando x tiende a a» obliga a acercarse a a por los dos lados a la vez. Pero hay funciones que se comportan de una manera por la izquierda y de otra por la derecha, y para hablar de ellas hace falta partir la definición en dos mitades.

Definición 6.4 (Límites laterales)**.** Sea f definida a la derecha de a. Se dice que el *límite por la derecha* de f en a vale L, y se escribe $\lim_{x\to a^+} f(x) = L$, si

$$\forall \varepsilon > 0 \;\; \exists \delta > 0 : \quad a < x < a + \delta \implies |f(x) - L| < \varepsilon.$$

Análogamente, el *límite por la izquierda* es $\lim_{x\to a^-} f(x) = L$ si

$$\forall \varepsilon > 0 \;\; \exists \delta > 0 : \quad a - \delta < x < a \implies |f(x) - L| < \varepsilon.$$

La única diferencia con la definición ordinaria de límite es que la condición $0 < |x - a| < \delta$ se ha sustituido por media condición: solo los x de un lado.

Idea clave

Qué dice esto en cristiano. El límite ordinario exige una promesa que valga para todos los puntos próximos a a, vengan de donde vengan. El lateral rebaja la exigencia a los que vienen de un lado concreto. Por eso puede existir un lateral sin que exista el límite, igual que uno puede saber a qué hora llega el tren desde el norte sin saber nada del tren del sur. Y por eso mismo, si los dos laterales existen y son *el mismo número*, ya no queda nadie de quien preocuparse y el límite existe.

Teorema 6.5. *Sea f definida a ambos lados de a. Entonces $\lim_{x\to a} f(x) = L$ si y solo si*

$$\lim_{x\to a^-} f(x) = L \qquad y \qquad \lim_{x\to a^+} f(x) = L.$$

Demostración. ($\Rightarrow$) Si el límite vale L, dado $\varepsilon > 0$ existe δ con $0 < |x-a| < \delta \Rightarrow |f(x)-L| < \varepsilon$. Los x con $a < x < a + \delta$ cumplen en particular $0 < |x - a| < \delta$, luego ese mismo δ sirve para el lateral derecho; y lo mismo para el izquierdo.

($\Leftarrow$) Supongamos que ambos laterales valen L. Dado $\varepsilon > 0$, el lateral izquierdo proporciona un δ_1 y el derecho un δ_2. Tomemos $\delta = \min\{\delta_1, \delta_2\}$, que es positivo por ser el mínimo de dos números positivos. Si $0 < |x - a| < \delta$, entonces x está o bien en $(a - \delta, a)$ o bien en $(a, a + \delta)$; en el primer caso $|f(x) - L| < \varepsilon$ por el lateral izquierdo, y en el segundo por el derecho. En ambos casos se cumple lo pedido. □

Ejemplo 6.6 (Tres funciones y sus dos mitades)**. (a) Un salto limpio.** Sea $f(x) = \dfrac{x^2-1}{|x-1|}$, no definida en $x = 1$. Para $x > 1$ es $|x-1| = x-1$ y la fracción se simplifica a $x+1$, luego $\lim_{x\to 1^+} f(x) = 2$. Para $x < 1$ es $|x-1| = 1-x$ y la fracción vale $-(x+1)$, luego $\lim_{x\to 1^-} f(x) = -2$. Comprobación numérica: $f(0{,}999) = -1{,}999$ y $f(1{,}001) = 2{,}001$. Los dos laterales existen, son distintos y, por el teorema 6.5, el límite ordinario *no* existe.

(b) Uno que ni siquiera es finito. Sea $g(x) = \mathrm{e}^{1/x}$ en $x = 0$. Por la izquierda el exponente $1/x$ se va a $-\infty$ y la exponencial se aplasta contra el cero: $g(-0{,}1) = 0{,}0000454$. Por la derecha el exponente se va a $+\infty$ y la función explota: $g(0{,}1) = 22\,026{,}47$. Aquí el lateral izquierdo existe y vale 0, y el derecho no existe ni siquiera como número.

(c) Uno donde sí coinciden. Sea $h(x) = |x|$ en $x = 0$. Por la derecha $h(x) = x \to 0$ y por la izquierda $h(x) = -x \to 0$: los dos laterales valen 0, luego el límite existe y vale 0. Como además $h(0) = 0$, la función es continua ahí, aunque no sea derivable.

Con los laterales sobre la mesa, la clasificación de las discontinuidades que el capítulo enuncia en prosa se puede escribir con precisión. Recordemos que f es continua en a cuando se cumplen *tres* cosas: que $f(a)$ exista, que $\lim_{x\to a} f(x)$ exista y que ambos coincidan. Fallar cada una da un tipo distinto de avería.

Definición 6.7 (Tipos de discontinuidad)**.** Sea f definida en un entorno de a, salvo quizá en el propio a, y supongamos que f no es continua en a. Se dice que la discontinuidad es:

(i) *evitable*, si $\lim_{x\to a} f(x)$ existe y es finito (pero $f(a)$ no está definido o vale otra cosa);

(ii) *de salto* o de primera especie, si los dos laterales existen, son finitos y son distintos; el número $\lim_{x\to a^+} f - \lim_{x\to a^-} f$ se llama *salto*;

(iii) *esencial* o de segunda especie, si alguno de los dos laterales no existe como número finito, ya sea por explosión a infinito o por oscilación.

Observación 6.8. El nombre de cada tipo indica exactamente qué remedio admite. La evitable se arregla *redefiniendo un solo punto*, poniendo $f(a) := \lim_{x\to a} f(x)$; por eso se llama así. La de salto no se arregla con ninguna redefinición, porque el problema no está en el valor sino en que no hay un destino único. Y la esencial ni siquiera tiene laterales que comparar. Los cuatro ejemplos numéricos del apartado de continuidad ilustran los tres tipos y no hace falta repetirlos; lo que faltaba era el enunciado formal que les pone nombre.

Ejemplo 6.9 (Clasificar cuatro discontinuidades, con el cálculo entero)**. (a)** $f(x) = \dfrac{x^2-4}{x-2}$ en $a = 2$. Factorizando, $x^2 - 4 = (x-2)(x+2)$, y para $x \neq 2$ la función vale $x+2$. Los dos laterales valen 4, luego el límite existe y vale 4; como $f(2)$ no está definido, la discontinuidad es *evitable*, y definiendo $f(2) = 4$ queda continua.

(b) La función parte entera $\lfloor x \rfloor$ en $a = 3$. Por la izquierda toma el valor 2 en todo $(2, 3)$, luego el lateral izquierdo es 2; por la derecha vale 3 en todo $[3, 4)$, luego el lateral derecho es 3. Ambos existen y son distintos: discontinuidad *de salto*, con salto $3 - 2 = 1$.

(c) $f(x) = \mathrm{e}^{1/x}$ en $a = 0$. Por el apartado (b) del ejemplo anterior, el lateral derecho no existe como número finito: discontinuidad *esencial*. Obsérvese que el lateral izquierdo sí existe y vale 0; basta con que falle uno.

(d) $f(x) = \operatorname{sin} \frac{1}{x}$ en $a = 0$. Ninguno de los dos laterales existe, porque en los puntos $x = \frac{1}{k\pi}$ la función vale 0 y en $x = \frac{2}{(4k+1)\pi}$ vale 1, para k tan grande como se quiera, es decir tan cerca de 0 como se quiera. Discontinuidad *esencial*, esta vez por oscilación y no por explosión.

Cuidado

[Evitable no significa inofensiva] Una discontinuidad evitable sigue siendo una discontinuidad mientras no se evite: la función $\frac{\operatorname{sin} x}{x}$ no es continua en 0, y solo pasa a serlo la función *nueva* que se define poniendo el valor 1 ahí. Segundo aviso: que $f(a)$ no esté definido no basta para llamar evitable a nada; hace falta además que el límite exista, y en el apartado (c) no existe. Y tercero: hay discontinuidades esenciales sin ningún infinito de por medio, como la del apartado (d), donde la función está acotada entre -1 y 1 todo el tiempo.

6.3. La derivada de la función inversa: demostración y de dónde salen las derivadas de los arcos y del logaritmo

El capítulo enuncia más adelante que la derivada de la inversa es el inverso de la derivada, y de ahí saca la tabla de derivadas de ln, arcsin y arctan. Ese enunciado merece demostración y, sobre todo, merece verse funcionando en un caso concreto, porque la fórmula tiene una trampa de notación: el punto donde se evalúa cada cosa no es el mismo.

Idea clave

Intuición primero. La gráfica de f^{-1} es la de f reflejada en la recta $y = x$. Al reflejar, lo que era «subir 2 por cada 1 que se avanza» pasa a ser «subir 1 por cada 2 que se avanza»: la pendiente se invierte. Eso es literalmente todo el teorema. Lo único que hay que cuidar es dónde se mide esa pendiente: la de f se mide en el punto a, y la de f^{-1} en su imagen $b = f(a)$, que en general es otro número.

Teorema 6.10 (Derivada de la función inversa)**.** *Sea f continua y estrictamente monótona en un intervalo abierto I, de modo que existe la inversa f^{-1} sobre $J = f(I)$. Si f es derivable en $a \in I$ y $f'(a) \neq 0$, entonces f^{-1} es derivable en $b = f(a)$ y*

$$(f^{-1})'(b) = \frac{1}{f'(a)} = \frac{1}{f'(f^{-1}(b))}.$$

Demostración. Escribamos $g = f^{-1}$. Como f es continua y estrictamente monótona en un intervalo, g también lo es, y en particular g es continua en b. Sea $k \neq 0$ un incremento pequeño y pongamos

$$h = g(b + k) - g(b),$$

de modo que $g(b+k) = a+h$ y, aplicando f, $b+k = f(a+h)$, es decir $k = f(a+h) - f(a)$. Dos observaciones: $h \neq 0$ siempre que $k \neq 0$, porque g es inyectiva; y $h \to 0$ cuando $k \to 0$, porque g es continua en b. Entonces

$$\frac{g(b+k)-g(b)}{k} = \frac{h}{f(a+h)-f(a)} = \frac{1}{\dfrac{f(a+h)-f(a)}{h}}.$$

Cuando $k \to 0$ se tiene $h \to 0$, el denominador tiende a $f'(a) \neq 0$, y el cociente tiende a $1/f'(a)$. Ese límite es, por definición, $g'(b)$. □

Ejemplo 6.11 (La fórmula, aplicada de principio a fin)**. (a) Un caso en el que la inversa no se puede escribir.** Sea $f(x) = x^3 + x$. Su derivada es $f'(x) = 3x^2 + 1$, siempre positiva, luego f es estrictamente creciente en todo $\mathbb{R}$ y tiene inversa. Despejar x de $y = x^3 + x$ exigiría la fórmula de Cardano, pero la derivada de la inversa no necesita despejar nada. Como $f(1) = 2$, tomamos $a = 1$ y $b = 2$:

$$(f^{-1})'(2) = \frac{1}{f'(1)} = \frac{1}{3 \cdot 1^2 + 1} = \frac{1}{4} = 0{,}25.$$

Comprobación numérica: resolviendo $x^3 + x = 2{,}001$ se obtiene $x = 1{,}000249953$, de modo que el cociente incremental de f^{-1} vale $\frac{1{,}000249953-1}{0{,}001} = 0{,}249953$, coherente con 0,25 hasta la cuarta cifra.

(b) El logaritmo. Sea $f(x) = \mathrm{e}^x$, con $f'(x) = \mathrm{e}^x$, y llamemos $g = \ln$. Para $b > 0$ es $a = \ln b$, luego

$$(\ln)'(b) = \frac{1}{f'(\ln b)} = \frac{1}{\mathrm{e}^{\ln b}} = \frac{1}{b}.$$

Ahí está, demostrada y no memorizada, la derivada del logaritmo.

(c) El arcoseno. Sea $f(x) = \sin x$ en $(-\frac{\pi}{2}, \frac{\pi}{2})$, donde es estrictamente creciente, y $g = \arcsin$. Con $a = \arcsin b$ se tiene

$$(\arcsin)'(b) = \frac{1}{\cos(\arcsin b)}.$$

Falta traducir ese coseno. Como $\sin a = b$ y a está en un intervalo donde el coseno es positivo, $\cos a = \sqrt{1 - \sin^2 a} = \sqrt{1 - b^2}$, y queda $(\arcsin)'(b) = \dfrac{1}{\sqrt{1-b^2}}$. Con $b = \frac{3}{5}$ sale $\frac{1}{\sqrt{1-9/25}} = \frac{1}{4/5} = \frac{5}{4} = 1{,}25$.

(d) La arcotangente. Con $f(x) = \tan x$ y $f'(x) = 1 + \tan^2 x$, para $a = \arctan b$ resulta

$$(\arctan)'(b) = \frac{1}{1 + \tan^2(\arctan b)} = \frac{1}{1+b^2},$$

que en $b = 1$ vale $\frac{1}{2}$. Los tres casos siguen el mismo guion: escribir $1/f'$, sustituir el punto por $f^{-1}(b)$, y usar una identidad para eliminar la función inversa que ha quedado dentro.

Cuidado

[La hipótesis $f'(a) \neq 0$ no es un adorno] Si $f'(a) = 0$, la inversa existe igualmente cuando f es estrictamente monótona, pero *no* es derivable en b. El ejemplo es $f(x) = x^3$ en $a = 0$: es estrictamente creciente, su inversa es $f^{-1}(y) = y^{1/3}$, y esa inversa tiene tangente vertical en el origen. Los cocientes incrementales valen $y^{-2/3}$, es decir 4,6416 para $y = 0{,}01$ y 21,544 para $y = 0{,}0001$: se disparan. Al reflejar, una tangente horizontal se convierte en una vertical, y las verticales no son derivadas. Segundo aviso, puramente de notación: la fórmula es $(f^{-1})'(b) = 1/f'(a)$ con $b = f(a)$, y escribirla como $1/f'(b)$ es el error más común del tema; en el apartado (a) daría $1/f'(2) = 1/13$ en vez de $1/4$.

6.4. La primitiva y la integral indefinida: definición, unicidad, y sustitución y partes con demostración

El capítulo habla de primitivas desde el momento en que aparece el teorema fundamental, y las técnicas de integración se presentan como un catálogo de recetas. Falta la definición del objeto y faltan las dos demostraciones, que son cortas y explican por qué las recetas funcionan.

Definición 6.12 (Primitiva e integral indefinida)**.** Sea f definida en un intervalo I. Una *primitiva* de f en I es una función $F : I \to \mathbb{R}$ derivable en todo I con

$$F'(x) = f(x) \qquad \text{para todo } x \in I.$$

El conjunto de todas las primitivas de f en I se llama *integral indefinida* de f y se denota $\int f(x)\,\mathrm{d}x$. El símbolo $\mathrm{d}x$ solo indica cuál es la variable respecto de la que se deriva.

Cuidado

[Dos objetos distintos con notación casi idéntica] La integral *indefinida* $\int f(x)\,\mathrm{d}x$ es una familia de funciones; la integral *definida* $\int_a^b f(x)\,\mathrm{d}x$ de la sección 6.10 es un número, definido con sumas superiores e inferiores y sin mencionar derivadas por ninguna parte. Son conceptos independientes, y lo que los une es un teorema —el fundamental del cálculo—, no la notación. Se parecen porque Leibniz eligió el mismo símbolo justamente porque conocía ese teorema.

Teorema 6.13 (Unicidad salvo constante)**.** *Sean F y G dos primitivas de la misma f en un* intervalo *I. Entonces existe una constante C con $G(x) = F(x) + C$ para todo $x \in I$.*

Demostración. Sea $H = G - F$. Es derivable en I y $H'(x) = G'(x) - F'(x) = f(x) - f(x) = 0$ en todo punto. Por el corolario 6.54, consecuencia directa del teorema del valor medio, una función con derivada nula en un intervalo es constante: si $x < y$ están en I, el teorema del valor medio da $H(y) - H(x) = H'(c)(y - x) = 0$. Luego H es una constante C y $G = F + C$. □

Cuidado

[La palabra «intervalo» es la hipótesis clave] Sobre un dominio que no es un intervalo, el teorema es falso. La función $f(x) = \frac{1}{x}$ está definida en $(-\infty, 0) \cup (0, \infty)$, y tanto $F(x) = \ln|x|$ como

$$G(x) = \begin{cases} \ln|x| + 7, & x < 0, \\ \ln|x| - 3, & x > 0, \end{cases}$$

son primitivas suyas: ambas tienen derivada $\frac{1}{x}$ en todo punto del dominio. Y su diferencia no es constante, sino que vale 7 a la izquierda y -3 a la derecha. La costumbre de escribir $\int \frac{\mathrm{d}x}{x} = \ln|x| + C$ con una sola constante es, estrictamente, un abuso: hay una constante por cada trozo conexo del dominio.

Teorema 6.14 (Integración por sustitución)**.** *Sea g derivable en $[\alpha, \beta]$ con g' continua, y sea f continua en un intervalo que contenga a $g([\alpha, \beta])$. Entonces*

$$\int_\alpha^\beta f(g(x))\, g'(x)\, \mathrm{d}x = \int_{g(\alpha)}^{g(\beta)} f(u)\, \mathrm{d}u.$$

Demostración. Sea F una primitiva de f, que existe por el teorema fundamental 6.66 al ser f continua. Por la regla de la cadena, la función $x \mapsto F(g(x))$ es derivable y

$$\frac{\mathrm{d}}{\mathrm{d}x} F(g(x)) = F'(g(x))g'(x) = f(g(x))g'(x),$$

es decir, $F \circ g$ es una primitiva del integrando de la izquierda. Aplicando la regla de Barrow a los dos miembros:

$$\int_\alpha^\beta f(g(x))g'(x)\, \mathrm{d}x = F(g(\beta)) - F(g(\alpha)) = \int_{g(\alpha)}^{g(\beta)} f(u)\, \mathrm{d}u.$$

□

Regla: cambiar de variable en una integral definida

1. **Elige** $u = g(x)$ de modo que el integrando quede escrito como «algo de u» multiplicado por $g'(x)$. Si el factor $g'(x)$ no está, el cambio no sirve tal cual.
2. **Calcula** $\mathrm{d}u = g'(x)\, \mathrm{d}x$ y sustituye ese bloque entero, no solo una parte.
3. **Cambia también los límites**: el α de abajo pasa a ser $g(\alpha)$ y el β de arriba pasa a ser $g(\beta)$. Este es el paso que más se olvida.
4. **No vuelvas a la variable original** si has cambiado los límites; ya no hace falta y es fuente de errores. Si prefieres no cambiarlos, calcula primero la primitiva en u, deshaz el cambio y evalúa después: las dos vías dan lo mismo.
5. **Comprueba que g es monótona** en el intervalo de trabajo si vas a deshacer el cambio; si no lo es, el mismo valor de u corresponde a dos x distintas y hay que partir la integral.

6. **Verifica** derivando la primitiva obtenida: debe reproducir el integrando de partida.

Ejemplo 6.15 (Tres sustituciones, una de ellas mal hecha)**. (a) El caso limpio.** $\displaystyle\int_0^2 x\,\mathrm{e}^{x^2}\,\mathrm{d}x$. Con $u = x^2$ se tiene $\mathrm{d}u = 2x\,\mathrm{d}x$, luego $x\,\mathrm{d}x = \frac{1}{2}\mathrm{d}u$; los límites pasan de 0 y 2 a 0 y 4:

$$\int_0^2 x\mathrm{e}^{x^2}\mathrm{d}x = \frac{1}{2}\int_0^4 \mathrm{e}^{u}\mathrm{d}u = \frac{\mathrm{e}^4 - 1}{2} = 26{,}799075.$$

Verificación derivando: $\frac{\mathrm{d}}{\mathrm{d}x}\left(\frac{1}{2}\mathrm{e}^{x^2}\right) = x\mathrm{e}^{x^2}$.

(b) Con función trigonométrica. $\displaystyle\int_0^{\pi/2} \sin^3 x\cos x\,\mathrm{d}x$. Con $u = \sin x$, $\mathrm{d}u = \cos x\,\mathrm{d}x$, y los límites pasan de 0 y $\frac{\pi}{2}$ a 0 y 1:

$$\int_0^1 u^3\mathrm{d}u = \left[\frac{u^4}{4}\right]_0^1 = \frac{1}{4} = 0{,}25.$$

(c) El cambio que destruye la integral. $\displaystyle\int_{-1}^{1} x^2\,\mathrm{d}x$, cuyo valor es $\frac{2}{3}$. Si alguien intenta $u = x^2$, los dos límites -1 y 1 se convierten *ambos* en 1 y la integral parece valer 0, lo cual es imposible porque el integrando es positivo. El fallo no está en cambiar los límites, sino en que $g(x) = x^2$ no es monótona en $[-1, 1]$ y además falta el factor $g'(x) = 2x$. La sustitución exige las dos cosas, y el teorema 6.14 las pide explícitamente.

Teorema 6.16 (Integración por partes)**.** *Sean u, v derivables en $[a, b]$ con derivadas continuas. Entonces*

$$\int_a^b u(x)v'(x)\,\mathrm{d}x = \big[u(x)v(x)\big]_a^b - \int_a^b u'(x)v(x)\,\mathrm{d}x.$$

Demostración. Por la regla del producto, $(uv)' = u'v + uv'$, y el miembro derecho es continuo, luego integrable. Integrando los dos lados en $[a, b]$ y aplicando Barrow al de la izquierda:

$$u(b)v(b) - u(a)v(a) = \int_a^b u'v + \int_a^b uv'.$$

Despejando la última integral se obtiene el enunciado. La versión indefinida $\int u\,\mathrm{d}v = uv - \int v\,\mathrm{d}u$ es la misma identidad sin evaluar. □

Ejemplo 6.17 (Dos integrales por partes, enteras)**. (a) El logaritmo, que se integra derivándolo.** $\displaystyle\int_1^{\mathrm{e}} \ln x\,\mathrm{d}x$. No hay producto a la vista, así que se fabrica uno: $u = \ln x$ y $v' = 1$, de donde $u' = \frac{1}{x}$ y $v = x$. Entonces

$$\int_1^{\mathrm{e}} \ln x\,\mathrm{d}x = \big[x\ln x\big]_1^{\mathrm{e}} - \int_1^{\mathrm{e}} x\cdot\frac{1}{x}\,\mathrm{d}x = (\mathrm{e} - 0) - (\mathrm{e} - 1) = 1.$$

El valor es exactamente 1; numéricamente $\mathrm{e} = 2{,}7182818$ y la resta da 1,0000000.

(b) Polinomio por trigonométrica. $\displaystyle\int_0^{\pi} x\sin x\,\mathrm{d}x$. Se deriva el polinomio, que es quien se simplifica: $u = x$, $v' = \sin x$, luego $u' = 1$ y $v = -\cos x$:

$$\int_0^{\pi} x\sin x\,\mathrm{d}x = \big[-x\cos x\big]_0^{\pi} + \int_0^{\pi} \cos x\,\mathrm{d}x = \pi + \big[\sin x\big]_0^{\pi} = \pi + 0 = \pi = 3{,}1415927.$$

La elección correcta de u es la que hace que la integral nueva sea más fácil que la vieja; si se hubiera tomado $u = \sin x$, el polinomio se habría convertido en $\frac{x^2}{2}$ y la cosa empeoraría.

Observación 6.18. La tercera técnica, la descomposición en fracciones simples, sí está desarrollada en el capítulo con ejemplos completos, y no se repite aquí. Lo único que conviene añadir es su fundamento: que toda función racional se descompone de esa manera no es un truco de cálculo sino un teorema de álgebra sobre la factorización de polinomios en $\mathbb{R}[x]$ en factores lineales y cuadráticos irreducibles.

6.5. El teorema fundamental del cálculo: sus dos partes por separado, y la regla de Barrow demostrada

El teorema fundamental se enuncia en este capítulo con dos apartados, y esa numeración esconde que son *dos teoremas distintos*, con hipótesis distintas, que dicen cosas distintas y se usan en direcciones opuestas. Conviene separarlos del todo.

Idea clave

Intuición primero. La primera parte dice que *acumular y luego medir el ritmo devuelve el ritmo original*: si vas anotando el área acumulada bajo una curva y preguntas a qué velocidad crece esa área, la respuesta es la altura de la curva. La segunda dice lo recíproco: para saber cuánta área hay entre a y b, basta con encontrar cualquier función cuyo ritmo de crecimiento sea la curva y restar sus valores en los extremos. La primera parte *construye* primitivas; la segunda las *usa*. Una demuestra que existen, la otra que sirven.

Teorema 6.19 (Parte I: la integral con extremo variable deriva a la función)**.** *Sea f integrable en $[a,b]$ y defínase $F(x) = \int_a^x f(t)\,\mathrm{d}t$ para $x \in [a,b]$. Entonces:*

(i) F es continua en $[a,b]$, incluso si f no lo es;

(ii) si f es continua en un punto c, entonces F es derivable en c y $F'(c) = f(c)$.

En particular, toda función continua en un intervalo tiene primitiva, a saber, F.

Teorema 6.20 (Parte II: regla de Barrow)**.** *Sea f integrable en $[a,b]$ y sea G una función continua en $[a,b]$, derivable en (a,b), con $G'(x) = f(x)$ para todo $x \in (a,b)$. Entonces*

$$\int_a^b f(x)\,\mathrm{d}x = G(b) - G(a).$$

Demostración. Sea $P = \{a = x_0 < x_1 < \cdots < x_n = b\}$ una partición cualquiera. Aplicando el teorema del valor medio 6.52 a G en cada subintervalo $[x_{i-1}, x_i]$, existe $t_i \in (x_{i-1}, x_i)$ con

$$G(x_i) - G(x_{i-1}) = G'(t_i)(x_i - x_{i-1}) = f(t_i)\,\Delta x_i.$$

Sumando en i, la izquierda telescopa:

$$G(b) - G(a) = \sum_{i=1}^{n} f(t_i)\,\Delta x_i.$$

Ahora bien, en cada subintervalo $m_i \le f(t_i) \le M_i$, luego esa suma está atrapada entre la suma inferior y la superior:

$$L(f,P) \;\le\; G(b) - G(a) \;\le\; U(f,P) \qquad \text{para } \textit{toda} \text{ partición } P.$$

Como f es integrable, el supremo de las L y el ínfimo de las U coinciden y valen $\int_a^b f$. Un número atrapado entre ambos para toda partición solo puede ser ese valor común, de donde $G(b) - G(a) = \int_a^b f$. □

Observación 6.21 (Qué pide cada parte, y por qué importa). La parte I necesita la continuidad de f *en el punto* para concluir que F es derivable ahí; sin ella, F sigue siendo continua pero puede tener un pico. La parte II, en cambio, no exige continuidad de f en absoluto: le basta con que f sea integrable y con que exista una G cuya derivada sea f en todo el intervalo abierto. Son hipótesis lógicamente independientes. Existen funciones integrables sin primitiva (una con un salto: por el teorema del valor intermedio de Darboux, ninguna derivada puede saltar) y funciones con primitiva que no son integrables Riemann. La versión que casi todo el mundo recita —«f continua»— es simplemente el caso cómodo en el que ambas partes se aplican a la vez.

Ejemplo 6.22 (Las dos partes trabajando, y la regla de Leibniz)**. (a) Parte I, en un caso sin fórmula.** Sea $F(x) = \int_0^x \mathrm{e}^{-t^2}\mathrm{d}t$. No existe ninguna expresión elemental para F, y sin embargo la parte I garantiza que F es derivable con $F'(x) = \mathrm{e}^{-x^2}$ en todo punto, porque el integrando es continuo. Se sabe todo sobre su derivada sin saber nada sobre ella: por ejemplo, F es estrictamente creciente (la derivada es positiva) y tiene inflexión en 0, pues $F''(x) = -2x\mathrm{e}^{-x^2}$ cambia de signo ahí.

(b) Parte I, reconociendo una función conocida. Sea $F(x) = \int_1^x \frac{\mathrm{d}t}{t}$ para $x > 0$. La parte I da $F'(x) = \frac{1}{x}$ y además $F(1) = 0$. Como el logaritmo cumple lo mismo, el teorema 6.13 obliga a que $F(x) = \ln x$. Esta es, de hecho, una manera de *definir* el logaritmo sin hablar de exponenciales.

(c) Parte II, el cálculo de todos los días. Para $\int_0^1 x^2\mathrm{d}x$ se busca una G con $G' = x^2$; sirve $G(x) = \frac{x^3}{3}$. Entonces la integral vale $\frac{1}{3} - 0 = \frac{1}{3} = 0{,}3333333$, que es exactamente el número al que convergían las sumas superiores e inferiores de la tabla del apartado de la integral de Riemann. Ahí se ve para qué sirve el teorema: el mismo número que costaba un límite de sumas sale ahora de una resta.

(d) Extremos variables, o regla de Leibniz. Si $H(x) = \int_0^{x^2} \sin t\,\mathrm{d}t$, se escribe $H = F \circ g$ con $F(y) = \int_0^y \sin t\,\mathrm{d}t$ y $g(x) = x^2$. Por la parte I, $F'(y) = \sin y$; por la regla de la cadena,

$$H'(x) = F'(g(x))\,g'(x) = \sin(x^2)\cdot 2x.$$

En $x=1$ vale $2\sin 1 = 1{,}6829419$. Comprobación directa: $H(x)=1-\cos(x^2)$, cuya derivada es $2x\sin(x^2)$. En general, $\frac{\mathrm{d}}{\mathrm{d}x}\int_{u(x)}^{v(x)} f = f(v(x))v'(x) - f(u(x))u'(x)$, por el mismo argumento aplicado a los dos extremos.

Cuidado

[Barrow aplicado donde no vale: el error que da un área negativa] Considérese $\int_{-1}^{1}\frac{\mathrm{d}x}{x^2}$. Aplicando Barrow con $G(x)=-\frac{1}{x}$ saldría

$$G(1)-G(-1) = -1-1 = -2,$$

un número negativo para la integral de una función que es positiva en todo su dominio: imposible. El fallo es que G no es derivable en 0, ni siquiera está definida ahí, de modo que la hipótesis «$G'=f$ en todo el intervalo» falla en un punto interior, y eso basta para invalidarlo todo. La integral es en realidad impropia y diverge a $+\infty$. Segundo aviso, el mismo error con otra cara: en $\int_0^{2\pi}\frac{\mathrm{d}x}{2+\cos x}$ el cambio $t=\tan\frac{x}{2}$ aplicado sin cuidado da 0, porque $\tan\frac{x}{2}$ no es continua en $x=\pi$; el valor correcto es $\frac{2\pi}{\sqrt{3}} = 3{,}6275987$. Antes de restar $G(b)-G(a)$ hay que mirar si G existe y es derivable en *todo* el trayecto.

6.6. La regla de L'Hôpital enunciada con todas sus hipótesis, y los cuatro modos de aplicarla mal

Ninguna herramienta del cálculo se usa peor que esta. El motivo es que su enunciado corriente —«se derivan numerador y denominador»— omite todas las condiciones, y sin ellas la regla produce respuestas falsas con total naturalidad. Aquí está el enunciado completo, y a continuación los cuatro abusos, cada uno con su contraejemplo numérico.

Teorema 6.23 (L'Hôpital, versión completa)**.** *Sea I un intervalo abierto que tenga a a como extremo o como punto interior, donde a puede ser un número real, $+\infty$ o $-\infty$. Supongamos que:*

(H1) f y g son derivables en I, salvo quizá en el propio a;

(H2) $g'(x)\neq 0$ para todo $x\in I$ distinto de a;

(H3) se da una de estas dos situaciones: o bien $f(x)\to 0$ y $g(x)\to 0$ cuando $x\to a$, o bien $|g(x)|\to+\infty$ cuando $x\to a$;

(H4) existe el límite $L=\lim_{x\to a}\dfrac{f'(x)}{g'(x)}$, finito o infinito.

Entonces existe también $\lim_{x\to a}\dfrac{f(x)}{g(x)}$ y vale L. El enunciado es válido igualmente para límites laterales $x\to a^+$ y $x\to a^-$.

Observación 6.24 (Dos finuras del enunciado que casi nunca se dicen). Primera: en el caso ∞/∞ basta con que *el denominador* se vaya a infinito; no hace falta suponer nada sobre f.

Por ejemplo, $\lim_{x\to\infty} \frac{\arctan x}{x}$ cae bajo la regla aunque el numerador esté acotado entre $-\frac{\pi}{2}$ y $\frac{\pi}{2}$: el cociente de derivadas es $\frac{1}{1+x^2} \to 0$, y en efecto el cociente original tiende a 0 (en $x = 1000$ vale 0,0015698). Segunda: no se exige que f y g estén definidas en a, ni que valgan 0 ahí; solo se exige el comportamiento *límite* de (H3). Por eso la regla se aplica a $\frac{\sin x}{x}$ en $x = 0$, donde el cociente no existe.

Regla: comprobar las hipótesis antes de derivar nada

1. **Sustituye primero.** Evalúa el límite del numerador y el del denominador por separado. Si no sale $\frac{0}{0}$ ni algo con $|g| \to \infty$, la regla *no* se aplica y el límite se calcula por sustitución directa o por otro camino.
2. **Identifica la forma indeterminada.** Si es una de las cinco restantes ($0\cdot\infty$, $\infty-\infty$, 1^∞, 0^0, ∞^0), transfórmala primero en un cociente; el taller de indeterminaciones del capítulo detalla cómo.
3. **Comprueba que g' no se anula** cerca de a. Con funciones elementales monótonas esto es automático, pero con senos y cosenos dentro no lo es en absoluto.
4. **Deriva por separado** numerador y denominador. No es la regla del cociente: no aparece ningún g^2.
5. **Vuelve al paso 1 con el nuevo cociente.** Si sigue siendo indeterminado, se repite; si ya no lo es, se evalúa y se para. Derivar una vez de más es un error tan grave como no derivar.
6. **Si el cociente de derivadas no tiene límite, abandona la regla**, no concluyas que el original tampoco lo tiene. La implicación va en un solo sentido.
7. **Verifica numéricamente** el resultado evaluando la expresión original en dos o tres puntos cada vez más próximos a a. Es barato y detecta casi todos los errores.

Ejemplo 6.25 (Los cuatro abusos, con números)**. (a) Aplicarla sin indeterminación.** Sea $\lim_{x\to 0} \dfrac{x+2}{x+1}$. Sustituyendo sale $\frac{2}{1} = 2$, sin ninguna indeterminación. Quien derive arriba y abajo obtiene $\frac{1}{1} = 1$, que es falso. La hipótesis (H3) es la que se ha saltado, y es la que más veces se salta.

(b) Seguir derivando cuando ya se ha terminado. Sea $\lim_{x\to 0} \frac{e^x-1}{x}$, de la forma $\frac{0}{0}$. Una aplicación da $\frac{e^x}{1} \to 1$, y ahí se acaba. Si se deriva otra vez sale $\frac{e^x}{0}$, que no significa nada. El paso 5 de la regla existe exactamente para esto.

(c) Concluir que no hay límite porque la regla no concluye. Sea $\lim_{x\to\infty} \frac{2x+\sin x}{3x+1}$, de la forma $\frac{\infty}{\infty}$. El cociente de derivadas es $\frac{2+\cos x}{3}$, que oscila entre $\frac{1}{3}$ y 1 y no tiene límite: la hipótesis (H4) falla y la regla no dice nada. Pero el límite original existe y vale $\frac{2}{3}$, como se ve dividiendo por x: en $x = 1000$ el cociente vale 0,6663. Que la herramienta no responda no significa que la respuesta no exista.

(d) El abuso fino: olvidar que g' no debe anularse. Tomemos, con $x \to +\infty$,

$$f(x) = x + \sin x \cos x, \qquad g(x) = e^{\sin x}\big(x + \sin x \cos x\big).$$

Ambas tienden a $+\infty$, luego (H3) se cumple. Las derivadas son $f'(x) = 2\cos^2 x$ y $g'(x) = \mathrm{e}^{\sin x}\cos x\,(x + \sin x\cos x + 2\cos x)$, de modo que

$$\frac{f'(x)}{g'(x)} = \frac{2\cos x}{\mathrm{e}^{\sin x}(x + \sin x\cos x + 2\cos x)} \longrightarrow 0,$$

porque el numerador está acotado y el denominador crece sin freno: vale 0,0155 en $x = 20$, 0,0283 en $x = 100$ y 0,00049 en $x = 1000$. Existe, pues, el límite de las derivadas y vale 0. Y sin embargo el cociente original se simplifica a

$$\frac{f(x)}{g(x)} = \mathrm{e}^{-\sin x},$$

que oscila para siempre entre $\mathrm{e}^{-1} = 0{,}3678794$ y $\mathrm{e} = 2{,}7182818$, y no tiene ningún límite: toma el primer valor en $x = \frac{\pi}{2} + 2k\pi$ y el segundo en $x = \frac{3\pi}{2} + 2k\pi$, para todo k. La conclusión del teorema es falsa, y lo único que ha fallado es (H2): g' se anula en todos los puntos donde $\cos x = 0$, es decir infinitas veces en cualquier entorno del infinito.

Cuidado

[Y el quinto pecado: la circularidad] Usar L'Hôpital para calcular $\lim_{x\to 0}\frac{\sin x}{x}$ da $\frac{\cos x}{1} \to 1$, resultado correcto y demostración nula: la derivada del seno se ha obtenido precisamente a partir de ese límite, de modo que el argumento se muerde la cola. Como comprobación numérica está bien; como prueba no vale. Lo mismo ocurre con $\lim_{x\to 0}\frac{\mathrm{e}^x - 1}{x}$. Y una recomendación general: cuando el límite se resiste, el desarrollo de Taylor del teorema 6.59 suele ser más seguro y siempre es más informativo, porque no da solo el límite sino la velocidad con que se alcanza.

6.7. Sucesiones

Idea clave

Intuición primero. Una sucesión es una lista infinita de números, y la única pregunta que de verdad importa sobre ella es si *se asienta* en algún sitio. La definición formal con ε y N parece hostil, pero dice algo muy concreto y muy razonable: "dame un margen de error, por pequeño que sea, y yo te diré a partir de qué término la sucesión ya no se sale de ese margen". Nada de "se acerca infinitamente", nada de movimiento: es una promesa que se puede verificar, para cada tolerancia, con un número concreto de términos.
El otro concepto de la sección, el de sucesión de Cauchy, cambia la pregunta de un modo sutil e importante: en vez de exigir que los términos se acerquen *a un límite*, exige que se acerquen *entre sí*. La ventaja es que se puede comprobar sin saber cuál es el límite —a menudo no lo conoces— y el teorema de completitud dice que en $\mathbb{R}$ eso basta: si los términos se apiñan, hay algo ahí a lo que se apiñan. En $\mathbb{Q}$ eso sería falso, y esa diferencia es literalmente la razón de ser de los números reales.

Definición 6.26. Una sucesión (a_n) *converge* a L (escrito $a_n \to L$) si para todo $\varepsilon > 0$ existe N tal que $n \geq N \implies |a_n - L| < \varepsilon$.

Propiedades básicas (todas se demuestran con la desigualdad triangular): el límite es único; toda sucesión convergente está acotada; límites de sumas, productos y cocientes (denominador no nulo) son suma, producto y cociente de límites; y el *teorema del sándwich*: si $a_n \le b_n \le c_n$ con $a_n \to L$ y $c_n \to L$, entonces $b_n \to L$.

Teorema 6.27 (Convergencia monótona)**.** *Toda sucesión creciente y acotada superiormente converge (a su supremo).*

Demostración. Sea $L = \sup\{a_n\}$, que existe por completitud. Dado $\varepsilon > 0$, $L - \varepsilon$ no es cota superior, así que existe N con $a_N > L - \varepsilon$. Por monotonía, para $n \ge N$: $L - \varepsilon < a_N \le a_n \le L$, es decir $|a_n - L| < \varepsilon$. □

Ejemplo 6.28 (El número e)**.** La sucesión $a_n = (1 + \frac{1}{n})^n$ es creciente y acotada (se comprueba con el binomio de Newton: $a_n = \sum_{k=0}^{n} \binom{n}{k} n^{-k} \le \sum_k \frac{1}{k!} < 3$), luego converge; su límite se llama $e = 2.71828\ldots$

Teorema 6.29 (Bolzano–Weierstrass)**.** *Toda sucesión acotada tiene una subsucesión convergente.*

Idea. Bisección: si (a_n) vive en $[A, B]$, alguna de las dos mitades del intervalo contiene infinitos términos; se elige esa mitad y se repite. Los intervalos encajados (cuya intersección es no vacía, por completitud) atrapan el límite de una subsucesión. □

Idea clave

Qué dice esto en cristiano. Una sucesión acotada puede no ir a ninguna parte, saltando sin descanso. Lo que el teorema afirma es que, por muy caótico que sea el salto, si todo ocurre dentro de una caja finita entonces *alguna parte* de la sucesión sí se está acercando a algún sitio: se puede pescar de ella una subsucesión convergente. La razón es un argumento de bisección: partes la caja por la mitad, una de las dos mitades contiene infinitos términos, repites, y las cajas se encogen hasta un punto. Es el teorema del que cuelga toda la compacidad.

Ejemplo 6.30 (Pescar subsucesiones que sí convergen)**. (a) Una sucesión que oscila.** Sea $a_n = (-1)^n + \frac{1}{n}$, cuyos primeros valores son 0; 1,5; −0,6667; 1,25; −0,8; 1,1667;… No converge, porque salta entre las cercanías de -1 y de 1. Pero los términos pares dan 1,5; 1,25; 1,1667; 1,125; 1,1, que convergen a 1, y los impares convergen a -1. Hay dos subsucesiones convergentes, con límites distintos.

(b) Un caso mucho peor. La sucesión $a_n = \sin n$ (con n en radianes) da 0,8415; 0,9093; 0,1411; −0,7568; −0,9589;… y no converge ni tiene un patrón periódico. Aun así está acotada entre -1 y 1, luego el teorema garantiza subsucesiones convergentes. Por ejemplo, para $n = 22, 44, 66, 88, 110$ el valor $|\sin n|$ es menor que 0,05: hay una subsucesión que tiende a 0. De hecho se puede pescar una que tienda a cualquier valor entre -1 y 1.

(c) La bisección, con números. Si toda la sucesión vive en $[0, 2]$, una de las mitades $[0, 1]$ o $[1, 2]$ contiene infinitos términos; se elige esa y se repite. Las longitudes son

1; 0,5; 0,25; 0,125; 0,0625; …, es decir $2/2^k$, que tiende a cero. Eligiendo un término en cada caja se obtiene la subsucesión buscada, y su límite es el punto común a todas.

(d) De dónde viene la fuerza. El paso "las cajas encajadas tienen un punto común" no es gratis: es exactamente el axioma de completitud. En $\mathbb{Q}$ el argumento se cae, y de hecho la sucesión de aproximaciones decimales de $\sqrt{2}$ está acotada, es de racionales, y ninguna subsucesión suya converge *dentro de* $\mathbb{Q}$.

Cuidado

[Contraejemplo: la acotación es imprescindible] Sin acotación el teorema es falso de la manera más simple posible: la sucesión $a_n = n$ no tiene ninguna subsucesión convergente, porque cualquier subsucesión suya también tiende a infinito. Segundo aviso: el teorema promete *una* subsucesión convergente, no que todas lo sean ni que el límite sea único; el apartado (a) tiene dos límites distintos, y el (b) tiene infinitos. Esos valores se llaman puntos de acumulación, y el teorema dice simplemente que hay al menos uno.

Definición 6.31. (a_n) es *de Cauchy* si para todo $\varepsilon > 0$ existe N tal que $m, n \geq N \implies |a_m - a_n| < \varepsilon$.

Teorema 6.32. *En $\mathbb{R}$, una sucesión converge si y solo si es de Cauchy.*

Este criterio permite probar convergencia *sin conocer el límite*; es la base de la convergencia de series y, mucho más allá, de la noción de espacio completo del análisis funcional.

Límites de referencia que conviene tener a mano:

$$\sqrt[n]{n} \to 1, \qquad \frac{c^n}{n!} \to 0, \qquad n^{1/n^2} \to 1, \qquad \left(1 + \frac{x}{n}\right)^n \to e^x, \qquad \frac{\ln n}{n^p} \to 0 \ (p > 0).$$

6.8. Límites de funciones y continuidad

Idea clave

Intuición primero. El los límites elementales presentó el límite como "hacia dónde se dirige la función". Aquí esa frase se convierte en una definición que se puede usar en una demostración, y el precio de la precisión es la aparición de ε y δ. Conviene leerla como un juego entre dos jugadores: el adversario elige un margen de error ε alrededor de L y tú debes responder con un radio δ alrededor de a tal que todo lo que entre por ese radio salga dentro del margen. Si tienes respuesta para *todo* ε, el límite es L. La definición no habla de movimiento ni de infinito: habla de una estrategia ganadora.

La continuidad es el caso especial en que ese destino coincide con el valor, y su importancia real no es estética. Es la hipótesis que aparece en los tres teoremas grandes de esta sección —valor intermedio, Weierstrass, Bolzano—, y cada uno de ellos falla estrepitosamente sin ella. Y en la continuidad uniforme, más adelante, hay una diferencia que cuesta ver la primera vez y vale la pena forzar: si el δ sirve para *todos* los puntos a la vez o si hay que rehacerlo en cada punto.

Definición 6.33 (Límite)**.** $\lim_{x\to a} f(x) = L$ si para todo $\varepsilon > 0$ existe $\delta > 0$ tal que $0 < |x-a| < \delta \implies |f(x)-L| < \varepsilon$. Equivalentemente (caracterización por sucesiones): $f(x_n) \to L$ para *toda* sucesión $x_n \to a$ con $x_n \neq a$.

Existen las variantes laterales ($x \to a^{\pm}$), en el infinito ($x \to \pm\infty$) y con valor infinito. El límite existe si y solo si los dos laterales existen y coinciden.

Ejemplo 6.34 (Un ε–δ completo)**.** Probemos $\lim_{x\to 3} x^2 = 9$. Dado $\varepsilon > 0$, notamos $|x^2-9| = |x-3||x+3|$. Si exigimos de entrada $|x-3| < 1$, entonces $|x+3| < 7$. Tomando $\delta = \min\{1, \varepsilon/7\}$: $|x-3| < \delta \implies |x^2-9| < 7\cdot\frac{\varepsilon}{7} = \varepsilon$. El truco de "acotar primero el factor feo" es completamente general.

Dos límites fundamentales, que son la fuente de todas las derivadas trigonométricas y exponenciales:

$$\lim_{x\to 0} \frac{\sin x}{x} = 1, \qquad \lim_{x\to 0} \frac{e^x - 1}{x} = 1.$$

(El primero se demuestra comparando áreas en el círculo unitario: $\sin x \le x \le \tan x$ para $0 < x < \pi/2$, y aplicando el sándwich.)

Definición 6.35. f es *continua* en a si $\lim_{x\to a} f(x) = f(a)$. Sumas, productos, cocientes (donde el denominador no se anula) y composiciones de continuas son continuas. Polinomios, racionales, exp, ln, potencias y trigonométricas son continuas en sus dominios.

Idea clave

Qué dice esto en cristiano. Continua en a significa que el valor que la función *tiene* en a coincide con el valor al que *tiende* cerca de a: no hay sorpresa al llegar. La definición junta tres exigencias en una sola línea: que $f(a)$ exista, que el límite exista, y que sean iguales. Fallar cualquiera de las tres produce un tipo distinto de discontinuidad, y conviene saber distinguirlas porque el remedio es distinto en cada caso: una se arregla redefiniendo un punto, otra no se arregla de ninguna manera.

Ejemplo 6.36 (Las cuatro maneras de fallar)**. (a) Evitable: existe el límite pero no coincide.** La función $\frac{\sin x}{x}$ no está definida en 0, pero sus valores son 0,99833 para $x = 0{,}1$, luego 0,99998 para $x = 0{,}01$ y 0,9999998 para $x = 0{,}001$: el límite es 1. Basta definir $f(0) = 1$ y la función pasa a ser continua. Se llama evitable porque se evita con un solo retoque.

(b) De salto: los dos límites laterales existen pero difieren. La función signo vale -1 a la izquierda de 0 y $+1$ a la derecha. Ambos límites laterales existen y valen -1 y 1; como no coinciden, no hay límite y ninguna redefinición de $f(0)$ arregla nada. La parte entera $\lfloor x \rfloor$ tiene un salto así en cada entero.

(c) Esencial por explosión y por oscilación. En $\frac{1}{x}$ los valores se disparan al acercarse a 0: no hay límite finito. En $\sin\frac{1}{x}$ ocurre algo peor: en los puntos $x = \frac{1}{k\pi}$ la función vale 0, y en $x = \frac{2}{(4k+1)\pi}$ vale 1, para todo k tan grande como se quiera. Oscila infinitas veces en cualquier entorno del origen y no se acerca a nada.

(d) El truco del factor amortiguador. La función $x \sin \frac{1}{x}$ oscila igual, pero la amplitud se apaga: como $\left|x \sin \frac{1}{x}\right| \leq |x|$, el valor tiende a 0. Numéricamente da $-0{,}0544$ en $x = 0{,}1$, luego $-0{,}0051$ en $x = 0{,}01$ y $0{,}00083$ en $x = 0{,}001$. Definiendo $f(0) = 0$ queda continua. Compárese con (c): la misma oscilación, pero multiplicada por algo que se va a cero, deja de ser un problema.

Cuidado

[Continua en cada punto no es lo mismo que continua sin levantar el lápiz] La imagen del trazo continuo es útil pero engañosa. La función de Dirichlet, que vale 1 en los racionales y 0 en los irracionales, es discontinua en *todos* los puntos y no se puede dibujar de ninguna manera. Y $x \cdot D(x)$, con D la de Dirichlet, es continua exactamente en un punto, el 0, y en ningún otro: una función continua en un solo punto del mundo. Segundo aviso: que existan los dos límites laterales no garantiza continuidad, hace falta que coincidan *entre sí y con* $f(a)$, tres cosas y no dos.

Las discontinuidades se clasifican: *evitable* (el límite existe pero difiere de $f(a)$ o $f(a)$ no está definida), *de salto* (laterales distintos), *esencial* (algún lateral no existe, como $\sin(1/x)$ en 0).

Teorema 6.37 (Bolzano / valor intermedio)**.** *Si f es continua en $[a, b]$ y $f(a)f(b) < 0$, existe $c \in (a, b)$ con $f(c) = 0$. En consecuencia, f alcanza todo valor entre $f(a)$ y $f(b)$.*

Demostración. Sea $c = \sup\{x \in [a, b] : f(x) < 0\}$ (supongamos $f(a) < 0 < f(b)$; el conjunto es no vacío y acotado, el supremo existe por completitud). Si fuera $f(c) > 0$, por continuidad $f > 0$ en un entorno de c, y $c - \delta$ sería cota superior menor: contradicción. Si fuera $f(c) < 0$, habría puntos a la derecha de c con $f < 0$, contradiciendo que c es cota superior. Luego $f(c) = 0$. □

Idea clave

Qué dice esto en cristiano. Si una función continua empieza por debajo del cero y acaba por encima, en algún momento tuvo que pasar por el cero: no puede saltárselo. Es el teorema más intuitivo del cálculo y a la vez uno de los más útiles, porque demuestra que una ecuación tiene solución sin decir cuál, y porque su demostración se convierte directamente en un algoritmo, el método de bisección. La versión general, el teorema del valor intermedio, dice que una función continua en un intervalo toma todos los valores entre dos cualesquiera de los suyos.

Ejemplo 6.38 (De la existencia al cálculo)**. (a) Una raíz que no tiene fórmula sencilla.** Sea $g(x) = x^3 - x - 1$. Como $g(1) = -1 < 0$ y $g(2) = 5 > 0$, hay una raíz entre 1 y 2. Nótese que el teorema no la calcula: solo garantiza que está.

(b) La bisección, que sí la calcula. Partiendo de $[1, 2]$ y quedándose siempre con la mitad donde cambia el signo: [1; 1,5], luego [1,25; 1,5], luego [1,25; 1,375], y tras doce pasos

[1,324707; 1,324951]. El valor exacto es 1,3247180, el llamado número plástico. Cada paso gana un factor 2 de precisión, es decir unas tres cifras decimales cada diez pasos.

(c) Ecuaciones que mezclan familias. Para resolver $\cos x = x$ se estudia $h(x) = \cos x - x$: como $h(0) = 1 > 0$ y $h(1) = \cos 1 - 1 = -0{,}4597 < 0$, hay solución en $(0, 1)$, y resulta ser $x = 0{,}7390851$. Ninguna manipulación algebraica despeja esa x; la única vía es este argumento más un método numérico.

(d) Una consecuencia general. Todo polinomio de grado impar tiene al menos una raíz real, porque sus dos brazos van a lados opuestos del eje. Por ejemplo $x^5 + x + 1$ vale -1 en $x = -1$ y 1 en $x = 0$, luego hay raíz entre ambos; es $x = -0{,}7548777$. Con grado par esto puede fallar: $x^2 + 1$ no tiene ninguna raíz real.

Cuidado

[Contraejemplo: sin continuidad se salta el cero] La función $f(x) = \frac{1}{x}$ cumple $f(-1) = -1$ y $f(1) = 1$, pasa de negativo a positivo, y sin embargo *nunca* vale cero. No hay contradicción: no es continua en $[-1, 1]$, porque ni siquiera está definida en 0. Segundo aviso: el teorema garantiza *al menos* una raíz, no exactamente una; la función $\operatorname{sin} x$ en $[0, 10]$ cambia de signo y tiene tres. Y tercero: que $f(a)$ y $f(b)$ tengan el mismo signo no demuestra que no haya raíces; $x^2 - 1$ vale 3 en $x = -2$ y 3 en $x = 2$, y tiene dos raíces en medio. El teorema da una condición suficiente, nunca necesaria.

Teorema 6.39 (Weierstrass)**.** *Si f es continua en $[a, b]$, entonces está acotada y alcanza máximo y mínimo absolutos.*

Idea. Si no estuviera acotada, habría x_n con $|f(x_n)| > n$; por Bolzano–Weierstrass una subsucesión $x_{n_k} \to c \in [a, b]$, y la continuidad daría $f(x_{n_k}) \to f(c)$, finito: contradicción. Para el máximo se aplica el mismo argumento a una sucesión con $f(x_n) \to \sup f$. □

Idea clave

Qué dice esto en cristiano. Una función continua sobre un intervalo *cerrado y acotado* no puede escaparse a infinito ni acercarse a su valor extremo sin llegar a tocarlo: alcanza de verdad un punto donde vale lo máximo y otro donde vale lo mínimo. Las dos hipótesis son igual de imprescindibles, y la mejor manera de recordarlo es tener a mano los contraejemplos de lo que ocurre al quitar cada una. Este teorema es el que legitima todos los problemas de optimización: sin él, buscar el máximo de una función podría ser buscar algo que no existe.

Ejemplo 6.40 (Localizar los extremos y ver qué se rompe al quitar hipótesis)**. (a) El procedimiento completo.** Sea $f(x) = x^3 - 3x$ en $[-2, 3]$. Los puntos críticos salen de $f'(x) = 3x^2 - 3 = 0$, es decir $x = \pm 1$. Se evalúa en los críticos y en los extremos del intervalo: $f(-2) = -2$, $f(-1) = 2$, $f(1) = -2$ y $f(3) = 18$. El máximo absoluto es 18, en $x = 3$, y el mínimo absoluto es -2, alcanzado en dos puntos, $x = -2$ y $x = 1$.

(b) Quitar "cerrado". La función $f(x) = x$ en el intervalo abierto $(0, 1)$ es continua y acotada, pero no alcanza ni máximo ni mínimo: el supremo es 1 y el ínfimo es 0, y ninguno pertenece al intervalo. Basta abrir los extremos para perder la conclusión.

(c) Quitar "acotado" por un lado. La función $f(x) = \frac{1}{x}$ en $(0, 1]$ es continua en todo su dominio y sin embargo ni siquiera está acotada: vale 10 en $x = 0{,}1$, 100 en $x = 0{,}01$ y 1000 en $x = 0{,}001$. El intervalo no es cerrado en 0, y por ahí se escapa.

(d) Quitar la continuidad. Sea $f(x) = x$ en $[0, 1)$ y $f(1) = 0$. El dominio ahora sí es cerrado y acotado, la función está acotada, y aun así el supremo 1 no se alcanza: la discontinuidad en $x = 1$ le arrebata el máximo justo en el último instante. Con las tres condiciones, en cambio, no hay escapatoria: $\operatorname{sin} x$ en $[0, 2\pi]$ alcanza su máximo 1 en $\frac{\pi}{2}$ y su mínimo -1 en $\frac{3\pi}{2}$.

Cuidado

[Alcanzar el máximo no es lo mismo que tenerlo en un punto crítico] El máximo de una función continua en un intervalo cerrado puede estar en un extremo del intervalo, donde la derivada no tiene por qué anularse; en el apartado (a) el máximo 18 está en $x = 3$ y $f'(3) = 24 \neq 0$. Por eso el procedimiento correcto evalúa *siempre* en los extremos además de en los críticos. Segundo aviso: el teorema es de existencia pura, no dice dónde está el extremo ni ofrece manera de calcularlo; es la derivada la que lo localiza, y solo en los puntos interiores.

Definición 6.41 (Continuidad uniforme)**.** f es *uniformemente continua* en D si el δ puede elegirse independiente del punto: para todo $\varepsilon > 0$ existe $\delta > 0$ tal que $|x - y| < \delta \implies |f(x) - f(y)| < \varepsilon$ para *todos* $x, y \in D$. **Teorema de Heine**: toda función continua en un compacto $[a, b]$ es uniformemente continua. (En abiertos puede fallar: $f(x) = 1/x$ en $(0, 1)$.)

Idea clave

Qué dice esto en cristiano. En la continuidad de siempre, el δ puede depender del punto donde estés: cerca de una zona muy empinada necesitas acercarte mucho más para que la imagen no se mueva. La continuidad uniforme exige un δ único que sirva *en todo el dominio a la vez*. La diferencia se ve mirando dónde está el cuantificador: en la continuidad normal el punto se fija antes de buscar δ; en la uniforme, δ se busca antes de saber en qué punto se va a usar. Y esa diferencia es exactamente la que separa lo que se puede aproximar de manera controlada de lo que no.

Ejemplo 6.42 (Uniforme y no uniforme, con números)**. (a) Sobre un compacto siempre lo es.** En $[0, 3]$, la función x^2 cumple $|x^2 - y^2| = |x + y|\,|x - y| \leq 6|x - y|$, porque $|x + y| \leq 6$. Para $\varepsilon = 0{,}01$ basta $\delta = \frac{0{,}01}{6} = 0{,}0016667$, válido en todo el intervalo. Ese es el contenido del teorema de Heine: continua en un compacto implica uniformemente continua.

(b) La misma función, en toda la recta, deja de serlo. Tomando $x = n$ e $y = n + \frac{1}{n}$, la distancia entre ellos es $\frac{1}{n}$, que tiende a cero, pero $y^2 - x^2 = 2 + \frac{1}{n^2}$, que se queda por encima de 2: para $n = 10$ da 2,01, para $n = 100$ da 2,0001 y para $n = 1000$ da 2,000001. Por muy

pequeño que sea δ, siempre hay dos puntos a distancia menor que δ cuyas imágenes distan más de 2.

(c) La pendiente que explota. En $(0,1]$, la función $\frac{1}{x}$ falla por la misma razón en el otro extremo. Con $x = \frac{1}{n}$ e $y = \frac{1}{2n}$ la distancia es $\frac{1}{2n}$, que tiende a cero, mientras que $\frac{1}{y} - \frac{1}{x} = n$: vale 10, luego 100, luego 1000. La función se vuelve arbitrariamente empinada cerca del origen.

(d) Un caso que sorprende. La función $\sqrt{x}$ en $[0,\infty)$ *sí* es uniformemente continua, pese a que su derivada $\frac{1}{2\sqrt{x}}$ no está acotada cerca de cero. La razón es la desigualdad $\left|\sqrt{a} - \sqrt{b}\right| \le \sqrt{|a-b|}$: para $a = 0{,}0001$ y $b = 0$, la diferencia de imágenes es 0,01, que sigue siendo pequeña. Derivada acotada implica uniformemente continua, pero el recíproco es falso, y este es el contraejemplo.

Cuidado

[Dónde está el cuantificador] La continuidad ordinaria se escribe "para todo a, para todo ε, existe δ"; la uniforme, "para todo ε, existe δ, para todo a". Solo se ha movido un cuantificador de sitio y el significado cambia por completo. Segundo aviso: la uniformidad es una propiedad del *par* función–dominio, no de la función sola; x^2 es uniformemente continua en $[0,3]$ y no lo es en $\mathbb{R}$, y es exactamente la misma fórmula. Y tercero: el teorema de Heine exige compacidad; en un intervalo abierto o no acotado puede fallar, como muestran (b) y (c).

La continuidad uniforme es la hipótesis que hace funcionar la integrabilidad de las funciones continuas (sección 6.10).

6.9. La derivada

Idea clave

Intuición primero. Hay dos maneras de leer la derivada y conviene tener las dos a mano. La primera es la de las reglas de derivación e integración: pendiente de la tangente, ritmo de cambio instantáneo. La segunda es más profunda y es la que se generaliza: derivar es *aproximar por una recta*. Decir que f es derivable en a es decir que existe una función lineal que se pega tanto a f cerca de a que el error es despreciable incluso comparado con la distancia recorrida, es decir, $f(a+h) = f(a) + f'(a)h + o(h)$.
Esa segunda lectura es la que sobrevive al cambio de dimensión: en el capítulo 9 la derivada será una matriz, y en variedades será una aplicación lineal entre espacios tangentes, pero la idea seguirá siendo la misma —la mejor aproximación lineal— mientras que "la pendiente" habrá dejado de significar nada. Si te llevas una sola frase del cálculo diferencial, que sea esta: derivar es linealizar.

Definición 6.43. f es *derivable* en a si existe

$$f'(a) = \lim_{h\to 0} \frac{f(a+h) - f(a)}{h} \quad \Longleftrightarrow \quad f(a+h) = f(a) + f'(a)h + o(h).$$

Idea clave

Qué dice esto en cristiano. El cociente $\frac{f(a+h)-f(a)}{h}$ es la pendiente de la recta que une dos puntos de la gráfica: cuánto sube dividido entre cuánto avanza. La derivada es lo que ese número se acerca cuando los dos puntos se juntan, es decir la pendiente de la recta tangente. Que el límite *exista* significa que la curva tiene una dirección bien definida en ese punto; si al acercarse por la izquierda y por la derecha salen pendientes distintas, la curva tiene un pico y no hay derivada. Derivar es medir una tendencia instantánea, no un cambio real.

Ejemplo 6.44 (Calcular la derivada desde la definición, y ver cuándo no existe)**. (a) El cálculo directo.** Para $f(x) = x^2$ en $a = 3$:

$$\frac{(3+h)^2 - 9}{h} = \frac{9 + 6h + h^2 - 9}{h} = 6 + h \longrightarrow 6.$$

Numéricamente el cociente vale 6,1 para $h = 0{,}1$, luego 6,01 y luego 6,001. La recta tangente es $y = 9 + 6(x - 3)$; en $x = 3{,}1$ predice 9,6 frente al valor real 9,61, con un error de una centésima.

(b) Un pico. Para $f(x) = |x|$ en $a = 0$ el cociente vale $\frac{|h|}{h}$, que es $+1$ para todo $h > 0$ y -1 para todo $h < 0$. Los dos límites laterales existen y son distintos, luego no hay derivada. La función es continua ahí: continuidad no implica derivabilidad, y este es el contraejemplo estándar.

(c) Una tangente vertical. Para $f(x) = x^{2/3}$ en $a = 0$ el cociente es $h^{-1/3}$, que vale 2,15 para $h = 0{,}1$, luego 10 para $h = 0{,}001$ y 100 para $h = 10^{-6}$: se dispara. No hay derivada finita, aunque la gráfica sí tiene una tangente, vertical.

(d) Derivable con derivada discontinua. Sea $f(x) = x^2 \sin\frac{1}{x}$ y $f(0) = 0$. El cociente en 0 es $h \sin\frac{1}{h}$, que vale $-0{,}0544$; $-0{,}0051$; $0{,}00083$ para $h = 0{,}1$; $0{,}01$; $0{,}001$, y tiende a 0: la función *es* derivable en 0 con $f'(0) = 0$. Pero $f'(x) = 2x \sin\frac{1}{x} - \cos\frac{1}{x}$ para $x \neq 0$ oscila sin límite. Ser derivable en todas partes no obliga a que la derivada sea continua.

Cuidado

[Los dos errores de lectura] El cociente incremental no es una fracción que se pueda simplificar cancelando: el símbolo $\frac{\mathrm{d}y}{\mathrm{d}x}$ es una notación heredada de Leibniz, no un cociente de dos números. Segundo aviso, y es el importante: derivable implica continua, pero *no* al revés. El apartado (b) da una función continua sin derivada en un punto, y Weierstrass construyó en 1872 una función continua en toda la recta y derivable en *ningún* punto, lo que en su momento se consideró casi indecente. La intuición de que "una curva continua tiene tangente casi siempre" es sencillamente falsa.

Proposición 6.45. *Derivable $\Longrightarrow$ continua. (De $f(a+h) - f(a) = h \cdot \frac{f(a+h)-f(a)}{h} \to 0 \cdot f'(a) = 0$.) El recíproco falla: $|x|$ es continua y no derivable en 0; existen incluso funciones continuas en todo $\mathbb{R}$ sin derivada en ningún punto (Weierstrass).*

6.9.1. Las reglas, con demostración

Teorema 6.46 (Reglas de derivación)**.** *Si f, g son derivables en x:*

$$(f+g)' = f' + g', \qquad (cf)' = cf',$$
$$(fg)' = f'g + fg', \qquad \left(\frac{f}{g}\right)' = \frac{f'g - fg'}{g^2} \ (g \neq 0).$$

Demostración de la regla del producto.

$$\frac{f(x+h)g(x+h) - f(x)g(x)}{h} = \frac{f(x+h)-f(x)}{h}\, g(x+h) + f(x)\, \frac{g(x+h)-g(x)}{h}.$$

(Se sumó y restó $f(x)g(x+h)$.) Al tomar $h \to 0$, el primer sumando tiende a $f'(x)g(x)$ (usando que g es continua) y el segundo a $f(x)g'(x)$. □

Teorema 6.47 (Regla de la cadena)**.** *Si f es derivable en a y g en $f(a)$: $(g \circ f)'(a) = g'(f(a))f'(a)$.*

Idea. Escribimos los incrementos en forma lineal: $f(a+h) = f(a) + [f'(a) + \alpha(h)]h$ y $g(y+k) = g(y) + [g'(y) + \beta(k)]k$ con $\alpha, \beta \to 0$. Componiendo, el incremento de $g \circ f$ es $[g'(f(a)) + \beta][f'(a) + \alpha]h$, y el corchete tiende a $g'(f(a))f'(a)$. (Este truco evita el problema de dividir por un incremento que podría anularse.) □

Idea clave

Qué dice esto en cristiano. Cuando una cantidad depende de otra que a su vez depende de una tercera, los ritmos de cambio se multiplican. Si un engranaje gira tres veces más rápido que otro, y ese otro cinco veces más rápido que un tercero, el primero gira quince veces más rápido que el tercero. La regla de la cadena es exactamente eso: $(g \circ f)' = g'(f(a)) \cdot f'(a)$, la derivada de fuera evaluada donde toca, por la derivada de dentro. El único detalle que hay que cuidar es *dónde* se evalúa cada factor, y es justo ahí donde se cometen todos los errores.

Ejemplo 6.48 (La cadena en cuatro niveles de dificultad)**. (a) Un caso limpio.** Para $\sin(x^2)$, la función de fuera es el seno y la de dentro es x^2, luego la derivada es $\cos(x^2) \cdot 2x$. En $x = 1$ vale $2\cos 1 = 1{,}0806046$. Comprobación numérica con incremento 10^{-6}: el cociente da 1,0806035, coherente hasta la sexta cifra.

(b) Potencias de una expresión. Para $(3x+1)^5$, la de fuera es la quinta potencia y la de dentro $3x+1$: sale $5(3x+1)^4 \cdot 3 = 15(3x+1)^4$. En $x = 0$ vale 15. Obsérvese que desarrollar el binomio primero daría el mismo resultado con muchísimo más trabajo.

(c) Exponencial compuesta. Para $\mathrm{e}^{\sin x}$ sale $\mathrm{e}^{\sin x}\cos x$; en $x = 0$ vale $\mathrm{e}^0 \cdot 1 = 1$. La derivada de fuera se evalúa en $\sin x$, no en x: escribir $\mathrm{e}^x \cos x$ sería el error típico.

(d) Tres eslabones. Para $\sin(\cos(x^2))$ se encadena tres veces: $\cos(\cos(x^2)) \cdot (-\sin(x^2)) \cdot 2x$. En $x = 1$ vale $-1{,}4432123$. La regla se aplica de fuera hacia dentro y cada factor nuevo se multiplica; no hay límite al número de eslabones.

Cuidado

[Evaluar cada factor donde corresponde] El factor g' se evalúa en $f(a)$, no en a. Para $\sin(x^2)$ en $x=2$, la derivada es $2\cdot 2\cdot\cos(4)$ y no $4\cos 2$: la diferencia entre $\cos 4 = -0{,}6536$ y $\cos 2 = -0{,}4161$ no es pequeña. Segundo aviso: hay que identificar bien quién es la función de dentro. En $(\sin x)^2$ la de dentro es $\sin x$ y sale $2\sin x\cos x$; en $\sin(x^2)$ la de dentro es x^2 y sale $2x\cos(x^2)$. Son funciones distintas con notaciones que solo difieren en la posición de un paréntesis, y confundirlas es el error más frecuente de todo el cálculo diferencial.

Teorema 6.49 (Derivada de la inversa)**.** *Si f es biyectiva, derivable en a con $f'(a)\neq 0$, entonces f^{-1} es derivable en $b=f(a)$ y*

$$(f^{-1})'(b)=\frac{1}{f'(f^{-1}(b))}.$$

De aquí salen las derivadas de ln, arcsin, arctan, etc. La tabla completa de derivadas que hay que dominar:

$$\begin{array}{llll}
(x^n)'=nx^{n-1} & (e^x)'=e^x & (a^x)'=a^x\ln a & (\ln x)'=1/x\\
(\sin x)'=\cos x & (\cos x)'=-\sin x & (\tan x)'=\sec^2 x & \\
(\arcsin x)'=\frac{1}{\sqrt{1-x^2}} & (\arc\cos x)'=\frac{-1}{\sqrt{1-x^2}} & (\arctan x)'=\frac{1}{1+x^2} & \\
(\sinh x)'=\cosh x & (\cosh x)'=\sinh x. & &
\end{array}$$

6.9.2. Los teoremas del valor medio

Teorema 6.50 (Fermat)**.** *Si f tiene un extremo local en un punto interior c y es derivable ahí, entonces $f'(c)=0$.*

Demostración. Si es máximo: para $h>0$, $\frac{f(c+h)-f(c)}{h}\leq 0$; para $h<0$, ≥ 0. El límite debe ser a la vez ≤ 0 y ≥ 0. □

Teorema 6.51 (Rolle)**.** *Si f es continua en $[a,b]$, derivable en (a,b) y $f(a)=f(b)$, existe $c\in(a,b)$ con $f'(c)=0$.*

Demostración. Por Weierstrass, f alcanza máximo y mínimo. Si ambos están en los extremos del intervalo, f es constante y cualquier c sirve. Si no, hay un extremo interior, y Fermat concluye. □

Teorema 6.52 (Valor medio de Lagrange)**.** *Si f es continua en $[a,b]$ y derivable en (a,b), existe $c\in(a,b)$ con*

$$f(b)-f(a)=f'(c)(b-a).$$

Versión de Cauchy: si además g cumple lo mismo y $g'\neq 0$ en (a,b), $\dfrac{f(b)-f(a)}{g(b)-g(a)}=\dfrac{f'(c)}{g'(c)}$.

Demostración. Aplicar Rolle a $h(x) = f(x) - f(a) - \frac{f(b)-f(a)}{b-a}(x-a)$ (y, para Cauchy, a $h(x) = [f(b)-f(a)]g(x) - [g(b)-g(a)]f(x)$). □

Idea clave

Qué dice esto en cristiano. Los tres teoremas son el mismo, en dificultad creciente. Fermat dice que en un pico o un valle interior la tangente es horizontal, porque si tuviera pendiente se podría subir más siguiendo la cuesta. Rolle añade que si sales y llegas a la misma altura, en algún punto intermedio la tangente fue horizontal: para volver hay que dar la vuelta. Y el teorema del valor medio es Rolle inclinado: en algún punto la tangente es paralela a la cuerda que une los extremos. Traducido a velocidades: si has recorrido 180 kilómetros en dos horas, en algún instante el velocímetro marcó exactamente 90.

Ejemplo 6.53 (Localizar el punto que los teoremas prometen)**. (a) Rolle con números.** Sea $f(x) = x^2 - 4x + 3$ en $[1,3]$. Se cumple $f(1) = 0 = f(3)$, luego Rolle garantiza un c interior con $f'(c) = 0$. Como $f'(x) = 2x - 4$, ese punto es $c = 2$, el vértice de la parábola. La hipótesis de igualdad en los extremos se cumple porque 1 y 3 son precisamente las dos raíces.

(b) El valor medio en una cúbica. Para $f(x) = x^3$ en $[0,2]$, la pendiente de la cuerda es $\frac{8-0}{2} = 4$. Buscando $3c^2 = 4$ sale $c = \frac{2}{\sqrt{3}} = 1{,}1547005$, que está dentro de $(0,2)$ como debe. Comprobación: $3 \cdot 1{,}1547005^2 = 4$ exactamente.

(c) Otro, con una raíz. Para $f(x) = \sqrt{x}$ en $[1,4]$, la pendiente de la cuerda es $\frac{2-1}{3} = 0{,}3333333$. Resolviendo $\frac{1}{2\sqrt{c}} = \frac{1}{3}$ sale $c = 2{,}25$, y en efecto $\frac{1}{2\sqrt{2{,}25}} = \frac{1}{3}$. El punto no es el punto medio del intervalo: el teorema no dice dónde está, solo que existe.

(d) Para qué sirve de verdad. El teorema del valor medio es la herramienta que convierte información sobre f' en información sobre f. De él sale que si $f' = 0$ en todo un intervalo entonces f es constante, que si $f' > 0$ entonces f crece, y desigualdades como $|\sin a - \sin b| \le |a - b|$, ya que $|\cos c| \le 1$; para $a = 1$ y $b = 1{,}5$ da $0{,}1560240 \le 0{,}5$. Casi todo lo que se afirma sobre funciones a partir de su derivada pasa por aquí.

Cuidado

[Contraejemplos: cada hipótesis hace falta] Sin derivabilidad, Rolle falla: $f(x) = |x|$ en $[-1,1]$ cumple $f(-1) = f(1) = 1$ y su derivada no se anula en ningún punto, porque en 0 ni siquiera existe. Sin continuidad en el intervalo cerrado también falla: $f(x) = x$ en $[0,1)$ con $f(1) = 0$ cumple $f(0) = f(1) = 0$ y su derivada vale 1 en todas partes. Segundo aviso: Fermat exige que el extremo sea *interior*; en $f(x) = x$ sobre $[0,1]$ el máximo está en $x = 1$ y ahí $f' = 1 \neq 0$. Y tercero: el recíproco de Fermat es falso, es decir $f'(c) = 0$ no implica extremo: en $f(x) = x^3$ la derivada se anula en 0 y la función es estrictamente creciente.

Corolario 6.54. *En un intervalo:* $f' > 0 \implies f$ *estrictamente creciente;* $f' \ge 0 \iff f$ *creciente;* $f' = 0 \implies f$ *constante;* $|f'| \le M \implies f$ *es* Lipschitz: $|f(x) - f(y)| \le M|x - y|$.

Teorema 6.55 (L'Hôpital)**.** *Sean* f, g *derivables cerca de* a *con* $g' \neq 0$*, y supongamos*

$f(x), g(x) \to 0$ (o ambos $\to \pm\infty$). Si $\lim_{x\to a} \frac{f'(x)}{g'(x)} = L$ (finito o infinito), entonces $\lim_{x\to a} \frac{f(x)}{g(x)} = L$.

Idea. Caso 0/0: por el TVM de Cauchy, $\frac{f(x)}{g(x)} = \frac{f(x)-f(a)}{g(x)-g(a)} = \frac{f'(\xi)}{g'(\xi)}$ con ξ entre a y x; cuando $x \to a$, también $\xi \to a$. □

Idea clave

Qué dice esto en cristiano. Cuando dos funciones se van juntas a cero (o juntas a infinito), su cociente puede tender a cualquier cosa, según cuál de las dos gane la carrera. La regla de L'Hôpital dice que la carrera la deciden las *velocidades*: se comparan las derivadas en vez de los valores. La justificación intuitiva es que, cerca del punto, cada función se parece a su recta tangente, y el cociente de dos rectas que pasan por el origen es el cociente de sus pendientes. Es una herramienta potentísima y también la más maltratada del cálculo, porque se aplica constantemente donde no toca.

Ejemplo 6.56 (Cuatro indeterminaciones resueltas)**. (a) El caso clásico.** $\lim_{x\to 0} \frac{\sin x}{x}$ es de la forma $\frac{0}{0}$; derivando arriba y abajo queda $\frac{\cos x}{1} \to 1$. Curiosamente es un uso circular si la derivada del seno se dedujo de este mismo límite, así que como demostración no vale, pero como comprobación es perfecta.

(b) Aplicar la regla dos veces. $\lim_{x\to 0} \frac{e^x-1-x}{x^2}$ es $\frac{0}{0}$; derivando queda $\frac{e^x-1}{2x}$, que sigue siendo $\frac{0}{0}$; derivando otra vez, $\frac{e^x}{2} \to \frac{1}{2}$. Numéricamente el cociente original vale 0,5170918 para $x = 0{,}1$, luego 0,5016708 y 0,5001667: se acerca a 0,5. Igual ocurre con $\frac{1-\cos x}{x^2}$, que también da $\frac{1}{2}$.

(c) La forma ∞/∞. $\lim_{x\to\infty} \frac{\ln x}{x}$ es $\frac{\infty}{\infty}$; derivando queda $\frac{1/x}{1} = \frac{1}{x} \to 0$. Es la manera rápida de ver que el logaritmo crece más despacio que cualquier potencia.

(d) Formas que hay que transformar primero. $\lim_{x\to 0^+} x \ln x$ es del tipo $0 \cdot \infty$, y la regla no se aplica directamente: hay que reescribirlo como $\frac{\ln x}{1/x}$, que ya es $\frac{-\infty}{\infty}$, y entonces sale $\frac{1/x}{-1/x^2} = -x \to 0$. Numéricamente da $-0{,}2302585$ en $x = 0{,}1$, $-0{,}0460517$ en $x = 0{,}01$ y $-0{,}0009210$ en $x = 0{,}0001$.

Cuidado

[Los dos abusos que invalidan el resultado] Primero: la regla solo se aplica a indeterminaciones. Para $\lim_{x\to 1} \frac{x^2+1}{2x^2-3}$ el valor es sencillamente $\frac{2}{-1} = -2$, porque no hay indeterminación; derivar arriba y abajo daría $\frac{2x}{4x} = \frac{1}{2}$, que es falso. Antes de derivar hay que *comprobar* que numerador y denominador tienden ambos a cero o ambos a infinito. Segundo: la regla dice que si el límite de las derivadas existe, entonces el original vale lo mismo; *no* dice lo contrario. Para $\lim_{x\to\infty} \frac{x+\sin x}{x}$ el valor es 1 (los cocientes valen 0,9456, 0,9949 y 1,0008 para $x = 10, 100, 1000$), pero el cociente de derivadas es $1 + \cos x$, que oscila para siempre y no tiene límite. Que L'Hôpital no concluya no significa que no haya límite.

Cuidado

L'Hôpital solo aplica a indeterminaciones $\frac{0}{0}$ o $\frac{\infty}{\infty}$; usarlo en otro caso da respuestas falsas. Y si $\lim f'/g'$ no existe, no se puede concluir nada ($\frac{x+\sin x}{x} \to 1$ aunque $\frac{1+\cos x}{1}$ oscile). Ante la duda, Taylor es más seguro y más informativo.

6.9.3. Convexidad y trazado de gráficas

Definición 6.57. f es *convexa* en un intervalo si el segmento entre dos puntos cualesquiera de su gráfica queda por encima de la gráfica: $f(\lambda x + (1-\lambda)y) \le \lambda f(x) + (1-\lambda)f(y)$. Si f'' existe: f convexa $\iff$ $f'' \ge 0$. Un *punto de inflexión* separa tramos de convexidad opuesta.

Idea clave

Qué dice esto en cristiano. Convexa significa "con forma de cuenco": si tiendes una cuerda entre dos puntos de la gráfica, la cuerda queda por encima de la curva. Cóncava es lo contrario, forma de campana. Y aquí está el criterio práctico que se usa siempre: la segunda derivada mide la curvatura del dibujo, de modo que $f'' > 0$ es cuenco y $f'' < 0$ es campana. Un punto donde f'' cambia de signo es un punto de inflexión, el instante en que la curva deja de curvarse hacia un lado y empieza a hacerlo hacia el otro.

Ejemplo 6.58 (Reconocer la forma con la segunda derivada)**. (a) El cuenco por excelencia.** Para $f(x) = x^2$ se tiene $f'' = 2 > 0$ siempre: convexa en toda la recta. Comprobación con la cuerda entre $x = 1$ y $x = 3$: el punto medio de la cuerda está a altura $\frac{1+9}{2} = 5$, mientras que la función en $x = 2$ vale 4. La cuerda queda por encima, como debe.

(b) La campana. Para $f(x) = \ln x$ se tiene $f'' = -\frac{1}{x^2} < 0$: cóncava. Entre $x = 1$ y $x = 4$, el punto medio de la cuerda está a altura $\frac{0+\ln 4}{2} = 0{,}6931472$, mientras que $\ln 2{,}5 = 0{,}9162907$. Aquí la curva queda por encima de la cuerda, justo al revés.

(c) Un punto de inflexión. Para $f(x) = x^3$ se tiene $f'' = 6x$, negativa a la izquierda de 0 y positiva a la derecha: cóncava y luego convexa, con inflexión en el origen. Nótese que $f'(0) = 0$ y sin embargo no hay extremo, precisamente porque la curvatura cambia ahí.

(d) Una desigualdad que sale gratis. La convexidad da la desigualdad de Jensen: para f convexa, $f\left(\frac{a+b}{2}\right) \le \frac{f(a)+f(b)}{2}$. Con $f(x) = x^2$, $a = 1$ y $b = 3$: $4 \le 5$. Aplicándola al logaritmo, que es cóncavo, sale la desigualdad entre medias: $\ln \frac{4+9}{2} = 1{,}8718022 \ge \frac{\ln 4 + \ln 9}{2} = 1{,}7917595$, y exponenciando queda $6{,}5 \ge 6$, es decir la media aritmética por encima de la geométrica.

Cuidado

[Convexa no significa creciente, ni la inflexión exige $f'' = 0$] Son propiedades independientes: $f(x) = x^2$ es convexa y decrece en $(-\infty, 0)$, mientras que $\ln x$ es creciente y cóncava. La primera derivada habla de subir o bajar; la segunda, de la forma. Segundo aviso: que $f''(c) = 0$ no basta para que haya inflexión, hace falta que *cambie de signo*. En $f(x) = x^4$ se tiene $f''(0) = 0$ y sin embargo la función es convexa a ambos lados: no hay inflexión, hay un mínimo. Y tercero: la convención sobre los nombres varía, y

algunos textos llaman "cóncava hacia arriba" a lo que aquí es convexa; conviene fijarse en el signo de f'' y no en la palabra.

Para clasificar puntos críticos: si $f'(c) = 0$ y $f''(c) > 0$, mínimo local; $f''(c) < 0$, máximo local; $f''(c) = 0$, hay que mirar derivadas superiores (la primera no nula: orden par $\Rightarrow$ extremo, orden impar $\Rightarrow$ inflexión con tangente horizontal, como x^3 en 0).

La convexidad da desigualdades importantes: de la convexidad de e^x y la concavidad de ln salen $e^x \geq 1 + x$, $\ln x \leq x - 1$, y la desigualdad entre las medias aritmética y geométrica:

$$\sqrt[n]{a_1 \cdots a_n} \leq \frac{a_1 + \cdots + a_n}{n}.$$

6.9.4. El teorema de Taylor

Idea clave

Intuición primero. Las derivadas en un punto son el ADN local de la función: posición, velocidad, aceleración, tirón… El polinomio de Taylor es el *mejor impostor* de grado n: la única función polinómica con exactamente el mismo ADN hasta el orden n. El resto mide qué tan lejos del punto sigue funcionando la impostura, y las tres fórmulas del resto son tres maneras de facturar el primer rasgo genético que el impostor *no* copió: la derivada $n{+}1$. Por eso todas llevan $f^{(n+1)}$: el error de una copia perfecta hasta orden n solo puede venir del orden $n + 1$.

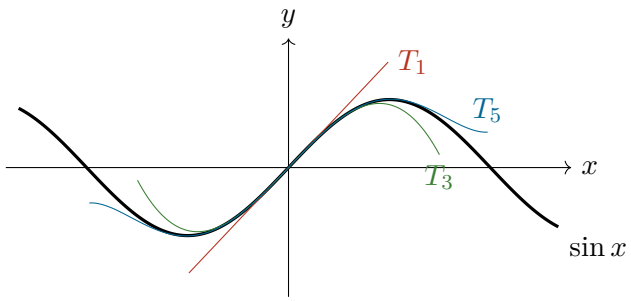


Figura 6.1: Los impostores de Taylor de $\sin x$: cada grado extra compra un tramo más de fidelidad alrededor del origen; lejos del punto, todos acaban traicionando.

Teorema 6.59 (Taylor con resto de Lagrange)**.** *Si f es $n+1$ veces derivable en un entorno de a, para cada x existe ξ entre a y x con*

$$f(x) = \underbrace{\sum_{k=0}^{n} \frac{f^{(k)}(a)}{k!}(x-a)^k}_{P_n(x)} + \underbrace{\frac{f^{(n+1)}(\xi)}{(n+1)!}(x-a)^{n+1}}_{R_n(x)}.$$

Idea. Se fija x y se define $g(t) = f(x) - \sum_{k\leq n} \frac{f^{(k)}(t)}{k!}(x-t)^k - M(x-t)^{n+1}$, eligiendo M para que $g(a) = 0$. Como también $g(x) = 0$, Rolle da ξ con $g'(\xi) = 0$; al derivar, la suma es telescópica y la condición $g'(\xi) = 0$ despeja exactamente $M = \frac{f^{(n+1)}(\xi)}{(n+1)!}$. □

Ejemplo 6.60 (Estimación de errores)**.** ¿Cuántos términos de la serie de e hacen falta para 6 decimales? Con $a=0$, $x=1$: $R_n = \frac{e^\xi}{(n+1)!} \le \frac{3}{(n+1)!}$. Para $n=9$: $\frac{3}{10!} \approx 8\times 10^{-7} < 10^{-6}$. Nueve términos bastan. Así se construyen las tablas y algoritmos numéricos: Taylor *con control del resto*.

Los desarrollos de memoria, centrados en 0:

$$e^x = \sum_{n\ge 0} \frac{x^n}{n!} \qquad \sin x = \sum_{n\ge 0} \frac{(-1)^n x^{2n+1}}{(2n+1)!} \qquad \cos x = \sum_{n\ge 0} \frac{(-1)^n x^{2n}}{(2n)!}$$

$$\frac{1}{1-x} = \sum_{n\ge 0} x^n \ (|x|<1) \qquad \ln(1+x) = \sum_{n\ge 1} \frac{(-1)^{n-1} x^n}{n} \qquad \arctan x = \sum_{n\ge 0} \frac{(-1)^n x^{2n+1}}{2n+1}$$

$$(1+x)^\alpha = \textstyle\sum_{n\ge 0} \binom{\alpha}{n} x^n$$

donde $\binom{\alpha}{n} = \frac{\alpha(\alpha-1)\cdots(\alpha-n+1)}{n!}$ (serie binomial, válida para $|x|<1$).

6.10. La integral de Riemann

Idea clave

Intuición primero. La construcción de Riemann responde a "¿cómo se define el área bajo una curva sin dar por sentado qué es el área?". La estrategia es la de Arquímedes: aproximar por arriba y por abajo con figuras que sí sabemos medir —rectángulos— y llamar área al número al que ambas aproximaciones convergen, *si* convergen al mismo. Las sumas inferiores siempre se quedan cortas, las superiores siempre se pasan, y la función es integrable exactamente cuando el hueco entre ambas se puede hacer tan pequeño como se quiera afinando la partición.
Esa definición tiene una consecuencia que conviene ver desde el principio: la integrabilidad no depende de que la función sea bonita, sino de que sus saltos no ocupen demasiado. Una función con infinitas discontinuidades puede ser integrable, y la función de Dirichlet no lo es en absoluto. Ahí está el límite del método de Riemann, y ese límite es exactamente lo que motivará la integral de Lebesgue en el capítulo 31.

Definición 6.61. Sea f acotada en $[a,b]$ y $P=\{a=x_0<x_1<\cdots<x_n=b\}$ una partición. Con $m_i = \inf_{[x_{i-1},x_i]} f$ y $M_i = \sup_{[x_{i-1},x_i]} f$, se definen las sumas inferior y superior

$$L(f,P) = \sum m_i\,\Delta x_i, \qquad U(f,P) = \sum M_i\,\Delta x_i.$$

f es *integrable* si $\sup_P L(f,P) = \inf_P U(f,P)$; ese valor común es $\int_a^b f$.

Idea clave

Qué dice esto en cristiano. Para medir el área bajo una curva, Riemann propone encerrarla entre dos aproximaciones: rectángulos que se quedan cortos (tomando el ínfimo en cada trozo) y rectángulos que se pasan (tomando el supremo). La suma inferior

siempre subestima y la superior siempre sobreestima, de modo que el área verdadera está atrapada entre ambas. Refinando la partición, la inferior sube y la superior baja. La función es integrable cuando ambas se encuentran en un único número, y ese número es la integral.

Ejemplo 6.62 (Atrapar el área entre dos sumas)**. (a) La parábola, con particiones cada vez más finas.** Para $f(x) = x^2$ en $[0, 1]$ con n subintervalos iguales:

n	$L(f, P)$	$U(f, P)$	$U - L$
2	0,125000	0,625000	0,500000
4	0,218750	0,468750	0,250000
10	0,285000	0,385000	0,100000
100	0,328350	0,338350	0,010000
1000	0,332834	0,333834	0,001000

(b) La lectura de la tabla. Las inferiores suben, las superiores bajan y ambas se aproximan a $\frac{1}{3} = 0{,}3333333$. La diferencia es exactamente $\frac{1}{n}$, luego se puede hacer tan pequeña como se quiera eligiendo n grande: eso es precisamente lo que exige el criterio de integrabilidad, y por tanto x^2 es integrable con $\int_0^1 x^2 \, \mathrm{d}x = \frac{1}{3}$.

(c) Por qué $U - L$ es el objeto correcto. No hace falta conocer el valor de la integral para decidir si existe; basta con que la horquilla se cierre. Ese es todo el contenido del criterio, y por eso se demuestra con él que toda función continua en un compacto es integrable: la continuidad uniforme permite hacer $U - L$ pequeño de golpe en todos los subintervalos.

(d) Una función con salto, integrable. Sea $f = 0$ en $[0, 1)$ y $f = 1$ en $[1, 2]$. No es continua, pero es monótona, y toda función monótona es integrable. La horquilla se cierra porque la discontinuidad está en un solo punto, que se puede encerrar en un subintervalo de longitud despreciable. La integral vale 1.

Cuidado

[Contraejemplo: la función que no se deja atrapar] La función de Dirichlet, que vale 1 en los racionales y 0 en los irracionales, no es integrable en el sentido de Riemann sobre $[0, 1]$. En cualquier subintervalo, por diminuto que sea, hay racionales e irracionales, así que el ínfimo es 0 y el supremo es 1 *siempre*. Por tanto $L(f, P) = 0$ y $U(f, P) = 1$ para toda partición, y la horquilla nunca se cierra: la diferencia se queda clavada en 1. Es el ejemplo que motivó la integral de Lebesgue, donde esta función sí es integrable y su integral vale 0.

Teorema 6.63 (Criterio de integrabilidad)**.** *f es integrable $\iff$ para todo $\varepsilon > 0$ existe una partición con $U(f, P) - L(f, P) < \varepsilon$.*

Teorema 6.64. *Toda función continua en $[a, b]$ es integrable; también toda función monótona, y toda función acotada con una cantidad finita de discontinuidades.*

Demostración para continuas. Por Heine, f es uniformemente continua: dado ε, hay δ tal que $|x-y|<\delta \implies |f(x)-f(y)|<\frac{\varepsilon}{b-a}$. En cualquier partición con celdas de ancho $<\delta$, $M_i-m_i<\frac{\varepsilon}{b-a}$ y entonces $U-L<\varepsilon$. □

Idea clave

Qué dice esto en cristiano. El criterio libera de tener que calcular la integral para saber si existe: basta con comprobar que la horquilla entre la suma superior y la inferior se puede estrechar tanto como se quiera. Es la versión práctica de la definición, porque no obliga a conocer el número final. Y el segundo teorema aplica el criterio a las tres familias que cubren casi todo lo que uno se encuentra: las continuas, las monótonas y las que tienen finitas discontinuidades.

Ejemplo 6.65 (Aplicar el criterio a las tres familias)**. (a) Continuas.** Si f es continua en $[a,b]$, es uniformemente continua por el teorema de Heine, luego dado ε existe δ que sirve en todo el intervalo. Tomando una partición de paso menor que δ, en cada trozo se cumple $M_i-m_i<\frac{\varepsilon}{b-a}$, y sumando queda $U-L<\varepsilon$. La uniformidad es la que permite usar el mismo δ en todos los subintervalos a la vez.

(b) Monótonas, aunque salten mucho. Si f crece en $[a,b]$, con la partición uniforme de paso h se tiene $M_i=f(x_i)$ y $m_i=f(x_{i-1})$, de modo que la suma telescopa: $U-L=h\,(f(b)-f(a))$. Basta tomar h pequeño. Una función monótona puede tener infinitas discontinuidades (una en cada racional, por ejemplo) y aun así es integrable.

(c) Con finitas discontinuidades. La función $f=0$ en $[0,1)$ y $f=1$ en $[1,2]$ tiene un salto. Se encierra ese punto en un subintervalo de longitud η; ahí la contribución a $U-L$ es a lo sumo $\eta\cdot 1$, y fuera la función es continua. Eligiendo $\eta<\varepsilon$ queda $U-L<\varepsilon$. La integral vale 1.

(d) La medida cero, que es el criterio definitivo. El teorema de Lebesgue afirma que f acotada es integrable Riemann exactamente cuando su conjunto de discontinuidades tiene medida nula. Las tres familias anteriores son casos particulares: las continuas no tienen ninguna, las monótonas tienen a lo sumo una cantidad numerable, y las de (c) tienen finitas. Y la de Dirichlet falla porque es discontinua en *todos* los puntos.

Cuidado

[Acotada es una hipótesis, no un detalle] Todo lo anterior exige que f esté acotada, porque de lo contrario el supremo en algún subintervalo es infinito y las sumas de Darboux ni siquiera están definidas. La función $\frac{1}{\sqrt{x}}$ en $(0,1]$ tiene integral finita, 2, pero no es integrable Riemann en el sentido propio: es una *integral impropia*, definida como límite. Segundo aviso: continua implica integrable, pero no al revés; el apartado (c) es integrable sin ser continua. Y tercero: cambiar el valor de una función en finitos puntos no altera su integral, así que dos funciones distintas pueden tener exactamente la misma.

La función de Dirichlet (1 en los racionales, 0 en los irracionales) *no* es integrable-Riemann: toda suma inferior es 0 y toda superior es $b-a$. (Su tratamiento requiere la integral de Lebesgue.)

Propiedades: linealidad, monotonía ($f \le g \implies \int f \le \int g$), aditividad en el intervalo, y $|\int f| \le \int |f|$. Además el *teorema del valor medio integral*: si f es continua, existe c con $\int_a^b f = f(c)(b-a)$.

Teorema 6.66 (Fundamental del cálculo)**.** *Sea f integrable en $[a,b]$ y $F(x) = \int_a^x f$.*

(i) F es continua; y en cada punto donde f es continua, $F'(x) = f(x)$.

(ii) (Barrow) Si f es continua y $G' = f$, entonces $\int_a^b f = G(b) - G(a)$.

Demostración de (i) en puntos de continuidad.

$$\frac{F(x+h)-F(x)}{h} - f(x) = \frac{1}{h}\int_x^{x+h} \big(f(t)-f(x)\big)\,\mathrm{d}t,$$

y si $|f(t)-f(x)| < \varepsilon$ para $|t-x| < \delta$, el miembro derecho está acotado por ε cuando $|h| < \delta$. □

6.10.1. Técnicas de integración

Por partes ($\int u\,\mathrm{d}v = uv - \int v\,\mathrm{d}u$): para productos de polinomio con exponencial/trigonométrica (derivar el polinomio), para ln y arcos (derivarlos), y las cíclicas como $\int e^x \cos x\,\mathrm{d}x$ (dos pasadas y despejar).

Sustitución: $\int f(g(x))g'(x)\mathrm{d}x = \int f(u)\mathrm{d}u$. Las sustituciones clásicas:

integrando contiene	sustituir
$\sqrt{a^2-x^2}$	$x = a\sin\theta$
$\sqrt{a^2+x^2}$	$x = a\tan\theta$
$\sqrt{x^2-a^2}$	$x = a\sec\theta$
racional en $\sin,\cos$	$t = \tan(x/2)$: $\sin x = \frac{2t}{1+t^2}$, $\cos x = \frac{1-t^2}{1+t^2}$, $\mathrm{d}x = \frac{2\,\mathrm{d}t}{1+t^2}$
racional en e^x	$u = e^x$

Fracciones parciales: toda función racional $\frac{P}{Q}$ con $\deg P < \deg Q$ se descompone según los factores de Q: cada factor $(x-a)^m$ aporta $\frac{A_1}{x-a} + \cdots + \frac{A_m}{(x-a)^m}$; cada factor cuadrático irreducible $(x^2+bx+c)^m$ aporta términos $\frac{B_k x + C_k}{(x^2+bx+c)^k}$. Los términos lineales dan logaritmos; los cuadráticos, logaritmos y arcotangentes.

Ejemplo 6.67.

$$\int \frac{x+4}{x^2-x-2}\,\mathrm{d}x = \int \frac{x+4}{(x-2)(x+1)}\,\mathrm{d}x.$$

Con $\frac{x+4}{(x-2)(x+1)} = \frac{A}{x-2} + \frac{B}{x+1}$: multiplicando y evaluando en $x=2$: $A=2$; en $x=-1$: $B=-1$. Resultado: $2\ln|x-2| - \ln|x+1| + C$.

6.10.2. Integrales impropias

Definición 6.68. Sobre intervalos infinitos: $\int_a^\infty f = \lim_{R\to\infty}\int_a^R f$, si el límite existe. Con integrando no acotado en b: $\int_a^b f = \lim_{\epsilon\to 0^+}\int_a^{b-\epsilon} f$.

Las dos escalas patrón, que hay que memorizar:

$$\int_1^\infty \frac{\mathrm{d}x}{x^p} \text{ converge } \iff p>1, \qquad \int_0^1 \frac{\mathrm{d}x}{x^p} \text{ converge } \iff p<1.$$

Criterio de comparación: si $0 \le f \le g$ y $\int g$ converge, $\int f$ converge; en la práctica se usa la *comparación por límite* con estas escalas. Ejemplo notable: $\int_0^\infty e^{-x^2}\mathrm{d}x$ converge (comparar con e^{-x} para $x \ge 1$) y vale $\frac{\sqrt{\pi}}{2}$, como veremos en el capítulo 9. La función Gamma,

$$\Gamma(s) = \int_0^\infty x^{s-1}e^{-x}\,\mathrm{d}x \quad (s>0), \qquad \Gamma(n+1) = n!,$$

interpola el factorial y aparecerá en las EDO (funciones de Bessel).

6.11. Series numéricas

Idea clave

Intuición primero. Sumar infinitos números no es una operación que exista por sí sola: hay que *definirla*, y se define como el límite de las sumas parciales. Es decir, no sumas infinitos términos; sumas los primeros n y miras hacia dónde se dirige ese resultado cuando n crece. Todo el aparato de criterios de esta sección sirve para responder a una única pregunta con dos posibles respuestas: ¿ese límite existe o no?
Dos advertencias que ahorran mucho tiempo. La primera: que los términos tiendan a cero es *necesario* pero no suficiente, y la serie armónica es el contraejemplo que hay que llevar siempre encima —sus términos van a cero y aun así la suma diverge, muy despacio pero sin freno—. La segunda: la convergencia absoluta y la condicional no son grados de una misma cosa, son fenómenos distintos. Una serie absolutamente convergente se puede reordenar libremente; una condicionalmente convergente se puede reordenar para que sume *cualquier* número que se te antoje, y ese teorema de Riemann es una de las sorpresas genuinas del análisis.

Definición 6.69. $\sum a_n$ converge a S si $S_N = \sum_{n\le N} a_n \to S$. *Condición necesaria*: $a_n \to 0$ (¡no suficiente: la armónica $\sum \frac{1}{n}$ diverge!).

Ejemplo 6.70 (Las dos series patrón)**.** *Geométrica*: $\sum_{n\ge 0} r^n = \frac{1}{1-r}$ si $|r|<1$, divergente si $|r|\ge 1$ (la suma parcial es $\frac{1-r^{N+1}}{1-r}$). *Armónica generalizada*: $\sum \frac{1}{n^p}$ converge $\iff p>1$ (por el criterio integral con $\int_1^\infty x^{-p}\mathrm{d}x$, o por condensación de Cauchy). La divergencia de $\sum \frac{1}{n}$ se ve agrupando: $\frac{1}{3}+\frac{1}{4} > \frac{1}{2}$, $\frac{1}{5}+\cdots+\frac{1}{8} > \frac{1}{2}$, …

Teorema 6.71 (Criterios de convergencia para términos positivos)**.** *(i)* Comparación*:* $0 \le a_n \le b_n$ *y* $\sum b_n$ *converge* $\implies$ $\sum a_n$ *converge.*

(ii) Comparación por límite*: si* $a_n/b_n \to L \in (0,\infty)$*, ambas series comparten destino.*

(iii) Cociente (d'Alembert)*: si* $a_{n+1}/a_n \to L$*: converge si* $L<1$*, diverge si* $L>1$ *(*$L=1$ *no decide).*

(iv) Raíz (Cauchy)*: igual con* $\sqrt[n]{a_n} \to L$.

(v) Integral*: si* f *es positiva decreciente,* $\sum f(n)$ *y* $\int_1^\infty f$ *convergen o divergen juntas.*

Demostración del cociente. Si $L < 1$, elijamos r con $L < r < 1$: desde cierto N, $a_{n+1} < r a_n$, luego $a_{N+k} < r^k a_N$ y la serie queda dominada por una geométrica convergente. Si $L > 1$, los términos crecen y no tienden a 0. □

Idea clave

Qué dice esto en cristiano. Los tres criterios responden a la misma pregunta con estrategias distintas. La comparación dice "no calcules, mide contra una serie conocida"; las series patrón son la geométrica y las $\sum 1/n^p$. El cociente y la raíz dicen "mira si se parece a una geométrica de razón L". Y por eso ambos fallan cuando $L = 1$: ahí es donde vive la frontera, ocupada precisamente por las series p, que con el mismo $L = 1$ unas convergen y otras no. El orden práctico es: primero mirar si $a_n \to 0$, luego comparar, y solo después derivar cocientes.

Ejemplo 6.72 (Las series patrón y los tres criterios)**. (a) La frontera de las series** p**.** Sumando cien mil términos de $\sum 1/n^p$:

p	suma parcial	
0,5	630,996759	diverge
1	12,090146	diverge
1,5	2,606051	converge
2	1,644924	converge
3	1,202057	converge

Los límites exactos para $p = 2$ y $p = 3$ son $\zeta(2) = 1{,}644934$ y $\zeta(3) = 1{,}202057$. Obsérvese lo despacio que crece el caso $p = 1$: cien mil términos y solo va por 12.

(b) Cociente, con factoriales o potencias. Para $\sum \frac{n}{2^n}$ la razón $\frac{a_{n+1}}{a_n} = \frac{n+1}{2n}$ tiende a $\frac{1}{2} < 1$: converge. La suma exacta es 2, y las parciales de 5, 10 y 20 términos dan 1,78125, 1,98828 y 1,99998.

(c) Raíz, cuando todo está elevado a n**.** Para $\sum \left(\frac{n}{2n+1}\right)^n$ la raíz n-ésima es $\frac{n}{2n+1} \to \frac{1}{2} < 1$: converge. Aquí el criterio del cociente daría una expresión mucho más fea, así que la elección importa.

(d) Condensación, para los casos límite. Cuando a_n decrece, $\sum a_n$ converge si y solo si $\sum 2^k a_{2^k}$ converge. Para $a_n = \frac{1}{n}$ eso da $\sum \frac{2^k}{2^k} = \sum 1$, que diverge. Para $a_n = \frac{1}{n \ln n}$ da $\sum \frac{1}{k \ln 2}$, que también diverge. Pero para $a_n = \frac{1}{n(\ln n)^2}$ da $\sum \frac{1}{k^2 (\ln 2)^2}$, que converge. Ahí está la verdadera frontera, mucho más fina que la de las series p.

Cuidado

[$L = 1$ significa "no lo sé", y $a_n \to 0$ no demuestra nada] Cuando el cociente o la raíz devuelven 1, el criterio no concluye: $\sum \frac{1}{n}$ y $\sum \frac{1}{n^2}$ dan ambas $L = 1$ y se comportan de manera opuesta. Hay que cambiar de herramienta. Segundo aviso, y es el error más repetido del tema: que el término general tienda a cero es condición *necesaria* y en absoluto suficiente; la armónica lo cumple y diverge. Y tercero: al comparar, la desigualdad debe ir en el sentido correcto. Como $\frac{1}{\sqrt{n}} > \frac{1}{n}$ y la armónica diverge, $\sum \frac{1}{\sqrt{n}}$ diverge también; compararla con $\sum \frac{1}{n^2}$ "porque se parece" no demuestra nada.

Teorema 6.73 (Series alternadas, Leibniz)**.** *Si $b_n \downarrow 0$ (decreciente con límite 0), entonces $\sum(-1)^{n-1}b_n$ converge, y el error de truncar está acotado por el primer término omitido: $|S - S_N| \le b_{N+1}$.*

Idea clave

Qué dice esto en cristiano. Cuando los signos se alternan y los tamaños menguan, las sumas parciales van saltando por encima y por debajo del valor final, encerrándolo cada vez más. De ahí sale lo mejor del criterio: no solo garantiza que la serie converge, sino que regala una *cota del error* gratis. Si cortas en el término n, te has equivocado en menos que el primer término que has dejado fuera. Ningún otro criterio elemental ofrece eso.

Ejemplo 6.74 (La convergencia y la cota, comprobadas)**. (a) La serie armónica alternada.** Con $b_n = \frac{1}{n}$, decreciente y con límite cero, la serie $1 - \frac{1}{2} + \frac{1}{3} - \frac{1}{4} + \cdots$ converge, y su suma es $\ln 2 = 0{,}6931472$.

(b) Las sumas parciales encerrando el límite.

n	S_n	error real	cota $1/(n+1)$
1	1,0000000	0,3068528	0,5000000
2	0,5000000	0,1931472	0,3333333
3	0,8333333	0,1401862	0,2500000
5	0,7833333	0,0901862	0,1666667
10	0,6456349	0,0475123	0,0909091

Las impares quedan por encima, las pares por debajo, y en todos los casos el error real es menor que la cota, como promete el teorema.

(c) Convergencia lentísima. La serie de Leibniz $\frac{\pi}{4} = 1 - \frac{1}{3} + \frac{1}{5} - \cdots$ converge, pero con mil términos da 0,7851482 frente al valor exacto 0,7853982: apenas tres cifras correctas. La cota lo anticipa: el error es del orden de $\frac{1}{2n+1}$, así que para diez cifras harían falta unos cinco mil millones de términos.

(d) Cuando sí converge deprisa. Para $\sum \frac{(-1)^{n+1}}{n^2}$, cuya suma es $\frac{\pi^2}{12} = 0{,}8224670$, diez términos ya dan 0,8179622, con error 0,0045, por debajo de la cota $\frac{1}{121} = 0{,}00826$. La rapidez la marca lo deprisa que decrezca b_n.

Cuidado

[Contraejemplo: hace falta que b_n sea decreciente] No basta con que $b_n \to 0$ y los signos alternen. Si se toma $b_n = \frac{1}{n}$ para n impar y $b_n = \frac{1}{n^2}$ para n par, la sucesión tiende a cero pero no es monótona, y la serie alternada correspondiente diverge, porque los términos positivos forman esencialmente una armónica que los negativos, mucho menores, no consiguen compensar. Segundo aviso: convergencia condicional no es lo mismo que absoluta. La armónica alternada converge, pero $\sum \left|\frac{(-1)^{n+1}}{n}\right| = \sum \frac{1}{n}$ diverge; por el teorema de reordenación de Riemann, sus términos se pueden reordenar para que la suma dé *cualquier* número real que se desee.

Definición 6.75. $\sum a_n$ converge *absolutamente* si $\sum |a_n| < \infty$ (lo que implica convergencia), y *condicionalmente* si converge pero no absolutamente, como $\sum \frac{(-1)^n}{n} = -\ln 2$.

Teorema 6.76 (Riemann)**.** *Una serie condicionalmente convergente puede reordenarse para converger a* cualquier *valor prefijado, o para divergir. Las absolutamente convergentes, en cambio, tienen suma invariante bajo reordenación.*

Este resultado sorprendente marca la diferencia práctica: solo con convergencia absoluta se puede operar con series como si fueran sumas finitas (reordenar, multiplicar, agrupar).

6.12. Series de potencias y funciones analíticas

Idea clave

Intuición primero. Una serie de potencias es un polinomio de infinitos términos, y la idea es que muchas funciones importantes —la exponencial, el seno, el logaritmo— *son* exactamente eso, si te mantienes lo bastante cerca del centro. Eso convierte funciones trascendentes en objetos que se manipulan como polinomios: se derivan término a término, se integran término a término, se multiplican, se componen.
El "lo bastante cerca" se llama radio de convergencia, y tiene una propiedad geométrica curiosa: la serie converge en un intervalo simétrico alrededor del centro, y el radio se decide por la singularidad más cercana. Aquí eso parece arbitrario —¿por qué la serie de $\frac{1}{1+x^2}$ deja de funcionar en $|x| = 1$ si la función es perfectamente lisa en toda la recta real?— y no tiene respuesta en el mundo real. La tiene en el complejo: hay polos en $\pm i$, a distancia 1 del origen. Es uno de los mejores argumentos para leer el capítulo 30.

Teorema 6.77 (Radio de convergencia)**.** *Para $\sum c_n(x-a)^n$ existe $R \in [0,\infty]$ (fórmula de Cauchy–Hadamard: $\frac{1}{R} = \limsup \sqrt[n]{|c_n|}$, o $\frac{1}{R} = \lim |c_{n+1}/c_n|$ cuando existe) tal que la serie converge absolutamente si $|x-a| < R$ y diverge si $|x-a| > R$. En los bordes $x = a \pm R$ puede pasar cualquier cosa y se estudia aparte.*

Teorema 6.78. *En el interior del intervalo de convergencia, una serie de potencias puede derivarse e integrarse término a término, y el radio no cambia. En consecuencia, $f(x) =$*

$\sum c_n(x-a)^n$ *es infinitamente derivable y* $c_n = \frac{f^{(n)}(a)}{n!}$*:* ***la serie de potencias de una función, si existe, es su serie de Taylor****, y los coeficientes son únicos.*

Ejemplo 6.79 (Cosechando series)**.** De $\frac{1}{1-x} = \sum x^n$: sustituyendo $x \to -x^2$ e integrando,

$$\arctan x = \int_0^x \frac{\mathrm{d}t}{1+t^2} = \sum_{n\geq 0} \frac{(-1)^n x^{2n+1}}{2n+1},$$

y evaluando en $x = 1$ (justificable por Abel): $\frac{\pi}{4} = 1 - \frac{1}{3} + \frac{1}{5} - \cdots$. Del mismo modo se obtiene $\ln(1+x)$ integrando la geométrica. Manipular series conocidas es casi siempre mejor que calcular derivadas.

Cuidado

No toda función C^∞ es analítica: $f(x) = e^{-1/x^2}$ (con $f(0) = 0$) tiene *todas* sus derivadas nulas en 0, así que su serie de Taylor es 0, pero $f \neq 0$. La igualdad $f =$ su Taylor requiere que el resto $R_n \to 0$, y hay que comprobarlo.

6.13. Fundamentos: la construcción de los números reales

Hasta aquí hemos *postulado* la completitud. Pero ¿existe realmente un cuerpo ordenado completo? La respuesta afirmativa se construye, y hay dos rutas clásicas.

Cortaduras de Dedekind. Una *cortadura* es un subconjunto $A \subset \mathbb{Q}$ no vacío, distinto de $\mathbb{Q}$, cerrado hacia abajo ($q < a \in A \Rightarrow q \in A$) y sin máximo. La idea: cada cortadura "es" el número real $\sup A$. Se define $\mathbb{R}$ como el conjunto de las cortaduras, con el orden $A \leq B \iff A \subseteq B$, la suma $A + B = \{a + b : a \in A, b \in B\}$ y un producto definido con cuidado de signos. La completitud es entonces casi gratis: dado un conjunto acotado $\mathcal{F}$ de cortaduras, su supremo es simplemente la unión $\bigcup_{A\in\mathcal{F}} A$, que se verifica ser cortadura. Por ejemplo, $\sqrt{2}$ *es* la cortadura $\{q \in \mathbb{Q} : q \leq 0 \text{ o } q^2 < 2\}$: el agujero de $\mathbb{Q}$ se rellena con el conjunto de sus aproximaciones.

Sucesiones de Cauchy (Cantor). Alternativamente, $\mathbb{R}$ se define como el conjunto de sucesiones de Cauchy de racionales, identificando dos sucesiones cuando su diferencia tiende a cero (formalmente: el cociente del anillo de sucesiones de Cauchy por el ideal de las que tienden a 0 —lenguaje del capítulo 25—). Cada real "es" la clase de todas las sucesiones racionales que lo aproximan; $\sqrt{2}$ es la clase de $(1, 1.4, 1.41, 1.414, \dots)$. Esta construcción es la que se generaliza (la *completación* de un espacio métrico, y también los números p-ádicos de la teoría de números).

Teorema 6.80 (Unicidad)**.** *Todo cuerpo ordenado completo es isomorfo a* $\mathbb{R}$ *(con isomorfismo único que respeta el orden). Por eso hablar de "los" reales tiene sentido: las dos construcciones producen el mismo objeto.*

Ejemplo 6.81 (Cuerpos ordenados que no son $\mathbb{R}$)**. (a)** $\mathbb{Q}$**.** Es ordenado y no completo: el conjunto $\{x \in \mathbb{Q} : x^2 < 2\}$ está acotado superiormente y no tiene supremo racional. Le falta exactamente una propiedad, y esa carencia es la que $\mathbb{R}$ viene a reparar.

(b) Los números algebraicos reales. Forman un cuerpo ordenado, numerable, donde todo polinomio con coeficientes en él tiene sus raíces reales dentro. Tampoco es completo: π no está.

(c) Un cuerpo ordenado no arquimediano. El cuerpo de funciones racionales $\mathbb{R}(t)$, ordenado declarando t mayor que toda constante, contiene "infinitos". Su subconjunto de elementos acotados no tiene supremo, y el enunciado "para todo x existe n con $n > x$" falla. Los hiperreales del análisis no estándar son de esta clase.

(d) $\mathbb{C}$ no es ordenable. Si lo fuera, $\mathrm{i}^2 = -1$ tendría que ser positivo por ser un cuadrado, y a la vez $-1 < 0$. La contradicción muestra que el orden no es un adorno que se pueda añadir a cualquier cuerpo: es una estructura que restringe fuertemente.

Teorema 6.82 (Equivalencias de la completitud)**.** *En un cuerpo ordenado arquimediano, son equivalentes: (i) todo conjunto acotado tiene supremo; (ii) toda sucesión monótona acotada converge; (iii) intervalos cerrados encajados con longitudes $\to 0$ tienen intersección de un punto; (iv) Bolzano–Weierstrass; (v) toda sucesión de Cauchy converge.*

Demostración del ciclo (iv)⇒(v) y (v)⇒(ii). (iv)⇒(v): una sucesión de Cauchy está acotada (desde cierto N todos los términos distan < 1 de a_N); por B–W tiene subsucesión $a_{n_k} \to L$; y una Cauchy con subsucesión convergente converge entera: $|a_n - L| \le |a_n - a_{n_k}| + |a_{n_k} - L|$, ambos pequeños para n, n_k grandes. *(v)⇒(ii)*: si (a_n) es creciente, acotada y *no* fuera de Cauchy, existiría $\varepsilon > 0$ con pares $m_k < n_k$ arbitrariamente grandes y $a_{n_k} - a_{m_k} \ge \varepsilon$; encadenando bloques disjuntos, la sucesión creciente subiría al menos ε infinitas veces y no estaría acotada: contradicción. Los demás eslabones son similares. □

Ejemplo 6.83 (Las equivalencias, vistas por dentro)**. (a) Supremo implica encajados.** Dados intervalos cerrados encajados $[a_n, b_n]$, el conjunto $\{a_n\}$ está acotado por cualquier b_m, luego tiene supremo c, y ese c está en todos los intervalos. Con $[0, 1/n]$ el punto común es el 0.

(b) Encajados implica Bolzano–Weierstrass. Se biseccIona repetidamente quedándose con la mitad que contenga infinitos términos: las longitudes son $2/2^k$, que valen 1; 0,5; 0,25; 0,125 y tienden a cero, y el punto común es el límite de una subsucesión.

(c) Bolzano–Weierstrass implica Cauchy. Una sucesión de Cauchy está acotada; por Bolzano–Weierstrass tiene una subsucesión convergente; y una sucesión de Cauchy con subsucesión convergente converge al mismo punto.

(d) Dónde se rompe la cadena en $\mathbb{Q}$. Todas las implicaciones son válidas en cualquier cuerpo ordenado; lo que falla en $\mathbb{Q}$ es el punto de partida. La sucesión 1; 1,4; 1,41; 1,414; 1,4142 es de Cauchy en $\mathbb{Q}$ y no converge en $\mathbb{Q}$. Por eso las cinco propiedades son equivalentes *y* las cinco fallan a la vez.

Estas equivalencias explican por qué los teoremas de existencia del capítulo son "el mismo teorema": todos son la completitud con distintos disfraces. Sobre $\mathbb{Q}$, *todos* fallan simultáneamente.

6.14. Demostraciones completas de los teoremas de compacidad

Teorema 6.84 (Bolzano–Weierstrass, demostración completa)**.** *Toda sucesión (a_n) en $[A,B]$ tiene una subsucesión convergente.*

Demostración. Bisección. Sea $I_0=[A,B]$. Al menos una mitad de I_0 contiene a_n para infinitos índices n; llamémosla I_1, y repitamos: obtenemos intervalos encajados $I_0\supset I_1\supset I_2\supset\cdots$ con $|I_k|=\frac{B-A}{2^k}\to 0$, cada uno visitado por infinitos términos. Elegimos índices $n_1<n_2<\cdots$ con $a_{n_k}\in I_k$ (posible porque cada I_k tiene infinitos). Por el principio de intervalos encajados (Teorema 6.82(iii)), $\bigcap I_k=\{L\}$; y como a_{n_k} y L están ambos en I_k, $|a_{n_k}-L|\le|I_k|\to 0$. □

Teorema 6.85 (Heine: continuidad uniforme, demostración completa)**.** *Toda f continua en $[a,b]$ es uniformemente continua.*

Demostración. Por contradicción: si no, existe $\varepsilon_0>0$ y, para cada n, puntos $x_n,y_n\in[a,b]$ con $|x_n-y_n|<\frac{1}{n}$ pero $|f(x_n)-f(y_n)|\ge\varepsilon_0$. Por B–W, una subsucesión $x_{n_k}\to c\in[a,b]$; como $|x_n-y_n|\to 0$, también $y_{n_k}\to c$. La continuidad en c da $f(x_{n_k})\to f(c)$ y $f(y_{n_k})\to f(c)$, luego $|f(x_{n_k})-f(y_{n_k})|\to 0$: contradice $\ge\varepsilon_0$. □

Este patrón de demostración —negar, extraer subsucesión convergente por compacidad, contradecir con continuidad— es una de las técnicas más repetidas del análisis; vale la pena interiorizarlo.

Teorema 6.86 (Darboux: las derivadas tienen la propiedad del valor intermedio)**.** *Si f es derivable en $[a,b]$ y $f'(a)<\lambda<f'(b)$, existe $c\in(a,b)$ con $f'(c)=\lambda$* —aunque f' no sea continua—*.*

Demostración. Sea $g(x)=f(x)-\lambda x$: $g'(a)<0<g'(b)$. Por Weierstrass, g alcanza su mínimo en $[a,b]$; $g'(a)<0$ impide que sea en a (justo a la derecha de a, g decrece) y $g'(b)>0$ impide que sea en b. El mínimo es interior, y por Fermat $g'(c)=0$, es decir $f'(c)=\lambda$. □

Idea clave

Qué dice esto en cristiano. Uno esperaría que una derivada, para tomar todos los valores intermedios, tuviera que ser continua. El teorema de Darboux dice que no hace falta: aunque f' tenga discontinuidades, nunca puede *saltarse* un valor. Dicho al revés, y esta es la lectura más útil: hay funciones perfectamente inofensivas que no son la derivada de nadie. La función escalón, por ejemplo, no tiene primitiva, no porque nadie la haya encontrado, sino porque no puede existir.

Ejemplo 6.87 (Una derivada discontinua que aun así no salta)**. (a) El ejemplo canónico.** Sea $f(x)=x^2\sin\frac{1}{x}$ con $f(0)=0$. Ya vimos que es derivable en todas partes, con $f'(0)=0$. Fuera del origen, $f'(x)=2x\sin\frac{1}{x}-\cos\frac{1}{x}$.

(b) La derivada oscila salvajemente. Evaluando: $f'(0{,}1) = 0{,}730267$, $f'(0{,}01) = -0{,}872446$ y $f'(0{,}001) = -0{,}560725$. Los valores no se acercan a $f'(0) = 0$; de hecho f' toma todos los valores entre -1 y 1 en cualquier entorno del origen. Es una derivada discontinua en un punto.

(c) Y sin embargo cumple Darboux. Precisamente porque oscila tomando *todos* los valores intermedios, no se salta ninguno. La discontinuidad de una derivada nunca puede ser de salto; solo puede ser del tipo oscilante como esta. Es una restricción muy fuerte sobre qué funciones pueden aparecer como derivadas.

(d) La consecuencia práctica. La función signo, que vale -1 para $x < 0$, 0 en el origen y 1 para $x > 0$, no es la derivada de ninguna función, porque pasa de -1 a 1 sin tomar, por ejemplo, el valor $\frac{1}{2}$. Lo mismo ocurre con la parte entera. Ninguna de las dos tiene primitiva en un intervalo que contenga el salto, y el teorema lo demuestra sin buscarla.

Cuidado

[No confundir con el teorema del valor intermedio de Bolzano] Bolzano habla de funciones *continuas*; Darboux habla de *derivadas*, que pueden no serlo. Son enunciados con la misma conclusión y con hipótesis muy distintas, y precisamente por eso Darboux resulta sorprendente. Segundo aviso: el teorema no dice que toda derivada sea continua, ni que las derivadas discontinuas sean raras; solo prohíbe un tipo concreto de discontinuidad. Y tercero: que una función tenga primitiva no implica que esa primitiva se pueda escribir con funciones elementales; e^{-x^2} es continua, luego tiene primitiva por el teorema fundamental, y aun así esa primitiva no admite expresión elemental.

Consecuencia curiosa: una derivada no puede tener discontinuidades de salto; si f' existe en todo punto, sus discontinuidades son necesariamente esenciales (como $x^2 \sin(1/x)$).

6.15. Sucesiones y series de funciones: convergencia uniforme

Idea clave

Intuición primero. Convergencia *puntual*: cada espectador del estadio acaba sentándose, pero cada uno a su hora; siempre hay algún rezagado de pie. Convergencia *uniforme*: se anuncia una hora y a partir de ella *todo el estadio* está sentado a la vez. La diferencia parece burocrática y es todo: las propiedades que queremos heredar en el límite (continuidad, integral) son afirmaciones sobre *todos los puntos a la vez*, así que solo sobreviven si la aproximación es buena *en todos los puntos a la vez*. El dibujo clave: la banda de tolerancia $f \pm \varepsilon$; uniforme significa que la gráfica entera de f_n acaba entrando en la banda, no punto por punto sino de cuerpo entero.

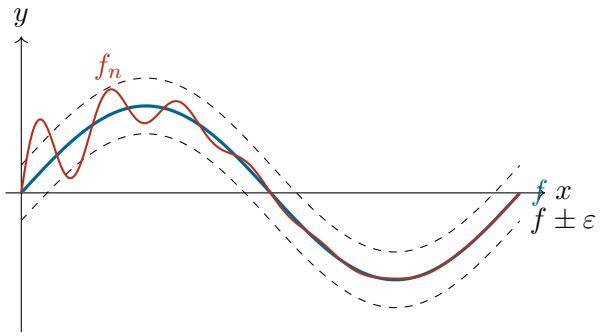


Figura 6.2: Convergencia uniforme: la gráfica de f_n (roja) acaba entrando *entera* en la banda $f \pm \varepsilon$; a la izquierda aún se sale, a la derecha ya vive dentro.

Falta el fundamento que sostiene a las series de potencias y de Fourier: ¿cuándo se puede intercambiar un límite con una integral, una derivada u otro límite?

Definición 6.88. $f_n \to f$ *puntualmente* en D si $f_n(x) \to f(x)$ para cada x. La convergencia es *uniforme* si

$$\sup_{x \in D} |f_n(x) - f(x)| \longrightarrow 0 :$$

el mismo N sirve para todos los x a la vez. Gráficamente: toda la gráfica de f_n entra en la banda $f \pm \varepsilon$.

Ejemplo 6.89 (La convergencia puntual no basta)**.** $f_n(x) = x^n$ en $[0,1]$ converge puntualmente a la función que vale 0 en $[0,1)$ y 1 en $x = 1$: *el límite puntual de continuas puede ser discontinuo.* Y $g_n =$ triángulo de base $[0, \frac{1}{n}]$ y altura n converge puntualmente a 0, pero $\int_0^1 g_n = \frac{1}{2}$ para todo n: *el límite no conmuta con la integral.* En ambos casos la convergencia no es uniforme ($\sup |f_n - f| = 1$ y $\sup |g_n| = n$).

Teorema 6.90 (Lo que la uniformidad compra)**.** *Si $f_n \to f$ uniformemente en $[a,b]$:*

(i) si las f_n son continuas, f es continua;

(ii) si las f_n son integrables, $\int_a^b f_n \to \int_a^b f$;

(iii) si las f_n son C^1, $f_n \to f$ al menos puntualmente, y $f_n' \to g$ uniformemente*, entonces f es derivable con $f' = g$ (nótese: la uniformidad se necesita en las* derivadas*).*

Demostración de (i) y (ii). (i) El argumento "$\varepsilon/3$": dado x_0 y $\varepsilon > 0$, elegimos n con $\sup |f_n - f| < \frac{\varepsilon}{3}$, y δ de la continuidad de f_n en x_0. Para $|x - x_0| < \delta$:

$$|f(x) - f(x_0)| \le \underbrace{|f(x) - f_n(x)|}_{<\varepsilon/3} + \underbrace{|f_n(x) - f_n(x_0)|}_{<\varepsilon/3} + \underbrace{|f_n(x_0) - f(x_0)|}_{<\varepsilon/3} < \varepsilon.$$

(ii) $\left| \int f_n - \int f \right| \le \int |f_n - f| \le (b-a) \sup |f_n - f| \to 0.$ □

Idea clave

Qué dice esto en cristiano. Hay dos maneras de que una sucesión de funciones se acerque a otra. En la convergencia *puntual*, cada punto va por su cuenta y unos pueden tardar muchísimo más que otros. En la *uniforme*, todos los puntos convergen al mismo ritmo: existe un n a partir del cual la gráfica entera queda dentro de una franja de anchura ε alrededor del límite. La diferencia parece técnica y no lo es: solo la uniforme permite pasar el límite dentro de una integral o conservar la continuidad. La puntual no conserva casi nada.

Ejemplo 6.91 (Lo que se pierde sin uniformidad)**. (a) La continuidad se rompe.** Sea $f_n(x) = x^n$ en $[0,1]$. Puntualmente converge a 0 para $x < 1$ y a 1 para $x = 1$: el límite es discontinuo pese a que todas las f_n son polinomios. En $x = 0{,}9$ los valores son 0,348678, 0,005154 y 0,000027 para $n = 10, 50, 100$; pero en $x = 0{,}99$ son 0,904382, 0,605006 y 0,366032: mucho más lentos. Cuanto más cerca de 1, más tarda, y por eso no hay uniformidad.

(b) La medida del fallo. El criterio operativo es mirar $\sup_x |f_n(x) - f(x)|$. Aquí ese supremo vale 1 para todo n, porque siempre hay puntos cerca de 1 donde f_n aún vale casi 1 mientras el límite ya vale 0. Como no tiende a cero, la convergencia no es uniforme.

(c) Un caso que sí lo es. Para $f_n(x) = \frac{x}{n}$ en $[0,1]$, el supremo de $|f_n - 0|$ vale $\frac{1}{n}$, que sí tiende a cero: convergencia uniforme. Aquí el límite es continuo y se puede integrar término a término sin problema.

(d) La integral se escapa. Sea $f_n(x) = n\,x\,\mathrm{e}^{-nx^2}$ en $[0,1]$. Puntualmente $f_n \to 0$ en todo punto. Pero $\int_0^1 f_n = \frac{1-\mathrm{e}^{-n}}{2}$, que vale 0,499977 para $n = 10$ y 0,5 para $n = 100$: tiende a $\frac{1}{2}$, no a 0. El límite de las integrales no es la integral del límite. La razón se ve en el máximo de f_n, que vale 1,3562 para $n = 10$, luego 4,2888 y luego 13,5624: una joroba cada vez más alta y más estrecha que se desliza hacia el origen sin perder área.

Cuidado

[La uniformidad es sobre el conjunto, y no es necesaria para todo] Una misma sucesión puede converger uniformemente en un trozo y no en otro: x^n converge uniformemente en $[0;\, 0{,}9]$, donde el supremo es $0{,}9^n \to 0$, y no en $[0,1]$. Siempre hay que decir *dónde*. Segundo aviso: la uniformidad es condición suficiente para intercambiar límite e integral, no necesaria; hay sucesiones que convergen solo puntualmente y cuyas integrales sí convergen correctamente. El teorema de la convergencia dominada de Lebesgue da una condición mucho más débil y mucho más útil. Y tercero: la convergencia uniforme conserva la continuidad y la integral, pero *no* la derivabilidad; hace falta que converjan uniformemente las derivadas.

Teorema 6.92 (Criterio M de Weierstrass)**.** *Si* $|u_n(x)| \le M_n$ *para todo* $x \in D$ *y* $\sum M_n < \infty$, *entonces* $\sum u_n(x)$ *converge absoluta y* uniformemente *en* D.

Demostración. Las sumas parciales son uniformemente de Cauchy: $\left|\sum_{n=p}^{q} u_n(x)\right| \le \sum_{n=p}^{q} M_n$, que es pequeño para p, q grandes independientemente de x; y el criterio de Cauchy uniforme implica convergencia uniforme. □

Idea clave

Qué dice esto en cristiano. El criterio M es la manera más cómoda que existe de demostrar convergencia uniforme, y consiste en olvidarse de la variable. Si cada término está acotado por un número M_n que no depende de x, y esos números forman una serie convergente, entonces la serie de funciones converge uniformemente en todo el conjunto. Se cambia un problema sobre funciones por uno sobre números, que es infinitamente más fácil. Es el caballo de batalla para justificar que las series de Fourier y las de potencias se pueden manipular término a término.

Ejemplo 6.93 (Aplicar el criterio en tres situaciones)**. (a) Una serie trigonométrica en toda la recta.** Sea $\sum \frac{\sin(nx)}{n^2}$. Como $|\sin(nx)| \leq 1$, se puede tomar $M_n = \frac{1}{n^2}$, y $\sum \frac{1}{n^2} = \frac{\pi^2}{6} =$ 1,6449341 converge. Luego la serie converge uniformemente en $\mathbb{R}$ entero, y su suma es una función continua. Obsérvese que la cota no depende de x en absoluto: ese es el punto.

(b) Una serie de potencias en un cerrado. Sea $\sum \frac{x^n}{n^2}$ en $[-1, 1]$. Como $|x| \leq 1$, otra vez $M_n = \frac{1}{n^2}$ sirve, y la convergencia es uniforme en todo el intervalo cerrado, extremos incluidos. Eso permite integrar y evaluar en $x = 1$ sin ningún cuidado especial.

(c) Lo que garantiza de propina. El criterio da convergencia *absoluta* y *uniforme* a la vez, y por tanto la suma es continua, se puede integrar término a término y se puede reordenar libremente. Tres conclusiones fuertes a cambio de encontrar unos números M_n.

(d) Cuándo no se puede aplicar. Para $\sum x^n$ en $[0, 1)$ no hay ningún M_n válido: el supremo de $|x^n|$ sobre ese intervalo es 1 para todo n, y $\sum 1$ diverge. La serie converge puntualmente a $\frac{1}{1-x}$, pero no uniformemente, cosa coherente con que la suma se dispare cerca de $x = 1$. En $[0; 0{,}9]$, en cambio, sí vale $M_n = 0{,}9^n$, cuya serie converge.

Cuidado

[El criterio es suficiente, no necesario] Que no se encuentre una sucesión M_n válida no demuestra que la convergencia no sea uniforme. La serie $\sum \frac{(-1)^n}{n} x^n$ converge uniformemente en $[0, 1]$ por el criterio de Abel, y sin embargo el criterio M falla, porque las cotas $\frac{1}{n}$ dan una serie divergente. Segundo aviso: los M_n deben ser constantes, independientes de x; escribir "$M_n = \frac{x^n}{n^2}$" no es aplicar el criterio, es no haber hecho nada. Y tercero: el criterio da convergencia absoluta, así que nunca servirá para series que solo convergen condicionalmente, como la del apartado anterior.

Corolario 6.94 (Justificación de las series de potencias)**.** *Una serie de potencias con radio R converge uniformemente en cada intervalo $[a - r, a + r]$ con $r < R$ (comparar con $\sum |c_n| r^n$, convergente, vía el criterio M). Por el Teorema 6.90, su suma es continua en $(a - R, a + R)$, se integra término a término, y —aplicando el criterio M también a la serie derivada, que tiene el mismo radio— se deriva término a término.* Toda *la teoría de funciones analíticas de la sección anterior descansa en estas tres líneas.*

Observación 6.95. La convergencia uniforme es también la clave oculta del capítulo 12: las series de Fourier y las soluciones en serie de las EDO se derivan término a término precisamente

bajo estas hipótesis. Y el "espacio de funciones con la norma sup" donde la convergencia uniforme es *la* convergencia es el primer ejemplo de espacio de Banach: la puerta del análisis funcional.

6.16. Complementos: hiperbólicas, complejos y técnicas finas

6.16.1. Funciones hiperbólicas

Definición 6.96. $\sinh x = \frac{e^x - e^{-x}}{2}$, $\cosh x = \frac{e^x + e^{-x}}{2}$, $\tanh x = \frac{\sinh x}{\cosh x}$. Cumplen la "trigonometría con signos cambiados":

$$\cosh^2 x - \sinh^2 x = 1, \qquad \sinh(a+b) = \sinh a \cosh b + \cosh a \sinh b,$$

y parametrizan la hipérbola $x^2 - y^2 = 1$ como seno y coseno parametrizan el círculo. Derivadas: $(\sinh)' = \cosh$, $(\cosh)' = \sinh$ (sin signo menos). Inversas con fórmula logarítmica:

$$\operatorname{arcsinh} x = \ln\left(x + \sqrt{x^2+1}\right), \qquad \operatorname{arctanh} x = \tfrac{1}{2}\ln\tfrac{1+x}{1-x},$$

que se obtienen resolviendo la cuadrática en e^y. Aparecen al integrar $\int \frac{\mathrm{d}x}{\sqrt{x^2 \pm a^2}}$, en la catenaria $y = a\cosh(x/a)$ (la forma de un cable colgante) y en la relatividad (las "rotaciones hiperbólicas" de Lorentz).

6.16.2. Números complejos y la fórmula de Euler, demostrada

Un complejo es $z = x + iy$ con $i^2 = -1$; módulo $|z| = \sqrt{x^2+y^2}$, conjugado $\bar z = x - iy$, forma polar $z = r(\cos\theta + i\sin\theta)$. Las series de potencias convergen igual con argumento complejo (mismo criterio del cociente sobre módulos), lo que permite *definir* $e^z = \sum \frac{z^n}{n!}$.

Teorema 6.97 (Fórmula de Euler)**.** *$e^{i\theta} = \cos\theta + i\sin\theta$ para todo $\theta \in \mathbb{R}$.*

Demostración. Sustituyendo $z = i\theta$ en la serie y separando pares e impares (reordenación lícita: convergencia absoluta):

$$e^{i\theta} = \sum_{n\ge 0} \frac{(i\theta)^n}{n!} = \underbrace{\sum_{k\ge 0} \frac{(-1)^k \theta^{2k}}{(2k)!}}_{\cos\theta} + i \underbrace{\sum_{k\ge 0} \frac{(-1)^k \theta^{2k+1}}{(2k+1)!}}_{\sin\theta},$$

usando $i^{2k} = (-1)^k$: exactamente los desarrollos de coseno y seno. □

Dividendos: $e^{i\pi} + 1 = 0$; De Moivre $(\cos\theta + i\sin\theta)^n = \cos n\theta + i\sin n\theta$, que genera las identidades de ángulos múltiples; las raíces de la unidad $\zeta_k = e^{2\pi i k/n}$ (protagonistas del capítulo 26); y el truco de integrar $\int e^{ax}\cos bx\,\mathrm{d}x$ como parte real de $\int e^{(a+ib)x}\mathrm{d}x$.

6.16.3. Derivación bajo el signo integral

Teorema 6.98 (Regla de Leibniz)**.** *Si $f(x,t)$ y $\frac{\partial f}{\partial t}$ son continuas, $F(t) = \int_a^b f(x,t)\,\mathrm{d}x$ es derivable con $F'(t) = \int_a^b \frac{\partial f}{\partial t}\,\mathrm{d}x$; con límites variables se suman $f(b(t),t)b'(t) - f(a(t),t)a'(t)$.*

Ejemplo 6.99 (El "truco de Feynman"). Para $I = \int_0^1 \frac{x-1}{\ln x}\,\mathrm{d}x$ (sin primitiva elemental), sea $I(t) = \int_0^1 \frac{x^t-1}{\ln x}\mathrm{d}x$:

$$I'(t) = \int_0^1 x^t\,\mathrm{d}x = \frac{1}{t+1} \;\Rightarrow\; I(t) = \ln(t+1), \qquad I = I(1) = \ln 2.$$

Derivar respecto de un parámetro bien plantado convierte integrales imposibles en EDOs triviales.

6.16.4. Criterio de Dirichlet para integrales oscilantes

Si $F(R) = \int_a^R f$ está acotada y $g \downarrow 0$, entonces $\int_a^\infty fg$ converge. Ejemplo canónico: $\int_1^\infty \frac{\sin x}{x}\mathrm{d}x$ converge, pero *no* absolutamente (como la armónica alternada). Es el análogo integral de Leibniz, y da sentido a las integrales de la teoría de señales.

6.16.5. La irracionalidad de e

Teorema 6.100. *$e \notin \mathbb{Q}$.*

Demostración. Si $e = \frac{p}{q}$, multiplicando la serie por $q!$: $q!\,e = \underbrace{\sum_{n\le q} \frac{q!}{n!}}_{\text{entero}} + R$ con

$$0 < R = \frac{1}{q+1} + \frac{1}{(q+1)(q+2)} + \cdots < \sum_{k\ge 1} \frac{1}{(q+1)^k} = \frac{1}{q} \le 1 :$$

R sería un entero estrictamente entre 0 y 1. Contradicción. □

Con la misma filosofía (aproximar demasiado bien es imposible para un racional) se demuestran la irracionalidad de π y —capítulo 26— la trascendencia vía Liouville.

6.17. Cierre del temario I: aplicaciones de la derivada

6.17.1. Asíntotas y trazado completo de curvas

Asíntotas: verticales en los límites infinitos laterales; horizontales $y = L$ si $f \to L$ en $\pm\infty$; oblicuas $y = mx + b$ con $m = \lim \frac{f}{x}$, $b = \lim(f - mx)$.

Ejemplo 6.101 (Trazado de $f(x) = \frac{x^2}{x-1}$). Dominio $x \neq 1$; división: $f = x+1+\frac{1}{x-1}$: asíntotas $x = 1$ y $y = x + 1$. $f' = \frac{x(x-2)}{(x-1)^2}$: máximo local $f(0) = 0$, mínimo local $f(2) = 4$. $f'' = \frac{2}{(x-1)^3}$: concavidades opuestas a cada lado, sin inflexión. La gráfica queda determinada: dos ramas pegadas a sus asíntotas con la banda $0 < y < 4$ prohibida. Protocolo: dominio, asíntotas, f', f'', valores clave.

6.17.2. Tasas relacionadas

Ejemplo 6.102 (La escalera). Escalera de 5 m; la base se aleja a 1 m/s. De $x^2 + y^2 = 25$: $\dot y = -\frac{x}{y}\dot x$; con $x = 3$, $y = 4$: $\dot y = -\frac{3}{4}$ m/s. Cuando $y \to 0$, $\dot y \to -\infty$: todo modelo tiene dominio de validez.

Ejemplo 6.103 (El cono de agua)**.** Depósito cónico (radio 2, altura 4) llenándose a 3 m^3/min: por semejanza $r = \frac{h}{2}$, $V = \frac{\pi h^3}{12}$, y $\dot V = \frac{\pi h^2}{4}\dot h$; con $h = 2$: $\dot h = \frac{3}{\pi}$ m/min. Receta: geometría a una variable, cadena, evaluar.

6.17.3. Diferenciales, aproximación lineal y método de Newton

$f(a + \mathrm{d}x) \approx f(a) + f'(a)\mathrm{d}x$: $\sqrt{4.1} \approx 2.025$. Errores relativos con derivada logarítmica: $V = \frac{4}{3}\pi r^3 \Rightarrow \frac{\mathrm{d}V}{V} = 3\frac{\mathrm{d}r}{r}$. **Newton**: $x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)}$, convergencia *cuadrática* cerca de raíces simples ($e_{n+1} \approx \frac{|f''|}{2|f'|}e_n^2$: los dígitos se duplican).

Ejemplo 6.104. $\sqrt{2}$ con $f = x^2 - 2$: $x_{n+1} = \frac{x_n}{2} + \frac{1}{x_n}$ (el método babilónico). Desde 1: 1.5, $1.41\overline{6}$, 1.414215…: 2, 5, 12 dígitos en tres pasos. Advertencias: puede divergir con mal inicio; raíces múltiples degradan a convergencia lineal.

6.18. Cierre del temario II: aplicaciones de la integral

6.18.1. Áreas y volúmenes

Área entre curvas: $\int(f_{\text{arriba}} - f_{\text{abajo}})$. Volumen por secciones: $V = \int A(x)\mathrm{d}x$ (Cavalieri). Revolución:

Discos: $V = \pi\displaystyle\int f^2\,\mathrm{d}x$, **Arandelas:** $\pi\displaystyle\int (R^2 - r^2)$, **Cascarones:** $V = 2\pi\displaystyle\int x\,f(x)\,\mathrm{d}x$.

Ejemplo 6.105 (El mismo sólido por dos métodos)**.** Región bajo $y = x^2$ en $[0,1]$ girada alrededor del eje y. Cascarones: $2\pi\int_0^1 x\cdot x^2\mathrm{d}x = \frac{\pi}{2}$. Arandelas en y: $\pi\int_0^1(1-y)\mathrm{d}y = \frac{\pi}{2}$ ✓. Elegir método = elegir la variable de secciones simples. De estas fórmulas: esfera $\frac{4}{3}\pi R^3$, cono $\frac{1}{3}\pi R^2 h$, toro $2\pi^2 R r^2$.

6.18.2. Longitud de arco y superficies de revolución

$$L = \int \sqrt{1 + f'^2}\,\mathrm{d}x, \qquad S = 2\pi\int f\sqrt{1 + f'^2}\,\mathrm{d}x.$$

Ejemplo 6.106. $y = \frac{2}{3}x^{3/2}$ en $[0,3]$: $L = \int_0^3 \sqrt{1+x}\mathrm{d}x = \frac{14}{3}$. Esfera: con $f = \sqrt{R^2 - x^2}$ el integrando colapsa a R: $S = 4\pi R^2$, con el corolario de Arquímedes (rebanadas de igual altura, igual área: los mapas de Lambert).

Cuidado

[La trompeta de Gabriel] Girando $y = \frac{1}{x}$, $x \geq 1$: volumen $\pi\int x^{-2} = \pi$ finito, área $\geq 2\pi\int\frac{\mathrm{d}x}{x}$ infinita: se llena con pintura finita pero no se pinta.

6.18.3. Valor medio y trabajo

$\bar{f} = \frac{1}{b-a}\int f$ (alcanzado: TVM integral). Trabajo $W = \int F\,\mathrm{d}x$: resorte $\frac{1}{2}kx^2$; bombeo por capas.

Ejemplo 6.107 (Vaciar un tanque semiesférico)**.** Radio R, capa a altura y: radio $\sqrt{2Ry - y^2}$, sube $R - y$:

$$W = \rho g\pi \int_0^R (2R^2 y - 3Ry^2 + y^3)\mathrm{d}y = \rho g\pi\Big(R^4 - R^4 + \frac{R^4}{4}\Big) = \frac{\rho g\pi R^4}{4}.$$

6.18.4. Potencias trigonométricas y estrategia general

$\int \sin^m \cos^n$: exponente impar ⇒ apartar factor y sustituir; ambos pares ⇒ ángulo doble. Estrategia general: (1) simplificar, (2) sustitución evidente, (3) partes (LIATE), (4) racionales/raíces/trig con su cambio, (5) tabla, (6) rendirse con elegancia: e^{-x^2}, $\frac{\sin x}{x}$, $\frac{1}{\ln x}$ *no tienen* primitiva elemental (Liouville): series, parámetros o numéricos.

6.19. Profundización final: herramientas finas del análisis

6.19.1. Límite superior e inferior

$\limsup a_n = \lim_n \sup_{k\geq n} a_k$ y $\liminf$ análogo (existen por monotonía): el mayor y menor límite de subsucesiones; convergencia $\iff$ coinciden. Con ellos, la forma siempre válida del criterio de la raíz es Cauchy–Hadamard: $\frac{1}{R} = \limsup \sqrt[n]{|c_n|}$.

6.19.2. Condensación de Cauchy

Teorema 6.108. *Si $a_n \downarrow 0$: $\sum a_n$ converge $\iff$ $\sum 2^k a_{2^k}$ converge. (Bloques diádicos: cada bloque está entre $\frac{1}{2}$ y 1 veces $2^k a_{2^k}$.)*

Ejemplo 6.109 (El criterio de condensación aplicado)**. (a) La serie armónica.** Con $a_n = 1/n$, la condensada es $\sum 2^k \cdot 2^{-k} = \sum 1$, que diverge. Luego $\sum 1/n$ diverge, y la demostración cabe en una línea.

(b) Las p-series. Con $a_n = n^{-p}$, la condensada es $\sum 2^k \cdot 2^{-kp} = \sum 2^{k(1-p)}$, una geométrica de razón 2^{1-p}. Converge si y sólo si $2^{1-p} < 1$, es decir, $p > 1$. Todo el criterio de las p-series en dos renglones. Para $p = 2$ la condensada suma exactamente 2.

(c) El caso fino. Con $a_n = 1/(n \ln n)$, la condensada es $\sum 2^k/(2^k \cdot k \ln 2) = \frac{1}{\ln 2}\sum 1/k$, que diverge. Y con $a_n = 1/(n(\ln n)^2)$ la condensada es $\frac{1}{(\ln 2)^2}\sum 1/k^2$, que converge. El criterio distingue casos que el de comparación no ve.

(d) La hipótesis de monotonía. Es imprescindible. La serie $\sum a_n$ con $a_n = 1/n$ si n es potencia de dos y $a_n = 0$ en otro caso converge ($\sum 2^{-k} = 2$), mientras que su condensada es $\sum 1$, divergente. Sin decrecimiento el criterio miente.

Rinde la familia logarítmica entera: $\sum n^{-p}$ (geométrica condensada), $\sum \frac{1}{n \ln n}$ diverge, $\sum \frac{1}{n \ln^2 n}$ converge: la escalera infinita de logaritmos, peldaño a peldaño.

6.19.3. Criterios de Abel y Dirichlet para series

Con suma por partes de Abel: *(Dirichlet)* sumas parciales de $\sum b_n$ acotadas y $a_n \downarrow 0$ $\Rightarrow \sum a_n b_n$ converge; *(Abel)* $\sum b_n$ convergente y a_n monótona acotada $\Rightarrow$ converge. Con $b_n = \sin n\theta$: la convergencia de $\sum \frac{\sin n\theta}{n}$, la serie del diente de sierra —los criterios finos existen porque Fourier los necesita—.

6.19.4. Fracciones continuas

$x = [a_0; a_1, a_2, \dots]$ (parte entera, invertir, repetir); convergentes p_k/q_k con $p_k = a_k p_{k-1} + p_{k-2}$: las *mejores* aproximaciones racionales, con $|x - \frac{p_k}{q_k}| < \frac{1}{q_k q_{k+1}}$. Ejemplos: $\sqrt{2} = [1; 2, 2, \dots]$ (convergentes $\frac{3}{2}, \frac{7}{5}, \frac{17}{12}$: ¡las soluciones de Pell!), $\varphi = [1; 1, 1, \dots]$ (el número *peor* aproximable: por eso los girasoles), $\pi = [3; 7, 15, 1, 292, \dots]$ (el 292 explica la calidad de $\frac{355}{113}$). Teorema de Lagrange: expansión periódica $\iff$ irracional cuadrático —la periodicidad de $[1; 2, 2, \dots]$ *es* $x = 1 + \frac{1}{1+x}$—.

6.19.5. Euler–Maclaurin

$$\sum_{k=a}^{b} f(k) = \int_a^b f + \frac{f(a)+f(b)}{2} + \frac{f'(b)-f'(a)}{12} - \cdots$$

(números de Bernoulli). Da $H_n = \ln n + \gamma + \frac{1}{2n} - \frac{1}{12n^2} + \cdots$ (la definición fina de γ), la serie de Stirling completa, y acelera series lentas: $\sum \frac{1}{n^2}$ a 6 decimales pide 10^6 términos directos o 20 con dos correcciones.

6.20. Última capa: los teoremas que faltaban

6.20.1. El teorema de Abel del borde

Teorema 6.110 (Abel)**.** *Si* $\sum a_n$ *converge, entonces* $f(x) = \sum a_n x^n$ *cumple* $\lim_{x\to 1^-} f(x) = \sum a_n$.

Idea. Suma por partes: $f(x) = (1-x)\sum A_n x^n$ con A_n las sumas parciales; los pesos $(1-x)x^n$ suman 1 y la media ponderada hereda el límite de A_n. □

Ejemplo 6.111 (El teorema de Abel y el borde del disco)**. (a) El logaritmo.** De $\ln(1+x) = \sum_{n\geq 1}(-1)^{n+1}x^n/n$ para $|x| < 1$ y de que la serie converge en $x = 1$, el teorema da $\sum_{n\geq 1}(-1)^{n+1}/n = \ln 2 = 0{,}693147$. Las sumas parciales oscilan 1; 0,5; 0,833333; 0,583333; 0,783333; 0,616667 y convergen despacio.

(b) El arcotangente. De $\arctan x = \sum(-1)^n x^{2n+1}/(2n+1)$ y la convergencia en $x = 1$ sale la fórmula de Leibniz $\pi/4 = 1 - \frac{1}{3} + \frac{1}{5} - \cdots$. La serie converge tan despacio que hacen falta unos quinientos términos para dos decimales correctos.

(c) El recíproco es falso. Que exista $\lim_{x\to 1^-} f(x)$ no implica que $\sum a_n$ converja: con $a_n = (-1)^n$, la serie $\sum(-1)^n x^n = 1/(1+x)$ tiende a $1/2$ cuando $x \to 1^-$, y sin embargo $\sum(-1)^n$ no converge. El valor $1/2$ es la suma de Abel, no la suma ordinaria.

(d) Cuándo sí hay recíproco. Los teoremas tauberianos dan condiciones extra que lo arreglan: el de Tauber pide $na_n \to 0$; el de Littlewood, sólo na_n acotada. Son mucho más difíciles que el de Abel, y esa asimetría de dificultad es característica del análisis.

Es el teorema que legitimó $\ln 2 = 1 - \frac{1}{2} + \frac{1}{3} - \cdots$ y $\frac{\pi}{4} = 1 - \frac{1}{3} + \frac{1}{5} - \cdots$. El recíproco es falso; los recíprocos con hipótesis (Tauber) fundaron la teoría tauberiana con la que se demostró el teorema de los números primos.

6.20.2. Criterios de Raabe y Gauss

Si $\frac{a_{n+1}}{a_n} = 1 - \frac{\beta}{n} + O(n^{-2})$: converge si $\beta > 1$, diverge si $\beta \leq 1$ (comparación con $n^{-\beta}$). Es el criterio para la serie binomial en el borde y las hipergeométricas —Gauss lo inventó para eso—. Con él, el catálogo queda completo: escala geométrica (cociente), polinómica (Raabe), logarítmica (condensación).

6.20.3. Productos infinitos

$\prod(1 + a_n)$ converge (a límite no nulo) absolutamente $\iff \sum |a_n| < \infty$ (logaritmos). El ejemplo supremo (Euler):

$$\frac{\sin \pi x}{\pi x} = \prod_{n \geq 1} \left(1 - \frac{x^2}{n^2}\right) :$$

el seno factorizado por sus raíces. Comparar el coeficiente de x^2 da $\sum \frac{1}{n^2} = \frac{\pi^2}{6}$ (la solución original de Basilea); $x = \frac{1}{2}$ da Wallis; y el molde $\zeta(s) = \prod_p (1 - p^{-s})^{-1}$ (factorización única hecha análisis) abre la teoría analítica de números: la hipótesis de Riemann vive en esta fórmula.

6.20.4. Sumabilidad de series divergentes

Cesàro: límite de los promedios de sumas parciales ($\sum(-1)^n \mapsto \frac{1}{2}$). *Abel*: límite radial. Ambos regulares (extienden la suma ordinaria) y jerárquicos: convergencia $\Rightarrow$ Cesàro $\Rightarrow$ Abel. No es un juego: Fejér dice que toda serie de Fourier de una continua es Cesàro-sumable a la función —la sumabilidad es la noción correcta para Fourier—, y las regularizaciones de la física cuántica ($-\frac{1}{12}$ del efecto Casimir) son sumabilidades bien definidas.

6.20.5. La exponencial construida desde cero

Defínase $\exp(x) = \sum \frac{x^n}{n!}$: de la serie salen $\exp' = \exp$, $\exp(x + y) = \exp x \exp y$ (Cauchy), positividad, monotonía, biyectividad sobre $(0, \infty)$; ln es su inversa con $\ln' = \frac{1}{x}$; $e := \exp(1)$ coincide con $(1 + \frac{1}{n})^n$. Seno y coseno por series (o Re, Im de e^{ix}), π *definido* como doble del primer cero de cos. Ni una función del libro apoyada en un dibujo: todo cuelga de series y convergencia uniforme.

6.20.6. El teorema de aproximación de Weierstrass

Teorema 6.112. *Toda continua en $[a, b]$ es límite uniforme de polinomios.*

Idea constructiva de Bernstein. $B_n f(x) = \sum_k f(\frac{k}{n})\binom{n}{k}x^k(1-x)^{n-k}$ es la esperanza de $f(\frac{S_n}{n})$ con S_n binomial; la ley de los grandes números concentra $\frac{S_n}{n}$ en x y la continuidad uniforme (Heine) convierte concentración en convergencia uniforme. □

Taylor aproxima localmente lo muy suave; Bernstein globalmente cualquier continua. Las curvas de Bézier del diseño gráfico son polinomios de Bernstein: cada letra en pantalla es este teorema.

6.21. Ronda de temarios: cierre curricular de Cálculo

Cotejando con el temario oficial de Cálculo Diferencial e Integral II (Facultad de Ciencias, UNAM), cerramos los últimos puntos que aún no habíamos demostrado.

6.21.1. Las tres formas del resto de Taylor

Ya probamos el resto de Lagrange. El temario pide también las otras dos formas, y con razón: cada una gana batallas que las otras pierden.

Teorema 6.113 (Resto en forma integral)**.** *Si $f \in C^{n+1}$ en un intervalo que contiene a a y x, entonces*

$$R_n(x) = f(x) - \sum_{k=0}^{n} \frac{f^{(k)}(a)}{k!}(x-a)^k = \frac{1}{n!}\int_a^x (x-t)^n f^{(n+1)}(t)\,\mathrm{d}t.$$

Demostración. Inducción sobre n. Para $n=0$ es el teorema fundamental: $f(x)-f(a)=\int_a^x f'(t)\,\mathrm{d}t$. Paso inductivo: integrando por partes con $u = f^{(n+1)}(t)$, $\mathrm{d}v = \frac{(x-t)^n}{n!}\mathrm{d}t$ (de modo que $v = -\frac{(x-t)^{n+1}}{(n+1)!}$),

$$\frac{1}{n!}\int_a^x (x-t)^n f^{(n+1)}\,\mathrm{d}t = \frac{f^{(n+1)}(a)}{(n+1)!}(x-a)^{n+1} + \frac{1}{(n+1)!}\int_a^x (x-t)^{n+1} f^{(n+2)}\,\mathrm{d}t,$$

que es exactamente el término $n{+}1$ de Taylor más el resto siguiente. □ □

Corolario 6.114 (Formas de Lagrange y de Cauchy)**.** *Existen ξ, η entre a y x tales que*

$$R_n(x) = \frac{f^{(n+1)}(\xi)}{(n+1)!}(x-a)^{n+1} \quad \textit{(Lagrange)}, \qquad R_n(x) = \frac{f^{(n+1)}(\eta)}{n!}(x-\eta)^n(x-a) \quad \textit{(Cauchy)}.$$

Demostración. Ambas salen del teorema del valor medio *ponderado* para integrales (si $g \geq 0$, $\int fg = f(\xi)\int g$, consecuencia del teorema del valor intermedio aplicado a f continua): con peso $g(t) = \frac{(x-t)^n}{n!}$ (que no cambia de signo entre a y x) se obtiene Lagrange, pues $\int_a^x \frac{(x-t)^n}{n!}\mathrm{d}t = \frac{(x-a)^{n+1}}{(n+1)!}$; con peso $g \equiv 1$ se obtiene Cauchy. □ □

Ejemplo 6.115 (Por qué hace falta la forma de Cauchy)**.** Para $\ln(1+x)=\sum(-1)^{k+1}\frac{x^k}{k}$ con $-1<x<0$, el resto de Lagrange da $|R_n|\le\frac{|x|^{n+1}}{(n+1)(1+x)^{n+1}}$, que *no* tiende a cero si $x<-\frac12$ (el denominador $(1+x)^{n+1}$ se hunde). La forma de Cauchy da

$$|R_n|=\frac{|x-\eta|^n\,|x|}{(1+\eta)^{n+1}}\le\frac{|x|}{1+x}\left(\frac{|x-\eta|}{1+\eta}\right)^n,\qquad\text{y}\quad\frac{|x-\eta|}{1+\eta}<|x|$$

para η entre x y 0 (comprobación directa: equivale a $|x|(1+\eta)>\eta-x$ con $x<\eta\le0$), de modo que $R_n\to0$ en *todo* $(-1,0)$. La convergencia de la serie del logaritmo en todo su intervalo depende de elegir bien la forma del resto: no son adornos, son herramientas con dominios de eficacia distintos.

6.21.2. La irracionalidad de π (Niven, 1947)

Probamos que e es irracional; la demostración de Niven para π cabe en una página usando solo herramientas de este capítulo.

Teorema 6.116. *π es irracional.*

Demostración. Supongamos $\pi=\frac{a}{b}$ con $a,b\in\mathbb{N}$. Para n por elegir, sea

$$f(x)=\frac{x^n(a-bx)^n}{n!}.$$

Dos observaciones. (1) f y todas sus derivadas toman valores *enteros* en 0 y en π: desarrollando, $f(x)=\frac{1}{n!}\sum_{k=n}^{2n}c_kx^k$ con $c_k\in\mathbb{Z}$, así que $f^{(j)}(0)=0$ si $j<n$ y $f^{(j)}(0)=\frac{j!}{n!}c_j\in\mathbb{Z}$ si $j\ge n$ (pues $n!\mid j!$); y como $f(\pi-x)=f(x)$ (simetría de $x^n(a-bx)^n$ respecto de $x=\frac{a}{2b}=\frac{\pi}{2}$), lo mismo vale en π. (2) Sea

$$F=f-f^{(2)}+f^{(4)}-\cdots+(-1)^nf^{(2n)}.$$

Entonces $\frac{\mathrm{d}}{\mathrm{d}x}\big(F'\sin x-F\cos x\big)=(F^{(2)}+F)\sin x=f\sin x$ (la suma telescopia, pues $f^{(2n+2)}=0$), luego

$$\int_0^\pi f(x)\sin x\,\mathrm{d}x=\big[F'\sin x-F\cos x\big]_0^\pi=F(\pi)+F(0)\in\mathbb{Z}.$$

Pero en $(0,\pi)$: $0<f(x)\sin x<\frac{(\pi a)^n}{n!}$, así que $0<\int_0^\pi f\sin x\,\mathrm{d}x<\pi\cdot\frac{(\pi a)^n}{n!}\xrightarrow{n\to\infty}0$ (el factorial gana a la exponencial, como probamos en sucesiones). Para n grande, el entero $F(\pi)+F(0)$ estaría estrictamente entre 0 y 1: absurdo. □ □

Ejemplo 6.117 (Irracionalidad y trascendencia)**. (a) La idea de la demostración.** Se supone $\pi=p/q$ y se construye $f(x)=\frac{x^n(p-qx)^n}{n!}$; la integral $\int_0^\pi f(x)\sin x\,\mathrm{d}x$ resulta ser un entero positivo y a la vez menor que 1 para n grande, porque el factorial del denominador acaba ganando. Contradicción.

(b) Cuánto se sabe y cuánto no. Son irracionales π, e, $\ln 2$, $\sqrt{2}$ y $\zeta(3)$ (Apéry, 1978). Se ignora si lo son $\pi+\mathrm{e}$, $\pi\mathrm{e}$, $\zeta(5)$ o la constante de Euler–Mascheroni $\gamma=0{,}577216$. La frontera del conocimiento está sorprendentemente cerca.

(c) Trascendencia. π es además trascendente (Lindemann, 1882), lo que es mucho más fuerte: no es raíz de ningún polinomio con coeficientes racionales. De ahí sale que la cuadratura del círculo con regla y compás es imposible, cerrando un problema abierto desde la Grecia clásica.

(d) Una fuente de irracionales. Todo número cuyo desarrollo decimal no sea periódico a partir de un punto es irracional. La constante de Champernowne $0{,}123456789101112\ldots$ lo es por construcción, y de hecho es trascendente y normal en base diez.

Observación 6.118. La estructura es idéntica a la de la irracionalidad de e: construir una cantidad que *debe* ser un entero positivo y *debe* ser menor que 1. Es la plantilla de casi todas las pruebas elementales de irracionalidad, y la elección del polinomio f (con ceros de orden alto en ambos extremos) es el germen de los aproximantes con que Hermite probó la trascendencia de e (capítulo 26).

6.21.3. Trapecio y Simpson con cotas de error demostradas

Ya usamos las reglas; ahora las demostramos, porque sin cota de error un método numérico es una superstición.

Teorema 6.119 (Error del trapecio)**.** *Si $f \in C^2[x_0, x_1]$ y $h = x_1 - x_0$, entonces*

$$\int_{x_0}^{x_1} f - \frac{h}{2}\big(f(x_0) + f(x_1)\big) = -\frac{h^3}{12} f^{(2)}(\xi) \quad \text{para algún } \xi \in (x_0, x_1).$$

En n subintervalos iguales de $[a,b]$: error total $= -\frac{(b-a)h^2}{12} f^{(2)}(\xi)$.

Demostración. Sea $w(t) = (t - x_0)(x_1 - t) \geq 0$. Integrando por partes dos veces se comprueba la identidad

$$\int_{x_0}^{x_1} w(t)\, f^{(2)}(t)\,\mathrm{d}t = 2\left[\frac{h}{2}\big(f(x_0) + f(x_1)\big) - \int_{x_0}^{x_1} f\right]$$

(en la primera integración el término de borde muere porque w se anula en los extremos; en la segunda, $w' = x_0 + x_1 - 2t$ evaluado en los extremos produce exactamente la regla del trapecio). Como $w \geq 0$, el valor medio ponderado da $\int w f^{(2)} = f^{(2)}(\xi) \int_{x_0}^{x_1} w = f^{(2)}(\xi)\,\frac{h^3}{6}$, de donde la fórmula. Sumando sobre los n subintervalos y promediando los $f^{(2)}(\xi_i)$ (teorema del valor intermedio otra vez) sale la forma global con $h = \frac{b-a}{n}$. □ □

Teorema 6.120 (Error de Simpson)**.** *Si $f \in C^4$, en un par de subintervalos $[m-h, m+h]$:*

$$\int_{m-h}^{m+h} f - \frac{h}{3}\big(f(m-h) + 4f(m) + f(m+h)\big) = -\frac{h^5}{90} f^{(4)}(\xi); \qquad \textit{global: } -\frac{(b-a)h^4}{180} f^{(4)}(\xi).$$

Idea de la demostración. Fíjese el punto medio m y estúdiese el error como función del semiancho: $G(h) = \int_{m-h}^{m+h} f - \frac{h}{3}\big(f(m-h) + 4f(m) + f(m+h)\big)$. Derivando tres veces respecto de h (regla de Leibniz del capítulo) todo colapsa a

$$G^{(3)}(h) = -\frac{h}{3}\big(f^{(3)}(m+h) - f^{(3)}(m-h)\big) = -\frac{2h^2}{3} f^{(4)}(\zeta_h)$$

(valor medio en el último paso). Integrando tres veces hacia atrás, con $G(0) = G'(0) = G^{(2)}(0) = 0$, se recupera $G(h) = -\frac{h^5}{90} f^{(4)}(\xi)$. El exponente 5 (y no 4) refleja lo que ya sabíamos: Simpson es exacto para cúbicas por simetría, así que el primer término de error involucra la *cuarta* derivada. □ □

Ejemplo 6.121. Para $\int_0^1 e^{-x^2}\mathrm{d}x$, con $|f^{(4)}| \le 12$ en $[0,1]$: Simpson con $h = \frac{1}{4}$ *garantiza* error $\le \frac{(1/4)^4 \cdot 12}{180} \approx 2.6 \times 10^{-4}$; el error real observado (3×10^{-5}) es menor, como corresponde a una cota de peor caso. Para garantizar 10^{-8}: $h^4 \le \frac{180 \cdot 10^{-8}}{12}$ da $h \approx 0.02$, unas 50 evaluaciones; el trapecio (con $|f^{(2)}| \le 2$) necesitaría $h \le \sqrt{6 \times 10^{-8}} \approx 2.4 \times 10^{-4}$: unas cuatro mil. Ese cociente de precios es la razón de existir de los métodos de orden alto, y la pareja de cotas h^2 vs. h^4 es el primer ejemplo del *orden de convergencia* que gobierna el capítulo 12 numérico (Euler vs. RK4: los mismos exponentes por la misma razón, Simpson incluido en los pesos de RK4).

6.22. Cierre definitivo: Beta y Stirling demostrado

Dos piezas finales para que en este capítulo no quede tema con deuda: la función Beta (la hermana de Γ que organiza todas las integrales trigonométricas) y la fórmula de Stirling con demostración completa.

6.22.1. La función Beta

Definición 6.122. Para $p, q > 0$:

$$B(p,q) = \int_0^1 x^{p-1}(1-x)^{q-1}\,\mathrm{d}x.$$

Teorema 6.123. $B(p,q) = \dfrac{\Gamma(p)\,\Gamma(q)}{\Gamma(p+q)}$.

Demostración. El truco es calcular $\Gamma(p)\Gamma(q)$ como integral doble y cambiar variables:

$$\Gamma(p)\Gamma(q) = \int_0^\infty\!\!\int_0^\infty s^{p-1}t^{q-1}e^{-s-t}\,\mathrm{d}s\,\mathrm{d}t.$$

Con $s = ux$, $t = u(1-x)$ (masa total $u = s+t$ y proporción x; jacobiano u, cálculo directo):

$$= \int_0^\infty\!\!\int_0^1 (ux)^{p-1}\big(u(1-x)\big)^{q-1}e^{-u}\,u\,\mathrm{d}x\,\mathrm{d}u = \underbrace{\int_0^\infty u^{p+q-1}e^{-u}\mathrm{d}u}_{\Gamma(p+q)} \cdot \underbrace{\int_0^1 x^{p-1}(1-x)^{q-1}\mathrm{d}x}_{B(p,q)}.$$

La integral doble se separa porque el cambio de variables desacopla escala y proporción. □ □

Ejemplo 6.124 (Todas las integrales trigonométricas de una vez)**.** Con $x = \sin^2\theta$: $B(p,q) = 2\int_0^{\pi/2} \sin^{2p-1}\theta\cos^{2q-1}\theta\,\mathrm{d}\theta$. Así,

$$\int_0^{\pi/2} \sin^m\theta\cos^n\theta\,\mathrm{d}\theta = \frac{1}{2}B\Big(\frac{m+1}{2},\frac{n+1}{2}\Big) = \frac{\Gamma\big(\frac{m+1}{2}\big)\Gamma\big(\frac{n+1}{2}\big)}{2\,\Gamma\big(\frac{m+n}{2}+1\big)}\ :$$

las fórmulas de reducción que dedujimos por partes, empaquetadas en una sola. Con $m = n = 0$ recuperamos $\Gamma(\frac{1}{2})^2 = \pi$ (¡otra vez la gaussiana!); con $q = \frac{1}{2}$, el volumen de la bola n-dimensional del capítulo 9 se reduce a una recursión de Betas; y $B(p, 1-p) = \Gamma(p)\Gamma(1-p) = \frac{\pi}{\sin \pi p}$ conecta con la fórmula de reflexión del capítulo 30. Beta es el pegamento entre las integrales de este capítulo y las funciones especiales de los siguientes.

6.22.2. La fórmula de Stirling, demostrada

Teorema 6.125 (Stirling)**.**

$$n! \sim \sqrt{2\pi n}\left(\frac{n}{e}\right)^n \qquad \textit{(cociente} \to 1\textit{)}.$$

Demostración. Paso 1 (la forma, sin la constante). Sea $a_n = \dfrac{n!}{n^{n+1/2}e^{-n}}$. Entonces

$$\ln \frac{a_n}{a_{n+1}} = \left(n + \frac{1}{2}\right)\ln\left(1 + \frac{1}{n}\right) - 1.$$

La serie de $\ln \frac{1+x}{1-x}$ con $x = \frac{1}{2n+1}$ (¡serie de potencias del capítulo!) da

$$\left(n + \frac{1}{2}\right)\ln\left(1 + \frac{1}{n}\right) = 1 + \frac{1}{3(2n+1)^2} + \frac{1}{5(2n+1)^4} + \cdots,$$

así que $0 < \ln \frac{a_n}{a_{n+1}} < \frac{1}{3(2n+1)^2-3} = \frac{1}{12n(n+1)} = \frac{1}{12n} - \frac{1}{12(n+1)}$ (geométrica mayorante y fracciones parciales). Conclusión: a_n es decreciente y $a_n e^{-1/12n}$ creciente: ambas convergen al mismo límite $C > 0$ (monótona acotada: el axioma del supremo trabajando). Ya sabemos $n! \sim C\sqrt{n}\,(n/e)^n$, con el error atrapado: $C < a_n < Ce^{1/12n}$.

Paso 2 (la constante, vía Wallis). El producto de Wallis (ya demostrado) dice $\frac{2^{4n}(n!)^4}{((2n)!)^2(2n+1)} \to \frac{\pi}{2}$. Sustituyendo $n! \approx C\sqrt{n}(n/e)^n$ y $(2n)! \approx C\sqrt{2n}(2n/e)^{2n}$: las potencias de n y de e se cancelan en masa y queda $\frac{C^2}{4n+2} \cdot 2n \to \frac{C^2}{4} = \frac{\pi}{2} : C = \sqrt{2\pi}$. □ □

Ejemplo 6.126 (Stirling, con números)**. (a) La calidad de la aproximación.** El cociente entre $n!$ y $\sqrt{2\pi n}(n/\mathrm{e})^n$ vale 1,084438 para $n = 1$, 1,016784 para $n = 5$, 1,008365 para $n = 10$, 1,004175 para $n = 20$ y 1,000834 para $n = 100$. El error relativo se comporta como $1/(12n)$, de modo que ya con $n = 10$ se baja del uno por ciento.

(b) Un caso concreto. $10! = 3\,628\,800$ y la aproximación da 3 598 695,618741: la diferencia es de unas treinta mil unidades sobre tres millones y medio.

(c) La versión con corrección. Multiplicando por $\mathrm{e}^{1/(12n)}$ el error cae drásticamente. Con $n = 10$, $\mathrm{e}^{1/120} = 1{,}008368$, que reproduce el cociente observado hasta la quinta cifra.

(d) Para qué se usa. Para estimar coeficientes binomiales ($\binom{2n}{n} \sim 4^n/\sqrt{\pi n}$), para deducir el teorema de De Moivre–Laplace, para la entropía en física estadística y para contar particiones. Siempre que un factorial aparece dentro de un límite, Stirling es la manera de sacarlo.

Observación 6.127. La demostración deja un regalo cuantitativo: $\sqrt{2\pi n}(n/e)^n < n! < \sqrt{2\pi n}(n/e)^n e^{1/12n}$ *para todo* n, no solo asintóticamente. Con $n = 10$: las cotas dan 3 598 696 <

$10! = 3\,628\,800 < 3\,628\,810$: la superior es finísima (siete cifras) y la inferior yerra por menos del 1 %, exactamente el $e^{-1/120}$ que predice el teorema. De aquí viven la entropía de Boltzmann–Shannon ($\ln n! \approx n\ln n - n$), el análisis de algoritmos ($\log_2 n!$ comparaciones para ordenar) y el teorema del límite central visto por monedas: $\binom{2n}{n}2^{-2n} \sim \frac{1}{\sqrt{\pi n}}$, la gaussiana emergiendo del triángulo de Pascal. El $\sqrt{2\pi}$ que apareció en la gaussiana, en Fourier y aquí es la misma constante contando lo mismo: no hay coincidencias, hay un solo análisis.

6.23. Segunda ronda: más teoría y ejemplos a fondo

6.23.1. Integrales de Frullani

Teorema 6.128. *Si f es continua con límites $f(0)$ y $f(\infty)$:*

$$\int_0^\infty \frac{f(ax)-f(bx)}{x}\,\mathrm{d}x = (f(0)-f(\infty))\ln\frac{b}{a}.$$

Idea. Escribir el integrando como $\int_b^a f'(tx)\,\mathrm{d}t$, intercambiar integrales (Fubini) y usar $\int_0^\infty f'(tx)\mathrm{d}x = \frac{f(\infty)-f(0)}{t}$. □

Cosecha: $\int_0^\infty \frac{e^{-ax}-e^{-bx}}{x}\mathrm{d}x = \ln\frac{b}{a}$ y su prima con cosenos. La lección: *muchas integrales difíciles son casos de un teorema con parámetros*; reconocer el patrón vale más que la fuerza bruta.

6.23.2. La media aritmético-geométrica de Gauss

$a_{n+1} = \frac{a_n+b_n}{2}$, $b_{n+1} = \sqrt{a_n b_n}$: encajadas por AM–GM y con diferencia que se *eleva al cuadrado* por paso: convergencia cuadrática. El límite $\mathrm{AGM}(a,b)$ esconde un tesoro (Gauss a los 22: "un nuevo campo del análisis"):

$$\frac{1}{\mathrm{AGM}(1,\sqrt{2})} = \frac{2}{\pi}\int_0^1 \frac{\mathrm{d}t}{\sqrt{1-t^4}}\,:$$

la AGM computa integrales elípticas (capítulo 12) y sostiene los algoritmos que calculan π duplicando dígitos por iteración (Brent–Salamin). Numérico con $(1,2)$: $(1.5, 1.41421) \to (1.45711, 1.45648) \to (1.45679, 1.45679)$: seis cifras en tres pasos.

6.23.3. Ejemplos finos

Ejemplo 6.129 (Un límite 1^∞ delicado)**.** $\lim_{x\to 0}\left(\frac{\sin x}{x}\right)^{1/x^2}$: el logaritmo es $\frac{1}{x^2}\ln(1-\frac{x^2}{6}+O(x^4)) \to -\frac{1}{6}$: límite $e^{-1/6}$. Patrón general: $\lim f^g = e^{\lim g(f-1)}$ cuando $f\to 1$, con Taylor hasta el orden que el cero pida.

Ejemplo 6.130 (Integral por autorreferencia)**.** $I = \int_0^{\pi/2}\ln\sin x\,\mathrm{d}x$: por $x \to \frac{\pi}{2}-x$, $I = \int \ln\cos$; sumando y usando $\sin 2x$:

$$2I = \int_0^{\pi/2}\ln\frac{\sin 2x}{2}\,\mathrm{d}x = I - \frac{\pi}{2}\ln 2 \;\Rightarrow\; I = -\frac{\pi}{2}\ln 2.$$

Una ecuación *para* la integral: el truco de hacerla aparecer en ambos lados, elevado a arte.

6.24. Coda: Stieltjes y la variación acotada

La *integral de Riemann–Stieltjes* $\int f\,\mathrm{d}g$ pesa los incrementos con otra función: sumas $\sum f(x_i^*)[g(x_i)-g(x_{i-1})]$. Existe para f continua y g de *variación acotada* (suma de incrementos absolutos acotada; equivalentemente, diferencia de dos crecientes). Si g es C^1, $\int f\,\mathrm{d}g = \int fg'$; pero g puede saltar: con g escalonada, la integral es una *suma* $\sum f(x_k)\Delta g_k$. Esta unificación suma–integral es el lenguaje natural de la probabilidad ($E[X]=\int x\,\mathrm{d}F$ con F la distribución, tenga densidad o sea discreta) y de la física de masas puntuales; su culminación es la integral de Lebesgue, primera parada del análisis posterior a este libro. Por partes vale simétricamente: $\int f\,\mathrm{d}g + \int g\,\mathrm{d}f = [fg]$: la fórmula de Abel del capítulo era su caso discreto.

6.25. Temas extra y ejemplos desarrollados

6.25.1. Sumas de Riemann como límites

Ejemplo 6.131.

$$\lim_n \Big(\frac{1}{n+1}+\cdots+\frac{1}{2n}\Big) = \lim \frac{1}{n}\sum_{k=1}^{n}\frac{1}{1+k/n} = \int_0^1 \frac{\mathrm{d}x}{1+x} = \ln 2,$$

y $\lim \frac{\sqrt[n]{n!}}{n} = e^{\int_0^1 \ln x\,\mathrm{d}x} = e^{-1}$. Patrón: $\frac{1}{n}$ afuera, $\frac{k}{n}$ adentro.

6.25.2. Trapecio y Simpson

$$T_n:\ |E| \le \frac{(b-a)h^2}{12}\,\text{máx}\,|f''|; \qquad S_n\ (\text{pesos } 1,4,2,4,\dots):\ |E| \le \frac{(b-a)h^4}{180}\,\text{máx}\,|f^{(4)}|.$$

Simpson integra exactos los cúbicos (simetría). $\int_0^1 e^{-x^2}$ con $n=4$: 0.746855 (error 3×10^{-5}).

6.25.3. La ley de Snell por Fermat

Minimizar $T(x) = \frac{\sqrt{a^2+x^2}}{v_1} + \frac{\sqrt{b^2+(d-x)^2}}{v_2}$ da $\frac{\sin\theta_1}{v_1} = \frac{\sin\theta_2}{v_2}$: la luz elige el camino de tiempo mínimo; una derivada contiene la óptica geométrica y anticipa el cálculo de variaciones.

6.25.4. Producto de Cauchy

Con convergencia absoluta, $(\sum a_n)(\sum b_n) = \sum c_n$, $c_n = \sum_k a_k b_{n-k}$. Ejemplo: $e^x e^y = e^{x+y}$ por el binomio —la ecuación funcional de la exponencial es un producto de Cauchy—.

6.25.5. Wallis y Stirling

$$\frac{\pi}{2} = \prod \frac{2n}{2n-1}\frac{2n}{2n+1}, \qquad n! \sim \sqrt{2\pi n}\big(\tfrac{n}{e}\big)^n.$$

Wallis por las integrales $\int \sin^n$; Stirling ancla su constante en Wallis. Uso: $\binom{2n}{n}2^{-2n} \sim \frac{1}{\sqrt{\pi n}}$ (monedas al aire), $\ln n! \approx n\ln n - n$ (algoritmos, entropía).

6.26. Seno, coseno y π, construidos desde cero

La exponencial se construyó desde su serie, sin apoyarse en nada previo. Con las trigonométricas hemos hecho trampa: se han usado sus desarrollos para deducir la fórmula de Euler, y esos desarrollos se obtuvieron derivando funciones definidas *geométricamente*, es decir, apelando a la longitud de arco de una circunferencia, que a su vez se define con una integral que usa el seno. El círculo lógico es real y hay que romperlo. Se rompe empezando por las series.

Definición 6.132.

$$\cos x = \sum_{n=0}^{\infty} \frac{(-1)^n x^{2n}}{(2n)!}, \qquad \sin x = \sum_{n=0}^{\infty} \frac{(-1)^n x^{2n+1}}{(2n+1)!}.$$

Ambas convergen absolutamente para todo x (cociente de términos $\to 0$), luego definen funciones enteras y se pueden derivar término a término.

De ahí salen, en este orden y sin usar nada más:

Teorema 6.133 (Las propiedades básicas, en cascada)**.**

1. $\cos' = -\sin$ *y* $\sin' = \cos$*, por derivación término a término.*
2. $\sin^2 x + \cos^2 x = 1$ *para todo* x*. En efecto, la derivada de* $\sin^2 + \cos^2$ *es* $2\sin\cos - 2\cos\sin = 0$*, luego la función es constante, y en* $x = 0$ *vale* 1*.*
3. $\cos(a+b) = \cos a \cos b - \sin a \sin b$ *y* $\sin(a+b) = \sin a \cos b + \cos a \sin b$*, por producto de Cauchy de las series (o, más rápido, derivando* $g(x) = \cos(a+b-x)\cos x - \sin(a+b-x)\sin x$ *y viendo que* $g' = 0$*).*
4. $\cos$ *es par y* $\sin$ *impar, por los exponentes de la serie.*

Teorema 6.134 (Existencia de π)**.** $\cos$ *tiene ceros positivos, y el menor de ellos es un número que llamamos* $\pi/2$*.*

Demostración. La serie de $\cos 2$ es alternada con términos decrecientes en módulo a partir del primero —en efecto $\frac{2^{2n}}{(2n)!}$ decrece desde $n = 1$—, luego la suma está *por debajo* de cualquier suma parcial que termine en un término negativo:

$$\cos 2 \;\le\; 1 - \frac{2^2}{2!} + \frac{2^4}{4!} = 1 - 2 + \frac{2}{3} = -\frac{1}{3} < 0.$$

Y $\cos 0 = 1 > 0$. Por Bolzano hay un cero en $(0,2)$. El conjunto de ceros es cerrado (por continuidad) y no vacío en $(0,2)$, luego tiene mínimo; llámese $\pi/2$ a ese mínimo. Numéricamente, $\pi/2 = 1{,}570\,796\,3\ldots$, coherente con que $\cos 2 = -0{,}416\,147 < 0$ efectivamente. □

Teorema 6.135 (Periodicidad, deducida)**.** *Con esa definición de* π*:*

$$\cos\tfrac{\pi}{2} = 0, \quad \sin\tfrac{\pi}{2} = 1, \quad e^{i\pi/2} = i, \quad e^{i\pi} = -1, \quad e^{2\pi i} = 1,$$

y en consecuencia $\sin$ *y* $\cos$ *son* 2π*-periódicas.*

Demostración. En $[0,\pi/2]$ el coseno es positivo salvo en el extremo, luego sin crece ahí; como $\sin 0 = 0$, se tiene $\sin(\pi/2) > 0$, y de $\sin^2+\cos^2 = 1$ con $\cos(\pi/2) = 0$ sale $\sin(\pi/2) = 1$. Entonces $e^{i\pi/2} = \cos\frac{\pi}{2} + i\sin\frac{\pi}{2} = i$, y elevando al cuadrado y a la cuarta, $e^{i\pi} = -1$ y $e^{2\pi i} = 1$. Finalmente, $e^{i(x+2\pi)} = e^{ix}e^{2\pi i} = e^{ix}$, y separando partes real e imaginaria queda la periodicidad de ambas funciones. □

Idea clave

Repárese en el orden. *Primero* se definen las funciones por sus series, *después* aparece π como el doble del primer cero del coseno, y *al final* se demuestra la periodicidad. En ningún momento se ha hablado de circunferencias, ángulos ni longitudes de arco. La geometría llega después, como interpretación: la aplicación $t \mapsto (\cos t, \sin t)$ recorre la circunferencia unidad —porque $\cos^2+\sin^2 = 1$— a velocidad constante 1 —porque $\|(-\sin t, \cos t)\| = 1$—, luego t *es* la longitud de arco y 2π es el perímetro. Ahora la fórmula de Euler ya no es circular: es un teorema.

Observación 6.136. Merece la pena ver qué otras definiciones de π producen el mismo número, porque la coincidencia no es evidente:

$$\pi = 2\cdot(\text{primer cero de }\cos) = 4\arctan 1 = 2\int_{-1}^{1}\frac{\mathrm{d}x}{\sqrt{1-x^2}} = \sqrt{6\sum_{n\geq 1} n^{-2}}.$$

La segunda y la tercera necesitan el teorema fundamental y la inversa del seno; la cuarta es el problema de Basilea. Que las cuatro den 3,141 592 653 6… es un pequeño milagro que solo se aprecia cuando se ha visto que la primera basta para construirlo todo.

6.27. Tres piezas que faltaban: Bonnet, medida nula y la integral de Stieltjes

El segundo teorema del valor medio

El criterio de Dirichlet para integrales oscilantes se enunció sin demostración. Su prueba necesita una herramienta que es el análogo continuo exacto de la sumación de Abel.

Teorema 6.137 (Bonnet)**.** *Si f es integrable en $[a,b]$ y g es monótona decreciente y no negativa, existe $\xi \in [a,b]$ con*

$$\int_a^b f(x)g(x)\,\mathrm{d}x = g(a)\int_a^{\xi} f(x)\,\mathrm{d}x.$$

La demostración es sumación de Abel aplicada a las sumas de Riemann y paso al límite: la monotonía de g hace que los pesos sean no negativos y la suma quede entre el mínimo y el máximo de las integrales parciales de f.

Ejemplo 6.138 (La integral de Dirichlet, ahora sí)**. (a) Converge.** Para $1 \leq R < S$, aplicando Bonnet con $g(x) = 1/x$:

$$\Big|\int_R^S \frac{\sin x}{x}\mathrm{d}x\Big| = \Big|\frac{1}{R}\int_R^{\xi}\sin x\,\mathrm{d}x\Big| = \frac{|\cos R - \cos\xi|}{R} \leq \frac{2}{R}.$$

El criterio de Cauchy queda satisfecho, luego $\int_1^\infty \frac{\sin x}{x}\mathrm{d}x$ converge; su valor, junto con el trozo $[0,1]$, es $\pi/2$.

(b) No converge absolutamente. En cada intervalo $[(k-1)\pi, k\pi]$ se tiene $\int |\sin x| = 2$ y $1/x \geq 1/(k\pi)$, luego

$$\int_0^{N\pi} \frac{|\sin x|}{x}\mathrm{d}x \;\geq\; \frac{2}{\pi}\sum_{k=1}^{N}\frac{1}{k} \;\longrightarrow\; \infty.$$

Con $N = 1000$ la cota inferior ya vale 4,77 y sigue creciendo como $\frac{2}{\pi}\log N$. La integral es condicionalmente convergente, exactamente igual que la serie armónica alternada.

Medida nula y el criterio definitivo

Hemos citado el criterio de Lebesgue para clasificar funciones sin haber definido su ingrediente central.

Definición 6.139 (Medida nula)**.** $E \subset \mathbb{R}$ tiene **medida nula** si para todo $\varepsilon > 0$ existe una familia *numerable* de intervalos abiertos que lo recubre con longitud total menor que ε.

Teorema 6.140. *Todo conjunto numerable tiene medida nula; una unión numerable de conjuntos de medida nula tiene medida nula; y ningún intervalo de longitud positiva la tiene.*

Lo primero se prueba cubriendo el punto n-ésimo con un intervalo de longitud $\varepsilon 2^{-n-1}$, cuya suma es ε. Lo segundo, con el mismo truco aplicado a los recubrimientos. Y lo tercero es el teorema de Heine–Borel disfrazado.

Teorema 6.141 (Criterio de Lebesgue)**.** *Una función acotada en $[a,b]$ es Riemann-integrable si y solo si su conjunto de discontinuidades tiene medida nula.*

Ejemplo 6.142 (Las tres funciones patológicas, clasificadas de un vistazo)**.**

función	discontinuidades	integrable
Dirichlet ($\mathbf{1}_{\mathbb{Q}}$)	todo $[0,1]$	no
Thomae ($1/q$ en p/q)	$\mathbb{Q}\cap[0,1]$, numerable	sí, $\int = 0$
$\mathbf{1}_{\text{Cantor}}$	el Cantor, no numerable pero de medida 0	sí, $\int = 0$

La tercera línea es la interesante: el conjunto de Cantor no es numerable —tiene el cardinal del continuo— y sin embargo tiene medida nula, luego su indicadora sí es integrable. Ser pequeño en cardinal y ser pequeño en medida son cosas distintas, y el criterio de Lebesgue depende de la segunda.

Variación acotada y la integral de Stieltjes

Definición 6.143. La **variación total** de g en $[a,b]$ es

$$V_a^b(g) = \sup_P \sum_i |g(x_i) - g(x_{i-1})|,$$

sobre todas las particiones. Si es finita, g es de **variación acotada**.

Teorema 6.144 (Descomposición de Jordan)**.** *g es de variación acotada si y solo si es diferencia de dos funciones crecientes. Basta tomar $g = V(x) - (V(x) - g(x))$ con $V(x) = V_a^x(g)$, y comprobar que ambas crecen.*

Definición 6.145 (Integral de Riemann–Stieltjes)**.**

$$\int_a^b f \,\mathrm{d}g = \lim_{\|P\|\to 0} \sum_i f(\xi_i)\big[g(x_i) - g(x_{i-1})\big].$$

El límite existe si f es continua y g es de variación acotada, y entonces $\left|\int f \,\mathrm{d}g\right| \le \|f\|_\infty V_a^b(g)$.

Ejemplo 6.146 (Una parte continua y un salto, en la misma integral)**.** Sea $g(x) = x + \mathbf{1}_{[1/2,1]}(x)$ en $[0,1]$: crece con pendiente 1 y además da un salto de tamaño 1 en $x = 1/2$. Entonces

$$\int_0^1 x^2 \,\mathrm{d}g = \underbrace{\int_0^1 x^2 \,\mathrm{d}x}_{\text{parte continua}} + \underbrace{\left(\tfrac{1}{2}\right)^2 \cdot 1}_{\text{el salto}} = \frac{1}{3} + \frac{1}{4} = \frac{7}{12} = 0{,}583\,333.$$

Ahí está toda la gracia del invento: una sola integral recoge a la vez una densidad continua y una masa puntual. Si g es escalonada con saltos s_k en x_k, la integral se reduce a la suma $\sum f(x_k)s_k$; si g es derivable, a $\int fg'$. La suma y la integral dejan de ser operaciones distintas.

Observación 6.147. Ese es el motivo de que la integral de Stieltjes sea el lenguaje natural de la probabilidad: la esperanza $E[X] = \int x \,\mathrm{d}F(x)$ vale igual para variables discretas —donde F es escalonada— que para continuas —donde $F' = f$— sin necesidad de dos definiciones. Y es también el paso intermedio hacia Lebesgue: la medida de Lebesgue–Stieltjes asociada a g es la que asigna al intervalo $(a,b]$ el valor $g(b) - g(a)$. Además, la integración por partes vale en la forma simétrica

$$\int_a^b f \,\mathrm{d}g + \int_a^b g \,\mathrm{d}f = f(b)g(b) - f(a)g(a),$$

que es la sumación de Abel escrita en continuo.

6.28. Los criterios de convergencia, demostrados

Los criterios de series se enunciaron como un catálogo, y sólo el del cociente se demostró. Los demás salen todos del mismo principio —comparar con algo conocido— y sus demostraciones son cortas.

Teorema 6.148 (Criterio de comparación)**.** *Sean $0 \le a_n \le b_n$ para todo $n \ge N$. Si $\sum b_n$ converge, $\sum a_n$ converge; si $\sum a_n$ diverge, $\sum b_n$ diverge.*

Demostración. Las sumas parciales de una serie de términos no negativos forman una sucesión *creciente*, luego convergen si y sólo si están acotadas (teorema de la sucesión monótona). Si $\sum b_n = B < \infty$, entonces $\sum_{n\le M} a_n \le \sum_{n<N} a_n + B$ para todo M: acotada, luego convergente. La segunda afirmación es el contrarrecíproco. □

Teorema 6.149 (Comparación por paso al límite)**.** *Sean $a_n, b_n > 0$ con $\lim_{n\to\infty} \frac{a_n}{b_n} = L$.*

1. *Si $0 < L < \infty$, las dos series tienen el mismo carácter.*
2. *Si $L = 0$ y $\sum b_n$ converge, $\sum a_n$ converge.*
3. *Si $L = \infty$ y $\sum b_n$ diverge, $\sum a_n$ diverge.*

Demostración. (1) Por definición de límite con $\varepsilon = L/2$, existe N tal que $\frac{L}{2} < \frac{a_n}{b_n} < \frac{3L}{2}$ para $n \geq N$, es decir $\frac{L}{2} b_n < a_n < \frac{3L}{2} b_n$. Ambas desigualdades, con el teorema 6.148, dan las dos implicaciones. *(2)* Con $\varepsilon = 1$ se tiene $a_n < b_n$ para n grande. *(3)* Con la definición de límite infinito, $a_n > b_n$ para n grande. □

Teorema 6.150 (Criterio de la raíz, Cauchy)**.** *Sea $L = \limsup_{n\to\infty} \sqrt[n]{|a_n|}$. Si $L < 1$ la serie $\sum a_n$ converge absolutamente; si $L > 1$ diverge; si $L = 1$ no se concluye nada.*

Demostración. Si $L < 1$, tómese r con $L < r < 1$. Por definición de límite superior, existe N tal que $\sqrt[n]{|a_n|} < r$ para $n \geq N$, es decir $|a_n| < r^n$. Como $\sum r^n$ es una serie geométrica convergente, el teorema 6.148 concluye.

Si $L > 1$, hay infinitos n con $\sqrt[n]{|a_n|} > 1$, o sea $|a_n| > 1$: el término general no tiende a 0 y la serie no puede converger.

Para $L = 1$: $\sum \frac{1}{n}$ diverge y $\sum \frac{1}{n^2}$ converge, y en ambas $\sqrt[n]{a_n} \to 1$. □

Observación 6.151 (La raíz es estrictamente más fuerte que el cociente). Siempre se cumple $\liminf \frac{|a_{n+1}|}{|a_n|} \leq \liminf \sqrt[n]{|a_n|} \leq \limsup \sqrt[n]{|a_n|} \leq \limsup \frac{|a_{n+1}|}{|a_n|}$, de modo que si el criterio del cociente decide, el de la raíz también, y con el mismo valor. El recíproco es falso. Tómese la serie

$$\tfrac{1}{2} + \tfrac{1}{3} + \tfrac{1}{2^2} + \tfrac{1}{3^2} + \tfrac{1}{2^3} + \tfrac{1}{3^3} + \cdots,$$

es decir $a_{2k-1} = 2^{-k}$ y $a_{2k} = 3^{-k}$. Los cocientes consecutivos valen $\frac{a_{2k}}{a_{2k-1}} = \left(\frac{2}{3}\right)^k \to 0$ y $\frac{a_{2k+1}}{a_{2k}} = \frac{1}{2}\left(\frac{3}{2}\right)^k \to \infty$: el criterio del cociente no decide nada, porque su $\liminf$ es 0 y su $\limsup$ es ∞. En cambio $\sqrt[2k-1]{a_{2k-1}} \to 2^{-1/2}$ y $\sqrt[2k]{a_{2k}} \to 3^{-1/2}$, luego $\limsup \sqrt[n]{a_n} = 2^{-1/2} = 0{,}7071 < 1$ y la raíz concluye sin dudar (la serie converge, claro: es la suma de dos geométricas). En la práctica se usa el cociente porque es más cómodo de calcular, y se recurre a la raíz cuando aparece una potencia n-ésima.

Teorema 6.152 (Criterio integral, Cauchy–Maclaurin)**.** *Sea $f : [1, \infty) \to [0, \infty)$ continua, decreciente, y $a_n = f(n)$. Entonces*

$$\sum_{n\geq 1} a_n \text{ converge} \quad \Longleftrightarrow \quad \int_1^\infty f(x)\,\mathrm{d}x \text{ converge},$$

y además, para todo N,

$$\int_{N+1}^\infty f \;\leq\; \sum_{n>N} a_n \;\leq\; \int_N^\infty f.$$

Demostración. Como f decrece, para $x \in [n, n+1]$ se tiene $f(n+1) \le f(x) \le f(n)$, e integrando en ese intervalo de longitud 1:

$$a_{n+1} \;\le\; \int_n^{n+1} f(x)\,\mathrm{d}x \;\le\; a_n.$$

Sumando desde $n = 1$ hasta M:

$$\sum_{n=2}^{M+1} a_n \;\le\; \int_1^{M+1} f \;\le\; \sum_{n=1}^{M} a_n.$$

Si la integral converge, la primera desigualdad acota las sumas parciales, que son crecientes: la serie converge. Si la serie converge, la segunda desigualdad acota las integrales parciales, que también crecen: la integral converge. Sumando desde $N+1$ en lugar de desde 1 se obtienen las dos cotas del error. □

Ejemplo 6.153 (Las p-series, de un plumazo, y una cota de error)**.** Con $f(x) = x^{-p}$: $\int_1^\infty x^{-p}\mathrm{d}x$ converge exactamente cuando $p > 1$. Luego $\sum n^{-p}$ converge si y sólo si $p > 1$, lo que incluye la divergencia de la serie armónica ($p = 1$) y la convergencia de $\sum 1/n^2$. Y la cota del error es utilísima: para $p = 2$,

$$\frac{1}{N+1} \;\le\; \sum_{n>N} \frac{1}{n^2} \;\le\; \frac{1}{N}.$$

Para calcular $\sum 1/n^2 = \pi^2/6 = 1{,}644934$ con error $< 10^{-3}$ harían falta $N = 1000$ términos, y sumando los mil primeros se obtiene 1,643935: el error es $9{,}995 \cdot 10^{-4}$, dentro de la cota $1/1000$ y casi igual a ella. Ese es el diagnóstico de una convergencia lenta, y la razón por la que Euler necesitó una idea, y no un ordenador, para encontrar el valor exacto.

Teorema 6.154 (Criterio de Leibniz para alternadas)**.** *Sea (b_n) decreciente con $b_n \to 0$. Entonces $\sum_{n\ge1}(-1)^{n+1}b_n$ converge, y si S es su suma y S_N la N-ésima suma parcial,*

$$|S - S_N| \;\le\; b_{N+1}.$$

Además S está siempre entre dos sumas parciales consecutivas.

Demostración. Considérense las sumas parciales de índice par:

$$S_{2m+2} - S_{2m} = b_{2m+1} - b_{2m+2} \;\ge\; 0,$$

luego (S_{2m}) crece; y las de índice impar: $S_{2m+1} - S_{2m-1} = -b_{2m} + b_{2m+1} \le 0$, luego (S_{2m+1}) decrece. Además $S_{2m} = S_{2m-1} - b_{2m} \le S_{2m-1}$, de modo que

$$S_2 \le S_4 \le \cdots \le S_{2m} \le S_{2m+1} \le \cdots \le S_3 \le S_1.$$

Las dos subsucesiones son monótonas y acotadas, luego convergen; y su diferencia es $S_{2m+1} - S_{2m} = b_{2m+1} \to 0$, así que tienen el mismo límite S, que por tanto es el límite de toda la sucesión. Como S queda encerrado entre S_N y S_{N+1} para todo N, se tiene

$$|S - S_N| \le |S_{N+1} - S_N| = b_{N+1}.$$ □

Ejemplo 6.155 (Una cota que se puede usar)**.** Para $\log 2 = \sum_{n\geq 1} \frac{(-1)^{n+1}}{n} = 1 - \frac{1}{2} + \frac{1}{3} - \cdots$, la cota dice que sumando N términos el error es a lo sumo $\frac{1}{N+1}$. Con $N = 10$: $S_{10} = 0{,}645635$ frente a $\log 2 = 0{,}693147$, con error $0{,}047512 \leq \frac{1}{11} = 0{,}0909$. Correcto, pero la lentitud es brutal: para tres decimales harían falta mil términos. Compárese con la serie $\log 2 = \sum_{k\geq 1} \frac{1}{k\,2^k}$, que da tres decimales con nueve términos. Convergencia condicional frente a convergencia geométrica: la misma constante, mundos distintos.

Teorema 6.156 (Cauchy–Hadamard)**.** *El radio de convergencia de* $\sum_{n\geq 0} c_n(x-x_0)^n$ *es*

$$R = \frac{1}{\limsup_{n\to\infty} \sqrt[n]{|c_n|}},$$

con los convenios $1/0 = \infty$ *y* $1/\infty = 0$*. La serie converge absolutamente para* $|x - x_0| < R$ *y diverge para* $|x - x_0| > R$*.*

Demostración. Aplíquese el criterio de la raíz (teorema 6.150) a los términos $a_n = c_n(x-x_0)^n$:

$$\limsup \sqrt[n]{|a_n|} = |x - x_0| \cdot \limsup \sqrt[n]{|c_n|} = \frac{|x - x_0|}{R}.$$

Ese límite es < 1 exactamente cuando $|x - x_0| < R$, y > 1 cuando $|x - x_0| > R$. □

Teorema 6.157 (Derivación término a término)**.** *Sea* (f_n) *una sucesión de funciones de clase* C^1 *en* $[a, b]$ *tal que*

- $f_n(x_0)$ *converge para algún* $x_0 \in [a, b]$*, y*
- (f_n') *converge* uniformemente *a una función* g *en* $[a, b]$*.*

Entonces (f_n) *converge uniformemente a una función* f *derivable, y* $f' = g$*. Es decir,*

$$\Big(\lim_n f_n\Big)' = \lim_n f_n'.$$

Demostración. Por el teorema fundamental del cálculo, $f_n(x) = f_n(x_0) + \int_{x_0}^{x} f_n'(t)\,\mathrm{d}t$. El primer sumando converge por hipótesis; y en el segundo se puede pasar al límite bajo la integral, porque la convergencia de f_n' es uniforme sobre un intervalo acotado:

$$\left|\int_{x_0}^{x} f_n' - \int_{x_0}^{x} g\right| \leq (b-a)\,\|f_n' - g\|_\infty \xrightarrow[n\to\infty]{} 0,$$

y la cota no depende de x, luego la convergencia de f_n es uniforme. Llamando f al límite,

$$f(x) = \lim_n f_n(x_0) + \int_{x_0}^{x} g(t)\,\mathrm{d}t.$$

Como g es continua —límite uniforme de continuas—, el teorema fundamental del cálculo da que f es derivable con $f' = g$. □

Cuidado

[La hipótesis uniforme va sobre las *derivadas*] Es el punto que más se confunde. Que $f_n \to f$ uniformemente *no* basta para derivar término a término: $f_n(x) = \frac{\sin(n^2 x)}{n}$ converge uniformemente a 0 (su norma del supremo es $1/n$), pero $f_n'(x) = n\cos(n^2 x)$ no converge a nada. Lo que hay que controlar uniformemente es la sucesión de derivadas; el control sobre las funciones puede reducirse a un solo punto.

Corolario 6.158 (Las series de potencias se derivan e integran término a término)**.** *Si* $f(x) = \sum_{n\geq 0} c_n (x-x_0)^n$ *tiene radio* $R > 0$*, entonces* f *es infinitamente derivable en* $(x_0 - R, x_0 + R)$ *y*

$$f'(x) = \sum_{n\geq 1} n c_n (x-x_0)^{n-1}, \qquad \int_{x_0}^{x} f = \sum_{n\geq 0} \frac{c_n}{n+1}(x-x_0)^{n+1},$$

y ambas series tienen el mismo radio R*. En consecuencia* $c_n = \frac{f^{(n)}(x_0)}{n!}$*: toda serie de potencias es la serie de Taylor de su suma.*

Demostración. Que el radio no cambia se sigue de Cauchy–Hadamard y de $\sqrt[n]{n} \to 1$: $\limsup \sqrt[n]{n|c_n|} = \limsup \sqrt[n]{|c_n|}$. Fijado $r < R$, la serie derivada converge uniformemente en $[x_0 - r, x_0 + r]$ por el criterio M de Weierstrass, comparando con $\sum n|c_n| r^{n-1} < \infty$. El teorema 6.157 aplicado a las sumas parciales —que son polinomios, luego C^1— da la primera fórmula; iterando se obtiene la derivabilidad infinita, y evaluando la n-ésima derivada en x_0 sobreviven sólo los términos con $n!\, c_n$. □

6.29. Series de potencias con todas las definiciones: radio, intervalo y qué pasa en los extremos

Ya usamos las series de potencias como herramienta y enunciamos el teorema del radio de convergencia. Falta lo elemental: escribir las definiciones con todas sus letras y, sobre todo, tratar en serio el punto que los estudiantes fallan sistemáticamente en los exámenes, que es qué ocurre en los dos extremos del intervalo. El teorema del radio no dice nada sobre ellos, y no es un descuido: es que allí puede pasar cualquier cosa, y las cuatro combinaciones posibles se dan de verdad.

Definición 6.159 (Serie de potencias, radio e intervalo de convergencia)**.** Una *serie de potencias* centrada en $a \in \mathbb{R}$ es una expresión de la forma

$$\sum_{n\geq 0} c_n\,(x-a)^n, \qquad c_n \in \mathbb{R}.$$

Su *radio de convergencia* es

$$R = \frac{1}{\limsup_{n\to\infty} \sqrt[n]{|c_n|}} \in [0, +\infty]$$

(fórmula de Cauchy–Hadamard, con los convenios $1/0 = +\infty$ y $1/+\infty = 0$). La serie converge absolutamente si $|x-a| < R$ y diverge si $|x-a| > R$. El *intervalo de convergencia* es el

conjunto de todos los $x \in \mathbb{R}$ en los que la serie converge: es siempre un intervalo centrado en a, de longitud $2R$, que puede incluir o no cada uno de sus extremos $a-R$ y $a+R$. Los cuatro casos $(a-R, a+R)$, $[a-R, a+R)$, $(a-R, a+R]$ y $[a-R, a+R]$ ocurren realmente.

Definición 6.160 (Función analítica)**.** Una función f definida en un abierto $U \subset \mathbb{R}$ es *analítica* (real-analítica) si para cada $a \in U$ existe $r > 0$ y una serie de potencias tales que

$$f(x) = \sum_{n \geq 0} c_n (x-a)^n \qquad \text{para todo } x \text{ con } |x-a| < r.$$

En ese caso, los coeficientes están forzados: derivando n veces y evaluando en a se obtiene $c_n = f^{(n)}(a)/n!$, de modo que la serie es necesariamente la serie de Taylor de f en a. Ser analítica es, entonces, "ser localmente igual a la suma de su propia serie de Taylor".

Regla: Hallar el radio y decidir los extremos

Dada $\sum a_n(x)$ con término general $a_n(x) = c_n(x-a)^n$:

1. **Aplica el cociente o la raíz al término general completo**, con la x dentro:

$$L(x) = \lim_{n\to\infty} \left| \frac{a_{n+1}(x)}{a_n(x)} \right| = |x-a| \cdot \lim_{n\to\infty} \left| \frac{c_{n+1}}{c_n} \right|,$$

   o bien, con la raíz,

$$L(x) = \lim_{n\to\infty} \sqrt[n]{|a_n(x)|} = |x-a| \cdot \limsup_{n\to\infty} \sqrt[n]{|c_n|}.$$

   Usa la raíz cuando aparezcan potencias n-ésimas (como $c_n = (3+\frac{1}{n})^n$) y el cociente cuando aparezcan factoriales. Si solo hay potencias pares o el exponente no es n, aplica el criterio al término general tal cual y despeja: no intentes leer los c_n de una fórmula que no los tiene.

2. **Despeja el radio de la condición $L(x) < 1$.** Esa desigualdad es siempre de la forma $|x-a| < R$, y ese R es el radio. Los casos extremos también cuentan: si el límite de los cocientes es 0, entonces $R = +\infty$ y la serie converge en toda la recta; si es $+\infty$, entonces $R = 0$ y la serie solo converge en $x = a$.

3. **Estudia aparte cada extremo.** Este es el paso que no se puede saltar y que ningún criterio general resuelve. Sustituye $x = a+R$ en la serie *original*: queda una serie *numérica*, y se decide con las herramientas de series numéricas —comparación, integral, Leibniz para las alternadas, o simplemente el criterio del término general si no tiende a cero—. Repite con $x = a-R$. Advertencia importante: en los extremos el criterio del cociente y el de la raíz dan $L = 1$ y por tanto *no concluyen*; volver a aplicarlos allí es tiempo perdido.

4. **Escribe el intervalo con los corchetes correctos.** Corchete cerrado en el extremo donde la serie numérica converge, paréntesis donde diverge. Y no confundas los dos tipos de convergencia: en un extremo donde la convergencia es solo condicional (por Leibniz), la convergencia *no* es absoluta, y eso importa si luego se quiere reordenar o multiplicar series.

5. **Verifica.** Comprueba el radio con un valor interior cómodo (la serie numérica resultante debe converger visiblemente) y con uno exterior (el término general no debe tender a cero). Si conoces la suma en forma cerrada, evalúa numéricamente unas cuantas sumas parciales y compara.

Ejemplo 6.161 (La serie que converge en un extremo y diverge en el otro)**.** Sea

$$\sum_{n\geq 1} \frac{x^n}{n}.$$

Radio. Con el cociente, $\left|\frac{a_{n+1}}{a_n}\right| = |x| \cdot \frac{n}{n+1} \to |x|$ (numéricamente, $\frac{n}{n+1}$ vale 0,9090909091 para $n = 10$, 0,9900990099 para $n = 100$ y 0,9999990000 para $n = 10^6$: el límite es 1 sin ambigüedad). La condición $|x| < 1$ da $R = 1$.

Extremo $x = 1$. La serie se convierte en $\sum \frac{1}{n}$, la serie armónica, que **diverge**. Numéricamente se ve el crecimiento logarítmico: las sumas parciales valen 5,187378 para $N = 100$, 9,787606 para $N = 10^4$, 14,392727 para $N = 10^6$ y 16,695311 para $N = 10^7$, ajustándose a $\ln N + \gamma$ (que vale 5,182386, 9,787556, 14,392726 y 16,695311 respectivamente: las tres últimas coinciden en seis cifras). Diverge, pero tan despacio que ningún ordenador lo vería sin la teoría.

Extremo $x = -1$. La serie se convierte en $\sum \frac{(-1)^n}{n}$, alternada con términos decrecientes en valor absoluto hacia cero: por el criterio de Leibniz **converge**, y su suma es $-\ln 2 = -0{,}6931471806$. Las sumas parciales lo confirman: $-0{,}6881721793$ para $N = 100$ y $-0{,}6931421806$ para $N = 10^5$. La convergencia es condicional, no absoluta, porque en valor absoluto reaparece la armónica.

Intervalo. $[-1, 1)$: cerrado a la izquierda, abierto a la derecha. Y esto no es una casualidad de esta serie: es la serie de $-\ln(1-x)$, que en $x = 1$ tiene una singularidad logarítmica real y en $x = -1$ vale $-\ln 2$, perfectamente finito. La verificación interior confirma la identidad: en $x = 0{,}9$ la suma parcial hasta $N = 4000$ da 2,302585093, y $-\ln(0{,}1) = 2{,}302585093$; en $x = -0{,}5$ da $-0{,}4054651081$ frente a $-\ln(1{,}5) = -0{,}4054651081$.

Ejemplo 6.162 (Lo mismo con el centro desplazado)**.** Sea $\displaystyle\sum_{n\geq 1} \frac{(x-2)^n}{n\,3^n}$. El cociente da $|x-2| \cdot \frac{n}{3(n+1)} \to \frac{|x-2|}{3}$, y la condición de ser menor que 1 da $R = 3$: el intervalo abierto es $(-1, 5)$. En $x = 5$ queda $\sum \frac{1}{n}$, divergente (la suma parcial hasta 10^6 vale 14,392727, la misma armónica de antes); en $x = -1$ queda $\sum \frac{(-1)^n}{n}$, convergente a $-0{,}6931472$ (parcial hasta 10^5: $-0{,}6931422$). El intervalo de convergencia es $[-1, 5)$. Comprobación interior: en $x = 3$ la suma vale 0,4054651081 y $-\ln(1 - \frac{1}{3}) = 0{,}4054651081$; en $x = 0$ vale $-0{,}5108256238$ y $-\ln(1 + \frac{2}{3}) = -0{,}5108256238$. Todo el ejemplo anterior trasladado y reescalado, que es exactamente lo que hace el cambio de centro.

Cuidado

[Derivable infinitas veces no implica analítica] Ya lo avisamos al presentar las series de potencias, pero conviene ver el tamaño real del desastre. La función

$$f(x)=\begin{cases} e^{-1/x^2}, & x\neq 0,\\ 0, & x=0,\end{cases}$$

es de clase C^∞ en toda la recta y cumple $f^{(n)}(0)=0$ para *todo* n: por inducción, cada derivada es de la forma $P(1/x)e^{-1/x^2}$ con P polinómico, y la exponencial aplasta a cualquier polinomio cuando $x\to 0$. Su serie de Taylor en el origen es, por tanto, la serie idénticamente nula, cuyo radio de convergencia es $R=+\infty$ y cuya suma es la función cero. Pero f no es cero en ningún entorno del origen: $f(1)=0{,}3678794$, $f(0{,}5)=0{,}01831564$, $f(0{,}2)=1{,}388794\cdot 10^{-11}$ y $f(0{,}1)=3{,}720076\cdot 10^{-44}$. La serie converge perfectamente y converge a la función equivocada en todos los puntos salvo uno.

Lo que mide el fenómeno es el *aplanamiento*: $f(x)$ tiende a cero más deprisa que cualquier potencia. En $x=0{,}1$, el cociente $f(x)/x^{20}$ vale $3{,}72\cdot 10^{-24}$, todavía minúsculo pese a dividir por 10^{-20}. Por eso f es indistinguible de la función nula desde el punto de vista de las derivadas en el origen, y por eso la implicación "$C^\infty \Rightarrow$ analítica" es falsa en $\mathbb{R}$. En $\mathbb{C}$ es verdadera —una función holomorfa es automáticamente analítica, capítulo 30—, y esa es una de las diferencias más profundas entre las dos teorías. Añádase, para calibrar cuán falsa es la implicación, el teorema de Borel: dada *cualquier* sucesión (a_n) de números reales, existe una función C^∞ con $f^{(n)}(0)=a_n$ para todo n, aunque la serie $\sum \frac{a_n}{n!}x^n$ tenga radio cero. Las funciones C^∞ son enormemente más abundantes que las analíticas; la rigidez de estas últimas —quedar determinadas por su comportamiento en un punto— es una propiedad excepcional, no la norma.

6.30. Ejercicios resueltos

Ejercicio 6.1. Demuestra por definición que $a_n=\dfrac{2n+1}{n+3}\to 2$, hallando el N correspondiente a cada ε.

Solución. $\left|\frac{2n+1}{n+3}-2\right|=\left|\frac{2n+1-2n-6}{n+3}\right|=\frac{5}{n+3}$. Dado $\varepsilon>0$, basta $\frac{5}{n+3}<\varepsilon$, es decir $n>\frac{5}{\varepsilon}-3$. Tomando $N=\lceil 5/\varepsilon\rceil$, para $n\geq N$ se cumple $|a_n-2|<\varepsilon$. ■

Ejercicio 6.2. Estudia la sucesión definida por recurrencia $a_1=1$, $a_{n+1}=\sqrt{2+a_n}$.

Solución. *Acotación*: por inducción, $a_n<2$ (si $a_n<2$, $a_{n+1}=\sqrt{2+a_n}<\sqrt{4}=2$). *Monotonía*: $a_{n+1}>a_n \iff 2+a_n>a_n^2 \iff (2-a_n)(1+a_n)>0$, cierto pues $a_n<2$. Creciente y acotada: converge (Teorema 6.27). El límite L satisface $L=\sqrt{2+L}$, es decir $L^2-L-2=(L-2)(L+1)=0$; como $L>0$, $L=2$. ■

Ejercicio 6.3. Calcula $\displaystyle\lim_{x\to 0}\frac{\tan x-\sin x}{x^3}$ y $\displaystyle\lim_{x\to 0^+}x^x$.

Solución. Primero: $\tan x = x + \frac{x^3}{3} + o(x^3)$ y $\sin x = x - \frac{x^3}{6} + o(x^3)$, luego el numerador es $\frac{x^3}{3} + \frac{x^3}{6} + o(x^3) = \frac{x^3}{2} + o(x^3)$ y el límite es $\boxed{\frac{1}{2}}$. Segundo: $x^x = e^{x \ln x}$ y $x \ln x \to 0$ (pues $\frac{\ln x}{1/x} \to 0$ por L'Hôpital en la forma $\frac{-\infty}{\infty}$). Por continuidad de e^t: $x^x \to e^0 = 1$. ■

Ejercicio 6.4. Demuestra que $e^x \geq 1 + x$ para todo $x \in \mathbb{R}$, y deduce la desigualdad AM–GM para dos números: $\sqrt{ab} \leq \frac{a+b}{2}$.

Solución. Sea $g(x) = e^x - 1 - x$: $g'(x) = e^x - 1$ es negativa para $x < 0$ y positiva para $x > 0$, así que g tiene mínimo global en 0 con $g(0) = 0$: $g \geq 0$. Para AM–GM, con $m = \frac{a+b}{2}$ aplicamos la desigualdad a $x = \frac{a}{m} - 1$ y $x = \frac{b}{m} - 1$:

$$e^{a/m-1} \geq \frac{a}{m}, \qquad e^{b/m-1} \geq \frac{b}{m}.$$

Multiplicando: $e^{(a+b)/m-2} \geq \frac{ab}{m^2}$. Pero $\frac{a+b}{m} = 2$, así que el exponente es 0: $1 \geq \frac{ab}{m^2}$, es decir $\sqrt{ab} \leq m$. ■

Ejercicio 6.5. Halla el rectángulo de mayor área inscrito en la elipse $\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1$ con lados paralelos a los ejes.

Solución. El vértice en el primer cuadrante es (x, y) con $y = b\sqrt{1 - x^2/a^2}$, y el área es $A(x) = 4xy = 4bx\sqrt{1 - x^2/a^2}$, $x \in [0, a]$. Derivando (mejor: maximizar $A^2 = 16b^2x^2(1 - x^2/a^2)$, con $u = x^2$: cuadrática $16b^2(u - u^2/a^2)$ con máximo en $u = a^2/2$): $x = \frac{a}{\sqrt{2}}$, $y = \frac{b}{\sqrt{2}}$, y el área máxima es $A = 4 \cdot \frac{a}{\sqrt{2}} \cdot \frac{b}{\sqrt{2}} = 2ab$. Como A es continua en el compacto $[0, a]$ y se anula en los extremos, este crítico interior es el máximo (Weierstrass). ■

Ejercicio 6.6. Calcula $\displaystyle\int e^x \cos x \, \mathrm{d}x$, $\displaystyle\int_0^1 \frac{\mathrm{d}x}{(1+x^2)^2}$ y $\displaystyle\int \frac{\mathrm{d}x}{x^2\sqrt{4-x^2}}$.

Solución. (a) Dos veces por partes: $I = \int e^x \cos x \, \mathrm{d}x = e^x \sin x - \int e^x \sin x \, \mathrm{d}x = e^x \sin x + e^x \cos x - I$, de donde $I = \frac{e^x(\sin x + \cos x)}{2} + C$.

(b) Con $x = \tan\theta$, $\mathrm{d}x = \sec^2\theta \, \mathrm{d}\theta$, $1 + x^2 = \sec^2\theta$:

$$\int_0^{\pi/4} \frac{\sec^2\theta}{\sec^4\theta} \mathrm{d}\theta = \int_0^{\pi/4} \cos^2\theta \, \mathrm{d}\theta = \Big[\frac{\theta}{2} + \frac{\sin 2\theta}{4}\Big]_0^{\pi/4} = \frac{\pi}{8} + \frac{1}{4}.$$

(c) Con $x = 2\sin\theta$: $\sqrt{4 - x^2} = 2\cos\theta$, $\mathrm{d}x = 2\cos\theta \, \mathrm{d}\theta$:

$$\int \frac{2\cos\theta \, \mathrm{d}\theta}{4\sin^2\theta \cdot 2\cos\theta} = \frac{1}{4}\int \csc^2\theta \, \mathrm{d}\theta = -\frac{\cot\theta}{4} + C = -\frac{\sqrt{4-x^2}}{4x} + C,$$

usando $\cot\theta = \frac{\cos\theta}{\sin\theta} = \frac{\sqrt{4-x^2}}{x}$. ■

Ejercicio 6.7. Determina si convergen: (a) $\displaystyle\int_1^\infty \frac{\ln x}{x^2} \mathrm{d}x$; (b) $\displaystyle\sum_{n\geq 2} \frac{1}{n \ln n}$; (c) $\displaystyle\sum_{n\geq 1} \frac{n^2}{2^n}$ (calcula además su suma).

Solución. (a) Por partes: $\int_1^R \frac{\ln x}{x^2}\mathrm{d}x = \big[-\frac{\ln x}{x}\big]_1^R + \int_1^R \frac{\mathrm{d}x}{x^2} = -\frac{\ln R}{R} + 1 - \frac{1}{R} \to 1$. Converge y vale 1.

(b) Criterio integral: $\int_2^\infty \frac{\mathrm{d}x}{x \ln x} = \big[\ln(\ln x)\big]_2^\infty = \infty$. **Diverge** (aunque sus términos van a cero más rápido que $\frac{1}{n}$; las escalas logarítmicas son traicioneras).

(c) Cociente: $\frac{(n+1)^2/2^{n+1}}{n^2/2^n} = \frac{1}{2}\big(1 + \frac{1}{n}\big)^2 \to \frac{1}{2} < 1$: converge. Para sumarla, partimos de $\sum_{n\geq 0} x^n = \frac{1}{1-x}$; derivando y multiplicando por x: $\sum nx^n = \frac{x}{(1-x)^2}$; repitiendo la operación: $\sum n^2x^n = \frac{x(1+x)}{(1-x)^3}$. En $x = \frac{1}{2}$:

$$\sum_{n\geq 1} \frac{n^2}{2^n} = \frac{\frac{1}{2} \cdot \frac{3}{2}}{(\frac{1}{2})^3} = 6.$$

■

Ejercicio 6.8. Encuentra el intervalo de convergencia *completo* (incluyendo bordes) de $\displaystyle\sum_{n\geq 1} \frac{(x-2)^n}{n\,3^n}$.

Solución. Cociente: $\left|\frac{(x-2)^{n+1}}{(n+1)3^{n+1}} \cdot \frac{n3^n}{(x-2)^n}\right| = \frac{n}{n+1}\cdot\frac{|x-2|}{3} \to \frac{|x-2|}{3}$. Converge si $|x-2|<3$: intervalo abierto $(-1,5)$, radio $R=3$. *Bordes*: en $x=5$ la serie es $\sum \frac{1}{n}$: diverge. En $x=-1$ es $\sum \frac{(-1)^n}{n}$: converge (Leibniz), condicionalmente. Intervalo de convergencia: $[-1,5)$. ■

Ejercicio 6.9. Aproxima $\displaystyle\int_0^1 e^{-x^2}\,\mathrm{d}x$ con error menor que 10^{-3}.

Solución. Sustituyendo en la serie de e^t: $e^{-x^2} = \sum_{n\geq 0} \frac{(-1)^n x^{2n}}{n!}$, e integrando término a término (lícito dentro del radio, aquí infinito):

$$\int_0^1 e^{-x^2}\,\mathrm{d}x = \sum_{n\geq 0} \frac{(-1)^n}{n!\,(2n+1)} = 1 - \frac{1}{3} + \frac{1}{10} - \frac{1}{42} + \frac{1}{216} - \cdots$$

Serie alternada con términos decrecientes: el error de truncar está acotado por el primer término omitido. Como $\frac{1}{5!\cdot 11} = \frac{1}{1320} < 10^{-3}$, bastan los términos hasta $n=4$:

$$1 - 0.33333 + 0.1 - 0.02381 + 0.00463 \approx 0.7475.$$

(Valor exacto: $0.746824\ldots$; el error real es $\approx 7\times 10^{-4}$, dentro de lo prometido.) ■

6.31. Ejercicios adicionales resueltos

Ejercicio 6.10. Demuestra con ε–δ que $\lim_{x\to 2} \frac{1}{x} = \frac{1}{2}$.

Solución. $\left|\frac{1}{x} - \frac{1}{2}\right| = \frac{|x-2|}{2|x|}$; con $|x-2|<1$: $x>1$, factor $<\frac{1}{2}$. $\delta = \min\{1, 2\varepsilon\}$ funciona. ■

Ejercicio 6.11. Calcula $\lim_n \sum_{k=1}^{n} \frac{k}{n^2+k}$.

Solución. Sándwich entre $\frac{n(n+1)/2}{n^2+n} = \frac{1}{2}$ y $\frac{n(n+1)/2}{n^2} = \frac{n+1}{2n}$: ambos $\to \frac{1}{2}$. ■

Ejercicio 6.12. Calcula $\int e^{2x}\cos 3x\,\mathrm{d}x$ con complejos.

Solución. $\mathrm{Re}\int e^{(2+3i)x}\mathrm{d}x = \mathrm{Re}\,\frac{e^{(2+3i)x}(2-3i)}{13} = \frac{e^{2x}(2\cos 3x + 3\sin 3x)}{13} + C$, sin partes. ■

Ejercicio 6.13. Calcula $\int \frac{\mathrm{d}x}{x^3+1}$.

Solución. $x^3+1 = (x+1)(x^2-x+1)$; fracciones parciales con $A=\frac{1}{3}$, $B=-\frac{1}{3}$, $C=\frac{2}{3}$; separando derivada del denominador y completando cuadrado:

$$\frac{1}{3}\ln|x+1| - \frac{1}{6}\ln(x^2-x+1) + \frac{1}{\sqrt{3}}\arctan\frac{2x-1}{\sqrt{3}} + C.$$

■

Ejercicio 6.14. ¿Convergen $\sum(\sqrt[n]{2}-1)$ y $\sum \frac{1}{(\ln n)^{\ln n}}$?

Solución. $\sqrt[n]{2}-1 \sim \frac{\ln 2}{n}$: diverge (comparación con la armónica). $(\ln n)^{\ln n} = n^{\ln\ln n} \geq n^2$ desde $n \geq e^{e^2}$: converge. ■

Ejercicio 6.15. Calcula $F(b) = \int_0^\infty e^{-x^2}\cos(2bx)\,\mathrm{d}x$.

Solución. Derivando bajo la integral y por partes: $F' = -2bF$, EDO con $F(0) = \frac{\sqrt{\pi}}{2}$: $F(b) = \frac{\sqrt{\pi}}{2}e^{-b^2}$ —la gaussiana es su propia transformada: el principio de incertidumbre—. ■

Ejercicio 6.16. Suma $\sum_{n\geq 1} \frac{1}{n(n+1)(n+2)}$.

Solución. Forma telescópica $\frac{1}{2}\left[\frac{1}{n(n+1)} - \frac{1}{(n+1)(n+2)}\right]$: suma $\frac{1}{2}\cdot\frac{1}{2} = \frac{1}{4}$. ■

6.32. Taller de ejemplos explicados: ¿converge o no una serie?

Una *serie* es una suma infinita, como $1+\frac{1}{2}+\frac{1}{4}+\cdots$. La gran pregunta es si esa suma infinita da un número finito (decimos que *converge*) o si se dispara sin límite (*diverge*). Aquí aprendemos a decidirlo con ejemplos, despacio.

6.32.1. Series geométricas: las más fáciles

Regla: serie geométrica

Una serie donde cada término es el anterior multiplicado por una razón fija r:

$$a + ar + ar^2 + ar^3 + \cdots = \frac{a}{1-r} \quad \text{si } |r|<1 \quad \text{(converge)}.$$

Si $|r| \geq 1$, diverge. Aquí a es el primer término.

Ejemplo 6.163. Calcular $1+\frac{1}{2}+\frac{1}{4}+\frac{1}{8}+\cdots$.

Paso 1. Identificamos: el primer término es $a=1$, y cada término es la mitad del anterior, así que $r=\frac{1}{2}$.

Paso 2. Como $|r| = \frac{1}{2} < 1$, converge, y usamos la fórmula:

$$\frac{a}{1-r} = \frac{1}{1-\frac{1}{2}} = \frac{1}{\frac{1}{2}} = 2.$$

La suma infinita vale exactamente 2. (Tiene sentido: vas sumando mitades y te acercas a 2 sin pasarte.)

6.32.2. Series-p: la regla del exponente

Regla: series-p

La serie $\displaystyle\sum_{n=1}^{\infty}\frac{1}{n^p} = \frac{1}{1^p}+\frac{1}{2^p}+\frac{1}{3^p}+\cdots$ converge si $p>1$ y diverge si $p\leq 1$.

Ejemplo 6.164. La *serie armónica* $1+\frac{1}{2}+\frac{1}{3}+\frac{1}{4}+\cdots$ tiene $p=1$, así que **diverge**: se hace infinitamente grande, aunque sus términos se achiquen. En cambio $1+\frac{1}{4}+\frac{1}{9}+\frac{1}{16}+\cdots=\sum\frac{1}{n^2}$ tiene $p=2>1$, así que **converge** (de hecho a $\frac{\pi^2}{6}$, un resultado célebre). La frontera está justo en $p=1$.

6.32.3. El criterio del cociente (el más útil)

Regla: criterio del cociente

Para $\sum a_n$, calcula $L=\lim_{n\to\infty}\left|\frac{a_{n+1}}{a_n}\right|$ (cuánto crece cada término respecto del anterior). Entonces: si $L<1$ *converge*; si $L>1$ *diverge*; si $L=1$ el criterio no decide.

Ejemplo 6.165. Decidir si $\displaystyle\sum_{n=1}^{\infty}\frac{n}{2^n}$ converge.

Paso 1. El término general es $a_n=\dfrac{n}{2^n}$, y el siguiente $a_{n+1}=\dfrac{n+1}{2^{n+1}}$.

Paso 2. Formamos el cociente y simplificamos:

$$\frac{a_{n+1}}{a_n}=\frac{n+1}{2^{n+1}}\cdot\frac{2^n}{n}=\frac{n+1}{n}\cdot\frac{2^n}{2^{n+1}}=\frac{n+1}{n}\cdot\frac{1}{2}.$$

Paso 3. Tomamos el límite: $\dfrac{n+1}{n}\to 1$, así que $L=1\cdot\frac{1}{2}=\frac{1}{2}$.

Paso 4. Como $L=\frac{1}{2}<1$, la serie **converge**. El factor $\frac{1}{2}$ (la exponencial 2^n) gana a la n de arriba, y por eso la suma es finita.

6.33. Ampliación del temario: L'Hôpital, técnicas de integración y curvas paramétricas/polares

Esta sección desarrolla, con el detalle de un curso completo, tres bloques que redondean el cálculo de una variable: las formas indeterminadas y la regla de L'Hôpital, el catálogo explícito de técnicas de integración, y el cálculo sobre curvas paramétricas y en coordenadas polares.

6.33.1. Regla de L'Hôpital y las formas indeterminadas

Idea clave

Intuición primero. ¿Por qué funciona derivar arriba y abajo? Cerca del punto conflictivo, f y g valen casi cero, así que cada una se parece a su recta tangente: $f(x)\approx f'(a)(x-a)$ y $g(x)\approx g'(a)(x-a)$. Al dividir, el factor común $(x-a)$ se cancela y queda $\frac{f'(a)}{g'(a)}$. En otras palabras: cuando $\frac{0}{0}$ te bloquea, reemplaza cada función por su pendiente y la indeterminación se deshace. Esa es toda la magia.

Al calcular un límite de un cociente, a veces la sustitución directa produce una expresión sin sentido definido, como $\frac{0}{0}$ o $\frac{\infty}{\infty}$: una *forma indeterminada*. Hay siete:

$$\frac{0}{0}, \quad \frac{\infty}{\infty}, \quad 0\cdot\infty, \quad \infty-\infty, \quad 1^{\infty}, \quad 0^{0}, \quad \infty^{0}.$$

Regla: L'Hôpital

Si $\lim \dfrac{f(x)}{g(x)}$ es de la forma $\frac{0}{0}$ o $\frac{\infty}{\infty}$, entonces

$$\lim \frac{f(x)}{g(x)} = \lim \frac{f'(x)}{g'(x)},$$

siempre que el segundo límite exista. Se deriva *numerador y denominador por separado* (no es la regla del cociente).

Ejemplo 6.166 (Forma $\frac{0}{0}$). $\lim\limits_{x\to 0} \dfrac{e^x-1-x}{x^2}$ da $\frac{0}{0}$. Derivamos: $\dfrac{e^x-1}{2x}$, aún $\frac{0}{0}$; otra vez: $\dfrac{e^x}{2} \to \dfrac{1}{2}$.

Ejemplo 6.167 (Las demás formas se reducen a $\frac{0}{0}$ o $\frac{\infty}{\infty}$). *($0\cdot\infty$)* $\lim\limits_{x\to 0^+} x\ln x$: reescribimos como $\dfrac{\ln x}{1/x}$ (forma $\frac{-\infty}{\infty}$), L'Hôpital da $\dfrac{1/x}{-1/x^2} = -x \to 0$.

(1^{∞}) $\lim\limits_{x\to 0}(1+x)^{1/x}$: tomamos logaritmo, $\dfrac{\ln(1+x)}{x} \to 1$ (L'Hôpital), así que el límite original es $e^1 = e$. Este es, de hecho, el número e.

Cuidado

L'Hôpital *solo* se aplica a $\frac{0}{0}$ y $\frac{\infty}{\infty}$. Si sustituir da un número normal o algo como $\frac{5}{0}$, aplicarla da una respuesta equivocada. Comprueba siempre la forma primero.

6.33.2. Técnicas de integración, en detalle

Integración por partes.

Idea clave

Intuición primero. Es la regla del producto leída al revés. Como $(uv)' = u'v + uv'$, al integrar los dos lados queda $uv = \int u'v + \int uv'$, y despejando obtenemos la fórmula. Lo que estás haciendo es *mover la derivada de un factor al otro*: cambias una integral difícil por otra, con la esperanza de que la nueva sea más fácil. La regla LIATE solo te dice qué factor conviene derivar (el que se simplifica) y cuál integrar.

De la regla del producto al revés: $\displaystyle\int u\,dv = uv - \int v\,du$. Se elige u por la regla **LIATE** (Logaritmos, Inversas, Algebraicas, Trigonométricas, Exponenciales: lo que aparezca antes es u).

Ejemplo 6.168. $\int x^2 e^x \, dx$. Con $u = x^2$, $dv = e^x dx$: $= x^2 e^x - \int 2x e^x dx$. La integral que queda se hace de nuevo por partes ($\int x e^x = e^x(x-1)$), y al final

$$\int x^2 e^x \, dx = e^x(x^2 - 2x + 2) + C.$$

Fracciones parciales. Para integrar un cociente de polinomios, se descompone en fracciones simples. Los casos: factores lineales distintos ($\frac{A}{x-a}$), factores lineales repetidos ($\frac{A}{x-a} + \frac{B}{(x-a)^2}$), y factores cuadráticos irreducibles ($\frac{Ax+B}{x^2+bx+c}$).

Ejemplo 6.169 (Factores lineales distintos)**.** $\int \dfrac{5x-4}{x^2-x-2} \, dx$. Factorizamos $x^2-x-2 = (x-2)(x+1)$ y ponemos $\dfrac{5x-4}{(x-2)(x+1)} = \dfrac{A}{x-2} + \dfrac{B}{x+1}$. Multiplicando: $5x-4 = A(x+1)+B(x-2)$. Con $x=2$: $6 = 3A \Rightarrow A = 2$. Con $x=-1$: $-9 = -3B \Rightarrow B = 3$. Entonces

$$\int \frac{5x-4}{x^2-x-2} \, dx = 2\ln|x-2| + 3\ln|x+1| + C.$$

Sustitución trigonométrica.

Idea clave

Intuición primero. La idea es un triángulo rectángulo escondido. Si escribes $x = a\sin\theta$, estás diciendo "x es el cateto opuesto de un triángulo de hipotenusa a"; entonces $\sqrt{a^2-x^2}$ es simplemente el cateto adyacente $a\cos\theta$, y la raíz desaparece. Cada una de las tres sustituciones elige el triángulo que convierte la raíz molesta en un lado limpio, gracias a la identidad pitagórica.

Cuando aparece $\sqrt{a^2-x^2}$, $\sqrt{a^2+x^2}$ o $\sqrt{x^2-a^2}$, un cambio con seno, tangente o secante elimina la raíz usando la identidad pitagórica.

Regla: las tres sustituciones

$$\begin{aligned}
\sqrt{a^2-x^2}: &\quad x = a\sin\theta \quad (\text{usa } 1-\sin^2 = \cos^2)\\
\sqrt{a^2+x^2}: &\quad x = a\tan\theta \quad (\text{usa } 1+\tan^2 = \sec^2)\\
\sqrt{x^2-a^2}: &\quad x = a\sec\theta \quad (\text{usa } \sec^2 - 1 = \tan^2)
\end{aligned}$$

Ejemplo 6.170. $\int \dfrac{dx}{\sqrt{9-x^2}}$. Con $x = 3\sin\theta$, $dx = 3\cos\theta\, d\theta$ y $\sqrt{9-x^2} = 3\cos\theta$:

$$\int \frac{3\cos\theta\, d\theta}{3\cos\theta} = \int d\theta = \theta + C = \arcsin\left(\frac{x}{3}\right) + C.$$

6.33.3. Curvas paramétricas

Idea clave

Intuición primero. Piensa en una película: t es el tiempo y $(x(t), y(t))$ es dónde está la punta del lápiz en cada instante. La pendiente $\frac{\mathrm{d}y}{\mathrm{d}x}$ es (velocidad vertical)/(velocidad horizontal). Y la longitud se arma sumando pequeñas hipotenusas $\sqrt{\mathrm{d}x^2+\mathrm{d}y^2}$, que no es otra cosa que la rapidez por $\mathrm{d}t$: recorres la curva y vas midiendo cuánto avanzas en cada tic del reloj.

Una curva paramétrica se describe con $x = x(t)$, $y = y(t)$: un punto que se mueve con el tiempo t. La pendiente y la longitud se calculan con derivadas respecto de t.

Regla: tangente y longitud de arco (paramétrica)

$$\frac{\mathrm{d}y}{\mathrm{d}x} = \frac{\mathrm{d}y/\mathrm{d}t}{\mathrm{d}x/\mathrm{d}t}, \qquad L = \int_a^b \sqrt{\Big(\frac{\mathrm{d}x}{\mathrm{d}t}\Big)^2 + \Big(\frac{\mathrm{d}y}{\mathrm{d}t}\Big)^2}\,\mathrm{d}t.$$

Ejemplo 6.171 (La cicloide)**.** La curva que traza un punto de una rueda que gira: $x = t - \sin t$, $y = 1 - \cos t$. La longitud de un arco completo ($0 \le t \le 2\pi$):

$$L = \int_0^{2\pi} \sqrt{(1-\cos t)^2 + \sin^2 t}\,\mathrm{d}t = \int_0^{2\pi} \sqrt{2-2\cos t}\,\mathrm{d}t = \int_0^{2\pi} 2\Big|\sin\tfrac{t}{2}\Big|\,\mathrm{d}t = 8.$$

(Se usó $2 - 2\cos t = 4\sin^2(t/2)$.) La longitud del arco es exactamente 8 (cuatro diámetros).

6.33.4. Coordenadas polares

Idea clave

Intuición primero. En vez de decir "cuánto a la derecha y cuánto arriba", dices "cuán lejos del centro y en qué ángulo". El elemento de área ya no es un rectangulito: es una *rebanada de pastel*, un sector de radio r y ángulo $\mathrm{d}\theta$, cuya área es $\frac{1}{2}r^2\mathrm{d}\theta$. Por eso la fórmula del área lleva ese $\frac{1}{2}r^2$: estás sumando rebanadas, no cuadraditos.

En polares, un punto se ubica por su distancia al origen r y su ángulo θ: $x = r\cos\theta$, $y = r\sin\theta$. Muchas curvas (círculos, espirales, pétalos) son mucho más simples así.

Regla: área y longitud en polares

El área barrida por $r = r(\theta)$ entre α y β, y la longitud de arco, son

$$A = \frac{1}{2}\int_\alpha^\beta r^2\,\mathrm{d}\theta, \qquad L = \int_\alpha^\beta \sqrt{r^2 + \Big(\frac{\mathrm{d}r}{\mathrm{d}\theta}\Big)^2}\,\mathrm{d}\theta.$$

Ejemplo 6.172 (Área de una cardioide)**.** La cardioide $r = 1 + \cos\theta$ (forma de corazón) encierra un área

$$A = \frac{1}{2}\int_0^{2\pi} (1+\cos\theta)^2\,\mathrm{d}\theta = \frac{1}{2}\int_0^{2\pi} \left(1 + 2\cos\theta + \cos^2\theta\right)\mathrm{d}\theta = \frac{1}{2}\Big(2\pi + 0 + \pi\Big) = \frac{3\pi}{2}.$$

(Usamos $\int_0^{2\pi}\cos^2\theta\,\mathrm{d}\theta = \pi$.) El área es $\frac{3\pi}{2}$.

6.34. Taller de integrales impropias

Idea clave

Intuición primero. Una integral impropia es una integral a la que le falta algo: o el intervalo no termina, o la función se dispara en algún punto. En los dos casos el área que se quiere medir se escapa hacia el infinito en alguna dirección, y la única pregunta sensata es si esa fuga es lo bastante rápida como para que el área total siga siendo finita. Piensa en una cola que se estrecha: si se estrecha *deprisa*, el área bajo la cola es finita aunque la cola sea infinitamente larga; si se estrecha despacio, no. Toda la teoría se reduce a tener un patrón de medida —las potencias $1/x^p$— y saber comparar con él. Y hay una simetría preciosa que conviene fijar desde el principio: en el infinito convergen las potencias *grandes* ($p > 1$), y en una singularidad finita convergen las *pequeñas* ($p < 1$). Es el mismo criterio visto por los dos extremos del catalejo, y equivocarse de lado es el error que más puntos cuesta.

Regla: decidir si una integral impropia converge

(I) **Localiza *todos* los puntos conflictivos** antes de escribir nada: extremos infinitos, extremos donde el integrando se dispara y puntos interiores donde deja de estar acotado. Luego parte la integral de modo que cada trozo tenga exactamente *un* problema y en un solo extremo.

(II) **Exige convergencia de cada trozo por separado.** La integral completa converge si y solo si convergen todos los pedazos. En particular $\int_{-\infty}^{\infty} f$ es $\int_{-\infty}^{0} f + \int_0^{\infty} f$, no un único límite simétrico.

(III) **Compara con las dos escalas patrón,** que tiran en sentidos opuestos:

$$\int_1^{\infty} \frac{\mathrm{d}x}{x^p} \text{ converge } \iff p > 1, \qquad \int_0^{1} \frac{\mathrm{d}x}{x^p} \text{ converge } \iff p < 1.$$

(IV) **Usa la comparación por paso al límite.** Si $f, g > 0$ cerca del punto conflictivo y $f/g \to L$ con $0 < L < \infty$, entonces las dos integrales hacen lo mismo. En la práctica: busca el término dominante y quédate con él.

(V) **Recuerda la jerarquía.** En el infinito, la exponencial aplasta a cualquier potencia ($x^k e^{-x} \to 0$) y el logaritmo crece menos que cualquier x^ε. Los logaritmos solo deciden los casos de frontera, y allí el criterio fino es $\int_2^{\infty} \mathrm{d}x/(x(\ln x)^q)$, que converge si y solo si $q > 1$.

(vi) **Si f cambia de signo, estudia primero $\int |f|$.** Si $\int |f|$ converge, $\int f$ también (convergencia absoluta). Si $\int |f|$ diverge, todavía puede haber convergencia *condicional* por cancelación, y ahí hacen falta argumentos de tipo Dirichlet o integración por partes.

(vii) **No confundas valor principal con convergencia.** El valor principal de Cauchy, $\lim_{\varepsilon\to 0^+}\big(\int_a^{c-\varepsilon} + \int_{c+\varepsilon}^{b}\big)$, puede existir para integrales que sencillamente divergen. Si lo usas, dilo expresamente.

Ejemplo 6.173 (Las dos escalas y la trampa de juntarlas)**.** Directamente,

$$\int_1^\infty \frac{\mathrm{d}x}{x^{3/2}} = \Big[\frac{-2}{\sqrt{x}}\Big]_1^\infty = 2, \qquad \int_0^1 \frac{\mathrm{d}x}{\sqrt{x}} = \big[2\sqrt{x}\big]_0^1 = 2.$$

Las dos convergen, y las dos valen 2, pero por razones opuestas: la primera porque $3/2 > 1$, la segunda porque $1/2 < 1$. Ahora júntense: $\int_0^\infty \mathrm{d}x/x^p$ *diverge para todo* p, porque ningún exponente puede ser a la vez mayor y menor que 1. Es el ejemplo que hay que tener en la cabeza para no olvidar nunca el paso (i) de la regla.

Ejemplo 6.174 (Cuatro comparaciones por paso al límite)**.** (i) $\displaystyle\int_1^\infty \frac{\mathrm{d}x}{\sqrt{x^3+1}}$: para x grande el integrando se parece a $x^{-3/2}$, y $3/2 > 1$. Converge; numéricamente vale 1,894760.

(ii) $\displaystyle\int_1^\infty \frac{\mathrm{d}x}{\sqrt{x^2+1}}$: ahora el integrando se parece a $1/x$, con $p = 1$. Diverge, y de hecho la primitiva es $\operatorname{arcsinh} x$, que crece como $\ln x$.

(iii) $\displaystyle\int_2^\infty \frac{\mathrm{d}x}{x\ln x}$: el caso frontera. Con $u = \ln x$ se convierte en $\int_{\ln 2}^\infty \mathrm{d}u/u$, que diverge. El logaritmo no basta para salvar la situación.

(iv) $\displaystyle\int_2^\infty \frac{\mathrm{d}x}{x(\ln x)^2}$: el mismo cambio la convierte en $\int_{\ln 2}^\infty \mathrm{d}u/u^2$, que converge y vale $1/\ln 2 \approx 1{,}442695$. Un solo exponente de diferencia en el logaritmo decide el asunto.

Ejemplo 6.175 (Singularidades finitas que no lo parecen)**.** $\displaystyle\int_0^1 \frac{\mathrm{d}x}{\sqrt{1-x^2}}$ tiene el integrando disparado en $x = 1$. Cerca de ese punto, $1 - x^2 = (1-x)(1+x) \approx 2(1-x)$, de modo que el integrando se comporta como $(2(1-x))^{-1/2}$: una potencia con $p = 1/2 < 1$. Converge, y en efecto

$$\int_0^1 \frac{\mathrm{d}x}{\sqrt{1-x^2}} = \big[\arcsin x\big]_0^1 = \frac{\pi}{2}.$$

Con logaritmos ocurre lo mismo: $\int_0^1 \ln x\,\mathrm{d}x = [x\ln x - x]_0^1 = -1$ converge, pese a que $\ln x \to -\infty$, porque el logaritmo diverge muchísimo más despacio que cualquier potencia. Y $\int_0^1 x^{-1/2}\ln x\,\mathrm{d}x = -4$ también converge: el producto de dos singularidades suaves sigue siendo integrable.

Ejemplo 6.176 (Convergencia condicional: $\int_0^\infty \frac{\sin x}{x}\,\mathrm{d}x$)**.** En $x=0$ no hay problema, porque $\sin x/x \to 1$. En el infinito, el integrando no tiende a cero *en valor absoluto* lo bastante rápido: sobre el intervalo $[k\pi,(k+1)\pi]$ el área de $|\sin x|/x$ vale al menos $\frac{2}{(k+1)\pi}$, y la serie de esos números diverge como la armónica. Por tanto $\int_0^\infty |\sin x|/x\,\mathrm{d}x = \infty$: no hay convergencia absoluta. De hecho $\int_1^R |\sin x|/x\,\mathrm{d}x$ crece como $\frac{2}{\pi}\ln R$ —el valor medio de $|\sin|$ sobre un periodo es exactamente $2/\pi$—, y para $R=1000$ vale 4,5638 frente a la estimación $\frac{2}{\pi}\ln 1000 = 4{,}3976$. Y sin embargo la integral *con signo* sí converge, porque los lóbulos positivos y negativos se cancelan casi enteros: integrando por partes,

$$\int_1^R \frac{\sin x}{x}\,\mathrm{d}x = \Big[\frac{-\cos x}{x}\Big]_1^R - \int_1^R \frac{\cos x}{x^2}\,\mathrm{d}x,$$

y ahora el término de contorno tiende a $\cos 1$ y la integral que queda es absolutamente convergente, porque $|\cos x|/x^2 \le 1/x^2$. El valor total es $\pi/2$.

Cuidado

[El valor principal no salva nada] $\int_{-1}^1 \mathrm{d}x/x$ *diverge*: los dos trozos valen $-\infty$ y $+\infty$, y $-\infty+\infty$ no es un número. Alguien podrá objetar que "por simetría" debería dar 0, y en efecto $\lim_{\varepsilon\to 0^+}(\int_{-1}^{-\varepsilon}+\int_{\varepsilon}^{1})\mathrm{d}x/x = 0$; pero ese es el valor principal, que se obtiene recortando los dos lados *al mismo ritmo*. Si se recorta a ritmos distintos —por ejemplo $[-\varepsilon,\,2\varepsilon]$— el límite es $\ln 2$, o cualquier otra cosa según el ritmo elegido. Una integral que da respuestas distintas según cómo te acerques no converge; el valor principal es una convención útil (en análisis complejo, en transformadas), no una convergencia.

Pausa para pensar

La función Gamma $\Gamma(s)=\int_0^\infty x^{s-1}e^{-x}\,\mathrm{d}x$ es un buen examen de todo lo anterior: tiene dos puntos conflictivos, no uno. En el infinito no hay problema para ningún s, porque e^{-x} aplasta cualquier potencia (paso (v)); en el origen el integrando se comporta como x^{s-1}, y $\int_0^1 x^{s-1}\mathrm{d}x$ converge si y solo si $1-s<1$, es decir, $s>0$. De ahí sale la condición $s>0$ del enunciado clásico, que no es un capricho sino exactamente el paso (iii) aplicado en el extremo izquierdo. Compruébalo tú mismo con $s=1/2$: la integral vale $\Gamma(1/2)=\sqrt{\pi}$, y con el cambio $x=u^2$ se convierte en $2\int_0^\infty e^{-u^2}\mathrm{d}u = 2\cdot\frac{\sqrt{\pi}}{2}$, la integral de Gauss del capítulo 9.

6.35. Taller de indeterminaciones: las siete formas y cómo romperlas

Idea clave

Intuición primero. Una indeterminación no es un límite difícil: es un límite del que la aritmética de límites no dice *nada*. Cuando escribes $0/0$ no estás escribiendo un número, estás escribiendo una carrera entre dos cantidades que van a cero, y quién gana depende de a qué velocidad va cada una. $x^2/x\to 0$, $x/x\to 1$ y $x/x^2\to\infty$ son las tres $0/0$, y

dan tres respuestas distintas. Por eso el trabajo consiste siempre en lo mismo: medir la velocidad de cada participante.
Y hay un instrumento que mide velocidades mejor que ningún otro, el desarrollo de Taylor. L'Hôpital es cómodo y muchas veces suficiente, pero es un martillo; Taylor es un microscopio que te dice, término a término, cuánto vale cada uno de los competidores. Cuando dudes, desarrolla.

Regla: romper una indeterminación

(i) **Identifica la forma.** Solo hay siete indeterminaciones:

$$\frac{0}{0},\quad \frac{\infty}{\infty},\quad 0\cdot\infty,\quad \infty-\infty,\quad 1^{\infty},\quad 0^{0},\quad \infty^{0}.$$

Todo lo demás *no* es indeterminado y se evalúa sin más: $0^{\infty}=0$, $\infty^{\infty}=\infty$, $1/0^{+}=+\infty$. Antes de trabajar, comprueba que realmente tienes una de las siete.

(ii) **Prueba primero el álgebra.** Factoriza y simplifica, racionaliza multiplicando por el conjugado, saca denominador común, o divide numerador y denominador por la potencia dominante. Muchísimos límites de examen caen aquí, y es siempre más rápido y más seguro que derivar.

(iii) **Reduce a un cociente.** La forma $0\cdot\infty$ se escribe $f\cdot g=f/(1/g)$; la forma $\infty-\infty$ se ataca con denominador común, con el conjugado si hay raíces, o sacando factor común el término dominante.

(iv) **Para las tres formas exponenciales, toma logaritmos.** Si $L=\lim f^{g}$, entonces $\ln L=\lim g\ln f$, que es del tipo $0\cdot\infty$, y al final $L=e^{\lim g\ln f}$. No olvides deshacer la exponencial: es el despiste más frecuente.

(v) **Usa infinitésimos equivalentes, pero solo como *factores*.** Cuando $u\to 0$:

$$\sin u\sim u,\quad \tan u\sim u,\quad 1-\cos u\sim \tfrac{u^2}{2},\quad \ln(1+u)\sim u,\quad e^{u}-1\sim u,\quad (1+u)^{a}-1\sim au.$$

Sustituir dentro de una *suma* donde hay cancelación es un error grave: $\sin x-x$ no es 0, es $-x^3/6$.

(vi) **L'Hôpital, con sus condiciones.** Solo vale para $0/0$ e ∞/∞; hay que comprobar que el límite de f'/g' existe, y hay que saber parar. En $x/\sqrt{x^2+1}$ el método se muerde la cola: derivando se vuelve a lo mismo, indefinidamente. Dividir por x lo resuelve de un golpe.

(vii) **Taylor como recurso definitivo.** Desarrolla numerador y denominador hasta el primer orden en que dejen de cancelarse, y lee el cociente de los coeficientes. Es el único método que nunca falla y el único que te dice, además del límite, *a qué velocidad* se alcanza.

Ejemplo 6.177 (Por qué los equivalentes no valen dentro de una resta). $\lim_{x\to 0} \dfrac{\sin x - x}{x^3}$. Quien sustituya $\sin x$ por x obtiene $0/x^3 = 0$, y se equivoca. La razón es que la sustitución tira precisamente el término que decide el resultado. Con Taylor, $\sin x = x - \frac{x^3}{6} + \frac{x^5}{120} - \cdots$, luego

$$\frac{\sin x - x}{x^3} = \frac{-\frac{x^3}{6} + \frac{x^5}{120} - \cdots}{x^3} \longrightarrow -\frac{1}{6}.$$

Compárese con el trabajo que cuesta por L'Hôpital: tres derivaciones sucesivas ($\frac{\cos x - 1}{3x^2}$, $\frac{-\sin x}{6x}$, $\frac{-\cos x}{6}$) para llegar al mismo $-1/6$. El mismo fenómeno, con el mismo desenlace, en

$$\lim_{x\to 0} \frac{\tan x - \sin x}{x^3} = \frac{1}{2}, \qquad \lim_{x\to 0} \frac{e^x - 1 - x}{x^2} = \frac{1}{2}, \qquad \lim_{x\to 0} \frac{\ln(1+x) - x}{x^2} = -\frac{1}{2},$$

que salen de $\tan x = x + \frac{x^3}{3} + \cdots$, $e^x = 1 + x + \frac{x^2}{2} + \cdots$ y $\ln(1+x) = x - \frac{x^2}{2} + \cdots$ sin derivar ni una sola vez.

Ejemplo 6.178 (Cuando el equivalente sí vale). $\lim_{x\to 0} \dfrac{1 - \cos x}{x^2}$. Aquí $1 - \cos x$ aparece como factor entero, no dentro de una resta que se cancela, así que el equivalente $1 - \cos x \sim x^2/2$ se aplica sin miedo y el límite es $\frac{1}{2}$. Igual de inmediato:

$$\lim_{x\to 0} \frac{\sin(3x)\ln(1+2x)}{x\,(e^{5x} - 1)} = \lim_{x\to 0} \frac{3x \cdot 2x}{x \cdot 5x} = \frac{6}{5},$$

donde se han sustituido cuatro equivalentes de golpe porque los cuatro son factores. La regla práctica es sencilla: *en un producto o un cociente, sí; dentro de una suma que se cancela, no.*

Ejemplo 6.179 (Las tres formas exponenciales).

(i) $\lim_{x\to\infty} \left(1 + \dfrac{3}{x}\right)^{2x}$ es del tipo 1^∞. Tomando logaritmos, $2x\ln(1 + 3/x) \sim 2x \cdot \frac{3}{x} = 6$, luego el límite es e^6.

(ii) $\lim_{x\to 0^+} x^x$ es del tipo 0^0. El logaritmo es $x \ln x$, del tipo $0 \cdot \infty$; escribiéndolo como $\ln x/(1/x)$ y aplicando L'Hôpital sale $\lim(-x) = 0$, de modo que $x^x \to e^0 = 1$.

(iii) $\lim_{x\to 0} (\cos x)^{1/x^2}$ es otro 1^∞. El logaritmo es $\frac{\ln(\cos x)}{x^2}$; usando $\cos x = 1 - \frac{x^2}{2} + \cdots$ y $\ln(1+u) \sim u$ con $u = -\frac{x^2}{2}$, queda $-\frac{1}{2}$. El límite es $e^{-1/2}$, que no es ni 0 ni 1: buena prueba de que 1^∞ es de verdad indeterminado.

Ejemplo 6.180 (Restas de infinitos). $\lim_{x\to\infty} \left(\sqrt{x^2 + x} - x\right)$ es del tipo $\infty - \infty$. Multiplicando y dividiendo por el conjugado,

$$\sqrt{x^2 + x} - x = \frac{x}{\sqrt{x^2 + x} + x} = \frac{1}{\sqrt{1 + 1/x} + 1} \longrightarrow \frac{1}{2}.$$

Sin el conjugado no hay nada que hacer, y L'Hôpital no se puede ni empezar a aplicar porque no hay cociente. Otro caso de la misma familia: $\lim_{x\to 0} \left(\dfrac{1}{x} - \dfrac{1}{\sin x}\right)$, que con denominador común se convierte en

$$\frac{\sin x - x}{x \sin x} = \frac{-\frac{x^3}{6} + \cdots}{x^2 + \cdots} \longrightarrow 0.$$

Obsérvese que el numerador es exactamente el del primer ejemplo: el mismo cálculo de Taylor sirve dos veces.

Cuidado

[L'Hôpital que se muerde la cola] $\lim_{x\to\infty} \dfrac{x}{\sqrt{x^2+1}}$ es del tipo ∞/∞, así que L'Hôpital es aplicable. Derivando:

$$\frac{1}{x/\sqrt{x^2+1}} = \frac{\sqrt{x^2+1}}{x},$$

que es el inverso del original y vuelve a ser ∞/∞. Derivando otra vez se regresa al punto de partida, y así indefinidamente. La regla es aplicable pero inútil, y quien no se dé cuenta puede pasarse la tarde girando. Sacando factor común, en cambio, el límite es inmediato: $x/\sqrt{x^2+1} = 1/\sqrt{1+1/x^2} \to 1$. Moraleja: L'Hôpital es el paso (vi), no el paso (ii).

Pausa para pensar

Vuelve a mirar los tres resultados $-\frac{1}{6}$, $\frac{1}{2}$ y $-\frac{1}{2}$ de los ejemplos. No son números caprichosos: son exactamente los coeficientes del primer término no nulo del desarrollo de Taylor. Es decir, calcular un límite indeterminado y calcular un coeficiente de Taylor son la misma operación. Esa es la razón profunda de por qué el paso (vii) nunca falla, y también de por qué L'Hôpital funciona: derivar k veces y evaluar en el punto es, literalmente, leer el k-ésimo coeficiente. Los dos métodos hacen lo mismo; Taylor lo hace todo de una vez y L'Hôpital de uno en uno.

6.36. Centroides y los teoremas de Pappus

Idea clave

Intuición primero. El centroide de una región plana es el punto donde la región se equilibraría si estuviera recortada en cartón homogéneo. Los teoremas de Pappus dicen algo que parece demasiado bueno para ser cierto: si uno conoce ese punto de equilibrio, obtiene gratis el volumen y la superficie del sólido que se genera al girar la región alrededor de un eje. La razón es sencilla: al girar, cada trocito de la región recorre una circunferencia cuyo radio es su distancia al eje; sumar todos esos recorridos es exactamente promediar la distancia, y ese promedio es el centroide. El sólido pesa, literalmente, lo que pesa la región multiplicado por lo que viaja su centro.

Regla: centroides y giros por el método de Pappus

(i) **Dibujar y fijar la región.** Escribir la región como $\{(x,y) : a \le x \le b,\ g(x) \le y \le f(x)\}$ con $f \ge g$ en $[a,b]$. Todo lo que sigue supone que la región es la de *arriba menos abajo*.

(ii) **Calcular primero el área.** $A = \displaystyle\int_a^b (f(x) - g(x))\mathrm{d}x$. Sin A no hay centroide, porque el centroide es un promedio y A es el denominador.

(iii) **Los dos momentos.**

$$\bar{x} = \frac{1}{A}\int_a^b x(f(x)-g(x))\mathrm{d}x, \qquad \bar{y} = \frac{1}{A}\int_a^b \tfrac{1}{2}(f(x)^2-g(x)^2)\mathrm{d}x.$$

El $\frac{1}{2}(f^2-g^2)$ no es un capricho: es la altura media $\frac{1}{2}(f+g)$ del trozo vertical multiplicada por su longitud $f-g$.

(iv) **Usar la simetría antes de integrar.** Si la región es simétrica respecto de una recta, el centroide está sobre esa recta. Eso puede ahorrar una de las dos integrales por completo.

(v) **Pappus I (volumen).** Al girar la región alrededor de un eje del plano *que no la corte*, el sólido tiene volumen

$$V = 2\pi d\, A,$$

donde d es la distancia del centroide al eje. Si el eje es OY, entonces $d=\bar{x}$; si es OX, entonces $d=\bar{y}$.

(vi) **Pappus II (superficie).** Al girar una *curva* de longitud L alrededor de un eje que no la corte, la superficie generada es

$$S = 2\pi d_{\text{curva}}\, L,$$

con d_{curva} la distancia al eje del centroide *de la curva*, que en general **no** coincide con el de la región que encierra.

(vii) **Calibrar con un caso conocido.** Antes de fiarse del resultado, rehacerlo por discos o capas en un caso que uno ya sepa. Si Pappus y el método directo no coinciden, el error casi siempre está en d o en haber mezclado el centroide de la región con el de su borde.

Ejemplo 6.181 (Dos centroides calculados a mano)**. (a) El triángulo de vértices** $(0,0)$**,** $(1,0)$**,** $(0,1)$**.** Aquí $f(x)=1-x$ y $g(x)=0$ en $[0,1]$, de modo que

$$A = \int_0^1 (1-x)\mathrm{d}x = \tfrac{1}{2}, \qquad \bar{x} = \frac{1}{1/2}\int_0^1 x(1-x)\mathrm{d}x = 2\cdot\tfrac{1}{6} = \tfrac{1}{3},$$

$$\bar{y} = \frac{1}{1/2}\int_0^1 \tfrac{1}{2}(1-x)^2\mathrm{d}x = 2\cdot\tfrac{1}{6} = \tfrac{1}{3}.$$

Sale $(\bar{x},\bar{y}) = (1/3,1/3)$, que es el promedio de los tres vértices: para un triángulo el centroide es siempre el baricentro, y no hace falta integrar.

(b) La región bajo $y=x^2$ **sobre** $[0,1]$**.** Ahora $f(x)=x^2$, $g=0$:

$$A = \int_0^1 x^2\mathrm{d}x = \tfrac{1}{3}, \qquad \bar{x} = 3\int_0^1 x^3\mathrm{d}x = 3\cdot\tfrac{1}{4} = \tfrac{3}{4},$$

$$\bar{y} = 3\int_0^1 \tfrac{1}{2}x^4\mathrm{d}x = \tfrac{3}{2}\cdot\tfrac{1}{5} = \tfrac{3}{10}.$$

El centroide (3/4, 3/10) está desplazado hacia la derecha y hacia abajo, como debe ser: casi toda el área de la región se acumula cerca de $x = 1$ y pegada al eje OX.

Ejemplo 6.182 (Pappus contrastado con discos y capas)**.** Seguimos con la región bajo $y = x^2$ sobre $[0, 1]$, cuyo centroide acabamos de calcular, y la hacemos girar de las dos maneras posibles.

(a) Giro alrededor de OY. La distancia del centroide al eje es $d = \bar{x} = 3/4$, así que Pappus da

$$V = 2\pi \cdot \tfrac{3}{4} \cdot \tfrac{1}{3} = \frac{\pi}{2}.$$

Por capas cilíndricas, $V = \int_0^1 2\pi x \cdot x^2 \mathrm{d}x = 2\pi \cdot \frac{1}{4} = \pi/2$. Coinciden.

(b) Giro alrededor de OX. Ahora $d = \bar{y} = 3/10$ y

$$V = 2\pi \cdot \tfrac{3}{10} \cdot \tfrac{1}{3} = \frac{\pi}{5}.$$

Por discos, $V = \int_0^1 \pi (x^2)^2 \mathrm{d}x = \pi/5$. Vuelven a coincidir.

Obsérvese lo que ha ocurrido: con *un solo* par de integrales (las del centroide) hemos obtenido *dos* volúmenes de revolución. Ésa es toda la economía del método.

Ejemplo 6.183 (La esfera, dos veces, y el toro)**. (a) El semidisco.** Sea la región $x^2+y^2 \le R^2$ con $y \ge 0$. Por simetría $\bar{x} = 0$. Para $\bar{y}$ tomamos $f(x) = \sqrt{R^2 - x^2}$, $g = 0$ y $A = \pi R^2/2$:

$$\bar{y} = \frac{2}{\pi R^2} \int_{-R}^{R} \tfrac{1}{2}(R^2 - x^2)\mathrm{d}x = \frac{1}{\pi R^2} \cdot \frac{4R^3}{3} = \frac{4R}{3\pi}.$$

Como $4/(3\pi) \approx 0{,}4244$, el centroide queda algo por debajo de la mitad del radio, lo cual es razonable: hay más área cerca del diámetro.

(b) La esfera por Pappus I. Girando ese semidisco alrededor de su diámetro (el eje OX, que sólo toca el borde de la región, no su interior):

$$V = 2\pi \cdot \frac{4R}{3\pi} \cdot \frac{\pi R^2}{2} = \frac{4\pi R^3}{3}.$$

(c) El arco semicircular, que es otro centroide. La *curva* $x^2 + y^2 = R^2$ con $y \ge 0$ tiene longitud $L = \pi R$ y, parametrizada por $(R\cos t, R\sin t)$ con $\mathrm{d}s = R\mathrm{d}t$,

$$\bar{y}_{\text{arco}} = \frac{1}{\pi R} \int_0^{\pi} R\sin t \cdot R\mathrm{d}t = \frac{R^2 \cdot 2}{\pi R} = \frac{2R}{\pi}.$$

Nótese que $2R/\pi \approx 0{,}6366R$ está *más arriba* que el $0{,}4244R$ del semidisco: la curva no tiene masa cerca del centro, sólo en el borde.

(d) La superficie de la esfera por Pappus II.

$$S = 2\pi \cdot \frac{2R}{\pi} \cdot \pi R = 4\pi R^2.$$

Si uno hubiera usado por descuido el centroide del semidisco habría obtenido $2\pi \cdot \frac{4R}{3\pi} \cdot \pi R = \frac{8\pi R^2}{3}$, que no es la superficie de ninguna esfera.

(e) El toro. Un disco de radio r cuyo centro dista $R > r$ del eje tiene, por simetría, su centroide en el propio centro. Entonces

$$V = 2\pi R \cdot \pi r^2 = 2\pi^2 R r^2, \qquad S = 2\pi R \cdot 2\pi r = 4\pi^2 R r.$$

Estas dos fórmulas, obtenidas en dos líneas, son bastante penosas de deducir por discos.

Ejemplo 6.184 (El cono, con volumen y superficie lateral)**.** Tomemos el triángulo de vértices $(0,0)$, $(R,0)$ y $(0,h)$, es decir la región bajo la recta $y = h(1 - \frac{x}{R})$ sobre $[0,R]$, y girémoslo alrededor de OY.

(a) El centroide. $A = \dfrac{Rh}{2}$ y

$$\bar{x} = \frac{2}{Rh}\int_0^R x\, h\Big(1 - \frac{x}{R}\Big)\mathrm{d}x = \frac{2}{R}\Big(\frac{R^2}{2} - \frac{R^2}{3}\Big) = \frac{R}{3},$$

otra vez el baricentro $\frac{1}{3}(0 + R + 0)$.

(b) El volumen. Pappus I:

$$V = 2\pi \cdot \frac{R}{3} \cdot \frac{Rh}{2} = \frac{\pi R^2 h}{3},$$

la fórmula del cono, deducida sin integrar ni un solo disco.

(c) La superficie lateral. La generatriz es el segmento de $(R,0)$ a $(0,h)$, de longitud $L = \sqrt{R^2 + h^2}$, y el centroide de un segmento es su punto medio, con $\bar{x} = R/2$. Pappus II:

$$S = 2\pi \cdot \frac{R}{2} \cdot \sqrt{R^2 + h^2} = \pi R\sqrt{R^2 + h^2},$$

que es la fórmula $\pi R g$ que se aprende en el colegio, con g la generatriz. Aquí se ve otra vez la diferencia entre los dos centroides: para el volumen hizo falta $R/3$ (el de la región), para la superficie $R/2$ (el del borde que gira).

Cuidado

[El eje no puede atravesar la región] Pappus I exige que el eje de giro no corte el interior de la región. Si lo corta, partes de la región se superponen al girar y $2\pi d A$ cuenta ese volumen dos veces (o, peor, lo cancela). El caso extremo es un eje de simetría que pasa por el centroide: entonces $d = 0$ y la fórmula predice $V = 0$, cosa manifiestamente falsa. Por ejemplo, el disco completo $x^2 + y^2 \le R^2$ girado alrededor de OX tiene $\bar{y} = 0$, y Pappus daría $V = 0$, cuando el sólido generado es la esfera de volumen $4\pi R^3/3$. El remedio correcto es el que usamos arriba: quedarse sólo con el semidisco $y \ge 0$, que toca el eje pero no lo cruza.

Pausa para pensar

Las fórmulas de $\bar{x}$ y $\bar{y}$ que hemos usado son un caso particular de las integrales dobles

$$\bar{x} = \frac{1}{A}\iint_D x\mathrm{d}A, \qquad \bar{y} = \frac{1}{A}\iint_D y\mathrm{d}A,$$

del capítulo 9: basta integrar primero en y entre $g(x)$ y $f(x)$ para recuperar el $\frac{1}{2}(f^2-g^2)$ que aparece en $\bar{y}$. Y si en vez de densidad constante uno pone una densidad $\rho(x,y)$, el mismo promedio pasa a llamarse *centro de masa*; el centroide es simplemente el centro de masa del caso homogéneo. ¿Sabría el lector rehacer el apartado (b) del primer ejemplo con densidad $\rho(x,y)=x$ y explicar por qué $\bar{x}$ tiene que salir mayor que $3/4$?

6.37. Ejercicios propuestos

Sin solución desarrollada: solo enunciado y respuesta al final. Úsalos para comprobar que lo del capítulo se sostiene sin la muleta del texto delante.

6.1. Calcula $\displaystyle\lim_{n\to\infty}\frac{3n^2+1}{2n^2-n}$.

6.2. Da el supremo y el ínfimo de $\{1-\frac{1}{n} : n\in\mathbb{N}\}$ y di cuáles se alcanzan.

6.3. ¿Converge $\displaystyle\sum_{n\ge 1} n^{-3/2}$? ¿Y $\displaystyle\sum_{n\ge 2}\frac{1}{n\ln n}$?

6.4. Calcula $\displaystyle\sum_{n\ge 1}\frac{n}{2^n}$.

6.5. Calcula $\displaystyle\sum_{n\ge 1}\frac{1}{n(n+1)}$.

6.6. Da el radio de convergencia de $\displaystyle\sum_{n\ge 1}\frac{x^n}{n^2}$ y di qué pasa en los extremos.

6.7. Escribe el polinomio de Taylor de $\ln(1+x)$ en 0 hasta grado 3.

6.8. Calcula $\displaystyle\lim_{n\to\infty} n\big(\sqrt{n^2+1}-n\big)$.

6.9. ¿Es $f(x)=x^2$ uniformemente continua en $\mathbb{R}$? ¿Y en $[0,1]$?

6.10. Aplica el teorema del valor medio a $f(x)=x^3$ en $[0,2]$ y halla el c que garantiza.

6.11. ¿A qué converge $\displaystyle\sum_{n\ge 1}\frac{(-1)^{n+1}}{n}$, y por qué el reordenamiento de esa serie es peligroso?

6.12. Da un ejemplo de sucesión de Cauchy en $\mathbb{Q}$ que no converja en $\mathbb{Q}$.

Respuestas

1. $\frac{3}{2}$. **2.** Supremo 1 (no se alcanza), ínfimo 0 (sí, en $n=1$). **3.** La primera sí ($p=\frac{3}{2}>1$); la segunda no (criterio integral: $\ln\ln n\to\infty$). **4.** 2. **5.** 1 (serie telescópica). **6.** Radio 1; converge también en $x=\pm 1$, absolutamente. **7.** $x-\frac{x^2}{2}+\frac{x^3}{3}$. **8.** $\frac{1}{2}$. **9.** En $\mathbb{R}$ no; en $[0,1]$ sí (Heine–Cantor). **10.** $3c^2=4$, luego $c=\frac{2}{\sqrt{3}}\approx 1{,}155$. **11.** A $\ln 2$; converge pero no absolutamente, así que por el teorema de reordenamiento de Riemann (capítulo 29) se puede reordenar para que sume lo que uno quiera. **12.** Por ejemplo $a_n=(1+\frac{1}{n})^n$, o las truncaciones decimales de $\sqrt{2}$.

Capítulo 7

Álgebra lineal

Lo que necesitas traer

Del capítulo 1: conjuntos y funciones. Del capítulo 4: los números complejos, porque los valores propios de una matriz real pueden ser complejos. Del capítulo 6: nada esencial, aunque las secciones sobre normas y convergencia de métodos iterativos usan el concepto de límite.

El álgebra lineal estudia los espacios vectoriales y las funciones que respetan su estructura. Es el lenguaje común de todo lo que sigue: los sistemas de ecuaciones diferenciales se resuelven con valores propios, los tensores son objetos multilineales sobre espacios vectoriales, y la teoría de Galois trata las extensiones de cuerpos como espacios vectoriales.

7.1. Sistemas lineales y eliminación de Gauss

Idea clave

Intuición primero. La eliminación de Gauss es el método de eliminación del álgebra elemental escrito de forma sistemática para que lo pueda ejecutar una máquina sin pensar. Lo único que se hace es sumar múltiplos de una ecuación a otra para ir borrando incógnitas, y lo único que hay que creerse es que esas operaciones *no cambian el conjunto de soluciones*: la nueva ecuación es consecuencia de las viejas y viceversa, así que ambos sistemas tienen exactamente los mismos puntos.
La forma escalonada final no es un capricho de presentación: es la disposición en la que la respuesta se lee sin más trabajo. Los pivotes marcan las variables que quedan determinadas; las columnas sin pivote marcan las variables libres, y cada variable libre es una dirección entera de soluciones. De ahí sale, de un vistazo, la trilogía de siempre —una solución, ninguna, o infinitas—, que en dos variables eran dos rectas cortándose, siendo paralelas o coincidiendo.

Un sistema de m ecuaciones con n incógnitas se escribe $A\mathbf{x} = \mathbf{b}$ con $A \in M_{m\times n}$. Las *operaciones elementales de fila* —intercambiar dos filas, multiplicar una fila por un escalar no nulo, sumar a una fila un múltiplo de otra— no cambian el conjunto de soluciones, y permiten

llevar la matriz a *forma escalonada*: cada fila empieza con más ceros que la anterior. El primer coeficiente no nulo de cada fila es un *pivote*; las columnas sin pivote corresponden a *variables libres*.

Teorema 7.1 (Rouché–Frobenius)**.** $A\mathbf{x} = \mathbf{b}$ *es compatible* $\iff \operatorname{rango}(A) = \operatorname{rango}(A|\mathbf{b})$. *En tal caso: solución única si* $\operatorname{rango}(A) = n$*; infinitas soluciones con* $n - \operatorname{rango}(A)$ *parámetros si* $\operatorname{rango}(A) < n$.

Idea clave

Qué dice esto en cristiano. El rango cuenta cuántas ecuaciones *independientes* hay de verdad. Si al añadir la columna de términos independientes el rango sube, es que esa columna aporta información nueva, y esa información nueva solo puede ser una contradicción: el sistema no tiene solución. Si no sube, hay solución. Y entonces la cuenta de las incógnitas libres es inmediata: quedan $n - \operatorname{rango}(A)$ parámetros, porque las ecuaciones independientes fijan $\operatorname{rango}(A)$ de las n incógnitas y el resto queda al aire.

Ejemplo 7.2 (Los tres casos, con matrices de dos por dos)**. (a) Determinado.** El sistema $x + y = 3$, $x - y = 1$ tiene $\operatorname{rango}(A) = \operatorname{rango}(A|\mathbf{b}) = 2 = n$: solución única, $(x, y) = (2, 1)$.

(b) Indeterminado. El sistema $x + y = 3$, $2x + 2y = 6$ tiene $\operatorname{rango}(A) = \operatorname{rango}(A|\mathbf{b}) = 1$ y $n = 2$: hay $2 - 1 = 1$ parámetro libre, y las soluciones son $(3 - t,\, t)$. La segunda ecuación era la primera multiplicada por dos, es decir, no aportaba nada.

(c) Incompatible. El sistema $x+y = 3$, $2x+2y = 7$ tiene $\operatorname{rango}(A) = 1$ pero $\operatorname{rango}(A|\mathbf{b}) = 2$. La segunda ecuación dice $2(x + y) = 7$, es decir $x + y = 3{,}5$, que contradice a la primera. El salto de rango *es* la contradicción.

(d) Un sistema de tres por tres. Para $x + y + z = 6$, $2x - y + z = 3$, $x + 2y - z = 2$ el determinante vale $7 \neq 0$, luego el rango es 3 por ambos lados y la solución es única: $(1, 2, 3)$. Comprobación directa: $1 + 2 + 3 = 6$, $2 - 2 + 3 = 3$ y $1 + 4 - 3 = 2$.

Cuidado

[El rango se calcula, no se adivina] Contar ecuaciones no es contar rango: en (b) hay dos ecuaciones y rango 1. Segundo aviso: un sistema homogéneo ($\mathbf{b} = \mathbf{0}$) es *siempre* compatible, porque $\mathbf{x} = \mathbf{0}$ lo resuelve y la columna de ceros nunca sube el rango; la pregunta interesante ahí es si tiene soluciones distintas de la trivial, y las tiene exactamente cuando $\operatorname{rango}(A) < n$. Y tercero: con más incógnitas que ecuaciones, un sistema homogéneo tiene siempre infinitas soluciones, porque el rango no puede superar el número de filas.

Observación 7.3 (Estructura de las soluciones). Si $\mathbf{x}_p$ es una solución particular de $A\mathbf{x} = \mathbf{b}$, toda solución es $\mathbf{x}_p + \mathbf{x}_h$ con $A\mathbf{x}_h = \mathbf{0}$: *particular más homogénea*. Este patrón —idéntico al de las ecuaciones diferenciales lineales— es un hecho general sobre aplicaciones lineales: la preimagen de un punto es un trasladado del núcleo.

7.1.1. Factorización LU y coste computacional

Si en la eliminación no hacen falta intercambios, el proceso equivale a factorizar $A = LU$ con L triangular inferior (unos en la diagonal, y debajo los multiplicadores usados) y U triangular superior (la escalonada). Resolver $A\mathbf{x} = \mathbf{b}$ se reduce entonces a dos sustituciones triangulares ($L\mathbf{y} = \mathbf{b}$ hacia adelante, $U\mathbf{x} = \mathbf{y}$ hacia atrás). La eliminación cuesta $\sim \frac{2}{3}n^3$ operaciones; cada resolución posterior con la misma A, solo $\sim 2n^2$. Por eso los paquetes numéricos factorizan una vez y resuelven muchas.

7.1.2. Matrices y sus operaciones

El producto $(AB)_{ij} = \sum_k A_{ik}B_{kj}$ está definido cuando el número de columnas de A iguala al de filas de B. No es conmutativo, pero sí asociativo y distributivo, con $(AB)^{\mathsf{T}} = B^{\mathsf{T}}A^{\mathsf{T}}$.

A cuadrada es *invertible* si existe A^{-1} con $AA^{-1} = A^{-1}A = I$; entonces $(AB)^{-1} = B^{-1}A^{-1}$. Cálculo: Gauss–Jordan sobre $[A \,|\, I] \to [I \,|\, A^{-1}]$. Para 2×2:

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix}^{-1} = \frac{1}{ad - bc}\begin{pmatrix} d & -b \\ -c & a \end{pmatrix}.$$

Teorema 7.4 (Caracterizaciones de invertibilidad)**.** *Para A de tamaño $n \times n$, son equivalentes: (i) A es invertible; (ii) $A\mathbf{x} = \mathbf{0}$ solo tiene la solución trivial; (iii)* rango $A = n$*; (iv) las columnas son linealmente independientes; (v)* $\det A \neq 0$*; (vi) 0 no es valor propio de A; (vii) $A\mathbf{x} = \mathbf{b}$ tiene solución única para todo $\mathbf{b}$.*

Idea clave

Qué dice esto en cristiano. Este teorema dice que una docena de preguntas aparentemente distintas son *la misma pregunta.* ¿Tiene inversa? ¿Es su determinante no nulo? ¿Tienen sus columnas rango máximo? ¿Es su núcleo trivial? ¿Es cero un autovalor? Todas se responden a la vez y con el mismo sí o el mismo no. En la práctica esto es enormemente cómodo: se comprueba la que sea más barata de calcular en cada caso, normalmente el determinante en tamaños pequeños y una eliminación gaussiana en tamaños grandes.

Ejemplo 7.5 (Una invertible y una que no, comprobadas por todos los caminos)**. (a) Una matriz invertible.** Sea $A = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}$. Su determinante es $1 \cdot 4 - 2 \cdot 3 = -2 \neq 0$; su rango es 2; su núcleo es solo el vector cero; y su inversa es $\begin{pmatrix} -2 & 1 \\ 3/2 & -1/2 \end{pmatrix}$, como se comprueba multiplicando y obteniendo la identidad. Los cinco criterios dicen lo mismo.

(b) Una singular. Sea $B = \begin{pmatrix} 1 & 2 \\ 2 & 4 \end{pmatrix}$. Su determinante es $1 \cdot 4 - 2 \cdot 2 = 0$; su rango es 1, porque la segunda fila es el doble de la primera; su núcleo contiene al vector $(-2, 1)$, ya que $1 \cdot (-2) + 2 \cdot 1 = 0$; y sus autovalores son 5 y 0. Otra vez, todos los criterios coinciden.

(c) El autovalor cero es la clave. B tiene el autovalor 0 precisamente porque tiene núcleo no trivial: un vector del núcleo cumple $B\mathbf{v} = \mathbf{0} = 0 \cdot \mathbf{v}$. Por eso "0 no es autovalor" y "el núcleo es trivial" son la misma afirmación escrita de dos maneras.

(d) La lectura como sistema. Que A sea invertible equivale a que $A\mathbf{x} = \mathbf{b}$ tenga solución única *para todo* $\mathbf{b}$; que B no lo sea equivale a que, según el $\mathbf{b}$, el sistema no tenga solución o tenga infinitas. Es Rouché–Frobenius visto desde la matriz en vez de desde el sistema.

Cuidado

[Estas equivalencias solo valen para matrices cuadradas] Una matriz rectangular nunca tiene inversa, aunque su rango sea máximo; a lo sumo tiene inversa por un lado. La matriz $\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}$ tiene rango 2 y no es invertible, porque su núcleo tiene dimensión 1. Segundo aviso: que el determinante sea "pequeño" no es lo mismo que ser singular, pero sí es un aviso serio; la matriz $\begin{pmatrix} 1 & 1 \\ 1 & 1{,}0001 \end{pmatrix}$ tiene determinante 10^{-4}, es invertible, y su número de condición es 40 002: numéricamente se comporta casi como si no lo fuera.

Esta lista de equivalencias es el "diccionario" central del álgebra lineal; conviene poder justificar cada flecha.

Regla: resolver un sistema lineal y leer correctamente el resultado

La eliminación de Gauss es un algoritmo, y como tal se ejecuta sin pensar. Lo que sí exige pensar es la lectura del resultado, y ahí es donde se pierden los puntos.

1. **Escribe la matriz ampliada.** Coloca $[A \mid b]$ y no vuelvas a escribir las incógnitas hasta el final. Todo el trabajo es sobre números.

2. **Baja escalonando, columna por columna.** En cada columna elige un pivote no nulo (intercambiando filas si hace falta) y usa la fila del pivote para anular todo lo que hay debajo. Las tres operaciones permitidas son: intercambiar dos filas, multiplicar una fila por un escalar no nulo y sumar a una fila un múltiplo de otra. *Nunca* operes con columnas: eso mezcla las incógnitas.

3. **Diagnostica antes de resolver.** Con la matriz ya escalonada, compara rangos:

   - si $\operatorname{rango}(A) < \operatorname{rango}([A \mid b])$ — es decir, si aparece una fila $(0\ 0 \cdots 0 \mid c)$ con $c \neq 0$ — el sistema es **incompatible**;
   - si $\operatorname{rango}(A) = \operatorname{rango}([A \mid b]) = n$ (número de incógnitas), hay **solución única**;
   - si $\operatorname{rango}(A) = \operatorname{rango}([A \mid b]) = r < n$, hay **infinitas soluciones** con $n - r$ parámetros libres.

   Este es el teorema de Rouché–Frobenius, y se aplica *mirando*, sin cuentas adicionales.

4. **Identifica variables pivote y variables libres.** Las columnas con pivote corresponden a las incógnitas determinadas; las columnas sin pivote, a los parámetros.

El número de parámetros es exactamente $n - r$, y coincide con la dimensión del núcleo.

5. **Sube despejando, o mejor, reduce del todo.** Puedes hacer sustitución hacia atrás, o seguir eliminando hacia arriba hasta la forma escalonada reducida, donde la solución se lee directamente. La segunda opción cuesta un poco más pero se equivoca uno mucho menos.

6. **Escribe la solución en forma vectorial.** Solución general = solución particular + combinación de vectores del núcleo:

$$x = x_p + t_1 v_1 + \cdots + t_{n-r} v_{n-r}.$$

Esta forma es la única que exhibe la estructura y la única que sirve para lo que viene después.

7. **Control obligatorio.** Sustituye la solución en las ecuaciones *originales*. Cuesta treinta segundos y detecta el noventa por ciento de los errores aritméticos.

Ejemplo 7.6 (Un sistema con infinitas soluciones, resuelto y leído)**.** Resolvamos

$$\begin{cases} x + 2y - z = 3 \\ 2x + 5y + z = 8 \\ 3x + 7y \qquad = 11 \end{cases}$$

Eliminación. Matriz ampliada y operaciones $F_2 \to F_2 - 2F_1$, $F_3 \to F_3 - 3F_1$:

$$\left[\begin{array}{ccc|c} 1 & 2 & -1 & 3 \\ 2 & 5 & 1 & 8 \\ 3 & 7 & 0 & 11 \end{array}\right] \longrightarrow \left[\begin{array}{ccc|c} 1 & 2 & -1 & 3 \\ 0 & 1 & 3 & 2 \\ 0 & 1 & 3 & 2 \end{array}\right] \longrightarrow \left[\begin{array}{ccc|c} 1 & 2 & -1 & 3 \\ 0 & 1 & 3 & 2 \\ 0 & 0 & 0 & 0 \end{array}\right].$$

La tercera ecuación era la suma de las dos primeras; no aportaba información.

Diagnóstico. $\operatorname{rango}(A) = \operatorname{rango}([A \mid b]) = 2$ y hay $n = 3$ incógnitas. Sistema compatible indeterminado con $3 - 2 = 1$ parámetro.

Reducción completa. Con $F_1 \to F_1 - 2F_2$:

$$\left[\begin{array}{ccc|c} 1 & 0 & -7 & -1 \\ 0 & 1 & 3 & 2 \\ 0 & 0 & 0 & 0 \end{array}\right].$$

Las columnas con pivote son la 1 y la 2, luego x e y son variables pivote y z es libre. Poniendo $z = t$:

$$x = -1 + 7t, \qquad y = 2 - 3t, \qquad z = t.$$

Forma vectorial.

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} -1 \\ 2 \\ 0 \end{pmatrix} + t \begin{pmatrix} 7 \\ -3 \\ 1 \end{pmatrix}.$$

Geométricamente: la solución es una recta en $\mathbb{R}^3$, la intersección de dos planos (el tercero era redundante). El vector $(7,-3,1)$ genera el núcleo de A, de dimensión $3-\text{rango}(A)=1$, como manda el teorema del rango.

Control. Con $t=1$ sale $(6,-1,1)$; se sustituye: $6-2-1=3$ ✓, $12-5+1=8$ ✓, $18-7=11$ ✓.

Cuidado

El error más común al leer un sistema escalonado es contar mal los parámetros. El número de parámetros libres es $n-\text{rango}(A)$, donde n es el número de *incógnitas*, no el de ecuaciones. Un sistema con cinco ecuaciones y tres incógnitas puede perfectamente tener infinitas soluciones, y uno con dos ecuaciones y dos incógnitas puede ser incompatible.

7.2. Matrices: qué son, cómo se operan y cómo escriben un sistema

Hasta este punto hemos hablado de matrices como si el lector ya supiera qué son. Hemos escrito $A\mathbf{x}=\mathbf{b}$, hemos manipulado matrices ampliadas y hemos usado el símbolo A^{T} sin haber dicho nunca qué significa ninguna de esas cosas. Esta sección salda esa deuda desde cero. No supone absolutamente nada: empieza definiendo qué es una tabla de números, termina con el algoritmo de Gauss escrito con todo detalle, y por el camino define cada operación, demuestra cada propiedad y calcula entero cada ejemplo. Quien ya sepa todo esto puede saltársela; quien no, encontrará aquí el suelo sobre el que se apoya el resto del capítulo.

Idea clave

Intuición primero. Una matriz no es más que una tabla rectangular de números. Lo que la convierte en un objeto matemático interesante no es la tabla, sino las *operaciones* que se definen sobre ella, y en particular una: el producto. El producto de matrices tiene una regla que a primera vista parece arbitraria y retorcida —cada entrada del resultado es una suma de productos cruzados—, y casi todo el mundo la aprende de memoria sin entender de dónde sale. Sale de un sitio muy concreto: de *encadenar dos cambios de variable*. Si unas variables se escriben en función de otras, y esas otras en función de unas terceras, sustituir y simplificar produce exactamente esa regla y ninguna otra. El producto de matrices es la sustitución hecha aritmética.

Con esa idea en la mano, todo lo demás cae por su peso: que el producto no sea conmutativo es que hacer dos cambios de variable en un orden no es lo mismo que hacerlos en el otro; que sí sea asociativo es que encadenar tres sustituciones da lo mismo agrupándolas como se quiera; y que un sistema de ecuaciones se escriba $A\mathbf{x}=\mathbf{b}$ es simplemente que un sistema *es* un cambio de variable del que conocemos la salida y buscamos la entrada.

7.2.1. Qué es exactamente una matriz

Muchos problemas producen de manera natural una tabla de números con dos índices: los coeficientes de un sistema de ecuaciones (una fila por ecuación, una columna por incógnita), los

precios de varios productos en varias tiendas, las distancias entre ciudades, las probabilidades de pasar de un estado a otro. En todos esos casos el dato no es un número ni una lista, sino una retícula: hace falta decir en qué fila y en qué columna está cada número. Ese objeto es una matriz.

Definición 7.7. Sea K un cuerpo (para todo este capítulo basta pensar en $K = \mathbb{R}$ o $K = \mathbb{C}$). Una *matriz* de tamaño $m \times n$ sobre K es una tabla rectangular de $m \cdot n$ elementos de K dispuestos en m *filas* y n *columnas*:

$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix}.$$

El elemento a_{ij} situado en la fila i y la columna j se llama *entrada* de lugar (i, j); el primer índice es *siempre* el de la fila y el segundo *siempre* el de la columna. Se abrevia $A = (a_{ij})$, y también se escribe A_{ij} o $(A)_{ij}$ para la entrada. El par (m, n) es el *tamaño* o *dimensión* de la matriz. El conjunto de todas las matrices $m \times n$ sobre K se denota $M_{m\times n}(K)$; si $m = n$ la matriz se llama *cuadrada* de orden n y se escribe $M_n(K)$.

Ejemplo 7.8 (Leer una matriz entrada a entrada)**.** Sea

$$A = \begin{pmatrix} 4 & -1 & 0 \\ 7 & 2 & -5 \end{pmatrix}.$$

Tiene dos filas y tres columnas, luego su tamaño es 2×3 y $A \in M_{2\times 3}(\mathbb{R})$. Sus seis entradas son $a_{11} = 4$, $a_{12} = -1$, $a_{13} = 0$, $a_{21} = 7$, $a_{22} = 2$ y $a_{23} = -5$. La primera fila es $(4, -1, 0)$ y la segunda es $(7, 2, -5)$; la primera columna es $\left(\begin{smallmatrix} 4 \\ 7 \end{smallmatrix}\right)$, la segunda $\left(\begin{smallmatrix} -1 \\ 2 \end{smallmatrix}\right)$ y la tercera $\left(\begin{smallmatrix} 0 \\ -5 \end{smallmatrix}\right)$. Obsérvese que $a_{23} = -5$ y que a_{32} *no existe*: no hay tercera fila.

Cuidado

[El orden de los índices no es negociable] En a_{ij}, el primer índice cuenta filas (hacia abajo) y el segundo columnas (hacia la derecha). Es una convención universal, y confundirla es equivalente a trabajar todo el rato con la matriz traspuesta sin darse cuenta, lo que produce resultados que casi cuadran y nunca cuadran del todo. Regla mnemotécnica: se lee como se lee una página, primero la línea y luego la posición dentro de ella. Segundo aviso: en "$m \times n$" el primer número es el de filas, con la misma lógica.

Dos matrices son el mismo objeto cuando lo son casilla por casilla, y para eso hace falta antes que las casillas se correspondan, es decir, que los tamaños coincidan.

Definición 7.9. Dos matrices $A = (a_{ij})$ y $B = (b_{ij})$ son *iguales*, y se escribe $A = B$, cuando tienen el mismo tamaño $m \times n$ y además $a_{ij} = b_{ij}$ para todos los índices $1 \le i \le m$, $1 \le j \le n$. Una igualdad entre matrices $m \times n$ equivale, por tanto, a $m \cdot n$ igualdades numéricas simultáneas.

Esta última frase es más útil de lo que parece: cada vez que en el resto del capítulo se demuestre una identidad entre matrices, la demostración consistirá en comprobar que las entradas de lugar (i, j) de ambos lados coinciden, para i y j arbitrarios. No hay otro camino, y siempre es ese.

Dos tamaños merecen nombre propio. Una matriz $1 \times n$, es decir, con una sola fila, se llama *matriz fila*; una matriz $m \times 1$, con una sola columna, se llama *matriz columna*. A lo largo de todo el libro adoptamos el **convenio de que un vector de K^n es una matriz columna** $n \times 1$:

$$\mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} \in M_{n\times 1}(K).$$

Cuando por comodidad tipográfica escribamos $\mathbf{x} = (x_1, \ldots, x_n)$ en una línea de texto, hay que entender siempre la columna. El convenio no es un capricho: es lo que hace que la expresión $A\mathbf{x}$ tenga sentido con la regla del producto que veremos enseguida, y es lo que obliga a escribir las matrices de las transformaciones lineales por columnas y no por filas.

7.2.2. Un catálogo de matrices con nombre

Ciertas formas de tabla aparecen tantas veces que tienen nombre propio. Conviene fijarlas ahora, porque el resto del capítulo las usa constantemente y porque casi todas ellas se reconocen de un vistazo, sin calcular nada.

Definición 7.10. En una matriz cuadrada $A = (a_{ij})$ de orden n, las entradas $a_{11}, a_{22}, \ldots, a_{nn}$ forman la *diagonal principal*. Las entradas con $i < j$ están *por encima* de la diagonal y las de $i > j$, *por debajo*.

Definición 7.11. La *matriz nula* $0_{m\times n}$ es la matriz de tamaño $m \times n$ con todas sus entradas iguales a 0; se escribe simplemente 0 cuando el tamaño se deduce del contexto. La *matriz identidad* de orden n, denotada I_n o simplemente I, es la matriz cuadrada con unos en la diagonal principal y ceros fuera de ella; en símbolos, $(I_n)_{ij} = \delta_{ij}$, donde δ_{ij} vale 1 si $i = j$ y 0 si $i \neq j$ (la *delta de Kronecker*).

Definición 7.12. Sea A cuadrada de orden n. Se dice que A es: *diagonal* si $a_{ij} = 0$ siempre que $i \neq j$, y entonces se escribe $A = \operatorname{diag}(a_{11}, \ldots, a_{nn})$; *escalar* si es diagonal y todas las entradas de la diagonal son iguales, es decir, si $A = \lambda I$ para algún $\lambda \in K$; *triangular superior* si $a_{ij} = 0$ siempre que $i > j$ (todo lo que hay debajo de la diagonal es cero); *triangular inferior* si $a_{ij} = 0$ siempre que $i < j$ (todo lo que hay encima es cero); *simétrica* si $a_{ij} = a_{ji}$ para todos i, j; y *antisimétrica* si $a_{ij} = -a_{ji}$ para todos i, j.

Nótese una consecuencia inmediata de la última definición: si A es antisimétrica, tomando $i = j$ queda $a_{ii} = -a_{ii}$, es decir $2a_{ii} = 0$, luego *la diagonal de una matriz antisimétrica es toda de ceros* (siempre que en K el 2 sea invertible, lo cual ocurre en $\mathbb{R}$ y en $\mathbb{C}$).

Ejemplo 7.13 (Un ejemplar de cada especie)**.** De izquierda a derecha: la nula 2×3, la identidad de orden 3 y una diagonal de orden 3.

$$0 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \qquad I_3 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \qquad \operatorname{diag}(2,-1,5) = \begin{pmatrix} 2 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 5 \end{pmatrix}.$$

Una escalar, una triangular superior y una triangular inferior:

$$3I_2 = \begin{pmatrix} 3 & 0 \\ 0 & 3 \end{pmatrix}, \qquad \begin{pmatrix} 1 & 4 & -2 \\ 0 & 5 & 7 \\ 0 & 0 & 3 \end{pmatrix}, \qquad \begin{pmatrix} 1 & 0 & 0 \\ 6 & 2 & 0 \\ -3 & 8 & 4 \end{pmatrix}.$$

Una simétrica y una antisimétrica:

$$S = \begin{pmatrix} 1 & 2 & -3 \\ 2 & 0 & 4 \\ -3 & 4 & 5 \end{pmatrix}, \qquad N = \begin{pmatrix} 0 & 2 & -3 \\ -2 & 0 & 4 \\ 3 & -4 & 0 \end{pmatrix}.$$

En S, cada entrada coincide con su reflejada respecto de la diagonal: $s_{12} = s_{21} = 2$, $s_{13} = s_{31} = -3$, $s_{23} = s_{32} = 4$. En N ocurre lo mismo pero con el signo cambiado: $n_{12} = 2$ y $n_{21} = -2$, y la diagonal es forzosamente nula, como acabamos de razonar. Obsérvese además que toda matriz diagonal es a la vez triangular superior e inferior, y también simétrica: los nombres no son excluyentes.

7.2.3. Sumar matrices y multiplicarlas por un número

Las dos primeras operaciones son las que uno esperaría, y se definen casilla a casilla. Si dos tablas registran las ventas de enero y las de febrero, la tabla del primer trimestre parcial se obtiene sumando cada casilla con la casilla correspondiente; y si se quiere duplicar toda la producción, se multiplica cada casilla por dos. Nada más.

Definición 7.14. Sean $A = (a_{ij})$ y $B = (b_{ij})$ dos matrices *del mismo tamaño* $m \times n$, y sea $\lambda \in K$. Se definen la *suma* $A + B$ y el *producto por el escalar* λA como las matrices $m \times n$ dadas por

$$(A + B)_{ij} = a_{ij} + b_{ij}, \qquad (\lambda A)_{ij} = \lambda\, a_{ij}.$$

Se escribe $-A$ por $(-1)A$ y $A - B$ por $A + (-B)$.

Ejemplo 7.15 (Suma y escalado, con todas las casillas)**.** Con $A = \begin{pmatrix} 4 & -1 & 0 \\ 7 & 2 & -5 \end{pmatrix}$ y $B = \begin{pmatrix} 1 & 3 & -2 \\ 0 & -2 & 6 \end{pmatrix}$, ambas 2×3:

$$A + B = \begin{pmatrix} 4+1 & -1+3 & 0+(-2) \\ 7+0 & 2+(-2) & -5+6 \end{pmatrix} = \begin{pmatrix} 5 & 2 & -2 \\ 7 & 0 & 1 \end{pmatrix},$$

$$3A = \begin{pmatrix} 3\cdot 4 & 3\cdot(-1) & 3\cdot 0 \\ 3\cdot 7 & 3\cdot 2 & 3\cdot(-5) \end{pmatrix} = \begin{pmatrix} 12 & -3 & 0 \\ 21 & 6 & -15 \end{pmatrix},$$

$$A - B = \begin{pmatrix} 4-1 & -1-3 & 0+2 \\ 7-0 & 2+2 & -5-6 \end{pmatrix} = \begin{pmatrix} 3 & -4 & 2 \\ 7 & 4 & -11 \end{pmatrix}.$$

Proposición 7.16. *Para matrices $A, B, C \in M_{m\times n}(K)$ y escalares $\lambda, \mu \in K$ se cumplen:*

$A + B = B + A$	*(conmutativa)*
$(A + B) + C = A + (B + C)$	*(asociativa)*
$A + 0 = A$	*(elemento neutro: la matriz nula)*
$A + (-A) = 0$	*(elemento opuesto)*
$\lambda(A + B) = \lambda A + \lambda B$	*(distributiva respecto de la suma de matrices)*
$(\lambda + \mu)A = \lambda A + \mu A$	*(distributiva respecto de la suma de escalares)*
$(\lambda\mu)A = \lambda(\mu A), \quad 1A = A$	*(compatibilidad con el producto de escalares)*

Demostración. Todas se demuestran igual, mirando la entrada de lugar (i, j) y usando la propiedad correspondiente en el cuerpo K. Por ejemplo, para la primera: $(A+B)_{ij} = a_{ij}+b_{ij} = b_{ij} + a_{ij} = (B + A)_{ij}$, donde la igualdad central es la conmutatividad de la suma en K. Como las entradas coinciden y los tamaños también, las matrices son iguales por la definición 7.9. Las demás son idénticas palabra por palabra. □

Estas siete propiedades son exactamente los axiomas de espacio vectorial, y por eso $M_{m\times n}(K)$ aparecerá en la próxima sección como uno de los ejemplos básicos: las matrices de un tamaño fijo forman un espacio vectorial de dimensión mn.

Cuidado

[Solo se suman matrices del mismo tamaño] La expresión $A + B$ carece de sentido si A es 2×3 y B es 3×2: no hay manera de emparejar las casillas. No es que la suma valga cero o que haya que rellenar con ceros; es que la operación *no está definida*. Lo mismo ocurre con $A - B$. En cambio λA está siempre definido, sea cual sea el tamaño, porque no hay nada que emparejar. Y ojo con una tentación frecuente: el producto de matrices, que viene a continuación, *no* se define casilla a casilla, y sus condiciones de tamaño son otras completamente distintas.

7.2.4. La traspuesta

Una tabla se puede leer por filas o por columnas, y a veces la información pide un giro. Si A registra "lo que la tienda i vende del producto j", puede convenir la tabla "lo que del producto j vende la tienda i", que es la misma información con los papeles de fila y columna intercambiados. Esa operación de girar la tabla sobre su diagonal principal es la trasposición, y es el símbolo A^{T} que hemos venido usando sin haberlo definido.

Definición 7.17. Sea $A = (a_{ij}) \in M_{m\times n}(K)$. Su *traspuesta* es la matriz $A^{\mathsf{T}} \in M_{n\times m}(K)$ cuya entrada de lugar (i, j) es

$$(A^{\mathsf{T}})_{ij} = a_{ji}.$$

Es decir: la fila i de A^{T} es la columna i de A, y la columna j de A^{T} es la fila j de A. Nótese el cambio de tamaño: si A es $m \times n$, entonces A^{T} es $n \times m$.

Ejemplo 7.18 (Trasponer una 2×3 y obtener una 3×2)**.** Sea

$$A = \begin{pmatrix} 1 & -2 & 4 \\ 0 & 3 & 5 \end{pmatrix} \in M_{2\times 3}(\mathbb{R}).$$

Sus entradas son $a_{11} = 1$, $a_{12} = -2$, $a_{13} = 4$, $a_{21} = 0$, $a_{22} = 3$, $a_{23} = 5$. Aplicando la definición, $(A^{\mathsf{T}})_{11} = a_{11} = 1$, $(A^{\mathsf{T}})_{12} = a_{21} = 0$, $(A^{\mathsf{T}})_{21} = a_{12} = -2$, $(A^{\mathsf{T}})_{22} = a_{22} = 3$, $(A^{\mathsf{T}})_{31} = a_{13} = 4$ y $(A^{\mathsf{T}})_{32} = a_{23} = 5$. Colocando esas seis entradas en su sitio:

$$A^{\mathsf{T}} = \begin{pmatrix} 1 & 0 \\ -2 & 3 \\ 4 & 5 \end{pmatrix} \in M_{3\times 2}(\mathbb{R}).$$

La primera fila de A, que era $(1, -2, 4)$, se ha convertido en la primera columna de A^{T}; y la primera columna de A, que era $\left(\begin{smallmatrix} 1 \\ 0 \end{smallmatrix}\right)$, es ahora la primera fila. Trasponiendo otra vez se recupera A, como es evidente al mirar el dibujo y como demostraremos en un momento.

Con la traspuesta a mano, dos de los nombres del catálogo se reescriben de forma mucho más limpia y mucho más útil: son las formulaciones que se usan en la práctica.

Proposición 7.19. *Una matriz cuadrada A es simétrica si y solo si $A = A^{\mathsf{T}}$, y es antisimétrica si y solo si $A^{\mathsf{T}} = -A$.*

Demostración. Por la definición 7.9, la igualdad $A = A^{\mathsf{T}}$ significa que $a_{ij} = (A^{\mathsf{T}})_{ij} = a_{ji}$ para todos i, j, que es literalmente la condición de simetría de la definición 7.12. Y $A^{\mathsf{T}} = -A$ significa $a_{ji} = -a_{ij}$ para todos i, j, que es la condición de antisimetría. □

Proposición 7.20. *Para $A, B \in M_{m\times n}(K)$ y $\lambda \in K$:*

$$(A^{\mathsf{T}})^{\mathsf{T}} = A, \qquad (A + B)^{\mathsf{T}} = A^{\mathsf{T}} + B^{\mathsf{T}}, \qquad (\lambda A)^{\mathsf{T}} = \lambda\, A^{\mathsf{T}}.$$

Demostración. Entrada a entrada, que es el único método. Para la primera, aplicando dos veces la definición 7.17: $\left((A^{\mathsf{T}})^{\mathsf{T}}\right)_{ij} = (A^{\mathsf{T}})_{ji} = a_{ij}$, y los tamaños son los de A porque trasponer dos veces devuelve $m \times n$. Para la segunda: $\left((A + B)^{\mathsf{T}}\right)_{ij} = (A + B)_{ji} = a_{ji} + b_{ji} = (A^{\mathsf{T}})_{ij} + (B^{\mathsf{T}})_{ij} = (A^{\mathsf{T}} + B^{\mathsf{T}})_{ij}$. Y la tercera: $\left((\lambda A)^{\mathsf{T}}\right)_{ij} = (\lambda A)_{ji} = \lambda a_{ji} = \lambda(A^{\mathsf{T}})_{ij}$. □

Falta la cuarta propiedad, la del producto, que es la más importante y la más sorprendente porque invierte el orden de los factores. No podemos demostrarla todavía por la razón evidente de que aún no hemos definido el producto; lo haremos en cuanto lo definamos.

7.2.5. Por qué el producto de matrices es como es

Antes de escribir la fórmula conviene ver de dónde sale, porque escrita en frío es incomprensible. Supongamos que dos variables x_1, x_2 se expresan en función de otras dos y_1, y_2, y que estas a su vez se expresan en función de z_1, z_2:

$$\begin{cases} x_1 = \ \ y_1 + 4y_2 \\ x_2 = 5y_1 - 2y_2 \end{cases} \qquad \begin{cases} y_1 = 2z_1 + 3z_2 \\ y_2 = \ \ z_1 - \ \ z_2. \end{cases}$$

Es natural preguntarse cómo dependen las x de las z directamente. Se sustituye y se agrupa:

$$x_1 = (2z_1 + 3z_2) + 4(z_1 - z_2) = 6z_1 - z_2,$$

$$x_2 = 5(2z_1 + 3z_2) - 2(z_1 - z_2) = (10 - 2)z_1 + (15 + 2)z_2 = 8z_1 + 17z_2.$$

Ahora escribamos los tres bloques de coeficientes como tablas:

$$A = \begin{pmatrix} 1 & 4 \\ 5 & -2 \end{pmatrix}, \qquad B = \begin{pmatrix} 2 & 3 \\ 1 & -1 \end{pmatrix}, \qquad C = \begin{pmatrix} 6 & -1 \\ 8 & 17 \end{pmatrix}.$$

La tabla C, la de la sustitución encadenada, se obtiene de A y B por una regla concretísima: el coeficiente de z_2 en x_1 salió de recorrer la *primera fila* de A y la *segunda columna* de B multiplicando término a término y sumando, $1 \cdot 3 + 4 \cdot (-1) = -1$; el coeficiente de z_1 en x_2 salió de la segunda fila de A y la primera columna de B, $5 \cdot 2 + (-2) \cdot 1 = 8$. Esa es la regla, y no hay margen: la sustitución la impone. Si en lugar de eso definiéramos el producto multiplicando casilla con casilla —lo que daría $\left(\begin{smallmatrix} 2 & 12 \\ 5 & 2 \end{smallmatrix}\right)$— obtendríamos una operación perfectamente respetable pero completamente inútil, porque no representaría nada: no sería la tabla de ninguna sustitución encadenada.

Observación 7.21 (La condición de tamaños). Para que la sustitución tenga sentido, el número de variables intermedias tiene que casar: si A expresa m variables en función de n intermedias, B tiene que expresar esas mismas n intermedias en función de p finales. De ahí la regla de compatibilidad, que conviene visualizar así:

$$\underbrace{(m \times \boxed{n})}_{A} \cdot \underbrace{(\boxed{n} \times p)}_{B} = \underbrace{(m \times p)}_{AB}.$$

Los dos números centrales deben coincidir y *desaparecen*; los dos exteriores dan el tamaño del resultado. Si no coinciden, el producto no existe.

Definición 7.22. Sean $A = (a_{ik}) \in M_{m\times n}(K)$ y $B = (b_{kj}) \in M_{n\times p}(K)$. Su *producto* es la matriz $AB \in M_{m\times p}(K)$ cuya entrada de lugar (i, j) es

$$(AB)_{ij} = \sum_{k=1}^{n} a_{ik}b_{kj} = a_{i1}b_{1j} + a_{i2}b_{2j} + \cdots + a_{in}b_{nj},$$

es decir, la fila i de A recorrida a la vez que la columna j de B, multiplicando los términos que se encuentran y sumando todos los productos.

Ejemplo 7.23 (Un producto 2×3 por 3×2, entrada por entrada)**.** Sean

$$A = \begin{pmatrix} 2 & -1 & 3 \\ 0 & 4 & 1 \end{pmatrix} \ (2 \times 3), \qquad B = \begin{pmatrix} 1 & 2 \\ 5 & 0 \\ -1 & 3 \end{pmatrix} \ (3 \times 2).$$

Los tamaños casan: $(2 \times \boxed{3})(\boxed{3} \times 2)$, luego AB existe y es 2×2. Hay cuatro entradas que calcular, y cada una es una suma de tres productos:

$$(AB)_{11} = 2 \cdot 1 + (-1) \cdot 5 + 3 \cdot (-1) = 2 - 5 - 3 = -6,$$

$$(AB)_{12} = 2 \cdot 2 + (-1) \cdot 0 + 3 \cdot 3 = 4 + 0 + 9 = 13,$$

$$(AB)_{21} = 0 \cdot 1 + 4 \cdot 5 + 1 \cdot (-1) = 0 + 20 - 1 = 19,$$

$$(AB)_{22} = 0 \cdot 2 + 4 \cdot 0 + 1 \cdot 3 = 0 + 0 + 3 = 3.$$

Por tanto

$$AB = \begin{pmatrix} -6 & 13 \\ 19 & 3 \end{pmatrix}.$$

El producto en el otro orden también existe, porque $(3 \times \boxed{2})(\boxed{2} \times 3)$ también casa, pero da una matriz 3×3, que ni siquiera tiene el mismo tamaño:

$$BA = \begin{pmatrix} 1 \cdot 2 + 2 \cdot 0 & 1 \cdot (-1) + 2 \cdot 4 & 1 \cdot 3 + 2 \cdot 1 \\ 5 \cdot 2 + 0 \cdot 0 & 5 \cdot (-1) + 0 \cdot 4 & 5 \cdot 3 + 0 \cdot 1 \\ -1 \cdot 2 + 3 \cdot 0 & -1 \cdot (-1) + 3 \cdot 4 & -1 \cdot 3 + 3 \cdot 1 \end{pmatrix} = \begin{pmatrix} 2 & 7 & 5 \\ 10 & -5 & 15 \\ -2 & 13 & 0 \end{pmatrix}.$$

Ejemplo 7.24 (Dos matrices 2×2 multiplicadas en los dos órdenes)**.** Sean

$$C = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}, \qquad D = \begin{pmatrix} 0 & -1 \\ 2 & 1 \end{pmatrix}.$$

Aquí ambos productos existen y ambos son 2×2, así que la comparación es limpia. Primero CD, entrada por entrada:

$$(CD)_{11} = 1 \cdot 0 + 2 \cdot 2 = 4, \qquad (CD)_{12} = 1 \cdot (-1) + 2 \cdot 1 = -1 + 2 = 1,$$

$$(CD)_{21} = 3 \cdot 0 + 4 \cdot 2 = 8, \qquad (CD)_{22} = 3 \cdot (-1) + 4 \cdot 1 = -3 + 4 = 1,$$

de donde $CD = \begin{pmatrix} 4 & 1 \\ 8 & 1 \end{pmatrix}$. Ahora DC:

$$(DC)_{11} = 0 \cdot 1 + (-1) \cdot 3 = -3, \qquad (DC)_{12} = 0 \cdot 2 + (-1) \cdot 4 = -4,$$

$$(DC)_{21} = 2 \cdot 1 + 1 \cdot 3 = 5, \qquad (DC)_{22} = 2 \cdot 2 + 1 \cdot 4 = 8,$$

de donde $DC = \begin{pmatrix} -3 & -4 \\ 5 & 8 \end{pmatrix}$. Las dos matrices son distintas en las cuatro entradas: $CD \neq DC$. Este par es el contraejemplo que conviene tener memorizado.

Proposición 7.25. *Siempre que los tamaños hagan definidos los productos que aparecen:*

$$(AB)C = A(BC), \qquad A(B+C) = AB + AC, \qquad (A+B)C = AC + BC,$$

$$\lambda(AB) = (\lambda A)B = A(\lambda B), \qquad I_m A = A = A I_n \ \text{ si } A \text{ es } m \times n, \qquad 0 \cdot A = 0.$$

Demostración. Demostramos la asociativa, que es la única con algo de trabajo; las demás salen del mismo modo y más rápido. Sean A de tamaño $m \times n$, B de tamaño $n \times p$ y C de tamaño $p \times q$, de modo que ambos lados son $m \times q$. Calculamos la entrada (i,j) del lado izquierdo aplicando dos veces la definición 7.22:

$$\big((AB)C\big)_{ij} = \sum_{l=1}^{p} (AB)_{il}\, c_{lj} = \sum_{l=1}^{p} \left(\sum_{k=1}^{n} a_{ik} b_{kl} \right) c_{lj} = \sum_{l=1}^{p} \sum_{k=1}^{n} a_{ik} b_{kl} c_{lj},$$

y ahora la del lado derecho:

$$\big(A(BC)\big)_{ij} = \sum_{k=1}^{n} a_{ik} (BC)_{kj} = \sum_{k=1}^{n} a_{ik} \left(\sum_{l=1}^{p} b_{kl} c_{lj} \right) = \sum_{k=1}^{n} \sum_{l=1}^{p} a_{ik} b_{kl} c_{lj}.$$

Ambas expresiones son la misma suma finita de los np productos $a_{ik} b_{kl} c_{lj}$, recorrida en distinto orden; como la suma en K es conmutativa y asociativa, coinciden. Para la identidad: $(AI_n)_{ij} = \sum_k a_{ik}\delta_{kj} = a_{ij}$, porque todos los sumandos con $k \neq j$ se anulan y queda uno solo. □

Teorema 7.26. *Si A es $m \times n$ y B es $n \times p$, entonces*

$$(AB)^{\mathsf{T}} = B^{\mathsf{T}} A^{\mathsf{T}}.$$

Demostración. Empecemos por comprobar que los tamaños son coherentes, cosa que ya es reveladora: AB es $m \times p$, luego $(AB)^{\mathsf{T}}$ es $p \times m$; por otro lado B^{T} es $p \times n$ y A^{T} es $n \times m$, así que $B^{\mathsf{T}} A^{\mathsf{T}}$ es $p \times m$. Coinciden. En cambio $A^{\mathsf{T}} B^{\mathsf{T}}$ sería $(n \times m)(p \times n)$, que en general ni siquiera está definido: la inversión del orden no es una elección, es una obligación.

Ahora las entradas. Para $1 \le i \le p$ y $1 \le j \le m$, usando la definición 7.17 y luego la 7.22:

$$\big((AB)^{\mathsf{T}}\big)_{ij} = (AB)_{ji} = \sum_{k=1}^{n} a_{jk} b_{ki}.$$

Por el otro lado, y usando de nuevo la definición de traspuesta en cada factor:

$$\big(B^{\mathsf{T}} A^{\mathsf{T}}\big)_{ij} = \sum_{k=1}^{n} (B^{\mathsf{T}})_{ik} (A^{\mathsf{T}})_{kj} = \sum_{k=1}^{n} b_{ki}\, a_{jk}.$$

Las dos sumas tienen los mismos n sumandos, solo que escritos con los factores en distinto orden, y el producto en K es conmutativo. Luego coinciden para todos los i, j, y las matrices son iguales. □

Ejemplo 7.27 (Comprobación numérica de $(AB)^{\mathsf{T}} = B^{\mathsf{T}}A^{\mathsf{T}}$)**.** Con las matrices del ejemplo 7.23 teníamos $AB = \left(\begin{smallmatrix} -6 & 13 \\ 19 & 3 \end{smallmatrix}\right)$, luego $(AB)^{\mathsf{T}} = \left(\begin{smallmatrix} -6 & 19 \\ 13 & 3 \end{smallmatrix}\right)$. Por el otro camino,

$$B^{\mathsf{T}} = \begin{pmatrix} 1 & 5 & -1 \\ 2 & 0 & 3 \end{pmatrix}, \qquad A^{\mathsf{T}} = \begin{pmatrix} 2 & 0 \\ -1 & 4 \\ 3 & 1 \end{pmatrix},$$

y multiplicando $(2 \times 3)(3 \times 2)$:

$$(B^{\mathsf{T}}A^{\mathsf{T}})_{11} = 1 \cdot 2 + 5 \cdot (-1) + (-1) \cdot 3 = 2 - 5 - 3 = -6,$$
$$(B^{\mathsf{T}}A^{\mathsf{T}})_{12} = 1 \cdot 0 + 5 \cdot 4 + (-1) \cdot 1 = 0 + 20 - 1 = 19,$$
$$(B^{\mathsf{T}}A^{\mathsf{T}})_{21} = 2 \cdot 2 + 0 \cdot (-1) + 3 \cdot 3 = 4 + 0 + 9 = 13,$$
$$(B^{\mathsf{T}}A^{\mathsf{T}})_{22} = 2 \cdot 0 + 0 \cdot 4 + 3 \cdot 1 = 3,$$

es decir $B^{\mathsf{T}}A^{\mathsf{T}} = \left(\begin{smallmatrix} -6 & 19 \\ 13 & 3 \end{smallmatrix}\right)$, que es exactamente $(AB)^{\mathsf{T}}$.

Cuidado

[Tres costumbres de los números que aquí no valen] **(1) El producto no es conmutativo.** En general $AB \neq BA$, y a menudo uno de los dos ni siquiera está definido. El ejemplo 7.24 lo exhibe con dos matrices 2×2 donde ambos productos existen y son distintos. Consecuencia inmediata: la identidad $(A+B)^2 = A^2 + 2AB + B^2$ es *falsa*; lo correcto es $(A+B)^2 = A^2 + AB + BA + B^2$, y solo cuando A y B conmutan se pueden juntar los dos términos centrales.

(2) Un producto nulo no obliga a que un factor sea nulo. Con

$$P = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}, \qquad Q = \begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix}$$

se tiene $(PQ)_{11} = 1 \cdot 1 + 1 \cdot (-1) = 0$, $(PQ)_{12} = 1 \cdot (-1) + 1 \cdot 1 = 0$, y lo mismo en la segunda fila, de modo que $PQ = 0$ aunque ni P ni Q son la matriz nula. En el lenguaje del capítulo 24, las matrices tienen *divisores de cero*.

(3) No se puede cancelar. De $AB = AC$ con $A \neq 0$ no se sigue $B = C$. Con la misma P de antes, tomemos

$$B = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}, \qquad C = \begin{pmatrix} 2 & 3 \\ 2 & 3 \end{pmatrix}.$$

Entonces PB tiene primera fila $(1 \cdot 1 + 1 \cdot 3,\ 1 \cdot 2 + 1 \cdot 4) = (4, 6)$ y segunda fila idéntica, mientras que PC tiene primera fila $(1 \cdot 2 + 1 \cdot 2,\ 1 \cdot 3 + 1 \cdot 3) = (4, 6)$ y segunda fila idéntica. Es decir $PB = PC = \left(\begin{smallmatrix} 4 & 6 \\ 4 & 6 \end{smallmatrix}\right)$ con $B \neq C$. La cancelación solo es lícita cuando A es invertible, y de eso trata el teorema 7.4.

7.2.6. Potencias de una matriz cuadrada

Como el producto es asociativo, multiplicar una matriz cuadrada por sí misma repetidamente no requiere paréntesis y define una potencia sin ambigüedad.

Definición 7.28. Sea $A \in M_n(K)$. Se define $A^0 = I_n$ y, para $k \geq 1$, $A^k = A^{k-1}A$. De la asociatividad se sigue $A^r A^s = A^{r+s}$ y $(A^r)^s = A^{rs}$ para todos $r, s \geq 0$. Una matriz A se llama *nilpotente* si $A^k = 0$ para algún $k \geq 1$.

Obsérvese que solo las matrices *cuadradas* tienen potencias: si A es $m \times n$ con $m \neq n$, el producto AA ni siquiera está definido. Y obsérvese también que $(AB)^k$ no es $A^k B^k$ salvo que A y B conmuten, por la misma razón de siempre.

Ejemplo 7.29 (Una potencia que se adivina y una matriz que se anula)**. (a) La cizalla.** Sea $U = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$. Entonces

$$U^2 = \begin{pmatrix} 1\cdot 1 + 1\cdot 0 & 1\cdot 1 + 1\cdot 1 \\ 0\cdot 1 + 1\cdot 0 & 0\cdot 1 + 1\cdot 1 \end{pmatrix} = \begin{pmatrix} 1 & 2 \\ 0 & 1 \end{pmatrix}, \qquad U^3 = U^2 U = \begin{pmatrix} 1 & 3 \\ 0 & 1 \end{pmatrix},$$

y en general $U^k = \begin{pmatrix} 1 & k \\ 0 & 1 \end{pmatrix}$, como se comprueba por inducción: si vale para k, entonces $U^{k+1} = U^k U$ tiene entrada $(1,2)$ igual a $1 \cdot 1 + k \cdot 1 = k + 1$ y las demás sin cambio. Para $k = 5$ sale $\left(\begin{smallmatrix} 1 & 5 \\ 0 & 1 \end{smallmatrix}\right)$.

(b) Una nilpotente. Sea

$$N = \begin{pmatrix} 0 & 1 & 2 \\ 0 & 0 & 3 \\ 0 & 0 & 0 \end{pmatrix}.$$

Calculamos N^2 entrada por entrada. La primera fila de N es $(0,1,2)$; contra las tres columnas de N, que son $(0,0,0)$, $(1,0,0)$ y $(2,3,0)$, da $0\cdot 0 + 1\cdot 0 + 2\cdot 0 = 0$, $0\cdot 1 + 1\cdot 0 + 2\cdot 0 = 0$ y $0\cdot 2 + 1\cdot 3 + 2\cdot 0 = 3$. La segunda fila de N es $(0,0,3)$ y produce $0,0,0$ contra las tres columnas; la tercera es nula. Luego

$$N^2 = \begin{pmatrix} 0 & 0 & 3 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \qquad N^3 = N^2 N = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} = 0.$$

La matriz no es nula, ni lo es su cuadrado, y sin embargo su cubo sí: nada parecido ocurre con los números, donde $x^3 = 0$ obliga a $x = 0$. El patrón es general y se entenderá del todo en la subsección sobre la forma de Jordan: cada producto por N empuja la información una diagonal hacia arriba, hasta que se cae de la matriz.

7.2.7. Un sistema de ecuaciones escrito como $A\mathbf{x} = \mathbf{b}$

Ya tenemos todo lo necesario para justificar la notación con la que se abrió el capítulo. La clave es un caso particular del producto: cuando el segundo factor es una matriz columna. Si A es $m \times n$ y $\mathbf{x}$ es $n \times 1$, los tamaños casan y $A\mathbf{x}$ es una matriz $m \times 1$, es decir, un vector de K^m, cuya componente i-ésima es

$$(A\mathbf{x})_i = \sum_{k=1}^{n} a_{ik} x_k = a_{i1}x_1 + a_{i2}x_2 + \cdots + a_{in}x_n,$$

que es exactamente el miembro izquierdo de la ecuación i-ésima de un sistema lineal. Esa coincidencia no es casual: es la razón de ser de la notación.

Regla: pasar de un sistema de ecuaciones a la forma matricial, y volver

El paso es puramente mecánico en los dos sentidos, y hacerlo despacio la primera vez evita errores para siempre.

1. **Ordena el sistema antes de tocar nada.** Escribe las m ecuaciones con las incógnitas siempre en el mismo orden y siempre en el mismo lado, y los términos independientes en el otro. Si una incógnita falta en una ecuación, su coeficiente es 0 y hay que escribirlo; si aparece sin número delante, su coeficiente es 1.
2. **La matriz de coeficientes.** $A = (a_{ij})$ es la matriz $m \times n$ cuya entrada (i, j) es el coeficiente de la incógnita j-ésima en la ecuación i-ésima. Cada *fila* es una ecuación; cada *columna* es una incógnita.
3. **El vector de incógnitas y el de términos independientes.** $\mathbf{x}$ es la matriz columna $n \times 1$ con las incógnitas en el orden fijado, y $\mathbf{b}$ es la matriz columna $m \times 1$ con los términos independientes en el orden de las ecuaciones. El sistema completo es entonces la única igualdad matricial $A\mathbf{x} = \mathbf{b}$.
4. **La matriz ampliada.** Para trabajar solo con números se escribe $[A \mid \mathbf{b}]$, la matriz $m \times (n+1)$ que se obtiene pegando la columna $\mathbf{b}$ a la derecha de A, con una barra vertical que recuerda dónde acababan los coeficientes. La barra no es un objeto matemático, es un recordatorio tipográfico.
5. **Para volver, lee cada fila como una ecuación.** Dada A y $\mathbf{b}$, la fila i de A da los coeficientes y b_i el término independiente: la ecuación es $a_{i1}x_1 + \cdots + a_{in}x_n = b_i$. Los nombres de las incógnitas los eliges tú; lo que la matriz guarda es el orden, no los nombres.
6. **Control de tamaños.** A es $m \times n$ con m el número de ecuaciones y n el de incógnitas; $\mathbf{x}$ es $n \times 1$; $\mathbf{b}$ es $m \times 1$; la ampliada es $m \times (n+1)$. Si alguno de estos números no encaja, hay una ecuación o una incógnita mal contada.

Ejemplo 7.30 (Un sistema 3×3 recorrido en los dos sentidos)**. (a) Del sistema a la matriz.** Consideremos

$$\begin{cases} 2x + \ y - \ z = 3 \\ \ x - 3y + 2z = -3 \\ 3x + 2y + \ z = 8. \end{cases}$$

Las incógnitas están en el orden x, y, z en las tres ecuaciones y todos los coeficientes son visibles. Por tanto

$$A = \begin{pmatrix} 2 & 1 & -1 \\ 1 & -3 & 2 \\ 3 & 2 & 1 \end{pmatrix}, \qquad \mathbf{x} = \begin{pmatrix} x \\ y \\ z \end{pmatrix}, \qquad \mathbf{b} = \begin{pmatrix} 3 \\ -3 \\ 8 \end{pmatrix},$$

y el sistema entero es la igualdad $A\mathbf{x} = \mathbf{b}$ entre dos matrices 3×1. La matriz ampliada es

$$[A \mid \mathbf{b}] = \left[\begin{array}{ccc|c} 2 & 1 & -1 & 3 \\ 1 & -3 & 2 & -3 \\ 3 & 2 & 1 & 8 \end{array}\right].$$

(b) De la matriz al sistema. Si alguien nos entrega esa misma A y ese mismo $\mathbf{b}$ y nos pide el sistema, leemos fila a fila: la primera fila $(2, 1, -1)$ con término independiente 3 da $2x_1 + x_2 - x_3 = 3$; la segunda, $x_1 - 3x_2 + 2x_3 = -3$; la tercera, $3x_1 + 2x_2 + x_3 = 8$. Es el sistema de partida con las incógnitas rebautizadas.

(c) Comprobar una solución es multiplicar. La terna $x = 1$, $y = 2$, $z = 1$ es solución, y verificarlo es hacer un producto matricial. La primera componente de $A\mathbf{x}$ es $2 \cdot 1 + 1 \cdot 2 + (-1) \cdot 1 = 2 + 2 - 1 = 3$; la segunda, $1 \cdot 1 + (-3) \cdot 2 + 2 \cdot 1 = 1 - 6 + 2 = -3$; la tercera, $3 \cdot 1 + 2 \cdot 2 + 1 \cdot 1 = 3 + 4 + 1 = 8$. Luego $A\mathbf{x} = (3, -3, 8)^{\mathsf{T}} = \mathbf{b}$, como se quería.

Observación 7.31 ($A\mathbf{x}$ es una combinación de las columnas de A)*.* Hay una segunda manera de leer el producto $A\mathbf{x}$, y es la más fértil de las dos. Llamemos $\mathbf{c}_1, \ldots, \mathbf{c}_n$ a las columnas de A, cada una un vector de K^m. Entonces

$$A\mathbf{x} = x_1\mathbf{c}_1 + x_2\mathbf{c}_2 + \cdots + x_n\mathbf{c}_n,$$

porque la componente i-ésima del miembro derecho es $x_1a_{i1} + \cdots + x_na_{in}$, que es justamente $(A\mathbf{x})_i$. Dicho en palabras: *multiplicar A por un vector es combinar las columnas de A usando las componentes del vector como coeficientes.* Con el ejemplo anterior:

$$1\begin{pmatrix}2\\1\\3\end{pmatrix} + 2\begin{pmatrix}1\\-3\\2\end{pmatrix} + 1\begin{pmatrix}-1\\2\\1\end{pmatrix} = \begin{pmatrix}2+2-1\\1-6+2\\3+4+1\end{pmatrix} = \begin{pmatrix}3\\-3\\8\end{pmatrix} = \mathbf{b}.$$

La consecuencia conceptual es grande: **el sistema $A\mathbf{x} = \mathbf{b}$ tiene solución si y solo si $\mathbf{b}$ se puede escribir como combinación lineal de las columnas de A**, y resolverlo es exactamente encontrar los coeficientes de esa combinación. Todo el aparato de espacios vectoriales, envolvente lineal, rango y dimensión que viene a continuación no es más que la maquinaria para responder a esa única pregunta con precisión.

7.2.8. Operaciones de fila, forma escalonada, pivotes y rango

La sección primera del capítulo usó estos términos apoyándose en la intuición. Aquí los fijamos con precisión, porque son los cimientos de todos los cálculos posteriores.

Definición 7.32. Las *operaciones elementales de fila* sobre una matriz son tres: *(I)* intercambiar dos filas, $F_i \leftrightarrow F_j$; *(II)* multiplicar una fila por un escalar $\lambda \neq 0$, $F_i \to \lambda F_i$; *(III)* sumar a una fila un múltiplo de otra, $F_i \to F_i + \lambda F_j$ con $i \neq j$. Dos matrices son *equivalentes por filas* si se pasa de una a otra mediante una sucesión finita de estas operaciones.

Las tres son reversibles (la inversa de un intercambio es el mismo intercambio; la de $F_i \to \lambda F_i$ es $F_i \to \lambda^{-1}F_i$; la de $F_i \to F_i + \lambda F_j$ es $F_i \to F_i - \lambda F_j$), y esa reversibilidad es la

razón exacta por la que no alteran el conjunto de soluciones de un sistema: cada ecuación nueva es consecuencia de las viejas y al revés.

Definición 7.33. Una matriz está en *forma escalonada por filas* si: las filas nulas, si las hay, están todas al final; y en cada fila no nula, la primera entrada no nula empezando por la izquierda —llamada *pivote* de esa fila— está estrictamente más a la derecha que el pivote de la fila anterior. Está en *forma escalonada reducida* si además todos los pivotes valen 1 y cada pivote es la única entrada no nula de su columna. Las columnas que contienen un pivote se llaman *columnas pivote*.

Definición 7.34. El *rango* de una matriz A, denotado $\operatorname{rango}(A)$, es el número de pivotes de cualquier forma escalonada de A, es decir, el número de filas no nulas que quedan al escalonar.

Que este número no dependa del camino de escalonamiento elegido no es evidente y se demuestra más adelante, cuando el rango se reinterprete como la dimensión del espacio generado por las filas (y también por las columnas). De momento basta con saber calcularlo.

Regla: escalonar una matriz por el algoritmo de Gauss

El algoritmo trabaja columna por columna, de izquierda a derecha, y de arriba abajo.

1. **Sitúate en la primera columna que no sea entera de ceros.** Esa columna contendrá el primer pivote.
2. **Consigue un pivote no nulo en la fila de trabajo.** Si la entrada de la fila actual en esa columna es 0, intercámbiala con una fila de más abajo que tenga ahí algo distinto de cero (operación I). Si todas son cero, esa columna no lleva pivote: pasa a la siguiente.
3. **Anula todo lo que hay debajo del pivote.** Para cada fila F_i por debajo, aplica $F_i \to F_i - \frac{a_{ij}}{p} F_{\text{piv}}$, donde p es el pivote (operación III). Nunca operes con columnas: eso mezclaría las incógnitas.
4. **Baja una fila y una columna, y repite.** El bloque que queda por debajo y a la derecha se trata igual, ignorando lo ya hecho. Cuando se acaben las filas o las columnas, la matriz está escalonada.
5. **Si quieres la forma reducida, sigue hacia arriba.** Divide cada fila por su pivote (operación II) para dejarlo en 1, y con cada pivote anula también lo que tiene *encima* (operación III), empezando por el pivote de más a la derecha.
6. **Cuenta los pivotes: ese número es el rango.** Y si estabas escalonando una matriz ampliada, compara el rango de la parte de coeficientes con el de la ampliada completa: eso es Rouché–Frobenius.

Ejemplo 7.35 (Un escalonamiento completo, con todas las matrices intermedias)**.** Escalone-

mos hasta la forma reducida la matriz ampliada del sistema

$$\begin{cases} x + 2y - z = 2 \\ 2x + 3y + z = 11 \\ 3x + 5y + 2z = 19. \end{cases}$$

La matriz ampliada de partida es

$$M_0 = \left[\begin{array}{ccc|c} 1 & 2 & -1 & 2 \\ 2 & 3 & 1 & 11 \\ 3 & 5 & 2 & 19 \end{array}\right].$$

Paso 1. El pivote de la primera columna ya es 1 y está en su sitio. Anulamos lo de debajo con $F_2 \to F_2 - 2F_1$ y $F_3 \to F_3 - 3F_1$. Fila a fila: $F_2 - 2F_1 = (2-2,\ 3-4,\ 1+2 \mid 11-4) = (0, -1, 3 \mid 7)$ y $F_3 - 3F_1 = (3-3,\ 5-6,\ 2+3 \mid 19-6) = (0, -1, 5 \mid 13)$. Queda

$$M_1 = \left[\begin{array}{ccc|c} 1 & 2 & -1 & 2 \\ 0 & -1 & 3 & 7 \\ 0 & -1 & 5 & 13 \end{array}\right].$$

Paso 2. Bajamos a la segunda fila y segunda columna: el pivote es -1, que no es cero, así que no hace falta intercambiar. Anulamos lo de debajo con $F_3 \to F_3 - F_2$, que da $(0,\ -1+1,\ 5-3 \mid 13-7) = (0, 0, 2 \mid 6)$:

$$M_2 = \left[\begin{array}{ccc|c} 1 & 2 & -1 & 2 \\ 0 & -1 & 3 & 7 \\ 0 & 0 & 2 & 6 \end{array}\right].$$

Esta matriz ya está en forma escalonada: los pivotes son 1, -1 y 2, situados en las columnas 1, 2 y 3, y cada uno está más a la derecha que el anterior. Hay tres pivotes, luego $\text{rango}(A) = 3$ y también $\text{rango}([A \mid \mathbf{b}]) = 3$: sistema compatible determinado, con $n = 3$ incógnitas y ningún parámetro libre.

Paso 3. Normalizamos los pivotes con $F_2 \to -F_2$ y $F_3 \to \frac{1}{2}F_3$:

$$M_3 = \left[\begin{array}{ccc|c} 1 & 2 & -1 & 2 \\ 0 & 1 & -3 & -7 \\ 0 & 0 & 1 & 3 \end{array}\right].$$

Paso 4. Limpiamos hacia arriba empezando por el pivote de la tercera columna, con $F_2 \to F_2 + 3F_3$ (que da $(0, 1, -3+3 \mid -7+9) = (0, 1, 0 \mid 2)$) y $F_1 \to F_1 + F_3$ (que da $(1, 2, -1+1 \mid 2+3) = (1, 2, 0 \mid 5)$):

$$M_4 = \left[\begin{array}{ccc|c} 1 & 2 & 0 & 5 \\ 0 & 1 & 0 & 2 \\ 0 & 0 & 1 & 3 \end{array}\right].$$

Paso 5. Queda el pivote de la segunda columna, con $F_1 \to F_1 - 2F_2$, que da $(1,\, 2-2,\, 0 \mid 5-4) = (1,0,0 \mid 1)$:

$$M_5 = \left[\begin{array}{ccc|c} 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 2 \\ 0 & 0 & 1 & 3 \end{array}\right].$$

Esta es la forma escalonada reducida, y la solución se lee sin hacer nada más: $x = 1$, $y = 2$, $z = 3$.

Control. Sustituimos en las ecuaciones *originales*: $1 + 4 - 3 = 2$ ✓; $2 + 6 + 3 = 11$ ✓; $3 + 10 + 6 = 19$ ✓.

Ejemplo 7.36 (Rango de una matriz con filas dependientes)**.** Sea

$$B = \begin{pmatrix} 1 & 2 & -1 & 3 \\ 2 & 4 & 1 & 0 \\ 3 & 6 & 0 & 3 \end{pmatrix},$$

de tamaño 3×4. Con $F_2 \to F_2 - 2F_1$ obtenemos $(0,0,3,-6)$ y con $F_3 \to F_3 - 3F_1$ obtenemos $(0,0,3,-6)$ también:

$$\begin{pmatrix} 1 & 2 & -1 & 3 \\ 0 & 0 & 3 & -6 \\ 0 & 0 & 3 & -6 \end{pmatrix} \xrightarrow{F_3 \to F_3 - F_2} \begin{pmatrix} 1 & 2 & -1 & 3 \\ 0 & 0 & 3 & -6 \\ 0 & 0 & 0 & 0 \end{pmatrix}.$$

Ya está escalonada, con pivotes 1 (columna 1) y 3 (columna 3): hay dos pivotes, luego $\operatorname{rango}(B) = 2$. Nótese que la segunda columna no lleva pivote pese a no ser nula; lo que decide es la posición del primer elemento no nulo de cada fila, no el contenido de las columnas. Normalizando con $F_2 \to \frac{1}{3}F_2$ y limpiando hacia arriba con $F_1 \to F_1 + F_2$ se llega a la forma reducida

$$\begin{pmatrix} 1 & 2 & 0 & 1 \\ 0 & 0 & 1 & -2 \\ 0 & 0 & 0 & 0 \end{pmatrix},$$

donde los dos pivotes valen 1 y son los únicos elementos no nulos de sus columnas. La tercera fila se anuló porque era redundante: en efecto, $F_3 = F_1 + F_2$ en la matriz original, ya que $(1,2,-1,3) + (2,4,1,0) = (3,6,0,3)$.

Cuidado

[Errores típicos al escalonar] El primero es operar con columnas. Las tres operaciones de la definición 7.32 son de *fila*; intercambiar dos columnas de una matriz ampliada equivale a renombrar incógnitas y, si se hace sin avisar, la solución sale permutada. El segundo es multiplicar una fila por cero, que no es una operación elemental (se exige $\lambda \neq 0$) y que destruye información irreversiblemente. El tercero es olvidar que la columna de términos independientes participa en todas las operaciones: si se aplica $F_2 \to F_2 - 2F_1$ hay que aplicarla también a la última columna. Y el cuarto es confundir "número de filas no

nulas de la matriz de partida" con el rango: el rango se lee *después* de escalonar, nunca antes.

Pausa para pensar

Merece la pena mirar hacia atrás y ver cuánto se ha conseguido con tan poco. Se ha definido una tabla, dos operaciones obvias y una tercera menos obvia, y con eso ya está escrito todo el álgebra lineal computacional: un sistema es $A\mathbf{x} = \mathbf{b}$, resolverlo es escalonar $[A \mid \mathbf{b}]$, y la pregunta de si tiene solución es la pregunta de si $\mathbf{b}$ está en el conjunto de combinaciones de las columnas de A. Esa última frase contiene ya, en germen, los conceptos de subespacio generado, independencia lineal, base y dimensión, que la sección siguiente desarrolla con nombre propio. Y contiene también el motivo de que el producto de matrices merezca su extraña regla: porque es la operación que corresponde a encadenar transformaciones, y en cuanto se mira una matriz como una transformación en vez de como una tabla, todo el capítulo cambia de color.

7.3. Espacios vectoriales

Idea clave

Intuición primero. Un espacio vectorial es cualquier conjunto donde tenga sentido *sumar* elementos y *escalarlos*, con las reglas de siempre. La lista de axiomas parece un trámite burocrático, pero lo que compra es enorme: todo lo que se demuestre usando solo esos axiomas vale automáticamente para las flechas del plano, para las matrices, para los polinomios, para las funciones continuas y para las soluciones de una ecuación diferencial. Una demostración, muchos teoremas.
Los conceptos de esta sección son todos variaciones sobre una sola pregunta: *¿cuánta información hace falta para describir un elemento?* Generar es poder alcanzarlo todo; ser independiente es no tener repeticiones; una base es lo justo, sin sobras ni faltas; y la dimensión es cuántos elementos tiene una base, que resulta ser un número que no depende de la base elegida. Esa unicidad es lo que convierte la dimensión en una propiedad del espacio y no de la descripción.

Definición 7.37. Un *espacio vectorial* sobre un cuerpo K es un conjunto V con suma y producto por escalares tales que $(V, +)$ es grupo abeliano y valen $c(u + v) = cu + cv$, $(c + d)v = cv + dv$, $c(dv) = (cd)v$, $1v = v$. De los axiomas se deducen $0v = \mathbf{0}$, $(-1)v = -v$, etc.

Ejemplo 7.38. $\mathbb{R}^n$ y $\mathbb{C}^n$; matrices $M_{m\times n}$; polinomios $\mathbb{R}[x]$ y los de grado $\leq n$ (denotados $\mathbb{R}_n[x]$); funciones continuas $C[a, b]$; sucesiones; soluciones de una EDO lineal homogénea; y —clave para el capítulo 26— cualquier cuerpo L que contenga a K es un K-espacio vectorial.

Definición 7.39. $W \subseteq V$ es *subespacio* si es no vacío y cerrado bajo las operaciones. La *envolvente lineal* $\text{gen}(S)$ es el conjunto de combinaciones lineales de S, el menor subespacio que contiene a S. La *suma* de subespacios es $U + W = \{u + w\}$, y es *directa* $(U \oplus W)$

cuando $U \cap W = \{\mathbf{0}\}$: entonces cada vector se descompone de forma única. Vale la fórmula de Grassmann:

$$\dim(U+W) = \dim U + \dim W - \dim(U \cap W).$$

7.3.1. Independencia, bases, dimensión

Definición 7.40. $v_1, \dots, v_k$ son *linealmente independientes* si $c_1 v_1 + \cdots + c_k v_k = \mathbf{0}$ obliga $c_1 = \cdots = c_k = 0$; equivalentemente, ninguno es combinación de los demás. Una *base* es un conjunto independiente que genera V.

Teorema 7.41 (Steinitz / dimensión)**.** *Si V tiene una base de n elementos, todo conjunto independiente tiene a lo más n vectores, todo generador tiene al menos n, y toda base tiene exactamente n: la* dimensión *está bien definida. Además: todo conjunto independiente se extiende a una base, y de todo generador se extrae una base.*

En la práctica: para decidir independencia de vectores en $\mathbb{R}^n$ se colocan como columnas y se escalona; los pivotes marcan una base de la envolvente.

7.3.2. Coordenadas y cambio de base

Fijada una base ordenada $\mathcal{B} = \{v_1, \dots, v_n\}$, cada $v \in V$ se escribe de forma única $v = c_1 v_1 + \cdots + c_n v_n$; el vector $[v]_{\mathcal{B}} = (c_1, \dots, c_n)$ son sus *coordenadas*. Si $\mathcal{B}'$ es otra base y P es la matriz cuyas columnas son $[v'_j]_{\mathcal{B}}$ (las coordenadas de la base nueva en la vieja), entonces

$$[v]_{\mathcal{B}} = P\,[v]_{\mathcal{B}'}, \qquad [v]_{\mathcal{B}'} = P^{-1}[v]_{\mathcal{B}}.$$

Ejemplo 7.42. En $\mathbb{R}^2$, sean $\mathcal{B}$ la canónica y $\mathcal{B}' = \{(1,1),(1,-1)\}$. Entonces $P = \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}$. El vector $v = (3,1)$ tiene $[v]_{\mathcal{B}'} = P^{-1}(3,1)^{\mathsf{T}} = (2,1)^{\mathsf{T}}$: en efecto $2(1,1) + 1(1,-1) = (3,1)$.

Regla: leer los cuatro subespacios fundamentales en la forma escalonada reducida

Toda matriz A de tamaño $m \times n$ lleva asociados cuatro subespacios: el núcleo $\operatorname{Nul} A \subseteq \mathbb{R}^n$, el espacio columna $\operatorname{Col} A \subseteq \mathbb{R}^m$, el espacio fila $\operatorname{Fil} A \subseteq \mathbb{R}^n$ y el núcleo por la izquierda $\operatorname{Nul} A^T \subseteq \mathbb{R}^m$. Una sola eliminación de Gauss–Jordan produce bases de los cuatro. Este es el procedimiento.

1. **Reduce A a su forma escalonada reducida R, sin trasponer ni permutar columnas.** Anota las operaciones elementales de fila que usas; te harán falta al final para el cuarto subespacio. La forma reducida tiene un 1 como primera entrada no nula de cada fila no nula, y ceros encima y debajo.
2. **Localiza las columnas pivote.** Son las que contienen esos unos principales. Su número es el rango $r = \operatorname{rango} A$, que es a la vez la dimensión de $\operatorname{Col} A$ y la de $\operatorname{Fil} A$.

3. **Base de $\operatorname{Col} A$: las columnas pivote *de* A, no de R.** Este es el error más frecuente de todo el tema. La eliminación cambia el espacio columna, pero *no* cambia qué columnas son combinación de cuáles; por eso los índices se leen en R y los vectores se copian de A.

4. **Base de $\operatorname{Fil} A$: las filas no nulas *de* R.** Aquí ocurre lo contrario: la eliminación no cambia el espacio fila (cada operación elemental sustituye una fila por una combinación de filas), así que las filas de R sirven y además son la base más limpia posible.

5. **Base de $\operatorname{Nul} A$: una solución básica por cada variable libre.** Las variables libres son las de las columnas sin pivote; hay $n - r$. Para cada una, ponla igual a 1, las demás libres a 0, y despeja las pivote en R. Los vectores así obtenidos son automáticamente independientes, porque cada uno tiene un 1 donde los otros tienen 0.

6. **Base de $\operatorname{Nul} A^T$: las combinaciones de filas que se anularon.** Si al reducir apareció una fila de ceros, el registro de operaciones te dice qué combinación de las filas originales la produjo, y sus coeficientes forman un vector del núcleo por la izquierda. Alternativamente, reduce A^T y aplica el paso 5.

7. **Verifica con las dimensiones y con la ortogonalidad.** Deben cumplirse

$$\dim \operatorname{Col} A = \dim \operatorname{Fil} A = r, \quad \dim \operatorname{Nul} A = n - r, \quad \dim \operatorname{Nul} A^T = m - r,$$

y además $\operatorname{Fil} A \perp \operatorname{Nul} A$ en $\mathbb{R}^n$ y $\operatorname{Col} A \perp \operatorname{Nul} A^T$ en $\mathbb{R}^m$. Multiplicar un par de vectores de cada lado y obtener 0 es una comprobación de diez segundos.

Ejemplo 7.43 (Una matriz 3×5 desmontada en sus cuatro subespacios)**.** Tomemos

$$A = \begin{pmatrix} 1 & 2 & 2 & -1 & 3 \\ 2 & 4 & 5 & 0 & 7 \\ 3 & 6 & 7 & -1 & 10 \end{pmatrix}, \qquad m = 3,\ n = 5.$$

(a) Eliminación. Con $F_2 \to F_2 - 2F_1$ y $F_3 \to F_3 - 3F_1$:

$$\begin{pmatrix} 1 & 2 & 2 & -1 & 3 \\ 0 & 0 & 1 & 2 & 1 \\ 0 & 0 & 1 & 2 & 1 \end{pmatrix},$$

y con $F_3 \to F_3 - F_2$ y $F_1 \to F_1 - 2F_2$ se llega a

$$R = \begin{pmatrix} 1 & 2 & 0 & -5 & 1 \\ 0 & 0 & 1 & 2 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}.$$

Los pivotes están en las columnas 1 y 3, luego $r = \operatorname{rango} A = 2$.

(b) Espacio columna. Copiamos de A las columnas 1 y 3:

$$\operatorname{Col} A = \operatorname{gen}\left\{\begin{pmatrix}1\\2\\3\end{pmatrix},\ \begin{pmatrix}2\\5\\7\end{pmatrix}\right\}, \qquad \dim = 2.$$

Es un plano en $\mathbb{R}^3$. Compruébese que la columna 2 de A es 2 veces la primera, que la columna 4 es $-5c_1 + 2c_3 = (-5,-10,-15) + (4,10,14) = (-1,0,-1)$ y que la columna 5 es $c_1 + c_3 = (3,7,10)$: exactamente los coeficientes que se leen en R.

(c) Espacio fila. Las filas no nulas de R:

$$\operatorname{Fil} A = \operatorname{gen}\{(1,2,0,-5,1),\ (0,0,1,2,1)\}, \qquad \dim = 2.$$

(d) Núcleo. Las variables libres son x_2, x_4, x_5; hay $5-2=3$. Las ecuaciones de R son $x_1 + 2x_2 - 5x_4 + x_5 = 0$ y $x_3 + 2x_4 + x_5 = 0$.

- $x_2 = 1$, $x_4 = x_5 = 0$: $x_1 = -2$, $x_3 = 0$, y sale $\mathbf{v}_1 = (-2,1,0,0,0)$.
- $x_4 = 1$, $x_2 = x_5 = 0$: $x_1 = 5$, $x_3 = -2$, y sale $\mathbf{v}_2 = (5,0,-2,1,0)$.
- $x_5 = 1$, $x_2 = x_4 = 0$: $x_1 = -1$, $x_3 = -1$, y sale $\mathbf{v}_3 = (-1,0,-1,0,1)$.

Luego $\dim \operatorname{Nul} A = 3$, y en efecto $2+3=5=n$: el teorema del rango se cumple.

(e) Núcleo por la izquierda. La fila nula apareció al hacer $F_3 - 3F_1 - (F_2 - 2F_1) = F_3 - F_2 - F_1$. Por tanto

$$-1\cdot F_1 - 1\cdot F_2 + 1\cdot F_3 = \mathbf{0} \quad\Longrightarrow\quad \operatorname{Nul} A^T = \operatorname{gen}\{(-1,-1,1)\},$$

de dimensión $3-2=1$. Verificación inmediata: $F_1 + F_2 = (1,2,2,-1,3) + (2,4,5,0,7) = (3,6,7,-1,10) = F_3$. La tercera fila era redundante desde el principio.

(f) Las dos ortogonalidades. Probemos una de cada:

$$(1,2,0,-5,1)\cdot(5,0,-2,1,0) = 5+0+0-5+0 = 0, \qquad (1,2,3)\cdot(-1,-1,1) = -1-2+3 = 0.$$

Ambas dan cero, como debía ser. Y geométricamente el cuadro es completo: en $\mathbb{R}^5$, un plano ($\operatorname{Fil} A$) y su complemento ortogonal de dimensión 3 ($\operatorname{Nul} A$); en $\mathbb{R}^3$, un plano ($\operatorname{Col} A$) y su recta normal ($\operatorname{Nul} A^T$).

(g) Qué significa para el sistema $A\mathbf{x} = \mathbf{b}$. Como $\operatorname{Col} A$ es un plano de $\mathbb{R}^3$ y no todo $\mathbb{R}^3$, el sistema *no* tiene solución para toda $\mathbf{b}$: la tiene si y solo si $\mathbf{b} \perp (-1,-1,1)$, es decir si $-b_1 - b_2 + b_3 = 0$. Con $\mathbf{b} = (1,1,2)$ hay solución; con $\mathbf{b} = (1,1,3)$ no la hay. Y cuando la hay, hay infinitas: la solución particular más el subespacio de dimensión 3 del apartado (d).

Cuidado

Repetimos el error del paso 3 porque cuesta muchos exámenes: la base de $\operatorname{Col} A$ se toma de A, la de $\operatorname{Fil} A$ se toma de R. En este ejemplo, las columnas pivote de R son $(1,0,0)$ y $(0,1,0)$, que generan un plano *distinto* del que generan $(1,2,3)$ y $(2,5,7)$. La eliminación por filas preserva el espacio fila y el núcleo, pero destroza el espacio columna;

conserva únicamente las *relaciones de dependencia* entre columnas, que es justo lo que necesitamos para elegir cuáles copiar.

Idea clave

Los cuatro subespacios no son cuatro objetos sueltos: son dos pares de complementos ortogonales, y la aplicación $\mathbf{x} \mapsto A\mathbf{x}$ es una biyección entre $\operatorname{Fil} A$ y $\operatorname{Col} A$. Esa biyección, invertida y extendida por cero sobre $\operatorname{Nul} A^T$, es exactamente la pseudoinversa de Moore–Penrose, y sus factores de escala son los valores singulares. Dicho de otro modo: el diagrama de los cuatro subespacios es la SVD antes de conocer los números.

7.4. Transformaciones lineales

Idea clave

Intuición primero. Una transformación lineal es una función que respeta la estructura: transformar una suma es sumar los transformados, y transformar un escalado es escalar el transformado. Geométricamente eso significa que manda rectas a rectas y deja el origen quieto: estira, rota, refleja, proyecta o aplasta, pero nunca curva ni desplaza.
La consecuencia práctica es de las más rentables del libro: como todo vector se escribe con la base, y la transformación respeta sumas y escalados, *basta saber qué hace con los vectores de la base* para saber qué hace con todos los demás. Esa información cabe en una tabla de números, y esa tabla es la matriz. Por eso las matrices no son un objeto aparte que haya que aprender: son transformaciones lineales escritas en coordenadas. Y desde esa lectura, el producto de matrices deja de ser una regla rara de filas por columnas y pasa a ser lo obvio: es la matriz de la composición, es decir, hacer una transformación después de la otra.

Definición 7.44. $T : V \to W$ es *lineal* si $T(u+v) = T(u) + T(v)$ y $T(cv) = cT(v)$. El *núcleo* $\ker T$ y la *imagen* $\operatorname{im} T$ son subespacios. T es inyectiva $\iff \ker T = \{\mathbf{0}\}$. Un *isomorfismo* es una lineal biyectiva; dos espacios de la misma dimensión finita son isomorfos.

Una transformación lineal queda determinada por sus valores en una base. Fijadas bases $\mathcal{B}$ de V y $\mathcal{C}$ de W, la matriz $[T]_{\mathcal{CB}}$ tiene por columnas las imágenes de la base: $[T(v)]_{\mathcal{C}} = [T]_{\mathcal{CB}}[v]_{\mathcal{B}}$, y la composición corresponde al producto de matrices: *esa* es la razón de ser del producto matricial.

Ejemplo 7.45 (Geometría de $\mathbb{R}^2$)**.** Rotación de ángulo θ: $R_\theta = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}$; reflexión respecto del eje x: $\operatorname{diag}(1,-1)$; proyección sobre el eje x: $\operatorname{diag}(1,0)$; cizalla: $\begin{pmatrix} 1 & k \\ 0 & 1 \end{pmatrix}$. Componer rotaciones suma ángulos: $R_\alpha R_\beta = R_{\alpha+\beta}$, que encierra las fórmulas de adición de seno y coseno.

Teorema 7.46 (Rango–nulidad)**.** $\dim \ker T + \dim \operatorname{im} T = \dim V$.

Demostración. Sea $u_1, \dots, u_k$ base de $\ker T$, extendida a base $u_1, \dots, u_k, w_1, \dots, w_r$ de V ($k + r = \dim V$). Afirmamos que $T(w_1), \dots, T(w_r)$ es base de $\operatorname{im} T$. *Generan*: $T(\sum a_i u_i + \sum b_j w_j) = \sum b_j T(w_j)$. *Independientes*: si $\sum b_j T(w_j) = \mathbf{0}$, entonces $\sum b_j w_j \in \ker T$, luego $\sum b_j w_j = \sum a_i u_i$ para ciertos a_i; la independencia de la base completa fuerza todos los coeficientes a cero. □

Idea clave

Qué dice esto en cristiano. Una aplicación lineal reparte las dimensiones del espacio de partida en dos montones: las que sobreviven al otro lado (la imagen) y las que se aplastan hasta el cero (el núcleo). El teorema dice que no se pierde ni se gana nada: la suma de los dos montones es la dimensión de partida. Es una ley de conservación, y sirve para calcular una de las dos cantidades cuando la otra es fácil. Casi siempre lo fácil es el rango, y el núcleo sale por diferencia.

Ejemplo 7.47 (La conservación de dimensiones, en tres casos)**. (a) Una matriz con núcleo.** Sea $T = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 4 & 6 \\ 1 & 1 & 1 \end{pmatrix}$. La segunda fila es el doble de la primera, así que el rango es 2. El teorema predice núcleo de dimensión $3 - 2 = 1$, y en efecto el núcleo está generado por $(1, -2, 1)$: comprobando, $1 - 4 + 3 = 0$ y $1 - 2 + 1 = 0$.

(b) Una proyección. La aplicación $\mathbb{R}^3 \to \mathbb{R}^2$ que olvida la tercera coordenada tiene imagen todo $\mathbb{R}^2$, luego rango 2, y su núcleo es el eje OZ, de dimensión 1: otra vez $2 + 1 = 3$. La proyección aplasta exactamente una dimensión, y eso es lo que mide el núcleo.

(c) Derivar polinomios. En el espacio de polinomios de grado a lo sumo 2, que tiene dimensión 3, la derivada tiene por imagen los polinomios de grado a lo sumo 1, de dimensión 2, y por núcleo las constantes, de dimensión 1. De nuevo $2 + 1 = 3$. Es el mismo teorema aplicado a un espacio que no es de vectores columna.

(d) La consecuencia más usada. Si $\dim V = \dim U$ y $T : V \to U$ es lineal, entonces T es inyectiva si y solo si es sobreyectiva. En dimensión finita no hace falta comprobar las dos cosas: basta con una. Eso explica por qué para matrices cuadradas "núcleo trivial" e "imagen total" son equivalentes, y falla estrepitosamente en dimensión infinita, donde el operador "desplazar a la derecha" es inyectivo sin ser sobreyectivo.

Cuidado

[La dimensión que se conserva es la de partida] En $\dim \ker T + \dim \operatorname{im} T = \dim V$, el V es el espacio de *salida*, no el de llegada. Para una matriz $m \times n$ la fórmula es nulidad + rango $= n$, con n el número de *columnas*. Confundirlo con m es el error habitual. Segundo aviso: el rango de las filas y el de las columnas coinciden siempre, y eso no es evidente; es un teorema, y de él sale que $\operatorname{rango} A = \operatorname{rango} A^{\mathsf{T}}$.

Corolario 7.48. *El rango por filas y el rango por columnas de una matriz coinciden. Para A cuadrada: inyectiva $\iff$ sobreyectiva.*

Si P es el cambio de base en V, la matriz de $T : V \to V$ en la base nueva es $A' = P^{-1}AP$: matrices *semejantes* representan el mismo operador. Las cantidades que no cambian bajo semejanza (determinante, traza, polinomio característico, valores propios) son las *invariantes* del operador, propiedades genuinas de la transformación y no de la base.

Regla: escribir la matriz de una transformación lineal y cambiar de base

Una transformación lineal es un objeto geométrico; su matriz es solo su descripción *en unas coordenadas concretas*. Cambiar de base no cambia la transformación, cambia el idioma en que se la cuenta. Confundir ambas cosas es la fuente principal de errores en este tema, y la receta siguiente lo evita.

1. **Fija las dos bases y respeta el orden.** Si $T : V \to W$, necesitas una base $\mathcal{B} = \{\mathbf{b}_1, \dots, \mathbf{b}_n\}$ de la *salida* y una base $\mathcal{C} = \{\mathbf{c}_1, \dots, \mathbf{c}_m\}$ de la *llegada*. Cuando $V = W$ suele tomarse $\mathcal{C} = \mathcal{B}$, pero conviene escribirlo explícitamente.

2. **Aplica T a cada vector de $\mathcal{B}$, uno por uno.** Obtienes n vectores de W: $T(\mathbf{b}_1), \dots, T(\mathbf{b}_n)$. Ni más ni menos; la linealidad hace el resto.

3. **Expresa cada imagen en la base $\mathcal{C}$ y ponla *en columna*.** La matriz es

$$[T]_{\mathcal{C}\leftarrow\mathcal{B}} = \Big(\, [T(\mathbf{b}_1)]_{\mathcal{C}} \;\Big|\; \cdots \;\Big|\; [T(\mathbf{b}_n)]_{\mathcal{C}} \, \Big).$$

   Columnas, no filas. Si las pones en filas obtienes la traspuesta y todo lo demás sale al revés.

4. **Comprueba el tamaño.** La matriz es $m \times n$: tantas filas como vectores tiene la base de llegada, tantas columnas como tiene la de salida. Si los números no cuadran, has confundido las bases.

5. **Para cambiar de base, construye la matriz de paso.** Si $\mathcal{B}' = \{\mathbf{b}'_1, \dots, \mathbf{b}'_n\}$ es otra base y todo se refiere a la canónica, la matriz P que tiene a los $\mathbf{b}'_i$ *como columnas* cumple

$$[\mathbf{x}]_{\text{can}} = P\,[\mathbf{x}]_{\mathcal{B}'}.$$

   Es decir, P traduce *de* la base nueva *a* la canónica, aunque se la llame "matriz de cambio a la base nueva"; la dirección es la fuente de la mitad de los errores de signo.

6. **Aplica la fórmula de semejanza.** Si M es la matriz de T en la base canónica, su matriz en la base $\mathcal{B}'$ es

$$[T]_{\mathcal{B}'} = P^{-1} M\, P.$$

   Léela de derecha a izquierda: se pasa de coordenadas nuevas a canónicas (P), se aplica T (M) y se vuelve a las nuevas (P^{-1}).

7. **Verifica con los invariantes.** La traza, el determinante, el polinomio característico, el rango y los autovalores *no* dependen de la base. Si tu matriz nueva tiene otra traza que la vieja, hay un error. Y si eliges como base una de autovectores, la matriz debe salir diagonal.

Ejemplo 7.49 (Una reflexión, un operador sobre polinomios y una derivada nilpotente)**. (a) La reflexión respecto del plano $x+y+z=0$.** La fórmula general de la reflexión respecto del plano de normal $\mathbf{n}$ es $M = I - \dfrac{2\,\mathbf{n}\mathbf{n}^T}{\mathbf{n}^T\mathbf{n}}$. Con $\mathbf{n} = (1,1,1)^T$ se tiene $\mathbf{n}^T\mathbf{n} = 3$ y

$$M = I - \frac{2}{3}\begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{pmatrix} = \begin{pmatrix} \frac{1}{3} & -\frac{2}{3} & -\frac{2}{3} \\ -\frac{2}{3} & \frac{1}{3} & -\frac{2}{3} \\ -\frac{2}{3} & -\frac{2}{3} & \frac{1}{3} \end{pmatrix}.$$

Comprobaciones de rutina: $\operatorname{tr} M = 1$, $\det M = -1$ y $M^2 = I$, las tres firmas inconfundibles de una reflexión en $\mathbb{R}^3$.

(b) La misma reflexión, en una base adaptada. Tomemos

$$\mathbf{b}_1 = (1,-1,0), \qquad \mathbf{b}_2 = (1,1,-2), \qquad \mathbf{b}_3 = (1,1,1).$$

Los dos primeros están en el plano (suman cero sus coordenadas) y el tercero es la normal. Sin calcular nada podemos anticipar la respuesta: $T(\mathbf{b}_1) = \mathbf{b}_1$, $T(\mathbf{b}_2) = \mathbf{b}_2$, $T(\mathbf{b}_3) = -\mathbf{b}_3$, así que

$$[T]_{\mathcal{B}'} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{pmatrix}.$$

Y en efecto, con $P = \begin{pmatrix} 1 & 1 & 1 \\ -1 & 1 & 1 \\ 0 & -2 & 1 \end{pmatrix}$ el cálculo $P^{-1}MP$ devuelve exactamente esa matriz diagonal. Los invariantes coinciden: traza $1+1-1=1$ y determinante -1, como en (a). Toda la geometría de la reflexión cabe en tres números cuando se eligen bien las coordenadas.

(c) Un operador sobre polinomios. Sea $V = P_2$, el espacio de polinomios de grado ≤ 2, con la base $\mathcal{B} = \{1,\ x,\ x^2\}$, y sea

$$T(p) = p'' + x\,p'.$$

Aplicamos T a cada elemento de la base:

$$T(1) = 0, \qquad T(x) = 0 + x\cdot 1 = x, \qquad T(x^2) = 2 + x\cdot 2x = 2 + 2x^2.$$

Escribimos cada imagen en coordenadas respecto de $\{1, x, x^2\}$ y las ponemos en columnas:

$$[T]_{\mathcal{B}} = \begin{pmatrix} 0 & 0 & 2 \\ 0 & 1 & 0 \\ 0 & 0 & 2 \end{pmatrix}.$$

Es triangular superior, así que los autovalores se leen en la diagonal: 0, 1 y 2. Los autovectores correspondientes son $p = 1$ (porque $T(1) = 0$), $p = x$ (porque $T(x) = x$) y $p = 1 + x^2$, ya que

$$T(1 + x^2) = 0 + (2 + 2x^2) = 2(1 + x^2). \quad \checkmark$$

En la base $\{1,\ x,\ 1 + x^2\}$ el operador se escribe $\operatorname{diag}(0, 1, 2)$.

(d) La derivada, que no se puede diagonalizar. Sea $D : P_3 \to P_3$, $D(p) = p'$, con la base $\{1, x, x^2, x^3\}$. Entonces $D(1) = 0$, $D(x) = 1$, $D(x^2) = 2x$, $D(x^3) = 3x^2$, y

$$[D] = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 3 \\ 0 & 0 & 0 & 0 \end{pmatrix}.$$

Es nilpotente: $[D]^4 = 0$ pero $[D]^3 \neq 0$ (contiene un $1 \cdot 2 \cdot 3 = 6$ en la esquina). Su único autovalor es 0 con multiplicidad algebraica 4 y geométrica 1, así que D *no* es diagonalizable en ninguna base: su forma de Jordan es un único bloque 4×4. Ningún cambio de base puede arreglarlo, porque el rango y la nilpotencia son invariantes.

(e) Un caso rectangular, para fijar el tamaño. Sea $S : P_2 \to \mathbb{R}^2$ dada por $S(p) = (p(0),\ p(1))$, con $\mathcal{B} = \{1, x, x^2\}$ y $\mathcal{C}$ la base canónica de $\mathbb{R}^2$. Entonces $S(1) = (1, 1)$, $S(x) = (0, 1)$, $S(x^2) = (0, 1)$, y

$$[S]_{\mathcal{C} \leftarrow \mathcal{B}} = \begin{pmatrix} 1 & 0 & 0 \\ 1 & 1 & 1 \end{pmatrix},$$

de tamaño 2×3, como anuncia el paso 4. Su núcleo tiene dimensión $3 - 2 = 1$ y está generado por $x^2 - x$, el único polinomio de P_2 (salvo múltiplos) que se anula en 0 y en 1.

Cuidado

Cuidado con la dirección de P. Si P tiene a los vectores de la base nueva como columnas, entonces P va *de* coordenadas nuevas *a* canónicas, y la fórmula correcta es $[T]_{\text{nueva}} = P^{-1}MP$. Escribir PMP^{-1} da la matriz de T "vista al revés", que en general es distinta. Un control rápido: aplica la fórmula a la base de autovectores; si no sale la diagonal de autovalores, tienes P invertida.

Pausa para pensar

Dos matrices semejantes, $B = P^{-1}AP$, describen la misma transformación en coordenadas distintas. Por eso comparten traza, determinante, polinomio característico, rango y forma de Jordan: todo eso es propiedad de la transformación, no del sistema de coordenadas. La pregunta "¿cuál es la base en la que esta matriz se ve lo más simple posible?" tiene por respuesta la diagonalización cuando se puede, y la forma de Jordan cuando no; y la clasificación de todas las matrices salvo semejanza es exactamente la clasificación de las formas de Jordan. Lo que en el capítulo 25 se llamará "clasificar objetos salvo isomorfismo" es esta misma idea, un piso más arriba.

7.5. Determinantes

Idea clave

Intuición primero. El determinante es el *factor de escala de volúmenes*: la transformación manda el cubo unitario a un paralelepípedo, y $|\det|$ es su volumen; el signo dice si además hubo un espejo (orientación invertida). Todo lo demás son corolarios de esta imagen: $\det = 0$ significa que el cubo quedó aplastado en algo plano (dimensión perdida: no invertible); $\det(AB) = \det A \det B$ porque escalar volúmenes dos veces multiplica los factores; y las filas repetidas dan cero porque un paralelepípedo con dos aristas iguales es plano.

Definición 7.50. $\det : M_n(K) \to K$ es la única función multilineal y alternada en las filas con $\det I = 1$. Fórmula explícita (Leibniz): $\det A = \sum_{\sigma \in S_n} \operatorname{sgn}(\sigma) \prod_i A_{i\sigma(i)}$, suma sobre las $n!$ permutaciones —de aquí la conexión con el capítulo 25—.

Consecuencias operativas: el determinante cambia de signo al intercambiar filas, no cambia al sumar a una fila un múltiplo de otra, se multiplica por c al escalar una fila, y el de una triangular es el producto de la diagonal. **El método práctico es escalonar y llevar la cuenta**: coste $O(n^3)$ contra $O(n!)$ de la fórmula.

Teorema 7.51. $\det(AB) = \det A \det B$; $\det A^{\mathsf{T}} = \det A$; *$A$ invertible* $\iff \det A \neq 0$; $\det(A^{-1}) = 1/\det A$; *matrices semejantes tienen el mismo determinante.*

Idea clave

Qué dice esto en cristiano. El determinante es un volumen con signo. En dimensión dos es el área del paralelogramo que forman las filas; en dimensión tres, el volumen del paralelepípedo. Las tres propiedades de la definición son justo las que tiene un volumen: crece proporcionalmente al estirar un lado (multilineal), se anula si dos lados coinciden porque la figura se aplasta (alternada), y el cubo unidad mide 1. De ahí sale todo lo demás: que $\det(AB) = \det A \det B$ significa que aplicar dos transformaciones multiplica los factores de escala, y que $\det A = 0$ signifique no invertible significa que la transformación aplasta el volumen a cero.

Ejemplo 7.52 (Calcular determinantes y comprobar las propiedades)**. (a) El producto.** Con $A = \begin{pmatrix} 2 & 1 \\ 1 & 3 \end{pmatrix}$ y $B = \begin{pmatrix} 1 & 4 \\ 0 & 2 \end{pmatrix}$: $\det A = 5$, $\det B = 2$ y $\det(AB) = 10 = 5 \cdot 2$. También $\det A^{\mathsf{T}} = 5$ y $\det(A^{-1}) = \frac{1}{5}$.

(b) Un tres por tres. Para $C = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 10 \end{pmatrix}$ el determinante vale -3. Al intercambiar las dos primeras filas pasa a valer 3: cambia de signo, como exige la alternancia. Y una matriz con dos filas iguales tiene determinante 0, porque el paralelogramo se aplasta.

(c) La interpretación geométrica. El paralelogramo de lados $(1,2)$ y $(3,1)$ tiene área $\left|\det \begin{pmatrix} 1 & 2 \\ 3 & 1 \end{pmatrix}\right| = |1-6| = 5$. El paralelepípedo de lados $(1,0,0)$, $(1,1,0)$ y $(1,1,1)$ tiene volumen 1: es un cubo deformado sin cambiar de volumen, porque las operaciones fueron de cizalla.

(d) Por qué el signo importa. El signo del determinante dice si la transformación conserva la orientación o la invierte. Una rotación tiene determinante $+1$; una reflexión, -1. Por eso ninguna combinación de rotaciones puede producir una imagen especular: el determinante, que siempre vale 1 en las rotaciones, tendría que valer -1.

Cuidado

[El determinante no es lineal, ni suma bien] $\det(A+B)$ no tiene nada que ver con $\det A + \det B$: con $A = B = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$ se tiene $\det(A+B) = \det \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix} = 4$, mientras que $\det A + \det B = 2$. La multilinealidad es *fila a fila*, no sobre la matriz entera. Segundo aviso: multiplicar toda una matriz $n \times n$ por un escalar multiplica el determinante por c^n, no por c; para $n = 3$ y $c = 2$ el factor es 8. Y tercero: la fórmula de Leibniz con $n!$ sumandos es inútil para calcular; ya con $n = 20$ tendría más de dos billones de términos, y por eso siempre se usa eliminación gaussiana.

Desarrollo por cofactores: $\det A = \sum_j (-1)^{i+j} A_{ij} M_{ij}$ (fila i cualquiera, M_{ij} el menor complementario). De ahí la *regla de Cramer* para sistemas $n \times n$ con $\det A \neq 0$: $x_i = \frac{\det A_i}{\det A}$, donde A_i reemplaza la columna i por $\mathbf{b}$ (útil en teoría; en la práctica, Gauss es más barato).

Interpretación geométrica: $|\det A|$ es el volumen del paralelepípedo generado por las columnas de A; el signo indica si se preserva la orientación. Por eso el jacobiano aparece en el cambio de variables de las integrales múltiples (capítulo 9): una aplicación lineal multiplica todos los volúmenes por $|\det|$.

Regla: calcular un determinante sin equivocarse

Hay tres métodos y conviene saber cuál toca. Desarrollar por cofactores en un 4×4 genérico son 24 productos de cuatro factores: se falla seguro. Escalonar son cuatro divisiones. La elección no es de gusto, es de supervivencia.

1. **Tamaño 2×2 y 3×3: fórmula directa.** Para 2×2, $ad - bc$. Para 3×3, la regla de Sarrus. *Sarrus no funciona en 4×4 ni en ningún tamaño mayor*: es un error muy extendido y da resultados sin sentido.

2. **Busca ceros antes de nada.** Si alguna fila o columna tiene muchos ceros, desarrolla por ella: cada cero elimina un sumando entero. Y si no los hay, créalos con operaciones elementales, que es el paso siguiente.

3. **Tamaño ≥ 4: escalona.** Sumar a una fila un múltiplo de otra **no cambia** el determinante. Es la operación gratis, y con ella se lleva la matriz a forma triangular; entonces el determinante es el producto de la diagonal.

4. **Contabiliza las otras dos operaciones.** Intercambiar dos filas **cambia el signo**. Multiplicar una fila por λ **multiplica el determinante por** λ — y por tanto, si sacas factor común de una fila, el factor sale fuera *una sola vez*, no elevado a n. Lleva la cuenta en un margen; olvidar un cambio de signo es el error número uno.

5. **Usa las propiedades multiplicativas como atajo.** $\det(AB) = \det A \det B$, $\det(A^T) = \det A$, $\det(\lambda A) = \lambda^n \det A$ (aquí sí aparece la potencia, porque se multiplican las n filas), $\det(A^{-1}) = 1/\det A$. Una matriz triangular o triangular por bloques se lee directamente.

6. **Recuerda qué significa el resultado.** $\det A = 0$ equivale a: columnas linealmente dependientes, A no invertible, núcleo no trivial, sistema homogéneo con soluciones no triviales, 0 es valor propio. Son cinco formas de decir lo mismo, y saltar entre ellas es medio curso de álgebra lineal.

7. **Control.** Si la matriz tiene dos filas iguales o proporcionales, el determinante es 0 y no hay que calcular nada. Y para una matriz de enteros el determinante es un entero: si te sale una fracción, hay un error.

Ejemplo 7.53 (Un 3×3 por dos caminos y un 4×4 por escalonamiento)**. (a) Un 3×3, comprobado dos veces.** Sea

$$M = \begin{pmatrix} 2 & -1 & 3 \\ 1 & 4 & -2 \\ 3 & 0 & 1 \end{pmatrix}.$$

Por Sarrus:

$$\det M = (2)(4)(1)+(-1)(-2)(3)+(3)(1)(0)-(3)(4)(3)-(2)(-2)(0)-(-1)(1)(1) = 8+6+0-36-0+1 = -21.$$

Por cofactores a lo largo de la tercera fila, que tiene un cero:

$$\det M = 3\begin{vmatrix} -1 & 3 \\ 4 & -2 \end{vmatrix} - 0 + 1\begin{vmatrix} 2 & -1 \\ 1 & 4 \end{vmatrix} = 3(2-12) + (8+1) = -30+9 = -21.$$

Coinciden. Nótese cuánto más corto fue el segundo camino gracias al cero, y que hay que respetar el patrón de signos $+,-,+$ de la tercera fila.

(b) Un 4×4: aquí Sarrus no existe. Sea

$$N = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 2 & 3 & 1 & 2 \\ 1 & 1 & 1 & -1 \\ 1 & 0 & -2 & -6 \end{pmatrix}.$$

El pivote 1 de la esquina es ideal. Con $F_2 \to F_2 - 2F_1$, $F_3 \to F_3 - F_1$, $F_4 \to F_4 - F_1$ (tres operaciones gratis, el determinante no cambia):

$$\begin{pmatrix} 1 & 2 & 3 & 4 \\ 0 & -1 & -5 & -6 \\ 0 & -1 & -2 & -5 \\ 0 & -2 & -5 & -10 \end{pmatrix}.$$

Ahora el pivote es -1. Con $F_3 \to F_3 - F_2$ y $F_4 \to F_4 - 2F_2$:

$$\begin{pmatrix} 1 & 2 & 3 & 4 \\ 0 & -1 & -5 & -6 \\ 0 & 0 & 3 & 1 \\ 0 & 0 & 5 & 2 \end{pmatrix}.$$

Y finalmente $F_4 \to F_4 - \frac{5}{3}F_3$ deja $0, 0, 0, 2 - \frac{5}{3} = \frac{1}{3}$. La matriz es triangular y el determinante es el producto de la diagonal:

$$\det N = 1 \cdot (-1) \cdot 3 \cdot \tfrac{1}{3} = -1.$$

No hubo ni un solo cambio de signo (no se intercambiaron filas) ni ningún factor sacado fuera: todas las operaciones fueron del tipo gratis. Y el control del paso 7 se cumple: la matriz era de enteros y el resultado es entero.

Pausa para pensar

Compárese el coste. Por cofactores, un 4×4 se reduce a cuatro determinantes 3×3, cada uno con seis productos: 24 productos de cuatro factores, más los signos. Por escalonamiento fueron seis operaciones de fila y un producto de cuatro números. Para un 10×10 la diferencia es entre $10! \approx 3{,}6$ millones de productos y unas 300 operaciones. Esa brecha es la razón de que todo programa de cálculo numérico use eliminación y ninguno use cofactores.

Regla: invertir una matriz: tres caminos y la decisión de no recorrerlos

Invertir una matriz es una de esas operaciones que se piden mucho más de lo que se necesitan. Antes de calcular A^{-1} conviene preguntarse si de verdad hace falta, y si hace falta, elegir bien el método. Estos son los siete pasos.

1. **Pregúntate primero si necesitas la inversa o solo resolver un sistema.** Para $A\mathbf{x} = \mathbf{b}$ *nunca* conviene calcular A^{-1} y multiplicar: la eliminación directa cuesta unas tres veces menos y es numéricamente más estable. La inversa solo se justifica cuando el objeto que se busca es la matriz en sí (por ejemplo, una matriz de covarianzas o el operador inverso de una teoría).

2. **Comprueba que exista: $\det A \neq 0$.** Equivalentemente, rango $A = n$, o el sistema homogéneo solo tiene la solución trivial, o ningún autovalor es cero. Si el determinante

es muy pequeño comparado con las entradas, la matriz es *mal condicionada* y la inversa numérica no será de fiar aunque exista.

3. **Método 1, Gauss–Jordan: reduce $[\,A \mid I\,]$ hasta $[\,I \mid A^{-1}\,]$.** Es el método de elección en la práctica y a mano para $n \geq 3$. Coste del orden de n^3 operaciones. Cada operación elemental que apliques a la izquierda aplícala también a la derecha, sin excepción.

4. **Método 2, matriz adjunta: $A^{-1} = \dfrac{1}{\det A}\operatorname{adj}(A)$.** La adjunta es la traspuesta de la matriz de cofactores: $\operatorname{adj}(A)_{ij} = (-1)^{i+j} M_{ji}$, donde M_{ji} es el menor que se obtiene tachando la fila j y la columna i. *Nótese la trasposición de índices*: es el error clásico. Solo es práctico para $n = 2$ y $n = 3$.

5. **Método 3, factorización LU: resuelve n sistemas de golpe.** Si $A = LU$, entonces la columna j de A^{-1} es la solución de $A\mathbf{x} = \mathbf{e}_j$, y con LU ya calculada cada una cuesta solo $O(n^2)$. Es lo que hace cualquier biblioteca numérica seria.

6. **Aprovecha la estructura si la hay.** Para matrices triangulares por bloques,

$$\begin{pmatrix} A & B \\ 0 & D \end{pmatrix}^{-1} = \begin{pmatrix} A^{-1} & -A^{-1}BD^{-1} \\ 0 & D^{-1} \end{pmatrix}.$$

   Para una perturbación de rango uno, la fórmula de Sherman–Morrison:

$$(A + \mathbf{u}\mathbf{v}^T)^{-1} = A^{-1} - \frac{A^{-1}\mathbf{u}\mathbf{v}^T A^{-1}}{1 + \mathbf{v}^T A^{-1}\mathbf{u}},$$

   válida siempre que el denominador no se anule.

7. **Verifica multiplicando, en los dos órdenes.** Calcula AA^{-1} y $A^{-1}A$: ambos deben dar I. Para matrices cuadradas basta uno de los dos, pero hacer los dos detecta errores de trasposición al instante.

Ejemplo 7.54 (Una matriz 3×3 invertida de tres maneras, y dos atajos estructurales)**.** Trabajamos con

$$A = \begin{pmatrix} 2 & 1 & 1 \\ 1 & 3 & 2 \\ 1 & 0 & 0 \end{pmatrix}.$$

(a) El determinante. Desarrollando por la tercera fila, que tiene dos ceros:

$$\det A = 1 \cdot \det \begin{pmatrix} 1 & 1 \\ 3 & 2 \end{pmatrix} = 1 \cdot (2 - 3) = -1.$$

Es distinto de cero, así que la inversa existe. Y como vale -1, la inversa tendrá entradas enteras: buena señal para trabajar a mano.

(b) Método de la adjunta. Los nueve cofactores $C_{ij} = (-1)^{i+j} M_{ij}$:

$$C_{11} = 0, \quad C_{12} = -(0-2) = 2, \quad C_{13} = -3,$$

$$C_{21} = -(0-0) = 0, \quad C_{22} = -1, \quad C_{23} = 1,$$

$$C_{31} = (2-3) = -1, \quad C_{32} = -(4-1) = -3, \quad C_{33} = 5.$$

La adjunta es la *traspuesta* de esta matriz de cofactores:

$$\operatorname{adj}(A) = \begin{pmatrix} 0 & 0 & -1 \\ 2 & -1 & -3 \\ -3 & 1 & 5 \end{pmatrix},$$

y dividiendo entre $\det A = -1$:

$$A^{-1} = \begin{pmatrix} 0 & 0 & 1 \\ -2 & 1 & 3 \\ 3 & -1 & -5 \end{pmatrix}.$$

(c) Método de Gauss–Jordan. Partimos de $[\,A \mid I\,]$ e intercambiamos $F_1 \leftrightarrow F_3$ para tener un pivote 1:

$$\left(\begin{array}{ccc|ccc} 1 & 0 & 0 & 0 & 0 & 1 \\ 1 & 3 & 2 & 0 & 1 & 0 \\ 2 & 1 & 1 & 1 & 0 & 0 \end{array}\right) \xrightarrow[F_3 - 2F_1]{F_2 - F_1} \left(\begin{array}{ccc|ccc} 1 & 0 & 0 & 0 & 0 & 1 \\ 0 & 3 & 2 & 0 & 1 & -1 \\ 0 & 1 & 1 & 1 & 0 & -2 \end{array}\right).$$

Ahora $F_2 \leftrightarrow F_3$, luego $F_3 - 3F_2$, y queda un -1 en la posición $(3,3)$ que se normaliza y se elimina hacia arriba. El resultado final es la misma matriz del apartado (b), como debía ser. La ventaja del método es que no crece el trabajo tan deprisa: para $n = 6$ la adjunta pediría 36 determinantes 5×5 y Gauss–Jordan sigue siendo una eliminación normal.

(d) Factorización LU y resolución de un sistema. Sin permutar,

$$L = \begin{pmatrix} 1 & 0 & 0 \\ \tfrac{1}{2} & 1 & 0 \\ \tfrac{1}{2} & -\tfrac{1}{5} & 1 \end{pmatrix}, \qquad U = \begin{pmatrix} 2 & 1 & 1 \\ 0 & \tfrac{5}{2} & \tfrac{3}{2} \\ 0 & 0 & -\tfrac{1}{5} \end{pmatrix}, \qquad LU = A.$$

Nótese que $\det A = \det U = 2 \cdot \tfrac{5}{2} \cdot (-\tfrac{1}{5}) = -1$: la factorización regala el determinante. Para resolver $A\mathbf{x} = (1,2,3)^T$ se hace descenso $L\mathbf{y} = \mathbf{b}$ y luego remonte $U\mathbf{x} = \mathbf{y}$, y sale

$$\mathbf{x} = (3,\ 9,\ -14)^T.$$

Comprobación: $2(3) + 9 + (-14) = 1$; $3 + 3(9) + 2(-14) = 3 + 27 - 28 = 2$; $3 + 0 + 0 = 3$. Correcto, y sin haber calculado la inversa en ningún momento, tal como recomienda el paso 1.

(e) Sherman–Morrison en acción. Sea $\mathbf{u} = (1,1,1)^T$ y

$$B = I + \mathbf{u}\mathbf{u}^T = \begin{pmatrix} 2 & 1 & 1 \\ 1 & 2 & 1 \\ 1 & 1 & 2 \end{pmatrix}.$$

Con $A = I$ la fórmula del paso 6 da directamente, sin eliminar nada,

$$B^{-1} = I - \frac{\mathbf{u}\mathbf{u}^T}{1 + \mathbf{u}^T\mathbf{u}} = I - \frac{\mathbf{u}\mathbf{u}^T}{4} = \begin{pmatrix} \frac{3}{4} & -\frac{1}{4} & -\frac{1}{4} \\ -\frac{1}{4} & \frac{3}{4} & -\frac{1}{4} \\ -\frac{1}{4} & -\frac{1}{4} & \frac{3}{4} \end{pmatrix}.$$

Verificación de una entrada: la fila 1 de B por la columna 1 de B^{-1} es $2\cdot\frac{3}{4}+1\cdot(-\frac{1}{4})+1\cdot(-\frac{1}{4}) = \frac{3}{2}-\frac{1}{2} = 1$. Correcto. Esta fórmula es la que permite actualizar un modelo estadístico cuando llega un dato nuevo sin rehacer toda la inversión.

(f) Por qué Cramer no se usa. La regla de Cramer es una fórmula preciosa y una catástrofe computacional. Resolver un sistema 20×20 con ella exige 21 determinantes de orden 20; desarrollados por la definición son $21! \approx 5.11 \times 10^{19}$ productos de 19 factores, unas 9.7×10^{20} operaciones. A un billón de operaciones por segundo eso son *casi treinta y un años*. La eliminación de Gauss para el mismo sistema necesita del orden de $\frac{2}{3}n^3 \approx 5\,300$ operaciones: unos microsegundos. La diferencia no es de constantes, es de $n!$ frente a n^3.

Cuidado

Que $\det A \neq 0$ garantice la existencia de A^{-1} no significa que el cálculo sea fiable. La matriz de Hilbert $H_{ij} = 1/(i+j-1)$ tiene determinante no nulo para todo n, pero su número de condición crece exponencialmente: para $n = 10$ supera 10^{13}, de modo que en aritmética de doble precisión la inversa calculada no conserva ni una cifra significativa. En problemas numéricos, "invertible" y "invertible en la práctica" son cosas distintas, y quien decide es el número de condición, no el determinante.

Idea clave

La adjunta parece un método peor en todos los sentidos, y computacionalmente lo es. Pero tiene una virtud que Gauss–Jordan no tiene: es una fórmula *polinómica* en las entradas. Por eso la identidad $A \cdot \operatorname{adj}(A) = (\det A)\, I$ vale sobre cualquier anillo conmutativo, no solo sobre un cuerpo, y de ahí sale el criterio de que una matriz con entradas enteras tiene inversa entera si y solo si su determinante es ± 1. Ese es el origen del grupo $GL_n(\mathbb{Z})$ y de las matrices unimodulares que aparecen en el capítulo 24. Los métodos rápidos calculan mejor; las fórmulas cerradas demuestran más.

7.6. Valores y vectores propios

Idea clave

Intuición primero. Una transformación lineal revuelve el espacio, pero suele tener *direcciones privilegiadas* donde no rota nada: solo estira o encoge. Esos son los vectores propios, y el factor de estiramiento es el valor propio. Diagonalizar es ponerse los lentes correctos: en la base propia, la matriz entera se reduce a n perillas de escala independientes. Y como aplicar la transformación mil veces es girar cada perilla mil veces (λ^{1000}), los valores propios gobiernan todo comportamiento a largo plazo: por eso

aparecen en la estabilidad de las EDO, en PageRank, en las potencias de Fibonacci y en cada resonancia de este libro.

Definición 7.55. $\lambda \in K$ es *valor propio* de A si existe $v \neq \mathbf{0}$ con $Av = \lambda v$. El conjunto de tales v (más $\mathbf{0}$) es el *espacio propio* $E_\lambda = \ker(A - \lambda I)$. Los valores propios son las raíces del *polinomio característico* $p_A(\lambda) = \det(A - \lambda I)$, de grado n; su *multiplicidad algebraica* es la multiplicidad como raíz, y la *geométrica* es $\dim E_\lambda$. Siempre $1 \leq \text{geom} \leq \text{alg}$.

Datos útiles: la suma de los valores propios (con multiplicidad) es la traza $\operatorname{tr} A$; el producto es $\det A$. Para 2×2: $p(\lambda) = \lambda^2 - (\operatorname{tr} A)\lambda + \det A$.

Teorema 7.56. *Vectores propios de valores propios distintos son linealmente independientes.*

Demostración. Inducción. Si $c_1 v_1 + \cdots + c_k v_k = \mathbf{0}$ con λ_i distintos, aplicamos $A - \lambda_k I$: el último término muere y queda $\sum_{i<k} c_i(\lambda_i - \lambda_k)v_i = \mathbf{0}$; por hipótesis inductiva $c_i(\lambda_i - \lambda_k) = 0$, y como $\lambda_i \neq \lambda_k$, $c_i = 0$ para $i < k$; entonces también $c_k = 0$. □

Teorema 7.57 (Criterio de diagonalización)**.** *A es diagonalizable $\iff$ para cada valor propio, multiplicidad geométrica = algebraica $\iff$ existe base de vectores propios. En particular, n valores propios distintos garantizan diagonalizabilidad.*

Ejemplo 7.58. $A = \begin{pmatrix} 4 & 1 \\ 2 & 3 \end{pmatrix}$: $p(\lambda) = \lambda^2 - 7\lambda + 10 = (\lambda - 5)(\lambda - 2)$. Para $\lambda = 5$: $(A - 5I)v = 0$ da $v = (1, 1)$; para $\lambda = 2$: $v = (1, -2)$. Con $P = \begin{pmatrix} 1 & 1 \\ 1 & -2 \end{pmatrix}$, $P^{-1}AP = \operatorname{diag}(5, 2)$, y

$$A^k = P \operatorname{diag}(5^k, 2^k)\, P^{-1} \quad \text{para todo } k :$$

diagonalizar convierte potencias de matrices en potencias de números.

Ejemplo 7.59 (Fibonacci en forma cerrada)**.** La recurrencia $F_{n+1} = F_n + F_{n-1}$ se escribe $\binom{F_{n+1}}{F_n} = A\binom{F_n}{F_{n-1}}$ con $A = \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix}$. Los valores propios son las raíces de $\lambda^2 - \lambda - 1$: $\varphi = \frac{1+\sqrt{5}}{2}$ y $\psi = \frac{1-\sqrt{5}}{2}$. Diagonalizando y aplicando A^n al vector inicial se obtiene la fórmula de Binet:

$$F_n = \frac{\varphi^n - \psi^n}{\sqrt{5}}.$$

Como $|\psi| < 1$, F_n es el entero más cercano a $\varphi^n/\sqrt{5}$: la razón áurea gobierna el crecimiento.

Teorema 7.60 (Cayley–Hamilton)**.** *Toda matriz anula su polinomio característico: $p_A(A) = 0$.*

Consecuencia práctica: A^n (y A^{-1} si existe) se expresan como combinación de $I, A, \ldots, A^{n-1}$. Para la A del Ejemplo 7.58: $A^2 = 7A - 10I$, y $A^{-1} = \frac{7I - A}{10}$.

Regla: diagonalizar una matriz paso a paso

Diagonalizar es encontrar P invertible y D diagonal con $A = PDP^{-1}$, es decir, encontrar una base en la que la transformación solo estire cada eje. El procedimiento es rígido y hay que seguirlo en orden.

1. **Calcula el polinomio característico $p(\lambda) = \det(A - \lambda I)$ y sus raíces.** Esas raíces son los valores propios. Anota la *multiplicidad algebraica* $m_a(\lambda)$ de cada una, que es su multiplicidad como raíz.
2. **Para cada λ, resuelve el sistema homogéneo $(A - \lambda I)v = 0$.** El espacio de soluciones es el subespacio propio E_λ, y su dimensión es la *multiplicidad geométrica* $m_g(\lambda) = n - \text{rango}(A - \lambda I)$. Los vectores de una base de E_λ son los vectores propios que buscas.
3. **Compara las dos multiplicidades.** Siempre se cumple $1 \le m_g(\lambda) \le m_a(\lambda)$. Y el criterio es exacto:

   $$A \text{ es diagonalizable} \iff m_g(\lambda) = m_a(\lambda) \text{ para todo valor propio.}$$

   Si en algún valor propio $m_g < m_a$, la matriz es *defectuosa* y no hay nada que hacer: hay que ir a la forma de Jordan.
4. **Dos atajos que ahorran todo el trabajo.** Si hay n valores propios *distintos*, es diagonalizable sin más comprobación. Si la matriz es *simétrica real*, es diagonalizable y además con P ortogonal: eso es el teorema espectral, y no admite excepciones.
5. **Monta P y D en el mismo orden.** Las columnas de P son los vectores propios; las entradas de D son los valores propios *en las posiciones correspondientes*. Si intercambias dos columnas de P tienes que intercambiar las dos entradas de D: el orden es libre pero tiene que ser el mismo en las dos matrices.
6. **Verifica con $AP = PD$.** Esta comprobación no requiere invertir nada y es la más rápida. Columna a columna dice exactamente $Av_k = \lambda_k v_k$, que es la definición.
7. **Controles numéricos.** La traza es la suma de los valores propios (con multiplicidad) y el determinante es su producto. Si tus valores propios no reproducen esos dos números, están mal antes de seguir.

Ejemplo 7.61 (Una simétrica con valor propio doble, y una defectuosa)**. (a) Diagonalizable, con multiplicidad 2.** Sea

$$A = \begin{pmatrix} 5 & 4 & 2 \\ 4 & 5 & 2 \\ 2 & 2 & 2 \end{pmatrix},$$

simétrica, luego el paso 4 ya garantiza el resultado; hagamos las cuentas de todos modos. El polinomio característico es $p(\lambda) = -(\lambda - 10)(\lambda - 1)^2$, así que $\lambda = 10$ con $m_a = 1$ y $\lambda = 1$ con $m_a = 2$. Control del paso 7: $\operatorname{tr} A = 5 + 5 + 2 = 12$ y $10 + 1 + 1 = 12$ ✓; $\det A = 10 \cdot 1 \cdot 1 = 10$

✓.

Para $\lambda = 10$ el sistema $(A - 10I)v = 0$ tiene matriz $\left(\begin{smallmatrix} -5 & 4 & 2 \\ 4 & -5 & 2 \\ 2 & 2 & -8 \end{smallmatrix}\right)$, de rango 2, luego $m_g = 3 - 2 = 1 = m_a$, y una base del núcleo es $v_1 = (2, 2, 1)^T$.

Para $\lambda = 1$ el sistema tiene matriz $\left(\begin{smallmatrix} 4 & 4 & 2 \\ 4 & 4 & 2 \\ 2 & 2 & 1 \end{smallmatrix}\right)$: las tres filas son proporcionales, el rango es 1, luego $m_g = 3 - 1 = 2 = m_a$. Una base del núcleo (todo el plano $4x + 4y + 2z = 0$) es

$$v_2 = (-1, 1, 0)^T, \qquad v_3 = (-1, 0, 2)^T.$$

Como $m_g = m_a$ en ambos casos, A es diagonalizable, con

$$P = \begin{pmatrix} 2 & -1 & -1 \\ 2 & 1 & 0 \\ 1 & 0 & 2 \end{pmatrix}, \qquad D = \begin{pmatrix} 10 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

Verificación por el paso 6: la primera columna de AP es $Av_1 = (10 \cdot 2, 10 \cdot 2, 10 \cdot 1)^T = (20, 20, 10)^T$, que es $10\, v_1$ ✓, y las otras dos columnas de AP son v_2 y v_3 sin cambio ✓. Obsérvese además que v_1 es ortogonal a v_2 y a v_3, como predice el teorema espectral para una matriz simétrica.

(b) Defectuosa: la que no se puede. Sea

$$B = \begin{pmatrix} 3 & 1 \\ 0 & 3 \end{pmatrix}.$$

El polinomio característico es $(\lambda - 3)^2$, así que $\lambda = 3$ con $m_a = 2$. Pero

$$B - 3I = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$$

tiene rango 1, luego $m_g = 2 - 1 = 1 < 2 = m_a$. *No* es diagonalizable: solo hay una dirección propia, la del eje x. Si lo fuera, sería $B = P(3I)P^{-1} = 3I$, lo cual es falso porque $B \neq 3I$. Este razonamiento vale en general: *una matriz con un único valor propio es diagonalizable si y solo si ya es un múltiplo escalar de la identidad.*

Cuidado

Que un valor propio esté repetido no impide la diagonalización: en el ejemplo (a) el valor 1 aparece dos veces y todo funciona. Lo que la impide es que *falten vectores propios*, es decir, $m_g < m_a$. Confundir "valor propio repetido" con "no diagonalizable" es el malentendido más frecuente de la sección, y el ejemplo (a) está aquí precisamente para desmentirlo.

7.6.1. Cuando no se puede diagonalizar: Jordan

Idea clave

Intuición primero. Un bloque de Jordan es "escala más cinta transportadora": la diagonal λ estira, y los unos de arriba hacen que cada vector de la base *gotee* hacia el anterior en cada aplicación. Esa gotera es la que fabrica los factores t: en e^{tA} el bloque produce $te^{\lambda t}$, la respuesta "resonante" de una raíz repetida en las EDO. La matriz no diagonalizable no es una patología exótica: es una transformación con memoria de un paso, y Jordan dice que *toda* la no-diagonalizabilidad del mundo se reduce a cadenas de este único mecanismo.

$\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$ tiene $\lambda = 1$ doble con espacio propio de dimensión 1: no diagonalizable. El **teorema de la forma canónica de Jordan** garantiza (sobre $\mathbb{C}$) que toda matriz es semejante a una diagonal por bloques

$$J_k(\lambda) = \begin{pmatrix} \lambda & 1 & & \\ & \lambda & \ddots & \\ & & \ddots & 1 \\ & & & \lambda \end{pmatrix},$$

con tantos bloques por valor propio como su multiplicidad geométrica. Equivalentemente, $A = D + N$ con D diagonalizable, N nilpotente ($N^m = 0$) y $DN = ND$. Para las EDO bastará esto: los bloques de Jordan producen soluciones con factores polinomiales $t^k e^{\lambda t}$.

Regla: calcular la forma de Jordan de una matriz concreta

Cuando una matriz no se deja diagonalizar, la forma de Jordan es lo más cerca que se puede llegar. El cálculo es mecánico si se sigue el orden.

1. **Halla el polinomio característico y sus raíces.** Calcula $p(\lambda) = \det(A - \lambda I)$ y factorízalo. La multiplicidad de λ como raíz se llama *multiplicidad algebraica* $m_a(\lambda)$: es el tamaño total que ocupará λ en la matriz de Jordan.

2. **Calcula la multiplicidad geométrica.** Es $m_g(\lambda) = \dim \ker(A - \lambda I) = n - \operatorname{rango}(A - \lambda I)$. Este número dice *cuántos bloques* lleva ese autovalor, uno por cada vector propio independiente.

3. **Compara los dos números.** Si $m_g = m_a$ para todos los autovalores, la matriz es diagonalizable y ya terminaste. Si $m_g < m_a$ para alguno, ese autovalor tiene bloques de tamaño mayor que 1 y hay que seguir.

4. **Determina los tamaños con los rangos sucesivos.** El número de bloques de λ de tamaño $\geq k$ es
$$\operatorname{rango}(A - \lambda I)^{k-1} - \operatorname{rango}(A - \lambda I)^k.$$
Ve subiendo k hasta que el rango se estabilice. Como control: la suma de todos los tamaños de bloque de λ debe dar $m_a(\lambda)$.

5. **Escribe J.** Un bloque de tamaño k para λ es la matriz $k \times k$ con λ en la diagonal y unos en la superdiagonal. Coloca todos los bloques en diagonal. El orden de los bloques es convención; cualquier orden da una forma de Jordan válida.

6. **Construye las cadenas para hallar P.** Para cada bloque de tamaño k, busca un vector w con $(A-\lambda I)^k w = 0$ pero $(A-\lambda I)^{k-1} w \neq 0$, y forma la cadena

$$w, \quad (A-\lambda I)w, \quad (A-\lambda I)^2 w, \ \ldots$$

   Las columnas de P son estas cadenas escritas *del final al principio*: primero el vector propio, después sus antecesores.

7. **Verifica multiplicando.** Comprueba $AP = PJ$, que evita invertir P. Controles rápidos: la traza de J debe ser la de A y el determinante también.

Ejemplo 7.62 (Tres matrices que no se diagonalizan)**. (a) El caso 2×2 mínimo.** Sea $E = \begin{pmatrix} 1 & 1 \\ -1 & 3 \end{pmatrix}$. Su polinomio característico es $\lambda^2 - 4\lambda + 4 = (\lambda - 2)^2$, luego $m_a(2) = 2$. Ahora

$$E - 2I = \begin{pmatrix} -1 & 1 \\ -1 & 1 \end{pmatrix},$$

que tiene rango 1, así que $m_g(2) = 2 - 1 = 1$. Un solo bloque, de tamaño 2:

$$J = \begin{pmatrix} 2 & 1 \\ 0 & 2 \end{pmatrix}.$$

El vector propio sale de $-x + y = 0$, es decir $v = (1, 1)$. Para el generalizado buscamos w con $(E - 2I)w = v$: tomando $w = (-1, 0)$ se obtiene $(1, 1) = v$. Correcto. Entonces $P = \begin{pmatrix} 1 & -1 \\ 1 & 0 \end{pmatrix}$ y se comprueba $PJP^{-1} = E$.

(b) Una 3×3 con un autovalor repetido defectuoso. Sea

$$A = \begin{pmatrix} 5 & 4 & 3 \\ -1 & 0 & -3 \\ 1 & -2 & 1 \end{pmatrix}.$$

Su polinomio característico es $\lambda^3 - 6\lambda^2 + 32 = (\lambda + 2)(\lambda - 4)^2$, así que los autovalores son -2 (simple) y 4 (doble). Control: $\operatorname{tr} A = 5 + 0 + 1 = 6$ y $-2 + 4 + 4 = 6$; además $\det A = (-2) \cdot 4 \cdot 4 = -32$, coherente con el término independiente 32 del polinomio.

Para $\lambda = -2$ no hay nada que decidir: bloque de tamaño 1, con vector propio $(-1, 1, 1)$.

Para $\lambda = 4$:

$$A - 4I = \begin{pmatrix} 1 & 4 & 3 \\ -1 & -4 & -3 \\ 1 & -2 & -3 \end{pmatrix},$$

de rango 2, luego $m_g(4) = 3 - 2 = 1 < 2 = m_a(4)$: no diagonalizable, un solo bloque de tamaño 2. Por tanto

$$J = \begin{pmatrix} -2 & 0 & 0 \\ 0 & 4 & 1 \\ 0 & 0 & 4 \end{pmatrix}.$$

El vector propio de 4 es $v = (1, -1, 1)$. Para la cadena buscamos w con $(A-4I)w = v$; probando $w = (1, 0, 0)$ obtenemos la primera columna de $A - 4I$, que es justamente $(1, -1, 1) = v$. Así $P = \begin{pmatrix} -1 & 1 & 1 \\ 1 & -1 & 0 \\ 1 & 1 & 0 \end{pmatrix}$ y se verifica $PJP^{-1} = A$ exactamente.

(c) Cuando los rangos sucesivos son imprescindibles. Sea $C = \operatorname{diag}(J_2(3), 3, 4)$, es decir la matriz 4×4 con $J = \left(\begin{smallmatrix} 3 & 1 \\ 0 & 3 \end{smallmatrix}\right)$ arriba a la izquierda, un 3 suelto y un 4. Aquí $m_a(3) = 3$, y $\operatorname{rango}(C-3I) = 2$, de modo que $m_g(3) = 4-2 = 2$: dos bloques para el autovalor 3. Los tamaños salen del paso 4: bloques de tamaño ≥ 1 hay $\operatorname{rango}(C-3I)^0 - \operatorname{rango}(C-3I)^1 = 4-2 = 2$; de tamaño ≥ 2 hay $\operatorname{rango}(C-3I)^1 - \operatorname{rango}(C-3I)^2 = 2-1 = 1$; de tamaño ≥ 3 hay $1-1 = 0$. Conclusión: un bloque de tamaño 2 y uno de tamaño 1, que suman $3 = m_a$. Sin los rangos sucesivos, saber que hay dos bloques no bastaría para saber si son $2+1$ o algo distinto.

Cuidado

La multiplicidad geométrica cuenta *bloques*, no tamaños. Es el error más frecuente. Si $m_a(\lambda) = 5$ y $m_g(\lambda) = 2$, sabes que hay dos bloques que suman 5, pero eso deja dos posibilidades, $4+1$ y $3+2$, y solo los rangos de las potencias sucesivas las distinguen.

Idea clave

Un bloque de Jordan de tamaño k es $\lambda I + N$ con N nilpotente de orden k. Toda la teoría de Jordan se reduce a esta frase: cada autovalor aporta una parte escalar, que es la que se ve, y una parte nilpotente, que es la que causa problemas. Y la nilpotente desaparece al elevar: por eso $(A - \lambda I)^n = 0$ sobre el subespacio generalizado, y por eso el polinomio mínimo tiene, en cada λ, exponente igual al *mayor* tamaño de bloque.

7.7. Producto interno y ortogonalidad

Idea clave

Intuición primero. Hasta aquí un espacio vectorial no tenía geometría: se podía sumar y escalar, pero no había manera de decir cuánto mide un vector ni qué ángulo forman dos. El producto interno es exactamente el ingrediente que añade eso, y con una sola operación aparecen de golpe la longitud ($\|v\| = \sqrt{\langle v, v\rangle}$), el ángulo y, sobre todo, la perpendicularidad.

Lo importante es que "perpendicular" se vuelve calculable en sitios donde no hay nada que dibujar. Dos polinomios pueden ser ortogonales; dos funciones también. Y de ahí sale la técnica más útil de la sección: proyectar. Proyectar un vector sobre un subespacio es

encontrar el punto de ese subespacio más cercano a él, y eso es exactamente lo que hace el método de mínimos cuadrados cuando ajusta una recta a datos que no son colineales. Ajustar es proyectar.

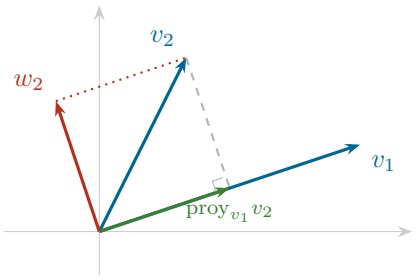


Figura 7.1: Un paso de Gram–Schmidt. Se proyecta v_2 sobre v_1 y se resta: lo que queda es perpendicular a v_1 por construcción. Con $v_1 = (3,1)$ y $v_2 = (1,2)$ sale proy $= (1{,}5;\ 0{,}5)$ y $w_2 = (-0{,}5;\ 1{,}5)$, cuyo producto escalar con v_1 es $-1{,}5 + 1{,}5 = 0$.

Definición 7.63. Un *producto interno* en un $\mathbb{R}$-espacio V es $\langle \cdot, \cdot \rangle$ bilineal, simétrico, definido positivo. Norma: $\|v\| = \sqrt{\langle v, v \rangle}$. Ejemplos: el estándar en $\mathbb{R}^n$; $\langle f, g \rangle = \int_a^b fg$ en $C[a,b]$ (el que hace funcionar las series de Fourier); $\langle A, B \rangle = \operatorname{tr}(A^{\mathsf{T}} B)$ en matrices. En espacios complejos se pide $\langle u, v \rangle = \overline{\langle v, u \rangle}$ y linealidad en una variable.

Teorema 7.64 (Cauchy–Schwarz y triangular)**.** $|\langle u, v \rangle| \le \|u\|\|v\|$ *y* $\|u + v\| \le \|u\| + \|v\|$.

Demostración. Para $t \in \mathbb{R}$: $0 \le \|u + tv\|^2 = \|u\|^2 + 2t\langle u, v \rangle + t^2 \|v\|^2$, un polinomio cuadrático en t sin dos raíces distintas: discriminante ≤ 0, que es Cauchy–Schwarz. La triangular sale de expandir $\|u + v\|^2$ y aplicar C–S. □

El *ángulo* entre vectores se define por $\cos\theta = \frac{\langle u, v \rangle}{\|u\|\|v\|}$; ortogonalidad es $\langle u, v \rangle = 0$, y entonces vale Pitágoras: $\|u + v\|^2 = \|u\|^2 + \|v\|^2$.

7.7.1. Gram–Schmidt, proyecciones y mínimos cuadrados

Definición 7.65. Una base es *ortonormal* si sus vectores son unitarios y ortogonales dos a dos. En base ortonormal, las coordenadas se calculan sin resolver sistemas: $v = \sum \langle v, e_i \rangle e_i$.

Gram–Schmidt: dado $\{v_1, \dots, v_k\}$ independiente,

$$u_j = v_j - \sum_{i<j} \frac{\langle v_j, u_i \rangle}{\langle u_i, u_i \rangle} u_i,$$

produce ortogonales con la misma envolvente; normalizando, ortonormales. En forma matricial, esto es la *factorización* $A = QR$ (Q de columnas ortonormales, R triangular superior), pilar de los algoritmos numéricos de valores propios.

Teorema 7.66 (Proyección ortogonal)**.** *Sea $W \leq V$ de dimensión finita. Todo v se descompone de forma única $v = w + w^{\perp}$ con $w \in W$, $w^{\perp} \in W^{\perp}$. Además $w = \mathrm{proy}_W(v)$ es el punto de W* más cercano *a v.*

Ejemplo 7.67 (Mínimos cuadrados)**.** Cuando $A\mathbf{x} = \mathbf{b}$ no tiene solución (típico: más ecuaciones que incógnitas, datos con ruido), se busca minimizar $\|A\mathbf{x} - \mathbf{b}\|$. El mínimo se alcanza cuando $A\mathbf{x}$ es la proyección de $\mathbf{b}$ sobre el espacio de columnas, es decir cuando el residuo es ortogonal a las columnas:

$$A^{\mathsf{T}} A\,\hat{\mathbf{x}} = A^{\mathsf{T}}\mathbf{b} \qquad \text{(ecuaciones normales).}$$

Recta de regresión: para ajustar $y = \alpha + \beta x$ a los puntos $(0,1), (1,3), (2,4)$:

$$A = \begin{pmatrix} 1 & 0 \\ 1 & 1 \\ 1 & 2 \end{pmatrix}, \quad A^{\mathsf{T}}A = \begin{pmatrix} 3 & 3 \\ 3 & 5 \end{pmatrix}, \quad A^{\mathsf{T}}\mathbf{b} = \begin{pmatrix} 8 \\ 11 \end{pmatrix} \Rightarrow \hat{\alpha} = \frac{7}{6},\ \hat{\beta} = \frac{3}{2}.$$

La recta $y = \frac{7}{6} + \frac{3}{2}x$ es la que minimiza la suma de los cuadrados de los errores verticales.

Regla: ajustar por mínimos cuadrados

Cuando un sistema $Ax = b$ no tiene solución — más ecuaciones que incógnitas, datos con ruido — la pregunta correcta ya no es "¿cuál es la solución?" sino "¿cuál es el x que menos se equivoca?". La respuesta es geométrica y siempre la misma: proyectar.

1. **Entiende qué se está pidiendo.** Ax recorre el espacio columna de A. Si b no está ahí, el sistema no tiene solución. Se busca entonces el punto de $\mathrm{Col}(A)$ más cercano a b, que es su proyección ortogonal $\hat{b}$, y se resuelve $A\hat{x} = \hat{b}$.

2. **Traduce "más cercano" a "residuo perpendicular".** El residuo $b - A\hat{x}$ debe ser ortogonal a todas las columnas de A, es decir $A^T(b - A\hat{x}) = 0$. Reordenando queda la *ecuación normal*

   $$A^T A\,\hat{x} = A^T b.$$

   Toda la teoría de mínimos cuadrados cabe en esa línea.

3. **Resuélvela.** Si las columnas de A son linealmente independientes, $A^T A$ es invertible y $\hat{x} = (A^T A)^{-1} A^T b$ es único. Si no lo son, hay infinitas soluciones óptimas y se elige la de norma mínima, que es la que da la pseudoinversa.

4. **Monta A correctamente para un ajuste de curvas.** Para ajustar $y = a + bx$ a los puntos (x_i, y_i), la matriz tiene una columna de unos y una columna de x_i; para $y = a + bx + cx^2$, se añade la columna de x_i^2. Las incógnitas son siempre los *coeficientes*, nunca los datos, y por eso el problema es lineal aunque el modelo sea una parábola.

5. **Mide la calidad con el residuo.** $\|b - A\hat{x}\|^2$ es la suma de los cuadrados de los errores, y es el número que el método minimiza. Si vale 0, el sistema sí tenía solución exacta.

6. **Si el problema está mal condicionado, no uses las ecuaciones normales.** A^TA eleva al cuadrado el número de condición. En la práctica numérica se usa la factorización QR (que es Gram–Schmidt hecho matriz) o la SVD. Para ejercicios a mano, las ecuaciones normales están bien.
7. **Control.** Comprueba que el residuo es ortogonal a cada columna de A: los productos escalares deben dar 0. Es el mismo control que en cualquier proyección.

Ejemplo 7.68 (La recta que mejor ajusta cuatro puntos)**.** Ajustemos $y = a + bx$ a los datos $(1,1)$, $(2,3)$, $(3,4)$, $(4,6)$. No hay ninguna recta que pase por los cuatro (los incrementos son $2, 1, 2$, no constantes), así que el sistema

$$\begin{pmatrix} 1 & 1 \\ 1 & 2 \\ 1 & 3 \\ 1 & 4 \end{pmatrix} \begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} 1 \\ 3 \\ 4 \\ 6 \end{pmatrix}$$

es incompatible: cuatro ecuaciones, dos incógnitas.

Ecuaciones normales. Se calcula

$$A^TA = \begin{pmatrix} 4 & 10 \\ 10 & 30 \end{pmatrix}, \qquad A^Tb = \begin{pmatrix} 14 \\ 43 \end{pmatrix}.$$

(El 4 es el número de puntos, el 10 es $\sum x_i$, el 30 es $\sum x_i^2$, el 14 es $\sum y_i$ y el 43 es $\sum x_iy_i$: las ecuaciones normales son exactamente las fórmulas de la regresión lineal que se ven en estadística, deducidas aquí sin derivar nada.) Resolviendo el sistema 2×2:

$$\hat{a} = -\tfrac{1}{2}, \qquad \hat{b} = \tfrac{8}{5}.$$

La recta de mejor ajuste es

$$y = -\tfrac{1}{2} + \tfrac{8}{5}x = -0{,}5 + 1{,}6\,x.$$

Residuos. Los valores predichos son 1,1, 2,7, 4,3 y 5,9, luego los residuos son

$$-0{,}1, \quad +0{,}3, \quad -0{,}3, \quad +0{,}1,$$

y la suma de sus cuadrados vale $\frac{1}{100} + \frac{9}{100} + \frac{9}{100} + \frac{1}{100} = \frac{20}{100} = \frac{1}{5} = 0{,}2$. Ninguna otra recta consigue bajar de 0,2; esa es la afirmación exacta que hace el método.

Control del paso 7. El residuo es $r = (-0{,}1,\ 0{,}3,\ -0{,}3,\ 0{,}1)$. Su producto escalar con la primera columna de A (los unos) es $-0{,}1 + 0{,}3 - 0{,}3 + 0{,}1 = 0$ ✓, y con la segunda columna $(1, 2, 3, 4)$ es $-0{,}1 + 0{,}6 - 0{,}9 + 0{,}4 = 0$ ✓. Perpendicular a ambas, como exige la geometría.

Idea clave

Obsérvese lo que acaba de ocurrir: la fórmula de la regresión lineal, que en estadística se presenta como un resultado de cálculo diferencial (derivar la suma de cuadrados e igualar a cero), aquí ha salido de una sola frase geométrica — *el error debe ser perpendicular a lo que se puede alcanzar* — sin derivar absolutamente nada. Esa es la ventaja de mirar el problema en el espacio de columnas.

7.7.2. Matrices ortogonales y el teorema espectral

Definición 7.69. Q es *ortogonal* si $Q^{\mathsf{T}}Q = I$ (columnas ortonormales). Conserva productos internos, normas y ángulos: son las rotaciones y reflexiones. $\det Q = \pm 1$.

Idea clave

Qué dice esto en cristiano. Una matriz ortogonal es la matriz de un movimiento rígido: gira o refleja, pero no estira, no encoge y no deforma ángulos. La condición $Q^{\mathsf{T}}Q = I$ dice exactamente que las columnas son vectores unitarios perpendiculares entre sí, es decir, que la matriz manda una base ortonormal a otra base ortonormal. Su gran ventaja práctica es que la inversa es gratis: $Q^{-1} = Q^{\mathsf{T}}$, sin resolver ningún sistema. Por eso todos los algoritmos numéricos serios se apoyan en ellas.

Ejemplo 7.70 (Las dos familias y lo que conservan)**. (a) Una rotación.** Con $\theta = 30°$, la matriz $Q = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}$ cumple $Q^{\mathsf{T}}Q = I$ y $\det Q = \cos^2\theta + \sin^2\theta = 1$. El determinante positivo es la firma de que conserva la orientación.

(b) Una reflexión. La matriz $\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$ también cumple $Q^{\mathsf{T}}Q = I$, pero su determinante es -1: invierte la orientación. En dimensión dos, toda matriz ortogonal es una de las dos cosas, y el determinante decide cuál.

(c) Conservan las normas. Para $\mathbf{v} = (3, 4)$, cuya norma es 5, la imagen $Q\mathbf{v}$ vuelve a tener norma 5, sea cual sea el ángulo. Es consecuencia directa de $\|Q\mathbf{v}\|^2 = \mathbf{v}^{\mathsf{T}}Q^{\mathsf{T}}Q\mathbf{v} = \mathbf{v}^{\mathsf{T}}\mathbf{v}$: el $Q^{\mathsf{T}}Q$ se convierte en la identidad y desaparece.

(d) Una que no parece ortogonal y lo es. La matriz $\frac{1}{\sqrt{2}}\begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}$ cumple $Q^{\mathsf{T}}Q = I$. Es la matriz de Hadamard normalizada, y es la base de la transformada rápida de Walsh–Hadamard. El factor $\frac{1}{\sqrt{2}}$ es imprescindible: sin él las columnas tendrían norma $\sqrt{2}$ y no serían unitarias.

Cuidado

[Ortogonal exige columnas *ortonormales*, no solo perpendiculares] La matriz $\begin{pmatrix} 2 & 0 \\ 0 & 3 \end{pmatrix}$ tiene columnas perpendiculares y no es ortogonal, porque no son unitarias: $Q^{\mathsf{T}}Q$ da

$\begin{pmatrix} 4 & 0 \\ 0 & 9 \end{pmatrix}$ y no la identidad. El nombre es desafortunado y confunde a todo el mundo; algunos textos las llaman *ortonormales*, que sería lo correcto. Segundo aviso: $\det Q = \pm 1$ es condición necesaria y no suficiente; $\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$ tiene determinante 1 y deforma. Y tercero: sobre $\mathbb{C}$ el análogo son las matrices *unitarias*, con $U^*U = I$ usando el conjugado traspuesto, no el traspuesto a secas.

Teorema 7.71 (Espectral, caso real simétrico)**.** *Si $A = A^{\mathsf{T}}$, todos sus valores propios son reales, espacios propios de valores propios distintos son ortogonales, y existe Q ortogonal con*

$$Q^{\mathsf{T}} A Q = \operatorname{diag}(\lambda_1, \ldots, \lambda_n).$$

Equivalentemente, $A = \sum_i \lambda_i \, q_i q_i^{\mathsf{T}}$: suma de proyecciones sobre ejes ortogonales.

Demostración de las dos primeras afirmaciones. Si $Av = \lambda v$ con $v \neq 0$ (quizá complejos), entonces $\bar{v}^{\mathsf{T}} A v = \lambda \|v\|^2$; tomando conjugada transpuesta y usando $A^{\mathsf{T}} = A$ real: $\bar{\lambda}\|v\|^2 = \lambda\|v\|^2$, luego $\lambda \in \mathbb{R}$. Si $Av = \lambda v$, $Aw = \mu w$ con $\lambda \neq \mu$: $\lambda\langle v, w\rangle = \langle Av, w\rangle = \langle v, Aw\rangle = \mu\langle v, w\rangle$, de donde $\langle v, w\rangle = 0$. La existencia de la base ortonormal completa se prueba por inducción, restringiendo A al complemento ortogonal de un vector propio. □

7.8. Formas cuadráticas

Idea clave

Intuición primero. Una forma cuadrática es lo que sale al elevar al cuadrado en varias variables: $q(x) = x^T A x$, con A simétrica. Aparece en tres sitios reconocibles: en la energía de un sistema físico, en el término de segundo orden del desarrollo de Taylor multivariable (y por eso decide si un punto crítico es máximo o mínimo, como se verá en el capítulo 9) y en las cónicas y cuádricas de la geometría analítica.
La pregunta central es el *signo*: si q es siempre positiva, siempre negativa, o cambia. Y el teorema espectral la contesta de forma limpia y definitiva: en la base de vectores propios la forma se vuelve una suma de cuadrados con los valores propios como coeficientes, y entonces el signo se lee mirando los signos de esos valores propios. Los términos cruzados que ensucian la expresión no eran esenciales: eran un artefacto de haber elegido mal los ejes.

Definición 7.72. Una *forma cuadrática* es $q(\mathbf{x}) = \mathbf{x}^{\mathsf{T}} A \mathbf{x}$ con A simétrica. Es *definida positiva* si $q > 0$ fuera del origen; *semidefinida* si $q \geq 0$; *indefinida* si toma ambos signos.

Teorema 7.73 (Clasificación)**.** *Diagonalizando ortogonalmente ($\mathbf{x} = Q\mathbf{y}$): $q = \sum \lambda_i y_i^2$. Por tanto: definida positiva $\iff$ todos los $\lambda_i > 0$ $\iff$ (criterio de Sylvester) todos los menores principales dominantes son positivos. Indefinida $\iff$ valores propios de ambos signos. Los*

*números de coeficientes positivos y negativos no dependen de la diagonalización (*ley de inercia de Sylvester*).*

Ejemplo 7.74 (Cónicas)**.** La ecuación $5x^2-4xy+8y^2=36$ tiene matriz $A=\begin{pmatrix}5&-2\\-2&8\end{pmatrix}$, con valores propios 4 y 9 (raíces de $\lambda^2-13\lambda+36$). En los ejes propios la ecuación es $4x'^2+9y'^2=36$, es decir $\frac{x'^2}{9}+\frac{y'^2}{4}=1$: una elipse de semiejes 3 y 2, rotada. Diagonalizar la forma cuadrática *es* clasificar la cónica.

Regla: clasificar una forma cuadrática

Dada $q(\mathbf{x})=\mathbf{x}^\top A\,\mathbf{x}$ con A simétrica, la pregunta es siempre la misma: ¿qué signo toma q?

1. **Escribe la matriz simétrica.** Los coeficientes de x_i^2 van a la diagonal tal cual; los de x_ix_j con $i\neq j$ se reparten *a la mitad* entre las posiciones (i,j) y (j,i). Si te saltas la mitad, todo lo demás sale mal.
2. **Decide qué criterio usar.** Hay dos, y son equivalentes: el de los autovalores (siempre funciona) y el de los menores principales (más rápido, pero solo concluyente en los casos definidos).
3. **Criterio de los autovalores.** Como A es simétrica, el teorema espectral garantiza que todos sus autovalores son reales. Entonces q es definida positiva si todos son >0; definida negativa si todos son <0; semidefinida si alguno es 0 y el resto comparte signo; e *indefinida* si hay autovalores de ambos signos.
4. **Criterio de Sylvester (menores principales líderes).** Sean Δ_k los determinantes de las submatrices $k\times k$ de la esquina superior izquierda. Entonces q es definida positiva si y solo si $\Delta_1,\dots,\Delta_n>0$; y definida negativa si y solo si los signos alternan empezando por $\Delta_1<0$, es decir $(-1)^k\Delta_k>0$.
5. **No abuses de Sylvester.** Si algún $\Delta_k=0$, el criterio de los menores líderes *no decide*: hay que pasar a los autovalores o a todos los menores principales, no solo a los líderes.
6. **Reduce por Lagrange si quieres verlo.** Completando cuadrados escribes q como suma de cuadrados con signos. El número de signos $+$, de $-$ y de ceros no depende del método usado: es la *ley de inercia* de Sylvester, y esos tres números son la firma de q.
7. **Controla con la traza y el determinante.** $\operatorname{tr}A$ es la suma de autovalores y $\det A$ su producto. Si $\det A<0$ en dimensión 2, la forma es indefinida sin más cálculo.

Ejemplo 7.75 (Cinco formas clasificadas por dos caminos)**. (a)** $q(x,y)=2x^2+2xy+2y^2$**.** La matriz es $A=\left(\begin{smallmatrix}2&1\\1&2\end{smallmatrix}\right)$: el $2xy$ se reparte como 1 y 1. Menores: $\Delta_1=2>0$, $\Delta_2=4-1=3>0$, luego definida positiva. Autovalores: $\lambda=3$ y $\lambda=1$, ambos positivos. Los dos caminos coinciden.

Geométricamente, $q = 1$ es una elipse.

(b) $q(x, y) = x^2 + 4xy + y^2$**.** Ahora $A = \left(\begin{smallmatrix} 1 & 2 \\ 2 & 1 \end{smallmatrix}\right)$. Menores: $\Delta_1 = 1 > 0$ pero $\Delta_2 = 1 - 4 = -3 < 0$: no es definida positiva. Autovalores: 3 y -1, de signos opuestos, luego *indefinida*. La curva $q = 1$ es una hipérbola. Nótese que el determinante negativo bastaba para concluir.

(c) $q(x, y) = x^2 - 2xy + y^2$**.** Aquí $A = \left(\begin{smallmatrix} 1 & -1 \\ -1 & 1 \end{smallmatrix}\right)$, con $\Delta_1 = 1 > 0$ y $\Delta_2 = 1 - 1 = 0$. Sylvester no decide. Los autovalores son 2 y 0: semidefinida positiva. Y en efecto $q = (x - y)^2 \geq 0$, con igualdad en toda la recta $x = y$. Este es exactamente el caso que el paso 5 advertía.

(d) Una 3×3 definida positiva. Sea $q(x, y, z) = 2x^2 + 2y^2 + 2z^2 - 2xy - 2yz$, con

$$A = \begin{pmatrix} 2 & -1 & 0 \\ -1 & 2 & -1 \\ 0 & -1 & 2 \end{pmatrix}.$$

Menores líderes: $\Delta_1 = 2$, $\Delta_2 = 4 - 1 = 3$, $\Delta_3 = 2(4 - 1) - (-1)(-2 - 0) = 6 - 2 = 4$. Los tres positivos: definida positiva. Autovalores: 2 y $2 \pm \sqrt{2}$, es decir 0,5858, 2 y 3,4142, todos positivos. Control: su suma es $6 = \operatorname{tr} A$ y su producto $4 = \det A$. Esta matriz es el laplaciano discreto en una dimensión, y su positividad es lo que hace estable el método numérico correspondiente.

(e) Cuando Sylvester engaña si se lee mal. Sea $A = \left(\begin{smallmatrix} 1 & 2 & 0 \\ 2 & 1 & 0 \\ 0 & 0 & 3 \end{smallmatrix}\right)$. Menores líderes: $\Delta_1 = 1 > 0$, $\Delta_2 = 1 - 4 = -3 < 0$, $\Delta_3 = 3 \cdot (-3) = -9 < 0$. Los signos ni son todos positivos ni alternan: indefinida. Autovalores 3, 3 y -1; firma $(2, 1, 0)$, o sea dos cuadrados positivos y uno negativo. Completando cuadrados: $q = (x + 2y)^2 - 3y^2 + 3z^2$, donde se leen directamente los signos $+, -, +$.

Cuidado

El criterio de los menores *líderes* solo caracteriza los casos definidos. Para semidefinida positiva no basta con $\Delta_k \geq 0$: hace falta que *todos* los menores principales (no solo los de la esquina) sean ≥ 0. Contraejemplo clásico: $\left(\begin{smallmatrix} 0 & 0 \\ 0 & -1 \end{smallmatrix}\right)$ tiene $\Delta_1 = 0$ y $\Delta_2 = 0$, ambos ≥ 0, y sin embargo la forma $-y^2$ es semidefinida *negativa*.

Pausa para pensar

¿Por qué la ley de inercia es un teorema y no una obviedad? Porque uno puede reducir la misma forma a suma de cuadrados de infinitas maneras distintas, con cambios de variable diferentes, y a priori nada impide que dos reducciones den distinto número de signos positivos. El teorema afirma que no: la firma es un invariante de la forma, no del método. Es el análogo, para formas cuadráticas, de que la dimensión de un espacio no dependa de la base elegida.

7.9. Descomposición en valores singulares (SVD)

Idea clave

Intuición primero. La SVD dice algo asombrosamente general: *toda* matriz, sea cuadrada o no, invertible o no, se descompone en tres pasos elementales —una rotación, un estirado a lo largo de ejes perpendiculares, y otra rotación—. Eso es lo que significa $A = U\Sigma V^T$. Geométricamente: cualquier transformación lineal convierte la esfera unitaria en un elipsoide, y los valores singulares son las longitudes de los semiejes de ese elipsoide.
De ahí sale su utilidad práctica. Los valores singulares vienen ordenados de mayor a menor, y quedarse con los primeros k da la mejor aproximación posible de rango k —el teorema de Eckart–Young—. Eso es, literalmente, lo que hace la compresión de imágenes, el análisis de componentes principales y buena parte de los sistemas de recomendación: quedarse con las direcciones donde de verdad hay información y tirar el resto. Donde la diagonalización solo funciona para matrices especiales, la SVD funciona siempre; por eso es la herramienta que de verdad se usa fuera de los libros.

No toda matriz es cuadrada ni simétrica, pero *toda* matriz admite la factorización más importante del álgebra lineal aplicada:

Teorema 7.76 (SVD)**.** *Toda $A \in M_{m\times n}$ se factoriza $A = U\Sigma V^{\mathsf{T}}$ con U $(m\times m)$ y V $(n\times n)$ ortogonales y Σ diagonal rectangular con entradas $\sigma_1 \geq \sigma_2 \geq \cdots \geq 0$ (los* valores singulares*, raíces cuadradas de los valores propios de $A^{\mathsf{T}}A$).*

Idea clave

Qué dice esto en cristiano. La descomposición en valores singulares dice que *cualquier* matriz, sea cuadrada o no, invertible o no, hace siempre lo mismo en tres pasos: gira (o refleja), estira cada eje por un factor, y vuelve a girar. Los factores de estiramiento son los valores singulares, y son números no negativos ordenados de mayor a menor. Es la factorización más general y más útil del álgebra lineal: donde la diagonalización exige matrices cuadradas y bien portadas, la SVD no exige nada. De ella salen el rango, el número de condición, la pseudoinversa y la mejor aproximación de rango bajo.

Ejemplo 7.77 (Leer una matriz a través de sus valores singulares)**. (a) Un caso concreto.** Sea $A = \begin{pmatrix} 3 & 0 \\ 4 & 5 \end{pmatrix}$. Sus valores singulares son 6,708204 y 2,236068. Su producto es 15, exactamente $|\det A|$: los factores de estiramiento multiplicados dan el factor de cambio de volumen.

(b) De dónde salen. Son las raíces cuadradas de los autovalores de $A^{\mathsf{T}}A$, que aquí valen 45 y 5; en efecto $\sqrt{45} = 6{,}708204$ y $\sqrt{5} = 2{,}236068$. Esa matriz $A^{\mathsf{T}}A$ es siempre simétrica y semidefinida positiva, de modo que sus autovalores son reales y no negativos, y la raíz tiene sentido.

(c) El número de condición. Es el cociente entre el mayor y el menor valor singular:

aquí $\frac{6{,}708204}{2{,}236068} = 3$, un valor excelente. Compárese con $\begin{pmatrix} 1 & 1 \\ 1 & 1{,}0001 \end{pmatrix}$, cuyos valores singulares son 2,00005 y 0,00005, con cociente 40 002: casi singular.

(d) Funciona con matrices rectangulares. Para $\begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix}$, de tamaño 2×3, los valores singulares son 9,508032 y 0,772870, ambos no nulos, luego el rango es 2. Contar valores singulares no nulos es la manera numéricamente fiable de calcular el rango, mucho más que mirar determinantes de submatrices.

Cuidado

[Valores singulares no son autovalores] Para $\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$ los autovalores son 0 y 0, mientras que los valores singulares son 1 y 0. Coinciden solo cuando la matriz es simétrica y semidefinida positiva. Segundo aviso: los valores singulares son siempre reales y no negativos, aunque los autovalores sean complejos; por eso la SVD existe para toda matriz real y la diagonalización no. Y tercero: U y V no son únicas cuando hay valores singulares repetidos, así que "la" SVD es en realidad una familia; lo único unívoco son los valores singulares.

Lectura geométrica: toda transformación lineal es *rotación–escalado–rotación*. El número de σ_i no nulos es el rango; truncar la suma $A = \sum \sigma_i u_i v_i^{\mathsf{T}}$ a los k mayores da la mejor aproximación de rango k (Eckart–Young), base de la compresión de datos y del análisis de componentes principales.

Regla: calcular una descomposición en valores singulares a mano

La SVD escribe cualquier matriz A de tamaño $m \times n$ como $A = U\Sigma V^{\top}$, con U y V ortogonales y Σ diagonal no negativa. A mano se construye siempre desde $A^{\top}A$.

1. **Forma $A^{\top}A$.** Es una matriz $n \times n$ simétrica y semidefinida positiva, así que sus autovalores son reales y ≥ 0. Si $n > m$ conviene trabajar con la más pequeña de $A^{\top}A$ y $AA^{\top}$: tienen los mismos autovalores no nulos.

2. **Diagonaliza $A^{\top}A$.** Halla sus autovalores $\lambda_1 \geq \cdots \geq \lambda_n \geq 0$ y una base *ortonormal* de autovectores $\mathbf{v}_1, \ldots, \mathbf{v}_n$. El teorema espectral garantiza que existe. Estos son los *vectores singulares por la derecha* y forman las columnas de V.

3. **Extrae los valores singulares.** Define $\sigma_i = \sqrt{\lambda_i}$, siempre con el signo positivo, y ordénalos de mayor a menor. El número de σ_i no nulos es exactamente $\operatorname{rango}(A)$.

4. **Genera las columnas de U con la fórmula.** Para cada $\sigma_i > 0$ pon
$$\mathbf{u}_i = \frac{1}{\sigma_i} A\mathbf{v}_i.$$
Estos vectores salen automáticamente ortonormales; no hay que normalizarlos otra vez, y si no salen unitarios es que hubo un error de cuenta antes.

5. **Completa U si hace falta.** Si $\operatorname{rango}(A) = r < m$, faltan $m - r$ columnas: complétalas con cualquier base ortonormal del complemento ortogonal de $\{\mathbf{u}_1, \dots, \mathbf{u}_r\}$, que es precisamente $\ker(A^\top)$.
6. **Ensambla y comprueba.** Escribe Σ de tamaño $m \times n$ con los σ_i en la diagonal y ceros en el resto, y verifica $A = U\Sigma V^\top$ multiplicando. Control barato: $\sum \sigma_i^2 = \sum_{ij} a_{ij}^2$, la norma de Frobenius al cuadrado.
7. **Lee la información geométrica.** $\sigma_1 = \|A\|_2$ es el máximo estiramiento; σ_1/σ_r es el número de condición; las primeras columnas de U y V dan la mejor aproximación de rango bajo por el teorema de Eckart–Young.

Ejemplo 7.78 (Dos descomposiciones completas)**. (a) Una matriz cuadrada:** $A = \left(\begin{smallmatrix} 3 & 0 \\ 4 & 5 \end{smallmatrix}\right)$**.** Primero

$$A^\top A = \begin{pmatrix} 3 & 4 \\ 0 & 5 \end{pmatrix} \begin{pmatrix} 3 & 0 \\ 4 & 5 \end{pmatrix} = \begin{pmatrix} 25 & 20 \\ 20 & 25 \end{pmatrix}.$$

Sus autovalores son 25 ± 20, es decir 45 y 5, con autovectores $(1,1)$ y $(1,-1)$ (los de toda matriz de la forma $aI + bJ$). Normalizando,

$$\mathbf{v}_1 = \tfrac{1}{\sqrt{2}}(1,1), \qquad \mathbf{v}_2 = \tfrac{1}{\sqrt{2}}(1,-1),$$

y los valores singulares son $\sigma_1 = \sqrt{45} = 3\sqrt{5} \approx 6{,}7082$ y $\sigma_2 = \sqrt{5} \approx 2{,}2361$.

Ahora las columnas de U:

$$\mathbf{u}_1 = \frac{1}{3\sqrt{5}} A\mathbf{v}_1 = \frac{1}{\sqrt{10}}(1,3), \qquad \mathbf{u}_2 = \frac{1}{\sqrt{5}} A\mathbf{v}_2 = \frac{1}{\sqrt{10}}(3,-1).$$

Se comprueba que son unitarios y ortogonales entre sí, como prometía el paso 4. Control de Frobenius: $\sigma_1^2 + \sigma_2^2 = 45 + 5 = 50$ y $3^2 + 0^2 + 4^2 + 5^2 = 9 + 16 + 25 = 50$. Coincide.

Obsérvese además que $\det A = 15$ y $\sigma_1\sigma_2 = 3\sqrt{5} \cdot \sqrt{5} = 15$: para matrices cuadradas, el producto de los valores singulares es el valor absoluto del determinante.

(b) Una matriz rectangular alta: $C = \left(\begin{smallmatrix} 1 & 1 \\ 0 & 1 \\ 1 & 0 \end{smallmatrix}\right)$**.** Aquí $m = 3$, $n = 2$, así que conviene $C^\top C$, que es 2×2:

$$C^\top C = \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix},$$

con autovalores 3 y 1 y los mismos autovectores $(1,1)$ y $(1,-1)$ normalizados. Luego $\sigma_1 = \sqrt{3} \approx 1{,}7321$ y $\sigma_2 = 1$. Como hay dos valores singulares no nulos, $\operatorname{rango}(C) = 2$, lo que era de esperar porque las dos columnas son independientes.

Las columnas de U:

$$\mathbf{u}_1 = \frac{1}{\sqrt{3}} C\mathbf{v}_1 = \frac{1}{\sqrt{3}\cdot\sqrt{2}}(2,1,1) = \frac{1}{\sqrt{6}}(2,1,1), \qquad \mathbf{u}_2 = \frac{1}{1} C\mathbf{v}_2 = \frac{1}{\sqrt{2}}(0,-1,1).$$

Falta una tercera columna para completar U a una matriz 3×3 ortogonal: hay que tomar un vector unitario ortogonal a los dos anteriores, por ejemplo $\mathbf{u}_3=\frac{1}{\sqrt{3}}(-1,1,1)$, que genera $\ker(C^\top)$. Entonces

$$\Sigma=\begin{pmatrix}\sqrt{3} & 0\\ 0 & 1\\ 0 & 0\end{pmatrix},$$

rectangular igual que C, con la fila de ceros que corresponde a la dimensión sobrante. Control de Frobenius: $3+1=4$ y la suma de los cuadrados de las entradas de C es $1+1+1+1=4$.

(c) Qué se lee de todo esto. En (a), el número de condición es $\sigma_1/\sigma_2=3\sqrt{5}/\sqrt{5}=3$: la matriz es benigna numéricamente. La mejor aproximación de rango 1 a A es $\sigma_1\mathbf{u}_1\mathbf{v}_1^\top$, y el error cometido en norma espectral es exactamente $\sigma_2=2{,}2361$; en norma de Frobenius, $\sqrt{\sigma_2^2}=2{,}2361$ también, porque solo se descarta un valor singular.

Cuidado

Los valores singulares no son los autovalores, salvo cuando A es simétrica semidefinida positiva. En (a), los autovalores de A son 3 y 5 (la matriz es triangular), mientras que sus valores singulares son 6,7082 y 2,2361. Ni siquiera coinciden los conjuntos. La relación correcta es $\sigma_i(A)=\sqrt{\lambda_i(A^\top A)}$, y para A simétrica se reduce a $\sigma_i=|\lambda_i|$.

Idea clave

La SVD dice que *toda* aplicación lineal es, en las bases adecuadas, un simple reescalado de ejes. Se gira el espacio de partida con $V^\top$, se estiran los ejes por $\sigma_1,\dots,\sigma_r$, se aplasta lo que sobra, y se gira el resultado con U. La imagen de la esfera unidad es siempre un elipsoide (posiblemente degenerado) cuyos semiejes son los σ_i y cuyas direcciones son los $\mathbf{u}_i$. Que esto valga para cualquier matriz, cuadrada o no, invertible o no, real o compleja, es lo que hace de la SVD la herramienta más usada del álgebra lineal aplicada.

7.10. Fundamentos: las demostraciones estructurales

Esta sección demuestra los teoremas que antes solo enunciamos, y desarrolla la teoría fina de operadores que los sustenta.

7.10.1. Existencia de bases y el lema de intercambio

Lema 7.79 (Intercambio de Steinitz)**.** *Sea $\{v_1,\dots,v_m\}$ un generador de V y $\{w_1,\dots,w_k\}$ un conjunto independiente. Entonces $k\le m$, y se pueden reemplazar k de los v_i por los w_j manteniendo un generador.*

Demostración. Inducción en k. Agregamos w_1 al generador: como $w_1=\sum c_i v_i$ con algún $c_i\neq 0$ (pues $w_1\neq 0$), podemos despejar ese v_i y eliminarlo: $\{w_1,v_1,\dots,\hat{v}_i,\dots,v_m\}$ sigue generando. Repetimos con w_2: al escribirlo en el nuevo generador, algún coeficiente sobre los v restantes es no nulo (si solo dependiera de w_1, contradiría la independencia de los w), y

despejamos otro v. Tras k pasos hemos insertado todos los w_j; si fuera $k > m$, los v se habrían agotado antes y algún w dependería de los anteriores: absurdo. □

Idea clave

Qué dice esto en cristiano. El lema de intercambio es el motor que hace funcionar toda la teoría de la dimensión. Dice que un conjunto independiente nunca puede tener más elementos que un generador, porque cada vector independiente se puede colar en el generador expulsando a uno de los suyos. De ahí sale de golpe que todas las bases de un espacio tienen el mismo número de elementos, que es exactamente lo que permite *definir* la dimensión sin ambigüedad. Sin este lema, la palabra dimensión no significaría nada.

Ejemplo 7.80 (Intercambiar vectores paso a paso)**. (a) El intercambio en $\mathbb{R}^2$.** Sea el generador $\{(1,0),(0,1)\}$ y el independiente $\{(1,1)\}$. Como $(1,1) = 1\cdot(1,0) + 1\cdot(0,1)$ y el coeficiente de $(1,0)$ no es cero, se puede expulsar a $(1,0)$: el conjunto $\{(1,1),(0,1)\}$ sigue generando $\mathbb{R}^2$. Comprobación: $(1,0) = (1,1) - (0,1)$, luego se recupera lo perdido.

(b) La cota que impone. En $\mathbb{R}^3$, que está generado por tres vectores, no puede haber cuatro vectores independientes. El lema lo demuestra: cada uno de los cuatro exigiría expulsar a uno de los tres, y a la cuarta ronda ya no quedaría a quién expulsar. La conclusión es $k \le m$, es decir, independiente $\le$ generador.

(c) La consecuencia: todas las bases miden lo mismo. Si B_1 y B_2 son bases, entonces B_1 es independiente y B_2 generador, luego $|B_1| \le |B_2|$; e intercambiando los papeles, $|B_2| \le |B_1|$. Por tanto son iguales, y ese número común es la dimensión.

(d) Comprobación en un caso concreto. El espacio de las matrices 2×2 tiene la base evidente de cuatro matrices con un 1 y el resto ceros, luego dimensión 4. También

$$\left\{\begin{pmatrix}1&0\\0&1\end{pmatrix},\begin{pmatrix}1&0\\0&-1\end{pmatrix},\begin{pmatrix}0&1\\1&0\end{pmatrix},\begin{pmatrix}0&1\\-1&0\end{pmatrix}\right\}$$

es base, y tiene cuatro elementos. El lema garantiza que ninguna base puede tener tres ni cinco.

Cuidado

[Independiente y generador no se pueden confundir] Un conjunto puede ser independiente sin generar: $\{(1,0)\}$ en $\mathbb{R}^2$. Y puede generar sin ser independiente: $\{(1,0),(0,1),(1,1)\}$. Base es el punto exacto donde ambas cosas ocurren a la vez, y por el lema es también el punto donde el número de elementos es simultáneamente el máximo de un independiente y el mínimo de un generador. Segundo aviso: en dimensión infinita el lema falla como herramienta práctica y el concepto de base cambia; en un espacio de Hilbert, la "base" habitual permite sumas infinitas, cosa que aquí está prohibida.

De aquí, en tres líneas: dos bases cualesquiera tienen el mismo tamaño (cada una es independiente y la otra genera: $k \le m$ y $m \le k$); todo independiente se extiende a base (se insertan sus vectores en cualquier generador); de todo generador se extrae una base (se

descartan los dependientes uno a uno). La *dimensión* queda bien definida y con ella todo el edificio.

7.10.2. El determinante: existencia y unicidad

Teorema 7.81. *Existe exactamente una función* $\det : M_n(K) \to K$ *multilineal y alternada en las filas con* $\det I = 1$*, y es la fórmula de Leibniz* $\det A = \sum_{\sigma \in S_n} \operatorname{sgn}(\sigma) \prod_i A_{i\sigma(i)}$.

Demostración (unicidad y existencia). Unicidad: escribamos cada fila en la base canónica, $F_i = \sum_j A_{ij} e_j$, y expandamos por multilinealidad: $\det A = \sum_{j_1,\dots,j_n} A_{1j_1} \cdots A_{nj_n} \det(e_{j_1}, \dots, e_{j_n})$. Por alternancia, los términos con índices repetidos mueren; quedan solo las permutaciones σ, y $\det(e_{\sigma(1)}, \dots, e_{\sigma(n)}) = \operatorname{sgn}(\sigma)$ (cada trasposición de filas cambia el signo, y se necesitan tantas como la paridad de σ; aquí se usa que el signo está bien definido, Teorema del capítulo 25). Con $\det I = 1$, la fórmula queda forzada. *Existencia*: se verifica directamente que la fórmula de Leibniz es multilineal (cada sumando es lineal en cada fila), alternada (intercambiar dos filas biyecta $\sigma \leftrightarrow \sigma \circ (\text{trasposición})$, cambiando todos los signos) y vale 1 en I. □

La multiplicatividad $\det(AB) = \det A \det B$ sale gratis de la unicidad: la función $B \mapsto \det(AB)$ es multilineal alternada en las filas de B con valor $\det A$ en $B = I$; por unicidad debe ser $\det A \cdot \det B$. Este estilo de argumento —"caracterizar y comparar"— es marca de la casa del álgebra lineal abstracta.

7.10.3. Polinomio mínimo y descomposición primaria

Definición 7.82. El *polinomio mínimo* m_A de una matriz A es el mónico de menor grado con $m_A(A) = 0$ (existe: las potencias $I, A, A^2, \dots$ viven en un espacio de dimensión n^2 y deben ser dependientes). Divide a todo polinomio que anule a A —en particular, por Cayley–Hamilton, $m_A \mid p_A$— y tiene las mismas raíces que p_A: los valores propios.

Teorema 7.83 (Criterio limpio de diagonalización)**.** *A es diagonalizable $\iff$ m_A no tiene raíces repetidas ($m_A = \prod(\lambda - \lambda_i)$ con λ_i distintos).*

Idea clave

Qué dice esto en cristiano. El polinomio característico dice *cuáles* son los autovalores y con qué multiplicidad aparecen en el determinante. El polinomio mínimo dice algo más fino: cuántas veces hay que aplicar $(A - \lambda I)$ para aniquilar de verdad ese autovalor. Y ahí está el criterio: si con una sola vez basta para todos, es decir si el mínimo no tiene raíces repetidas, la matriz diagonaliza. Si hace falta aplicarlo dos veces para algún autovalor, es que faltan autovectores y aparece un bloque de Jordan.

Ejemplo 7.84 (Tres matrices y sus dos polinomios)**. (a) Una diagonal.** $A = \begin{pmatrix} 2 & 0 \\ 0 & 3 \end{pmatrix}$ tiene característico $(\lambda - 2)(\lambda - 3)$ y mínimo igual. Sin raíces repetidas, luego diagonalizable, cosa que ya era evidente.

(b) Un bloque de Jordan. $B = \begin{pmatrix} 2 & 1 \\ 0 & 2 \end{pmatrix}$ tiene característico $(\lambda-2)^2$, con el autovalor 2 de multiplicidad algebraica 2. Pero al resolver $(B-2I)\mathbf{v}=\mathbf{0}$ sale un único autovector independiente, $(1,0)$: la multiplicidad geométrica es 1. El mínimo es $(\lambda-2)^2$, con raíz repetida, y por tanto B *no* diagonaliza. Falta un autovector.

(c) La rotación de noventa grados. $C = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$ tiene característico λ^2+1, sin raíces reales: sobre $\mathbb{R}$ no diagonaliza, y era esperable, porque una rotación no deja ninguna dirección fija. Sobre $\mathbb{C}$, en cambio, sus autovalores son i y $-$i, distintos, y sí diagonaliza. El cuerpo importa.

(d) El criterio rápido. Si el característico tiene n raíces distintas, la matriz diagonaliza sin más comprobaciones, porque autovectores de autovalores distintos son siempre independientes. Solo cuando hay raíces repetidas hay que mirar el mínimo o contar autovectores, como en (b).

Cuidado

[Multiplicidad algebraica y geométrica no coinciden siempre] La geométrica (dimensión del espacio de autovectores) es siempre menor o igual que la algebraica (multiplicidad como raíz del característico), y la matriz diagonaliza exactamente cuando coinciden para todos los autovalores. En (b) valen 1 y 2: por eso falla. Segundo aviso: el mínimo divide siempre al característico (teorema de Cayley–Hamilton) y tiene las mismas raíces, pero con exponentes posiblemente menores; no son intercambiables. Y tercero: "no diagonalizable" depende del cuerpo, como muestra (c); sobre $\mathbb{C}$ toda matriz es al menos triangularizable.

Teorema 7.85 (Descomposición primaria)**.** *Si $m_A = \prod_i (\lambda-\lambda_i)^{r_i}$, el espacio se parte en suma directa de los* espacios propios generalizados

$$V = \bigoplus_i V_i, \qquad V_i = \ker(A-\lambda_i I)^{r_i},$$

cada uno invariante bajo A. Sobre cada V_i, $A = \lambda_i I + N_i$ con N_i nilpotente: esta es exactamente la estructura $D+N$ que anunciamos, y el refinamiento de la acción de cada N_i en cadenas produce los bloques de Jordan.

Idea. Los polinomios $q_i = m_A/(\lambda-\lambda_i)^{r_i}$ son coprimos en conjunto, así que por Bézout (capítulo 25) existen u_i con $\sum u_i q_i = 1$. Evaluando en A: $\sum u_i(A)q_i(A) = I$, y los operadores $P_i = u_i(A)q_i(A)$ resultan ser proyecciones sobre los V_i que suman la identidad. La aritmética de polinomios fabrica la geometría de la descomposición. □

Idea clave

Qué dice esto en cristiano. Cuando una matriz no diagonaliza, no hay que rendirse: se puede partir el espacio en trozos, uno por autovalor, de modo que dentro de cada trozo la matriz solo tenga ese autovalor. Esos trozos son los espacios generalizados, y se obtienen no con $(A-\lambda I)\mathbf{v}=\mathbf{0}$, sino con $(A-\lambda I)^r\mathbf{v}=\mathbf{0}$ para r suficientemente

grande. Es el paso previo imprescindible a la forma de Jordan: primero se separan los autovalores, y luego se ordena cada trozo por dentro.

Ejemplo 7.86 (Separar por autovalores)**. (a) Cuando ya está separado.** Para $\begin{pmatrix} 2 & 0 \\ 0 & 3 \end{pmatrix}$ los espacios generalizados son los dos ejes coordenados, de dimensión 1 cada uno, y su suma directa es $\mathbb{R}^2$. Con exponente $r = 1$ basta.

(b) Cuando hace falta subir el exponente. Para $\begin{pmatrix} 2 & 1 \\ 0 & 2 \end{pmatrix}$ el único autovalor es 2. El espacio de autovectores propiamente dichos tiene dimensión 1, pero $(A - 2I)^2 = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$, así que el espacio generalizado es todo $\mathbb{R}^2$, de dimensión 2. Con $r = 2$ se recupera la dimensión completa.

(c) Un caso mixto. Sea A de tamaño 3 con autovalores 2 (doble) y 5 (simple), y supongamos que el 2 solo aporta un autovector. Entonces el espacio generalizado del 2 tiene dimensión 2 y el del 5 dimensión 1; su suma directa es $\mathbb{R}^3$. La matriz se escribe como un bloque de Jordan de tamaño 2 y otro de tamaño 1.

(d) Para qué sirve. Una vez separado el espacio, en cada trozo la matriz es $\lambda I + N$ con N nilpotente, y las potencias y exponenciales se calculan de forma cerrada porque $N^k = 0$ a partir de cierto k. Ese es el mecanismo que permite resolver sistemas de ecuaciones diferenciales lineales con coeficientes constantes aunque la matriz no diagonalice.

Cuidado

[El espacio generalizado no es el de autovectores] Confundirlos es el error que hace que las cuentas no cuadren. En (b), el espacio de autovectores tiene dimensión 1 y el generalizado dimensión 2; solo el segundo tiene la dimensión que exige la multiplicidad algebraica. Segundo aviso: el exponente r_i que hay que usar es el del polinomio *mínimo*, no el del característico; usar el del característico funciona pero puede ser innecesariamente grande. Y tercero: la descomposición exige que el polinomio característico factorice completamente en el cuerpo, lo que sobre $\mathbb{R}$ puede fallar y sobre $\mathbb{C}$ nunca.

7.10.4. Triangularización de Schur y demostración de Cayley–Hamilton

Teorema 7.87 (Schur)**.** *Toda matriz compleja es* unitariamente triangularizable*: existe U con $U^*U = I$ tal que $U^*AU = T$ es triangular superior (con los valores propios en la diagonal).*

Demostración. Inducción en n. Sobre $\mathbb{C}$, A tiene al menos un valor propio λ_1 con vector propio unitario u_1 (teorema fundamental del álgebra aplicado a p_A). Completamos u_1 a base ortonormal (Gram–Schmidt): en esa base, la primera columna de A es $(\lambda_1, 0, \dots, 0)^{\mathsf{T}}$ y la matriz tiene forma de bloques $\left(\begin{smallmatrix} \lambda_1 & * \\ 0 & A' \end{smallmatrix}\right)$. Por hipótesis inductiva, A' se triangulariza unitariamente, y componiendo las transformaciones se concluye. □

Idea clave

Qué dice esto en cristiano. El teorema de Schur dice que aunque una matriz compleja no se pueda diagonalizar, siempre se puede al menos *triangularizar*, y además con un cambio de base ortonormal, que es el mejor tipo de cambio de base posible porque no deforma nada y su inversa es gratis. Es el resultado que salva la situación cuando la diagonalización falla, y de él se deducen casi todos los teoremas espectrales: el de matrices simétricas, el de normales y la propia descomposición en valores singulares.

Ejemplo 7.88 (Triangularizar cuando no se puede diagonalizar)**. (a) El caso fácil.** Si la matriz ya es diagonalizable con autovectores ortogonales, la U de Schur es la matriz de esos autovectores normalizados y U^*AU sale directamente diagonal. Toda matriz simétrica real está en este caso.

(b) El caso interesante. Para $A = \begin{pmatrix} 2 & 1 \\ 0 & 2 \end{pmatrix}$, que no diagonaliza, Schur no promete nada nuevo: la matriz ya es triangular superior, con la diagonal $2, 2$. El contenido del teorema es que *siempre* se puede llegar a esa forma, para cualquier matriz compleja, por muy desagradable que sea.

(c) Lo que se lee en la diagonal. En una triangular, los elementos de la diagonal son exactamente los autovalores, porque el determinante de $\lambda I - T$ es el producto de $\lambda - t_{ii}$. Como U^*AU y A tienen los mismos autovalores, Schur exhibe todos los autovalores de A en la diagonal de una matriz triangular.

(d) La consecuencia estrella. Si además A es *normal* ($AA^* = A^*A$), la triangular de Schur resulta ser diagonal, y se obtiene el teorema espectral: toda matriz normal diagonaliza en una base ortonormal. Las simétricas reales, las hermíticas y las unitarias son normales, así que todas ellas diagonalizan ortogonalmente. Media docena de teoremas por el precio de uno.

Cuidado

[Sobre los reales el teorema no vale tal cual] La rotación $\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$ no es triangularizable sobre $\mathbb{R}$, porque una triangular real tendría autovalores reales en la diagonal y esta matriz tiene $\pm$i. La versión real del teorema admite bloques 2×2 en la diagonal, uno por cada par de autovalores complejos conjugados, y se llama forma real de Schur. Segundo aviso: la U de Schur no es única, ni siquiera el orden en que aparecen los autovalores en la diagonal está determinado; se puede elegir el orden que convenga, y los algoritmos numéricos lo aprovechan.

Demostración de Cayley–Hamilton. Por Schur, basta probarlo para T triangular (pues $p_A(A) = U p_T(T) U^*$). Con diagonal $\lambda_1, \ldots, \lambda_n$: $p_T(T) = \prod_i (T - \lambda_i I)$. Afirmamos que el producto parcial $\prod_{i \le k} (T - \lambda_i I)$ anula los primeros k vectores de la base canónica, por inducción: $(T - \lambda_k I) e_k$ es combinación de $e_1, \ldots, e_{k-1}$ (triangularidad con diagonal λ_k en la posición k), y esos ya son anulados por los factores anteriores; los factores conmutan entre sí, así que el orden no estorba. Con $k = n$: $p_T(T) = 0$. □

Este mismo teorema de Schur, aplicado a una matriz *normal* ($AA^* = A^*A$), fuerza a la triangular T a ser diagonal (T triangular y normal $\Rightarrow$ diagonal, comparando normas de filas y columnas): eso demuestra de una vez el teorema espectral complejo, del que el caso real simétrico es corolario.

7.10.5. Existencia de la SVD, con demostración

Idea clave

Intuición primero. Cualquier matriz, por fea que sea, hace geométricamente una sola cosa: *rotar, estirar a lo largo de ejes perpendiculares, y volver a rotar* (V^T, Σ, U). La imagen de la esfera unitaria es siempre un elipsoide; los valores singulares son sus semiejes. Donde los valores propios preguntan "¿qué direcciones no rotan?" (pregunta que puede no tener respuesta real), los singulares preguntan "¿cuánto estira, como máximo y como mínimo?" — pregunta que *siempre* tiene respuesta, y por eso la SVD existe para toda matriz, cuadrada o no, y es la herramienta universal de datos, imágenes y mínimos cuadrados.

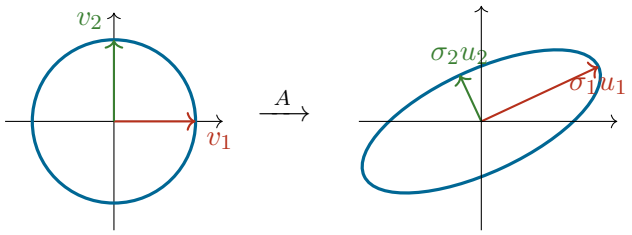


Figura 7.2: La SVD en una imagen: toda matriz manda el círculo unitario a una elipse; los valores singulares $\sigma_1 \geq \sigma_2$ son los semiejes y las u_i, v_i las direcciones privilegiadas de salida y de entrada.

Teorema 7.89. *Toda $A \in M_{m\times n}(\mathbb{R})$ admite $A = U\Sigma V^\mathsf{T}$ con U, V ortogonales y Σ diagonal con $\sigma_1 \geq \cdots \geq \sigma_r > 0$ ($r = \operatorname{rango} A$).*

Demostración. $A^\mathsf{T}A$ es simétrica y semidefinida positiva ($\mathbf{x}^\mathsf{T}A^\mathsf{T}A\mathbf{x} = \|A\mathbf{x}\|^2 \geq 0$): por el teorema espectral tiene base ortonormal de vectores propios $v_1, \ldots, v_n$ con valores propios $\sigma_1^2 \geq \cdots \geq \sigma_r^2 > 0 = \cdots = 0$. Definimos $u_i = \frac{Av_i}{\sigma_i}$ para $i \leq r$: son ortonormales, pues

$$u_i^\mathsf{T} u_j = \frac{v_i^\mathsf{T} A^\mathsf{T} A v_j}{\sigma_i \sigma_j} = \frac{\sigma_j^2\, v_i^\mathsf{T} v_j}{\sigma_i \sigma_j} = \delta_{ij}.$$

Completamos $\{u_i\}$ a base ortonormal de $\mathbb{R}^m$. Entonces $Av_i = \sigma_i u_i$ ($i \leq r$) y $Av_i = 0$ ($i > r$), que en forma matricial es exactamente $AV = U\Sigma$, es decir $A = U\Sigma V^\mathsf{T}$. □

Obsérvese la economía: la SVD es *el teorema espectral aplicado a* $A^\mathsf{T}A$. De paso, la demostración muestra qué son los cuatro subespacios fundamentales: $\{v_1, \ldots, v_r\}$ genera el espacio fila, $\{v_{r+1}, \ldots, v_n\}$ el núcleo, $\{u_1, \ldots, u_r\}$ el espacio columna y $\{u_{r+1}, \ldots, u_m\}$ el núcleo izquierdo, con las ortogonalidades (fila $\perp$ núcleo, columna $\perp$ núcleo izquierdo) a la vista.

7.10.6. El espacio dual

Definición 7.90. El *dual* de V es $V^* = \{\varphi : V \to K \text{ lineales}\}$ (los *funcionales*). Si $\dim V = n$, también $\dim V^* = n$: a cada base $\{v_i\}$ le corresponde la *base dual* $\{\varphi^j\}$ con $\varphi^j(v_i) = \delta_i^j$ (cada φ^j "lee la coordenada j"). Toda aplicación lineal $T : V \to W$ induce la *traspuesta* $T^* : W^* \to V^*$, $T^*(\psi) = \psi \circ T$, cuya matriz es la transpuesta de la de T: la transposición de matrices, que parecía un truco notacional, es un funtor.

Idea clave

Qué dice esto en cristiano. Un funcional lineal es una máquina que come vectores y escupe números, respetando sumas y escalas: medir una coordenada, sumar componentes, evaluar un polinomio en un punto. El dual es el conjunto de todas esas máquinas, y resulta ser otro espacio vectorial de la misma dimensión que el original. Que tengan la misma dimensión significa que son isomorfos, pero el isomorfismo depende de la base elegida, y esa dependencia es justo lo que distingue un vector de un covector, distinción que en el capítulo de tensores se vuelve esencial.

Ejemplo 7.91 (Funcionales concretos y la base dual)**. (a) Los más simples.** En $\mathbb{R}^3$, la aplicación $\varphi(x, y, z) = x$ es lineal y va a $\mathbb{R}$: es un funcional. También lo son $\psi(x, y, z) = 2x - y + 3z$ y, en general, todo producto escalar por un vector fijo. En $\mathbb{R}^n$ todos los funcionales son de esa forma.

(b) La base dual. Si $\{e_1, e_2, e_3\}$ es la base canónica, la base dual $\{e^1, e^2, e^3\}$ está definida por $e^i(e_j) = 1$ si $i = j$ y 0 si no. Es decir, e^i es "devuélveme la coordenada i-ésima". Cualquier funcional se escribe como combinación de esos tres, y sus coeficientes son precisamente los valores sobre la base: para ψ de (a), los coeficientes son 2, -1 y 3.

(c) Un funcional que no es un producto escalar. En el espacio de polinomios de grado a lo sumo 2, la evaluación $\varphi(p) = p(3)$ es lineal, y también lo es $\varphi(p) = \int_0^1 p$. Ambas comen polinomios y devuelven números. La dimensión del espacio es 3 y la del dual también.

(d) El cambio de base va al revés. Si se cambia la base multiplicando por una matriz P, las coordenadas de un vector cambian con P^{-1} y las de un funcional con P^{T}. Por eso los vectores se llaman contravariantes y los funcionales covariantes. Con una base ortonormal ambos cambian igual, y por eso en $\mathbb{R}^n$ con el producto escalar habitual la distinción parece invisible.

Cuidado

[La identificación $V \cong V^*$ no es canónica] En dimensión finita V y V^* tienen la misma dimensión, luego son isomorfos, pero para exhibir el isomorfismo hace falta *elegir* una base o un producto interno; no hay ninguno preferido. En cambio V y V^{**} sí son canónicamente isomorfos, sin elegir nada. Segundo aviso: en dimensión infinita esto se rompe del todo; el dual de un espacio de dimensión infinita es estrictamente mayor, y por eso en análisis funcional se trabaja con el dual *topológico*, solo los funcionales continuos.

Dos advertencias que preparan el capítulo 36. Primera: $V \cong V^*$ en dimensión finita, pero *no canónicamente* —el isomorfismo depende de elegir base o producto interno—, mientras que $V \cong V^{**}$ sí es canónico ($v \mapsto$ "evaluar en v"). Segunda: bajo cambio de base, las coordenadas de los vectores y las de los funcionales se transforman con matrices *inversas* entre sí. Esa asimetría es exactamente la distinción contravariante/covariante que el cálculo tensorial convertirá en sistema.

7.11. Complementos: bloques, Cholesky, normas y pseudoinversa

7.11.1. Matrices por bloques

Se operan por bloques como números (respetando orden). Triangular por bloques: $\det\left(\begin{smallmatrix} A & B \\ 0 & D \end{smallmatrix}\right) = \det A \det D$; en general, con A invertible, el *complemento de Schur*:

$$\det \begin{pmatrix} A & B \\ C & D \end{pmatrix} = \det A \cdot \det(D - CA^{-1}B),$$

por eliminación por bloques. $D - CA^{-1}B$ es lo que queda al eliminar variables: covarianza condicional en estadística, impedancia equivalente en circuitos.

7.11.2. Cholesky

Teorema 7.92. *A simétrica es definida positiva $\iff$ $A = LL^{\mathsf{T}}$ con L triangular inferior de diagonal positiva, única.*

Idea clave

Qué dice esto en cristiano. Cholesky es la raíz cuadrada de una matriz: escribe A como LL^{T}, igual que un número positivo se escribe como $s \cdot s$. Solo funciona con matrices simétricas definidas positivas, y de hecho el intento de calcularla *es* el test más barato para saber si una matriz lo es: si en algún paso hay que sacar la raíz de un número negativo, la matriz no era definida positiva. Cuesta la mitad que una eliminación gaussiana y es numéricamente muy estable, por lo que se usa en todas partes: mínimos cuadrados, simulación, estadística multivariante.

Ejemplo 7.93 (Factorizar y detectar)**. (a) Una que sí se puede.** Sea $A = \begin{pmatrix} 4 & 2 \\ 2 & 3 \end{pmatrix}$. Su factor de Cholesky es $L = \begin{pmatrix} 2 & 0 \\ 1 & 1{,}414214 \end{pmatrix}$, y en efecto $LL^{\mathsf{T}} = \begin{pmatrix} 4 & 2 \\ 2 & 3 \end{pmatrix}$. Los autovalores de A son 1,438447 y 5,561553, ambos positivos, coherente con que sea definida positiva.

(b) Una que no. Para $B = \begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix}$ los autovalores son -1 y 3: hay uno negativo, luego no es definida positiva y Cholesky falla. El fallo se produce en el segundo paso, al intentar $\sqrt{1-4}$.

(c) El número de condición como amplificador. La matriz de Hilbert $\begin{pmatrix} 1 & 1/2 \\ 1/2 & 1/3 \end{pmatrix}$ es definida positiva y admite Cholesky, pero su número de condición es 19,28: un error relativo del uno por mil en los datos puede convertirse en un error del dos por ciento en la solución. La factorización existe; lo que no existe es la precisión.

(d) La lectura de la cota. La desigualdad $\frac{\|\delta \mathbf{x}\|}{\|\mathbf{x}\|} \leq \kappa(A) \frac{\|\delta \mathbf{b}\|}{\|\mathbf{b}\|}$ dice exactamente cuánto puede amplificarse el error relativo. Con $\kappa = 1$ (matrices ortogonales) no se amplifica nada, y por eso los algoritmos numéricos las prefieren; con $\kappa = 40\,002$, como en el ejemplo casi singular de antes, un error en la sexta cifra puede comerse la segunda.

Cuidado

[Simétrica no basta, y el condicionamiento no es culpa del método] Toda matriz simétrica tiene autovalores reales, pero pueden ser negativos; definida positiva exige que *todos* lo sean, o equivalentemente que todos los menores principales sean positivos. El apartado (b) es simétrico y no cumple. Segundo aviso, y es conceptual: un número de condición alto no significa que el algoritmo esté mal programado; significa que el *problema* amplifica los errores, y ningún método puede evitarlo. Cambiar de algoritmo no ayuda; hay que reformular el problema.

Fórmulas columna a columna: $\ell_{jj} = \sqrt{a_{jj} - \sum_{k<j} \ell_{jk}^2}$, $\ell_{ij} = \frac{a_{ij} - \sum_k \ell_{ik}\ell_{jk}}{\ell_{jj}}$. Mitad del coste de LU y test numérico de positividad; caballo de batalla de mínimos cuadrados y simulación gaussiana.

7.11.3. Normas matriciales y condicionamiento

$\|A\| = \sigma_1$ (máximo estiramiento), $\kappa(A) = \frac{\sigma_1}{\sigma_n}$.

Teorema 7.94. $\frac{\|\delta \mathbf{x}\|}{\|\mathbf{x}\|} \leq \kappa(A) \frac{\|\delta \mathbf{b}\|}{\|\mathbf{b}\|}$ *(multiplicar* $\|\delta x\| \leq \|A^{-1}\|\|\delta b\|$ *por* $\|b\| \leq \|A\|\|x\|$*).*

Ejemplo 7.95 (Un sistema mal condicionado)**. (a) La matriz peligrosa.** Sea $A = \left(\begin{smallmatrix} 1 & 1 \\ 1 & 1{,}0001 \end{smallmatrix}\right)$, cuyo número de condición vale $\kappa(A) = 40002{,}000075$. Con $b = (2;\ 2{,}0001)$ la solución exacta es $(1, 1)$.

(b) La catástrofe. Cambiando b a $(2;\ 2{,}0002)$, es decir, con una perturbación relativa de $3{,}5 \cdot 10^{-5}$, la solución pasa a ser $(0, 2)$: un error relativo del cien por cien. La cota del teorema lo predecía: $40002 \cdot 3{,}5 \cdot 10^{-5} \approx 1{,}4$.

(c) La matriz de Hilbert. $H_{ij} = 1/(i + j - 1)$ tiene condición $1{,}551374 \cdot 10^4$ ya con 4×4, y crece exponencialmente con el tamaño: para $n = 10$ supera 10^{13}, de modo que en doble precisión no queda ninguna cifra correcta. Es el ejemplo clásico de sistema que hay que resolver con aritmética exacta o no resolverlo.

(d) Qué no significa. Un κ grande no dice que el algoritmo sea malo ni que la matriz esté cerca de ser singular en valor absoluto: dice que el *problema* es sensible. Ningún método, por bueno que sea, puede devolver más precisión de la que los datos contienen.

$\kappa \approx 10^k$ come k dígitos; y $\kappa(A^\mathsf{T}A) = \kappa(A)^2$: por eso el software serio usa QR o SVD y no las ecuaciones normales.

7.11.4. Pseudoinversa de Moore–Penrose

$A^+ = V\Sigma^+U^\mathsf{T}$: $\mathbf{x}^* = A^+\mathbf{b}$ es la solución de mínimos cuadrados *de norma mínima*; con columnas independientes, $A^+ = (A^\mathsf{T}A)^{-1}A^\mathsf{T}$; invertible, $A^+ = A^{-1}$. Existe siempre: resuelve, aproxima u optimiza según el caso.

7.12. Cierre del temario: los subtemas restantes

7.12.1. Matrices elementales

Cada operación de fila es multiplicar por una *matriz elemental* invertible; toda invertible es producto de elementales; Gauss–Jordan sobre $[A|I]$ acumula exactamente A^{-1}. La forma escalonada *reducida* es única: representante canónico de la equivalencia por filas.

7.12.2. Los cuatro subespacios fundamentales

Teorema 7.96. *Para A de rango r: fila y columna de dimensión r; núcleo $n-r$; núcleo izquierdo $m-r$; con $N(A) = C(A^\mathsf{T})^\perp$ y $N(A^\mathsf{T}) = C(A)^\perp$.*

Ejemplo 7.97 (Los cuatro subespacios fundamentales)**. (a) Un caso 3×4.** Sea A con filas $(1,2,0,1)$, $(2,4,1,3)$, $(3,6,1,4)$. La tercera fila es la suma de las dos primeras, luego $r = 2$. Entonces el espacio columna tiene dimensión 2 dentro de $\mathbb{R}^3$, el espacio fila dimensión 2 dentro de $\mathbb{R}^4$, el núcleo dimensión $4-2=2$ y el núcleo por la izquierda dimensión $3-2=1$.

(b) Las dos ortogonalidades. El núcleo es el ortogonal del espacio fila en $\mathbb{R}^4$, y el núcleo por la izquierda es el ortogonal del espacio columna en $\mathbb{R}^3$. Aquí, el vector $(1,1,-1)$ genera el núcleo izquierdo, y en efecto es ortogonal a las tres columnas.

(c) Existencia y unicidad de soluciones. $A\mathbf{x} = \mathbf{b}$ tiene solución si y sólo si $\mathbf{b}$ está en el espacio columna, y la solución es única si y sólo si el núcleo es trivial. Con $r = 2 < 4$ el núcleo no lo es, así que cuando hay solución hay infinitas, formando un plano de dimensión 2.

(d) Rango fila igual a rango columna. Es el hecho menos obvio del teorema: las filas viven en $\mathbb{R}^4$ y las columnas en $\mathbb{R}^3$, espacios distintos, y sin embargo los dos subespacios tienen la misma dimensión. La demostración corta pasa por la factorización $A = CR$ con C de r columnas y R de r filas.

A manda el espacio fila biyectivamente al columna (ahí vive la SVD) y aplasta el núcleo. Alternativa de Fredholm: $A\mathbf{x} = \mathbf{b}$ resoluble $\iff \mathbf{b} \perp N(A^\mathsf{T})$ —o resuelves o hay certificado de imposibilidad: la semilla de la dualidad—.

7.12.3. Espacios complejos: hermitianas, unitarias, normales

Producto $\langle u, v\rangle = \sum \bar{u}_i v_i$, adjunta $A^* = \bar{A}^\mathsf{T}$. Diccionario: simétrica→hermitiana (autovalores reales), ortogonal→unitaria ($|\lambda| = 1$), y *normal* ($AA^* = A^*A$) $\iff$ unitariamente

diagonalizable (vía Schur). En cuántica: observables = hermitianas, evolución = unitaria, e^{iH} las conecta.

7.12.4. Formas bilineales y congruencia

$B(u,v) = u^{\mathsf{T}}Av$ cambia por *congruencia* $P^{\mathsf{T}}AP$ (no semejanza). Invariante completo real: la *signatura* (ley de inercia). Congruencia clasifica geometrías (euclídea, Minkowski); semejanza clasifica dinámicas: distinguirlas evita "diagonalizar" métricas con autovalores cuando toca completar cuadrados.

7.12.5. Matrices y grafos

$(M^k)_{ij}$ cuenta caminos de longitud k (inducción = producto de matrices). Triángulo completo: $M = J - I$, espectro $\{2,-1,-1\}$: caminos cerrados $= 2^k + 2(-1)^k$. Sobre esto: rankings espectrales, laplaciano de grafo $L = D - M$ y la conectividad algebraica de Fiedler.

7.13. Profundización final: teoría espectral fina

7.13.1. Espacios cociente

V/W: clases $v+W$, $\dim = \dim V - \dim W$; $V/\ker T \cong \operatorname{im} T$ (isomorfía lineal; rango–nulidad es su sombra numérica).

7.13.2. Desigualdades de rango

$\operatorname{rango}(AB) \le$ mín; Sylvester: $\operatorname{rango} A + \operatorname{rango} B - n \le \operatorname{rango}(AB)$; $\operatorname{rango}(A+B) \le \operatorname{rango} A + \operatorname{rango} B$; $\operatorname{rango}(A^{\mathsf{T}}A) = \operatorname{rango} A$ (por $\|Ax\|^2$) —la que garantiza las ecuaciones normales—.

7.13.3. Funciones de matrices

$f(A) = P\,\mathrm{diag}(f(\lambda_i))P^{-1}$: $\sqrt{A}$, $\cos(At)$ resolviendo $\ddot{x} = -A^2x$, $e^{-iHt/\hbar}$. En bloques de Jordan entran derivadas: $f(J) = f(\lambda)I + f'(\lambda)N + \cdots$ —por eso los $te^{\lambda t}$—.

7.13.4. Courant–Fischer y entrelazado

$$\lambda_k = \max_{\dim S = k} \min_{0 \neq x \in S} \frac{x^{\mathsf{T}}Ax}{x^{\mathsf{T}}x} :$$

caracterización variacional del espectro (cociente de Rayleigh). Corolario: *entrelazado de Cauchy* al borrar fila y columna. Permite estimar espectros con vectores de prueba y fundamenta los elementos finitos para autovalores.

7.13.5. Perron–Frobenius e iteración de potencias

Matriz positiva: autovalor dominante simple > 0 con autovector positivo (único). Algoritmo: $x_{k+1} = \frac{Ax_k}{\|Ax_k\|}$ converge al dominante con razón $|\lambda_2/\lambda_1|$. Para $\left(\begin{smallmatrix} 4 & 1 \\ 2 & 3 \end{smallmatrix}\right)$: razón $\frac{2}{5}$ hacia $(1,1)$. Google: esto, en dimensión 10^{10}.

7.13.6. Definidas positivas: las siete caracterizaciones

(i) $x^{\mathsf{T}}Ax > 0$; (ii) autovalores > 0; (iii) menores dominantes > 0; (iv) $A = B^{\mathsf{T}}B$, B invertible; (v) Cholesky; (vi) matriz de Gram de una base; (vii) pivotes de Gauss positivos. Saber transitar la lista es elegir el test más barato: (v) numérico, (iii) a mano, (vi) para positividad gratis (covarianzas).

7.14. Última capa: formas canónicas y transformadas

7.14.1. La forma de Jordan, demostrada

Falta el teorema completo. Por la descomposición primaria basta el caso nilpotente: si $N^m = 0$, se eligen vectores de altura máxima y sus cadenas $\{v, Nv, \dots, N^{k-1}v\}$; un argumento inductivo (cociente por la imagen, o el lema de intercambio aplicado a las alturas) muestra que V es suma directa de cadenas: en esa base, N es suma de bloques nilpotentes elementales, y $A = \lambda I + N$ da los bloques de Jordan. *Unicidad*: el número de bloques de tamaño $\geq k$ para λ es $\operatorname{rango}(A-\lambda I)^{k-1} - \operatorname{rango}(A-\lambda I)^k$: los rangos de las potencias determinan la forma —y dan el algoritmo práctico para calcularla—.

7.14.2. Forma canónica racional

Sin salir del cuerpo base (sin autovalores): toda matriz es semejante a una suma de *matrices compañeras* de sus *factores invariantes* $d_1 \mid d_2 \mid \cdots \mid d_k$ (con d_k = polinomio mínimo y $\prod d_i$ = característico). Dos matrices son semejantes sobre $K \iff$ mismos factores invariantes: la clasificación completa de la semejanza, válida sobre $\mathbb{Q}$ sin tocar $\mathbb{C}$. (Es el teorema de estructura de módulos sobre el DIP $K[\lambda]$ —capítulo 25— con el espacio visto como módulo vía $p(\lambda)\cdot v = p(A)v$.)

7.14.3. Descomposición polar

Teorema 7.98. *Toda A invertible se factoriza únicamente $A = QS$ con Q ortogonal y S simétrica definida positiva ($S = \sqrt{A^{\mathsf{T}}A}$ vía cálculo funcional, $Q = AS^{-1}$).*

Ejemplo 7.99 (La descomposición polar)**. (a) La analogía.** Para un número complejo, $z = re^{i\theta}$ separa módulo y giro. La descomposición polar hace lo mismo con matrices: $S = \sqrt{A^{\top}A}$ es la parte de "estiramiento" (simétrica definida positiva) y Q la de "giro" (ortogonal). El nombre no es una metáfora: es literalmente la generalización.

(b) Un ejemplo. Sea $A = \left(\begin{smallmatrix} 0 & -2 \\ 3 & 0 \end{smallmatrix}\right)$. Entonces $A^{\top}A = \operatorname{diag}(9,4)$, luego $S = \operatorname{diag}(3,2)$ y $Q = AS^{-1} = \left(\begin{smallmatrix} 0 & -1 \\ 1 & 0 \end{smallmatrix}\right)$, un giro de noventa grados. La matriz estira por 3 y por 2 y luego gira.

(c) Relación con la SVD. Si $A = U\Sigma V^\top$, entonces $Q = UV^\top$ y $S = V\Sigma V^\top$. Las dos descomposiciones contienen la misma información reorganizada, y la polar es la que usan la mecánica de medios continuos (para separar deformación de rotación) y los algoritmos de alineación de estructuras moleculares.

(d) Sin invertibilidad, se pierde la unicidad. Para $A = 0$ cualquier Q ortogonal sirve con $S = 0$. La descomposición sigue existiendo para matrices singulares, pero Q deja de estar determinada en el núcleo, y por eso el enunciado exige que A sea invertible.

Es el análogo matricial de $z = e^{i\theta}r$: toda transformación invertible es rotación por estiramiento. En mecánica de medios continuos es *el* teorema de descomposición de deformaciones (rotación local × estiramiento puro), y la Q es la solución del problema de Procrustes (la rotación que mejor alinea dos nubes de puntos, con SVD).

7.14.4. Cauchy–Binet

Para A de $m \times n$ y B de $n \times m$ ($m \le n$):

$$\det(AB) = \sum_S \det(A_{\cdot S}) \det(B_{S\cdot}),$$

suma sobre los subconjuntos S de m columnas. Generaliza $\det(AB) = \det A \det B$ y da geometría: $\det(A^\mathsf{T}A) = \sum_S (\det A_S)^2$ es el cuadrado del volumen m-dimensional del paralelepípedo generado —el "Pitágoras de los volúmenes": el área de una superficie inclinada al cuadrado es la suma de los cuadrados de sus sombras coordenadas, que es exactamente el $\|\mathbf{r}_u \times \mathbf{r}_v\|$ del capítulo 11—.

7.14.5. Circulantes y la transformada discreta de Fourier

Una *circulante* (cada fila es la anterior rotada) es un polinomio en la matriz de rotación cíclica P; como $P^n = I$, sus autovalores son las raíces de la unidad y *toda* circulante se diagonaliza por la matriz de Fourier $F_{jk} = \frac{1}{\sqrt{n}}\omega^{jk}$ ($\omega = e^{2\pi i/n}$), con autovalores el "espectro" del primer renglón. Consecuencia práctica: convolución circular = producto de circulantes = multiplicación de espectros: multiplicar polinomios grandes, filtrar señales y resolver EDP periódicas cuesta $O(n \log n)$ vía FFT. Es el teorema espectral del grupo $\mathbb{Z}_n$ (¡capítulo 25!): los caracteres ω^{jk} diagonalizan todo lo que conmuta con las rotaciones.

7.15. Ronda de temarios: cociente, bidual y anuladores

El temario oficial de Álgebra Lineal II (Facultad de Ciencias, UNAM) dedica una unidad a la dualidad que aún no habíamos cerrado del todo: espacio cociente, bidual y anuladores, el trío que convierte la dualidad en un diccionario exacto.

7.15.1. El espacio cociente y el primer teorema de isomorfismo

Definición 7.100. Si $W \subseteq V$ es un subespacio, el *cociente* V/W es el conjunto de clases $v + W = \{v + w : w \in W\}$ con las operaciones $(u + W) + (v + W) = (u + v) + W$

y $\lambda(v+W)=\lambda v+W$ (bien definidas: cambiar de representante altera el resultado en un elemento de W, es decir, en nada). Geométricamente, los elementos de V/W son los trasladados paralelos de W, y el cociente "aplasta" W a cero.

Idea clave

Qué dice esto en cristiano. Cocientar por W es decidir que los vectores de W "no cuentan": dos vectores pasan a ser el mismo si difieren en algo de W. Si W es una recta en el plano, el cociente identifica cada recta paralela a W con un solo punto, y lo que queda es una recta de clases. Es la manera algebraica de mirar un espacio ignorando una dirección, exactamente como en aritmética modular se ignoran los múltiplos de m.

Ejemplo 7.101 (Cocientes concretos)**. (a) El plano módulo una recta.** $V=\mathbb{R}^2$, $W=\{(t,0)\}$. Las clases son las rectas horizontales, cada una determinada por su altura, luego $V/W\cong\mathbb{R}$. La proyección es $(x,y)\mapsto y$.

(b) Los polinomios. $V=\mathbb{R}[x]$, W los múltiplos de x^2+1. El cociente tiene dimensión 2 con base $\{1,x\}$, y como $x^2\equiv -1$, resulta ser $\mathbb{C}$. La construcción de los complejos por cociente es literalmente este ejemplo.

(c) El caso trivial en los dos extremos. Con $W=\{0\}$ se recupera $V/W\cong V$; con $W=V$ el cociente es el espacio nulo. Los dos extremos son coherentes con la fórmula de dimensiones.

(d) No es un subespacio. El cociente V/W no es un subespacio de V: sus elementos son *conjuntos* de vectores, no vectores. Se le puede identificar con un complemento de W, pero esa identificación depende del complemento elegido y no es canónica, mientras que el cociente sí lo es.

Proposición 7.102. *Si* $\dim V=n$ *y* $\dim W=k$*, entonces* $\dim V/W=n-k$.

Demostración. Extiéndase una base $w_1,\dots,w_k$ de W a una de V con $v_{k+1},\dots,v_n$. Las clases $v_{k+1}+W,\dots,v_n+W$ generan V/W (todo vector es combinación de la base y los w_i mueren en el cociente) y son independientes: si $\sum_{i>k}\lambda_i v_i\in W$, escribiéndolo en la base de V la unicidad de coordenadas fuerza $\lambda_i=0$. □ □

Ejemplo 7.103 (Contar dimensiones al cocientar)**. (a) Casos inmediatos.** En $\mathbb{R}^5$, cocientar por un plano deja dimensión 3; cocientar por un hiperplano deja dimensión 1; cocientar por el subespacio nulo deja 5.

(b) La demostración en una frase. Se toma una base $w_1,\dots,w_k$ de W, se amplía a una base $w_1,\dots,w_k,v_1,\dots,v_{n-k}$ de V, y las clases $[v_1],\dots,[v_{n-k}]$ forman una base del cociente, porque las $[w_i]$ son todas nulas.

(c) La relación con el rango. Aplicándolo a $W=\ker T$ se obtiene $\dim(V/\ker T)=n-\dim\ker T=\operatorname{rango} T$, que es el teorema del rango disfrazado. Las dos formulaciones son la misma cuenta.

(d) En dimensión infinita. La fórmula deja de tener sentido literal, pero sobrevive la

noción de *codimensión*: en ℓ^2, el subespacio de las sucesiones con primera coordenada nula tiene codimensión 1 aunque tanto él como el total sean de dimensión infinita. La codimensión, no la dimensión, es lo que se conserva.

Teorema 7.104 (Primer teorema de isomorfismo vectorial)**.** *Si $T : V \to U$ es lineal, entonces $\overline{T} : V/\ker T \to \operatorname{im} T$, $\overline{T}(v + \ker T) = T(v)$, es un isomorfismo.*

Demostración. Bien definida y lineal: si $v - v' \in \ker T$, $Tv = Tv'$. Sobreyectiva sobre la imagen por construcción. Inyectiva: $\overline{T}(v + \ker T) = 0$ significa $v \in \ker T$, la clase cero. □ □

Ejemplo 7.105 (El primer teorema de isomorfismo en acción)**. (a) Una proyección.** Sea $T\colon \mathbb{R}^3 \to \mathbb{R}^2$, $T(x, y, z) = (x, y)$. El núcleo es el eje z y la imagen es todo $\mathbb{R}^2$, de modo que $\mathbb{R}^3/\text{eje } z \cong \mathbb{R}^2$. La dimensión cuadra: $3 - 1 = 2$.

(b) La derivada. Sea $D\colon \mathbb{R}[x]_{\le 3} \to \mathbb{R}[x]_{\le 2}$ la derivación. Su núcleo son las constantes (dimensión 1) y su imagen es todo $\mathbb{R}[x]_{\le 2}$ (dimensión 3). El teorema da $\mathbb{R}[x]_{\le 3}/\mathbb{R} \cong \mathbb{R}[x]_{\le 2}$, y $4 - 1 = 3$.

(c) La evaluación. Sea $T\colon \mathbb{R}[x] \to \mathbb{R}$, $p \mapsto p(2)$. El núcleo son los múltiplos de $x - 2$ y la imagen es $\mathbb{R}$ entero, luego $\mathbb{R}[x]/(x - 2) \cong \mathbb{R}$. Ésta es la versión vectorial del hecho de que evaluar en un punto es lo mismo que reducir módulo $x - a$.

(d) Por qué importa. El teorema dice que toda aplicación lineal es, en esencia, "aplastar el núcleo y luego identificar". Cualquier estudio de una aplicación se reduce al de su imagen, y el cociente es el precio contable exacto de ese aplastamiento. La misma frase, palabra por palabra, vale para grupos, anillos y módulos.

Observación 7.106. Tomando dimensiones: $\dim V - \text{nulidad}(T) = \text{rango}(T)$: *el teorema rango–nulidad es el primer teorema de isomorfismo con los ojos entrecerrados.* Compárese con el capítulo 25: mismo enunciado, misma prueba, otra categoría; ver el patrón aquí, donde todo es explícito, hace transparente el caso de grupos, anillos y módulos.

7.15.2. El bidual, o por qué "canónico" es una palabra técnica

Teorema 7.107. *La aplicación $\mathrm{ev} : V \to V^{**}$ dada por $\mathrm{ev}(v)(\varphi) = \varphi(v)$ es lineal e inyectiva; si $\dim V < \infty$, es un isomorfismo.*

Demostración. La linealidad es inmediata. Inyectividad: si $v \neq 0$, extiéndase $\{v\}$ a una base y tómese el funcional coordenado φ con $\varphi(v) = 1$: entonces $\mathrm{ev}(v)(\varphi) = 1 \neq 0$. En dimensión finita $\dim V^{**} = \dim V^* = \dim V$, y una inyección lineal entre espacios de igual dimensión finita es un isomorfismo. □ □

Idea clave

$V \cong V^*$ también es cierto en dimensión finita, pero *ese* isomorfismo exige elegir una base (o un producto interno): elecciones distintas dan isomorfismos distintos. En cambio ev no menciona base alguna: por eso se llama *canónico*. La distinción parece filosófica y es operativa: lo canónico sobrevive a los cambios de coordenadas. En el capítulo 36,

vectores y covectores se transforman de manera distinta *precisamente* porque $V \cong V^*$ no es canónico (hace falta la métrica g para subir y bajar índices), mientras que la identificación $V \cong V^{**}$ es gratis y silenciosa. En dimensión infinita ev puede no ser sobreyectiva: los espacios donde sí lo es (*reflexivos*, como los L^p con $1 < p < \infty$) son los bien portados del análisis funcional.

7.15.3. Anuladores: la correspondencia exacta

Definición 7.108. Para $S \subseteq V$, el *anulador* es $S^0 = \{\varphi \in V^* : \varphi(s) = 0 \ \forall s \in S\}$, un subespacio de V^*.

Teorema 7.109. *Sea* $\dim V = n$ *y* $W \subseteq V$ *un subespacio. Entonces*

$$\dim W^0 = n - \dim W, \qquad (W^0)^0 = W$$

(la segunda, identificando $V^{**} \cong V$ *vía* ev*). Además* $(U + W)^0 = U^0 \cap W^0$ *y* $(U \cap W)^0 = U^0 + W^0$.

Demostración. Base $w_1, \dots, w_k$ de W extendida a una de V; sea $\varphi^1, \dots, \varphi^n$ la base dual. Un funcional $\sum a_i \varphi^i$ anula a W si y solo si $a_1 = \cdots = a_k = 0$: así $W^0 = \operatorname{gen}(\varphi^{k+1}, \dots, \varphi^n)$, de dimensión $n - k$. Doble anulador: claramente $W \subseteq (W^0)^0$ y las dimensiones coinciden $(n - (n - k) = k)$, luego son iguales. La ley $(U + W)^0 = U^0 \cap W^0$ es directa de la definición (anular la suma es anular ambos); aplicando esa ley a U^0, W^0 y tomando anuladores con $(W^0)^0 = W$ sale la otra. □ □

Ejemplo 7.110 (El anulador es el espacio de ecuaciones)**.** Si $W \subseteq \mathbb{R}^4$ es el plano generado por $(1, 0, 1, 0)$ y $(0, 1, 0, 1)$, resolver $\varphi(w) = 0$ da la base $\{x_1 - x_3,\ x_2 - x_4\}$ de W^0. Es decir: W^0 *es el espacio de ecuaciones lineales que definen a* W, y $\dim W^0 = n - \dim W$ dice que un subespacio de dimensión k requiere exactamente $n - k$ ecuaciones independientes. La dualidad parametrización↔ecuaciones que usamos desde el capítulo de sistemas era teoría de anuladores sin decir su nombre; y la transpuesta cierra el círculo con la identidad $(\operatorname{im} T)^0 = \ker T^*$, que es la alternativa de Fredholm en lenguaje limpio.

7.16. Cierre definitivo: diagonalización simultánea

La última pieza estructural del capítulo: ¿cuándo dos operadores comparten una base de vectores propios? La respuesta es un si-y-solo-si limpio, y es la matemática detrás de frases físicas como "observables compatibles".

Teorema 7.111 (Diagonalización simultánea)**.** *Sean* A, B *diagonalizables. Existe una base que diagonaliza a* ambos *si y solo si conmutan:* $AB = BA$.

Demostración. ($\Rightarrow$) En una base propia común ambos son diagonales, y las diagonales conmutan. ($\Leftarrow$) El paso clave: B *preserva los espacios propios de* A. Si $Av = \lambda v$, entonces

$$A(Bv) = B(Av) = \lambda(Bv) :$$

Bv vive en el mismo espacio propio E_λ. Restrinjamos B a cada E_λ: sigue siendo diagonalizable (su polinomio mínimo divide al de B, que no tiene raíces repetidas: el criterio del capítulo), así que dentro de cada E_λ hay una base de vectores propios de B —que son automáticamente propios de A, pues todo E_λ es propio de A—. Uniendo esas bases sobre todos los λ (suma directa por diagonalizabilidad de A): base propia común. □ □

Ejemplo 7.112 (Tres lecturas del mismo teorema)**.** *(a) Las circulantes, explicadas.* Todas las matrices circulantes conmutan entre sí (son polinomios en el desplazamiento cíclico S): el teorema *predice* una base propia común, y es la base de Fourier discreta que encontramos: la DFT no era un truco sino la consecuencia estructural de la conmutación con S.

(b) Cuántica. Dos observables (hermitianos) admiten estados con ambos valores definidos exactamente cuando conmutan: "compatibles". Posición y momento tienen $[X, P] = i\hbar I \neq 0$: *no* hay base común, y esa obstrucción algebraica es el principio de incertidumbre en su forma más desnuda (la analítica, con la desigualdad, está en el capítulo 31: son el mismo hecho en dos idiomas).

(c) Con simetrías se calcula. Si A conmuta con una simetría P (permutación, reflexión), los espacios propios de P parten el problema en bloques pequeños. El grafo camino con su reflexión izquierda–derecha: los modos se separan en pares e impares y el cálculo espectral se parte a la mitad: así se diagonalizó el laplaciano discreto, y así explotan las simetrías todos los métodos numéricos y la química cuántica (los orbitales moleculares se etiquetan por las simetrías de la molécula: teoría de caracteres del capítulo 25 aplicada a este teorema).

Observación 7.113. La versión con repetidos: una *familia* arbitraria de diagonalizables que conmutan dos a dos tiene base propia común (inducción sobre la dimensión, partiendo por un operador no escalar de la familia). Y la versión hermitiana es la joya práctica: familias que conmutan de matrices hermitianas se diagonalizan simultáneamente con una base *ortonormal* (espectral más el argumento de arriba, con restricciones que siguen siendo hermitianas). El fallo cuando no conmutan no es raro sino típico: casi ningún par de matrices conmuta, y por eso el mundo cuántico está lleno de incompatibilidades: la no conmutatividad, que en el capítulo parecía una molestia contable, es contenido físico.

7.17. Segunda ronda: Householder, tridiagonales y el laplaciano discreto

7.17.1. Reflexiones de Householder

$H = I - 2\frac{\mathbf{v}\mathbf{v}^\top}{\|\mathbf{v}\|^2}$ (ortogonal, simétrica, involutiva) manda $\mathbf{x} \mapsto \|\mathbf{x}\|\mathbf{e}_1$ con $\mathbf{v} = \mathbf{x} - \|\mathbf{x}\|\mathbf{e}_1$: aplicadas por columnas triangularizan con estabilidad perfecta ($\kappa = 1$): así se calcula QR en serio (Gram–Schmidt pierde ortogonalidad numérica). Ejemplo: $\mathbf{x} = (3, 4)$: $H = \frac{1}{5}\left(\begin{smallmatrix} 3 & 4 \\ 4 & -3 \end{smallmatrix}\right)$, $H\mathbf{x} = (5, 0)$ ✓.

7.17.2. Tridiagonales y el laplaciano discreto

Desarrollo por la primera fila: $D_n = a_n D_{n-1} - b_{n-1}c_{n-1}D_{n-2}$: determinantes en tiempo lineal (y con $a = 1$, $bc = -1$: ¡Fibonacci!).

Ejemplo 7.114 (El laplaciano discreto, diagonalizado exacto)**.** La matriz de segunda diferencia (2 en la diagonal, -1 al lado, tamaño n) tiene

$$\lambda_k = 4\sin^2\frac{k\pi}{2(n+1)}, \qquad v_k = \Big(\sin\frac{jk\pi}{n+1}\Big)_j,$$

verificable con $\sin(j\pm1)\theta$. Es el puente exacto entre capítulos: los autovectores son los modos de Fourier discretos (capítulo 12), los autovalores discretizan $(\frac{k\pi}{L})^2$, y $\kappa \sim \frac{4}{\pi^2}n^2$ explica por qué las mallas finas son duras. Álgebra numérica, EDP y Fourier en un ejemplo.

7.17.3. La matriz de Hilbert

$H_{ij} = \frac{1}{i+j-1}$: Gram de $\{x^k\}$ en $[0,1]$ (definida positiva por la caracterización (vi)) con $\kappa \sim e^{3.5n}$: ajustar polinomios en la base $\{x^k\}$ es suicidio numérico —por eso existen los polinomios ortogonales de Legendre: la misma proyección en base sana—. Elegir base no es estética: es supervivencia.

7.18. Coda: perturbación de autovalores

¿Cuánto se mueven los autovalores al perturbar la matriz? Para simétricas, la respuesta es limpia (*teorema de Weyl*, corolario de Courant–Fischer): $|\lambda_k(A+E)-\lambda_k(A)| \le \|E\|$: *los autovalores simétricos son Lipschitz en la matriz*: perfectamente condicionados. Para no simétricas puede ser catastrófico: en un bloque de Jordan de tamaño n, perturbar la esquina con ε mueve los autovalores $\varepsilon^{1/n}$ (¡la raíz n-ésima!): con $n=16$ y $\varepsilon = 10^{-16}$, los autovalores se mueven 0.1. Por eso el software nunca calcula la forma de Jordan numéricamente (es discontinua) y trabaja con Schur; y el *pseudoespectro* (los autovalores de todas las perturbaciones pequeñas) es el objeto correcto para operadores no normales. Complemento algorítmico: la *iteración del cociente de Rayleigh* (invertir sobre $A - R(x_k)I$) converge *cúbicamente* para simétricas: triplica dígitos por paso: el refinamiento final de la iteración de potencias.

7.19. Temas extra y ejemplos desarrollados

7.19.1. Cadenas de Markov y PageRank

Ejemplo 7.115 (Clima)**.** $P = \left(\begin{smallmatrix} 0.9 & 0.5 \\ 0.1 & 0.5 \end{smallmatrix}\right)$: autovalores 1 y 0.4; estacionaria $(\frac{5}{6}, \frac{1}{6})$; memoria semanal $0.4^7 \approx 0.2$ %. PageRank es esto a escala web, con Perron–Frobenius garantizando unicidad.

7.19.2. Vandermonde e interpolación

$\det V = \prod_{i>j}(x_i - x_j) \neq 0$: existencia y unicidad del interpolador (inducción restando columnas). La fórmula de Lagrange lo da sin sistemas; Simpson y el secreto compartido de Shamir viven aquí. Nótese: $\det V$ es, salvo cuadrado, el discriminante del capítulo 26.

7.19.3. Círculos de Gershgorin

Todo autovalor vive en $\bigcup_i \{|z - A_{ii}| \le \sum_{j\neq i} |A_{ij}|\}$ (mirar la fila de la componente máxima del autovector). Dominancia diagonal ⇒ invertible gratis; estocástica ⇒ espectro en el disco unitario.

7.19.4. La matriz de proyección

$P = A(A^{\mathsf{T}}A)^{-1}A^{\mathsf{T}}$: $P^2 = P = P^{\mathsf{T}}$ (y esas dos propiedades caracterizan); espectro $\{0, 1\}$; mínimos cuadrados: $\hat{y} = Pb$; en estadística es la matriz sombrero y su traza son los grados de libertad.

7.19.5. Rotaciones en 3D

Ortogonal con det = 1: eje = autovector de $\lambda = 1$ (Euler) y $\operatorname{tr} R = 1 + 2\cos\theta$. Ejemplo: componer giros de 90° en x y luego z da la permutación cíclica $\left(\begin{smallmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{smallmatrix}\right)$, traza 0: $\cos\theta = -\frac{1}{2}$: ¡un giro de 120° alrededor de $(1, 1, 1)$!

7.20. La mejor aproximación de rango bajo: Eckart–Young y el análisis de componentes principales

Hemos afirmado varias veces que truncar la descomposición en valores singulares da "la mejor aproximación de rango k" y lo hemos usado con números. Es un teorema, y con Courant–Fischer ya demostrado su prueba es corta.

Teorema 7.116 (Eckart–Young–Mirsky)**.** *Sea $A = \sum_{i=1}^{r} \sigma_i u_i v_i^{\top}$ la SVD, con $\sigma_1 \ge \cdots \ge \sigma_r > 0$, y sea $A_k = \sum_{i\le k} \sigma_i u_i v_i^{\top}$. Entonces, para toda B de rango $\le k$,*

$$\|A - B\|_2 \;\ge\; \|A - A_k\|_2 = \sigma_{k+1}, \qquad \|A - B\|_F^2 \;\ge\; \|A - A_k\|_F^2 = \sum_{i>k} \sigma_i^2.$$

Caso espectral. Que A_k alcanza esos valores es inmediato: $A - A_k = \sum_{i>k} \sigma_i u_i v_i^{\top}$ tiene valores singulares $\sigma_{k+1}, \ldots, \sigma_r$. Para la cota inferior, sea B de rango $\le k$; entonces $\ker B$ tiene dimensión $\ge n - k$, y el subespacio $W = \operatorname{span}(v_1, \ldots, v_{k+1})$ tiene dimensión $k + 1$. Como $(n-k)+(k+1) > n$, las dimensiones obligan a que se corten: hay $w \in \ker B \cap W$ con $\|w\| = 1$. Para ese w,

$$\|A - B\|_2^2 \;\ge\; \|(A - B)w\|^2 = \|Aw\|^2 = \sum_{i\le k+1} \sigma_i^2 \langle v_i, w\rangle^2 \;\ge\; \sigma_{k+1}^2. \qquad \square$$

Ejemplo 7.117 (Un caso mínimo, comprobado)**.** Sea

$$A = \begin{pmatrix} 3 & 0 \\ 0 & 2 \\ 0 & 0 \end{pmatrix}, \qquad \sigma_1 = 3,\ \sigma_2 = 2.$$

La mejor aproximación de rango 1 es $A_1 = 3u_1 v_1^\top$, es decir, A con el 2 puesto a cero. Y en efecto

$$\|A - A_1\|_2 = 2 = \sigma_2, \qquad \|A - A_1\|_F = 2 = \sqrt{\sigma_2^2}.$$

Ninguna matriz de rango 1, por ingeniosa que sea, se acerca más.

Componentes principales

Definición 7.118. Dada una nube de puntos centrada, con matriz de datos X (n filas, p columnas) y covarianza muestral $S = \frac{1}{n-1} X^\top X$, las **componentes principales** son los autovectores de S ordenados por autovalor decreciente. La *varianza explicada* por las k primeras es $\sum_{i \le k} \lambda_i / \sum_i \lambda_i$.

Y ahí está la conexión: los autovectores de $X^\top X$ son los v_i de la SVD de X, y $\lambda_i = \sigma_i^2/(n-1)$. Por Eckart–Young, proyectar sobre las k primeras componentes es *exactamente* la proyección lineal de dimensión k que menos error cuadrático comete. El análisis de componentes principales no es una heurística: es un teorema de aproximación matricial.

Ejemplo 7.119 (Dos variables correlacionadas)**.** Sea la covarianza

$$S = \begin{pmatrix} 4 & 2 \\ 2 & 3 \end{pmatrix}.$$

Su polinomio característico es $\lambda^2 - 7\lambda + 8$, con raíces

$$\lambda_1 = \frac{7 + \sqrt{17}}{2} = 5{,}561\,553, \qquad \lambda_2 = \frac{7 - \sqrt{17}}{2} = 1{,}438\,447.$$

La primera componente explica

$$\frac{5{,}561\,553}{7} = 0{,}794\,5,$$

es decir, el 79,45 por ciento de la variabilidad total. Proyectando los datos sobre esa única dirección se pierde solo un 20,5 por ciento de la varianza, y a cambio se pasa de dos dimensiones a una. Ese es el compromiso que el método propone, y Eckart–Young garantiza que no hay ninguna dirección mejor.

Obsérvese además que la traza se conserva: $\lambda_1 + \lambda_2 = 7 = 4 + 3$. La varianza total es un invariante que las componentes se limitan a *repartir* de la manera más desigual posible.

7.21. Autovalores complejos de matrices reales: complexificación y forma canónica real

Todo el capítulo trabaja sobre $\mathbb{R}$, y ahí hay matrices sin ningún autovalor: un giro de 90° no deja ninguna dirección invariante. La solución es subir a $\mathbb{C}$, hacer el trabajo y volver.

Definición 7.120 (Complexificación)**.** Para un $\mathbb{R}$-espacio V, se define $V_{\mathbb{C}} = V \oplus iV$, con la multiplicación por escalares complejos obvia. Toda base real de V es base compleja de $V_{\mathbb{C}}$, y toda aplicación lineal T se extiende a $T_{\mathbb{C}}$ con la *misma* matriz.

Teorema 7.121 (Los autovalores van en pares)**.** *Si A es real y $Av = \lambda v$ con $\lambda = a+bi$ y $b \neq 0$, entonces $A\bar{v} = \bar{\lambda}\bar{v}$: los autovalores complejos aparecen en pares conjugados, con autovectores conjugados. Además,* $\operatorname{Re} v$ *e* $\operatorname{Im} v$ *son* reales *y linealmente independientes, y generan un plano invariante por A.*

Teorema 7.122 (Forma canónica real)**.** *En la base* $\{\operatorname{Re} v,\ \operatorname{Im} v\}$ *de ese plano, la matriz de A es*

$$\begin{pmatrix} a & b \\ -b & a \end{pmatrix} = r \begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix}, \qquad r = |\lambda|,\ \theta = \arg\lambda.$$

Toda matriz real es, pues, semejante sobre $\mathbb{R}$ a una matriz por bloques con escalares en los autovalores reales y bloques de giro-escalado 2×2 en cada par conjugado.

Ejemplo 7.123 (Un giro escondido)**.** Sea

$$A = \begin{pmatrix} 3 & -2 \\ 4 & -1 \end{pmatrix}.$$

Su polinomio característico es $\lambda^2 - 2\lambda + 5$, con raíces $\lambda = 1 \pm 2i$: no hay autovalores reales, y sobre $\mathbb{R}$ la matriz parece indescifrable.

Un autovector para $\lambda = 1+2i$ es $v = (1,1)+i(0,-1)$, es decir $\operatorname{Re} v = (1,1)$ y $\operatorname{Im} v = (0,-1)$. Con $P = \left(\begin{smallmatrix} 1 & 0 \\ 1 & -1 \end{smallmatrix}\right)$:

$$P^{-1}AP = \begin{pmatrix} 1 & 2 \\ -2 & 1 \end{pmatrix} = \sqrt{5} \begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix}, \qquad \theta = \arctan 2 = 1{,}107\,149 \text{ rad} = 63{,}43^\circ.$$

Así que A *es* un giro de $63{,}43^\circ$ compuesto con una dilatación de factor $\sqrt{5} = 2{,}236\,068$, escrito en una base oblicua. Todo lo que parecía irregular era un giro mal mirado.

Observación 7.124 (Por qué importa en las ecuaciones diferenciales)*.* La solución de $\dot{\mathbf{x}} = A\mathbf{x}$ es $e^{tA}\mathbf{x}_0$, y con la forma real

$$e^{tA} = e^{at} \begin{pmatrix} \cos bt & \sin bt \\ -\sin bt & \cos bt \end{pmatrix}$$

en cada bloque. De ahí sale, sin números complejos en la respuesta final, que las trayectorias sean espirales que se abren si $a > 0$, se cierran si $a < 0$ y son elipses si $a = 0$. La parte real del autovalor gobierna la estabilidad y la imaginaria la frecuencia de giro: los dos hechos que el capítulo de ecuaciones diferenciales usa constantemente, ahora justificados.

7.22. Producto de Kronecker, el operador vec y las ecuaciones de Sylvester y Lyapunov

El producto tensorial se presentó de forma abstracta. En coordenadas tiene una cara muy concreta y sorprendentemente útil.

Definición 7.125. Para A de tamaño $m\times n$ y B de tamaño $p\times q$, el **producto de Kronecker** es la matriz $mp \times nq$ por bloques

$$A \otimes B = \begin{pmatrix} a_{11}B & \cdots & a_{1n}B \\ \vdots & & \vdots \\ a_{m1}B & \cdots & a_{mn}B \end{pmatrix},$$

y $\operatorname{vec} X$ es el vector que se obtiene apilando las columnas de X.

Teorema 7.126 (Las tres identidades que hacen falta)**.**

$$(A \otimes B)(C \otimes D) = AC \otimes BD, \qquad (A \otimes B)^{-1} = A^{-1} \otimes B^{-1},$$

$$\operatorname{vec}(AXB) = (B^{\top} \otimes A) \operatorname{vec} X.$$

Si A tiene autovalores λ_i y B tiene μ_j, entonces $A\otimes B$ tiene autovalores $\lambda_i\mu_j$ y $A\otimes I + I\otimes B$ tiene $\lambda_i + \mu_j$.

La tercera identidad es la que convierte una ecuación *matricial* en un sistema lineal ordinario, y por tanto en algo que las técnicas del capítulo ya resuelven.

Ejemplo 7.127 (Sylvester: cuándo hay solución y cuándo no)**.** La ecuación $AX + XB = C$ se vectoriza como

$$\left(I \otimes A + B^{\top} \otimes I\right) \operatorname{vec} X = \operatorname{vec} C,$$

cuyos autovalores son $\lambda_i + \mu_j$. Luego *hay solución única si y solo si A y $-B$ no comparten ningún autovalor.*

(a) Caso regular. Con $A = \operatorname{diag}(1, 2)$, $B = (3)$ y $C = \binom{1}{1}$, la matriz del sistema es $\operatorname{diag}(1+3,\ 2+3) = \operatorname{diag}(4, 5)$, invertible. La solución es

$$X = \begin{pmatrix} 0{,}25 \\ 0{,}2 \end{pmatrix}, \qquad \text{y en efecto } AX + XB = \begin{pmatrix} 1 \\ 1 \end{pmatrix}.\checkmark$$

(b) Caso singular. Con $B = (-1)$, los autovalores del sistema son $1 - 1 = 0$ y $2 - 1 = 1$: hay un cero, la matriz es singular y la ecuación o no tiene solución o tiene infinitas. El criterio espectral lo predice sin intentar resolver nada.

Observación 7.128 (Lyapunov y la estabilidad)*.* El caso $B = A^{\top}$ tiene nombre propio:

$$A^{\top}P + PA = -Q.$$

El teorema de Lyapunov dice que si Q es definida positiva, la ecuación tiene solución P definida positiva *si y solo si* todos los autovalores de A tienen parte real negativa. Es decir: la estabilidad de $\dot{\mathbf{x}} = A\mathbf{x}$ se decide resolviendo una ecuación lineal, sin calcular autovalores. Con $V(\mathbf{x}) = \mathbf{x}^{\top}P\mathbf{x}$ se tiene $\dot{V} = -\mathbf{x}^{\top}Q\mathbf{x} < 0$: la forma cuadrática V es una función de energía que decrece siempre, y esa es la demostración. Toda la teoría de control lineal está construida sobre esta ecuación.

7.23. Proyecciones oblicuas, la fórmula de Grassmann y las bases de Hamel

Definición 7.129 (Idempotente)**.** $P : V \to V$ lineal es una **proyección** si $P^2 = P$. Entonces

$$V = \operatorname{im} P \ \oplus \ \ker P,$$

y P es la proyección sobre $\operatorname{im} P$ *a lo largo de* $\ker P$. Es *ortogonal* si además $P = P^\top$; en otro caso se llama oblicua.

La descomposición sale de escribir $v = Pv + (v - Pv)$: el primero está en la imagen y el segundo en el núcleo, porque $P(v - Pv) = Pv - P^2v = 0$. Y la suma es directa porque un vector en ambos cumple $v = Pw$ y $Pv = 0$, luego $v = Pw = P^2w = Pv = 0$.

Ejemplo 7.130 (Una proyección que no es ortogonal)**.**

$$P = \begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix}, \qquad P^2 = P, \qquad P \neq P^\top.$$

Su imagen es $\langle e_1 \rangle$ y su núcleo es $\langle (1, -1) \rangle$, pues $P(1, -1)^\top = (0, 0)^\top$. Proyecta sobre el eje X *a lo largo* de la recta $y = -x$, no perpendicularmente. Ambas proyecciones sobre el mismo subespacio existen y son distintas; lo que fija cuál es el *complemento* elegido, y solo la ortogonal minimiza la distancia.

Teorema 7.131 (Fórmula de Grassmann)**.** *Para subespacios de dimensión finita,*

$$\dim(U + W) = \dim U + \dim W - \dim(U \cap W).$$

Demostración. Tómese una base $\{c_1, \dots, c_k\}$ de $U \cap W$ y complétese a una base de U con $\{u_1, \dots, u_r\}$ y a una de W con $\{w_1, \dots, w_s\}$. La unión de las tres familias genera $U + W$; y es independiente, porque una combinación nula obliga a que $\sum \beta_i u_i = -\sum c - \sum \gamma_j w_j \in U \cap W$, lo que fuerza $\beta = 0$ por la elección de la base de U, y luego el resto también. □

Ejemplo 7.132. En $\mathbb{R}^4$, con $U = \langle e_1, e_2, e_3 \rangle$ y $W = \langle e_3, e_4 \rangle$: $\dim(U + W) = 4$, luego $\dim(U \cap W) = 3 + 2 - 4 = 1$, y en efecto la intersección es $\langle e_3 \rangle$. Y si dos planos de $\mathbb{R}^3$ son distintos, $3 \geq 2 + 2 - \dim(U \cap W)$ obliga a $\dim(U \cap W) \geq 1$: dos planos distintos del espacio se cortan siempre en al menos una recta. La geometría elemental como corolario del recuento de dimensiones.

Dimensión infinita: bases de Hamel

Teorema 7.133. *Todo espacio vectorial tiene base, y todo conjunto linealmente independiente se extiende a una.*

Demostración. Considérese el conjunto de los subconjuntos linealmente independientes que contienen al dado, ordenado por inclusión. La unión de una cadena de independientes es

independiente —toda combinación lineal involucra un número *finito* de vectores, luego cabe en algún eslabón—, así que toda cadena tiene cota superior. Por el lema de Zorn hay un maximal, y un independiente maximal es generador: si algún v no estuviera en su envolvente, añadirlo daría uno mayor. □

Cuidado

El teorema es un puro existencial y no exhibe nada. $\mathbb{R}$ como espacio vectorial sobre $\mathbb{Q}$ tiene base —necesariamente no numerable—, pero nadie ha escrito ni escribirá un solo elemento de ella; de hecho, su existencia depende del axioma de elección y de ella salen las funciones aditivas no lineales y los conjuntos no medibles.

Y hay que distinguir dos nociones que en dimensión finita coinciden. Una **base de Hamel** exige combinaciones *finitas*: $\{1, x, x^2, \dots\}$ es base de Hamel de $K[x]$, pero *no* de $K[[x]]$, porque una serie formal no es una combinación finita de monomios. Una **base topológica** admite series convergentes, y es la noción útil en análisis funcional: $\{e_n\}$ es base ortonormal de ℓ^2 en ese sentido, y no es base de Hamel.

De ahí también que en dimensión infinita V y V^{**} dejen de ser isomorfos, aunque la inyección canónica siga existiendo. La palabra "canónico" del teorema de doble dualidad significa precisamente eso: la aplicación $v \mapsto \mathrm{ev}_v$ no necesita elegir base, y por eso sobrevive; la biyección con V^* sí la necesita, y por eso no.

7.24. Las demostraciones pendientes: espectral, rangos y Jordan

Tres resultados centrales quedaron con la demostración a medias o esbozada. Se completan aquí, porque de ellos dependen las formas cuadráticas, la SVD y todo el cálculo con matrices no diagonalizables.

Teorema 7.134 (Espectral real, versión completa)**.** *Sea A una matriz real simétrica $n \times n$. Entonces existe una matriz* ortogonal *Q (es decir, $Q^{\top}Q = I$) tal que*

$$Q^{\top} A Q = \operatorname{diag}(\lambda_1, \dots, \lambda_n),$$

con todos los λ_i reales. Es decir, $\mathbb{R}^n$ tiene una base ortonormal de autovectores de A.

Demostración. Por inducción sobre n. El caso $n = 1$ es trivial.

Paso 1: existe un autovalor real con autovector real. Considerada como matriz compleja, A tiene un autovalor $\lambda \in \mathbb{C}$ con autovector $v \neq 0$. Entonces

$$\lambda\, \overline{v}^{\top} v = \overline{v}^{\top} A v = \overline{(A\overline{v})}^{\top} v = \overline{\overline{(Av)}}^{\top} v = \overline{\lambda}\, \overline{v}^{\top} v,$$

usando que A es real y simétrica; como $\overline{v}^{\top} v = \|v\|^2 > 0$, resulta $\lambda = \overline{\lambda}$, es decir $\lambda \in \mathbb{R}$. Y siendo λ real, el sistema $(A - \lambda I)x = 0$ tiene coeficientes reales y solución no trivial, luego tiene una solución real v_1, que se normaliza a $\|v_1\| = 1$.

Paso 2: el complemento ortogonal es invariante. Sea $W = v_1^{\perp}$, de dimensión $n - 1$. Para $w \in W$,

$$\langle Aw, v_1 \rangle = \langle w, A^{\top} v_1 \rangle = \langle w, A v_1 \rangle = \lambda_1 \langle w, v_1 \rangle = 0,$$

donde la primera igualdad es la definición de traspuesta y la segunda usa $A^\top = A$. Luego $Aw \in W$: la restricción $A|_W$ está bien definida.

Paso 3: inducción. Tómese una base ortonormal de W; en ella, la matriz de $A|_W$ es simétrica (porque $\langle Aw_1, w_2\rangle = \langle w_1, Aw_2\rangle$ se hereda al restringir). Por hipótesis de inducción, W tiene una base ortonormal $v_2, \dots, v_n$ de autovectores de $A|_W$, que lo son también de A. Entonces $v_1, \dots, v_n$ es una base ortonormal de $\mathbb{R}^n$ formada por autovectores, y la matriz Q que los tiene por columnas es ortogonal y cumple $AQ = Q\Lambda$. □

Corolario 7.135 (Formas cuadráticas y ejes principales)**.** *Toda forma cuadrática $q(x) = x^\top Ax$ con A simétrica se escribe, en las coordenadas $y = Q^\top x$, como $q = \sum_i \lambda_i y_i^2$. En consecuencia q es definida positiva si y sólo si todos los $\lambda_i > 0$, y los valores extremos de q sobre la esfera unidad son $\lambda_{\text{mín}}$ y $\lambda_{\text{máx}}$.*

Teorema 7.136 (Rango fila igual a rango columna)**.** *Para toda matriz $A \in \mathbb{R}^{m\times n}$, la dimensión del espacio generado por sus filas coincide con la del generado por sus columnas.*

Demostración. Sea r la dimensión del espacio de columnas y $c_1, \dots, c_r$ una base suya. Fórmese la matriz $C \in \mathbb{R}^{m\times r}$ con esas columnas. Como cada columna a_j de A es combinación de la base, existen escalares r_{ij} con $a_j = \sum_i r_{ij} c_i$, es decir

$$A = C\,R, \qquad R = (r_{ij}) \in \mathbb{R}^{r\times n}.$$

Ahora léase la misma factorización *por filas*: la fila k de A es $\sum_i C_{ki}$ (fila i de R), luego el espacio de filas de A está contenido en el espacio generado por las r filas de R, y por tanto tiene dimensión $\le r$. Es decir,

$$\dim(\text{filas}) \le \dim(\text{columnas}).$$

Aplicando lo mismo a $A^\top$, que intercambia filas y columnas, se obtiene la desigualdad contraria. □

Observación 7.137. La factorización $A = CR$ no es un truco auxiliar: es la forma más económica de guardar una matriz de rango bajo ($mr + rn$ números en vez de mn), y es el esqueleto de la compresión por rango reducido y de la aproximación de rango k que la SVD optimiza. Que las dos dimensiones coincidan es, mirado así, la afirmación de que hay una única noción de "cuánta información contiene A".

Teorema 7.138 (Forma de Jordan: existencia)**.** *Sea $T : V \to V$ un endomorfismo de un espacio vectorial complejo de dimensión finita. Entonces existe una base de V en la que la matriz de T es diagonal por bloques, con bloques de Jordan*

$$J_k(\lambda) = \begin{pmatrix} \lambda & 1 & & \\ & \lambda & \ddots & \\ & & \ddots & 1 \\ & & & \lambda \end{pmatrix} \in \mathbb{C}^{k\times k}.$$

La lista de bloques, con sus tamaños y autovalores, es única salvo el orden.

Demostración. Paso 1: reducción al caso nilpotente. El polinomio característico se factoriza sobre $\mathbb{C}$ como $\prod_i (x-\lambda_i)^{m_i}$ con los λ_i distintos. Los subespacios *generalizados* $V_i = \ker(T-\lambda_i I)^{m_i}$ cumplen $V = \bigoplus_i V_i$ y $T(V_i) \subset V_i$ —es la descomposición primaria, consecuencia de que los factores $(x-\lambda_i)^{m_i}$ son coprimos dos a dos y de la identidad de Bézout aplicada a polinomios—. Sobre cada V_i, el operador $N_i = T - \lambda_i I$ es nilpotente. Basta pues probar el teorema para operadores nilpotentes: los bloques que salgan, sumados λ_i en la diagonal, son los de T.

Paso 2: el caso nilpotente, por inducción sobre $\dim V$. Sea N nilpotente en V, con $V \neq 0$. Como N no es inyectivo, $W = N(V)$ tiene dimensión estrictamente menor, y $N(W) \subset W$. Por hipótesis de inducción, W admite una base de Jordan, es decir, se descompone en cadenas

$$w_j,\ Nw_j,\ N^2 w_j,\ \ldots,\ N^{k_j-1} w_j \qquad (j = 1, \ldots, s),$$

donde $N^{k_j} w_j = 0$, y esos vectores juntos forman una base de W.

Paso 3: levantar las cadenas. Cada w_j está en $W = N(V)$, luego existe $v_j \in V$ con $Nv_j = w_j$. Las cadenas alargadas

$$v_j,\ Nv_j = w_j,\ \ldots,\ N^{k_j} v_j$$

siguen siendo independientes entre sí (una relación entre ellas, aplicada N, daría una relación en W).

Paso 4: completar el núcleo. Los vectores $N^{k_j} v_j$ construidos están en $\ker N$, y son s vectores independientes de él. Complétense a una base de $\ker N$ con vectores $u_1, \ldots, u_t$; cada uno de ellos es una cadena de longitud 1.

Paso 5: recuento. La colección obtenida tiene $\sum_j (k_j+1) + t = \dim W + s + t$ vectores. Por el teorema del rango, $\dim V = \dim W + \dim \ker N$, y $\dim \ker N = s + t$ por construcción; luego el número de vectores es exactamente $\dim V$. Siendo independientes, forman una base, y en ella N actúa como una suma de bloques de Jordan nilpotentes.

Unicidad. El número de bloques de tamaño $\geq k$ asociados a λ es $\operatorname{rango}(T-\lambda I)^{k-1} - \operatorname{rango}(T-\lambda I)^k$, cantidad que no depende de ninguna elección; de ahí se despeja el número de bloques de cada tamaño exacto. □

Observación 7.139 (Por qué el algoritmo de cálculo termina). La receta práctica —calcular los rangos de las potencias $(A-\lambda I)^k$ hasta que se estabilicen, y leer de ahí los tamaños de los bloques— queda ahora justificada por la fórmula de unicidad: los rangos determinan la estructura, y la estabilización ocurre a más tardar en $k = m_\lambda$, la multiplicidad algebraica, porque el subespacio generalizado tiene esa dimensión. Sin el teorema de existencia, el algoritmo podría en principio dar números incoherentes; con él, está garantizado que la lista de tamaños que se lee es la de una descomposición real.

7.25. Condicionamiento, pseudoinversa y Gram–Schmidt, explícitos

Tres objetos que las reglas del capítulo mandan usar y que sólo aparecían en prosa, a veces cientos de líneas después de haberse necesitado.

Definición 7.140 (Número de condición)**.** Para una matriz invertible A y una norma matricial subordinada, el *número de condición* es

$$\kappa(A) = \|A\| \cdot \|A^{-1}\| \;\geq\; 1.$$

Con la norma euclídea, $\kappa_2(A) = \sigma_{\text{máx}}/\sigma_{\text{mín}}$, el cociente entre el mayor y el menor valor singular. Si A no es invertible se conviene $\kappa = \infty$.

Teorema 7.141 (Qué mide, exactamente)**.** *Si $Ax = b$ y $A(x + \Delta x) = b + \Delta b$, entonces*

$$\frac{\|\Delta x\|}{\|x\|} \;\leq\; \kappa(A)\,\frac{\|\Delta b\|}{\|b\|},$$

y la cota se alcanza. Es decir, $\kappa(A)$ es el factor máximo por el que un error relativo en los datos se amplifica en la solución.

Demostración. De $A\Delta x = \Delta b$ sale $\|\Delta x\| \leq \|A^{-1}\|\,\|\Delta b\|$, y de $b = Ax$ sale $\|b\| \leq \|A\|\,\|x\|$, o sea $\frac{1}{\|x\|} \leq \frac{\|A\|}{\|b\|}$. Multiplicando ambas desigualdades se obtiene el enunciado. La igualdad se alcanza eligiendo b en la dirección del valor singular máximo y Δb en la del mínimo. □

Ejemplo 7.142 (Un sistema de dos ecuaciones que no se puede resolver)**.** Sea

$$A = \begin{pmatrix} 1 & 1 \\ 1 & 1{,}0001 \end{pmatrix}, \qquad b = \begin{pmatrix} 2 \\ 2{,}0001 \end{pmatrix},$$

cuya solución exacta es $x = (1, 1)$. Perturbando b a $(2;\, 2{,}0002)$ —un cambio relativo de $5 \cdot 10^{-5}$— la solución pasa a ser $(0, 2)$: ha cambiado por completo. Aquí $\kappa_2(A) \approx 4 \cdot 10^4$, y la regla práctica es que *se pierden* $\log_{10}\kappa$ *cifras significativas*: con $\kappa = 4 \cdot 10^4$ se pierden entre cuatro y cinco, de modo que en doble precisión (16 cifras) quedan unas once buenas. Con $\kappa \approx 10^{16}$ no queda ninguna, y el ordenador devuelve ruido con toda naturalidad.

Definición 7.143 (Pseudoinversa de Moore–Penrose)**.** Dada $A \in \mathbb{R}^{m\times n}$ con descomposición en valores singulares $A = U\Sigma V^{\top}$, su *pseudoinversa* es

$$A^{+} = V\,\Sigma^{+}\,U^{\top},$$

donde Σ^{+} se obtiene de Σ transponiendo e invirtiendo los valores singulares *no nulos* y dejando en cero los demás. Es la única matriz que cumple las cuatro condiciones de Penrose:

$$AA^{+}A = A, \qquad A^{+}AA^{+} = A^{+}, \qquad (AA^{+})^{\top} = AA^{+}, \qquad (A^{+}A)^{\top} = A^{+}A.$$

Observación 7.144 (Qué resuelve)*.* El vector $x = A^{+}b$ es, entre todos los que minimizan $\|Ax - b\|$, el de *norma mínima*. Eso cubre los tres casos de un solo golpe: si A es invertible, $A^{+} = A^{-1}$; si el sistema es sobredeterminado y de rango completo, $A^{+} = (A^{\top}A)^{-1}A^{\top}$ y se recupera la solución de mínimos cuadrados; y si está indeterminado, elige entre las infinitas soluciones la más corta. Calcularla por la SVD y no por las ecuaciones normales es lo que recomienda la regla del capítulo, y la razón es cuantitativa: $\kappa(A^{\top}A) = \kappa(A)^2$, de modo que formar $A^{\top}A$ *duplica* el número de cifras que se pierden.

Regla: Ortonormalizar por Gram–Schmidt

Dados vectores independientes $v_1, \dots, v_k$:

1. $u_1 = v_1$, $e_1 = u_1/\|u_1\|$.
2. Para $j = 2, \dots, k$: resta a v_j sus proyecciones sobre lo ya construido,
$$u_j = v_j - \sum_{i<j} \langle v_j, e_i \rangle \, e_i,$$
y normaliza: $e_j = u_j/\|u_j\|$.
3. Si algún u_j sale nulo, los vectores no eran independientes: ese v_j está en el generado por los anteriores, y se descarta.
4. **Comprueba** que $\langle e_i, e_j \rangle = \delta_{ij}$ y que $\operatorname{gen}(e_1, \dots, e_j) = \operatorname{gen}(v_1, \dots, v_j)$ para cada j: esa igualdad de subespacios *parciales* es la propiedad característica del método, y la que hace que $A = QR$ con R triangular superior.

En cálculo numérico, usa la versión modificada. En lugar de restar todas las proyecciones de golpe usando v_j, réstalas una a una actualizando el vector:

$$w \leftarrow v_j; \qquad \text{para } i = 1, \dots, j-1: \quad w \leftarrow w - \langle w, e_i \rangle e_i.$$

Matemáticamente es lo mismo; numéricamente, la versión clásica pierde la ortogonalidad de forma catastrófica cuando los vectores están casi alineados, y la modificada no. En bibliotecas serias ni siquiera se usa Gram–Schmidt, sino reflexiones de Householder.

Ejemplo 7.145 (Tres vectores en $\mathbb{R}^3$)**.** Sean $v_1 = (1,1,0)$, $v_2 = (1,0,1)$, $v_3 = (0,1,1)$.

$$u_1 = (1,1,0), \qquad e_1 = \tfrac{1}{\sqrt{2}}(1,1,0).$$

$\langle v_2, e_1 \rangle = \frac{1}{\sqrt{2}}$, luego $u_2 = (1,0,1) - \frac{1}{2}(1,1,0) = (\frac{1}{2}, -\frac{1}{2}, 1)$, con $\|u_2\| = \sqrt{3/2}$ y $e_2 = \frac{1}{\sqrt{6}}(1,-1,2)$. Finalmente $\langle v_3, e_1 \rangle = \frac{1}{\sqrt{2}}$ y $\langle v_3, e_2 \rangle = \frac{1}{\sqrt{6}}$, de donde $u_3 = (0,1,1) - \frac{1}{2}(1,1,0) - \frac{1}{6}(1,-1,2) = (-\frac{2}{3}, \frac{2}{3}, \frac{2}{3})$ y $e_3 = \frac{1}{\sqrt{3}}(-1,1,1)$. Comprobación: los tres son unitarios y ortogonales dos a dos.

7.26. Ejercicios resueltos

Ejercicio 7.1. Resuelve y clasifica: $\begin{cases} x + 2y - z = 3 \\ 2x + 5y + z = 10 \\ x + 4y + 5z = 11 \end{cases}$, y halla la factorización LU de su matriz de coeficientes.

Solución. Escalonando la ampliada con $F_2 - 2F_1$, $F_3 - F_1$ y luego $F_3 - 2F_2$:

$$\left(\begin{array}{ccc|c} 1 & 2 & -1 & 3 \\ 2 & 5 & 1 & 10 \\ 1 & 4 & 5 & 11 \end{array}\right) \to \left(\begin{array}{ccc|c} 1 & 2 & -1 & 3 \\ 0 & 1 & 3 & 4 \\ 0 & 0 & 0 & 0 \end{array}\right).$$

Rangos iguales a $2 < 3$: compatible indeterminado. Con $z = t$: $y = 4 - 3t$, $x = -5 + 7t$: $(x, y, z) = (-5, 4, 0) + t(7, -3, 1)$. Los multiplicadores usados $(2, 1, 2)$ dan

$$L = \begin{pmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ 1 & 2 & 1 \end{pmatrix}, \qquad U = \begin{pmatrix} 1 & 2 & -1 \\ 0 & 1 & 3 \\ 0 & 0 & 0 \end{pmatrix}, \qquad A = LU \text{ (verificable multiplicando).}$$

■

Ejercicio 7.2. Sea $W = \operatorname{gen}\{(1, 2, 0, 1), (2, 4, 1, 3), (0, 0, 1, 1), (1, 2, 1, 2)\} \subset \mathbb{R}^4$. Halla una base de W y su dimensión.

Solución. Colocamos los vectores como filas y escalonamos:

$$\begin{pmatrix} 1 & 2 & 0 & 1 \\ 2 & 4 & 1 & 3 \\ 0 & 0 & 1 & 1 \\ 1 & 2 & 1 & 2 \end{pmatrix} \xrightarrow[F_4 - F_1]{F_2 - 2F_1} \begin{pmatrix} 1 & 2 & 0 & 1 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 1 & 1 \end{pmatrix} \to \begin{pmatrix} 1 & 2 & 0 & 1 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}.$$

Dos filas no nulas: $\dim W = 2$, base $\{(1, 2, 0, 1), (0, 0, 1, 1)\}$. Los otros dos vectores originales son combinaciones de estos. ■

Ejercicio 7.3. Sea $T : \mathbb{R}_2[x] \to \mathbb{R}_2[x]$ dada por $T(p) = p + p'$. Halla su matriz en la base $\{1, x, x^2\}$, demuestra que es un isomorfismo y calcula $T^{-1}(x^2)$.

Solución. $T(1) = 1$, $T(x) = x + 1$, $T(x^2) = x^2 + 2x$. La matriz (columnas = imágenes):

$$M = \begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 2 \\ 0 & 0 & 1 \end{pmatrix}.$$

Triangular con diagonal sin ceros: $\det M = 1 \neq 0$, isomorfismo. Para $T^{-1}(x^2)$ resolvemos $M\mathbf{c} = (0, 0, 1)^{\mathsf{T}}$: de abajo hacia arriba, $c_3 = 1$, $c_2 = -2$, $c_1 = 2$. Así $T^{-1}(x^2) = x^2 - 2x + 2$. Comprobación: $(x^2 - 2x + 2) + (2x - 2) = x^2$. *Nota*: sobre polinomios, "$I + \frac{d}{dx}$" es invertible con inversa $I - \frac{d}{dx} + \frac{d^2}{dx^2} - \cdots$ (serie geométrica de operadores, finita aquí porque derivar es nilpotente). ■

Ejercicio 7.4. Calcula $\det A$ para $A = \begin{pmatrix} 2 & 1 & 0 & 1 \\ 1 & 3 & 1 & 0 \\ 0 & 1 & 2 & 1 \\ 1 & 0 & 1 & 2 \end{pmatrix}$ escalonando.

Solución. $F_2 - \frac{1}{2}F_1$, $F_4 - \frac{1}{2}F_1$: filas $(2, 1, 0, 1), (0, \frac{5}{2}, 1, -\frac{1}{2}), (0, 1, 2, 1), (0, -\frac{1}{2}, 1, \frac{3}{2})$. $F_3 - \frac{2}{5}F_2$, $F_4 + \frac{1}{5}F_2$: $(0, 0, \frac{8}{5}, \frac{6}{5})$, $(0, 0, \frac{6}{5}, \frac{7}{5})$. $F_4 - \frac{3}{4}F_3$: $(0, 0, 0, \frac{7}{5} - \frac{9}{10}) = (0, 0, 0, \frac{1}{2})$. Sin intercambios de filas: $\det A = 2 \cdot \frac{5}{2} \cdot \frac{8}{5} \cdot \frac{1}{2} = 4$. ■

Ejercicio 7.5. Diagonaliza $A = \begin{pmatrix} 2 & 1 & 1 \\ 1 & 2 & 1 \\ 1 & 1 & 2 \end{pmatrix}$ (simétrica) con una base ortonormal de vectores propios.

Solución. Truco: $A = I + J$ donde J es la matriz de unos. Como $J\mathbf{1} = 3 \cdot \mathbf{1}$ (con $\mathbf{1} = (1,1,1)$) y $Jv = 0$ para todo $v \perp \mathbf{1}$, los valores propios de A son $1+3 = 4$ (simple, dirección $\mathbf{1}$) y $1+0 = 1$ (doble, el plano $x+y+z = 0$). Base ortonormal: $q_1 = \frac{1}{\sqrt{3}}(1,1,1)$; en el plano, elegimos $q_2 = \frac{1}{\sqrt{2}}(1,-1,0)$ y $q_3 = \frac{1}{\sqrt{6}}(1,1,-2)$ (ortogonales entre sí y a q_1). Con $Q = [q_1\, q_2\, q_3]$: $Q^{\mathsf{T}}AQ = \operatorname{diag}(4,1,1)$, confirmando el teorema espectral: valores propios reales y ejes ortogonales. ■

Ejercicio 7.6. Verifica Cayley–Hamilton para $A = \begin{pmatrix} 4 & 1 \\ 2 & 3 \end{pmatrix}$ y úsalo para expresar A^4 como combinación de A e I.

Solución. $p(\lambda) = \lambda^2 - 7\lambda + 10$. En efecto: $A^2 = \begin{pmatrix} 18 & 7 \\ 14 & 11 \end{pmatrix}$ y $A^2 - 7A + 10I = \begin{pmatrix} 18-28+10 & 7-7 \\ 14-14 & 11-21+10 \end{pmatrix} = 0$. De $A^2 = 7A - 10I$:

$$A^3 = 7A^2 - 10A = 7(7A - 10I) - 10A = 39A - 70I,$$

$$A^4 = 39A^2 - 70A = 39(7A - 10I) - 70A = 203A - 390I = \begin{pmatrix} 422 & 203 \\ 406 & 219 \end{pmatrix}.$$

■

Ejercicio 7.7. Encuentra la parábola $y = a + bx + cx^2$ que mejor ajusta (mínimos cuadrados) los puntos $(-1,1)$, $(0,0)$, $(1,1)$, $(2,5)$.

Solución.

$$A = \begin{pmatrix} 1 & -1 & 1 \\ 1 & 0 & 0 \\ 1 & 1 & 1 \\ 1 & 2 & 4 \end{pmatrix}, \quad \mathbf{b} = \begin{pmatrix} 1 \\ 0 \\ 1 \\ 5 \end{pmatrix}, \quad A^{\mathsf{T}}A = \begin{pmatrix} 4 & 2 & 6 \\ 2 & 6 & 8 \\ 6 & 8 & 18 \end{pmatrix}, \quad A^{\mathsf{T}}\mathbf{b} = \begin{pmatrix} 7 \\ 10 \\ 22 \end{pmatrix}.$$

Resolviendo las ecuaciones normales (Gauss): la solución exacta es $a = \frac{1}{20}$, $b = \frac{1}{10}$, $c = \frac{5}{4}$; conviene verificar: $A^{\mathsf{T}}A(a,b,c)^{\mathsf{T}} = (\frac{1}{5} + \frac{1}{5} + \frac{15}{2}, \dots)$ — hagámoslo con fracciones: $4a + 2b + 6c = \frac{1}{5} + \frac{1}{5} + \frac{15}{2} = \frac{2}{10} + \frac{2}{10} + \frac{75}{10} = \frac{79}{10} \neq 7$. Corrijamos resolviendo con cuidado: $\begin{cases} 4a + 2b + 6c = 7 \\ 2a + 6b + 8c = 10 \\ 6a + 8b + 18c = 22 \end{cases}$ De la 1^{a}: $2a + b + 3c = \frac{7}{2}$. De la 2^{a}: $a + 3b + 4c = 5$. De la 3^{a}: $3a + 4b + 9c = 11$. Eliminando a ($2^{\text{a}} \times 2 - 1^{\text{a}}$): $5b + 5c = \frac{13}{2} \Rightarrow b + c = \frac{13}{10}$. ($3^{\text{a}} \times 2 - 1^{\text{a}} \times 3$): $2b + 0c = \frac{1}{2} \cdot \dots$ mejor: $6a + 8b + 18c = 22$ menos $3 \times (2a + b + 3c = \frac{7}{2})$: $5b + 9c = \frac{23}{2}$. Restando $5 \times (b + c = \frac{13}{10})$: $4c = \frac{23}{2} - \frac{13}{2} = 5 \Rightarrow c = \frac{5}{4}$. Entonces $b = \frac{13}{10} - \frac{5}{4} = \frac{1}{20}$, y $a = 5 - 3b - 4c = 5 - \frac{3}{20} - 5 = -\frac{3}{20}$. La parábola de mejor ajuste es

$$y = -\frac{3}{20} + \frac{1}{20}x + \frac{5}{4}x^2.$$

Comprobación en las normales, 1^{a} ecuación: $4(-\frac{3}{20}) + 2(\frac{1}{20}) + 6(\frac{5}{4}) = -\frac{12}{20} + \frac{2}{20} + \frac{150}{20} = \frac{140}{20} = 7$. ✓■

Ejercicio 7.8. Clasifica la forma cuadrática $q(x,y,z) = x^2 + 2y^2 + 2z^2 + 2xy + 2yz$.

Solución. Matriz $A = \begin{pmatrix} 1 & 1 & 0 \\ 1 & 2 & 1 \\ 0 & 1 & 2 \end{pmatrix}$. Menores dominantes: $\Delta_1 = 1 > 0$; $\Delta_2 = \det \begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix} = 1 > 0$; $\Delta_3 = \det A = 1 \cdot (4-1) - 1 \cdot (2-0) = 1 > 0$. Por Sylvester, **definida positiva**. (Alternativa: completar cuadrados: $q = (x+y)^2 + (y+z)^2 + z^2$, suma de tres cuadrados independientes.) ■

7.27. Ejercicios adicionales resueltos

Ejercicio 7.9. Cholesky de $A = \begin{pmatrix} 4 & 2 & 2 \\ 2 & 5 & 3 \\ 2 & 3 & 6 \end{pmatrix}$.

Solución. $L = \left(\begin{smallmatrix} 2 & 0 & 0 \\ 1 & 2 & 0 \\ 1 & 1 & 2 \end{smallmatrix}\right)$; $LL^{\mathsf{T}} = A$ ✓; $\det A = (\det L)^2 = 64$; positividad confirmada. ■

Ejercicio 7.10. Jordan de $A = \begin{pmatrix} 5 & 1 \\ -1 & 3 \end{pmatrix}$.

Solución. $(\lambda - 4)^2$; espacio propio 1D con $v = (1, -1)$; generalizado $w = (1, 0)$; $P^{-1}AP = \left(\begin{smallmatrix} 4 & 1 \\ 0 & 4 \end{smallmatrix}\right)$; y $e^{At} = e^{4t}(I + Nt)$ trae el factor t. ■

Ejercicio 7.11. $\det M$ por Schur para $M = \begin{pmatrix} 1 & 0 & 1 & 2 \\ 0 & 1 & 3 & 1 \\ 1 & 1 & 2 & 0 \\ 2 & 0 & 1 & 3 \end{pmatrix}$.

Solución. $A = I$: $\det M = \det(D - CB) = \det\left(\begin{smallmatrix} -2 & -3 \\ -1 & -1 \end{smallmatrix}\right) = -1$. ■

Ejercicio 7.12. κ de $\begin{pmatrix} 1 & 1 \\ 1 & 1.0001 \end{pmatrix}$.

Solución. $\lambda \approx 2$ y 5×10^{-5}: $\kappa \approx 4 \times 10^4$: con 4 cifras de datos, ninguna correcta en la solución. Columnas casi paralelas: base pésima; motivo de la regularización. ■

Ejercicio 7.13. Norma mínima para $x + y + z = 3$.

Solución. $A^+ = \frac{1}{3}(1, 1, 1)^{\mathsf{T}}$: $\mathbf{x}^* = (1, 1, 1)$, la proyección del origen al plano; coincide con Lagrange. ■

Ejercicio 7.14. $T(X) = X^{\mathsf{T}}$ en $M_2(\mathbb{R})$: espectro y determinante.

Solución. $T^2 = \mathrm{id}$: $\lambda = \pm 1$; simétricas (dim 3) y antisimétricas (dim 1): $\det T = -1$, $\operatorname{tr} T = 2$. ■

7.28. Taller de ejemplos explicados: sistemas y valores propios

Aquí resolvemos, con todos los pasos, los dos cálculos que más se piden en álgebra lineal: resolver un sistema por eliminación de Gauss y hallar los valores y vectores propios de una matriz.

7.28.1. Resolver un sistema por eliminación de Gauss

Ejemplo 7.146. Resolver

$$\begin{cases} x + y + z = 6 \\ 2x + y + z = 7 \\ x + 2y + 3z = 14. \end{cases}$$

Paso 1. Eliminar x de las ecuaciones 2 y 3. Restamos la ecuación 1 de la 2:

$$(2x + y + z) - (x + y + z) = 7 - 6 \implies x = 1.$$

¡Suerte! Ya tenemos $x = 1$ directamente.

Paso 2. Sustituir $x = 1$ en las otras dos. La ecuación 1 se vuelve $1 + y + z = 6$, o sea $y + z = 5$. La ecuación 3 se vuelve $1 + 2y + 3z = 14$, o sea $2y + 3z = 13$.

Paso 3. Resolver ese sistema 2×2. De $y + z = 5$ despejamos $y = 5 - z$ y sustituimos en $2y + 3z = 13$:

$$2(5 - z) + 3z = 13 \implies 10 - 2z + 3z = 13 \implies z = 3.$$

Entonces $y = 5 - 3 = 2$.

Solución: $x = 1$, $y = 2$, $z = 3$. (Comprueba en la tercera: $1 + 2(2) + 3(3) = 1 + 4 + 9 = 14$. Correcto.)

7.28.2. Valores y vectores propios de una matriz 2×2

Recuerda: un *vector propio* es una dirección que la matriz solo estira (sin girar), y el *valor propio* λ es cuánto la estira.

Ejemplo 7.147. Hallar los valores y vectores propios de $A = \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}$.

Paso 1. Plantear la ecuación característica $\det(A - \lambda I) = 0$. Restamos λ en la diagonal:

$$\det \begin{pmatrix} 2 - \lambda & 1 \\ 1 & 2 - \lambda \end{pmatrix} = (2 - \lambda)^2 - (1)(1) = 0.$$

Paso 2. Resolver para λ. Desarrollamos: $(2 - \lambda)^2 - 1 = 0$, así que $(2 - \lambda)^2 = 1$, es decir $2 - \lambda = \pm 1$. Esto da $\lambda = 1$ o $\lambda = 3$. Esos son los dos valores propios.

Paso 3. Hallar el vector propio de $\lambda = 3$. Resolvemos $(A - 3I)v = 0$:

$$\begin{pmatrix} -1 & 1 \\ 1 & -1 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = 0 \implies -x + y = 0 \implies y = x.$$

Un vector que cumple $y = x$ es $\begin{pmatrix} 1 \\ 1 \end{pmatrix}$.

Paso 4. Hallar el vector propio de $\lambda = 1$. Resolvemos $(A - I)v = 0$:

$$\begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = 0 \implies x + y = 0 \implies y = -x,$$

con vector propio $\begin{pmatrix} 1 \\ -1 \end{pmatrix}$.

En resumen: A estira por 3 en la dirección $(1,1)$ y por 1 en la dirección $(1,-1)$. Esas dos direcciones son perpendiculares, algo típico de las matrices simétricas.

7.29. Ampliación del temario: geometría en R3, espacios afines, programación lineal y producto tensorial

7.29.1. Geometría analítica en $\mathbb{R}^2$ y $\mathbb{R}^3$

El producto cruz.

Idea clave

Intuición primero. El producto cruz responde dos preguntas a la vez: "¿cuál es la dirección perpendicular a estos dos vectores?" (la que apunta según la regla de la mano derecha) y "¿qué tan grande es el parche que abarcan?" (el área del paralelogramo, que es la longitud del resultado). Por eso vive solo en $\mathbb{R}^3$: es donde "perpendicular a un plano" señala una única dirección.

En $\mathbb{R}^3$, el producto cruz $\mathbf{a} \times \mathbf{b}$ da un vector *perpendicular* a ambos, con magnitud igual al área del paralelogramo que forman.

Regla: producto cruz

$$\mathbf{a} \times \mathbf{b} = \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \end{vmatrix} = (a_2 b_3 - a_3 b_2,\ a_3 b_1 - a_1 b_3,\ a_1 b_2 - a_2 b_1),$$

y $\|\mathbf{a} \times \mathbf{b}\| = \|\mathbf{a}\|\,\|\mathbf{b}\| \sin\theta$ es el área del paralelogramo.

Ejemplo 7.148. $(1,2,3) \times (4,5,6) = (2\cdot 6 - 3\cdot 5,\ 3\cdot 4 - 1\cdot 6,\ 1\cdot 5 - 2\cdot 4) = (-3, 6, -3)$. Comprobación: el producto punto con $(1,2,3)$ da $-3 + 12 - 9 = 0$, así que es perpendicular, como debe.

Rectas y planos. Una *recta* en $\mathbb{R}^3$ se escribe paramétricamente $\mathbf{r} = \mathbf{r}_0 + t\mathbf{v}$ (un punto más múltiplos de una dirección). Un *plano* se escribe con su normal $\mathbf{n}$: $\mathbf{n} \cdot (\mathbf{r} - \mathbf{r}_0) = 0$, es decir $ax + by + cz = d$.

Regla: distancia de un punto a un plano

La distancia del punto $P = (x_0, y_0, z_0)$ al plano $ax + by + cz = d$ es

$$\text{dist} = \frac{|ax_0 + by_0 + cz_0 - d|}{\sqrt{a^2 + b^2 + c^2}}.$$

Ejemplo 7.149. Distancia del origen al plano $x + 2y + 2z = 9$: $\dfrac{|0+0+0-9|}{\sqrt{1+4+4}} = \dfrac{9}{3} = 3$.

7.29.2. Espacios afines: puntos frente a vectores

Idea clave

Intuición primero. Imagina una pizarra: hay *posiciones* (puntos) y *flechas* (vectores). Posición menos posición da una flecha (el desplazamiento entre ellas); posición más flecha da otra posición; pero posición más posición no significa nada, porque no hay un origen privilegiado. Lo que sí tiene sentido es el *promedio* de posiciones, como el punto medio: esa es la idea de combinación afín.

Un *espacio afín* distingue formalmente dos cosas que en $\mathbb{R}^n$ solemos confundir: los *puntos* (posiciones) y los *vectores* (desplazamientos). La resta de dos puntos da un vector ($Q - P = \vec{PQ}$), y un punto más un vector da otro punto ($P + \mathbf{v}$); pero *sumar dos puntos no tiene sentido* por sí solo. Lo que sí tiene sentido es una *combinación afín* $\sum \lambda_i P_i$ con $\sum \lambda_i = 1$ (por ejemplo, el punto medio $\frac{1}{2}P + \frac{1}{2}Q$).

Observación 7.150. Esta distinción, que parece pedante, es la base de los gráficos por computadora y la geometría proyectiva: un modelo 3D guarda puntos, pero las transformaciones (mover, rotar) actúan sobre vectores. El origen no es un punto "especial" en un espacio afín: todos los puntos son iguales, y por eso el espacio afín captura la geometría "sin origen privilegiado", a diferencia del espacio vectorial, donde el $\mathbf{0}$ es especial.

7.29.3. Optimización y programación lineal: el método Simplex

Idea clave

Intuición primero. La región permitida es un polígono, y el objetivo es como una regla recta que deslizas por encima: el último rincón que toca es el máximo. Por eso el óptimo siempre está en un *vértice*. El método Simplex no revisa todos los puntos: camina de esquina en esquina por las aristas, subiendo siempre, hasta que ningún vecino esté más alto. Ahí paró: ese es el óptimo.

La *programación lineal* busca maximizar o minimizar una función lineal (el *objetivo*) sujeta a restricciones lineales (desigualdades). El conjunto factible es un poliedro, y el óptimo siempre está en un *vértice*.

Ejemplo 7.151 (Un problema completo)**.** Maximizar $P = 3x + 5y$ sujeto a $x \le 4$, $2y \le 12$, $3x + 2y \le 18$, con $x, y \ge 0$.

Paso 1. Los vértices del polígono factible son $(0,0)$, $(4,0)$, $(4,3)$, $(2,6)$ y $(0,6)$.

Paso 2. Evaluar el objetivo en cada uno: $P(0,0) = 0$, $P(4,0) = 12$, $P(4,3) = 27$, $P(2,6) = 36$, $P(0,6) = 30$.

Paso 3. El máximo es $P = 36$ en $(2,6)$. El *método Simplex* es el algoritmo que, en vez de probar todos los vértices, camina de vértice en vértice mejorando el objetivo en cada paso,

hasta llegar al óptimo: es lo que hace posible resolver problemas con miles de variables, y es uno de los algoritmos más usados de la historia (logística, dietas, producción, redes).

7.29.4. Producto tensorial y álgebra multilineal

Idea clave

Intuición primero. Una función que es lineal en *cada* uno de dos vectores por separado (bilineal) es incómoda de manejar. El producto tensorial la reempaqueta como una función lineal de *un solo* objeto combinado $v \otimes w$. Pensado con coordenadas: $e_i \otimes f_j$ es simplemente "la casilla (i, j)", y por eso una matriz *es* un tensor: la tabla de casillas de un producto tensorial.

El *producto tensorial* $V \otimes W$ de dos espacios vectoriales es el espacio donde viven las *aplicaciones bilineales*: convierte "una función de dos vectores lineal en cada uno" en "una función lineal de un solo objeto". Su dimensión es el producto:

$$\dim(V \otimes W) = \dim V \cdot \dim W.$$

Si $\{e_i\}$ es base de V y $\{f_j\}$ de W, entonces $\{e_i \otimes f_j\}$ es base de $V \otimes W$.

Ejemplo 7.152. $\mathbb{R}^2 \otimes \mathbb{R}^3$ tiene dimensión 6, y sus elementos son exactamente las matrices 2×3: el tensor $e_i \otimes f_j$ es la matriz con un 1 en la posición (i, j). De hecho, *una matriz es un tensor*: el espacio de aplicaciones lineales de W a V es $V \otimes W^*$. Esta es la puerta al álgebra multilineal y al cálculo tensorial del capítulo 36, donde los objetos con varios índices T^{ij}_k son elementos de productos tensoriales de copias de V y su dual V^*.

7.30. Ejercicios propuestos

Sin solución desarrollada: solo enunciado y respuesta al final. Úsalos para comprobar que lo del capítulo se sostiene sin la muleta del texto delante.

7.1. Calcula el determinante y la inversa de $A = \begin{pmatrix} 2 & 1 \\ 3 & 4 \end{pmatrix}$.

7.2. Halla el rango de $\begin{pmatrix} 1 & 2 & 3 \\ 2 & 4 & 6 \\ 1 & 0 & 1 \end{pmatrix}$.

7.3. Da los autovalores de $\begin{pmatrix} 3 & 1 \\ 0 & 2 \end{pmatrix}$.

7.4. Da autovalores y autovectores de $\begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}$.

7.5. ¿Es $\{(1, 0, 1), (0, 1, 1), (1, 1, 2)\}$ base de $\mathbb{R}^3$?

7.6. Resuelve $x + 2y = 5$, $3x - y = 1$ escribiéndolo como sistema matricial.

7.7. Ortonormaliza $\{(1,1),(1,0)\}$ por Gram–Schmidt.

7.8. Da la traza y el determinante de $\begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}$ y comprueba que coinciden con la suma y el producto de sus autovalores.

7.9. Proyecta $(3,4)$ sobre la recta generada por $(1,0)$ y sobre la generada por $(1,1)$.

7.10. Da una base del núcleo de $\begin{pmatrix} 1 & 2 & 3 \\ 2 & 4 & 6 \end{pmatrix}$ y comprueba el teorema del rango.

7.11. ¿Es $\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$ diagonalizable? Justifícalo con la dimensión del autoespacio.

7.12. Si A es 5×7 y $\operatorname{rango} A = 4$, ¿cuánto vale $\operatorname{nulidad} A$?

Respuestas

1. $\det A = 5$, $A^{-1} = \frac{1}{5}\begin{pmatrix} 4 & -1 \\ -3 & 2 \end{pmatrix}$. **2.** 2. **3.** 3 y 2 (es triangular). **4.** $\lambda = 3$ con $(1,1)$ y $\lambda = 1$ con $(1,-1)$. **5.** No: el tercero es la suma de los dos primeros, el rango es 2. **6.** $x = 1$, $y = 2$. **7.** $u_1 = \frac{1}{\sqrt{2}}(1,1)$, $u_2 = \frac{1}{\sqrt{2}}(1,-1)$. **8.** $\operatorname{tr} = 5$, $\det = -2$; los autovalores son $\frac{5 \pm \sqrt{33}}{2}$, cuya suma es 5 y cuyo producto es -2. **9.** $(3,0)$ y $\left(\frac{7}{2}, \frac{7}{2}\right)$. **10.** Base $\{(-2,1,0),(-3,0,1)\}$; rango 1 más nulidad 2 igual a 3 columnas. **11.** No: el único autovalor es 1 y su autoespacio tiene dimensión 1, no 2. **12.** $\operatorname{nulidad} A = 7 - 4 = 3$.

Capítulo 8

Grafos: la matemática de las conexiones

Lo que necesitas traer

Del capítulo 1: conjuntos, relaciones e inducción. Del capítulo 5: conteo básico y coeficientes binomiales. Del capítulo 7: matriz de adyacencia, potencias de matrices y valores propios, que son la herramienta con la que se cuentan caminos y se estudian grafos regulares.

El capítulo 5 enseñó a contar objetos. Este enseña a estudiar *relaciones* entre objetos, que es una idea distinta y quizá más moderna: en un grafo lo importante no es qué son las cosas, sino quién está conectado con quién. Es el capítulo del libro que más de cerca toca la informática, las redes de transporte, la logística y las redes sociales, y sin embargo no necesita más que saber contar.

Idea clave

Intuición primero. Dibuje usted unos puntos en un papel y una raya entre los que estén relacionados. Ya está: eso es un grafo. Los puntos pueden ser ciudades y las rayas carreteras, o personas y amistades, o átomos y enlaces, o páginas web y enlaces de hipertexto.
Lo asombroso es que muchísimas preguntas prácticas —¿puedo recorrer todas las calles sin repetir ninguna?, ¿cuál es la ruta más barata?, ¿cuántos colores necesito para que dos países vecinos no tengan el mismo?— dependen *solo* del dibujo de puntos y rayas, y no de qué representen. Estirar el papel, mover los puntos o alargar las rayas no cambia nada. Por eso la teoría de grafos es prima hermana de la topología (capítulo 28): las dos estudian lo que sobrevive a la deformación.

8.1. Qué es un grafo

Definición 8.1. Un **grafo** $G = (V, A)$ consta de un conjunto finito V de **vértices** (o nodos) y un conjunto A de **aristas**, cada una de las cuales une dos vértices. Se dice que dos vértices

son **adyacentes** si hay una arista entre ellos, y que una arista **incide** en los vértices que une.

Un grafo es **simple** si no hay aristas de un vértice a sí mismo (*lazos*) ni dos aristas distintas uniendo la misma pareja (*aristas múltiples*). Si las aristas tienen sentido se llama **dirigido** (una calle de sentido único, un enlace de una web a otra); si además llevan un número asociado —kilómetros, precio, tiempo— se llama **ponderado**.

Ejemplo 8.2. El **grafo completo** K_n es el que tiene n vértices y *todas* las aristas posibles entre ellos. ¿Cuántas son? Tantas como parejas de vértices, es decir $\binom{n}{2} = \frac{n(n-1)}{2}$, que es exactamente la combinatoria del capítulo 5. Así K_5 tiene 10 aristas y K_6 tiene 15.

Observación 8.3. El dibujo de un grafo no es el grafo. El mismo grafo puede dibujarse de infinitas maneras, con las aristas rectas o curvas, cruzándose o no. Lo único que cuenta es qué parejas están unidas. Dos dibujos muy distintos pueden ser el mismo grafo (se dice entonces que son *isomorfos*), y decidir si dos grafos dados son isomorfos es, en general, un problema computacionalmente difícil y hasta hoy no del todo entendido.

8.2. El vocabulario básico: familias de grafos, subgrafos y las tres maneras de escribir uno

La sección anterior ha dicho qué es un grafo: unos puntos y unas rayas. Con eso basta para entender el dibujo, pero no para leer el resto del capítulo, porque a partir de aquí van a aparecer sin previo aviso dos docenas de palabras —simple, regular, subgrafo inducido, componente, cuello, bosque, matriz de incidencia— que no significan lo que uno adivinaría. Esta sección las fija todas de golpe sobre un mismo ejemplo, y termina explicando cómo se escribe un grafo cuando no se puede dibujar: de las tres formas que usa cualquier programa de ordenador, y cómo pasar de una a otra sin equivocarse.

Anticipamos una sola noción de la sección siguiente, porque hace falta ya para nombrar las familias: el *grado* de un vértice es el número de aristas que inciden en él. La sección siguiente lo estudia a fondo, con el lema del apretón de manos; aquí sólo hay que saber contarlo.

Definición 8.4 (Orden, tamaño y las cuatro variantes del concepto de grafo)**.** Sea $G = (V, A)$ un grafo. Su **orden** es el número de vértices $n = |V|$ y su **tamaño** es el número de aristas $m = |A|$. Una arista que une un vértice consigo mismo es un **lazo**; dos o más aristas distintas que unen la misma pareja de vértices son **aristas paralelas** o *múltiples*.

(i) Un grafo es **simple** si no tiene lazos ni aristas paralelas. Entonces cada arista queda determinada por el par no ordenado $\{u, v\}$ de sus extremos, y A es un conjunto de pares.

(ii) Un **multigrafo** es un grafo en el que se permiten lazos y aristas paralelas. Ahí A ya no es un conjunto de pares sino una lista, porque hay que distinguir dos aristas con los mismos extremos.

(iii) Un grafo **dirigido** tiene **arcos** (u, v) con sentido, del origen u al destino v. Cada vértice tiene entonces dos grados: el **grado de entrada** $\deg^-(v)$, número de arcos que terminan en v, y el **grado de salida** $\deg^+(v)$, número de arcos que salen de v.

(iv) Un grafo **ponderado** lleva asociada una función $w : A \to \mathbb{R}$ que asigna a cada arista un número: su *peso*, *coste* o *longitud*.

El **vecindario** de un vértice es $N(v) = \{u \in V : \{u, v\} \in A\}$, el conjunto de sus vecinos. En un grafo simple, $\deg(v) = |N(v)|$.

Observación 8.5. Las palabras *simple*, *dirigido* y *ponderado* ya se habían usado en la sección anterior, en prosa y sin más explicación; lo único que hace la definición es darles forma precisa y añadir lo que allí faltaba: el orden, el tamaño, el vecindario y los dos grados de un vértice dirigido. Adóptese además el convenio habitual, que se seguirá en todo el capítulo salvo aviso expreso: **cuando se diga "grafo" a secas se entenderá grafo simple, finito y no dirigido**. Es un convenio, no un teorema: los puentes de Königsberg de dos secciones más adelante son un multigrafo, y allí se dirá.

Ejemplo 8.6 (Cuatro objetos que la palabra "grafo" abarca)**.** (a) *Un grafo simple.* Sea H con $V = \{1, 2, 3, 4, 5, 6\}$ y

$$A = \{12,\ 13,\ 23,\ 34,\ 45,\ 46,\ 56\}.$$

Su orden es $n = 6$ y su tamaño $m = 7$. No hay lazos ni parejas repetidas: es simple. Los vecindarios son $N(1) = \{2, 3\}$, $N(2) = \{1, 3\}$, $N(3) = \{1, 2, 4\}$, $N(4) = \{3, 5, 6\}$, $N(5) = \{4, 6\}$ y $N(6) = \{4, 5\}$, de donde los grados son $2, 2, 3, 3, 2, 2$, que suman $14 = 2m$. Este grafo —dos triángulos, 123 y 456, unidos por la arista puente 34— será el ejemplo de trabajo de toda la sección.

(b) *Un multigrafo.* Sobre $V = \{A, B, C, D\}$, tómense *dos* aristas entre A y B, *dos* entre A y C, y una entre A y D, otra entre B y D y otra entre C y D. Son siete aristas y sólo cinco parejas distintas de vértices: no es simple. Los grados valen $\deg(A) = 5$ y $\deg(B) = \deg(C) = \deg(D) = 3$, y suman $14 = 2 \cdot 7$. Es el multigrafo de los puentes de Königsberg, que reaparecerá dos secciones más adelante; obsérvese que sin aristas paralelas sería imposible modelar la ciudad, porque de verdad había dos puentes entre las mismas dos orillas.

(c) *Un grafo dirigido.* Sobre $V = \{1, 2, 3, 4\}$ tómense los arcos $1 \to 2$, $1 \to 3$, $2 \to 4$, $3 \to 2$, $3 \to 4$, $4 \to 1$. Los grados de entrada son $\deg^- = 1, 2, 1, 2$ y los de salida $\deg^+ = 2, 1, 2, 1$. Cada lista suma 6, que es el número de arcos: cada arco aporta 1 a un grado de salida y 1 a un grado de entrada. Esa es la versión dirigida del lema del apretón de manos, y aquí no aparece ningún factor 2.

(d) *Un grafo ponderado.* El mismo H del apartado (a) con los pesos $w(12) = 4$, $w(13) = 2$, $w(23) = 1$, $w(34) = 7$, $w(45) = 3$, $w(46) = 5$, $w(56) = 2$. La estructura no ha cambiado en absoluto —mismos vértices, mismas aristas—; lo único nuevo es una etiqueta numérica por arista, y con ella tienen sentido preguntas que antes no lo tenían: cuál es la ruta más barata, o cuál es la red conexa más económica.

Definición 8.7 (Familias de grafos con nombre propio)**.** Sea G un grafo simple de orden n.

(i) G es **regular** de grado r, o **r-regular**, si todos sus vértices tienen el mismo grado r.

(ii) El **grafo completo** K_n tiene n vértices y todas las aristas posibles; es $(n-1)$-regular.

(iii) El **ciclo** C_n (con $n \geq 3$) tiene vértices $v_1, \ldots, v_n$ y aristas $v_1v_2,\ v_2v_3,\ \ldots,\ v_{n-1}v_n,\ v_nv_1$; es 2-regular.

(iv) El **camino** P_n es el ciclo C_n al que se le ha quitado una arista: n vértices en fila y $n-1$ aristas.

(v) El **bipartito completo** $K_{p,q}$ tiene $p+q$ vértices repartidos en dos bandos y todas las aristas de un bando al otro, ninguna dentro. La **estrella** es $K_{1,q}$.

(vi) El **complementario** $\overline{G}$ tiene los mismos vértices que G, y dos vértices son adyacentes en $\overline{G}$ exactamente cuando *no* lo son en G.

Ejemplo 8.8 (Las cuentas de aristas de cada familia)**.** (a) *El completo.* Cada arista es una pareja de vértices, luego K_n tiene $\binom{n}{2} = \frac{n(n-1)}{2}$ aristas: K_4 tiene 6, K_5 tiene 10 y K_6 tiene 15. Comprobación por grados: los n vértices tienen grado $n-1$, la suma es $n(n-1)$ y el número de aristas es la mitad. Las dos cuentas coinciden, como debe ser.

(b) *El ciclo y el camino.* C_n tiene n aristas y todos los grados iguales a 2: suma $2n$, mitad n ✓. P_n tiene $n-1$ aristas, dos vértices de grado 1 (los extremos) y $n-2$ de grado 2: suma $2 + 2(n-2) = 2n-2$, mitad $n-1$ ✓.

(c) *El bipartito completo.* $K_{p,q}$ tiene pq aristas, una por cada elección de un vértice de cada bando. En $K_{3,4}$ son 12: los tres vértices del bando pequeño tienen grado 4 y los cuatro del grande grado 3, la suma es $3 \cdot 4 + 4 \cdot 3 = 24$ y la mitad es 12 ✓. Nótese que $K_{p,q}$ es regular sólo si $p = q$.

(d) *Regular no significa completo.* El grafo de Petersen tiene 10 vértices y es 3-regular, luego tiene $\frac{10 \cdot 3}{2} = 15$ aristas; el completo K_{10} tendría $\binom{10}{2} = 45$. Y el H del ejemplo anterior no es regular: sus grados son $2, 2, 3, 3, 2, 2$.

(e) *El complementario.* Como cada pareja de vértices está unida en G o en $\overline{G}$ y nunca en ambos, se cumple siempre

$$m(G) + m(\overline{G}) = \binom{n}{2}.$$

Para H, con $n = 6$ y $m = 7$: el complementario tiene $15 - 7 = 8$ aristas, y en efecto son 14, 15, 16, 24, 25, 26, 35 y 36. Un caso llamativo: el complementario de C_5 vuelve a ser 2-regular con cinco vértices y cinco aristas, y de hecho *es* un C_5 —el pentágono se convierte en la estrella de cinco puntas—. Se dice que C_5 es *autocomplementario*.

Definición 8.9 (Subgrafo, subgrafo inducido y subgrafo generador)**.** Un **subgrafo** de $G = (V, A)$ es un grafo $G' = (V', A')$ con $V' \subseteq V$ y $A' \subseteq A$, donde además toda arista de A' tiene sus dos extremos en V'. Dado $S \subseteq V$, el **subgrafo inducido** $G[S]$ es el que tiene por vértices S y por aristas *todas* las de G con ambos extremos en S. Un subgrafo es **generador** (o abarcador) si $V' = V$, es decir, si conserva todos los vértices y sólo se le quitan aristas.

Ejemplo 8.10 (Los subgrafos de H, contados y exhibidos)**.** Sobre el grafo H del ejemplo anterior:

(a) *Un inducido conexo.* Con $S = \{1, 2, 3, 4\}$, el subgrafo $H[S]$ tiene las aristas 12, 13, 23 y 34: cuatro. No se elige cuáles se toman; se toman todas las que quedan dentro. El resultado es un triángulo con un vértice colgando.

(b) *Un inducido desconectado.* Con $S = \{1, 2, 4, 5\}$ sólo sobreviven 12 y 45: el subgrafo inducido tiene dos piezas separadas, aunque en H esos cuatro vértices estaban todos conectados entre sí por caminos. Quitar vértices puede romper la conexión, y un subgrafo inducido de un grafo conexo no tiene por qué serlo.

(c) *Subgrafo que no es inducido.* El grafo con vértices $\{1, 2, 3\}$ y única arista 12 es un subgrafo perfectamente legítimo de H, pero no es $H[\{1, 2, 3\}]$, que sería el triángulo completo 12, 13, 23. La diferencia entre "subgrafo" e "inducido" es exactamente esa libertad de tirar aristas.

(d) *Cuántos hay.* Subgrafos inducidos hay uno por cada subconjunto de vértices: $2^6 = 64$, contando el vacío y el propio H. Subgrafos generadores hay uno por cada subconjunto de aristas: $2^7 = 128$. Y subgrafos a secas hay muchos más, porque se puede elegir primero el conjunto de vértices y después cualquier subconjunto de las aristas admisibles.

Definición 8.11 (Recorridos, caminos, ciclos, componentes y bosques)**.** Sea G un grafo.

(i) Un **recorrido** (o *paseo*) de longitud k es una sucesión de vértices $v_0, v_1, \ldots, v_k$ en la que v_i y v_{i+1} son adyacentes para todo i. Se permite repetir vértices y aristas. Su **longitud** es el número de aristas recorridas, k, no el de vértices.

(ii) Un **camino simple** es un recorrido que no repite ningún vértice.

(iii) Un **circuito** es un recorrido cerrado, es decir, con $v_k = v_0$.

(iv) Un **ciclo** es un circuito de longitud ≥ 3 que no repite ningún vértice salvo el primero, que coincide con el último. El **cuello** (o *cintura*) de G es la longitud del ciclo más corto, y es ∞ si no hay ninguno.

(v) La **distancia** $d(u, v)$ es la longitud del camino más corto entre u y v, e ∞ si no hay ninguno. El **diámetro** es el mayor valor de $d(u, v)$.

(vi) Una **componente conexa** es un subgrafo inducido conexo maximal: se toma un vértice, se le añaden todos los vértices a los que se llega por algún camino, y se induce. Un grafo es conexo si tiene una sola componente.

(vii) Un **bosque** es un grafo sin ciclos; sus componentes conexas son árboles. Un árbol es un bosque conexo.

Observación 8.12. Aquí hay una trampa de vocabulario que causa la mitad de los errores del principiante, y conviene verla escrita: *recorrido*, *paseo* y *camino* no son sinónimos, aunque el castellano corriente los trate como tales. La sección siguiente sobre caminos y conexión usa "camino" en el sentido amplio de recorrido, y la regla de la matriz de adyacencia dirá explícitamente que las potencias A^k cuentan *recorridos* y no caminos simples: esa distinción es la que hace que una cuenta sea fácil y la otra endiabladamente difícil. Léase siempre con atención si se permite repetir o no.

Ejemplo 8.13 (Todo el vocabulario, medido sobre el mismo grafo)**.**

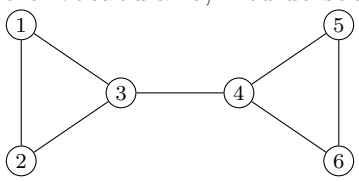


Figura 8.1: El grafo H: dos triángulos unidos por la arista puente 34. Seis vértices, siete aristas, grados $2, 2, 3, 3, 2, 2$ y cuello 3.

(a) *Un recorrido que no es camino.* La sucesión $1, 2, 3, 1, 2$ tiene longitud 4 y es un recorrido legítimo: cada vértice es adyacente al siguiente. No es camino simple, porque repite 1 y 2, ni circuito, porque no vuelve al principio.

(b) *Un camino simple.* La sucesión $1, 2, 3, 4, 5$ tiene longitud 4 y no repite ningún vértice: es un camino simple. Otro es $1, 3, 4, 6$, de longitud 3.

(c) *Ciclos y cuello.* Los ciclos de H son exactamente dos: $1, 2, 3, 1$ y $4, 5, 6, 4$, ambos de longitud 3. Por tanto el cuello vale 3. La arista 34 no pertenece a ningún ciclo; se la llama *puente*, precisamente porque quitarla desconecta el grafo.

(d) *Distancias y diámetro.* La tabla de distancias es

d	1	2	3	4	5	6
1	0	1	1	2	3	3
2	1	0	1	2	3	3
3	1	1	0	1	2	2
4	2	2	1	0	1	1
5	3	3	2	1	0	1
6	3	3	2	1	1	0

El mayor valor es 3, alcanzado por las parejas $\{1, 5\}$, $\{1, 6\}$, $\{2, 5\}$ y $\{2, 6\}$: el diámetro de H es 3.

(e) *Componentes.* H es conexo: una sola componente. Si se borra la arista 34 quedan dos componentes, $H[\{1, 2, 3\}]$ y $H[\{4, 5, 6\}]$, cada una un triángulo. Si en cambio se borra la arista 12, el grafo sigue siendo conexo, porque 1 y 2 siguen comunicados a través de 3. Ninguna arista de un ciclo desconecta; las que no están en ningún ciclo, sí.

(f) *Bosque, no.* H tiene ciclos, luego no es un bosque ni un árbol. Pero quitándole una arista de cada triángulo —por ejemplo 12 y 56— quedan $7 - 2 = 5$ aristas y seis vértices sin ciclos: un árbol generador, con $n - 1 = 5$ aristas, exactamente como exige el teorema de los árboles de más adelante.

Definición 8.14 (Las tres representaciones de un grafo)**.** Sea G un grafo con vértices numerados $v_1, \ldots, v_n$ y aristas numeradas $a_1, \ldots, a_m$.

(i) La **matriz de adyacencia** es la matriz A de tamaño $n \times n$ con $A_{ij} = 1$ si v_i y v_j son adyacentes y $A_{ij} = 0$ si no. En un multigrafo se pone en A_{ij} el número de aristas entre v_i y v_j.

(II) La **matriz de incidencia** es la matriz B de tamaño $n \times m$ con $B_{ij} = 1$ si el vértice v_i es un extremo de la arista a_j, y $B_{ij} = 0$ si no.

(III) La **lista de adyacencia** es la colección de los n vecindarios: para cada vértice, la lista de sus vecinos.

Idea clave

Las tres dicen exactamente lo mismo y sirven para cosas distintas. La matriz de adyacencia es cuadrada, simétrica en el caso no dirigido, y convierte el grafo en un objeto de álgebra lineal: sus potencias cuentan recorridos y sus autovalores son invariantes, como se verá en la regla sobre potencias de A. La matriz de incidencia es rectangular y pone el foco en las aristas; es la que aparece en las leyes de Kirchhoff y en los flujos. Y la lista de adyacencia es la que usa cualquier programa real, porque ocupa espacio proporcional a $n + m$ en lugar de a n^2: en una red social con mil millones de personas y unos pocos cientos de amigos cada una, la matriz sería inconcebible y la lista cabe holgadamente.

Regla: escribir un grafo en las tres formas y pasar de una a otra

1. **Numera los vértices y fija el orden.** Todo lo que sigue depende de esa numeración, así que escríbela antes de dibujar nada. Numera también las aristas si vas a usar la matriz de incidencia.
2. **Del dibujo a la lista.** Recorre los vértices en orden y anota, para cada uno, sus vecinos. Cada arista debe aparecer *dos* veces en la lista, una en cada extremo: es la primera comprobación gratuita.
3. **De la lista a la matriz de adyacencia.** Pon un 1 en la casilla (i, j) por cada j que aparezca en la lista de i. La matriz sale simétrica y con la diagonal nula si el grafo es simple y no dirigido; una diagonal no nula significa un lazo.
4. **De la lista a la matriz de incidencia.** Una columna por arista, con exactamente dos unos, en las filas de sus dos extremos. Un lazo es la excepción: se marca con un 2.
5. **Comprueba con las sumas.** La suma de la fila i de A es $\deg(v_i)$; la suma de todas las entradas de A es $2m$; la suma de la fila i de B es también $\deg(v_i)$, y la suma de cada columna de B vale 2. Si alguna de las cuatro falla, hay un error de transcripción y no merece la pena seguir.
6. **De la matriz al dibujo.** Recorre sólo el triángulo superior de A —el inferior es su reflejo— y traza una raya por cada 1. Coloca los vértices donde quieras: el dibujo no es el grafo.
7. **De la matriz de incidencia a la de adyacencia.** Lee columna a columna: cada columna te da la pareja de vértices de una arista. Un atajo algebraico: el producto BB^{T} vale A fuera de la diagonal y $\deg(v_i)$ en la diagonal, de modo que $BB^{\mathsf{T}} = A + D$ con D la matriz diagonal de los grados.

8. **Verifica el resultado contando aristas de tres maneras:** rayas del dibujo, mitad de la suma de la lista y columnas de B. Los tres números tienen que ser m.

Ejemplo 8.15 (El grafo H escrito de las tres maneras, y las vueltas de ida y vuelta)**.** Se toma el grafo de la figura 8.1, con el orden de vértices $1, 2, 3, 4, 5, 6$ y el orden de aristas

$$a_1 = 12, \quad a_2 = 13, \quad a_3 = 23, \quad a_4 = 34, \quad a_5 = 45, \quad a_6 = 46, \quad a_7 = 56.$$

(a) *Lista de adyacencia.* Recorriendo los vértices en orden y anotando vecinos:

$$1:\ 2,3 \qquad 2:\ 1,3 \qquad 3:\ 1,2,4 \qquad 4:\ 3,5,6 \qquad 5:\ 4,6 \qquad 6:\ 4,5.$$

Los tamaños de las seis listas son $2, 2, 3, 3, 2, 2$: son los grados. Suman 14, que es $2m$ con $m = 7$ ✓. Cada arista aparece dos veces: el 2 de la lista de 1 y el 1 de la lista de 2 son la misma arista a_1 vista desde sus dos extremos.

(b) *Matriz de adyacencia.*

$$A = \begin{pmatrix} 0 & 1 & 1 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 1 & 1 & 0 \end{pmatrix}.$$

Es simétrica, tiene diagonal nula y sus filas suman $2, 2, 3, 3, 2, 2$ ✓. La suma total de sus 36 entradas es $14 = 2m$ ✓. Obsérvese cómo se lee la estructura a simple vista: el bloque de arriba a la izquierda y el de abajo a la derecha son los dos triángulos, y el único 1 que los cruza es la arista puente 34.

(c) *Matriz de incidencia.* Siete columnas, una por arista, cada una con exactamente dos unos:

$$B = \begin{pmatrix} 1 & 1 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 1 & 1 \end{pmatrix}.$$

Las filas suman $2, 2, 3, 3, 2, 2$ —otra vez los grados— y las siete columnas suman 2 cada una ✓. Nótese que B no es cuadrada: tiene una fila por vértice y una columna por arista, y por tanto ni siquiera tiene sentido preguntarse si es simétrica.

(d) *La vuelta: de B a A.* La columna a_4 tiene unos en las filas 3 y 4, luego la arista une 3 con 4; repitiendo con las siete columnas se recupera la lista de aristas inicial. Por el atajo del paso 7, $BB^{\mathsf{T}} = A + D$, donde D es la matriz diagonal con los grados $2, 2, 3, 3, 2, 2$; en particular la entrada $(3, 4)$ de BB^{T} vale 1, porque 3 y 4 comparten exactamente la arista a_4, y la entrada $(1, 4)$ vale 0, porque no comparten ninguna.

(e) *Lo que la matriz regala.* Elevando al cuadrado,

$$A^2 = \begin{pmatrix} 2 & 1 & 1 & 1 & 0 & 0 \\ 1 & 2 & 1 & 1 & 0 & 0 \\ 1 & 1 & 3 & 0 & 1 & 1 \\ 1 & 1 & 0 & 3 & 1 & 1 \\ 0 & 0 & 1 & 1 & 2 & 1 \\ 0 & 0 & 1 & 1 & 1 & 2 \end{pmatrix},$$

cuya diagonal $(2, 2, 3, 3, 2, 2)$ vuelve a dar los grados, como tiene que ser: los recorridos de longitud 2 que salen y vuelven a un vértice son "ir a un vecino y volver". La entrada $(1, 4) = 1$ dice que hay un solo recorrido de longitud 2 de 1 a 4, y es $1, 3, 4$. La entrada $(1, 5) = 0$ confirma que $d(1, 5) > 2$; en efecto la tabla de distancias daba 3. Y $\operatorname{tr}(A^3) = 12$, luego el grafo tiene $12/6 = 2$ triángulos, que son los dos que se ven.

(f) *Los tres recuentos de aristas.* Rayas del dibujo: 7. Mitad de la suma de los tamaños de las listas: $14/2 = 7$. Columnas de B: 7. Los tres coinciden y la transcripción es correcta.

Regla: describir un grafo desconocido con una ficha de siete casillas

Cuando llega un grafo del que no se sabe nada —un dibujo, una matriz o una lista—, este es el orden que más información da con menos esfuerzo, y con él se contesta la mayoría de las preguntas de examen antes de empezar a pensar.

1. **Orden y tamaño:** cuenta n y m. Si te dan la matriz, m es la mitad de la suma de todas las entradas.
2. **Sucesión de grados**, ordenada de menor a mayor. Comprueba que suma $2m$; si no, has contado mal.
3. **¿Simple?** Mira si hay lazos (diagonal no nula) o aristas paralelas (entradas mayores que 1). Muchos teoremas del capítulo, como la cota $m \leq 3n - 6$ de los grafos planos, exigen que lo sea.
4. **¿Regular, completo, ciclo, camino, estrella?** Si todos los grados coinciden es regular; si además valen $n - 1$, es K_n; si todos valen 2 y es conexo, es C_n.
5. **¿Conexo?** Recorre el grafo desde un vértice; lo que no alcances es otra componente. Cuenta cuántas hay.
6. **¿Bipartito? ¿Con ciclos? ¿Cuál es el cuello?** Busca el ciclo más corto: si hay alguno de longitud impar, no es bipartito; si no hay ninguno, es un bosque, y es un árbol si además es conexo, lo que equivale a $m = n - 1$.
7. **Cierra con los invariantes finos si hacen falta:** número de triángulos $\frac{1}{6}\operatorname{tr}(A^3)$, diámetro y complementario. Y verifica siempre el conjunto de la ficha con el lema del apretón de manos, que es la comprobación más barata que existe.

Ejemplo 8.16 (Cuatro grafos pasados por la ficha)**.** (a) *El grafo H.* $n = 6$, $m = 7$, grados

$(2,2,2,2,3,3)$ que suman $14 = 2m$ ✓; simple; no regular; conexo con una componente; tiene ciclos —dos triángulos—, luego no es bosque ni bipartito, y su cuello es 3; diámetro 3; complementario con $15 - 7 = 8$ aristas.

(b) *El ciclo C_6.* $n = 6$, $m = 6$, grados $(2,2,2,2,2,2)$ que suman $12 = 2m$ ✓; simple, 2-regular, conexo; su único ciclo tiene longitud 6, que es par, así que es bipartito con los bandos $\{1,3,5\}$ y $\{2,4,6\}$; cuello 6; no tiene triángulos, y en efecto $\operatorname{tr}(A^3) = 0$; diámetro 3.

(c) *El bipartito completo $K_{3,4}$.* $n = 7$, $m = 12$, grados $(3,3,3,3,4,4,4)$ que suman $24 = 2m$ ✓; simple, no regular porque los dos bandos tienen tamaños distintos; conexo; bipartito por construcción, luego sin ciclos impares y sin triángulos; su cuello es 4; diámetro 2.

(d) *El grafo de Petersen.* $n = 10$, $m = 15$, grados todos 3 que suman $30 = 2m$ ✓; simple, 3-regular, conexo; no es completo, porque K_{10} tendría 45 aristas; su cuello es 5, impar, así que *no* es bipartito, y como $5 > 3$ no tiene ningún triángulo: $\operatorname{tr}(A^3) = 0$. Este par de datos —3-regular con cuello 5— es el que lo caracteriza entre todos los grafos de diez vértices, según se vio al hablar de invariantes.

Cuidado

Tres confusiones que cuestan puntos y que conviene tener presentes.

Primera: subgrafo no es subgrafo inducido. Al inducido no se le puede quitar ninguna arista de las que quedan dentro; al subgrafo a secas, sí. Cuando un enunciado dice "contiene un triángulo" se refiere a un subgrafo cualquiera; cuando dice "el subgrafo inducido por los vecinos de v", hay que tomar todas las aristas entre ellos, sin excepción.

Segunda: la longitud se mide en aristas, no en vértices. El camino $1,2,3,4,5$ tiene longitud 4 y cinco vértices. Contar vértices en lugar de aristas desplaza en uno todas las distancias y todos los diámetros.

Tercera: la matriz depende del orden elegido. Renumerar los vértices cambia A por otra matriz distinta que representa el mismo grafo. Por eso las entradas de A no son invariantes del grafo, mientras que sus autovalores sí lo son; la sección sobre isomorfismo vuelve sobre este punto con detalle.

Pausa para pensar

¿Por qué tres representaciones y no una? Porque el coste de las preguntas cambia radicalmente según cómo esté guardado el grafo. Preguntar "¿son adyacentes u y v?" cuesta una sola consulta con la matriz de adyacencia y hasta $\deg(u)$ comparaciones con la lista. En cambio recorrer todos los vecinos de u cuesta $\deg(u)$ con la lista y siempre n con la matriz, aun cuando u tenga un único vecino. Como la mayoría de los grafos reales son *dispersos* —m del orden de n y no de n^2—, la lista gana casi siempre, y por eso los algoritmos de recorrido de la sección siguiente cuestan $\mathcal{O}(n+m)$ y no $\mathcal{O}(n^2)$. La matriz se reserva para cuando lo que se quiere es álgebra: potencias, autovalores, PageRank. La lección general vale mucho más allá de los grafos: elegir la estructura de datos *es* media solución del problema.

8.3. Grado y el lema del apretón de manos

Definición 8.17. El **grado** de un vértice v, escrito $\deg(v)$, es el número de aristas que inciden en él.

Teorema 8.18 (Lema del apretón de manos)**.** *En todo grafo, la suma de los grados de todos los vértices es el doble del número de aristas:*

$$\sum_{v \in V} \deg(v) = 2|A|.$$

Demostración. Cada arista tiene dos extremos, luego contribuye exactamente 1 al grado de cada uno de ellos, es decir, 2 a la suma total. Sumando sobre todas las aristas se obtiene $2|A|$. □

Idea clave

El nombre viene de la lectura social del enunciado: si en una fiesta cada apretón de manos involucra a dos personas, el total de manos estrechadas contadas persona a persona es el doble del número de apretones. Es una demostración por *doble conteo*: se cuenta lo mismo de dos maneras y se igualan los resultados. Esa técnica, que aquí parece trivial, es una de las más potentes de toda la combinatoria.

Corolario 8.19. *En todo grafo, el número de vértices de grado impar es* par*.*

Demostración. La suma total de grados es par, por el lema. Los vértices de grado par aportan una suma par. Luego lo que aportan los de grado impar también es par, y una suma de números impares solo es par si hay una cantidad par de sumandos. □

Ejemplo 8.20. ¿Existe un grafo con cinco vértices, todos de grado 3? La suma de grados sería $5 \cdot 3 = 15$, que es impar, y por el lema tendría que ser el doble de un entero. Imposible.

En cambio con ocho vértices de grado 3 sí: la suma es 24 y habría 12 aristas. Un ejemplo concreto es el grafo del cubo, cuyos ocho vértices tienen tres aristas cada uno.

Cuidado

Que la suma de grados sea par es condición *necesaria* para que exista el grafo, pero no siempre suficiente. Por ejemplo, no existe grafo simple con cuatro vértices de grados $3, 3, 3, 1$ pese a que suman 10: si tres vértices tienen grado 3 en un grafo de cuatro vértices, cada uno está unido a todos los demás, y entonces el cuarto tendría grado 3 y no 1. Comprobar la paridad descarta casos, no los confirma.

Regla: decidir si dos grafos son isomorfos

Dos grafos son *el mismo* si existe una biyección entre sus vértices que respete las aristas. El dibujo no cuenta: un grafo no es un dibujo, es una lista de quién está unido con quién.

1. **Contar vértices y aristas.** Si no coinciden, no son isomorfos y se acabó.
2. **Comparar las sucesiones de grados** ordenadas de menor a mayor. Si difieren, no son isomorfos.
3. **Comparar el número de componentes conexas.**
4. **Comparar el cuello** (la longitud del ciclo más corto) y el **número de triángulos**.
5. **Comparar propiedades cualitativas:** bipartito o no, plano o no, número cromático, existencia de ciclo euleriano o hamiltoniano.
6. **Si todos los invariantes coinciden, hay que construir la biyección explícita.** Se empieza emparejando vértices de grado poco frecuente, que son los que dejan menos opciones, y se va propagando a los vecinos.
7. **Verificar la biyección arista por arista**, en los dos sentidos: cada arista de G debe ir a una arista de H y cada arista de H debe venir de una de G.

La asimetría es esencial: para demostrar que *no* son isomorfos basta un invariante que difiera; para demostrar que *sí* lo son no basta ningún número de invariantes, hace falta la biyección.

Ejemplo 8.21 (seis parejas y por qué los invariantes nunca bastan)**.** (a) *El pentágono y la estrella de cinco puntas.* G es el ciclo $1-2-3-4-5-1$; H es la estrella dibujada uniendo a con c, c con e, e con b, b con d y d con a. Los dibujos no se parecen en nada, pero ambos tienen 5 vértices, 5 aristas y todos los grados iguales a 2. La biyección

$$1 \mapsto a, \quad 2 \mapsto c, \quad 3 \mapsto e, \quad 4 \mapsto b, \quad 5 \mapsto d$$

convierte cada arista de G en una de H: son el mismo grafo, C_5. Toda estrella de cinco puntas es un pentágono mal dibujado.

(b) *Dos grafos 3-regulares con seis vértices.* $K_{3,3}$ (tres casas y tres suministros) y el prisma triangular tienen ambos 6 vértices, 9 aristas, todos los grados iguales a 3, y ambos son conexos. Cuatro invariantes coincidentes, y aun así no son isomorfos: el prisma tiene 2 triángulos y $K_{3,3}$ no tiene ninguno, porque es bipartito. Equivalentemente, el cuello del prisma es 3 y el de $K_{3,3}$ es 4.

(c) *Cuándo se responde en un segundo.* Si G tiene 12 aristas y H tiene 11, o si la sucesión de grados de G es $(1,2,2,3,4)$ y la de H es $(2,2,2,3,3)$, no hay nada que pensar. Conviene siempre empezar por aquí; el lema del apretón de manos da además una comprobación gratis, porque la suma de los grados debe ser $2m$ en los dos casos.

(d) *El grafo de Petersen.* Tiene 10 vértices, 15 aristas, es 3-regular y su cuello es 5. Resulta que *es el único* grafo con esas cuatro propiedades, de modo que cualquier grafo 3-regular de cuello 5 con diez vértices que aparezca por ahí, dibujado como sea, es él. Es el ejemplo perfecto de invariantes que, excepcionalmente, sí determinan el grafo.

(e) *Ni siquiera el espectro basta.* A cada grafo se le puede asociar la lista de autovalores de su matriz de adyacencia, un invariante mucho más fino que los grados. Pues bien: la

estrella $K_{1,4}$ y la unión de un cuadrado con un vértice aislado tienen ambas el espectro $\{-2, 0, 0, 0, 2\}$ y no son isomorfas —la primera es conexa y la segunda no, y sus sucesiones de grados son $(1, 1, 1, 1, 4)$ y $(0, 2, 2, 2, 2)$—. Se llaman grafos *coespectrales*, y son la prueba de que ninguna lista finita de invariantes fáciles zanjará la cuestión.

(f) *Cuántos grafos distintos hay realmente.* Con 4 vértices etiquetados hay $2^6 = 64$ grafos posibles, uno por cada subconjunto de las $\binom{4}{2} = 6$ aristas. Pero muchos son el mismo grafo con otros nombres: salvo isomorfismo solo hay 11. Contar objetos "salvo isomorfismo" es contar clases de equivalencia, exactamente la operación del capítulo 23, y de hecho el grupo que actúa aquí es el grupo simétrico S_4 permutando las etiquetas.

Cuidado

El error más frecuente es dar por buena la respuesta "sí son isomorfos" porque coinciden los grados. El caso (b) lo desmiente con seis vértices, y hay parejas mucho peores. La sucesión de grados sirve para decir que *no*; nunca para decir que *sí*.

Pausa para pensar

¿Cuántas biyecciones hay que probar en el peor caso? Todas las de n elementos, es decir $n!$, que para $n = 20$ ya supera los dos trillones. Y sin embargo nadie ha demostrado que el problema del isomorfismo de grafos sea NP-completo, ni tampoco que se pueda resolver en tiempo polinómico: vive en una tierra de nadie que lo ha convertido en uno de los problemas abiertos más peculiares de la informática teórica. En la práctica los programas resuelven casos con miles de vértices sin despeinarse, usando exactamente la estrategia de la regla: refinar por invariantes locales hasta que casi no queden opciones.

8.4. Caminos, conexión y los puentes de Königsberg

Definición 8.22. Un **camino** es una sucesión de vértices en la que cada uno es adyacente al siguiente. Es **cerrado** (o un *circuito*) si empieza y termina en el mismo vértice. Un grafo es **conexo** si entre cualesquiera dos vértices hay algún camino.

Ejemplo 8.23 (Caminos, paseos y ciclos)**. (a) En el cuadrado.** Sea el grafo con vértices $1, 2, 3, 4$ y aristas $12, 23, 34, 41$. La sucesión $1, 2, 3, 4$ es un camino simple de longitud 3; la sucesión $1, 2, 3, 4, 1$ es un ciclo de longitud 4; y $1, 2, 1, 2$ es un paseo que repite aristas y por tanto no es camino simple.

(b) Contar caminos con matrices. Si A es la matriz de adyacencia, la entrada (i, j) de A^k cuenta los paseos de longitud k de i a j. En el cuadrado anterior, A^2 tiene un 2 en la diagonal (ir y volver por cualquiera de los dos vecinos) y un 2 en la posición $(1, 3)$: los paseos $1, 2, 3$ y $1, 4, 3$.

(c) La distancia. La distancia entre dos vértices es la longitud del camino más corto; en el cuadrado, $d(1, 3) = 2$. El *diámetro* es la mayor de todas las distancias y vale 2 aquí. En un camino de n vértices el diámetro es $n - 1$, y en el grafo completo K_n vale 1.

(d) Conexión. Un grafo es conexo si hay camino entre dos vértices cualesquiera. El grafo con aristas $12, 34$ sobre cuatro vértices no lo es: tiene dos componentes. Una manera mecánica de detectarlo es sumar $I + A + A^2 + \cdots + A^{n-1}$ y comprobar que ninguna entrada es cero.

En 1736, en la ciudad prusiana de Königsberg, siete puentes unían cuatro zonas de tierra separadas por el río Pregel. Los vecinos se preguntaban si era posible dar un paseo cruzando cada puente exactamente una vez. Leonhard Euler demostró que no, y al hacerlo inventó de paso la teoría de grafos: fue el primero en darse cuenta de que la geometría del asunto —las distancias, las formas de las islas— no importaba nada, y de que todo se reducía a un dibujo de cuatro puntos y siete rayas.

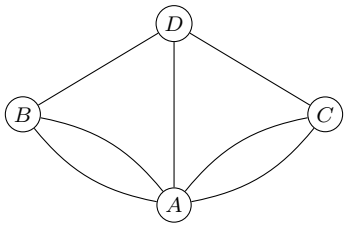


Figura 8.2: Los siete puentes de Königsberg reducidos a cuatro vértices y siete aristas. Los cuatro vértices tienen grado impar, y por eso el paseo que Euler declaró imposible lo es.

Los grados son $\deg(A) = 5$ y $\deg(B) = \deg(C) = \deg(D) = 3$: suman 14, es decir, hay 7 aristas. ✓

Teorema 8.24 (Euler, 1736)**.** *Un grafo conexo admite un* ***camino euleriano*** *—uno que recorra cada arista exactamente una vez— si y solo si tiene cero o dos vértices de grado impar. Si tiene cero, el camino puede además cerrarse (****circuito euleriano****) y empezar donde se quiera; si tiene dos, el camino debe empezar en uno de ellos y terminar en el otro.*

Idea de la demostración. La necesidad es la parte fácil y la que resuelve Königsberg: cada vez que el paseo *pasa* por un vértice, gasta dos aristas, una para entrar y otra para salir. Luego todo vértice intermedio ha de tener grado par. Solo el vértice de salida y el de llegada pueden quedar descompensados, y solo si son distintos. La suficiencia se prueba construyendo el camino: se recorre hasta atascarse, y se demuestra que lo que queda sin recorrer forma circuitos que pueden insertarse en el camino ya construido. □

En Königsberg los cuatro vértices tienen grado impar. Como $4 > 2$, el paseo es imposible. No es que nadie haya sido lo bastante hábil: es que no existe.

Idea clave

Este es exactamente el juego infantil de "dibujar una figura sin levantar el lápiz y sin repetir ninguna línea". La regla completa cabe en una frase: se puede si y solo si hay cero o dos puntos de los que salga un número impar de líneas; y si hay dos, hay que empezar obligatoriamente en uno de ellos. Con esa frase se resuelven de un vistazo todos los acertijos de ese tipo, incluido el clásico sobre de carta —que sí se puede— y la casita con la cruz en el tejado —que no—.

Ejemplo 8.25. Considérese el grafo de la figura, con $V = \{A, \ldots, F\}$ y aristas AB, AC, BC, BD, CD, CE, DE, DF, EF.

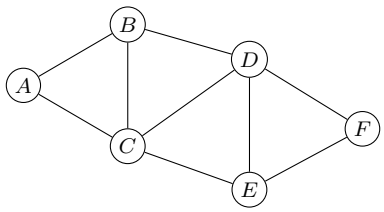


Figura 8.3: El grafo con $V = \{A, \ldots, F\}$ y nueve aristas sobre el que se calculan grados, caminos y ciclos a lo largo de la sección.

Los grados son $\deg A = 2$, $\deg B = 3$, $\deg C = 4$, $\deg D = 4$, $\deg E = 3$, $\deg F = 2$. Suman 18, luego hay 9 aristas ✓, y hay exactamente *dos* vértices impares, B y E. Por el teorema existe camino euleriano, y debe ir de B a E. Uno es

$$B \to D \to F \to E \to D \to C \to B \to A \to C \to E,$$

que recorre los nueve tramos sin repetir ninguno. No hay circuito euleriano, porque los impares no son cero.

Regla: decidir si un grafo se puede dibujar de un solo trazo

Éste es el problema con el que nació la teoría de grafos, y tiene una solución completa que cabe en tres líneas: solo hay que contar grados impares.

1. **Comprueba primero la conexión.** Todas las aristas tienen que estar en una sola pieza (los vértices aislados sin aristas no molestan). Si el grafo está partido en dos trozos con aristas, es imposible: el lápiz no salta.

2. **Escribe el grado de cada vértice**, es decir, cuántas aristas salen de él. Un lazo cuenta dos veces.

3. **Cuenta cuántos grados impares hay.** Por el lema del apretón de manos ese número siempre es par: 0, 2, 4, ...

4. **Aplica el criterio de Euler.** Si hay **cero** impares, existe un *circuito* de Euler: se recorre todo y se vuelve al punto de partida. Si hay **exactamente dos**, existe un *camino* de Euler, y forzosamente empieza en uno de esos dos vértices impares y acaba en el otro. Si hay **cuatro o más**, es imposible.

5. **Entiende el porqué, que es lo que hace inolvidable la regla.** Cada vez que el trazo pasa por un vértice intermedio, gasta dos aristas: una para entrar y otra para salir. Los únicos vértices que pueden quedar con un número impar son el de salida y el de llegada.

6. **Si es imposible, calcula cuántos trazos hacen falta:** con $2k$ vértices impares ($k \geq 1$) el grafo se recorre con exactamente k trazos, y no menos.

7. **No lo confundas con Hamilton.** Euler recorre todas las *aristas*; Hamilton visita todos los *vértices*. El primero tiene criterio sencillo, el segundo no tiene ninguno conocido.

Ejemplo 8.26 (Cinco grafos ante el criterio de Euler)**. (a) Los puentes de Königsberg.** Los cuatro trozos de tierra tienen grados 3, 3, 3 y 5: los cuatro son impares. Con cuatro vértices impares el paso 4 declara imposible el paseo, y el paso 6 dice además que harían falta $4/2 = 2$ trazos. Ésta es la respuesta que Euler dio en 1736, y obsérvese que no requiere probar ningún recorrido: basta con contar.

(b) El sobre abierto (el dibujo clásico de la casita: un rectángulo con las dos diagonales y un tejado triangular encima). Sus cinco vértices tienen grados 2, 3, 3, 4 y 4. Exactamente dos impares, luego *sí* se puede de un trazo, y hay que empezar en uno de los dos vértices de grado 3 —las esquinas inferiores— y terminar en el otro. Si alguien empieza por el pico del tejado, fracasará siempre.

(c) El grafo completo K_5. Cada uno de los cinco vértices se une a los otros cuatro, así que todos los grados valen 4: cero impares. Hay circuito de Euler, y recorre las $\binom{5}{2} = 10$ aristas volviendo al inicio.

(d) El grafo completo K_6. Ahora cada vértice tiene grado 5, impar, y hay seis vértices impares. Imposible de un trazo; por el paso 6 hacen falta $6/2 = 3$ trazos para cubrir sus 15 aristas. En general K_n admite circuito de Euler exactamente cuando n es impar.

(e) El grafo bipartito completo $K_{3,3}$. Los seis vértices tienen grado 3. Seis impares: imposible, y hacen falta 3 trazos para sus 9 aristas.

Cuidado

El criterio es sobre el número de vértices impares, no sobre el número total de vértices ni sobre el de aristas. En (c) hay 10 aristas y sí se puede; en (e) hay 9 y no. Y ojo con el paso 1: un grafo con todos los grados pares pero desconectado en dos ciclos separados tampoco se dibuja de un trazo, aunque el recuento de impares dé cero.

Pausa para pensar

¿Por qué el número de vértices de grado impar es siempre par? Porque la suma de todos los grados cuenta cada arista dos veces, una por cada extremo, y por tanto vale $2|E|$, que es par. Si hubiera un número impar de sumandos impares, la suma total sería impar. Éste es el lema del apretón de manos: en cualquier reunión, el número de personas que han dado un número impar de apretones es par.

8.5. Recorrer un grafo: anchura y profundidad

Casi todo algoritmo sobre grafos empieza por visitar sus vértices en algún orden sistemático. Hay dos maneras canónicas de hacerlo, y elegir bien entre ellas resuelve de un plumazo problemas que planteados a mano parecen difíciles: encontrar las componentes conexas,

detectar ciclos, hallar el camino más corto cuando todas las aristas cuestan igual o salir de un laberinto.

Definición 8.27. El *recorrido en anchura* (o BFS) desde un vértice s visita primero s, luego todos sus vecinos, luego los vecinos de estos aún no visitados, y así sucesivamente; se implementa con una *cola*. El *recorrido en profundidad* (o DFS) avanza todo lo que puede por un camino antes de retroceder, y se implementa con una *pila* o con recursión. Ambos generan un *árbol de recorrido* formado por las aristas que descubren vértices nuevos. En un grafo con n vértices, m aristas y c componentes conexas, el número de aristas independientes que cierran ciclo es $m - n + c$, llamado *número ciclomático*.

Idea clave

La única diferencia entre los dos algoritmos es la estructura donde se guardan los vértices pendientes: una cola atiende al que lleva más tiempo esperando, una pila al último que llegó. Ese detalle mínimo cambia por completo lo que sirven para hacer. La cola garantiza que los vértices salen ordenados por distancia al origen, y por eso BFS resuelve el camino más corto en grafos sin pesos. La pila hace que el recorrido se hunda por una rama hasta el fondo, y por eso DFS es el que sirve para detectar ciclos, ordenar tareas o explorar un laberinto con la regla de la mano derecha. Ambos cuestan $\mathcal{O}(n + m)$, es decir, lo mínimo imaginable: mirar cada vértice y cada arista una vez.

Regla: recorrer un grafo y sacarle partido al recorrido

El esquema es común a los dos métodos y solo cambia la estructura de datos:

1. **Marca el vértice inicial como visitado** y colócalo en la cola (BFS) o en la pila (DFS).
2. **Extrae un vértice** de la estructura y examina todos sus vecinos.
3. **Los vecinos no visitados se marcan e insertan**, anotando de qué vértice vienen: ese dato es el árbol de recorrido y permite reconstruir caminos.
4. **En BFS, anota además la distancia** como la del padre más uno; queda calculado el camino mínimo en número de aristas.
5. **Repite hasta vaciar la estructura**; lo que quede sin visitar pertenece a otra componente conexa.
6. **Relanza el recorrido desde un vértice no visitado** para encontrar la siguiente componente, y cuenta cuántas veces has tenido que relanzarlo.
7. **Interpreta las aristas que no entraron en el árbol**: cada una cierra un ciclo, y hay exactamente $m - n + c$ de ellas.

Ejemplo 8.28 (Un grafo recorrido de las dos maneras)**. (a) El grafo.** Sea G con vértices $A, \ldots, I$ y aristas AB, AC, BD, CD, CE, DF, EF, FG, HI. Hay $n = 9$ vértices y $m = 9$

aristas; los grados son $2, 2, 3, 3, 2, 3, 1, 1, 1$ y suman $18 = 2m$, como manda el lema del apretón de manos. Hay dos componentes conexas: $\{A, B, C, D, E, F, G\}$ y $\{H, I\}$.

(b) Anchura desde A. Tomando los vecinos en orden alfabético, el orden de visita es A, B, C, D, E, F, G y las distancias resultan

v	A	B	C	D	E	F	G
$d(A, v)$	0	1	1	2	2	3	4

El árbol BFS está formado por las aristas AB, AC, BD, CE, DF y FG: seis aristas para siete vértices, como corresponde a un árbol. Las tres aristas restantes del grafo (CD, EF y HI) no entraron; las dos primeras cierran ciclo, y en efecto $m - n + c = 9 - 9 + 2 = 2$.

(c) Profundidad desde A. Con una pila y los mismos criterios, el orden es A, B, D, C, E, F, G y el árbol lo forman AB, BD, DC, CE, EF, FG: un camino, no una estrella. Compárese con el árbol BFS, que era mucho más achatado. Esa es la diferencia visual entre los dos recorridos: BFS produce árboles anchos y bajos, DFS produce árboles estrechos y profundos.

(d) Lo que cada uno resuelve. Con BFS sabemos ya que la distancia de A a G es 4 y podemos exhibir el camino $A \to B \to D \to F \to G$ reconstruyéndolo hacia atrás por los padres. Con DFS sabemos que al llegar a C desde D y encontrar que A ya estaba visitado se ha cerrado un ciclo, y podemos exhibirlo: $A \to B \to D \to C \to A$. Y con cualquiera de los dos, relanzando desde H, descubrimos la segunda componente.

Ejemplo 8.29 (Árboles, laberintos y tareas)**. (a) Reconocer un árbol.** El grafo con aristas AB, AC, BD, BE, CF tiene $n = 6$ y $m = 5$: cumple $m = n - 1$, es conexo y por tanto es un árbol. Sus hojas (vértices de grado 1) son D, E y F. Enraizado en A, su altura es 2. En un árbol, BFS y DFS recorren los mismos vértices pero en órdenes que corresponden respectivamente al recorrido por niveles y al recorrido en preorden.

(b) Un laberinto es un grafo rejilla. La cuadrícula de 4×4 celdas, con aristas entre celdas contiguas, tiene 16 vértices y 24 aristas. La distancia de la esquina $(0, 0)$ a la $(3, 3)$ es 6, que es lo mínimo posible porque hay que dar tres pasos en cada dirección. El número de caminos mínimos distintos es $\binom{6}{3} = 20$, y BFS los encuentra todos de golpe porque descubre la esquina opuesta exactamente en el nivel 6.

(c) Detectar bipartición con BFS. Si al hacer BFS coloreamos cada nivel alternando dos colores, el grafo es bipartito si y solo si ninguna arista une dos vértices del mismo nivel de paridad. El grafo del ejemplo anterior sí lo es. Este es el algoritmo estándar para comprobar la bipartición, y cuesta lo mismo que un recorrido.

(d) Ordenar tareas con DFS. En el grafo dirigido con arcos $a \to b$, $a \to c$, $b \to d$, $c \to d$, $d \to e$, $c \to e$ no hay ciclos, luego es un DAG y admite orden topológico: a, b, c, d, e es uno válido. El algoritmo clásico es un DFS que va apilando cada vértice cuando termina de explorarlo, y luego lee la pila al revés. Si durante el DFS se encuentra un arco hacia un vértice todavía en proceso, hay un ciclo y el orden topológico no existe.

Cuidado

Marcar el vértice como visitado hay que hacerlo *al insertarlo* en la cola, no al extraerlo; si se hace al extraerlo, el mismo vértice puede entrar varias veces y el coste se dispara. Segundo aviso: BFS da el camino mínimo en *número de aristas*, no en peso; en cuanto las aristas tienen costes distintos hay que usar Dijkstra, y aplicar BFS da una respuesta plausible y equivocada. Y tercero: DFS implementado con recursión puede agotar la pila del sistema en grafos grandes; la versión con pila explícita hace exactamente lo mismo sin ese riesgo, aunque el orden de visita puede diferir según cómo se apilen los vecinos.

Pausa para pensar

El recorrido en profundidad aparece por primera vez, con todas sus ideas, en un trabajo de 1882 del matemático francés Charles Pierre Trémaux sobre cómo salir de un laberinto: su regla era marcar cada pasillo al recorrerlo y no tomar nunca un pasillo marcado dos veces. La versión moderna, con su análisis de coste y su uso para descomponer grafos, es de Hopcroft y Tarjan en 1973, y le valió a Tarjan el premio Turing. El BFS es más joven en su formulación explícita (Moore, 1959, buscando caminos en laberintos para un ordenador), aunque Zuse lo había programado ya en 1945 sin publicarlo. Dos algoritmos que hoy se enseñan en la primera semana de cualquier curso tardaron casi un siglo en cristalizar.

8.6. Ciclos hamiltonianos: el problema gemelo que nadie sabe resolver

Definición 8.30. Un **ciclo hamiltoniano** es un camino cerrado que pasa por cada *vértice* exactamente una vez.

Idea clave

Qué dice esto en cristiano. Euler y Hamilton se parecen en el enunciado y no tienen nada que ver en la dificultad. Un circuito euleriano recorre todas las *aristas* una vez, y hay un criterio inmediato para saber si existe: que todos los grados sean pares. Un ciclo hamiltoniano recorre todos los *vértices* una vez, y *no* se conoce ningún criterio eficiente: decidir si existe es un problema NP-completo. Un cambio de una palabra en el enunciado convierte un ejercicio de un minuto en uno de los problemas abiertos centrales de la informática teórica.

Ejemplo 8.31 (Euleriano frente a hamiltoniano)**. (a) Los puentes de Kœnigsberg.** El multigrafo de Euler tiene cuatro vértices de grados $3, 3, 3, 5$: los cuatro impares. Como un circuito euleriano exige todos pares y un recorrido abierto admite exactamente dos impares, no existe ningún paseo que cruce los siete puentes una sola vez. La respuesta cabe en una línea.

(b) El grafo completo. K_5 tiene todos los grados iguales a 4, pares, luego es euleriano; y también es hamiltoniano, porque cualquier ordenación de los cinco vértices sirve. En general

K_n es hamiltoniano para $n \geq 3$ y euleriano si y sólo si n es impar.

(c) Euleriano sin ser hamiltoniano y viceversa. El grafo "pajarita" (dos triángulos pegados por un vértice) tiene todos los grados pares, luego es euleriano, y no es hamiltoniano, porque el vértice central habría que visitarlo dos veces. Al revés, el cuadrado C_4 con una diagonal añadida es hamiltoniano y tiene dos vértices de grado 3, luego no es euleriano.

(d) El criterio de Dirac. Aunque no haya caracterización, sí hay condiciones suficientes: si $n \geq 3$ y todo vértice tiene grado al menos $n/2$, el grafo es hamiltoniano. Con $n = 6$ eso pide grado ≥ 3. La condición es suficiente y no necesaria: el ciclo C_6 tiene todos los grados iguales a 2 y es hamiltoniano por definición.

El cambio de "arista" a "vértice" parece inocente y no lo es en absoluto.

Cuidado

No existe ningún criterio como el de Euler para los ciclos hamiltonianos. Hay condiciones suficientes —por ejemplo, el teorema de Dirac: basta con que todos los grados sean al menos $n/2$—, pero ningún criterio que sea a la vez necesario y suficiente y fácil de comprobar. Decidir si un grafo dado tiene ciclo hamiltoniano es un problema *NP-completo*: nadie conoce un algoritmo esencialmente mejor que probar las posibilidades, y saber si existe uno es la formulación más famosa de la pregunta abierta "¿$P = NP$?".

Observación 8.32. La versión ponderada —encontrar el ciclo hamiltoniano de menor peso total— es el **problema del viajante de comercio**, probablemente el problema de optimización más estudiado de la historia. Con n ciudades hay $(n-1)!/2$ recorridos distintos; para $n = 20$ eso ya son unos 6×10^{16}, y examinarlos todos a mil millones por segundo llevaría casi dos años. Aquí se ve por qué la programación lineal de la programación lineal es tan valiosa: cuando un problema *sí* tiene estructura lineal y continua, se resuelve rápido; cuando exige soluciones enteras o combinatorias, la dificultad explota.

8.7. Árboles

Definición 8.33. Un **árbol** es un grafo conexo sin ciclos. Sus vértices de grado 1 se llaman **hojas**.

Ejemplo 8.34 (Árboles y sus caracterizaciones)**. (a) Cinco definiciones equivalentes.** Para un grafo G con n vértices, son equivalentes: es conexo y sin ciclos; es conexo con $n-1$ aristas; no tiene ciclos y tiene $n-1$ aristas; entre dos vértices cualesquiera hay un único camino; es conexo y quitar cualquier arista lo desconecta. Elegir la más cómoda en cada demostración ahorra mucho trabajo.

(b) Contar árboles. La fórmula de Cayley dice que hay n^{n-2} árboles etiquetados con n vértices: 1 para $n = 2$, 3 para $n = 3$, 16 para $n = 4$ y 125 para $n = 5$. Los tres de $n = 3$ son los tres caminos según cuál sea el vértice central.

(c) Hojas. Todo árbol con al menos dos vértices tiene al menos dos hojas. La razón es de conteo: si todos los grados fuesen ≥ 2, la suma de grados sería $\geq 2n$, pero vale $2(n-1) < 2n$. Es el argumento que arranca casi todas las inducciones sobre árboles.

(d) Árboles de expansión. Todo grafo conexo contiene un árbol que toca todos sus vértices. El cuadrado C_4 tiene cuatro árboles de expansión, uno por cada arista que se elimine. El algoritmo de Kruskal encuentra el de peso mínimo ordenando las aristas y añadiendo las que no cierren ciclo.

Teorema 8.35. *Un árbol con n vértices tiene exactamente $n-1$ aristas.*

Demostración. Por inducción sobre n. Para $n = 1$ el árbol es un punto suelto y tiene $0 = 1-1$ aristas. Supóngase cierto para n y sea un árbol con $n+1$ vértices. Todo árbol con al menos dos vértices tiene alguna hoja (si todos los vértices tuvieran grado ≥ 2 podría seguirse caminando sin volver atrás y, al ser el grafo finito, se repetiría un vértice, formando un ciclo). Quítese una hoja junto con su única arista: lo que queda sigue siendo conexo y sin ciclos, es decir, un árbol con n vértices, que por hipótesis tiene $n-1$ aristas. Devolviendo la hoja se obtienen n aristas y $n+1$ vértices. □

Ejemplo 8.36 (La cuenta de aristas en árboles)**. (a) Casos pequeños.** Un vértice suelto: 0 aristas. Dos vértices unidos: 1. Un camino de cuatro vértices: 3. Una estrella con un centro y cinco hojas: 6 vértices y 5 aristas. Siempre uno menos.

(b) La suma de grados. En un árbol con n vértices, $\sum \deg(v) = 2(n-1)$. Para la estrella de seis vértices: el centro aporta 5 y las cinco hojas aportan 1 cada una, total $10 = 2 \cdot 5$. Cuadra.

(c) Una consecuencia inmediata. Un grafo conexo con n vértices y m aristas tiene exactamente $m - n + 1$ ciclos independientes (su *número ciclomático*). El cuadrado C_4 tiene $4 - 4 + 1 = 1$; el grafo completo K_4 tiene $6 - 4 + 1 = 3$. Ese número es la dimensión del espacio de ciclos y, en teoría de circuitos, el número de ecuaciones de malla independientes de Kirchhoff.

(d) El contraejemplo obligado. Un grafo con n vértices y $n-1$ aristas no tiene por qué ser un árbol: sobre cuatro vértices, un triángulo más un vértice aislado tiene $3 = 4-1$ aristas y no es conexo ni acíclico. Hace falta añadir una de las dos condiciones; con cualquiera de ellas, la otra sale gratis.

Idea clave

Los árboles son los grafos "justos": tienen las aristas mínimas para estar conectados y las máximas para no tener ciclos. Quitar cualquier arista de un árbol lo desconecta; añadir cualquier arista le crea un ciclo. Esa doble rigidez es lo que los hace aparecer en todas partes: los organigramas, los índices de un libro, la jerarquía de carpetas de un ordenador, el árbol genealógico y la estructura de decisiones de un algoritmo son todos árboles.

Observación 8.37. Cayley demostró en 1889 que el número de árboles distintos que pueden formarse sobre n vértices *etiquetados* es n^{n-2}. Para $n = 4$ eso da 16; para $n = 10$, cien millones. La fórmula es de una simplicidad sospechosa y admite media docena de demostraciones muy distintas, ninguna de ellas evidente.

8.8. Grafos planos y la fórmula de Euler

Definición 8.38. Un grafo es **plano** si puede dibujarse en el papel sin que dos aristas se crucen. Un dibujo así divide al plano en regiones llamadas **caras**, contando siempre la región exterior ilimitada como una cara más.

Teorema 8.39 (Fórmula de Euler)**.** *En todo grafo plano conexo dibujado sin cruces, con V vértices, A aristas y C caras,*

$$V - A + C = 2.$$

Idea de la demostración. Por inducción sobre el número de aristas. Si el grafo es un árbol, no hay ciclos y por tanto solo hay una cara: $V - (V-1) + 1 = 2$ ✓. Si hay algún ciclo, quítese una de sus aristas: el grafo sigue siendo conexo, tiene una arista menos y, como esa arista separaba dos caras que ahora se funden en una, tiene también una cara menos. La cantidad $V - A + C$ no cambia, y por inducción vale 2. □

Ejemplo 8.40. El cubo, visto como grafo, tiene $V = 8$ vértices, $A = 12$ aristas y $C = 6$ caras: $8 - 12 + 6 = 2$ ✓. Lo mismo vale para el tetraedro ($4 - 6 + 4 = 2$), el octaedro ($6 - 12 + 8 = 2$) y cualquier poliedro convexo, porque todos ellos pueden "aplastarse" sobre el plano. Esta fórmula, escrita para poliedros, es de 1750 y es una de las primeras verdades topológicas de la historia: no depende de longitudes ni de ángulos, solo de cómo están pegadas las piezas.

Corolario 8.41. *En un grafo plano simple y conexo con $V \geq 3$ se cumple $A \leq 3V - 6$.*

Idea de la demostración. Cada cara está bordeada por al menos tres aristas, y cada arista bordea a lo sumo dos caras, luego $3C \leq 2A$. Sustituyendo $C = 2 - V + A$ en esa desigualdad y despejando sale $A \leq 3V - 6$. □

Ejemplo 8.42. K_5 no es plano: tiene $V = 5$ y $A = 10$, pero $3 \cdot 5 - 6 = 9 < 10$. Por muchas vueltas que se le dé al dibujo, dos aristas tendrán que cruzarse.

$K_{3,3}$ —tres casas y tres suministros, cada casa conectada a cada suministro— tampoco lo es. Aquí $A = 9$ y $3V - 6 = 12$, de modo que la cota anterior no lo descarta; hace falta afinar: en $K_{3,3}$ no hay triángulos, así que cada cara necesita al menos *cuatro* aristas, lo que da $4C \leq 2A$ y de ahí $A \leq 2V - 4 = 8 < 9$. Contradicción.

Observación 8.43. El teorema de Kuratowski (1930) cierra la cuestión de manera espectacular: un grafo es plano *si y solo si* no contiene, escondida dentro, una copia de K_5 ni de $K_{3,3}$. Es decir, hay exactamente dos obstrucciones en el universo entero de los grafos, y toda no planaridad se debe a una de las dos. Resultados de esa forma —"una lista finita de casos prohibidos caracteriza toda una propiedad"— son rarísimos y muy apreciados en matemática.

Regla: usar la fórmula de Euler para grafos planos

La fórmula $V - E + F = 2$ relaciona vértices, aristas y caras de cualquier dibujo plano conexo, y de ella salen unas cotas que permiten demostrar que ciertos grafos *no* se pueden dibujar sin cruces, sin intentarlo nunca.

1. **Cuenta bien las caras.** F incluye la cara exterior, la región infinita de alrededor. Olvidarla es el error número uno y desplaza la fórmula en una unidad.
2. **Verifica la conexión.** Para un grafo plano conexo, $V - E + F = 2$. Si tiene c componentes, la fórmula correcta es $V - E + F = 1 + c$.
3. **Deduce la cota general.** Si el grafo es simple y tiene $V \geq 3$, cada cara está bordeada por al menos 3 aristas y cada arista bordea a lo sumo 2 caras, luego $3F \leq 2E$. Sustituyendo en la fórmula sale
$$E \leq 3V - 6.$$
4. **Si además no hay triángulos** (por ejemplo, si el grafo es bipartito), cada cara necesita al menos 4 aristas y la cota se afina a
$$E \leq 2V - 4.$$
5. **Úsalas al revés, como test de imposibilidad.** Si un grafo viola la cota, no es plano. Si la cumple, el test no concluye nada: hay que buscar un dibujo o invocar el teorema de Kuratowski.
6. **Recuerda a Kuratowski para el caso general:** un grafo es plano si y solo si no contiene una subdivisión de K_5 ni de $K_{3,3}$. Esos dos son los únicos obstáculos posibles.
7. **Control:** en todo grafo plano simple existe siempre algún vértice de grado ≤ 5. Se deduce de la cota del paso 3, y es la base de la demostración de que cinco colores bastan.

Ejemplo 8.44 (La fórmula en acción)**. (a) El cubo.** Tiene $V = 8$ vértices, $E = 12$ aristas y $F = 6$ caras. Comprobación: $8 - 12 + 6 = 2$. ✓ Y para los cinco sólidos platónicos sale siempre lo mismo: tetraedro $4 - 6 + 4 = 2$, octaedro $6 - 12 + 8 = 2$, dodecaedro $20 - 30 + 12 = 2$, icosaedro $12 - 30 + 20 = 2$.

(b) Un grafo plano cualquiera. Si sabes que un grafo plano conexo tiene $V = 6$ y $E = 9$, ya sabes cuántas caras tendrá cualquier dibujo suyo sin cruces: $F = 2 - V + E = 2 - 6 + 9 = 5$. Y esto es cierto *independientemente* de cómo lo dibujes: el número de regiones no depende del trazado. Es un invariante topológico.

(c) K_5 no es plano. Tiene $V = 5$ y $E = \binom{5}{2} = 10$. La cota del paso 3 da $3 \cdot 5 - 6 = 9$, y $10 > 9$. Contradicción: no existe ningún dibujo de K_5 sin cruces. Nótese la economía del argumento: no hemos intentado dibujarlo ni una sola vez.

(d) $K_{3,3}$ tampoco. Tiene $V = 6$ y $E = 9$, y la cota del paso 3 da $3 \cdot 6 - 6 = 12$, que se

cumple: ese test no decide. Pero $K_{3,3}$ es bipartito y por tanto no tiene triángulos, así que se aplica el paso 4: $2 \cdot 6 - 4 = 8$, y $9 > 8$. No es plano. Éste es el famoso problema de las tres casas y los tres servicios (agua, gas y luz): no hay manera de conectarlos todos sin que dos tuberías se crucen.

(e) El grafo de Petersen. Tiene $V = 10$ y $E = 15$; la cota da $3 \cdot 10 - 6 = 24$, que se cumple holgadamente. Y sin embargo no es plano: contiene una subdivisión de $K_{3,3}$. Es el ejemplo que muestra por qué el paso 5 avisa de que la cota es solo una condición *necesaria*.

Idea clave

La fórmula $V - E + F = 2$ no es un hecho sobre dibujos, sino sobre la esfera. El número $V - E + F$ se llama característica de Euler y vale 2 para todo lo que sea topológicamente una esfera, pero 0 para el toro (la rosquilla). Por eso K_5 y $K_{3,3}$, que no caben en el plano, sí se dibujan sin cruces sobre un toro: hay un asa por la que pasar por debajo.

Cuidado

La cota $E \leq 3V - 6$ exige que el grafo sea *simple*: sin lazos ni aristas múltiples. Con aristas repetidas se pueden tener tantas como se quiera entre dos vértices y el dibujo sigue siendo plano, aunque la cota se viole por completo. Comprueba siempre esa hipótesis antes de concluir que algo no es plano.

Regla: contar caminos con la matriz de adyacencia

Un grafo se puede guardar como una matriz, y entonces todo el álgebra lineal del cálculo matricial elemental se vuelve teoría de grafos.

1. **Fijar un orden de los vértices** $v_1, \ldots, v_n$ y escribir la **matriz de adyacencia** A, con $a_{ij} = 1$ si v_i y v_j son adyacentes y $a_{ij} = 0$ si no. En un grafo no dirigido A es simétrica y su diagonal es nula.
2. **El resultado central:** la entrada (i, j) de A^k cuenta los *recorridos* de longitud exactamente k que van de v_i a v_j. Recorrido significa que se permite repetir vértices y aristas.
3. **Por qué:** $(A^k)_{ij} = \sum_r (A^{k-1})_{ir} a_{rj}$, que es literalmente "para llegar a j en k pasos, llego a algún vecino r en $k - 1$ pasos y doy el último paso". La multiplicación de matrices *es* el principio multiplicativo.
4. **Diagonal de A^2:** $(A^2)_{ii} = \deg(v_i)$, porque los recorridos de longitud 2 que salen y vuelven a v_i son "ir a un vecino y volver".
5. **Diagonal de A^3:** $(A^3)_{ii}$ cuenta los triángulos que pasan por v_i, cada uno dos veces (una por sentido). Por tanto el número total de triángulos del grafo es $\operatorname{tr}(A^3)/6$.
6. **Conexión:** el grafo es conexo si y solo si $(A + I)^{n-1}$ no tiene ninguna entrada nula. Y la distancia entre v_i y v_j es el menor k con $(A^k)_{ij} > 0$.

7. **Verificar** sumando: la suma de la fila i de A debe ser $\deg(v_i)$, y la suma de todas las entradas debe ser $2m$.

Ejemplo 8.45 (una matriz 5×5 que lo dice todo)**.** Sea G el grafo con vértices $1, \dots, 5$ y aristas $12, 13, 23, 24, 34, 45$. Con ese orden,

$$A = \begin{pmatrix} 0 & 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 1 & 0 \\ 1 & 1 & 0 & 1 & 0 \\ 0 & 1 & 1 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 \end{pmatrix}.$$

(a) *Comprobación previa.* Las sumas de las filas son $2, 3, 3, 3, 1$, que son los grados; su suma es $12 = 2 \cdot 6$, y en efecto hay 6 aristas. ✓

(b) *Recorridos de longitud* 2*.*

$$A^2 = \begin{pmatrix} 2 & 1 & 1 & 2 & 0 \\ 1 & 3 & 2 & 1 & 1 \\ 1 & 2 & 3 & 1 & 1 \\ 2 & 1 & 1 & 3 & 0 \\ 0 & 1 & 1 & 0 & 1 \end{pmatrix}.$$

La diagonal $(2, 3, 3, 3, 1)$ reproduce los grados, como anunciaba la regla. La entrada $(1, 4) = 2$ dice que hay exactamente dos maneras de ir de 1 a 4 en dos pasos: $1{-}2{-}4$ y $1{-}3{-}4$. La entrada $(1, 5) = 0$ dice que no hay ninguna, y es cierto: 5 solo se toca desde 4, que está a distancia 2 de 1.

(c) *Recorridos de longitud* 3 *y triángulos.*

$$A^3 = \begin{pmatrix} 2 & 5 & 5 & 2 & 2 \\ 5 & 4 & 5 & 6 & 1 \\ 5 & 5 & 4 & 6 & 1 \\ 2 & 6 & 6 & 2 & 3 \\ 2 & 1 & 1 & 3 & 0 \end{pmatrix}, \qquad \operatorname{tr}(A^3) = 2 + 4 + 4 + 2 + 0 = 12.$$

Entonces el grafo tiene $12/6 = 2$ triángulos, y se ven a simple vista: 123 y 234. Nótese que el vértice 5 tiene un 0 en la diagonal, como debe ser: no pertenece a ningún triángulo.

(d) *Distancias.* La entrada $(1, 5)$ vale 0 en A y en A^2, y vale 2 en A^3: la distancia de 1 a 5 es 3, y hay dos caminos mínimos, $1{-}2{-}4{-}5$ y $1{-}3{-}4{-}5$. Este es, en pequeño, el mismo dato que produce el algoritmo de Dijkstra cuando todos los pesos valen 1.

(e) *Conexión.* La matriz $(A+I)^4$ resulta ser

$$\begin{pmatrix} 31 & 39 & 39 & 32 & 10 \\ 39 & 51 & 51 & 44 & 16 \\ 39 & 51 & 51 & 44 & 16 \\ 32 & 44 & 44 & 42 & 18 \\ 10 & 16 & 16 & 18 & 10 \end{pmatrix},$$

sin ningún cero: el grafo es conexo. El I está ahí para permitir "quedarse quieto", de modo que la potencia $n-1$ recoge caminos de longitud *hasta* 4 y no exactamente 4.

(f) *Un aviso sobre las palabras.* A^k cuenta *recorridos*, no caminos simples. En la entrada $(1,1)$ de A^3 hay un 2, y corresponden a $1{-}2{-}3{-}1$ y $1{-}3{-}2{-}1$; pero en $(1,1)$ de A^4 habría también cosas como $1{-}2{-}1{-}2{-}1$, que repite aristas. Contar caminos *sin* repeticiones es un problema muchísimo más difícil, y no hay ninguna fórmula matricial que lo haga.

Cuidado

La matriz depende del orden elegido para los vértices: renumerar los vértices sustituye A por $P^{-1}AP = P^{\mathsf{T}}AP$ con P una matriz de permutación. Dos grafos son isomorfos exactamente cuando sus matrices son conjugadas por una permutación, y por eso los autovalores —que son invariantes por conjugación— son un invariante del grafo, mientras que las entradas de A no lo son.

Idea clave

Esto es una máquina de Markov disfrazada. Si en lugar de A se usa la matriz que divide cada fila por el grado, las potencias dejan de contar recorridos y pasan a dar *probabilidades* de un paseo aleatorio. El vector propio dominante de esa matriz mide entonces la importancia de cada vértice, y esa es, literalmente, la idea del PageRank con el que Google ordenó la web: un grafo, una matriz, un vector propio.

8.9. Grafos bipartitos y emparejamientos

Muchos problemas de la vida diaria tienen la misma forma: hay dos colecciones de objetos —trabajadores y puestos, estudiantes y proyectos, órganos y receptores— y solo ciertos cruces son admisibles. El grafo que los describe es **bipartito**: sus vértices se reparten en dos bandos X e Y y toda arista va de un bando al otro, nunca dentro del mismo. Lo que se busca en él es un **emparejamiento**: un conjunto de aristas sin vértices repetidos.

Definición 8.46. Un grafo es **bipartito** si sus vértices pueden repartirse en dos conjuntos disjuntos X e Y de manera que toda arista tenga un extremo en X y el otro en Y. Un **emparejamiento** es un conjunto M de aristas dos a dos sin extremos comunes; es **máximo** si ningún otro emparejamiento tiene más aristas, y **perfecto** si cubre todos los vértices.

Ejemplo 8.47 (Grafos bipartitos por todas partes)**. (a) El cuadrado y el pentágono.** C_4 es bipartito con $X = \{1, 3\}$ e $Y = \{2, 4\}$. C_5 no lo es: al intentar colorear alternando se llega al quinto vértice con un conflicto, porque 5 es impar.

(b) El bipartito completo. $K_{3,3}$ tiene $3 + 3 = 6$ vértices y 9 aristas. Es el grafo de las "tres casas y tres pozos", y es uno de los dos grafos prohibidos del teorema de Kuratowski: un grafo es plano si y sólo si no contiene una subdivisión de $K_{3,3}$ ni de K_5.

(c) Emparejamientos. En un bipartito, el teorema de Hall dice que existe un emparejamiento que satura X si y sólo si todo subconjunto $S \subseteq X$ tiene al menos $|S|$ vecinos. Con X = tres trabajadores y Y = tres tareas, si dos de los trabajadores sólo saben hacer la misma única tarea, la condición falla para ese par y no hay asignación posible.

(d) Árboles y cuadrículas. Todo árbol es bipartito (colorear por paridad de la distancia a la raíz), y toda cuadrícula también, con el patrón del tablero de ajedrez. De ahí sale el clásico problema del tablero mutilado: quitando dos esquinas opuestas quedan 32 casillas de un color y 30 del otro, y como cada ficha de dominó cubre una de cada, el recubrimiento es imposible.

Teorema 8.48 (Criterio de bipartición)**.** *Un grafo es bipartito si y solo si no contiene ningún ciclo de longitud impar.*

Demostración. Si es bipartito, todo ciclo alterna entre X e Y, así que para volver al vértice de partida hacen falta un número par de pasos. Recíprocamente, supóngase que no hay ciclos impares y tómese el grafo conexo (si no lo es, se razona en cada componente). Fíjese un vértice v_0 y colóquese en X todo vértice cuya distancia a v_0 sea par, y en Y los de distancia impar. Si hubiera una arista uw con u, w en el mismo bando, los caminos mínimos de v_0 a u y a w más esa arista producirían un ciclo cerrado de longitud impar. □

Ejemplo 8.49 (Aplicar el criterio de bipartición)**. (a) El algoritmo.** Se recorre el grafo en anchura desde un vértice, asignando a cada uno el color de la paridad de su nivel. Si alguna arista une dos vértices del mismo nivel o de niveles de la misma paridad, hay ciclo impar y no es bipartito. El coste es lineal en el número de aristas.

(b) El ciclo par. C_6 no contiene ningún ciclo impar (su único ciclo tiene longitud 6), luego es bipartito: $\{1, 3, 5\}$ contra $\{2, 4, 6\}$.

(c) El triángulo escondido. El grafo de Petersen tiene 10 vértices, 15 aristas y su ciclo más corto tiene longitud 5: impar, luego no es bipartito, aunque a simple vista no se vea ningún triángulo. El criterio no exige buscar triángulos, sino ciclos impares de cualquier longitud.

(d) La consecuencia sobre colores. Bipartito equivale a ser 2-coloreable. Como todo ciclo impar necesita tres colores, el número cromático de un grafo con ciclos impares es al menos 3. Para C_5 vale exactamente 3, y para K_n vale n.

Regla: decidir si un grafo es bipartito y hallar un emparejamiento máximo

1. **Colorea con dos colores y sal de dudas.** Empieza en un vértice cualquiera, píntalo de blanco, pinta de negro a todos sus vecinos, de blanco a los vecinos de

estos, y así sucesivamente. Si en algún momento una arista une dos vértices del mismo color, has encontrado un ciclo impar y el grafo *no* es bipartito. Si terminas sin conflicto, los dos colores son los bandos X e Y.

2. **Anota los tamaños.** Si $|X| \neq |Y|$ no puede haber emparejamiento perfecto, y el máximo posible es mín$(|X|, |Y|)$. Este control de un segundo resuelve muchos problemas antes de empezar.

3. **Construye un emparejamiento inicial a ojo.** Ve tomando aristas libres mientras puedas, sin pensarlo mucho. Casi nunca será el máximo, pero es un punto de partida perfectamente válido.

4. **Busca un camino aumentante.** Parte de un vértice de X sin emparejar y avanza alternando: arista fuera de M, arista de M, arista fuera de M... Si llegas a un vértice de Y sin emparejar, has encontrado un camino aumentante.

5. **Inviértelo.** Cambia en ese camino lo que estaba emparejado por lo que no lo estaba. Como el camino tiene una arista más fuera de M que dentro, el emparejamiento crece exactamente en una unidad.

6. **Repite hasta que no haya camino aumentante** (teorema de Berge): un emparejamiento es máximo *si y solo si* no admite ningún camino aumentante. Ese es el criterio de parada, y también la demostración de que has terminado.

7. **Certifica el resultado con un cubrimiento** (teorema de König): en un grafo bipartito, el tamaño del emparejamiento máximo coincide con el del *cubrimiento por vértices* más pequeño, es decir, del menor conjunto de vértices que toca todas las aristas. Si exhibes un emparejamiento de k aristas y un cubrimiento de k vértices, tienes la prueba cerrada: nadie podrá superar k.

Ejemplo 8.50 (Siete emparejamientos, y uno que resulta imposible)**.** (a) **Dos ciclos y un veredicto.** El hexágono C_6 con vértices $1, 2, 3, 4, 5, 6$ y aristas consecutivas es bipartito: pintando alternativamente sale $X = \{1, 3, 5\}$, $Y = \{2, 4, 6\}$. El pentágono C_5 no lo es: al dar la vuelta completa con cinco pasos se regresa al color contrario del de partida, y la última arista une dos vértices del mismo color. La longitud impar es el único obstáculo.

(b) **Una cuerda inocente y otra fatal.** Añádase al hexágono la cuerda 1–4: como $1 \in X$ y $4 \in Y$, la bipartición sobrevive intacta. Añádase en cambio la cuerda 1–3: ambos están en X, aparece el triángulo $1, 2, 3$ y el grafo deja de ser bipartito. La misma operación, un color de diferencia y dos resultados opuestos.

(c) **Cuatro trabajadores, cuatro tareas.** Sean $X = \{a, b, c, d\}$ y $Y = \{1, 2, 3, 4\}$ con las compatibilidades $a1$, $a2$, $b1$, $c2$, $c3$, $d3$, $d4$. Empezando a ojo con $M = \{a1, c2, d3\}$, el trabajador b queda suelto y solo sabe hacer la tarea 1, ya ocupada por a. Camino aumentante desde b:

$$b \;-\; 1 \;=\; a \;-\; 2 \;=\; c \;-\; 3 \;=\; d \;-\; 4,$$

donde − marca aristas fuera de M y = aristas de M. Termina en 4, que está libre: se invierte todo y queda $M' = \{b1,\ a2,\ c3,\ d4\}$, de tamaño 4. Es perfecto, así que desde luego es máximo.

(d) **El mismo problema, sin solución.** Cámbiense las compatibilidades a $a1$, $a2$, $b1$, $b2$, $c1$, $c2$, $d3$. El emparejamiento máximo tiene tamaño 3 —por ejemplo $\{c1,\ b2,\ d3\}$— y uno de los tres primeros trabajadores se queda siempre sin tarea. La razón se ve en un instante tomando $S = \{a, b, c\}$: el conjunto de tareas que ese trío sabe hacer es $N(S) = \{1, 2\}$, y *tres* personas no caben en *dos* puestos.

(e) **La condición de Hall.** El apartado anterior es el caso general: existe un emparejamiento que satura todo X si y solo si $|N(S)| \geq |S|$ para *todo* subconjunto $S \subseteq X$. Para demostrar que un emparejamiento perfecto es imposible no hace falta ensayar combinaciones: basta exhibir un solo S que incumpla la desigualdad, y ese S es una prueba completa, breve y verificable por cualquiera.

(f) **El certificado de König.** En el apartado (d), el conjunto de vértices $\{1, 2, d\}$ toca todas las aristas del grafo: es un cubrimiento de tamaño 3. Como cada arista de un emparejamiento necesita al menos un vértice propio del cubrimiento, ningún emparejamiento puede tener más de 3 aristas. Emparejamiento de tamaño 3 y cubrimiento de tamaño 3: los dos números coinciden y el problema queda cerrado por arriba y por abajo.

(g) **El tablero mutilado.** Quítense a un tablero de ajedrez de 8×8 las dos casillas de *esquinas opuestas* y trátese de cubrir lo que queda con 31 fichas de dominó, cada una tapando dos casillas contiguas. Quedan 62 casillas, luego los números cuadran. Y sin embargo es imposible: constrúyase el grafo cuyos vértices son las casillas y cuyas aristas unen casillas contiguas. Es bipartito, porque casillas contiguas tienen colores distintos, y las dos esquinas retiradas son del *mismo* color, de modo que quedan 32 casillas de un color y solo 30 del otro. Un dominó es exactamente una arista, y un emparejamiento no puede saturar el bando de 32 contra un bando de 30. El máximo alcanzable es 30 fichas, dejando siempre dos casillas huérfanas.

Idea clave

Los tres teoremas de esta sección —Berge, Hall y König— dicen lo mismo desde tres alturas distintas, y conviene verlos juntos. Berge da un *algoritmo*: mientras haya camino aumentante, sigue. Hall da un *criterio*: la única manera de que falle un emparejamiento total es que algún grupo tenga demasiados pocos vecinos. König da un *certificado*: un número pequeño de vértices que demuestra la imposibilidad sin más argumentos. En matemática aplicada esa tercera forma es oro puro, porque convierte una afirmación negativa —"no existe"—, que en principio exigiría examinar todos los casos, en un objeto concreto que cualquiera puede comprobar en un minuto.

Cuidado

El error habitual al buscar caminos aumentantes es olvidar que deben *alternar* y que deben *empezar y terminar* en vértices no emparejados. Un camino que arranca en un

vértice libre y acaba en uno ya emparejado no sirve para nada: al invertirlo no se gana ninguna arista. Y tampoco vale el atajo de "emparejar primero al que menos opciones tiene": es una heurística razonable que a menudo funciona, pero no garantiza el máximo y no sustituye al criterio de Berge.

Pausa para pensar

Cuenta los emparejamientos perfectos de $K_{3,3}$, el grafo bipartito completo con tres vértices en cada bando. Cada uno equivale a decidir qué elemento de Y le toca a cada elemento de X sin repetir, es decir, a una permutación de tres elementos: hay $3! = 6$. En general, $K_{n,n}$ tiene $n!$ emparejamientos perfectos, y esa identificación entre emparejamientos y permutaciones es la puerta de entrada a toda la combinatoria del capítulo 5 y al grupo simétrico del capítulo 23. Un emparejamiento perfecto de $K_{n,n}$ *es* una permutación, disfrazada de dibujo.

8.10. Coloración y el teorema de los cuatro colores

Definición 8.51. Una **coloración** de un grafo asigna un color a cada vértice de modo que dos vértices adyacentes nunca tengan el mismo. El **número cromático** $\chi(G)$ es la menor cantidad de colores con la que se puede lograr.

Ejemplo 8.52. $\chi(K_n) = n$, porque todos los vértices son adyacentes entre sí y no hay manera de repetir ningún color.

Para un ciclo de longitud par, $\chi = 2$: se alternan dos colores y encajan. Para un ciclo de longitud *impar*, $\chi = 3$: al dar la vuelta, el último vértice choca con el primero y hay que gastar un tercer color. Un ciclo de 5 vértices necesita 3; uno de 6, solo 2.

El problema nació de los mapas. Si se pone un vértice en cada país y una arista entre países vecinos, el grafo resultante es plano, y colorear el mapa sin que dos vecinos compartan color es exactamente colorear ese grafo.

Teorema 8.53 (Cuatro colores; Appel y Haken, 1976)**.** *Todo grafo plano admite una coloración con cuatro colores.*

Idea clave

Qué dice esto en cristiano. Cualquier mapa dibujado en un papel, por retorcidas que sean sus fronteras, se puede colorear con cuatro colores sin que dos países vecinos compartan color. Que cinco colores bastan es un ejercicio de una página; que cuatro bastan resistió ciento veinticuatro años y sólo se demostró comprobando por ordenador más de mil quinientas configuraciones. Fue la primera demostración importante que ningún humano puede verificar a mano, y abrió un debate sobre qué significa "demostrar" que todavía colea.

Ejemplo 8.54 (Colorear mapas y grafos)**. (a) Cuándo hacen falta cuatro.** El grafo K_4 es plano y necesita cuatro colores, así que la cota no se puede mejorar. En términos de mapas: cuatro países mutuamente fronterizos, como se puede dibujar con tres países alrededor de uno central.

(b) Cuándo bastan dos o tres. Un mapa hecho de rectas que atraviesan el plano de lado a lado siempre se colorea con dos colores, porque cada recta invierte la paridad. Un mapa cuyas fronteras se cortan siempre de tres en tres necesita como mucho tres.

(c) En otras superficies el número cambia. Sobre un toro hacen falta siete colores, y esa cota se alcanza: se puede dibujar en el toro un mapa de siete países mutuamente vecinos, es decir, un K_7. La fórmula de Heawood da $\lfloor (7+\sqrt{49-24\chi})/2 \rfloor$, que con $\chi = 0$ da 7 y con $\chi = 2$ daría 4.

(d) La versión fácil. Cinco colores siempre bastan, y la demostración es de una página: por la fórmula de Euler todo grafo plano tiene un vértice de grado ≤ 5; se elimina, se colorea el resto por inducción y se recoloca intercambiando colores en una cadena de Kempe. El salto de cinco a cuatro es donde el argumento se rompe y donde hizo falta el ordenador.

Observación 8.55. La conjetura se planteó en 1852 y resistió ciento veinticuatro años. La demostración de 1976 fue la primera de un teorema importante que requirió un ordenador: reduce el problema a comprobar 1936 configuraciones inevitables, una por una, cosa imposible a mano. Aquello abrió un debate que aún colea: ¿es una demostración algo que ningún ser humano puede verificar por completo? Hoy existe una versión formalizada y verificada por un asistente de pruebas (Coq, 2005), lo que traslada la confianza del ordenador al programa que comprueba al ordenador, pero no la elimina.

Curiosamente, la versión *fácil* del teorema —que bastan cinco colores— se demuestra en una página con la fórmula de Euler, y era conocida desde 1890. El salto de cinco a cuatro costó otros ochenta y seis años.

Regla: colorear un grafo y acotar el número cromático

El **número cromático** $\chi(G)$ es el mínimo de colores necesarios para pintar los vértices sin que dos vecinos compartan color. Calcularlo exactamente es difícil; acorralarlo entre dos números es rutina.

1. **Cota inferior por clique.** Si G contiene k vértices mutuamente adyacentes, esos k necesitan k colores distintos: $\chi(G) \geq \omega(G)$, donde ω es el tamaño del mayor subgrafo completo.
2. **Cota superior por el algoritmo voraz.** Se recorren los vértices en algún orden y a cada uno se le da el menor color no usado por sus vecinos ya pintados. Como cada vértice tiene a lo sumo Δ vecinos, nunca hacen falta más de $\Delta + 1$ colores: $\chi(G) \leq \Delta(G) + 1$.
3. **Afinar la cota superior con Brooks:** si G es conexo y no es ni un grafo completo ni un ciclo impar, entonces $\chi(G) \leq \Delta(G)$.
4. **Comprobar primero si es bipartito.** $\chi(G) = 2$ si y solo si G tiene alguna arista y no contiene ciclos de longitud impar. Se detecta pintando a dos colores por capas

desde un vértice.

5. **Si el grafo es plano**, el teorema de los cuatro colores da $\chi(G) \leq 4$ gratis, sea cual sea Δ.
6. **Cerrar el intervalo.** Si la cota inferior y la superior coinciden, ya está. Si no, hay que exhibir un coloreo con k colores (cota superior mejorada) o demostrar que $k-1$ es imposible.
7. **Verificar el coloreo** arista por arista: recorrer las m aristas y comprobar que ninguna une dos vértices del mismo color.

Ejemplo 8.56 (seis grafos, y un orden que arruina el algoritmo voraz)**.** (a) *Ciclos.* C_6 tiene $\Delta = 2$ y no contiene triángulos: $\omega = 2$, y el coloreo alterno funciona, luego $\chi(C_6) = 2$. En cambio C_5 es impar: al alternar colores alrededor del ciclo, el último vértice choca con el primero, y $\chi(C_5) = 3$. La regla general es $\chi(C_n) = 2$ si n es par y 3 si es impar, un ejemplo perfecto de la excepción de Brooks.

(b) *Grafos completos.* En K_4 todos los vértices se ven entre sí, de modo que $\omega = 4$ y $\chi = 4$; en general $\chi(K_n) = n$. Junto con los ciclos impares, los completos son la otra excepción del teorema de Brooks, y en ambos casos hace falta exactamente $\Delta + 1$.

(c) *Bipartitos.* $K_{3,3}$ tiene $\Delta = 3$, pero al ser bipartito basta con 2 colores, uno para cada lado. Aquí la cota $\Delta + 1 = 4$ se queda muy corta de puntería: no toda cota superior es informativa.

(d) *El grafo de Petersen.* Es 3-regular, conexo, no completo y no es un ciclo impar, así que Brooks da $\chi \leq 3$. Y como contiene ciclos de longitud 5 —su cuello es exactamente 5— no puede ser bipartito, luego $\chi \geq 3$. Las dos cotas se juntan: $\chi = 3$, sin necesidad de probar coloreos. Obsérvese que aquí $\omega = 2$, así que la cota por clique habría sido inútil; hizo falta el argumento del ciclo impar.

(e) *Un horario de exámenes.* Cinco asignaturas A, B, C, D, E; hay alumnos matriculados a la vez en AB, AC, BC, BD, CD y DE, y además en AE. Se construye el *grafo de conflictos* con una arista por cada par incompatible y se colorea: cada color es una franja horaria. Como A, B y C se ven mutuamente, $\omega = 3$ y hacen falta al menos tres franjas; el coloreo

$$A \to 1, \quad B \to 2, \quad C \to 3, \quad D \to 1, \quad E \to 2$$

usa tres y es válido —se comprueban las siete aristas una por una— luego $\chi = 3$: tres franjas, ni una más. Aquí $\Delta = 3$ y la cota voraz habría dicho 4, una franja de más y por tanto un día de exámenes desperdiciado.

(f) *El algoritmo voraz puede salir malísimo.* Tómense los vértices u_0, u_1, u_2, u_3 y v_0, v_1, v_2, v_3, con u_i unido a v_j siempre que $i \neq j$. El grafo es bipartito, así que $\chi = 2$. Pero si el voraz recorre los vértices en el orden $u_0, v_0, u_1, v_1, u_2, v_2, u_3, v_3$, el resultado es

$$u_0, v_0 \to \text{color } 0; \quad u_1, v_1 \to 1; \quad u_2, v_2 \to 2; \quad u_3, v_3 \to 3,$$

porque cada nuevo par ya tiene vecinos de todos los colores anteriores: **cuatro colores para un grafo que necesita dos**. El algoritmo voraz es correcto pero no óptimo, y la diferencia puede hacerse tan grande como se quiera aumentando el tamaño.

Cuidado

Un coloreo válido con k colores demuestra $\chi \leq k$, nunca $\chi = k$. Para la igualdad hace falta además una cota inferior, y esta casi siempre viene de exhibir un clique o un ciclo impar. Decir "lo he coloreado con cuatro y no se me ocurre hacerlo con tres, luego $\chi = 4$" no demuestra absolutamente nada.

Pausa para pensar

Ninguna de las dos cotas de la regla es cara de calcular, pero $\chi(G)$ sí lo es: decidir si un grafo se puede colorear con tres colores es NP-completo, mientras que decidir si se puede con dos —o sea, si es bipartito— se resuelve en tiempo lineal. El salto de dificultad ocurre entre 2 y 3, no entre 3 y 100. Es uno de los umbrales más bruscos de toda la complejidad computacional.

Idea clave

Colorear es repartir recursos escasos. Los vértices son tareas, las aristas dicen "estas dos no pueden coincidir" y los colores son el recurso: franjas horarias, aulas, frecuencias de radio para antenas vecinas, registros del procesador para variables vivas a la vez. El compilador que traduce cualquier programa resuelve un problema de coloración cada vez que asigna registros, y usa exactamente el algoritmo voraz de la regla, con una heurística de orden mejor que la del caso (f).

8.11. Caminos mínimos: el algoritmo de Dijkstra

Cuando las aristas llevan pesos —kilómetros, minutos, euros— la pregunta práctica por excelencia es: ¿cuál es la ruta más barata de A a F? Es lo que hace el navegador del teléfono cada vez que se le pide una dirección.

Regla: algoritmo de Dijkstra

1. Asignar distancia 0 al vértice de partida e ∞ a todos los demás. Ninguno está "cerrado".
2. Elegir el vértice abierto con la distancia provisional *menor* y cerrarlo: su distancia ya es definitiva.
3. Desde él, actualizar a sus vecinos abiertos: si llegar por aquí es más barato que la distancia que tenían anotada, se sustituye y se anota también de dónde se vino.
4. Repetir desde el paso 2 hasta cerrar el destino.

Ejemplo 8.57. Con los pesos $AB = 4$, $AC = 2$, $BC = 1$, $BD = 5$, $CD = 8$, $CE = 10$, $DE = 2$, $DF = 6$, $EF = 3$, búsquese la ruta más barata de A a F.

Paso	Se cierra	Distancia	Se actualiza
1	A	0	B: 4, C: 2
2	C	2	B: $2+1=3$ (mejora 4), D: 10, E: 12
3	B	3	D: $3+5=8$ (mejora 10)
4	D	8	E: $8+2=10$ (mejora 12), F: $8+6=14$
5	E	10	F: $10+3=13$ (mejora 14)
6	F	**13**	—

La ruta óptima cuesta 13 y, deshaciendo las anotaciones de procedencia, es

$$A \to C \to B \to D \to E \to F.$$

Obsérvese el detalle instructivo del paso 2: la mejor manera de ir de A a B *no* es la arista directa AB, que cuesta 4, sino el rodeo por C, que cuesta 3. Los caminos mínimos casi nunca son los evidentes.

Cuidado

Dijkstra exige que *todos* los pesos sean no negativos. Con pesos negativos el algoritmo falla, porque su idea entera se apoya en que cerrar el vértice más cercano es seguro: una arista de peso negativo podría abaratar después un camino ya dado por definitivo. Para ese caso hay otros algoritmos (Bellman–Ford), más lentos.

Pausa para pensar

¿Por qué es correcto Dijkstra? La clave es un argumento de mínimo: cuando se cierra un vértice u con distancia provisional d, cualquier otro camino hasta u tendría que salir en algún momento de la zona ya cerrada por una arista hacia un vértice abierto, y ese vértice abierto tiene distancia provisional $\geq d$ —porque u era el menor—. Como los pesos son no negativos, ese camino alternativo ya no puede bajar de d. Es la misma estructura lógica que la del principio del buen orden: se avanza tomando siempre el menor elemento disponible y se demuestra que esa avaricia local no arruina el resultado global. Que la avaricia funcione es la excepción, no la regla; en el problema del viajante, por ejemplo, no funciona en absoluto.

Regla: construir un árbol generador mínimo

Un árbol generador mínimo es la red más barata que mantiene todo conectado: conecta todos los vértices, no tiene ciclos, y entre todas las que cumplen eso es la de peso total más pequeño. Se construye con un algoritmo tan sencillo que parece trampa.

1. **Comprueba que el grafo es conexo.** Si no lo es, no hay árbol generador: lo que se obtiene es un bosque, uno por componente.

2. **Cuenta cuántas aristas necesitarás:** un árbol con n vértices tiene exactamente $n-1$ aristas, ni una más ni una menos. Es el criterio de parada del algoritmo.
3. **Kruskal: ordena todas las aristas de menor a mayor peso.**
4. **Recórrelas en ese orden y añade cada arista, salvo si cierra un ciclo** con las ya elegidas. Para saberlo, basta con mirar si sus dos extremos ya están conectados entre sí por lo elegido hasta ahora.
5. **Para cuando tengas $n-1$ aristas.** El resultado es un árbol generador de peso mínimo, garantizado.
6. **Alternativa (Prim), útil si prefieres crecer desde un punto:** empieza en un vértice cualquiera y añade repetidamente la arista más barata que salga del conjunto ya construido hacia fuera. Da el mismo peso total, aunque a veces otro árbol si hay empates.
7. **Control:** un árbol generador mínimo *no* es lo mismo que el árbol de caminos mínimos de Dijkstra. El primero minimiza el coste total de la red; el segundo minimiza la distancia desde un vértice fijo a cada uno de los demás. Casi nunca coinciden.

Ejemplo 8.58 (Un árbol mínimo y un árbol de Dijkstra sobre el mismo grafo)**.** Tomemos cinco ciudades A, B, C, D, E con estas ocho conexiones y sus costes:

arista	peso	arista	peso
AC	1	BD	5
BC	2	BE	7
DE	2	CD	8
AB	4	CE	10

Kruskal, paso a paso. Ordenadas: $AC(1)$, $BC(2)$, $DE(2)$, $AB(4)$, $BD(5)$, $BE(7)$, $CD(8)$, $CE(10)$.

- $AC(1)$: se añade. Grupos: $\{A, C\}$, $\{B\}$, $\{D\}$, $\{E\}$.
- $BC(2)$: se añade. Grupos: $\{A, B, C\}$, $\{D\}$, $\{E\}$.
- $DE(2)$: se añade. Grupos: $\{A, B, C\}$, $\{D, E\}$.
- $AB(4)$: *se descarta*, porque A y B ya están en el mismo grupo y cerraría el ciclo A–C–B–A.
- $BD(5)$: se añade. Ya hay un solo grupo y $4 = 5 - 1$ aristas: se para.

El árbol generador mínimo es $\{AC, BC, DE, BD\}$ con peso total

$$1 + 2 + 2 + 5 = 10.$$

Dijkstra desde A, sobre el mismo grafo. Se obtienen las distancias mínimas

$$d(A) = 0, \quad d(C) = 1, \quad d(B) = 3, \quad d(D) = 8, \quad d(E) = 10,$$

alcanzadas por los caminos $A-C$ (1), $A-C-B$ ($1+2=3$), $A-C-B-D$ ($3+5=8$) y $A-C-B-E$ ($3+7=10$), este último empatado con $A-C-B-D-E$ ($8+2=10$).

La comparación. El árbol de Dijkstra usa las aristas AC, CB, BD y BE, con peso total $1+2+5+7=15$, mientras que el árbol mínimo pesa 10. Son árboles distintos: el de Kruskal prefiere la arista barata $DE(2)$, y el de Dijkstra la evita porque llegar a D ya cuesta 8 y no ayuda a acortar el camino hasta E.

Idea clave

La lección del paso 7 es la del ejemplo: *barato para la red* y *corto para el viajero* son objetivos distintos. Una compañía que tiende cable elige Kruskal, porque paga por metro instalado; un conductor que va de A a E elige Dijkstra, porque paga por metro recorrido. El mismo mapa, dos respuestas correctas a dos preguntas distintas.

Cuidado

El error clásico en Kruskal es descartar una arista "porque toca un vértice ya usado". La condición correcta es que cierre un ciclo, es decir, que *ambos* extremos estén ya conectados entre sí. En el ejemplo, $BD(5)$ toca a B, que ya estaba usado, y aun así se añade: D estaba en otro grupo.

Pausa para pensar

¿Por qué funciona un algoritmo tan ingenuo? Por la llamada propiedad del corte: si partes los vértices en dos bandos, la arista más barata que los cruza pertenece a *algún* árbol mínimo. Kruskal no hace más que aplicar esa propiedad una y otra vez, y por eso la elección voraz —siempre la más barata disponible— nunca se arrepiente. Con Dijkstra ocurre algo parecido, pero solo si todos los pesos son no negativos; con pesos negativos el algoritmo falla y hay que usar otro.

8.12. Flujo máximo y corte mínimo

Un grafo dirigido cuyas aristas llevan una **capacidad** —litros por segundo en una tubería, coches por hora en una carretera, bits por segundo en un cable— se llama una **red**. La pregunta es cuánto puede mandarse desde un vértice de origen s hasta un vértice de destino t sin que ningún tramo rebase su capacidad. La respuesta tiene una forma tan limpia que es uno de los teoremas más citados de toda la matemática aplicada.

Definición 8.59. Un **flujo** asigna a cada arista uv un valor $f(uv)$ con $0 \le f(uv) \le c(uv)$, de modo que en todo vértice distinto de s y t lo que entra es igual a lo que sale. Su **valor** es el total que sale de s. Un **corte** es un reparto de los vértices en dos bandos S (con $s \in S$) y T (con $t \in T$); su **capacidad** es la suma de las capacidades de las aristas que van de S a T, contando solo esas y no las de vuelta.

Teorema 8.60 (Flujo máximo – corte mínimo)**.** *En toda red, el valor del flujo máximo*

coincide con la capacidad del corte mínimo.

Idea de la demostración. Que todo flujo es menor o igual que la capacidad de cualquier corte es evidente: todo lo que llega a t tuvo que cruzar el corte en algún momento. La igualdad es lo profundo, y sale del propio algoritmo: cuando ya no queda ningún camino por el que enviar más, el conjunto S de vértices alcanzables desde s en la red residual define un corte cuyas aristas de salida están todas saturadas y cuyas aristas de entrada están todas vacías. Su capacidad es exactamente el flujo alcanzado. □

Regla: calcular el flujo máximo y exhibir el corte que lo limita

1. **Dibuja la red y marca capacidades.** Escribe en cada arista f/c, empezando con $f = 0$ en todas. Comprueba que hay un único origen y un único destino; si hubiera varios, añade un origen ficticio conectado a todos con capacidad infinita, y lo mismo con el destino.
2. **Busca un camino de s a t con holgura** en el que toda arista tenga $f < c$. Vale cualquiera; no hace falta que sea el mejor.
3. **Envía por él el cuello de botella.** La cantidad que puedes empujar es el mínimo de las holguras $c - f$ del camino. Súmala a todas sus aristas.
4. **No olvides las aristas de retroceso.** En la red residual, cada arista con $f > 0$ permite *deshacer* flujo yendo hacia atrás con holgura f. Los caminos aumentantes pueden y deben usar esas marchas atrás; sin ellas el algoritmo se queda corto.
5. **Repite hasta que no quede ningún camino con holgura.** Ese es el criterio de parada de Ford–Fulkerson.
6. **Extrae el corte.** Sea S el conjunto de vértices a los que aún se puede llegar desde s en la red residual, y T el resto. Ese corte tiene capacidad igual al flujo obtenido, y es la prueba de que no se puede hacer mejor.
7. **Contrasta los dos números.** Flujo hallado y capacidad del corte deben coincidir hasta la última unidad. Si no coinciden, o has dejado un camino sin explorar, o has contado en el corte alguna arista que va de T a S, que no suma.

Ejemplo 8.61 (Una red pequeña resuelta hasta el final)**.** Considérese la red con vértices s, a, b, c, d, t y capacidades

$$sa\colon 10,\quad sb\colon 10,\quad ab\colon 2,\quad ac\colon 4,\quad ad\colon 8,\quad bd\colon 9,\quad dc\colon 6,\quad ct\colon 10,\quad dt\colon 10.$$

(a) **Cota superior de un vistazo.** Todo lo que sale de s pasa por sa y sb: como mucho $10 + 10 = 20$. Y todo lo que entra en t pasa por ct y dt: como mucho $10 + 10 = 20$. Sabemos ya que el flujo máximo está entre 0 y 20; toca estrechar el intervalo por los dos lados.

(b) **Primer camino: $s \to a \to c \to t$.** Las holguras son 10, 4 y 10; el cuello de botella es la arista ac con 4. Flujo acumulado: 4.

(c) **Segundo camino:** $s \to a \to d \to t$**.** Holguras $10-4=6$, 8 y 10; el cuello de botella es ahora sa con 6. Flujo acumulado: $4+6=10$, y la arista sa queda saturada a $10/10$.

(d) **Tercer camino:** $s \to b \to d \to t$**.** Holguras 10, 9 y $10-6=4$; pasa 4 y la arista dt queda saturada a $10/10$. Flujo acumulado: 14.

(e) **Cuarto camino, el que casi nadie ve:** $s \to b \to d \to c \to t$**.** Aunque dt está lleno, todavía se puede desviar por dc: las holguras son $10-4=6$, $9-4=5$, 6 y $10-4=6$, con cuello de botella 5 en bd. Pasa 5 y el flujo acumulado sube a $14+5=19$.

(f) **El estado final.** Las aristas quedan

$$sa = 10/10, \quad sb = 9/10, \quad ab = 0/2, \quad ac = 4/4, \quad ad = 6/8,$$

$$bd = 9/9, \quad dc = 5/6, \quad ct = 9/10, \quad dt = 10/10.$$

Compruébese la conservación vértice a vértice: en a entran 10 y salen $4+6=10$ ✓; en b entran 9 y salen 9 ✓; en c entran $4+5=9$ y salen 9 ✓; en d entran $6+9=15$ y salen $5+10=15$ ✓. Y a t llegan $9+10=19$, que es el valor del flujo.

(g) **El corte que cierra el caso.** ¿Adónde se llega desde s en la red residual? La arista sa está saturada, así que a queda fuera; por sb aún cabe 1, luego b entra; pero desde b la única salida, bd, está saturada a $9/9$, y la marcha atrás $b \to s$ no aporta nada nuevo. Por tanto $S=\{s,b\}$ y $T=\{a,c,d,t\}$. Las aristas que van de S a T son sa y bd, con capacidad

$$10+9=19,$$

exactamente el flujo obtenido. La arista ab va de T a S y no cuenta. El problema queda cerrado: 19 es el máximo, y ninguna mejora en las demás tuberías lo cambiará mientras sa y bd sigan como están.

Cuidado

Los dos errores clásicos de esta sección son gemelos. El primero es olvidar las aristas de retroceso: sin ellas, un reparto inicial desafortunado puede dejar el algoritmo atascado por debajo del máximo, y uno concluye erróneamente que ha terminado. El segundo es contar en el corte las aristas que van de T a S: la capacidad de un corte suma *solo* las que lo cruzan en el sentido $S \to T$. En el ejemplo, incluir ab daría 21 y rompería la igualdad con el flujo.

Pausa para pensar

Fíjate en el apartado (e). Después de tres caminos el flujo era 14, y las dos aristas que entran en t parecían ofrecer todavía mucho margen (ct tenía 6 libres). Lo que faltaba no era capacidad al final, sino un *camino* que la alcanzara. Esa es la moraleja de todo el capítulo de redes: la limitación nunca está en una arista aislada, sino en un conjunto de aristas que, juntas, separan el origen del destino. El teorema del corte mínimo dice exactamente eso, y por eso su enunciado es una igualdad y no una desigualdad: siempre hay un cuello de botella *global* que explica el resultado.

Idea clave

El teorema del flujo máximo y el corte mínimo es el ejemplo por excelencia de *dualidad*: un problema de maximizar (el flujo) y otro de minimizar (el corte) tienen la misma respuesta, y la solución de cada uno certifica la del otro. Es la misma estructura de la dualidad en programación lineal del la programación lineal, y no es casualidad: el flujo máximo puede plantearse como un programa lineal cuyo dual es precisamente el corte mínimo. Y todavía más: el teorema de König de la sección anterior es un caso particular —emparejamiento máximo y cubrimiento mínimo— que se obtiene convirtiendo el grafo bipartito en una red con capacidades 1. Tres teoremas que parecían independientes son el mismo teorema visto desde tres sitios.

8.13. Grafos dirigidos: orden topológico y camino crítico

Cuando las aristas tienen sentido único, el grafo se llama **dirigido**. El caso más útil en la práctica es aquel en que además no hay ciclos: un **DAG**, o grafo dirigido acíclico. Todo proyecto con tareas que dependen unas de otras, toda hoja de cálculo con fórmulas encadenadas, todo sistema de paquetes de software con dependencias y todo plan de estudios con asignaturas de las que otras son prerrequisito es un DAG, y las dos preguntas que se le hacen son siempre las mismas: ¿en qué orden puedo hacer las cosas?, y ¿cuánto tardaré como mínimo?

Definición 8.62. Un **orden topológico** de un grafo dirigido es una lista de todos sus vértices tal que toda arista va de un vértice anterior a uno posterior de la lista.

Ejemplo 8.63 (Ordenar tareas con dependencias)**. (a) Un ejemplo pequeño.** Con las dependencias $a \to b$, $a \to c$, $b \to d$, $c \to d$, los órdenes topológicos válidos son a, b, c, d y a, c, b, d: dos. El primero y el último están forzados; los del medio son intercambiables porque no hay arista entre ellos.

(b) El algoritmo de Kahn. Se toma repetidamente un vértice sin aristas entrantes, se emite y se borra junto con sus aristas salientes. Si en algún momento no queda ninguno y todavía hay vértices, hay un ciclo. El coste es lineal.

(c) Cuántos órdenes hay. Si no hay ninguna arista, todos los $n!$ ordenamientos valen. Si el grafo es una cadena $v_1 \to v_2 \to \cdots \to v_n$, hay exactamente uno. Contar los órdenes topológicos de un grafo cualquiera es, curiosamente, un problema difícil: es #P-completo.

(d) Para qué se usa. Compilar módulos que se importan unos a otros, instalar paquetes con dependencias, planificar tareas de un proyecto, evaluar una hoja de cálculo con fórmulas encadenadas y decidir el orden de ejecución de las capas de una red neuronal. En todos los casos, el fallo del algoritmo es el aviso de una dependencia circular.

Teorema 8.64. *Un grafo dirigido admite un orden topológico si y solo si no tiene ciclos dirigidos.*

Demostración. Si hubiera un ciclo dirigido, cada vértice del ciclo tendría que ir antes que

el siguiente y, dando la vuelta, antes que sí mismo: imposible. Recíprocamente, si no hay ciclos, siempre existe algún vértice sin aristas entrantes —si todos tuvieran alguna, se podría caminar hacia atrás indefinidamente y, al ser el grafo finito, se repetiría un vértice, formando un ciclo—. Colóquese ese vértice el primero, quítese del grafo, y repítase el argumento con lo que queda. □

Ejemplo 8.65 (Ciclos y su detección)**. (a) La implicación fácil.** Si hay un ciclo $v_1 \to v_2 \to \cdots \to v_k \to v_1$, cualquier orden debería poner v_1 antes que v_2, éste antes que v_3, y así hasta concluir que v_1 va antes que sí mismo. Imposible.

(b) La implicación útil. Si no hay ciclos, siempre existe un vértice sin aristas entrantes: si todos tuvieran alguna, se podría retroceder indefinidamente y, al haber un número finito de vértices, se repetiría uno, formando un ciclo. Ese vértice se emite y se repite el argumento.

(c) Un caso con ciclo. Con $a \to b$, $b \to c$, $c \to a$, el algoritmo de Kahn se atasca en el primer paso: los tres vértices tienen grado de entrada 1 y no hay por dónde empezar. La detección es inmediata y además señala dónde está el problema.

(d) Un aviso sobre grafos no dirigidos. El enunciado exige ciclos *dirigidos*. El grafo con $a \to b$, $a \to c$, $b \to d$, $c \to d$ contiene un cuadrilátero si se ignoran las flechas, y sin embargo admite orden topológico sin ningún problema, porque recorriéndolo con las direcciones puestas nunca se vuelve al punto de partida. Sólo la orientación importa: la palabra "acíclico" significa cosas distintas según haya flechas o no.

Regla: ordenar tareas y encontrar el camino crítico de un proyecto

1. **Modela.** Un vértice por tarea, con su duración escrita al lado; una arista $u \to v$ por cada relación "v no puede empezar hasta que u termine". Comprueba que no hay ciclos: un ciclo de dependencias significa que el proyecto, tal como está descrito, es irrealizable.

2. **Ordena topológicamente.** Ve retirando repetidamente una tarea sin predecesores pendientes y añádela a la lista. Si en algún momento no queda ninguna y todavía hay tareas, hay un ciclo. Suele haber muchos órdenes válidos; cualquiera sirve para lo que viene.

3. **Pasada hacia adelante: las fechas más tempranas.** Recorriendo la lista en orden, calcula el inicio temprano $\mathrm{ES}(v) = \max\{\mathrm{EF}(u) : u \to v\}$ (y 0 si v no tiene predecesores) y el fin temprano $\mathrm{EF}(v) = \mathrm{ES}(v) + d(v)$. Ojo: es un *máximo*, porque una tarea espera a la última de sus dependencias.

4. **Lee la duración total.** Es el mayor de todos los EF. Ese número es la duración mínima del proyecto por muchos recursos que se tengan: no depende de cuánta gente haya, sino de la cadena de dependencias.

5. **Pasada hacia atrás: las fechas más tardías.** Recorriendo la lista al revés, calcula el fin tardío $\mathrm{LF}(v) = \min\{\mathrm{LS}(w) : v \to w\}$ (y la duración total T si v no tiene sucesores) y el inicio tardío $\mathrm{LS}(v) = \mathrm{LF}(v) - d(v)$. Aquí es un *mínimo*: retrasarse más estropearía a la más urgente de las tareas que dependen de ella.

6. **Calcula la holgura** $\mathrm{LS}(v) - \mathrm{ES}(v)$ de cada tarea: cuántas unidades de tiempo puede retrasarse sin alargar el proyecto.
7. **Identifica el camino crítico:** las tareas de holgura cero. Forman siempre al menos un camino completo del principio al final, y son las únicas en las que invertir esfuerzo acorta el proyecto. Acelerar una tarea con holgura no sirve absolutamente de nada.

Ejemplo 8.66 (Un proyecto de ocho tareas, medido hasta la última holgura)**.** Ocho tareas $A, \ldots, H$ con duraciones (en días) y dependencias:

Tarea	Duración	Depende de
A	3	—
B	2	A
C	4	A
D	5	B
E	1	C
F	2	D, E
G	3	C
H	2	F, G

(a) **Es un DAG.** Toda dependencia va de una letra anterior a una posterior del alfabeto, así que no puede haber ciclos: ya tenemos gratis un orden topológico, A, B, C, D, E, F, G, H. No es el único: A, C, G, B, E, D, F, H también sirve, y en total hay 26 órdenes válidos. Que haya varios refleja algo real: hay tareas que pueden hacerse en cualquier orden entre sí, o incluso a la vez.

(b) **Pasada hacia adelante.** $\mathrm{ES}(A) = 0$ y $\mathrm{EF}(A) = 3$. De ahí $\mathrm{ES}(B) = \mathrm{ES}(C) = 3$, con $\mathrm{EF}(B) = 5$ y $\mathrm{EF}(C) = 7$. Luego $\mathrm{ES}(D) = 5$, $\mathrm{EF}(D) = 10$; $\mathrm{ES}(E) = 7$, $\mathrm{EF}(E) = 8$; $\mathrm{ES}(G) = 7$, $\mathrm{EF}(G) = 10$. La tarea F depende de D y E: $\mathrm{ES}(F) = \max(10, 8) = 10$ y $\mathrm{EF}(F) = 12$. Y H depende de F y G: $\mathrm{ES}(H) = \max(12, 10) = 12$, $\mathrm{EF}(H) = 14$.

(c) **La duración total es** 14 **días.** Nótese que la suma de todas las duraciones es $3+2+4+5+1+2+3+2 = 22$: el proyecto se hace en 14 y no en 22 porque varias tareas corren en paralelo. Pero tampoco baja de 14 por mucho personal que se contrate, porque A, B, D, F y H tienen que ir una tras otra y suman exactamente $3+2+5+2+2 = 14$.

(d) **Pasada hacia atrás.** $\mathrm{LF}(H) = 14$ y $\mathrm{LS}(H) = 12$. Como F y G desembocan en H, ambas tienen $\mathrm{LF} = 12$, de donde $\mathrm{LS}(F) = 10$ y $\mathrm{LS}(G) = 9$. Siguiendo: $\mathrm{LF}(D) = \mathrm{LF}(E) = 10$, luego $\mathrm{LS}(D) = 5$ y $\mathrm{LS}(E) = 9$. Para C, que alimenta a E y a G, $\mathrm{LF}(C) = \min(9, 9) = 9$ y $\mathrm{LS}(C) = 5$. Para B, $\mathrm{LF}(B) = 5$ y $\mathrm{LS}(B) = 3$. Y para A, $\mathrm{LF}(A) = \min(3, 5) = 3$ con $\mathrm{LS}(A) = 0$.

(e) **La tabla completa.**

Tarea	ES	EF	LS	LF	Holgura
A	0	3	0	3	**0**
B	3	5	3	5	**0**
C	3	7	5	9	2
D	5	10	5	10	**0**
E	7	8	9	10	2
F	10	12	10	12	**0**
G	7	10	9	12	2
H	12	14	12	14	**0**

(f) **El camino crítico.** Las tareas de holgura cero son A, B, D, F, H, y forman en efecto un camino completo $A \to B \to D \to F \to H$ de duración 14. Las tres restantes —C, E, G— tienen dos días de margen cada una: pueden retrasarse hasta dos días sin que el proyecto se alargue ni un minuto.

(g) **Dónde invertir el dinero.** Supóngase que se puede acortar una tarea en un día pagando lo mismo en cualquiera de ellas. Acortar C, E o G no cambia nada: el proyecto sigue durando 14, y lo único que se consigue es aumentar su holgura a tres días. Acortar en cambio cualquiera de las críticas —A, B, D, F o H— lleva el total a 13 días. Pero cuidado con insistir: si se acorta D en dos días, el camino $A \to C \to G \to H$, que sumaba $3 + 4 + 3 + 2 = 12$, pasa a ser competitivo, y a partir de cierto punto el camino crítico se muda a otro sitio. Por eso la tabla hay que recalcularla después de cada cambio.

Cuidado

El error más común es confundir los dos operadores de las pasadas. Hacia adelante se toma el *máximo* de los fines tempranos de los predecesores, porque una tarea espera a que terminen todas; hacia atrás se toma el *mínimo* de los inicios tardíos de los sucesores, porque no puede hacer esperar a ninguna. Intercambiarlos da holguras negativas o caminos críticos rotos, y es fácil detectarlo: en un cálculo correcto, las tareas de holgura cero forman siempre un camino continuo de principio a fin.

Pausa para pensar

Compárese este algoritmo con el de Dijkstra de la sección anterior. Los dos recorren un grafo con pesos actualizando distancias, pero aquí se busca el camino *más largo* y allí el más corto. ¿Por qué el camino más largo, que en un grafo general es un problema tan difícil como el del viajante, se resuelve aquí en una sola pasada? Porque el grafo es acíclico: el orden topológico garantiza que cuando se calcula $\mathrm{ES}(v)$ ya están calculados todos sus predecesores, sin necesidad de volver atrás nunca. Con ciclos, el camino más largo ni siquiera está bien definido —podría darse vueltas indefinidamente—, y ahí es donde la dificultad reaparece de golpe.

Idea clave

El orden topológico es la versión finita y computable de la idea de *orden parcial*: una relación en la que algunos pares de elementos son comparables y otros no. En este proyecto, B va antes que D, pero B y G no guardan ninguna relación de precedencia: son incomparables. Lo que hace el algoritmo es *extender* ese orden parcial a un orden total, y el hecho de que haya 26 maneras de hacerlo mide exactamente cuánta libertad deja la estructura. Esa misma idea —extender un orden parcial a uno total— reaparece en lógica y en teoría de conjuntos, y es también la razón por la que un DAG puede dibujarse siempre en capas de izquierda a derecha, algo que ningún grafo con ciclos permite.

8.14. Redes reales: centralidad, PageRank y el mundo pequeño

Los grafos dejaron de ser un juego de salón cuando se descubrió que las redes sociales, la web, las redes eléctricas y las de contagio se describen todas con el mismo lenguaje. La pregunta típica ya no es "¿existe un camino?" sino "¿quién es importante aquí?", y hay varias respuestas distintas y todas razonables.

Definición 8.67. En un grafo conexo con n vértices, la *centralidad de grado* de v es $\dfrac{\deg(v)}{n-1}$; la *centralidad de cercanía* es $\dfrac{n-1}{\sum_u d(v,u)}$; y la *centralidad de intermediación* es la proporción de caminos mínimos entre pares de vértices que pasan por v. El *coeficiente de agrupamiento* mide qué fracción de los caminos de longitud dos se cierran en triángulo. En un grafo dirigido, el *PageRank* es el vector $\vec{p}$ que cumple $\vec{p} = d\,M\vec{p} + \frac{1-d}{n}\vec{1}$, donde M reparte por igual la importancia de cada página entre sus enlaces salientes y $d \approx 0{,}85$ es el factor de amortiguación.

Idea clave

Hay al menos tres maneras distintas de ser importante en una red, y conviene no confundirlas. Se puede tener mucha gente alrededor (grado): es el popular. Se puede estar cerca de todos en pocos saltos (cercanía): es el bien situado. Y se puede ser el paso obligado entre dos regiones que sin ti quedarían lejos (intermediación): es el puente. Un vértice puede tener grado bajísimo y una intermediación enorme si es el único enlace entre dos comunidades, y ese es exactamente el perfil del intermediario en sociología, del router crítico en una red o del aeropuerto de conexión. El PageRank añade una cuarta idea, recursiva: eres importante si te enlazan páginas importantes, lo que obliga a resolver un problema de punto fijo en lugar de contar.

Regla: medir la importancia de un vértice

Para analizar una red, este es el orden que aporta más información con menos cálculo:

1. **Calcula la distribución de grados**, que es lo más barato y ya distingue una red homogénea de una con concentradores.

2. **Ejecuta un BFS desde cada vértice** para obtener de una vez todas las distancias; de ahí salen la cercanía, el diámetro y el camino medio.
3. **Calcula el camino medio y el diámetro**: el primero dice cómo de conectada está la red en general, el segundo cuál es el peor caso.
4. **Cuenta triángulos** para obtener el coeficiente de agrupamiento; en redes sociales suele ser alto y en redes aleatorias, bajísimo.
5. **Estima la intermediación** acumulando, para cada par, los caminos mínimos que pasan por cada vértice intermedio.
6. **Para grafos dirigidos, itera el PageRank** partiendo del vector uniforme; converge deprisa porque el factor de contracción es d.
7. **Compara los rankings**: si las distintas centralidades ordenan igual, la red es homogénea; si difieren mucho, cada una está detectando un papel distinto.

Ejemplo 8.68 (Una red social de siete personas)**. (a) La red.** Las amistades son Ana–Beto, Ana–Carla, Beto–Carla, Carla–Dora, Dora–Elsa, Dora–Fito, Elsa–Fito y Fito–Gil: siete personas y ocho relaciones. Los grados son $2, 2, 3, 3, 2, 3, 1$, y suman $16 = 2 \cdot 8$. Hay dos triángulos, Ana–Beto–Carla y Dora–Elsa–Fito.

(b) Las tres centralidades. Por grado empatan Carla, Dora y Fito con $\frac{3}{6} = 0{,}5$, y Gil queda último con 0,166667. Por cercanía gana Dora con 0,666667, seguida de Carla con 0,6 y Fito con 0,545455: Dora está mejor situada aunque tenga el mismo grado. Y por intermediación vuelve a ganar Dora, con 0,6, seguida de Carla con 0,533333 y Fito con 0,333333; Ana, Beto, Elsa y Gil tienen intermediación cero.

(c) Qué significa cada resultado. Ana y Beto tienen grado 2 e intermediación nula porque están dentro de un triángulo: todo lo que pasa por ellos tiene una ruta alternativa. Gil, con grado 1, es una hoja. Y Carla y Dora forman el puente entre las dos comunidades: si desaparece Dora, la red se parte. Esa es exactamente la información que la centralidad de intermediación está diseñada para detectar, y que el grado no ve.

(d) Las medidas globales. El diámetro es 4 (de Ana o Beto hasta Gil) y el camino medio 2,047619. El coeficiente de agrupamiento global vale 0,5 y el medio 0,571429: alto, como es típico en redes de amistad, donde los amigos de mis amigos suelen ser amigos entre sí.

Ejemplo 8.69 (PageRank y el fenómeno del mundo pequeño)**. (a) Cuatro páginas.** Sea la web con enlaces $1 \to 2$, $1 \to 3$, $2 \to 3$, $3 \to 1$ y $4 \to 3$. Con $d = 0{,}85$, iterando desde el vector uniforme se llega a $p_1 = 0{,}372526$, $p_2 = 0{,}195824$, $p_3 = 0{,}394150$ y $p_4 = 0{,}037500$, que suman 1. La página 3 gana porque la enlazan tres páginas; la 1 queda casi igual de alta porque, aunque solo la enlaza una página, esa página es la 3, que es la mejor de todas.

(b) Por qué la página 4 queda en el fondo. Nadie la enlaza, así que solo recibe el término $\frac{1-d}{n} = \frac{0{,}15}{4} = 0{,}0375$, exactamente su valor final. Ese término es la razón de ser del factor de amortiguación: sin él, las páginas sin enlaces entrantes valdrían cero y el sistema podría no tener solución única. La interpretación clásica es la del *navegante aleatorio*: con probabilidad d sigue un enlace y con probabilidad $1 - d$ salta a una página cualquiera.

(c) El experimento del mundo pequeño. Tomemos un anillo de 1000 vértices en el que cada uno se une a sus seis vecinos más próximos. Su camino medio es 83,7508 y su coeficiente de agrupamiento 0,6: muy agrupado y muy lento de recorrer. Si ahora reconectamos al azar solo el 5 % de las aristas, el camino medio se desploma a 7,8163 mientras que el agrupamiento apenas baja a 0,517567.

(d) La conclusión. Con unos pocos atajos aleatorios se obtiene lo mejor de los dos mundos: la estructura local de comunidades del anillo y la accesibilidad global de una red aleatoria, cuyo camino medio sería 4,0978 pero con un agrupamiento ridículo de 0,004994. Esa combinación es la firma de casi todas las redes reales, y explica por qué en una red de mil millones de personas dos cualesquiera están a media docena de saltos: el orden de magnitud es $\log n$, no n.

Cuidado

La centralidad de cercanía tal como se ha definido exige que el grafo sea conexo; si no lo es, alguna distancia es infinita y la fórmula se rompe, de modo que hay que calcularla dentro de cada componente. Segundo aviso: un grado alto no implica una intermediación alta ni viceversa, así que hablar de "el vértice más importante" sin decir según qué medida no significa nada. Y tercero: el PageRank depende del factor d; con d próximo a 1 el cálculo converge muy despacio y se vuelve sensible a los vértices sin salida, razón por la que el valor 0,85 se ha convertido en estándar de hecho.

Pausa para pensar

El experimento del mundo pequeño lo hizo Stanley Milgram en 1967: pidió a unos voluntarios de Nebraska que hicieran llegar una carta a un desconocido de Boston pasándola solo a conocidos directos. De las cartas que llegaron, la mediana fue de unos cinco intermediarios, de donde salió la expresión "seis grados de separación". El modelo matemático que lo explica es de Watts y Strogatz, publicado en *Nature* en 1998, y su gracia está en lo poco que hace falta: basta un puñado de atajos aleatorios. El PageRank, por su parte, es de Brin y Page en 1998 y no es más que el método de la potencia aplicado a una matriz gigantesca; el mismo algoritmo elemental que se estudia con matrices 2×2 movió, durante años, la empresa más valiosa del mundo.

8.15. Conjuntos independientes, cliques y recubrimientos

Los tres conceptos de este apartado son los mismos vistos desde tres ángulos, y juntos forman el núcleo de los problemas de grafos que nadie sabe resolver deprisa. Conviene entender bien las relaciones entre ellos, porque permiten traducir un problema en otro y porque su dificultad es la puerta de entrada a la teoría de la complejidad.

Definición 8.70. Un *conjunto independiente* de un grafo G es un subconjunto de vértices sin ninguna arista entre ellos; su tamaño máximo se denota $\alpha(G)$ y se llama *número de independencia*. Una *clique* es un subconjunto de vértices todos adyacentes entre sí, con tamaño

máximo $\omega(G)$. Un *recubrimiento por vértices* es un conjunto de vértices que toca a todas las aristas, con tamaño mínimo $\tau(G)$. Se cumplen siempre

$$\alpha(G) + \tau(G) = n, \qquad \omega(G) = \alpha(\overline{G}),$$

donde $\overline{G}$ es el grafo *complementario*. Además $\omega(G) \leq \chi(G)$, siendo χ el número cromático.

Idea clave

La primera identidad se ve de un vistazo: un conjunto S es independiente si y solo si su complementario toca todas las aristas, porque una arista sin ningún extremo fuera de S sería una arista dentro de S. Por tanto encontrar el mayor independiente y el menor recubrimiento son literalmente el mismo problema con la respuesta escrita al revés. La segunda identidad es todavía más simple: una clique en G es un conjunto de vértices sin aristas en $\overline{G}$. Así que los tres problemas son uno solo con tres disfraces, y por eso todo lo que se descubre sobre cualquiera de ellos se traduce inmediatamente a los otros dos, incluida la mala noticia: los tres son NP-completos.

Regla: acotar y calcular estos tres números

No hay algoritmo rápido conocido, pero sí un procedimiento razonable para grafos pequeños y buenas cotas para los grandes:

1. **Empieza por la cota inferior** $\alpha(G) \geq \dfrac{n}{1+\bar{d}}$, donde $\bar{d}$ es el grado medio; es gratis y suele ser sorprendentemente buena.
2. **Aplica el algoritmo voraz**: elige repetidamente el vértice de menor grado, añádelo al conjunto y elimina a él y a sus vecinos.
3. **Usa la coloración como cota superior**: en cada clase de color los vértices son independientes, luego $\alpha(G) \geq \frac{n}{\chi(G)}$, y también $\omega(G) \leq \chi(G)$.
4. **Traduce al complementario** si el grafo original es denso: buscar una clique grande en un grafo denso equivale a buscar un independiente en uno disperso, que suele ser más fácil de manejar a ojo.
5. **Recuerda que** $\tau = n - \alpha$, de modo que calcular uno resuelve el otro sin trabajo adicional.
6. **En grafos bipartitos, aplica el teorema de König**: allí el recubrimiento mínimo coincide con el emparejamiento máximo, y eso sí se calcula deprisa.
7. **Verifica siempre el resultado** comprobando que no hay ninguna arista dentro del conjunto propuesto; es la única parte del problema que resulta rápida.

Ejemplo 8.71 (Los tres números en grafos conocidos)**. (a) El ciclo** C_5**.** En un pentágono, dos vértices no consecutivos forman un independiente y no se puede hacer mejor: $\alpha = 2$. Por tanto $\tau = 5 - 2 = 3$, y en efecto tres vértices bien elegidos tocan las cinco aristas. La clique máxima es una arista, $\omega = 2$, mientras que $\chi(C_5) = 3$: es el ejemplo más pequeño de grafo

donde $\omega < \chi$, y por eso los ciclos impares aparecen en todos los contraejemplos de la teoría de coloración.

(b) El grafo completo y el bipartito completo. En K_6 hay 15 aristas, $\alpha = 1$, $\tau = 5$ y $\omega = 6$: todo es adyacente a todo. En $K_{3,4}$, con 7 vértices y 12 aristas, el mayor independiente es el lado grande, $\alpha = 4$, luego $\tau = 3$; y como no hay triángulos, $\omega = 2$. El teorema de König se comprueba aquí: el emparejamiento máximo tiene tamaño 3, igual que el recubrimiento mínimo.

(c) El grafo de Petersen. Tiene 10 vértices, 15 aristas y todos los grados iguales a 3. Su número de independencia es $\alpha = 4$, luego $\tau = 6$. No contiene triángulos (su cintura es 5), así que $\omega = 2$, mientras que $\chi = 3$: otra vez la brecha entre clique y coloración, esta vez en un grafo sin ningún triángulo. La cota inferior $\frac{10}{1+3} = 2{,}5$ da $\alpha \geq 3$, correcta pero floja.

(d) Un grafo pequeño resuelto a mano. Sea G con vértices $1, \ldots, 5$ y aristas 12, 23, 34, 45, 51, 13. El conjunto $\{1, 4\}$ es independiente y no hay ninguno de tamaño tres (compruébese caso por caso), luego $\alpha = 2$ y $\tau = 3$. La clique máxima es $\{1, 2, 3\}$, de tamaño 3, y el número cromático es también 3. La cota $\frac{5}{1+2{,}4} = 1{,}470588$ solo garantiza $\alpha \geq 2$, que aquí resulta ser exacto.

Ejemplo 8.72 (Los mismos problemas, disfrazados de la vida real)**. (a) Repartir tareas incompatibles.** Si cada vértice es una tarea y cada arista une dos tareas que no pueden hacerse a la vez, el mayor conjunto de tareas simultáneas es exactamente un conjunto independiente máximo. Con la coloración se obtiene además una planificación completa: cada color es un turno, y el número cromático es el mínimo de turnos necesarios.

(b) Vigilar una red. Si cada arista es un tramo de tubería que hay que monitorizar y los sensores se colocan en los nodos, cubrir todos los tramos con el menor número de sensores es el problema del recubrimiento por vértices. La solución voraz de "coge siempre el vértice de mayor grado" no es óptima, pero el algoritmo que toma ambos extremos de una arista no cubierta, repetidamente, garantiza a lo sumo el doble del óptimo: es el ejemplo más limpio de algoritmo de aproximación con factor demostrado.

(c) Buscar grupos en una red social. Un grupo de personas todas amigas entre sí es una clique. Encontrar la mayor es NP-completo, y por eso los programas de análisis de redes buscan estructuras más laxas (comunidades, cliques parciales) que sí se pueden calcular. Nótese que en el complementario, "todos amigos" se convierte en "ninguno amigo": el mismo cálculo con los papeles cambiados.

(d) Por qué esto importa tanto. La clique fue uno de los veintiún problemas que Karp demostró NP-completos en 1972, y arrastró consigo al independiente y al recubrimiento por las equivalencias de arriba. Que sean equivalentes significa que un algoritmo rápido para cualquiera de ellos los resolvería todos, y de hecho resolvería los miles de problemas NP-completos conocidos, zanjando la pregunta de si P = NP. Nadie espera que exista, pero tampoco nadie ha demostrado que no.

Cuidado

La identidad $\alpha + \tau = n$ vale siempre, pero la igualdad entre emparejamiento máximo y recubrimiento mínimo (König) vale *solo* en grafos bipartitos; en C_5, por ejemplo, el emparejamiento máximo tiene tamaño 2 y el recubrimiento mínimo tamaño 3. Segundo aviso: la desigualdad $\omega \leq \chi$ puede ser muy holgada, y creer que colorear un grafo cuesta tantos colores como su mayor clique es un error clásico; el grafo de Petersen y los ciclos impares lo desmienten. Y tercero: el algoritmo voraz para el conjunto independiente da una respuesta correcta como conjunto pero rara vez máxima, así que conviene decir siempre si el número obtenido es exacto o solo una cota inferior.

Pausa para pensar

La palabra *clique* viene del francés y entró en la literatura matemática en 1949, de la mano de un trabajo de sociología: Luce y Perry estudiaban grupos de personas mutuamente conocidas y tomaron prestado el término coloquial para camarilla. Es un buen recordatorio de que la teoría de grafos ha crecido tanto desde la sociología y la informática como desde las matemáticas puras. La lista de Karp de 1972 fue el momento en que quedó claro que decenas de problemas aparentemente inconexos (rutas, horarios, empaquetado, coloración) eran en realidad el mismo problema con otra ropa; medio siglo después, sigue sin saberse si todos son difíciles o todos son fáciles, y es el problema abierto más famoso de la informática.

8.16. Errores típicos con grafos

1. Confundir camino euleriano (todas las *aristas*) con ciclo hamiltoniano (todos los *vértices*).
2. Creer que la condición de Euler exige que *todos* los grados sean pares; solo lo exige si se quiere un circuito cerrado.
3. Olvidar comprobar que el grafo es conexo antes de aplicar el teorema de Euler.
4. Olvidar la cara exterior al usar la fórmula $V - A + C = 2$.
5. Aplicar la fórmula de Euler a un dibujo con cruces: solo vale para dibujos *sin* cruces.
6. Deducir que un grafo es plano porque cumple $A \leq 3V - 6$; esa desigualdad solo sirve para descartar, no para confirmar.
7. Contar mal las aristas de K_n usando $n(n-1)$ en lugar de $n(n-1)/2$, es decir, contando cada pareja dos veces.
8. Afirmar que existe un grafo cuya suma de grados es impar.
9. Suponer que un árbol tiene siempre una raíz. Solo la tiene si se elige una: un árbol, como grafo, no tiene arriba ni abajo.
10. Usar Dijkstra con pesos negativos.

11. En Dijkstra, cerrar un vértice antes de haber examinado todas las mejoras posibles hacia él.
12. Creer que el número cromático de un grafo plano es siempre exactamente 4; el teorema dice "a lo sumo 4". Un árbol se colorea con 2.
13. Confundir el grafo con su dibujo, y concluir que dos grafos son distintos porque los dibujos lo parecen.

8.17. Isomorfismo de grafos: cuándo dos dibujos son el mismo objeto

Al principio del capítulo se dio un procedimiento para decidir si dos grafos coinciden, y con él se vio ya una pareja coespectral, la estrella $K_{1,4}$ y el cuadrado con un vértice suelto. Faltan dos cosas. La primera, la definición formal de isomorfismo, que es la que permite hablar de invariantes con precisión y la que da sentido al grupo de automorfismos. La segunda, entender hasta qué punto la lista de invariantes no se cierra: aquella pareja coespectral se distinguía a simple vista por los grados, y cabía la esperanza de que combinando invariantes se acabara zanjando la cuestión. Esta sección construye la pareja que mata esa esperanza.

Definición 8.73 (Isomorfismo de grafos)**.** Sean $G = (V_G, A_G)$ y $H = (V_H, A_H)$ dos grafos simples. Un *isomorfismo* de G en H es una biyección $\varphi : V_G \to V_H$ tal que, para todos $u, v \in V_G$,

$$\{u, v\} \in A_G \quad \Longleftrightarrow \quad \{\varphi(u), \varphi(v)\} \in A_H.$$

Si existe alguno, G y H son *isomorfos* y se escribe $G \cong H$. La doble implicación es esencial: pedir solo que las aristas vayan a aristas admitiría añadir aristas de más, y entonces todo grafo sería "isomorfo" al completo con el mismo número de vértices.

Un isomorfismo de G en sí mismo se llama *automorfismo*, y el conjunto de todos ellos forma un grupo, $\mathrm{Aut}(G)$, con la composición. Una magnitud asociada a un grafo que toma el mismo valor en grafos isomorfos se llama *invariante*.

Observación 8.74. La relación $\cong$ es de equivalencia: la identidad es un isomorfismo, la inversa de una biyección que preserva adyacencia también la preserva, y la composición de dos isomorfismos lo es. Por eso tiene sentido hablar de "el" grafo con tales propiedades salvo isomorfismo, y por eso contar grafos es contar clases de equivalencia. En términos de matrices, φ es un isomorfismo exactamente cuando la matriz de permutación P asociada cumple $P^{\mathsf{T}} A_G P = A_H$.

Regla: afinar los invariantes cuando la lista básica no decide

La regla "decidir si dos grafos son isomorfos", al principio del capítulo, agota los invariantes baratos —vértices, aristas, grados, componentes, cuello, bipartición— y con ellos se resuelve casi cualquier ejercicio. Cuando todos coinciden y aun así hay que decidir, se continúa con estos tres, ordenados de nuevo de más barato a más caro.

1. **Circuitos cerrados de cada longitud.** El número de recorridos cerrados de

longitud k es $\operatorname{tr}(A^k)$; en particular el número de triángulos es $\frac{1}{6}\operatorname{tr}(A^3)$, porque cada triángulo se recorre en dos sentidos desde cada uno de sus tres vértices. Coste: unos pocos productos de matrices.

2. **Espectro**, es decir, la lista de autovalores de la matriz de adyacencia con sus multiplicidades. Es un invariante porque las matrices de dos grafos isomorfos son conjugadas por una matriz de permutación, y matrices conjugadas tienen los mismos autovalores. Coste $O(n^3)$, y es con mucho el más fino de la lista. Nótese que el espectro determina todos los $\operatorname{tr}(A^k)$ a la vez, de modo que este escalón subsume el anterior.

3. **Refinamiento por vecindarios.** En vez de comparar los grados a secas, compárense los *multiconjuntos de grados de los vecinos* de cada vértice; eso parte los vértices en clases. Repítase usando las clases en lugar de los grados, y otra vez, hasta que la partición se estabilice. Dos grafos isomorfos han de tener particiones estables idénticas, y si lo son, la biyección solo puede emparejar vértices de la misma clase: lo que quede se prueba por fuerza bruta sobre un conjunto ya muy reducido. Es la estrategia real de los programas de isomorfismo.

La asimetría de fondo no cambia por afinar: un invariante que discrepa es una *demostración* de que no son isomorfos; un invariante que coincide, por fino que sea, es solo ausencia de información. El ejemplo siguiente agota los tres escalones sin decidir nada.

Ejemplo 8.75 (Dos grafos que superan todos los invariantes de la lista)**.** El ejemplo clásico son el *grafo de la torre* 4×4 y el *grafo de Shrikhande*.

Construcción. Los vértices de ambos son los 16 pares (i, j) con $i, j \in \{0, 1, 2, 3\}$. En el grafo de la torre, (i, j) y (k, l) son adyacentes si comparten fila o columna: son las casillas de un tablero 4×4 unidas cuando una torre de ajedrez va de una a otra en un movimiento. En el de Shrikhande, (i, j) y (k, l) son adyacentes cuando la diferencia $(i - k, j - l)$, tomada módulo 4, pertenece al conjunto

$$\{(1,0),\ (3,0),\ (0,1),\ (0,3),\ (1,1),\ (3,3)\},$$

que es la construcción anterior con las columnas "torcidas" a lo largo de una diagonal.

Todo coincide. Ambos tienen 16 vértices y 48 aristas; ambos son 6-regulares, luego comparten sucesión de grados; ambos son conexos, de una sola componente; ninguno es bipartito; ambos tienen cuello 3 y exactamente 32 triángulos; y el espectro de ambos es el mismo,

$$6 \text{ (una vez)}, \qquad 2 \text{ (seis veces)}, \qquad -2 \text{ (nueve veces)}.$$

Los siete primeros escalones de la regla se agotan sin decidir nada. De hecho ambos son grafos fuertemente regulares con los mismos parámetros: dos vértices adyacentes tienen siempre 2 vecinos comunes, y dos no adyacentes también.

Y sin embargo no son isomorfos. Lo que los separa es la estructura *local*. Tómese un vértice y mírense sus seis vecinos con las aristas que hay entre ellos. En el grafo de la torre ese subgrafo son dos triángulos disjuntos —los tres compañeros de fila por un lado y los tres

de columna por otro, sin ninguna arista entre ambos grupos—; en el de Shrikhande es un hexágono, un único ciclo de seis vértices. Seis vértices y seis aristas en los dos casos, todos de grado 2 en los dos casos, pero uno es conexo y el otro no. Como un isomorfismo llevaría vecindarios a vecindarios, no puede existir.

Observación 8.76. Los grafos coespectrales, que ya aparecieron con $K_{1,4}$, no son una rareza de dos casos aislados. El ejemplo más pequeño con forma de árbol aparece ya con ocho vértices: el árbol que se obtiene colgando cuatro hojas de un vértice y una cadena de longitud tres del mismo vértice, y el árbol formado por dos vértices unidos entre sí de los que cuelgan tres hojas cada uno, tienen ambos polinomio característico

$$x^8 - 7x^6 + 9x^4$$

y no son isomorfos; aquí basta mirar los grados, que son $(1,1,1,1,1,2,2,5)$ y $(1,1,1,1,1,1,4,4)$. Y se sabe además que la proporción de árboles que tienen un gemelo coespectral no tiende a cero al crecer el número de vértices: lo raro es el árbol determinado por su espectro, no el que no lo está. La moraleja es doble. Primera: el espectro, siendo mucho más fino que los grados, tampoco basta, y la torre y Shrikhande muestran que añadir los grados tampoco cierra el problema. Segunda: los invariantes se complementan de forma imprevisible, y cuál falla depende del par concreto, de modo que la única certeza posible sigue siendo exhibir la biyección o demostrar que no existe.

8.18. Ejercicios resueltos

Ejercicio 8.1. Un grafo tiene 6 vértices de grados $2, 2, 3, 3, 4, 4$. ¿Cuántas aristas tiene?

Solución. La suma de grados es $2+2+3+3+4+4 = 18$, luego por el lema del apretón de manos $|A| = 18/2 = 9$ aristas. Además hay dos vértices de grado impar, un número par ✓, como exige el corolario. ■

Ejercicio 8.2. ¿Existe un grafo simple con 7 vértices, todos de grado 3?

Solución. La suma de grados sería 21, impar, y debería ser el doble del número de aristas. No existe. En general, no puede haber un número impar de vértices de grado impar. ■

Ejercicio 8.3. Un grafo plano conexo tiene 10 vértices y 15 aristas. ¿Cuántas caras tiene su dibujo sin cruces?

Solución. De $V - A + C = 2$ sale $C = 2 - V + A = 2 - 10 + 15 = 7$ caras, incluida la exterior. Comprobación de coherencia: $A = 15 \leq 3V - 6 = 24$ ✓, así que la planaridad no queda descartada. ■

Ejercicio 8.4. Decide si la figura de un sobre de carta —un rectángulo con las dos diagonales del triángulo del cierre— puede dibujarse sin levantar el lápiz.

Solución. El sobre tiene 5 vértices: las cuatro esquinas del rectángulo y el pico del cierre, y 6 aristas: los cuatro lados del rectángulo y las dos que suben desde las esquinas superiores hasta el pico. Los grados son: 3 en cada esquina superior (dos lados del rectángulo más la arista al pico), 2 en cada esquina inferior y 2 en el pico. Suman $3+3+2+2+2 = 12$, es decir 6 aristas ✓, y hay exactamente dos vértices impares.

Por el teorema de Euler sí se puede, y además el trazo debe empezar en una esquina superior y terminar en la otra. Ese es justamente el motivo por el que el acertijo "sale" cuando uno empieza por el sitio adecuado y "no sale" cuando empieza por otro. ■

Ejercicio 8.5. ¿Cuál es el número cromático de un ciclo de 7 vértices? ¿Y de 8?

Solución. Con 7 (impar) hace falta $\chi = 3$: al alternar dos colores alrededor del ciclo, el séptimo vértice queda adyacente al primero y ambos tendrían el mismo. Con 8 (par) bastan $\chi = 2$, alternando. La regla general es que un ciclo tiene número cromático 2 si su longitud es par y 3 si es impar; y un grafo se puede colorear con dos colores exactamente cuando no tiene ningún ciclo de longitud impar. ■

Ejercicio 8.6. En el grafo ponderado con $AB = 2$, $AC = 5$, $BC = 1$, $BD = 4$, $CD = 2$, halla el camino más barato de A a D.

Solución. Se cierra A con 0 y se actualizan $B\colon 2$, $C\colon 5$. Se cierra B con 2 y se actualizan $C\colon 2+1=3$ (mejora el 5) y $D\colon 2+4=6$. Se cierra C con 3 y se actualiza $D\colon 3+2=5$ (mejora el 6). Se cierra D con **5**.

La ruta es $A \to B \to C \to D$, de coste 5, mejor que $A \to B \to D$ (6) y que $A \to C \to D$ (7). ■

8.19. Ejercicios propuestos

8.1. ¿Cuántas aristas tiene K_6? ¿Y K_{10}?

8.2. ¿Existe un grafo con 5 vértices todos de grado 3? Razona la respuesta.

8.3. Un grafo tiene 12 aristas. ¿Cuánto vale la suma de los grados de sus vértices?

8.4. ¿Cuántas aristas tiene un árbol de 20 vértices?

8.5. Un grafo plano conexo tiene 9 vértices y 6 caras. ¿Cuántas aristas tiene?

8.6. Comprueba la fórmula de Euler en el octaedro ($V = 6$, $A = 12$, $C = 8$).

8.7. ¿Tiene circuito euleriano un grafo cuyos grados son $4, 4, 4, 2$? ¿Y camino euleriano?

8.8. Explica por qué K_4 sí es plano y dibújalo sin cruces.

8.9. ¿Cuál es el número cromático de un árbol con más de un vértice?

8.10. Aplica Dijkstra al grafo con $AB = 3$, $AC = 1$, $BC = 1$, $BD = 2$, $CD = 5$ para ir de A a D.

8.11. ¿Cuántos árboles distintos hay sobre 4 vértices etiquetados?

8.12. Un mapa tiene diez países. ¿Se puede asegurar que basten cuatro colores? ¿Y que hagan falta cuatro?

Respuestas

1. $\binom{6}{2} = 15$ y $\binom{10}{2} = 45$. **2.** No: la suma de grados sería 15, impar, y debe ser par por el lema del apretón de manos. **3.** $2 \cdot 12 = 24$. **4.** 19, porque un árbol de n vértices tiene $n-1$ aristas. **5.** De $V - A + C = 2$: $A = V + C - 2 = 9 + 6 - 2 = 13$. **6.** $6 - 12 + 8 = 2$ ✓. **7.** Sí tiene circuito euleriano, porque todos los grados son pares (y suponiendo el grafo conexo); teniendo circuito, tiene camino euleriano trivialmente. **8.** $V = 4$, $A = 6$ y $3V - 6 = 6$, luego la cota no lo descarta; y de hecho se dibuja como un triángulo con un cuarto vértice *dentro*, unido a los tres, sin ningún cruce. Resulta $C = 4$ y $4 - 6 + 4 = 2$ ✓. **9.** 2: se colorean alternando por

niveles, y no puede ser 1 porque hay al menos una arista. Un árbol nunca tiene ciclos impares. **10.** Se cierra $A:0$ y se actualizan $B:3$, $C:1$; se cierra $C:1$ y se actualizan $B:2$ (mejora el 3) y $D:6$; se cierra $B:2$ y se actualiza $D:4$; se cierra $D:\mathbf{4}$, por la ruta $A \to C \to B \to D$. **11.** $4^{4-2} = 16$, por la fórmula de Cayley. **12.** Sí se puede asegurar que bastan cuatro, por el teorema de los cuatro colores, que vale para cualquier mapa. No se puede asegurar que hagan falta: un mapa de diez países en franjas horizontales se colorea con dos.

Capítulo 9

Cálculo multivariable

Lo que necesitas traer

Del capítulo 6: límites, derivadas, la regla de la cadena, integrales y el teorema del valor medio. Del capítulo 7: matrices, determinantes, valores propios y formas cuadráticas, sin los cuales la matriz hessiana y el jacobiano no significan nada.

Extendemos el cálculo a funciones de varias variables: campos escalares $f : \mathbb{R}^n \to \mathbb{R}$ y campos vectoriales $\mathbf{F} : \mathbb{R}^n \to \mathbb{R}^m$. La idea rectora viene del capítulo 6: *la derivada es la mejor aproximación lineal*; ahora "lineal" significa una transformación lineal del capítulo 7, y la matriz jacobiana toma el relevo de la derivada numérica.

9.1. Topología de $\mathbb{R}^n$

Idea clave

Intuición primero. Antes de derivar en varias variables hay que resolver un problema que en una variable no existía. En la recta, "acercarse a un punto" solo se puede hacer por dos lados, y por eso bastaba con límites laterales. En el plano hay infinitas maneras de acercarse a un punto: por rectas de cualquier pendiente, por parábolas, en espiral. El límite existe solo si *todas* esas maneras dan el mismo resultado, y por eso hay funciones cuyos límites a lo largo de cada recta coinciden y que aun así no tienen límite (basta acercarse por una parábola bien elegida para arruinarlo).

Esta sección monta el vocabulario mínimo para hablar de eso con precisión: bolas, abiertos, cerrados, frontera, compacidad. No es material de relleno: la compacidad —cerrado y acotado, por Heine–Borel— es exactamente la hipótesis que garantiza que una función continua alcanza su máximo, y sin ella la optimización de la sección siguiente no tendría teoremas. El capítulo 28 retoma estas ideas quitándoles la distancia.

En $\mathbb{R}^n$ usamos la norma euclídea $\|\mathbf{x}\| = \sqrt{x_1^2 + \cdots + x_n^2}$, con la distancia $d(\mathbf{x}, \mathbf{y}) = \|\mathbf{x} - \mathbf{y}\|$.

Definición 9.1. La *bola abierta* es $B(\mathbf{a}, r) = \{\mathbf{x} : \|\mathbf{x} - \mathbf{a}\| < r\}$. Un conjunto U es *abierto* si cada punto suyo tiene una bola contenida en U; es *cerrado* si su complemento es abierto

(equivalentemente: contiene los límites de sus sucesiones convergentes). La *frontera* ∂S son los puntos cuyas bolas cortan a S y a su complemento. S es *acotado* si cabe en alguna bola, y *compacto* si es cerrado y acotado (Heine–Borel). Un abierto es *conexo* si no se parte en dos abiertos disjuntos no vacíos; para abiertos de $\mathbb{R}^n$ equivale a que dos puntos cualesquiera se unan por un camino.

Los teoremas de una variable se trasladan: una sucesión en un compacto tiene subsucesión convergente (Bolzano–Weierstrass), y

Teorema 9.2 (Weierstrass en $\mathbb{R}^n$)**.** *Toda función continua sobre un compacto está acotada y alcanza máximo y mínimo. Este teorema es el que garantiza que los problemas de optimización de la sección 9.4 tengan solución.*

9.1.1. Límites: el peligro de las trayectorias

Definición 9.3. $\lim_{\mathbf{x}\to\mathbf{a}} f(\mathbf{x}) = L$ si para todo $\varepsilon > 0$ existe $\delta > 0$ tal que $0 < \|\mathbf{x}-\mathbf{a}\| < \delta \implies |f(\mathbf{x}) - L| < \varepsilon$.

Cuidado

En $\mathbb{R}^n$ hay infinitas maneras de acercarse a un punto. Que el límite exista a lo largo de todas las rectas **no** garantiza que exista: hace falta el mismo valor por *toda* trayectoria. Estrategia doble: *para refutar*, encontrar dos caminos con límites distintos (probar rectas $y = mx$, parábolas $y = x^2$, etc.); *para demostrar*, acotar $|f(\mathbf{x}) - L|$ por algo que tienda a cero, típicamente usando polares ($x = r\cos\theta$, $y = r\sin\theta$) y comprobando que la cota no depende de θ.

Ejemplo 9.4. $\lim_{(x,y)\to(0,0)} \frac{x^2y}{x^2+y^2} = 0$: en polares, $\left|\frac{r^3\cos^2\theta\sin\theta}{r^2}\right| = r|\cos^2\theta\sin\theta| \le r \to 0$. En cambio $\frac{xy}{x^2+y^2}$ no tiene límite: sobre $y = mx$ vale $\frac{m}{1+m^2}$, distinto para cada pendiente.

Regla: decidir si un límite en varias variables existe

En una variable solo había dos maneras de acercarse a un punto: por la izquierda y por la derecha. En el plano hay infinitas, y esa es toda la dificultad. La estrategia siempre es la misma: *primero se intenta destruir el límite, y solo si no se consigue se intenta demostrarlo.*

1. **Sustituir.** Si la función es continua en el punto (cociente con denominador no nulo, composición de continuas), el límite es el valor. No hay más que hacer. Todo lo que sigue supone una indeterminación real, casi siempre $0/0$.
2. **Probar con rectas $y = mx$.** Se sustituye y se simplifica. Si el resultado *depende de* m, el límite no existe y se acabó: dos rectas distintas ya son dos trayectorias con límites distintos.
3. **Si todas las rectas dan lo mismo, no cantar victoria.** Este es el error central

del tema. Que todas las rectas coincidan *no* demuestra nada. Hay que probar además con parábolas $y = kx^2$, con cúbicas $y = kx^3$, y en general con la curva que iguale los grados del numerador y del denominador.

4. **Cómo elegir la curva asesina.** Se mira el denominador: si tiene $x^4 + y^2$, la curva que hace competir los dos términos es $y = x^2$; si tiene $x^2 + y^6$, es $y^3 = x$, o sea $x = y^3$. La receta es igualar los pesos de los dos sumandos.

5. **Si el límite parece existir, pasar a polares.** Con $x = r\cos\theta$, $y = r\sin\theta$, el punto tiende al origen exactamente cuando $r \to 0$. Si se consigue una cota $|f - L| \le g(r)$ con $g(r) \to 0$ *y* g *sin ningún* θ *dentro*, el límite existe y vale L.

6. **O acotar directamente (el sandwich).** Las dos desigualdades que resuelven la mayoría de los ejercicios son $|xy| \le \frac{x^2+y^2}{2}$ y $\frac{|x|}{\sqrt{x^2+y^2}} \le 1$. Con ellas casi todo se reduce a "algo acotado por algo que tiende a cero".

7. **Control de coherencia.** Los límites iterados $\lim_{x\to 0}\lim_{y\to 0} f$ y $\lim_{y\to 0}\lim_{x\to 0} f$ son gratis de calcular. Si salen distintos, el límite doble no existe. Si salen iguales, no se ha demostrado nada: solo se ha ganado confianza.

Ejemplo 9.5 (Cuatro límites en el origen)**. (a) Muere con rectas.** $\lim\limits_{(x,y)\to(0,0)} \dfrac{xy}{x^2+y^2}$. Sobre $y = mx$:

$$\frac{x \cdot mx}{x^2 + m^2x^2} = \frac{m}{1+m^2},$$

que depende de m: vale 0 sobre el eje x ($m = 0$) y $\frac{1}{2}$ sobre la diagonal ($m = 1$). El límite *no existe*. Los dos límites iterados, en cambio, valen ambos 0: buen recordatorio de que el control del paso 7 no es una demostración.

(b) El clásico traicionero. $\lim\limits_{(x,y)\to(0,0)} \dfrac{x^2y}{x^4+y^2}$. Sobre $y = mx$ sale $\frac{mx^3}{x^4+m^2x^2} = \frac{mx}{x^2+m^2} \to 0$: *todas* las rectas dan cero. Pero el denominador tiene x^4 frente a y^2, así que la curva a probar es $y = x^2$:

$$\frac{x^2 \cdot x^2}{x^4 + x^4} = \frac{1}{2}.$$

Cero por todas las rectas y un medio por una parábola: el límite no existe. Este ejemplo es la razón de ser del paso 3.

(c) Sí existe, por polares. $\lim\limits_{(x,y)\to(0,0)} \dfrac{x^3}{x^2+y^2}$. En polares el cociente es

$$\frac{r^3\cos^3\theta}{r^2} = r\cos^3\theta,$$

y como $|\cos^3\theta| \le 1$, se tiene $|f| \le r \to 0$ *uniformemente en* θ. El límite existe y vale 0.

(d) Sí existe, por sandwich. $\lim\limits_{(x,y)\to(0,0)} \dfrac{x^2y^2}{x^2+y^2}$. Usando $x^2 \le x^2 + y^2$:

$$0 \le \frac{x^2y^2}{x^2+y^2} \le \frac{(x^2+y^2)y^2}{x^2+y^2} = y^2 \to 0.$$

El límite es 0, sin necesidad de polares ni de probar curva alguna.

Cuidado

En polares hay una trampa fina: no basta con que el límite en $r \to 0$ salga 0 *para cada* θ *fijo*. La cota tiene que ser independiente de θ. Para $f = \frac{x^2 y}{x^4+y^2}$ del apartado (b), en polares queda $\frac{r\cos^2\theta \sin\theta}{r^2\cos^4\theta + \sin^2\theta}$, que para cada θ fijo tiende a 0 cuando $r \to 0$ (con $\sin\theta \neq 0$) y sin embargo el límite no existe: el denominador se hace pequeño justo cuando θ se acerca a 0 al mismo ritmo que r. Trayectoria radial fijada no es lo mismo que trayectoria cualquiera.

9.2. La derivada parcial: definición, todas sus notaciones, y el aviso que hay que aprenderse de memoria

La sección anterior montó la topología de $\mathbb{R}^n$ y los límites. A partir de aquí el capítulo deriva, integra y optimiza sin parar, y lo hace con un puñado de símbolos —$\frac{\partial f}{\partial x}$, $\operatorname{grad} f$, H_f, $\mathrm{d}A$, $\rho^2 \sin\varphi$— que conviene tener definidos uno por uno antes de usarlos. Esta sección hace exactamente eso: define la derivada parcial con su límite, enumera las cuatro notaciones con que aparece en la literatura, fija qué significa una derivada segunda cruzada, y precisa qué es un punto crítico, una hessiana, una integral iterada, un jacobiano y cada uno de los tres sistemas de coordenadas. Nada de esto es difícil; lo que sí es peligroso es usarlo sin haberlo escrito.

Idea clave

Intuición primero. Derivar una función de una variable era medir una pendiente: cuánto sube f por cada unidad que avanza x. Con dos variables el problema es que "avanzar" ya no significa una sola cosa: puedes avanzar hacia el este, hacia el norte, o en cualquier diagonal. La salida más barata es congelar todas las variables menos una y medir la pendiente en esa única dirección. Eso es una derivada parcial: una derivada de toda la vida, disfrazada, en la que las demás letras se comportan como si fueran números fijos.

Lo barato tiene su precio, y este capítulo lo cobra entero. Congelar variables solo mira a lo largo de dos avenidas, y una superficie puede estar perfectamente lisa sobre las dos avenidas y ser un desastre en las diagonales. Al final de esta sección hay una función concreta —la más importante del capítulo— cuyas dos parciales existen y valen cero en el origen, y que ni siquiera es continua ahí.

Definición 9.6 (Derivada parcial y sus notaciones)**.** Sea $U \subseteq \mathbb{R}^n$ abierto, $f\colon U \to \mathbb{R}$ y $\mathbf{a} = (a_1, \dots, a_n) \in U$. La *derivada parcial* de f respecto de la variable x_i en el punto $\mathbf{a}$ es el límite

$$\frac{\partial f}{\partial x_i}(\mathbf{a}) \;=\; \lim_{h\to 0} \frac{f(a_1, \dots, a_{i-1},\, a_i + h,\, a_{i+1}, \dots, a_n) - f(\mathbf{a})}{h},$$

cuando ese límite existe y es finito. Escribiendo $\mathbf{e}_i$ para el i-ésimo vector de la base canónica,

la misma fórmula es

$$\frac{\partial f}{\partial x_i}(\mathbf{a}) = \lim_{h\to 0} \frac{f(\mathbf{a}+h\mathbf{e}_i) - f(\mathbf{a})}{h}.$$

Las cuatro notaciones que se usan indistintamente son

$$\frac{\partial f}{\partial x_i}(\mathbf{a}), \qquad f_{x_i}(\mathbf{a}), \qquad \partial_{x_i} f(\mathbf{a}), \qquad D_i f(\mathbf{a}),$$

y para $n = 2$, con las variables llamadas x e y, se escriben f_x y f_y. El símbolo ∂ es una "d" redondeada y no significa nada por sí solo: $\frac{\partial f}{\partial x}$ es *un único símbolo*, no un cociente de dos cantidades ∂f y ∂x, y no se puede simplificar como una fracción.

Observación 9.7. La definición dice literalmente esto: fija todas las coordenadas salvo la i-ésima, con lo que queda una función de una sola variable real, $t \mapsto f(a_1, \dots, t, \dots, a_n)$, y derívala en $t = a_i$ con la derivada de siempre. Por eso no hace falta ninguna regla nueva: valen la regla del producto, la del cociente y la de la cadena del capítulo 6, aplicadas tratando las demás letras como constantes.

Regla: calcular una derivada parcial sin equivocarse

1. **Decide primero cuál es la variable viva.** Escríbelo aparte: "derivo respecto de x, luego y y z son números". La mitad de los errores del tema nacen de olvidar cuál se movía.
2. **Marca las constantes.** Antes de escribir nada, rodea mentalmente cada aparición de las variables congeladas. En x^2y, respecto de x, el factor y es un número que acompaña, exactamente igual que el 5 de $5x^2$.
3. **Deriva con las reglas de una variable, todas.** Producto, cociente, cadena, exponenciales, logaritmos: no cambia ninguna. Si aparece una función compuesta, la derivada de lo de dentro es también una parcial.
4. **Ojo con los términos que no contienen la variable viva.** Su derivada es cero, entera. Un sumando y^4 desaparece al derivar respecto de x; no se queda a medias.
5. **Sustituye el punto solo al final.** Deriva en forma simbólica y evalúa después. Sustituir antes de derivar es el error que convierte una parcial en un cero espurio.
6. **Si la fórmula se rompe en el punto, vuelve a la definición.** En un punto donde el denominador se anula o la función está definida a trozos, la fórmula general no vale y hay que calcular el límite del cociente incremental a mano. Este paso es el que separa el ejercicio rutinario del contraejemplo importante.
7. **Control barato.** Las unidades y los grados deben cuadrar: si f es un polinomio homogéneo de grado d, cada parcial es homogénea de grado $d-1$. Y la suma $xf_x + yf_y$ debe dar $d\,f$ por el teorema de Euler.

Ejemplo 9.8 (Cuatro parciales calculadas del todo)**. (a) Un polinomio.** Sea $f(x, y) =$

$x^3y^2 - 2xy^4 + 5$. Respecto de x, el factor y^2 es constante, el y^4 también, y el 5 desaparece:

$$\frac{\partial f}{\partial x} = 3x^2y^2 - 2y^4.$$

Respecto de y, ahora los constantes son x^3 y x:

$$\frac{\partial f}{\partial y} = 2x^3y - 8xy^3.$$

En el punto $(2,1)$ sale $\frac{\partial f}{\partial x}(2,1) = 3\cdot 4\cdot 1 - 2\cdot 1 = 10$ y $\frac{\partial f}{\partial y}(2,1) = 2\cdot 8\cdot 1 - 8\cdot 2\cdot 1 = 16 - 16 = 0$. Control del paso 7: f no es homogénea (el 5 lo estropea), pero su parte de grado 5 sí lo es, y en efecto $x(3x^2y^2 - 2y^4) + y(2x^3y - 8xy^3) = 5x^3y^2 - 10xy^4$, que es 5 veces la parte homogénea. ✓

(b) Un cociente. Sea $f(x,y) = \dfrac{x}{y}$ con $y \neq 0$. Respecto de x el denominador es una constante, así que $\frac{\partial f}{\partial x} = \dfrac{1}{y}$. Respecto de y, la función es $x\cdot y^{-1}$ con x constante, luego $\frac{\partial f}{\partial y} = -\dfrac{x}{y^2}$. Nótese la asimetría total entre las dos: una crece con y y la otra decrece, y ninguna se parece a la otra.

(c) Una exponencial compuesta. Sea $f(x,y) = \mathrm{e}^{xy}$. La cadena da $\frac{\partial f}{\partial x} = y\,\mathrm{e}^{xy}$ y $\frac{\partial f}{\partial y} = x\,\mathrm{e}^{xy}$: lo de fuera se deriva igual y lo de dentro aporta su propia parcial. En $(0,3)$ valen 3 y 0 respectivamente.

(d) Una raíz, y un punto donde la fórmula falla. Sea $f(x,y) = \sqrt{x^2+y^2}$. Fuera del origen,

$$\frac{\partial f}{\partial x} = \frac{x}{\sqrt{x^2+y^2}}, \qquad \frac{\partial f}{\partial y} = \frac{y}{\sqrt{x^2+y^2}}.$$

En $(3,4)$ valen $3/5 = 0{,}6$ y $4/5 = 0{,}8$. Pero en el origen la fórmula da $0/0$, así que hay que ir a la definición (paso 6):

$$\lim_{h\to 0} \frac{f(h,0) - f(0,0)}{h} = \lim_{h\to 0} \frac{|h|}{h},$$

que vale $+1$ por la derecha y -1 por la izquierda: *el límite no existe* y $\frac{\partial f}{\partial x}(0,0)$ no está definida. Es el cono con el pico en el vértice, y la parcial es la derivada de $|h|$ en cero.

Definición 9.9 (Derivadas parciales de orden superior)**.** Si $\frac{\partial f}{\partial x_i}$ existe en todo un abierto, es a su vez una función de n variables y se le pueden volver a tomar parciales. La derivada segunda que primero deriva respecto de x_i y después respecto de x_j se escribe

$$\frac{\partial^2 f}{\partial x_j\,\partial x_i} \;=\; \frac{\partial}{\partial x_j}\Big(\frac{\partial f}{\partial x_i}\Big), \qquad \text{y también} \qquad f_{x_i x_j}, \qquad D_{ij}f.$$

Atención al orden, que es la fuente inagotable de confusiones: en la notación con ∂ el índice que se lee *primero* (el más cercano a la f, o sea el de la derecha) es el que se deriva *primero*, mientras que en la notación con subíndices f_{xy} se lee de izquierda a derecha, o sea primero x y luego y. Por tanto

$$f_{xy} = \frac{\partial^2 f}{\partial y\,\partial x}.$$

Cuando $i = j$ se escribe $\dfrac{\partial^2 f}{\partial x_i^2}$ o $f_{x_i x_i}$. Se llaman *cruzadas* a las derivadas segundas con $i \neq j$. Las derivadas de orden k se definen igual, iterando.

Ejemplo 9.10 (Las cuatro segundas del polinomio de antes)**.** Con $f(x,y) = x^3y^2 - 2xy^4 + 5$, partiendo de $f_x = 3x^2y^2 - 2y^4$ y $f_y = 2x^3y - 8xy^3$:

$$f_{xx} = 6xy^2, \qquad f_{xy} = 6x^2y - 8y^3, \qquad f_{yx} = 6x^2y - 8y^3, \qquad f_{yy} = 2x^3 - 24xy^2.$$

Las dos cruzadas coinciden, como anuncia el teorema de Clairaut–Schwarz (que se enuncia y se demuestra más adelante en este mismo capítulo). En el punto $(2,1)$ los cuatro números son $f_{xx} = 12$, $f_{xy} = f_{yx} = 24 - 8 = 16$ y $f_{yy} = 16 - 48 = -32$.

Ejemplo 9.11 (El contraejemplo que hay que saber de memoria)**.** Sea

$$f(x,y) = \begin{cases} \dfrac{xy}{x^2+y^2}, & (x,y) \neq (0,0), \\ 0, & (x,y) = (0,0). \end{cases}$$

Vamos a calcular *todo* sobre esta función en el origen, porque cada pieza desmonta una creencia distinta.

(1) Las dos parciales existen en el origen y valen cero. Sobre el eje x la función es idénticamente nula: $f(h,0) = \dfrac{h \cdot 0}{h^2 + 0} = 0$ para todo $h \neq 0$, y $f(0,0) = 0$. Por tanto

$$\frac{\partial f}{\partial x}(0,0) = \lim_{h\to 0} \frac{f(h,0) - f(0,0)}{h} = \lim_{h\to 0} \frac{0-0}{h} = \lim_{h\to 0} 0 = 0.$$

Por la simetría $f(x,y) = f(y,x)$, el mismo cálculo sobre el eje y da $\frac{\partial f}{\partial y}(0,0) = 0$. Las dos parciales existen, son finitas y valen cero. El gradiente en el origen es $\operatorname{grad} f(0,0) = (0,0)$.

(2) Y sin embargo la función no es continua en el origen. Sobre la recta $y = mx$ (con $x \neq 0$),

$$f(x, mx) = \frac{x \cdot mx}{x^2 + m^2x^2} = \frac{mx^2}{x^2(1+m^2)} = \frac{m}{1+m^2},$$

un valor *constante* sobre cada recta y *distinto* para cada pendiente: vale 0 sobre el eje x ($m = 0$), $\frac{1}{2}$ sobre la diagonal $y = x$ ($m = 1$) y $-\frac{1}{2}$ sobre $y = -x$. Como el valor límite depende del camino, $\lim_{(x,y)\to(0,0)} f$ no existe, y en particular no vale $f(0,0) = 0$: la función no es continua en el origen. En polares la cuenta es todavía más elocuente:

$$f(r\cos\theta,\, r\sin\theta) = \frac{r^2 \cos\theta \sin\theta}{r^2} = \cos\theta\sin\theta = \frac{\sin(2\theta)}{2},$$

que *no depende de* r: por más que uno se acerque al origen, el valor de f sigue oscilando entre $-\frac{1}{2}$ y $\frac{1}{2}$ según por dónde se acerque. La superficie no tiene un único valor límite en el origen; tiene un abanico entero de ellos.

(3) Las derivadas direccionales en diagonal ni siquiera existen. Para una dirección unitaria $\mathbf{u} = (a,b)$ con $a \neq 0$ y $b \neq 0$, el cociente incremental es

$$\frac{f(ta,tb) - f(0,0)}{t} = \frac{1}{t} \cdot \frac{t^2 ab}{t^2(a^2+b^2)} = \frac{ab}{t\,(a^2+b^2)},$$

que tiende a $\pm\infty$ cuando $t \to 0$. Es decir: de las infinitas derivadas direccionales, *solo existen las dos de los ejes*, que son las parciales. En dirección $(1,1)/\sqrt{2}$ el cociente vale $\dfrac{1/2}{t}$ y se dispara.

(4) Fuera del origen todo es normal. Ahí la fórmula es un cociente de polinomios con denominador no nulo, luego f es de clase C^∞, y las parciales son

$$\frac{\partial f}{\partial x} = \frac{y\,(y^2 - x^2)}{(x^2+y^2)^2}, \qquad \frac{\partial f}{\partial y} = \frac{x\,(x^2 - y^2)}{(x^2+y^2)^2}.$$

Comprobación en $(1,1)$: las dos valen 0, coherente con que sobre la diagonal f sea constante e igual a $\frac{1}{2}$.

Cuidado

Esta es la advertencia central del capítulo, y conviene copiarla a mano. **Que existan todas las derivadas parciales en un punto no implica que la función sea diferenciable ahí, y ni siquiera implica que sea continua.** El ejemplo 9.11 lo demuestra con el mínimo de esfuerzo: dos parciales que existen y valen 0, y una función que no tiene límite en el punto.

La razón es geométrica y se resume en una frase: *las parciales solo miran a lo largo de los dos ejes*. La función $xy/(x^2+y^2)$ vale idénticamente cero sobre los dos ejes —por eso sus parciales salen nulas sin esfuerzo— y vale $\frac{1}{2}$ sobre toda la diagonal. Las avenidas están lisas; el descampado entre ellas, no. La diferenciabilidad, en cambio, exige que la aproximación lineal funcione en *todas* las direcciones a la vez y con error $o(\|\mathbf{h}\|)$, que es una condición mucho más fuerte.

En la práctica, la implicación que sí funciona es la de la sección siguiente: si las parciales existen y son *continuas* en un entorno, entonces la función es diferenciable. Esa continuidad es exactamente la hipótesis que impide la arruga entre avenidas, y es la que cumplen todas las funciones razonables. Pero es una hipótesis, no un regalo.

Definición 9.12 (Punto crítico, matriz hessiana y punto de silla)**.** Sea $f\colon U \subseteq \mathbb{R}^n \to \mathbb{R}$ con U abierto. Un punto $\mathbf{a} \in U$ es un *punto crítico* (o *estacionario*) de f si f es diferenciable en $\mathbf{a}$ y $\operatorname{grad} f(\mathbf{a}) = \mathbf{0}$, o bien si alguna parcial de f no existe en $\mathbf{a}$. Si f es de clase C^2, su *matriz hessiana* en $\mathbf{a}$ es la matriz $n \times n$ de las derivadas segundas,

$$H_f(\mathbf{a}) = \Big(\frac{\partial^2 f}{\partial x_i\,\partial x_j}(\mathbf{a})\Big)_{i,j=1}^{n},$$

simétrica por el teorema de Clairaut–Schwarz. Un punto crítico se llama *punto de silla* si no es ni máximo ni mínimo local, es decir, si en todo entorno suyo hay puntos donde f es mayor y puntos donde f es menor que $f(\mathbf{a})$.

Definición 9.13 (Diferencial total y plano tangente)**.** Si f es diferenciable en $\mathbf{a}$, su *diferencial* en $\mathbf{a}$ es la aplicación lineal $\mathrm{d}f_{\mathbf{a}}\colon \mathbb{R}^n \to \mathbb{R}$ dada por

$$\mathrm{d}f_{\mathbf{a}}(\mathbf{h}) = \operatorname{grad} f(\mathbf{a}) \cdot \mathbf{h} = \sum_{i=1}^{n} \frac{\partial f}{\partial x_i}(\mathbf{a})\, h_i,$$

que en la notación clásica se escribe $\mathrm{d}f = \frac{\partial f}{\partial x}\mathrm{d}x + \frac{\partial f}{\partial y}\mathrm{d}y$. Para $n = 2$, el *plano tangente* a la gráfica $z = f(x, y)$ en el punto $(a, b, f(a, b))$ es el plano de ecuación

$$z = f(a, b) + \frac{\partial f}{\partial x}(a, b)\,(x - a) + \frac{\partial f}{\partial y}(a, b)\,(y - b),$$

es decir, la gráfica de la aproximación lineal de f en (a, b). La *recta normal* es la que pasa por el punto con vector director $\big(f_x(a, b),\ f_y(a, b),\ -1\big)$.

Ejemplo 9.14 (Un plano tangente y la calidad de la aproximación)**.** Sea $f(x, y) = x^2 + 3xy - y^2$ en el punto $(1, 2)$. Se tiene $f(1, 2) = 1 + 6 - 4 = 3$, $f_x = 2x + 3y$ con $f_x(1, 2) = 8$, y $f_y = 3x - 2y$ con $f_y(1, 2) = -1$. El plano tangente es

$$z = 3 + 8(x - 1) - 1\,(y - 2) = 8x - y - 3.$$

Comprobación: en $(1, 2)$ da $8 - 2 - 3 = 3$. ✓Usémoslo para estimar $f(1{,}1;\ 1{,}9)$, con $\mathbf{h} = (0{,}1;\ -0{,}1)$:

$$f(1{,}1;\ 1{,}9) \approx 3 + 8 \cdot (0{,}1) + (-1) \cdot (-0{,}1) = 3 + 0{,}8 + 0{,}1 = 3{,}9.$$

El valor exacto es $1{,}21 + 6{,}27 - 3{,}61 = 3{,}87$, con un error de $0{,}03$; y en efecto $\|\mathbf{h}\|^2 = 0{,}02$, del mismo orden que el error, como debe ser en una aproximación de primer orden. El diferencial es $\mathrm{d}f = 8\,\mathrm{d}x - \mathrm{d}y$ en ese punto: un movimiento de una milésima hacia el este cambia f ocho veces más que el mismo movimiento hacia el norte.

Definición 9.15 (Integral iterada)**.** Sea f continua sobre una región del plano descrita por $a \le x \le b$ y $\varphi_1(x) \le y \le \varphi_2(x)$. La *integral iterada* en el orden $\mathrm{d}y\,\mathrm{d}x$ es el número que se obtiene en dos etapas: primero se calcula, para cada x fijo, la integral de una variable

$$G(x) = \int_{\varphi_1(x)}^{\varphi_2(x)} f(x, y)\,\mathrm{d}y,$$

en la que x actúa como una constante; y después se integra el resultado,

$$\int_a^b \int_{\varphi_1(x)}^{\varphi_2(x)} f(x, y)\,\mathrm{d}y\,\mathrm{d}x := \int_a^b G(x)\,\mathrm{d}x.$$

El orden $\mathrm{d}x\,\mathrm{d}y$ se define simétricamente. Obsérvese la regla estructural: *los límites de la integral exterior son siempre constantes*, y los de la interior pueden depender de la variable exterior, nunca al revés. El teorema de Fubini, más abajo, dice cuándo las dos integrales iteradas coinciden con la integral doble.

Ejemplo 9.16 (Una integral iterada, en los dos órdenes)**.** Sea D el triángulo $\{0 \le x \le 2,\ 0 \le y \le x\}$ y $f(x, y) = x + 2y$. En el orden $\mathrm{d}y\,\mathrm{d}x$, la etapa interior es

$$G(x) = \int_0^x (x + 2y)\,\mathrm{d}y = \big[xy + y^2\big]_0^x = x^2 + x^2 = 2x^2,$$

y la exterior, $\displaystyle\int_0^2 2x^2\,\mathrm{d}x = \frac{2 \cdot 8}{3} = \frac{16}{3}$.

Descrito al revés, el mismo triángulo es $\{0 \le y \le 2,\ y \le x \le 2\}$. Ahora la etapa interior es

$$\int_y^2 (x+2y)\,\mathrm{d}x = \Big[\frac{x^2}{2} + 2yx\Big]_y^2 = (2+4y) - \Big(\frac{y^2}{2} + 2y^2\Big) = 2 + 4y - \frac{5y^2}{2},$$

y la exterior,

$$\int_0^2 \Big(2 + 4y - \frac{5y^2}{2}\Big)\mathrm{d}y = 4 + 8 - \frac{5}{2}\cdot\frac{8}{3} = 12 - \frac{20}{3} = \frac{16}{3}.\checkmark$$

Los dos órdenes dan $16/3 \approx 5{,}333$, como garantiza Fubini. Que coincidan es el mejor control de que la región se ha leído bien.

Definición 9.17 (Jacobiano de un cambio de variable)**.** Sea $T\colon S \subseteq \mathbb{R}^n \to \mathbb{R}^n$ una aplicación de clase C^1, escrita $T(u_1,\dots,u_n) = (x_1,\dots,x_n)$. Su *matriz jacobiana* es $DT = (\frac{\partial x_i}{\partial u_j})$, y su *jacobiano* es el determinante de esa matriz, que se denota

$$\det DT = \frac{\partial(x_1,\dots,x_n)}{\partial(u_1,\dots,u_n)}.$$

Para $n = 2$ es simplemente

$$\frac{\partial(x,y)}{\partial(u,v)} = \begin{vmatrix} \frac{\partial x}{\partial u} & \frac{\partial x}{\partial v} \\ \frac{\partial y}{\partial u} & \frac{\partial y}{\partial v} \end{vmatrix} = \frac{\partial x}{\partial u}\frac{\partial y}{\partial v} - \frac{\partial x}{\partial v}\frac{\partial y}{\partial u}.$$

Su significado es geométrico y hay que retenerlo: $|\det DT|$ es el *factor por el que T multiplica los volúmenes* cerca del punto. Un cuadradito de lados $\mathrm{d}u$ y $\mathrm{d}v$ se convierte en un paralelogramo de área $|\det DT|\,\mathrm{d}u\,\mathrm{d}v$. Cuando el cambio se conoce en la dirección contraria, es útil la identidad

$$\frac{\partial(x,y)}{\partial(u,v)} = \Big(\frac{\partial(u,v)}{\partial(x,y)}\Big)^{-1},$$

consecuencia de que las jacobianas de T y T^{-1} son matrices inversas.

Definición 9.18 (Coordenadas polares, cilíndricas y esféricas)**.** Los tres cambios estándar, con los rangos de los parámetros que los hacen biyectivos, son:

- **Polares** en $\mathbb{R}^2$: $x = r\cos\theta$, $y = r\sin\theta$, con $r > 0$ y $\theta \in [0, 2\pi)$. Aquí $r = \sqrt{x^2+y^2}$ es la distancia al origen y θ el ángulo con el semieje x positivo.
- **Cilíndricas** en $\mathbb{R}^3$: $x = r\cos\theta$, $y = r\sin\theta$, $z = z$, con $r > 0$, $\theta \in [0, 2\pi)$ y $z \in \mathbb{R}$. Son las polares del plano xy con la altura sin tocar.
- **Esféricas** en $\mathbb{R}^3$: $x = \rho\sin\varphi\cos\theta$, $y = \rho\sin\varphi\sin\theta$, $z = \rho\cos\varphi$, con $\rho > 0$, $\varphi \in (0, \pi)$ el ángulo *desde el semieje z positivo* (colatitud) y $\theta \in [0, 2\pi)$ el ángulo azimutal. Aquí $\rho = \sqrt{x^2+y^2+z^2}$.

Sus jacobianos, calculados con la definición 9.17, son

$$\frac{\partial(x,y)}{\partial(r,\theta)} = r, \qquad \frac{\partial(x,y,z)}{\partial(r,\theta,z)} = r, \qquad \frac{\partial(x,y,z)}{\partial(\rho,\varphi,\theta)} = \rho^2\sin\varphi.$$

Ejemplo 9.19 (El jacobiano de las polares, calculado a mano)**.** Con $x = r\cos\theta$ e $y = r\sin\theta$ las cuatro parciales son $\frac{\partial x}{\partial r} = \cos\theta$, $\frac{\partial x}{\partial \theta} = -r\sin\theta$, $\frac{\partial y}{\partial r} = \sin\theta$, $\frac{\partial y}{\partial \theta} = r\cos\theta$, y el determinante es

$$\begin{vmatrix} \cos\theta & -r\sin\theta \\ \sin\theta & r\cos\theta \end{vmatrix} = r\cos^2\theta + r\sin^2\theta = r\,(\cos^2\theta + \sin^2\theta) = r.$$

De ahí la fórmula $\mathrm{d}A = r\,\mathrm{d}r\,\mathrm{d}\theta$ que se usará en todo el capítulo. Y de ahí también la única trampa: el factor r *no es opcional*, y olvidarlo es el error más frecuente de las integrales múltiples. El jacobiano se anula en $r = 0$, lo cual es coherente con que el origen sea el único punto donde el cambio deja de ser biyectivo: todos los ángulos θ le corresponden a la vez.

Regla: no perder el jacobiano al cambiar de coordenadas

1. **Escribe el cambio en la dirección** $(u, v) \mapsto (x, y)$, es decir, las viejas en función de las nuevas. Es la dirección en la que la fórmula del cambio de variable está enunciada.
2. **Calcula la matriz jacobiana entera antes de tomar el determinante.** Cuatro parciales en el plano, nueve en el espacio. Escribirlas todas evita confundir filas con columnas.
3. **Toma el valor absoluto del determinante.** La fórmula lleva $|\det DT|$, no $\det DT$. Un cambio que invierte la orientación tiene jacobiano negativo y la integral seguiría siendo positiva.
4. **Sustituye las tres cosas a la vez:** el integrando, los límites y el elemento de área o de volumen. Cambiar el integrando y olvidar el r es exactamente el error del que avisa el paso anterior.
5. **Si te dan el cambio al revés, invierte el número, no la fórmula.** Si es fácil calcular $\partial(u, v)/\partial(x, y)$, calcúlalo y toma su inverso. Es mucho más rápido que despejar x e y.
6. **Comprueba dónde falla la biyectividad.** El teorema pide $\det DT \neq 0$ y T inyectiva. En polares eso excluye el origen; como es un conjunto de área nula, la integral no se entera, pero conviene saber por qué se puede ignorar.
7. **Control dimensional.** Integra la función constante 1 con tu cambio y comprueba que sale el área o el volumen conocido de la región. Si sale otra cosa, el jacobiano está mal.

Ejemplo 9.20 (Dos cambios de variable con el jacobiano explícito)**. (a) Polares.** Calculemos la masa de un disco de radio 2 cuya densidad en cada punto es la distancia al centro, $\rho(x, y) = \sqrt{x^2 + y^2}$. En polares el integrando es r y el elemento de área es $r\,\mathrm{d}r\,\mathrm{d}\theta$, de modo que aparecen *dos* factores r:

$$m = \int_0^{2\pi}\int_0^2 r \cdot r\,\mathrm{d}r\,\mathrm{d}\theta = \int_0^{2\pi}\Big[\frac{r^3}{3}\Big]_0^2 \mathrm{d}\theta = 2\pi \cdot \frac{8}{3} = \frac{16\pi}{3} \approx 16{,}755.$$

Control del paso 7: la densidad media sobre el disco es algo entre 0 y 2, y el área es $4\pi \approx 12{,}57$; el resultado 16,76 corresponde a una densidad media de $4/3 \approx 1{,}33$, perfectamente razonable.

(b) Un cambio lineal inventado. Sea P el paralelogramo definido por $1 \le x+y \le 2$ y $0 \le x-2y \le 3$, y queremos su área. Poniendo $u = x+y$ y $v = x-2y$, el paso 5 aconseja calcular el jacobiano en esta dirección:

$$\frac{\partial(u,v)}{\partial(x,y)} = \begin{vmatrix} 1 & 1 \\ 1 & -2 \end{vmatrix} = -2-1 = -3,$$

luego $\dfrac{\partial(x,y)}{\partial(u,v)} = -\dfrac{1}{3}$ y $\mathrm{d}x\,\mathrm{d}y = \frac{1}{3}\,\mathrm{d}u\,\mathrm{d}v$. La región se convierte en el rectángulo $[1,2]\times[0,3]$, y

$$\mathrm{Area}(P) = \int_0^3 \int_1^2 \frac{1}{3}\,\mathrm{d}u\,\mathrm{d}v = \frac{1}{3}\cdot 1 \cdot 3 = 1.$$

Un paralelogramo de área exactamente 1, obtenido sin dibujar nada: el cambio de variable convirtió una región oblicua en un rectángulo, y el jacobiano se encargó de la contabilidad.

Pausa para pensar

Vuelve un momento al ejemplo 9.11 y compáralo con lo que acabas de hacer. El jacobiano mide cómo se deforma el área bajo un cambio de coordenadas, y para calcularlo hacen falta las cuatro parciales; el teorema del cambio de variable, sin embargo, exige que la aplicación sea de clase C^1, no solo que tenga parciales. Ahora sabes por qué la distinción no es una manía de los libros: una aplicación puede tener las cuatro parciales en un punto y no ser siquiera continua ahí, con lo que hablar de "cómo deforma el área cerca de ese punto" no tendría sentido alguno. Todas las hipótesis de regularidad de este capítulo existen para excluir ese tipo de patología.

9.3. Diferenciabilidad

Pausa para pensar

Antes de seguir, un aviso que conviene tener presente todo el capítulo: en varias variables, *tener todas las derivadas parciales no basta* para ser diferenciable. Las parciales solo miran lo que ocurre a lo largo de los ejes, y una función puede portarse bien en las dos direcciones cardinales y ser un desastre en las diagonales. Diferenciable significa algo más fuerte: que existe un plano tangente que aproxima la función en *todas* las direcciones a la vez, con error despreciable. La condición práctica que casi siempre se usa es que las parciales existan y sean continuas; eso sí implica diferenciabilidad, y es lo que se cumple en toda función razonable que encontrarás.

Idea clave

Intuición primero. Ser diferenciable es *verse plana bajo el microscopio*: al hacer zoom en el punto, la gráfica se confunde con un plano (el tangente), y la diferencial es ese

plano. Las derivadas parciales solas prometen menos: garantizan planitud solo a lo largo de las dos avenidas coordenadas, y una sábana puede estar lisa sobre las dos avenidas y arrugada en las diagonales (los contraejemplos de esta sección son exactamente eso). Por eso el teorema fuerte pide parciales *continuas*: es la hipótesis que impide la arruga entre avenidas.

Definición 9.21. La *derivada parcial* $\frac{\partial f}{\partial x_i}(\mathbf{a})$ es la derivada ordinaria congelando las demás variables. La *derivada direccional* en la dirección unitaria $\mathbf{u}$ es $D_{\mathbf{u}}f(\mathbf{a}) = \lim_{h\to 0} \frac{f(\mathbf{a}+h\mathbf{u})-f(\mathbf{a})}{h}$.

Definición 9.22 (Diferenciabilidad)**.** $f:\mathbb{R}^n\to\mathbb{R}$ es *diferenciable* en $\mathbf{a}$ si

$$f(\mathbf{a}+\mathbf{h}) = f(\mathbf{a}) + \nabla f(\mathbf{a})\cdot\mathbf{h} + o(\|\mathbf{h}\|), \qquad \nabla f = \Big(\frac{\partial f}{\partial x_1},\dots,\frac{\partial f}{\partial x_n}\Big).$$

Para $\mathbf{F}:\mathbb{R}^n\to\mathbb{R}^m$: $\mathbf{F}(\mathbf{a}+\mathbf{h}) = \mathbf{F}(\mathbf{a}) + D\mathbf{F}(\mathbf{a})\mathbf{h} + o(\|\mathbf{h}\|)$, con la *jacobiana* $D\mathbf{F} = \big(\frac{\partial F_i}{\partial x_j}\big) \in M_{m\times n}$.

Teorema 9.23 (Jerarquía de regularidad)**.** Parciales continuas $\Longrightarrow$ diferenciable $\Longrightarrow$ continua y con todas las derivadas direccionales*, con $D_{\mathbf{u}}f = \nabla f\cdot\mathbf{u}$. Ninguna flecha se invierte: $f = \frac{x^2y}{x^2+y^2}$ (con $f(0,0)=0$) tiene todas las direccionales en el origen pero no es diferenciable ahí.*

Geometría del gradiente: de $D_{\mathbf{u}}f = \nabla f\cdot\mathbf{u} = \|\nabla f\|\cos\theta$ se lee todo:

- ∇f apunta en la dirección de **máximo crecimiento**, con tasa $\|\nabla f\|$;
- $-\nabla f$ es la dirección de máximo descenso (de ahí el algoritmo de *descenso por gradiente*);
- ∇f es **ortogonal a los conjuntos de nivel** $\{f=c\}$: moverse por el nivel no cambia f, así que la direccional tangente es nula.

Plano tangente: para una superficie $z=f(x,y)$ en (a,b):

$$z = f(a,b) + f_x(a,b)(x-a) + f_y(a,b)(y-b);$$

para una superficie de nivel $F(x,y,z)=c$ en $\mathbf{p}$: $\nabla F(\mathbf{p})\cdot(\mathbf{x}-\mathbf{p}) = 0$.

Teorema 9.24 (Regla de la cadena)**.** *Si $\mathbf{g}:\mathbb{R}^k\to\mathbb{R}^n$ y $\mathbf{F}:\mathbb{R}^n\to\mathbb{R}^m$ son diferenciables,*

$$D(\mathbf{F}\circ\mathbf{g})(\mathbf{a}) = D\mathbf{F}(\mathbf{g}(\mathbf{a}))\,D\mathbf{g}(\mathbf{a}):$$

las jacobianas se multiplican. *Casos de uso diario: para $z=f(x(t),y(t))$, $\frac{\mathrm{d}z}{\mathrm{d}t} = f_x\dot{x} + f_y\dot{y}$; para $z=f(u(s,t),v(s,t))$, $\frac{\partial z}{\partial s} = f_u u_s + f_v v_s$.*

Ejemplo 9.25 (El laplaciano en polares)**.** Con $x = r\cos\theta$, $y = r\sin\theta$, la regla de la cadena (aplicada dos veces, con paciencia) da la fórmula

$$\Delta f = f_{xx} + f_{yy} = f_{rr} + \frac{1}{r}f_r + \frac{1}{r^2}f_{\theta\theta},$$

esencial para resolver EDP en discos y anillos, y un anticipo del laplaciano covariante del capítulo 36.

Teorema 9.26 (Clairaut–Schwarz)**.** *Si las segundas parciales son continuas, $f_{xy} = f_{yx}$. (Sin continuidad puede fallar.) En consecuencia la matriz hessiana es simétrica y le aplica el teorema espectral.*

Ejemplo 9.27 (Cuándo las cruzadas coinciden y cuándo no)**. (a) Un caso normal.** Con $f = x^3y^2$: $f_x = 3x^2y^2$, $f_{xy} = 6x^2y$; y $f_y = 2x^3y$, $f_{yx} = 6x^2y$. Coinciden, como debe ocurrir con cualquier polinomio.

(b) Con exponenciales. Con $f = \mathrm{e}^{xy}$: $f_{xy} = f_{yx} = (1+xy)\mathrm{e}^{xy}$. Evaluado en $(1,1)$ da $2\mathrm{e} = 5{,}436564$ por los dos caminos.

(c) El contraejemplo de Peano. Sea $f(x,y) = xy\dfrac{x^2-y^2}{x^2+y^2}$ con $f(0,0) = 0$. Entonces $f_{xy}(0,0) = -1$ y $f_{yx}(0,0) = 1$: no coinciden. La razón es que las segundas parciales existen pero no son continuas en el origen, que es justo la hipótesis del teorema.

(d) Por qué importa. La igualdad de cruzadas es lo que hace que la matriz hessiana sea simétrica, y por tanto diagonalizable con valores propios reales. Sin ella, el criterio de clasificación de puntos críticos por el signo de los valores propios se vendría abajo. También es lo que garantiza que un campo con potencial tenga rotacional nulo.

9.3.1. Taylor en varias variables

Teorema 9.28 (Taylor de orden 2)**.** *Si f es C^2 cerca de $\mathbf{a}$:*

$$f(\mathbf{a}+\mathbf{h}) = f(\mathbf{a}) + \nabla f(\mathbf{a})\cdot\mathbf{h} + \frac{1}{2}\,\mathbf{h}^{\mathsf{T}} H_f(\mathbf{a})\,\mathbf{h} + o(\|\mathbf{h}\|^2),$$

donde $H_f = \left(\frac{\partial^2 f}{\partial x_i \partial x_j}\right)$ es la hessiana*. El término cuadrático es una forma cuadrática del capítulo 7: toda la teoría de clasificación se aplica.*

Regla: aplicar la regla de la cadena con el árbol de dependencias

En una variable la regla de la cadena es una multiplicación. En varias variables es una *suma* de multiplicaciones, y el error más común es olvidar uno de los sumandos. El árbol de dependencias existe justamente para que eso no pase.

1. **Dibuja el árbol antes de derivar.** Pon arriba la variable dependiente, debajo las variables intermedias de las que depende directamente, y debajo de cada una las variables finales. Cada flecha del árbol es una derivada parcial.

2. **Cuenta los caminos.** La derivada de la variable de arriba respecto de una variable de abajo es la *suma sobre todos los caminos* que las conectan, y cada camino aporta el *producto* de las derivadas que hay a lo largo de él. Tantos sumandos como caminos: ni uno más, ni uno menos.

3. **Distingue $\frac{\mathrm{d}}{\mathrm{d}t}$ de $\frac{\partial}{\partial t}$.** Si abajo hay una sola variable final, el resultado es una derivada total y se escribe con d. Si hay varias, es parcial y se escribe con ∂. Escribir d donde va ∂ suele ser el síntoma de haber perdido un camino.

4. **El caso estándar de dos intermedias.** Si $z = f(x, y)$ con $x = x(t)$, $y = y(t)$, hay dos caminos y por tanto
$$\frac{\mathrm{d}z}{\mathrm{d}t} = \frac{\partial f}{\partial x}\frac{\mathrm{d}x}{\mathrm{d}t} + \frac{\partial f}{\partial y}\frac{\mathrm{d}y}{\mathrm{d}t}.$$
Si en cambio $x = x(u, v)$ e $y = y(u, v)$, se repite lo mismo una vez por cada variable final.

5. **En forma matricial es una multiplicación.** Todo lo anterior es $D(f \circ g)(p) = Df(g(p)) \cdot Dg(p)$: el jacobiano de la composición es el producto de los jacobianos. Comprobar las dimensiones de esas matrices es la forma más rápida de detectar un error.

6. **Para segundas derivadas, vuelve a aplicar la regla a cada factor.** $\frac{\partial}{\partial t}(\frac{\partial f}{\partial x})$ *no* es $\frac{\partial^2 f}{\partial x^2}\frac{\mathrm{d}x}{\mathrm{d}t}$ a secas: $\frac{\partial f}{\partial x}$ vuelve a ser una función de x e y, así que hay que aplicarle otra vez el paso 4 completo. Aquí es donde se pierden más términos.

7. **Verifica sustituyendo.** Si el problema es explícito, sustituye y deriva directamente: los dos resultados deben coincidir. Si es implícito, prueba con una función homogénea, donde el teorema de Euler da una respuesta conocida de antemano.

Ejemplo 9.29 (Seis cadenas, de la más directa a la que engaña)**.** Cada apartado añade una rama al árbol.

(a) Una intermedia doble, una final. Sea $z = x^2 y$ con $x = \cos t$, $y = \sin t$. El árbol tiene dos caminos de z a t, y
$$\frac{\mathrm{d}z}{\mathrm{d}t} = 2xy \cdot (-\sin t) + x^2 \cdot \cos t = -2\cos t \sin^2 t + \cos^3 t = (1 - 3\sin^2 t)\cos t.$$
Verificación por sustitución (paso 7): $z = \cos^2 t \sin t$, y derivando directamente sale lo mismo. En $t = \pi/4$ ambos dan $-\sqrt{2}/4 \approx -0{,}35355$.

(b) Dos intermedias y dos finales: la aplicación conforme. Sea $w = f(u, v)$ con $u = x^2 - y^2$, $v = 2xy$. Cuatro caminos en total, dos por cada variable final:
$$\frac{\partial w}{\partial x} = 2x\, f_u + 2y\, f_v, \qquad \frac{\partial w}{\partial y} = -2y\, f_u + 2x\, f_v.$$
Elevando al cuadrado y sumando, los términos cruzados $\pm 8xy\, f_u f_v$ se cancelan y queda
$$\Big(\frac{\partial w}{\partial x}\Big)^2 + \Big(\frac{\partial w}{\partial y}\Big)^2 = 4(x^2 + y^2)(f_u^2 + f_v^2).$$
El factor $4(x^2 + y^2) = |2z|^2$ es exactamente el módulo al cuadrado de la derivada de $z \mapsto z^2$: el cambio de variable estira el gradiente por un factor que no depende de la dirección. Eso es lo que significa que la aplicación sea conforme (capítulo 30).

(c) Segundas derivadas: el laplaciano en polares. Con $x = r\cos\theta$, $y = r\sin\theta$, aplicando el paso 6 dos veces y agrupando se llega a
$$\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} = \frac{\partial^2 u}{\partial r^2} + \frac{1}{r}\frac{\partial u}{\partial r} + \frac{1}{r^2}\frac{\partial^2 u}{\partial \theta^2}.$$

Los dos términos con $1/r$ y $1/r^2$ no salen de ninguna parte «mágica»: son precisamente los que aparecen al derivar $\frac{\partial u}{\partial r}$ respecto de x recordando que r y θ dependen ambos de x. Comprobación rápida con $u = \ln r$: el miembro derecho da $-\frac{1}{r^2} + \frac{1}{r}\cdot\frac{1}{r} + 0 = 0$, y en efecto $\ln\sqrt{x^2+y^2}$ es armónica fuera del origen.

Cuidado

Si al calcular (c) obtienes solo $u_{rr} + \frac{1}{r^2}u_{\theta\theta}$, has aplicado el paso 6 a medias: derivaste u_r como si fuera una función de r solamente. Es el error más frecuente de todo el capítulo, y se detecta al instante probando con $u = \ln r$, que dejaría de ser armónica.

(d) Derivación implícita, y una identidad que parece falsa. Si $F(x,y,z) = 0$ define z como función de x e y, derivar la identidad respecto de x da $F_x + F_z\frac{\partial z}{\partial x} = 0$, es decir

$$\frac{\partial z}{\partial x} = -\frac{F_x}{F_z}.$$

Con $F = x^2+y^2+z^2-1$ sale $\frac{\partial z}{\partial x} = -x/z$, que es lo que se obtiene despejando $z = \sqrt{1-x^2-y^2}$. Pero ahora multiplica las tres derivadas cíclicas:

$$\frac{\partial x}{\partial y}\cdot\frac{\partial y}{\partial z}\cdot\frac{\partial z}{\partial x} = \Big(-\frac{F_y}{F_x}\Big)\Big(-\frac{F_z}{F_y}\Big)\Big(-\frac{F_x}{F_z}\Big) = -1.$$

No $+1$. Las «fracciones» $\partial z/\partial x$ no se cancelan como fracciones, porque cada una se calcula manteniendo fija una variable *distinta*. Esta identidad, con el signo menos, es la que en termodinámica relaciona presión, volumen y temperatura.

(e) El teorema de Euler para funciones homogéneas. Si $f(tx,ty) = t^n f(x,y)$ para todo $t>0$, deriva ambos lados respecto de t usando el paso 4 y evalúa en $t=1$:

$$x\,f_x + y\,f_y = n\,f.$$

Con $f = x^3+xy^2$, que es homogénea de grado 3: $x(3x^2+y^2)+y(2xy) = 3x^3+3xy^2 = 3f$. ✓Este resultado es el que hace funcionar el análisis dimensional y, en economía, los rendimientos constantes a escala.

(f) Un cambio de variable que resuelve una ecuación en derivadas parciales. Para la ecuación de ondas $u_{tt} = c^2u_{xx}$, introduce $\xi = x-ct$, $\eta = x+ct$. Aplicando el paso 4 dos veces:

$$u_{tt} = c^2(u_{\xi\xi} - 2u_{\xi\eta} + u_{\eta\eta}), \qquad u_{xx} = u_{\xi\xi} + 2u_{\xi\eta} + u_{\eta\eta},$$

de modo que la ecuación se reduce a $u_{\xi\eta} = 0$. Integrando dos veces, $u = F(\xi) + G(\eta)$, es decir

$$u(x,t) = F(x-ct) + G(x+ct)\,:$$

una onda que viaja a la derecha más otra que viaja a la izquierda, ambas a velocidad c. La solución general de la ecuación de ondas, obtenida únicamente con la regla de la cadena.

Pausa para pensar

Vuelve al paso 5 y mira (f) desde ahí. El cambio $(x,t)\mapsto(\xi,\eta)$ tiene jacobiano constante $\left(\begin{smallmatrix}1 & -c\\ 1 & c\end{smallmatrix}\right)$, y lo que hemos hecho es conjugar el operador $\partial_t^2-c^2\partial_x^2$ por esa matriz hasta convertirlo en $\partial_\xi\partial_\eta$. Es la misma idea que diagonalizar una matriz (capítulo 7), aplicada a operadores diferenciales en lugar de a vectores.

Idea clave

La factorización $\partial_t^2-c^2\partial_x^2=(\partial_t-c\partial_x)(\partial_t+c\partial_x)$ es la clave de (f), y explica por qué la ecuación de ondas se comporta tan distinto de la del calor $u_t=ku_{xx}$: aquella se factoriza sobre los reales y esta no. La clasificación de las ecuaciones en derivadas parciales de segundo orden en hiperbólicas, parabólicas y elípticas es exactamente la clasificación de una forma cuadrática por su signatura (capítulo 7).

9.4. Optimización

9.4.1. Extremos libres

Teorema 9.30. *Si f es diferenciable y tiene extremo local en un punto interior* $\mathbf{a}$*, entonces* $\nabla f(\mathbf{a})=\mathbf{0}$ *(punto* crítico*). Clasificación por la hessiana $H=H_f(\mathbf{a})$:*

- *H definida positiva (todos los $\lambda_i>0$) $\Rightarrow$ mínimo local;*
- *H definida negativa $\Rightarrow$ máximo local;*
- *H indefinida (signos mezclados) $\Rightarrow$* punto de silla*;*
- *H singular: el criterio no decide; hay que examinar f directamente.*

En dos variables, con $D=f_{xx}f_{yy}-f_{xy}^2$: $D>0$ y $f_{xx}>0$ mínimo; $D>0$ y $f_{xx}<0$ máximo; $D<0$ silla; $D=0$ sin decisión.

Ejemplo 9.31 (Puntos críticos y sus trampas)**. (a) Un mínimo.** $f=x^2+y^2$ tiene $\nabla f=(2x,2y)$, que se anula sólo en el origen, donde hay mínimo. La hessiana es $2I$, definida positiva.

(b) Una silla. $f=x^2-y^2$ también tiene su único punto crítico en el origen, y allí no hay ni máximo ni mínimo: subiendo por el eje x crece y por el eje y decrece. El gradiente nulo es necesario, no suficiente.

(c) Cuando el criterio no decide. $f=x^4+y^4$ y $f=x^3$ tienen ambas hessiana nula en el origen; la primera tiene mínimo y la segunda no tiene extremo. Con hessiana degenerada hay que mirar términos de orden superior.

(d) En el borde no vale. $f=x$ en $[0,1]^2$ alcanza máximo y mínimo en la frontera, donde el gradiente vale $(1,0)$ y no se anula nunca. El teorema exige punto *interior*, y por eso los problemas con restricciones necesitan multiplicadores de Lagrange.

Idea clave

La demostración es Taylor: cerca del crítico, $f(\mathbf{a}+\mathbf{h}) - f(\mathbf{a}) \approx \frac{1}{2}\mathbf{h}^{\mathsf{T}} H \mathbf{h}$, y el signo de esa forma cuadrática (clasificada con los valores propios, capítulo 7) decide si todas las direcciones suben, todas bajan, o hay de ambas.

Para *extremos absolutos* en una región compacta: comparar los críticos interiores con el comportamiento en la frontera (parametrizándola o con Lagrange). Weierstrass garantiza que existen.

Regla: clasificar los puntos críticos de una función de dos o más variables

1. **Hallar los críticos resolviendo $\nabla f = \mathbf{0}$.** Es un sistema *no lineal*: hay que resolverlo con cuidado y, sobre todo, encontrar *todas* las soluciones. Truco habitual: despejar una variable en una ecuación y sustituirla en la otra, lo que deja una sola ecuación polinómica.
2. **Calcular la hessiana, una sola vez, en forma simbólica.** $H = \begin{pmatrix} f_{xx} & f_{xy} \\ f_{xy} & f_{yy} \end{pmatrix}$. Es simétrica por el teorema de Schwarz; si sale asimétrica, hay un error de derivación.
3. **Evaluar H en cada crítico por separado** y calcular $D = \det H = f_{xx} f_{yy} - f_{xy}^2$.
4. **Leer el resultado con la tabla de dos variables:** $D > 0$ y $f_{xx} > 0$ mínimo; $D > 0$ y $f_{xx} < 0$ máximo; $D < 0$ silla; $D = 0$ el criterio calla.
5. **Con más de dos variables, usar los menores principales líderes** (criterio de Sylvester): todos positivos $\Rightarrow$ mínimo; alternando $-, +, -, \ldots$ empezando en negativo $\Rightarrow$ máximo; cualquier otro patrón con todos no nulos $\Rightarrow$ silla. Alternativa equivalente: calcular los valores propios de H y mirar sus signos.
6. **Si $D = 0$, atacar la función a mano.** Se estudia f sobre rectas y curvas que pasen por el punto, exactamente igual que con los límites. Basta con encontrar una dirección donde suba y otra donde baje para concluir silla.
7. **Para extremos absolutos en un compacto, la frontera es obligatoria.** Weierstrass garantiza que existen máximo y mínimo, pero pueden estar en el borde, donde ∇f no tiene por qué anularse. Se comparan los valores en los críticos interiores con los del borde (parametrizado o vía Lagrange).

Control gratuito: la traza de H es $f_{xx} + f_{yy}$ y debe coincidir con la suma de los valores propios; el determinante, con su producto.

Ejemplo 9.32 (Tres funciones clasificadas)**. (a) La silla del mono domesticada.** $f(x, y) = x^3 + y^3 - 3xy$. El gradiente es $(3x^2 - 3y,\ 3y^2 - 3x)$; igualando a cero, $y = x^2$ y $x = y^2$, luego

$x = x^4$, es decir $x(x^3-1)=0$ y $x=0$ o $x=1$. Críticos: $(0,0)$ y $(1,1)$. La hessiana es

$$H = \begin{pmatrix} 6x & -3 \\ -3 & 6y \end{pmatrix}.$$

En $(0,0)$: $H = \begin{pmatrix} 0 & -3 \\ -3 & 0 \end{pmatrix}$ y $D=-9<0$, *silla*. En $(1,1)$: $H = \begin{pmatrix} 6 & -3 \\ -3 & 6 \end{pmatrix}$, $D = 36-9 = 27>0$ con $f_{xx}=6>0$, *mínimo local*, de valor $f(1,1)=1+1-3=-1$. Nótese que no es mínimo absoluto: por el eje x negativo, $f=x^3\to-\infty$.

(b) Tres críticos, dos mínimos. $f(x,y)=x^4+y^4-4xy$. Gradiente: $(4x^3-4y,\ 4y^3-4x)$, luego $y=x^3$ y $x=y^3=x^9$, es decir $x(x^8-1)=0$: $x=0,1,-1$. Críticos $(0,0)$, $(1,1)$, $(-1,-1)$. Con $H = \begin{pmatrix} 12x^2 & -4 \\ -4 & 12y^2 \end{pmatrix}$: en el origen $D=0-16=-16<0$, silla; en $(1,1)$ y en $(-1,-1)$, $D=144-16=128>0$ con $f_{xx}=12>0$, *mínimos locales* de valor $1+1-4=-2$. La simetría $f(-x,-y)=f(x,y)$ explica por qué los dos mínimos son gemelos.

(c) Tres variables. $f(x,y,z)=x^2+y^2+z^2+xy$. El único crítico es el origen y

$$H = \begin{pmatrix} 2 & 1 & 0 \\ 1 & 2 & 0 \\ 0 & 0 & 2 \end{pmatrix}.$$

Menores líderes: $2>0$, $\det\left(\begin{smallmatrix} 2 & 1 \\ 1 & 2 \end{smallmatrix}\right)=3>0$, $\det H = 6>0$. Todos positivos: *mínimo*. Control por valores propios: son $1, 2, 3$, cuya suma 6 es la traza y cuyo producto 6 es el determinante. Todos positivos, misma conclusión.

Cuidado

Cuando $D=0$ el criterio no dice "es degenerado, no se puede saber": dice "con segundo orden no basta, mira el cuarto". Las tres funciones x^2+y^4, x^2-y^4 y x^2+y^3 tienen todas la misma hessiana en el origen, $\left(\begin{smallmatrix} 2 & 0 \\ 0 & 0 \end{smallmatrix}\right)$, con $D=0$; y sin embargo la primera tiene un mínimo, y las otras dos una silla (en la segunda, f baja por el eje y; en la tercera también, porque y^3 cambia de signo). El mismo segundo orden, tres comportamientos distintos.

9.4.2. Extremos condicionados: multiplicadores de Lagrange

Idea clave

Intuición primero. Caminas por un sendero (la restricción) en una ladera (la función). Mientras el sendero *cruce* las curvas de nivel, sigues subiendo o bajando: no estás en el óptimo. El máximo llega exactamente cuando el sendero se vuelve *tangente* a una curva de nivel: ahí avanzar ya no cambia la altura a primer orden. Tangencia de niveles es paralelismo de gradientes: eso es $\nabla f = \lambda \nabla g$, todo el método en una imagen. Y el multiplicador no es un fantasma auxiliar: λ mide cuánto mejora el óptimo si aflojas la restricción una unidad —el "precio sombra" con que los economistas cobran los recursos escasos—.

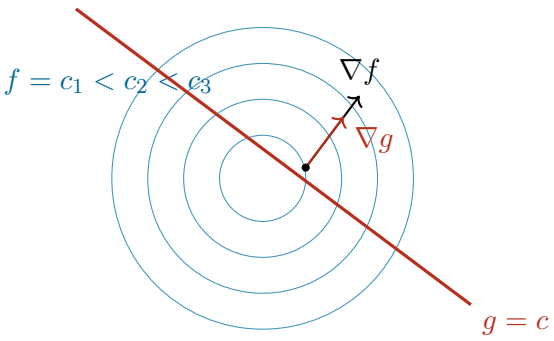


Figura 9.1: Multiplicadores de Lagrange: el óptimo restringido ocurre donde el sendero $g = c$ es *tangente* a una curva de nivel de f: allí los dos gradientes son paralelos.

Teorema 9.33 (Lagrange)**.** *Sea* $\mathbf{a}$ *extremo de* f *sujeto a* $g = c$*, con* $\nabla g(\mathbf{a}) \neq \mathbf{0}$*. Entonces existe* λ *(el* multiplicador*) con*

$$\nabla f(\mathbf{a}) = \lambda \nabla g(\mathbf{a}).$$

Con dos restricciones $g = c_1, h = c_2$ *(gradientes independientes):* $\nabla f = \lambda \nabla g + \mu \nabla h$*.*

Idea geométrica. Sobre la superficie de restricción, en un extremo la derivada direccional de f a lo largo de cualquier dirección *tangente* debe anularse: ∇f es ortogonal al tangente. Pero el ortogonal al tangente de $\{g = c\}$ está generado por ∇g: luego $\nabla f \parallel \nabla g$. □

Observación 9.34. El multiplicador tiene significado: si el óptimo depende del nivel c, entonces $\frac{\mathrm{d}}{\mathrm{d}c} f_{\text{ópt}}(c) = \lambda$. En economía, λ es el "precio sombra" de la restricción; en mecánica, la fuerza de ligadura.

Ejemplo 9.35 (Desigualdad AM–GM por Lagrange)**.** Maximicemos $f = x_1 x_2 \cdots x_n$ sujeto a $x_1 + \cdots + x_n = s$ $(x_i > 0)$. Lagrange: $\frac{f}{x_i} = \lambda$ para todo i, luego todos los x_i son iguales a $\frac{s}{n}$. El máximo es $(\frac{s}{n})^n$, es decir, para cualesquiera positivos:

$$x_1 \cdots x_n \leq \left(\frac{x_1 + \cdots + x_n}{n}\right)^n \iff \sqrt[n]{x_1 \cdots x_n} \leq \frac{x_1 + \cdots + x_n}{n}.$$

La media geométrica nunca supera la aritmética, con igualdad solo si todos son iguales.

Regla: resolver un problema de extremos condicionados con multiplicadores

1. **Separar bien los papeles.** Escribir explícitamente cuál es la función objetivo f (lo que se optimiza) y cuál la restricción, siempre en la forma $g = c$ con todo lo variable en g. Confundirlas es el error más caro del tema.

2. **Plantear el sistema completo.** $\nabla f = \lambda \nabla g$ junto con $g = c$. Son $n + 1$ ecuaciones con $n+1$ incógnitas $(x_1, \ldots, x_n, \lambda)$. La restricción *es una de las ecuaciones*: olvidarla deja el sistema indeterminado.

3. **Resolver sin dividir a lo loco.** La técnica que casi siempre funciona: dividir una ecuación entre otra para eliminar λ de un plumazo. Pero antes hay que tratar aparte el caso en que el denominador se anula, porque ahí suelen esconderse soluciones.

4. **Con dos restricciones,** $\nabla f = \lambda\nabla g + \mu\nabla h$, más las dos restricciones: $n+2$ ecuaciones y $n+2$ incógnitas. Geométricamente, ∇f debe estar en el plano generado por los dos gradientes de restricción.

5. **Comprobar la hipótesis de regularidad.** El teorema exige $\nabla g \neq \mathbf{0}$ (y, con varias restricciones, gradientes linealmente independientes). Los puntos donde falla son *puntos singulares* y hay que examinarlos aparte, a mano: Lagrange no los ve.

6. **Evaluar f en todos los candidatos y comparar.** Lagrange produce *candidatos*, no clasificaciones. Si la restricción define un conjunto compacto (una esfera, una elipse), Weierstrass garantiza que el mayor valor es el máximo y el menor el mínimo, y no hace falta nada más. Si no es compacto (un plano, una hipérbola), hay que estudiar además el comportamiento en el infinito.

7. **Interpretar λ como control.** Se cumple $\frac{\mathrm{d}f_{\text{óptimo}}}{\mathrm{d}c} = \lambda$: si al aflojar la restricción el óptimo debería mejorar y el λ obtenido tiene el signo contrario, hay un error de signos en alguna parte.

Ejemplo 9.36 (Del caso de libro al de dos restricciones)**. (a) Sobre la circunferencia.** Extremos de $f = x^2 + 2y^2$ sujeto a $x^2 + y^2 = 1$. El sistema es $2x = 2\lambda x$, $4y = 2\lambda y$, $x^2 + y^2 = 1$. La primera dice $x(1-\lambda) = 0$ y la segunda $y(2-\lambda) = 0$; como x e y no pueden anularse a la vez (la restricción lo impide), hay dos ramas:

- $\lambda = 1$: entonces $y = 0$ y $x = \pm 1$, con $f = 1$;
- $\lambda = 2$: entonces $x = 0$ e $y = \pm 1$, con $f = 2$.

La circunferencia es compacta, así que el mínimo es 1 y el máximo 2. Y hay un premio conceptual: $f = \mathbf{x}^{\mathsf{T}} A\mathbf{x}$ con $A = \left(\begin{smallmatrix} 1 & 0 \\ 0 & 2 \end{smallmatrix}\right)$, y los multiplicadores que han salido, 1 y 2, son exactamente los valores propios de A. No es casualidad: $\nabla f = 2A\mathbf{x}$ y $\nabla g = 2\mathbf{x}$, de modo que la condición de Lagrange *es* la ecuación $A\mathbf{x} = \lambda\mathbf{x}$. Optimizar una forma cuadrática sobre la esfera y diagonalizar una matriz simétrica son el mismo problema.

(b) Un producto con suma fija. Maximizar $f = xyz$ sujeto a $x+y+z = 3$ con $x, y, z > 0$. Las ecuaciones son $yz = \lambda$, $xz = \lambda$, $xy = \lambda$. Restando las dos primeras, $z(y-x) = 0$, y como $z > 0$ sale $x = y$; igual con las otras, $y = z$. Entonces $3x = 3$, es decir $x = y = z = 1$ y $f = 1$: la caja de volumen máximo con perímetro de aristas fijo es el cubo. Aquí la restricción *no* es compacta (es un triángulo abierto), pero $f \to 0$ al acercarse a cualquier borde, así que el único candidato interior es el máximo.

(c) Dos restricciones a la vez. Extremos de $f = x + y + z$ sujeto a $x^2 + y^2 = 2$ y $x + z = 1$. Con $g = x^2 + y^2$ y $h = x + z$:

$$(1, 1, 1) = \lambda(2x, 2y, 0) + \mu(1, 0, 1).$$

La tercera componente da $\mu = 1$ de inmediato. La primera queda $1 = 2\lambda x + 1$, luego $\lambda x = 0$; y la segunda, $1 = 2\lambda y$, obliga a $\lambda \neq 0$, de modo que $x = 0$. La primera restricción da $y = \pm\sqrt{2}$ y la segunda $z = 1$. Los dos candidatos son $(0, \pm\sqrt{2}, 1)$, con

$$f = 1 + \sqrt{2} \approx 2{,}414 \qquad \text{y} \qquad f = 1 - \sqrt{2} \approx -0{,}414.$$

La intersección de un cilindro con un plano es una elipse, que es compacta: máximo y mínimo, respectivamente.

Cuidado

Lagrange encuentra puntos donde los gradientes son paralelos, y eso incluye mínimos, máximos *y* puntos que no son ni una cosa ni la otra. Nunca hay que declarar "máximo" solo porque un candidato dé el valor más alto de la lista, salvo que se haya comprobado que la restricción define un compacto. En $f = x$ sujeto a $xy = 1$, el sistema de Lagrange da $(1, 0) = \lambda(y, x)$, es decir $\lambda x = 0$ con $\lambda y = 1$; de ahí $\lambda \neq 0$ y $x = 0$, incompatible con $xy = 1$: no hay ningún candidato, y en efecto f no está acotada sobre la hipérbola. La ausencia de soluciones también es información.

Regla: gradiente, derivada direccional, plano tangente y aproximación lineal

Estos cuatro objetos son en realidad uno solo visto desde cuatro ángulos: la derivada de f en un punto es una aplicación lineal, y el gradiente es su vector de coeficientes.

1. **Calcula el gradiente y evalúa.** $\operatorname{grad} f = (f_x, f_y, \ldots)$, y luego sustituye el punto. Un gradiente sin punto de evaluación es un campo vectorial; lo que hace falta para todo lo demás es un vector concreto.

2. **Normaliza la dirección antes de usarla.** La derivada direccional es $D_{\mathbf{u}} f(p) = \operatorname{grad} f(p) \cdot \mathbf{u}$ *con* $\|\mathbf{u}\| = 1$. Si te dan una dirección sin normalizar, divídela por su norma primero; si no, el número que salga no mide nada.

3. **Lee las tres consecuencias geométricas.** El crecimiento es máximo en la dirección de $\operatorname{grad} f$ y vale $\|\operatorname{grad} f\|$; es mínimo en la dirección opuesta y vale $-\|\operatorname{grad} f\|$; y es nulo en las direcciones perpendiculares al gradiente, que son las tangentes a la curva o superficie de nivel.

4. **Para una gráfica $z = f(x, y)$, el plano tangente es la aproximación lineal.** En $p = (a, b)$,
$$z = f(a, b) + f_x(a, b)(x - a) + f_y(a, b)(y - b),$$
y la recta normal tiene dirección $(f_x, f_y, -1)$.

5. **Para una superficie de nivel $g(x, y, z) = k$, usa el gradiente de g.** El plano tangente es $\operatorname{grad} g(p) \cdot (\mathbf{x} - p) = 0$ y la recta normal tiene dirección $\operatorname{grad} g(p)$. Es el mismo enunciado del paso 4 aplicado a $g = z - f(x, y)$.

6. **Usa la aproximación lineal para estimar y para propagar errores.** $f(p+\mathbf{h}) \approx f(p) + \operatorname{grad} f(p) \cdot \mathbf{h}$, y por tanto $|\Delta f| \lesssim \|\operatorname{grad} f(p)\|\, \|\mathbf{h}\|$. El error de la aproximación es de orden $\|\mathbf{h}\|^2$.

7. **Verifica que la función sea diferenciable antes de creer en el gradiente.** Que existan las parciales no basta. Si f no es diferenciable en p, la fórmula $D_{\mathbf{u}} f = \operatorname{grad} f \cdot \mathbf{u}$ del paso 2 puede ser falsa aunque todas las derivadas direccionales existan.

Ejemplo 9.37 (Cinco cálculos geométricos, y una función que tiene todas las direccionales sin ser diferenciable)**.** Todo el trabajo se hace con productos escalares.

(a) Plano tangente a un paraboloide. Sea $f(x,y) = x^2 + y^2$ en $p = (1,2)$, donde $f(p) = 5$ y $\operatorname{grad} f = (2x, 2y) = (2,4)$. Por el paso 4,

$$z = 5 + 2(x-1) + 4(y-2) = 2x + 4y - 5.$$

Comprobación: en $(1,2)$ da $2 + 8 - 5 = 5$. ✓La recta normal tiene dirección $(2,4,-1)$ y ecuaciones paramétricas $(1+2t,\ 2+4t,\ 5-t)$.

(b) Derivada direccional y dirección óptima. En el mismo punto, la derivada en la dirección de $(3,4)$ exige normalizar: $\mathbf{u} = (3/5, 4/5)$, y

$$D_{\mathbf{u}} f(1,2) = (2,4) \cdot \Big(\frac{3}{5}, \frac{4}{5}\Big) = \frac{6}{5} + \frac{16}{5} = \frac{22}{5} = 4{,}4.$$

El crecimiento máximo se da en la dirección $\operatorname{grad} f / \|\operatorname{grad} f\| = (1,2)/\sqrt{5}$ y vale $\|\operatorname{grad} f\| = \sqrt{20} \approx 4{,}47214$. Y en la dirección $(2,-1)/\sqrt{5}$, perpendicular al gradiente, la derivada es cero: es la tangente a la circunferencia de nivel $x^2 + y^2 = 5$.

(c) Plano tangente a una superficie de nivel: el cono. Sea $g(x,y,z) = x^2 + y^2 - z^2$ y el punto $(3,4,5)$, que cumple $g = 9 + 16 - 25 = 0$. Entonces $\operatorname{grad} g = (2x, 2y, -2z) = (6, 8, -10)$ y el plano tangente es

$$6(x-3) + 8(y-4) - 10(z-5) = 0 \iff 3x + 4y - 5z = 0.$$

El plano pasa por el origen, como debe ser: en un cono, el plano tangente a lo largo de una generatriz contiene a esa generatriz entera. Nótese que en el vértice $(0,0,0)$ el gradiente se anula y no hay plano tangente, lo cual coincide con que ahí el cono tenga un pico.

(d) Aproximación lineal y estimación de errores. Sea $f(x,y) = \sqrt{x^2 + y^2}$ y $p = (3,4)$, donde $f(p) = 5$ y $\operatorname{grad} f = (\frac{x}{r}, \frac{y}{r}) = (0{,}6,\ 0{,}8)$. Para estimar $f(3{,}1,\ 3{,}9)$ tomamos $\mathbf{h} = (0{,}1,\ -0{,}1)$:

$$f(3{,}1,\ 3{,}9) \approx 5 + 0{,}6(0{,}1) + 0{,}8(-0{,}1) = 5 + 0{,}06 - 0{,}08 = 4{,}98.$$

El valor exacto es $\sqrt{9{,}61 + 15{,}21} = \sqrt{24{,}82} \approx 4{,}9819675$, con un error de $0{,}00197$, coherente con el orden $\|\mathbf{h}\|^2 = 0{,}02$ que anuncia el paso 6. Como $\|\operatorname{grad} f\| = 1$ aquí, un error de $0{,}1$ en los datos produce a lo sumo un error del mismo orden en el resultado: la función está bien condicionada en ese punto.

(e) El gradiente es perpendicular a las curvas de nivel. Sea $f = xy$ y el punto $(1,1)$ sobre la hipérbola $xy = 1$. Entonces $\operatorname{grad} f = (y, x) = (1,1)$. Parametrizando la hipérbola como $(t, 1/t)$, el vector tangente en $t = 1$ es $(1, -1/t^2) = (1,-1)$, y

$$(1,1) \cdot (1,-1) = 0. \checkmark$$

Esta perpendicularidad es lo que hace funcionar los multiplicadores de Lagrange: en un extremo condicionado, las curvas de nivel de la función y de la restricción son tangentes, luego sus gradientes son paralelos.

(f) Todas las derivadas direccionales existen, y la función no es diferenciable. Sea

$$f(x,y)=\begin{cases}\dfrac{x^3}{x^2+y^2} & (x,y)\neq(0,0),\\ 0 & (x,y)=(0,0).\end{cases}$$

Para una dirección unitaria $\mathbf{u}=(a,b)$, la definición da

$$D_{\mathbf{u}}f(0,0)=\lim_{t\to 0}\frac{f(ta,tb)}{t}=\lim_{t\to 0}\frac{1}{t}\cdot\frac{t^3a^3}{t^2}=a^3.$$

Existe para *toda* dirección. En particular $f_x(0,0)=1$ (con $a=1$) y $f_y(0,0)=0$ (con $a=0$), luego $\operatorname{grad} f(0,0)=(1,0)$. Pero si f fuera diferenciable, el paso 2 daría $D_{\mathbf{u}}f=\operatorname{grad} f\cdot\mathbf{u}=a$, y $a^3\neq a$: con $a=b=1/\sqrt{2}$ sale $a^3\approx 0{,}35355$ frente a $a\approx 0{,}70711$. Luego f no es diferenciable en el origen.

Cuidado

Este ejemplo desmonta la creencia de que «existen las parciales, luego el gradiente sirve». La diferenciabilidad es una condición estrictamente más fuerte: pide que la aproximación lineal funcione *uniformemente* en todas las direcciones, no solo que exista un límite a lo largo de cada recta. El criterio práctico es el de siempre: si las parciales son continuas en un entorno, entonces f es diferenciable ahí, y ese criterio cubre casi todos los casos que aparecen en la práctica.

Pausa para pensar

Fíjate en que la f del apartado (f) es continua en el origen (su valor está acotado por $|x|$) y sus derivadas direccionales son todas finitas, y aun así el plano $z=x$ no aproxima bien a la superficie. Dibuja la restricción de f a la recta $y=x$: vale $x/2$, con pendiente $1/2$, mientras que el «plano tangente» $z=x$ tiene pendiente $1/\sqrt{2}\approx 0{,}707$ en esa dirección. La discrepancia no se va al acercarse al origen.

Idea clave

La verdadera definición es esta: f es diferenciable en p si existe una aplicación *lineal* L con $f(p+\mathbf{h})=f(p)+L(\mathbf{h})+o(\|\mathbf{h}\|)$. El gradiente es solo la matriz de L en la base canónica. Formulado así, el concepto se traslada sin cambios a espacios de dimensión infinita (derivada de Fréchet, capítulo 32) y es lo que permite hablar de la derivada de un funcional respecto de una función entera, base del cálculo de variaciones.

9.5. Optimización numérica: descenso por gradiente y Newton

Los métodos del capítulo localizan los puntos críticos resolviendo $\nabla f=\mathbf{0}$ con lápiz y papel. Pero en cuanto la función tiene cinco variables, o un logaritmo mezclado con un polinomio, ese sistema no se resuelve en forma cerrada. La salida práctica es abandonar la fórmula exacta y *caminar* hacia el mínimo: partir de un punto cualquiera y repetir un paso que garantice

bajar. Los dos caminantes clásicos son el descenso por gradiente, que solo mira la primera derivada, y el método de Newton, que además usa la hessiana.

Definición 9.38 (Los dos iterantes)**.** Sea $f\colon \mathbb{R}^n \to \mathbb{R}$ de clase C^2. El *descenso por gradiente* con paso $\alpha > 0$ genera la sucesión

$$\mathbf{x}_{k+1} = \mathbf{x}_k - \alpha\, \nabla f(\mathbf{x}_k),$$

y el *método de Newton* para minimizar genera

$$\mathbf{x}_{k+1} = \mathbf{x}_k - \big[Hf(\mathbf{x}_k)\big]^{-1} \nabla f(\mathbf{x}_k),$$

donde Hf es la matriz hessiana. Para *resolver* un sistema $\mathbf{F}(\mathbf{x}) = \mathbf{0}$ con $\mathbf{F}\colon \mathbb{R}^n \to \mathbb{R}^n$, el mismo método toma la forma $\mathbf{x}_{k+1} = \mathbf{x}_k - J\mathbf{F}(\mathbf{x}_k)^{-1}\mathbf{F}(\mathbf{x}_k)$ con $J\mathbf{F}$ la matriz jacobiana.

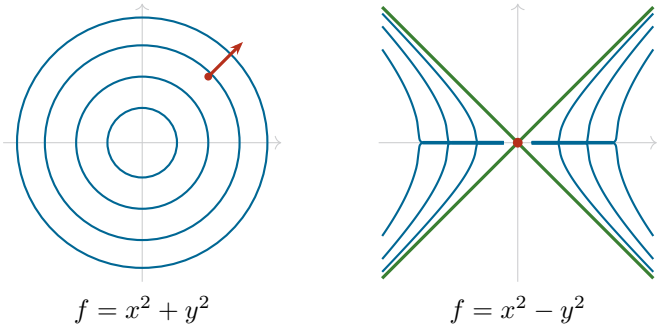


Figura 9.2: Curvas de nivel. A la izquierda, circunferencias concéntricas alrededor de un mínimo, y el gradiente apuntando radialmente hacia afuera, siempre perpendicular a la curva. A la derecha, un punto de silla: las curvas son hipérbolas y el nivel cero se cruza consigo mismo.

Ejemplo 9.39 (Descenso y Newton comparados)**. (a) Descenso en un cuenco.** Para $f = x^2$ con paso 0,1 desde $x_0 = 1$, la iteración $x_{k+1} = x_k - 0{,}2x_k = 0{,}8x_k$ da 0,8; 0,64; 0,512; 0,4096; 0,32768; 0,262144. Converge geométricamente, y la razón 0,8 depende del paso elegido.

(b) Newton en el mismo cuenco. Aquí $f'' = 2$, y la iteración $x_{k+1} = x_k - f'/f'' = x_k - x_k = 0$ llega al óptimo en *un* paso. Newton usa la curvatura; el gradiente sólo la pendiente.

(c) El precio de Newton. Cada paso exige resolver un sistema con la hessiana, de coste $O(n^3)$, mientras que un paso de gradiente cuesta $O(n)$. Con un millón de variables, Newton es impracticable y por eso el aprendizaje automático usa variantes del gradiente.

(d) El paso mal elegido. Para $f = x^2$ con paso 1,1 la iteración es $x_{k+1} = -1{,}2x_k$, que *diverge* oscilando. La condición de estabilidad es $\alpha < 2/L$ con L la constante de Lipschitz del gradiente; aquí $L = 2$, luego el paso debe ser menor que 1.

Idea clave

El gradiente apunta en la dirección de máximo crecimiento, así que $-\nabla f$ apunta cuesta abajo: el descenso por gradiente es un excursionista en la niebla que solo palpa la pendiente bajo sus pies y da un paso en el sentido más empinado hacia abajo. Newton es más ambicioso: en cada punto sustituye f por su polinomio de Taylor de grado dos —un paraboloide— y salta directamente al vértice de ese paraboloide. Cuando la función *es* un paraboloide, Newton acierta en un solo paso; cuando no lo es, acierta casi, y por eso converge tan rápido cerca del mínimo. El precio es que hay que calcular e invertir la hessiana, mientras que el gradiente solo pide n derivadas.

Regla: elegir el método y el paso, y saber cuándo parar

El fallo típico no es programar mal la fórmula, sino elegir un paso disparatado o no reconocer que el método ha dejado de converger.

1. **Escríbanse a mano ∇f y, si se va a usar Newton, Hf.** Un error de signo en una parcial es indetectable a ojo en la salida numérica: conviene comprobar las derivadas contra un cociente incremental $(f(\mathbf{x}+h\mathbf{e}_i)-f(\mathbf{x}))/h$ con $h=10^{-6}$.
2. **Estímese el rango admisible del paso.** Si f es cuadrática con hessiana de autovalores entre $m>0$ y L, el descenso converge si y solo si $0<\alpha<2/L$, y el paso óptimo es $\alpha=2/(L+m)$. Para funciones generales, L es una cota de la norma de la hessiana en la región de trabajo.
3. **Arránquese de un punto razonable.** Newton converge muy rápido *cerca* de la solución, pero puede divergir o saltar a otra raíz si el punto inicial está lejos. Una estrategia sensata es dar unos pasos de gradiente y rematar con Newton.
4. **Ejecútese la iteración vigilando $f(\mathbf{x}_k)$.** El valor de la función debe decrecer en cada paso del descenso; si sube, el paso es demasiado grande y debe reducirse a la mitad. Esa comprobación es la versión mínima de la *búsqueda en línea*.
5. **Párese con un criterio doble.** Se detiene cuando $\|\nabla f(\mathbf{x}_k)\|<\varepsilon$ y $\|\mathbf{x}_{k+1}-\mathbf{x}_k\|<\varepsilon$. Solo el segundo criterio es engañoso: un paso pequeño también ocurre cuando el método se ha atascado.
6. **Verifíquese que el punto hallado es mínimo y no otra cosa.** Ambos métodos buscan $\nabla f=\mathbf{0}$, y eso incluye máximos y puntos de silla; Newton, de hecho, converge hacia sillas con la misma alegría. Calcúlese la hessiana en el punto final y compruébese que es definida positiva (véase la sección 9.4).
7. **Repítase desde varios puntos iniciales.** Ninguno de los dos métodos distingue un mínimo local de uno global. Media docena de arranques dispersos es la comprobación más barata que existe contra ese error.

Ejemplo 9.40 (El paso lo es todo: cuatro ejecuciones sobre $f(x,y)=x^2+10y^2$)**.** La

función tiene un único mínimo en el origen, gradiente $\nabla f = (2x,\, 20y)$ y hessiana constante $Hf = \begin{pmatrix} 2 & 0 \\ 0 & 20 \end{pmatrix}$, de autovalores $m = 2$ y $L = 20$. Su número de condición es $\kappa = L/m = 10$: las curvas de nivel son elipses diez veces más largas que anchas. Arráncase siempre de $\mathbf{x}_0 = (1, 1)$, donde $f = 11$.

(a) El paso óptimo $\alpha = 2/(L + m) = 1/11$. La iteración factoriza por coordenadas: $x_{k+1} = (1 - 2\alpha)x_k$ e $y_{k+1} = (1 - 20\alpha)y_k$, es decir factores $9/11$ y $-9/11$. Los valores de f son

$$11, \quad 7{,}3636, \quad 4{,}9294, \quad 3{,}2998, \quad 2{,}2090, \quad 1{,}4787, \quad 0{,}9899, \ \ldots$$

Cada paso multiplica f por $(9/11)^2 \approx 0{,}6694$. Nótese que *y cambia de signo* en cada paso: el excursionista zigzaguea de una ladera a la otra del valle estrecho, y ese zigzag es exactamente lo que hace lento al método cuando κ es grande.

(b) El paso $\alpha = 1/20 = 1/L$. Ahora el factor de y es $1 - 20/20 = 0$: la coordenada estrecha se anula *de un solo golpe*, y a partir de ahí el problema es unidimensional en x con factor $1 - 2/20 = 9/10$. Los valores de f son

$$11, \quad 0{,}81, \quad 0{,}6561, \quad 0{,}5314, \quad 0{,}4305, \quad 0{,}3487, \quad 0{,}2824, \ \ldots$$

Muchísimo mejor en las primeras iteraciones que el paso "óptimo", pero asintóticamente más lento (0,81 por paso frente a 0,6694). La optimalidad de $2/(L + m)$ es una garantía de *peor caso*, no una promesa de superioridad en cada trayectoria.

(c) El paso $\alpha = 1/10$. El factor de y es $1 - 20/10 = -1$: la coordenada y oscila entre $+1$ y -1 eternamente, sin crecer ni decrecer. La función se estanca cerca de 10 mientras x sí converge. Es el caso frontera, y muestra que el criterio de parada por $\|\mathbf{x}_{k+1} - \mathbf{x}_k\|$ nunca se activaría.

(d) El paso $\alpha = 0{,}11 > 2/L = 0{,}1$. El factor de y pasa a ser $1 - 2{,}2 = -1{,}2$, de módulo mayor que uno. Las coordenadas y sucesivas son

$$1, \quad -1{,}2, \quad 1{,}44, \quad -1{,}728, \quad 2{,}0736, \quad -2{,}48832, \quad 2{,}985984, \ \ldots$$

y el método *diverge* en espiral, aunque la función sea convexa y tenga un mínimo perfectamente definido. Un 10 % de más en el paso basta para arruinarlo todo.

Ejemplo 9.41 (Newton: convergencia cuadrática y sus límites)**. (a) Un sistema no lineal resuelto en cuatro pasos.** Búsquese la intersección de la circunferencia $x^2 + y^2 = 4$ con la parábola $y = x^2$ en el primer cuadrante. Póngase $\mathbf{F}(x, y) = (x^2 + y^2 - 4,\ y - x^2)$, cuya jacobiana es

$$J\mathbf{F} = \begin{pmatrix} 2x & 2y \\ -2x & 1 \end{pmatrix}.$$

Partiendo de (1,5, 1,5), las abscisas sucesivas son

$$1{,}270833, \quad 1{,}2497982, \quad 1{,}24962108, \quad 1{,}2496210677,$$

y las normas del residuo $\|\mathbf{F}\|$ valen $7{,}7 \cdot 10^{-2}$, $6{,}3 \cdot 10^{-4}$, $4{,}4 \cdot 10^{-8}$ y $2{,}2 \cdot 10^{-16}$. Obsérvese el patrón: el exponente se *duplica* en cada paso, que es la firma de la convergencia cuadrática.

La solución exacta es $y = (-1+\sqrt{17})/2$ y $x = \sqrt{y}$, es decir $x = 1{,}2496210676876531$: las doce cifras coinciden.

(b) Newton sobre una cuadrática acierta en un paso. Vuélvase a $f = x^2 + 10y^2$ desde $(1,1)$. Como $Hf^{-1}\nabla f = \begin{pmatrix} 1/2 & 0 \\ 0 & 1/20 \end{pmatrix}\begin{pmatrix} 2 \\ 20 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$, el primer paso lleva exactamente a $(0,0)$. Compárese con las decenas de iteraciones que necesitaba el descenso: Newton *no ve* el número de condición, porque la hessiana inversa lo corrige.

(c) La función de Rosenbrock, el banco de pruebas clásico. Sea $f(x,y) = 100(y - x^2)^2 + (1-x)^2$, con

$$\nabla f = \big(-400x(y-x^2) + 2x - 2, \quad 200(y-x^2)\big).$$

Su único mínimo está en $(1,1)$, donde $f = 0$. Allí la hessiana es $\begin{pmatrix} 802 & -400 \\ -400 & 200 \end{pmatrix}$, con autovalores $501 \pm \sqrt{250601}$, o sea 0,39936 y 1001,60: número de condición $\kappa \approx 2508$. El descenso por gradiente necesita miles de pasos para recorrer ese valle en forma de plátano. Newton, en cambio, arrancando en $(0,0)$ llega a $(1,0)$ con $f = 100$ y de ahí a $(1,1)$ con $f = 0$: dos pasos.

(d) Dónde falla Newton. En $(0,0)$ la hessiana de Rosenbrock es $\begin{pmatrix} 2 & 0 \\ 0 & 200 \end{pmatrix}$, definida positiva, y por eso el paso funcionó. Pero si en algún iterado la hessiana es indefinida — por ejemplo en un punto de silla, donde tiene autovalores de ambos signos — el "paso de Newton" apunta cuesta *arriba* en alguna dirección y el método puede converger al propio punto de silla. Si la hessiana es singular, ni siquiera está definido. Los métodos serios (Levenberg–Marquardt, regiones de confianza) resuelven esto sustituyendo Hf por $Hf + \lambda I$ con λ suficientemente grande, lo cual interpola continuamente entre Newton ($\lambda = 0$) y el descenso por gradiente ($\lambda \to \infty$).

Cuidado

Un error muy común es aplicar Newton *para resolver* $\nabla f = \mathbf{0}$ y creer que se está minimizando. No es lo mismo: resolver $\nabla f = \mathbf{0}$ encuentra cualquier punto crítico, y sin comprobar la hessiana al final uno puede entregar orgullosamente un máximo. Un segundo error es confiar en el criterio $\|\mathbf{x}_{k+1} - \mathbf{x}_k\| < \varepsilon$ a solas: en el caso (c) del primer ejemplo los pasos son grandes y el método no avanza, y en un descenso con paso diminuto los pasos son pequeños desde el principio aunque falte muchísimo para el mínimo.

Pausa para pensar

En dimensión n, un paso de descenso cuesta $O(n)$ operaciones más el coste de evaluar el gradiente; un paso de Newton exige montar una matriz $n \times n$ y resolver un sistema, es decir $O(n^3)$. Con $n = 10^6$ — tamaño habitual en aprendizaje automático — la hessiana ni siquiera cabe en memoria. ¿Qué haría el lector para conservar algo de la información de segundo orden sin construir la matriz? La respuesta que dio la comunidad numérica se llama *métodos quasi-Newton*: aproximar Hf^{-1} acumulando, paso a paso, los pares

$(\mathbf{x}_{k+1} - \mathbf{x}_k,\ \nabla f(\mathbf{x}_{k+1}) - \nabla f(\mathbf{x}_k))$ observados durante la propia iteración.

9.6. Integrales múltiples

Idea clave

Intuición primero. Una integral doble es lo mismo que una simple, pero cortando en cajitas en vez de en rebanadas: sumas f por el área de cada cajita y afinas. Lo nuevo no es el concepto, es la *contabilidad*: hay que describir la región de integración, y esa descripción es la que decide los límites.
Dos ideas hacen todo el trabajo. La primera es el teorema de Fubini: una integral en varias variables se calcula haciendo integrales de una variable, una detrás de otra, y se puede elegir el orden. Elegir bien el orden convierte a veces una integral imposible en una trivial, así que siempre vale la pena dibujar la región y preguntarse cómo conviene barrerla. La segunda es el cambio de variables, con su jacobiano: al deformar el dominio, las cajitas cambian de tamaño, y el jacobiano es exactamente el factor por el que cambia el volumen de cada cajita. El r que aparece en coordenadas polares no es un truco que memorizar: es ese factor, y dice que las cajitas polares son más grandes lejos del origen, cosa que se ve en cualquier dibujo de una malla polar.

Definición 9.42. Sobre un rectángulo R, $\iint_R f \,\mathrm{d}A$ se define con sumas de Riemann sobre subdivisiones, como en una variable; sobre una región general acotada, se extiende f por cero a un rectángulo que la contenga. Toda función continua sobre una región "razonable" (frontera formada por gráficas de funciones continuas) es integrable.

Teorema 9.43 (Fubini)**.** *Si f es continua en $R = [a,b] \times [c,d]$:*

$$\iint_R f \,\mathrm{d}A = \int_a^b \int_c^d f(x,y)\,\mathrm{d}y\,\mathrm{d}x = \int_c^d \int_a^b f(x,y)\,\mathrm{d}x\,\mathrm{d}y.$$

Sobre regiones tipo I *($a \le x \le b$, $\varphi_1(x) \le y \le \varphi_2(x)$):* $\displaystyle\iint f \,\mathrm{d}A = \int_a^b \int_{\varphi_1(x)}^{\varphi_2(x)} f \,\mathrm{d}y\,\mathrm{d}x$, *y análogamente tipo II intercambiando papeles.*

Ejemplo 9.44 (Cambiar el orden salva integrales)**.** $\displaystyle\int_0^1 \int_y^1 e^{x^2} \mathrm{d}x\,\mathrm{d}y$ es incalculable tal cual (e^{x^2} no tiene primitiva elemental). La región es $0 \le y \le x \le 1$; invirtiendo el orden:

$$\int_0^1 \int_0^x e^{x^2} \mathrm{d}y\,\mathrm{d}x = \int_0^1 x e^{x^2} \mathrm{d}x = \frac{e-1}{2}.$$

Dibujar la región antes de integrar no es opcional.

Teorema 9.45 (Cambio de variables)**.** *Si $T : (u,v) \mapsto (x,y)$ es biyección C^1 entre S y R con* $\det DT \neq 0$*:*

$$\iint_R f(x,y)\,\mathrm{d}x\,\mathrm{d}y = \iint_S f(T(u,v))\,\big|\det DT(u,v)\big|\,\mathrm{d}u\,\mathrm{d}v.$$

La razón: T manda un rectángulo infinitesimal $\mathrm{d}u\,\mathrm{d}v$ *a un paralelogramo de área* $|\det DT|\mathrm{d}u\,\mathrm{d}v$ *—el determinante como factor de volumen del capítulo 7.*

Los tres cambios estándar y sus jacobianos:

$$\underbrace{x=r\cos\theta,\ y=r\sin\theta}_{\text{polares: } J=r} \qquad \underbrace{+\,z=z}_{\text{cilíndricas: } J=r} \qquad \underbrace{x=\rho\sin\phi\cos\theta,\ y=\rho\sin\phi\sin\theta,\ z=\rho\cos\phi}_{\text{esféricas: } J=\rho^2\sin\phi}$$

Ejemplo 9.46 (La integral de Gauss)**.** $I=\int_{-\infty}^{\infty}e^{-x^2}\mathrm{d}x$ cumple

$$I^2=\iint_{\mathbb{R}^2}e^{-(x^2+y^2)}\mathrm{d}A=\int_0^{2\pi}\int_0^{\infty}e^{-r^2}r\,\mathrm{d}r\,\mathrm{d}\theta=2\pi\cdot\frac{1}{2}=\pi\implies I=\sqrt{\pi}.$$

Esta integral es la base de la distribución normal en probabilidad y de la función Gamma en $s=\frac{1}{2}$: $\Gamma(\frac{1}{2})=\sqrt{\pi}$.

9.6.1. Aplicaciones físicas

Para una lámina con densidad $\rho(x,y)$ sobre D:

$$m=\iint_D\rho\,\mathrm{d}A,\qquad \bar{x}=\frac{1}{m}\iint_D x\rho\,\mathrm{d}A,\qquad \bar{y}=\frac{1}{m}\iint_D y\rho\,\mathrm{d}A,$$

y los *momentos de inercia* $I_x=\iint y^2\rho\,\mathrm{d}A$, $I_y=\iint x^2\rho\,\mathrm{d}A$, $I_0=I_x+I_y$. En 3D las fórmulas son análogas con integrales triples. El *valor medio* de f sobre D es $\frac{1}{\text{área}(D)}\iint_D f\,\mathrm{d}A$.

Regla: montar una integral múltiple: orden de integración y coordenadas

Lo difícil de una integral múltiple casi nunca es integrar: es *describir la región*. El orden de trabajo es este.

1. **Dibujar la región.** Siempre, aunque sea un croquis torcido a mano. Las intersecciones de las curvas frontera son los datos que fijan los límites exteriores; sin dibujo se sacan mal casi siempre.
2. **Elegir el orden.** Para $\mathrm{d}y\,\mathrm{d}x$ se barre en vertical: x recorre un intervalo fijo $[a,b]$ y, para cada x, la y va de la curva de abajo a la de arriba. Para $\mathrm{d}x\,\mathrm{d}y$, en horizontal. **Regla de oro: los límites exteriores son siempre constantes**; si al terminar quedan variables en los límites de fuera, hay un error.
3. **Cambiar el orden si la región lo pide.** Si un barrido obliga a partir la región en dos o tres trozos y el otro no, se usa el otro. Para invertir el orden se despejan las mismas curvas respecto de la otra variable, siempre releyendo el dibujo.
4. **Cambiar el orden si el integrando lo pide.** A veces la región no es el problema sino la primitiva: si $\int e^{x^2}\mathrm{d}x$ aparece, es imposible en términos elementales, y lo único que salva el ejercicio es integrar primero en y. Es una señal muy reconocible.
5. **Considerar el cambio de coordenadas.** Círculos, sectores y coronas piden *polares*, con el factor r: $\mathrm{d}A=r\,\mathrm{d}r\,\mathrm{d}\theta$. Cilindros y sólidos de revolución piden *cilíndricas*, con

$\mathrm{d}V = r\,\mathrm{d}r\,\mathrm{d}\theta\,\mathrm{d}z$. Esferas, conos y bolas piden *esféricas*, con $\mathrm{d}V = \rho^2 \sin\varphi\,\mathrm{d}\rho\,\mathrm{d}\varphi\,\mathrm{d}\theta$. Olvidar el jacobiano ($r$, r, $\rho^2\sin\varphi$) es el error número uno del capítulo.

6. **Si la región es rara pero el integrando sugiere un cambio, inventarlo.** Se define u, v de modo que la región se vuelva un rectángulo y se usa $\mathrm{d}x\,\mathrm{d}y = |\det J|\,\mathrm{d}u\,\mathrm{d}v$, con J el jacobiano del cambio. Atajo útil: es más fácil calcular $\partial(u,v)/\partial(x,y)$ e invertir el resultado que despejar x, y.

7. **Controles de sensatez.** Si el integrando es positivo, el resultado debe salir positivo. Si hay simetría impar respecto de un eje y la región es simétrica, la integral es cero sin calcular nada. Y el orden de magnitud debe parecerse a (valor típico del integrando) × (área o volumen de la región).

Ejemplo 9.47 (Tres integrales que se resuelven eligiendo bien)**. (a) Cambiar el orden porque la primitiva no existe.** Calcular

$$\int_0^1 \int_y^1 e^{x^2}\,\mathrm{d}x\,\mathrm{d}y.$$

Tal como está es imposible: e^{x^2} no tiene primitiva elemental. Leamos la región: $0 \le y \le 1$ y $y \le x \le 1$, es decir el triángulo de vértices $(0,0)$, $(1,0)$, $(1,1)$, el que queda *bajo* la diagonal. Descrito al revés: $0 \le x \le 1$ y $0 \le y \le x$. Entonces

$$\int_0^1 \int_0^x e^{x^2}\,\mathrm{d}y\,\mathrm{d}x = \int_0^1 x e^{x^2}\,\mathrm{d}x = \Big[\tfrac{1}{2}e^{x^2}\Big]_0^1 = \frac{e-1}{2} \approx 0{,}8591.$$

El x que aparece al integrar en y es justo lo que faltaba para que la primitiva existiera. Control: el integrando vale entre 1 y e sobre un triángulo de área $\frac{1}{2}$, así que el resultado debe estar entre 0,5 y 1,36. Lo está.

(b) Polares. $\displaystyle\iint_{x^2+y^2\le 1} e^{-(x^2+y^2)}\,\mathrm{d}A$. En cartesianas es intratable; en polares el integrando es e^{-r^2} y la región es $0 \le r \le 1$, $0 \le \theta \le 2\pi$:

$$\int_0^{2\pi}\int_0^1 e^{-r^2}\, r\,\mathrm{d}r\,\mathrm{d}\theta = 2\pi\Big[-\tfrac{1}{2}e^{-r^2}\Big]_0^1 = \pi(1-e^{-1}) \approx 1{,}9858.$$

De nuevo el jacobiano r es precisamente el factor que hace la integral elemental. Control: el integrando va de $e^{-1} \approx 0{,}37$ a 1 sobre un disco de área $\pi \approx 3{,}14$, luego el resultado debe estar entre 1,16 y 3,14.

(c) Cilíndricas. Volumen de la región encerrada entre el paraboloide $z = x^2 + y^2$ y el plano $z = 4$. En cilíndricas el paraboloide es $z = r^2$ y las dos superficies se cortan donde $r^2 = 4$, es decir $r = 2$:

$$V = \int_0^{2\pi}\int_0^2\int_{r^2}^4 r\,\mathrm{d}z\,\mathrm{d}r\,\mathrm{d}\theta = 2\pi\int_0^2 (4r - r^3)\,\mathrm{d}r = 2\pi\Big[2r^2 - \tfrac{r^4}{4}\Big]_0^2 = 2\pi(8-4) = 8\pi.$$

Control precioso: el cilindro que contiene la región tiene volumen $\pi\cdot 2^2\cdot 4 = 16\pi$, y el resultado es exactamente la mitad. El paraboloide parte el cilindro en dos mitades iguales, un hecho que ya conocía Arquímedes.

Pausa para pensar

¿Por qué el jacobiano de las polares es r y no 1? Porque un "rectangulito" en coordenadas polares no es un rectángulo: al avanzar $\mathrm{d}\theta$, el arco recorrido mide $r\,\mathrm{d}\theta$, y ese arco crece conforme uno se aleja del origen. El pedacito tiene lados $\mathrm{d}r$ y $r\,\mathrm{d}\theta$, de modo que su área es $r\,\mathrm{d}r\,\mathrm{d}\theta$. Lejos del centro los pedacitos son más grandes, y el r es la contabilidad exacta de esa deformación. La misma lectura vale para $\rho^2 \sin\varphi$ en esféricas: $\rho\,\mathrm{d}\varphi$ por un lado y $\rho \sin\varphi\,\mathrm{d}\theta$ por el otro —el $\sin\varphi$ es el radio del paralelo, que se encoge hacia los polos—.

Regla: cambiar el orden de integración en una integral doble o triple

Muchas integrales iteradas son imposibles en el orden en que vienen escritas y elementales en el orden contrario. El cambio no es un truco: es reescribir la *misma* región de otra manera.

1. **Lee la región en los límites, no en el integrando.** En $\int_a^b \int_{g_1(x)}^{g_2(x)} f\,\mathrm{d}y\,\mathrm{d}x$ la región es $D = \{(x,y) : a \le x \le b,\ g_1(x) \le y \le g_2(x)\}$. Los límites exteriores son siempre constantes; los interiores pueden depender de la variable exterior, nunca al revés.

2. **Dibuja D.** Grafica las curvas $y = g_1(x)$, $y = g_2(x)$ y las rectas $x = a$, $x = b$, y sombrea. Este paso no es opcional: el 90 % de los errores nacen de saltárselo.

3. **Cambia de tipo.** La región estaba descrita como *tipo I* (verticales: x fijo, y varía). Descríbela ahora como *tipo II* (horizontales: y fijo, x varía): busca el rango total de y, digamos $c \le y \le d$, y para cada y los extremos izquierdo y derecho $h_1(y) \le x \le h_2(y)$.

4. **Despeja las fronteras respecto de la otra variable.** Si la frontera era $y = x^2$ con $x \ge 0$, ahora es $x = \sqrt{y}$. Cuida el signo y el dominio al despejar.

5. **Parte la región si hace falta.** Si al barrer horizontalmente el extremo derecho cambia de fórmula a media altura, la integral nueva es una *suma* de dos integrales. Un orden puede necesitar un solo pedazo y el otro tres.

6. **Escribe la integral nueva e intégrala.** Ahora el integrando debe volverse elemental: típicamente el factor que faltaba (x, y, o un d del cambio) aparece solo al integrar la variable interior.

7. **Verifica.** Comprueba el área: $\iint_D 1$ debe dar lo mismo en ambos órdenes. Y evalúa numéricamente la integral original: si el nuevo resultado no coincide, la región se leyó mal.

En triples es igual, solo que hay $3! = 6$ órdenes posibles y la proyección de la región sobre cada plano coordenado dicta los límites de los dos exteriores.

Ejemplo 9.48 (Cuatro integrales imposibles que se vuelven elementales, y una que no se deja)**.** (a)

La primitiva que no existe. Queremos

$$I = \int_0^1 \int_x^1 e^{y^2}\,\mathrm{d}y\,\mathrm{d}x.$$

La integral interior es imposible: e^{y^2} no tiene primitiva elemental. Pero la región es $D = \{0 \le x \le 1,\ x \le y \le 1\}$, es decir el triángulo de vértices $(0,0)$, $(0,1)$, $(1,1)$. Barriéndolo horizontalmente: $0 \le y \le 1$ y, para cada y, $0 \le x \le y$. Entonces

$$I = \int_0^1 \int_0^y e^{y^2}\,\mathrm{d}x\,\mathrm{d}y = \int_0^1 y\,e^{y^2}\,\mathrm{d}y = \Big[\tfrac{1}{2}e^{y^2}\Big]_0^1 = \frac{e-1}{2} \approx 0{,}859141.$$

El factor y que salió de integrar $\mathrm{d}x$ es exactamente el que pedía la sustitución $u = y^2$. El integrando no cambió; cambió quién es la variable interior.

(b) **El mismo truco con seno.** Sea

$$J = \int_0^{\sqrt{\pi}} \int_y^{\sqrt{\pi}} \sin(x^2)\,\mathrm{d}x\,\mathrm{d}y.$$

La región es el triángulo $0 \le y \le x \le \sqrt{\pi}$. Invirtiendo: $0 \le x \le \sqrt{\pi}$, $0 \le y \le x$, y

$$J = \int_0^{\sqrt{\pi}} x\,\sin(x^2)\,\mathrm{d}x = \Big[-\tfrac{1}{2}\cos(x^2)\Big]_0^{\sqrt{\pi}} = -\tfrac{1}{2}(\cos\pi - \cos 0) = 1.$$

Un número redondo salido de una integral que parecía intratable.

(c) **Una frontera con pendiente.** Sea

$$K = \int_0^1 \int_{3y}^3 e^{x^2}\,\mathrm{d}x\,\mathrm{d}y.$$

Región: $0 \le y \le 1$, $3y \le x \le 3$, o sea el triángulo de vértices $(0,0)$, $(3,0)$, $(3,1)$. Leído verticalmente: $0 \le x \le 3$ y $0 \le y \le x/3$. Luego

$$K = \int_0^3 \frac{x}{3}\,e^{x^2}\,\mathrm{d}x = \frac{1}{6}\Big[e^{x^2}\Big]_0^3 = \frac{e^9 - 1}{6} \approx 1350{,}347321.$$

Nota que aquí el despeje del paso 4 fue $x = 3y \Rightarrow y = x/3$, y que el $1/3$ sobrevive hasta el resultado final.

(d) **Una triple: el tetraedro y sus seis órdenes.** Sea T el sólido limitado por los tres planos coordenados y $x + y + z = 1$. En el orden $\mathrm{d}z\,\mathrm{d}y\,\mathrm{d}x$,

$$\mathrm{Vol}(T) = \int_0^1 \int_0^{1-x} \int_0^{1-x-y} 1\,\mathrm{d}z\,\mathrm{d}y\,\mathrm{d}x = \int_0^1 \int_0^{1-x} (1-x-y)\,\mathrm{d}y\,\mathrm{d}x = \int_0^1 \frac{(1-x)^2}{2}\,\mathrm{d}x = \frac{1}{6}.$$

Por la simetría total del tetraedro en (x, y, z), los seis órdenes dan integrales *idénticas* salvo el nombre de las letras, y todas valen $1/6$. Es el mejor sólido para practicar: la respuesta se sabe de antemano.

(e) **La que sí depende del orden.** Sobre el cuadrado $[0,1]^2$ tomemos

$$f(x,y)=\frac{x^2-y^2}{(x^2+y^2)^2}, \qquad f(0,0) \text{ sin definir.}$$

Integrando primero en y y luego en x el resultado es $+\pi/4 \approx 0{,}785398$; integrando primero en x y luego en y el resultado es $-\pi/4$. Los dos cálculos son correctos: los *números* son distintos porque la integral doble de $|f|$ sobre $[0,1]^2$ es infinita, y el teorema de Fubini exige precisamente esa hipótesis. La antisimetría $f(y,x)=-f(x,y)$ explica el cambio de signo.

Cuidado

Cambiar el orden *no* es cambiar los límites de sitio. Escribir $\int_0^1\int_x^1 e^{y^2}\,dy\,dx$ como $\int_x^1\int_0^1 e^{y^2}\,dx\,dy$ es un sinsentido: el límite exterior quedó dependiendo de x, que ya no existe fuera. Los límites exteriores siempre son constantes.

Pausa para pensar

En el caso (e), ¿por qué no contradice esto al teorema de Fubini? Porque Fubini pide $\iint |f| < \infty$. Aquí f explota como $1/r^2$ cerca del origen y en polares $\int\int |f|\, r\,dr\,d\theta$ diverge logarítmicamente. Sin integrabilidad absoluta, «la integral doble» ni siquiera está definida: solo existen dos integrales iteradas que resultan ser distintas. Compara con la reordenación de series condicionalmente convergentes en el capítulo 29.

Idea clave

La regla práctica: si la integral interior es imposible, casi seguro el problema pide invertir el orden, y el enunciado fue diseñado para que el factor que aparece al integrar la variable interior complete exactamente una sustitución. Si además la región es un sector o un anillo, el cambio correcto no es de orden sino a polares (sección de coordenadas polares de este mismo capítulo). Y el enfoque general —qué hipótesis hace falta para intercambiar dos operaciones de paso al límite— es el tema del capítulo 31.

9.7. El área de una superficie parametrizada

Una integral doble mide el área de una región *plana*. Pero muchas superficies interesantes — una esfera, un toro, el techo curvo de un edificio — no son planas, y medirlas exige una idea nueva: partir la superficie en paralelogramos infinitesimales y sumar sus áreas. La sorpresa agradable es que ese elemento de área se calcula con un producto vectorial, de modo que todo el asunto se reduce a derivar, hacer un producto vectorial y calcular un módulo.

Definición 9.49 (Superficie parametrizada y elemento de área)**.** Una *superficie parametrizada* es una aplicación C^1

$$\mathbf{r}\colon D\subseteq\mathbb{R}^2 \longrightarrow \mathbb{R}^3, \qquad \mathbf{r}(u,v)=\big(x(u,v),\,y(u,v),\,z(u,v)\big),$$

con D una región medible del plano de parámetros. Los vectores $\mathbf{r}_u = \partial\mathbf{r}/\partial u$ y $\mathbf{r}_v = \partial\mathbf{r}/\partial v$ generan el plano tangente donde no se anula su producto vectorial; se dice entonces que la parametrización es *regular*. El *elemento de área* es $\mathrm{d}S = \|\mathbf{r}_u \times \mathbf{r}_v\|\,\mathrm{d}u\,\mathrm{d}v$, y el área de la superficie es

$$\mathrm{A}(S) = \iint_D \|\mathbf{r}_u \times \mathbf{r}_v\|\,\mathrm{d}u\,\mathrm{d}v.$$

Las tres cantidades $E = \mathbf{r}_u \cdot \mathbf{r}_u$, $F = \mathbf{r}_u \cdot \mathbf{r}_v$, $G = \mathbf{r}_v \cdot \mathbf{r}_v$ forman la *primera forma fundamental*, y la identidad de Lagrange da $\|\mathbf{r}_u \times \mathbf{r}_v\| = \sqrt{EG - F^2}$.

Ejemplo 9.50 (Superficies y sus elementos de área)**. (a) La esfera.** Con

$$\mathbf{r}(\theta, \varphi) = R(\sin\theta\cos\varphi,\ \sin\theta\sin\varphi,\ \cos\theta)$$

sale $\|\mathbf{r}_\theta \times \mathbf{r}_\varphi\| = R^2\sin\theta$, y el área es $\int_0^{2\pi}\!\int_0^{\pi} R^2\sin\theta\,\mathrm{d}\theta\,\mathrm{d}\varphi = 4\pi R^2$.

(b) Una gráfica. Para $z = g(x, y)$ el elemento es $\mathrm{d}S = \sqrt{1 + g_x^2 + g_y^2}\,\mathrm{d}x\,\mathrm{d}y$. Con $g = x + y$ sobre el cuadrado unidad el área es $\sqrt{3} = 1{,}732051$: el plano inclinado tiene más superficie que su sombra.

(c) Un cilindro. Con $\mathbf{r}(u, v) = (\cos u, \sin u, v)$, el producto vectorial tiene norma 1 y el área de la banda $0 \le v \le h$ es $2\pi h$. Coincide con desenrollar el cilindro en un rectángulo: el cilindro es desarrollable.

(d) Parametrizaciones degeneradas. En los polos de la esfera, $\|\mathbf{r}_\theta \times \mathbf{r}_\varphi\| = R^2\sin\theta$ se anula: la parametrización deja de ser regular allí. No es un defecto de la esfera sino de las coordenadas, y por eso las integrales se calculan sin problema pese a la degeneración en un conjunto de medida nula.

Idea clave

El rectangulito $[u, u + \mathrm{d}u] \times [v, v + \mathrm{d}v]$ del dominio de parámetros se estira, al pasar por $\mathbf{r}$, hasta convertirse en un paralelogramo curvo con lados aproximadamente $\mathbf{r}_u\,\mathrm{d}u$ y $\mathbf{r}_v\,\mathrm{d}v$. Y el área de un paralelogramo generado por dos vectores es el módulo de su producto vectorial. Eso es todo: $\|\mathbf{r}_u \times \mathbf{r}_v\|$ es el *factor de estiramiento* del área, exactamente el papel que el jacobiano desempeña en el cambio de variables — de hecho, si la superficie es un trozo de plano, ese módulo *es* el jacobiano.

Regla: calcular el área de una superficie

Todo el procedimiento es mecánico una vez elegida la parametrización; el arte está en elegirla bien.

1. **Elíjase la parametrización y su dominio.** Dos vías: si la superficie es un grafo $z = f(x, y)$, tómese $\mathbf{r}(x, y) = (x, y, f(x, y))$ directamente. Si tiene simetría de revolución o esférica, úsense ángulos. Y anótese *siempre* el dominio D de los parámetros antes de seguir; es donde más se falla.

2. **Calcúlense $\mathbf{r}_u$ y $\mathbf{r}_v$.** Son dos derivadas parciales vectoriales, es decir seis derivadas escalares. Conviene escribirlas en columna para el paso siguiente.

3. **Hágase el producto vectorial y su módulo.** $\mathbf{r}_u \times \mathbf{r}_v$ es un vector normal a la superficie; su módulo es el integrando. Si el determinante sale feo, pruébese la vía alternativa $\sqrt{EG - F^2}$: con frecuencia $F = 0$ y entonces el módulo es simplemente $\sqrt{EG}$.

4. **Simplifíquese antes de integrar.** Aquí se aplican las identidades trigonométricas, se sacan factores positivos del radical y se vigilan los valores absolutos: $\sqrt{\sin^2\varphi} = |\sin\varphi|$, que vale $\sin\varphi$ solo si $\varphi \in [0, \pi]$.

5. **Intégrese sobre D con las técnicas del capítulo.** Si D es un rectángulo de parámetros, Fubini separa la integral en dos de una variable. Si es un disco, pásese a polares en el plano de parámetros.

6. **Compruébese con un caso conocido.** Antes de dar el resultado por bueno, particularícense los datos a una superficie de área sabida — esfera, cilindro, plano — y véase si sale lo que debe.

7. **Compruébese la unidad y el orden de magnitud.** El área tiene dimensiones de longitud al cuadrado; si en el resultado aparece un radio elevado a 1 o a 3, hay un error. Y el área debe crecer si la región crece: un signo negativo o una función decreciente en el radio delatan un fallo de valor absoluto en el paso 4.

Ejemplo 9.51 (Cuatro áreas calculadas de principio a fin)**. (a) La esfera de radio R.** Parametrícese con

$$\mathbf{r}(\varphi, \theta) = (R\sin\varphi\cos\theta,\ R\sin\varphi\sin\theta,\ R\cos\varphi), \qquad (\varphi, \theta) \in [0, \pi] \times [0, 2\pi].$$

Un cálculo directo da $E = R^2$, $F = 0$ y $G = R^2\sin^2\varphi$, luego

$$\|\mathbf{r}_\varphi \times \mathbf{r}_\theta\| = \sqrt{EG - F^2} = R^2|\sin\varphi| = R^2\sin\varphi$$

en el rango indicado. Por tanto

$$\mathrm{A} = \int_0^{2\pi}\int_0^{\pi} R^2\sin\varphi\,\mathrm{d}\varphi\,\mathrm{d}\theta = 2\pi R^2\big[-\cos\varphi\big]_0^{\pi} = 4\pi R^2,$$

la fórmula de siempre. Obsérvese de paso el *teorema de Arquímedes*: como el integrando no depende de θ y $\mathrm{d}(\cos\varphi)$ es lineal en la altura, el área de una banda esférica entre dos planos horizontales es $2\pi Rh$, con h la separación entre los planos, sea la banda ecuatorial o polar. El casquete $\varphi \in [0, \pi/3]$, por ejemplo, mide $2\pi R^2(1 - \cos\frac{\pi}{3}) = \pi R^2$: exactamente un cuarto de la esfera.

(b) El toro de radios $R = 3$ y $r = 1$. Con

$$\mathbf{r}(u, v) = \big((3 + \cos u)\cos v,\ (3 + \cos u)\sin v,\ \sin u\big), \qquad u, v \in [0, 2\pi],$$

sale $E = 1$, $F = 0$ y $G = (3 + \cos u)^2$, de modo que $\|\mathbf{r}_u \times \mathbf{r}_v\| = 3 + \cos u$. Entonces

$$\mathrm{A} = \int_0^{2\pi}\int_0^{2\pi} (3 + \cos u)\,\mathrm{d}u\,\mathrm{d}v = 2\pi \cdot 6\pi = 12\pi^2 \approx 118{,}4352,$$

que coincide con la fórmula general $4\pi^2 Rr$ y con el teorema de Pappus: la circunferencia generatriz, de longitud $2\pi r = 2\pi$, recorre un camino de longitud $2\pi R = 6\pi$, y el producto es el área.

(c) El paraboloide $z = x^2 + y^2$ sobre el disco unidad. Como es un grafo, sirve la fórmula abreviada

$$\mathrm{A} = \iint_D \sqrt{1 + f_x^2 + f_y^2}\,\mathrm{d}x\,\mathrm{d}y = \iint_{x^2+y^2\le 1} \sqrt{1 + 4x^2 + 4y^2}\,\mathrm{d}x\,\mathrm{d}y.$$

En polares el integrando es $\sqrt{1+4\rho^2}$ y el jacobiano aporta un ρ, así que

$$\mathrm{A} = \int_0^{2\pi}\int_0^1 \rho\sqrt{1+4\rho^2}\,\mathrm{d}\rho\,\mathrm{d}\theta = 2\pi \cdot \frac{(1+4\rho^2)^{3/2}}{12}\Big|_0^1 = \frac{\pi}{6}(5\sqrt{5}-1) \approx 5{,}3304.$$

Compárese con el área del disco de abajo, $\pi \approx 3{,}1416$: el paraboloide es aproximadamente un 70 % más grande, que es lo razonable para una superficie que se empina.

(d) El helicoide, una rampa de caracol. Con $\mathbf{r}(u,v) = (u\cos v,\ u\sin v,\ v)$, $u \in [0,1]$, $v \in [0, 2\pi]$, resulta $E = 1$, $F = 0$, $G = 1 + u^2$ y por tanto $\|\mathbf{r}_u \times \mathbf{r}_v\| = \sqrt{1+u^2}$. La integral se separa:

$$\mathrm{A} = 2\pi \int_0^1 \sqrt{1+u^2}\,\mathrm{d}u = 2\pi \cdot \frac{u\sqrt{1+u^2} + \operatorname{arcsinh} u}{2}\Big|_0^1 = \pi\big(\sqrt{2} + \ln(1+\sqrt{2})\big) \approx 7{,}2118.$$

Nótese que $2\pi \cdot 1 = 6{,}2832$ sería el área si la rampa fuese plana (un anillo completo): el exceso 0,93 es lo que cuesta la inclinación.

Ejemplo 9.52 (Cuándo la parametrización engaña, y cómo detectarlo)**. (a) Recorrer dos veces.** Si en la esfera se toma $\theta \in [0, 4\pi]$ en vez de $[0, 2\pi]$, la integral da $8\pi R^2$: el doble. La fórmula del área no sabe si la parametrización es inyectiva; simplemente suma. Regla práctica: antes de integrar hay que asegurarse de que $\mathbf{r}$ es inyectiva en el interior de D, y admitir solapes solo en conjuntos de medida nula (un meridiano, un polo).

(b) Puntos singulares que no importan. En la esfera, $\mathbf{r}_\varphi \times \mathbf{r}_\theta = \mathbf{0}$ cuando $\varphi = 0$ o $\varphi = \pi$: los polos. La parametrización no es regular ahí, y sin embargo el área sale perfecta, porque los polos forman un conjunto de medida nula en el plano de parámetros. Lo mismo ocurre con el vértice del cono $z = \sqrt{x^2+y^2}$: el integrando es la constante $\sqrt{2}$ en todo el disco menos el origen, así que sobre el disco de radio a el área es $\sqrt{2}\,\pi a^2$, y para $a = 1$ vale $\pi\sqrt{2} \approx 4{,}4429$.

(c) El olvido del valor absoluto. Si en el paso 4 se escribe $\sqrt{\sin^2\varphi} = \sin\varphi$ sin mirar el rango, y se integra φ entre 0 y 2π creyendo cubrir la esfera dos veces, el resultado es 0 en lugar de $8\pi R^2$: las dos mitades se cancelan. Un área nunca puede salir negativa ni nula; ese resultado es el aviso de que falta un valor absoluto.

(d) Comparar dos parametrizaciones de la misma superficie. El hemisferio superior de la esfera unidad admite la parametrización esférica, con área $\int_0^{2\pi}\int_0^{\pi/2} \sin\varphi\,\mathrm{d}\varphi\,\mathrm{d}\theta = 2\pi$, y también la de grafo $z = \sqrt{1-x^2-y^2}$ sobre el disco unidad, con integrando

$$\sqrt{1 + \frac{x^2}{1-x^2-y^2} + \frac{y^2}{1-x^2-y^2}} = \frac{1}{\sqrt{1-\rho^2}},$$

cuya integral en polares es $\int_0^{2\pi}\int_0^1 \rho(1-\rho^2)^{-1/2}\mathrm{d}\rho\,\mathrm{d}\theta = 2\pi$. Las dos coinciden, como manda el teorema de invariancia del área frente a reparametrizaciones. La segunda tiene un integrando impropio en el borde, y aun así converge; la primera no tiene esa dificultad. Ese es exactamente el criterio para elegir parametrización: la que convierta la frontera en algo manso.

Cuidado

La fórmula $\mathrm{A} = \iint \sqrt{1+f_x^2+f_y^2}$ solo vale para *grafos*, es decir, superficies que se proyectan biyectivamente sobre una región del plano xy. La esfera completa no lo es — cada punto del disco tiene dos preimágenes, una arriba y otra abajo —, y aplicarla directamente daría la mitad del área. La forma correcta es partir la superficie en trozos que sí sean grafos y sumar, o bien parametrizar con ángulos desde el principio. Este error es hermano del de olvidar los dos signos al despejar de $x^2+y^2+z^2=R^2$.

Pausa para pensar

Hay un fenómeno que conviene conocer: a diferencia de la longitud de una curva, el área de una superficie *no* se puede definir como el límite de las áreas de poliedros inscritos. La *linterna de Schwarz* es el contraejemplo: una sucesión de poliedros triangulados inscritos en un cilindro cuyas áreas tienden a infinito, aunque los poliedros converjan al cilindro punto a punto. El fallo está en que los triángulos, aun siendo pequeños, se inclinan cada vez más respecto de la superficie. Por eso la definición correcta pasa por el plano tangente — es decir, por las derivadas — y no por aproximaciones poligonales. Es un recordatorio de que en varias variables la intuición de una variable no siempre sobrevive al ascenso de dimensión.

9.8. Derivación bajo el signo integral y las funciones Gamma y Beta

Hay integrales que resisten cualquier primitiva y sin embargo se rinden ante un truco de una elegancia asombrosa: meter un parámetro donde no lo había, derivar respecto de ese parámetro *dentro* de la integral, resolver la integral más fácil que resulta, e integrar de vuelta. Feynman popularizó el método hasta el punto de que se le llama a menudo "el truco de Feynman", aunque es la regla de Leibniz de toda la vida. Y de esa misma familia de integrales con parámetro nacen las dos funciones especiales que aparecen en todo el análisis: Gamma y Beta.

Teorema 9.53 (Regla de Leibniz)**.** *Sea $f(x,t)$ continua con $\partial f/\partial x$ continua en $[a,b]\times[c,d]$. Entonces $I(x)=\int_c^d f(x,t)\,\mathrm{d}t$ es derivable y*

$$I'(x) = \int_c^d \frac{\partial f}{\partial x}(x,t)\,\mathrm{d}t.$$

Si además los límites dependen de x y son derivables, con $G(x) = \int_{u(x)}^{v(x)} f(x,t)\,\mathrm{d}t$, entonces

$$G'(x) = f(x, v(x))v'(x) - f(x, u(x))u'(x) + \int_{u(x)}^{v(x)} \frac{\partial f}{\partial x}(x,t)\,\mathrm{d}t.$$

Para intervalos no acotados la conclusión sigue valiendo si $\int \frac{\partial f}{\partial x}$ converge uniformemente en x, lo cual se comprueba en la práctica con el criterio de Weierstrass.

Idea clave

La integral es una suma y la derivada es un límite; el teorema dice cuándo se pueden intercambiar. La versión con límites variables tiene una lectura geométrica inmediata: al mover x un poquito, el área bajo la curva cambia por tres motivos independientes que se suman. Se gana una tirita de anchura $v'(x)\mathrm{d}x$ y altura $f(x, v(x))$ por el borde derecho; se pierde otra por el borde izquierdo; y toda la curva sube o baja un poco, lo cual aporta la integral de $\partial f/\partial x$. Los dos primeros términos son el teorema fundamental del cálculo; el tercero es la novedad.

Regla: resolver una integral con el truco del parámetro

El paso creativo es el primero; los demás son mecánicos.

1. **Introdúzcase un parámetro donde la integral se vuelva fácil.** Si aparece x^a, tómese a como parámetro; si aparece e^{-x^2}, póngase e^{-ax^2}; si hay un $\ln x$ en el denominador, recuérdese que $\int_0^1 x^a\,\mathrm{d}a$ genera precisamente ese logaritmo.

2. **Anótese un valor del parámetro en el que la integral sea conocida.** Casi siempre es un valor que anula el integrando o lo hace trivial: ese será la constante de integración del final. Sin este dato el método no cierra.

3. **Compruébese la hipótesis de derivación.** Sobre un intervalo compacto basta la continuidad de $\partial f/\partial a$. Sobre un intervalo infinito hay que acotar $|\partial f/\partial a| \le g(t)$ con g integrable e *independiente de* a, en un entorno del parámetro.

4. **Derívese respecto del parámetro dentro de la integral.** Con suerte el integrando se simplifica drásticamente: el logaritmo desaparece, una potencia baja de grado o aparece una exponencial pura.

5. **Resuélvase la integral resultante, que ya debe ser elemental.** Si no lo es, o el parámetro se eligió mal o hace falta derivar una segunda vez.

6. **Intégrese en el parámetro y ajústese la constante.** Aquí entra el valor del paso 2. Es el punto donde más se pierde el resultado por olvido.

7. **Particularícese y compruébese numéricamente.** Sustitúyase el valor del parámetro que interesaba, y contrástese el número con una integración numérica en dos o tres casos. Una discrepancia en el tercer decimal delata un signo perdido.

Ejemplo 9.54 (Cuatro integrales que se rinden ante el parámetro)**. (a) La integral con logaritmo en el denominador.** Calcúlese $\displaystyle\int_0^1 \frac{x^a-1}{\ln x}\,\mathrm{d}x$ para $a>-1$. Llámese $I(a)$ a esa integral. Como $\frac{\partial}{\partial a}\frac{x^a-1}{\ln x}=\frac{x^a\ln x}{\ln x}=x^a$, resulta

$$I'(a)=\int_0^1 x^a\,\mathrm{d}x=\frac{1}{a+1}.$$

El logaritmo, que era el estorbo, se ha cancelado. Integrando, $I(a)=\ln(a+1)+C$, y como $I(0)=\int_0^1 0\,\mathrm{d}x=0$ se obtiene $C=0$. Luego $I(a)=\ln(a+1)$. Comprobación numérica: $I(1)=0{,}6931472=\ln 2$, $I(2)=1{,}0986123=\ln 3$ y $I(3)=1{,}3862944=\ln 4$.

(b) Los momentos de la gaussiana, gratis. Se sabe que $J(a)=\displaystyle\int_0^\infty \mathrm{e}^{-ax^2}\mathrm{d}x=\frac{\sqrt{\pi}}{2\sqrt{a}}$ para $a>0$. Derivando respecto de a bajo el signo:

$$-\int_0^\infty x^2\mathrm{e}^{-ax^2}\mathrm{d}x=J'(a)=-\frac{\sqrt{\pi}}{4a^{3/2}}\implies\int_0^\infty x^2\mathrm{e}^{-ax^2}\mathrm{d}x=\frac{\sqrt{\pi}}{4a^{3/2}},$$

y en $a=1$ eso vale 0,4431134627. Otra derivada da $\int_0^\infty x^4\mathrm{e}^{-ax^2}\mathrm{d}x=\dfrac{3\sqrt{\pi}}{8a^{5/2}}$. Toda la familia de momentos pares sale de una sola integral conocida, sin integrar por partes ni una vez.

(c) Límites variables. Sea $G(x)=\displaystyle\int_x^{x^2}\frac{\sin(xt)}{t}\,\mathrm{d}t$ para $x>0$. Los tres términos de la regla completa dan

$$G'(x)=\frac{\sin(x\cdot x^2)}{x^2}\cdot 2x-\frac{\sin(x\cdot x)}{x}\cdot 1+\int_x^{x^2}\cos(xt)\,\mathrm{d}t.$$

La última integral vale $\big[\sin(xt)/x\big]_{t=x}^{t=x^2}=\big(\sin(x^3)-\sin(x^2)\big)/x$, y al sumar todo se simplifica a

$$G'(x)=\frac{3\sin(x^3)-2\sin(x^2)}{x}.$$

Comprobación por cociente incremental: en $x=1{,}3$ la fórmula da 0,3422827 y la derivada numérica 0,3422827; en $x=2$, ambas dan 2,2408399.

(d) Cuándo el intercambio es ilegítimo. Considérese $f(a,t)=a\,t\,\mathrm{e}^{-at^2}$ en $t\in[0,\infty)$. Se tiene $\int_0^\infty f(a,t)\,\mathrm{d}t=1/2$ para todo $a>0$, luego la derivada de la integral es 0. Pero si se deja $a\to 0^+$ la cota uniforme del paso 3 no existe: no hay ninguna $g(t)$ integrable que domine $|\partial f/\partial a|$ cerca de $a=0$, y ahí el método deja de garantizar nada. La moraleja: cuando el intervalo es infinito, saltarse la comprobación de la dominación no es una formalidad, es la diferencia entre un resultado y un disparate.

Ejemplo 9.55 (Las funciones Gamma y Beta y su tabla de valores)**. (a) Definición y la propiedad que lo explica todo.** Para $s>0$ se define

$$\Gamma(s)=\int_0^\infty t^{s-1}\mathrm{e}^{-t}\,\mathrm{d}t.$$

Integrando por partes sale $\Gamma(s+1)=s\,\Gamma(s)$, y como $\Gamma(1)=\int_0^\infty \mathrm{e}^{-t}\mathrm{d}t=1$, por inducción $\Gamma(k)=(k-1)!$ para todo entero $k\geq 1$. Es decir: Γ interpola el factorial en los valores no enteros.

(b) El valor $\Gamma(1/2)$ y la gaussiana. Con el cambio $t = x^2$,

$$\Gamma(\tfrac{1}{2}) = \int_0^\infty t^{-1/2} \mathrm{e}^{-t}\,\mathrm{d}t = 2\int_0^\infty \mathrm{e}^{-x^2}\,\mathrm{d}x = \sqrt{\pi}.$$

De la relación de recurrencia se deduce entonces $\Gamma(3/2) = \tfrac{1}{2}\sqrt{\pi}$ y $\Gamma(5/2) = \tfrac{3}{4}\sqrt{\pi} = 1{,}3293404$. Este es el eslabón que conecta el factorial con la integral de Gauss y, por tanto, con el volumen de la bola n-dimensional.

(c) La función Beta y su fórmula. Para $p, q > 0$,

$$B(p,q) = \int_0^1 x^{p-1}(1-x)^{q-1}\,\mathrm{d}x = \frac{\Gamma(p)\Gamma(q)}{\Gamma(p+q)}.$$

La demostración es un cambio de variable en la integral doble $\Gamma(p)\Gamma(q)$ a coordenadas $t = u+v$, $x = u/(u+v)$, y es uno de los usos más bonitos del jacobiano. Cuatro valores para calibrar: $B(2,3) = \Gamma(2)\Gamma(3)/\Gamma(5) = 1\cdot 2/24 = 1/12$; $B(1/2,1/2) = \pi$; $B(3/2,1/2) = \pi/2$; y $B(4,2) = \Gamma(4)\Gamma(2)/\Gamma(6) = 6/120 = 1/20$. Los cuatro coinciden con la integral calculada directamente.

(d) La aplicación práctica: potencias de senos y cosenos. Con el cambio $x = \sin^2\theta$ la fórmula Beta se convierte en

$$\int_0^{\pi/2} \sin^m\theta\,\cos^n\theta\,\mathrm{d}\theta = \frac{1}{2}\,B\Big(\frac{m+1}{2}, \frac{n+1}{2}\Big),$$

que resuelve de un golpe toda una familia que en el bachillerato se atacaba con fórmulas de reducción. Cuatro casos: $(m,n) = (2,0)$ da $\pi/4$; $(4,0)$ da $3\pi/16$; $(2,2)$ da $\pi/16$; y $(3,1)$ da $1/4$. Los cuatro se han comprobado contra la integral directa.

Cuidado

La regla de Leibniz sobre un intervalo *compacto* solo pide continuidad, y ahí se puede aplicar sin remordimientos. Sobre un intervalo infinito, o cuando el integrando tiene una singularidad en un extremo, hace falta la dominación uniforme del paso 3, y omitirla produce resultados falsos con toda facilidad — el ejemplo clásico es una familia de funciones cuya masa se escapa al infinito mientras la integral permanece constante. Cuando se disponga de la integral de Lebesgue (capítulo 31), el teorema de convergencia dominada es la herramienta correcta y hace esta comprobación mucho más cómoda.

Pausa para pensar

La fórmula $\Gamma(s+1) = s\Gamma(s)$ permite *extender* Γ a los reales negativos: si $-1 < s < 0$, defínase $\Gamma(s) = \Gamma(s+1)/s$, y repítase. ¿Qué ocurre en $s = 0, -1, -2, \ldots$? En cada uno aparece una división por cero, de modo que Γ tiene polos exactamente en los enteros no positivos y en ningún otro sitio. Esa extensión es la que hace de Γ una función meromorfa en todo el plano complejo, y con ella el factorial deja de ser una operación sobre enteros para convertirse en un objeto analítico global (capítulo 30).

9.9. Funciones inversas e implícitas

Idea clave

Intuición primero. ¿Cuándo se puede "despejar"? La respuesta del teorema es: cuando *la versión linealizada* se puede despejar. Cerca de un punto, la función es su diferencial más pelusa; si la diferencial (una matriz) es invertible, invertir la función es invertir la matriz y absorber la pelusa con un punto fijo. La moraleja es el eslogan de todo el capítulo: *las preguntas locales sobre funciones suaves se responden en la matriz jacobiana.* Despejar variables, enderezar coordenadas, definir superficies: todo es la misma pregunta, "¿es invertible la parte lineal?".

Teorema 9.56 (Función inversa)**.** *Si* $\mathbf{F}:\mathbb{R}^n\to\mathbb{R}^n$ *es* C^1 *y* $\det D\mathbf{F}(\mathbf{a})\neq 0$*, existe un entorno donde* $\mathbf{F}$ *es biyectiva con inversa* C^1 *y* $D(\mathbf{F}^{-1})(\mathbf{F}(\mathbf{a}))=[D\mathbf{F}(\mathbf{a})]^{-1}$*.*

Idea clave

Qué dice esto en cristiano. El teorema de la función inversa dice que si el determinante jacobiano no se anula en un punto, entonces cerca de ese punto la aplicación se puede deshacer. La idea es que la derivada es la mejor aproximación lineal, y si esa aproximación lineal es invertible (determinante no nulo), la función original hereda la propiedad en un entorno. El de la función implícita es la misma idea aplicada a despejar: una ecuación $F(x,y)=0$ define y como función de x cerca de un punto siempre que $\frac{\partial F}{\partial y}$ no se anule ahí.

Ejemplo 9.57 (Invertir y despejar localmente)**. (a) Las coordenadas polares.** Para $x=r\cos\theta$, $y=r\sin\theta$ el jacobiano vale r, no nulo salvo en el origen. Luego el cambio es invertible cerca de cualquier punto distinto del origen, y ahí es donde falla: en $r=0$ el ángulo queda indeterminado.

(b) Local no significa global. Ese mismo cambio de polares *no* es globalmente invertible: θ y $\theta+2\pi$ dan el mismo punto. El teorema promete un entorno, y ese entorno puede ser pequeñísimo. La exponencial compleja es el ejemplo extremo: su derivada no se anula nunca, es localmente invertible en todo punto, y es periódica de período $2\pi\mathrm{i}$, luego nada inyectiva.

(c) Despejar de una circunferencia. Sea $F(x,y)=x^2+y^2-1$. En el punto $(0,1)$ se tiene $F_y=2y=2\neq 0$, luego y se despeja localmente como $y=\sqrt{1-x^2}$. En cambio en $(1,0)$ se tiene $F_y=0$ y el teorema no dice nada; y en efecto ahí la circunferencia es vertical y y no es función de x en ningún entorno.

(d) Derivar sin despejar. La fórmula $\frac{\mathrm{d}y}{\mathrm{d}x}=-\frac{F_x}{F_y}$ da la pendiente sin necesidad de resolver. En el punto $(0{,}6;\,0{,}8)$ de la circunferencia sale $-\frac{0{,}6}{0{,}8}=-0{,}75$, y derivando $\sqrt{1-x^2}$ en $x=0{,}6$ sale exactamente $-0{,}75$. Ese es todo el valor práctico de la derivación implícita.

Cuidado

[La condición es suficiente, no necesaria] Que el jacobiano se anule no impide que la función sea invertible: $f(x) = x^3$ tiene $f'(0) = 0$ y es biyectiva de $\mathbb{R}$ en $\mathbb{R}$. Lo que se pierde es la garantía, y también la suavidad de la inversa, pues $\sqrt[3]{x}$ no es derivable en el origen. Segundo aviso: la conclusión es siempre *local*, y confundirla con invertibilidad global es el error habitual, como muestran los apartados (b). Y tercero: en el teorema de la función implícita hay que derivar respecto de la variable que se quiere *despejar*, no respecto de la otra; invertir los papeles de F_x y F_y da una condición distinta y una fórmula equivocada.

Teorema 9.58 (Función implícita)**.** *Sea $F : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}^m$ de clase C^1, $F(\mathbf{a}, \mathbf{b}) = \mathbf{0}$, y supongamos que el bloque jacobiano $\partial F/\partial \mathbf{y}$ es invertible en $(\mathbf{a}, \mathbf{b})$. Entonces cerca de ese punto el sistema $F(\mathbf{x}, \mathbf{y}) = \mathbf{0}$ define $\mathbf{y} = \varphi(\mathbf{x})$ con*

$$D\varphi = -\Big(\frac{\partial F}{\partial \mathbf{y}}\Big)^{-1} \frac{\partial F}{\partial \mathbf{x}}.$$

Caso escalar: $F(x, y) = 0$ con $F_y \neq 0$ define $y(x)$ con $y' = -F_x/F_y$.

Idea clave

Ambos teoremas (equivalentes entre sí) justifican rigurosamente "despejar variables" y son la maquinaria detrás de las superficies de nivel, de las coordenadas curvilíneas (capítulo 36) y del método de Lagrange. El criterio siempre es el mismo: *una jacobiana invertible en el punto.*

9.10. Fundamentos: las demostraciones del cálculo multivariable

9.10.1. Igualdad de las parciales cruzadas

Teorema 9.59 (Clairaut–Schwarz, demostración)**.** *Si f_{xy} y f_{yx} existen y son continuas en un entorno de (a, b), entonces $f_{xy}(a, b) = f_{yx}(a, b)$.*

Demostración. Consideremos el *doble incremento*

$$\Delta(h, k) = f(a + h, b + k) - f(a + h, b) - f(a, b + k) + f(a, b).$$

Fijado k, sea $\varphi(x) = f(x, b + k) - f(x, b)$: entonces $\Delta = \varphi(a + h) - \varphi(a)$, y por el teorema del valor medio (dos veces):

$$\Delta = h\,\varphi'(\xi) = h\big[f_x(\xi, b + k) - f_x(\xi, b)\big] = hk\, f_{xy}(\xi, \eta)$$

con ξ entre $a, a+h$ y η entre $b, b+k$. Simétricamente, agrupando primero en y: $\Delta = hk\, f_{yx}(\xi', \eta')$. Igualando y dividiendo por hk: $f_{xy}(\xi, \eta) = f_{yx}(\xi', \eta')$; al hacer $(h, k) \to (0, 0)$, ambos puntos tienden a (a, b) y la continuidad concluye. □

Idea clave

Qué dice esto en cristiano. Derivar primero respecto de x y luego de y da lo mismo que hacerlo al revés, siempre que las segundas parciales sean continuas. Parece obvio y no lo es: es un teorema con una hipótesis que hace falta de verdad, y existen funciones perfectamente definidas donde las dos derivadas cruzadas dan números distintos. La consecuencia práctica es enorme, porque legitima escribir $\frac{\partial^2 f}{\partial x \partial y}$ sin especificar el orden, y de ahí salen las identidades $\mathrm{rot}(\nabla f) = \mathbf{0}$ y la simetría de la matriz hessiana.

Ejemplo 9.60 (Cuándo coinciden y el caso en que no)**. (a) El caso normal.** Para $f(x,y) = x^2 y + \sin(xy)$, ambas derivadas cruzadas valen $2x + \cos(xy) - xy\sin(xy)$. Coinciden, como era de esperar, porque todo es continuo.

(b) El contraejemplo clásico. Sea $f(x,y) = \dfrac{xy(x^2-y^2)}{x^2+y^2}$ con $f(0,0) = 0$. Sobre el eje vertical se tiene $f_x(0,y) = -y$, y sobre el horizontal $f_y(x,0) = x$. Derivando otra vez en el origen: $f_{xy}(0,0) = \frac{\partial}{\partial y}(-y) = -1$, mientras que $f_{yx}(0,0) = \frac{\partial}{\partial x}(x) = +1$.

(c) Por qué no hay contradicción. Ambas derivadas cruzadas *existen* en el origen y valen -1 y $+1$; lo que falla es que no son *continuas* ahí. La hipótesis del teorema no se cumple, y por eso la conclusión tampoco. Es el ejemplo que demuestra que la continuidad no es un adorno técnico.

(d) Lo que se apoya en este teorema. La matriz hessiana es simétrica gracias a él, y por tanto diagonalizable con autovalores reales, que es lo que permite clasificar los puntos críticos por el signo de esos autovalores. También la igualdad $\mathrm{rot}(\nabla f) = \mathbf{0}$ y la condición $\frac{\partial P}{\partial y} = \frac{\partial Q}{\partial x}$ para que un campo sea conservativo: las dos son Clairaut disfrazado.

Cuidado

[La hipótesis es sobre las segundas parciales, no sobre f] Que f sea continua, o incluso que sus primeras parciales existan en todas partes, no basta. La función del apartado (b) es continua y tiene primeras parciales continuas en todo el plano, y aun así falla. Segundo aviso: el teorema se generaliza a derivadas de cualquier orden y a cualquier número de variables, de modo que en una función C^k se puede reordenar libremente cualquier combinación de hasta k derivadas. Y tercero: en la práctica casi todo lo que aparece es C^∞, así que el contraejemplo es una curiosidad; pero es la curiosidad que explica por qué el teorema se enuncia con hipótesis y no como una obviedad.

Sin continuidad el teorema falla: el contraejemplo clásico es $f = \frac{xy(x^2-y^2)}{x^2+y^2}$ (con $f(0,0) = 0$), donde $f_{xy}(0,0) = -1 \neq 1 = f_{yx}(0,0)$.

9.10.2. Parciales continuas implican diferenciabilidad

Demostración (caso $n = 2$). Queremos ver que $R = f(a+h,b+k) - f(a,b) - f_x(a,b)h - f_y(a,b)k$ es $o(\|(h,k)\|)$. Partimos el incremento en dos pasos coordenados:

$$f(a+h,b+k) - f(a,b) = \underbrace{\big[f(a+h,b+k) - f(a,b+k)\big]}_{\text{TVM en } x:\ = f_x(\xi,b+k)\,h} + \underbrace{\big[f(a,b+k) - f(a,b)\big]}_{\text{TVM en } y:\ = f_y(a,\eta)\,k}.$$

Entonces $R = [f_x(\xi,b+k) - f_x(a,b)]h + [f_y(a,\eta) - f_y(a,b)]k$. Por la *continuidad* de f_x, f_y, los corchetes tienden a 0 cuando $(h,k) \to 0$, y como $|h|, |k| \le \|(h,k)\|$, resulta $|R| \le \varepsilon \cdot 2\|(h,k)\|$: exactamente $o(\|(h,k)\|)$. □

9.10.3. Taylor multivariable, con demostración

Demostración del Taylor de orden 2. El truco es reducir a una variable: fijado $\mathbf{h}$, sea $g(t) = f(\mathbf{a} + t\mathbf{h})$, $t \in [0,1]$. Por la regla de la cadena,

$$g'(t) = \nabla f(\mathbf{a}+t\mathbf{h}) \cdot \mathbf{h}, \qquad g''(t) = \mathbf{h}^{\mathsf{T}} H_f(\mathbf{a}+t\mathbf{h})\,\mathbf{h}.$$

El Taylor de una variable (capítulo 6) con resto de Lagrange aplicado a g en $[0,1]$ da

$$f(\mathbf{a}+\mathbf{h}) = g(1) = g(0) + g'(0) + \tfrac{1}{2}g''(\theta) = f(\mathbf{a}) + \nabla f(\mathbf{a})\cdot\mathbf{h} + \tfrac{1}{2}\,\mathbf{h}^{\mathsf{T}} H_f(\mathbf{a}+\theta\mathbf{h})\,\mathbf{h}$$

para algún $\theta \in (0,1)$; la continuidad de H_f convierte el último término en $\frac{1}{2}\mathbf{h}^{\mathsf{T}}H_f(\mathbf{a})\mathbf{h} + o(\|\mathbf{h}\|^2)$. El mismo mecanismo (derivar g m veces) produce el Taylor de cualquier orden, con los coeficientes multinomiales emergiendo de la regla de la cadena iterada. □

Con esto, el criterio de la hessiana queda demostrado de verdad: en un crítico, $f(\mathbf{a}+\mathbf{h}) - f(\mathbf{a}) = \frac{1}{2}\mathbf{h}^{\mathsf{T}}H\mathbf{h} + o(\|\mathbf{h}\|^2)$, y si H es definida positiva, su menor valor propio $\lambda_{\min} > 0$ da $\mathbf{h}^{\mathsf{T}}H\mathbf{h} \ge \lambda_{\min}\|\mathbf{h}\|^2$ (teorema espectral), que domina al resto: mínimo local estricto.

9.10.4. El teorema de la función inversa, demostrado

Teorema 9.61 (Punto fijo de Banach)**.** *Sea X completo (por ejemplo, un cerrado de $\mathbb{R}^n$) y $T : X \to X$ una* contracción*: $\|T(x) - T(y)\| \le c\|x-y\|$ con $c < 1$. Entonces T tiene exactamente un punto fijo, y la iteración $x_{k+1} = T(x_k)$ converge a él desde cualquier inicio.*

Demostración. $\|x_{k+1} - x_k\| \le c^k\|x_1 - x_0\|$, así que (x_k) es de Cauchy (serie geométrica) y converge a un x^* con $T(x^*) = x^*$ por continuidad. Unicidad: dos puntos fijos cumplirían $\|x^* - y^*\| \le c\|x^* - y^*\|$, forzando $x^* = y^*$. □

Ejemplo 9.62 (Contracciones en varias variables)**. (a) Un sistema lineal.** Sea $T(\mathbf{x}) = A\mathbf{x}+\mathbf{b}$ con $A = \left(\begin{smallmatrix} 0{,}5 & 0{,}2 \\ 0{,}1 & 0{,}3 \end{smallmatrix}\right)$. Como $\|A\|_\infty = 0{,}7 < 1$, T es contracción y el sistema $\mathbf{x} = A\mathbf{x} + \mathbf{b}$ tiene solución única, alcanzable iterando desde cualquier punto. Es el método de Jacobi.

(b) El criterio práctico. Si T es de clase C^1 y $\|DT\| \le k < 1$ en un convexo, entonces T es contracción de factor k allí, por el teorema del valor medio. Comprobar una derivada es mucho más fácil que verificar la definición.

(c) Hay que quedarse dentro. La hipótesis $T(X) \subseteq X$ es tan importante como la contractividad. La aplicación $T(x) = x/2 + 3$ contrae, y si se toma $X = [0, 1]$ resulta que $T(0) = 3 \notin X$: no hay punto fijo en X. Con $X = [0, 10]$ sí lo hay, en $x = 6$.

(d) Para qué se usa. Existencia y unicidad de soluciones de ecuaciones diferenciales (Picard–Lindelöf), el teorema de la función implícita, la convergencia de métodos iterativos y la existencia de fractales autosemejantes. Es probablemente el teorema con mejor relación entre sencillez de la demostración y número de aplicaciones.

Demostración del teorema de la función inversa. Sea $A = D\mathbf{F}(\mathbf{a})$, invertible. Normalizando ($\mathbf{G} = A^{-1}\mathbf{F}$, que basta invertir), podemos suponer $D\mathbf{F}(\mathbf{a}) = I$. Para cada $\mathbf{y}$ cercano a $\mathbf{F}(\mathbf{a})$, resolver $\mathbf{F}(\mathbf{x}) = \mathbf{y}$ equivale a hallar un punto fijo de

$$T_{\mathbf{y}}(\mathbf{x}) = \mathbf{x} + \mathbf{y} - \mathbf{F}(\mathbf{x}).$$

Su diferencial es $I - D\mathbf{F}(\mathbf{x})$, que en $\mathbf{a}$ vale 0; por continuidad de $D\mathbf{F}$, en una bola cerrada pequeña $\|I - D\mathbf{F}\| \leq \frac{1}{2}$, y el teorema del valor medio vectorial hace de $T_{\mathbf{y}}$ una contracción de constante $\frac{1}{2}$ (que además, para $\mathbf{y}$ suficientemente próximo, manda la bola en sí misma). Banach da un único punto fijo: existe una única preimagen local $\mathbf{x} = \mathbf{F}^{-1}(\mathbf{y})$. La estimación de contracción da también la continuidad de la inversa, y la diferenciabilidad con $D(\mathbf{F}^{-1}) = (D\mathbf{F})^{-1}$ sale de pasar al límite en la identidad de aproximación lineal. □

Observación 9.63. La misma iteración de punto fijo, aplicada al problema $F(\mathbf{x}, \mathbf{y}) = \mathbf{0}$ congelando $\mathbf{x}$, demuestra el teorema de la función implícita (o se deduce del inverso aplicado a $(\mathbf{x}, \mathbf{y}) \mapsto (\mathbf{x}, F(\mathbf{x}, \mathbf{y}))$). Y es la *misma* iteración contractiva que demostrará Picard–Lindelöf en el capítulo 12: tres teoremas centrales, un solo mecanismo.

9.10.5. Lagrange, demostrado con la función implícita

Demostración de los multiplicadores. Sea $\mathbf{a}$ extremo de f sobre $M = \{g = c\}$ con $\nabla g(\mathbf{a}) \neq \mathbf{0}$; digamos $g_{x_n}(\mathbf{a}) \neq 0$. Por la función implícita, cerca de $\mathbf{a}$ la restricción es una gráfica $x_n = \varphi(x_1, \ldots, x_{n-1})$, y el problema restringido se vuelve *libre* para $h = f(x_1, \ldots, x_{n-1}, \varphi)$. En el extremo, $\nabla h = 0$; por la regla de la cadena,

$$0 = \frac{\partial h}{\partial x_i} = \frac{\partial f}{\partial x_i} + \frac{\partial f}{\partial x_n}\frac{\partial \varphi}{\partial x_i} = \frac{\partial f}{\partial x_i} - \frac{\partial f}{\partial x_n}\frac{g_{x_i}}{g_{x_n}} \qquad (i < n),$$

usando $\varphi_{x_i} = -g_{x_i}/g_{x_n}$ (fórmula implícita). Con $\lambda = f_{x_n}/g_{x_n}$, esto dice $f_{x_i} = \lambda g_{x_i}$ para todo i: $\nabla f = \lambda \nabla g$. □

9.10.6. Fubini y el cambio de variables: por qué son ciertos

Fubini (idea de la demostración): para f continua en un rectángulo, las sumas de Riemann dobles sobre una malla se pueden agrupar "por filas" o "por columnas"; cada agrupación es una suma de Riemann iterada, y la continuidad uniforme (Heine, capítulo 6) controla el error de ambas agrupaciones simultáneamente, de modo que los tres límites (doble, iterada en un orden, iterada en el otro) coinciden.

Cambio de variables (idea): la clave es local. Cerca de cada punto, T se aproxima por su diferencial: $T(\mathbf{u}_0 + \mathbf{h}) \approx T(\mathbf{u}_0) + DT(\mathbf{u}_0)\mathbf{h}$, así que T manda un cuadradito de área ε^2 en

(casi) un paralelogramo de área $|\det DT(\mathbf{u}_0)|\,\varepsilon^2$ —el determinante como factor de volumen, demostrado en el capítulo 7 para aplicaciones lineales—. Se cubre S con una malla fina, se aplica esta aproximación en cada celda con error controlado por la continuidad uniforme de DT, y se suma: las sumas de Riemann de ambos lados coinciden en el límite. Los tres jacobianos clásicos (polares, cilíndricas, esféricas) son casos de este único teorema.

9.10.7. Convexidad en varias variables

Definición 9.64. $f:\mathbb{R}^n\to\mathbb{R}$ es *convexa* en un convexo D si $f(\lambda\mathbf{x}+(1-\lambda)\mathbf{y})\le\lambda f(\mathbf{x})+(1-\lambda)f(\mathbf{y})$. Para $f\in C^2$: convexa $\iff$ $H_f(\mathbf{x})$ es semidefinida positiva en todo D (por Taylor con resto, exactamente como en una variable).

Teorema 9.65. *Para f convexa: todo mínimo local es global; el conjunto de minimizadores es convexo; y si f es* estrictamente *convexa, el mínimo es único. Además todo punto crítico es mínimo global: para funciones convexas, $\nabla f=\mathbf{0}$ no es solo necesario sino* suficiente.

Demostración. Si $\nabla f(\mathbf{a})=\mathbf{0}$ y f es convexa C^1, la desigualdad del gradiente $f(\mathbf{y})\ge f(\mathbf{x})+\nabla f(\mathbf{x})\cdot(\mathbf{y}-\mathbf{x})$ (que se obtiene haciendo $\lambda\to0$ en la definición) da $f(\mathbf{y})\ge f(\mathbf{a})$ para todo $\mathbf{y}$. Para un mínimo local $\mathbf{a}$ de una convexa cualquiera: dado $\mathbf{y}$, los puntos $\mathbf{a}+t(\mathbf{y}-\mathbf{a})$ con t pequeño tienen $f\ge f(\mathbf{a})$, y la convexidad "estira" esa desigualdad hasta $\mathbf{y}$. □

Esta es la razón por la que la optimización convexa es un mundo aparte (y una industria): sin sillas ni mínimos locales engañosos, los métodos de descenso convergen al óptimo global. Los mínimos cuadrados del capítulo 7 son el ejemplo primordial: $\|A\mathbf{x}-\mathbf{b}\|^2$ es convexa (hessiana $2A^{\mathsf T}A\succeq0$).

9.11. Complementos: parámetros, cuádricas e impropias múltiples

9.11.1. Integrales con parámetro y la integral de Dirichlet

Leibniz + Fubini + convergencia uniforme = el arte de intercambiar operaciones.

Ejemplo 9.66. $F(t)=\int_0^\infty e^{-tx}\frac{\sin x}{x}\mathrm{d}x$: $F'(t)=-\int_0^\infty e^{-tx}\sin x\,\mathrm{d}x=-\frac{1}{1+t^2}$, $F(\infty)=0$: $F(t)=\frac{\pi}{2}-\arctan t$, y con $t\to0^+$:

$$\int_0^\infty\frac{\sin x}{x}\,\mathrm{d}x=\frac{\pi}{2}.$$

El regularizador e^{-tx} es una idea ubicua, y F es la transformada de Laplace de $\frac{\sin x}{x}$.

9.11.2. Cuádricas clasificadas

$\mathbf{x}^{\mathsf T}A\mathbf{x}+\mathbf{b}^{\mathsf T}\mathbf{x}+c=0$: teorema espectral + completar cuadrados; la firma decide: $(+,+,+)$ elipsoide; $(+,+,-)$ hiperboloide de una hoja (reglado: torres de enfriamiento); $(+,-,-)$ dos hojas; con ceros, paraboloides elíptico (antenas) o hiperbólico ($z=x^2-y^2$: la silla). La geometría de grado 2 es el capítulo 7 más traslaciones.

9.11.3. Impropias múltiples

Por exhaución; para $f \geq 0$ el límite no depende de ella. Escalas en dimensión n:

$$\int_{\|\mathbf{x}\|\geq 1} \|\mathbf{x}\|^{-p} \text{ converge } \iff p > n, \qquad \int_{\|\mathbf{x}\|\leq 1} \|\mathbf{x}\|^{-p} \text{ converge } \iff p < n$$

(el jacobiano r^{n-1} corre la frontera). El potencial newtoniano $\int \frac{\rho}{\|\mathbf{x}-\mathbf{y}\|}\mathrm{d}V$ converge dentro de los cuerpos: $p = 1 < 3$.

9.11.4. Jensen

φ convexa: $\varphi\big(\frac{1}{b-a}\int f\big) \leq \frac{1}{b-a}\int \varphi(f)$. De aquí, en una línea, Cauchy–Schwarz integral, AM–GM y las desigualdades de Hölder y Minkowski de los espacios L^p.

9.12. Cierre del temario: los subtemas restantes

9.12.1. Cinemática: componentes de la aceleración

$$\mathbf{a} = \dot{v}\,\mathbf{T} + \kappa v^2\,\mathbf{N}, \qquad a_T = \frac{\mathbf{v}\cdot\mathbf{a}}{\|\mathbf{v}\|}, \quad a_N = \frac{\|\mathbf{v}\times\mathbf{a}\|}{\|\mathbf{v}\|}\,:$$

la tangencial cambia rapidez, la normal dirección. Circular uniforme: $a_N = \frac{v^2}{R}$ (centrípeta *deducida*); curvas cerradas exigen frenar.

Ejemplo 9.67 (Proyectil)**.** Alcance $\frac{v_0^2 \sin 2\alpha}{g}$ (máximo a 45°); en el vértice toda la gravedad es normal: $\kappa = \frac{g}{v_0^2\cos^2\alpha}$ —Frenet leyendo la física—.

9.12.2. Límites iterados vs. doble

$\frac{xy}{x^2+y^2}$: iterados 0 y 0, doble no existe. $\frac{x^2-y^2}{x^2+y^2}$: iterados $1 \neq -1$. Regla: si el doble existe, los iterados que existan coinciden con él; nada se invierte. Intercambiar límites siempre exige un teorema (Fubini, uniformidad, Riemann).

9.12.3. Taylor con multi-índices

$$f(\mathbf{a}+\mathbf{h}) = \sum_{|\alpha|\leq m} \frac{\partial^\alpha f(\mathbf{a})}{\alpha!}\mathbf{h}^\alpha + O(\|\mathbf{h}\|^{m+1}).$$

El término cúbico decide cuando la hessiana degenera: x^2+y^4 (mínimo) vs. x^2+y^3 (pliegue).

9.12.4. Probabilidad con integrales dobles

Densidad conjunta $f \geq 0$, $\iint f = 1$: $P((X,Y)\in D) = \iint_D f$; marginales por Fubini; esperanza $E[g] = \iint gf$; independencia = factorización; cambios de densidad = jacobianos.

Ejemplo 9.68. Exponenciales independientes: $P(X<Y)=\int_0^\infty \lambda e^{-(\lambda+\mu)x}\mathrm{d}x=\frac{\lambda}{\lambda+\mu}$: las tasas compiten proporcionalmente. Y la normal: su constante es *nuestra* $\sqrt{\pi}$, y las polares la resuelven porque $e^{-(x^2+y^2)/2}$ es radial.

9.12.5. Sistemas implícitos

$\frac{\partial(u,v)}{\partial(x,y)}=-\big(\frac{\partial F}{\partial(u,v)}\big)^{-1}\frac{\partial F}{\partial(x,y)}$: jacobianos por bloques; el determinante que no se anula delimita dónde el despeje es legal, y sus ceros son los pliegues (catástrofes en germen).

9.13. Profundización final: teoremas selectos

9.13.1. Funciones homogéneas y el teorema de Euler

Teorema 9.69. $f(t\mathbf{x})=t^k f(\mathbf{x}) \iff \mathbf{x}\cdot\nabla f=kf$. *(Derivar en t y evaluar en 1; recíproco: $t^{-k}f(t\mathbf{x})$ tiene derivada nula.)*

Corolarios: agotamiento del producto en economía ($k=1$); teorema del virial $2\langle T\rangle=-\langle U\rangle$ para potenciales $1/r$ —con el que se pesó la materia oscura—.

9.13.2. Taylor con resto integral

$R_n=\frac{1}{n!}\int_a^x (x-t)^n f^{(n+1)}(t)\mathrm{d}t$ (por partes iteradas): resto explícito, acotable y derivable —la forma de las demostraciones serias—.

9.13.3. Condiciones KKT

Para mín f con $g_i\le 0$: $\nabla f+\sum \mu_i\nabla g_i=0$, $\mu_i\ge 0$, $\mu_i g_i=0$ (holgura: no se paga por restricciones que no muerden).

Ejemplo 9.70. Proyectar $(2,1)$ al triángulo $x+y\le 2$, $x,y\ge 0$: la restricción $x+y=2$ activa da $(\frac{3}{2},\frac{1}{2})$ con $\mu=1>0$; las otras, inactivas con holgura. En convexos, KKT es además suficiente: sobre estas condiciones corre la optimización moderna (SVM, control).

9.13.4. Transporte de Reynolds

$$\frac{\mathrm{d}}{\mathrm{d}t}\iiint_{\Omega(t)} f\,\mathrm{d}V=\iiint_{\Omega(t)}\Big(\frac{\partial f}{\partial t}+\operatorname{div}(f\mathbf{v})\Big)\mathrm{d}V:$$

cambio local más barrido de la frontera móvil. Cierra la caja de derivar integrales y deduce sin heurística la continuidad y las ecuaciones de Euler de fluidos.

9.14. Última capa: hessiana orlada, Morse y rango

9.14.1. La hessiana orlada: el segundo orden de Lagrange

Falta el criterio de segundo orden para extremos *condicionados*: en un candidato de Lagrange, se clasifica la forma cuadrática del lagrangiano $L = f - \lambda g$ *restringida al tangente de la restricción*. El test práctico es la *hessiana orlada*

$$\bar{H} = \begin{pmatrix} 0 & \nabla g^{\mathsf{T}} \\ \nabla g & H_L \end{pmatrix}:$$

con una restricción en $\mathbb{R}^2$, un solo número decide: $\det \bar{H} > 0 \Rightarrow$ máximo condicionado, $< 0 \Rightarrow$ mínimo (¡signos invertidos respecto del caso libre: la orla voltea el criterio!); en general se examinan los signos de los últimos $n - m$ menores principales de $\bar{H}$. Ejemplo: en nuestro problema de $f = xy$ sobre $x^2 + y^2 = 2$, en $(1,1)$ con $\lambda = \frac{1}{2}$: $H_L = \left(\begin{smallmatrix} -1 & 1 \\ 1 & -1 \end{smallmatrix}\right)$, $\nabla g = (2,2)$, y $\det \bar{H} = \det \left(\begin{smallmatrix} 0 & 2 & 2 \\ 2 & -1 & 1 \\ 2 & 1 & -1 \end{smallmatrix}\right) = 16 > 0$: máximo condicionado, confirmando lo que la compacidad ya decía —pero ahora con un criterio local que funciona también en problemas no compactos—.

9.14.2. El lema de Morse

Teorema 9.71. *Cerca de un punto crítico* no degenerado *(*$\det H \neq 0$*), existen coordenadas en las que*

$$f = f(\mathbf{a}) - y_1^2 - \cdots - y_k^2 + y_{k+1}^2 + \cdots + y_n^2 :$$

la función es *exactamente su forma cuadrática, sin resto. El entero* k *(el* índice*) cuenta las direcciones de bajada.*

Consecuencias: los críticos no degenerados son aislados y su tipo es estable ante perturbaciones; y contar críticos por índice ata el análisis a la topología (desigualdades de Morse: una superficie con g agujeros obliga a toda función genérica sobre ella a tener al menos $2g + 2$ críticos —el toro exige mínimo, máximo y dos sillas—). La teoría de Morse es el puente moderno entre optimización y forma global del espacio.

9.14.3. El teorema del rango

Si $D\mathbf{F}$ tiene rango constante r cerca de un punto, existen coordenadas en dominio e imagen donde $\mathbf{F}(x_1, \ldots, x_n) = (x_1, \ldots, x_r, 0, \ldots, 0)$: *toda aplicación de rango constante es, localmente, la proyección lineal estándar*. Unifica inversa ($r = n = m$) e implícita, y es el teorema que garantiza que los conjuntos de nivel de rango máximo son superficies suaves (variedades): la justificación final de todas las "superficies de restricción" de Lagrange.

9.15. Ronda de temarios: envolventes de familias de curvas

Tema clásico de los temarios de Cálculo III que aún nos debíamos, con dividendos inmediatos en el capítulo 12.

Definición 9.72. Dada una familia de curvas $F(x, y, c) = 0$ (una por cada valor del parámetro c), su *envolvente* es una curva tangente a *todas* las curvas de la familia.

Teorema 9.73 (Ecuaciones de la envolvente)**.** *Los puntos de la envolvente satisfacen el sistema*

$$F(x, y, c) = 0, \qquad \frac{\partial F}{\partial c}(x, y, c) = 0,$$

del que se elimina c.

Demostración. Parametrícese la envolvente como $(x(c), y(c))$, tocando en cada instante a la curva de parámetro c. Derivando la identidad $F(x(c), y(c), c) = 0$ respecto de c (regla de la cadena):

$$F_x\, x' + F_y\, y' + F_c = 0.$$

La tangencia dice que la velocidad (x', y') de la envolvente es tangente también a la curva de la familia, cuyo vector normal es (F_x, F_y) (gradiente): así $F_x x' + F_y y' = 0$, y queda $F_c = 0$. □ □

Ejemplo 9.74 (Tres envolventes con nombre)**.** *(a) Rectas a distancia 1 del origen*: $F = x\cos c + y\sin c - 1 = 0$. Entonces $F_c = -x\sin c + y\cos c = 0$; el sistema da $x = \cos c$, $y = \sin c$: la envolvente es el círculo unitario, como la intuición exigía.

(b) La parábola de seguridad: un proyectil lanzado desde el origen con rapidez v y ángulo c recorre $y = x\tan c - \frac{gx^2}{2v^2}(1 + \tan^2 c)$. Con $u = \tan c$ la familia es cuadrática en u, y $F_u = 0$ da $u = \frac{v^2}{gx}$; sustituyendo,

$$y = \frac{v^2}{2g} - \frac{g}{2v^2}\, x^2 :$$

la *parábola de seguridad*. Ningún punto por encima de ella es alcanzable con rapidez v; todo punto por debajo lo es con exactamente dos ángulos (tiro tenso y tiro parabólico): la envolvente resuelve de una vez el problema completo del alcance en balística.

(c) La escalera que resbala: los segmentos de longitud L con extremos en los ejes son $\frac{x}{L\cos c} + \frac{y}{L\sin c} = 1$; el sistema $F = F_c = 0$ (un poco más de álgebra: de $F_c = 0$ sale $\tan^3 c = \frac{y}{x}$) produce la *astroide*

$$x^{2/3} + y^{2/3} = L^{2/3} :$$

la curva que barre una escalera al caer deslizándose, frontera exacta de la región peligrosa.

Observación 9.75. La condición $F = F_c = 0$ dice que la envolvente es el *discriminante* de la familia: el lugar donde curvas vecinas se cortan "a primer orden" ($F(x, y, c)$ y $F(x, y, c + \mathrm{d}c)$ compatibles). En el capítulo 12 este mismo sistema producirá las *soluciones singulares* de la ecuación de Clairaut: soluciones que no salen de la fórmula general porque son la envolvente de todas las demás, el ejemplo canónico de por qué la unicidad de Picard necesita sus hipótesis. Y en óptica, la envolvente de una familia de rayos reflejados es la *cáustica*: la curva brillante en el fondo de la taza de café es la envolvente de los rayos que la pared reflejó.

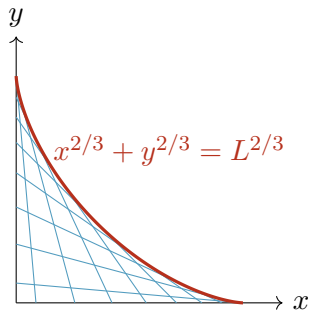


Figura 9.3: La escalera que resbala: cada posición es una recta de la familia; la *astroide* (roja) es la envolvente que todas rozan, calculada con $F = F_c = 0$.

9.16. Cierre definitivo: el teorema del valor medio multivariable

Lo usamos dentro de la demostración de la función inversa; ahora lo enunciamos, demostramos y delimitamos, porque su frontera exacta (funciona escalar, falla vectorial) es de las lecciones más finas del cálculo.

Teorema 9.76 (Valor medio en varias variables)**.** *Si $f : \mathbb{R}^n \to \mathbb{R}$ es diferenciable en un abierto que contiene al segmento $[\mathbf{a}, \mathbf{b}]$, existe $\boldsymbol{\xi}$ en el segmento con*

$$f(\mathbf{b}) - f(\mathbf{a}) = \nabla f(\boldsymbol{\xi}) \cdot (\mathbf{b} - \mathbf{a}).$$

Demostración. Restringir a la recta: $\varphi(t) = f(\mathbf{a} + t(\mathbf{b} - \mathbf{a}))$ es derivable en $[0, 1]$ con $\varphi'(t) = \nabla f \cdot (\mathbf{b} - \mathbf{a})$ (regla de la cadena). El teorema del valor medio del capítulo 6 aplicado a φ es el enunciado. Todo teorema multivariable que se precie tiene un alma unidimensional. □ □

Ejemplo 9.77 (El valor medio en varias variables)**. (a) Un cálculo.** Sea $f(x, y) = x^2 + y^2$ de $(0, 0)$ a $(1, 1)$. Entonces $f(1, 1) - f(0, 0) = 2$, y hay que encontrar c en el segmento con $\nabla f(c) \cdot (1, 1) = 2$, es decir $2c_1 + 2c_2 = 2$. El punto medio $c = (0{,}5;\ 0{,}5)$ lo cumple.

(b) La consecuencia útil. De aquí sale $|f(\mathbf{b}) - f(\mathbf{a})| \le M\|\mathbf{b} - \mathbf{a}\|$ con M una cota del gradiente en el segmento. Ésa es la desigualdad que se usa de verdad; el punto c concreto casi nunca importa.

(c) El dominio debe contener el segmento. En un dominio con forma de herradura, dos puntos cercanos pueden estar unidos sólo por un camino largo, y la cota se aplica a la longitud de ese camino, no a la distancia en línea recta. Por eso los enunciados se hacen en convexos.

(d) Para funciones vectoriales, falla. Con $\mathbf{f}(t) = (\cos t, \sin t)$ de $t = 0$ a $t = 2\pi$ se tiene $\mathbf{f}(2\pi) - \mathbf{f}(0) = \mathbf{0}$ mientras que $\|\mathbf{f}'(t)\| = 1$ nunca se anula: no existe ningún c con $\mathbf{f}'(c) \cdot 2\pi = \mathbf{0}$. Lo que sobrevive es la desigualdad, no la igualdad.

Cuidado

[Para funciones vectoriales es falso] Para $\mathbf{F}:\mathbb{R}\to\mathbb{R}^2$, $\mathbf{F}(t)=(\cos t,\sin t)$: $\mathbf{F}(2\pi)-\mathbf{F}(0)=\mathbf{0}$, pero $\mathbf{F}'(t)=(-\sin t,\cos t)$ *nunca* se anula ($\|\mathbf{F}'\|=1$): no existe el ξ prometido. La razón: cada componente tiene su propio ξ_i y no hay motivo para que coincidan. Lo que sobrevive —y basta para todo— es la *desigualdad* del valor medio:

$$\|\mathbf{F}(\mathbf{b})-\mathbf{F}(\mathbf{a})\|\le\Big(\sup_{[\mathbf{a},\mathbf{b}]}\|D\mathbf{F}\|\Big)\,\|\mathbf{b}-\mathbf{a}\|$$

(demostración: aplicar el teorema escalar a $\varphi(t)=\mathbf{u}\cdot\mathbf{F}(\mathbf{a}+t(\mathbf{b}-\mathbf{a}))$ con $\mathbf{u}$ el unitario en la dirección de $\mathbf{F}(\mathbf{b})-\mathbf{F}(\mathbf{a})$, y Cauchy–Schwarz). La moraleja es general y vale memorizarla: *las igualdades de valor medio son un lujo escalar; las desigualdades son el patrimonio universal.*

Corolario 9.78 (C^1 implica localmente Lipschitz)**.** *Si $\mathbf{F}\in C^1$, en cada bola cerrada $\|\mathbf{F}(\mathbf{x})-\mathbf{F}(\mathbf{y})\|\le L\|\mathbf{x}-\mathbf{y}\|$ con $L=\max\|D\mathbf{F}\|$ (máximo alcanzado: compacidad, capítulo 6).*

Observación 9.79. Este corolario es un conector maestro del libro entero: es *exactamente* la hipótesis que Picard necesita en el capítulo 12 (por eso "$f\in C^1$ garantiza unicidad" en las EDO), es lo que hizo contractiva a la iteración en la demostración de la función inversa, y es la estimación con que el teorema del transporte de Reynolds y todos los teoremas de dependencia continua controlan sus errores. Además regala el criterio práctico: si $\nabla f=\mathbf{0}$ en un abierto *conexo*, f es constante (unir dos puntos por una poligonal y aplicar el teorema tramo a tramo): la unicidad de los potenciales del capítulo 11, demostrada aquí de una vez por todas.

9.17. Segunda ronda: gaussianas n-dimensionales y la bola

9.17.1. La gaussiana multivariable

Teorema 9.80. *Para A simétrica definida positiva:* $\displaystyle\int_{\mathbb{R}^n}e^{-\frac12\mathbf{x}^{\mathsf T}A\mathbf{x}}\mathrm{d}\mathbf{x}=\frac{(2\pi)^{n/2}}{\sqrt{\det A}}$.

Demostración. Diagonalizar ortogonalmente (jacobiano 1): producto de gaussianas $\sqrt{2\pi/\lambda_i}$, y $\prod\lambda_i=\det A$. □

La fórmula madre de la normal multivariante ($A=\Sigma^{-1}$: por eso las covarianzas se simulan con Cholesky) y de las integrales gaussianas de la física; con término lineal, completar cuadrados da $e^{\frac12\mathbf{b}^{\mathsf T}A^{-1}\mathbf{b}}$: la generatriz de momentos.

9.17.2. El volumen de la bola n-dimensional

Ejemplo 9.81 (Doble evaluación)**.** $\int_{\mathbb{R}^n} e^{-\|\mathbf{x}\|^2} = \pi^{n/2}$ (producto) $= \frac{S_{n-1}}{2}\Gamma(\frac{n}{2})$ (cáscaras): despejando,

$$S_{n-1} = \frac{2\pi^{n/2}}{\Gamma(\frac{n}{2})}, \qquad V_n = \frac{\pi^{n/2}}{\Gamma(\frac{n}{2}+1)}:$$

$V_2 = \pi$, $V_3 = \frac{4}{3}\pi$ ✓. Sorpresa: V_n crece hasta $n = 5$ y decae a 0: la bola es minúscula frente al cubo 2^n, y su volumen se concentra junto al borde: la *concentración de la medida*, geometría sobre la que trabaja la estadística en dimensión alta: en mil variables, "el centro está vacío".

9.17.3. Newton multivariable

$\mathbf{x}_{k+1} = \mathbf{x}_k - [D\mathbf{F}]^{-1}\mathbf{F}$ (resolviendo el sistema, no invirtiendo): cuadrático por Taylor. Para optimizar: $\mathbf{x}_{k+1} = \mathbf{x}_k - H_f^{-1}\nabla f$: el gradiente con la métrica correcta, que redondea los valles de κ grande; los cuasi-Newton (BFGS) son el compromiso costo/velocidad que domina la práctica.

9.18. Coda: dualidad de Lagrange

Todo problema de minimización con restricciones tiene una sombra: el *dual*. Con $L(\mathbf{x}, \mu) = f(\mathbf{x}) + \sum \mu_i g_i(\mathbf{x})$, la función dual $q(\mu) = \min_{\mathbf{x}} L$ es *cóncava siempre* (ínfimo de funciones lineales en μ) y satisface la *dualidad débil*: $q(\mu) \le f(\mathbf{x})$ para todo par factible: *cualquier* evaluación dual certifica una cota inferior del óptimo. Para problemas convexos (con regularidad), la brecha se cierra: $\max q = \min f$ (*dualidad fuerte*), y los multiplicadores óptimos son los precios sombra de KKT. Ejemplo mínimo: minimizar x^2 con $x \ge 1$: $q(\mu) = \min_x(x^2 + \mu(1-x)) = \mu - \frac{\mu^2}{4}$, maximizada en $\mu = 2$ con $q = 1 = f(1)$ ✓. Esta simetría primal–dual es el motor del "certificado": la programación lineal, las máquinas de soporte vectorial y los solvers modernos optimizan ambos lados a la vez y se detienen cuando las cotas se tocan —saber *que* terminaste, no solo terminar—.

9.19. Temas extra y ejemplos desarrollados

9.19.1. D'Alembert por cambio de variables

Con $\xi = x - ct$, $\eta = x + ct$: $u_{tt} - c^2 u_{xx} = -4c^2 u_{\xi\eta} = 0$, luego $u = F(x-ct) + G(x+ct)$: toda solución es dos ondas viajeras. Elegir coordenadas donde el operador se desacopla = diagonalizar con la regla de la cadena.

9.19.2. La derivada material

$\frac{Df}{Dt} = f_t + \mathbf{v}\cdot\nabla f$: cambio local más advección. Sobre la velocidad, el término $(\mathbf{v}\cdot\nabla)\mathbf{v}$ —pura regla de la cadena— es el no lineal de Navier–Stokes: la turbulencia y un problema del milenio. Río que se estrecha: la boya acelera aunque el campo sea estacionario.

9.19.3. Descenso por gradiente en cuadráticas

$x_{k+1} = (I - \eta A)x_k$: converge si $\eta < \frac{2}{\lambda_{\max}}$, con razón óptima $\rho^* = \frac{\kappa-1}{\kappa+1}$: el mal condicionamiento ralentiza el aprendizaje ($\kappa = 100 \Rightarrow \rho \approx 0.98$) —teorema espectral explicando el entrenamiento de redes—.

9.19.4. El teorema de la envolvente

$F(t) = \max_x f(x,t) \Rightarrow F'(t) = \frac{\partial f}{\partial t}(x^*(t), t)$: solo la dependencia explícita (la condición de primer orden mata el resto). Generaliza el significado de λ; precios sombra y Hellmann–Feynman.

9.19.5. Curvatura de curvas implícitas

$$\kappa = \frac{|F_{xx}F_y^2 - 2F_{xy}F_xF_y + F_{yy}F_x^2|}{(F_x^2 + F_y^2)^{3/2}}.$$

Elipse en $(a,0)$: $\kappa = \frac{a}{b^2}$, radio osculador $\frac{b^2}{a}$ = semilatus rectum de Kepler.

9.20. Cuándo existe de verdad una integral múltiple: contenido de Jordan

Se dijo que "toda función continua sobre una región razonable es integrable", con las comillas puestas a propósito. Toca quitarlas.

Definición 9.82 (Contenido exterior e interior)**.** Para $S \subset \mathbb{R}^n$ acotado y una malla de cubos de lado 2^{-k}, sean $\overline{c}_k(S)$ el volumen total de los cubos que *cortan* a S y $\underline{c}_k(S)$ el de los que están *contenidos* en S. Sus límites son el **contenido exterior** $\overline{c}(S)$ y el **interior** $\underline{c}(S)$. Si coinciden, S es **medible-Jordan** y ese valor común es su contenido (área, volumen).

Teorema 9.83 (El criterio)**.** *S acotado es medible-Jordan si y solo si ∂S tiene contenido nulo. Y si f es acotada en S medible-Jordan y continua salvo en un conjunto de contenido nulo, entonces f es integrable sobre S.*

Teorema 9.84 (La gráfica de una continua no ocupa volumen)**.** *Si $g : K \to \mathbb{R}$ es continua sobre un compacto $K \subset \mathbb{R}^{n-1}$, su gráfica tiene contenido nulo en $\mathbb{R}^n$.*

Demostración. Por compacidad, g es uniformemente continua: dado ε, hay δ tal que la oscilación de g en cada cubo de lado δ es menor que ε. Cubriendo K con N cubos de lado δ, la gráfica cabe en N cajas de volumen $\delta^{n-1}\varepsilon$, con total $\varepsilon \cdot N\delta^{n-1} \approx \varepsilon \cdot \mathrm{vol}(K)$, tan pequeño como se quiera. □

Eso es exactamente lo que legitima la palabra "razonable": las regiones del capítulo tienen frontera formada por un número finito de gráficas de funciones continuas, luego de contenido nulo, luego son medibles-Jordan.

Ejemplo 9.85 (Uno que sí y otro que no)**. (a) El disco unidad.** Su frontera es la circunferencia, unión de dos gráficas $y = \pm\sqrt{1-x^2}$ sobre $[-1,1]$, ambas continuas. Contenido nulo, luego el disco es medible-Jordan. (Directamente: la circunferencia se cubre con $2n$ cuadraditos de lado $1/n$, de área total $2/n \to 0$.)

(b) Los racionales del cuadrado. Sea $S = \mathbb{Q}^2 \cap [0,1]^2$. Ningún cubo está contenido en S, luego $\underline{c}(S) = 0$; y todo cubo lo corta, luego $\overline{c}(S) = 1$. No es medible-Jordan, y $\mathbf{1}_S$ no es integrable en el sentido de Riemann: es el Dirichlet bidimensional.

(c) La diferencia con Lebesgue. S es numerable, luego tiene *medida* de Lebesgue nula y es Lebesgue-integrable con integral 0. Contenido de Jordan y medida de Lebesgue no son lo mismo: el primero solo admite recubrimientos *finitos*, y esa es toda la diferencia. Un conjunto de contenido nulo tiene medida nula, y el recíproco falla exactamente en ejemplos como este.

9.21. El cambio de variables, y las integrales impropias múltiples

El teorema del cambio de variables es el más usado del capítulo y el único gran resultado del que solo dimos "la idea". Aquí está el esquema completo.

Teorema 9.86. *Sea $T : U \to V$ un difeomorfismo de clase C^1 entre abiertos de $\mathbb{R}^n$ y f integrable en V. Entonces*

$$\int_V f(\mathbf{y})\,\mathrm{d}\mathbf{y} = \int_U f\big(T(\mathbf{u})\big)\,\big|\det DT(\mathbf{u})\big|\,\mathrm{d}\mathbf{u}.$$

Esquema en cuatro pasos. **(1) El caso lineal.** Si $T = A$ es lineal, $\mathrm{vol}(A(Q)) = |\det A|\,\mathrm{vol}(Q)$ para todo cubo Q. Se prueba descomponiendo A en matrices elementales: las de intercambio y suma no cambian el volumen y tienen determinante ± 1; las de escalado lo multiplican por el factor, igual que su determinante.

(2) Localización. Ambos miembros son aditivos respecto de particiones del dominio, así que basta probar la igualdad en un entorno de cada punto.

(3) Factorización en primitivos. Cerca de un punto, el teorema de la función inversa —ya demostrado— permite escribir T como composición de aplicaciones *primitivas*, que cambian una sola coordenada:

$$(u_1,\dots,u_n) \longmapsto (u_1,\dots,u_{k-1},\varphi(\mathbf{u}),u_{k+1},\dots,u_n).$$

Como el determinante es multiplicativo bajo composición y ambos lados de la fórmula también, basta con el caso primitivo.

(4) El caso primitivo. Ahí Fubini reduce el problema a una integral unidimensional en la coordenada afectada, y esa es la fórmula de sustitución de una variable, que ya está demostrada. Inducción en n. □

Ejemplo 9.87 (Comprobación con una transformación no lineal)**.** Sea $T(u,v) = (u^2 - v^2,\ 2uv)$ —que es $z \mapsto z^2$ en complejos— sobre $R = [1,2] \times [0,1]$. Su jacobiano es

$$\det DT = \begin{vmatrix} 2u & -2v \\ 2v & 2u \end{vmatrix} = 4(u^2+v^2) > 0.$$

La fórmula predice que el área de $T(R)$ vale

$$\int_1^2\int_0^1 4(u^2+v^2)\,\mathrm{d}v\,\mathrm{d}u = \frac{32}{3} = 10{,}666\,667,$$

y el mismo número se obtiene calculando $\frac{1}{2}\oint_{\partial T(R)}(x\,\mathrm{d}y - y\,\mathrm{d}x)$ con el teorema de Green sobre la imagen. Que dos métodos independientes coincidan es la comprobación de que el factor jacobiano es exactamente el correcto: mide cuánto dilata T el área *localmente*, y al integrarlo se recupera la dilatación global.

Integrales impropias múltiples

Definición 9.88 (Exhaución)**.** Si $f \geq 0$ en un abierto Ω posiblemente no acotado, se toma una sucesión creciente de compactos medibles $K_1 \subset K_2 \subset \cdots$ con $\bigcup K_j = \Omega$ y se define

$$\int_\Omega f = \lim_j \int_{K_j} f.$$

Para $f \geq 0$ el límite existe (es monótono) y *no depende* de la exhaución elegida.

Teorema 9.89 (Tonelli)**.** *Si $f \geq 0$ es medible, las dos integrales iteradas y la doble coinciden siempre, finitas o infinitas. Si f cambia de signo, hace falta comprobar primero que $\iint |f| < \infty$ (Fubini); sin esa comprobación, las iteradas pueden diferir.*

Ejemplo 9.90 (La integral de Gauss, ahora legítima)**.** El cálculo clásico

$$\Big(\int_{\mathbb{R}} e^{-x^2}\mathrm{d}x\Big)^2 = \iint_{\mathbb{R}^2} e^{-(x^2+y^2)}\mathrm{d}A = \int_0^{2\pi}\int_0^{\infty} e^{-r^2}\,r\,\mathrm{d}r\,\mathrm{d}\theta = \pi$$

usa tres cosas prohibidas sin permiso: pasar de iteradas a doble, cambiar variables en un dominio no acotado y permutar límites. Las tres quedan autorizadas de golpe porque el integrando es *no negativo*: Tonelli da el primer paso, la exhaución por discos $K_j = \{r \leq j\}$ da el segundo, y la monotonía el tercero. Sin la hipótesis de signo, el argumento sería inválido.

Cuidado

[Cuando las iteradas discrepan] Sea

$$f(x,y) = \frac{x^2-y^2}{(x^2+y^2)^2} \qquad \text{en } [0,1]^2.$$

Integrando en un orden y luego en el otro:

$$\int_0^1\int_0^1 f\,\mathrm{d}y\,\mathrm{d}x = \frac{\pi}{4}, \qquad \int_0^1\int_0^1 f\,\mathrm{d}x\,\mathrm{d}y = -\frac{\pi}{4}.$$

Los dos valores existen, son finitos, y son *distintos*. La razón es que $\iint |f| = \infty$: la singularidad en el origen no es absolutamente integrable. Cambiar el orden de integración sin comprobar la integrabilidad absoluta no es un descuido menor, es un error que puede cambiar el signo del resultado.

9.22. Las tres demostraciones que sostienen el capítulo

Tres resultados se han usado en cada página sin demostrarse: la regla de la cadena, el teorema de la función implícita y el criterio de la hessiana. Los tres se prueban aquí.

Teorema 9.91 (Regla de la cadena, versión general)**.** *Sean $g:\mathbb{R}^n\to\mathbb{R}^m$ diferenciable en a y $f:\mathbb{R}^m\to\mathbb{R}^p$ diferenciable en $b=g(a)$. Entonces $f\circ g$ es diferenciable en a y*

$$D(f\circ g)(a)=Df(b)\cdot Dg(a),$$

producto de las matrices jacobianas.

Demostración. Escríbanse las dos hipótesis de diferenciabilidad con sus restos:

$$g(a+h)=g(a)+Dg(a)h+\varepsilon_1(h)\,|h|,\qquad f(b+k)=f(b)+Df(b)k+\varepsilon_2(k)\,|k|,$$

donde $\varepsilon_1(h)\to 0$ cuando $h\to 0$ y $\varepsilon_2(k)\to 0$ cuando $k\to 0$, con $\varepsilon_2(0)=0$ (lo cual se puede imponer, y es lo que evita el problema clásico de dividir entre $|k|$ cuando k se anula).

Póngase $k(h)=g(a+h)-g(a)=Dg(a)h+\varepsilon_1(h)|h|$. Nótese que

$$|k(h)|\ \le\ \big(\|Dg(a)\|+|\varepsilon_1(h)|\big)\,|h|\ \le\ C\,|h|$$

para $|h|$ pequeño, con $C=\|Dg(a)\|+1$. Entonces

$$f\big(g(a+h)\big)=f\big(b+k(h)\big)=f(b)+Df(b)\,k(h)+\varepsilon_2\big(k(h)\big)\,|k(h)|.$$

Sustituyendo $k(h)$ y agrupando,

$$f\big(g(a+h)\big)-f(b)-Df(b)Dg(a)h=\underbrace{Df(b)\,\varepsilon_1(h)\,|h|}_{\text{(I)}}+\underbrace{\varepsilon_2\big(k(h)\big)\,|k(h)|}_{\text{(II)}}.$$

Falta ver que ambos términos son $o(|h|)$. Para (I): $|Df(b)\varepsilon_1(h)|h||\le\|Df(b)\|\,|\varepsilon_1(h)|\,|h|$, y $|\varepsilon_1(h)|\to 0$. Para (II): $|\varepsilon_2(k(h))|\,|k(h)|\le|\varepsilon_2(k(h))|\cdot C|h|$, y cuando $h\to 0$ se tiene $k(h)\to 0$, luego $\varepsilon_2(k(h))\to 0$ (aquí se usa $\varepsilon_2(0)=0$ y la continuidad en 0). Dividiendo todo entre $|h|$ y haciendo $h\to 0$, el cociente tiende a 0: eso es exactamente la diferenciabilidad de $f\circ g$ en a con derivada $Df(b)Dg(a)$. □

Cuidado

[No basta con que existan las parciales] La demostración usa la *diferenciabilidad* de g en a, no sólo la existencia de sus derivadas parciales. La diferencia no es académica: la función

$$g(x,y)=\begin{cases}\dfrac{xy}{x^2+y^2}, & (x,y)\neq(0,0)\\[2mm] 0, & (x,y)=(0,0)\end{cases}$$

tiene ambas parciales nulas en el origen y ni siquiera es continua allí (sobre la recta $y=x$ vale constantemente $\tfrac12$). Componerla con cualquier cosa y aplicar la regla de la cadena produciría un disparate. La condición práctica que sí basta —y que se cumple

casi siempre— es que las parciales existan y sean *continuas* en un entorno: entonces la función es diferenciable.

Teorema 9.92 (Función implícita)**.** *Sea $F:\mathbb{R}^n\times\mathbb{R}^m\to\mathbb{R}^m$ de clase C^1 en un entorno de (a,b) con $F(a,b)=0$, y supóngase que la matriz $m\times m$ de derivadas respecto de las últimas variables, $D_yF(a,b)$, es* invertible*. Entonces existen entornos $U\ni a$ y $V\ni b$ y una única función $\varphi:U\to V$ de clase C^1 tal que*

$$\{(x,y)\in U\times V:\ F(x,y)=0\}=\{(x,\varphi(x)):x\in U\},$$

y además

$$D\varphi(x)=-\big[D_yF(x,\varphi(x))\big]^{-1}D_xF(x,\varphi(x)).$$

Demostración. Existencia y unicidad, por contracción. Sea $A=D_yF(a,b)$ y defínase, para cada x fijo, la aplicación

$$T_x(y)=y-A^{-1}F(x,y).$$

Un punto fijo de T_x es exactamente una solución de $F(x,y)=0$. Se comprueba que T_x es contractiva cerca de b: su derivada respecto de y es

$$D_yT_x(y)=I-A^{-1}D_yF(x,y),$$

que se anula en (a,b) y, por continuidad de D_yF (aquí se usa que F es C^1), tiene norma $\le\frac12$ en una bola $\overline{B}(b,r)$ y para x en un entorno U_0 de a. Por el teorema del valor medio, $|T_x(y_1)-T_x(y_2)|\le\frac12|y_1-y_2|$ allí.

Falta que T_x aplique la bola en sí misma. Como $T_a(b)=b$ y $T_x(b)$ depende continuamente de x, existe $U\subset U_0$ tal que $|T_x(b)-b|<r/2$ para $x\in U$; entonces, para $y\in\overline{B}(b,r)$,

$$|T_x(y)-b|\le|T_x(y)-T_x(b)|+|T_x(b)-b|\le\tfrac12|y-b|+\tfrac{r}{2}\le r.$$

El teorema del punto fijo de Banach (capítulo 12) da, para cada $x\in U$, un único $y=\varphi(x)\in\overline{B}(b,r)$ con $F(x,\varphi(x))=0$.

Continuidad de φ. De $|\varphi(x_1)-\varphi(x_2)|=|T_{x_1}(\varphi(x_1))-T_{x_2}(\varphi(x_2))|\le\frac12|\varphi(x_1)-\varphi(x_2)|+|T_{x_1}(\varphi(x_2))-T_{x_2}(\varphi(x_2))|$ se despeja $|\varphi(x_1)-\varphi(x_2)|\le2|A^{-1}|\,|F(x_1,\varphi(x_2))-F(x_2,\varphi(x_2))|$, y el miembro derecho tiende a 0 con $x_1\to x_2$ por continuidad de F.

Derivabilidad y fórmula. Derívese la identidad $F(x,\varphi(x))=0$ respecto de x usando el teorema 9.91:

$$D_xF+D_yF\cdot D\varphi=0\quad\Longrightarrow\quad D\varphi=-\big[D_yF\big]^{-1}D_xF,$$

válido porque $D_yF(x,\varphi(x))$ sigue siendo invertible en un entorno (el determinante es continuo y no nulo en (a,b)). Que φ sea efectivamente derivable —y no sólo que la fórmula sea consistente— se obtiene repitiendo el argumento del resto: se escribe $0=F(x+h,\varphi(x+h))-F(x,\varphi(x))=D_xF\,h+D_yF\,\Delta\varphi+o(|h|+|\Delta\varphi|)$, y la cota de Lipschitz del párrafo anterior garantiza $|\Delta\varphi|=O(|h|)$, con lo que el resto es $o(|h|)$ y se despeja $\Delta\varphi=-[D_yF]^{-1}D_xF\,h+o(|h|)$. □

Observación 9.93 (Implícita e inversa son el mismo teorema). El teorema de la función inversa se obtiene aplicando el implícito a $F(x,y)=x-f(y)$: la hipótesis $D_yF=-Df(b)$ invertible

es exactamente que el jacobiano de f no se anule, y la φ que sale es f^{-1}. Recíprocamente, el implícito se deduce del inverso aplicándolo a $(x,y)\mapsto(x,F(x,y))$. Son dos formas del mismo hecho, y ese hecho es, en el fondo, el teorema del punto fijo de Banach: *una perturbación pequeña de algo invertible sigue siendo invertible.*

Teorema 9.94 (Criterio de la hessiana)**.** *Sea f de clase C^2 en un entorno de un punto crítico a (es decir, $\nabla f(a)=0$) y sea $H=D^2f(a)$ su matriz hessiana, que es simétrica por el teorema de Clairaut. Entonces:*

1. *si H es definida positiva, a es un mínimo local estricto;*
2. *si H es definida negativa, a es un máximo local estricto;*
3. *si H tiene autovalores de los dos signos, a es un punto de silla;*
4. *si H es semidefinida y singular, el criterio no decide.*

Demostración. Por el desarrollo de Taylor de orden dos con resto de Peano, y usando $\nabla f(a)=0$,

$$f(a+h)-f(a)=\tfrac12\,h^{\top}Hh+R(h),\qquad \frac{R(h)}{|h|^2}\xrightarrow[h\to0]{}0.$$

(1) Si H es definida positiva, por el teorema espectral (capítulo 7) sus autovalores $\lambda_1\le\cdots\le\lambda_n$ son todos positivos, y para todo h

$$h^{\top}Hh\ \ge\ \lambda_1|h|^2$$

—basta escribir h en la base ortonormal de autovectores: $h^{\top}Hh=\sum\lambda_i c_i^2\ge\lambda_1\sum c_i^2$—. Tómese $\delta>0$ tal que $|R(h)|<\frac{\lambda_1}{4}|h|^2$ para $0<|h|<\delta$, lo cual es posible porque $R(h)/|h|^2\to0$. Entonces

$$f(a+h)-f(a)\ \ge\ \frac{\lambda_1}{2}|h|^2-\frac{\lambda_1}{4}|h|^2=\frac{\lambda_1}{4}|h|^2>0,$$

y a es mínimo local estricto.

(2) Aplíquese (1) a $-f$.

(3) Sean v_+ y v_- autovectores unitarios de autovalores $\lambda_+>0$ y $\lambda_-<0$. A lo largo de la recta $h=tv_+$,

$$f(a+tv_+)-f(a)=\tfrac12\lambda_+t^2+R(tv_+)\ \ge\ \tfrac14\lambda_+t^2>0$$

para t pequeño, y análogamente $f(a+tv_-)-f(a)<0$. En cualquier entorno de a hay puntos por encima y por debajo: no es ni máximo ni mínimo.

(4) Los ejemplos $f=x^4+y^4$ (mínimo), $f=-x^4-y^4$ (máximo) y $f=x^3$ (ni una cosa ni la otra) tienen los tres la misma hessiana nula en el origen, luego ninguna conclusión puede extraerse de H sola. □

Regla: Clasificar un punto crítico en la práctica

1. Resolver $\nabla f=0$ para localizar los puntos críticos.
2. Calcular H en cada uno.

3. En dimensión 2, con $H = \begin{pmatrix} A & B \\ B & C \end{pmatrix}$ y $\det H = AC - B^2$: si $\det H > 0$ y $A > 0$, mínimo; si $\det H > 0$ y $A < 0$, máximo; si $\det H < 0$, silla; si $\det H = 0$, indeterminado.
4. En dimensión cualquiera, usar el *criterio de los menores principales de Sylvester*: H es definida positiva si y sólo si todos los menores principales dominantes son positivos; definida negativa si alternan de signo empezando por negativo.
5. Si el criterio no decide, estudiar f directamente a lo largo de curvas: rectas, parábolas $y = cx^2$. Un solo par de direcciones con signos opuestos basta para concluir silla.
6. No olvidar la frontera ni el infinito: el criterio sólo detecta extremos *locales* e *interiores*.

9.23. Dominio, gráfica, conjuntos de nivel y continuidad: el vocabulario que faltaba

La primera sección de este capítulo montó la topología de $\mathbb{R}^n$ —bolas, abiertos, compactos— y la segunda pasó directamente a los límites. Entre medias quedó sin decir algo elemental y que el capítulo usa en cada página: qué es exactamente el *dominio* de una función de varias variables, qué es su *gráfica*, qué son los *conjuntos de nivel* que aparecen en los multiplicadores de Lagrange y en el dibujo del gradiente, y qué significa que una función de varias variables sea *continua*. Los teoremas de Weierstrass, de Clairaut, de Fubini y de la función implícita tienen todos la palabra "continua" en sus hipótesis; conviene que esa palabra esté definida.

Idea clave

Intuición primero. Una función de dos variables es una *altura* definida sobre un trozo de suelo. El dominio es el trozo de suelo donde la fórmula tiene sentido; la gráfica es la sábana que queda por encima; y los conjuntos de nivel son las curvas que dibujaría un cartógrafo al cortar esa sábana con planos horizontales, exactamente igual que las curvas de nivel de un mapa topográfico. Todo lo que hace este capítulo se lee en ese mapa: el gradiente es perpendicular a las curvas de nivel y apunta cuesta arriba, los máximos y mínimos son curvas de nivel que se cierran alrededor de un punto, y un puerto de montaña es un punto de silla.

Definición 9.95 (Campo escalar, dominio, imagen y gráfica)**.** Un *campo escalar* de n variables es una función $f\colon D \subseteq \mathbb{R}^n \to \mathbb{R}$. El conjunto D es su *dominio*: cuando la función viene dada por una fórmula y no se especifica nada más, el dominio es por convenio el mayor subconjunto de $\mathbb{R}^n$ en el que la fórmula produce un número real. La *imagen* es $f(D) = \{f(\mathbf{x}) : \mathbf{x} \in D\} \subseteq \mathbb{R}$, y la *gráfica* de f es el subconjunto de $\mathbb{R}^{n+1}$

$$\operatorname{Graf}(f) = \{(\mathbf{x}, f(\mathbf{x})) : \mathbf{x} \in D\}.$$

Para $n = 2$ la gráfica es una superficie de $\mathbb{R}^3$; para $n = 3$ vive en $\mathbb{R}^4$ y ya no se dibuja, razón por la cual los conjuntos de nivel dejan de ser un adorno y pasan a ser la única herramienta visual disponible.

Regla: determinar el dominio de una función de varias variables

El dominio no se adivina mirando el resultado: se construye *intersecando* las condiciones que impone cada pieza de la fórmula.

1. **Localiza todos los denominadores** e impón que no se anulen. Cada denominador quita del plano una curva entera, no un punto: $x-y$ quita la recta $y = x$, y x^2+y^2-1 quita la circunferencia unidad.
2. **Localiza las raíces de índice par** e impón que el radicando sea ≥ 0. Aquí es donde aparecen los discos, los semiplanos y las regiones limitadas por parábolas.
3. **Localiza los logaritmos** e impón que el argumento sea estrictamente positivo. Ojo a la diferencia con el punto anterior: el logaritmo pide > 0 y la raíz cuadrada admite el $= 0$, y esa diferencia es exactamente la que separa un dominio abierto de uno cerrado.
4. **Localiza los arcosenos y arcocosenos** e impón que el argumento esté en $[-1, 1]$: dos desigualdades, no una.
5. **Interseca todas las condiciones.** El dominio es la intersección, nunca la unión. Si una sola condición falla en un punto, el punto está fuera.
6. **Dibuja la región y decide si es abierta, cerrada, acotada o compacta.** Esta lectura no es decorativa: la compacidad es la hipótesis del teorema de Weierstrass, y sin ella los problemas de extremos absolutos pueden no tener solución.
7. **Comprueba con puntos de prueba.** Toma un punto claramente dentro y otro claramente fuera, y sustitúyelos. Media docena de sustituciones detectan al instante una desigualdad girada del revés.

Ejemplo 9.96 (Cinco dominios determinados por escrito)**. (a) Un disco cerrado.** Sea $f(x, y) = \sqrt{9 - x^2 - y^2}$. La única condición es $9 - x^2 - y^2 \geq 0$, es decir $x^2 + y^2 \leq 9$: el disco cerrado de centro el origen y radio 3. Es cerrado y acotado, luego *compacto*: toda función continua sobre él alcanzará máximo y mínimo. La imagen es $[0, 3]$.

(b) Un semiplano abierto. Sea $f(x, y) = \ln(x + y - 1)$. La condición es $x + y - 1 > 0$, el semiplano abierto que queda por encima de la recta $x + y = 1$. Es abierto y no acotado; el punto (0,5; 0,5) está fuera porque ahí $x + y - 1 = 0$ y el logaritmo no existe, mientras que $(1, 1)$ está dentro con $f = \ln 1 = 0$.

(c) Dos condiciones que se cortan. Sea $f(x, y) = \dfrac{\sqrt{y - x^2}}{x - 1}$. La raíz pide $y \geq x^2$ (la región cerrada por encima de la parábola) y el denominador pide $x \neq 1$. El dominio es la región parabólica cerrada *menos* la recta vertical $x = 1$, que la atraviesa. El punto $(1, 4)$

cumple la primera condición y falla la segunda: está fuera. El punto $(2,3)$ falla la primera, porque $3<4$: también fuera. Y $(0,1)$ cumple ambas.

(d) Una corona. Sea $f(x,y)=\arcsin(x^2+y^2-2)$. La condición es $-1\le x^2+y^2-2\le 1$, es decir $1\le x^2+y^2\le 3$: la corona circular cerrada entre los radios 1 y $\sqrt{3}\approx 1{,}732051$. Compacta otra vez, pero no simplemente conexa: tiene un agujero.

(e) Tres variables. Sea $f(x,y,z)=\dfrac{1}{\sqrt{1-x^2-y^2-z^2}}$. Ahora la raíz está en el denominador, así que hace falta $1-x^2-y^2-z^2>0$ con desigualdad *estricta*: la bola abierta de radio 1. Ni cerrada ni compacta, y en efecto f no está acotada: tiende a $+\infty$ al acercarse a la esfera frontera.

Definición 9.97 (Conjuntos de nivel, curvas de nivel y trazas)**.** Dado $f\colon D\subseteq\mathbb{R}^n\to\mathbb{R}$ y un número $c\in\mathbb{R}$, el *conjunto de nivel* de altura c es

$$N_c=\{\mathbf{x}\in D:\ f(\mathbf{x})=c\}.$$

Para $n=2$ se le llama *curva de nivel* y para $n=3$ *superficie de nivel*. Se llama *traza* o *sección* de la gráfica al corte de $\mathrm{Graf}(f)$ con un plano coordenado: los cortes con $x=a$ o $y=b$ dan las curvas $z=f(a,y)$ y $z=f(x,b)$, cuyas pendientes son precisamente las derivadas parciales. Dos conjuntos de nivel de alturas distintas son siempre disjuntos, porque una función asigna un solo valor a cada punto: *las curvas de nivel nunca se cortan*, salvo en puntos donde f no esté definida.

Regla: dibujar el mapa de niveles de una función de dos variables

1. **Determina primero el dominio** con la regla anterior. Fuera de él no hay niveles que dibujar.
2. **Escribe la ecuación $f(x,y)=c$ y despeja o reconoce la curva.** Casi siempre sale una cónica, una recta o una exponencial disfrazada; el objetivo es identificar la familia, no resolver nada.
3. **Averigua para qué valores de c el nivel es no vacío.** Si $f=x^2+y^2$, no hay niveles con $c<0$; el rango de c admisible *es* la imagen de f.
4. **Dibuja varios niveles con alturas equiespaciadas** ($c=0,1,2,3,\dots$). El equiespaciado es lo que hace legible el mapa: donde las curvas se apiñan, la función es empinada; donde se separan, es plana.
5. **Marca los niveles degenerados.** Un nivel puede reducirse a un punto (el fondo de un cuenco) o cruzarse consigo mismo (un punto de silla). Esos dos accidentes señalan los puntos críticos antes de derivar nada.
6. **Comprueba la ortogonalidad con el gradiente** en un punto concreto: $\operatorname{grad} f$ debe salir perpendicular al nivel que pasa por él y apuntar hacia los valores mayores de c.
7. **Reconstruye la gráfica mentalmente** apilando los niveles a las alturas correspondientes. Es exactamente lo que hace un cartógrafo, al revés.

Ejemplo 9.98 (Cuatro mapas de niveles leídos entero)**. (a) El cuenco.** $f(x,y) = x^2+y^2$. El nivel c es $x^2+y^2 = c$: vacío si $c<0$, el punto $(0,0)$ si $c=0$, y la circunferencia de radio $\sqrt{c}$ si $c>0$. Para $c=5$ el radio es $\sqrt{5} \approx 2{,}236068$. Los radios $\sqrt{c}$ crecen *cada vez más despacio*, de modo que a alturas equiespaciadas las circunferencias se apiñan al alejarse: la superficie se empina, como corresponde a un paraboloide. El gradiente $(2x, 2y)$ es radial, y en efecto es perpendicular a las circunferencias y apunta hacia afuera, hacia los c mayores.

(b) La silla. $f(x,y) = x^2 - y^2$. El nivel $c \neq 0$ es la hipérbola $x^2-y^2 = c$, con las ramas abiertas a izquierda y derecha si $c>0$ y arriba y abajo si $c<0$. Y el nivel $c=0$ es $x^2 = y^2$, es decir el *par de rectas* $y = \pm x$: un nivel que se cruza consigo mismo. Ese cruce es la firma visual de un punto de silla, y ocurre exactamente en el punto crítico $(0,0)$.

(c) Un plano inclinado. $f(x,y) = 2x+3y$. Los niveles son las rectas paralelas $2x+3y = c$, todas con la misma pendiente $-2/3$, y equiespaciadas para c equiespaciado: la función crece a ritmo constante, como debe hacer una función lineal. El gradiente es el vector constante $(2,3)$, perpendicular a todas esas rectas, y $\|\operatorname{grad} f\| = \sqrt{13} \approx 3{,}605551$ es la tasa de crecimiento en la dirección de máxima pendiente.

(d) Una elipse por nivel. $f(x,y) = x^2+3y^2$. El nivel $c>0$ es $\dfrac{x^2}{c} + \dfrac{y^2}{c/3} = 1$, una elipse de semiejes $\sqrt{c}$ y $\sqrt{c/3}$. Para $c=12$ los semiejes son $\sqrt{12} \approx 3{,}464102$ y 2: la elipse es $\sqrt{3}$ veces más ancha que alta, siempre, sea cual sea c. Que todos los niveles sean semejantes entre sí es lo que caracteriza a las formas cuadráticas, y la razón $\sqrt{3}$ entre los semiejes es la raíz del cociente de los autovalores 3 y 1 de la matriz asociada.

Definición 9.99 (Continuidad de un campo escalar y de un campo vectorial)**.** Sea $f\colon D \subseteq \mathbb{R}^n \to \mathbb{R}$ y $\mathbf{a} \in D$. Se dice que f es *continua en* $\mathbf{a}$ si para todo $\varepsilon > 0$ existe $\delta > 0$ tal que

$$\mathbf{x} \in D,\ \|\mathbf{x}-\mathbf{a}\| < \delta \implies |f(\mathbf{x}) - f(\mathbf{a})| < \varepsilon;$$

equivalentemente, si $\mathbf{a}$ es punto de acumulación de D, si $\lim_{\mathbf{x}\to\mathbf{a}} f(\mathbf{x}) = f(\mathbf{a})$. Se dice que f es continua en D si lo es en todos sus puntos. Una aplicación $\mathbf{F}\colon D \subseteq \mathbb{R}^n \to \mathbb{R}^m$ es continua en $\mathbf{a}$ si lo es cada una de sus m componentes, lo cual equivale a la misma definición con $\|\mathbf{F}(\mathbf{x}) - \mathbf{F}(\mathbf{a})\|$ en lugar del valor absoluto.

Obsérvese que la definición es *literalmente* la de una variable con $|x-a|$ sustituido por $\|\mathbf{x}-\mathbf{a}\|$. Todo lo que cambia es la forma del entorno: en la recta era un intervalo, aquí es una bola.

Teorema 9.100 (Álgebra de las funciones continuas)**.** *Si f y g son continuas en $\mathbf{a}$, también lo son $f+g$, $f-g$, fg y, si $g(\mathbf{a}) \neq 0$, el cociente f/g. En consecuencia, todo polinomio en n variables es continuo en todo $\mathbb{R}^n$, y toda función racional es continua en todos los puntos donde el denominador no se anula.*

Demostración. Basta repetir las demostraciones de una variable, que no usan de $\mathbb{R}$ nada más que la desigualdad triangular. Para la suma: dado $\varepsilon > 0$, tómense δ_1 y δ_2 que fuercen $|f(\mathbf{x}) - f(\mathbf{a})| < \varepsilon/2$ y $|g(\mathbf{x}) - g(\mathbf{a})| < \varepsilon/2$, y tómese $\delta = \min(\delta_1, \delta_2)$; entonces

$$|(f+g)(\mathbf{x}) - (f+g)(\mathbf{a})| \leq |f(\mathbf{x}) - f(\mathbf{a})| + |g(\mathbf{x}) - g(\mathbf{a})| < \varepsilon.$$

Para el producto se escribe

$$f(\mathbf{x})g(\mathbf{x}) - f(\mathbf{a})g(\mathbf{a}) = f(\mathbf{x})\big(g(\mathbf{x}) - g(\mathbf{a})\big) + g(\mathbf{a})\big(f(\mathbf{x}) - f(\mathbf{a})\big),$$

y se usa que f está acotada cerca de $\mathbf{a}$ (consecuencia inmediata de su continuidad, tomando $\varepsilon = 1$). El cociente se reduce al producto probando primero que $1/g$ es continua donde $g(\mathbf{a}) \neq 0$, lo cual sale de $\left|\frac{1}{g(\mathbf{x})} - \frac{1}{g(\mathbf{a})}\right| = \frac{|g(\mathbf{a}) - g(\mathbf{x})|}{|g(\mathbf{x})g(\mathbf{a})|}$ y de que $|g(\mathbf{x})| \geq |g(\mathbf{a})|/2$ cerca de $\mathbf{a}$. Finalmente, las proyecciones $\mathbf{x} \mapsto x_i$ son continuas (con $\delta = \varepsilon$, pues $|x_i - a_i| \leq \|\mathbf{x} - \mathbf{a}\|$) y las constantes también; sumando y multiplicando se obtienen todos los polinomios. □

Teorema 9.101 (Composición)**.** *Si* $\mathbf{g}\colon D \subseteq \mathbb{R}^n \to \mathbb{R}^m$ *es continua en* $\mathbf{a}$ *y* $h\colon E \subseteq \mathbb{R}^m \to \mathbb{R}$ *es continua en* $\mathbf{b} = \mathbf{g}(\mathbf{a})$*, con* $\mathbf{g}(D) \subseteq E$*, entonces* $h \circ \mathbf{g}$ *es continua en* $\mathbf{a}$*.*

Demostración. Dado $\varepsilon > 0$, la continuidad de h en $\mathbf{b}$ da un $\eta > 0$ con $\|\mathbf{y} - \mathbf{b}\| < \eta \Rightarrow |h(\mathbf{y}) - h(\mathbf{b})| < \varepsilon$. La continuidad de $\mathbf{g}$ en $\mathbf{a}$ da entonces un $\delta > 0$ con $\|\mathbf{x} - \mathbf{a}\| < \delta \Rightarrow \|\mathbf{g}(\mathbf{x}) - \mathbf{b}\| < \eta$. Encadenando las dos implicaciones se obtiene la definición para $h \circ \mathbf{g}$. □

Estos dos teoremas son los que justifican el paso 1 de la regla para decidir límites: en cuanto la función es un cociente de polinomios, una exponencial, un seno o una composición de todo eso, y el denominador no se anula, el límite es el valor y no hay nada más que hacer. El trabajo duro solo empieza en los puntos donde la fórmula se rompe.

Ejemplo 9.102 (Continuidad decidida en cinco funciones)**. (a) Continuidad gratis.** $f(x,y) = \dfrac{x^2 + \sin(xy)}{1 + x^2 + y^2}$ es continua en todo $\mathbb{R}^2$: numerador y denominador son continuos por los teoremas 9.100 y 9.101, y el denominador vale al menos 1, luego nunca se anula.

(b) Continua tras rellenar el agujero. Sea $f(x,y) = \dfrac{x^3 - y^3}{x^2 + y^2}$ con $f(0,0) = 0$. Fuera del origen es un cociente de polinomios con denominador no nulo: continua. En el origen, con polares,

$$|f| = \left|\frac{r^3(\cos^3\theta - \sin^3\theta)}{r^2}\right| = r\,|\cos^3\theta - \sin^3\theta| \leq 2r \xrightarrow[r\to 0]{} 0,$$

y la cota $2r$ no depende de θ: el límite es $0 = f(0,0)$ y la función es continua en todo el plano.

(c) Un agujero que no se puede rellenar. $f(x,y) = \dfrac{xy}{x^2 + y^2}$ para $(x,y) \neq (0,0)$. Sobre $y = mx$ vale $\frac{m}{1+m^2}$, que depende de m: el límite en el origen no existe y *ningún* valor que se le asigne a $f(0,0)$ la hace continua. La discontinuidad es esencial, no evitable.

(d) Continua y sin embargo no diferenciable. $f(x,y) = \sqrt{x^2 + y^2}$ es continua en todo el plano (composición de la raíz, continua en $[0,\infty)$, con un polinomio), y en el origen no es diferenciable: su gráfica es un cono con un pico. Continuidad y diferenciabilidad son dos escalones distintos, y el segundo es estrictamente más alto.

(e) Discontinua en toda una recta. $f(x,y) = \dfrac{1}{x - y}$ es continua en su dominio, que es $\mathbb{R}^2$ menos la recta $y = x$. Decir "f es discontinua en $y = x$" es un abuso de lenguaje: en esos puntos f ni siquiera está definida, y la continuidad solo se predica de puntos del dominio.

Cuidado

[Ser continua en cada variable no es ser continua] La función $f(x,y)=\frac{xy}{x^2+y^2}$ del apartado (c), con $f(0,0)=0$, es continua *en x con y fijo* y continua *en y con x fijo*, en todo punto, incluido el origen: sobre el eje x vale idénticamente 0, y sobre el eje y también. Y aun así no es continua en el origen como función de dos variables. La razón es la misma que con las derivadas parciales: mirar solo a lo largo de los ejes no dice nada sobre las diagonales, donde f vale constantemente $\frac{1}{2}$. La continuidad de la definición 9.99 es *conjunta*, no separada, y es la única que sirve para los teoremas.

Definición 9.103 (Las clases C^k)**.** Sea $U\subseteq\mathbb{R}^n$ abierto. Se dice que $f\colon U\to\mathbb{R}$ es de *clase* C^0 si es continua; de *clase* C^1 si existen todas sus derivadas parciales de primer orden en U y son continuas; y, recursivamente, de *clase* C^k si existen y son continuas todas sus derivadas parciales hasta el orden k. Es de clase C^∞ si lo es de clase C^k para todo k. Una aplicación $\mathbf{F}\colon U\to\mathbb{R}^m$ es de clase C^k si lo es cada componente. Con este vocabulario, los tres resultados centrales del capítulo se enuncian en una línea: $C^1\Rightarrow$ diferenciable; $C^2\Rightarrow$ hessiana simétrica (Clairaut); y C^1 con jacobiana invertible $\Rightarrow$ inversa local C^1.

Pausa para pensar

Merece la pena fijar la jerarquía completa, porque cada flecha es un teorema distinto y ninguna se invierte:

$$C^\infty\subset\cdots\subset C^2\subset C^1\subset\{\text{diferenciables}\}\subset C^0\subset\{\text{con todas las parciales}\}?$$

La última inclusión está escrita con interrogante a propósito: es *falsa*. Existir todas las parciales no implica ser continua —el apartado (c) del ejemplo anterior tiene las dos parciales nulas en el origen y ni siquiera es continua ahí—, de modo que ese último conjunto no contiene a C^0 ni está contenido en él. Es la única irregularidad de la cadena, y es exactamente la que hace que el cálculo en varias variables no sea el cálculo en una variable repetido n veces.

9.24. Weierstrass demostrado, y el procedimiento completo para los extremos absolutos sobre un compacto

El teorema de Weierstrass en $\mathbb{R}^n$ se enunció en la primera sección sin demostración, y es el que sostiene todo lo demás: sin él, un problema de optimización puede no tener solución y buscarla es perder el tiempo. Su demostración es corta y solo usa Bolzano–Weierstrass, así que aquí está. Y a continuación está el procedimiento completo para hallar los extremos absolutos, que hasta ahora aparecía repartido en un paso suelto de dos reglas distintas.

Teorema 9.104 (Bolzano–Weierstrass en $\mathbb{R}^n$)**.** *Toda sucesión acotada de $\mathbb{R}^n$ tiene una subsucesión convergente. En consecuencia, toda sucesión contenida en un compacto K tiene una subsucesión que converge* a un punto de K.

Demostración. Sea $(\mathbf{x}_k)$ acotada, con $\mathbf{x}_k = (x_k^{(1)}, \dots, x_k^{(n)})$. Como $|x_k^{(i)}| \le \|\mathbf{x}_k\| \le M$, cada sucesión de coordenadas es una sucesión real acotada. Por el Bolzano–Weierstrass de una variable, hay una subsucesión $(\mathbf{x}_{k_j})$ cuya primera coordenada converge. De ella se extrae otra subsucesión cuya segunda coordenada converge —y la primera sigue convergiendo, porque toda subsucesión de una sucesión convergente converge al mismo límite—, y se repite n veces. Tras n extracciones queda una subsucesión en la que las n coordenadas convergen a la vez, digamos a $\ell_1, \dots, \ell_n$; y como

$$\|\mathbf{x}_{k_j} - \boldsymbol{\ell}\| \le \sum_{i=1}^{n} |x_{k_j}^{(i)} - \ell_i| \xrightarrow[j\to\infty]{} 0,$$

esa subsucesión converge a $\boldsymbol{\ell}$ en $\mathbb{R}^n$. El número de extracciones es finito, y ahí está toda la gracia del argumento. Si además todos los términos están en un compacto K, que es cerrado, el límite pertenece a K. □

Teorema 9.105 (Weierstrass, demostrado)**.** *Sea $K \subset \mathbb{R}^n$ compacto no vacío y $f\colon K \to \mathbb{R}$ continua. Entonces f está acotada en K y alcanza en K su máximo y su mínimo: existen $\mathbf{p}, \mathbf{q} \in K$ con $f(\mathbf{q}) \le f(\mathbf{x}) \le f(\mathbf{p})$ para todo $\mathbf{x} \in K$.*

Demostración. Acotación. Supongamos que f no está acotada superiormente. Entonces para cada $k \in \mathbb{N}$ existe $\mathbf{x}_k \in K$ con $f(\mathbf{x}_k) > k$. Por el teorema 9.104 hay una subsucesión $\mathbf{x}_{k_j} \to \mathbf{a} \in K$, y por continuidad $f(\mathbf{x}_{k_j}) \to f(\mathbf{a})$, un número real. Pero $f(\mathbf{x}_{k_j}) > k_j \to \infty$: una sucesión no puede converger y tender a infinito a la vez. Contradicción, luego f está acotada superiormente; con $-f$ se obtiene la cota inferior.

El supremo se alcanza. Sea $S = \sup_K f$, que existe y es finito por lo anterior. Por definición de supremo, para cada k hay $\mathbf{y}_k \in K$ con $S - \frac{1}{k} < f(\mathbf{y}_k) \le S$. De nuevo por Bolzano–Weierstrass hay $\mathbf{y}_{k_j} \to \mathbf{p} \in K$, y por continuidad $f(\mathbf{p}) = \lim f(\mathbf{y}_{k_j}) = S$ por el teorema del sándwich. Es decir, el supremo es un valor efectivamente alcanzado, y por tanto un máximo. Aplicando esto a $-f$ se obtiene el mínimo. □

Cuidado

Las dos hipótesis se usan por separado y ninguna sobra. Sin *acotado*: $f(x, y) = x$ sobre todo $\mathbb{R}^2$ es continua y no tiene máximo. Sin *cerrado*: $f(x, y) = x$ sobre el disco abierto $x^2 + y^2 < 1$ es continua y acotada, con supremo 1, y no lo alcanza en ningún punto del disco abierto. Sin *continua*: la función que vale 0 en el disco cerrado salvo en el origen, donde vale 1... está acotada y alcanza sus extremos, pero basta cambiarla a $f(x, y) = 1/(x^2 + y^2)$ fuera del origen y $f(0, 0) = 0$ para tener un compacto y una función no acotada. Cada contraejemplo se construye rompiendo exactamente una pieza.

Regla: hallar los extremos absolutos de una función continua sobre un compacto

El error típico no es derivar mal: es olvidarse de la frontera, o tratar la frontera como si fuera interior.

1. **Comprueba primero que hay solución.** Verifica que el dominio es cerrado y

acotado y que f es continua en él. Solo entonces el teorema 9.105 garantiza que el máximo y el mínimo existen, y solo entonces tiene sentido comparar una lista finita de candidatos.

2. **Busca los puntos críticos del interior** resolviendo $\operatorname{grad} f = \mathbf{0}$, y *descarta los que caigan fuera* de la región o justo sobre la frontera. Este descarte se olvida a menudo y produce respuestas absurdas.

3. **No los clasifiques.** Aquí no hace falta la hessiana: solo se va a evaluar f en ellos y compararla. Calcular hessianas en este problema es trabajo tirado.

4. **Descompón la frontera en trozos suaves.** Un disco tiene un solo trozo; un cuadrado tiene cuatro lados y cuatro vértices; un triángulo, tres lados y tres vértices; una región mixta puede tener un arco y un segmento.

5. **Reduce cada trozo a una variable.** O bien se despeja ($y =$ algo, y se sustituye), o bien se parametriza ($x = \cos t$, $y = \sin t$). En cualquier caso queda una función de una variable sobre un intervalo cerrado, y se le aplica el procedimiento de una variable: derivar, igualar a cero, y añadir los extremos del intervalo.

6. **Añade siempre las esquinas.** Los puntos donde dos trozos de frontera se juntan son candidatos por derecho propio: ahí la parametrización se rompe y ninguna derivada los detecta.

7. **Evalúa f en todos los candidatos y ordena la lista.** El mayor valor es el máximo absoluto y el menor el mínimo absoluto. No hay más criterio ni hace falta ninguno: Weierstrass ya hizo el trabajo conceptual.

8. **Controla el resultado.** El máximo absoluto nunca puede ser menor que f en un punto cualquiera elegido al azar de la región; evalúa f en dos o tres puntos sueltos y comprueba que quedan entre el mínimo y el máximo hallados.

Alternativa al paso 5: usar multiplicadores de Lagrange sobre la frontera, con g la ecuación del trozo. Es equivalente, y conviene cuando despejar es incómodo.

Ejemplo 9.106 (Un triángulo, resuelto entero)**.** Hállense los extremos absolutos de $f(x, y) = x^2 + y^2 - 2x - 4y$ sobre el triángulo cerrado T de vértices $(0, 0)$, $(4, 0)$ y $(0, 4)$, es decir sobre $\{x \geq 0,\ y \geq 0,\ x + y \leq 4\}$.

Paso 1. T es cerrado (intersección de tres semiplanos cerrados) y acotado, luego compacto; f es un polinomio, luego continua. Existen máximo y mínimo.

Paso 2, interior. $\operatorname{grad} f = (2x - 2,\ 2y - 4) = \mathbf{0}$ da el único punto crítico $(1, 2)$, que cumple $1 > 0$, $2 > 0$ y $1 + 2 = 3 < 4$: está *dentro*. Su valor es

$$f(1, 2) = 1 + 4 - 2 - 8 = -5.$$

Pasos 4 y 5, frontera. Tres lados. *Lado $y = 0$, con $0 \leq x \leq 4$:* $f = x^2 - 2x$, cuya derivada $2x - 2$ se anula en $x = 1$, con $f(1, 0) = -1$. *Lado $x = 0$, con $0 \leq y \leq 4$:* $f = y^2 - 4y$,

con derivada $2y-4$ nula en $y=2$ y $f(0,2)=-4$. *Hipotenusa* $x+y=4$, *con* $0\le x\le 4$: sustituyendo $y=4-x$,

$$f = x^2+(4-x)^2-2x-4(4-x) = 2x^2-6x,$$

cuya derivada $4x-6$ se anula en $x=\frac{3}{2}$, con $y=\frac{5}{2}$ y valor $2\cdot\frac{9}{4}-9=-\frac{9}{2}=-4{,}5$.

Paso 6, esquinas. $f(0,0)=0$, $f(4,0)=16-8=8$, $f(0,4)=16-16=0$.

Paso 7, la lista. Los candidatos y sus valores son

$$(1,2)\to -5,\qquad (1,0)\to -1,\qquad (0,2)\to -4,\qquad \left(\tfrac{3}{2},\tfrac{5}{2}\right)\to -4{,}5,$$

$$(0,0)\to 0,\qquad (4,0)\to 8,\qquad (0,4)\to 0.$$

El **mínimo absoluto** es -5, alcanzado en $(1,2)$, y el **máximo absoluto** es 8, alcanzado en el vértice $(4,0)$.

Paso 8, control. Completando cuadrados, $f=(x-1)^2+(y-2)^2-5$: la función es la distancia al cuadrado al punto $(1,2)$ menos 5. El mínimo debe estar en $(1,2)$ con valor -5 ✓, y el máximo en el punto de T más lejano de $(1,2)$, que es el vértice $(4,0)$ a distancia $\sqrt{9+4}=\sqrt{13}$, con valor $13-5=8$ ✓. Los dos resultados salen de nuevo, por un camino completamente distinto. Nótese además que el máximo está en una esquina, donde no se anula ninguna derivada: sin el paso 6 se habría perdido.

Ejemplo 9.107 (Un disco, con la frontera parametrizada)**.** Hállense los extremos absolutos de $f(x,y)=x^2-y^2+2x$ sobre el disco cerrado $x^2+y^2\le 4$.

Interior. $\operatorname{grad} f=(2x+2,\ -2y)=\mathbf{0}$ en $(-1,0)$, que está dentro porque $1<4$. Su valor es $f(-1,0)=1-0-2=-1$. La hessiana es $\left(\begin{smallmatrix}2&0\\0&-2\end{smallmatrix}\right)$, indefinida: es una silla. Detalle importante: *una silla también entra en la lista de candidatos*, aunque no vaya a ganar.

Frontera. Se parametriza con $x=2\cos t$, $y=2\sin t$, $t\in[0,2\pi]$:

$$\varphi(t)=4\cos^2 t-4\sin^2 t+4\cos t=4\cos 2t+4\cos t.$$

Conviene escribirlo en función de $c=\cos t\in[-1,1]$ usando $\cos 2t=2c^2-1$:

$$\varphi = 8c^2+4c-4,$$

una parábola en c con vértice en $c=-\frac{4}{16}=-\frac{1}{4}$. Los tres candidatos son $c=-\frac{1}{4}$, con valor $8\cdot\frac{1}{16}-1-4=-\frac{9}{2}=-4{,}5$; $c=1$, con valor $8+4-4=8$; y $c=-1$, con valor $8-4-4=0$. El caso $c=-\frac{1}{4}$ corresponde a $x=-\frac{1}{2}$ e $y=\pm\sqrt{4-\frac{1}{4}}=\pm\frac{\sqrt{15}}{2}\approx\pm 1{,}936492$, dos puntos simétricos; $c=1$ es $(2,0)$ y $c=-1$ es $(-2,0)$.

La lista. -1 en $(-1,0)$; $-4{,}5$ en $\left(-\frac{1}{2},\pm\frac{\sqrt{15}}{2}\right)$; 8 en $(2,0)$; 0 en $(-2,0)$. El **máximo absoluto** es 8 en $(2,0)$ y el **mínimo absoluto** es $-\frac{9}{2}$ en los dos puntos $\left(-\frac{1}{2},\pm\frac{\sqrt{15}}{2}\right)$.

Control. Un barrido numérico de la frontera con doscientos mil valores de t da un mínimo de $-4{,}499\,999\,99$ y un máximo de $8{,}000\,000$: coincide con lo calculado a mano hasta el último decimal disponible. Y el valor $f(0,0)=0$ de un punto interior elegido al azar queda, como debe, entre $-4{,}5$ y 8.

Comparación con Lagrange. La misma frontera con multiplicadores: $\operatorname{grad} f = \lambda \operatorname{grad} g$ con $g = x^2 + y^2$ da $2x + 2 = 2\lambda x$ y $-2y = 2\lambda y$. La segunda ecuación se parte en dos ramas, $y = 0$ o $\lambda = -1$. Con $y = 0$ salen $(\pm 2, 0)$; con $\lambda = -1$, la primera ecuación da $2x + 2 = -2x$, es decir $x = -\frac{1}{2}$, y la restricción devuelve $y = \pm\frac{\sqrt{15}}{2}$. Los mismos cuatro candidatos, obtenidos sin parametrizar nada. Que dos métodos independientes produzcan exactamente la misma lista es la mejor comprobación posible.

Pausa para pensar

Compárense los dos ejemplos. En el primero, el máximo estaba en una *esquina* de la frontera; en el segundo, la frontera no tenía esquinas y el máximo estaba en un punto donde la restricción era tangente a una curva de nivel. Esa es la diferencia esencial entre optimizar sobre un polígono y optimizar sobre una curva suave, y es también la razón de que la programación lineal tenga un teorema propio: cuando la función objetivo es lineal y la región es un polígono, el óptimo está *siempre* en un vértice, y basta con evaluar en ellos. El método del símplex no es más que una manera inteligente de recorrer esos vértices sin visitarlos todos.

9.25. Ejercicios resueltos

Ejercicio 9.1. Estudia $\displaystyle\lim_{(x,y)\to(0,0)} \frac{x^2 y}{x^4 + y^2}$ y $\displaystyle\lim_{(x,y)\to(0,0)} \frac{\sin(xy)}{x^2 + y^2}$.

Solución. *(a)* Sobre toda recta $y = mx$: $\frac{mx^3}{x^4 + m^2x^2} \to 0$; sobre los ejes, 0. Pero sobre $y = x^2$: $\frac{x^4}{2x^4} = \frac{1}{2}$. Dos caminos, dos valores: **no existe**. (Las rectas no bastan nunca.) *(b)* Sobre $y = x$: $\frac{\sin(x^2)}{2x^2} \to \frac{1}{2}$; sobre $y = 0$: 0. **No existe** tampoco. ■

Ejercicio 9.2. La temperatura en el plano es $T(x, y) = 20 + xe^{y - x^2}$. Estando en $(1, 1)$: ¿en qué dirección conviene moverse para calentarse lo más rápido posible, y a qué tasa? ¿En qué direcciones la temperatura no cambia?

Solución. $\nabla T = \big(e^{y-x^2}(1 - 2x^2),\ xe^{y-x^2}\big)$, y en $(1, 1)$: $\nabla T = (-1, 1)$ (pues $e^0 = 1$). Máximo crecimiento en la dirección unitaria $\frac{1}{\sqrt{2}}(-1, 1)$, con tasa $\|\nabla T\| = \sqrt{2}$ grados por unidad de longitud. La temperatura no cambia en las direcciones ortogonales al gradiente: $\pm\frac{1}{\sqrt{2}}(1, 1)$, tangentes a la isoterma que pasa por el punto. ■

Ejercicio 9.3. Halla y clasifica todos los puntos críticos de $f(x, y) = x^3 + y^3 - 3xy$.

Solución. $\nabla f = (3x^2 - 3y,\ 3y^2 - 3x) = \mathbf{0}$: $y = x^2$ y $x = y^2$, luego $x = x^4$: $x = 0$ o $x = 1$. Críticos: $(0, 0)$ y $(1, 1)$. Hessiana: $H = \begin{pmatrix} 6x & -3 \\ -3 & 6y \end{pmatrix}$. En $(0, 0)$: $H = \begin{pmatrix} 0 & -3 \\ -3 & 0 \end{pmatrix}$, $\det = -9 < 0$: **silla**. En $(1, 1)$: $H = \begin{pmatrix} 6 & -3 \\ -3 & 6 \end{pmatrix}$, $\det = 27 > 0$, $f_{xx} = 6 > 0$: **mínimo local**, $f(1, 1) = -1$. No es global: $f(x, 0) = x^3 \to -\infty$. ■

Ejercicio 9.4. Halla los extremos absolutos de $f(x, y) = x^2 + 2y^2 - x$ sobre el disco $x^2 + y^2 \le 1$.

Solución. Interior: $\nabla f = (2x-1, 4y) = \mathbf{0}$ da $(\frac{1}{2}, 0)$, dentro del disco, con $f = -\frac{1}{4}$. *Frontera* $x^2+y^2=1$: sustituimos $y^2 = 1-x^2$: $g(x) = x^2 + 2(1-x^2) - x = 2 - x - x^2$, $x \in [-1,1]$. $g' = -1 - 2x = 0$ en $x = -\frac{1}{2}$: $g(-\frac{1}{2}) = \frac{9}{4}$; extremos del intervalo: $g(1) = 0$, $g(-1) = 2$. Comparando: mínimo absoluto $-\frac{1}{4}$ en $(\frac{1}{2}, 0)$; máximo absoluto $\frac{9}{4}$ en $(-\frac{1}{2}, \pm\frac{\sqrt{3}}{2})$. (Compacidad + Weierstrass garantizaban su existencia.) ■

Ejercicio 9.5. Con multiplicadores de Lagrange, halla la distancia mínima del punto $(1, 0)$ a la parábola $y^2 = x$.

Solución. Minimizamos $f = (x-1)^2 + y^2$ con $g = y^2 - x = 0$. Sistema: $2(x-1) = -\lambda$, $2y = 2\lambda y$, $y^2 = x$. La segunda: $y(1-\lambda) = 0$. *Caso* $y = 0$: $x = 0$, $f = 1$. *Caso* $\lambda = 1$: $2(x-1) = -1$: $x = \frac{1}{2}$, $y^2 = \frac{1}{2}$, $f = \frac{1}{4} + \frac{1}{2} = \frac{3}{4}$. Mínimo: $f = \frac{3}{4}$, distancia $\frac{\sqrt{3}}{2}$ en los puntos $(\frac{1}{2}, \pm\frac{1}{\sqrt{2}})$. (El caso $y = 0$ es un máximo local de la distancia restringida.) ■

Ejercicio 9.6. Calcula $\iiint_E z \, \mathrm{d}V$ donde E es la región sobre el cono $z = \sqrt{x^2+y^2}$ y bajo la esfera $x^2 + y^2 + z^2 = 1$.

Solución. En esféricas, el cono es $\phi = \frac{\pi}{4}$ y la esfera $\rho = 1$; la región es $0 \le \rho \le 1$, $0 \le \phi \le \frac{\pi}{4}$. Con $z = \rho\cos\phi$ y $J = \rho^2 \sin\phi$:

$$\int_0^{2\pi}\int_0^{\pi/4}\int_0^1 \rho\cos\phi \cdot \rho^2 \sin\phi \, \mathrm{d}\rho \, \mathrm{d}\phi \, \mathrm{d}\theta = 2\pi \cdot \frac{1}{4} \cdot \Big[\frac{\sin^2\phi}{2}\Big]_0^{\pi/4} = 2\pi \cdot \frac{1}{4} \cdot \frac{1}{4} = \frac{\pi}{8}.$$

■

Ejercicio 9.7. Calcula $\iint_R (x+y)^2 \, \mathrm{d}A$ donde R es el cuadrado de vértices $(1,0), (0,1), (-1,0), (0,-1)$, con el cambio $u = x + y$, $v = x - y$.

Solución. El cambio inverso es $x = \frac{u+v}{2}$, $y = \frac{u-v}{2}$, con jacobiano $\det \begin{pmatrix} 1/2 & 1/2 \\ 1/2 & -1/2 \end{pmatrix} = -\frac{1}{2}$, luego $\mathrm{d}A = \frac{1}{2}\mathrm{d}u\,\mathrm{d}v$. El cuadrado $|x| + |y| \le 1$ se convierte en $|u| \le 1$, $|v| \le 1$. Entonces

$$\iint_R (x+y)^2 \mathrm{d}A = \frac{1}{2}\int_{-1}^1\int_{-1}^1 u^2 \, \mathrm{d}u \, \mathrm{d}v = \frac{1}{2} \cdot \frac{2}{3} \cdot 2 = \frac{2}{3}.$$

El cambio de variables convirtió un cuadrado girado en uno alineado: elegir coordenadas adaptadas a la región es la mitad del trabajo. ■

Ejercicio 9.8. El sistema $\begin{cases} u = x^2 - y^2 \\ v = 2xy \end{cases}$ (¡la función compleja z^2!) ¿es invertible cerca de $(x, y) = (1, 1)$? Calcula la jacobiana de la inversa en la imagen de ese punto.

Solución. $D\mathbf{F} = \begin{pmatrix} 2x & -2y \\ 2y & 2x \end{pmatrix}$, $\det = 4(x^2+y^2)$, que en $(1,1)$ vale $8 \neq 0$: el teorema de la función inversa garantiza inversa local C^1 cerca de $\mathbf{F}(1,1) = (0,2)$. Su jacobiana:

$$D(\mathbf{F}^{-1})(0,2) = \begin{pmatrix} 2 & -2 \\ 2 & 2 \end{pmatrix}^{-1} = \frac{1}{8}\begin{pmatrix} 2 & 2 \\ -2 & 2 \end{pmatrix} = \begin{pmatrix} 1/4 & 1/4 \\ -1/4 & 1/4 \end{pmatrix}.$$

(Globalmente z^2 no es inyectiva —$\pm z$ tienen igual cuadrado—: la inversa es solo *local*, como promete el teorema y no más.) ■

Ejercicio 9.9. Verifica la fórmula del laplaciano en polares (Ejemplo 9.25) para $f = x^2 - y^2$.

Solución. Directamente: $f_{xx} + f_{yy} = 2 - 2 = 0$. En polares, $f = r^2\cos^2\theta - r^2\sin^2\theta = r^2\cos 2\theta$:

$$f_r = 2r\cos 2\theta, \quad f_{rr} = 2\cos 2\theta, \quad f_{\theta\theta} = -4r^2\cos 2\theta,$$

$$f_{rr} + \frac{1}{r}f_r + \frac{1}{r^2}f_{\theta\theta} = 2\cos 2\theta + 2\cos 2\theta - 4\cos 2\theta = 0. \checkmark$$

Las funciones con $\Delta f = 0$ (*armónicas*) reaparecerán en el capítulo 12 como estados estacionarios del calor. ■

9.26. Ejercicios adicionales resueltos

Ejercicio 9.10. $\int_0^1 \frac{x^3-1}{\ln x}\mathrm{d}x$.

Solución. $I(t) = \ln(t+1)$ por Feynman: $I(3) = 2\ln 2$. ■

Ejercicio 9.11. Clasifica el origen para $f = x^2 + y^2 + z^2 - 2xyz$ y decide si es extremo global.

Solución. $H(0) = 2I$: mínimo local, $f = 0$; no global: sobre $x = y = z = t$, $f = 3t^2 - 2t^3 \to -\infty$. ■

Ejercicio 9.12. Maximiza $x + 2y + 3z$ en $\{x + y + z = 1\} \cap \{x^2 + y^2 + z^2 = 1\}$.

Solución. Lagrange doble: progresión aritmética con $y = \frac{1}{3}$, $d = \pm\frac{1}{\sqrt{3}}$: máximo $2 + \frac{2}{\sqrt{3}}$; compacidad (circunferencia) garantiza. ■

Ejercicio 9.13. Centroide del semidisco de radio R.

Solución. $\bar{y} = \frac{2R^3/3}{\pi R^2/2} = \frac{4R}{3\pi}$; con Pappus regala $V_{\text{esfera}} = \frac{4}{3}\pi R^3$. ■

Ejercicio 9.14. $\iint_{\mathbb{R}^2}(1 + x^2 + y^2)^{-p}$: ¿cuándo converge y cuánto vale con $p = 2$?

Solución. Polares: converge $\iff p > 1$; con $p = 2$ vale π. ■

Ejercicio 9.15. Verifica $f_{xy}(0,0) = -1 \neq 1 = f_{yx}(0,0)$ para $f = \frac{xy(x^2-y^2)}{x^2+y^2}$.

Solución. $f_x(0,y) = -y$ y $f_y(x,0) = x$ por definición de parcial; derivando, -1 y $+1$. Falla la *continuidad* de las segundas cruzadas: la hipótesis de Clairaut no es decorativa. ■

9.27. Taller de ejemplos explicados: parciales, gradiente y Lagrange

9.27.1. Derivadas parciales

Derivar *parcialmente* respecto de x significa tratar a y como si fuera una constante, y derivar normal.

Ejemplo 9.108. Hallar las dos parciales de $f(x,y) = x^2y + \sin(xy)$.

Parcial respecto de x (con y constante): la derivada de x^2y es $2xy$ (el y es constante que acompaña); la de $\sin(xy)$ es $\cos(xy)\cdot y$ por la regla de la cadena (lo de dentro xy derivado respecto de x da y):

$$\frac{\partial f}{\partial x} = 2xy + y\cos(xy).$$

Parcial respecto de y (con x constante): la derivada de x^2y es x^2; la de $\sin(xy)$ es $\cos(xy)\cdot x$:

$$\frac{\partial f}{\partial y} = x^2 + x\cos(xy).$$

9.27.2. El gradiente

El *gradiente* $\nabla f = \left(\dfrac{\partial f}{\partial x}, \dfrac{\partial f}{\partial y}\right)$ es el vector que apunta hacia donde f crece más rápido.

Ejemplo 9.109. Para $f(x,y) = x^2+y^2$, el gradiente es $\nabla f = (2x, 2y)$. En el punto $(1,2)$ vale $(2,4)$: desde ahí, la dirección de máximo ascenso es $(2,4)$ (hacia afuera, alejándose del origen, que es donde f es mínima). Tiene sentido: x^2+y^2 es un tazón y sube al alejarse del centro.

9.27.3. Multiplicadores de Lagrange: un problema completo

Ejemplo 9.110. Hallar el máximo de $f(x,y) = xy$ sujeto a la restricción $x + y = 10$. (Es decir: de todos los rectángulos con perímetro fijo, ¿cuál tiene área máxima?)

Paso 1. Plantear $\nabla f = \lambda \nabla g$, donde $g(x,y) = x+y$. Calculamos: $\nabla f = (y, x)$ y $\nabla g = (1,1)$. La condición da

$$(y,x) = \lambda(1,1) \implies y = \lambda \text{ y } x = \lambda.$$

Paso 2. Deducir que $x = y$**.** Como ambos valen λ, tenemos $x = y$.

Paso 3. Usar la restricción. Sustituimos $x = y$ en $x + y = 10$: $2x = 10$, así que $x = 5$ y $y = 5$.

Paso 4. Calcular el máximo: $f(5,5) = 25$. El área máxima es 25, y se logra con el cuadrado (¡$x = y$!). Este es el famoso resultado de que, a perímetro fijo, el cuadrado maximiza el área.

9.28. Ampliación del temario: coordenadas cilíndricas/esféricas y aplicaciones físicas

9.28.1. Integrales múltiples en coordenadas cilíndricas y esféricas

Idea clave

Intuición primero. El jacobiano es el "factor de estiramiento" de la rejilla de coordenadas. En polares, una celdita que abarca $\mathrm{d}r$ y $\mathrm{d}\theta$ no es un cuadrado: es un rectangulito curvo de lados $\mathrm{d}r$ y $r\,\mathrm{d}\theta$ (el arco crece con el radio), así que su área es $r\,\mathrm{d}r\,\mathrm{d}\theta$. Ese r extra no es un adorno: es cuánto se abre la rejilla lejos del centro. En esféricas, $\rho^2\sin\varphi$ hace el mismo trabajo en 3D.

Cuando una región tiene simetría de eje o de centro, cambiar a coordenadas cilíndricas o esféricas simplifica enormemente una integral triple. La clave es el *jacobiano*, el factor que corrige cómo se estira el elemento de volumen.

Regla: elementos de volumen

$$\text{Cilíndricas } (r,\theta,z): \quad \mathrm{d}V = r\,\mathrm{d}r\,\mathrm{d}\theta\,\mathrm{d}z$$

$$\text{Esféricas } (\rho,\varphi,\theta): \quad \mathrm{d}V = \rho^2\sin\varphi\,\mathrm{d}\rho\,\mathrm{d}\varphi\,\mathrm{d}\theta$$

(en esféricas, ρ es la distancia al origen, φ el ángulo desde el eje z, y θ el azimutal).

Ejemplo 9.111 (Volumen de una esfera por esféricas)**.** El volumen de la bola de radio R:

$$V = \int_0^{2\pi}\int_0^{\pi}\int_0^{R} \rho^2\sin\varphi\,\mathrm{d}\rho\,\mathrm{d}\varphi\,\mathrm{d}\theta = \underbrace{\int_0^{2\pi}\mathrm{d}\theta}_{2\pi}\underbrace{\int_0^{\pi}\sin\varphi\,\mathrm{d}\varphi}_{2}\underbrace{\int_0^{R}\rho^2\,\mathrm{d}\rho}_{R^3/3} = 2\pi\cdot 2\cdot\frac{R^3}{3} = \frac{4}{3}\pi R^3.$$

La fórmula clásica, deducida en tres integrales que se separan gracias a la simetría.

9.28.2. Aplicaciones físicas de la integral múltiple

Idea clave

Intuición primero. Las tres integrales responden preguntas físicas concretas. La *masa* suma densidad por volumen. El *centro de masa* es el punto de equilibrio: el promedio de las posiciones pesado por la masa (donde el cuerpo se balancearía sobre un dedo). El *momento de inercia* pesa por la distancia *al cuadrado* porque la masa lejos del eje es muchísimo más difícil de girar: al rotar, su velocidad es $v=\omega r$, y la energía crece como r^2.

Las integrales triples calculan cantidades físicas de cuerpos sólidos con densidad $\rho(x,y,z)$.

Regla: masa, centro de masa y momento de inercia

$$M = \iiint_V \rho\,\mathrm{d}V, \quad \bar{x} = \frac{1}{M}\iiint_V x\,\rho\,\mathrm{d}V, \quad I_{\text{eje}} = \iiint_V (\text{dist. al eje})^2\,\rho\,\mathrm{d}V.$$

El *centroide* es el centro de masa cuando la densidad es constante (el centro geométrico). El *momento de inercia* I mide la resistencia a girar: cuanta más masa lejos del eje, mayor I.

Ejemplo 9.112 (Momento de inercia de un disco)**.** Un disco uniforme de masa M y radio R, girando respecto de su centro. Con densidad superficial $\sigma = \dfrac{M}{\pi R^2}$ y usando polares (la distancia al eje es r):

$$I = \int_0^{2\pi}\int_0^{R} r^2\,\sigma\,r\,\mathrm{d}r\,\mathrm{d}\theta = \sigma\cdot 2\pi\cdot\frac{R^4}{4} = \frac{M}{\pi R^2}\cdot\frac{\pi R^4}{2} = \frac{1}{2}MR^2.$$

El resultado $I = \frac{1}{2}MR^2$ es el que aparece en toda la física de rotación (ruedas, poleas, engranajes).

9.29. Ejercicios propuestos

Sin solución desarrollada: solo enunciado y respuesta al final. Úsalos para comprobar que lo del capítulo se sostiene sin la muleta del texto delante.

9.1. Calcula $\frac{\partial f}{\partial x}$ y $\frac{\partial f}{\partial y}$ para $f(x,y) = x^2y + y^3$.

9.2. Halla $\operatorname{grad} f$ en $(1,2)$ para $f(x,y) = x^2 + y^2$ y di en qué dirección crece más rápido.

9.3. Calcula la derivada direccional de $f(x,y) = xy$ en $(1,1)$ según $\frac{1}{\sqrt{2}}(1,1)$.

9.4. Escribe el plano tangente a $z = x^2 + y^2$ en $(1,1,2)$.

9.5. Halla y clasifica los puntos críticos de $f(x,y) = x^2 + y^2 - 4x$.

9.6. Clasifica el punto crítico de $f(x,y) = x^2 - y^2$ en el origen.

9.7. Calcula $\displaystyle\iint_{[0,1]^2} xy \,\mathrm{d}A$.

9.8. Calcula $\displaystyle\iint_{x^2+y^2\le 1} (x^2+y^2)\,\mathrm{d}A$ pasando a polares.

9.9. ¿Cuál es el jacobiano del cambio a coordenadas polares? ¿Y a esféricas?

9.10. Maximiza xy sujeto a $x + y = 10$ con multiplicadores de Lagrange.

9.11. Calcula el volumen de la bola unidad de $\mathbb{R}^3$ con una integral triple.

9.12. Si $z = f(x,y)$ con $x = t^2$, $y = t^3$, escribe $\frac{\mathrm{d}z}{\mathrm{d}t}$ por la regla de la cadena.

Respuestas

1. $2xy$ y $x^2 + 3y^2$. **2.** $(2,4)$; crece más rápido en esa dirección, con rapidez $\sqrt{20}$. **3.** $\sqrt{2}$. **4.** $z = 2x + 2y - 2$. **5.** Único punto crítico $(2,0)$; mínimo (hessiana definida positiva). **6.** Punto silla. **7.** $\frac{1}{4}$. **8.** $\frac{\pi}{2}$. **9.** r y $\rho^2 \sin\varphi$. **10.** $x = y = 5$, máximo 25. **11.** $\frac{4\pi}{3}$. **12.** $\frac{\mathrm{d}z}{\mathrm{d}t} = 2t\,\frac{\partial f}{\partial x} + 3t^2\,\frac{\partial f}{\partial y}$.

Capítulo 10

Coordenadas polares y ecuaciones paramétricas

Lo que necesitas traer

Del capítulo 6: integral definida, longitud de arco y la regla de la cadena. Del capítulo 4: la forma polar de un número complejo, que es literalmente el mismo cambio de coordenadas. Del capítulo 9: el jacobiano, para entender de dónde sale el factor r al cambiar a polares.

Idea clave

Hasta ahora, describir un punto ha significado siempre lo mismo: decir cuánto hay que avanzar a la derecha y cuánto hacia arriba. Es un buen sistema, pero no es el único ni siempre el mejor. Cuando un radar detecta un avión no dice "está 40 km al este y 30 al norte": dice "está a 50 km, en el rumbo 53 grados". Y cuando queremos describir el movimiento de una piedra lanzada al aire, ni las coordenadas cartesianas ni las polares bastan por sí solas: hace falta una tercera idea, la de que x e y dependen ambas del *tiempo*. Este capítulo presenta esas dos maneras alternativas de mirar el plano. Ninguna sustituye a la cartesiana; las tres juntas son mucho más potentes que cualquiera por separado.

Este capítulo se apoya en la geometría analítica de la geometría analítica y en la trigonometría de la trigonometría elemental, y a su vez prepara el terreno para el cálculo multivariable del capítulo 9 y para el cálculo vectorial del capítulo 11, donde las curvas parametrizadas son el objeto central y las coordenadas polares el cambio de variable más usado del libro.

10.1. Dónde está un punto: dos respuestas posibles

Definición 10.1 (Coordenadas polares)**.** Fijamos un punto O (el **polo**, que será el origen) y una semirrecta que sale de él (el **eje polar**, que será el semieje positivo de las x). Las coordenadas polares de un punto $P \neq O$ son el par (r, θ) donde

$r = \overline{OP} > 0$ es la distancia al polo, θ es el ángulo que forma $\overline{OP}$ con el eje polar.

La relación con las cartesianas es inmediata y sale de la definición misma de seno y coseno en la circunferencia:

Regla: conversión entre polares y cartesianas

$$\boxed{x = r\cos\theta, \qquad y = r\sin\theta} \qquad \boxed{r = \sqrt{x^2+y^2}, \qquad \tan\theta = \frac{y}{x}}$$

De polares a cartesianas nunca hay problema: se sustituye y ya está. **De cartesianas a polares sí lo hay**: la tangente no distingue el cuadrante. El procedimiento correcto es:

1. Calcule $r = \sqrt{x^2+y^2}$.
2. Calcule el ángulo de referencia $\arctan|y/x|$.
3. **Dibuje el punto** y corrija el ángulo según el cuadrante: sume π en el segundo y tercer cuadrante, sume 2π en el cuarto si quiere el ángulo positivo.
4. Si $x = 0$, no divida: el ángulo es $\pi/2$ (si $y > 0$) o $3\pi/2$ (si $y < 0$).

Ejemplo 10.2. ▪ $(1,1)$: $r = \sqrt{2}$ y el punto está en el primer cuadrante, luego $\theta = \pi/4$. Polares: $(\sqrt{2},\, \pi/4)$.

- $(-1,1)$: $r = \sqrt{2}$; $\arctan|1/(-1)| = \pi/4$, pero el punto está en el segundo cuadrante, así que $\theta = \pi - \pi/4 = 3\pi/4$. Si uno se fía ciegamente de la calculadora, $\arctan(-1) = -\pi/4$, que apunta al cuarto cuadrante: mal.
- $(0,-3)$: $r = 3$, $\theta = 3\pi/2$ (o $-\pi/2$).
- $(\,2,\, \pi/3\,)$ en polares es $x = 2\cos(\pi/3) = 1$, $y = 2\sin(\pi/3) = \sqrt{3}$.

Cuidado

[Las polares no son únicas] El punto $(\sqrt{2}, \pi/4)$ es también $(\sqrt{2},\, \pi/4+2\pi)$, y $(\sqrt{2},\, \pi/4-2\pi)$, y así infinitas veces. Peor todavía: si se admite $r < 0$ —cosa que se hace a veces, entendiendo que un radio negativo significa "avance en el sentido opuesto"— entonces también es $(-\sqrt{2},\, 5\pi/4)$. Y el origen tiene $r = 0$ con *cualquier* θ. En cartesianas cada punto tiene un nombre y solo uno; en polares tiene infinitos. Esto no es un defecto que haya que arreglar, es el precio de la comodidad, pero obliga a un cuidado especial al resolver ecuaciones: dos curvas polares pueden cortarse en un punto sin que exista ningún θ que satisfaga las dos ecuaciones a la vez.

10.2. Coordenadas polares, curvas polares y curvas paramétricas: el vocabulario básico, definido desde cero

La sección anterior ha dado la definición de coordenadas polares con $r > 0$ y ha avisado, en un recuadro, de que un mismo punto tiene infinitos nombres. Ese aviso es correcto pero no es una definición, y el resto del capítulo va a usar constantemente cosas que todavía no se han dicho: que r puede ser negativo y qué significa entonces, cómo se calcula el ángulo sin

equivocarse de cuadrante, qué es exactamente una curva polar, qué es una simetría de una curva polar, y cómo se derivan y se miden las curvas dadas por un parámetro. Esta sección cierra ese hueco. Todo lo que viene después se apoya en ella.

10.2.1. El nombre de un punto no es único

Definición 10.3 (Coordenadas polares, con radio negativo permitido)**.** Fijados el polo O y el eje polar, a cada par de números reales (r, θ) se le asigna el punto del plano de coordenadas cartesianas

$$x = r\cos\theta, \qquad y = r\sin\theta.$$

Si $r > 0$, el punto está a distancia r del polo en la dirección θ. Si $r < 0$, el punto está a distancia $|r|$ en la dirección *opuesta*, es decir en $\theta + \pi$. Si $r = 0$, el punto es el polo, sea cual sea θ.

Se llaman *coordenadas polares de un punto* a cualquier par (r, θ) que produzca ese punto por la fórmula anterior. Un punto no tiene unas coordenadas polares: tiene infinitas.

Proposición 10.4 (Todos los nombres de un punto)**.** *Sea $P \neq O$ un punto con un par de coordenadas polares (r_0, θ_0), con $r_0 > 0$. Entonces los pares (r, θ) que representan a P son exactamente*

$$(r_0,\ \theta_0 + 2k\pi) \qquad y \qquad (-r_0,\ \theta_0 + \pi + 2k\pi), \qquad k \in \mathbb{Z}.$$

Para el polo, los pares que lo representan son exactamente los de la forma $(0, \theta)$ con θ cualquiera.

Demostración. Que todos esos pares dan P se comprueba sustituyendo: sumar $2k\pi$ al ángulo no cambia ni el seno ni el coseno, y cambiar el signo de r a la vez que se suma π al ángulo cambia dos signos que se cancelan, porque $\cos(\theta + \pi) = -\cos\theta$ y $\sin(\theta + \pi) = -\sin\theta$.

Recíprocamente, si (r, θ) representa a P, entonces $r^2 = r^2\cos^2\theta + r^2\sin^2\theta = x^2 + y^2 = r_0^2$, luego $r = \pm r_0$. Si $r = r_0$, dividiendo por $r_0 > 0$ queda $\cos\theta = \cos\theta_0$ y $\sin\theta = \sin\theta_0$, lo que obliga a $\theta = \theta_0 + 2k\pi$. Si $r = -r_0$, el mismo razonamiento da $\cos\theta = -\cos\theta_0$ y $\sin\theta = -\sin\theta_0$, es decir $\theta = \theta_0 + \pi + 2k\pi$. El caso del polo es inmediato, porque $r = 0$ anula las dos ecuaciones sea cual sea θ. □

Regla: convertir entre cartesianas y polares con el cuadrante correcto

De polares a cartesianas no hay dificultad ninguna: se sustituye en $x = r\cos\theta$, $y = r\sin\theta$ y se acabó. El problema está en el otro sentido, y su origen es que $\tan\theta = y/x$ no distingue un punto de su simétrico respecto del origen: los pares $(3, 4)$ y $(-3, -4)$ tienen el mismo cociente y/x y ángulos que difieren en π.

1. **Calcula el radio, que nunca da problemas:** $r = \sqrt{x^2 + y^2} \geq 0$.
2. **Trata los casos con $x = 0$ aparte, sin dividir.** Si $x = 0$ e $y > 0$, entonces $\theta = \pi/2$; si $x = 0$ e $y < 0$, entonces $\theta = -\pi/2$; si $x = y = 0$, el punto es el polo y θ queda libre.
3. **Con $x \neq 0$, usa la función** atan2, que es exactamente la arcotangente ya corregida

por cuadrantes:

$$\operatorname{atan2}(y,x)=\begin{cases}\arctan\dfrac{y}{x} & \text{si } x>0,\\[2mm] \arctan\dfrac{y}{x}+\pi & \text{si } x<0 \text{ e } y\geq 0,\\[2mm] \arctan\dfrac{y}{x}-\pi & \text{si } x<0 \text{ e } y<0.\end{cases}$$

Devuelve siempre un ángulo en el intervalo $(-\pi,\pi]$, y está en todas las calculadoras y lenguajes de programación con ese mismo nombre.

4. **Si prefieres el ángulo en $[0,2\pi)$, súmale 2π cuando salga negativo.** Es la única corrección que hace falta, y es puramente cosmética.

5. **Comprueba el signo de x y de y por separado.** El cuadrante lo deciden esos dos signos, no su cociente: $(+,+)$ es el primero, $(-,+)$ el segundo, $(-,-)$ el tercero y $(+,-)$ el cuarto. Si el ángulo obtenido no cae en el cuadrante que marcan los signos, hay un error.

6. **Verifica volviendo.** Sustituye el (r,θ) obtenido en $x=r\cos\theta$, $y=r\sin\theta$ y comprueba que salen los datos de partida. Cuesta diez segundos y detecta cualquier despiste de signo.

Ejemplo 10.5 (Cinco conversiones, con el cuadrante vigilado)**. (a) Segundo cuadrante:** $(x,y)=(-3,4)$**.** Radio: $r=\sqrt{9+16}=5$. Como $x<0$ e $y\geq 0$, el paso 3 da $\theta=\arctan(\frac{4}{-3})+\pi=-0{,}9273+\pi=2{,}2143$ radianes, es decir 126,8699°: segundo cuadrante, como marcan los signos. Verificación: $5\cos(2{,}2143)=-3$ y $5\sin(2{,}2143)=4$. Quien se hubiera fiado de la calculadora sin corregir habría obtenido $-0{,}9273$, que apunta al cuarto cuadrante.

(b) Tercer cuadrante: $(x,y)=(-1,-\sqrt{3})$**.** Radio $r=\sqrt{1+3}=2$. Con $x<0$ e $y<0$, el paso 3 resta π: $\theta=\arctan(\sqrt{3})-\pi=1{,}0472-\pi=-2{,}0944$ radianes, o sea $-120°$. En $[0,2\pi)$ es $\theta=4\pi/3=240°$. Comprobación: $2\cos 240°=-1$ y $2\sin 240°=-\sqrt{3}$.

(c) Sobre el eje: $(x,y)=(0,-5)$**.** Aquí no se divide: el paso 2 da directamente $r=5$ y $\theta=-\pi/2$, es decir $3\pi/2$ si se quiere positivo.

(d) Cuarto cuadrante: $(x,y)=(1,-1)$**.** Radio $\sqrt{2}\approx 1{,}4142$ y, como $x>0$, $\theta=\arctan(-1)=-\pi/4=-45°$, que en $[0,2\pi)$ es $7\pi/4$. Aquí la calculadora acierta sola, precisamente porque $x>0$.

(e) El mismo punto con radio negativo. El punto de (a) también admite el par $(-5,\ 2{,}2143-\pi)=(-5,\ -0{,}9273)$, por la proposición 10.4. Los dos pares son igual de legítimos; el primero es el que da la fórmula, el segundo aparecerá solo cuando una ecuación polar produzca radios negativos.

10.2.2. Qué es una curva polar y qué es una simetría

Definición 10.6 (Curva polar)**.** Dada una función f definida en un conjunto de ángulos, la *curva polar* de ecuación $r=f(\theta)$ es el conjunto de todos los puntos del plano cuyas

coordenadas cartesianas son

$$\big(f(\theta)\cos\theta,\ f(\theta)\sin\theta\big)$$

para algún θ del dominio. Si $f(\theta)<0$, el punto correspondiente se dibuja en la dirección $\theta+\pi$, según la definición 10.3.

Cuando la ecuación viene dada en la forma $r^2=g(\theta)$, la curva solo tiene puntos en los ángulos con $g(\theta)\geq 0$, y en cada uno de ellos hay *dos* valores de r, opuestos.

Cuidado

Un punto puede pertenecer a la curva $r=f(\theta)$ sin que sus coordenadas "naturales" cumplan la ecuación, porque el punto tiene otros nombres y basta con que uno de ellos la cumpla. Esta es la razón de fondo de todas las rarezas del capítulo: de que el origen se escape al buscar intersecciones, de que las rosas tengan a veces k pétalos y a veces $2k$, y de que haya que vigilar el intervalo de integración al calcular áreas.

Proposición 10.7 (Tres criterios de simetría)**.** *Sea la curva $r=f(\theta)$.*

(i) Si $f(-\theta)=f(\theta)$ para todo θ, la curva es simétrica respecto del ***eje polar*** *(el eje X).*

(ii) Si $f(\pi-\theta)=f(\theta)$ para todo θ, la curva es simétrica respecto de la ***recta*** *$\theta=\pi/2$ (el eje Y).*

(iii) Si $f(\theta+\pi)=f(\theta)$ para todo θ, la curva es simétrica respecto del ***polo****.*

En los tres casos la condición es suficiente, pero no necesaria: una curva puede tener la simetría sin que su ecuación la delate, porque los puntos pueden alcanzarse con otros nombres.

Demostración. (i) El punto de ángulo $-\theta$ y radio $f(-\theta)=f(\theta)$ tiene coordenadas cartesianas $\big(f(\theta)\cos(-\theta),\ f(\theta)\sin(-\theta)\big)=\big(f(\theta)\cos\theta,\ -f(\theta)\sin\theta\big)$, que es el reflejado respecto del eje X del punto de ángulo θ. Como esto vale para todo θ, la curva contiene el reflejo de cada uno de sus puntos.

(ii) Análogamente, con $\cos(\pi-\theta)=-\cos\theta$ y $\sin(\pi-\theta)=\sin\theta$, el punto de ángulo $\pi-\theta$ es el reflejado respecto del eje Y.

(iii) Con $\cos(\theta+\pi)=-\cos\theta$ y $\sin(\theta+\pi)=-\sin\theta$, el punto de ángulo $\theta+\pi$ y el mismo radio es el simétrico respecto del origen.

Que las condiciones no sean necesarias se ve con $r=2\sin(3\theta)$: no cumple (i), y sin embargo su dibujo sí es simétrico respecto de la recta $\theta=\pi/2$ por el criterio (ii), como se comprueba abajo. □

Regla: dibujar a mano una curva polar apoyándose en las simetrías

El objetivo es trazar la curva calculando la cuarta parte de los puntos. El orden de trabajo es siempre el mismo.

1. **Aplica los tres criterios de la proposición 10.7** antes de calcular ningún punto. Cada simetría que se encuentre reduce a la mitad el trabajo restante.

2. **Determina el dominio.** Si la ecuación es $r^2 = g(\theta)$, quédate solo con los ángulos donde $g(\theta) \geq 0$; en los demás no hay curva, y dibujar algo ahí es el error más grave posible.
3. **Resuelve $f(\theta) = 0$.** Esos ángulos son los pasos por el polo y, además, las direcciones de las tangentes en el polo. Son el esqueleto del dibujo.
4. **Localiza los máximos de $|f|$.** Dan los puntos más alejados, que son los vértices o las puntas de la curva.
5. **Tabula solo el intervalo mínimo.** Toma ángulos cómodos, de 15° en 15° o de 30° en 30°, y para cada uno anota r y el punto cartesiano $(r\cos\theta, r\sin\theta)$.
6. **Cuando salga $r < 0$, dibuja en la dirección opuesta.** Ni se ignora el punto ni se toma el valor absoluto: se marca a distancia $|r|$ en el ángulo $\theta + \pi$.
7. **Refleja usando las simetrías del paso 1** y une los puntos siguiendo el orden creciente de θ, que es el orden en que se recorre la curva.
8. **Comprueba un punto suelto** pasándolo a cartesianas y verificando que cae donde lo has dibujado.

Ejemplo 10.8 (Dos curvas trazadas punto a punto)**. (a) La rosa $r = 2\sin(3\theta)$.**

Paso 1. Con $\sin\big(3(\pi-\theta)\big) = \sin(3\pi - 3\theta) = \sin(3\theta)$, la curva cumple el criterio (ii): es simétrica respecto del eje Y. El criterio (i) falla, porque $\sin(-3\theta) = -\sin(3\theta)$.

Paso 3. $\sin(3\theta) = 0$ en $\theta = 0,\ \pi/3,\ 2\pi/3,\ \pi$: cuatro pasos por el polo en el intervalo $[0, \pi]$, que delimitan tres arcos.

Paso 4. $|r| = 2$ en $\theta = \pi/6$, $\theta = \pi/2$ y $\theta = 5\pi/6$.

Paso 5, la tabla completa de 0 *a* π*, de* 15 *en* 15 *grados:*

θ	0°	15°	30°	45°	60°	75°	90°	105°	120°	135°	150°	165°	180°
r	0	1,414	2	1,414	0	−1,414	−2	−1,414	0	1,414	2	1,414	0

Paso 6. Entre 60° y 120° el radio es negativo, así que ese arco *no* se dibuja arriba a la izquierda sino abajo: el máximo $r = -2$ en $\theta = 90°$ se marca en la dirección 270°, es decir en el punto cartesiano $(0, -2)$.

Paso 8, verificación. Los tres vértices son

$$\theta = 30°:\ (2\cos 30°, 2\sin 30°) = (\sqrt{3},\ 1) \approx (1{,}7321;\ 1),$$

$$\theta = 90°:\ (0, -2), \qquad \theta = 150°:\ (-\sqrt{3},\ 1) \approx (-1{,}7321;\ 1).$$

Están separados exactamente 120° entre sí, como corresponde a tres pétalos iguales, y el dibujo es simétrico respecto del eje Y, como anunciaba el paso 1. Con θ entre π y 2π no aparece nada nuevo: se repiten los mismos tres pétalos. *Control numérico:* el área de un pétalo es $\frac{1}{2}\int_0^{\pi/3} 4\sin^2(3\theta)\,\mathrm{d}\theta = \pi/3 \approx 1{,}0472$.

(b) La lemniscata $r^2 = 4\cos(2\theta)$.

Paso 1. $\cos(2(-\theta)) = \cos(2\theta)$, luego hay simetría respecto del eje polar; y $\cos(2(\theta+\pi)) = \cos(2\theta)$, luego también respecto del polo. Con las dos, basta calcular el primer octante.

Paso 2, el dominio. Hace falta $\cos(2\theta) \geq 0$, es decir $\theta \in [-\pi/4, \pi/4]$ y $\theta \in [3\pi/4, 5\pi/4]$. En el resto del plano **no hay curva**: entre 45° y 135° la figura está vacía, y ahí está la mitad de la gracia del dibujo.

Pasos 3 y 4. $r = 0$ en $\theta = \pm\pi/4$, donde la curva vuelve al origen; el máximo es $r = 2$ en $\theta = 0$.

Paso 5, la tabla:

θ	0°	15°	30°	45°
$r^2 = 4\cos 2\theta$	4	3,4641	2	0
r	2	1,8612	1,4142	0
punto (x, y)	(2; 0)	(1,7978; 0,4817)	(1,2247; 0,7071)	(0; 0)

Paso 7. Reflejando respecto del eje X se completa el lazo derecho, y reflejando el lazo derecho respecto del polo aparece el izquierdo, que es el que ocupa $\theta \in [3\pi/4, 5\pi/4]$. Sale el ocho tumbado. Las dos ramas llegan al origen con tangentes $\theta = \pi/4$ y $\theta = -\pi/4$, que forman entre sí un ángulo recto. *Control numérico:* el área de un lazo es $\frac{1}{2}\int_{-\pi/4}^{\pi/4} 4\cos(2\theta)\,\mathrm{d}\theta = 2$, y la de los dos, 4.

Definición 10.9 (Las cinco curvas polares con nombre propio)**.** Con $a > 0$ y k entero positivo se llaman:

- *caracol* a la curva $r = b + a\cos\theta$, y *cardioide* al caso $b = a$, es decir $r = a(1 + \cos\theta)$, que pasa por el polo una sola vez y allí forma una cúspide;
- *rosa* a la curva $r = a\cos(k\theta)$ o $r = a\sin(k\theta)$, formada por arcos que salen del polo y vuelven a él;
- *lemniscata* a la curva $r^2 = a^2\cos(2\theta)$ o $r^2 = a^2\sin(2\theta)$, que solo existe en la mitad de los ángulos y tiene forma de ocho tumbado;
- *espiral de Arquímedes* a la curva $r = a\theta$ con $\theta \geq 0$, cuyas vueltas distan siempre $2\pi a$;
- *espiral logarítmica* a la curva $r = \mathrm{e}^{a\theta}$, cuyas vueltas se separan en progresión geométrica.

10.2.3. Curvas paramétricas: la pendiente y la concavidad

La sección titulada "Ecuaciones paramétricas: cuando hace falta el tiempo" define formalmente qué es una curva paramétrica $x = x(t)$, $y = y(t)$ y qué es el parámetro; no hace falta repetirlo. Lo que allí se usa sin demostrar, y se demuestra aquí, son las dos fórmulas de derivación y el vocabulario de las tangentes.

Teorema 10.10 (Pendiente de una curva paramétrica)**.** *Sea una curva $x = x(t)$, $y = y(t)$ derivable, y sea t_0 con $x'(t_0) \neq 0$. Entonces cerca de t_0 la curva es la gráfica de una función de x, y su pendiente vale*

$$\frac{\mathrm{d}y}{\mathrm{d}x} = \frac{y'(t)}{x'(t)}.$$

Demostración. Como $x'(t_0) \neq 0$ y x' es continua, la función $x(t)$ es estrictamente monótona en un entorno de t_0, luego admite inversa derivable $t = t(x)$ con $\frac{\mathrm{d}t}{\mathrm{d}x} = 1/x'(t)$. Componiendo, $y = y(t(x))$ y por la regla de la cadena

$$\frac{\mathrm{d}y}{\mathrm{d}x} = y'(t)\,\frac{\mathrm{d}t}{\mathrm{d}x} = \frac{y'(t)}{x'(t)}.$$

□

Definición 10.11 (Tangente horizontal, tangente vertical y punto singular)**.** En un punto de parámetro t_0 se dice que la curva tiene tangente *horizontal* si $y'(t_0) = 0$ y $x'(t_0) \neq 0$, y tangente *vertical* si $x'(t_0) = 0$ e $y'(t_0) \neq 0$. Si $x'(t_0) = y'(t_0) = 0$, el punto se llama *singular* y el teorema anterior no dice nada sobre él: hay que estudiar el límite de $y'(t)/x'(t)$ cuando $t \to t_0$.

Teorema 10.12 (Segunda derivada)**.** *Con las hipótesis del teorema 10.10 y suponiendo x e y dos veces derivables,*

$$\frac{\mathrm{d}^2y}{\mathrm{d}x^2} = \frac{\dfrac{\mathrm{d}}{\mathrm{d}t}\left(\dfrac{y'(t)}{x'(t)}\right)}{x'(t)} = \frac{x'(t)\,y''(t) - y'(t)\,x''(t)}{x'(t)^3}.$$

Demostración. Llamemos $m(t) = y'(t)/x'(t)$ a la pendiente. La segunda derivada es la derivada de m respecto de x, no respecto de t, así que hay que volver a aplicar el teorema anterior con m en lugar de y:

$$\frac{\mathrm{d}^2y}{\mathrm{d}x^2} = \frac{\mathrm{d}m}{\mathrm{d}x} = \frac{m'(t)}{x'(t)}.$$

Derivando el cociente, $m' = \dfrac{y''x' - y'x''}{(x')^2}$, y dividiendo entre x' sale la fórmula del enunciado.

□

Cuidado

De aquí sale el error más caro de todo el tema: la segunda derivada **no** es $y''(t)/x''(t)$. La primera derivada sí era un cociente, y por eso la tentación es enorme, pero la demostración anterior muestra dónde está la trampa: hay que dividir entre x', no entre x''. El ejemplo siguiente da los dos números y se ve la diferencia.

Ejemplo 10.13 (Una curva con nodo, estudiada entera)**.** Sea $x(t) = t^2 - 1$, $y(t) = t^3 - t$, con $t \in \mathbb{R}$.

(a) Derivadas y pendiente. $x' = 2t$ e $y' = 3t^2 - 1$, luego

$$\frac{\mathrm{d}y}{\mathrm{d}x} = \frac{3t^2 - 1}{2t} \qquad (t \neq 0).$$

(b) Tangentes horizontales. $y' = 0$ en $t = \pm\frac{1}{\sqrt{3}} = \pm 0{,}5774$, donde $x' \neq 0$. Los puntos son

$$x = \tfrac{1}{3} - 1 = -\tfrac{2}{3} \approx -0{,}6667, \qquad y = t(t^2 - 1) = \mp\tfrac{2\sqrt{3}}{9} \approx \mp 0{,}3849.$$

(c) Tangente vertical. $x' = 0$ solo en $t = 0$, donde $y' = -1 \neq 0$: por la definición 10.11 hay tangente vertical en el punto $(-1, 0)$, que es el extremo izquierdo de la curva. No es un punto singular.

(d) El nodo. Los valores $t = 1$ y $t = -1$ dan el *mismo* punto: $x = 0$, $y = 0$. Pero las pendientes son distintas,

$$m(1) = \frac{3-1}{2} = 1, \qquad m(-1) = \frac{3-1}{-2} = -1.$$

La curva pasa dos veces por el origen con dos tangentes distintas, perpendiculares entre sí. Cerca del origen la curva *no* es la gráfica de ninguna función de x, y sin embargo cada una de las dos ramas sí lo es: por eso el teorema 10.10 habla de un entorno del parámetro y no de un entorno del punto.

(e) La concavidad, bien calculada. Simplificando la pendiente como $m(t) = \frac{3t}{2} - \frac{1}{2t}$, se tiene $m'(t) = \frac{3}{2} + \frac{1}{2t^2}$, y dividiendo entre $x' = 2t$,

$$\frac{\mathrm{d}^2 y}{\mathrm{d}x^2} = \frac{3t^2 + 1}{4t^3}.$$

En $t = 1$ vale $\frac{4}{4} = 1 > 0$: la rama es cóncava hacia arriba. En $t = -1$ vale $-1 < 0$: la otra rama es cóncava hacia abajo, y por eso las dos ramas se cruzan en aspa. *El error del recuadro anterior:* $y''/x'' = 6t/2 = 3t$, que en $t = 1$ daría 3 en lugar de 1. Ni el valor ni, en general, el signo tienen nada que ver.

10.2.4. La cicloide, deducida en lugar de anunciada

Definición 10.14 (Cicloide)**.** Se llama *cicloide* a la curva que describe un punto fijo del borde de una circunferencia de radio R que rueda sin deslizar sobre una recta.

Proposición 10.15 (Su parametrización)**.** *Si la recta es el eje X, la circunferencia rueda hacia la derecha y el punto marcado parte del origen, la cicloide es*

$$x(t) = R\,(t - \sin t), \qquad y(t) = R\,(1 - \cos t),$$

donde t es el ángulo, en radianes, que la rueda ha girado.

Demostración. La condición de rodar *sin deslizar* dice que la longitud del arco de rueda que ha tocado el suelo coincide con la distancia recorrida: tras girar un ángulo t, ese arco mide Rt, luego el centro de la rueda está en $(Rt,\ R)$, siempre a altura R.

El punto marcado partió del origen, es decir, del punto más bajo de la rueda, que respecto del centro está en la posición $(0, -R)$. Al girar la rueda un ángulo t *en sentido horario* (que es como gira al avanzar hacia la derecha), ese vector se transforma en $(-R\sin t,\ -R\cos t)$: en efecto, para $t = 0$ da $(0, -R)$, para $t = \pi/2$ da $(-R, 0)$, es decir el punto ha pasado a la parte trasera de la rueda, y para $t = \pi$ da $(0, R)$, el punto más alto. Sumando la posición del centro y la del punto respecto del centro,

$$(x, y) = (Rt,\ R) + (-R\sin t,\ -R\cos t) = \big(R(t - \sin t),\ R(1 - \cos t)\big). \qquad \square$$

Observación 10.16. Dos lecturas inmediatas de la fórmula. El punto toca el suelo ($y = 0$) exactamente cuando $\cos t = 1$, es decir en $t = 0, 2\pi, 4\pi, \ldots$, y allí $x' = R(1 - \cos t) = 0$ y $y' = R\sin t = 0$: son puntos singulares en el sentido de la definición 10.11, y de hecho son las cúspides. El punto más alto se alcanza en $t = \pi$, con $y = 2R$, el diámetro de la rueda, como tiene que ser.

10.2.5. Medir una curva paramétrica: longitud y superficie

Las fórmulas de área, longitud y superficie se usan en varias secciones de este capítulo. Las de área en polares y en paramétricas se demuestran más adelante, la primera a partir de la fórmula de Green; aquí quedan enunciadas y demostradas las dos que nunca se justifican.

Proposición 10.17 (Longitud de arco)**.** *Si x e y son de clase C^1 en $[t_0, t_1]$, la longitud de la curva recorrida es*

$$L = \int_{t_0}^{t_1} \sqrt{x'(t)^2 + y'(t)^2}\ \mathrm{d}t.$$

Demostración. Se aproxima la curva por la poligonal que une los puntos correspondientes a una partición $t_0 < t_1 < \cdots < t_N$. El lado k mide $\sqrt{(\Delta x_k)^2 + (\Delta y_k)^2}$, y por el teorema del valor medio aplicado a cada componente existen ξ_k, η_k en el subintervalo con $\Delta x_k = x'(\xi_k)\Delta t_k$ y $\Delta y_k = y'(\eta_k)\Delta t_k$. La suma de las longitudes es entonces $\sum \sqrt{x'(\xi_k)^2 + y'(\eta_k)^2}\,\Delta t_k$, que es una suma de Riemann de la función continua $\sqrt{x'^2 + y'^2}$; al afinar la partición converge a la integral del enunciado. □

Definición 10.18 (Área de una superficie de revolución)**.** Al girar la curva $x = x(t)$, $y = y(t)$, con $t \in [t_0, t_1]$, alrededor del eje X, cada trocito de curva de longitud $\mathrm{d}s = \sqrt{x'^2 + y'^2}\,\mathrm{d}t$ describe una circunferencia de radio $|y(t)|$ y por tanto de longitud $2\pi|y(t)|$. El área de la superficie engendrada es

$$S = 2\pi \int_{t_0}^{t_1} |y(t)| \sqrt{x'(t)^2 + y'(t)^2}\ \mathrm{d}t,$$

y girando alrededor del eje Y se cambia $|y|$ por $|x|$. Los valores absolutos son imprescindibles: un radio nunca es negativo.

Ejemplo 10.19 (Tres superficies, dos de ellas de calibración)**. (a) La esfera.** La semicircunferencia $x = R\cos t$, $y = R\sin t$ con $t \in [0, \pi]$ tiene $\sqrt{x'^2 + y'^2} = R$, luego

$$S = 2\pi \int_0^{\pi} R\sin t \cdot R\,\mathrm{d}t = 2\pi R^2 \big[-\cos t\big]_0^{\pi} = 4\pi R^2,$$

la fórmula conocida. La primera calibración sale bien.

(b) El cono. El segmento $x = t$, $y = mt$ con $t \in [0, h]$ y $m > 0$ engendra un cono de radio de base $\rho = mh$ y generatriz $g = h\sqrt{1 + m^2}$. La fórmula da

$$S = 2\pi \int_0^{h} mt\sqrt{1 + m^2}\ \mathrm{d}t = \pi m h^2 \sqrt{1 + m^2} = \pi \rho g.$$

Con $m = 0{,}75$ y $h = 4$: $\rho = 3$, $g = 4 \cdot 1{,}25 = 5$ y $S = 15\pi \approx 47{,}1239$. Coincide con la fórmula $\pi\rho g$ de la geometría elemental: segunda calibración superada.

(c) Un caso nuevo. El arco de parábola $x = t$, $y = t^2$ con $t \in [0, 1]$, girado alrededor del eje Y, tiene $|x| = t$ y $\sqrt{x'^2 + y'^2} = \sqrt{1 + 4t^2}$, luego

$$S = 2\pi \int_0^1 t\sqrt{1 + 4t^2}\, \mathrm{d}t = 2\pi \left[\frac{(1 + 4t^2)^{3/2}}{12}\right]_0^1 = \frac{\pi\,(5\sqrt{5} - 1)}{6} \approx 5{,}3304.$$

La sustitución $u = 1 + 4t^2$, $\mathrm{d}u = 8t\,\mathrm{d}t$ es la que hace elemental la integral; el factor t que sobra del radio es justamente el que se necesita. Con la misma curva girada alrededor del eje X la integral ya no es elemental y aparece un arcsinh: la diferencia entre un ejercicio y un problema está muchas veces en el eje elegido.

Pausa para pensar

Conviene fijarse en lo que tienen en común las tres fórmulas de esta subsección y las de área que vienen después. Todas son integrales de algo por $\mathrm{d}s$ o por $\mathrm{d}\theta$: la longitud integra 1, la superficie integra el perímetro $2\pi|y|$, el área en polares integra $\frac{1}{2}r^2$. Cambiar de curva no cambia el método, solo el integrando; y el trabajo de verdad casi nunca está en integrar, sino en dos cosas más humildes: fijar bien el intervalo en el que la curva se recorre una sola vez, y no perder ningún valor absoluto por el camino.

10.3. Ecuaciones sencillas en polares

Lo que hace útiles a las polares es que ciertas curvas, complicadísimas en cartesianas, tienen ecuaciones de una línea.

En polares	Qué es	En cartesianas
$r = a$	circunferencia centrada en el origen	$x^2 + y^2 = a^2$
$\theta = \alpha$	recta que pasa por el origen	$y = (\tan\alpha)\,x$
$r = 2a\cos\theta$	circunferencia de centro $(a, 0)$ y radio $\lvert a\rvert$	$(x - a)^2 + y^2 = a^2$
$r = 2a\sin\theta$	circunferencia de centro $(0, a)$ y radio $\lvert a\rvert$	$x^2 + (y - a)^2 = a^2$
$r\cos\theta = d$	recta vertical	$x = d$
$r = \dfrac{d}{\cos(\theta - \alpha)}$	recta a distancia d del origen	$x\cos\alpha + y\sin\alpha = d$

Teorema 10.20. *La ecuación polar $r = 2a\cos\theta$ describe la circunferencia de centro $(a, 0)$ y radio $|a|$, que pasa por el origen.*

Demostración. Multiplicando los dos miembros por r se obtiene $r^2 = 2ar\cos\theta$, es decir $x^2 + y^2 = 2ax$. Completando el cuadrado, $(x - a)^2 + y^2 = a^2$. El paso de multiplicar por r solo puede añadir la solución $r = 0$, que es el origen, y el origen ya está en la circunferencia ($\theta = \pi/2$ lo produce). □

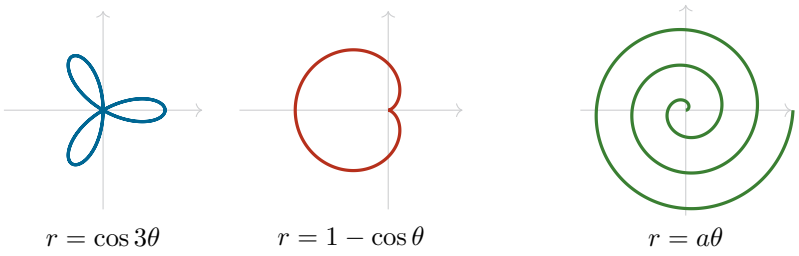


Figura 10.1: Tres curvas que en cartesianas serían monstruosas y en polares caben en una línea: la rosa de tres pétalos, la cardioide y la espiral de Arquímedes.

Ejemplo 10.21 (Circunferencias en polares)**. (a) La comprobación.** Multiplicando por r: $r^2 = 2ar\cos\theta$, es decir $x^2 + y^2 = 2ax$, que completando cuadrados da $(x-a)^2 + y^2 = a^2$. Centro $(a, 0)$ y radio $|a|$, como anuncia el enunciado.

(b) Las cuatro orientaciones. $r = 2a\cos\theta$ da una circunferencia tangente al eje y en el origen, desplazada hacia la derecha si $a > 0$ y hacia la izquierda si $a < 0$; $r = 2a\sin\theta$ la da tangente al eje x, desplazada hacia arriba o hacia abajo. Con $a = 3$: centro $(3, 0)$, radio 3, y pasa por el origen y por $(6, 0)$.

(c) El recorrido. Basta con $\theta \in [-\pi/2, \pi/2]$ para trazar la circunferencia entera; dejando correr θ hasta 2π se recorre dos veces, porque para $\theta \in (\pi/2, 3\pi/2)$ el radio sale negativo y el punto se refleja al otro lado. Es la trampa clásica al calcular áreas con estas curvas.

(d) El área. $A = \frac{1}{2}\int_{-\pi/2}^{\pi/2}(2a\cos\theta)^2 \mathrm{d}\theta = 2a^2\int_{-\pi/2}^{\pi/2}\cos^2\theta\,\mathrm{d}\theta = \pi a^2$. Coincide con la fórmula del círculo de radio a, y si se integrase de 0 a 2π saldría el doble: precisamente el error que avisa el apartado anterior.

Observación 10.22. Compare las dos descripciones de esa misma circunferencia: $r = 2a\cos\theta$ frente a $x^2 + y^2 = 2ax$. La primera es explícita —se le da un ángulo y devuelve un radio—, la segunda es implícita. Ese es el patrón general: **las polares convierten en explícito todo lo que tiene al origen como punto privilegiado**. Por eso son las coordenadas naturales de la gravitación, del sonar, del radar y de cualquier fenómeno que se propague desde un centro.

Regla: cómo dibujar una curva en polares sin morir en el intento

Dibujar $r = f(\theta)$ a base de tabla de valores funciona, pero es lento y se cometen errores. Con simetrías y ceros se hace en la cuarta parte del tiempo.

1. **Busca las simetrías antes de calcular nada.** Si al cambiar θ por $-\theta$ la ecuación no cambia, la curva es simétrica respecto del eje polar (el eje X). Si no cambia al sustituir θ por $\pi - \theta$, es simétrica respecto del eje Y. Si no cambia al sustituir r por $-r$ (o θ por $\theta + \pi$), es simétrica respecto del origen. Con una sola simetría te ahorras la mitad de la tabla.

2. **Localiza dónde $r = 0$.** Resolviendo $f(\theta) = 0$ sabes en qué ángulos la curva pasa por el origen, y esos son los ángulos de las rectas tangentes ahí. En las rosas, son exactamente los ángulos que separan un pétalo del siguiente.

3. **Localiza los máximos y mínimos de r.** Los puntos más alejados del origen están donde f alcanza su máximo. Marcarlos primero te da el «esqueleto» de la curva.
4. **Haz la tabla solo en el intervalo que hace falta.** Toma ángulos cómodos (0, $\pi/6$, $\pi/4$, $\pi/3$, $\pi/2$, …) y aprovecha las simetrías del paso 1 para reflejar el resto.
5. **Interpreta bien los valores negativos de r.** El punto (r,θ) con $r<0$ se dibuja a distancia $|r|$ en la dirección *opuesta*, es decir en $\theta+\pi$. No se ignora ni se toma el valor absoluto: se dibuja al otro lado. Es lo que genera los bucles interiores.
6. **Comprueba el dominio.** Si la ecuación es del tipo $r^2=f(\theta)$, solo existen los ángulos con $f(\theta)\geq 0$; en el resto no hay curva. Y si f toma valores negativos con r despejada, mira el paso 5.
7. **Verifica pasando a cartesianas un punto.** Elige un θ concreto, calcula r, pasa a $(x,y)=(r\cos\theta,\ r\sin\theta)$ y comprueba que ese punto cae donde lo has dibujado.

Ejemplo 10.23 (Seis curvas, de la circunferencia disfrazada al bucle interior)**.** (a) **Una circunferencia disfrazada: $r=2\cos\theta$.** Simetría respecto del eje polar (paso 1, porque $\cos(-\theta)=\cos\theta$). Ceros: $r=0$ en $\theta=\pi/2$. Máximo: $r=2$ en $\theta=0$. El truco definitivo es multiplicar por r:

$$r^2=2r\cos\theta \implies x^2+y^2=2x \implies (x-1)^2+y^2=1.$$

Es la circunferencia de centro $(1,0)$ y radio 1. Aviso importante: se recorre *entera* con $\theta\in[0,\pi]$; si dejas θ llegar a 2π la recorres dos veces, y eso estropeará el cálculo de áreas de la sección siguiente.

(b) **La cardioide $r=1+\cos\theta$.** Simétrica respecto del eje polar. Cero único en $\theta=\pi$, máximo $r=2$ en $\theta=0$. Tabla con la mitad de los ángulos:

θ	0	$\pi/3$	$\pi/2$	$2\pi/3$	$5\pi/6$	π
r	2	$1{,}5$	1	$0{,}5$	$\approx 0{,}134$	0

Se dibuja de 0 a π y se refleja. El pico en el origen no es un fallo del dibujo: es una *cúspide* de verdad.

(c) **La rosa $r=\cos(2\theta)$.** Ceros en $\theta=\pi/4,\ 3\pi/4,\ 5\pi/4,\ 7\pi/4$; máximos $|r|=1$ en $\theta=0,\ \pi/2,\ \pi,\ 3\pi/2$. Entre $\theta=\pi/4$ y $\theta=3\pi/4$ el coseno es negativo, así que ese pétalo se dibuja *al otro lado* por el paso 5. Resultado: cuatro pétalos, con los ejes de los pétalos a 0°, 90°, 180° y 270°.

(d) **La espiral de Arquímedes $r=\theta$, con $\theta\geq 0$.** No tiene ninguna simetría, $r=0$ solo en el origen y r crece sin parar. Cada vuelta completa aleja la curva exactamente 2π del trazo anterior: por eso las espirales de este tipo tienen los surcos igual de separados, como un disco de vinilo.

(e) **Un bucle interior: $r = 1 - 2\cos\theta$.** Aquí $r < 0$ cuando $\cos\theta > \frac{1}{2}$, es decir para $-\pi/3 < \theta < \pi/3$. En $\theta = 0$ sale $r = -1$, que por el paso 5 se dibuja a distancia 1 en la dirección π: el punto cartesiano $(-1, 0)$. Ese tramo con r negativa es el que forma el lazo pequeño dentro de la curva grande. Los ceros están en $\theta = \pm\pi/3$, y son donde el lazo se cierra.

(f) **Una curva con agujeros: la lemniscata $r^2 = \cos(2\theta)$.** Por el paso 6 solo hay curva donde $\cos 2\theta \geq 0$, o sea en $\theta \in [-\pi/4, \pi/4]$ y en $\theta \in [3\pi/4, 5\pi/4]$. En el resto del plano, *nada*. Salen dos lazos simétricos con forma de ocho tumbado, y en $\theta = \pm\pi/4$ la curva llega al origen con tangentes que forman 90° entre sí.

Cuidado

El error más caro es el de (a): dejar que θ recorra $[0, 2\pi]$ cuando la curva ya se ha cerrado en $[0, \pi]$. Con el dibujo no se nota —la curva sale igual— pero al calcular el área o la longitud sale exactamente el doble. Antes de integrar, comprueba siempre en qué intervalo se recorre la curva *una sola vez*.

Pausa para pensar

El mismo punto del plano tiene infinitas coordenadas polares: $(1, \pi/4)$, $(1, \pi/4 + 2\pi)$, $(-1, 5\pi/4)$... son todas el mismo punto. Esto tiene una consecuencia incómoda: un punto puede estar en una curva $r = f(\theta)$ *sin* que sus coordenadas «naturales» cumplan la ecuación. Por eso, para hallar las intersecciones de dos curvas en polares, resolver el sistema no basta: hay que comprobar además si el origen pertenece a las dos, y probar las representaciones alternativas.

Idea clave

Las polares no son un capricho: son las coordenadas en las que la simetría del problema es evidente. Una circunferencia centrada en el origen es $r =$ constante, una recta por el origen es $\theta =$ constante, y una espiral logarítmica es $r = e^{a\theta}$. La regla general —y vale para todo lo que verás después— es que el sistema de coordenadas correcto es aquel en el que las simetrías del problema se convierten en «una variable constante». Esa idea, llevada a su conclusión, es la que separa un cálculo de tres líneas de uno de tres páginas, y reaparece en las integrales múltiples del capítulo 9 y en la separación de variables de las ecuaciones diferenciales del capítulo 12.

10.4. Curvas que solo existen cómodamente en polares

10.4.1. La cardioide

Ejemplo 10.24. La curva $r = 1 + \cos\theta$. Tabulamos:

θ	0	$\pi/4$	$\pi/2$	$3\pi/4$	π	$5\pi/4$	$3\pi/2$	$7\pi/4$	2π
r	2	1,707	1	0,293	0	0,293	1	1,707	2

El radio va decreciendo de 2 a 0 mientras el ángulo recorre media vuelta, y luego vuelve a crecer simétricamente. El resultado tiene forma de corazón, y de ahí el nombre. En cartesianas su ecuación es $(x^2+y^2-x)^2 = x^2+y^2$: una curva de grado cuatro con la que nadie querría trabajar.

10.4.2. Las rosas

Regla: el número de pétalos de $r = \cos(k\theta)$

Para k entero:

- si k es **impar**, la curva tiene exactamente k pétalos y se recorre entera con $\theta \in [0, \pi]$;
- si k es **par**, tiene $2k$ pétalos y hace falta $\theta \in [0, 2\pi]$.

La razón de esta rareza es la ambigüedad de las polares: cuando k es impar, el tramo $\theta \in [\pi, 2\pi]$ vuelve a dibujar los mismos pétalos que ya estaban, mientras que si k es par dibuja pétalos nuevos intercalados. Para $r = \cos 3\theta$, por ejemplo, r vale 1 en $\theta = 0$, 0 en $\theta = 30°$, -1 en $\theta = 60°$, y ya se ha completado un pétalo y medio.

10.4.3. La espiral de Arquímedes

$r = \theta$ (con $\theta \geq 0$) es la curva que se aleja del origen a velocidad constante mientras gira a velocidad constante. Sus vueltas están *equiespaciadas*: la distancia entre dos vueltas consecutivas es siempre 2π. Es exactamente la forma de un rollo de papel o de una cuerda enrollada, y por eso aparece en la ingeniería más que en los libros.

Pausa para pensar

Compárela con la **espiral logarítmica** $r = \mathrm{e}^{a\theta}$, donde la distancia entre vueltas crece geométricamente. Esta segunda es la que aparece en las conchas de los nautilos, en los brazos de las galaxias y en el vuelo de un halcón hacia su presa. La diferencia entre las dos es la diferencia entre crecer sumando y crecer multiplicando, que es la misma distinción entre progresión aritmética y geométrica del las sucesiones y series elementales. Jakob Bernoulli quedó tan impresionado con la propiedad de que la espiral logarítmica es semejante a sí misma tras un giro, que pidió que se grabara en su tumba con la inscripción *eadem mutata resurgo*: "transformada, resurjo la misma".

10.4.4. Las cónicas vistas desde el foco

Teorema 10.25 (Ecuación focal de las cónicas)**.** *Con el origen en un foco, toda cónica tiene ecuación polar*

$$r = \frac{\ell}{1 + e\cos\theta},$$

donde e es la excentricidad y ℓ el semilado recto. La curva es una elipse si $0 \leq e < 1$ (circunferencia si $e = 0$), una parábola si $e = 1$ y una hipérbola si $e > 1$.

Idea clave

Esta única fórmula es la razón de que las polares sean las coordenadas de la astronomía. En cartesianas, elipse, parábola e hipérbola tienen tres ecuaciones distintas y una tabla de casos; aquí son *la misma curva* con un parámetro que varía. Y el origen no está en el centro de la elipse, sino en el foco, que es exactamente donde está el Sol. La primera ley de Kepler, escrita en polares, es esta fórmula y nada más.

Ejemplo 10.26. La órbita de la Tierra tiene $e = 0{,}0167$ y semieje mayor $a = 149{,}6$ millones de km. Como el punto más cercano al foco es $a(1-e)$ y el más lejano $a(1+e)$, la Tierra oscila entre 147,1 y 152,1 millones de km del Sol: una variación de apenas el 3,3 %. La órbita es prácticamente una circunferencia, y las estaciones *no* se deben a esa diferencia sino a la inclinación del eje. En cambio el cometa Halley tiene $e = 0{,}967$ y $a = 17{,}8$ unidades astronómicas: se acerca hasta 0,59 UA (más cerca que Venus) y se aleja hasta 35 UA (más allá de Neptuno).

Regla: calcular áreas y longitudes de curvas dadas en polares

En cartesianas el elemento de área es un rectangulito; en polares es un *sector circular*, y de ahí sale el $\frac{1}{2}r^2$ que sorprende la primera vez.

1. **Recuerda de dónde sale la fórmula.** El sector de radio r y ángulo $\mathrm{d}\theta$ tiene área $\frac{1}{2}r^2\,\mathrm{d}\theta$ (es la fracción $\mathrm{d}\theta/2\pi$ del círculo completo πr^2). Sumando todos los sectores:
$$A = \frac{1}{2}\int_\alpha^\beta r^2\,\mathrm{d}\theta.$$
El $\frac{1}{2}$ no es un factor de conversión: es la mitad del sector.

2. **Elige bien los límites α y β.** Deben recorrer la curva *exactamente una vez*. Los ceros de r suelen ser los límites naturales de un pétalo o de un lazo.

3. **Para el área *entre* dos curvas, resta los cuadrados.** Si la exterior es r_2 y la interior r_1 en el mismo rango de ángulos: $A = \frac{1}{2}\int (r_2^2 - r_1^2)\,\mathrm{d}\theta$. Nunca restes las áreas de rangos angulares distintos sin dibujarlo antes.

4. **Halla los límites resolviendo $r_1(\theta) = r_2(\theta)$.** Y comprueba aparte si el origen pertenece a las dos curvas, porque las intersecciones en el origen no siempre aparecen al resolver la ecuación.

5. **Para la longitud, usa** $L = \displaystyle\int_\alpha^\beta \sqrt{r^2 + \left(\frac{\mathrm{d}r}{\mathrm{d}\theta}\right)^2}\,\mathrm{d}\theta$. Sale del teorema de Pitágoras aplicado al desplazamiento radial $\mathrm{d}r$ y al transversal $r\,\mathrm{d}\theta$, que son perpendiculares.

6. **Simplifica el radicando antes de integrar.** Casi siempre aparece una identidad trigonométrica que lo convierte en un cuadrado perfecto: $2 + 2\cos\theta = 4\cos^2(\theta/2)$ y $2 - 2\cos\theta = 4\sin^2(\theta/2)$ resuelven la mayoría de los ejercicios clásicos.

7. **Verifica con un caso conocido.** Aplica tu fórmula a $r = R$ constante: debe dar $A = \pi R^2$ y $L = 2\pi R$. Si no, hay un error de límites o un factor 2 perdido.

Ejemplo 10.27 (Cinco áreas y dos longitudes, con todas las cuentas)**.** (a) **La verificación del paso 7.** Para $r = R$ constante: $A = \frac{1}{2}\int_0^{2\pi} R^2 \mathrm{d}\theta = \frac{1}{2}R^2 \cdot 2\pi = \pi R^2$ ✓, y como $\mathrm{d}r/\mathrm{d}\theta = 0$, $L = \int_0^{2\pi} \sqrt{R^2}\,\mathrm{d}\theta = 2\pi R$ ✓.

(b) **La circunferencia disfrazada.** Con $r = 2\cos\theta$, que recorre toda la circunferencia una sola vez para $\theta \in [-\pi/2, \pi/2]$:

$$A = \frac{1}{2}\int_{-\pi/2}^{\pi/2} 4\cos^2\theta\,\mathrm{d}\theta = 2\int_{-\pi/2}^{\pi/2} \frac{1+\cos 2\theta}{2}\,\mathrm{d}\theta = \Big[\theta + \frac{\sin 2\theta}{2}\Big]_{-\pi/2}^{\pi/2} = \pi.$$

Y en efecto es la circunferencia de radio 1, cuya área es π. Si por descuido hubieras integrado de 0 a 2π habrías obtenido 2π: el doble, exactamente el error que anunciaba la advertencia anterior.

(c) **El área de la cardioide.** Con $r = 1 + \cos\theta$ y $\theta \in [0, 2\pi]$:

$$A = \frac{1}{2}\int_0^{2\pi} (1+\cos\theta)^2 \mathrm{d}\theta = \frac{1}{2}\int_0^{2\pi} (1+2\cos\theta+\cos^2\theta)\mathrm{d}\theta = \frac{1}{2}(2\pi+0+\pi) = \frac{3\pi}{2} \approx 4{,}7124.$$

Los tres sumandos salen de: $\int 1 = 2\pi$, $\int \cos\theta = 0$ sobre un periodo completo, y $\int \cos^2\theta = \pi$ (la mitad del periodo, porque el valor medio de $\cos^2$ es $\frac{1}{2}$).

(d) **Un pétalo de la rosa.** El pétalo de $r = \cos 2\theta$ centrado en $\theta = 0$ va de cero a cero, es decir de $-\pi/4$ a $\pi/4$:

$$A = \frac{1}{2}\int_{-\pi/4}^{\pi/4} \cos^2(2\theta)\,\mathrm{d}\theta = \frac{1}{4}\int_{-\pi/4}^{\pi/4} (1+\cos 4\theta)\mathrm{d}\theta = \frac{1}{4}\Big[\theta + \frac{\sin 4\theta}{4}\Big]_{-\pi/4}^{\pi/4} = \frac{\pi}{8}.$$

Los cuatro pétalos suman $\pi/2$, es decir la mitad del círculo de radio 1 que los contiene. Un dato que sirve para comprobar el resultado de un vistazo.

(e) **Área entre dos curvas.** ¿Qué área hay dentro de $r = 3\cos\theta$ y fuera de $r = 1 + \cos\theta$? Paso 4: $3\cos\theta = 1 + \cos\theta$ da $\cos\theta = \frac{1}{2}$, o sea $\theta = \pm\pi/3$. Paso 3:

$$A = \frac{1}{2}\int_{-\pi/3}^{\pi/3} \Big(9\cos^2\theta - (1+\cos\theta)^2\Big)\mathrm{d}\theta.$$

Desarrollando el integrando: $9\cos^2\theta - 1 - 2\cos\theta - \cos^2\theta = 8\cos^2\theta - 2\cos\theta - 1 = 3 + 4\cos 2\theta - 2\cos\theta$. Integrando entre $-\pi/3$ y $\pi/3$, el término $4\cos 2\theta$ aporta $4\sqrt{3}/2 \cdot 2 = 2\sqrt{3}$ y el término $-2\cos\theta$ aporta $-2\sqrt{3}$: se cancelan exactamente. Queda $3 \cdot \frac{2\pi}{3} = 2\pi$, y por tanto

$$A = \frac{2\pi}{2} = \pi \approx 3{,}1416.$$

(f) **La longitud de la cardioide.** Con $r = 1 + \cos\theta$ se tiene $\mathrm{d}r/\mathrm{d}\theta = -\sin\theta$, luego

$$r^2 + \Big(\frac{\mathrm{d}r}{\mathrm{d}\theta}\Big)^2 = 1 + 2\cos\theta + \cos^2\theta + \sin^2\theta = 2 + 2\cos\theta = 4\cos^2\frac{\theta}{2},$$

que es justo la identidad del paso 6. Por tanto

$$L = \int_0^{2\pi} 2\Big|\cos\frac{\theta}{2}\Big|\,\mathrm{d}\theta = 8.$$

Un número entero exacto, sin ningún π: una de esas coincidencias que hacen sospechar que la cardioide es más especial de lo que parece. (Las barras de valor absoluto son imprescindibles: $\cos(\theta/2)$ es negativo en la segunda mitad del recorrido, y sin ellas saldría 0.)

(g) **Una longitud que no es tan bonita.** Para la espiral $r = \theta$ con $\theta \in [0, 2\pi]$:

$$L = \int_0^{2\pi} \sqrt{\theta^2 + 1}\, \mathrm{d}\theta = \Big[\frac{\theta\sqrt{\theta^2+1}}{2} + \frac{1}{2}\ln\left(\theta + \sqrt{\theta^2+1}\right)\Big]_0^{2\pi} \approx 21{,}2563.$$

Compáralo con $2\pi \cdot \pi \approx 19{,}74$, la longitud de la circunferencia de radio medio π: la espiral es algo más larga, como cabía esperar por su componente radial.

Cuidado

En el apartado (e), restar «área de la primera menos área de la segunda» calculadas cada una en *todo* su rango daría $\frac{9\pi}{4} - \frac{3\pi}{2} = \frac{3\pi}{4}$, un número que no responde a la pregunta. Las áreas en polares solo se restan cuando el rango de ángulos es el mismo y una curva está por fuera de la otra en todo ese rango. Si no, hay que partir la integral en trozos.

Pausa para pensar

¿Por qué en cartesianas el área es $\int y\,\mathrm{d}x$ y en polares $\frac{1}{2}\int r^2\mathrm{d}\theta$, con ese cuadrado? Porque el elemento de área no es el mismo: al aumentar θ un poquito, la porción de plano barrida es un triángulo delgadísimo cuya base mide $r\,\mathrm{d}\theta$ y cuya altura es r, de donde $\frac{1}{2}r \cdot r\,\mathrm{d}\theta$. Cuanto más lejos estás del origen, más área barre el mismo ángulo. Es exactamente lo mismo que dice la segunda ley de Kepler cuando afirma que un planeta barre áreas iguales en tiempos iguales.

Idea clave

El factor r que aparece en el elemento de área es el *jacobiano* del cambio a polares, y por eso en las integrales dobles del capítulo 9 se escribe $\mathrm{d}A = r\,\mathrm{d}r\,\mathrm{d}\theta$. La fórmula $\frac{1}{2}\int r^2\mathrm{d}\theta$ de este capítulo no es otra cosa que esa integral doble con la variable r ya integrada de 0 a $f(\theta)$: $\int_0^{f(\theta)} r\,\mathrm{d}r = \frac{1}{2}f(\theta)^2$. Las dos fórmulas que parecen distintas son la misma, vista con un paso de más o de menos.

10.5. Dónde se cortan dos curvas polares

Ya se ha avisado dos veces en este capítulo: resolver el sistema de dos ecuaciones polares no basta para hallar todos los puntos comunes. Ahora toca convertir ese aviso en un procedimiento que no falle.

Regla: hallar todas las intersecciones de dos curvas en polares

El problema no es el álgebra, sino que un mismo punto tiene infinitos nombres. La receta consiste en buscar los puntos por *todos* sus nombres, no solo por el más cómodo.

1. **Resuelve primero el sistema tal cual.** Iguala $f(\theta) = g(\theta)$ y despeja los ángulos en $[0, 2\pi)$. Cada solución θ_0 da el punto $(f(\theta_0), \theta_0)$. Estos son los cortes «honrados», los que se ven a la primera.

2. **Comprueba el origen aparte, siempre.** El origen es el único punto cuyo ángulo es completamente libre. Resuelve $f(\theta) = 0$ y $g(\theta) = 0$ por separado: si *ambas* ecuaciones tienen alguna solución, el origen es un punto común, aunque los ángulos no coincidan y aunque el sistema del paso 1 no lo haya sacado.

3. **Prueba también la rama con radio negativo.** El punto (r, θ) es el mismo que $(-r, \theta + \pi)$. Por eso hay que resolver además la ecuación $f(\theta) = -g(\theta + \pi)$. En las curvas con pétalos o bucles, esta segunda ecuación suele dar tantos puntos como la primera.

4. **Y las representaciones desplazadas 2π, si los periodos no coinciden.** Si f tiene periodo 2π y g periodo π —como pasa con las rosas de índice par—, el mismo lugar del plano se alcanza con ángulos que difieren en 2π, y hay que recorrer un intervalo suficientemente largo.

5. **Traduce cada solución a cartesianas.** Con $(x, y) = (r\cos\theta,\, r\sin\theta)$ se detectan inmediatamente los duplicados: dos parejas (r, θ) distintas que den el mismo (x, y) son un solo punto y hay que contarlo una vez.

6. **Dibuja y cuenta.** Un esbozo, aunque sea tosco, dice cuántos cortes debe haber. Si el álgebra da menos, falta alguno de los pasos 2, 3 o 4; si da más, hay duplicados que el paso 5 debe eliminar.

7. **Si todo falla, pasa las dos curvas a cartesianas.** En cartesianas cada punto tiene un solo nombre y el problema de la ambigüedad desaparece por completo. El precio es un sistema algebraico más feo, pero es un método que no puede perder soluciones.

Ejemplo 10.28 (Seis intersecciones, y en la mitad falta algo)**.** (a) **La cardioide y la circunferencia grande.** Sean $r = 1 + \cos\theta$ y $r = 3\cos\theta$. Paso 1: $1 + \cos\theta = 3\cos\theta$ da $\cos\theta = \frac{1}{2}$, o sea $\theta = \pi/3$ y $\theta = 5\pi/3$, con $r = \frac{3}{2}$ en ambos. En cartesianas son

$$\left(\tfrac{3}{4},\ \tfrac{3\sqrt{3}}{4}\right) \approx (0{,}75;\ 1{,}2990) \quad \text{y} \quad \left(\tfrac{3}{4},\ -\tfrac{3\sqrt{3}}{4}\right).$$

Paso 2: la cardioide se anula en $\theta = \pi$ y la circunferencia en $\theta = \pi/2$. Ángulos distintos, pero *las dos pasan por el origen.* Hay por tanto **tres** intersecciones, no dos, y el sistema solo encontró dos de ellas.

(b) **Dos circunferencias sin sorpresas.** Sean $r = 2$ y $r = 4\cos\theta$. Paso 1: $\cos\theta = \frac{1}{2}$, luego $\theta = \pm\pi/3$ y $r = 2$; los puntos son $(1, \sqrt{3})$ y $(1, -\sqrt{3})$. Paso 2: $r = 2$ nunca se anula, así que el origen no está en la primera curva y no aporta nada. Paso 3: $2 = -4\cos(\theta + \pi) = 4\cos\theta$, la misma ecuación de antes. Total: exactamente **dos** puntos. Aquí el método ingenuo acertaba, y conviene saber por qué: porque ninguna de las dos curvas pasa por el origen y ninguna admite radios negativos.

(c) **La circunferencia y la rosa: el sistema encuentra la mitad.** Sean $r = 1$ y $r = 2\sin 2\theta$. Paso 1: $\sin 2\theta = \frac{1}{2}$ da

$$\theta = \frac{\pi}{12},\ \frac{5\pi}{12},\ \frac{13\pi}{12},\ \frac{17\pi}{12},$$

cuatro puntos. Paso 3: la otra rama es $1 = -2\sin\big(2(\theta+\pi)\big) = -2\sin 2\theta$, es decir $\sin 2\theta = -\frac{1}{2}$, que da

$$\theta = \frac{7\pi}{12},\ \frac{11\pi}{12},\ \frac{19\pi}{12},\ \frac{23\pi}{12},$$

otros cuatro. En total **ocho** intersecciones, y el dibujo lo confirma de un vistazo: la rosa tiene cuatro pétalos de radio máximo 2, y la circunferencia de radio 1 entra y sale de cada pétalo, cortándolo dos veces. Quien se quede con el paso 1 pierde la mitad de la respuesta.

(d) **Dos circunferencias que se tocan en el origen.** Sean $r = \cos\theta$ y $r = \sin\theta$. Paso 1: $\tan\theta = 1$, luego $\theta = \pi/4$ y $r = \frac{\sqrt{2}}{2}$, es decir el punto cartesiano $\big(\frac{1}{2}, \frac{1}{2}\big)$. Paso 2: la primera se anula en $\theta = \pi/2$ y la segunda en $\theta = 0$. Otra vez ángulos distintos y otra vez origen común. Son **dos** puntos, y tiene todo el sentido geométrico: dos circunferencias de radio $\frac{1}{2}$ centradas en $\big(\frac{1}{2}, 0\big)$ y $\big(0, \frac{1}{2}\big)$ se cortan en dos sitios.

(e) **Dos cardioides enfrentadas.** Sean $r = 1 - \cos\theta$ y $r = 1 + \cos\theta$. Paso 1: $\cos\theta = 0$, luego $\theta = \pi/2$ y $\theta = 3\pi/2$, con $r = 1$: los puntos $(0, 1)$ y $(0, -1)$. Paso 2: la primera se anula en $\theta = 0$ y la segunda en $\theta = \pi$, así que el origen también es común. **Tres** puntos. Obsérvese que las dos curvas son simétricas una de la otra respecto del eje Y, y que el origen es la cúspide de las dos.

(f) **Cuando conviene rendirse y pasar a cartesianas.** Para $r = 1 + \cos\theta$ y $r = 1 + \sin\theta$ el paso 1 da $\cos\theta = \sin\theta$, o sea $\theta = \pi/4$ y $\theta = 5\pi/4$; el primero da $r = 1 + \frac{\sqrt{2}}{2} \approx 1{,}7071$ y el segundo $r = 1 - \frac{\sqrt{2}}{2} \approx 0{,}2929$. Ambos son puntos válidos, sobre la bisectriz. Y por el paso 2 el origen vuelve a ser común, porque la primera se anula en $\theta = \pi$ y la segunda en $\theta = 3\pi/2$. Son **tres** puntos. Quien intente esto en cartesianas se encontrará con dos curvas de grado cuatro; aquí el paso 7 es el remedio de última instancia, no el primer recurso.

Cuidado

El origen se escapa del sistema de ecuaciones con una regularidad exasperante: aparece en cuatro de los seis apartados anteriores y en ninguno de ellos lo detecta el paso 1. La razón es puramente formal: el sistema $f(\theta) = g(\theta)$ busca un *ángulo común*, y el origen no tiene ángulo propio. Adquiere el hábito de escribir, debajo de cada sistema resuelto, la línea «¿pasan las dos por el origen?». Cuesta diez segundos y evita el error más frecuente del tema.

Pausa para pensar

¿Por qué en cartesianas no pasa nada de esto? Porque la aplicación que a cada punto le asigna su par (x, y) es biyectiva, mientras que la que le asigna su par (r, θ) no lo es: es infinito a uno, y encima de manera irregular —el origen tiene un comportamiento

distinto del resto—. Resolver un sistema es buscar coincidencias de *nombres*; si un objeto tiene varios nombres, comparar nombres deja de ser lo mismo que comparar objetos. Toda la dificultad de este apartado cabe en esa frase.

Idea clave

Esta patología tiene una versión respetable en matemáticas superiores. Las coordenadas polares son una *carta* que cubre el plano menos el origen, y el origen es exactamente el punto donde la carta degenera: allí el jacobiano r se anula. Lo mismo ocurre con la latitud y la longitud en los polos de la Tierra, donde la longitud deja de estar definida. La lección general —que se formaliza en el capítulo 9 y se lleva a su conclusión en el 36— es que ningún sistema de coordenadas cubre bien todo el espacio, y que los puntos donde falla hay que tratarlos a mano.

10.6. Taller: exprimir la ecuación focal de una cónica

El teorema anterior dice que toda cónica se escribe $r = \ell/(1 + e\cos\theta)$ con el foco en el origen. Eso está muy bien, pero en la práctica uno se encuentra la ecuación escrita de cualquier otra manera y necesita extraer de ella todo lo que hay dentro: qué cónica es, hacia dónde apunta, cuánto miden sus ejes y a qué distancia pasa del foco. Este apartado es ese taller.

Regla: identificar y medir una cónica dada en polares

1. **Normaliza: el denominador debe empezar por un 1.** Si la ecuación es $r = \dfrac{A}{B \pm C\cos\theta}$, divide numerador y denominador entre B:

$$r = \frac{A/B}{1 \pm (C/B)\cos\theta}.$$

   Solo entonces se puede leer nada. Leer la excentricidad sin normalizar es el error número uno del tema.

2. **Lee e y ℓ directamente.** Ya normalizada, el coeficiente del coseno (o del seno) es la excentricidad e y el numerador es el semilado recto ℓ. Clasifica en el acto: elipse si $e < 1$, parábola si $e = 1$, hipérbola si $e > 1$.

3. **Mira si es coseno o seno, y con qué signo, para saber la orientación.** Con cos el eje de la cónica es horizontal y con sin es vertical. El signo + coloca la directriz del lado positivo y acerca la curva por ahí; el signo − la coloca del lado negativo. Los cuatro casos son $\ell/(1 \pm e\cos\theta)$ y $\ell/(1 \pm e\sin\theta)$.

4. **Calcula los dos ápsides evaluando en los ángulos extremos.** El denominador es máximo y mínimo cuando el coseno vale ± 1, de modo que las distancias mínima y máxima al foco son

$$r_{\text{mín}} = \frac{\ell}{1+e}, \qquad r_{\text{máx}} = \frac{\ell}{1-e} \quad (e < 1).$$

En astronomía se llaman perihelio y afelio, y son los datos que de verdad se miden.

5. **Recupera los semiejes.** Para la elipse,
$$a = \frac{\ell}{1-e^2}, \qquad b = \frac{\ell}{\sqrt{1-e^2}}, \qquad c = ea,$$
y para la hipérbola las mismas con e^2-1 en lugar de $1-e^2$. Se cumplen siempre $r_{\text{mín}} = a-c$ y $r_{\text{máx}} = a+c$, y también la relación $\ell = b^2/a$: úsalas como comprobación.

6. **Sitúa la directriz.** La cónica es el lugar de los puntos cuya distancia al foco es e veces la distancia a una recta fija, y esa recta está en $x = \ell/e$ (en el caso del coseno con signo +). Si $e = 0$ no hay directriz: es una circunferencia.

7. **Para una hipérbola, comprueba dónde hay curva.** Solo existe punto si $1 + e\cos\theta > 0$, es decir $\cos\theta > -1/e$. Fuera de ese rango de ángulos el radio saldría negativo y estarías dibujando la *otra* rama en el sitio equivocado.

Ejemplo 10.29 (Seis cónicas medidas de arriba abajo)**.** (a) **Una elipse escondida:** $r = \dfrac{6}{2+\cos\theta}$**.** Paso 1: dividiendo entre 2,
$$r = \frac{3}{1+\frac{1}{2}\cos\theta}.$$
Paso 2: $e = \frac{1}{2} < 1$, elipse, con $\ell = 3$. Paso 3: coseno con signo +, eje horizontal, curva más cerca del foco por la derecha. Paso 4: $r_{\text{mín}} = 3/(3/2) = 2$ y $r_{\text{máx}} = 3/(1/2) = 6$. Paso 5: $a = 3/(1-\frac{1}{4}) = 4$, $b = 3/\sqrt{3/4} = 2\sqrt{3} \approx 3{,}4641$ y $c = \frac{1}{2}\cdot 4 = 2$. Comprobaciones: $a-c = 2$ ✓, $a+c = 6$ ✓, y $b^2/a = 12/4 = 3 = \ell$ ✓. Paso 6: la directriz es $x = \ell/e = 6$. Quien hubiera leído «$e = 1$, parábola» en la ecuación sin normalizar se habría equivocado en todo.

(b) **Una elipse vertical:** $r = \dfrac{12}{3-2\sin\theta}$**.** Normalizando, $r = \dfrac{4}{1-\frac{2}{3}\sin\theta}$: $e = \frac{2}{3}$, $\ell = 4$, elipse. Por el paso 3 el eje es *vertical* y el signo menos indica que la curva se acerca al foco por abajo. Ápsides: $r_{\text{mín}} = 4/(1+\frac{2}{3}) = \frac{12}{5} = 2{,}4$ (alcanzado en $\theta = 3\pi/2$) y $r_{\text{máx}} = 4/(1-\frac{2}{3}) = 12$ (en $\theta = \pi/2$). Semiejes: $a = 4/(1-\frac{4}{9}) = \frac{36}{5} = 7{,}2$, $b = 4/\sqrt{5/9} = \frac{12}{\sqrt{5}} \approx 5{,}3666$ y $c = \frac{2}{3}\cdot 7{,}2 = 4{,}8$. De nuevo $a-c = 2{,}4$ y $a+c = 12$ ✓.

(c) **Una parábola:** $r = \dfrac{4}{1+\cos\theta}$**.** Ya está normalizada: $e = 1$, $\ell = 4$. El paso 4 falla a medias y eso es informativo: $r_{\text{mín}} = 4/2 = 2$ existe, pero $r_{\text{máx}}$ sería $4/0$, es decir, no existe. Justamente: la parábola es la cónica abierta que no tiene afelio. El vértice está a distancia 2 del foco sobre el eje X positivo, la directriz es $x = \ell/e = 4$, y en efecto el vértice $(2,0)$ equidista del foco $(0,0)$ y de la recta $x = 4$, las dos distancias valen 2 ✓.

(d) **Una hipérbola:** $r = \dfrac{8}{1+3\cos\theta}$**.** Aquí $e = 3 > 1$ y $\ell = 8$. Paso 5 con $e^2-1 = 8$: $a = 8/8 = 1$, $c = ea = 3$ y $b = 8/\sqrt{8} = 2\sqrt{2} \approx 2{,}8284$. El vértice más cercano está a

$r_{\text{mín}} = 8/4 = 2$ del foco, y en efecto $c - a = 3 - 1 = 2$ ✓. Paso 7: hay curva solo si $\cos\theta > -\frac{1}{3}$, o sea para $|\theta| < 109{,}47°$ aproximadamente; en el resto de ángulos no hay nada de esta rama.

(e) **La Tierra y el cometa Halley, con números reales.** Para la Tierra, $e = 0{,}0167$ y $a = 149{,}6$ millones de km, luego $r_{\text{mín}} = a(1-e) = 147{,}10$ y $r_{\text{máx}} = a(1+e) = 152{,}10$. Para el Halley, $e = 0{,}967$ y $a = 17{,}8$ UA, luego $r_{\text{mín}} = 0{,}5874$ UA y $r_{\text{máx}} = 35{,}0126$ UA, con semilado recto $\ell = a(1-e^2) = 1{,}1554$ UA. La órbita de la Tierra es prácticamente circular; la del Halley es un puro cigarro que va de dentro de la órbita de Venus a más allá de Neptuno.

(f) **Un regalo: la tercera ley de Kepler sale del semieje mayor.** Midiendo el tiempo en años y las distancias en unidades astronómicas, $T^2 = a^3$. Con los a del apartado anterior: para el Halley, $T = 17{,}8^{3/2} = 75{,}10$ años, que es el periodo observado; para Mercurio, con $a = 0{,}387$ UA, $T = 0{,}387^{3/2} = 0{,}2408$ años $= 87{,}9$ días, exactamente el valor tabulado; y para Marte, con $a = 1{,}524$, $T = 1{,}8814$ años $= 687{,}2$ días. Tres cifras significativas correctas con una potencia de tres medios.

Cuidado

El paso 1 no es un formalismo. En $r = \dfrac{6}{2+\cos\theta}$ la tentación de leer «$e = 1$, luego parábola» es enorme, y el resultado es una cónica equivocada, unos semiejes inventados y una directriz que no existe. Comprueba siempre que el denominador empieza con un 1 solitario antes de leer cualquier cosa. Y ojo con el numerador: después de dividir, el semilado recto es A/B, no A.

Pausa para pensar

¿Por qué ℓ se llama semilado recto? Porque es exactamente el valor de r cuando $\theta = \pi/2$, es decir, la distancia del foco a la curva medida *perpendicularmente* al eje. La cuerda completa perpendicular al eje que pasa por el foco mide 2ℓ y se llama lado recto. En la elipse del apartado (a), $r(\pi/2) = 3$, así que esa cuerda mide 6: un dato que se puede comprobar en el dibujo con una regla y que sirve para verificar que la ecuación es la correcta.

Idea clave

Merece la pena ver qué gana uno con esta fórmula frente a la geometría analítica de la geometría analítica. Allí, elipse, parábola e hipérbola tenían tres ecuaciones distintas, tres juegos de fórmulas y una tabla de casos; el paso de una a otra era un salto cualitativo. Aquí son una sola ecuación y el paso de una a otra es simplemente mover el número e de forma continua a través de 1. Ésa es la razón profunda de que Newton pudiera demostrar que la fuerza gravitatoria produce cónicas sin tener que estudiar tres casos: la ecuación diferencial de la órbita, resuelta en el capítulo 12, entrega esta fórmula de golpe, y el tipo de cónica lo decide después la energía total del cuerpo.

10.7. Ecuaciones paramétricas: cuando hace falta el tiempo

Una ecuación como $y = f(x)$ solo puede describir curvas que cortan cada vertical una vez: no sirve para una circunferencia. Una ecuación implícita $F(x, y) = 0$ sí describe la circunferencia, pero no dice *en qué orden* se recorre. La solución a ambos problemas es la misma idea: introducir una variable auxiliar.

Definición 10.30. Una **curva paramétrica** en el plano es un par de funciones

$$x = x(t), \qquad y = y(t), \qquad t \in I,$$

definidas en un intervalo I. La variable t se llama **parámetro**, y en la mayoría de las aplicaciones representa el tiempo.

Parametrización	Curva
$x = x_0 + at$, $y = y_0 + bt$	recta por (x_0, y_0) con dirección (a, b)
$x = R\cos t$, $y = R\sin t$	circunferencia de radio R, sentido antihorario
$x = a\cos t$, $y = b\sin t$	elipse de semiejes a y b
$x = t$, $y = f(t)$	la gráfica de f de toda la vida
$x = a\cosh t$, $y = b\sinh t$	una rama de hipérbola

Regla: eliminar el parámetro

Para pasar de paramétricas a cartesianas:

1. Si es fácil, despeje t de una ecuación y sustitúyalo en la otra.
2. Si aparecen seno y coseno, no despeje: use $\cos^2 t + \sin^2 t = 1$.
3. **Compruebe el recorrido**. La cartesiana resultante suele describir *más* curva que la paramétrica original.

Ejemplo 10.31. $x = t + 1$, $y = t^2$. Despejando, $t = x - 1$, luego $y = (x - 1)^2$: una parábola con vértice en $(1, 0)$. Aquí no se pierde nada porque t recorre todo $\mathbb{R}$.

Ejemplo 10.32. $x = 3\cos t$, $y = 2\sin t$. Entonces $\dfrac{x^2}{9} + \dfrac{y^2}{4} = \cos^2 t + \sin^2 t = 1$: la elipse de semiejes 3 y 2. En $t = \pi/3$ el punto es $(1{,}5;\ \sqrt{3}\,)$, y se comprueba que $\frac{2{,}25}{9} + \frac{3}{4} = 1$. ✓

Cuidado

[Lo que se pierde al eliminar el parámetro] Considere $x = \cos t$, $y = \cos^2 t$ con $t \in \mathbb{R}$. Eliminando, $y = x^2$. Pero la curva paramétrica **no** es la parábola entera: como $\cos t \in [-1, 1]$, solo se recorre el trocito con $-1 \le x \le 1$, y además se recorre infinitas veces, de ida y de vuelta. La ecuación cartesiana ha perdido tres informaciones: el dominio, la orientación y el número de pasadas. Por eso una parametrización es *más* información

que una ecuación, no menos.

10.8. Tiro parabólico: la parametrización que sí importa

Regla: movimiento de un proyectil

Un objeto lanzado desde el origen con rapidez v_0 y ángulo α sobre la horizontal, despreciando el rozamiento, describe

$$x(t) = v_0 \cos\alpha \, t, \qquad y(t) = v_0 \sin\alpha \, t - \tfrac{1}{2} g t^2.$$

De ahí salen, sin más que igualar $y = 0$ o derivar:

$$T = \frac{2 v_0 \sin\alpha}{g}, \qquad R = \frac{v_0^2 \sin 2\alpha}{g}, \qquad H = \frac{v_0^2 \sin^2\alpha}{2g}.$$

Ejemplo 10.33. Con $v_0 = 20$ m/s, $\alpha = 30°$ y $g = 9{,}8$ m/s^2:

$$T = \frac{2 \cdot 20 \cdot 0{,}5}{9{,}8} = 2{,}04 \text{ s}, \quad R = \frac{400 \cdot \sin 60°}{9{,}8} = 35{,}35 \text{ m}, \quad H = \frac{400 \cdot 0{,}25}{19{,}6} = 5{,}10 \text{ m}.$$

Idea clave

La fórmula del alcance contiene $\sin 2\alpha$, que vale 1 cuando $2\alpha = 90°$: el alcance máximo se consigue con 45°, y vale $v_0^2/g = 40{,}8$ m en el ejemplo anterior. Y como $\sin 2\alpha = \sin(180° - 2\alpha)$, los ángulos α y $90° - \alpha$ dan el mismo alcance: 30° y 60° llegan igual de lejos, aunque por trayectorias muy distintas —una tensa y rápida, otra alta y lenta—. Esa observación, tan escolar, es la que permitía a los artilleros del siglo XVII elegir entre dos elevaciones para el mismo blanco.

Cuidado

Todo esto vale *sin rozamiento*. Con resistencia del aire la trayectoria deja de ser una parábola, el alcance máximo se consigue con ángulos bastante menores de 45° (unos 35° para una bala real) y la curva de bajada es más empinada que la de subida. El modelo escolar es útil precisamente porque es falso de una manera controlada: conviene saber en qué dirección miente.

10.9. La cicloide y sus dos milagros

Clave una chincheta en el borde de una rueda y hágala rodar sin deslizar sobre una recta. La curva que describe la chincheta es la **cicloide**, y su parametrización natural es

$$x(t) = R(t - \sin t), \qquad y(t) = R(1 - \cos t),$$

donde t es el ángulo girado por la rueda. En cartesianas es imposible de escribir razonablemente; en paramétricas, dos líneas.

Observación 10.34. Un arco completo (de $t = 0$ a $t = 2\pi$) tiene longitud exactamente $8R$ y encierra un área de exactamente $3\pi R^2$, es decir el triple del área del círculo que rueda. Los dos resultados son de Roberval y de Wren, del siglo XVII, y las dos demostraciones son ejercicios estándar del capítulo 6 en cuanto se dispone de la integral.

Idea clave

La cicloide tiene dos propiedades que parecen imposibles y que la convirtieron en la "Helena de la geometría", por las peleas que causó entre matemáticos:

- **Braquistócrona**: entre todas las curvas que unen dos puntos, la cicloide invertida es la que una bola recorre en *menos tiempo* bajo la gravedad. No es la recta. Bajar por un tobogán curvo, si la curva es la correcta, es más rápido que bajar por la rampa directa.
- **Tautócrona**: una bola soltada en cualquier punto de una cicloide invertida tarda *lo mismo* en llegar al fondo, suelte usted desde arriba o desde muy cerca del final.

Huygens usó la segunda propiedad en 1673 para diseñar un péndulo cuyo periodo no dependiera de la amplitud, y el problema de la primera, planteado por Johann Bernoulli en 1696, fue el acta de nacimiento del cálculo de variaciones.

Regla: derivar una curva dada en paramétricas: pendiente, concavidad y puntos raros

Cuando la curva viene como $x = x(t)$, $y = y(t)$, no hay una y en función de x que derivar. Pero la regla de la cadena resuelve el problema sin necesidad de despejar.

1. **La pendiente de la tangente es el cociente de las derivadas.** Como $\dfrac{\mathrm{d}y}{\mathrm{d}t} = \dfrac{\mathrm{d}y}{\mathrm{d}x} \cdot \dfrac{\mathrm{d}x}{\mathrm{d}t}$, despejando queda
$$\frac{\mathrm{d}y}{\mathrm{d}x} = \frac{\mathrm{d}y/\mathrm{d}t}{\mathrm{d}x/\mathrm{d}t}, \qquad \text{siempre que } \frac{\mathrm{d}x}{\mathrm{d}t} \neq 0.$$
2. **Tangentes horizontales y verticales, por separado.** Horizontal donde $\mathrm{d}y/\mathrm{d}t = 0$ y $\mathrm{d}x/\mathrm{d}t \neq 0$; vertical donde $\mathrm{d}x/\mathrm{d}t = 0$ y $\mathrm{d}y/\mathrm{d}t \neq 0$. Si las dos se anulan a la vez, el punto es *singular* y hay que estudiarlo aparte (paso 6).
3. **Para la segunda derivada, vuelve a aplicar la regla, no la copies.** El error clásico es escribir $\ddot{y}/\ddot{x}$, que está mal. Lo correcto es derivar respecto de t la pendiente ya calculada y volver a dividir:
$$\frac{\mathrm{d}^2y}{\mathrm{d}x^2} = \frac{\dfrac{\mathrm{d}}{\mathrm{d}t}\left(\dfrac{\mathrm{d}y}{\mathrm{d}x}\right)}{\dfrac{\mathrm{d}x}{\mathrm{d}t}}.$$

4. **Escribe la recta tangente con el punto y la pendiente.** Evalúa $x(t_0)$, $y(t_0)$ y la pendiente en t_0, y usa $y - y_0 = m(x - x_0)$ como siempre.

5. **Separa velocidad de pendiente.** El vector velocidad es $(\dot{x}, \dot{y})$ y su módulo, la rapidez, es $\sqrt{\dot{x}^2 + \dot{y}^2}$. La pendiente solo depende de la *dirección*; dos parametrizaciones distintas de la misma curva dan la misma pendiente y rapideces distintas.

6. **Estudia los puntos donde $\dot{x} = \dot{y} = 0$.** Ahí la fórmula del paso 1 no dice nada. Calcula el límite de la pendiente cuando $t \to t_0$: si existe, hay tangente; si además la curva vuelve sobre sí misma, es una cúspide.

7. **Verifica.** Si puedes eliminar el parámetro, deriva implícitamente la ecuación cartesiana y comprueba que sale la misma pendiente. Y la longitud de arco es $L = \int \sqrt{\dot{x}^2 + \dot{y}^2}\,\mathrm{d}t$: es la integral de la rapidez, o sea el espacio recorrido.

Ejemplo 10.35 (Seis curvas, un punto donde la curva se corta a sí misma y una cúspide)**.** (a) **La elipse, con verificación cruzada.** Sea $x = 3\cos t$, $y = 2\sin t$. Entonces $\dot{x} = -3\sin t$, $\dot{y} = 2\cos t$ y

$$\frac{\mathrm{d}y}{\mathrm{d}x} = \frac{2\cos t}{-3\sin t}.$$

En $t = \pi/4$: el punto es $\left(\frac{3\sqrt{2}}{2},\ \sqrt{2}\right) \approx (2{,}1213;\ 1{,}4142)$ y la pendiente es $\dfrac{2 \cdot \frac{\sqrt{2}}{2}}{-3 \cdot \frac{\sqrt{2}}{2}} = -\dfrac{2}{3}$.

Verificación del paso 7: eliminando el parámetro, $\dfrac{x^2}{9} + \dfrac{y^2}{4} = 1$, y derivando implícitamente, $\dfrac{2x}{9} + \dfrac{2y}{4}y' = 0$, de donde $y' = -\dfrac{4x}{9y} = -\dfrac{4 \cdot \frac{3\sqrt{2}}{2}}{9\sqrt{2}} = -\dfrac{6\sqrt{2}}{9\sqrt{2}} = -\dfrac{2}{3}$ ✓. Tangentes horizontales en $t = \pi/2$ y $t = 3\pi/2$ (los extremos del eje menor), verticales en $t = 0$ y $t = \pi$ (los del eje mayor), como manda el paso 2.

(b) **La cicloide y su cúspide.** Sea $x = t - \sin t$, $y = 1 - \cos t$. Entonces $\dot{x} = 1 - \cos t$, $\dot{y} = \sin t$ y

$$\frac{\mathrm{d}y}{\mathrm{d}x} = \frac{\sin t}{1 - \cos t} = \cot\frac{t}{2},$$

usando las identidades del ángulo mitad. En $t = \pi/2$ la pendiente es $\cot(\pi/4) = 1$, y en efecto $\sin(\pi/2) = 1$ y $1 - \cos(\pi/2) = 1$. En $t = \pi$, punto más alto, la pendiente es $\cot(\pi/2) = 0$: tangente horizontal ✓. Y cuando $t \to 0^+$ la pendiente tiende a $+\infty$: la rueda toca el suelo con tangente *vertical*. Justo ahí $\dot{x} = \dot{y} = 0$, el caso singular del paso 6: es una cúspide.

(c) **La longitud de un arco de cicloide.** Por el paso 7,

$$\dot{x}^2 + \dot{y}^2 = (1 - \cos t)^2 + \sin^2 t = 2 - 2\cos t = 4\sin^2\frac{t}{2},$$

luego $L = \int_0^{2\pi} 2\left|\sin\frac{t}{2}\right| \mathrm{d}t = 8$. El arco de cicloide mide exactamente 8 veces el radio de la rueda: cuatro diámetros. Y el área bajo el arco es $\int_0^{2\pi}(1 - \cos t)^2 \mathrm{d}t = 3\pi$, tres veces el

área del círculo que la genera. Dos resultados exactos y elementales para una curva que ni siquiera tiene ecuación cartesiana decente.

(d) **Una curva que se corta a sí misma.** Sea $x = t^2$, $y = t^3 - 3t$. La pendiente es

$$\frac{\mathrm{d}y}{\mathrm{d}x} = \frac{3t^2 - 3}{2t} = \frac{3(t^2-1)}{2t}.$$

Se anula en $t = \pm 1$: tangentes horizontales en $(1,-2)$ y $(1,2)$. Pero además, $t = \sqrt{3}$ y $t = -\sqrt{3}$ dan *el mismo punto* $(3,0)$, porque x solo depende de t^2 y $y(\pm\sqrt{3}) = \pm(3\sqrt{3} - 3\sqrt{3}) = 0$. La curva pasa dos veces por ahí, con pendientes $\dfrac{3(3-1)}{2\sqrt{3}} = \sqrt{3}$ y $-\sqrt{3}$: dos tangentes distintas en el mismo punto, un *nodo*. La función y de x no existe cerca de $(3,0)$, y sin embargo la pendiente en cada rama sí.

(e) **La segunda derivada, bien hecha.** Con la misma curva, escribimos la pendiente como $\frac{3t}{2} - \frac{3}{2t}$, la derivamos respecto de t —$\frac{3}{2} + \frac{3}{2t^2}$— y dividimos entre $\dot{x} = 2t$:

$$\frac{\mathrm{d}^2y}{\mathrm{d}x^2} = \frac{\frac{3}{2} + \frac{3}{2t^2}}{2t} = \frac{3(t^2+1)}{4t^3}.$$

En $t = 1$ vale $\frac{3\cdot 2}{4} = \frac{3}{2} > 0$: cóncava hacia arriba, y como la pendiente era 0, el punto $(1,-2)$ es un mínimo local de la rama. En $t = -1$ vale $-\frac{3}{2} < 0$: el punto $(1,2)$ es un máximo de la otra rama. Nota que el signo lo decide t^3, es decir el sentido en que se recorre la curva. *El error que evita el paso 3:* el cociente $\ddot{y}/\ddot{x} = 6t/2 = 3t$ da en $t = 1$ el valor 3, que no es la concavidad ni se le parece.

(f) **Misma curva, distinta rapidez.** Las parametrizaciones $x = \cos t$, $y = \sin t$ y $x = \cos 2t$, $y = \sin 2t$ describen la misma circunferencia. La pendiente en el punto $(0,1)$ es 0 en ambas. Pero la rapidez es 1 en la primera y 2 en la segunda, y la longitud recorrida en $t \in [0, 2\pi]$ es 2π en la primera y 4π en la segunda, porque da dos vueltas. La geometría es la misma; el movimiento, no.

Cuidado

Escribir $\dfrac{\mathrm{d}^2y}{\mathrm{d}x^2} = \dfrac{\ddot{y}}{\ddot{x}}$ es el error más frecuente de todo el tema, y es tentador porque la fórmula de la primera derivada sí es un cociente. La razón de que falle es que $\mathrm{d}y/\mathrm{d}x$ ya es una función nueva, y para derivarla respecto de x hay que volver a pasar por t con la regla de la cadena, exactamente como en el paso 3. Compruébalo en el apartado (e): la fórmula correcta da $3/2$ y la falsa da 3.

Pausa para pensar

Los apartados (b) y (d) muestran dos maneras distintas de que una curva perfectamente lisa como objeto geométrico deje de ser gráfica de una función. En (d) el problema es que hay dos ramas superpuestas; en (b), que la tangente se pone vertical y además la curva retrocede. En los dos casos la parametrización sigue funcionando sin ninguna dificultad. Esa es exactamente la ventaja de las paramétricas: describen la curva como un *recorrido*,

y un recorrido puede cruzarse, dar la vuelta o pararse sin dejar de estar bien definido.

Idea clave

La condición $\dot{x} = \dot{y} = 0$ del paso 6 es la que separa las parametrizaciones *regulares* de las demás, y es la hipótesis que hay que pedir para que la curva tenga tangente en todos sus puntos. Es también la hipótesis del teorema de la función implícita, que garantiza que cerca de un punto donde el gradiente no se anula la curva sí es localmente la gráfica de una función. Todo el capítulo 9 descansa sobre esa misma condición, y su versión en dimensión superior es la definición de *variedad diferenciable*, que es donde empieza la geometría del capítulo 36.

10.10. Velocidad, aceleración y curvatura de una curva paramétrica

Si el parámetro es el tiempo, las derivadas de una parametrización dejan de ser un ejercicio y se convierten en física: la primera es la velocidad y la segunda la aceleración. Y hay una tercera magnitud, la curvatura, que no depende de la velocidad con que se recorra la curva sino solo de su forma, y que mide exactamente lo cerrada que es una curva en cada punto.

Definición 10.36. Dada una curva $\vec{c}(t) = (x(t), y(t))$ de clase C^2, el *vector velocidad* es $\vec{v} = (x'(t), y'(t))$, la *rapidez* es $\|\vec{v}\| = \sqrt{x'^2 + y'^2}$ y el *vector aceleración* es $\vec{a} = (x''(t), y''(t))$. La *curvatura* en un punto es

$$\kappa = \frac{|x'y'' - y'x''|}{(x'^2 + y'^2)^{3/2}},$$

y su inverso $R = 1/\kappa$ es el *radio de curvatura*: el radio de la circunferencia que mejor se ajusta a la curva en ese punto, llamada *circunferencia osculatriz*. Un punto donde $\vec{v} = \vec{0}$ se llama *singular*, y allí la curva puede tener una *cúspide*.

Idea clave

Conviene separar dos cosas que la notación mezcla. La velocidad depende de *cómo* se recorre la curva: el mismo círculo recorrido despacio o deprisa da vectores velocidad distintos. La curvatura, en cambio, depende solo del *dibujo*: si se reparametriza la curva, κ no cambia en cada punto geométrico. Por eso la fórmula de la curvatura lleva un $(x'^2 + y'^2)^{3/2}$ en el denominador, que es precisamente lo que hace falta para cancelar el efecto de la velocidad. La forma más limpia de verlo es imaginar que se recorre la curva a rapidez constante 1: entonces la aceleración es puramente centrípeta y su módulo *es* la curvatura. Todo lo demás es corrección por ir más deprisa o más despacio.

Regla: analizar el movimiento sobre una curva paramétrica

Ante una curva dada por $x(t)$ e $y(t)$, este es el orden de trabajo:

1. **Deriva las dos componentes** para obtener $\vec{v}$, y vuelve a derivar para obtener $\vec{a}$; son dos derivadas ordinarias, nada más.
2. **Calcula la rapidez** $\|\vec{v}\|$ y localiza dónde se anula: esos son los puntos singulares, candidatos a cúspide.
3. **Obtén la pendiente de la tangente** como $\dfrac{dy}{dx} = \dfrac{y'(t)}{x'(t)}$ cuando $x'(t) \neq 0$; donde $x' = 0$ y $y' \neq 0$ la tangente es vertical.
4. **Aplica la fórmula de la curvatura** sin simplificar a medias: el numerador es un determinante 2×2 y el denominador el cubo de la rapidez.
5. **Interpreta** $R = 1/\kappa$: cuanto mayor es R, más recta es la curva allí; en una recta $\kappa = 0$ y R es infinito.
6. **Descompón la aceleración** en su parte tangencial (que cambia la rapidez) y su parte normal (que cambia la dirección y vale $\kappa\|\vec{v}\|^2$).
7. **Calcula la longitud recorrida** con $\displaystyle\int_a^b \|\vec{v}(t)\|\,\mathrm{d}t$ cuando haga falta, como se hace en el capítulo 9.

Ejemplo 10.37 (Cuatro curvas y sus tres magnitudes)**. (a) La circunferencia, el caso de referencia.** Con $\vec{c}(t) = (3\cos t, 3\sin t)$ se tiene $\vec{v} = (-3\sin t, 3\cos t)$, de rapidez constante 3, y $\vec{a} = (-3\cos t, -3\sin t) = -\vec{c}(t)$, siempre apuntando al centro. La curvatura es $\kappa = \frac{9}{27} = \frac{1}{3}$ en todo punto, y $R = 3$: la circunferencia osculatriz de una circunferencia es ella misma. Obsérvese que $\vec{v} \cdot \vec{a} = 0$ siempre, señal de que la rapidez no cambia.

(b) La parábola $y = x^2$. Con $\vec{c}(t) = (t, t^2)$ resulta $\vec{v} = (1, 2t)$, $\|\vec{v}\| = \sqrt{1 + 4t^2}$, $\vec{a} = (0, 2)$ constante, y $\kappa = \dfrac{2}{(1 + 4t^2)^{3/2}}$. En el vértice $t = 0$ vale $\kappa = 2$ y $R = 0{,}5$: la circunferencia osculatriz tiene centro en $(0; 0{,}5)$. En $t = 1$ la curvatura ya ha bajado a 0,178885 y $R = 5{,}590170$; en $t = 2$, a 0,028534. La parábola se endereza muy deprisa al alejarse del vértice.

(c) La cicloide y su cúspide. Con $\vec{c}(t) = (t - \sin t,\ 1 - \cos t)$ la rapidez es $\sqrt{2 - 2\cos t}$, que se anula en $t = 0$ y $t = 2\pi$: ahí está la cúspide, el punto en que la rueda toca el suelo y su velocidad instantánea es cero. En $t = \pi$ (lo más alto) la rapidez es 2, el doble de la del centro de la rueda, y la curvatura vale 0,25. En $t = \pi/2$ la rapidez es 1,414214 y la curvatura 0,353553.

(d) Una cúspide sin rueda. La curva $\vec{c}(t) = (t^2, t^3)$ tiene $\vec{v} = (2t, 3t^2)$, que se anula en $t = 0$: el origen es un punto singular y en él la curva presenta una punta afilada, aunque ambas componentes sean polinomios perfectamente suaves. Es el aviso de que la suavidad de $x(t)$ e $y(t)$ no garantiza la suavidad del *dibujo*: en $t = 1$ todo es normal ($\|\vec{v}\| = 3{,}605551$, $\kappa = 0{,}128008$), pero en el origen la parametrización se detiene y gira.

Ejemplo 10.38 (Longitudes y un proyectil medido a fondo)**. (a) Longitud de un arco de cicloide.** Con radio 1,

$$L = \int_0^{2\pi} \sqrt{2 - 2\cos t}\,\mathrm{d}t = 8.$$

Ocho veces el radio, sin ningún π a la vista pese a tratarse de una curva engendrada por una circunferencia: es uno de los dos "milagros" de la cicloide, y fue el primer ejemplo histórico de rectificación exacta de una curva trascendente.

(b) Longitud de un arco de parábola. Para $\vec{c}(t) = (t, t^2)$ entre $t = 0$ y $t = 1$,

$$L = \int_0^1 \sqrt{1 + 4t^2}\,\mathrm{d}t = \frac{\sqrt{5}}{2} + \frac{\operatorname{arcsinh} 2}{4} = 1{,}478943,$$

algo más que la cuerda, que mide $\sqrt{2} = 1{,}414214$. Aquí sí aparece una función trascendente: casi ninguna curva algebraica tiene longitud elemental, y la cicloide es la excepción llamativa.

(c) Un proyectil, con todos sus números. Lanzamos a 20 m/s con un ángulo de 40°: $v_x = 15{,}320889$ y $v_{y0} = 12{,}855752$ m/s. La altura máxima es $\frac{v_{y0}^2}{2g} = 8{,}432161$ m, alcanzada en $t = 1{,}311811$ s; el tiempo total de vuelo es 2,623623 s y el alcance 40,196235 m. En $t = 1$ s la rapidez ya ha bajado a 15,622652 m/s, porque la componente vertical se ha reducido a 3,055752 m/s.

(d) La curvatura del proyectil en su punto más alto. Ahí la velocidad es horizontal y vale v_x, y la aceleración es vertical y vale g; por tanto $\kappa = \dfrac{g}{v_x^2} = \dfrac{9{,}8}{234{,}729} = 0{,}041750$ y el radio de curvatura es 23,952004 m. Es decir: en la cima, la trayectoria se confunde con una circunferencia de casi veinticuatro metros de radio. Y como κ solo depende de v_x, dos disparos con la misma componente horizontal tienen exactamente la misma curvatura en la cima aunque sus alturas sean muy distintas.

Cuidado

La curvatura no es la segunda derivada: para $y = f(x)$ vale $\kappa = \dfrac{|f''|}{(1 + f'^2)^{3/2}}$, y solo coincide con $|f''|$ donde la tangente es horizontal. Segundo aviso: el módulo del vector aceleración no mide la variación de la rapidez; en un movimiento circular uniforme la rapidez es constante y la aceleración no es nula en absoluto. Y tercero: un punto singular ($\vec{v} = \vec{0}$) no siempre es una cúspide, pero toda cúspide de una parametrización suave es un punto singular; conviene comprobar el comportamiento de la tangente a ambos lados antes de afirmar nada.

Pausa para pensar

La circunferencia osculatriz la introdujo Leibniz en 1686 con ese nombre latino, *osculum*, beso: la circunferencia que besa a la curva. Newton había llegado a la misma idea unos años antes por el camino de la mecánica, buscando la fuerza centrípeta necesaria para mantener un cuerpo en una trayectoria dada. Que dos formulaciones tan distintas (una geométrica y otra dinámica) produjeran la misma magnitud no es casual: la curvatura es la traducción exacta de "cuánta fuerza hace falta para no seguir recto". La generalización

a superficies, obra de Gauss en 1827, abrió la puerta a la geometría diferencial y, por ese camino, a la relatividad general.

10.11. La familia de las ruletas: astroide, deltoide, epicicloides y caracoles

La cicloide se obtiene haciendo rodar una circunferencia sobre una recta. Si en vez de una recta la hacemos rodar sobre otra circunferencia (por dentro o por fuera), aparece una familia entera de curvas con nombres propios que llenaron los tratados del siglo XVII y que hoy siguen dibujando los engranajes, los espirógrafos y los perfiles de leva.

Definición 10.39. Si una circunferencia de radio r rueda sin deslizar por el *interior* de otra de radio R, un punto de la pequeña describe una *hipocicloide*

$$x = (R-r)\cos t + r\cos\left(\frac{R-r}{r}t\right), \qquad y = (R-r)\sin t - r\sin\left(\frac{R-r}{r}t\right);$$

si rueda por el *exterior*, describe una *epicicloide*, con $R+r$ en lugar de $R-r$ y el signo del segundo seno cambiado. El caso $R/r = 4$ se llama *astroide* y el caso $R/r = 3$, *deltoide*. Por su parte, el *caracol de Pascal* es la curva polar $r = b + a\cos\theta$, que para $b = a$ es la cardioide.

Idea clave

Toda esta familia se genera con la misma receta: sumar dos movimientos circulares. Uno es el del centro de la rueda pequeña, que da el término $(R \mp r)$; el otro es el giro del punto marcado alrededor de ese centro, que da el término con r. Cambiar la razón R/r cambia la frecuencia relativa de los dos giros, y con ella el número de puntas. El resultado es sorprendentemente rígido: si $R/r = p/q$ en forma irreducible, la curva tiene exactamente p cúspides y se cierra tras q vueltas. Cuando R/r es irracional la curva no se cierra nunca y acaba llenando una corona circular entera, que es justamente lo que se ve cuando un espirógrafo tiene engranajes mal emparejados.

Regla: identificar y medir una ruleta

Ante una curva de esta familia, el procedimiento es:

1. **Decide si rueda por dentro o por fuera**, es decir si es hipocicloide o epicicloide; el signo del término $R \mp r$ lo delata.
2. **Simplifica la fracción** $R/r = p/q$ y anota que habrá p cúspides y que hará falta recorrer q vueltas para cerrarla.
3. **Localiza las cúspides** resolviendo $\|\vec{v}(t)\| = 0$; están igualmente espaciadas en ángulo.
4. **Busca una ecuación implícita** si la hay: la astroide cumple $x^{2/3} + y^{2/3} = a^{2/3}$, que es mucho más manejable que la paramétrica.

5. **Calcula la longitud** con $\int \|\vec{v}\|\,\mathrm{d}t$ aprovechando la simetría: basta un arco y multiplicar.
6. **Calcula el área** con la fórmula de Green $A = \frac{1}{2}\left|\oint (x\,\mathrm{d}y - y\,\mathrm{d}x)\right|$ o, en polares, con $A = \frac{1}{2}\int r^2\,\mathrm{d}\theta$.
7. **Comprueba con dos o tres puntos concretos** sustituyendo valores del parámetro; es la manera más rápida de detectar un signo equivocado.

Ejemplo 10.40 (Astroide y deltoide, medidas de arriba abajo)**. (a) La astroide.** Con $R = 4$, $r = 1$ la parametrización se simplifica a $x = \cos^3 t$, $y = \sin^3 t$. Los puntos $t = 0, \frac{\pi}{2}, \pi, \frac{3\pi}{2}$ dan las cuatro cúspides $(\pm 1, 0)$ y $(0, \pm 1)$, y en $t = \pi/4$ la curva pasa por (0,353553; 0,353553). La ecuación implícita $x^{2/3} + y^{2/3} = 1$ se comprueba ahí: $0{,}353553^{2/3} \cdot 2 = 1$ exactamente.

(b) Su longitud y su área. La rapidez es $3|\sin t \cos t|$, y por simetría

$$L = 4\int_0^{\pi/2} 3\sin t\cos t\,\mathrm{d}t = 6,$$

seis veces el radio del círculo grande dividido entre cuatro, o dicho de otro modo: $6a$ con $a = 1$. El área encerrada, con la fórmula de Green, resulta ser $\frac{3\pi}{8} = 1{,}178097$, algo más de un tercio del área del círculo circunscrito, que vale π.

(c) La deltoide. Con $R = 3$, $r = 1$ queda $x = 2\cos t + \cos 2t$, $y = 2\sin t - \sin 2t$. Las tres cúspides están en $t = 0, \frac{2\pi}{3}, \frac{4\pi}{3}$, es decir en $(3, 0)$, $(-1{,}5;\ 2{,}598076)$ y $(-1{,}5;\ -2{,}598076)$: forman un triángulo equilátero, como el nombre sugiere. Es la curva que describe el punto medio del lado de un triángulo de Morley y la que se usa como perfil en ciertos mecanismos de pistón rotativo.

(d) La epicicloide con $R = r$. Rodando por fuera con radios iguales sale $x = 2\cos t - \cos 2t$, $y = 2\sin t - \sin 2t$, que es la cardioide: una sola cúspide, en $(1, 0)$, y el punto más alejado en $(-3, 0)$. Su longitud es 16, es decir $16r$, y su forma es la misma que la de la curva polar $\rho = 1 + \cos\theta$ salvo un giro y una escala. Que la cardioide aparezca tanto como epicicloide como en el conjunto de Mandelbrot y como diagrama de un micrófono no es casualidad: es la curva más simple que tiene exactamente una punta.

Ejemplo 10.41 (Caracoles de Pascal y figuras de Lissajous)**. (a) Las cuatro formas del caracol.** La curva $r = b + a\cos\theta$ cambia de aspecto según la razón b/a. Con $b = 1, a = 2$ (razón $0{,}5 < 1$) hay un lazo interior, y en efecto $r(\pi) = -1 < 0$. Con $b = a = 1$ es la cardioide y $r(\pi) = 0$: la cúspide. Con $1 < b/a < 2$ aparece un hoyuelo, y con $b/a \geq 2$ (por ejemplo $b = 3, a = 1$) la curva es convexa, sin entrantes. En los cuatro casos $r(0) = b + a$ marca el punto más lejano.

(b) El área de un caracol. Para $r = 2 + \cos\theta$, que es convexo,

$$A = \frac{1}{2}\int_0^{2\pi} (2+\cos\theta)^2\,\mathrm{d}\theta = \frac{1}{2}\int_0^{2\pi} \left(4 + 4\cos\theta + \cos^2\theta\right)\mathrm{d}\theta = \frac{1}{2}\left(8\pi + 0 + \pi\right) = \frac{9\pi}{2} = 14{,}137167.$$

El truco es recordar que $\int_0^{2\pi}\cos\theta\,\mathrm{d}\theta = 0$ y $\int_0^{2\pi}\cos^2\theta\,\mathrm{d}\theta = \pi$; con esos dos valores se resuelven casi todas las áreas polares de este tipo.

(c) Cuándo hay lazo y cómo medirlo. Si $b < a$, el valor r se vuelve negativo para los θ que cumplen $\cos\theta < -b/a$, y esos puntos se dibujan en la dirección opuesta: eso es el lazo. Los límites del lazo son $\theta = \pm \arccos(-b/a)$, y el área del lazo se calcula integrando entre ellos. Para $b = 1, a = 2$ los límites son $\theta = \pm 120°$.

(d) Figuras de Lissajous. Si en vez de sumar dos giros multiplicamos las frecuencias en cada eje por separado, $x = \sin(3t)$, $y = \sin(2t)$, aparece una curva cerrada con tres lóbulos en una dirección y dos en la otra; en $t = 0{,}5$ pasa por $(0{,}997495;\ 0{,}841471)$ y en $t = 1$ por $(0{,}141120;\ 0{,}909297)$. Se cierra en $t = 2\pi$ porque la razón 3:2 es racional. Estas figuras se veían en los osciloscopios antiguos y servían para medir frecuencias comparando dos señales: si la figura quedaba quieta, la razón era exacta.

Cuidado

En una hipocicloide, la frecuencia del segundo término es $\frac{R-r}{r}$ y no $\frac{R}{r}$; equivocarse ahí produce una curva parecida pero con otro número de puntas. Segundo aviso: en polares, un valor negativo de r *no* significa que no haya punto, sino que el punto se dibuja en la dirección opuesta; ignorarlo hace perder el lazo del caracol. Y tercero: para calcular el área de una curva polar con lazo hay que integrar por tramos y decidir qué se quiere medir (el lazo, la región exterior o toda la mancha), porque la integral ciega de $\frac{1}{2}r^2$ cuenta dos veces la parte solapada.

Pausa para pensar

El caracol lleva el nombre de Étienne Pascal, padre de Blaise, que lo estudió hacia 1630; fue Roberval quien le puso el nombre. Las epicicloides tienen una historia mucho más antigua y menos inocente: los astrónomos griegos, y luego Ptolomeo en el siglo II, describían el movimiento de los planetas como círculos rodando sobre círculos, precisamente para explicar el movimiento retrógrado sin abandonar la circularidad perfecta. El modelo era falso pero muy preciso, y no por casualidad: sumar giros de distintas frecuencias es exactamente lo que hoy llamamos serie de Fourier, y con suficientes epiciclos se puede aproximar cualquier órbita cerrada. Ptolomeo estaba haciendo análisis armónico mil setecientos años antes de Fourier.

10.12. Áreas, longitudes y superficies de una curva paramétrica

Una curva dada en paramétricas no necesita despejarse para poder medirse. Todas las magnitudes geométricas (área encerrada, longitud, superficie y volumen de revolución) se calculan directamente con $x(t)$, $y(t)$ y sus derivadas. Esta sección reúne las cuatro fórmulas y las pone a trabajar sobre dos curvas clásicas que en cartesianas serían inmanejables: la astroide y la cicloide.

Regla: medir una curva paramétrica: área, longitud, superficie y volumen

1. **Fija el intervalo y comprueba qué tipo de curva tienes.** Anota $t \in [t_0, t_1]$. Pregúntate tres cosas: ¿es cerrada, es decir $x(t_0) = x(t_1)$ y $y(t_0) = y(t_1)$?; ¿es simple, o se corta a sí misma?; ¿en qué sentido se recorre? Las fórmulas de área exigen curva cerrada y simple; el sentido decide el signo.

2. **Área encerrada por una curva cerrada: fórmula de Green.** Si la curva es cerrada y simple,
$$A = \frac{1}{2}\oint (x\,y' - y\,x')\,\mathrm{d}t = \frac{1}{2}\int_{t_0}^{t_1} (x(t)y'(t) - y(t)x'(t))\,\mathrm{d}t.$$
Sale positiva si se recorre en sentido antihorario y negativa en el otro; si dudas del sentido, toma el valor absoluto.

3. **Área bajo un arco no cerrado.** Si la curva no se cierra pero avanza de izquierda a derecha sobre el eje OX (es decir $x'(t) \geq 0$), el área entre la curva y ese eje es $A = \int_{t_0}^{t_1} y(t)\,x'(t)\,\mathrm{d}t$, que no es más que $\int y\,\mathrm{d}x$ con el cambio de variable ya hecho.

4. **Longitud de arco.** $L = \displaystyle\int_{t_0}^{t_1} \sqrt{x'(t)^2 + y'(t)^2}\,\mathrm{d}t$. El paso decisivo casi nunca es integrar, sino *simplificar el radicando*: saca factor común, usa $\sin^2 + \cos^2 = 1$ y recuerda $1 - \cos t = 2\sin^2(t/2)$. Si el radicando no se convierte en un cuadrado perfecto, la integral suele no ser elemental.

5. **Superficie de revolución.** Girando alrededor de OX, $S = 2\pi \int |y| \sqrt{x'^2 + y'^2}\,\mathrm{d}t$; alrededor de OY, se cambia $|y|$ por $|x|$. Es *longitud por perímetro recorrido*: cada trocito $\mathrm{d}s$ describe una circunferencia de radio $|y|$.

6. **Volumen de revolución.** Girando alrededor de OX, $V = \pi \int y(t)^2\,x'(t)\,\mathrm{d}t$, con el intervalo recorrido en el sentido en que x crece. Es el método de discos con el cambio de variable incorporado.

7. **Contrasta siempre con un caso conocido y con las dimensiones.** La circunferencia y la esfera son el patrón de calibración: si tu fórmula no devuelve πR^2, $2\pi R$ y $4\pi R^2$, está mal. Y comprueba las potencias: una longitud debe salir proporcional a a, un área y una superficie a a^2, un volumen a a^3.

Ejemplo 10.42 (Siete medidas sobre tres curvas)**.** (a) **Calibrando con la circunferencia.** Con $x = R\cos t$, $y = R\sin t$, $t \in [0, 2\pi]$, se tiene $x' = -R\sin t$, $y' = R\cos t$, luego
$$x\,y' - y\,x' = R^2\cos^2 t + R^2\sin^2 t = R^2, \qquad A = \frac{1}{2}\int_0^{2\pi} R^2\,\mathrm{d}t = \pi R^2.$$
El radicando de la longitud es $R^2\sin^2 t + R^2\cos^2 t = R^2$, así que $L = \int_0^{2\pi} R\,\mathrm{d}t = 2\pi R$. Y girando la semicircunferencia superior ($t \in [0, \pi]$) alrededor de OX, $S = 2\pi\int_0^{\pi} R\sin t \cdot R\,\mathrm{d}t = 2\pi R^2\big[-\cos t\big]_0^{\pi} = 4\pi R^2$: la superficie de la esfera. Las tres fórmulas quedan calibradas.

(b) **La elipse, donde Green enseña los dientes.** Con $x = a\cos t$, $y = b\sin t$,

$$x\,y' - y\,x' = a\cos t\cdot b\cos t - b\sin t\cdot(-a\sin t) = ab(\cos^2 t + \sin^2 t) = ab,$$

constante otra vez, y por tanto $A = \frac{1}{2}\,ab\cdot 2\pi = \pi ab$. Obsérvese lo que ha pasado: la integral que en cartesianas obliga a $4\int_0^a b\sqrt{1-x^2/a^2}\,\mathrm{d}x$ con cambio trigonométrico, aquí es la integral de una constante. La longitud, en cambio, no se simplifica: el radicando $a^2\sin^2 t + b^2\cos^2 t$ no es un cuadrado perfecto salvo si $a = b$, y esa integral elíptica es precisamente la que dio nombre a toda una familia de funciones.

(c) **Área de la astroide.** La astroide $x = a\cos^3 t$, $y = a\sin^3 t$, $t \in [0, 2\pi]$, es la estrella de cuatro puntas de ecuación cartesiana $x^{2/3} + y^{2/3} = a^{2/3}$. Derivando, $x' = -3a\cos^2 t\sin t$ y $y' = 3a\sin^2 t\cos t$, de modo que

$$x\,y' - y\,x' = 3a^2\cos^4 t\sin^2 t + 3a^2\sin^4 t\cos^2 t = 3a^2\sin^2 t\cos^2 t.$$

Usando $\sin^2 t\cos^2 t = \frac{1}{4}\sin^2(2t)$ y $\int_0^{2\pi}\sin^2(2t)\,\mathrm{d}t = \pi$,

$$A = \frac{1}{2}\int_0^{2\pi} 3a^2\sin^2 t\cos^2 t\,\mathrm{d}t = \frac{3a^2}{8}\int_0^{2\pi}\sin^2(2t)\,\mathrm{d}t = \frac{3\pi a^2}{8} \approx 1{,}1781\,a^2.$$

Menos de la mitad del círculo de radio a (πa^2), como corresponde a una figura que se hunde hacia el centro.

(d) **Longitud de la astroide.** El radicando se factoriza limpiamente:

$$x'^2 + y'^2 = 9a^2\cos^4 t\sin^2 t + 9a^2\sin^4 t\cos^2 t = 9a^2\sin^2 t\cos^2 t, \qquad \sqrt{x'^2+y'^2} = 3a\,|\sin t\cos t|.$$

El valor absoluto es imprescindible: sin él la integral daría cero. Por simetría basta el primer cuadrante:

$$L = 4\int_0^{\pi/2} 3a\sin t\cos t\,\mathrm{d}t = 12a\Big[\frac{\sin^2 t}{2}\Big]_0^{\pi/2} = 6a.$$

Un resultado sorprendentemente redondo: el perímetro de la astroide es exactamente seis veces a, sin rastro de π.

(e) **Superficie que genera la astroide al girar.** Girando la mitad superior ($t \in [0, \pi]$) alrededor de OX,

$$S = 2\pi\int_0^{\pi} a\sin^3 t\cdot 3a|\sin t\cos t|\,\mathrm{d}t = 6\pi a^2\cdot 2\int_0^{\pi/2}\sin^4 t\cos t\,\mathrm{d}t = 12\pi a^2\Big[\frac{\sin^5 t}{5}\Big]_0^{\pi/2} = \frac{12\pi a^2}{5}.$$

De nuevo el valor absoluto ha obligado a partir en $t = \pi/2$ y duplicar. El resultado, $2{,}4\pi a^2$, es menor que los $4\pi a^2$ de la esfera de radio a: la superficie generada está *abollada* hacia dentro.

(f) **La cicloide: área y longitud de un arco.** Con $x = R(t - \sin t)$, $y = R(1 - \cos t)$ y $t \in [0, 2\pi]$ se tiene $x' = R(1 - \cos t) \geq 0$, así que el arco avanza siempre hacia la derecha y vale la fórmula del área bajo el arco:

$$A = \int_0^{2\pi} y\,x'\,\mathrm{d}t = R^2\int_0^{2\pi}(1-\cos t)^2\,\mathrm{d}t = R^2\int_0^{2\pi}(1-2\cos t+\cos^2 t)\mathrm{d}t = R^2(2\pi+\pi) = 3\pi R^2.$$

Exactamente el triple del círculo que rueda: uno de los resultados favoritos del siglo XVII. Para la longitud, el radicando se simplifica con la identidad del ángulo mitad:

$$x'^2 + y'^2 = R^2[(1-\cos t)^2 + \sin^2 t] = R^2(2 - 2\cos t) = 4R^2 \sin^2(t/2),$$

luego $\sqrt{x'^2 + y'^2} = 2R\sin(t/2)$ (positivo en todo $[0, 2\pi]$) y

$$L = \int_0^{2\pi} 2R\sin(t/2)\,\mathrm{d}t = 2R\big[-2\cos(t/2)\big]_0^{2\pi} = 8R.$$

Ocho radios, sin π: la longitud de un arco de cicloide es cuatro veces el diámetro de la rueda.

(g) **La cicloide girando: superficie y volumen.** Con el mismo $\mathrm{d}s = 2R\sin(t/2)\,\mathrm{d}t$ y $y = R(1-\cos t) = 2R\sin^2(t/2)$, la superficie que engendra el arco al girar alrededor de OX es

$$S = 2\pi \int_0^{2\pi} 2R\sin^2(t/2)\cdot 2R\sin(t/2)\,\mathrm{d}t = 8\pi R^2 \int_0^{2\pi} \sin^3(t/2)\,\mathrm{d}t = 16\pi R^2 \int_0^{\pi} \sin^3 u\,\mathrm{d}u = \frac{64\pi R^2}{3},$$

donde se ha usado $u = t/2$ y $\int_0^{\pi} \sin^3 u\,\mathrm{d}u = 4/3$. El volumen, por discos,

$$V = \pi \int_0^{2\pi} y^2\,x'\,\mathrm{d}t = \pi R^3 \int_0^{2\pi} (1-\cos t)^3\,\mathrm{d}t = \pi R^3(2\pi + 3\pi) = 5\pi^2 R^3,$$

porque $\int_0^{2\pi}(1 - 3\cos t + 3\cos^2 t - \cos^3 t)\,\mathrm{d}t = 2\pi + 3\pi = 5\pi$ (los términos con $\cos t$ y $\cos^3 t$ tienen integral nula sobre un periodo completo). Repara en la forma del resultado: un π^2, no un π, señal inconfundible de que se han integrado dos oscilaciones encadenadas.

Cuidado

En la longitud de arco, $\sqrt{x'^2 + y'^2}$ es una raíz cuadrada y por tanto *siempre* es positiva; escribir $3a\sin t\cos t$ en lugar de $3a|\sin t\cos t|$ para la astroide es el error más frecuente de toda esta sección, y no da un resultado *algo desviado*: da $L = 0$, porque el seno por coseno cambia de signo cuatro veces en $[0, 2\pi]$ y las contribuciones se cancelan. Regla práctica: si una longitud te sale cero o negativa, has perdido un valor absoluto. La misma precaución vale en la fórmula de la superficie, donde aparece $|y|$ y no y.

Pausa para pensar

¿Por qué la fórmula del área $A = \frac{1}{2}\oint (x\,y' - y\,x')\,\mathrm{d}t$ funciona también para curvas tan retorcidas como la astroide? Piensa en el integrando: $\frac{1}{2}(x\,y' - y\,x')\,\mathrm{d}t$ es el área del triángulo que forman el origen, el punto (x, y) y el punto vecino $(x + \mathrm{d}x,\, y + \mathrm{d}y)$; es el determinante $\frac{1}{2}\left|\begin{smallmatrix} x & \mathrm{d}x \\ y & \mathrm{d}y \end{smallmatrix}\right|$ que ya viste en 6. Al recorrer la curva entera, esos triángulos barren el interior: los tramos en que el radio vector gira hacia adelante suman y los que giran hacia atrás restan, y lo que queda es exactamente el área encerrada. Por eso el signo depende del sentido del recorrido, y por eso la fórmula falla si la curva se corta a sí misma: un trozo se contaría dos veces.

Idea clave

Las cuatro fórmulas de esta sección son casos particulares de dos ideas mucho más generales. La del área es el teorema de Green del plano ($\oint_{\partial D} P\,\mathrm{d}x + Q\,\mathrm{d}y = \iint_D (\partial_x Q - \partial_y P)\,\mathrm{d}x\,\mathrm{d}y$ con $P = -y/2$, $Q = x/2$), que verás con todo detalle en 11; las de longitud y superficie son integrales de línea del elemento de arco $\mathrm{d}s$, y las de volumen, integrales dobles disfrazadas de simples, tal como se estudian en 9. Merece la pena quedarse con esta lectura: la parametrización no es un truco para dibujar curvas raras, es la manera natural de integrar *sobre* un objeto geométrico en lugar de *bajo* una gráfica. Cuando el objeto deja de ser una curva plana y pasa a ser una superficie en el espacio, cambian los nombres pero no la estructura.

10.13. Polares en el espacio: coordenadas cilíndricas y esféricas

Las coordenadas polares resuelven en el plano el problema de describir situaciones con simetría de giro. En el espacio hay dos maneras naturales de extenderlas, según si la simetría es alrededor de un eje o alrededor de un punto, y ambas son de uso constante en física, en cartografía y en integración múltiple.

Definición 10.43. Las *coordenadas cilíndricas* (r, θ, z) se relacionan con las cartesianas por

$$x = r\cos\theta, \qquad y = r\sin\theta, \qquad z = z,$$

con $r \geq 0$ y $\theta \in [0, 2\pi)$: son las polares del plano XY más la altura sin tocar. Las *coordenadas esféricas* (ρ, θ, φ), con $\rho \geq 0$, $\theta \in [0, 2\pi)$ el *acimut* y $\varphi \in [0, \pi]$ la *colatitud*, se relacionan por

$$x = \rho\sin\varphi\cos\theta, \qquad y = \rho\sin\varphi\sin\theta, \qquad z = \rho\cos\varphi,$$

de donde $\rho = \sqrt{x^2 + y^2 + z^2}$ y $\varphi = \arccos(z/\rho)$. La curva $\vec{c}(t) = (a\cos t,\, a\sin t,\, bt)$ es una *hélice* circular.

Idea clave

La elección entre unas coordenadas y otras no es cuestión de gusto sino de qué superficies se vuelven triviales. En cartesianas, las superficies "una coordenada constante" son tres familias de planos. En cilíndricas son cilindros (r constante), semiplanos (θ constante) y planos horizontales (z constante). En esféricas son esferas (ρ constante), semiplanos (θ constante) y conos (φ constante). Por eso un problema con simetría axial (un cable, un tubo, un remolino) se escribe en dos líneas en cilíndricas, y uno con simetría radial (un átomo, un planeta, una explosión) se escribe en dos líneas en esféricas. Elegir mal las coordenadas no impide resolver el problema, solo lo hace diez veces más largo.

Regla: trabajar con coordenadas cilíndricas y esféricas

Para pasar de unas a otras y no perderse:

1. **Identifica la simetría del problema** antes de nada: eje, punto o ninguna de las

dos.

2. **Para ir a cilíndricas, aplica las polares al par** (x, y) y deja z intacta; $r = \sqrt{x^2 + y^2}$ y $\theta = \operatorname{atan2}(y, x)$.
3. **Para ir a esféricas, calcula primero** ρ, luego φ con $\cos\varphi = z/\rho$, y por último θ con el mismo atan2 que en el plano.
4. **Cuida el orden y el nombre de los ángulos**: hay dos convenios opuestos en la literatura, y conviene declarar cuál se usa antes de escribir una sola fórmula.
5. **Escribe las superficies coordenadas** para verificar que has entendido el sistema: cilindro, semiplano y plano; esfera, semiplano y cono.
6. **Recuerda los elementos de volumen**: $\mathrm{d}V = r\,\mathrm{d}r\,\mathrm{d}\theta\,\mathrm{d}z$ en cilíndricas y $\mathrm{d}V = \rho^2 \sin\varphi\,\mathrm{d}\rho\,\mathrm{d}\theta\,\mathrm{d}\varphi$ en esféricas, como se deduce en el capítulo 9.
7. **Comprueba el resultado volviendo a cartesianas**; una conversión de ida y vuelta detecta cualquier error de signo o de convenio.

Ejemplo 10.44 (Convertir puntos y reconocer superficies)**. (a) De cilíndricas a cartesianas y vuelta.** El punto $(2, \frac{\pi}{3}, 5)$ en cilíndricas es $(2\cos 60°, 2\sin 60°, 5) = (1;\ 1{,}732051;\ 5)$. Al revés, el punto cartesiano $(3, 4, 5)$ tiene $r = \sqrt{9+16} = 5$, $\theta = \arctan\frac{4}{3} = 0{,}927295$ rad $= 53{,}1301°$ y $z = 5$. Nótese que r no es la distancia al origen sino al *eje*: la distancia al origen sería $\sqrt{50}$.

(b) De esféricas a cartesianas. El punto cartesiano $(1, 1, \sqrt{2})$ tiene $\rho = \sqrt{1+1+2} = 2$, $\varphi = \arccos\frac{\sqrt{2}}{2} = 45°$ y $\theta = 45°$. La conversión de vuelta, $(2\sin 45°\cos 45°,\ 2\sin 45°\sin 45°,\ 2\cos 45°) = (1, 1, 1{,}414214)$, devuelve el punto de partida. Para $(0, 0, 7)$ se tiene $\rho = 7$, $\varphi = 0$ y θ indeterminado: en el polo, el acimut no significa nada, igual que en el origen de las polares del plano.

(c) Superficies escritas en una línea. La esfera de radio R centrada en el origen es $\rho = R$; el cono de semiángulo $30°$ es $\varphi = \frac{\pi}{6}$; el cilindro de radio 2 alrededor del eje OZ es $r = 2$; el plano XZ es $\theta = 0$ junto con $\theta = \pi$. Compárese con sus versiones cartesianas, $x^2 + y^2 + z^2 = R^2$ y $3(x^2 + y^2) = z^2$: la ventaja es evidente.

(d) Coordenadas geográficas. La latitud λ y la colatitud se relacionan por $\varphi = 90° - \lambda$, y la longitud es el acimut. Ciudad de Guatemala está a $\lambda = 14{,}6349°$ norte y $-90{,}5069°$ de longitud; con $R = 6371$ km eso da las coordenadas cartesianas $(-54{,}5353;\ -6164{,}0545;\ 1609{,}6890)$ km, cuya norma es exactamente 6371. La distancia sobre la superficie hasta Madrid ($40{,}4168°$ norte, $-3{,}7038°$) se obtiene del ángulo entre los dos vectores unitarios, que vale $78{,}1769°$, y resulta ser $6371 \cdot 1{,}364444 = 8692{,}8758$ km.

Ejemplo 10.45 (La hélice, la curva que sube girando)**. (a) La parametrización más simple.** Sea $\vec{c}(t) = (\cos t, \sin t, t)$. En cilíndricas es sencillamente $r = 1$, $\theta = t$, $z = t$, es decir la recta $z = \theta$ dibujada sobre el cilindro unidad. La velocidad es $(-\sin t, \cos t, 1)$, de módulo constante $\sqrt{2}$, así que la curva se recorre a rapidez uniforme.

(b) Paso y longitud. Cada vuelta completa (t de 0 a 2π) eleva la curva $2\pi = 6{,}283185$ unidades: eso es el *paso* de la hélice. La longitud recorrida en esa vuelta es $\sqrt{2} \cdot 2\pi = 8{,}885766$,

mayor que la circunferencia 2π y que el paso, como debe ser: es la hipotenusa del triángulo que se forma al desenrollar el cilindro.

(c) Curvatura y torsión. Para la hélice general $\vec{c}(t) = (a\cos t,\ a\sin t,\ bt)$ se obtiene

$$\kappa = \frac{a}{a^2+b^2}, \qquad \tau = \frac{b}{a^2+b^2},$$

ambas constantes: la hélice es la única curva del espacio con curvatura y torsión constantes no nulas. Con $a = b = 1$ vale $\kappa = \tau = 0{,}5$. Con $a = 3$, $b = 4$ la rapidez es 5, la curvatura 0,12 y la torsión 0,16, y una vuelta mide $10\pi = 31{,}415927$.

(d) Los dos casos límite. Si $b \to 0$ la hélice se aplasta hasta la circunferencia de radio a, y en efecto $\kappa \to \frac{1}{a}$ y $\tau \to 0$: una curva plana tiene torsión nula. Si en cambio $a \to 0$ queda la recta vertical, con $\kappa \to 0$. La torsión mide precisamente cuánto se sale la curva de su plano instantáneo, y es lo único que distingue una hélice de una circunferencia vista desde dentro de la propia curva. Por eso las escaleras de caracol, los muelles y el ADN comparten forma: es la única manera de avanzar y girar a ritmo constante a la vez.

Cuidado

En cilíndricas, r es la distancia al eje OZ, no al origen; confundirla con ρ es el error más frecuente al pasar entre los dos sistemas. Segundo aviso: existen dos convenios para las esféricas, uno que llama φ a la colatitud (el que usamos aquí, habitual en física) y otro que la llama θ (habitual en algunos textos de matemáticas); antes de copiar una fórmula de otra fuente hay que verificar cuál emplea. Y tercero: el elemento de volumen no es $\mathrm{d}\rho\,\mathrm{d}\theta\,\mathrm{d}\varphi$; lleva el factor $\rho^2\sin\varphi$, y olvidarlo invalida cualquier integral triple.

Pausa para pensar

Las coordenadas esféricas son mucho más viejas que el sistema cartesiano que pretenden complementar: los astrónomos las usaban en forma de latitud y longitud celestes desde Hiparco, en el siglo II antes de nuestra era. Lo que Descartes aportó en 1637 no fue la idea de localizar con números, sino la de *tratar esos números algebraicamente*. Y hay una ironía final: la esfera es la única superficie cerrada que no admite un sistema de coordenadas válido en todos sus puntos a la vez, hecho topológico conocido como teorema de la bola peluda. Los polos, donde el acimut se vuelve indeterminado, no son un defecto de nuestra elección: son inevitables, y la razón se explica en el capítulo 28.

10.14. Errores típicos

1. Convertir de cartesianas a polares con $\arctan(y/x)$ sin mirar el cuadrante. La calculadora siempre devuelve un ángulo entre $-\pi/2$ y $\pi/2$.
2. Olvidar que $r = \sqrt{x^2+y^2}$ es no negativo por definición cuando se acaba de escribir $r = -3$.
3. Creer que un punto tiene unas únicas coordenadas polares.

4. Buscar los cortes de dos curvas polares resolviendo solo el sistema de ecuaciones: el origen puede ser común a las dos con ángulos distintos y no aparecer.
5. Contar los pétalos de $r = \cos(k\theta)$ sin distinguir k par de k impar.
6. Trabajar en grados cuando la fórmula está en radianes (o al revés) al tabular una curva polar.
7. Eliminar el parámetro y quedarse tan tranquilo, sin comprobar qué trozo de la curva cartesiana se recorre realmente.
8. Perder la orientación: $x = \cos t, y = \sin t$ y $x = \cos t, y = -\sin t$ dan la misma circunferencia recorrida en sentidos opuestos.
9. Pensar que t es siempre el tiempo. Es una variable auxiliar; a veces es un ángulo, a veces una longitud de arco.
10. En el tiro parabólico, usar $\sin\alpha$ donde va $\sin 2\alpha$ en el alcance.
11. Poner g con signo positivo en $y(t)$ en lugar de restar el término $\frac{1}{2}gt^2$.
12. Creer que la cicloide es un arco de circunferencia. No lo es, y ninguna cónica se le parece.
13. Confundir la espiral de Arquímedes ($r = \theta$, vueltas equiespaciadas) con la logarítmica ($r = \mathrm{e}^{a\theta}$, vueltas que se separan cada vez más).

10.15. Rectas tangentes a una curva polar y el ángulo ψ

Para curvas paramétricas hemos deducido $\dfrac{\mathrm{d}y}{\mathrm{d}x} = \dfrac{\dot y}{\dot x}$ y hasta la derivada segunda. Falta el análogo polar, que hasta ahora hemos usado sin justificar cada vez que decíamos "ahí la tangente es la recta $\theta = \theta_0$".

Teorema 10.46 (Pendiente de una curva polar)**.** *Si $r = r(\theta)$, entonces $x = r\cos\theta$, $y = r\sin\theta$ y*

$$\frac{\mathrm{d}y}{\mathrm{d}x} = \frac{r'\sin\theta + r\cos\theta}{r'\cos\theta - r\sin\theta}.$$

Demostración. Derivando cada coordenada respecto de θ y usando la regla de la cadena: $\dot y = r'\sin\theta + r\cos\theta$ y $\dot x = r'\cos\theta - r\sin\theta$. □

Observación 10.47 (La tangente en el polo). Si $r(\theta_0) = 0$ y $r'(\theta_0) \neq 0$, la fórmula se reduce a

$$\frac{\mathrm{d}y}{\mathrm{d}x} = \frac{r'\sin\theta_0}{r'\cos\theta_0} = \tan\theta_0,$$

es decir, la tangente en el origen es exactamente la recta $\theta = \theta_0$. Ahora sí está justificado: los ángulos en que r se anula *son* las direcciones de las tangentes en el polo, y por eso la lemniscata $r^2 = \cos 2\theta$, que se anula en $\theta = \pm\pi/4$, se corta a sí misma en el origen formando un ángulo recto.

Ejemplo 10.48 (Las tangentes de la cardioide)**.** Sea $r = 1 + \cos\theta$.

(a) Horizontales. $\dot{y} = r'\sin\theta + r\cos\theta = -\sin^2\theta + (1+\cos\theta)\cos\theta$. Usando $\sin^2 = 1-\cos^2$ queda $2\cos^2\theta + \cos\theta - 1 = (2\cos\theta - 1)(\cos\theta + 1)$, que se anula en

$$\theta = \pm\frac{\pi}{3} \qquad \text{y} \qquad \theta = \pi.$$

Los dos primeros son tangentes horizontales de verdad; en $\theta = \pi$ también se anula $\dot{x}$, y ese es el punto anguloso de la cúspide.

(b) Verticales. $\dot{x} = r'\cos\theta - r\sin\theta = -\sin\theta(1 + 2\cos\theta)$, que se anula en

$$\theta = 0 \qquad \text{y} \qquad \theta = \pm\frac{2\pi}{3}.$$

Definición 10.49 (El ángulo ψ)**.** Sea ψ el ángulo entre el radio vector y la tangente a la curva. Entonces

$$\tan\psi = \frac{r}{r'}.$$

Ejemplo 10.50 (Por qué la espiral logarítmica se llama equiangular)**.** Para $r = e^{a\theta}$ se tiene $r' = a\,e^{a\theta} = ar$, luego

$$\tan\psi = \frac{r}{ar} = \frac{1}{a},$$

constante, independientemente del punto. La espiral logarítmica corta a todos sus radios vectores con el mismo ángulo, y de ahí su otro nombre: espiral equiangular. Es la razón de que la use un halcón que se acerca a su presa manteniendo un ángulo fijo de visión, y de que aparezca en la concha del nautilo, donde cada tramo nuevo es una copia a escala del anterior. La propiedad de autosemejanza y la de ángulo constante son la misma propiedad escrita dos veces.

Contrástese con la espiral de Arquímedes $r = a\theta$, donde $\tan\psi = \theta$ crece sin parar: el ángulo se abre hacia el exterior y la espiral tiende a parecerse a una circunferencia.

10.16. El triedro de Frenet, la torsión y la segunda ley de Kepler

Definición 10.51 (El triedro móvil)**.** Para una curva regular parametrizada por el arco s,

$$\vec{T} = \frac{\mathrm{d}\vec{r}}{\mathrm{d}s}, \qquad \vec{N} = \frac{1}{\kappa}\frac{\mathrm{d}\vec{T}}{\mathrm{d}s}, \qquad \vec{B} = \vec{T}\times\vec{N},$$

y las tres derivadas se cierran entre sí en las **fórmulas de Frenet–Serret**:

$$\vec{T}' = \kappa\vec{N}, \qquad \vec{N}' = -\kappa\vec{T} + \tau\vec{B}, \qquad \vec{B}' = -\tau\vec{N}.$$

La curvatura κ mide cuánto se dobla la curva dentro de su plano osculador; la **torsión** τ mide cuánto se sale de él. Una curva plana tiene $\tau \equiv 0$, y recíprocamente.

Teorema 10.52 (Fórmulas operativas)**.** *Sin necesidad de reparametrizar por el arco,*

$$\kappa = \frac{\|\vec{r}' \times \vec{r}''\|}{\|\vec{r}'\|^3}, \qquad \tau = \frac{(\vec{r}' \times \vec{r}'') \cdot \vec{r}'''}{\|\vec{r}' \times \vec{r}''\|^2}.$$

Y el ***teorema fundamental de las curvas*** *dice que $\kappa(s) > 0$ y $\tau(s)$ determinan la curva salvo movimientos rígidos: dos curvas con la misma curvatura y la misma torsión son superponibles.*

Ejemplo 10.53 (La hélice y la cúbica alabeada)**. (a) Hélice** $\vec{r} = (a\cos t,\ a\sin t,\ bt)$**.** Aplicando las fórmulas,

$$\kappa = \frac{a}{a^2+b^2}, \qquad \tau = \frac{b}{a^2+b^2},$$

ambas constantes. Con $a = b = 1$ valen las dos $\frac{1}{2}$. Por el teorema fundamental, la hélice es *la única* curva del espacio con curvatura y torsión constantes no nulas; con $\tau = 0$ se obtiene la circunferencia y con $\kappa = 0$ la recta.

(b) Cúbica alabeada $\vec{r} = (t, t^2, t^3)$**.** Ahora

$$\kappa(t) = \frac{2\sqrt{9t^4+9t^2+1}}{(9t^4+4t^2+1)^{3/2}}, \qquad \tau(t) = \frac{3}{9t^4+9t^2+1},$$

y ninguna es constante: en $t = 0$ valen $\kappa = 2$ y $\tau = 3$, y en $t = 1$ valen 0,166 424 y 0,157 895. La curva se aplana y se destuerce a la vez al alejarse del origen.

Teorema 10.54 (Descomposición de la aceleración)**.**

$$\vec{a} = \frac{\mathrm{d}v}{\mathrm{d}t}\,\vec{T} + \kappa v^2\,\vec{N}.$$

Es el resultado que da sentido físico a todo lo anterior: la aceleración no tiene componente en $\vec{B}$ nunca, la tangencial cambia la *rapidez* y la normal cambia la *dirección*. Como $\kappa = 1/R$ con R el radio de curvatura, la componente normal es v^2/R: la aceleración centrípeta de siempre, ahora deducida y no postulada.

Ejemplo 10.55 (La parábola $\vec{r}(t) = (t, t^2)$ en $t = 1$)**.** $\vec{r}' = (1, 2t)$ y $\vec{r}'' = (0, 2)$, luego en $t = 1$:

$$v = \sqrt{5} = 2{,}236\,068, \qquad a_T = \frac{\mathrm{d}v}{\mathrm{d}t} = \frac{4t}{\sqrt{1+4t^2}} = \frac{4}{\sqrt{5}} = 1{,}788\,854,$$

$$a_N = \kappa v^2 = \frac{2}{5^{3/2}} \cdot 5 = \frac{2}{\sqrt{5}} = 0{,}894\,427.$$

Comprobación: $a_T^2 + a_N^2 = \frac{16}{5} + \frac{4}{5} = 4 = \|\vec{a}\|^2$, y en efecto $\vec{a} = (0, 2)$ tiene módulo 2. ✓

La segunda ley de Kepler, deducida

Teorema 10.56 (Velocidad areolar)**.** *Si la fuerza es* central *—es decir, $\vec{a}$ es paralelo a $\vec{r}$— entonces $\vec{r} \times \vec{v}$ es constante, y en coordenadas polares*

$$r^2\dot{\theta} = h \quad \textit{(constante)}, \qquad \frac{\mathrm{d}A}{\mathrm{d}t} = \tfrac{1}{2}r^2\dot{\theta} = \tfrac{h}{2}.$$

Demostración. $\dfrac{\mathrm{d}}{\mathrm{d}t}(\vec{r}\times\vec{v}) = \vec{v}\times\vec{v} + \vec{r}\times\vec{a} = \vec{0}$, porque el primer término es nulo y el segundo lo es por ser $\vec{a} \parallel \vec{r}$. El módulo de $\vec{r}\times\vec{v}$ en polares es exactamente $r^2\dot\theta$. Y el área barrida en polares es $A = \frac{1}{2}\int r^2 \mathrm{d}\theta$, cuya derivada temporal es $\frac{1}{2}r^2\dot\theta$. □

Ahí está la segunda ley de Kepler —el radio vector barre áreas iguales en tiempos iguales— convertida en un corolario de dos líneas. Y obsérvese lo que *no* se ha usado: la ley del inverso del cuadrado. La segunda ley vale para *cualquier* fuerza central; solo la primera (órbitas cónicas) y la tercera necesitan el $1/r^2$.

Ejemplo 10.57 (La Tierra, con números)**.** La órbita terrestre tiene excentricidad $e = 0{,}0167$.

(a) Velocidades. Como $r^2\dot\theta$ es constante y $r_{\text{afelio}}/r_{\text{perihelio}} = (1+e)/(1-e) = 1{,}033\,967$, la velocidad angular en el perihelio es un 3,4 por ciento mayor. La Tierra corre más en enero que en julio.

(b) Reparto del año. Como consecuencia, las dos mitades de la órbita separadas por el eje menor no tardan lo mismo. A primer orden en e, la diferencia es

$$\frac{2eT}{\pi} = \frac{2\cdot 0{,}0167\cdot 365{,}256}{\pi} = 3{,}88 \text{ días}.$$

El invierno del hemisferio norte es casi cuatro días más corto que el verano, y la razón no es astrológica ni climática: es que en enero la Tierra está en el perihelio y va más deprisa.

10.17. Cálculo de variaciones: la braquistócrona y la tautócrona, demostradas

La sección de la cicloide se titula “sus dos milagros” y no demuestra ninguno. Los dos son consecuencia de una misma herramienta —la ecuación de Euler–Lagrange—, que además es la puerta de entrada a la mecánica analítica y al método directo del cálculo de variaciones.

Teorema 10.58 (Ecuación de Euler–Lagrange)**.** *Sea $F(x,y,y')$ de clase C^2 y considérese el funcional*

$$J[y] = \int_a^b F(x, y(x), y'(x))\,\mathrm{d}x$$

sobre las funciones y de clase C^2 con extremos fijos $y(a)=y_a$, $y(b)=y_b$. Si y es un extremal de J, entonces

$$\frac{\partial F}{\partial y} - \frac{\mathrm{d}}{\mathrm{d}x}\left(\frac{\partial F}{\partial y'}\right) = 0.$$

Demostración. Sea η cualquier función C^2 con $\eta(a) = \eta(b) = 0$, y considérese la familia $y_\varepsilon = y + \varepsilon\eta$, que respeta las condiciones de contorno. La función $\varphi(\varepsilon) = J[y_\varepsilon]$ tiene un extremo en $\varepsilon = 0$, luego $\varphi'(0) = 0$. Derivando bajo el signo integral,

$$0 = \varphi'(0) = \int_a^b \left(\frac{\partial F}{\partial y}\,\eta + \frac{\partial F}{\partial y'}\,\eta'\right)\mathrm{d}x.$$

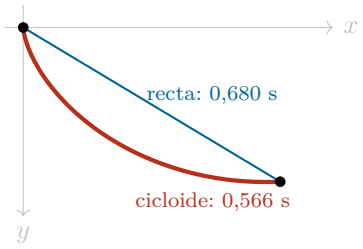


Figura 10.2: Descenso sin rozamiento desde el origen hasta el punto (1; 0,6) en metros. La recta es el camino más corto, pero no el más rápido: la cicloide baja antes, gana velocidad pronto y llega en 0,566 segundos frente a los 0,680 de la recta, un 17 % menos. Obsérvese que la solución óptima llega incluso a pasar por debajo del punto de destino cuando éste está lo bastante lejos.

Intégrese por partes el segundo sumando:

$$\int_a^b \frac{\partial F}{\partial y'}\eta' \,\mathrm{d}x = \left[\frac{\partial F}{\partial y'}\eta\right]_a^b - \int_a^b \frac{\mathrm{d}}{\mathrm{d}x}\left(\frac{\partial F}{\partial y'}\right)\eta \,\mathrm{d}x,$$

y el término de frontera se anula porque $\eta(a) = \eta(b) = 0$. Así

$$\int_a^b \left[\frac{\partial F}{\partial y} - \frac{\mathrm{d}}{\mathrm{d}x}\left(\frac{\partial F}{\partial y'}\right)\right]\eta(x)\,\mathrm{d}x = 0$$

para *toda* η admisible. Por el lema fundamental del cálculo de variaciones —si una función continua g cumple $\int g\eta = 0$ para toda η que se anule en los extremos, entonces $g \equiv 0$; se prueba tomando η concentrada en un entorno de un punto donde $g \neq 0$ y con el mismo signo—, el corchete es idénticamente nulo. □

Corolario 10.59 (Identidad de Beltrami)**.** *Si F no depende explícitamente de x, entonces*

$$F - y'\frac{\partial F}{\partial y'} = \text{constante}$$

a lo largo de los extremales.

Demostración. Derívese el miembro izquierdo respecto de x:

$$\frac{\mathrm{d}}{\mathrm{d}x}\left(F - y'F_{y'}\right) = F_y y' + F_{y'}y'' - y''F_{y'} - y'\frac{\mathrm{d}}{\mathrm{d}x}F_{y'} = y'\left(F_y - \frac{\mathrm{d}}{\mathrm{d}x}F_{y'}\right) = 0$$

por Euler–Lagrange (nótese que el término F_x no aparece por hipótesis). □

Teorema 10.60 (La braquistócrona es una cicloide)**.** *Entre todas las curvas que unen el origen con un punto (x_1, y_1) situado más abajo, la que minimiza el tiempo de descenso de una partícula que parte del reposo y desliza sin rozamiento bajo la gravedad es un arco de* cicloide*, de ecuaciones paramétricas*

$$x = \frac{C}{2}(\theta - \sin\theta), \qquad y = \frac{C}{2}(1 - \cos\theta).$$

Demostración. Tómese el eje y hacia abajo. Por conservación de la energía, la partícula que ha descendido una altura y tiene rapidez $v = \sqrt{2gy}$. El tiempo total es

$$T[y] = \int \frac{\mathrm{d}s}{v} = \int_0^{x_1} \frac{\sqrt{1+y'^2}}{\sqrt{2gy}}\,\mathrm{d}x,$$

de modo que $F(y,y') = \sqrt{\dfrac{1+y'^2}{2gy}}$ no depende de x: se aplica Beltrami. Con $F_{y'} = \dfrac{y'}{\sqrt{2gy}\sqrt{1+y'^2}}$,

$$F - y'F_{y'} = \frac{\sqrt{1+y'^2}}{\sqrt{2gy}} - \frac{y'^2}{\sqrt{2gy}\sqrt{1+y'^2}} = \frac{1}{\sqrt{2gy}\sqrt{1+y'^2}} = \text{cte},$$

es decir

$$y\,(1+y'^2) = C$$

para cierta constante $C > 0$. Esa es la ecuación diferencial de la braquistócrona. Para resolverla, hágase la sustitución $y = \frac{C}{2}(1-\cos\theta) = C\sin^2\frac{\theta}{2}$. Entonces

$$y'^2 = \frac{C}{y} - 1 = \frac{C-y}{y} = \frac{\cos^2(\theta/2)}{\sin^2(\theta/2)}, \qquad \text{luego} \qquad y' = \frac{\cos(\theta/2)}{\sin(\theta/2)}.$$

Por otro lado $\mathrm{d}y = C\sin\frac{\theta}{2}\cos\frac{\theta}{2}\,\mathrm{d}\theta$, y de $\mathrm{d}x = \mathrm{d}y/y'$ resulta

$$\mathrm{d}x = C\sin^2\tfrac{\theta}{2}\,\mathrm{d}\theta = \frac{C}{2}(1-\cos\theta)\,\mathrm{d}\theta,$$

que integrado con $x(0) = 0$ da $x = \frac{C}{2}(\theta - \sin\theta)$. Junto con la expresión de y, son las ecuaciones de la cicloide generada por un círculo de radio $C/2$ que rueda por debajo del eje x. La constante C y el ángulo final se determinan imponiendo que la curva pase por (x_1, y_1). □

Teorema 10.61 (La cicloide es tautócrona)**.** *Sobre un arco de cicloide invertido, el tiempo que tarda una partícula en llegar al punto más bajo partiendo del reposo* no depende *del punto de partida.*

Demostración. Parametrícese la cicloide invertida como $x = R(\theta - \sin\theta)$, $y = -R(1-\cos\theta)$ con $\theta \in [0, 2\pi]$ y el punto más bajo en $\theta = \pi$. La longitud de arco cumple

$$\mathrm{d}s^2 = \mathrm{d}x^2 + \mathrm{d}y^2 = R^2\big[(1-\cos\theta)^2 + \sin^2\theta\big]\mathrm{d}\theta^2 = 2R^2(1-\cos\theta)\,\mathrm{d}\theta^2,$$

luego $\mathrm{d}s = 2R\sin\frac{\theta}{2}\,\mathrm{d}\theta$ y, midiendo s desde el punto más bajo,

$$s = 4R\cos\frac{\theta_0}{2} - 4R\cos\frac{\theta}{2}.$$

La clave está en la altura. Tomando la altura h medida *desde el punto más bajo*, $h = R(1+\cos\theta) = 2R\cos^2\frac{\theta}{2}$, y llamando u a la distancia *a lo largo del arco* desde el punto más bajo,

$$u = \int_\theta^\pi 2R\sin\tfrac{\sigma}{2}\,\mathrm{d}\sigma = \Big[-4R\cos\tfrac{\sigma}{2}\Big]_\theta^\pi = 4R\cos\frac{\theta}{2},$$

se obtiene la relación notable

$$h = 2R\cos^2\frac{\theta}{2} = 2R\left(\frac{u}{4R}\right)^2 = \frac{u^2}{8R}.$$

Es decir: la altura es proporcional al *cuadrado* de la longitud de arco recorrida. Ahora, por conservación de la energía, la rapidez a lo largo del arco cuando se ha caído desde u_0 hasta u es

$$\left(\frac{\mathrm{d}u}{\mathrm{d}t}\right)^2 = 2g\,(h_0 - h) = \frac{g}{4R}(u_0^2 - u^2),$$

que es exactamente la ecuación del *oscilador armónico*: derivando, $\ddot{u} = -\frac{g}{4R}u$. Su periodo es

$$T = 2\pi\sqrt{\frac{4R}{g}} = 4\pi\sqrt{\frac{R}{g}},$$

independiente de la amplitud u_0. El tiempo de caída hasta el punto más bajo es un cuarto de periodo, $\pi\sqrt{R/g}$, y no depende del punto de partida. □

Observación 10.62 (Por qué Huygens no consiguió el reloj perfecto). El resultado es exacto y espectacular: un péndulo obligado a describir una cicloide es rigurosamente isócrono, mientras que el péndulo circular sólo lo es aproximadamente para oscilaciones pequeñas —su periodo real crece con la amplitud, y la corrección de primer orden es un factor $1 + \theta_0^2/16$—. Huygens construyó en 1673 un péndulo cicloidal haciendo que el hilo se enrollara entre dos mejillas cicloidales, aprovechando que la evoluta de una cicloide es otra cicloide igual. En la práctica el rozamiento del hilo contra las mejillas introdujo más error del que corregía, y el invento fracasó; la solución real llegó por otro camino, con el escape de áncora. Es un buen recordatorio de que un teorema exacto no garantiza una máquina mejor.

10.18. Curvatura y torsión: las fórmulas operativas, deducidas

Las fórmulas que permiten calcular κ y τ sin reparametrizar por arco se han usado en todos los ejemplos sin obtenerlas. Salen de derivar tres veces y ordenar.

Teorema 10.63 (Fórmulas sin reparametrizar). *Sea $\vec{r}(t)$ una curva regular de clase C^3 en $\mathbb{R}^3$. Entonces*

$$\kappa = \frac{|\vec{r}\,' \times \vec{r}\,''|}{|\vec{r}\,'|^3}, \qquad \tau = \frac{(\vec{r}\,' \times \vec{r}\,'') \cdot \vec{r}\,'''}{|\vec{r}\,' \times \vec{r}\,''|^2},$$

y la aceleración se descompone como

$$\vec{a} = \dot{v}\,\vec{T} + \kappa v^2\,\vec{N}, \qquad v = |\vec{r}\,'|.$$

Demostración. Escríbase $s(t)$ para la longitud de arco, con $\dot{s} = v = |\vec{r}\,'|$, y $\vec{T} = \vec{r}\,'/v$.

Descomposición de la aceleración. De $\vec{r}\,' = v\vec{T}$, derivando respecto de t y usando la primera fórmula de Frenet $\frac{\mathrm{d}\vec{T}}{\mathrm{d}s} = \kappa\vec{N}$:

$$\vec{a} = \vec{r}\,'' = \dot{v}\,\vec{T} + v\,\frac{\mathrm{d}\vec{T}}{\mathrm{d}t} = \dot{v}\,\vec{T} + v\cdot\frac{\mathrm{d}\vec{T}}{\mathrm{d}s}\cdot\dot{s} = \dot{v}\,\vec{T} + \kappa v^2\,\vec{N}.$$

Esta es la tercera fórmula, y de paso explica los dos efectos que siente un pasajero en un coche: la componente tangencial es la aceleración del acelerador, y la normal —proporcional al *cuadrado* de la velocidad y a la curvatura— es la que empuja hacia fuera en las curvas.

Curvatura. Multiplíquese vectorialmente $\vec{r}\,' = v\vec{T}$ por la expresión anterior:

$$\vec{r}\,' \times \vec{r}\,'' = v\vec{T} \times \left(\dot{v}\vec{T} + \kappa v^2 \vec{N}\right) = \kappa v^3 \left(\vec{T} \times \vec{N}\right) = \kappa v^3 \, \vec{B},$$

porque $\vec{T} \times \vec{T} = 0$. Tomando módulos y usando que $\vec{B}$ es unitario, $|\vec{r}\,' \times \vec{r}\,''| = \kappa v^3$, que es la primera fórmula.

Torsión. Derívese una vez más la descomposición de $\vec{a}$, usando ahora las tres fórmulas de Frenet ($\vec{T}\,' = \kappa v \vec{N}$, $\vec{N}\,' = v(-\kappa\vec{T} + \tau\vec{B})$, $\vec{B}\,' = -\tau v \vec{N}$):

$$\vec{r}\,''' = \left(\ddot{v} - \kappa^2 v^3\right)\vec{T} + \left(3\kappa v \dot{v} + \kappa' v^2\right)\vec{N} + \kappa\tau v^3 \, \vec{B}.$$

Los coeficientes exactos de $\vec{T}$ y $\vec{N}$ no importan: al multiplicar escalarmente por $\vec{r}\,' \times \vec{r}\,'' = \kappa v^3 \vec{B}$, ambos desaparecen por ortogonalidad, y queda

$$\left(\vec{r}\,' \times \vec{r}\,''\right) \cdot \vec{r}\,''' = \kappa v^3 \cdot \kappa\tau v^3 = \tau \left(\kappa v^3\right)^2 = \tau \left|\vec{r}\,' \times \vec{r}\,''\right|^2. \qquad \square$$

Ejemplo 10.64 (La hélice, con las fórmulas)**.** Sea $\vec{r}(t) = (a\cos t,\ a\sin t,\ bt)$. Entonces $\vec{r}\,' = (-a\sin t, a\cos t, b)$, $\vec{r}\,'' = (-a\cos t, -a\sin t, 0)$, $\vec{r}\,''' = (a\sin t, -a\cos t, 0)$, y

$$\vec{r}\,' \times \vec{r}\,'' = (ab\sin t,\ -ab\cos t,\ a^2), \qquad |\vec{r}\,' \times \vec{r}\,''| = a\sqrt{a^2+b^2}, \qquad |\vec{r}\,'| = \sqrt{a^2+b^2}.$$

Luego

$$\kappa = \frac{a\sqrt{a^2+b^2}}{(a^2+b^2)^{3/2}} = \frac{a}{a^2+b^2}, \qquad \tau = \frac{(ab\sin t)(a\sin t) + (-ab\cos t)(-a\cos t)}{a^2(a^2+b^2)} = \frac{a^2 b}{a^2(a^2+b^2)} = \frac{b}{a^2+b^2}.$$

Ambas son *constantes*, y de hecho la hélice es la única curva con curvatura y torsión constantes no nulas: eso es el teorema fundamental de las curvas, según el cual $\kappa(s)$ y $\tau(s)$ determinan la curva salvo movimiento rígido. Con $b = 0$ se recupera $\kappa = 1/a$ y $\tau = 0$: la circunferencia.

Teorema 10.65 (Teorema fundamental de las curvas)**.** *Dadas funciones continuas $\kappa(s) > 0$ y $\tau(s)$ en un intervalo, existe una curva parametrizada por arco con esas curvatura y torsión, y es única salvo un movimiento rígido de $\mathbb{R}^3$.*

Demostración. Existencia. Las fórmulas de Frenet forman un sistema lineal de ecuaciones diferenciales ordinarias de primer orden para la matriz $(\vec{T}, \vec{N}, \vec{B})$:

$$\frac{\mathrm{d}}{\mathrm{d}s}\begin{pmatrix}\vec{T}\\ \vec{N}\\ \vec{B}\end{pmatrix} = \begin{pmatrix}0 & \kappa & 0\\ -\kappa & 0 & \tau\\ 0 & -\tau & 0\end{pmatrix}\begin{pmatrix}\vec{T}\\ \vec{N}\\ \vec{B}\end{pmatrix}.$$

Con coeficientes continuos, el teorema de existencia y unicidad para sistemas lineales (capítulo 12) da una solución única en todo el intervalo para cada dato inicial. Si el dato inicial es una base ortonormal directa, la solución sigue siéndolo para todo s: la matriz del sistema es *antisimétrica*, y por tanto $\frac{\mathrm{d}}{\mathrm{d}s}\langle \vec{e}_i, \vec{e}_j\rangle = 0$ para todos i, j, de modo que los productos escalares

se conservan. Definiendo $\vec{r}(s) = \vec{r}_0 + \int_0^s \vec{T}$, se obtiene una curva parametrizada por arco cuyas curvatura y torsión son las dadas.

Unicidad. Si dos curvas tienen las mismas κ y τ, llévese por un movimiento rígido el triedro de una sobre el de la otra en $s = 0$. Entonces ambos triedros resuelven el mismo problema de valores iniciales, luego coinciden para todo s por unicidad; y al integrar $\vec{T}$ desde el mismo punto, las curvas coinciden. □

10.19. La ecuación focal de las cónicas, deducida

La ecuación $r = \ell/(1 + e\cos\theta)$ se enunció como teorema y se usó en los ejemplos de Halley y de la órbita terrestre sin obtenerla, y la definición foco–directriz de la que sale aparecía enterrada después. Se pone aquí el orden correcto.

Definición 10.66 (Cónica por foco y directriz)**.** Sean un punto F (el *foco*), una recta d que no pasa por él (la *directriz*) y un número $e > 0$ (la *excentricidad*). La cónica asociada es el conjunto de puntos P tales que

$$\frac{\operatorname{dist}(P,F)}{\operatorname{dist}(P,d)} = e.$$

Se llama *elipse* si $e < 1$, *parábola* si $e = 1$ e *hipérbola* si $e > 1$.

Teorema 10.67 (Ecuación polar con foco en el origen)**.** *Colóquese el foco en el origen y la directriz en la recta vertical $x = D$ con $D > 0$. Entonces la cónica tiene ecuación polar*

$$r = \frac{\ell}{1 + e\cos\theta}, \qquad \ell = eD,$$

donde ℓ es el semilado recto*: el valor de r cuando $\theta = \pi/2$.*

Demostración. Sea P un punto de coordenadas polares (r, θ), de modo que su abscisa es $x = r\cos\theta$. Entonces

$$\operatorname{dist}(P,F) = r, \qquad \operatorname{dist}(P,d) = D - r\cos\theta.$$

La condición de la definición dice $r = e\,(D - r\cos\theta)$, es decir

$$r + er\cos\theta = eD \quad\Longrightarrow\quad r = \frac{eD}{1 + e\cos\theta}.$$

Y evaluando en $\theta = \pi/2$ se obtiene $r = eD = \ell$, que justifica el nombre. □

Corolario 10.68 (Relación con los semiejes)**.** *Para $e < 1$, los valores extremos de r son*

$$r_{\text{mín}} = \frac{\ell}{1+e} \quad (\theta = 0), \qquad r_{\text{máx}} = \frac{\ell}{1-e} \quad (\theta = \pi),$$

de donde el semieje mayor es $a = \frac{r_{\text{mín}} + r_{\text{máx}}}{2} = \frac{\ell}{1-e^2}$, el semieje menor $b = a\sqrt{1-e^2}$, y la distancia del centro al foco $c = ae$. En particular $\ell = b^2/a$.

Demostración. La media de $\frac{\ell}{1+e}$ y $\frac{\ell}{1-e}$ es $\frac{\ell}{1-e^2}$. Para b, obsérvese que en $\theta=\pi/2$ el punto de la cónica está a distancia ℓ del foco y a distancia horizontal $c=ae$ del centro; el teorema de Pitágoras con la definición bifocal da $b^2=a^2-c^2=a^2(1-e^2)$, y $\ell=a(1-e^2)=b^2/a$. □

Ejemplo 10.69 (El cometa Halley, con los números correctos)**.** La órbita de Halley tiene semieje mayor $a=17{,}834$ unidades astronómicas y excentricidad $e=0{,}967\,1$. De aquí:

$$r_{\text{mín}}=a(1-e)=17{,}834\cdot 0{,}032\,9=0{,}587 \text{ ua}, \qquad r_{\text{máx}}=a(1+e)=35{,}081 \text{ ua},$$

es decir, el perihelio queda entre Mercurio y Venus, y el afelio más allá de Neptuno. El semilado recto es $\ell=a(1-e^2)=17{,}834\cdot 0{,}064\,7=1{,}154$ ua, y la tercera ley de Kepler da el periodo $T=a^{3/2}=17{,}834^{3/2}=75{,}3$ años. El valor observado es de unos 75 a 76 años, y la discrepancia se debe a las perturbaciones de Júpiter y Saturno: la elipse kepleriana es una primera aproximación excelente, no la última palabra.

10.20. Las fórmulas polares, demostradas: área, longitud, ángulo ψ y curvatura

La sección anterior dedujo la pendiente $\mathrm{d}y/\mathrm{d}x$ de una curva polar y enunció, dentro de una definición, que $\tan\psi=r/r'$. Antes, la regla de calcular áreas y longitudes había dado por buenas las fórmulas $A=\frac12\int r^2\mathrm{d}\theta$ y $L=\int\sqrt{r^2+r'^2}\,\mathrm{d}\theta$ con un argumento de sectores circulares que es una buena intuición pero no una demostración. Y de la curvatura de una curva polar el capítulo no ha dicho una sola palabra, pese a haberla calculado una y otra vez para curvas paramétricas. Esta sección salda las cuatro deudas de golpe, y lo notable es que las cuatro salen del mismo cálculo de tres líneas.

Lema 10.70 (Las tres identidades polares)**.** *Sea $r=f(\theta)$ una función de clase C^2 y considérese la parametrización $x(\theta)=f(\theta)\cos\theta$, $y(\theta)=f(\theta)\sin\theta$, con las derivadas respecto de θ denotadas con punto. Escribiendo $r=f(\theta)$, $r'=f'(\theta)$ y $r''=f''(\theta)$ se cumplen, en todo punto:*

$$\dot x^2+\dot y^2=r^2+r'^2, \qquad x\dot y-y\dot x=r^2, \qquad \dot x\ddot y-\dot y\ddot x=r^2+2r'^2-r\,r''.$$

Demostración. Derivando con la regla del producto,

$$\dot x=r'\cos\theta-r\sin\theta, \qquad \dot y=r'\sin\theta+r\cos\theta.$$

Para la primera identidad, elévense al cuadrado y súmense: los dobles productos son $-2rr'\sin\theta\cos\theta$ y $+2rr'\sin\theta\cos\theta$, que se cancelan, y quedan $r'^2(\cos^2\theta+\sin^2\theta)+r^2(\sin^2\theta+\cos^2\theta)=r'^2+r^2$.

Para la segunda,

$$x\dot y-y\dot x=r\cos\theta(r'\sin\theta+r\cos\theta)-r\sin\theta(r'\cos\theta-r\sin\theta)=r^2\cos^2\theta+r^2\sin^2\theta=r^2,$$

porque los dos términos con $rr'\sin\theta\cos\theta$ vuelven a cancelarse.

Para la tercera hay que derivar otra vez:

$$\ddot x=(r''-r)\cos\theta-2r'\sin\theta, \qquad \ddot y=(r''-r)\sin\theta+2r'\cos\theta.$$

Ahora

$$\dot{x}\ddot{y}-\dot{y}\ddot{x} = (r'\cos\theta - r\sin\theta)\big[(r''-r)\sin\theta + 2r'\cos\theta\big] - (r'\sin\theta + r\cos\theta)\big[(r''-r)\cos\theta - 2r'\sin\theta\big].$$

Al desarrollar, los cuatro términos que llevan $r'(r''-r)\sin\theta\cos\theta$ se cancelan dos a dos, y sobreviven $2r'^2(\cos^2\theta+\sin^2\theta) = 2r'^2$ y $-r(r''-r)(\sin^2\theta+\cos^2\theta) = r^2 - rr''$. Sumando, $\dot{x}\ddot{y}-\dot{y}\ddot{x} = r^2 + 2r'^2 - rr''$. □

Teorema 10.71 (Área, longitud y curvatura en polares)**.** *Sea $r = f(\theta)$ de clase C^2 en $[\alpha,\beta]$. Entonces:*

(i) el área barrida por el radio vector es $A = \dfrac{1}{2}\displaystyle\int_\alpha^\beta r^2\,\mathrm{d}\theta$;

(ii) la longitud de la curva es $L = \displaystyle\int_\alpha^\beta \sqrt{r^2+r'^2}\,\mathrm{d}\theta$;

(iii) la curvatura en el punto de ángulo θ, si $r^2+r'^2 \neq 0$, vale

$$\kappa = \frac{\left|r^2 + 2r'^2 - r\,r''\right|}{\left(r^2+r'^2\right)^{3/2}}.$$

Demostración. Todo es el lema 10.70 sustituido en las fórmulas paramétricas que ya se demostraron en este capítulo. Para (i), la fórmula de Green da $A = \frac{1}{2}\int(x\dot{y}-y\dot{x})\,\mathrm{d}\theta$, y la segunda identidad convierte el integrando en r^2. Para (ii), la longitud de arco es $\int\sqrt{\dot{x}^2+\dot{y}^2}\,\mathrm{d}\theta$, y la primera identidad convierte el radicando en $r^2+r'^2$. Para (iii), la curvatura de una curva plana es $\kappa = |\dot{x}\ddot{y}-\dot{y}\ddot{x}|/(\dot{x}^2+\dot{y}^2)^{3/2}$, y basta meter la tercera identidad en el numerador y la primera en el denominador. □

Observación 10.72. Merece la pena ver qué acaba de ocurrir con el famoso $\frac{1}{2}$ del área. No es un factor de conversión ni una corrección: es literalmente el $\frac{1}{2}$ de la fórmula de Green, y el r^2 es el determinante $x\dot{y}-y\dot{x}$, que en polares resulta ser constante en el sentido de que solo depende de r. El argumento del sector circular que se daba en la regla es correcto como intuición, pero esta demostración tiene la ventaja de no apelar a ningún "elemento infinitesimal": es una identidad algebraica seguida de un teorema ya probado.

Teorema 10.73 (El ángulo entre el radio vector y la tangente)**.** *Sea ψ el ángulo, medido en el sentido positivo, que va del radio vector a la tangente de la curva $r = f(\theta)$ en un punto con $r \neq 0$. Entonces*

$$\tan\psi = \frac{r}{r'},$$

entendiendo $\psi = \pi/2$ cuando $r' = 0$.

Demostración. Sea ϕ el ángulo que forma la tangente con el eje polar, de modo que $\tan\phi = \dot{y}/\dot{x}$. Por construcción $\psi = \phi - \theta$, y aplicando la fórmula de la tangente de una diferencia,

$$\tan\psi = \frac{\tan\phi - \tan\theta}{1+\tan\phi\tan\theta} = \frac{\dot{y}\cos\theta - \dot{x}\sin\theta}{\dot{x}\cos\theta + \dot{y}\sin\theta},$$

donde el segundo paso se obtiene multiplicando numerador y denominador por $\dot{x}\cos\theta$. Ahora se sustituyen $\dot{x} = r'\cos\theta - r\sin\theta$ y $\dot{y} = r'\sin\theta + r\cos\theta$. En el numerador,

$$(r'\sin\theta + r\cos\theta)\cos\theta - (r'\cos\theta - r\sin\theta)\sin\theta = r\cos^2\theta + r\sin^2\theta = r,$$

y en el denominador,

$$(r'\cos\theta - r\sin\theta)\cos\theta + (r'\sin\theta + r\cos\theta)\sin\theta = r'\cos^2\theta + r'\sin^2\theta = r'.$$

Luego $\tan\psi = r/r'$, que es lo que la definición de la sección anterior afirmaba sin justificar. □

Teorema 10.74 (El número de pétalos de una rosa)**.** *Sea k un entero positivo y considérese $r = \cos(k\theta)$. Si k es impar, la curva se recorre entera con $\theta \in [0,\pi]$ y tiene exactamente k pétalos. Si k es par, hace falta $\theta \in [0,2\pi]$ y la curva tiene exactamente $2k$ pétalos.*

Demostración. El punto de coordenadas polares (r,θ) es el mismo que $(-r,\theta+\pi)$. Al sustituir θ por $\theta+\pi$ en la ecuación se obtiene $\cos\big(k(\theta+\pi)\big) = (-1)^k\cos(k\theta)$.

Si k es impar, ese valor es $-\cos(k\theta)$, así que el par $\big(-\cos(k\theta),\theta+\pi\big)$ representa *el mismo punto del plano* que $\big(\cos(k\theta),\theta\big)$: la aplicación que a cada θ le asigna su punto tiene periodo π, y $[0,\pi]$ agota la curva. Si k es par, en cambio, el valor es $+\cos(k\theta)$, y el punto $\big(\cos(k\theta),\theta+\pi\big)$ es el simétrico respecto del origen del anterior: un punto nuevo salvo que $\cos(k\theta) = 0$. Ahí el periodo es 2π.

En cuanto al recuento, la curva pasa por el origen exactamente en los ceros de $\cos(k\theta)$, que son $\theta = (2j+1)\pi/(2k)$ y están igualmente espaciados a distancia π/k. En $[0,2\pi)$ hay $2k$ de ellos, luego $2k$ arcos que salen del origen y vuelven a él: $2k$ pétalos contados con multiplicidad. Si k es par, todos son distintos y hay $2k$. Si k es impar, el argumento anterior empareja el arco que empieza en θ_0 con el que empieza en $\theta_0+\pi$, de modo que cada pétalo se dibuja dos veces y quedan k. □

Ejemplo 10.75 (Cuatro curvaturas polares, con la calibración incluida)**.** (a) **La circunferencia $r = R$, que es la calibración obligatoria.** Aquí $r' = r'' = 0$, de modo que $\kappa = R^2/(R^2)^{3/2} = 1/R$, constante, como tiene que ser. Y el radio de curvatura es R: la circunferencia osculatriz de una circunferencia es ella misma.

(b) **La circunferencia disfrazada $r = 2a\cos\theta$.** Ahora $r' = -2a\sin\theta$ y $r'' = -2a\cos\theta$, luego

$$r^2 + 2r'^2 - rr'' = 4a^2\cos^2\theta + 8a^2\sin^2\theta + 4a^2\cos^2\theta = 8a^2,$$

mientras que $r^2 + r'^2 = 4a^2$. Por tanto $\kappa = 8a^2/(4a^2)^{3/2} = 8a^2/(8|a|^3) = 1/|a|$, constante. Es exactamente lo que debía salir: esa curva es la circunferencia de radio $|a|$, como el capítulo demostró al principio. Que la fórmula devuelva una constante donde la ecuación no lo parece es la mejor prueba de que está bien deducida.

(c) **La cardioide $r = a(1+\cos\theta)$.** Con $r' = -a\sin\theta$ y $r'' = -a\cos\theta$,

$$r^2 + 2r'^2 - rr'' = a^2\big[(1+\cos\theta)^2 + 2\sin^2\theta + (1+\cos\theta)\cos\theta\big] = 3a^2(1+\cos\theta),$$

y $r^2 + r'^2 = a^2(1+\cos\theta)^2 + a^2\sin^2\theta = 2a^2(1+\cos\theta)$. Dividiendo,

$$\kappa = \frac{3a^2(1+\cos\theta)}{\left[2a^2(1+\cos\theta)\right]^{3/2}} = \frac{3}{2\sqrt{2a^3(1+\cos\theta)}} = \frac{3}{2\sqrt{2ar}}.$$

Con $a = 1$: en $\theta = 0$, donde $r = 2$, sale $\kappa = 3/4 = 0{,}75$ y radio de curvatura 1,333333; en $\theta = \pi/2$, donde $r = 1$, sale $\kappa = 3\sqrt{2}/4 = 1{,}060660$ y radio 0,942809. Y cuando $\theta \to \pi$ se tiene $r \to 0$ y $\kappa \to \infty$: la cúspide de la cardioide es un punto de curvatura infinita, cosa coherente con que allí la curva forma una punta y no admite circunferencia osculatriz.

(d) **Las dos espirales.** Para la de Arquímedes $r = \theta$ se tiene $r' = 1$, $r'' = 0$, y

$$\kappa = \frac{\theta^2 + 2}{(\theta^2+1)^{3/2}}.$$

En $\theta = 0$ vale 2 (radio 0,5), en $\theta = \pi$ vale 0,331219 (radio 3,019148) y en $\theta = 2\pi$ vale 0,161060 (radio 6,208878). La espiral se va enderezando, y el radio de curvatura crece casi como θ: en la vuelta n se parece cada vez más a una circunferencia. Para la logarítmica $r = \mathrm{e}^{\theta}$ resulta $r' = r'' = r$ y

$$\kappa = \frac{r^2 + 2r^2 - r^2}{(2r^2)^{3/2}} = \frac{2r^2}{2\sqrt{2}\,r^3} = \frac{1}{\sqrt{2}\,r},$$

es decir $\kappa = \sqrt{2}\,\mathrm{e}^{-\theta}/2$, que en $\theta = 0$ vale 0,707107. Obsérvese la diferencia de fondo: en la espiral logarítmica el radio de curvatura es *proporcional* a r, y por eso la curva es semejante a sí misma en cada vuelta; en la de Arquímedes no lo es.

Y el ángulo ψ, con los mismos ejemplos. Para la cardioide, $\tan\psi = (1+\cos\theta)/(-\sin\theta) = -\cot(\theta/2) = \tan(\theta/2 + \pi/2)$, de donde $\psi = \theta/2 + \pi/2$: el ángulo entre el radio y la tangente crece linealmente con θ, y vale $\pi/2$ en $\theta = 0$ (la tangente es perpendicular al eje en el punto más alejado) y tiende a π cuando $\theta \to \pi$. Para la espiral logarítmica $r = \mathrm{e}^{a\theta}$ se obtiene $\tan\psi = 1/a$, constante, que es la propiedad equiangular que ya se comentó; con $a = 1$ el ángulo vale exactamente $45°$.

Cuidado

Tres avisos sobre la fórmula de la curvatura polar. Primero: **no es** $|r''|/(1+r'^2)^{3/2}$; esa sería la fórmula cartesiana aplicada a la gráfica de f en el plano (θ, r), que es una curva distinta. La aparición del término $-r\,r''$ *restando* y de un $2r'^2$ con coeficiente dos es lo que distingue una fórmula de la otra, y confundirlas da resultados que ni siquiera calibran bien con la circunferencia. Segundo: en los puntos donde $r = r' = 0$ la fórmula no dice nada, porque el denominador se anula; ahí hay que estudiar el comportamiento a mano, como en la cúspide de la cardioide. Y tercero: el valor absoluto del numerador es imprescindible si se quiere una curvatura no negativa; si se prescinde de él se obtiene la *curvatura con signo*, que es útil para detectar puntos de inflexión pero no es lo mismo.

Pausa para pensar

La identidad $x\dot{y} - y\dot{x} = r^2$ del lema es, ella sola, la segunda ley de Kepler. En efecto, si el parámetro es el tiempo en lugar del ángulo, $x\dot{y} - y\dot{x}$ es el módulo del producto vectorial $\vec{r} \times \vec{v}$, y la identidad dice que ese módulo es $r^2\dot{\theta}$: justo la cantidad que la sección de Frenet demuestra que es constante para fuerzas centrales. Es decir, las tres fórmulas de esta sección y la segunda ley de Kepler son consecuencias del mismo par de derivadas. Cuando dos resultados aparentemente lejanos comparten el cálculo intermedio, casi siempre significa que hay un solo teorema debajo.

10.21. Curvas cerradas, simples y regulares: el vocabulario que las fórmulas dan por sabido

Este capítulo ha exigido, en varios sitios, que la curva sea "cerrada y simple" para aplicar la fórmula de Green, ha hablado de "parametrizaciones regulares" al estudiar las cúspides y ha supuesto en el triedro de Frenet que la curva está "parametrizada por el arco". Ninguna de esas tres expresiones se ha definido. Son las hipótesis de los teoremas del capítulo, de modo que sin ellas no se puede saber cuándo se tiene derecho a aplicar una fórmula y cuándo no.

Definición 10.76 (Curva paramétrica, regular, simple y cerrada)**.** Una *curva paramétrica de clase* C^k en $\mathbb{R}^n$ es una aplicación $\vec{c}\colon [a,b] \to \mathbb{R}^n$ cuyas componentes tienen k derivadas continuas. Se dice que:

- $\vec{c}$ es *regular* si $\vec{c}\,'(t) \neq \vec{0}$ para todo t; un t_0 con $\vec{c}\,'(t_0) = \vec{0}$ se llama *punto singular* o *parada*;
- $\vec{c}$ es *cerrada* si $\vec{c}(a) = \vec{c}(b)$;
- $\vec{c}$ es *simple* si es inyectiva en $[a,b)$, es decir, si no se corta a sí misma (se permite el único encuentro $\vec{c}(a) = \vec{c}(b)$ de una curva cerrada);
- $\vec{c}$ es *regular a trozos* si $[a,b]$ se puede partir en un número finito de subintervalos en cada uno de los cuales es regular.

La *traza* de la curva es su imagen $\vec{c}([a,b])$, es decir el dibujo; la curva es la aplicación, no el dibujo.

Definición 10.77 (Reparametrización y orientación)**.** Sea $\vec{c}\colon [a,b] \to \mathbb{R}^n$ y sea $\varphi : [\alpha,\beta] \to [a,b]$ biyectiva, de clase C^1 y con φ' que no se anula. La curva $\vec{d} = \vec{c} \circ \varphi$ se llama *reparametrización* de $\vec{c}$. Si $\varphi' > 0$ la reparametrización *conserva la orientación* (se recorre la traza en el mismo sentido) y si $\varphi' < 0$ la *invierte*.

Teorema 10.78 (La longitud no depende de la parametrización)**.** *Si $\vec{d} = \vec{c} \circ \varphi$ es una reparametrización de $\vec{c}$ en el sentido de la definición 10.77, entonces las dos curvas tienen la misma longitud:*

$$\int_\alpha^\beta \|\vec{d}\,'(u)\|\,\mathrm{d}u = \int_a^b \|\vec{c}\,'(t)\|\,\mathrm{d}t.$$

Demostración. Por la regla de la cadena, $\vec{d}\,'(u) = \vec{c}\,'(\varphi(u))\varphi'(u)$, y tomando normas, $\|\vec{d}\,'(u)\| = \|\vec{c}\,'(\varphi(u))\|\,|\varphi'(u)|$. Supóngase primero $\varphi' > 0$, de modo que $|\varphi'| = \varphi'$, $\varphi(\alpha) = a$ y $\varphi(\beta) = b$. El cambio de variable $t = \varphi(u)$, $\mathrm{d}t = \varphi'(u)\,\mathrm{d}u$, transforma

$$\int_\alpha^\beta \|\vec{c}\,'(\varphi(u))\|\varphi'(u)\,\mathrm{d}u = \int_a^b \|\vec{c}\,'(t)\|\,\mathrm{d}t.$$

Si $\varphi' < 0$ entonces $\varphi(\alpha) = b$ y $\varphi(\beta) = a$, el mismo cambio de variable da $-\int_b^a \|\vec{c}\,'\| = \int_a^b \|\vec{c}\,'\|$, y el signo menos del valor absoluto se compensa exactamente con el intercambio de los límites. □

Teorema 10.79 (Toda curva regular admite parámetro arco)**.** *Sea $\vec{c}\colon [a,b] \to \mathbb{R}^n$ regular de clase C^k con $k \geq 1$, y defínase*

$$s(t) = \int_a^t \|\vec{c}\,'(\tau)\|\,\mathrm{d}\tau, \qquad L = s(b).$$

Entonces s es una biyección estrictamente creciente de $[a,b]$ en $[0,L]$, de clase C^k, con inversa de clase C^k; y la reparametrización $\vec{\gamma}(s) = \vec{c}(t(s))$ cumple $\|\vec{\gamma}\,'(s)\| = 1$ para todo s.

Demostración. Por el teorema fundamental del cálculo, $s'(t) = \|\vec{c}\,'(t)\|$, que es continua y *estrictamente positiva* por la hipótesis de regularidad. Luego s es estrictamente creciente, y por tanto inyectiva; y como es continua con $s(a) = 0$ y $s(b) = L$, es una biyección sobre $[0,L]$. El teorema de la función inversa garantiza que $t(s)$ es derivable con $t'(s) = 1/\|\vec{c}\,'(t(s))\|$, que a su vez es de clase C^{k-1}, de modo que t es C^k. Finalmente, por la regla de la cadena,

$$\vec{\gamma}\,'(s) = \vec{c}\,'(t(s))\,t'(s) = \frac{\vec{c}\,'(t(s))}{\|\vec{c}\,'(t(s))\|},$$

que es un vector unitario. La regularidad se usa exactamente una vez, y sin ella todo se cae: en un punto singular $s'(t) = 0$ y s deja de ser invertible. □

Regla: comprobar las hipótesis antes de aplicar una fórmula

Delante de una curva paramétrica y de una fórmula que se quiere usar, la comprobación es siempre la misma y cuesta un minuto.

1. **Deriva y busca los ceros de $\|\vec{c}\,'\|$.** Si no hay ninguno, la curva es regular y todo funciona. Si los hay, anótalos: son cúspides potenciales, y la curva solo es *regular a trozos*. Las fórmulas de longitud y de área siguen valiendo (la integral no se entera de un punto aislado), pero el triedro de Frenet y la curvatura *no* están definidos ahí.
2. **Comprueba si es cerrada** evaluando en los dos extremos: $\vec{c}(a) = \vec{c}(b)$. Sin eso, la fórmula de Green para el área no dice nada.
3. **Comprueba si es simple** buscando parejas $t_1 \neq t_2$ en $[a,b)$ con $\vec{c}(t_1) = \vec{c}(t_2)$. Si las hay, la curva se corta a sí misma, la fórmula de Green cuenta dos veces el trozo solapado y hay que partir la integral por tramos.

4. **Determina la orientación** calculando el signo de $\frac{1}{2}\oint(x\,\mathrm{d}y - y\,\mathrm{d}x)$ o, más rápido, mirando hacia dónde va la curva cerca de un punto conocido. Positivo es antihorario.

5. **Si el problema pide algo geométrico, reparametriza mentalmente.** La curvatura, la longitud y la traza no cambian al reparametrizar; la velocidad, la aceleración y el tiempo de recorrido sí. Antes de responder, pregúntate a cuál de los dos grupos pertenece lo que te piden.

6. **Si necesitas el parámetro arco, calcúlalo con el teorema 10.79** y no de oído: integra la rapidez desde el extremo inicial y despeja t en función de s. Casi nunca se puede despejar en forma cerrada, y eso no es un problema teórico sino práctico: el teorema garantiza que $t(s)$ existe aunque no sepas escribirlo.

Ejemplo 10.80 (Las curvas del capítulo, clasificadas una por una)**.** (a) **La circunferencia $\vec{c}(t) = (R\cos t, R\sin t)$ en $[0, 2\pi]$.** Rapidez constante $R > 0$: regular. $\vec{c}(0) = \vec{c}(2\pi) = (R, 0)$: cerrada. Inyectiva en $[0, 2\pi)$: simple. Cumple las tres condiciones, y por eso Green da πR^2 sin ninguna precaución.

(b) **La astroide $\vec{c}(t) = (\cos^3 t, \sin^3 t)$.** La rapidez es $3|\sin t\cos t|$, que se anula en $t = 0, \pi/2, \pi, 3\pi/2$: *no* es regular, solo regular a trozos, y esos cuatro puntos son las cuatro cúspides. Es cerrada y simple. Conclusión práctica: la fórmula de Green sí se le puede aplicar (y da $3\pi/8 = 1{,}178097$ con $a = 1$), la fórmula de la longitud también (y da exactamente 6, comprobado numéricamente), pero la curvatura no está definida en las puntas. Ese es el caso más frecuente en todo el capítulo.

(c) **Un arco de cicloide $\vec{c}(t) = (t - \sin t,\ 1 - \cos t)$ en $[0, 2\pi]$.** Rapidez $2|\sin(t/2)|$, nula en $t = 0$ y $t = 2\pi$: regular a trozos. Y *no es cerrada*, porque $\vec{c}(0) = (0, 0)$ y $\vec{c}(2\pi) = (2\pi, 0)$ son puntos distintos. Por eso el área bajo el arco no se calculó con Green sino con $\int y\,x'\,\mathrm{d}t$: la fórmula de Green habría sido inaplicable, y quien la use de todos modos obtiene un número sin significado. Su longitud es 8, comprobada por integración numérica.

(d) **La curva $\vec{c}(t) = (t^2,\ t^3 - 3t)$.** Es regular salvo donde $2t = 0$ y $3t^2 - 3 = 0$ simultáneamente, cosa que no ocurre nunca: es regular en todo $\mathbb{R}$. Pero *no es simple*: $t = \sqrt{3}$ y $t = -\sqrt{3}$ dan los dos el punto $(3, 0)$, porque x depende solo de t^2 y $y(\pm\sqrt{3}) = \pm(3\sqrt{3} - 3\sqrt{3}) = 0$. Es el ejemplo que muestra que regular y simple son condiciones *independientes*: aquí se cumple la primera y falla la segunda, y en la astroide ocurre justo al revés.

(e) **La figura de Lissajous $x = \sin 3t$, $y = \sin 2t$ en $[0, 2\pi]$.** Es cerrada ($\sin 3\cdot 2\pi = 0 = \sin 0$ y lo mismo con el otro seno) y regular, pero no es simple: los valores $t = 0$ y $t = \pi$ dan los dos el origen, porque $\sin 3\pi = 0$ y $\sin 2\pi = 0$. Aplicarle la fórmula de Green sin más devolvería una suma algebraica de las áreas de los lóbulos, con signos distintos según el sentido en que se recorra cada uno, y no el área de la mancha.

Ejemplo 10.81 (Una reparametrización por el arco hecha entera)**.** Tómese la espiral logarítmica $\vec{c}(t) = (\mathrm{e}^t\cos t,\ \mathrm{e}^t\sin t)$ con $t \in [0, 2\pi]$, que es la curva polar $r = \mathrm{e}^\theta$ vista como parametrización.

Paso 1: la rapidez. Derivando, $\vec{c}\,'(t) = \big(\mathrm{e}^t(\cos t - \sin t),\ \mathrm{e}^t(\sin t + \cos t)\big)$, y

$$\|\vec{c}\,'(t)\|^2 = \mathrm{e}^{2t}\big[(\cos t - \sin t)^2 + (\sin t + \cos t)^2\big] = \mathrm{e}^{2t}\cdot 2,$$

porque los dobles productos se cancelan. Luego $\|\vec{c}\,'(t)\| = \sqrt{2}\,\mathrm{e}^t$, que nunca se anula: la curva es regular y el teorema 10.79 se puede aplicar.

Paso 2: la función arco.

$$s(t) = \int_0^t \sqrt{2}\,\mathrm{e}^\tau \,\mathrm{d}\tau = \sqrt{2}(\mathrm{e}^t - 1).$$

Es estrictamente creciente, como anunciaba el teorema, y la longitud total es $L = s(2\pi) = \sqrt{2}(\mathrm{e}^{2\pi} - 1) = 755{,}885348$, valor comprobado por integración numérica.

Paso 3: la inversa. Despejando, $\mathrm{e}^t = 1 + s/\sqrt{2}$, es decir

$$t(s) = \ln\!\left(1 + \frac{s}{\sqrt{2}}\right), \qquad s \in [0, L].$$

Aquí, excepcionalmente, la inversa se escribe en forma cerrada; en la inmensa mayoría de las curvas no se puede, y el teorema sigue garantizando que existe.

Paso 4: la curva reparametrizada y la comprobación. Sustituyendo,

$$\vec{\gamma}(s) = \left(\left(1 + \tfrac{s}{\sqrt{2}}\right)\cos t(s),\quad \left(1 + \tfrac{s}{\sqrt{2}}\right)\sin t(s)\right).$$

Para comprobar que la rapidez es 1 basta con la regla de la cadena: $\|\vec{\gamma}\,'(s)\| = \|\vec{c}\,'(t)\|\cdot|t'(s)| = \sqrt{2}\,\mathrm{e}^t \cdot \frac{1}{\sqrt{2}\,\mathrm{e}^t} = 1$ ✓.

Paso 5: la invariancia de la longitud, comprobada a lo bruto. Tómese ahora una reparametrización arbitraria, $\varphi(u) = u^3$ con $u \in [0, (2\pi)^{1/3}]$. La nueva rapidez es $\sqrt{2}\,\mathrm{e}^{u^3}\cdot 3u^2$, y la integral correspondiente vale 755,885348: exactamente la misma longitud, como garantiza el teorema 10.78. Lo que sí cambia es todo lo demás: con $\varphi(u) = u^3$ la curva se recorre despacio al principio y muy deprisa al final, y su vector velocidad en cada punto es distinto del original. La traza es la misma; el viaje, no.

Idea clave

Conviene fijar de una vez la distinción que atraviesa todo el capítulo, y que ahora tiene nombres precisos. Hay magnitudes *geométricas* —longitud, área encerrada, curvatura, torsión, traza— que no cambian al reparametrizar, y magnitudes *cinemáticas* —velocidad, rapidez, aceleración, tiempo de recorrido— que dependen de la parametrización elegida. El parámetro arco es simplemente la elección canónica que hace coincidir las dos familias, porque al recorrer la curva a rapidez 1 el tiempo transcurrido *es* la distancia recorrida. Por eso el triedro de Frenet se define con él, y por eso todo lo que se deduce en esa notación es automáticamente geométrico. Cuando en el capítulo 36 se hable de invariancia frente a cambios de coordenadas, la estructura del argumento será exactamente esta.

10.22. Ejercicios resueltos

Ejercicio 10.1. Pase el punto $(-2,-2)$ a coordenadas polares.

Solución. $r=\sqrt{4+4}=2\sqrt{2}$. El ángulo de referencia es $\arctan(2/2)=\pi/4$, pero el punto está en el tercer cuadrante, luego $\theta=\pi+\pi/4=5\pi/4$. Respuesta: $(2\sqrt{2},\ 5\pi/4)$. ■

Ejercicio 10.2. Escriba en polares la recta $x+y=2$.

Solución. Sustituyendo, $r\cos\theta+r\sin\theta=2$, es decir $r=\dfrac{2}{\cos\theta+\sin\theta}$. Usando $\cos\theta+\sin\theta=\sqrt{2}\cos(\theta-\pi/4)$ queda $r=\dfrac{\sqrt{2}}{\cos(\theta-\pi/4)}$, que es la forma normal: la recta está a distancia $\sqrt{2}$ del origen en la dirección $\pi/4$. ■

Ejercicio 10.3. Identifique la curva $r=4\sin\theta$.

Solución. Multiplicando por r: $r^2=4r\sin\theta$, o sea $x^2+y^2=4y$, o $x^2+(y-2)^2=4$. Es la circunferencia de centro $(0,2)$ y radio 2, que pasa por el origen. ■

Ejercicio 10.4. Elimine el parámetro en $x=2+3\cos t$, $y=-1+3\sin t$.

Solución. $\dfrac{x-2}{3}=\cos t$ y $\dfrac{y+1}{3}=\sin t$. Elevando al cuadrado y sumando, $(x-2)^2+(y+1)^2=9$: circunferencia de centro $(2,-1)$ y radio 3, recorrida entera y en sentido antihorario. ■

Ejercicio 10.5. Una pelota sale con $v_0=15$ m/s y $\alpha=45^\circ$. ¿A qué distancia cae y qué altura máxima alcanza? ($g=9{,}8$.)

Solución. $R=\dfrac{v_0^2\sin 90^\circ}{g}=\dfrac{225}{9{,}8}=22{,}96$ m. $H=\dfrac{v_0^2\sin^2 45^\circ}{2g}=\dfrac{225\cdot 0{,}5}{19{,}6}=5{,}74$ m. Observe que $H=R/4$: siempre ocurre a 45°, porque $H/R=\tan\alpha/4$. ■

Ejercicio 10.6. Halle el punto de la cicloide $x=t-\sin t$, $y=1-\cos t$ correspondiente a $t=\pi/2$, y compruebe que está a altura 1.

Solución. $x=\pi/2-1\approx 0{,}571$ y $y=1-0=1$. En efecto, cuando la rueda de radio 1 ha girado un cuarto de vuelta, la chincheta está exactamente a la altura del centro. El punto de contacto con el suelo, en cambio, está en $x=\pi/2\approx 1{,}571$: la chincheta va *retrasada* respecto del centro, y esa diferencia es justamente el $\sin t$ de la fórmula. ■

10.23. Ejercicios propuestos

10.1. Pase a polares: $(0,5)$, $(-3,0)$, $(1,-\sqrt{3})$.

10.2. Pase a cartesianas: $(4,\pi/6)$, $(2,\pi)$, $(1,\ 3\pi/2)$.

10.3. ¿Qué curva es $r=6\cos\theta$?

10.4. Escriba en polares la circunferencia $x^2+y^2=25$ y la recta $y=x$.

10.5. ¿Cuántos pétalos tiene $r=\sin 4\theta$? ¿Y $r=\sin 5\theta$?

10.6. Tabule $r=1-\sin\theta$ para $\theta=0,\pi/2,\pi,3\pi/2$ y diga qué forma tiene.

10.7. Elimine el parámetro en $x=t^2$, $y=t^4$, e indique qué trozo de curva se obtiene.

10.8. Elimine el parámetro en $x = 2\sec t$, $y = 3\tan t$.

10.9. Parametrice el segmento que une $(1, 2)$ con $(5, -2)$ de modo que $t \in [0, 1]$.

10.10. Un proyectil se lanza a 30 m/s con $\alpha = 60°$. Calcule T, R y H (tome $g = 9{,}8$).

10.11. ¿Con qué otro ángulo se consigue el mismo alcance que con 25°?

10.12. Halle la excentricidad y el semilado recto de $r = \dfrac{6}{2 + \cos\theta}$.

Respuestas

1. $(5, \pi/2)$; $(3, \pi)$; $(2,\ 5\pi/3)$. **2.** $(2\sqrt{3},\ 2)$; $(-2, 0)$; $(0, -1)$. **3.** Circunferencia de centro $(3, 0)$ y radio 3. **4.** $r = 5$ y $\theta = \pi/4$. **5.** Ocho pétalos y cinco pétalos, respectivamente. **6.** 1, 0, 1, 2: una cardioide con la punta hacia arriba (apuntando a $\theta = \pi/2$). **7.** $y = x^2$, pero solo el trozo con $x \geq 0$. **8.** $\dfrac{x^2}{4} - \dfrac{y^2}{9} = 1$, una hipérbola. **9.** $x = 1 + 4t$, $y = 2 - 4t$. **10.** $T = 5{,}30$ s, $R = 79{,}53$ m, $H = 34{,}44$ m. **11.** Con 65°. **12.** Dividiendo entre 2, $r = \dfrac{3}{1 + \frac{1}{2}\cos\theta}$: excentricidad $e = \frac{1}{2}$ (elipse) y semilado recto $\ell = 3$.

Capítulo 11

Cálculo vectorial

Lo que necesitas traer

Del capítulo 9: derivadas parciales, gradiente, integrales dobles y triples, y el cambio de variable. Del capítulo 7: producto escalar, producto vectorial y determinantes. Del capítulo 10: parametrización de curvas.

Este capítulo estudia integrales sobre curvas y superficies, y culmina en los teoremas de Green, Stokes y Gauss: versiones multidimensionales del teorema fundamental del cálculo. Es el lenguaje del electromagnetismo y de la mecánica de fluidos, y el puente directo al cálculo tensorial.

11.1. Curvas y su geometría

Idea clave

Intuición primero. Una curva parametrizada no es un dibujo: es un *recorrido*. La misma circunferencia se puede recorrer despacio o deprisa, una vez o tres, en un sentido o en el otro, y todas esas son parametrizaciones distintas del mismo dibujo. Esa distinción entre el camino y la manera de recorrerlo organiza toda la sección: la velocidad, la aceleración y el parámetro dependen del recorrido, mientras que la longitud, la curvatura y la torsión son propiedades del dibujo.
Para separar limpiamente ambas cosas se usa el truco de parametrizar por longitud de arco, que es recorrer la curva a velocidad constante igual a uno: entonces el tiempo y la distancia coinciden y todo lo que quede depende solo de la forma. Con esa normalización, la curvatura mide cuánto se dobla la curva —el inverso del radio del círculo que mejor se le pega— y la torsión mide cuánto se sale del plano. El triedro de Frenet es simplemente un sistema de ejes que viaja pegado a la curva, y sus fórmulas cuentan cómo gira ese sistema al avanzar.

Una *curva parametrizada* es $\mathbf{r} : [a, b] \to \mathbb{R}^3$, $\mathbf{r}(t) = (x(t), y(t), z(t))$; su velocidad es $\mathbf{r}'(t)$, su

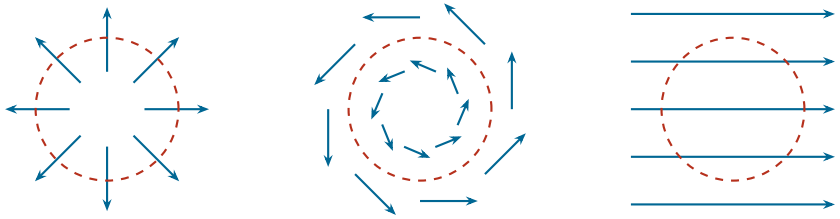


Figura 11.1: Los dos operadores, vistos. La divergencia mide cuánto sale de una región pequeña: es positiva en la fuente de la izquierda y nula en las otras dos. El rotacional mide cuánto gira alrededor de un punto: sólo el campo central da vueltas.

rapidez $\|\mathbf{r}'(t)\|$, y es *regular* si $\mathbf{r}' \neq \mathbf{0}$. La *longitud de arco* es

$$L = \int_a^b \|\mathbf{r}'(t)\|\,\mathrm{d}t, \qquad s(t) = \int_a^t \|\mathbf{r}'(u)\|\,\mathrm{d}u,$$

y s es el parámetro "natural": reparametrizada por s, la curva se recorre con rapidez 1.

Definición 11.1 (Aparato de Frenet)**.** Para una curva regular parametrizada por arco: el *tangente* $\mathbf{T} = \mathbf{r}'(s)$; la *curvatura* $\kappa = \|\mathbf{T}'(s)\|$ (cuánto se dobla); el *normal* $\mathbf{N} = \mathbf{T}'/\kappa$; el *binormal* $\mathbf{B} = \mathbf{T} \times \mathbf{N}$; y la *torsión* τ (cuánto sale del plano), definida por $\mathbf{B}' = -\tau\mathbf{N}$. Fórmula práctica sin reparametrizar:

$$\kappa = \frac{\|\mathbf{r}' \times \mathbf{r}''\|}{\|\mathbf{r}'\|^3}.$$

Ejemplo 11.2. La hélice $\mathbf{r}(t) = (a\cos t, a\sin t, bt)$ tiene $\|\mathbf{r}'\| = \sqrt{a^2+b^2}$ constante, curvatura $\kappa = \frac{a}{a^2+b^2}$ y torsión $\tau = \frac{b}{a^2+b^2}$, ambas constantes: es la única curva con esa propiedad (recta y circunferencia aparte, que son casos límite $a = 0$ o $b = 0$).

Regla: calcular el triedro de Frenet, la curvatura y la torsión

Toda curva del espacio lleva pegado un sistema de referencia propio que viaja con ella. Dos números —cuánto se dobla y cuánto se sale del plano— la determinan por completo salvo posición.

1. **Deriva tres veces** la parametrización $\mathbf{r}(t)$ y guarda $\mathbf{r}'$, $\mathbf{r}''$ y $\mathbf{r}'''$. Casi todo se construye con estas tres.
2. **Calcula la rapidez** $v = \|\mathbf{r}'\|$ y el vector tangente unitario $\mathbf{T} = \mathbf{r}'/v$. La longitud de arco es $s(t) = \int v\,\mathrm{d}t$.
3. **Usa las fórmulas con producto vectorial**, que evitan tener que reparametrizar por arco:
$$\kappa = \frac{\|\mathbf{r}' \times \mathbf{r}''\|}{\|\mathbf{r}'\|^3}, \qquad \tau = \frac{(\mathbf{r}' \times \mathbf{r}'') \cdot \mathbf{r}'''}{\|\mathbf{r}' \times \mathbf{r}''\|^2}.$$
Éste es el atajo que hay que memorizar: reparametrizar por longitud de arco casi nunca es viable a mano.

4. **Obtén los otros dos vectores del triedro:** la binormal $\mathbf{B} = \dfrac{\mathbf{r}' \times \mathbf{r}''}{\|\mathbf{r}' \times \mathbf{r}''\|}$ y la normal principal $\mathbf{N} = \mathbf{B} \times \mathbf{T}$. Los tres forman una base ortonormal directa en cada punto.
5. **Interpreta cada uno. T** apunta hacia donde vas; **N** hacia el centro de la curvatura, es decir, hacia donde giras; **B** es perpendicular al plano en el que la curva "vive localmente" (el plano osculador).
6. **Lee los dos números.** κ mide cuánto se aparta la curva de su recta tangente: $\kappa = 0$ significa recta, y $1/\kappa$ es el radio del círculo que mejor la aproxima. τ mide cuánto se sale del plano osculador: $\tau = 0$ en todo punto significa curva plana.
7. **Controles rápidos:** para la circunferencia de radio R debe salir $\kappa = 1/R$ y $\tau = 0$; y en toda curva $\kappa \geq 0$ siempre, mientras que τ puede tener cualquier signo.

Ejemplo 11.3 (Tres curvas y sus dos números)**. (a) La hélice circular** $\mathbf{r}(t) = (a\cos t,\ a\sin t,\ bt)$ con $a > 0$. Derivando,

$$\mathbf{r}' = (-a\sin t,\ a\cos t,\ b), \qquad \mathbf{r}'' = (-a\cos t,\ -a\sin t,\ 0), \qquad \mathbf{r}''' = (a\sin t,\ -a\cos t,\ 0).$$

La rapidez es constante, $v = \sqrt{a^2+b^2}$, y el producto vectorial $\mathbf{r}' \times \mathbf{r}'' = (ab\sin t,\ -ab\cos t,\ a^2)$ tiene norma $a\sqrt{a^2+b^2}$. Sustituyendo en el paso 3:

$$\kappa = \frac{a\sqrt{a^2+b^2}}{(a^2+b^2)^{3/2}} = \frac{a}{a^2+b^2}, \qquad \tau = \frac{a^2 b}{a^2(a^2+b^2)} = \frac{b}{a^2+b^2}.$$

Ambos son *constantes*: la hélice es la única curva del espacio con curvatura y torsión constantes no nulas, y por eso todos los muelles se parecen. Control del paso 7: haciendo $b \to 0$ la hélice se aplasta en la circunferencia de radio a y en efecto $\kappa \to 1/a$ y $\tau \to 0$. Para $a = b = 1$ sale $\kappa = \tau = \frac{1}{2}$.

(b) La cúbica alabeada $\mathbf{r}(t) = (t,\ t^2,\ t^3)$. Ahora $\mathbf{r}' = (1, 2t, 3t^2)$, $\mathbf{r}'' = (0, 2, 6t)$ y $\mathbf{r}''' = (0, 0, 6)$, con

$$\mathbf{r}' \times \mathbf{r}'' = (6t^2,\ -6t,\ 2), \qquad \|\mathbf{r}'\| = \sqrt{9t^4+4t^2+1}.$$

De ahí

$$\kappa(t) = \frac{2\sqrt{9t^4+9t^2+1}}{(9t^4+4t^2+1)^{3/2}}, \qquad \tau(t) = \frac{3}{9t^4+9t^2+1}.$$

En el origen, $\kappa(0) = 2$ y $\tau(0) = 3$: la curva se dobla y además se sale del plano. En $t = 1$, $\kappa(1) = \frac{2\sqrt{19}}{14^{3/2}} = \frac{\sqrt{266}}{98} \approx 0{,}16642$ y $\tau(1) = \frac{3}{19} \approx 0{,}15789$: lejos del origen la curva se ha estirado y ambas cantidades caen. La torsión nunca se anula, así que ningún trozo de esta curva es plano.

(c) Una curva plana, $y = x^2$**.** Parametrizando $\mathbf{r}(t) = (t, t^2, 0)$, el producto $\mathbf{r}' \times \mathbf{r}'' = (0, 0, 2)$ es constante y perpendicular al plano $z = 0$, luego $\mathbf{B}$ es fijo y $\tau = 0$ en todo punto: la curva es plana, como debía. La curvatura queda

$$\kappa(t) = \frac{2}{(1+4t^2)^{3/2}},$$

máxima en el vértice, $\kappa(0) = 2$, y tendiendo a 0 en las ramas, donde la parábola se parece cada vez más a una recta. El radio de curvatura en el vértice es $1/2$.

Idea clave

El teorema fundamental de las curvas dice que si te dan dos funciones $\kappa(s) > 0$ y $\tau(s)$, existe una curva con esa curvatura y esa torsión, y es única salvo un movimiento rígido. Es decir: κ y τ contienen *toda* la información geométrica de la curva. La parametrización concreta es un disfraz; estos dos números son la curva.

Cuidado

Las fórmulas del paso 3 exigen $\mathbf{r}' \neq \mathbf{0}$ (curva regular) y, para la torsión, además $\mathbf{r}' \times \mathbf{r}'' \neq \mathbf{0}$, es decir $\kappa \neq 0$. En un punto de inflexión donde $\kappa = 0$ la torsión sencillamente no está definida: sin plano osculador no hay nada de lo que salirse. Un ejemplo es $\mathbf{r}(t) = (t, t^3, 0)$ en $t = 0$.

11.2. Cimientos del cálculo vectorial: campos, curvas, superficies y las integrales que se hacen sobre ellos

Idea clave

De qué va todo esto. El cálculo de una variable integraba funciones sobre *intervalos*. El cálculo vectorial hace lo mismo cambiando el intervalo por objetos más interesantes —una curva torcida en el espacio, un trozo de superficie, un sólido— y cambiando también lo que se integra: ya no solo números pegados a cada punto, sino *flechas*.
Esa doble generalización produce exactamente cuatro objetos nuevos, y toda la dificultad del tema consiste en no confundirlos entre sí:

$$\int_{\mathcal{C}} f \, \mathrm{d}s, \qquad \int_{\mathcal{C}} \vec{F} \cdot \mathrm{d}\vec{r}, \qquad \iint_{S} f \, \mathrm{d}S, \qquad \iint_{S} \vec{F} \cdot \mathrm{d}\vec{S}.$$

Los primeros de cada pareja miden *cantidades* (la masa de un alambre, el área de una chapa); los segundos miden *cuánto empuja o cuánto atraviesa* un campo de flechas (trabajo, flujo). Los primeros no distinguen el sentido del recorrido; los segundos cambian de signo si se recorre al revés. Esta sección construye desde cero las cuatro, junto con los cuatro operadores diferenciales —gradiente, divergencia, rotacional y laplaciano— que las relacionan. Todo lo demás del capítulo (Green, Stokes, Gauss) son igualdades entre estos objetos, y no se pueden entender antes de tenerlos claros uno por uno.

11.2.1. Campos escalares y campos vectoriales

Antes de integrar sobre curvas hay que decir qué se integra. En física uno no mide una función de una variable: mide una magnitud *en cada punto del espacio*. La temperatura de una habitación es un número por punto; la velocidad del aire en esa misma habitación es una flecha por punto. Esos dos tipos de objeto —número por punto y flecha por punto— son los

protagonistas del capítulo, y conviene bautizarlos con precisión, porque casi todos los errores del tema nacen de aplicarle a uno lo que solo vale para el otro.

Definición 11.4 (Campo escalar)**.** Sea $D \subset \mathbb{R}^n$ un conjunto abierto. Un *campo escalar* sobre D es una función $f : D \to \mathbb{R}$, es decir, una regla que asigna a cada punto $\vec{p} \in D$ un *número* $f(\vec{p})$. Se escribe $f(x, y)$ en el plano y $f(x, y, z)$ en el espacio. Diremos que f es de clase C^k si todas sus derivadas parciales hasta el orden k existen y son continuas.

Definición 11.5 (Campo vectorial)**.** Sea $D \subset \mathbb{R}^n$ abierto. Un *campo vectorial* sobre D es una función $\vec{F} : D \to \mathbb{R}^n$ que asigna a cada punto $\vec{p} \in D$ un *vector* $\vec{F}(\vec{p})$ del mismo espacio. En el plano se escribe $\vec{F}(x, y) = \big(P(x, y),\, Q(x, y)\big)$ y en el espacio

$$\vec{F}(x, y, z) = \big(P(x, y, z),\, Q(x, y, z),\, R(x, y, z)\big),$$

donde P, Q y R se llaman las *componentes* de $\vec{F}$ y son, cada una de ellas, campos escalares. El campo es de clase C^k si lo son sus componentes.

Observación 11.6 (Notación). En este libro un vector se escribe indistintamente $\vec{F}$ o $\mathbf{F}$: ambas grafías significan lo mismo y conviven en el capítulo. Los vectores de la base canónica son $\vec{\imath} = (1, 0, 0)$, $\vec{\jmath} = (0, 1, 0)$ y $\vec{k} = (0, 0, 1)$, de modo que $\vec{F} = P\,\vec{\imath} + Q\,\vec{\jmath} + R\,\vec{k}$ es otra forma de escribir la misma terna. El símbolo $\|\vec{v}\|$ es la longitud del vector: $\|(a, b, c)\| = \sqrt{a^2 + b^2 + c^2}$.

Cómo se dibujan. Un campo escalar se representa por sus *curvas de nivel* (en el plano) o *superficies de nivel* (en el espacio): los conjuntos $f = c$ donde el campo toma un valor fijo. Un mapa del tiempo con isotermas es exactamente eso. Un campo vectorial se representa pinchando, en una retícula de puntos, la flecha $\vec{F}(\vec{p})$ con origen en $\vec{p}$; para que el dibujo sea legible se suele reducir la escala de todas las flechas por igual. La figura 11.1 de este capítulo es justamente un retrato de tres campos vectoriales dibujados así.

Ejemplo 11.7 (Los dos tipos, con números)**. (a) Un campo escalar.** Sea $T(x, y, z) = x^2 + y^2 + z^2$, que podemos leer como la temperatura de una habitación cuyo punto más frío está en el origen. En $(1, 2, 2)$ vale $T = 1 + 4 + 4 = 9$, y en $(0, 0, 3)$ vale 9 también. Sus superficies de nivel $T = c$ son las esferas centradas en el origen: todos los puntos de una misma esfera están igual de calientes. Es un número por punto y nada más; preguntar "¿hacia dónde apunta la temperatura?" no tiene sentido.

(b) Un campo vectorial. Sea $\vec{F}(x, y) = (-y,\, x)$, que podemos leer como la velocidad del agua en un remolino. En el punto $(1, 0)$ vale $\vec{F} = (0, 1)$: una flecha que apunta hacia arriba, de longitud 1. En $(0, 1)$ vale $(-1, 0)$: apunta hacia la izquierda. En $(2, 0)$ vale $(0, 2)$: hacia arriba y el doble de larga. Dibujando unas cuantas se ve el patrón: en cada punto la flecha es perpendicular al radio y su longitud crece con la distancia al origen, es decir, el fluido gira alrededor del origen como si fuera un disco rígido. Obsérvese que $\vec{F}(\vec{p}) \cdot \vec{p} = -yx + xy = 0$ en todo punto: eso *demuestra* la perpendicularidad que el dibujo solo sugería.

(c) Ejemplos físicos que conviene tener en la cabeza. Campos escalares: la temperatura, la presión, la densidad, el potencial eléctrico, la altura del terreno. Campos vectoriales: la velocidad de un fluido, la fuerza gravitatoria, el campo eléctrico $\vec{E}$, el campo magnético $\vec{B}$,

el viento. La distinción no es un capricho matemático: un termómetro devuelve un número, una veleta devuelve una dirección.

11.2.2. Curvas parametrizadas, vector tangente y elemento de longitud

Para integrar a lo largo de una curva hace falta primero describirla con una fórmula. La manera útil no es dar una ecuación implícita como $x^2+y^2=1$, sino *recorrerla*: dar la posición en cada instante. Esa es la idea de parametrización, y trae de regalo la velocidad, que es lo que convierte una integral sobre la curva en una integral ordinaria en una variable.

Definición 11.8 (Curva parametrizada, regularidad y vector tangente)**.** Una *curva parametrizada* en $\mathbb{R}^n$ es una aplicación continua y de clase C^1 a trozos

$$\vec{r}\colon [a,b] \longrightarrow \mathbb{R}^n, \qquad \vec{r}(t) = \big(x(t), y(t), z(t)\big).$$

Su imagen $\mathcal{C} = \vec{r}([a,b])$ es el *trazo* de la curva: el dibujo. El vector $\vec{r}\,'(t) = \big(x'(t),\, y'(t),\, z'(t)\big)$ se llama *vector velocidad* o *vector tangente* en el punto $\vec{r}(t)$, y su norma $\|\vec{r}\,'(t)\|$ es la *rapidez*. La parametrización es *regular* si $\vec{r}\,'(t) \neq \vec{0}$ para todo t; en ese caso el vector tangente unitario es $\vec{T}(t) = \vec{r}\,'(t)/\|\vec{r}\,'(t)\|$. La curva es *cerrada* si $\vec{r}(a) = \vec{r}(b)$, y *simple* si no se corta a sí misma (salvo, en el caso cerrado, por los extremos).

La regularidad no es un tecnicismo: es lo que garantiza que la curva tiene una dirección bien definida en cada punto. La parametrización $\vec{r}(t) = (t^2, t^3)$ tiene $\vec{r}\,'(0) = \vec{0}$, y en efecto su trazo presenta un pico en el origen.

Definición 11.9 (Longitud de arco y elemento de longitud)**.** Para una curva parametrizada $\vec{r}$ sobre $[a,b]$, la *longitud* es $L = \int_a^b \|\vec{r}\,'(t)\|\,\mathrm{d}t$, y se escribe simbólicamente

$$\boxed{\mathrm{d}s = \|\vec{r}\,'(t)\|\,\mathrm{d}t}$$

para el *elemento de longitud de arco*. La razón del nombre es que en un intervalo de tiempo diminuto $\mathrm{d}t$ el punto se desplaza aproximadamente $\vec{r}\,'(t)\,\mathrm{d}t$, y la longitud de ese desplazamiento es $\|\vec{r}\,'(t)\|\,\mathrm{d}t$: espacio igual a rapidez por tiempo. Sumando todos esos trocitos sale la longitud total.

Definición 11.10 (Orientación de una curva)**.** Una parametrización regular induce en el trazo un *sentido de recorrido*, dado por la dirección del vector tangente: eso es la *orientación* de la curva. Una curva simple admite exactamente dos orientaciones. Si $\vec{r}$ está definida en $[a,b]$, la parametrización $\vec{r}^{\,-}(t) = \vec{r}(a+b-t)$ con $t \in [a,b]$ recorre el mismo trazo en sentido contrario; la curva así orientada se denota $-\mathcal{C}$ y cumple $(\vec{r}^{\,-})'(t) = -\vec{r}\,'(a+b-t)$: el vector tangente se da la vuelta, mientras que su *norma* no cambia. Esta observación tan tonta es la que gobierna todos los signos del capítulo.

Ejemplo 11.11 (Cuatro parametrizaciones y sus elementos de longitud)**. (a) Un segmento.** Para ir de $\vec{A} = (1,2)$ a $\vec{B} = (4,6)$ se usa siempre $\vec{r}(t) = \vec{A} + t(\vec{B}-\vec{A}) = (1+3t,\ 2+4t)$ con $t \in [0,1]$. Entonces $\vec{r}\,'(t) = (3,4)$, constante, y $\|\vec{r}\,'\| = \sqrt{9+16} = 5$, luego $\mathrm{d}s = 5\,\mathrm{d}t$ y $L = \int_0^1 5\,\mathrm{d}t = 5$. En efecto, la distancia de $\vec{A}$ a $\vec{B}$ es 5. Es regular porque $(3,4) \neq \vec{0}$.

(b) Una circunferencia. Para la de centro el origen y radio 2 tomamos $\vec{r}(t) = (2\cos t,\ 2\sin t)$ con $t \in [0, 2\pi]$, de donde $\vec{r}\,'(t) = (-2\sin t,\ 2\cos t)$ y su norma es

$$\|\vec{r}\,'(t)\| = \sqrt{4\sin^2 t + 4\cos^2 t} = 2.$$

Por tanto $\mathrm{d}s = 2\,\mathrm{d}t$ y $L = \int_0^{2\pi} 2\,\mathrm{d}t = 4\pi$, que es $2\pi R$ con $R = 2$. ✓ El vector tangente $(-2\sin t, 2\cos t)$ es perpendicular al radio $(2\cos t, 2\sin t)$, como se comprueba haciendo el producto escalar: $-4\sin t\cos t + 4\cos t\sin t = 0$.

(c) La misma circunferencia, recorrida el doble de rápido. Con $\vec{q}(t) = (2\cos 2t,\ 2\sin 2t)$ y $t \in [0, \pi]$ se obtiene $\|\vec{q}\,'\| = 4$ y $L = \int_0^{\pi} 4\,\mathrm{d}t = 4\pi$: la misma longitud. La longitud depende del *trazo*, no de la prisa con que se recorra; el factor 2 que gana la rapidez lo pierde el intervalo.

(d) Un trozo de gráfica. Para $y = x^2$ con $x \in [0, 1]$ se toma siempre $\vec{r}(t) = (t,\, t^2)$, de donde $\vec{r}\,' = (1, 2t)$ y $\mathrm{d}s = \sqrt{1 + 4t^2}\,\mathrm{d}t$. La longitud es

$$L = \int_0^1 \sqrt{1 + 4t^2}\,\mathrm{d}t = \frac{\sqrt{5}}{2} + \frac{1}{4}\ln\bigl(2 + \sqrt{5}\bigr) \approx 1{,}4789,$$

un poco mayor que la cuerda $\sqrt{2} \approx 1{,}4142$, como debía ser.

11.2.3. Las dos integrales de línea, que son cosas distintas

Y aquí llega el punto donde se pierde media clase. Sobre una misma curva se pueden integrar dos cosas radicalmente distintas: un campo escalar y un campo vectorial. Los símbolos se parecen, los nombres se parecen y, sin embargo, miden magnitudes diferentes y se comportan de forma opuesta ante un cambio de sentido. Vamos a definirlas por separado y luego a enfrentarlas.

Definición 11.12 (Integral de línea de un campo escalar)**.** Sea $\mathcal{C}$ una curva con parametrización regular $\vec{r}\colon [a, b] \to \mathbb{R}^n$ y sea f un campo escalar continuo definido sobre el trazo de $\mathcal{C}$. La *integral de línea de f respecto de la longitud de arco* es

$$\int_{\mathcal{C}} f\,\mathrm{d}s \;:=\; \int_a^b f\bigl(\vec{r}(t)\bigr)\,\|\vec{r}\,'(t)\|\,\mathrm{d}t.$$

Su significado: si f es la densidad lineal de un alambre con forma de $\mathcal{C}$, esta integral es la *masa* del alambre. Con $f \equiv 1$ devuelve la longitud.

Definición 11.13 (Integral de línea de un campo vectorial: trabajo y circulación)**.** Sea $\mathcal{C}$ una curva *orientada* con parametrización regular $\vec{r}\colon [a, b] \to \mathbb{R}^n$ compatible con la orientación, y sea $\vec{F}$ un campo vectorial continuo sobre el trazo. Escribiendo $\mathrm{d}\vec{r} = \vec{r}\,'(t)\,\mathrm{d}t$, la *integral de línea de $\vec{F}$ a lo largo de $\mathcal{C}$* es

$$\int_{\mathcal{C}} \vec{F}\cdot\mathrm{d}\vec{r} \;:=\; \int_a^b \vec{F}\bigl(\vec{r}(t)\bigr)\cdot\vec{r}\,'(t)\,\mathrm{d}t = \int_{\mathcal{C}} P\,\mathrm{d}x + Q\,\mathrm{d}y + R\,\mathrm{d}z.$$

Si $\vec{F}$ es un campo de fuerzas, esta integral es el *trabajo* realizado al recorrer $\mathcal{C}$. Cuando la curva es cerrada se llama *circulación* de $\vec{F}$ y se escribe $\oint_{\mathcal{C}} \vec{F}\cdot\mathrm{d}\vec{r}$.

Merece la pena entender de dónde sale el producto escalar. Sobre un trocito de curva el desplazamiento es $\mathrm{d}\vec{r}$; de la fuerza $\vec{F}$ solo trabaja la componente que va en la dirección del desplazamiento, y esa componente vale $\|\vec{F}\|\cos\theta$, donde θ es el ángulo entre ambos. Multiplicando por la longitud $\|\mathrm{d}\vec{r}\|$ del desplazamiento queda $\|\vec{F}\|\,\|\mathrm{d}\vec{r}\|\cos\theta = \vec{F}\cdot\mathrm{d}\vec{r}$. El producto escalar es exactamente el descuento por dirección: si empujas perpendicularmente al movimiento, no trabajas.

Proposición 11.14 (Comportamiento frente a la orientación)**.** *Sea $\mathcal{C}$ una curva y $-\mathcal{C}$ la misma curva recorrida en sentido contrario. Entonces*

$$\int_{-\mathcal{C}} f\,\mathrm{d}s = \int_{\mathcal{C}} f\,\mathrm{d}s \qquad y \qquad \int_{-\mathcal{C}} \vec{F}\cdot\mathrm{d}\vec{r} = -\int_{\mathcal{C}} \vec{F}\cdot\mathrm{d}\vec{r}.$$

Además, ninguna de las dos cambia si se reparametriza la curva conservando el sentido.

Demostración. Con $\vec{r}^{\,-}(t) = \vec{r}(a+b-t)$ de la definición 11.10 y el cambio de variable $u = a+b-t$ (con $\mathrm{d}u = -\mathrm{d}t$, que además intercambia los límites de integración), la integral escalar lleva $\|(\vec{r}^{\,-})'(t)\| = \|-\vec{r}\,'(u)\| = \|\vec{r}\,'(u)\|$, y los dos signos negativos —el de $\mathrm{d}u$ y el del intercambio de límites— se cancelan: sale la misma integral. La vectorial lleva $(\vec{r}^{\,-})'(t) = -\vec{r}\,'(u)$, y ese menos *no* se cancela con nada: sobrevive y cambia el signo del resultado. La invariancia frente a una reparametrización con $\varphi' > 0$ es el mismo cambio de variable, ahora sin intercambio de límites. □

Cuidado

[No confundir $\mathrm{d}s$ con $\mathrm{d}\vec{r}$] Es el error más silencioso del capítulo, porque produce un número plausible en vez de un mensaje de error. Reténgase la tabla:

objeto	$\int_{\mathcal{C}} f\,\mathrm{d}s$	$\int_{\mathcal{C}} \vec{F}\cdot\mathrm{d}\vec{r}$
qué se integra	un campo *escalar*	un campo *vectorial*
por qué se multiplica	$\|\vec{r}\,'(t)\|$ (un número ≥ 0)	$\vec{r}\,'(t)$ (un vector con signo)
depende de la orientación	no	sí, cambia de signo
qué mide	masa, longitud, promedio	trabajo, circulación

El elemento $\mathrm{d}s = \|\vec{r}\,'\|\,\mathrm{d}t$ es una *longitud*: siempre positivo, sin flecha. El elemento $\mathrm{d}\vec{r} = \vec{r}\,'\,\mathrm{d}t$ es un *vector*: lleva flecha, y por eso al dar la vuelta apunta al revés. Regla práctica para saber cuál toca: si el integrando es una cantidad (una densidad, una temperatura, el número 1), es la escalar; si es una flecha (una fuerza, una velocidad, un campo), es la vectorial, y entonces hay que decir en voz alta en qué sentido se recorre la curva.

Ejemplo 11.15 (Dos integrales de línea de campo escalar, enteras)**. (a) La masa de un alambre recto.** Un alambre ocupa el segmento de $(0,0)$ a $(3,4)$ y su densidad lineal en cada punto es $f(x,y) = x+y$. Queremos su masa.

Paso 1, parametrizar. $\vec{r}(t) = (3t,\,4t)$ con $t \in [0,1]$.

Paso 2, el elemento de longitud. $\vec{r}\,'(t) = (3,4)$, luego $\|\vec{r}\,'(t)\| = \sqrt{9+16} = 5$ y $\mathrm{d}s = 5\,\mathrm{d}t$.

Paso 3, sustituir el integrando. $f(\vec{r}(t)) = 3t + 4t = 7t$. Ya no queda ninguna x ni ninguna y: solo t.

Paso 4, integrar.

$$M = \int_{\mathcal{C}} (x+y)\,\mathrm{d}s = \int_0^1 7t \cdot 5\,\mathrm{d}t = 35\int_0^1 t\,\mathrm{d}t = \frac{35}{2} = 17{,}5.$$

Control. La longitud del alambre es 5, así que la densidad media es $17{,}5/5 = 3{,}5$. Y en efecto f vale 0 en un extremo y 7 en el otro y crece linealmente, de modo que su media es $(0+7)/2 = 3{,}5$. ✓

(b) Un arco de circunferencia, y la orientación no importa. Calculemos $\int_{\mathcal{C}} xy\,\mathrm{d}s$, donde $\mathcal{C}$ es el cuarto de circunferencia de radio 3 centrada en el origen que va de $(3,0)$ a $(0,3)$. Parametrizamos $\vec{r}(t) = (3\cos t,\ 3\sin t)$ con $t \in [0, \pi/2]$. Entonces $\vec{r}\,'(t) = (-3\sin t,\ 3\cos t)$ y $\|\vec{r}\,'(t)\| = 3$, luego $\mathrm{d}s = 3\,\mathrm{d}t$. Sustituyendo, $f(\vec{r}(t)) = 3\cos t \cdot 3\sin t = 9\sin t\cos t$, y

$$\int_{\mathcal{C}} xy\,\mathrm{d}s = \int_0^{\pi/2} 9\sin t\cos t \cdot 3\,\mathrm{d}t = 27\int_0^{\pi/2} \sin t\cos t\,\mathrm{d}t = 27\left[\frac{\sin^2 t}{2}\right]_0^{\pi/2} = \frac{27}{2} = 13{,}5.$$

Y ahora al revés. Si recorremos el mismo cuarto de $(0,3)$ a $(3,0)$, la parametrización es $\vec{q}(u) = (3\cos(\frac{\pi}{2}-u),\ 3\sin(\frac{\pi}{2}-u)) = (3\sin u,\, 3\cos u)$ con $u \in [0,\pi/2]$, otra vez con $\|\vec{q}\,'\| = 3$, y el integrando queda $9\sin u\cos u$: literalmente la misma integral, y sale otra vez $27/2$. El resultado no ha cambiado, tal como anunciaba la proposición 11.14: la masa de un alambre no depende de por qué punta se empiece a pesarlo.

Ejemplo 11.16 (Dos integrales de línea de campo vectorial, enteras)**. (a) Circulación en un remolino, y el signo de la orientación.** Sea $\vec{F} = (-y,\, x)$ y sea $\mathcal{C}$ la circunferencia de radio a centrada en el origen, recorrida en sentido *antihorario*.

Paso 1. $\vec{r}(t) = (a\cos t,\ a\sin t)$ con $t \in [0, 2\pi]$. Este sentido es el antihorario: al crecer t se pasa de $(a,0)$ a $(0,a)$.

Paso 2. $\vec{r}\,'(t) = (-a\sin t,\ a\cos t)$. Aquí **no** se normaliza: la integral vectorial usa $\vec{r}\,'$ entero.

Paso 3. $\vec{F}(\vec{r}(t)) = (-a\sin t,\ a\cos t)$, y el producto escalar es

$$\vec{F}(\vec{r}(t)) \cdot \vec{r}\,'(t) = (-a\sin t)(-a\sin t) + (a\cos t)(a\cos t) = a^2\sin^2 t + a^2\cos^2 t = a^2.$$

Paso 4. $\oint_{\mathcal{C}} \vec{F}\cdot\mathrm{d}\vec{r} = \int_0^{2\pi} a^2\,\mathrm{d}t = 2\pi a^2$. Para $a = 2$ da $8\pi \approx 25{,}13$. El resultado es positivo porque el campo, que gira en sentido antihorario, empuja siempre a favor del recorrido.

Al revés. Con $\vec{q}(t) = (a\cos t,\, -a\sin t)$, que recorre la misma circunferencia en sentido horario, se obtiene $\vec{q}\,' = (-a\sin t,\, -a\cos t)$ y $\vec{F}(\vec{q}) = (a\sin t,\, a\cos t)$, cuyo producto escalar vale $-a^2$: la integral da $-2\pi a^2$. Ha cambiado el signo, como exigía la proposición 11.14, porque ahora se avanza contra el campo.

El contraste. Sobre esa misma circunferencia, la integral escalar $\int_{\mathcal{C}} 1\,\mathrm{d}s$ vale $2\pi a$ en los dos sentidos. Dos integrales sobre la misma curva, dos números distintos ($2\pi a$ frente a $2\pi a^2$) y dos comportamientos opuestos ante el cambio de sentido: son objetos diferentes.

(b) El trabajo depende del camino. Sea $\vec{F} = (x^2,\ xy)$ y calculemos el trabajo de $(0,0)$ a $(1,1)$ por dos caminos.

Camino 1: la recta $y = x$. Con $\vec{r}(t) = (t, t)$, $\vec{r}\,' = (1, 1)$ y $\vec{F}(\vec{r}(t)) = (t^2,\ t^2)$:

$$\vec{F} \cdot \vec{r}\,' = t^2 + t^2 = 2t^2, \qquad W_1 = \int_0^1 2t^2 \,\mathrm{d}t = \frac{2}{3} \approx 0{,}6667.$$

Camino 2: la parábola $y = x^2$. Con $\vec{r}(t) = (t, t^2)$, $\vec{r}\,' = (1, 2t)$ y $\vec{F}(\vec{r}(t)) = (t^2,\ t\cdot t^2) = (t^2,\ t^3)$:

$$\vec{F} \cdot \vec{r}\,' = t^2 + 2t^4, \qquad W_2 = \int_0^1 (t^2 + 2t^4) \,\mathrm{d}t = \frac{1}{3} + \frac{2}{5} = \frac{11}{15} \approx 0{,}7333.$$

Mismos extremos, resultados distintos: $\frac{11}{15} \neq \frac{2}{3}$. El trabajo de este campo *sí* depende del camino, y esa es la situación normal. La subsección siguiente estudia la excepción, que es la interesante.

11.2.4. Independencia del camino, campos conservativos y función potencial

El ejemplo anterior deja una pregunta en el aire: ¿hay campos para los que el trabajo *no* dependa del camino? Los hay, son los importantes en física, y tienen una caracterización sencilla.

Definición 11.17 (Dominio, conexo y simplemente conexo)**.** Un *dominio* es un subconjunto abierto y conexo de $\mathbb{R}^n$. Es *conexo por caminos* si dos puntos cualesquiera se pueden unir mediante una curva contenida en él, es decir, si es "de una sola pieza". Es *simplemente conexo* si, además, todo lazo contenido en él se puede contraer con continuidad a un punto sin salirse: intuitivamente, si no tiene agujeros. El disco, el semiplano, el interior de un rectángulo y $\mathbb{R}^2$ entero son simplemente conexos; $\mathbb{R}^2$ menos un punto y una corona circular no lo son, porque un lazo que rodee el agujero se queda enganchado. La definición formal, con la homotopía escrita entera, está en la definición 11.108.

Definición 11.18 (Campo conservativo y función potencial)**.** Un campo vectorial $\vec{F}$ continuo en un dominio D se llama *conservativo* (o *de gradiente*) si existe un campo escalar $f : D \to \mathbb{R}$ de clase C^1 tal que $\vec{F} = \operatorname{grad} f$ en todo D, es decir, $P = \frac{\partial f}{\partial x}$, $Q = \frac{\partial f}{\partial y}$ y $R = \frac{\partial f}{\partial z}$. Esa f se llama *función potencial* de $\vec{F}$, y está determinada salvo una constante aditiva. En física se prefiere la energía potencial $U = -f$, y entonces $\vec{F} = -\operatorname{grad} U$; el signo es una convención, no matemática.

Se dice que la integral de $\vec{F}$ es *independiente del camino* en D si para cualesquiera dos puntos $\vec{A}, \vec{B} \in D$ y cualesquiera dos curvas $\mathcal{C}_1, \mathcal{C}_2$ contenidas en D que vayan de $\vec{A}$ a $\vec{B}$ se cumple $\int_{\mathcal{C}_1} \vec{F} \cdot \mathrm{d}\vec{r} = \int_{\mathcal{C}_2} \vec{F} \cdot \mathrm{d}\vec{r}$.

Teorema 11.19 (Fundamental de las integrales de línea)**.** *Sea f de clase C^1 en un dominio D y sea $\vec{F} = \operatorname{grad} f$. Si $\mathcal{C} \subset D$ es una curva C^1 a trozos que va de $\vec{A}$ a $\vec{B}$, entonces*

$$\int_{\mathcal{C}} \vec{F} \cdot \mathrm{d}\vec{r} = f(\vec{B}) - f(\vec{A}).$$

En particular, la integral solo depende de los extremos, y sobre toda curva cerrada *contenida en D vale cero.*

Demostración. Sea $\vec{r} : [a,b] \to D$ una parametrización de $\mathcal{C}$ con $\vec{r}(a) = \vec{A}$ y $\vec{r}(b) = \vec{B}$. Considérese la función de una sola variable $g(t) = f(\vec{r}(t))$. Por la regla de la cadena para funciones de varias variables (capítulo 9),

$$g'(t) = \frac{\partial f}{\partial x}\,x'(t) + \frac{\partial f}{\partial y}\,y'(t) + \frac{\partial f}{\partial z}\,z'(t) = \operatorname{grad} f(\vec{r}(t)) \cdot \vec{r}\,'(t) = \vec{F}(\vec{r}(t)) \cdot \vec{r}\,'(t).$$

El miembro de la derecha es exactamente el integrando de la definición 11.13. Por tanto, aplicando el teorema fundamental del cálculo de una variable a g, que es continua en $[a,b]$ y tiene derivada continua,

$$\int_{\mathcal{C}} \vec{F} \cdot \mathrm{d}\vec{r} = \int_a^b g'(t)\,\mathrm{d}t = g(b) - g(a) = f(\vec{B}) - f(\vec{A}).$$

Si la curva es C^1 solo a trozos, se aplica lo anterior en cada trozo y se suma: los valores intermedios de f aparecen dos veces con signos opuestos y se cancelan telescópicamente. Y si $\mathcal{C}$ es cerrada, $\vec{A} = \vec{B}$ y la diferencia es 0. □

Observación 11.20. Este teorema es literalmente el teorema fundamental del cálculo con una curva metida dentro: "la integral de una derivada es la diferencia de la primitiva entre los extremos". El gradiente hace de derivada, el potencial de primitiva y los dos extremos de la curva hacen de a y b. La consecuencia física es la conservación de la energía: en un campo conservativo, dar una vuelta completa y volver al punto de partida no cuesta trabajo neto, y por eso no existen las máquinas de movimiento perpetuo gravitatorias. El recíproco —independencia del camino implica existencia de potencial— se demuestra construyendo f como una integral desde un punto base fijo, y el enunciado completo con todas las equivalencias es el teorema 11.40, demostrado del todo en el corolario 11.103.

Falta la parte práctica: dado un campo, ¿cómo se sabe si es conservativo y cómo se encuentra su potencial? En el plano, el procedimiento cabe en seis pasos.

Regla: decidir si un campo del plano es conservativo y hallar su potencial

1. **Escribe las componentes.** $\vec{F} = (P, Q)$ con $P = P(x,y)$ y $Q = Q(x,y)$.
2. **Aplica el test de las cruzadas.** Calcula $\frac{\partial P}{\partial y}$ y $\frac{\partial Q}{\partial x}$. Si son distintas, el campo **no** es conservativo y el problema termina aquí. Si coinciden, sigue, pero sin cantar victoria todavía.
3. **Mira el dominio.** El test del paso 2 solo garantiza la existencia de potencial si el dominio es *simplemente conexo*. Si al campo hay que quitarle algún punto porque un denominador se anula, enciende la alarma: ahí es donde el criterio falla.
4. **Integra P respecto de x.** Escribe $f(x,y) = \int P\,\mathrm{d}x$ tratando y como una constante. La constante de integración no es un número: es una función $g(y)$.
5. **Deriva respecto de y e iguala a Q.** De $\frac{\partial f}{\partial y} = Q$ despeja $g'(y)$; todo lo que dependa de x debe cancelarse solo, y si no lo hace es que el paso 2 estaba mal. Integra g' y toma la constante final igual a cero.

6. **Verifica.** Calcula grad f y comprueba que reproduce $\vec{F}$ componente a componente. Cuesta treinta segundos y detecta cualquier error acumulado.

La versión para $\mathbb{R}^3$ es la regla sobre campos conservativos de la sección de integrales de línea: la mecánica es idéntica, solo que la "constante" del paso 4 es una función $g(y,z)$ y hace falta una tercera vuelta imponiendo $\frac{\partial f}{\partial z} = R$.

Ejemplo 11.21 (Un potencial en el plano, y la comprobación por tres caminos)**.** Sea $\vec{F} = (3x^2y + y^3,\ x^3 + 3xy^2)$ en todo $\mathbb{R}^2$.

Paso 2, el test. Con $P = 3x^2y+y^3$ y $Q = x^3+3xy^2$ resulta $\frac{\partial P}{\partial y} = 3x^2+3y^2$ y $\frac{\partial Q}{\partial x} = 3x^2+3y^2$: coinciden en todo punto.

Paso 3, el dominio. Es $\mathbb{R}^2$ entero, simplemente conexo. Luego el potencial existe.

Paso 4. Integramos P respecto de x, con y congelada:

$$f(x,y) = \int (3x^2y + y^3)\,\mathrm{d}x = x^3y + xy^3 + g(y).$$

Paso 5. Derivamos respecto de y e igualamos a Q:

$$\frac{\partial f}{\partial y} = x^3 + 3xy^2 + g'(y) \overset{!}{=} x^3 + 3xy^2 \implies g'(y) = 0 \implies g = \text{cte} = 0.$$

Nótese cómo los términos en x se han cancelado solos: es la señal de que el test del paso 2 era correcto. El potencial es

$$\boxed{f(x,y) = x^3y + xy^3}.$$

Paso 6, verificación. $\frac{\partial f}{\partial x} = 3x^2y + y^3$ ✓ y $\frac{\partial f}{\partial y} = x^3 + 3xy^2$ ✓.

El pago. El trabajo de $\vec{F}$ desde $(0,0)$ hasta $(1,2)$ vale, por el teorema 11.19, $f(1,2)-f(0,0) = (1\cdot 2 + 1\cdot 8) - 0 = 10$, *por cualquier camino.* Comprobémoslo por tres:

- *El segmento.* $\vec{r}(t) = (t, 2t)$ con $\vec{r}\,' = (1,2)$. Entonces $P = 6t^3+8t^3 = 14t^3$ y $Q = t^3+12t^3 = 13t^3$, luego $\vec{F}\cdot\vec{r}\,' = 14t^3 + 26t^3 = 40t^3$ e $\int_0^1 40t^3\,\mathrm{d}t = 10$. ✓
- *El camino en escuadra* $(0,0)\to(1,0)\to(1,2)$. En el tramo horizontal $y = 0$ y $\mathrm{d}y = 0$, y además $P(t,0) = 0$: no aporta nada. En el vertical $x = 1$ y $\mathrm{d}x = 0$, de modo que solo cuenta $Q(1,y) = 1 + 3y^2$, y $\int_0^2 (1+3y^2)\,\mathrm{d}y = 2 + 8 = 10$. ✓
- *La parábola* $y = 2x^2$, es decir $\vec{r}(t) = (t, 2t^2)$ con $\vec{r}\,' = (1, 4t)$. Ahora $P = 6t^4 + 8t^6$ y $Q = t^3+12t^5$, luego $\vec{F}\cdot\vec{r}\,' = (6t^4+8t^6)+4t(t^3+12t^5) = 10t^4+56t^6$, e $\int_0^1 (10t^4+56t^6)\,\mathrm{d}t = \frac{10}{5} + \frac{56}{7} = 2 + 8 = 10$. ✓

Tres caminos completamente distintos, el mismo número. Eso *es* la independencia del camino, y el ahorro es evidente: con el potencial en la mano, la tercera integral no había ni que plantearla.

Proposición 11.22 (La condición necesaria)**.** *Si $\vec{F} = (P,Q)$ es conservativo y de clase C^1 en un dominio D del plano, entonces $\frac{\partial P}{\partial y} = \frac{\partial Q}{\partial x}$ en todo D. En el espacio, la condición análoga es* $\operatorname{rot}\vec{F} = \vec{0}$.

Demostración. Si $\vec{F} = \operatorname{grad} f$, entonces $P = f_x$ y $Q = f_y$, luego $P_y = f_{xy}$ y $Q_x = f_{yx}$. Como $\vec{F}$ es de clase C^1, la función f es de clase C^2, y el teorema de Clairaut–Schwarz sobre la igualdad de las derivadas cruzadas da $f_{xy} = f_{yx}$. La versión espacial es idéntica, componente a componente, y quedará demostrada en la proposición 11.30. □

Cuidado

[La condición es necesaria pero no suficiente: el contraejemplo obligatorio] El paso 3 de la regla no es una prudencia retórica. Considérese el campo

$$\vec{F}(x,y) = \left(\frac{-y}{x^2+y^2},\ \frac{x}{x^2+y^2}\right), \qquad \text{definido en } \Omega = \mathbb{R}^2 \setminus \{(0,0)\}.$$

Escribiendo $D = x^2 + y^2$, las derivadas cruzadas son

$$\frac{\partial P}{\partial y} = \frac{-D + y\cdot 2y}{D^2} = \frac{y^2 - x^2}{D^2}, \qquad \frac{\partial Q}{\partial x} = \frac{D - x\cdot 2x}{D^2} = \frac{y^2 - x^2}{D^2},$$

iguales en *todo* punto de Ω. El campo pasa el test. Y sin embargo no es conservativo: su circulación por la circunferencia unidad, parametrizada por $\vec{r}(t) = (\cos t, \sin t)$, vale

$$\oint \vec{F}\cdot \mathrm{d}\vec{r} = \int_0^{2\pi} (-\sin t,\ \cos t)\cdot(-\sin t,\ \cos t)\,\mathrm{d}t = \int_0^{2\pi} (\sin^2 t + \cos^2 t)\,\mathrm{d}t = 2\pi \neq 0,$$

mientras que el teorema 11.19 obliga a que la circulación de un campo conservativo por cualquier curva cerrada sea 0.

¿Dónde está la trampa? En el dominio. Al quitar el origen, Ω deja de ser simplemente conexo: la circunferencia unidad no se puede contraer a un punto sin cruzar el agujero. Localmente sí hay potencial —en el semiplano $x > 0$ funciona el ángulo $\varphi = \arctan(y/x)$—, pero el ángulo no se puede definir de manera continua dando la vuelta entera: al completar el lazo ha crecido en 2π en vez de volver a su valor. La condición de las cruzadas es *local*; ser conservativo es *global*, y entre ambas cosas se cuela la topología del dominio. El estudio completo de este campo, con lazos que dan varias vueltas y dominios con varios agujeros, está en el ejemplo del campo del ángulo de la sección 11.7.

11.2.5. Superficies parametrizadas, vector normal y las dos integrales de superficie

Todo lo anterior se repite un piso más arriba cambiando curvas por superficies. Una curva se describía con un parámetro; una superficie necesita dos, porque tiene dos direcciones independientes en las que moverse.

Definición 11.23 (Superficie parametrizada, vectores coordenados y vector normal)**.** Sea $D \subset \mathbb{R}^2$ una región. Una *superficie parametrizada* es una aplicación de clase C^1

$$\vec{r}: D \longrightarrow \mathbb{R}^3, \qquad \vec{r}(u,v) = \big(x(u,v),\ y(u,v),\ z(u,v)\big),$$

cuya imagen S es el trazo de la superficie. Los *vectores coordenados* son las derivadas parciales

$$\vec{r}_u = \frac{\partial \vec{r}}{\partial u} = \Big(\frac{\partial x}{\partial u}, \frac{\partial y}{\partial u}, \frac{\partial z}{\partial u}\Big), \qquad \vec{r}_v = \frac{\partial \vec{r}}{\partial v} = \Big(\frac{\partial x}{\partial v}, \frac{\partial y}{\partial v}, \frac{\partial z}{\partial v}\Big);$$

son tangentes a S, porque $\vec{r}_u$ es el vector velocidad de la curva que se obtiene congelando v y moviendo u, y viceversa. La parametrización es *regular* si $\vec{r}_u \times \vec{r}_v \neq \vec{0}$ en todo punto; entonces esos dos vectores generan el *plano tangente* y el producto vectorial $\vec{N} = \vec{r}_u \times \vec{r}_v$ es un *vector normal* a S. El *elemento de área* es

$$\boxed{\mathrm{d}S = \|\vec{r}_u \times \vec{r}_v\| \,\mathrm{d}u\,\mathrm{d}v}, \qquad \text{y el elemento de área orientado es} \qquad \mathrm{d}\vec{S} = \vec{N}\,\mathrm{d}u\,\mathrm{d}v = \vec{n}\,\mathrm{d}S,$$

con $\vec{n} = \vec{N}/\|\vec{N}\|$. Elegir de manera continua uno de los dos normales unitarios posibles ($\vec{n}$ o $-\vec{n}$) es *orientar* la superficie; la definición precisa, y el motivo de que no siempre sea posible, están en la definición 11.104.

De dónde sale el elemento de área: un rectangulito de lados $\mathrm{d}u$ y $\mathrm{d}v$ en el dominio de parámetros se transforma, en primer orden, en el paralelogramo generado por $\vec{r}_u\,\mathrm{d}u$ y $\vec{r}_v\,\mathrm{d}v$, y el área de un paralelogramo es la norma del producto vectorial de sus lados. El módulo $\|\vec{r}_u \times \vec{r}_v\|$ es, pues, el factor por el que la parametrización estira el área, exactamente igual que $\|\vec{r}\,'(t)\|$ lo era para longitudes.

Definición 11.24 (Las dos integrales de superficie)**.** Sea S una superficie parametrizada regular por $\vec{r} : D \to \mathbb{R}^3$.

- Si f es un campo escalar continuo sobre S, su *integral de superficie* es

$$\iint_S f\,\mathrm{d}S := \iint_D f(\vec{r}(u,v))\,\|\vec{r}_u \times \vec{r}_v\|\,\mathrm{d}u\,\mathrm{d}v,$$

  que con $f \equiv 1$ devuelve el *área* de S y con f igual a una densidad superficial devuelve la *masa* de una chapa.

- Si $\vec{F}$ es un campo vectorial continuo sobre S y S está *orientada*, el *flujo* de $\vec{F}$ a través de S es

$$\iint_S \vec{F}\cdot\mathrm{d}\vec{S} := \iint_S \vec{F}\cdot\vec{n}\,\mathrm{d}S = \iint_D \vec{F}(\vec{r}(u,v))\cdot(\vec{r}_u \times \vec{r}_v)\,\mathrm{d}u\,\mathrm{d}v.$$

  Mide cuánto atraviesa el campo la superficie por unidad de tiempo, y cambia de signo al invertir la orientación.

Cuidado

[$\mathrm{d}S$ frente a $\mathrm{d}\vec{S}$: la misma trampa un piso más arriba] Vuelve a haber dos objetos que se parecen y no lo son. En $\iint_S f\,\mathrm{d}S$ el factor $\|\vec{r}_u \times \vec{r}_v\|$ es un número positivo y el resultado no depende de la orientación: el área de una chapa es la misma se mire por delante o por detrás. En $\iint_S \vec{F}\cdot\mathrm{d}\vec{S}$ el factor es el vector $\vec{r}_u \times \vec{r}_v$ entero, con su dirección, y al elegir la normal opuesta el flujo cambia de signo, porque entrar y salir tienen que contar distinto.

La correspondencia con las integrales de línea es exacta:

$$\mathrm{d}s \;\longleftrightarrow\; \mathrm{d}S \quad \text{(escalares, sin signo)}, \qquad \mathrm{d}\vec{r} \;\longleftrightarrow\; \mathrm{d}\vec{S} \quad \text{(vectoriales, con signo)}.$$

Y un aviso de cálculo: al montar el flujo, **no** normalices $\vec{r}_u \times \vec{r}_v$. Su módulo es justamente el factor de área, y si lo divides tendrás que volver a multiplicarlo. Escribe $\vec{F} \cdot (\vec{r}_u \times \vec{r}_v)$ y punto.

Ejemplo 11.25 (Integral de superficie de un campo escalar: la altura media de una media esfera)**.** Calculemos $\iint_S z \,\mathrm{d}S$, donde S es el hemisferio superior de la esfera de radio a centrada en el origen.

Paso 1, parametrizar. En coordenadas esféricas, con φ el ángulo desde el eje z y θ el ángulo alrededor de él,

$$\vec{r}(\varphi,\theta) = (a\sin\varphi\cos\theta,\ a\sin\varphi\sin\theta,\ a\cos\varphi), \qquad \varphi \in [0, \tfrac{\pi}{2}],\ \theta \in [0, 2\pi].$$

Paso 2, los vectores coordenados y la normal.

$$\vec{r}_\varphi = (a\cos\varphi\cos\theta,\ a\cos\varphi\sin\theta,\ -a\sin\varphi), \qquad \vec{r}_\theta = (-a\sin\varphi\sin\theta,\ a\sin\varphi\cos\theta,\ 0),$$

y el producto vectorial resulta ser

$$\vec{r}_\varphi \times \vec{r}_\theta = a\sin\varphi\,(a\sin\varphi\cos\theta,\ a\sin\varphi\sin\theta,\ a\cos\varphi) = a\sin\varphi\ \vec{r}(\varphi,\theta),$$

un vector radial, como tenía que ser en una esfera. Su norma es $a\sin\varphi \cdot a = a^2\sin\varphi$ (positiva porque $\sin\varphi \geq 0$ en $[0,\pi]$), de modo que $\mathrm{d}S = a^2\sin\varphi\,\mathrm{d}\varphi\,\mathrm{d}\theta$.

Control previo. Con $f \equiv 1$ el área del hemisferio sale $\int_0^{2\pi}\int_0^{\pi/2} a^2\sin\varphi\,\mathrm{d}\varphi\,\mathrm{d}\theta = 2\pi a^2$, que es la mitad de $4\pi a^2$. ✓

Paso 3, sustituir e integrar. Sobre la superficie, $z = a\cos\varphi$, luego

$$\iint_S z\,\mathrm{d}S = \int_0^{2\pi}\int_0^{\pi/2} a\cos\varphi\cdot a^2\sin\varphi\,\mathrm{d}\varphi\,\mathrm{d}\theta = 2\pi a^3\left[\frac{\sin^2\varphi}{2}\right]_0^{\pi/2} = \pi a^3.$$

Para $a = 2$ vale $8\pi \approx 25{,}13$.

Lectura. Dividiendo por el área, la altura media de un casquete hemisférico es $\pi a^3/(2\pi a^2) = a/2$: exactamente la mitad del radio. La orientación no ha intervenido en ningún momento, porque esta integral no la necesita.

Ejemplo 11.26 (Flujo a través de un cilindro, con comprobación independiente)**.** Calculemos el flujo del campo $\vec{F} = (x,\, y,\, 0)$ a través de la superficie lateral del cilindro $x^2 + y^2 = a^2$ con $0 \leq z \leq h$, orientada con la normal que apunta hacia *fuera* del eje.

Paso 1, parametrizar. Los parámetros naturales son el ángulo y la altura: $\vec{r}(\theta, z) = (a\cos\theta,\ a\sin\theta,\ z)$ con $\theta \in [0, 2\pi]$ y $z \in [0, h]$.

Paso 2, la normal. Con $\vec{r}_\theta = (-a\sin\theta,\ a\cos\theta,\ 0)$ y $\vec{r}_z = (0, 0, 1)$,

$$\vec{N} = \vec{r}_\theta \times \vec{r}_z = (a\cos\theta\cdot 1 - 0,\quad 0 - (-a\sin\theta)\cdot 1,\quad 0) = (a\cos\theta,\ a\sin\theta,\ 0).$$

Paso 3, comprobar la orientación. En $\theta = 0$ el punto de la superficie es $(a, 0, z)$ y la normal vale $(a, 0, 0)$: apunta alejándose del eje, que es lo que pedía el enunciado. No hay que cambiar ningún signo; si hubiera apuntado hacia dentro, bastaría con usar $-\vec{N}$.

Paso 4, sustituir. Sobre la superficie, $\vec{F}(\vec{r}(\theta, z)) = (a\cos\theta,\ a\sin\theta,\ 0)$, de modo que

$$\vec{F} \cdot \vec{N} = a^2\cos^2\theta + a^2\sin^2\theta = a^2,$$

constante. Tiene sentido: el campo es radial y la normal también, así que el producto escalar es el producto de los módulos, que valen ambos a.

Paso 5, integrar. $\displaystyle\iint_S \vec{F}\cdot \mathrm{d}\vec{S} = \int_0^h\int_0^{2\pi} a^2\,\mathrm{d}\theta\,\mathrm{d}z = 2\pi a^2 h$. Con $a = 1$ y $h = 3$ da $6\pi \approx 18{,}85$, positivo, como corresponde a un campo que sale por todas partes.

Comprobación por otro camino. Cerremos el cilindro con sus dos tapas. En la tapa superior la normal exterior es $(0, 0, 1)$ y $\vec{F}\cdot(0, 0, 1) = 0$, porque la tercera componente de $\vec{F}$ es nula; en la inferior, lo mismo con $(0, 0, -1)$. Las tapas no aportan nada. Por el teorema de la divergencia (teorema 11.51), el flujo total por la superficie cerrada vale

$$\iiint_E \operatorname{div}\vec{F}\,\mathrm{d}V = \iiint_E 2\,\mathrm{d}V = 2\cdot\pi a^2 h = 2\pi a^2 h,$$

que coincide con lo obtenido a mano por el lateral. ✓ Dos métodos independientes que aterrizan en el mismo número.

11.2.6. Los cuatro operadores diferenciales

Falta la última pieza elemental: las operaciones que fabrican campos nuevos a partir de campos viejos. Son cuatro, y lo primero que hay que fijar de cada una es *qué come y qué devuelve*; esa contabilidad de tipos descarta por sí sola la mayoría de los disparates.

Definición 11.27 (Gradiente, divergencia, rotacional y laplaciano)**.** Sea f un campo escalar de clase C^1 y $\vec{F} = (P, Q, R)$ un campo vectorial de clase C^1 en un abierto de $\mathbb{R}^3$.

- El *gradiente* de f es el campo **vectorial** $\operatorname{grad} f = \left(\frac{\partial f}{\partial x},\ \frac{\partial f}{\partial y},\ \frac{\partial f}{\partial z}\right)$.
- La *divergencia* de $\vec{F}$ es el campo **escalar** $\operatorname{div}\vec{F} = \frac{\partial P}{\partial x} + \frac{\partial Q}{\partial y} + \frac{\partial R}{\partial z}$.
- El *rotacional* de $\vec{F}$ es el campo **vectorial**

$$\operatorname{rot}\vec{F} = \Big(\frac{\partial R}{\partial y} - \frac{\partial Q}{\partial z},\ \ \frac{\partial P}{\partial z} - \frac{\partial R}{\partial x},\ \ \frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y}\Big).$$

- El *laplaciano* de f (que ha de ser de clase C^2) es el campo **escalar**

$$\Delta f = \operatorname{div}(\operatorname{grad} f) = \frac{\partial^2 f}{\partial x^2} + \frac{\partial^2 f}{\partial y^2} + \frac{\partial^2 f}{\partial z^2}.$$

En el plano se usan las mismas fórmulas suprimiendo la tercera coordenada; el rotacional se reduce entonces al único número $\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y}$, llamado *rotacional escalar*.

Qué mide cada uno, en palabras. El **gradiente** apunta en la dirección de máximo crecimiento de f, y su módulo es ese ritmo de crecimiento; es siempre perpendicular a las curvas (o superficies) de nivel, porque a lo largo de un nivel la función no cambia. Si f es la altura del terreno, $\operatorname{grad} f$ marca la línea de máxima pendiente. La **divergencia** mide en cada punto cuánto sale del campo por unidad de volumen: es positiva en las *fuentes*, negativa en los *sumideros* y nula donde el fluido solo pasa de largo (campo *incompresible* o solenoidal). El **rotacional** mide cuánto gira el campo alrededor de cada punto: si se coloca un molinillo diminuto en el fluido, $\operatorname{rot} \vec{F}$ apunta en su eje de giro y su módulo es el doble de la velocidad angular; un campo con $\operatorname{rot} \vec{F} = \vec{0}$ se llama *irrotacional*. El **laplaciano** compara el valor de f en un punto con el promedio de sus vecinos: si es cero (función *armónica*), cada punto vale exactamente la media de su entorno, que es la firma del equilibrio térmico o electrostático.

Observación 11.28 (El operador ∇ y sus abusos). Es cómodo escribir los operadores anteriores mediante el símbolo $\nabla = \left(\frac{\partial}{\partial x}, \frac{\partial}{\partial y}, \frac{\partial}{\partial z}\right)$, tratándolo como si fuera un vector cuyas componentes son órdenes de derivar. Entonces

$$\operatorname{grad} f = \nabla f, \qquad \operatorname{div} \vec{F} = \nabla \cdot \vec{F}, \qquad \operatorname{rot} \vec{F} = \nabla \times \vec{F}, \qquad \Delta f = \nabla \cdot \nabla f = \nabla^2 f,$$

y el rotacional se recuerda con el determinante simbólico

$$\operatorname{rot} \vec{F} = \begin{vmatrix} \vec{i} & \vec{j} & \vec{k} \\ \frac{\partial}{\partial x} & \frac{\partial}{\partial y} & \frac{\partial}{\partial z} \\ P & Q & R \end{vmatrix}.$$

Cuidado

[∇ no es un vector de verdad] Es una regla mnemotécnica excelente y una fuente inagotable de errores. El símbolo ∇ no tiene componentes numéricas: tiene *operaciones*, y las operaciones no conmutan con lo que hay a su derecha. Consecuencias prácticas:

- $\nabla \cdot \vec{F}$ y $\vec{F} \cdot \nabla$ no son lo mismo. El primero es un número (la divergencia); el segundo es un *operador* que todavía espera algo que derivar.
- ∇f no se puede "simplificar" como si ∇ fuese un factor que se cancela.
- Las identidades del álgebra vectorial no se trasladan sin más. Por ejemplo $\vec{a} \times \vec{a} = \vec{0}$ para todo vector, y en efecto $\nabla \times \nabla f = \vec{0}$; pero eso no es por antisimetría, sino por la igualdad de las derivadas cruzadas, y hay que demostrarlo.
- En coordenadas curvilíneas el símbolo deja de funcionar directamente, porque los vectores de la base cambian de dirección de punto a punto y hay que derivarlos también.

Úsese ∇ como recordatorio de la fórmula, nunca como argumento de una demostración.

Regla: aplicar los cuatro operadores sin equivocarse de tipo

1. **Mira qué te dan.** Si es una función escalar, el único operador de primer orden aplicable es el gradiente. Si es un campo vectorial, puedes aplicarle la divergencia o el rotacional. A este nivel no existen "el gradiente de un campo vectorial" ni "la divergencia de una función".
2. **Gradiente: deriva y apila.** Una parcial por coordenada, en orden. Sale un vector.
3. **Divergencia: cada componente con su propia variable, y suma.** P se deriva respecto de x, Q respecto de y y R respecto de z; nunca se cruzan. Sale un escalar.
4. **Rotacional: monta el determinante** de la observación 11.28 y desarrolla por la primera fila. Atención al signo de la componente central, que lleva un menos delante y por eso queda $P_z - R_x$ y no $R_x - P_z$: es el error más frecuente del tema. Sale un vector.
5. **Laplaciano: no lo memorices, compónlo** como $\operatorname{div}(\operatorname{grad} f)$. Así nunca fallará ni el número de términos ni el signo.
6. **Comprueba el tipo del resultado.** Gradiente y rotacional deben darte tres números; divergencia y laplaciano, uno solo. Si sale un vector donde esperabas un escalar, has mezclado operadores.

Los controles finos de cálculo —qué verificar para descartar errores de signo— están en la regla sobre gradiente, divergencia y rotacional de la sección siguiente.

Ejemplo 11.29 (Los cuatro operadores sobre un mismo par de campos)**.** Sean $f(x,y,z) = x^2y + yz^2$ y $\vec{G}(x,y,z) = (x^2z,\ xy^2,\ yz^2)$.

Gradiente de f. Una parcial por coordenada: $\frac{\partial f}{\partial x} = 2xy$, $\frac{\partial f}{\partial y} = x^2 + z^2$ y $\frac{\partial f}{\partial z} = 2yz$, luego $\operatorname{grad} f = (2xy,\ x^2+z^2,\ 2yz)$. En el punto $(1,1,1)$ vale $(2,2,2)$: allí f crece más deprisa en la dirección de la diagonal, a un ritmo $\|(2,2,2)\| = 2\sqrt{3} \approx 3{,}46$ por unidad de longitud.

Laplaciano de f. Componemos: $\Delta f = \operatorname{div}(2xy,\ x^2+z^2,\ 2yz)$, es decir

$$\Delta f = \frac{\partial}{\partial x}(2xy) + \frac{\partial}{\partial y}(x^2+z^2) + \frac{\partial}{\partial z}(2yz) = 2y + 0 + 2y = 4y.$$

No es armónica: en el semiespacio $y > 0$ el laplaciano es positivo y cada punto vale menos que el promedio de sus vecinos.

Divergencia de $\vec{G}$. Cada componente con su variable:

$$\operatorname{div}\vec{G} = \frac{\partial}{\partial x}(x^2z) + \frac{\partial}{\partial y}(xy^2) + \frac{\partial}{\partial z}(yz^2) = 2xz + 2xy + 2yz.$$

Es un escalar. En $(1,1,1)$ vale $6 > 0$: allí hay una fuente.

Rotacional de $\vec{G}$. Con $P = x^2z$, $Q = xy^2$ y $R = yz^2$:

$$\frac{\partial R}{\partial y} - \frac{\partial Q}{\partial z} = z^2 - 0 = z^2, \qquad \frac{\partial P}{\partial z} - \frac{\partial R}{\partial x} = x^2 - 0 = x^2, \qquad \frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y} = y^2 - 0 = y^2,$$

luego $\operatorname{rot}\vec{G} = (z^2,\ x^2,\ y^2)$, que en $(1,1,1)$ vale $(1,1,1)$.

Controles. Gradiente y rotacional han dado ternas; divergencia y laplaciano, números sueltos. ✓ Y dos controles de fondo, que la proposición siguiente demuestra en general: $\operatorname{rot}(\operatorname{grad} f)$ debe ser $\vec{0}$ y $\operatorname{div}(\operatorname{rot}\vec{G})$ debe ser 0. En efecto, la primera componente de $\operatorname{rot}(\operatorname{grad} f)$ es $\frac{\partial}{\partial y}(2yz) - \frac{\partial}{\partial z}(x^2+z^2) = 2z - 2z = 0$, y las otras dos también se anulan; y $\operatorname{div}(z^2, x^2, y^2) = \frac{\partial}{\partial x}(z^2) + \frac{\partial}{\partial y}(x^2) + \frac{\partial}{\partial z}(y^2) = 0+0+0 = 0$. ✓

Proposición 11.30 (Las dos identidades básicas)**.** *Para todo campo escalar f de clase C^2 y todo campo vectorial $\vec{F} = (P,Q,R)$ de clase C^2:*

$$\operatorname{rot}(\operatorname{grad} f) = \vec{0}, \qquad \operatorname{div}(\operatorname{rot}\vec{F}) = 0.$$

Demostración. Ambas son cálculo directo, y en las dos la clave es la misma: el teorema de Clairaut–Schwarz, que para funciones de clase C^2 garantiza $\partial_i\partial_j = \partial_j\partial_i$.

Primera identidad. Aplicando la definición de rotacional al campo $\operatorname{grad} f = (f_x,\ f_y,\ f_z)$, sus tres componentes son

$$\frac{\partial}{\partial y}(f_z) - \frac{\partial}{\partial z}(f_y) = f_{zy} - f_{yz}, \qquad \frac{\partial}{\partial z}(f_x) - \frac{\partial}{\partial x}(f_z) = f_{xz} - f_{zx},$$
$$\frac{\partial}{\partial x}(f_y) - \frac{\partial}{\partial y}(f_x) = f_{yx} - f_{xy}.$$

Cada una es la diferencia de las dos derivadas cruzadas de un mismo par de variables, y por Clairaut las tres se anulan. Luego $\operatorname{rot}(\operatorname{grad} f) = \vec{0}$.

Segunda identidad. El rotacional de $\vec{F}$ es $(R_y - Q_z,\ P_z - R_x,\ Q_x - P_y)$, y su divergencia es

$$\begin{aligned}\operatorname{div}(\operatorname{rot}\vec{F}) &= \frac{\partial}{\partial x}(R_y - Q_z) + \frac{\partial}{\partial y}(P_z - R_x) + \frac{\partial}{\partial z}(Q_x - P_y)\\ &= \underbrace{R_{yx} - R_{xy}}_{0} + \underbrace{P_{zy} - P_{yz}}_{0} + \underbrace{Q_{xz} - Q_{zx}}_{0} = 0,\end{aligned}$$

donde se han agrupado los seis términos por parejas, una por cada letra: cada pareja consta de las dos derivadas cruzadas de una misma componente y se cancela por Clairaut. □

Idea clave

Estas dos identidades no son curiosidades de cálculo: son el motivo de que existan los potenciales. La primera dice que *ningún gradiente rota*, y por eso $\operatorname{rot}\vec{F} = \vec{0}$ es condición necesaria para que $\vec{F}$ tenga potencial escalar (proposición 11.22). La segunda dice que *ningún rotacional tiene fuentes*, y por eso $\operatorname{div}\vec{B} = 0$ es condición necesaria para que $\vec{B}$ tenga potencial vectorial. Ambas se resumen en un eslogan: derivar dos veces seguidas en el sentido del cálculo vectorial siempre da cero, porque las derivadas cruzadas coinciden. En el lenguaje de las formas diferenciales las dos son la misma ecuación $\mathrm{d}\mathrm{d} = 0$, y de ahí arranca la topología algebraica. Las identidades restantes —$\operatorname{div}(f\vec{F})$ y $\operatorname{rot}(\operatorname{rot}\vec{F})$— y su uso como filtro para descartar candidatos están en la proposición 11.31.

Pausa para pensar

Repasa mentalmente el mapa antes de seguir. Un campo escalar y un campo vectorial son objetos distintos; una curva se parametriza con un parámetro y una superficie con dos; sobre cada uno de esos objetos hay una integral *sin* orientación (masa, área) y otra *con* orientación (trabajo, flujo); y los cuatro operadores relacionan campos escalares con campos vectoriales en un sentido o en el otro. Con ese mapa en la cabeza, los tres grandes teoremas que vienen después dejan de ser tres fórmulas que memorizar y pasan a ser tres afirmaciones sobre cómo lo que ocurre dentro de una región se lee entero en su borde.

11.3. Campos vectoriales y operadores diferenciales

Idea clave

Intuición primero. Un campo vectorial pone una flecha en cada punto del espacio: la velocidad del agua en cada punto de un río, la fuerza en cada punto de un campo eléctrico, el viento en cada punto del mapa. Los tres operadores de esta sección son tres preguntas distintas sobre ese campo, y cada una tiene una lectura física directa.
El *gradiente* se aplica a una función escalar y produce el campo que apunta en la dirección de máximo crecimiento, con longitud igual a ese ritmo de crecimiento; si la función es la altura del terreno, el gradiente es la dirección de máxima pendiente y es siempre perpendicular a las curvas de nivel. La *divergencia* se aplica a un campo y mide cuánto sale de cada punto: es positiva donde hay una fuente, negativa donde hay un sumidero, cero donde el fluido solo pasa. El *rotacional* mide cuánto gira el campo alrededor de cada punto: si pusieras una ruedita diminuta en el fluido, el rotacional dice a qué velocidad giraría y alrededor de qué eje. Con esas tres imágenes, las identidades $\operatorname{rot}(\operatorname{grad} f) = 0$ y $\operatorname{div}(\operatorname{rot} F) = 0$ dejan de ser cálculos y pasan a ser afirmaciones creíbles: un campo de máxima pendiente no puede girar, y un campo de rotación pura no tiene fuentes.

Un *campo vectorial* asigna un vector a cada punto: $\mathbf{F} = (P, Q, R)$. Con el operador $\nabla = (\partial_x, \partial_y, \partial_z)$:

$$\begin{aligned}
\operatorname{grad} f = \nabla f &= (f_x, f_y, f_z) && \text{escalar} \to \text{vector}\\
\operatorname{div} \mathbf{F} = \nabla \cdot \mathbf{F} &= P_x + Q_y + R_z && \text{vector} \to \text{escalar}\\
\operatorname{rot} \mathbf{F} = \nabla \times \mathbf{F} &= (R_y - Q_z,\ P_z - R_x,\ Q_x - P_y) && \text{vector} \to \text{vector}\\
\Delta f = \nabla \cdot \nabla f &= f_{xx} + f_{yy} + f_{zz} && \text{laplaciano.}
\end{aligned}$$

Interpretaciones: si $\mathbf{F}$ es el campo de velocidades de un fluido, $\operatorname{div} \mathbf{F}(\mathbf{p})$ mide el flujo neto *saliente* por unidad de volumen en $\mathbf{p}$ (fuentes > 0, sumideros < 0; $\operatorname{div} \mathbf{F} = 0$ es un fluido *incompresible*); $\operatorname{rot} \mathbf{F}(\mathbf{p})$ apunta en el eje de rotación local, con módulo el doble de la velocidad angular (un molinillo en $\mathbf{p}$ gira si $\operatorname{rot} \neq \mathbf{0}$; $\operatorname{rot} \mathbf{F} = \mathbf{0}$ es un campo *irrotacional*).

Proposición 11.31 (Identidades fundamentales)**.** *Para campos C^2:*

$$\operatorname{rot}(\nabla f) = \mathbf{0}, \qquad \operatorname{div}(\operatorname{rot}\mathbf{F}) = 0,$$
$$\operatorname{div}(f\mathbf{F}) = f\operatorname{div}\mathbf{F} + \nabla f \cdot \mathbf{F}, \qquad \operatorname{rot}(\operatorname{rot}\mathbf{F}) = \nabla(\operatorname{div}\mathbf{F}) - \Delta\mathbf{F}.$$

Demostración de las dos primeras. Primera componente de $\operatorname{rot}(\nabla f)$: $\partial_y f_z - \partial_z f_y = 0$ por Clairaut. Y $\operatorname{div}(\operatorname{rot}\mathbf{F}) = \partial_x(R_y - Q_z) + \partial_y(P_z - R_x) + \partial_z(Q_x - P_y) = 0$: los seis términos se cancelan por pares, otra vez por igualdad de cruzadas. □

Idea clave

Qué dice esto en cristiano. Las dos identidades $\operatorname{rot}(\nabla f) = \mathbf{0}$ y $\operatorname{div}(\operatorname{rot}\mathbf{F}) = 0$ dicen lo mismo desde dos pisos distintos: derivar dos veces seguidas en el sentido del cálculo vectorial siempre da cero. La razón profunda es la igualdad de las derivadas cruzadas, $\partial_x\partial_y = \partial_y\partial_x$, que hace que todos los términos se cancelen por parejas. En el lenguaje de formas diferenciales las dos identidades son la misma, $\mathrm{dd} = 0$, y esa única ecuación es el punto de partida de la topología algebraica.

Ejemplo 11.32 (Comprobar las identidades y usarlas como filtro)**. (a) Un gradiente no rota.** Para $f = x^2y + z$ se tiene $\nabla f = (2xy,\ x^2,\ 1)$, y su rotacional es $(0-0,\ 0-0,\ 2x-2x) = \mathbf{0}$. Cada componente es una diferencia de derivadas cruzadas, y por eso se anula.

(b) Un rotacional no diverge. Para $\mathbf{F} = (xy,\ yz,\ zx)$ el rotacional es $(-y,\ -z,\ -x)$, y su divergencia es $0 + 0 + 0 = 0$. Otra vez todo se cancela.

(c) El uso práctico: descartar candidatos. Si alguien propone que $\mathbf{F} = (y,\ -x,\ 0)$ es un gradiente, basta calcular su rotacional: sale $(0, 0, -2) \neq \mathbf{0}$, luego no lo es y no hace falta buscar el potencial. De igual modo, un campo con divergencia no nula no puede ser el rotacional de nada.

(d) El recíproco necesita hipótesis sobre el dominio. Que $\operatorname{rot}\mathbf{F} = \mathbf{0}$ no basta para que $\mathbf{F}$ sea un gradiente: hace falta que el dominio no tenga agujeros. El contraejemplo clásico es $\mathbf{F} = \left(\frac{-y}{x^2+y^2}, \frac{x}{x^2+y^2}\right)$ en $\mathbb{R}^2 \setminus \{\mathbf{0}\}$, cuyo rotacional es cero y cuya circulación por el círculo unidad vale $2\pi \neq 0$. Si fuera un gradiente, esa circulación cerrada tendría que ser cero.

Cuidado

[Las identidades son necesarias, nunca suficientes] Que el rotacional se anule es condición *necesaria* para ser gradiente, y solo es suficiente en dominios simplemente conexos. El apartado (d) enseña exactamente dónde se rompe. Segundo aviso: las identidades exigen que el campo sea de clase C^2; con derivadas cruzadas que no coincidan (cosa posible en funciones patológicas) la cancelación no ocurre. Y tercero: $\operatorname{div}(\operatorname{rot}\mathbf{F}) = 0$ no implica que todo campo de divergencia nula sea un rotacional; hace falta, otra vez, una condición topológica sobre el dominio.

Las identidades $\operatorname{rot} \circ \operatorname{grad} = 0$ y $\operatorname{div} \circ \operatorname{rot} = 0$ forman una "cadena" $f \xrightarrow{\text{grad}} \mathbf{F} \xrightarrow{\text{rot}} \mathbf{G} \xrightarrow{\text{div}} g$ donde componer dos flechas seguidas da cero: esta estructura (un *complejo*) es la versión vectorial de $\mathrm{d} \circ \mathrm{d} = 0$ de las formas diferenciales (sección 11.7).

Regla: calcular gradiente, divergencia y rotacional sin equivocarse

Los tres operadores se confunden porque los tres llevan ∇. Lo primero es mirar *qué entra y qué sale*; eso ya descarta dos de los tres.

1. **Identifica el tipo del objeto.** Si te dan una función escalar $f(x,y,z)$, el único operador aplicable es el gradiente. Si te dan un campo $\mathbf{F} = (P,Q,R)$, puedes aplicarle divergencia (sale un escalar) o rotacional (sale un vector). No existe "el gradiente de un campo" ni "la divergencia de una función".

2. **Gradiente: deriva y apila.** $\operatorname{grad} f = (f_x, f_y, f_z)$: una derivada parcial por coordenada, en orden.

3. **Divergencia: deriva cada componente respecto de su propia variable y suma.** $\operatorname{div}\mathbf{F} = P_x + Q_y + R_z$. La primera componente se deriva respecto de x, la segunda respecto de y, la tercera respecto de z: nunca se cruzan.

4. **Rotacional: monta el determinante.** La forma segura es

$$\operatorname{rot}\mathbf{F} = \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ \partial_x & \partial_y & \partial_z \\ P & Q & R \end{vmatrix} = (R_y - Q_z,\ P_z - R_x,\ Q_x - P_y).$$

   Fíjate en el signo de la componente central: al desarrollar el determinante lleva un menos delante, y por eso queda $P_z - R_x$ y no $R_x - P_z$. Es el error más frecuente.

5. **Comprueba el tipo del resultado.** Gradiente y rotacional deben darte tres números; divergencia, uno solo. Si te sale un vector donde esperabas un escalar, has mezclado operadores.

6. **Verificación gratuita.** Si el campo venía de un gradiente, su rotacional *tiene* que ser $\mathbf{0}$; si venía de un rotacional, su divergencia *tiene* que ser 0 (Proposición 11.31). Aplicar esas dos identidades como control cuesta treinta segundos y detecta casi cualquier error de signo.

Ejemplo 11.33 (Los tres operadores sobre el mismo problema)**.** Sea $f(x,y,z) = x^2y + yz^3$ y sea $\mathbf{F} = (xz,\ xyz,\ -y^2)$.

Gradiente de f. Una derivada por coordenada:

$$f_x = 2xy, \qquad f_y = x^2 + z^3, \qquad f_z = 3yz^2,$$

así que $\operatorname{grad} f = (2xy,\ x^2+z^3,\ 3yz^2)$. En el punto $(1,1,1)$ vale $(2,2,3)$: esa es la dirección de máximo crecimiento de f allí, y su módulo $\sqrt{4+4+9} = \sqrt{17} \approx 4{,}12$ es el ritmo de crecimiento en esa dirección.

Divergencia de $\mathbf{F}$. Cada componente con su propia variable:

$$\operatorname{div}\mathbf{F} = \frac{\partial}{\partial x}(xz) + \frac{\partial}{\partial y}(xyz) + \frac{\partial}{\partial z}(-y^2) = z + xz + 0 = z(1+x).$$

Es un escalar, como debe ser. En $(1,1,1)$ vale $2>0$: allí el campo es una fuente. En todo el plano $z=0$ la divergencia se anula, y para $x<-1$ se invierte el signo: hay sumideros.

Rotacional de $\mathbf{F}$. Con $P=xz$, $Q=xyz$, $R=-y^2$:

$$R_y-Q_z=-2y-xy, \qquad P_z-R_x=x-0=x, \qquad Q_x-P_y=yz-0=yz,$$

luego $\operatorname{rot}\mathbf{F}=(-y(2+x),\ x,\ yz)$. Obsérvese que la componente central salió de P_z-R_x: si por descuido se hubiera escrito R_x-P_z el signo saldría cambiado y el resto del ejercicio arrastraría el error.

Control. ¿Es $\mathbf{F}$ un gradiente? Su rotacional no es nulo (en $(1,1,1)$ vale $(-3,1,1)$), así que no: no existe ninguna g con $\operatorname{grad} g=\mathbf{F}$. En cambio, el rotacional de $\operatorname{grad} f$ sí debe anularse; comprobémoslo con la primera componente:

$$\frac{\partial}{\partial y}(3yz^2)-\frac{\partial}{\partial z}(x^2+z^3)=3z^2-3z^2=0,$$

y lo mismo ocurre con las otras dos.

Ejemplo 11.34 (Laplaciano y funciones armónicas)**.** Sea $f(x,y)=x^2-y^2$. Entonces $\operatorname{grad} f=(2x,-2y)$ y

$$\Delta f=f_{xx}+f_{yy}=2+(-2)=0.$$

Una función con laplaciano nulo se llama *armónica*, y modela un estado de equilibrio: temperatura estacionaria, potencial electrostático en el vacío, altura de una membrana en reposo. La lectura intuitiva es que el laplaciano compara el valor de f en un punto con el promedio de sus vecinos; si es cero, cada punto vale exactamente el promedio de su entorno, y por tanto no hay ningún punto interior más caliente ni más frío que todos los de alrededor. De ahí el principio del máximo: una función armónica alcanza sus extremos en la frontera.

En cambio $g(x,y)=x^2+y^2$ tiene $\Delta g=2+2=4>0$: en cada punto el valor está *por debajo* del promedio de sus vecinos, que es justo lo que ocurre en el fondo de un cuenco.

11.4. Integrales de línea

Idea clave

Intuición primero. Una integral de línea es el trabajo. Si empujas algo a lo largo de un camino con una fuerza que cambia punto a punto, el trabajo total es la suma de "fuerza por desplazamiento" a lo largo del camino, contando solo la componente de la fuerza que va en la dirección del movimiento; el producto escalar $F\cdot dr$ hace precisamente ese descuento.

La pregunta interesante llega enseguida: ¿depende el resultado del camino elegido, o solo de dónde empiezas y dónde acabas? Para casi todos los campos depende del camino. Pero hay una familia especial —los campos conservativos, los que son gradiente de algo— para los que no depende, y entonces la integral es simplemente la diferencia del potencial entre los extremos. Es la versión vectorial del teorema fundamental del cálculo, y su significado físico es la conservación de la energía: dar una vuelta completa y volver al mismo sitio no cuesta trabajo neto. Por eso no existen las máquinas de movimiento

perpetuo en un campo gravitatorio.

Definición 11.35. Sea $\mathcal{C}$ parametrizada por $\mathbf{r}(t)$, $t \in [a,b]$.

- De campo escalar (respecto de longitud de arco; no depende de la orientación):

$$\int_{\mathcal{C}} f \,\mathrm{d}s = \int_a^b f(\mathbf{r}(t))\,\|\mathbf{r}'(t)\|\,\mathrm{d}t$$

  (masa de un alambre si f es densidad; longitud si $f = 1$).

- De campo vectorial (*trabajo*; cambia de signo con la orientación):

$$\int_{\mathcal{C}} \mathbf{F}\cdot\mathrm{d}\mathbf{r} = \int_a^b \mathbf{F}(\mathbf{r}(t))\cdot\mathbf{r}'(t)\,\mathrm{d}t = \int_{\mathcal{C}} P\,\mathrm{d}x + Q\,\mathrm{d}y + R\,\mathrm{d}z.$$

El valor no depende de la parametrización elegida (con la misma orientación): es una propiedad de la curva.

Idea clave

Qué dice esto en cristiano. Hay dos integrales de línea y conviene no confundirlas. La *escalar* suma una magnitud a lo largo del alambre: si f es la densidad, $\int f\,\mathrm{d}s$ es la masa. La *vectorial* suma sólo la componente del campo que va en la dirección del movimiento: si $\mathbf{F}$ es una fuerza, $\int \mathbf{F}\cdot\mathrm{d}\mathbf{r}$ es el trabajo. La primera no distingue el sentido del recorrido; la segunda cambia de signo al recorrer al revés, porque empujar a favor no es lo mismo que empujar en contra.

Ejemplo 11.36 (Las dos integrales de línea, calculadas)**. (a) Una hélice.** Sea $\mathbf{r}(t) = (\cos t, \sin t, t)$ con $t \in [0, 2\pi]$. Entonces $\mathbf{r}' = (-\sin t, \cos t, 1)$ y $\|\mathbf{r}'\| = \sqrt{2}$, constante. La longitud es $\int_0^{2\pi} \sqrt{2}\,\mathrm{d}t = 2\sqrt{2}\,\pi$. Si la densidad es $f = z = t$, la masa es $\int_0^{2\pi} t\sqrt{2}\,\mathrm{d}t = 2\sqrt{2}\,\pi^2$.

(b) Un campo conservativo. Con $\mathbf{F} = (y, x)$ y potencial $f = xy$, el trabajo de $(0,0)$ a $(1,1)$ vale $f(1,1) - f(0,0) = 1$ *por cualquier camino*: por el segmento recto, por la parábola $y = x^2$ o por el camino en escuadra. Merece la pena comprobarlo con la parábola: $\int_0^1 (x^2 \cdot 1 + x \cdot 2x)\,\mathrm{d}x = \int_0^1 3x^2 \mathrm{d}x = 1$.

(c) Un campo que no lo es. Con $\mathbf{F} = (-y, x)$ sobre la circunferencia unidad recorrida en sentido positivo, la circulación vale $\int_0^{2\pi} (\sin^2 t + \cos^2 t)\,\mathrm{d}t = 2\pi \neq 0$. Como hay un lazo cerrado con circulación no nula, no existe potencial en todo el plano menos el origen, por más que $\partial_x(x) - \partial_y(-y)$ sugiera lo contrario.

(d) La orientación. Recorriendo la misma circunferencia al revés, la integral vectorial da -2π, mientras que la escalar $\int 1\,\mathrm{d}s$ sigue valiendo 2π en los dos sentidos. Ésa es la diferencia práctica entre $\mathrm{d}s$ y $\mathrm{d}\mathbf{r}$: la primera es una longitud y la segunda lleva flecha.

Regla: calcular una integral de línea en cinco pasos

Toda integral de línea, escalar o vectorial, se reduce siempre a una integral ordinaria en una sola variable t. El procedimiento no cambia nunca:

1. **Parametriza la curva.** Escribe $\mathbf{r}(t)$ y el intervalo $t \in [a, b]$. Recorridos habituales: un segmento de $\mathbf{A}$ a $\mathbf{B}$ es $\mathbf{r}(t) = \mathbf{A} + t(\mathbf{B} - \mathbf{A})$ con $t \in [0, 1]$; una circunferencia de radio a es $(a\cos t, a\sin t)$ con $t \in [0, 2\pi]$; el trozo de una gráfica $y = g(x)$ es $(t, g(t))$.

2. **Deriva la parametrización.** Calcula $\mathbf{r}'(t)$. Si la integral es escalar necesitarás además $\|\mathbf{r}'(t)\|$; si es vectorial, necesitarás $\mathbf{r}'(t)$ entero, sin normalizar.

3. **Sustituye el integrando.** Reemplaza x, y, z por las componentes de $\mathbf{r}(t)$ *en todas partes*. Después de este paso no puede quedar ninguna x, y o z en el papel: solo t.

4. **Monta la integral correcta.**

$$\int_{\mathcal{C}} f \, \mathrm{d}s = \int_a^b f(\mathbf{r}(t)) \, \|\mathbf{r}'(t)\| \, \mathrm{d}t, \qquad \int_{\mathcal{C}} \mathbf{F} \cdot \mathrm{d}\mathbf{r} = \int_a^b \mathbf{F}(\mathbf{r}(t)) \cdot \mathbf{r}'(t) \, \mathrm{d}t.$$

   La escalar lleva un módulo (siempre positivo); la vectorial, un producto escalar (con signo).

5. **Integra y comprueba el signo.** Si la integral era escalar con $f \geq 0$, el resultado tiene que salir positivo: un resultado negativo significa que has usado $\mathbf{r}'$ en lugar de $\|\mathbf{r}'\|$. Si era vectorial, invertir la orientación cambia el signo del resultado; conviene decir en voz alta en qué sentido se recorrió la curva.

Cuidado

Las dos integrales se parecen en el papel pero no son la misma cosa, y el error de confundirlas es silencioso.

$$\int_{\mathcal{C}} f \, \mathrm{d}s \quad \text{(escalar)} \qquad \text{frente a} \qquad \int_{\mathcal{C}} \mathbf{F} \cdot \mathrm{d}\mathbf{r} \quad \text{(vectorial)}.$$

La primera no depende de la orientación: la masa de un alambre es la misma se recorra de izquierda a derecha o al revés. La segunda cambia de signo: el trabajo de empujar un carro cuesta arriba es el opuesto del de bajarlo. Regla mnemotécnica: si el integrando es una *cantidad* (densidad, temperatura, longitud), es escalar y lleva $\|\mathbf{r}'\|$; si es una *flecha* (fuerza, velocidad, campo), es vectorial y lleva el producto escalar con $\mathbf{r}'$.

Ejemplo 11.37 (Integral escalar: la masa de un alambre helicoidal)**.** Un alambre tiene la forma de la hélice $\mathbf{r}(t) = (\cos t,\ \sin t,\ t)$ con $t \in [0, 2\pi]$, y su densidad lineal en cada punto es igual a la altura, $f(x, y, z) = z$. Queremos su longitud y su masa.

Paso 2. $\mathbf{r}'(t) = (-\sin t,\ \cos t,\ 1)$, de modo que

$$\|\mathbf{r}'(t)\| = \sqrt{\sin^2 t + \cos^2 t + 1} = \sqrt{2},$$

constante: la hélice se recorre a rapidez uniforme.

Longitud. Con $f = 1$:

$$L = \int_0^{2\pi} \sqrt{2}\,\mathrm{d}t = 2\sqrt{2}\,\pi \approx 8{,}886.$$

Comprobación intuitiva: la hélice sube $2\pi \approx 6{,}28$ mientras da una vuelta de longitud horizontal 2π; por Pitágoras la longitud es $\sqrt{(2\pi)^2 + (2\pi)^2} = 2\sqrt{2}\pi$. Coincide.

Masa. Aquí $f(\mathbf{r}(t)) = z = t$, luego

$$M = \int_0^{2\pi} t\,\sqrt{2}\,\mathrm{d}t = \sqrt{2}\,\frac{t^2}{2}\Big|_0^{2\pi} = 2\sqrt{2}\,\pi^2 \approx 27{,}92.$$

El resultado es positivo, como debe ser una masa. Y la densidad media $M/L = \pi \approx 3{,}14$ es exactamente el valor medio de z sobre $[0, 2\pi]$, lo cual confirma el cálculo.

Ejemplo 11.38 (Integral vectorial: el trabajo sí depende del camino)**.** Sea $\mathbf{F} = (y,\ 2x)$ y calculemos el trabajo de $(0,0)$ a $(1,1)$ por dos caminos distintos.

Camino 1: la recta $y = x$. Parametrizamos $\mathbf{r}(t) = (t,t)$, $t \in [0,1]$, con $\mathbf{r}' = (1,1)$. Sustituyendo, $\mathbf{F}(\mathbf{r}(t)) = (t, 2t)$, luego

$$\mathbf{F}\cdot\mathbf{r}' = t\cdot 1 + 2t\cdot 1 = 3t, \qquad W_1 = \int_0^1 3t\,\mathrm{d}t = \frac{3}{2} = 1{,}5.$$

Camino 2: la parábola $y = x^2$. Ahora $\mathbf{r}(t) = (t, t^2)$, $\mathbf{r}' = (1, 2t)$ y $\mathbf{F}(\mathbf{r}(t)) = (t^2, 2t)$, de modo que

$$\mathbf{F}\cdot\mathbf{r}' = t^2\cdot 1 + 2t\cdot 2t = 5t^2, \qquad W_2 = \int_0^1 5t^2\,\mathrm{d}t = \frac{5}{3} \approx 1{,}667.$$

Los dos caminos unen los mismos extremos y dan resultados distintos: $\frac{5}{3} \neq \frac{3}{2}$. Por el Teorema 11.40 eso demuestra que $\mathbf{F}$ *no* es conservativo, y en efecto

$$Q_x - P_y = \frac{\partial}{\partial x}(2x) - \frac{\partial}{\partial y}(y) = 2 - 1 = 1 \neq 0.$$

El circuito cerrado. Si vamos por la parábola y volvemos por la recta, el trabajo neto es $\frac{5}{3} - \frac{3}{2} = \frac{1}{6}$, distinto de cero: dar una vuelta completa en este campo produce energía. Ningún campo de fuerzas físico real se comporta así, y esa es exactamente la razón por la que la gravedad y la electrostática son conservativas.

11.4.1. Campos conservativos

Teorema 11.39 (Fundamental de las integrales de línea)**.** *Si $\mathbf{F} = \nabla f$ en un dominio D y $\mathcal{C} \subset D$ va de $\mathbf{A}$ a $\mathbf{B}$:*

$$\int_{\mathcal{C}} \mathbf{F}\cdot\mathrm{d}\mathbf{r} = f(\mathbf{B}) - f(\mathbf{A}).$$

Demostración. $\frac{\mathrm{d}}{\mathrm{d}t} f(\mathbf{r}(t)) = \nabla f \cdot \mathbf{r}'(t)$ (regla de la cadena), e integrar esta derivada da el incremento de f: es el teorema fundamental del cálculo compuesto con la curva. □

Teorema 11.40 (Caracterización). *Para* $\mathbf{F}$ *de clase* C^1 *en un abierto* conexo D*, son equivalentes: (i)* $\mathbf{F} = \nabla f$ *para algún potencial* f*; (ii)* $\int_{\mathcal{C}} \mathbf{F} \cdot \mathrm{d}\mathbf{r}$ *solo depende de los extremos; (iii)* $\oint_{\mathcal{C}} \mathbf{F} \cdot \mathrm{d}\mathbf{r} = 0$ *sobre toda curva cerrada. Estas condiciones implican* $\operatorname{rot} \mathbf{F} = \mathbf{0}$*; y si además* D *es* simplemente conexo *(sin agujeros: toda curva cerrada se contrae a un punto dentro de* D*), entonces* $\operatorname{rot} \mathbf{F} = \mathbf{0}$ *implica que* $\mathbf{F}$ *es conservativo.*

Idea clave

Qué dice esto en cristiano. Un campo es conservativo cuando el trabajo que realiza al ir de un punto a otro no depende del camino, solo de los extremos. Las cuatro condiciones del teorema son cuatro maneras de decirlo: existe potencial, la integral no depende del camino, toda circulación cerrada es nula y el rotacional se anula. La última es la única que se comprueba derivando, y por eso es la que se usa, pero solo es equivalente a las otras si el dominio no tiene agujeros. En física esto es la conservación de la energía: el trabajo es una diferencia de energía potencial y no importa por dónde vayas.

Ejemplo 11.41 (Decidir si un campo es conservativo). **(a) Uno que sí lo es.** Sea $\mathbf{F} = (2xy,\, x^2)$. Su rotacional escalar es $\frac{\partial}{\partial x}(x^2) - \frac{\partial}{\partial y}(2xy) = 2x - 2x = 0$, y el dominio es todo $\mathbb{R}^2$, que es simplemente conexo. Luego es conservativo, con potencial $f = x^2 y$: en efecto $\nabla f = (2xy, x^2)$.

(b) Calcular el trabajo sin integrar. Con ese potencial, el trabajo de $(0,0)$ a $(2,3)$ vale $f(2,3) - f(0,0) = 4 \cdot 3 - 0 = 12$, sea cual sea el camino: recto, en zigzag o dando vueltas. Ese es todo el ahorro que ofrece el teorema fundamental de las integrales de línea.

(c) Uno que no lo es. Sea $\mathbf{F} = (y, -x)$. Su rotacional escalar es $-1 - 1 = -2 \neq 0$, luego no es conservativo. Su circulación por el círculo unidad vale -2π, y no cero, como confirma Green.

(d) El caso traicionero: rotacional cero y no conservativo. Sea $\mathbf{F} = \left(\frac{-y}{x^2+y^2}, \frac{x}{x^2+y^2}\right)$ en $\mathbb{R}^2 \setminus \{\mathbf{0}\}$. Su rotacional escalar es cero en todo punto del dominio, y sin embargo su circulación por el círculo unidad vale 2π. No es conservativo. El culpable es el agujero: el dominio no es simplemente conexo y la hipótesis del teorema falla.

Cuidado

[La hipótesis sobre el dominio es la que se olvida] El criterio "rotacional nulo implica conservativo" es falso sin simple conexión, como demuestra (d). En la práctica esto importa mucho: el campo magnético alrededor de un hilo de corriente es exactamente el del apartado (d), y su circulación no nula es la ley de Ampère. Segundo aviso: para $\mathbb{R}^3$ el rotacional es vectorial y hay que comprobar sus *tres* componentes, no una. Y tercero: conexo no es lo mismo que simplemente conexo; el plano sin el origen es conexo (de una pieza) y no simplemente conexo (tiene un agujero).

En física: los campos de fuerza conservativos (gravitatorio, electrostático) tienen energía potencial $U = -f$, y el trabajo entre dos puntos no depende del camino: de ahí la conservación de la energía mecánica.

Cuidado

La topología importa. $\mathbf{F} = \left(\frac{-y}{x^2+y^2}, \frac{x}{x^2+y^2}\right)$ cumple $Q_x = P_y$ en $\mathbb{R}^2 \setminus \{0\}$, pero $\oint \mathbf{F}\cdot \mathrm{d}\mathbf{r} = 2\pi$ sobre la circunferencia unitaria: no es conservativo, porque el dominio tiene un agujero. De hecho $\frac{1}{2\pi}\oint$ cuenta las vueltas de la curva alrededor del origen (el *índice*): la integral de línea detecta la topología del dominio. Esta observación es la puerta de entrada a la topología algebraica (cohomología de de Rham).

Cálculo de potenciales: para $\mathbf{F} = (P, Q)$ con $Q_x = P_y$ en un simplemente conexo, se integra $f = \int P\,\mathrm{d}x + g(y)$ y se ajusta g con $f_y = Q$ (ver ejercicios).

Regla: decidir si un campo es conservativo y hallar su potencial

Este es el cálculo que más aparece en los exámenes y en la física elemental. Consta de un test y una construcción.

1. **Test rápido (necesario).** Calcula $\operatorname{rot}\mathbf{F}$. En el plano basta comprobar $\frac{\partial Q}{\partial x} = \frac{\partial P}{\partial y}$; en el espacio hay que verificar las tres igualdades $R_y = Q_z$, $P_z = R_x$, $Q_x = P_y$. Si alguna falla, el campo **no** es conservativo y aquí termina el problema.

2. **Mira el dominio (imprescindible).** Que $\operatorname{rot}\mathbf{F} = \mathbf{0}$ solo garantiza la existencia de potencial si el dominio es *simplemente conexo*. Si al campo se le ha quitado un punto, una recta o un eje, hay que sospechar: el contraejemplo del ángulo polar de la advertencia anterior pasa el test del rotacional y aun así no tiene potencial global.

3. **Integra la primera componente.** Escribe $f = \int P\,\mathrm{d}x$ tratando y y z como constantes. La "constante de integración" no es un número: es una función $g(y, z)$ de las variables que no intervinieron.

4. **Deriva respecto de la segunda variable y compara.** Impón $f_y = Q$. Todo lo que dependa solo de x debe cancelarse; lo que sobrevive determina $\frac{\partial g}{\partial y}$. Intégralo para obtener g, arrastrando ahora una nueva función $h(z)$.

5. **Repite con la tercera.** Impón $f_z = R$ y despeja $h'(z)$; intégralo. Ahora sí, lo que queda es una constante numérica C, que puedes tomar igual a cero: el potencial está definido salvo constante aditiva.

6. **Verifica.** Calcula $\operatorname{grad} f$ y comprueba que reproduce $\mathbf{F}$ componente a componente. Este control cuesta un minuto y detecta cualquier error acumulado.

7. **Úsalo.** Una vez tienes f, toda integral de línea entre $\mathbf{A}$ y $\mathbf{B}$ vale $f(\mathbf{B}) - f(\mathbf{A})$, sin parametrizar nada, por retorcido que sea el camino.

Ejemplo 11.42 (Un potencial en el espacio, paso a paso)**.** Sea $\mathbf{F} = (2xy + z^2,\ x^2 + 2yz,\ 2xz + y^2)$, definido en todo $\mathbb{R}^3$.

Paso 1: el test. Con $P = 2xy + z^2$, $Q = x^2 + 2yz$, $R = 2xz + y^2$:

$$R_y = 2y = Q_z, \qquad P_z = 2z = R_x, \qquad Q_x = 2x = P_y.$$

Las tres coinciden, así que $\operatorname{rot}\mathbf{F} = \mathbf{0}$.

Paso 2: el dominio. Es $\mathbb{R}^3$ entero, que es simplemente conexo. Luego el potencial existe.

Paso 3. Integramos P respecto de x:

$$f = \int (2xy + z^2)\,\mathrm{d}x = x^2y + xz^2 + g(y,z).$$

Paso 4. Derivamos respecto de y e igualamos a Q:

$$f_y = x^2 + \frac{\partial g}{\partial y} \overset{!}{=} x^2 + 2yz \implies \frac{\partial g}{\partial y} = 2yz \implies g = y^2z + h(z).$$

Nótese cómo el término x^2 se cancela solo: si no lo hiciera, sería señal de que el paso 1 estaba mal.

Paso 5. Ahora $f = x^2y + xz^2 + y^2z + h(z)$; derivamos respecto de z:

$$f_z = 2xz + y^2 + h'(z) \overset{!}{=} 2xz + y^2 \implies h'(z) = 0 \implies h = C.$$

Tomamos $C = 0$:

$$\boxed{f(x,y,z) = x^2y + xz^2 + y^2z}.$$

Paso 6: verificación. $f_x = 2xy + z^2$ ✓, $f_y = x^2 + 2yz$ ✓, $f_z = 2xz + y^2$ ✓.

Paso 7: el pago. El trabajo de $\mathbf{F}$ desde el origen hasta $(1,2,3)$, *por cualquier camino*, es

$$f(1,2,3) - f(0,0,0) = (1\cdot 2 + 1\cdot 9 + 4\cdot 3) - 0 = 2 + 9 + 12 = 23.$$

Hacer esto parametrizando una curva concreta habría llevado media página; con el potencial son dos sustituciones.

Ejemplo 11.43 (Un campo que no lo es, y qué hacer entonces)**.** Sea $\mathbf{F} = (x^2y,\ xy^2)$. El test plano:

$$Q_x = y^2, \qquad P_y = x^2,$$

y $y^2 \neq x^2$ salvo sobre las rectas $y = \pm x$. El campo no es conservativo, así que no hay potencial y no hay atajo: cada integral de línea hay que calcularla parametrizando, y el resultado dependerá del camino elegido.

Sin embargo, para *curvas cerradas* sí queda un atajo, y es el teorema de Green de la sección siguiente: en lugar de recorrer el contorno, se integra $Q_x - P_y = y^2 - x^2$ sobre la región encerrada. Sobre un disco centrado en el origen esa integral vale cero por simetría (lo que y^2 aporta de más en la dirección vertical lo compensa x^2 en la horizontal), de modo que la circulación de este campo alrededor de cualquier circunferencia centrada en el origen es nula *sin que el campo sea conservativo*. Circulación nula sobre *algunas* curvas cerradas no basta; el Teorema 11.40 exige que lo sea sobre *todas*.

Regla: cambio de variable en integrales múltiples con el jacobiano

Cuando una integral doble o triple resulta imposible en las coordenadas originales, casi siempre existe un cambio de variable que la vuelve trivial. El procedimiento es mecánico si se respetan los siete pasos.

1. **Mira la región antes que el integrando.** Si el recinto es un disco, un anillo o un sector, piensa en polares; si es un elipsoide, en esféricas escaladas; si está limitado por rectas del tipo $y - x = c$ y $y + x = d$, define $u = y - x$, $v = y + x$. La región dicta el cambio; el integrando solo lo confirma.

2. **Escribe la transformación en la dirección que te conviene.** Para el jacobiano directo necesitas $x = x(u, v)$, $y = y(u, v)$. Si solo tienes $u = u(x, y)$, $v = v(x, y)$, calcula el jacobiano inverso y usa

$$\frac{\partial(x, y)}{\partial(u, v)} = \left(\frac{\partial(u, v)}{\partial(x, y)}\right)^{-1},$$

   evaluado en el punto correspondiente.

3. **Calcula el determinante jacobiano.** Es el determinante de la matriz de derivadas parciales

$$J = \frac{\partial(x, y)}{\partial(u, v)} = \det \begin{pmatrix} \frac{\partial x}{\partial u} & \frac{\partial x}{\partial v} \\ \frac{\partial y}{\partial u} & \frac{\partial y}{\partial v} \end{pmatrix}.$$

   En tres variables la matriz es 3×3 y el procedimiento es idéntico.

4. **Toma valor absoluto.** La fórmula correcta es

$$\iint_R f(x, y)\, dA = \iint_S f\big(x(u, v), y(u, v)\big)\, |J|\, du\, dv.$$

   El valor absoluto no es un adorno: un cambio que invierte la orientación da $J < 0$ y el área no puede ser negativa. Esto distingue el cambio de variable de la integral de línea, donde el signo sí importa.

5. **Traduce la región, no solo la fórmula.** El nuevo recinto S es la preimagen de R. Sustituye las ecuaciones de las fronteras de R en el cambio y despeja: cada curva de R se convierte en una curva de S. Si el cambio está bien elegido, S es un rectángulo o un triángulo.

6. **Comprueba que la transformación sea inyectiva en el interior.** Basta con que $J \neq 0$ ahí. En polares el punto $r = 0$ y la identificación $\theta \sim \theta + 2\pi$ forman un conjunto de medida nula y por eso no estorban; pero si tu cambio recubre la región dos veces, la integral saldrá justo el doble.

7. **Verifica con un caso conocido.** Aplica tu cambio a $f \equiv 1$: debe devolver el área de R. Es la comprobación más barata y detecta casi todos los errores de jacobiano.

Ejemplo 11.44 (Cinco cambios de variable, del más simple al más astuto)**. (a) El jacobiano de una transformación cuadrática.** Sea $u = x^2 - y^2$, $v = 2xy$ (es decir, $u + iv = (x+iy)^2$). Aquí es más fácil el jacobiano inverso:

$$\frac{\partial(u,v)}{\partial(x,y)} = \det\begin{pmatrix} 2x & -2y \\ 2y & 2x \end{pmatrix} = 4x^2 + 4y^2 = 4(x^2+y^2).$$

Es positivo salvo en el origen, así que la transformación es localmente invertible en todas partes menos ahí, y por el paso 2

$$\frac{\partial(x,y)}{\partial(u,v)} = \frac{1}{4(x^2+y^2)}.$$

Obsérvese que $4(x^2+y^2) = |2(x+iy)|^2$: para una función holomorfa $w = f(z)$ el jacobiano real es siempre $|f'(z)|^2$, que es la razón geométrica de que las aplicaciones holomorfas conserven ángulos. Volveremos sobre esto en el capítulo 30.

(b) La integral de Gauss, hecha en polares. La integral

$$I = \iint_{\mathbb{R}^2} e^{-(x^2+y^2)}\, dA$$

es inabordable en cartesianas porque e^{-x^2} no tiene primitiva elemental. En polares $x = r\cos\theta$, $y = r\sin\theta$ el jacobiano es

$$J = \det\begin{pmatrix} \cos\theta & -r\sin\theta \\ \sin\theta & r\cos\theta \end{pmatrix} = r\cos^2\theta + r\sin^2\theta = r,$$

y la región $\mathbb{R}^2$ se convierte en la banda $r \in [0,\infty)$, $\theta \in [0, 2\pi)$. Entonces

$$I = \int_0^{2\pi}\int_0^{\infty} e^{-r^2}\, r\, dr\, d\theta = 2\pi \cdot \Big[-\tfrac{1}{2}e^{-r^2}\Big]_0^{\infty} = 2\pi \cdot \tfrac{1}{2} = \pi.$$

El factor r del jacobiano es exactamente lo que hace falta para que $e^{-r^2} r\, dr$ se integre a ojo: sin él, la integral sería tan imposible como la original. Sobre un disco de radio R el mismo cálculo da

$$\iint_{x^2+y^2\le R^2} e^{-(x^2+y^2)}\, dA = \pi\big(1 - e^{-R^2}\big),$$

que tiende a π cuando $R \to \infty$. Con $R = 1$ vale $\pi(1-e^{-1}) = 1.98587$; con $R = 2$, $\pi(1-e^{-4}) = 3.08404$; con $R = 3$ ya es 3.14120, a cuatro decimales de π.

De aquí sale, de regalo, la integral unidimensional: como $I = \left(\int_{-\infty}^{\infty} e^{-x^2} dx\right)^2$ por Fubini, se concluye

$$\int_{-\infty}^{\infty} e^{-x^2}\, dx = \sqrt{\pi} = 1.7724539.$$

(c) Un cambio lineal que rectifica un triángulo. Calculemos

$$K = \iint_T e^{(y-x)/(y+x)}\, dA, \qquad T = \text{triángulo de vértices } (0,0),\ (1,0),\ (0,1).$$

El integrando pide a gritos $u = y - x$, $v = y + x$. Despejando: $x = \dfrac{v-u}{2}$, $y = \dfrac{v+u}{2}$, y

$$J = \det \begin{pmatrix} -\frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} \end{pmatrix} = -\tfrac{1}{4} - \tfrac{1}{4} = -\tfrac{1}{2}, \qquad |J| = \tfrac{1}{2}.$$

La región se traduce lado a lado: la hipotenusa $x + y = 1$ pasa a $v = 1$; el cateto $x = 0$ pasa a $u = v$; el cateto $y = 0$ pasa a $u = -v$. Luego

$$S = \{(u, v) : 0 \le v \le 1,\ -v \le u \le v\},$$

otro triángulo, pero ahora con el integrando separado. Integramos primero en u, tratando v como constante:

$$\int_{-v}^{v} e^{u/v}\, du = v\left[e^{u/v}\right]_{u=-v}^{u=v} = v\left(e - e^{-1}\right).$$

Por tanto

$$K = \frac{1}{2}\int_0^1 v\,(e - e^{-1})\, dv = \frac{e - e^{-1}}{2}\cdot\frac{1}{2} = \frac{e - e^{-1}}{4} = \frac{2.3504024}{4} = 0.5876006.$$

Una integración numérica directa sobre el triángulo original devuelve 0.5876006: coinciden en los siete decimales. Y el paso 7 también se cumple: con $f \equiv 1$ el cambio da $\frac{1}{2}\int_0^1 2v\, dv = \frac{1}{2}$, que es el área del triángulo.

(d) Coordenadas esféricas y el volumen de la bola. Con $x = \rho\sin\varphi\cos\theta$, $y = \rho\sin\varphi\sin\theta$, $z = \rho\cos\varphi$ el jacobiano es $\rho^2\sin\varphi$, positivo si $\varphi \in (0, \pi)$. Entonces

$$\mathrm{Vol}(B_R) = \int_0^{2\pi}\int_0^{\pi}\int_0^{R} \rho^2\sin\varphi\, d\rho\, d\varphi\, d\theta = 2\pi\cdot 2\cdot\frac{R^3}{3} = \frac{4\pi R^3}{3}.$$

Eso es literalmente el paso 7: probamos el cambio con $f \equiv 1$ y salió la fórmula que ya sabíamos, señal de que el jacobiano está bien.

(e) Un elipsoide por escalado. Para $\dfrac{x^2}{a^2} + \dfrac{y^2}{b^2} + \dfrac{z^2}{c^2} \le 1$ pon $x = au$, $y = bv$, $z = cw$. El jacobiano es la constante abc y la región se convierte en la bola unidad, de modo que

$$\mathrm{Vol} = abc\cdot\frac{4\pi}{3}.$$

Con $a = b = c = R$ se recupera la esfera; con $a = 3$, $b = 2$, $c = 1$ el volumen es $6\cdot\frac{4\pi}{3} = 8\pi = 25.13274$. El mismo truco en dos dimensiones da el área de la elipse: πab.

Cuidado

Dos errores clásicos. El primero: olvidar el jacobiano. Escribir $\iint f\, dr\, d\theta$ en vez de $\iint f\, r\, dr\, d\theta$ es el fallo más frecuente de todo el cálculo en varias variables, y produce resultados que ni siquiera tienen las dimensiones correctas. El segundo: ponerlo al revés. Si calculaste $\partial(u,v)/\partial(x,y)$ pero vas a integrar en $du\, dv$, tienes que invertirlo. Truco de memoria: el dA nuevo debe estar multiplicado por algo que, al integrar $f \equiv 1$, devuelva el área vieja.

Pausa para pensar

¿Por qué el jacobiano y no otra cosa? Porque un rectángulo infinitesimal de lados du y dv se transforma, en primer orden, en el paralelogramo generado por $\frac{\partial \mathbf{r}}{\partial u}\,du$ y $\frac{\partial \mathbf{r}}{\partial v}\,dv$, y el área de un paralelogramo es el valor absoluto del determinante de sus lados. El jacobiano no es una convención: es exactamente el factor por el que la transformación estira el área en cada punto, y su signo dice si además da la vuelta al plano.

11.5. Integrales de superficie

Idea clave

Intuición primero. Si la integral de línea era trabajo, la integral de superficie de un campo es *flujo*: cuánto atraviesa la superficie por unidad de tiempo. Piensa en una red puesta en un río; el flujo cuenta cuánta agua pasa a través de ella. Y de nuevo hay un descuento por dirección: lo que va paralelo a la superficie no la atraviesa, así que solo cuenta la componente perpendicular, que es lo que hace $F \cdot n$.
De ahí sale la necesidad de orientar: hay que decidir cuál es el lado positivo, porque "entrar" y "salir" deben tener signos opuestos. Es una elección arbitraria, pero hay que hacerla y mantenerla. Que existan superficies donde esa elección es imposible —la banda de Möbius, que no tiene dos lados— no es una curiosidad de feria: es la razón por la que los teoremas de la sección siguiente piden explícitamente superficies orientables.

Una *superficie parametrizada* es $\mathbf{r}(u,v)$, $(u,v) \in D$; los vectores $\mathbf{r}_u, \mathbf{r}_v$ generan el plano tangente y $\mathbf{r}_u \times \mathbf{r}_v$ es normal, con $\mathrm{d}S = \|\mathbf{r}_u \times \mathbf{r}_v\|\,\mathrm{d}u\,\mathrm{d}v$ el elemento de área.

- Campo escalar: $\displaystyle\iint_S f\,\mathrm{d}S = \iint_D f(\mathbf{r}(u,v))\|\mathbf{r}_u \times \mathbf{r}_v\|\,\mathrm{d}u\,\mathrm{d}v$.
- *Flujo* (requiere orientación, es decir, elección continua de normal):

$$\iint_S \mathbf{F}\cdot\mathrm{d}\mathbf{S} = \iint_S \mathbf{F}\cdot\mathbf{n}\,\mathrm{d}S = \iint_D \mathbf{F}\cdot(\mathbf{r}_u\times\mathbf{r}_v)\,\mathrm{d}u\,\mathrm{d}v.$$

Para un grafo $z = g(x,y)$: $\mathrm{d}S = \sqrt{1+g_x^2+g_y^2}\,\mathrm{d}A$ y, con normal hacia arriba, $\iint \mathbf{F}\cdot\mathrm{d}\mathbf{S} = \iint_D(-Pg_x - Qg_y + R)\,\mathrm{d}A$.

No toda superficie es orientable: la banda de Möbius no admite normal continua, y sobre ella el flujo no está definido.

Regla: calcular el flujo de un campo a través de una superficie

El esquema es el mismo que en las integrales de línea, con una variable más y un cuidado extra con la orientación.

1. **Parametriza la superficie** con dos parámetros: $\mathbf{r}(u,v)$ con (u,v) recorriendo una región plana D. Las tres parametrizaciones que resuelven casi todo son: la gráfica $z = g(x,y)$, con $\mathbf{r} = (x, y, g(x,y))$; la esfera de radio a, con $\mathbf{r} = (a\sin\varphi\cos\theta,\ a\sin\varphi\sin\theta,\ a\cos\varphi)$; y el cilindro de radio a, con $\mathbf{r} =$

$(a\cos\theta,\ a\sin\theta,\ z)$.

2. **Calcula el vector normal $\mathbf{N} = \mathbf{r}_u \times \mathbf{r}_v$.** No lo normalices: el módulo que lleva dentro es justamente el factor de área $\mathrm{d}S = \|\mathbf{N}\|\,\mathrm{d}u\,\mathrm{d}v$, y normalizarlo te obligaría a volver a multiplicar por él.

3. **Decide la orientación y ajústala.** Mira hacia dónde apunta $\mathbf{N}$ en un punto concreto y compáralo con lo que pide el enunciado (hacia arriba, hacia fuera, alejándose del eje). Si apunta al revés, cambia el signo de $\mathbf{N}$ entero. Este es el paso que más resultados arruina, y es el único que no se detecta después: el número sale bien pero con el signo cambiado.

4. **Sustituye el campo** evaluándolo sobre la superficie: $\mathbf{F}(\mathbf{r}(u,v))$. Como antes, tras este paso solo pueden quedar u y v.

5. **Monta la integral doble** $\displaystyle\iint_S \mathbf{F}\cdot\mathrm{d}\mathbf{S} = \iint_D \mathbf{F}(\mathbf{r}(u,v))\cdot\mathbf{N}\,\mathrm{d}u\,\mathrm{d}v$ y resuélvela con las técnicas del capítulo 9 (Fubini, cambio a polares si D es un disco).

6. **Atajo para gráficas.** Si S es $z = g(x,y)$ con normal hacia arriba, se puede saltar directamente a
$$\iint_S \mathbf{F}\cdot\mathrm{d}\mathbf{S} = \iint_D (-P\,g_x - Q\,g_y + R)\,\mathrm{d}A,$$
donde D es la sombra de S sobre el plano xy.

7. **Si la superficie es cerrada, plantéate Gauss.** Antes de parametrizar seis caras de un cubo o dos casquetes de una esfera, calcula $\operatorname{div}\mathbf{F}$: si sale simple, el teorema de la divergencia convierte todo el trabajo en una sola integral triple.

Ejemplo 11.45 (Flujo a través de un paraboloide, con y sin atajo)**.** Calculemos el flujo de $\mathbf{F} = (x,\ y,\ z)$ a través de la porción del paraboloide $z = 4 - x^2 - y^2$ situada sobre el plano xy, orientada hacia arriba.

Vía la parametrización general. Tomamos $\mathbf{r}(x,y) = (x,\ y,\ 4 - x^2 - y^2)$ sobre el disco $D:\ x^2 + y^2 \le 4$. Entonces

$$\mathbf{r}_x = (1,0,-2x), \qquad \mathbf{r}_y = (0,1,-2y), \qquad \mathbf{N} = \mathbf{r}_x \times \mathbf{r}_y = (2x,\ 2y,\ 1).$$

La tercera componente es $+1$, es decir positiva: la normal apunta hacia arriba, que es lo que pedía el enunciado. No hay que cambiar ningún signo.

Sustituyendo el campo sobre la superficie, $\mathbf{F}(\mathbf{r}) = (x,\ y,\ 4 - x^2 - y^2)$, y

$$\mathbf{F}\cdot\mathbf{N} = 2x^2 + 2y^2 + 4 - x^2 - y^2 = x^2 + y^2 + 4.$$

Pasando a polares, con $x^2 + y^2 = \rho^2$ y $\mathrm{d}A = \rho\,\mathrm{d}\rho\,\mathrm{d}\theta$:

$$\iint_S \mathbf{F}\cdot\mathrm{d}\mathbf{S} = \int_0^{2\pi}\!\!\int_0^{2} (\rho^2 + 4)\rho\,\mathrm{d}\rho\,\mathrm{d}\theta = 2\pi\Big[\frac{\rho^4}{4} + 2\rho^2\Big]_0^2 = 2\pi(4 + 8) = 24\pi.$$

Vía el atajo de la gráfica. Con $g = 4 - x^2 - y^2$ se tiene $g_x = -2x$, $g_y = -2y$, y

$$-Pg_x - Qg_y + R = -x(-2x) - y(-2y) + (4 - x^2 - y^2) = x^2 + y^2 + 4,$$

el mismo integrando. Es el mismo cálculo escrito más corto; conviene saber hacer los dos, porque el atajo solo sirve cuando la superficie es una gráfica.

El resultado $24\pi \approx 75{,}4$ es positivo, coherente con el hecho de que $\mathbf{F} = (x, y, z)$ es un campo que apunta radialmente hacia fuera desde el origen y la superficie está orientada hacia arriba, es decir, alejándose del origen.

Pausa para pensar

En el ejemplo anterior el paraboloide es una superficie *abierta*: tiene borde, la circunferencia $x^2+y^2 = 4$ del plano $z = 0$. Si la tapamos con el disco de ese borde obtenemos una superficie cerrada, y entonces el flujo total se puede calcular con Gauss: $\operatorname{div}(x, y, z) = 3$, y el volumen encerrado entre el paraboloide y el disco es $\int_0^{2\pi}\int_0^2 (4 - \rho^2)\rho \,\mathrm{d}\rho\,\mathrm{d}\theta = 8\pi$, de modo que el flujo cerrado vale $3 \cdot 8\pi = 24\pi$. Pero por la tapa (el disco $z = 0$, con normal exterior $-\mathbf{k}$) el flujo es $\iint (x, y, 0) \cdot (0, 0, -1)\,\mathrm{d}A = 0$. Restando, el flujo por el paraboloide es $24\pi - 0 = 24\pi$: exactamente lo que habíamos obtenido a mano. Dos caminos completamente distintos que aterrizan en el mismo número son la mejor verificación disponible.

11.6. Los tres grandes teoremas

Teorema 11.46 (Green)**.** *Sea $D \subset \mathbb{R}^2$ una región cuya frontera ∂D es una curva cerrada simple a trozos, recorrida dejando D a la izquierda (antihoraria). Para P, Q de clase C^1:*

$$\oint_{\partial D} P\,\mathrm{d}x + Q\,\mathrm{d}y = \iint_D \Big(\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y}\Big)\mathrm{d}A.$$

Demostración para regiones tipo I/II. Basta probar $\oint P\,\mathrm{d}x = -\iint P_y\,\mathrm{d}A$ en una región tipo I ($a \le x \le b$, $\varphi_1 \le y \le \varphi_2$) y la fórmula análoga para Q en tipo II, y sumar. En tipo I:

$$\iint_D P_y\,\mathrm{d}A = \int_a^b \big[P(x, \varphi_2(x)) - P(x, \varphi_1(x))\big]\mathrm{d}x,$$

por el teorema fundamental del cálculo en la variable y. Por otro lado, $\oint P\,\mathrm{d}x$ recorre la gráfica de φ_1 de izquierda a derecha ($\int_a^b P(x, \varphi_1)\mathrm{d}x$), la de φ_2 de derecha a izquierda ($-\int_a^b P(x, \varphi_2)\mathrm{d}x$), y los lados verticales no aportan ($\mathrm{d}x = 0$). Comparando, $\oint P\,\mathrm{d}x = -\iint P_y \mathrm{d}A$. Una región general se trocea en piezas de estos tipos; las fronteras interiores se recorren dos veces en sentidos opuestos y se cancelan. □

Idea clave

Qué dice esto en cristiano. Green, Stokes y Gauss son *el mismo teorema* escrito en tres dimensiones distintas, y todos dicen lo que dice el teorema fundamental del cálculo: lo que pasa dentro de una región queda determinado por lo que pasa en su borde. En

una dimensión, la integral de f' entre a y b es $f(b)-f(a)$: el interior se lee en los dos extremos. En dos, la integral del rotacional sobre una región es la circulación por su curva frontera. En tres, la integral de la divergencia sobre un sólido es el flujo por su superficie. Una sola idea, tres disfraces.

Ejemplo 11.47 (Los tres teoremas comprobados con números)**. (a) Green para calcular áreas.** Con $\mathbf{F}=(-\frac{y}{2},\frac{x}{2})$ el rotacional escalar vale 1, luego $\oint_{\partial D}\mathbf{F}\cdot\mathrm{d}\mathbf{r}$ es exactamente el área de D. Para el círculo de radio 2 da $4\pi=12{,}566371$, y para la elipse de semiejes 3 y 2 da $6\pi=18{,}849556=\pi ab$. Es el principio del planímetro, el aparato que mide áreas recorriendo el contorno.

(b) Stokes en un disco. Sea $\mathbf{F}=(-y,x,0)$, cuyo rotacional es $(0,0,2)$. El flujo del rotacional por el disco unidad del plano XY es $2\cdot\pi\cdot1^2=2\pi=6{,}283185$. Por el otro lado, la circulación directa a lo largo del círculo es $\int_0^{2\pi}(\sin^2 t+\cos^2 t)\,\mathrm{d}t=2\pi$. Coinciden.

(c) Gauss en una bola. Sea $\mathbf{F}=(x,y,z)$, cuya divergencia es 3. Sobre la bola de radio 2, la integral de la divergencia es $3\cdot\frac{4}{3}\pi\cdot8=32\pi=100{,}530965$. Por el otro lado, en la esfera el campo es radial de módulo 2 y el área es $4\pi\cdot4=16\pi$, luego el flujo es $2\cdot16\pi=32\pi$. Coinciden otra vez.

(d) El caso singular que da la ley de Gauss. Para $\mathbf{F}=\frac{\mathbf{r}}{r^3}$ la divergencia es cero en todo punto salvo el origen, y sin embargo el flujo por cualquier esfera centrada en el origen vale 4π, independientemente del radio. El teorema no se aplica porque el campo no está definido en el origen, y esa excepción es precisamente la ley de Gauss del electromagnetismo.

Cuidado

[La orientación no es un detalle] Si la curva se recorre en sentido horario en vez de antihorario, Green cambia de signo y el área sale negativa. En Gauss la normal debe ser *exterior*; con la interior el flujo cambia de signo. Segundo aviso: la frontera debe ser cerrada. Stokes se aplica a superficies *con* borde y la integral es sobre ese borde; si la superficie es cerrada, no tiene borde y la circulación es cero, lo que da otra vez $\operatorname{div}(\operatorname{rot}\mathbf{F})=0$. Y tercero: el campo debe ser C^1 en toda la región cerrada, incluida la frontera; el apartado (d) muestra lo que ocurre al olvidar un solo punto.

Corolario 11.48 (Área encerrada)**.** *Con $P=-y/2$, $Q=x/2$: Área$(D)=\dfrac{1}{2}\displaystyle\oint_{\partial D}(x\,\mathrm{d}y-y\,\mathrm{d}x)$. Para la elipse $(a\cos t,b\sin t)$: $\frac{1}{2}\int_0^{2\pi}ab\,\mathrm{d}t=\pi ab$.*

Teorema 11.49 (Stokes)**.** *Sea S una superficie orientada con frontera ∂S orientada compatiblemente (regla de la mano derecha). Entonces*

$$\oint_{\partial S}\mathbf{F}\cdot\mathrm{d}\mathbf{r}=\iint_S(\operatorname{rot}\mathbf{F})\cdot\mathrm{d}\mathbf{S}.$$

Green es el caso plano ($S\subset$ plano xy, normal $\mathbf{k}$: $(\operatorname{rot}\mathbf{F})\cdot\mathbf{k}=Q_x-P_y$). Consecuencia: el flujo del rotacional solo depende de la frontera; dos superficies con el mismo borde dan el mismo valor —se puede "deformar la superficie" para facilitar el cálculo—.

Idea clave

Qué dice esto en cristiano. Stokes es Green subido a tres dimensiones: la circulación de un campo alrededor del borde de una superficie es igual al flujo de su rotacional a través de ella. Lo notable es que la superficie no importa: dos superficies distintas con el mismo borde dan el mismo resultado, porque la diferencia entre ambas es una superficie cerrada y allí el flujo del rotacional es cero. El teorema de Gauss dice lo análogo con sólidos: lo que sale por la frontera es lo que se ha generado dentro.

Ejemplo 11.50 (La misma frontera, distintas superficies)**. (a) Independencia de la superficie.** Para $\mathbf{F} = (-y, x, 0)$, cuyo rotacional es $(0, 0, 2)$, el flujo a través del disco unidad del plano XY vale 2π. Si en vez del disco se toma el hemisferio superior de la esfera unidad, que tiene el mismo borde, el flujo vuelve a valer 2π. La circulación por el círculo también: $\int_0^{2\pi} (\sin^2 t + \cos^2 t)\, \mathrm{d}t = 2\pi$.

(b) Superficie cerrada, circulación nula. Si S es cerrada no tiene borde, luego Stokes da $\iint_S \operatorname{rot} \mathbf{F} \cdot \mathrm{d}\mathbf{S} = 0$ para todo campo. Eso es otra manera de ver $\operatorname{div}(\operatorname{rot} \mathbf{F}) = 0$, ahora combinando Stokes con Gauss.

(c) Gauss con un campo radial. Para $\mathbf{F} = (x, y, z)$, con $\operatorname{div} \mathbf{F} = 3$, sobre la bola de radio 2: la integral de la divergencia es $3 \cdot \frac{4}{3}\pi \cdot 8 = 32\pi$, y el flujo directo es $2 \cdot 4\pi \cdot 4 = 32\pi$. Coinciden.

(d) La lectura física. La divergencia mide la producción de fluido por unidad de volumen y el rotacional mide la rotación local. Gauss dice que todo el fluido creado dentro tiene que salir por la superficie, y Stokes dice que la rotación total del interior se manifiesta como circulación en el borde. Las ecuaciones de Maxwell en forma integral son exactamente estos dos teoremas aplicados a los campos eléctrico y magnético.

Cuidado

[Las orientaciones deben ser compatibles] El sentido de recorrido del borde y el de la normal de la superficie se relacionan por la regla de la mano derecha; invertir uno sin invertir el otro cambia el signo del resultado y hace que el teorema parezca falso. En Gauss, la normal debe apuntar hacia *fuera* del sólido. Segundo aviso: la superficie debe ser orientable. Sobre la banda de Möbius no se puede aplicar Stokes, porque no hay manera coherente de elegir la normal, y ese es el ejemplo estándar de por qué la hipótesis está en el enunciado.

Teorema 11.51 (Gauss / divergencia)**.** *Sea $E \subset \mathbb{R}^3$ un sólido cuya frontera ∂E se orienta con la normal* exterior*. Entonces*

$$\iint_{\partial E} \mathbf{F} \cdot \mathrm{d}\mathbf{S} = \iiint_E \operatorname{div} \mathbf{F}\, \mathrm{d}V.$$

Idea. Como en Green: se prueba $\iint R\,\mathbf{k} \cdot \mathrm{d}\mathbf{S} = \iiint R_z\, \mathrm{d}V$ en regiones limitadas por dos gráficas $z = g_1, g_2$ (teorema fundamental en z), se hace lo análogo en x e y, y se trocea el sólido general. □

Ejemplo 11.52 (El teorema de la divergencia en tres casos)**. (a) La bola.** Sea $\mathbf{F} = (x, y, z)$ y E la bola de radio $a = 2$. La divergencia es $\operatorname{div}\mathbf{F} = 3$, constante, luego el flujo vale $3 \cdot \frac{4}{3}\pi a^3 = 4\pi a^3 = 32\pi \approx 100{,}530965$. Comprobación directa: sobre la esfera la normal es $\mathbf{n} = \mathbf{r}/a$, de modo que $\mathbf{F} \cdot \mathbf{n} = a$ en todo punto, y el flujo es $a \cdot 4\pi a^2 = 4\pi a^3$. Coincide.

(b) El cubo. Sea $\mathbf{F} = (x^2, y^2, z^2)$ y $E = [0, 1]^3$. Entonces $\operatorname{div}\mathbf{F} = 2(x + y + z)$ y la integral triple vale 3. Calcularlo cara a cara exigiría seis integrales dobles; con Gauss basta una triple de un integrando trivial. Ése es el sentido práctico del teorema: cambiar seis cuentas por una.

(c) Volumen a partir del flujo. Como $\operatorname{div}(x, y, z) = 3$, el volumen de cualquier sólido se puede calcular como $\frac{1}{3}\iint_{\partial E}(x, y, z) \cdot \mathbf{n}\,\mathrm{d}S$. Para el cubo unidad, tres de las seis caras contribuyen $1/3$ cada una y las otras tres nada, dando 1. Es el mismo truco que en el plano usa el teorema de Green para calcular áreas recorriendo el borde.

(d) Cuando el campo no está definido dentro. Con $\mathbf{F} = \mathbf{r}/\|\mathbf{r}\|^3$, se tiene $\operatorname{div}\mathbf{F} = 0$ fuera del origen, y sin embargo el flujo por cualquier esfera centrada en el origen vale 4π, no 0. No hay contradicción: el teorema exige que $\mathbf{F}$ sea suave en *todo* E, y aquí falla en un punto. Ese punto es la carga, y ésta es exactamente la ley de Gauss del electromagnetismo.

Idea clave

Los cuatro teoremas (TFC, Green, Stokes, Gauss) son uno:

$$\int_{\Omega} \mathrm{d}\omega = \int_{\partial\Omega} \omega :$$

la integral de una derivada en una región es la integral del objeto en la frontera. La "derivada" correcta según la dimensión es f', el rotacional o la divergencia; el lenguaje unificador son las formas diferenciales (sección 11.7).

Regla: elegir cuál de los tres teoremas aplicar

Los tres dicen lo mismo en dimensiones distintas, y por eso se confunden. La forma segura de decidir es mirar *qué te dan* y *qué te piden*, no el nombre del capítulo.

1. **¿El problema es plano o espacial?** Si todo ocurre en $\mathbb{R}^2$, el único candidato es Green. Si hay tres coordenadas, quedan Stokes y Gauss.
2. **¿El objeto de partida es una curva o una superficie?**
   - Curva cerrada en el plano $\longleftrightarrow$ región plana: **Green**.
   - Curva cerrada en el espacio $\longleftrightarrow$ superficie con ese borde: **Stokes**.
   - Superficie cerrada en el espacio $\longleftrightarrow$ sólido que encierra: **Gauss**.
3. **Comprueba la hipótesis de cierre.** Green y Gauss exigen que la curva o la superficie sea *cerrada*; Stokes no, pero exige que la curva sea el borde completo de la superficie. Si el enunciado te da un arco abierto, ninguno de los tres se aplica directamente: hay que cerrarlo añadiendo un tramo auxiliar, calcular sobre el circuito cerrado y restar después lo que aporta el tramo añadido.

4. **Decide en qué dirección conviene usarlo.** El teorema es una igualdad, y se puede leer en los dos sentidos. Si el contorno es feo (varios tramos, esquinas) y la región es simple, ve de la izquierda a la derecha: convierte el contorno en área. Si la región es fea y el contorno es simple, ve al revés.

5. **Verifica la orientación antes de calcular.** Green pide recorrido antihorario (la región queda a la izquierda); Stokes pide la regla de la mano derecha entre la normal de S y el sentido de ∂S; Gauss pide normal *exterior*. Si el enunciado da la orientación opuesta, calcula con la buena y cambia el signo al final.

6. **Aprovecha la libertad de Stokes.** Como el flujo del rotacional solo depende del borde, puedes sustituir una superficie complicada por cualquier otra con el mismo borde: casi siempre, el disco plano.

Ejemplo 11.53 (Green: un contorno convertido en un área)**.** Calculemos $\oint_{\mathcal{C}} y^3\,\mathrm{d}x - x^3\,\mathrm{d}y$, donde $\mathcal{C}$ es la circunferencia $x^2+y^2=a^2$ recorrida en sentido antihorario.

Por definición habría que sustituir $x=a\cos t$, $y=a\sin t$ e integrar $-a^4(\sin^4 t+\cos^4 t)$ entre 0 y 2π, lo que exige las fórmulas de reducción de potencias. *Por Green* es inmediato: con $P=y^3$ y $Q=-x^3$,

$$\frac{\partial Q}{\partial x}-\frac{\partial P}{\partial y}=-3x^2-3y^2=-3(x^2+y^2),$$

y en polares $x^2+y^2=\rho^2$, $\mathrm{d}A=\rho\,\mathrm{d}\rho\,\mathrm{d}\theta$:

$$\oint_{\mathcal{C}} y^3\,\mathrm{d}x-x^3\,\mathrm{d}y=\iint_D -3\rho^2\,\rho\,\mathrm{d}\rho\,\mathrm{d}\theta=-3\int_0^{2\pi}\!\!\int_0^{a}\rho^3\,\mathrm{d}\rho\,\mathrm{d}\theta=-3\cdot 2\pi\cdot\frac{a^4}{4}=-\frac{3\pi a^4}{2}.$$

Para $a=1$ da $-3\pi/2\approx -4{,}712$. El signo negativo era previsible: el integrando Q_x-P_y es negativo en todo el disco.

Quien tenga paciencia puede comprobar el camino largo usando $\int_0^{2\pi}\sin^4 t\,\mathrm{d}t=\int_0^{2\pi}\cos^4 t\,\mathrm{d}t=\frac{3\pi}{4}$, lo que da $-a^4\cdot\frac{3\pi}{2}$: el mismo resultado.

Ejemplo 11.54 (Green al revés: el área de un triángulo desde su contorno)**.** El corolario del área dice $\text{Área}(D)=\frac{1}{2}\oint_{\partial D}(x\,\mathrm{d}y-y\,\mathrm{d}x)$. Apliquémoslo al triángulo de vértices $A=(0,0)$, $B=(4,0)$, $C=(0,3)$, recorrido en ese orden (antihorario).

Sobre el lado AB, $y=0$ y $\mathrm{d}y=0$, así que $x\,\mathrm{d}y-y\,\mathrm{d}x=0$: no aporta nada. Sobre CA, $x=0$ y $\mathrm{d}x=0$, así que el integrando también se anula. Solo cuenta el lado BC, que parametrizamos como $\mathbf{r}(t)=(4-4t,\ 3t)$, $t\in[0,1]$, con $\mathrm{d}x=-4\,\mathrm{d}t$ y $\mathrm{d}y=3\,\mathrm{d}t$:

$$x\,\mathrm{d}y-y\,\mathrm{d}x=(4-4t)(3)\,\mathrm{d}t-(3t)(-4)\,\mathrm{d}t=(12-12t+12t)\,\mathrm{d}t=12\,\mathrm{d}t.$$

El integrando es constante, y

$$\text{Área}=\frac{1}{2}\int_0^1 12\,\mathrm{d}t=6,$$

que coincide con $\frac{1}{2}\cdot 4\cdot 3=6$. Esta es exactamente la idea del planímetro: recorrer el borde de una figura y obtener su área sin medir nada del interior.

Ejemplo 11.55 (Stokes: cambiar la superficie por otra más cómoda)**.** Sea $\mathbf{F} = (z,\ x,\ y)$ y sea $\mathcal{C}$ el triángulo de vértices $(1,0,0)$, $(0,1,0)$, $(0,0,1)$, recorrido en ese orden. Queremos $\oint_{\mathcal{C}} \mathbf{F}\cdot \mathrm{d}\mathbf{r}$.

Por Stokes. Primero el rotacional:

$$\operatorname{rot}\mathbf{F} = (R_y - Q_z,\ P_z - R_x,\ Q_x - P_y) = (1-0,\ 1-0,\ 1-0) = (1,1,1).$$

La superficie más cómoda con ese borde es el propio triángulo plano, contenido en el plano $x+y+z=1$, cuya normal unitaria es $\mathbf{n} = \frac{1}{\sqrt{3}}(1,1,1)$; el sentido de recorrido indicado corresponde a esa normal por la regla de la mano derecha. El área del triángulo es la de un triángulo equilátero de lado $\sqrt{2}$, es decir $\frac{\sqrt{3}}{4}(\sqrt{2})^2 = \frac{\sqrt{3}}{2}$. Por tanto

$$\oint_{\mathcal{C}} \mathbf{F}\cdot \mathrm{d}\mathbf{r} = \iint_S (\operatorname{rot}\mathbf{F})\cdot \mathbf{n}\,\mathrm{d}S = \frac{(1,1,1)\cdot(1,1,1)}{\sqrt{3}}\cdot\frac{\sqrt{3}}{2} = \frac{3}{\sqrt{3}}\cdot\frac{\sqrt{3}}{2} = \frac{3}{2}.$$

Comprobación directa. Los tres lados son simétricos. En el primero, $\mathbf{r}(t) = (1-t,\ t,\ 0)$ con $\mathbf{r}' = (-1,1,0)$ y $\mathbf{F} = (0,\ 1-t,\ t)$, luego $\mathbf{F}\cdot\mathbf{r}' = 1-t$ y la integral vale $\frac{1}{2}$. Los otros dos lados dan exactamente lo mismo por la simetría cíclica $x\to y\to z\to x$ que respetan a la vez el campo y el triángulo. Total: $3\cdot\frac{1}{2} = \frac{3}{2}$. ✓

Ejemplo 11.56 (Gauss: seis caras reducidas a una integral triple)**.** Calculemos el flujo saliente de $\mathbf{F} = (x^2,\ y^2,\ z^2)$ a través de la superficie del cubo unidad $[0,1]^3$.

Por definición habría que hacer seis integrales dobles, una por cara, cuidando la normal exterior de cada una. *Por Gauss*, una sola:

$$\operatorname{div}\mathbf{F} = 2x+2y+2z,$$

$$\iint_{\partial E}\mathbf{F}\cdot\mathrm{d}\mathbf{S} = \iiint_E (2x+2y+2z)\,\mathrm{d}V = 2\iiint_E x\,\mathrm{d}V + 2\iiint_E y\,\mathrm{d}V + 2\iiint_E z\,\mathrm{d}V.$$

Cada una de las tres integrales vale $\frac{1}{2}$ (el valor medio de una coordenada en el cubo unidad es $\frac{1}{2}$ y el volumen es 1), luego el flujo total es $2\cdot\frac{1}{2}\cdot 3 = 3$.

Comprobación cara a cara. En la cara $x=1$ la normal exterior es $\mathbf{i}$ y $\mathbf{F}\cdot\mathbf{n} = x^2 = 1$, que sobre un área 1 aporta 1. En la cara $x=0$ la normal exterior es $-\mathbf{i}$ y $\mathbf{F}\cdot\mathbf{n} = -x^2 = 0$: no aporta nada. Lo mismo ocurre en las direcciones y y z, así que el total es $1+1+1=3$. ✓

La moraleja es de contabilidad: el campo entra por tres caras con intensidad nula y sale por las otras tres con intensidad 1; el desequilibrio es lo que la divergencia mide punto a punto.

Ejemplo 11.57 (Gauss con simetría esférica)**.** Flujo saliente de $\mathbf{F} = (x^3,\ y^3,\ z^3)$ a través de la esfera de radio a centrada en el origen. La divergencia es

$$\operatorname{div}\mathbf{F} = 3x^2+3y^2+3z^2 = 3r^2,$$

que depende solo de la distancia al origen. En una situación así conviene integrar por capas: el volumen de la cáscara esférica entre r y $r+\mathrm{d}r$ es $4\pi r^2\,\mathrm{d}r$, de modo que

$$\iiint_E 3r^2\,\mathrm{d}V = \int_0^a 3r^2\cdot 4\pi r^2\,\mathrm{d}r = 12\pi\int_0^a r^4\,\mathrm{d}r = \frac{12\pi a^5}{5}.$$

Para $a=1$ da $12\pi/5 \approx 7{,}54$.

Este truco —cambiar $\mathrm{d}V$ por $4\pi r^2\,\mathrm{d}r$ cuando el integrando solo depende de r— convierte una integral triple en una integral simple, y es el que está detrás de casi todos los cálculos de campo gravitatorio y electrostático de cuerpos con simetría esférica.

11.6.1. Aplicaciones físicas

Ecuación de continuidad: si ρ es densidad y $\mathbf{v}$ velocidad de un fluido, la conservación de la masa en un volumen arbitrario E dice $\frac{\mathrm{d}}{\mathrm{d}t}\iiint_E \rho\,\mathrm{d}V = -\iint_{\partial E}\rho\mathbf{v}\cdot\mathrm{d}\mathbf{S}$; por Gauss, $\iiint_E (\rho_t + \operatorname{div}(\rho\mathbf{v}))\mathrm{d}V = 0$ para todo E, luego

$$\rho_t + \operatorname{div}(\rho\mathbf{v}) = 0.$$

El paso "integral sobre todo volumen ⇒ ecuación puntual" es el mecanismo estándar para derivar las EDP de la física.

Electromagnetismo: las ecuaciones de Maxwell en forma integral (ley de Gauss: flujo de $\mathbf{E}$ = carga encerrada/ε_0; ley de Faraday: circulación de $\mathbf{E}$ = −variación del flujo magnético; etc.) se convierten, con Gauss y Stokes, en su forma diferencial:

$$\operatorname{div}\mathbf{E} = \frac{\rho}{\varepsilon_0},\quad \operatorname{div}\mathbf{B} = 0,\quad \operatorname{rot}\mathbf{E} = -\frac{\partial\mathbf{B}}{\partial t},\quad \operatorname{rot}\mathbf{B} = \mu_0\mathbf{J} + \mu_0\varepsilon_0\frac{\partial\mathbf{E}}{\partial t}.$$

De ellas se deduce (con la identidad $\operatorname{rot}\operatorname{rot} = \operatorname{grad}\operatorname{div} - \Delta$ de la Proposición 11.31) que en el vacío $\mathbf{E}$ satisface la ecuación de onda con velocidad $c = 1/\sqrt{\mu_0\varepsilon_0}$: la luz es una onda electromagnética.

Regla: hallar un potencial vectorial: resolver $\mathbf{B} = \operatorname{rot}\mathbf{A}$

Igual que un campo irrotacional en un dominio simplemente conexo es el gradiente de un potencial escalar, un campo solenoidal en un dominio adecuado es el rotacional de un *potencial vectorial*. El problema inverso $\mathbf{B} = \operatorname{rot}\mathbf{A}$ tiene una receta tan mecánica como la del potencial escalar, pero con una diferencia importante: la solución nunca es única.

1. **Verifica primero la condición necesaria: $\operatorname{div}\mathbf{B} = 0$.** Como $\operatorname{div}(\operatorname{rot}\mathbf{A}) = 0$ siempre, un campo con divergencia no nula no puede ser un rotacional. Este paso cuesta tres derivadas y te ahorra media hora de trabajo inútil.

2. **Comprueba la topología del dominio.** La condición $\operatorname{div}\mathbf{B} = 0$ basta si el dominio es *estrellado* (todo punto se ve desde un punto fijo por un segmento contenido en el dominio), o más generalmente si su segundo grupo de cohomología de De Rham se anula. En $\mathbb{R}^3$ entero, en una bola o en un cubo, siempre basta. En $\mathbb{R}^3\setminus\{0\}$ *no* basta.

3. **Fija un gauge para eliminar la ambigüedad.** Si $\mathbf{A}$ es solución, también lo es $\mathbf{A} + \operatorname{grad} f$ para cualquier f, porque $\operatorname{rot}(\operatorname{grad} f) = \mathbf{0}$. La elección más cómoda a mano es imponer una componente nula, por ejemplo $A_1 = 0$.

4. **Integra dos de las tres ecuaciones.** Con $A_1 = 0$, las componentes del rotacional dan
$$B_2 = -\frac{\partial A_3}{\partial x},\qquad B_3 = \frac{\partial A_2}{\partial x}.$$
Integra ambas respecto de x; aparecen dos "constantes" de integración que son funciones de y y z.

5. **Impón la tercera ecuación para fijar esas funciones.** Sustituye en $B_1 = \frac{\partial A_3}{\partial y} - \frac{\partial A_2}{\partial z}$. Aquí es donde interviene $\operatorname{div}\mathbf{B} = 0$: sin ella, la ecuación resultante sería incompatible. Si todo cuadra, casi siempre puedes tomar las funciones sobrantes iguales a cero.

6. **O usa la fórmula de Poincaré, que no requiere integrar a mano.** En un dominio estrellado respecto del origen,

$$\mathbf{A}(\mathbf{r}) = \int_0^1 t\,\mathbf{B}(t\mathbf{r}) \times \mathbf{r}\,dt$$

resuelve $\operatorname{rot}\mathbf{A} = \mathbf{B}$. Es la versión vectorial del lema de Poincaré y sirve como fórmula cerrada cuando la integración paso a paso se complica.

7. **Verifica calculando el rotacional.** Es una comprobación de tres derivadas cruzadas y es obligatoria: si dos potenciales tuyos difieren, debes poder escribir su diferencia como un gradiente, y si no puedes, uno de los dos está mal.

Ejemplo 11.58 (Cinco potenciales vectoriales, dos gauges y un imposible)**. (a) El campo magnético uniforme.** Sea $\mathbf{B} = (0, 0, B_0)$, el campo dentro de un solenoide ideal. Su divergencia es cero. Dos potenciales muy usados en física:

$$\mathbf{A}_{\text{sim}} = \frac{B_0}{2}(-y,\, x,\, 0), \qquad \mathbf{A}_{\text{Lan}} = (-B_0\, y,\, 0,\, 0).$$

Comprobación del primero: $\operatorname{rot}\mathbf{A}_{\text{sim}} = \big(0-0,\ 0-0,\ \frac{B_0}{2} - (-\frac{B_0}{2})\big) = (0, 0, B_0)$. Comprobación del segundo: $\operatorname{rot}\mathbf{A}_{\text{Lan}} = \big(0,\ 0,\ 0-(-B_0)\big) = (0, 0, B_0)$. Ambos correctos, y su diferencia es

$$\mathbf{A}_{\text{sim}} - \mathbf{A}_{\text{Lan}} = \left(\frac{B_0 y}{2},\ \frac{B_0 x}{2},\ 0\right) = \operatorname{grad}\left(\frac{B_0\, xy}{2}\right),$$

exactamente como predice el paso 7. En mecánica cuántica el primero se llama *gauge simétrico* y el segundo *gauge de Landau*; dan la misma física pero funciones de onda distintas.

(b) Un campo no constante, por integración directa. Sea $\mathbf{B} = (x,\, y,\, -2z)$. Primero el paso 1:

$$\operatorname{div}\mathbf{B} = 1 + 1 - 2 = 0. \quad \checkmark$$

Imponemos $A_1 = 0$. De $B_2 = -\frac{\partial A_3}{\partial x} = y$ sale $A_3 = -xy + f(y, z)$. De $B_3 = \frac{\partial A_2}{\partial x} = -2z$ sale $A_2 = -2xz + g(y, z)$. Ahora la tercera:

$$B_1 = \frac{\partial A_3}{\partial y} - \frac{\partial A_2}{\partial z} = (-x + \partial_y f) - (-2x + \partial_z g) = x + \partial_y f - \partial_z g.$$

Como debe valer x, basta $\partial_y f = \partial_z g$, y lo más simple es $f = g = 0$. Resultado:

$$\mathbf{A} = (0,\, -2xz,\, -xy), \qquad \operatorname{rot}\mathbf{A} = (-x + 2x,\ 0 + y,\ -2z - 0) = (x, y, -2z). \quad \checkmark$$

(c) El mismo campo, por la fórmula de Poincaré. Como $\mathbf{B}$ es lineal, $\mathbf{B}(t\mathbf{r}) = t\,\mathbf{B}(\mathbf{r})$, y la integral se hace sola:

$$\mathbf{A}(\mathbf{r}) = \int_0^1 t \cdot t\,\mathbf{B}(\mathbf{r}) \times \mathbf{r}\,dt = \frac{1}{3}\,\mathbf{B}(\mathbf{r}) \times \mathbf{r}.$$

El producto vectorial es

$$(x, y, -2z) \times (x, y, z) = \big(yz - (-2z)y,\ (-2z)x - xz,\ xy - yx\big) = (3yz,\ -3xz,\ 0),$$

luego $\mathbf{A}_{\mathrm{P}} = (yz,\ -xz,\ 0)$. Verificación:

$$\operatorname{rot} \mathbf{A}_{\mathrm{P}} = \big(0 - (-x),\ y - 0,\ -z - z\big) = (x,\ y,\ -2z). \quad \checkmark$$

Y ahora el control del paso 7: la diferencia entre los dos potenciales es

$$\mathbf{A}_{\mathrm{P}} - \mathbf{A} = (yz,\ -xz + 2xz,\ 0 + xy) = (yz,\ xz,\ xy) = \operatorname{grad}(xyz),$$

un gradiente perfecto. Los dos métodos son consistentes.

(d) Un campo que no puede tener potencial: divergencia no nula. Sea $\mathbf{B} = (x, 0, 0)$, con $\operatorname{div} \mathbf{B} = 1$. El paso 1 ya lo descarta, pero conviene ver *por qué* en términos de flujo. Si existiera $\mathbf{A}$ con $\mathbf{B} = \operatorname{rot} \mathbf{A}$, el flujo de $\mathbf{B}$ a través de cualquier superficie cerrada sería cero (Stokes aplicado a una superficie sin borde). Sin embargo, por el teorema de la divergencia,

$$\iint_{S^2} \mathbf{B} \cdot d\mathbf{S} = \iiint_{B_1} 1 \, dV = \frac{4\pi}{3} = 4.18879 \neq 0.$$

Contradicción. No hay potencial vectorial, ni local ni global.

(e) El monopolo: divergencia nula y aun así imposible. Sea, en $\mathbb{R}^3 \setminus \{0\}$,

$$\mathbf{B} = \frac{\mathbf{r}}{r^3} = \frac{(x, y, z)}{(x^2 + y^2 + z^2)^{3/2}}.$$

Un cálculo directo da $\operatorname{div} \mathbf{B} = \dfrac{3}{r^3} - \dfrac{3(x^2 + y^2 + z^2)}{r^5} = 0$ en todo punto con $r > 0$. Pasa el paso 1. Pero el flujo a través de la esfera unidad es

$$\iint_{S^2} \mathbf{B} \cdot d\mathbf{S} = \iint_{S^2} 1 \, dS = 4\pi \neq 0,$$

y por el mismo argumento de (d) no puede existir un $\mathbf{A}$ definido en todo $\mathbb{R}^3 \setminus \{0\}$. El obstáculo es puramente topológico: el dominio tiene un "agujero" de dimensión dos. Si se retira además una semirrecta desde el origen (la famosa *cuerda de Dirac*), el dominio se vuelve estrellado y el potencial vuelve a existir. Este es el análogo exacto, un grado más arriba, del campo con circulación no nula alrededor de un agujero.

Cuidado

No confundas los dos problemas inversos. Para $\mathbf{F} = \operatorname{grad} \varphi$ la condición es $\operatorname{rot} \mathbf{F} = \mathbf{0}$ y el obstáculo es que el dominio no sea simplemente conexo (agujeros de dimensión uno, tipo toro o plano perforado). Para $\mathbf{B} = \operatorname{rot} \mathbf{A}$ la condición es $\operatorname{div} \mathbf{B} = 0$ y el obstáculo son las cavidades (agujeros de dimensión dos, tipo esfera perforada). Un mismo dominio puede fallar en un caso y no en el otro: $\mathbb{R}^3 \setminus \{0\}$ es simplemente conexo pero tiene una cavidad; $\mathbb{R}^3$ menos un eje tiene el defecto contrario.

Idea clave

Las dos condiciones $\mathrm{rot} \circ \mathrm{grad} = \mathbf{0}$ y $\mathrm{div} \circ \mathrm{rot} = 0$ son la misma identidad escrita dos veces: $d \circ d = 0$ para formas diferenciales. Y las dos preguntas "¿todo campo irrotacional es un gradiente?" y "¿todo campo solenoidal es un rotacional?" son la misma pregunta en grados distintos: ¿toda forma cerrada es exacta? La respuesta la mide la cohomología de De Rham, y el hecho de que en un dominio estrellado siempre sea afirmativa es el lema de Poincaré. La descomposición de Helmholtz, que escribe cualquier campo razonable como $\mathrm{grad}\,\varphi + \mathrm{rot}\,\mathbf{A}$, es la traducción física de ese mismo teorema.

11.7. Formas diferenciales: el lenguaje unificado

Idea clave

Intuición primero. Una k-forma es un *aparato de medición*: come paralelogramitos k-dimensionales y devuelve números (una 1-forma mide segmentitos orientados: trabajo; una 2-forma mide parchecitos orientados: flujo). La derivada exterior d es el *detector de fronteras*: mide cuánto marca el aparato alrededor del borde de una celdita infinitesimal. Con ese diccionario, el teorema de Stokes $\int_M \mathrm{d}\omega = \int_{\partial M} \omega$ es casi una tautología telescópica: sumar las circulaciones de todas las celditas interiores hace que cada pared compartida se recorra dos veces en sentidos opuestos y se cancele; solo sobrevive la frontera exterior. Green, Stokes clásico y Gauss son tres disfraces de esa única cancelación burocrática.

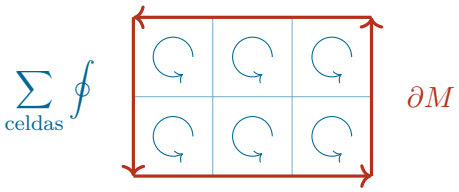


Figura 11.2: El teorema de Stokes como cancelación: cada pared interna se recorre dos veces en sentidos opuestos; de la suma de todas las circulaciones pequeñas solo sobrevive la frontera roja.

Sin pretender un desarrollo completo (que pertenece al capítulo 36 y a la geometría diferencial), esbozamos el lenguaje. Una *k-forma* es un objeto que se integra sobre objetos k-dimensionales:

$$\underbrace{f}_{\text{0-forma}}, \qquad \underbrace{P\,\mathrm{d}x + Q\,\mathrm{d}y + R\,\mathrm{d}z}_{\text{1-forma: se integra en curvas}}, \qquad \underbrace{A\,\mathrm{d}y\wedge\mathrm{d}z + B\,\mathrm{d}z\wedge\mathrm{d}x + C\,\mathrm{d}x\wedge\mathrm{d}y}_{\text{2-forma: se integra en superficies}},$$

y las 3-formas $g\,\mathrm{d}x\wedge\mathrm{d}y\wedge\mathrm{d}z$ se integran en volúmenes. El producto $\wedge$ es *anticonmutativo* ($\mathrm{d}x\wedge\mathrm{d}y = -\mathrm{d}y\wedge\mathrm{d}x$, $\mathrm{d}x\wedge\mathrm{d}x = 0$): de ahí salen los determinantes jacobianos automáticamente al cambiar variables. La *derivada exterior* d sube el grado en 1 y unifica los tres operadores:

$$\mathrm{d}(\text{0-forma}) \leftrightarrow \mathrm{grad}, \qquad \mathrm{d}(\text{1-forma}) \leftrightarrow \mathrm{rot}, \qquad \mathrm{d}(\text{2-forma}) \leftrightarrow \mathrm{div},$$

y cumple $\mathrm{d}\circ\mathrm{d} = 0$ (que condensa rot grad $= 0$ y div rot $= 0$). El teorema de Stokes generalizado,

$$\int_M \mathrm{d}\omega = \int_{\partial M} \omega,$$

contiene como casos particulares los cuatro teoremas de este capítulo. Una forma con $\mathrm{d}\omega = 0$ se llama *cerrada*; una que es $\omega = \mathrm{d}\eta$, *exacta*. Toda exacta es cerrada; el recíproco vale en dominios simplemente conexos (lema de Poincaré) y su fallo en dominios con agujeros —como el campo del índice— mide la topología del dominio.

Regla: operar con formas diferenciales: producto exterior y derivada exterior

Las formas diferenciales son la notación que hace que gradiente, rotacional y divergencia sean *un solo* operador, y que los tres teoremas integrales sean *un solo* teorema. Manejarlas es cuestión de dos reglas de cálculo.

1. **Identifica el grado.** Una 0-forma es una función; una 1-forma es $P\,\mathrm{d}x+Q\,\mathrm{d}y+R\,\mathrm{d}z$; una 2-forma es $P\,\mathrm{d}y\wedge\mathrm{d}z + Q\,\mathrm{d}z\wedge\mathrm{d}x + R\,\mathrm{d}x\wedge\mathrm{d}y$; una 3-forma es $f\,\mathrm{d}x\wedge\mathrm{d}y\wedge\mathrm{d}z$. En $\mathbb{R}^3$ no hay más: toda 4-forma es cero.

2. **Multiplica con la regla del signo:** el producto exterior es antisimétrico, $\mathrm{d}x\wedge\mathrm{d}y = -\,\mathrm{d}y\wedge\mathrm{d}x$, y en consecuencia $\mathrm{d}x\wedge\mathrm{d}x = 0$. Al intercambiar dos factores adyacentes cambia el signo; en general $\alpha\wedge\beta = (-1)^{pq}\beta\wedge\alpha$ para grados p y q.

3. **Deriva con la regla mecánica:** para derivar $f\,\mathrm{d}x_{i_1}\wedge\cdots$, escribe $\mathrm{d}f\wedge\mathrm{d}x_{i_1}\wedge\cdots$, donde $\mathrm{d}f = \sum \frac{\partial f}{\partial x_j}\mathrm{d}x_j$. Luego simplifica con el paso 2, que mata todos los términos repetidos.

4. **Recuerda la regla del producto con signo:** $\mathrm{d}(\alpha\wedge\beta) = \mathrm{d}\alpha\wedge\beta + (-1)^p\alpha\wedge\mathrm{d}\beta$ si α tiene grado p.

5. **Usa la identidad central** $\mathrm{d}\circ\mathrm{d} = 0$**.** No es una curiosidad: es simultáneamente $\mathrm{rot}(\nabla f) = \mathbf{0}$ (aplicada a una 0-forma) y $\mathrm{div}(\mathrm{rot}\,\mathbf{F}) = 0$ (aplicada a una 1-forma). Un solo hecho en lugar de dos.

6. **Distingue cerrada de exacta.** Cerrada significa $\mathrm{d}\omega = 0$; exacta significa $\omega = \mathrm{d}\eta$ para alguna η. Toda exacta es cerrada (paso 5); el recíproco vale si el dominio es *contráctil* (lema de Poincaré), y puede fallar si tiene agujeros.

7. **Control:** el teorema de Stokes general, $\int_M \mathrm{d}\omega = \int_{\partial M}\omega$, contiene como casos particulares el teorema fundamental del cálculo, Green, Stokes clásico y la divergencia. Si tu cuenta no encaja en ese esquema, revisa el grado de la forma.

Ejemplo 11.59 (Cuatro cálculos con formas)**. (a) La derivada exterior de una 1-forma es el rotacional.** Sea $\omega = xy\,\mathrm{d}x + z\,\mathrm{d}y$. Aplicando el paso 3,

$$\mathrm{d}\omega = (y\,\mathrm{d}x + x\,\mathrm{d}y)\wedge\mathrm{d}x + \mathrm{d}z\wedge\mathrm{d}y = x\,\mathrm{d}y\wedge\mathrm{d}x - \mathrm{d}y\wedge\mathrm{d}z = -x\,\mathrm{d}x\wedge\mathrm{d}y - \mathrm{d}y\wedge\mathrm{d}z,$$

donde el término $y\,\mathrm{d}x\wedge\mathrm{d}x$ murió por el paso 2. Traduciendo con el diccionario del paso 1, la 2-forma resultante corresponde al campo $(-1, 0, -x)$. Y en efecto, el rotacional del campo

$\mathbf{F} = (xy,\ z,\ 0)$ es

$$\operatorname{rot}\mathbf{F} = \left(\frac{\partial 0}{\partial y} - \frac{\partial z}{\partial z},\ \frac{\partial (xy)}{\partial z} - \frac{\partial 0}{\partial x},\ \frac{\partial z}{\partial x} - \frac{\partial (xy)}{\partial y}\right) = (-1,\ 0,\ -x).\checkmark$$

(b) La derivada exterior de una 2-forma es la divergencia. Sea

$$\omega = x\,\mathrm{d}y \wedge \mathrm{d}z + y\,\mathrm{d}z \wedge \mathrm{d}x + z\,\mathrm{d}x \wedge \mathrm{d}y,$$

que corresponde al campo radial (x, y, z). Derivando cada sumando, solo sobrevive la derivada respecto de la variable que *falta*:

$$\mathrm{d}\omega = \mathrm{d}x \wedge \mathrm{d}y \wedge \mathrm{d}z + \mathrm{d}y \wedge \mathrm{d}z \wedge \mathrm{d}x + \mathrm{d}z \wedge \mathrm{d}x \wedge \mathrm{d}y = 3\,\mathrm{d}x \wedge \mathrm{d}y \wedge \mathrm{d}z,$$

porque las tres permutaciones son cíclicas y por tanto pares. El 3 es exactamente $\operatorname{div}(x, y, z) = 3$. ✓

(c) $\mathrm{d} \circ \mathrm{d} = 0$, comprobado. Partamos de la 0-forma $f = x^2yz$. Entonces $\mathrm{d}f = 2xyz\,\mathrm{d}x + x^2z\,\mathrm{d}y + x^2y\,\mathrm{d}z$, y derivando otra vez:

$$\mathrm{d}(\mathrm{d}f) = (2xz\,\mathrm{d}y + 2xy\,\mathrm{d}z) \wedge \mathrm{d}x + (2xz\,\mathrm{d}x + x^2\mathrm{d}z) \wedge \mathrm{d}y + (2xy\,\mathrm{d}x + x^2\mathrm{d}y) \wedge \mathrm{d}z.$$

Agrupando por parejas, el término en $\mathrm{d}x \wedge \mathrm{d}y$ recibe $-2xz$ del primer paréntesis y $+2xz$ del segundo; el de $\mathrm{d}x \wedge \mathrm{d}z$ recibe $+2xy$ y $-2xy$; y el de $\mathrm{d}y \wedge \mathrm{d}z$ recibe $+x^2$ y $-x^2$. Todo se cancela: $\mathrm{d}\mathrm{d}f = 0$. La razón de fondo es la igualdad de las derivadas cruzadas, $\partial_i\partial_j f = \partial_j\partial_i f$, contra la antisimetría de $\wedge$.

(d) Cerrada pero no exacta. En $\mathbb{R}^2 \setminus \{(0,0)\}$ considérese

$$\omega = \frac{-y\,\mathrm{d}x + x\,\mathrm{d}y}{x^2 + y^2}.$$

Un cálculo directo da $\mathrm{d}\omega = 0$: la forma es cerrada. Pero integrándola sobre la circunferencia unidad parametrizada por $(\cos t, \sin t)$, el integrando se simplifica a $\mathrm{d}t$ y

$$\oint_{|z|=1} \omega = \int_0^{2\pi} \mathrm{d}t = 2\pi \neq 0.$$

Si fuera exacta, $\omega = \mathrm{d}f$, la integral sobre una curva cerrada valdría 0 por el paso 7. Luego *no* es exacta. El culpable es el agujero: en cualquier disco que no contenga al origen sí existe una primitiva local, el ángulo $\theta = \arctan(y/x)$, pero ese ángulo no se puede definir de forma continua alrededor de todo el agujero, porque al dar la vuelta se ha incrementado en 2π.

Idea clave

El ejemplo (d) es el nacimiento de la cohomología. La pregunta "¿cuántas formas cerradas no exactas hay?" resulta ser una pregunta topológica: cuenta los agujeros del dominio. En $\mathbb{R}^2 \setminus \{0\}$ la respuesta es "una familia de dimensión 1", generada precisamente por esta ω, y por eso mismo aparecen en análisis complejo el índice de una curva y el teorema de los residuos.

Cuidado

La hipótesis del lema de Poincaré es sobre el *dominio*, no sobre la forma. La misma expresión ω de (d), restringida al semiplano $x > 0$, sí es exacta: allí $\omega = \mathrm{d}(\arctan(y/x))$. Cerrada más dominio contráctil implica exacta; cerrada sola, no. Antes de escribir "existe potencial", dibuja el dominio.

Regla: campos cerrados que no son exactos: agujeros, número de vueltas y periodos

El criterio "$\operatorname{rot}\mathbf{F} = \mathbf{0} \Rightarrow \mathbf{F}$ es conservativo" lleva una letra pequeña que casi nadie lee: el dominio debe ser simplemente conexo. Cuando no lo es, aparecen campos cerrados pero no exactos, y con ellos un invariante entero —el número de vueltas— que mide el agujero. Así se maneja la situación.

1. **Calcula el rotacional y anótalo, pero no concluyas todavía.** En el plano, el rotacional escalar es $\frac{\partial F_2}{\partial x} - \frac{\partial F_1}{\partial y}$. Si sale cero, el campo es *cerrado*. Eso es una condición local y solo garantiza que existe un potencial *en un entorno de cada punto*.
2. **Identifica el dominio real de definición.** No es $\mathbb{R}^2$: es $\mathbb{R}^2$ menos los puntos donde el campo o sus derivadas explotan. Localiza esos puntos singulares; son los agujeros.
3. **Decide si el dominio es simplemente conexo.** Lo es si toda curva cerrada puede contraerse a un punto sin salirse. Un disco, un semiplano, un rectángulo y $\mathbb{R}^2$ entero lo son. El plano menos un punto, una corona y el plano menos k puntos no lo son.
4. **Si no lo es, calcula la circulación en un lazo alrededor de cada agujero.** Este número —el *periodo* del campo en ese lazo— es lo que decide todo. Si todos los periodos son cero, el campo sí es conservativo pese al agujero. Si alguno no lo es, el campo es cerrado y no exacto.
5. **Usa la invariancia por deformación para simplificar el cálculo.** Dos lazos homótopos dentro del dominio dan la misma circulación (es Green aplicado a la región entre ambos, donde $\operatorname{rot}\mathbf{F} = \mathbf{0}$). Por eso puedes sustituir cualquier curva fea por la circunferencia más cómoda.
6. **Cuenta las vueltas.** Para una curva cerrada γ que no pasa por el origen,
$$n(\gamma, 0) = \frac{1}{2\pi} \oint_\gamma \frac{x\,dy - y\,dx}{x^2 + y^2} \in \mathbb{Z}$$
es el número de vueltas. La circulación de un campo cerrado sobre γ es la combinación entera de sus periodos con estos coeficientes.
7. **No apliques Green a través de la singularidad.** El teorema de Green exige $\mathbf{F}$ de clase C^1 en *toda* la región encerrada. Si el agujero está dentro, la hipótesis falla y la conclusión "la integral es cero" es falsa. Este es el error más común del tema.

Ejemplo 11.60 (El campo del ángulo, y cinco lazos que lo interrogan)**.** Trabajamos con el campo estrella del tema, definido en $\Omega = \mathbb{R}^2 \setminus \{(0,0)\}$:

$$\mathbf{F}(x,y) = \left(\frac{-y}{x^2+y^2},\ \frac{x}{x^2+y^2} \right).$$

(a) Es cerrado. Con $D = x^2+y^2$,

$$\frac{\partial F_2}{\partial x} = \frac{D - x\cdot 2x}{D^2} = \frac{y^2-x^2}{D^2}, \qquad \frac{\partial F_1}{\partial y} = \frac{-D + y\cdot 2y}{D^2} = \frac{y^2-x^2}{D^2},$$

y la diferencia es 0 en todo punto de Ω. Localmente sí hay potencial: $\varphi = \arctan(y/x)$ funciona en el semiplano $x>0$, y allí $\operatorname{grad}\varphi = \mathbf{F}$ tal cual. El problema es que φ es el *ángulo*, y el ángulo no se puede definir de forma continua en toda la corona alrededor del origen: al dar una vuelta completa aumenta en 2π en vez de volver a su valor.

(b) Circulación en una circunferencia centrada en el origen. Con $\gamma(t) = (R\cos t,\ R\sin t)$, $t\in[0,2\pi]$:

$$\mathbf{F}(\gamma(t)) = \frac{1}{R}(-\sin t,\ \cos t), \qquad \gamma'(t) = R(-\sin t,\ \cos t),$$

luego $\mathbf{F}\cdot\gamma' = \sin^2 t + \cos^2 t = 1$ y

$$\oint_\gamma \mathbf{F}\cdot d\mathbf{r} = \int_0^{2\pi} 1\,dt = 2\pi.$$

Y lo notable: el resultado *no depende de* R. El campo decae como $1/R$ y la longitud crece como R; los dos efectos se cancelan exactamente. Este 2π es el periodo del campo, y su no anulación demuestra de golpe que $\mathbf{F}$ no es conservativo en Ω.

(c) Una elipse, para ver la invariancia por deformación. Tomemos $\gamma(t) = (3\cos t,\ 2\sin t)$. Ahora

$$-y\,x' + x\,y' = -2\sin t\cdot(-3\sin t) + 3\cos t\cdot 2\cos t = 6\sin^2 t + 6\cos^2 t = 6,$$

y $x^2+y^2 = 9\cos^2 t + 4\sin^2 t$, de modo que

$$\oint_\gamma \mathbf{F}\cdot d\mathbf{r} = \int_0^{2\pi} \frac{6\,dt}{9\cos^2 t + 4\sin^2 t} = 6\cdot\frac{2\pi}{3\cdot 2} = 2\pi,$$

usando la fórmula estándar $\int_0^{2\pi} \frac{dt}{a^2\cos^2 t + b^2\sin^2 t} = \frac{2\pi}{ab}$. Una integración numérica confirma 6.2831853. Otra vez 2π: la elipse es deformable a la circunferencia sin cruzar el origen, y el paso 5 garantiza que el valor no cambia.

(d) Un lazo que no encierra el origen. Sea la circunferencia de centro $(2,0)$ y radio 1, es decir $\gamma(t) = (2+\cos t, \sin t)$. Ahora sí se puede aplicar Green: el disco encerrado está contenido en el semiplano $x>0$, donde $\mathbf{F}$ es de clase C^1 y cerrado. Por tanto la circulación es 0, y el cálculo directo lo confirma. Moraleja del paso 7: Green no es falso, simplemente hay que respetar sus hipótesis.

(e) Dar dos vueltas. Con $\gamma(t) = (\cos 2t, \sin 2t)$, $t\in[0,2\pi]$, la curva recorre la circunferencia unidad *dos veces*, y la circulación vale $4\pi = 2\cdot 2\pi$. Con n vueltas en sentido positivo

da $2\pi n$; en sentido negativo, $-2\pi n$. La circulación de $\mathbf{F}$ dividida por 2π es precisamente el número de vueltas.

(f) Un campo cerrado que sí es exacto, en el mismo dominio. Para contrastar, sea

$$\mathbf{G} = \left(\frac{x}{x^2+y^2},\ \frac{y}{x^2+y^2} \right).$$

También está definido solo en Ω y también es cerrado: $\frac{\partial G_2}{\partial x} - \frac{\partial G_1}{\partial y} = \frac{-2xy}{D^2} + \frac{2xy}{D^2} = 0$. Pero ahora

$$\mathbf{G} = \operatorname{grad}\left(\ln \sqrt{x^2+y^2} \right) = \operatorname{grad}(\ln r),$$

y $\ln r$ sí es una función bien definida y continua en todo Ω. Su circulación en cualquier lazo es cero; en la circunferencia de radio R se ve al instante, porque $\mathbf{G}$ es radial y $d\mathbf{r}$ tangente, así que el producto escalar es idénticamente nulo. Conclusión: en un dominio no simplemente conexo hay campos cerrados exactos y campos cerrados no exactos; el agujero *permite* la anomalía pero no la obliga.

(g) Dos agujeros, dos periodos independientes. En $\mathbb{R}^2 \setminus \{a, b\}$ sean $\mathbf{F}_a$ y $\mathbf{F}_b$ el campo del ángulo trasladado a cada punto, y considérese $\mathbf{H} = \mathbf{F}_a + 2\,\mathbf{F}_b$. Un lazo que rodea solo a a da 2π; uno que rodea solo a b da 4π; uno que los rodea a ambos da 6π. Los periodos $(2\pi, 4\pi)$ determinan por completo la circulación de $\mathbf{H}$ sobre cualquier lazo, porque toda curva cerrada se clasifica por su par de números de vueltas.

Cuidado

Que un campo tenga circulación nula en *una* curva concreta no demuestra nada. En el ejemplo (d) obtuvimos 0 y el campo no es conservativo. El criterio correcto es: la circulación debe anularse en *todo* lazo, y para verificarlo basta comprobarla en un lazo alrededor de cada agujero, no en infinitas curvas. Del mismo modo, encontrar un potencial en un trozo del dominio (como $\arctan(y/x)$ en $x > 0$) tampoco basta: hace falta que el potencial se extienda de forma continua a todo el dominio.

Idea clave

Este ejemplo es la puerta de entrada a tres teorías. En cohomología de De Rham, el espacio de campos cerrados módulo exactos en $\mathbb{R}^2$ menos k puntos es $\mathbb{R}^k$, con base los k campos angulares: el álgebra lineal cuenta los agujeros. En variable compleja, el mismo cálculo se escribe $\oint_{|z|=R} \frac{dz}{z} = 2\pi i$, y es el corazón del teorema de los residuos (capítulo 30); de hecho $\mathbf{F}$ y $\mathbf{G}$ son las partes imaginaria y real de $1/\bar{z}$ bien colocadas, y $\ln r + i\theta$ es el logaritmo complejo con su multivaluación. Y en física, este es el efecto Aharonov–Bohm: un electrón que rodea un solenoide adquiere una fase proporcional al flujo encerrado, aunque el campo magnético sea idénticamente nulo en toda la región por la que viaja. El potencial vectorial resulta ser más real que el campo.

11.8. Fundamentos: la teoría detrás de los teoremas integrales

11.8.1. Las fórmulas de Frenet, deducidas

Teorema 11.61 (Fórmulas de Frenet–Serret)**.** *Para una curva parametrizada por arco con* $\kappa > 0$*:*

$$\mathbf{T}' = \kappa\mathbf{N}, \qquad \mathbf{N}' = -\kappa\mathbf{T} + \tau\mathbf{B}, \qquad \mathbf{B}' = -\tau\mathbf{N}.$$

Demostración. $\mathbf{T}' = \kappa\mathbf{N}$ es la definición de $\mathbf{N}$. Como $\|\mathbf{T}\| = 1$, derivando $\mathbf{T}\cdot\mathbf{T} = 1$: $\mathbf{T}'\cdot\mathbf{T} = 0$ —*todo vector de norma constante es ortogonal a su derivada*, el lema que se usa tres veces aquí—. Para $\mathbf{B} = \mathbf{T}\times\mathbf{N}$: $\mathbf{B}' = \mathbf{T}'\times\mathbf{N} + \mathbf{T}\times\mathbf{N}' = \kappa\underbrace{\mathbf{N}\times\mathbf{N}}_{0} + \mathbf{T}\times\mathbf{N}'$, que es ortogonal a $\mathbf{T}$; y por norma constante, ortogonal a $\mathbf{B}$: luego $\mathbf{B}'$ es múltiplo de $\mathbf{N}$, y se *define* τ por $\mathbf{B}' = -\tau\mathbf{N}$. Finalmente $\mathbf{N} = \mathbf{B}\times\mathbf{T}$ se deriva con la regla del producto y las dos fórmulas ya probadas dan la tercera. □

Idea clave

Qué dice esto en cristiano. El triedro de Frenet es un sistema de referencia que viaja pegado a la curva: $\mathbf{T}$ apunta hacia donde se va, $\mathbf{N}$ hacia dónde se gira y $\mathbf{B}$ completa la terna. Las fórmulas dicen cómo giran esos tres ejes al avanzar, y solo hacen falta dos números para describirlo todo: la curvatura κ, que mide cuánto se cierra la curva, y la torsión τ, que mide cuánto se sale del plano. El teorema fundamental añade la noticia decisiva: esos dos números *determinan* la curva salvo su posición en el espacio.

Ejemplo 11.62 (Curvatura y torsión de curvas conocidas)**. (a) Una recta.** $\kappa = 0$ y τ no está definida. Es el único caso degenerado, y por eso el teorema exige $\kappa > 0$.

(b) Una circunferencia. De radio a: $\kappa = \frac{1}{a}$ constante y $\tau = 0$, porque es plana. Para $a = 3$ se tiene $\kappa = \frac{1}{3}$. Cuanto mayor el radio, menor la curvatura: una circunferencia enorme se parece localmente a una recta.

(c) Una hélice. Para $\mathbf{r}(t) = (\cos t, \sin t, t)$ la rapidez es $\sqrt{2}$ constante, y resultan $\kappa = \tau = \frac{1}{2}$, ambas constantes. La hélice es la *única* curva del espacio con curvatura y torsión constantes y no nulas: eso lo garantiza el teorema fundamental.

(d) Qué significa el teorema fundamental. Dadas dos funciones $\kappa(s) > 0$ y $\tau(s)$, existe una curva con esa curvatura y esa torsión, y es única salvo un movimiento rígido. Es decir, κ y τ son el ADN de la curva: dos curvas con las mismas funciones son la misma curva colocada en otro sitio. Es el análogo, para curvas, de que dos triángulos con los mismos lados sean congruentes.

Cuidado

[Las fórmulas exigen parametrización por arco] Tal como están escritas, $\mathbf{T}' = \kappa\mathbf{N}$ y las demás valen cuando el parámetro es la longitud de arco, es decir cuando $\|\mathbf{r}'\| = 1$. Con otro parámetro aparecen factores $\|\mathbf{r}'\|$ por todas partes, y ese es el error de cálculo más frecuente del tema. Segundo aviso: la torsión puede ser negativa (hélice a izquierdas)

mientras que la curvatura se define siempre positiva. Y tercero: en los puntos donde $\kappa = 0$ el vector $\mathbf{N}$ no está definido y el triedro se rompe; por eso una curva con tramos rectos no admite triedro de Frenet global.

Teorema 11.63 (Fundamental de la teoría de curvas)**.** *Dadas funciones continuas $\kappa(s) > 0$ y $\tau(s)$, existe una curva parametrizada por arco con esas curvatura y torsión, única salvo movimiento rígido.* La forma de una curva en el espacio está completamente codificada en dos funciones escalares*: las fórmulas de Frenet son un sistema lineal de EDOs (capítulo 12) cuya solución reconstruye la curva.*

Idea clave

Qué dice esto en cristiano. La curvatura dice cuánto se dobla la curva y la torsión cuánto se sale del plano. El teorema afirma que esos dos números, dados en función de la longitud recorrida, determinan la curva por completo: no hay dos curvas distintas con la misma curvatura y la misma torsión salvo que una sea la otra movida de sitio. Es como decir que basta con anotar cuánto se gira el volante y cuánto se inclina la moto en cada instante para reconstruir el recorrido exacto.

Ejemplo 11.64 (Curvatura y torsión de curvas concretas)**. (a) La recta.** $\kappa \equiv 0$. Es el único caso excluido del enunciado, que pide $\kappa > 0$, porque cuando la curvatura se anula el vector normal deja de estar definido y el triedro de Frenet se rompe.

(b) La circunferencia de radio R. $\kappa = 1/R$ y $\tau = 0$. Curvatura constante y torsión nula caracterizan exactamente a las circunferencias: cuanto menor el radio, más cerrada la curva, como dice el cociente.

(c) La hélice. Para $\mathbf{r}(t) = (a\cos t, a\sin t, bt)$ se obtiene $\kappa = \dfrac{a}{a^2+b^2}$ y $\tau = \dfrac{b}{a^2+b^2}$, ambas constantes. Con $a = b = 1$ salen $\kappa = \tau = 0{,}5$. Recíprocamente, toda curva con κ y τ constantes no nulas es una hélice: los dos números determinan a y b.

(d) Curvas planas. $\tau \equiv 0$ si y sólo si la curva vive en un plano. La demostración es inmediata con las ecuaciones de Frenet: $\tau = 0$ implica que el binormal $\mathbf{B}$ es constante, y entonces $\mathbf{r}\cdot\mathbf{B}$ tiene derivada nula, de modo que la curva está contenida en el plano perpendicular a $\mathbf{B}$.

11.8.2. Existencia del potencial: demostración constructiva

Demostración de que independencia del camino $\Rightarrow$ potencial. Fijemos $\mathbf{x}_0 \in D$ (conexo) y definamos

$$f(\mathbf{x}) = \int_{\mathbf{x}_0}^{\mathbf{x}} \mathbf{F}\cdot \mathrm{d}\mathbf{r}$$

por cualquier camino (bien definido por hipótesis). Para calcular $\frac{\partial f}{\partial x}$ en $\mathbf{x}$, comparamos $f(\mathbf{x}+h\mathbf{e}_1)$ con $f(\mathbf{x})$ usando el camino que va primero a $\mathbf{x}$ y luego el segmento recto horizontal:

$$f(\mathbf{x}+h\mathbf{e}_1) - f(\mathbf{x}) = \int_0^h P(\mathbf{x}+t\mathbf{e}_1)\,\mathrm{d}t,$$

y por el teorema fundamental del cálculo (capítulo 6), la derivada en $h = 0$ es $P(\mathbf{x})$. Igual con las demás componentes: $\nabla f = \mathbf{F}$. □

La parte topológica (irrotacional en simplemente conexo ⇒ independencia del camino) se demuestra deformando continuamente un camino en otro y aplicando Green/Stokes en la superficie barrida por la deformación: ahí es exactamente donde se usa que no hay agujeros.

11.8.3. Demostración del teorema de Stokes

Demostración (superficie parametrizada). Sea S la imagen de $\mathbf{r} : D \to \mathbb{R}^3$ con $D \subset \mathbb{R}^2$ y frontera ∂D. La estrategia: *trasladar todo al plano y usar Green*. Para el término $P\,\mathrm{d}x$ (los otros son análogos):

$$\oint_{\partial S} P\,\mathrm{d}x = \oint_{\partial D} P(\mathbf{r}(u,v))\Big(\frac{\partial x}{\partial u}\,\mathrm{d}u + \frac{\partial x}{\partial v}\,\mathrm{d}v\Big),$$

por la regla de la cadena sobre la parametrización del borde. Aplicando Green en D a este campo plano:

$$= \iint_D \Big[\frac{\partial}{\partial u}\Big(P\frac{\partial x}{\partial v}\Big) - \frac{\partial}{\partial v}\Big(P\frac{\partial x}{\partial u}\Big)\Big]\mathrm{d}u\,\mathrm{d}v.$$

Al expandir con la regla del producto, los términos $P\,x_{vu}$ y $P\,x_{uv}$ se cancelan (¡Clairaut otra vez!), y queda

$$\iint_D \Big(\frac{\partial P}{\partial u}\frac{\partial x}{\partial v} - \frac{\partial P}{\partial v}\frac{\partial x}{\partial u}\Big)\mathrm{d}u\,\mathrm{d}v = \iint_D \big[(\nabla P \times \nabla_{(u,v)} x)\cdot \dots\big],$$

que, tras expandir $\frac{\partial P}{\partial u} = P_x x_u + P_y y_u + P_z z_u$ y agrupar, es exactamente la contribución de P al flujo $\iint_S (\operatorname{rot}\mathbf{F})\cdot \mathrm{d}\mathbf{S}$ escrito en la parametrización. Sumando las tres letras P, Q, R se obtiene el teorema. Para superficies no cubiertas por una sola parametrización, se trocean: las costuras interiores se recorren dos veces con signos opuestos y desaparecen. □

Nótese la anatomía: *Stokes = Green + regla de la cadena + Clairaut*, y Green a su vez fue *teorema fundamental del cálculo + Fubini*. Toda la cadena de los grandes teoremas cuelga del TFC de una variable.

11.8.4. Formas diferenciales con rigor

Formalicemos el esbozo de la sección 11.7 usando el álgebra multilineal del capítulo 7 (y anticipando el 6). Una *k-forma* en $\mathbb{R}^n$ asigna a cada punto una función multilineal *alternada* de k vectores. Las k-formas básicas son $\mathrm{d}x^{i_1} \wedge \dots \wedge \mathrm{d}x^{i_k}$ $(i_1 < \dots < i_k)$, que actúan sobre k vectores como el determinante $k \times k$ de sus coordenadas seleccionadas: la alternancia de las formas *es* la alternancia del determinante. Hay $\binom{n}{k}$ básicas: por eso en $\mathbb{R}^3$ las 1-formas y las 2-formas tienen ambas 3 componentes y se camuflan de campos vectoriales —una coincidencia de $\binom{3}{1} = \binom{3}{2}$ que solo ocurre en dimensión 3, donde vive el rotacional—.

Definición 11.65 (Derivada exterior)**.** Sobre una k-forma $\omega = \sum_I a_I\,\mathrm{d}x^I$ se define $\mathrm{d}\omega = \sum_I \mathrm{d}a_I \wedge \mathrm{d}x^I$, donde $\mathrm{d}a = \sum_j \frac{\partial a}{\partial x_j}\mathrm{d}x^j$.

Teorema 11.66. $\mathrm{d}(\mathrm{d}\omega) = 0$ *para toda forma* ω.

Demostración. Basta en las funciones (el caso general itera):

$$\mathrm{d}(\mathrm{d}a) = \sum_{i,j} \frac{\partial^2 a}{\partial x_j \partial x_i}\, \mathrm{d}x^j \wedge \mathrm{d}x^i.$$

El coeficiente es *simétrico* en (i,j) (Clairaut) y el producto $\mathrm{d}x^j \wedge \mathrm{d}x^i$ es *antisimétrico*: al sumar sobre todos los pares, cada término se cancela con su espejo. Un objeto simétrico contraído con uno antisimétrico se anula: este pequeño argumento es una de las cancelaciones más repetidas de la matemática (reaparecerá con los tensores del capítulo 36). □

Teorema 11.67 (Lema de Poincaré)**.** *En un abierto* estrellado *(en particular, convexo o todo* $\mathbb{R}^n$*), toda forma cerrada (*$\mathrm{d}\omega = 0$*) es exacta (*$\omega = \mathrm{d}\eta$*).*

Idea. Se construye η integrando ω a lo largo de los rayos desde el centro del estrellado (el *operador de homotopía* K), y una verificación directa da $\omega = \mathrm{d}(K\omega) + K(\mathrm{d}\omega) = \mathrm{d}(K\omega)$ cuando $\mathrm{d}\omega = 0$. Para 1-formas, esta η es exactamente el potencial construido arriba con integrales de línea radiales. □

Idea clave

Qué dice esto en cristiano. La derivada exterior es la operación que unifica el gradiente, el rotacional y la divergencia en una sola. Aplicada a una función (una 0-forma) da el gradiente; a una 1-forma en $\mathbb{R}^3$ da el rotacional; a una 2-forma da la divergencia. Y la identidad $\mathrm{dd} = 0$ recoge de golpe las dos identidades del cálculo vectorial. El lema de Poincaré es el recíproco: en un dominio sin agujeros, toda forma cerrada ($\mathrm{d}\omega = 0$) es exacta ($\omega = \mathrm{d}\eta$). La distancia entre cerrado y exacto mide, literalmente, los agujeros del dominio.

Ejemplo 11.68 (Cerrado, exacto y la diferencia entre ambos)**. (a) El diccionario.** En $\mathbb{R}^3$, si $\omega = f$ es una 0-forma, $\mathrm{d}f$ tiene por coeficientes las componentes de ∇f. Si $\omega = P\mathrm{d}x + Q\mathrm{d}y + R\mathrm{d}z$, los coeficientes de $\mathrm{d}\omega$ son las componentes de $\mathrm{rot}(P,Q,R)$. Y si ω es una 2-forma, $\mathrm{d}\omega$ da la divergencia. Tres operadores, una sola definición.

(b) $\mathrm{dd} = 0$ **recupera lo conocido.** Aplicado a una función da $\mathrm{rot}(\nabla f) = \mathbf{0}$; aplicado a una 1-forma da $\mathrm{div}(\mathrm{rot}\,\mathbf{F}) = 0$. Las dos identidades del cálculo vectorial son el mismo enunciado en grados distintos.

(c) El lema de Poincaré funcionando. En $\mathbb{R}^3$ entero, que es convexo y por tanto estrellado, todo campo con rotacional nulo es un gradiente. Para $\mathbf{F} = (2xy, x^2, 0)$ el rotacional es cero y el potencial es $f = x^2 y$, que se puede construir integrando a lo largo del segmento que une el origen con el punto, que es exactamente la receta que da la demostración del lema.

(d) Donde falla, y por qué es interesante. La 1-forma $\omega = \frac{-y\,\mathrm{d}x + x\,\mathrm{d}y}{x^2+y^2}$ en $\mathbb{R}^2 \setminus \{\mathbf{0}\}$ cumple $\mathrm{d}\omega = 0$ y no es exacta: su integral por el círculo unidad vale 2π, y una forma exacta integrada sobre una curva cerrada daría siempre 0. El plano perforado no es estrellado, y el fallo del lema mide exactamente ese agujero. Toda 1-forma cerrada en ese dominio se escribe como $c\,\omega + \mathrm{d}f$, con el número c contando cuántas vueltas da.

Cuidado

[Cerrado y exacto no son sinónimos] Exacto siempre implica cerrado, porque $\mathrm{d}\mathrm{d} = 0$. El recíproco es el que necesita hipótesis sobre el dominio, y el apartado (d) enseña un contraejemplo con un solo punto quitado. Segundo aviso: estrellado es más débil que convexo pero más fuerte que simplemente conexo, y basta para el lema; lo que no basta es "conexo" a secas. Y tercero: el cociente entre formas cerradas y exactas tiene nombre, cohomología de De Rham, y es un invariante topológico: cuenta los agujeros del dominio sin mirarlos, solo derivando.

El fallo del lema en dominios con agujeros no es un defecto: es información. El espacio vectorial de las formas cerradas módulo las exactas (la *cohomología de de Rham* $H^k(D)$) mide los agujeros k-dimensionales de D: para el plano agujereado, $H^1 \cong \mathbb{R}$, generado por nuestra vieja conocida $\frac{-y\,\mathrm{d}x + x\,\mathrm{d}y}{x^2+y^2}$. El análisis vectorial, llevado al límite, se convierte en topología.

11.8.5. Aplicación: unicidad del problema de Dirichlet

Como muestra del poder de los teoremas integrales, demostremos un resultado que usaremos en el capítulo 12. De la identidad $\operatorname{div}(u\nabla u) = \|\nabla u\|^2 + u\,\Delta u$ (Proposición 11.31) y Gauss se obtiene la *primera identidad de Green*:

$$\iint_{\partial E} u\,\nabla u \cdot \mathrm{d}\mathbf{S} = \iiint_E (\|\nabla u\|^2 + u\,\Delta u)\mathrm{d}V.$$

Teorema 11.69 (Unicidad)**.** *Dos funciones armónicas en E que coinciden en ∂E son iguales.*

Demostración. Sea u la diferencia: $\Delta u = 0$ en E y $u = 0$ en ∂E. La identidad de Green da $0 = \iiint_E \|\nabla u\|^2\,\mathrm{d}V$; el integrando es continuo y ≥ 0, luego $\nabla u \equiv \mathbf{0}$: u es constante, y como vale 0 en el borde, $u \equiv 0$. □

Ejemplo 11.70 (Unicidad y sus consecuencias)**. (a) Un caso trivial pero instructivo.** Si u es armónica en la bola y vale 0 en toda la esfera frontera, entonces $u \equiv 0$. Por tanto *no existe* ninguna función armónica no nula que se anule en todo el borde: no hay "modos propios gratis". Compárese con la ecuación de ondas, donde sí los hay.

(b) Resolver adivinando. Si en el disco unidad se pide $u = x^2 - y^2$ en la circunferencia, la respuesta es $u = x^2 - y^2$ en todo el disco, porque esa función es armónica ($2 - 2 = 0$) y cumple el dato. La unicidad convierte una conjetura afortunada en una demostración: basta comprobar que el candidato es armónico y encaja en el borde.

(c) Estabilidad. Del principio del máximo se sigue algo más fuerte que la unicidad: si dos datos de frontera difieren en menos de ε, las soluciones difieren en menos de ε en todo el dominio. El problema de Dirichlet está bien planteado, y por eso se puede resolver numéricamente con confianza.

(d) Dónde falla. En dominios no acotados la unicidad se pierde si no se impone una condición en el infinito: en el semiplano superior, tanto $u = 0$ como $u = y$ son armónicas y valen 0 en el eje real. Hace falta añadir "u acotada" para recuperar la unicidad.

Este *método de energía* —integrar una identidad, obtener una cantidad ≥ 0 forzada a ser 0— es una de las herramientas maestras de las EDP, y volverá a aparecer con el calor y la onda.

11.9. Complementos: coordenadas curvilíneas, Helmholtz y Green

11.9.1. Operadores en coordenadas curvilíneas ortogonales

Con factores de escala $h_i = \|\partial \mathbf{r}/\partial q_i\|$ (cilíndricas: $1, r, 1$; esféricas: $1, \rho, \rho\sin\phi$):

$$\nabla f = \sum_i \frac{1}{h_i}\frac{\partial f}{\partial q_i}\hat{e}_i, \qquad \operatorname{div}\mathbf{F} = \frac{1}{h_1h_2h_3}\sum_i \frac{\partial}{\partial q_i}\Big(\frac{h_1h_2h_3}{h_i}F_i\Big), \qquad \Delta f = \frac{1}{h_1h_2h_3}\sum_i \frac{\partial}{\partial q_i}\Big(\frac{h_1h_2h_3}{h_i^2}\frac{\partial f}{\partial q_i}\Big).$$

Todas son la divergencia covariante del capítulo 36 con $g = \operatorname{diag}(h_i^2)$, $\sqrt{g} = h_1h_2h_3$. El laplaciano esférico que esto produce explica que $\frac{1}{\rho}$ sea armónica.

Regla: escribir gradiente, divergencia, rotacional y laplaciano en coordenadas curvilíneas

En cilíndricas y esféricas los operadores no son "lo mismo con otras letras": aparecen factores de escala. Todos ellos salen de una sola cuenta, así que no hay que memorizar cuatro fórmulas por sistema, sino tres números.

1. **Escribe el elemento de longitud** y lee de él los factores de escala h_1, h_2, h_3:
$$\mathrm{d}s^2 = h_1^2\,\mathrm{d}u_1^2 + h_2^2\,\mathrm{d}u_2^2 + h_3^2\,\mathrm{d}u_3^2.$$
Todo lo demás se deduce de aquí.

2. **Anota los tres casos estándar.** Cartesianas: $h_x = h_y = h_z = 1$. Cilíndricas (r, θ, z): $h_r = 1$, $h_\theta = r$, $h_z = 1$. Esféricas (ρ, φ, θ) con φ el ángulo desde el eje z: $h_\rho = 1$, $h_\varphi = \rho$, $h_\theta = \rho\sin\varphi$.

3. **Gradiente: divide cada derivada por su factor.**
$$\nabla f = \frac{1}{h_1}\frac{\partial f}{\partial u_1}\,\mathbf{e}_1 + \frac{1}{h_2}\frac{\partial f}{\partial u_2}\,\mathbf{e}_2 + \frac{1}{h_3}\frac{\partial f}{\partial u_3}\,\mathbf{e}_3.$$
La razón es física: la derivada mide cambio por unidad de *parámetro*, y el gradiente lo quiere por unidad de *longitud*.

4. **Divergencia: mete todo dentro y saca el producto.**
$$\operatorname{div}\mathbf{F} = \frac{1}{h_1h_2h_3}\left[\frac{\partial(h_2h_3F_1)}{\partial u_1} + \frac{\partial(h_3h_1F_2)}{\partial u_2} + \frac{\partial(h_1h_2F_3)}{\partial u_3}\right].$$

5. **Rotacional: un determinante con los factores repartidos.**
$$\operatorname{rot}\mathbf{F} = \frac{1}{h_1h_2h_3}\begin{vmatrix} h_1\mathbf{e}_1 & h_2\mathbf{e}_2 & h_3\mathbf{e}_3 \\ \partial_{u_1} & \partial_{u_2} & \partial_{u_3} \\ h_1F_1 & h_2F_2 & h_3F_3 \end{vmatrix}.$$

6. **Laplaciano: no lo memorices, compónlo** como $\nabla^2 f = \operatorname{div}(\nabla f)$ usando los pasos 3 y 4. En cilíndricas y esféricas eso da respectivamente
$$\nabla^2 f = \frac{1}{r}\frac{\partial}{\partial r}\left(r\frac{\partial f}{\partial r}\right) + \frac{1}{r^2}\frac{\partial^2 f}{\partial \theta^2} + \frac{\partial^2 f}{\partial z^2}, \qquad \nabla^2 f = \frac{1}{\rho^2}\frac{\partial}{\partial \rho}\left(\rho^2\frac{\partial f}{\partial \rho}\right) + (\text{parte angular}).$$
7. **Controles:** el elemento de volumen es $\mathrm{d}V = h_1 h_2 h_3\,\mathrm{d}u_1\mathrm{d}u_2\mathrm{d}u_3$, que en cilíndricas da $r\,\mathrm{d}r\,\mathrm{d}\theta\,\mathrm{d}z$ y en esféricas $\rho^2\sin\varphi\,\mathrm{d}\rho\,\mathrm{d}\varphi\,\mathrm{d}\theta$: los jacobianos de siempre. Y las identidades $\operatorname{rot}(\nabla f) = \mathbf{0}$ y $\operatorname{div}(\operatorname{rot}\mathbf{F}) = 0$ deben seguir cumpliéndose; si no, hay un factor mal puesto.

Ejemplo 11.71 (Los operadores puestos a prueba)**. (a) El campo radial del plano.** Sea $\mathbf{F} = r\,\mathbf{e}_r$ en cilíndricas, que en cartesianas es $(x, y, 0)$. Con $h_r = 1$, $h_\theta = r$, $h_z = 1$ y $F_r = r$, el paso 4 da
$$\operatorname{div}\mathbf{F} = \frac{1}{r}\frac{\partial(r\cdot r)}{\partial r} = \frac{2r}{r} = 2,$$
que coincide con la cuenta cartesiana $\partial_x x + \partial_y y = 2$. ✓ Obsérvese que quien olvide el factor $1/r$ obtendría 1, y quien olvide el r de dentro obtendría 1 también: los dos errores no se cancelan por casualidad.

(b) Dos funciones armónicas. En cilíndricas, con $f = \ln r$ el paso 6 da
$$\nabla^2 f = \frac{1}{r}\frac{\partial}{\partial r}\left(r\cdot\frac{1}{r}\right) = \frac{1}{r}\frac{\partial(1)}{\partial r} = 0,$$
así que $\ln r$ es armónica en el plano menos el origen: es el potencial de una carga lineal. En esféricas, con $f = 1/\rho$,
$$\nabla^2 f = \frac{1}{\rho^2}\frac{\partial}{\partial \rho}\left(\rho^2\cdot\left(-\frac{1}{\rho^2}\right)\right) = \frac{1}{\rho^2}\frac{\partial(-1)}{\partial \rho} = 0,$$
y reaparece el potencial newtoniano y coulombiano. Que el potencial sea $\ln r$ en dimensión 2 y $1/\rho$ en dimensión 3 es consecuencia directa de que los factores de escala son distintos.

(c) Controles de consistencia. Con $f = r^2$ en cilíndricas, $\nabla^2 f = \frac{1}{r}\partial_r(r\cdot 2r) = \frac{4r}{r} = 4$; en cartesianas $f = x^2 + y^2$ y $\nabla^2 f = 2 + 2 = 4$. ✓ Con $f = \rho^2$ en esféricas, $\nabla^2 f = \frac{1}{\rho^2}\partial_\rho(2\rho^3) = \frac{6\rho^2}{\rho^2} = 6$; en cartesianas $f = x^2 + y^2 + z^2$ y $\nabla^2 f = 2 + 2 + 2 = 6$. ✓

(d) El rotacional del campo azimutal. Sea $\mathbf{F} = \frac{1}{r}\,\mathbf{e}_\theta$, el campo magnético de un hilo. Aplicando el paso 5, la única componente posible es la z, y vale
$$(\operatorname{rot}\mathbf{F})_z = \frac{1}{r}\frac{\partial(r\cdot\frac{1}{r})}{\partial r} = \frac{1}{r}\frac{\partial(1)}{\partial r} = 0$$
en todo punto con $r \neq 0$. El campo es irrotacional fuera del eje y, sin embargo, su circulación a lo largo de una circunferencia que rodee el eje vale $2\pi \neq 0$. No hay contradicción con el teorema de Stokes: el dominio tiene un agujero y ninguna superficie con ese borde evita el eje.

Cuidado

Los vectores $\mathbf{e}_r$, $\mathbf{e}_\theta$, $\mathbf{e}_\varphi$ *no son constantes*: cambian de dirección de punto a punto. Por eso no se pueden sacar de una derivada. Escribir $\frac{\partial}{\partial\theta}(F_r\mathbf{e}_r) = \frac{\partial F_r}{\partial\theta}\mathbf{e}_r$ es falso, y ése es el origen de los términos "extra" de las fórmulas de los pasos 4 y 5: no son adornos, son las derivadas de la base.

Pausa para pensar

¿Por qué $h_\theta = r$ en cilíndricas? Porque avanzar $\mathrm{d}\theta$ en ángulo recorre un arco de longitud $r\,\mathrm{d}\theta$, no $\mathrm{d}\theta$: cuanto más lejos del eje, más camino por el mismo ángulo. Y por eso mismo $h_\theta = \rho\sin\varphi$ en esféricas: el radio del paralelo a latitud φ es $\rho\sin\varphi$, y se anula en los polos, donde toda la longitud se colapsa.

11.9.2. Helmholtz

Todo campo que decae se parte único como $\mathbf{F} = -\nabla\varphi + \operatorname{rot}\mathbf{A}$ (irrotacional + solenoidal), reconstruible desde $\operatorname{div}\mathbf{F}$ y $\operatorname{rot}\mathbf{F}$ con integrales newtonianas: *un campo es su divergencia y su rotacional*. Por eso Maxwell son cuatro ecuaciones: prescriben ambas para $\mathbf{E}$ y $\mathbf{B}$.

11.9.3. Identidades de Green y unicidad

$$\iint_{\partial E} u\frac{\partial v}{\partial n}\,\mathrm{d}S = \iiint_E (u\Delta v + \nabla u\cdot\nabla v)\mathrm{d}V, \qquad \iint_{\partial E}\Big(u\frac{\partial v}{\partial n} - v\frac{\partial u}{\partial n}\Big)\mathrm{d}S = \iiint_E (u\Delta v - v\Delta u)\mathrm{d}V.$$

La segunda dice que Δ es autoadjunto (Sturm–Liouville dimensional). Con la solución fundamental $\frac{1}{4\pi r}$ sale la fórmula de representación (funciones de Green, jaulas de Faraday).

Teorema 11.72 (Valor medio y máximo)**.** *u armónica: su valor central es el promedio esférico ($\varphi'(r) = \frac{1}{4\pi r^2}\iint \Delta u = 0$); en consecuencia, sin extremos interiores estrictos.*

Ejemplo 11.73 (La propiedad del valor medio en acción)**. (a) Comprobación directa.** Sea $u = x^2 - y^2$, armónica. Sobre la circunferencia unidad centrada en el origen, $u = \cos^2 t - \sin^2 t = \cos 2t$, cuya media en $[0, 2\pi]$ es 0, que es justamente $u(0,0)$. Con $u = xy$ ocurre lo mismo: la media de $\cos t\sin t$ es 0 y $u(0,0) = 0$.

(b) Cuando la función no es armónica. Sea $u = x^2 + y^2$, con $\Delta u = 4 > 0$ (subarmónica). Su media sobre la circunferencia unidad vale 1, estrictamente mayor que $u(0,0) = 0$. La desigualdad va siempre en ese sentido: las subarmónicas quedan por debajo de sus promedios, las superarmónicas por encima, y sólo las armónicas empatan.

(c) El principio del máximo. De la propiedad del valor medio se sigue que una función armónica no constante no puede tener un máximo interior: si lo tuviera, sería igual al promedio de valores todos menores o iguales, lo que fuerza a que todos sean iguales, y se propaga por conexión hasta la frontera. Por eso el calor estacionario en una placa siempre alcanza su temperatura extrema en el borde.

(d) El teorema de Liouville. Si u es armónica y acotada en todo $\mathbb{R}^n$, entonces es constante. La razón es que el valor en dos puntos cualesquiera se puede escribir como promedio sobre bolas gigantescas que casi coinciden, y la diferencia tiende a cero. Es el análogo exacto del teorema de Liouville para funciones enteras del análisis complejo, y no es casualidad: la parte real de una función holomorfa es armónica.

11.9.4. Ángulo sólido

$\Omega = \iint_S \frac{\mathbf{x}\cdot d\mathbf{S}}{\|\mathbf{x}\|^3}$: 4π si la superficie cerrada encierra el origen, 0 si no —la ley de Gauss es un invariante topológico, hermano 3D del índice—.

11.10. Cierre del temario: los subtemas restantes

11.10.1. Forma normal de Green

$$\oint_{\partial D} \mathbf{F}\cdot\mathbf{n}\,ds = \iint_D \operatorname{div}\mathbf{F}\,dA$$

(la tangencial aplicada al campo girado $(Q,-P)$): el teorema de la divergencia plano, y el indicio de que rotacional y divergencia son sombras del único d.

11.10.2. Deformación de contornos

Si $\operatorname{rot}\mathbf{F}=0$ entre $\mathcal{C}_1$ y $\mathcal{C}_2$: $\oint_{\mathcal{C}_1} = \oint_{\mathcal{C}_2}$ (Green en la corona). El campo del índice sobre cualquier ameba que rodee el origen: 2π, deformando a la circunferencia. Corazón del teorema de los residuos y de la clasificación de campos por períodos.

11.10.3. Parametrizaciones estándar

grafo: $\sqrt{1+g_x^2+g_y^2}$; esfera: $R^2\sin\phi$; cilindro: a; revolución: $f\sqrt{1+f'^2}$ (¡demuestra la fórmula del capítulo 6!); toro $(R+r\cos v,\dots)$: $r(R+r\cos v)$, de donde

$$S_{\text{toro}} = \int_0^{2\pi}\int_0^{2\pi} r(R+r\cos v)\,du\,dv = 4\pi^2 Rr$$

confirmando a Pappus. Patrón: el factor es siempre $\sqrt{\det g}$ de la métrica inducida.

11.10.4. Independencia de la parametrización

Reparametrizar con $\varphi'>0$ no cambia las integrales (cambio de variables 1D/2D); invertir orientación cambia el signo de trabajo y flujo pero no de ds, dS: la tabla de signos completa.

11.11. Profundización final: Kepler, potenciales vectoriales e índices

11.11.1. Potencial vector explícito

Para $\mathbf{B} = B\mathbf{e}_3$: $\mathbf{A} = \frac{1}{2}\mathbf{B} \times \mathbf{r}$ cumple $\operatorname{rot} \mathbf{A} = \mathbf{B}$; también $B(-y, 0, 0)$: libertad de gauge $\mathbf{A} \to \mathbf{A} + \nabla\chi$. En cuántica el gauge deja huella medible ($\oint \mathbf{A} \cdot \mathrm{d}\mathbf{r}$: Aharonov–Bohm): topología hecha física.

11.11.2. El problema de Kepler completo

Con $u = \frac{1}{r}$ y θ como variable ($L = r^2\dot{\theta}$ conservado): $\dot{r} = -Lu'$, $\ddot{r} = -L^2u^2u''$, y la ecuación radial se convierte en

$$u'' + u = \frac{GM}{L^2} \implies r(\theta) = \frac{L^2/GM}{1 + e\cos(\theta - \theta_0)} :$$

¡el oscilador armónico disfrazado, y una cónica polar! Primera ley de Kepler demostrada; la tercera sale del área de la elipse: $T^2 = \frac{4\pi^2}{GM}a^3$. Setenta años de datos de Tycho comprimidos en un cambio de variables.

11.11.3. Poincaré–Hopf

Índice de un cero = vueltas del campo alrededor (fuentes/sumideros/centros $+1$, sillas -1).

Teorema 11.74. *En una superficie cerrada:* $\sum \mathrm{ind} = \chi(S)$ *(2 esfera, 0 toro).*

Ejemplo 11.75 (Índices y el teorema de la bola peluda)**. (a) Los índices básicos en el plano.** Una fuente o un sumidero tienen índice $+1$; un centro (órbitas circulares) también $+1$; un punto de silla tiene índice -1. El índice se calcula contando cuántas vueltas netas da el campo al recorrer una curva pequeña alrededor de la singularidad.

(b) La esfera. $\chi(S^2) = 2$, luego todo campo tangente continuo en la esfera tiene singularidades cuyos índices suman 2. Con dos fuentes (polos) sale $1 + 1 = 2$; con una única singularidad, ésta ha de tener índice 2, como el campo que "sale del polo norte y entra por el mismo sitio". Lo que es imposible es no tener ninguna: ése es el teorema de la bola peluda, y su versión coloquial es que no se puede peinar una esfera sin dejar un remolino.

(c) El toro. $\chi(T^2) = 0$, así que sí existe un campo tangente sin ceros: el que apunta siempre en la dirección del paralelo. Un donut sí se puede peinar. Ésta es la diferencia topológica entre un balón y un flotador, expresada en términos de campos vectoriales.

(d) Una consecuencia meteorológica. Como el viento en la superficie terrestre es un campo tangente continuo, en todo momento existe al menos un punto del planeta donde el viento horizontal es exactamente nulo. No es una observación empírica: es un teorema.

Bola peluda: no se peina un coco sin remolino; siempre hay un punto de la Tierra sin viento. El pariente suave de Gauss–Bonnet: topología imponiendo cuotas a lo local.

11.11.4. Maxwell rematado: ondas y Poynting

rot de Faraday + identidades: $\Delta\mathbf{E} = \mu_0\varepsilon_0\mathbf{E}_{tt}$: ondas a $c = \frac{1}{\sqrt{\mu_0\varepsilon_0}}$ —Maxwell, 1864: la luz es electromagnetismo—. Energía: con $u = \frac{\varepsilon_0 E^2}{2} + \frac{B^2}{2\mu_0}$ y $\mathbf{S} = \frac{\mathbf{E}\times\mathbf{B}}{\mu_0}$:

$$u_t + \operatorname{div}\mathbf{S} = -\mathbf{J}\cdot\mathbf{E},$$

continuidad de la energía; la luz solar es el flujo de $\mathbf{S}$.

11.12. Última capa: Gauss demostrado, Biot–Savart y cohomología

11.12.1. Demostración del teorema de la divergencia

Idea clave

Intuición primero. Divide el sólido en cajitas. El flujo que sale de cada cajita por una pared interna *entra* a la cajita vecina por la misma pared: en la suma total, todas las paredes internas se cancelan por pares y solo queda la piel exterior. Por otro lado, el flujo neto de cada cajita diminuta es (por definición de divergencia) $\operatorname{div}\mathbf{F}$ por su volumen. Igualar las dos contabilidades —la de la piel y la de las fuentes— *es* el teorema; la demostración de abajo solo pone ε donde este párrafo pone "diminuta".

Para la componente $R\,\mathbf{e}_3$ en una región z-simple ($g_1(x,y) \le z \le g_2(x,y)$ sobre D):

$$\iiint_E \frac{\partial R}{\partial z}\,\mathrm{d}V = \iint_D \big[R(x,y,g_2) - R(x,y,g_1)\big]\mathrm{d}A$$

por el teorema fundamental en z. En la tapa superior, la normal exterior tiene componente $\mathbf{k}$ positiva y $\iint R\,\mathbf{k}\cdot\mathrm{d}\mathbf{S} = \iint_D R(x,y,g_2)\mathrm{d}A$ (parametrización como grafo); en la inferior, signo opuesto; en las paredes laterales, $\mathbf{k}\cdot\mathbf{n} = 0$. Sumando: el flujo de $R\,\mathbf{e}_3$ iguala la integral de R_z. Se repite en x e y para regiones simples en cada eje, y una región general se corta en trozos simples: las caras internas se recorren con normales opuestas y se cancelan. □ La misma cirugía de Green, un piso más arriba: TFC + Fubini + recortar.

11.12.2. Biot–Savart

El campo magnético de una corriente estacionaria:

$$\mathbf{B}(\mathbf{r}) = \frac{\mu_0 I}{4\pi}\oint \frac{\mathrm{d}\boldsymbol{\ell}\times(\mathbf{r}-\mathbf{r}')}{\|\mathbf{r}-\mathbf{r}'\|^3},$$

la "ley de Coulomb magnética" (se verifica que da $\operatorname{div}\mathbf{B} = 0$ y Ampère). Ejemplo canónico, el hilo recto infinito: por simetría e integrando,

$$B = \frac{\mu_0 I}{2\pi d},$$

círculos alrededor del hilo decayendo como $\frac{1}{d}$ —exactamente el campo del índice del plano, con intensidad $\mu_0 I$: la circulación $\oint \mathbf{B}\cdot\mathrm{d}\mathbf{r} = \mu_0 I$ (Ampère) es el "período" topológico alrededor del agujero que el hilo perfora en el espacio—. Dos siglos de electromagnetismo y la cohomología del plano agujereado resultan ser el mismo cálculo.

11.12.3. El primer grupo de cohomología, calculado

Teorema 11.76. *En $\mathbb{R}^2 \setminus \{0\}$, toda 1-forma cerrada ω se escribe de forma única*

$$\omega = c\,\omega_0 + \mathrm{d}f, \qquad c = \frac{1}{2\pi}\oint_{S^1} \omega,$$

con $\omega_0 = \frac{-y\,\mathrm{d}x + x\,\mathrm{d}y}{x^2+y^2}$: es decir, $H^1(\mathbb{R}^2 \setminus \{0\}) \cong \mathbb{R}$, generado por el campo del índice.

Idea. Restando $c\,\omega_0$ se obtiene una forma cerrada con período cero alrededor del agujero; entonces la integral desde un punto base no depende del camino (todo lazo es combinación de lazos pequeños contractibles y vueltas al agujero, y ambas dan cero) y define el potencial f. La unicidad: ω_0 no es exacta (su período es $2\pi \neq 0$). □

Idea clave

Qué dice esto en cristiano. Una forma cerrada es un campo que "parece" tener potencial: todas sus derivadas cruzadas cuadran. Una forma exacta es un campo que *de verdad* lo tiene. En un dominio sin agujeros ambas cosas coinciden; en el plano agujereado no, y el teorema mide exactamente el desfase: hay una única forma cerrada no exacta esencialmente distinta, la que da vueltas alrededor del agujero, y todo lo demás es exacto. El coeficiente c es el número de vueltas. En una frase: el álgebra de las formas está contando los agujeros del dominio.

Ejemplo 11.77 (La forma que no tiene potencial)**. (a) La protagonista.** $\omega_0 = \dfrac{-y\,\mathrm{d}x + x\,\mathrm{d}y}{x^2+y^2}$ es cerrada: las dos derivadas cruzadas valen $\dfrac{y^2-x^2}{(x^2+y^2)^2}$. Su integral sobre la circunferencia unidad vale 2π, luego no es exacta.

(b) El falso potencial. La función $\theta = \arctan(y/x)$ cumple $\mathrm{d}\theta = \omega_0$ allí donde está definida, pero el ángulo no es una función continua en todo el plano agujereado: al dar una vuelta completa aumenta en 2π. La obstrucción no es local sino global, y por eso ninguna cuenta con derivadas la detecta.

(c) Descomponer una forma cualquiera. Sea $\omega = \dfrac{-y\,\mathrm{d}x + x\,\mathrm{d}y}{x^2+y^2} + 2x\,\mathrm{d}x + 2y\,\mathrm{d}y$. Su integral sobre la circunferencia unidad vale 2π, luego $c = 1$, y la diferencia $\omega - \omega_0 = \mathrm{d}(x^2+y^2)$ es exacta. La descomposición del teorema es, en la práctica, "integra sobre un lazo para leer c y luego busca potencial para el resto".

(d) Con dos agujeros. En $\mathbb{R}^2$ menos dos puntos hacen falta *dos* generadores, uno por cada agujero, y el espacio de formas cerradas módulo exactas tiene dimensión 2. En general la dimensión es el número de agujeros: eso es la cohomología de De Rham, y el enunciado es un caso particular del teorema de De Rham.

Es el primer cálculo cohomológico de la historia (implícito en Gauss y Riemann): *el espacio vectorial de las obstrucciones a integrar mide los agujeros.* Con k pinchazos, $H^1 \cong \mathbb{R}^k$ (un período por agujero); en la esfera, $H^1 = 0$ (todo cerrado es exacto: no hay dónde enrollarse);

en el toro, $H^1 \cong \mathbb{R}^2$ (dos lazos independientes). El análisis vectorial, mirado con suficiente atención, cuenta agujeros: ese es el mensaje final del capítulo.

11.13. Ronda de temarios: Reynolds y el ángulo sólido

Dos piezas que los temarios de cálculo vectorial aplicado exigen y que completan el puente hacia la mecánica de fluidos y el electromagnetismo.

11.13.1. Derivada material y teorema del transporte de Reynolds

En un fluido con campo de velocidades $\mathbf{v}(\mathbf{x},t)$, una magnitud $f(\mathbf{x},t)$ vista *desde una partícula que viaja con el fluido* cambia a razón de

$$\frac{Df}{Dt} = \frac{\partial f}{\partial t} + \mathbf{v}\cdot\nabla f$$

(regla de la cadena del capítulo 9 aplicada a $t \mapsto f(\mathbf{x}(t),t)$ con $\dot{\mathbf{x}} = \mathbf{v}$): la *derivada material*. El primer término es el cambio local; el segundo, el cambio por ser arrastrado a donde f vale otra cosa: un termómetro fijo en un río mide $\partial_t f$; uno que flota mide Df/Dt.

Teorema 11.78 (Transporte de Reynolds)**.** *Si $\Omega(t)$ es una región material (que se mueve con el fluido),*

$$\frac{\mathrm{d}}{\mathrm{d}t}\int_{\Omega(t)} f\,\mathrm{d}V = \int_{\Omega(t)} \frac{\partial f}{\partial t}\,\mathrm{d}V + \oint_{\partial\Omega(t)} f\,\mathbf{v}\cdot\mathbf{n}\,\mathrm{d}S = \int_{\Omega(t)} \Big(\frac{\partial f}{\partial t} + \operatorname{div}(f\mathbf{v})\Big)\mathrm{d}V.$$

Demostración. Trasládese todo al dominio fijo $\Omega(0)$ mediante el flujo $\boldsymbol{\varphi}_t$: por cambio de variables (capítulo 9), $\int_{\Omega(t)} f\,\mathrm{d}V = \int_{\Omega(0)} f(\boldsymbol{\varphi}_t(\mathbf{y}),t)\,J(\mathbf{y},t)\,\mathrm{d}V$ con $J = \det D\boldsymbol{\varphi}_t > 0$. El ingrediente clave es la *fórmula de Liouville* $\partial_t J = (\operatorname{div}\mathbf{v})\,J$: es exactamente el lema de Abel–Liouville del capítulo 12 (la derivada de un determinante que fluye es la traza del generador por el determinante) aplicado a la ecuación variacional $\partial_t D\boldsymbol{\varphi}_t = D\mathbf{v}\cdot D\boldsymbol{\varphi}_t$. Derivando bajo la integral (¡capítulo 6!):

$$\frac{\mathrm{d}}{\mathrm{d}t}\int_{\Omega(t)} f\,\mathrm{d}V = \int_{\Omega(0)} \Big(\frac{Df}{Dt} + f\,\operatorname{div}\mathbf{v}\Big)J\,\mathrm{d}V = \int_{\Omega(t)} \Big(\frac{Df}{Dt} + f\,\operatorname{div}\mathbf{v}\Big)\mathrm{d}V,$$

y la identidad $\frac{Df}{Dt} + f\operatorname{div}\mathbf{v} = \partial_t f + \operatorname{div}(f\mathbf{v})$ (regla del producto para la divergencia) da la tercera forma; la segunda es Gauss. □ □

Ejemplo 11.79 (La ecuación de continuidad, deducida)**.** Con $f = \rho$ (densidad): la masa de una porción material de fluido no cambia, así que $\frac{\mathrm{d}}{\mathrm{d}t}\int_{\Omega(t)} \rho\,\mathrm{d}V = 0$ para *toda* región material. Como el integrando resultante es continuo y la región es arbitraria (lema de localización: si $\int_\Omega g = 0$ para toda bola Ω y g es continua, entonces $g \equiv 0$; si $g(p) > 0$, una bolita alrededor de p daría integral positiva),

$$\frac{\partial \rho}{\partial t} + \operatorname{div}(\rho\mathbf{v}) = 0.$$

Fluido incompresible ($\frac{D\rho}{Dt} = 0$ con ρ constante): $\operatorname{div}\mathbf{v} = 0$: la condición que llevamos todo el capítulo llamando "campo sin fuentes" es exactamente la conservación de la masa. La misma

plantilla con la densidad de carga da la continuidad eléctrica (que ya vimos escondida en las ecuaciones de Maxwell), y con la cantidad de movimiento $\rho\mathbf{v}$ más las fuerzas, las ecuaciones de Euler de los fluidos: Reynolds es la máquina de fabricar leyes de conservación *locales* a partir de principios *globales*.

11.13.2. El ángulo sólido

Definición 11.80. El *ángulo sólido* que una superficie S subtiende desde el origen es

$$\Omega = \iint_S \frac{\hat{\mathbf{r}}\cdot\mathbf{n}}{r^2}\,\mathrm{d}S$$

(el área de la proyección radial de S sobre la esfera unitaria; se mide en *estereorradianes*, y la esfera completa mide 4π).

Teorema 11.81. *Para una superficie cerrada S que no pasa por el origen:*

$$\oint_S \frac{\hat{\mathbf{r}}\cdot\mathbf{n}}{r^2}\,\mathrm{d}S = \begin{cases} 4\pi & \text{si el origen está dentro,} \\ 0 & \text{si está fuera.} \end{cases}$$

Demostración. El campo $\mathbf{F} = \frac{\hat{\mathbf{r}}}{r^2}$ tiene divergencia nula fuera del origen (cálculo hecho en este capítulo). Si el origen queda fuera, Gauss da 0. Si queda dentro, exclúyase una bolita B_ε: en la región intermedia Gauss da $\oint_S = \oint_{\partial B_\varepsilon} = \frac{1}{\varepsilon^2}\cdot 4\pi\varepsilon^2 = 4\pi$. Es el mismo argumento "residual" con que calculamos $H^1(\mathbb{R}^2\setminus\{0\})$: la singularidad deja una huella que ninguna deformación borra, un capítulo antes de llamarla residuo en el 9. □ □

Idea clave

Qué dice esto en cristiano. El ángulo sólido mide qué porción del cielo ocupa una superficie vista desde un punto, igual que el ángulo ordinario mide qué porción de la circunferencia ocupa un arco. Se mide en estereorradianes y el cielo entero son 4π. El teorema dice algo muy limpio: el flujo del campo radial inverso al cuadrado a través de una superficie cerrada vale 4π si el origen está dentro y 0 si está fuera. Esa dicotomía es la ley de Gauss del electromagnetismo y de la gravitación.

Ejemplo 11.82 (Contar el cielo)**. (a) Los casos básicos.** Toda la esfera subtiende $4\pi = 12{,}566371$ estereorradianes; un hemisferio, $2\pi = 6{,}283185$; y un octante, $\frac{\pi}{2} = 1{,}570796$.

(b) Un cono. El ángulo sólido de un cono de semiángulo α es $2\pi(1-\cos\alpha)$. Para $\alpha = 60°$ da $2\pi\left(1-\frac{1}{2}\right) = \pi = 3{,}141593$, la cuarta parte del cielo. Para α pequeño se aproxima por $\pi\alpha^2$, que es el área del disco correspondiente.

(c) Dentro o fuera, la dicotomía. Para una superficie cerrada que encierre al origen, el flujo de $\frac{\hat{\mathbf{r}}}{r^2}$ vale 4π sea cual sea la forma: una esfera, un cubo o una patata deformada dan lo mismo. Si el origen queda fuera, el flujo es exactamente 0, porque la divergencia es nula en todo el interior y se aplica Gauss sin problemas.

(d) Por qué esto es la ley de Gauss. El campo eléctrico de una carga puntual es $\frac{q}{4\pi\varepsilon_0}\frac{\hat{\mathbf{r}}}{r^2}$, luego su flujo por cualquier superficie cerrada vale $\frac{q}{\varepsilon_0}$ si la carga está dentro y cero si está

fuera. La ley de Gauss no es un hecho experimental independiente: es geometría del inverso del cuadrado, y la misma cuenta vale para la gravitación de Newton.

Cuidado

[El teorema de la divergencia no se aplica directamente si el origen está dentro] Cuando la carga está dentro, el campo no es C^1 en todo el sólido (explota en el origen), así que Gauss no es aplicable tal cual. La demostración correcta rodea el origen con una esfera pequeña, aplica Gauss en la región entre ambas superficies (donde la divergencia sí es cero) y calcula directamente el flujo por la esfera interior, que vale 4π sin importar su radio. Segundo aviso: la independencia respecto del radio es una peculiaridad del exponente -2; con $\frac{\hat{\mathbf{r}}}{r^3}$ el flujo dependería del tamaño de la esfera y no habría ley de Gauss.

Observación 11.83. Esto *es* la ley de Gauss del electromagnetismo: el flujo de $\mathbf{E} = \frac{q}{4\pi\varepsilon_0}\frac{\hat{\mathbf{r}}}{r^2}$ por una superficie cerrada vale $\frac{q}{\varepsilon_0}$ o 0 según la carga esté dentro o fuera, *sin importar la forma de la superficie*: toda la utilidad práctica de la ley (esferas, cilindros, placas) es esta libertad. En versión abierta, el ángulo sólido calcula flujos sin integrar: el flujo por un casquete es $\frac{q}{4\pi\varepsilon_0}\,\Omega$. Y mide cuánto cielo tapa un objeto: la Luna subtiende $\Omega \approx 6\times10^{-5}$ sr —y sin embargo tapa exactamente al Sol en un eclipse: mismo Ω desde la Tierra, pura coincidencia cósmica—.

11.14. Cierre definitivo: las identidades de Green

Las dos identidades de Green son "integración por partes en varias variables": el dividendo más rentable del teorema de la divergencia, y la herramienta con que la teoría del potencial demuestra sus teoremas de unicidad. Con ellas el capítulo no deja promesa sin pagar.

Teorema 11.84 (Identidades de Green)**.** *Para u, v suficientemente suaves en Ω con frontera $\partial\Omega$ (normal exterior $\mathbf{n}$, derivada normal $\partial_n u = \nabla u\cdot\mathbf{n}$):*

$$\textit{(I)}\quad \int_\Omega (u\,\Delta v + \nabla u\cdot\nabla v)\,\mathrm{d}V = \oint_{\partial\Omega} u\,\partial_n v\,\mathrm{d}S, \qquad \textit{(II)}\quad \int_\Omega (u\,\Delta v - v\,\Delta u)\,\mathrm{d}V = \oint_{\partial\Omega} (u\,\partial_n v - v\,\partial_n u)\,\mathrm{d}S.$$

Demostración. (I) es la divergencia aplicada al campo $u\nabla v$: $\operatorname{div}(u\nabla v) = u\Delta v + \nabla u\cdot\nabla v$ (regla del producto). (II) es (I) menos (I) con los papeles cambiados. Compárese con $\int uv' + \int u'v = [uv]$: son la misma fórmula con Δ en el papel de la segunda derivada y la frontera en el papel de los extremos. □ □

Ejemplo 11.85 (Las identidades de Green y para qué sirven)**. (a) Integración por partes en varias variables.** La primera identidad, $\iiint_\Omega (u\Delta v + \nabla u\cdot\nabla v) = \iint_{\partial\Omega} u\,\partial_n v$, es literalmente la fórmula $\int uv'' = -\int u'v' + [uv']$ del cálculo de una variable, escrita en dimensión n. Todo lo que se hace con ella es lo mismo que se hacía con aquélla: pasar derivadas de un factor al otro.

(b) Energía y unicidad. Tomando $u = v$ armónica y nula en la frontera, la primera identidad da $\iiint_\Omega \|\nabla u\|^2 = 0$, luego $\nabla u \equiv 0$ y u es constante, y como vale cero en el borde, $u \equiv 0$. Ésta es la demostración de unicidad del problema de Dirichlet, y cabe en dos renglones.

(c) Una condición de compatibilidad. Con $v = 1$ en la segunda identidad sale $\iiint_\Omega \Delta u = \iint_{\partial\Omega} \partial_n u$. Por tanto el problema de Neumann $\Delta u = f$ con $\partial_n u = g$ sólo puede tener solución si $\iiint f = \iint g$: si se calienta un cuerpo aislado, la temperatura no puede quedarse estacionaria. Es una obstrucción física evidente que aquí aparece como un cálculo.

(d) La solución fundamental. Aplicando la segunda identidad con $v = 1/(4\pi\|x - x_0\|)$, que es armónica salvo en x_0, se obtiene la representación integral que expresa $u(x_0)$ en función de sus valores y derivadas en la frontera. De ahí salen la función de Green y, con ella, la fórmula de Poisson para el disco.

Teorema 11.86 (Unicidad del problema de Dirichlet)**.** *Dos soluciones de $\Delta u = f$ en Ω con los mismos valores de frontera coinciden.*

Demostración. La diferencia w cumple $\Delta w = 0$ y $w = 0$ en $\partial\Omega$. Green (I) con $u = v = w$:

$$\int_\Omega \|\nabla w\|^2 \,\mathrm{d}V = \oint_{\partial\Omega} w\, \partial_n w \,\mathrm{d}S = 0,$$

así que $\nabla w \equiv 0$: w es constante (corolario del valor medio, capítulo 9), y en la frontera vale 0: $w \equiv 0$. □ □

Ejemplo 11.87 (Unicidad, existencia y sus límites)**. (a) La cuenta.** Si u_1, u_2 resuelven el mismo problema, $w = u_1 - u_2$ cumple $\Delta w = 0$ y $w = 0$ en $\partial\Omega$. Por el argumento de energía del apartado anterior, $w \equiv 0$.

(b) Unicidad no es existencia. Este teorema no dice que haya solución. La existencia para dominios razonables la dieron Perron y otros, y para dominios malos puede fallar: el ejemplo clásico es el disco punteado $\{0 < \|x\| < 1\}$ con dato 0 en la circunferencia exterior y 1 en el origen, donde no hay ninguna función armónica acotada que cumpla ambas cosas, porque un punto aislado no es "visible" para el laplaciano en dimensión 2.

(c) Neumann es distinto. Con datos de Neumann la unicidad sólo vale salvo constante aditiva: si u resuelve $\Delta u = f$ con $\partial_n u = g$, también lo hace $u + 7$. Físicamente es obvio: fijar los flujos de calor por el borde no determina el cero de la escala de temperaturas.

(d) El método numérico. La unicidad es lo que legitima resolver el problema discretizando: si el esquema en diferencias finitas converge a *alguna* solución, esa es *la* solución. Sin unicidad, un método numérico podría dar una respuesta correcta y otro método otra respuesta también correcta, y no habría manera de compararlos.

Observación 11.88. El mismo cálculo con condiciones de Neumann ($\partial_n u$ dado) da unicidad *salvo constante* —físicamente obvio: fijar el flujo de calor por las paredes determina las temperaturas salvo el cero de la escala— y de paso la *condición de compatibilidad* $\oint \partial_n u \,\mathrm{d}S = \int_\Omega f \,\mathrm{d}V$: no se puede pedir cualquier flujo (el calor que entra debe igualar al que se genera: Green (I) con $u = 1$). Este argumento de "energía" $\int \|\nabla w\|^2$ es la plantilla de la mitad de las demostraciones de unicidad de las EDP: calor, onda, elasticidad, Navier–Stokes; el capítulo 12 lo usó en versión unidimensional para Sturm–Liouville sin decir su nombre.

Teorema 11.89 (Propiedad del valor medio, demostrada en 3D)**.** *Si $\Delta u = 0$ en una bola $B_R(\mathbf{p})$, entonces $u(\mathbf{p})$ es el promedio de u sobre cada esfera $\partial B_r(\mathbf{p})$, $r < R$.*

Demostración. Sea $\phi(r) = \frac{1}{4\pi r^2}\oint_{\partial B_r} u\,\mathrm{d}S = \frac{1}{4\pi}\oint_{\|\boldsymbol{\omega}\|=1} u(\mathbf{p}+r\boldsymbol{\omega})\,\mathrm{d}\omega$ (promedio en la esfera unitaria: el cambio de escala elimina el r^2). Derivando bajo la integral,

$$\phi'(r) = \frac{1}{4\pi}\oint \nabla u\cdot\boldsymbol{\omega}\,\mathrm{d}\omega = \frac{1}{4\pi r^2}\oint_{\partial B_r}\partial_n u\,\mathrm{d}S = \frac{1}{4\pi r^2}\int_{B_r}\Delta u\,\mathrm{d}V = 0$$

(divergencia en el penúltimo paso). Así ϕ es constante, y $\phi(0^+) = u(\mathbf{p})$ por continuidad. □ □

Idea clave

Qué dice esto en cristiano. Una función armónica es la que cumple $\Delta u = 0$, y su propiedad definitoria es de una sencillez sorprendente: el valor en cada punto es exactamente el promedio de los valores en cualquier esfera centrada ahí. De eso se deduce todo lo demás. Como ningún punto interior puede superar a su propio promedio, los máximos y mínimos tienen que estar en la frontera; y como una función armónica que vale cero en la frontera vale cero dentro, dos soluciones con los mismos datos de frontera coinciden. Es el modelo del equilibrio térmico: la temperatura de un punto es la media de la de sus vecinos.

Ejemplo 11.90 (Verificar la propiedad del valor medio)**. (a) Una armónica sencilla.** Sea $u = x^2 - y^2$. Su laplaciano es $2 + (-2) = 0$: es armónica. En el círculo de radio 1 centrado en el origen, el promedio de u es $\frac{1}{2\pi}\int_0^{2\pi}(\cos^2 t - \sin^2 t)\,\mathrm{d}t = 0$, que es exactamente $u(0,0)$.

(b) El máximo está en el borde. En el disco unidad, esa misma u vale 1 en $(1,0)$, -1 en $(0,1)$ y 0 en el centro. Su máximo y su mínimo están ambos en la circunferencia frontera, como exige el principio del máximo. Una función armónica no constante no puede tener un pico interior.

(c) La unicidad, en una línea. Si u_1 y u_2 son armónicas y coinciden en la frontera, la diferencia $u = u_1 - u_2$ es armónica y vale cero en la frontera. Por el principio del máximo, su máximo y su mínimo están en la frontera, luego valen cero, luego $u \equiv 0$. La demostración cabe en tres renglones y es el argumento estándar de unicidad en ecuaciones en derivadas parciales.

(d) La lectura física. En equilibrio térmico, la temperatura es armónica. La propiedad del valor medio dice que la temperatura de cada punto es el promedio de la de su entorno; el principio del máximo dice que el punto más caliente de una placa en equilibrio está siempre en el borde, nunca dentro. Si estuviera dentro, ese punto sería más caliente que su propio promedio, y el calor fluiría desde él, rompiendo el equilibrio.

Cuidado

[La unicidad exige un dominio acotado] En un dominio no acotado el principio del máximo puede fallar: la función $u = y$ es armónica en el semiplano superior, vale 0 en toda la frontera $y = 0$ y no es idénticamente nula. Hace falta añadir una condición de comportamiento en el infinito. Segundo aviso: la propiedad del valor medio *caracteriza* a las armónicas; si una función continua cumple la propiedad para todas las esferas pequeñas, entonces es armónica y además automáticamente C^∞. Y tercero: en dimensión uno las armónicas son solo las rectas, así que toda la riqueza del tema aparece a partir de dimensión dos.

Observación 11.91. Con esto queda demostrado lo que el capítulo 12 enunció y el capítulo 30 probó en el disco plano: ahora en cualquier dimensión. Y cae gratis el *principio del máximo fuerte*: una armónica que alcanza su máximo en un punto interior es constante cerca (es promedio de sus vecinos: si iguala al máximo, todos los vecinos lo igualan), y por conexidad, constante en todo el dominio. Temperatura estacionaria sin fuentes: los extremos viven en la frontera, siempre. La cadena completa —divergencia → Green → valor medio → máximo → unicidad— es la columna vertebral de la teoría del potencial, y cada eslabón quedó demostrado en este libro.

11.15. Segunda ronda: el planímetro, el dipolo y el generador

11.15.1. El planímetro: Green hecho latón

El planímetro de Amsler (1854): dos brazos y una ruedita que, recorriendo el borde, marca el área. La ruedita integra $\oint \mathbf{F}\cdot \mathrm{d}\mathbf{r}$ para un campo con $\operatorname{rot}\mathbf{F}\equiv 1$ impuesto por la geometría del brazo: por Green, la lectura *es* $\iint 1\,\mathrm{d}A$. Un siglo de agrimensores y cardiólogos midiendo áreas con el teorema de Green sin saberlo.

11.15.2. El dipolo, trabajado

Cargas $\pm q$ a distancia $\mathrm{d}\to 0$ con $p=q\mathrm{d}$: el potencial es la derivada direccional del de Coulomb:

$$\varphi_{\mathrm{dip}}=\frac{p\cos\theta}{4\pi\varepsilon_0 r^2}\,:$$

el monopolo cancelado deja la siguiente derivada (el multipolar de Legendre, capítulo 12). El agua es un dipolo (así calienta el microondas), las antenas radian con este $\sin\theta$, y el flujo por cualquier esfera que encierre ambas cargas es 0: Gauss cuenta carga neta.

11.15.3. Faraday cuantitativo

Espira de área A girando a ω en campo B: $\Phi=BA\cos\omega t$ y

$$\mathcal{E}=-\dot{\Phi}=BA\omega\sin\omega t\,:$$

la corriente alterna. El signo de Lenz garantiza oposición (sin él, energía gratis); todo el parque eléctrico mundial es esta derivada a 50–60 Hz.

11.16. Coda: potenciales retardados

Con fuentes que dependen del tiempo, las ecuaciones de los potenciales electromagnéticos se vuelven ondas forzadas $\Box\varphi=\rho/\varepsilon_0$, y su solución es la fórmula del *potencial retardado*:

$$\varphi(\mathbf{r},t)=\frac{1}{4\pi\varepsilon_0}\iiint\frac{\rho\big(\mathbf{r}',\,t-\frac{\|\mathbf{r}-\mathbf{r}'\|}{c}\big)}{\|\mathbf{r}-\mathbf{r}'\|}\,\mathrm{d}V'\,:$$

el potencial de Coulomb del capítulo, pero evaluando la fuente *en el pasado justo*: la noticia viaja a velocidad c. Vemos el Sol como era hace 8 minutos y las galaxias como eran hace

millones de años: la astronomía es arqueología por potenciales retardados. De esta fórmula, dipolos oscilantes radian (las antenas), con la potencia de Larmor $\sim \ddot{p}^2$: la razón por la que el cielo es azul (la luz azul, de mayor frecuencia, se re-radia más: ω^4 de Rayleigh) cabe en esta coda.

11.17. Temas extra y ejemplos desarrollados

11.17.1. Trabajo–energía y conservación

$\int \mathbf{F}\cdot \mathrm{d}\mathbf{r} = \Delta(\frac{mv^2}{2})$ (cadena + TFC); con $\mathbf{F} = -\nabla U$: $\frac{mv^2}{2} + U = \text{cte}$. La física de bachillerato con cimientos.

11.17.2. Kepler II para toda fuerza central

$\dot{\mathbf{L}} = \mathbf{r} \times f(r)\hat{\mathbf{r}} = \mathbf{0}$: órbita plana y $\dot{A} = \frac{L}{2m}$ constante: áreas iguales en tiempos iguales, para *cualquier* fuerza central.

11.17.3. Deducción de la ecuación del calor

Balance en E arbitrario + ley de Fourier $\mathbf{q} = -k\nabla u$ + Gauss: $c\rho u_t = k\Delta u$. El mismo molde: difusión de Fick, Black–Scholes, Schrödinger.

11.17.4. Pappus–Guldin

$S = 2\pi \bar{y}\, \ell$, $V = 2\pi \bar{y}\, A$. El toro sin integrales: $V = 2\pi^2 R r^2$, $S = 4\pi^2 R r$.

11.17.5. Vorticidad y Kelvin

$\boldsymbol{\omega} = \operatorname{rot} \mathbf{v}$; circulación sobre lazos materiales se conserva (fluido ideal): anillos de humo persistentes, y la sustentación de un ala es una circulación ($L = \rho v \Gamma$, Kutta–Yukovski): los aviones vuelan sobre una integral de línea.

11.18. Líneas de campo, el flujo φ_t y la divergencia como dilatación de volumen

La demostración del teorema del transporte usaba "el flujo $\boldsymbol{\varphi}_t$" sin haberlo definido. Es un objeto sencillo y merece sección propia, porque convierte la divergencia de fórmula en significado.

Definición 11.92 (Curva integral y flujo)**.** Una **curva integral** de $\mathbf{F}$ es una solución de $\dot{\mathbf{r}}(t) = \mathbf{F}(\mathbf{r}(t))$. Si $\mathbf{F}$ es de clase C^1, por cada punto pasa una y solo una —es Picard—, y se define el **flujo**

$$\boldsymbol{\varphi}_t(\mathbf{x}_0) = \mathbf{r}(t) \text{ con } \mathbf{r}(0) = \mathbf{x}_0.$$

Cumple $\boldsymbol{\varphi}_0 = \mathrm{id}$ y $\boldsymbol{\varphi}_{t+s} = \boldsymbol{\varphi}_t \circ \boldsymbol{\varphi}_s$: es un grupo uniparamétrico de difeomorfismos.

Teorema 11.93 (Fórmula de Liouville)**.** *Sea* $J(t) = \det D\boldsymbol{\varphi}_t$*. Entonces*

$$\frac{\partial J}{\partial t} = (\operatorname{div}\mathbf{F})\, J, \qquad \textit{luego} \qquad J(t) = \exp\Big(\int_0^t \operatorname{div}\mathbf{F}\,\mathrm{d}s\Big).$$

Demostración. Derivando la ecuación del flujo respecto de la condición inicial, $\partial_t D\boldsymbol{\varphi}_t = D\mathbf{F}\cdot D\boldsymbol{\varphi}_t$. La derivada del determinante es $\partial_t \det M = \det M \cdot \operatorname{tr}(M^{-1}\partial_t M)$, y aquí $M^{-1}\partial_t M = D\boldsymbol{\varphi}_t^{-1} D\mathbf{F}\, D\boldsymbol{\varphi}_t$, que tiene la misma traza que $D\mathbf{F}$, es decir, $\operatorname{div}\mathbf{F}$. □

Idea clave

Esa es *la* definición conceptual de la divergencia: es la tasa relativa de cambio de volumen a lo largo del flujo. Un campo con divergencia nula transporta las regiones deformándolas pero conservando su volumen exactamente; uno con divergencia positiva las dilata. Todo lo que el capítulo dice sobre fuentes y sumideros deja de ser metáfora.

Ejemplo 11.94 (Dos campos, dos comportamientos)**. (a)** $\mathbf{F} = (x, -y)$**.** La ecuación se separa: $\dot{x} = x$ y $\dot{y} = -y$, luego

$$\boldsymbol{\varphi}_t(x_0, y_0) = (x_0 e^{t},\ y_0 e^{-t}), \qquad J = e^{t}\cdot e^{-t} = 1.$$

Y en efecto $\operatorname{div}\mathbf{F} = 1 - 1 = 0$. Las líneas de campo son las hipérbolas $xy = \text{cte}$: un cuadrado se estira en horizontal y se aplasta en vertical, y su área no cambia ni un ápice. Es un flujo incompresible con deformación pura.

(b) $\mathbf{F} = (x, y)$**.** Ahora $\boldsymbol{\varphi}_t(x_0, y_0) = e^{t}(x_0, y_0)$ y $J = e^{2t}$, con $\operatorname{div}\mathbf{F} = 2$. Se comprueba $\partial_t J = 2e^{2t} = 2J$ ✓. Las líneas de campo son las rectas por el origen y todo se expande: hay una fuente en el origen.

Observación 11.95. De aquí sale de inmediato el teorema del transporte de Reynolds, ahora sí justificado:

$$\frac{\mathrm{d}}{\mathrm{d}t}\int_{\Omega(t)} f = \int_{\Omega(t)} \Big(\frac{\partial f}{\partial t} + \operatorname{div}(f\mathbf{v})\Big),$$

porque trasladar la integral al dominio fijo introduce el factor J y derivar produce el término $\operatorname{div}\mathbf{v}$ vía Liouville. Y de ahí, con $f = \rho$, la ecuación de continuidad. En mecánica hamiltoniana, el mismo teorema aplicado a $\mathbf{F} = (\partial_p H, -\partial_q H)$, que tiene divergencia idénticamente nula, da la conservación del volumen en el espacio de fases: el teorema de Liouville de la física estadística es este mismo enunciado.

11.19. Solución fundamental, fórmula de representación y la descomposición de Helmholtz

El teorema de Helmholtz se enunció en dos frases: todo campo se descompone en una parte que solo tiene divergencia y otra que solo tiene rotacional. Falta *construirlas*, y para eso hace falta saber invertir el laplaciano.

Teorema 11.96 (Solución fundamental)**.** *En $\mathbb{R}^3$, la función $\Phi(\mathbf{x}) = \dfrac{1}{4\pi|\mathbf{x}|}$ cumple $\Delta\Phi = -\delta_0$. Concretamente, es armónica fuera del origen y, para toda ψ suave de soporte compacto,*

$$\int_{\mathbb{R}^3} \Phi\, \Delta\psi = -\psi(0).$$

Con la bola excluida. $\Delta(1/r) = 0$ para $r > 0$: en coordenadas esféricas $\Delta u = \frac{1}{r^2}(r^2u')'$, y con $u = 1/r$ resulta $r^2u' = -1$, constante, cuya derivada es cero. Para el comportamiento en el origen se aplica la segunda identidad de Green en $\{\varepsilon < |\mathbf{x}| < R\}$; el término de volumen se anula, y en la esfera pequeña queda

$$\int_{|\mathbf{x}|=\varepsilon} \Big(\Phi \frac{\partial\psi}{\partial\nu} - \psi \frac{\partial\Phi}{\partial\nu}\Big)\, \mathrm{d}S.$$

El primer sumando es $O(\varepsilon)$; el segundo, como $\partial_\nu\Phi = 1/(4\pi\varepsilon^2)$ y la esfera tiene área $4\pi\varepsilon^2$, tiende exactamente a $-\psi(0)$. □

Teorema 11.97 (Potencial newtoniano)**.** *Si ρ tiene soporte compacto, la función*

$$\varphi(\mathbf{x}) = \int_{\mathbb{R}^3} \frac{\rho(\mathbf{y})}{4\pi|\mathbf{x}-\mathbf{y}|}\, \mathrm{d}\mathbf{y}$$

resuelve $-\Delta\varphi = \rho$, y es la única solución que tiende a 0 en el infinito.

Ejemplo 11.98 (La bola uniforme, dentro y fuera)**.** Sea una bola de radio a con densidad constante y carga total Q, es decir $\rho = \dfrac{3Q}{4\pi a^3}$. Su potencial es

$$\varphi(r) = \begin{cases} \dfrac{Q(3a^2 - r^2)}{8\pi a^3} & r \le a, \\[2ex] \dfrac{Q}{4\pi r} & r \ge a. \end{cases}$$

Comprobación fuera: $\Delta\varphi = \frac{1}{r^2}(r^2\varphi')'$ con $\varphi' = -Q/(4\pi r^2)$ da $r^2\varphi' = -Q/(4\pi)$, constante, luego $\Delta\varphi = 0$. ✓ Eso es el teorema de la cáscara de Newton: desde fuera, una bola atrae como si toda su masa estuviera en el centro.

Comprobación dentro: $\varphi' = -Qr/(4\pi a^3)$, luego $r^2\varphi' = -Qr^3/(4\pi a^3)$ y

$$\Delta\varphi = \frac{1}{r^2}\cdot\Big(-\frac{3Qr^2}{4\pi a^3}\Big) = -\frac{3Q}{4\pi a^3} = -\rho. ✓$$

Y en $r = a$ las dos expresiones y sus derivadas coinciden —ambas dan $Q/(4\pi a)$ y $-Q/(4\pi a^2)$—, así que φ es de clase C^1 pese a que ρ salta.

Teorema 11.99 (Descomposición de Helmholtz)**.** *Todo campo $\mathbf{F}$ de clase C^2 que decaiga suficientemente en el infinito se escribe de manera única como*

$$\mathbf{F} = -\nabla\varphi + \operatorname{rot}\mathbf{A},$$

con

$$\varphi(\mathbf{x}) = \int \frac{\operatorname{div}\mathbf{F}(\mathbf{y})}{4\pi|\mathbf{x}-\mathbf{y}|}\mathrm{d}\mathbf{y}, \qquad \mathbf{A}(\mathbf{x}) = \int \frac{\operatorname{rot}\mathbf{F}(\mathbf{y})}{4\pi|\mathbf{x}-\mathbf{y}|}\mathrm{d}\mathbf{y}.$$

Idea. Aplíquese la identidad $\Delta\mathbf{F} = \nabla(\operatorname{div}\mathbf{F}) - \operatorname{rot}(\operatorname{rot}\mathbf{F})$ e invíértase el laplaciano componente a componente con el potencial newtoniano. La unicidad sale de que la diferencia de dos descomposiciones sería un campo armónico que decae, y por Liouville es nulo. □

Ejemplo 11.100 (Un campo de cizalla, descompuesto)**.** Sea $\mathbf{F} = (y, 0, 0)$: el flujo de un fluido entre dos placas. Entonces

$$\operatorname{div}\mathbf{F} = 0, \qquad \operatorname{rot}\mathbf{F} = (0, 0, -1).$$

La parte gradiente es nula —no hay fuentes— y todo el campo es rotacional. Un potencial vector explícito es

$$\mathbf{A} = \Big(0,\ 0,\ \tfrac{y^2}{2}\Big), \qquad \text{pues} \qquad \operatorname{rot}\mathbf{A} = \Big(\frac{\partial}{\partial y}\frac{y^2}{2},\ 0,\ 0\Big) = (y, 0, 0) = \mathbf{F}.\checkmark$$

La cizalla pura es, pese a las apariencias, un campo enteramente de vorticidad: cada partícula gira aunque las trayectorias sean rectas. Ese es el contenido físico de $\operatorname{rot}\mathbf{F} \neq 0$ con trayectorias rectilíneas, y la razón de que una ruedecilla puesta en el flujo empiece a girar.

Idea clave

Con esto, la lectura de las ecuaciones de Maxwell queda cerrada. Helmholtz dice que un campo está *determinado* por su divergencia y su rotacional (más las condiciones en el infinito). Maxwell prescribe exactamente esas cuatro cosas para $\mathbf{E}$ y $\mathbf{B}$:

$$\operatorname{div}\mathbf{E} = \rho/\epsilon_0, \quad \operatorname{rot}\mathbf{E} = -\partial_t\mathbf{B}, \quad \operatorname{div}\mathbf{B} = 0, \quad \operatorname{rot}\mathbf{B} = \mu_0\mathbf{J} + \mu_0\epsilon_0\partial_t\mathbf{E}.$$

No son cuatro ecuaciones sueltas: son la lista completa de lo que hay que decir de un campo vectorial para determinarlo. Y $\operatorname{div}\mathbf{B} = 0$ es lo que garantiza que exista el potencial vector $\mathbf{A}$ con $\mathbf{B} = \operatorname{rot}\mathbf{A}$, del que sale Biot–Savart al aplicarle el potencial newtoniano.

11.20. Homotopía y campos conservativos: la implicación que faltaba

El criterio que el capítulo manda comprobar en cada ejercicio —irrotacional más dominio simplemente conexo implica conservativo— se despachó en una frase. El lema de Poincaré demostrado sólo cubría dominios estrellados. Aquí está el argumento completo.

Definición 11.101 (Caminos homótopos)**.** Dos caminos $\gamma_0, \gamma_1 : [0,1] \to \Omega$ con los mismos extremos son *homótopos en* Ω si existe una aplicación continua $H : [0,1] \times [0,1] \to \Omega$ con

$$H(t,0) = \gamma_0(t), \quad H(t,1) = \gamma_1(t), \quad H(0,s) = \gamma_0(0), \quad H(1,s) = \gamma_0(1) \ \text{ para todo } s.$$

El dominio Ω es *simplemente conexo* si es conexo por caminos y todo lazo es homótopo a un punto.

Teorema 11.102 (Invariancia homotópica de la integral de línea)**.** *Sea $\vec{F}$ un campo de clase C^1 e irrotacional ($\operatorname{rot}\vec{F}=0$) en un abierto $\Omega\subset\mathbb{R}^3$. Si γ_0 y γ_1 son homótopos en Ω mediante una homotopía H de clase C^2, entonces*

$$\int_{\gamma_0}\vec{F}\cdot\mathrm{d}\vec{r}=\int_{\gamma_1}\vec{F}\cdot\mathrm{d}\vec{r}.$$

Demostración. Defínase, para cada $s\in[0,1]$,

$$I(s)=\int_0^1\vec{F}(H(t,s))\cdot\frac{\partial H}{\partial t}(t,s)\,\mathrm{d}t,$$

que es la integral de línea sobre el camino intermedio $\gamma_s=H(\cdot,s)$. Basta ver que $I'(s)=0$. Derivando bajo el signo integral,

$$I'(s)=\int_0^1\Big[\Big((\partial_sH\cdot\nabla)\vec{F}\Big)\cdot\partial_tH+\vec{F}\cdot\partial_s\partial_tH\Big]\,\mathrm{d}t=\int_0^1\left[\sum_{i,j}\frac{\partial F_i}{\partial x_j}\,\partial_sH_j\,\partial_tH_i+\vec{F}\cdot\partial_{ts}H\right]\mathrm{d}t.$$

Por otro lado, considérese la función auxiliar $g(t,s)=\vec{F}(H(t,s))\cdot\partial_sH(t,s)$, cuya derivada respecto de t es

$$\partial_tg=\sum_{i,j}\frac{\partial F_i}{\partial x_j}\,\partial_tH_j\,\partial_sH_i+\vec{F}\cdot\partial_{ts}H.$$

Réstense ambas expresiones. Los segundos sumandos se cancelan (las derivadas cruzadas coinciden por ser H de clase C^2), y en los primeros queda

$$I'(s)-\int_0^1\partial_tg\,\mathrm{d}t=\int_0^1\sum_{i,j}\left(\frac{\partial F_i}{\partial x_j}-\frac{\partial F_j}{\partial x_i}\right)\partial_sH_j\,\partial_tH_i\,\mathrm{d}t=0,$$

porque el paréntesis es, componente a componente, el rotacional de $\vec{F}$, que se anula por hipótesis. Luego

$$I'(s)=\int_0^1\partial_tg\,\mathrm{d}t=g(1,s)-g(0,s).$$

Pero los extremos están fijos: $H(0,s)$ y $H(1,s)$ no dependen de s, luego $\partial_sH(0,s)=\partial_sH(1,s)=0$ y ambos términos se anulan. Así $I'\equiv0$ e $I(0)=I(1)$. □

Corolario 11.103 (El criterio, ahora demostrado)**.** *Sea $\vec{F}$ de clase C^1 en un abierto Ω* simplemente conexo*. Entonces son equivalentes:*

1. $\operatorname{rot}\vec{F}=0$ *en Ω;*
2. $\oint_\gamma\vec{F}\cdot\mathrm{d}\vec{r}=0$ *para todo lazo γ en Ω;*
3. $\vec{F}=\nabla f$ *para alguna función f definida en todo Ω.*

Demostración. (3)⇒(1) es el hecho de que $\operatorname{rot}\nabla f=0$, o sea la igualdad de las derivadas cruzadas. *(3)⇒(2)* es el teorema fundamental para integrales de línea. *(1)⇒(2)*: por simple conexión, todo lazo es homótopo a un punto, y por el teorema 11.102 su integral coincide con la integral sobre el camino constante, que es 0. *(2)⇒(3)*: fíjese $p_0\in\Omega$ y defínase $f(p)=\int_{p_0}^{p}\vec{F}\cdot\mathrm{d}\vec{r}$ a lo largo de cualquier camino; la hipótesis (2) garantiza que el valor no depende del camino, y derivando en la dirección $\vec{e}_i$ mediante un segmento se obtiene $\partial_if=F_i$. □

Cuidado

[Sin simple conexión, todo se cae] El campo del capítulo

$$\vec{F} = \left(\frac{-y}{x^2+y^2},\ \frac{x}{x^2+y^2},\ 0 \right)$$

es irrotacional en $\Omega = \mathbb{R}^3 \setminus \{\text{eje } z\}$, y sin embargo su integral sobre la circunferencia unidad vale $2\pi \neq 0$: no es conservativo. No hay contradicción, porque ese Ω no es simplemente conexo —la circunferencia no se puede contraer sin cruzar el eje—. El número 2π no es un accidente: dividido entre 2π es el *índice* de la curva alrededor del eje, el mismo invariante que en el capítulo 30 gobierna el teorema de los residuos, y en el capítulo 37 genera el grupo fundamental del plano perforado. Las tres apariciones son el mismo objeto visto con tres lenguajes.

11.21. Tres definiciones que los teoremas daban por sabidas: orientación, campos solenoidales y conexión simple

Los tres grandes teoremas de este capítulo llevan hipótesis que hemos usado como si fueran obvias: "superficie orientable", "con la orientación inducida", "en un dominio simplemente conexo". Ninguna de esas expresiones estaba definida en su momento, y las tres son exactamente el lugar por donde los teoremas se rompen cuando se rompen. Esta sección las define y, en cada caso, exhibe el contraejemplo que muestra que la hipótesis no sobra.

11.21.1. Superficies orientables y orientación inducida en el borde

Definición 11.104 (Superficie orientable, orientación y orientación inducida)**.** Una superficie regular $S \subset \mathbb{R}^3$ es *orientable* si admite un campo continuo de vectores normales unitarios, es decir, una aplicación continua $\mathbf{N} : S \to \mathbb{R}^3$ con $\|\mathbf{N}(p)\| = 1$ y $\mathbf{N}(p) \perp T_pS$ para todo $p \in S$. Cada una de esas aplicaciones (hay exactamente dos en cada componente conexa, $\mathbf{N}$ y $-\mathbf{N}$) se llama una *orientación* de S; una superficie con una orientación elegida se dice *orientada*.

Si S es orientada y tiene borde ∂S, la *orientación inducida* en el borde es la que cumple la *regla de la mano derecha*: recorriendo ∂S en el sentido positivo, con la cabeza apuntando en el sentido de $\mathbf{N}$, la superficie queda a la izquierda. Dicho con vectores: si $\mathbf{T}$ es el tangente unitario al borde en el sentido positivo y $\boldsymbol{\nu}$ es el vector unitario tangente a S, normal a ∂S y apuntando hacia *fuera* de S, entonces $\{\boldsymbol{\nu}, \mathbf{T}, \mathbf{N}\}$ debe ser una base positivamente orientada de $\mathbb{R}^3$, es decir, $\boldsymbol{\nu} \times \mathbf{T} = \mathbf{N}$.

La superficie no orientable de referencia es la *banda de Möbius*: tómese un rectángulo y péguense dos lados opuestos tras darle media vuelta a uno de ellos. Empiécese en un punto con una normal elegida y transpórtese con continuidad a lo largo de la curva central; al volver al punto de partida, después de recorrer la banda entera, la normal apunta al revés. No hay contradicción con la continuidad —el transporte fue continuo en todo momento—, pero sí con la existencia de un campo normal *global*: la única forma de que $\mathbf{N}$ sea continua en todo el

cierre del recorrido sería $\mathbf{N} = -\mathbf{N}$, imposible para un vector unitario. La banda de Möbius no tiene dos lados: tiene uno.

Cuidado

[Qué falla exactamente sin orientabilidad] El teorema de Stokes afirma que $\iint_S \operatorname{rot}\mathbf{F}\cdot\mathbf{N}\,\mathrm{d}S = \oint_{\partial S}\mathbf{F}\cdot\mathrm{d}\mathbf{r}$ para S orientable con la orientación inducida en el borde. Si S no es orientable, el fallo no es que la igualdad sea falsa: es que **el miembro izquierdo no está definido**. Escribir $\mathbf{N}$ presupone haber elegido un lado, y en la banda de Möbius esa elección no existe globalmente; el integrando $\operatorname{rot}\mathbf{F}\cdot\mathbf{N}$ cambiaría de signo al dar la vuelta, y la integral dependería de por dónde se empiece a parametrizar. Lo mismo ocurre con el flujo de cualquier campo: sobre una superficie de un solo lado, "entrar" y "salir" no son conceptos distinguibles, y el flujo carece de sentido.
Hay una segunda razón, más sutil, para exigir orientabilidad, y es la coherencia del borde. La banda de Möbius tiene por borde *una sola* curva cerrada (recórrase con el dedo: se dan dos vueltas antes de volver al punto de partida), de modo que ni siquiera la orientación inducida tendría cómo definirse de manera consistente. La hipótesis no es un tecnicismo que los autores añaden por prudencia: es la condición que hace que los dos miembros de la igualdad existan.

11.21.2. Campos solenoidales

Definición 11.105 (Campo solenoidal)**.** Un campo vectorial $\mathbf{B}$ de clase C^1 en un abierto $\Omega\subset\mathbb{R}^3$ es *solenoidal* (o incompresible, o de divergencia nula) si $\operatorname{div}\mathbf{B}=0$ en todo Ω. Un campo $\mathbf{A}$ con $\mathbf{B}=\operatorname{rot}\mathbf{A}$ se llama *potencial vectorial* de $\mathbf{B}$.

Proposición 11.106 (Caracterización)**.** *Si* $\mathbf{B}=\operatorname{rot}\mathbf{A}$ *para algún* $\mathbf{A}$ *de clase* C^2*, entonces* $\operatorname{div}\mathbf{B}=0$*, porque* $\operatorname{div}\circ\operatorname{rot}=0$ *idénticamente. El recíproco vale en dominios adecuados: si* Ω *es estrellado (o, más generalmente, si* Ω *no tiene cavidades, es decir si su segundo grupo de cohomología de De Rham es nulo), todo campo solenoidal en* Ω *es el rotacional de algún potencial vectorial. El potencial nunca es único:* $\mathbf{A}+\operatorname{grad}\varphi$ *sirve igual para cualquier* φ*, y esa libertad es lo que en electromagnetismo se llama* invariancia gauge.

Ejemplo 11.107 (Un solenoidal y su potencial)**.** El campo $\mathbf{B}=(x,\,y,\,-2z)$ tiene $\operatorname{div}\mathbf{B}=1+1-2=0$: es solenoidal en todo $\mathbb{R}^3$, que es estrellado, luego debe existir potencial vectorial. Buscándolo con la tercera componente nula se llega a $\mathbf{A}=(yz,\,-xz,\,0)$, y en efecto

$$\operatorname{rot}\mathbf{A}=\Big(\frac{\partial A_3}{\partial y}-\frac{\partial A_2}{\partial z},\,\frac{\partial A_1}{\partial z}-\frac{\partial A_3}{\partial x},\,\frac{\partial A_2}{\partial x}-\frac{\partial A_1}{\partial y}\Big)=(0+x,\;y-0,\;-z-z)=(x,\,y,\,-2z)=\mathbf{B},$$

comprobado símbolo a símbolo. Nótese que $\mathbf{A}+\operatorname{grad}\varphi$ funciona igual: por ejemplo $\mathbf{A}'=(yz+x,\,-xz+y,\,z)$, con $\varphi=\frac{x^2+y^2+z^2}{2}$, tiene el mismo rotacional.

Por qué importa el nombre: "solenoidal" viene de solenoide, y la razón es que un campo de divergencia nula organiza el espacio en *tubos de flujo*. Aplíquese el teorema de la divergencia

al trozo de tubo comprendido entre dos secciones S_1 y S_2: como no hay lateral que aporte (la superficie lateral está formada por líneas de campo, donde $\mathbf{B}\cdot\mathbf{N}=0$) y la integral de volumen es nula, el flujo entrante por S_1 iguala al saliente por S_2. El flujo a través de cualquier sección del tubo es el mismo; las líneas de campo no nacen ni mueren; si se estrechan, el campo se intensifica. Eso es literalmente la ley $\operatorname{div}\mathbf{B}=0$ de Maxwell (no hay monopolos magnéticos) y la ecuación de continuidad de un fluido incompresible, escritas en el mismo renglón.

11.21.3. Dominios simplemente conexos

Definición 11.108 (Dominio simplemente conexo)**.** Un abierto $\Omega\subset\mathbb{R}^n$ es *simplemente conexo* si es conexo por caminos y, además, todo lazo contenido en Ω se puede contraer a un punto sin salirse de Ω: formalmente, para toda curva cerrada continua $\gamma:[0,1]\to\Omega$ con $\gamma(0)=\gamma(1)=p$ existe una aplicación continua $H:[0,1]\times[0,1]\to\Omega$ con $H(s,0)=\gamma(s)$, $H(s,1)=p$ y $H(0,t)=H(1,t)=p$ para todo t. En el lenguaje del capítulo 37: Ω es conexo por caminos y $\pi_1(\Omega)$ es trivial.

Los ejemplos que hay que tener memorizados, porque son los que deciden en la práctica, son estos tres:

- $\mathbb{R}^3\setminus\{0\}$ **sí** es simplemente conexo. Cualquier lazo que evite el origen se puede contraer: primero se empuja radialmente hasta la esfera unidad, y en la esfera todo lazo se contrae deslizándolo hacia un polo.
- $\mathbb{R}^3$ menos una recta (por ejemplo el eje z) **no** lo es. Un lazo que rodea el eje no se puede contraer: la recta es un obstáculo infinito que no se puede esquivar ni por arriba ni por abajo.
- $\mathbb{R}^2\setminus\{0\}$ **no** lo es. La circunferencia unidad no se puede contraer dentro del plano perforado, y el número de vueltas alrededor del agujero es un invariante entero.

Idea clave

[Por qué el punto estorba en el plano y no en el espacio] Que un mismo obstáculo —un único punto— arruine la conexión simple del plano y no la del espacio suele sorprender, y la razón es puro recuento de dimensiones. Contraer un lazo significa tapar la curva con un disco, un objeto de dimensión 2, dentro del dominio. En $\mathbb{R}^2$, un disco cuyo borde rodea al origen *tiene que contener* al origen: no hay hueco por donde escaparse, porque el disco ya ocupa toda la dimensión disponible. En $\mathbb{R}^3$, en cambio, sobra una dimensión: el disco puede abombarse hacia arriba y esquivar el punto, igual que una película de jabón se levanta para evitar un obstáculo puntual. La regla general es que un obstáculo de dimensión k en $\mathbb{R}^n$ estorba a los lazos si $k+2>n$, es decir si su codimensión es ≤ 2. Un punto en el plano tiene codimensión 2: estorba. Un punto en el espacio tiene codimensión 3: no estorba. Una recta en el espacio tiene codimensión 2: estorba, y por eso se comporta como un punto del plano.
La moraleja práctica para este capítulo es que "quitar un punto" es inofensivo en $\mathbb{R}^3$ para los gradientes —un campo irrotacional en $\mathbb{R}^3\setminus\{0\}$ sí tiene potencial— pero fatal en $\mathbb{R}^2$, como demuestra el campo angular. Y que el obstáculo que sí importa en $\mathbb{R}^3$ para

los gradientes es el de dimensión uno: quitar un eje. Cuál es el defecto relevante depende de qué problema inverso se esté resolviendo, y esa distinción ya la vimos en la sección de potenciales vectoriales: $\mathbb{R}^3 \setminus \{0\}$ es simplemente conexo pero tiene una cavidad, y $\mathbb{R}^3$ menos un eje tiene el defecto contrario.

11.22. Ejercicios resueltos

Ejercicio 11.1. Calcula la masa de un alambre con forma de hélice $\mathbf{r}(t) = (\cos t, \sin t, t)$, $t \in [0, 2\pi]$, con densidad $\rho(x, y, z) = z$.

Solución. $\|\mathbf{r}'\| = \sqrt{\sin^2 t + \cos^2 t + 1} = \sqrt{2}$. Entonces

$$m = \int_{\mathcal{C}} z \, \mathrm{d}s = \int_0^{2\pi} t\sqrt{2} \, \mathrm{d}t = \sqrt{2} \cdot \frac{(2\pi)^2}{2} = 2\sqrt{2}\,\pi^2.$$

■

Ejercicio 11.2. Sea $\mathbf{F} = (2xy + z^2,\ x^2 + 2yz,\ y^2 + 2xz)$. Demuestra que es conservativo en $\mathbb{R}^3$, halla un potencial y calcula $\int_{\mathcal{C}} \mathbf{F} \cdot \mathrm{d}\mathbf{r}$ para cualquier curva de $(0,0,0)$ a $(1,1,1)$.

Solución. $\operatorname{rot}\mathbf{F}$: primera componente $\partial_y(y^2 + 2xz) - \partial_z(x^2 + 2yz) = 2y - 2y = 0$; las otras dos, análogas, también nulas. $\mathbb{R}^3$ es simplemente conexo: existe potencial. Integrando $f_x = 2xy + z^2$: $f = x^2y + xz^2 + g(y, z)$. De $f_y = x^2 + g_y = x^2 + 2yz$: $g = y^2z + h(z)$. De $f_z = 2xz + y^2 + h'(z) = y^2 + 2xz$: $h' = 0$. Potencial: $f = x^2y + xz^2 + y^2z$. La integral vale $f(1,1,1) - f(0,0,0) = 3$. ■

Ejercicio 11.3. Calcula $\oint_{\mathcal{C}} (x^4 - y^3)\,\mathrm{d}x + (x^3 + y^5)\,\mathrm{d}y$ donde $\mathcal{C}$ es la circunferencia unitaria antihoraria.

Solución. Por Green con $Q_x - P_y = 3x^2 + 3y^2$:

$$\oint = \iint_{x^2+y^2\le 1} 3(x^2 + y^2)\,\mathrm{d}A = 3\int_0^{2\pi}\int_0^1 r^2 \cdot r \,\mathrm{d}r\,\mathrm{d}\theta = 3 \cdot 2\pi \cdot \frac{1}{4} = \frac{3\pi}{2}.$$

Parametrizar directamente habría sido un mar de senos y cosenos; Green lo vuelve rutina. ■

Ejercicio 11.4. Verifica el teorema de Stokes para $\mathbf{F} = (z, x, y)$ y S el hemisferio $x^2+y^2+z^2 = 1$, $z \ge 0$, con normal exterior.

Solución. Lado de la curva: ∂S es la circunferencia unitaria en $z = 0$, antihoraria vista desde arriba: $\mathbf{r}(t) = (\cos t, \sin t, 0)$.

$$\oint \mathbf{F} \cdot \mathrm{d}\mathbf{r} = \int_0^{2\pi} (0, \cos t, \sin t) \cdot (-\sin t, \cos t, 0)\,\mathrm{d}t = \int_0^{2\pi} \cos^2 t \,\mathrm{d}t = \pi.$$

Lado de la superficie: $\operatorname{rot}\mathbf{F} = (1, 1, 1)$. Como el flujo del rotacional solo depende del borde, podemos reemplazar el hemisferio por el disco unitario plano con normal $\mathbf{k}$:

$$\iint_S (1,1,1) \cdot \mathbf{k}\,\mathrm{d}S = \iint_{\text{disco}} 1\,\mathrm{d}A = \pi. \ \checkmark$$

(Este reemplazo de superficie es exactamente la libertad que Stokes otorga.) ■

Ejercicio 11.5. Calcula el flujo de $\mathbf{F} = (x^3, y^3, z^3)$ a través de la esfera $x^2 + y^2 + z^2 = R^2$ con normal exterior.

Solución. Por Gauss, $\operatorname{div}\mathbf{F} = 3(x^2+y^2+z^2)$:

$$\iint_{\partial E} \mathbf{F}\cdot \mathrm{d}\mathbf{S} = 3\iiint_E (x^2+y^2+z^2)\,\mathrm{d}V = 3\int_0^{2\pi}\int_0^{\pi}\int_0^{R} \rho^2\cdot\rho^2\sin\phi\,\mathrm{d}\rho\,\mathrm{d}\phi\,\mathrm{d}\theta$$

$$= 3\cdot 2\pi\cdot 2\cdot\frac{R^5}{5} = \frac{12\pi R^5}{5}.$$

■

Ejercicio 11.6. Sea $\mathbf{F} = \dfrac{(x,y,z)}{(x^2+y^2+z^2)^{3/2}}$ (el campo de una carga puntual). Demuestra que $\operatorname{div}\mathbf{F} = 0$ fuera del origen, pero que su flujo por cualquier esfera centrada en el origen es 4π. ¿Contradice esto a Gauss?

Solución. Con $r = \|\mathbf{x}\|$, $\mathbf{F} = r^{-3}\mathbf{x}$. Entonces $\operatorname{div}\mathbf{F} = r^{-3}\operatorname{div}\mathbf{x} + \nabla(r^{-3})\cdot\mathbf{x} = 3r^{-3} - 3r^{-5}\,\mathbf{x}\cdot\mathbf{x} = 3r^{-3} - 3r^{-3} = 0$ para $r \neq 0$ (usando $\nabla(r^{-3}) = -3r^{-5}\mathbf{x}$). Sobre la esfera de radio R: $\mathbf{n} = \mathbf{x}/R$ y $\mathbf{F}\cdot\mathbf{n} = R^{-3}\cdot R^2/R\cdot R = R^{-2}$... calculemos con cuidado: $\mathbf{F}\cdot\mathbf{n} = \frac{\mathbf{x}}{R^3}\cdot\frac{\mathbf{x}}{R} = \frac{R^2}{R^4} = \frac{1}{R^2}$, luego el flujo es $\frac{1}{R^2}\cdot 4\pi R^2 = 4\pi$, independiente de R. No hay contradicción: Gauss exige que $\mathbf{F}$ sea C^1 en *todo* el sólido, y el origen (donde vive la carga) está excluido. La enseñanza física es la ley de Gauss del electromagnetismo: el flujo mide la carga encerrada, y la divergencia de este campo es una "delta de Dirac" concentrada en el origen. ■

Ejercicio 11.7. Usando que el flujo del rotacional a través de una superficie *cerrada* es cero, demuestra que $\iint_{\partial E}\operatorname{rot}\mathbf{F}\cdot\mathrm{d}\mathbf{S} = 0$ de dos maneras (Gauss y Stokes).

Solución. Con Gauss: $\iint_{\partial E}\operatorname{rot}\mathbf{F}\cdot\mathrm{d}\mathbf{S} = \iiint_E \operatorname{div}(\operatorname{rot}\mathbf{F})\,\mathrm{d}V = 0$ por la identidad $\operatorname{div}\operatorname{rot} = 0$. *Con Stokes*: cortamos la superficie cerrada en dos mitades S_1, S_2 con borde común $\mathcal{C}$. Con las orientaciones inducidas, $\partial S_1 = \mathcal{C}$ y $\partial S_2 = -\mathcal{C}$; entonces $\iint_{S_1}\operatorname{rot}\mathbf{F}\cdot\mathrm{d}\mathbf{S} + \iint_{S_2}\operatorname{rot}\mathbf{F}\cdot\mathrm{d}\mathbf{S} = \oint_{\mathcal{C}}\mathbf{F}\cdot\mathrm{d}\mathbf{r} - \oint_{\mathcal{C}}\mathbf{F}\cdot\mathrm{d}\mathbf{r} = 0$. Ambos caminos expresan lo mismo: "la frontera de una frontera es vacía" ($\partial\partial = \emptyset$), la contraparte geométrica de $\mathrm{d}\circ\mathrm{d} = 0$. ■

Ejercicio 11.8. Calcula el área de la parte del paraboloide $z = x^2 + y^2$ con $z \le 1$, y el flujo de $\mathbf{F} = (0,0,z)$ a través de ella (normal hacia arriba).

Solución. Como grafo con $g = x^2+y^2$: $\mathrm{d}S = \sqrt{1+4x^2+4y^2}\,\mathrm{d}A$ sobre el disco unitario. En polares:

$$A = 2\pi\int_0^1 \sqrt{1+4r^2}\,r\,\mathrm{d}r = 2\pi\cdot\frac{1}{12}\big[(1+4r^2)^{3/2}\big]_0^1 = \frac{\pi}{6}\big(5\sqrt{5}-1\big).$$

Para el flujo, con normal arriba: $\iint(-Pg_x - Qg_y + R)\mathrm{d}A = \iint_{x^2+y^2\le 1} z\,\mathrm{d}A$ con $z = r^2$:

$$\int_0^{2\pi}\int_0^1 r^2\cdot r\,\mathrm{d}r\,\mathrm{d}\theta = 2\pi\cdot\frac{1}{4} = \frac{\pi}{2}.$$

■

11.23. Ejercicios adicionales resueltos

Ejercicio 11.9. Armónicas radiales en $\mathbb{R}^3$ y $\mathbb{R}^2$.

Solución. $(\rho^2 u')' = 0 \Rightarrow u = \frac{A}{\rho} + B$ (Coulomb); en el plano, $(ru')' = 0 \Rightarrow u = A \ln r + B$: la dimensión cambia la ley física. ■

Ejercicio 11.10. Flujo de (x^2, y^2, z^2) por el cubo unitario.

Solución. Gauss: $\iiint 2(x+y+z) = 3$; directo: cada par de caras aporta 1. ✓ ■

Ejercicio 11.11. Helmholtz para $\mathbf{F} = (x+y,\, y-x,\, 0)$.

Solución. $(x, y, 0) = \nabla \frac{x^2+y^2}{2}$ (div 2, rot 0) y $(y, -x, 0) = \operatorname{rot}(0, 0, -\frac{x^2+y^2}{2})$ (div 0, rot $(0, 0, -2)$): expansión pura + rotación pura. ■

Ejercicio 11.12. Compatibilidad y unicidad del problema de Neumann.

Solución. Necesario: $\iint g \, \mathrm{d}S = \iiint \Delta u = 0$. Unicidad salvo constante: Green con $w = u_1 - u_2$ da $\iiint \|\nabla w\|^2 = 0$. ■

Ejercicio 11.13. Circulación de (z^2, x^2, y^2) sobre el triángulo $(1,0,0), (0,1,0), (0,0,1)$.

Solución. $\operatorname{rot} = (2y, 2z, 2x)$, $\cdot \mathbf{n} = \frac{2}{\sqrt{3}}$ constante en el plano $x+y+z=1$; área $\frac{\sqrt{3}}{2}$: circulación $= 1$. ■

Ejercicio 11.14. Arquímedes con Gauss.

Solución. $B_z = -\iint p\, \mathbf{e}_3 \cdot \mathbf{n} \, \mathrm{d}S = -\iiint p_z = \rho_0 g V$; horizontales nulas: empuje = peso del fluido desplazado. ■

11.24. Taller de ejemplos explicados: integrales de línea y Green

11.24.1. Una integral de línea

Una integral de línea $\displaystyle\int_C \mathbf{F} \cdot \mathrm{d}\mathbf{r}$ suma el "empuje" de un campo $\mathbf{F}$ a lo largo de un camino C; físicamente, es el *trabajo* que hace una fuerza al mover algo por ese camino.

Ejemplo 11.109. Calcular $\displaystyle\int_C \mathbf{F} \cdot \mathrm{d}\mathbf{r}$ con $\mathbf{F} = (y, x)$ a lo largo del segmento recto de $(0,0)$ a $(1,1)$.

Paso 1. Parametrizar el camino. La recta se describe con $x = t$, $y = t$, mientras t va de 0 a 1. Entonces $\mathrm{d}\mathbf{r} = (\mathrm{d}x, \mathrm{d}y) = (1,1)\,\mathrm{d}t$.

Paso 2. Escribir F sobre el camino. Como $x = t$, $y = t$: $\mathbf{F} = (y, x) = (t, t)$.

Paso 3. Hacer el producto punto: $\mathbf{F} \cdot \mathrm{d}\mathbf{r} = (t,t) \cdot (1,1)\,\mathrm{d}t = (t+t)\,\mathrm{d}t = 2t\,\mathrm{d}t$.

Paso 4. Integrar de 0 a 1:

$$\int_0^1 2t \, \mathrm{d}t = \left[t^2\right]_0^1 = 1.$$

El trabajo vale 1.

11.24.2. El teorema de Green (convertir un contorno en un área)

Regla: teorema de Green

Para una curva cerrada C que rodea una región D:

$$\oint_C P\,\mathrm{d}x + Q\,\mathrm{d}y = \iint_D \left(\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y}\right)\mathrm{d}A.$$

Convierte una integral *sobre el borde* en una *sobre el área*, que a veces es mucho más fácil.

Ejemplo 11.110. Calcular $\oint_C (-y\,\mathrm{d}x + x\,\mathrm{d}y)$ donde C es el borde del cuadrado unitario $[0,1]\times[0,1]$.

Paso 1. Identificar P y Q: aquí $P = -y$ y $Q = x$.

Paso 2. Calcular las parciales que pide Green: $\dfrac{\partial Q}{\partial x} = 1$ y $\dfrac{\partial P}{\partial y} = -1$. Su resta es $1-(-1) = 2$.

Paso 3. Integrar esa constante sobre el cuadrado. El área del cuadrado unitario es 1, así que

$$\iint_D 2\,\mathrm{d}A = 2\cdot(\text{área}) = 2\cdot 1 = 2.$$

Rodear el borde entero se redujo a multiplicar 2 por el área. (De hecho, $\frac{1}{2}\oint(-y\,\mathrm{d}x + x\,\mathrm{d}y)$ siempre da el área encerrada: una fórmula que usan los planímetros.)

11.25. Ampliación del temario: introducción al cálculo tensorial

Idea clave

Intuición primero. Un tensor es una *máquina* que come vectores y devuelve un número, y que es lineal en cada ranura donde metes un vector. Su *rango* es cuántas ranuras tiene: cero ranuras es un escalar (ya es un número), una ranura es un vector o covector, dos ranuras es una matriz (como el tensor de esfuerzos). La gracia es que es el *mismo* objeto físico visto en cualquier sistema de coordenadas, gracias a una ley de transformación fija: por eso las leyes de la física se escriben con tensores.

Este capítulo trabajó con campos escalares (temperatura) y vectoriales (velocidad). Hay un tercer tipo, más general, que los engloba: los *campos tensoriales*. Un tensor en un punto es una máquina multilineal que come vectores y devuelve un número; su *rango* es cuántos vectores come.

Regla: la jerarquía de tensores

Rango 0	escalar (un número): temperatura, presión
Rango 1	vector (o covector): velocidad, gradiente
Rango 2	matriz que transforma vectores: tensor de esfuerzos, métrica

Ejemplo 11.111 (El tensor de esfuerzos)**.** En un material sólido bajo carga, el *tensor de esfuerzos* σ_{ij} es un tensor de rango 2: dada una superficie con normal $\mathbf{n}$, el vector $\sigma\,\mathbf{n}$ da la fuerza por unidad de área sobre esa superficie. Necesita dos índices porque relaciona *dos* direcciones (la de la superficie y la de la fuerza). Es el objeto central de la ingeniería estructural: predice cuándo una viga se rompe.

Observación 11.112 (Operadores tensoriales elementales). Los operadores del capítulo son casos de operaciones tensoriales. El *gradiente* de un escalar produce un covector (rango 1). La *divergencia* contrae un vector a un escalar (baja el rango). El *producto tensorial* de dos vectores $\mathbf{a}\otimes\mathbf{b}$ da un tensor de rango 2 (la matriz a_ib_j). Lo esencial de un tensor es su *ley de transformación*: al cambiar de coordenadas, sus componentes se transforman de manera precisa (un índice "arriba" como un vector, uno "abajo" como un covector), y esa ley es lo que garantiza que las leyes de la física escritas con tensores valgan en cualquier sistema de coordenadas. El desarrollo completo —derivada covariante, curvatura, la métrica que sube y baja índices— es el capítulo 36.

11.26. Ejercicios propuestos

Sin solución desarrollada: solo enunciado y respuesta al final. Úsalos para comprobar que lo del capítulo se sostiene sin la muleta del texto delante.

11.1. Calcula $\operatorname{div}\mathbf{F}$ para $\mathbf{F}=(x^2,y^2,z^2)$.

11.2. Calcula $\operatorname{rot}\mathbf{F}$ para $\mathbf{F}=(-y,x,0)$.

11.3. ¿Es $\mathbf{F}=(2xy,\,x^2)$ conservativo? Si lo es, da un potencial.

11.4. Con ese mismo $\mathbf{F}$, calcula $\int_C \mathbf{F}\cdot\mathrm{d}\mathbf{r}$ desde $(0,0)$ hasta $(1,2)$ por cualquier camino.

11.5. Usa Green para calcular el área del disco unidad mediante $\frac{1}{2}\oint(x\,\mathrm{d}y-y\,\mathrm{d}x)$.

11.6. Calcula el flujo de $\mathbf{F}=(x,y,z)$ a través de la esfera unidad.

11.7. Demuestra que $\operatorname{rot}(\operatorname{grad} f)=\mathbf{0}$ para f de clase C^2.

11.8. Demuestra que $\operatorname{div}(\operatorname{rot}\mathbf{F})=0$.

11.9. Halla la longitud de la hélice $(\cos t,\sin t,t)$ para $t\in[0,2\pi]$.

11.10. Calcula $\displaystyle\oint_{|z|=1}\frac{-y\,\mathrm{d}x+x\,\mathrm{d}y}{x^2+y^2}$ y explica por qué no es cero pese a que el campo tiene rotacional nulo.

11.11. Verifica Stokes para $\mathbf{F}=(-y,x,0)$ sobre el disco unidad en el plano $z=0$.

11.12. Calcula el laplaciano de $f(x,y,z)=x^2+y^2+z^2$.

Respuestas

1. $2(x+y+z)$. **2.** $(0,0,2)$. **3.** Sí: $\varphi = x^2 y$. **4.** 2. **5.** π. **6.** $\operatorname{div}\mathbf{F} = 3$ y el volumen es $\frac{4\pi}{3}$, luego el flujo es 4π. **7.** Sale de la igualdad de las derivadas cruzadas (Schwarz). **8.** Igual: las seis derivadas cruzadas se cancelan por pares. **9.** $2\sqrt{2}\,\pi$. **10.** 2π; el campo no está definido en el origen, y el dominio $\mathbb{R}^2 \setminus \{0\}$ no es simplemente conexo —esto es exactamente $H^1 \neq 0$. **11.** Ambos lados dan 2π. **12.** 6.

Capítulo 12

Ecuaciones diferenciales

Lo que necesitas traer

Del capítulo 6: derivadas, integrales, series de potencias y la exponencial. Del capítulo 7: valores y vectores propios, diagonalización y la exponencial de una matriz, que es la que resuelve los sistemas lineales. Del capítulo 4: las raíces complejas del polinomio característico.

Una *ecuación diferencial ordinaria* (EDO) relaciona una función $y(t)$ con sus derivadas; su *orden* es el de la derivada más alta. Son el idioma de las leyes de la naturaleza: $F = ma$, los circuitos, las poblaciones, la desintegración radiactiva. Este capítulo desarrolla las técnicas clásicas de primer y segundo orden, la teoría de sistemas lineales (donde el álgebra lineal cobra vida), la transformada de Laplace, los métodos de series, el análisis cualitativo y una introducción sustancial a las ecuaciones en derivadas parciales.

12.1. Ecuaciones de primer orden

Idea clave

Intuición primero. Una ecuación diferencial no te da la función: te da una *ley de cambio* y te pide reconstruir la función a partir de ella. Eso es exactamente cómo funciona la física: nadie observa la ley "la posición del planeta es tal fórmula"; lo que se observa es la ley que relaciona la aceleración con la fuerza, y la trayectoria hay que deducirla. Por eso la solución nunca es única sin información adicional: hay toda una familia de funciones con la misma ley de cambio, y la condición inicial es la que selecciona una. Geométricamente, la ecuación $y' = f(x, y)$ dibuja una pequeña flecha de pendiente $f(x, y)$ en cada punto del plano —el campo de direcciones— y resolverla es encontrar las curvas que en cada punto van en la dirección que marca la flecha. La condición inicial dice por dónde entras a ese campo; a partir de ahí, la curva ya no tiene elección. Todos los métodos de esta sección (separables, lineales, exactas) son recetas para recorrer esa curva con álgebra en vez de con dibujos.

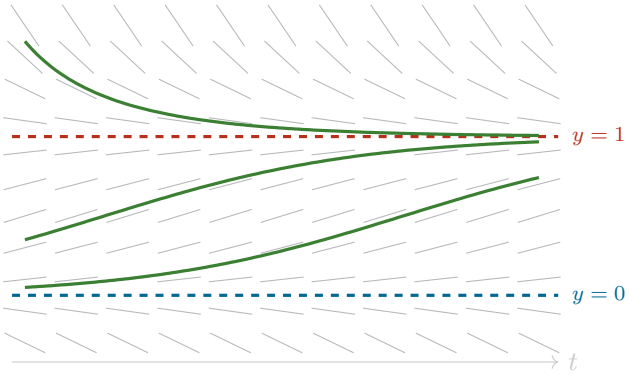


Figura 12.1: Campo de direcciones de la ecuación logística $y' = y(1-y)$ con tres soluciones. Se lee de un vistazo lo que costaría deducir de la fórmula: $y = 0$ es un equilibrio inestable y $y = 1$ es asintóticamente estable, tanto por arriba como por abajo.

12.1.1. Separables

Si $y' = f(t)\,g(y)$, se separa e integra: $\int \frac{\mathrm{d}y}{g(y)} = \int f(t)\,\mathrm{d}t + C$. (Cuidado: las raíces de $g(y) = 0$ dan *soluciones de equilibrio* constantes que la separación puede perder.)

Ejemplo 12.1 (Modelos de población)**.** *Malthus*: $y' = ky$ da $y = y_0 e^{kt}$; con $k < 0$ es la desintegración radiactiva, con vida media $\tau = \ln 2/|k|$ (así funciona la datación por carbono-14). *Logístico*: $y' = ry(1 - y/K)$; separando y con fracciones parciales,

$$y(t) = \frac{K}{1 + (\frac{K}{y_0} - 1)e^{-rt}}\,:$$

crecimiento casi exponencial al principio, saturación en la *capacidad de carga* K. Los equilibrios son $y = 0$ (inestable) y $y = K$ (estable).

12.1.2. Lineales: el factor integrante

Para $y' + p(t)y = q(t)$, multiplicamos por $\mu(t) = e^{\int p\,\mathrm{d}t}$; entonces el lado izquierdo se convierte exactamente en $(\mu y)'$, y

$$y = \frac{1}{\mu}\Big(\int \mu q\,\mathrm{d}t + C\Big).$$

La solución es (como siempre en lo lineal) *homogénea + particular*: $Ce^{-\int p}$ más una solución cualquiera.

Ejemplo 12.2 (Mezclas)**.** Un tanque con 100 L de agua pura recibe salmuera (0.5 kg/L) a 3 L/min, bien mezclada, y desagua a 3 L/min. Si $y(t)$ es la sal en el tanque:

$$y' = \underbrace{1.5}_{\text{entra}} - \underbrace{\frac{3y}{100}}_{\text{sale}}, \qquad y(0) = 0.$$

Lineal con $\mu = e^{0.03t}$: $y = 50 + Ce^{-0.03t}$, y la condición inicial da $y(t) = 50\,(1 - e^{-0.03t})$: la sal tiende al equilibrio de 50 kg exponencialmente.

12.1.3. Exactas y factores integrantes

$M(x,y)\,\mathrm{d}x + N(x,y)\,\mathrm{d}y = 0$ es *exacta* si $M_y = N_x$ (en un dominio simplemente conexo): entonces existe F con $F_x = M$, $F_y = N$, y las soluciones son las curvas de nivel $F(x,y) = C$. Es el criterio de campo conservativo del capítulo 11 con otro sombrero. Si no es exacta, a veces un factor μ la vuelve exacta: si $\frac{M_y - N_x}{N}$ depende solo de x, sirve $\mu(x) = e^{\int \frac{M_y - N_x}{N}\mathrm{d}x}$; simétricamente con y.

12.1.4. Sustituciones clásicas

- *Homogénea* ($y' = F(y/x)$): con $v = y/x$ se separa: $xv' = F(v) - v$.
- *Bernoulli* ($y' + p\,y = q\,y^n$): con $u = y^{1-n}$ se vuelve lineal: $u' + (1-n)pu = (1-n)q$.
- *Riccati* ($y' = a + by + cy^2$): conocida una solución particular y_1, el cambio $y = y_1 + \frac{1}{u}$ la vuelve lineal.

Regla: clasificar una EDO de primer orden y elegir el método

Ante $y' = f(x,y)$ lo primero no es calcular: es *reconocer*. Se recorre esta lista en orden y se para en la primera casilla que encaje.

1. **¿Se separan las variables?** Si la ecuación se puede escribir como $y' = g(x)h(y)$, se pasa $h(y)$ al otro lado y se integra: $\int \frac{\mathrm{d}y}{h(y)} = \int g(x)\,\mathrm{d}x$. Es el método más barato; siempre se prueba primero.
2. **¿Es lineal en y?** Es decir, ¿tiene la forma $y' + p(x)y = q(x)$, con y e y' apareciendo solos, en primera potencia y sin multiplicarse entre sí? Entonces factor integrante $\mu = e^{\int p}$ y la fórmula $y = \frac{1}{\mu}(\int \mu q + C)$.
3. **¿Es exacta?** Se escribe como $M\,\mathrm{d}x + N\,\mathrm{d}y = 0$ y se comprueba $M_y = N_x$. Si sale, se busca F con $F_x = M$ y $F_y = N$; la solución es $F(x,y) = C$. Si no sale, todavía queda el factor integrante: si $\frac{M_y - N_x}{N}$ solo depende de x, hay $\mu(x)$; si $\frac{N_x - M_y}{M}$ solo depende de y, hay $\mu(y)$.
4. **¿Es homogénea de grado cero?** Prueba rápida: sustituir $x \mapsto \lambda x$, $y \mapsto \lambda y$ y ver si f no cambia. Si no cambia, el cambio $v = y/x$ la vuelve separable: $y = vx$, $y' = v + xv'$.
5. **¿Es de Bernoulli?** Forma $y' + p\,y = q\,y^n$ con $n \neq 0, 1$. El cambio $u = y^{1-n}$ la linealiza. Si $n = 0$ ya era lineal y si $n = 1$ ya era separable: por eso se excluyen.
6. **¿Es de Riccati?** Forma $y' = a + by + cy^2$. Aquí hace falta *regalada* una solución particular y_1; con ella, $y = y_1 + 1/u$ la vuelve lineal. Sin esa solución regalada no hay método elemental.
7. **Control final.** Sea cual sea el camino, se deriva la respuesta y se mete en la ecuación original. Y si hay condición inicial, se sustituye *antes* de simplificar, para no perder el signo de una raíz.

Un aviso de contabilidad: los pasos 1 y 3 pueden *perder* soluciones. Al dividir entre $h(y)$ se está suponiendo $h(y) \neq 0$, de modo que cada raíz constante $y \equiv c$ con $h(c) = 0$ es una solución legítima que hay que añadir a mano.

Ejemplo 12.3 (Cinco ecuaciones, cinco métodos distintos)**. (a) Separable.** $y' = xy^2$ con $y(0) = 1$. Se separa: $\int y^{-2}\mathrm{d}y = \int x\,\mathrm{d}x$, es decir $-\frac{1}{y} = \frac{x^2}{2} + C$. La condición inicial da $-1 = C$, luego

$$y = \frac{1}{1 - \frac{x^2}{2}} = \frac{2}{2 - x^2}.$$

Control: $y' = \frac{4x}{(2-x^2)^2}$ y $xy^2 = x\frac{4}{(2-x^2)^2}$: coinciden. Obsérvese que la solución explota en $x = \sqrt{2}$ aunque la ecuación no tenga nada raro ahí: existencia *local*, como avisa Picard–Lindelöf. Y no olvidemos $y \equiv 0$, que se perdió al dividir entre y^2.

(b) Lineal. $y' + \frac{2}{x}y = x^3$. Aquí $p = 2/x$, así que $\mu = e^{\int 2/x\,\mathrm{d}x} = e^{2\ln x} = x^2$. Multiplicando: $(x^2 y)' = x^5$, luego $x^2 y = \frac{x^6}{6} + C$ y

$$y = \frac{x^4}{6} + \frac{C}{x^2}.$$

La estructura es la de siempre: una particular $x^4/6$ más la homogénea C/x^2.

(c) Exacta. $(2xy + 3x^2)\,\mathrm{d}x + (x^2 - 1)\,\mathrm{d}y = 0$. Test: $M_y = 2x$ y $N_x = 2x$, coinciden, es exacta. Buscamos F: de $F_x = 2xy + 3x^2$ sale $F = x^2 y + x^3 + g(y)$; derivando, $F_y = x^2 + g'(y)$ debe valer $x^2 - 1$, luego $g(y) = -y$. La solución implícita es

$$x^2 y + x^3 - y = C \quad \Longleftrightarrow \quad y = \frac{C - x^3}{x^2 - 1}.$$

(d) Homogénea. $y' = \frac{x^2+y^2}{xy}$. Al cambiar $x \mapsto \lambda x$, $y \mapsto \lambda y$ el λ^2 se cancela arriba y abajo: es homogénea. Con $v = y/x$:

$$v + xv' = \frac{1 + v^2}{v} \implies xv' = \frac{1 + v^2}{v} - v = \frac{1}{v} \implies v\,\mathrm{d}v = \frac{\mathrm{d}x}{x}.$$

Integrando, $\frac{v^2}{2} = \ln x + C$, es decir $y^2 = x^2(2\ln x + C_1)$.

(e) Bernoulli. $y' + y = xy^2$, con $n = 2$. El cambio es $u = y^{1-2} = 1/y$, y entonces $u' = -y'/y^2$. Dividiendo la ecuación entre y^2 y sustituyendo:

$$u' - u = -x.$$

Ya es lineal, con $\mu = e^{-x}$: $(ue^{-x})' = -xe^{-x}$, y como $\int -xe^{-x}\mathrm{d}x = (x+1)e^{-x}$, sale $u = x + 1 + Ce^x$. Deshaciendo el cambio,

$$y = \frac{1}{x + 1 + Ce^x}.$$

Otra vez, $y \equiv 0$ es solución y se había perdido al dividir entre y^2.

Cuidado

Una misma ecuación puede caer en varias casillas, y eso no es un problema sino una comodidad: se elige la más barata. Por ejemplo $y' = \frac{y}{x}$ es a la vez separable, lineal y homogénea, y las tres rutas dan $y = Cx$. El error de verdad es el contrario: forzar una casilla que no toca. Antes de escribir $\mu = e^{\int p}$ hay que estar seguro de que la ecuación es lineal *en* y; en $y' = x + y^2$ el término y^2 lo arruina todo, y ahí no hay factor integrante que valga (es una Riccati sin solución particular a la vista, y de hecho su solución general se escribe con funciones de Airy).

12.1.5. Existencia, unicidad y el método de Picard

Teorema 12.4 (Picard–Lindelöf)**.** *Si f y $\frac{\partial f}{\partial y}$ son continuas en un rectángulo alrededor de (t_0, y_0), el problema $y' = f(t, y)$, $y(t_0) = y_0$ tiene solución única en un intervalo alrededor de t_0.*

Idea: iteración de Picard. El problema equivale a la ecuación integral $y(t) = y_0 + \int_{t_0}^{t} f(s, y(s))\,\mathrm{d}s$. Se define la sucesión $y_{n+1}(t) = y_0 + \int_{t_0}^{t} f(s, y_n(s))\,\mathrm{d}s$ partiendo de $y_0(t) \equiv y_0$; la condición sobre $\frac{\partial f}{\partial y}$ (Lipschitz) hace que la iteración sea *contractiva* y converja a un único punto fijo. Para $y' = y$, $y(0) = 1$: $y_1 = 1 + t$, $y_2 = 1 + t + \frac{t^2}{2}$, … ¡la iteración construye la serie de e^t! □

Idea clave

Qué dice esto en cristiano. Picard responde a la pregunta previa a cualquier cálculo: ¿esta ecuación diferencial tiene solución, y es la única? La respuesta es sí a las dos si f y su derivada respecto de y son continuas cerca del punto inicial. Y la palabra clave del enunciado es *alrededor*: la garantía es local, en un intervalo posiblemente diminuto. Que la derivada parcial exista y sea continua es lo que impide que dos soluciones distintas salgan del mismo punto, porque acota cuánto pueden separarse dos trayectorias vecinas.

Ejemplo 12.5 (Cuándo hay unicidad y cuándo no)**. (a) El caso bueno, con las iteradas.** Para $y' = y$ con $y(0) = 1$, las iteradas de Picard $y_{n+1} = 1 + \int_0^t y_n$ dan sucesivamente $1 + t$, luego $1 + t + \frac{t^2}{2}$, luego $1 + t + \frac{t^2}{2} + \frac{t^3}{6}$: son las sumas parciales de e^t. La iteración *construye* la solución, no solo demuestra que existe.

(b) Sin unicidad. Para $y' = y^{2/3}$ con $y(0) = 0$, la derivada parcial $\frac{2}{3}y^{-1/3}$ no es continua en $y = 0$. Y en efecto hay al menos dos soluciones: $y \equiv 0$ y $y = \left(\frac{t}{3}\right)^3$, que también cumple la ecuación y la condición inicial. De hecho hay infinitas, pegando trozos de una y de otra.

(c) La solución existe solo un rato. Para $y' = 1 + y^2$ con $y(0) = 0$, tanto f como $\frac{\partial f}{\partial y} = 2y$ son continuas en todo el plano, y sin embargo la solución es $y = \tan t$, que explota en $t = \frac{\pi}{2} = 1{,}5707963$. La hipótesis es global y la conclusión sigue siendo local: no hay contradicción.

(d) Cuándo sí es global. Si $|f(t, y)| \le a(t)|y| + b(t)$ con a y b continuas (crecimiento a lo sumo lineal en y), la solución existe en todo el intervalo. La desigualdad de Gronwall da

la cota: de $u(t) \le C + K\int u$ se deduce $u(t) \le C\,\mathrm{e}^{K(t-t_0)}$, que es finita en cualquier intervalo acotado. Es lo que impide la explosión de (c), donde el crecimiento era cuadrático.

Cuidado

[Existencia y unicidad son dos cosas distintas] La continuidad de f sola ya garantiza *existencia* (teorema de Peano); lo que exige la derivada parcial continua es la *unicidad*. El ejemplo (b) tiene existencia de sobra y ninguna unicidad. Segundo aviso: en la práctica basta con que f sea Lipschitz en y, condición algo más débil que tener derivada continua, y es la que se usa en la demostración vía punto fijo de Banach. Y tercero: sin unicidad, los métodos numéricos no fallan con un error, simplemente convergen a *una* de las soluciones sin avisar de que hay otras.

Cuidado

Sin la condición de Lipschitz, la unicidad puede fallar: $y' = 3y^{2/3}$, $y(0) = 0$ admite $y \equiv 0$ y $y = t^3$ (y de hecho infinitas soluciones). Y la existencia es solo *local*: $y' = y^2$, $y(0) = 1$ da $y = \frac{1}{1-t}$, que explota en $t = 1$ aunque la ecuación sea perfectamente regular.

Regla: reconocer una ecuación exacta y fabricar un factor integrante

Una ecuación de primer orden escrita en forma diferencial

$$M(x,y)\,dx + N(x,y)\,dy = 0$$

es *exacta* si el lado izquierdo es la diferencial total de una función $F(x,y)$; entonces la solución general es simplemente $F(x,y) = C$. Cuando no lo es, a veces se la puede volver exacta multiplicándola por una función bien elegida. El procedimiento completo es este.

1. **Escribe la ecuación en forma diferencial y separa M y N.** Si te dan $y' = f(x,y)$, pasa a $f(x,y)\,dx - dy = 0$ o multiplica por lo que haga falta para tener una forma cómoda. Identifica sin ambigüedad quién acompaña a dx y quién a dy.
2. **Aplica el test de exactitud: $\frac{\partial M}{\partial y} = \frac{\partial N}{\partial x}$.** Es la igualdad de las derivadas cruzadas de F, y en un dominio simplemente conexo es necesaria y suficiente. Dos derivadas parciales y ya sabes si tienes trabajo fácil o difícil.
3. **Si es exacta, integra M respecto de x.** Obtienes $F(x,y) = \int M\,dx + g(y)$, donde g es una función desconocida *de y sola*: la "constante" de integración cuando se integra en x.
4. **Determina $g(y)$ derivando respecto de y.** Impón $\frac{\partial F}{\partial y} = N$. Todos los términos con x deben cancelarse; si no lo hacen, hay un error de cuentas o la ecuación no era exacta. Integra lo que quede y escribe la solución $F(x,y) = C$.
5. **Si no es exacta, busca un factor integrante de una variable.** Calcula

$$\frac{1}{N}\left(\frac{\partial M}{\partial y} - \frac{\partial N}{\partial x}\right).$$

Si el resultado depende *solo de* x, llámalo $h(x)$ y toma $\mu = e^{\int h\,dx}$. Si no, prueba con

$$\frac{1}{M}\left(\frac{\partial N}{\partial x} - \frac{\partial M}{\partial y}\right);$$

si depende solo de y, llámalo $k(y)$ y toma $\mu = e^{\int k\,dy}$.

6. **Multiplica todo por μ y vuelve al paso 2.** La ecuación $\mu M\,dx + \mu N\,dy = 0$ debe pasar ahora el test. Compruébalo antes de seguir: cuesta dos derivadas y evita arrastrar un error.

7. **Verifica la solución y vigila las soluciones perdidas.** Deriva implícitamente $F(x,y) = C$ y comprueba que reproduce la ecuación. Y recuerda que multiplicar por μ puede introducir o eliminar soluciones donde $\mu = 0$ o μ no está definido; revisa esos valores por separado.

Ejemplo 12.6 (Cinco ecuaciones: dos exactas, tres rescatadas)**. (a) Exacta de entrada.** Sea

$$(2xy + 3)\,dx + (x^2 - 1)\,dy = 0.$$

Test: $\frac{\partial M}{\partial y} = 2x$ y $\frac{\partial N}{\partial x} = 2x$. Coinciden, es exacta. Integramos M en x:

$$F = \int (2xy + 3)\,dx = x^2y + 3x + g(y).$$

Derivamos en y: $\frac{\partial F}{\partial y} = x^2 + g'(y)$, y esto debe ser $N = x^2 - 1$. Luego $g'(y) = -1$, $g(y) = -y$, y la solución general es

$$\boxed{x^2y + 3x - y = C} \quad\Longrightarrow\quad y = \frac{C - 3x}{x^2 - 1}.$$

Verificación: $\frac{\partial F}{\partial x} = 2xy + 3$ y $\frac{\partial F}{\partial y} = x^2 - 1$, exactamente M y N.

(b) Factor integrante que solo depende de x. Sea

$$(3x^2y + 2xy + y^3)\,dx + (x^2 + y^2)\,dy = 0.$$

Test: $\frac{\partial M}{\partial y} = 3x^2 + 2x + 3y^2$ y $\frac{\partial N}{\partial x} = 2x$; no coinciden. Paso 5:

$$\frac{1}{N}\left(\frac{\partial M}{\partial y} - \frac{\partial N}{\partial x}\right) = \frac{3x^2 + 3y^2}{x^2 + y^2} = 3,$$

constante, y por tanto función de x sola. Así que $\mu = e^{\int 3\,dx} = e^{3x}$. La nueva ecuación es

$$e^{3x}(3x^2y + 2xy + y^3)\,dx + e^{3x}(x^2 + y^2)\,dy = 0,$$

y ahora sí: $\frac{\partial}{\partial y}\left[e^{3x}(3x^2y+2xy+y^3)\right] = e^{3x}(3x^2+2x+3y^2)$ y $\frac{\partial}{\partial x}\left[e^{3x}(x^2+y^2)\right] = e^{3x}(3x^2+3y^2+2x)$, que son iguales. Integrando N en y (más cómodo aquí):

$$F = e^{3x}\left(x^2y + \frac{y^3}{3}\right) + h(x),$$

y al derivar en x sale exactamente μM con $h' = 0$. Solución:

$$e^{3x}\left(x^2y + \frac{y^3}{3}\right) = C.$$

(c) Factor integrante que solo depende de y. Sea

$$y\,dx + (2x - y\,e^y)\,dy = 0.$$

Test: $\frac{\partial M}{\partial y} = 1$, $\frac{\partial N}{\partial x} = 2$. No exacta. El primer cociente da $\frac{1-2}{2x-ye^y}$, que depende de las dos variables; probamos el segundo:

$$\frac{1}{M}\left(\frac{\partial N}{\partial x} - \frac{\partial M}{\partial y}\right) = \frac{2-1}{y} = \frac{1}{y},$$

función de y sola. Luego $\mu = e^{\int dy/y} = y$. La ecuación pasa a ser

$$y^2\,dx + (2xy - y^2e^y)\,dy = 0,$$

con $\frac{\partial}{\partial y}(y^2) = 2y = \frac{\partial}{\partial x}(2xy - y^2e^y)$. Integrando M en x: $F = xy^2 + g(y)$. Derivando: $2xy + g'(y) = 2xy - y^2e^y$, de donde $g'(y) = -y^2e^y$ y, por partes dos veces,

$$g(y) = -e^y\left(y^2 - 2y + 2\right).$$

Solución general:

$$xy^2 - e^y\left(y^2 - 2y + 2\right) = C.$$

Comprobación directa: $\frac{\partial F}{\partial x} = y^2$ y $\frac{\partial F}{\partial y} = 2xy - y^2e^y$, como debía ser. Nótese que $\mu = y$ se anula en $y = 0$, y en efecto $y \equiv 0$ es solución de la ecuación original y hay que anotarla aparte (paso 7).

(d) La ecuación lineal, vista como problema de exactitud. La conocida $y' + p(x)y = q(x)$ se escribe

$$(p(x)y - q(x))\,dx + dy = 0,$$

con $\frac{\partial M}{\partial y} = p$ y $\frac{\partial N}{\partial x} = 0$. El cociente del paso 5 es $\frac{p-0}{1} = p(x)$, función de x sola, así que $\mu = e^{\int p\,dx}$: *el factor integrante de la ecuación lineal no es un truco memorizado, es un caso particular de esta regla.* Con $y' + \frac{2}{x}y = x$ se obtiene $\mu = e^{2\ln x} = x^2$, la ecuación se vuelve $(x^2y)' = x^3$ y

$$x^2y = \frac{x^4}{4} + C \quad\Longrightarrow\quad y = \frac{x^2}{4} + \frac{C}{x^2},$$

que es justo lo que devuelve cualquier sistema de cálculo simbólico.

(e) Un factor integrante que no es único. Sea $y\,dx - x\,dy = 0$, con $\frac{\partial M}{\partial y} = 1$ y $\frac{\partial N}{\partial x} = -1$. Aquí funcionan al menos cuatro factores distintos:

$$\mu_1 = \frac{1}{y^2} \;\to\; \frac{x}{y} = C, \qquad \mu_2 = \frac{1}{x^2} \;\to\; \frac{y}{x} = C,$$

$$\mu_3 = \frac{1}{xy} \;\to\; \ln\left|\frac{x}{y}\right| = C, \qquad \mu_4 = \frac{1}{x^2+y^2} \;\to\; \arctan\left(\frac{y}{x}\right) = C.$$

Las cuatro soluciones describen la misma familia de rectas por el origen, escritas de cuatro maneras. Que haya infinitos factores integrantes no es una anomalía: si μ es uno de ellos y F es el potencial correspondiente, entonces $\mu \cdot \Phi(F)$ también lo es para cualquier función Φ.

Cuidado

El test de exactitud detecta un problema, no lo resuelve. Si ninguno de los dos cocientes del paso 5 sale de una sola variable, no significa que no exista factor integrante: existen *siempre* en teoría (localmente, toda ecuación de primer orden admite uno), pero encontrarlos puede ser tan difícil como resolver la ecuación original. En ese caso conviene volver a la regla de clasificación y probar si la ecuación es separable, homogénea, de Bernoulli o de Riccati antes de insistir.

Pausa para pensar

¿Por qué el test es la igualdad de derivadas cruzadas? Porque $M\,dx+N\,dy$ es exactamente una 1-forma diferencial, y la condición $\frac{\partial M}{\partial y} = \frac{\partial N}{\partial x}$ dice que esa forma es *cerrada*; ser exacta significa ser la diferencial de una función. Es literalmente la misma pregunta que en el capítulo 11 decidía si un campo plano es conservativo, con M y N en el papel de las componentes del campo. Y el factor integrante es el ajuste que convierte una forma cerrada en una exacta cuando la geometría lo permite.

12.2. Cimientos: qué es una ecuación diferencial, qué es una solución y cómo se reconoce cada tipo

Las páginas anteriores han resuelto ecuaciones dando por sabido el vocabulario. Esta sección levanta ese vocabulario desde cero, para quien nunca ha visto una ecuación diferencial: qué objeto es, qué significa resolverlo, cuántas soluciones hay, cómo se seleccionan y, sobre todo, cómo se mira una ecuación y se decide qué método toca. Ninguna de estas definiciones exige nada más que saber derivar e integrar.

12.2.1. La incógnita es una función

Cuando en el instituto se escribe $x^2 - 5x + 6 = 0$, la incógnita es un *número*: hay que decir qué números puestos en el lugar de x convierten la igualdad en verdad. La respuesta es un conjunto finito de números, $\{2, 3\}$, y comprobarla consiste en sustituir y ver que sale $0 = 0$.

En una ecuación diferencial la incógnita no es un número: es una *función entera*. La ecuación no dice cuánto vale la función, dice cómo se relaciona la función con su propia velocidad de cambio, y hay que reconstruir la función a partir de esa relación. La respuesta ya no es un conjunto finito de números, sino una familia infinita de funciones. Y comprobarla sigue consistiendo en sustituir y ver que sale una identidad, solo que ahora la sustitución obliga a derivar antes.

Definición 12.7 (Ecuación diferencial ordinaria, orden y grado)**.** Sea y una función de una sola variable real. Una *ecuación diferencial ordinaria* (abreviado EDO) es una igualdad de la forma

$$F(x,\ y,\ y',\ y'',\ \ldots,\ y^{(n)}) = 0,$$

donde F es una expresión dada y en la que aparece efectivamente al menos una derivada de y. La variable x se llama *variable independiente* y la función incógnita y se llama *variable*

dependiente. El *orden* de la ecuación es n, el de la derivada más alta que aparece. Si además F es un polinomio en las derivadas, el *grado* de la ecuación es la potencia a la que está elevada la derivada de orden máximo, una vez quitados radicales y denominadores.

La palabra *ordinaria* está ahí para excluir un vecino muy próximo. Si la función incógnita depende de *dos o más* variables y en la ecuación aparecen derivadas parciales respecto de varias de ellas, la ecuación se llama *en derivadas parciales* (EDP), y su teoría es otra: no basta con una constante para fijar la solución, hacen falta funciones enteras de datos. Esa es la única frontera, y es nítida: cuántas variables tiene la incógnita.

Ejemplo 12.8 (Cinco ecuaciones, sus órdenes y un intruso)**. (a)** $y' = ky$. Variable independiente x (o t), variable dependiente y. Orden 1, grado 1. Es la ley del interés compuesto y de la desintegración radiactiva.

(b) $y'' + \omega^2 y = 0$. Orden 2, grado 1. Es el muelle, el péndulo pequeño y el circuito LC.

(c) $(y')^2 = 4y$. Orden 1, grado 2: la derivada de orden máximo, que es y', está elevada al cuadrado. Que el grado sea 2 y no 1 tiene una consecuencia inmediata: la ecuación no está *resuelta* en y', y al despejar aparecen dos ramas, $y' = 2\sqrt{y}$ e $y' = -2\sqrt{y}$.

(d) $y^{(4)} = \dfrac{w}{EI}$. Orden 4, grado 1: la viga apoyada bajo carga uniforme. Que el orden sea cuatro anticipa que harán falta cuatro datos para fijar la solución, y en efecto una viga se sujeta por dos extremos con dos condiciones cada uno.

(e) $y' = \sqrt{1 + (y'')^3}$. Elevando al cuadrado queda $(y')^2 = 1 + (y'')^3$: orden 2, grado 3. El grado se lee *después* de quitar la raíz.

El intruso. $\frac{\partial u}{\partial t} = k\,\dfrac{\partial^2 u}{\partial x^2}$ no es una EDO: la incógnita u depende de dos variables, x y t, y aparecen derivadas respecto de las dos. Es la ecuación del calor, una EDP. Lo que la separa de las cinco anteriores no es su dificultad ni su aspecto, es el número de variables de la incógnita. El capítulo la trata más adelante, y el método que la resuelve consiste precisamente en convertirla en dos EDO.

Idea clave

Por qué esto es el idioma de la física. Nadie observa nunca la ley "la posición del planeta en el instante t es esta fórmula". Lo que se observa es una relación local: la aceleración es proporcional a la fuerza, la velocidad de enfriamiento es proporcional a la diferencia de temperatura, la tasa de desintegración es proporcional a la cantidad presente. Todas esas frases son ecuaciones diferenciales, y la fórmula de la trayectoria es lo que hay que *deducir*. Por eso las EDO no son un capítulo más del cálculo: son el puente entre una ley observable y una predicción.

12.2.2. Qué significa ser solución

Definición 12.9 (Solución de una EDO)**.** Una función φ, definida y derivable n veces en un intervalo abierto I, es *solución* de la EDO de orden n dada por $F(x, y, y', \dots, y^{(n)}) = 0$ si al sustituir $y = \varphi(x)$, $y' = \varphi'(x)$, $\dots$, $y^{(n)} = \varphi^{(n)}(x)$ la igualdad se cumple *para todo* $x \in I$.

El intervalo I forma parte de la respuesta: una función no es solución "a secas", lo es en un intervalo.

Se dice que la solución es *explícita* si viene dada como $y = \varphi(x)$, e *implícita* si viene dada por una relación $\Phi(x, y) = 0$ de la que y no se ha despejado, pero que define localmente una función derivable que sí es solución.

Dos avisos que están dentro de la definición y se pasan por alto. El primero: "para todo $x \in I$" significa que no basta con que la igualdad se cumpla en algunos puntos; una identidad es una igualdad de funciones. El segundo: el intervalo importa de verdad. La función $y = 1/(1-x)$ es solución de $y' = y^2$ en $(-\infty, 1)$ y también en $(1, +\infty)$, pero no en ningún intervalo que contenga a 1, porque allí ni siquiera está definida.

Regla: comprobar que una función dada es solución de una EDO dada

Este es el procedimiento que convierte "creo que es solución" en "es solución". Nunca se salta, y en un examen vale tanto como resolver.

1. **Anota el intervalo.** Mira dónde está definida la función candidata y dónde es derivable tantas veces como pida el orden. Ese conjunto, o el mayor intervalo suyo que interese, es donde vas a trabajar. Si hay denominadores o raíces, aquí es donde se detectan.
2. **Calcula todas las derivadas que la ecuación pide.** Si la ecuación es de orden dos, necesitas φ' y φ''. Escríbelas aparte, sin sustituir todavía.
3. **Sustituye en el miembro izquierdo, no en los dos a la vez.** Pon las derivadas en su sitio y *simplifica el lado izquierdo solo*, dejando el derecho quieto. Así el cálculo tiene una meta: llegar a lo que pone a la derecha.
4. **Simplifica hasta una identidad.** El final aceptable es $0 = 0$, o bien "lado izquierdo = lado derecho" escrito con los dos miembros idénticos carácter a carácter. Si te quedas a medias, no has terminado.
5. **Si hay condición inicial, compruébala aparte.** Ser solución de la ecuación y cumplir la condición son dos exigencias distintas; la segunda se verifica evaluando, no derivando.
6. **Si la solución es implícita, deriva la relación.** De $\Phi(x, y) = 0$ se obtiene $\Phi_x + \Phi_y\, y' = 0$ por la regla de la cadena; despeja y' y compara con la ecuación. No hace falta despejar y.

Ejemplo 12.10 (Dos comprobaciones hechas enteras)**. (a) Una solución explícita de una ecuación de primer orden.** ¿Es $y = x^2 + 3$ solución de

$$x\,y' - 2y + 6 = 0 \quad \text{en } \mathbb{R}?$$

Paso 1. $y = x^2 + 3$ es un polinomio: está definida y es derivable en todo $\mathbb{R}$, así que $I = \mathbb{R}$.

Paso 2. $y' = 2x$.

Paso 3. Sustituimos en el miembro izquierdo:

$$x\,y' - 2y + 6 = x \cdot 2x - 2\big(x^2 + 3\big) + 6.$$

Paso 4. Simplificamos:

$$2x^2 - 2x^2 - 6 + 6 = 0.$$

El miembro izquierdo vale 0 para todo $x \in \mathbb{R}$, que es exactamente lo que pide el miembro derecho. Luego $y = x^2 + 3$ *es* solución en $\mathbb{R}$. Obsérvese que la comprobación no ha requerido saber resolver la ecuación: comprobar es siempre más fácil que resolver, y por eso conviene hacerlo siempre.

(b) Una solución de una ecuación de segundo orden. ¿Es $y = e^{-t}\cos 2t$ solución de

$$y'' + 2y' + 5y = 0 \quad \text{en } \mathbb{R}?$$

Paso 1. Producto de funciones derivables infinitas veces: $I = \mathbb{R}$.

Paso 2. Derivamos con la regla del producto:

$$y' = -e^{-t}\cos 2t - 2e^{-t}\sin 2t = -e^{-t}\big(\cos 2t + 2\sin 2t\big),$$

$$y'' = e^{-t}\big(\cos 2t + 2\sin 2t\big) - e^{-t}\big(-2\sin 2t + 4\cos 2t\big) = e^{-t}\big(-3\cos 2t + 4\sin 2t\big).$$

Paso 3. Sustituimos en el miembro izquierdo y sacamos e^{-t} factor común:

$$y'' + 2y' + 5y = e^{-t}\Big[\big(-3\cos 2t + 4\sin 2t\big) - 2\big(\cos 2t + 2\sin 2t\big) + 5\cos 2t\Big].$$

Paso 4. Agrupamos por separado los cosenos y los senos:

$$\cos 2t: \quad -3 - 2 + 5 = 0, \qquad \sin 2t: \quad 4 - 4 = 0.$$

Luego el corchete es idénticamente nulo y $y'' + 2y' + 5y = e^{-t}\cdot 0 = 0$ para todo t. Es solución.

12.2.3. Solución general, particular y singular

Al integrar $y' = 2x$ no sale una función, salen todas las de la forma $y = x^2 + C$: una *familia* de parábolas, una por cada valor de la constante. Este fenómeno es la regla, no la excepción, y tiene un nombre.

Definición 12.11 (Familia uniparamétrica, solución general y solución particular)**.** Una *familia uniparamétrica* de curvas es un conjunto de curvas del plano descrito por una relación $\Phi(x, y, C) = 0$ en la que C recorre un intervalo de números reales: cada valor de C da una curva del conjunto.

Dada una EDO de primer orden, se llama *solución general* a una familia uniparamétrica $\Phi(x, y, C) = 0$ cuyos miembros son todos soluciones de la ecuación y que las agota salvo, a lo sumo, las soluciones singulares definidas más abajo. Una *solución particular* es la que se obtiene fijando un valor concreto del parámetro. Para una EDO de orden n la solución general depende de n constantes arbitrarias $C_1, \ldots, C_n$.

Geométricamente conviene imaginar la familia dibujada: para $y' = 2x$, un haz de parábolas iguales apiladas verticalmente, que llenan el plano sin cortarse, una por cada punto. Para $y' = y$, un haz de exponenciales $y = Ce^x$ que llenan el plano y tampoco se cortan. Que las curvas de la familia no se corten no es casualidad: es la unicidad de la que hablaremos al final de la sección.

Hay, sin embargo, un tercer tipo de solución que no aparece dando valores a C.

Definición 12.12 (Solución singular, versión elemental)**.** Una solución de la ecuación que no se obtiene de la familia $\Phi(x, y, C) = 0$ para ningún valor de C se llama *singular*. Geométricamente suele ser la *envolvente* de la familia: una curva tangente en cada uno de sus puntos a algún miembro de la familia, sin coincidir con ninguno.

La formulación completa, con la discusión de por qué la envolvente delata el fallo de la unicidad, está en la definición 12.106; aquí solo se necesita el ejemplo mínimo.

Ejemplo 12.13 (Una familia de parábolas y su envolvente)**.** Consideremos la ecuación de grado dos $(y')^2 = 4y$, que ya apareció en el ejemplo 12.8(c). Probemos la familia $y = (x + C)^2$. Su derivada es $y' = 2(x + C)$, luego

$$(y')^2 = 4(x + C)^2 = 4y.$$

Es solución para todo C: una familia uniparamétrica de parábolas, todas iguales, desplazadas horizontalmente, todas tangentes al eje x por abajo.

Ahora probemos la función constante $y \equiv 0$. Entonces $y' = 0$ y $(y')^2 = 0 = 4 \cdot 0$: también es solución. ¿Se obtiene de la familia? Habría que resolver $(x + C)^2 = 0$ *para todo* x, y eso solo se cumple si $C = -x$, que no es una constante. Por tanto **no**: la recta $y = 0$ es una solución singular. Y es exactamente la envolvente del haz de parábolas, la recta a la que todas ellas son tangentes.

La moraleja práctica es directa: escribir "la solución general es $y = (x + C)^2$" y parar es perder una solución. Las soluciones singulares no se buscan dando valores a la constante, se buscan mirando dónde el método usado dividió entre algo que podía anularse (aquí, al despejar $y' = \pm 2\sqrt{y}$ se dividió entre $\sqrt{y}$) o calculando la envolvente.

12.2.4. Problema de valor inicial y problema de contorno

Una familia infinita de soluciones no sirve para predecir nada. Para quedarse con una hay que añadir datos, y hay dos formas muy distintas de añadirlos.

Definición 12.14 (Problema de valor inicial y problema de contorno)**.** Un *problema de valor inicial* (PVI, o problema de Cauchy) consiste en una EDO de orden n junto con n condiciones impuestas *todas en un mismo punto* x_0:

$$y(x_0) = a_0, \quad y'(x_0) = a_1, \quad \ldots, \quad y^{(n-1)}(x_0) = a_{n-1}.$$

Un *problema de contorno* consiste en una EDO de orden $n \geq 2$ junto con condiciones repartidas entre *dos puntos distintos* $x = \alpha$ y $x = \beta$, los extremos del intervalo de trabajo.

La diferencia parece menor y es enorme. En el PVI se conoce todo el estado del sistema en un instante y se pregunta por su futuro: es la situación de la mecánica, donde posición y velocidad iniciales determinan el movimiento. En el problema de contorno se conoce un poco en cada extremo y se pregunta por lo que pasa entre medias: es la situación de una viga sujeta por sus dos apoyos o de una barra con las dos puntas a temperatura fijada.

Ejemplo 12.15 (Un PVI y tres problemas de contorno con la misma ecuación)**.** Tomamos siempre $y'' + y = 0$, cuya solución general es $y = C_1 \cos t + C_2 \sin t$ (se deducirá en la sección siguiente; aquí basta comprobar que $y'' = -C_1 \cos t - C_2 \sin t = -y$).

(a) Problema de valor inicial. $y'' + y = 0$ con $y(0) = 1$, $y'(0) = 0$. De $y(0) = C_1 = 1$ y $y'(0) = C_2 = 0$ sale $y = \cos t$, y *solo* $\cos t$. Exactamente una solución.

(b) Contorno con infinitas soluciones. $y'' + y = 0$ con $y(0) = 0$, $y(\pi) = 0$. La primera condición da $C_1 = 0$, luego $y = C_2 \sin t$. La segunda pide $C_2 \sin \pi = 0$, que se cumple *sea cual sea* C_2 porque $\sin \pi = 0$. Hay infinitas soluciones, una recta entera de ellas.

(c) Contorno sin ninguna solución. $y'' + y = 0$ con $y(0) = 0$, $y(\pi) = 1$. Igual que antes, $C_1 = 0$ e $y = C_2 \sin t$; ahora la segunda condición pide $C_2 \cdot 0 = 1$, imposible. El problema no tiene solución.

(d) Contorno con solución única. $y'' + y = 0$ con $y(0) = 0$, $y(\pi/2) = 3$. De nuevo $C_1 = 0$; ahora $C_2 \sin(\pi/2) = C_2 = 3$, luego $y = 3 \sin t$ y nada más.

Cuidado

[Un PVI y un problema de contorno no se parecen en nada] En condiciones razonables —las del teorema de Picard–Lindelöf que veremos enseguida— un problema de valor inicial tiene *exactamente una* solución. Un problema de contorno, en cambio, puede tener una, ninguna o infinitas, y los tres casos se dan con la misma ecuación sin más que mover los extremos, como acaba de verse. La razón es estructural: el PVI se resuelve avanzando paso a paso desde x_0, mientras que el problema de contorno es una condición *global* que puede ser incompatible. Esos valores excepcionales del intervalo en los que aparecen infinitas soluciones son los *autovalores* del problema, y de ahí nace toda la teoría de Sturm–Liouville y la separación de variables de las EDP. Cuando en un problema de contorno le salgan a alguien infinitas soluciones, no ha hecho nada mal: ha encontrado un modo propio.

12.2.5. El vocabulario que clasifica: lineal, homogénea, coeficientes constantes

Definición 12.16 (Ecuación lineal, término independiente y coeficientes)**.** Una EDO de orden n es *lineal* si se puede escribir en la forma

$$a_n(x)\, y^{(n)} + a_{n-1}(x)\, y^{(n-1)} + \cdots + a_1(x)\, y' + a_0(x)\, y = g(x),$$

es decir, si y y todas sus derivadas aparecen en *primera potencia*, sin multiplicarse entre sí y sin estar dentro de ninguna función (nada de y^2, $y\,y'$, $\sin y$, e^y, $\sqrt{y'}$). Los a_k son los *coeficientes* y g es el *término independiente*. Obsérvese que sobre la variable independiente x no se impone

ninguna restricción: puede aparecer como quiera. Si no admite esa forma, la ecuación es *no lineal*.

Una ecuación lineal se llama *homogénea* cuando $g \equiv 0$, y *completa* (o no homogénea) cuando $g \not\equiv 0$. Se dice de *coeficientes constantes* si todos los a_k son números, y de *coeficientes variables* si alguno depende de x.

Dos consecuencias inmediatas de la linealidad, que son la razón de que esta clase sea la clase importante.

Proposición 12.17 (Operador diferencial lineal y principio de superposición)**.** *Escribamos* $L[y] = a_n(x)y^{(n)} + \cdots + a_1(x)y' + a_0(x)y$. *Entonces* L *es un* operador diferencial lineal*: para cualesquiera funciones* u, v *suficientemente derivables y cualesquiera constantes* α, β*,*

$$L[\alpha u + \beta v] = \alpha\, L[u] + \beta\, L[v].$$

En consecuencia:

1. *(*Principio de superposición homogéneo*) Si* $L[u] = 0$ *y* $L[v] = 0$*, entonces* $L[\alpha u + \beta v] = 0$*: las soluciones de la ecuación homogénea forman un espacio vectorial.*
2. *(*Estructura de la completa*) Si* y_p *cumple* $L[y_p] = g$ *y* y_h *cumple* $L[y_h] = 0$*, entonces* $L[y_h + y_p] = g$*. Y recíprocamente, si* $L[y_1] = L[y_2] = g$ *entonces* $L[y_1 - y_2] = 0$*. Por tanto la solución general de la completa es*

$$y = \underbrace{y_h}_{\text{general de la homogénea}} + \underbrace{y_p}_{\text{una particular cualquiera}}.$$

Demostración. La linealidad de L es la linealidad de la derivada aplicada término a término: cada $(\alpha u + \beta v)^{(k)} = \alpha u^{(k)} + \beta v^{(k)}$, y al multiplicar por $a_k(x)$ y sumar se saca α y β factor común. Los dos apartados son sustituciones directas: $L[\alpha u + \beta v] = \alpha \cdot 0 + \beta \cdot 0 = 0$ para el primero, y $L[y_h + y_p] = 0 + g = g$ para el segundo. Para el recíproco, $L[y_1 - y_2] = L[y_1] - L[y_2] = g - g = 0$. □

Este es el motivo de que toda la teoría lineal tenga la misma pinta: resolver la homogénea (un espacio vectorial cuya dimensión es el orden) y añadir una particular cualquiera. Nada de esto vale para las ecuaciones no lineales, donde sumar dos soluciones da casi siempre algo que no es solución.

Regla: clasificar una EDO cualquiera y decidir qué método toca

Ante una ecuación desconocida, se responde a estas preguntas *en este orden*. Cada respuesta recorta el terreno, y al final del recorrido el método está decidido.

1. **¿Es ordinaria o en derivadas parciales?** Mira cuántas variables tiene la incógnita. Si hay derivadas parciales respecto de dos variables, esto no es una EDO y ninguna de las recetas siguientes sirve.
2. **¿Cuál es el orden?** La derivada más alta. Es el dato que dice cuántas constantes llevará la solución general y cuántas condiciones hacen falta.

3. **¿Cuál es el grado?** Si la ecuación es polinómica en las derivadas, mira la potencia de la derivada de orden máximo tras quitar raíces y denominadores. Grado mayor que uno avisa de dos cosas: la ecuación no está resuelta en $y^{(n)}$ y es muy probable que haya soluciones singulares.

4. **¿Es lineal?** Comprueba que $y, y', \dots, y^{(n)}$ aparecen en primera potencia, sin productos entre ellas y sin meterse dentro de funciones. Es la pregunta que más terreno recorta, porque la respuesta afirmativa abre toda la maquinaria de la proposición 12.17.

5. **Si es lineal: ¿homogénea o completa?** Mira el término independiente g. Si $g = 0$ es homogénea; si no, es completa y habrá que resolver la homogénea y buscar además una particular.

6. **Si es lineal: ¿coeficientes constantes o variables?** Si son números, la homogénea se resuelve con la ecuación característica. Si dependen de x, no hay característica que valga: toca Euler–Cauchy si la ecuación tiene esa forma, series de potencias o Frobenius si no.

7. **Si es de primer orden, recorre las casillas elementales.** ¿Separable ($y' = g(x)h(y)$)? ¿Lineal ($y' + p(x)y = q(x)$)? ¿Exacta ($M_y = N_x$)? ¿Homogénea de grado cero ($f(\lambda x, \lambda y) = f(x, y)$)? ¿Bernoulli? ¿Riccati? La lista detallada, con las sustituciones de cada casilla, está en la regla de la página anterior; lo que se añade aquí es que estas preguntas van *después* de las seis anteriores, no antes.

8. **Anota lo que puedes perder.** Cada división por una expresión que pueda anularse (por $h(y)$ al separar, por y^n en Bernoulli, por $\sqrt{y}$ al despejar en una ecuación de grado dos) es una solución potencialmente perdida. Escríbela en el margen ahora, antes de olvidarla.

Ejemplo 12.18 (La regla aplicada a seis ecuaciones, una por una)**. (1)** $y' + 2xy = x$. Ordinaria; orden 1; grado 1; *lineal* (y e y' en primera potencia; que el coeficiente sea $2x$ da igual, la restricción es sobre y, no sobre x); *completa*, porque $g(x) = x \neq 0$; *coeficientes variables*. Además, de primer orden y lineal, luego factor integrante $\mu = e^{\int 2x\,dx} = e^{x^2}$. *Conclusión:* método del factor integrante, y la solución es $y = \frac{1}{2} + Ce^{-x^2}$. (Comprobación: $y' = -2Cxe^{-x^2}$ y $2xy = x + 2Cxe^{-x^2}$, cuya suma es x.) Un detalle: también era separable, porque $y' = x(1 - 2y)$; las dos rutas dan lo mismo y se elige la más cómoda.

(2) $y'' - 5y' + 6y = 0$. Orden 2; grado 1; lineal; *homogénea*; *coeficientes constantes*. *Conclusión:* ecuación característica $r^2 - 5r + 6 = 0$, con raíces 2 y 3, luego $y = C_1e^{2x} + C_2e^{3x}$. Por la proposición 12.17 esas dos exponenciales generan un espacio de dimensión dos, y no hay más soluciones.

(3) $y' = \dfrac{x + y}{x - y}$. Orden 1; grado 1; *no lineal*, porque y aparece en el denominador. Probamos la casilla de homogénea de grado cero sustituyendo $x \mapsto \lambda x$, $y \mapsto \lambda y$:

$$\frac{\lambda x + \lambda y}{\lambda x - \lambda y} = \frac{\lambda(x + y)}{\lambda(x - y)} = \frac{x + y}{x - y},$$

y el segundo miembro no cambia: sí es homogénea de grado cero. *Conclusión:* cambio $v = y/x$, que la vuelve separable. El resultado es la familia implícita

$$\arctan\frac{y}{x} - \tfrac{1}{2}\ln(x^2+y^2) = C,$$

que en polares es $\theta - \ln r = C$, es decir, espirales logarítmicas. (Comprobación: derivando la relación implícita respecto de x y despejando sale exactamente $y' = (x+y)/(x-y)$.)

(4) $(2xy+y^2)\,\mathrm{d}x + (x^2+2xy)\,\mathrm{d}y = 0$. Orden 1; grado 1; no lineal (hay y^2 y productos xy multiplicando a $\mathrm{d}y$). Escrita en forma diferencial, probamos el test de exactitud con $M = 2xy+y^2$ y $N = x^2+2xy$:

$$\frac{\partial M}{\partial y} = 2x+2y, \qquad \frac{\partial N}{\partial x} = 2x+2y.$$

Coinciden: es *exacta*. *Conclusión:* se busca F con $F_x = M$. Integrando en x, $F = x^2y + xy^2 + g(y)$; derivando en y, $F_y = x^2 + 2xy + g'(y)$, que debe ser N, luego $g' = 0$. La solución general es $x^2y + xy^2 = C$. Con la condición $y(1) = 1$ sale $C = 2$.

(5) $x^2y'' + xy' + (x^2-1)y = 0$. Orden 2; grado 1; *lineal*; *homogénea*; *coeficientes variables*, y además no es de tipo Euler–Cauchy porque el coeficiente de y no es constante. *Conclusión:* no hay ecuación característica ni fórmula elemental; toca el método de Frobenius en torno al punto singular regular $x = 0$. Es la ecuación de Bessel de orden uno, y sus soluciones son $y = C_1J_1(x) + C_2Y_1(x)$: funciones nuevas, definidas por sus series. Este es el ejemplo que enseña por qué la pregunta 6 de la regla no es un capricho: cambiar coeficientes constantes por variables cambia el método de raíz.

(6) $(y')^2 = 4y$. Orden 1; *grado* 2; no lineal; no está resuelta en y'. *Conclusión:* primero se despeja, $y' = \pm 2\sqrt{y}$, y cada rama se trata por separado como ecuación separable; y por ser de grado mayor que uno hay que buscar solución singular. Como se vio en el ejemplo 12.13, la solución general es $y = (x+C)^2$ y además está la singular $y \equiv 0$, que ningún C produce.

Cuidado

[La palabra "homogénea" significa dos cosas distintas en este capítulo] Esta es, con diferencia, la mayor fuente de confusión del tema, y no es culpa del lector: es una colisión histórica de vocabulario. Los dos sentidos son:

- **Homogénea = término independiente nulo.** Se aplica solo a ecuaciones *lineales*. La ecuación $a_ny^{(n)} + \cdots + a_0y = g$ es homogénea cuando $g \equiv 0$. Su contrario es *completa*. Este sentido es el de "$y''+3y'+2y = 0$ es homogénea, y $y''+3y'+2y = e^x$ es completa". Lo relevante aquí es la proposición 12.17: las soluciones forman un espacio vectorial.
- **Homogénea de grado cero = invariante por dilataciones.** Se aplica a ecuaciones de primer orden $y' = f(x,y)$, *casi siempre no lineales*, y significa que $f(\lambda x, \lambda y) = f(x,y)$ para todo $\lambda \neq 0$. Su contrario no tiene nombre. Lo relevante aquí es que el cambio $v = y/x$ la convierte en separable.

Los dos sentidos son *independientes*, y se dan las cuatro combinaciones posibles. La ecuación $y' = y/x$ es homogénea en los dos sentidos a la vez: es lineal con término

independiente nulo, y al dilatar, $\lambda y/(\lambda x) = y/x$ no cambia. La ecuación $y' = y/x + x$ es lineal *completa* (término independiente $x \neq 0$) y *no* es homogénea de grado cero, porque al dilatar el segundo sumando pasa a valer λx. La ecuación $y' = (x^2 + y^2)/(xy)$ sí es homogénea de grado cero, y ni siquiera es lineal, de modo que la palabra solo puede tener ahí el segundo sentido. Y $y'' + y = \sin x$ es lineal no homogénea en el primer sentido, mientras que el segundo sentido ni se le aplica, porque no es de primer orden.

Cómo no equivocarse nunca. Pregúntese siempre: ¿estoy hablando de un *término independiente* o de una *dilatación*? Si la ecuación no es lineal, la palabra "homogénea" solo puede referirse a la dilatación. Si la ecuación es de orden dos o más, solo puede referirse al término independiente. La ambigüedad real se da únicamente en las lineales de primer orden, y allí conviene decir la frase completa: "homogénea de grado cero" o "con término independiente nulo".

12.2.6. Ecuaciones de variables separables

Definición 12.19 (Ecuación de variables separables)**.** Una EDO de primer orden es de *variables separables* si puede escribirse como

$$\frac{\mathrm{d}y}{\mathrm{d}x} = g(x)\,h(y),$$

es decir, si el segundo miembro se factoriza en un trozo que solo depende de x por otro que solo depende de y. Equivalentemente, en forma diferencial, $p(x)\,\mathrm{d}x + q(y)\,\mathrm{d}y = 0$.

El método se enuncia siempre igual: "se pasan las y a un lado, las x al otro y se integra". Esa frase esconde una manipulación de diferenciales que merece explicación, porque hecha sin entenderla es magia.

Regla: resolver una ecuación de variables separables

1. **Comprueba que factoriza.** Escribe el segundo miembro como $g(x)h(y)$. Si no factoriza, esta no es la casilla.
2. **Anota las soluciones de equilibrio.** Resuelve $h(y) = 0$. Cada raíz $y = c$ da la solución constante $y \equiv c$ (en efecto, $y' = 0 = g(x)h(c)$). Estas soluciones se van a perder en el paso siguiente, así que se apuntan *ahora*.
3. **Separa e integra.** Dividiendo entre $h(y)$ y usando la notación de diferenciales,

$$\frac{\mathrm{d}y}{h(y)} = g(x)\,\mathrm{d}x \quad \Longrightarrow \quad \int \frac{\mathrm{d}y}{h(y)} = \int g(x)\,\mathrm{d}x + C.$$

   Basta *una* constante, en un solo lado.
4. **Despeja y si se puede.** Si no se puede, la solución implícita es una respuesta perfectamente válida; no se fuerza el despeje.
5. **Impón la condición inicial.** Sustituye el punto dado y calcula C. Hazlo antes de simplificar radicales, para no perder el signo de la rama correcta.

6. **Determina el intervalo y comprueba.** Mira dónde está definida la respuesta y qué intervalo alrededor del punto inicial se puede tomar. Después deriva y sustituye en la ecuación original.

Pausa para pensar

¿Por qué se pueden "separar los diferenciales"? Porque no se están separando: se está usando la regla de la cadena al revés. Sea H una primitiva de $1/h$, es decir $H' = 1/h$. Si $y(x)$ es solución de $y' = g(x)h(y)$ con $h(y(x)) \neq 0$, considérese la función compuesta $x \mapsto H(y(x))$. Por la regla de la cadena,

$$\frac{\mathrm{d}}{\mathrm{d}x} H\big(y(x)\big) = H'\big(y(x)\big)\, y'(x) = \frac{y'(x)}{h(y(x))} = g(x).$$

O sea: $H(y(x))$ y una primitiva de g tienen la misma derivada, luego se diferencian en una constante:

$$H\big(y(x)\big) = \int g(x)\,\mathrm{d}x + C,$$

que es exactamente el resultado de "separar e integrar". La notación $\mathrm{d}y/h(y) = g(x)\,\mathrm{d}x$ es una taquigrafía cómoda de este razonamiento; lo que se hace de verdad es integrar por sustitución. Y ahora se ve por qué el paso 2 de la regla es obligatorio: todo el argumento supone $h(y(x)) \neq 0$, así que los ceros de h hay que tratarlos aparte.

Ejemplo 12.20 (Dos separables resueltas enteras)**. (a)** $y' = 3x^2 y$ con $y(0) = 5$.

Paso 1. Factoriza con $g(x) = 3x^2$ y $h(y) = y$.

Paso 2. $h(y) = 0$ solo en $y = 0$: apuntamos la solución de equilibrio $y \equiv 0$.

Paso 3. Separando e integrando,

$$\int \frac{\mathrm{d}y}{y} = \int 3x^2\,\mathrm{d}x + C \quad \Longrightarrow \quad \ln|y| = x^3 + C.$$

Paso 4. Exponenciando, $|y| = e^C e^{x^3}$, y absorbiendo el signo y la constante en un único $A = \pm e^C$ queda $y = A\,e^{x^3}$. Nótese que $A = 0$ recupera la solución de equilibrio: aquí no se pierde nada, pero solo lo sabemos porque la hemos apuntado.

Paso 5. Condición inicial: $y(0) = A\,e^0 = A = 5$, luego

$$y = 5\,e^{x^3}.$$

Paso 6. Está definida en todo $\mathbb{R}$. Comprobación: $y' = 5\,e^{x^3} \cdot 3x^2 = 3x^2\big(5e^{x^3}\big) = 3x^2 y$. Correcto. Y $y(0) = 5$. Correcto.

(b) $y' = \dfrac{1+y^2}{1+x^2}$ con $y(0) = 1$.

Paso 1. Factoriza con $g(x) = 1/(1+x^2)$ y $h(y) = 1 + y^2$.

Paso 2. $h(y) = 1 + y^2$ no se anula nunca: no hay soluciones de equilibrio y no se pierde nada al dividir.

Paso 3. Separando,

$$\int \frac{\mathrm{d}y}{1+y^2} = \int \frac{\mathrm{d}x}{1+x^2} + C \quad \Longrightarrow \quad \arctan y = \arctan x + C.$$

Paso 5 (aquí conviene adelantarlo). Con $x = 0$, $y = 1$: $\arctan 1 = \arctan 0 + C$, luego $C = \pi/4$.

Paso 4. Despejamos tomando tangente y usando la fórmula de la tangente de una suma con $\tan(\pi/4) = 1$:

$$y = \tan\Big(\arctan x + \frac{\pi}{4}\Big) = \frac{\tan(\arctan x) + 1}{1 - \tan(\arctan x)} = \frac{1+x}{1-x}.$$

Paso 6. El intervalo: la expresión explota en $x = 1$, y el punto inicial es $x = 0$, así que la solución del problema de valor inicial vive en $(-\infty, 1)$. Comprobación:

$$y' = \frac{(1-x)+(1+x)}{(1-x)^2} = \frac{2}{(1-x)^2},$$

mientras que

$$\frac{1+y^2}{1+x^2} = \frac{1+\frac{(1+x)^2}{(1-x)^2}}{1+x^2} = \frac{\frac{(1-x)^2+(1+x)^2}{(1-x)^2}}{1+x^2} = \frac{\frac{2+2x^2}{(1-x)^2}}{1+x^2} = \frac{2}{(1-x)^2}.$$

Coinciden. Y $y(0) = 1$. Correcto.

Este segundo ejemplo enseña algo que conviene interiorizar pronto: la ecuación es perfectamente apacible en todo el plano (el segundo miembro es continuo y con derivadas continuas en todas partes) y aun así la solución muere en $x = 1$. La existencia garantizada es siempre *local*.

12.2.7. Ecuaciones lineales de primer orden: de dónde sale el factor integrante

Definición 12.21 (Forma estándar de la lineal de primer orden)**.** Una EDO lineal de primer orden está en *forma estándar* cuando se ha escrito como

$$y' + p(x)\, y = q(x),$$

con el coeficiente de y' igual a 1. Si viene como $a(x)y' + b(x)y = c(x)$, se divide entre $a(x)$ en los intervalos donde a no se anula, y esa división forma parte del método.

Ahora la deducción, que es el corazón del asunto y no debe presentarse como una receta caída del cielo. La idea es esta: si el miembro izquierdo fuese la derivada de un producto, la ecuación se resolvería integrando de un tirón. Busquemos entonces una función $\mu(x)$ tal que al multiplicar por ella el miembro izquierdo se convierta exactamente en $(\mu y)'$. Se impone

$$\mu\, y' + \mu\, p\, y \overset{!}{=} (\mu\, y)' = \mu\, y' + \mu'\, y.$$

Los términos con y' ya son iguales, así que la condición se reduce a

$$\mu'\, y = \mu\, p\, y \qquad \text{para toda } y, \qquad \text{es decir} \qquad \mu' = p\, \mu.$$

Pero $\mu' = p\mu$ es una ecuación separable en μ, y sabemos resolverla: $\mathrm{d}\mu/\mu = p\,\mathrm{d}x$, luego $\ln|\mu| = \int p\,\mathrm{d}x$ y

$$\boxed{\mu(x) = e^{\int p(x)\,\mathrm{d}x}}$$

(no hace falta constante: cualquier factor sirve, y se toma el más simple). Con ese μ la ecuación multiplicada se lee $(\mu y)' = \mu q$, e integrando,

$$\mu(x)\,y = \int \mu(x)\,q(x)\,\mathrm{d}x + C \qquad \Longrightarrow \qquad y = \frac{1}{\mu(x)}\left(\int \mu(x)\,q(x)\,\mathrm{d}x + C\right).$$

El factor integrante no es un truco: es la solución de una ecuación separable auxiliar, y por eso existe siempre.

Regla: resolver una lineal de primer orden con factor integrante

1. **Forma estándar.** Divide hasta dejar $y' + p(x)y = q(x)$, y anota en qué intervalo es lícita esa división.
2. **Calcula $\mu = e^{\int p\,\mathrm{d}x}$.** Sin constante de integración. Simplifica exponenciales de logaritmos: $e^{2\ln x} = x^2$, $e^{-\ln x} = 1/x$.
3. **Multiplica y reconoce la derivada del producto.** Escribe directamente $(\mu y)' = \mu q$. Si al derivar μy no sale el miembro izquierdo, hay un error en μ.
4. **Integra los dos lados.** $\mu y = \int \mu q\,\mathrm{d}x + C$. Aquí sí va la constante, y va una sola.
5. **Despeja y dividiendo entre μ.** Observa la estructura: sale $y = y_p + C/\mu$, es decir, una particular más la solución general de la homogénea, tal como anticipaba la proposición 12.17.
6. **Condición inicial y comprobación.** Sustituye el punto, calcula C, deriva y verifica.

Ejemplo 12.22 (Dos lineales resueltas enteras)**. (a)** $x\,y' + y = 3x^2$ para $x > 0$, con $y(1) = 2$.

Paso 1. Dividimos entre x (lícito porque $x > 0$):

$$y' + \frac{1}{x}\,y = 3x, \qquad p(x) = \frac{1}{x}, \quad q(x) = 3x.$$

Paso 2. $\mu = e^{\int \mathrm{d}x/x} = e^{\ln x} = x$.

Paso 3. Multiplicando: $x\,y' + y = 3x^2$, y en efecto $xy' + y = (xy)'$. Así que $(xy)' = 3x^2$.

Paso 4. Integrando, $xy = x^3 + C$.

Paso 5. Despejando,

$$y = x^2 + \frac{C}{x}.$$

La particular es x^2 y la homogénea es C/x, como anunciaba la estructura.

Paso 6. Condición inicial: $y(1) = 1 + C = 2$, luego $C = 1$ y

$$y = x^2 + \frac{1}{x} = \frac{x^3 + 1}{x}.$$

Comprobación: $y' = 2x - 1/x^2$, y entonces

$$y' + \frac{y}{x} = 2x - \frac{1}{x^2} + \frac{x^2 + \frac{1}{x}}{x} = 2x - \frac{1}{x^2} + x + \frac{1}{x^2} = 3x.$$

Correcto, y $y(1) = 2$.

(b) $y' - 2y = e^{3t}$ con $y(0) = 2$.

Pasos 1 y 2. Ya está en forma estándar, con $p(t) = -2$ y $q(t) = e^{3t}$; por tanto $\mu = e^{\int (-2)\,\mathrm{d}t} = e^{-2t}$.

Pasos 3 y 4. Multiplicando, $\left(e^{-2t}y\right)' = e^{-2t}e^{3t} = e^{t}$; integrando, $e^{-2t}y = e^{t} + C$.

Paso 5. Despejando, $y = e^{3t} + C\,e^{2t}$.

Paso 6. $y(0) = 1 + C = 2$ da $C = 1$, luego $y = e^{3t} + e^{2t}$. Comprobación: $y' = 3e^{3t} + 2e^{2t}$ y

$$y' - 2y = 3e^{3t} + 2e^{2t} - 2e^{3t} - 2e^{2t} = e^{3t}.$$

Correcto, y $y(0) = 2$. Obsérvese de paso el peligro clásico: si la condición inicial hubiera sido $y(0) = 1$, saldría $C = 0$ y la respuesta sería $y = e^{3t}$ sin rastro de la homogénea; eso no significa que la homogénea no exista, significa que su coeficiente es cero en ese caso concreto.

12.2.8. Ecuaciones exactas: la definición que faltaba

El método de las exactas, con su regla de siete pasos y sus cinco ejemplos, ya se ha desarrollado más arriba en este mismo capítulo. Lo que faltaba era decir con precisión qué es una ecuación exacta y por qué el criterio de las derivadas cruzadas es el criterio correcto; eso es lo que se hace aquí, sin repetir el procedimiento.

Definición 12.23 (Ecuación exacta y factor integrante)**.** Sea la ecuación escrita en forma diferencial

$$M(x,y)\,\mathrm{d}x + N(x,y)\,\mathrm{d}y = 0$$

con M, N de clase C^1 en un abierto $D \subset \mathbb{R}^2$. La ecuación es *exacta* en D si existe una función $F : D \to \mathbb{R}$ de clase C^2, llamada *función potencial*, tal que

$$\frac{\partial F}{\partial x} = M \qquad \text{y} \qquad \frac{\partial F}{\partial y} = N \quad \text{en todo } D.$$

En tal caso las soluciones son exactamente las curvas de nivel $F(x,y) = C$.

Si la ecuación no es exacta, una función $\mu(x,y)$ no nula tal que $\mu M\,\mathrm{d}x + \mu N\,\mathrm{d}y = 0$ sí lo sea se llama *factor integrante* de la ecuación.

Proposición 12.24 (Criterio de exactitud)**.** *Con las hipótesis anteriores y D simplemente conexo, la ecuación es exacta en D si y solo si*

$$\frac{\partial M}{\partial y} = \frac{\partial N}{\partial x} \qquad \textit{en todo } D.$$

Demostración. Si existe F con $F_x = M$ y $F_y = N$, entonces $M_y = F_{xy}$ y $N_x = F_{yx}$, iguales por el teorema de Schwarz al ser F de clase C^2. Esa es la implicación fácil, y es la que se usa

como test: si las derivadas cruzadas no coinciden, no puede haber potencial. El recíproco es el teorema del campo conservativo del capítulo 11 aplicado al campo plano (M, N), y necesita la hipótesis topológica sobre D. □

Que las soluciones sean las curvas de nivel es inmediato: si $F(x, y(x)) = C$, derivando por la regla de la cadena queda $F_x + F_y\, y' = 0$, es decir $M + N\, y' = 0$, que es la ecuación de partida.

Ejemplo 12.25 (Una exacta y una rescatada con factor integrante)**. (a) Exacta de entrada.** Ya se resolvió en el ejemplo 12.18(4): $(2xy+y^2)\,\mathrm{d}x+(x^2+2xy)\,\mathrm{d}y = 0$, con $M_y = N_x = 2x+2y$ y potencial $F = x^2y + xy^2$. La solución general es $x^2y + xy^2 = C$. Se puede comprobar sin despejar: derivando implícitamente, $(2xy + y^2) + (x^2 + 2xy)y' = 0$, que es la ecuación original.

(b) No exacta, rescatada. Sea

$$(2y^2 + 3x)\,\mathrm{d}x + 2xy\,\mathrm{d}y = 0.$$

Test: $M_y = 4y$ y $N_x = 2y$, distintos, luego no es exacta. Buscamos factor integrante de una sola variable:

$$\frac{1}{N}\left(\frac{\partial M}{\partial y} - \frac{\partial N}{\partial x}\right) = \frac{4y - 2y}{2xy} = \frac{2y}{2xy} = \frac{1}{x},$$

que depende solo de x. Por tanto $\mu(x) = e^{\int \mathrm{d}x/x} = x$. Multiplicando:

$$(2xy^2 + 3x^2)\,\mathrm{d}x + 2x^2y\,\mathrm{d}y = 0,$$

y ahora $\frac{\partial}{\partial y}(2xy^2 + 3x^2) = 4xy = \frac{\partial}{\partial x}(2x^2y)$: ya es exacta. Integramos M en x: $F = x^2y^2 + x^3 + g(y)$; derivando en y, $F_y = 2x^2y + g'(y)$, que debe ser $2x^2y$, luego $g' = 0$. La solución general es

$$x^2y^2 + x^3 = C.$$

Comprobación: derivando implícitamente, $2xy^2 + 2x^2yy' + 3x^2 = 0$, y dividiendo entre x (para $x \neq 0$) queda $2y^2 + 2xyy' + 3x = 0$, que es la ecuación original. Como $\mu = x$ se anula en $x = 0$, la recta $x = 0$ hay que revisarla aparte, según el aviso habitual.

12.2.9. Existencia y unicidad: qué garantiza y qué no

Todo lo anterior supone tácitamente que la solución existe y es la única. Conviene saber cuándo eso es cierto, porque a veces no lo es.

Definición 12.26 (Condición de Lipschitz respecto de y)**.** Sea f definida en un rectángulo $R = [x_0 - a, x_0 + a] \times [y_0 - b, y_0 + b]$. Se dice que f es *lipschitziana respecto de y* en R si existe una constante $L \geq 0$ tal que

$$|f(x, y_1) - f(x, y_2)| \leq L\,|y_1 - y_2| \qquad \text{para todos } (x, y_1), (x, y_2) \in R.$$

La constante L se llama constante de Lipschitz.

Traducido: la condición dice que f no puede cambiar arbitrariamente deprisa al mover y; el cociente incremental en la dirección y está acotado por L. Geométricamente, para cada x fijo la gráfica de $y \mapsto f(x,y)$ no puede tener pendientes mayores que L en valor absoluto, tampoco en los puntos donde no sea derivable. De ahí el criterio práctico más usado: *si $\frac{\partial f}{\partial y}$ existe y es continua en R, entonces f es lipschitziana en R*, con $L = \max_R |\frac{\partial f}{\partial y}|$, por el teorema del valor medio aplicado a $y \mapsto f(x,y)$. Lo recíproco es falso: $f(x,y) = |y|$ es lipschitziana con $L = 1$ y no es derivable en $y = 0$.

Teorema 12.27 (Picard–Lindelöf, enunciado elemental)**.** *Sea f continua en un rectángulo R centrado en (x_0, y_0) y lipschitziana respecto de y en R. Entonces el problema de valor inicial*

$$y' = f(x,y), \qquad y(x_0) = y_0$$

tiene una y solo una solución, definida en algún intervalo $(x_0 - h, x_0 + h)$ con $h > 0$.

El teorema está demostrado con todo detalle más adelante en este capítulo (teorema 12.4 y la sección de fundamentos); lo que interesa ahora es leerlo bien. Dice tres cosas y calla una cuarta. Dice que hay solución; dice que es única; dice que la garantía es *local*, en un intervalo posiblemente diminuto. Y calla cuánto vale h: puede ser mucho menor que el rectángulo, como enseñó el ejemplo 12.20(b), donde la solución moría en $x = 1$ pese a que f era buena en todo el plano.

Ejemplo 12.28 (El contraejemplo obligatorio: $y' = y^{2/3}$ con $y(0) = 0$)**.** Aquí $f(x,y) = y^{2/3}$, que es continua en todo el plano. Pero

$$\frac{\partial f}{\partial y} = \frac{2}{3}\, y^{-1/3},$$

que no está acotada cerca de $y = 0$. Comprobemos que además falla la condición de Lipschitz: tomando $y_1 = h > 0$ e $y_2 = 0$,

$$\frac{|f(x,h) - f(x,0)|}{|h - 0|} = \frac{h^{2/3}}{h} = h^{-1/3} \xrightarrow[h \to 0^+]{} +\infty,$$

así que ninguna constante L puede servir en un rectángulo que contenga al origen. Sin Lipschitz, Picard no dice nada sobre la unicidad. Y en efecto la unicidad falla, de manera espectacular.

Primera solución. La función constante $y \equiv 0$: cumple $y' = 0$ y $y^{2/3} = 0$, y también $y(0) = 0$.

Segunda solución. Separando variables en la región $y > 0$:

$$\int y^{-2/3}\, \mathrm{d}y = \int \mathrm{d}x \quad \Longrightarrow \quad 3\, y^{1/3} = x + K \quad \Longrightarrow \quad y = \Big(\frac{x + K}{3}\Big)^3.$$

Con $y(0) = 0$ sale $K = 0$, es decir $y = x^3/27$. Comprobación: $y' = 3x^2/27 = x^2/9$, y por otro lado $y^{2/3} = (x^3/27)^{2/3} = x^2/9$. Coinciden.

Infinitas soluciones. Lo grave es que se pueden *pegar* las dos anteriores. Fijados dos números $a \le 0 \le b$, defínase

$$y_{a,b}(x) = \begin{cases} \left(\dfrac{x-a}{3}\right)^3, & x < a, \\ 0, & a \le x \le b, \\ \left(\dfrac{x-b}{3}\right)^3, & x > b. \end{cases}$$

Esta función es derivable en todo $\mathbb{R}$ (en $x = b$, por ejemplo, la derivada por la izquierda es 0 y la derivada por la derecha es $\lim_{x\to b^+}\left((x-b)/3\right)^2 = 0$), cumple $y_{a,b}(0) = 0$ y verifica la ecuación en los tres tramos: en el central porque $0' = 0 = 0^{2/3}$, y en los laterales porque

$$y' = 3 \cdot \left(\frac{x-c}{3}\right)^2 \cdot \frac{1}{3} = \left(\frac{x-c}{3}\right)^2 = \left[\left(\frac{x-c}{3}\right)^3\right]^{2/3} = y^{2/3},$$

con $c = a$ o $c = b$ según el tramo (y la igualdad final vale también con $x - c < 0$, porque el cubo negativo elevado a $2/3$ devuelve el cuadrado, que es positivo). Como a y b se eligen libremente, hay un continuo biparamétrico de soluciones distintas del mismo problema de valor inicial.

Qué significa esto. La solución puede quedarse pegada al eje el tiempo que quiera y "despegar" cuando le apetezca: el pasado no determina el futuro. No es una patología artificial: el vaciado de un depósito, $h' = -k\sqrt{h}$, tiene exactamente el mismo tipo de raíz no lipschitziana en $h = 0$, y por eso mirando un depósito vacío es imposible saber cuándo se vació. La hipótesis de Lipschitz es justo lo que impide esto.

12.2.10. Campo de direcciones e isoclinas: ver la ecuación antes de resolverla

Definición 12.29 (Campo de direcciones)**.** Dada la ecuación $y' = f(x, y)$ con f definida en una región D del plano, el *campo de direcciones* (o campo de pendientes) asociado es la asignación, a cada punto $(x_0, y_0) \in D$, del segmento corto centrado en ese punto y de pendiente $f(x_0, y_0)$.

Una solución $y = \varphi(x)$ de la ecuación es entonces, exactamente, una curva del plano que en cada uno de sus puntos es tangente al segmento que allí marca el campo.

Esta lectura es la más honesta que existe de una ecuación de primer orden: la ecuación no da la función, da la *dirección* en cada punto, y resolverla es dejarse llevar. La condición inicial dice por qué punto se entra al campo; a partir de ahí la curva no tiene elección. Y de paso explica la figura 12.1 y por qué en ella las soluciones no se cruzan: cruzarse querría decir que por un punto pasan dos curvas con la misma pendiente, y eso es precisamente lo que Picard–Lindelöf prohíbe.

Dibujar un segmento en cada punto es imposible a mano. El truco que lo hace factible es agrupar los puntos donde la pendiente es la misma.

La definición de isoclina está unas páginas más adelante ("la isoclina de pendiente m es la curva $f(x, y) = m$"); lo que se añade aquí es el procedimiento manual completo.

Regla: dibujar a mano el campo de direcciones de $y' = f(x, y)$

1. **Elige unos pocos valores de m.** Bastan cinco o seis, por ejemplo $m = -2, -1, 0, 1, 2$. No hacen falta más para el retrato cualitativo.
2. **Para cada m, dibuja la isoclina $f(x, y) = m$.** Es una curva del plano, y en los casos elementales suele ser una recta o una cónica, que se dibuja de memoria.
3. **Sobre cada isoclina, traza segmentitos paralelos de pendiente m.** Todos iguales, todos con la misma inclinación: por eso agrupar por isoclinas ahorra el trabajo.
4. **Empieza siempre por la isoclina $m = 0$.** Marca dónde las soluciones tienen tangente horizontal, es decir, sus máximos y mínimos, y separa la región donde crecen de aquella donde decrecen. Es la línea más informativa del dibujo.
5. **Comprueba si alguna isoclina es a la vez solución.** Si la isoclina $f(x, y) = m$ es una recta de pendiente exactamente m, entonces esa recta es solución de la ecuación: su pendiente coincide en cada punto con la que marca el campo. Estas rectas actúan como separadores y organizan todo el dibujo.
6. **Traza las curvas solución siguiendo las flechas.** Se entra por un punto y se avanza siempre tangente a los segmentos, sin cruzar nunca otra curva solución.

Ejemplo 12.30 (El campo de direcciones de $y' = x - y$, a mano)**.** *Paso 1–2.* Las isoclinas son $x - y = m$, es decir, las rectas $y = x - m$: todas paralelas a la bisectriz, una por cada m.

- $m = 0$: la recta $y = x$. Sobre ella los segmentos son horizontales, y allí están los mínimos de las soluciones.
- $m = 1$: la recta $y = x - 1$. Sobre ella los segmentos tienen pendiente 1.
- $m = 2$: la recta $y = x - 2$, con segmentos de pendiente 2.
- $m = -1$: la recta $y = x + 1$, con segmentos de pendiente -1.

Paso 5. La isoclina $m = 1$ es la recta $y = x - 1$, que tiene pendiente 1: coincide con m. Luego $y = x - 1$ *es* una solución. Comprobación: si $y = x - 1$ entonces $y' = 1$ y $x - y = x - (x - 1) = 1$. Correcto.

Paso 6. Con eso el dibujo se organiza solo: por encima de $y = x$ el campo apunta hacia abajo ($x - y < 0$), por debajo apunta hacia arriba, y todas las curvas se van pegando a la recta $y = x - 1$ conforme x crece. Resolviendo (es lineal, con $\mu = e^x$) se confirma:

$$y = x - 1 + C\,e^{-x},$$

y el término Ce^{-x} se desvanece al crecer x, de modo que *todas* las soluciones se aproximan a la recta $y = x - 1$, que es la que corresponde a $C = 0$. El dibujo lo había anticipado sin integrar nada. Con la condición inicial $y(0) = 2$ sale $C = 3$ y la solución concreta $y = x - 1 + 3e^{-x}$; comprobación: $y' = 1 - 3e^{-x}$ y $x - y = x - (x - 1 + 3e^{-x}) = 1 - 3e^{-x}$. Coinciden.

Idea clave

Lo que hay que llevarse de esta sección. Una EDO es una relación entre una función desconocida y sus derivadas; resolverla es encontrar funciones, no números; hay familias enteras de soluciones y hace falta un dato extra para elegir una; ese dato puede darse en un punto (valor inicial, y entonces la solución suele ser única) o repartido entre dos (contorno, y entonces puede haber una, ninguna o infinitas). Antes de calcular hay que clasificar: orden, grado, lineal o no, homogénea o completa, coeficientes constantes o variables, y luego la casilla concreta. Y antes de todo eso, si se quiere entender lo que va a salir, se dibuja el campo de direcciones.

12.3. Ecuaciones lineales de segundo orden

Idea clave

Intuición primero. Estas son las ecuaciones de todo lo que oscila: un muelle, un péndulo pequeño, un circuito eléctrico, una molécula vibrando, un puente balanceándose. La estructura de su solución es siempre la misma y merece la pena tenerla clara antes de cualquier cálculo: *solución general = solución de la homogénea + una solución particular.* La homogénea describe cómo se comporta el sistema por su cuenta, sin que nadie lo empuje; la particular describe su respuesta al empuje externo. En un muelle real, la homogénea es cómo vibra al soltarlo y se apaga con el rozamiento; la particular es cómo se acopla a la fuerza que lo agita, y es lo que queda al cabo de un rato.
La ecuación característica, que convierte el problema diferencial en una ecuación de segundo grado del álgebra elemental, funciona porque las exponenciales son las funciones que al derivarse se reproducen. Y sus tres casos —dos raíces reales, una doble, dos complejas— son exactamente los tres comportamientos físicos que se observan: se apaga sin oscilar, se apaga justo en el límite, u oscila apagándose. El discriminante decide cuál, y por eso los amortiguadores de un coche se diseñan afinando ese número.

La forma general es $y'' + p(t)y' + q(t)y = f(t)$; *homogénea* si $f = 0$.

Teorema 12.31 (Estructura)**.** *Las soluciones de la homogénea forman un espacio vectorial de dimensión 2. Dos soluciones y_1, y_2 son base (*sistema fundamental*) si y solo si su* wronskiano

$$W(t) = \det \begin{pmatrix} y_1 & y_2 \\ y_1' & y_2' \end{pmatrix} = y_1 y_2' - y_2 y_1'$$

no se anula. (Por la fórmula de Abel, $W(t) = W(t_0)e^{-\int p}$: o nunca se anula o es idénticamente cero.) La solución general de la no homogénea es $y = C_1 y_1 + C_2 y_2 + y_p$.

12.3.1. Coeficientes constantes

Para $ay'' + by' + cy = 0$, el ansatz $y = e^{rt}$ da la *ecuación característica* $ar^2 + br + c = 0$:

raíces reales $r_1 \neq r_2$	$y = C_1 e^{r_1 t} + C_2 e^{r_2 t}$
raíz doble r	$y = (C_1 + C_2 t)\, e^{rt}$
complejas $\alpha \pm i\beta$	$y = e^{\alpha t}(C_1 \cos \beta t + C_2 \sin \beta t)$

El caso complejo usa la **fórmula de Euler** $e^{i\theta} = \cos\theta + i \sin\theta$; el de raíz doble se obtiene por *reducción de orden*: si y_1 es solución, se busca $y_2 = v(t)y_1$ y v satisface una ecuación de primer orden (aquí, $v = t$).

Ejemplo 12.32 (El oscilador armónico, la EDO más importante de la física)**.** $my'' + cy' + ky = F(t)$: masa, amortiguador, resorte, fuerza externa. *Sin fricción ni fuerza* ($c = F = 0$): $y = A\cos\omega_0 t + B\sin\omega_0 t$, con frecuencia natural $\omega_0 = \sqrt{k/m}$. *Con fricción*: raíces $r = \frac{-c \pm \sqrt{c^2 - 4mk}}{2m}$. Si $c^2 < 4mk$ (*subamortiguado*), oscila con amplitud $e^{-ct/2m}$; si $c^2 > 4mk$ (*sobreamortiguado*), decae sin oscilar; el caso crítico $c^2 = 4mk$ es la vuelta al equilibrio más rápida sin oscilación (así se calibran amortiguadores y galvanómetros). *Forzado periódico* $F = F_0 \cos\omega t$ sin fricción: si $\omega \neq \omega_0$, la respuesta tiene amplitud $\frac{F_0/m}{\omega_0^2 - \omega^2}$, que crece al acercarse las frecuencias; si $\omega = \omega_0$, hay **resonancia**: la solución particular es $\frac{F_0}{2m\omega_0}\, t \sin\omega_0 t$, de amplitud creciente sin cota (puentes que colapsan, copas que estallan).

12.3.2. Búsqueda de soluciones particulares

Coeficientes indeterminados (para f = polinomio, exponencial, seno/coseno o productos): se propone y_p del mismo tipo con coeficientes libres; *si la propuesta resuelve la homogénea, se multiplica por* t (y por t^2 si aún la resuelve). Ejemplos de propuesta: $f = t^2 \Rightarrow y_p = At^2 + Bt + C$; $f = e^{3t} \Rightarrow Ae^{3t}$; $f = e^t \cos 2t \Rightarrow e^t(A\cos 2t + B\sin 2t)$.

Variación de parámetros (general, para cualquier f continua): con sistema fundamental y_1, y_2 de la homogénea y wronskiano W,

$$y_p = -y_1 \int \frac{y_2 f}{W}\, \mathrm{d}t + y_2 \int \frac{y_1 f}{W}\, \mathrm{d}t$$

(para la ecuación normalizada $y'' + py' + qy = f$).

Regla: hallar una solución particular por coeficientes indeterminados

Para $y'' + by' + cy = f(t)$ con b, c constantes y f de tipo "bueno" (polinomio, exponencial, seno/coseno o producto de esos tres), la particular se *adivina con plantilla*. El procedimiento completo es:

1. **Resolver primero la homogénea.** Se calculan las raíces r_1, r_2 de $r^2 + br + c = 0$. No es un paso opcional que se pueda dejar para el final: la lista de raíces es exactamente lo que decide si la plantilla se puede usar tal cual o hay que corregirla.
2. **Escribir la plantilla según f.**
   - f = polinomio de grado $n \Rightarrow y_p = A_n t^n + \cdots + A_1 t + A_0$ (*todos* los grados intermedios, aunque f no los tenga).
   - $f = Ke^{\alpha t} \Rightarrow y_p = Ae^{\alpha t}$.

- $f = K\cos\beta t$ o $K\sin\beta t \Rightarrow y_p = A\cos\beta t + B\sin\beta t$ (siempre los dos, aunque f solo traiga uno).
- Producto de los anteriores $\Rightarrow$ producto de las plantillas.

3. **Comprobar resonancia.** Se mira el "exponente escondido" de la plantilla: α para $e^{\alpha t}$, 0 para un polinomio, $\pm i\beta$ para senos y cosenos. Si ese número es raíz del polinomio característico, la plantilla ya resuelve la homogénea y daría $0 = f$: hay que multiplicarla por t. Si es raíz doble, por t^2. Esta es *la* regla que más se olvida.
4. **Derivar y sustituir.** Se calculan y_p', y_p'', se meten en la ecuación y se agrupa.
5. **Igualar coeficientes.** Los coeficientes de cada potencia de t, de cada exponencial y de cada cos/sin se igualan por separado: funciones distintas son linealmente independientes, así que la identidad debe cumplirse término a término. Sale un sistema lineal pequeño.
6. **Superposición si f es una suma.** Si $f = f_1 + f_2$ de tipos distintos, se resuelve una particular para cada uno y se suman. No hay interferencia: el operador es lineal.
7. **Escribir la general y solo entonces imponer las condiciones iniciales.** $y = y_h + y_p$. Aplicar $y(0)$ e $y'(0)$ a y_h sola es el error clásico que arruina todo el ejercicio.

Ejemplo 12.33 (Cuatro forzamientos, con y sin resonancia)**. (a) Sin resonancia, polinomio.** $y'' - 3y' + 2y = 4t^2$. Raíces: $r^2 - 3r + 2 = 0$ da $r = 1, 2$. El exponente escondido de un polinomio es 0, que no es raíz: plantilla limpia $y_p = At^2 + Bt + C$. Entonces $y_p' = 2At + B$, $y_p'' = 2A$, y

$$2A - 3(2At + B) + 2(At^2 + Bt + C) = 2At^2 + (2B - 6A)t + (2A - 3B + 2C).$$

Igualando a $4t^2 + 0t + 0$: $2A = 4$, $2B - 6A = 0$, $2A - 3B + 2C = 0$, es decir $A = 2$, $B = 6$, $C = 7$. Luego

$$y = C_1 e^t + C_2 e^{2t} + 2t^2 + 6t + 7.$$

Control rápido: en $t = 0$ la particular vale 7 y $y_p'' - 3y_p' + 2y_p = 4 - 18 + 14 = 0$, que es $4t^2$ en $t = 0$. Correcto.

(b) Resonancia simple. $y'' - 3y' + 2y = e^t$. Mismas raíces $1, 2$, y ahora $\alpha = 1$ *sí* es raíz: la plantilla Ae^t moriría. Se multiplica por t: $y_p = Ate^t$. Con $y_p' = A(1 + t)e^t$ y $y_p'' = A(2 + t)e^t$,

$$A(2 + t) - 3A(1 + t) + 2At = A(2 + t - 3 - 3t + 2t) = -A,$$

así que $-A = 1$ y $y_p = -te^t$.

(c) Resonancia doble. $y'' - 2y' + y = e^t$. Ahora $r = 1$ es raíz *doble*, así que hay que subir hasta t^2: $y_p = At^2e^t$. Sustituyendo sale $2A = 1$, es decir

$$y = (C_1 + C_2 t)e^t + \tfrac{1}{2}t^2 e^t.$$

(d) Resonancia trigonométrica. $y'' + 4y = \cos 2t$. Las raíces son $\pm 2i$ y el exponente escondido de $\cos 2t$ es justamente $\pm 2i$: resonancia. Plantilla $y_p = t(A\cos 2t + B\sin 2t)$. Derivando

dos veces y sumando $4y_p$, todos los términos con t se cancelan y queda $-4A\sin 2t+4B\cos 2t=\cos 2t$, luego $A=0$, $B=\frac{1}{4}$:

$$y_p=\frac{t}{4}\sin 2t.$$

Esta y_p crece sin cota: es la *resonancia* física, empujar un columpio exactamente a su frecuencia propia. Sin amortiguamiento, la amplitud es una recta que sube para siempre.

Ejemplo 12.34 (Un forzamiento oscilante amortiguado)**.** $y''+2y'+5y=10\cos t$. Raíces: $r=\frac{-2\pm\sqrt{4-20}}{2}=-1\pm 2i$, así que $y_h=e^{-t}(C_1\cos 2t+C_2\sin 2t)$. El exponente escondido del forzamiento es $\pm i$, que *no* coincide con $-1\pm 2i$: sin resonancia. Plantilla $y_p=A\cos t+B\sin t$; sustituyendo,

$$(-A\cos t-B\sin t)+2(-A\sin t+B\cos t)+5(A\cos t+B\sin t)=(4A+2B)\cos t+(4B-2A)\sin t.$$

Igualando a $10\cos t+0\sin t$: $4A+2B=10$ y $4B-2A=0$, luego $A=2$, $B=1$ y

$$y=e^{-t}(C_1\cos 2t+C_2\sin 2t)+2\cos t+\sin t.$$

Idea clave

Aquí se ve el reparto de papeles que gobierna toda la ingeniería de vibraciones: la parte homogénea lleva el factor e^{-t} y por tanto *muere* (es el *transitorio*, y depende de las condiciones iniciales), mientras que $2\cos t+\sin t$ sobrevive para siempre (es el *régimen permanente*, y no depende de las condiciones iniciales en absoluto). Al cabo de unos pocos segundos el sistema ha olvidado cómo empezó y solo recuerda quién lo está empujando. Su amplitud es $\sqrt{2^2+1^2}=\sqrt{5}\approx 2{,}236$, frente a la amplitud 10 del forzamiento: el sistema responde con un factor de $\sqrt{5}/10\approx 0{,}224$.

Cuidado

La regla de multiplicar por t se aplica *solo* al bloque de la plantilla que resuena, no a toda la particular. En $y''-y=e^t+t$ hay que tratar los dos trozos por separado: para e^t hay resonancia (raíz $r=1$) y toca Ate^t; para t el exponente escondido es 0, que no es raíz, así que basta $Bt+C$. La respuesta es $y_p=\frac{1}{2}te^t-t$, y multiplicar el $-t$ también por t habría dado un sistema incompatible.

12.3.3. Euler–Cauchy

$t^2y''+aty'+by=0$ (coeficientes variables pero con estructura): el ansatz $y=t^m$ da $m(m-1)+am+b=0$. Raíces reales distintas: $y=C_1t^{m_1}+C_2t^{m_2}$; doble: $(C_1+C_2\ln t)t^m$; complejas $\alpha\pm i\beta$: $t^\alpha(C_1\cos(\beta\ln t)+C_2\sin(\beta\ln t))$. (Equivale a coeficientes constantes tras el cambio $t=e^s$.) Esta ecuación reaparece al separar variables en el laplaciano polar.

12.4. Sistemas lineales: el álgebra lineal en acción

Toda EDO de orden n se convierte en un sistema de primer orden con el vector $\mathbf{x}=(y,y',\dots,y^{(n-1)})$; por eso la teoría de sistemas es la teoría general.

12.4.1. El caso diagonalizable

Para $\mathbf{x}' = A\mathbf{x}$ con A constante: si $Av = \lambda v$, entonces $e^{\lambda t}v$ es solución. Con base de vectores propios (λ_i, v_i):

$$\mathbf{x}(t) = C_1 e^{\lambda_1 t} v_1 + \cdots + C_n e^{\lambda_n t} v_n.$$

Valores propios complejos $\lambda = \alpha \pm i\beta$ con vector $v = \mathbf{a} \pm i\mathbf{b}$: se toman las partes real e imaginaria de $e^{\lambda t}v$, que dan las dos soluciones reales

$$e^{\alpha t}(\mathbf{a}\cos\beta t - \mathbf{b}\sin\beta t), \qquad e^{\alpha t}(\mathbf{a}\sin\beta t + \mathbf{b}\cos\beta t)\,:$$

espirales (si $\alpha \neq 0$) o elipses (si $\alpha = 0$) en el plano de $\mathbf{a}, \mathbf{b}$.

Valores propios repetidos sin base propia (bloques de Jordan): aparecen soluciones con factor polinomial; para un bloque 2×2 con vector propio v y *vector propio generalizado* w (con $(A - \lambda I)w = v$):

$$\mathbf{x}_1 = e^{\lambda t}v, \qquad \mathbf{x}_2 = e^{\lambda t}(w + tv).$$

12.4.2. La exponencial de matriz

Definición 12.35. $e^{At} = \sum_{k\geq 0} \frac{(At)^k}{k!}$ (converge siempre). La solución de $\mathbf{x}' = A\mathbf{x}$, $\mathbf{x}(0) = \mathbf{x}_0$ es exactamente $\mathbf{x}(t) = e^{At}\mathbf{x}_0$.

Idea clave

Qué dice esto en cristiano. La exponencial de una matriz es la misma serie de siempre, $\sum \frac{(At)^k}{k!}$, con matrices en lugar de números, y resuelve los sistemas lineales exactamente igual que e^{at} resuelve $x' = ax$. Converge siempre, sin condiciones. Lo único que hay que calcular es la serie, y para eso se usa la forma de la matriz: si diagonaliza, la exponencial es diagonal; si tiene un bloque de Jordan, aparecen factores t; y si tiene autovalores complejos, aparecen senos y cosenos.

Ejemplo 12.36 (Calcular e^{At} según el tipo de matriz)**. (a) Diagonal.** Para $A = \begin{pmatrix} 1 & 0 \\ 0 & 3 \end{pmatrix}$ sale $\mathrm{e}^{At} = \begin{pmatrix} \mathrm{e}^{t} & 0 \\ 0 & \mathrm{e}^{3t} \end{pmatrix}$: cada autovalor va a su exponencial y no se mezclan.

(b) Rotación. Para $A = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$, cuyos autovalores son $\pm$i, sale $\mathrm{e}^{At} = \begin{pmatrix} \cos t & \sin t \\ -\sin t & \cos t \end{pmatrix}$. Los autovalores imaginarios producen oscilación pura, sin crecimiento ni decaimiento; es la fórmula de Euler en versión matricial.

(c) Bloque de Jordan. Para $B = \begin{pmatrix} 2 & 1 \\ 0 & 2 \end{pmatrix}$ sale $\mathrm{e}^{Bt} = \begin{pmatrix} \mathrm{e}^{2t} & t\,\mathrm{e}^{2t} \\ 0 & \mathrm{e}^{2t} \end{pmatrix}$. El factor t que aparece es exactamente el que en las ecuaciones de segundo orden con raíz doble obliga a escribir la segunda solución como $t\,\mathrm{e}^{\lambda t}$.

(d) El wronskiano por Abel–Liouville. Se cumple $W' = \operatorname{tr}(A)\,W$, luego $W(t) = W(t_0)\mathrm{e}^{\int \operatorname{tr} A}$. Para la rotación de (b), $\operatorname{tr} A = 0$ y el wronskiano es constante: el flujo conserva

áreas. Para la diagonal de (a), $\operatorname{tr} A = 4$ y $W(t) = W(0)\mathrm{e}^{4t}$: las áreas crecen. Nótese la consecuencia útil: si W no es cero en un punto, no lo es en ninguno.

Cuidado

[La regla $\mathrm{e}^{A+B} = \mathrm{e}^{A}\mathrm{e}^{B}$ es falsa] Solo vale si $AB = BA$. Con $P = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$ y $Q = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}$ se obtiene $\mathrm{e}^{P}\mathrm{e}^{Q} = \begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix}$, mientras que $\mathrm{e}^{P+Q} = \begin{pmatrix} \cosh 1 & \sinh 1 \\ \sinh 1 & \cosh 1 \end{pmatrix}$, es decir $\begin{pmatrix} 1{,}5431 & 1{,}1752 \\ 1{,}1752 & 1{,}5431 \end{pmatrix}$. Son matrices distintas. Segundo aviso: e^{At} nunca es singular, porque su inversa es e^{-At}; por eso las soluciones de un sistema lineal homogéneo no pueden anularse en un punto sin ser idénticamente nulas.

Cálculo: si $A = PDP^{-1}$, $e^{At} = Pe^{Dt}P^{-1}$ con $e^{Dt} = \operatorname{diag}(e^{\lambda_i t})$; para bloques de Jordan, $e^{(D+N)t} = e^{Dt}e^{Nt}$ y la serie de N es finita. Para el no homogéneo $\mathbf{x}' = A\mathbf{x} + \mathbf{f}(t)$, la *fórmula de variación de constantes*:

$$\mathbf{x}(t) = e^{At}\mathbf{x}_0 + \int_0^t e^{A(t-s)}\mathbf{f}(s)\,\mathrm{d}s.$$

12.4.3. Retratos de fase y estabilidad

Idea clave

Para $\mathbf{x}' = A\mathbf{x}$ en el plano, el par de valores propios decide la geometría cerca del origen: reales negativos: *nodo estable* (todo cae al origen); reales positivos: nodo inestable; signos opuestos: *silla* (entra por una dirección propia, escapa por la otra); complejos con $\alpha < 0$: *espiral estable*; $\alpha > 0$: espiral inestable; imaginarios puros: *centro* (órbitas cerradas). Regla resumen con $\tau = \operatorname{tr} A$, $\Delta = \det A$: silla si $\Delta < 0$; estable si $\Delta > 0, \tau < 0$; espirales si $\tau^2 < 4\Delta$.

Para sistemas *no lineales* $\mathbf{x}' = \mathbf{F}(\mathbf{x})$: en cada *equilibrio* ($\mathbf{F}(\mathbf{x}^*) = 0$) se *linealiza* con la jacobiana $D\mathbf{F}(\mathbf{x}^*)$, y (teorema de Hartman–Grobman) si ningún valor propio tiene parte real nula, el retrato local es el del sistema lineal. Así se analizan los modelos depredador–presa, péndulos y epidemias.

Regla: resolver un sistema lineal y leer su retrato de fase

Para $\mathbf{x}' = A\mathbf{x}$ con A constante, la solución y su forma geométrica se leen ambas de los autovalores.

1. **Calcula los autovalores.** De $\det(A - \lambda I) = 0$. Para 2×2 hay un atajo: $\lambda^2 - (\operatorname{tr} A)\lambda + \det A = 0$, así que basta con la traza y el determinante.

2. **Caso autovalores reales y distintos.** Si $A\mathbf{v}_i = \lambda_i \mathbf{v}_i$, la solución general es
$$\mathbf{x}(t) = c_1 e^{\lambda_1 t}\mathbf{v}_1 + c_2 e^{\lambda_2 t}\mathbf{v}_2.$$
Las rectas generadas por los $\mathbf{v}_i$ son invariantes: si empiezas sobre una, no la abandonas nunca.

3. **Caso autovalores complejos $\alpha \pm i\beta$.** Toma un autovector complejo $\mathbf{v} = \mathbf{p} + i\mathbf{q}$ y separa parte real e imaginaria:
$$\mathbf{x}(t) = e^{\alpha t}\big[c_1(\mathbf{p}\cos\beta t - \mathbf{q}\sin\beta t) + c_2(\mathbf{p}\sin\beta t + \mathbf{q}\cos\beta t)\big].$$
El α controla si la espiral crece o decrece; el β, la velocidad de giro.

4. **Caso defectuoso.** Si λ es doble con un solo autovector $\mathbf{v}$, busca $\mathbf{w}$ con $(A-\lambda I)\mathbf{w} = \mathbf{v}$ y escribe
$$\mathbf{x}(t) = c_1 e^{\lambda t}\mathbf{v} + c_2 e^{\lambda t}(t\mathbf{v} + \mathbf{w}).$$
El factor t es la marca del bloque de Jordan.

5. **Clasifica con traza y determinante.** Si $\det A < 0$: *punto de silla*, siempre inestable. Si $\det A > 0$ y $\operatorname{tr} A < 0$: sumidero estable; si $\operatorname{tr} A > 0$: fuente inestable. El discriminante $(\operatorname{tr} A)^2 - 4\det A$ decide entre nodo (positivo) y espiral (negativo); si vale cero, nodo impropio.

6. **Dibuja el retrato.** Marca las rectas invariantes, orienta las flechas según el signo de λ (hacia el origen si $\lambda < 0$), y en el caso complejo determina el sentido de giro evaluando el campo en un punto concreto, por ejemplo $(1, 0)$.

7. **Verifica derivando.** Sustituye la solución en el sistema. Control adicional: la exponencial de matriz debe cumplir $e^{A\cdot 0} = I$ y $\frac{\mathrm{d}}{\mathrm{d}t}e^{At} = Ae^{At}$.

Ejemplo 12.37 (Cuatro sistemas y sus cuatro geometrías)**. (a) Punto de silla.** Sea $A = \left(\begin{smallmatrix} 1 & 2 \\ 3 & 2 \end{smallmatrix}\right)$. Aquí $\operatorname{tr} A = 3$ y $\det A = 2 - 6 = -4 < 0$, luego es una silla sin más cálculo. Los autovalores salen de $\lambda^2 - 3\lambda - 4 = 0$: son $\lambda = 4$ y $\lambda = -1$. Los autovectores son $\mathbf{v}_1 = (2, 3)$ para $\lambda = 4$ y $\mathbf{v}_2 = (-1, 1)$ para $\lambda = -1$. Solución general:
$$\mathbf{x}(t) = c_1 e^{4t}\begin{pmatrix} 2 \\ 3 \end{pmatrix} + c_2 e^{-t}\begin{pmatrix} -1 \\ 1 \end{pmatrix}.$$
La recta generada por $(-1, 1)$ es la *variedad estable*: sobre ella las trayectorias caen al origen. La generada por $(2, 3)$ es la inestable. Cualquier otra condición inicial acaba escapando en la dirección de $(2, 3)$, porque e^{4t} domina a e^{-t}.

(b) Centro: giro puro. Sea $A = \left(\begin{smallmatrix} 0 & -1 \\ 1 & 0 \end{smallmatrix}\right)$, con $\operatorname{tr} A = 0$ y $\det A = 1$. Autovalores $\pm i$: parte real nula. La exponencial es exactamente la matriz de rotación
$$e^{At} = \begin{pmatrix} \cos t & -\sin t \\ \sin t & \cos t \end{pmatrix},$$

así que las órbitas son circunferencias recorridas en sentido antihorario, con periodo 2π. Es el oscilador armónico escrito como sistema. Nótese que el centro es un caso frágil: cualquier perturbación de la traza lo convierte en espiral.

(c) Foco estable. Sea $A = \left(\begin{smallmatrix} -1 & -2 \\ 2 & -1 \end{smallmatrix}\right)$, con $\operatorname{tr} A = -2 < 0$ y $\det A = 1 + 4 = 5 > 0$, y discriminante $4 - 20 = -16 < 0$: espiral estable. Los autovalores son $-1 \pm 2i$, y la exponencial resulta

$$e^{At} = e^{-t} \begin{pmatrix} \cos 2t & -\sin 2t \\ \sin 2t & \cos 2t \end{pmatrix} :$$

una rotación de velocidad angular 2 multiplicada por una contracción e^{-t}. Las trayectorias son espirales que se cierran sobre el origen, dando una vuelta completa cada π unidades de tiempo mientras su radio se reduce por el factor $e^{-\pi} \approx 0{,}0432$.

(d) Nodo impropio (caso defectuoso). Sea $A = \left(\begin{smallmatrix} 3 & 1 \\ 0 & 3 \end{smallmatrix}\right)$. El único autovalor es 3 doble, y $A - 3I = \left(\begin{smallmatrix} 0 & 1 \\ 0 & 0 \end{smallmatrix}\right)$ tiene rango 1, luego solo hay un autovector, $\mathbf{v} = (1, 0)$. Buscamos $\mathbf{w}$ con $(A - 3I)\mathbf{w} = \mathbf{v}$: sirve $\mathbf{w} = (0, 1)$. Entonces

$$e^{At} = e^{3t} \begin{pmatrix} 1 & t \\ 0 & 1 \end{pmatrix}, \qquad \mathbf{x}(t) = c_1 e^{3t} \begin{pmatrix} 1 \\ 0 \end{pmatrix} + c_2 e^{3t} \left[t \begin{pmatrix} 1 \\ 0 \end{pmatrix} + \begin{pmatrix} 0 \\ 1 \end{pmatrix} \right].$$

Todas las trayectorias escapan al infinito (porque $\lambda = 3 > 0$) y todas lo hacen tangentes a la única dirección propia $(1, 0)$: el factor t hace que la primera componente crezca más deprisa. Es el nodo impropio, la geometría intermedia entre nodo y espiral.

Idea clave

Todo el retrato de fase cabe en un plano de dos ejes: la traza en horizontal y el determinante en vertical. La parábola $(\operatorname{tr} A)^2 = 4 \det A$ separa nodos de espirales; el eje $\det A = 0$ separa sillas de lo demás; el eje $\operatorname{tr} A = 0$ separa lo estable de lo inestable. Un sistema es un punto en ese plano, y su comportamiento cualitativo es la región donde cae. Cambiar un parámetro del modelo es mover ese punto, y las *bifurcaciones* ocurren justo al cruzar una de las fronteras.

Cuidado

El caso del centro, $\operatorname{tr} A = 0$ y $\det A > 0$, no es estructuralmente estable: no basta con la parte lineal para concluir nada sobre el sistema no lineal original. El teorema de Hartman–Grobman garantiza que el retrato lineal describe el no lineal solo cuando ningún autovalor tiene parte real nula. En un centro, términos no lineales arbitrariamente pequeños pueden convertirlo en espiral atractora o repulsora.

Regla: variación de parámetros y el wronskiano: cuando los coeficientes indeterminados fallan

El método de coeficientes indeterminados solo sirve si el término independiente es de un tipo muy concreto (polinomio, exponencial, seno, coseno y sus productos). Para $\sec x$, $\ln x$, $1/x$ o cualquier otra cosa, hace falta variación de parámetros, que funciona *siempre*

a cambio de tener que integrar. Estos son los siete pasos.

1. **Normaliza la ecuación.** Escríbela como $y''+p(x)y'+q(x)y=g(x)$, con coeficiente 1 en y''. Si venía como $a(x)y''+\dots$, divide todo entre $a(x)$. *El g que se usa en las fórmulas es el de la ecuación normalizada*, y olvidarlo es el error número uno del método.

2. **Resuelve la homogénea y obtén y_1, y_2.** Si los coeficientes son constantes, por la ecuación característica. Si es de Euler–Cauchy $x^2y''+axy'+by=0$, prueba $y=x^r$ y resuelve $r(r-1)+ar+b=0$. Si solo conoces una solución, obtén la otra por reducción de orden (paso 6).

3. **Calcula el wronskiano.**

$$W(x)=\det\begin{pmatrix} y_1 & y_2 \\ y_1' & y_2' \end{pmatrix}=y_1y_2'-y_2y_1'.$$

   Debe ser distinto de cero en el intervalo de trabajo; si sale cero, tus dos soluciones eran dependientes y hay que rehacer el paso 2.

4. **Escribe las derivadas de los parámetros.** Se busca $y_p=u_1y_1+u_2y_2$ con

$$u_1'=-\frac{y_2\,g}{W},\qquad u_2'=\frac{y_1\,g}{W}.$$

   La asimetría del signo es real y hay que respetarla: el subíndice 1 lleva el menos.

5. **Integra y monta la particular.** Calcula u_1 y u_2 sin añadir constantes (las constantes solo reproducen la solución homogénea) y forma $y_p=u_1y_1+u_2y_2$. Si en y_p aparece un múltiplo de y_1 o de y_2, puedes tirarlo: ya está en la homogénea.

6. **Si te falta una segunda solución homogénea, úsala por reducción de orden.** Conocida y_1, la fórmula es

$$y_2=y_1\int\frac{e^{-\int p\,dx}}{y_1^2}\,dx.$$

   Es la misma idea de variación de parámetros aplicada a la homogénea.

7. **Verifica sustituyendo, y comprueba el wronskiano con la fórmula de Abel.** Debe cumplirse $W(x)=W(x_0)\,e^{-\int_{x_0}^{x}p(t)\,dt}$. Es una comprobación independiente que detecta errores de signo en W sin volver a derivar nada.

Ejemplo 12.38 (Cuatro ecuaciones que los coeficientes indeterminados no saben resolver)**. (a) La secante, el ejemplo canónico.** Resolvamos $y''+y=\sec x$ en $(-\frac{\pi}{2},\frac{\pi}{2})$. La homogénea da $y_1=\cos x$, $y_2=\sin x$, y

$$W=\cos x\cdot\cos x-\sin x\cdot(-\sin x)=\cos^2x+\sin^2x=1.$$

Con $g = \sec x$:

$$u_1' = -\frac{\sin x \cdot \sec x}{1} = -\tan x \implies u_1 = \ln|\cos x|,$$
$$u_2' = \frac{\cos x \cdot \sec x}{1} = 1 \implies u_2 = x.$$

Por tanto

$$y_p = \cos x \, \ln|\cos x| + x \sin x,$$

y la solución general es $y = C_1 \cos x + C_2 \sin x + \cos x \ln|\cos x| + x \sin x$. Ningún método de coeficientes indeterminados puede producir un logaritmo: por eso hacía falta este.

(b) Raíz doble y término e^x/x. Sea $y'' - 2y' + y = \dfrac{e^x}{x}$ en $x > 0$. La característica $r^2 - 2r + 1 = 0$ tiene la raíz doble $r = 1$, así que $y_1 = e^x$, $y_2 = xe^x$. El wronskiano:

$$W = e^x(e^x + xe^x) - xe^x \cdot e^x = e^{2x} + xe^{2x} - xe^{2x} = e^{2x}.$$

Con $g = e^x/x$:

$$u_1' = -\frac{xe^x \cdot e^x/x}{e^{2x}} = -1 \implies u_1 = -x, \qquad u_2' = \frac{e^x \cdot e^x/x}{e^{2x}} = \frac{1}{x} \implies u_2 = \ln x.$$

Entonces $y_p = -xe^x + xe^x \ln x$. El primer sumando es un múltiplo de y_2, así que lo absorbemos en la homogénea (paso 5) y queda

$$y = (C_1 + C_2 x)e^x + xe^x \ln x.$$

(c) Una de Euler–Cauchy. Resolvamos $x^2y'' - 2xy' + 2y = x^3 \ln x$ en $x > 0$. Homogénea: con $y = x^r$, $r(r-1) - 2r + 2 = r^2 - 3r + 2 = 0$, luego $r = 1, 2$ y $y_1 = x$, $y_2 = x^2$. El wronskiano es

$$W = x \cdot 2x - x^2 \cdot 1 = x^2.$$

Normalizamos (paso 1) dividiendo entre x^2: la ecuación es $y'' - \frac{2}{x}y' + \frac{2}{x^2}y = x \ln x$, de modo que $g = x \ln x$ y no $x^3 \ln x$. Ahora

$$u_1' = -\frac{x^2 \cdot x \ln x}{x^2} = -x \ln x \implies u_1 = -\frac{x^2}{2} \ln x + \frac{x^2}{4},$$

$$u_2' = \frac{x \cdot x \ln x}{x^2} = \ln x \implies u_2 = x \ln x - x.$$

Montamos:

$$y_p = x\left(-\frac{x^2}{2} \ln x + \frac{x^2}{4}\right) + x^2(x \ln x - x) = \frac{x^3}{2} \ln x - \frac{3x^3}{4}.$$

Verificación directa: con $y_p = \frac{x^3}{2} \ln x - \frac{3}{4}x^3$ se obtiene, tras derivar dos veces y sustituir, exactamente $x^3 \ln x$. La solución general es $y = C_1 x + C_2 x^2 + \frac{x^3}{2} \ln x - \frac{3}{4}x^3$. Y si alguien prefiere, el término $-\frac{3}{4}x^3$ no es homogéneo y no se puede tirar.

(d) Reducción de orden. Sea $x^2y'' + xy' - y = 0$ con la solución evidente $y_1 = x$. Normalizando, $p = \frac{1}{x}$, luego $e^{-\int p\,dx} = e^{-\ln x} = \frac{1}{x}$ y

$$y_2 = x \int \frac{1/x}{x^2}\,dx = x \int x^{-3}\,dx = x \cdot \left(-\frac{1}{2x^2}\right) = -\frac{1}{2x}.$$

Como las constantes multiplicativas dan igual, tomamos $y_2 = \frac{1}{x}$. Comprobación: $y_2'' = \frac{2}{x^3}$, $y_2' = -\frac{1}{x^2}$, y

$$x^2 \cdot \frac{2}{x^3} + x \cdot \left(-\frac{1}{x^2}\right) - \frac{1}{x} = \frac{2}{x} - \frac{1}{x} - \frac{1}{x} = 0. \quad \checkmark$$

El wronskiano es $W = x \cdot (-x^{-2}) - x^{-1} \cdot 1 = -\frac{2}{x}$, no nulo en $x > 0$, y coincide con la fórmula de Abel: $W(x) = W(1)e^{-\int_1^x dt/t} = -2 \cdot \frac{1}{x}$.

Cuidado

Dos trampas del wronskiano. Primera: $W \equiv 0$ *no* implica dependencia lineal para funciones cualesquiera. El par $y_1 = x^2$, $y_2 = x|x|$ en $[-1, 1]$ tiene wronskiano idénticamente nulo y sin embargo es linealmente independiente. La implicación solo vale para soluciones de una misma ecuación lineal con coeficientes continuos, y es entonces cuando la fórmula de Abel la garantiza: si W se anula en un punto, se anula en todo el intervalo. Segunda: no olvides normalizar. En el apartado (c), usar $g = x^3 \ln x$ en vez de $g = x \ln x$ produce una respuesta que difiere en un factor x^2 y no satisface la ecuación.

Idea clave

Variación de parámetros no es un truco: es la fórmula de la función de Green disfrazada. Al escribir u_1 y u_2 como integrales definidas desde x_0 y sustituir, se obtiene

$$y_p(x) = \int_{x_0}^{x} \frac{y_1(t)y_2(x) - y_1(x)y_2(t)}{W(t)}\, g(t)\, dt,$$

es decir, $y_p(x) = \int G(x,t)\, g(t)\, dt$ con un núcleo G que no depende del término independiente. La ecuación diferencial se ha invertido como si fuera una matriz, y G es la "matriz inversa". Esa es exactamente la idea que reaparece en las ecuaciones integrales del capítulo 32 y en toda la física matemática.

12.5. La transformada de Laplace

Idea clave

Intuición primero. Los logaritmos convirtieron multiplicar en sumar; Laplace convierte *derivar en multiplicar por s*. Una ecuación diferencial, al otro lado del espejo, es una ecuación *algebraica*: se despeja $Y(s)$ con álgebra de bachillerato y se regresa. ¿Por qué funciona? Porque las exponenciales e^{st} son los vectores propios del operador "derivar", y la transformada escribe todo en esa base propia: es la diagonalización del capítulo 7, aplicada al operador $\frac{\mathrm{d}}{\mathrm{d}t}$. Los polos de $Y(s)$ son las frecuencias naturales del sistema: donde el denominador muere, el sistema resuena.

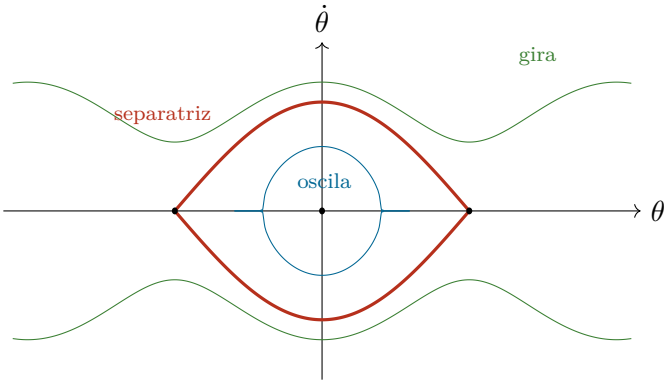


Figura 12.2: El plano de fase del péndulo: órbitas cerradas (oscilación), curvas abiertas (rotación) y la separatriz roja que las divide, pasando por los equilibrios inestables $\theta = \pm\pi$.

Definición 12.39. $\mathcal{L}\{f\}(s) = F(s) = \displaystyle\int_0^\infty e^{-st} f(t)\,\mathrm{d}t$, definida para s grande si f es de orden exponencial. Es lineal e inyectiva (sobre funciones continuas): se puede antitransformar.

Tabla esencial:

$$\mathcal{L}\{1\} = \tfrac{1}{s} \qquad \mathcal{L}\{t^n\} = \tfrac{n!}{s^{n+1}} \qquad \mathcal{L}\{e^{at}\} = \tfrac{1}{s-a} \qquad \mathcal{L}\{t^n e^{at}\} = \tfrac{n!}{(s-a)^{n+1}}$$
$$\mathcal{L}\{\sin bt\} = \tfrac{b}{s^2+b^2} \qquad \mathcal{L}\{\cos bt\} = \tfrac{s}{s^2+b^2} \qquad \mathcal{L}\{e^{at}\sin bt\} = \tfrac{b}{(s-a)^2+b^2} \qquad \mathcal{L}\{e^{at}\cos bt\} = \tfrac{s-a}{(s-a)^2+b^2}$$

Reglas operativas:

$$\mathcal{L}\{f'\} = sF(s) - f(0), \qquad \mathcal{L}\{f''\} = s^2F(s) - sf(0) - f'(0),$$

$$\mathcal{L}\{e^{at}f(t)\} = F(s-a) \text{ (desplazamiento en } s), \qquad \mathcal{L}\{u_a(t)f(t-a)\} = e^{-as}F(s),$$

donde $u_a(t)$ es la *función escalón* (0 antes de a, 1 después): con ella se modelan fuerzas que se encienden y apagan. La *delta de Dirac* $\delta(t-a)$ (impulso instantáneo) tiene $\mathcal{L}\{\delta(t-a)\} = e^{-as}$.

Teorema 12.40 (Convolución)**.** *Con* $(f*g)(t) = \int_0^t f(\tau)g(t-\tau)\,\mathrm{d}\tau$*:* $\mathcal{L}\{f*g\} = F(s)\,G(s)$.

Idea clave

Qué dice esto en cristiano. La transformada de Laplace convierte ecuaciones diferenciales en ecuaciones algebraicas: derivar en t se transforma en multiplicar por s. Se aplica la transformada, se despeja como si fuera álgebra de bachillerato, y se deshace el cambio con una tabla. Y la propiedad de convolución cierra el círculo: multiplicar dos transformadas corresponde a convolucionar las funciones originales, lo que permite invertir productos que de otro modo no estarían en ninguna tabla.

Ejemplo 12.41 (La tabla básica y el uso de la convolución)**. (a) Las transformadas que hay que saber.** $\mathcal{L}\{1\} = \frac{1}{s}$; $\mathcal{L}\{t\} = \frac{1}{s^2}$; $\mathcal{L}\{t^2\} = \frac{2}{s^3}$; $\mathcal{L}\{\mathrm{e}^{3t}\} = \frac{1}{s-3}$; y $\mathcal{L}\{\sin 2t\} = \frac{2}{s^2+4}$. Todas se obtienen integrando directamente $\int_0^\infty \mathrm{e}^{-st} f(t)\,\mathrm{d}t$.

(b) Resolver una ecuación. Para $y' + 2y = 0$ con $y(0) = 3$, transformando queda $sY - 3 + 2Y = 0$, es decir $Y = \frac{3}{s+2}$, y deshaciendo el cambio $y = 3\mathrm{e}^{-2t}$. La condición inicial entra sola en la transformada, sin necesidad de imponerla al final.

(c) La convolución en acción. Como $\mathcal{L}\{t\} = \frac{1}{s^2}$, el producto $\frac{1}{s^2}\cdot\frac{1}{s^2} = \frac{1}{s^4}$ debe corresponder a $t * t$. Calculando: $\int_0^t \tau(t-\tau)\,\mathrm{d}\tau = \frac{t^3}{6}$, y en efecto $\mathcal{L}\left\{\frac{t^3}{6}\right\} = \frac{1}{s^4}$. Coincide.

(d) Para qué sirve de verdad. Con la convolución se resuelven sistemas con entrada arbitraria: si $\mathcal{L}\{h\}$ es la función de transferencia y u la entrada, la salida es $h * u$. Toda la teoría de sistemas lineales y de filtros se apoya en esto, y explica por qué en ingeniería se habla de la respuesta al impulso: convolucionar con ella reconstruye la respuesta a cualquier cosa.

Cuidado

[La transformada no siempre existe, y la convolución no es el producto] Hace falta que f sea de orden exponencial; e^{t^2} no tiene transformada de Laplace, porque la integral diverge para todo s. Segundo aviso, y es el error clásico: $\mathcal{L}\{fg\}$ *no* es $F(s)G(s)$. Lo que vale es $\mathcal{L}\{f * g\} = F\,G$, con la convolución, que es una integral y no un producto puntual. Y tercero: la transformada solo mira $t \geq 0$, así que dos funciones que difieran para $t < 0$ tienen la misma transformada; por eso el método es natural para problemas de valor inicial y no para problemas de contorno.

La convolución expresa la respuesta de un sistema lineal como superposición de respuestas a impulsos: si h es la respuesta al impulso ($\mathcal{L}\{h\} = H$ es la *función de transferencia*), la respuesta a una entrada f es $h * f$. Este punto de vista domina la ingeniería de control y el procesamiento de señales.

Método: transformar la EDO (las condiciones iniciales entran solas), despejar $Y(s)$ algebraicamente, descomponer en fracciones parciales, antitransformar con la tabla.

Regla: resolver un problema de valor inicial con la transformada de Laplace

El método es siempre el mismo triángulo: transformar, despejar con álgebra, antitransformar.

1. **Transformar toda la ecuación de golpe**, usando las dos fórmulas que lo hacen todo:
$$\mathcal{L}\{y'\} = sY(s) - y(0), \qquad \mathcal{L}\{y''\} = s^2Y(s) - sy(0) - y'(0).$$
Aquí está la gran ventaja frente a los métodos anteriores: las condiciones iniciales entran *desde el primer renglón*, no al final. Nunca hay que resolver un sistema para C_1 y C_2.

2. **Despejar $Y(s)$.** Queda una ecuación algebraica lineal en Y. Siempre resulta
$$Y(s) = \frac{(\text{datos iniciales}) + F(s)}{p(s)},$$
donde $p(s) = s^2 + bs + c$ es el mismo polinomio característico de siempre y $F = \mathcal{L}\{f\}$.

3. **Descomponer en fracciones simples.** Es el paso que consume el tiempo. Si p tiene raíces reales distintas, salen sumandos $\frac{A}{s-r}$; si hay raíz doble, $\frac{A}{(s-r)^2}$; si hay raíces complejas $\alpha \pm i\beta$, *no* se separan: se completa el cuadrado y se deja $\frac{A(s-\alpha)+B\beta}{(s-\alpha)^2+\beta^2}$.

4. **Antitransformar con la tabla.** Cada trozo se lee al revés: $\frac{1}{s-r} \mapsto e^{rt}$, $\frac{n!}{(s-r)^{n+1}} \mapsto t^n e^{rt}$, $\frac{\beta}{(s-\alpha)^2+\beta^2} \mapsto e^{\alpha t}\sin\beta t$, $\frac{s-\alpha}{(s-\alpha)^2+\beta^2} \mapsto e^{\alpha t}\cos\beta t$.

5. **Si aparece e^{-as}, es un retraso.** El segundo teorema de traslación dice $\mathcal{L}^{-1}\{e^{-as}F(s)\} = u_a(t)\,f(t-a)$, con u_a el escalón de Heaviside. Regla práctica: se antitransforma F ignorando el e^{-as}, y al final se reemplaza t por $t-a$ y se multiplica por $u_a(t)$.

6. **Si el trozo no está en la tabla, convolución.** $\mathcal{L}^{-1}\{F(s)G(s)\} = (f*g)(t) = \int_0^t f(\tau)g(t-\tau)\,\mathrm{d}\tau$. En particular $\mathcal{L}^{-1}\{F(s)/p(s)\}$ es la convolución del forzamiento con la respuesta al impulso.

7. **Controlar en $t=0$.** La respuesta debe cumplir $y(0)$ e $y'(0)$; es una comprobación de dos segundos que atrapa casi todos los errores de fracciones simples. Otro control gratis: el teorema del valor inicial, $y(0) = \lim_{s\to\infty} sY(s)$.

Cuándo conviene Laplace y cuándo no: es *el* método si el forzamiento es discontinuo, a trozos o impulsivo (un interruptor, un martillazo), porque los otros métodos exigirían resolver por tramos y pegar. Para un forzamiento suave y sencillo, coeficientes indeterminados suele ser más rápido.

Ejemplo 12.42 (Tres problemas de valor inicial resueltos con Laplace)**. (a) El caso básico.** $y''+3y'+2y=0$, con $y(0)=1$, $y'(0)=0$. Transformando:

$$(s^2Y - s\cdot 1 - 0) + 3(sY-1) + 2Y = 0 \implies (s^2+3s+2)Y = s+3.$$

Como $s^2+3s+2=(s+1)(s+2)$,

$$Y = \frac{s+3}{(s+1)(s+2)} = \frac{2}{s+1} - \frac{1}{s+2},$$

donde los residuos se leen tapando: en $s=-1$, $\frac{-1+3}{-1+2}=2$; en $s=-2$, $\frac{-2+3}{-2+1}=-1$. Antitransformando,

$$y = 2e^{-t} - e^{-2t}.$$

Control: $y(0)=2-1=1$ y $y'(0)=-2+2=0$. Correcto.

(b) Resonancia pura, sin adivinar plantillas. $y''+y=\sin t$, con $y(0)=y'(0)=0$. Transformando, $(s^2+1)Y = \frac{1}{s^2+1}$, es decir

$$Y = \frac{1}{(s^2+1)^2}.$$

Esto no está en la tabla básica, pero es un producto: $Y = F(s)\cdot F(s)$ con $F = \frac{1}{s^2+1}$, cuya antitransformada es $\sin t$. Por convolución,

$$y = \int_0^t \sin\tau\,\sin(t-\tau)\,\mathrm{d}\tau = \frac{\sin t - t\cos t}{2},$$

usando $\sin\tau\sin(t-\tau)=\frac{1}{2}[\cos(2\tau-t)-\cos t]$ e integrando cada trozo. Aparece el término $t\cos t$, que crece sin cota: la misma resonancia del ejemplo anterior, pero deducida sin necesidad de saber de antemano que había que multiplicar por t. Laplace *descubre* la resonancia; no hay que anticiparla.

(c) Un interruptor. $y'+y=u_1(t)$ con $y(0)=0$: el sistema está en reposo y en $t=1$ alguien enciende una fuente constante. Como $\mathcal{L}\{u_1\}=\frac{e^{-s}}{s}$,

$$(s+1)Y=\frac{e^{-s}}{s}\implies Y=e^{-s}\cdot\frac{1}{s(s+1)}=e^{-s}\Big(\frac{1}{s}-\frac{1}{s+1}\Big).$$

Ignorando el e^{-s}, la antitransformada de $\frac{1}{s}-\frac{1}{s+1}$ es $1-e^{-t}$. Ahora se aplica el retraso: se cambia t por $t-1$ y se multiplica por el escalón:

$$y(t)=(1-e^{-(t-1)})u_1(t)=\begin{cases}0, & t<1,\\ 1-e^{-(t-1)}, & t\geq 1.\end{cases}$$

La solución es idénticamente cero hasta $t=1$ y luego sube hacia 1. Nótese que es continua en $t=1$ aunque el forzamiento salte: integrar suaviza.

Cuidado

Dos trampas muy frecuentes. La primera: $\mathcal{L}\{y'\}=sY-y(0)$ lleva $y(0)$, no $y(0)$ multiplicado por s; y en $\mathcal{L}\{y''\}=s^2Y-sy(0)-y'(0)$ el que va acompañado de s es $y(0)$, no $y'(0)$. Un intercambio ahí cambia por completo la respuesta y el control en $t=0$ lo delata inmediatamente. La segunda: al aplicar el retraso hay que sustituir $t\mapsto t-a$ en *toda* la función antitransformada, no solo en el primer término. En el ejemplo (c) el resultado es $1-e^{-(t-1)}$, no $1-e^{-t}$.

Pausa para pensar

¿Por qué no aparecen constantes arbitrarias en ningún punto del método de Laplace? Porque la transformada no ve la ecuación diferencial *sola*: ve el problema de valor inicial completo. La familia infinita de soluciones que producían los otros métodos existía precisamente porque no se había dicho aún de dónde arranca el sistema. Laplace exige ese dato en el primer paso, y a cambio devuelve una única función, ya lista.

12.6. Soluciones por series de potencias

Idea clave

Intuición primero. La mayoría de las ecuaciones diferenciales no se resuelven con fórmulas cerradas. La salida no es rendirse: es cambiar de idea sobre qué significa "resolver". Si propones que la solución es una serie de potencias con coeficientes desconocidos, sustituyes en la ecuación e igualas coeficientes, obtienes una relación de recurrencia que los va produciendo uno a uno. Eso es una solución: te permite calcular la función con la precisión que quieras y demostrar propiedades sobre ella.

De hecho, así es como nacieron casi todas las funciones especiales de la física. Las funciones de Bessel, los polinomios de Legendre, los de Hermite no se descubrieron por gusto: son las soluciones en serie de las ecuaciones que salen al separar variables en un tambor, en una esfera o en un oscilador cuántico. No son funciones "raras"; son tan legítimas como el seno, solo que menos familiares porque su ecuación aparece más tarde.

12.6.1. Puntos ordinarios

Si en $y'' + p(x)y' + q(x)y = 0$ los coeficientes son analíticos en x_0 (punto *ordinario*), toda solución es analítica: se sustituye $y = \sum c_n (x - x_0)^n$ y se obtiene una *recurrencia* para los c_n, con c_0, c_1 libres.

Ejemplo 12.43 (Ecuación de Hermite)**.** $y'' - 2xy' + 2\lambda y = 0$ (oscilador armónico cuántico). Sustituyendo la serie en 0:

$$c_{n+2} = \frac{2(n - \lambda)}{(n+1)(n+2)}\, c_n.$$

Si $\lambda = n$ entero, una de las dos series *se corta* y queda un polinomio de grado n: los *polinomios de Hermite* $H_0 = 1$, $H_1 = 2x$, $H_2 = 4x^2 - 2$,... La cuantización de la energía en mecánica cuántica es, matemáticamente, esta condición de corte.

12.6.2. Puntos singulares regulares: Frobenius

Si x_0 es singular pero $(x - x_0)p$ y $(x - x_0)^2 q$ son analíticas (singularidad *regular*), se busca $y = x^r \sum c_n x^n$; el exponente r satisface la *ecuación indicial*. La estrella es la *ecuación de Bessel*

$$x^2 y'' + xy' + (x^2 - \nu^2)y = 0,$$

con indicial $r^2 = \nu^2$; la solución acotada es la función de Bessel

$$J_\nu(x) = \sum_{k \geq 0} \frac{(-1)^k}{k!\,\Gamma(k + \nu + 1)} \Big(\frac{x}{2}\Big)^{2k+\nu},$$

que describe las vibraciones de una membrana circular (tambor), la difracción y la conducción de calor en cilindros. Junto con Legendre (esferas) y Hermite, forman el zoológico de *funciones especiales*: todas nacen separando variables en las EDP de la física en distintas geometrías.

Regla: resolver una EDO por series de potencias y por Frobenius

Cuando los coeficientes son variables y no hay fórmula cerrada, se busca la solución como serie. El procedimiento cambia según el punto sea ordinario o singular.

1. **Normaliza y clasifica el punto.** Escribe la ecuación como $y'' + p(x)y' + q(x)y = 0$. El punto x_0 es *ordinario* si p y q son analíticas allí; es *singular regular* si no lo son pero $(x - x_0)p(x)$ y $(x - x_0)^2 q(x)$ sí lo son; y singular irregular en cualquier otro caso.

2. **Punto ordinario: propón una serie ordinaria.** Escribe $y = \sum_{n \geq 0} a_n x^n$, deriva

término a término y sustituye. El teorema de Fuchs garantiza que la serie converge al menos hasta la singularidad más cercana en el plano complejo.

3. **Iguala coeficientes y obtén la recurrencia.** Reindexa las sumas para que todas lleven la misma potencia x^n, agrupa, e iguala a cero el coeficiente de cada potencia. Sale una relación que expresa a_{n+2} (o a_{n+k}) en función de anteriores.

4. **Genera las dos soluciones independientes.** Toma $(a_0, a_1) = (1, 0)$ para una y $(0, 1)$ para la otra. En un punto ordinario siempre se obtienen exactamente dos soluciones independientes de este modo.

5. **Punto singular regular: usa Frobenius.** Propón $y = x^r \sum_{n\geq 0} a_n x^n$ con $a_0 \neq 0$. El exponente r no se elige: sale de la *ecuación indicial*

$$r(r-1) + p_0 r + q_0 = 0, \qquad p_0 = \lim_{x\to 0} xp(x), \quad q_0 = \lim_{x\to 0} x^2 q(x).$$

6. **Trata los casos degenerados de Frobenius.** Si las dos raíces indiciales difieren en un entero (o coinciden), la segunda solución puede necesitar un término logarítmico $y_2 = Cy_1 \ln x + x^{r_2} \sum b_n x^n$. Empieza siempre por la raíz *mayor*, que nunca da problemas.

7. **Verifica los primeros términos.** Sustituye la serie truncada en la ecuación y comprueba que los coeficientes de x^0, x^1, x^2 se anulan. Si la ecuación tiene solución elemental conocida, compara los desarrollos.

Ejemplo 12.44 (Tres desarrollos en serie)**. (a) Un caso con solución conocida: $y'' + y = 0$.** El punto $x = 0$ es ordinario. Con $y = \sum a_n x^n$ se tiene $y'' = \sum_{n\geq 0}(n+2)(n+1)a_{n+2}x^n$, luego la ecuación exige

$$(n+2)(n+1)a_{n+2} + a_n = 0 \implies a_{n+2} = -\frac{a_n}{(n+2)(n+1)}.$$

Con $(a_0, a_1) = (1, 0)$: $a_2 = -\frac{1}{2}$, $a_4 = \frac{1}{24}$, y obtenemos $1 - \frac{x^2}{2} + \frac{x^4}{24} - \cdots = \cos x$. Con $(a_0, a_1) = (0, 1)$: $a_3 = -\frac{1}{6}$, $a_5 = \frac{1}{120}$, y sale $\sin x$. El método reconstruye lo que ya sabíamos, que es la mejor manera de confiar en él.

(b) Un caso sin solución elemental: la ecuación de Airy $y'' - xy = 0$. Sigue siendo un punto ordinario. Sustituyendo,

$$\sum_{n\geq 0}(n+2)(n+1)a_{n+2}x^n - \sum_{n\geq 0} a_n x^{n+1} = 0.$$

Reindexando la segunda suma con $m = n + 1$ y agrupando, el coeficiente de x^0 da $2a_2 = 0$, luego $a_2 = 0$; y para $n \geq 1$,

$$a_{n+2} = \frac{a_{n-1}}{(n+2)(n+1)}.$$

Con $(a_0, a_1) = (1, 0)$ los índices no múltiplos de 3 se anulan y queda

$$y_1 = 1 + \frac{x^3}{6} + \frac{x^6}{180} + \frac{x^9}{12\,960} + \cdots,$$

donde $\frac{1}{6}\cdot\frac{1}{30}=\frac{1}{180}$ confirma el segundo paso. Con $(a_0,a_1)=(0,1)$,

$$y_2 = x + \frac{x^4}{12} + \frac{x^7}{504} + \cdots,$$

con $\frac{1}{12}\cdot\frac{1}{42}=\frac{1}{504}$. Ninguna de las dos es una función elemental: son las funciones de Airy, que describen el comportamiento de la luz cerca de una cáustica y de la mecánica cuántica cerca de un punto de retorno. Como p y q son enteras, ambas series convergen en todo $\mathbb{R}$.

(c) Frobenius con raíces que difieren en un entero: $2x^2y''+xy'+(x-1)y=0$. Dividiendo por $2x^2$:

$$y'' + \frac{1}{2x}y' + \frac{x-1}{2x^2}y = 0,$$

de donde $p_0=\lim x\cdot\frac{1}{2x}=\frac{1}{2}$ y $q_0=\lim x^2\cdot\frac{x-1}{2x^2}=-\frac{1}{2}$. El punto es singular regular. La ecuación indicial es

$$r(r-1)+\tfrac{1}{2}r-\tfrac{1}{2}=0 \iff r^2-\tfrac{1}{2}r-\tfrac{1}{2}=0,$$

con raíces $r_1=1$ y $r_2=-\frac{1}{2}$. Difieren en $\frac{3}{2}$, que no es entero, así que ambas raíces dan series de Frobenius limpias, sin logaritmo. La solución general es

$$y = c_1 x\sum_{n\ge0} a_n x^n + c_2 x^{-1/2}\sum_{n\ge0} b_n x^n,$$

y la segunda rama es no acotada cerca del origen: exactamente el tipo de información que un punto singular regular codifica y que una serie ordinaria jamás podría capturar.

Cuidado

En un punto singular no se puede proponer $y=\sum a_n x^n$ directamente: el método falla o produce una sola solución, perdiendo la mitad del espacio de soluciones. El caso más famoso es la ecuación de Bessel de orden 0, $x^2y''+xy'+x^2y=0$, cuyas raíces indiciales son ambas 0: la primera solución es $J_0(x)=1-\frac{x^2}{4}+\frac{x^4}{64}-\frac{x^6}{2304}+\cdots$, y la segunda necesita obligatoriamente un $\ln x$. Esa segunda solución, Y_0, diverge en el origen, y por eso los modos de vibración de un tambor circular la descartan por razones físicas, no matemáticas.

Pausa para pensar

¿Por qué el radio de convergencia de la serie en un punto ordinario llega exactamente hasta la singularidad más cercana *en el plano complejo*, aunque la ecuación sea real? Piensa en $y'=\frac{-2xy}{1+x^2}$, cuya solución es $\frac{1}{1+x^2}$. Sobre la recta real nada llamativo ocurre, y sin embargo la serie $1-x^2+x^4-\cdots$ solo converge para $|x|<1$. La razón está en $x=\pm i$, invisible desde $\mathbb{R}$. Las ecuaciones diferenciales reales son, en el fondo, ecuaciones complejas mirando de perfil.

12.7. Series de Fourier

Idea clave

Intuición primero. El oído hace series de Fourier desde antes de Fourier: el tímpano recibe una sola presión oscilante y la cóclea la descompone en tonos puros: eso que oyes como "acorde" es una descomposición espectral en tiempo real. La afirmación matemática es la misma: toda señal periódica es una superposición de senos y cosenos, y los coeficientes dicen *cuánto de cada nota* lleva. Las fórmulas integrales de los coeficientes no son un truco: son proyecciones ortogonales del capítulo 7 (los senos son perpendiculares entre sí bajo $\int fg$), es decir, "¿cuánta sombra proyecta mi señal sobre esta nota?".

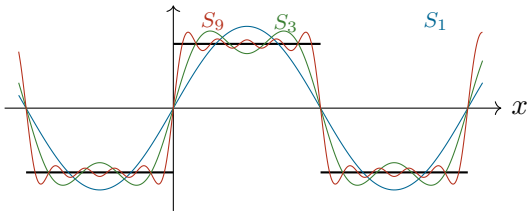


Figura 12.3: Sumas parciales de Fourier de la onda cuadrada: cada armónico nuevo endereza los hombros, y las orejitas persistentes junto al salto son el fenómeno de Gibbs.

Definición 12.45. Para f definida en $[-L, L]$ (o periódica de período $2L$):

$$f(x) \sim \frac{a_0}{2} + \sum_{n\geq 1} \Big(a_n \cos\frac{n\pi x}{L} + b_n \sin\frac{n\pi x}{L} \Big),$$

$$a_n = \frac{1}{L}\int_{-L}^{L} f(x)\cos\frac{n\pi x}{L}\,\mathrm{d}x, \qquad b_n = \frac{1}{L}\int_{-L}^{L} f(x)\sin\frac{n\pi x}{L}\,\mathrm{d}x.$$

Las fórmulas salen de la *ortogonalidad*: $\int_{-L}^{L} \sin\frac{n\pi x}{L}\sin\frac{m\pi x}{L}\mathrm{d}x = L\,\delta_{nm}$, etc.: los senos y cosenos son una "base ortogonal" del espacio de funciones, y los coeficientes son proyecciones, exactamente como en el capítulo 7.

Teorema 12.46 (Convergencia de Dirichlet)**.** *Si f es suave a trozos, la serie converge a $f(x)$ en los puntos de continuidad y al promedio $\frac{f(x^+)+f(x^-)}{2}$ en los saltos. (En los saltos aparece además el* fenómeno de Gibbs*: un sobresalto del ~ 9 % que no desaparece al sumar más términos.)*

Funciones pares solo llevan cosenos; impares, solo senos. Para f definida en $[0, L]$ se usa la *extensión* par o impar según las condiciones de contorno del problema.

Ejemplo 12.47 (La onda cuadrada y un regalo de π)**.** Para $f(x) = 1$ en $(0,\pi)$, -1 en $(-\pi, 0)$ (impar): $b_n = \frac{2}{\pi}\int_0^{\pi} \sin nx\,\mathrm{d}x = \frac{2(1-\cos n\pi)}{n\pi}$, es decir $b_n = \frac{4}{n\pi}$ para n impar y 0 para n par:

$$f(x) = \frac{4}{\pi}\Big(\sin x + \frac{\sin 3x}{3} + \frac{\sin 5x}{5} + \cdots \Big).$$

En $x=\frac{\pi}{2}$: $1=\frac{4}{\pi}(1-\frac{1}{3}+\frac{1}{5}-\cdots)$, recuperando la serie de Leibniz $\frac{\pi}{4}=1-\frac{1}{3}+\frac{1}{5}-\cdots$. Y la identidad de Parseval $\frac{1}{\pi}\int_{-\pi}^{\pi} f^2=\frac{a_0^2}{2}+\sum(a_n^2+b_n^2)$ aplicada aquí da $\sum_{n\ \text{impar}}\frac{1}{n^2}=\frac{\pi^2}{8}$, de donde $\sum_{n\geq 1}\frac{1}{n^2}=\frac{\pi^2}{6}$: el famoso problema de Basilea.

12.8. Ecuaciones en derivadas parciales

Las tres EDP clásicas, cada una prototipo de una familia:

$$\underbrace{u_t=k\,u_{xx}}_{\text{calor (parabólica)}}\qquad\underbrace{u_{tt}=c^2u_{xx}}_{\text{onda (hiperbólica)}}\qquad\underbrace{u_{xx}+u_{yy}=0}_{\text{Laplace (elíptica)}}$$

12.8.1. La ecuación del calor

Problema modelo: barra de longitud L, extremos a temperatura 0, perfil inicial $f(x)$:

$$u_t=k\,u_{xx},\qquad u(0,t)=u(L,t)=0,\qquad u(x,0)=f(x).$$

Separación de variables: $u=X(x)T(t)$ da $\frac{T'}{kT}=\frac{X''}{X}=-\lambda$. Las condiciones de contorno fuerzan $X_n=\operatorname{sin}\frac{n\pi x}{L}$ con $\lambda_n=(\frac{n\pi}{L})^2$, y $T_n=e^{-k\lambda_n t}$. Superponiendo y ajustando la condición inicial con la serie de senos de f:

$$u(x,t)=\sum_{n\geq 1} b_n\,e^{-k(n\pi/L)^2 t}\,\operatorname{sin}\frac{n\pi x}{L},\qquad b_n=\frac{2}{L}\int_0^L f\operatorname{sin}\frac{n\pi x}{L}\,\mathrm{d}x.$$

Los modos altos mueren como e^{-n^2}: el calor *suaviza* instantáneamente. Con extremos aislados ($u_x=0$) se usan cosenos, y la temperatura tiende al promedio inicial.

Idea clave

Separar variables es diagonalizar: los $\operatorname{sin}\frac{n\pi x}{L}$ son *vectores propios* del operador $-\frac{\mathrm{d}^2}{\mathrm{d}x^2}$ con esas condiciones de contorno, y la serie de Fourier es el desarrollo en esa base propia, donde la EDP se desacopla en EDOs triviales, una por modo. La teoría general (operadores de *Sturm–Liouville*) garantiza que estos problemas siempre producen bases ortogonales de funciones propias: Fourier, Bessel y Legendre son casos de un mismo teorema espectral en dimensión infinita.

12.8.2. La ecuación de onda

Cuerda de longitud L fija en los extremos, forma inicial f, velocidad inicial g:

$$u(x,t)=\sum_{n\geq 1}\Big(a_n\cos\frac{n\pi ct}{L}+b_n\operatorname{sin}\frac{n\pi ct}{L}\Big)\operatorname{sin}\frac{n\pi x}{L},$$

con a_n los coeficientes de f y $b_n=\frac{2}{n\pi c}\int_0^L g\operatorname{sin}\frac{n\pi x}{L}\mathrm{d}x$. Cada modo vibra con frecuencia $\frac{nc}{2L}$: la fundamental y sus *armónicos* —la física de la música: acortar la cuerda (menor L) o tensarla

(mayor c) agudiza el tono—. A diferencia del calor, la onda *no* pierde los modos altos: conserva la energía y no suaviza. Sobre la recta infinita vale la solución de *d'Alembert*:

$$u(x,t) = \frac{f(x-ct)+f(x+ct)}{2} + \frac{1}{2c}\int_{x-ct}^{x+ct} g(s)\,\mathrm{d}s\,:$$

dos ondas viajeras a velocidad c, hacia la derecha y hacia la izquierda.

12.8.3. La ecuación de Laplace

$\Delta u = 0$ (funciones *armónicas*) describe estados estacionarios: temperatura final, potencial electrostático, flujo ideal. Dos propiedades notables: el *principio del máximo* (una armónica en una región alcanza su máximo en la frontera) y la *propiedad del valor medio* (el valor en un punto es el promedio sobre cualquier circunferencia centrada en él). En un rectángulo se resuelve por separación con senos en una dirección y senos hiperbólicos en la otra; en el disco, usando el laplaciano polar del capítulo 9 y separando $u = R(r)\Theta(\theta)$, aparecen la ecuación de Euler–Cauchy para R y la solución general

$$u(r,\theta) = \frac{a_0}{2} + \sum_{n\geq 1} r^n(a_n\cos n\theta + b_n \sin n\theta).$$

Regla: resolver una EDP por separación de variables

El método convierte una ecuación en derivadas parciales en dos ordinarias. Funciona cuando el dominio y las condiciones de contorno son compatibles con las coordenadas elegidas.

1. **Comprueba que las condiciones son homogéneas.** Las condiciones de *contorno* deben ser homogéneas (del tipo $u = 0$ o $u_x = 0$ en los bordes). Si no lo son, resta primero una solución estacionaria que las absorba; la condición *inicial* sí puede ser cualquiera.

2. **Propón el producto.** Escribe $u(x,t) = X(x)T(t)$, sustituye y separa de manera que cada miembro dependa de una sola variable. Como una función de x solo puede igualar a una de t si ambas son constantes, aparece la constante de separación, que se escribe $-\lambda$ por conveniencia.

3. **Resuelve primero el problema espacial.** Queda un problema de Sturm–Liouville, por ejemplo $X'' + \lambda X = 0$ con $X(0) = X(L) = 0$. Analiza los tres casos $\lambda < 0$, $\lambda = 0$ y $\lambda > 0$: normalmente solo el tercero da soluciones no triviales.

4. **Extrae los autovalores y autofunciones.** Con las condiciones anteriores salen $\lambda_n = \left(\frac{n\pi}{L}\right)^2$ y $X_n(x) = \sin\frac{n\pi x}{L}$, con $n = 1, 2, 3, \ldots$. Estos son los *modos* del problema, y el conjunto de λ_n es su espectro.

5. **Resuelve la parte temporal para cada modo.** Sustituye cada λ_n en la ecuación de T. Para el calor sale $T_n = e^{-k\lambda_n t}$; para la onda, $T_n = A_n\cos(c\sqrt{\lambda_n}\,t) + B_n\sin(c\sqrt{\lambda_n}\,t)$.

6. **Superpone y ajusta con Fourier.** Escribe $u = \sum_n c_n X_n(x) T_n(t)$ e impón la condición inicial en $t = 0$. Los coeficientes salen de la ortogonalidad de las autofunciones:
$$c_n = \frac{2}{L}\int_0^L f(x)\sin\frac{n\pi x}{L}\,\mathrm{d}x.$$
7. **Verifica el comportamiento asintótico.** Para el calor, todos los modos decaen y $u \to 0$; el modo $n = 1$ es el que sobrevive más tiempo y determina el régimen final. Para la onda no hay decaimiento y la solución es periódica si las frecuencias son conmensurables.

Ejemplo 12.48 (La barra, la cuerda y la placa)**. (a) La ecuación del calor en una barra.** Resolvamos $u_t = k u_{xx}$ en $0 < x < L$ con $u(0,t) = u(L,t) = 0$ y $u(x,0) = f(x)$. Separando, $\frac{T'}{kT} = \frac{X''}{X} = -\lambda$. La parte espacial $X'' + \lambda X = 0$ con extremos nulos obliga a $\lambda_n = \left(\frac{n\pi}{L}\right)^2$ y $X_n = \sin\frac{n\pi x}{L}$, y la temporal da $T_n = e^{-k(n\pi/L)^2 t}$. Por tanto
$$u(x,t) = \sum_{n\geq 1} c_n\, e^{-k(n\pi/L)^2 t}\sin\frac{n\pi x}{L}.$$
Si la condición inicial es $f(x) = \sin\frac{\pi x}{L}$, entonces $c_1 = 1$ y todos los demás se anulan, y la solución es simplemente $e^{-k\pi^2 t/L^2}\sin\frac{\pi x}{L}$: el perfil no cambia de forma, solo baja de amplitud. Con $L = \pi$ y $k = 1$, la amplitud se reduce a la mitad en $t = \ln 2 \approx 0{,}693$.

Obsérvese el factor n^2 en el exponente: el modo $n = 3$ decae 9 veces más deprisa que el fundamental. Por eso el calor *suaviza*: los detalles finos de la condición inicial, que son los modos altos, desaparecen casi instantáneamente, y a tiempos moderados solo queda el primer armónico.

(b) La cuerda vibrante. Con $u_{tt} = c^2 u_{xx}$ y los mismos extremos fijos, la parte espacial es idéntica, pero la temporal da $T_n'' + c^2\lambda_n T_n = 0$, es decir oscilaciones de frecuencia $\omega_n = \frac{n\pi c}{L}$. La solución es
$$u(x,t) = \sum_{n\geq 1}\left[A_n\cos\frac{n\pi c t}{L} + B_n\sin\frac{n\pi c t}{L}\right]\sin\frac{n\pi x}{L}.$$
Las frecuencias son múltiplos enteros de la fundamental ω_1: esa es la razón física de que una cuerda produzca una nota definida y no un ruido. Y como no hay decaimiento, la solución es periódica de periodo $2L/c$: la cuerda vuelve exactamente a su estado inicial. Compárese con (a), donde nada vuelve nunca.

(c) Laplace en un rectángulo. Para $u_{xx} + u_{yy} = 0$ en $[0,a]\times[0,b]$ con $u = 0$ en tres lados y $u(x,b) = g(x)$ en el cuarto, la separación da $X'' + \lambda X = 0$ con $X(0) = X(a) = 0$, o sea de nuevo $X_n = \sin\frac{n\pi x}{a}$, y $Y'' - \lambda Y = 0$ con $Y(0) = 0$, cuya solución es $Y_n = \sinh\frac{n\pi y}{a}$. Entonces
$$u(x,y) = \sum_{n\geq 1} c_n \sinh\frac{n\pi y}{a}\sin\frac{n\pi x}{a}, \qquad c_n = \frac{2}{a\sinh\frac{n\pi b}{a}}\int_0^a g(x)\sin\frac{n\pi x}{a}\,\mathrm{d}x.$$
Aquí el seno hiperbólico crece en vez de decaer: la información del borde caliente se propaga hacia el interior atenuándose, y la atenuación es más brusca para los modos altos. Es el mismo fenómeno de suavizado de (a), pero en el espacio en vez del tiempo.

Idea clave

Los tres problemas comparten exactamente la misma parte espacial, $X'' + \lambda X = 0$ con extremos nulos, y difieren solo en qué hace el tiempo (o la segunda variable espacial) con cada autovalor: decaimiento exponencial en el calor, oscilación en la onda, crecimiento hiperbólico en Laplace. La separación de variables es, en el fondo, diagonalizar el operador $-\mathrm{d}^2/\mathrm{d}x^2$ y resolver una ecuación escalar independiente por cada autovalor. Es el teorema espectral del capítulo de álgebra lineal, aplicado a un espacio de dimensión infinita.

Cuidado

La separación de variables no siempre funciona. Exige que el dominio sea un producto en las coordenadas elegidas: un rectángulo en cartesianas, un disco en polares, una bola en esféricas. Sobre un dominio en forma de L, o sobre un triángulo genérico, el método no arranca, y hay que recurrir a funciones de Green, a métodos variacionales o a elementos finitos. Además, la separación exige que las condiciones de contorno sean homogéneas: con $u(0,t) = T_1 \neq 0$ el paso 2 falla, y hay que restar antes el perfil estacionario lineal que interpola los dos extremos.

12.9. Fundamentos: la teoría que sostiene las EDO

12.9.1. Picard–Lindelöf, demostrado de verdad

Demostración completa. Suponemos $|f(t,y)| \le M$ y la condición de Lipschitz $|f(t,y_1) - f(t,y_2)| \le L|y_1 - y_2|$ en un rectángulo alrededor de (t_0, y_0) (garantizada por la continuidad de $\frac{\partial f}{\partial y}$). Sea $I = [t_0 - h, t_0 + h]$ con $h < \min(1/L,\, a/M, \dots)$ pequeño, y sea X el espacio de funciones continuas $y : I \to \mathbb{R}$ con la norma $\|y\|_\infty = \sup_I |y|$: un espacio *completo* (la convergencia en esta norma es la convergencia uniforme del capítulo 6, y el límite uniforme de continuas es continuo). El operador

$$(Ty)(t) = y_0 + \int_{t_0}^{t} f(s, y(s))\,\mathrm{d}s$$

manda X (restringido a una bola adecuada) en sí mismo, y es contractivo:

$$|(Ty_1)(t) - (Ty_2)(t)| \le \int_{t_0}^{t} |f(s,y_1) - f(s,y_2)|\mathrm{d}s \le L\,h\,\|y_1 - y_2\|_\infty,$$

con $Lh < 1$. El teorema del punto fijo de Banach (demostrado en el capítulo 9) da un único punto fijo y, que es exactamente la única solución del problema de valor inicial: la ecuación integral $y = Ty$ equivale al problema por el teorema fundamental del cálculo. *Los tres capítulos anteriores trabajando juntos en cinco líneas.* □

Teorema 12.49 (Desigualdad de Gronwall)**.** *Si $u \ge 0$ es continua y $u(t) \le C + K\int_{t_0}^{t} u(s)\,\mathrm{d}s$ para $t \ge t_0$, entonces*

$$u(t) \le C\,e^{K(t-t_0)}.$$

Demostración. Sea $V(t) = C + K\int_{t_0}^{t} u$: entonces $V' = Ku \le KV$, y $(Ve^{-Kt})' = e^{-Kt}(V' - KV) \le 0$: Ve^{-Kt} decrece, luego $V(t) \le V(t_0)e^{K(t-t_0)} = Ce^{K(t-t_0)}$, y $u \le V$ concluye. □

Ejemplo 12.50 (Gronwall, la herramienta que acota todo)**. (a) El caso básico.** Si $u(t) \le 1 + 2\int_0^t u$, entonces $u(t) \le \mathrm{e}^{2t}$; en $t = 1$ eso da $u(1) \le 7{,}389056$. La desigualdad integral se convierte en una cota explícita.

(b) Unicidad de soluciones. Si x e y resuelven la misma ecuación con la misma condición inicial y f es Lipschitz de constante K, entonces $|x - y|(t) \le K\int_0^t |x - y|$, es decir, $C = 0$. Gronwall da $|x - y| \le 0$, luego coinciden. La unicidad de Picard es exactamente este corolario.

(c) Dependencia continua. Con condiciones iniciales que difieren en ε, la misma cuenta da $|x-y|(t) \le \varepsilon\,\mathrm{e}^{Kt}$. Las soluciones se separan como mucho exponencialmente, y ese exponente es el que en sistemas caóticos se llama exponente de Lyapunov.

(d) La versión con C variable. Si C depende de t y es creciente, la conclusión pasa a ser $u(t) \le C(t)\mathrm{e}^{K(t-t_0)}$. La monotonía de C es imprescindible: sin ella el argumento de comparación se rompe.

Gronwall es la navaja suiza de la teoría: da de nuevo la unicidad (dos soluciones cumplen $u = |y_1 - y_2| \le L\int u$, es decir $C = 0$: $u \equiv 0$), y además la *dependencia continua de las condiciones iniciales*: soluciones que empiezan a distancia δ se separan a lo más δe^{Lt}. Este crecimiento exponencial de la separación es real, no un artefacto: su tasa optimizada (el *exponente de Lyapunov*) es la definición cuantitativa del caos determinista.

12.9.2. Teoría de sistemas lineales: la matriz fundamental

Teorema 12.51 (Estructura global de los sistemas lineales)**.** *Para $\mathbf{x}' = A(t)\mathbf{x}$ con A continua en un intervalo I: por Picard (versión vectorial, misma demostración), cada problema de valor inicial tiene solución única definida en* todo *I (lo lineal no explota: Gronwall acota $\|\mathbf{x}\|$ por una exponencial). El conjunto de soluciones es un espacio vectorial de dimensión n: la aplicación $\mathbf{x} \mapsto \mathbf{x}(t_0)$ es un isomorfismo con $\mathbb{R}^n$ (inyectiva por unicidad, sobreyectiva por existencia). Una* matriz fundamental *$\Phi(t)$ tiene por columnas una base de soluciones, y la solución general es $\mathbf{x} = \Phi(t)\mathbf{c}$.*

Ejemplo 12.52 (La estructura del espacio de soluciones)**. (a) Dimensión.** Para $\mathbf{x}' = A\mathbf{x}$ con A de tamaño 2×2, el espacio de soluciones tiene dimensión 2: dos constantes libres, una por cada componente del dato inicial. Con $A = \left(\begin{smallmatrix} 0 & 1 \\ -1 & 0 \end{smallmatrix}\right)$ una base es $(\cos t, -\sin t)$ y $(\sin t, \cos t)$.

(b) No homogéneo. Las soluciones de $\mathbf{x}' = A\mathbf{x} + \mathbf{g}$ son una particular más el espacio homogéneo: una variedad afín, no un subespacio. Es la misma estructura que en los sistemas lineales algebraicos.

(c) La matriz fundamental. Si Φ tiene por columnas una base de soluciones, toda solución es $\Phi(t)\mathbf{c}$ y el problema con dato $\mathbf{x}_0$ en t_0 se resuelve como $\mathbf{x} = \Phi(t)\Phi(t_0)^{-1}\mathbf{x}_0$. Cuando A es constante, $\Phi(t) = \mathrm{e}^{At}$.

(d) La linealidad es esencial. Para $x' = x^2$ con $x(0) = 1$ la solución es $x = 1/(1 - t)$, que explota en $t = 1$: el intervalo de existencia depende del dato inicial. En el caso lineal las

soluciones existen en *todo* el intervalo donde A es continua, sin excepción. Ésa es la diferencia práctica más importante entre lineal y no lineal.

Teorema 12.53 (Fórmula de Abel–Liouville)**.** *$W(t) = \det \Phi(t)$ satisface la EDO escalar $W' = \operatorname{tr} A(t)\, W$, y por tanto*

$$W(t) = W(t_0)\, e^{\int_{t_0}^{t} \operatorname{tr} A(s)\, \mathrm{d}s}\,:$$

el wronskiano o nunca se anula o es idénticamente nulo, como afirmamos en su momento.

Idea. det es multilineal en las columnas; derivarlo produce una suma de determinantes donde cada columna se deriva una vez, $\Phi_j' = A\Phi_j$; expandiendo A en cada uno y usando la alternancia, solo sobreviven los términos diagonales de A, que suman $\operatorname{tr} A \cdot \det \Phi$. (Para la ecuación de segundo orden $y'' + py' + qy = 0$, la matriz compañera tiene traza $-p$: se recupera $W = W_0 e^{-\int p}$.) □

Ejemplo 12.54 (El wronskiano y lo que dice)**. (a) Un oscilador.** Para $x'' + x = 0$, escrito como sistema con $A = \left(\begin{smallmatrix} 0 & 1 \\ -1 & 0 \end{smallmatrix}\right)$, se tiene $\operatorname{tr} A = 0$, luego W es constante. Con las soluciones $\cos t$ y $\sin t$, $W = \cos^2 t + \sin^2 t = 1$ para todo t. Cuadra.

(b) Con amortiguamiento. Para $x'' + 2x' + x = 0$ se tiene $\operatorname{tr} A = -2$, luego $W(t) = W(0)\mathrm{e}^{-2t}$. Con las soluciones e^{-t} y $t\mathrm{e}^{-t}$ sale $W = \mathrm{e}^{-2t}$, coherente.

(c) Todo o nada. De la fórmula se sigue que W es idénticamente nulo o no se anula nunca, porque una exponencial jamás vale cero. Por eso basta comprobar la independencia en *un* punto: si $W(t_0) \neq 0$, las soluciones son independientes en todo el intervalo.

(d) El aviso. Ese "todo o nada" vale para soluciones de una misma ecuación lineal, no para funciones cualesquiera. Las funciones x^2 y $x|x|$ tienen wronskiano idénticamente nulo en $\mathbb{R}$ y son linealmente independientes; lo que ocurre es que no son soluciones de ninguna ecuación lineal de segundo orden con coeficientes continuos.

La fórmula de *variación de constantes* también se demuestra limpiamente con la matriz fundamental: buscando $\mathbf{x} = \Phi(t)\mathbf{c}(t)$ en el sistema no homogéneo, los términos con Φ' se cancelan y queda $\mathbf{c}' = \Phi^{-1}\mathbf{f}$, de donde

$$\mathbf{x}(t) = \Phi(t)\Phi(t_0)^{-1}\mathbf{x}_0 + \int_{t_0}^{t} \Phi(t)\Phi(s)^{-1}\mathbf{f}(s)\,\mathrm{d}s,$$

que con A constante es la fórmula $e^{At}\mathbf{x}_0 + \int e^{A(t-s)}\mathbf{f}(s)\mathrm{d}s$ ya vista; la variación de parámetros de segundo orden es su caso 2×2.

12.9.3. La transformada de Laplace, con fundamento

Teorema 12.55 (Existencia e inyectividad)**.** *Si f es continua a trozos y de* orden exponencial *($|f(t)| \leq Me^{ct}$), entonces $F(s) = \int_0^\infty e^{-st} f\,\mathrm{d}t$ converge (absolutamente) para $s > c$, define una función con $F(s) \to 0$ cuando $s \to \infty$, y la transformada es inyectiva sobre funciones continuas (teorema de Lerch): F determina f. La inyectividad es la que autoriza el método de "antitransformar con la tabla".*

Demostración de la convergencia y de la regla de derivación. $\int_0^\infty |e^{-st}f|\mathrm{d}t \le M\int_0^\infty e^{-(s-c)t}\mathrm{d}t = \frac{M}{s-c} < \infty$ (comparación, capítulo 6: nótese que la teoría de integrales impropias es exactamente lo que está trabajando aquí). Para la regla clave, integración por partes:

$$\int_0^R e^{-st}f'(t)\,\mathrm{d}t = \Big[e^{-st}f\Big]_0^R + s\int_0^R e^{-st}f\,\mathrm{d}t \xrightarrow{R\to\infty} -f(0) + sF(s),$$

donde el término de borde en R muere porque $|e^{-sR}f(R)| \le Me^{-(s-c)R} \to 0$. □

Ejemplo 12.56 (La transformada de Laplace y su inversa)**. (a) Las básicas.** $\mathcal{L}\{1\} = 1/s$, $\mathcal{L}\{t\} = 1/s^2$, $\mathcal{L}\{\mathrm{e}^{at}\} = 1/(s-a)$, $\mathcal{L}\{\sin(\omega t)\} = \omega/(s^2+\omega^2)$. Con estas cuatro y la linealidad se cubre la mayoría de los ejercicios.

(b) La inyectividad en acción. Si $\mathcal{L}\{f\} = \dfrac{2}{s^2+4}$, entonces $f(t) = \sin 2t$ y no hay otra opción entre las funciones continuas. Ésa es la propiedad que legitima el método: resolver la ecuación algebraica en s y "deshacer" la transformada sin ambigüedad.

(c) La letra pequeña. La unicidad es *en casi todo punto*: dos funciones que difieran en un conjunto de medida nula tienen la misma transformada. Entre funciones continuas eso obliga a la igualdad, pero con funciones a trozos el valor exacto en los saltos no queda determinado.

(d) Cuándo no existe. $f(t) = \mathrm{e}^{t^2}$ crece más deprisa que cualquier exponencial y su transformada no converge para ningún s. La hipótesis de orden exponencial no es un tecnicismo: delimita el dominio de la herramienta.

Demostración del teorema de convolución.

$$F(s)G(s) = \int_0^\infty\!\!\int_0^\infty e^{-s(u+v)}f(u)g(v)\,\mathrm{d}u\,\mathrm{d}v,$$

y el cambio de variables $t = u+v$, $\tau = u$ (jacobiano 1, capítulo 9) transforma la región $\{u, v \ge 0\}$ en $\{0 \le \tau \le t\}$:

$$= \int_0^\infty e^{-st}\Big(\underbrace{\int_0^t f(\tau)g(t-\tau)\,\mathrm{d}\tau}_{(f*g)(t)}\Big)\mathrm{d}t = \mathcal{L}\{f*g\}(s).$$

El teorema de convolución *es* Fubini más un cambio de variables. □

12.9.4. Series de Fourier: la teoría de la convergencia

¿Por qué converge una serie de Fourier? El análisis pasa por el *núcleo de Dirichlet*. Sustituyendo las fórmulas de los coeficientes en la suma parcial S_N e intercambiando suma e integral:

$$S_Nf(x) = \frac{1}{2\pi}\int_{-\pi}^{\pi} f(x-t)\,D_N(t)\,\mathrm{d}t, \qquad D_N(t) = \sum_{n=-N}^{N} e^{int} = \frac{\sin\left((N+\frac12)t\right)}{\sin(t/2)},$$

donde la suma geométrica cerró el núcleo. La convergencia $S_Nf(x) \to f(x)$ equivale entonces a que estas "medias oscilantes" se concentren en $t = 0$, y el ingrediente decisivo es el *lema de*

Riemann–Lebesgue: $\int g(t)\sin(\lambda t)\,\mathrm{d}t \to 0$ cuando $\lambda \to \infty$ para g integrable (las oscilaciones rápidas se cancelan). Si f es derivable a trozos, la función $\frac{f(x-t)-f(x)}{\sin(t/2)}$ es integrable y el lema da el teorema de Dirichlet que enunciamos, incluida la convergencia al punto medio en los saltos.

Dos complementos importantes. *Parseval*: con la teoría de espacios con producto interno del capítulo 7 aplicada a $\langle f, g\rangle = \frac{1}{\pi}\int_{-\pi}^{\pi} fg$, los senos y cosenos son ortonormales y Parseval es exactamente "Pitágoras en dimensión infinita": la norma al cuadrado es la suma de los cuadrados de las coordenadas. *Fejér*: si en lugar de S_N se toman sus promedios (sumas de Cesàro), la convergencia es *uniforme* para toda f continua —los promedios suavizan el fenómeno de Gibbs—; este es el análogo trigonométrico del teorema de aproximación de Weierstrass (todo continuo en $[a,b]$ es límite uniforme de polinomios).

12.9.5. Sturm–Liouville: el teorema espectral de las EDP

La teoría general detrás de la separación de variables. Un *problema de Sturm–Liouville* es

$$-(p(x)y')' + q(x)y = \lambda\, w(x)\, y, \qquad x \in [a,b],$$

con condiciones de contorno (por ejemplo $y(a) = y(b) = 0$), $p, w > 0$. El operador $Ly = \frac{1}{w}[-(py')' + qy]$ es *autoadjunto* respecto del producto interno $\langle f, g\rangle_w = \int_a^b fg\,w\,\mathrm{d}x$:

$$\langle Lf, g\rangle_w - \langle f, Lg\rangle_w = \Big[p\,(fg' - f'g)\Big]_a^b = 0,$$

donde la integración por partes (¡dos veces!) produjo el término de borde, anulado por las condiciones de contorno. De la autoadjunción salen, con las mismas demostraciones del teorema espectral del capítulo 7:

Teorema 12.57. *Los valores propios de un problema regular de Sturm–Liouville son reales, forman una sucesión $\lambda_1 < \lambda_2 < \cdots \to \infty$, las funciones propias de valores propios distintos son ortogonales respecto de $\langle\cdot,\cdot\rangle_w$, y el sistema de funciones propias es* completo*: toda función razonable se expande en serie de ellas, con coeficientes $c_n = \frac{\langle f, y_n\rangle_w}{\langle y_n, y_n\rangle_w}$.*

Este único teorema contiene todos los desarrollos del capítulo: Fourier en senos ($p = w = 1$, $q = 0$), Bessel en el disco ($p = x$, $w = x$: por eso la ortogonalidad de las Bessel lleva peso x), Legendre en la esfera ($p = 1 - x^2$), Hermite en el oscilador cuántico. Separar variables funciona *siempre* que la geometría produzca un problema de Sturm–Liouville: es el teorema espectral en dimensión infinita, tal como anticipamos.

12.9.6. Métodos de energía: unicidad para el calor y la onda

Teorema 12.58. *El problema del calor $u_t = ku_{xx}$ en $[0,L]$ con condiciones de contorno e inicial dadas tiene a lo sumo una solución.*

Demostración. Sea u la diferencia de dos soluciones: cumple todo con datos cero. Sea $E(t) = \frac{1}{2}\int_0^L u^2\,\mathrm{d}x \geq 0$. Derivando bajo la integral e integrando por partes:

$$E'(t) = \int_0^L u u_t\,\mathrm{d}x = k\int_0^L u u_{xx}\,\mathrm{d}x = k\Big[u u_x\Big]_0^L - k\int_0^L u_x^2\,\mathrm{d}x = -k\int_0^L u_x^2\,\mathrm{d}x \leq 0.$$

E decrece, $E(0) = 0$ y $E \geq 0$: $E \equiv 0$, luego $u \equiv 0$. □

Para la onda, la cantidad correcta es la *energía física* $E = \frac{1}{2}\int (u_t^2 + c^2 u_x^2)\,\mathrm{d}x$, cuya derivada es exactamente cero (se conserva): de nuevo datos nulos fuerzan solución nula. El contraste $E' \leq 0$ (calor: disipa) contra $E' = 0$ (onda: conserva) es la versión matemática precisa de la diferencia física entre difusión y vibración que observamos en los modos.

12.10. Complementos: análisis cualitativo, Lyapunov y numéricos

12.10.1. Ecuaciones autónomas y línea de fases

Para $y' = f(y)$: ceros = equilibrios; signo de f = dirección; $f'(y^*) < 0$ estable, > 0 inestable.

Ejemplo 12.59 (Bifurcación silla–nodo)**.** $y' = \mu + y^2$: dos equilibrios si $\mu < 0$, colisión en $\mu = 0$, ninguno si $\mu > 0$: cambio cualitativo al variar el parámetro. Las bifurcaciones (silla–nodo, horquilla, Hopf) son el esqueleto de los sistemas dinámicos.

12.10.2. Resonancia amortiguada y factor Q

Estacionaria de $my'' + cy' + ky = F_0 \cos \omega t$: amplitud

$$R(\omega) = \frac{F_0}{\sqrt{m^2(\omega_0^2 - \omega^2)^2 + c^2\omega^2}},$$

máximo finito cerca de ω_0, más agudo cuanto menor c; $Q = \frac{m\omega_0}{c}$ mide la agudeza. Radios (RLC), copas, amortiguadores sísmicos: la ingeniería de vibraciones vive aquí.

12.10.3. Lyapunov

Si $V > 0$ alrededor del equilibrio y $\dot{V} = \nabla V \cdot \mathbf{F} \leq 0$: estable; < 0: asintóticamente estable. Sin resolver ni linealizar.

Ejemplo 12.60 (Péndulo amortiguado)**.** $V = \frac{v^2}{2} + (1 - \cos\theta)$: $\dot{V} = -bv^2 \leq 0$: el péndulo con fricción siempre termina colgando (con LaSalle).

12.10.4. Euler y Runge–Kutta

Euler: error global $O(h)$. RK4 (pesos $\frac{1}{6}(1, 2, 2, 1)$ = Simpson): $O(h^4)$: mitad de paso, error entre 16. Rígidos: implícitos; hamiltonianos: simplécticos.

Regla: integrar una EDO numéricamente: Euler, Heun y Runge–Kutta con control del error

La inmensa mayoría de las ecuaciones diferenciales no tiene solución en términos elementales. Cuando hay que dar números, se sustituye la derivada por un cociente y se avanza a pasitos. Estos siete pasos convierten ese eslogan en un método fiable.

1. **Pon la ecuación en forma normal de primer orden.** Necesitas $\mathbf{y}' = \mathbf{f}(x, \mathbf{y})$ con condición inicial $\mathbf{y}(x_0) = \mathbf{y}_0$. Una ecuación de orden n se reduce a un sistema de n ecuaciones de primer orden con $y_1 = y$, $y_2 = y'$, …, $y_n = y^{(n-1)}$. Todos los métodos que siguen valen igual para sistemas: basta leer $\mathbf{y}$ como vector.

2. **Elige el paso h y el número de pasos N.** Para llegar de x_0 a b hacen falta $N = (b - x_0)/h$ pasos, y $x_n = x_0 + nh$.

3. **Euler explícito: el método de un solo término.** La fórmula es

$$y_{n+1} = y_n + h\, f(x_n, y_n).$$

   Es la recta tangente extendida un paso. Error local $O(h^2)$, error global $O(h)$: se dice que tiene *orden 1*. Es fácil de programar y casi siempre demasiado impreciso.

4. **Heun (Runge–Kutta de orden 2): predecir y corregir.** Se calcula

$$k_1 = f(x_n, y_n), \qquad k_2 = f(x_n + h,\ y_n + hk_1), \qquad y_{n+1} = y_n + \frac{h}{2}(k_1 + k_2).$$

   Primero se predice con Euler y luego se promedia la pendiente al principio y al final. Orden 2 por el coste de dos evaluaciones.

5. **Runge–Kutta clásico de orden 4: el caballo de batalla.**

$$\begin{aligned} k_1 &= f(x_n, y_n), & k_2 &= f\Big(x_n + \tfrac{h}{2},\ y_n + \tfrac{h}{2}k_1\Big), \\ k_3 &= f\Big(x_n + \tfrac{h}{2},\ y_n + \tfrac{h}{2}k_2\Big), & k_4 &= f(x_n + h,\ y_n + hk_3), \end{aligned}$$

$$y_{n+1} = y_n + \frac{h}{6}(k_1 + 2k_2 + 2k_3 + k_4).$$

   Cuatro evaluaciones por paso, orden 4. Es la relación calidad–precio que domina la práctica.

6. **Estima el orden empíricamente dividiendo el paso a la mitad.** Si el error global es $O(h^p)$, al pasar de h a $h/2$ el error debe dividirse por 2^p. Así, un cociente de errores cercano a 2 delata orden 1; cercano a 4, orden 2; cercano a 16, orden 4. *Esta es la única verificación que no requiere conocer la solución exacta*, si se usa la solución con paso $h/2$ como referencia (extrapolación de Richardson).

7. **Vigila la estabilidad, que es distinta de la precisión.** Para el modelo $y' = \lambda y$ con $\lambda < 0$, Euler explícito da $y_{n+1} = (1 + h\lambda)y_n$ y solo decae si $|1 + h\lambda| < 1$, es decir $h < 2/|\lambda|$. Si el problema es *rígido* (constantes de tiempo muy distintas), esa restricción puede ser brutal y conviene un método implícito.

Ejemplo 12.61 (El mismo problema resuelto tres veces, y la tabla que revela el orden)**.** Tomamos el problema más simple posible con solución conocida:

$$y' = y, \qquad y(0) = 1, \qquad \text{solución exacta } y(x) = e^x, \quad y(1) = e = 2.718281828.$$

Integramos hasta $x = 1$ con los tres métodos y con cuatro pasos distintos.

(a) Euler a mano. Con $h = 0.1$, la recurrencia $y_{n+1} = y_n + 0.1\,y_n = 1.1\,y_n$ es geométrica, así que

$$y_{10} = (1.1)^{10} = 2.5937425, \qquad \text{error} = 0.1245393.$$

Con $h = 0.05$ sale $(1.05)^{20} = 2.6532977$, error 0.0649841. El error se ha dividido por 1.92, muy cerca de 2: orden 1 confirmado.

(b) Heun a mano. Para $f(x, y) = y$ resulta $k_1 = y_n$, $k_2 = y_n + hy_n = (1 + h)y_n$, luego

$$y_{n+1} = y_n + \frac{h}{2}\big(y_n + (1 + h)y_n\big) = \left(1 + h + \frac{h^2}{2}\right) y_n.$$

Con $h = 0.1$ el factor es 1.105 y $y_{10} = (1.105)^{10} = 2.7140808$, error 0.0042010. Obsérvese que el factor $1 + h + h^2/2$ es la serie de e^h truncada en el término cuadrático: eso *es* tener orden 2.

(c) Runge–Kutta 4 a mano. El mismo cálculo da el factor de amplificación

$$1 + h + \frac{h^2}{2} + \frac{h^3}{6} + \frac{h^4}{24},$$

la serie de e^h truncada en grado 4. Con $h = 0.1$ vale 1.1051708333 y $y_{10} = 2.7182797$, con error 2.08×10^{-6}. Con solo diez pasos ya hay seis cifras correctas.

(d) La tabla completa. Errores absolutos $|y_N - e|$ en $x = 1$:

h	Euler	Heun	RK4
0.2	2.300×10^{-1}	1.557×10^{-2}	3.069×10^{-5}
0.1	1.245×10^{-1}	4.201×10^{-3}	2.084×10^{-6}
0.05	6.498×10^{-2}	1.091×10^{-3}	1.358×10^{-7}
0.025	3.322×10^{-2}	2.779×10^{-4}	8.666×10^{-9}
cociente	≈ 1.95	≈ 3.93	≈ 15.7

Los cocientes de la última fila son la medida experimental del orden: 2^1, 2^2 y 2^4. Y el mensaje práctico: para alcanzar la precisión que RK4 logra con $h = 0.2$ y cinco pasos, Euler necesitaría un paso unas 8800 veces menor, es decir casi cuarenta mil pasos en vez de cinco.

(e) Un problema rígido, donde Euler explota. Sea ahora

$$y' = -50y, \qquad y(0) = 1, \qquad y(1) = e^{-50} = 1.93 \times 10^{-22}.$$

Euler explícito da $y_{n+1} = (1 - 50h)y_n$. Con $h = 0.1$ el factor es $1 - 5 = -4$, y tras diez pasos

$$y_{10} = (-4)^{10} = 1\,048\,576,$$

oscilando y creciendo, cuando la solución verdadera es prácticamente cero. Con $h = 0.05$ el factor es -1.5 y todavía diverge. La condición del paso 7 exige $h < 2/50 = 0.04$; con $h = 0.01$ el factor es 0.5 y $y_{100} = 2^{-100} = 7.9 \times 10^{-31}$, que ya decae correctamente. En cambio Euler *implícito*, $y_{n+1} = y_n/(1 + 50h)$, tiene factor en $(0, 1)$ para *todo* $h > 0$ y nunca explota: es incondicionalmente estable. Ese es el motivo de que los solucionadores profesionales incluyan métodos implícitos.

Cuidado

Precisión y estabilidad son dos propiedades distintas y no se implican. En el apartado (e), Euler con $h = 0.1$ es de orden 1 —tan "preciso" como siempre en el sentido asintótico— y aun así da un resultado catastróficamente falso, porque h está fuera de la región de estabilidad. Reducir el paso arregla la estabilidad, pero a costa de multiplicar el tiempo de cálculo y de acumular más error de redondeo. La regla práctica: si al reducir h la solución cambia de forma cualitativa (deja de oscilar, deja de crecer), el paso anterior era inestable, no simplemente impreciso.

Pausa para pensar

¿De dónde salen los coeficientes $\frac{1}{6}(1, 2, 2, 1)$ de Runge–Kutta? De imponer que el desarrollo de Taylor de y_{n+1} coincida con el de la solución exacta hasta el término en h^4. Es un sistema de ecuaciones sobre los coeficientes del método, y resulta tener más incógnitas que ecuaciones: por eso hay familias enteras de métodos de orden 4 y el "clásico" es solo el más simétrico. La regla de Simpson aparece como caso particular cuando f no depende de y, lo cual es una buena manera de recordar los pesos: son los de Simpson con el punto medio contado dos veces.

12.11. Cierre del temario: los subtemas restantes

12.11.1. Orden n

Espacio de soluciones de dimensión n, wronskiano $n \times n$ con Abel, característica de grado n (multiplicidad m: factores $t^j e^{rt}$). Ejemplo: $y^{(4)} - y = 0$: $y = C_1 e^t + C_2 e^{-t} + C_3 \cos t + C_4 \sin t$ —la viga vibrante: senos con senos hiperbólicos, frecuencias no armónicas—.

12.11.2. Fourier complejo

$f \sim \sum c_n e^{in\pi x/L}$, $c_n = \frac{1}{2L}\int f e^{-in\pi x/L}$; $c_{\pm n} = \frac{a_n \mp i b_n}{2}$; Parseval $\frac{1}{2L}\int |f|^2 = \sum |c_n|^2$. Derivar = multiplicar por $\frac{in\pi}{L}$; productos = convoluciones; y es la forma que discretiza la FFT: el algoritmo del audio, el vídeo y la resonancia magnética.

12.11.3. Clasificación de EDP

$Au_{xx}+Bu_{xy}+Cu_{yy}+\cdots$: $B^2-4AC > 0$ hiperbólica (onda: dos características, propagación), $= 0$ parabólica (calor: difusión), < 0 elíptica (Laplace: equilibrio). Las características son las curvas por donde viaja la información ($x \pm ct$ para la onda); cada tipo pide sus datos (Hadamard).

12.11.4. Dirichlet, Neumann, Robin

$u = g$ (temperatura fijada, cuerda sujeta); $\frac{\partial u}{\partial n} = g$ (flujo impuesto, extremo libre, cosenos con modo constante); $\alpha u + \beta \frac{\partial u}{\partial n} = g$ (enfriamiento de Newton: autovalores trascendentes). Sturm–Liouville sobrevive: solo pide que el término de borde se anule.

12.11.5. Legendre y armónicos esféricos

$(1-x^2)y''-2xy'+\ell(\ell+1)y=0$ (laplaciano esférico con $x=\cos\phi$): soluciones acotadas solo para ℓ entero: $P_0=1$, $P_1=x$, $P_2=\frac{3x^2-1}{2}$, $P_3=\frac{5x^3-3x}{2}$; Rodrigues $P_\ell=\frac{1}{2^\ell \ell!}\frac{\mathrm{d}^\ell}{\mathrm{d}x^\ell}(x^2-1)^\ell$; ortogonales con $\int P_\ell P_m=\frac{2}{2\ell+1}\delta_{\ell m}$. Potenciales axiales: $\sum(A_\ell r^\ell+B_\ell r^{-\ell-1})P_\ell$: monopolo, dipolo, cuadrupolo. Con $e^{im\theta}$: los armónicos esféricos Y_ℓ^m —los orbitales s,p,d,f: la tabla periódica organizada por un Sturm–Liouville—.

12.11.6. El tambor circular

Membrana en el disco: parte radial de Bessel, borde fijo $J_m(kR)=0$: frecuencias $\omega_{mn}=\frac{cj_{m,n}}{R}$ con $j_{0,1}\approx 2.405$, $j_{1,1}\approx 3.832$. Cocientes $1,\ 1.593,\ 2.295,\ldots$ *no enteros*: los sobretonos del tambor no son armónicos: por eso golpea y no canta. Los nodos son las figuras de Chladni.

12.12. Profundización final: Green, el péndulo exacto y Hamilton

12.12.1. Funciones de Green

Para $-y''=f$, $y(0)=y(1)=0$:

$$y(x)=\int_0^1 G(x,s)f(s)\,\mathrm{d}s,\qquad G(x,s)=\begin{cases}s(1-x) & s\le x\\ x(1-s) & s\ge x\end{cases}$$

($-G''=\delta(x-s)$): pegar soluciones homogéneas con salto -1 en la derivada, normalizado por el wronskiano). Verificación con $f=1$: $\int G=\frac{x(1-x)}{2}$ ✓. Respuesta a impulso + superposición = todas las respuestas: la convolución de Laplace adaptada a contornos.

12.12.2. El péndulo exacto

$T=\frac{4}{\omega}K(k)$, $k=\sin\frac{\theta_0}{2}$, con K la integral elíptica; por el binomio integrado:

$$T=\frac{2\pi}{\omega}\Big(1+\frac{\theta_0^2}{16}+\cdots\Big):$$

el período crece con la amplitud (7 % a 60°; diverge hacia 180°). Las elípticas llevan a las funciones elípticas, sus curvas ($y^2=$ cúbica) a la criptografía y a Fermat–Wiles: este péndulo se columpia lejos.

12.12.3. Estabilidad numérica

Euler en $y'=\lambda y$: estable si $|1+h\lambda|<1$ (disco): con λ muy negativo, paso diminuto aunque la solución sea casi cero: *rigidez*. Euler implícito: $\frac{1}{1-h\lambda}$, estable para todo h: pagar una ecuación por paso compra estabilidad incondicional. La disyuntiva organiza todo el software de EDO.

12.12.4. Sistemas hamiltonianos

$\dot q = H_p$, $\dot p = -H_q$: $\dot H = 0$ (cadena) y el flujo conserva área (Liouville: $\operatorname{div}(H_p, -H_q) = 0$: ¡Clairaut!). Péndulo: $H = \frac{p^2}{2} - \omega^2 \cos q$: libración dentro de la separatriz, rotación fuera —el retrato se dibuja sin resolver: curvas de nivel—. Los simplécticos conservan el área exactamente: órbitas planetarias estables millones de años; y $\{f, g\}$ es el ancestro del conmutador cuántico.

12.13. Última capa: características, ciclos límite, Floquet y caos

12.13.1. El método de las características

Para la EDP de primer orden $a(x,t)u_x + u_t = 0$: a lo largo de las curvas $\frac{\mathrm{d}x}{\mathrm{d}t} = a$, la solución es *constante* ($\frac{\mathrm{d}}{\mathrm{d}t}u(x(t),t) = u_x\dot x + u_t = 0$): la EDP se convierte en una familia de EDOs. Transporte puro ($a = c$): $u(x,t) = f(x - ct)$, el dato inicial viajando. Cuasilineal $u_t + u\,u_x = 0$ (Burgers): las características son rectas de pendiente el *valor inicial que transportan*: donde lo alto alcanza a lo bajo, las características se cruzan y la solución clásica muere: nace un *choque* (la ola que rompe, el frenazo en cadena del tráfico). La condición de salto (Rankine–Hugoniot) da la velocidad del choque desde la conservación integral: las ondas de choque de la aerodinámica supersónica salen de esta media página.

12.13.2. Poincaré–Bendixson y ciclos límite

Teorema 12.62. *En el plano, una trayectoria acotada que no converge a un equilibrio se enrolla en un* ciclo límite *(órbita periódica aislada).* Corolario*: el caos necesita dimensión ≥ 3; el plano solo permite equilibrios y ciclos.*

Ejemplo 12.63 (Poincaré–Bendixson y por qué el plano es especial)**. (a) Un ciclo límite.** El sistema $r' = r(1 - r^2)$, $\theta' = 1$ en polares tiene un equilibrio inestable en el origen y una órbita periódica en $r = 1$. Toda trayectoria que arranque fuera del origen converge a esa circunferencia: el teorema garantiza que no puede hacer otra cosa.

(b) Cómo se usa. Si se encuentra una región anular sin equilibrios dentro y en la que el campo apunta hacia adentro por ambos bordes, entonces hay una órbita periódica dentro. Es la manera estándar de demostrar la existencia de oscilaciones sostenidas, por ejemplo en el oscilador de Van der Pol.

(c) En dimensión tres, falla. El sistema de Lorenz tiene trayectorias acotadas que ni convergen a un equilibrio ni son periódicas ni se acercan a un ciclo: llenan un atractor extraño. El caos determinista es imposible en el plano y ordinario en el espacio.

(d) La razón geométrica. En el plano, una curva cerrada separa el interior del exterior (teorema de Jordan), de modo que una trayectoria queda atrapada y no puede enredarse. En dimensión tres una curva no separa nada y la trayectoria puede pasar indefinidamente cerca de sí misma sin cerrarse. La topología decide la dinámica.

El ejemplo canónico es el oscilador de *van der Pol* $\ddot x + \mu(x^2 - 1)\dot x + x = 0$: fricción negativa cerca del origen (bombea energía) y positiva lejos (disipa): una región anular atrapante sin

equilibrios, y Poincaré–Bendixson garantiza el ciclo —el latido—. Van der Pol lo encontró en válvulas de radio; hoy modela marcapasos y neuronas: los relojes autónomos de la naturaleza son ciclos límite.

12.13.3. Teoría de Floquet y resonancia paramétrica

Para sistemas con coeficientes periódicos $\mathbf{x}' = A(t)\mathbf{x}$, $A(t+T) = A(t)$: la matriz fundamental cumple $\Phi(t+T) = \Phi(t)M$ con $M = \Phi(T)$ la *matriz de monodromía*; los autovalores de M (multiplicadores de Floquet) deciden la estabilidad: dentro del círculo unitario, estable. La ecuación modelo es la de *Mathieu* $\ddot{x} + (\omega_0^2 + \varepsilon\cos\omega t)x = 0$: bombear el parámetro a $\omega \approx 2\omega_0$ desestabiliza: *resonancia paramétrica* —el columpio que se impulsa poniéndose de pie y en cuclillas dos veces por vaivén, sin que nadie empuje; también las trampas de iones de los relojes atómicos y la inestabilidad de las mareas de los barcos—. A diferencia de la resonancia forzada, la paramétrica crece aun con fricción si ε supera un umbral: por eso es la peligrosa.

12.13.4. Bifurcación de Hopf

Cuando al variar μ un par de autovalores complejos de la linealización cruza el eje imaginario, el equilibrio entrega su estabilidad a un *ciclo límite* que nace de él con amplitud $\sim \sqrt{\mu - \mu_c}$: el encendido suave de una oscilación. Es el mecanismo por el que un flujo laminar empieza a oscilar, un láser arranca, una red neuronal entra en ritmo: el puente entre los equilibrios del capítulo y los relojes de la sección anterior.

12.13.5. Caos: Lorenz y el mapa logístico

Lorenz (1963), tres EDOs de convección atmosférica:

$$\dot{x} = \sigma(y - x), \quad \dot{y} = x(\rho - z) - y, \quad \dot{z} = xy - \beta z :$$

con $\sigma = 10, \beta = \frac{8}{3}, \rho = 28$, las trayectorias caen a un *atractor extraño* (la mariposa) y exhiben *sensibilidad a las condiciones iniciales*: errores iniciales crecen como $e^{\lambda t}$ con exponente de Lyapunov $\lambda \approx 0.9$ —Gronwall saturado: la predicción meteorológica tiene horizonte, no por ignorancia sino por estructura—. La versión discreta mínima es el *mapa logístico* $x_{n+1} = r x_n(1 - x_n)$: al crecer r, el equilibrio se desdobla en ciclo de 2, luego 4, 8, … (cascada de duplicación con la constante universal de Feigenbaum $\delta \approx 4.669$) hasta el caos en $r \approx 3.5699$. Determinismo sin predictibilidad: el mensaje final de la teoría de EDO, y la frontera con la matemática del siglo XXI.

12.13.6. Comparación de Sturm

Teorema 12.64. *Si $q_1(x) \le q_2(x)$, entre dos ceros de una solución de $y'' + q_1 y = 0$ hay al menos un cero de toda solución de $y'' + q_2 y = 0$:* más rigidez, más oscilación.

Corolarios: las soluciones de $y'' + qy = 0$ con $q \le 0$ no oscilan y con $q \ge m > 0$ oscilan con ceros a distancia $\le \frac{\pi}{\sqrt{m}}$; los ceros de las funciones de Bessel se entrelazan y espacian asintóticamente como π (comparar con $y'' + y = 0$ tras normalizar): la afinación del tambor, cuantificada sin resolver nada.

12.14. Ronda de temarios: Clairaut y las soluciones singulares

Los temarios de Ecuaciones Diferenciales I incluyen un tema clásico que aún no habíamos tocado y que conecta directamente con las envolventes del capítulo 9: las ecuaciones no resueltas en y' y sus soluciones singulares.

12.14.1. La ecuación de Clairaut

Definición 12.65. La *ecuación de Clairaut* es $y = xy' + f(y')$, con f dada.

Teorema 12.66. *Escribiendo $p = y'$: (i) para cada constante c, la recta $y = cx + f(c)$ es solución (¡sustituir y verificar!); (ii) además existe la* solución singular *dada paramétricamente por*

$$x = -f'(p), \qquad y = f(p) - p\,f'(p),$$

que es la envolvente de la familia de rectas y no se obtiene de la fórmula general para ningún c.

Demostración. Derivando $y = xp + f(p)$ respecto de x: $p = p + (x + f'(p))\frac{\mathrm{d}p}{\mathrm{d}x}$, es decir

$$(x + f'(p))\frac{\mathrm{d}p}{\mathrm{d}x} = 0.$$

O bien $p' = 0$ ($p = c$ constante: las rectas), o bien $x = -f'(p)$: sustituyendo en la ecuación queda la curva paramétrica del enunciado. Que esta curva es la envolvente de las rectas es exactamente el cálculo del capítulo 9: con $F(x, y, c) = cx + f(c) - y$, el sistema $F = 0$, $F_c = x + f'(c) = 0$ es *el mismo sistema.* □ □

Ejemplo 12.67. $y = xy' + (y')^2$. Rectas: $y = cx + c^2$. Singular: $x = -2p$, $y = p^2 - 2p^2 = -p^2$, es decir $y = -\frac{x^2}{4}$: una parábola tangente a todas las rectas de la familia. Por el punto $(0, 1)$, que está *encima* de la parábola, pasan dos rectas de la familia ($c = \pm 1$, discriminante $x_0^2 + 4y_0 > 0$); por los puntos estrictamente debajo no pasa *ninguna* solución recta (discriminante negativo); y por cada punto de la parábola pasan exactamente una recta tangente y la propia parábola, con la misma pendiente: la unicidad falla exactamente sobre la solución singular. La razón está en Picard: despejando, $y' = \frac{-x \pm \sqrt{x^2+4y}}{2}$, y en $x^2 + 4y = 0$ (¡la parábola!) la raíz cuadrada tiene derivada infinita respecto de y: la constante de Lipschitz explota justo ahí. La geometría (envolvente), el álgebra (discriminante de la cuadrática en p) y el análisis (hipótesis de Picard) señalan el mismo lugar con tres dedos distintos.

Observación 12.68 (La ecuación de d'Alembert–Lagrange). La generalización $y = x\,\varphi(p) + \psi(p)$ con $\varphi(p) \neq p$ se trata igual: derivando y agrupando queda

$$(\varphi(p) - p)\frac{\mathrm{d}x}{\mathrm{d}p} + \varphi'(p)\,x = -\psi'(p),$$

una ecuación *lineal* para $x(p)$: se resuelve con factor integrante (¡primer método del capítulo!) y la solución queda en forma paramétrica $(x(p), y(p))$. La moraleja del par Clairaut–d'Alembert: cuando la ecuación no está resuelta en y', tratar a p como *variable independiente* y a x como

incógnita suele linealizar el problema: el mismo truco de intercambiar papeles que usamos con $\frac{\mathrm{d}x}{\mathrm{d}y} = 1/\frac{\mathrm{d}y}{\mathrm{d}x}$.

Idea clave

Interpretación por dualidad (la transformada de Legendre del capítulo 9 y de los hamiltonianos): una recta $y = cx + f(c)$ está determinada por su pendiente c y su ordenada $f(c)$; la ecuación de Clairaut dice que la curva solución tiene, en cada punto, una tangente cuya ordenada al origen es función de la pendiente. La solución singular es la curva reconstruida desde sus tangentes: la envolvente *es* la transformada de Legendre de $-f$. Clairaut, la mecánica hamiltoniana y la termodinámica usan la misma dualidad punto–tangente.

12.15. Cierre definitivo: Peano, intervalos maximales y existencia global

El capítulo demostró Picard (existencia y unicidad con Lipschitz) y mostró de pasada los contraejemplos. Cerramos ahora la teoría fundamental con las tres preguntas restantes: ¿qué pasa sin Lipschitz?, ¿hasta dónde vive una solución?, ¿cuándo vive para siempre?

12.15.1. Existencia sin unicidad: el teorema de Peano

Teorema 12.69 (Peano)**.** *Si f es solo* continua *cerca de (t_0, y_0), el problema $y' = f(t, y)$, $y(t_0) = y_0$ tiene* al menos una *solución local.*

Idea de la demostración. Constrúyanse las *poligonales de Euler* y_h con paso h (¡el método numérico como herramienta teórica!). En un rectángulo donde $|f| \le M$, todas las poligonales tienen pendientes acotadas por M: son *equilipschitzianas* y uniformemente acotadas. El teorema de *Arzelà–Ascoli* —el Bolzano–Weierstrass de los espacios de funciones: toda familia acotada y equicontinua tiene subsucesión uniformemente convergente; se demuestra con el argumento diagonal de Cantor del capítulo 6 sobre los racionales— extrae $y_{h_k} \to y$ uniforme. El límite satisface la ecuación integral $y(t) = y_0 + \int_{t_0}^{t} f(s, y(s))\mathrm{d}s$ (la convergencia uniforme permite pasar al límite bajo la integral: capítulo 6), y por el teorema fundamental es solución. Nótese qué se perdió respecto de Picard: la contracción daba *un* punto fijo con unicidad y convergencia de todas las iteradas; la compacidad da *algún* límite de *alguna* subsucesión: existencia sin exclusividad. □ □

Ejemplo 12.70 (El abanico completo de $y' = 3y^{2/3}$)**.** Con $y(0) = 0$: además de $y \equiv 0$ y $y = t^3$, para cualesquiera $a \le 0 \le b$ la función que vale $(t-b)^3$ para $t > b$, cero en $[a, b]$, y $(t-a)^3$ para $t < a$ es solución: *un continuo bi-paramétrico* de soluciones, todas tangentes al eje. La solución puede "despegar del reposo" cuando quiera: sin Lipschitz, el pasado no determina el futuro. (Física del fenómeno: el tanque que se vacía, $h' = -k\sqrt{h}$, tiene la misma raíz no Lipschitz en $h = 0$: mirando un tanque vacío es imposible saber *cuándo* se vació: la no unicidad hacia atrás es la pérdida de memoria del sistema.)

12.15.2. El intervalo maximal y la alternativa de escape

Teorema 12.71. *Toda solución (con f localmente Lipschitz) se extiende a un único* intervalo maximal (α, β)*, y si $\beta < \infty$, entonces la solución* abandona todo compacto*: $|y(t)| \to \infty$ cuando $t \to \beta^-$ (o se acerca a la frontera del dominio de f).*

Idea. Pegar soluciones locales: mientras $(t, y(t))$ permanezca en un compacto K donde f está acotada por M y es Lipschitz, cada punto permite avanzar un paso de tamaño *uniforme* $\delta(K) > 0$ (el radio de Picard depende solo de M y L); si $\beta < \infty$ y la trayectoria volviera a K en tiempos arbitrariamente próximos a β, un paso más la extendería más allá: contradicción con la maximalidad. Escapar de todo compacto es la única forma de morir en tiempo finito. □ □

Ejemplo 12.72. $y' = y^2$, $y(0) = 1$: $y = \frac{1}{1-t}$, maximal en $(-\infty, 1)$: la solución escapa a ∞ en $t = 1$ cumpliendo la alternativa al pie de la letra. Obsérvese la geometría: la ecuación es impecable ($f(y) = y^2$ es C^∞ en todo el plano), y aun así la solución tiene un horizonte. *El dominio de una solución no se lee en la ecuación*: es un resultado, no un dato: el error conceptual más común de todo el tema, ahora con teorema que lo explica.

12.15.3. Existencia global: el criterio de crecimiento lineal

Teorema 12.73. *Si $|f(t, y)| \le a(t)|y| + b(t)$ con a, b continuas (crecimiento a lo más* lineal *en y), toda solución es global: $\beta = +\infty$.*

Demostración. De la ecuación integral, $|y(t)| \le |y_0| + \int_{t_0}^{t} (a|y| + b)$; Gronwall (¡ya demostrado!) da

$$|y(t)| \le \Big(|y_0| + \int_{t_0}^{t} b\Big)\, e^{\int_{t_0}^{t} a}\,:$$

acotada en cada intervalo finito. La alternativa de escape queda bloqueada: la solución no puede irse a infinito en tiempo finito, luego vive para siempre. □ □

Ejemplo 12.74 (Cuándo la solución no explota)**. (a) Crecimiento lineal, todo bien.** Para $y' = ty + 1$ se cumple la hipótesis con $a(t) = |t|$ y $b = 1$, luego la solución existe para todo t por más que crezca deprisa.

(b) Crecimiento cuadrático, explota. Para $y' = y^2$ con $y(0) = 1$ la solución es $y = 1/(1 - t)$, que se va a infinito en $t = 1$. La cota $|f| \le a|y| + b$ falla, y el resultado falla con ella. Con $y(0) = c$ la explosión ocurre en $t = 1/c$: cuanto mayor el dato, antes.

(c) La frontera. Para $y' = y^p$ con $p > 1$ siempre hay explosión en tiempo finito; para $p \le 1$ no la hay. El exponente 1 es exactamente el umbral, y el teorema lo captura en la forma de la hipótesis.

(d) La ecuación logística. $y' = y(1 - y)$ tiene un término cuadrático y sin embargo sus soluciones están acotadas por 1 y existen para todo t: la no linealidad es estabilizadora en vez de explosiva. La hipótesis del teorema es suficiente y no necesaria, y conviene no leerla al revés.

Observación 12.75. Esto explica de un plumazo por qué *todas* las ecuaciones lineales del capítulo (incluidos los sistemas $\mathbf{y}' = A(t)\mathbf{y} + \mathbf{g}$) tienen soluciones globales sin que jamás nos preocupara el dominio, mientras que las no lineales (y^2, Riccati, Euler de los fluidos…) viven bajo amenaza. La frontera exacta es el exponente: $y' = y^p$ explota si $p > 1$. Y sitúa el problema abierto más caro del área: para Navier–Stokes en 3D, decidir si las soluciones suaves son globales o pueden explotar es uno de los problemas del milenio (un millón de dólares): la alternativa de escape de esta sección, planteada a la ecuación que gobierna el agua.

12.16. Segunda ronda: perturbaciones, diferencias y resonancia modal

12.16.1. Lindstedt–Poincaré

En el péndulo débilmente no lineal, la expansión ingenua $y_0 + \varepsilon y_1$ genera términos *seculares* $t \sin t$ (confunde cambio de frecuencia con crecimiento). La cura: expandir también $\omega = 1 + \varepsilon\omega_1 + \cdots$ y elegir ω_1 para matarlos. Resultado: $\omega \approx 1 - \frac{\theta_0^2}{16}$, reproduciendo el desarrollo elíptico exacto sin conocerlo. Esta maquinaria (y escalas múltiples, promediado) es la lengua franca de la mecánica celeste y la física de lo "casi resoluble".

12.16.2. Ecuaciones en diferencias

El gemelo discreto: $x_{n+2} = bx_{n+1} + cx_n$ con característica $r^2 = br + c$: $x_n = C_1 r_1^n + C_2 r_2^n$ (doble: $(C_1 + C_2 n)r^n$): Fibonacci–Binet era esta teoría; estabilidad $|r| < 1$ (el disco: exactamente la frontera de Euler numérico, que *es* una ecuación en diferencias). Hipotecas ($x_{n+1} = (1+i)x_n - P$), filtros digitales IIR y el logístico caótico: este párrafo aplicado. El diccionario $e^{\lambda t} \leftrightarrow r^n$, Laplace↔Z, semiplano↔disco es de ida y vuelta.

12.16.3. Resonancia modal completa

$\ddot{\mathbf{x}} = -\left(\begin{smallmatrix} 2 & -1 \\ -1 & 2 \end{smallmatrix}\right)\mathbf{x} + \cos(\omega t)\binom{1}{0}$: en la base modal, el forzante se proyecta en *ambos* modos: resonancia si $\omega = 1$ o $\omega = \sqrt{3}$: n grados de libertad, n frecuencias críticas. Por eso se calculan *todos* los modos de un puente: el del Milenio (Londres, 2000) resonó lateralmente a 1 Hz con la zancada de los peatones.

12.17. Coda: el oscilador cuántico, resuelto

El problema central de la mecánica cuántica se rinde ante el capítulo: $-\frac{\hbar^2}{2m}\psi'' + \frac{1}{2}m\omega^2 x^2 \psi = E\psi$. Adimensionalizando ($\xi = \sqrt{\frac{m\omega}{\hbar}}x$) y extrayendo la gaussiana asintótica $\psi = H(\xi)e^{-\xi^2/2}$, la ecuación para H es *exactamente* la de Hermite con $\lambda = \frac{E}{\hbar\omega} - \frac{1}{2}$. La serie de potencias se corta (condición para que ψ sea normalizable: si no, $H \sim e^{\xi^2}$ y la gaussiana pierde) solo cuando $\lambda = n$ entero:

$$E_n = \hbar\omega\Big(n + \tfrac{1}{2}\Big), \qquad \psi_n = H_n(\xi)\, e^{-\xi^2/2}\,:$$

niveles *equiespaciados* (por eso los espectros vibracionales moleculares tienen líneas regulares) y energía de punto cero $\frac{\hbar\omega}{2}$ (el vacío vibra: efecto Casimir, estabilidad de la materia). La cuantización no fue un postulado: fue el criterio de corte de una serie de potencias del capítulo 6 aplicado a una EDO del capítulo 12 —la mecánica cuántica de primer curso es, técnicamente, este libro—.

12.18. Temas extra y ejemplos desarrollados

12.18.1. Diccionario mecánica–electricidad

$Lq'' + Rq' + \frac{q}{C} = V(t)$: masa↔inductancia, fricción↔resistencia, resorte↔ $1/C$; $\omega_0 = \frac{1}{\sqrt{LC}}$, sintonizar = resonar. Sistemas distintos, misma ecuación.

12.18.2. Modos normales

Dos masas y tres resortes: $\ddot{\mathbf{x}} = -\frac{k}{m}\left(\begin{smallmatrix} 2 & -1 \\ -1 & 2 \end{smallmatrix}\right)\mathbf{x}$; autovalores 1 y 3: modo simétrico $\omega_1 = \sqrt{k/m}$ y antisimétrico $\omega_2 = \sqrt{3k/m}$; todo movimiento es su superposición (cuasiperiódico: $\sqrt{3} \notin \mathbb{Q}$). El modelo de todas las vibraciones: moléculas, edificios, la cuerda (infinitos modos).

12.18.3. De la serie a la transformada de Fourier

$\hat{f}(\omega) = \int f e^{-i\omega t}\mathrm{d}t$ con inversión $\frac{1}{2\pi}\int \hat{f} e^{i\omega t}\mathrm{d}\omega$. Calor en la recta: cada frecuencia decae $e^{-\omega^2 t}$: el *núcleo del calor* $\frac{1}{\sqrt{4\pi t}} e^{-(x-y)^2/4t}$: dato inicial difuminado con gaussiana $\sqrt{t}$ —browniano, y el desenfoque gaussiano de los editores de imagen—.

12.18.4. El pozo cuántico

$-\frac{\hbar^2}{2m}\psi'' = E\psi$, $\psi(0) = \psi(L) = 0$: exactamente la cuerda: $E_n = \frac{\hbar^2\pi^2}{2mL^2}n^2$. La cuantización es condiciones de contorno; la ortogonalidad, Sturm–Liouville.

12.18.5. Laplace en el rectángulo

$u = \sum b_n \frac{\sinh(n\pi y/a)}{\sinh(n\pi b/a)} \sin\frac{n\pi x}{a}$: senos donde hay dos ceros, hiperbólicos en la otra dirección; los detalles finos del borde mueren como $e^{-n\pi(b-y)/a}$: suavización elíptica. Cuatro lados no nulos: superposición de cuatro problemas.

12.19. Geometría del primer orden: isoclinas y trayectorias ortogonales

Antes de resolver una ecuación conviene saber dibujarla, y de ese dibujo salen aplicaciones clásicas.

Definición 12.76 (Isoclinas)**.** Para $y' = f(x, y)$, la **isoclina** de pendiente m es la curva $f(x, y) = m$: el lugar de los puntos donde el campo de direcciones tiene esa inclinación.

Dibujando unas pocas isoclinas con sus segmentitos se obtiene el retrato cualitativo de todas las soluciones sin integrar nada.

La isoclina $m = 0$ es especialmente útil: marca dónde las soluciones tienen máximos o mínimos, y separa la región donde crecen de aquella donde decrecen.

Regla: trayectorias ortogonales en tres pasos

Dada una familia de curvas $F(x, y, c) = 0$:

1. Deriva respecto de x y *elimina* c entre las dos ecuaciones. Queda la ecuación diferencial $y' = f(x, y)$ que cumple toda la familia.
2. Sustituye y' por $-1/y'$: la nueva ecuación es la de las curvas que cortan a las primeras en ángulo recto.
3. Resuélvela. (Para cortes a un ángulo α cualquiera, la sustitución es $y' \mapsto \dfrac{y' - \tan\alpha}{1 + y'\tan\alpha}$.)

Ejemplo 12.77 (Dos haces de circunferencias que se cortan en ángulo recto)**.** Sea la familia $x^2 + y^2 = cx$: circunferencias que pasan por el origen con centro en el eje X.

(a) Su ecuación diferencial. Derivando, $2x + 2yy' = c$; y despejando $c = (x^2 + y^2)/x$ de la ecuación original y sustituyendo,

$$2x + 2yy' = \frac{x^2 + y^2}{x} \qquad \Longrightarrow \qquad y' = \frac{y^2 - x^2}{2xy}.$$

(b) Las ortogonales. Cambiando y' por $-1/y'$:

$$-\frac{1}{y'} = \frac{y^2 - x^2}{2xy} \qquad \Longrightarrow \qquad y' = \frac{2xy}{x^2 - y^2},$$

que es homogénea y se resuelve con $y = vx$. El resultado es

$$x^2 + y^2 = ky,$$

la familia de circunferencias que pasan por el origen con centro en el eje Y.

(c) La imagen. Dos haces de circunferencias, todas por el origen, unas con centro horizontal y otras vertical, cortándose perpendicularmente en todas partes. Es exactamente el retrato de las líneas de campo y las equipotenciales de un dipolo, y la razón profunda es que ambas familias son las curvas de nivel de la parte real y la parte imaginaria de $1/z$: dos familias conjugadas armónicas se cortan siempre en ángulo recto.

Observación 12.78. El mismo procedimiento resuelve el problema inverso de la física: dadas las isotermas de un cuerpo, las líneas de flujo de calor son sus trayectorias ortogonales; dadas las curvas de nivel de un terreno, las líneas de máxima pendiente lo son. Es el contenido geométrico de que el gradiente sea perpendicular a las curvas de nivel, ahora convertido en un método de cálculo.

12.20. Tres complementos: Bendixson–Dulac, Laplace de señales periódicas y la ecuación de Lagrange

Cómo demostrar que NO hay ciclos límite

Poincaré–Bendixson es la herramienta positiva: si hay una región anular atrapante, hay ciclo. Falta la negativa, que en la práctica se usa más.

Teorema 12.79 (Bendixson–Dulac)**.** *Sea $\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x})$ en una región $D \subset \mathbb{R}^2$ simplemente conexa. Si existe $B \in C^1(D)$ tal que $\operatorname{div}(B\mathbf{f})$ tiene signo constante y no se anula idénticamente en D, entonces* no *hay órbitas periódicas contenidas en D.*

Demostración. Si Γ fuera una órbita cerrada rodeando la región S, el teorema de Green daría

$$\iint_S \operatorname{div}(B\mathbf{f})\,\mathrm{d}A = \oint_\Gamma B\,(f_1\,\mathrm{d}y - f_2\,\mathrm{d}x) = \oint_\Gamma B\,(f_1 f_2 - f_2 f_1)\,\mathrm{d}t = 0,$$

porque sobre la órbita $\mathrm{d}x = f_1\mathrm{d}t$ y $\mathrm{d}y = f_2\mathrm{d}t$. Pero la integral de una función de signo constante y no idénticamente nula no puede ser cero. □

Ejemplo 12.80 (Dos sistemas sin ciclos)**. (a) Con $B = 1$.** Para $\dot{x} = y$, $\dot{y} = -x - y + x^3$,

$$\operatorname{div}\mathbf{f} = \frac{\partial}{\partial x}(y) + \frac{\partial}{\partial y}(-x - y + x^3) = 0 - 1 = -1 < 0$$

en todo el plano. No hay ninguna órbita cerrada, y el análisis se reduce por completo al de los tres equilibrios.

(b) Con B elegida a mano. El sistema de competencia de Lotka–Volterra

$$\dot{x} = x(a - bx - cy), \qquad \dot{y} = y(d - ex - fy)$$

tiene divergencia de signo variable, así que $B = 1$ no sirve. Con la función de Dulac $B = 1/(xy)$, en cambio,

$$\operatorname{div}(B\mathbf{f}) = \frac{\partial}{\partial x}\Big(\frac{a - bx - cy}{y}\Big) + \frac{\partial}{\partial y}\Big(\frac{d - ex - fy}{x}\Big) = -\frac{b}{y} - \frac{f}{x} < 0$$

en todo el cuadrante positivo. Conclusión: dos especies en competencia *nunca* coexisten oscilando; el sistema siempre va a un equilibrio. Que la competencia no produzca ciclos y la depredación sí es un hecho biológico, y aquí queda demostrado en tres líneas.

Laplace de una señal periódica

Teorema 12.81. *Si f es T-periódica en $[0,\infty)$,*

$$\mathcal{L}\{f\}(s) = \frac{1}{1 - e^{-sT}} \int_0^T e^{-st} f(t)\,\mathrm{d}t.$$

Demostración. Pártase la integral en los periodos y sustitúyase $t = nT + u$ en el n-ésimo:

$$\int_0^\infty e^{-st} f = \sum_{n \geq 0} e^{-snT} \int_0^T e^{-su} f(u)\,\mathrm{d}u,$$

y la serie geométrica de razón $e^{-sT}<1$ da el factor. □

Ejemplo 12.82 (Onda cuadrada, y la respuesta de un sistema de primer orden)**.** Sea f de periodo 2, con $f=1$ en $[0,1)$ y $f=0$ en $[1,2)$. Entonces

$$\mathcal{L}\{f\}=\frac{1}{1-e^{-2s}}\int_0^1 e^{-st}\mathrm{d}t=\frac{1-e^{-s}}{s(1-e^{-2s})}=\frac{1}{s\,(1+e^{-s})}.$$

Resolviendo $y'+y=f(t)$ con $y(0)=0$:

$$Y(s)=\frac{1}{s(s+1)(1+e^{-s})},$$

y desarrollando $\frac{1}{1+e^{-s}}=\sum_{n\ge 0}(-1)^n e^{-ns}$ se obtiene la solución como suma de escalones desplazados. Físicamente, la respuesta sube durante el primer segundo, baja durante el segundo, y tras unos pocos ciclos oscila de forma estable alrededor del valor medio de la entrada, que es $\frac12$. Esa es exactamente la manera en que un filtro de primer orden alisa una señal cuadrada, y el cálculo lo predice sin simular nada.

La ecuación de Lagrange

Clairaut, $y=xp+g(p)$, se trató a fondo. Su generalización usa la misma técnica.

Definición 12.83. La **ecuación de Lagrange** (o de d'Alembert) es

$$y=x\,f(p)+g(p),\qquad p=y'.$$

Derivando respecto de x y tomando p como variable independiente, se convierte en una ecuación *lineal* en $x(p)$.

Ejemplo 12.84 ($y=2xp-p^2$, resuelta)**.** Aquí $f(p)=2p$ y $g(p)=-p^2$. Derivando respecto de x:

$$p=2p+2x\frac{\mathrm{d}p}{\mathrm{d}x}-2p\frac{\mathrm{d}p}{\mathrm{d}x}\qquad\Longrightarrow\qquad -p=2(x-p)\frac{\mathrm{d}p}{\mathrm{d}x},$$

y dando la vuelta a la derivada,

$$\frac{\mathrm{d}x}{\mathrm{d}p}+\frac{2}{p}\,x=2,$$

lineal con factor integrante p^2. Su solución es

$$x(p)=\frac{2p}{3}+\frac{C}{p^2},\qquad y(p)=2xp-p^2=\frac{p^2}{3}+\frac{2C}{p},$$

una familia de curvas en paramétricas. Con $C=0$ se puede eliminar el parámetro: $p=3x/2$ y entonces $y=3x^2/4$.

La solución singular. Aparece donde $f(p)=p$, es decir $2p=p$, o sea $p=0$; sustituyendo, $y=0$. Y en efecto la recta $y\equiv 0$ satisface la ecuación y no se obtiene de la familia para ningún C.

Por qué Clairaut es el caso degenerado. Si $f(p)=p$ para *todo* p, el término $x\,\mathrm{d}p/\mathrm{d}x$ es lo único que queda y la ecuación se factoriza en $\mathrm{d}p/\mathrm{d}x=0$ —que da la familia de rectas— o $x+g'(p)=0$ —que da la envolvente—. La lineal en $x(p)$ se vuelve trivial, y por eso Clairaut se resuelve sin integrarla.

12.21. Oscilación y comparación: los teoremas de Sturm

Se afirmó, sin prueba, que los ceros de las funciones de Bessel se entrelazan y que la teoría de Sturm–Liouville produce bases ortogonales completas. Lo primero es un teorema de comparación de media página; lo segundo requiere una corrección importante.

Teorema 12.85 (Comparación de Sturm)**.** *Sean u y v soluciones no triviales de*

$$u'' + q_1(x)\,u = 0, \qquad v'' + q_2(x)\,v = 0$$

en $[a,b]$, con q_1, q_2 continuas y $q_2(x) \geq q_1(x)$ en todo el intervalo. Si $u(a) = u(b) = 0$, entonces v se anula en algún punto de $[a,b]$; y si además $q_2 > q_1$ en algún punto, el cero de v está en el interior (a,b).

Demostración. Supóngase, buscando contradicción, que v no se anula en (a,b). Tómese $a < b$ ceros *consecutivos* de u y, cambiando signos si hace falta, supóngase $u > 0$ y $v > 0$ en (a,b). Considérese el *wronskiano*

$$W = u\,v' - u'\,v.$$

Derivando y sustituyendo las dos ecuaciones,

$$W' = uv'' - u''v = -q_2uv + q_1uv = (q_1 - q_2)uv \;\leq\; 0$$

en todo (a,b), porque $q_1 \leq q_2$ y $u, v > 0$. Por tanto

$$W(b) - W(a) = \int_a^b (q_1 - q_2)uv\,\mathrm{d}x \;\leq\; 0.$$

Ahora evalúense los extremos usando $u(a) = u(b) = 0$:

$$W(a) = -u'(a)\,v(a), \qquad W(b) = -u'(b)\,v(b).$$

Como $u > 0$ inmediatamente a la derecha de a y $u(a) = 0$, se tiene $u'(a) > 0$ (no puede ser nulo: con $u(a) = u'(a) = 0$ la unicidad daría $u \equiv 0$); y $v(a) \geq 0$ por continuidad, luego $W(a) \leq 0$. Simétricamente $u'(b) < 0$ y $v(b) \geq 0$, luego $W(b) \geq 0$. Así

$$W(b) - W(a) \;\geq\; 0.$$

Las dos desigualdades obligan a $W(b) - W(a) = 0$, y por tanto a que el integrando se anule idénticamente: $q_1 \equiv q_2$ en (a,b). Si $q_2 > q_1$ en algún punto, ya hay contradicción. En el caso restante, $W(a) = W(b) = 0$ fuerza $v(a) = v(b) = 0$: v sí se anula, en los extremos del intervalo. □

Corolario 12.86 (Cotas para la separación de los ceros)**.** *Sea $u'' + q(x)u = 0$ con $m^2 \leq q(x) \leq M^2$ y $0 < m \leq M$. Entonces la distancia entre dos ceros consecutivos de cualquier solución no trivial está entre π/M y π/m.*

Demostración. Compárese con $v'' + M^2 v = 0$, cuyas soluciones $\sin(M(x-c))$ tienen ceros separados exactamente π/M. Como $M^2 \geq q$, el teorema 12.85 aplicado con $q_1 = q$, $q_2 = M^2$ dice que entre dos ceros de u hay un cero de v, luego la separación de los de u es al menos π/M. Intercambiando los papeles con $m^2 \leq q$ se obtiene la cota superior π/m. □

Ejemplo 12.87 (Los ceros de Bessel, entrelazados)**.** La ecuación de Bessel $x^2y'' + xy' + (x^2 - \nu^2)y = 0$ se lleva a forma normal con la sustitución $y = u/\sqrt{x}$, que elimina el término en y' y produce

$$u'' + \left(1 + \frac{\frac{1}{4} - \nu^2}{x^2}\right) u = 0.$$

Aquí $q_\nu(x) = 1 + \frac{1/4-\nu^2}{x^2}$ es *decreciente* en ν para $\nu \geq 0$: si $\mu < \nu$, entonces $q_\mu > q_\nu$. Por el teorema de comparación, entre dos ceros consecutivos de J_ν hay un cero de J_μ: los ceros de las Bessel de orden menor se *intercalan* entre los de orden mayor. Y como $q_\nu \to 1$ cuando $x \to \infty$, el corolario dice que la separación entre ceros consecutivos tiende a π: para J_0, los cuatro primeros ceros son 2,4048, 5,5201, 8,6537 y 11,7915, con separaciones 3,1153, 3,1336 y 3,1378, acercándose a $\pi = 3{,}1416$ por debajo. Esa es la razón por la que una membrana circular suena tan distinta de una cuerda: sus modos no están en progresión aritmética exacta, y por eso un tambor no tiene una nota clara.

Cuidado

[Una afirmación del capítulo que hay que corregir] Se dijo que la completitud del sistema de autofunciones de Sturm–Liouville sale "con las mismas demostraciones del teorema espectral" del capítulo de álgebra lineal. Eso no es cierto, y conviene decirlo con claridad: el teorema espectral para matrices simétricas usa que el espacio tiene dimensión finita, y su demostración por inducción sobre la dimensión no tiene análogo en $L^2(w)$. Que las autofunciones *existan* y sean ortogonales sí es elemental; que *generen todo el espacio* es un enunciado de otra naturaleza, y hace falta compacidad.

Teorema 12.88 (Completitud, por el operador de Green)**.** *Sea el problema regular de Sturm–Liouville*

$$-(p(x)y')' + q(x)y = \lambda\, w(x)\, y \quad \text{en } [a,b], \qquad \begin{cases} \alpha_1 y(a) + \alpha_2 y'(a) = 0 \\ \beta_1 y(b) + \beta_2 y'(b) = 0 \end{cases}$$

con $p \in C^1$, $p > 0$, q, w continuas y $w > 0$. Entonces los autovalores forman una sucesión real $\lambda_1 < \lambda_2 < \cdots \to +\infty$, cada uno simple, y las autofunciones normalizadas $\{y_n\}$ constituyen una base ortonormal *de $L^2([a,b], w\,\mathrm{d}x)$.*

Demostración. Paso 1: invertir el operador. Sea $Ly = -(py')' + qy$ con esas condiciones de contorno, y supóngase (trasladando q si hace falta) que 0 no es autovalor. Entonces, para cada f continua, el problema $Ly = f$ tiene solución única, dada por

$$y(x) = \int_a^b G(x,\xi)\, f(\xi)\, \mathrm{d}\xi,$$

donde G es la *función de Green*, que se construye explícitamente a partir de dos soluciones y_a, y_b de $Ly = 0$ que cumplen respectivamente la condición de la izquierda y la de la derecha:

$$G(x,\xi) = \frac{1}{c} \begin{cases} y_a(x)\, y_b(\xi), & x \leq \xi, \\ y_a(\xi)\, y_b(x), & x \geq \xi, \end{cases} \qquad c = p\,(y_a y_b' - y_a' y_b) = \text{cte} \neq 0.$$

Obsérvese que G es *continua*, *simétrica* ($G(x,\xi) = G(\xi,x)$) y acotada.

Paso 2: el operador de Green es compacto y autoadjunto. Defínase $T : L^2(w) \to L^2(w)$ por $(Tf)(x) = \int_a^b G(x,\xi) f(\xi) w(\xi) \mathrm{d}\xi$. Como G es continua sobre el compacto $[a,b]^2$, es de cuadrado integrable, y todo operador integral de núcleo L^2 —un operador de Hilbert–Schmidt— es *compacto* (capítulo 32). La simetría de G y la positividad de w dan $\langle Tf, g\rangle_w = \langle f, Tg\rangle_w$: T es autoadjunto. Y es inyectivo, porque $Tf = 0$ implica $L(0) = f$, o sea $f = 0$.

Paso 3: aplicar el teorema espectral de operadores compactos. El teorema espectral para operadores compactos autoadjuntos —ese sí válido en dimensión infinita, y demostrado en el capítulo 32— da una base ortonormal de $L^2(w)$ formada por autofunciones de T, con autovalores reales $\mu_n \to 0$. Como T es inyectivo, ningún μ_n es 0 y la base es completa. Finalmente, $Ty_n = \mu_n y_n$ equivale a $Ly_n = \mu_n^{-1} y_n$, de modo que las autofunciones de T son exactamente las del problema de Sturm–Liouville, con $\lambda_n = 1/\mu_n$; y $\mu_n \to 0$ se traduce en $|\lambda_n| \to \infty$.

Simplicidad. Dos autofunciones del mismo λ satisfacen la misma ecuación de segundo orden y la misma condición en a, luego son proporcionales por unicidad del problema de valores iniciales. □

Observación 12.89 (Dónde está el trabajo de verdad)*.* Todo el peso lo lleva la palabra *compacto*. La función de Green convierte un problema diferencial —donde el operador no está acotado— en un problema integral con un núcleo continuo, y los operadores integrales de núcleo continuo sobre un compacto son compactos. Esa es la traducción que hace posible la teoría, y es también la razón por la que el resultado falla en dominios no acotados: el operador de Schrödinger en toda la recta tiene espectro continuo, no una sucesión de autovalores, y sus "autofunciones" no son de cuadrado integrable. Fourier en $[0, L]$ da una serie; Fourier en $\mathbb{R}$ da una integral. La diferencia es exactamente la compacidad.

12.22. Poincaré–Bendixson: por qué el caos necesita tres dimensiones

Se usó este teorema para concluir que en el plano no hay caos, sin enunciarlo con precisión ni dar su demostración. Es el resultado más importante de la teoría cualitativa en dimensión dos.

Definición 12.90 (Conjunto límite)**.** Sea $\varphi_t(p)$ el flujo de un campo C^1 en el plano. El *conjunto ω-límite* de p es

$$\omega(p) = \Big\{ q : \ \exists\, t_n \to +\infty \text{ con } \varphi_{t_n}(p) \to q \Big\},$$

es decir, el conjunto de puntos a los que la órbita se acerca infinitas veces.

Teorema 12.91 (Poincaré–Bendixson)**.** *Sea $\vec{F}$ un campo de clase C^1 en un abierto del plano y sea γ^+ una semiórbita positiva contenida en un conjunto compacto K. Si $\omega(\gamma^+)$ no contiene ningún punto de equilibrio, entonces $\omega(\gamma^+)$ es una* órbita periódica*.*

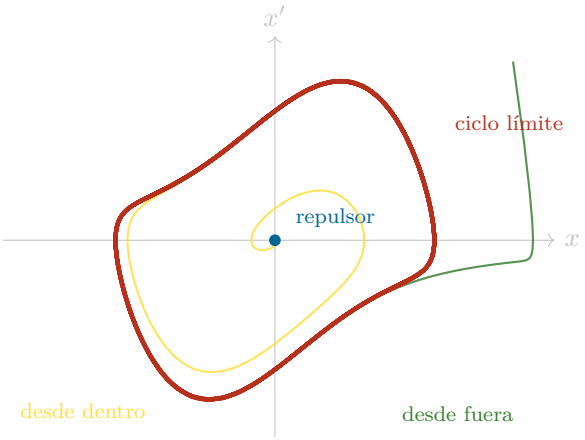


Figura 12.4: Retrato de fases del oscilador de Van der Pol con $\mu = 1$. El origen es un repulsor y las órbitas lejanas entran, de modo que el anillo entre ambas regiones es una región atrapante sin equilibrios; Poincaré–Bendixson garantiza entonces que contiene una órbita periódica. Las dos trayectorias —una que sale del origen y otra que viene de fuera— se enrollan sobre el mismo ciclo, que queda así demostrado sin resolver la ecuación.

En consecuencia, el conjunto límite de toda órbita acotada en el plano es un punto de equilibrio, una órbita periódica, o un conjunto formado por equilibrios y órbitas que los conectan (un gráfico*). No hay más posibilidades.*

Idea clave

La demostración, en cuatro pasos. La herramienta es la *sección transversal*: un segmento Σ que el campo cruza siempre en el mismo sentido.

1. *Monotonía sobre una transversal.* Si una órbita corta a Σ en los puntos $x_1, x_2, x_3, \ldots$ en ese orden temporal, entonces esos puntos están *ordenados monótonamente* a lo largo de Σ. La razón es puramente topológica: el arco de órbita entre x_1 y x_2 junto con el segmento de Σ entre ellos forma una curva de Jordan, que divide al plano en dentro y fuera; el campo no puede volver a cruzarla, luego x_3 queda al otro lado de x_2, en el mismo sentido. *Aquí es donde se usa que la dimensión es dos*, y es el único sitio.

2. *Una transversal corta al conjunto límite en un punto a lo sumo.* Si $q_1, q_2 \in \omega(\gamma^+) \cap \Sigma$ fueran distintos, la órbita tendría que acercarse alternadamente a ambos, contradiciendo la monotonía del paso 1.

3. *Existe una órbita periódica dentro del límite.* Tómese $q \in \omega(\gamma^+)$, que no es equilibrio por hipótesis, y trácese una transversal Σ por q. La órbita de q está contenida en $\omega(\gamma^+)$, que es cerrado e invariante; si esa órbita corta a Σ en dos instantes distintos, por el paso 2 los cortes coinciden, luego la órbita es *cerrada*.

4. *El límite es exactamente esa órbita.* La órbita periódica hallada es un subconjunto cerrado e invariante de $\omega(\gamma^+)$; un argumento de conexión —ω es conexo por ser límite de una órbita— muestra que no puede haber nada más.

El único ingrediente no elemental es el teorema de la curva de Jordan, usado en el paso 1.

Corolario 12.92 (Criterio de la región atrapante)**.** *Si existe una región anular A del plano tal que el campo apunta hacia dentro en ambas fronteras y A no contiene equilibrios, entonces A contiene al menos una órbita periódica.*

Ejemplo 12.93 (El oscilador de Van der Pol tiene ciclo límite)**.** Sea $x'' - \mu(1-x^2)x' + x = 0$ con $\mu > 0$, escrito como sistema. El único equilibrio es el origen, y su linealización tiene autovalores de parte real positiva: es un repulsor, de modo que una circunferencia pequeña alrededor de él es cruzada hacia fuera. Por otro lado, para $|x|$ grande el término $-\mu(1-x^2)x'$ actúa como amortiguamiento fuerte y las órbitas entran: una curva grande es cruzada hacia dentro. El anillo entre ambas es una región atrapante sin equilibrios, y Poincaré–Bendixson garantiza una órbita periódica. Es el *ciclo límite* de Van der Pol, y su existencia queda así demostrada sin resolver la ecuación —cosa que, por lo demás, nadie sabe hacer—.

Observación 12.94 (Por qué el caos necesita dimensión tres)*.* El paso 1 de la demostración es una consecuencia del teorema de la curva de Jordan: en el plano, una curva cerrada separa. Por eso una órbita plana acotada no puede vagar indefinidamente sin repetirse: queda encerrada por su propio pasado. En $\mathbb{R}^3$ ese argumento se derrumba, porque una curva cerrada *no* separa el espacio: una órbita puede pasar por encima o por debajo de su recorrido anterior, y volver a hacerlo, indefinidamente, sin cortarse nunca. Eso es exactamente lo que hace el atractor de Lorenz. La frase "el caos necesita al menos tres dimensiones" no es una observación empírica: es un corolario del teorema de la curva de Jordan.

Teorema 12.95 (Criterio negativo de Bendixson–Dulac)**.** *Sea $\vec{F} = (P, Q)$ de clase C^1 en una región D simplemente conexa del plano. Si existe una función B de clase C^1 y de signo constante en D tal que*

$$\frac{\partial(BP)}{\partial x} + \frac{\partial(BQ)}{\partial y}$$

no cambia de signo en D y no es idénticamente nula, entonces no hay *órbitas periódicas enteramente contenidas en D.*

Demostración. Supóngase que γ es una órbita periódica en D, y sea R la región que encierra, contenida en D por simple conexión. Por el teorema de Green (capítulo 11),

$$\iint_R \left(\frac{\partial(BP)}{\partial x} + \frac{\partial(BQ)}{\partial y} \right) \mathrm{d}A = \oint_\gamma (BP\,\mathrm{d}y - BQ\,\mathrm{d}x).$$

Sobre la órbita, $\mathrm{d}x = P\,\mathrm{d}t$ y $\mathrm{d}y = Q\,\mathrm{d}t$, luego el integrando de la derecha es $B(PQ - QP)\mathrm{d}t = 0$: la integral de línea se anula. Pero la integral doble tiene integrando de signo constante y no idénticamente nulo sobre un abierto, luego es distinta de cero. Contradicción. □

Ejemplo 12.96 (Descartar ciclos en un modelo de competencia)**.** Para el sistema $\dot{x} = x(a - bx - cy)$, $\dot{y} = y(d - ex - fy)$ con todos los parámetros positivos y $x, y > 0$, tómese

$B = 1/(xy)$. Entonces

$$\frac{\partial}{\partial x}\Big(\frac{a-bx-cy}{y}\Big)+\frac{\partial}{\partial y}\Big(\frac{d-ex-fy}{x}\Big) = -\frac{b}{y}-\frac{f}{x} \;<\; 0$$

en todo el cuadrante positivo. Por Bendixson–Dulac no hay ciclos: dos especies que compiten con este modelo nunca oscilan indefinidamente, sino que van a un equilibrio. El truco está en dar con la B adecuada, y $1/(xy)$ es la que funciona en casi todos los modelos de Lotka–Volterra generalizados.

12.23. Definiciones que faltaban: estabilidad, línea de fases, wronskiano y contornos de Robin

A lo largo del capítulo hemos usado con soltura palabras que nunca definimos con precisión: dijimos que un equilibrio era "estable", que el wronskiano "mide la independencia", que las condiciones de Robin "dan autovalores trascendentes". Todo eso era verdad, pero era vocabulario prestado. Esta sección paga la deuda: fija las definiciones exactas, demuestra el criterio de la línea de fases que veníamos aplicando y explica, en cada caso, qué se rompe cuando la definición no se cumple. No es pedantería: en estabilidad, el orden de los cuantificadores decide entre un puente que aguanta y un puente que se cae.

12.23.1. Ecuaciones autónomas, equilibrios y estabilidad

Definición 12.97 (Ecuación autónoma, punto de equilibrio y estabilidad de Lyapunov)**.** Un sistema $\mathbf{y}' = \mathbf{f}(\mathbf{y})$, con $\mathbf{f}:\Omega\subset\mathbb{R}^n\to\mathbb{R}^n$ localmente lipschitziana, se llama *autónomo* cuando el segundo miembro no depende explícitamente de t. Su rasgo característico es la *invariancia por traslación temporal*: si $\mathbf{y}(t)$ es solución, también lo es $\mathbf{y}(t+c)$ para cualquier constante c. Un punto $\mathbf{y}^*\in\Omega$ es un *punto de equilibrio* (o punto crítico, o punto fijo) si $\mathbf{f}(\mathbf{y}^*)=\mathbf{0}$; equivalentemente, si la función constante $\mathbf{y}(t)\equiv\mathbf{y}^*$ es solución para todo $t\in\mathbb{R}$.

Sea $\mathbf{y}(t;\mathbf{y}_0)$ la solución maximal con $\mathbf{y}(0)=\mathbf{y}_0$. El equilibrio $\mathbf{y}^*$ es:

- *estable* (en el sentido de Lyapunov) si

$$\forall\varepsilon>0\;\;\exists\delta>0:\quad \|\mathbf{y}_0-\mathbf{y}^*\|<\delta \implies \|\mathbf{y}(t;\mathbf{y}_0)-\mathbf{y}^*\|<\varepsilon \quad \text{para todo } t\geq 0$$

  (en particular la solución está definida para todo $t\geq 0$);

- *asintóticamente estable* si es estable y además existe $\delta_0>0$ tal que $\|\mathbf{y}_0-\mathbf{y}^*\|<\delta_0$ implica $\lim_{t\to+\infty}\mathbf{y}(t;\mathbf{y}_0)=\mathbf{y}^*$;

- *inestable* si no es estable, es decir, si existe $\varepsilon_0>0$ tal que para todo $\delta>0$ hay algún dato inicial con $\|\mathbf{y}_0-\mathbf{y}^*\|<\delta$ y algún $t>0$ con $\|\mathbf{y}(t;\mathbf{y}_0)-\mathbf{y}^*\|\geq\varepsilon_0$.

Léase despacio el orden de los cuantificadores, porque ahí está todo. Primero se fija la tolerancia ε (cuánto nos permitimos alejarnos) y *después* se busca la precisión δ (cuánto hay que afinar el dato inicial). Es la misma estructura que la definición de límite del capítulo 6, con una diferencia crucial: aquí la desigualdad $\|\mathbf{y}(t)-\mathbf{y}^*\|<\varepsilon$ debe valer *para todo* $t\geq 0$

simultáneamente, no en el límite. Estabilidad significa "si empiezas cerca, te quedas cerca siempre", y eso es una afirmación uniforme en el tiempo.

Cuidado

[Atraer no es lo mismo que estabilizar] Es tentador definir "estable" como "todas las soluciones cercanas convergen a $\mathbf{y}^*$". Sería un error: esa propiedad se llama *atractividad* y no implica estabilidad. El contraejemplo clásico vive en la circunferencia: la ecuación $\theta' = 1 - \cos\theta$ tiene un único equilibrio, $\theta^* = 0$, y como $1 - \cos\theta \geq 0$ toda solución avanza y converge a $2\pi \equiv 0$; el equilibrio atrae a *todo* el círculo. Pero un dato inicial en $\theta_0 = 0{,}01$ —arbitrariamente cerca de 0— se ve obligado a recorrer la circunferencia entera antes de volver: se aleja hasta distancia π. Con $\varepsilon_0 = 1$ la definición falla, y el equilibrio es *inestable y atractor a la vez*. Por eso la definición pide las dos cosas por separado, y por eso "asintóticamente estable" se define como "estable *y* atractor", con la conjunción explícita.

12.23.2. La línea de fases y el criterio de la derivada

Definición 12.98 (Línea de fases)**.** Para una ecuación escalar autónoma $y' = f(y)$ con $f \in C^1(I)$, la *línea de fases* es el intervalo I del eje y decorado con dos datos: los ceros de f, marcados como puntos (los equilibrios), y una flecha en cada subintervalo entre ceros consecutivos, dirigida hacia la derecha donde $f > 0$ y hacia la izquierda donde $f < 0$.

La línea de fases es el diagrama de fase completo de la ecuación, y es legítimo porque en dimensión uno las soluciones no se pueden cruzar (unicidad) ni oscilar: entre dos equilibrios consecutivos f tiene signo constante, luego $y(t)$ es *monótona*, y al estar acotada por los equilibrios que la encierran converge necesariamente a uno de ellos. Toda la dinámica escalar autónoma cabe en esa frase; no hay ciclos, ni caos, ni sorpresas. Dibujar la línea de fases es más informativo que resolver, y cuesta treinta segundos.

Proposición 12.99 (Criterio de la derivada en el equilibrio)**.** *Sea $f \in C^1$ y $f(y^*) = 0$, y escríbase $\lambda = f'(y^*)$.*

1. *Si $\lambda < 0$, el equilibrio y^* es asintóticamente estable, y la convergencia es exponencial: $|y(t) - y^*| \leq |y_0 - y^*|\, e^{(\lambda+\varepsilon)t}$ para datos suficientemente próximos y cualquier $\varepsilon > 0$ pequeño.*
2. *Si $\lambda > 0$, el equilibrio es inestable.*
3. *Si $\lambda = 0$, el criterio* no decide.

Demostración. Sea $u = y - y^*$. Por el teorema de Taylor con resto,

$$u' = f(y^* + u) = \lambda u + R(u), \qquad \frac{R(u)}{u} \xrightarrow[u\to 0]{} 0.$$

Fíjese ε con $0<\varepsilon<|\lambda|$ y tómese $\delta>0$ tal que $|R(u)|\le\varepsilon|u|$ siempre que $|u|\le\delta$. La cuenta decisiva es la derivada del cuadrado de la distancia:

$$\frac{\mathrm{d}}{\mathrm{d}t}\Big(\frac{u^2}{2}\Big)=u\,u'=\lambda u^2+u\,R(u).$$

Si $\lambda<0$, mientras $|u|\le\delta$ se tiene $uR(u)\le\varepsilon u^2$ y por tanto $\frac{\mathrm{d}}{\mathrm{d}t}(u^2/2)\le(\lambda+\varepsilon)u^2<0$ salvo en $u=0$. En consecuencia u^2 es decreciente: la trayectoria nunca abandona el intervalo $|u|\le\delta$ (esto es exactamente la estabilidad, tomando $\delta(\varepsilon)=\min\{\delta,\varepsilon\}$), y la desigualdad diferencial $\frac{\mathrm{d}}{\mathrm{d}t}u^2\le 2(\lambda+\varepsilon)u^2$ integrada —Gronwall en su versión más simple— da $|u(t)|\le|u(0)|e^{(\lambda+\varepsilon)t}\to 0$.

Si $\lambda>0$, la misma cuenta con las desigualdades invertidas da $\frac{\mathrm{d}}{\mathrm{d}t}(u^2/2)\ge(\lambda-\varepsilon)u^2>0$ mientras $0<|u|\le\delta$, de modo que $|u(t)|\ge|u(0)|e^{(\lambda-\varepsilon)t}$ crece hasta salir del intervalo de radio δ en tiempo finito. Con $\varepsilon_0=\delta$, ningún $\delta'>0$ sirve en la definición: inestable. □

Obsérvese qué se ha hecho realmente: sustituir la ecuación por su *linealización* $u'=\lambda u$ y comprobar que el término no lineal, siendo $o(u)$, es demasiado pequeño para cambiar el signo de la derivada de u^2. Esa es la idea que en dimensión superior se llama teorema de Hartman–Grobman, y esta demostración de tres líneas es su caso escalar completo.

Ejemplo 12.100 (Un pozo entre dos cimas, y tres casos degenerados)**. (a) El caso genérico.** Sea $y'=y^3-y=y(y-1)(y+1)$. Equilibrios: -1, 0, 1. Como $f'(y)=3y^2-1$, resulta $f'(-1)=2>0$, $f'(0)=-1<0$ y $f'(1)=2>0$: las dos cimas son inestables y el pozo central es asintóticamente estable. La línea de fases lo confirma sin derivar nada: $f<0$ en $(-1,0)$ y $f>0$ en $(0,1)$, luego ambas flechas apuntan hacia el 0. La integración numérica desde $y_0=0{,}5$ da $y(4)=0{,}010574$, del orden de $0{,}5\,e^{-4}=0{,}009158$ que predice la tasa $\lambda=-1$; desde $y_0=1{,}5$ la solución escapa a infinito (de hecho en tiempo finito).

(b) $\lambda=0$ con estabilidad. Para $y'=-y^3$ se tiene $f'(0)=0$ y el criterio calla, pero la línea de fases decide: $f>0$ a la izquierda y $f<0$ a la derecha, luego 0 es asintóticamente estable. La solución explícita es $y(t)=y_0/\sqrt{1+2y_0^2t}$, y la convergencia es *algebraica*, no exponencial: con $y_0=1$, en $t=50$ vale $0{,}09950372$ (la integración numérica da $0{,}09950372$, coincidiendo en nueve cifras), mientras que $e^{-50}\approx 1{,}93\cdot10^{-22}$. La estabilidad se conserva; la velocidad, no.

(c) $\lambda=0$ con inestabilidad, y el caso semiestable. Para $y'=y^3$ el origen es inestable, y para $y'=y^2$ es *semiestable*: atrae por la izquierda y repele por la derecha. Los tres ejemplos tienen la misma linealización, $u'=0$, y comportamientos distintos. Eso es lo que significa exactamente "el criterio no decide".

12.23.3. El wronskiano, con su definición y su ley de evolución

Definición 12.101 (Wronskiano)**.** Dadas n funciones $y_1,\dots,y_n$ derivables $n-1$ veces en un intervalo I, su *wronskiano* es el determinante

$$W(x)=W[y_1,\dots,y_n](x)=\begin{vmatrix} y_1 & y_2 & \cdots & y_n\\ y_1' & y_2' & \cdots & y_n'\\ \vdots & \vdots & & \vdots\\ y_1^{(n-1)} & y_2^{(n-1)} & \cdots & y_n^{(n-1)}\end{vmatrix}.$$

Para $n=2$, $W = y_1y_2' - y_1'y_2$. Si además $y_1,\dots,y_n$ son soluciones de una misma ecuación lineal homogénea de orden n con coeficientes continuos en I, forman un *sistema fundamental* (es decir, una base del espacio de soluciones) si y solo si $W(x_0)\neq 0$ para algún $x_0\in I$.

Proposición 12.102 (Fórmula de Abel para el orden dos)**.** *Si y_1, y_2 resuelven $y''+p(x)y'+q(x)y=0$ con p,q continuas en I, su wronskiano satisface la ecuación diferencial de primer orden*

$$W' = -p(x)\,W, \qquad \textit{luego} \qquad W(x) = W(x_0)\,\exp\Big(-\int_{x_0}^{x} p(t)\,\mathrm{d}t\Big).$$

En consecuencia, W o no se anula nunca en I, o es idénticamente nulo*: no hay término medio.*

Demostración. Derívese $W = y_1y_2' - y_1'y_2$: los términos $y_1'y_2'$ se cancelan y queda $W' = y_1y_2'' - y_1''y_2$. Sustituyendo $y_i'' = -py_i' - qy_i$, los términos con q se cancelan por simetría y sobrevive $W' = -p(y_1y_2' - y_1'y_2) = -pW$. La exponencial nunca se anula, así que el signo de W lo fija por completo el valor $W(x_0)$. □

La consecuencia es la que hace útil la fórmula. Para verificar que dos soluciones son independientes basta evaluar el wronskiano en *un* punto cómodo, normalmente donde las condiciones iniciales son limpias; no hace falta comprobarlo en todo el intervalo. Y a la inversa: si al calcular W sale una función que se anula en un punto interior y no en otro, hay un error de cálculo, garantizado. Es el mejor control de errores gratuito del capítulo. La versión matricial general, $W' = \operatorname{tr} A(t)\,W$, ya se demostró en la sección de fundamentos; la de arriba es su caso $n=2$, donde la matriz compañera tiene traza $-p$.

Ejemplo 12.103 (Un wronskiano que explota pero no se anula)**.** Sea la ecuación de Legendre con $\ell=1$ en el intervalo $(-1,1)$:

$$(1-x^2)y'' - 2xy' + 2y = 0.$$

Dos soluciones son $y_1 = x$ y $y_2 = \frac{x}{2}\ln\frac{1+x}{1-x} - 1$ (compruébese sustituyendo: el cálculo simbólico confirma que ambas anulan la ecuación). Normalizando, $p(x) = \frac{-2x}{1-x^2}$, y la fórmula de Abel predice

$$W(x) = W(0)\exp\Big(\int_0^x \frac{2t}{1-t^2}\,\mathrm{d}t\Big) = \frac{W(0)}{1-x^2}.$$

El cálculo directo da $W(x) = \frac{1}{1-x^2}$, o sea $W(0)=1$; en $x=\frac12$ vale $\frac43 = 1{,}333333$ y en $x=0{,}9$ vale $\frac{100}{19} = 5{,}263158$, exactamente lo que predice la fórmula. El wronskiano crece sin límite al acercarse a ±1 —donde el coeficiente principal $1-x^2$ se anula y la ecuación deja de ser regular— pero jamás se hace cero dentro del intervalo: las dos soluciones son independientes en todo $(-1,1)$, y la segunda es la que la física descarta por no ser acotada.

12.23.4. Dirichlet, Neumann y Robin: la tercera condición de contorno

Definición 12.104 (Condiciones de contorno de Dirichlet, Neumann y Robin)**.** Sea u definida en un dominio Ω con frontera $\partial\Omega$ y normal exterior unitaria $\mathbf{n}$. Una condición de contorno prescribe el comportamiento de u sobre $\partial\Omega$:

- *Dirichlet*: $u = g$ sobre $\partial\Omega$ (se impone el valor);

- *Neumann*: $\frac{\partial u}{\partial n} = \nabla u \cdot \mathbf{n} = g$ (se impone el flujo);
- *Robin* (o de tercera especie, o mixta): con α, β constantes no ambas nulas,

$$\alpha\, u + \beta\, \frac{\partial u}{\partial n} = g \qquad \text{sobre } \partial\Omega.$$

Dirichlet es el caso $\beta = 0$ y Neumann el caso $\alpha = 0$: Robin las contiene a las dos y las interpola.

El caso físico que la produce es el *enfriamiento de Newton*. Una barra cuyo extremo no está ni a temperatura fijada (eso sería un termostato perfecto: Dirichlet) ni perfectamente aislada (eso sería flujo nulo: Neumann), sino en contacto con un ambiente a temperatura u_∞, pierde calor a un ritmo proporcional a la diferencia de temperatura. La ley de Fourier iguala ese ritmo con el flujo conductivo y produce exactamente $-k\,\frac{\partial u}{\partial n} = h\,(u - u_\infty)$, es decir $h\,u + k\,\frac{\partial u}{\partial n} = h\,u_\infty$: una condición de Robin, con $\alpha/\beta = h/k$ midiendo cuánto pesa el intercambio con el ambiente frente a la conducción interna. Los dos casos límite son los esperados: $h \to \infty$ (contacto perfecto) da Dirichlet, y $h \to 0$ (aislamiento) da Neumann.

Ejemplo 12.105 (Por qué Robin produce autovalores trascendentes)**.** Sepárense variables en la ecuación del calor en $[0,1]$ con extremo izquierdo a temperatura fija y extremo derecho enfriándose al ambiente: $X'' + \lambda X = 0$, $X(0) = 0$, $X'(1) + X(1) = 0$. Con $\lambda = \mu^2 > 0$ se tiene $X = \sin(\mu x)$, y la condición de Robin en $x = 1$ obliga a

$$\mu \cos\mu + \sin\mu = 0, \qquad \text{es decir} \qquad \tan\mu = -\mu.$$

Y aquí está la diferencia. Con Dirichlet en ambos extremos, la condición sería $\sin\mu = 0$, cuyas raíces son $\mu_k = k\pi$: fórmula cerrada, exacta, y con ella las series de Fourier de siempre. Con Robin, la ecuación mezcla una función trigonométrica con un polinomio y *no tiene solución en forma cerrada*: hay que resolverla numéricamente. Las cinco primeras raíces son

$$\mu_1 = 2{,}028758, \quad \mu_2 = 4{,}913180, \quad \mu_3 = 7{,}978666, \quad \mu_4 = 11{,}085538, \quad \mu_5 = 14{,}207437,$$

con autovalores $\lambda_k = \mu_k^2$ iguales a 4,115858, 24,139342, 63,659107, 122,889162 y 201,851258. Compárense con los de Dirichlet, $\pi, 2\pi, 3\pi, \cdots = 3{,}141593,\ 6{,}283185,\ 9{,}424778, \ldots$: los de Robin están sistemáticamente por debajo, y de hecho μ_k decrece hacia $(k - \frac{1}{2})\pi$, al que se acerca desde arriba con diferencias 0,458, 0,201, 0,125, 0,090, 0,070. Lo importante es que la teoría de Sturm–Liouville sobrevive intacta —los autovalores siguen siendo reales, simples, crecientes al infinito, y las autofunciones $\sin(\mu_k x)$ siguen siendo ortogonales y completas—, porque lo único que exige la simetría del operador es que el término de borde se anule, y la condición de Robin lo consigue igual que las otras dos. Lo que se pierde no es la teoría: es la fórmula.

12.23.5. Soluciones singulares

Definición 12.106 (Solución singular)**.** Sea $\Phi(x, y, C) = 0$ una familia uniparamétrica que representa la solución general de una ecuación de primer orden. Una solución $y = \varphi(x)$ de esa ecuación se llama *singular* si no se obtiene de la familia para ningún valor del parámetro C (ni

siquiera como límite $C \to \infty$ bajo cambios de parametrización). Típicamente es la *envolvente* de la familia: la curva tangente en cada uno de sus puntos a algún miembro de la familia, sin coincidir con ninguno.

La existencia de soluciones singulares no contradice el teorema de Picard–Lindelöf: lo delata. En cada punto de la envolvente pasan al menos dos soluciones con la misma pendiente —el miembro tangente de la familia y la propia envolvente—, luego la unicidad falla exactamente allí, y por tanto la hipótesis de Lipschitz tiene que fallar allí también. La envolvente es, entonces, el lugar geométrico donde la ecuación resuelta en y' deja de ser lipschitziana: los tres lenguajes —geometría (tangencia), álgebra (discriminante) y análisis (Picard)— señalan el mismo conjunto. El caso de Clairaut, tratado en la sección correspondiente de este capítulo, es el ejemplo canónico; la definición de arriba es la que allí usábamos sin enunciar.

12.23.6. Clasificar equilibrios en dimensión dos

Regla: Clasificar los equilibrios de un sistema no lineal

Para $\mathbf{x}' = \mathbf{f}(\mathbf{x})$ en el plano, con $\mathbf{f} = (f_1, f_2)$ de clase C^1:

1. **Halla todos los ceros de f.** Resuelve el sistema $f_1(x, y) = 0$, $f_2(x, y) = 0$. Casi siempre conviene factorizar cada ecuación y combinar los casos; no te saltes los equilibrios de frontera (los que tienen alguna coordenada nula), que en los modelos de poblaciones son los que dicen si una especie se extingue.

2. **Calcula la matriz jacobiana y evalúala en cada equilibrio.**

$$D\mathbf{f}(x, y) = \begin{pmatrix} \frac{\partial f_1}{\partial x} & \frac{\partial f_1}{\partial y} \\ \frac{\partial f_2}{\partial x} & \frac{\partial f_2}{\partial y} \end{pmatrix}, \qquad A = D\mathbf{f}(\mathbf{x}^*).$$

   Deriva la expresión general una sola vez y sustituye después: derivar en cada punto por separado multiplica los errores.

3. **Obtén los autovalores.** Con $\lambda^2 - (\operatorname{tr} A)\lambda + \det A = 0$ basta; muchas veces ni hace falta resolver, porque el signo de $\det A$ y el de $\operatorname{tr} A$ ya deciden.

4. **Clasifica.** Si $\det A < 0$: *punto de silla* (inestable). Si $\det A > 0$, mira $(\operatorname{tr} A)^2 - 4 \det A$: positivo da *nodo* y negativo da *foco* o espiral; y en ambos casos el signo de $\operatorname{tr} A$ decide, negativo estable y positivo inestable. Si $\operatorname{tr} A = 0$ y $\det A > 0$, la linealización es un *centro*.

5. **Invoca Hartman–Grobman.** Si *ningún* autovalor tiene parte real nula (el equilibrio se llama entonces *hiperbólico*), existe un homeomorfismo local que transforma las órbitas del sistema no lineal en las del lineal: el retrato de fase cerca de $\mathbf{x}^*$ es el del sistema linealizado, deformado pero no cambiado de tipo. Solo entonces tienes derecho a trasladar la conclusión.

6. **Detente si aparece parte real nula.** Si $\operatorname{tr} A = 0$ con $\det A > 0$ —el caso centro— o si algún autovalor es cero, la linealización *no* decide: el sistema no lineal puede

tener ahí un centro genuino, un foco atractor o un foco repulsor. Hace falta otra herramienta: una integral primera, una función de Lyapunov, simetría, o el cálculo de coeficientes de Lyapunov de orden superior.

7. **Verifica.** Comprueba que $\mathbf{f}(\mathbf{x}^*) = \mathbf{0}$ numéricamente, que la traza y el determinante que has usado coinciden con los de la matriz, e integra el sistema desde un dato cercano: un foco estable debe acercarse girando, con periodo aproximado $2\pi/|\operatorname{Im}\lambda|$.

Ejemplo 12.107 (Depredador y presa con presa logística, resuelto entero)**.** Consideremos el modelo

$$\begin{cases} x' = x(1-x) - xy, \\ y' = y\left(x - \tfrac{1}{2}\right), \end{cases}$$

donde $x \geq 0$ es la presa (que crece logísticamente hacia su capacidad 1 y es depredada) e $y \geq 0$ el depredador (que solo crece si la presa supera el umbral $\tfrac{1}{2}$).

Paso 1: equilibrios. De $y(x - \tfrac{1}{2}) = 0$ sale $y = 0$ o $x = \tfrac{1}{2}$. Si $y = 0$, la primera ecuación da $x(1-x) = 0$, o sea $x = 0$ o $x = 1$. Si $x = \tfrac{1}{2}$, la primera da $\tfrac{1}{2}(1 - \tfrac{1}{2}) - \tfrac{1}{2}y = 0$, luego $y = \tfrac{1}{2}$. Hay tres equilibrios:

$$E_1 = (0,0), \qquad E_2 = (1,0), \qquad E_3 = \left(\tfrac{1}{2}, \tfrac{1}{2}\right).$$

La evaluación numérica de $\mathbf{f}$ en los tres da $(0,0)$ exactamente.

Paso 2: la jacobiana. Con $f_1 = x - x^2 - xy$ y $f_2 = xy - \tfrac{1}{2}y$,

$$D\mathbf{f}(x,y) = \begin{pmatrix} 1 - 2x - y & -x \\ y & x - \tfrac{1}{2} \end{pmatrix}.$$

Paso 3 y 4: los tres retratos locales.

- En $E_1 = (0,0)$: $A = \left(\begin{smallmatrix} 1 & 0 \\ 0 & -\frac{1}{2} \end{smallmatrix}\right)$, diagonal, con autovalores 1 y $-\tfrac{1}{2}$. Como $\det A = -\tfrac{1}{2} < 0$, es un **punto de silla**: la extinción total es inestable, y lo es en la dirección de la presa (autovalor positivo 1), como debe ser.

- En $E_2 = (1,0)$: $A = \left(\begin{smallmatrix} -1 & -1 \\ 0 & \frac{1}{2} \end{smallmatrix}\right)$, triangular, con autovalores -1 y $\tfrac{1}{2}$. De nuevo $\det A = -\tfrac{1}{2} < 0$: **silla**. La presa sola en su capacidad de carga es inestable frente a la introducción de depredadores, porque $x = 1 > \tfrac{1}{2}$ y al depredador le compensa entrar.

- En $E_3 = (\tfrac{1}{2}, \tfrac{1}{2})$: $A = \left(\begin{smallmatrix} -\frac{1}{2} & -\frac{1}{2} \\ \frac{1}{2} & 0 \end{smallmatrix}\right)$, con $\operatorname{tr} A = -\tfrac{1}{2}$ y $\det A = \tfrac{1}{4}$. El discriminante es $(\operatorname{tr} A)^2 - 4\det A = \tfrac{1}{4} - 1 = -\tfrac{3}{4} < 0$, luego los autovalores son complejos:

$$\lambda_\pm = \frac{-\tfrac{1}{2} \pm i\sqrt{3}/2}{2} = -\frac{1}{4} \pm \frac{\sqrt{3}}{4}\, i = -0{,}25 \pm 0{,}433013\, i.$$

  Parte real negativa: **foco (espiral) asintóticamente estable**.

Paso 5: Hartman–Grobman. Los tres equilibrios son hiperbólicos —ninguno de los seis autovalores tiene parte real nula—, así que las tres conclusiones son válidas para el sistema no lineal, no solo para su linealización.

Paso 7: verificación numérica. Integrando con tolerancia 10^{-10} desde $(0{,}6,\ 0{,}4)$, la trayectoria describe espirales cada vez menores y en $t = 200$ está en $(0{,}500000,\ 0{,}500000)$: converge a E_3, como predice el foco estable. El periodo de giro predicho es $2\pi/0{,}433013 = 14{,}510395$ unidades de tiempo, y el módulo $|\lambda_\pm| = \frac{1}{2}$ dice que la amplitud se reduce por el factor $e^{-0{,}25\cdot 14{,}51} = e^{-3{,}628} \approx 0{,}0266$ en cada vuelta. La conclusión biológica: este modelo predice coexistencia estable con oscilaciones amortiguadas, no ciclos perpetuos.

Cuidado

[El centro es precisamente el caso que no se puede decidir] El modelo de Lotka–Volterra clásico, $x' = x - xy$, $y' = -y + xy$, tiene en $(1,1)$ la jacobiana $\left(\begin{smallmatrix} 0 & -1 \\ 1 & 0 \end{smallmatrix}\right)$, con autovalores $\pm i$: parte real nula, y Hartman–Grobman no aplica. Aquí resulta que sí es un centro, pero eso *no* se deduce de la linealización, sino de que existe una integral primera, $H(x,y) = x - \ln x + y - \ln y$, constante sobre las órbitas (la integración numérica en $[0, 50]$ desde $(1{,}5,\ 1)$ conserva $H = 2{,}0945348919$ con doce cifras). Que exista esa cantidad conservada es un accidente afortunado del modelo, no una regla.
Para ver que el peligro es real, considérese

$$x' = -y + a\,x(x^2+y^2), \qquad y' = x + a\,y(x^2+y^2).$$

Sea cual sea a, la jacobiana en el origen es siempre $\left(\begin{smallmatrix} 0 & -1 \\ 1 & 0 \end{smallmatrix}\right)$: la linealización es idéntica y dice "centro". Pero en polares el sistema es $r' = a\,r^3$, $\theta' = 1$, y la verdad depende del signo de a. Partiendo de $r(0) = 0{,}3$, la integración numérica en $t = 4$ da $r = 0{,}228748$ si $a = -1$ (foco estable), $r = 0{,}300000$ si $a = 0$ (centro genuino) y $r = 0{,}566947$ si $a = 1$ (foco inestable; de hecho la solución explota en tiempo finito $t = 1/(2r_0^2) = 5{,}5556$). Las tres cifras coinciden en seis decimales con la solución exacta $r(t) = r_0/\sqrt{1 - 2ar_0^2 t}$. Un mismo retrato lineal, tres destinos distintos: eso es lo que significa que la linealización no decida.

12.24. Ejercicios resueltos

Ejercicio 12.1. Resuelve $y' = \dfrac{x(1+y^2)}{y}$, $y(0) = 1$, e indica dónde vive la solución.

Solución. Separando: $\frac{y\,dy}{1+y^2} = x\,dx$; integrando, $\frac{1}{2}\ln(1+y^2) = \frac{x^2}{2} + C$. Con $y(0) = 1$: $C = \frac{1}{2}\ln 2$, luego $1 + y^2 = 2e^{x^2}$ y (rama positiva) $y = \sqrt{2e^{x^2} - 1}$, definida para todo x (el radicando es ≥ 1). ■

Ejercicio 12.2. Resuelve la ecuación de Bernoulli $y' - \dfrac{y}{x} = x\,y^2$ (para $x > 0$).

Solución. Aquí $n = 2$: con $u = y^{-1}$, $u' = -y^{-2}y'$, y dividiendo la ecuación por $-y^2$: $u' + \frac{u}{x} = -x$. Lineal con $\mu = x$: $(xu)' = -x^2$, $xu = -\frac{x^3}{3} + C$, $u = \frac{C}{x} - \frac{x^2}{3}$. Entonces

$$y = \frac{1}{u} = \frac{3x}{3C - x^3}.$$

(También está la solución perdida $y \equiv 0$, aniquilada al dividir por y^2.) ■

Ejercicio 12.3. Resuelve $(2xy + 3y^2)\,\mathrm{d}x + (x^2 + 6xy - 2y)\,\mathrm{d}y = 0$.

Solución. $M_y = 2x + 6y = N_x$: exacta en $\mathbb{R}^2$. Integramos $F = \int M\,\mathrm{d}x = x^2y + 3xy^2 + g(y)$; de $F_y = x^2 + 6xy + g'(y) = x^2 + 6xy - 2y$: $g' = -2y$, $g = -y^2$. Solución implícita:

$$x^2y + 3xy^2 - y^2 = C.$$

■

Ejercicio 12.4. Resuelve $y'' + 4y' + 13y = 0$ con $y(0) = 1$, $y'(0) = 1$, y describe el movimiento.

Solución. Característica: $r^2 + 4r + 13 = 0$, $r = -2 \pm 3i$. Solución general: $y = e^{-2t}(C_1 \cos 3t + C_2 \sin 3t)$. Condiciones: $C_1 = 1$; $y' = e^{-2t}[(-2C_1 + 3C_2)\cos 3t + \cdots]$ da $-2 + 3C_2 = 1$, $C_2 = 1$. Así

$$y = e^{-2t}(\cos 3t + \sin 3t) = \sqrt{2}\, e^{-2t} \cos\left(3t - \tfrac{\pi}{4}\right):$$

oscilación subamortiguada de frecuencia 3 cuya amplitud decae como e^{-2t}. ■

Ejercicio 12.5. Resuelve $y'' + y = \tan t$ (donde coeficientes indeterminados no sirve) por variación de parámetros.

Solución. Sistema fundamental $y_1 = \cos t$, $y_2 = \sin t$, $W = \cos^2 t + \sin^2 t = 1$. Entonces

$$y_p = -\cos t \int \sin t \tan t\,\mathrm{d}t + \sin t \int \cos t \tan t\,\mathrm{d}t.$$

La segunda integral: $\int \sin t\,\mathrm{d}t = -\cos t$. La primera: $\int \frac{\sin^2 t}{\cos t}\mathrm{d}t = \int \frac{1-\cos^2 t}{\cos t}\mathrm{d}t = \ln|\sec t + \tan t| - \sin t$. Por tanto

$$y_p = -\cos t\, \ln|\sec t + \tan t| + \cos t \sin t - \sin t \cos t = -\cos t\, \ln|\sec t + \tan t|,$$

y la general es $y = C_1 \cos t + C_2 \sin t - \cos t \ln|\sec t + \tan t|$. ■

Ejercicio 12.6. Resuelve el sistema $\mathbf{x}' = \begin{pmatrix} 1 & -2 \\ 2 & 1 \end{pmatrix} \mathbf{x}$ y describe su retrato de fase.

Solución. $p(\lambda) = \lambda^2 - 2\lambda + 5$: $\lambda = 1 \pm 2i$. Para $\lambda = 1 + 2i$: $(A - \lambda I)v = 0$ con primera fila $(-2i, -2)$: $v = (1, -i) = \underbrace{(1,0)}_{\mathbf{a}} + i\,\underbrace{(0,-1)}_{\mathbf{b}}$. Soluciones reales:

$$\mathbf{x}_1 = e^t\big(\cos 2t\,(1,0) - \sin 2t\,(0,-1)\big) = e^t(\cos 2t,\ \sin 2t), \qquad \mathbf{x}_2 = e^t(\sin 2t,\ -\cos 2t),$$

y la general es combinación de ambas. Parte real $\alpha = 1 > 0$ con parte imaginaria no nula: **espiral inestable** — las trayectorias giran (a frecuencia 2) mientras se alejan exponencialmente. ■

Ejercicio 12.7. Un circuito se enciende en $t = 1$: resuelve con Laplace $y'' + y = u_1(t)$, $y(0) = y'(0) = 0$, donde u_1 es el escalón en $t = 1$.

Solución. Transformando: $(s^2 + 1)Y = \frac{e^{-s}}{s}$, luego $Y = e^{-s} \cdot \frac{1}{s(s^2+1)}$. Fracciones parciales: $\frac{1}{s(s^2+1)} = \frac{1}{s} - \frac{s}{s^2+1}$, cuya antitransformada es $1 - \cos t$. Por la regla del desplazamiento temporal:

$$y(t) = u_1(t)\big(1 - \cos(t-1)\big):$$

nada ocurre hasta $t=1$; después, el sistema oscila alrededor del nuevo equilibrio $y=1$. Las funciones escalón + Laplace manejan discontinuidades sin partir el problema en trozos. ■

Ejercicio 12.8. Resuelve por series $y''-xy=0$ (ecuación de Airy) alrededor de $x=0$, dando los primeros términos de las dos soluciones básicas.

Solución. Con $y=\sum c_n x^n$: $y''=\sum_{n\ge 0}(n+2)(n+1)c_{n+2}x^n$ y $xy=\sum_{n\ge 1}c_{n-1}x^n$. Igualando coeficientes: $2c_2=0$ y $(n+2)(n+1)c_{n+2}=c_{n-1}$ para $n\ge 1$:

$$c_{n+2}=\frac{c_{n-1}}{(n+1)(n+2)}.$$

Los coeficientes se encadenan de 3 en 3. Con $c_0=1, c_1=0$: $y_1=1+\frac{x^3}{6}+\frac{x^6}{180}+\cdots$; con $c_0=0, c_1=1$: $y_2=x+\frac{x^4}{12}+\frac{x^7}{504}+\cdots$. La solución general es $C_1y_1+C_2y_2$ (radio de convergencia infinito). Las funciones de Airy —combinaciones especiales de estas— describen la difracción de la luz cerca de una cáustica y el arcoíris. ■

Ejercicio 12.9. Una barra de longitud π con extremos a 0° tiene perfil inicial $u(x,0)=\sin x+\frac{1}{2}\sin 3x$. Resuelve $u_t=u_{xx}$ y calcula cuánto tarda el segundo modo en reducirse al 1 % de su amplitud.

Solución. Por la fórmula general con $L=\pi$, $k=1$, y la condición inicial ya en la base de senos ($b_1=1$, $b_3=\frac{1}{2}$):

$$u(x,t)=e^{-t}\sin x+\tfrac{1}{2}e^{-9t}\sin 3x.$$

El modo $n=3$ decae como e^{-9t}: se reduce al 1 % cuando $e^{-9t}=0.01$, es decir $t=\frac{\ln 100}{9}\approx 0.51$. (El modo fundamental tardaría $\ln 100\approx 4.6$: nueve veces más. Tras un instante, toda solución del calor "parece" su primer modo.) ■

Ejercicio 12.10. Modelo depredador–presa de Lotka–Volterra: $\begin{cases} x'=x(3-y) \\ y'=y(x-2) \end{cases}$ (presas x, depredadores y). Halla los equilibrios y clasifícalos por linealización.

Solución. Equilibrios: $(0,0)$ y $(2,3)$. Jacobiana: $D\mathbf{F}=\begin{pmatrix} 3-y & -x \\ y & x-2 \end{pmatrix}$. En $(0,0)$: $\operatorname{diag}(3,-2)$: valores propios $3,-2$ de signos opuestos, **silla** (sin presas ni depredadores, cada población por separado crece o se extingue). En $(2,3)$: $\begin{pmatrix} 0 & -2 \\ 3 & 0 \end{pmatrix}$, con $\lambda^2+6=0$: $\lambda=\pm i\sqrt{6}$, imaginarios puros: **centro** en la linealización. (Aquí Hartman–Grobman no decide, pero para Lotka–Volterra existe una cantidad conservada, $H=x-2\ln x+y-3\ln y$... comprobemos: $\dot H=\dot x(1-\frac{2}{x})+\dot y(1-\frac{3}{y})=x(3-y)\frac{x-2}{x}+y(x-2)\frac{y-3}{y}=(x-2)(3-y)+(x-2)(y-3)=0$ ✓, así que las órbitas son curvas cerradas de verdad: poblaciones que oscilan cíclicamente, con período $\approx\frac{2\pi}{\sqrt{6}}$ cerca del equilibrio.) ■

12.25. Ejercicios adicionales resueltos

Ejercicio 12.11. Cosecha $y'=y(1-y)-h$: línea de fases y cosecha máxima.

Solución. Equilibrios $\frac{1\pm\sqrt{1-4h}}{2}$: silla–nodo en $h^*=\frac{1}{4}$; sobre él, extinción en tiempo finito. Operar en el umbral es colapso ante cualquier fluctuación: gestión pesquera real. ■

Ejercicio 12.12. e^{At} para $A = \begin{pmatrix} 4 & 1 \\ 0 & 4 \end{pmatrix}$ y solución con $\mathbf{x}(0) = (1, 1)$.

Solución. $e^{At} = e^{4t}(I + Nt)$: $\mathbf{x} = e^{4t}(1 + t,\, 1)$: el factor t de Jordan. ■

Ejercicio 12.13. $y'' + 2y' + y = e^{-t}$ (resonancia doble).

Solución. Raíz doble -1: $y_p = \frac{t^2}{2}e^{-t}$; general $(C_1 + C_2 t + \frac{t^2}{2})e^{-t}$. ■

Ejercicio 12.14. $y'' + 4y = \delta(t - \pi)$, condiciones nulas.

Solución. $Y = \frac{e^{-\pi s}}{s^2+4}$: $y = u_\pi(t)\frac{\sin 2(t-\pi)}{2} = u_\pi(t)\frac{\sin 2t}{2}$: nada hasta el golpe, luego oscilación libre; $h = \frac{\sin 2t}{2}$ es la respuesta a impulso que resuelve todo por convolución. ■

Ejercicio 12.15. Euler vs RK4 con $h = 0.1$ en $y' = y$, $y(0) = 1$.

Solución. Euler: 1.1 (error 5×10^{-3}); RK4: 1.1051708 (error 2×10^{-7}): cuatro órdenes por cuatro evaluaciones. ■

Ejercicio 12.16. Onda en la recta con $f = e^{-x^2}$, $g = 0$.

Solución. $u = \frac{e^{-(x-ct)^2} + e^{-(x+ct)^2}}{2}$: la campana se parte en dos mitades viajeras. La onda transporta; el calor difumina. ■

Ejercicio 12.17. Sistema gradiente con $V = x^4 - x^2 + y^2$.

Solución. $\dot V = -\|\nabla V\|^2$: todo cae a un crítico. Mínimos $(\pm\frac{1}{\sqrt{2}}, 0)$, silla en el origen cuya variedad estable (el eje y) separa las dos cuencas: el doble pozo, modelo de biestabilidad (interruptores genéticos, memorias). ■

12.26. Taller de ejemplos explicados: resolver ecuaciones diferenciales

Una ecuación diferencial relaciona una función con su derivada. "Resolverla" es hallar la función. Veamos los tres tipos que más aparecen, cada uno con su receta.

12.26.1. Tipo 1: variables separables

Ejemplo 12.108. Resolver $\dfrac{\mathrm{d}y}{\mathrm{d}x} = xy$.

Paso 1. Separar las variables: juntamos todo lo de y a un lado y lo de x al otro:

$$\frac{\mathrm{d}y}{y} = x\,\mathrm{d}x.$$

Paso 2. Integrar ambos lados:

$$\int \frac{\mathrm{d}y}{y} = \int x\,\mathrm{d}x \implies \ln|y| = \frac{x^2}{2} + C.$$

Paso 3. Despejar y aplicando la exponencial:

$$y = e^{x^2/2 + C} = e^C\, e^{x^2/2} = A\, e^{x^2/2},$$

donde $A = e^C$ es una constante nueva. Esa es la solución general.

12.26.2. Tipo 2: lineal de primer orden (factor integrante)

Regla: ecuación lineal $y' + p(x)\,y = q(x)$

Se multiplica todo por el *factor integrante* $\mu = e^{\int p\,dx}$, que convierte el lado izquierdo en la derivada de un producto: $(\mu y)' = \mu q$. Luego se integra.

Ejemplo 12.109. Resolver $y' + 2y = e^x$.

Paso 1. Factor integrante. Aquí $p = 2$, así que $\mu = e^{\int 2\,dx} = e^{2x}$.

Paso 2. Multiplicar todo por μ: $e^{2x}y' + 2e^{2x}y = e^{2x}e^x = e^{3x}$. El lado izquierdo es, por diseño, la derivada de $e^{2x}y$:

$$\left(e^{2x}y\right)' = e^{3x}.$$

Paso 3. Integrar: $e^{2x}y = \displaystyle\int e^{3x}\,dx = \frac{e^{3x}}{3} + C$.

Paso 4. Despejar y dividiendo entre e^{2x}:

$$y = \frac{e^{3x}}{3e^{2x}} + \frac{C}{e^{2x}} = \frac{e^x}{3} + Ce^{-2x}.$$

12.26.3. Tipo 3: lineal de segundo orden con coeficientes constantes

Regla: ecuación $y'' + by' + cy = 0$

Se escribe la *ecuación característica* $r^2 + br + c = 0$ y se resuelve. Si tiene dos raíces reales r_1, r_2, la solución general es $y = C_1e^{r_1x} + C_2e^{r_2x}$.

Ejemplo 12.110. Resolver $y'' - 3y' + 2y = 0$.

Paso 1. Ecuación característica: cambiamos $y'' \to r^2$, $y' \to r$, $y \to 1$:

$$r^2 - 3r + 2 = 0.$$

Paso 2. Resolverla (factorizando): $(r-1)(r-2) = 0$, así que $r = 1$ y $r = 2$.

Paso 3. Escribir la solución general con esas dos raíces:

$$y = C_1e^x + C_2e^{2x}.$$

Las constantes C_1, C_2 se fijan si te dan condiciones iniciales (el valor de y y de y' en un punto).

12.27. Ampliación del temario: Bessel, transformada Z y métodos numéricos para EDP

12.27.1. La ecuación de Bessel y las funciones cilíndricas

Idea clave

Intuición primero. Esta ecuación es lo que aparece cuando separas variables en un tambor circular: $J_n(x)$ describe la *forma radial* de cada modo de vibración de la membrana. La diferencia con una cuerda es clave: los armónicos de una cuerda son senos, con frecuencias en proporción $1, 2, 3, \ldots$ (por eso "canta" una nota), mientras que las frecuencias del tambor son los *ceros* de J_n, espaciados de forma irregular. Por eso un tambor suena a golpe y no a nota afinada.

Al separar variables en problemas con simetría cilíndrica (una membrana circular, el calor en un tubo), aparece la *ecuación de Bessel*:

$$x^2 y'' + x\,y' + (x^2 - n^2)\,y = 0.$$

Es un punto singular regular, así que se resuelve por el método de Frobenius (serie $y = x^r \sum c_k x^k$). Sus soluciones son las *funciones de Bessel* $J_n(x)$, las "funciones cilíndricas":

$$J_n(x) = \sum_{k=0}^{\infty} \frac{(-1)^k}{k!\,(k+n)!} \Big(\frac{x}{2}\Big)^{2k+n}.$$

Observación 12.111. $J_0(0) = 1$ y $J_n(0) = 0$ para $n \geq 1$; las J_n oscilan como senos amortiguados, pero con ceros *no* periódicos. Esos ceros dan las frecuencias de vibración de un tambor: por eso el sonido de un tambor no es armónico (sus sobretonos no son múltiplos enteros del fundamental), a diferencia de una cuerda. Junto con los polinomios de Legendre (simetría esférica) y de Hermite (oscilador cuántico), las Bessel forman el "zoológico de funciones especiales" que resuelve las EDP de la física en cada geometría.

12.27.2. La transformada Z: Laplace para sistemas discretos

Idea clave

Intuición primero. Es la transformada de Laplace, pero para señales que avanzan a saltos (muestra a muestra) en vez de de forma continua. Donde Laplace convertía "derivar" en "multiplicar por s", la transformada Z convierte "retroceder un paso" ($x[n-1]$) en "multiplicar por z^{-1}". Con ese diccionario, una ecuación en diferencias (una receta recursiva) se vuelve álgebra que despejas para $X(z)$, igual que Laplace volvía álgebra una ecuación diferencial.

Así como la transformada de Laplace convierte ecuaciones diferenciales (tiempo continuo) en álgebra, la *transformada Z* hace lo mismo para *ecuaciones en diferencias* (tiempo discreto: señales muestreadas, filtros digitales, economía por periodos).

Regla: transformada Z

Para una sucesión $x[n]$ $(n = 0, 1, 2, \dots)$:

$$X(z) = \sum_{n=0}^{\infty} x[n]\, z^{-n}.$$

Convierte el *desplazamiento* $x[n-1]$ en multiplicar por z^{-1}, tal como Laplace convierte derivar en multiplicar por s.

Ejemplo 12.112. La sucesión $x[n] = a^n$ (crecimiento geométrico) tiene transformada

$$X(z) = \sum_{n=0}^{\infty} a^n z^{-n} = \sum_{n=0}^{\infty} \left(\frac{a}{z}\right)^n = \frac{1}{1 - a/z} = \frac{z}{z-a} \qquad (|z| > |a|),$$

usando la serie geométrica. El polo en $z = a$ contiene toda la información del comportamiento: $|a| < 1$ significa una señal que decae (sistema estable), $|a| > 1$ una que explota. Es el análogo discreto exacto de que los polos de Laplace en el semiplano izquierdo dan estabilidad. Con esta herramienta se diseñan los filtros digitales de todo el audio y video que consumes.

12.27.3. Métodos numéricos para EDP: diferencias finitas y elementos finitos

Idea clave

Intuición primero. Una computadora no puede guardar una función continua, así que la reemplaza por sus valores en una malla de puntos, y cada derivada por una diferencia entre puntos vecinos. Fíjate en el esquema del calor: el valor nuevo de cada punto es un *promedio* de sus vecinos, que es exactamente lo que hace el calor de verdad (difundir, suavizar). Los elementos finitos aplican la misma idea sobre triangulitos en vez de una malla regular, para poder tratar formas complicadas como el ala de un avión.

Casi ninguna EDP realista se resuelve con fórmula; se resuelve en la computadora. Los dos grandes métodos:

Diferencias finitas. Se reemplaza el continuo por una malla de puntos y las derivadas por cocientes. La segunda derivada, por ejemplo, se aproxima con la fórmula que sale de Taylor:

$$f''(x) \approx \frac{f(x+h) - 2f(x) + f(x-h)}{h^2}.$$

Ejemplo 12.113 (La ecuación del calor, explícita)**.** Para $u_t = \alpha u_{xx}$, discretizando el tiempo en pasos Δt y el espacio en pasos Δx, el esquema explícito actualiza cada punto como

$$u_i^{n+1} = u_i^n + \frac{\alpha\, \Delta t}{\Delta x^2}(u_{i+1}^n - 2u_i^n + u_{i-1}^n).$$

Cada punto nuevo es un promedio ponderado de sus vecinos: el calor se difunde numéricamente igual que en la realidad. (Hay una condición de estabilidad, $\frac{\alpha \Delta t}{\Delta x^2} \le \frac{1}{2}$, o la solución numérica explota.)

Elementos finitos (FEM). En geometrías complicadas (el ala de un avión, un hueso), en vez de una malla regular se divide el dominio en pequeños triángulos (o tetraedros), y la solución se aproxima por funciones simples (lineales) en cada elemento. En vez de la EDP directamente, se resuelve su *forma débil* (una versión integral que solo pide una derivada, obtenida multiplicando por una función de prueba e integrando por partes). El resultado es un enorme sistema lineal $K\mathbf{u} = \mathbf{f}$ que la computadora resuelve. El FEM es la herramienta con que se diseña prácticamente toda la ingeniería moderna: estructuras, aerodinámica, electromagnetismo, biomecánica.

12.28. Ejercicios propuestos

Sin solución desarrollada: solo enunciado y respuesta al final. Úsalos para comprobar que lo del capítulo se sostiene sin la muleta del texto delante.

12.1. Resuelve $y' = 3y$ con $y(0) = 2$.

12.2. Resuelve $y' + 2y = 0$.

12.3. Resuelve $y' = \dfrac{y}{t}$ para $t > 0$.

12.4. Resuelve $y'' - 5y' + 6y = 0$.

12.5. Resuelve $y'' + 4y = 0$.

12.6. Resuelve $y'' + 2y' + y = 0$ (raíz doble).

12.7. Resuelve $y' + y = e^{t}$.

12.8. Resuelve $y' = y^2$ con $y(0) = 1$ y di dónde deja de existir la solución.

12.9. Halla una solución particular de $y'' + 4y = \sin t$.

12.10. Calcula el wronskiano de e^{t} y e^{2t} y di qué concluyes.

12.11. Calcula la transformada de Laplace de $t\,e^{at}$.

12.12. Halla los equilibrios de $y' = y(1 - y)$ y clasifícalos por estabilidad.

Respuestas

1. $y = 2e^{3t}$. **2.** $y = Ce^{-2t}$. **3.** $y = Ct$. **4.** $y = C_1e^{2t} + C_2e^{3t}$. **5.** $y = C_1\cos 2t + C_2\sin 2t$. **6.** $y = (C_1 + C_2t)e^{-t}$. **7.** $y = Ce^{-t} + \frac{1}{2}e^{t}$. **8.** $y = \dfrac{1}{1-t}$: explota en $t = 1$, aunque la ecuación no tenga nada singular ahí (explosión en tiempo finito). **9.** $y_p = \frac{1}{3}\sin t$. **10.** $W = e^{3t} \neq 0$: son linealmente independientes y forman base del espacio de soluciones. **11.** $\dfrac{1}{(s-a)^2}$. **12.** $y = 0$ inestable, $y = 1$ asintóticamente estable.

Capítulo 13

Ecuaciones en derivadas parciales

Lo que necesitas traer

Del capítulo 9: derivadas parciales y la regla de la cadena en varias variables. Del capítulo 11: divergencia, laplaciano y el teorema de la divergencia. Del capítulo 12: ecuaciones diferenciales ordinarias, porque la separación de variables las produce.

Una ecuación diferencial ordinaria describe algo que depende de una sola variable: la posición de un móvil a lo largo del tiempo, la carga de un condensador, la población de una especie. Pero casi todo lo que ocurre en el mundo depende a la vez del tiempo y del espacio: la temperatura de una barra en cada punto y en cada instante, la altura de una cuerda vibrante, la presión del aire alrededor de un ala. Describir eso exige derivadas respecto de varias variables a la vez, y las ecuaciones que resultan son las ecuaciones en derivadas parciales.

El salto de dificultad es enorme y conviene decirlo desde el principio. Para las ecuaciones ordinarias hay un teorema general de existencia y unicidad, el de Picard, que cubre prácticamente todos los casos razonables. Para las ecuaciones en derivadas parciales *no existe* ningún teorema análogo: cada familia se estudia por separado, con técnicas propias, y hay ecuaciones de aspecto inofensivo cuya resolución sigue abierta. Este capítulo se concentra en las tres ecuaciones clásicas (ondas, calor y Laplace), que son a la vez las más importantes de la física y los tres representantes de los tres comportamientos posibles.

13.1. Qué es una EDP y por qué hay tres tipos

Definición 13.1. Una *ecuación en derivadas parciales* es una relación entre una función incógnita u de varias variables y sus derivadas parciales. Su *orden* es el de la derivada más alta que aparece, y es *lineal* si u y sus derivadas aparecen solo con exponente uno y sin multiplicarse entre sí. Para una ecuación lineal de segundo orden en dos variables,

$$A\,u_{xx} + B\,u_{xy} + C\,u_{yy} + (\text{términos de orden menor}) = 0,$$

se define el *discriminante* $\Delta = B^2 - 4AC$, y la ecuación se llama *hiperbólica* si $\Delta > 0$, *parabólica* si $\Delta = 0$ y *elíptica* si $\Delta < 0$.

Idea clave

Qué dice esto en cristiano. La clasificación no es una manía taxonómica: los tres tipos se comportan de maneras radicalmente distintas y hay que tratarlos con herramientas distintas. Las hiperbólicas *transportan* información a velocidad finita y conservan las irregularidades: una esquina en el dato inicial sigue siendo una esquina para siempre. Las parabólicas *difunden*: suavizan todo instantáneamente y son irreversibles, como el calor que se reparte y no vuelve. Y las elípticas describen *equilibrio*, no evolución: no hay tiempo, y el valor en cada punto es el promedio de sus vecinos. Saber de qué tipo es una ecuación dice de antemano qué esperar de sus soluciones, y hasta qué datos hay que dar para que el problema tenga sentido.

Regla: clasificar una EDP y decidir qué esperar de ella

Ante una ecuación nueva conviene contestar estas preguntas en orden antes de intentar resolverla:

1. **Anota el orden** y comprueba si es lineal; si no lo es, casi todo lo de este capítulo deja de aplicarse y hay que buscar técnicas específicas.
2. **Identifica A, B y C** como los coeficientes de u_{xx}, u_{xy} y u_{yy}, con el signo que tengan al pasar todo a un lado.
3. **Calcula $\Delta = B^2 - 4AC$** y clasifica; si los coeficientes dependen del punto, la ecuación puede cambiar de tipo de una región a otra.
4. **Deduce el comportamiento esperado**: propagación con velocidad finita, difusión instantánea o equilibrio sin tiempo.
5. **Decide qué datos hacen falta.** Hiperbólica de segundo orden: posición y velocidad iniciales. Parabólica: solo el dato inicial. Elíptica: valores en toda la frontera y ningún dato inicial.
6. **Comprueba que el problema está bien planteado**, es decir, que hay solución, que es única y que depende de forma continua de los datos.
7. **Elige la técnica**: características para las hiperbólicas de primer orden, separación de variables en dominios acotados, transformada de Fourier en dominios infinitos.

Ejemplo 13.2 (Las tres ecuaciones clásicas, clasificadas)**. (a) La ecuación de ondas.** $u_{tt} - c^2 u_{xx} = 0$ tiene $A = 1$, $B = 0$ y $C = -c^2$, luego $\Delta = 4c^2 > 0$: hiperbólica. Y en efecto describe una perturbación que viaja a velocidad c sin deformarse.

(b) La ecuación del calor. $u_t - k\,u_{xx} = 0$ no tiene término u_{tt}, así que $A = -k$ visto como coeficiente de u_{xx}, $B = 0$ y $C = 0$: $\Delta = 0$, parabólica. La ausencia de u_{tt} es exactamente lo que rompe la simetría entre pasado y futuro.

(c) La ecuación de Laplace. $u_{xx} + u_{yy} = 0$ tiene $A = C = 1$ y $B = 0$, luego $\Delta = -4 < 0$: elíptica. Aquí ninguna de las dos variables hace de tiempo, y por eso no tiene sentido dar un dato inicial: hay que dar toda la frontera.

(d) Una que cambia de tipo. La ecuación de Tricomi, $y\,u_{xx}+u_{yy}=0$, tiene $\Delta=-4y$: es elíptica para $y>0$, parabólica en $y=0$ e hiperbólica para $y<0$. No es una rareza académica: modela el flujo de aire alrededor de un ala en régimen transónico, donde la región subsónica es elíptica y la supersónica hiperbólica, y la línea sónica las separa.

Cuidado

[No existe un teorema general de existencia] Para las ecuaciones ordinarias, Picard cubre casi todo. Para las parciales no hay nada parecido, y el teorema de Cauchy–Kovalévskaya, que es lo más cercano, exige analiticidad y da solo existencia local. Peor aún: Hans Lewy construyó en 1957 una ecuación lineal de primer orden con coeficientes C^∞ que *no tiene ninguna solución* en ningún entorno, ni siquiera en el sentido generalizado. Segundo aviso: la clasificación es local, y el apartado (d) muestra que una misma ecuación puede ser de tipos distintos en regiones distintas del mismo problema.

13.2. El alfabeto de una ecuación en derivadas parciales: la notación de subíndices, el orden, la linealidad, qué es una solución y qué datos hay que dar

La sección anterior ha dicho qué es una ecuación en derivadas parciales y ha repartido las de segundo orden en tres cajones. Pero para leer con soltura el resto del capítulo hace falta antes el vocabulario básico, y conviene no darlo por sabido: la notación u_x, u_{xy}, u_{tt} aparece en cada página del tema y significa algo muy concreto; las palabras *lineal*, *cuasilineal* y *homogénea* deciden qué técnicas se pueden usar; y la palabra *solución*, que parece obvia, tiene aquí un contenido preciso. Esta sección lo monta todo desde cero, con un ejemplo resuelto de cada cosa, y termina con dos recetas: una para clasificar cualquier ecuación que a uno le pongan delante y otra para resolver por separación de variables, aplicada entera a un problema de calor.

La notación de subíndices

Definición 13.3 (Notación de subíndices para las derivadas parciales)**.** Sea u una función de varias variables. Se escribe como subíndice la variable respecto de la cual se deriva:

$$u_x=\frac{\partial u}{\partial x},\qquad u_y=\frac{\partial u}{\partial y},\qquad u_t=\frac{\partial u}{\partial t},$$

y los subíndices se acumulan de izquierda a derecha en el orden en que se deriva:

$$u_{xx}=\frac{\partial^2 u}{\partial x^2},\qquad u_{tt}=\frac{\partial^2 u}{\partial t^2},\qquad u_{xy}=\frac{\partial^2 u}{\partial y\,\partial x}=\frac{\partial}{\partial y}\left(\frac{\partial u}{\partial x}\right),$$

es decir, u_{xy} significa "deriva primero respecto de x y después respecto de y". La notación se prolonga sin cambios: u_{xxx}, u_{xxyy}, u_{xxt}. Si u tiene derivadas segundas continuas, el teorema de Schwarz del capítulo 9 garantiza que $u_{xy}=u_{yx}$, de modo que el orden de los subíndices es irrelevante; por eso una ecuación de segundo orden en dos variables tiene *tres* coeficientes de segundo orden y no cuatro.

Ejemplo 13.4 (Todas las derivadas de una función concreta)**.** Sea $u(x,y) = x^2y^3 + \sin(xy)$. Al derivar respecto de x, la variable y se trata como una constante, y al revés:

$$u_x = 2xy^3 + y\cos(xy), \qquad u_y = 3x^2y^2 + x\cos(xy).$$

Derivando otra vez,

$$u_{xx} = 2y^3 - y^2\sin(xy), \qquad u_{yy} = 6x^2y - x^2\sin(xy), \qquad u_{xy} = 6xy^2 + \cos(xy) - xy\sin(xy),$$

y partiendo de u_y para derivar respecto de x sale exactamente esa misma u_{xy}, como anuncia Schwarz. En el punto $(1,2)$: $u_x = 16 + 2\cos 2 = 15{,}167706$, $u_y = 12 + \cos 2 = 11{,}583853$, $u_{xx} = 16 - 4\sin 2 = 12{,}362810$, $u_{xy} = 24 + \cos 2 - 2\sin 2 = 21{,}765258$ y $u_{yy} = 12 - \sin 2 = 11{,}090703$.

Cuidado

[Tres lecturas erróneas del subíndice] Primera: u_x *no* es un producto de u por x; es una derivada, y el subíndice sólo dice respecto de qué variable. Segunda: u_{xy} *no* es $u_x \cdot u_y$; es una derivada segunda mixta, y confundirlas cambia el orden de la ecuación. Y tercera: u_x^2 significa $(u_x)^2$, el cuadrado de la derivada primera, mientras que u_{xx} es la derivada segunda; son objetos distintos, y la diferencia entre una ecuación lineal y una que no lo es puede estar exactamente ahí.

Ecuación ordinaria, ecuación en derivadas parciales, orden y solución

Definición 13.5 (EDO, EDP, orden y solución)**.** Una *ecuación diferencial ordinaria* (EDO) relaciona una función incógnita de *una sola* variable con sus derivadas. Una *ecuación en derivadas parciales* (EDP) relaciona una función incógnita de *dos o más* variables independientes con sus derivadas parciales. El *orden* de una u otra es el de la derivada más alta que aparece efectivamente.

Se llama *solución* de una EDP en una región Ω a toda función u definida en Ω, con tantas derivadas continuas como el orden de la ecuación, que al sustituirla convierte la ecuación en una identidad válida en todo punto de Ω. Cuando se quiera subrayar esa regularidad se dirá *solución clásica*, para distinguirla de las soluciones débiles que aparecen más adelante en el capítulo.

Ejemplo 13.6 (Seis ecuaciones: cuáles son EDP y de qué orden)**.** (a) $y'' + 4y = 0$, con $y = y(t)$: una sola variable independiente, luego EDO, de orden dos.

(b) $u_t = k\,u_{xx}$: la incógnita $u(x,t)$ depende de dos variables, luego EDP; el orden es dos.

(c) $u_t + c\,u_x = 0$: EDP de orden uno. Tiene dos variables, pero ninguna derivada segunda.

(d) $u_{xx} + u_{yy} + u_{zz} = 0$: EDP de orden dos en tres variables.

(e) $u_{tt} + u_{xxxx} = 0$, la ecuación de la viga vibrante: EDP de orden cuatro.

(f) $u_x + u_y = u^5$: EDP de orden uno. La potencia quinta afecta a la incógnita, no a las derivadas, así que no cambia el orden; sí cambiará, como veremos, la linealidad.

Ejemplo 13.7 (Comprobar que una función es solución)**.** Comprobar es sustituir; no hace falta saber de dónde salió la fórmula.

(a) $u = \mathrm{e}^{-t}\sin x$ **y el calor** $u_t = u_{xx}$**.** Se tiene $u_t = -\mathrm{e}^{-t}\sin x$ y $u_{xx} = -\mathrm{e}^{-t}\sin x$: coinciden, luego es solución. Y cumple $u(0,t) = u(\pi,t) = 0$, así que sirve para una barra $[0,\pi]$ con los extremos a cero.

(b) $u = \mathrm{e}^{-4t}\sin 2x$ **y la misma ecuación.** Ahora $u_t = -4\mathrm{e}^{-4t}\sin 2x = u_{xx}$: también lo es. Al doblar la frecuencia espacial el ritmo de decaimiento se ha multiplicado por cuatro.

(c) $u = \sin x\cos 3t$ **y las ondas** $u_{tt} = 9u_{xx}$**.** Aquí $u_{tt} = -9\sin x\cos 3t$ y $9u_{xx} = -9\sin x\cos 3t$. Es un *modo estacionario*: oscila en el tiempo sin desplazarse.

(d) $u = F(x-2t)$**, con** F **derivable cualquiera, y** $u_t + 2u_x = 0$**.** Por la regla de la cadena $u_t = -2F'(x-2t)$ y $u_x = F'(x-2t)$, luego se cumple sea cual sea F. Una EDP de primer orden tiene por solución general una *función* arbitraria, no una constante arbitraria: ahí está la diferencia esencial con las ordinarias.

(e) Una que no lo es. $u = x^2 + y^2$ no resuelve $u_{xx} + u_{yy} = 0$, porque su laplaciano vale $2 + 2 = 4 \neq 0$. En cambio $u = x^2 - y^2$ sí, con laplaciano $2 - 2 = 0$.

Lineal, semilineal, cuasilineal, no lineal, homogénea

Definición 13.8 (Los cuatro grados de linealidad, y la homogeneidad)**.** Una EDP de orden m es:

- **Lineal** si se puede escribir como $L[u] = f$, donde L combina u y sus derivadas con coeficientes que dependen sólo de las variables independientes, nunca de u. En una lineal, u y sus derivadas aparecen con exponente uno y sin multiplicarse entre sí.
- **Semilineal** si las derivadas de orden máximo aparecen linealmente y con coeficientes que dependen sólo de las variables independientes, pero los términos de orden menor pueden ser cualquier cosa.
- **Cuasilineal** si las derivadas de orden máximo aparecen linealmente pero sus coeficientes dependen de u o de derivadas de orden menor.
- **No lineal** (o totalmente no lineal) si ni siquiera las derivadas de orden máximo aparecen linealmente.

Una ecuación lineal $L[u] = f$ se llama *homogénea* si $f \equiv 0$, y *no homogénea* en caso contrario. Cada clase contiene a la anterior: toda lineal es semilineal y toda semilineal es cuasilineal.

Ejemplo 13.9 (Seis ecuaciones clasificadas por linealidad)**.** (a) $u_t - k\,u_{xx} = 0$: *lineal homogénea*; los coeficientes son 1 y $-k$, constantes.

(b) $u_{tt} - c^2 u_{xx} = \sin(\omega t)$: *lineal no homogénea*. El segundo miembro no depende de u, así que la linealidad no se pierde; lo que se pierde es la homogeneidad, y con ella el que la suma de dos soluciones vuelva a serlo.

(c) $u_t - u_{xx} = u^2$: *semilineal*. La derivada máxima entra con coeficiente constante; el estropicio está en un término de orden cero.

(d) $u_t = (1+u^2)u_{xx}$: *cuasilineal*. La derivada máxima entra linealmente, pero su coeficiente depende de u: la difusividad cambia según lo caliente que esté el material.

(e) $u_t + u\,u_x = 0$, la ecuación de Burgers: *cuasilineal* de primer orden, con el coeficiente de u_x igual a la propia u.

(f) $u_x^2 + u_y^2 = 1$, la ecuación eikonal de la óptica: *no lineal*. Las derivadas de orden máximo aparecen al cuadrado.

Idea clave

Por qué se hacen estas distinciones. Porque marcan hasta dónde llegan las herramientas. Con una lineal homogénea se pueden sumar soluciones, y de ahí salen la separación de variables y las series de Fourier, que son media asignatura. Con una lineal no homogénea todavía se puede, restando primero una solución particular. Con una semilineal o una cuasilineal ya no: se conservan las características, porque las derivadas de orden máximo siguen entrando linealmente, pero se pierde la superposición, y aparecen choques y tiempos de rotura. Y con una totalmente no lineal se pierde hasta la noción sencilla de característica. Cada peldaño que se baja cuesta una técnica.

Teorema 13.10 (Principio de superposición)**.** *Sea L un operador lineal, es decir, con $L[\alpha u + \beta v] = \alpha L[u] + \beta L[v]$ para todas las constantes α, β y todas las funciones admisibles u, v. Si $u_1, \dots, u_N$ resuelven la ecuación homogénea $L[u] = 0$, entonces cualquier combinación lineal $u = \sum_{j=1}^{N} c_j\, u_j$ también la resuelve. Y si u_p resuelve la no homogénea $L[u] = f$, todas las soluciones de $L[u] = f$ son de la forma $u_p + w$ con $L[w] = 0$.*

Demostración. La linealidad aplicada repetidamente da $L[\sum_j c_j u_j] = \sum_j c_j L[u_j] = \sum_j c_j \cdot 0 = 0$. Para lo segundo, si $L[u] = f$ y $L[u_p] = f$, entonces $w = u - u_p$ cumple $L[w] = f - f = 0$, luego $u = u_p + w$; y recíprocamente $L[u_p + w] = f + 0 = f$. □

Observación 13.11 (La versión infinita necesita un permiso)*.* En la práctica se superponen *infinitas* soluciones, $u = \sum_{n \geq 1} c_n u_n$, y ahí el teorema no basta: hay que poder derivar la serie término a término, lo cual exige convergencia uniforme de la serie derivada. Para el calor el factor $\mathrm{e}^{-n^2 t}$ lo garantiza con enorme holgura en cuanto $t > 0$; para las ondas hay que ser más cuidadoso. Es el único punto delicado del método que cierra la sección.

Ejemplo 13.12 (La superposición en acción)**.** Como $\mathrm{e}^{-t}\sin x$ y $\mathrm{e}^{-9t}\sin 3x$ resuelven ambas $u_t = u_{xx}$ y ambas se anulan en $x = 0$ y en $x = \pi$, también lo hace $u(x,t) = 3\,\mathrm{e}^{-t}\sin x - 2\,\mathrm{e}^{-9t}\sin 3x$, como se comprueba sustituyendo: el residuo $u_t - u_{xx}$ es idénticamente nulo. Su dato inicial es $3\sin x - 2\sin 3x$, que en $x = \frac{\pi}{2}$ vale $3 \cdot 1 - 2 \cdot (-1) = 5$. En $t = 0{,}5$ ese mismo punto vale 1,841810: el segundo modo, que decae nueve veces más deprisa, ya casi ha desaparecido.

El laplaciano

Definición 13.13 (Operador laplaciano)**.** El *laplaciano* de una función u es la suma de sus derivadas segundas puras respecto de cada variable espacial:

$$\Delta u = u_{xx} + u_{yy} \ \text{(en el plano)}, \qquad \Delta u = u_{xx} + u_{yy} + u_{zz} \ \text{(en el espacio)}.$$

Se escribe también $\nabla^2 u$, y equivale a $\operatorname{div}(\operatorname{grad} u)$: primero el gradiente, que es un campo vectorial, y después su divergencia, que vuelve a ser un número. Una función con $\Delta u = 0$ se llama *armónica*. En coordenadas polares, con $x = r\cos\theta$ e $y = r\sin\theta$, la misma operación se escribe

$$\Delta u = u_{rr} + \frac{1}{r}\,u_r + \frac{1}{r^2}\,u_{\theta\theta}.$$

Idea clave

Qué mide el laplaciano. Mide cuánto se separa el valor de u en un punto del promedio de u en los puntos de alrededor. Si $\Delta u(P) > 0$, el punto está por debajo de su vecindario —es un valle—, y si $\Delta u(P) < 0$ está por encima. Con eso a mano, las tres ecuaciones canónicas se leen solas: $u_t = k\Delta u$ dice "la temperatura sube donde el punto está más frío que sus vecinos"; $u_{tt} = c^2\Delta u$ dice "la aceleración apunta hacia el promedio de los vecinos", que es la ley de Hooke de una red de muelles; y $\Delta u = 0$ dice "cada punto vale ya el promedio de sus vecinos", que es el equilibrio.

Ejemplo 13.14 (Cuatro laplacianos calculados)**.** $u = x^2 - y^2$ da $\Delta u = 2 - 2 = 0$: armónica. $u = \mathrm{e}^x\cos y$ da $\Delta u = \mathrm{e}^x\cos y - \mathrm{e}^x\cos y = 0$: armónica. $u = \log r$ da, con la fórmula polar, $\Delta u = -\frac{1}{r^2} + \frac{1}{r}\cdot\frac{1}{r} = 0$: armónica en todo el plano salvo el origen, donde ni está definida. Y $u = x^2 + y^2$ da $\Delta u = 4 \neq 0$: no lo es, y en efecto tiene un mínimo interior, cosa que ninguna armónica puede permitirse.

Condiciones iniciales y condiciones de contorno

Definición 13.15 (Condición inicial y los tres tipos de condición de contorno)**.** Sea u definida en $\Omega \times [0, T)$, con Ω el dominio espacial y $\partial\Omega$ su frontera, y sea $\partial u/\partial n$ la derivada de u en la dirección normal exterior a $\partial\Omega$.

- Una *condición inicial* prescribe el estado en $t = 0$: $u(x, 0) = f(x)$ para todo $x \in \Omega$. Si la ecuación es de segundo orden en el tiempo hacen falta dos, la segunda $u_t(x, 0) = g(x)$.
- Una condición de *Dirichlet* prescribe el *valor* de la incógnita en la frontera: $u = g$ sobre $\partial\Omega$. Se llama *homogénea* si $g \equiv 0$.
- Una condición de *Neumann* prescribe el *flujo*, es decir la derivada normal: $\partial u/\partial n = h$ sobre $\partial\Omega$. La homogénea significa que nada entra ni sale: la frontera está aislada.
- Una condición de *Robin*, o de tercera especie, prescribe una combinación de ambas: $\alpha\,u + \beta\,\frac{\partial u}{\partial n} = h$ sobre $\partial\Omega$, con α y β no simultáneamente nulos.

Idea clave

Las tres, en una barra de metal. Dirichlet es sumergir el extremo en un baño térmico que lo mantiene a temperatura fija, pase lo que pase. Neumann es controlar cuánto calor se inyecta por ese extremo; el caso homogéneo es envolverlo en un aislante perfecto. Y Robin es lo que ocurre de verdad casi siempre: el extremo está al aire, y la ley de enfriamiento de Newton dice que el flujo que sale es proporcional a la diferencia entre

la temperatura del extremo y la del ambiente, $-\frac{\partial u}{\partial n} = \kappa\,(u - u_{\text{amb}})$, que reordenada es exactamente una condición de Robin.

Ejemplo 13.16 (Las tres condiciones, sobre una barra en equilibrio)**.** Una barra ocupa $[0, 1]$ y ha alcanzado el equilibrio, de modo que $u_t = 0$ y la ecuación del calor se reduce a $u'' = 0$: la temperatura es una recta, $u(x) = a + bx$. Se fija el extremo izquierdo con $u(0) = 10$ y se prueban tres cosas distintas en el derecho.

(a) Dirichlet: $u(1) = 4$**.** Entonces $a = 10$ y $a + b = 4$, luego $b = -6$ y $u(x) = 10 - 6x$.

(b) Neumann: $u'(1) = -3$**.** Como $u' = b$ en todo punto, sale $b = -3$ y $u(x) = 10 - 3x$, con $u(1) = 7$. Aquí no se ha impuesto ninguna temperatura en el extremo derecho, sino el ritmo al que sale el calor; la temperatura del extremo es una consecuencia.

(c) Robin: $u'(1) + 2(u(1) - 5) = 0$, es decir, un extremo al aire a 5 grados con coeficiente de intercambio 2. Sustituyendo $u = 10 + bx$ queda $b + 2(10 + b - 5) = 0$, o sea $3b = -10$ y $b = -\frac{10}{3}$:

$$u(x) = 10 - \tfrac{10}{3}x, \qquad u(1) = \tfrac{20}{3} = 6{,}666667.$$

Comprobación: $u'(1) = -3{,}333333$ y $2\,(6{,}666667 - 5) = 3{,}333333$, cuya suma es cero ✓. El extremo se ha quedado a 6,67 grados, entre los 5 del ambiente y los 10 de la fuente: ni pegado al ambiente (que sería Dirichlet con $u(1) = 5$) ni aislado.

Observación 13.17 (Qué datos pide cada tipo, y el problema bien planteado)*.* La combinación de datos no se elige al gusto: cada tipo de ecuación admite la suya. Las hiperbólicas de segundo orden piden dos condiciones iniciales más las de contorno; las parabólicas, una condición inicial más las de contorno; y las elípticas no admiten condición inicial ninguna, sino los valores en *toda* la frontera. Poner los datos equivocados no produce un problema difícil, sino un problema sin solución o con infinitas. El criterio preciso —existencia, unicidad y dependencia continua de los datos— es el de Hadamard y tiene sección propia más adelante en este capítulo; conviene leerla después de haber visto las tres ecuaciones en acción.

Las tres ecuaciones canónicas

Definición 13.18 (Calor, ondas y Laplace)**.** Con $k > 0$ y $c > 0$ constantes:

$$\text{calor: } u_t = k\,\Delta u, \qquad \text{ondas: } u_{tt} = c^2\,\Delta u, \qquad \text{Laplace: } \Delta u = 0.$$

La versión con segundo miembro de la última, $\Delta u = f$, se llama *ecuación de Poisson*. La constante k es la *difusividad*, con dimensiones de longitud al cuadrado partido por tiempo, y c es una *velocidad*.

Ejemplo 13.19 (Qué modela cada una)**. (a) El calor** describe la temperatura de una barra o una placa, la concentración de una sustancia que se difunde, la densidad de probabilidad de una partícula browniana y el precio de una opción financiera, que es la misma ecuación tras un cambio de variable: todos son procesos que reparten y suavizan.

(b) Las ondas describen la cuerda de un violín, la membrana de un tambor, el sonido en el aire y las ondas electromagnéticas en el vacío. La constante c es la velocidad de propagación: 340 metros por segundo en el aire, $3 \cdot 10^8$ para la luz.

(c) Laplace describe el potencial gravitatorio y el electrostático fuera de las masas y las cargas, la temperatura de equilibrio de una placa y el flujo de un fluido incompresible e irrotacional. Aquí no hay tiempo.

(d) Y cómo se relacionan. Si en la del calor o en la de ondas se busca una solución que no dependa del tiempo, los miembros izquierdos se anulan y ambas se convierten en $\Delta u = 0$: la ecuación de Laplace es el estado estacionario de las otras dos, y ese es el motivo de que aparezca en todas partes.

Dos recetas

Regla: clasificar una EDP cualquiera: orden, linealidad, homogeneidad y tipo

Ante una ecuación que uno no ha visto nunca, conviene contestar estas seis preguntas en este orden. Se usa aquí el convenio de la sección anterior, con el término cruzado escrito $B\,u_{xy}$ —sin el factor 2— y discriminante $\Delta = B^2 - 4AC$. Más adelante, en la sección de formas canónicas, se adopta el otro convenio habitual, con $2B\,u_{xy}$ y discriminante $B^2 - AC$, y allí hay una advertencia que explica por qué los dos clasifican exactamente igual. Lo inaceptable es mezclarlos sin darse cuenta.

1. **Cuenta las variables independientes.** Si hay una sola es una EDO y este capítulo no es el sitio; si hay dos o más, es una EDP.
2. **Anota el orden**: la derivada más alta que aparece de verdad.
3. **Decide la linealidad** con la definición 13.8: mira primero si los coeficientes de las derivadas de orden máximo dependen de u —entonces es cuasilineal como mucho— y después si los términos de orden menor son lineales en u.
4. **Si es lineal, mira si es homogénea**: pasa todo lo que lleve u a la izquierda y comprueba si a la derecha queda cero. Sólo entonces podrás superponer soluciones.
5. **Si es de segundo orden en dos variables, identifica A, B y C** como los coeficientes de u_{xx}, u_{xy} y u_{yy}, con el signo que tengan al pasar todo a un miembro, y calcula $\Delta = B^2 - 4AC$: hiperbólica si $\Delta > 0$, parabólica si $\Delta = 0$, elíptica si $\Delta < 0$. Si los coeficientes dependen del punto, la respuesta puede cambiar de una región a otra.
6. **Deduce qué datos hacen falta** y qué comportamiento esperar: propagación a velocidad finita, difusión instantánea o equilibrio sin tiempo.

Ejemplo 13.20 (La receta aplicada entera a tres ecuaciones)**. (a)** $2u_{xx} - 3u_{xy} + u_{yy} + 5u_x - u = \mathrm{e}^x$**.** Dos variables independientes, luego EDP; orden dos; coeficientes constantes y u y sus derivadas con exponente uno, luego lineal; el segundo miembro no es cero y no depende de u, luego *lineal no homogénea*. Con $A = 2$, $B = -3$ y $C = 1$,

$$\Delta = (-3)^2 - 4 \cdot 2 \cdot 1 = 9 - 8 = 1 > 0 :$$

hiperbólica en todo el plano. Habrá dos familias de características y la información viajará por ellas; para resolverla convendría reducirla a la forma canónica $U_{\xi\eta} = \cdots$, como se hace

más adelante en el capítulo.

(b) $u_t = (1+u^2)u_{xx}$**.** EDP de orden dos. No es lineal ni semilineal, porque el coeficiente de la derivada de orden máximo depende de u: es *cuasilineal*. La pregunta sobre la homogeneidad ni se plantea, porque esa palabra sólo tiene sentido para ecuaciones lineales. Para el tipo, en las variables (x,t) el coeficiente de u_{xx} es $A = 1+u^2$, el de u_{xt} es $B=0$ y el de u_{tt} es $C=0$, luego $\Delta = 0$: parabólica. Difunde como el calor, pero al no ser lineal la separación de variables está descartada de entrada.

(c) $u_{xx} - 6u_{xy} + 9u_{yy} = 0$**.** Lineal homogénea de orden dos, con $A=1$, $B=-6$, $C=9$ y $\Delta = 36-36=0$: parabólica. El discriminante nulo delata que el operador es un cuadrado perfecto, $(\partial_x - 3\partial_y)^2$, y en efecto la solución general resulta ser $u = F(y+3x) + x\,G(y+3x)$: una sola familia de características, $y+3x = \text{cte}$, y un factor x acompañando a la segunda función arbitraria, igual que la raíz doble de una ecuación ordinaria de coeficientes constantes.

Regla: resolver un problema de contorno por separación de variables, paso a paso

El método sirve para ecuaciones *lineales y homogéneas*, en dominios acotados de geometría sencilla y con condiciones de contorno *homogéneas*. Son ocho pasos:

1. **Comprueba las tres hipótesis**: ecuación lineal, ecuación homogénea, contorno homogéneo. Si el contorno no lo es, resta primero la solución estacionaria que lo absorba.
2. **Ensaya** $u(x,t) = X(x)\,T(t)$ y sustituye en la ecuación.
3. **Separa**: deja a un lado lo que depende de x y al otro lo que depende de t. Como son variables independientes, ambos miembros han de valer una misma constante, que se escribe $-\lambda$ por comodidad de signos.
4. **Traduce las condiciones de contorno a** X: de $u(0,t)=0$ para todo t, con $T \not\equiv 0$, sale $X(0)=0$, y análogamente en el otro extremo.
5. **Resuelve el problema de autovalores** $X'' + \lambda X = 0$ con esas condiciones, *discutiendo los tres casos* $\lambda<0$, $\lambda=0$ y $\lambda>0$. Sólo ciertos λ dan soluciones no nulas: son los autovalores, y las X correspondientes, las autofunciones.
6. **Resuelve la parte temporal** para cada autovalor: sale una exponencial en el caso del calor y una combinación de seno y coseno en el de ondas.
7. **Superpone**, con un coeficiente libre por cada modo, apoyándote en el teorema 13.10.
8. **Ajusta los coeficientes al dato inicial** desarrollándolo en serie de Fourier, y **comprueba** el resultado: que cada término resuelva la ecuación, que se cumplan las condiciones de contorno y que la serie reproduzca el dato en $t=0$.

Ejemplo 13.21 (Un problema de calor con Dirichlet homogéneo, resuelto entero)**.** Resolvamos

$$u_t = u_{xx} \ \text{ en } 0<x<\pi,\ t>0, \qquad u(0,t)=u(\pi,t)=0, \qquad u(x,0)=x(\pi-x).$$

Es una barra de longitud π y difusividad 1, con los dos extremos sumergidos en hielo y con

un perfil inicial en forma de parábola, nulo en los extremos y máximo en el centro, donde vale $\frac{\pi^2}{4} = 2{,}467401$.

Paso 1. La ecuación es lineal y homogénea y el contorno es Dirichlet homogéneo: las tres hipótesis se cumplen.

Pasos 2 y 3. Con $u = X(x)T(t)$ se tiene $u_t = XT'$ y $u_{xx} = X''T$, luego $XT' = X''T$ y, dividiendo entre XT donde no se anule,

$$\frac{T'}{T} = \frac{X''}{X} = -\lambda,$$

con λ constante, porque el miembro izquierdo no depende de x y el derecho no depende de t.

Paso 4. De $u(0,t) = X(0)T(t) = 0$ para todo t, y como $T \not\equiv 0$ (si no, $u \equiv 0$), sale $X(0) = 0$; igualmente $X(\pi) = 0$.

Paso 5: los tres casos. Hay que resolver $X'' + \lambda X = 0$ con $X(0) = X(\pi) = 0$.

- Si $\lambda < 0$, escríbase $\lambda = -\mu^2$ con $\mu > 0$; la solución general es $X = A\mathrm{e}^{\mu x} + B\mathrm{e}^{-\mu x}$. De $X(0) = 0$ sale $B = -A$, y de $X(\pi) = 0$ sale $A(\mathrm{e}^{\mu\pi} - \mathrm{e}^{-\mu\pi}) = 0$; como el paréntesis no se anula para $\mu > 0$, queda $A = 0$ y $X \equiv 0$. No hay autovalores negativos.
- Si $\lambda = 0$, entonces $X'' = 0$ y $X = A + Bx$; de $X(0) = 0$ sale $A = 0$ y de $X(\pi) = 0$ sale $B = 0$. Tampoco.
- Si $\lambda > 0$, escríbase $\lambda = \mu^2$ con $\mu > 0$; entonces $X = A\cos(\mu x) + B\sin(\mu x)$. De $X(0) = 0$ sale $A = 0$, y de $X(\pi) = B\sin(\mu\pi) = 0$ con $B \neq 0$ sale $\sin(\mu\pi) = 0$, es decir $\mu = n$ entero positivo.

Los autovalores son, pues, $\lambda_n = n^2$ y las autofunciones $X_n(x) = \sin(nx)$, con $n = 1, 2, 3, \ldots$

Paso 6. Para cada n, la ecuación $T' = -n^2 T$ da $T_n(t) = \mathrm{e}^{-n^2 t}$ salvo constante.

Paso 7. Superponiendo, $u(x,t) = \sum_{n\geq 1} b_n\, \mathrm{e}^{-n^2 t}\sin(nx)$.

Paso 8: los coeficientes. En $t = 0$ hay que tener $\sum_n b_n \sin(nx) = x(\pi - x)$, y por la ortogonalidad de los senos en $[0, \pi]$,

$$b_n = \frac{2}{\pi}\int_0^\pi x(\pi - x)\sin(nx)\,\mathrm{d}x.$$

Integrando por partes con $v = -\frac{\cos(nx)}{n}$, el término de frontera se anula porque $\pi x - x^2$ vale 0 en los dos extremos, y queda $\frac{1}{n}\int_0^\pi (\pi - 2x)\cos(nx)\,\mathrm{d}x$. Una segunda integración por partes anula también el nuevo término de frontera, porque $\sin(n\pi) = \sin 0 = 0$, y deja

$$\frac{1}{n}\cdot\frac{2}{n}\int_0^\pi \sin(nx)\,\mathrm{d}x = \frac{2}{n^2}\cdot\frac{1-(-1)^n}{n} = \frac{2\big(1-(-1)^n\big)}{n^3},$$

de modo que

$$b_n = \frac{4\big(1-(-1)^n\big)}{\pi\, n^3} = \begin{cases} \dfrac{8}{\pi n^3} & n \text{ impar},\\[2mm] 0 & n \text{ par}. \end{cases}$$

Numéricamente $b_1 = 2{,}546479$, $b_3 = 0{,}094314$, $b_5 = 0{,}020372$ y $b_7 = 0{,}007424$, y la solución completa es

$$u(x,t) = \frac{8}{\pi}\sum_{n \text{ impar}} \frac{1}{n^3}\,\mathrm{e}^{-n^2 t}\sin(nx).$$

Comprobaciones. Cada término resuelve la ecuación, porque $\partial_t(\mathrm{e}^{-n^2t}\sin nx) = -n^2\mathrm{e}^{-n^2t}\sin nx$ coincide con su u_{xx}; cada término se anula en $x=0$ y en $x=\pi$; y en $t=0$ la serie reproduce la parábola, pues sumando cien términos impares en $x=\frac{\pi}{2}$ se obtiene 2,467402 frente al valor exacto $\frac{\pi^2}{4}=2{,}467401$. Hay además una comprobación independiente muy barata: como $u_t=u_{xx}$ y en $t=0$ se tiene $u_{xx}=-2$ en toda la barra, para tiempos pequeños debe cumplirse $u\approx\frac{\pi^2}{4}-2t$ en el centro; y en efecto la serie da 2,447401 en $t=0{,}01$ y 2,367401 en $t=0{,}05$, exactamente los valores de esa recta.

Qué dice la solución. En $x=\frac{\pi}{2}$ la temperatura vale 2,467401 en $t=0$, 2,267422 en $t=0{,}1$, 1,543470 en $t=0{,}5$ y 0,936786 en $t=1$. A partir de tiempos moderados sólo sobrevive el primer modo: en $t=0{,}5$ el término $n=3$ ya es 0,00068 veces el primero, y el primer término solo, 1,544518, difiere del total en menos de una milésima. La barra olvida su forma inicial y se convierte en una campana $\sin x$ que se aplana; y cuando $t\to\infty$ toda la solución tiende a cero, porque el hielo de los extremos acaba llevándose todo el calor.

Cuidado

[Los tres casos del paso 5 no se pueden saltar] Escribir directamente $X=A\cos(\mu x)+B\sin(\mu x)$ es suponer sin decirlo que $\lambda>0$, y con otras condiciones de contorno eso cuesta autovalores. Con Neumann homogéneo, $X'(0)=X'(\pi)=0$, el caso $\lambda=0$ *sí* da una autofunción, la constante $X_0=1$, que corresponde a la temperatura media de la barra y es justamente la que no decae nunca: en una barra aislada por los dos extremos el calor no se escapa y la solución tiende al promedio del dato inicial en vez de a cero. Y con ciertas condiciones de Robin los autovalores ya no son n^2, sino las raíces de una ecuación trascendente que hay que localizar numéricamente. El esqueleto del método es el mismo; lo que cambia es el problema de autovalores del paso 5.

13.3. Transporte y el método de las características

Teorema 13.22 (Ecuación de transporte)**.** *La solución de $u_t+c\,u_x=0$ con $u(x,0)=f(x)$ es*

$$u(x,t)=f(x-ct),$$

es decir, el perfil inicial se traslada rígidamente con velocidad c sin cambiar de forma.

Idea clave

Qué dice esto en cristiano. Las líneas $x-ct=$ constante se llaman *características*, y a lo largo de ellas la solución no cambia. La ecuación dice justamente eso: la derivada de u en la dirección $(1,c)$ del plano (t,x) es nula. Resolver la ecuación se reduce entonces a un problema geométrico: para saber cuánto vale u en un punto, se retrocede por la característica que pasa por él hasta llegar al instante inicial, y se lee ahí el dato. Toda la información viaja por esas líneas y no se mezcla.

Ejemplo 13.23 (Seguir las características)**. (a) Un pulso que viaja.** Con $c=2$ y $f(x)=\mathrm{e}^{-x^2}$, la solución es $u(x,t)=\mathrm{e}^{-(x-2t)^2}$. En $t=3$ el pico, que estaba en $x=0$, se encuentra en $x=6$, y la forma es idéntica: ni se ensancha ni se aplana.

(b) Retroceder por la característica. Para saber $u(10;\,4)$ con $c=2$, se retrocede: $x-ct=10-8=2$, luego $u(10;\,4)=f(2)=\mathrm{e}^{-4}=0{,}018316$. No hace falta resolver nada, solo seguir una recta.

(c) Las esquinas sobreviven. Si el dato inicial es la función escalón, que vale 1 para $x<0$ y 0 para $x>0$, la solución es el mismo escalón desplazado. La discontinuidad se propaga intacta a velocidad c: las ecuaciones hiperbólicas no suavizan nada.

(d) Con coeficiente variable. Para $u_t + x\,u_x = 0$ las características cumplen $\frac{\mathrm{d}x}{\mathrm{d}t}=x$, es decir $x=x_0\mathrm{e}^{t}$, que ya no son rectas sino exponenciales. La solución es $u(x,t)=f(x\mathrm{e}^{-t})$. El método sigue funcionando; solo cambian las curvas por las que viaja la información.

Cuidado

[Con ecuaciones no lineales las características chocan] En la ecuación de Burgers $u_t + u\,u_x = 0$, la velocidad de cada característica es el propio valor de u, de modo que las partes altas del perfil avanzan más deprisa que las bajas y acaban alcanzándolas. Cuando dos características se cortan, la solución dejaría de ser una función, y lo que ocurre en realidad es que se forma una *onda de choque*: una discontinuidad que se propaga. Ese es el mecanismo del estampido sónico y de los atascos de tráfico, y obliga a ampliar la noción de solución para admitir funciones discontinuas.

13.4. La ecuación de ondas y la fórmula de d'Alembert

Teorema 13.24 (d'Alembert)**.** *La solución de $u_{tt}=c^2u_{xx}$ en toda la recta con $u(x,0)=f(x)$ y $u_t(x,0)=g(x)$ es*

$$u(x,t)=\frac{f(x-ct)+f(x+ct)}{2}+\frac{1}{2c}\int_{x-ct}^{x+ct} g(s)\,\mathrm{d}s.$$

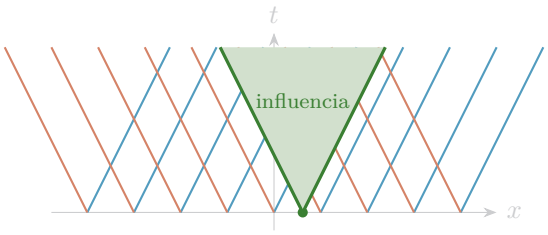


Figura 13.1: Las características de la ecuación de ondas. La información viaja por las rectas $x\pm ct$ a velocidad finita, de modo que lo que ocurre en un punto sólo afecta al cono sombreado. La ecuación del calor no tiene características reales y propaga instantáneamente.

Idea clave

Qué dice esto en cristiano. La ecuación de ondas se factoriza: $\partial_{tt}-c^2\partial_{xx}=(\partial_t-c\partial_x)(\partial_t+c\partial_x)$, es decir, es un producto de dos ecuaciones de transporte, una hacia la derecha y otra hacia la izquierda. Por eso la solución es la suma de dos ondas viajeras.

La fórmula dice además algo muy concreto: el valor en (x,t) depende *únicamente* de lo que ocurriera inicialmente en el intervalo $[x-ct,\,x+ct]$. Fuera de ese intervalo, lo que pasara no ha tenido tiempo de llegar. Es la velocidad finita de propagación, escrita como fórmula.

Ejemplo 13.25 (Una cuerda pulsada y el dominio de dependencia)**. (a) Un pulso que se parte en dos.** Con $c=1$, $g=0$ y $f(x)=\mathrm{e}^{-x^2}$, la solución es $u=\frac{\mathrm{e}^{-(x-t)^2}+\mathrm{e}^{-(x+t)^2}}{2}$: el pulso inicial se divide en dos mitades que viajan en sentidos opuestos. En $x=0$ la altura vale $\mathrm{e}^{-1}=0{,}367879$ para $t=1$ y $\mathrm{e}^{-4}=0{,}018316$ para $t=2$: el centro se vacía a medida que las dos mitades se alejan.

(b) El dominio de dependencia. El valor de $u(0;\,3)$ con $c=1$ depende solo de f y g en $[-3,3]$. Cambiar el dato inicial en $x=10$ no altera nada en $(0,3)$: la señal viajaría a velocidad 1 y necesitaría diez unidades de tiempo para llegar.

(c) El cono de influencia. Al revés, un dato inicial concentrado en $x=0$ influye en t solo dentro de $[-ct,\,ct]$. En el plano (x,t) eso dibuja un cono, y ese cono es literalmente el cono de luz de la relatividad cuando c es la velocidad de la luz: la ecuación de ondas es la que impone que nada viaje más deprisa que c.

(d) Conservación de la energía. La cantidad $E(t)=\frac{1}{2}\int\left(u_t^2+c^2u_x^2\right)\mathrm{d}x$ es constante en el tiempo, como se comprueba derivando bajo la integral e integrando por partes. Físicamente es la suma de la energía cinética y la potencial de la cuerda, y su conservación es lo que impide que la vibración se amortigüe: la ecuación de ondas pura no tiene rozamiento.

Cuidado

[Hacen falta dos datos iniciales, y la ecuación es reversible] Al ser de segundo orden en el tiempo, la ecuación de ondas necesita la posición y la velocidad iniciales; con solo una de las dos el problema queda indeterminado. Segundo aviso, y es lo que la separa del calor: la ecuación de ondas es *reversible*. Cambiando t por $-t$ la ecuación no cambia, así que se puede integrar hacia atrás sin ninguna dificultad y recuperar el pasado a partir del presente. Y tercero: en dimensión tres la fórmula es distinta y aparece el principio de Huygens, según el cual la señal llega y se va limpiamente; en dimensión dos, en cambio, deja una cola que nunca se apaga del todo, razón por la cual las ondas en la superficie de un estanque siguen agitándose después de pasar el frente.

13.5. La ecuación del calor: difusión e irreversibilidad

Teorema 13.26 (Núcleo del calor)**.** *La solución de $u_t=k\,u_{xx}$ en toda la recta con $u(x,0)=f(x)$ es*

$$u(x,t)=\int_{\mathbb{R}} G(x-y,t)\,f(y)\,\mathrm{d}y,\qquad G(x,t)=\frac{1}{\sqrt{4\pi kt}}\,\mathrm{e}^{-x^2/(4kt)},$$

donde G es el núcleo del calor*, una gaussiana de anchura $\sqrt{2kt}$ e integral total 1.*

Idea clave

Qué dice esto en cristiano. La solución es el dato inicial *emborronado* con una campana de Gauss que se va ensanchando. Cada punto reparte su calor entre sus vecinos, dando más a los cercanos que a los lejanos, con un reparto que sigue exactamente la curva normal. Y como la integral del núcleo vale 1 en todo instante, el calor total se conserva: no se crea ni se destruye, solo se reparte. La anchura crece como $\sqrt{t}$ y no como t, que es la firma inconfundible de la difusión frente al transporte: para recorrer el doble de distancia hace falta cuatro veces más tiempo.

Ejemplo 13.27 (Cómo se ensancha una gota de calor)**. (a) El ensanchamiento.** Con $k = 1$, la altura del núcleo en el origen vale 0,892062 para $t = 0{,}1$, luego 0,398942 para $t = 0{,}5$ y 0,199471 para $t = 2$. Las anchuras correspondientes, $\sqrt{2kt}$, son 0,447214, 1 y 2. La campana baja y se ensancha conservando el área.

(b) La raíz cuadrada del tiempo. Para que la anchura pase de 1 a 2 hace falta pasar de $t = 0{,}5$ a $t = 2$: cuatro veces más tiempo para el doble de distancia. Por eso el humo de un cigarrillo tarda segundos en llenar un palmo y horas en llenar la habitación, y por eso la difusión es un mecanismo de transporte pésimo a escalas grandes.

(c) Suavizado instantáneo. Aunque f sea discontinua, la solución es C^∞ para todo $t > 0$: la convolución con una gaussiana suaviza cualquier cosa. Las esquinas del dato inicial desaparecen inmediatamente, justo al revés que en la ecuación de ondas.

(d) Velocidad infinita de propagación. Como $G(x,t) > 0$ para todo x en cuanto $t > 0$, un dato inicial concentrado en el origen produce, un instante después, una temperatura estrictamente positiva a un millón de kilómetros. Es minúscula, pero no es cero. Físicamente eso es un defecto del modelo, y las correcciones relativistas lo arreglan; matemáticamente es una propiedad inevitable de las ecuaciones parabólicas.

Cuidado

[El calor hacia atrás está mal planteado] Integrar la ecuación del calor hacia el pasado es un problema *mal planteado* en el sentido de Hadamard: la solución existe y es única, pero no depende de forma continua de los datos. Con el dato $u(x,0) = \varepsilon \sin(nx)$ la solución hacia atrás es $\varepsilon \sin(nx)\, \mathrm{e}^{n^2 t}$. Tomando $\varepsilon = 10^{-6}$ y $t = 0{,}01$: para $n = 10$ la amplitud es $2{,}718 \cdot 10^{-6}$, inofensiva; para $n = 50$ ya vale $7{,}200 \cdot 10^{4}$; y para $n = 100$ alcanza $2{,}688 \cdot 10^{37}$. Una perturbación invisible en el dato produce una respuesta astronómica. Por eso no se puede reconstruir el pasado térmico de un objeto midiendo su temperatura actual, y por eso la difusión es irreversible.

13.6. La ecuación de Laplace: el equilibrio

Teorema 13.28 (Principio del máximo)**.** *Si u es armónica ($\Delta u = 0$) en un dominio acotado Ω y continua en $\overline{\Omega}$, entonces alcanza su máximo y su mínimo en la frontera $\partial\Omega$. En consecuencia, dos soluciones del problema de Dirichlet con los mismos valores de frontera coinciden.*

Idea clave

Qué dice esto en cristiano. Una función armónica es la que cumple que su valor en cada punto es el promedio de los valores en cualquier esfera centrada ahí. De esa propiedad sale todo: ningún punto interior puede ser mayor que su propio promedio, luego no puede haber picos dentro y los extremos tienen que estar en el borde. Físicamente es la temperatura de equilibrio de una placa: el punto más caliente está siempre en el contorno, porque si estuviera dentro el calor fluiría desde él y no habría equilibrio. Y de ahí sale la unicidad en tres renglones, sin construir ninguna solución.

Ejemplo 13.29 (Armónicas y el principio del máximo)**. (a) Reconocer armónicas.** $u = x^2 - y^2$ tiene laplaciano $2 - 2 = 0$: armónica. También $u = xy$, con laplaciano $0 + 0 = 0$, y $u = \mathrm{e}^x \cos y$. En cambio $u = x^2 + y^2$ tiene laplaciano $4 \neq 0$: no lo es, y en efecto tiene un mínimo interior en el origen, cosa que ninguna armónica puede permitirse.

(b) El máximo en el borde. En el disco unidad, $u = x^2 - y^2$ vale 1 en $(1, 0)$, -1 en $(0, 1)$ y 0 en el centro. Máximo y mínimo están en la circunferencia, como manda el teorema.

(c) La unicidad, entera. Si u_1 y u_2 resuelven $\Delta u = f$ con los mismos valores de frontera, la diferencia $u = u_1 - u_2$ es armónica y vale cero en el borde. Por el principio del máximo su máximo y su mínimo valen cero, luego $u \equiv 0$. Tres líneas, sin resolver nada.

(d) La conexión con variable compleja. Si $f = u + \mathrm{i}v$ es holomorfa, las ecuaciones de Cauchy–Riemann implican que u y v son armónicas. Por eso $x^2 - y^2$ y $2xy$ lo son: son la parte real e imaginaria de z^2. Toda la teoría de funciones armónicas en el plano es análisis complejo disfrazado, y esa es la razón de que las transformaciones conformes sirvan para resolver problemas de Dirichlet en regiones complicadas.

Cuidado

[La acotación del dominio es imprescindible] En un dominio no acotado el principio del máximo puede fallar: la función $u = y$ es armónica en el semiplano superior, vale 0 en toda su frontera y no es idénticamente nula. Hace falta añadir una condición de comportamiento en el infinito para recuperar la unicidad. Segundo aviso: el problema de *Neumann*, en el que se dan las derivadas normales en vez de los valores, solo tiene solución si el flujo total por la frontera es cero (no se puede meter calor indefinidamente en una placa y esperar equilibrio), y entonces la solución es única salvo una constante aditiva.

13.7. Separación de variables y series de Fourier

Regla: resolver una EDP por separación de variables

El método funciona en dominios acotados con geometría sencilla y condiciones de contorno homogéneas. El procedimiento completo es:

1. **Busca soluciones de la forma** $u(x, t) = X(x)T(t)$ y sustituye en la ecuación.

2. **Separa las variables** dejando todo lo que depende de x a un lado y de t al otro; como son variables independientes, ambos lados deben valer la misma constante.
3. **Resuelve el problema espacial** con las condiciones de contorno: sale un problema de autovalores, y solo ciertos valores de la constante admiten solución no nula.
4. **Anota los autovalores y las autofunciones**, que para una barra con extremos fijos son $\lambda_n = \left(\frac{n\pi}{L}\right)^2$ y $\operatorname{sin} \frac{n\pi x}{L}$.
5. **Resuelve el problema temporal** para cada autovalor; sale una exponencial en el caso del calor y un seno o coseno en el de ondas.
6. **Superpone todas las soluciones** con coeficientes libres, aprovechando que la ecuación es lineal.
7. **Ajusta los coeficientes al dato inicial** desarrollándolo en serie de Fourier, lo que se hace con la fórmula de los coeficientes gracias a la ortogonalidad de las autofunciones.

Ejemplo 13.30 (Una barra que se enfría, paso a paso)**. (a) El planteamiento.** Barra $[0, \pi]$ con $k = 1$ y extremos mantenidos a temperatura cero. Separando, $\frac{T'}{T} = \frac{X''}{X} = -\lambda$. Las condiciones $X(0) = X(\pi) = 0$ obligan a $\lambda_n = n^2$ y $X_n = \operatorname{sin}(nx)$, y entonces $T_n = \mathrm{e}^{-n^2 t}$.

(b) Un dato inicial que ya es una autofunción. Si $u(x, 0) = \operatorname{sin} x$, la solución es directamente $u = \operatorname{sin}(x)\, \mathrm{e}^{-t}$, sin ninguna serie. En $x = \frac{\pi}{2}$ y $t = 1$ vale $\mathrm{e}^{-1} = 0{,}367879$. La forma no cambia; solo se aplasta.

(c) Un dato inicial cualquiera. Si $u(x, 0) = 1$ en toda la barra, los coeficientes son $b_n = \frac{4}{n\pi}$ para n impar y 0 para n par: $b_1 = 1{,}273240$, $b_3 = 0{,}424413$, $b_5 = 0{,}254648$, $b_7 = 0{,}181891$. Sumando cien términos impares en $x = \frac{\pi}{2}$ y $t = 0$ se obtiene 0,996817, acercándose al valor correcto 1.

(d) Los modos altos mueren primero. El término n-ésimo decae como $\mathrm{e}^{-n^2 t}$, así que el modo $n = 5$ se apaga veinticinco veces más rápido que el fundamental. A tiempos moderados solo sobrevive el primer término, y la temperatura tiende a la forma $\operatorname{sin} x$ sea cual sea el dato inicial. Esa es la razón de que la difusión olvide los detalles: son precisamente los modos de frecuencia alta.

Cuidado

[El método exige linealidad, homogeneidad y geometría compatible] La superposición del paso 6 solo es legítima porque la ecuación es lineal; con una ecuación no lineal el método se derrumba en el primer paso. Segundo aviso: las condiciones de contorno deben ser homogéneas (valor cero en los extremos); si no lo son, hay que restar primero una solución estacionaria que las absorba. Y tercero: la geometría debe ser compatible con la separación. Un rectángulo, un disco o una esfera sí lo son; un dominio con forma de L, no, y ahí no queda más remedio que ir a métodos numéricos.

13.8. Problemas bien planteados

Definición 13.31. Un problema está *bien planteado* en el sentido de Hadamard si cumple tres condiciones: existe solución, es única, y depende de forma *continua* de los datos, es decir, un cambio pequeño en el dato produce un cambio pequeño en la solución. Si falla cualquiera de las tres, el problema está *mal planteado*.

Idea clave

Qué dice esto en cristiano. La tercera condición es la que se olvida y la que de verdad importa en la práctica. Un problema con solución única pero que amplifica las perturbaciones es inútil, porque los datos reales siempre tienen error de medida y ese error se convierte en basura. Es exactamente el mismo fenómeno que el número de condición en álgebra lineal, llevado a dimensión infinita. Y cada tipo de ecuación tiene su combinación de datos que la hace bien planteada: dar los datos equivocados no produce un problema difícil, produce un problema sin sentido.

Ejemplo 13.32 (La combinación correcta para cada tipo)**. (a) Lo que necesita cada una.** Ondas: posición y velocidad iniciales en toda la recta, o bien esos datos más condiciones en los extremos si el dominio es acotado. Calor: solo el dato inicial, más los extremos. Laplace: valores en *toda* la frontera y ningún dato inicial.

(b) Calor hacia atrás. Como se vio, existe y es única, pero amplifica el modo n por $e^{n^2 t}$: falla la tercera condición. Es el ejemplo canónico de problema mal planteado, y explica por qué restaurar una fotografía desenfocada o deducir la temperatura pasada de un objeto son problemas tan difíciles.

(c) Laplace con datos iniciales. El propio Hadamard dio este ejemplo: resolver $u_{xx} + u_{yy} = 0$ con $u(x,0) = 0$ y $u_y(x,0) = \frac{\sin(nx)}{n}$ da $u = \frac{\sin(nx)\sinh(ny)}{n^2}$. El dato tiende a cero cuando n crece, y la solución crece como e^{ny}: sin dependencia continua. Las ecuaciones elípticas no admiten problemas de valor inicial.

(d) Sin unicidad. El problema $u_{xx} + u_{yy} = 0$ en el disco unidad sin ninguna condición de frontera tiene infinitas soluciones: todas las armónicas. Y el problema de Neumann puro no tiene ninguna solución si el flujo total no se anula. Falla la existencia por un motivo puramente físico: lo que entra tiene que salir.

Cuidado

[Mal planteado no significa inútil] Muchísimos problemas de interés real están mal planteados: la tomografía, la deconvolución de imágenes, la inversión sísmica. Lo que se hace no es rendirse, sino *regularizar*: se añade un término que penaliza las soluciones muy oscilantes (regularización de Tíjonov) y se obtiene un problema cercano que sí está bien planteado, a costa de una solución algo suavizada. Segundo aviso: que un problema esté bien planteado no dice nada sobre si se puede resolver explícitamente; casi ninguna EDP tiene solución cerrada, y por eso los métodos numéricos del capítulo 14 son la herramienta habitual.

13.9. Errores típicos con ecuaciones en derivadas parciales

Cuidado

[Las ocho confusiones más frecuentes] Aplicar intuiciones de ecuaciones ordinarias es el origen de casi todos los errores del tema:

1. **Creer que hay un teorema general de existencia.** No lo hay; cada familia se estudia aparte, y hay ecuaciones lineales suaves sin ninguna solución.
2. **Dar el número equivocado de datos.** Ondas necesita dos condiciones iniciales por ser de segundo orden en el tiempo; el calor solo una.
3. **Poner datos iniciales a una ecuación elíptica.** Laplace no tiene tiempo; exige frontera cerrada, y con datos iniciales el problema queda mal planteado.
4. **Suponer que el calor se puede integrar hacia atrás.** Se puede en teoría y no en la práctica, porque amplifica el ruido exponencialmente en el cuadrado de la frecuencia.
5. **Separar variables en una ecuación no lineal.** La superposición de soluciones exige linealidad y sin ella el método no significa nada.
6. **Olvidar homogeneizar las condiciones de contorno** antes de separar; con extremos a temperatura no nula hay que restar primero la solución estacionaria.
7. **Confundir suavizado con estabilidad.** La ecuación del calor suaviza hacia adelante y explota hacia atrás: son la misma propiedad vista en los dos sentidos.
8. **Esperar que las soluciones sean funciones clásicas.** En las ecuaciones hiperbólicas no lineales aparecen choques, y hay que trabajar con soluciones débiles o distribuciones.

Pausa para pensar

Las tres ecuaciones de este capítulo nacieron casi a la vez y de problemas muy concretos. La de ondas la escribió d'Alembert en 1747 estudiando la cuerda de un violín, y su discusión con Euler y Daniel Bernoulli sobre qué funciones podían ser datos iniciales duró décadas y obligó a repensar qué era una función. La del calor la publicó Fourier en 1822, y para resolverla inventó las series que llevan su nombre; la Academia de París le había rechazado el trabajo en 1807 porque nadie creía que una función discontinua pudiera escribirse como suma de senos. Y la de Laplace venía de la mecánica celeste. Es notable que tres problemas tan distintos (una cuerda, una barra caliente y el potencial gravitatorio) resultaran ser los tres representantes exactos de los tres tipos posibles, y que buena parte del análisis moderno (series de Fourier, espacios de Hilbert, teoría de distribuciones) se inventara para responderlas.

13.10. Soluciones débiles, formulación variacional y el teorema de Lax–Milgram

Las ecuaciones diferenciales ordinarias tienen a Picard: un teorema que garantiza existencia y unicidad antes de calcular nada. Para las EDP no hay Picard. Lo que hay —y es enormemente potente en el caso elíptico— es un cambio de punto de vista: dejar de exigir que la solución sea dos veces derivable y pedirle solo que se comporte bien *al integrarla contra funciones de prueba*.

Idea clave

El truco es viejo y consiste en pasar las derivadas al otro lado. Si $-u'' = f$ en $(0,1)$ con $u(0) = u(1) = 0$, multiplíquese por una función v que se anule en los extremos e intégrese por partes:

$$\int_0^1 u'v'\,\mathrm{d}x = \int_0^1 fv\,\mathrm{d}x \qquad \text{para toda } v.$$

En la izquierda ya no aparece u'': solo u'. Una función que cumpla esa igualdad para toda v admisible es una **solución débil**, y puede tener un pico donde la clásica no existiría. Como la física real está llena de picos —materiales con dos conductividades pegadas, cargas puntuales—, la noción débil no es una concesión técnica: es la correcta.

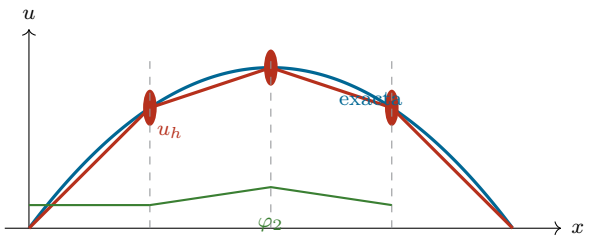


Figura 13.2: Galerkin con tres funciones sombrero sobre $-u'' = 1$. La solución discreta es una poligonal y sin embargo pasa *exactamente* por la solución verdadera en los nodos: 0,09375, 0,125 y 0,09375. La matriz que se ensambla al integrar los productos $\int \varphi_i'\varphi_j'$ es la misma tridiagonal de las diferencias finitas, obtenida por otro camino.

Definición 13.33 (Derivada débil y espacios de Sobolev)**.** g es la **derivada débil** de u si $\int u\varphi' = -\int g\varphi$ para toda φ infinitamente derivable con soporte compacto. El espacio $H^1(\Omega)$ está formado por las funciones de L^2 cuya derivada débil también está en L^2, con la norma

$$\|u\|_{H^1}^2 = \|u\|_{L^2}^2 + \|\nabla u\|_{L^2}^2,$$

y $H_0^1(\Omega)$ es el cierre de las funciones de soporte compacto: las que "se anulan en la frontera" en sentido débil.

Teorema 13.34 (Desigualdad de Poincaré)**.** *Si Ω está acotado, existe C tal que $\|u\|_{L^2} \le C\|\nabla u\|_{L^2}$ para todo $u \in H_0^1(\Omega)$.*

Esto es lo que hace que $\|\nabla u\|_{L^2}$ sea por sí sola una norma equivalente en H_0^1, y es la pieza que faltaba para el teorema central.

Teorema 13.35 (Lax–Milgram)**.** *Sea H un Hilbert, $a(\cdot,\cdot)$ una forma bilineal* acotada *($|a(u,v)| \le M\|u\|\|v\|$) y* coerciva *($a(u,u) \ge \alpha\|u\|^2$ con $\alpha > 0$), y L un funcional lineal continuo. Entonces existe un único $u \in H$ con*

$$a(u,v) = L(v) \qquad \text{para todo } v \in H,$$

y además $\|u\| \le \|L\|/\alpha$.

Ejemplo 13.36 (El problema de Dirichlet, resuelto de un plumazo)**.** Para $-\Delta u = f$ en Ω con $u = 0$ en $\partial\Omega$, tómese $H = H_0^1(\Omega)$,

$$a(u,v) = \int_\Omega \nabla u \cdot \nabla v, \qquad L(v) = \int_\Omega fv.$$

Acotada: por Cauchy–Schwarz. Coerciva: $a(u,u) = \|\nabla u\|^2$, que por Poincaré domina la norma completa. Continua: $|L(v)| \le \|f\|_{L^2}\|v\|_{L^2} \le C\|f\|\,\|v\|_{H^1}$. Lax–Milgram concluye: existe una única solución débil, y depende continuamente del dato f. El problema está bien planteado, y no hemos calculado nada.

Observación 13.37. Cuando a es además simétrica, Lax–Milgram se reduce al teorema de representación de Riesz y la solución tiene una segunda lectura: es el mínimo del funcional de energía

$$J(v) = \tfrac{1}{2}a(v,v) - L(v).$$

Esa es la formulación variacional propiamente dicha, y es la que hace que la solución de $-\Delta u = f$ sea la configuración de mínima energía elástica de una membrana cargada. El principio de Dirichlet que Riemann usó sin demostrar y Weierstrass criticó con razón queda aquí demostrado, ochenta años después, con las herramientas correctas.

Del teorema al ordenador: el método de Galerkin

La formulación débil no solo demuestra existencia: se calcula. Se sustituye H por un subespacio V_h de dimensión finita y se resuelve el mismo problema ahí.

Ejemplo 13.38 ($-u'' = 1$ resuelto con funciones sombrero)**.** Sea $-u'' = 1$ en $(0,1)$, $u(0) = u(1) = 0$, cuya solución exacta es $u(x) = x(1-x)/2$.

Tómense n nodos interiores equiespaciados con paso $h = 1/(n+1)$ y las funciones sombrero φ_i, que valen 1 en el nodo i y 0 en los demás. Con $u_h = \sum c_j\varphi_j$, el sistema $a(u_h,\varphi_i) = L(\varphi_i)$ tiene matriz tridiagonal

$$K = \frac{1}{h}\begin{pmatrix} 2 & -1 & & \\ -1 & 2 & -1 & \\ & \ddots & \ddots & \ddots \\ & & -1 & 2 \end{pmatrix}, \qquad F_i = h,$$

que es exactamente la matriz de diferencias finitas, obtenida aquí por un camino completamente distinto.

n	nodos	u_h en los nodos
1	0,5	0,125
2	0,3333; 0,6667	0,111111; 0,111111
3	0,25; 0,5; 0,75	0,09375; 0,125; 0,09375

Compárese con la exacta: $u(0{,}5) = 0{,}125$, $u(1/3) = 0{,}111111$, $u(0{,}25) = 0{,}09375$. Coinciden *exactamente*, con un solo nodo ya. No es casualidad ni suerte: para este operador en una dimensión, la solución de Galerkin es interpolante exacta de la verdadera en los nodos, un fenómeno de superconvergencia que se pierde en cuanto el coeficiente deja de ser constante o se sube de dimensión. Pero la estructura —ensamblar una matriz a partir de integrales de funciones base— es la misma que usa todo programa de elementos finitos del mundo.

Teorema 13.39 (Lema de Céa)**.** *Si u es la solución débil y u_h la de Galerkin en $V_h \subset H$, entonces*

$$\|u - u_h\| \le \frac{M}{\alpha} \inf_{v \in V_h} \|u - v\|.$$

En palabras: la solución numérica es, salvo una constante, tan buena como la mejor aproximación que el espacio V_h permita. El problema del análisis numérico se reduce entonces a un problema de teoría de la aproximación, y por eso los elementos finitos tienen una teoría de errores tan limpia. Con sombreros lineales, $\inf \|u - v\|_{H^1} = O(h)$, y de ahí la convergencia de primer orden en energía y de segundo en L^2.

13.11. Solución fundamental, funciones de Green y el núcleo de Poisson

El principio del máximo dice cómo se comportan las soluciones de Laplace, pero no produce ninguna. Para eso hace falta una fórmula, y la fórmula sale siempre del mismo objeto: la respuesta del operador a una carga puntual.

Definición 13.40 (Solución fundamental)**.** Φ es solución fundamental de $-\Delta$ si $-\Delta\Phi = \delta_0$ en sentido de distribuciones. Explícitamente,

$$\Phi(x) = \begin{cases} -\dfrac{1}{2\pi} \log |x| & \text{en } \mathbb{R}^2, \\[2mm] \dfrac{1}{4\pi |x|} & \text{en } \mathbb{R}^3. \end{cases}$$

En el espacio, $1/(4\pi r)$ es el potencial newtoniano de una masa puntual: la ley de la gravitación de Newton *es* la solución fundamental del laplaciano, y la ley de Coulomb también. La convolución $\Phi * f$ resuelve $-\Delta u = f$ en todo el espacio: el potencial creado por una distribución de carga es la superposición de los potenciales de sus partes.

Definición 13.41 (Función de Green)**.** Para un dominio Ω, $G(x,y) = \Phi(x-y) - h^y(x)$, donde h^y es armónica en Ω y coincide con $\Phi(\cdot - y)$ en $\partial\Omega$. Entonces G se anula en la frontera y

$$u(x) = -\int_{\partial\Omega} \frac{\partial G}{\partial \nu}(x,y)\, g(y)\, \mathrm{d}S(y) + \int_\Omega G(x,y) f(y)\, \mathrm{d}y$$

resuelve $-\Delta u = f$ con $u = g$ en la frontera.

Ejemplo 13.42 (La Green del semiplano por reflexión)**.** En $\{y > 0\} \subset \mathbb{R}^2$, la función armónica que iguala a Φ en la frontera se obtiene poniendo una carga *imagen* de signo opuesto en el punto reflejado $\bar{y} = (y_1, -y_2)$:

$$G(x,y) = -\frac{1}{2\pi}\log|x-y| + \frac{1}{2\pi}\log|x-\bar{y}|.$$

Sobre el eje real, $|x-y| = |x-\bar{y}|$ y $G = 0$ automáticamente. El método de las imágenes de la electrostática es exactamente esto: para saber qué hace una carga junto a una placa conductora, se sustituye la placa por una carga espejo.

Teorema 13.43 (Fórmula de Poisson en el disco)**.** *La solución de $\Delta u = 0$ en el disco unidad con $u = g$ en el borde es*

$$u(r,\theta) = \frac{1}{2\pi}\int_{-\pi}^{\pi} \frac{1-r^2}{1-2r\cos(\theta-\varphi)+r^2}\, g(\varphi)\, \mathrm{d}\varphi.$$

Ejemplo 13.44 (Un dato de contorno concreto, por dos caminos)**.** Sea $g(\varphi) = \cos^2\varphi$.

(a) Por series. La identidad $\cos^2\varphi = \frac{1}{2} + \frac{1}{2}\cos 2\varphi$ y el hecho de que $r^n\cos n\theta$ es armónica dan de inmediato

$$u(r,\theta) = \tfrac{1}{2} + \tfrac{1}{2}r^2\cos 2\theta.$$

(b) Por la integral de Poisson. Evaluando numéricamente en $r = 0{,}5$, $\theta = 0$ se obtiene $0{,}625\,000\,00$, y la serie da $\frac{1}{2}+\frac{1}{2}(0{,}25)(1) = 0{,}625$. En $r = 0{,}5$, $\theta = \pi/4$ la integral da $0{,}500\,000\,00$ y la serie $\frac{1}{2} + \frac{1}{2}(0{,}25)(0) = 0{,}5$. Coinciden hasta la octava cifra.

(c) El valor en el centro. Con $r = 0$ el núcleo vale 1 y la fórmula se reduce a

$$u(0) = \frac{1}{2\pi}\int_{-\pi}^{\pi} g = \text{media de } g \text{ sobre el borde} = \tfrac{1}{2}.$$

Ahí está la propiedad del valor medio, no como intuición sino como caso particular de una fórmula.

Idea clave

El núcleo de Poisson es una *aproximación de la identidad*: al acercarse al borde ($r \to 1$) se concentra alrededor de $\varphi = \theta$ y reproduce g. De la fórmula salen en cascada casi todas las propiedades de las funciones armónicas: la propiedad del valor medio (poniendo $r = 0$), la desigualdad de Harnack (acotando el núcleo por $\frac{1-r}{1+r}$ y $\frac{1+r}{1-r}$), el teorema de Liouville (una función armónica acotada en todo el plano es constante, porque Harnack aplicada en discos cada vez mayores aplasta las diferencias) y la analiticidad (el núcleo

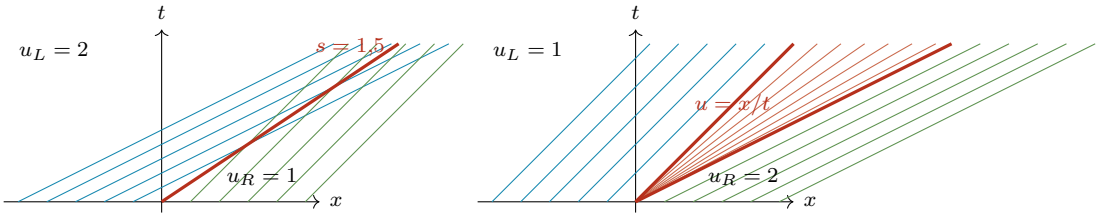
es analítico en x, y derivar bajo la integral está permitido).
Compárese con lo que ocurre en variable compleja: la fórmula integral de Cauchy hace exactamente el mismo trabajo, y por la misma razón. La parte real de una función holomorfa es armónica, y el núcleo de Poisson es la parte real del núcleo de Cauchy.

13.12. Leyes de conservación no lineales: choques, Rankine–Hugoniot y entropía

La ecuación de Burgers $u_t + uu_x = 0$ es el ejemplo más simple posible de ecuación no lineal, y ya exhibe todo lo que puede salir mal.

Idea clave

Léase la ecuación como "cada punto del perfil viaja a velocidad igual a su propia altura". Las crestas van más deprisa que los valles. Si al principio la altura decrece hacia la derecha, las crestas alcanzan a los valles y el perfil se vuelca: la solución dejaría de ser una función. En ese instante deja de existir la solución clásica, y no por falta de técnica, sino porque físicamente ahí se forma una *onda de choque*.



(a) choque: las características chocan

(b) rarefacción: se separan

Figura 13.3: Las características de $u_t + uu_x = 0$ son rectas de pendiente $1/u$. Si el dato decrece se cruzan y nace un choque que viaja a la media de las dos velocidades; si crece dejan un hueco que se rellena con un abanico. Los dos casos son simétricos en el dato inicial y radicalmente distintos en la solución: es la condición de entropía la que descarta el choque en el segundo.

Teorema 13.45 (Tiempo de rotura)**.** *Para $u_t + uu_x = 0$ con dato $u(x,0) = u_0(x)$, las características son rectas y la solución es implícita, $u = u_0(x - ut)$. La primera singularidad aparece en*

$$t^* = \frac{-1}{\min_x u_0'(x)},$$

siempre que u_0 decrezca en algún punto.

Con $u_0(x) = -\arctan x$, por ejemplo, $\min u_0' = -1$ y el choque nace en $t^* = 1$.

Teorema 13.46 (Condición de salto de Rankine–Hugoniot)**.** *Si $u_t + f(u)_x = 0$ admite una solución débil con una discontinuidad que se mueve a velocidad s, con valores u_L y u_R a cada lado, entonces*

$$s = \frac{f(u_L) - f(u_R)}{u_L - u_R}.$$

Para Burgers, con $f(u) = u^2/2$, esto da $s = \dfrac{u_L + u_R}{2}$: el choque viaja a la media *de las dos velocidades.*

Ejemplo 13.47 (Dos escalones simétricos, dos destinos opuestos)**. (a) Escalón decreciente: $u_L = 2$, $u_R = 1$.** Las características de la izquierda (velocidad 2) alcanzan a las de la derecha (velocidad 1): se cruzan de inmediato. La solución es un choque desde $t = 0$, con velocidad

$$s = \frac{2+1}{2} = 1{,}5.$$

Es decir, la discontinuidad avanza más despacio que el fluido de detrás y más deprisa que el de delante: por eso lo alcanza y lo empuja. Un atasco de tráfico se comporta exactamente así.

(b) Escalón creciente: $u_L = 1$, $u_R = 2$. Ahora las características se separan y dejan un sector vacío entre $x = t$ y $x = 2t$. La solución física lo rellena con un *abanico de rarefacción*:

$$u(x,t) = \begin{cases} 1 & x \le t, \\ x/t & t \le x \le 2t, \\ 2 & x \ge 2t. \end{cases}$$

Es continua, y resuelve la ecuación en cada trozo.

(c) La solución espuria. Y sin embargo, el escalón quieto —mantener la discontinuidad y hacerla viajar a $s = (1+2)/2 = 1{,}5$— *también* satisface Rankine–Hugoniot y también es solución débil. Hay dos, y la ecuación por sí sola no distingue entre ellas. Ese es el escándalo: la formulación débil compra existencia al precio de perder unicidad.

Teorema 13.48 (Condición de entropía de Lax)**.** *Una discontinuidad es admisible solo si las características* entran *en ella por ambos lados:*

$$f'(u_L) > s > f'(u_R), \qquad \text{para Burgers: } u_L > u_R.$$

Con esta restricción añadida, el problema de valores iniciales recupera existencia y unicidad para datos acotados.

En el caso (c), $u_L = 1 < 2 = u_R$: la condición falla y el choque queda descartado. Sobrevive el abanico, que es la solución correcta. La condición de entropía no se deduce de la ecuación: se *añade*, y su justificación es física —es el límite cuando la viscosidad tiende a cero de la ecuación con difusión $u_t + uu_x = \varepsilon u_{xx}$, que sí tiene solución única—. La irreversibilidad termodinámica entra en las matemáticas por esta puerta y no por otra.

Observación 13.49. Esta es la razón profunda de que las EDP no lineales sean un mundo aparte. Todo el aparato del capítulo —superposición, separación de variables, series de Fourier— descansa en la linealidad. En cuanto se pierde, aparecen choques, hay que redefinir qué es

una solución, la unicidad exige un criterio extra, y los métodos numéricos ingenuos producen oscilaciones espurias junto a las discontinuidades. Los esquemas conservativos de Godunov, Lax–Friedrichs y sus descendientes con limitadores de flujo existen precisamente para respetar Rankine–Hugoniot y la entropía a nivel discreto.

13.13. Sturm–Liouville: por qué la separación de variables funciona, y qué sale en el disco

La separación de variables se ha ejecutado siempre con senos en un intervalo. Eso no es una casualidad ni una limitación: los senos son las autofunciones de un problema de contorno concreto, y en cuanto se cambia de geometría aparecen otras.

Definición 13.50 (Problema regular de Sturm–Liouville)**.**

$$-\big(p(x)u'\big)' + q(x)u = \lambda\, w(x)\, u \quad \text{en } [a,b],$$

con condiciones de contorno separadas, $p > 0$ y $w > 0$ continuas.

Teorema 13.51 (Lo que garantiza)**.** *1. Los autovalores son reales, simples y forman una sucesión creciente $\lambda_1 < \lambda_2 < \cdots \to \infty$.*

2. *Las autofunciones son ortogonales con el peso w: $\int_a^b u_n u_m\, w\,\mathrm{d}x = 0$ si $n \neq m$.*
3. *El sistema $\{u_n\}$ es* completo *en $L^2(w)$: toda función de cuadrado integrable admite desarrollo.*
4. *El menor autovalor es el mínimo del cociente de Rayleigh* $\lambda_1 = \min_{u} \dfrac{\int p|u'|^2 + q|u|^2}{\int w|u|^2}$.

Ahí está la respuesta completa a por qué la separación de variables funciona: produce un problema de Sturm–Liouville, y los cuatro puntos garantizan que existe una base ortogonal en la que desarrollar el dato inicial. Con $p = w = 1$, $q = 0$ y condiciones nulas en los extremos, las autofunciones son $\sin(n\pi x/L)$ y se recupera Fourier.

Ejemplo 13.52 (La membrana circular, y por qué un tambor no suena a nota)**.** En el disco, separar variables en coordenadas polares para $\Delta u + \lambda u = 0$ deja la ecuación radial

$$r^2 R'' + rR' + (\lambda r^2 - n^2)R = 0,$$

que es la ecuación de Bessel. Las soluciones regulares en el origen son $J_n(\sqrt{\lambda}\, r)$, y la condición de borde $R(1) = 0$ obliga a que $\sqrt{\lambda}$ sea un cero de J_n.

Los primeros ceros son

$$j_{0,1} = 2{,}404\,826, \qquad j_{1,1} = 3{,}831\,706, \qquad j_{0,2} = 5{,}520\,078,$$

y las frecuencias del tambor son proporcionales a ellos. Los cocientes con el fundamental:

$$\frac{j_{1,1}}{j_{0,1}} = 1{,}593\,341, \qquad \frac{j_{0,2}}{j_{0,1}} = 2{,}295\,417.$$

No son enteros. En la cuerda vibrante los armónicos están en razón $1:2:3:\ldots$, y por eso una cuerda suena a una nota definida. En la membrana circular las razones son 1,59, 2,30, 2,92, … y el resultado es un sonido percusivo sin altura clara. La diferencia entre un violín y un timbal está en la lista de ceros de una función de Bessel.

Observación 13.53. En la esfera, el mismo procedimiento da los **armónicos esféricos** $Y_\ell^m(\theta,\varphi)$, con autovalores $\ell(\ell+1)$ y multiplicidad $2\ell+1$. Son la base de todo desarrollo sobre la esfera, y por eso aparecen en el mapa del fondo cósmico de microondas, en los orbitales atómicos y en el potencial gravitatorio terrestre. La multiplicidad $2\ell+1$ —autovalores no simples— es lo que distingue el caso multidimensional del teorema unidimensional de arriba, y es exactamente la dimensión de una representación irreducible de $SO(3)$: la geometría fija la degeneración.

Regla: problemas no homogéneos

Dos técnicas cierran los casos que la separación de variables no cubre directamente.

1. **Restar el estado estacionario.** Si las condiciones de contorno no son nulas pero sí constantes, se busca primero la solución estacionaria v (que resuelve $v''=0$ con esos valores, o sea una recta) y se estudia $w=u-v$, que ya tiene contorno nulo. Para una barra de $[0,1]$ con extremos a 20 y 80 grados e inicial nula, el estacionario es $v(x)=20+60x$, y la temperatura del centro sube de 0 hacia 50 con la ley $50-\sum_{n\ \text{impar}}\frac{200}{n\pi}\sin(n\pi/2)e^{-n^2\pi^2kt}$: con $k=1$ vale 1,24 en $t=0{,}02$, 11,38 en $t=0{,}05$ y 26,28 en $t=0{,}1$.

2. **Principio de Duhamel.** Para $u_t-ku_{xx}=F(x,t)$ con dato nulo, la solución es la superposición de las respuestas a impulsos instantáneos:

$$u(x,t)=\int_0^t S(t-\sigma)\,[F(\cdot,\sigma)]\,(x)\,\mathrm{d}\sigma,$$

donde $S(t)$ es el operador que resuelve el problema homogéneo. Es exactamente la fórmula de variación de las constantes de las ecuaciones ordinarias, subida un piso.

13.14. Clasificación general de segundo orden y formas canónicas

Hasta aquí los tres tipos se han presentado con tres ejemplos. Conviene saber qué hacer con una ecuación cualquiera, porque en la práctica las ecuaciones no llegan ya clasificadas.

Definición 13.54 (Ecuación lineal de segundo orden en dos variables)**.** Una ecuación de la forma

$$A(x,y)\,u_{xx}+2B(x,y)\,u_{xy}+C(x,y)\,u_{yy}+(\text{términos de orden}\le 1)=0$$

se llama *elíptica* en un punto si $B^2-AC<0$ allí, *parabólica* si $B^2-AC=0$ e *hiperbólica* si $B^2-AC>0$.

Cuidado

[Dos convenios para el mismo discriminante] Al comienzo del capítulo el término cruzado se escribió $B\,u_{xy}$ —sin el factor 2— y el discriminante quedó $\Delta = B^2 - 4AC$. Aquí se escribe $2B\,u_{xy}$ y el discriminante es $B^2 - AC$. Los dos convenios son estándar y *clasifican igual*: si $B_{\text{viejo}} = 2B$, entonces $B^2_{\text{viejo}} - 4AC = 4(B^2 - AC)$, que tiene el mismo signo. El convenio con $2B$ es el que hace que la matriz $\left(\begin{smallmatrix} A & B \\ B & C \end{smallmatrix}\right)$ tenga a B en las dos posiciones cruzadas, y por eso es el que se usa a partir de aquí. Lo único inaceptable es mezclarlos sin darse cuenta: antes de calcular, hay que mirar si el término cruzado lleva el 2 delante o no.

El nombre viene de que $A\xi^2 + 2B\xi\eta + C\eta^2$ es la ecuación de una elipse, una parábola o una hipérbola según el signo del mismo discriminante: la clasificación de las cónicas del capítulo 10 reaparece aquí como clasificación de ecuaciones.

Teorema 13.55 (Forma canónica)**.** *Sea $\Phi(x,y) =$ constante una curva a lo largo de la cual se busca simplificar. Si se sustituye $\xi = \Phi_1(x,y)$, $\eta = \Phi_2(x,y)$ con jacobiano no nulo, el coeficiente de $u_{\xi\xi}$ en la ecuación transformada es*

$$\widetilde{A} = A\,\xi_x^2 + 2B\,\xi_x\xi_y + C\,\xi_y^2.$$

En consecuencia $\widetilde{A} = 0$ exactamente cuando ξ es constante sobre las soluciones de la ecuación característica

$$A\,(\mathrm{d}y)^2 - 2B\,\mathrm{d}x\,\mathrm{d}y + C\,(\mathrm{d}x)^2 = 0, \qquad \textit{es decir} \qquad \frac{\mathrm{d}y}{\mathrm{d}x} = \frac{B \pm \sqrt{B^2 - AC}}{A}.$$

Eligiendo ξ, η constantes sobre las dos familias de características, la ecuación se reduce a $u_{\xi\eta} =$ (orden ≤ 1) en el caso hiperbólico; en el parabólico (una sola familia) a $u_{\eta\eta} =$ (orden ≤ 1); y en el elíptico las características son complejas conjugadas y sus partes real e imaginaria dan $u_{\alpha\alpha} + u_{\beta\beta} =$ (orden ≤ 1).

Demostración. Por la regla de la cadena, $u_x = U_\xi\xi_x + U_\eta\eta_x$ y

$$u_{xx} = U_{\xi\xi}\xi_x^2 + 2U_{\xi\eta}\xi_x\eta_x + U_{\eta\eta}\eta_x^2 + U_\xi\xi_{xx} + U_\eta\eta_{xx},$$

y análogamente para u_{xy} y u_{yy}. Al sustituir y agrupar, los términos de segundo orden que multiplican a $U_{\xi\xi}$ son exactamente $A\xi_x^2 + 2B\xi_x\xi_y + C\xi_y^2$, que es la expresión del enunciado; los que multiplican a $U_{\eta\eta}$ son la misma con η, y los términos con ξ_{xx}, η_{xx}, etc., son de orden uno. Ahora, si ξ es constante sobre una curva $y = y(x)$, derivando $\xi(x, y(x)) =$ cte se obtiene $\xi_x + \xi_y\,y' = 0$, o sea $\xi_x/\xi_y = -y'$. Dividiendo $\widetilde{A}$ entre ξ_y^2:

$$\frac{\widetilde{A}}{\xi_y^2} = A\,(y')^2 - 2B\,y' + C,$$

que se anula precisamente en las raíces de la ecuación característica. □

Ejemplo 13.56 (Reducción explícita de una hiperbólica)**.** Sea $u_{xx} + 4u_{xy} + 3u_{yy} = 0$. Aquí $A = 1$, $B = 2$, $C = 3$, luego $B^2 - AC = 4 - 3 = 1 > 0$: hiperbólica en todo el plano. La

ecuación característica es $(y')^2 - 4y' + 3 = 0$, con raíces $y' = 1$ e $y' = 3$; las características son $y - x = \text{cte}$ e $y - 3x = \text{cte}$. Tómese entonces

$$\xi = y - x, \qquad \eta = y - 3x.$$

Con $\xi_x = -1$, $\xi_y = 1$, $\eta_x = -3$, $\eta_y = 1$ resulta

$$u_{xx} = U_{\xi\xi} + 6U_{\xi\eta} + 9U_{\eta\eta}, \quad u_{xy} = -U_{\xi\xi} - 4U_{\xi\eta} - 3U_{\eta\eta}, \quad u_{yy} = U_{\xi\xi} + 2U_{\xi\eta} + U_{\eta\eta},$$

y al combinar: $(1 - 4 + 3)U_{\xi\xi} + (6 - 16 + 6)U_{\xi\eta} + (9 - 12 + 3)U_{\eta\eta} = -4U_{\xi\eta}$. La ecuación es pues $U_{\xi\eta} = 0$, cuya solución general es $U = F(\xi) + G(\eta)$, es decir

$$u(x, y) = F(y - x) + G(y - 3x).$$

Dos ondas viajeras, exactamente como en d'Alembert, pero con velocidades 1 y 1/3 en lugar de $\pm c$.

Ejemplo 13.57 (Una ecuación que cambia de tipo: Tricomi)**.** La ecuación de Tricomi $u_{xx} + y\, u_{yy} = 0$ tiene $A = 1$, $B = 0$, $C = y$, luego $B^2 - AC = -y$. Es elíptica en el semiplano $y > 0$, hiperbólica en $y < 0$ y parabólica sobre la recta $y = 0$. No es una curiosidad: modela el flujo transónico, y la recta $y = 0$ es la barrera del sonido. Que una misma ecuación sea de dos tipos en dos regiones explica por qué el vuelo supersónico fue un problema matemático antes que un problema de ingeniería: a un lado de la barrera la información se propaga instantáneamente por todo el dominio (elíptico), y al otro viaja por características a velocidad finita, con ondas de choque.

Regla: Clasificar y reducir una EDP de segundo orden

1. Identificar A, B, C *con el convenio* $2B$ *para el término cruzado*: es el error número uno.
2. Calcular $B^2 - AC$ y su signo, que puede depender del punto.
3. Resolver $A(y')^2 - 2By' + C = 0$: dos familias de características (hiperbólico), una (parabólico), ninguna real (elíptico).
4. Tomar ξ, η constantes sobre esas familias y sustituir.
5. Leer la forma canónica y, con ella, qué se puede esperar: propagación a velocidad finita y datos de Cauchy en el caso hiperbólico; suavizado instantáneo en el parabólico; datos sólo en la frontera en el elíptico.

13.15. Las demostraciones que faltaban: d'Alembert, energía y unicidad

El teorema 13.24 se enunció sin prueba. Con la forma canónica del apartado anterior la demostración es de tres líneas.

Demostración del teorema 13.24. Para $u_{tt} = c^2 u_{xx}$ se tiene $A = c^2$ (coeficiente de u_{xx}), $B = 0$ y $C = -1$ (coeficiente de u_{tt}), con discriminante $c^2 > 0$: hiperbólica. Las características son $x \pm ct = \text{cte}$, de modo que se sustituye

$$\xi = x - ct, \qquad \eta = x + ct.$$

Entonces $u_x = U_\xi + U_\eta$, $u_t = c(-U_\xi + U_\eta)$, y

$$u_{xx} = U_{\xi\xi} + 2U_{\xi\eta} + U_{\eta\eta}, \qquad u_{tt} = c^2\big(U_{\xi\xi} - 2U_{\xi\eta} + U_{\eta\eta}\big),$$

luego $u_{tt} - c^2 u_{xx} = -4c^2 U_{\xi\eta}$. La ecuación equivale por tanto a $U_{\xi\eta} = 0$, cuya solución general es $U = F(\xi) + G(\eta)$:

$$u(x,t) = F(x - ct) + G(x + ct).$$

Falta ajustar los datos. De $u(x,0) = f(x)$ sale $F(x) + G(x) = f(x)$; de $u_t(x,0) = g(x)$ sale $-cF'(x) + cG'(x) = g(x)$, e integrando entre 0 y x,

$$-F(x) + G(x) = \frac{1}{c}\int_0^x g(s)\,\mathrm{d}s + K$$

con $K = -F(0) + G(0)$ constante. Sumando y restando las dos igualdades:

$$G(x) = \frac{f(x)}{2} + \frac{1}{2c}\int_0^x g + \frac{K}{2}, \qquad F(x) = \frac{f(x)}{2} - \frac{1}{2c}\int_0^x g - \frac{K}{2}.$$

Al evaluar $F(x-ct)+G(x+ct)$ las constantes $K/2$ se cancelan y las dos integrales se combinan en $\frac{1}{2c}\int_{x-ct}^{x+ct} g$, que es la fórmula del enunciado. Recíprocamente, esa expresión es de la forma $F(x-ct)+G(x+ct)$, luego resuelve la ecuación, y cumple los dos datos por construcción. □

Observación 13.58 (Regularidad: la onda no suaviza). De la fórmula se lee que u tiene exactamente la regularidad de f (y una derivada más que g). Si f tiene un pico, el pico sigue ahí para siempre, viajando. Compárese con el calor, donde el dato inicial se vuelve analítico en cuanto $t > 0$. Esta es la diferencia cualitativa fundamental entre lo hiperbólico y lo parabólico, y explica por qué el sonido transmite información y la difusión la destruye.

La unicidad se puede obtener de la fórmula, pero el argumento de *energía* es más robusto: funciona en cualquier dimensión y en dominios acotados, donde no hay fórmula explícita.

Teorema 13.59 (Conservación de la energía y unicidad)**.** *Sea u solución de $u_{tt} = c^2\Delta u$ en $\Omega \times (0,T)$, con $\Omega \subset \mathbb{R}^n$ acotado y $u = 0$ en $\partial\Omega$ (o bien $\partial u/\partial n = 0$). Entonces*

$$E(t) = \frac{1}{2}\int_\Omega \Big(u_t^2 + c^2|\nabla u|^2\Big)\mathrm{d}x$$

es constante en el tiempo. En consecuencia, la solución con datos $u(\cdot,0) = f$, $u_t(\cdot,0) = g$ y frontera dada es única.

Demostración. Derivando bajo el signo integral y usando la ecuación,

$$E'(t) = \int_\Omega (u_t u_{tt} + c^2 \nabla u \cdot \nabla u_t)\mathrm{d}x = \int_\Omega (c^2 u_t\,\Delta u + c^2 \nabla u \cdot \nabla u_t)\mathrm{d}x.$$

La primera identidad de Green (capítulo 11) da $\int_\Omega u_t \Delta u = \int_{\partial\Omega} u_t \frac{\partial u}{\partial n} - \int_\Omega \nabla u_t \cdot \nabla u$, de modo que los dos términos de volumen se cancelan y queda

$$E'(t) = c^2 \int_{\partial\Omega} u_t \, \frac{\partial u}{\partial n} \, \mathrm{d}S.$$

Con condición de Dirichlet, $u = 0$ en la frontera para todo t, luego $u_t = 0$ allí; con condición de Neumann, $\partial u/\partial n = 0$. En ambos casos la integral se anula y $E' \equiv 0$.

Para la unicidad, si u_1, u_2 son dos soluciones con los mismos datos, $w = u_1 - u_2$ resuelve la ecuación con $w(\cdot,0) = 0$, $w_t(\cdot,0) = 0$ y frontera nula, luego $E_w(0) = 0$ y por lo tanto $E_w(t) = 0$ para todo t. Como el integrando es no negativo, $w_t \equiv 0$ y $\nabla w \equiv 0$: w es constante en el tiempo y en el espacio, y vale 0 en $t = 0$. Así $u_1 = u_2$. □

Idea clave

Por qué la energía es la herramienta correcta. El truco es siempre el mismo: multiplicar la ecuación por la derivada temporal (aquí u_t), integrar en el dominio y usar Green para pasar las derivadas espaciales de un lado a otro. Lo que sale es la derivada de una cantidad positiva. Si esa cantidad se conserva, la ecuación es reversible y hay unicidad; si decrece, la ecuación es disipativa. Para el calor el mismo truco con multiplicador u da $\frac{\mathrm{d}}{\mathrm{d}t}\frac{1}{2}\int u^2 = -k\int |\nabla u|^2 \le 0$: la energía *decrece*, y ahí está, en una línea, la irreversibilidad.

13.16. El núcleo del calor, deducido con Fourier

El teorema 13.26 apareció con la gaussiana ya escrita, como si alguien la hubiera adivinado. No hace falta adivinar: la transformada de Fourier del capítulo 31 convierte la ecuación en derivadas parciales en una ecuación diferencial ordinaria trivial.

Demostración del teorema 13.26. Transfórmese en la variable espacial, con $\widehat{u}(\omega,t) = \int_{\mathbb{R}} u(x,t)\mathrm{e}^{-\mathrm{i}\omega x}\mathrm{d}x$. Como la transformada convierte ∂_x en multiplicación por $\mathrm{i}\omega$, se tiene $\widehat{u_{xx}} = -\omega^2\widehat{u}$, y la ecuación $u_t = ku_{xx}$ se vuelve

$$\frac{\partial \widehat{u}}{\partial t}(\omega,t) = -k\omega^2\,\widehat{u}(\omega,t), \qquad \widehat{u}(\omega,0) = \widehat{f}(\omega).$$

Para cada ω fijo esto es una ecuación lineal de primer orden, con solución

$$\widehat{u}(\omega,t) = \widehat{f}(\omega)\,\mathrm{e}^{-k\omega^2 t}.$$

Sólo falta deshacer la transformada. Un producto de transformadas es la transformada de una convolución, de modo que $u = f * G_t$ donde G_t es la función cuya transformada es $\mathrm{e}^{-k\omega^2 t}$. Esa función se calcula completando el cuadrado:

$$G_t(x) = \frac{1}{2\pi}\int_{\mathbb{R}} \mathrm{e}^{-k\omega^2 t}\mathrm{e}^{\mathrm{i}\omega x}\mathrm{d}\omega = \frac{1}{2\pi}\,\mathrm{e}^{-x^2/(4kt)}\int_{\mathbb{R}} \mathrm{e}^{-kt\left(\omega - \frac{\mathrm{i}x}{2kt}\right)^2}\mathrm{d}\omega = \frac{1}{2\pi}\,\mathrm{e}^{-x^2/(4kt)}\sqrt{\frac{\pi}{kt}},$$

donde la última integral es la gaussiana desplazada al plano complejo, cuyo valor $\sqrt{\pi/(kt)}$ se justifica con el teorema de Cauchy del capítulo 30 (el integrando es entero y decae en las bandas horizontales, así que el contorno se puede devolver al eje real). Simplificando,

$$G_t(x) = \frac{1}{\sqrt{4\pi kt}}\,\mathrm{e}^{-x^2/(4kt)},$$

que es el núcleo anunciado. Que G resuelva la ecuación se comprueba además derivando directamente:

$$\frac{\partial G}{\partial t} = G\left(\frac{x^2}{4kt^2} - \frac{1}{2t}\right), \qquad \frac{\partial^2 G}{\partial x^2} = G\left(\frac{x^2}{4k^2t^2} - \frac{1}{2kt}\right),$$

y en efecto $G_t = kG_{xx}$. Finalmente, $\int_{\mathbb{R}} G_t = 1$ para todo $t > 0$ (es una densidad normal de varianza $2kt$), y G_t concentra toda su masa cerca de 0 cuando $t \to 0^+$, de modo que $f * G_t \to f$ en los puntos de continuidad: el dato inicial se recupera. □

Observación 13.60 (La velocidad de propagación es infinita)*.* La fórmula tiene una consecuencia que conviene mirar de frente. Si $f \geq 0$ es continua, no nula y de soporte compacto —pongamos que la temperatura inicial es 0 salvo en un intervalito—, entonces para *todo* x y *todo* $t > 0$,

$$u(x,t) = \int_{\mathbb{R}} \frac{\mathrm{e}^{-(x-y)^2/(4kt)}}{\sqrt{4\pi kt}} f(y)\,\mathrm{d}y > 0,$$

porque el integrando es estrictamente positivo donde f lo es. Es decir: un segundo después de encender una cerilla en Madrid, la ecuación del calor predice que Sídney está más caliente. La cantidad es absurdamente pequeña (del orden de $\mathrm{e}^{-10^{13}}$), pero no es cero, y eso señala el límite físico del modelo: la ecuación del calor es una aproximación de un proceso de transporte que sí tiene velocidad finita. Contrástese con d'Alembert, donde $u(x,t)$ sólo depende de f y g en el intervalo $[x-ct, x+ct]$: propagación estrictamente a velocidad c.

Ejemplo 13.61 (Una barra con un escalón de temperatura)**.** Sea $f = \mathbf{1}_{[-a,a]}$, una barra que está a temperatura 1 en el centro y 0 fuera. Entonces

$$u(x,t) = \int_{-a}^{a} \frac{\mathrm{e}^{-(x-y)^2/(4kt)}}{\sqrt{4\pi kt}}\mathrm{d}y = \frac{1}{2}\left[\operatorname{erf}\left(\frac{a-x}{\sqrt{4kt}}\right) + \operatorname{erf}\left(\frac{a+x}{\sqrt{4kt}}\right)\right],$$

con $\operatorname{erf}(z) = \frac{2}{\sqrt{\pi}}\int_0^z \mathrm{e}^{-s^2}\mathrm{d}s$. Para $a = 1$, $k = 1$ y $x = 0$: en $t = 0{,}01$ el centro está a 0,99999999; en $t = 0{,}25$ a 0,6827; en $t = 1$ a 0,3829; en $t = 4$ a 0,1974. La escala de tiempo del suavizado es $t \sim a^2/k$, que es la razón por la que duplicar el grosor de un aislante multiplica por cuatro el tiempo de respuesta.

13.17. Principios del máximo: elíptico y parabólico, demostrados

El teorema 13.28 se enunció sin prueba, y además falta su hermano parabólico, que es el que da unicidad y comparación para el calor en dominios acotados. Los dos salen del mismo truco.

Demostración del teorema 13.28. Paso 1: el caso estrictamente subarmónico. Supóngase primero $\Delta v > 0$ en todo Ω. Si v tuviera un máximo en un punto interior x_0, allí la matriz hessiana sería semidefinida negativa (capítulo 9), luego todas sus entradas diagonales serían ≤ 0 y por tanto $\Delta v(x_0) = \sum_i \partial_{ii} v(x_0) \leq 0$, contradiciendo $\Delta v > 0$. Así que el máximo de v sobre $\overline{\Omega}$ —que existe por compacidad— está en $\partial\Omega$.

Paso 2: el caso armónico, por perturbación. Sea ahora $\Delta u = 0$ y, para $\varepsilon > 0$,

$$v_\varepsilon(x) = u(x) + \varepsilon |x|^2, \qquad \text{de modo que} \quad \Delta v_\varepsilon = 0 + 2n\varepsilon > 0.$$

Por el paso 1, $\max_{\overline{\Omega}} v_\varepsilon = \max_{\partial\Omega} v_\varepsilon$. Sea $R = \max_{x \in \overline{\Omega}} |x|$, finito porque Ω está acotado. Entonces, para todo $x \in \overline{\Omega}$,

$$u(x) \le v_\varepsilon(x) \le \max_{\partial\Omega} v_\varepsilon \le \max_{\partial\Omega} u + \varepsilon R^2.$$

Haciendo $\varepsilon \to 0^+$ se obtiene $\max_{\overline{\Omega}} u \le \max_{\partial\Omega} u$, y la desigualdad contraria es obvia. Para el mínimo basta aplicar lo anterior a $-u$, que también es armónica.

Unicidad. Si u_1, u_2 son armónicas en Ω, continuas en $\overline{\Omega}$ y coinciden en $\partial\Omega$, entonces $w = u_1 - u_2$ es armónica con $w = 0$ en la frontera; su máximo y su mínimo valen 0, luego $w \equiv 0$. □

Teorema 13.62 (Principio del máximo parabólico)**.** *Sea $\Omega \subset \mathbb{R}^n$ acotado, $Q_T = \Omega \times (0, T]$ y llámese* frontera parabólica *al conjunto*

$$\Gamma_T = \big(\overline{\Omega} \times \{0\}\big) \;\cup\; \big(\partial\Omega \times [0, T]\big),$$

es decir, la tapa inferior más las paredes laterales, pero no *la tapa superior. Si $u_t = k\Delta u$ en Q_T con u continua en $\overline{Q_T}$, entonces*

$$\max_{\overline{Q_T}} u = \max_{\Gamma_T} u, \qquad \textit{y análogamente para el mínimo.}$$

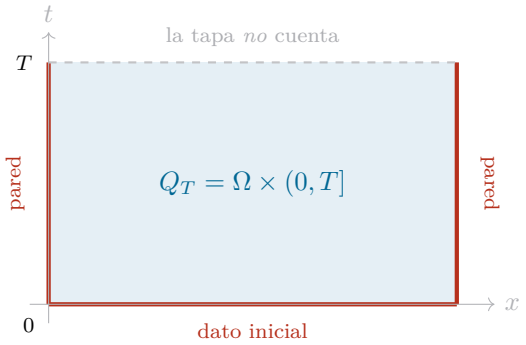


Figura 13.4: La frontera parabólica Γ_T (en rojo) está formada por el dato inicial y las paredes laterales, pero *no* por la tapa superior. El máximo de una solución de la ecuación del calor se alcanza siempre en el rojo. Que la tapa quede fuera no es un tecnicismo: es la irreversibilidad escrita en la geometría del enunciado, pues una barra que se calienta desde los bordes sí alcanza su máximo en el instante final.

Demostración. Sea $\varepsilon > 0$ y $v = u - \varepsilon t$, de modo que

$$v_t - k\Delta v = (u_t - \varepsilon) - k\Delta u = -\varepsilon < 0.$$

Supóngase que v alcanza su máximo sobre $\overline{Q_T}$ en un punto (x_0, t_0) con $x_0 \in \Omega$ y $0 < t_0 \le T$, es decir, fuera de Γ_T. Como x_0 es interior en el espacio, la hessiana espacial en ese punto

es semidefinida negativa y $\Delta v(x_0,t_0)\le 0$. En cuanto al tiempo: si $t_0<T$, el punto es un máximo interior en t y $v_t(x_0,t_0)=0$; si $t_0=T$, la función $t\mapsto v(x_0,t)$ alcanza su máximo en el extremo derecho del intervalo, luego $v_t(x_0,T)\ge 0$. En los dos casos $v_t(x_0,t_0)\ge 0$, y por lo tanto

$$v_t(x_0,t_0)-k\Delta v(x_0,t_0)\ \ge\ 0,$$

contradiciendo que esa expresión vale $-\varepsilon<0$. Así que el máximo de v está en Γ_T, y entonces para todo $(x,t)\in\overline{Q_T}$

$$u(x,t)=v(x,t)+\varepsilon t\le \max_{\Gamma_T} v+\varepsilon T\le \max_{\Gamma_T} u+\varepsilon T.$$

Haciendo $\varepsilon\to 0^+$ se concluye. Para el mínimo, aplíquese lo anterior a $-u$. □

Observación 13.63 (Por qué la tapa superior queda fuera). La exclusión de $\Omega\times\{T\}$ no es un tecnicismo: es la irreversibilidad escrita en la geometría del enunciado. El calor puede alcanzar su máximo en el instante final —una barra que se está calentando desde los extremos lo hace— y en cambio nunca puede aparecer, en el interior y a mitad de camino, un punto más caliente que todo lo que hubo antes y que todo lo que hay en el borde. Comparado con la ecuación de ondas, donde un pico puede formarse en el interior por interferencia de dos frentes, es una restricción severa: la difusión sólo sabe aplanar.

Corolario 13.64 (Unicidad, comparación y dependencia continua)**.** *Sean u_1,u_2 soluciones de $u_t=k\Delta u$ en Q_T, continuas en $\overline{Q_T}$.*

1. *Si coinciden en Γ_T, coinciden en todo $\overline{Q_T}$.*
2. *Si $u_1\le u_2$ en Γ_T, entonces $u_1\le u_2$ en todo $\overline{Q_T}$.*
3. $\max_{\overline{Q_T}}|u_1-u_2|=\max_{\Gamma_T}|u_1-u_2|$*: un error de δ en los datos produce un error de a lo sumo δ en la solución, en cualquier instante. El problema está bien planteado.*

Demostración. Aplíquese el teorema 13.62 a $w=u_1-u_2$, que también resuelve la ecuación. Para (1), $\max w=\max_{\Gamma_T} w=0$ y $\min w=\min_{\Gamma_T} w=0$. Para (2), $\max_{\overline{Q_T}} w=\max_{\Gamma_T} w\le 0$. Para (3), el máximo de $|w|$ es el mayor entre $\max w$ y $-\min w$, y ambos se realizan en Γ_T. □

Ejemplo 13.65 (Una cota sin resolver nada)**.** Una placa ocupa el cuadrado unidad. Inicialmente está a $20\,^\circ$C y sus bordes se mantienen entre 10 y $80\,^\circ$C durante todo el experimento. ¿Cuánto puede llegar a valer la temperatura en el centro al cabo de una hora? Sin resolver ninguna ecuación: los datos sobre Γ_T están entre 10 y 80, luego la solución está entre 10 y 80 en todo punto y todo instante. El principio del máximo es la única herramienta del capítulo que da respuestas cuantitativas sin calcular la solución, y por eso es la que usan de verdad los analistas numéricos: un esquema que produzca $85\,^\circ$C está mal, se mire como se mire.

Cuidado

[La versión fuerte dice bastante más] El principio *débil* demostrado arriba dice que el máximo *se alcanza* en la frontera; el principio *fuerte* (Hopf) dice que si además se alcanza

en un punto interior, entonces la función es constante. Su demostración es más delicada —usa el lema de Hopf sobre la derivada normal en el punto de máximo— y no se incluye aquí; lo que sí conviene retener es la consecuencia: una función armónica no constante *jamás* toca su máximo dentro, y una solución del calor no constante tampoco, ni siquiera tangencialmente. Es el mismo enunciado que en el capítulo 30 se llama principio del módulo máximo, y no es casualidad: la parte real de una función holomorfa es armónica.

13.18. El método de las imágenes: problemas en la semirrecta

Las fórmulas de d'Alembert y del núcleo del calor resuelven el problema en *toda* la recta. En la práctica los dominios tienen bordes: una cuerda con un extremo fijo, una barra aislada por un lado. El método de las imágenes convierte el problema con borde en un problema sin borde, extendiendo el dato de modo que la condición de frontera se cumpla sola por simetría.

Regla: Reflejar según la condición de frontera

1. Si la condición es de *Dirichlet* ($u(0,t)=0$), extiéndase el dato de forma **impar**: $\tilde f(-x) = -\tilde f(x)$. Toda función impar se anula en 0, y la ecuación preserva la imparidad, luego la solución extendida se anula en $x=0$ automáticamente.
2. Si la condición es de *Neumann* ($u_x(0,t)=0$), extiéndase de forma **par**. La derivada de una función par se anula en 0.
3. Resuélvase el problema extendido en toda la recta con la fórmula que corresponda.
4. Restrínjase la solución a $x>0$ y simplifíquese usando la simetría.

La razón por la que esto funciona es que la ecuación es invariante bajo $x \mapsto -x$ (sólo aparecen derivadas de orden par en x), de modo que si el dato es impar la solución es impar en todo instante, y si es par, par.

Teorema 13.66 (Ondas en la semirrecta con extremo fijo)**.** *La solución de $u_{tt}=c^2u_{xx}$ en $x>0$, $t>0$, con $u(x,0)=f(x)$, $u_t(x,0)=g(x)$ y $u(0,t)=0$ (siendo $f(0)=g(0)=0$ para compatibilidad) es*

$$u(x,t)=\begin{cases} \dfrac{f(x+ct)+f(x-ct)}{2}+\dfrac{1}{2c}\displaystyle\int_{x-ct}^{x+ct} g, & ct<x,\\[2mm] \dfrac{f(x+ct)-f(ct-x)}{2}+\dfrac{1}{2c}\displaystyle\int_{ct-x}^{ct+x} g, & ct>x. \end{cases}$$

Demostración. Sean $\tilde f,\tilde g$ las extensiones impares. Por d'Alembert, $u(x,t)=\frac12[\tilde f(x-ct)+\tilde f(x+ct)]+\frac{1}{2c}\int_{x-ct}^{x+ct}\tilde g$. Si $ct<x$, todos los argumentos son positivos y $\tilde f=f$, $\tilde g=g$: sale la primera línea. Si $ct>x$, entonces $x-ct<0$ y $\tilde f(x-ct)=-f(ct-x)$; para la integral, se parte en $\int_{x-ct}^{ct-x}+\int_{ct-x}^{x+ct}$, y la primera es nula porque $\tilde g$ es impar y el intervalo es simétrico respecto del origen. Queda la segunda línea. □

Ejemplo 13.67 (La onda que vuelve del revés)**.** Una cuerda semiinfinita fija en $x = 0$ recibe un pulso triangular centrado en $x = 5$ que viaja hacia el origen, con $c = 1$. Hasta $t = 4$ el pulso avanza intacto. A partir de $t \approx 5$ el pulso alcanza el extremo, y la fórmula dice que lo que emerge es $-f(ct - x)$: el pulso reflejado, con el signo cambiado. Un pulso hacia arriba vuelve hacia abajo. Eso es exactamente lo que se ve en una cuerda atada a una pared, y la razón está en una línea de álgebra: la extensión impar. Si en vez de estar fija la cuerda tuviera el extremo libre (Neumann), la extensión sería par y el pulso volvería con el mismo signo.

Teorema 13.68 (Calor en la semirrecta)**.** *La solución de $u_t = k u_{xx}$ en $x > 0$ con $u(x,0) = f(x)$ es*

$$u(x,t) = \int_0^\infty \big[G(x-y,t) \mp G(x+y,t)\big] f(y)\,\mathrm{d}y,$$

con el signo $-$ si $u(0,t) = 0$ (Dirichlet) y el signo $+$ si $u_x(0,t) = 0$ (Neumann), donde G es el núcleo del calor.

Demostración. Extiéndase f de forma impar (Dirichlet) o par (Neumann) y aplíquese la fórmula de convolución en toda la recta: $u(x,t) = \int_{-\infty}^{\infty} G(x-y,t)\tilde f(y)\mathrm{d}y$. Partiendo la integral en $(-\infty, 0)$ y $(0,\infty)$ y sustituyendo $y \mapsto -y$ en la primera,

$$\int_{-\infty}^0 G(x-y,t)\tilde f(y)\mathrm{d}y = \int_0^\infty G(x+y,t)\tilde f(-y)\mathrm{d}y = \mp\int_0^\infty G(x+y,t) f(y)\mathrm{d}y,$$

que sumada a la otra da el enunciado. La expresión $G(x-y,t) - G(x+y,t)$ se llama *función de Green del semiespacio*: es el núcleo del calor menos el núcleo de una fuente *imagen* colocada en $-y$, que es de donde viene el nombre del método. □

13.19. Ondas en dimensión dos y tres: Kirchhoff, Poisson y Huygens

Se anunció, sin más, que el principio de Huygens vale en dimensión tres y falla en dimensión dos. Aquí están las fórmulas que lo demuestran. Escríbase, para el promedio esférico,

$$\mathcal{M}_r[h](x) = \frac{1}{4\pi r^2}\int_{\partial B(x,r)} h(y)\,\mathrm{d}S(y) \qquad (\text{en } \mathbb{R}^3),$$

que es el *promedio esférico* de h sobre la esfera de centro x y radio r.

Teorema 13.69 (Fórmula de Kirchhoff, $n = 3$)**.** *La solución de $u_{tt} = c^2\Delta u$ en $\mathbb{R}^3$ con $u(\cdot,0) = f$ y $u_t(\cdot,0) = g$ (de clase C^3 y C^2 respectivamente) es*

$$u(x,t) = \frac{1}{4\pi c^2 t^2}\int_{\partial B(x,ct)} \Big[\, t\,g(y) + f(y) + \nabla f(y)\cdot(y-x)\,\Big]\mathrm{d}S(y),$$

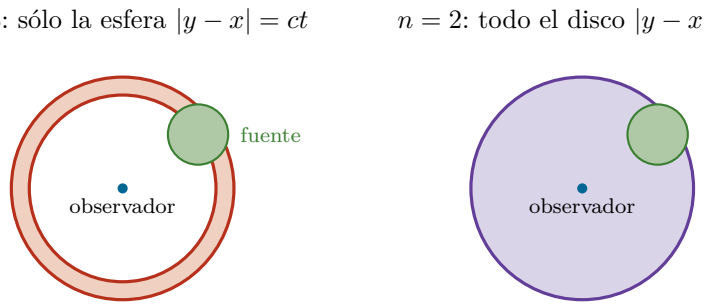


Figura 13.5: Dominio de dependencia en dimensión tres y en dimensión dos. En $\mathbb{R}^3$ la solución en un punto depende sólo de los datos sobre la *esfera* de radio ct: en cuanto la esfera sobrepasa a la fuente, el observador deja de oírla. En $\mathbb{R}^2$ depende de todo el *disco*, que una vez engulle a la fuente ya no la suelta nunca: queda una cola que decae pero no termina. Por eso se puede hablar en el aire y por eso un estanque sigue agitándose.

Idea clave

De dónde sale. Se define el promedio esférico $U(r,t) = \mathcal{M}_r[u(\cdot,t)](x)$ y se comprueba, con el teorema de la divergencia, que U satisface la ecuación de Euler–Poisson–Darboux $U_{tt} = c^2(U_{rr} + \frac{2}{r}U_r)$. El milagro de la dimensión tres es que la sustitución $V = rU$ convierte eso en la ecuación de ondas *en una dimensión*, $V_{tt} = c^2 V_{rr}$, resoluble por d'Alembert en la semirrecta $r > 0$ con $V(0,t) = 0$. Deshaciendo el cambio y haciendo $r \to 0^+$ —donde $U(r,t) \to u(x,t)$ por continuidad— se obtiene la fórmula. La verificación numérica es inmediata: con $f(y) = \sin(k \cdot y)$, $|k| = 1$, $g = 0$ y $c = 1$, la solución exacta es $u = \sin(k \cdot x)\cos t$, y la integral de Kirchhoff evaluada por Monte Carlo en $x = (0{,}4, -0{,}2, 1{,}1)$ da 0,342 frente a 0,342 en $t = 0{,}5$, y $-0{,}354$ frente a $-0{,}352$ en $t = 2{,}7$.

Teorema 13.70 (Fórmula de Poisson, $n = 2$, por descenso)**.** *En* $\mathbb{R}^2$,

$$u(x,t) = \frac{1}{2\pi c^2 t^2} \int_{B(x,ct)} \frac{t\, f(y) + t^2 g(y) + t\, \nabla f(y) \cdot (y - x)}{\left(c^2 t^2 - |y - x|^2\right)^{1/2}} \, \mathrm{d}y,$$

donde ahora la integral es sobre el disco *lleno, no sobre la circunferencia.*

Idea clave

El método del descenso de Hadamard. Un problema en el plano es un problema en el espacio cuyos datos no dependen de la tercera coordenada. Se aplica Kirchhoff en $\mathbb{R}^3$ a $F(y_1, y_2, y_3) = f(y_1, y_2)$, y la integral sobre la esfera se proyecta sobre el disco ecuatorial; el factor $(c^2t^2 - |y-x|^2)^{-1/2}$ es exactamente el jacobiano de esa proyección, y el 2 del denominador cuenta que cada punto del disco recibe dos casquetes, el de arriba y el de abajo.

Observación 13.71 (Huygens, por fin explicado)*.* Compárense los dominios de integración. En $\mathbb{R}^3$, $u(x,t)$ depende de los datos *sólo sobre la esfera* $|y - x| = ct$: una cáscara de espesor nulo. En $\mathbb{R}^2$, depende de los datos sobre *todo el disco* $|y - x| \le ct$. La consecuencia es inmediata: si

el dato tiene soporte en una región K y se observa desde un punto x a distancia d de K, en dimensión tres la señal empieza en $t = d/c$ y *termina* cuando la esfera ha sobrepasado K; en dimensión dos, una vez que el disco toca K ya nunca deja de contenerlo, y la señal no termina jamás: queda una cola decreciente pero eterna.

Por eso se puede hablar y escuchar música: en el aire, tres dimensiones, cada sílaba llega limpia y se va. Y por eso una piedra en un estanque —ondas efectivamente bidimensionales— deja la superficie agitándose mucho después de que pase el frente. El principio de Huygens vale en toda dimensión impar ≥ 3 y falla en toda dimensión par; que vivamos en tres dimensiones espaciales es la condición para que exista la comunicación acústica nítida.

Ejemplo 13.72 (Una explosión localizada)**.** Sea $f = 0$ y g la función que vale 1 en la bola $|y| \leq a$ y 0 fuera, con $c = 1$. Kirchhoff da

$$u(x,t) = t\,\mathcal{M}_t[g](x) = t \cdot \frac{\text{área de } \partial B(x,t) \cap \{|y| \leq a\}}{4\pi t^2}.$$

Para un observador a distancia $R > a$ del origen, la esfera de radio t centrada en él corta la bola sólo cuando $R - a \leq t \leq R + a$. Antes: silencio absoluto. Después: silencio absoluto otra vez. La señal dura exactamente $2a/c$ segundos, la anchura de la fuente dividida por la velocidad, con independencia de lo lejos que esté el observador. En el plano, con la fórmula de Poisson, el integrando nunca vuelve a anularse una vez que el disco engulle la fuente, y el observador oye una cola que decae como t^{-1} para siempre.

13.20. Lax–Milgram y Poincaré, demostrados

Los dos teoremas sobre los que descansa la formulación variacional se enunciaron sin prueba. Las dos demostraciones son cortas y sólo usan el teorema de representación de Riesz del capítulo 32.

Teorema 13.73 (Desigualdad de Poincaré, con constante)**.** *Sea $\Omega \subset \mathbb{R}^n$ contenido en una banda $\{0 < x_1 < d\}$. Entonces, para toda $u \in H^1_0(\Omega)$,*

$$\|u\|_{L^2(\Omega)} \leq d\,\|\nabla u\|_{L^2(\Omega)}.$$

Demostración. Basta probarlo para $u \in C_c^\infty(\Omega)$, porque esas funciones son densas en H^1_0 por definición y ambos lados son continuos para la norma de H^1. Extiéndase u por cero a toda la banda y escríbase $x = (x_1, x')$. Como $u(0, x') = 0$,

$$u(x_1, x') = \int_0^{x_1} \partial_1 u(s, x')\,\mathrm{d}s.$$

Por Cauchy–Schwarz,

$$|u(x_1,x')|^2 \leq x_1 \int_0^{x_1} |\partial_1 u(s,x')|^2 \mathrm{d}s \leq d \int_0^{d} |\partial_1 u(s,x')|^2 \mathrm{d}s.$$

El miembro derecho no depende de x_1; integrando en $x_1 \in (0,d)$,

$$\int_0^d |u(x_1,x')|^2 \mathrm{d}x_1 \leq d^2 \int_0^d |\partial_1 u(s,x')|^2 \mathrm{d}s,$$

e integrando ahora en $x' \in \mathbb{R}^{n-1}$ y usando Fubini,

$$\|u\|_{L^2}^2 \le d^2 \|\partial_1 u\|_{L^2}^2 \le d^2 \|\nabla u\|_{L^2}^2.$$

Tomando raíces se obtiene el enunciado. □

Cuidado

[La hipótesis de anularse en la frontera es esencial] Sin ella la desigualdad es falsa de forma tonta: tómese $u \equiv 1$ en un dominio acotado. Entonces $\nabla u = 0$ pero $\|u\|_{L^2} = |\Omega|^{1/2} \neq 0$, y ninguna constante sirve. Lo que hace funcionar el argumento es disponer de un punto donde u vale 0 para empezar a integrar. La versión sin condición de frontera existe, pero hay que restar la media: $\|u - \bar{u}\|_{L^2} \le C\|\nabla u\|_{L^2}$, y es la desigualdad de Poincaré–Wirtinger.

Demostración del teorema de Lax–Milgram. Paso 1: la forma bilineal define un operador. Fíjese $u \in H$. La aplicación $v \mapsto a(u,v)$ es lineal y acotada, pues $|a(u,v)| \le M\|u\|\,\|v\|$. Por el teorema de representación de Riesz existe un único vector, que se llama Au, con

$$a(u,v) = \langle Au, v\rangle \quad \text{para todo } v \in H, \qquad \|Au\| \le M\|u\|.$$

La unicidad en Riesz garantiza que A es lineal, y la desigualdad, que es continuo.

Paso 2: A es inyectivo con inversa acotada sobre su imagen. La coercividad da

$$\alpha\|u\|^2 \le a(u,u) = \langle Au, u\rangle \le \|Au\|\,\|u\|,$$

luego $\|Au\| \ge \alpha\|u\|$ para todo u. En particular $Au = 0$ implica $u = 0$.

Paso 3: la imagen es cerrada. Si $Au_n \to w$, la desigualdad anterior aplicada a $u_n - u_m$ da $\|u_n - u_m\| \le \alpha^{-1}\|Au_n - Au_m\|$, de modo que (u_n) es de Cauchy; sea $u = \lim u_n$. Por continuidad de A, $Au = w$, así que $w \in \operatorname{im} A$.

Paso 4: la imagen es densa. Sea $w \perp \operatorname{im} A$. En particular $\langle Aw, w\rangle = 0$, y por coercividad $\alpha\|w\|^2 \le \langle Aw, w\rangle = 0$, luego $w = 0$. El complemento ortogonal de la imagen es trivial.

Paso 5: conclusión. Por los pasos 3 y 4, $\operatorname{im} A$ es un subespacio cerrado y denso, es decir, todo H; con el paso 2, A es biyectivo y $\|A^{-1}\| \le 1/\alpha$. Ahora, dado el funcional continuo L, Riesz da un $\ell \in H$ con $L(v) = \langle \ell, v\rangle$ y $\|\ell\| = \|L\|$. El problema $a(u,v) = L(v)$ para todo v equivale a $\langle Au - \ell, v\rangle = 0$ para todo v, o sea a $Au = \ell$, que tiene la solución única $u = A^{-1}\ell$, con

$$\|u\| \le \frac{\|\ell\|}{\alpha} = \frac{\|L\|}{\alpha}.$$

□

Observación 13.74 (Lax–Milgram es Riesz con la simetría rota). Si a fuera además simétrica, sería un producto escalar equivalente al de H (acotada y coerciva significa justamente eso), y el teorema sería literalmente Riesz en ese nuevo producto escalar. El contenido real del teorema es que la simetría *no hace falta*, lo cual importa mucho: las ecuaciones con término de transporte, como $-\Delta u + b \cdot \nabla u = f$, dan formas bilineales no simétricas, y sin Lax–Milgram habría que tratarlas de otro modo.

13.21. Diferencias finitas: consistencia, estabilidad de von Neumann y la condición CFL

Casi ninguna EDP realista se resuelve con fórmula. Se discretiza. Y al discretizar aparece un fenómeno que no tiene análogo en el mundo continuo: un esquema perfectamente razonable puede explotar. Esta sección explica por qué, y da el criterio que lo predice.

Definición 13.75 (Esquema explícito para el calor)**.** Sobre una malla $x_j = j\,\Delta x$, $t_n = n\,\Delta t$, y escribiendo $u_j^n \approx u(x_j, t_n)$, se sustituye u_t por una diferencia hacia adelante y u_{xx} por la diferencia centrada de segundo orden:

$$\frac{u_j^{n+1} - u_j^n}{\Delta t} = k\,\frac{u_{j+1}^n - 2u_j^n + u_{j-1}^n}{\Delta x^2},$$

es decir

$$u_j^{n+1} = u_j^n + r\big(u_{j+1}^n - 2u_j^n + u_{j-1}^n\big), \qquad r = \frac{k\,\Delta t}{\Delta x^2},$$

donde r se llama *número de malla*.

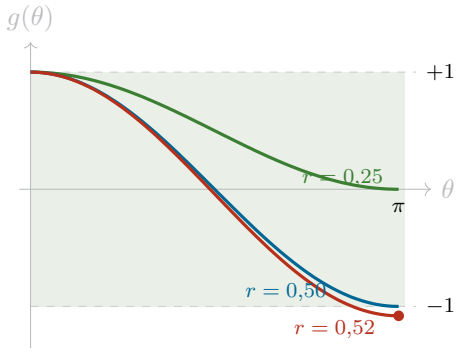


Figura 13.6: Factor de amplificación $g(\theta) = 1 - 4r\sin^2(\theta/2)$ del esquema explícito para el calor. Todo el problema está en el modo más rápido que la malla admite, $\theta = \pi$: mientras $g(\pi) = 1 - 4r$ se mantenga dentro de la banda $[-1, 1]$ (sombreada) el esquema es estable, y en cuanto se sale —punto rojo, $r = 0{,}52$— ese modo se multiplica por 1,08 en cada paso y arruina el cálculo.

Teorema 13.76 (Consistencia)**.** *Si u es de clase C^4, el error de truncamiento local del esquema es $O(\Delta t) + O(\Delta x^2)$.*

Demostración. Por Taylor, $\frac{u(x,t+\Delta t)-u(x,t)}{\Delta t} = u_t + \frac{\Delta t}{2}u_{tt} + O(\Delta t^2)$ y

$$\frac{u(x+\Delta x,t) - 2u(x,t) + u(x-\Delta x,t)}{\Delta x^2} = u_{xx} + \frac{\Delta x^2}{12}u_{xxxx} + O(\Delta x^4),$$

donde los términos impares se cancelan por simetría. Restando y usando $u_t = k u_{xx}$, lo que queda es $\frac{\Delta t}{2}u_{tt} - \frac{k\Delta x^2}{12}u_{xxxx} + \cdots$. □

La consistencia dice que el esquema *aproxima* la ecuación correcta. No dice que funcione.

Regla: Análisis de estabilidad de von Neumann

1. Sustitúyase en el esquema un modo de Fourier discreto $u_j^n = g^n \mathrm{e}^{\mathrm{i}\theta j}$, con $\theta \in [-\pi, \pi]$ la frecuencia adimensional.
2. Despéjese el *factor de amplificación* $g = g(\theta)$, que es lo que multiplica al modo en cada paso.
3. El esquema es estable si y sólo si $|g(\theta)| \le 1$ para todo θ. Si algún $|g(\theta_0)| > 1$, ese modo crece como $|g|^n$ y la solución numérica explota.

El método es legítimo porque el esquema es lineal y de coeficientes constantes: los modos de Fourier son sus autovectores, y basta controlar los autovalores.

Teorema 13.77 (Estabilidad del esquema explícito)**.** *El esquema explícito para el calor es estable si y sólo si*

$$r = \frac{k\,\Delta t}{\Delta x^2} \le \frac{1}{2}.$$

Demostración. Sustitúyase $u_j^n = g^n \mathrm{e}^{\mathrm{i}\theta j}$:

$$g^{n+1}\mathrm{e}^{\mathrm{i}\theta j} = g^n \mathrm{e}^{\mathrm{i}\theta j}\Big[1 + r(\mathrm{e}^{\mathrm{i}\theta} - 2 + \mathrm{e}^{-\mathrm{i}\theta})\Big],$$

y como $\mathrm{e}^{\mathrm{i}\theta} + \mathrm{e}^{-\mathrm{i}\theta} = 2\cos\theta$ y $2\cos\theta - 2 = -4\sin^2(\theta/2)$,

$$g(\theta) = 1 - 4r\,\sin^2\frac{\theta}{2}.$$

Este número es real y decrece desde $g(0) = 1$ hasta $g(\pi) = 1 - 4r$. La condición $|g| \le 1$ para todo θ equivale entonces a $1 - 4r \ge -1$, es decir a $r \le 1/2$. Si $r > 1/2$, el modo $\theta = \pi$ —el diente de sierra $(-1)^j$, la oscilación más rápida que la malla admite— se multiplica por $|1 - 4r| > 1$ en cada paso. □

Ejemplo 13.78 (La explosión, medida)**.** Tómese $k = 1$, $\Delta x = 1/40$ y dato inicial $\sin(\pi x)$ contaminado con un diente de sierra de amplitud 10^{-6}, del tamaño del ruido de redondeo. Con $r = 0{,}5$ la solución numérica decae como debe: $\max|u|$ vale 0,734 tras 100 pasos, 0,539 tras 200, 0,291 tras 400 y 0,157 tras 600. Con $r = 0{,}52$ —un 4 % más de paso temporal— se tiene $|g(\pi)| = 1{,}08$, y la contaminación de 10^{-6} se multiplica por $1{,}08^n$:

pasos	100	200	400	600
$\max\|u\|$	0,727	3,91	$9{,}1 \cdot 10^{6}$	$2{,}4 \cdot 10^{13}$

A los 100 pasos nada parece ir mal. A los 200 la solución ya no significa nada. Y obsérvese que no hay ningún error de programación: el esquema es consistente, converge en el límite formal, y aun así produce basura. La estabilidad es una condición independiente, y hay que comprobarla.

Cuidado

[El precio de la condición $r \le 1/2$] La restricción no es inocua. Como $\Delta t \le \Delta x^2/(2k)$, refinar la malla espacial a la mitad obliga a dividir el paso temporal *por cuatro*, y por tanto a cuadruplicar el número de pasos; el coste total se multiplica por ocho. Ese es el motivo por el que en problemas parabólicos serios se usan esquemas *implícitos*. El de Euler hacia atrás,

$$u_j^{n+1} - r(u_{j+1}^{n+1} - 2u_j^{n+1} + u_{j-1}^{n+1}) = u_j^n,$$

tiene factor de amplificación $g = (1 + 4r\sin^2\frac{\theta}{2})^{-1}$, que cumple $|g| \le 1$ *para todo* $r > 0$: es incondicionalmente estable. El precio es resolver un sistema tridiagonal en cada paso, lo cual, con el algoritmo de Thomas, cuesta $O(N)$ operaciones: barato. El esquema de Crank–Nicolson promedia ambos, con $g = (1 - 2r\sin^2\frac{\theta}{2})/(1 + 2r\sin^2\frac{\theta}{2})$, también incondicionalmente estable y además de orden dos en el tiempo.

Teorema 13.79 (Condición CFL para el transporte)**.** *Para $u_t + a u_x = 0$ con $a > 0$, el esquema* descentrado a barlovento

$$u_j^{n+1} = u_j^n - \nu(u_j^n - u_{j-1}^n), \qquad \nu = \frac{a\,\Delta t}{\Delta x},$$

es estable si y sólo si $0 \le \nu \le 1$. El número ν se llama número de Courant*, y la condición $\nu \le 1$ es la* condición CFL *(Courant–Friedrichs–Lewy).*

Demostración. Con $u_j^n = g^n \mathrm{e}^{\mathrm{i}\theta j}$ se obtiene $g = 1 - \nu(1 - \mathrm{e}^{-\mathrm{i}\theta})$, luego

$$|g|^2 = (1 - \nu + \nu\cos\theta)^2 + \nu^2\sin^2\theta = 1 - 2\nu(1-\nu)(1-\cos\theta).$$

Como $1 - \cos\theta \ge 0$, se tiene $|g| \le 1$ para todo θ exactamente cuando $\nu(1-\nu) \ge 0$, es decir cuando $0 \le \nu \le 1$. □

Idea clave

La CFL, sin cuentas. La solución exacta en (x_j, t_{n+1}) es el dato en $x_j - a\Delta t$. El esquema calcula ese valor a partir de u_j^n y u_{j-1}^n, o sea, a partir del intervalo $[x_j - \Delta x,\ x_j]$. Para que el esquema tenga alguna posibilidad de acertar, el punto del que la ecuación toma la información debe estar *dentro* del intervalo del que el esquema toma la suya:

$$x_j - a\Delta t \ \in\ [x_j - \Delta x,\ x_j] \quad\Longleftrightarrow\quad a\,\Delta t \le \Delta x \quad\Longleftrightarrow\quad \nu \le 1.$$

Enunciado en general: *el dominio de dependencia numérico debe contener al dominio de dependencia de la ecuación.* Es una condición necesaria para cualquier esquema explícito, y para los buenos también es suficiente. Numéricamente, con un pulso gaussiano y 200 pasos: $\nu = 0{,}9$ da $\max|u| = 0{,}958$, $\nu = 1$ da exactamente 1 (el esquema transporta el perfil sin error), $\nu = 1{,}05$ da $1{,}4 \cdot 10^3$ y $\nu = 1{,}2$ da $7 \cdot 10^{23}$.

Teorema 13.80 (Equivalencia de Lax)**.** *Para un problema lineal bien planteado y un esquema* consistente*, se cumple*

$$\textit{estabilidad} \iff \textit{convergencia}.$$

Idea clave

Qué hay que retener. Que un esquema sea consistente —que aproxime la ecuación cuando los pasos tienden a cero— y que converja —que su solución tienda a la verdadera— son cosas distintas, y lo único que separa a la una de la otra es la estabilidad. Como la consistencia se comprueba con Taylor en tres líneas y la convergencia es lo que uno realmente quiere, el teorema de Lax convierte el trabajo en una sola tarea factible: analizar $|g(\theta)|$. Ese es el motivo por el que el análisis de von Neumann, que parece un truco, es en realidad *la* herramienta del área.

Regla: Diseñar un esquema en diferencias finitas y no equivocarse

1. Elegir las diferencias (adelante, atrás, centrada) para cada derivada y comprobar por Taylor el orden de consistencia.
2. Escribir el factor de amplificación sustituyendo $u_j^n = g^n \mathrm{e}^{\mathrm{i}\theta j}$.
3. Imponer $|g(\theta)| \leq 1$ en todo $\theta \in [-\pi, \pi]$; de ahí sale la restricción sobre Δt.
4. Si la restricción resulta cara (caso parabólico, $\Delta t \sim \Delta x^2$), pasar a un esquema implícito y resolver el sistema tridiagonal.
5. Verificar el resultado contra una solución exacta conocida y comprobar que el error decae con el orden previsto al refinar. Si no decae con ese orden, hay un error de programación, no de teoría.

13.22. Curvas características, dominio de dependencia y dominio de influencia

La palabra *característica* ha aparecido ya dos veces en este capítulo con dos significados aparentemente distintos: en la ecuación de transporte designaba las rectas $x - ct = \text{cte}$ por las que viaja la información, y en el teorema 13.55 designaba las curvas que anulan el coeficiente de $u_{\xi\xi}$ al cambiar de variables. Y el "dominio de dependencia" se ha usado en los ejemplos sin definirlo. Las tres cosas son la misma, y conviene decirlo de una vez, porque la definición unificada es la que explica —en el apartado final— por qué existe la condición CFL.

Definición 13.81 (Curva característica)**.** Sea Γ una curva regular del plano y considérese el *problema de Cauchy* consistente en prescribir sobre Γ el valor de u y, si la ecuación es de segundo orden, también el de su derivada normal. La curva Γ se llama *característica* para la ecuación cuando esos datos, junto con la propia ecuación, *no* determinan las derivadas de orden máximo de u sobre Γ; en consecuencia, no permiten prolongar la solución fuera de ella.

Concretamente:

- **Primer orden.** Para $a\,u_x + b\,u_y = c$, los datos sobre Γ dan la derivada tangencial de u y la ecuación da la derivada en la dirección (a, b). El par (u_x, u_y) queda determinado salvo

si (a, b) es tangente a Γ. Las características son, pues, las curvas integrales de $\frac{\mathrm{d}y}{\mathrm{d}x} = b/a$: exactamente las del método de las características.

- **Segundo orden.** Para $A\,u_{xx} + 2B\,u_{xy} + C\,u_{yy} + (\text{orden} \le 1) = 0$, el mismo cálculo con (u_{xx}, u_{xy}, u_{yy}) conduce a un sistema 3×3 cuyo determinante es $A\,(y')^2 - 2B\,y' + C$ evaluado sobre Γ. Las características son las soluciones de la *ecuación característica*

$$A\,(y')^2 - 2B\,y' + C = 0, \qquad \text{es decir} \qquad \frac{\mathrm{d}y}{\mathrm{d}x} = \frac{B \pm \sqrt{B^2 - AC}}{A},$$

que son las mismas curvas del teorema 13.55. Hay dos familias reales si $B^2 - AC > 0$, una sola si $B^2 - AC = 0$ y ninguna real si $B^2 - AC < 0$.

Idea clave

Qué dice esto en cristiano. Una característica es una curva por la que la información *viaja* y a través de la cual no *atraviesa*. Dar datos sobre una curva cualquiera determina la solución en un entorno; darlos justo sobre una característica no determina nada, porque la ecuación, restringida a esa curva, se convierte en una relación entre los propios datos en lugar de en una prescripción para salir de ella. De ahí la otra cara de la moneda: las características son también las únicas curvas por las que puede propagarse una discontinuidad de las derivadas. Una esquina del dato inicial no se queda quieta ni se difumina: se desplaza siguiendo una característica, y por eso el número de familias características —dos, una o ninguna— es lo que separa a los tres tipos de ecuación.

Ejemplo 13.82 (Las características de $u_{tt} = 9u_{xx}$, y las de las otras dos clásicas)**. (a) La onda con $c = 3$.** Escríbase $u_{tt} - 9u_{xx} = 0$ en las variables (x, t), con t haciendo el papel de y. El coeficiente de u_{xx} es $A = -9$, el del término cruzado es $2B = 0$ y el de u_{tt} es $C = 1$, luego $B^2 - AC = 0 - (-9)(1) = 9 > 0$: hiperbólica, dos familias. La ecuación característica es

$$-9\Big(\frac{\mathrm{d}t}{\mathrm{d}x}\Big)^2 + 1 = 0 \iff \frac{\mathrm{d}t}{\mathrm{d}x} = \pm\frac{1}{3} \iff t = \pm\frac{x}{3} + k,$$

o sea

$$x - 3t = \text{cte} \qquad \text{y} \qquad x + 3t = \text{cte}.$$

Son rectas de pendiente $\pm 1/3 = \pm 0{,}333333$ en el plano (x, t), es decir, curvas recorridas a velocidad $\frac{\mathrm{d}x}{\mathrm{d}t} = \pm 3$. El resultado era previsible sin cuentas, porque $\partial_{tt} - 9\partial_{xx} = (\partial_t - 3\partial_x)(\partial_t + 3\partial_x)$: la ecuación es el producto de dos transportes, uno a velocidad $+3$ y otro a velocidad -3, y cada factor aporta su familia. La solución general es $u = F(x - 3t) + G(x + 3t)$.

(b) El calor. Para $u_t = k\,u_{xx}$ se tiene $A = -k$, $B = 0$ y $C = 0$, luego $B^2 - AC = 0$ y la ecuación característica es $-k\,(t')^2 = 0$: una raíz doble, $t' = 0$, y una sola familia, las rectas $t = \text{cte}$. Esa es la traducción geométrica de la propagación instantánea: la única dirección privilegiada es la del instante entero, así que un dato en un punto se reparte de golpe por toda la recta $t = \text{cte}$ siguiente.

(c) Laplace. Para $u_{xx} + u_{yy} = 0$, $B^2 - AC = -1 < 0$ y las raíces de $(y')^2 + 1 = 0$ son $y' = \pm\mathrm{i}$: *ninguna* característica real. No hay ninguna dirección por la que la información viaje, y por eso las funciones armónicas son analíticas y no admiten esquinas propagándose: cualquier irregularidad de la frontera se suaviza instantáneamente hacia el interior.

Definición 13.83 (Dominio de dependencia y dominio de influencia)**.** Fijado un problema de evolución con datos sobre el instante inicial, el *dominio de dependencia* del punto (x_0, t_0) es el conjunto de puntos del dato inicial cuyo valor puede alterar $u(x_0, t_0)$; y el *dominio de influencia* del punto inicial x_0 es el conjunto de puntos (x, t) cuyo valor puede ser alterado por el dato en x_0. Uno es el otro leído al revés.

Para la ecuación de ondas $u_{tt} = c^2 u_{xx}$, la fórmula de d'Alembert (teorema 13.24) los da explícitamente: el dominio de dependencia de (x_0, t_0) es el intervalo

$$D(x_0, t_0) = \big[\, x_0 - ct_0,\ x_0 + ct_0 \,\big],$$

que es el triángulo característico truncado en $t = 0$, y el dominio de influencia de x_0 es el *cono de luz*

$$I(x_0) = \{(x, t) :\ |x - x_0| \le ct,\ t \ge 0\},$$

la región limitada por las dos características que salen de $(x_0, 0)$.

Ejemplo 13.84 (Los tres tipos, medidos)**. (a) Ondas con $c = 3$.** El valor $u(4;\, 2)$ depende *únicamente* de f y g en $[4 - 6,\ 4 + 6] = [-2, 10]$, un intervalo de longitud $2ct_0 = 12$. Cambiar el dato en $x = 20$ no altera nada allí. Y la señal que parte de $x = 10$ tarda $10/3 = 3{,}333333$ unidades de tiempo en llegar al origen: antes de ese instante, $u(0, t)$ no sabe siquiera que existía. Con $f(x) = \mathrm{e}^{-x^2}$ y $g = 0$, por ejemplo,

$$u(0;\, 1) = \tfrac{1}{2}\big(\mathrm{e}^{-9} + \mathrm{e}^{-9}\big) = \mathrm{e}^{-9} = 1{,}23410 \cdot 10^{-4},$$

y las dos mitades del pulso ya están en $x = \pm 3$: el centro casi se ha vaciado.

(b) Calor: dependencia total. El núcleo $G(x, t) = (4\pi kt)^{-1/2}\mathrm{e}^{-x^2/(4kt)}$ es estrictamente positivo para todo x en cuanto $t > 0$, así que el dominio de dependencia de *cualquier* punto es *toda* la recta, y el dominio de influencia de cualquier punto es todo el semiplano $t > 0$. La magnitud, eso sí, es ridícula: con $k = 1$, $t = 0{,}001$ y $x = 10$ se obtiene $G = 3{,}88 \cdot 10^{-10857}$. No es cero —y esa es toda la diferencia matemática—, pero está unos diez mil órdenes de magnitud por debajo de cualquier cosa medible, que es la razón por la que el modelo sigue siendo utilísimo pese a violar la relatividad.

(c) Laplace: ni dependencia ni influencia, porque no hay tiempo. Aquí no existe instante inicial y la pregunta cambia de forma: el valor en un punto interior depende de *toda* la frontera, y de toda ella con peso estrictamente positivo, como muestra el núcleo de Poisson. Modificar los datos en un arco arbitrariamente pequeño de $\partial\Omega$ modifica la solución en todos los puntos de Ω a la vez. Hablar de velocidad de propagación en una ecuación elíptica no tiene sentido: no describe un proceso, describe un equilibrio ya alcanzado.

Observación 13.85 (Esto es exactamente lo que mide la condición CFL)*.* La definición anterior deja de ser un adorno geométrico en cuanto se discretiza. Un esquema explícito calcula u_j^{n+1} a partir de un puñado de vecinos del nivel anterior, así que tiene su propio dominio de dependencia: tras n pasos con una plantilla de un vecino a cada lado, el valor u_j^n depende únicamente de los nodos iniciales del intervalo $[\, x_j - n\Delta x,\ x_j + n\Delta x \,]$, de semianchura $n\Delta x$. La ecuación, en cambio, exige la semianchura $c\,t_n = c\,n\Delta t$. Si el intervalo numérico no contiene al

analítico, el esquema está ignorando datos que sí determinan la respuesta, y ninguna cantidad de refinamiento lo salvará: no puede converger a lo que ni siquiera mira. La condición

$$c\,n\Delta t \le n\Delta x \quad\Longleftrightarrow\quad \nu = \frac{c\,\Delta t}{\Delta x} \le 1$$

es precisamente la CFL del teorema 13.79, aquí deducida sin ningún análisis de Fourier: *el dominio de dependencia numérico debe contener al de la ecuación.*

Con $c = 3$ y $\Delta x = 0{,}01$ la restricción es $\Delta t \le 0{,}01/3 = 0{,}003333$. Tomando el valor límite $\Delta t = 1/300$ hacen falta 300 pasos para llegar a $t = 1$, y ambas semianchuras coinciden exactamente: $300 \cdot 0{,}01 = 3 = c \cdot 1$. Tomando en cambio $\Delta t = 0{,}004$, apenas un 20 % más, resulta $\nu = 1{,}2$; en 250 pasos se alcanza el mismo $t = 1$, pero la semianchura numérica es solo $250 \cdot 0{,}01 = 2{,}5$, frente a los 3 que exige la ecuación. Todo lo que ocurriera inicialmente en $2{,}5 < |x - x_0| \le 3$ influye en la respuesta verdadera y el esquema jamás lo lee. Conviene subrayar la jerarquía lógica: la CFL así entendida es una condición *necesaria* para cualquier esquema explícito, sea cual sea su plantilla; que además sea *suficiente* depende del esquema concreto y eso sí hay que comprobarlo con von Neumann.

13.23. Ejercicios resueltos

Ejercicio 13.1. Clasifica $u_{xx} + 4u_{xy} + 3u_{yy} = 0$ y resuélvela.

Solución. Aquí $A = 1$, $B = 4$, $C = 3$, luego $\Delta = 16 - 12 = 4 > 0$: hiperbólica. Factorizando el operador, $\partial_{xx} + 4\partial_{xy} + 3\partial_{yy} = (\partial_x + \partial_y)(\partial_x + 3\partial_y)$, de modo que la solución general es $u = F(y - x) + G(y - 3x)$ con F y G arbitrarias de clase C^2. Las características son las rectas $y - x =$ constante y $y - 3x =$ constante. ∎

Ejercicio 13.2. Una barra $[0, \pi]$ con $k = 1$ y extremos a cero parte de $u(x, 0) = 3\sin x - \sin(4x)$. Halla la temperatura en $x = \frac{\pi}{2}$ para $t = 1$.

Solución. El dato ya está desarrollado en autofunciones, así que la solución es $u = 3\sin(x)\mathrm{e}^{-t} - \sin(4x)\mathrm{e}^{-16t}$. En $x = \frac{\pi}{2}$ se tiene $\sin\frac{\pi}{2} = 1$ y $\sin 2\pi = 0$, luego $u = 3\mathrm{e}^{-1} = 1{,}103638$. El segundo modo no contribuye por estar en un nodo, y de todos modos su amplitud ya se habría reducido por el factor $\mathrm{e}^{-16} = 1{,}125 \cdot 10^{-7}$. ∎

Ejercicio 13.3. Resuelve $u_{tt} = u_{xx}$ con $u(x, 0) = 0$ y $u_t(x, 0) = 1$ si $|x| < 1$ y 0 en otro caso.

Solución. Por d'Alembert con $f = 0$, la solución es $u = \frac{1}{2}\int_{x-t}^{x+t} g$, es decir la mitad de la longitud de la intersección de $[x - t, x + t]$ con $[-1, 1]$. En el origen, para $t \le 1$ el intervalo $[-t, t]$ está dentro y $u(0, t) = t$; para $t \ge 1$ la intersección es todo $[-1, 1]$, de longitud 2, y $u(0, t) = 1$ constante. La cuerda sube linealmente y se queda quieta: es el rastro que deja una onda plana en dimensión uno, donde no hay principio de Huygens. ∎

Ejercicio 13.4. Comprueba que $u(x, y) = \mathrm{e}^x \cos y$ es armónica y localiza sus extremos en el cuadrado $[0, 1] \times [0, \pi]$.

Solución. Se tiene $u_{xx} = \mathrm{e}^x \cos y$ y $u_{yy} = -\mathrm{e}^x \cos y$, cuya suma es cero. Por el principio del máximo, los extremos están en la frontera. Sobre $y = 0$ la función es e^x, que alcanza su máximo $\mathrm{e} = 2{,}718282$ en $x = 1$; sobre $y = \pi$ es $-\mathrm{e}^x$, con mínimo $-\mathrm{e}$ en $x = 1$. Sobre $x = 0$ vale $\cos y \in [-1, 1]$ y sobre $x = 1$ vale $\mathrm{e}\cos y$. El máximo absoluto es e en $(1, 0)$ y el mínimo $-\mathrm{e}$ en $(1, \pi)$, ambos en el borde, como debía ser. ∎

Ejercicio 13.5. Resuelva la ecuación de ondas $u_{tt} = c^2 u_{xx}$ en toda la recta con $u(x,0) = f(x)$ y $u_t(x,0) = 0$, y explique qué dice la fórmula sobre la velocidad de propagación.

Solución. El cambio $\xi = x - ct$, $\eta = x + ct$ convierte la ecuación en $u_{\xi\eta} = 0$, cuya solución general es $u = F(\xi) + G(\eta)$. Imponiendo $u(x,0) = F(x) + G(x) = f(x)$ y $u_t(x,0) = -cF'(x) + cG'(x) = 0$, la segunda da $F' = G'$, y resulta la fórmula de d'Alembert

$$u(x,t) = \tfrac{1}{2}\big(f(x-ct) + f(x+ct)\big).$$

El perfil inicial se parte en dos mitades que viajan en sentidos opuestos a velocidad c. La consecuencia física es que el valor de u en (x,t) depende *sólo* de f en los dos puntos $x \pm ct$: la información no viaja más rápido que c, y un dato inicial concentrado nunca se siente antes de tiempo. Compárese con el calor, donde eso falla. ■

Ejercicio 13.6. Resuelva $u_t = \alpha u_{xx}$ en $[0,L]$ con $u(0,t) = u(L,t) = 0$ y $u(x,0) = f(x)$ por separación de variables, y diga qué le ocurre a la solución cuando $t \to \infty$.

Solución. Ensayando $u = X(x)T(t)$ queda $T'/(\alpha T) = X''/X = -\lambda$. Las condiciones de contorno obligan a $X_n = \sin(n\pi x/L)$ con $\lambda_n = (n\pi/L)^2$, y entonces $T_n = \mathrm{e}^{-\alpha\lambda_n t}$. Superponiendo,

$$u(x,t) = \sum_{n\geq 1} b_n \,\mathrm{e}^{-\alpha(n\pi/L)^2 t} \sin\frac{n\pi x}{L}, \qquad b_n = \frac{2}{L}\int_0^L f(x)\sin\frac{n\pi x}{L}\,\mathrm{d}x.$$

Cada armónico decae exponencialmente, y tanto más deprisa cuanto mayor es n. En consecuencia, las irregularidades finas del dato inicial desaparecen casi de inmediato y la solución se vuelve infinitamente derivable para todo $t > 0$ por áspera que fuese f. Cuando $t \to \infty$, $u \to 0$: el calor se escapa por los extremos. Ese suavizado instantáneo es también el motivo de que la ecuación del calor no se pueda resolver hacia atrás en el tiempo. ■

Ejercicio 13.7. Resuelva el problema de Dirichlet $\Delta u = 0$ en el disco unidad con $u = g(\theta)$ en el borde, y deduzca la propiedad del valor medio.

Solución. En polares, $\Delta u = u_{rr} + \frac{1}{r}u_r + \frac{1}{r^2}u_{\theta\theta}$. Separando variables se obtienen las soluciones $r^n\cos n\theta$ y $r^n \sin n\theta$, y superponiendo,

$$u(r,\theta) = \frac{a_0}{2} + \sum_{n\geq 1} r^n\big(a_n\cos n\theta + b_n \sin n\theta\big),$$

donde a_n, b_n son los coeficientes de Fourier de g. Evaluando en $r = 0$ sobrevive sólo el primer término, y $a_0/2 = \frac{1}{2\pi}\int_0^{2\pi} g(\theta)\mathrm{d}\theta$: el valor en el centro es el promedio de los valores en el borde.

De ahí sale casi todo lo demás. El principio del máximo, porque una función que promedia no puede tener un máximo interior estricto. La unicidad, porque la diferencia de dos soluciones es armónica y nula en el borde, luego nula. Y la analiticidad, porque la serie converge y define una función analítica en el interior. Las ecuaciones elípticas premian con regularidad todo lo que las hiperbólicas se niegan a dar. ■

13.24. Ejercicios propuestos

13.1. Clasifique en hiperbólica, parabólica o elíptica: $u_{xx} - 4u_{yy} = 0$, $u_{xx} + u_{yy} = 0$, $u_{xx} - u_y = 0$ y $u_{xx} + 4u_{xy} + 4u_{yy} = 0$.

13.2. Escriba la solución de d'Alembert para $f(x) = \mathrm{e}^{-x^2}$ y $u_t(x,0) = 0$ con $c = 2$.

13.3. ¿Cuáles son las características de $u_{tt} = 9u_{xx}$?

13.4. Resuelva $u_t = u_{xx}$ en $[0, \pi]$ con extremos a cero y $u(x, 0) = 3 \sin 2x$.

13.5. Para esa solución, ¿en qué instante la amplitud se ha reducido a la mitad?

13.6. Halle los autovalores y autofunciones de $X'' + \lambda X = 0$ con $X(0) = X(L) = 0$.

13.7. ¿Por qué la ecuación del calor no se puede resolver hacia atrás en el tiempo?

13.8. Verifique que $u(x, t) = \frac{1}{\sqrt{4\pi\alpha t}} \mathrm{e}^{-x^2/(4\alpha t)}$ resuelve $u_t = \alpha u_{xx}$ y diga qué representa.

13.9. Enuncie el principio del máximo para funciones armónicas y dé una consecuencia sobre la unicidad.

13.10. ¿Es armónica $u(x, y) = x^2 - y^2$? ¿Y $x^2 + y^2$? Relaciónelo con la parte real de una función holomorfa.

13.11. Escriba la ecuación de Laplace en coordenadas polares.

13.12. Explique por qué el problema de Cauchy para la ecuación de Laplace está mal planteado, y cite el ejemplo de Hadamard.

Capítulo 14

Métodos numéricos: cuando no hay fórmula

Lo que necesitas traer

Del capítulo 6: Taylor y el teorema del valor medio, de donde salen todas las cotas de error. Del capítulo 7: eliminación gaussiana, normas y el número de condición. Del capítulo 12: los problemas de valor inicial que resuelven los métodos de Runge–Kutta.

Los capítulos anteriores enseñan a resolver problemas *exactamente*: despejar la incógnita, hallar la primitiva, escribir la solución. Este capítulo empieza reconociendo algo incómodo: la inmensa mayoría de los problemas reales no se pueden resolver así, y no por falta de habilidad, sino porque se ha demostrado que la fórmula no existe. Lo que sí existe siempre es un procedimiento para acercarse tanto como se quiera. Ese cambio de ambición —de la respuesta exacta a la respuesta con error controlado— es el nacimiento del cálculo numérico, y con él el de casi todo lo que hoy hace un ordenador cuando se le pide "resolver" algo.

Idea clave

Intuición primero. Suponga que quiere hallar $\sqrt{2}$ y no dispone de la tecla. Pruebe con 1,5: al cuadrado da 2,25, se ha pasado. Pruebe con 1,4: da 1,96, se ha quedado corto. La respuesta está entre ambos. Repita partiendo el intervalo por la mitad y ya tiene un método, tosco pero infalible.

Todo el capítulo es una elaboración de esa idea, con dos preguntas siempre presentes: ¿me estoy acercando de verdad?, y ¿cómo de cerca estoy ya? La segunda pregunta es la difícil, y es la que separa un cálculo aproximado de una simple corazonada.

14.1. Por qué hacen falta

Tres ejemplos, de tres tipos distintos de imposibilidad:

La ecuación $x^3 - x - 1 = 0$ tiene una raíz real perfectamente definida, cercana a 1,3247. Existe una fórmula de Cardano para las cúbicas, pero para el grado cinco no hay ninguna: el teorema de Abel–Ruffini, que se demuestra en el capítulo 26, dice que *ninguna* fórmula con

radicales puede resolver la ecuación general de quinto grado. La imposibilidad es un teorema, no una limitación técnica.

La ecuación $\cos x = x$ tiene también una solución única y evidentísima si se dibujan las dos gráficas, y sin embargo no hay manera de despejar x: la incógnita aparece dentro y fuera de una función trascendente.

Y la integral $\int_0^1 \mathrm{e}^{-x^2}\,\mathrm{d}x$, que es el corazón del cálculo de probabilidades del capítulo 16, tiene un valor perfectamente definido y ninguna primitiva expresable con funciones elementales. Eso también es un teorema, el de Liouville.

Observación 14.1. Conviene tener claro qué se está aceptando y qué no. No se renuncia a la existencia ni a la unicidad de la solución —esas se demuestran con las herramientas de siempre, típicamente con el teorema de Bolzano— sino solo a escribirla en forma cerrada. La solución existe; lo que se calcula es una aproximación con error acotado, que para cualquier uso práctico es indistinguible de tenerla.

14.2. Cuánto vale «aproximadamente»: error absoluto, error relativo y cifras significativas

Todo este capítulo consiste en dar respuestas que no son exactas, y lo que separa un cálculo numérico de una corazonada es que en el primero se sabe *cuánto* se ha fallado. Antes de ningún método, pues, hace falta el vocabulario para medir el fallo. Es un vocabulario de tres palabras y se usará en todas las páginas siguientes.

Definición 14.2 (Error absoluto, relativo y porcentual)**.** Sea x el valor exacto de una cantidad y $\tilde{x}$ una aproximación suya. Se llaman

$$E_{\mathrm{abs}} = |\tilde{x} - x|, \qquad E_{\mathrm{rel}} = \frac{|\tilde{x} - x|}{|x|} \quad (x \neq 0), \qquad E_{\mathrm{pct}} = 100\, E_{\mathrm{rel}}$$

el *error absoluto*, el *error relativo* y el *error porcentual* (en tanto por ciento) de la aproximación. La cantidad $\tilde{x} - x$, con su signo, es el *error* propiamente dicho; cuando es positivo se dice que la aproximación está *por exceso* y cuando es negativo, *por defecto*.

Idea clave

Qué dice esto en cristiano. El error absoluto responde a «¿cuánto me he desviado?» y se mide en las mismas unidades que la cantidad: metros, segundos, euros. El relativo responde a «¿cuánto me he desviado *en proporción*?» y no tiene unidades, porque es un cociente entre dos cantidades homogéneas. La diferencia no es un matiz: equivocarse en un metro midiendo la distancia a la Luna es un prodigio de precisión, y equivocarse en un metro midiendo una mesa es un desastre. El error absoluto vale lo mismo en los dos casos; el relativo, no. Por eso todo lo que este capítulo dice sobre cifras perdidas, condicionamiento y estabilidad se formula siempre en términos relativos.

Ejemplo 14.3 (Los mismos números, dos lecturas)**. (a) Aproximando π.** El valor exacto es $\pi = 3{,}14159265$. La célebre aproximación $\frac{22}{7} = 3{,}14285714$ tiene

$$E_{\text{abs}} = 0{,}00126449, \qquad E_{\text{rel}} = 0{,}00040250, \qquad E_{\text{pct}} = 0{,}0402.$$

La aproximación 3,14 tiene $E_{\text{abs}} = 0{,}00159265$ y $E_{\text{rel}} = 0{,}00050696$: es peor que $\frac{22}{7}$, aunque parezca más inocente. Y 3,1416 tiene $E_{\text{abs}} = 0{,}0000073464$ y $E_{\text{rel}} = 0{,}0000023384$.

(b) La escala lo cambia todo. Medir una mesa de 2 metros con un error de 1 centímetro da $E_{\text{rel}} = 0{,}005$, o sea un 0,5 %. Medir la distancia Madrid–Barcelona, 505 000 metros, con el mismo error de 1 centímetro da $E_{\text{rel}} = 1{,}98 \cdot 10^{-8}$. El error absoluto es idéntico y la calidad de las dos medidas no tiene nada que ver.

(c) Cuando el relativo no sirve. Si el valor exacto es $x = 0$, el error relativo no está definido, y si x está muy cerca de cero, se dispara aunque el resultado sea excelente. Por eso los criterios de parada de los métodos de este capítulo miden $|x_{k+1} - x_k|$ y no $|f(x_k)|/|f(x_0)|$: cerca de una raíz, exigir un error relativo pequeño en el valor de la función es exigir un imposible.

Con el error relativo bien definido se puede por fin decir con precisión qué significa «esta aproximación tiene tres cifras buenas», frase que el capítulo usará constantemente al comparar métodos.

Definición 14.4 (Cifras significativas correctas)**.** Las *cifras significativas* de un número escrito en decimal son todas sus cifras a partir de la primera no nula, contando desde la izquierda: 0,004 520 tiene cuatro cifras significativas (4, 5, 2 y 0), y 3 700 tiene dos, tres o cuatro según lo que la notación pretenda afirmar (por eso se escribe $3{,}7 \cdot 10^3$ cuando se quiere ser inequívoco).

Sea ahora $x \neq 0$ con $10^e \leq |x| < 10^{e+1}$. Se dice que la aproximación $\tilde{x}$ tiene d *cifras significativas correctas* si el error absoluto no supera media unidad en la posición de la cifra d-ésima:

$$|\tilde{x} - x| \;\leq\; \tfrac{1}{2} \cdot 10^{\,e-d+1}.$$

Proposición 14.5. *Si $\tilde{x}$ tiene d cifras significativas correctas, entonces $E_{\text{rel}} \leq 5 \cdot 10^{-d}$. Recíprocamente, si $E_{\text{rel}} \leq 5 \cdot 10^{-(d+1)}$, entonces $\tilde{x}$ tiene al menos d cifras significativas correctas.*

Demostración. Escribamos $10^e \leq |x| < 10^{e+1}$. Para la ida, de $|\tilde{x} - x| \leq \frac{1}{2}10^{e-d+1}$ y $|x| \geq 10^e$ se sigue

$$E_{\text{rel}} = \frac{|\tilde{x} - x|}{|x|} \leq \frac{\frac{1}{2}10^{e-d+1}}{10^e} = \tfrac{1}{2} \cdot 10^{1-d} = 5 \cdot 10^{-d}.$$

Para la vuelta, de $E_{\text{rel}} \leq 5 \cdot 10^{-(d+1)}$ y $|x| < 10^{e+1}$ se obtiene

$$|\tilde{x} - x| = E_{\text{rel}}|x| < 5 \cdot 10^{-d-1} \cdot 10^{e+1} = 5 \cdot 10^{e-d} = \tfrac{1}{2} \cdot 10^{\,e-d+1},$$

que es la condición de la definición 14.4. □

Observación 14.6. El hueco de una cifra entre la ida y la vuelta no es un descuido de la demostración: es inevitable, porque un número puede empezar por 1 o por 9 y eso cambia en un factor 10 el peso relativo de la misma cifra. La consecuencia práctica es la regla de bolsillo que se usa en todo el capítulo: *un error relativo del orden de* 10^{-d} *corresponde a unas* d *cifras correctas*, con una cifra de incertidumbre en el recuento.

Ejemplo 14.7 (Contar cifras correctas de verdad)**.** Tomamos $x = \pi$, para el que $e = 0$, de modo que la condición es $|\tilde{x} - \pi| \leq \frac{1}{2} \cdot 10^{1-d}$.

$\tilde{x}$	E_{abs}	E_{rel}	cifras correctas
3,14	0,00159265	0,00050696	3
$22/7 = 3{,}142857$	0,00126449	0,00040250	3
3,141	0,00059265	0,00018865	3
3,1416	0,0000073464	0,0000023384	5
3,14159	0,0000026536	0,00000084466	6

Lectura de la primera fila: $\frac{1}{2} \cdot 10^{1-3} = 0{,}005$ y el error es 0,00159, menor, luego hay tres cifras; con $d = 4$ haría falta un error por debajo de 0,0005 y no lo es. Obsérvese la cuarta fila: 3,1416 tiene cinco cifras correctas pese a estar redondeada por exceso en la última, porque redondear al más próximo es precisamente lo que garantiza el «media unidad» de la definición. Y obsérvese la trampa de la segunda fila: $\frac{22}{7}$ escribe muchos decimales (3,142857…) y solo tres son de fiar. Escribir cifras que no se tienen es la manera más extendida de mentir con números correctos.

Regla: decidir cuántas cifras dar en un resultado

1. **Estima el error absoluto** del resultado por el camino que sea: la cota del método (bisección, cuadratura), la diferencia entre dos iteraciones consecutivas, o el condicionamiento del problema.
2. **Pásalo a relativo** dividiendo por el valor obtenido, salvo que ese valor esté cerca de cero, en cuyo caso quédate con el absoluto.
3. **Cuenta las cifras** con la proposición 14.5: unas d cifras si el error relativo es del orden de 10^{-d}.
4. **Escribe solo esas cifras y ni una más**, redondeando al más próximo. Añadir decimales que el cálculo no sostiene no es más información: es información falsa.
5. **Al operar, no redondees los pasos intermedios.** Se arrastran todas las cifras que dé la máquina y solo se redondea al final; redondear por el camino introduce errores que se acumulan.
6. **Verifica** recalculando con más precisión o por otra vía: si las cifras que declaras buenas cambian, la estimación del error estaba mal.

Cuidado

No hay que confundir las cifras *decimales* con las *significativas*: 0,00031 tiene cinco decimales y dos cifras significativas, y 31 000 tiene cero decimales y dos cifras significativas. Los métodos de este capítulo controlan unas veces unas y otras veces otras: la bisección acota el error *absoluto* y por tanto los decimales, mientras que el número de condición del apartado final habla de error *relativo* y por tanto de significativas. Segundo aviso: la aritmética del ordenador trabaja siempre en relativo, porque el punto flotante guarda un número fijo de cifras significativas y no un número fijo de decimales.

14.3. El orden de convergencia: qué se está diciendo exactamente al llamar a un método lineal, cuadrático o de cuarto orden

El capítulo compara métodos todo el rato con frases como «gana una cifra binaria por paso», «duplica las cifras correctas», «converge con orden 1,618» o «es de cuarto orden». Todas esas frases tienen una definición precisa detrás, y de hecho tienen *dos* definiciones distintas que conviene no mezclar: una para los métodos que generan una sucesión, y otra para los que dependen de un paso h.

Definición 14.8 (Orden de convergencia de una sucesión)**.** Sea (x_k) una sucesión que converge a p, con errores $e_k = x_k - p$ no nulos. Se dice que la convergencia tiene *orden* $q \geq 1$ con *constante asintótica* $C > 0$ si

$$\lim_{k\to\infty} \frac{|e_{k+1}|}{|e_k|^q} = C.$$

Los casos con nombre propio son: $q = 1$ con $0 < C < 1$, *convergencia lineal*; $q = 2$, *convergencia cuadrática*; $q = 3$, *cúbica*. Si $|e_{k+1}|/|e_k| \to 0$ sin que haya un $q > 1$ con límite finito no nulo, se habla de convergencia *superlineal*.

Idea clave

Qué dice esto en cristiano. El orden mide cómo se transforma el error de un paso al siguiente. Con orden 1 el error se multiplica por un factor fijo: si $C = 0{,}5$, cada paso gana una cifra binaria, y hacen falta algo más de tres pasos por cifra decimal. Con orden 2 el error se *eleva al cuadrado*: si tienes tres cifras correctas, el paso siguiente te da seis, y el siguiente doce. Y con orden 1,618, el número de cifras se multiplica por 1,618 cada vez. Traducido a decimales correctos $d_k \approx -\log_{10}|e_k|$, la relación es $d_{k+1} \approx q\, d_k - \log_{10} C$: el orden es el factor por el que se multiplican las cifras buenas.

Ejemplo 14.9 (Medir el orden de Newton sobre un caso concreto)**.** Tomamos $f(x) = x^3 - x - 1$, la ecuación del apartado anterior, cuya raíz es $p = 1{,}324717957244746$, y arrancamos Newton en $x_0 = 1{,}5$. Los errores $e_k = x_k - p$ y los cocientes de la definición son:

k	$\lvert e_k\rvert$	$\lvert e_{k+1}\rvert/e_k^2$
0	$1{,}75282\cdot 10^{-1}$	0,752125
1	$2{,}31081\cdot 10^{-2}$	0,903473
2	$4{,}82442\cdot 10^{-4}$	0,931275
3	$2{,}16754\cdot 10^{-7}$	0,931886
4	$4{,}37823\cdot 10^{-14}$	0,931886
5	$1{,}78632\cdot 10^{-27}$	

La tercera columna se estabiliza en 0,9318865. El teorema 14.63 predice que ese límite debe ser

$$\left|\frac{f''(p)}{2f'(p)}\right| = \frac{6p}{2(3p^2-1)} = \frac{3\cdot 1{,}3247180}{3\cdot 1{,}7548777-1} = 0{,}9318865,$$

y coincide hasta la séptima cifra ✓. Obsérvese también la columna de errores: los exponentes van -1, -2, -4, -7, -14, -27, aproximadamente duplicándose. Eso es la convergencia cuadrática vista con los ojos.

Regla: estimar el orden a partir de una tabla de errores

Si se conocen tres errores consecutivos, el orden se despeja sin saber nada del método.

1. **Consigue el valor exacto o uno mucho más preciso** (por ejemplo, iterando bastante más y tomando el último valor como si fuera exacto) y calcula los errores $|e_k|$.
2. **Aplica la fórmula**
$$q \;\approx\; \frac{\ln(|e_{k+1}|/|e_k|)}{\ln(|e_k|/|e_{k-1}|)},$$
que sale de suponer $|e_{k+1}| \approx C|e_k|^q$ en dos pasos consecutivos y dividir para eliminar C.
3. **Repite con varias ternas** y mira si el valor se estabiliza; una sola terna no dice nada, porque el régimen asintótico tarda unos pasos en instalarse.
4. **Descarta las últimas ternas** si los errores han bajado al nivel del redondeo de la máquina: ahí la tabla mide ruido, no convergencia.
5. **Estima después la constante** con $C \approx |e_{k+1}|/|e_k|^q$.
6. **Contrasta con lo esperado**: si el método debía ser cuadrático y sale $q \approx 1$, lo más probable es que la raíz sea múltiple o que haya un error de programación.

Ejemplo 14.10 (El orden de la secante, medido y no supuesto)**.** Con la misma $f(x) = x^3 - x - 1$ y los arranques $x_0 = 1$, $x_1 = 2$, la secante produce los errores

$1{,}58051\cdot 10^{-1}$, $7{,}16059\cdot 10^{-2}$, $1{,}24885\cdot 10^{-2}$, $8{,}67861\cdot 10^{-4}$, $1{,}00207\cdot 10^{-5}$, $8{,}10907\cdot 10^{-9}$, $7{,}57244\cdot 10^{-14}$.

Aplicando la fórmula del paso 2 a las ternas sucesivas se obtienen las estimaciones

$$1{,}5269, \quad 1{,}6731, \quad 1{,}5958, \quad 1{,}6267, \quad 1{,}6147,$$

que oscilan alrededor de $\varphi = \dfrac{1+\sqrt{5}}{2} = 1{,}6180340$ y se van cerrando sobre él. Esa oscilación es típica de la secante y se debe a que su error depende de los *dos* pasos anteriores, no de uno; la sucesión de exponentes obedece a una recurrencia de Fibonacci, y de ahí sale el número áureo. El método no gana tanto como Newton por paso, pero cada paso cuesta una sola evaluación de f en vez de dos —f y f'—, de modo que por evaluación su rendimiento efectivo es $\varphi \approx 1{,}618$ frente a $\sqrt{2} \approx 1{,}414$: en la práctica la secante suele salir ganando.

Definición 14.11 (Orden de un método de discretización)**.** Un método que aproxima una cantidad exacta I mediante un valor $I(h)$ dependiente de un paso $h > 0$ se dice de *orden* q si su error se comporta como

$$|I(h) - I| \;\le\; C\,h^{\,q} \qquad \text{para } h \text{ pequeño},$$

lo que se abrevia escribiendo $E(h) = O(h^q)$. En ese caso, al dividir el paso por dos el error se divide aproximadamente por 2^q, y por tanto

$$q \;\approx\; \log_2 \frac{E(h)}{E(h/2)}.$$

Cuidado

[Son dos conceptos distintos con la misma palabra] El orden de la definición 14.8 se refiere a una sucesión de iterados y compara cada error con el anterior; el de la definición 14.11 se refiere a un parámetro de discretización y compara el error con el tamaño del paso. Newton es de orden 2 en el primer sentido; Simpson es de orden 4 en el segundo; y no significan lo mismo en absoluto. La confusión tiene consecuencias prácticas: en un método iterativo, un orden alto es gratis porque las iteraciones son baratas; en uno de discretización, un orden alto se paga con más evaluaciones por paso.

Ejemplo 14.12 (Verificar el orden de cuatro métodos con la tabla en la mano)**. (a) Cuadratura, sobre** $\int_1^2 \frac{\mathrm{d}x}{x} = \ln 2 = 0{,}69314718$**.** Los errores son

n	error trapecios	cociente	error Simpson	cociente
2	$1{,}51862 \cdot 10^{-2}$		$1{,}29726 \cdot 10^{-3}$	
4	$3{,}87663 \cdot 10^{-3}$	3,917	$1{,}06788 \cdot 10^{-4}$	12,15
8	$9{,}74670 \cdot 10^{-4}$	3,977	$7{,}35009 \cdot 10^{-6}$	14,53
16	$2{,}44022 \cdot 10^{-4}$	3,994	$4{,}72259 \cdot 10^{-7}$	15,56

Los cocientes tienden a $4 = 2^2$ y a $16 = 2^4$, confirmando los órdenes 2 y 4 que el teorema 14.66 demuestra. En logaritmo base dos, las estimaciones de q son 1,970, 1,992, 1,998 para el trapecio y 3,603, 3,861, 3,960 para Simpson.

(b) Ecuaciones diferenciales, sobre $y' = y$**,** $y(0) = 1$**, evaluando en** $t = 1$**,** donde la solución exacta es $\mathrm{e} = 2{,}718281828$. Con n pasos:

n	error de Euler	cociente	error de Runge–Kutta 4	cociente
10	$1{,}24539 \cdot 10^{-1}$		$2{,}08432 \cdot 10^{-6}$	
20	$6{,}49841 \cdot 10^{-2}$	1,916	$1{,}35803 \cdot 10^{-7}$	15,35
40	$3{,}32180 \cdot 10^{-2}$	1,956	$8{,}66619 \cdot 10^{-9}$	15,67
80	$1{,}67969 \cdot 10^{-2}$	1,978	$5{,}47306 \cdot 10^{-10}$	15,83

Los cocientes tienden a $2 = 2^1$ y a $16 = 2^4$: Euler es de primer orden y Runge–Kutta clásico de cuarto, exactamente como anuncia la regla de integración numérica de ecuaciones diferenciales. Nótese la magnitud: con solo diez pasos, Runge–Kutta ya acierta seis cifras y Euler ni una.

Observación 14.13 (El catálogo del capítulo, en una tabla).

método	orden	observación
bisección	1, con $C = \frac{1}{2}$	sobre la cota, no sobre el error real
punto fijo con $\lvert g'(r)\rvert = k$	1, con $C = k$	diverge si $k > 1$
secante	$\varphi = 1{,}618$	una evaluación por paso
Newton, raíz simple	2	constante $\lvert f''(p)/2f'(p)\rvert$
Newton, raíz doble	1, con $C = \frac{1}{2}$	se recupera el orden 2 multiplicando por m
trapecio compuesto	$O(h^2)$	discretización
Simpson compuesto	$O(h^4)$	discretización
Euler	$O(h)$	discretización
Runge–Kutta clásico	$O(h^4)$	discretización

Las dos últimas columnas hablan idiomas distintos, y esa es justamente la moraleja de la advertencia anterior.

14.4. Por qué convergen Jacobi y Gauss–Seidel: matriz de iteración, radio espectral y dominancia diagonal

El capítulo enuncia que ambos métodos convergen cuando la matriz es estrictamente diagonal dominante, y menciona el radio espectral sin definirlo. Las dos cosas son un mismo teorema visto desde dos alturas, y ambas tienen demostración corta. Usaremos las normas vectoriales y matriciales ya definidas en la definición 14.81; no hace falta nada más.

Definición 14.14 (Matriz de iteración)**.** Escríbase $A = D + L + U$, donde D es la diagonal de A, L su parte estrictamente triangular inferior y U la estrictamente superior, y supóngase $a_{ii} \neq 0$. Entonces:

$$\text{Jacobi:}\quad \vec{x}^{(k+1)} = B_J\vec{x}^{(k)} + \vec{c}_J, \qquad B_J = -D^{-1}(L+U), \quad \vec{c}_J = D^{-1}\vec{b};$$

$$\text{Gauss–Seidel:}\quad \vec{x}^{(k+1)} = B_G\vec{x}^{(k)} + \vec{c}_G, \qquad B_G = -(D+L)^{-1}U, \quad \vec{c}_G = (D+L)^{-1}\vec{b}.$$

La matriz B que multiplica al iterado anterior se llama *matriz de iteración* del método.

Definición 14.15 (Radio espectral)**.** El *radio espectral* de una matriz cuadrada B es

$$\rho(B) = \max\{|\lambda| : \lambda \text{ autovalor de } B\},$$

donde los autovalores se toman en $\mathbb{C}$ aunque B sea real.

Teorema 14.16 (Criterio general de convergencia)**.** *Sea $\vec{x}^*$ la solución de $\vec{x} = B\vec{x} + \vec{c}$. La iteración $\vec{x}^{(k+1)} = B\vec{x}^{(k)} + \vec{c}$ converge a $\vec{x}^*$ para* todo *vector inicial si y solo si $\rho(B) < 1$.*

Demostración. Restando $\vec{x}^* = B\vec{x}^* + \vec{c}$ de la recurrencia, el error $\vec{e}^{(k)} = \vec{x}^{(k)} - \vec{x}^*$ cumple $\vec{e}^{(k+1)} = B\vec{e}^{(k)}$ y por tanto $\vec{e}^{(k)} = B^k\vec{e}^{(0)}$. Todo se reduce, pues, a si $B^k\vec{e}^{(0)} \to \vec{0}$.

($\Rightarrow$) Si $\rho(B) \geq 1$, tómese un autovalor λ con $|\lambda| \geq 1$ y un autovector $\vec{v} \neq \vec{0}$. Eligiendo $\vec{x}^{(0)} = \vec{x}^* + \vec{v}$ se tiene $\vec{e}^{(0)} = \vec{v}$ y $\vec{e}^{(k)} = \lambda^k\vec{v}$, cuya norma es $|\lambda|^k\|\vec{v}\| \geq \|\vec{v}\| > 0$ y no tiende a cero. La iteración no converge para ese arranque.

($\Leftarrow$) Supóngase $\rho(B) < 1$ y, para fijar ideas, que B es diagonalizable: $B = PDP^{-1}$ con D diagonal de autovalores. Entonces $B^k = PD^kP^{-1}$, y como cada entrada de D^k es λ_i^k con $|\lambda_i| \leq \rho(B) < 1$, se tiene $D^k \to 0$ y por tanto $B^k \to 0$: el error tiende a cero sea cual sea $\vec{e}^{(0)}$. En el caso general se usa que para cada $\varepsilon > 0$ existe una norma subordinada con $\|B\| \leq \rho(B) + \varepsilon$; tomando $\varepsilon = \frac{1-\rho(B)}{2}$ queda $\|B\| < 1$ y entonces $\|\vec{e}^{(k)}\| \leq \|B\|^k\|\vec{e}^{(0)}\| \to 0$. □

Proposición 14.17. *Para toda matriz B y toda norma subordinada, $\rho(B) \leq \|B\|$.*

Demostración. Sea λ un autovalor con autovector $\vec{v} \neq \vec{0}$. De $B\vec{v} = \lambda\vec{v}$ se obtiene $\|B\vec{v}\| = |\lambda|\,\|\vec{v}\|$, y por definición de norma subordinada $\|B\vec{v}\| \leq \|B\|\,\|\vec{v}\|$. Dividiendo entre $\|\vec{v}\| > 0$ resulta $|\lambda| \leq \|B\|$ para todo autovalor, y en particular para el de módulo máximo. □

Teorema 14.18 (La dominancia diagonal basta)**.** *Si A es estrictamente diagonal dominante por filas, es decir $|a_{ii}| > \sum_{j\neq i}|a_{ij}|$ para todo i, entonces $\rho(B_J) < 1$ y $\rho(B_G) < 1$; por tanto Jacobi y Gauss–Seidel convergen desde cualquier vector inicial.*

Demostración. Jacobi. Las entradas de B_J son $(B_J)_{ij} = -a_{ij}/a_{ii}$ para $j \neq i$ y 0 en la diagonal. La suma de módulos de la fila i vale

$$\sum_{j\neq i}\frac{|a_{ij}|}{|a_{ii}|} = \frac{1}{|a_{ii}|}\sum_{j\neq i}|a_{ij}| < 1$$

por la hipótesis. Como $\|B_J\|_\infty$ es el mayor de esos números, $\|B_J\|_\infty < 1$, y por la proposición 14.17, $\rho(B_J) \leq \|B_J\|_\infty < 1$.

Gauss–Seidel. Sea λ un autovalor de B_G con autovector $\vec{v}$ normalizado de modo que $\|\vec{v}\|_\infty = 1$, y sea i un índice con $|v_i| = 1$. De $B_G\vec{v} = \lambda\vec{v}$ se sigue $-U\vec{v} = \lambda(D+L)\vec{v}$, cuya fila i-ésima es

$$-\sum_{j>i}a_{ij}v_j = \lambda\Big(a_{ii}v_i + \sum_{j<i}a_{ij}v_j\Big).$$

Despejando y acotando con $|v_j| \leq 1$:

$$|\lambda|\,|a_{ii}| \;\leq\; |\lambda|\sum_{j<i}|a_{ij}| + \sum_{j>i}|a_{ij}|.$$

Si fuese $|\lambda| \geq 1$, dividiendo entre $|\lambda|$ el segundo sumando solo puede crecer y quedaría $|a_{ii}| \leq \sum_{j<i}|a_{ij}| + \sum_{j>i}|a_{ij}| = \sum_{j\neq i}|a_{ij}|$, en contradicción con la dominancia estricta. Luego $|\lambda| < 1$ para todo autovalor. □

Ejemplo 14.19 (Los radios espectrales del sistema del capítulo)**.** Retomamos el sistema $4x - y = 2$, $-x + 4y - z = 4$, $-y + 4z = 10$, de solución $(1, 2, 3)$, cuya matriz es

$$A = \begin{pmatrix} 4 & -1 & 0 \\ -1 & 4 & -1 \\ 0 & -1 & 4 \end{pmatrix}, \qquad B_J = \begin{pmatrix} 0 & 0{,}25 & 0 \\ 0{,}25 & 0 & 0{,}25 \\ 0 & 0{,}25 & 0 \end{pmatrix}.$$

(a) La cota grosera. $\|B_J\|_\infty = \max\{0{,}25;\ 0{,}5;\ 0{,}25\} = 0{,}5 < 1$: por el teorema 14.18 Jacobi converge, y la cota predice que el error se multiplica a lo sumo por 0,5 en cada paso.

(b) El valor exacto. Los autovalores de B_J son $-\frac{\sqrt{2}}{4}$, 0 y $\frac{\sqrt{2}}{4}$, de modo que

$$\rho(B_J) = \frac{\sqrt{2}}{4} = 0{,}3535534.$$

La tabla de errores máximos del ejemplo de Jacobi era 1; 0,25; 0,125; 0,03125; 0,015625; 0,00390625, con cocientes que alternan 0,25 y 0,5. Su media geométrica es $\sqrt{0{,}125} = 0{,}3535534$: exactamente el radio espectral ✓. El radio espectral predice el comportamiento *a largo plazo*, no paso a paso, y esta alternancia es la razón de que la frase «el error se divide entre dos» sea solo una aproximación grosera.

(c) Gauss–Seidel. Los autovalores de B_G son 0, 0 y 0,125, luego

$$\rho(B_G) = \frac{1}{8} = 0{,}125 = \rho(B_J)^2.$$

Los errores observados eran 0,875; 0,21875; 0,0273438; 0,0034180; 0,0004272; 0,0000534, con cociente constante 0,125 ✓. Aquí sí se cumple limpiamente «el error se divide entre ocho», y la relación $\rho(B_G) = \rho(B_J)^2$ —cierta para las matrices tridiagonales con esta estructura— explica en una línea por qué Gauss–Seidel necesita la mitad de iteraciones: cada paso suyo equivale a dos de Jacobi.

Ejemplo 14.20 (Dominancia suficiente, no necesaria)**. (a) Converge sin ser dominante.** Sea $A = \begin{pmatrix} 1 & 2 \\ 2 & 5 \end{pmatrix}$ con $\vec{b} = (5, 12)$, de solución exacta $(1, 2)$. La primera fila incumple la dominancia ($1 < 2$), así que el teorema 14.18 no dice nada; de hecho $\|B_J\|_\infty = 2 > 1$. Pero

$$B_J = \begin{pmatrix} 0 & -2 \\ -0{,}4 & 0 \end{pmatrix}, \qquad \rho(B_J) = \sqrt{0{,}8} = 0{,}8944272 < 1,$$

y el teorema 14.16 garantiza convergencia. Arrancando en $(0, 0)$, los errores máximos son 4; 1,6; 3,2; 1,28; 2,56; 1,024; 2,048; 0,8192: oscilan, pero cada dos pasos se multiplican por $\rho^2 = 0{,}8$, y acaban en cero. La convergencia es real y lentísima.

(b) No dominante y divergente. Con $A = \begin{pmatrix} 1 & 2 \\ 3 & 1 \end{pmatrix}$ se obtiene $B_J = \begin{pmatrix} 0 & -2 \\ -3 & 0 \end{pmatrix}$ y $\rho(B_J) = \sqrt{6} = 2{,}4494897 > 1$: la iteración se dispara desde cualquier arranque que no sea la solución exacta. Basta intercambiar las dos filas para obtener $\begin{pmatrix} 3 & 1 \\ 1 & 2 \end{pmatrix}$, que sí es estrictamente dominante y converge sin problemas. Ese intercambio es el paso 1 de la regla de resolución iterativa del capítulo, y aquí se ve por qué está ahí.

Regla: decidir de antemano si un método iterativo va a converger, y cómo de rápido

1. **Mira primero la dominancia diagonal**, que es gratis: una suma de módulos por fila. Si la hay, se acabó la discusión, ambos métodos convergen.
2. **Si no la hay, reordena las filas** e inténtalo otra vez; muchas matrices que parecen malas solo están mal escritas.
3. **Si sigue sin haberla, calcula o estima $\rho(B)$.** Es el criterio exacto: converge si y solo si $\rho(B) < 1$, con dominancia o sin ella.
4. **Usa $\|B\|_\infty$ como cota rápida.** Si $\|B\|_\infty < 1$, ya está: por la proposición 14.17 también $\rho(B) < 1$. Si $\|B\|_\infty > 1$, no has demostrado nada, ni a favor ni en contra.
5. **Traduce ρ a velocidad**: el error se multiplica por ρ por iteración a la larga, de modo que ganar una cifra decimal cuesta unas $-1/\log_{10}\rho$ iteraciones. Con $\rho = 0{,}3536$ son 2,2 iteraciones por cifra; con $\rho = 0{,}894$, veinte.
6. **No confundas lento con divergente.** Un ρ muy próximo a 1 da una convergencia tan lenta que en la pantalla parece un estancamiento; el remedio no es cambiar de método sino precondicionar el sistema.
7. **Verifica** calculando el residuo $\|A\vec{x}^{(k)} - \vec{b}\|$ además del cambio entre iteraciones, por las razones que explica el apartado del número de condición.

Cuidado

El radio espectral *no* es una norma y no cumple la desigualdad triangular: hay matrices con $\rho(B) = 0$ que no son nulas, como $\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$, cuyos dos autovalores son cero. Por eso la desigualdad de la proposición 14.17 va en un solo sentido y puede ser muy holgada: en el ejemplo del sistema del capítulo, $\|B_J\|_\infty = 0{,}5$ frente a $\rho(B_J) = 0{,}354$. Segundo aviso: $\rho(B) < 1$ describe el comportamiento *asintótico*. En los primeros pasos el error puede crecer, y de hecho crece en el apartado (a) del último ejemplo; que las dos o tres primeras iteraciones empeoren no demuestra que el método diverja.

14.5. Bisección: el método que nunca falla

Teorema 14.21 (Bolzano)**.** *Si f es continua en $[a, b]$ y $f(a)$ y $f(b)$ tienen signos opuestos, existe al menos un $c \in (a, b)$ con $f(c) = 0$.*

Este teorema, ya visto en los límites elementales, no solo garantiza la raíz: sugiere cómo encontrarla.

Regla: Método de bisección

1. Localiza $a < b$ con $f(a)$ y $f(b)$ de signos opuestos.
2. Calcula el punto medio $m = \frac{a+b}{2}$ y evalúa $f(m)$.
3. Si $f(a)$ y $f(m)$ tienen signos opuestos, la raíz está en $[a, m]$: haz $b = m$. En caso contrario está en $[m, b]$: haz $a = m$.
4. Repite hasta que la longitud del intervalo sea menor que la precisión deseada.
5. Da como respuesta el punto medio del último intervalo.

Ejemplo 14.22 (La raíz de $x^3 - x - 1$)**.** Aquí $f(1) = -1 < 0$ y $f(2) = 5 > 0$, luego hay raíz en $[1, 2]$. Las primeras iteraciones son:

Paso	a	b	m	$f(m)$
1	1,0000	2,0000	1,5000	+0,8750
2	1,0000	1,5000	1,2500	−0,2969
3	1,2500	1,5000	1,3750	+0,2246
4	1,2500	1,3750	1,3125	−0,0515
5	1,3125	1,3750	1,3438	+0,0826
6	1,3125	1,3438	1,3281	+0,0146
7	1,3125	1,3281	1,3203	−0,0187
8	1,3203	1,3281	1,3242	−0,0021

La raíz verdadera es 1,3247179…, de modo que tras ocho pasos ya se tienen tres cifras correctas.

Proposición 14.23 (Cota de error)**.** *Tras n pasos, el error cometido al dar el punto medio es a lo sumo*

$$\frac{b-a}{2^{n+1}}.$$

En particular el método siempre converge, y se sabe de antemano cuántos pasos harán falta.

Demostración. Cada paso reduce la longitud del intervalo exactamente a la mitad, luego tras n pasos mide $(b-a)/2^n$. La raíz y el punto medio están ambos dentro de él, y el punto medio dista a lo sumo la mitad de su longitud de cualquier punto suyo. □

Ejemplo 14.24. Partiendo de $[1, 2]$, para garantizar un error menor que 0,001 basta con $2^{-n} < 10^{-3}$, es decir $n \geq 10$, pues $2^{-10} = 0{,}0009765\ldots$. Diez pasos, y ni uno más, sea cual sea la función.

Idea clave

La bisección es el método más lento y el más fiable. Gana *una cifra binaria por paso*, o lo que es igual, algo más de tres pasos por cada cifra decimal; nunca diverge, nunca depende de una buena suposición inicial y no necesita derivadas.
Es la estrategia de "adivina el número que estoy pensando entre 1 y 1000": preguntando siempre por la mitad se acierta en diez preguntas, y esa es exactamente la razón de que $2^{10} \approx 1000$ aparezca aquí y en los sistemas de numeración.

Cuidado

La bisección necesita un cambio de signo. Si la raíz es doble —como el 0 de $f(x) = x^2$—, la función toca el eje sin cruzarlo, no hay dos signos opuestos y el método no arranca. Tampoco distingue una raíz de una asíntota vertical: $f(x) = 1/x$ cambia de signo entre -1 y 1 y ahí no hay ninguna raíz, sino una discontinuidad. La hipótesis de continuidad no es decorativa.

14.6. Newton: seguir la tangente

Regla: Método de Newton–Raphson

Partiendo de una estimación x_0, se itera

$$x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)}.$$

Se para cuando dos iteraciones consecutivas coinciden en las cifras deseadas.

De dónde sale la fórmula. Sustituya la curva por su recta tangente en x_n, que es $y = f(x_n) + f'(x_n)(x - x_n)$ —la aproximación lineal del estudio de funciones—. Esa recta corta al eje X cuando $f(x_n) + f'(x_n)(x - x_n) = 0$, es decir, en $x = x_n - f(x_n)/f'(x_n)$. Ese corte es el siguiente candidato. □

Idea clave

Newton dice: *si no sé resolver la ecuación, resuelvo la de la tangente, que es de primer grado, y confío en que el corte de la tangente esté más cerca de la raíz que mi punto de partida.* Repetido, funciona espectacularmente bien.
La bisección camina; Newton salta. Y salta tan bien que, cerca de la raíz, *duplica el número de cifras correctas en cada paso.* A eso se le llama convergencia cuadrática, y es la razón de que sea el método que está dentro de casi todas las calculadoras.

Ejemplo 14.25 (La misma raíz, con Newton)**.** Para $f(x) = x^3 - x - 1$ se tiene $f'(x) = 3x^2 - 1$, luego

$$x_{n+1} = x_n - \frac{x_n^3 - x_n - 1}{3x_n^2 - 1}.$$

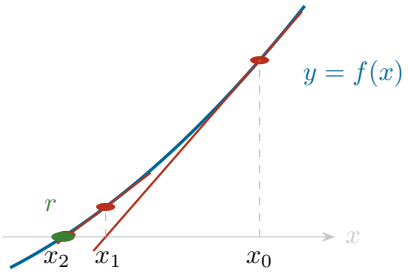


Figura 14.1: El método de Newton: desde x_n se traza la tangente y se toma su corte con el eje como x_{n+1}. Partiendo de $x_0 = 3$ se obtiene $x_1 = 1{,}9815$, $x_2 = 1{,}7197$ y $x_3 = 1{,}6998$, frente a la raíz $r = 1{,}6997$: cerca de una raíz simple el número de cifras correctas se duplica en cada paso.

Partiendo de $x_0 = 1{,}5$:

n	x_n
0	1,5
1	1,347826086957
2	1,325200398951
3	1,324718173999
4	1,324717957245

Cuatro pasos dan doce cifras correctas; la bisección necesitaría unas cuarenta iteraciones para lo mismo. Obsérvese la duplicación: el paso 2 acierta tres cifras, el 3 acierta seis, el 4 acierta doce.

Ejemplo 14.26 (La raíz cuadrada, o el método de los babilonios)**.** Aplicar Newton a $f(x) = x^2 - A$ da $x_{n+1} = x_n - \frac{x_n^2 - A}{2x_n}$, que simplificado es

$$x_{n+1} = \frac{1}{2}\left(x_n + \frac{A}{x_n}\right).$$

Es decir: *promedia tu estimación con A dividido por tu estimación.* Si x_n se queda corto, A/x_n se pasa, y la media queda entre medias.

Con $A = 2$ y $x_0 = 1$: 1,5; $1{,}41\overline{6}$; 1,4142156862745; 1,4142135623746; 1,4142135623731. El valor real es 1,4142135623731. Cinco pasos, quince cifras. Este algoritmo aparece en tablillas babilónicas de hace casi cuatro mil años, dos milenios y medio antes que la derivada que lo justifica.

Cuidado

Newton es rápido pero no es infalible, y falla de tres maneras:

- Si $f'(x_n) = 0$ la fórmula divide entre cero: la tangente es horizontal y no corta al eje.
- Si el punto de partida es malo, puede alejarse en vez de acercarse.

- Puede *ciclar*. Para $f(x) = x^3 - 2x + 2$ con $x_0 = 0$ se obtiene $x_1 = 1$, luego $x_2 = 0$, luego $x_3 = 1\ldots$ para siempre, sin acercarse jamás a la raíz, que está cerca de $-1{,}77$.

Por eso en la práctica se combinan los dos métodos: unos pasos de bisección para acorralar la raíz, y luego Newton para afinar.

14.7. La secante: Newton sin derivada

Si la derivada es difícil o directamente desconocida —porque f viene de un experimento y no de una fórmula—, se sustituye la tangente por la recta que pasa por los dos últimos puntos:

$$x_{n+1} = x_n - f(x_n)\,\frac{x_n - x_{n-1}}{f(x_n) - f(x_{n-1})}.$$

Ejemplo 14.27. Con $f(x) = x^3 - x - 1$ y los puntos de partida $x_0 = 1$, $x_1 = 2$: 1,166667; 1,253112; 1,337206; 1,323850; 1,324708; 1,324718. Seis pasos para lo que Newton hace en cuatro, pero sin calcular ni una derivada.

Observación 14.28. La secante converge con orden $\varphi = \frac{1+\sqrt{5}}{2} \approx 1{,}618$, el número áureo de las sucesiones y series elementales: cada paso multiplica por 1,618 el número de cifras correctas, frente al 2 de Newton y al 1 de la bisección. Que el número áureo aparezca aquí no es una casualidad pintoresca: sale de la misma recurrencia de Fibonacci, porque el error de un paso depende de los *dos* anteriores.

14.8. Iteración de punto fijo

Muchos problemas se dejan escribir en la forma $x = g(x)$. Entonces se puede probar lo más ingenuo del mundo: elegir un x_0, calcular $x_1 = g(x_0)$, luego $x_2 = g(x_1)$, y ver qué pasa.

Ejemplo 14.29 ($\cos x = x$)**.** Tome la calculadora en radianes, escriba 1 y pulse la tecla del coseno muchas veces:

$$0{,}5403,\ \ 0{,}8576,\ \ 0{,}6543,\ \ 0{,}7935,\ \ 0{,}7014,\ \ \ldots,\ \ 0{,}74424\ (n=10),\ \ \ldots,\ \ 0{,}739071\ (n=25).$$

Los valores oscilan alrededor de $0{,}7390851332\ldots$ y se van estrechando. Es un experimento que cualquiera puede hacer en treinta segundos y que sale siempre, se empiece donde se empiece.

Teorema 14.30 (Criterio de convergencia)**.** *Si g es derivable cerca de un punto fijo p (es decir, $g(p) = p$) y $|g'(p)| < 1$, entonces la iteración $x_{n+1} = g(x_n)$ converge a p para todo punto de partida suficientemente cercano. Si $|g'(p)| > 1$, diverge.*

Idea de la demostración. Por el teorema del valor medio del estudio de funciones, $|x_{n+1} - p| = |g(x_n) - g(p)| = |g'(\xi)|\,|x_n - p|$ para algún ξ intermedio. Si $|g'| \le k < 1$ en un entorno, el error se multiplica por k en cada paso, luego tiende a cero geométricamente. Ese argumento, llevado a espacios métricos completos, es el *teorema del punto fijo de Banach* del capítulo 32,

y es la herramienta con la que allí se demuestra la existencia de soluciones de ecuaciones diferenciales. □

Ejemplo 14.31. En el caso anterior $g(x) = \cos x$, luego $g'(x) = -\sin x$ y en el punto fijo $|g'| = \sin(0{,}73909) \approx 0{,}674 < 1$. Converge, y además se ve por qué converge *despacio*: cada paso solo reduce el error a dos tercios, lo que exige unos cinco pasos por cifra decimal.

En cambio, si se intentase resolver la misma ecuación escrita como $x = \arccos x$, se tendría $|g'| \approx 1/0{,}674 > 1$ y la iteración se escaparía. La misma ecuación, dos maneras de despejar, dos destinos opuestos.

Pausa para pensar

Newton es un caso particular de esto, con $g(x) = x - f(x)/f'(x)$. Derivando y evaluando en la raíz sale $g'(p) = 0$: no es que el error se multiplique por un número menor que uno, es que el término lineal del error desaparece por completo y solo queda el cuadrático. Ahí está, en una línea, la explicación de por qué Newton duplica cifras y los demás métodos no.

Regla: Iteración de punto fijo: elegir bien la g

Toda ecuación $f(x) = 0$ se puede reescribir como $x = g(x)$, y entonces iterar $x_{n+1} = g(x_n)$. Lo delicado no es iterar, sino elegir la g: la misma ecuación admite infinitas reescrituras y unas convergen mientras otras estallan.

1. **Despejar x de alguna manera** para pasar de $f(x) = 0$ a $x = g(x)$. Toda solución de una ecuación es punto fijo de la otra, y al revés.
2. **Localizar la raíz aproximadamente** (por bisección, por una gráfica o por cambio de signo) para saber en qué entorno se va a trabajar.
3. **Calcular $g'(x)$ y evaluarla cerca de la raíz.** Este es el paso que decide todo.
4. **Criterio de convergencia:** si $|g'(r)| < 1$ en un entorno de la raíz r, la iteración converge para todo arranque suficientemente próximo; si $|g'(r)| > 1$, diverge por muy bueno que sea el arranque.
5. **Velocidad:** el error se multiplica aproximadamente por $|g'(r)|$ en cada paso, es decir convergencia *lineal*. Cuanto menor sea $|g'(r)|$, más rápido.
6. **Si $g'(r) = 0$,** la convergencia pasa a ser cuadrática: el número de decimales correctos se duplica en cada paso. El método de Newton es exactamente el caso $g(x) = x - f(x)/f'(x)$, construido a propósito para que $g'(r) = 0$.
7. **Si la g elegida diverge, no se abandona: se cambia de g.** Despejar por otro lado, o incluso invertir la relación, suele bastar.

Ejemplo 14.32 (una ecuación, tres despejes, tres destinos)**.** Se busca la raíz de $x^3 + x - 1 = 0$, que existe y es única porque la función es estrictamente creciente y cambia de signo entre 0 y

1. Su valor, con quince cifras, es

$$r = 0{,}682327803828019.$$

(a) *Primer despeje:* $x = 1 - x^3$. Aquí $g_1(x) = 1 - x^3$ y $g_1'(x) = -3x^2$, de modo que

$$|g_1'(r)| = 3 \cdot (0{,}6823)^2 = 1{,}3967 > 1.$$

El criterio predice divergencia, y arrancando en $x_0 = 0{,}5$ eso es exactamente lo que pasa:

$$0{,}5 \to 0{,}875 \to 0{,}3301 \to 0{,}9640 \to 0{,}1041 \to 0{,}9989 \to 0{,}0034 \to 1 \to 0 \to 1 \to 0 \to \cdots$$

La sucesión no solo no converge: acaba oscilando entre 0 y 1 para siempre. Y sin embargo cualquiera de esos números es un despeje perfectamente legítimo de la ecuación original.

(b) *Segundo despeje:* $x = \sqrt[3]{1-x}$. Ahora $g_2'(x) = -\dfrac{1}{3(1-x)^{2/3}}$ y

$$|g_2'(r)| = 0{,}7160 < 1.$$

Converge, y arrancando en 0,5:

$$0{,}5 \to 0{,}7937 \to 0{,}5909 \to 0{,}7424 \to 0{,}6363 \to 0{,}7138 \to 0{,}6590 \to 0{,}6986 \to 0{,}6704 \to \cdots$$

Se ve la oscilación alrededor de la raíz —el signo negativo de g'— y la reducción del error en un factor 0,72 por paso. Para ganar un decimal hacen falta unos siete pasos.

(c) *Tercer despeje: el de Newton.* Con $g_3(x) = x - \dfrac{x^3 + x - 1}{3x^2 + 1} = \dfrac{2x^3 + 1}{3x^2 + 1}$, arrancando también en 0,5:

$$0{,}5 \to 0{,}714285714286 \to 0{,}683179723502 \to 0{,}682328423305 \to 0{,}682327803828.$$

Cuatro pasos para doce cifras exactas, frente a los cientos que necesitaría el despeje (b). El motivo es $g_3'(r) = 0$: la convergencia es cuadrática, y los decimales correctos van $1, 3, 6, 12$, duplicándose.

(d) *El caso más famoso.* La ecuación $x = \cos x$ ya viene despejada. Su punto fijo es $r = 0{,}739085133215161$, y $g'(r) = -\sin(r) = -0{,}6736$, con módulo menor que 1: converge. Arrancando en $x_0 = 1$ y pulsando la tecla del coseno repetidamente en una calculadora en radianes,

$$1 \to 0{,}5403 \to 0{,}8576 \to 0{,}6543 \to 0{,}7935 \to 0{,}7014 \to 0{,}7640 \to 0{,}7221 \to 0{,}7504 \to \cdots$$

Es el experimento de sobremesa clásico: la calculadora acaba clavada en 0,7390851 salga de donde salga.

(e) *Verificación.* Sustituyendo el resultado de (c) en la ecuación original:

$$0{,}682327803828^3 + 0{,}682327803828 - 1 = -4{,}6 \cdot 10^{-14} \approx 0. \checkmark$$

Toda iteración debe terminar con este paso; el criterio de parada $|x_{n+1} - x_n| < \varepsilon$ dice que la sucesión se ha estabilizado, no que se haya estabilizado en la raíz correcta.

Cuidado

El caso (a) es la lección entera del apartado: haber despejado bien el álgebra no garantiza nada. Si al iterar los valores se disparan, oscilan o saltan al azar, no es que la ecuación no tenga solución ni que el ordenador falle: es que $|g'| > 1$. La respuesta es despejar por otro lado, y hay una receta general para ello, porque si $|g'(r)| > 1$ entonces la función inversa cumple $|(g^{-1})'(r)| < 1$.

Pausa para pensar

El teorema que hay detrás es el *teorema del punto fijo de Banach*, y no es exclusivo de los números: vale en cualquier espacio métrico completo, con $|g(x) - g(y)| \leq k|x - y|$ y $k < 1$ en lugar de la derivada. Con él se demuestran la existencia y unicidad de soluciones de ecuaciones diferenciales (capítulo 12), el teorema de la función implícita (capítulo 9) y la existencia de fractales autosemejantes. Esta humilde iteración de calculadora es la versión escolar de uno de los teoremas más productivos del análisis (capítulo 32).

14.9. Interpolación

Interpolar es estimar el valor de una función en un punto intermedio a partir de valores conocidos. La versión más simple sustituye la curva por la recta que une los dos datos vecinos:

$$f(x) \approx f(x_0) + \frac{f(x_1) - f(x_0)}{x_1 - x_0}\,(x - x_0).$$

Ejemplo 14.33. Sabiendo que $\sqrt{4} = 2$ y $\sqrt{9} = 3$, la interpolación lineal estima

$$\sqrt{7} \approx 2 + \frac{7-4}{9-4}(3-2) = 2 + 0{,}6 = 2{,}6,$$

frente al valor real 2,6458. El error es de un 1,7 %, razonable para una tabla y muy malo para un cálculo científico. Con datos más juntos el error cae deprisa: interpolando entre $\sqrt{6{,}25} = 2{,}5$ y $\sqrt{9} = 3$ ya se acierta bastante más.

Cuidado

Interpolar es estimar *entre* los datos; **extrapolar** es estimar fuera de ellos, y es mucho más peligroso. Una recta ajustada a las temperaturas de una semana no dice nada sobre el mes que viene, y una curva ajustada a diez datos puede oscilar salvajemente justo después del último. La regresión de la estadística descriptiva lleva la misma advertencia por la misma razón.

Regla: Construir el polinomio interpolador

Dados $n + 1$ puntos con abscisas distintas, existe un único polinomio de grado $\leq n$ que pasa por todos. Hay dos maneras de escribirlo, y conviene conocer las dos.

1. **Comprobar que las abscisas son distintas.** Si se repite una x_i con dos y distintas, no hay ninguna función que pase por ambos puntos, y menos un polinomio.
2. **Forma de Lagrange.** Para cada i se construye
$$L_i(x) = \prod_{j \neq i} \frac{x - x_j}{x_i - x_j},$$
que vale 1 en x_i y 0 en los demás nodos, y entonces $P(x) = \sum_i y_i L_i(x)$. Es transparente pero incómoda: añadir un punto obliga a recalcularlo todo.
3. **Forma de Newton con diferencias divididas.** Se define $f[x_i] = y_i$ y, recursivamente,
$$f[x_i, \dots, x_{i+k}] = \frac{f[x_{i+1}, \dots, x_{i+k}] - f[x_i, \dots, x_{i+k-1}]}{x_{i+k} - x_i}.$$
Entonces $P(x) = f[x_0] + f[x_0, x_1](x - x_0) + f[x_0, x_1, x_2](x - x_0)(x - x_1) + \cdots$
4. **Ventaja de Newton:** añadir un punto nuevo solo añade un sumando; lo anterior no se toca.
5. **Estimar el error** en un punto x interior: $|f(x) - P(x)| \leq \dfrac{M_{n+1}}{(n+1)!} \prod_i |x - x_i|$, con M_{n+1} una cota de la derivada $(n+1)$-ésima.
6. **No interpolar fuera del rango de los nodos.** Extrapolar con un polinomio es de las cosas más peligrosas que se pueden hacer con un ordenador.
7. **Verificar** evaluando P en cada nodo: debe reproducir exactamente los datos.

Ejemplo 14.34 (cuatro interpolaciones y el desastre de Runge)**.** (a) *Tres puntos, una parábola.* Se quiere el polinomio que pasa por $(0,1)$, $(1,3)$ y $(3,2)$. Con diferencias divididas:
$$f[x_0] = 1, \qquad f[x_0, x_1] = \frac{3-1}{1-0} = 2, \qquad f[x_1, x_2] = \frac{2-3}{3-1} = -\tfrac{1}{2},$$
$$f[x_0, x_1, x_2] = \frac{-\frac{1}{2} - 2}{3 - 0} = -\tfrac{5}{6}.$$
Por tanto
$$P(x) = 1 + 2x - \tfrac{5}{6}\, x(x-1) = -\tfrac{5}{6}x^2 + \tfrac{17}{6}x + 1.$$
Verificación: $P(0) = 1$, $P(1) = -\frac{5}{6} + \frac{17}{6} + 1 = 3$, $P(3) = -\frac{15}{2} + \frac{17}{2} + 1 = 2$. ✓

(b) *Interpolar es estimar entre datos.* El valor $P(2) = -\frac{10}{3} + \frac{17}{3} + 1 = \frac{10}{3} = 3{,}3333$ es la estimación en un punto donde no había dato. Nótese que queda *por encima* de los dos valores vecinos 3 y 2: la parábola se abomba, y eso puede ser realista o absurdo según de qué sean los datos.

(c) *Aproximar una raíz cuadrada.* Con los cuatro puntos $(1,1)$, $(4,2)$, $(9,3)$, $(16,4)$ de $\sqrt{x}$, las diferencias divididas son
$$1, \qquad \tfrac{1}{3}, \qquad -\tfrac{1}{60}, \qquad \tfrac{1}{1260},$$

y el polinomio de Newton evaluado en 7 da

$$P(7) = 1 + \tfrac{1}{3}\cdot 6 - \tfrac{1}{60}\cdot 6\cdot 3 + \tfrac{1}{1260}\cdot 6\cdot 3\cdot(-2) = 2{,}671429.$$

El valor exacto es $\sqrt{7} = 2{,}645751$: un error de 0,0257, algo menos del 1 %, con solo cuatro datos y sin ninguna calculadora científica.

(d) *El fenómeno de Runge.* Parece razonable pensar que más nodos dan mejor aproximación. Con la función $f(x) = \dfrac{1}{1+25x^2}$ en $[-1,1]$ e interpolando en nodos *equiespaciados*, el error máximo es:

grado n	error máximo
5	0,4327
10	1,9156
15	2,1076

El error *crece* con el grado. Cerca de los extremos el polinomio oscila violentamente y se despega de la función, aunque en los nodos coincida exactamente con ella. Es el ejemplo de Runge de 1901, y desmiente de un golpe la intuición de que interpolar más es interpolar mejor.

(e) *La cura: nodos de Chebyshev.* Si en lugar de repartir los nodos uniformemente se toman

$$x_k = \cos\!\left(\frac{(2k+1)\pi}{2n+2}\right), \qquad k = 0,\dots,n,$$

que se concentran cerca de los extremos, el error de la misma función pasa a

$$n = 10:\ 0{,}1092, \qquad n = 15:\ 0{,}0831,$$

y ahora sí decrece. La diferencia con la tabla anterior es de más de un orden de magnitud, y lo único que ha cambiado es *dónde* se han tomado las medidas.

Cuidado

Interpolar y ajustar por mínimos cuadrados no son lo mismo. La interpolación obliga al polinomio a pasar *exactamente* por todos los puntos, lo cual es lo correcto cuando los datos son exactos (una tabla matemática) y lo peor posible cuando tienen ruido experimental: el polinomio persigue el error de medida y oscila. Con datos ruidosos hay que ajustar, no interpolar; el la estadística descriptiva explica la regresión.

Pausa para pensar

Existe una manera puramente algebraica de ver la unicidad: si dos polinomios P y Q de grado $\le n$ coinciden en $n+1$ puntos, entonces $P-Q$ es un polinomio de grado $\le n$ con $n+1$ raíces, luego es idénticamente nulo. Es el mismo argumento con el que el capítulo 24 demuestra que un polinomio no nulo de grado n sobre un cuerpo tiene a lo sumo n raíces, y aquí lleva a un resultado de análisis numérico. Un teorema de álgebra

abstracta puesto a trabajar en una tabla de datos.

14.10. Ajuste por mínimos cuadrados: la recta que mejor pasa entre los puntos

La interpolación obliga a la curva a pasar exactamente por todos los datos, y eso es justo lo que no se quiere cuando los datos vienen de medidas con error: forzar el paso por cada punto haría que el modelo copiara también el ruido. La alternativa es pedir que la curva pase *cerca* de todos ellos, y "cerca" se cuantifica de una manera concreta que además tiene solución cerrada.

Definición 14.35. Dados n pares (x_i, y_i), la *recta de regresión* por *mínimos cuadrados* es la recta $y = a + bx$ que minimiza $\sum_i (y_i - a - bx_i)^2$. Derivando e igualando a cero se obtienen las *ecuaciones normales*, cuya solución es

$$b = \frac{n\sum x_i y_i - (\sum x_i)(\sum y_i)}{n\sum x_i^2 - (\sum x_i)^2}, \qquad a = \bar{y} - b\,\bar{x}.$$

Los $r_i = y_i - (a + bx_i)$ son los *residuos*, y el *coeficiente de determinación* $R^2 = 1 - \dfrac{\sum r_i^2}{\sum (y_i - \bar{y})^2}$ mide qué fracción de la variación de y explica el modelo.

Idea clave

Se minimizan los cuadrados de los residuos y no sus valores absolutos, y hay dos razones. La primera es práctica: la suma de cuadrados es una función derivable de a y b, así que igualar a cero las dos derivadas parciales da un sistema lineal de dos ecuaciones que se resuelve a mano. Con valores absolutos la función tiene picos y no hay fórmula cerrada. La segunda es geométrica: minimizar la suma de cuadrados equivale a proyectar ortogonalmente el vector de datos sobre el subespacio generado por las columnas del modelo, y esa proyección es única y fácil de calcular. Los mínimos cuadrados no son una elección arbitraria; son la única que convierte el problema en álgebra lineal.

Regla: Ajustar una recta o una curva por mínimos cuadrados

El procedimiento, con las comprobaciones que casi nadie hace y que evitan la mayoría de los disparates:

1. **Dibuja los puntos antes de ajustar nada.** Si la nube no tiene forma de recta, la recta de regresión existirá igualmente y no significará nada.
2. **Calcula las cuatro sumas** $\sum x_i$, $\sum y_i$, $\sum x_i y_i$ y $\sum x_i^2$; son los únicos datos que necesitan las fórmulas.
3. **Aplica la fórmula de la pendiente** y luego la de la ordenada, en ese orden, porque la segunda usa la primera.

4. **Comprueba que la recta pasa por** $(\bar{x}, \bar{y})$; siempre lo hace, y es la verificación más rápida de que no hay error aritmético.
5. **Calcula los residuos y comprueba que suman cero**, otra propiedad automática del ajuste que sirve de control.
6. **Evalúa** R^2 e interpreta: cerca de 1 el modelo explica casi toda la variación; cerca de 0, prácticamente nada.
7. **Si la relación no es lineal, linealiza primero**: $y = Ae^{kx}$ se convierte en $\ln y = \ln A + kx$, y $y = Ax^p$ en $\ln y = \ln A + p \ln x$; ajusta la recta sobre las variables transformadas y deshaz el cambio al final.

Ejemplo 14.36 (Una recta ajustada paso a paso)**. (a) Los datos y las sumas.** Tomemos (1; 2,1), (2; 3,9), (3; 6,2), (4; 7,8), (5; 10,1), (6; 12,2). Con $n = 6$: $\sum x_i = 21$, $\sum y_i = 42{,}3$, $\sum x_i y_i = 183{,}4$ y $\sum x_i^2 = 91$.

(b) Los coeficientes. Aplicando las fórmulas,

$$b = \frac{6 \cdot 183{,}4 - 21 \cdot 42{,}3}{6 \cdot 91 - 21^2} = \frac{1100{,}4 - 888{,}3}{546 - 441} = \frac{212{,}1}{105} = 2{,}02,$$

y $a = \frac{42{,}3}{6} - 2{,}02 \cdot \frac{21}{6} = 7{,}05 - 7{,}07 = -0{,}02$. La recta es $y = 2{,}02x - 0{,}02$, y pasa por (3,5; 7,05), que es el centro de gravedad de la nube, como debe.

(c) Residuos y bondad del ajuste. Las predicciones son 2,00, 4,02, 6,04, 8,06, 10,08 y 12,10, con residuos 0,10, −0,12, 0,16, −0,26, 0,02 y 0,10, que suman cero. La suma de cuadrados de los residuos es 0,128 frente a una variación total de 71,535, luego $R^2 = 1 - \frac{0{,}128}{71{,}535} = 0{,}998211$: el modelo explica un 99,82 % de la variación. El coeficiente de correlación es $r = 0{,}999105$, y en efecto $r^2 = R^2$.

(d) Usar el modelo, con prudencia. Para $x = 7$ la recta predice $2{,}02 \cdot 7 - 0{,}02 = 14{,}12$. Es una *extrapolación*, y solo es defendible si hay razones para creer que la relación lineal se mantiene fuera del rango medido; dentro del rango, entre $x = 1$ y $x = 6$, la interpolación es mucho más segura. La estadística no sabe nada del fenómeno: predecirá con la misma serenidad valores absurdos si se le piden.

Ejemplo 14.37 (Ajustes no lineales por linealización)**. (a) Un crecimiento exponencial.** Los datos (0; 5,0), (1; 8,2), (2; 13,4), (3; 22,0), (4; 36,1) crecen multiplicativamente. Tomando logaritmos y ajustando una recta a $(x_i, \ln y_i)$ sale $\ln y = 1{,}609109 + 0{,}494062x$, de donde $A = e^{1{,}609109} = 4{,}998355$ y $k = 0{,}494062$. El modelo es $y = 4{,}998355\, e^{0{,}494062x}$.

(b) Comprobación del ajuste exponencial. Las predicciones son 4,998355, 8,192103, 13,426528, 22,005540 y 36,066197, prácticamente indistinguibles de los datos; el R^2 sobre los logaritmos es 0,999998. Obsérvese que $e^{0{,}494062} = 1{,}639$: cada unidad de x multiplica y por algo más de 1,6, que es la lectura útil del parámetro.

(c) Una ley potencial. Con (1; 3), (2; 8,5), (4; 24), (8; 68), (16; 192), el ajuste sobre $(\ln x, \ln y)$ da pendiente $p = 1{,}500000$ y $A = 3{,}002080$. El modelo es $y \approx 3x^{1{,}5}$, es decir $y \propto x^{3/2}$: exactamente la forma de la tercera ley de Kepler. Este es el gran uso del papel

logarítmico, hoy sustituido por los ejes logarítmicos de cualquier programa: una ley potencial se ve como una recta y su exponente es la pendiente.

(d) Ajustar una parábola. Nada obliga a que el modelo sea una recta. Con los datos del primer ejemplo, el mejor polinomio de grado dos es $y = 0{,}026786x^2 + 1{,}8325x + 0{,}23$, con suma de cuadrados 0,101214, algo menor que los 0,128 de la recta. Pero la mejora es minúscula y ha costado un parámetro extra: casi siempre conviene el modelo más simple que ajuste razonablemente, porque añadir grados siempre reduce el error y acaba ajustando el ruido.

Cuidado

Un R^2 alto no demuestra que la relación sea lineal ni, mucho menos, que sea causal: existen conjuntos de datos con formas radicalmente distintas y el mismo R^2, por lo que dibujar la nube sigue siendo imprescindible. Segundo aviso: al linealizar un ajuste exponencial se minimizan los errores sobre $\ln y$, no sobre y, lo que pondera más los valores pequeños; el resultado es excelente como punto de partida, pero no coincide exactamente con el mejor ajuste exponencial. Y tercero: aumentar el grado del polinomio siempre reduce la suma de cuadrados, hasta llegar a cero con grado $n-1$; eso no es un ajuste mejor sino el fenómeno del *sobreajuste*, y produce modelos que predicen fatal.

Pausa para pensar

Legendre publicó el método en 1805 y Gauss respondió en 1809 que él lo usaba desde 1795, desatando una de las disputas de prioridad más agrias de la época. Lo que zanjó el asunto a favor de Gauss ante la opinión pública fue un éxito espectacular: en 1801 el asteroide Ceres se había perdido tras unas pocas observaciones, y Gauss, con esos datos escasos y su método, predijo dónde volvería a aparecer. Los astrónomos lo encontraron casi exactamente donde él dijo. Fue la demostración de que ajustar datos con error no es una chapuza sino una técnica con poder predictivo real, y el punto de partida de toda la estadística moderna.

14.11. Integración numérica

Si no hay primitiva, se aproxima el área directamente. Se parte $[a, b]$ en n trozos iguales de anchura $h = (b-a)/n$ y se sustituye la función en cada trozo por algo integrable.

Regla: Regla de los trapecios

$$\int_a^b f(x)\,\mathrm{d}x \approx h\left[\frac{f(x_0)}{2} + f(x_1) + f(x_2) + \cdots + f(x_{n-1}) + \frac{f(x_n)}{2}\right].$$

Se sustituye la curva por segmentos: los extremos pesan la mitad porque solo pertenecen a un trapecio.

Regla: Regla de Simpson (con n par)

$$\int_a^b f(x)\,\mathrm{d}x \approx \frac{h}{3}\Big[f(x_0)+4f(x_1)+2f(x_2)+4f(x_3)+\cdots+4f(x_{n-1})+f(x_n)\Big],$$

con los coeficientes alternando 4 y 2 en el interior. Se sustituye la curva por *parábolas* que pasan por cada tres puntos consecutivos.

Ejemplo 14.38 (Calculando $\ln 2$ sin logaritmos)**.** Como $\int_1^2 \frac{\mathrm{d}x}{x} = \ln 2 = 0{,}69314718\ldots$, la integral sirve de banco de pruebas:

n	Trapecios	Simpson
2	0,70833333	0,69444444
4	0,69702381	0,69325397
8	0,69412185	0,69315453

Con $n = 8$, los trapecios aciertan tres cifras y Simpson acierta cinco, con exactamente el mismo número de evaluaciones de la función.

Observación 14.39. El error de los trapecios es del orden de h^2 y el de Simpson del orden de h^4. Se ve en la tabla: al pasar de $n = 4$ a $n = 8$ el error de los trapecios se divide por 4 y el de Simpson por 16.

Y hay un detalle elegante: Simpson es *exacto* para polinomios de grado hasta tres, aunque solo use parábolas. Por ejemplo $\int_0^1 x^2\mathrm{d}x = 1/3$, y Simpson con $n = 4$ da 0,33333333 exacto, mientras que los trapecios dan 0,34375. La razón es que el error de la parábola por encima se cancela con el de por debajo cuando el término cúbico es simétrico respecto del punto medio.

14.12. El error de verdad: redondeo y cancelación

Hay dos fuentes de error y conviene no confundirlas. El **error de truncamiento** es el del método: la parábola no es la curva. El **error de redondeo** es el de la máquina, y viene del punto flotante de los sistemas de numeración. El primero baja al afinar la partición; el segundo *sube*, porque hay más operaciones que redondear. Por eso afinar indefinidamente no mejora el resultado: hay un punto óptimo y pasado ese punto se empeora.

Cuidado

[Cancelación catastrófica] Restar dos números casi iguales destruye cifras significativas. Resolvamos $x^2 + 10^8 x + 1 = 0$ con la fórmula de siempre. La raíz pequeña sale de

$$x = \frac{-10^8 + \sqrt{10^{16} - 4}}{2},$$

donde se restan dos números que coinciden en dieciséis cifras. Con aritmética de doble precisión el resultado es $-7{,}45 \cdot 10^{-9}$, y el valor correcto es $-1{,}00 \cdot 10^{-8}$: *no hay ni una sola cifra correcta*, ni siquiera la primera.

El arreglo es puramente algebraico. Multiplicando y dividiendo por el conjugado,

$$x = \frac{-2}{10^8 + \sqrt{10^{16} - 4}},$$

que ya no resta nada y da $-1{,}0000000 \cdot 10^{-8}$, correcto hasta la última cifra. La misma fórmula, escrita de otro modo, y toda la diferencia del mundo. Racionalizar, que en los límites elementales era un truco para levantar indeterminaciones, es aquí una necesidad de ingeniería.

Idea clave

La moraleja del capítulo entero: en cálculo numérico no basta con que la fórmula sea correcta; tiene que ser correcta *y* estable. Dos expresiones matemáticamente idénticas pueden comportarse de forma radicalmente distinta en una máquina finita, y elegir entre ellas es una decisión matemática, no informática. El análisis numérico es, en el fondo, el capítulo del análisis real que se toma en serio que los recursos son limitados.

Regla: Reconocer y evitar la cancelación catastrófica

Un ordenador guarda los números con unas 16 cifras decimales, y eso basta casi siempre. El problema no es la precisión: es *restar dos números casi iguales*, porque entonces las cifras buenas se anulan entre sí y afloran las malas.

1. **Localizar en la fórmula toda resta $a - b$ con $a \approx b$.** Esos son los únicos puntos peligrosos; las sumas, los productos y las divisiones son inofensivos.
2. **Estimar cuántas cifras se pierden:** si a y b coinciden en k cifras significativas, la resta pierde esas k cifras. Restando dos números que coinciden en 8 cifras, de las 16 disponibles quedan 8.
3. **Reescribir la expresión para que la resta desaparezca.** Las tres herramientas habituales son multiplicar y dividir por el conjugado, usar una identidad trigonométrica, o sustituir por el desarrollo de Taylor cuando el argumento es pequeño.
4. **En la ecuación de segundo grado**, si $b^2 \gg 4ac$ una de las dos raíces se calcula mal. Se obtiene la "buena" con la fórmula habitual y la otra por la relación de Vieta $x_1 x_2 = c/a$.
5. **Al sumar muchos términos de magnitudes distintas**, sumar de menor a mayor: si no, los sumandos pequeños se pierden enteros contra un acumulador grande.
6. **No comparar números en coma flotante con igualdad.** En lugar de `a == b`, exigir $|a - b| < \varepsilon$ con una tolerancia razonable.

7. **Verificar** recalculando con más precisión o con la fórmula reescrita: si los dos resultados difieren en cifras que se esperaban buenas, la culpable es siempre una resta.

Ejemplo 14.40 (cinco cuentas que un ordenador hace mal)**.** (a) *La ecuación de segundo grado.* Se resuelve $x^2 - 10^8 x + 1 = 0$. Con la fórmula de siempre,

$$x_2 = \frac{10^8 - \sqrt{10^{16} - 4}}{2}.$$

El radicando es $99\,999\,999\,999\,999\,996$ y su raíz es prácticamente 10^8: se están restando dos números que coinciden en dieciséis cifras. En doble precisión el resultado que sale es

$$x_2 = 7{,}450580596923828 \cdot 10^{-9},$$

cuando el valor verdadero es 10^{-8}. Un error del 25 %, en una máquina con dieciséis cifras exactas. Reescribiendo con el conjugado,

$$x_2 = \frac{2}{10^8 + \sqrt{10^{16} - 4}} = 1{,}000000000000000 \cdot 10^{-8},$$

exacto. La resta ha desaparecido y con ella el problema.

(b) *El límite* $\dfrac{1 - \cos x}{x^2} \to \dfrac{1}{2}$. Al evaluarlo numéricamente cerca de cero:

x	fórmula directa	$\dfrac{2\sin^2(x/2)}{x^2}$
10^{-4}	0,4999999970	0,4999999996
10^{-6}	0,5000444503	0,5000000000
10^{-8}	0,0000000000	0,5000000000

Con $x = 10^{-8}$ el ordenador da cero: $\cos x$ vale 1 hasta la última cifra guardada, la resta da exactamente 0 y el límite desaparece. La versión con la identidad $1 - \cos x = 2\sin^2(x/2)$ no resta nada y acierta siempre.

(c) *El orden de una suma importa.* Sumando $\sum_{k=1}^{10^6} 1/k$ en precisión simple:

$$\text{de } k = 1 \text{ a } 10^6 : \ 14{,}357358, \qquad \text{de } k = 10^6 \text{ a } 1 : \ 14{,}392652,$$

frente al valor correcto 14,392727. Sumando hacia adelante, cuando el acumulador ya vale 14 y el término vale 10^{-6}, la suma es tan desigual que el término se pierde entero. Hacia atrás los pequeños se juntan primero y llegan a pesar. La suma no es asociativa en coma flotante, y este ejemplo lo demuestra con dos líneas de código.

(d) *El épsilon de máquina.* Es el menor ε con $1 + \varepsilon \neq 1$ en la aritmética de la máquina:

$$\varepsilon_{\text{doble}} = 2{,}22 \cdot 10^{-16}, \qquad \varepsilon_{\text{simple}} = 1{,}19 \cdot 10^{-7}.$$

Es el tamaño relativo del "ruido de fondo" de todas las operaciones, y explica por qué la precisión simple pierde la mitad de las cifras del apartado (c).

(e) *Lo que todo el mundo ha visto alguna vez.* En cualquier lenguaje de programación,

$$0{,}1 + 0{,}2 = 0{,}30000000000000004 \neq 0{,}3.$$

No es un error del ordenador: 0,1 no tiene representación finita en base 2, igual que 1/3 no la tiene en base 10. El los sistemas de numeración explica el mecanismo. La consecuencia práctica es la del paso 6 de la regla: nunca comparar decimales con igualdad exacta.

Cuidado

El síntoma es traicionero, porque el programa no falla ni avisa: devuelve un número perfectamente presentable, con todos sus decimales, y equivocado. En (a) el resultado tenía dieciséis cifras impresas y solo la primera era correcta. Por eso la verificación independiente —recalcular por otra vía y comparar— no es una manía de perfeccionistas, es la única defensa.

Pausa para pensar

Conviene separar dos cosas que suelen confundirse. El *condicionamiento* es una propiedad del problema: mide cuánto amplifica la respuesta un error en los datos, y un problema mal condicionado lo está para cualquier algoritmo. La *estabilidad* es una propiedad del algoritmo: mide si el método añade error propio. Los apartados (a) y (b) son problemas bien condicionados resueltos por algoritmos inestables, y por eso tienen arreglo: basta cambiar de fórmula. Cuando lo que está mal es el condicionamiento, no hay fórmula que salve nada, y la única salida es reformular el problema.

Idea clave

La aritmética del ordenador no es la de $\mathbb{R}$. No es asociativa, no es distributiva, tiene un mayor elemento y un menor positivo, y entre dos reales consecutivos representables hay un hueco proporcional a su tamaño. Es un conjunto finito de racionales con operaciones redondeadas, y programar cálculo numérico consiste en buena medida en recordar esa diferencia. Todo el análisis de los capítulos anteriores vive en $\mathbb{R}$; el ordenador vive en otro sitio.

14.13. Sistemas lineales: eliminación con pivoteo y condicionamiento

Resolver $Ax = b$ es, en teoría, un problema cerrado: en el capítulo el cálculo matricial elemental vimos la eliminación de Gauss y en el 7 el marco conceptual. Pero en el ordenador aparecen dos fenómenos que la teoría no ve: el *orden* en que se eliminan las incógnitas cambia el resultado, y hay matrices para las que ningún método puede dar una respuesta fiable. Este apartado explica ambos.

Regla: Resolver un sistema lineal en el ordenador

1. **Elimina por columnas, de izquierda a derecha.** En la columna k se anula todo lo que hay debajo del pivote a_{kk} restando a cada fila $i > k$ la fila k multiplicada por $m_{ik} = a_{ik}/a_{kk}$. Al final queda un sistema triangular superior que se resuelve hacia atrás.
2. **Nunca uses el pivote que te toque: usa el mayor.** Antes de eliminar en la columna k, busca entre las filas $k, k+1, \dots, n$ la que tenga $|a_{ik}|$ máximo e intercámbiala con la fila k. Esto es el *pivoteo parcial* y no es un lujo: sin él, un pivote pequeño produce multiplicadores enormes que destruyen la información de las demás filas.
3. **Cuenta el trabajo antes de lanzarlo.** La eliminación cuesta del orden de $n^3/3$ operaciones y la sustitución hacia atrás sólo $n^2/2$. Duplicar n multiplica el coste por ocho. Para $n = 1000$ son unos $3{,}3 \times 10^8$ productos.
4. **Reaprovecha: si vas a resolver con varios b, guarda la factorización.** Los multiplicadores m_{ik} y la triangular final son la descomposición $A = LU$ (con permutaciones, $PA = LU$). Guardarla cuesta lo mismo que una eliminación y luego cada nuevo b sale en n^2 operaciones. El determinante es el producto de los pivotes, con un signo por cada intercambio.
5. **Mide el condicionamiento antes de creerte el resultado.** El número $\kappa(A) = \|A\|\,\|A^{-1}\|$ mide cuánto amplifica A los errores: si los datos tienen un error relativo δ, la solución puede tenerlo de $\kappa(A)\cdot\delta$. Con $\kappa(A) \approx 10^k$ pierdes unas k cifras significativas.
6. **No confundas residuo pequeño con solución buena.** El residuo $r = b - A\hat{x}$ puede ser diminuto y el error $\hat{x} - x$ enorme; eso es exactamente lo que significa estar mal condicionado. Comprobar que $A\hat{x} \approx b$ no demuestra nada por sí solo.
7. **Para sistemas grandes y dispersos, itera en vez de eliminar.** Jacobi actualiza todas las incógnitas con los valores viejos; Gauss–Seidel usa ya los nuevos y suele converger el doble de rápido. Ambos convergen seguro si A es de diagonal estrictamente dominante ($|a_{ii}| > \sum_{j\neq i} |a_{ij}|$); si no lo es, reordena las filas antes de rendirte.

Ejemplo 14.41 (Seis sistemas, seis lecciones)**.** (a) *La eliminación en limpio.* Resolvamos

$$\begin{cases} 2x + y - z = 8, \\ -3x - y + 2z = -11, \\ -2x + y + 2z = -3. \end{cases}$$

Pivote $a_{11} = 2$. Los multiplicadores son $m_{21} = -3/2$ y $m_{31} = -1$; restando obtenemos las filas $(0, \frac{1}{2}, \frac{1}{2} \mid 1)$ y $(0, 2, 1 \mid 5)$. Nuevo pivote $\frac{1}{2}$, multiplicador $m_{32} = 2/(1/2) = 4$, y la tercera fila queda $(0, 0, -1 \mid 1)$. Hacia atrás: $z = -1$, luego $\frac{1}{2}y + \frac{1}{2}(-1) = 1$ da $y = 3$, y por fin $2x + 3 + 1 = 8$ da $x = 2$. La solución es $(2, 3, -1)$ y el determinante es el producto de pivotes, $2 \cdot \frac{1}{2} \cdot (-1) = -1$.

(b) *Un pivote pequeño arruina el cálculo.* Consideremos

$$0{,}0001\,x + y = 1, \qquad x + y = 2,$$

cuya solución exacta es $x = 1{,}000100010001\ldots$, $y = 0{,}999899989999\ldots$ Trabajemos con tres cifras significativas, como haría una máquina muy pobre. Sin pivoteo el multiplicador es $m = 1/0{,}0001 = 10\,000$; la segunda ecuación pasa a $(1 - 10\,000)y = 2 - 10\,000$, es decir $-9999\,y = -9998$, que redondeado a tres cifras es $-10\,000\,y = -10\,000$ y da $y = 1{,}00$. Sustituyendo hacia atrás, $x = (1 - 1{,}00)/0{,}0001 = 0$. ¡El resultado es $x = 0$ en lugar de $x \approx 1$: un error del cien por cien! Con pivoteo intercambiamos las filas primero: $m = 0{,}0001$, la segunda ecuación queda $0{,}9999\,y = 0{,}9998$, o sea $1{,}00\,y = 1{,}00$, luego $y = 1{,}00$ y $x = 2 - 1{,}00 = 1{,}00$. Ahora sí, tres cifras correctas.

(c) *Mal condicionamiento: dos rectas casi paralelas.* Sea

$$A = \begin{pmatrix} 1 & 1 \\ 1 & 1{,}0001 \end{pmatrix}, \qquad \kappa(A) \approx 4{,}0002 \times 10^{4}.$$

Con $b = (2,\ 2{,}0001)$ la solución es $(1, 1)$. Si movemos el segundo dato una diezmilésima, $b = (2,\ 2{,}0002)$, la solución salta a $(0, 2)$. Un cambio relativo de 5×10^{-5} en los datos produce un cambio del cien por cien en la respuesta; el factor de amplificación es precisamente del orden de $\kappa(A)$.

(d) *Residuo minúsculo, error enorme.* Con el mismo A y $b = (2,\ 2{,}0001)$, propongamos la solución falsa $\hat{x} = (0, 2)$. Entonces $A\hat{x} = (2,\ 2{,}0002)$ y el residuo es $r = b - A\hat{x} = (0,\ -10^{-4})$, de norma 10^{-4}: parece un éxito. Pero el error verdadero es $\hat{x} - x = (-1, 1)$, de norma $\sqrt{2} \approx 1{,}414$. La comprobación "sustituyo y sale casi b" es aquí completamente engañosa.

(e) *Las matrices de Hilbert, el peor enemigo clásico.* La matriz H_n con $(H_n)_{ij} = 1/(i + j - 1)$ tiene

$$\kappa(H_3) \approx 5{,}24 \times 10^{2}, \quad \kappa(H_5) \approx 4{,}77 \times 10^{5}, \quad \kappa(H_{10}) \approx 1{,}60 \times 10^{13}.$$

Tomemos $n = 5$, la solución exacta $x = (1, 1, 1, 1, 1)$ y $b = H_5 x$. Si perturbamos sólo la primera componente de b en 10^{-6}, la solución calculada pasa a ser

$$(1{,}000025,\ 0{,}9997,\ 1{,}00105,\ 0{,}9986,\ 1{,}00063).$$

Una perturbación de una millonésima ha movido la respuesta en más de una milésima: mil veces más, coherente con $\kappa \approx 5 \times 10^{5}$ actuando sobre datos de tamaño 1. Con H_{10} y aritmética de doble precisión ($\varepsilon \approx 2{,}2 \times 10^{-16}$) no queda ninguna cifra fiable.

(f) *Jacobi y Gauss–Seidel comparados.* Para

$$\begin{pmatrix} 4 & -1 & 0 \\ -1 & 4 & -1 \\ 0 & -1 & 4 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 2 \\ 4 \\ 10 \end{pmatrix}, \qquad \text{solución } (1, 2, 3),$$

la matriz es de diagonal estrictamente dominante, así que ambos métodos convergen. Partiendo de $(0, 0, 0)$:

k	Jacobi			Gauss–Seidel		
1	0,5000	1,0000	2,5000	0,5000	1,1250	2,78125
2	0,7500	1,7500	2,7500	0,78125	1,890625	2,972656
3	0,9375	1,8750	2,9375	0,972656	1,986328	2,996582
4	0,96875	1,96875	2,96875	0,996582	1,998291	2,999573
5	0,992188	1,984375	2,992188	0,999573	1,999786	2,999947
6	0,996094	1,996094	2,996094	0,999947	1,999973	2,999993

En seis pasos Gauss–Seidel lleva cinco cifras correctas y Jacobi apenas dos. La explicación está en el radio espectral de la matriz de iteración: $\sqrt{2}/4 \approx 0{,}3536$ para Jacobi y $1/8 = 0{,}125$ para Gauss–Seidel; como $0{,}125 = (0{,}3536)^2$, cada paso de Gauss–Seidel vale por dos de Jacobi. Jacobi necesita ocho pasos para llegar a $(0{,}999512,\ 1{,}999512,\ 2{,}999512)$.

Cuidado

La dominancia diagonal no es un adorno. Para

$$x + 3y = 7, \qquad 4x + y = 9, \qquad \text{solución } (20/11,\ 19/11) = (1{,}8182,\ 1{,}7273),$$

Gauss–Seidel desde $(0, 0)$ produce $(7, -19)$, $(64, -247)$, $(748, -2983)$, $(8956, -35815)\ldots$ diverge sin remedio, porque el radio espectral de su matriz de iteración es 12. Basta *intercambiar las dos ecuaciones* —el sistema es el mismo— para que la matriz $\left(\begin{smallmatrix} 4 & 1 \\ 1 & 3 \end{smallmatrix}\right)$ sea dominante: entonces el radio baja a $1/12$ y la iteración da (2,25; 1,5833), (1,8542; 1,7153), (1,8212; 1,7263), (1,8184; 1,7272), convergiendo a toda velocidad. Reordenar filas es gratis; no hacerlo puede costar todo.

Pausa para pensar

¿Por qué el pivoteo parcial funciona? Al elegir como pivote el elemento de mayor módulo de la columna, todos los multiplicadores cumplen $|m_{ik}| \le 1$. Así, al restar m_{ik} veces la fila del pivote, nunca se amplifica lo que ya había en la fila i: los errores de redondeo acumulados se transportan, pero no crecen. Sin pivoteo, un multiplicador de 10 000 como el del apartado (b) hace que el valor original de la fila —el 1 y el 2— se pierda por completo frente a los términos gigantes que se le suman.

Idea clave

Hay una simetría interesante entre este apartado y el de la cancelación catastrófica. Allí el problema era la *aritmética*: una fórmula mal escrita perdía cifras aunque el problema fuese bueno. Aquí el problema puede ser el *problema mismo*: si $\kappa(A)$ es grande, ninguna reescritura salva el cálculo, porque la respuesta verdaderamente depende de forma brutal de los datos. La distinción se llama *estabilidad del algoritmo* frente a *condicionamiento del problema*, y es la dicotomía central del análisis numérico. En el lenguaje del capítulo 32, $\kappa(A)$ no es más que la norma del operador A multiplicada por la de su inverso.

14.14. Métodos iterativos: Jacobi, Gauss–Seidel y el método de la potencia

La eliminación gaussiana resuelve cualquier sistema en un número finito de pasos, pero su coste crece como n^3 y destruye la estructura de ceros de la matriz. Para los sistemas enormes y dispersos que aparecen al discretizar una ecuación en derivadas parciales, la alternativa es renunciar a la solución exacta y acercarse a ella iterando, que es mucho más barato.

Definición 14.42. Dado $A\vec{x} = \vec{b}$ con $a_{ii} \neq 0$, el *método de Jacobi* genera la sucesión

$$x_i^{(k+1)} = \frac{1}{a_{ii}} \left(b_i - \sum_{j \neq i} a_{ij} x_j^{(k)} \right),$$

mientras que el *método de Gauss–Seidel* usa ya los valores nuevos en cuanto están disponibles. Ambos convergen para cualquier vector inicial si A es *estrictamente diagonal dominante*, es decir si $|a_{ii}| > \sum_{j \neq i} |a_{ij}|$ para toda fila. El *método de la potencia* aproxima el autovalor de mayor módulo iterando $\vec{v}^{(k+1)} = A\vec{v}^{(k)}$ y normalizando en cada paso.

Idea clave

Los tres métodos comparten la misma idea, que es la del punto fijo: reescribir el problema como "$\vec{x}$ es lo que sale al aplicar cierta transformación a $\vec{x}$" y luego aplicarla una y otra vez. En Jacobi se despeja cada incógnita de su propia ecuación, tratando las demás como conocidas; el método converge si la diagonal es lo bastante grande como para que cada despeje apenas dependa de los otros. Gauss–Seidel es la mejora obvia: si ya has calculado el valor nuevo de x_1, sería absurdo seguir usando el viejo al calcular x_2. Y el método de la potencia explota que al multiplicar repetidamente por A, la componente del autovector dominante crece más deprisa que las demás y acaba comiéndose el vector entero.

Regla: Resolver un sistema iterando y estimar un autovalor

El procedimiento común a los tres métodos:

1. **Comprueba la dominancia diagonal** antes de empezar; si no la hay, reordena las filas para conseguirla si es posible.
2. **Despeja cada incógnita de su ecuación**, la i-ésima de la fila i-ésima, y escribe la fórmula de recurrencia.
3. **Elige un vector inicial**, típicamente el nulo, porque la convergencia no depende de él cuando hay dominancia.
4. **Itera y mide el cambio** $\|\vec{x}^{(k+1)} - \vec{x}^{(k)}\|$ en cada paso; ese es el criterio de parada práctico, pues el error verdadero no se conoce.
5. **Usa Gauss–Seidel si puedes**: converge aproximadamente al doble de velocidad que Jacobi en los casos habituales, y no cuesta más.

6. **Para el autovalor dominante, normaliza en cada paso** dividiendo por la componente de mayor módulo, y observa cómo ese factor converge al autovalor.
7. **Refina con el cociente de Rayleigh** $\lambda \approx \dfrac{\vec{v}^T A\vec{v}}{\vec{v}^T\vec{v}}$, que da más cifras correctas que el propio factor de normalización.

Ejemplo 14.43 (Un sistema resuelto por los dos métodos)**. (a) El sistema.** Sea

$$\begin{cases} 4x - y = 2, \\ -x + 4y - z = 4, \\ -y + 4z = 10, \end{cases}$$

cuya solución exacta es $(1, 2, 3)$. La matriz es estrictamente diagonal dominante: en cada fila, 4 supera a la suma de los valores absolutos de los demás coeficientes, que vale 1 o 2.

(b) Jacobi. Despejando, $x = \frac{2+y}{4}$, $y = \frac{4+x+z}{4}$, $z = \frac{10+y}{4}$. Partiendo de $(0, 0, 0)$ se obtiene sucesivamente (0,5; 1; 2,5), (0,75; 1,75; 2,75), (0,9375; 1,875; 2,9375), (0,96875; 1,96875; 2,96875). Los errores máximos son 1, 0,25, 0,125, 0,03125: se dividen aproximadamente entre dos en cada paso, es decir, convergencia lineal. Tras ocho iteraciones el error es 0,000488.

(c) Gauss–Seidel. Con la misma recurrencia pero usando los valores recién calculados, la primera iteración da $x = 0{,}5$, luego $y = \frac{4+0{,}5+0}{4} = 1{,}125$ y después $z = \frac{10+1{,}125}{4} = 2{,}78125$. Los errores máximos son 0,875, 0,21875, 0,027344, 0,003418, 0,000427, 0,000053: cada paso divide el error entre ocho, y en seis iteraciones se ha llegado más lejos que Jacobi en ocho.

(d) Por qué importa. Con una matriz 3×3 la eliminación gaussiana es obviamente preferible. Pero al discretizar la ecuación del calor en una malla de 1000×1000 puntos aparece un sistema de un millón de incógnitas cuya matriz tiene solo cinco elementos no nulos por fila. La eliminación exigiría del orden de 10^{18} operaciones y llenaría de ceros que dejan de serlo toda la memoria disponible; Gauss–Seidel cuesta cinco millones de multiplicaciones por iteración y no almacena nada nuevo. Ahí la elección no es de gusto sino de posibilidad.

Ejemplo 14.44 (Condicionamiento y el autovalor dominante)**. (a) Un sistema traicionero.** La matriz $\begin{pmatrix} 1 & 1 \\ 1 & 1{,}0001 \end{pmatrix}$ tiene número de condición 40 002. Con $\vec{b} = (2;\ 2{,}0001)$ la solución es exactamente $(1, 1)$; si el segundo dato cambia a 2,0002, es decir una milésima de tanto por ciento, la solución salta a $(0, 2)$. No hay ningún error de programación: el problema mismo amplifica el error de entrada, y ningún método puede evitarlo.

(b) Las matrices de Hilbert. Con $H_{ij} = \frac{1}{i+j-1}$, el número de condición vale 19,2815 para $n = 2$ y 524,0568 para $n = 3$, y sigue creciendo de forma explosiva; para $n = 10$ supera 10^{13}. Son el ejemplo clásico de matriz mal condicionada, y aparecen justo al ajustar polinomios de grado alto por mínimos cuadrados, lo que explica por qué ese ajuste es numéricamente delicado.

(c) El método de la potencia. Sea $A = \begin{pmatrix} 2 & 1 \\ 1 & 3 \end{pmatrix}$, con autovalores $\frac{5 \pm \sqrt{5}}{2}$, es decir 3,618034 y 1,381966. Partiendo de $\vec{v} = (1, 0)$ y normalizando por la componente mayor, los factores

obtenidos son 2, 2,5, 4, 3,75, 3,666667, y tras diez iteraciones 3,618421: converge al autovalor dominante.

(d) La velocidad de esa convergencia. El error se reduce en cada paso por el factor $\left|\frac{\lambda_2}{\lambda_1}\right| = \frac{1{,}381966}{3{,}618034} = 0{,}381966$, que es lo que gobierna todo el método: cuanto más separados están los dos autovalores mayores, más rápido converge, y si coinciden en módulo el método no converge. El vector se estabiliza en $(0{,}618182;\ 1)$, muy cerca del autovector exacto, cuya primera componente es $\frac{\sqrt{5}-1}{2} = 0{,}618034$, la razón áurea. Y el cociente de Rayleigh sobre ese vector da 3,618034, con más cifras correctas que el propio factor de normalización.

Cuidado

La dominancia diagonal es condición *suficiente* para la convergencia, no necesaria: hay sistemas que convergen sin ella y sistemas que, sin ella, divergen alegremente. Comprobarla es barato y da tranquilidad. Segundo aviso: el criterio de parada basado en $\|\vec{x}^{(k+1)} - \vec{x}^{(k)}\|$ puede engañar cuando la convergencia es muy lenta, porque dos iteraciones consecutivas se parecen mucho aunque ambas estén lejos de la solución; conviene comprobar además el residuo $\|A\vec{x}^{(k)} - \vec{b}\|$. Y tercero: un residuo pequeño no garantiza un error pequeño si la matriz está mal condicionada, que es exactamente la lección del primer ejemplo.

Pausa para pensar

Gauss describió su método iterativo en una carta de 1823 a su alumno Gerling, con una frase que merece ser citada: decía que el procedimiento se podía hacer "medio dormido, o pensando en otras cosas". Seidel lo publicó en 1874, y el nombre conjunto se impuso después. Lo curioso es que Gauss lo consideraba un método de segunda, útil solo cuando la eliminación resultaba pesada; no podía sospechar que siglo y medio después, con matrices de millones de filas, esa herramienta "de segunda" sería la única practicable y que la eliminación exacta quedaría relegada a los problemas pequeños.

14.15. Ecuaciones diferenciales: de Euler a Runge–Kutta

En el capítulo 12 resolvimos ecuaciones diferenciales con fórmulas. Pero la inmensa mayoría de los problemas de valor inicial

$$y' = f(t, y), \qquad y(t_0) = y_0$$

no tiene solución expresable con funciones elementales. Lo que sí se puede siempre es avanzar la solución paso a paso, y de eso trata este apartado: cómo se avanza, cuánto error se comete y cuándo el método explota.

Regla: Integrar numéricamente una ecuación diferencial

1. **Escribe el problema en forma normal.** Necesitas y' despejada: $y' = f(t, y)$ más el dato inicial. Si la ecuación es de orden n, conviértela antes en un sistema de n ecuaciones de primer orden llamando $u_1 = y$, $u_2 = y'$, …; todos los métodos de

aquí funcionan igual sobre vectores.

2. **Elige el paso h y la malla.** Con n pasos sobre $[a,b]$ es $h=(b-a)/n$ y $t_k=a+kh$. El paso pequeño da más precisión pero más operaciones y más redondeo acumulado; hay un punto de equilibrio, igual que en la derivación numérica.
3. **Euler: avanza por la tangente.** $y_{k+1}=y_k+h\,f(t_k,y_k)$. Es el método más simple posible y el patrón de todos los demás: estimar la pendiente en el intervalo y multiplicar por h. Su error global es de orden h, es decir, de *primer* orden.
4. **Heun: promedia dos pendientes.** Calcula $k_1=f(t_k,y_k)$, predice $\tilde{y}=y_k+hk_1$, calcula $k_2=f(t_k+h,\tilde{y})$ y avanza con la media: $y_{k+1}=y_k+\frac{h}{2}(k_1+k_2)$. Cuesta dos evaluaciones de f por paso y sube a *segundo* orden.
5. **Runge–Kutta clásico: cuatro pendientes ponderadas.** Con $k_1=f(t_k,y_k)$, $k_2=f(t_k+\frac{h}{2},y_k+\frac{h}{2}k_1)$, $k_3=f(t_k+\frac{h}{2},y_k+\frac{h}{2}k_2)$ y $k_4=f(t_k+h,y_k+hk_3)$, se toma $y_{k+1}=y_k+\frac{h}{6}(k_1+2k_2+2k_3+k_4)$. Cuatro evaluaciones por paso, *cuarto* orden: es el estándar de la práctica.
6. **Verifica el orden empíricamente.** Repite el cálculo con h y con $h/2$ y divide los errores. Si el cociente tiende a 2 el método es de orden 1; si tiende a 4, de orden 2; si tiende a 16, de orden 4. Es la comprobación más barata de que has programado bien el método.
7. **Vigila la estabilidad, no sólo la precisión.** Para $y'=\lambda y$ con $\lambda<0$, Euler multiplica por $1+h\lambda$ en cada paso, así que sólo decae si $|1+h\lambda|<1$, o sea $h<2/|\lambda|$. Si $|\lambda|$ es grande el problema es *rígido* y el paso queda limitado por la estabilidad, no por la exactitud; entonces hay que usar métodos implícitos como el Euler regresivo $y_{k+1}=y_k+hf(t_{k+1},y_{k+1})$.

Ejemplo 14.45 (La exponencial y otros seis experimentos)**.** Tomamos como banco de pruebas $y'=y$, $y(0)=1$ en $[0,1]$, cuya solución exacta es $y(t)=e^t$ y por tanto $y(1)=e=2{,}718281828459045$.

(a) *Euler a mano con cuatro pasos.* Con $h=0{,}25$ la recurrencia es $y_{k+1}=y_k+0{,}25\,y_k=1{,}25\,y_k$:

t	Euler	exacto e^t	error
0,25	1,250000	1,284025	0,034025
0,50	1,562500	1,648721	0,086221
0,75	1,953125	2,117000	0,163875
1,00	2,441406	2,718282	0,276876

El error crece a cada paso: no sólo se comete uno nuevo, sino que se arrastra el anterior. Obsérvese además que Euler siempre se queda *corto* aquí, porque e^t es convexa y la tangente va por debajo.

(b) *Heun a mano con dos pasos.* Con $h = 0{,}5$: primero $k_1 = 1$, $\tilde{y} = 1 + 0{,}5 = 1{,}5$, $k_2 = 1{,}5$, y

$$y_1 = 1 + \tfrac{0{,}5}{2}(1 + 1{,}5) = 1{,}625.$$

Segundo paso: $k_1 = 1{,}625$, $\tilde{y} = 1{,}625 + 0{,}8125 = 2{,}4375$, $k_2 = 2{,}4375$, y

$$y_2 = 1{,}625 + 0{,}25\,(1{,}625 + 2{,}4375) = 2{,}640625.$$

Con dos pasos Heun ya supera a Euler con dieciséis.

(c) *Los tres métodos en la misma tabla.* Valor calculado de $y(1)$:

n	Euler	Heun	RK4
1	2,00000000	2,50000000	2,7083333333
2	2,25000000	2,64062500	2,7173461914
4	2,44140625	2,69485569	2,7182099392
8	2,56578451	2,71184124	2,7182768444
16	2,63792850	2,71659352	2,7182815003

RK4 con *un solo paso* es más preciso que Euler con dieciséis. Y RK4 con cuatro pasos —dieciséis evaluaciones de f, el mismo coste que Euler con dieciséis pasos— acierta seis cifras frente a la única de Euler.

(d) *La confirmación del orden.* Errores absolutos frente a e:

n	Euler	Heun	RK4
2	$4{,}68 \times 10^{-1}$	$7{,}77 \times 10^{-2}$	$9{,}36 \times 10^{-4}$
4	$2{,}77 \times 10^{-1}$	$2{,}34 \times 10^{-2}$	$7{,}19 \times 10^{-5}$
8	$1{,}52 \times 10^{-1}$	$6{,}44 \times 10^{-3}$	$4{,}98 \times 10^{-6}$
16	$8{,}04 \times 10^{-2}$	$1{,}69 \times 10^{-3}$	$3{,}28 \times 10^{-7}$

Los cocientes entre filas consecutivas son 1,90 para Euler, 3,81 para Heun y 15,19 para RK4: exactamente los 2, 4 y 16 anunciados, con la aproximación esperable para h todavía no muy pequeño.

(e) *Un problema no lineal.* Para $y' = -y^2$, $y(0) = 1$, la solución exacta es $y(t) = 1/(1 + t)$, luego $y(1) = 0{,}5$:

n	Euler	Heun	RK4
2	0,375000	0,518447	0,50002881
4	0,449837	0,504562	0,50001002
8	0,476811	0,501066	0,50000072

Nada cambia: la no linealidad no molesta a ninguno de los tres métodos, porque sólo necesitan evaluar f, no resolver la ecuación.

(f) *Rigidez: cuando el paso lo manda la estabilidad.* Sea $y' = -15y$, $y(0) = 1$, con solución e^{-15t} y $y(1) = 3{,}06 \times 10^{-7}$. Euler multiplica por $1 + 15h$ con signo, es decir por $1 - 15h$:

h	factor $1 - 15h$	primeros valores	Euler en $t = 1$
0,100	$-0{,}5$	$-0{,}5$; 0,25; $-0{,}125$	$9{,}77 \times 10^{-4}$
0,125	$-0{,}875$	$-0{,}875$; 0,766; $-0{,}670$	0,343609
0,200	-2	-2; 4; -8	$-32{,}0$
0,250	$-2{,}75$	$-2{,}75$; 7,56; $-20{,}80$	57,19

La frontera está en $h = 2/15 \approx 0{,}1333$: por debajo la solución numérica decae (oscilando, porque el factor es negativo) y por encima explota con signos alternos. Ningún refinamiento del *método* arregla esto; hay que bajar h o cambiar a un método implícito.

(g) *RK4 tampoco es inmune.* En el mismo problema rígido, RK4 con $h = 0{,}2$ da $y(1) \approx 4{,}91$, un disparate, mientras que con $h = 0{,}1$ da $2{,}34 \times 10^{-6}$, razonable. La región de estabilidad de RK4 sobre el eje real negativo llega hasta $|h\lambda| \approx 2{,}785$; con $h = 0{,}2$ es $h\lambda = -3$, fuera; con $h = 0{,}1$ es $-1{,}5$, dentro. Ser de orden cuatro dice cómo de rápido converge el método, no si converge.

Cuidado

Es un error clásico creer que "más orden" significa "mejor siempre". El orden describe el comportamiento del error *cuando* $h \to 0$; con h grande un método de orden alto puede ser peor, e incluso inestable, como acabamos de ver. La regla práctica es: elige el paso mirando la estabilidad, y sólo después, dentro de la zona estable, elige el método mirando la precisión.

Pausa para pensar

¿Por qué Euler es de orden 1 globalmente si el desarrollo de Taylor dice que el error de *un* paso es $O(h^2)$? Porque hay $n = (b-a)/h$ pasos. Si cada uno aporta un error del orden de h^2 y se acumulan, el total es del orden de $n \cdot h^2 = (b-a)h$. El mismo argumento explica el resto: error local $O(h^{p+1})$ acumulado n veces da error global $O(h^p)$. Por eso a p se le llama el orden del método, y por eso se pierde siempre exactamente una potencia de h al pasar de lo local a lo global.

Idea clave

Los tres métodos son, en el fondo, tres cuadraturas disfrazadas. Integrando $y' = f$ entre t_k y t_{k+1} se obtiene

$$y(t_{k+1}) = y(t_k) + \int_{t_k}^{t_{k+1}} f(s, y(s))\,\mathrm{d}s,$$

y aproximar esa integral por el rectángulo izquierdo da Euler, por el trapecio da Heun, y por una variante de Simpson da RK4. No es casualidad que los órdenes 1, 2 y 4 coincidan con los del rectángulo, el trapecio y Simpson del apartado de integración numérica: es la

misma matemática vista desde otro ángulo.

14.16. Derivación numérica: el paso que no debe ser demasiado pequeño

La derivada se define en el capítulo los límites elementales como el límite del cociente incremental cuando $h \to 0$. La traducción ingenua al ordenador —"tomo h muy pequeño"— es una de las peores ideas del análisis numérico, y entender por qué es una de las lecciones más útiles de todo el capítulo.

Regla: Derivar numéricamente sin que el redondeo lo estropee

1. **Usa la diferencia centrada siempre que puedas.** La fórmula adelantada
$$f'(x) \approx \frac{f(x+h) - f(x)}{h}$$
tiene error $O(h)$, mientras que la centrada
$$f'(x) \approx \frac{f(x+h) - f(x-h)}{2h}$$
tiene error $O(h^2)$ con el mismo número de evaluaciones. Reserva la adelantada para los extremos de un intervalo o cuando no puedas evaluar a la izquierda.

2. **Conoce el término de error, no sólo su orden.** Por Taylor,
$$\frac{f(x+h) - f(x)}{h} = f'(x) + \frac{h}{2} f''(\xi), \qquad \frac{f(x+h) - f(x-h)}{2h} = f'(x) - \frac{h^2}{6} f'''(\xi).$$
En la centrada los términos impares se cancelan: ésa es exactamente la razón de que gane un orden gratis.

3. **Recuerda que el redondeo aporta un error que *crece* al bajar h.** Cada valor $f(x \pm h)$ se calcula con un error absoluto del orden de $\varepsilon|f(x)|$, y al dividir por h ese error se multiplica por $1/h$. La contribución del redondeo es del orden de $\varepsilon|f(x)|/h$.

4. **Suma los dos errores y minimiza.** El error total se comporta como $C_1 h^p + C_2 \varepsilon/h$. Derivando e igualando a cero se obtiene el paso óptimo: $h^* \sim \sqrt{\varepsilon} \approx 1{,}5 \times 10^{-8}$ para la fórmula adelantada y $h^* \sim \varepsilon^{1/3} \approx 6 \times 10^{-6}$ para la centrada, en doble precisión.

5. **No bajes de ahí.** Por debajo del paso óptimo el error *aumenta*: $f(x+h)$ y $f(x-h)$ se parecen tanto que su resta es pura cancelación catastrófica, y el cociente amplifica la basura. Con $h = 10^{-16}$ la respuesta puede ser exactamente 0.

6. **Para la segunda derivada, usa la diferencia segunda y sube el listón.** La fórmula
$$f''(x) \approx \frac{f(x+h) - 2f(x) + f(x-h)}{h^2}$$

tiene error de truncamiento $-\frac{h^2}{12}f^{(4)}(\xi)$, pero el redondeo entra dividido por h^2, así que el paso óptimo sube a $h^* \sim \varepsilon^{1/4} \approx 10^{-4}$: mucho más grande de lo que la intuición sugiere.

7. **Exprime el orden con extrapolación de Richardson.** Si $D(h) = f'(x) + Ch^2 + O(h^4)$, la combinación
$$\frac{4D(h/2) - D(h)}{3}$$
elimina el término en h^2 y deja error $O(h^4)$, sin necesidad de un h menor. Es el mismo truco que convierte trapecios en el método de Romberg.

Ejemplo 14.46 (Derivar $\sin x$ en $x = 1$ y ver el desastre en directo)**.** Queremos $f'(1)$ con $f(x) = \sin x$. El valor exacto es $\cos 1 = 0{,}5403023058681398$.

(a) *La fórmula adelantada, paso a paso hacia abajo.*

h	valor calculado	error
10^{-1}	0,497363752535	$4{,}294 \times 10^{-2}$
10^{-3}	0,539881480360	$4{,}208 \times 10^{-4}$
10^{-5}	0,540298098506	$4{,}207 \times 10^{-6}$
10^{-7}	0,540302264040	$4{,}183 \times 10^{-8}$
10^{-8}	0,540302302898	$2{,}970 \times 10^{-9}$
10^{-9}	0,540302358409	$5{,}254 \times 10^{-8}$
10^{-11}	0,540301137164	$1{,}169 \times 10^{-6}$

Hasta $h = 10^{-7}$ el error se divide por diez cada vez que h se divide por diez: es el orden 1 funcionando. El mínimo está en $h = 10^{-8}$, justo donde predice $\sqrt{\varepsilon} \approx 1{,}5 \times 10^{-8}$. Y a partir de ahí el error *sube*: con $h = 10^{-11}$ es cuatrocientas veces peor que con $h = 10^{-8}$.

(b) *Comprobemos que el término de error es el anunciado.* La cota dice error $\approx \frac{h}{2}|f''(1)| = \frac{h}{2}\sin 1 = 0{,}42074\,h$. Para $h = 10^{-3}$ eso da $4{,}207 \times 10^{-4}$, y la tabla marca $4{,}208 \times 10^{-4}$. La teoría no es una estimación vaga: acierta cuatro cifras.

(c) *La fórmula centrada, mucho mejor y con el mismo esfuerzo.*

h	valor calculado	error
10^{-1}	0,539402252170	$9{,}001 \times 10^{-4}$
10^{-2}	0,540293300875	$9{,}005 \times 10^{-6}$
10^{-3}	0,540302215818	$9{,}005 \times 10^{-8}$
10^{-4}	0,540302304968	$9{,}004 \times 10^{-10}$
10^{-5}	0,540302305857	$1{,}114 \times 10^{-11}$
10^{-6}	0,540302305896	$2{,}772 \times 10^{-11}$
10^{-8}	0,540302308449	$2{,}581 \times 10^{-9}$
10^{-11}	0,540301137164	$1{,}169 \times 10^{-6}$

Ahora el error se divide por cien cada vez que h se divide por diez: orden 2. La cota teórica es $\frac{h^2}{6}|f'''(1)| = \frac{h^2}{6}\cos 1 = 0{,}09005\,h^2$, que para $h = 10^{-2}$ da $9{,}005 \times 10^{-6}$, clavado. El mínimo cae en $h = 10^{-5}$, junto al $\varepsilon^{1/3} \approx 6{,}1 \times 10^{-6}$ previsto, y allí el error es 10^{-11}: cien veces mejor que el mejor resultado de la fórmula adelantada.

(d) *La comparación que importa.* Con la fórmula adelantada, por mucho que se afine, no se pasa de unas ocho cifras correctas. Con la centrada se llega a once. La diferencia no viene de calcular más, sino de calcular *mejor*: las dos fórmulas usan dos evaluaciones de f por punto.

(e) *La segunda derivada y su paso óptimo sorprendente.* Con $f''(1) = -\sin 1 = -0{,}8414709848$:

h	valor calculado	error
10^{-2}	$-0{,}8414639726$	$7{,}012 \times 10^{-6}$
10^{-3}	$-0{,}8414709147$	$7{,}008 \times 10^{-8}$
10^{-4}	$-0{,}8414709818$	$3{,}025 \times 10^{-9}$
10^{-5}	$-0{,}8414702268$	$7{,}580 \times 10^{-7}$
10^{-6}	$-0{,}8415490527$	$7{,}807 \times 10^{-5}$

El óptimo está en $h = 10^{-4}$, un paso que a ojo parece "poco pequeño". Con $h = 10^{-6}$ —tres órdenes por debajo— ya se han perdido cuatro cifras. Y en efecto, la cota de truncamiento $\frac{h^2}{12}\sin 1 = 0{,}07012\,h^2$ da $7{,}012 \times 10^{-6}$ para $h = 10^{-2}$, exactamente lo observado.

(f) *Richardson: cuarto orden sin bajar el paso.* Con la fórmula centrada,

$$D(0{,}1) = 0{,}539402252170, \qquad D(0{,}05) = 0{,}540077208046,$$

con errores $9{,}0 \times 10^{-4}$ y $2{,}25 \times 10^{-4}$. Combinándolos,

$$\frac{4D(0{,}05) - D(0{,}1)}{3} = 0{,}540302193339,$$

con error $1{,}13 \times 10^{-7}$: dos mil veces mejor que el mejor de los dos ingredientes, y sin evaluar f en ningún punto más cercano que 0,05. Justo lo contrario de la estrategia ingenua de achicar h.

Cuidado

Aquí está el error conceptual más frecuente de todo el capítulo: creer que "h más pequeño" es siempre "más exacto", porque así funciona el límite matemático. En el papel, $h \to 0$ da la derivada exacta. En la máquina, $h \to 0$ da primero mejores resultados, luego peores, y al final $0/h$, es decir, cero. La derivada es un problema mal condicionado en el sentido del apartado anterior: pequeñas perturbaciones de f producen perturbaciones arbitrariamente grandes de f', y el ordenador perturba f siempre, aunque sea en la decimosexta cifra.

Pausa para pensar

Compárese esta situación con la de la integración numérica. Allí, refinar el paso mejoraba el resultado durante muchísimo tiempo antes de que el redondeo se notara, porque integrar *promedia* los errores en lugar de amplificarlos. Derivar es la operación inversa y se comporta al revés: amplifica. Ésa es la razón profunda de que, cuando se puede elegir, un método numérico que integre sea preferible a uno que derive; y de que en el capítulo 12 los métodos de paso avancen integrando f, nunca derivando y.

Idea clave

Si lo que se busca es la derivada de una función de la que se tiene la *fórmula* —y no sólo una tabla de valores—, existen dos alternativas que evitan el problema por completo. La primera es la derivación simbólica: aplicar las reglas del capítulo el estudio de funciones de forma automática y evaluar la expresión resultante, exacta salvo redondeo final. La segunda es la *diferenciación automática*, que propaga pares (f, f') a través de cada operación aritmética usando la regla del producto y la de la cadena; da la derivada con precisión de máquina, sin ningún h, y es la técnica que hace posible el entrenamiento de las redes neuronales modernas. Su generalización a varias variables produce directamente el gradiente y la jacobiana del capítulo 9.

14.17. Monte Carlo: cuando lanzar dados calcula mejor que una cuadratura

Todos los métodos vistos hasta aquí son deterministas: con los mismos datos dan siempre el mismo resultado. Existe una familia completamente distinta que usa números aleatorios y que, siendo mucho peor que Simpson en una dimensión, se vuelve imbatible en cuanto la dimensión crece.

Definición 14.47. El *método de Monte Carlo* para estimar $I = \int_D f$ consiste en generar N puntos aleatorios uniformes en D y aproximar

$$I \approx \operatorname{vol}(D) \cdot \frac{1}{N} \sum_{k=1}^{N} f(P_k).$$

Por la ley de los grandes números el estimador converge a I, y por el teorema central del límite su error típico es $\dfrac{\sigma \operatorname{vol}(D)}{\sqrt{N}}$, donde σ es la desviación típica de f sobre D. En el caso particular en que f es la función indicadora de una región, el método se reduce a *contar aciertos*: la fracción de puntos que caen dentro estima la fracción de volumen.

Idea clave

Lo decisivo de Monte Carlo es que su error, $\mathcal{O}(N^{-1/2})$, *no depende de la dimensión*. La regla de Simpson en una dimensión tiene error $\mathcal{O}(h^4)$, muchísimo mejor; pero en

dimensión d hace falta una malla de n^d puntos, y con $N = n^d$ evaluaciones el error pasa a ser $\mathcal{O}(N^{-4/d})$, que a partir de $d = 8$ ya es peor que $N^{-1/2}$. En dimensión 100 ninguna cuadratura clásica es siquiera imaginable: una malla con solo dos puntos por eje exigiría 2^{100} evaluaciones. Monte Carlo sigue necesitando el mismo millón de puntos que en dimensión dos. Es un mal método que resulta ser el único posible, y por eso lo usan la física de partículas, las finanzas y la renderización de imágenes.

Regla: Estimar una cantidad con Monte Carlo

El esquema es siempre el mismo, y la parte difícil no es programarlo sino saber cuánto fiarse del resultado:

1. **Escribe la cantidad buscada como una media** de algo que se pueda evaluar en un punto al azar; casi todo lo es.
2. **Elige una región sencilla que contenga al dominio**, típicamente un cubo, para poder generar puntos uniformes sin esfuerzo.
3. **Genera N puntos con un generador de calidad** y evalúa la función en cada uno; no reutilices puntos entre experimentos que quieras comparar.
4. **Promedia y multiplica por el volumen** de la región de muestreo.
5. **Estima el error** con $\sigma/\sqrt{N}$, calculando σ a partir de la propia muestra; sin esta barra de error el número obtenido no significa nada.
6. **Recuerda la regla de las cuatro veces**: para ganar un decimal hay que multiplicar N por cien, porque el error va con $\sqrt{N}$.
7. **Fija la semilla del generador** si quieres poder reproducir el cálculo; un resultado que no se puede repetir no se puede depurar.

Ejemplo 14.48 (Estimar π y una integral)**. (a) Puntos en un cuadrado.** Generamos puntos uniformes en $[0,1]^2$ y contamos cuántos caen dentro del cuarto de círculo $x^2 + y^2 \le 1$. La fracción estima $\pi/4$, así que multiplicamos por cuatro. Los resultados de una simulación fueron: 3,16 con $N = 100$, 3,16 con $N = 1000$, 3,1408 con $N = 10\,000$, 3,1396 con $N = 100\,000$ y 3,141768 con $N = 1\,000\,000$.

(b) La lentitud, cuantificada. Los errores fueron 0,018407, 0,018407, 0,000793, 0,001993 y 0,000175. Nótese que *no* decrecen monótonamente: el error es aleatorio y solo su *tamaño típico* decrece, como $1/\sqrt{N}$. Para pasar de dos decimales correctos a cuatro hay que multiplicar N por diez mil. Cualquier método determinista humilla a Monte Carlo en este problema.

(c) Una integral en una dimensión. Para $\int_0^1 \mathrm{e}^{-x^2}\mathrm{d}x = 0{,}746824$, Monte Carlo con $N = 1000$ dio 0,737310 (error 0,009514) y con $N = 100\,000$ dio 0,746141 (error 0,000683). La regla del trapecio con solo diez subintervalos da 0,7462108, con error $6{,}13 \cdot 10^{-4}$, y con cien subintervalos baja a $6{,}13 \cdot 10^{-6}$. Cien evaluaciones deterministas baten a cien mil aleatorias.

(d) La barra de error, que es lo importante. La desviación típica de e^{-x^2} sobre $[0,1]$ vale 0,201004, luego el error típico con $N = 10\,000$ es $\frac{0{,}201004}{100} = 0{,}002010$. Ese número,

calculable de la propia muestra, es lo que convierte a Monte Carlo en un método serio: no solo da una estimación, sino que dice cuánto vale esa estimación. Ninguna cuadratura clásica ofrece eso sin trabajo adicional.

Ejemplo 14.49 (Donde Monte Carlo gana: dimensiones altas)**. (a) El volumen de la bola unidad.** En dimensión d vale $V_d = \dfrac{\pi^{d/2}}{\Gamma\left(\frac{d}{2}+1\right)}$, que da 3,141593 para $d=2$, 4,188790 para $d=3$, 5,263789 para $d=5$ y 2,550164 para $d=10$. Estimándolo con 200 000 puntos en el cubo $[-1,1]^d$ se obtuvo 3,136380, 4,177560, 5,295040 y 2,375680: aceptable hasta $d=5$ y ya flojo en $d=10$.

(b) Por qué se degrada. La fracción del cubo ocupada por la bola es $V_d/2^d$: vale 0,785398 en dimensión 2, 0,523599 en dimensión 3, 0,164493 en dimensión 5 y apenas 0,002490 en dimensión 10. De doscientos mil puntos, en dimensión diez solo unos quinientos caen dentro, y con quinientos aciertos no se puede tener mucha precisión. El fenómeno tiene nombre: en dimensión alta, casi todo el volumen de un cubo está en sus esquinas.

(c) La comparación decisiva. Aun así, una cuadratura en dimensión 10 con solo diez puntos por eje exigiría 10^{10} evaluaciones para un resultado mediocre, mientras que Monte Carlo con 10^6 da una estimación con barra de error honesta. Y en dimensión 100, que es lo normal al valorar una cartera de derivados o al simular un detector de partículas, la cuadratura sencillamente no existe. La regla práctica es: hasta dimensión cuatro o cinco, cuadratura; a partir de ahí, Monte Carlo.

(d) La aguja de Buffon. El experimento más antiguo del género: se lanza una aguja de longitud L sobre un suelo con líneas paralelas separadas $d \geq L$, y la probabilidad de que corte una línea es $\frac{2L}{\pi d}$. Con $L=d$ eso vale 0,636620. Simulando medio millón de lanzamientos se obtuvo una frecuencia de 0,635736, de donde $\pi \approx 2/0{,}635736 = 3{,}1460$. Es una estimación pésima de π, pero como demostración de que un experimento aleatorio puede calcular una constante matemática resulta insuperable.

Cuidado

Duplicar N no reduce el error a la mitad, solo lo divide entre $\sqrt{2} \approx 1{,}41$; para ganar un decimal hacen falta cien veces más puntos. Segundo aviso: un resultado de Monte Carlo sin barra de error es inútil, porque el propio resultado es una variable aleatoria y podría estar mucho más lejos de lo que su aspecto sugiere. Y tercero: los generadores pseudoaleatorios tienen período y estructura; los antiguos generadores lineales congruenciales colocan los puntos sobre un número reducido de hiperplanos, lo que arruinó bastantes simulaciones de los años setenta. Conviene usar el generador que ofrezca la biblioteca moderna del lenguaje y no improvisar uno.

Pausa para pensar

El nombre lo puso Nicholas Metropolis en 1946, en Los Álamos, por el casino de Mónaco donde jugaba un tío de Stanisław Ulam. Ulam había tenido la idea mientras convalecía

de una enfermedad, jugando al solitario: se preguntó qué probabilidad tenía de ganar y, tras intentar contarlo combinatoriamente, cayó en que era mucho más fácil jugar cien partidas y contar. Von Neumann vio al instante que la misma idea servía para la difusión de neutrones y la programó en el ENIAC. Es uno de los pocos casos en que un método numérico de primera fila nace literalmente de un juego de cartas, y la única familia de métodos donde el error se estima con estadística en vez de con análisis; su fundamento teórico es el del capítulo 16.

14.18. Errores típicos

1. Aplicar bisección sin comprobar que hay cambio de signo. Sin él, no hay garantía de nada.
2. Creer que un cambio de signo garantiza una raíz aunque la función sea discontinua. Bolzano exige continuidad; $1/x$ es el contraejemplo.
3. Usar bisección para raíces dobles. No hay cambio de signo, y el método no arranca.
4. Aplicar Newton sin mirar si f' se anula cerca del punto de partida.
5. Suponer que Newton siempre converge. Puede divergir o ciclar; conviene acotar el número de iteraciones.
6. Iterar con la calculadora en grados cuando la fórmula es de radianes. Con $\cos x = x$ en grados el punto fijo es otro por completo.
7. Escribir la ecuación como $x = g(x)$ sin comprobar $|g'| < 1$. Hay despejes que convergen y despejes que se escapan.
8. Parar cuando $|f(x_n)|$ es pequeño en vez de cuando lo es $|x_{n+1} - x_n|$. Con funciones muy planas, f puede ser diminuta lejos de la raíz.
9. Dar más cifras de las que el método garantiza. Escribir diez decimales tras tres iteraciones de bisección es mentir con precisión.
10. Usar Simpson con un número impar de subintervalos. Necesita n par, porque agrupa los puntos de tres en tres.
11. Olvidar dividir por 2 los extremos en la regla de los trapecios.
12. Creer que aumentar n indefinidamente mejora siempre el resultado. El redondeo acaba dominando.
13. Restar cantidades casi iguales sin reescribir la expresión. Es la vía más rápida a un resultado sin ninguna cifra correcta.
14. Confundir interpolar con extrapolar y fiarse de un valor fuera del rango de los datos.

14.19. Integración de altas prestaciones, y el suelo que impone el punto flotante

Teorema 14.50 (Errores de las fórmulas elementales)**.** *Para n subintervalos de longitud $h = (b-a)/n$,*

$$E_{trapecios} = -\frac{(b-a)h^2}{12} f''(\xi), \qquad E_{Simpson} = -\frac{(b-a)h^4}{180} f^{(4)}(\xi).$$

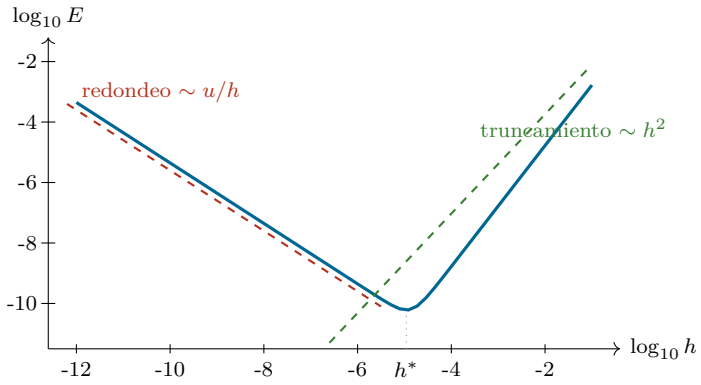


Figura 14.2: El error total de la diferencia centrada es la suma de dos efectos que tiran en sentidos opuestos: el de truncamiento baja como h^2 y el de redondeo sube como u/h, porque restar números casi iguales pierde cifras. El mínimo está en $h^* \approx u^{1/3} \approx 6 \times 10^{-6}$, y por debajo de ahí afinar el paso *empeora* el resultado. En doble precisión no se pueden obtener más de unas once cifras.

Estas fórmulas permiten decidir *antes de calcular* cuántos subintervalos hacen falta, cosa que la tabla empírica no permite.

Definición 14.51 (Extrapolación de Romberg)**.** Sea $T(h)$ la regla del trapecio con paso h. Como su error se desarrolla en potencias *pares* de h, la combinación

$$R_{j,k} = \frac{4^k R_{j,k-1} - R_{j-1,k-1}}{4^k - 1}$$

cancela sucesivamente los términos $h^2, h^4, h^6, \ldots$ Cada columna nueva gana dos órdenes.

Ejemplo 14.52 ($\int_1^2 \mathrm{d}x/x = \ln 2$, con la tabla completa)**.**

h	trapecios	col. 1	col. 2	col. 3	col. 4
1	0,750 000 0000				
1/2	0,708 333 3333	0,694 444 4444			
1/4	0,697 023 8095	0,693 253 9683	0,693 174 6032		
1/8	0,694 121 8504	0,693 154 5307	0,693 147 9015	0,693 147 4776	
1/16	0,693 391 2022	0,693 147 6528	0,693 147 1943	0,693 147 1831	0,693 147 1819

El valor exacto es $\ln 2 = 0{,}693\,147\,180\,6$. La primera columna, con dieciséis subintervalos, acierta cuatro cifras; la última, con las *mismas* evaluaciones, acierta nueve. El error pasa de $5{,}7 \times 10^{-2}$ (primer trapecio) a $1{,}4 \times 10^{-9}$ sin evaluar la función ni una vez más.

Obsérvese además que la columna 1 *es* la regla de Simpson: extrapolar trapecios una vez da Simpson, y por eso Simpson es de orden 4 y no de orden 2.

Definición 14.53 (Cuadratura de Gauss–Legendre)**.** En lugar de fijar los nodos y calcular los pesos, se optimizan ambos. Con n nodos se consigue integrar exactamente todos los polinomios de grado $\le 2n-1$, el doble de lo que permitiría cualquier fórmula de nodos prefijados. Los nodos son las raíces del polinomio de Legendre P_n; para $n=2$, $x = \pm 1/\sqrt{3}$ con pesos 1; para $n=3$, $x = 0, \pm\sqrt{3/5}$ con pesos $8/9$ y $5/9$.

Ejemplo 14.54 (Dos y tres evaluaciones)**.** Sobre el mismo $\int_1^2 \mathrm{d}x/x$, trasladando el intervalo a $[-1,1]$:

$$\text{Gauss con 2 nodos} = 0{,}692\,307\,6923 \quad (\text{error } 8{,}4 \times 10^{-4}),$$

$$\text{Gauss con 3 nodos} = 0{,}693\,121\,6931 \quad (\text{error } 2{,}5 \times 10^{-5}).$$

Compárese: Simpson con ocho subintervalos —nueve evaluaciones— comete $7{,}4 \times 10^{-6}$. Gauss con tres evaluaciones ya está en el mismo orden de magnitud, y con cinco lo supera holgadamente. Cuando evaluar f es caro (una simulación, un experimento), esa diferencia lo es todo.

El suelo: aritmética de punto flotante

Definición 14.55 (IEEE 754 doble precisión)**.** Un número se guarda como signo (1 bit), exponente (11 bits) y mantisa (52 bits). Todo real representable cumple

$$\mathrm{fl}(x) = x(1+\delta), \qquad |\delta| \le u = 2^{-53} \approx 1{,}11 \times 10^{-16},$$

y el **épsilon de máquina** —el menor ε con $1 + \varepsilon \neq 1$— es $2^{-52} = 2{,}220\,446 \times 10^{-16}$.

Ejemplo 14.56 (Por qué $0{,}1 + 0{,}2 \neq 0{,}3$)**.** En binario, 0,1 es periódico ($0{,}0\overline{0011}$), igual que $1/3$ lo es en decimal. Lo que la máquina guarda es

$$0{,}1 \;\longrightarrow\; 0{,}100\,000\,000\,000\,000\,005\,55\ldots$$

Al sumarlo con la representación de 0,2, el resultado es $0{,}300\,000\,000\,000\,000\,04$, que es un número de máquina distinto del más cercano a 0,3. La comparación con igualdad falla. La regla de oro es no comparar nunca reales con `==`, sino con una tolerancia.

Observación 14.57 (El h óptimo, deducido y no observado)*.* En la diferencia centrada, el error de truncamiento es $O(h^2)$ y el de redondeo es $O(u/h)$, porque al restar dos números casi iguales se pierden cifras. El total,

$$E(h) \approx Ch^2 + \frac{2u}{h},$$

se minimiza en $h^* \approx u^{1/3}$, es decir

$$h^* \approx (2{,}22 \times 10^{-16})^{1/3} = 6{,}06 \times 10^{-6},$$

con un error mínimo del orden de $u^{2/3} \approx 4 \times 10^{-11}$. Ese es el suelo: por mucho que se afine, la derivada numérica no puede dar más de once cifras en doble precisión, y afinar h por debajo de 10^{-6} *empeora* el resultado. La tabla empírica del capítulo lo mostraba; ahora se predice.

Observación 14.58 (Orden de convergencia y raíces múltiples). Un método converge con **orden** p si $|e_{k+1}|/|e_k|^p \to C$ con $0 < C < \infty$: bisección es lineal con $C = 1/2$, la secante tiene $p = \varphi = 1{,}618$, y Newton $p = 2$. Pero eso presupone una raíz simple.

Con $f(x) = (x-1)^3$, que tiene una raíz triple, Newton desde $x_0 = 2$ da errores

$$1, \quad 0{,}666\,667, \quad 0{,}444\,444, \quad 0{,}296\,296, \quad 0{,}197\,531, \ \ldots$$

El cociente es exactamente $2/3$ en cada paso: convergencia *lineal* con razón $1 - 1/m$, donde $m = 3$ es la multiplicidad. Newton pierde su virtud principal.

El remedio es conocer m y usar $x_{k+1} = x_k - m\, f(x_k)/f'(x_k)$: aquí eso da $x_1 = 1$ exactamente, en una sola iteración desde $x_0 = 2$. Si m no se conoce, se aplica Newton a $u = f/f'$, que tiene una raíz simple donde f la tiene múltiple.

14.20. Splines cúbicos, y Newton en varias variables

El fenómeno de Runge dejó una pregunta abierta: si subir el grado empeora las cosas, ¿qué se hace? La respuesta que usa todo el software del mundo es la contraria: bajar el grado y subir el número de trozos.

Definición 14.59 (Spline cúbico)**.** Dados $n+1$ nodos, un spline cúbico es una función que en cada uno de los n subintervalos es un polinomio de grado ≤ 3, y que es globalmente de clase C^2.

El recuento de grados de libertad explica la construcción: hay $4n$ coeficientes; la interpolación impone $2n$ condiciones (cada trozo pasa por sus dos extremos) y la continuidad de f' y f'' en los $n-1$ nodos interiores impone $2(n-1)$ más. Total, $4n-2$: faltan dos condiciones, y de ahí las dos variantes clásicas. El spline **natural** pone $f'' = 0$ en los extremos; el **sujeto** fija ahí la derivada primera.

Ejemplo 14.60 (Cuando el spline no gana, y cuando gana por goleada)**. (a) Cuatro puntos de $\sqrt{x}$.** Con los nodos $(1,1)$, $(4,2)$, $(9,3)$, $(16,4)$ y estimando $\sqrt{7}$:

spline natural: 2,674 811, polinomio de grado 3: 2,671 429, exacto: 2,645 751.

Los errores son 0,029 1 y 0,025 7: el polinomio gana *por poco*. Con cuatro nodos sobre una función suave no hay ninguna razón para preferir el spline, y conviene decirlo.

(b) La función de Runge. Ahora $f(x) = 1/(1+25x^2)$ en $[-1,1]$ con nodos equiespaciados. Error máximo:

nodos	polinomio interpolante	spline cúbico natural
11	1,915 6	0,021 97
21	59,82	0,003 18

Con once nodos el polinomio ya comete un error mayor que el propio rango de la función; con veintiuno, el error se multiplica por treinta. El spline, en cambio, mejora un factor siete al doblar los nodos, tal como predice su orden $O(h^4)$.

Esa es la lección completa: la interpolación polinómica global es inestable, y la solución no es más grado sino más trozos. Con nodos de Chebyshev el polinomio se recupera, pero eso exige poder *elegir* dónde medir, y en la práctica los datos vienen donde vienen.

Newton en varias variables

Definición 14.61. Para $F : \mathbb{R}^n \to \mathbb{R}^n$, la iteración de Newton es

$$\mathbf{x}_{k+1} = \mathbf{x}_k - J(\mathbf{x}_k)^{-1} F(\mathbf{x}_k),$$

que en la práctica *nunca* se implementa invirtiendo: se resuelve el sistema lineal

$$J(\mathbf{x}_k)\,\Delta_k = -F(\mathbf{x}_k), \qquad \mathbf{x}_{k+1} = \mathbf{x}_k + \Delta_k,$$

con las técnicas de la sección de sistemas lineales.

Ejemplo 14.62 (Intersección de una circunferencia y una exponencial)**.** Resolvamos

$$\begin{cases} x^2 + y^2 = 4, \\ y = e^x - 1, \end{cases} \qquad J(x,y) = \begin{pmatrix} 2x & 2y \\ -e^x & 1 \end{pmatrix},$$

desde $\mathbf{x}_0 = (1,1)$.

k	x_k	y_k	$\|F(\mathbf{x}_k)\|$	error
1	1,075 766	1,924 234	$8{,}6 \times 10^{-1}$	$2{,}1 \times 10^{-1}$
2	1,009 471	1,737 845	$4{,}0 \times 10^{-2}$	$9{,}8 \times 10^{-3}$
3	1,004 189	1,729 653	$1{,}0 \times 10^{-4}$	$2{,}6 \times 10^{-5}$
4	1,004 169	1,729 637	$8{,}4 \times 10^{-10}$	$2{,}3 \times 10^{-10}$

La columna del error se lee sola: 2×10^{-1}, 10^{-2}, 10^{-5}, 10^{-10}. El número de cifras correctas se *dobla* en cada paso, exactamente como en una variable. La convergencia cuadrática sobrevive intacta al pasar a $\mathbb{R}^n$, siempre que la raíz sea simple, es decir, que J sea no singular ahí.

Cuidado

Los tres problemas de Newton en una variable se agravan en varias. Primero, hay que *arrancar cerca*: la cuenca de atracción puede ser un fractal, y este mismo sistema desde $(-1,-1)$ converge a la otra raíz. Segundo, cada paso cuesta evaluar n^2 derivadas y resolver un sistema $n \times n$: de ahí los métodos *cuasi-Newton* (Broyden, BFGS), que actualizan una aproximación de J en lugar de recalcularla. Y tercero, si J es casi singular el paso Δ_k se dispara; el remedio estándar es amortiguar, $\mathbf{x}_{k+1} = \mathbf{x}_k + \lambda_k \Delta_k$ con $\lambda_k \le 1$ elegido para que $\|F\|$ baje de verdad, o pasar a una región de confianza como en Levenberg–Marquardt.

14.21. Los teoremas de error: Newton, interpolación y cuadratura

El capítulo ha usado tres cotas de error como recetas. Las tres son teoremas con demostración corta, y sin ellas no se puede saber cuándo un método es aplicable.

Teorema 14.63 (Convergencia cuadrática de Newton)**.** *Sea f de clase C^2 en un entorno de una raíz p con $f(p) = 0$ y $f'(p) \neq 0$. Entonces existe $\delta > 0$ tal que, si $x_0 \in [p-\delta, p+\delta]$, la iteración $x_{k+1} = x_k - \frac{f(x_k)}{f'(x_k)}$ está bien definida, converge a p, y los errores $e_k = x_k - p$ cumplen*

$$e_{k+1} = \frac{f''(\xi_k)}{2f'(x_k)}\, e_k^2 \qquad \text{para algún } \xi_k \text{ entre } x_k \text{ y } p.$$

En particular $\dfrac{|e_{k+1}|}{e_k^2} \longrightarrow \left|\dfrac{f''(p)}{2f'(p)}\right|$.

Demostración. Desarróllese f por Taylor alrededor de x_k evaluando en p, con resto de Lagrange:

$$0 = f(p) = f(x_k) + f'(x_k)(p - x_k) + \tfrac{1}{2} f''(\xi_k)(p - x_k)^2.$$

Divídase entre $f'(x_k)$, que no se anula en un entorno de p por continuidad:

$$0 = \frac{f(x_k)}{f'(x_k)} + (p - x_k) + \frac{f''(\xi_k)}{2f'(x_k)} e_k^2.$$

Sumando $x_k - p = e_k$ a ambos lados y reconociendo $x_k - \frac{f(x_k)}{f'(x_k)} = x_{k+1}$:

$$e_{k+1} = x_{k+1} - p = \frac{f''(\xi_k)}{2f'(x_k)}\, e_k^2.$$

Para la convergencia, tómense δ y M tales que en $[p-\delta, p+\delta]$ se cumpla $|f''| \le 2M_2$, $|f'| \ge m_1 > 0$ y $C\delta \le \frac{1}{2}$, con $C = M_2/m_1$. Entonces $|e_{k+1}| \le C|e_k|^2 \le \frac{1}{2}|e_k|$, luego los iterados no salen del intervalo y $e_k \to 0$; y de $C|e_{k+1}| \le (C|e_k|)^2$ se obtiene $C|e_k| \le (C|e_0|)^{2^k}$: el número de cifras correctas se duplica en cada paso. □

Cuidado

[Las dos hipótesis son las que fallan en la práctica] La condición $f'(p) \neq 0$ excluye las raíces múltiples. Si p es raíz doble, Newton converge sólo *linealmente*, con razón $1/2$: para $f(x) = (x-1)^2$ y $x_0 = 2$, los errores son $1,\ 0{,}5,\ 0{,}25,\ 0{,}125, \ldots$ El remedio es la variante de Schröder $x_{k+1} = x_k - m\frac{f(x_k)}{f'(x_k)}$ con m la multiplicidad, que restaura el orden dos. Y la condición de partir *cerca* de la raíz no es negociable: para $f(x) = \arctan x$ y $x_0 = 2$ la iteración diverge oscilando, con $|x_k| \to \infty$. La convergencia de Newton es local, y ese es su único defecto serio.

Teorema 14.64 (Error de interpolación)**.** *Sea f de clase C^{n+1} en $[a,b]$ y P el polinomio de grado $\leq n$ que interpola a f en los nodos distintos $x_0, \ldots, x_n \in [a,b]$. Entonces, para cada $x \in [a,b]$, existe $\xi = \xi(x) \in (a,b)$ con*

$$f(x) - P(x) = \frac{f^{(n+1)}(\xi)}{(n+1)!} \prod_{i=0}^{n} (x - x_i).$$

Demostración. Si x coincide con algún nodo, ambos lados son 0. Sea pues x distinto de todos los nodos y defínase $\omega(t) = \prod_i (t - x_i)$ y la constante $K = \dfrac{f(x) - P(x)}{\omega(x)}$, de modo que la función auxiliar

$$g(t) = f(t) - P(t) - K\,\omega(t)$$

se anula en los $n+1$ nodos y en $t = x$: en total $n+2$ ceros distintos en $[a,b]$. Por el teorema de Rolle, g' tiene al menos $n+1$ ceros; g'', al menos n; y, iterando, $g^{(n+1)}$ tiene al menos un cero ξ. Ahora bien,

$$g^{(n+1)}(t) = f^{(n+1)}(t) - 0 - K\,(n+1)!,$$

porque P tiene grado $\leq n$ y ω es mónico de grado $n+1$. Evaluando en ξ y despejando, $K = \frac{f^{(n+1)}(\xi)}{(n+1)!}$, que es la fórmula del enunciado. □

Observación 14.65 (De aquí sale todo lo demás). La fórmula explica las tres cosas que el capítulo daba por sabidas. *Primera:* por qué interpolar con muchos nodos equiespaciados es mala idea —el fenómeno de Runge—: el factor $f^{(n+1)}$ puede crecer más deprisa que $(n+1)!$, y $\omega(x)$ es enorme cerca de los extremos. *Segunda:* por qué los nodos de Chebyshev son óptimos: minimizan $\max_{[a,b]} |\omega(x)|$, cuyo valor mínimo posible es $2\left(\frac{b-a}{4}\right)^{n+1}$. Y *tercera:* las cotas de las fórmulas de cuadratura, que se obtienen integrando la fórmula de error.

Teorema 14.66 (Error de las fórmulas de cuadratura)**.** *Para f suficientemente regular en*

$[a,b]$ *y* h *el paso de una malla uniforme:*

$$\begin{aligned}
&\textit{Trapecio simple} && \int_a^b f - \frac{b-a}{2}\big(f(a)+f(b)\big) = -\frac{(b-a)^3}{12} f''(\xi)\\
&\textit{Simpson simple} && \int_a^b f - \frac{b-a}{6}\Big(f(a)+4f\big(\tfrac{a+b}{2}\big)+f(b)\Big) = -\frac{(b-a)^5}{2880} f^{(4)}(\xi)\\
&\textit{Trapecio compuesto} && E = -\frac{(b-a)h^2}{12} f''(\xi) = O(h^2)\\
&\textit{Simpson compuesto} && E = -\frac{(b-a)h^4}{180} f^{(4)}(\xi) = O(h^4)
\end{aligned}$$

Demostración para el trapecio. Con $n=1$ y nodos a, b, el teorema 14.64 da $f(x) - P_1(x) = \frac{f''(\xi(x))}{2}(x-a)(x-b)$. El polinomio $(x-a)(x-b)$ no cambia de signo en $[a,b]$ (es ≤ 0), de modo que el teorema del valor medio para integrales permite sacar f'' de la integral en un punto:

$$\int_a^b (f - P_1) = \frac{f''(\xi)}{2}\int_a^b (x-a)(x-b)\,\mathrm{d}x = \frac{f''(\xi)}{2}\cdot\Big(-\frac{(b-a)^3}{6}\Big) = -\frac{(b-a)^3}{12} f''(\xi),$$

y $\int_a^b P_1$ es exactamente la regla del trapecio. El caso compuesto se obtiene sumando sobre los $N = (b-a)/h$ subintervalos y aplicando el teorema del valor intermedio a la media de los $f''(\xi_i)$. □

Observación 14.67 (Por qué Simpson gana dos órdenes y no uno). Simpson se construye interpolando con un polinomio de grado 2, de modo que cabría esperar un error $O(h^3)$. Sin embargo integra exactamente los polinomios de grado 3: por simetría, el término cúbico $(x - \frac{a+b}{2})^3$ tiene integral nula sobre el intervalo y también la regla le asigna 0. Ese grado extra de exactitud gratuito es lo que eleva el orden a $O(h^4)$. La misma idea, llevada al extremo eligiendo también los *nodos*, produce la cuadratura de Gauss–Legendre: con n nodos se integran exactamente los polinomios de grado $2n-1$, porque los nodos se toman como raíces del polinomio de Legendre P_n, que es ortogonal a todos los de grado $< n$; el resto de la división de f por P_n se integra exactamente y el cociente se anula contra P_n por ortogonalidad.

Definición 14.68 (Región de estabilidad absoluta)**.** Aplíquese un método de un paso a la ecuación de prueba $y' = \lambda y$ con $\lambda \in \mathbb{C}$. Todo método explícito produce una recurrencia $y_{n+1} = R(h\lambda)\, y_n$, donde R es la *función de estabilidad*. La *región de estabilidad absoluta* es

$$\mathcal{R} = \{z \in \mathbb{C} : \ |R(z)| \le 1\},$$

el conjunto de valores de $z = h\lambda$ para los que la solución numérica no crece.

Ejemplo 14.69 (Las regiones de los métodos usuales, calculadas)**.** Para Euler explícito, $y_{n+1} = y_n + h\lambda y_n$, luego $R(z) = 1+z$ y $\mathcal{R}$ es el disco $|1+z| \le 1$: un círculo de radio 1 centrado en -1. Sobre el eje real negativo llega hasta $z = -2$.

Para Runge–Kutta de orden 4 aplicado a la ecuación de prueba, las cuatro etapas dan

$$R(z) = 1 + z + \frac{z^2}{2} + \frac{z^3}{6} + \frac{z^4}{24},$$

que es el desarrollo de e^z truncado en grado 4. Resolviendo $|R(x)| = 1$ sobre el eje real negativo se obtiene $x = -2{,}785\,3$: ese es el famoso número, y ahora se puede comprobar. La consecuencia práctica: para $y' = -100y$, RK4 exige $h < 2{,}785/100 = 0{,}0279$, y con $h = 0{,}03$ la solución numérica explota aunque la exacta decaiga.

Para Euler *implícito*, $y_{n+1} = y_n + h\lambda y_{n+1}$ da $R(z) = \frac{1}{1-z}$, y $|R(z)| \le 1$ para todo z con parte real ≤ 0: el método es *A-estable*, sin restricción de paso. Ese es el motivo por el que los problemas rígidos —los que tienen escalas de tiempo muy dispares— se integran con métodos implícitos.

Teorema 14.70 (Error local y error global)**.** *Sea un método de un paso $y_{n+1} = y_n + h\Phi(t_n, y_n, h)$ con error local de truncamiento $O(h^{p+1})$ y con Φ lipschitziana en y con constante L. Entonces el error global en un intervalo fijo $[t_0, T]$ es $O(h^p)$; concretamente*

$$|y(t_n) - y_n| \;\le\; \frac{Ch^p}{L}\Big(\mathrm{e}^{L(t_n - t_0)} - 1\Big).$$

Demostración. Sea $e_n = y(t_n) - y_n$ y τ_n el error local, con $|\tau_n| \le Ch^{p+1}$. Restando la recurrencia numérica de la exacta,

$$|e_{n+1}| \le |e_n| + h\,L\,|e_n| + Ch^{p+1} = (1 + hL)|e_n| + Ch^{p+1}.$$

Esta recurrencia lineal, con $e_0 = 0$, se resuelve:

$$|e_n| \le Ch^{p+1} \sum_{j=0}^{n-1} (1 + hL)^j = Ch^{p+1}\,\frac{(1+hL)^n - 1}{hL} \le \frac{Ch^p}{L}(\mathrm{e}^{nhL} - 1),$$

usando $1 + hL \le \mathrm{e}^{hL}$ y $nh = t_n - t_0$. □

Observación 14.71 (Se pierde exactamente un orden)*.* El error local es $O(h^{p+1})$, se acumula sobre $n = (T - t_0)/h$ pasos, y $h^{p+1} \cdot h^{-1} = h^p$. Esa es toda la contabilidad, y explica la nomenclatura: RK4 tiene error local $O(h^5)$ y se llama "de orden 4" porque su error global es $O(h^4)$. El factor $\mathrm{e}^{L(T-t_0)}$ es el precio de la propagación: en problemas caóticos, con L grande, la cota se vuelve inútil mucho antes de que h pueda hacerse suficientemente pequeño, y ahí está el límite de la predicción numérica a largo plazo.

14.22. La propagación del error de redondeo: el modelo estándar y la suma compensada

El apartado anterior mostró *qué* puede salir mal y dio recetas para esquivarlo. Falta el marco que permite predecirlo en lugar de descubrirlo: una hipótesis sobre lo que hace la máquina en cada operación aislada, y dos teoremas que dicen cuánto error se acumula al encadenar muchas. Con eso, la afirmación «sumar de menor a mayor» deja de ser un consejo y pasa a ser un corolario.

Definición 14.72 (Modelo estándar de la aritmética de punto flotante)**.** Sea u la unidad de redondeo del formato ($u = 2^{-53} \approx 1{,}1102 \cdot 10^{-16}$ en doble precisión). Se dice que una

aritmética cumple el **modelo estándar** si, para cualesquiera números de máquina a y b y cualquier operación $* \in \{+, -, \times, /\}$ cuyo resultado exacto esté dentro del rango representable,

$$\mathrm{fl}(a * b) = (a * b)(1 + \delta), \qquad |\delta| \le u.$$

Es decir: cada operación aislada devuelve el resultado exacto redondeado, y por tanto comete un error *relativo* no mayor que u.

Observación 14.73. El estándar IEEE 754 garantiza el modelo, y lo garantiza también para la raíz cuadrada. Lo que *no* garantiza es nada sobre una cadena de operaciones: el modelo se aplica a cada paso por separado, y el problema del análisis numérico es precisamente componer esas cotas. Obsérvese además que el modelo no dice que la aritmética sea asociativa —no lo es— ni distributiva —tampoco—; solo dice que cada paso individual es lo mejor posible.

Teorema 14.74 (Error de una suma de n términos)**.** *Sean $x_1, \dots, x_n$ números de máquina y sea $\hat{s}$ el resultado de sumarlos en el orden dado, es decir $\hat{s}_1 = x_1$ y $\hat{s}_k = \mathrm{fl}(\hat{s}_{k-1} + x_k)$. Si $(n-1)u < 1$, entonces*

$$\Big|\hat{s}_n - \sum_{i=1}^{n} x_i\Big| \;\le\; \gamma_{n-1} \sum_{i=1}^{n} |x_i|, \qquad \textit{donde} \quad \gamma_k = \frac{ku}{1-ku} \approx ku.$$

Demostración. Por el modelo estándar, $\hat{s}_k = (\hat{s}_{k-1} + x_k)(1+\delta_k)$ con $|\delta_k| \le u$, para $k = 2, \dots, n$. Desarrollando la recurrencia,

$$\hat{s}_n = x_1 \prod_{k=2}^{n}(1+\delta_k) + x_2 \prod_{k=2}^{n}(1+\delta_k) + x_3 \prod_{k=3}^{n}(1+\delta_k) + \cdots + x_n(1+\delta_n).$$

Cada x_i aparece multiplicado por un producto de a lo sumo $n-1$ factores $(1+\delta)$. Ahora bien, si $|\delta_j| \le u$ para $j = 1, \dots, k$ y $ku < 1$, entonces

$$\Big|\prod_{j=1}^{k}(1+\delta_j) - 1\Big| \le (1+u)^k - 1 \le \frac{ku}{1-ku} = \gamma_k,$$

donde la última desigualdad se comprueba con el desarrollo del binomio, acotando $\binom{k}{j}u^j \le (ku)^j$ y sumando la serie geométrica. Por tanto cada producto se escribe como $1+\theta_i$ con $|\theta_i| \le \gamma_{n-1}$, y

$$\hat{s}_n - \sum_i x_i = \sum_i x_i\theta_i, \qquad \Big|\hat{s}_n - \sum_i x_i\Big| \le \gamma_{n-1} \sum_i |x_i|.$$

□

Corolario 14.75. *La cota del teorema 14.74 es pesimista, pero el examen de la demostración da una cota más fina: el sumando x_i aparece afectado por $n-i+1$ redondeos si $i \ge 2$, y por $n-1$ si $i = 1$. Conviene, pues, que los sumandos que soporten más redondeos sean los más pequeños, es decir,* sumar de menor a mayor. *Si todos los x_i son del mismo signo, la cota relativa es γ_{n-1}, porque entonces $\sum |x_i| = |\sum x_i|$: sumar términos de igual signo es una operación bien condicionada.*

Ejemplo 14.76 (La serie armónica en precisión simple)**.** Se calcula $\sum_{k=1}^{10^6} 1/k$ con $u = 2^{-24} = 5{,}96 \cdot 10^{-8}$. El valor correcto, obtenido con precisión suficiente, es 14,392 726 72.

orden de suma	resultado	error relativo
de $k = 1$ a $k = 10^6$	14,357 358	$2{,}5 \cdot 10^{-3}$
de $k = 10^6$ a $k = 1$	14,392 652	$5{,}2 \cdot 10^{-6}$
suma compensada (adelante)	14,392 727	$1{,}2 \cdot 10^{-8}$

La primera fila pierde casi tres cifras y la segunda casi ninguna: es el corolario 14.75 en acción. La cota del teorema con $n = 10^6$ da $\gamma_{n-1} \approx 0{,}06$, es decir, algo más de un 6 %; el error real es de 2,5 por mil, unas veinticinco veces menor, porque los errores individuales tienen signos que se compensan en parte. La cota es una garantía, no una predicción.

Regla: Predecir cómo se propaga el error en las cuatro operaciones

Escriba cada dato como $\tilde{a} = a(1 + \varepsilon_a)$, con ε_a su error relativo. Despreciando términos de segundo orden:

1. **Producto y cociente: los errores relativos se suman.** $\varepsilon_{ab} \approx \varepsilon_a + \varepsilon_b$ y $\varepsilon_{a/b} \approx \varepsilon_a - \varepsilon_b$. Multiplicar y dividir son inofensivos: nunca amplifican más allá de un factor 2.
2. **Suma de términos del mismo signo: inofensiva también.** $\varepsilon_{a+b} = \dfrac{a\varepsilon_a + b\varepsilon_b}{a+b}$, y si $a, b > 0$ ese cociente es una media ponderada, luego $|\varepsilon_{a+b}| \le \max(|\varepsilon_a|, |\varepsilon_b|)$.
3. **Resta de términos parecidos: aquí está todo el peligro.** Con la misma fórmula, si $a \approx b$ el denominador $a - b$ es diminuto y el error relativo se multiplica por
$$\kappa_- = \frac{|a| + |b|}{|a - b|},$$
que puede ser tan grande como se quiera. Ese número es el factor de amplificación exacto y conviene calcularlo antes de programar nada.
4. **Funciones: multiplique por la derivada logarítmica.** $\varepsilon_{f(a)} \approx \dfrac{a f'(a)}{f(a)}\, \varepsilon_a$. Para $f = \sqrt{\ }$ el factor es $1/2$ (inofensiva); para $f = \ln$ cerca de 1 el factor es $1/\ln a$, que explota.
5. **Encadene multiplicando los factores.** El error relativo final es la suma de las contribuciones de cada dato multiplicadas por los factores de amplificación de todos los pasos posteriores.
6. **Añada u por cada operación efectuada**, que es el error que introduce la propia máquina además del que traían los datos.
7. **Verifique** recalculando en precisión doble o con una reescritura equivalente. Si la discrepancia supera lo previsto, hay una resta escondida que no se ha contabilizado.

Ejemplo 14.77 (Una resta con su factor de amplificación calculado de antemano)**.** Sean $a = 1{,}234\,567\,890\,123$ y $b = 1{,}234\,567\,890\,000$, cuya diferencia exacta es $1{,}23 \cdot 10^{-10}$. El factor del paso 3 es

$$\kappa_- = \frac{|a| + |b|}{|a - b|} = \frac{2{,}469\,135\,78}{1{,}23 \cdot 10^{-10}} = 2{,}0074 \cdot 10^{10}.$$

Como los datos, al guardarse en doble precisión, ya traen un error relativo de hasta $u = 1{,}11 \cdot 10^{-16}$, se predice para la diferencia un error relativo de hasta $\kappa_- \, u = 2{,}23 \cdot 10^{-6}$. Y en efecto: la resta calculada en doble precisión devuelve $1{,}230\,000\,5459 \cdot 10^{-10}$, con un error relativo de $4{,}44 \cdot 10^{-7}$, por debajo de la cota prevista y del mismo orden ✓. De las dieciséis cifras iniciales quedan seis: se han perdido $\log_{10} \kappa_- \approx 10$, exactamente las diez cifras en que a y b coincidían.

Lema 14.78 (Suma exacta de dos números, o lema de Dekker)**.** *Sean s e y números de máquina con $|s| \geq |y|$, y sea $t = \mathrm{fl}(s + y)$ sin desbordamiento. Entonces la cantidad $(t - s) - y$ es exactamente representable y se calcula sin error alguno:*

$$\mathrm{fl}(\mathrm{fl}(t - s) - y) = t - (s + y),$$

es decir, el error de redondeo cometido al formar t se recupera exactamente con dos restas.

Esquema de la demostración. La clave es que ambas restas son de las que *no* pierden nada. Sea e_s el exponente de s. Como $|s| \geq |y|$, se tiene $|s + y| \leq 2|s|$, luego t tiene exponente e_s o $e_s + 1$, y en cualquier caso t es un múltiplo entero de $2^{\,e_s - p + 1}$ o de $2^{\,e_s - p + 2}$. Por otro lado $|t - s| \leq |y| + |t - (s + y)| \leq |y| + \frac{1}{2}\,\mathrm{ulp}(t) \leq 2|s|$: la diferencia $t - s$ es un múltiplo entero de la separación de la retícula de s y no excede el rango de esa retícula, así que es un número de máquina y la primera resta es exacta (éste es el lema de Sterbenz en su forma habitual: si $s/2 \leq t \leq 2s$ la resta $t - s$ es exacta). Para la segunda, $t - s$ e y son ambos múltiplos enteros de $2^{\,e_y - p + 1}$, siendo e_y el exponente de y, y su diferencia $t - (s + y)$ tiene módulo $\leq \frac{1}{2}\,\mathrm{ulp}(t)$, que cabe holgadamente en p bits sobre esa retícula; luego también es exacta. □

Regla: Suma compensada de Kahan

Para sumar $x_1, \ldots, x_n$ sin que el orden importe casi nada:

1. **Inicialice** el acumulador $s = 0$ y el compensador $c = 0$. El compensador guardará, con signo cambiado, la parte que la última suma no pudo retener.
2. **Para cada x_k, corrija el sumando primero**: $y = x_k - c$. Aquí se devuelve al total lo que se perdió en el paso anterior.
3. **Sume**: $t = s + y$. Éste es el único paso que redondea de verdad.
4. **Recupere lo perdido**: $c = (t - s) - y$. Por el lema 14.78, si $|s| \geq |y|$ esta cuenta es exacta y c es exactamente el error cometido en el paso 3.
5. **Actualice** $s = t$ y pase al siguiente término.
6. **Devuelva** s al terminar. No devuelva $s - c$ salvo que quiera la última corrección incluida.

7. **Verifique** comparando con la suma ingenua hecha en el orden inverso: si las tres cifras últimas de la compensada coinciden con las de un cálculo en precisión doble y las de la ingenua no, el método está funcionando.

Teorema 14.79 (Cota de la suma compensada)**.** *Con la notación anterior, y si* $nu \le \frac{1}{10}$*, el resultado* $\hat{s}$ *de la suma compensada cumple*

$$\left|\hat{s} - \sum_{i=1}^{n} x_i\right| \le (2u + O(nu^2)) \sum_{i=1}^{n} |x_i|.$$

La cota es esencialmente independiente de n*, frente al* $\gamma_{n-1} \approx nu$ *de la suma ingenua.*

Ejemplo 14.80 (La compensación seguida paso a paso)**.** Trabajemos en precisión simple, donde $2^{24} = 16\,777\,216$ y, por el corolario sobre los enteros representables, $2^{24}+1$ no es número de máquina. Sumemos la lista 2^{24}, 1, 1, 1, 1, cuyo total exacto es $16\,777\,220$.

Suma ingenua: $2^{24} + 1$ da 2^{24} por redondeo al par, y lo mismo ocurre las cuatro veces. Resultado: $16\,777\,216$. Se han perdido las cuatro unidades enteras.

Suma compensada: en la retícula de $[2^{24}, 2^{25})$ la separación es 2.

x_k	$y = x_k - c$	$t = \mathrm{fl}(s + y)$	$c = (t - s) - y$	s nuevo
16 777 216	16 777 216	16 777 216	0	16 777 216
1	1	16 777 216	-1	16 777 216
1	2	16 777 218	0	16 777 218
1	1	16 777 220	1	16 777 220
1	0	16 777 220	0	16 777 220

Resultado: $16\,777\,220$, exacto. Merece la pena seguir la segunda fila: la suma $16\,777\,216 + 1$ se redondea a $16\,777\,216$, de modo que $t - s = 0$ y $c = 0 - 1 = -1$; en el paso siguiente ese -1 se resta del nuevo sumando, $y = 1 - (-1) = 2$, y esa vez sí cabe en la retícula. La cuarta fila muestra el desempate al par actuando en sentido contrario: $16\,777\,218 + 1$ está en el punto medio y sube a $16\,777\,220$, con lo que $c = 2 - 1 = +1$ y el exceso queda anotado para descontarlo después. La contabilidad cuadra exactamente.

Cuidado

La suma compensada se estropea si el compilador «optimiza». Un compilador que aplique las reglas del álgebra sobre los reales simplificará $c = (t - s) - y$ a $c = (s + y - s) - y = 0$, y el algoritmo se convertirá en la suma ingenua sin que nada avise. Por eso los lenguajes serios ofrecen modos de compilación que prohíben reasociar operaciones en punto flotante, y por eso conviene no activar las opciones de aritmética «rápida» en un programa que dependa de este tipo de trucos.

Pausa para pensar

El lema 14.78 tiene una consecuencia filosófica bonita: el punto flotante no *destruye* información en una suma, la *desplaza* a un sitio donde deja de ser visible. El par (t, c) contiene exactamente la misma información que $s + y$; lo que ocurre es que un único número de máquina no puede guardarla y hace falta usar dos. Sobre esa idea se construye la aritmética de doble-doble, que representa un número por la suma de dos flotantes y consigue treinta y dos cifras decimales sin salir del hardware ordinario.

14.23. El número de condición: cuánto amplifica un problema los errores de sus datos

El apartado siguiente usará repetidamente el símbolo $\kappa(A)$ y la frase «con $\kappa \approx 10^k$ pierdes unas k cifras». Conviene tener antes la definición, la demostración de la cota que la justifica y —esto es lo que casi nunca se dice— la versión escalar del mismo concepto, que se aplica a cualquier problema, no solo a los sistemas lineales.

Definición 14.81. Para $\vec{x} = (x_1, \dots, x_n) \in \mathbb{R}^n$ se definen

$$\|\vec{x}\|_1 = \sum_i |x_i|, \qquad \|\vec{x}\|_2 = \Big(\sum_i x_i^2\Big)^{1/2}, \qquad \|\vec{x}\|_\infty = \max_i |x_i|.$$

Dada una de ellas, la **norma matricial subordinada** de una matriz A de orden n es

$$\|A\| = \max_{\vec{x} \neq \vec{0}} \frac{\|A\vec{x}\|}{\|\vec{x}\|},$$

es decir, el máximo factor por el que A puede alargar un vector. Se calcula con estas fórmulas: $\|A\|_1$ es el mayor de los sumatorios de módulos por *columnas*, $\|A\|_\infty$ es el mayor de los sumatorios de módulos por *filas*, y $\|A\|_2$ es el mayor valor singular de A. De la definición se sigue $\|A\vec{x}\| \le \|A\|\,\|\vec{x}\|$ y $\|AB\| \le \|A\|\,\|B\|$.

Definición 14.82. Si A es invertible, su **número de condición** respecto de una norma subordinada es

$$\kappa(A) = \|A\| \cdot \|A^{-1}\|.$$

Si A es singular se conviene $\kappa(A) = \infty$. Con la norma euclídea, $\kappa_2(A) = \sigma_{\text{máx}}/\sigma_{\text{mín}}$.

Proposición 14.83. *Para toda matriz invertible y toda norma subordinada, $\kappa(A) \ge 1$; y $\kappa(A) = 1$ para las matrices ortogonales en la norma euclídea.*

Demostración. $1 = \|I\| = \|AA^{-1}\| \le \|A\|\,\|A^{-1}\| = \kappa(A)$, usando que la norma subordinada de la identidad vale 1 y la submultiplicatividad. Si A es ortogonal, $\|A\vec{x}\|_2 = \|\vec{x}\|_2$ para todo $\vec{x}$, luego $\|A\|_2 = 1$, y lo mismo para $A^{-1} = A^{\mathsf{T}}$. □

Teorema 14.84 (Perturbación de los datos)**.** *Sea A invertible, $A\vec{x} = \vec{b}$ con $\vec{b} \neq \vec{0}$, y sea $\vec{x} + \Delta\vec{x}$ la solución de $A(\vec{x} + \Delta\vec{x}) = \vec{b} + \Delta\vec{b}$. Entonces*

$$\frac{\|\Delta\vec{x}\|}{\|\vec{x}\|} \;\leq\; \kappa(A)\,\frac{\|\Delta\vec{b}\|}{\|\vec{b}\|}.$$

Demostración. Restando las dos ecuaciones, $A\Delta\vec{x} = \Delta\vec{b}$, luego $\Delta\vec{x} = A^{-1}\Delta\vec{b}$ y

$$\|\Delta\vec{x}\| \leq \|A^{-1}\|\,\|\Delta\vec{b}\|.$$

Por otro lado $\|\vec{b}\| = \|A\vec{x}\| \leq \|A\|\,\|\vec{x}\|$, de donde $\dfrac{1}{\|\vec{x}\|} \leq \dfrac{\|A\|}{\|\vec{b}\|}$. Multiplicando las dos desigualdades miembro a miembro se obtiene el enunciado. □

Teorema 14.85 (Residuo pequeño no implica error pequeño)**.** *Sea $\hat{\vec{x}}$ una solución aproximada de $A\vec{x} = \vec{b}$ y sea $\vec{r} = \vec{b} - A\hat{\vec{x}}$ su residuo. Entonces*

$$\frac{\|\hat{\vec{x}} - \vec{x}\|}{\|\vec{x}\|} \;\leq\; \kappa(A)\,\frac{\|\vec{r}\|}{\|\vec{b}\|},$$

y la desigualdad puede alcanzarse. Por tanto un residuo relativo de 10^{-12} solo garantiza un error relativo de $10^{-12}\kappa(A)$, que con $\kappa(A) = 10^{13}$ no garantiza absolutamente nada.

Demostración. De $\vec{r} = \vec{b} - A\hat{\vec{x}} = A\vec{x} - A\hat{\vec{x}} = A(\vec{x} - \hat{\vec{x}})$ sale $\hat{\vec{x}} - \vec{x} = -A^{-1}\vec{r}$, luego $\|\hat{\vec{x}} - \vec{x}\| \leq \|A^{-1}\|\,\|\vec{r}\|$. Se combina con $1/\|\vec{x}\| \leq \|A\|/\|\vec{b}\|$ igual que antes. □

Proposición 14.86 (Cuántas cifras se pierden)**.** *Si los datos se conocen con d cifras significativas correctas, la solución de $A\vec{x} = \vec{b}$ solo puede garantizarse con $d - \log_{10}\kappa(A)$ cifras.*

Demostración. Conocer $\vec{b}$ con d cifras significa $\|\Delta\vec{b}\|/\|\vec{b}\| \approx 10^{-d}$. Por el teorema 14.84, $\|\Delta\vec{x}\|/\|\vec{x}\| \leq \kappa(A)10^{-d} = 10^{\log_{10}\kappa(A) - d}$, que es un error relativo del orden de $10^{-(d - \log_{10}\kappa)}$, es decir, $d - \log_{10}\kappa(A)$ cifras correctas. □

Definición 14.87. Sea f derivable y $f(x) \neq 0$. El **número de condición relativo del problema** «calcular f en x» es

$$\kappa_f(x) = \left|\frac{x\,f'(x)}{f(x)}\right|.$$

Es el factor por el que un error relativo en x se traduce en error relativo en $f(x)$: si $\tilde{x} = x(1+\varepsilon)$, entonces $f(\tilde{x}) \approx f(x)\big(1 + \kappa_f(x)\,\varepsilon\big)$ salvo términos de orden ε^2.

Proposición 14.88. *El problema «restar b de a», visto como función de los dos datos, tiene número de condición*

$$\kappa_-(a, b) = \frac{|a| + |b|}{|a - b|},$$

que tiende a infinito cuando $b \to a$. En cambio $\kappa_{\sqrt{\ }}(x) = \frac{1}{2}$ para todo $x > 0$, y $\kappa_\times = 2$ para el producto.

Demostración. Con $\tilde a = a(1+\varepsilon_a)$ y $\tilde b = b(1+\varepsilon_b)$,

$$\tilde a - \tilde b = (a-b) + a\varepsilon_a - b\varepsilon_b = (a-b)\Big(1 + \frac{a\varepsilon_a - b\varepsilon_b}{a-b}\Big),$$

y el error relativo del resultado está acotado por $\dfrac{|a|+|b|}{|a-b|}\,\max(|\varepsilon_a|,|\varepsilon_b|)$, con igualdad cuando los signos se alinean. Para $f(x)=\sqrt{x}$, $\kappa_f = |x\cdot \frac{1}{2\sqrt{x}}/\sqrt{x}| = \frac12$. Para el producto, $\widetilde{ab} = ab(1+\varepsilon_a)(1+\varepsilon_b) \approx ab(1+\varepsilon_a+\varepsilon_b)$, luego el factor es 2. □

Regla: Estimar el condicionamiento antes de creerse un resultado

1. **Identifique cuál es el problema y cuáles los datos.** Condicionamiento es una propiedad del par problema-datos, no del programa. Cambiar de algoritmo no lo cambia.
2. **Si el problema es escalar, derive**: calcule $\kappa_f(x) = |xf'(x)/f(x)|$ en el punto en cuestión. Si sale del orden de la unidad, el problema es benigno.
3. **Si el problema es un sistema lineal, calcule o estime $\kappa(A)$.** A mano, con la norma del infinito: mayor suma de módulos por filas de A, multiplicada por la de A^{-1}. Con ordenador, cualquier biblioteca lo estima sin invertir la matriz.
4. **Traduzca a cifras**: reste $\log_{10}\kappa$ a las cifras que traigan los datos. Si el resultado es negativo, no hay ninguna cifra fiable y el cálculo carece de sentido tal como está planteado.
5. **No confunda condicionamiento con estabilidad.** Un problema bien condicionado resuelto con un algoritmo inestable se arregla cambiando de fórmula; un problema mal condicionado no se arregla con nada, y hay que replantearlo (cambiar de base, escalar las variables, reformular).
6. **Escale antes de medir.** Multiplicar una fila entera del sistema por 10^6 cambia $\kappa(A)$ sin cambiar el problema. Conviene normalizar las filas antes de sacar conclusiones del número de condición.
7. **Verifique** perturbando los datos a propósito en la última cifra y volviendo a resolver: si la respuesta se mueve mucho más de lo que predice κ, hay un error en el cálculo; si se mueve como predice κ, el diagnóstico es correcto y la culpa es del problema.

Ejemplo 14.89 (Cinco condicionamientos calculados)**.** (a) *Una matriz inofensiva, a mano.* Sea $A = \begin{pmatrix} 2 & 1 \\ 1 & 3 \end{pmatrix}$, con $\det A = 5$ y

$$A^{-1} = \frac{1}{5}\begin{pmatrix} 3 & -1 \\ -1 & 2 \end{pmatrix}.$$

Sumas de módulos por filas de A: 3 y 4, luego $\|A\|_\infty = 4$. Para A^{-1}: $\frac45$ y $\frac35$, luego $\|A^{-1}\|_\infty = \frac45$. Por tanto $\kappa_\infty(A) = 4\cdot\frac45 = \frac{16}{5} = 3{,}2$. Como A es simétrica, sus valores

singulares son los módulos de sus autovalores, que son las raíces de $\lambda^2 - 5\lambda + 5 = 0$, es decir $\frac{5\pm\sqrt{5}}{2}$, y

$$\kappa_2(A) = \frac{5+\sqrt{5}}{5-\sqrt{5}} = \frac{3+\sqrt{5}}{2} = 2{,}618\,034.$$

Los dos números son del mismo orden; con $\kappa \approx 3$ no se pierde ni media cifra y el sistema es completamente fiable.

(b) *La matriz de las dos rectas casi paralelas, exactamente.* Sea $A = \begin{pmatrix} 1 & 1 \\ 1 & 1{,}0001 \end{pmatrix}$, con $\det A = 10^{-4}$ y

$$A^{-1} = 10^4 \begin{pmatrix} 1{,}0001 & -1 \\ -1 & 1 \end{pmatrix} = \begin{pmatrix} 10\,001 & -10\,000 \\ -10\,000 & 10\,000 \end{pmatrix}.$$

Entonces $\|A\|_\infty = 2{,}0001$ y $\|A^{-1}\|_\infty = 20\,001$, luego

$$\kappa_\infty(A) = 2{,}0001 \cdot 20\,001 = 40\,004{,}0001,$$

y en la norma euclídea $\kappa_2(A) = 40\,002{,}00008$ —el valor que se citó antes—. Ambos dicen lo mismo: $\log_{10}\kappa \approx 4{,}6$, se pierden entre cuatro y cinco cifras. Comprobación del paso 7: con $\vec{b} = (2;\, 2{,}0001)$ la solución es $(1, 1)$; perturbando a $\vec{b} = (2;\, 2{,}0002)$, es decir con $\|\Delta\vec{b}\|_\infty/\|\vec{b}\|_\infty = 10^{-4}/2{,}0001 = 4{,}99975 \cdot 10^{-5}$, la solución pasa a $(0, 2)$ y el cambio relativo es $\|\Delta\vec{x}\|_\infty/\|\vec{x}\|_\infty = 1$. La cota del teorema 14.84 predice como máximo $40\,004 \cdot 4{,}99975 \cdot 10^{-5} = 2{,}0001$: se cumple, y con poco margen ✓.

(c) *Residuo engañoso, con la cota puesta.* Con la misma A y $\vec{b} = (2;\, 2{,}0001)$, propóngase $\hat{\vec{x}} = (0, 2)$. El residuo es $\vec{r} = (0;\, -10^{-4})$, con $\|\vec{r}\|_\infty/\|\vec{b}\|_\infty = 4{,}99975 \cdot 10^{-5}$. El teorema 14.85 solo garantiza $\|\hat{\vec{x}} - \vec{x}\|_\infty/\|\vec{x}\|_\infty \le 2{,}0001$, es decir, no garantiza nada útil; y el error real es 1. El residuo minúsculo era compatible, desde el principio, con una respuesta completamente equivocada.

(d) *Un problema escalar mal condicionado sin matrices de por medio.* Calcular $\ln x$ cerca de $x = 1$. Aquí $\kappa_f(x) = |x \cdot (1/x)/\ln x| = 1/|\ln x|$. En $x = 1{,}0001$ es $\ln x = 9{,}99950 \cdot 10^{-5}$ y

$$\kappa_f = 10\,000{,}5.$$

Comprobación: perturbando x en un 10^{-8} por ciento, o sea con error relativo 10^{-10}, el logaritmo cambia en $1{,}000\,05 \cdot 10^{-6}$ relativo, que es exactamente $\kappa_f \cdot 10^{-10}$ ✓. No hay algoritmo que evite esto: el problema mismo amplifica por diez mil. La única salida es reformular, y aquí la reformulación existe y tiene nombre propio: calcular $\ln(1+t)$ con $t = x - 1$ como dato, función para la que $\kappa = |t/((1+t)\ln(1+t))| \to 1$ cuando $t \to 0$. Por eso las bibliotecas ofrecen `log1p` como función aparte.

(e) *Raíz cuadrada: el problema perfecto.* Para $f(x) = \sqrt{x}$ se tiene $\kappa_f = \frac{1}{2}$ en todo punto: la raíz cuadrada *reduce* a la mitad el error relativo de su dato. Verificación: $\sqrt{2 \cdot (1 + 10^{-8})}$ difiere de $\sqrt{2}$ en un $5 \cdot 10^{-9}$ relativo, la mitad justa ✓. Es la razón de que los algoritmos que pueden expresarse con raíces y evitar restas —la norma euclídea calculada con la fórmula de Moler y Morrison, por ejemplo— sean tan estables.

Cuidado

El número de condición *no* es una medida de lo cerca que está la matriz de ser singular en términos de su determinante. La matriz $10^{-8}I$ de orden 2 tiene determinante 10^{-16}, prácticamente cero, y sin embargo $\kappa = 1$: es perfectamente fiable. Y una matriz de determinante 1 puede tener $\kappa = 10^{12}$. El determinante depende de la escala y el número de condición no; usar el determinante como diagnóstico de mal condicionamiento es uno de los errores más extendidos, y siempre acaba mal.

Observación 14.90. La proposición 14.88 cierra el círculo con el apartado de la cancelación catastrófica: aquello no era un capricho de la aritmética, sino la manifestación de que *la resta de números parecidos es un problema mal condicionado*. Lo que ocurre en los ejemplos de la ecuación de segundo grado y del límite de $(1-\cos x)/x^2$ es que el problema global —calcular la raíz pequeña, calcular el límite— está *bien* condicionado, pero la fórmula elegida lo descompone en pasos de los cuales uno está mal condicionado. Reescribir la fórmula es reorganizar la descomposición para que ningún paso intermedio lo esté. Ésa es, en una frase, la definición de algoritmo estable.

Pausa para pensar

Hay una asimetría que conviene tener presente. El condicionamiento se refiere siempre a errores *relativos*, y eso presupone que la respuesta no es cero. Cerca de una raíz, calcular $f(x)$ está infinitamente mal condicionado en sentido relativo, porque $f(x) \to 0$ y cualquier error absoluto es infinitamente grande en proporción. Por eso los criterios de parada de los métodos de este capítulo no exigen $|f(x_k)|$ pequeño en términos relativos, sino $|x_{k+1} - x_k|$ pequeño: se mide el condicionamiento donde tiene sentido medirlo.

14.24. Ejercicios resueltos

Ejercicio 14.1. Comprueba que $f(x) = x^3 - x - 1$ tiene una raíz en $[1, 2]$ y da las tres primeras iteraciones de bisección.

Solución. $f(1) = 1 - 1 - 1 = -1 < 0$ y $f(2) = 8 - 2 - 1 = 5 > 0$, y f es continua por ser polinómica. Por Bolzano hay raíz.

Paso 1: $m = 1{,}5$, $f(1{,}5) = 3{,}375 - 1{,}5 - 1 = 0{,}875 > 0$, luego la raíz está en $[1;\, 1{,}5]$.

Paso 2: $m = 1{,}25$, $f(1{,}25) = 1{,}953125 - 2{,}25 = -0{,}296875 < 0$, luego está en $[1{,}25;\, 1{,}5]$.

Paso 3: $m = 1{,}375$, $f(1{,}375) = +0{,}2246 > 0$, luego está en $[1{,}25;\, 1{,}375]$. El error de dar 1,3125 es a lo sumo 0,0625. ■

Ejercicio 14.2. Aplica dos pasos de Newton a $f(x) = x^2 - 5$ partiendo de $x_0 = 2$.

Solución. La iteración es $x_{n+1} = \frac{1}{2}\left(x_n + \frac{5}{x_n}\right)$.

$x_1 = \frac{1}{2}(2 + 2{,}5) = 2{,}25$. $\quad x_2 = \frac{1}{2}\left(2{,}25 + \frac{5}{2{,}25}\right) = \frac{1}{2}(2{,}25 + 2{,}2\overline{2}) = 2{,}236111\ldots$

El valor real es $\sqrt{5} = 2{,}2360679\ldots$: dos pasos y cinco cifras correctas. Un tercer paso daría diez. ■

Ejercicio 14.3. ¿Cuántos pasos de bisección hacen falta para localizar una raíz en $[2, 3]$ con error menor que 10^{-4}?

Solución. Hay que imponer $\frac{3-2}{2^{n+1}} < 10^{-4}$, es decir $2^{n+1} > 10\,000$.

Como $2^{13} = 8192$ y $2^{14} = 16\,384$, hace falta $n+1 \geq 14$, o sea $n \geq 13$. Trece pasos. Obsérvese que la respuesta no depende de la función, solo del intervalo: es la gran virtud del método. ■

Ejercicio 14.4. Aproxima $\int_0^1 x^2\,\mathrm{d}x$ con la regla de los trapecios y $n = 4$, y compara con el valor exacto.

Solución. Con $h = 0{,}25$ los valores son $f(0) = 0$, $f(0{,}25) = 0{,}0625$, $f(0{,}5) = 0{,}25$, $f(0{,}75) = 0{,}5625$, $f(1) = 1$.

$$0{,}25\left[\tfrac{0}{2} + 0{,}0625 + 0{,}25 + 0{,}5625 + \tfrac{1}{2}\right] = 0{,}25 \cdot 1{,}375 = 0{,}34375.$$

El valor exacto es $1/3 = 0{,}3\overline{3}$, luego el error es $+0{,}01042$. El signo positivo no es casual: la parábola x^2 es convexa, y los segmentos que la aproximan quedan siempre *por encima* de ella, de modo que los trapecios sobrestiman toda integral de función convexa. ■

Ejercicio 14.5. Repite el cálculo anterior con Simpson y explica el resultado.

Solución.

$$\frac{0{,}25}{3}\left[0 + 4(0{,}0625) + 2(0{,}25) + 4(0{,}5625) + 1\right] = \frac{0{,}25}{3} \cdot 4 = 0{,}3\overline{3}.$$

Sale exacto. Y tenía que salir: Simpson sustituye la función por parábolas, y aquí la función *es* una parábola, de modo que la aproximación coincide con el original. Por la misma razón sería exacto para cualquier polinomio de grado ≤ 3. ■

Ejercicio 14.6. Explica por qué la iteración $x_{n+1} = x_n^2$ no sirve para hallar el punto fijo $x = 1$, y qué ocurre según dónde se empiece.

Solución. Aquí $g(x) = x^2$ y $g'(x) = 2x$, luego en el punto fijo $p = 1$ se tiene $|g'(1)| = 2 > 1$. El criterio predice divergencia, y así es.

Empezando en $x_0 = 0{,}9$ la sucesión da 0,81, 0,656, 0,430… y se va al otro punto fijo, el 0, donde $|g'(0)| = 0 < 1$ y que por tanto sí es atractivo. Empezando en $x_0 = 1{,}1$ da 1,21, 1,46, 2,14… y se escapa a infinito. Solo empezando exactamente en 1 se queda ahí, en un equilibrio inestable: es el lápiz de pie sobre su punta. ■

14.25. Ejercicios propuestos

14.1. Comprueba que $x^3 + x - 3 = 0$ tiene raíz en $[1, 2]$ y haz dos pasos de bisección.

14.2. ¿Cuántos pasos de bisección aseguran error menor que 0,01 partiendo de $[0, 4]$?

14.3. Escribe la iteración de Newton para $f(x) = x^2 - 7$ y haz dos pasos desde $x_0 = 3$.

14.4. Escribe la iteración de Newton para $f(x) = \cos x - x$.

14.5. ¿Por qué falla Newton para $f(x) = x^3$ empezando en $x_0 = 0$?

14.6. Estima $\sqrt{10}$ por interpolación lineal entre $\sqrt{9} = 3$ y $\sqrt{16} = 4$, y compara con el valor real.

14.7. Aplica la regla de los trapecios a $\int_0^2 x^3\mathrm{d}x$ con $n = 2$ y compara con el valor exacto.

14.8. Aplica Simpson a la misma integral con $n = 2$. ¿Qué esperas obtener y por qué?

14.9. Para $g(x) = \frac{x}{2} + 1$, halla el punto fijo y decide si la iteración converge.

14.10. Para $g(x) = 3x - 1$, haz lo mismo.

14.11. Reescribe $\sqrt{x+1}-\sqrt{x}$ de forma que no haya cancelación para x grande.

14.12. Si un método triplica el número de cifras correctas en cada paso, ¿cuántos pasos necesita para pasar de 1 cifra a 30?

Respuestas

1. $f(1)=-1<0$ y $f(2)=7>0$. Paso 1: $f(1{,}5)=1{,}875>0$, raíz en $[1;1{,}5]$. Paso 2: $f(1{,}25)=0{,}203>0$, raíz en $[1;1{,}25]$. **2.** $\frac{4}{2^{n+1}}<0{,}01$ exige $2^{n+1}>400$, y como $2^9=512$, basta $n\geq 8$. **3.** $x_{n+1}=\frac{1}{2}(x_n+7/x_n)$; desde 3: $x_1=\frac{1}{2}(3+2{,}3\overline{3})=2{,}6\overline{6}$ y $x_2=2{,}645833$, frente a $\sqrt{7}=2{,}6457513$. **4.** $x_{n+1}=x_n-\dfrac{\cos x_n-x_n}{-\sin x_n-1}=x_n+\dfrac{\cos x_n-x_n}{\sin x_n+1}$. **5.** Porque $f'(0)=0$: la tangente en el origen es el propio eje X y la fórmula divide entre cero. (Empezando en cualquier otro punto sí converge, aunque despacio, por ser raíz triple.) **6.** $3+\frac{10-9}{16-9}=3+\frac{1}{7}\approx 3{,}1429$, frente a 3,16228: error del 0,6 %. **7.** $h=1$ y valores $0,1,8$: $1\left[\frac{0}{2}+1+\frac{8}{2}\right]=5$, frente al valor exacto 4. Sobrestima, por ser x^3 convexa en $[0,2]$. **8.** $\frac{1}{3}[0+4\cdot 1+8]=4$, exacto, porque Simpson integra sin error los polinomios de grado hasta tres. **9.** Punto fijo $p=2$; $|g'|=\frac{1}{2}<1$, luego converge desde cualquier punto de partida. **10.** Punto fijo $p=\frac{1}{2}$; $|g'|=3>1$, luego diverge salvo si se empieza exactamente en $\frac{1}{2}$. **11.** Multiplicando por el conjugado: $\dfrac{1}{\sqrt{x+1}+\sqrt{x}}$, que para x grande vale aproximadamente $\frac{1}{2\sqrt{x}}$ y no resta nada. **12.** Las cifras van $1,3,9,27,81$: al cuarto paso ya se superan las 30. Con orden de convergencia 3 el crecimiento es exponencial, igual que la duplicación de Newton pero más rápido.

Capítulo 15

Optimización y análisis convexo

Lo que necesitas traer

Del capítulo 6: derivadas y el criterio de la segunda derivada. Del capítulo 7: matrices simétricas, valores propios y formas definidas positivas. Del capítulo 9: gradiente, hessiana y multiplicadores de Lagrange.

Optimizar es buscar el mejor elemento de un conjunto según un criterio. La dificultad no está en que haya muchos candidatos, sino en la posibilidad de quedarse atrapado en un óptimo local creyendo que es el global. La convexidad es exactamente la hipótesis que elimina ese peligro, y esa es la razón de que separe los problemas tratables de los que no lo son. La frontera relevante en optimización no es lineal frente a no lineal, sino convexo frente a no convexo.

15.1. Convexidad y condiciones de optimalidad

Definición 15.1. Un conjunto C es *convexo* si contiene el segmento entre dos puntos cualesquiera suyos. Una función f es *convexa* si $f(\lambda x + (1-\lambda)y) \le \lambda f(x) + (1-\lambda) f(y)$; si es C^2, equivale a que su hessiana sea semidefinida positiva. Un *problema convexo* minimiza una función convexa sobre un conjunto convexo.



Figura 15.1: Izquierda: por cada punto de la frontera de un convexo pasa un hiperplano que lo deja entero a un lado. Derecha: en cuanto el conjunto tiene una entrante, hay pares de puntos cuyo segmento se escapa.

Idea clave

Qué dice esto en cristiano. En un problema convexo *todo mínimo local es global*, y esa única frase es la que lo cambia todo. Si un algoritmo baja por la pendiente y se detiene, ha terminado: no hay ningún valle mejor escondido en otra parte. En un problema no convexo, en cambio, detenerse no significa nada, y la única manera de estar seguro sería explorar todo el espacio. Además, la convexidad regala un *certificado* de optimalidad: la teoría de la dualidad permite demostrar que la solución encontrada es la mejor, sin explorar el resto.

Regla: plantear y resolver un problema de optimización

1. **Escribe la función objetivo y las restricciones** por separado, distinguiendo las de igualdad de las de desigualdad.
2. **Comprueba la convexidad**: objetivo convexo, restricciones de desigualdad convexas y las de igualdad afines. Si falla, cambia de expectativas.
3. **Sin restricciones, iguala el gradiente a cero** y clasifica con la hessiana.
4. **Con restricciones de igualdad, usa multiplicadores de Lagrange**: el gradiente del objetivo debe ser combinación de los de las restricciones.
5. **Con desigualdades, aplica las condiciones de Karush–Kuhn–Tucker**, que añaden la holgura complementaria: cada multiplicador es cero o su restricción está activa.
6. **Plantea el problema dual** y compara: en el caso convexo con una condición de cualificación, los valores óptimos coinciden y esa igualdad certifica la solución.
7. **Verifica numéricamente** con un método de descenso, comprobando que converge al mismo punto desde varios inicios distintos.

Ejemplo 15.2 (Optimización con y sin restricciones)**. (a) Un caso sin restricciones.** Minimizar $f(x,y) = x^2 + 2y^2 - 2x - 4y$. El gradiente $(2x-2,\ 4y-4)$ se anula en $(1,1)$, y la hessiana $\begin{pmatrix} 2 & 0 \\ 0 & 4 \end{pmatrix}$ es definida positiva, luego la función es convexa y ese punto es el mínimo *global*, con valor $f(1,1) = -3$.

(b) Con una restricción de igualdad. Minimizar $x^2 + y^2$ sujeto a $x + y = 2$. Lagrange da $(2x, 2y) = \lambda(1,1)$, luego $x = y$, y con la restricción $x = y = 1$ y valor 2. Geométricamente: el punto de la recta más cercano al origen es el pie de la perpendicular, y el multiplicador vale $\lambda = 2$.

(c) Con desigualdad, holgura complementaria. Minimizar x^2 sujeto a $x \geq 1$. La condición KKT es $2x = \mu$ con $\mu \geq 0$ y $\mu(x-1) = 0$. Si $\mu = 0$ saldría $x = 0$, que viola la restricción; luego $\mu > 0$ y la restricción está activa, $x = 1$, con $\mu = 2$ y valor 1. El multiplicador mide cuánto mejoraría el óptimo si se relajara la restricción.

(d) Dualidad con cifras. En programación lineal, maximizar $3x + 2y$ sujeto a $x + y \leq 4$,

$x \le 3$ y $x, y \ge 0$ da el óptimo en $(3, 1)$ con valor 11. El problema dual, minimizar $4u + 3v$ sujeto a $u + v \ge 3$ y $u \ge 2$, alcanza también 11 en $(2, 1)$. Esa coincidencia es el teorema de dualidad fuerte, y el par de soluciones se certifica mutuamente.

Cuidado

[Sin convexidad todo lo anterior deja de valer] Con una función no convexa, un punto crítico puede ser un mínimo local pésimo, y ningún algoritmo de descenso distingue eso de un óptimo global. El entrenamiento de redes neuronales es el caso masivo: el problema es no convexo, no hay certificado de optimalidad, y todo lo que se obtiene es un punto razonablemente bueno al que se llegó. Segundo aviso: los multiplicadores de Lagrange dan condiciones *necesarias*, no suficientes, y solo bajo convexidad son también suficientes. Y tercero: las condiciones KKT requieren una cualificación de restricciones (por ejemplo la de Slater, que exista un punto estrictamente factible); sin ella pueden fallar incluso en problemas convexos.

Pausa para pensar

Durante décadas se enseñó que la frontera de la dificultad estaba entre problemas lineales y no lineales, porque el simplex de Dantzig resolvía los primeros desde 1947. La reformulación moderna, popularizada por Rockafellar y luego por Boyd, es que la frontera real está entre convexo y no convexo: hay problemas no lineales que se resuelven en segundos con métodos de punto interior y problemas de aspecto inocente que son NP-difíciles. El cambio de perspectiva tuvo consecuencias prácticas enormes, porque muchos problemas de ingeniería que parecían intratables resultaron ser convexos tras un cambio de variable adecuado, y pasaron de irresolubles a rutinarios.

15.2. Los cimientos: combinaciones convexas, epigrafo, Jensen y los criterios de primer y segundo orden

La sección anterior definió convexo y convexa en dos líneas y se lanzó a usarlas. Pero la definición de conjunto convexo habla de segmentos entre *dos* puntos y luego el capítulo maneja envolventes convexas de familias enteras; la de función convexa aparece sin decir qué es una función estrictamente convexa ni qué es un epigrafo, pese a que ambas cosas se citan más adelante; y la afirmación de que la hessiana semidefinida positiva equivale a la convexidad se enuncia de pasada, sin definir qué es semidefinida positiva, sin decir cómo se comprueba y sin demostrar nada. Esta sección pone esos cimientos. Todo lo que viene después del capítulo descansa sobre ellos.

Notación: producto escalar, gradiente, hessiana

Definición 15.3 (La notación que el capítulo da por sabida)**.** Para vectores columna $x, y \in \mathbb{R}^n$, el *producto escalar* es

$$\langle x, y \rangle \;=\; x^{\mathsf{T}} y \;=\; \sum_{i=1}^{n} x_i y_i,$$

donde x^{T} denota el vector *traspuesto*, es decir, el mismo vector escrito en fila; el libro escribe indistintamente x^{T} y $x^{\top}$, y para una matriz A la traspuesta A^{T} es la que resulta de intercambiar filas por columnas. La *norma euclídea* es $\|x\| = \sqrt{\langle x, x\rangle}$.

Si $f : \mathbb{R}^n \to \mathbb{R}$ es derivable, su *gradiente* en x es el vector de derivadas parciales $\nabla f(x) = (\partial f/\partial x_1, \dots, \partial f/\partial x_n)$, y si es dos veces derivable, su *hessiana* $\nabla^2 f(x)$ es la matriz $n \times n$ de segundas derivadas $(\partial^2 f/\partial x_i \partial x_j)$, que es simétrica cuando f es de clase C^2.

Una matriz simétrica H es *semidefinida positiva*, escrito $H \succeq 0$, si $v^{\mathsf{T}} H v \geq 0$ para todo $v \in \mathbb{R}^n$; y *definida positiva*, escrito $H \succ 0$, si $v^{\mathsf{T}} H v > 0$ para todo $v \neq 0$. Se escribe $A \preceq B$ para decir $B - A \succeq 0$.

Combinaciones convexas y envolvente

Definición 15.4 (Combinación convexa y envolvente convexa)**.** Una *combinación convexa* de los puntos $x_1, \dots, x_m \in \mathbb{R}^n$ es cualquier punto de la forma

$$\lambda_1 x_1 + \lambda_2 x_2 + \cdots + \lambda_m x_m, \qquad \lambda_i \geq 0, \qquad \sum_{i=1}^{m} \lambda_i = 1.$$

Es una media ponderada: los pesos no son negativos y suman uno. Con $m = 2$ se recupera el segmento $\lambda x + (1-\lambda)y$ con $\lambda \in [0,1]$ de la definición de conjunto convexo.

La *envolvente convexa* de un conjunto $S \subseteq \mathbb{R}^n$, escrita $\operatorname{conv}(S)$, es el conjunto de todas las combinaciones convexas de puntos de S. Equivalentemente, y esta es la caracterización útil, es el menor convexo que contiene a S.

Ejemplo 15.5 (Combinaciones y envolventes con cifras)**. (a) Una combinación concreta.** Con $x_1 = (1,0)$, $x_2 = (0,4)$, $x_3 = (2,2)$ y pesos $\lambda = (0{,}25;\ 0{,}25;\ 0{,}5)$, que suman 1, la combinación es

$$0{,}25\,(1,0) + 0{,}25\,(0,4) + 0{,}5\,(2,2) = (0{,}25 + 0 + 1,\ \ 0 + 1 + 1) = (1{,}25;\ 2).$$

Ese punto está dentro del triángulo de vértices x_1, x_2, x_3, como tiene que estar.

(b) La envolvente de tres puntos es el triángulo lleno, vértices, lados e interior. La de dos puntos es el segmento que los une. La de una circunferencia es el disco cerrado.

(c) Un punto que sobra. $\operatorname{conv}\{(0,0),(2,0),(0,2),(1,1)\}$ es el mismo triángulo que $\operatorname{conv}\{(0,0),(2,0),(0,2)\}$, porque $(1,1) = \frac{1}{2}(2,0) + \frac{1}{2}(0,2)$ ya es combinación convexa de los otros: añadir un punto que estaba dentro no agranda nada.

(d) Un conjunto no convexo y su envolvente. Si S son los cuatro vértices de un cuadrado, S no es convexo (no contiene ni un solo punto de sus lados) mientras que $\operatorname{conv}(S)$ es el cuadrado macizo.

Teorema 15.6 (Un convexo contiene todas sus combinaciones convexas)**.** *Un conjunto $C \subseteq \mathbb{R}^n$ es convexo si y solo si toda combinación convexa de puntos de C pertenece a C.*

Demostración. La implicación de derecha a izquierda es inmediata: el caso $m = 2$ es la propia definición de convexo. Para la otra, sea C convexo y procedamos por inducción sobre m. El caso $m = 1$ es trivial y el $m = 2$ es la definición. Supongamos el resultado cierto para m y

sea $x = \sum_{i=1}^{m+1} \lambda_i x_i$ una combinación convexa de $m+1$ puntos de C. Si $\lambda_{m+1} = 1$ entonces $x = x_{m+1} \in C$ y no hay nada que probar. En caso contrario $s = \sum_{i=1}^{m} \lambda_i = 1 - \lambda_{m+1} > 0$, y podemos escribir

$$x = s \underbrace{\sum_{i=1}^{m} \frac{\lambda_i}{s} x_i}_{=y} + \lambda_{m+1} x_{m+1}.$$

Los coeficientes λ_i/s son no negativos y suman 1, así que y es combinación convexa de m puntos de C y por hipótesis de inducción $y \in C$. Y $x = sy + (1-s)x_{m+1}$ es un punto del segmento entre y y x_{m+1}, luego está en C por convexidad. □

Proposición 15.7. *La intersección de cualquier familia de convexos es convexa.*

Demostración. Sea $C = \bigcap_{i \in I} C_i$ con cada C_i convexo, y sean $x, y \in C$ y $\lambda \in [0,1]$. Para cada i se tiene $x, y \in C_i$, y por convexidad de C_i el punto $\lambda x + (1-\lambda)y$ está en C_i. Como esto vale para todo i, el punto está en la intersección. □

Observación 15.8. De aquí sale, sin más trabajo, que el conjunto factible de un problema convexo es convexo: cada restricción $g_i(x) \le 0$ con g_i convexa define un conjunto convexo, cada igualdad afín $a^\mathsf{T} x = b$ define un hiperplano, que también lo es, y la región admisible es la intersección de todos ellos. La unión, en cambio, no conserva la convexidad: dos discos disjuntos lo demuestran.

Funciones convexas, epigrafo y Jensen

Definición 15.9 (Convexa, estrictamente convexa, cóncava, epigrafo)**.** Sea $C \subseteq \mathbb{R}^n$ convexo y $f : C \to \mathbb{R}$.

- f es *convexa* si para todos $x, y \in C$ y todo $\lambda \in [0,1]$

  $$f(\lambda x + (1-\lambda)y) \;\le\; \lambda f(x) + (1-\lambda) f(y).$$

  Geométricamente: la cuerda queda por encima de la gráfica.
- f es *estrictamente convexa* si la desigualdad es estricta siempre que $x \neq y$ y $\lambda \in (0,1)$: la cuerda no toca la gráfica salvo en los extremos.
- f es *cóncava* si $-f$ es convexa, es decir, si la desigualdad va al revés.
- El *epigrafo* de f es el conjunto de puntos que quedan por encima de su gráfica,

  $$\operatorname{epi} f = \{(x,t) \in C \times \mathbb{R} : \; f(x) \le t\} \subseteq \mathbb{R}^{n+1}.$$

Teorema 15.10 (Convexidad es convexidad del epigrafo)**.** *f es convexa si y solo si* $\operatorname{epi} f$ *es un conjunto convexo.*

Demostración. Supongamos f convexa y tomemos $(x,s), (y,t) \in \operatorname{epi} f$, es decir $f(x) \le s$ y $f(y) \le t$. Para $\lambda \in [0,1]$,

$$f(\lambda x + (1-\lambda)y) \le \lambda f(x) + (1-\lambda) f(y) \le \lambda s + (1-\lambda) t,$$

luego $\lambda(x,s)+(1-\lambda)(y,t)$ está en el epigrafo. Recíprocamente, si el epigrafo es convexo, dados x, y los puntos $(x, f(x))$ e $(y, f(y))$ están en él, luego también su combinación $(\lambda x+(1-\lambda)y,\ \lambda f(x)+(1-\lambda)f(y))$, y pertenecer al epigrafo significa exactamente que $f(\lambda x+(1-\lambda)y) \le \lambda f(x)+(1-\lambda)f(y)$. □

Observación 15.11. Este teorema es el que convierte propiedades de conjuntos en propiedades de funciones y viceversa, y explica de un plumazo dos afirmaciones que el capítulo usa más adelante. Que el máximo de convexas sea convexo es que $\operatorname{epi} \max\{f,g\} = \operatorname{epi} f \cap \operatorname{epi} g$, y la intersección de convexos es convexa. Y que la teoría de hiperplanos de apoyo se aplique a funciones es que se aplica a su epigrafo, que es un convexo de $\mathbb{R}^{n+1}$: de ahí nacen los subgradientes.

Teorema 15.12 (Desigualdad de Jensen, versión finita)**.** *Si f es convexa sobre C, entonces para todos $x_1,\dots,x_m \in C$ y todos pesos $\lambda_i \ge 0$ con $\sum \lambda_i = 1$,*

$$f\Big(\sum_{i=1}^{m} \lambda_i x_i\Big) \ \le\ \sum_{i=1}^{m} \lambda_i f(x_i).$$

Demostración. Inducción sobre m. El caso $m=2$ es la definición. Supuesto cierto para m, sea una combinación de $m+1$ puntos. Si $\lambda_{m+1}=1$ es trivial; si no, con $s = 1-\lambda_{m+1} > 0$ y $y = \sum_{i\le m}(\lambda_i/s)x_i$ se tiene, aplicando primero el caso $m=2$ y luego la hipótesis de inducción,

$$f(sy+(1-s)x_{m+1}) \le s\,f(y)+(1-s)f(x_{m+1}) \le s\sum_{i\le m}\frac{\lambda_i}{s}f(x_i)+\lambda_{m+1}f(x_{m+1}) = \sum_{i=1}^{m+1}\lambda_i f(x_i).$$

□

Ejemplo 15.13 (Jensen comprobado, y la desigualdad de las medias)**. (a) Con cifras.** Sea $f(x)=x^2$, que es convexa, y tomemos $x=(1,2,6)$ con pesos $(\frac12,\frac14,\frac14)$. La media ponderada es $\frac12\cdot1+\frac14\cdot2+\frac14\cdot6=\frac52=2{,}5$, y $f(2{,}5)=6{,}25$. La media ponderada de los valores es $\frac12\cdot1+\frac14\cdot4+\frac14\cdot36=\frac{21}{2}=10{,}5$. En efecto $6{,}25\le10{,}5$. La diferencia entre ambos lados es, en el lenguaje de la probabilidad, la varianza.

(b) La desigualdad de las medias sale de aquí. La función $-\ln$ es convexa en $x>0$, porque su segunda derivada es $1/x^2>0$. Jensen con pesos iguales $1/m$ da

$$-\ln\Big(\frac{x_1+\cdots+x_m}{m}\Big) \le \frac{-\ln x_1-\cdots-\ln x_m}{m},$$

y multiplicando por -1 y exponenciando queda $\frac{x_1+\cdots+x_m}{m} \ge \sqrt[m]{x_1\cdots x_m}$: la media aritmética domina a la geométrica. Comprobación con $1,4,16$: la aritmética es $21/3=7$ y la geométrica $\sqrt[3]{64}=4$, y en efecto $7\ge4$.

Los dos criterios de derivadas

Teorema 15.14 (Criterio de primer orden)**.** *Sea f derivable en un convexo abierto C. Entonces f es convexa si y solo si*

$$f(y)\ \ge\ f(x)+\nabla f(x)^{\mathsf T}(y-x) \qquad \textit{para todos } x,y\in C,$$

es decir, si y solo si la gráfica queda por encima de todos sus planos tangentes. En una variable esto es $f(y) \geq f(x) + f'(x)(y-x)$.

Demostración. Supongamos f convexa. Para $\lambda \in (0,1]$, escribiendo $x+\lambda(y-x) = \lambda y+(1-\lambda)x$, la definición da $f(x+\lambda(y-x)) \leq \lambda f(y) + (1-\lambda)f(x)$, y reordenando,

$$\frac{f(x+\lambda(y-x)) - f(x)}{\lambda} \;\leq\; f(y) - f(x).$$

Haciendo $\lambda \to 0^+$, el cociente de la izquierda tiende a la derivada direccional $\nabla f(x)^{\mathsf{T}}(y-x)$, y se obtiene la desigualdad buscada. Recíprocamente, si la desigualdad tangente vale siempre, sea $z = \lambda x + (1-\lambda)y$; aplicándola en z hacia x y hacia y,

$$f(x) \geq f(z) + \nabla f(z)^{\mathsf{T}}(x-z), \qquad f(y) \geq f(z) + \nabla f(z)^{\mathsf{T}}(y-z),$$

y multiplicando la primera por λ, la segunda por $1-\lambda$ y sumando, el término con el gradiente se anula porque $\lambda(x-z)+(1-\lambda)(y-z) = 0$, con lo que queda $\lambda f(x) + (1-\lambda) f(y) \geq f(z)$, que es la convexidad. □

Ejemplo 15.15 (El criterio de primer orden en dos casos)**. (a)** $f(x) = \mathrm{e}^x$ **en** $x = 0$**.** La tangente es $y \mapsto 1 + y$, y el criterio afirma $\mathrm{e}^y \geq 1 + y$ para todo y. Comprobación: en $y = -2$ da $0{,}135\ldots \geq -1$; en $y = -1$, $0{,}368\ldots \geq 0$; en $y = 0$, $1 \geq 1$ con igualdad, como debe ser en el punto de tangencia; en $y = 1$, $2{,}718\ldots \geq 2$; en $y = 2$, $7{,}389\ldots \geq 3$.

(b) $f(x) = x^2$ **en** $x = 1$**.** La tangente es $y \mapsto 1 + 2(y-1) = 2y - 1$, y el criterio afirma $y^2 \geq 2y - 1$, es decir $(y-1)^2 \geq 0$. En $y = -1$: $1 \geq -3$; en $y = 0$: $0 \geq -1$; en $y = 0{,}5$: $0{,}25 \geq 0$; en $y = 3$: $9 \geq 5$.

(c) Por qué esto importa. Tomando y arbitrario y x un punto con $\nabla f(x) = 0$, la desigualdad se convierte en $f(y) \geq f(x)$ para todo y: en una función convexa, gradiente nulo implica mínimo *global*, sin más comprobaciones. Ésta es la razón profunda de que la convexidad haga tratable la optimización, y es también el germen de la definición de subgradiente: cuando no hay derivada, se pide simplemente que exista algún vector g que cumpla la desigualdad tangente.

Teorema 15.16 (Criterio de segundo orden)**.** *Sea f de clase C^2 en un convexo abierto $C \subseteq \mathbb{R}^n$. Entonces*

1. *f es convexa si y solo si $\nabla^2 f(x) \succeq 0$ para todo $x \in C$;*
2. *si $\nabla^2 f(x) \succ 0$ para todo $x \in C$, entonces f es estrictamente convexa (el recíproco es falso).*

En una variable esto se reduce al criterio de la segunda derivada: f es convexa si y solo si $f'' \geq 0$ en todo el intervalo, y es estrictamente convexa si $f'' > 0$.

Idea de la demostración. Todo se reduce a una variable, porque f es convexa si y solo si lo es su restricción a cada recta, $\varphi(t) = f(x+tv)$, y la regla de la cadena da $\varphi''(t) = v^{\mathsf{T}} \nabla^2 f(x+tv)\, v$. Así que "la hessiana es semidefinida positiva en todas partes" equivale a "$\varphi'' \geq 0$ para toda recta". En una variable, si $f'' \geq 0$ entonces f' es no decreciente, y por el teorema del valor

medio, $f(y) - f(x) = f'(c)(y-x) \geq f'(x)(y-x)$ para algún c entre x e y —el signo funciona en los dos casos, $y > x$ e $y < x$—, que es exactamente el criterio de primer orden y por tanto la convexidad. Al revés, si f es convexa, f' es no decreciente por el mismo criterio y entonces $f'' \geq 0$. □

Regla: comprobar que una matriz simétrica es semidefinida positiva

1. **Calcula los autovalores.** Una matriz simétrica es semidefinida positiva si y solo si todos sus autovalores son ≥ 0, y definida positiva si y solo si todos son > 0. Es el criterio infalible, y en dimensión 2 sale de resolver una ecuación de segundo grado.
2. **Atajo en dimensión 2:** para $\begin{pmatrix} a & b \\ b & c \end{pmatrix}$ basta mirar la traza $a + c$ y el determinante $ac - b^2$, porque los autovalores suman la traza y multiplican el determinante. Es definida positiva si $a > 0$ y $ac - b^2 > 0$.
3. **Para *definida* positiva usa los menores principales dominantes:** los determinantes de las submatrices de las primeras k filas y columnas. Todos > 0 equivale a definida positiva (criterio de Sylvester).
4. **Para *semi*definida positiva ese atajo no vale** y hay que mirar *todos* los menores principales, no solo los dominantes: los determinantes de cualquier submatriz que use el mismo juego de filas y de columnas. El contraejemplo obligatorio es $\begin{pmatrix} 0 & 0 \\ 0 & -1 \end{pmatrix}$, cuyos menores dominantes son 0 y 0, ambos ≥ 0, y que sin embargo tiene autovalor -1; en efecto, con $v = (0, 1)$ sale $v^{\mathsf{T}} H v = -1 < 0$. El menor principal que la delata es el de la segunda fila y columna, que vale -1.
5. **Si aparece un menor negativo o un autovalor negativo, has terminado:** la función no es convexa, y el vector propio correspondiente indica la dirección en la que se curva hacia abajo.
6. **Truco de reconocimiento:** toda matriz de la forma $A^{\mathsf{T}} A$ es semidefinida positiva, porque $v^{\mathsf{T}} A^{\mathsf{T}} A v = \|Av\|^2 \geq 0$. Las hessianas de los problemas de mínimos cuadrados son de ese tipo, y por eso esos problemas son siempre convexos.

Ejemplo 15.17 (Cuatro hessianas, calculadas y clasificadas)**. (a)** $f(x, y) = x^2 + xy + y^2$**.** Las parciales segundas dan $\nabla^2 f = \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}$, constante. Su polinomio característico es $(2-\mu)^2 - 1 = 0$, de donde $\mu = 1$ y $\mu = 3$: ambos positivos, luego la hessiana es definida positiva y f es estrictamente convexa. Traza $4 = 1 + 3$ y determinante $3 = 1 \cdot 3$: cuadra.

(b) $f(x, y) = x^2 + 4xy + y^2$**.** Ahora $\nabla^2 f = \begin{pmatrix} 2 & 4 \\ 4 & 2 \end{pmatrix}$, con característica $(2 - \mu)^2 - 16 = 0$ y autovalores $\mu = 6$ y $\mu = -2$. Hay uno negativo, luego la matriz es indefinida y f *no* es convexa. El vector propio de -2 es $(1, -1)$, y siguiendo esa dirección se ve el fallo sin matrices: $f(t, -t) = t^2 - 4t^2 + t^2 = -2t^2$, de modo que $f(1, -1) = f(-1, 1) = -2$ mientras que en el

punto medio $f(0,0)=0$. La cuerda vale -2 y la gráfica 0: la gráfica está *por encima* de la cuerda, justo lo contrario de la convexidad.

(c) $f(x,y)=(x+y)^2$**.** Aquí $\nabla^2 f=\begin{pmatrix}2&2\\2&2\end{pmatrix}$, con determinante 0 y traza 4, luego autovalores 0 y 4. Semidefinida positiva pero no definida: f es convexa y *no* estrictamente convexa. Se ve en el resultado: el mínimo vale 0 y se alcanza en toda la recta $x+y=0$, no en un solo punto. La dirección de autovalor nulo, $(1,-1)$, es exactamente esa recta.

(d) Un caso 3×3**.** Para $f(x,y,z)=x^2+y^2+z^2-xy-yz$ la hessiana es

$$\nabla^2 f=\begin{pmatrix}2&-1&0\\-1&2&-1\\0&-1&2\end{pmatrix},$$

con menores principales dominantes 2, 3 y 4, todos positivos: definida positiva por Sylvester. Sus autovalores son $2-\sqrt{2}\approx 0{,}586$, 2 y $2+\sqrt{2}\approx 3{,}414$, los tres positivos, como debía ser. La función es estrictamente convexa.

Cuidado

Que $f''>0$ implique estrictamente convexa no se puede invertir. La función $f(x)=x^4$ es estrictamente convexa —entre dos puntos distintos la cuerda queda estrictamente por encima; por ejemplo con $x=-1$ e $y=1$ el punto medio de la gráfica vale 0 y el de la cuerda vale 1— y sin embargo $f''(x)=12x^2$ se anula en $x=0$. Definida positiva es condición suficiente para la convexidad estricta, no necesaria. Segundo aviso, en dirección contraria: $f(x)=x^3$ tiene $f''(x)=6x$, que es negativa para $x<0$, luego no es convexa en $\mathbb{R}$; sí lo es en $[0,\infty)$. El dominio forma parte del enunciado, y una función puede ser convexa en un trozo y no en otro.

La consecuencia que justifica todo el capítulo

Teorema 15.18 (Todo mínimo local es global)**.** *Sea f convexa sobre un convexo C. Si $x^\star\in C$ es un mínimo local —existe $r>0$ tal que $f(x^\star)\le f(x)$ para todo $x\in C$ con $\|x-x^\star\|<r$—, entonces es un mínimo global. Si además f es estrictamente convexa, el mínimo es único.*

Demostración. Sea $y\in C$ arbitrario con $f(y)<f(x^\star)$; llegaremos a contradicción. Para $\lambda\in(0,1)$ pequeño, el punto $z_\lambda=\lambda y+(1-\lambda)x^\star$ está en C por convexidad y cumple $\|z_\lambda-x^\star\|=\lambda\|y-x^\star\|<r$ sin más que tomar $\lambda<r/\|y-x^\star\|$. Por convexidad de f,

$$f(z_\lambda)\le \lambda f(y)+(1-\lambda)f(x^\star)<\lambda f(x^\star)+(1-\lambda)f(x^\star)=f(x^\star),$$

donde la desigualdad estricta usa $f(y)<f(x^\star)$ y $\lambda>0$. Pero z_λ está a distancia menor que r de $x^\star$, contra la minimalidad local. Luego no existe tal y y $x^\star$ es mínimo global.

Unicidad en el caso estricto: si $x^\star\neq y^\star$ fueran dos mínimos globales con el mismo valor v, el punto medio cumpliría $f(\tfrac12 x^\star+\tfrac12 y^\star)<\tfrac12 v+\tfrac12 v=v$, un valor menor que el mínimo, lo cual es absurdo. □

Ejemplo 15.19 (Un mínimo localizado y certificado)**.** Minimicemos $f(x,y) = x^2+xy+y^2-3x$ sobre todo $\mathbb{R}^2$. El gradiente es $\nabla f = (2x+y-3,\ x+2y)$, y anularlo da el sistema $2x+y=3$, $x+2y=0$; de la segunda ecuación $x=-2y$, y sustituyendo, $-4y+y=3$, luego $y=-1$ y $x=2$. La hessiana es la del apartado (a) del ejemplo anterior, $\begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}$, definida positiva en todo punto, así que f es estrictamente convexa. Por el teorema, el punto crítico $(2,-1)$ es el mínimo global y es el único. Su valor es

$$f(2,-1) = 4 + 2\cdot(-1) + 1 - 6 = 4 - 2 + 1 - 6 = -3.$$

Obsérvese lo que se ha hecho y lo que no: no se ha explorado el plano, no se ha comparado con otros candidatos y no se ha usado ningún algoritmo. Dos comprobaciones —gradiente nulo y hessiana semidefinida positiva— bastan para un certificado de optimalidad global. Sin convexidad, ese mismo punto crítico no probaría absolutamente nada, como muestra el ejemplo de x^3 que aparece más adelante al hablar de KKT.

Regla: aplicar el criterio de la segunda derivada paso a paso

1. **Comprueba que f es de clase C^2** en un abierto convexo. Si tiene esquinas —valores absolutos, máximos, normas—, este criterio no se aplica y hay que ir al catálogo de operaciones que preservan la convexidad o a los subgradientes.
2. **Calcula las parciales segundas y monta la hessiana**, que será simétrica.
3. **Decide el signo de la hessiana en *todo* el dominio**, no solo en el punto crítico: la convexidad es una propiedad global, y basta un punto con autovalor negativo para perderla.
4. **Si sale semidefinida positiva, f es convexa**, y a partir de ahí todo punto con gradiente nulo es mínimo global.
5. **Si sale definida positiva, además el mínimo es único**; si sale solo semidefinida, puede haber toda una recta o todo un subespacio de mínimos.
6. **Si es indefinida, busca la dirección mala**: el vector propio del autovalor negativo da una recta sobre la que la función es cóncava, y restringirse a ella produce el contraejemplo explícito que hace falta para afirmar que f no es convexa.

15.3. Algoritmos de descenso y desigualdades

Definición 15.20. El *descenso por gradiente* itera $\mathbf{x}_{k+1} = \mathbf{x}_k - \alpha\nabla f(\mathbf{x}_k)$ con paso $\alpha > 0$. El *método de Newton* usa además la curvatura, $\mathbf{x}_{k+1} = \mathbf{x}_k - H^{-1}\nabla f(\mathbf{x}_k)$. Para funciones convexas no derivables se usa el *subgradiente*: cualquier g con $f(\mathbf{y}) \geq f(\mathbf{x}) + g^{\mathsf{T}}(\mathbf{y}-\mathbf{x})$ para todo $\mathbf{y}$.

Idea clave

Qué dice esto en cristiano. El gradiente apunta hacia donde la función crece más deprisa, así que para minimizar se camina en sentido contrario. El tamaño del paso es todo el problema: si es pequeño la convergencia es lentísima, y si es grande el método oscila o diverge. Newton corrige eso usando la segunda derivada para estimar cuánto hay que avanzar, y por eso converge cuadráticamente cerca del óptimo, aunque cada paso cuesta mucho más. El subgradiente extiende la idea a funciones con esquinas, como el valor absoluto, que aparecen constantemente en cuanto se añade regularización.

Ejemplo 15.21 (Velocidades de convergencia, con números)**. (a) Descenso en un caso simple.** Para $f(x) = x^2$ con paso $\alpha = 0{,}1$, la iteración es $x_{k+1} = x_k - 0{,}2x_k = 0{,}8x_k$. Desde $x_0 = 1$ da 0,8; 0,64; 0,512; 0,4096; 0,32768; 0,262144. Es convergencia *lineal*: el error se multiplica por 0,8 en cada paso, y hacen falta unos treinta y un pasos para ganar tres cifras decimales.

(b) Newton, en comparación. Aplicado a $f(x) = x^2 - 2$ para hallar $\sqrt{2}$, la iteración es $x_{k+1} = \frac{x_k + 2/x_k}{2}$, que desde $x_0 = 1$ da 1,5; 1,4166666667; 1,4142156863; 1,4142135624. El número de cifras correctas se *duplica* en cada paso: cuatro iteraciones bastan para diez cifras.

(c) El paso importa más que el método. Con $\alpha = 1$ la iteración de (a) sería $x_{k+1} = -x_k$: oscila para siempre sin converger. Con $\alpha = 1{,}1$ diverge. La condición de convergencia es $\alpha < \frac{2}{L}$ con L la constante de Lipschitz del gradiente, aquí $L = 2$, luego $\alpha < 1$.

(d) Subgradiente. La función $|x|$ no es derivable en 0, y su subdiferencial ahí es todo el intervalo $[-1, 1]$: cualquier pendiente entre -1 y 1 da una recta que queda por debajo de la gráfica. La condición de optimalidad se convierte en $0 \in \partial f(\mathbf{x})$, que aquí se cumple en $x = 0$, el mínimo. Con esa generalización, la regularización ℓ^1 que produce soluciones dispersas queda dentro de la teoría.

Cuidado

[Jensen y la dirección de la desigualdad] Para f convexa vale $f(E[X]) \leq E[f(X)]$, y con f cóncava la desigualdad se invierte. Con $f(x) = x^2$ y X uniforme en $\{1, 3\}$: $f(2) = 4 \leq \frac{1+9}{2} = 5$. Confundir el sentido es el error más habitual, y tiene consecuencias: de ahí sale que $E[X^2] \geq (E[X])^2$, es decir, que la varianza sea no negativa. Segundo aviso: el descenso por gradiente converge al mínimo global *solo* si la función es convexa; en el caso general converge a un punto crítico cualquiera, y ni siquiera hay garantía de que sea un mínimo local, pues puede quedarse en un punto de silla, que en dimensión alta son muchísimo más abundantes que los mínimos.

15.4. Dualidad de Lagrange y las condiciones KKT

Todo problema de optimización con restricciones tiene un gemelo. El original se llama *primal* y pregunta cuál es el menor valor alcanzable; el gemelo, el *dual*, pregunta cuál es la mayor cota inferior demostrable. Bajo hipótesis de convexidad los dos coinciden, y esa coincidencia es lo que convierte la optimización en una teoría y no en una colección de trucos.

Definición 15.22. Para el problema $\min f(x)$ sujeto a $g_i(x) \le 0$ y $h_j(x) = 0$, la *función lagrangiana* es

$$L(x, \lambda, \nu) = f(x) + \sum_i \lambda_i g_i(x) + \sum_j \nu_j h_j(x), \qquad \lambda_i \ge 0,$$

y la *función dual* es $q(\lambda, \nu) = \inf_x L(x, \lambda, \nu)$. El *problema dual* consiste en maximizar q sobre $\lambda \ge 0$. Siempre se cumple la *dualidad débil* $q(\lambda, \nu) \le f(x)$ para todo par admisible; la diferencia entre el óptimo primal y el dual se llama *salto de dualidad*.

Idea clave

Qué dice esto en cristiano. Los multiplicadores λ_i son precios. Cada restricción prohíbe algo, y el multiplicador dice cuánto valdría poder saltársela un poquito: si $\lambda_i = 0$, la restricción no aprieta y relajarla no serviría de nada; si λ_i es grande, esa restricción es la que está costando dinero. La función dual toma esos precios y pregunta "con estos precios, ¿cuál es la mejor cota inferior que puedo demostrar?". La dualidad débil dice que ninguna cota se pasa nunca del óptimo verdadero, y la fuerte, que con convexidad la mejor cota *es* el óptimo: no hay hueco entre lo que se puede lograr y lo que se puede demostrar.

Teorema 15.23 (Karush–Kuhn–Tucker)**.** *Sea el problema anterior con f y las g_i convexas, las h_j afines, y supóngase que se cumple la condición de Slater (existe un punto estrictamente admisible). Entonces $x^\star$ es óptimo si y sólo si existen $\lambda^\star \ge 0$ y $\nu^\star$ tales que*

$$\nabla f(x^\star) + \sum_i \lambda_i^\star \nabla g_i(x^\star) + \sum_j \nu_j^\star \nabla h_j(x^\star) = 0, \qquad \lambda_i^\star g_i(x^\star) = 0 \ \text{ para todo } i,$$

además de la admisibilidad de $x^\star$. La segunda familia de igualdades se llama holgura complementaria.

Regla: Resolver con KKT en siete pasos

1. **Escribir el problema en forma estándar**: minimizar, con $g_i \le 0$ y $h_j = 0$. Un $\ge$ se convierte en $\le$ cambiando el signo; un máximo, en mínimo del opuesto.
2. **Comprobar la convexidad** de f y de las g_i, y que las h_j sean afines. Sin eso, KKT sólo da condiciones necesarias, no suficientes.
3. **Verificar Slater**: que exista un punto con todas las desigualdades estrictas. Es lo que garantiza que no haya salto de dualidad.
4. **Montar la lagrangiana** y anular su gradiente respecto de x.
5. **Aplicar la holgura complementaria** razonando por casos: para cada restricción, o bien $\lambda_i = 0$, o bien $g_i(x) = 0$. Con m restricciones hay 2^m combinaciones, pero casi todas se descartan de inmediato.
6. **Descartar los candidatos** que violen $\lambda_i \ge 0$ o la admisibilidad.

7. **Comprobar el valor dual**: si coincide con el primal, el óptimo está certificado y no hace falta ningún argumento adicional.

Ejemplo 15.24 (Cuatro problemas resueltos con multiplicadores)**. (a) Una restricción activa.** Minimizar $x^2 + y^2$ sujeto a $x + y \geq 2$. En forma estándar, $g = 2 - x - y \leq 0$. El gradiente de la lagrangiana da $2x = \lambda$ y $2y = \lambda$, luego $x = y$. Si $\lambda = 0$ saldría $x = y = 0$, que viola la restricción; luego la restricción está activa, $x + y = 2$, y el óptimo es $(1, 1)$ con $\lambda = 2$ y valor 2. Geométricamente: el punto de la recta más cercano al origen.

(b) Una restricción inactiva. El mismo objetivo sujeto a $x + y \geq -2$. Ahora el mínimo sin restricciones, $(0, 0)$, ya cumple la condición, así que $\lambda = 0$ y la restricción no pinta nada. La holgura complementaria es exactamente el mecanismo que detecta esto sin tener que probar.

(c) Programación lineal y su dual. Maximizar $3x + 2y$ sujeto a $x + y \leq 4$, $x + 3y \leq 6$, $x, y \geq 0$. El óptimo primal es $(4, 0)$ con valor 12. El dual es minimizar $4u + 6v$ sujeto a $u + v \geq 3$, $u + 3v \geq 2$, $u, v \geq 0$, y su óptimo es $(3, 0)$ con valor 12: coinciden, como garantiza la dualidad fuerte para programas lineales factibles. El precio $u = 3$ dice que aflojar la primera restricción en una unidad valdría tres unidades de beneficio; el precio $v = 0$, que la segunda restricción sobra.

(d) Máquinas de vectores soporte. El problema de clasificar dos nubes de puntos con el margen más ancho posible es un programa cuadrático convexo. Su dual sólo depende de los productos escalares entre pares de datos, y la holgura complementaria dice que $\lambda_i > 0$ únicamente para los puntos que tocan el margen: los *vectores soporte*. De ahí sale el nombre del método y también el truco del núcleo, que consiste en sustituir esos productos escalares por otros sin construir nunca el espacio donde viven.

Cuidado

[Contraejemplo: sin convexidad, KKT no basta] Para minimizar $f(x) = x^3$ sobre $\mathbb{R}$, la condición $f'(x) = 0$ se cumple en $x = 0$ y allí no hay mínimo: la función baja sin límite. En problemas no convexos, KKT identifica *candidatos*, igual que $f' = 0$ en el cálculo de una variable, y hace falta un argumento aparte para decidir cuáles son óptimos. Peor aún: puede haber mínimos locales que satisfacen KKT y no son globales, y ningún examen de los multiplicadores los distingue.

Cuidado

[Contraejemplo: sin Slater puede haber salto de dualidad] Considérese minimizar e^{-x} sujeto a $x^2/y \leq 0$ con $y > 0$. El conjunto admisible obliga a $x = 0$, luego el óptimo primal vale 1; y sin embargo la función dual es idénticamente 0, de modo que el salto de dualidad es 1. El problema es convexo y aun así falla, porque no existe ningún punto estrictamente admisible: la condición de Slater no es un tecnicismo decorativo. En programación lineal, en cambio, la dualidad fuerte vale sin necesidad de Slater siempre que alguno de los dos problemas sea factible.

Pausa para pensar

Kuhn y Tucker publicaron sus condiciones en 1951 y durante veinte años el resultado llevó sólo sus nombres. Después se descubrió que William Karush las había demostrado en 1939, en su tesis de máster en Chicago, y que nunca las publicó porque no le parecieron gran cosa. Hoy el teorema lleva las tres iniciales. La moraleja no es sólo sobre la justicia académica: la tesis de Karush estuvo doce años en una estantería mientras el mundo redescubría lo mismo, y eso ocurre más veces de las que la historia oficial de las matemáticas deja ver.

15.5. Funciones no suaves: subgradientes y métodos proximales

Casi todas las funciones convexas interesantes tienen esquinas. Las normas, los máximos, las pérdidas por tramos, cualquier cosa con un valor absoluto: ninguna es derivable en todas partes, y sin embargo todas son convexas. La teoría se reconstruye entera sustituyendo el gradiente por un conjunto.

Definición 15.25. Sea f convexa. Un vector g es un *subgradiente* de f en x si $f(y) \geq f(x) + g^{\top}(y - x)$ para todo y. El conjunto de todos ellos, $\partial f(x)$, es el *subdiferencial*. Es no vacío, convexo y compacto en el interior del dominio, y se reduce a $\{\nabla f(x)\}$ exactamente donde f es derivable.

Idea clave

Qué dice esto en cristiano. Un gradiente da la única recta tangente. En una esquina no hay una sola recta que toque sin cortar: hay todo un abanico de ellas. El subdiferencial es ese abanico. En $|x|$ y en el origen, cualquier pendiente entre -1 y 1 define una recta que pasa por el vértice y deja la gráfica por encima, de modo que $\partial|x|(0) = [-1, 1]$. Y la condición de mínimo se traduce sola: $x^\star$ minimiza f si y sólo si $0 \in \partial f(x^\star)$. Donde antes se pedía que la derivada valiera cero, ahora se pide que el cero esté entre las pendientes admisibles. Nótese que es una condición de pertenencia, no una ecuación, y que por eso los mínimos de funciones no suaves tienden a caer justo en las esquinas.

Definición 15.26. El *operador proximal* de f con paso $t > 0$ es

$$\operatorname{prox}_{tf}(v) = \arg\min_{x} \Big\{ f(x) + \frac{1}{2t}\|x - v\|^2 \Big\},$$

que siempre existe y es único si f es convexa cerrada. El *algoritmo proximal por gradiente* para minimizar $f + g$ con f suave y g no suave itera $x_{k+1} = \operatorname{prox}_{tg}(x_k - t\nabla f(x_k))$.

Ejemplo 15.27 (Subgradientes y proximales, calculados)**. (a) El valor absoluto.** $\partial|x| = \{\operatorname{sgn} x\}$ si $x \neq 0$ y $[-1, 1]$ si $x = 0$. Su operador proximal es el *umbral suave*: $\operatorname{prox}_{t|\cdot|}(v) = \operatorname{sgn}(v)\max(|v| - t, 0)$. Con $v = 3$ y $t = 1$ da 2; con $v = 0{,}5$ y $t = 1$ da exactamente 0.

(b) Por qué eso produce ceros. El umbral suave no encoge un poco los valores pequeños: los aniquila. Ésa es la razón matemática de que la regresión lasso, que minimiza

$\|Ax-b\|^2+\lambda\|x\|_1$, seleccione variables poniendo coeficientes exactamente a cero, mientras que la regresión de cresta, con $\|x\|_2^2$, los deja pequeños pero no nulos. La diferencia no es de grado: es la esquina de la norma ℓ^1 en los ejes.

(c) El máximo de dos funciones. Si $f=\max\{f_1,f_2\}$ y en x ambas coinciden, entonces $\partial f(x)$ es el segmento que une $\nabla f_1(x)$ con $\nabla f_2(x)$. En general, la envolvente convexa de los gradientes de las funciones activas.

(d) La proyección como caso particular. Si g es la función indicadora de un convexo C —cero dentro, infinito fuera— entonces prox_{tg} es exactamente la proyección sobre C, y el método proximal por gradiente se convierte en el gradiente proyectado. Optimizar con restricciones y optimizar con penalizaciones no suaves resultan ser el mismo algoritmo escrito de dos maneras.

Cuidado

[Contraejemplo: el subgradiente no siempre es dirección de descenso] Con el gradiente ordinario, moverse en dirección $-\nabla f$ siempre baja el valor si el paso es pequeño. Con subgradientes no: un paso demasiado grande cruza la esquina y sube. Por eso el método del subgradiente exige pasos que tiendan a cero y no garantiza descenso monótono: sólo converge en el mejor valor visto hasta el momento, y a velocidad $O(1/\sqrt{k})$, mucho peor que el $O(1/k)$ del caso suave.

Pausa para pensar

Moreau introdujo el operador proximal en 1962 estudiando mecánica de contactos, donde las fuerzas son discontinuas por naturaleza. Rockafellar lo convirtió en teoría en su *Convex Analysis* de 1970, un libro que durante veinte años fue considerado matemáticamente elegante y prácticamente irrelevante. Luego llegaron los problemas de gran escala en estadística y aprendizaje automático, con sus normas ℓ^1 por todas partes, y el análisis convexo no suave pasó de curiosidad a herramienta diaria en menos de una década.

15.6. Conos, programación semidefinida y relajaciones

Generalizar las desigualdades $x\geq 0$ a órdenes más ricos abre una familia de problemas que siguen siendo convexos, se resuelven en tiempo polinómico y capturan cosas que la programación lineal no alcanza. Es la frontera práctica de la optimización moderna.

Definición 15.28. Un *cono convexo* K es un convexo cerrado bajo multiplicación por escalares positivos. Un *programa cónico* es minimizar $c^\top x$ sujeto a $Ax=b$ y $x\in K$. Los tres casos importantes son el ortante positivo, que da la programación lineal; el cono de Lorentz $\{(t,u):\|u\|\leq t\}$, que da la programación cónica de segundo orden; y el cono de las matrices simétricas semidefinidas positivas $\mathbb{S}^n_+$, que da la *programación semidefinida*.

Idea clave

Qué dice esto en cristiano. En programación lineal, la restricción es que cada variable sea no negativa. En semidefinida, la variable es una matriz y la restricción es que sea semidefinida positiva, es decir, que todos sus autovalores lo sean. Es la misma idea —"estar en el lado bueno"— aplicada a un orden parcial mucho más rico, porque dos matrices simétricas no siempre son comparables. Lo notable es que el conjunto de las semidefinidas positivas es convexo, de modo que toda la teoría de dualidad sigue funcionando, y que existen algoritmos de punto interior que lo resuelven con la misma eficiencia práctica que un programa lineal grande.

Ejemplo 15.29 (Qué se puede escribir como un programa semidefinido)**. (a) Minimizar el mayor autovalor.** El problema $\min_x \lambda_{\max}(A_0 + \sum x_i A_i)$ se escribe como minimizar t sujeto a $tI - A_0 - \sum x_i A_i \succeq 0$. Un problema espectral, que parece intratable, se vuelve un programa convexo estándar.

(b) Cualquier cuadrática convexa. Por el complemento de Schur, $\|Ax - b\|^2 \le t$ equivale a que $\begin{pmatrix} I & Ax-b \\ (Ax-b)^\top & t \end{pmatrix} \succeq 0$. Los mínimos cuadrados son un caso particular.

(c) La relajación de maxcut. Partir los vértices de un grafo en dos grupos maximizando las aristas cortadas es NP-difícil. Sustituyendo cada variable ± 1 por un vector unitario y el producto $x_i x_j$ por $v_i^\top v_j$ se obtiene un programa semidefinido que sí se resuelve. Goemans y Williamson demostraron en 1995 que redondear la solución con un hiperplano aleatorio da un corte de valor al menos 0,8786 veces el óptimo, y esa constante es $\frac{2}{\pi}\min_{0<\theta\le\pi} \frac{\theta}{1-\cos\theta}$.

(d) El límite de la técnica. Bajo la conjetura de juegos únicos, ese 0,8786 es el mejor factor de aproximación alcanzable en tiempo polinómico para MAXCUT. No es que no se haya encontrado nada mejor: es que probablemente no lo hay. La relajación semidefinida no es un truco, es el algoritmo óptimo.

Cuidado

[Contraejemplo: la dualidad fuerte puede fallar en programas cónicos] En programación lineal, si el primal es factible y acotado, la dualidad fuerte se cumple sin más hipótesis. En semidefinida no. Existen programas semidefinidos con salto de dualidad estricto y otros donde el óptimo dual no se alcanza, y para descartarlo hace falta una condición de Slater: que exista un punto estrictamente factible, con la matriz definida positiva y no sólo semidefinida.

15.7. La conjugada de Fenchel y la dualidad de Fenchel–Rockafellar

Toda la dualidad que hemos usado —lagrangiana, precios sombra, holgura complementaria— es un caso particular de una operación única que actúa sobre funciones convexas y las envía a otras funciones convexas.

Definición 15.30 (Conjugada)**.** Para $f:\mathbb{R}^n\to\mathbb{R}\cup\{+\infty\}$,

$$f^*(y)=\sup_x\{\langle y,x\rangle - f(x)\}.$$

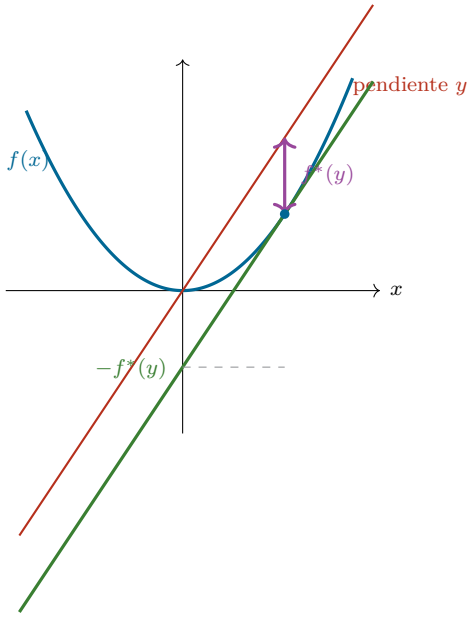


Figura 15.2: La conjugada $f^*(y)$ es la mayor separación vertical entre la recta de pendiente y que pasa por el origen y la gráfica de f; equivalentemente, $-f^*(y)$ es la ordenada en el origen de la recta soporte de esa pendiente. La función queda descrita por sus tangentes en lugar de por sus puntos, y la desigualdad $f(x)+f^*(y)\geq xy$ dice simplemente que la recta va por debajo.

La lectura geométrica es la que importa: $f^*(y)$ es la mayor distancia vertical entre el hiperplano de pendiente y que pasa por el origen y la gráfica de f; equivalentemente, $-f^*(y)$ es la ordenada en el origen del hiperplano de pendiente y que soporta a f desde abajo. La conjugada describe f *por sus tangentes* en lugar de por sus puntos: es la misma información codificada en el sistema dual.

Teorema 15.31 (Desigualdad de Fenchel–Young)**.** *Para todo x y todo y,*

$$f(x)+f^*(y)\ \geq\ \langle y,x\rangle,$$

con igualdad si y solo si $y\in\partial f(x)$.

La demostración es la definición leída al revés, y sin embargo casi todas las desigualdades clásicas caen de ella.

Ejemplo 15.32 (Cuatro conjugadas, y las desigualdades que salen)**. (a)** $f(x)=\frac12x^2$**.** $\sup_x(yx-\frac12x^2)$ se alcanza en $x=y$ y vale $\frac12y^2$. Luego $f^*=f$: la cuadrática es autoconjugada, la única función que lo es. Fenchel–Young dice $\frac12x^2+\frac12y^2\geq xy$, es decir $(x-y)^2\geq0$. Con $x=1{,}5$ e $y=0{,}8$: $1{,}445\geq1{,}2$.

(b) $f(x)=|x|$**.** Si $|y|\le 1$ el supremo de $yx-|x|$ es 0; si $|y|>1$ es $+\infty$. Luego $f^*=\iota_{[-1,1]}$, la indicadora del intervalo. Norma y bola son conjugadas la una de la otra, y en general $\|\cdot\|^*=\iota_{B^*}$ con B^* la bola dual. De ahí que la conjugada de $\|\cdot\|_1$ sea la indicadora de la bola de $\|\cdot\|_\infty$.

(c) $f(x)=e^x$**.** Derivando, $y=e^x$, luego $f^*(y)=y\log y-y$ para $y>0$ (y 0 en $y=0$, $+\infty$ si $y<0$). Fenchel–Young da $e^x+y\log y-y\ge xy$; con $x=2$, $y=3$: $7{,}389\,056+0{,}295\,837=7{,}684\,893\ge 6$.

(d) $f(x)=\frac{1}{p}|x|^p$ **con** $p>1$**.** Sale $f^*(y)=\frac{1}{q}|y|^q$ con $\frac{1}{p}+\frac{1}{q}=1$, y Fenchel–Young es exactamente la desigualdad de Young

$$xy\le \frac{|x|^p}{p}+\frac{|y|^q}{q},$$

de la que se deduce Hölder integrando. La desigualdad que sostiene la teoría de los espacios L^p no es más que la definición de conjugada.

Teorema 15.33 (Biconjugada, Fenchel–Moreau)**.** *$f^{**}\le f$ siempre, y $f^{**}=f$ si y solo si f es convexa, cerrada (semicontinua inferiormente) y propia.*

En general, f^{**} es la *envolvente convexa cerrada* de f: el mayor convexo que queda por debajo. Conjugar dos veces es la manera canónica de convexificar. Y para un conjunto C, la conjugada de su indicadora es la **función soporte** $\sigma_C(y)=\sup_{x\in C}\langle y,x\rangle$, que describe C por sus hiperplanos de apoyo: los convexos cerrados están en biyección con sus funciones soporte.

Teorema 15.34 (Dualidad de Fenchel–Rockafellar)**.** *Si f y g son convexas cerradas propias y A lineal,*

$$\inf_x\{f(x)+g(Ax)\}=\sup_y\{-f^*(-A^\top y)-g^*(y)\},$$

bajo una condición de cualificación (por ejemplo, que exista un punto donde f sea finita y g sea continua en Ax).

Ejemplo 15.35 (El lasso y su dual)**.** El problema $\min_x \frac{1}{2}\|Ax-b\|^2+\lambda\|x\|_1$ encaja en el esquema con $g(z)=\frac{1}{2}\|z-b\|^2$ y $f(x)=\lambda\|x\|_1$. Como f^* es la indicadora de $\{u:\|u\|_\infty\le\lambda\}$ y $g^*(y)=\frac{1}{2}\|y\|^2+\langle b,y\rangle$, el dual resulta ser

$$\max_y\ -\tfrac{1}{2}\|y\|^2-\langle b,y\rangle\quad\text{sujeto a}\quad\|A^\top y\|_\infty\le\lambda,$$

una proyección sobre un politopo. El primal tiene tantas variables como predictores y el dual tantas como observaciones: con diez mil genes y cincuenta pacientes, uno de los dos es mil veces más barato. Esa es la razón práctica de que la dualidad importe.

Idea clave

La dualidad lagrangiana que vimos antes es este teorema con g igual a la indicadora del conjunto factible. Lo que en aquel lenguaje eran multiplicadores, aquí son puntos del espacio dual; lo que era holgura complementaria, aquí es el caso de igualdad de

Fenchel–Young; y lo que era el salto de dualidad, aquí es la diferencia entre f y su biconjugada. Un solo mecanismo debajo de tres vocabularios.

15.8. Separación de convexos, lema de Farkas y teoremas minimax

Toda la dualidad descansa en un hecho geométrico que hasta ahora hemos usado sin enunciar: dos convexos disjuntos se pueden separar con un plano.

Teorema 15.36 (Proyección y separación estricta)**.** *Sea C convexo, cerrado y no vacío, y $x_0 \notin C$. Entonces existe un único $P_C(x_0) \in C$ a distancia mínima, y el vector $a = x_0 - P_C(x_0)$ separa estrictamente:*

$$\langle a, x\rangle \le \langle a, P_C(x_0)\rangle < \langle a, x_0\rangle \qquad \text{para todo } x \in C.$$

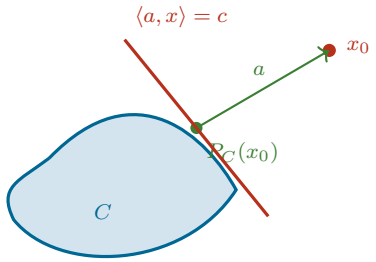


Figura 15.3: La proyección sobre un convexo cerrado existe, es única, y el vector que va de ella al punto exterior es la normal de un hiperplano que separa estrictamente. De este dibujo salen el teorema del hiperplano soporte, el lema de Farkas y, con ellos, toda la dualidad del capítulo.

Demostración. La existencia y unicidad de la proyección salen de que $\|x - x_0\|^2$ es estrictamente convexa y coerciva. Para la desigualdad, sea $p = P_C(x_0)$ y $x \in C$; como C es convexo, $p + t(x - p) \in C$ para $t \in [0, 1]$, y la función $t \mapsto \|p + t(x - p) - x_0\|^2$ tiene mínimo en $t = 0$. Derivando en 0: $\langle p - x_0, x - p\rangle \ge 0$, es decir $\langle a, x - p\rangle \le 0$. Y $\langle a, x_0 - p\rangle = \|a\|^2 > 0$. □

De ahí, pasando al límite, sale el **teorema del hiperplano soporte**: por cada punto de la frontera de un convexo pasa al menos un hiperplano que deja al convexo entero a un lado. Ese es el teorema que legitima hablar de subgradientes en todo punto y el que da sentido a la función soporte.

Teorema 15.37 (Lema de Farkas)**.** *Exactamente una de las dos alternativas se cumple:*

$$\textit{(i)}\ \exists\, x \ge 0 \ \textit{con}\ Ax = b, \qquad \textit{o} \qquad \textit{(ii)}\ \exists\, y \ \textit{con}\ A^{\top} y \ge 0 \ \textit{y}\ \langle b, y\rangle < 0.$$

La demostración es separación: el cono $K = \{Ax : x \ge 0\}$ es convexo y cerrado, y si $b \notin K$ hay un hiperplano que los separa; su normal es el y de la alternativa (ii).

Ejemplo 15.38 (Un certificado de infactibilidad, con números)**.** Sea

$$A = \begin{pmatrix} 1 & 1 & 1 \\ 2 & 2 & 2 \end{pmatrix}, \qquad b = \begin{pmatrix} 1 \\ 3 \end{pmatrix}.$$

El sistema $Ax = b$, $x \geq 0$ es infactible, y de hecho lo es sin la restricción de signo: la segunda fila es el doble de la primera, luego cualquier Ax cumple $(Ax)_2 = 2(Ax)_1$, y $3 \neq 2$.

El certificado de Farkas lo dice sin razonar: tómese $y = (2, -1)^\top$. Entonces

$$A^\top y = \begin{pmatrix} 1 & 2 \\ 1 & 2 \\ 1 & 2 \end{pmatrix} \begin{pmatrix} 2 \\ -1 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \geq 0, \qquad \langle b, y \rangle = 2 \cdot 1 - 1 \cdot 3 = -1 < 0.$$

Se cumple (ii), luego (i) es imposible. Y lo bonito es que *cualquiera* puede verificar el certificado en dos multiplicaciones, sin repetir la búsqueda. Farkas es a la infactibilidad lo que una factorización es a la no primalidad: convierte una afirmación negativa, difícil de comprobar por definición, en una afirmación positiva y barata.

Observación 15.39. La dualidad de la programación lineal se deduce de Farkas en media página, aplicándolo al sistema que codifica "existe un x factible con $c^\top x < z$". La alternativa que sale es precisamente una solución factible del dual con valor $\geq z$. Por eso el teorema de dualidad fuerte de la PL no necesita ninguna hipótesis de cualificación: en el caso poliédrico, Farkas la proporciona gratis.

Puntos de silla y minimax

Definición 15.40 (Punto de silla)**.** $(x^\star, \lambda^\star)$ es punto de silla de la lagrangiana L si

$$L(x^\star, \lambda) \leq L(x^\star, \lambda^\star) \leq L(x, \lambda^\star) \qquad \text{para todo } x, \lambda \geq 0.$$

Teorema 15.41. *$(x^\star, \lambda^\star)$ es punto de silla si y solo si $x^\star$ resuelve el primal, $\lambda^\star$ el dual, y no hay salto de dualidad. En problemas convexos con Slater, eso equivale a las condiciones KKT.*

Y ahí está la lectura definitiva de la dualidad: siempre se cumple $\sup_\lambda \inf_x L \leq \inf_x \sup_\lambda L$ —el que juega segundo tiene ventaja—, y la dualidad fuerte es exactamente la igualdad, es decir, la licencia para intercambiar el orden.

Ejemplo 15.42 (Un juego matricial resuelto de las dos maneras)**.** Dos jugadores; el pago que la fila recibe de la columna es

$$A = \begin{pmatrix} 3 & -1 \\ -2 & 4 \end{pmatrix}.$$

No hay punto de silla en estrategias puras: el máximo de los mínimos de fila es -1 y el mínimo de los máximos de columna es 3. Con estrategias mixtas, en cambio,

$$p^\star = (0{,}6;\ 0{,}4), \qquad q^\star = (0{,}5;\ 0{,}5), \qquad v = 1.$$

Comprobación: si la fila juega $p^\star$, su pago esperado es $0{,}6\cdot 3+0{,}4\cdot(-2)=1$ contra la primera columna y $0{,}6\cdot(-1)+0{,}4\cdot 4=1$ contra la segunda. Le da exactamente igual lo que haga el oponente: eso es lo que caracteriza a la estrategia óptima. Simétricamente, $q^\star$ deja a la fila indiferente entre sus dos opciones.

El teorema minimax de von Neumann (1928) garantiza que esto ocurre siempre: para todo juego matricial de suma cero, $\max_p \min_q p^\top A q = \min_q \max_p p^\top A q$. Es un caso particular de dualidad de la PL —cada jugador resuelve un programa lineal, y son duales el uno del otro—, y la generalización convexo-cóncava es el teorema de Sion. En piedra-papel-tijera la matriz es antisimétrica y el valor es 0, con la estrategia uniforme como única óptima.

15.9. Cuánto tardan de verdad: condicionamiento, tasas y la aceleración de Nesterov

Definición 15.43 (Suavidad y convexidad fuerte)**.** f es **L-suave** si ∇f es L-Lipschitz, y **μ-fuertemente convexa** si $f-\frac{\mu}{2}\|x\|^2$ es convexa. En el caso dos veces derivable, ambas cosas dicen $\mu I \preceq \nabla^2 f \preceq L I$. El cociente $\kappa = L/\mu$ es el **número de condición**.

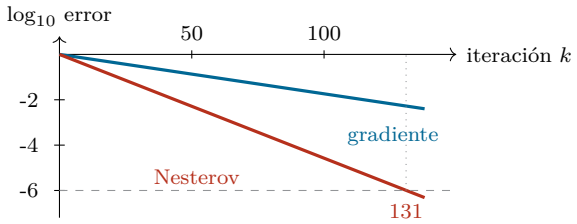


Figura 15.4: Con número de condición $\kappa=100$, el descenso de gradiente contrae el error por 0,9608 en cada paso y el método acelerado por 0,9. Para seis cifras hacen falta 345 iteraciones frente a 131. La ganancia es de orden $\sqrt{\kappa}$, así que crece justamente en los problemas mal condicionados.

Teorema 15.44 (Las cuatro tasas que hay que saber)**.** *Para el descenso de gradiente con paso $1/L$:*

convexa, no suave (subgradiente)	$f(x_k)-f^\star = O(1/\sqrt{k})$	*paso* $\propto 1/\sqrt{k}$
convexa y L-suave	$f(x_k)-f^\star \le \dfrac{2L\|x_0-x^\star\|^2}{k+4}$	$O(1/k)$
fuertemente convexa	$\|x_k-x^\star\|^2 \le \left(\dfrac{\kappa-1}{\kappa+1}\right)^{2k}\|x_0-x^\star\|^2$	*lineal*
acelerado (Nesterov)	$f(x_k)-f^\star = O(1/k^2)$ *y* $(1-1/\sqrt{\kappa})^k$	*óptimo*

El método acelerado consiste en añadir inercia:

$$y_k = x_k + \beta_k(x_k - x_{k-1}), \qquad x_{k+1} = y_k - \tfrac{1}{L}\nabla f(y_k),$$

con $\beta_k = \frac{\sqrt{\kappa}-1}{\sqrt{\kappa}+1}$ en el caso fuertemente convexo. Dos líneas de código más, y la raíz cuadrada del condicionamiento.

Ejemplo 15.45 (Una cuadrática mal condicionada, cronometrada)**.** Sea f cuadrática con $\kappa = 100$ (por ejemplo, autovalores del hessiano entre 1 y 100), y pidamos error relativo 10^{-6}.

(a) Gradiente. La razón de contracción por iteración es

$$\Big(\frac{99}{101}\Big)^2 = 0{,}960\,788,$$

luego hacen falta $\log(10^{-6})/\log(0{,}960\,788) = 345$ iteraciones.

(b) Acelerado. La razón es $1 - 1/\sqrt{100} = 0{,}9$, y hacen falta $\log(10^{-6})/\log(0{,}9) = 131$ iteraciones.

(c) La lectura. Un factor 2,6 con $\kappa = 100$. Con $\kappa = 10^4$ el gradiente necesitaría unas 34 500 iteraciones y el acelerado unas 1 375: un factor 25. La ganancia es de orden $\sqrt{\kappa}$, y por eso importa tanto justamente en los problemas difíciles.

Idea clave

La cota inferior de Nemirovski y Yudin cierra la historia: para cualquier método que solo use el gradiente y forme combinaciones lineales de los iterados, existe una función convexa L-suave donde el error no puede bajar de $\Omega(1/k^2)$. Nesterov alcanza esa cota, luego es *óptimo*: no hay ni habrá un método de primer orden esencialmente mejor. Ese es un resultado poco frecuente en optimización, donde casi todo es cuestión de heurísticas. La contrapartida es que el método acelerado no es monótono: $f(x_k)$ puede subir en algunas iteraciones. Quien programe un criterio de parada basado en "si el objetivo no baja, para" lo estropeará.

Observación 15.46 (A gran escala: estocástico y descomposición)*.* Con $n = 10^6$ datos, una sola iteración de gradiente completo cuesta n evaluaciones. El **gradiente estocástico** usa una sola muestra (o un mini-lote) por paso: cada paso es n veces más barato y apunta en la dirección correcta *en media*. La contrapartida es que su tasa es $O(1/\sqrt{k})$ en general y $O(1/k)$ con convexidad fuerte, siempre sublineal, porque el ruido del gradiente no desaparece. Los pasos deben cumplir las condiciones de Robbins–Monro, $\sum \alpha_k = \infty$ y $\sum \alpha_k^2 < \infty$ —por ejemplo $\alpha_k = 1/k$—, y promediar los iterados mejora la constante.

Cuando el objetivo es $f + g$ con ambos proximales fáciles pero la suma no, la respuesta es desdoblar: ADMM introduce una copia $z = x$ con un multiplicador y alterna un paso en cada variable. Es lo que permite resolver el lasso —donde el proximal de $\|\cdot\|_1$ es el umbral blando $\operatorname{sign}(x)\,\max(|x| - \lambda, 0)$, que anula coeficientes de un tajo— y lo que hace que el lasso produzca modelos ralos y no solo pequeños.

15.10. Cómo reconocer que algo es convexo sin derivar dos veces

El criterio de la hessiana exige que la función sea de clase C^2, y el capítulo trabaja constantemente con funciones que no lo son: valores absolutos, máximos, normas, la regularización ℓ^1. Hace falta un catálogo explícito.

Regla: Operaciones que preservan la convexidad

Si f y g son convexas, también lo son:

1. **Suma y múltiplo positivo:** $\alpha f + \beta g$ con $\alpha, \beta \geq 0$. (Con $\alpha < 0$ *no*: cambia el sentido.)
2. **Máximo puntual:** $\max\{f, g\}$, y en general el supremo de *cualquier* familia de convexas, $\sup_{i \in I} f_i$, incluso infinita. Ésta es la operación más útil del catálogo.
3. **Composición con una afín:** $x \mapsto f(Ax + b)$.
4. **Composición externa creciente:** $h \circ f$ si h es convexa y *no decreciente*. Por ejemplo e^f es convexa si f lo es; y f^2 lo es si además $f \geq 0$.
5. **Perspectiva:** $(x, t) \mapsto t\, f(x/t)$ para $t > 0$.
6. **Minimización parcial:** $g(x) = \inf_y f(x, y)$ es convexa si f es convexa *conjuntamente* en (x, y).
7. **Toda norma** es convexa, por la desigualdad triangular y la homogeneidad: $\|\lambda x + (1-\lambda)y\| \leq \lambda\|x\| + (1-\lambda)\|y\|$. En particular $|x|$, $\|x\|_1$, $\|x\|_2$ y $\|x\|_\infty$.
8. **Toda función afín** $a^\top x + b$ es a la vez convexa y cóncava.

Lo que NO preserva la convexidad: el producto de dos convexas (piénsese en $x \cdot x^{-1}$ sobre $x > 0$), el mínimo puntual, la composición con una función decreciente, y la composición $f \circ g$ con g no afín en general.

Ejemplo 15.47 (Cinco funciones no derivables, todas convexas)**.**

- $|x| = \max\{x, -x\}$: máximo de dos afines, luego convexa (punto 2).
- $\|x\|_1 = \sum_i |x_i|$: suma de convexas (puntos 1 y 2).
- $\max_i(a_i^\top x + b_i)$: máximo de afines; toda función convexa lineal a trozos es de esta forma.
- La *pérdida bisagra* $\max\{0,\, 1 - y\, a^\top x\}$ de las máquinas de vectores soporte: máximo de una constante y una afín.
- El *mayor autovalor* $\lambda_{\max}(X) = \sup_{\|v\|=1} v^\top X v$ de una matriz simétrica: supremo de funciones *lineales* en X, luego convexa aunque no sea derivable donde hay autovalores repetidos.

Ninguna de las cinco admite el criterio de la hessiana, y las cinco aparecen en los problemas del capítulo.

Regla: Comprobar convexidad en la práctica

Ante una función concreta, en este orden:

1. ¿Es afín? Entonces es convexa y cóncava a la vez, y el problema es lineal.

2. ¿Se construye a partir de piezas convexas conocidas mediante las operaciones de la regla anterior? Ése es el camino corto, y el que se usa el noventa por ciento de las veces.

3. Si es de clase C^2, calcula la hessiana y comprueba que es semidefinida positiva (todos los autovalores ≥ 0, o todos los menores principales ≥ 0).

4. Si nada de lo anterior sirve, vuelve a la definición: comprueba $f(\lambda x + (1-\lambda)y) \leq \lambda f(x) + (1-\lambda)f(y)$, o restringe a rectas —$f$ es convexa si y sólo si $t \mapsto f(x+tv)$ lo es para todos x, v—, que reduce el problema a una variable.

15.11. Ejercicios resueltos

Tres problemas que van de comprobar convexidad a certificar un óptimo con el dual.

Ejercicio 15.1. Decida si son convexas: $f(x) = x^4$, $g(x,y) = x^2 - y^2$, $h(x) = \max\{x, 0\}$ y $k(x) = -\ln x$ en $x > 0$. Enuncie el criterio que use en cada caso.

Solución. Para funciones de una variable dos veces derivables basta el signo de la segunda derivada. $f'' = 12x^2 \geq 0$, luego f es convexa. $k'' = 1/x^2 > 0$, luego $-\ln$ es convexa, y de ahí sale la desigualdad de las medias.

g no es convexa: su hessiano es $\operatorname{diag}(2, -2)$, indefinido. Es el ejemplo de silla, convexa en una dirección y cóncava en otra.

h no es derivable en 0 y aun así es convexa: el máximo de dos funciones convexas siempre lo es, porque el epigrafo de un máximo es la intersección de los epigrafos. Esto es importante en la práctica, porque las funciones convexas interesantes —normas, máximos, sumas de pérdidas— casi nunca son suaves, y la teoría se construye con subgradientes en lugar de gradientes. ■

Ejercicio 15.2. Minimice $x^2 + y^2$ sujeto a $2x + y = 10$ de tres maneras: sustituyendo, con multiplicadores y geométricamente. Compruebe que las tres dan lo mismo.

Solución. **Sustituyendo.** De $y = 10 - 2x$ sale $\varphi(x) = 5x^2 - 40x + 100$, con mínimo en $x = 4$, luego $y = 2$ y el valor es 20.

Con multiplicadores. $\nabla(x^2 + y^2) = \lambda \nabla(2x + y)$ da $2x = 2\lambda$ y $2y = \lambda$, luego $x = 2y$. Con la restricción, $y = 2$, $x = 4$, $\lambda = 4$.

Geométricamente. Se busca el punto de la recta más cercano al origen, que es el pie de la perpendicular. La distancia es $10/\sqrt{5}$, y su cuadrado es $100/5 = 20$.

Las tres coinciden, y la tercera explica por qué $\lambda = 4$: el multiplicador es la tasa a la que mejoraría el óptimo si se relajase la restricción, y en efecto cambiar el 10 por $10 + h$ multiplica el valor por $((10+h)/10)^2$, cuya derivada en $h = 0$ vale 4. ■

Ejercicio 15.3. Resuelva el programa lineal: maximizar $5x + 4y$ sujeto a $6x + 4y \leq 24$, $x + 2y \leq 6$, $x, y \geq 0$. Escriba el dual, resuélvalo y compruebe la dualidad fuerte y la holgura complementaria.

Solución. Los vértices de la región son $(0,0)$, $(4,0)$, $(3;1{,}5)$ y $(0,3)$, con valores 0, 20, 21 y 12. El óptimo es $(3;1{,}5)$ con valor 21.

El dual es minimizar $24u + 6v$ sujeto a $6u + v \geq 5$, $4u + 2v \geq 4$, $u, v \geq 0$. Su óptimo es $u = 0{,}75$ y $v = 0{,}5$, con valor $24 \cdot 0{,}75 + 6 \cdot 0{,}5 = 21$. Coinciden, como garantiza la dualidad fuerte.

La holgura complementaria se lee así: en el primal ambas restricciones están activas —$6\cdot 3+4\cdot 1{,}5 = 24$ y $3 + 2\cdot 1{,}5 = 6$—, y en consecuencia ambos precios duales son estrictamente positivos. Los precios dicen cuánto valdría una unidad más de cada recurso: la primera restricción vale 0,75 por unidad y la segunda 0,5. Si un proveedor ofreciese ampliar la primera capacidad por menos de 0,75 la unidad, convendría aceptar. ■

15.12. Ejercicios propuestos

15.1. ¿Es convexo el conjunto $\{(x,y) : x^2 + y^2 \le 4\}$? ¿Y $\{(x,y) : xy \ge 1,\ x > 0\}$? ¿Y la unión de dos discos disjuntos?

15.2. Demuestre que la intersección de convexos es convexa y dé un contraejemplo para la unión.

15.3. ¿Es convexa $f(x) = |x|$? ¿Y $f(x) = x^3$ en $\mathbb{R}$? ¿Y en $[0,\infty)$?

15.4. Calcule el subdiferencial de $|x|$ en $x = 0$.

15.5. Minimice $x^2 + 2y^2$ sujeto a $x + y = 3$.

15.6. Proyecte el punto $(3,4)$ sobre la bola unidad euclídea y sobre el semiplano $x \le 1$.

15.7. Escriba el dual de: minimizar $c^\top x$ sujeto a $Ax \ge b$, $x \ge 0$.

15.8. Para el programa lineal máximo $3x + 2y$ con $x + y \le 4$, $x \le 3$, $x, y \ge 0$: halle el óptimo y los precios sombra.

15.9. ¿Qué garantiza la condición de Slater y por qué no basta la convexidad sola?

15.10. Enuncie la holgura complementaria e interprete el caso $\lambda_i = 0$.

15.11. ¿Converge el descenso de gradiente con paso fijo para toda función convexa? ¿Qué hipótesis adicional hace falta y qué velocidad se obtiene?

15.12. Explique por qué en un problema convexo todo mínimo local es global, y dé un problema no convexo donde falle.

Parte III

Probabilidad y estadística

Capítulo 16

Variables aleatorias, distribuciones e inferencia

Lo que necesitas traer

Del capítulo 5: conteo, coeficientes binomiales y la regla de Laplace. Del capítulo 6: integrales (las variables continuas son integrales de densidades) y series (las discretas son sumas). Del capítulo 1: el lenguaje de conjuntos, que es el de los sucesos.

El capítulo 5 enseñó a calcular probabilidades contando casos, y el la estadística descriptiva enseñó a describir datos ya recogidos. Falta la pieza que une las dos mitades y que es, con diferencia, la más usada del mundo real: cómo se pasa de *una muestra* a *una afirmación sobre la población*, y con qué derecho. Ese salto es la **inferencia estadística**, y para darlo hace falta primero un objeto nuevo: la *variable aleatoria*. Con lo de la estadística descriptiva y la probabilidad elemental del capítulo 5 basta; el único cálculo que aparece —integrales para las variables continuas— es el de las reglas de derivación e integración, y quien no lo haya leído puede seguir el capítulo entero interpretando toda integral como un área.

Idea clave

Intuición primero. Toda la inferencia descansa en una sola observación, y conviene tenerla clara antes que ninguna fórmula:
Las muestras varían, pero varían de manera predecible.
Si mido la estatura media de 64 personas elegidas al azar y luego repito el experimento con otras 64, las dos medias no coincidirán. Pero la diferencia no es arbitraria: si la población tiene desviación típica σ, las medias muestrales se agolpan alrededor de la media verdadera con desviación típica $\sigma/\sqrt{n}$, cuatro veces menor si multiplico la muestra por dieciséis. Esa fórmula tan sencilla es la que convierte la estadística en ciencia: permite decir *cuánto* me puedo equivocar antes de haberme equivocado.

16.1. Espacio muestral, sucesos y la regla de Laplace

Antes de hablar de variables aleatorias conviene fijar el vocabulario básico con toda la precisión posible, porque casi todos los errores de probabilidad son en realidad errores de lectura del enunciado: no se ha identificado bien qué se está contando. Un cuarto de hora dedicado a escribir el espacio muestral ahorra muchas respuestas absurdas.

Definición 16.1. Un *experimento aleatorio* es aquel cuyo resultado no se puede predecir aunque se repita en idénticas condiciones. El conjunto de todos los resultados posibles es el *espacio muestral* E, y un *suceso* es cualquier subconjunto de E. Si E es finito y todos sus elementos son equiprobables, la *regla de Laplace* da

$$P(A) = \frac{\text{casos favorables a } A}{\text{casos posibles}} = \frac{|A|}{|E|}.$$

Para sucesos cualesquiera valen $P(\overline{A}) = 1 - P(A)$ y la *regla de la suma* $P(A \cup B) = P(A) + P(B) - P(A \cap B)$; los sucesos son *incompatibles* si $A \cap B = \varnothing$, en cuyo caso el último término desaparece.

Idea clave

El espacio muestral es una elección, no un dato del universo, y elegirlo bien es la mitad del problema. Al lanzar dos dados podemos describir el resultado como "la suma", y entonces $E = \{2, 3, \ldots, 12\}$ tiene once elementos, o como "el par ordenado de caras", y entonces E tiene treinta y seis. La diferencia crucial es que en el segundo caso todos los elementos son *equiprobables* y en el primero no: la suma 7 ocurre de seis maneras y la suma 2 de una sola. La regla de Laplace exige equiprobabilidad, y por eso el espacio bueno casi siempre es el más detallado, aquel en el que cada resultado es tan verosímil como cualquier otro. Cuando alguien dice que la probabilidad de que mañana llueva es un medio "porque o llueve o no llueve", está usando Laplace sobre un espacio que no es equiprobable.

Regla: montar y resolver un problema de probabilidad elemental

El guion siguiente resuelve la práctica totalidad de los ejercicios de este nivel:

1. **Describe el experimento en una frase** que empiece por un verbo: "lanzar dos dados y anotar las dos caras", "extraer dos bolas sin devolverlas".
2. **Escribe el espacio muestral de forma que sea equiprobable**, aunque tenga más elementos; cuenta cuántos son con las herramientas del capítulo 5.
3. **Traduce el suceso pedido a un subconjunto** y cuenta sus elementos, apoyándote en una tabla o en un diagrama si el enunciado es enrevesado.
4. **Aplica Laplace** y simplifica la fracción; deja el resultado también en decimal, porque es lo que permite juzgar si es razonable.

5. **Si aparece "al menos", pasa al suceso contrario**, que casi siempre es mucho más corto de contar.
6. **Si aparece "o", usa la regla de la suma** y no olvides restar la intersección; solo se puede omitir cuando los sucesos son incompatibles.
7. **Comprueba que el resultado está entre** 0 **y** 1 y que tiene sentido: una probabilidad de 1,2 o de $-0{,}3$ delata un error de conteo.

Ejemplo 16.2 (Dos dados, analizados a fondo)**. (a) El espacio bueno.** Lanzamos dos dados distinguibles y anotamos el par ordenado: hay $6 \cdot 6 = 36$ resultados equiprobables. La distribución de la suma es entonces

suma	2	3	4	5	6	7	8	9	10	11	12
casos	1	2	3	4	5	6	5	4	3	2	1

que suma 36 como debe. Así $P(\text{suma} = 7) = \frac{6}{36} = \frac{1}{6} \approx 0{,}1667$ y $P(\text{suma} = 2) = \frac{1}{36} \approx 0{,}0278$: son sucesos con probabilidades muy distintas, lo que confirma que el espacio $\{2, \dots, 12\}$ no es equiprobable.

(b) Sucesos con "al menos". $P(\text{al menos un 6})$ se cuenta mucho mejor por el contrario: no salir ningún seis tiene $5 \cdot 5 = 25$ casos, luego $P = 1 - \frac{25}{36} = \frac{11}{36} \approx 0{,}3056$. Contar directamente exige recordar que el par $(6, 6)$ no se debe contar dos veces, y ahí es donde suele fallar la cuenta directa.

(c) La regla de la suma con y sin intersección. $P(\text{suma} = 7) = \frac{6}{36}$ y $P(\text{doble}) = \frac{6}{36}$; como ninguna suma 7 es doble, son incompatibles y $P(\text{suma} = 7 \text{ o doble}) = \frac{12}{36} = \frac{1}{3}$. En cambio, si A es "suma par" (18 casos) y B es "suma mayor que 9" (6 casos), su intersección son las sumas 10 y 12, con $3 + 1 = 4$ casos, de modo que $P(A \cup B) = \frac{18+6-4}{36} = \frac{20}{36} = \frac{5}{9}$.

(d) Una baraja. En una baraja de 52 cartas, $P(\text{as}) = \frac{4}{52} = \frac{1}{13}$ y $P(\text{corazón}) = \frac{13}{52} = \frac{1}{4}$. Como el as de corazones es una sola carta, $P(\text{as o corazón}) = \frac{4+13-1}{52} = \frac{16}{52} = \frac{4}{13} \approx 0{,}3077$. Nótese además que $\frac{1}{13} \cdot \frac{1}{4} = \frac{1}{52} = P(\text{as y corazón})$: los dos sucesos son independientes, un hecho que no es casualidad sino consecuencia de que la baraja sea un rectángulo completo de cuatro palos por trece figuras.

Ejemplo 16.3 (Extracciones de una urna y la paradoja de los cumpleaños)**. (a) Sin reemplazamiento.** Una urna tiene 5 bolas rojas y 3 azules; sacamos dos sin devolver la primera. Entonces $P(\text{dos rojas}) = \frac{5}{8} \cdot \frac{4}{7} = \frac{20}{56} = \frac{5}{14} \approx 0{,}3571$ y $P(\text{dos azules}) = \frac{3}{8} \cdot \frac{2}{7} = \frac{6}{56} = \frac{3}{28} \approx 0{,}1071$. Lo que falta hasta 1 es $P(\text{una de cada}) = \frac{15}{28} \approx 0{,}5357$, que también se obtiene directamente como $2 \cdot \frac{5}{8} \cdot \frac{3}{7}$, con el factor 2 por los dos órdenes posibles.

(b) Con reemplazamiento. Si devolvemos la primera bola antes de sacar la segunda, las extracciones se vuelven independientes y $P(\text{dos rojas}) = \left(\frac{5}{8}\right)^2 = \frac{25}{64} \approx 0{,}3906$, algo mayor que antes. La razón es intuitiva: al no retirar la primera roja, la urna sigue teniendo la misma proporción de rojas en la segunda extracción.

(c) Una tabla de contingencia. En un instituto de 200 estudiantes, 120 estudian inglés, 60 estudian francés y 30 estudian los dos idiomas. Entonces $P(\text{inglés o francés}) = \frac{120+60-30}{200} =$

$\frac{150}{200} = 0{,}75$, y quienes no estudian ninguno son 50, es decir un 25 %. Colocar los cuatro números en una tabla de dos por dos (30 ambos, 90 solo inglés, 30 solo francés, 50 ninguno) hace que cualquier pregunta posterior se conteste leyendo una casilla.

(d) La paradoja de los cumpleaños. ¿Cuál es la probabilidad de que en un grupo de n personas al menos dos cumplan años el mismo día? Otra vez conviene el contrario: que todos los cumpleaños sean distintos tiene probabilidad $\frac{365}{365} \cdot \frac{364}{365} \cdots \frac{365-n+1}{365}$. Los valores son 0,1169 para $n = 10$, 0,4114 para $n = 20$, 0,5073 para $n = 23$, 0,7063 para $n = 30$ y 0,9704 para $n = 50$. Con solo veintitrés personas ya es más probable que haya coincidencia que que no la haya, un resultado que casi nadie adivina y que se explica porque lo que crece no es el número de personas sino el de *parejas*, que para $n = 23$ ya son 253.

Cuidado

No se puede aplicar Laplace sin comprobar la equiprobabilidad: el error de considerar $E = \{2, 3, \ldots, 12\}$ como equiprobable al lanzar dos dados llevaría a decir que la suma 7 y la suma 2 son igual de frecuentes, cosa que cualquier partida desmiente. Segundo aviso: en la regla de la suma hay que restar la intersección siempre que los sucesos puedan darse a la vez; sumar sin restar es el error más repetido del tema. Y tercero: "al menos uno" no es lo mismo que "exactamente uno", y "dos rojas" no es lo mismo que "la segunda roja". Conviene leer el enunciado dos veces y escribir el suceso con sus elementos antes de contar.

Pausa para pensar

La probabilidad nació de una consulta muy poco académica: en 1654 el caballero de Méré, jugador profesional, preguntó a Pascal por qué le salían las cuentas mal en cierta apuesta con dados. Pascal escribió a Fermat, y de aquella correspondencia de unos pocos meses salió el núcleo de la teoría, incluido el reparto justo de una apuesta interrumpida. Que una disciplina hoy central en física, medicina y economía empezara resolviendo dudas de casino es un recordatorio de que las matemáticas serias suelen brotar de problemas concretos. La formalización completa, con espacios de medida, tardó casi tres siglos más y llegó con Kolmogórov en 1933; se trata en el capítulo 31.

16.2. Los axiomas de Kolmogórov: qué es exactamente una probabilidad

La regla de Laplace de la sección anterior tiene un defecto de nacimiento: exige que el espacio muestral sea finito y que todos sus elementos sean igual de verosímiles. En cuanto el dado está cargado, o la moneda tiene más peso de un lado, o el experimento consiste en medir el tiempo que tarda en llegar un autobús —donde los resultados posibles son infinitos y ninguno tiene la más mínima probabilidad individual—, la fórmula «casos favorables entre casos posibles» se queda muda. Hace falta decir qué es una probabilidad *en general*, sin apoyarse en el recuento.

La respuesta que hoy usa todo el mundo la dio Andréi Kolmogórov en 1933 y consiste en

dar la vuelta al problema. En lugar de explicar de dónde salen las probabilidades, se enumeran las tres propiedades mínimas que cualquier asignación razonable debe cumplir, y se llama probabilidad a cualquier cosa que las cumpla. Todo lo demás —las reglas de la suma, del contrario, del producto, el teorema de Bayes— se *demuestra* a partir de ahí. La ventaja es enorme: un solo cuerpo de teoremas sirve para el dado honrado, para el dado cargado, para las alturas de una población y para la desintegración de un átomo.

Definición 16.4 (Álgebra y σ-álgebra de sucesos)**.** Sea E un espacio muestral. Una familia $\mathcal{A}$ de subconjuntos de E es un *álgebra de sucesos* si

(I) $E \in \mathcal{A}$;

(II) si $A \in \mathcal{A}$, entonces su complementario $\overline{A} = E \setminus A$ está en $\mathcal{A}$;

(III) si $A, B \in \mathcal{A}$, entonces $A \cup B \in \mathcal{A}$.

Si además la familia es cerrada para uniones *numerables* —es decir, si $A_1, A_2, A_3, \ldots \in \mathcal{A}$ implica $\bigcup_{n\geq 1} A_n \in \mathcal{A}$— se llama *$\sigma$-álgebra*. Los elementos de $\mathcal{A}$ son los *sucesos*, y son exactamente aquellos subconjuntos a los que se les va a poder asignar una probabilidad.

Idea clave

Por qué hace falta esta lista aparentemente burocrática: porque una probabilidad debe ser *compatible con el lenguaje*. Si tiene sentido preguntar por «sale par», tiene que tener sentido preguntar por «no sale par»; y si tienen sentido «sale par» y «sale mayor que cuatro», tiene que tenerlo «sale par o mayor que cuatro». Las tres condiciones son la traducción a conjuntos de las tres partículas del castellano: *todo*, *no* y *o*. La intersección no hace falta pedirla porque sale gratis por las leyes de De Morgan, $A \cap B = \overline{\overline{A} \cup \overline{B}}$, y lo mismo pasa con la diferencia $A \setminus B = A \cap \overline{B}$.

Ejemplo 16.5 (Tres álgebras sobre el mismo dado)**.** Sea $E = \{1, 2, 3, 4, 5, 6\}$ el lanzamiento de un dado.

(a) La grande. $\mathcal{A} = \mathcal{P}(E)$, todos los subconjuntos posibles. Son $2^6 = 64$ sucesos, desde $\varnothing$ hasta E. Es la elección habitual cuando E es finito: todo subconjunto es un suceso legítimo.

(b) La pequeña. $\mathcal{A} = \{\varnothing, E\}$. Cumple las tres condiciones y no sirve para nada, pero demuestra que la definición no obliga a incluir todos los subconjuntos.

(c) Una intermedia. Si solo interesa la paridad, basta con $\mathcal{A} = \{\varnothing, \{1, 3, 5\}, \{2, 4, 6\}, E\}$. Compruébese: el complementario de $\{1, 3, 5\}$ es $\{2, 4, 6\}$, que está; la unión de los dos es E, que está. Es un álgebra con cuatro elementos, y representa exactamente la información «solo sé si salió par o impar». Esta idea —que un álgebra codifica *cuánta información* se tiene— es la que en teoría de procesos estocásticos se convierte en las filtraciones.

Definición 16.6 (Axiomas de Kolmogórov)**.** Sea $\mathcal{A}$ una σ-álgebra sobre E. Una *probabilidad* es una función $P\colon \mathcal{A} \to \mathbb{R}$ que cumple los tres *axiomas de Kolmogórov*:

(K1) **No negatividad:** $P(A) \geq 0$ para todo $A \in \mathcal{A}$.

(K2) **Normalización:** $P(E) = 1$.

(K3) **Aditividad numerable:** si $A_1, A_2, \ldots$ son sucesos *incompatibles dos a dos* —es decir, $A_i \cap A_j = \varnothing$ siempre que $i \neq j$—, entonces

$$P\Big(\bigcup_{n\geq 1} A_n\Big) = \sum_{n\geq 1} P(A_n).$$

La terna $(E, \mathcal{A}, P)$ se llama *espacio probabilístico*.

Observación 16.7. Los tres axiomas son exactamente lo que uno esperaría de una *proporción*. El primero dice que no hay proporciones negativas; el segundo, que el todo vale uno; el tercero, que si se trocea un suceso en pedazos que no se solapan, las proporciones de los pedazos suman la del total. Lo único que no es evidente es por qué (K3) se pide para infinitas colecciones y no solo para dos sucesos: la razón es que sin la versión numerable no se pueden tratar preguntas como «¿cuál es la probabilidad de que la primera cara aparezca alguna vez?», que involucra infinitos casos disjuntos. Con solo la versión finita, la teoría se queda sin límites y sin variables continuas.

Teorema 16.8 (Propiedades que se deducen de los axiomas)**.** *Sea $(E, \mathcal{A}, P)$ un espacio probabilístico y sean A, B sucesos cualesquiera. Entonces*

(i) $P(\varnothing) = 0$*;*

(ii) aditividad finita*: si $A \cap B = \varnothing$, entonces $P(A \cup B) = P(A) + P(B)$;*

(iii) regla del contrario*: $P(\overline{A}) = 1 - P(A)$;*

(iv) monotonía*: si $A \subseteq B$, entonces $P(A) \leq P(B)$, y además $P(B \setminus A) = P(B) - P(A)$;*

(v) $0 \leq P(A) \leq 1$ *para todo suceso A;*

(vi) regla de la suma*: $P(A \cup B) = P(A) + P(B) - P(A \cap B)$;*

(vii) desigualdad de Boole*: $P\big(\bigcup_{n\geq 1} A_n\big) \leq \sum_{n\geq 1} P(A_n)$ para sucesos cualesquiera, sean o no incompatibles.*

Demostración. (i) Aplíquese (K3) a la sucesión $\varnothing, \varnothing, \varnothing, \ldots$, que es incompatible dos a dos porque $\varnothing \cap \varnothing = \varnothing$. Su unión es $\varnothing$, luego $P(\varnothing) = \sum_{n\geq 1} P(\varnothing)$. Un número no negativo que es igual a la suma de infinitas copias de sí mismo solo puede ser cero.

(ii) Aplíquese (K3) a $A, B, \varnothing, \varnothing, \ldots$ y úsese (i): los términos a partir del tercero no aportan nada.

(iii) A y $\overline{A}$ son incompatibles y su unión es E, luego por (ii) y (K2) $P(A)+P(\overline{A}) = P(E) = 1$.

(iv) Si $A \subseteq B$, entonces B se parte en los dos trozos disjuntos A y $B \setminus A$, así que $P(B) = P(A) + P(B \setminus A)$ por (ii). Como (K1) garantiza $P(B \setminus A) \geq 0$, resulta $P(B) \geq P(A)$; y despejando sale la segunda fórmula.

(v) La cota inferior es (K1); la superior sale de (iv) con $A \subseteq E$ y de (K2).

(vi) Pártase la unión en tres trozos disjuntos: $A \cup B = (A \setminus B) \cup (A \cap B) \cup (B \setminus A)$. Por (ii),

$$P(A \cup B) = P(A \setminus B) + P(A \cap B) + P(B \setminus A).$$

Por (iv), $P(A \setminus B) = P(A) - P(A \cap B)$ y $P(B \setminus A) = P(B) - P(A \cap B)$. Sustituyendo,

$$P(A \cup B) = P(A) - P(A \cap B) + P(A \cap B) + P(B) - P(A \cap B) = P(A) + P(B) - P(A \cap B).$$

El término que se resta es precisamente la zona que se ha contado dos veces al sumar $P(A)$ y $P(B)$.

(vii) Constrúyase la sucesión disjunta $B_1 = A_1$ y $B_n = A_n \setminus (A_1 \cup \cdots \cup A_{n-1})$ para $n \geq 2$. Los B_n son incompatibles dos a dos, su unión coincide con la de los A_n y $B_n \subseteq A_n$, luego por (K3) y (iv)

$$P\Big(\bigcup_n A_n\Big) = P\Big(\bigcup_n B_n\Big) = \sum_n P(B_n) \leq \sum_n P(A_n). \qquad \square$$

Observación 16.9. Merece la pena notar qué acaba de ocurrir. La regla del contrario y la regla de la suma, que en la sección anterior se enunciaron como recetas de conteo, no eran hechos sobre casos favorables: son *consecuencias lógicas* de los tres axiomas y valen igual para un dado cargado, para una variable continua o para cualquier otro modelo. Esa es toda la ganancia de axiomatizar.

Ejemplo 16.10 (Los axiomas en acción)**. (a) Laplace es un caso particular.** Si E es finito con N elementos y se define $P(A) = |A|/N$, los tres axiomas se cumplen: $|A|/N \geq 0$; $|E|/N = 1$; y si A y B son disjuntos, el cardinal de la unión es la suma de cardinales. Así que la regla de Laplace no compite con los axiomas, sino que es un ejemplo de ellos, el más sencillo posible.

(b) Un dado cargado. Supóngase $P(6) = 0{,}25$ y que las otras cinco caras son equiprobables: cada una vale $0{,}75/5 = 0{,}15$. Las probabilidades suman $5 \cdot 0{,}15 + 0{,}25 = 1$ ✓. Entonces $P(\text{par}) = 0{,}15 + 0{,}15 + 0{,}25 = 0{,}55$ y $P(\text{impar}) = 1 - 0{,}55 = 0{,}45$ por la regla del contrario. Laplace no sirve aquí y los axiomas sí: nada en ellos exige equiprobabilidad.

(c) La regla de la suma con tres sucesos. Se elige al azar un número entre 1 y 300. Sean A, B y C los sucesos «múltiplo de 2», «de 3» y «de 5». Contando: $|A| = 150$, $|B| = 100$, $|C| = 60$; los múltiplos de 6, de 10 y de 15 son $|A \cap B| = 50$, $|A \cap C| = 30$, $|B \cap C| = 20$; y los de 30 son $|A \cap B \cap C| = 10$. Aplicando dos veces el apartado (vi) del teorema 16.8 se obtiene la fórmula de *inclusión–exclusión*

$$P(A \cup B \cup C) = P(A) + P(B) + P(C) - P(A \cap B) - P(A \cap C) - P(B \cap C) + P(A \cap B \cap C),$$

que aquí da $\frac{150+100+60-50-30-20+10}{300} = \frac{220}{300} = \frac{11}{15} = 0{,}7333$. El recuento directo confirma que hay exactamente 220 números divisibles por 2, por 3 o por 5. La probabilidad de que el número no sea divisible por ninguno de los tres es, por el contrario, $\frac{80}{300} = 0{,}2667$.

(d) Boole, cuando no se sabe nada de las intersecciones. Un sistema tiene tres componentes que fallan con probabilidades 0,02, 0,03 y 0,05, y no se sabe si los fallos están relacionados. La probabilidad de que falle alguno es como mucho $0{,}02 + 0{,}03 + 0{,}05 = 0{,}10$. La cota es burda —vale 0,0966 si son independientes— pero se obtiene sin ninguna hipótesis, y en ingeniería de fiabilidad se usa exactamente así.

Regla: traducir un enunciado al álgebra de sucesos

Antes de calcular nada hay que convertir las frases en conjuntos. El diccionario completo es este:

1. «A **y** B» es $A \cap B$; «A **o** B» es $A \cup B$, siempre en sentido inclusivo (vale que ocurran los dos); «**no** A» es $\overline{A}$.
2. **«ninguno de los dos»** es $\overline{A} \cap \overline{B} = \overline{A \cup B}$, y **«no los dos a la vez»** es $\overline{A \cap B} = \overline{A} \cup \overline{B}$. Son las leyes de De Morgan y son cosas distintas: confundirlas es el error de traducción más frecuente.
3. **«al menos uno»** es $A \cup B$, y casi siempre conviene calcularlo por el contrario: $P(A \cup B) = 1 - P(\overline{A} \cap \overline{B})$.
4. **«exactamente uno»** es $(A \cap \overline{B}) \cup (\overline{A} \cap B)$, cuya probabilidad es $P(A) + P(B) - 2P(A \cap B)$.
5. «A **pero no** B» es $A \setminus B = A \cap \overline{B}$, con $P(A \setminus B) = P(A) - P(A \cap B)$.
6. **Comprueba el reparto:** los cuatro trozos $A \cap B$, $A \setminus B$, $B \setminus A$ y $\overline{A \cup B}$ son disjuntos y sus cuatro probabilidades deben sumar exactamente 1. Si no suman, hay un dato mal leído.
7. **Dibuja el diagrama de Venn y escribe un número en cada una de las cuatro regiones**, empezando siempre por la intersección. Con los cuatro números puestos, cualquier pregunta posterior se responde sumando casillas y sin fórmula ninguna.

Definición 16.11 (Probabilidad condicionada)**.** Si $P(B) > 0$, la *probabilidad de A condicionada a B* es

$$P(A \mid B) = \frac{P(A \cap B)}{P(B)}.$$

Teorema 16.12 (Condicionar produce otra probabilidad)**.** *Fijado un suceso B con $P(B) > 0$, la función $Q(A) = P(A \mid B)$ cumple los tres axiomas de Kolmogórov sobre la misma σ-álgebra.*

Demostración. (K1) $Q(A) = P(A \cap B)/P(B)$ es cociente de dos números no negativos con denominador positivo. (K2) $Q(E) = P(E \cap B)/P(B) = P(B)/P(B) = 1$. (K3) Si los A_n son incompatibles dos a dos, también lo son los $A_n \cap B$, y como $\left(\bigcup_n A_n\right) \cap B = \bigcup_n (A_n \cap B)$,

$$Q\Big(\bigcup_n A_n\Big) = \frac{1}{P(B)} \sum_n P(A_n \cap B) = \sum_n Q(A_n). \qquad \square$$

Idea clave

Este teorema, que parece un trámite, es lo que autoriza a manejar $P(\cdot \mid B)$ con total libertad: todas las propiedades del teorema 16.8 valen automáticamente para la probabilidad condicionada. En particular $P(\overline{A} \mid B) = 1 - P(A \mid B)$, que se usa

constantemente. Lo que *no* vale, y es el origen de innumerables errores, es «condicionar por el otro lado»: no hay ninguna regla que relacione $P(A \mid \overline{B})$ con $P(A \mid B)$, y $P(A \mid B) + P(A \mid \overline{B})$ no tiene por qué valer 1 ni ninguna otra cosa.

Teorema 16.13 (Producto, probabilidad total y Bayes)**.** *Sean A y B sucesos con probabilidad positiva, y sea $B_1, \dots, B_n$ una* partición *del espacio muestral, es decir, sucesos incompatibles dos a dos, de probabilidad positiva y con $B_1 \cup \cdots \cup B_n = E$. Entonces:*

(i) regla del producto*: $P(A \cap B) = P(B)\,P(A \mid B) = P(A)\,P(B \mid A)$;*

(ii) regla de la cadena*: $P(A_1 \cap \cdots \cap A_k) = P(A_1)P(A_2 \mid A_1)P(A_3 \mid A_1 \cap A_2)\cdots P(A_k \mid A_1 \cap \cdots \cap A_{k-1})$;*

(iii) teorema de la probabilidad total*: $P(A) = \sum_{k=1}^{n} P(B_k)\,P(A \mid B_k)$;*

(iv) teorema de Bayes*: para cada j,*

$$P(B_j \mid A) = \frac{P(B_j)\,P(A \mid B_j)}{\sum_{k=1}^{n} P(B_k)\,P(A \mid B_k)}.$$

Demostración. (i) Es la definición 16.11 despejada, escrita en los dos órdenes posibles; obsérvese que $A \cap B = B \cap A$, de donde la doble igualdad.

(ii) Por inducción sobre k. Para $k = 2$ es (i). Si vale para $k-1$, aplíquese (i) al par formado por el suceso $A_1 \cap \cdots \cap A_{k-1}$ y el suceso A_k.

(iii) Los sucesos $A \cap B_k$ son incompatibles dos a dos (porque los B_k lo son) y su unión es $A \cap E = A$. Por la aditividad finita y por (i),

$$P(A) = \sum_{k=1}^{n} P(A \cap B_k) = \sum_{k=1}^{n} P(B_k)P(A \mid B_k).$$

(iv) Por definición y por (i), $P(B_j \mid A) = \dfrac{P(A \cap B_j)}{P(A)} = \dfrac{P(B_j)P(A \mid B_j)}{P(A)}$, y basta sustituir el denominador por el valor que da (iii). □

Observación 16.14. La demostración deja ver la estructura: el numerador de Bayes es *uno* de los sumandos del denominador. Por eso la probabilidad a posteriori es siempre la fracción que ese camino representa dentro del total de caminos que llevan al mismo dato observado, y por eso nunca puede salir mayor que 1. La sección siguiente pone esta maquinaria a trabajar con diagramas de árbol y con un caso clínico completo; aquí interesa una segunda forma de escribirla, mucho más rápida cuando la partición tiene solo dos casos.

Regla: aplicar Bayes en forma de razón de verosimilitudes

Cuando hay dos hipótesis, D y $\overline{D}$, conviene trabajar con *ventajas* (*odds*) en lugar de con probabilidades. La ventaja de un suceso es $\text{odds}(D) = P(D)/P(\overline{D})$, y se recupera la probabilidad con $P = \text{odds}\,/(1 + \text{odds})$.

1. **Escribe la ventaja inicial** $\text{odds}_0 = \dfrac{P(D)}{1-P(D)}$. Con una prevalencia de uno entre 500, la ventaja inicial es $1/499$.
2. **Calcula la razón de verosimilitudes del dato observado**, $\text{LR} = \dfrac{P(\text{dato} \mid D)}{P(\text{dato} \mid \overline{D})}$. Para un positivo es $\text{LR}^+ = \dfrac{\text{sensibilidad}}{1-\text{especificidad}}$; para un negativo, $\text{LR}^- = \dfrac{1-\text{sensibilidad}}{\text{especificidad}}$.
3. **Multiplica**: $\text{odds}_1 = \text{odds}_0 \cdot \text{LR}$. Esta es toda la fórmula de Bayes: en ventajas, actualizar es multiplicar, y el denominador engorroso desaparece.
4. **Vuelve a probabilidad** con $P = \text{odds}_1/(1+\text{odds}_1)$.
5. **Encadena las pruebas independientes multiplicando sus** LR, sin volver a pasar por probabilidades intermedias. Dos pruebas con $\text{LR} = 30$ equivalen a una con $\text{LR} = 900$.
6. **Juzga la fuerza del dato solo por su** LR: por debajo de 5 apenas mueve la creencia, por encima de 20 la mueve mucho. La sensibilidad sola no informa de nada.
7. **Localiza la prevalencia de equilibrio**, la que hace que un positivo deje la probabilidad justo en un medio: es $p^* = 1/(1+\text{LR}^+)$. Por debajo de ella, la mayoría de los positivos serán falsos por muy bueno que sea el test.

Ejemplo 16.15 (Un test excelente para una enfermedad rara)**.** Un test tiene *sensibilidad* 0,98 (detecta al enfermo el 98 % de las veces) y *especificidad* 0,97 (deja tranquilo al sano el 97 % de las veces). La enfermedad afecta a uno de cada 500 habitantes, o sea $P(D) = 0{,}002$. Alguien da positivo. ¿Qué probabilidad tiene de estar enfermo?

(a) Por la fórmula de Bayes. Los datos son $P(D) = 0{,}002$, $P(+ \mid D) = 0{,}98$ y $P(+ \mid \overline{D}) = 1 - 0{,}97 = 0{,}03$. Por probabilidad total,

$$P(+) = 0{,}002 \cdot 0{,}98 + 0{,}998 \cdot 0{,}03 = 0{,}00196 + 0{,}02994 = 0{,}0319,$$

y por tanto

$$P(D \mid +) = \frac{0{,}00196}{0{,}0319} = 0{,}061442,$$

es decir, un 6,14 %. Dicho de otro modo: de cada dieciséis positivos, quince están sanos.

(b) Por ventajas, en dos multiplicaciones. La ventaja inicial es $\text{odds}_0 = 0{,}002/0{,}998 = 1/499 = 0{,}002004$ y la razón de verosimilitudes vale $\text{LR}^+ = 0{,}98/0{,}03 = 32{,}6667$. Multiplicando, $\text{odds}_1 = 0{,}065464$, y volviendo a probabilidad, $0{,}065464/1{,}065464 = 0{,}061442$: el mismo número, con la mitad de cuentas.

(c) Por qué ocurre, contado con personas. Tómense 100 000 habitantes. Enfermos hay 200 y sanos 99 800. El test acierta en $200 \cdot 0{,}98 = 196$ enfermos, pero marca por error a $99\,800 \cdot 0{,}03 = 2994$ sanos. Positivos totales: $196 + 2994 = 3190$, de los cuales solo 196 están

enfermos, o sea $196/3190 = 0{,}061442$. *Ahí está toda la explicación*: el 3 % de una multitud de casi cien mil sanos son casi tres mil personas, y el 98 % de un grupo de doscientos son menos de doscientas. Un porcentaje pequeño de un colectivo enorme aplasta a un porcentaje enorme de un colectivo pequeño. No es que el test sea malo; es que hay muy poca enfermedad que encontrar.

(d) El mismo test, cambiando solo la prevalencia. La fórmula $P(D \mid +) = \dfrac{0{,}98\,p}{0{,}98\,p + 0{,}03(1-p)}$ da

p	0,0001	0,001	0,002	0,01	0,05	0,10	0,50
$P(D \mid +)$	0,0033	0,0317	0,0614	0,2481	0,6323	0,7840	0,9703

El test no ha cambiado ni una décima: lo único que cambia es a quién se le aplica. La prevalencia de equilibrio es $p^* = 1/(1 + 32{,}6667) = 3/101 = 0{,}029703$, un 2,97 %; por debajo de esa cifra, la mayoría de los positivos son falsos. Esto es exactamente lo que separa un cribado poblacional de una prueba pedida por un médico ante síntomas: en el segundo caso la prevalencia de partida ya es alta, y el mismo test resulta concluyente.

(e) Repetir la prueba. Si se aplica un segundo test independiente y vuelve a salir positivo, se multiplica otra vez por LR^+: $\mathrm{odds}_2 = 0{,}065464 \cdot 32{,}6667 = 2{,}138499$, luego $P(D \mid ++) = 2{,}138499/3{,}138499 = 0{,}681376$, un 68,1 %. Con dos positivos la sospecha pasa de despreciable a mayoritaria, y por eso los protocolos de cribado siempre exigen confirmación.

(f) Y el negativo, que casi nunca se calcula. Con $\mathrm{LR}^- = 0{,}02/0{,}97 = 0{,}020619$, la ventaja pasa a $0{,}002004 \cdot 0{,}020619 = 0{,}0000413$, de modo que $P(D \mid -) = 0{,}0000413$: cuatro de cada cien mil. Un negativo en una enfermedad rara es una noticia excelente, y esa es la cara buena de la misma moneda.

Definición 16.16 (Independencia de sucesos)**.** Dos sucesos A y B son *independientes* si

$$P(A \cap B) = P(A)\,P(B),$$

lo que, cuando $P(B) > 0$, equivale a $P(A \mid B) = P(A)$: saber que ha ocurrido B no altera la probabilidad de A. Una familia $A_1, \dots, A_n$ es *independiente dos a dos* si cada par lo es, y es *mutuamente independiente* si para *todo* subconjunto de índices $\{i_1, \dots, i_k\}$ se cumple

$$P(A_{i_1} \cap \cdots \cap A_{i_k}) = P(A_{i_1}) \cdots P(A_{i_k}).$$

Para tres sucesos eso son cuatro igualdades, no tres.

Ejemplo 16.17 (Dos a dos no basta: el ejemplo de Bernstein)**.** Se lanza una moneda dos veces, con $E = \{cc, cx, xc, xx\}$ y los cuatro resultados equiprobables. Sean $A =$ «la primera es cara», $B =$ «la segunda es cara» y $C =$ «las dos coinciden». Cada uno tiene probabilidad $\frac{1}{2}$, y

$$P(A\cap B) = P(\{cc\}) = \tfrac{1}{4} = P(A)P(B), \quad P(A\cap C) = P(\{cc\}) = \tfrac{1}{4} = P(A)P(C), \quad P(B\cap C) = \tfrac{1}{4} = P(B)P(C$$

de modo que los tres sucesos son independientes dos a dos. Sin embargo

$$P(A \cap B \cap C) = P(\{cc\}) = \tfrac{1}{4} \neq \tfrac{1}{8} = P(A)P(B)P(C),$$

y por tanto *no* son mutuamente independientes. Y con razón: conocidos A y B, el suceso C queda completamente determinado. La moraleja es que la independencia de una familia no se comprueba mirando parejas; hay que verificar todas las intersecciones.

Cuidado

Incompatible e independiente son casi lo contrario, y se confunden siempre. Dos sucesos son *incompatibles* cuando no pueden ocurrir a la vez, $A \cap B = \varnothing$; son *independientes* cuando ocurrir uno no informa sobre el otro. Véase con un dado: sea $A =$ «sale par» $= \{2, 4, 6\}$, con $P(A) = \frac{1}{2}$, y $B =$ «sale 3», con $P(B) = \frac{1}{6}$. Son incompatibles, luego

$$P(A \cap B) = P(\varnothing) = 0, \qquad \text{mientras que} \qquad P(A)\,P(B) = \tfrac{1}{2} \cdot \tfrac{1}{6} = \tfrac{1}{12} = 0{,}0833 \neq 0.$$

No son independientes. Y la razón es transparente en cuanto se condiciona: $P(A \mid B) = \frac{0}{1/6} = 0$, frente a $P(A) = \frac{1}{2}$. Saber que salió el 3 no es que no informe sobre la paridad: informa del todo, porque garantiza que *no* es par. Ser incompatibles es la forma más extrema de dependencia que existe.

En general, si A y B son incompatibles *y* además independientes, entonces $P(A)P(B) = P(A \cap B) = 0$, así que uno de los dos tiene probabilidad nula. Es decir: salvo en el caso degenerado de un suceso imposible, **incompatibles implica dependientes**. Regla mnemotécnica: la incompatibilidad es una propiedad de los *conjuntos* (se ve en el diagrama de Venn, no se tocan); la independencia es una propiedad de los *números* (hay que multiplicar y comparar). Nunca se deduce una de la otra.

Pausa para pensar

Kolmogórov publicó los *Grundbegriffe der Wahrscheinlichkeitsrechnung* en 1933, un opúsculo de sesenta páginas que resolvió de golpe el sexto problema de Hilbert en su parte probabilística: axiomatizar el cálculo de probabilidades. La idea decisiva fue reconocer que una probabilidad no es más que una medida cuyo total vale 1, con lo que toda la teoría de la medida construida por Borel y Lebesgue entre 1898 y 1904 —véase el capítulo 31— quedaba disponible de inmediato. Antes de 1933 la probabilidad era una colección de recetas brillantes sin fundamento común; después, un capítulo del análisis. Vale la pena subrayar lo que eso costó: la aditividad numerable, que es el único axioma que no resulta obvio, es exactamente la bisagra que permite hablar de límites, de sucesiones infinitas de ensayos y de variables continuas, es decir, de todo lo que viene en el resto del capítulo.

16.3. Probabilidad condicionada, árboles y el teorema de Bayes

La mayoría de las preguntas interesantes no son "¿qué probabilidad tiene esto?" sino "¿qué probabilidad tiene esto *sabiendo* aquello?". Un diagnóstico médico, un control de calidad o un filtro de correo basura son exactamente eso: revisar una creencia inicial a la luz de una prueba. La herramienta es una fórmula de tres líneas que, sin embargo, produce resultados

que contradicen la intuición de casi todo el mundo.

Definición 16.18. Si $P(B) > 0$, la *probabilidad condicionada* de A dado B es

$$P(A \mid B) = \frac{P(A \cap B)}{P(B)},$$

de donde sale la *regla del producto* $P(A \cap B) = P(B)\,P(A \mid B)$. Dos sucesos son *independientes* si $P(A \cap B) = P(A)P(B)$, es decir si $P(A \mid B) = P(A)$. Si $B_1, \ldots, B_n$ forman una partición del espacio muestral, el *teorema de la probabilidad total* afirma que $P(A) = \sum_k P(B_k)P(A \mid B_k)$, y el *teorema de Bayes* invierte el condicionamiento:

$$P(B_j \mid A) = \frac{P(B_j)\,P(A \mid B_j)}{\sum_k P(B_k)\,P(A \mid B_k)}.$$

Idea clave

Condicionar es *cambiar de espacio muestral*. Cuando sabemos que ha ocurrido B, el mundo se encoge hasta B, y para volver a tener una probabilidad legítima (que sume 1) hay que renormalizar dividiendo entre $P(B)$: eso y no otra cosa es la fórmula. De ahí sale también la asimetría que tanta confusión causa: $P(A \mid B)$ y $P(B \mid A)$ son números distintos, porque se dividen entre denominadores distintos. La probabilidad de que un enfermo dé positivo es altísima; la probabilidad de que un positivo esté enfermo puede ser pequeñísima, y la diferencia la marca cuánta gente sana hay de partida. Bayes es simplemente el peaje que hay que pagar para pasar de una a la otra.

Regla: resolver un problema con árbol y con Bayes

Casi todo enunciado con la palabra "sabiendo", "si resulta que" o "dado que" se resuelve así:

1. **Identifica la partición**, es decir la primera pregunta que decide el caso: enfermo o sano, fábrica A o B, urna elegida.
2. **Dibuja el árbol con dos niveles**, poniendo en el primero las probabilidades de la partición y en el segundo las condicionadas.
3. **Comprueba que en cada nodo las ramas suman** 1; si no suman, hay un dato mal leído y todo lo que venga después será falso.
4. **Multiplica a lo largo de cada rama** para obtener la probabilidad de cada camino completo, aplicando la regla del producto.
5. **Suma los caminos que conducen al suceso observado**: eso es la probabilidad total, y es el denominador de Bayes.
6. **Divide el camino que interese entre esa suma** para obtener la probabilidad a posteriori.
7. **Contrasta el resultado con una población imaginaria** de 10 000 individuos; si los dos métodos no coinciden, hay un error en alguna parte.

Ejemplo 16.19 (El test médico, o por qué un positivo asusta menos de lo que parece)**. (a) Los datos.** Una enfermedad afecta al 1 % de la población (esa es la *prevalencia*). Un test detecta al enfermo el 99 % de las veces (*sensibilidad*) y da negativo en el sano el 95 % de las veces (*especificidad*). Con D el suceso "estar enfermo" y $+$ el suceso "dar positivo", los datos son $P(D) = 0{,}01$, $P(+ \mid D) = 0{,}99$ y $P(+ \mid \overline{D}) = 0{,}05$.

(b) Probabilidad total y Bayes. El positivo puede llegar por dos caminos:

$$P(+) = 0{,}01 \cdot 0{,}99 + 0{,}99 \cdot 0{,}05 = 0{,}0099 + 0{,}0495 = 0{,}0594.$$

Y entonces

$$P(D \mid +) = \frac{0{,}0099}{0{,}0594} = 0{,}1667.$$

Es decir: de cada seis personas que dan positivo, solo una está realmente enferma. El test no es malo; lo que ocurre es que hay muchísima más gente sana, y un 5 % de una multitud pesa más que un 99 % de un grupo diminuto.

(c) La comprobación con diez mil personas. De 10 000 individuos, 100 están enfermos y 9900 sanos. El test acierta en 99 de los enfermos (verdaderos positivos) y falla en 1. De los sanos, marca erróneamente a $9900 \cdot 0{,}05 = 495$ (falsos positivos) y deja correctamente en paz a 9405. Los positivos totales son $99 + 495 = 594$, y la proporción de enfermos entre ellos es $\frac{99}{594} = 0{,}1667$, idéntica a la anterior. Este recuento con personas concretas convence a mucha gente que la fórmula no convence.

(d) Qué pasa si la enfermedad es más común. Con una prevalencia del 10 % en vez del 1 %, la cuenta cambia radicalmente: $P(+) = 0{,}10 \cdot 0{,}99 + 0{,}90 \cdot 0{,}05 = 0{,}144$ y $P(D \mid +) = \frac{0{,}099}{0{,}144} = 0{,}6875$. El mismo test, con la misma sensibilidad y la misma especificidad, pasa de un valor predictivo del 17 a uno del 69 %. Por eso los cribados masivos sobre población general dan tantos sustos y por eso se repiten las pruebas: la segunda se aplica sobre un grupo con prevalencia ya muy alta.

Ejemplo 16.20 (Fábricas, urnas e independencia)**. (a) Dos líneas de producción.** Una fábrica produce el 60 % de las piezas en la línea A, con un 2 % de defectuosas, y el 40 % en la línea B, con un 5 %. La probabilidad de que una pieza al azar salga defectuosa es $P(\text{def}) = 0{,}6 \cdot 0{,}02 + 0{,}4 \cdot 0{,}05 = 0{,}012 + 0{,}020 = 0{,}032$. Si nos dan una pieza defectuosa, la probabilidad de que venga de A es $\frac{0{,}012}{0{,}032} = 0{,}375$ y de B es 0,625: aunque A fabrique más, B aporta la mayoría de los defectos.

(b) Tres urnas. Se elige al azar una de tres urnas, con composiciones $U_1 = (3 \text{ blancas}, 1 \text{ negra})$, $U_2 = (2, 2)$ y $U_3 = (1, 3)$, y se extrae una bola. Por probabilidad total, $P(\text{blanca}) = \frac{1}{3}\left(\frac{3}{4} + \frac{2}{4} + \frac{1}{4}\right) = \frac{1}{3} \cdot \frac{6}{4} = \frac{1}{2}$. Si la bola sale blanca, $P(U_1 \mid \text{blanca}) = \frac{\frac{1}{3} \cdot \frac{3}{4}}{\frac{1}{2}} = \frac{1}{2}$: observar una bola blanca eleva de $\frac{1}{3}$ a $\frac{1}{2}$ la sospecha de que la urna era la más rica en blancas. Eso es aprender de los datos, escrito en una línea.

(c) Independencia comprobada, no supuesta. Lanzamos una moneda tres veces. Sea $A =$ "la primera es cara" y $B =$ "la segunda es cara". Entonces $P(A) = P(B) = \frac{1}{2}$ y $P(A \cap B) = \frac{1}{4}$: independientes, como cabía esperar. Ahora sea $C =$ "las tres monedas salen iguales", con $P(C) = \frac{2}{8} = \frac{1}{4}$. Resulta que $P(A \cap C) = \frac{1}{8} = P(A)P(C)$, luego A y C *también* son independientes, cosa que la intuición no anticipa en absoluto. La independencia es una propiedad numérica que hay que verificar, no una impresión.

(d) El problema de Monty Hall. Hay tres puertas, un coche detrás de una y cabras detrás de las otras dos. Elegimos una puerta, el presentador (que sabe dónde está el coche) abre otra con una cabra y nos ofrece cambiar. Cambiar gana con probabilidad $\frac{2}{3}$ y quedarse con $\frac{1}{3}$. La clave es que la puerta que abre el presentador no es aleatoria: está condicionada por lo que él sabe, y por eso su acción transporta información. Nuestra elección inicial acierta con probabilidad $\frac{1}{3}$, y esa probabilidad no cambia por nada de lo que ocurra después; luego el resto, $\frac{2}{3}$, se concentra entero en la única puerta que queda.

Cuidado

Confundir $P(A \mid B)$ con $P(B \mid A)$ es el error más costoso de toda la estadística aplicada; en los tribunales tiene nombre propio, "falacia del fiscal", y ha provocado condenas injustas. Segundo aviso: incompatible no es lo mismo que independiente, sino casi lo contrario; si A y B son incompatibles con probabilidad no nula, saber que ocurrió B garantiza que no ocurrió A, lo cual es la máxima dependencia posible. Y tercero: al dibujar el árbol, las ramas que salen de un mismo nodo deben sumar exactamente 1; comprobarlo cuesta cinco segundos y detecta la mayoría de los errores de lectura del enunciado.

Pausa para pensar

Thomas Bayes fue un pastor presbiteriano inglés que no publicó su teorema: apareció en 1763, dos años después de su muerte, gracias a su amigo Richard Price. Laplace lo redescubrió por su cuenta hacia 1774 y le dio la forma general que hoy usamos, aplicándolo a problemas de mecánica celeste y de demografía. Durante buena parte del siglo XX el enfoque bayesiano estuvo marginado en favor de la estadística frecuentista, y solo la potencia de cálculo de los años noventa lo devolvió al centro. Hoy sostiene desde los filtros de correo basura hasta los sistemas de diagnóstico automático: una fórmula de tres líneas escrita por un clérigo aficionado que tardó doscientos años en imponerse.

16.4. Variables aleatorias discretas

Definición 16.21. Una **variable aleatoria** X es una función que asigna un número real a cada resultado de un experimento aleatorio. Se llama **discreta** si el conjunto de valores que puede tomar es finito o numerable.

Su **función de masa** o **distribución de probabilidad** es la tabla (o fórmula) que da $P(X = x_i) = p_i$ para cada valor posible. Siempre se cumple

$$p_i \geq 0 \qquad \text{y} \qquad \sum_i p_i = 1.$$

Observación 16.22. La palabra "variable" es un residuo histórico desafortunado: X no es una variable en el sentido del álgebra elemental, sino una *función*, en el sentido preciso de las funciones elementales. Lo que hace es traducir resultados —"sale cara", "el paciente mejora"— a números, que es lo único con lo que sabemos operar. En el capítulo 31 esta idea se formaliza del todo: una variable aleatoria es una función medible sobre un espacio de medida total 1,

y la probabilidad es literalmente la medida de un conjunto. Toda la teoría moderna de la probabilidad, desde Kolmogórov en 1933, es un capítulo de la teoría de la medida.

Ejemplo 16.23. Se lanza un dado y X es el resultado. Su distribución es $P(X=k)=1/6$ para $k=1,\dots,6$. Comprobación obligatoria: $6\cdot\frac{1}{6}=1$. ✓

Definición 16.24. La **esperanza** (o media, o valor esperado) de una variable aleatoria discreta es

$$\mu = E(X) = \sum_i x_i\, p_i,$$

y su **varianza** es

$$\sigma^2 = \operatorname{Var}(X) = E\big[(X-\mu)^2\big] = \sum_i (x_i-\mu)^2 p_i \;=\; E(X^2) - \mu^2.$$

La **desviación típica** es $\sigma = \sqrt{\operatorname{Var}(X)}$.

Idea clave

La esperanza es una media ponderada, no un pronóstico. La fórmula $\sum x_i p_i$ es exactamente la media de la estadística descriptiva con las frecuencias relativas sustituidas por probabilidades. De hecho es el mismo número al que tiende la media de los datos si repito el experimento muchísimas veces —eso es la *ley de los grandes números*, que se demuestra en el capítulo 29.
Lo que la esperanza *no* es: un valor que quepa esperar. La esperanza de un dado es 3,5, y el dado jamás saca 3,5. La esperanza del número de hijos por familia puede ser 1,8. Es un centro de gravedad, no una predicción.

Ejemplo 16.25. Para el dado:

$$E(X) = \frac{1+2+3+4+5+6}{6} = \frac{21}{6} = 3{,}5.$$

$$E(X^2) = \frac{1+4+9+16+25+36}{6} = \frac{91}{6}, \qquad \operatorname{Var}(X) = \frac{91}{6} - \left(\frac{7}{2}\right)^2 = \frac{182-147}{12} = \frac{35}{12} \approx 2{,}9167,$$

luego $\sigma = \sqrt{35/12} \approx 1{,}708$.

Regla: propiedades de esperanza y varianza

Para constantes a, b y variables aleatorias cualesquiera X, Y:

$$E(aX+b) = aE(X)+b, \qquad \operatorname{Var}(aX+b) = a^2 \operatorname{Var}(X),$$

$$E(X+Y) = E(X)+E(Y) \quad \text{(siempre)},$$

$$\operatorname{Var}(X+Y) = \operatorname{Var}(X)+\operatorname{Var}(Y) \quad \text{(solo si } X \text{ e } Y \text{ son independientes)}.$$

Cuidado

Las dos últimas líneas son distintas y confundirlas es un error clásico. *La esperanza siempre se suma*, aunque las variables estén relacionadísimas; la varianza solo se suma si son independientes. Y en $\mathrm{Var}(aX+b)$ el b desaparece —trasladar no dispersa— mientras que el a sale *al cuadrado*: si triplico todas las ganancias, la varianza se multiplica por nueve y la desviación típica por tres. Con $\mathrm{Var}(X)=4$ resulta $\mathrm{Var}(3X-2)=9\cdot 4=36$, no 12 ni 34.

Ejemplo 16.26 (Un juego, ¿justo o no?)**.** En la ruleta europea hay 37 casillas. Apostar 1 a un número paga 35 a 1: se gana 35 con probabilidad $1/37$ y se pierde 1 con probabilidad $36/37$. La ganancia esperada por jugada es

$$E(G)=35\cdot\frac{1}{37}+(-1)\cdot\frac{36}{37}=\frac{35-36}{37}=-\frac{1}{37}\approx -0{,}027.$$

Es decir, 2,7 céntimos perdidos por cada euro apostado, jugada tras jugada, sin excepción y sin sistema que lo evite. Un juego se llama **justo** si $E(G)=0$; ningún juego de casino lo es, y el margen de la casa es precisamente $|E(G)|$.

Pausa para pensar

Una aseguradora cubre un siniestro de 50 000 que ocurre con probabilidad 0,002. ¿Cuál es la prima mínima que no la lleva a la ruina a largo plazo?
La indemnización esperada es $0{,}002\cdot 50\,000=100$. Con una prima de 100 la compañía ni gana ni pierde *en media*; cobra más para cubrir gastos, beneficio y —sobre todo— la varianza, porque un año malo no se compensa con la esperanza sino con reservas. Toda la industria del seguro es esta cuenta, más la observación de que la varianza de la media de n pólizas independientes es n veces menor.

16.4.1. La distribución binomial, revisitada

El capítulo 5 obtuvo la distribución binomial contando: si se repiten n ensayos independientes con probabilidad de éxito p en cada uno, y X cuenta los éxitos,

$$P(X=k)=\binom{n}{k}p^k(1-p)^{n-k},\qquad k=0,1,\dots,n.$$

Ahora podemos añadir sus dos parámetros resumen.

Teorema 16.27. *Si $X\sim B(n,p)$ y $q=1-p$, entonces $E(X)=np$ y $\mathrm{Var}(X)=npq$.*

Demostración. Escribimos $X=X_1+\cdots+X_n$, donde X_i vale 1 si el ensayo i es éxito y 0 si no. Cada X_i tiene $E(X_i)=1\cdot p+0\cdot q=p$ y $E(X_i^2)=1^2\cdot p+0^2\cdot q=p$, luego $\mathrm{Var}(X_i)=p-p^2=pq$. Por la linealidad de la esperanza, $E(X)=np$; y como los ensayos son independientes, las varianzas se suman: $\mathrm{Var}(X)=npq$. □

Observación 16.28. Merece la pena detenerse en lo barato que ha salido. Calcular $\sum_k k\binom{n}{k}p^k q^{n-k}$ directamente es un ejercicio de manipulación penoso; descomponer X en indicadores lo convierte en una línea. Es un patrón que reaparece en toda la matemática: *el objeto complicado es una suma de objetos triviales*, exactamente como una función se descompone en senos en el capítulo 31 o un espacio vectorial en una base en el capítulo 7.

Ejemplo 16.29. Un examen tipo test tiene 10 preguntas. El alumno ha estudiado a medias: descarta con seguridad algunas opciones en cada pregunta y estima que acierta con probabilidad 0,3, igual en todas y sin que unas influyan en otras. Entonces el número de aciertos es $X \sim B(10; 0{,}3)$.

$$E(X) = 10 \cdot 0{,}3 = 3, \qquad \operatorname{Var}(X) = 10 \cdot 0{,}3 \cdot 0{,}7 = 2{,}1, \qquad \sigma \approx 1{,}449.$$

$$P(X = 3) = \binom{10}{3}(0{,}3)^3(0{,}7)^7 = 120 \cdot 0{,}027 \cdot 0{,}0823543 \approx 0{,}2668.$$

$$P(X \geq 1) = 1 - P(X = 0) = 1 - (0{,}7)^{10} \approx 1 - 0{,}0282 = 0{,}9718.$$

Obsérvese el truco de la última línea, que se usará constantemente: *"al menos uno" se calcula por el complementario*, nunca sumando diez términos.

16.5. La distribución de Poisson: contar sucesos raros

La binomial cuenta éxitos en un número fijo de intentos. Pero muchísimas situaciones reales no tienen "intentos": ¿cuántas llamadas entran en una centralita en una hora? ¿Cuántos coches pasan por un peaje en un minuto? ¿Cuántas erratas hay en una página? No hay un n que contar, solo un ritmo medio.

Definición 16.30. Una variable aleatoria X sigue una **distribución de Poisson** de parámetro $\lambda > 0$, y se escribe $X \sim P(\lambda)$, si toma los valores $0, 1, 2, \ldots$ con probabilidades

$$P(X = k) = \mathrm{e}^{-\lambda}\,\frac{\lambda^k}{k!}, \qquad k = 0, 1, 2, \ldots$$

Proposición 16.31. *Si $X \sim P(\lambda)$ entonces $E(X) = \lambda$ y $\operatorname{Var}(X) = \lambda$.*

Idea de la demostración. Usando la serie de la exponencial (las sucesiones y series elementales),

$$\sum_{k\geq 0} \mathrm{e}^{-\lambda}\frac{\lambda^k}{k!} = \mathrm{e}^{-\lambda}\mathrm{e}^{\lambda} = 1,$$

lo que confirma que es una distribución de probabilidad. Para la media, $E(X) = \sum_{k\geq 1} k\,\mathrm{e}^{-\lambda}\lambda^k/k! = \lambda\sum_{k\geq 1}\mathrm{e}^{-\lambda}\lambda^{k-1}/(k-1)! = \lambda$, porque la suma restante es de nuevo la distribución completa. La varianza sale igual calculando primero $E(X(X-1)) = \lambda^2$. □

Idea clave

Que la media y la varianza coincidan es la huella dactilar de Poisson. Si usted tiene datos de conteos y observa que la varianza muestral es muy parecida a la media, tiene delante un proceso de Poisson; si la varianza es mucho mayor, los sucesos vienen "a rachas" y hace falta otro modelo. Es un diagnóstico de un segundo.

Teorema 16.32 (Poisson como límite de la binomial)**.** *Si $n \to \infty$ y $p \to 0$ de modo que $np \to \lambda$, entonces*

$$\binom{n}{k} p^k (1-p)^{n-k} \longrightarrow \mathrm{e}^{-\lambda} \frac{\lambda^k}{k!}.$$

Ejemplo 16.33. Una fábrica produce tornillos con un 1 % de defectuosos. En una caja de 200, ¿cuál es la probabilidad de que no haya ninguno defectuoso? La respuesta exacta es binomial: $(0{,}99)^{200} = 0{,}1340$. La aproximación de Poisson con $\lambda = np = 2$ da $\mathrm{e}^{-2} = 0{,}1353$. Para exactamente uno: binomial 0,2707, Poisson 0,2707; son ya indistinguibles. **Regla práctica**: la aproximación es buena cuando $n \geq 50$ y $np \leq 5$.

Ejemplo 16.34. A una centralita llegan en promedio 3 llamadas por minuto. Entonces $X \sim P(3)$ y

$$P(X=0) = \mathrm{e}^{-3} = 0{,}0498, \quad P(X=2) = \mathrm{e}^{-3}\frac{9}{2} = 0{,}2240, \quad P(X \leq 2) = 0{,}0498{+}0{,}1494{+}0{,}2240 = 0{,}4232,$$

y por complementario $P(X \geq 1) = 1 - 0{,}0498 = 0{,}9502$.

Cuidado

λ va siempre referido a un *intervalo concreto*. Si llegan 3 llamadas por minuto y le preguntan por 5 minutos, el parámetro no es 3 sino $\lambda = 15$. Cambiar la unidad de tiempo sin cambiar λ es el error clásico de este tema.

Observación 16.35. El nombre viene de Siméon Denis Poisson (1837), pero la distribución se hizo famosa gracias a Ladislaus Bortkiewicz, que en 1898 la usó para modelar el número de soldados del ejército prusiano muertos por coces de caballo: 200 observaciones (10 cuerpos de ejército durante 20 años) con media 0,61, ajustadas por la fórmula con una precisión que asombró a sus contemporáneos. Es el ejemplo canónico de que un fenómeno sin ninguna estructura aparente obedece una ley exacta en cuanto se mira en conjunto.

16.6. Juegos de azar, esperanza de ganancia y la ley de los grandes números

La esperanza de una variable aleatoria se entiende mejor cuando lleva un signo de moneda delante. Este apartado la pone a trabajar en apuestas, seguros y loterías, y termina con el teorema que explica por qué la banca siempre gana aunque cada jugada por separado sea impredecible.

Definición 16.36. En un juego con premios $x_1, \dots, x_n$ y probabilidades $p_1, \dots, p_n$, la *esperanza de ganancia* del jugador es $E(X) = \sum_k x_k p_k$, contando como negativas las cantidades que se pagan. El juego es *justo* si $E(X) = 0$, favorable al jugador si $E(X) > 0$ y desfavorable si $E(X) < 0$. La *ley de los grandes números* afirma que si $X_1, X_2, \dots$ son repeticiones independientes de la misma variable con esperanza μ y varianza finita, entonces la media $\overline{X}_n = \frac{1}{n}\sum_{k=1}^{n} X_k$ se acerca a μ cuando n crece; más precisamente, la desviación típica de $\overline{X}_n$ es $\sigma/\sqrt{n}$.

Idea clave

La esperanza no predice lo que pasará en una jugada; predice el promedio de muchísimas jugadas. Un jugador que apuesta un euro a un número de la ruleta no puede perder nunca "dos céntimos y siete milésimas": o pierde el euro o gana treinta y cinco. Y sin embargo, a la larga, pierde justo esos 0,027 euros por jugada. Ahí está toda la asimetría del casino: el jugador juega unas pocas veces y para él manda la varianza, que es enorme y le permite soñar; la casa juega millones de veces y para ella manda la esperanza, que es pequeña, negativa para el cliente y absolutamente fiable. La ley de los grandes números es literalmente el modelo de negocio de un casino y de una aseguradora.

Regla: analizar un juego o un contrato de riesgo

Toda apuesta, seguro o sorteo se evalúa con el mismo procedimiento:

1. **Define la variable como ganancia neta**, restando ya lo que cuesta participar; así el juego justo se reconoce por $E(X) = 0$ sin más.
2. **Haz la tabla de valores y probabilidades** y comprueba que las probabilidades suman 1.
3. **Calcula** $E(X) = \sum x_k p_k$ y quédate con el signo antes que con el número: dice a quién favorece el juego.
4. **Calcula también la varianza**, $\mathrm{Var}(X) = E(X^2) - E(X)^2$, porque mide cuánto puede desviarse el resultado real del promedio.
5. **Multiplica la esperanza por el número de jugadas** para saber el resultado esperado a largo plazo; la esperanza es aditiva aunque las jugadas no sean independientes.
6. **Divide la desviación típica entre** $\sqrt{n}$ para saber cuánto se estabiliza la media al repetir.
7. **Interpreta en porcentaje sobre lo apostado**, que es como el sector expresa la ventaja de la casa y permite comparar juegos distintos.

Ejemplo 16.37 (Ruleta, dados y lotería)**. (a) La ruleta europea.** Hay 37 casillas (0 a 36). Apostar 1 a un número paga 35 a 1, de modo que la ganancia neta es $+35$ con probabilidad

$\frac{1}{37}$ y -1 con probabilidad $\frac{36}{37}$:

$$E(X) = 35 \cdot \frac{1}{37} - 1 \cdot \frac{36}{37} = \frac{35-36}{37} = -\frac{1}{37} = -0{,}027027.$$

Apostar a rojo (dieciocho casillas) da $E(X) = \frac{18}{37} - \frac{19}{37} = -\frac{1}{37}$: exactamente lo mismo. Todas las apuestas de la ruleta europea tienen la misma esperanza, y la ventaja de la casa es siempre $\frac{1}{37} = 2{,}7027\ \%$. En 100 jugadas de un euro se pierden, en promedio, 2,70 euros.

(b) Un dado y el precio justo. Un juego paga tantos euros como marque un dado. $E(X) = \frac{1+2+3+4+5+6}{6} = 3{,}5$, luego la entrada justa es de 3,50 euros. Si cobran 4, la ganancia neta esperada del jugador es $3{,}5 - 4 = -0{,}5$ euros por partida. La varianza es $\mathrm{Var}(X) = \frac{91}{6} - 3{,}5^2 = \frac{35}{12} = 2{,}9167$ y la desviación típica 1,7078: bastante mayor que el medio euro de desventaja, y por eso a corto plazo el jugador puede perfectamente ir ganando.

(c) Una lotería. Se venden dos millones de boletos a un euro y hay un único premio de un millón. La esperanza es $E(X) = 1\,000\,000 \cdot \frac{1}{2\,000\,000} - 1 = 0{,}5 - 1 = -0{,}5$ euros por boleto: el jugador pierde la mitad de lo que apuesta. Comprar diez boletos multiplica por diez tanto la probabilidad de ganar como la pérdida esperada, que pasa a ser de 5 euros.

(d) Un seguro visto desde los dos lados. Una aseguradora cobra 300 euros de prima por cubrir un siniestro de 20 000 euros que ocurre con probabilidad 0,01. Su gasto esperado por póliza es $0{,}01 \cdot 20\,000 = 200$, luego gana 100 euros esperados por póliza. Para el cliente la esperanza es negativa (-100 euros), y aun así contratar es razonable: nadie quiere exponerse a una pérdida de 20 000 euros aunque sea improbable. La diferencia entre ambas posturas no está en las matemáticas sino en que la utilidad del dinero no es lineal.

Ejemplo 16.38 (Por qué la banca siempre gana)**. (a) La media se estabiliza.** Al lanzar un dado, $\mu = 3{,}5$ y $\sigma = 1{,}7078$. La desviación típica de la media de n tiradas es $\sigma/\sqrt{n}$: vale 0,5401 para $n = 10$, 0,1708 para $n = 100$, 0,0540 para $n = 1000$ y 0,0171 para $n = 10\,000$. Cada vez que multiplicamos por cien el número de tiradas, la dispersión de la media se divide por diez. Eso, y nada más que eso, es la ley de los grandes números.

(b) Una cota explícita. La desigualdad de Chebyshev da $P(|\overline{X}_n - 3{,}5| \geq 0{,}1) \leq \frac{\sigma^2}{n \cdot 0{,}01}$, que vale 2,9167 para $n = 100$ (una cota inútil, mayor que 1), 0,2917 para $n = 1000$ y 0,0292 para $n = 10\,000$. La cota es muy generosa, pero tiene la virtud de no exigir ninguna hipótesis sobre la forma de la distribución.

(c) La aseguradora del ejemplo anterior. Con una sola póliza, la desviación típica de su gasto es $\sqrt{0{,}01 \cdot 0{,}99} \cdot 20\,000 = 1989{,}9749$ euros, veinte veces mayor que el beneficio esperado: un negocio suicida. Con 10 000 pólizas, el beneficio esperado es de un millón de euros y la desviación típica del gasto *medio* baja a $1989{,}9749/\sqrt{10\,000} = 19{,}8997$ euros por póliza. La empresa no ha reducido el riesgo de cada cliente; lo ha promediado sobre muchos, que es la única manera conocida de convertir la incertidumbre en una previsión.

(d) El jugador que quiere doblar su capital. Un jugador con 10 euros apuesta uno a rojo en cada tirada hasta llegar a 20 o arruinarse. Si el juego fuera justo, su probabilidad de éxito sería $\frac{10}{20} = 0{,}5$. Con la ventaja real de la casa ($p = \frac{18}{37}$), la fórmula de la ruina del jugador da

$$P(\text{éxito}) = \frac{1 - (q/p)^{10}}{1 - (q/p)^{20}} = 0{,}3680.$$

Una desventaja de apenas el 2,7 % por tirada convierte una apuesta al 50 % en una del 37 %. Y jugando indefinidamente contra una banca de capital ilimitado, la ruina llega con probabilidad 1.

Cuidado

La esperanza no es lo más probable ni tiene por qué ser un valor posible: la esperanza de un dado es 3,5, que ninguna cara muestra. Segundo aviso, y es el más importante de todo el capítulo: la ley de los grandes números *no* implica que después de muchas cruces "toque" una cara. La moneda no tiene memoria; lo que ocurre es que las desviaciones se diluyen al dividir entre n, no que se compensen con desviaciones opuestas. Creer lo contrario es la llamada falacia del jugador, y ha arruinado a mucha gente. Y tercero: duplicar la apuesta tras cada pérdida no convierte un juego desfavorable en favorable; solo cambia la forma de la distribución de pérdidas, haciendo muy probable ganar poco y poco probable perder muchísimo.

Pausa para pensar

La ley de los grandes números fue demostrada por Jacob Bernoulli y publicada en 1713 en su *Ars Conjectandi*, ocho años después de su muerte; él la llamó su "teorema áureo" y confesó haber tardado veinte años en encontrarla. Su enunciado responde a una pregunta muy concreta: si observo una proporción en una muestra, ¿en qué medida puedo fiarme de ella? Con esa pregunta nace la estadística inferencial entera. Poisson le puso el nombre actual en 1837, y Chebyshev dio en 1867 la demostración corta que hoy se estudia. La versión fuerte, mucho más delicada, es de Kolmogórov y ya requiere la teoría de la medida del capítulo 31.

16.7. Variables aleatorias continuas

Definición 16.39. Una variable aleatoria X es **continua** si existe una **función de densidad** $f(x) \geq 0$ tal que para todo intervalo

$$P(a \leq X \leq b) = \int_a^b f(x)\,\mathrm{d}x, \qquad \text{con} \qquad \int_{-\infty}^{\infty} f(x)\,\mathrm{d}x = 1.$$

Su esperanza y su varianza son

$$\mu = \int_{-\infty}^{\infty} x\, f(x)\,\mathrm{d}x, \qquad \sigma^2 = \int_{-\infty}^{\infty} (x-\mu)^2 f(x)\,\mathrm{d}x = E(X^2) - \mu^2.$$

Idea clave

Probabilidad es área. Todo el paso de lo discreto a lo continuo se resume en un cambio de símbolo: donde había $\sum p_i$ ahora hay $\int f(x)\,\mathrm{d}x$. La densidad $f(x)$ *no* es una probabilidad —puede valer 3 perfectamente—; lo que es probabilidad es el *área* bajo ella. Por eso la condición "las probabilidades suman 1" se convierte en "el área total es 1".

Cuidado

En una variable continua, $P(X = a) = 0$ para todo valor concreto a, porque un segmento de anchura nula no encierra área. Esto no significa que $X = a$ sea imposible: significa que la probabilidad de acertar un número exacto entre infinitos es cero. Consecuencia práctica cómoda: en el caso continuo da igual poner $<$ o $\leq$,

$$P(a < X < b) = P(a \leq X \leq b).$$

En el caso discreto esto es rotundamente falso: $P(X < 3)$ y $P(X \leq 3)$ difieren en $P(X = 3)$.

Ejemplo 16.40. Sea $f(x) = 3x^2$ en $[0, 1]$ y 0 fuera. Es densidad porque $\int_0^1 3x^2 \mathrm{d}x = [x^3]_0^1 = 1$. ✓

$$P(X \leq \tfrac{1}{2}) = \int_0^{1/2} 3x^2 \mathrm{d}x = \left(\tfrac{1}{2}\right)^3 = \tfrac{1}{8} = 0{,}125.$$

$$\mu = \int_0^1 x \cdot 3x^2 \mathrm{d}x = \left[\tfrac{3x^4}{4}\right]_0^1 = \tfrac{3}{4}, \qquad E(X^2) = \int_0^1 3x^4 \mathrm{d}x = \tfrac{3}{5},$$

$$\sigma^2 = \tfrac{3}{5} - \tfrac{9}{16} = \tfrac{48-45}{80} = \tfrac{3}{80} = 0{,}0375, \qquad \sigma \approx 0{,}1936.$$

La densidad crece hacia la derecha, y en efecto la media 0,75 está bastante por encima del centro 0,5 del intervalo: la fórmula confirma lo que dice el dibujo.

16.8. La distribución exponencial y la falta de memoria

Poisson cuenta *cuántos* sucesos ocurren; su compañera continua responde a la otra pregunta natural: *cuánto hay que esperar* hasta el siguiente.

Definición 16.41. X sigue una **distribución exponencial** de parámetro $\lambda > 0$ si su función de densidad es

$$f(x) = \lambda \, \mathrm{e}^{-\lambda x} \quad \text{para } x \geq 0, \qquad f(x) = 0 \text{ si } x < 0.$$

Su función de distribución es $F(x) = P(X \leq x) = 1 - \mathrm{e}^{-\lambda x}$, su media es $1/\lambda$ y su varianza $1/\lambda^2$.

Ejemplo 16.42. Si a la centralita anterior llegan 3 llamadas por minuto, el tiempo de espera hasta la siguiente llamada es exponencial con $\lambda = 3$ (por minuto) y media $1/3$ de minuto, es decir 20 segundos. Con otro ejemplo: si un autobús pasa cada 10 minutos en promedio, $\lambda = 0{,}1$ y

$$P(X > 15) = \mathrm{e}^{-1{,}5} = 0{,}2231, \qquad P(X < 5) = 1 - \mathrm{e}^{-0{,}5} = 0{,}3935.$$

La mediana es $\ln 2/\lambda = 6{,}93$ minutos: **menor que la media**, porque la distribución está sesgada a la derecha. Más de la mitad de las esperas son cortas y unas pocas son larguísimas, y esas pocas tiran de la media hacia arriba.

Teorema 16.43 (Falta de memoria)**.** *Si X es exponencial, entonces para todos $s, t > 0$*

$$P(X > s + t \mid X > s) = P(X > t).$$

Y la exponencial es la única *distribución continua con esta propiedad.*

Demostración. Por la definición de probabilidad condicionada,

$$P(X > s + t \mid X > s) = \frac{P(X > s + t)}{P(X > s)} = \frac{\mathrm{e}^{-\lambda(s+t)}}{\mathrm{e}^{-\lambda s}} = \mathrm{e}^{-\lambda t} = P(X > t).$$

□

Ejemplo 16.44 (La falta de memoria, con números)**. (a) Una bombilla.** Si la duración es exponencial de media 1000 horas, la probabilidad de que dure más de 500 es $\mathrm{e}^{-0,5} = 0{,}606531$. Y si ya lleva 800 horas encendida, la probabilidad de que aguante 500 más sigue siendo 0,606531: la bombilla no envejece.

(b) La única continua. La exponencial es la única distribución continua sin memoria, y la geométrica la única discreta. La demostración reduce la propiedad a la ecuación funcional $g(s + t) = g(s)g(t)$, cuyas soluciones medibles son las exponenciales.

(c) Cuándo el modelo miente. Las personas, los motores y los rodamientos sí envejecen: su tasa de fallo crece con el tiempo. Para ellos se usa la distribución de Weibull, que tiene un parámetro de forma k; con $k = 1$ se recupera la exponencial, con $k > 1$ hay desgaste y con $k < 1$ hay mortalidad infantil.

(d) La falacia del jugador. La misma propiedad, en versión discreta, dice que tras diez caras seguidas la probabilidad de cara sigue siendo $1/2$. La moneda no "debe" cruces. Lo que sí es cierto es que la *proporción* de caras tiende a $1/2$, pero por dilución de la ventaja, no por compensación.

Idea clave

Traducido: **el autobús no sabe cuánto lleva usted esperando**. Si lleva 10 minutos en la parada, la probabilidad de esperar 15 más es la misma que tenía al llegar. Esto choca de frente con la intuición —"ya me toca"— y es rigurosamente cierto para procesos sin desgaste: la desintegración de un átomo radiactivo, la llegada de un correo electrónico, el fallo de un componente electrónico de estado sólido. No lo es, en cambio, para cosas que envejecen: la vida de una persona, el desgaste de un neumático o un autobús con horario. Cuando alguien comete la falacia del jugador ("lleva cinco rojos, ahora toca negro"), está negándole a la ruleta una propiedad que sí tiene.

Observación 16.45. La pareja Poisson–exponencial es la misma moneda por sus dos caras: si los sucesos llegan según un proceso de Poisson de ritmo λ, el número de sucesos en un intervalo de longitud t es $P(\lambda t)$ y el tiempo entre sucesos consecutivos es exponencial de parámetro λ. Comprobarlo es inmediato: no haber tenido ningún suceso en $[0, t]$ es el caso $k = 0$ de Poisson, con probabilidad $\mathrm{e}^{-\lambda t}$, que es exactamente $P(X > t)$ para la exponencial.

16.9. La distribución normal

Definición 16.46. X sigue una **distribución normal** $N(\mu,\sigma)$ si su densidad es

$$f(x)=\frac{1}{\sigma\sqrt{2\pi}}\,e^{-\frac{(x-\mu)^2}{2\sigma^2}}.$$

La normal **tipificada** o **estándar** es $Z\sim N(0,1)$, y su función de distribución acumulada se denota $\Phi(z)=P(Z\le z)$.

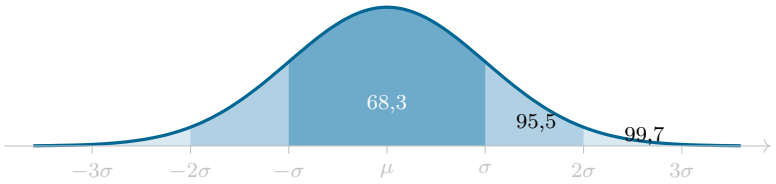


Figura 16.1: La regla empírica de la normal: entre $\mu\pm\sigma$ cae el 68,3 por ciento de la masa, entre $\mu\pm2\sigma$ el 95,5 y entre $\mu\pm3\sigma$ el 99,7. Estas tres cifras explican por qué σ es la unidad natural para medir cualquier desviación.

Regla: tipificación

Si $X\sim N(\mu,\sigma)$, entonces

$$Z=\frac{X-\mu}{\sigma}\sim N(0,1),$$

y por tanto

$$P(a\le X\le b)=\Phi\!\left(\frac{b-\mu}{\sigma}\right)-\Phi\!\left(\frac{a-\mu}{\sigma}\right).$$

Idea clave

Qué está haciendo la tipificación. Z mide *a cuántas desviaciones típicas del centro* está un dato. Un $z=2$ significa "dos σ por encima de la media", y eso significa lo mismo en estaturas, en notas o en piezas de motor. Tipificar es cambiar de unidades a una escala universal, exactamente como pasar de grados a radianes en la trigonometría elemental. Por eso basta *una* tabla —la de Φ— para todas las normales del mundo: una vez tipificadas, todas son la misma.
Las simetrías que hay que saber usar sin pensarlas:

$$\Phi(-z)=1-\Phi(z),\qquad P(Z\ge z)=1-\Phi(z),\qquad P(|Z|\le z)=2\Phi(z)-1.$$

Regla: la regla 68–95–99,7

En cualquier normal,

$$P(\mu-\sigma\le X\le\mu+\sigma)\approx 68\,\%,\quad P(\mu-2\sigma\le X\le\mu+2\sigma)\approx 95\,\%,\quad P(\mu-3\sigma\le X\le\mu+3\sigma)\approx 99{,}7\,\%.$$

Con más precisión, el 95 % central corresponde a $z=1{,}96$ y el 99 % a $z=2{,}576$.

Ejemplo 16.47. Las estaturas de una población siguen $N(170, 8)$ en centímetros.

(a) $P(X < 180)$: tipificamos, $z = (180 - 170)/8 = 1{,}25$, luego $P = \Phi(1{,}25) \approx 0{,}8944$, un 89,4 %.

(b) $P(165 < X < 180)$: $z_1 = (165 - 170)/8 = -0{,}625$ y $z_2 = 1{,}25$, luego

$$P = \Phi(1{,}25) - \Phi(-0{,}625) = 0{,}8944 - 0{,}2660 = 0{,}6284.$$

(c) ¿Por encima de qué estatura está el 10 % más alto? Buscamos z con $\Phi(z) = 0{,}90$, que es $z \approx 1{,}2816$; deshaciendo la tipificación, $x = \mu + z\sigma = 170 + 8 \cdot 1{,}2816 \approx 180{,}3$ cm.

Cuidado

El apartado (c) va *al revés* que los otros dos y es donde más se falla. En (a) se conoce el valor y se busca la probabilidad; en (c) se conoce la probabilidad y se busca el valor, así que hay que entrar en la tabla por dentro y despejar $x = \mu + z\sigma$. Además, "el 10 % más alto" deja 90 % por debajo: hay que buscar 0,90, no 0,10. Dibujar la campana y sombrear antes de tocar la tabla evita las dos trampas.

16.9.1. Aproximación normal a la binomial

Teorema 16.48 (de Moivre–Laplace)**.** *Si $X \sim B(n, p)$ con $np \geq 5$ y $nq \geq 5$, entonces X se aproxima bien por $N\big(np, \sqrt{npq}\big)$.*

Regla: corrección por continuidad

Al sustituir una variable discreta por una continua hay que ensanchar medio punto a cada lado:

$$P(X \geq k) \approx P\left(Z \geq \frac{k - 0{,}5 - np}{\sqrt{npq}}\right), \qquad P(X \leq k) \approx P\left(Z \leq \frac{k + 0{,}5 - np}{\sqrt{npq}}\right).$$

Ejemplo 16.49. Se lanza una moneda 100 veces; ¿cuál es la probabilidad de obtener 60 caras o más? Aquí $np = 50$ y $\sqrt{npq} = \sqrt{25} = 5$.

Con corrección: $z = (59{,}5 - 50)/5 = 1{,}9$, luego $P \approx 1 - \Phi(1{,}9) = 0{,}0287$.

El valor exacto, sumando los 41 términos binomiales, es 0,02844. Sin corrección por continuidad habríamos usado $z = 2$ y obtenido 0,0228, con un error relativo del 20 %. *El medio punto no es un adorno.*

Observación 16.50. Este teorema es el primer caso históricamente demostrado (de Moivre, 1733) del *teorema central del límite*, y explica por qué la campana aparece por todas partes: cualquier magnitud que sea la suma de muchos efectos pequeños e independientes —el error de una medición, la altura, el ruido eléctrico— tiende a ser normal, con independencia de cómo sea cada efecto por separado. Es lo más parecido que tiene la probabilidad a una ley de la naturaleza, y su demostración moderna, vía funciones características, es esencialmente un argumento de transformada de Fourier: véase el capítulo 31.

Regla: cómo reconocer qué distribución modela el problema

Casi todos los ejercicios de este capítulo se resuelven en cuanto se identifica la distribución. El diagnóstico es mecánico:

1. **¿Lo que cuento es discreto o mido algo continuo?** Si la respuesta es un número entero de cosas (aciertos, llamadas, piezas defectuosas), la distribución es discreta; si es una magnitud (peso, tiempo, longitud), es continua.
2. **Discreto, número *fijo* de intentos independientes con la misma probabilidad p.** Es **binomial** $B(n,p)$: $P(X=k)=\binom{n}{k}p^k(1-p)^{n-k}$, con $E(X)=np$ y $\mathrm{Var}(X)=np(1-p)$. La palabra clave es *de n intentos, cuántos*.
3. **Discreto, intentos hasta el primer éxito.** Es **geométrica**: $P(X=k)=(1-p)^{k-1}p$, con $E(X)=1/p$. La palabra clave es *hasta que*.
4. **Discreto, sucesos raros en un intervalo de tiempo o espacio, con media λ conocida.** Es **Poisson**: $P(X=k)=e^{-\lambda}\lambda^k/k!$, con $E(X)=\mathrm{Var}(X)=\lambda$. La palabra clave es *por hora*, *por metro*, *por página*: no hay un n a la vista.
5. **Discreto, pero extraigo *sin reponer* de una población pequeña.** Los intentos ya no son independientes: es **hipergeométrica**. Si la población es mucho mayor que la muestra (regla práctica: $n \le 0{,}1N$), la binomial la aproxima bien.
6. **Continuo, tiempo de espera hasta el siguiente suceso de un proceso de Poisson.** Es **exponencial** de media $1/\lambda$: $P(X>t)=e^{-\lambda t}$.
7. **Continuo, suma o media de muchos efectos pequeños.** Es **normal** $N(\mu,\sigma)$, y se resuelve tipificando. Verifique siempre el resultado: toda probabilidad debe caer en $[0,1]$, y una P muy próxima a 0 o a 1 suele delatar un error de signo en la tipificación.

Ejemplo 16.51 (siete enunciados, siete distribuciones)**.** El ejercicio real no dice qué distribución usar. Aquí está el diagnóstico y la cuenta completa en cada caso.

(a) *Un jugador encesta el* 80 *% de sus tiros libres y lanza* 12*. ¿Probabilidad de encestar al menos* 10*?*

Número fijo de intentos ($n=12$), misma probabilidad ($p=0{,}8$), independientes: **binomial** $B(12;0{,}8)$.

$$P(X\ge 10)=\sum_{k=10}^{12}\binom{12}{k}0{,}8^k0{,}2^{12-k}=0{,}2835+0{,}2062+0{,}0687=0{,}5583.$$

Apenas algo más de la mitad, aunque el jugador sea excelente: $E(X)=12\cdot 0{,}8=9{,}6$ y $\sigma=\sqrt{1{,}92}\approx 1{,}386$, así que pedir 10 es pedir por encima de la media.

(b) *A una centralita entran* 3 *llamadas por hora de media. ¿Probabilidad de que en la próxima hora no entre ninguna? ¿Y de que entren* 5 *o más?*

Sucesos raros por unidad de tiempo, sin n a la vista: **Poisson** con $\lambda = 3$.

$$P(X = 0) = e^{-3} = 0{,}0498, \qquad P(X \geq 5) = 1 - \sum_{k=0}^{4} \frac{e^{-3}3^k}{k!} = 0{,}1847.$$

(c) *Se lanza un dado hasta sacar un* 6. *¿Probabilidad de que el primer* 6 *salga justo en el cuarto lanzamiento?*

Hasta que: **geométrica** con $p = 1/6$.

$$P(X = 4) = \left(\tfrac{5}{6}\right)^3 \cdot \tfrac{1}{6} = \frac{125}{1296} = 0{,}0965,$$

y de media hacen falta $E(X) = 1/p = 6$ lanzamientos.

(d) *De una baraja de* 52 *se sacan* 5 *cartas. ¿Probabilidad de que haya exactamente dos ases?*

Sin reposición y la muestra no es despreciable frente a la población: **hipergeométrica**.

$$P(X = 2) = \frac{\binom{4}{2}\binom{48}{3}}{\binom{52}{5}} = \frac{6 \cdot 17\,296}{2\,598\,960} = 0{,}03993.$$

Si se tratara mal como binomial $B(5; 4/52)$ saldría 0,04654: un 17 % de error, y siempre por exceso, porque quitar un as de la baraja hace más difícil el siguiente.

(e) *Las estaturas de una población siguen* $N(170; 8)$. *¿Qué proporción mide más de* 180 *cm?*

Magnitud continua acumulada de muchos efectos: **normal**. Se tipifica:

$$z = \frac{180 - 170}{8} = 1{,}25, \qquad P(Z > 1{,}25) = 1 - 0{,}8944 = 0{,}1056.$$

(f) *Una bombilla dura* 12 *meses de media y no envejece. ¿Probabilidad de que pase de* 20 *meses? ¿Y de que falle antes de* 6*?*

Tiempo de espera sin memoria: **exponencial** con $\lambda = 1/12$.

$$P(X > 20) = e^{-20/12} = 0{,}1889, \qquad P(X \leq 6) = 1 - e^{-1/2} = 0{,}3935.$$

(g) *Una fábrica produce piezas con un* 0,2 % *de defectuosas y se examina un lote de* 1000. *¿Probabilidad de encontrar exactamente* 3 *defectuosas?*

Es binomial $B(1000; 0{,}002)$ en rigor, pero con n enorme y p diminuto la cuenta exacta es incómoda y la **Poisson** con $\lambda = np = 2$ la sustituye:

$$\text{exacto: } 0{,}18063, \qquad \text{Poisson: } e^{-2}\frac{2^3}{3!} = 0{,}18045.$$

Coinciden en tres cifras. Esta es la razón histórica del nombre *ley de los sucesos raros*.

Cuidado

El error de diagnóstico más frecuente es usar la binomial cuando no hay un número fijo de intentos. "Lanzo el dado 10 veces, ¿cuántos seises?" es binomial; "lanzo el dado hasta el primer seis" es geométrica; "cuántos seises salen en una hora de partida" no es ninguna de las dos si no se sabe cuántas tiradas caben en una hora. Antes de escribir una fórmula hay que poder señalar con el dedo dónde está n.

Pausa para pensar

Las distribuciones de la lista no son siete objetos sueltos: están emparentadas. La geométrica es lo que se ve si se mira la binomial "desde el primer éxito"; la Poisson es el límite de la binomial cuando $n \to \infty$ y $p \to 0$ con $np = \lambda$ fijo; la exponencial es la versión continua de la geométrica, y por eso ambas son las únicas sin memoria; y la normal es el límite de casi todas cuando se suman muchas copias, que es exactamente lo que dice el teorema central del límite. Aprender una es aprender media familia.

16.10. Muestreo: la distribución de la media muestral

Teorema 16.52 (central del límite, versión operativa)**.** *Sea una población con media μ y desviación típica σ, y sea $\bar{X}$ la media de una muestra aleatoria de tamaño n. Entonces*

$$E(\bar{X}) = \mu, \qquad \sigma_{\bar{X}} = \frac{\sigma}{\sqrt{n}},$$

y para n suficientemente grande (en la práctica $n \geq 30$) $\bar{X}$ sigue aproximadamente una $N(\mu, \sigma/\sqrt{n})$, sea cual sea la forma de la población de partida.

Idea clave

El $\sqrt{n}$ es la clave de todo. $\sigma_{\bar{X}} = \sigma/\sqrt{n}$ se llama *error típico* y dice tres cosas de golpe:

1. Promediar reduce la incertidumbre: la media de la muestra es más estable que un individuo.
2. Pero la reduce *despacio*. Para dividir el error entre dos hay que *cuadruplicar* la muestra; para dividirlo entre diez, multiplicarla por cien.
3. Y no depende del tamaño de la población. Una muestra de 1000 personas informa casi igual de bien sobre un país de un millón que sobre uno de trescientos millones. Es contraintuitivo y es cierto: lo que importa es n, no la fracción muestreada.

El punto 2 es la razón económica de que las encuestas tengan el margen de error que tienen y no menos: bajar del ± 3 % al ± 1 % multiplica el coste por nueve.

Ejemplo 16.53. Si $\sigma = 8$ cm y se toman muestras de $n = 64$ personas, el error típico es $8/\sqrt{64} = 1$ cm. Las medias muestrales se reparten, pues, como una $N(\mu, 1)$: el 95 % de ellas

caerá a menos de 1,96 cm de la media verdadera.

16.11. Intervalos de confianza

Definición 16.54. Un **intervalo de confianza** al nivel $1-\alpha$ para la media μ, conocida σ, es

$$\left(\bar{x} - z_{\alpha/2}\frac{\sigma}{\sqrt{n}}\ ,\ \ \bar{x} + z_{\alpha/2}\frac{\sigma}{\sqrt{n}}\right),$$

donde $z_{\alpha/2}$ cumple $P(|Z| \le z_{\alpha/2}) = 1-\alpha$. Los valores de uso corriente son

$$90\,\% \to z = 1{,}645, \qquad 95\,\% \to z = 1{,}96, \qquad 99\,\% \to z = 2{,}576.$$

La cantidad $E = z_{\alpha/2}\,\sigma/\sqrt{n}$ es el **error máximo admisible** o semiamplitud.

Ejemplo 16.55. Con $\sigma = 8$, una muestra de $n = 64$ da $\bar{x} = 172{,}4$ cm. El intervalo al 95 % es

$$172{,}4 \pm 1{,}96 \cdot \frac{8}{8} = 172{,}4 \pm 1{,}96 = (170{,}44;\ 174{,}36).$$

Al 99 % sería $172{,}4 \pm 2{,}576 = (169{,}82;\ 174{,}98)$: más confianza, intervalo más ancho. *Siempre* se paga la seguridad con precisión.

Cuidado

Lo que *no* significa un intervalo al 95 %: que μ tenga un 95 % de probabilidad de estar dentro. μ es un número fijo y desconocido; está dentro o no está, sin probabilidad de por medio. Lo aleatorio es el intervalo, que cambia con cada muestra. La frase correcta es: *si repitiera el muestreo indefinidamente, el 95 % de los intervalos construidos así contendrían a* μ. La confianza es una propiedad del procedimiento, no del intervalo concreto que tengo delante.

Regla: tamaño de muestra necesario

Despejando n de $E = z\,\sigma/\sqrt{n}$:

$$n \ge \left(\frac{z_{\alpha/2}\,\sigma}{E}\right)^2,$$

y se redondea siempre *hacia arriba*, porque un n menor no alcanza el error pedido.

Ejemplo 16.56. ¿Cuántas personas hay que medir para estimar μ con error máximo 0,5 cm al 95 %, sabiendo que $\sigma = 8$?

$$n \ge \left(\frac{1{,}96 \cdot 8}{0{,}5}\right)^2 = (31{,}36)^2 = 983{,}45 \ \Rightarrow\ n = 984.$$

Nótese el precio del rigor: para pasar de ± 2 cm a $\pm 0{,}5$ cm hacen falta dieciséis veces más personas.

Regla: intervalo de confianza para una proporción

La media no es lo único que se estima: las encuestas, los controles de calidad y los ensayos clínicos estiman *proporciones*. El procedimiento es el mismo con una sola sustitución.

1. Contar los éxitos x en la muestra de tamaño n y formar la proporción muestral $\hat{p} = x/n$.
2. Comprobar que la aproximación normal es lícita: hacen falta $n\hat{p} \geq 5$ y $n(1-\hat{p}) \geq 5$. Con muestras pequeñas o proporciones extremas hay que usar la binomial exacta.
3. Calcular el error típico $\sigma_{\hat{p}} = \sqrt{\dfrac{\hat{p}\,(1-\hat{p})}{n}}$. Es el análogo exacto de $\sigma/\sqrt{n}$, con la particularidad de que la propia muestra proporciona la σ.
4. Escribir el intervalo $\hat{p} \pm z_{\alpha/2}\,\sigma_{\hat{p}}$, con los mismos z de siempre: 1,645 al 90 %, 1,96 al 95 %, 2,576 al 99 %.
5. Para el tamaño de muestra necesario, despejar $n \geq z_{\alpha/2}^2\,\hat{p}(1-\hat{p})/E^2$ y redondear *hacia arriba*. Si no hay ninguna estimación previa de $\hat{p}$, usar el caso peor $\hat{p} = 0{,}5$, que maximiza $\hat{p}(1-\hat{p}) = 0{,}25$.
6. Interpretar en el lenguaje del problema y con el sujeto correcto: lo aleatorio es el intervalo, no la proporción verdadera.
7. Verificar: el intervalo debe quedar dentro de $[0, 1]$ y ser simétrico respecto de $\hat{p}$. Un extremo negativo o mayor que 1 delata que la aproximación normal no era aplicable.

Ejemplo 16.57 (cinco preguntas sobre la misma encuesta)**.** En una muestra de $n = 400$ votantes, 240 declaran apoyar una medida. Todo lo que sigue sale de ese único dato.

(a) *Estimación puntual y error típico.*

$$\hat{p} = \frac{240}{400} = 0{,}6, \qquad \sigma_{\hat{p}} = \sqrt{\frac{0{,}6 \cdot 0{,}4}{400}} = \sqrt{0{,}0006} = 0{,}024495.$$

Comprobación previa: $n\hat{p} = 240 \geq 5$ y $n(1-\hat{p}) = 160 \geq 5$. ✓

(b) *Intervalo al* 95 *%.*

$$0{,}6 \pm 1{,}96 \cdot 0{,}024495 = 0{,}6 \pm 0{,}04801 = (0{,}5520;\ 0{,}6480),$$

es decir, entre el 55,2 % y el 64,8 %. El titular de prensa sería "60 % de apoyo, con un margen de error de ±4,8 puntos".

(c) *El mismo dato al* 99 *%.*

$$0{,}6 \pm 2{,}576 \cdot 0{,}024495 = 0{,}6 \pm 0{,}06310 = (0{,}5369;\ 0{,}6631).$$

Más seguridad, intervalo más ancho: es el mismo canje que con la media, y no hay manera de evitarlo sin encuestar a más gente.

(d) *¿Se puede afirmar que la medida tiene mayoría?* La pregunta es si 0,5 cae dentro del intervalo. Como el extremo inferior al 95 % es $0{,}5520 > 0{,}5$, el valor 0,5 queda fuera, y también queda fuera del intervalo al 99 % (extremo inferior 0,5369). La conclusión es que sí, con esos datos la mayoría está sostenida.

(e) *¿Cuánta gente hay que encuestar para un margen de* ± 2 *puntos al* 95 *%?* Sin información previa se usa el caso peor:

$$n \geq \frac{1{,}96^2 \cdot 0{,}25}{0{,}02^2} = \frac{0{,}9604}{0{,}0004} = 2401.$$

Si se acepta la estimación previa $\hat{p} = 0{,}6$, el requisito baja a $n \geq 0{,}9604 \cdot 0{,}24/0{,}0004 = 2304{,}96$, o sea 2305 personas: apenas un 4 % menos, y por eso las encuestadoras suelen ir directamente al caso peor.

(f) *Por qué casi todas las encuestas usan mil personas.* Con $n = 1000$ y el caso peor,

$$E = 1{,}96\sqrt{\frac{0{,}25}{1000}} = 0{,}0310,$$

el famoso "± 3 puntos". Para bajar a ± 2 hacen falta 2401; para bajar a ± 1 harían falta 9604. El coste crece con el *cuadrado* de la precisión, y ahí es donde se detiene el presupuesto de cualquier medio de comunicación.

Cuidado

Un empate técnico no es un empate. Si una encuesta con $n = 1000$ da 52 % para un candidato, el intervalo al 95 % es (0,4890; 0,5510): contiene el 50 %, de modo que los datos no permiten afirmar que vaya en cabeza. Anunciar "gana por dos puntos" cuando el margen de error es de tres es el abuso estadístico más repetido de la prensa, y no es una cuestión de opinión: el intervalo está ahí y se calcula en dos líneas.

Pausa para pensar

Obsérvese la rareza de la fórmula: $\sigma_{\hat{p}}$ depende de n pero *no* del tamaño de la población, siempre que la población sea mucho mayor que la muestra. Para estimar la opinión de Guatemala y la de China hace falta el mismo número de encuestados. Suena imposible y no lo es: lo que se está midiendo es la variabilidad de la muestra, y una cucharada de sopa bien removida sabe igual venga de un plato o de una olla.

Idea clave

La misma maquinaria, tres disfraces. Estimar una media con σ conocida, estimar una proporción y contrastar una hipótesis son el mismo cálculo: se toma un estadístico, se divide su desviación respecto del centro entre su error típico y se compara con la normal. Cambia el error típico ($\sigma/\sqrt{n}$ o $\sqrt{\hat{p}(1-\hat{p})/n}$) y nada más. Quien vea esa unidad se ahorra memorizar tres tablas.

16.12. Contraste de hipótesis

Definición 16.58. Un **contraste de hipótesis** enfrenta dos afirmaciones:

- la **hipótesis nula** H_0, que es la situación por defecto, la que se mantiene mientras no haya pruebas en contra;
- la **hipótesis alternativa** H_1, que es lo que se quiere demostrar.

El **nivel de significación** α es la probabilidad que se acepta de rechazar H_0 siendo cierta. El p**-valor** es la probabilidad de observar un resultado tan extremo como el observado, o más, *suponiendo cierta* H_0. Se rechaza H_0 cuando $p < \alpha$.

Idea clave

Por qué el contraste está construido al revés. Un contraste es una demostración por reducción al absurdo (capítulo 1) con el absurdo debilitado a "improbable". Se supone H_0, se calcula qué tan raro sería lo observado bajo esa suposición, y si sale rarísimo se abandona H_0.
De ahí la asimetría que desconcierta a todo el mundo: se puede *rechazar* H_0, pero nunca *demostrarla*. Es el paralelo exacto de un juicio: la presunción de inocencia es H_0, y el veredicto no es "inocente" sino "no culpable". Falta de pruebas no es prueba de ausencia.

Regla: los dos errores

	H_0 cierta	H_0 falsa
Rechazo H_0	error tipo I (prob. α)	decisión correcta
No rechazo H_0	decisión correcta	error tipo II (prob. β)

Bajar α sube β y viceversa: la única manera de reducir ambos es aumentar n.

Regla: procedimiento para la media, σ conocida

1. Formular $H_0\colon \mu = \mu_0$ y H_1 (bilateral $\mu \neq \mu_0$, o unilateral $\mu < \mu_0$ o $\mu > \mu_0$).
2. Calcular el estadístico $z = \dfrac{\bar{x} - \mu_0}{\sigma/\sqrt{n}}$.
3. Comparar con el valor crítico ($\pm 1{,}96$ si bilateral al 5 %; $\pm 1{,}645$ si unilateral al 5 %), o equivalentemente calcular el p-valor.
4. Concluir *en el lenguaje del problema*, no en el de la tabla.

Ejemplo 16.59 (contraste unilateral)**.** Un fabricante afirma que sus paquetes pesan 500 g de media, con $\sigma = 12$ g. Una inspección toma $n = 36$ paquetes y obtiene $\bar{x} = 495{,}4$ g. ¿Hay evidencia, al 5 %, de que pesan menos de lo declarado?

$H_0\colon \mu = 500$ frente a $H_1\colon \mu < 500$. El error típico es $12/\sqrt{36} = 2$, y

$$z = \frac{495{,}4 - 500}{2} = -2{,}30.$$

El valor crítico unilateral es $-1{,}645$, y $-2{,}30 < -1{,}645$: se rechaza H_0. El p-valor es $\Phi(-2{,}30) = 0{,}0107$, muy por debajo de 0,05.

Traducido: si el fabricante dijera la verdad, ver una muestra tan baja o más ocurriría solo un 1,1 % de las veces. Hay evidencia razonable de que los paquetes pesan menos de lo declarado. La región crítica, en gramos, es $\bar{x} < 500 - 1{,}645 \cdot 2 = 496{,}71$.

Ejemplo 16.60 (contraste bilateral)**.** Una máquina debe cortar piezas de 250 mm con $\sigma = 6$ mm. Una muestra de $n = 25$ da $\bar{x} = 252{,}5$ mm. ¿Está descalibrada, al 5 %?

$H_0\colon \mu = 250$ frente a $H_1\colon \mu \neq 250$. Error típico $6/5 = 1{,}2$ y

$$z = \frac{252{,}5 - 250}{1{,}2} = 2{,}083.$$

Como $|z| > 1{,}96$, se rechaza H_0; el p-valor bilateral es $2\big(1 - \Phi(2{,}083)\big) = 0{,}0372$. Aquí la alternativa es bilateral porque desviarse por exceso es tan malo como por defecto: no interesa que la pieza sea larga, interesa que sea correcta. Elegir bilateral o unilateral se decide *antes* de mirar los datos, nunca después.

Cuidado

Tres frases que aparecen en exámenes, informes y periódicos, y las tres son falsas:

- "El p-valor es la probabilidad de que H_0 sea cierta." No: es la probabilidad de los *datos* suponiendo H_0, que es la condicional al revés (véase Bayes en el capítulo 5).
- "$p > 0{,}05$, luego H_0 es cierta." No: no se ha encontrado evidencia, que es otra cosa. Con n pequeño casi nunca se encuentra.
- "$p < 0{,}05$, luego el efecto es importante." No: significativo no es sinónimo de grande. Con $n = 10^6$, una diferencia irrelevante sale significativa. Significación estadística e importancia práctica son preguntas distintas, y solo la segunda decide algo.

Regla: cómo calcular e interpretar el p-valor

El valor crítico dice sí o no; el p-valor dice *cuánto*. Es la manera moderna de presentar un contraste, y la que aparece en cualquier programa de estadística.

1. Formular H_0 y H_1 *antes* de mirar los datos, y decidir si el contraste es bilateral o unilateral.
2. Calcular el estadístico, por ejemplo $z = (\bar{x} - \mu_0)/(\sigma/\sqrt{n})$.
3. Si el contraste es **unilateral**, el p-valor es la probabilidad de la cola que apunta hacia H_1: $p = P(Z \geq z)$ si H_1 dice "mayor", $p = P(Z \leq z)$ si dice "menor".
4. Si es **bilateral**, se toman las dos colas: $p = 2\,P(Z \geq |z|)$. Es exactamente el doble del unilateral, y olvidar ese factor 2 es el error de cuenta más común del capítulo.
5. Comparar con α: se rechaza H_0 si $p < \alpha$. Nada más.

6. Redactar la conclusión en el lenguaje del problema, incluyendo el p-valor y no solo la palabra "significativo".
7. Verificar la coherencia: un z grande en valor absoluto tiene que dar un p pequeño. Si sale al revés, hay un signo mal puesto o se ha usado la cola equivocada.

Ejemplo 16.61 (seis contrastes y los cuatro errores de lectura)**.** (a) *Bombillas.* El fabricante afirma $\mu = 1000$ horas, con $\sigma = 60$ conocida. Una muestra de $n = 36$ da $\bar{x} = 980$. Un consumidor sospecha que duran *menos*, así que $H_1 : \mu < 1000$, unilateral por la izquierda:

$$z = \frac{980 - 1000}{60/\sqrt{36}} = \frac{-20}{10} = -2, \qquad p = P(Z \le -2) = 0{,}0228.$$

Como $0{,}0228 < 0{,}05$, se rechaza H_0: los datos sostienen que duran menos de lo anunciado.

(b) *El mismo estadístico, otra pregunta.* Si el contraste fuera bilateral ($H_1 : \mu \neq 1000$, alguien que solo quiere saber si el fabricante miente en algún sentido), el p-valor sería

$$p = 2 \cdot 0{,}0228 = 0{,}0455,$$

que sigue por debajo de 0,05, pero por muy poco. *Los mismos datos, distinta pregunta, distinto p*: por eso la hipótesis hay que fijarla antes.

(c) *Un caso que se queda fuera.* Otra muestra da $\bar{x} = 1018$, y se contrasta bilateralmente:

$$z = \frac{1018 - 1000}{10} = 1{,}8, \qquad p = 2\,P(Z \ge 1{,}8) = 2 \cdot 0{,}0359 = 0{,}0719.$$

Como $0{,}0719 > 0{,}05$, *no* se rechaza H_0. Nótese la redacción: no se ha demostrado que $\mu = 1000$; simplemente no hay pruebas suficientes en contra.

(d) *Una moneda con* 60 *caras en* 100 *lanzamientos.* Aquí la distribución exacta es binomial $B(100; 0{,}5)$ bajo H_0:

$$P(X \ge 60) = 0{,}028444, \qquad p_{\text{bilateral}} = 2 \cdot 0{,}028444 = 0{,}056888.$$

Con la aproximación normal y corrección por continuidad, $z = (59{,}5 - 50)/5 = 1{,}9$ y $P(Z \ge 1{,}9) = 0{,}028717$: prácticamente el valor exacto. *Sin* corrección saldría $z = 2$ y 0,02275, un 20 % por debajo, y el contraste bilateral cruzaría el umbral del 5 % por el lado equivocado. La corrección no es un adorno.

(e) *El mismo experimento, menos desviación.* Con 55 caras en 100,

$$P(X \ge 55) = 0{,}1841, \qquad p_{\text{bilateral}} = 0{,}3682,$$

lejísimos de cualquier α razonable. Cinco caras de más en cien lanzamientos es un resultado perfectamente vulgar.

(f) *Significativo no es importante.* Se contrasta $H_0 : \mu = 100$ con $\sigma = 5$ y una diferencia real minúscula, $\bar{x} - \mu_0 = 0{,}2$. Según el tamaño de muestra:

n	z	p bilateral
25	0,20	0,8415
100	0,40	0,6892
2500	2,00	0,0455
10 000	4,00	0,00006

La diferencia es la misma en las cuatro filas: dos décimas sobre cien, irrelevante en cualquier aplicación. Con muestra suficiente, *cualquier* diferencia distinta de cero acaba siendo significativa. El p-valor mide la evidencia contra H_0, no el tamaño del efecto.

Cuidado

Los cuatro errores de lectura del p-valor, en orden de gravedad:

1. **Creer que p es la probabilidad de que H_0 sea cierta.** No lo es: p se calcula *suponiendo* H_0 cierta, así que no puede a la vez medir su probabilidad. Es $P(\text{datos} \mid H_0)$, jamás $P(H_0 \mid \text{datos})$; la diferencia entre ambas es exactamente el teorema de Bayes del capítulo 5.
2. **Creer que $p > \alpha$ demuestra H_0.** No rechazar es el "no culpable" del juicio, no el "inocente".
3. **Tratar 0,049 y 0,051 como mundos distintos.** El umbral 0,05 es una convención de Fisher, no una ley de la naturaleza. Dos estudios con esos p dicen esencialmente lo mismo.
4. **Elegir la hipótesis después de ver los datos.** Mirar la muestra, ver que $\bar{x}$ salió alto y *entonces* decidir contrastar $H_1\colon \mu > \mu_0$ duplica en la práctica el nivel real de significación.

Pausa para pensar

Si se hacen k contrastes independientes al 5 % sobre hipótesis todas ciertas, la probabilidad de que al menos uno salga "significativo" por puro azar es $1 - 0{,}95^k$:

$$k = 1 \to 5\,\%, \quad k = 5 \to 22{,}6\,\%, \quad k = 10 \to 40{,}1\,\%, \quad k = 20 \to 64{,}2\,\%, \quad k = 100 \to 99{,}4\,\%.$$

Un investigador que pruebe veinte variables y publique la única que salió con $p < 0{,}05$ no ha descubierto nada: era más probable que no que apareciera alguna. Este es el problema de las comparaciones múltiples, y la corrección más simple —la de Bonferroni— consiste en exigir $p < \alpha/k$ a cada contraste. Con $k = 20$ y $\alpha = 0{,}05$ eso significa exigir $p < 0{,}0025$, un listón bastante más alto.

Idea clave

Qué habría que publicar en lugar del p-valor. Un p solo responde "¿es distinguible de cero?". La pregunta interesante casi siempre es "¿de qué tamaño es el efecto y con

cuánta precisión lo conozco?", y esa la responde el intervalo de confianza. De hecho el intervalo contiene al contraste: rechazar $H_0\colon \mu = \mu_0$ al nivel α bilateral equivale exactamente a que μ_0 quede fuera del intervalo de confianza al $1-\alpha$. Dan siempre la misma decisión, pero el intervalo además dice cuánto.

16.13. Errores típicos de probabilidad e inferencia

1. Confundir la esperanza con un valor posible: $E(X) = 3{,}5$ en un dado que solo saca enteros.
2. Sumar varianzas de variables dependientes. $\operatorname{Var}(X+Y) = \operatorname{Var}(X) + \operatorname{Var}(Y)$ exige independencia; la esperanza no exige nada.
3. Olvidar el cuadrado en $\operatorname{Var}(aX) = a^2 \operatorname{Var}(X)$.
4. Tratar la densidad $f(x)$ como si fuera una probabilidad. Solo el área lo es; f puede superar 1.
5. Escribir $P(X = a) > 0$ en una variable continua.
6. Tipificar con la varianza en vez de con la desviación típica: es $(x-\mu)/\sigma$, no $(x-\mu)/\sigma^2$.
7. Usar σ donde va $\sigma/\sqrt{n}$. El error típico de la *media* lleva raíz de n; es el fallo más frecuente de todo el capítulo.
8. Olvidar la corrección por continuidad al aproximar una binomial por una normal.
9. Buscar en la tabla 0,10 cuando el enunciado pide "el 10 % superior".
10. Interpretar el nivel de confianza como probabilidad de que μ esté en el intervalo concreto obtenido.
11. Decidir si el contraste es uni o bilateral *después* de ver hacia dónde apuntan los datos: eso duplica la tasa real de falsos positivos.
12. Concluir "se acepta H_0". Se dice "no se rechaza".
13. Creer que una muestra grande corrige un sesgo de selección. No lo corrige: lo mide con más precisión.

16.14. Cuando σ no se conoce: la t de Student, y el contraste χ^2

Todos los intervalos y contrastes anteriores suponen σ *conocida*. En la práctica no lo está casi nunca, y sustituirla por la estimación muestral sin más tiene un coste que hay que cuantificar.

Definición 16.62 (Cuasivarianza y estadístico t)**.**

$$s^2 = \frac{1}{n-1}\sum_{i=1}^{n}(x_i - \bar{x})^2, \qquad t = \frac{\bar{x} - \mu}{s/\sqrt{n}} \;\sim\; t_{n-1}.$$

El denominador $n-1$ no es un capricho: $\bar{x}$ ya se ha estimado de los mismos datos, de modo que las desviaciones $x_i - \bar{x}$ están ligadas por la relación $\sum(x_i - \bar{x}) = 0$ y solo hay $n-1$ libres. Dividir entre n produce un estimador que subestima σ^2 de forma sistemática.

La distribución t tiene más cola que la normal porque acumula dos incertidumbres: la del numerador y la de haber estimado el denominador. Con $\nu \to \infty$ converge a la normal.

Ejemplo 16.63 (Diez piezas, y la diferencia entre z y t)**.** Sea $n = 10$, $\bar{x} = 172{,}4$ y $s = 8$.

(a) Con la t. $t_{0{,}025;9} = 2{,}262\,2$, luego el intervalo es

$$172{,}4 \pm 2{,}2622 \cdot \frac{8}{\sqrt{10}} = 172{,}4 \pm 5{,}723 = [166{,}68;\ 178{,}12].$$

(b) Con la normal (incorrecto). $172{,}4 \pm 1{,}96 \cdot 2{,}530 = [167{,}44;\ 177{,}36]$.

(c) La diferencia. El intervalo correcto es un 15,4 por ciento más ancho ($2{,}2622/1{,}96 = 1{,}154$). Eso basta para que un resultado en el filo cambie de veredicto: con $n = 10$, usar z en lugar de t declara significativo lo que no lo es.

(d) Cuándo da igual. Con $n = 30$, $t_{0{,}025;29} = 2{,}045$: solo un 4 por ciento más que 1,96. Con $n = 100$, un 1 por ciento. La regla práctica de "a partir de 30 da igual" es razonable, pero no hay motivo para aplicarla: cualquier calculadora da la t, y la t es correcta siempre.

Contrastar una distribución entera, no un parámetro

Definición 16.64 (Bondad de ajuste)**.** Con k categorías, frecuencias observadas O_i y esperadas E_i bajo H_0,

$$\chi^2 = \sum_{i=1}^{k} \frac{(O_i - E_i)^2}{E_i} \;\sim\; \chi^2_{k-1-p},$$

donde p es el número de parámetros estimados a partir de los datos. Se pide $E_i \geq 5$ en todas las casillas.

Ejemplo 16.65 (¿Es justo este dado?)**.** Sesenta lanzamientos, con frecuencias observadas $8, 9, 12, 15, 6, 10$. Bajo equiprobabilidad, $E_i = 10$ en las seis caras.

$$\chi^2 = \frac{4 + 1 + 4 + 25 + 16 + 0}{10} = \frac{50}{10} = 5{,}0, \qquad 5 \text{ grados de libertad}, \qquad p = 0{,}415\,9.$$

No se rechaza. El 15 del cuarto resultado llama la atención, pero con sesenta tiradas una desviación así es perfectamente rutinaria: la desviación típica de cada casilla es $\sqrt{10 \cdot \frac{5}{6}} = 2{,}9$, y el 15 está a 1,7 desviaciones.

Obsérvese la utilidad del contraste: es la primera herramienta del libro que permite *comprobar* la hipótesis de equiprobabilidad sobre la que descansa toda la regla de Laplace, en lugar de darla por buena.

Observación 16.66 (Tablas de contingencia). El mismo estadístico contrasta la independencia de dos variables categóricas. Las frecuencias esperadas se obtienen del producto de marginales, $E_{ij} = \frac{f_i\, c_j}{n}$, y los grados de libertad son $(f-1)(c-1)$: se pierde uno por fila y uno por columna porque los marginales se han fijado. Es el contraste más usado del mundo sobre datos de encuesta, y su hipótesis nula —independencia— es exactamente la que definimos al principio del capítulo con $P(A \cap B) = P(A)P(B)$.

16.15. Vectores aleatorios, la linealidad de la esperanza y las desigualdades de Jensen y Cauchy–Schwarz

Este capítulo lleva muchas páginas usando tres cosas que enunció como reglas y nunca demostró: que la esperanza de una suma es la suma de las esperanzas, que $\operatorname{Var}(aX+b) = a^2 \operatorname{Var}(X)$ y que la varianza de una suma lleva un término cruzado. También ha manejado «densidades conjuntas» sin haber dicho qué son, y ha afirmado que el coeficiente de correlación está entre -1 y 1 sin justificarlo. Esta sección cierra los cinco huecos, y añade la desigualdad de Jensen, que no aparecía en ninguna parte y que es la herramienta que explica por qué tantas medias salen «al revés» de lo que uno espera.

Vectores aleatorios y densidad conjunta

Definición 16.67 (Vector aleatorio y distribución conjunta continua)**.** Un *vector aleatorio* es un par (X, Y) —o una lista $(X_1, \dots, X_n)$— de variables aleatorias definidas sobre el *mismo* experimento. La condición «el mismo experimento» es esencial: sin ella no tiene sentido preguntar por X e Y a la vez.

El vector (X, Y) es *continuo* si existe una *función de densidad conjunta* $f(x, y) \geq 0$ tal que para toda región razonable $A \subseteq \mathbb{R}^2$

$$P\big((X,Y) \in A\big) = \iint_A f(x,y)\,\mathrm{d}x\,\mathrm{d}y, \qquad \text{con} \qquad \iint_{\mathbb{R}^2} f(x,y)\,\mathrm{d}x\,\mathrm{d}y = 1.$$

Sus *densidades marginales* se obtienen integrando la otra variable:

$$f_X(x) = \int_{-\infty}^{\infty} f(x,y)\,\mathrm{d}y, \qquad f_Y(y) = \int_{-\infty}^{\infty} f(x,y)\,\mathrm{d}x.$$

Las variables son *independientes* si $f(x,y) = f_X(x)\, f_Y(y)$ para todo (x, y). Y para cualquier función g razonable,

$$\mathrm{E}[g(X,Y)] = \iint_{\mathbb{R}^2} g(x,y)\, f(x,y)\,\mathrm{d}x\,\mathrm{d}y,$$

que es la versión de dos variables de la regla de que para calcular la esperanza de una función de X no hace falta hallar antes la distribución de esa función.

Idea clave

Densidad conjunta es masa por unidad de área. En una variable, la densidad era altura y la probabilidad, área bajo la curva. En dos variables, la densidad es una superficie sobre el plano y la probabilidad es el *volumen* que hay debajo, sobre la región

que interese. La marginal de X se obtiene «aplastando» esa superficie contra el eje x: se suma —se integra— toda la masa que hay encima de cada abscisa, sin importar el valor de y. Por eso la marginal se llama así: es lo que queda escrito en el margen de la tabla cuando se totaliza una fila o una columna.

Ejemplo 16.68 (Uniforme en un triángulo: marginales y dependencia)**.** Sea (X, Y) uniforme en el triángulo $T = \{(x, y) : 0 < x < y < 1\}$. Como el área de T es $\frac{1}{2}$, la densidad constante que integra 1 es $f(x, y) = 2$ dentro de T y 0 fuera. Las marginales son

$$f_X(x) = \int_x^1 2\,\mathrm{d}y = 2(1-x), \qquad f_Y(y) = \int_0^y 2\,\mathrm{d}x = 2y, \qquad 0 < x, y < 1.$$

Ambas integran 1. ✓Y no son independientes: el producto $f_X(x)f_Y(y) = 4y(1-x)$ no es $f(x, y)$; de hecho basta observar que el soporte no es un rectángulo —si X vale 0,9, entonces Y está forzada a estar entre 0,9 y 1—. *Cuando el soporte conjunto no es un producto de intervalos, las variables no pueden ser independientes*, sin necesidad de calcular nada.

Las esperanzas: $\mathrm{E}[X] = \int_0^1 x\,2(1-x)\mathrm{d}x = 1 - \frac{2}{3} = \frac{1}{3}$ y $\mathrm{E}[Y] = \int_0^1 y\,2y\,\mathrm{d}y = \frac{2}{3}$. Coherente con el dibujo: el triángulo tiene más masa a la izquierda en x y más arriba en y.

La linealidad de la esperanza, demostrada

Teorema 16.69 (Linealidad)**.** *Sean X e Y variables aleatorias con esperanza finita, definidas sobre el mismo experimento, y a, b, c constantes. Entonces*

$$\mathrm{E}[aX + bY + c] = a\,\mathrm{E}[X] + b\,\mathrm{E}[Y] + c.$$

***No se supone ninguna relación entre X e Y:** la fórmula vale aunque sean dependientes, aunque una determine completamente a la otra, y aunque su distribución conjunta sea desconocida.*

Demostración. Caso discreto. Sea $p(x, y) = P(X = x, Y = y)$ la función de masa conjunta. Sumando primero en y se obtiene la marginal, $\sum_y p(x, y) = P(X = x)$, y análogamente al revés. Entonces

$$\mathrm{E}[aX+bY+c] = \sum_x \sum_y (ax+by+c)\,p(x,y) = a\sum_x x \sum_y p(x,y) + b\sum_y y \sum_x p(x,y) + c\sum_x \sum_y p(x,y),$$

donde la reordenación es lícita porque las series convergen absolutamente. Los tres sumandos son $a\sum_x x\,P(X = x) = a\,\mathrm{E}[X]$, $b\,\mathrm{E}[Y]$ y $c \cdot 1$.

Caso continuo. Idéntico con integrales: con densidad conjunta f,

$$\mathrm{E}[aX + bY + c] = \iint (ax + by + c) f(x, y)\,\mathrm{d}x\,\mathrm{d}y = a\int x f_X(x)\mathrm{d}x + b\int y f_Y(y)\mathrm{d}y + c,$$

usando el teorema de Fubini (capítulo 31) para intercambiar el orden y las definiciones de las marginales.

Obsérvese dónde *no* ha hecho falta la independencia: en ningún sitio se ha factorizado $p(x, y)$ ni $f(x, y)$. Lo único que se ha usado es que al sumar o integrar en una variable aparece la marginal de la otra, y eso es cierto siempre. □

Teorema 16.70 (Producto de independientes)**.** *Si X e Y son independientes con esperanza finita, entonces* $\mathrm{E}[XY] = \mathrm{E}[X]\,\mathrm{E}[Y]$.

Demostración. Con densidad conjunta $f = f_X f_Y$,

$$\mathrm{E}[XY] = \iint xy\, f_X(x) f_Y(y)\,\mathrm{d}x\,\mathrm{d}y = \left(\int x f_X(x)\mathrm{d}x\right)\left(\int y f_Y(y)\mathrm{d}y\right),$$

y en el caso discreto igual con sumas. Aquí la independencia sí es imprescindible, y es lo que separa este resultado del anterior. □

Ejemplo 16.71 (Los sombreros: linealidad sin ninguna independencia)**.** n personas dejan su sombrero en el guardarropa y al salir se les devuelve uno al azar, en una permutación uniforme. Sea X el número de personas que recuperan el suyo. ¿Cuánto vale $\mathrm{E}[X]$?

Escríbase $X = I_1 + \cdots + I_n$, donde I_j vale 1 si la persona j acierta y 0 si no. Cada I_j vale 1 con probabilidad $1/n$ —su sombrero está en una posición uniforme—, luego $\mathrm{E}[I_j] = 1/n$ y, por linealidad,

$$\mathrm{E}[X] = n \cdot \frac{1}{n} = 1.$$

Sea cual sea n, en media acierta *una* persona. Y los I_j son claramente dependientes: si las $n-1$ primeras aciertan, la última acierta con probabilidad 1. La linealidad no se entera de esa dependencia y da igualmente el resultado exacto.

Comprobación completa con $n = 4$. De las $4! = 24$ permutaciones, 9 no dejan ningún punto fijo, 8 dejan uno, 6 dejan dos y 1 deja cuatro (ninguna deja exactamente tres). Luego

$$\mathrm{E}[X] = \frac{0\cdot 9 + 1\cdot 8 + 2\cdot 6 + 4\cdot 1}{24} = \frac{24}{24} = 1.\checkmark$$

Las propiedades de la varianza, demostradas

Definición 16.72 (Covarianza)**.** Para X e Y con varianza finita,

$$\mathrm{Cov}(X,Y) = \mathrm{E}\big[(X-\mu_X)(Y-\mu_Y)\big] = \mathrm{E}[XY] - \mathrm{E}[X]\,\mathrm{E}[Y],$$

donde la segunda igualdad sale de desarrollar el producto y aplicar linealidad. En particular $\mathrm{Cov}(X,X) = \mathrm{Var}(X)$, y si X e Y son independientes la covarianza es 0 por el teorema 16.70.

Teorema 16.73 (Las cuatro propiedades básicas)**.** *Para constantes a, b y variables con varianza finita:*

1. $\mathrm{Var}(X) = \mathrm{E}[X^2] - \mathrm{E}[X]^2$*;*
2. $\mathrm{Var}(aX + b) = a^2\,\mathrm{Var}(X)$*;*
3. $\mathrm{Var}(X+Y) = \mathrm{Var}(X) + \mathrm{Var}(Y) + 2\,\mathrm{Cov}(X,Y)$*;*
4. $\mathrm{Var}(X) \geq 0$*, con igualdad si y sólo si X es constante con probabilidad* 1*.*

Demostración. (1) Con $\mu = \mathrm{E}[X]$, desarróllese el cuadrado y aplíquese linealidad:

$$\mathrm{E}\big[(X-\mu)^2\big] = \mathrm{E}[X^2 - 2\mu X + \mu^2] = \mathrm{E}[X^2] - 2\mu\,\mathrm{E}[X] + \mu^2 = \mathrm{E}[X^2] - \mu^2.$$

(2) La esperanza de $aX+b$ es $a\mu+b$, luego la desviación respecto de su media es

$$(aX+b) - (a\mu+b) = a(X-\mu),$$

y el b desaparece antes de elevar al cuadrado. Elevando, $\mathrm{Var}(aX+b) = \mathrm{E}[a^2(X-\mu)^2] = a^2\,\mathrm{Var}(X)$. Esto explica de una vez las dos mitades de la regla: *trasladar no dispersa* y *escalar dispersa al cuadrado*. Con $\mathrm{Var}(X)=4$ resulta $\mathrm{Var}(3X-2) = 9\cdot 4 = 36$.

(3) Sean μ_X, μ_Y las medias. Como $\mathrm{E}[X+Y] = \mu_X + \mu_Y$ por linealidad,

$$\begin{aligned}\mathrm{Var}(X+Y) &= \mathrm{E}\big[\big((X-\mu_X)+(Y-\mu_Y)\big)^2\big] \\ &= \mathrm{E}\big[(X-\mu_X)^2\big] + \mathrm{E}\big[(Y-\mu_Y)^2\big] + 2\,\mathrm{E}\big[(X-\mu_X)(Y-\mu_Y)\big],\end{aligned}$$

que es justo lo afirmado. Si X e Y son independientes el último término se anula y se recupera la versión sin término cruzado que el capítulo venía usando.

(4) La variable $(X-\mu)^2$ es no negativa, luego su esperanza también. Si la esperanza de una variable no negativa es 0, la variable es 0 con probabilidad 1 —si tomara un valor $\varepsilon^2 > 0$ con probabilidad $q > 0$, su esperanza sería al menos $\varepsilon^2 q > 0$—, es decir $X = \mu$ casi seguramente. □

Proposición 16.74 (La covarianza es bilineal)**.** $\mathrm{Cov}(X,Y) = \mathrm{Cov}(Y,X)$, $\mathrm{Cov}(aX+b,\, cY+d) = ac\,\mathrm{Cov}(X,Y)$ *y* $\mathrm{Cov}(X+Z,\,Y) = \mathrm{Cov}(X,Y) + \mathrm{Cov}(Z,Y)$. *Todo sale de la linealidad de la esperanza aplicada a la definición.*

Ejemplo 16.75 (Los sombreros, ahora la varianza)**.** Con la notación de antes, $\mathrm{Var}(I_j) = \frac{1}{n} - \frac{1}{n^2}$. Para $i \neq j$, el suceso «ambos aciertan» tiene probabilidad $\frac{1}{n(n-1)}$, luego

$$\mathrm{Cov}(I_i, I_j) = \frac{1}{n(n-1)} - \frac{1}{n^2} = \frac{n-(n-1)}{n^2(n-1)} = \frac{1}{n^2(n-1)}.$$

Aplicando la versión con n sumandos de la propiedad (3) —hay n varianzas y $n(n-1)$ pares ordenados—,

$$\mathrm{Var}(X) = n\Big(\frac{1}{n} - \frac{1}{n^2}\Big) + n(n-1)\cdot\frac{1}{n^2(n-1)} = 1 - \frac{1}{n} + \frac{1}{n} = 1.$$

Media 1 y varianza 1, exactamente, para todo $n \geq 2$. **Comprobación con** $n=4$: con la tabla de antes, $\mathrm{E}[X^2] = \frac{0\cdot 9 + 1\cdot 8 + 4\cdot 6 + 16\cdot 1}{24} = \frac{48}{24} = 2$, luego $\mathrm{Var}(X) = 2 - 1^2 = 1$. ✓Obsérvese que las covarianzas eran positivas y minúsculas, y que sin el término cruzado el resultado habría sido $1 - \frac{1}{n}$: falso.

La desigualdad de Jensen

Definición 16.76 (Función convexa y recta soporte)**.** Una función $\phi : I \to \mathbb{R}$ definida en un intervalo es *convexa* si para todos $u, v \in I$ y $\lambda \in [0,1]$

$$\phi(\lambda u + (1-\lambda)v) \;\leq\; \lambda\phi(u) + (1-\lambda)\phi(v),$$

es decir, si la cuerda queda por encima de la gráfica. Es *cóncava* si vale la desigualdad contraria. Si ϕ es dos veces derivable, convexa equivale a $\phi'' \geq 0$.

Lema 16.77 (Recta soporte)**.** *Si ϕ es convexa en un intervalo I y a es un punto interior de I, existe una constante m tal que*

$$\phi(x) \;\geq\; \phi(a) + m\,(x-a) \qquad \textit{para todo } x \in I.$$

Demostración. La convexidad implica que la pendiente de la cuerda $\dfrac{\phi(x)-\phi(a)}{x-a}$ es una función *creciente* de x. Por tanto está acotada superiormente cuando $x<a$ e inferiormente cuando $x>a$, y existen los límites laterales

$$m_- = \lim_{x\to a^-} \frac{\phi(x)-\phi(a)}{x-a} \;\leq\; m_+ = \lim_{x\to a^+} \frac{\phi(x)-\phi(a)}{x-a}.$$

Tómese cualquier m con $m_- \leq m \leq m_+$ (si ϕ es derivable, $m=\phi'(a)$). Para $x>a$ se tiene $\frac{\phi(x)-\phi(a)}{x-a} \geq m_+ \geq m$ y multiplicar por $x-a>0$ da la desigualdad; para $x<a$ se tiene $\frac{\phi(x)-\phi(a)}{x-a} \leq m_- \leq m$ y multiplicar por $x-a<0$ invierte el sentido y da la misma conclusión. □

Teorema 16.78 (Desigualdad de Jensen)**.** *Sea X una variable aleatoria con esperanza finita y valores en un intervalo I, y sea ϕ convexa en I con $\mathrm{E}[\phi(X)]$ bien definida. Entonces*

$$\phi(\mathrm{E}[X]) \;\leq\; \mathrm{E}[\phi(X)].$$

Si ϕ es cóncava, la desigualdad se invierte. Y si ϕ es estrictamente *convexa, hay igualdad si y sólo si X es constante con probabilidad* 1*.*

Demostración. Sea $a=\mathrm{E}[X]$, que es un punto de I, y tómese la recta soporte del lema 16.77:

$$\phi(x) \;\geq\; \phi(a) + m(x-a) \qquad \text{para todo } x \in I.$$

Sustitúyase x por la variable X —la desigualdad vale punto a punto— y tómense esperanzas, que conservan el orden:

$$\mathrm{E}[\phi(X)] \;\geq\; \mathrm{E}[\phi(a)+m(X-a)] = \phi(a) + m(\mathrm{E}[X]-a) = \phi(a) = \phi(\mathrm{E}[X]),$$

donde se ha usado la linealidad de la esperanza y que $\mathrm{E}[X]-a=0$. Toda la demostración consiste en que la esperanza «no ve» la parte lineal y sí ve la curvatura.

Para el caso de igualdad con ϕ estrictamente convexa: la diferencia $D(x)=\phi(x)-\phi(a)-m(x-a)$ es no negativa y se anula únicamente en $x=a$. Si $\mathrm{E}[D(X)]=0$ con $D\geq 0$, entonces $D(X)=0$ con probabilidad 1, luego $X=a$ casi seguramente. □

Ejemplo 16.79 (Cuatro consecuencias inmediatas, con números)**. (a) La varianza es no negativa, otra vez.** Con $\phi(x)=x^2$, que es convexa, Jensen dice $\mathrm{E}[X]^2 \leq \mathrm{E}[X^2]$, y la diferencia es justamente $\mathrm{Var}(X)$.

(b) El inverso de la media no es la media de los inversos. $\phi(x) = 1/x$ es convexa en $x > 0$, luego $\mathrm{E}[1/X] \geq 1/\mathrm{E}[X]$. Con un dado equilibrado:

$$\mathrm{E}\Big[\frac{1}{X}\Big] = \frac{1}{6}\Big(1 + \frac{1}{2} + \frac{1}{3} + \frac{1}{4} + \frac{1}{5} + \frac{1}{6}\Big) = \frac{2{,}45}{6} = 0{,}408\,333 \quad \geq \quad \frac{1}{\mathrm{E}[X]} = \frac{1}{3{,}5} = 0{,}285\,714. \checkmark$$

Ésta es la razón de que la velocidad media de un trayecto de ida y vuelta sea siempre menor que la media de las dos velocidades, y de que las medias de tasas —consumo por kilómetro, precio por unidad— haya que calcularlas con cuidado.

(c) La media geométrica no supera a la aritmética. $\phi(x) = \ln x$ es *cóncava*, luego $\mathrm{E}[\ln X] \leq \ln \mathrm{E}[X]$; exponenciando y aplicándolo a una variable uniforme sobre $\{a_1, \ldots, a_n\}$ positivos queda $\sqrt[n]{a_1 \cdots a_n} \leq \frac{a_1 + \cdots + a_n}{n}$, la desigualdad entre medias, que aparece así como un caso particular de Jensen.

(d) La versión con dos puntos es la definición. Si X vale u con probabilidad λ y v con probabilidad $1 - \lambda$, Jensen es literalmente la desigualdad de la definición 16.76. Jensen es, por tanto, «la convexidad promediada sobre infinitos puntos a la vez».

Cuidado

[El sentido de la desigualdad hay que comprobarlo siempre] Convexa: $\phi(\mathrm{E}[X]) \leq \mathrm{E}[\phi(X)]$. Cóncava: al revés. El truco mnemotécnico seguro es fijar $\phi(x) = x^2$, que todo el mundo sabe que da $\mathrm{E}[X]^2 \leq \mathrm{E}[X^2]$, y recordar que x^2 es convexa. Segundo aviso: Jensen no dice nada sobre la mediana ni sobre la moda, sólo sobre la media; y exige que X tome valores en el intervalo donde ϕ es convexa. Aplicar la versión de $1/x$ a una variable que puede ser negativa produce disparates, porque $1/x$ no es convexa en toda la recta.

Cauchy–Schwarz y por qué $|\rho| \leq 1$

Teorema 16.80 (Desigualdad de Cauchy–Schwarz)**.** *Para variables aleatorias U, V con $\mathrm{E}[U^2] < \infty$ y $\mathrm{E}[V^2] < \infty$,*

$$\big(\mathrm{E}[UV]\big)^2 \;\leq\; \mathrm{E}[U^2] \cdot \mathrm{E}[V^2],$$

con igualdad si y sólo si existe una constante λ tal que $V = \lambda U$ con probabilidad 1 (o bien $U = 0$ casi seguramente).

Demostración. Supóngase $\mathrm{E}[U^2] > 0$, pues si no $U = 0$ casi seguramente y los dos miembros son 0. Considérese, para $t \in \mathbb{R}$, la función

$$q(t) = \mathrm{E}\big[(V - tU)^2\big] = \mathrm{E}[V^2] - 2t\,\mathrm{E}[UV] + t^2\,\mathrm{E}[U^2],$$

donde el desarrollo usa la linealidad de la esperanza. Por la propiedad (4) del teorema 16.73, $q(t) \geq 0$ para todo t, porque es la esperanza de un cuadrado. Es un polinomio de segundo grado en t con coeficiente principal positivo que nunca toma valores negativos, luego su discriminante no puede ser positivo:

$$4\big(\mathrm{E}[UV]\big)^2 - 4\,\mathrm{E}[U^2]\,\mathrm{E}[V^2] \;\leq\; 0,$$

que es la desigualdad. Hay igualdad exactamente cuando el discriminante es cero, es decir, cuando q tiene una raíz doble $t = \lambda$; y $q(\lambda) = \mathrm{E}[(V - \lambda U)^2] = 0$ con $(V - \lambda U)^2 \geq 0$ obliga a $V = \lambda U$ con probabilidad 1. □

Corolario 16.81 (El coeficiente de correlación vive entre -1 y 1)**.** *Sean X, Y con varianzas finitas y no nulas, y sea*

$$\rho = \rho(X, Y) = \frac{\mathrm{Cov}(X, Y)}{\sigma_X \, \sigma_Y}.$$

Entonces $|\mathrm{Cov}(X, Y)| \leq \sigma_X \sigma_Y$ y por tanto $-1 \leq \rho \leq 1$. Además $|\rho| = 1$ si y sólo si existe una relación lineal exacta $Y = aX + b$ con probabilidad 1; el signo de a coincide con el de ρ, de modo que $\rho = 1$ corresponde a una recta creciente y $\rho = -1$ a una decreciente.

Demostración. Aplíquese el teorema anterior a las variables centradas $U = X - \mu_X$ y $V = Y - \mu_Y$. Entonces $\mathrm{E}[UV] = \mathrm{Cov}(X, Y)$, $\mathrm{E}[U^2] = \mathrm{Var}(X)$ y $\mathrm{E}[V^2] = \mathrm{Var}(Y)$, y la desigualdad se lee

$$\mathrm{Cov}(X, Y)^2 \leq \mathrm{Var}(X)\,\mathrm{Var}(Y).$$

Tomando raíces y dividiendo entre $\sigma_X \sigma_Y > 0$ queda $|\rho| \leq 1$. La igualdad equivale a $Y - \mu_Y = a(X - \mu_X)$ casi seguramente, es decir $Y = aX + b$ con $b = \mu_Y - a\mu_X$; y en ese caso $\mathrm{Cov}(X, Y) = a\,\mathrm{Var}(X)$ y $\sigma_Y = |a|\sigma_X$, de donde $\rho = a/|a| = \pm 1$ según el signo de a. □

Ejemplo 16.82 (Los dos extremos)**.** Si $Y = 2X + 3$ con $\mathrm{Var}(X) = \sigma^2 > 0$, entonces $\mathrm{Cov}(X, Y) = 2\sigma^2$, $\sigma_Y = 2\sigma$ y $\rho = 2\sigma^2/(2\sigma^2) = 1$. Si $Y = -X$, entonces $\mathrm{Cov} = -\sigma^2$, $\sigma_Y = \sigma$ y $\rho = -1$. En cualquier otro caso $|\rho| < 1$ estrictamente, y cuanto más se aleja de ± 1 menos se parece la nube de puntos a una recta. Conviene retener la conclusión exacta del corolario: $|\rho| = 1$ caracteriza la dependencia *lineal* perfecta, no la dependencia a secas.

Regla: Explotar una densidad conjunta, paso a paso

1. **Dibuja el soporte** —la región donde $f > 0$— y comprueba la normalización $\iint f = 1$. Si no da 1, hay una constante que ajustar.
2. **Calcula las marginales** integrando la otra variable, con los límites que dicte el soporte. Comprueba que cada una integra 1.
3. **Decide la independencia** comparando $f(x, y)$ con $f_X(x) f_Y(y)$. Si el soporte no es un rectángulo, ya no lo son.
4. **Esperanzas y momentos** con las marginales: $\mathrm{E}[X] = \int x f_X$, $\mathrm{E}[X^2] = \int x^2 f_X$, y $\mathrm{Var} = \mathrm{E}[X^2] - \mathrm{E}[X]^2$.
5. **El momento cruzado** exige la conjunta: $\mathrm{E}[XY] = \iint xy\, f(x, y)$.
6. **Covarianza y correlación**: $\mathrm{Cov} = \mathrm{E}[XY] - \mathrm{E}[X]\,\mathrm{E}[Y]$ y $\rho = \mathrm{Cov}\,/(\sigma_X \sigma_Y)$.
7. **Comprueba** que $|\rho| \leq 1$ y que su signo cuadra con la forma de la densidad. Si sale $|\rho| > 1$, hay un error de cuenta seguro.

Ejemplo 16.83 (Una densidad conjunta, de principio a fin)**.** Sea $f(x,y) = x+y$ en el cuadrado unidad $0 \le x \le 1$, $0 \le y \le 1$, y 0 fuera.

(a) Es densidad. $\int_0^1\int_0^1 (x+y)\mathrm{d}x\,\mathrm{d}y = \int_0^1 (\frac{1}{2}+y)\mathrm{d}y = \frac{1}{2}+\frac{1}{2} = 1$. ✓

(b) Marginales. $f_X(x) = \int_0^1 (x+y)\mathrm{d}y = x+\frac{1}{2}$ para $0 \le x \le 1$, y por simetría $f_Y(y) = y+\frac{1}{2}$. Cada una integra 1. ✓

(c) No son independientes. $f_X(x)f_Y(y) = xy + \frac{x}{2} + \frac{y}{2} + \frac{1}{4} \neq x+y$. Aquí el soporte sí es un rectángulo, así que el criterio rápido no sirve y hay que multiplicar.

(d) Una probabilidad. $P(X+Y \le 1) = \int_0^1\int_0^{1-x}(x+y)\,\mathrm{d}y\,\mathrm{d}x = \int_0^1 \Big(x(1-x)+\frac{(1-x)^2}{2}\Big)\mathrm{d}x = \frac{1}{6}+\frac{1}{6} = \frac{1}{3} = 0{,}333\,333$.

(e) Esperanzas y varianzas.

$$\mathrm{E}[X] = \int_0^1 x\Big(x+\frac{1}{2}\Big)\mathrm{d}x = \frac{1}{3}+\frac{1}{4} = \frac{7}{12} = 0{,}583\,333 = \mathrm{E}[Y],$$

$$\mathrm{E}[X^2] = \int_0^1 x^2\Big(x+\frac{1}{2}\Big)\mathrm{d}x = \frac{1}{4}+\frac{1}{6} = \frac{5}{12}, \qquad \mathrm{Var}(X) = \frac{5}{12}-\frac{49}{144} = \frac{60-49}{144} = \frac{11}{144} = 0{,}076\,389,$$

y $\sigma_X = \sigma_Y = \sqrt{11}/12 = 0{,}276\,385$.

(f) Momento cruzado y covarianza.

$$\mathrm{E}[XY] = \int_0^1\int_0^1 xy(x+y)\mathrm{d}x\,\mathrm{d}y = \int_0^1 \Big(\frac{y}{3}+\frac{y^2}{2}\Big)\mathrm{d}y = \frac{1}{6}+\frac{1}{6} = \frac{1}{3},$$

$$\mathrm{Cov}(X,Y) = \frac{1}{3} - \Big(\frac{7}{12}\Big)^2 = \frac{48-49}{144} = -\frac{1}{144} = -0{,}006\,944.$$

(g) Correlación.

$$\rho = \frac{-1/144}{11/144} = -\frac{1}{11} = -0{,}090\,909.$$

Está en $[-1,1]$, como obliga el corolario 16.81. ✓

(h) Por qué sale negativa. La media condicionada es

$$\mathrm{E}[Y \mid X = x] = \frac{\int_0^1 y(x+y)\mathrm{d}y}{x+\frac{1}{2}} = \frac{3x+2}{3(2x+1)},$$

que vale $2/3 = 0{,}666\,667$ en $x=0$ y $5/9 = 0{,}555\,556$ en $x=1$: decrece. Cuando X es pequeña, la densidad $x+y$ depende sobre todo de y y empuja a Y hacia arriba; cuando X es grande, el término x domina y Y queda casi uniforme. De ahí una asociación negativa, débil —apenas un nueve por ciento de la correlación máxima— pero real.

Observación 16.84. El caso discreto de todo esto —función de masa conjunta $p(x,y)$, marginales por suma, independencia como factorización— es literalmente igual cambiando integrales por sumas, y es el que usa la sección siguiente para calcular covarianzas con dados y para introducir el coeficiente ρ. Con lo demostrado aquí, esa sección ya no necesita dar por buena ninguna de las tres fórmulas que empleaba: la linealidad, la varianza de la suma y el rango de ρ están probadas.

16.16. Dos variables a la vez, y el error que nadie calcula

Definición 16.85 (Conjunta, marginales y covarianza)**.** La función de masa conjunta es $p(x,y) = P(X = x,\ Y = y)$; las marginales se obtienen sumando en la otra variable. X e Y son independientes si $p(x,y) = p_X(x)p_Y(y)$ para todo par. Y

$$\operatorname{Cov}(X,Y) = E[XY] - E[X]E[Y], \qquad \rho = \frac{\operatorname{Cov}(X,Y)}{\sigma_X \sigma_Y} \in [-1,1].$$

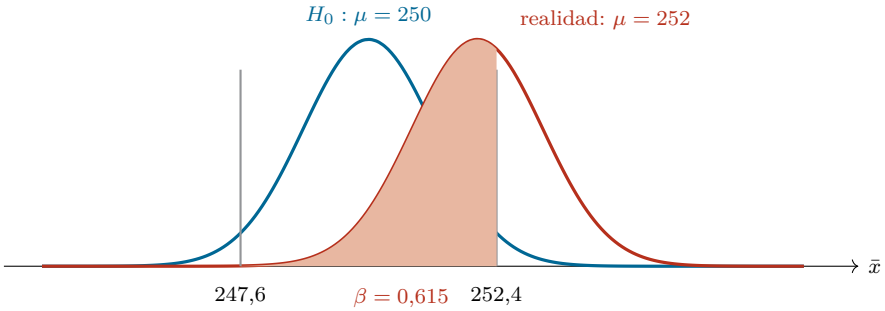


Figura 16.2: La región de aceptación se fija bajo H_0 (curva azul). Si la realidad es la curva roja, la parte sombreada —el 61,5 por ciento de su masa— cae dentro y no se rechaza: es el error de tipo II. La potencia es solo 0,385, así que el contraste pasa por alto la descalibración casi dos veces de cada tres. Con 71 paquetes en lugar de 25 subiría a 0,80.

Teorema 16.86 (La fórmula completa de la varianza de una suma)**.**

$$\operatorname{Var}(X+Y) = \operatorname{Var}(X) + \operatorname{Var}(Y) + 2\operatorname{Cov}(X,Y).$$

Ahora sí está claro por qué la versión sin el término cruzado exige independencia: la independencia hace nula la covarianza. Lo contrario no vale, y conviene tener el contraejemplo a mano.

Ejemplo 16.87 (Dos dados, y una relación perfecta con correlación cero)**. (a) X = primer dado, Y = suma de los dos.** Se tiene $E[X] = 3{,}5$, $E[Y] = 7$, y

$$\operatorname{Cov}(X,Y) = \operatorname{Cov}(X, X+Z) = \operatorname{Var}(X) = \frac{35}{12} = 2{,}916\,667,$$

porque X y el segundo dado Z son independientes. Como $\sigma_X = \sqrt{35/12}$ y $\sigma_Y = \sqrt{70/12}$,

$$\rho = \frac{35/12}{\sqrt{35/12}\sqrt{70/12}} = \frac{1}{\sqrt{2}} = 0{,}707\,107.$$

La suma está fuertemente correlacionada con cada sumando, pero no perfectamente: la mitad de su varianza viene del otro dado.

(b) El contraejemplo obligado. Sea X uniforme en $\{-1,0,1\}$ y $Y = X^2$. Entonces $E[X] = 0$ y $E[XY] = E[X^3] = 0$, luego

$$\operatorname{Cov}(X,Y) = 0, \qquad \rho = 0.$$

Y sin embargo Y está *completamente determinada* por X: son lo más dependientes que dos variables pueden ser. La correlación mide asociación *lineal* y nada más; ceros de correlación con dependencia total son la norma, no la excepción, en cuanto la relación es simétrica.

El error de tipo II, que casi nunca se calcula

Definición 16.88 (Potencia)**.** $\beta = P(\text{no rechazar} \mid H_1 \text{ cierta})$ es el error de tipo II, y la **potencia** es $1-\beta$: la probabilidad de detectar un efecto que existe de verdad. Depende de cuatro cosas: α, el tamaño real del efecto, la variabilidad y n.

Ejemplo 16.89 (Un contraste que falla dos de cada tres veces)**.** Una máquina debe llenar paquetes de 250 gramos, con $\sigma = 6$ conocida. Se toman $n = 25$ paquetes y se contrasta $H_0\colon \mu = 250$ al 5 por ciento bilateral. La región de aceptación es

$$250 \pm 1{,}96 \cdot \frac{6}{5} = [247{,}648;\ 252{,}352].$$

(a) Si la máquina está descalibrada a $\mu = 252$. La probabilidad de *no* rechazar es

$$\beta = \Phi\Big(\frac{252{,}352 - 252}{1{,}2}\Big) - \Phi\Big(\frac{247{,}648 - 252}{1{,}2}\Big) = 0{,}615\,2,$$

luego la potencia es 0,384 8. El contraste pasa por alto la descalibración casi dos veces de cada tres.

(b) Cuántos paquetes harían falta. Para potencia 0,80 frente a una desviación de 2 gramos:

$$n \ \geq\ \Big(\frac{(z_{0{,}975} + z_{0{,}80})\,\sigma}{\delta}\Big)^2 = \Big(\frac{(1{,}96 + 0{,}842)\cdot 6}{2}\Big)^2 = 70{,}64,$$

es decir, 71 paquetes en lugar de 25.

Cuidado

De aquí sale la lectura correcta de un contraste no significativo. "No se rechaza H_0" no significa "H_0 es cierta"; con potencia 0,38 significa poco más que "no había manera de verlo con estos datos". Un estudio sin cálculo previo de potencia y con resultado negativo es, casi siempre, un estudio que no ha medido nada.
Y el reverso: con n enormes, la potencia es tan alta que cualquier desviación —aunque sea de 0,01 gramos, irrelevante en la práctica— sale significativa. Por eso hay que informar siempre del *tamaño del efecto* junto al p-valor, y no del p-valor solo. La significación estadística y la importancia práctica son preguntas distintas, y solo una de las dos la responde el contraste.

16.17. Desigualdades que valen siempre: Markov, Chebyshev y Chernoff

Casi nunca se conoce la distribución exacta de lo que interesa. Lo notable es que se pueden dar cotas útiles sabiendo muy poco: sólo la media, o sólo la media y la varianza. Son las herramientas más rentables de toda la probabilidad, porque no exigen hipótesis.

Teorema 16.90 (Desigualdad de Markov)**.** *Sea $X \geq 0$ una variable aleatoria con esperanza finita. Entonces, para todo $a > 0$,*

$$P(X \geq a) \;\leq\; \frac{\mathrm{E}[X]}{a}.$$

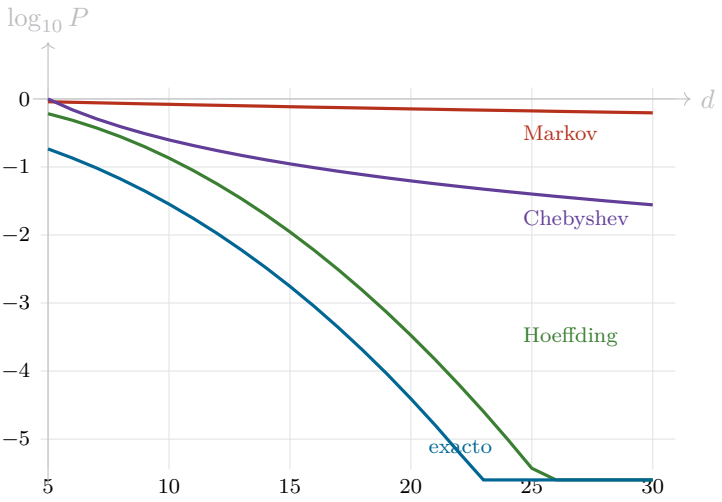


Figura 16.3: Cotas para $P(X \geq 50 + d)$ con $X \sim \mathrm{Binomial}(100, \frac{1}{2})$, en escala logarítmica. Markov y Chebyshev decaen polinómicamente y se quedan órdenes de magnitud por encima; Hoeffding decae exponencialmente y sigue de cerca al valor exacto. Es la diferencia entre usar dos momentos y usarlos todos a la vez.

Demostración. Considérese la variable indicadora del suceso $\{X \geq a\}$. Se cumple, punto por punto,

$$a \cdot \mathbf{1}_{\{X \geq a\}} \;\leq\; X,$$

porque donde el indicador vale 1 se tiene $X \geq a$, y donde vale 0 el miembro izquierdo es 0 y el derecho es ≥ 0 por hipótesis. Tomando esperanzas, que conservan las desigualdades, y usando $\mathrm{E}[\mathbf{1}_A] = P(A)$:

$$a\,P(X \geq a) \;\leq\; \mathrm{E}[X].$$

□

Teorema 16.91 (Desigualdad de Chebyshev)**.** *Si X tiene media μ y varianza σ^2 finitas, entonces para todo $k > 0$*

$$P(|X - \mu| \geq k) \;\leq\; \frac{\sigma^2}{k^2}, \qquad \text{o, con } k = t\sigma: \qquad P(|X - \mu| \geq t\sigma) \;\leq\; \frac{1}{t^2}.$$

Demostración. Aplíquese Markov a la variable no negativa $Y = (X - \mu)^2$ con umbral $a = k^2$:

$$P(|X - \mu| \geq k) = P(Y \geq k^2) \leq \frac{\mathrm{E}[Y]}{k^2} = \frac{\sigma^2}{k^2}.$$

□

Observación 16.92 (Son óptimas, y por eso son flojas)*.* Ambas desigualdades se alcanzan con igualdad para distribuciones concretas —Markov con la que vale 0 o a; Chebyshev con la que

vale μ, $\mu - k$ y $\mu + k$ con las probabilidades adecuadas—, de modo que *no se pueden mejorar* sin añadir hipótesis. Que resulten flojas en los casos habituales no es un defecto suyo: es el precio de no suponer nada. La lección práctica es que si se sabe más, hay que usar más, y ahí entra Chernoff.

Teorema 16.93 (Cota de Chernoff)**.** *Sea X una variable con función generadora de momentos $M_X(t) = \mathrm{E}[\mathrm{e}^{tX}]$ finita en un entorno de 0. Entonces, para todo a,*

$$P(X \geq a) \;\leq\; \inf_{t>0} \mathrm{e}^{-ta}\, M_X(t), \qquad P(X \leq a) \;\leq\; \inf_{t<0} \mathrm{e}^{-ta}\, M_X(t).$$

Demostración. Para $t > 0$, la función $x \mapsto \mathrm{e}^{tx}$ es estrictamente creciente, de modo que $\{X \geq a\} = \{\mathrm{e}^{tX} \geq \mathrm{e}^{ta}\}$. Aplicando Markov a la variable no negativa e^{tX}:

$$P(X \geq a) = P\big(\mathrm{e}^{tX} \geq \mathrm{e}^{ta}\big) \leq \frac{\mathrm{E}[\mathrm{e}^{tX}]}{\mathrm{e}^{ta}} = \mathrm{e}^{-ta} M_X(t).$$

Como vale para todo $t > 0$, vale para el ínfimo. Para la cola izquierda se usa $t < 0$, que invierte la desigualdad al exponenciar. □

Corolario 16.94 (Hoeffding para variables acotadas)**.** *Si $X_1, \dots, X_n$ son independientes con $a_i \leq X_i \leq b_i$ y $S = \sum X_i$, entonces*

$$P\big(S - \mathrm{E}[S] \geq s\big) \;\leq\; \exp\!\left(-\frac{2s^2}{\sum_i (b_i - a_i)^2}\right).$$

En particular, para n variables en $[0,1]$, $P(\bar{X} - \mu \geq \varepsilon) \leq \mathrm{e}^{-2n\varepsilon^2}$.

Ejemplo 16.95 (Las cuatro cotas, comparadas)**.** Lánzese una moneda equilibrada 100 veces y pregúntese por $P(X \geq 70)$, donde X es el número de caras. Aquí $\mathrm{E}[X] = 50$ y $\mathrm{Var}(X) = 25$.

Markov	$50/70$	$= 0{,}714$
Chebyshev	$25/20^2$	$= 0{,}0625$
Hoeffding	$\mathrm{e}^{-2\cdot 20^2/100}$	$= 3{,}35 \cdot 10^{-4}$
Chernoff óptimo	$\mathrm{e}^{-100\, D(0{,}7\,\|\|\,0{,}5)}$	$= 2{,}67 \cdot 10^{-4}$
Valor exacto		$= 3{,}93 \cdot 10^{-5}$

donde $D(p\|q) = p\log\frac{p}{q} + (1-p)\log\frac{1-p}{1-q}$. Markov se equivoca por un factor de 18 000; Chebyshev, por 1 600; las dos cotas exponenciales, por menos de 10. La moraleja es cuantitativa: Markov y Chebyshev sólo usan uno o dos momentos, y su decaimiento es polinómico; Chernoff usa *todos* los momentos a la vez —eso es la función generadora— y por eso decae exponencialmente. Cuando se necesita una cota que sirva de verdad, hay que pagar el precio de conocer la generatriz.

Idea clave

Dónde aparece esto en la práctica. Toda la teoría de la concentración descansa en Chernoff: el tamaño de muestra necesario para que una encuesta tenga error $\le \varepsilon$ con confianza $1-\delta$ sale de $\mathrm{e}^{-2n\varepsilon^2} \le \delta$, es decir $n \ge \frac{\log(1/\delta)}{2\varepsilon^2}$, y nótese que *no depende del tamaño de la población*. Con $\varepsilon = 0{,}03$ y $\delta = 0{,}05$ salen $n \ge 1\,665$ personas, tanto para un municipio como para un país entero. Esa es la respuesta matemática a la pregunta "¿cómo pueden mil personas representar a cuarenta millones?".

16.18. Funciones generadoras de momentos y funciones características

Una distribución se puede describir por su función de densidad, por su función de distribución, o por una tercera vía que convierte las convoluciones en productos y los momentos en derivadas. Es la más potente de las tres para trabajar con sumas.

Definición 16.96 (Función generadora de momentos)**.** La *función generadora de momentos* de X es

$$M_X(t) = \mathrm{E}\big[\mathrm{e}^{tX}\big],$$

definida para los t en que la esperanza es finita. Si M_X es finita en un entorno de 0, se dice que X *tiene generatriz*.

Teorema 16.97 (Propiedades)**.** *Supóngase que M_X es finita en $(-h,h)$ para algún $h>0$. Entonces:*

1. *M_X es infinitamente derivable en $(-h,h)$ y $\mathrm{E}[X^k] = M_X^{(k)}(0)$; de hecho $M_X(t) = \sum_{k\ge 0} \frac{\mathrm{E}[X^k]}{k!} t^k$.*
2. *$M_{aX+b}(t) = \mathrm{e}^{bt} M_X(at)$.*
3. *Si X e Y son* independientes*, entonces $M_{X+Y}(t) = M_X(t)\, M_Y(t)$.*
4. (Unicidad.) *Si $M_X = M_Y$ en un entorno de 0, entonces X e Y tienen la misma distribución.*

Demostración. (1) La finitud en $(-h,h)$ permite derivar bajo el signo de esperanza (convergencia dominada, con dominante $|X|^k \mathrm{e}^{t_0|X|}$ para $|t|<t_0<h$), y $\frac{\mathrm{d}^k}{\mathrm{d}t^k}\mathrm{e}^{tX} = X^k \mathrm{e}^{tX}$, que en $t=0$ vale X^k. La serie sale de desarrollar e^{tX} e intercambiar.

(2) $\mathrm{E}[\mathrm{e}^{t(aX+b)}] = \mathrm{e}^{tb}\,\mathrm{E}[\mathrm{e}^{(ta)X}]$.

(3) Si $X \perp Y$, también lo son e^{tX} y e^{tY}, y la esperanza de un producto de independientes es el producto de las esperanzas.

(4) Es el resultado profundo del grupo, y es esencialmente un teorema de inversión de la transformada de Laplace; no se demuestra aquí. Su análogo para funciones características se demuestra con la fórmula de inversión de Fourier del capítulo 31. □

Regla: Catálogo de generatrices

Distribución	$M_X(t)$	dominio
Bernoulli(p)	$1-p+p\,\mathrm{e}^t$	todo t
Binomial(n,p)	$\left(1-p+p\,\mathrm{e}^t\right)^n$	todo t
Poisson(λ)	$\exp(\lambda(\mathrm{e}^t-1))$	todo t
Geométrica(p)	$\dfrac{p\,\mathrm{e}^t}{1-(1-p)\mathrm{e}^t}$	$t<-\log(1-p)$
Exponencial(λ)	$\dfrac{\lambda}{\lambda-t}$	$t<\lambda$
Gamma(α,λ)	$\left(\dfrac{\lambda}{\lambda-t}\right)^\alpha$	$t<\lambda$
Normal(μ,σ^2)	$\exp(\mu t+\frac{1}{2}\sigma^2t^2)$	todo t
Uniforme(a,b)	$\dfrac{\mathrm{e}^{tb}-\mathrm{e}^{ta}}{t(b-a)}$	todo t

Ejemplo 16.98 (Cómo se deduce la de la normal, y qué se saca de ella)**.** Para $Z\sim N(0,1)$, completando el cuadrado en el exponente,

$$M_Z(t)=\int_{\mathbb{R}}\mathrm{e}^{tz}\frac{\mathrm{e}^{-z^2/2}}{\sqrt{2\pi}}\mathrm{d}z=\mathrm{e}^{t^2/2}\int_{\mathbb{R}}\frac{\mathrm{e}^{-(z-t)^2/2}}{\sqrt{2\pi}}\mathrm{d}z=\mathrm{e}^{t^2/2},$$

porque la última integral es la de una densidad normal de media t. Con la propiedad (2), $X=\mu+\sigma Z$ da $M_X(t)=\mathrm{e}^{\mu t+\sigma^2t^2/2}$. De aquí salen los momentos sin integrar nada: $M_Z(t)=\sum_k\frac{t^{2k}}{2^k k!}$, y comparando con $\sum_j\frac{\mathrm{E}[Z^j]}{j!}t^j$ resulta que todos los momentos impares son nulos y

$$\mathrm{E}[Z^{2k}]=\frac{(2k)!}{2^k k!}=(2k-1)!!=1\cdot3\cdot5\cdots(2k-1).$$

Así $\mathrm{E}[Z^4]=3$, $\mathrm{E}[Z^6]=15$: la curtosis de la normal es 3, que es el patrón contra el que se compara todo lo demás.

Ejemplo 16.99 (Sumas, en tres líneas)**.** La propiedad (3) más la unicidad (4) resuelven de un golpe media docena de resultados que directamente costarían páginas de convoluciones:

- *Poisson + Poisson.* Si $X\sim$ Poisson(λ) e $Y\sim$ Poisson(μ) independientes, entonces $M_{X+Y}(t)=\mathrm{e}^{\lambda(\mathrm{e}^t-1)}\mathrm{e}^{\mu(\mathrm{e}^t-1)}=\mathrm{e}^{(\lambda+\mu)(\mathrm{e}^t-1)}$: es Poisson$(\lambda+\mu)$.
- *Normal + Normal.* $M_{X+Y}(t)=\mathrm{e}^{(\mu_1+\mu_2)t+\frac{1}{2}(\sigma_1^2+\sigma_2^2)t^2}$: normal, con medias y *varianzas* que se suman.
- *Suma de n exponenciales(λ).* El producto de n copias de $\frac{\lambda}{\lambda-t}$ es $\left(\frac{\lambda}{\lambda-t}\right)^n$: es Gamma$(n,\lambda)$, la distribución del tiempo hasta el n-ésimo suceso de un proceso de Poisson.
- *Binomial + Binomial* con la misma p: binomial, con los n sumados. Con p distintas, *no*: los factores no se combinan, y de hecho la suma no es binomial.

Este último punto es el que más se equivoca en los exámenes, y la generatriz lo detecta sola.

Definición 16.100 (Función característica)**.** $\varphi_X(t) = \mathrm{E}\big[\mathrm{e}^{\mathrm{i}tX}\big]$.

Observación 16.101 (Por qué hace falta la versión compleja)*.* La función característica *siempre* existe, para toda variable aleatoria y todo t, porque $|\mathrm{e}^{\mathrm{i}tX}| = 1$ y la esperanza de algo de módulo 1 está acotada por 1. La generatriz, en cambio, puede no existir: la distribución de Cauchy, de densidad $\frac{1}{\pi(1+x^2)}$, tiene $M_X(t) = \infty$ para todo $t \neq 0$ —ni siquiera tiene media—, pero su función característica es $\varphi_X(t) = \mathrm{e}^{-|t|}$, perfectamente definida. Y la lognormal tiene todos los momentos finitos y aun así su generatriz diverge para $t > 0$, lo que muestra que "tener todos los momentos" no basta.

Además, φ_X no es otra cosa que la transformada de Fourier de la densidad (capítulo 31), con lo que todo el arsenal de aquel capítulo —inversión, Plancherel, el hecho de que la transformada de una convolución es un producto— se traslada íntegro a la probabilidad. El teorema de continuidad de Lévy dice además que $\varphi_{X_n} \to \varphi_X$ puntualmente equivale a $X_n \to X$ en distribución, y ese es el teorema que hace del método un método.

Definición 16.102 (Generatriz de cumulantes)**.** $K_X(t) = \log M_X(t)$. Sus derivadas en 0 son los *cumulantes*: $K'(0) = \mu$, $K''(0) = \sigma^2$, $K'''(0) = \mathrm{E}[(X-\mu)^3]$ (la asimetría sin normalizar), y $K^{(4)}(0)$ es la curtosis en exceso multiplicada por σ^4.

Observación 16.103. Los cumulantes son *aditivos* sobre sumas de independientes, porque el logaritmo convierte el producto de generatrices en suma. Para la normal, $K(t) = \mu t + \frac{1}{2}\sigma^2 t^2$: todos los cumulantes de orden ≥ 3 son nulos. Eso caracteriza a la normal, y es la manera más limpia de entender por qué es el atractor del teorema central del límite: al sumar y reescalar, los cumulantes de orden alto se desvanecen más deprisa que los de orden dos.

16.19. Modos de convergencia y las leyes de los grandes números

Hasta ahora "la media muestral se acerca a μ" se ha usado sin decir en qué sentido. Hay cuatro sentidos distintos, y no son equivalentes.

Definición 16.104 (Los cuatro modos)**.** Sea (X_n) una sucesión de variables aleatorias y X otra.

- *Casi segura*, $X_n \to X$ c.s.: $P\big(\lim_n X_n = X\big) = 1$.
- *En probabilidad*, $X_n \xrightarrow{P} X$: para todo $\varepsilon > 0$, $P(|X_n - X| > \varepsilon) \to 0$.
- *En media p-ésima*, $X_n \xrightarrow{L^p} X$: $\mathrm{E}[|X_n - X|^p] \to 0$.
- *En distribución*, $X_n \xrightarrow{d} X$: $F_{X_n}(x) \to F_X(x)$ en todo punto de continuidad de F_X.

Teorema 16.105 (Implicaciones)**.**

$$\textit{c.s.} \implies \textit{en probabilidad} \implies \textit{en distribución}, \qquad L^p \implies \textit{en probabilidad},$$

y ninguna implicación se invierte en general. Sí se invierten en dos casos: si el límite es una constante, convergencia en distribución implica convergencia en probabilidad; y de convergencia en probabilidad siempre se puede extraer una subsucesión que converge casi seguro.

Demostración de las implicaciones. $L^p \Rightarrow P$*:* es Markov aplicado a $|X_n - X|^p$: $P(|X_n - X| > \varepsilon) \le \mathrm{E}[|X_n - X|^p]/\varepsilon^p \to 0$.

c.s.$\Rightarrow P$*:* sea $A_n = \{|X_n - X| > \varepsilon\}$ y $B_n = \bigcup_{m\ge n} A_m$. Si $X_n \to X$ c.s., el conjunto $\bigcap_n B_n$ está contenido en el de no convergencia, que es nulo; como B_n decrece, $P(B_n) \to 0$, y $P(A_n) \le P(B_n)$.

$P \Rightarrow d$*:* sea x punto de continuidad de F_X y $\varepsilon > 0$. Entonces

$$F_{X_n}(x) = P(X_n \le x) \le P(X \le x + \varepsilon) + P(|X_n - X| > \varepsilon) \to F_X(x + \varepsilon),$$

y simétricamente $\liminf F_{X_n}(x) \ge F_X(x - \varepsilon)$. Haciendo $\varepsilon \to 0$ y usando la continuidad en x, ambos extremos convergen a $F_X(x)$. □

Ejemplo 16.106 (Los contraejemplos que hay que conocer)**.**

1. *En probabilidad pero no casi seguro (la "máquina de escribir").* En $[0,1]$ con la medida de Lebesgue, tómense los indicadores de los intervalos $[0,1], [0,\frac12], [\frac12,1], [0,\frac13], [\frac13,\frac23], \ldots$, recorriendo cada nivel de izquierda a derecha. Las longitudes tienden a 0, luego $X_n \xrightarrow{P} 0$. Pero en cada punto x la sucesión vale 1 infinitas veces, así que no converge en ningún punto.

2. *Casi seguro pero no en* L^1*.* $X_n = n \cdot \mathbf{1}_{[0,1/n]}$. Converge a 0 en todo punto salvo en 0, pero $\mathrm{E}[X_n] = 1$ para todo n. La masa se escapa hacia arriba.

3. *En distribución pero no en probabilidad.* Sea $X \sim N(0,1)$ y $X_n = -X$ para todo n. Como la normal es simétrica, X_n tiene exactamente la misma distribución que X, luego $X_n \xrightarrow{d} X$; pero $|X_n - X| = 2|X|$, que no tiende a 0 en probabilidad. La convergencia en distribución no dice nada sobre *qué* variable se aproxima, sólo sobre su ley.

Lema 16.107 (Borel–Cantelli)**.** *Sea* (A_n) *una sucesión de sucesos y* $A = \limsup_n A_n = \bigcap_n \bigcup_{m\ge n} A_m$ *el suceso "ocurren infinitos* A_n*".*

1. *Si* $\sum_n P(A_n) < \infty$*, entonces* $P(A) = 0$*.*
2. *Si los* A_n *son* independientes *y* $\sum_n P(A_n) = \infty$*, entonces* $P(A) = 1$*.*

Demostración. (1) $P(A) \le P\big(\bigcup_{m\ge n} A_m\big) \le \sum_{m\ge n} P(A_m)$, que es la cola de una serie convergente y tiende a 0.

(2) Basta ver que $P\big(\bigcap_{m\ge n} A_m^c\big) = 0$ para cada n. Por independencia y usando $1 - x \le \mathrm{e}^{-x}$,

$$P\Big(\bigcap_{m=n}^{N} A_m^c\Big) = \prod_{m=n}^{N} (1 - P(A_m)) \le \exp\Big(-\sum_{m=n}^{N} P(A_m)\Big) \xrightarrow[N\to\infty]{} 0. \qquad \square$$

Teorema 16.108 (Ley débil de los grandes números)**.** *Sean* $X_1, X_2, \ldots$ *independientes e idénticamente distribuidas con media* μ *y varianza* $\sigma^2 < \infty$*, y* $\bar X_n = \frac1n \sum_{i\le n} X_i$*. Entonces* $\bar X_n \xrightarrow{P} \mu$*.*

Demostración. $\mathrm{E}[\bar X_n]=\mu$ y, por independencia, $\mathrm{Var}(\bar X_n)=\frac{\sigma^2}{n}$. Chebyshev da

$$P\big(|\bar X_n-\mu|\ge\varepsilon\big)\ \le\ \frac{\sigma^2}{n\varepsilon^2}\xrightarrow[n\to\infty]{}0.$$ □

Teorema 16.109 (Ley fuerte de los grandes números)**.** *Con X_i independientes e idénticamente distribuidas y $\mathrm{E}\,|X_1|<\infty$ (¡sin hipótesis sobre la varianza!), se tiene $\bar X_n\to\mu$ casi seguro.*

Demostración bajo la hipótesis extra de cuarto momento finito. Supóngase $\mu=0$ sin pérdida de generalidad y sea $S_n=\sum_{i\le n}X_i$. Al desarrollar $\mathrm{E}[S_n^4]$, los términos con algún factor de exponente 1 tienen esperanza nula por independencia y centrado; sobreviven los n términos $\mathrm{E}[X_i^4]$ y los $3n(n-1)$ términos $\mathrm{E}[X_i^2X_j^2]=\sigma^4$ con $i\neq j$. Por tanto

$$\mathrm{E}[S_n^4]=n\,\mathrm{E}[X_1^4]+3n(n-1)\sigma^4\ \le\ Cn^2.$$

Ahora Markov: $P(|\bar X_n|>\varepsilon)=P(S_n^4>n^4\varepsilon^4)\le\frac{Cn^2}{n^4\varepsilon^4}=\frac{C}{n^2\varepsilon^4}$, cuya serie en n converge. Por el lema 16.107(1), con probabilidad 1 sólo ocurren finitos sucesos $\{|\bar X_n|>\varepsilon\}$, es decir $\limsup|\bar X_n|\le\varepsilon$ casi seguro. Intersecando sobre $\varepsilon=1/k$ se obtiene $\bar X_n\to0$ c.s.

La versión general, con sólo $\mathrm{E}\,|X_1|<\infty$, es el teorema de Kolmogórov; su demostración usa truncamiento y la desigualdad maximal de Kolmogórov, y es considerablemente más larga. □

Observación 16.110 (Qué separa la débil de la fuerte). La ley débil dice que *para cada n grande*, es improbable que $\bar X_n$ esté lejos de μ. La fuerte dice que, con probabilidad 1, *la sucesión entera* acaba entrando en cualquier banda alrededor de μ y no vuelve a salir. La primera admite que $\bar X_n$ se desvíe infinitas veces, siempre que sea cada vez más raro; la segunda lo prohíbe. Para un casino, la diferencia es la que hay entre "cada noche concreta probablemente gano" y "a lo largo de mi vida, con certeza, gano".

16.20. El teorema central del límite, demostrado

Se ha citado varias veces. Aquí está el enunciado preciso y la demostración.

Teorema 16.111 (Central del límite, Lindeberg–Lévy)**.** *Sean $X_1,X_2,\dots$ independientes e idénticamente distribuidas con media μ y varianza $\sigma^2\in(0,\infty)$. Entonces*

$$Z_n=\frac{S_n-n\mu}{\sigma\sqrt{n}}=\frac{\bar X_n-\mu}{\sigma/\sqrt{n}}\ \xrightarrow{\ d\ }\ N(0,1).$$

Demostración con la generatriz. Supóngase primero que M_X es finita en un entorno de 0. Sea $Y_i=\frac{X_i-\mu}{\sigma}$, de modo que $\mathrm{E}[Y_i]=0$, $\mathrm{Var}(Y_i)=1$ y $Z_n=\frac{1}{\sqrt n}\sum_{i\le n}Y_i$. Por las propiedades (2) y (3) del teorema 16.97,

$$M_{Z_n}(t)=\Big[M_Y\Big(\frac{t}{\sqrt n}\Big)\Big]^n.$$

Tómense logaritmos y úsese la generatriz de cumulantes $K_Y=\log M_Y$, que cumple $K_Y(0)=0$, $K_Y'(0)=\mathrm{E}[Y]=0$ y $K_Y''(0)=\mathrm{Var}(Y)=1$. Por Taylor con resto de Peano,

$$K_Y(s)=\frac{s^2}{2}+o(s^2)\qquad(s\to0).$$

Sustituyendo $s = t/\sqrt{n}$:

$$\log M_{Z_n}(t) = n\,K_Y\Big(\frac{t}{\sqrt{n}}\Big) = n\left[\frac{t^2}{2n} + o\Big(\frac{t^2}{n}\Big)\right] = \frac{t^2}{2} + o(1) \xrightarrow[n\to\infty]{} \frac{t^2}{2}.$$

Es decir, $M_{Z_n}(t) \to \mathrm{e}^{t^2/2}$, que es la generatriz de la $N(0,1)$. El teorema de continuidad —la convergencia puntual de las generatrices en un entorno de 0 implica convergencia en distribución— concluye. □

Cuidado

[Lo que esta demostración esconde] La prueba anterior supone que existe la generatriz, lo cual el enunciado del teorema *no* exige: sólo se pide varianza finita. La demostración honesta se hace idéntica pero con la función característica $\varphi_Y(t)$, que siempre existe, usando el desarrollo $\varphi_Y(s) = 1 - \frac{s^2}{2} + o(s^2)$ —que sólo requiere dos momentos— y el teorema de continuidad de Lévy en lugar del de las generatrices. La estructura del argumento es exactamente la misma; lo que cambia es el trabajo técnico de justificar el desarrollo con dos momentos y de probar Lévy, que descansa en la fórmula de inversión de Fourier. Se ha presentado la versión con generatrices porque es la que hace visible *por qué* el teorema es cierto: al normalizar, todos los cumulantes de orden ≥ 3 se dividen por potencias de n mayores que la que los mantendría vivos, y sólo sobrevive el de orden 2. La normal es lo único que queda cuando se borran todos los cumulantes altos.

Teorema 16.112 (Berry–Esseen: la velocidad)**.** *Si además* $\rho = \mathrm{E}\,|X_1 - \mu|^3 < \infty$*, entonces*

$$\sup_{x\in\mathbb{R}} \big|P(Z_n \leq x) - \Phi(x)\big| \;\leq\; \frac{C\rho}{\sigma^3\sqrt{n}},$$

con una constante universal C *(se sabe que* $C \leq 0{,}4748$*).*

Ejemplo 16.113 (La cota, medida)**.** Para X_i de Bernoulli con $p = \frac{1}{2}$ se tiene $\sigma = \frac{1}{2}$ y $\rho = \frac{1}{8}$, luego la cota es $0{,}4748/\sqrt{n}$. El error máximo real, comparando la binomial exacta con la normal:

n	10	100	1000
error real	0,1231	0,0398	0,0126
cota $0{,}4748/\sqrt{n}$	0,1501	0,0475	0,0150

La cota se cumple y no está lejos de ser fina. Y obsérvese la velocidad: para dividir el error entre 10 hacen falta 100 veces más datos. Ese $\sqrt{n}$ es el impuesto que paga toda la estadística.

Ejemplo 16.114 (Doce uniformes y una normal)**.** Un truco clásico para simular normales antes de que existiera el método de Box–Muller: si $U_1, \ldots, U_{12}$ son uniformes en $(0,1)$ independientes, entonces $S = \sum U_i - 6$ tiene media 0 y varianza $12 \cdot \frac{1}{12} = 1$ exactamente. Con 400 000 simulaciones se obtiene media $-0{,}001$ y varianza 1,002, y las colas:

t	1	2	3
$P(\lvert S\rvert > t)$ simulado	0,3217	0,0449	0,00197
$P(\lvert Z\rvert > t)$ normal	0,3173	0,0455	0,00270

Excelente en el centro, y visiblemente peor en la cola: la suma de doce uniformes está acotada entre -6 y 6, de modo que a partir de ahí la aproximación es cero y la normal no. Es la advertencia general: el teorema central del límite es un enunciado sobre el *centro* de la distribución, y aproximar colas lejanas con él es un error frecuente y caro. Para las colas están las cotas de Chernoff.

Corolario 16.115 (De Moivre–Laplace y la corrección de continuidad)**.** *Si $X \sim \mathrm{Binomial}(n,p)$, entonces $\frac{X-np}{\sqrt{np(1-p)}} \xrightarrow{d} N(0,1)$. Al aproximar $P(a \le X \le b)$ conviene usar $P\big(\frac{a-0{,}5-np}{\sqrt{npq}} \le Z \le \frac{b+0{,}5-np}{\sqrt{npq}}\big)$: la* corrección de continuidad*, que reparte cada barra del histograma medio punto a cada lado y reduce el error de forma apreciable para n moderado.*

16.21. Transformar y sumar variables aleatorias

Dos operaciones aparecen sin cesar: aplicar una función a una variable, y sumar dos variables. Cada una tiene su fórmula.

Teorema 16.116 (Cambio de variable)**.** *Sea X continua con densidad f_X y sea g derivable y estrictamente monótona sobre el soporte de X. Entonces $Y = g(X)$ tiene densidad*

$$f_Y(y) = f_X(g^{-1}(y))\left|\frac{\mathrm{d}}{\mathrm{d}y} g^{-1}(y)\right| = \left.\frac{f_X(x)}{|g'(x)|}\right|_{x=g^{-1}(y)}.$$

Si g no es monótona, se suman las contribuciones de cada rama.

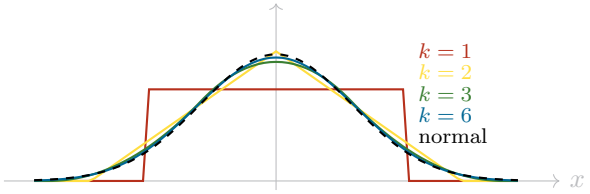


Figura 16.4: Densidad de la suma de k uniformes, tipificada a media 0 y varianza 1. Con $k = 1$ es un rectángulo; con $k = 2$, un triángulo; con $k = 3$ ya es suave; con $k = 6$ resulta prácticamente indistinguible de la normal (a trazos). Cada convolución añade una derivada de regularidad: el teorema central del límite, visto como un proceso de suavizado.

Demostración. Si g es creciente, $F_Y(y) = P(g(X) \le y) = P(X \le g^{-1}(y)) = F_X(g^{-1}(y))$, y derivando por la regla de la cadena sale la fórmula. Si es decreciente, $F_Y(y) = 1 - F_X(g^{-1}(y))$ y el signo del factor se compensa con el valor absoluto. □

Ejemplo 16.117 (Tres transformaciones que hay que tener a mano)**.** 1. *Cuadrado de una normal.* Si $Z \sim N(0,1)$ e $Y = Z^2$, la función no es monótona y hay dos ramas, $\pm\sqrt{y}$, cada una con $|g'| = 2\sqrt{y}$:

$$f_Y(y) = 2 \cdot \frac{1}{2\sqrt{y}} \cdot \frac{\mathrm{e}^{-y/2}}{\sqrt{2\pi}} = \frac{y^{-1/2}\mathrm{e}^{-y/2}}{\sqrt{2\pi}}, \qquad y > 0,$$

que es la densidad χ^2 con un grado de libertad, es decir Gamma$(\frac{1}{2}, \frac{1}{2})$.

2. *Exponencial de una normal.* Si $Y = \mathrm{e}^X$ con $X \sim N(\mu, \sigma^2)$, sale la *lognormal*, $f_Y(y) = \frac{1}{y\sigma\sqrt{2\pi}} \exp(-\frac{(\log y - \mu)^2}{2\sigma^2})$. Es el modelo estándar para precios de activos y para magnitudes que resultan de *multiplicar* muchos factores pequeños, igual que la normal lo es de sumarlos.

3. *Transformación integral de probabilidad.* Si F es continua y estrictamente creciente y $U = F(X)$, entonces $F_U(u) = P(F(X) \le u) = P(X \le F^{-1}(u)) = u$: U es *uniforme* en $(0,1)$, sea cual sea X. Leído al revés, $X = F^{-1}(U)$ genera cualquier distribución a partir de números aleatorios uniformes, y ese es el método de la transformada inversa con el que se simula casi todo.

Teorema 16.118 (Convolución)**.** *Si X e Y son independientes con densidades f_X y f_Y, entonces $S = X + Y$ tiene densidad*

$$f_S(s) = \int_{-\infty}^{\infty} f_X(x)\, f_Y(s-x)\,\mathrm{d}x = (f_X * f_Y)(s).$$

En el caso discreto, $P(S = s) = \sum_k P(X = k)P(Y = s - k)$.

Demostración. Por independencia, la densidad conjunta es el producto. Entonces

$$F_S(s) = \iint_{\{x+y\le s\}} f_X(x) f_Y(y)\,\mathrm{d}x\,\mathrm{d}y = \int_{\mathbb{R}} f_X(x) \left(\int_{-\infty}^{s-x} f_Y(y)\mathrm{d}y \right) \mathrm{d}x = \int_{\mathbb{R}} f_X(x) F_Y(s-x)\,\mathrm{d}x,$$

por Fubini (capítulo 31), y derivando respecto de s bajo el signo integral se obtiene la fórmula. □

Ejemplo 16.119 (Dos uniformes dan un triángulo)**.** Si $X, Y \sim U(0,1)$ independientes, entonces para $0 \le s \le 1$

$$f_S(s) = \int_0^1 \mathbf{1}_{[0,1]}(s-x)\mathrm{d}x = \int_0^s \mathrm{d}x = s,$$

y por simetría $f_S(s) = 2 - s$ para $1 \le s \le 2$: la densidad triangular. Con tres uniformes salen tres trozos de parábola; con doce, algo indistinguible de una campana, que es lo que hacía utilizable el truco de la sección anterior. La convolución *suaviza*, y ese es exactamente el mecanismo que hay detrás del teorema central del límite: cada convolución añade una derivada de regularidad y acerca el perfil a la gaussiana, que es el único punto fijo de la operación "convolucionar y reescalar".

Regla: Sumar variables: qué método usar

1. Si sólo se necesitan la media y la varianza, no hace falta la distribución: $\mathrm{E}[X+Y] = \mathrm{E}\,X + \mathrm{E}\,Y$ *siempre*, y $\mathrm{Var}(X+Y) = \mathrm{Var}\,X + \mathrm{Var}\,Y + 2\,\mathrm{Cov}(X,Y)$, que se simplifica sólo si son incorreladas.

2. Si las distribuciones están en el catálogo de generatrices, multiplíquense las

generatrices y reconózcase el resultado. Es el camino más corto.

3. Si no, convolución directa. Es un cálculo, a veces desagradable.
4. Si sólo se necesita una aproximación y n es grande, teorema central del límite; y para las colas, Chernoff.
5. Para más de dos variables con transformación conjunta, cambio de variable con jacobiano (capítulo 9): $f_{U,V}(u,v) = f_{X,Y}(x,y)\,|\det J|^{-1}$.

Ejemplo 16.120 (Box–Muller, con jacobiano)**.** Sean U_1, U_2 uniformes independientes en $(0,1)$ y

$$Z_1 = \sqrt{-2\log U_1}\,\cos(2\pi U_2), \qquad Z_2 = \sqrt{-2\log U_1}\,\sin(2\pi U_2).$$

Entonces Z_1, Z_2 son normales estándar independientes. La razón se ve al revés: en coordenadas polares (R,Θ) con $Z_1 = R\cos\Theta$, $Z_2 = R\sin\Theta$, la densidad conjunta de dos normales estándar es $\frac{1}{2\pi}\mathrm{e}^{-(z_1^2+z_2^2)/2}$, y el jacobiano del cambio a polares aporta un factor r:

$$f_{R,\Theta}(r,\theta) = \frac{r\,\mathrm{e}^{-r^2/2}}{2\pi},$$

que se factoriza: Θ es uniforme en $(0,2\pi)$ y R tiene densidad $r\mathrm{e}^{-r^2/2}$, es decir $R^2 \sim$ Exponencial$(\frac{1}{2})$, que se genera como $-2\log U_1$. La simetría radial de la gaussiana —que es la única densidad con simetría radial y coordenadas independientes— es lo que hace funcionar el método.

16.22. La uniforme discreta, la geométrica y la hipergeométrica

La regla de diagnóstico de la sección anterior nombraba siete distribuciones, pero solo tres de ellas —la binomial, la de Poisson y la exponencial— habían recibido una definición formal con su esperanza y su varianza. De la geométrica se dio la fórmula de la masa sin decir de dónde salía y sin su varianza; de la hipergeométrica, únicamente el nombre y un ejemplo con cartas; y la uniforme discreta, que es la que rige el lanzamiento de un dado y aparece en la primera página del capítulo, no se definió nunca. Esta sección cierra ese catálogo: define las tres que faltan, demuestra sus parámetros y las relaciona entre sí, de modo que la tabla de reconocimiento pueda usarse sin recurrir a ninguna otra fuente.

Definición 16.121 (Distribución uniforme discreta)**.** Una variable aleatoria X sigue una **distribución uniforme discreta** sobre los enteros $1, 2, \ldots, N$ si

$$P(X = k) = \frac{1}{N}, \qquad k = 1, 2, \ldots, N.$$

Es el modelo de todo mecanismo perfectamente equilibrado: un dado, una ruleta con casillas idénticas, un sorteo entre N papeletas.

Proposición 16.122. *Si X es uniforme discreta sobre $\{1,\dots,N\}$, entonces*

$$E(X)=\frac{N+1}{2},\qquad \operatorname{Var}(X)=\frac{N^2-1}{12}.$$

Demostración. La esperanza es la media aritmética de los N primeros naturales:

$$E(X)=\frac{1}{N}\sum_{k=1}^{N}k=\frac{1}{N}\cdot\frac{N(N+1)}{2}=\frac{N+1}{2}.$$

Para la varianza usamos $\sum_{k=1}^{N}k^2=\frac{N(N+1)(2N+1)}{6}$, que es la suma de cuadrados de las sucesiones y series elementales:

$$E(X^2)=\frac{1}{N}\cdot\frac{N(N+1)(2N+1)}{6}=\frac{(N+1)(2N+1)}{6},$$

y por tanto

$$\begin{aligned}\operatorname{Var}(X)&=\frac{(N+1)(2N+1)}{6}-\frac{(N+1)^2}{4}=\frac{(N+1)\big[2(2N+1)-3(N+1)\big]}{12}\\&=\frac{(N+1)(N-1)}{12}=\frac{N^2-1}{12}.\end{aligned}$$

□

Ejemplo 16.123. Para el dado, $N=6$ y las fórmulas dan $E(X)=7/2=3{,}5$ y $\operatorname{Var}(X)=35/12\approx 2{,}9167$, exactamente los dos números que se habían obtenido antes sumando término a término. La ventaja de la fórmula es que ahora una ruleta de 37 casillas se despacha sin sumar nada: $E=19$ y $\operatorname{Var}=(37^2-1)/12=114$, con $\sigma=\sqrt{114}\approx 10{,}677$.

Definición 16.124 (Distribución geométrica)**.** Se repite un ensayo con probabilidad de éxito $p\in(0,1)$ hasta obtener el primer éxito, y X cuenta el *número de ensayos realizados*, contando el que resultó exitoso. Entonces X sigue una **distribución geométrica** de parámetro p, escrito $X\sim G(p)$, y con $q=1-p$

$$P(X=k)=q^{k-1}p,\qquad k=1,2,3,\dots$$

Proposición 16.125. *Si $X\sim G(p)$, entonces la masa suma* 1*, la* cola *vale $P(X>k)=q^{\,k}$ para todo entero $k\geq 0$, y*

$$E(X)=\frac{1}{p},\qquad \operatorname{Var}(X)=\frac{q}{p^2}.$$

Demostración. La suma es una progresión geométrica de razón $q<1$: $\sum_{k\geq 1}q^{k-1}p=p\cdot\frac{1}{1-q}=\frac{p}{p}=1$. La cola es aún más directa: decir $X>k$ es decir que los k primeros ensayos fallaron, y por independencia eso tiene probabilidad q^k. De ahí sale la esperanza sin derivar ninguna serie, usando que para una variable con valores enteros positivos $E(X)=\sum_{k\geq 0}P(X>k)$:

$$E(X)=\sum_{k\geq 0}q^k=\frac{1}{1-q}=\frac{1}{p}.$$

Para la varianza calculamos $E[X(X-1)]$. Derivando dos veces la serie geométrica $\sum_{k\geq 0} x^k = (1-x)^{-1}$ se obtiene $\sum_{k\geq 2} k(k-1)x^{k-2} = 2(1-x)^{-3}$, de modo que

$$E[X(X-1)] = \sum_{k\geq 2} k(k-1)q^{k-1}p = pq\sum_{k\geq 2} k(k-1)q^{k-2} = pq\cdot\frac{2}{p^3} = \frac{2q}{p^2}.$$

Entonces $E(X^2) = \frac{2q}{p^2} + \frac{1}{p}$ y

$$\operatorname{Var}(X) = \frac{2q}{p^2} + \frac{1}{p} - \frac{1}{p^2} = \frac{2q+p-1}{p^2} = \frac{2q-q}{p^2} = \frac{q}{p^2},$$

porque $p - 1 = -q$. □

Teorema 16.126 (Falta de memoria de la geométrica)**.** *Si $X \sim G(p)$, entonces para todos los enteros $m, n \geq 0$*

$$P(X > m+n \mid X > m) = P(X > n),$$

y la geométrica es la única distribución sobre $\{1, 2, 3, \ldots\}$ con esa propiedad.

Demostración. Con la fórmula de la cola,

$$P(X > m+n \mid X > m) = \frac{P(X > m+n)}{P(X > m)} = \frac{q^{m+n}}{q^m} = q^n = P(X > n).$$

Para la unicidad, sea $g(n) = P(X > n)$ con $g(0) = 1$. La propiedad dice $g(m+n) = g(m)g(n)$, luego $g(n) = g(1)^n$ por inducción, y llamando $q = g(1)$ resulta $P(X = n) = g(n-1) - g(n) = q^{n-1}(1-q)$, que es la geométrica de parámetro $p = 1 - q$. Es el análogo discreto exacto de la ecuación funcional que caracterizaba a la exponencial en la sección anterior. □

Regla: resolver un problema de "¿cuántos intentos hacen falta?"

1. **Comprueba que la pregunta es *cuándo* ocurre el éxito**, no cuántos éxitos hay: si es lo segundo, la distribución es binomial. Y usa $P(X = k) = q^{k-1}p$ para una probabilidad puntual: $k-1$ fracasos seguidos y luego un éxito.
2. **Fija el convenio.** Aquí X cuenta los ensayos *incluido* el exitoso, así que $X \geq 1$. Hay libros que cuentan los fracasos previos, con lo que la variable empieza en 0 y la media es q/p en vez de $1/p$: las dos versiones difieren en una unidad y confundirlas es el error de este tema.
3. **Para "más de k intentos" usa la cola** $P(X > k) = q^k$, y para "a lo sumo k" su complementario $P(X \leq k) = 1 - q^k$. Nunca sumes k términos.
4. **Para la mediana**, resuelve $1 - q^k \geq \frac{1}{2}$, es decir $k \geq \ln(1/2)/\ln q$, y toma el primer entero que lo cumpla.
5. **Usa la falta de memoria para descartar información inútil.** «Llevo diez fracasos, ¿y ahora?» tiene siempre la misma respuesta que «acabo de empezar».

Ejemplo 16.127 (Lanzar un dado hasta el primer seis, con todas las cuentas)**.** Aquí $p = 1/6$ y $q = 5/6$.

(a) Los parámetros. $E(X) = 6$ lanzamientos y $\mathrm{Var}(X) = \dfrac{5/6}{1/36} = 30$, luego $\sigma = \sqrt{30} \approx 5{,}4772$: la desviación típica es casi tan grande como la media, de modo que «de media seis tiradas» no significa que suelan hacer falta seis.

(b) Una probabilidad puntual. $P(X = 4) = \left(\frac{5}{6}\right)^3 \frac{1}{6} = \frac{125}{1296} \approx 0{,}0965$, el valor que ya se había usado sin justificar.

(c) Colas. $P(X > 6) = \left(\frac{5}{6}\right)^6 = \frac{15\,625}{46\,656} \approx 0{,}3349$, de modo que en un tercio largo de las partidas hacen falta más de seis tiradas. Y $P(X \le 6) = 1 - 0{,}3349 = 0{,}6651$.

(d) La mediana. Buscamos el menor k con $1-(5/6)^k \ge 0{,}5$, es decir $k \ge \ln(0{,}5)/\ln(5/6) = 3{,}8018$: la mediana es 4. Es menor que la media 6, señal de la asimetría hacia la derecha; la mayoría de las partidas son cortas y unas pocas, larguísimas.

(e) Falta de memoria, con números. $P(X > 10) = (5/6)^{10} \approx 0{,}1615$ y $P(X > 4) = (5/6)^4 \approx 0{,}4823$, luego

$$P(X > 10 \mid X > 4) = \frac{0{,}1615}{0{,}4823} = 0{,}3349 = P(X > 6).$$

Cuatro tiradas fallidas no acercan el seis: quien ya ha fallado cuatro veces está exactamente donde estaba al empezar.

Definición 16.128 (Distribución hipergeométrica)**.** Una población de N individuos contiene K del tipo que interesa. Se extrae una muestra de n individuos *sin reposición* y X cuenta cuántos de la muestra son del tipo bueno. Entonces X sigue una **distribución hipergeométrica** y

$$P(X = k) = \frac{\binom{K}{k}\binom{N-K}{n-k}}{\binom{N}{n}},$$

para los enteros k con $\max(0,\, n - N + K) \le k \le \min(n, K)$.

Proposición 16.129. *Con $p = K/N$ y $q = 1 - p$,*

$$E(X) = n\,p, \qquad \mathrm{Var}(X) = n\,p\,q\,\frac{N-n}{N-1}.$$

El factor $\frac{N-n}{N-1}$ se llama corrección por población finita.

Demostración. Escribimos $X = I_1 + \cdots + I_n$, donde I_j vale 1 si el j-ésimo individuo extraído es del tipo bueno. Aunque las extracciones no sean independientes, cada individuo de la población tiene la misma probabilidad de ocupar la posición j, luego $P(I_j = 1) = K/N = p$ y, por linealidad de la esperanza —que no exige independencia—, $E(X) = np$. Para la varianza necesitamos las covarianzas. Para $i \ne j$,

$$E(I_i I_j) = P(\text{los dos buenos}) = \frac{K}{N} \cdot \frac{K-1}{N-1},$$

de modo que

$$\operatorname{Cov}(I_i, I_j) = \frac{K(K-1)}{N(N-1)} - p^2 = p\left(\frac{K-1}{N-1} - p\right) = -\frac{p\,q}{N-1}.$$

Como cada I_j tiene varianza pq y hay $n(n-1)$ pares ordenados distintos,

$$\operatorname{Var}(X) = n\,pq + n(n-1)\left(-\frac{pq}{N-1}\right) = n\,pq\left(1 - \frac{n-1}{N-1}\right) = n\,pq\,\frac{N-n}{N-1}. \qquad \square$$

Idea clave

La corrección por población finita explica de un golpe dos cosas. Primero, por qué la binomial aproxima bien a la hipergeométrica cuando la población es grande: si N es enorme frente a n, el factor $\frac{N-n}{N-1}$ vale casi 1 y las dos varianzas coinciden. Segundo, por qué muestrear sin reposición es *más* preciso que hacerlo con reposición: el factor es siempre menor o igual que 1, así que la varianza disminuye. En el caso extremo $n = N$ vale exactamente 0, y en efecto si se examina la población entera no queda ninguna incertidumbre.

Regla: elegir entre binomial e hipergeométrica

1. **Pregunta si el individuo extraído vuelve a la población.** Si vuelve, o si se extrae de un proceso que no se agota, es binomial; si no vuelve, es hipergeométrica.
2. **Identifica los cuatro números de la hipergeométrica:** tamaño de la población N, número de elementos buenos K, tamaño de la muestra n y número de aciertos k. Si el enunciado no permite señalar los cuatro, la distribución no es esta.
3. **Escribe la fracción de combinatorios** eligiendo k buenos entre los K y $n-k$ malos entre los $N-K$, dividido por todas las muestras posibles.
4. **Comprueba el rango de k:** no puede haber más aciertos que buenos hay ni que elementos se extraen.
5. **Usa $E(X) = nK/N$ sin miedo** —es idéntica a la de la binomial, porque la falta de independencia no afecta a la esperanza— y **corrige la varianza** multiplicando la binomial npq por $\frac{N-n}{N-1}$.
6. **Aproxima por la binomial solo si** $n \le 0{,}1N$, y en caso contrario resiste la tentación: el error va siempre en el mismo sentido, sobrestimando la dispersión.

Ejemplo 16.130 (Cartas, lotería y un lote de piezas)**. (a) Dos ases en cinco cartas.** Con $N = 52$, $K = 4$, $n = 5$,

$$P(X = 2) = \frac{\binom{4}{2}\binom{48}{3}}{\binom{52}{5}} = \frac{6 \cdot 17\,296}{2\,598\,960} = 0{,}039930.$$

La media es $E(X) = 5 \cdot \frac{4}{52} = \frac{5}{13} = 0{,}384615$ y la varianza

$$\operatorname{Var}(X) = 5 \cdot \frac{1}{13} \cdot \frac{12}{13} \cdot \frac{47}{51} = 0{,}327184,$$

frente a la varianza binomial $5 \cdot \frac{1}{13} \cdot \frac{12}{13} = 0{,}355030$. La corrección rebaja la dispersión casi un 8 %. La probabilidad calculada mal como binomial habría sido 0,046540, un 17 % de exceso.

(b) La lotería 6 de 49. Es hipergeométrica con $N = 49$, $K = 6$, $n = 6$, y hay $\binom{49}{6} = 13\,983\,816$ boletos posibles, de donde

$$P(\text{6 aciertos}) = \frac{1}{13\,983\,816} = 7{,}1511 \cdot 10^{-8}, \qquad P(\text{3 aciertos}) = \frac{\binom{6}{3}\binom{43}{3}}{\binom{49}{6}} = 0{,}017650.$$

Uno de cada cincuenta y siete boletos acierta tres números y uno de cada catorce millones acierta los seis; la esperanza es $6 \cdot 6/49 = 0{,}7347$ aciertos.

(c) Control de calidad. Un lote de 20 piezas contiene 3 defectuosas y se inspeccionan 5 sin reponer. La probabilidad de no detectar ninguna defectuosa es

$$P(X = 0) = \frac{\binom{17}{5}}{\binom{20}{5}} = \frac{6188}{15\,504} = 0{,}399123,$$

de modo que el control detecta algún defecto el 60,09 % de las veces. Aquí $n = 5$ es la cuarta parte de $N = 20$, muy lejos del 10 % que autorizaría la aproximación binomial, y usarla sería un error de bulto.

Cuidado

Los tres errores que se repiten. El primero es confundir la geométrica con la binomial: «lanzo diez veces, ¿cuántos seises?» es binomial, «lanzo hasta el primer seis» es geométrica, y la palabra que las separa es *hasta*. El segundo es usar la geométrica con el convenio equivocado: si el libro de tablas cuenta fracasos y no ensayos, todas las respuestas salen desplazadas en una unidad. Y el tercero es aplicar la binomial a extracciones sin reposición cuando la muestra no es despreciable frente a la población; el resultado no es una aproximación aceptable sino un número equivocado, y siempre por el mismo lado.

Pausa para pensar

Las cuatro distribuciones de esta sección y las dos anteriores forman una sola familia, y merece la pena ver el mapa completo. Fijado el número de ensayos y contando éxitos, sale la binomial; fijado el número de éxitos y contando ensayos, sale la geométrica. Quitar la reposición convierte la binomial en hipergeométrica, y devolver la población al infinito deshace el cambio. Hacer $n \to \infty$ con np constante convierte la binomial en Poisson, y mirar el tiempo entre sucesos convierte la Poisson en exponencial, que es la versión continua de la geométrica: por eso ambas, y solo ellas, carecen de memoria. Cinco distribuciones, un solo esquema, tres preguntas: ¿cuántos, en cuántos intentos, cada cuánto?

16.23. La función de distribución, los cuantiles y el método de la transformada inversa

Hasta aquí el capítulo ha descrito cada variable aleatoria por su masa o por su densidad, que son objetos distintos según el caso. Existe, sin embargo, una descripción única que sirve para

las dos y que es la que de verdad usan los ordenadores y las tablas: la función de distribución. Y de ella sale la noción de *cuantil*, que el capítulo lleva usando desde la primera tabla de la normal —los $z_{\alpha/2}$ de los intervalos de confianza y los valores críticos de los contrastes son cuantiles— sin haberla definido en ningún sitio.

Definición 16.131 (Función de distribución)**.** La **función de distribución** (acumulada) de una variable aleatoria X es

$$F(x) = P(X \leq x), \qquad x \in \mathbb{R}.$$

Si X es discreta, $F(x) = \sum_{x_i \leq x} p_i$ y es una función escalonada que salta p_i en cada valor posible. Si X es continua con densidad f, entonces $F(x) = \int_{-\infty}^{x} f(t)\,\mathrm{d}t$ y F es continua.

Teorema 16.132 (Propiedades de la función de distribución)**.** *Para cualquier variable aleatoria X y cualesquiera $a < b$:*

(i) *F es no decreciente y toma valores en $[0, 1]$;*

(ii) *$\lim_{x\to-\infty} F(x) = 0$ y $\lim_{x\to+\infty} F(x) = 1$;*

(iii) *$P(a < X \leq b) = F(b) - F(a)$;*

(iv) *$P(X = a) = F(a) - F(a^-)$, donde $F(a^-)$ es el límite por la izquierda; en particular F es continua en a si y solo si $P(X = a) = 0$;*

(v) *si X es continua con densidad f continua, entonces $F'(x) = f(x)$.*

Demostración. (i) Si $a < b$, el suceso $\{X \leq a\}$ está contenido en $\{X \leq b\}$, y la probabilidad de un subconjunto no supera a la del conjunto; que esté entre 0 y 1 es inmediato porque es una probabilidad.

(ii) Los sucesos $\{X \leq -n\}$ decrecen hacia el suceso imposible y los $\{X \leq n\}$ crecen hacia el seguro, y la probabilidad respeta esos límites monótonos.

(iii) Los sucesos $\{X \leq a\}$ y $\{a < X \leq b\}$ son incompatibles y su unión es $\{X \leq b\}$, luego $F(a) + P(a < X \leq b) = F(b)$, que es la fórmula. Éste es el motivo por el que F merece la pena: una sola función responde a *todas* las preguntas sobre intervalos.

(iv) Tomando $a_n \uparrow a$, los sucesos $\{a_n < X \leq a\}$ decrecen hacia $\{X = a\}$, y por (iii) sus probabilidades son $F(a) - F(a_n) \to F(a) - F(a^-)$.

(v) Es el teorema fundamental del cálculo aplicado a $F(x) = \int_{-\infty}^{x} f$, tal como se vio en las reglas de derivación e integración. □

Observación 16.133. La propiedad (iv) es la que explica de una vez la advertencia de la sección de variables continuas. En una variable continua F no da saltos, luego $P(X = a) = 0$ para todo a y da igual escribir $<$ que $\leq$; en una discreta, F salta exactamente p_i en cada valor, y ese salto *es* la probabilidad puntual. Las dos frases que parecían reglas sueltas son la misma propiedad de F leída en los dos casos.

Ejemplo 16.134 (Tres funciones de distribución)**. (a) El dado.** F es una escalera con seis peldaños de altura 1/6: vale 0 para $x < 1$, 1/6 en $[1, 2)$, 2/6 en $[2, 3)$ y así hasta valer 1 desde

$x = 6$. Y en efecto $P(2 < X \le 5) = F(5) - F(2) = \frac{5}{6} - \frac{2}{6} = \frac{1}{2}$, que son los tres resultados 3, 4 y 5.

(b) La densidad $f(x) = 3x^2$ en $[0,1]$. Integrando, $F(x) = \int_0^x 3t^2 \mathrm{d}t = x^3$ para $x \in [0,1]$, con $F(x) = 0$ antes y $F(x) = 1$ después. Comprobación de la propiedad (v): $F'(x) = 3x^2 = f(x)$ ✓. Y $P(X \le \frac{1}{2}) = (\frac{1}{2})^3 = 0{,}125$, el mismo número que antes se obtuvo integrando.

(c) La exponencial. $F(x) = 1 - \mathrm{e}^{-\lambda x}$ para $x \ge 0$. La sección anterior la había dado ya, y ahora se ve que no era un dato adicional de la distribución sino simplemente la integral de su densidad.

Definición 16.135 (Cuantil, mediana y percentiles)**.** Dado $p \in (0,1)$, el **cuantil** de orden p de la variable X es

$$q_p = \min\{x \in \mathbb{R} : \ F(x) \ge p\},$$

es decir el primer valor en el que la probabilidad acumulada alcanza p. El cuantil $q_{1/2}$ se llama **mediana**; $q_{1/4}$ y $q_{3/4}$ son el primer y el tercer **cuartil**, y su diferencia es el *recorrido intercuartílico*; los cuantiles $q_{k/100}$ se llaman **percentiles**. Para la normal tipificada se emplea además la notación z_α para el valor que deja *a la derecha* una probabilidad α, es decir $z_\alpha = q_{1-\alpha}$.

Regla: calcular un cuantil

1. **Dibuja siempre la acumulada o la densidad y sombrea** el área p empezando por la izquierda. La mitad de los errores de este tema son de orientación, no de cálculo.
2. **Si X es continua y F es estrictamente creciente**, el cuantil es la única solución de la ecuación $F(x) = p$; despéjala con la función inversa correspondiente —logaritmo, raíz, tabla—.
3. **Si X es discreta, no resuelvas ninguna ecuación:** recorre la tabla acumulada y quédate con el primer valor cuya F llega o supera a p. La ecuación $F(x) = p$ casi nunca tiene solución, y por eso la definición usa un mínimo y no una igualdad.
4. **Traduce el enunciado con cuidado.** «El 10 % superior» empieza en el percentil 90, no en el 10; «el valor que solo supera uno de cada cuatro» es el tercer cuartil.
5. **Para la normal, tipifica al revés:** busca z con $\Phi(z) = p$ en la tabla y deshaz con $x = \mu + z\sigma$.
6. **Comprueba con la propia definición:** evalúa F en el cuantil obtenido y verifica que da p (caso continuo) o que lo alcanza por primera vez (caso discreto).
7. **Contrasta con la forma de la distribución.** En una distribución asimétrica hacia la derecha la mediana queda por debajo de la media; si te sale por encima, revisa la cuenta.

Ejemplo 16.136 (Cuantiles de cuatro distribuciones conocidas)**. (a) La exponencial del autobús.** Con $\lambda = 0{,}1$ por minuto, resolver $1 - \mathrm{e}^{-\lambda x} = p$ da la fórmula general $q_p =$

$-\ln(1-p)/\lambda$. De ahí

$$q_{0,25} = 2{,}8768, \qquad q_{0,5} = \frac{\ln 2}{\lambda} = 6{,}9315, \qquad q_{0,75} = 13{,}8629, \qquad q_{0,9} = 23{,}0259$$

minutos. La mediana 6,93 es la que se había usado sin justificar en la sección de la exponencial: ahora se ve que es el cuantil de orden $1/2$ y de dónde sale su fórmula. Nótese que la mediana queda muy por debajo de la media 10 y que el 10 % de las esperas superan los 23 minutos: eso es una cola pesada.

(b) La densidad $3x^2$. Aquí $F(x) = x^3$, luego $q_p = \sqrt[3]{p}$ y

$$q_{0,25} = 0{,}6300, \qquad q_{0,5} = 0{,}7937, \qquad q_{0,75} = 0{,}9086.$$

El recorrido intercuartílico es 0,2786, y la mediana 0,7937 vuelve a caer por encima del centro del intervalo, coherente con una densidad que crece hacia la derecha. Comprobación: $F(0{,}7937) = 0{,}7937^3 = 0{,}5$ ✓.

(c) La normal. Los cuantiles de $N(0,1)$ que el capítulo usa una y otra vez son

$$q_{0,90} = 1{,}2816, \quad q_{0,95} = 1{,}6449, \quad q_{0,975} = 1{,}9600, \quad q_{0,99} = 2{,}3263, \quad q_{0,995} = 2{,}5758.$$

Los tres subrayados son los de los intervalos de confianza: un intervalo al 95 % usa $z_{0,025} = q_{0,975} = 1{,}96$ porque deja un 2,5 % en cada cola. Para las estaturas $N(170, 8)$, el percentil 90 es $170 + 8 \cdot 1{,}2816 = 180{,}25$ cm —el número que ya había salido al preguntar por el 10 % más alto— y los cuartiles son 164,60 y 175,40 cm.

(d) Una binomial, para ver por qué la definición lleva un mínimo. Sea $X \sim B(10; 0{,}3)$, con acumulada

$$F(0) = 0{,}0282, \quad F(1) = 0{,}1493, \quad F(2) = 0{,}3828,$$
$$F(3) = 0{,}6496, \quad F(4) = 0{,}8497, \quad F(5) = 0{,}9527.$$

La mediana es 3, porque es el primer valor cuya acumulada llega a 0,5, y sin embargo $F(3) = 0{,}6496$ no vale 0,5 ni por asomo: no existe ningún x con $F(x) = 0{,}5$. El primer cuartil es 2 y el percentil 90 es 5. Quien intente «resolver la ecuación» aquí no encontrará nada que resolver.

Teorema 16.137 (De la transformada inversa)**.** *Sea F la función de distribución de una variable continua estrictamente creciente en el conjunto donde toma valores entre 0 y 1, y sea U una variable uniforme en $(0,1)$. Entonces la variable $X = F^{-1}(U)$ tiene por función de distribución exactamente F. Recíprocamente, si X tiene función de distribución F continua, entonces $F(X)$ es uniforme en $(0,1)$.*

Demostración. Para la primera parte, y como F es creciente, la desigualdad $F^{-1}(U) \le x$ equivale a $U \le F(x)$. Por tanto

$$P(X \le x) = P\big(U \le F(x)\big) = F(x),$$

donde la última igualdad es la definición de uniforme en $(0,1)$: la probabilidad de que U caiga por debajo de un número de $[0,1]$ es ese número. La segunda parte es la misma cadena leída al revés: $P\big(F(X) \le u\big) = P\big(X \le F^{-1}(u)\big) = F\big(F^{-1}(u)\big) = u$. □

Regla: simular una variable aleatoria cualquiera

1. **Consigue un número u al azar entre 0 y 1**: es lo único que sabe hacer un ordenador, y cualquier hoja de cálculo lo ofrece.
2. **Escribe la función de distribución F** de la variable que quieras simular.
3. **Despeja x en $F(x) = u$**. Esa x es tu valor simulado.
4. **Si la variable es discreta**, sustituye el despeje por una búsqueda: recorre la tabla acumulada y devuelve el primer valor cuya acumulada supere a u.
5. **Repite tantas veces como necesites, y valida** comparando la media, la mediana y alguna cola de la muestra con sus valores teóricos; si no se parecen, el error está en el paso 3.
6. **Recuerda el orden de magnitud del error:** por el teorema central del límite, con m repeticiones la media simulada tiene un error típico de $\sigma/\sqrt{m}$, así que multiplicar la precisión por diez cuesta cien veces más cálculo.

Ejemplo 16.138 (Fabricar esperas exponenciales a partir de números al azar)**.** Queremos simular la espera del autobús, exponencial con $\lambda = 0{,}1$. De $1 - \mathrm{e}^{-\lambda x} = u$ se despeja

$$x = -\frac{\ln(1-u)}{\lambda}.$$

Con $u = 0{,}37$ sale $x = -\ln(0{,}63)/0{,}1 = 4{,}6204$ minutos. Repitiendo el procedimiento 200 000 veces con números al azar se obtiene una muestra cuya media es 9,989 frente al valor teórico 10, cuya mediana es 6,956 frente a $\ln 2/\lambda = 6{,}931$, y en la que la proporción de esperas superiores a 15 minutos es 0,2233 frente al valor exacto $\mathrm{e}^{-1{,}5} = 0{,}2231$. Las tres coincidencias son la comprobación del paso 6, y confirman que el método no es un truco de programador sino el teorema 16.137 en acción.

Cuidado

Tres avisos. El primero: en el caso discreto conviven varios convenios —algunos programas devuelven 3 como mediana del dado y otros 3,5—; la definición con mínimo dada aquí es la estándar, y es la que hace funcionar el teorema de la transformada inversa. El segundo: q_p es un valor de la variable y se mide en sus unidades —minutos, centímetros—, nunca en tanto por ciento. Y el tercero: el cuantil de una suma no es la suma de los cuantiles, ni siquiera para variables independientes; solo la esperanza tiene esa propiedad.

Idea clave

La función de distribución es el objeto que unifica todo el capítulo. Vale igual para variables discretas y continuas, determina la variable por completo, convierte cualquier pregunta sobre intervalos en una resta, produce los cuantiles con los que se construyen todos los intervalos de confianza y todos los contrastes, y —por el teorema de la transformada inversa— permite fabricar cualquier distribución del mundo a partir de

un simple generador de números al azar. Esa última consecuencia es la base de los métodos de Montecarlo, que hoy resuelven por simulación integrales, modelos financieros y problemas de física estadística cuya solución exacta nadie sabe escribir.

16.24. Ejercicios resueltos

Ejercicio 16.1. Sea X el número de caras al lanzar cuatro monedas. Halla su distribución, $E(X)$ y σ.

Solución. $X \sim B(4;0{,}5)$, luego $P(X=k)=\binom{4}{k}/16$, es decir 1/16, 4/16, 6/16, 4/16, 1/16 para $k=0,1,2,3,4$; suman $16/16=1$. ✓

$E(X)=np=4\cdot 0{,}5=2$ y $\mathrm{Var}(X)=npq=4\cdot 0{,}25=1$, luego $\sigma=1$. ■

Ejercicio 16.2. Una variable continua tiene densidad $f(x)=2x$ en $[0,1]$. Comprueba que lo es y halla μ, σ^2 y $P(X>0{,}5)$.

Solución. $\int_0^1 2x\,\mathrm{d}x=[x^2]_0^1=1$. ✓

$\mu=\int_0^1 2x^2\mathrm{d}x=\frac{2}{3}$; $E(X^2)=\int_0^1 2x^3\mathrm{d}x=\frac{1}{2}$; $\sigma^2=\frac{1}{2}-\frac{4}{9}=\frac{9-8}{18}=\frac{1}{18}\approx 0{,}0556$.

$P(X>0{,}5)=\int_{1/2}^1 2x\,\mathrm{d}x=1-\frac{1}{4}=\frac{3}{4}$. ■

Ejercicio 16.3. La duración de unas bombillas es $N(1000,50)$ horas. Halla $P(X>1100)$, $P(950<X<1050)$ y la duración que garantiza el fabricante si acepta reponer el 5 % que menos dura.

Solución. $P(X>1100)$: $z=(1100-1000)/50=2$, luego $1-\Phi(2)=0{,}0228$, un 2,3 %.

$P(950<X<1050)$: es exactamente $\mu\pm\sigma$, luego $\approx 68{,}3$ %.

Garantía: buscamos el percentil 5, con $z=-1{,}645$; $x=1000-1{,}645\cdot 50\approx 917{,}8$ h. El fabricante puede garantizar 918 horas. ■

Ejercicio 16.4. Con $\sigma=5$ conocida, una muestra de $n=100$ da $\bar{x}=48{,}2$. Construye el intervalo de confianza al 95 % e interpreta su amplitud.

Solución. Error típico $5/\sqrt{100}=0{,}5$; semiamplitud $1{,}96\cdot 0{,}5=0{,}98$. El intervalo es $(47{,}22;\ 49{,}18)$, de amplitud 1,96. Para reducirlo a la mitad, a $\pm 0{,}49$, haría falta $n=400$. ■

Ejercicio 16.5. Se sospecha que una moneda está trucada a favor de la cara. En 400 lanzamientos salen 220 caras. Contrasta al 5 %.

Solución. $H_0\colon p=0{,}5$ frente a $H_1\colon p>0{,}5$. Bajo H_0, $X\sim B(400;0{,}5)$ con $np=200$ y $\sqrt{npq}=\sqrt{100}=10$. Con corrección por continuidad,

$$z=\frac{219{,}5-200}{10}=1{,}95.$$

El crítico unilateral al 5 % es 1,645, y $1{,}95>1{,}645$: se rechaza H_0 (p-valor $\approx 0{,}026$). Hay evidencia, al 5 %, de sesgo hacia la cara —aunque nótese que al 1 % ($z_c=2{,}326$) *no* se rechazaría. El umbral 0,05 no tiene nada de sagrado. ■

Ejercicio 16.6. Demuestra la fórmula abreviada $\mathrm{Var}(X)=E(X^2)-\mu^2$.

Solución. Desarrollando el cuadrado y usando la linealidad de la esperanza:

$$E\big[(X-\mu)^2\big] = E\big[X^2 - 2\mu X + \mu^2\big] = E(X^2) - 2\mu E(X) + \mu^2 = E(X^2) - 2\mu^2 + \mu^2 = E(X^2) - \mu^2,$$

donde se ha usado que μ es una constante y que $E(X) = \mu$. Es literalmente la misma cuenta que la fórmula abreviada de la varianza de la estadística descriptiva, con $\sum$ sustituido por E. ■

16.25. Ejercicios propuestos

16.1. Sea X la suma de dos dados. Halla $E(X)$ y $\mathrm{Var}(X)$.

16.2. Si $X \sim B(8; 0{,}25)$, halla $P(X = 2)$, $E(X)$ y σ.

16.3. Determina k para que $f(x) = kx$ sea densidad en $[0, 2]$ y calcula μ.

16.4. Si $X \sim N(60, 10)$, halla $P(X > 75)$ y $P(45 < X < 70)$.

16.5. Para esa misma normal, halla el primer cuartil.

16.6. Si $\mathrm{Var}(X) = 4$, ¿cuánto valen $\mathrm{Var}(3X - 2)$ y la desviación típica de $3X - 2$?

16.7. Con $\sigma = 5$, ¿qué tamaño de muestra hace falta para estimar μ con error máximo 1 al 99 %?

16.8. Se contrasta $H_0\colon \mu = 2{,}5$ frente a $H_1\colon \mu \neq 2{,}5$ con $\sigma = 0{,}3$, $n = 36$ y $\bar{x} = 2{,}62$. Decide al 5 % y da el p-valor.

16.9. Si $X \sim B(400; 0{,}4)$, aproxima $P(X \leq 150)$.

16.10. Un juego consiste en lanzar un dado y cobrar en euros el resultado, pagando 4 por jugar. ¿Conviene jugar?

16.11. Explica en una línea por qué un intervalo al 99 % es más ancho que uno al 95 %.

16.12. Una encuesta con $n = 1000$ da un margen de ± 3 %. ¿Cuánta gente haría falta para ± 1 %?

Respuestas

1. $E(X) = 7$ por simetría (o sumando); $\mathrm{Var}(X) = 2 \cdot \frac{35}{12} = \frac{35}{6} \approx 5{,}83$, pues los dos dados son independientes. **2.** $P(X = 2) = \binom{8}{2}(0{,}25)^2(0{,}75)^6 = 28 \cdot 0{,}0625 \cdot 0{,}17798 \approx 0{,}3115$; $E(X) = 2$; Var $= 1{,}5$, $\sigma \approx 1{,}225$. **3.** $\int_0^2 kx\,\mathrm{d}x = 2k = 1 \Rightarrow k = \frac{1}{2}$; $\mu = \int_0^2 \frac{x^2}{2}\mathrm{d}x = \frac{8}{6} = \frac{4}{3}$. **4.** $z = 1{,}5 \Rightarrow P(X > 75) = 1 - \Phi(1{,}5) = 0{,}0668$; $P(45 < X < 70) = \Phi(1) - \Phi(-1{,}5) = 0{,}8413 - 0{,}0668 = 0{,}7745$. **5.** $z_{0{,}25} = -0{,}6745$, luego $Q_1 = 60 - 6{,}745 \approx 53{,}26$. **6.** $\mathrm{Var}(3X - 2) = 9 \cdot 4 = 36$ y $\sigma = 6$; el -2 no influye. **7.** $n \geq (2{,}576 \cdot 5/1)^2 = 165{,}89 \Rightarrow n = 166$. **8.** Error típico $0{,}3/6 = 0{,}05$; $z = (2{,}62 - 2{,}5)/0{,}05 = 2{,}4 > 1{,}96$, se rechaza H_0; $p = 2(1 - \Phi(2{,}4)) \approx 0{,}0164$. **9.** $np = 160$, $\sqrt{npq} = \sqrt{96} \approx 9{,}798$; con corrección $z = (150{,}5 - 160)/9{,}798 = -0{,}97$, luego $P \approx \Phi(-0{,}97) = 0{,}166$ (el valor exacto es 0,1662). **10.** No: $E(\text{ganancia}) = 3{,}5 - 4 = -0{,}5$ por jugada. A largo plazo se pierden 50 céntimos cada vez. **11.** Porque exigir que el procedimiento acierte más veces obliga a abarcar más valores; con los mismos datos, confianza y precisión son intercambiables. **12.** Nueve veces más: el error va como $1/\sqrt{n}$, así que dividirlo entre tres exige multiplicar n por 9, es decir, unas 9000 personas.

Capítulo 17

Estadística descriptiva y análisis exploratorio

Lo que necesitas traer

Del capítulo 16: variables aleatorias, media y varianza, aunque aquí se trabaja con datos y no con modelos. Del capítulo 6: logaritmos, para las transformaciones, y el concepto de cuantil.

Antes de contrastar hipótesis o construir modelos hay que mirar los datos, y mirarlos bien. La estadística descriptiva no hace inferencia: no pretende decir nada sobre una población, solo resumir con honestidad lo que hay delante. Parece la parte fácil de la asignatura y es donde se cometen la mayoría de los errores graves, porque un resumen mal elegido puede ocultar exactamente lo que importaba. Este capítulo trata las medidas de resumen desde el punto de vista de su *robustez*, la estructura de las relaciones entre variables, y los casos clásicos en que resumir engaña.

17.1. Medidas de centro y de dispersión: la cuestión de la robustez

Definición 17.1. Para una muestra $x_1, \dots, x_n$ se definen la *media* $\bar{x} = \frac{1}{n}\sum x_i$, la *mediana* (el valor central una vez ordenados) y la *cuasivarianza* $s^2 = \frac{1}{n-1}\sum (x_i - \bar{x})^2$. La *desviación absoluta mediana* es $\mathrm{MAD} = \mathrm{med}\,(|x_i - \mathrm{med}(x)|)$. El *punto de ruptura* de un estadístico es la proporción de datos que hay que corromper para que su valor se vaya a infinito.

Idea clave

Qué dice esto en cristiano. La media y la mediana intentan responder a la misma pregunta y se comportan de manera opuesta ante un dato extremo. La media tiene punto de ruptura 0: basta con que *un solo* dato se vaya a infinito para arrastrarla consigo. La mediana tiene punto de ruptura $\frac{1}{2}$: hay que corromper la mitad de la muestra para

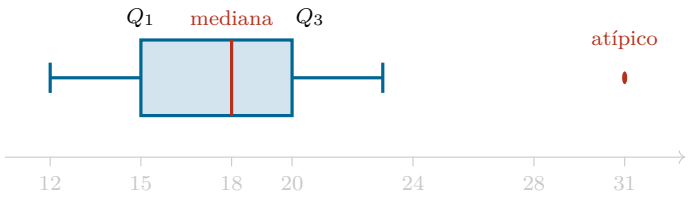


Figura 17.1: Diagrama de caja de trece datos. La caja va del primer al tercer cuartil, la raya interior es la mediana, los bigotes llegan al dato más extremo que aún queda dentro de $Q_3 + 1{,}5\,(Q_3 - Q_1) = 27{,}5$, y el 31 queda fuera y se dibuja aparte.

moverla. Lo mismo ocurre con la desviación típica frente a la MAD. Elegir entre ellas no es una cuestión de gusto sino de decidir si los valores extremos son información valiosa o contaminación, y esa decisión es sustantiva, no estadística.

Regla: describir una muestra sin engañar

El orden importa, porque cada paso condiciona la interpretación del siguiente:

1. **Dibuja antes de calcular.** Un histograma y un diagrama de caja cuestan segundos y revelan asimetrías, bimodalidades y atípicos que ningún número resume.
2. **Da siempre una medida de centro junto a una de dispersión.** Una media sin desviación no significa nada; con $\bar{x} = 24$ podría haber datos entre 23 y 25 o entre 0 y 100.
3. **Comprueba la simetría** comparando media y mediana: si difieren mucho, la distribución está sesgada y la media deja de representar al individuo típico.
4. **Localiza los atípicos** con el criterio $Q_1 - 1{,}5\,\mathrm{IQR}$ y $Q_3 + 1{,}5\,\mathrm{IQR}$, y *investígalos* antes de decidir nada; un atípico puede ser un error de tecleo o el hallazgo del estudio.
5. **Repite los cálculos sin ellos** y compara: si las conclusiones cambian, hay que decirlo en el informe.
6. **Usa medidas robustas si la contaminación es plausible**: mediana en vez de media, MAD o rango intercuartílico en vez de desviación típica.
7. **Informa de n siempre.** Una media de tres datos y otra de tres mil son objetos completamente distintos aunque el número coincida.

Ejemplo 17.2 (Una muestra con un dato extremo)**. (a) Los datos y los resúmenes.** Sea la muestra $12, 15, 15, 17, 18, 18, 19, 21, 22, 24, 25, 28, 31, 35, 64$, con $n = 15$. La media es 24,266667 y la mediana 21. La cuasivarianza vale 160,780952 y la desviación típica muestral 12,679943. El coeficiente de variación es $\frac{s}{\bar{x}} = 0{,}522525$, es decir, la dispersión es más de la mitad del centro.

(b) El diagrama de caja. Los cuartiles son $Q_1 = 17{,}5$, $Q_2 = 21$ y $Q_3 = 26{,}5$, con

IQR $= 9$. Los bigotes llegan hasta $17{,}5 - 13{,}5 = 4$ y $26{,}5 + 13{,}5 = 40$, de modo que el 64 queda fuera: es el único atípico según el criterio estándar.

(c) Robustez, medida. Quitando el 64, la media baja a 21,428571, es decir 2,838095 menos, mientras que la mediana pasa de 21 a 20: apenas se mueve. La MAD vale 4, y multiplicada por el factor de consistencia 1,4826 da 5,930400 como estimación robusta de la desviación típica, menos de la mitad que los 12,68 de la clásica. El único dato extremo ha inflado la desviación típica más del doble.

(d) La forma. El coeficiente de asimetría vale 2,418201, claramente positivo: cola larga hacia la derecha. La curtosis en exceso es 7,079733, muy por encima de cero, lo que confirma que hay más masa en las colas de la que tendría una normal. Con esos dos números ya se sabe, sin mirar el dibujo, que resumir con la media es una mala idea.

Cuidado

[Los tres errores que más se ven] Primero: usar la media con distribuciones asimétricas. El salario *medio* de una empresa donde el director gana veinte veces que los demás no describe a nadie; para eso está la mediana, y por eso los organismos oficiales publican la mediana de renta. Segundo: dividir entre n en vez de entre $n-1$. La cuasivarianza usa $n-1$ porque $\bar{x}$ ya ha consumido un grado de libertad, y dividir entre n produce un estimador *sesgado* que subestima sistemáticamente. Y tercero: eliminar atípicos sin justificarlo. Es una decisión que cambia los resultados y debe explicarse; el agujero de ozono sobre la Antártida pasó años sin detectarse porque el programa de análisis descartaba automáticamente las lecturas "imposiblemente" bajas.

17.2. Los cimientos: tipos de variable estadística, tablas de frecuencias, tipificación y las dos rectas de regresión

El capítulo arranca calculando medias y medianas de una muestra de números. Antes de eso hay que decidir si los datos *son* números, porque la mitad de los errores de un informe estadístico consisten en aplicar a una variable una operación que su naturaleza no admite: promediar códigos postales, calcular la mediana de colores, tratar como continua una escala de cinco puntos. Y una vez decidido, hay un paso que precede a cualquier resumen y que casi nunca se escribe: organizar los datos en una *tabla de frecuencias*. Esta sección cubre esos cimientos, añade la *tipificación* —que es la operación que permite comparar cosas medidas en unidades distintas— y cierra con un hecho de la regresión que el capítulo usa a medias: que las rectas de regresión no son una, sino dos.

Qué clase de variable tenemos delante

Definición 17.3 (Variable estadística y sus tipos)**.** Una *variable estadística* es la característica que se observa en cada individuo de una población o muestra. Según los valores que toma y las operaciones que admite se clasifica así:

- **Cualitativa nominal**: sus valores son categorías sin orden natural. Color de ojos, provincia, grupo sanguíneo. Sólo admite contar cuántos hay de cada clase; la única medida de centro es la moda.
- **Cualitativa ordinal**: categorías con un orden, pero sin distancias medibles entre ellas. Nivel de estudios, grado de dolor de 1 a 5, "muy en desacuerdo" a "muy de acuerdo". Admite orden, y por tanto mediana y cuantiles, pero no media: la distancia entre "bueno" y "excelente" no es comparable con la que hay entre "malo" y "regular".
- **Cuantitativa discreta**: números que proceden de contar, con valores aislados. Número de hijos, de averías, de coches. Admite todas las operaciones aritméticas, aunque su media puede ser un valor imposible.
- **Cuantitativa continua**: números que proceden de medir y que, al menos en teoría, pueden tomar cualquier valor de un intervalo. Estatura, tiempo, temperatura. En la práctica siempre se registran redondeados, y por eso se agrupan en clases.

Idea clave

Qué dice esto en cristiano. La clasificación no es taxonomía por gusto: decide qué se puede calcular. Con una variable nominal, la media no existe —"la media de las provincias" no significa nada— y el único resumen es la tabla de frecuencias con su moda. Con una ordinal se puede ordenar, y por tanto tienen sentido la mediana y los cuartiles, pero la media supone que los saltos entre categorías son iguales, cosa que nadie ha comprobado. Con una cuantitativa vale todo. Y hay una trampa clásica: una variable puede estar codificada con números y ser nominal ("1 = soltero, 2 = casado"), en cuyo caso el ordenador calculará encantado una media de 1,63 que no significa absolutamente nada.

Ejemplo 17.4 (Ocho variables clasificadas, y dos casos discutibles)**. (a) Las fáciles.** "Grupo sanguíneo" es nominal; "número de hermanos" es cuantitativa discreta; "tiempo en recorrer cien metros" es cuantitativa continua; "nivel de estudios terminados" es ordinal.

(b) Las que engañan. "Número de la camiseta de un jugador" parece cuantitativa discreta y es *nominal*: el 10 no vale el doble que el 5, y su media no significa nada. "Código postal" es igual. Son etiquetas escritas con dígitos.

(c) El caso discutible por excelencia. Una escala de satisfacción de 1 a 5 es estrictamente ordinal, y sin embargo se le calculan medias en todos los informes del mundo. La práctica es defendible si la escala tiene bastantes puntos y se interpreta con cautela, y es indefendible si se sacan conclusiones de una diferencia de 0,1 entre dos medias de escalas de cinco puntos.

(d) La temperatura en grados Celsius. Es cuantitativa, pero su cero es convencional: se pueden sumar y restar temperaturas, y no tiene sentido decir que 20 grados es "el doble de caliente" que 10. Los cocientes exigen un cero absoluto, y por eso el coeficiente de variación de una temperatura en Celsius es un número sin significado.

Tablas de frecuencias

Definición 17.5 (Frecuencias)**.** Sea una muestra de tamaño n en la que la variable toma k valores distintos $x_1 < x_2 < \cdots < x_k$ (o k categorías, o k clases). Para cada uno se definen:

- la *frecuencia absoluta* n_i: el número de veces que aparece x_i; cumple $\sum_{i=1}^{k} n_i = n$;
- la *frecuencia relativa* $f_i = \dfrac{n_i}{n}$, que cumple $\sum_i f_i = 1$ y suele darse en tanto por ciento;
- la *frecuencia absoluta acumulada* $N_i = n_1 + \cdots + n_i$, con $N_k = n$;
- la *frecuencia relativa acumulada* $F_i = \dfrac{N_i}{n} = f_1 + \cdots + f_i$, con $F_k = 1$.

La *tabla de frecuencias* es la tabla que recoge las cuatro columnas. Las acumuladas sólo tienen sentido si la variable está ordenada, es decir, nunca para una nominal.

Ejemplo 17.6 (Veinticinco familias, con la tabla completa)**.** Se pregunta a 25 familias por su número de hijos y se obtiene

$$0, 1, 2, 1, 3, 0, 2, 1, 1, 4, 2, 0, 1, 2, 3, 1, 0, 2, 1, 1, 4, 2, 3, 1, 2.$$

(a) La tabla.

x_i	n_i	f_i	N_i	F_i
0	4	0,16	4	0,16
1	9	0,36	13	0,52
2	7	0,28	20	0,80
3	3	0,12	23	0,92
4	2	0,08	25	1,00
total	25	1,00		

Las dos comprobaciones obligatorias: $\sum n_i = 25 = n$ y $\sum f_i = 1$.

(b) Qué se lee directamente. La moda es 1, el valor más repetido. Como $F_2 = 0{,}52$ supera por primera vez a 0,5 en el valor $x = 1$, la mediana es 1. El 80 % de las familias tiene dos hijos o menos, que es literalmente la fila $F_3 = 0{,}80$. Éstas son lecturas que la lista original de veinticinco números no permite hacer de un vistazo, y ésa es toda la utilidad de la tabla.

(c) Los resúmenes, calculados con la tabla. Con $\sum n_i x_i = 40$ y $\sum n_i x_i^2 = 96$:

$$\bar{x} = \frac{40}{25} = 1{,}6, \qquad s^2 = \frac{1}{n-1}\Big(\sum n_i x_i^2 - n\bar{x}^2\Big) = \frac{96 - 25 \cdot 2{,}56}{24} = 1{,}333333,$$

luego $s = 1{,}154701$ y el coeficiente de variación es $\mathrm{CV} = \frac{1{,}154701}{1{,}6} = 0{,}721688$, es decir, un 72 % de dispersión relativa: mucha. La media 1,6 es un valor imposible para una familia concreta, cosa perfectamente normal en variables discretas.

(d) El resumen de cinco números. Los cuartiles son $Q_1 = 1$, $Q_2 = 1$ y $Q_3 = 2$, con $\mathrm{RIC} = 1$; las vallas de Tukey caen en $1 - 1{,}5 = -0{,}5$ y $2 + 1{,}5 = 3{,}5$, de modo que las dos familias con cuatro hijos quedan marcadas como atípicas. Con datos discretos y muy concentrados esto ocurre a menudo: el criterio marca valores que no tienen nada de raro, y por eso marca candidatos, no errores.

Ejemplo 17.7 (Datos continuos agrupados en clases)**.** Cuarenta estaturas, en centímetros, resumidas en cuatro clases de amplitud 10. La *marca de clase* c_i es el punto medio del intervalo, y es el valor con el que se hacen todos los cálculos.

clase	c_i	n_i	f_i	N_i	F_i
$[150, 160)$	155	6	0,150	6	0,150
$[160, 170)$	165	14	0,350	20	0,500
$[170, 180)$	175	15	0,375	35	0,875
$[180, 190)$	185	5	0,125	40	1,000

(a) La media. $\sum n_i c_i = 6790$, luego $\bar{x} = \frac{6790}{40} = 169{,}75$ cm.

(b) La dispersión. $\sum n_i c_i^2 = 1\,155\,800$, de donde $\sum n_i (c_i - \bar{x})^2 = 1\,155\,800 - 40 \cdot 169{,}75^2 = 3197{,}5$; dividiendo entre $n - 1 = 39$ sale $s^2 = 81{,}987179$ y $s = 9{,}054677$ cm.

(c) La mediana interpolada. La clase mediana es la primera con $F_i \geq 0{,}5$; aquí $F_2 = 0{,}500$ justo, y la fórmula de interpolación $L + \frac{n/2 - N_{i-1}}{n_i} \cdot a$ da $160 + \frac{20-6}{14} \cdot 10 = 170$ cm. La clase modal es $[170, 180)$, la de mayor frecuencia.

(d) El precio de agrupar. Todos estos números son aproximados: se ha supuesto que los seis individuos de la primera clase miden exactamente 155 cm. Agrupar pierde información y suaviza; con clases demasiado anchas se borran los detalles y con clases demasiado estrechas reaparece el ruido de los datos individuales. Elegir la amplitud es la única decisión seria de todo este proceso, y la regla de Freedman–Diaconis del capítulo la resuelve razonablemente.

Regla: construir una tabla de frecuencias

1. **Decide el tipo de variable** antes de tocar nada; de ahí sale si las columnas acumuladas tienen sentido.
2. **Si es discreta con pocos valores o categórica**, una fila por valor y se acabó.
3. **Si es continua o discreta con muchos valores**, agrupa en clases: entre 5 y 15, de amplitud constante siempre que se pueda, sin huecos y sin solapes.
4. **Fija el convenio de los extremos** —aquí, cerrado por la izquierda y abierto por la derecha— y decláralo.
5. **Calcula n_i, f_i, N_i y F_i** y comprueba que $\sum n_i = n$, $\sum f_i = 1$, $N_k = n$ y $F_k = 1$. Si alguna falla, hay un dato perdido o contado dos veces.
6. **Añade la marca de clase** si hay agrupación, porque es con ella con lo que se calculan media y varianza.
7. **Dibuja**: diagrama de barras si es discreta o categórica, histograma con áreas proporcionales a las frecuencias si es continua. Con clases de amplitud desigual hay que dibujar la *densidad* n_i/a_i en el eje vertical, o el histograma miente.

Tipificar: poner todo en la misma escala

Definición 17.8 (Tipificación y puntuación típica)**.** Dada una muestra $x_1, \dots, x_n$ con media $\bar{x}$ y desviación típica $s > 0$, la *tipificación* o *estandarización* consiste en sustituir cada dato por su *puntuación típica*

$$z_i = \frac{x_i - \bar{x}}{s}.$$

El valor z_i mide a cuántas desviaciones típicas por encima o por debajo de la media está el dato i, y es un número *sin unidades*.

Proposición 17.9 (La muestra tipificada tiene media cero y desviación típica uno)**.** *Con las notaciones anteriores, $\bar{z} = 0$ y $s_z = 1$. Además la tipificación es invariante frente a cambios de unidad: si $y_i = a + bx_i$ con $b > 0$, las puntuaciones típicas de los y_i coinciden exactamente con las de los x_i.*

Demostración. Para la media,

$$\bar{z} = \frac{1}{n}\sum_i \frac{x_i - \bar{x}}{s} = \frac{1}{ns}\Big(\sum_i x_i - n\bar{x}\Big) = 0.$$

Para la dispersión, y usando ya $\bar{z} = 0$,

$$s_z^2 = \frac{1}{n-1}\sum_i (z_i - \bar{z})^2 = \frac{1}{n-1}\sum_i \frac{(x_i - \bar{x})^2}{s^2} = \frac{1}{s^2}\cdot\frac{\sum_i (x_i - \bar{x})^2}{n-1} = \frac{s^2}{s^2} = 1.$$

Para la invariancia: si $y_i = a + bx_i$ entonces $\bar{y} = a + b\bar{x}$ y $s_y = b\,s_x$, luego

$$\frac{y_i - \bar{y}}{s_y} = \frac{b(x_i - \bar{x})}{b\,s_x} = \frac{x_i - \bar{x}}{s_x}.$$

□

Ejemplo 17.10 (Comparar una nota de matemáticas con una de historia)**.** Diez alumnos obtienen en matemáticas

$$7{,}5;\ 8{,}0;\ 5{,}5;\ 9{,}0;\ 6{,}5;\ 4{,}0;\ 8{,}5;\ 7{,}0;\ 6{,}0;\ 8{,}0$$

y en historia

$$6{,}0;\ 6{,}5;\ 5{,}0;\ 7{,}0;\ 8{,}5;\ 5{,}5;\ 6{,}0;\ 7{,}5;\ 6{,}5;\ 6{,}5.$$

(a) Los resúmenes. En matemáticas $\bar{x} = 7$ y $s = 1{,}527525$; en historia $\bar{x} = 6{,}5$ y $s = 1$.

(b) Las puntuaciones típicas de matemáticas. Aplicando la fórmula salen

$$0{,}3273;\ 0{,}6547;\ -0{,}9820;\ 1{,}3093;\ -0{,}3273;\ -1{,}9640;\ 0{,}9820;\ 0;\ -0{,}6547;\ 0{,}6547.$$

Su media es cero exactamente y su cuasivarianza uno, como garantiza la proposición 17.9. El alumno con un 7,0 tiene $z = 0$: está justo en la media.

(c) La comparación que interesa. El mejor de matemáticas sacó 9,0, con $z = 1{,}3093$. El mejor de historia sacó 8,5, con $z = \frac{8{,}5 - 6{,}5}{1} = 2{,}0$. En nota bruta gana el de matemáticas; en posición relativa a su grupo gana claramente el de historia, que está a dos desviaciones

típicas del centro frente a 1,31. Ésa es exactamente la pregunta que la tipificación responde y la nota bruta no.

(d) Para qué más sirve. Tipificar es el paso previo a comparar variables con unidades distintas (kilos con centímetros), a detectar atípicos con el criterio $|z| > 3$, a calcular la correlación —que no es otra cosa que la media de los productos de las puntuaciones típicas— y a usar tablas de la normal. En el capítulo de inferencia reaparece con σ en lugar de s y recibe el nombre de tipificación de una variable aleatoria; la operación es la misma.

Definición 17.11 (Recorrido y resumen de cinco números)**.** El *recorrido* o *rango* de una muestra es $R = x_{(n)} - x_{(1)}$, la diferencia entre el máximo y el mínimo. El *resumen de cinco números* de Tukey es la lista $(x_{(1)}, Q_1, Q_2, Q_3, x_{(n)})$, y es lo que dibuja exactamente un *diagrama de caja*: una caja de Q_1 a Q_3 con la mediana marcada dentro, y dos bigotes que llegan hasta el dato más extremo que aún cae dentro de las vallas $Q_1 - 1{,}5\,\mathrm{RIC}$ y $Q_3 + 1{,}5\,\mathrm{RIC}$; los datos fuera de las vallas se dibujan sueltos. El recorrido usa sólo dos datos y por tanto tiene punto de ruptura cero: un único valor erróneo lo determina por completo.

Las dos rectas de regresión

El capítulo define la recta de regresión de y sobre x como la que minimiza $\sum(y_i - a - bx_i)^2$, es decir, la suma de los cuadrados de los errores *verticales*. Esa elección no es simétrica: si lo que se quiere es predecir x a partir de y, hay que minimizar los errores *horizontales*, y sale otra recta distinta. Hay pues dos rectas de regresión, y la relación entre ellas resume de golpe el significado de la correlación.

Definición 17.12 (Las dos rectas)**.** Dados los pares (x_i, y_i), la *recta de regresión de y sobre x* es

$$y - \bar{y} = b_{yx}\,(x - \bar{x}), \qquad b_{yx} = \frac{s_{xy}}{s_x^2},$$

que minimiza $\sum_i (y_i - a - bx_i)^2$ y se usa para predecir y conocido x. La *recta de regresión de x sobre y* es

$$x - \bar{x} = b_{xy}\,(y - \bar{y}), \qquad b_{xy} = \frac{s_{xy}}{s_y^2},$$

que minimiza $\sum_i (x_i - a' - b'y_i)^2$ y se usa para predecir x conocido y. Las dos pendientes b_{yx} y b_{xy} se llaman *coeficientes de regresión*.

Teorema 17.13 (Relación entre las dos rectas)**.** *Las dos rectas de regresión se cortan siempre en el centro de gravedad $(\bar{x}, \bar{y})$, sus pendientes tienen el mismo signo que s_{xy}, y su producto es*

$$b_{yx} \cdot b_{xy} = r^2.$$

En consecuencia: las dos rectas coinciden si y sólo si $|r| = 1$, es decir, si los puntos están alineados; y son perpendiculares —una horizontal y la otra vertical— si y sólo si $r = 0$. Cuanto menor es $|r|$, más se abren.

Demostración. Que ambas pasan por $(\bar{x},\bar{y})$ es inmediato de las ecuaciones, que están escritas en forma punto-pendiente con ese punto. La derivación de b_{xy} es la misma del teorema 17.44 sin más que intercambiar los papeles de x e y. Y para el producto,

$$b_{yx}\cdot b_{xy}=\frac{s_{xy}}{s_x^2}\cdot\frac{s_{xy}}{s_y^2}=\frac{s_{xy}^2}{s_x^2 s_y^2}=r^2.$$

Si $|r|=1$ entonces $b_{xy}=1/b_{yx}$, y una recta que pasa por $(\bar{x},\bar{y})$ con pendiente b_{yx} es la misma que la de pendiente inversa vista desde el otro eje: coinciden. Si $r=0$ entonces $s_{xy}=0$, luego $b_{yx}=0$ (la recta $y=\bar{y}$) y $b_{xy}=0$ (la recta $x=\bar{x}$), que son perpendiculares. □

Idea clave

Qué dice esto en cristiano. El ángulo que forman las dos rectas *es* la correlación, dibujada. Con datos casi alineados las dos rectas se superponen y da igual cuál se use; con datos desparramados forman una equis abierta, y entonces predecir y a partir de x y predecir x a partir de y son operaciones que dan resultados incompatibles entre sí. Ninguna de las dos rectas es "la relación entre las variables": cada una es la mejor respuesta a una pregunta distinta, y la pregunta la decide qué variable se conoce y cuál se quiere adivinar. Confundirlas —ajustar y sobre x y luego despejar x— es un error frecuente que sesga sistemáticamente hacia pendientes demasiado grandes.

Ejemplo 17.14 (Ocho puntos y las dos rectas, con todos los números)**.** Sean

x	2	4	5	6	8	9	11	13
y	3	7	4	9	7	12	9	15

(a) Los cimientos. Con $n=8$: $\sum x_i=58$, $\sum y_i=66$, luego $\bar{x}=7{,}25$ y $\bar{y}=8{,}25$. Las sumas centradas son

$$S_{xx}=\sum(x_i-\bar{x})^2=95{,}5,\qquad S_{yy}=\sum(y_i-\bar{y})^2=109{,}5,\qquad S_{xy}=\sum(x_i-\bar{x})(y_i-\bar{y})=87{,}5.$$

De ahí $s_x=3{,}693624$, $s_y=3{,}955105$, $s_{xy}=12{,}5$ y

$$r=\frac{87{,}5}{\sqrt{95{,}5\cdot 109{,}5}}=0{,}855656,\qquad r^2=0{,}732148.$$

(b) La recta de y sobre x. $b_{yx}=\frac{87{,}5}{95{,}5}=0{,}916230$ y $a=8{,}25-0{,}916230\cdot 7{,}25=1{,}607330$, es decir

$$\hat{y}=1{,}607330+0{,}916230\,x.$$

Los valores ajustados son 3,4398; 5,2723; 6,1885; 7,1047; 8,9372; 9,8534; 11,6859; 13,5183, y los residuos $-0{,}4398$; 1,7277; $-2{,}1885$; 1,8953; $-1{,}9372$; 2,1466; $-2{,}6859$; 1,4817, que suman cero.

(c) La recta de x sobre y. $b_{xy}=\frac{87{,}5}{109{,}5}=0{,}799087$ y $a'=7{,}25-0{,}799087\cdot 8{,}25=0{,}657534$, es decir $\hat{x}=0{,}657534+0{,}799087\,y$. Para dibujarla en los mismos ejes hay que despejar y:

$$y=\frac{x-0{,}657534}{0{,}799087}=-0{,}822857+1{,}251429\,x.$$

Nótese la diferencia: pendiente 0,916 frente a pendiente 1,251, un 37 % mayor. Las dos rectas se cruzan en (7,25; 8,25) y forman un ángulo de casi nueve grados.

(d) La comprobación del teorema. $b_{yx} \cdot b_{xy} = 0{,}916230 \cdot 0{,}799087 = 0{,}732148 = r^2$, exactamente. Y el coeficiente de determinación de la primera recta lo confirma por otra vía: SCT = 109,5, SCR = 80,170157 y SCE = 29,329843, de donde $R^2 = \frac{80{,}170157}{109{,}5} = 0{,}732148$. Es decir, la recta explica el 73,2 % de la variabilidad de y y deja sin explicar el 26,8 % restante; y ese mismo número mide cuánto se parecen entre sí las dos rectas.

(e) El aviso sobre la predicción. Para $x = 10$ la primera recta predice $\hat{y} = 10{,}770$. Si alguien quisiera ahora saber qué x corresponde a $y = 10$, la respuesta *no* se obtiene despejando en esa recta —daría 9,15— sino usando la segunda, que da $\hat{x} = 0{,}657534 + 0{,}799087 \cdot 10 = 8{,}648$. Media unidad de diferencia sobre un recorrido de once: no es un detalle.

Regla: elegir e interpretar la recta

1. **Decide primero qué variable vas a conocer** y cuál quieres predecir. La que se predice va en el eje vertical y es sobre la que se minimizan los cuadrados.
2. **Dibuja la nube antes de ajustar.** La recta se calcula igual de bien sobre una parábola, y sale igual de inútil.
3. **Interpreta la pendiente con sus unidades**: b_{yx} es el incremento medio de y por unidad de x, y cambia si cambias de metros a centímetros.
4. **Interpreta la ordenada en el origen con cautela**: es el valor ajustado en $x = 0$, y si $x = 0$ está fuera del recorrido de los datos no significa nada.
5. **Informa de R^2 y de r**, recordando que $R^2 = r^2$ sólo en la regresión simple.
6. **No extrapoles** fuera del recorrido observado, y no despejes una recta para usarla en el sentido contrario: para eso está la otra.
7. **No leas causalidad** en ninguna de las dos. Las dos rectas existen siempre, incluso cuando las variables no tienen nada que ver la una con la otra.

17.3. Dos variables: covarianza, correlación y regresión descriptiva

Definición 17.15. La *covarianza* muestral es $s_{xy} = \frac{1}{n-1}\sum(x_i - \bar{x})(y_i - \bar{y})$, y el *coeficiente de correlación de Pearson* es $r = \dfrac{s_{xy}}{s_x s_y} \in [-1, 1]$. La *recta de regresión* de y sobre x es la que minimiza $\sum(y_i - a - bx_i)^2$, con $b = \dfrac{s_{xy}}{s_x^2}$ y $a = \bar{y} - b\bar{x}$, y el *coeficiente de determinación* R^2 mide la fracción de la variación de y que explica el modelo. Para una regresión lineal simple, $R^2 = r^2$.

Idea clave

Qué dice esto en cristiano. La covarianza mide si las dos variables se desvían de su media en el mismo sentido, pero su valor depende de las unidades y por eso no se puede interpretar sola: cambiar de metros a centímetros la multiplica por cien. La correlación es la covarianza estandarizada, sin unidades, y por eso siempre cae entre -1 y 1. Ahora bien, mide únicamente relación *lineal*: una relación perfecta pero curva puede dar correlación cero. Y no dice nada sobre causalidad, ni sobre la pendiente: una correlación de 0,99 es compatible con una pendiente de 0,001.

Ejemplo 17.16 (Una regresión completa, con sus errores estándar)**. (a) Los datos.** Sean $x = 1, \ldots, 10$ e $y = 2{,}1;\ 3{,}9;\ 6{,}2;\ 7{,}8;\ 10{,}1;\ 12{,}2;\ 13{,}8;\ 16{,}1;\ 18{,}0;\ 20{,}2$. La covarianza muestral vale 18,4 y la correlación $r = 0{,}999670$, con $r^2 = 0{,}999340$.

(b) La recta. Sale $b = 2{,}007273$ y $a = 0$ (hasta la sexta cifra). La suma de cuadrados de los residuos es 0,219636 frente a una variación total de 332,624, luego $R^2 = 0{,}999340$, que coincide con r^2 como debe.

(c) La incertidumbre de la pendiente. El error típico residual es $s_e = \sqrt{\frac{\mathrm{SS_{res}}}{n-2}} = 0{,}165694$, y el error estándar de la pendiente $\mathrm{EE}(b) = \frac{s_e}{\sqrt{\sum (x_i - \bar{x})^2}} = 0{,}018242$. El estadístico t vale $\frac{2{,}007273}{0{,}018242} = 110{,}03$, con un p-valor de $5{,}2 \cdot 10^{-14}$: la pendiente es inequívocamente distinta de cero.

(d) El intervalo de confianza. Con $t_{8;\,0{,}975} = 2{,}306004$, el intervalo al 95 % para la pendiente es $2{,}007273 \pm 2{,}306004 \cdot 0{,}018242$, es decir $[1{,}965206;\ 2{,}049340]$. Nótese que un R^2 altísimo no exime de dar este intervalo: son informaciones distintas, una sobre el ajuste y otra sobre la precisión del coeficiente.

Cuidado

[Contraejemplo: el cuarteto de Anscombe] Francis Anscombe construyó en 1973 cuatro conjuntos de datos con *exactamente* los mismos resúmenes numéricos y gráficas radicalmente distintas. Los tres primeros tienen media de y igual a 7,50, desviación típica 2,03, correlación 0,816 y recta de regresión $y = 3{,}00 + 0{,}50x$, y sin embargo el primero es una nube lineal razonable, el segundo es una parábola perfecta y el tercero es una recta exacta con un único punto desplazado. Ningún resumen numérico los distingue; el dibujo los separa al instante. Es el argumento definitivo a favor del paso 1 de la regla anterior.

17.4. Cuando el resumen miente: agregación y confusión

Teorema 17.17 (Paradoja de Simpson)**.** *Existen conjuntos de datos en los que una asociación entre dos variables tiene un signo dentro de cada estrato de una tercera variable, y el signo* opuesto *al agregar todos los estratos.*

Idea clave

Qué dice esto en cristiano. Agregar datos de grupos distintos puede invertir por completo la conclusión, y no por un error de cálculo: los dos análisis son aritméticamente correctos. Lo que ocurre es que la variable de agrupación está asociada a la vez con la causa y con el efecto, de modo que la comparación agregada mezcla manzanas con naranjas. La lección no es que la estadística no sirva, sino que *ningún* procedimiento puramente numérico puede decidir cuál de los dos análisis es el correcto: eso depende de qué causa qué, que es una pregunta sobre el mundo y no sobre los datos.

Ejemplo 17.18 (La paradoja con números concretos)**. (a) Dos tratamientos y dos tipos de paciente.** El tratamiento A cura a 81 de 87 pacientes leves (93,1 %) y a 192 de 263 graves (73,0 %). El tratamiento B cura a 234 de 270 leves (86,7 %) y a 55 de 80 graves (68,8 %). En *ambos* grupos A es mejor que B.

(b) La agregación. Sumando, A cura 273 de 350, es decir un 78,0 %, y B cura 289 de 350, un 82,6 %. Agregado, B parece mejor. Las dos cuentas son correctas.

(c) De dónde sale la inversión. A se administró mayoritariamente a pacientes graves (263 de 350) y B a leves (270 de 350). Como los graves se curan menos con cualquier tratamiento, el promedio de A arrastra el peso de los casos difíciles. La variable "gravedad" es una *confusora*: influye en la asignación del tratamiento y en el resultado.

(d) Cuál es el análisis correcto. Aquí, el estratificado: la gravedad es anterior al tratamiento y hay que comparar dentro de cada nivel. Pero si la variable de agrupación fuera posterior al tratamiento (por ejemplo, un efecto secundario que él mismo provoca), estratificar sería el error. La regla no es "estratifica siempre" sino "dibuja primero el grafo causal".

Cuidado

[Correlación, causalidad y el rango de los datos] Que r sea alto no implica causalidad, y los ejemplos abundan: el consumo de helados correlaciona con los ahogamientos porque ambos dependen del calor. Segundo aviso: la correlación mide relación *lineal*. Para x uniformemente repartido en $[-1, 1]$ y $y = x^2$, la relación es perfecta y determinista, y sin embargo $r = 0$. Y tercero: r depende del rango muestreado. Restringiendo el rango de x, la correlación baja aunque la relación sea la misma; por eso las correlaciones calculadas sobre grupos seleccionados (los admitidos en una universidad, los contratados en una empresa) están sistemáticamente atenuadas, fenómeno conocido como restricción de rango.

17.5. La forma de una distribución: histogramas, cuantiles y colas

Centro y dispersión son dos números; la forma es todo lo demás. Dos conjuntos de datos pueden tener la misma media y la misma desviación típica y ser completamente distintos: uno simétrico y compacto, el otro con una cola larga hacia la derecha que arrastra a media

docena de valores enormes. Esta sección se ocupa de describir esa forma y, sobre todo, de las herramientas gráficas que la revelan.

Definición 17.19. Dada una muestra con media $\bar{x}$ y desviación típica s, el *coeficiente de asimetría* de Fisher y el *exceso de curtosis* son

$$g_1 = \frac{\frac{1}{n}\sum_i (x_i - \bar{x})^3}{\left(\frac{1}{n}\sum_i (x_i - \bar{x})^2\right)^{3/2}}, \qquad g_2 = \frac{\frac{1}{n}\sum_i (x_i - \bar{x})^4}{\left(\frac{1}{n}\sum_i (x_i - \bar{x})^2\right)^{2}} - 3.$$

El *recorrido intercuartílico* es $\mathrm{RIC} = Q_3 - Q_1$, y se llaman *atípicos* los datos fuera del intervalo $[Q_1 - 1{,}5\,\mathrm{RIC},\ Q_3 + 1{,}5\,\mathrm{RIC}]$.

Idea clave

Qué dice esto en cristiano. La asimetría mide hacia qué lado se estira la distribución: positiva significa cola larga a la derecha (unos pocos valores muy grandes, como en los salarios), negativa a la izquierda. La curtosis no mide "lo puntiaguda" que es la campana, como suele decirse mal, sino cuánto pesan las colas: un exceso positivo avisa de que los valores extremos son más frecuentes de lo que una normal produciría, y esa es exactamente la situación en la que las fórmulas basadas en la normal fallan. El -3 del final es una convención: la normal tiene curtosis bruta 3, así que restando tres se consigue que el exceso de la normal sea cero y sirva de referencia.

Regla: Describir la forma de unos datos en siete pasos

1. **Dibujar antes de calcular.** Un histograma y un diagrama de caja cuestan un segundo y evitan la mitad de los errores.
2. **Elegir el número de clases con criterio**: la regla de Sturges da $k = 1 + \log_2 n$ y funciona con datos casi normales; la de Freedman–Diaconis usa un ancho $h = 2\,\mathrm{RIC}/n^{1/3}$ y aguanta mucho mejor las colas largas.
3. **Comparar media y mediana**: si $\bar{x} > Q_2$ hay cola derecha; si $\bar{x} < Q_2$, cola izquierda; si coinciden, simetría (aunque la recíproca no es cierta).
4. **Calcular** g_1 para cuantificar esa asimetría, y g_2 para las colas.
5. **Marcar los atípicos** con la regla del $1{,}5\,\mathrm{RIC}$, pero *no borrarlos*: un atípico es una observación que hay que explicar, no que eliminar.
6. **Probar una transformación** si la asimetría es fuerte: el logaritmo para datos positivos con cola derecha, la raíz cuadrada para conteos.
7. **Comprobar la normalidad con un gráfico cuantil-cuantil**, no con la mera inspección del histograma: el ojo humano es pésimo juzgando campanas y excelente juzgando si unos puntos caen sobre una recta.

Ejemplo 17.20 (Doce datos con cola derecha, descritos por completo)**.** Tomamos

$$x = (2, 3, 3, 4, 4, 4, 5, 5, 6, 9, 12, 20), \qquad n = 12.$$

(a) Centro, dispersión y forma. La media es 6,416667 y la mediana 4,5: la media está muy por encima, señal inequívoca de cola derecha. La desviación típica muestral es $s = 5{,}107184$, de modo que el coeficiente de variación vale 0,795925, casi un 80 %. La asimetría de Fisher es $g_1 = 1{,}770185$ y el exceso de curtosis $g_2 = 2{,}221882$: cola derecha marcada y colas pesadas. La asimetría de Pearson $3(\bar{x} - Q_2)/s = 1{,}125865$ apunta en la misma dirección con menos aritmética.

(b) El diagrama de caja. Los cuartiles son $Q_1 = 3{,}75$, $Q_2 = 4{,}5$, $Q_3 = 6{,}75$, luego $\mathrm{RIC} = 3$ y los bigotes teóricos llegan hasta $3{,}75 - 4{,}5 = -0{,}75$ por abajo y $6{,}75 + 4{,}5 = 11{,}25$ por arriba. Quedan fuera el 12 y el 20: dos atípicos, ambos por la derecha, coherentes con $g_1 > 0$. Nótese que la caja es estrecha (3 unidades) mientras que el recorrido total es de 18: la mayor parte de los datos está apiñada y unos pocos se disparan.

(c) El histograma. Sturges pide $k = 1 + \log_2 12 = 4{,}584963$, es decir 5 clases. Freedman–Diaconis pide anchura $2 \cdot 3/12^{1/3} = 2{,}620741$, que sobre un recorrido de 18 da unas 7 clases. Con datos asimétricos como éstos, la segunda regla es preferible: Sturges supone normalidad y tiende a suavizar de más justo donde está la información.

(d) La transformación logarítmica. Tomando $\ln x$, la asimetría cae de 1,770185 a 0,707403 y la media pasa a 1,644813 frente a una mediana de 1,497866: mucho más cerca. La media geométrica es $\mathrm{e}^{1{,}644813} = 5{,}180040$ y la armónica 4,399185, ambas por debajo de la aritmética 6,416667, como garantiza la desigualdad de las medias. La correlación del gráfico cuantil-cuantil normal sube de 0,861188 en los datos originales a 0,967872 en los transformados: el logaritmo enderezó la nube.

Definición 17.21. El *gráfico cuantil-cuantil* frente a una distribución F representa los pares $\big(F^{-1}(p_i),\, x_{(i)}\big)$, donde $x_{(1)} \le \cdots \le x_{(n)}$ son los datos ordenados y $p_i = (i - 0{,}5)/n$ son las posiciones acumuladas. Si los datos proceden de F salvo cambio de escala y de origen, los puntos se alinean sobre una recta.

Ejemplo 17.22 (Leer un gráfico cuantil-cuantil)**. (a) Datos normales.** Los puntos caen sobre una recta cuya pendiente estima la desviación típica y cuya ordenada en el origen estima la media. La correlación entre los cuantiles observados y los teóricos supera 0,98 ya con muestras moderadas.

(b) Cola derecha larga. Los puntos se curvan hacia arriba en el extremo derecho, porque los datos grandes son mayores que lo previsto. Es lo que ocurre con los doce datos del ejemplo anterior, cuya correlación cuantil-cuantil es sólo 0,861188 y sube a 0,967872 tras tomar logaritmos.

(c) Colas pesadas por los dos lados. Los puntos forman una ese alargada: bajos por la izquierda y altos por la derecha. Es la firma de la t de Student con pocos grados de libertad, y avisa de que los contrastes basados en la normal serán poco fiables.

(d) Datos discretos o redondeados. Aparecen escalones horizontales, porque muchos datos comparten valor. No es una desviación de la normalidad sino un artefacto de la medición, y confundir ambas cosas es un error frecuente al informar.

Idea clave

Qué dice esto en cristiano. La idea es sencillísima: si los datos siguieran una normal, el segundo dato más pequeño de doce debería estar más o menos donde la normal pone su segundo doceavo. Se dibuja "dónde está" contra "dónde debería estar". Cuando los puntos se curvan hacia arriba en el extremo derecho, es que la cola derecha real es más larga que la normal; cuando se curvan en los dos extremos formando una ese, las dos colas son pesadas. Ningún resumen numérico transmite esta información con tanta claridad, y por eso es el gráfico que un estadístico dibuja primero cuando duda de la normalidad.

Cuidado

[Contraejemplo: la misma media y la misma varianza, distribuciones opuestas] Los conjuntos $A = (-1, -1, 1, 1)$ y $B = (0, 0, 0, 0, 0, 0, 0, 0, -3, 3)$ tienen ambos media 0; el primero es bimodal y no contiene ningún valor central, mientras que el segundo pone el 80 % de los datos exactamente en el centro y manda el resto a los extremos. Sus excesos de curtosis son -2 y $+2$, opuestos en signo, y sin embargo un informe que sólo diera media y desviación típica los presentaría como parientes cercanos. Más grave: la distribución de Cauchy tiene una forma perfectamente definida y simétrica, y *no tiene* media ni varianza; calcularlas sobre una muestra suya produce números que no convergen a nada por más datos que se acumulen.

Cuidado

[Contraejemplo: la curtosis no mide el pico] Durante décadas se enseñó que la curtosis mide lo apuntada que es la distribución. Es falso y hay contraejemplos explícitos: se pueden construir distribuciones con un pico altísimo y curtosis menor que la normal, y distribuciones planas en el centro con curtosis enorme. Lo que g_2 mide es el peso de las colas, porque la cuarta potencia amplifica brutalmente los valores lejanos y aplasta los cercanos: un dato a tres desviaciones aporta 81 veces más que uno a una desviación. Un solo atípico puede disparar la curtosis de una muestra sin que la forma central cambie en absoluto.

Pausa para pensar

La regla del 1,5 RIC la propuso Tukey en 1977 y su justificación es más sutil de lo que parece: para datos exactamente normales, ese umbral deja fuera aproximadamente el 0,7 % de las observaciones, lo que en una muestra de cien datos significa esperar cerca de un atípico *aunque no haya nada raro*. Es decir, el diagrama de caja no marca errores: marca candidatos a mirar. Con mil datos normales saldrán unos siete puntos fuera de los bigotes, y borrarlos sería destruir datos perfectamente buenos. La distinción entre "dato extremo" y "dato erróneo" no la puede tomar ninguna fórmula; la toma quien conoce de dónde salieron los números.

17.6. Errores típicos de estadística descriptiva

Cuidado

[Ocho maneras de resumir mal]

1. **Dar la media sin la dispersión**, o cualquiera de las dos sin el tamaño muestral.
2. **Usar la media en distribuciones muy asimétricas**, donde no representa a ningún individuo típico.
3. **Dividir entre n al calcular la varianza muestral** en vez de entre $n-1$.
4. **Eliminar atípicos automáticamente** sin investigar si son errores o hallazgos.
5. **Interpretar r como causalidad**, o como pendiente, o como bondad del modelo en general.
6. **Calcular r sin dibujar la nube**, lo que hace invisible cualquier no linealidad, como enseña el cuarteto de Anscombe.
7. **Agregar grupos heterogéneos** sin comprobar si la conclusión sobrevive dentro de cada uno.
8. **Extrapolar la recta fuera del rango de los datos**, donde nada garantiza que la relación siga siendo lineal.

Pausa para pensar

El diagrama de caja lo inventó John Tukey en 1970, y con él una manera entera de entender la estadística que llamó *análisis exploratorio de datos*. Su tesis era que la práctica se había volcado en confirmar hipótesis y había olvidado la fase previa de mirar sin prejuicios, con herramientas resistentes a la contaminación. La frase que resume su postura sigue siendo la mejor defensa de este capítulo: es preferible una respuesta aproximada a la pregunta correcta que una respuesta exacta a la pregunta equivocada. Tukey inventó también las palabras *bit* y *software*, y el algoritmo de la transformada rápida de Fourier.

17.7. Dos variables a la vez: correlación y la paradoja de Simpson

Resumir una variable es fácil. Resumir la relación entre dos abre la puerta a los errores más caros de la estadística aplicada, porque el número que mide la asociación es sencillo y lo que significa no lo es.

Definición 17.23. La *covarianza* muestral y el *coeficiente de correlación de Pearson* son

$$s_{xy} = \frac{1}{n-1}\sum_i (x_i - \bar{x})(y_i - \bar{y}), \qquad r = \frac{s_{xy}}{s_x s_y} \in [-1, 1].$$

El valor r mide exclusivamente la fuerza de la relación *lineal*, es adimensional e invariante frente a cambios de escala y de origen.

Idea clave

Qué dice esto en cristiano. La covarianza pregunta si, cuando x está por encima de su media, y también suele estarlo. El problema de la covarianza es que sus unidades son el producto de las unidades, de modo que su valor no se puede interpretar sin contexto: dividir por las dos desviaciones típicas lo arregla y produce un número entre -1 y 1 que ya se puede comparar entre estudios. Lo que no arregla es la palabra "lineal": r no mide asociación en general, mide alineamiento. Una relación perfecta y evidente puede dar $r = 0$.

Ejemplo 17.24 (Cuatro maneras de que la correlación engañe)**. (a) Relación perfecta con $r = 0$.** Tómese $x = -2, -1, 0, 1, 2$ e $y = x^2$. La dependencia es exacta y determinista, y sin embargo $r = 0$ porque la nube es simétrica. Cualquier conclusión del tipo "no hay relación" sería falsa.

(b) Un solo dato manda. Con nueve puntos en $(1, 1), \ldots, (9, 9)$ más un décimo en $(100, 1)$, la correlación se desploma. Un valor extremo puede mover r de 1 a casi cero, y por eso conviene mirar el diagrama de dispersión antes que el coeficiente.

(c) La paradoja de Simpson. En un proceso de admisión, el departamento A admite al 80 por ciento de los 500 hombres que solicitan y al 90 por ciento de las 100 mujeres; el departamento B admite al 10 por ciento de los 200 hombres y al 14,3 por ciento de las 700 mujeres. En ambos departamentos las mujeres son admitidas en mayor proporción. Y en el total, los hombres son admitidos en un 60 por ciento y las mujeres en un 23,75. La desigualdad se invierte al agregar.

(d) Por qué se invierte. Porque las mujeres solicitan mayoritariamente en el departamento difícil y los hombres en el fácil. La variable "departamento" es una confusora, y agregar sin tenerla en cuenta mezcla dos poblaciones distintas. Lo inquietante es que ninguno de los dos análisis es un error de cálculo: los dos números son correctos y responden a preguntas diferentes. Decidir cuál es la pregunta relevante no es estadística, es causalidad.

Cuidado

[Contraejemplo: agregar no siempre es lo malo] Se enseña la paradoja de Simpson como "hay que desagregar siempre", y eso también es falso. Si la variable por la que se separa es un efecto del tratamiento y no una causa previa —por ejemplo, separar pacientes según una respuesta que el propio tratamiento provocó— entonces desagregar introduce el sesgo en lugar de quitarlo. La regla correcta no es "desagrega" ni "agrega", sino "ajusta por las causas comunes previas y no por lo que está en el camino causal". Sin un modelo de cómo se generaron los datos, los números solos no deciden.

Pausa para pensar

El caso de las admisiones no es hipotético: es Berkeley, curso de 1973, y llegó a los tribunales. Los datos globales mostraban una diferencia de admisión aparentemente flagrante entre hombres y mujeres, y el análisis departamental de Bickel, Hammel y O'Connell, publicado en *Science* en 1975, mostró que departamento a departamento la diferencia se invertía ligeramente a favor de las mujeres. El fenómeno lo había descrito Yule en 1903 y Simpson en 1951; hicieron falta unas admisiones universitarias y una demanda para que dejara de ser una curiosidad de manual.

17.8. Asociación cuando los datos no son números: rangos y variables categóricas

El coeficiente de Pearson mide una cosa muy concreta: cuánto se parece la nube a una *recta*. Si la relación es perfecta pero curva, Pearson la infravalora; y si las variables no son números sino categorías, ni siquiera se puede calcular. Hay dos remedios, uno para cada caso.

Correlación de rangos: Spearman y Kendall

Definición 17.25 (Coeficiente de Spearman)**.** ρ_S es el coeficiente de Pearson calculado sobre los *rangos* en lugar de sobre los valores. Sin empates admite la forma abreviada

$$\rho_S = 1 - \frac{6\sum_{i=1}^{n} d_i^2}{n(n^2-1)}, \qquad d_i = \operatorname{rango}(x_i) - \operatorname{rango}(y_i).$$

Definición 17.26 (Tau de Kendall)**.** Se recorren los $\binom{n}{2}$ pares de observaciones y se clasifica cada uno como *concordante* (las dos variables se mueven en el mismo sentido) o *discordante*. Entonces

$$\tau = \frac{C - D}{\binom{n}{2}}.$$

Ejemplo 17.27 (Una relación perfecta que Pearson no ve)**.** Sean $x = 1, 2, 3, 4, 5, 6, 7$ e $y = 1;\ 1{,}4;\ 2{,}1;\ 3{,}5;\ 6{,}5;\ 13;\ 30$: cada valor de y es mayor que el anterior, sin excepción, pero la relación es exponencial, no lineal.

(a) Pearson. $r = 0{,}843940$. Un lector apresurado diría "relación fuerte pero con ruido". No hay ruido ninguno: la relación es *determinista*.

(b) Rangos. Los rangos de x son $1, \dots, 7$ y los de y también, en el mismo orden. Luego todos los d_i valen 0 y

$$\rho_S = 1 - \frac{6 \cdot 0}{7 \cdot 48} = 1.$$

Y los $\binom{7}{2} = 21$ pares son todos concordantes, luego $\tau = 21/21 = 1$.

(c) Qué ha pasado. ρ_S y τ valen 1 exactamente cuando y es una función *monótona creciente* de x, sea o no lineal. Son invariantes frente a cualquier transformación monótona de cualquiera de las dos variables: tomar logaritmos, elevar al cuadrado los positivos o pasar a escala logarítmica no los mueve ni una milésima. Pearson, en cambio, cambia con cada una de esas operaciones.

Ejemplo 17.28 (Un caso con desorden, para ver la diferencia entre los dos)**.** Con $x = 1, \dots, 6$ e $y = 2, 1, 5, 4, 3, 7$ los rangos de y son $2, 1, 5, 4, 3, 7$ sobre $1, \dots, 6 \dots$ salvo que y toma seis valores, así que sus rangos son $2, 1, 5, 4, 3, 6$. Las diferencias son $-1, 1, -2, 0, 2, 0$, de cuadrados $1, 1, 4, 0, 4, 0$, que suman 10:

$$\rho_S = 1 - \frac{6 \cdot 10}{6 \cdot 35} = 1 - \frac{60}{210} = 0{,}714286.$$

Contando pares, hay 11 concordantes y 4 discordantes de los 15 posibles, luego $\tau = (11 - 4)/15 = 0{,}466667$. Y Pearson da $r = 0{,}742316$.

Los tres números dicen "asociación positiva moderada", pero no son comparables entre sí: τ es siempre más pequeño que ρ_S ante los mismos datos (el valor 2/3 es el cociente típico), porque cuenta inversiones y no cuadrados de distancias. Al informar de una correlación de rangos hay que decir cuál de las dos es.

Idea clave

Elige Pearson si te interesa la *recta*; Spearman o Kendall si te interesa el *orden*. Y si los datos tienen atípicos, los rangos los desactivan automáticamente: cambiar el mayor valor de 30 a 3000 no altera ρ_S ni en una cifra, mientras que r se dispara.

Variables categóricas: tablas de contingencia

Cuando ambas variables son cualitativas no hay rangos que valgan: hay que contar. La herramienta es la **tabla de contingencia**, con las categorías de una variable en las filas y las de la otra en las columnas.

Ejemplo 17.29 (Tres centros y dos resultados, descritos por completo)**.** Ciento cincuenta pacientes, tres hospitales, dos desenlaces:

	mejora	no mejora	total
A	30	20	50
B	25	25	50
C	10	40	50
total	65	85	150

(a) Perfiles fila. Dividiendo cada fila por su total: 60 por ciento, 50 por ciento y 20 por ciento de mejora. El perfil marginal es $65/150 = 43{,}3$ por ciento. Comparar cada perfil con el marginal es toda la lectura descriptiva: A está por encima, C muy por debajo.

(b) Frecuencias esperadas bajo independencia. Si el hospital no influyera, cada casilla valdría total fila $\times$ total columna$/n$. Aquí todas las filas suman 50, luego se espera $50 \cdot 65/150 = 21{,}667$ mejoras y $50 \cdot 85/150 = 28{,}333$ no mejoras en cada hospital.

(c) La medida de fuerza. Se acumula la discrepancia:

$$\chi^2 = \sum \frac{(O - E)^2}{E} = 17{,}647059,$$

y como χ^2 crece con el tamaño de la muestra, se normaliza con la V de Cramér:

$$V = \sqrt{\frac{\chi^2}{n \cdot \min(f-1, c-1)}} = \sqrt{\frac{17{,}647059}{150 \cdot 1}} = 0{,}342997.$$

V vive siempre en $[0, 1]$: 0 es independencia perfecta y 1 es determinación completa. Un 0,34 es una asociación apreciable.

(d) Razón de momios. Comparando A con C, los momios de mejorar son $30/20 = 1{,}5$ y $10/40 = 0{,}25$, y su cociente es

$$\text{OR} = \frac{1{,}5}{0{,}25} = 6.$$

Un paciente de A tiene seis veces los momios de mejorar de uno de C. La razón de momios es la medida que usan la epidemiología y la regresión logística, porque no cambia si se intercambian filas por columnas, cosa que la razón de riesgos ($0{,}60/0{,}20 = 3$) no cumple.

Cuidado

V y OR describen; no demuestran nada. Con estos mismos datos, el χ^2 de 17,65 con 2 grados de libertad daría un p-valor de 0,000147, pero eso ya es inferencia y necesita las hipótesis del capítulo correspondiente. Y sobre todo: la asociación entre hospital y desenlace puede deberse por completo a la gravedad de los pacientes que recibe cada uno. Es la paradoja de Simpson esperando su turno.

17.9. Concentración y desigualdad: la curva de Lorenz y el índice de Gini

La media y la desviación típica describen dónde está el centro y cuánto se dispersan los datos, pero no dicen *cómo se reparte el total*. Dos plantillas con la misma nómina media pueden tener repartos irreconciliables. La herramienta que falta es la curva de Lorenz.

Definición 17.30 (Curva de Lorenz)**.** Ordénense los datos de menor a mayor. Para cada k se representa el punto (p_k, q_k), donde $p_k = k/n$ es la proporción acumulada de individuos y q_k es la proporción acumulada de la masa total que esos k individuos reúnen. La poligonal que une $(0, 0)$, esos puntos y $(1, 1)$ es la **curva de Lorenz**. La diagonal $q = p$ es la igualdad perfecta.

Definición 17.31 (Índice de Gini)**.** G es el doble del área comprendida entre la diagonal y la curva de Lorenz. Vale 0 en la igualdad perfecta y tiende a 1 cuando uno solo se lo lleva todo. Para datos ordenados $x_{(1)} \le \cdots \le x_{(n)}$,

$$G = \frac{2\sum_{i=1}^{n} i\, x_{(i)}}{n \sum_i x_{(i)}} - \frac{n+1}{n}.$$

Ejemplo 17.32 (Diez salarios y un ejecutivo)**.** Salarios anuales en miles de euros:

$$18,\ 19,\ 20,\ 21,\ 22,\ 24,\ 26,\ 28,\ 32,\ 190,$$

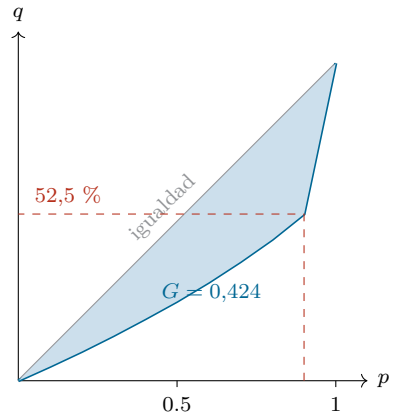


Figura 17.2: La curva de Lorenz de diez salarios con un ejecutivo. El área sombreada entre la diagonal de igualdad perfecta y la curva, multiplicada por dos, es el índice de Gini. El 90 por ciento peor pagado reúne solo el 52,5 por ciento de la masa salarial.

que suman 400.

(a) La curva, punto a punto.

p	0,1	0,2	0,3	0,4	0,5	0,6	0,7	0,8	0,9	1
q	0,045	0,0925	0,1425	0,195	0,25	0,31	0,375	0,445	0,525	1

Léase la columna $p = 0{,}9$: el 90 por ciento peor pagado reúne el 52,5 por ciento de la masa salarial. El 10 por ciento restante —una persona— se lleva el 47,5 por ciento.

(b) El índice. Aplicando la fórmula, $G = 0{,}424$.

(c) Quitando al ejecutivo. Con los nueve salarios restantes, $G = 0{,}103704$. La desigualdad se ha dividido por cuatro al retirar un solo dato: el Gini, como la media, tiene punto de ruptura 0 y un único valor extremo lo domina. Por eso conviene informarlo siempre junto a alguna medida robusta.

(d) La ratio $S80/S20$. El 20 por ciento mejor pagado reúne $32 + 190 = 222$; el 20 por ciento peor pagado, $18 + 19 = 37$. El cociente es exactamente 6. Es más burda que el Gini pero se explica en una frase, y es la que publican los organismos oficiales.

Idea clave

La curva de Lorenz responde a "¿qué fracción de la tarta se lleva qué fracción de la gente?", y el Gini la resume en un número. Su virtud es que no depende de las unidades: multiplicar todos los salarios por dos deja el Gini intacto, igual que el coeficiente de variación y a diferencia de la desviación típica. Su defecto es que dos curvas de Lorenz muy distintas pueden encerrar la misma área, y por tanto tener el mismo Gini: el número resume, la curva describe. Cuando dos curvas no se cortan, la de dentro es inequívocamente más igualitaria; cuando se cortan, ningún índice único puede zanjar la comparación.

Observación 17.33. El mismo aparato se usa fuera de la economía con otro nombre. En control de calidad, la curva de Lorenz invertida es el diagrama de Pareto ("el 20 por ciento de las

causas produce el 80 por ciento de los defectos"); en biodiversidad, el Gini de las abundancias mide la dominancia de una especie; y en aprendizaje automático, el coeficiente de Gini de un clasificador es exactamente $2 \cdot \mathrm{AUC} - 1$.

17.10. Estimadores robustos y atípicos que ninguna variable delata

Toda una escala entre la media y la mediana

La media tiene punto de ruptura 0 (un solo dato la arrastra a donde se quiera) y la mediana lo tiene 0,5 (hace falta corromper la mitad de la muestra). Entre ambas hay una familia continua.

Definición 17.34 (Media recortada y winsorizada)**.** La **media recortada** al α descarta el α por ciento de datos por cada extremo y promedia el resto. La **winsorizada** no los descarta: los *sustituye* por el valor superviviente más cercano. Ambas tienen punto de ruptura α.

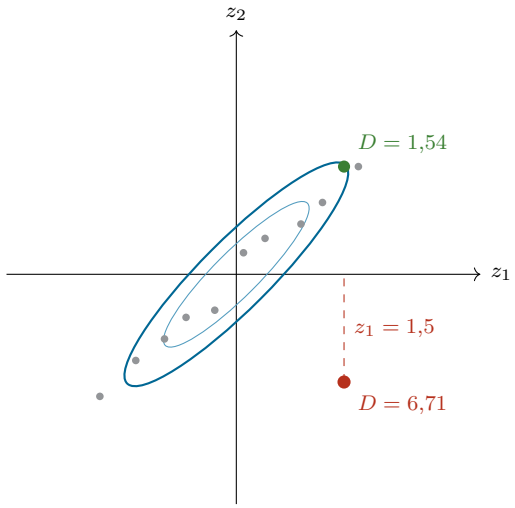


Figura 17.3: Dos puntos con coordenadas igual de extremas ($|z_i| = 1{,}5$) y destinos opuestos. Con correlación 0,9 los datos viven en la elipse alargada: el verde está donde toca y el rojo la rompe. Ninguna regla univariante distingue entre ellos; la distancia de Mahalanobis sí.

Ejemplo 17.35 (Quince datos con un 64, medidos de seis maneras)**.** Sobre los datos

$$12, 15, 15, 17, 18, 18, 19, 21, 22, 24, 25, 28, 31, 35, 64 :$$

media	24,266667	desviación típica s	12,679943
recortada al 10 %	22,153846	$1{,}4826 \cdot \mathrm{MAD}$	5,930400
recortada al 20 %	21,333333	RIC/1,349	6,671609
winsorizada al 10 %	22,533333		
winsorizada al 20 %	21,800000		
mediana	21		

Léase la primera columna de arriba abajo: al ir recortando, la estimación del centro baja monótonamente de 24,3 hacia el 21 de la mediana. El 64 empuja la media casi tres unidades y media por encima de donde está la masa de los datos.

La segunda columna es aún más elocuente. Las dos estimaciones robustas de la escala coinciden entre sí (5,93 y 6,67) y ambas son menos de la mitad de $s = 12{,}68$. Las constantes 1,4826 y 1,349 están puestas precisamente para que, si los datos fueran normales, las tres columnas estimaran el mismo σ; que no lo hagan es la prueba cuantitativa de que estos datos no son normales.

Regla: cómo elegir

- Si la muestra es limpia y aproximadamente simétrica, la media es la más eficiente y no hay razón para renunciar a ella.
- Con colas pesadas o atípicos aislados, la recortada al 10–20 por ciento conserva casi toda la eficiencia y compra mucha protección. Es la opción por defecto razonable.
- Si la contaminación puede superar el 20 por ciento, mediana y MAD.
- Informa siempre de las dos: si media y recortada coinciden, no hay nada de qué preocuparse; si difieren, la diferencia *es* el hallazgo.

El atípico que no lo parece en ninguna variable

Definición 17.36 (Distancia de Mahalanobis)**.** Para un vector x con vector de medias μ y matriz de covarianzas Σ,

$$D^2(x) = (x-\mu)^\top \Sigma^{-1} (x-\mu).$$

Es la distancia euclídea medida *después* de estandarizar y desacoplar las variables. Si los datos son normales multivariantes, D^2 sigue una χ^2 con tantos grados de libertad como dimensiones.

Ejemplo 17.37 (Un punto normal en cada eje y absurdo en el plano)**.** Dos variables estandarizadas con correlación $r = 0{,}9$: los datos viven pegados a la diagonal. Consideremos tres puntos.

(a) $z = (1{,}5;\ 1{,}5)$**.** Cada coordenada está a una vez y media su desviación típica: ninguna regla univariante lo señala. Y en efecto

$$D^2 = \frac{z_1^2 - 2r z_1 z_2 + z_2^2}{1-r^2} = \frac{2{,}25 - 2\cdot 0{,}9 \cdot 2{,}25 + 2{,}25}{0{,}19} = 2{,}368421,$$

con $D = 1{,}54$. No es atípico: está sobre la diagonal, que es donde toca.

(b) $z = (1{,}5;\ -1{,}5)$**.** Cada coordenada está exactamente igual de lejos del centro que antes, así que ninguna variable por separado lo delata. Pero ahora

$$D^2 = \frac{2{,}25 + 2 \cdot 0{,}9 \cdot 2{,}25 + 2{,}25}{0{,}19} = 45,$$

con $D = 6{,}708$. El umbral al 97,5 por ciento es $\chi^2_2 = 7{,}378$, es decir $D = 2{,}716$. Este punto lo supera con holgura: es un atípico flagrante, y ninguna caja, ningún z-score y ningún histograma marginal lo habría encontrado.

(c) $z = (2{,}5;\ 2{,}5)$**.** $D^2 = 6{,}579$, justo por debajo del umbral, aunque cada coordenada es más extrema que en los casos anteriores.

Idea clave

Mirar las variables de una en una no basta. Lo que hace atípico al punto (b) no es su posición en ningún eje, sino que *rompe la relación* que cumplen todos los demás: es alto en la primera variable y bajo en la segunda, cuando en estos datos ambas suben y bajan juntas. Un paciente de 1,90 de estatura no es raro, y 55 kilos tampoco; las dos cosas a la vez, sí. Esa es exactamente la información que la matriz Σ^{-1} aporta y que las marginales tiran a la basura.

Cuidado

μ y Σ se estiman de los mismos datos que se quieren depurar, así que un atípico severo contamina la propia regla que debería detectarlo: infla Σ y se enmascara a sí mismo. Es el efecto de *enmascaramiento*. El remedio estándar es calcular D^2 con estimaciones robustas de centro y covarianza —el determinante de covarianza mínima, MCD— en lugar de con la media y la covarianza muestrales.

17.11. Transformaciones de Box–Cox y el problema de los datos faltantes

La escalera de potencias

Definición 17.38 (Familia de Box–Cox)**.**

$$x^{(\lambda)} = \begin{cases} \dfrac{x^\lambda - 1}{\lambda} & \lambda \neq 0, \\ \log x & \lambda = 0, \end{cases}$$

para $x > 0$. La escalera de Tukey ordena las potencias $\ldots, -1, -\frac{1}{2}, 0, \frac{1}{2}, 1, 2, \ldots$: *bajar* por ella comprime la cola derecha, *subir* comprime la izquierda.

Ejemplo 17.39 (Doce datos con cola derecha, bajando la escalera)**.** Sobre $2, 3, 3, 4, 4, 4, 5, 5, 6, 9, 12, 20$:

λ	asimetría
1 (sin transformar)	2,033 8
0,5 (raíz)	1,477 4
0 (logaritmo)	0,812 7
$-0{,}5$	0,063 9

El máximo de verosimilitud de perfil da $\lambda = -0{,}528$, y en efecto es donde la asimetría prácticamente se anula. Los datos, que parecían desesperadamente sesgados, son casi simétricos en la escala $-1/\sqrt{x}$.

Regla: transformaciones estabilizadoras

Hay tres que aparecen una y otra vez, y las tres tienen una razón teórica.

- **Raíz cuadrada para conteos.** En una Poisson, Var $= \mu$; con $\sqrt{X}$ la varianza es aproximadamente $1/4$, constante.
- **Arcoseno de la raíz para proporciones.** En una binomial, Var $= p(1-p)/n$; con $\arcsin\sqrt{\hat{p}}$ la varianza es $1/(4n)$.
- **Logaritmo cuando la desviación típica es proporcional a la media**, es decir, cuando el error es multiplicativo y no aditivo. Es el caso de casi todo dato económico o biológico.

Cuidado

Interpretar en la escala transformada y luego destransformar sin cuidado es un error frecuente y silencioso. La media de los logaritmos, al exponenciarla, da la *media geométrica*, no la aritmética: para 2 y 8 da 4 y no 5. Si lo que se quiere informar es una media aritmética en la escala original, hay que corregir con la varianza (para una lognormal, $E[X] = e^{\mu+\sigma^2/2}$) o no transformar y usar errores robustos.

Datos faltantes

Definición 17.40 (Los tres mecanismos)**.**

- **MCAR**: la ausencia es independiente de todo. Un tubo se rompe en el laboratorio.
- **MAR**: la ausencia depende de las variables *observadas*. Los hombres contestan menos a la pregunta del peso, y el sexo está registrado.
- **MNAR**: la ausencia depende del valor *no observado*. Los que ganan mucho no declaran su renta.

Solo bajo MCAR el análisis de casos completos es insesgado. Bajo MAR es sesgado pero recuperable si se modela; bajo MNAR no hay solución sin supuestos externos, y lo único honrado es hacer un análisis de sensibilidad.

Ejemplo 17.41 (Cuánto sesga borrar los valores altos)**.** Sobre los doce datos anteriores, con media 6,417 y desviación típica 5,107: si se pierden los tres mayores (9, 12 y 20), la media baja a 4,0 y la desviación típica a 1,225. El centro cae un 38 por ciento y la dispersión se divide por más de cuatro. La pérdida de *tres* datos de doce ha destruido la descripción entera, y ningún resumen calculado sobre los nueve supervivientes avisa de ello.

Regla: qué hacer

1. **Dibuja el patrón de ausencias** antes que nada: una matriz de presencia/ausencia ordenada revela si faltan al azar o por bloques.
2. **No imputes por la media.** Rellena el hueco con $\bar{x}$ y habrás contraído artificialmente la varianza y atenuado todas las correlaciones hacia cero: el remedio introduce un sesgo nuevo, y encima invisible.
3. **Imputación por regresión** es mejor pero también contrae, porque pone cada valor exactamente sobre la recta. La corrección es añadir un residuo aleatorio (imputación estocástica).
4. **Imputación múltiple**: genera m conjuntos completos con imputación estocástica, analiza los m por separado y combina con las reglas de Rubin, que suman la varianza *dentro* y *entre* imputaciones. Es lo único que propaga honestamente la incertidumbre de no saber.

17.12. Por qué la recta de mínimos cuadrados es esa, y por qué $R^2 = r^2$

Tres fórmulas del capítulo se han dado ya resueltas: que la correlación está entre -1 y 1, que la recta de regresión tiene pendiente $b = s_{xy}/s_x^2$, y que el coeficiente de determinación coincide con el cuadrado de la correlación. Las tres se demuestran en cinco líneas cada una, y sin ellas la regresión es magia.

Teorema 17.42 (La correlación está entre -1 y 1)**.** *Para cualesquiera datos $(x_i, y_i)_{i\le n}$ no constantes,*

$$r = \frac{s_{xy}}{s_x s_y} \in [-1, 1],$$

y $|r| = 1$ si y sólo si los puntos están exactamente sobre una recta.

Demostración. Escríbanse $u_i = x_i - \bar{x}$ y $v_i = y_i - \bar{y}$, de modo que $s_{xy} = \frac{1}{n-1}\sum u_i v_i$, $s_x^2 = \frac{1}{n-1}\sum u_i^2$ y $s_y^2 = \frac{1}{n-1}\sum v_i^2$. La desigualdad de Cauchy–Schwarz en $\mathbb{R}^n$ dice

$$\Big(\sum_i u_i v_i\Big)^2 \;\le\; \Big(\sum_i u_i^2\Big)\Big(\sum_i v_i^2\Big),$$

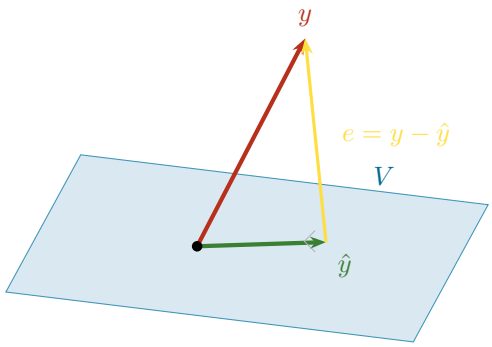


Figura 17.4: Los tres teoremas de la sección son un solo hecho de geometría en $\mathbb{R}^n$. El vector de datos y se proyecta ortogonalmente sobre el subespacio V generado por la constante y el predictor; el ajuste $\hat{y}$ es esa proyección, las ecuaciones normales son la condición de que el residuo e sea perpendicular a V, la descomposición de la varianza es el teorema de Pitágoras, y R^2 es el coseno al cuadrado del ángulo entre y y $\hat{y}$.

que dividida entre $(n-1)^2$ es exactamente $s_{xy}^2 \le s_x^2 s_y^2$, o sea $r^2 \le 1$. La igualdad en Cauchy–Schwarz se da si y sólo si los vectores u y v son proporcionales, es decir $y_i - \bar{y} = c\,(x_i - \bar{x})$ para todo i: los puntos están alineados. □

Observación 17.43 (Cauchy–Schwarz, la desigualdad que aparece en todas partes)*.* Es la misma desigualdad que en el capítulo 7 dice $|\langle u, v\rangle| \le \|u\|\|v\|$ y en el 31 da $\|fg\|_1 \le \|f\|_2\|g\|_2$. Aquí adquiere un significado geométrico concreto: la correlación es el *coseno del ángulo* entre los vectores de datos centrados,

$$r = \cos\theta, \qquad \theta = \angle(u, v).$$

Correlación 0 significa vectores perpendiculares; correlación ± 1, vectores paralelos. Y así se entiende de golpe por qué la correlación no detecta relaciones no lineales: sólo mide el ángulo con una recta.

Teorema 17.44 (La recta de mínimos cuadrados)**.** *Entre todas las rectas $y = a + bx$, la que minimiza la suma de cuadrados de los residuos*

$$Q(a, b) = \sum_{i=1}^{n} \big(y_i - a - bx_i\big)^2$$

es la única con

$$b = \frac{s_{xy}}{s_x^2}, \qquad a = \bar{y} - b\bar{x}.$$

En particular, la recta pasa siempre por el punto $(\bar{x}, \bar{y})$.

Demostración. Q es una función cuadrática de dos variables, y es convexa: su hessiana es $2\begin{pmatrix} n & \sum x_i \\ \sum x_i & \sum x_i^2 \end{pmatrix}$, cuyo determinante vale $4\big(n\sum x_i^2 - (\sum x_i)^2\big) = 4n(n-1)s_x^2 > 0$ mientras los datos no sean todos iguales. Basta pues anular el gradiente:

$$\frac{\partial Q}{\partial a} = -2\sum_i (y_i - a - bx_i) = 0 \quad \Longrightarrow \quad \bar{y} = a + b\bar{x},$$

que es la segunda fórmula y dice que la recta pasa por el centro de gravedad. Y

$$\frac{\partial Q}{\partial b} = -2\sum_i x_i(y_i - a - bx_i) = 0.$$

Sustituyendo $a = \bar{y} - b\bar{x}$ en esta segunda ecuación y agrupando en torno a las medias,

$$\sum_i x_i[(y_i - \bar{y}) - b(x_i - \bar{x})] = 0 \quad \Longrightarrow \quad \sum_i (x_i - \bar{x})(y_i - \bar{y}) = b\sum_i (x_i - \bar{x})^2,$$

donde se ha podido restar $\bar{x}$ dentro del corchete porque $\sum_i[(y_i - \bar{y}) - b(x_i - \bar{x})] = 0$ por la primera ecuación. Dividiendo entre $n-1$: $s_{xy} = b\, s_x^2$. □

Teorema 17.45 (Descomposición de la varianza y $R^2 = r^2$)**.** *Sean $\hat{y}_i = a + bx_i$ los valores ajustados por la recta de mínimos cuadrados. Entonces*

$$\underbrace{\sum_i (y_i - \bar{y})^2}_{total} = \underbrace{\sum_i (\hat{y}_i - \bar{y})^2}_{explicada} + \underbrace{\sum_i (y_i - \hat{y}_i)^2}_{residual},$$

y en consecuencia

$$R^2 = \frac{explicada}{total} = r^2.$$

Demostración. Descomposición. Escríbase $y_i - \bar{y} = (\hat{y}_i - \bar{y}) + (y_i - \hat{y}_i)$ y elévese al cuadrado sumando:

$$\sum (y_i - \bar{y})^2 = \sum (\hat{y}_i - \bar{y})^2 + \sum (y_i - \hat{y}_i)^2 + 2\sum (\hat{y}_i - \bar{y})(y_i - \hat{y}_i).$$

El término cruzado se anula. En efecto, $\hat{y}_i - \bar{y} = b(x_i - \bar{x})$, luego ese término vale $2b\sum (x_i - \bar{x})e_i$ con $e_i = y_i - \hat{y}_i$ el residuo; y las dos ecuaciones normales de la demostración anterior dicen exactamente $\sum e_i = 0$ y $\sum x_i e_i = 0$, de donde $\sum (x_i - \bar{x})e_i = 0$. Geométricamente: *el vector de residuos es ortogonal al de predictores*, que es la definición misma de proyección ortogonal.

La identidad $R^2 = r^2$. Usando $\hat{y}_i - \bar{y} = b(x_i - \bar{x})$ y $b = s_{xy}/s_x^2$,

$$R^2 = \frac{b^2 \sum (x_i - \bar{x})^2}{\sum (y_i - \bar{y})^2} = \frac{b^2\, s_x^2}{s_y^2} = \frac{s_{xy}^2}{s_x^4} \cdot \frac{s_x^2}{s_y^2} = \frac{s_{xy}^2}{s_x^2 s_y^2} = r^2.$$ □

Observación 17.46 (Regresión es proyección)*.* Los tres teoremas anteriores son un solo hecho de álgebra lineal. Considérese el vector $y \in \mathbb{R}^n$ y el subespacio V generado por el vector de unos y el vector x. La recta de mínimos cuadrados es la *proyección ortogonal* de y sobre V; las ecuaciones normales son la condición de ortogonalidad del residuo; el teorema de Pitágoras en $\mathbb{R}^n$ es la descomposición de la varianza; y $R^2 = \cos^2$ del ángulo entre $y - \bar{y}\mathbf{1}$ y su proyección. Entender esto —que la estadística de la regresión *es* geometría euclídea en dimensión n— vuelve innecesario memorizar las fórmulas, y es la puerta de entrada al modelo lineal general del capítulo 19, donde V tiene más dimensiones pero no ocurre nada nuevo.

17.13. Los conceptos que faltaban por escrito

Definición 17.47 (Cuantiles, cuartiles y percentiles)**.** Dada una muestra ordenada $x_{(1)} \le \cdots \le x_{(n)}$ y $p \in (0,1)$, el *cuantil muestral* de orden p se calcula así (convenio de interpolación lineal, que es el que usa este libro y el que traen por defecto la mayoría de programas):

1. se forma el índice $h = (n-1)p + 1$;
2. si h es entero, $Q_p = x_{(h)}$;
3. si no, siendo $\lfloor h \rfloor$ su parte entera y $\{h\}$ su parte decimal,

$$Q_p = x_{(\lfloor h \rfloor)} + \{h\} \cdot \big(x_{(\lfloor h \rfloor + 1)} - x_{(\lfloor h \rfloor)}\big).$$

Los *cuartiles* son $Q_1 = Q_{0{,}25}$, $Q_2 = Q_{0{,}5}$ (la mediana) y $Q_3 = Q_{0{,}75}$; el *percentil* k es $Q_{k/100}$. El *rango intercuartílico* es $\mathrm{RIC} = Q_3 - Q_1$.

Cuidado

[Hay al menos nueve convenios, y no coinciden] Con los datos $2, 4, 6, 8$ y $p = 0{,}25$: el convenio de arriba da $h = (4-1)(0{,}25) + 1 = 1{,}75$, luego $Q_1 = 2 + 0{,}75 \cdot (4-2) = 3{,}5$. Otros programas devuelven 3 (mediana de la mitad inferior, con Tukey), 2 (el orden estadístico más próximo por abajo) o 4. Ninguno está mal: son convenios distintos, todos con el mismo límite cuando n crece. Lo que *sí* está mal es no decir cuál se usa, y por eso conviene declararlo en cualquier informe.

Definición 17.48 (Moda y coeficiente de variación)**.** La *moda* de un conjunto de datos es el valor (o los valores) que más veces se repite; si ninguno se repite, se dice que no hay moda, y si hay dos máximos empatados la distribución es *bimodal*. En datos agrupados se habla de *clase modal*, el intervalo de mayor frecuencia.

El *coeficiente de variación* es

$$\mathrm{CV} = \frac{s}{\bar{x}} \qquad \text{(a menudo expresado en tanto por ciento)},$$

la desviación típica medida en unidades de la media. Sirve para comparar la dispersión de magnitudes con unidades distintas, y *sólo* tiene sentido si los datos son positivos y la media está lejos de cero.

Ejemplo 17.49 (Cuándo el coeficiente de variación dice algo y cuándo no)**.** Los pesos de un grupo de personas tienen $\bar{x} = 70$ kg y $s = 10$ kg, luego $\mathrm{CV} = 0{,}143$; y sus alturas, $\bar{x} = 1{,}70$ m y $s = 0{,}09$ m, con $\mathrm{CV} = 0{,}053$. La conclusión —el peso varía casi tres veces más que la altura, en términos relativos— es legítima y no dependería de que se midiera en libras y pulgadas.

En cambio, para una variable centrada en cero —una diferencia de temperaturas, un rendimiento que puede ser negativo— el coeficiente de variación es basura: la media puede acercarse a cero y dispararlo sin que la dispersión haya cambiado. Ésa es la razón por la que en finanzas se usa el cociente de Sharpe con un exceso de rentabilidad, y no el CV a secas.

Regla: Describir una tabla de contingencia $r \times c$

1. **Frecuencias marginales.** Suma por filas (f_i) y por columnas (c_j); comprueba que ambas sumas totales dan n.
2. **Perfiles.** Divide cada fila por su marginal (perfil fila) o cada columna por la suya (perfil columna). *La elección no es libre*: se condiciona por la variable explicativa. Si dos perfiles fila son casi iguales, no hay asociación.
3. **Frecuencias esperadas bajo independencia:** $e_{ij} = \dfrac{f_i\, c_j}{n}$. Es lo que habría si las dos variables no tuvieran nada que ver.
4. **Medir la asociación.** Con $\chi^2 = \sum_{i,j} \dfrac{(o_{ij} - e_{ij})^2}{e_{ij}}$, la *V de Cramér* es
$$V = \sqrt{\frac{\chi^2}{n\,\min(r-1,\, c-1)}} \in [0,1],$$
que vale 0 si hay independencia perfecta y 1 si la asociación es total. A diferencia de χ^2, no crece con el tamaño de muestra.
5. **En tablas** 2×2, la *razón de momios* (*odds ratio*)
$$\mathrm{OR} = \frac{o_{11}\, o_{22}}{o_{12}\, o_{21}}$$
compara las probabilidades relativas: $\mathrm{OR} = 1$ es independencia, > 1 asociación positiva. Es la medida que usan la epidemiología y la regresión logística, y no depende de qué variable se tome como explicativa.
6. **Comprueba** que ninguna e_{ij} baje de 5 antes de tomarse en serio la χ^2.

17.14. Ejercicios resueltos

Tres problemas sobre el mismo conjunto de datos, para ver cómo cada resumen cuenta una parte y esconde otra.

Ejercicio 17.1. Para los datos $4, 7, 7, 9, 12, 13, 15, 20, 20, 33$ calcule media, mediana, moda, rango, cuartiles, recorrido intercuartílico, varianza muestral, desviación típica y coeficiente de variación.

Solución. La media es $140/10 = 14$. La mediana es el promedio de los dos centrales, $(12+13)/2 = 12{,}5$. Hay dos modas, 7 y 20: la distribución es bimodal. El rango es $33 - 4 = 29$.

Con el criterio de los percentiles interpolados, $Q_1 = 7{,}5$ y $Q_3 = 18{,}75$, luego RIC $= 11{,}25$. La varianza muestral es 73,556 y la desviación típica 8,577. El coeficiente de variación es $8{,}577/14 = 0{,}613$, es decir, un 61 por ciento: mucha dispersión relativa.

Lo interesante es la discrepancia entre media y mediana. La media supera a la mediana en 1,5 porque el 33 tira de ella hacia arriba; la mediana no se entera. Es la primera señal de asimetría por la derecha, y el coeficiente de Fisher lo confirma: vale 1,205, claramente positivo. ■

Ejercicio 17.2. Con los mismos datos, decida si el 33 es un valor atípico según el criterio de las vallas de Tukey. Discuta si eso significa que deba eliminarse.

Solución. Las vallas son $Q_1 - 1{,}5\,\mathrm{RIC} = -9{,}375$ y $Q_3 + 1{,}5\,\mathrm{RIC} = 35{,}625$. Como $33 < 35{,}625$, el criterio *no* marca el 33 como atípico, aunque a ojo lo parezca.

Y aunque lo marcase, atípico no significa erróneo. El criterio de Tukey es una convención —en una normal marca alrededor del 0,7 por ciento de los datos, así que en una muestra grande siempre marca algunos— y su función es llamar la atención, no autorizar un borrado. Eliminar un dato sólo está justificado si hay una razón externa para creerlo mal medido: un sensor averiado, un error de transcripción. Quitarlo porque estorba es fabricar el resultado. ■

Ejercicio 17.3. Dos grupos tienen la misma media, 50, y desviaciones típicas 2 y 18. Explique qué información añade la desviación típica y qué situación real puede describir cada caso.

Solución. La media localiza el centro y no dice nada sobre la homogeneidad. Con $\sigma = 2$, la regla empírica sitúa casi todos los datos entre 44 y 56: un grupo compacto. Con $\sigma = 18$, el mismo argumento da un rango de -4 a 104, es decir, un grupo en el que hay de todo y en el que la media describe a muy poca gente.

Un caso real: dos clases con la misma nota media de 50. En la primera, todos rondan el aprobado justo y el problema es colectivo. En la segunda hay un grupo que domina la materia y otro que no se ha enterado de nada, y la intervención necesaria es completamente distinta. El promedio es idéntico y la conclusión pedagógica es opuesta: por eso ningún resumen numérico debería presentarse solo, y por eso conviene mirar siempre el histograma antes que la tabla. ■

17.15. Ejercicios propuestos

17.1. Para $2, 4, 4, 5, 9$: calcule media, mediana, moda y desviación típica muestral.

17.2. Si a todos los datos se les suma 10, ¿qué le pasa a la media, a la mediana y a la desviación típica?

17.3. Si todos se multiplican por 3, ¿qué le pasa a las tres?

17.4. ¿Cuándo es preferible la mediana a la media? Dé un ejemplo con salarios.

17.5. Calcule el coeficiente de variación de un conjunto con media 80 y desviación 12, y compárelo con otro de media 8 y desviación 2,4.

17.6. ¿Qué significa que la asimetría sea negativa? Dibuje mentalmente el histograma.

17.7. Con $Q_1 = 15$ y $Q_3 = 27$, ¿a partir de qué valores marca Tukey los atípicos?

17.8. ¿Puede un conjunto de datos no tener moda? ¿Puede tener tres?

17.9. Explique la diferencia entre dividir por n y por $n - 1$ al calcular la varianza.

17.10. Un histograma con dos picos claros, ¿qué suele indicar sobre la población?

17.11. ¿Por qué un diagrama de caja no revela la bimodalidad y un histograma sí?

17.12. El cuarteto de Anscombe: cuatro conjuntos con idénticas media, varianza, correlación y recta de regresión. ¿Qué moraleja se extrae?

Capítulo 18

Inferencia estadística

Lo que necesitas traer

Del capítulo 16: distribuciones, esperanza, varianza, el teorema central del límite y la distribución de la media muestral. Del capítulo 6: derivadas, para maximizar la verosimilitud. Del capítulo 17: cuantiles y medidas de dispersión.

La estadística descriptiva resume lo que hay; la inferencia da el salto de la muestra a la población. Ese salto nunca es seguro, y toda la disciplina consiste en cuantificar exactamente cuán inseguro es. Este capítulo trata las tres preguntas centrales: cómo construir un buen estimador, cómo dar un intervalo en vez de un número, y cómo decidir entre dos hipótesis controlando la probabilidad de equivocarse.

18.1. Estimadores: sesgo, varianza y error cuadrático medio

Definición 18.1. Un *estimador* $\hat{\theta}$ de un parámetro θ es una función de la muestra. Su *sesgo* es $\operatorname{sesgo}(\hat{\theta}) = E(\hat{\theta}) - \theta$, y es *insesgado* si vale cero. El *error cuadrático medio* es $\operatorname{ECM}(\hat{\theta}) = E\left[(\hat{\theta}-\theta)^2\right] = \operatorname{Var}(\hat{\theta}) + \operatorname{sesgo}(\hat{\theta})^2$. Es *consistente* si $\hat{\theta} \to \theta$ en probabilidad al crecer n, y *eficiente* si alcanza la cota inferior de la varianza que permite el teorema de Cramér–Rao.

Idea clave

Qué dice esto en cristiano. Un estimador es una receta para adivinar, y hay dos maneras de fallar: apuntar sistemáticamente mal (sesgo) o apuntar bien pero con pulso tembloroso (varianza). La descomposición del error cuadrático medio dice que el error total es la suma de ambos efectos, y de ahí sale una consecuencia que sorprende a todo el mundo: a veces conviene aceptar un poco de sesgo a cambio de reducir mucho la varianza. Insesgado no significa mejor, y buscar la insesgadez a toda costa es un reflejo que la estadística moderna ha abandonado.

Ejemplo 18.2 (Estimar la varianza: el precio de la insesgadez)**. (a) El estimador natural es sesgado.** Si se divide entre n, $\hat{\sigma}^2_n = \frac{1}{n}\sum(x_i - \bar{x})^2$ cumple $E(\hat{\sigma}^2_n) = \frac{n-1}{n}\sigma^2$: para $n = 10$

subestima sistemáticamente en un factor 0,9. El sesgo es $-\frac{\sigma^2}{n}$.

(b) La corrección. Dividiendo entre $n-1$ se obtiene s^2, que sí es insesgado. La razón intuitiva es que $\bar{x}$ está calculado *a partir de los mismos datos* y por construcción queda más cerca de ellos que la verdadera media, de modo que las desviaciones salen algo pequeñas; el $n-1$ compensa ese grado de libertad consumido.

(c) Y sin embargo el sesgado gana en ECM. Bajo normalidad, $\mathrm{Var}(s^2) = \frac{2\sigma^4}{n-1}$, mientras que el estimador dividido entre n tiene varianza menor. Calculando el error cuadrático medio total, el ganador no es s^2 sino el que divide entre $n+1$. La insesgadez tiene un precio, y a veces no compensa pagarlo.

(d) Consistencia. Ambos son consistentes, porque el sesgo $\frac{\sigma^2}{n}$ tiende a cero y la varianza también. Consistencia y insesgadez son propiedades independientes: hay estimadores insesgados no consistentes (tomar siempre x_1 como estimador de la media es insesgado y nunca mejora) y sesgados consistentes, como el de (a).

Cuidado

[Insesgado no implica bueno, ni siquiera sensato] Un estimador puede ser insesgado y absurdo. Para estimar $\mathrm{e}^{-2\lambda}$ en una Poisson con una sola observación X, el único estimador insesgado es $(-1)^X$, que vale $+1$ o -1 alternativamente y estima una cantidad que está entre 0 y 1. Es insesgado y es una tontería. Segundo aviso: la insesgadez no se conserva por transformaciones. Si $\hat{\theta}$ es insesgado para θ, entonces $\hat{\theta}^2$ *no* lo es para θ^2, porque $E(\hat{\theta}^2) = \mathrm{Var}(\hat{\theta}) + \theta^2 > \theta^2$. Por eso s no es un estimador insesgado de σ aunque s^2 lo sea de σ^2.

18.2. De la población a la muestra: el estadístico, su distribución en el muestreo y lo que de verdad promete un intervalo de confianza

El capítulo empieza hablando de estimadores como si ya estuviera claro qué se estima, a partir de qué y por qué eso funciona. No lo está. Antes del primer estimador hay tres objetos que conviene tener escritos con todas las letras: la *población* de la que se quiere saber algo, la *muestra* que efectivamente se observa, y el *estadístico*, que es cualquier número calculado a partir de la muestra. El salto de la muestra a la población es todo el capítulo; y ese salto se apoya en una idea que se enuncia aquí y que después se usa constantemente sin volver a mencionarse: un estadístico, por ser función de datos que salieron al azar, es él mismo una variable aleatoria, y tiene por tanto su propia distribución de probabilidad. Esa distribución se llama *distribución en el muestreo*, y es el objeto sobre el que descansa absolutamente todo lo demás.

Población, muestra y muestreo aleatorio simple

Definición 18.3 (Población, individuo y muestra)**.** La *población* es el conjunto completo de elementos —llamados *individuos* o *unidades*— sobre los que se quiere concluir algo. Su tamaño se denota N, y puede ser finito o infinito. Una *muestra* es cualquier subconjunto de n

individuos de la población que efectivamente se observa; n es el *tamaño muestral*. Observar la población entera se llama *censo*. La lista concreta de la que se sortea la muestra se llama *marco muestral*.

Una característica numérica de la población —su media μ, su varianza σ^2, una proporción p— se llama *parámetro*, se escribe con letra griega y es un número fijo y desconocido. La misma característica calculada sobre la muestra —$\bar{x}$, s^2, $\hat{p}$— se llama *estadístico* y es un número conocido pero aleatorio.

Definición 18.4 (Muestreo aleatorio simple)**.** Un *muestreo aleatorio simple* de tamaño n es un procedimiento de selección en el que *todas* las muestras posibles de tamaño n tienen la misma probabilidad de ser elegidas. Hay dos versiones:

- *con reemplazamiento*: cada individuo se devuelve a la población antes de extraer el siguiente, de modo que las extracciones son independientes y un individuo puede salir varias veces;
- *sin reemplazamiento*: no se devuelve, de modo que las extracciones ya no son independientes, aunque la dependencia es despreciable cuando n es pequeño frente a N.

En este capítulo, y salvo aviso, "muestra" significa muestra aleatoria simple con reemplazamiento, o equivalentemente extraída de una población infinita.

Idea clave

Qué dice esto en cristiano. La palabra "aleatorio" no significa "sin criterio". Significa exactamente lo contrario: significa que el criterio es un sorteo con probabilidades conocidas y controladas por quien diseña el estudio. Preguntar a los primeros cien transeúntes no es muestreo aleatorio, es muestreo de conveniencia, y ninguna de las fórmulas de este capítulo vale para él. Toda la maquinaria posterior —errores típicos, intervalos, p-valores— calcula probabilidades *sobre el sorteo*; si no hubo sorteo, esas probabilidades no significan nada, por muy bien impresas que salgan. Un sesgo de selección no se corrige con más datos: se agranda.

Definición 18.5 (Muestra aleatoria como vector de variables aleatorias)**.** Sea X una variable aleatoria con función de distribución F —el modelo probabilístico de la población—. Una *muestra aleatoria de tamaño n* de F es un vector

$$(X_1, X_2, \ldots, X_n)$$

de variables aleatorias *independientes* y con *la misma distribución* F que X; se abrevia diciendo que las X_i son *independientes e idénticamente distribuidas*, o *iid*. Una vez realizado el experimento, los valores concretos $x_1, \ldots, x_n$ que se observan se llaman *realización* de la muestra.

Cuidado

[Mayúsculas y minúsculas no son un capricho tipográfico] X_i (mayúscula) es la variable aleatoria: el resultado *antes* de mirar, con toda su incertidumbre. x_i (minúscula) es el número concreto *después* de mirar, sin ninguna incertidumbre. Del mismo modo, $\bar{X} = \frac{1}{n}\sum X_i$ es una variable aleatoria con su distribución, su media y su varianza, mientras que $\bar{x} = 102{,}5$ es un número. Todos los teoremas del capítulo hablan de las mayúsculas; todas las cuentas de los ejemplos, de las minúsculas. Confundirlas es el origen del error clásico "la probabilidad de que μ esté en mi intervalo", porque el intervalo concreto ya no tiene nada de aleatorio.

Ejemplo 18.6 (Una urna con cinco fichas, agotada por completo)**.** Una población de $N = 5$ fichas marcadas con 2, 4, 6, 8 y 10. Aquí se conoce todo, y por eso sirve para ver el mecanismo entero.

(a) Los parámetros. La media poblacional es $\mu = 6$ y la varianza poblacional $\sigma^2 = \frac{1}{5}\sum(x_i - \mu)^2 = 8$, luego $\sigma = 2{,}828427$. Son dos números fijos, sin nada aleatorio.

(b) Todas las muestras de tamaño 2 con reemplazamiento. Hay $5^2 = 25$, todas con probabilidad $\frac{1}{25}$. La media muestral $\bar{X}$ toma estos valores con estas frecuencias:

$\bar{x}$	2	3	4	5	6	7	8	9	10
casos	1	2	3	4	5	4	3	2	1

Ésa —y no otra cosa— es la *distribución en el muestreo* de $\bar{X}$. Obsérvese que es triangular y simétrica alrededor de 6, mientras que la población era plana: promediar ya concentra.

(c) Su media y su varianza, contadas a mano. Sumando $\sum \bar{x} \cdot \text{casos}/25$ sale $\mathrm{E}(\bar{X}) = 6 = \mu$, y del mismo modo $\mathrm{Var}(\bar{X}) = 4 = \frac{\sigma^2}{2} = \frac{\sigma^2}{n}$. La media muestral está centrada en el parámetro y es *la mitad* de dispersa que un dato suelto en varianza, es decir $\sqrt{2}$ veces menos en desviación típica. Además, promediando las 25 cuasivarianzas s^2 se obtiene exactamente $8 = \sigma^2$: es la insesgadez del divisor $n - 1$, comprobada por enumeración y no por fe.

(d) Y sin reemplazamiento. Las muestras posibles pasan a ser $\binom{5}{2} = 10$, y la varianza de $\bar{X}$ cae de 4 a 3. La fórmula general es $\mathrm{Var}(\bar{X}) = \frac{\sigma^2}{n} \cdot \frac{N-n}{N-1}$, y el factor $\frac{N-n}{N-1} = \frac{3}{4}$ se llama *corrección por población finita*. Muestrear sin reemplazamiento es más preciso porque no se malgasta información repitiendo individuos; cuando N es enorme frente a n ese factor vale casi 1 y se ignora, que es lo que se hace en todo el resto del capítulo.

Estadísticos y distribución en el muestreo

Definición 18.7 (Estadístico y distribución en el muestreo)**.** Un *estadístico* es cualquier función $T = T(X_1, \ldots, X_n)$ de la muestra que *no dependa de parámetros desconocidos*. Por ser función de variables aleatorias, T es a su vez una variable aleatoria; su distribución de probabilidad se llama *distribución en el muestreo* de T. Un estadístico que se usa para aproximar un parámetro θ se llama *estimador* de θ y se escribe $\hat{\theta}$.

La exigencia de que T no dependa de parámetros desconocidos es lo que separa un estadístico de una expresión inútil: $\frac{\bar{X}-\mu}{\sigma/\sqrt{n}}$ no es un estadístico si μ y σ son desconocidos, porque no se

puede calcular; $\frac{\bar X-\mu_0}{s/\sqrt{n}}$ sí lo es, porque μ_0 es un número que fija quien contrasta y s se calcula con los datos.

Proposición 18.8 (Media y varianza de la media muestral)**.** *Sea $X_1,\dots,X_n$ una muestra aleatoria de una distribución con media μ y varianza σ^2 finitas. Entonces*

$$\mathrm{E}(\bar X)=\mu,\qquad \mathrm{Var}(\bar X)=\frac{\sigma^2}{n},\qquad \mathrm{E}(s^2)=\sigma^2,$$

donde $\bar X=\frac{1}{n}\sum_i X_i$ y $s^2=\frac{1}{n-1}\sum_i (X_i-\bar X)^2$.

Demostración. La esperanza es lineal, luego $\mathrm{E}(\bar X)=\frac{1}{n}\sum \mathrm{E}(X_i)=\frac{1}{n}\cdot n\mu=\mu$. Para la varianza se usa la independencia, que permite sumar varianzas:

$$\mathrm{Var}(\bar X)=\frac{1}{n^2}\sum_i \mathrm{Var}(X_i)=\frac{1}{n^2}\cdot n\sigma^2=\frac{\sigma^2}{n}.$$

Para la tercera, obsérvese la identidad algebraica $\sum_i (X_i-\bar X)^2=\sum_i X_i^2-n\bar X^2$. Tomando esperanzas y usando $\mathrm{E}(Y^2)=\mathrm{Var}(Y)+\mathrm{E}(Y)^2$ en cada término,

$$\mathrm{E}\Big[\sum_i (X_i-\bar X)^2\Big]=n(\sigma^2+\mu^2)-n\Big(\frac{\sigma^2}{n}+\mu^2\Big)=n\sigma^2-\sigma^2=(n-1)\sigma^2.$$

Dividiendo entre $n-1$ queda $\mathrm{E}(s^2)=\sigma^2$. Aquí está, en una línea, la razón del divisor $n-1$: el σ^2 que se resta es la varianza de $\bar X$, es decir, el grado de libertad que se gasta al centrar los datos en su propia media en lugar de en μ. □

Definición 18.9 (Error típico de un estimador)**.** El *error típico* —o *error estándar*— de un estimador $\hat\theta$ es la desviación típica de su distribución en el muestreo, $\mathrm{ee}(\hat\theta)=\sqrt{\mathrm{Var}(\hat\theta)}$. Cuando depende de parámetros desconocidos se sustituyen por sus estimaciones y se habla de error típico *estimado*. Para la media muestral,

$$\mathrm{ee}(\bar X)=\frac{\sigma}{\sqrt{n}},\qquad \text{estimado por}\qquad \widehat{\mathrm{ee}}(\bar X)=\frac{s}{\sqrt{n}},$$

y para una proporción, $\mathrm{ee}(\hat p)=\sqrt{p(1-p)/n}$, estimado con $\hat p$ en lugar de p.

Cuidado

[La desviación típica y el error típico miden cosas distintas] s mide cuánto se dispersan *los datos* y no baja al aumentar n: si la población es heterogénea, seguirá siéndolo con un millón de observaciones. $s/\sqrt{n}$ mide cuánto se dispersa *la media muestral* de repetición en repetición, y sí baja, como $1/\sqrt{n}$. Informar s cuando se quería $s/\sqrt{n}$ exagera la incertidumbre; hacerlo al revés la esconde, que es peor. En las gráficas con barras de error hay que decir siempre cuál de las dos se ha dibujado, porque se diferencian en un factor $\sqrt{n}$ que puede ser de un orden de magnitud.

Regla: pasar de unos datos a una afirmación sobre la población

1. **Escribe la población** en una frase, con su unidad y su marco temporal. “Los votantes censados en mayo”, no “la gente”.
2. **Escribe el parámetro** que interesa, con su letra griega: μ, p, σ^2.
3. **Describe el mecanismo de muestreo** y comprueba que fue aleatorio. Si no lo fue, detente aquí: lo que sigue no se aplica.
4. **Elige el estadístico** $\hat{\theta}$ y calcúlalo con los datos.
5. **Identifica su distribución en el muestreo**, exacta si la hay o aproximada por el teorema central del límite si n es grande.
6. **Calcula el error típico** y decide si es aceptable; si no lo es, el remedio es más muestra, y crece con el cuadrado de la precisión exigida.
7. **Traduce a un intervalo**, no a un número solo. Un estimador puntual sin medida de incertidumbre es una opinión con decimales.

Con este vocabulario, las propiedades de la sección anterior de este capítulo se leen sin esfuerzo: el *sesgo* es la distancia entre el centro de la distribución en el muestreo y el parámetro; la *varianza* del estimador es la anchura de esa distribución; el *error cuadrático medio* los combina; la *consistencia* dice que esa distribución se concentra sobre θ al crecer n; y la *eficiencia* compara su anchura con la mínima posible. No hace falta repetir aquí ninguna de esas definiciones: lo único que faltaba era el objeto del que todas hablan.

Intervalos de confianza, y qué significa exactamente el nivel

Definición 18.10 (Intervalo de confianza para la media)**.** Con una muestra aleatoria de una población normal de media μ desconocida, el intervalo

$$\left[\ \bar{x} - t_{n-1;\,1-\alpha/2}\,\frac{s}{\sqrt{n}},\quad \bar{x} + t_{n-1;\,1-\alpha/2}\,\frac{s}{\sqrt{n}}\ \right]$$

se llama *intervalo de confianza* de nivel $1-\alpha$ para μ. La mitad de su anchura se llama *margen de error*. La definición general —cobertura, nivel nominal frente a nivel real— está en la definición 18.64, más adelante en el capítulo; aquí sólo se necesita la construcción concreta y su lectura correcta.

Ejemplo 18.11 (Diez mediciones, con su intervalo entero)**.** Se miden diez piezas y salen, en milímetros,

$$98,\ 104,\ 112,\ 95,\ 101,\ 107,\ 99,\ 110,\ 103,\ 96.$$

(a) Los estadísticos. La suma es 1025, luego $\bar{x} = 102{,}5$. Las desviaciones respecto de la media suman cero, como siempre, y sus cuadrados suman 302,5; por tanto $s^2 = \frac{302{,}5}{9} = 33{,}611111$ y $s = 5{,}797509$.

(b) El error típico. $\widehat{\mathrm{ee}}(\bar{X}) = \frac{5{,}797509}{\sqrt{10}} = 1{,}833333$. Nótese la diferencia con $s = 5{,}80$: los datos se dispersan casi seis milímetros, la media de diez de ellos apenas 1,8.

(c) El intervalo al 95 %. Con 9 grados de libertad, $t_{9;\,0{,}975} = 2{,}262157$, el margen es $2{,}262157 \cdot 1{,}833333 = 4{,}147288$ y el intervalo

$$[\,98{,}352712;\ \ 106{,}647288\,].$$

Si σ fuese conocida se usaría $z = 1{,}959964$ y el margen bajaría a 3,593: la t es más ancha porque paga el no conocer σ, y esa penalización desaparece al crecer n.

(d) Lectura. El valor 100 está dentro del intervalo, de modo que estos datos son compatibles con que la media verdadera sea 100; pero también lo son con 99, con 104 y con cualquier otro valor del intervalo. El intervalo no dice cuál es más probable: no es una distribución, es un conjunto.

Cuidado

[Qué significa y qué no significa un intervalo al 95 % de confianza] Ésta es, con diferencia, la frase peor entendida de toda la estadística, y conviene desmontarla despacio.

Lo que la gente cree que dice. "Hay una probabilidad de 0,95 de que μ esté entre 98,35 y 106,65." Esta afirmación es *falsa* en el marco frecuentista, y no por un tecnicismo. El parámetro μ es un número fijo, desconocido pero no aleatorio: la media verdadera de las piezas no cambia de valor según qué muestra saquemos. Los extremos 98,35 y 106,65 también son números fijos, ya calculados. Una afirmación sobre dos números fijos y un tercer número fijo es verdadera o falsa, sin término medio: o μ está ahí dentro, o no está. No hay ninguna probabilidad de 0,95 que repartir, del mismo modo que no tiene sentido decir que hay un 95 % de probabilidad de que 7 sea menor que 10.

Lo que sí dice. La probabilidad está en el *procedimiento*, antes de mirar los datos. Lo aleatorio son los extremos $\bar{X} \pm t\,s/\sqrt{n}$, que cambian con cada muestra. La afirmación correcta es: si se repitiera indefinidamente el proceso completo —extraer una muestra de diez piezas, calcular $\bar{x}$ y s, construir el intervalo—, el 95 % de los intervalos así construidos contendría al verdadero μ, y el 5 % restante no. La confianza es una propiedad de la fábrica de intervalos, no del intervalo que salió hoy.

La comprobación, contada. Se simula exactamente eso. Población normal con $\mu = 100$ y $\sigma = 15$; se extraen 100 muestras independientes de tamaño $n = 10$; con cada una se construye su intervalo al 95 %. Resultado del recuento: 94 **de los** 100 **intervalos contienen al valor** 100 **y** 6 **no lo contienen**. De los seis fallos, cuatro quedan enteros por debajo de 100 y dos enteros por encima, repartidos casi a partes iguales como corresponde a un intervalo simétrico. Los seis fallos son las muestras número 20, 32, 34, 39, 69 y 97; el peor es la 34, con $\bar{x} = 82{,}787$ e intervalo [69,425; 96,149], que no roza el 100 ni de lejos. Repitiendo el experimento con 200 000 muestras en lugar de 100, la proporción de aciertos es 0,949810, indistinguible del 0,95 prometido.

Tres lecciones del recuento. Primera: 94 de 100 no es 95 de 100, y no tiene por qué serlo; el número de aciertos es él mismo aleatorio, con distribución binomial de media 95 y desviación típica $\sqrt{100 \cdot 0{,}95 \cdot 0{,}05} = 2{,}18$, así que valores entre 91 y 99 son perfectamente normales. Segunda: los intervalos no miden todos lo mismo. En esta simulación la anchura media fue de 21,1 unidades, pero la más estrecha midió 9,07 y la más ancha 34,26, porque s también es aleatoria; un intervalo estrecho no es un intervalo mejor, es a menudo un intervalo desafortunado. Y tercera: mirando un intervalo concreto

es imposible saber si es de los 94 o de los 6. Ésa es exactamente la información que no se tiene y que la palabra "confianza" no proporciona.
El detalle que remata el argumento. Si en la misma simulación se sustituye el cuantil $t_{9;\,0{,}975} = 2{,}262$ por el normal 1,960 —el error de usar la normal con muestras pequeñas—, la cobertura real cae a 0,917950: se anuncia un 95 % y se entrega un 92 %. El nivel de confianza es una promesa que el método puede incumplir, y comprobarlo es tan sencillo como contar.

Contrastes: las piezas, con nombre

Definición 18.12 (Hipótesis, estadístico de contraste y región crítica)**.** Un *contraste de hipótesis* enfrenta dos afirmaciones sobre el parámetro: la *hipótesis nula* H_0, que es la que se supone cierta mientras no haya evidencia en contra, y la *hipótesis alternativa* H_1, que es lo que se quiere demostrar. Una hipótesis es *simple* si fija un único valor del parámetro y *compuesta* si abarca varios.

El *estadístico de contraste* es un estadístico T cuya distribución en el muestreo se conoce *bajo* H_0, y que toma valores tanto más extremos cuanto menos creíble sea H_0. La *región crítica* o *región de rechazo* R es el conjunto de valores de T para los que se decide rechazar H_0; su complementario es la región de aceptación. El *nivel de significación* es

$$\alpha = P(T \in R \;\mid\; H_0 \text{ cierta}),$$

y se fija *antes* de ver los datos. Finalmente, el *p-valor* es la probabilidad, calculada bajo H_0, de obtener un valor de T tan extremo o más que el observado t_{obs}:

$$p = P(|T| \geq |t_{\text{obs}}| \;\mid\; H_0) \quad \text{en un contraste bilateral.}$$

Rechazar cuando $p \leq \alpha$ es exactamente lo mismo que rechazar cuando $t_{\text{obs}} \in R$.

Los errores de tipo I y II, la potencia y el cálculo del tamaño muestral ya están definidos y discutidos más adelante en este mismo capítulo; no se repiten aquí. Lo único que faltaba por escribir eran las tres piezas anteriores, que se usan en todas partes sin haber sido nombradas.

Ejemplo 18.13 (El mismo conjunto de datos, ahora como contraste)**.** Con las diez piezas del ejemplo 18.11 se contrasta $H_0\colon \mu = 100$ frente a $H_1\colon \mu \neq 100$ con $\alpha = 0{,}05$.

(a) El estadístico. $t_{\text{obs}} = \dfrac{102{,}5 - 100}{1{,}833333} = 1{,}363636.$

(b) La región crítica. Bajo H_0, ese estadístico sigue una t_9, cuyo cuantil 0,975 vale 2,262157; la región crítica es $\{|t| > 2{,}262157\}$, que en unidades originales significa rechazar si $\bar{x} < 95{,}852712$ o $\bar{x} > 104{,}147288$.

(c) La decisión y el *p*-valor. Como $1{,}364 < 2{,}262$, no se rechaza. El p-valor es 0,205813: si la media verdadera fuese 100, en un 20,6 % de las muestras se vería una discrepancia igual o mayor que la observada. Nada llamativo.

(d) La coherencia con el intervalo. El intervalo del ejemplo anterior contenía a 100, y aquí no se rechaza $H_0\colon \mu = 100$. No es casualidad: el intervalo al 95 % es *exactamente* el

conjunto de valores μ_0 que no se rechazarían al 5 %. Intervalo y contraste son el mismo objeto visto de dos maneras, y por eso informar el intervalo es siempre más útil: contiene la decisión del contraste y además dice cuánto.

Cuidado

[Lo que un p-valor no es, con la cuenta que lo demuestra] La confusión estándar es leer $p = 0{,}05$ como "hay un 5 % de probabilidad de que H_0 sea cierta". Las dos cantidades son distintas y ni siquiera se parecen; puede comprobarse con un cálculo exacto de cinco líneas.

El montaje. Sea $X \sim N(\mu, 1)$ y una muestra de $n = 100$, de modo que $\bar{X} \sim N(\mu;\, 0{,}1^2)$. Se contrasta $H_0\colon \mu = 0$ frente a $H_1\colon \mu = 0{,}5$, y se observa $\bar{x} = 0{,}195996$, es decir, justo el valor que produce un p-valor bilateral de 0,05 exactamente.

La cuenta. La probabilidad de que H_0 sea cierta *dados los datos* sólo tiene sentido si se le asigna a H_0 una probabilidad a priori. Tómese la más neutra posible, $P(H_0) = P(H_1) = \frac{1}{2}$. Las densidades de $\bar{x}$ bajo cada hipótesis son proporcionales a $\exp\bigl(-\frac{1}{2}z^2\bigr)$ con z la distancia tipificada al centro correspondiente:

$$\text{bajo } H_0:\ z = \frac{0{,}195996}{0{,}1} = 1{,}959964, \qquad \text{bajo } H_1:\ z = \frac{0{,}195996 - 0{,}5}{0{,}1} = -3{,}040036.$$

El cociente de verosimilitudes es

$$\frac{L(H_1)}{L(H_0)} = \exp\Bigl(-\tfrac{1}{2}(3{,}040036)^2 + \tfrac{1}{2}(1{,}959964)^2\Bigr) = \mathrm{e}^{-2{,}700180} = 0{,}067193,$$

y por el teorema de Bayes, con priores iguales,

$$P(H_0 \mid \bar{x}) = \frac{1}{1 + 0{,}067193} = 0{,}937037.$$

El veredicto. El p-valor vale 0,05 y la probabilidad de que H_0 sea cierta vale 0,94. No es que difieran un poco: es que los datos, que "rechazan H_0 al 5 %", apoyan a H_0 frente a H_1 con un cociente de casi 15 a 1. La razón es sencilla: p mide sólo cuán raros son los datos bajo H_0, y no compara con nada; pero $\bar{x} = 0{,}196$ es todavía más raro bajo H_1, que predecía 0,5. Un contraste que no mira la alternativa no puede decir cuál de las dos hipótesis explica mejor lo visto.

Que no depende del ejemplo. Podría objetarse que la alternativa elegida es demasiado concreta. Existe una cota universal: para cualquier alternativa razonable, el cociente de verosimilitudes a favor de H_0 no puede bajar de $-\mathrm{e}\,p\ln p$. Con $p = 0{,}05$ eso da 0,407162, y con priores iguales $P(H_0 \mid \text{datos}) \geq 0{,}289350$: al menos un 29 %, seis veces el 5 % que se le atribuye. Con $p = 0{,}01$ la cota es 0,111254, todavía once veces mayor que 0,01. La distancia entre el p-valor y la probabilidad de la hipótesis no es un descuido corregible: es estructural.

Compárese con el contraejemplo por frecuencias de la advertencia sobre falsos positivos que aparece más adelante en el capítulo: allí se llega a la misma conclusión contando estudios, aquí calculando una probabilidad. Los dos argumentos son independientes y apuntan al mismo sitio.

Regla: informar un contraste sin decir tonterías

1. **Enuncia H_0 y H_1 antes de ver los datos**, y di si el contraste es uni o bilateral. Cambiarlo después de mirar invalida el α.
2. **Da el estadístico, sus grados de libertad y el p-valor exacto**, no "$p < 0{,}05$".
3. **Da el tamaño del efecto** en unidades del problema, y su intervalo de confianza. Es lo único que responde a la pregunta "¿cuánto?".
4. **No escribas nunca "se acepta H_0".** Escribe "no hay evidencia suficiente para rechazar H_0", que es lo que ocurrió.
5. **No traduzcas el p-valor a probabilidad de hipótesis** ni a probabilidad de haberte equivocado. Es $P(\text{datos} \mid H_0)$, no $P(H_0 \mid \text{datos})$.
6. **Declara cuántos contrastes hiciste** en total, incluidos los que no salieron.
7. **Si no rechazas, informa de la potencia planificada**, no de la calculada a posteriori con el efecto observado.

Pausa para pensar

Merece la pena mirar una vez más el 94 de la simulación. Si esos cien intervalos fuesen cien artículos científicos publicados, seis de ellos afirmarían con toda corrección metodológica algo falso sobre el valor de μ, y no habría manera de saber cuáles. El sistema funciona exactamente como se anunció, y aun así produce errores; lo que la estadística ofrece no es certeza sino una tasa de error conocida y controlada. Toda la discusión posterior sobre replicación, preregistro y p-hacking consiste en discutir si esa tasa es la que se cree, y casi nunca lo es, porque los seis fallos rara vez llegan a publicarse mientras que los noventa y cuatro aciertos —y algún falso positivo espectacular— sí.

18.3. Máxima verosimilitud

Definición 18.14. Dada una muestra $x_1, \dots, x_n$ de una distribución con densidad $f(x;\theta)$, la *función de verosimilitud* es $L(\theta) = \prod_i f(x_i;\theta)$, y el *estimador de máxima verosimilitud* es el $\hat\theta$ que la maximiza. En la práctica se maximiza $\ell(\theta) = \log L(\theta)$, que tiene el mismo máximo y convierte productos en sumas.

Idea clave

Qué dice esto en cristiano. La pregunta que responde el método es: de todos los valores posibles del parámetro, ¿cuál hace que los datos observados sean lo más probables posible? No es la probabilidad del parámetro dados los datos (eso sería bayesiano), sino al revés: la probabilidad de los datos dado el parámetro, vista como función del parámetro. La receta es mecánica (derivar el logaritmo e igualar a cero) y tiene propiedades asintóticas excelentes: es consistente, es asintóticamente normal y alcanza la cota de Cramér–Rao.

Es el método por defecto de casi toda la estadística aplicada.

Regla: calcular un estimador de máxima verosimilitud

El procedimiento es siempre el mismo:

1. **Escribe la densidad conjunta** como producto, aprovechando que la muestra es independiente.
2. **Toma logaritmos** para convertir el producto en suma; el máximo no se mueve porque el logaritmo es creciente.
3. **Deriva respecto del parámetro** e iguala a cero: esa es la *ecuación de verosimilitud*.
4. **Despeja** y comprueba con la segunda derivada que se trata de un máximo y no de un mínimo.
5. **Revisa la frontera del espacio de parámetros**, porque el máximo puede estar allí y entonces la derivada no se anula.
6. **Comprueba el sesgo**, que no está garantizado: el método optimiza verosimilitud, no insesgadez.
7. **Usa la invariancia** si necesitas otro parámetro: si $\hat{\theta}$ es el estimador de θ, entonces $g(\hat{\theta})$ lo es de $g(\theta)$, propiedad que la insesgadez no tiene.

Ejemplo 18.15 (Máxima verosimilitud en tres modelos)**. (a) Bernoulli.** Con k éxitos en n pruebas, $\ell(p) = k\log p + (n-k)\log(1-p)$, cuya derivada $\frac{k}{p} - \frac{n-k}{1-p}$ se anula en $\hat{p} = \frac{k}{n}$: la proporción muestral. Es insesgado y eficiente.

(b) Exponencial. Con densidad $\lambda \mathrm{e}^{-\lambda x}$, $\ell(\lambda) = n\log\lambda - \lambda\sum x_i$, y derivando sale $\hat{\lambda} = \frac{1}{\bar{x}}$. Aquí el estimador *es* sesgado, con $E(\hat{\lambda}) = \frac{n}{n-1}\lambda$, aunque el sesgo desaparece al crecer n.

(c) Normal. Salen $\hat{\mu} = \bar{x}$ y $\hat{\sigma}^2 = \frac{1}{n}\sum(x_i - \bar{x})^2$, es decir, el estimador *sesgado* de la varianza del apartado anterior. La máxima verosimilitud no busca la insesgadez y aquí se ve con claridad.

(d) El caso en que la derivada no sirve. Para una uniforme en $[0, \theta]$, la verosimilitud es θ^{-n} si $\theta \geq \max x_i$ y cero si no. Es decreciente donde está definida, así que el máximo está en la frontera: $\hat{\theta} = \max x_i$. Derivar habría dado una ecuación sin solución, y ese es el motivo del paso 5 de la regla.

Cuidado

[Verosimilitud no es probabilidad del parámetro] $L(\theta)$ no es una densidad en θ y no integra 1; decir "la probabilidad de que θ valga esto" es un error conceptual que solo tiene sentido en el marco bayesiano, donde hace falta además una distribución a priori. Segundo aviso: la ecuación de verosimilitud puede tener varias soluciones o ninguna en forma cerrada, y entonces hay que maximizar numéricamente, típicamente con Newton

o con el algoritmo EM. Y tercero: las buenas propiedades del método son *asintóticas*; con muestras pequeñas el estimador puede comportarse mal, y el apartado (d) da un ejemplo donde el sesgo es del orden de $\frac{\theta}{n+1}$, nada despreciable.

18.4. Intervalos de confianza

Teorema 18.16 (Intervalo para la media)**.** *Si $X_1, \dots, X_n$ son normales independientes con media μ y varianza desconocida, entonces $\frac{\bar{X}-\mu}{s/\sqrt{n}}$ sigue una distribución t de Student con $n-1$ grados de libertad, y el intervalo*

$$\left[\bar{x} - t_{n-1;\,1-\alpha/2}\,\frac{s}{\sqrt{n}},\quad \bar{x} + t_{n-1;\,1-\alpha/2}\,\frac{s}{\sqrt{n}}\right]$$

contiene a μ con probabilidad $1-\alpha$.

Idea clave

Qué dice esto en cristiano. Lo aleatorio es el intervalo, no el parámetro. La afirmación correcta es: si repitiéramos el muestreo muchas veces, el 95 % de los intervalos construidos así contendrían al verdadero μ. Una vez calculado un intervalo concreto, ese intervalo contiene o no contiene a μ, y ya no hay probabilidad de por medio. La confusión con "hay un 95 % de probabilidad de que μ esté aquí" es universal y es exactamente la interpretación bayesiana, que requiere otro aparato. La anchura del intervalo decrece como $\frac{1}{\sqrt{n}}$, y esa raíz es lo que hace tan cara la precisión.

Ejemplo 18.17 (Construir intervalos en tres situaciones)**. (a) Media con varianza desconocida.** Con $\bar{x} = 68{,}4$, $s = 5{,}2$ y $n = 25$, el valor crítico es $t_{24;\,0{,}975} = 2{,}063899$ y el margen $2{,}063899 \cdot \frac{5{,}2}{5} = 2{,}146455$. El intervalo al 95 % es $[66{,}253545;\ 70{,}546455]$. Si la varianza fuera conocida se usaría $z = 1{,}959964$ y el margen bajaría a 2,038363: la t es siempre algo más ancha, y ese exceso es el precio de estimar σ.

(b) Proporción. Con $\hat{p} = 0{,}52$ y $n = 1000$, el margen es $1{,}959964\sqrt{\frac{0{,}52\cdot 0{,}48}{1000}} = 0{,}030965$, y el intervalo $[0{,}489035;\ 0{,}550965]$. Como incluye a 0,5, esa encuesta *no* permite afirmar que la opción esté por delante, aunque el 52 % lo sugiera.

(c) El tamaño muestral necesario. Para un margen de 0,03 al 95 % con el caso peor $p = 0{,}5$, hace falta $n \geq \left(\frac{1{,}959964}{0{,}03}\right)^2 \cdot 0{,}25 = 1067{,}1$, es decir 1068 encuestados. Ese es el origen del famoso "margen de error de tres puntos" con muestras de mil personas.

(d) La raíz cuadrada manda. Para reducir el margen a la mitad hay que *cuadruplicar* la muestra: de 1068 a 4272. Duplicar la muestra solo mejora el margen en un factor $\sqrt{2} \approx 1{,}41$. Esa es la razón económica de que las encuestas casi nunca pasen de unos pocos miles de entrevistas.

Cuidado

[Tres malentendidos sobre los intervalos] Primero, el ya mencionado: la probabilidad se refiere al procedimiento, no al intervalo concreto. Segundo: el nivel de confianza no dice nada sobre *dónde* dentro del intervalo es más probable que esté μ; el intervalo no es una distribución. Y tercero: que dos intervalos al 95 % se solapen *no* implica que la diferencia entre las medias sea no significativa; el contraste correcto es sobre la diferencia y puede dar significativo aun con solapamiento moderado. Comparar intervalos a ojo es un atajo que falla en la dirección conservadora.

18.5. Contrastes de hipótesis

Definición 18.18. Un *contraste* enfrenta una hipótesis nula H_0 con una alternativa H_1. El *error de tipo I* es rechazar H_0 siendo cierta, y su probabilidad es el *nivel de significación* α; el *error de tipo II* es no rechazarla siendo falsa, con probabilidad β. La *potencia* es $1-\beta$. El *p-valor* es la probabilidad de obtener un estadístico tan extremo o más que el observado *suponiendo cierta* H_0.

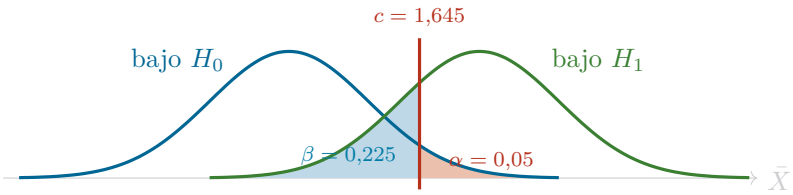


Figura 18.1: Los dos errores de un contraste. El área roja a la derecha del valor crítico es la probabilidad de rechazar H_0 siendo cierta; el área azul a su izquierda, la de no rechazarla siendo cierta H_1. Mover el corte reduce una y agranda la otra. Aquí la potencia vale $1-\beta = 0{,}775$.

Idea clave

Qué dice esto en cristiano. Un contraste es un juicio con presunción de inocencia: H_0 es la hipótesis por defecto y solo se abandona ante evidencia fuerte. Por eso los dos errores no son simétricos: α se fija de antemano y es la probabilidad de condenar a un inocente, que se mantiene baja a propósito. El p-valor mide cuán raros serían los datos si H_0 fuera cierta, y nada más: no es la probabilidad de que H_0 sea cierta, ni el tamaño del efecto, ni la importancia práctica del hallazgo. El lema de Neyman–Pearson dice cuál es el contraste de máxima potencia para un α dado, y resulta ser el basado en el cociente de verosimilitudes.

Ejemplo 18.19 (Un contraste completo, con su potencia)**. (a) El planteamiento.** Se quiere contrastar $H_0 : \mu = 70$ frente a $H_1 : \mu \neq 70$ con $\bar{x} = 68{,}4$, $s = 5{,}2$ y $n = 25$. El estadístico es $t = \frac{68{,}4-70}{5{,}2/5} = -1{,}538462$.

(b) La decisión. El valor crítico bilateral al 5 % es 2,063899, y $|t| = 1{,}538$ no lo supera; el p-valor es 0,137017. No se rechaza H_0. Obsérvese la formulación: *no se rechaza*, que no es lo mismo que *se acepta*.

(c) La potencia, que es lo que casi nadie calcula. Si la media real fuera 67, el parámetro de no centralidad es $\frac{70-67}{5{,}2/5} = 2{,}884615$ y la potencia del contraste vale 0,790339. Es decir, incluso con una diferencia de tres unidades, este diseño la detectaría solo en el 79 % de las repeticiones. Con muestras menores la potencia caería mucho más.

(d) Bondad de ajuste. Para comprobar si un dado es equilibrado con 60 tiradas y frecuencias observadas $8, 12, 9, 11, 14, 6$, el estadístico es $\chi^2 = \sum \frac{(o_i - e_i)^2}{e_i} = 4{,}2$ con 5 grados de libertad. El valor crítico al 5 % es 11,070498 y el p-valor 0,520995: los datos son perfectamente compatibles con un dado equilibrado.

Cuidado

[Lo que un p-valor no es] No es la probabilidad de que H_0 sea cierta: es una probabilidad calculada *suponiendo* que lo es. No es la probabilidad de haberse equivocado. No mide el tamaño del efecto: con una muestra suficientemente grande, cualquier diferencia por minúscula e irrelevante que sea acaba dando $p < 0{,}05$. Y un p grande no demuestra que H_0 sea cierta, solo que los datos no bastan para descartarla, cosa que puede deberse simplemente a poca potencia. Segundo aviso, y es el problema metodológico más serio de la ciencia actual: probar muchos contrastes y publicar solo el que sale significativo (lo que se llama p-hacking) garantiza falsos positivos, porque con veinte contrastes independientes y $\alpha = 0{,}05$ se espera uno significativo por puro azar.

18.6. Potencia, tamaño muestral y comparaciones múltiples

Un contraste de hipótesis controla el error de tipo I: la probabilidad de gritar "¡hallazgo!" cuando no hay nada. Pero eso es sólo la mitad del asunto. La otra mitad, la que decide si el estudio merece la pena hacerse, es la potencia: la probabilidad de detectar el efecto cuando sí existe. Y cuando en lugar de un contraste se hacen cincuenta, aparece un tercer problema que ha llegado a poner en duda literatura entera.

Definición 18.20. En un contraste con hipótesis nula H_0 y alternativa H_1, el *error de tipo I* tiene probabilidad $\alpha = P(\text{rechazar} \mid H_0)$ y el *error de tipo II* probabilidad $\beta = P(\text{no rechazar} \mid H_1)$. La *potencia* es $1 - \beta$. Para la media de una normal con σ conocida y contraste bilateral de nivel α, si la media verdadera es μ_1 y $\delta = \mu_1 - \mu_0$,

$$1 - \beta = \Phi\Big(-z_{\alpha/2} + \frac{\delta\sqrt{n}}{\sigma}\Big) + \Phi\Big(-z_{\alpha/2} - \frac{\delta\sqrt{n}}{\sigma}\Big),$$

y el tamaño muestral necesario para alcanzar una potencia dada es

$$n \;\geq\; \frac{(z_{\alpha/2} + z_{1-\beta})^2\,\sigma^2}{\delta^2}.$$

Idea clave

Qué dice esto en cristiano. Hay cuatro cantidades atadas entre sí: el nivel α, la potencia $1-\beta$, el tamaño del efecto que se quiere detectar δ, y el número de datos n. Fijadas tres, la cuarta queda determinada. La fórmula del tamaño muestral dice lo esencial: n crece con el *cuadrado* inverso del efecto, de modo que detectar un efecto la mitad de grande cuesta cuatro veces más datos. Ésa es la razón económica de que los estudios sobre efectos pequeños sean carísimos, y la razón estadística de que los estudios pequeños sobre efectos pequeños no sirvan para nada: no es que "no encuentren", es que no podían encontrar.

Regla: Diseñar un estudio con potencia suficiente

1. **Decidir el efecto mínimo que importa** δ, en unidades del problema y *antes* de ver los datos. Ésta es una decisión científica, no estadística.
2. **Estimar** σ de estudios previos, de un piloto o de la peor cota razonable.
3. **Fijar** α **y la potencia deseada**, habitualmente 0,05 y 0,80.
4. **Aplicar la fórmula** $n \geq (z_{\alpha/2} + z_{1-\beta})^2\sigma^2/\delta^2$ y redondear siempre hacia arriba.
5. **Añadir margen** por pérdidas de seguimiento y por el uso de t en lugar de z cuando σ se estima.
6. **No calcular la potencia después** con el efecto observado: la potencia retrospectiva es función determinista del p-valor y no aporta ninguna información nueva.
7. **Informar del intervalo de confianza**, no sólo del p-valor: un intervalo estrecho alrededor de cero es evidencia de ausencia de efecto; uno ancho es evidencia de ignorancia, y los dos producen el mismo "no significativo".

Ejemplo 18.21 (Cuánta gente hace falta para detectar cinco puntos de cociente intelectual)**.** Se contrasta $H_0\colon \mu = 100$ frente a $H_1\colon \mu \neq 100$ con $\sigma = 15$ y $\alpha = 0{,}05$, y el efecto real es $\mu_1 = 105$.

(a) Con $n = 36$**.** El error típico es $15/6 = 2{,}5$, el valor crítico es $z_{0{,}025} = 1{,}959964$ y se rechaza cuando $|\bar{x} - 100| > 4{,}899910$. Como el efecto es $\delta = 5$, apenas por encima de ese umbral, la potencia sale 0,516005: el estudio tiene poco más de una probabilidad entre dos de detectar un efecto que sí existe. Es tirar una moneda al aire.

(b) Cómo crece con n**.** Para $n = 16, 36, 64, 100, 144$ la potencia vale 0,265946; 0,516005; 0,760126; 0,915181; 0,979327. Obsérvese el rendimiento decreciente: pasar de 16 a 36 gana veinticinco puntos de potencia, pasar de 100 a 144 gana sólo seis.

(c) El tamaño necesario. Para potencia 0,80 hace falta $n \geq (1{,}959964 + 0{,}841621)^2 \cdot 225/25 = 70{,}639918$, es decir 71 personas; para potencia 0,90, $n \geq 94{,}566808$, es decir 95. Subir del 80 % al 90 % de potencia cuesta un tercio más de muestra.

(d) El tamaño del efecto. La d de Cohen es $\delta/\sigma = 5/15 = 0{,}333333$, lo que en la escala convencional es un efecto pequeño-mediano. La fórmula se reescribe como $n \geq (z_{\alpha/2}+z_{1-\beta})^2/d^2$,

lo que deja claro que sólo importa el efecto *relativo a la variabilidad*: cinco puntos son mucho si $\sigma = 2$ y nada si $\sigma = 50$.

Definición 18.22. Al realizar m contrastes simultáneos, la *tasa de error por familia* es la probabilidad de cometer al menos un error de tipo I. La corrección de *Bonferroni* consiste en usar el nivel α/m en cada contraste. El procedimiento de *Benjamini–Hochberg* ordena los p-valores $p_{(1)} \le \cdots \le p_{(m)}$, busca el mayor k con $p_{(k)} \le \frac{k}{m} q$ y rechaza las k hipótesis más pequeñas, controlando así la *tasa de falsos descubrimientos* en q.

Idea clave

Qué dice esto en cristiano. Si se hacen veinte contrastes independientes sin efecto real, la probabilidad de que alguno salga significativo al 5 % no es del 5 %: es del 64 %. Buscar en muchos sitios garantiza encontrar algo. Bonferroni lo arregla por la vía dura, exigiendo a cada contraste un nivel m veces más estricto, lo que controla el error pero destruye la potencia cuando m es grande. Benjamini–Hochberg cambia el objetivo: en vez de asegurar que *ningún* hallazgo sea falso, asegura que la *proporción* de falsos entre los declarados no pase de q. Es mucho menos conservador y es lo razonable cuando se exploran miles de hipótesis, como en genómica, donde nadie espera acertar en todas.

Ejemplo 18.23 (Diez p-valores, dos correcciones)**.** Sean

$$p = (0{,}001;\ 0{,}008;\ 0{,}012;\ 0{,}03;\ 0{,}04;\ 0{,}06;\ 0{,}2;\ 0{,}35;\ 0{,}6;\ 0{,}9),$$

con $m = 10$ y $\alpha = q = 0{,}05$.

(a) Sin corregir. Cinco p-valores bajan de 0,05: cinco "hallazgos".

(b) Bonferroni. El umbral pasa a $0{,}05/10 = 0{,}005$ y sólo sobrevive el primero. Se pasa de cinco descubrimientos a uno.

(c) Benjamini–Hochberg. Los umbrales $\frac{k}{10} \cdot 0{,}05$ valen 0,005; 0,01; 0,015; 0,02; 0,025; 0,03; 0,035; 0,04; 0,045; 0,05. Comparando término a término, se cumple $p_{(k)} \le \frac{k}{10} q$ para $k = 1, 2, 3$ y falla a partir de ahí; el mayor k válido es 3, así que se rechazan las tres primeras. Un resultado intermedio entre los otros dos, y el más sensato de los tres en un contexto exploratorio.

(d) Cómo escala el problema. La probabilidad de al menos un falso positivo con m contrastes independientes es $1 - 0{,}95^m$: vale 0,226219 con $m = 5$, 0,401263 con $m = 10$, 0,641514 con $m = 20$ y 0,994079 con $m = 100$. Con cien contrastes es prácticamente seguro encontrar algo aunque no haya absolutamente nada.

Cuidado

[Contraejemplo: el p-valor no es la probabilidad de que H_0 sea cierta] Supóngase que la mitad de las hipótesis que se contrastan son ciertas, con $\alpha = 0{,}05$ y potencia 0,80. De cada mil contrastes salen $0{,}05 \cdot 500 = 25$ falsos positivos y $0{,}80 \cdot 500 = 400$ verdaderos, de modo que la proporción de falsos entre los declarados significativos es $25/425 = 0{,}058824$: razonable. Pero si sólo el 10 % de las hipótesis fuese cierta, salen $0{,}05 \cdot 900 = 45$ falsos

frente a $0{,}80 \cdot 100 = 80$ verdaderos, y la proporción sube a $45/125 = 0{,}36$: más de un tercio de los "descubrimientos" con $p < 0{,}05$ serían falsos. El p-valor no cambió; lo que cambió fue la proporción de hipótesis ciertas en el pozo del que se pesca, y ése es un dato que el contraste no ve.

Cuidado

[Contraejemplo: la potencia retrospectiva no informa de nada] Terminado un estudio no significativo, es habitual calcular la potencia usando el efecto *observado* y concluir "teníamos poca potencia, por eso no salió". El razonamiento es circular: con esa definición, la potencia observada es una función decreciente del p-valor, de manera que todo estudio no significativo tendrá siempre potencia baja, por construcción y sin excepciones. No es evidencia de nada. Lo que sí informa es el intervalo de confianza: si el intervalo para la diferencia es $[-0{,}3;\ 0{,}4]$ y el efecto mínimo relevante era 2, el estudio descartó el efecto; si el intervalo es $[-15;\ 18]$, el estudio no descartó nada y simplemente no sabemos.

Pausa para pensar

En 2005 John Ioannidis publicó un artículo titulado "Por qué la mayoría de los hallazgos publicados son falsos" cuyo argumento es exactamente la aritmética del primer contraejemplo, agravada por dos factores: los investigadores prueban muchas hipótesis y publican la que sale, y las revistas prefieren los resultados positivos. El resultado es que el conjunto de hallazgos publicados está sesgado hacia el ruido afortunado. Las respuestas de la comunidad han sido el preregistro de los análisis antes de recoger los datos, la publicación de resultados nulos y la exigencia de replicación. Ninguna es estadística: son cambios en el procedimiento social. Los teoremas de esta sección eran correctos desde 1933; lo que fallaba era cómo se usaban.

18.7. Errores típicos de inferencia

Cuidado

[Ocho confusiones que invalidan un análisis]

1. **Leer el p-valor como probabilidad de la hipótesis**, en vez de como probabilidad de los datos bajo la hipótesis.
2. **Confundir significación estadística con relevancia práctica**: con n enorme todo sale significativo.
3. **Concluir que H_0 es cierta** al no rechazarla, sin haber calculado la potencia.
4. **Ignorar la potencia al diseñar**, con lo que un resultado negativo no significa nada.

5. **Elegir la hipótesis después de ver los datos**, o hacer muchos contrastes y publicar el mejor.
6. **Interpretar el intervalo de confianza como una distribución** del parámetro.
7. **Aplicar la t sin normalidad ni muestra grande**, sobre todo con muestras pequeñas y asimétricas.
8. **Olvidar que las observaciones deben ser independientes**: con datos agrupados o temporales, las fórmulas de este capítulo subestiman gravemente la incertidumbre.

Pausa para pensar

La distribución t la publicó en 1908 William Gosset bajo el seudónimo *Student*, porque trabajaba en la cervecería Guinness de Dublín y la empresa prohibía a sus empleados publicar. El problema que le llevó a inventarla era muy concreto: controlar la calidad de la cebada con muestras pequeñas, donde la aproximación normal fallaba. Fisher desarrolló después la teoría de la estimación y el análisis de la varianza, y Neyman y Pearson formalizaron los contrastes con sus dos tipos de error en los años treinta, en una disputa con Fisher que duró décadas y que sigue sin zanjarse del todo: el marco que hoy se enseña es una mezcla incómoda de las dos escuelas, y buena parte de la confusión con los p-valores viene de ahí.

18.8. Máxima verosimilitud y la información de Fisher

Detrás de casi todos los estimadores que aparecen en este capítulo hay un único principio, y no es "tómese la media". Es este: elíjase el valor del parámetro que hace más probable lo que se ha observado. De ahí salen la media, la proporción muestral y la recta de mínimos cuadrados, y además sale automáticamente cuánto vale su precisión.

Definición 18.24. Dada una muestra $x_1, \dots, x_n$ de una densidad $f(x \mid \theta)$, la *verosimilitud* es $L(\theta) = \prod_i f(x_i \mid \theta)$ y el *estimador de máxima verosimilitud* es $\hat{\theta} = \arg\max_\theta L(\theta)$, que en la práctica se busca maximizando $\ell(\theta) = \log L(\theta)$. La *información de Fisher* es

$$I(\theta) = -E\Big[\frac{\partial^2 \ell}{\partial \theta^2}\Big],$$

y bajo condiciones de regularidad, $\hat{\theta}$ es asintóticamente normal con media θ y varianza $1/I(\theta)$.

Idea clave

Qué dice esto en cristiano. La verosimilitud no es una probabilidad sobre θ: es la probabilidad de los datos, vista como función del parámetro. Maximizarla es preguntarse qué mundo explica mejor lo que se ha visto. La información de Fisher mide cuán picuda es esa función: si la log-verosimilitud tiene una curvatura pronunciada en el máximo, valores vecinos del parámetro explican los datos mucho peor y la estimación es precisa;

si es una meseta, muchos valores explican los datos casi igual de bien y la estimación es mala. Por eso la varianza sale como el inverso de la curvatura, y por eso el término "información" es literal: mide cuánto dicen los datos sobre el parámetro.

Regla: Estimar por máxima verosimilitud en siete pasos

1. **Escribir la verosimilitud** como producto sobre las observaciones, suponiendo independencia.
2. **Tomar logaritmos**, que convierte el producto en suma y no cambia dónde está el máximo.
3. **Derivar e igualar a cero**, obteniendo la *ecuación de verosimilitud*.
4. **Comprobar que es un máximo** con la segunda derivada, y no un mínimo o un punto de inflexión.
5. **Revisar la frontera**: a veces el máximo está en un extremo del dominio y la derivada nunca se anula.
6. **Calcular la información de Fisher** para obtener el error típico asintótico.
7. **Construir el intervalo** como $\hat{\theta} \pm z_{\alpha/2}/\sqrt{I(\hat{\theta})}$, o mejor, invirtiendo el contraste de razón de verosimilitudes.

Ejemplo 18.25 (Máxima verosimilitud en cuatro modelos)**. (a) Poisson.** Con $\ell = \sum(x_i \log \lambda - \lambda) + c$, derivar da $\hat{\lambda} = \bar{x}$. Para la muestra $3, 5, 4, 6, 2$ sale $\hat{\lambda} = 4$. La información es $I(\lambda) = n/\lambda = 1{,}25$, luego el error típico es $\sqrt{\lambda/n} = 0{,}894$ y el intervalo al 95 por ciento es $(2{,}25;\ 5{,}75)$, muy ancho porque cinco datos son pocos.

(b) Normal. Sale $\hat{\mu} = \bar{x}$, lo esperable, y $\hat{\sigma}^2 = \frac{1}{n}\sum(x_i - \bar{x})^2$, dividiendo por n y no por $n-1$. Ese estimador está sesgado: su esperanza es $\frac{n-1}{n}\sigma^2$, un 10 por ciento por debajo con $n = 10$. La máxima verosimilitud no garantiza insesgadez, sólo buenas propiedades asintóticas.

(c) Uniforme, donde el método se rompe. Para $U(0, \theta)$, la verosimilitud vale θ^{-n} si $\theta \geq \max x_i$ y cero si no. No se deriva: crece hasta el máximo muestral y allí salta a cero. El estimador es $\hat{\theta} = \max x_i$, siempre menor que el verdadero, y su distribución no es normal ni su varianza es $1/I$. Las condiciones de regularidad fallan porque el soporte depende del parámetro.

(d) La cota de Cramér–Rao. Ningún estimador insesgado puede tener varianza menor que $1/I(\theta)$. Cuando un estimador la alcanza se llama eficiente, y la media muestral lo es para la normal. La cota es el motivo por el que la máxima verosimilitud no se puede mejorar asintóticamente: alcanza el límite teórico de precisión.

Cuidado

[Contraejemplo: máxima verosimilitud puede ser inconsistente] Se enseña que el estimador de máxima verosimilitud converge al valor verdadero, y hay contraejemplos famosos.

El de Neyman y Scott, de 1948: obsérvense n parejas, cada una con su propia media μ_i y una varianza común σ^2. Al crecer n crece también el número de parámetros, y el estimador de σ^2 converge a $\sigma^2/2$, la mitad del valor correcto, por muchos datos que se recojan. La consistencia exige que el número de parámetros esté fijo; con parámetros incidentales que se multiplican con la muestra, el método falla de forma sistemática y silenciosa.

Pausa para pensar

Fisher introdujo la verosimilitud en 1912, siendo estudiante, y pasó décadas insistiendo en que no era una probabilidad y en que confundirlas era el error central de la estadística bayesiana de su tiempo. La discusión con Jeffreys sobre ese punto duró años y fue notoriamente áspera. Hoy la verosimilitud es el objeto común de ambas escuelas: el frecuentista la maximiza y el bayesiano la multiplica por una priori, pero la escriben igual.

18.9. Suficiencia, familias exponenciales y estimadores de varianza mínima

La cota de Cramér–Rao dice cuán buena *puede* llegar a ser una estimación insesgada. Falta la maquinaria que dice cómo alcanzarla, y esa maquinaria se apoya en una idea sola: casi todos los datos son irrelevantes.

Definición 18.26 (Estadístico suficiente)**.** $T(X_1, \dots, X_n)$ es **suficiente** para θ si la distribución condicionada de la muestra dado $T = t$ no depende de θ. Dicho de otro modo: una vez conocido T, los datos originales no contienen ni una gota más de información sobre el parámetro.

Teorema 18.27 (Criterio de factorización, Fisher–Neyman)**.** *T es suficiente si y solo si la verosimilitud se factoriza como*

$$L(\theta; x) = g(T(x), \theta)\, h(x),$$

donde h no depende de θ.

El criterio convierte la búsqueda en un ejercicio de mirar la fórmula.

Ejemplo 18.28 (Tres suficiencias leídas de un vistazo)**. (a) Bernoulli.** $L = p^{\sum x_i}(1-p)^{n-\sum x_i}$: todo depende de los datos a través de $\sum x_i$. Suficiente. Saber *cuántos* éxitos hubo basta; en qué orden salieron es información desechable.

(b) Normal con σ conocida. $L \propto \exp(-\sum (x_i - \mu)^2/2\sigma^2)$, y desarrollando el cuadrado el único término con μ es $\mu \sum x_i$. Suficiente: $\sum x_i$, o equivalentemente $\bar{X}$.

(c) Uniforme $U(0, \theta)$. $L = \theta^{-n}\mathbf{1}\{\max x_i \le \theta\}$. Suficiente: $\max x_i$. Aquí la suficiencia es espectacular: de n números, $n-1$ son basura.

Definición 18.29 (Familia exponencial)**.** Una familia de densidades es **exponencial** de k parámetros si se escribe

$$f(x;\theta) = h(x)\exp\Big(\sum_{j=1}^{k}\eta_j(\theta)T_j(x) - A(\theta)\Big).$$

Entonces $(\sum_i T_1(x_i),\dots,\sum_i T_k(x_i))$ es suficiente, y además *completo*: la única función suya con esperanza idénticamente nula es la nula.

Bernoulli, binomial, Poisson, normal, gamma, beta y exponencial son todas exponenciales, y su estadístico suficiente natural es siempre una suma. La uniforme no lo es (su soporte depende del parámetro), y por eso su suficiente es un máximo y no una suma.

Teorema 18.30 (Rao–Blackwell)**.** *Si $\hat\theta$ es insesgado y T suficiente, entonces $\hat\theta^* = E[\hat\theta \mid T]$ es insesgado y*

$$\operatorname{Var}(\hat\theta^*) \le \operatorname{Var}(\hat\theta),$$

con igualdad solo si $\hat\theta$ ya era función de T.

Teorema 18.31 (Lehmann–Scheffé)**.** *Si además T es completo, ese $\hat\theta^*$ es* el *estimador insesgado de varianza uniformemente mínima: no hay otro mejor, y es único.*

Ejemplo 18.32 (Rao-blackwellizar un estimador burdo, con las varianzas al lado)**.** Sea $X_1,\dots,X_n \sim U(0,\theta)$.

(a) El estimador ingenuo. Como $E[X]=\theta/2$, el estimador $\hat\theta_1 = 2\bar X$ es insesgado. Su varianza:

$$\operatorname{Var}(2\bar X) = \frac{4}{n}\cdot\frac{\theta^2}{12} = \frac{\theta^2}{3n}.$$

(b) El estimador rao-blackwellizado. El suficiente es $M = \max X_i$, con $E[M] = \frac{n}{n+1}\theta$, luego $\hat\theta_2 = \frac{n+1}{n}M$ es insesgado. Su varianza, tras un cálculo directo con la densidad nm^{n-1}/θ^n:

$$\operatorname{Var}\Big(\tfrac{n+1}{n}M\Big) = \frac{\theta^2}{n(n+2)}.$$

(c) La comparación, con $n=10$.

$$\operatorname{Var}(\hat\theta_1) = \frac{\theta^2}{30} = 0{,}033333\,\theta^2, \qquad \operatorname{Var}(\hat\theta_2) = \frac{\theta^2}{120} = 0{,}008333\,\theta^2.$$

El segundo tiene exactamente la cuarta parte de varianza. Y no es cuestión de suerte ni de ingenio: por Lehmann–Scheffé, $\hat\theta_2$ es el mejor estimador insesgado que existe, porque M es suficiente y completo.

(d) El detalle que sorprende. El cociente de varianzas es $(n+2)/3$, que *crece* con n: cuantos más datos, mayor la ventaja relativa. La razón es que $\hat\theta_2$ converge a velocidad $1/n$ y $\hat\theta_1$ solo a $1/\sqrt{n}$. La uniforme es uno de los pocos modelos donde el estimador óptimo es superconsistente, y por eso la cota de Cramér–Rao no se le aplica: sus hipótesis de regularidad exigen soporte independiente del parámetro.

Idea clave

El programa completo de la estimación clásica cabe en tres pasos: encuentra un suficiente completo (si la familia es exponencial, es la suma que aparece en el exponente); encuentra cualquier estimador insesgado, por burdo que sea; condiciónalo al suficiente. Lo que sale es óptimo, y no hay que probar nada más.

18.10. Los tres contrastes asintóticos: razón de verosimilitudes, Wald y puntuación

Con la verosimilitud ya construida, hay tres maneras naturales de preguntarle si θ_0 es creíble, y las tres miden lo mismo desde tres sitios distintos de la curva.

Definición 18.33 (Los tres estadísticos)**.** Sea $\hat\theta$ el estimador de máxima verosimilitud, $\ell = \log L$ la log-verosimilitud, $U = \ell'$ la puntuación e I la información de Fisher.

$$\underbrace{\Lambda = -2\log\frac{\sup_{H_0} L}{\sup L}}_{\text{razón de verosimilitudes}}, \qquad \underbrace{W = (\hat\theta-\theta_0)^2 I(\hat\theta)}_{\text{Wald}}, \qquad \underbrace{S = \frac{U(\theta_0)^2}{I(\theta_0)}}_{\text{puntuación}}.$$

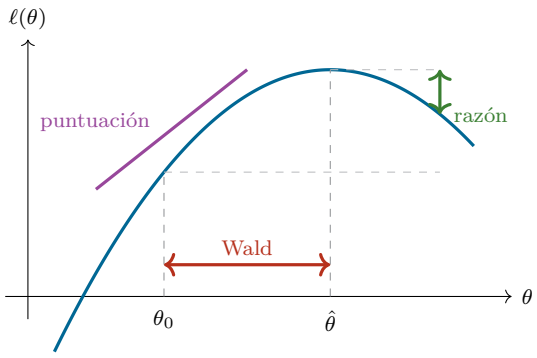


Figura 18.2: Los tres contrastes miden la misma distancia desde tres sitios distintos de la log-verosimilitud: Wald en horizontal ($\hat\theta-\theta_0$), la razón de verosimilitudes en vertical (cuánta verosimilitud se pierde) y la puntuación por la pendiente en θ_0. Coinciden asintóticamente y divergen cerca de la frontera.

Teorema 18.34 (Wilks)**.** *Bajo H_0 y condiciones de regularidad, los tres estadísticos convergen en distribución a una χ^2_r, donde r es el número de restricciones que impone H_0. Además son asintóticamente equivalentes entre sí.*

La interpretación geométrica es lo que hace que se recuerden: los tres miden la distancia entre θ_0 y $\hat\theta$, pero *Wald* la mide sobre el eje horizontal, *la razón de verosimilitudes* la mide en vertical (cuánta verosimilitud se pierde al obligarse a θ_0) y *la puntuación* la mide por la pendiente en θ_0 (si la curva todavía sube mucho ahí, es que θ_0 no es el máximo).

Ejemplo 18.35 (Sesenta caras de cien, con los tres estadísticos)**.** Contrastamos $H_0\colon p = 0{,}5$ con $k = 60$ éxitos en $n = 100$, luego $\hat{p} = 0{,}6$.

(a) Wald. $I(\hat{p}) = n/(\hat{p}(1-\hat{p}))$, así que

$$W = \frac{(0{,}6 - 0{,}5)^2}{0{,}6 \cdot 0{,}4/100} = \frac{0{,}01}{0{,}0024} = 4{,}166667.$$

(b) Puntuación. Idéntico pero evaluando la información en θ_0:

$$S = \frac{(0{,}6 - 0{,}5)^2}{0{,}5 \cdot 0{,}5/100} = \frac{0{,}01}{0{,}0025} = 4.$$

Este es exactamente el contraste z para una proporción de toda la vida, elevado al cuadrado: $z = 2$, $z^2 = 4$.

(c) Razón de verosimilitudes.

$$\Lambda = 2\Big[60 \log \frac{0{,}6}{0{,}5} + 40 \log \frac{0{,}4}{0{,}5}\Big] = 2\big[10{,}939\,3 - 8{,}925\,7\big] = 4{,}027\,1.$$

(d) Veredicto. Los tres rondan el 4, y el p-valor de χ^2_1 es 0,0448 para Λ: se rechaza por poco al 5 por ciento. Que los tres coincidan es la situación normal cuando n es grande y $\hat{\theta}$ no está en la frontera.

Cuidado

No siempre coinciden, y cuando divergen hay uno preferible. Con $k = 1$ éxito en $n = 10$, Wald da $W = (0{,}1-0{,}5)^2/(0{,}1\cdot 0{,}9/10) = 17{,}8$, la puntuación da $S = (0{,}4)^2/(0{,}025) = 6{,}4$ y la razón de verosimilitudes da $\Lambda = 7{,}36$: Wald más que dobla a los otros dos y rechazaría con una contundencia falsa. El motivo es estructural: Wald estima la curvatura en $\hat{\theta}$, y cerca del borde del espacio paramétrico esa estimación se desmadra.

Peor aún, Wald *no es invariante frente a reparametrizaciones*: contrastar $p = 0{,}5$ y contrastar $\log \frac{p}{1-p} = 0$ son la misma hipótesis, y Wald da números distintos según cuál se escriba. La razón de verosimilitudes da el mismo, siempre, porque el cociente de máximos no sabe en qué coordenadas está escrito el parámetro. Por eso, cuando los tres no coinciden, hay que fiarse de Λ.

Idea clave

Cada uno tiene su nicho de comodidad. La puntuación solo necesita ajustar el modelo *bajo* H_0, y por eso es la elección cuando el modelo alternativo es caro de ajustar (es el origen del contraste de Breusch–Pagan y del contraste de multiplicadores de Lagrange en econometría). Wald solo necesita el ajuste *sin restricciones*, y por eso es lo que imprime cualquier programa junto a cada coeficiente. La razón de verosimilitudes necesita los dos ajustes, y es la más fiable: es la que se usa para comparar modelos anidados.

18.11. Comparar dos grupos: independientes, apareados y la corrección de Welch

Regla: el árbol de decisión con dos muestras

1. ¿Cada dato del primer grupo está *emparejado* con uno del segundo (mismo individuo antes y después, gemelos, dos medidas del mismo objeto)? Si sí, calcula las diferencias y aplica una t de una muestra sobre ellas. Fin.
2. Si son independientes: ¿las varianzas son parecidas? Si sí, t con varianza combinada, $n_1 + n_2 - 2$ grados de libertad.
3. Si no lo son o no lo sabes, t de Welch. En la práctica, empieza siempre aquí.

Definición 18.36 (Los tres estadísticos)**.** Con varianza común, se combina $s_p^2 = \dfrac{(n_1-1)s_1^2 + (n_2-1)s_2^2}{n_1+n_2-2}$ y

$$t = \frac{\bar{x}_1 - \bar{x}_2}{s_p\sqrt{1/n_1 + 1/n_2}} \sim t_{n_1+n_2-2}.$$

Sin suponerla, la corrección de Welch usa

$$t = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{s_1^2/n_1 + s_2^2/n_2}}, \qquad \nu = \frac{(s_1^2/n_1 + s_2^2/n_2)^2}{\dfrac{(s_1^2/n_1)^2}{n_1-1} + \dfrac{(s_2^2/n_2)^2}{n_2-1}},$$

con grados de libertad fraccionarios. Y con datos apareados, sobre las diferencias d_i:

$$t = \frac{\bar{d}}{s_d/\sqrt{n}} \sim t_{n-1}.$$

Ejemplo 18.37 (Los mismos ocho pacientes, analizados bien y analizados mal)**.** Peso en kilos antes y después de ocho semanas de dieta:

antes	95	82	110	74	88	101	79	93
después	93	80	107	73	85	99	77	91

(a) Análisis correcto (apareado). Las diferencias son $2, 2, 3, 1, 3, 2, 2, 2$, con $\bar{d} = 2{,}125$ y $s_d = 0{,}640870$. Entonces

$$t = \frac{2{,}125}{0{,}640870/\sqrt{8}} = 9{,}378\,5, \qquad 7 \text{ grados de libertad}, \qquad p = 0{,}000033.$$

Aplastantemente significativo. El intervalo al 95 por ciento para la pérdida media es $2{,}125 \pm 2{,}3646 \cdot 0{,}2266 = [1{,}589;\ 2{,}661]$ kilos.

(b) Análisis incorrecto (como si fueran independientes). Las medias son 90,25 y 88,125, con desviaciones típicas 11,949 y 11,557. Entonces

$$t = 0{,}361\,6, \qquad p = 0{,}723.$$

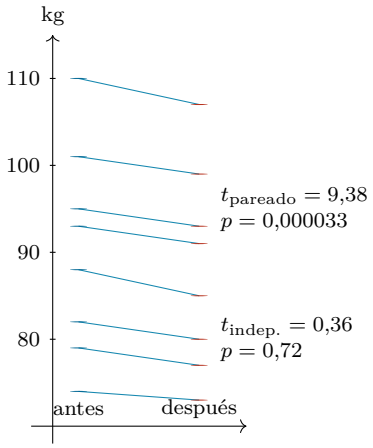


Figura 18.3: Los mismos ocho pacientes, antes y después. Todas las líneas bajan, y bajan lo mismo: el efecto es evidente. Pero la variabilidad *entre* pacientes (de 74 a 110 kilos) es dieciocho veces mayor que la de dentro, y por eso el análisis que ignora el emparejamiento no detecta absolutamente nada.

Nada. Cero evidencia.

(c) Qué ha pasado. La variabilidad *entre* pacientes es enorme (de 74 a 110 kilos, $s \approx 12$), y esa variabilidad ahoga por completo un efecto de dos kilos. El análisis apareado la elimina de un plumazo: cada paciente es su propio control, y lo único que queda en el denominador es la variabilidad *dentro* del paciente, que es de 0,64. Al dividir por un error típico dieciocho veces menor, el mismo efecto pasa de invisible a incontestable.

Ignorar el emparejamiento no es un tecnicismo: es tirar a la basura la mitad de la información del diseño.

Cuidado

El error inverso también existe y es peor: tratar como apareados datos que no lo están, o emparejar *después* de ver los resultados. El emparejamiento tiene que estar en el diseño, decidido antes de mirar. Y ojo con el contraste F de igualdad de varianzas como paso previo: es extremadamente sensible a la falta de normalidad, tanto que suele fallar más a menudo que el problema que pretende detectar. La recomendación moderna es saltárselo y usar Welch siempre: cuando las varianzas sí son iguales, Welch pierde muy poca potencia, y cuando no lo son, salva el contraste.

18.12. El contraste óptimo y la construcción general de intervalos

Teorema 18.38 (Lema de Neyman–Pearson)**.** *Para contrastar dos hipótesis* simples $H_0\colon \theta = \theta_0$ *frente a* $H_1\colon \theta = \theta_1$*, la región crítica más potente de nivel* α *es*

$$R = \left\{x : \frac{L(\theta_1; x)}{L(\theta_0; x)} > k\right\},$$

con k *elegido para que* $P_{\theta_0}(R) = \alpha$.

Demostración. Sea R^* cualquier otra región con $P_{\theta_0}(R^*) \le \alpha$. En $R \setminus R^*$ se cumple $L_1 > kL_0$ y en $R^* \setminus R$ se cumple $L_1 \le kL_0$. Integrando,

$$P_{\theta_1}(R) - P_{\theta_1}(R^*) = \int_{R\setminus R^*} L_1 - \int_{R^*\setminus R} L_1 \ge k\Big(\int_{R\setminus R^*} L_0 - \int_{R^*\setminus R} L_0\Big) = k(\alpha - P_{\theta_0}(R^*)) \ge 0. \ \square$$

Diez líneas, y de ahí sale toda la teoría de contrastes óptimos.

Ejemplo 18.39 (La región óptima resulta ser la de siempre)**.** Normal con $\sigma = 2$ conocida, $n = 9$, $H_0\colon \mu = 0$ frente a $H_1\colon \mu = 1$, $\alpha = 0{,}05$.

El cociente de verosimilitudes es $\exp\big(\frac{n}{\sigma^2}(\mu_1 - \mu_0)\bar{x} - \text{cte}\big)$, creciente en $\bar{x}$: la región óptima es $\{\bar{x} > c\}$, exactamente el contraste unilateral habitual. Con $c = z_{0{,}95}\sigma/\sqrt{n} = 1{,}645 \cdot 2/3 = 1{,}096\,6$, la potencia es

$$1 - \Phi\Big(\frac{1{,}0966 - 1}{2/3}\Big) = 1 - \Phi(0{,}145) = 0{,}442\,4.$$

Es decir: aun usando el mejor contraste posible, con estos datos se detecta el efecto menos de la mitad de las veces. Para alcanzar potencia 0,9 harían falta

$$n \ge \Big(\frac{(z_{0{,}95} + z_{0{,}90})\sigma}{\mu_1 - \mu_0}\Big)^2 = \big((1{,}645 + 1{,}282)\cdot 2\big)^2 = 34{,}26,$$

o sea 35 observaciones. Ese cálculo, hecho *antes* de recoger los datos, es la única defensa real contra un estudio condenado de antemano.

Observación 18.40. Neyman–Pearson pide hipótesis simples, pero como la región no depende del valor concreto de μ_1 (solo de que sea mayor que μ_0), el mismo contraste es óptimo para *toda* la alternativa $\mu > 0$ a la vez: es uniformemente más potente. El teorema de Karlin–Rubin generaliza el argumento a cualquier familia con razón de verosimilitudes monótona. Para alternativas bilaterales no existe ningún UMP —lo que es óptimo por la derecha es pésimo por la izquierda—, y hay que conformarse con la clase más pequeña de los contrastes insesgados.

Intervalos: el método general y por qué el de siempre está mal

Regla: construir un intervalo en dos pasos

1. *Por pivote.* Encuentra una cantidad $Q(X, \theta)$ cuya distribución no dependa de θ; despeja θ de $P(a \le Q \le b) = 1 - \alpha$. Así salen $\bar{x} \pm t\, s/\sqrt{n}$ y el intervalo para σ^2 a partir de $(n-1)s^2/\sigma^2 \sim \chi^2_{n-1}$, que es asimétrico porque la χ^2 lo es.
2. *Por inversión.* El conjunto de los θ_0 que *no* se rechazan al nivel α es un intervalo de confianza $1 - \alpha$. Siempre funciona, aunque no haya pivote.

Ejemplo 18.41 (Tres intervalos para una proporción, y uno de ellos es indefendible)**. (a) Caso benigno: $\hat{p} = 0{,}52$ con $n = 1000$.**

$$\text{Wald: } [0{,}4890;\ 0{,}5510], \qquad \text{Wilson: } [0{,}4890;\ 0{,}5508].$$

Idénticos hasta la tercera cifra. Con muestras grandes y p lejos de los bordes, da igual cuál se use.

(b) Caso letal: $x = 0$ éxitos de $n = 20$. El intervalo de Wald es

$$\hat{p} \pm 1{,}96\sqrt{\hat{p}(1-\hat{p})/n} = 0 \pm 0 = [0,\ 0].$$

Afirma con un 95 por ciento de confianza que la probabilidad de éxito es exactamente cero, tras veinte intentos. Es absurdo, y no es un caso raro de laboratorio: ocurre siempre que no se observa ningún éxito.

(c) La reparación. El intervalo de Wilson se obtiene por inversión —resolviendo $|\hat{p} - p| = z\sqrt{p(1-p)/n}$ en p en lugar de en $\hat{p}$— y da

$$[0;\ 0{,}161\,1],$$

que es una afirmación razonable: con 0 de 20, la probabilidad podría perfectamente ser del 10 por ciento. El de Agresti–Coull, que consiste en sumar dos éxitos y dos fracasos ficticios y aplicar Wald, da $[-0{,}027;\ 0{,}194]$, que truncado en 0 es igual de aceptable y se calcula mentalmente.

Cuidado

La cobertura real del intervalo de Wald para proporciones oscila muy por debajo del 95 por ciento nominal, y no mejora suavemente con n: salta arriba y abajo según dónde caigan los valores discretos posibles de $\hat{p}$. Para p cerca de 0 o de 1 puede quedarse por debajo del 80 por ciento con muestras de varios cientos. Wilson y Agresti–Coull no tienen ese problema. El de Wald sigue enseñándose por inercia; conviene saber que hay consenso desde hace décadas en que no debería usarse.

Ejemplo 18.42 (Método delta: cómo transformar un intervalo)**.** Si $\hat{\theta}$ es asintóticamente normal, entonces $g(\hat{\theta})$ también lo es, con

$$\operatorname{Var}(g(\hat{\theta})) \approx g'(\theta)^2 \operatorname{Var}(\hat{\theta}).$$

Con $\hat{p} = 0{,}6$ y $n = 100$, el error típico de $\hat{p}$ es 0,048 99. Para los momios $g(p) = p/(1-p)$ se tiene $g'(p) = 1/(1-p)^2 = 6{,}25$, luego

$$\widehat{\mathrm{OR}} = 1{,}5, \qquad \mathrm{ee} = 6{,}25 \cdot 0{,}048\,99 = 0{,}306\,2, \qquad \mathrm{IC} = [0{,}900;\ 2{,}100].$$

La alternativa —transformar los extremos del intervalo de $\hat{p}$, que es $[0{,}504;\ 0{,}696]$— da $[1{,}016;\ 2{,}290]$. Los dos son válidos asintóticamente y no coinciden: el segundo respeta que los momios son positivos y hereda la asimetría de la transformación, mientras que el primero es simétrico por construcción. Cuando la transformación es muy no lineal, el segundo es preferible; y en la práctica se trabaja en la escala donde la normalidad es mejor —aquí el logaritmo de los momios— y se destransforma al final.

18.13. Las distribuciones del muestreo: χ^2, t y F, y el teorema de Fisher

Todos los intervalos y contrastes del capítulo usan tres distribuciones y un teorema que aún no se han definido ni demostrado. Sin ellos, las recetas funcionan pero no se entienden.

Definición 18.43 (Chi cuadrado)**.** Si $Z_1, \dots, Z_k$ son normales estándar independientes, la variable $Q = Z_1^2 + \cdots + Z_k^2$ tiene la distribución *chi cuadrado con k grados de libertad*, $Q \sim \chi^2_k$, de densidad

$$f_Q(q) = \frac{1}{2^{k/2}\Gamma(k/2)}\, q^{k/2-1} \mathrm{e}^{-q/2}, \qquad q > 0.$$

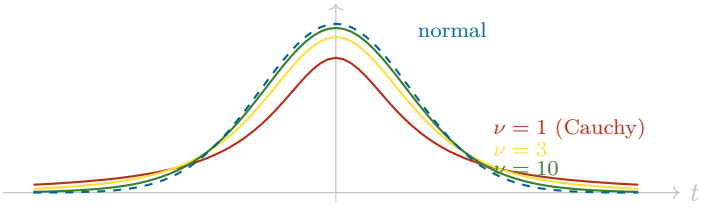


Figura 18.4: La densidad de la t_ν para tres valores de ν, junto a la normal estándar. Dividir por una desviación típica *estimada* —que a veces sale pequeña por azar— engorda las colas, y por eso el cuantil al 97,5 % pasa de 1,960 a 12,706 cuando $\nu = 1$. Con $\nu = 10$ la diferencia ya es pequeña, y desaparece al crecer ν.

Proposición 18.44 (Propiedades de la χ^2)**.** $\chi^2_k = \mathit{Gamma}(k/2, 1/2)$*; su generatriz es* $M(t) = (1-2t)^{-k/2}$ *para* $t < \frac{1}{2}$*; tiene media* k *y varianza* $2k$*; y es* aditiva*: si* $Q_1 \sim \chi^2_{k_1}$ *y* $Q_2 \sim \chi^2_{k_2}$ *son independientes,* $Q_1 + Q_2 \sim \chi^2_{k_1+k_2}$.

Demostración. Para $k = 1$, la densidad de Z^2 se calculó en el teorema del cambio de variable del capítulo anterior y es la de Gamma$(\frac{1}{2}, \frac{1}{2})$, con generatriz $\left(\frac{1/2}{1/2-t}\right)^{1/2} = (1-2t)^{-1/2}$. Por independencia, la generatriz de la suma de k copias es $(1-2t)^{-k/2}$, que es la de Gamma$(\frac{k}{2}, \frac{1}{2})$; la unicidad de la generatriz cierra el argumento. La media y la varianza salen de derivar: $M'(0) = k$, $M''(0) = k(k+2)$, luego Var $= k(k+2) - k^2 = 2k$. La aditividad es el producto de generatrices. □

Definición 18.45 (t de Student y F de Fisher)**.** Si $Z \sim N(0,1)$ y $Q \sim \chi^2_\nu$ son *independientes*, entonces

$$T = \frac{Z}{\sqrt{Q/\nu}} \;\sim\; t_\nu, \qquad \text{de densidad} \qquad f_T(t) = \frac{\Gamma\left(\frac{\nu+1}{2}\right)}{\sqrt{\nu\pi}\,\Gamma\left(\frac{\nu}{2}\right)} \left(1 + \frac{t^2}{\nu}\right)^{-\frac{\nu+1}{2}}.$$

Si $Q_1 \sim \chi^2_m$ y $Q_2 \sim \chi^2_n$ son independientes, $F = \dfrac{Q_1/m}{Q_2/n} \sim F_{m,n}$.

Observación 18.46 (Por qué la t tiene colas gruesas). La t_ν es una normal cuya escala está a su vez sujeta a azar: al dividir por $\sqrt{Q/\nu}$, que a veces sale pequeño, se producen valores

grandes con más frecuencia de la que una normal admitiría. Cuantitativamente: $\mathrm{E}[T^2] = \frac{\nu}{\nu-2}$ para $\nu > 2$, y la varianza es infinita para $\nu \le 2$. El caso $\nu = 1$ es exactamente la distribución de Cauchy, que no tiene ni media. En el otro extremo, cuando $\nu \to \infty$ la ley de los grandes números hace que $Q/\nu \to 1$, y $t_\nu \to N(0,1)$: por eso las tablas de la t terminan en ∞ con la fila de la normal. Los valores críticos al 95 % ilustran la transición: 12,706 para $\nu = 1$, 2,776 para $\nu = 4$, 2,228 para $\nu = 10$, 2,042 para $\nu = 30$ y 1,960 para la normal.

Teorema 18.47 (Fisher)**.** *Sean $X_1, \dots, X_n$ independientes $N(\mu, \sigma^2)$, y sean $\bar X = \frac{1}{n}\sum X_i$ y $s^2 = \frac{1}{n-1}\sum (X_i - \bar X)^2$. Entonces:*

1. $\bar X \sim N\!\left(\mu, \frac{\sigma^2}{n}\right)$;
2. $\dfrac{(n-1)s^2}{\sigma^2} \sim \chi^2_{n-1}$;
3. $\bar X$ *y* s^2 *son* independientes;
4. *en consecuencia,* $\dfrac{\bar X - \mu}{s/\sqrt{n}} \sim t_{n-1}$.

Demostración. Sin pérdida de generalidad, $\mu = 0$ y $\sigma = 1$ (basta tipificar). Sea H una matriz *ortogonal* $n \times n$ cuya primera fila sea $(\frac{1}{\sqrt n}, \dots, \frac{1}{\sqrt n})$ —existe: se completa ese vector unitario a una base ortonormal por Gram–Schmidt (capítulo 7); la elección clásica es la matriz de Helmert—. Póngase $Y = HX$.

Paso 1: Y también es normal estándar. El vector X tiene densidad conjunta $(2\pi)^{-n/2}\mathrm{e}^{-\|x\|^2/2}$, que sólo depende de la norma. Al aplicar una transformación ortogonal, la norma no cambia y el jacobiano vale $|\det H| = 1$, de modo que Y tiene exactamente la misma densidad: sus componentes son n normales estándar independientes. (Esta es la propiedad que hace única a la gaussiana: es la única distribución con componentes independientes e invariante por rotaciones.)

Paso 2: identificar las piezas. Por construcción, $Y_1 = \frac{1}{\sqrt n}\sum_i X_i = \sqrt{n}\,\bar X$, lo que da (1). Y como H es ortogonal conserva la norma:

$$\sum_{i=1}^n Y_i^2 = \sum_{i=1}^n X_i^2 \quad\Longrightarrow\quad \sum_{i=2}^n Y_i^2 = \sum_i X_i^2 - Y_1^2 = \sum_i X_i^2 - n\bar X^2 = \sum_i (X_i - \bar X)^2 = (n-1)s^2.$$

Paso 3: conclusión. El miembro derecho es suma de $n-1$ normales estándar independientes $Y_2, \dots, Y_n$, luego $(n-1)s^2 \sim \chi^2_{n-1}$: eso es (2). Y como $\bar X$ es función sólo de Y_1 mientras que s^2 es función sólo de $Y_2, \dots, Y_n$, y todas las Y_i son independientes, $\bar X$ y s^2 son independientes: eso es (3). Finalmente,

$$\frac{\bar X - \mu}{s/\sqrt n} = \frac{(\bar X - \mu)/(\sigma/\sqrt n)}{\sqrt{\frac{(n-1)s^2/\sigma^2}{n-1}}} = \frac{Z}{\sqrt{Q/(n-1)}}$$

con $Z \sim N(0,1)$, $Q \sim \chi^2_{n-1}$ e independientes: es exactamente la definición de t_{n-1}. □

Ejemplo 18.48 (La independencia, comprobada)**.** El punto (3) es el que más sorprende: la media y la dispersión de la misma muestra no se informan mutuamente. Con 200 000

muestras simuladas de tamaño $n = 8$ de una $N(0,1)$, la correlación entre $\bar{X}$ y s^2 sale +0,002 —compatible con cero—, la media de $(n-1)s^2$ vale 7,002 frente al valor teórico 7, su varianza 13,88 frente a 14, y el contraste de Kolmogórov–Smirnov contra la χ^2_7 da un estadístico de 0,0027. Conviene subrayar que la independencia *sólo* vale para la normal: es, de hecho, una caracterización de la normalidad (teorema de Geary), y por eso el aparato de la t de Student es tan frágil ante desviaciones fuertes de la normalidad en muestras pequeñas.

18.14. Suficiencia, Rao–Blackwell y Lehmann–Scheffé, demostrados

Los tres resultados que gobiernan la estimación óptima se enunciaron como recetas. Se demuestran aquí, y las demostraciones son cortas.

Definición 18.49 (Estadístico suficiente)**.** Un estadístico $T = T(X_1, \dots, X_n)$ es *suficiente* para θ si la distribución condicionada de la muestra dado $T = t$ no depende de θ.

Traducido: una vez conocido T, los datos originales no contienen ninguna información adicional sobre θ. Se puede tirar la muestra y quedarse con el resumen.

Teorema 18.50 (Criterio de factorización, Fisher–Neyman)**.** *T es suficiente para θ si y sólo si la verosimilitud se factoriza como*

$$L(\theta; x) = g\big(T(x), \theta\big) \cdot h(x),$$

donde h no depende de θ y g depende de los datos sólo a través de T.

Demostración en el caso discreto. ($\Rightarrow$) Si T es suficiente,

$$P_\theta(X = x) = P_\theta\big(T = T(x)\big) \cdot P\big(X = x \mid T = T(x)\big),$$

y el segundo factor no depende de θ por hipótesis: basta tomar $g = P_\theta(T = T(x))$ y $h = P(X = x \mid T = T(x))$.

($\Leftarrow$) Supóngase la factorización. Para t con $P_\theta(T = t) > 0$,

$$P_\theta(X = x \mid T = t) = \frac{P_\theta(X = x)}{P_\theta(T = t)} = \frac{g(t,\theta)h(x)}{\sum_{y:T(y)=t} g(t,\theta)h(y)} = \frac{h(x)}{\sum_{y:T(y)=t} h(y)}$$

para x con $T(x) = t$ (y 0 en otro caso), expresión en la que $g(t,\theta)$ se ha cancelado: no depende de θ. □

Ejemplo 18.51 (Suficiencia en tres modelos)**.**

- *Bernoulli(p):* $L = p^{\sum x_i}(1-p)^{n-\sum x_i}$, que depende de los datos sólo a través de $T = \sum x_i$: el número de éxitos es suficiente, y el orden en que salieron es irrelevante.
- *Normal(μ, σ^2) con ambos desconocidos:* $L \propto \sigma^{-n} \exp\big(-\frac{1}{2\sigma^2}(\sum x_i^2 - 2\mu \sum x_i + n\mu^2)\big)$, luego $T = (\sum x_i, \sum x_i^2)$ —equivalentemente $(\bar{X}, s^2)$— es suficiente. Dos números resumen cualquier muestra normal, por grande que sea.

- *Uniforme*$(0,\theta)$*:* $L = \theta^{-n}\mathbf{1}_{\{\max x_i \le \theta\}}$, y el indicador depende de los datos sólo a través de $T = \max x_i$: el máximo es suficiente. Nótese que aquí la media muestral *no* lo es, lo cual desconcierta a quien espera que la media sirva siempre.

Teorema 18.52 (Rao–Blackwell)**.** *Sea $\hat\theta$ un estimador insesgado de θ con varianza finita y T un estadístico suficiente. Defínase $\tilde\theta = \mathrm{E}[\hat\theta \mid T]$. Entonces $\tilde\theta$ es un estadístico legítimo (no depende de θ), es insesgado, y*

$$\mathrm{Var}(\tilde\theta) \;\le\; \mathrm{Var}(\hat\theta),$$

con igualdad si y sólo si $\hat\theta$ ya era función de T.

Demostración. Que $\tilde\theta$ no dependa de θ es exactamente la definición de suficiencia: la distribución condicionada dado T no involucra a θ, luego tampoco la esperanza condicionada. Es insesgado por la ley de la esperanza total: $\mathrm{E}[\tilde\theta] = \mathrm{E}\big[\mathrm{E}[\hat\theta \mid T]\big] = \mathrm{E}[\hat\theta] = \theta$. Para la varianza, la descomposición

$$\mathrm{Var}(\hat\theta) = \mathrm{E}\big[\mathrm{Var}(\hat\theta \mid T)\big] + \mathrm{Var}\big(\mathrm{E}[\hat\theta \mid T]\big) = \mathrm{E}\big[\mathrm{Var}(\hat\theta \mid T)\big] + \mathrm{Var}(\tilde\theta)$$

tiene el primer sumando ≥ 0, y vale 0 exactamente cuando $\hat\theta$ es constante dado T, es decir cuando es función de T. □

Idea clave

Qué dice esto en cristiano. Promediar un estimador sobre todas las muestras que comparten el mismo resumen suficiente nunca empeora, y casi siempre mejora. La receta es constructiva: si se tiene un estimador cualquiera —por tonto que sea— y un estadístico suficiente, condicionar produce uno al menos igual de bueno. El ejemplo canónico: para estimar $\mathrm{e}^{-\lambda} = P(X=0)$ en un modelo de Poisson, el estimador "tonto" $\mathbf{1}_{\{X_1=0\}}$ es insesgado; condicionándolo sobre el suficiente $T = \sum X_i$ se obtiene $\left(1-\frac{1}{n}\right)^T$, insesgado y con varianza muchísimo menor.

Definición 18.53 (Familia completa)**.** La familia de distribuciones de T es *completa* si la única función g con $\mathrm{E}_\theta[g(T)] = 0$ para todo θ es $g = 0$ casi seguro.

Teorema 18.54 (Lehmann–Scheffé)**.** *Si T es suficiente y completo y $\tilde\theta = g(T)$ es insesgado para θ, entonces $\tilde\theta$ es el* único *estimador insesgado de varianza mínima (UMVU).*

Demostración. Sea $\hat\theta$ cualquier otro estimador insesgado. Por Rao–Blackwell, $\mathrm{E}[\hat\theta \mid T]$ es insesgado, función de T, y con varianza $\le \mathrm{Var}(\hat\theta)$. Ahora bien, la diferencia $d(T) = \mathrm{E}[\hat\theta \mid T] - g(T)$ es función de T con $\mathrm{E}_\theta[d(T)] = \theta - \theta = 0$ para todo θ; por completitud, $d = 0$ c.s. Luego $\mathrm{E}[\hat\theta \mid T] = g(T) = \tilde\theta$, y por tanto $\mathrm{Var}(\tilde\theta) \le \mathrm{Var}(\hat\theta)$. La unicidad sale del mismo argumento aplicado a dos estimadores UMVU. □

Ejemplo 18.55 (Por qué s^2 lleva $n-1$ y no n)**.** En el modelo normal, $(\bar X, s^2)$ es suficiente y completo, y ambos son insesgados para μ y σ^2. Por Lehmann–Scheffé son los UMVU, y en particular s^2 con denominador $n-1$ es el *único* estimador insesgado de varianza mínima de

σ^2. Con denominador n el estimador tiene sesgo $-\sigma^2/n$ —y de hecho menor error cuadrático medio, lo cual muestra que "insesgado" y "bueno" no son sinónimos—, pero deja de ser insesgado, y toda la teoría de la t y de la F está construida sobre el insesgado.

18.15. Información de Fisher y la cota de Cramér–Rao

Se ha afirmado que ningún estimador insesgado puede bajar de cierta varianza, y que la máxima verosimilitud alcanza esa cota. Aquí está el teorema.

Definición 18.56 (Score e información de Fisher)**.** Sea $\ell(\theta; x) = \log L(\theta; x)$. El *score* es $U(\theta) = \frac{\partial \ell}{\partial \theta}$, y la *información de Fisher* es

$$I(\theta) = \mathrm{E}_\theta\big[U(\theta)^2\big] = -\,\mathrm{E}_\theta\left[\frac{\partial^2 \ell}{\partial \theta^2}\right].$$

Proposición 18.57 (El score tiene media cero, y las dos fórmulas coinciden)**.** *Bajo condiciones de regularidad —el soporte de la densidad no depende de θ, y se puede derivar bajo el signo integral—:*

$$\mathrm{E}_\theta[U(\theta)] = 0 \qquad y \qquad \mathrm{E}_\theta\big[U^2\big] = -\,\mathrm{E}_\theta\big[U'\big].$$

Demostración. Derívese la identidad $\int f(x;\theta)\mathrm{d}x = 1$ respecto de θ e intercámbiense derivada e integral:

$$0 = \int \frac{\partial f}{\partial \theta}\mathrm{d}x = \int \frac{\partial \log f}{\partial \theta}\, f\,\mathrm{d}x = \mathrm{E}_\theta[U].$$

Derivando otra vez,

$$0 = \int \left(\frac{\partial^2 \log f}{\partial \theta^2} f + \frac{\partial \log f}{\partial \theta}\cdot\frac{\partial f}{\partial \theta}\right)\mathrm{d}x = \mathrm{E}[U'] + \mathrm{E}[U^2],$$

donde se ha usado $\frac{\partial f}{\partial \theta} = U f$. □

Teorema 18.58 (Cota de Cramér–Rao)**.** *Bajo las condiciones de regularidad anteriores, todo estimador insesgado $\hat\theta$ de θ satisface*

$$\mathrm{Var}(\hat\theta) \;\geq\; \frac{1}{I_n(\theta)} = \frac{1}{n\, I_1(\theta)},$$

donde I_n es la información de la muestra completa. Más generalmente, si $\mathrm{E}[\hat\theta] = g(\theta)$, la cota es $g'(\theta)^2/I_n(\theta)$.

Demostración. Derívese la identidad de insesgadez $\int \hat\theta(x) f(x;\theta)\mathrm{d}x = \theta$ respecto de θ:

$$1 = \int \hat\theta\, \frac{\partial f}{\partial \theta}\mathrm{d}x = \int \hat\theta\, U\, f\,\mathrm{d}x = \mathrm{E}\big[\hat\theta\, U\big].$$

Como $\mathrm{E}[U] = 0$, esto dice que $\mathrm{Cov}(\hat\theta, U) = 1$. Aplíquese ahora la desigualdad de Cauchy–Schwarz a la covarianza:

$$1 = \mathrm{Cov}(\hat\theta, U)^2 \;\leq\; \mathrm{Var}(\hat\theta)\cdot\mathrm{Var}(U) = \mathrm{Var}(\hat\theta)\cdot I_n(\theta),$$

y despejando se obtiene la cota. La aditividad $I_n = nI_1$ para muestras independientes sale de que el score de la muestra es la suma de los scores individuales, que son independientes y de media cero, luego sus varianzas se suman. El caso general se obtiene derivando $\mathrm{E}[\hat\theta] = g(\theta)$ en lugar de $= \theta$. □

Observación 18.59 (Cuándo se alcanza). La igualdad en Cauchy–Schwarz exige que $\hat\theta - \theta$ y U sean proporcionales, es decir

$$U(\theta) = k(\theta)(\hat\theta - \theta)$$

para alguna función k. Integrando, esto ocurre exactamente cuando la densidad pertenece a la *familia exponencial* $f(x;\theta) = h(x)\exp(\eta(\theta)T(x) - A(\theta))$, y en tal caso el estimador eficiente es función del estadístico suficiente natural T. Fuera de la familia exponencial la cota generalmente no se alcanza para n finito, aunque sí asintóticamente.

Ejemplo 18.60 (Poisson: la cota se alcanza)**.** Para $X_i \sim \mathrm{Poisson}(\lambda)$, $\ell = -n\lambda + (\sum x_i)\log\lambda - \sum \log(x_i!)$, luego $U = -n + \frac{\sum x_i}{\lambda}$ y $U' = -\frac{\sum x_i}{\lambda^2}$, de donde $I_n(\lambda) = \frac{n\lambda}{\lambda^2} = \frac{n}{\lambda}$. La cota es λ/n, y en efecto $\mathrm{Var}(\bar X) = \lambda/n$: el estimador $\bar X$ es eficiente. Con $\lambda = 3$ y $n = 20$: cota $= 0{,}15$, varianza real $= 0{,}15$. Obsérvese además que $U = \frac{n}{\lambda}(\bar X - \lambda)$, exactamente la forma de proporcionalidad de la observación anterior.

Cuidado

[La uniforme rompe la cota, y no es un error] Para $X_i \sim U(0,\theta)$, el estimador $\hat\theta = \frac{n+1}{n}\max X_i$ es insesgado y tiene varianza $\frac{\theta^2}{n(n+2)}$, que es de orden n^{-2}: *mucho menor* que cualquier cota de orden n^{-1}. No hay contradicción, porque el soporte $[0,\theta]$ depende de θ y falla la condición de regularidad que permite derivar bajo el signo integral. Es el ejemplo que hay que tener presente: la cota de Cramér–Rao no es una ley de la naturaleza, sino un teorema con hipótesis, y los modelos con soporte variable las incumplen sistemáticamente. En esos modelos la convergencia puede ser a velocidad n^{-1} en lugar de $n^{-1/2}$, y la teoría asintótica estándar no se aplica.

18.16. Máxima verosimilitud: propiedades asintóticas y el teorema de Wilks

Teorema 18.61 (Consistencia y normalidad asintótica del EMV)**.** *Bajo condiciones de regularidad —identificabilidad, soporte independiente de θ, θ_0 interior al espacio de parámetros, y tres derivadas de ℓ con dominación local— el estimador de máxima verosimilitud $\hat\theta_n$ cumple*

$$\hat\theta_n \xrightarrow{P} \theta_0 \qquad y \qquad \sqrt{n}\,(\hat\theta_n - \theta_0) \xrightarrow{d} N\!\left(0,\ \frac{1}{I_1(\theta_0)}\right).$$

Idea clave

La demostración de la normalidad, en cuatro pasos. Sea $U_n(\theta) = \sum_i \frac{\partial}{\partial\theta}\log f(X_i;\theta)$ el score de la muestra.

1. Por definición, $U_n(\hat\theta_n) = 0$ (el EMV anula la derivada de la log-verosimilitud).
2. Taylor alrededor de θ_0: $0 = U_n(\theta_0) + U_n'(\bar\theta)(\hat\theta_n - \theta_0)$ para algún $\bar\theta$ entre ambos.
3. Despejando y dividiendo por $\sqrt{n}$:
$$\sqrt{n}(\hat\theta_n - \theta_0) = \frac{U_n(\theta_0)/\sqrt{n}}{-U_n'(\bar\theta)/n}.$$
4. El numerador es una suma de n variables independientes de media 0 y varianza $I_1(\theta_0)$, dividida por $\sqrt{n}$: por el teorema central del límite converge a $N(0, I_1(\theta_0))$. El denominador es una media de n variables independientes: por la ley de los grandes números converge a $-\mathrm{E}[\ell''] = I_1(\theta_0)$. El teorema de Slutsky combina ambos y da $N(0, I_1)/I_1 = N(0, 1/I_1)$.

La conclusión operativa es que el EMV es *asintóticamente eficiente*: alcanza la cota de Cramér–Rao en el límite, aunque no la alcance para n finito.

Teorema 18.62 (Wilks)**.** *Sea $\Theta_0 \subset \Theta$ definido por r restricciones, y sea*
$$\Lambda = \frac{\sup_{\theta\in\Theta_0} L(\theta)}{\sup_{\theta\in\Theta} L(\theta)}.$$
Entonces, bajo H_0 y las condiciones de regularidad,
$$-2\log\Lambda \xrightarrow{d} \chi^2_r.$$

Idea clave

De dónde sale el χ^2. Desarróllese ℓ por Taylor de segundo orden alrededor del máximo global $\hat\theta$:
$$\ell(\theta) \approx \ell(\hat\theta) - \tfrac{1}{2}(\theta - \hat\theta)^\top I_n(\hat\theta)(\theta - \hat\theta).$$
Entonces $-2\log\Lambda = 2[\ell(\hat\theta) - \ell(\hat\theta_0)] \approx (\hat\theta - \hat\theta_0)^\top I_n(\hat\theta - \hat\theta_0)$, que es una forma cuadrática en un vector asintóticamente normal, con la matriz de precisión correcta. Una forma cuadrática $Z^\top Z$ con Z normal estándar de dimensión r es precisamente una χ^2_r, y r es el número de direcciones que la hipótesis nula fija: los grados de libertad son la diferencia de dimensiones entre los dos modelos.

Observación 18.63 (Los tres contrastes asintóticos son el mismo teorema). Wald, score (Rao) y razón de verosimilitudes miden lo mismo de tres maneras: Wald mide la distancia horizontal $\hat\theta - \theta_0$ ponderada por la curvatura; el score mide la pendiente $U(\theta_0)$ en el punto nulo; y la razón de verosimilitudes mide la caída vertical de ℓ. Sobre una parábola exacta los tres coinciden numéricamente; en general sólo coinciden asintóticamente, y para n moderado pueden discrepar. Reglas prácticas: el de razón de verosimilitudes suele ser el más fiable y es invariante frente a reparametrizaciones; el de Wald es el más cómodo (sale de la salida estándar de cualquier programa) pero *no* es invariante, y puede fallar espectacularmente cerca

del borde del espacio de parámetros —el llamado efecto Hauck–Donner—; el score es el único que no requiere ajustar el modelo completo, y por eso es el que se usa cuando el modelo alternativo es caro de estimar.

18.17. Cobertura, y el método delta

Definición 18.64 (Intervalo de confianza, nivel y cobertura)**.** Un *intervalo de confianza* de nivel $1-\alpha$ para un parámetro θ es un par de estadísticos $L(X) \le U(X)$ —funciones de la muestra, no de θ— tales que

$$P_\theta\big(L(X) \le \theta \le U(X)\big) = 1-\alpha \qquad \text{para todo } \theta.$$

Esa probabilidad se llama *cobertura* del procedimiento. Si la igualdad sólo se cumple aproximadamente o en el límite, se habla de *cobertura real* frente a *cobertura nominal*, y el intervalo es *conservador* si la real supera a la nominal y *anticonservador* si queda por debajo.

Cuidado

[Lo aleatorio es el intervalo, no el parámetro] La probabilidad de la definición se toma *sobre las muestras*, con θ fijo. Una vez calculado el intervalo concreto $[2{,}1;\ 3{,}4]$, ya no hay nada aleatorio: o contiene a θ o no lo contiene, y no tiene sentido decir "hay un 95 % de probabilidad de que θ esté ahí". Lo correcto es: *el procedimiento* acierta el 95 % de las veces. La afirmación sobre el parámetro sí es legítima en el marco bayesiano del capítulo 20, con intervalos de credibilidad, y ésa es la diferencia conceptual más importante entre las dos escuelas.

Ejemplo 18.65 (Una cobertura real que no es la nominal)**.** El intervalo de Wald para una proporción, $\hat p \pm z_{\alpha/2}\sqrt{\hat p(1-\hat p)/n}$, tiene nivel nominal 95 %. Su cobertura real con $n=30$ y $p=0{,}1$ es de aproximadamente 0,88: falla más del doble de lo anunciado, porque cerca del borde $\hat p$ subestima la varianza y el intervalo sale demasiado corto. El intervalo de Wilson corrige esto y mantiene la cobertura cerca del 95 % en casi todo el rango. Es un buen recordatorio de que "nivel 95 %" es una promesa del método, y las promesas hay que comprobarlas.

Teorema 18.66 (Método delta)**.** *Sea $\hat\theta_n$ un estimador con*

$$\sqrt{n}\,(\hat\theta_n - \theta) \ \xrightarrow{d}\ N(0, \sigma^2),$$

y sea g derivable en θ con $g'(\theta) \neq 0$. Entonces

$$\sqrt{n}\,(g(\hat\theta_n) - g(\theta)) \ \xrightarrow{d}\ N\big(0,\ g'(\theta)^2\sigma^2\big).$$

En la práctica: $\operatorname{Var}(g(\hat\theta)) \approx g'(\theta)^2 \operatorname{Var}(\hat\theta)$.

Demostración. Desarróllese g por Taylor alrededor de θ:

$$g(\hat\theta_n) = g(\theta) + g'(\xi_n)(\hat\theta_n - \theta)$$

con ξ_n entre $\hat\theta_n$ y θ. Multiplicando por $\sqrt{n}$,

$$\sqrt{n}(g(\hat\theta_n)-g(\theta)) = g'(\xi_n)\cdot\sqrt{n}(\hat\theta_n-\theta).$$

Como $\hat\theta_n \xrightarrow{P} \theta$ —consecuencia de la convergencia en distribución a una normal centrada, dividida por $\sqrt{n}$—, también $\xi_n \xrightarrow{P} \theta$, y por continuidad de g' se tiene $g'(\xi_n)\xrightarrow{P} g'(\theta)$. El teorema de Slutsky permite entonces sustituir el factor por su límite, y queda $g'(\theta)$ multiplicando a una $N(0,\sigma^2)$. □

Ejemplo 18.67 (Error típico de un logaritmo y de una razón de momios)**.** Si $\hat p$ estima una proporción con $\operatorname{Var}(\hat p)=p(1-p)/n$ y se quiere el error de $\log\hat p$, entonces $g(p)=\log p$, $g'(p)=1/p$ y

$$\operatorname{Var}(\log\hat p)\approx \frac{1}{p^2}\cdot\frac{p(1-p)}{n}=\frac{1-p}{np}.$$

Con $p=0{,}2$ y $n=100$ sale 0,04, es decir un error típico de 0,2. Para la razón de momios $\mathrm{OR}=\frac{p/(1-p)}{q/(1-q)}$ el mismo cálculo da la fórmula que usa toda la epidemiología:

$$\operatorname{Var}\big(\log\widehat{\mathrm{OR}}\big)\approx \frac{1}{a}+\frac{1}{b}+\frac{1}{c}+\frac{1}{d},$$

con a,b,c,d las cuatro casillas de la tabla 2×2. Se construye el intervalo *en escala logarítmica* —donde la aproximación normal es buena— y se exponencia al final; hacerlo directamente sobre la OR produce intervalos asimétricos y a veces negativos.

18.18. Ejercicios resueltos

Tres problemas que separan lo que un intervalo dice de lo que la gente cree que dice.

Ejercicio 18.1. Con $\bar x=52$, $\sigma=8$ conocida y $n=64$, construya un intervalo de confianza al 95 por ciento para μ, y enuncie con precisión qué significa el 95 por ciento.

Solución. El error típico es $8/\sqrt{64}=1$, y el intervalo es $52\pm 1{,}96\cdot 1$, es decir, $(50{,}04;\ 53{,}96)$.

Lo que *no* significa: que haya un 95 por ciento de probabilidad de que μ esté ahí dentro. El parámetro μ es un número fijo; o está o no está, y no hay probabilidad que repartir. Lo que significa es que el *procedimiento* acierta el 95 por ciento de las veces: si se repitiese el muestreo muchas veces y se construyese cada vez el intervalo, el 95 por ciento de esos intervalos contendría a μ. La probabilidad está en el método, no en el resultado concreto. Quien quiera decir "hay un 95 por ciento de probabilidad de que μ esté aquí" tiene que hacerse bayesiano y calcular un intervalo creíble, que es otra cosa y a menudo sale parecido. ■

Ejercicio 18.2. Una muestra de $n=25$ da $\bar x=10{,}8$ y $s=2$. Contraste $H_0\colon \mu=10$ frente a $H_1\colon \mu\neq 10$ al 5 por ciento y calcule el p-valor. Interprete el resultado.

Solución. Como σ es desconocida y n es pequeño, se usa la t con 24 grados de libertad:

$$t=\frac{10{,}8-10}{2/\sqrt{25}}=\frac{0{,}8}{0{,}4}=2{,}00.$$

El valor crítico es $t_{24;\,0{,}975}=2{,}0639$, mayor que 2, luego *no* se rechaza H_0. El p-valor es 0,0569, ligeramente por encima de 0,05.

La interpretación honesta es que los datos no aportan evidencia suficiente contra H_0 al nivel fijado, y nada más. No se ha demostrado que $\mu = 10$: con $n = 25$ el contraste tiene poca potencia, y una diferencia real de 0,8 podría pasar desapercibida con facilidad. Nótese también lo arbitrario de la frontera: con $p = 0{,}0569$ se "acepta" y con $p = 0{,}0499$ se rechaza, cuando la evidencia es prácticamente la misma. Informar el p-valor y el tamaño del efecto dice mucho más que anunciar una decisión binaria. ■

Ejercicio 18.3. Se lanza un dado 120 veces y salen las frecuencias $22, 18, 20, 14, 26, 20$. Contraste si el dado es equilibrado.

Solución. Bajo H_0 cada cara esperaría $120/6 = 20$ veces. El estadístico es

$$X^2 = \sum \frac{(O_i - E_i)^2}{E_i} = \frac{4 + 4 + 0 + 36 + 36 + 0}{20} = 4{,}00,$$

con $6 - 1 = 5$ grados de libertad. El valor crítico al 5 por ciento es 11,07 y el p-valor es 0,549: no hay evidencia contra la equidad del dado.

Conviene fijarse en que la cara 4 salió seis veces menos de lo esperado y la 5 seis veces más, y que a ojo eso parece mucho. No lo es: con 120 tiradas la desviación típica del recuento de cada cara es $\sqrt{120 \cdot \frac{1}{6} \cdot \frac{5}{6}} \approx 4{,}08$, así que desviarse seis es menos de una vez y media esa cantidad. La intuición humana sobrestima sistemáticamente la regularidad que debería tener el azar, y ése es exactamente el trabajo que hace el contraste. ■

18.19. Ejercicios propuestos

18.1. Con $\bar{x} = 120$, $\sigma = 15$ y $n = 100$, dé el intervalo al 95 y al 99 por ciento.

18.2. Con $n = 16$, $\bar{x} = 12{,}4$ y $s = 2{,}1$, dé el intervalo al 95 por ciento usando la t.

18.3. En una encuesta, 208 de 400 responden que sí. Dé el intervalo al 95 por ciento para la proporción.

18.4. ¿Qué tamaño muestral hace falta para estimar una proporción con margen $\pm 0{,}02$ al 95 por ciento en el caso más desfavorable?

18.5. Explique la diferencia entre error de tipo I y de tipo II, y diga cuál controla α.

18.6. ¿Qué es la potencia de un contraste y de qué tres cosas depende?

18.7. Contraste $H_0\colon \mu_1 = \mu_2$ con $n_1 = n_2 = 30$, $\bar{x}_1 = 45$, $\bar{x}_2 = 41$ y $s = 6$ en ambos grupos.

18.8. ¿Qué significa exactamente un p-valor de 0,03? Dé también dos interpretaciones erróneas frecuentes.

18.9. ¿Por qué un resultado "estadísticamente significativo" puede ser irrelevante en la práctica?

18.10. Si se hacen 20 contrastes independientes al 5 por ciento sobre hipótesis todas ciertas, ¿cuántos rechazos falsos cabe esperar? ¿Qué corrección lo remedia?

18.11. ¿Cuándo se usa la t y cuándo la normal? ¿A partir de qué n deja de importar?

18.12. Explique por qué no se puede "aceptar H_0" y qué habría que hacer para concluir equivalencia.

Capítulo 19

Modelos lineales, regresión múltiple y análisis de la varianza

Lo que necesitas traer

Del capítulo 7: proyecciones ortogonales, rango, mínimos cuadrados e inversión de matrices. Del capítulo 18: estimadores, intervalos de confianza y contrastes. Del capítulo 16: la distribución normal y las distribuciones t, χ^2 y F.

La regresión simple relaciona una variable con otra. En cuanto hay más de un predictor aparecen fenómenos nuevos que no se pueden anticipar desde el caso de dos variables: coeficientes que cambian de signo al añadir una variable, predictores que resultan redundantes entre sí, y la necesidad de decidir cuántas variables merecen quedarse. Este capítulo presenta el modelo lineal en forma matricial, que unifica de golpe la regresión múltiple, el análisis de la varianza y buena parte de lo que se hace en estadística aplicada.

19.1. El modelo lineal en forma matricial

Definición 19.1. El *modelo lineal* escribe $\mathbf{y} = X\boldsymbol{\beta} + \boldsymbol{\varepsilon}$, donde X es la *matriz de diseño* de tamaño $n \times p$ (con una primera columna de unos para el término independiente), $\boldsymbol{\beta}$ el vector de coeficientes y $\boldsymbol{\varepsilon}$ los errores, que se suponen independientes con media cero y varianza común σ^2. El estimador de mínimos cuadrados es

$$\hat{\boldsymbol{\beta}} = (X^{\mathsf{T}}X)^{-1}X^{\mathsf{T}}\mathbf{y},$$

con matriz de covarianzas $\sigma^2(X^{\mathsf{T}}X)^{-1}$, y σ^2 se estima por $s^2 = \frac{\mathrm{SS}_{\mathrm{res}}}{n-p}$.

Idea clave

Qué dice esto en cristiano. La fórmula parece salida de la nada y es pura geometría: $X\hat{\boldsymbol{\beta}}$ es la *proyección ortogonal* de $\mathbf{y}$ sobre el subespacio generado por las columnas de X. Minimizar la suma de cuadrados de los residuos es exactamente buscar el punto de

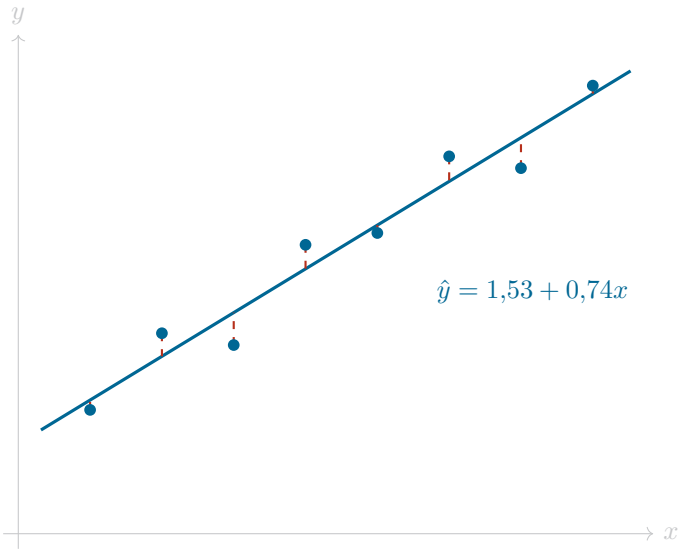


Figura 19.1: Ajuste por mínimos cuadrados. Los segmentos rojos son los residuos, y la recta elegida es la única que hace mínima la suma de sus cuadrados. Aquí $R^2 = 0{,}953$. Obsérvese que se miden distancias verticales, no perpendiculares: el modelo supone que el error está en y.

ese subespacio más cercano a $\mathbf{y}$, y el punto más cercano a un subespacio es siempre la proyección. De ahí sale que los residuos sean ortogonales a todas las columnas de X, que es lo que dicen las ecuaciones normales $X^{\mathsf{T}}(\mathbf{y} - X\hat{\boldsymbol{\beta}}) = \mathbf{0}$. Toda la regresión es álgebra lineal con nombres estadísticos.

Regla: ajustar e interpretar un modelo lineal múltiple

El análisis completo tiene siete pasos, y saltarse los tres últimos es el origen de la mayoría de los modelos publicados que luego no se replican:

1. **Monta la matriz de diseño** con la columna de unos y una columna por predictor; codifica las variables categóricas con variables indicadoras.
2. **Resuelve las ecuaciones normales** para obtener $\hat{\boldsymbol{\beta}}$; en la práctica se usa la factorización QR y no la inversa explícita, por estabilidad numérica.
3. **Calcula s^2 y los errores estándar** de cada coeficiente como la raíz de la diagonal de $s^2(X^{\mathsf{T}}X)^{-1}$.
4. **Contrasta cada coeficiente** con $t = \hat{\beta}_j / \mathrm{EE}(\hat{\beta}_j)$ y $n - p$ grados de libertad, y el modelo entero con el estadístico F global.
5. **Informa de R^2 ajustado** y no del R^2 crudo, porque este último siempre sube al añadir variables aunque sean ruido.
6. **Diagnostica los residuos**: dibújalos frente a los valores ajustados buscando curvatura o embudo, y frente a los cuantiles normales buscando colas.

7. **Comprueba la colinealidad** con el factor de inflación de la varianza; por encima de 5 o 10 los coeficientes individuales dejan de ser interpretables.

Ejemplo 19.2 (Una regresión con dos predictores, completa)**. (a) El ajuste.** Con ocho observaciones y dos predictores, el ajuste da $\hat\beta_0 = 0{,}757799$, $\hat\beta_1 = 1{,}586568$ y $\hat\beta_2 = 0{,}872097$. La suma de cuadrados residual es 0,050485 frente a una variación total de 261,315, luego $R^2 = 0{,}999807$ y R^2 ajustado 0,999730.

(b) La significación de cada coeficiente. Con $s^2 = 0{,}010097$, los errores estándar son 0,111524, 0,041780 y 0,038340, y los estadísticos t valen 6,79, 37,97 y 22,75, con p-valores 0,001051, prácticamente cero y 0,000003. Los tres coeficientes son claramente distintos de cero.

(c) El contraste global. El estadístico F vale 12 937,66 con 2 y 5 grados de libertad, y su p-valor es indistinguible de cero. El F global contrasta si *algún* coeficiente (aparte del término independiente) es no nulo; los t individuales contrastan cada uno por separado, y las dos preguntas son distintas.

(d) La colinealidad. La correlación entre los dos predictores es 0,928587, muy alta. El factor de inflación de la varianza es $\frac{1}{1-0{,}928587^2} = 7{,}260832$, por encima del umbral habitual de 5. Los coeficientes siguen siendo insesgados y el modelo predice bien, pero sus errores estándar están inflados y su interpretación individual es frágil: pequeños cambios en los datos los moverían mucho.

Cuidado

[El coeficiente de un predictor depende de qué otros haya en el modelo] En regresión múltiple, $\hat\beta_j$ mide el efecto de x_j *manteniendo constantes los demás predictores del modelo*, no el efecto total de x_j. Añadir o quitar una variable puede cambiar el valor de los otros coeficientes, e incluso su signo: es la paradoja de Simpson en versión continua. Segundo aviso: el R^2 crudo nunca baja al añadir una variable, ni siquiera si la variable es ruido puro, de modo que maximizarlo lleva siempre a sobreajustar; por eso existe el R^2 ajustado y los criterios AIC y BIC, que penalizan el número de parámetros. Y tercero: con colinealidad perfecta $X^\mathsf{T}X$ no es invertible y el modelo ni siquiera está determinado.

19.2. Antes de la forma matricial: la regresión lineal simple, las cuatro hipótesis del modelo y el diagnóstico por residuos

El capítulo entra directamente en la forma matricial $\mathbf{y} = X\boldsymbol{\beta} + \boldsymbol{\varepsilon}$, que es la manera correcta de organizar el asunto en cuanto hay más de un predictor pero que oculta de qué se está hablando a quien llega por primera vez. Esta sección recorre el caso de un solo predictor con todo el detalle, enuncia las cuatro hipótesis que el modelo pide —y que después se citan de pasada como "sus tres supuestos"—, define con precisión los residuos y los valores ajustados, demuestra la descomposición de la suma de cuadrados en forma general, y termina con lo que hay que mirar en los residuos y por qué.

El modelo y sus cuatro hipótesis

Definición 19.3 (Modelo de regresión lineal simple)**.** Sean n pares de observaciones (x_i, y_i). El *modelo de regresión lineal simple* postula que

$$Y_i = \beta_0 + \beta_1 x_i + \varepsilon_i, \qquad i = 1, \dots, n,$$

donde x_i es la variable *explicativa* o *regresor* (tratada como fija y medida sin error), Y_i la variable *respuesta*, β_0 y β_1 dos parámetros desconocidos, y ε_i el *error* o *perturbación*: la parte de Y_i que la recta no explica. El parámetro β_1 es el incremento medio de la respuesta por unidad de regresor, y β_0 el valor medio de la respuesta cuando el regresor vale cero.

Definición 19.4 (Las cuatro hipótesis del modelo lineal clásico)**.** El modelo anterior —y su versión múltiple $\mathbf{y} = X\boldsymbol{\beta} + \boldsymbol{\varepsilon}$— va acompañado de cuatro supuestos sobre los errores, que conviene tener numerados porque cada uno sirve para una cosa distinta y falla de manera distinta:

1. **Linealidad.** $\mathrm{E}[\varepsilon_i] = 0$ para todo i, es decir, la relación entre la media de la respuesta y los regresores es exactamente la que dice el modelo. Si falla, todos los coeficientes están sesgados y no hay remedio inferencial: el modelo es el equivocado.
2. **Independencia.** Los errores $\varepsilon_1, \dots, \varepsilon_n$ son mutuamente incorrelados, $\mathrm{Cov}(\varepsilon_i, \varepsilon_j) = 0$ para $i \neq j$. Si falla —datos temporales, medidas repetidas del mismo sujeto, alumnos de la misma clase—, los coeficientes siguen siendo insesgados pero los errores típicos quedan gravemente subestimados y todo sale significativo.
3. **Homocedasticidad.** $\mathrm{Var}(\varepsilon_i) = \sigma^2$, la misma para todos, sin depender de x_i ni de nada. Si falla, mínimos cuadrados deja de ser eficiente y los errores típicos dejan de ser correctos, aunque los coeficientes sigan siendo insesgados.
4. **Normalidad.** $\varepsilon_i \sim N(0, \sigma^2)$. Es la única de las cuatro que *no* hace falta para estimar: sólo se usa para que los contrastes t y F y los intervalos sean exactos con muestras pequeñas. Con n grande, el teorema central del límite la hace prescindible.

Las tres primeras se resumen en $\mathrm{E}[\boldsymbol{\varepsilon}] = \mathbf{0}$ y $\mathrm{Cov}(\boldsymbol{\varepsilon}) = \sigma^2 I$, que son exactamente las hipótesis del teorema de Gauss–Markov; la cuarta añade la distribución exacta de $\hat{\boldsymbol{\beta}}$.

Idea clave

Qué dice esto en cristiano. Las cuatro hipótesis no son igual de importantes ni fallan igual de grave, y ése es el orden en que están puestas. La linealidad es la más seria: si la relación es curva, no hay corrección de errores típicos que salve nada, porque el modelo está contestando a otra pregunta. La independencia es la más traicionera, porque no se ve en ningún gráfico de residuos contra ajustados y sin embargo puede multiplicar por tres la significación aparente. La homocedasticidad es visible de un vistazo —el embudo— y tiene remedios cómodos. Y la normalidad, que es la que más ansiedad produce en los cursos, es la menos importante de las cuatro. La regla mnemotécnica: preocúpate en el orden en que están numeradas, no al revés.

Mínimos cuadrados, valores ajustados y residuos

Teorema 19.5 (Estimadores de mínimos cuadrados en el caso simple)**.** *Los valores de β_0 y β_1 que minimizan $Q(\beta_0,\beta_1)=\sum_i (y_i-\beta_0-\beta_1 x_i)^2$ son*

$$\hat\beta_1=\frac{S_{xy}}{S_{xx}},\qquad \hat\beta_0=\bar y-\hat\beta_1\bar x,$$

con $S_{xy}=\sum_i (x_i-\bar x)(y_i-\bar y)$ y $S_{xx}=\sum_i (x_i-\bar x)^2$. Bajo las hipótesis 1 a 3 son insesgados, con

$$\operatorname{Var}(\hat\beta_1)=\frac{\sigma^2}{S_{xx}},\qquad \operatorname{Var}(\hat\beta_0)=\sigma^2\Big(\frac1n+\frac{\bar x^2}{S_{xx}}\Big).$$

Demostración. La minimización es el teorema 17.44 del capítulo 17, cuya demostración —anular las dos derivadas parciales y comprobar que la hessiana es definida positiva— no se repite aquí. Para la varianza basta escribir $\hat\beta_1$ como combinación lineal de las respuestas:

$$\hat\beta_1=\sum_i c_i Y_i,\qquad c_i=\frac{x_i-\bar x}{S_{xx}},$$

donde se ha usado que $\sum_i (x_i-\bar x)\bar y=0$. Los coeficientes c_i son constantes conocidas, luego, por la independencia y la varianza común de los errores,

$$\operatorname{Var}(\hat\beta_1)=\sum_i c_i^2\sigma^2=\frac{\sigma^2}{S_{xx}^2}\sum_i (x_i-\bar x)^2=\frac{\sigma^2}{S_{xx}}.$$

La insesgadez sale de $\mathrm{E}[Y_i]=\beta_0+\beta_1 x_i$ y de $\sum_i c_i=0$, $\sum_i c_i x_i=1$. El cálculo para $\hat\beta_0=\bar Y-\hat\beta_1\bar x$ es análogo, teniendo en cuenta que $\bar Y$ y $\hat\beta_1$ son incorrelados. □

La fórmula de $\operatorname{Var}(\hat\beta_1)$ contiene toda una teoría del diseño experimental en cuatro símbolos: la precisión de la pendiente depende de S_{xx}, es decir, de cuánto se *separen* los valores del regresor. Medir cien veces con x siempre parecido informa mucho menos que medir veinte veces con x bien repartido. Por eso los experimentos ponen los puntos en los extremos del rango, y por eso una regresión sobre datos donde x apenas varía tiene errores típicos enormes por muchas observaciones que se acumulen.

Definición 19.6 (Valores ajustados y residuos)**.** Con los estimadores anteriores se definen los *valores ajustados* —o predichos—

$$\hat y_i=\hat\beta_0+\hat\beta_1 x_i\qquad (\text{en general, } \hat{\mathbf{y}}=X\hat{\boldsymbol\beta}=H\mathbf{y})$$

y los *residuos*

$$e_i=y_i-\hat y_i\qquad (\text{en general, } \mathbf{e}=(I-H)\mathbf{y}),$$

donde $H=X(X^{\mathsf T}X)^{-1}X^{\mathsf T}$ es la matriz de proyección. El residuo es la parte observada del error: ε_i no se puede ver nunca, y e_i es lo más parecido que hay. Si el modelo incluye término independiente, los residuos cumplen siempre

$$\sum_i e_i=0,\qquad \sum_i x_i e_i=0,\qquad \sum_i \hat y_i e_i=0,$$

identidades que son las *ecuaciones normales* y que se cumplen por construcción, tanto si el modelo es correcto como si es un disparate.

Cuidado

[El residuo no es el error, y las tres identidades no validan nada] Es tentador comprobar que los residuos suman cero y concluir que el modelo está bien. No prueba absolutamente nada: esas tres igualdades salen de la fórmula de mínimos cuadrados y se cumplirían igual ajustando una recta a una circunferencia. Segundo aviso: los residuos *no* son independientes ni tienen todos la misma varianza, aunque los errores sí. En realidad $\operatorname{Var}(e_i) = \sigma^2(1-h_{ii})$, de modo que los puntos con mucho apalancamiento tienen residuos artificialmente pequeños —el modelo se ve obligado a pasar cerca de ellos— y por tanto se disfrazan justamente los datos más peligrosos. Ésa, y no otra, es la razón de que se usen residuos estandarizados en lugar de los brutos.

Teorema 19.7 (Descomposición de la suma de cuadrados)**.** *En un modelo lineal con término independiente,*

$$\underbrace{\sum_i (y_i-\bar y)^2}_{\text{SCT}} = \underbrace{\sum_i (\hat y_i-\bar y)^2}_{\text{SCR}} + \underbrace{\sum_i (y_i-\hat y_i)^2}_{\text{SCE}}.$$

Demostración. Escríbase $y_i - \bar y = (\hat y_i - \bar y) + e_i$, elévese al cuadrado y súmese:

$$\text{SCT} = \text{SCR} + \text{SCE} + 2\sum_i (\hat y_i - \bar y)e_i.$$

El término cruzado se anula: $\sum_i \hat y_i e_i = 0$ y $\sum_i e_i = 0$ por las ecuaciones normales de la definición 19.6, luego $\sum_i (\hat y_i - \bar y)e_i = 0$.

En lenguaje matricial la misma cuenta es el teorema de Pitágoras. El vector centrado $\mathbf{y} - \bar y\mathbf{1}$ se descompone como suma de su proyección $\hat{\mathbf{y}} - \bar y\mathbf{1}$ sobre el subespacio generado por las columnas de X y del residuo $\mathbf{e} = (I - H)\mathbf{y}$, que vive en el complemento ortogonal; como los dos sumandos son perpendiculares, las normas al cuadrado se suman. La hipótesis de que haya término independiente es la que garantiza que el vector $\mathbf{1}$ está en el subespacio y por tanto que centrar no rompe la ortogonalidad; sin constante, la identidad es falsa y el R^2 puede salir negativo. □

Observación 19.8 (Dos convenios para el R^2 ajustado, y cómo no equivocarse). El R^2 ajustado se escribe en el capítulo de dos formas, $1 - (1 - R^2)\frac{n-1}{n-p}$ y $1 - (1 - R^2)\frac{n-1}{n-p-1}$, y no son contradictorias: en la primera p es el número *total* de parámetros, constante incluida, y en la segunda es el número de *regresores*, constante excluida. Con dos predictores y término independiente, $p = 3$ en el primer convenio y $p = 2$ en el segundo, y ambas fórmulas dan lo mismo. La manera de no equivocarse nunca es recordar que el denominador debe ser siempre los *grados de libertad del residuo*, es decir, n menos el número de coeficientes estimados. Este libro usa el primer convenio.

Ejemplo 19.9 (Una regresión simple entera, de los datos al intervalo)**.** Diez estudiantes, con x las horas semanales de estudio e y la nota obtenida:

x	2	3	4	5	6	7	8	9	10	11
y	3,1	3,9	5,2	5,4	6,6	6,9	7,6	8,4	8,7	9,6

(a) Los estimadores. Con $\bar{x} = 6{,}5$ y $\bar{y} = 6{,}54$ salen $S_{xx} = 82{,}5$, $S_{yy} = 40{,}644$ y $S_{xy} = 57{,}5$, luego

$$\hat{\beta}_1 = \frac{57{,}5}{82{,}5} = 0{,}696970, \qquad \hat{\beta}_0 = 6{,}54 - 0{,}696970 \cdot 6{,}5 = 2{,}009697.$$

La lectura: cada hora semanal adicional de estudio se asocia a 0,70 puntos más de nota.

(b) Ajustados y residuos. Los valores ajustados son 3,4036; 4,1006; 4,7976; 5,4945; 6,1915; 6,8885; 7,5855; 8,2824; 8,9794; 9,6764, y los residuos

$$-0{,}3036;\ -0{,}2006;\ 0{,}4024;\ -0{,}0945;\ 0{,}4085;\ 0{,}0115;\ 0{,}0145;\ 0{,}1176;\ -0{,}2794;\ -0{,}0764,$$

que suman cero y cuyo producto escalar con x también es cero, como obliga la definición 19.6.

(c) Las sumas de cuadrados. SCT = 40,644, SCR = 40,075758 y SCE = 0,568242, cuya suma es exactamente SCT: el teorema 19.7, comprobado. De ahí

$$R^2 = \frac{40{,}075758}{40{,}644} = 0{,}986019, \qquad R^2_{\text{aj}} = 1 - (1 - 0{,}986019) \cdot \frac{9}{8} = 0{,}984271,$$

y $r = \sqrt{R^2} = 0{,}992985$, con signo positivo por serlo la pendiente.

(d) La inferencia. La varianza residual es $s^2 = \frac{0{,}568242}{8} = 0{,}071030$, luego $s = 0{,}266515$. El error típico de la pendiente es $\frac{s}{\sqrt{S_{xx}}} = 0{,}029342$, de donde

$$t = \frac{0{,}696970}{0{,}029342} = 23{,}753031, \qquad p = 1{,}05 \cdot 10^{-8},$$

con 8 grados de libertad. Como $t_{8;\,0{,}975} = 2{,}306004$, el intervalo al 95 % para β_1 es $[0{,}629306;\ 0{,}764633]$. El contraste global da $F = \frac{40{,}075758/1}{0{,}071030} = 564{,}206485$, que es exactamente t^2: con un solo regresor, el F global y el t de la pendiente son el mismo contraste.

Los dos contrastes, y la multicolinealidad

Definición 19.10 (Contraste individual y contraste global)**.** En un modelo con p parámetros ajustado sobre n observaciones:

- El *contraste individual* sobre el coeficiente j enfrenta $H_0\colon \beta_j = 0$ frente a $H_1\colon \beta_j \neq 0$ mediante $t = \hat{\beta}_j / \operatorname{ee}(\hat{\beta}_j)$, que bajo H_0 sigue una t_{n-p}. Responde a: *¿aporta este regresor algo, dado que todos los demás ya están en el modelo?*
- El *contraste global* enfrenta $H_0\colon \beta_1 = \cdots = \beta_{p-1} = 0$ —todos los coeficientes salvo la constante— frente a la alternativa de que alguno no sea nulo, mediante

  $$F = \frac{\text{SCR}/(p-1)}{\text{SCE}/(n-p)},$$

  que bajo H_0 sigue una $F_{p-1,\,n-p}$. Responde a: *¿explica algo el modelo en bloque?*

Las dos preguntas son distintas y sus respuestas pueden discrepar en ambos sentidos.

Definición 19.11 (Multicolinealidad)**.** Hay *multicolinealidad exacta* cuando alguna columna de la matriz de diseño X es combinación lineal de las demás; entonces X no tiene rango completo, $X^{\mathsf{T}}X$ no es invertible y $\hat{\boldsymbol{\beta}}$ ni siquiera está definido. Hay *multicolinealidad aproximada* —que es el caso real— cuando alguna columna es casi combinación lineal de las demás: el modelo se puede ajustar, pero los coeficientes individuales se vuelven inestables. Se mide con el factor de inflación de la varianza VIF_j definido en la sección de diagnóstico, y su efecto es multiplicar por VIF_j la varianza del coeficiente j respecto de la que tendría con regresores ortogonales.

Ejemplo 19.12 (Dos predictores muy correlacionados, con todo el diagnóstico)**.** Diez observaciones con x_1 las horas de estudio, x_2 el porcentaje de asistencia a clase e y la puntuación final sobre cien:

x_1	2	3	4	5	6	7	8	9	10	11
x_2	60	72	68	80	75	88	84	95	90	98
y	52	58	61	68	70	79	80	88	89	96

(a) El ajuste. Resolviendo las ecuaciones normales,

$$\hat{y} = 23{,}534062 + 3{,}498609\,x_1 + 0{,}343518\,x_2.$$

Con $n = 10$ y $p = 3$: SCT $= 1926{,}9$, SCE $= 2{,}782378$ y SCR $= 1924{,}117622$, luego $R^2 = 0{,}998556$ y $R^2_{\text{aj}} = 1 - (1 - 0{,}998556)\frac{9}{7} = 0{,}998143$. La varianza residual es $s^2 = \frac{2{,}782378}{7} = 0{,}397483$ y $s = 0{,}630462$.

(b) Los dos contrastes. Los errores típicos son 2,874398, 0,203757 y 0,050333, de modo que $t_0 = 8{,}187476$, $t_1 = 17{,}170476$ y $t_2 = 6{,}824927$, todos por encima de $t_{7;\,0{,}975} = 2{,}364624$, con p-valores 0,000079, 0,00000056 y 0,000248. El contraste global da

$$F = \frac{1924{,}117622/2}{0{,}397483} = 2420{,}379917,$$

frente al valor crítico $F_{0{,}95;\,2,7} = 4{,}737414$. Todo significativo.

(c) La colinealidad. La correlación entre x_1 y x_2 es 0,940187, de donde

$$\mathrm{VIF} = \frac{1}{1 - 0{,}940187^2} = 8{,}617120.$$

Está por encima del umbral de 5 y cerca del de 10: los errores típicos de las dos pendientes son $\sqrt{8{,}617} = 2{,}94$ veces mayores de lo que serían con regresores ortogonales. Aquí los coeficientes siguen saliendo significativos porque el ajuste es extraordinariamente bueno, pero la interpretación individual —"cada punto de asistencia aporta 0,34 puntos manteniendo fijas las horas de estudio"— descansa sobre muy pocos datos que separen ambas cosas.

(d) El contraste parcial. El modelo con sólo x_1 da $\hat{y} = 42{,}860606 + 4{,}806061\,x_1$ con $\mathrm{SCE}_0 = 21{,}296970$ y $R^2_0 = 0{,}988948$, que visto en solitario ya parece excelente. Pero

$$F = \frac{(21{,}296970 - 2{,}782378)/1}{0{,}397483} = 46{,}579633 = t_2^2,$$

así que x_2 aporta de forma inequívoca. Obsérvese además cómo cambia el coeficiente de x_1: pasa de 4,81 a 3,50 al entrar x_2. Es el fenómeno del que avisa el capítulo, y es la consecuencia directa de la colinealidad: parte de lo que en el modelo pequeño se atribuía a las horas de estudio era en realidad asistencia disfrazada.

Diagnóstico por residuos

Definición 19.13 (Residuos estandarizados y estudentizados)**.** Con h_{ii} el apalancamiento de la observación i y s la desviación típica residual, el *residuo estandarizado* —o internamente estudentizado— es

$$r_i = \frac{e_i}{s\sqrt{1-h_{ii}}},$$

que corrige el hecho de que $\operatorname{Var}(e_i) = \sigma^2(1-h_{ii})$ no sea constante. El *residuo estudentizado* externamente es

$$t_i = \frac{e_i}{s_{(i)}\sqrt{1-h_{ii}}},$$

donde $s_{(i)}$ es la desviación típica residual calculada *sin* la observación i; bajo el modelo sigue exactamente una t_{n-p-1}, y es por tanto el que se usa para decidir si un punto es atípico. Como referencia rápida, valores de $|r_i|$ o $|t_i|$ mayores que 2 merecen una mirada y mayores que 3 una explicación.

Ejemplo 19.14 (Diagnóstico de la regresión anterior)**.** Sobre el ajuste con dos predictores del ejemplo 19.12:

(a) Los residuos brutos. Valen 0,8576; −0,7632; 0,1123; −0,5086; −0,2896; 0,7461; −0,3785; 0,3442; −0,4368; 0,3164. Suman cero y no muestran ningún patrón evidente de signos.

(b) Los apalancamientos. h_{ii} vale 0,4410; 0,3675; 0,2532; 0,2686; 0,2100; 0,2686; 0,1740; 0,3040; 0,3675; 0,3456, y suman $3 = p$ como obliga la traza de la matriz sombrero. El umbral $2p/n = 0{,}6$ no lo cruza ninguno: no hay puntos tiranos, aunque el primero, con 0,44, es el que más manda.

(c) Los residuos corregidos. Estandarizados: 1,8195; −1,5221; 0,2060; −0,9432; −0,5168; 1,3837; −0,6605; 0,6544; −0,8711; 0,6205. Estudentizados externamente, el primero pasa de 1,82 a 2,3203, que ya roza el umbral de atención. La diferencia entre las dos versiones es justamente que la segunda no deja que el propio punto sospechoso infle la s con la que se juzga a sí mismo.

(d) La influencia. Las distancias de Cook valen como mucho 0,8707, en la primera observación, por debajo del umbral 1 pero muy por encima de $4/n = 0{,}4$; la segunda observación llega a 0,4487. Conclusión honrada: no hay ningún dato que invalide el ajuste, pero la primera observación merece que se compruebe de dónde salió, porque combina el mayor apalancamiento con el mayor residuo, que es exactamente la receta de la influencia.

Regla: diagnosticar un modelo lineal con los residuos

Cuatro gráficos y un orden, uno por hipótesis:

1. **Residuos frente a valores ajustados** —la hipótesis de linealidad—. Debe verse una nube horizontal sin forma. Una curva en U significa que falta un término cuadrático o una transformación; si aparece, para y arregla el modelo antes de mirar nada más.
2. **El mismo gráfico, para la homocedasticidad.** Un embudo que se abre hacia la

derecha es varianza creciente. Remedios: transformar la respuesta, ponderar, o usar errores típicos robustos.

3. **Residuos frente al orden de recogida o al tiempo** —la independencia—. Rachas largas del mismo signo delatan autocorrelación; el contraste de Durbin–Watson la cuantifica. Éste es el gráfico que casi nadie dibuja y el que detecta el fallo más caro.
4. **Gráfico cuantil-cuantil de los residuos** —la normalidad—. Se juzga si los puntos caen sobre una recta, no si el histograma parece una campana. Desviaciones moderadas con n grande son irrelevantes.
5. **Apalancamientos h_{ii} y distancias de Cook**, para localizar puntos con demasiado poder. Recuerda que atípico, palanca e influyente son tres cosas distintas.
6. **Nunca borres un punto porque estorbe.** Ajusta con él y sin él, y si las conclusiones cambian, dilo en el informe: eso *es* el resultado.
7. **Comprueba los VIF** antes de interpretar coeficientes individuales, y no antes de predecir: para predecir, la colinealidad es inofensiva.

Cuidado

[El diagnóstico no se hace con contrastes, se hace con dibujos] Existe la tentación de sustituir los cuatro gráficos por cuatro contrastes —Shapiro–Wilk para la normalidad, Breusch–Pagan para la homocedasticidad, Durbin–Watson para la independencia— y declarar el modelo válido si ninguno rechaza. Es mala idea por dos motivos simétricos. Con n pequeño, ninguno tiene potencia para detectar nada, y el "no rechazo" es puro desconocimiento. Con n grande, todos rechazan, porque ninguna hipótesis se cumple exactamente en el mundo real y basta muestra suficiente para probarlo; y sin embargo con n grande es cuando menos importan las desviaciones de la normalidad. La pregunta relevante nunca es si la hipótesis se cumple exactamente —no se cumple—, sino si se incumple lo bastante como para cambiar las conclusiones, y eso lo contesta un gráfico y un poco de criterio, no un p-valor.

19.3. Análisis de la varianza

Teorema 19.15 (Descomposición de la variabilidad)**.** *Para k grupos con N observaciones en total,*

$$\underbrace{\sum_{i,j}(x_{ij}-\bar{x})^2}_{\mathrm{SS_T}} = \underbrace{\sum_{i} n_i(\bar{x}_i-\bar{x})^2}_{\mathrm{SS_B}} + \underbrace{\sum_{i,j}(x_{ij}-\bar{x}_i)^2}_{\mathrm{SS_W}},$$

y bajo la hipótesis de igualdad de medias el cociente $F = \dfrac{\mathrm{SS_B}/(k-1)}{\mathrm{SS_W}/(N-k)}$ sigue una distribución F de Fisher con $k-1$ y $N-k$ grados de libertad.

Idea clave

Qué dice esto en cristiano. El análisis de la varianza tiene un nombre desconcertante, porque sirve para comparar *medias*. La lógica es esta: si todos los grupos tuvieran la misma media, la dispersión que se observa *entre* las medias de los grupos debería ser del mismo orden que la dispersión que hay *dentro* de cada grupo, ya que ambas vendrían solo del azar. Si la dispersión entre grupos es mucho mayor, es que las medias son de verdad distintas. El estadístico F es justamente ese cociente, y compararlo con 1 es comparar señal con ruido.

Ejemplo 19.16 (Tres tratamientos comparados)**. (a) Los datos.** Tres grupos de cinco observaciones con medias 23,8, 29 y 22, y media global 24,933333.

(b) La descomposición. $\mathrm{SS_B} = 132{,}133333$ con 2 grados de libertad y $\mathrm{SS_W} = 34{,}8$ con 12, cuya suma es $\mathrm{SS_T} = 166{,}933333$. Los cuadrados medios son $\mathrm{MS_B} = 66{,}066667$ y $\mathrm{MS_W} = 2{,}9$.

(c) La decisión. $F = \frac{66{,}066667}{2{,}9} = 22{,}781609$, muy por encima del valor crítico 3,885294, con p-valor $8{,}21 \cdot 10^{-5}$. Se rechaza la igualdad de medias. El tamaño del efecto es $\eta^2 = \frac{\mathrm{SS_B}}{\mathrm{SS_T}} = 0{,}791534$: el factor explica casi el 80 % de la variabilidad.

(d) Qué falta después. El F dice que *alguna* media difiere, no cuál. Para averiguarlo hay que hacer comparaciones por parejas, y ahí aparece el problema de la multiplicidad: con tres comparaciones y $\alpha = 0{,}05$ cada una, la probabilidad de algún falso positivo sube. La corrección de Bonferroni divide el nivel entre el número de comparaciones, usando $\frac{0{,}05}{3} = 0{,}016667$ en cada una.

Cuidado

[El ANOVA es un modelo lineal, con sus mismas hipótesis] Todo el análisis de la varianza se obtiene del modelo lineal codificando el grupo con variables indicadoras, así que hereda sus tres supuestos: independencia, normalidad de los residuos y *homocedasticidad*, es decir, igual varianza en todos los grupos. Este último es el más frágil, y con grupos de tamaños muy distintos su incumplimiento distorsiona seriamente el p-valor; en ese caso conviene la corrección de Welch. Segundo aviso: rechazar la hipótesis nula no dice qué grupos difieren ni por cuánto; sin las comparaciones múltiples y sin el tamaño del efecto, un F significativo informa de muy poco.

Pausa para pensar

El análisis de la varianza lo desarrolló Ronald Fisher en los años veinte en la estación experimental de Rothamsted, comparando rendimientos de cultivos bajo distintos fertilizantes. Allí inventó también dos ideas que hoy parecen obvias y entonces no lo eran: la *aleatorización* de las parcelas, para que ninguna diferencia sistemática del terreno se confundiera con el efecto del tratamiento, y el *diseño factorial*, que permite estudiar varios factores a la vez y sus interacciones en vez de uno cada vez. Su libro de 1935 sobre el diseño de experimentos empieza con el célebre ejemplo de la señora que afirmaba distinguir si el té se había servido antes o después que la leche, y que sirve

para explicar de qué manera un experimento bien diseñado convierte una anécdota en una pregunta contrastable.

19.4. Diagnóstico del modelo: qué mirar antes de creerse el resultado

Ajustar un modelo lineal es fácil: la fórmula $\hat{\beta} = (X^\top X)^{-1} X^\top y$ siempre devuelve algo mientras $X^\top X$ sea invertible. Lo difícil es decidir si ese algo significa lo que uno cree. Esta sección recorre las cuatro preguntas que un estadístico se hace después de ajustar y antes de publicar: ¿explica algo el modelo?, ¿aporta cada variable?, ¿hay observaciones que mandan demasiado?, ¿están las variables explicativas pisándose entre sí?

Definición 19.17. Con n observaciones y p parámetros (incluida la constante), se definen la *suma de cuadrados total* $\mathrm{SCT} = \sum_i (y_i - \bar{y})^2$, la *explicada por la regresión* $\mathrm{SCR} = \sum_i (\hat{y}_i - \bar{y})^2$ y la *residual* $\mathrm{SCE} = \sum_i (y_i - \hat{y}_i)^2$, que cumplen $\mathrm{SCT} = \mathrm{SCR} + \mathrm{SCE}$. El *coeficiente de determinación* y su versión corregida son

$$R^2 = \frac{\mathrm{SCR}}{\mathrm{SCT}}, \qquad R^2_{\mathrm{aj}} = 1 - (1 - R^2)\,\frac{n-1}{n-p}.$$

La *varianza residual* estimada es $s^2 = \mathrm{SCE}/(n-p)$, y el error típico del coeficiente j-ésimo es $\mathrm{ee}(\hat{\beta}_j) = s\sqrt{[(X^\top X)^{-1}]_{jj}}$.

Idea clave

Qué dice esto en cristiano. La variable y se mueve; esa agitación total es SCT. El modelo consigue reproducir una parte de ese movimiento (SCR) y el resto se lo queda sin explicar (SCE). El R^2 es simplemente la fracción del movimiento que el modelo reprodujo, un número entre 0 y 1. El problema del R^2 crudo es que *nunca baja* al añadir variables, aunque sean ruido puro: con suficientes columnas se puede reproducir cualquier y exactamente. El R^2 corregido cobra un peaje por cada parámetro gastado, de modo que sólo sube si la variable nueva aporta más de lo que cuesta.

Regla: Leer una salida de regresión en siete pasos

1. **Mirar el contraste global** F**.** $F = \dfrac{\mathrm{SCR}/(p-1)}{s^2}$ contra $F_{p-1,\,n-p}$. Si no se rechaza, el modelo entero no dice nada y lo demás sobra.
2. **Mirar** $R^2_{\mathbf{aj}}$, no R^2: es el que penaliza el exceso de variables.
3. **Mirar los** $t = \hat{\beta}_j/\mathrm{ee}(\hat{\beta}_j)$ de cada coeficiente, contra t_{n-p}. Un $|t|$ grande dice que esa variable aporta *dado que las demás ya están*.
4. **Mirar el signo** de cada $\hat{\beta}_j$: si contradice a la teoría, casi siempre hay colinealidad o una variable omitida.

5. **Dibujar los residuos contra $\hat{y}$.** Deben parecer una nube sin forma. Un embudo indica heterocedasticidad; una curva, que falta un término no lineal.
6. **Mirar los apalancamientos** $h_{ii} = [X(X^\top X)^{-1}X^\top]_{ii}$. Suman p, así que el valor típico es p/n; un h_{ii} mayor que $2p/n$ señala un punto con demasiado poder.
7. **Mirar los VIF** de cada columna. Por encima de 10 (algunos dicen 5) los coeficientes individuales dejan de ser interpretables aunque el modelo prediga bien.

Ejemplo 19.18 (Una regresión de ocho puntos, diagnosticada de cabo a rabo)**.** Tomamos $n = 8$ observaciones con dos regresores:

x_1	1	2	3	4	5	6	7	8
x_2	2	1	4	3	6	5	8	7
y	5	6	9	10	14	13	18	17

(a) El ajuste. Resolviendo las ecuaciones normales sale $\hat{\beta} = (2{,}5;\ 1;\ 1)$, es decir $\hat{y} = 2{,}5 + x_1 + x_2$. Los residuos son $-0{,}5;\ 0{,}5;\ -0{,}5;\ 0{,}5;\ 0{,}5;\ -0{,}5;\ 0{,}5;\ -0{,}5$: alternan de signo y tienen todos el mismo tamaño, cosa que en datos reales nunca pasa y que aquí delata que los números fueron construidos a mano. Suman cero, como obliga la constante del modelo.

(b) Las sumas de cuadrados. SCT $= 162$, SCE $= 2$, luego SCR $= 160$ y

$$R^2 = \frac{160}{162} = 0{,}987654, \qquad R^2_{\text{aj}} = 1 - (1 - 0{,}987654)\frac{7}{5} = 0{,}982716.$$

La varianza residual es $s^2 = 2/5 = 0{,}4$, de donde $s = 0{,}632456$. El contraste global da $F = \frac{160/2}{0{,}4} = 200$, muy por encima del valor crítico $F_{0{,}95;\,2,5} = 5{,}786135$: el modelo, en bloque, explica algo.

(c) Los coeficientes uno a uno. Los errores típicos salen 0,502494 para la constante y 0,229129 para cada pendiente, de modo que

$$t_0 = 4{,}975186, \qquad t_1 = t_2 = 4{,}364358,$$

todos por encima de $t_{0{,}975;\,5} = 2{,}570582$. Las dos variables aportan. Nótese que t_1 y t_2 coinciden: las dos columnas entran de forma simétrica en este conjunto de datos artificial.

(d) Apalancamiento y colinealidad. Los h_{ii} valen 0,475 para las cuatro observaciones de los extremos y 0,275 para las cuatro centrales; suman $3 = p$, como debe ser, y el umbral $2p/n = 0{,}75$ no lo cruza ninguna: no hay puntos tiranos. En cambio la correlación entre x_1 y x_2 es 0,904762, lo que da

$$\text{VIF} = \frac{1}{1 - 0{,}904762^2} = 5{,}5125.$$

Es alto. El modelo predice estupendamente, pero separar el efecto de x_1 del de x_2 es delicado: los dos errores típicos son $\sqrt{5{,}5125} \approx 2{,}35$ veces mayores de lo que serían con regresores ortogonales.

Definición 19.19. El *factor de inflación de la varianza* de la columna j es $\text{VIF}_j = 1/(1 - R_j^2)$, donde R_j^2 es el coeficiente de determinación de la regresión de x_j sobre todas las demás columnas explicativas.

Idea clave

Qué dice esto en cristiano. Si x_2 es casi predecible a partir de x_1, los datos apenas contienen información sobre qué parte del efecto le toca a cada una: se puede subir el coeficiente de una y bajar el de la otra sin que el ajuste empeore casi nada. Esa indeterminación se traduce en errores típicos enormes. El VIF mide exactamente cuántas veces se multiplica la varianza del coeficiente por culpa de esa redundancia: VIF $= 1$ significa columna ortogonal a las demás; VIF $= 100$ significa que el coeficiente es prácticamente inestimable por separado.

Ejemplo 19.20 (Comparar modelos: contraste parcial, AIC y BIC)**.** Con los mismos datos, ajustemos ahora el modelo reducido $y = \beta_0 + \beta_1 x_1$.

(a) El modelo reducido. Sale $\hat{y} = 2{,}928571 + 1{,}904762\, x_1$ con $\mathrm{SCE}_0 = 9{,}619048$, es decir $R^2_0 = 0{,}940623$. Visto en solitario parece excelente.

(b) El contraste parcial F. La pregunta "¿aporta x_2 una vez que x_1 ya está?" se responde con

$$F = \frac{(\mathrm{SCE}_0 - \mathrm{SCE})/1}{s^2} = \frac{9{,}619048 - 2}{0{,}4} = 19{,}047619,$$

con p-valor 0,00726 frente a $F_{1,5}$. Sí aporta. Obsérvese que con un solo grado de libertad en el numerador este F es exactamente el cuadrado del t del coeficiente: $4{,}364358^2 = 19{,}047619$.

(c) AIC y BIC. Con la fórmula gaussiana $\mathrm{AIC} = n \ln(\mathrm{SCE}/n) + 2p$ obtenemos $-5{,}090355$ para el modelo completo y 5,474430 para el reducido: el completo gana por más de diez unidades, diferencia que se considera decisiva. El BIC del modelo completo es $-4{,}852030$, y penaliza más los parámetros porque cambia el 2 por $\ln n = \ln 8 = 2{,}079442$.

(d) Qué no se puede comparar así. AIC y BIC sólo son comparables entre modelos ajustados *a los mismos datos y con la misma variable respuesta*. Comparar el AIC de un modelo para y con el de un modelo para $\ln y$ es un error frecuente y grave: las verosimilitudes viven en escalas distintas.

Cuidado

[Contraejemplo: un R^2 alto no valida nada] El cuarteto de Anscombe consiste en cuatro conjuntos de once puntos con *la misma* recta de regresión, el mismo $R^2 = 0{,}666$, las mismas medias y las mismas varianzas, y sin embargo: el primero es una nube razonable, el segundo es una parábola perfecta (donde la recta es absurda), el tercero es una recta exacta estropeada por un solo punto atípico, y el cuarto tiene todas las x iguales salvo una, que decide ella sola la pendiente. Ningún estadístico numérico distingue los cuatro casos; el dibujo de los residuos los distingue de un vistazo. Por eso el paso 5 de la regla no es opcional.

Cuidado

[Contraejemplo: significativo no es lo mismo que importante] Con $n = 100\,000$ observaciones, un coeficiente de 0,0003 con error típico 0,00001 da $t = 30$ y un p-valor

astronómicamente pequeño. Es significativo y es irrelevante: el efecto que describe no cambia nada en la práctica. El t mide cuán seguros estamos de que el efecto no es cero; el tamaño de $\hat{\beta}_j$ en unidades del problema mide si nos importa. Son dos preguntas distintas y sólo la segunda tiene contenido científico.

Pausa para pensar

El apalancamiento h_{ii} tiene una lectura preciosa: como $\hat{y} = Hy$ con H la matriz de proyección, se cumple $\partial \hat{y}_i / \partial y_i = h_{ii}$. Es decir, h_{ii} es literalmente cuánto se mueve la predicción en el punto i cuando se mueve el dato i. Un punto con $h_{ii} = 1$ arrastra al modelo consigo y su residuo es forzosamente cero: el modelo pasa por él porque no tiene más remedio. La traza de H vale p, así que el apalancamiento medio es siempre p/n pase lo que pase con los datos, y ese hecho es lo que convierte a $2p/n$ en un umbral universal en lugar de un número arbitrario.

19.5. El análisis de la varianza y los modelos lineales generalizados

Comparar dos medias es un contraste t; comparar cinco es un problema distinto, porque hacer diez contrastes por parejas dispara los falsos positivos. La solución de Fisher fue tratar la comparación como un modelo lineal, y de ahí salió media estadística aplicada del siglo pasado.

Definición 19.21. El *análisis de la varianza* de un factor descompone la variabilidad total en la debida a las diferencias entre grupos y la interna de cada grupo:

$$\underbrace{\sum_{ij} (y_{ij} - \bar{y})^2}_{SC_{\text{total}}} = \underbrace{\sum_{i} n_i (\bar{y}_i - \bar{y})^2}_{SC_{\text{entre}}} + \underbrace{\sum_{ij} (y_{ij} - \bar{y}_i)^2}_{SC_{\text{dentro}}},$$

y contrasta $H_0 \colon \mu_1 = \cdots = \mu_k$ mediante $F = \dfrac{SC_{\text{entre}}/(k-1)}{SC_{\text{dentro}}/(N-k)}$, que bajo H_0 sigue una $F_{k-1,\,N-k}$.

Idea clave

Qué dice esto en cristiano. El nombre engaña: el análisis de la varianza sirve para comparar *medias*, y la varianza es sólo el instrumento. La lógica es la de una carrera entre dos fuentes de dispersión. Si los grupos fuesen todos iguales, la separación entre sus medias sería sólo ruido, del mismo tamaño que el ruido interno de cada grupo, y el cociente F rondaría 1. Si F sale mucho mayor que 1, es que los grupos se separan más de lo que el ruido explica. Y es un modelo lineal disfrazado: basta codificar la pertenencia a cada grupo con variables indicadoras y el análisis de la varianza es la regresión de y sobre ellas.

Ejemplo 19.22 (Un análisis de la varianza completo)**.** Tres tratamientos con cinco observaciones cada uno: $A = \{23, 25, 21, 24, 22\}$, $B = \{28, 30, 27, 29, 31\}$ y $C = \{22, 24, 23, 21, 25\}$.

(a) Las medias. $\bar{y}_A = 23$, $\bar{y}_B = 29$, $\bar{y}_C = 23$, con media global 25.

(b) La tabla.

Fuente	SC	gl	CM	F
Entre	120	2	60	24,0
Dentro	30	12	2,5	
Total	150	14		

(c) La decisión. El valor crítico es $F_{2,12;\,0,95} = 3{,}885$ y el p-valor es 0,000064: se rechaza con holgura. La proporción de varianza explicada es $\eta^2 = 120/150 = 0{,}80$, un efecto enorme.

(d) Qué queda por hacer. El análisis dice que *alguna* media difiere, no cuál. En general hace falta una prueba *a posteriori* —Tukey, Bonferroni, Scheffé— que compare por parejas controlando la tasa global de error. Rechazar y luego mirar las medias sin corregir es exactamente el error que el análisis de la varianza venía a evitar.

Definición 19.23. Un *modelo lineal generalizado* consta de tres piezas: una distribución de la familia exponencial para Y, un predictor lineal $\eta = X\beta$ y una función de enlace g con $g(E[Y]) = \eta$. Con enlace identidad y normal se recupera la regresión ordinaria; con enlace logit y binomial se obtiene la *regresión logística* $\log \frac{p}{1-p} = X\beta$.

Ejemplo 19.24 (Leer una regresión logística)**. (a) Qué significa un coeficiente.** En logística, β_j no es un incremento de probabilidad sino de log-odds. Un $\beta = 0{,}7$ significa que las probabilidades a favor se multiplican por $\mathrm{e}^{0{,}7} = 2{,}014$.

(b) Por qué eso no es un incremento fijo de probabilidad. Si la probabilidad de partida es 0,20, los odds valen 0,25; multiplicados por 2,014 dan 0,503, que corresponde a una probabilidad de 0,335. El aumento es de 13,5 puntos. Pero si se parte de 0,80, el mismo coeficiente lleva la probabilidad a 0,890, apenas nueve puntos. El efecto en probabilidad depende de dónde se esté; el efecto en odds no.

(c) Por qué no se usa la recta. Ajustar una recta a datos 0/1 produce probabilidades predichas negativas o mayores que uno, y errores manifiestamente heterocedásticos. El enlace logit resuelve ambas cosas mapeando $(0, 1)$ sobre toda la recta real.

(d) Cómo se ajusta. No hay fórmula cerrada: la verosimilitud se maximiza por mínimos cuadrados ponderados iterativos, que es el método de Newton aplicado a la log-verosimilitud. Es una de las aplicaciones cotidianas de lo que en el capítulo 14 era un método para resolver ecuaciones.

Cuidado

[Contraejemplo: la separación perfecta rompe la logística] Si existe un hiperplano que separa exactamente los ceros de los unos, la verosimilitud no tiene máximo: se aproxima a 1 cuando los coeficientes tienden a infinito. El programa no falla, y ése es el problema: devuelve coeficientes enormes con errores típicos aún más enormes, y quien no esté

avisado leerá una relación fortísima donde en realidad no hay estimación posible. Ocurre a menudo con muestras pequeñas y predictores categóricos. El remedio habitual es penalizar.

Pausa para pensar

Fisher desarrolló el análisis de la varianza en la estación experimental de Rothamsted en los años veinte, comparando rendimientos de parcelas de trigo. Allí inventó también la aleatorización y el diseño de bloques, y allí escribió que "llamar al estadístico después del experimento es como pedirle que haga una autopsia: quizá pueda decir de qué murió". La estadística moderna nació en un campo de cultivo, y su vocabulario —tratamientos, bloques, parcelas— sigue delatándolo.

19.6. Por qué mínimos cuadrados: el teorema de Gauss–Markov

Hasta aquí hemos calculado $\hat{\beta} = (X^\top X)^{-1} X^\top y$ como si fuera la única opción concebible. No lo es: cualquier matriz A con $AX = I$ produce un estimador lineal insesgado $\tilde{\beta} = Ay$. La pregunta correcta es por qué elegimos precisamente ese, y la respuesta es un teorema.

Teorema 19.25 (Gauss–Markov)**.** *Supóngase $y = X\beta + \varepsilon$ con $E[\varepsilon] = 0$ y $\mathrm{Cov}(\varepsilon) = \sigma^2 I$. Entonces, entre todos los estimadores lineales e insesgados de $c^\top \beta$, el de mínimos cuadrados $c^\top \hat{\beta}$ tiene varianza mínima.*

Demostración. Sea $\tilde{\beta} = Ay$ insesgado, es decir $AX = I$. Escribiendo $A = (X^\top X)^{-1} X^\top + D$, la condición $AX = I$ obliga a $DX = 0$. Entonces

$$\mathrm{Cov}(\tilde{\beta}) = \sigma^2 AA^\top = \sigma^2 \left[(X^\top X)^{-1} + DD^\top \right],$$

porque los términos cruzados llevan $DX = 0$. Como $DD^\top$ es semidefinida positiva, $\mathrm{Var}(c^\top \tilde{\beta}) \geq \mathrm{Var}(c^\top \hat{\beta})$ para todo c, con igualdad solo si $D = 0$. □

Idea clave

Léase con cuidado lo que el teorema *no* pide: no pide normalidad. Solo pide media cero, varianza constante y errores incorrelados. Ese es todo el precio de la optimalidad. Y léase también lo que *sí* pide: varianza constante. En cuanto hay heterocedasticidad, mínimos cuadrados deja de ser óptimo, aunque siga siendo insesgado. Esa es la razón teórica de que la heterocedasticidad importe.

También conviene saber lo que el teorema deja fuera: estimadores *sesgados*. La regresión contraída y el lasso son sesgados y sin embargo pueden tener error cuadrático medio mucho menor. Gauss–Markov es óptimo dentro de una clase, no en absoluto.

Teorema 19.26 (Distribución exacta bajo normalidad)**.** *Si además* $\varepsilon \sim N(0, \sigma^2 I)$*, entonces*

$$\hat{\beta} \sim N\big(\beta,\ \sigma^2 (X^\top X)^{-1}\big), \qquad \frac{(n-p)s^2}{\sigma^2} \sim \chi^2_{n-p},$$

y ambos son independientes*. En consecuencia*

$$\frac{\hat{\beta}_j - \beta_j}{s\sqrt{[(X^\top X)^{-1}]_{jj}}} \sim t_{n-p}.$$

La independencia no es un detalle decorativo: es lo que permite dividir una normal por una raíz de χ^2 y obtener una t. Y sale de la geometría: $\hat{y} = Hy$ vive en el subespacio generado por las columnas de X, mientras que los residuos $e = (I - H)y$ viven en su ortogonal. Dos proyecciones ortogonales de un vector gaussiano esférico son independientes. El teorema de Cochran es exactamente esa afirmación, generalizada a cualquier descomposición del espacio en subespacios ortogonales, y es el que reparte los grados de libertad de la tabla ANOVA: $n - 1 = (k - 1) + (n - k)$ no es una casualidad aritmética, es una descomposición de $\mathbb{R}^n$.

Ejemplo 19.27 (Un estimador insesgado peor, para tocarlo con las manos)**.** Con $n = 8$ datos y el modelo $y = \beta_0 + \beta_1 x + \varepsilon$, considérese el estimador que *ignora las dos últimas observaciones* y ajusta la recta con las seis primeras. Es lineal en y y es insesgado. Pero su varianza es $\sigma^2 / S_{xx}^{(6)}$ frente a $\sigma^2 / S_{xx}^{(8)}$, y como S_{xx} solo puede crecer al añadir puntos, la varianza es estrictamente mayor. Gauss–Markov garantiza que *cualquier* atajo de este tipo —tirar datos, promediar a mano, ponderar por corazonada— paga en varianza.

19.7. Después de rechazar el F: Tukey, Bonferroni y Scheffé

El contraste F del ANOVA responde a una pregunta pobre: "¿son todas las medias iguales?". Cuando se rechaza, uno quiere saber *cuáles* difieren, y ahí empieza el problema de las comparaciones múltiples: con k grupos hay $\binom{k}{2}$ parejas, y si cada comparación se hace al 5 por ciento, la probabilidad de al menos un falso positivo se dispara.

Regla: los cuatro umbrales, de menos a más exigente

Todos tienen la forma "rechaza si $|\bar{y}_i - \bar{y}_j| >$ umbral". Con n observaciones por grupo y CME el cuadrado medio del error:

- **LSD de Fisher**: $t_{\alpha/2,\,\nu}\sqrt{2\,\mathrm{CME}/n}$. No corrige nada; solo es defendible con $k = 3$ y tras un F significativo.
- **Tukey**: $q_{\alpha,k,\nu}\sqrt{\mathrm{CME}/n}$, con q el rango studentizado. Controla exactamente el error por familia para *todas* las parejas. Es la elección estándar.
- **Bonferroni**: $t_{\alpha/(2m),\,\nu}\sqrt{2\,\mathrm{CME}/n}$ con m el número de comparaciones planeadas. Simple y válido siempre; conservador si m es grande.
- **Scheffé**: $\sqrt{(k-1)F_{\alpha,k-1,\nu}}\sqrt{2\,\mathrm{CME}/n}$. Cubre *todos* los contrastes imaginables, incluidos los que a uno se le ocurran después de ver los datos. Por eso es el más

conservador.

Ejemplo 19.28 (Tres tratamientos, cuatro umbrales y una conclusión)**.** Cinco pacientes por tratamiento:

						media
A	55	58	52	60	55	56
B	64	68	61	70	67	66
C	72	75	70	78	75	74

(a) La tabla ANOVA.

$$\mathrm{SC}_{\text{entre}} = 813{,}33 \text{ (2 gl)}, \qquad \mathrm{SC}_{\text{dentro}} = 126{,}00 \text{ (12 gl)}, \qquad \mathrm{CME} = 10{,}5,$$

$$F = \frac{813{,}33/2}{10{,}5} = 38{,}730\,2, \qquad p = 0{,}000\,006.$$

Se rechaza sin discusión la igualdad de las tres medias. Y ahora empieza el trabajo de verdad.

(b) Los cuatro umbrales. Con $\mathrm{CME} = 10{,}5$, $n = 5$ y $\nu = 12$:

método	constante	umbral
LSD	$t_{0{,}025;12} = 2{,}179$	4,465 2
Tukey	$q_{0{,}05;3,12} = 3{,}772\,9$	5,467 5
Bonferroni	$t_{0{,}05/6;12} = 2{,}779\,5$	5,696 2
Scheffé	$\sqrt{2F_{0{,}05;2,12}}$	5,712 8

Obsérvese el orden: LSD $<$ Tukey $<$ Bonferroni $<$ Scheffé, que es el orden de generalidad de lo que cada uno protege.

(c) Las tres comparaciones. Las diferencias son $B - A = 10$, $C - B = 8$ y $C - A = 18$. Las tres superan los cuatro umbrales, así que aquí los cuatro métodos coinciden: los tres tratamientos difieren entre sí dos a dos.

(d) Dónde habrían discrepado. Si la diferencia $C - B$ hubiera sido de 5 en lugar de 8, LSD la habría declarado significativa ($5 > 4{,}47$) y los otros tres no ($5 < 5{,}47$). Ese es exactamente el terreno donde la elección del método decide la conclusión publicada, y por eso hay que fijarla *antes* de mirar los datos.

Observación 19.29. Bonferroni y Tukey están tan cerca (5,70 frente a 5,47) porque con $k = 3$ solo hay tres comparaciones. Con $k = 6$ hay quince, y entonces Bonferroni se vuelve claramente más conservador que Tukey: la desigualdad de Bonferroni ignora que las comparaciones están correlacionadas (todas comparten el mismo CME y las mismas medias), mientras que el rango studentizado lo tiene en cuenta exactamente. La regla práctica es: para todas las parejas, Tukey; para un puñado de comparaciones decididas de antemano, Bonferroni; para contrastes exploratorios inventados a posteriori, Scheffé y solo Scheffé.

Idea clave

Hay una alternativa más potente que todas estas cuando las comparaciones se planifican de antemano: los **contrastes ortogonales**. Con k grupos caben $k-1$ contrastes mutuamente ortogonales, y sus sumas de cuadrados *parten exactamente* $\mathrm{SC}_{\text{entre}}$. Aquí, por ejemplo, $L_1 = \bar{y}_C - \bar{y}_A$ y $L_2 = \bar{y}_B - \frac{\bar{y}_A+\bar{y}_C}{2}$ son ortogonales, y sus sumas de cuadrados suman 813,33. Cada contraste responde a una pregunta científica concreta con un grado de libertad, y eso es siempre preferible a pescar entre parejas.

19.8. Dos intervalos que no son el mismo: la respuesta media y la observación futura

La regresión sirve para dos cosas distintas, y confundirlas es el error más común de todo el capítulo. Una cosa es preguntar *dónde está la recta verdadera* en x_0; otra, *dónde caerá el próximo dato* en x_0.

Definición 19.30 (Los dos intervalos)**.** Sea $\hat{y}_0 = x_0^\top \hat{\beta}$ y $h_0 = x_0^\top (X^\top X)^{-1} x_0$. Entonces

$$\text{respuesta media:}\quad \hat{y}_0 \pm t_{\alpha/2,n-p}\, s\sqrt{h_0}, \qquad \text{predicción:}\quad \hat{y}_0 \pm t_{\alpha/2,n-p}\, s\sqrt{1+h_0}.$$

En regresión simple, $h_0 = \dfrac{1}{n} + \dfrac{(x_0-\bar{x})^2}{S_{xx}}$.

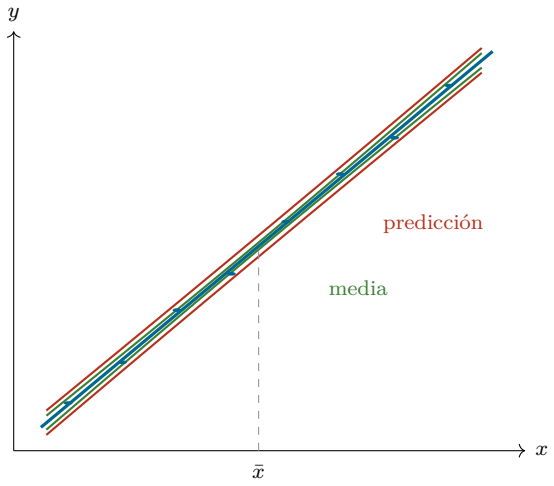


Figura 19.2: Las dos bandas al 95 por ciento. La estrecha acota dónde está la *recta* verdadera; la ancha, dónde caerá la *próxima observación*. Las dos tienen forma de reloj de arena, con cintura en $\bar{x}$, pero la de predicción arrastra además el ruido irreductible de la observación nueva y por eso nunca se estrecha por debajo de $\pm t\, s$.

El 1 de más lo dice todo: al predecir una observación nueva hay dos fuentes de error, la incertidumbre sobre la recta (h_0) y el ruido irreductible de esa observación concreta (1). Por muchos datos que se recojan, $h_0 \to 0$ pero el 1 se queda: el intervalo de predicción nunca se estrecha por debajo de $\pm t\, s$.

Ejemplo 19.31 (Ocho puntos y tres posiciones)**.** Con $x = 1, \dots, 8$ y los valores

$$y = 2{,}1;\quad 3{,}9;\quad 6{,}2;\quad 7{,}8;\quad 10{,}1;\quad 12{,}2;\quad 13{,}8;\quad 16{,}1,$$

el ajuste da

$$\hat\beta_0 = 0{,}035\,714, \qquad \hat\beta_1 = 1{,}997\,619, \qquad s = 0{,}180\,167,$$
$$S_{xx} = 42, \qquad \bar x = 4{,}5.$$

Con $t_{0{,}025;6} = 2{,}447$:

x_0	$\hat y_0$	$\pm$ respuesta media	$\pm$ predicción
2	4,031 0	0,230 7	0,497 6
4,5	9,025 0	0,155 9	0,467 6
8	16,016 7	0,284 6	0,524 7

Tres lecturas:

Primera. El intervalo de predicción es entre dos y tres veces más ancho que el de la respuesta media. Con ocho datos limpios, la recta se conoce muy bien; el próximo punto, no tanto.

Segunda. Ambos son mínimos en $x_0 = \bar x = 4{,}5$ y crecen hacia los extremos. La banda de confianza tiene forma de reloj de arena, porque un error en la pendiente se amplifica con el brazo de palanca $(x_0 - \bar x)$. En $x_0 = \bar x$ la pendiente no interviene y solo queda el $1/n$.

Tercera. La banda de predicción se ensancha mucho menos en términos relativos (0,468 $\to$ 0,525, un 12 por ciento) que la de la media (0,156 $\to$ 0,285, un 83 por ciento), justamente porque el 1 domina.

Cuidado

Estas bandas son *puntuales*: cada una tiene confianza 95 por ciento por separado. Dibujarlas a lo largo de todo el recorrido y afirmar que la recta verdadera está dentro en todas partes a la vez es un abuso. Para eso hacen falta las bandas simultáneas de Working–Hotelling, que sustituyen $t_{\alpha/2,n-p}$ por $\sqrt{p\,F_{\alpha,p,n-p}}$; aquí eso sería $\sqrt{2 \cdot 5{,}143} = 3{,}207$ en lugar de 2,447, un 31 por ciento más anchas.

Y la advertencia mayor: extrapolar. En $x_0 = 20$ las fórmulas siguen dando un número, con su intervalo y todo, pero ese número no descansa en ningún dato. Los intervalos cuantifican la incertidumbre *dentro* del modelo; no cuantifican la posibilidad de que el modelo no valga ahí fuera.

19.9. Análisis de la varianza de dos factores y la interacción

Con un solo factor, la pregunta es si los grupos difieren. Con dos, aparece una pregunta nueva que no existía antes y que casi siempre es la interesante: si el efecto de uno *depende* del nivel del otro.

Definición 19.32 (Modelo con interacción)**.** Con a niveles del factor A, b del factor B y n réplicas por casilla,

$$y_{ijk} = \mu + \alpha_i + \beta_j + (\alpha\beta)_{ij} + \varepsilon_{ijk},$$

y la suma total de cuadrados se parte en cuatro trozos ortogonales,

$$\mathrm{SC}_T = \mathrm{SC}_A + \mathrm{SC}_B + \mathrm{SC}_{AB} + \mathrm{SC}_E,$$

con $a-1$, $b-1$, $(a-1)(b-1)$ y $ab(n-1)$ grados de libertad. Cada efecto se contrasta dividiendo su cuadrado medio por el del error.

Ejemplo 19.33 (Un diseño 2×3 en el que el efecto principal miente)**.** Dos tratamientos (A) y tres dosis (B), con dos réplicas por casilla:

	b_1	b_2	b_3	media de fila
a_1	22, 24	28, 30	35, 37	29,33
a_2	30, 32	29, 31	26, 28	29,33
media de columna	27	29,5	31,5	29,33

(a) La tabla.

fuente	SC	gl	CM	F	p
A	0,00	1	0,00	0,00	1,000
B	40,67	2	20,33	10,17	0,0118
$A \times B$	146,00	2	73,00	36,50	0,00044
error	12,00	6	2,00		
total	198,67	11			

(b) La lectura ingenua sería desastrosa. El efecto principal de A es *exactamente* cero: las dos medias de fila valen 29,33. Quien mirase solo esa línea concluiría que el tratamiento no hace nada.

(c) Lo que pasa de verdad. Las medias de casilla son

	b_1	b_2	b_3
a_1	23	29	36
a_2	31	30	27

Con el tratamiento a_1 la respuesta *sube* con la dosis; con a_2 *baja*. Los dos efectos son fuertes y de signos opuestos, y al promediarlos se cancelan. La interacción, con $F = 36{,}5$ y $p = 0{,}00044$, es de largo el efecto más significativo de la tabla.

(d) La regla. Cuando la interacción es significativa, los efectos principales *no se interpretan*. Lo que hay que hacer es analizar el efecto de un factor por separado dentro de cada nivel del otro —los llamados efectos simples—, y presentar un gráfico de perfiles: dos rectas que se cruzan son la firma visual de la interacción, y dos rectas paralelas su ausencia.

Regla: diseños que conviene conocer

- **Factorial completo** 2^k: todos los factores a dos niveles. Con $k = 5$ son 32 combinaciones y se estiman los cinco efectos principales, las diez interacciones dobles y el resto. Es el caballo de batalla de la mejora de procesos.

- **Bloques aleatorizados**: cuando hay una fuente de variabilidad conocida y molesta (el día, el lote, el operario), se convierte en factor de bloque. Es la misma idea que el análisis apareado, generalizada.
- **Medidas repetidas**: cada sujeto pasa por todos los tratamientos. Exige tener en cuenta la correlación dentro del sujeto; ignorarla infla la significación.
- **Efectos fijos frente a aleatorios**: si los niveles son *los* que interesan (tres fármacos concretos), son fijos y se estiman medias. Si son una muestra de una población de niveles (diez lotes al azar de una producción), son aleatorios y lo que se estima es una *componente de varianza*. El denominador del contraste F cambia según el caso, y equivocarse ahí es un error frecuente.

Cuidado

Todo lo anterior supone un diseño **equilibrado**: el mismo número de observaciones en cada casilla. Si no lo está, los efectos dejan de ser ortogonales, las sumas de cuadrados ya no se reparten limpiamente y el resultado *depende del orden* en que se introduzcan los factores. De ahí las sumas de cuadrados de tipo I (secuenciales), II y III (marginales), que dan respuestas distintas sobre los mismos datos. Con diseños equilibrados las tres coinciden; con desequilibrados hay que decir cuál se ha usado.

19.10. Remedios: mínimos cuadrados ponderados, errores robustos y contracción

El capítulo enseña a *detectar* heterocedasticidad y colinealidad. Faltan los remedios, y hay uno para cada diagnóstico.

Definición 19.34 (Mínimos cuadrados generalizados y ponderados)**.** Si $\operatorname{Cov}(\varepsilon) = \sigma^2\Sigma$ con Σ conocida, se premultiplica todo por $\Sigma^{-1/2}$ y se aplica mínimos cuadrados ordinarios al modelo transformado:

$$\hat\beta_{\text{MCG}} = (X^\top \Sigma^{-1} X)^{-1} X^\top \Sigma^{-1} y.$$

Cuando Σ es diagonal se llama *ponderado*, con pesos $w_i = 1/\sigma_i^2$.

Y ese estimador es el que Gauss–Markov corona: bajo covarianza $\sigma^2\Sigma$, el óptimo entre los lineales insesgados es el de MCG, no el ordinario.

Ejemplo 19.35 (Varianza proporcional a x)**.** Ocho observaciones con $\operatorname{Var}(\varepsilon_i) \propto x_i$, es decir, pesos $w_i = 1/x_i$.

	$\hat\beta_0$	$\hat\beta_1$	$\text{ee}(\hat\beta_0)$	$\text{ee}(\hat\beta_1)$
MCO	0,4786	1,9131	1,0440	0,2067
MCP	0,2765	1,9580	0,5848	0,1607
MCO con ee robustos	0,4786	1,9131	0,6152	0,1859

Tres lecturas. Primera: los coeficientes apenas cambian —MCO sigue siendo insesgado—, pero los errores típicos sí: el de la pendiente baja un 22 por ciento al ponderar. Segunda: ponderar no es cosmética, es recuperar eficiencia; los intervalos son genuinamente más cortos porque se está usando la información de que unas observaciones son más fiables que otras. Tercera: los errores típicos robustos de White quedan en medio. No arreglan la ineficiencia, pero sí corrigen la *inferencia*: son válidos aunque la forma de la heterocedasticidad se desconozca, que es la situación normal.

Regla: el catálogo de remedios

- **Heterocedasticidad.** Contrástala con Breusch–Pagan o White. Si conoces la forma, pondera; si no, usa errores robustos de White (o de Huber–White agrupados si hay conglomerados). Transformar la respuesta —logaritmo, raíz— a menudo la elimina de raíz.
- **Errores correlacionados.** Detéctalos con Durbin–Watson en series temporales. El remedio es MCG con una estructura autorregresiva, o modelos de series temporales propiamente dichos.
- **Colinealidad.** Con VIF alto hay tres salidas: quitar variables redundantes, usar componentes principales como predictores, o contraer. La **regresión contraída** minimiza
$$\|y - X\beta\|^2 + \lambda\|\beta\|^2, \qquad \hat{\beta}_\lambda = (X^\top X + \lambda I)^{-1} X^\top y,$$
que siempre es invertible aunque $X^\top X$ no lo sea. El precio es un sesgo, y la ganancia es una varianza mucho menor: con el λ adecuado el error cuadrático medio total *baja*. Es el compromiso sesgo-varianza en su forma más limpia, y es también la razón de que Gauss–Markov no sea la última palabra.
- **Selección ralas.** El **lasso** sustituye $\|\beta\|^2$ por $\|\beta\|_1$, y la esquina de la bola ℓ^1 hace que la solución tenga coeficientes exactamente nulos: selecciona y ajusta a la vez.

19.11. Diseño de experimentos: bloques, cuadrados latinos y factoriales fraccionados

El análisis de la varianza contesta preguntas sobre datos ya recogidos. El diseño de experimentos decide *cómo recogerlos*, y es la parte con mayor rentabilidad por unidad de esfuerzo de toda la estadística: un diseño bueno hace trivial el análisis, y ninguno arregla un diseño malo.

Regla: los tres principios de Fisher

1. **Aleatorización.** Asignar los tratamientos al azar es lo único que convierte la asociación en causalidad, porque neutraliza *todas* las variables de confusión a la vez, incluidas las que no se conocen. Es la diferencia entre un experimento y un estudio observacional.

2. **Repetición.** Sin réplicas no hay estimación del error, y sin error no hay contraste posible. Ojo con la pseudorreplicación: tres medidas del mismo animal son una réplica, no tres.

3. **Bloqueo.** Toda fuente de variabilidad conocida y molesta debe convertirse en un factor de bloque. El lema es: *bloquea lo que puedas controlar, aleatoriza lo que no.*

Definición 19.36 (Bloques completos aleatorizados)**.** Cada bloque contiene una vez cada tratamiento, asignados al azar dentro del bloque. El modelo es $y_{ij} = \mu + \tau_i + \beta_j + \varepsilon_{ij}$, y la suma de cuadrados de los bloques se extrae del error, reduciéndolo.

Es la generalización exacta del análisis apareado: con dos tratamientos y bloques de tamaño dos, el contraste F del diseño en bloques *es* el cuadrado de la t apareada.

Definición 19.37 (Cuadrado latino)**.** Con dos fuentes de variabilidad a bloquear, se dispone una tabla $t \times t$ en la que cada tratamiento aparece exactamente una vez en cada fila y en cada columna:

	C_1	C_2	C_3	C_4
F_1	A	B	C	D
F_2	B	C	D	A
F_3	C	D	A	B
F_4	D	A	B	C

Ejemplo 19.38 (Lo que se ahorra, contado en grados de libertad)**.** Con $t = 4$ tratamientos, un factorial completo que cruzara tratamiento $\times$ fila $\times$ columna exigiría $4^3 = 64$ observaciones. El cuadrado latino usa 16 y reparte:

fuente	gl
filas	3
columnas	3
tratamientos	3
error	6
total	15

Una cuarta parte del coste. El precio está en los seis grados de libertad del error, que son pocos, y sobre todo en el supuesto de *no interacción*: el diseño confunde las interacciones con los efectos principales, y si existen, los resultados no son interpretables. Con $t = 3$ el error tendría solo 2 grados de libertad, y el diseño se vuelve inútil sin réplicas.

Definición 19.39 (Factorial fraccionado 2^{k-p})**.** Se estudian k factores a dos niveles usando solo 2^{k-p} de las 2^k combinaciones, elegidas mediante p *generadores* que definen la relación de definición. La **resolución** del diseño es la longitud de la palabra más corta de esa relación.

Regla: cómo leer la resolución

- **Resolución III**: los efectos principales se confunden con interacciones de dos factores. Sirve para cribar muchos factores y quedarse con los pocos que importan.
- **Resolución IV**: los efectos principales están limpios de interacciones dobles, pero las dobles se confunden entre sí. Es el punto dulce habitual.
- **Resolución V**: efectos principales e interacciones dobles todos limpios. Caro pero interpretable del todo.

Con siete factores, el completo pide $2^7 = 128$ corridas; el 2^{7-4} pide 8 (resolución III) y el 2^{7-3} pide 16 (resolución IV). Pasar de 128 a 16 experimentos sin perder los efectos principales es la razón de que estos diseños dominen la industria.

Observación 19.40 (El principio que lo justifica)*.* Todo esto descansa en tres supuestos empíricos que casi siempre se cumplen. *Escasez de efectos*: de veinte factores, solo unos pocos importan. *Jerarquía*: los efectos principales suelen ser mayores que las interacciones dobles, y estas mayores que las triples. *Herencia*: una interacción relevante casi siempre involucra factores cuyos efectos principales también lo son. Si esos tres se cumplen, confundir efectos altos con bajos es un riesgo asumible; si no, el diseño fraccionado engaña.

El paso siguiente, cuando ya se sabe qué factores importan y se busca el óptimo, es la **metodología de superficie de respuesta**: se ajusta un modelo cuadrático con un diseño central compuesto, se calcula el punto estacionario igualando el gradiente a cero, y se clasifica con el hessiano —máximo, mínimo o silla— exactamente como en el capítulo de cálculo multivariable. El experimento y la optimización acaban siendo el mismo problema.

19.12. La distribución exacta de los estimadores y el teorema de Cochran

Toda la inferencia del capítulo —los contrastes t de los coeficientes, la F del modelo, los intervalos— descansa en dos hechos distribucionales que se citaron de oídas. Se demuestran aquí.

Teorema 19.41 (Distribución exacta en el modelo normal)**.** *Sea $y = X\beta + \varepsilon$ con $X \in \mathbb{R}^{n\times p}$ de rango p y $\varepsilon \sim N(0, \sigma^2 I_n)$. Sean $\hat\beta = (X^\top X)^{-1}X^\top y$ y $s^2 = \frac{\|y - X\hat\beta\|^2}{n-p}$. Entonces:*

1. $\hat\beta \sim N_p(\beta,\ \sigma^2(X^\top X)^{-1})$*;*
2. $\dfrac{(n-p)s^2}{\sigma^2} \sim \chi^2_{n-p}$*;*
3. $\hat\beta$ *y* s^2 *son independientes;*
4. *en consecuencia,* $\dfrac{\hat\beta_j - \beta_j}{s\sqrt{[(X^\top X)^{-1}]_{jj}}} \sim t_{n-p}$.

Demostración. Sea $P = X(X^\top X)^{-1}X^\top$ la matriz de proyección ortogonal sobre el espacio de columnas $\mathcal{C}(X)$, de dimensión p; es simétrica e idempotente ($P^2 = P$), y $I - P$ proyecta sobre el complemento ortogonal, de dimensión $n - p$.

(1) $\hat\beta$ es una función *lineal* de y, que es normal multivariante; luego $\hat\beta$ es normal. Su media es $(X^\top X)^{-1}X^\top X\beta = \beta$, y su matriz de covarianzas

$$(X^\top X)^{-1}X^\top(\sigma^2 I)X(X^\top X)^{-1} = \sigma^2(X^\top X)^{-1}.$$

(2) y (3). Tómese una base ortonormal de $\mathbb{R}^n$ cuyos primeros p vectores generen $\mathcal{C}(X)$ y los $n - p$ restantes su ortogonal, y sea Q la matriz ortogonal correspondiente. Póngase $z = Q^\top\varepsilon$. Como en el teorema de Fisher del capítulo 16, una transformación ortogonal de un vector $N(0, \sigma^2 I)$ vuelve a ser $N(0, \sigma^2 I)$: las componentes $z_1, \dots, z_n$ son normales independientes de varianza σ^2.

Ahora bien, $\hat\beta - \beta = (X^\top X)^{-1}X^\top\varepsilon$ depende sólo de la proyección de ε sobre $\mathcal{C}(X)$, es decir sólo de $z_1, \dots, z_p$; mientras que el vector de residuos es

$$y - X\hat\beta = (I - P)y = (I - P)\varepsilon,$$

que depende sólo de $z_{p+1}, \dots, z_n$. Dos funciones de bloques disjuntos de variables independientes son independientes: eso es (3). Y como las proyecciones ortogonales conservan normas,

$$(n-p)s^2 = \|(I-P)\varepsilon\|^2 = \sum_{i=p+1}^{n} z_i^2,$$

suma de $n - p$ normales independientes de varianza σ^2, luego $(n-p)s^2/\sigma^2 \sim \chi^2_{n-p}$: eso es (2).

(4) Por (1), $\hat\beta_j - \beta_j \sim N(0, \sigma^2 v_{jj})$ con $v_{jj} = [(X^\top X)^{-1}]_{jj}$; al dividir entre $\sigma\sqrt{v_{jj}}$ queda una $N(0,1)$, y al sustituir σ por s aparece el cociente $Z/\sqrt{Q/(n-p)}$ con Z y Q independientes por (3): es la definición de t_{n-p}. □

Teorema 19.42 (Cochran)**.** *Sea $z \sim N_n(0, I)$ y supóngase que*

$$\|z\|^2 = Q_1 + \cdots + Q_k, \qquad Q_i = z^\top A_i z,$$

con matrices simétricas A_i de rangos r_i tales que $\sum_i r_i = n$. Entonces los Q_i son independientes y $Q_i \sim \chi^2_{r_i}$.

Demostración. Como $\sum A_i = I$ y $\sum \operatorname{rango} A_i = n$, cada A_i es necesariamente una matriz de *proyección ortogonal* y las imágenes son mutuamente ortogonales. En efecto: si V_i es la imagen de A_i, entonces $\sum_i \dim V_i = n$ y los V_i generan $\mathbb{R}^n$ (porque $\sum A_i = I$), luego la suma es directa y $\mathbb{R}^n = \bigoplus V_i$; escribiendo $z = \sum z_i$ con $z_i \in V_i$ y usando $\sum z^\top A_i z = \|z\|^2$ para todo z, se comprueba que la descomposición debe ser ortogonal y que A_i actúa como la proyección sobre V_i (las formas cuadráticas coinciden, y una forma cuadrática determina la matriz simétrica).

Tómese entonces una base ortonormal adaptada a la descomposición $\mathbb{R}^n = \bigoplus V_i$, y sea Q la matriz ortogonal del cambio. El vector $w = Q^\top z$ sigue siendo $N(0, I)$, y cada Q_i se convierte en la suma de los r_i cuadrados w_j^2 correspondientes al bloque i. Bloques disjuntos de variables independientes: los Q_i son independientes y cada uno es una χ^2 con r_i grados de libertad. □

Corolario 19.43 (La tabla ANOVA, justificada)**.** *En el análisis de la varianza de un factor con k grupos y n observaciones totales, bajo la hipótesis nula de igualdad de medias se tiene*

$$\underbrace{\frac{SC_{entre}}{\sigma^2}}_{\chi^2_{k-1}} + \underbrace{\frac{SC_{dentro}}{\sigma^2}}_{\chi^2_{n-k}} = \frac{SC_{total}}{\sigma^2} \sim \chi^2_{n-1},$$

con los dos sumandos independientes, y por tanto

$$F = \frac{SC_{entre}/(k-1)}{SC_{dentro}/(n-k)} \;\sim\; F_{k-1,\,n-k}.$$

Demostración. Los dos sumandos son formas cuadráticas asociadas a proyecciones sobre subespacios ortogonales —el generado por los indicadores de grupo menos la constante, y su complemento— de dimensiones $k-1$ y $n-k$, cuya suma es $n-1$: exactamente las hipótesis del teorema 19.42 aplicado dentro del subespacio ortogonal al vector de unos. El cociente de dos χ^2 independientes divididas por sus grados de libertad es, por definición, una F. □

Observación 19.44 (Los grados de libertad no son un convenio)*.* La igualdad $(k-1)+(n-k) = n-1$ que aparece en toda tabla ANOVA es una identidad de *dimensiones de subespacios*, no una regla mnemotécnica: la variabilidad total vive en un espacio de dimensión $n-1$ (se ha perdido una por centrar), y se reparte en dos direcciones ortogonales. Cada vez que un modelo lineal se amplía con un predictor, se roba un grado de libertad al residuo porque se le quita una dimensión a su subespacio. Ver la tabla ANOVA como el teorema de Pitágoras en $\mathbb{R}^n$, con las dimensiones de los catetos anotadas al margen, es la manera de no volver a equivocarse con los grados de libertad.

19.13. Diagnóstico y selección: las definiciones explícitas

Definición 19.45 (Apalancamiento y distancia de Cook)**.** Sea $H = X(X^\top X)^{-1}X^\top$ la matriz de proyección (*matriz sombrero*). El *apalancamiento* de la observación i es

$$h_{ii} = \big[H\big]_{ii} \;\in [0,1], \qquad \sum_i h_{ii} = p.$$

Mide cuánto se aleja x_i del centro de los predictores; su valor medio es p/n, y se considera *alto* si $h_{ii} > 2p/n$.

La *distancia de Cook* de la observación i es

$$D_i = \frac{e_i^2}{p\,s^2} \cdot \frac{h_{ii}}{(1-h_{ii})^2},$$

donde $e_i = y_i - \hat{y}_i$. Equivale a medir cuánto cambian *todas* las predicciones si se elimina esa observación. Se investiga si $D_i > 1$, y con más razón si $D_i > 4/n$.

Observación 19.46 (Atípico, influyente y palanca no son lo mismo)*.* Son tres cosas distintas y conviene no confundirlas. Un punto es *atípico* si su residuo es grande, es decir, si el modelo lo predice mal. Tiene *apalancamiento* alto si su x está lejos del resto, con independencia de su

y. Y es *influyente* si al quitarlo cambian los coeficientes, lo que requiere *las dos cosas a la vez*: residuo apreciable y apalancamiento apreciable. Eso es exactamente lo que multiplica la fórmula de Cook. Un punto con h_{ii} enorme pero residuo nulo no molesta —está lejos, pero sobre la recta—, y un residuo grande en el centro de los datos apenas mueve la pendiente.

Definición 19.47 (AIC y BIC)**.** Para un modelo ajustado por máxima verosimilitud con p parámetros y verosimilitud maximizada $\hat{L}$:

$$\mathrm{AIC} = -2\log\hat{L} + 2p, \qquad \mathrm{BIC} = -2\log\hat{L} + p\log n.$$

En regresión lineal normal, salvo constantes comunes a todos los modelos,

$$\mathrm{AIC} = n\log\left(\frac{\mathrm{SCE}}{n}\right) + 2p, \qquad \mathrm{BIC} = n\log\left(\frac{\mathrm{SCE}}{n}\right) + p\log n.$$

En ambos casos, *menor es mejor*, y sólo son comparables entre modelos ajustados sobre los *mismos* datos.

Observación 19.48 (Por qué eligen modelos distintos)*.* La única diferencia es la penalización por parámetro: 2 frente a $\log n$. Como $\log n > 2$ en cuanto $n \geq 8$, el BIC penaliza más y elige modelos *más pequeños*; con $n = 100$ la penalización del BIC es 4,6 por parámetro, más del doble. La diferencia no es de gusto: el AIC busca minimizar el error de predicción y no es consistente —tiende a incluir variables de más aunque $n \to \infty$—, mientras que el BIC es consistente para seleccionar el modelo verdadero si éste está en la lista, pero puede quedarse corto en predicción. Si el objetivo es predecir, AIC; si es identificar qué variables importan de verdad, BIC.

Teorema 19.49 (Contraste F parcial, caso general)**.** *Sean dos modelos anidados: el reducido, con p_0 parámetros y suma de cuadrados del error SCE_0, y el completo, con p parámetros y SCE, donde el reducido se obtiene imponiendo $q = p - p_0$ restricciones lineales. Bajo la hipótesis nula de que esas restricciones son ciertas,*

$$F = \frac{(\mathrm{SCE}_0 - \mathrm{SCE})/q}{\mathrm{SCE}/(n-p)} \;\sim\; F_{q,\,n-p}.$$

Demostración. Es el teorema de Cochran. El vector y se descompone ortogonalmente en la proyección sobre el subespacio del modelo reducido, la componente en el complemento de aquél dentro del modelo completo —de dimensión q—, y el residuo —de dimensión $n-p$—. Bajo H_0 las dos últimas son sumas de cuadrados de normales independientes con esas dimensiones, luego χ^2_q y χ^2_{n-p} independientes, y el cociente de ambas dividido por sus grados de libertad es una F. □

Cuidado

[Con $q = 1$ es el t al cuadrado, y sólo entonces] Si se elimina una sola variable, $q = 1$ y se cumple $F = t^2$, donde t es el estadístico del coeficiente correspondiente: los dos contrastes coinciden. Con $q \geq 2$ *no* hay tal identidad, y mirar los t individuales uno por uno no sustituye al F conjunto. De hecho es frecuente que ningún coeficiente sea

significativo por separado y el F conjunto sí lo sea: ocurre cuando las variables están muy correlacionadas entre sí, y es el síntoma clásico de la colinealidad.

19.14. Ejercicios resueltos

Tres problemas sobre la misma recta, para separar el ajuste de la inferencia y ambos de la interpretación causal.

Ejercicio 19.1. Ajuste por mínimos cuadrados la recta a los datos $x = 1, \dots, 6$ con

$$y = 2{,}0;\ 4{,}1;\ 5{,}9;\ 8{,}2;\ 9{,}8;\ 12{,}1.$$

Dé $\hat\beta_0$, $\hat\beta_1$, R^2 y el error típico de la regresión.

Solución. Con $\bar x = 3{,}5$ y $\bar y = 7{,}0167$, las fórmulas $\hat\beta_1 = S_{xy}/S_{xx}$ y $\hat\beta_0 = \bar y - \hat\beta_1 \bar x$ dan $\hat\beta_1 = 1{,}9971$ y $\hat\beta_0 = 0{,}0267$.

La suma de cuadrados de los residuos es 0,1082 y, con $n - 2 = 4$ grados de libertad, el error típico de la regresión es $s = \sqrt{0{,}1082/4} = 0{,}1645$. El coeficiente de determinación es $R^2 = 0{,}9985$.

Merece la pena mirar los números con desconfianza sana. Un R^2 de 0,9985 es altísimo, pero era de esperar: los datos se generaron alrededor de una recta con muy poco ruido. En datos reales de ciencias sociales, un R^2 así suele indicar que se ha metido por error una variable que es prácticamente la misma que y. ■

Ejercicio 19.2. Con el ajuste anterior, contraste $H_0\colon \beta_1 = 0$ y dé un intervalo de confianza al 95 por ciento para la pendiente. Prediga el valor en $x = 7$ y comente el riesgo.

Solución. El error típico de la pendiente es $s/\sqrt{S_{xx}} = 0{,}1645/\sqrt{17{,}5} = 0{,}0393$, luego $t = 1{,}9971/0{,}0393 = 50{,}80$ con 4 grados de libertad: se rechaza H_0 de forma abrumadora. El intervalo al 95 por ciento es $1{,}9971 \pm 2{,}776 \cdot 0{,}0393 = (1{,}888;\ 2{,}106)$.

La predicción en $x = 7$ es $0{,}0267 + 1{,}9971 \cdot 7 = 14{,}01$. El riesgo es que $x = 7$ cae fuera del rango observado $[1, 6]$: extrapolar supone que la relación lineal sigue valiendo donde no se ha mirado, y eso es una hipótesis, no un resultado. Además el intervalo de predicción se ensancha al alejarse de $\bar x$, de modo que la incertidumbre real es mayor de lo que sugiere el número. ■

Ejercicio 19.3. En una regresión múltiple, al añadir una variable el R^2 sube de 0,62 a 0,64 pero el R^2 ajustado baja. Explique qué está ocurriendo y qué debe hacerse.

Solución. El R^2 ordinario nunca baja al añadir regresores: aunque la variable nueva sea ruido puro, el ajuste por mínimos cuadrados siempre encuentra algo que explicar con ella. Por eso no sirve para comparar modelos con distinto número de variables.

El R^2 ajustado penaliza la complejidad, con $\bar R^2 = 1 - (1 - R^2)\frac{n-1}{n-p-1}$. Que baje significa que la mejora de 0,02 es menor que lo que cabía esperar por azar al gastar un grado de libertad más. La conclusión es que la variable no aporta, y lo razonable es no incluirla, o comprobarlo con un criterio como el AIC o con validación cruzada. Añadir regresores hasta que el R^2 quede bonito es la manera más rápida de construir un modelo que funciona de maravilla con los datos que ya se tienen y falla con los siguientes. ■

19.15. Ejercicios propuestos

19.1. Con $S_{xx} = 40$, $S_{xy} = 100$, $\bar x = 5$ y $\bar y = 20$, halle la recta de regresión.

19.2. Si $r = 0{,}8$, ¿cuánto vale R^2 y cómo se interpreta?

19.3. ¿Puede ser $r = 0$ y haber una relación fuerte entre x e y? Dé un ejemplo.

19.4. Enumere los cuatro supuestos del modelo lineal clásico sobre los errores.

19.5. ¿Qué gráfico se usa para detectar heterocedasticidad y qué aspecto tiene el problema?

19.6. ¿Qué es la multicolinealidad, cómo se detecta con el VIF y por qué molesta?

19.7. Interprete el coeficiente de una variable indicadora que vale 1 para mujeres y 0 para hombres.

19.8. ¿Qué diferencia hay entre un intervalo de confianza para la media condicional y un intervalo de predicción para una observación nueva?

19.9. ¿Por qué el R^2 nunca baja al añadir variables y el ajustado sí?

19.10. Un residuo con gran distancia de Cook, ¿qué indica y qué debe hacerse?

19.11. ¿Qué hipótesis contrasta el estadístico F global de la regresión?

19.12. “La correlación no implica causalidad.” Dé un ejemplo de correlación alta por variable confusora y otro por causalidad inversa.

Capítulo 20

Estadística bayesiana

Lo que necesitas traer

Del capítulo 16: probabilidad condicionada y el teorema de Bayes. Del capítulo 18: verosimilitud. Del capítulo 6: integrales, porque toda la maquinaria bayesiana consiste en integrar densidades.

Toda la inferencia del capítulo anterior trataba al parámetro como una constante desconocida y a los datos como aleatorios. La estadística bayesiana invierte los papeles: los datos son lo que son, y lo que se describe con una distribución de probabilidad es el *conocimiento* sobre el parámetro. Ese cambio de punto de vista responde de forma directa a la pregunta que todo el mundo quiere hacer y que la inferencia clásica se niega a contestar: dados estos datos, ¿qué valores del parámetro son creíbles?

20.1. El teorema de Bayes como maquinaria de aprendizaje

Definición 20.1. Se parte de una *distribución a priori* $\pi(\theta)$ que codifica el conocimiento antes de ver los datos. Tras observar x, la *distribución a posteriori* es

$$\pi(\theta \mid x) = \frac{L(x \mid \theta)\,\pi(\theta)}{\int L(x \mid \theta')\pi(\theta')\,\mathrm{d}\theta'} \;\propto\; L(x \mid \theta)\,\pi(\theta).$$

Una familia de prioris es *conjugada* para una verosimilitud si la posteriori pertenece a la misma familia. Un *intervalo creíble* al 95 % es cualquier región que acumule esa probabilidad posteriori.

Idea clave

Qué dice esto en cristiano. La fórmula dice: creencia final $\propto$ verosimilitud $\times$ creencia inicial. Los datos no sustituyen a lo que se sabía, lo actualizan. Y aquí está la diferencia práctica con el intervalo de confianza: el intervalo creíble sí admite la lectura que todo el mundo hace, "hay un 95 % de probabilidad de que el parámetro esté aquí", porque el parámetro sí tiene distribución en este marco. El precio es que hay que elegir una

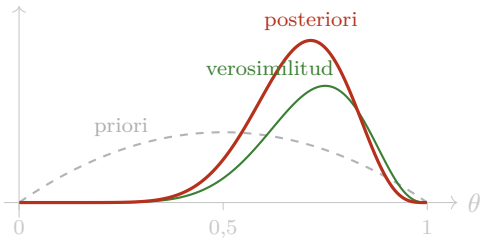


Figura 20.1: Cómo aprende el teorema de Bayes. Con priori Beta$(2,2)$ y 9 éxitos de 12 intentos, la posteriori es Beta$(11,5)$, con media 0,688 y moda 0,714. Queda entre lo que uno creía y lo que dicen los datos, más cerca de estos porque son doce y la priori pesa como dos.

priori, y esa elección es subjetiva; la defensa bayesiana es que la elección se declara explícitamente en vez de esconderse en el diseño.

Ejemplo 20.2 (Actualizar una proporción)**. (a) Priori uniforme.** Con priori Beta$(1,1)$, que es la uniforme en $[0,1]$, y 7 éxitos en 10 pruebas, la posteriori es Beta$(8,4)$. Su media es $\frac{8}{12} = 0{,}666667$ y su moda $\frac{7}{10} = 0{,}7$, que coincide con el estimador de máxima verosimilitud. El intervalo creíble al 95 % es $[0{,}390257;\ 0{,}890737]$.

(b) Priori informativa. Si antes había razones para creer que p ronda 0,5, se puede usar Beta$(20,20)$. Con los mismos datos la posteriori es Beta$(27,23)$, con media 0,54 e intervalo $[0{,}4023;\ 0{,}6747]$. Es más estrecho y está tirado hacia el 0,5 de la priori: diez observaciones no bastan para vencer una creencia previa equivalente a cuarenta.

(c) La conjugación, que es lo que hace factible el cálculo. La regla es Beta(a,b) más k éxitos de n da Beta$(a+k,\ b+n-k)$: actualizar consiste en sumar. Los parámetros de la priori se interpretan como observaciones ficticias previas, de ahí que Beta$(20,20)$ pese como cuarenta datos.

(d) Comparar hipótesis con el factor de Bayes. Para decidir entre $p = 0{,}5$ y $p = 0{,}7$ con 7 éxitos de 10: las verosimilitudes son 0,117187 y 0,266828, y su cociente es 2,276932. Los datos favorecen $p = 0{,}7$ por algo más de dos a uno, lo que en las escalas habituales se considera evidencia *débil*. Nótese la diferencia con un p-valor: aquí se compara una hipótesis con otra, no una con su negación.

Cuidado

[La priori importa, y con pocos datos importa mucho] Los apartados (a) y (b) usan los mismos datos y llegan a conclusiones distintas. Con muestras grandes la verosimilitud domina y la priori se vuelve irrelevante (es la convergencia bayesiana), pero con muestras pequeñas la elección de la priori *es* la conclusión. Por eso la práctica honesta exige declarar la priori y hacer un análisis de sensibilidad probando varias. Segundo aviso: las prioris llamadas "no informativas" no son inocuas; la uniforme en p no es uniforme en $\log\frac{p}{1-p}$, así que la ignorancia no tiene una representación única. Y tercero: fuera de las familias conjugadas la posteriori no tiene forma cerrada, y hay que recurrir a métodos de Monte Carlo con cadenas de Markov.

Pausa para pensar

El enfoque bayesiano es más antiguo que el clásico: Bayes escribió su ensayo antes de 1761 y Laplace lo desarrolló a fondo, usándolo para estimar la masa de Saturno y para calcular si nacen más niños que niñas. Durante el siglo XX quedó marginado, en buena medida por la hostilidad explícita de Fisher hacia la idea de asignar probabilidades a hipótesis, y solo la potencia de cálculo de los años noventa lo devolvió al centro: los métodos de Monte Carlo hicieron computable lo que antes era solo elegante. Hoy sostiene desde los filtros de correo basura hasta los sistemas de recomendación, y buena parte del debate metodológico sobre la crisis de replicación en ciencia consiste en discutir si el marco frecuentista de los p-valores era el adecuado desde el principio.

20.2. Los tres objetos del oficio: verosimilitud, distribución a priori y distribución a posteriori

La sección anterior ha enunciado la fórmula de Bayes y la ha usado para actualizar una proporción. En esa fórmula aparecen tres objetos —$L(x \mid \theta)$, $\pi(\theta)$ y $\pi(\theta \mid x)$— y una familia de distribuciones, la Beta(a, b), que el capítulo emplea una y otra vez sin haber escrito nunca su fórmula. Quien llegue aquí sin traer nada de casa se encuentra con un símbolo, Beta$(11, 5)$, del que sabe que tiene media 0,688 porque se lo han dicho. Esta sección pone las definiciones que faltan, demuestra las dos o tres cosas que hay que demostrar, y luego hace una actualización completa con números, desde el enunciado hasta el intervalo, sin saltarse ni un paso.

La verosimilitud: la misma fórmula leída al revés

Definición 20.3 (Modelo estadístico y función de verosimilitud)**.** Un *modelo estadístico paramétrico* es una familia de distribuciones $\{p(\cdot \mid \theta) : \theta \in \Theta\}$ sobre el conjunto de datos posibles, indexada por un *parámetro* θ que recorre el *espacio de parámetros* Θ. Observado un dato concreto x, la *función de verosimilitud* es la función

$$L(x \mid \theta) = p(x \mid \theta),$$

mirada como función de θ con x fijo. Es decir: la misma expresión que da la probabilidad del dato cuando el parámetro está fijo, pero recorriendo ahora los valores del parámetro con el dato ya observado y quieto.

Idea clave

Qué dice esto en cristiano. Una densidad y una verosimilitud son la misma fórmula con dos lecturas distintas. Si se fija θ y se deja correr x, la fórmula reparte probabilidad entre los datos posibles y suma uno. Si se fija x —el que de verdad se ha observado— y se deja correr θ, la fórmula mide *cuánto favorece cada valor del parámetro al dato que ha salido*, y ya no suma uno ni tiene por qué: no es una densidad en θ. Ese es exactamente el motivo por el que hace falta multiplicarla por una priori y volver a normalizar para obtener algo que sí sea una distribución sobre θ.

Ejemplo 20.4 (La verosimilitud binomial, evaluada)**.** Se lanzan $n = 20$ veces una chincheta y cae con la punta hacia arriba $k = 13$ veces. Si los lanzamientos son independientes y cada uno tiene probabilidad θ de caer así, la probabilidad del resultado observado es

$$L(x \mid \theta) = \binom{20}{13} \theta^{13}(1-\theta)^{7}.$$

Evaluándola en cuatro valores del parámetro:

θ	0,30	0,50	0,65	0,80
$L(x \mid \theta)$	0,001 018	0,073 929	0,184 401	0,054 550

El máximo está en $\theta = 13/20 = 0{,}65$, que es el estimador de máxima verosimilitud: el valor del parámetro que hace más probable lo que se ha visto. Nótese que los cuatro números no suman uno ni se acercan: la verosimilitud no reparte probabilidad entre los θ, solo los compara.

Observación 20.5 (Lo que la verosimilitud sí y no permite decir)*.* Que $L(x \mid 0{,}65)$ sea 2,5 veces mayor que $L(x \mid 0{,}5)$ significa que el dato observado es dos veces y media más probable bajo $\theta = 0{,}65$ que bajo $\theta = 0{,}5$. No significa que $\theta = 0{,}65$ sea dos veces y media más probable: para afirmar eso hace falta una distribución sobre θ, y para tenerla hace falta una priori. La estadística clásica se queda en la primera frase; la bayesiana da el paso a la segunda, y el precio de ese paso es exactamente la elección de $\pi(\theta)$.

Definición 20.6 (Las tres distribuciones y la evidencia)**.** Con la notación anterior:

- la *distribución a priori* $\pi(\theta)$ es una densidad (o función de masa) sobre Θ que describe la incertidumbre *antes* de ver los datos;
- la *evidencia* o *verosimilitud marginal* es el número $m(x) = \int_\Theta L(x \mid \theta)\,\pi(\theta)\,\mathrm{d}\theta$, que no depende de θ;
- la *distribución a posteriori* es $\pi(\theta \mid x) = L(x \mid \theta)\pi(\theta)/m(x)$, la incertidumbre *después*.

La evidencia es la constante que hace que la posteriori integre uno; su papel como criterio para comparar modelos se trata más adelante en el capítulo.

La distribución beta, que es la que se usa todo el rato

Definición 20.7 (Función beta y distribución beta)**.** Para $a, b > 0$, la *función beta* es

$$B(a,b) = \int_0^1 t^{a-1}(1-t)^{b-1}\,\mathrm{d}t = \frac{\Gamma(a)\Gamma(b)}{\Gamma(a+b)},$$

y la *distribución beta* $\mathrm{Beta}(a,b)$ es la distribución sobre el intervalo $[0,1]$ de densidad

$$f(\theta \mid a,b) = \frac{1}{B(a,b)}\,\theta^{a-1}(1-\theta)^{b-1}, \qquad 0 < \theta < 1.$$

Proposición 20.8 (Media, moda y varianza de una beta)**.** *Si* $\theta \sim \mathrm{Beta}(a,b)$ *entonces*

$$\mathrm{E}[\theta] = \frac{a}{a+b}, \qquad \mathrm{Var}(\theta) = \frac{ab}{(a+b)^2(a+b+1)},$$

y si $a>1$ *y* $b>1$ *la densidad tiene un único máximo, la* moda*, en* $(a-1)/(a+b-2)$*.*

Demostración. Para la media basta reconocer otra beta dentro de la integral:

$$\mathrm{E}[\theta] = \frac{1}{B(a,b)}\int_0^1 \theta^{a}(1-\theta)^{b-1}\mathrm{d}\theta = \frac{B(a+1,b)}{B(a,b)} = \frac{\Gamma(a+1)\Gamma(b)}{\Gamma(a+b+1)}\cdot\frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)} = \frac{a}{a+b},$$

donde se ha usado $\Gamma(z+1)=z\Gamma(z)$ dos veces. El mismo cálculo con θ^2 da $\mathrm{E}[\theta^2] = \frac{a(a+1)}{(a+b)(a+b+1)}$, y restando el cuadrado de la media sale la varianza anunciada. Para la moda, derívese el logaritmo de la densidad:

$$\frac{\mathrm{d}}{\mathrm{d}\theta}\Big[(a-1)\log\theta + (b-1)\log(1-\theta)\Big] = \frac{a-1}{\theta} - \frac{b-1}{1-\theta},$$

que se anula en $\theta=(a-1)/(a+b-2)$; como la segunda derivada es negativa en todo el intervalo cuando $a,b>1$, ese punto crítico es el máximo. □

Ejemplo 20.9 (Cuatro betas y lo que dicen)**. (a)** $\mathrm{Beta}(1,1)$**.** La densidad vale $1/B(1,1)=1$ en todo el intervalo: es la distribución uniforme en $[0,1]$. Media $1/2$, varianza $1/12 = 0{,}083\,333$.

(b) $\mathrm{Beta}(2,6)$**.** Media $2/8 = 0{,}25$, moda $1/6 = 0{,}166\,667$, desviación típica $\sqrt{12/(64\cdot 9)} = 0{,}144\,338$. Es una creencia previa que apuesta por valores bajos pero deja abierto casi todo el intervalo: su intervalo central al 95 % es $[0{,}036\,7;\ 0{,}578\,7]$.

(c) $\mathrm{Beta}(20,20)$**.** Media 0,5 y desviación típica 0,078 1: una creencia previa muy concentrada alrededor de la mitad.

(d) La regla de lectura. La suma $a+b$ mide la *fuerza* de la creencia y el cociente $a/(a+b)$ su *centro*. Cuanto mayor es $a+b$, más estrecha es la distribución: la varianza lleva $a+b+1$ en el denominador. Esa lectura se justifica sola en cuanto se vea, dos páginas más abajo, que observar un éxito suma uno a a y observar un fracaso suma uno a b.

La actualización beta–binomial, demostrada

Teorema 20.10 (Conjugación beta–binomial)**.** *Sea* $\theta \sim \mathrm{Beta}(a,b)$ *a priori y supóngase que, dado* θ*, se observan* k *éxitos en* n *pruebas independientes con probabilidad* θ *cada una. Entonces*

$$\theta \mid (k,n) \ \sim\ \mathrm{Beta}\big(a+k,\ b+n-k\big),$$

y la evidencia vale

$$m(k) = \binom{n}{k}\frac{B(a+k,\ b+n-k)}{B(a,b)}.$$

Demostración. Por la definición 20.6, la posteriori es proporcional al producto de verosimilitud por priori, y se pueden tirar todos los factores que no contengan θ:

$$\pi(\theta \mid k) \ \propto\ \underbrace{\binom{n}{k}\theta^k(1-\theta)^{n-k}}_{\text{verosimilitud}}\cdot\underbrace{\frac{\theta^{a-1}(1-\theta)^{b-1}}{B(a,b)}}_{\text{priori}} \ \propto\ \theta^{(a+k)-1}(1-\theta)^{(b+n-k)-1}.$$

El miembro derecho es exactamente el núcleo de una Beta$(a+k,\, b+n-k)$. Y como una densidad queda determinada por su núcleo —la constante es la que hace que integre uno—, la posteriori es esa beta y la constante que falta es $1/B(a+k, b+n-k)$. Para la evidencia se integra el producto sin tirar nada:

$$m(k) = \int_0^1 \binom{n}{k}\theta^k(1-\theta)^{n-k}\frac{\theta^{a-1}(1-\theta)^{b-1}}{B(a,b)}\mathrm{d}\theta = \frac{\binom{n}{k}}{B(a,b)}\int_0^1 \theta^{a+k-1}(1-\theta)^{b+n-k-1}\mathrm{d}\theta,$$

y la última integral es $B(a+k, b+n-k)$ por la definición 20.7. □

Observación 20.11 (Por qué esto se llama conjugación y por qué importa). La posteriori ha resultado ser de la misma familia que la priori, con los parámetros movidos. Eso es lo que significa que la beta sea *conjugada* de la binomial, y su valor práctico es doble: la integral del denominador —que en general es imposible— aquí se conoce en forma cerrada, y actualizar se reduce a dos sumas, $a \mapsto a+k$ y $b \mapsto b+n-k$. De paso queda clara la interpretación de los parámetros de la priori: $a-1$ éxitos y $b-1$ fracasos imaginarios, observados antes de empezar.

Ejemplo 20.12 (Una actualización completa, con todos los números)**.** Una fábrica prueba un procedimiento nuevo. La experiencia previa con procedimientos parecidos sugiere que la tasa de éxito ronda el 25 por ciento y desde luego no llega al 60; se traduce eso en una priori Beta$(2, 6)$. Se hacen $n = 20$ pruebas y salen $k = 13$ éxitos.

(a) Lo que dice la priori. Media $2/8 = 0{,}25$, y $a+b = 8$: la creencia previa pesa como ocho observaciones imaginarias. Su intervalo central al 95 % es $[0{,}036\,7;\ 0{,}578\,7]$, de modo que la probabilidad a priori de que θ supere 0,5 es $1 - F_{\mathrm{Beta}(2,6)}(0{,}5) = 0{,}062\,5$: bastante escéptica.

(b) Lo que dicen los datos. La verosimilitud es la del ejemplo 20.4, con máximo en $\hat\theta = 13/20 = 0{,}65$.

(c) La posteriori, por el teorema 20.10.

$$a' = a+k = 2+13 = 15, \qquad b' = b+n-k = 6+7 = 13, \qquad \text{luego} \qquad \theta \mid x \sim \mathrm{Beta}(15, 13).$$

(d) Sus resúmenes, por la proposición 20.8.

$$\mathrm{E}[\theta \mid x] = \frac{15}{28} = 0{,}535\,714, \qquad \text{moda} = \frac{14}{26} = 0{,}538\,462,$$

$$\mathrm{Var}(\theta \mid x) = \frac{15\cdot 13}{28^2\cdot 29} = 0{,}008\,577, \qquad \text{desviación típica} = 0{,}092\,611.$$

La incertidumbre se ha reducido mucho: la desviación típica ha pasado de 0,144 a 0,093.

(e) El intervalo de credibilidad central al 95 %. Son los cuantiles 0,025 y 0,975 de la Beta$(15, 13)$:

$$[\,0{,}353\,3;\ 0{,}713\,3\,].$$

(f) La pregunta que interesa de verdad. ¿Es el procedimiento mejor que echarlo a suertes?

$$P(\theta > 0{,}5 \mid x) = 1 - F_{\mathrm{Beta}(15,13)}(0{,}5) = 0{,}649\,4.$$

Ha subido del 6,25 por ciento previo al 64,9 por ciento, pero sigue lejos de ser concluyente: con veinte pruebas y una priori escéptica, lo honesto es decir que hay indicios y no certeza.

(g) El tira y afloja entre priori y datos. La media posterior 0,535 7 queda entre la priori (0,25) y el dato (0,65). Y no queda en cualquier sitio: se comprueba que

$$\frac{a+k}{a+b+n} = \frac{a+b}{a+b+n}\cdot\frac{a}{a+b} + \frac{n}{a+b+n}\cdot\frac{k}{n} = \frac{8}{28}\cdot 0{,}25 + \frac{20}{28}\cdot 0{,}65 = 0{,}535\,714,$$

es decir, la media posterior es *exactamente* el promedio ponderado de la media a priori y de la proporción observada, con pesos proporcionales al número de observaciones —ocho imaginarias frente a veinte reales—. Con doscientas pruebas en vez de veinte, el peso de la priori bajaría a $8/208 = 0{,}038$ y la posteriori sería prácticamente el dato.

(h) Actualizar de golpe o por partes da lo mismo. Supóngase que las veinte pruebas se hicieron en dos tandas de diez, con 6 éxitos la primera y 7 la segunda. Tras la primera tanda la posteriori es $\text{Beta}(2+6,\,6+4) = \text{Beta}(8, 10)$, de media 0,444 4; usándola como priori para la segunda tanda queda $\text{Beta}(8+7,\,10+3) = \text{Beta}(15, 13)$, que es justo lo obtenido en (c). Esa coherencia no es una casualidad de este ejemplo: es una propiedad general del teorema de Bayes, y es la que permite ir incorporando datos según llegan sin rehacer nada.

(i) La evidencia. Por el teorema 20.10,

$$m(13) = \binom{20}{13}\frac{B(15,13)}{B(2,6)} = 0{,}012\,486.$$

Comparada con la del modelo rival que fija $\theta = 0{,}5$, cuya "evidencia" es sencillamente $\binom{20}{13}0{,}5^{20} = 0{,}073\,929$, el factor de Bayes a favor del modelo simple resulta $0{,}073\,929/0{,}012\,486 = 5{,}92$. Con la priori uniforme $\text{Beta}(1,1)$ en lugar de la escéptica, la evidencia sería $1/21 = 0{,}047\,619$ y el factor de Bayes bajaría a 1,55. Los datos son los mismos y el veredicto cambia de "evidencia moderada a favor de la moneda justa" a "prácticamente nada": el factor de Bayes depende de la priori de una manera que la estimación puntual no.

Regla: Hacer una actualización conjugada sin equivocarse

1. **Escribe la verosimilitud como función de** θ y tira todos los factores que no dependan de θ: en el caso binomial, el $\binom{n}{k}$ sobra desde el principio.
2. **Traduce la información previa a parámetros.** Con la beta: elige el centro con $a/(a+b)$ y la fuerza con $a+b$, que es el número de observaciones imaginarias que estás metiendo en la balanza. Si no sabes cuántas, ponlas pocas.
3. **Suma**: $a \mapsto a+k$, $b \mapsto b+n-k$. Comprueba que $a'+b' = a+b+n$; si no cuadra, has contado mal los fracasos.
4. **Calcula los resúmenes con las fórmulas de la familia**, no volviendo a integrar: media $a'/(a'+b')$, moda $(a'-1)/(a'+b'-2)$, y los cuantiles con la función de distribución de la beta.
5. **Contesta la pregunta original en forma de probabilidad posterior**: $P(\theta > c \mid x)$, $P(\theta_1 > \theta_2 \mid x)$, lo que sea. Es lo único que el marco clásico no puede dar y es la razón de estar aquí.
6. **Repite con otra priori** razonablemente distinta. Si el resultado no se mueve, dilo;

si se mueve, esa dependencia *es* parte del resultado.

Resumir la posteriori: MAP, media y credibilidad

Definición 20.13 (Estimadores puntuales bayesianos)**.** Sea $\pi(\theta \mid x)$ la posteriori. Se llaman:

- *media a posteriori* a $\mathrm{E}[\theta \mid x] = \int \theta\, \pi(\theta \mid x)\, \mathrm{d}\theta$;
- *mediana a posteriori* al valor que deja probabilidad posterior $1/2$ a cada lado;
- *estimador de máxima a posteriori* o *MAP* al maximizador $\hat{\theta}_{\mathrm{MAP}} = \arg\max_\theta \pi(\theta \mid x)$, es decir, a la moda de la posteriori.

Observación 20.14 (El MAP y la máxima verosimilitud)*.* Como $\pi(\theta \mid x) \propto L(x \mid \theta)\pi(\theta)$, maximizar la posteriori es maximizar el producto de verosimilitud por priori. Si la priori es constante, ambas maximizaciones coinciden y el MAP *es* el estimador de máxima verosimilitud: la máxima verosimilitud es el caso particular bayesiano de priori uniforme, mirando solo el máximo. En el ejemplo 20.12 el MAP vale 0,538 5 y el de máxima verosimilitud 0,65: la diferencia es el tirón de la priori. Y obsérvese que el MAP, por ser un máximo de una densidad, *no* es invariante frente a reparametrizaciones —el máximo de la densidad de $\log\theta$ no está en el logaritmo del máximo, porque el cambio de variable introduce un jacobiano—, mientras que la media y la mediana no sufren ese problema del mismo modo. Cuál usar depende de qué se pierda al equivocarse, y eso se resuelve más adelante con la teoría de la decisión.

Definición 20.15 (Intervalo de credibilidad)**.** Un *intervalo de credibilidad* al nivel $1-\alpha$ para θ es cualquier conjunto $C \subset \Theta$ con

$$P(\theta \in C \mid x) = \int_C \pi(\theta \mid x)\, \mathrm{d}\theta = 1 - \alpha.$$

El *intervalo central* o *de colas iguales* es el que deja probabilidad posterior $\alpha/2$ a cada lado, es decir, el que va del cuantil $\alpha/2$ al cuantil $1-\alpha/2$ de la posteriori.

Cuidado

[Credibilidad no es confianza, y esta es la diferencia exacta] Con los datos del ejemplo 20.12 —13 éxitos de 20— se pueden calcular cinco intervalos al 95 %:

intervalo	extremos
credibilidad, priori Beta$(2,6)$	[0,353 3; 0,713 3]
credibilidad, priori uniforme	[0,430 3; 0,818 9]
confianza de Wald, $\hat{p} \pm 1{,}96\sqrt{\hat{p}(1-\hat{p})/n}$	[0,441 0; 0,859 0]
confianza de Wilson	[0,432 9; 0,818 8]
confianza exacta de Clopper–Pearson	[0,407 8; 0,846 1]

Los números se parecen; las frases que los acompañan, no. Del primero se puede decir: *dada esta priori y estos datos, la probabilidad de que* θ *esté entre* 0,353 *y* 0,713 *es* 0,95.

Del tercero *no* se puede decir nada parecido: ahí θ es una constante desconocida, el intervalo es lo aleatorio, y la afirmación correcta es que el *procedimiento*, repetido en muestras hipotéticas, acierta el 95 % de las veces. Sobre el intervalo concreto que ha salido, el frecuentista no puede poner ninguna probabilidad, porque θ o está o no está. Dos observaciones que impiden convertir esto en una guerra de banderas. La primera: los dos enfoques convergen. El intervalo de credibilidad con priori uniforme, $[0{,}430\,3;\ 0{,}818\,9]$, y el de Wilson, $[0{,}432\,9;\ 0{,}818\,8]$, coinciden en dos decimales sin haber sido diseñados juntos. Es más, la cobertura frecuentista del intervalo de credibilidad con priori uniforme —calculada sumando las probabilidades binomiales de los k para los que el intervalo cubre al verdadero θ— vale 0,959 cuando $\theta = 0{,}5$ y 0,968 cuando $\theta = 0{,}65$: es un intervalo de confianza perfectamente decente sin haberlo pretendido. La segunda: la diferencia sí es real cuando la priori aporta información, como en la primera fila, donde el intervalo está claramente desplazado hacia abajo respecto de todos los demás. Ahí la interpretación probabilista está condicionada a la priori, y quien no acepte la priori no acepta el intervalo. Declararla es obligatorio, no cortés.

Definición 20.16 (Prioris no informativas y débilmente informativas)**.** Una priori se llama *no informativa* u *objetiva* cuando se elige por una regla formal que pretende no aportar información sobre θ, y no por conocimiento sustantivo: la uniforme sobre un parámetro acotado, $\pi(\theta) \propto 1$; la de Haldane Beta$(0,0)$; o la de Jeffreys $\pi_J(\theta) \propto \sqrt{I(\theta)}$, que se trata más adelante. Se llama *débilmente informativa* a la que descarta lo absurdo sin comprometerse con nada dentro de lo plausible, como una Beta$(1,1)$ o una normal muy ancha.

Observación 20.17 (La ignorancia no tiene una representación única)*.* La uniforme parece la traducción evidente de "no sé nada", y no lo es: si θ es uniforme en $[0,1]$, entonces θ^2 no lo es, ni tampoco $\log \frac{\theta}{1-\theta}$. Ser ignorante sobre θ y serlo sobre una función de θ son cosas distintas, de modo que "no informativa" es siempre relativa a una parametrización. Esa es precisamente la incomodidad que la priori de Jeffreys resuelve, al construirse de forma que se transforma correctamente al cambiar de parámetro.

Proposición 20.18 (La predictiva de un solo dato)**.** *Con priori* Beta(a,b) *y antes de observar nada, la probabilidad de que la primera prueba sea un éxito es* $a/(a+b)$*; y tras haber observado* k *éxitos en* n *pruebas, la probabilidad de que la siguiente lo sea es*

$$P(\tilde{x} = 1 \mid x) = \frac{a+k}{a+b+n}.$$

Demostración. En ambos casos se integra el parámetro contra la distribución que corresponda: $P(\tilde{x}=1) = \int_0^1 \theta\,\pi(\theta)\mathrm{d}\theta = \mathrm{E}[\theta]$, que es $a/(a+b)$ para la priori y $(a+k)/(a+b+n)$ para la posteriori Beta$(a+k, b+n-k)$, por la proposición 20.8. □

Observación 20.19 (La regla de sucesión de Laplace)*.* Con priori uniforme, $a = b = 1$, la fórmula anterior es $(k+1)/(n+2)$: la célebre regla de sucesión que Laplace usó para preguntarse por la probabilidad de que salga el sol mañana. Su virtud práctica es que nunca da 0 ni 1. Si en diez pruebas ha habido diez éxitos, la máxima verosimilitud dice que la probabilidad

de éxito es exactamente 1 —una afirmación insostenible con diez datos—, mientras que la predictiva bayesiana da $11/12 = 0{,}916\,7$, que ya es una respuesta con la que se puede trabajar. La versión general de esto, integrar el parámetro en vez de sustituirlo por su estimación, es la distribución predictiva que se estudia más adelante.

Regla: Informar de un análisis bayesiano en cinco líneas

1. **La priori, con su justificación**, y su equivalente en datos imaginarios si es conjugada. Sin esto lo demás no es interpretable.
2. **Los datos y el modelo de verosimilitud**, incluida la hipótesis de independencia, que es la que se cuela sin avisar.
3. **La posteriori entera** si es de una familia conocida: decir Beta$(15, 13)$ es más informativo que dar tres resúmenes.
4. **Un intervalo de credibilidad**, diciendo si es central o de máxima densidad, y **la probabilidad posterior de la hipótesis que interesa**, que suele ser lo que de verdad preguntaba quien encargó el análisis.
5. **El análisis de sensibilidad**: qué pasa con otra priori razonable. Una línea basta.

Observación 20.20 (Lo que viene después). Todo lo anterior ha sido posible porque la beta y la binomial encajan. Las secciones siguientes hacen tres cosas con eso: enseñan las otras dos conjugaciones que se usan en la práctica —normal con normal y gamma con Poisson—, explican qué hacer cuando no hay conjugada, que es lo habitual, y vuelven sobre la evidencia $m(x)$ para comparar modelos enteros en lugar de estimar parámetros. La estructura, sin embargo, no cambia nunca: verosimilitud por priori, normalizar, y leer la posteriori.

20.3. Otras conjugaciones y la elección de la priori

Teorema 20.21 (Modelo normal–normal)**.** *Si la priori de μ es $N(\mu_0, \tau^2)$ y se observan n datos con media $\bar{x}$ de una $N(\mu, \sigma^2)$ con σ^2 conocida, la posteriori es normal con*

$$\frac{1}{\tau^2_{post}} = \frac{1}{\tau^2} + \frac{n}{\sigma^2}, \qquad \mu_{post} = \tau^2_{post}\left(\frac{\mu_0}{\tau^2} + \frac{n\bar{x}}{\sigma^2}\right).$$

Idea clave

Qué dice esto en cristiano. Las *precisiones* (inversos de las varianzas) se suman, y la media posteriori es el promedio ponderado de la priori y de los datos, con pesos proporcionales a sus precisiones. Es exactamente cómo combinaría dos mediciones cualquier ingeniero: se hace más caso a la más precisa. Y como la precisión de los datos crece con n, al aumentar la muestra el peso de la priori se diluye: con muchos datos, el bayesiano y el frecuentista acaban en el mismo sitio, y la discusión sobre la priori solo importa cuando los datos son escasos.

Ejemplo 20.22 (Combinar creencia previa y evidencia)**. (a) Los números.** Priori $N(100, 25)$, es decir $\mu_0 = 100$ y $\tau = 5$; se observan $n = 16$ datos con $\bar{x} = 106$ y $\sigma^2 = 100$. Las precisiones son $\frac{1}{25} = 0{,}04$ para la priori y $\frac{16}{100} = 0{,}16$ para los datos.

(b) La posteriori. La varianza posteriori es $\frac{1}{0{,}04+0{,}16} = 5$, con desviación 2,236068, y la media $5 \cdot (4 + 16{,}96) = 104{,}8$. El intervalo creíble al 95 % es [100,417387; 109,182613].

(c) Los pesos, explícitos. Los datos pesan $\frac{0{,}16}{0{,}20} = 0{,}8$ y la priori 0,2, de donde $104{,}8 = 0{,}8 \cdot 106 + 0{,}2 \cdot 100$. La media posteriori está más cerca de los datos porque su precisión es cuatro veces mayor. Ese tirón hacia la priori se llama *contracción* y es lo que hace robustas a las estimaciones bayesianas con muestras pequeñas.

(d) Otra conjugación útil. Para datos de Poisson, la priori Gamma(a, b) da posteriori Gamma$(a + \sum x_i,\ b + n)$. Con $a = 2$, $b = 1$ y diez observaciones que suman 23, la posteriori es Gamma$(25, 11)$, de media $\frac{25}{11} = 2{,}272727$, frente al estimador clásico $\bar{x} = 2{,}3$. Otra vez, la priori tira ligeramente del resultado.

Cuidado

[Contracción no es sesgo indeseable] Que la estimación se desplace hacia la priori parece un defecto y con frecuencia es una virtud: en problemas con muchos parámetros y pocos datos por parámetro, los estimadores contraídos tienen *menor* error cuadrático medio que los clásicos. Ese es el contenido del célebre resultado de Stein de 1956, según el cual en dimensión tres o más el estimador de máxima verosimilitud de un vector de medias normales es inadmisible: siempre existe otro que lo mejora en todo punto. La contracción bayesiana es la explicación natural de ese fenómeno.

20.4. Cuando no hay conjugada: muestrear la posteriori

Las familias conjugadas son cómodas y escasas. En cuanto el modelo se complica un poco —una regresión con varios parámetros, una jerarquía de grupos, una verosimilitud que no es de la familia exponencial— la posteriori deja de tener forma cerrada y la integral del denominador de Bayes se vuelve imposible. La salida no es calcularla: es *fabricar muestras* de la posteriori sin conocerla, y estimar con ellas todo lo que haga falta.

Definición 20.23. Un algoritmo de *Monte Carlo por cadenas de Markov* construye una cadena $\theta_0, \theta_1, \theta_2, \ldots$ cuya distribución estacionaria es la posteriori $\pi(\theta \mid x)$. El *algoritmo de Metropolis–Hastings* propone θ' desde una densidad $q(\theta' \mid \theta_t)$ y lo acepta con probabilidad

$$\alpha = \min\Big\{1,\ \frac{\pi(\theta' \mid x)\, q(\theta_t \mid \theta')}{\pi(\theta_t \mid x)\, q(\theta' \mid \theta_t)}\Big\};$$

si lo rechaza, repite $\theta_{t+1} = \theta_t$.

Idea clave

Qué dice esto en cristiano. Lo maravilloso de esa fórmula es lo que *no* aparece en ella: la constante de normalización. Como la posteriori entra dos veces, arriba y abajo,

la integral imposible se cancela y basta con saber evaluar el producto verosimilitud por priori, que siempre se sabe. El algoritmo es un paseo aleatorio con sesgo: propone un salto, y si el destino es más probable lo acepta siempre, mientras que si es menos probable lo acepta a veces, con probabilidad igual al cociente. Ese "a veces" es lo que impide que la cadena se quede atrapada en la moda y lo que hace que, a la larga, visite cada región exactamente en proporción a su probabilidad.

Regla: Ejecutar un muestreador y no engañarse

1. **Trabajar en escala logarítmica**: comparar $\log \pi(\theta') - \log \pi(\theta_t)$ en vez de cocientes, porque las verosimilitudes se van a cero y desbordan el punto flotante.
2. **Elegir la anchura de la propuesta**: demasiado estrecha y la cadena avanza a paso de tortuga; demasiado ancha y casi todo se rechaza. La tasa de aceptación de referencia es alrededor de 0,44 en una dimensión y 0,23 en muchas.
3. **Descartar el calentamiento**: los primeros cientos o miles de estados dependen del punto inicial y no representan a la posteriori.
4. **Correr varias cadenas desde puntos distintos** y comparar. Una sola cadena que parece estable puede estar atrapada en una moda.
5. **Medir la convergencia** con el estadístico $\hat{R}$ de Gelman–Rubin, que compara la varianza entre cadenas con la varianza dentro de cada una; valores por encima de 1,01 piden más iteraciones.
6. **Contar el tamaño efectivo de muestra**: las iteraciones están correlacionadas, así que diez mil estados pueden valer lo que doscientos independientes.
7. **Mirar las trazas**. Una cadena sana parece ruido alrededor de un nivel; una enferma tiene mesetas largas, tendencias o saltos entre dos niveles.

Ejemplo 20.24 (Dos muestreadores, comprobados)**. (a) Metropolis para una normal.** Objetivo $N(0,1)$, propuesta uniforme en $(\theta-1, \theta+1)$, que es simétrica y por tanto simplifica α a $\min\{1, \mathrm{e}^{-(\theta'^2-\theta_t^2)/2}\}$. Con veinte mil iteraciones y dos mil de calentamiento, la media muestral sale $-0{,}071708$ y la varianza 0,997528, frente a los valores teóricos 0 y 1. La cadena reproduce la distribución sin haber usado nunca la constante $1/\sqrt{2\pi}$.

(b) Gibbs para una normal bivariante. Con correlación $\rho = 0{,}8$, las condicionales son $\theta_1 \mid \theta_2 \sim N(\rho\theta_2,\ 1-\rho^2)$ y simétricamente. Alternando extracciones de una y otra durante veinte mil pasos, la correlación estimada resulta 0,803801. Gibbs no rechaza nunca: es el caso particular de Metropolis–Hastings en que la propuesta es la condicional exacta y la probabilidad de aceptación vale siempre 1.

(c) Qué se estima con las muestras. Cualquier cosa. La media posterior es el promedio de la muestra; un intervalo creíble al 95 % son los percentiles 2,5 y 97,5; la probabilidad de que $\theta > 0$ es la proporción de estados positivos. Ninguna requiere integrar nada: se cuenta.

(d) Cuándo Gibbs se atasca. Con $\rho = 0{,}99$ las condicionales son estrechísimas y la

cadena avanza en zigzag por una cresta diagonal, moviéndose una distancia del orden de $\sqrt{1-\rho^2} = 0{,}141$ por paso. Necesita cientos de veces más iteraciones que con $\rho = 0$. El remedio habitual es reparametrizar para descorrelacionar, o usar métodos que aprovechan el gradiente, como el Monte Carlo hamiltoniano.

Cuidado

[Contraejemplo: converger no es haber convergido] Una cadena de Metropolis sobre una posteriori bimodal con las modas muy separadas puede pasar cien mil iteraciones dentro de una sola de ellas, con trazas de aspecto impecable y $\hat{R}$ prácticamente igual a 1 si todas las cadenas arrancaron cerca del mismo sitio. Los diagnósticos detectan la no convergencia, nunca la certifican: sirven para descartar cadenas malas, no para aprobar las buenas. Por eso se arranca desde puntos dispersos y por eso conviene saber de antemano si el modelo puede tener varias modas.

Cuidado

[Contraejemplo: una priori impropia puede dar una posteriori impropia] Es tentador usar prioris "no informativas" del tipo $\pi(\sigma) \propto 1/\sigma$, que no integran a uno. A veces funciona y la posteriori sí es propia; a veces no. En el modelo jerárquico normal con priori $\pi(\tau) \propto 1/\tau$ sobre la desviación entre grupos, la posteriori es impropia y no existe. Lo grave es que el muestreador *no se entera*: produce números, las trazas parecen razonables y las estimaciones derivan lentamente hacia cero sin que nada avise. La única defensa es comprobar analíticamente que la posteriori es propia antes de muestrear.

Pausa para pensar

El algoritmo apareció en 1953 en un artículo firmado por Metropolis, los Rosenbluth y los Teller, escrito para simular un gas en Los Álamos con la computadora MANIAC. Hastings lo generalizó en 1970 y luego el resultado durmió veinte años, hasta que Gelfand y Smith mostraron en 1990 que servía para hacer estadística bayesiana. Marshall Rosenbluth contó décadas después que Metropolis, que aparecía primero por orden alfabético, no había contribuido gran cosa. La estadística bayesiana llevaba dos siglos siendo elegante e impracticable; lo que la volvió utilizable no fue un teorema, sino una máquina y un truco de simulación prestado de la física.

20.5. Comparar modelos: evidencia, factores de Bayes y la paradoja de Lindley

Estimar un parámetro es una cosa y decidir entre dos hipótesis es otra. El aparato bayesiano para lo segundo es elegante, tiene una interpretación limpia y produce, en un caso concreto, un desacuerdo frontal con el contraste clásico que conviene entender bien.

Definición 20.25. La *evidencia* o verosimilitud marginal de un modelo M es $P(D \mid M) =$

$\int P(D \mid \theta, M)\,\pi(\theta \mid M)\,\mathrm{d}\theta$, y el *factor de Bayes* entre dos modelos es el cociente

$$B_{10} = \frac{P(D \mid M_1)}{P(D \mid M_0)}, \qquad \text{de modo que} \qquad \frac{P(M_1 \mid D)}{P(M_0 \mid D)} = B_{10} \cdot \frac{P(M_1)}{P(M_0)}.$$

Idea clave

Qué dice esto en cristiano. El factor de Bayes es el número por el que hay que multiplicar las creencias previas para obtener las posteriores. Su virtud es que separa limpiamente los datos de la opinión: B_{10} sólo depende de lo observado y del modelo, y la opinión entra aparte. Y trae incorporada una navaja de Ockham automática: la evidencia es un promedio de la verosimilitud sobre *toda* la priori, de modo que un modelo con muchos parámetros libres reparte su masa de probabilidad entre muchas predicciones y acierta poco en cada una. Nadie penaliza la complejidad a mano; sale sola de integrar.

Ejemplo 20.26 (La paradoja de Lindley)**. (a) El montaje.** Se contrasta $H_0\colon \mu = 0$ frente a $H_1\colon \mu \neq 0$ con σ conocida y una muestra enorme, $n = 100\,000$. Sale $\bar{x}$ tal que el estadístico vale exactamente $z = 1{,}96$, es decir, $p = 0{,}05$.

(b) Las dos respuestas. El contraste clásico rechaza H_0: el p-valor está justo en el umbral. El análisis bayesiano con una priori difusa sobre μ bajo H_1 da un factor de Bayes a favor de H_0 del orden de $\sqrt{n}\,\mathrm{e}^{-z^2/2} \approx 46$: evidencia *fuerte* a favor de la hipótesis nula. Los dos métodos no discrepan un poco; discrepan por completo.

(c) Quién tiene razón. Los dos, porque responden a preguntas distintas. El p-valor pregunta si los datos son raros bajo H_0, y con n enorme una desviación diminuta ya lo es. El factor de Bayes pregunta qué modelo predijo mejor lo observado, y H_1, al repartir su probabilidad sobre todos los valores posibles de μ, apostó poquísimo por el valor minúsculo que realmente se observó. Con muestras gigantescas, un efecto estadísticamente significativo puede ser evidencia a favor de que el efecto es nulo.

(d) La lección práctica. El factor de Bayes depende de la priori de forma que no desaparece al crecer n, a diferencia de la estimación puntual. Duplicar la anchura de la priori divide el factor de Bayes por dos, y con prioris impropias directamente no está definido. Por eso en contraste de hipótesis la elección de la priori sí es una decisión sustantiva que hay que declarar y justificar.

Cuidado

[Contraejemplo: la evidencia no es la verosimilitud maximizada] Confundir $P(D \mid M)$ con $\max_\theta P(D \mid \theta, M)$ es el error más común, y anula la navaja de Ockham. La verosimilitud maximizada nunca baja al añadir parámetros, exactamente como el R^2; la evidencia sí baja, porque promedia en vez de maximizar. Un modelo con veinte parámetros ajustará mejor cualquier conjunto de datos y tendrá menos evidencia, y esa distinción es la razón entera de que el criterio bayesiano funcione para seleccionar modelos.

20.6. Lo que de verdad se quiere saber: la distribución predictiva

Un cliente no pregunta por θ. Pregunta por lo que va a pasar. La respuesta bayesiana a esa pregunta tiene nombre propio y es una de las cosas que el enfoque hace mejor.

Definición 20.27 (Predictivas a priori y a posteriori)**.**

$$p(\tilde{x}) = \int p(\tilde{x} \mid \theta)\, \pi(\theta)\, \mathrm{d}\theta \qquad \text{(a priori)},$$

$$p(\tilde{x} \mid x) = \int p(\tilde{x} \mid \theta)\, \pi(\theta \mid x)\, \mathrm{d}\theta \qquad \text{(a posteriori)}.$$

En ambas, el parámetro se integra: no se elige un valor, se promedian todos con su credibilidad.

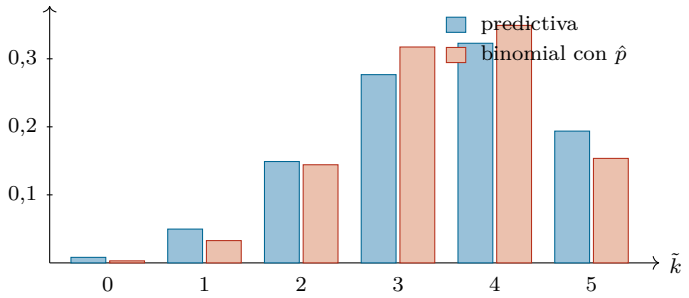


Figura 20.2: Éxitos en los próximos cinco intentos, calculados de dos maneras. Las dos tienen la misma media, 3,4375, pero la predictiva —que integra la incertidumbre sobre p en lugar de fijarla en $\hat{p}$— tiene un 24 por ciento más de varianza, y la diferencia está toda en las colas: asigna al "ningún éxito" casi el triple de probabilidad.

Ejemplo 20.28 (Cinco intentos más, con y sin honestidad sobre la incertidumbre)**.** Partimos de la posteriori Beta$(11, 5)$ obtenida antes, cuya media es $\hat{p} = 11/16 = 0{,}6875$. ¿Cuántos éxitos habrá en los próximos $m = 5$ intentos?

(a) El atajo incorrecto. Enchufar $\hat{p}$ y usar Bin$(5; 0{,}6875)$.

(b) La predictiva correcta. Integrando, sale una beta-binomial:

$$p(\tilde{k}) = \binom{5}{\tilde{k}} \frac{B(11 + \tilde{k},\ 5 + 5 - \tilde{k})}{B(11, 5)}.$$

(c) Las dos, lado a lado.

$\tilde{k}$	0	1	2	3	4	5
predictiva	0,0081	0,0497	0,1490	0,2767	0,3228	0,1937
binomial con $\hat{p}$	0,0030	0,0328	0,1442	0,3173	0,3491	0,1536

Las dos tienen la misma media, $5 \cdot 0{,}6875 = 3{,}4375$. Pero sus varianzas no coinciden:

$$\mathrm{Var}_{\text{predictiva}} = 1{,}327\,0, \qquad \mathrm{Var}_{\text{binomial}} = 1{,}074\,2.$$

La predictiva es un 24 por ciento más dispersa. Mírese dónde está la diferencia: en las colas. La predictiva asigna a "ningún éxito" una probabilidad de 0,0081, casi el triple que el atajo. Y esa es exactamente la clase de suceso que arruina a quien lo subestima.

Idea clave

Enchufar la estimación puntual finge que θ se conoce con exactitud. La predictiva reconoce que no, y paga el reconocimiento ensanchando las colas. La regla general: *integrar siempre da más incertidumbre que sustituir*, y la diferencia crece cuanto menos datos haya. Con $n = 10\,000$ observaciones las dos curvas serían indistinguibles; con las 14 de este ejemplo, no lo son.

En los casos conjugados la predictiva tiene forma cerrada y conocida: beta-binomial en el caso binomial, binomial negativa en el caso Poisson-gamma, t de Student en el caso normal con varianza desconocida. Que la t aparezca aquí de forma natural, como predictiva de una normal con incertidumbre sobre σ, es una de las coincidencias más bonitas entre los dos enfoques.

Observación 20.29 (Comprobación predictiva posterior). La predictiva también sirve para juzgar el modelo. Se simulan conjuntos de datos completos desde $p(\tilde{x} \mid x)$ y se compara alguna característica de los datos reales —el máximo, el número de ceros, la asimetría— con su distribución simulada. Si el dato real cae en el 0,1 por ciento de la cola, el modelo no sabe generar datos como los observados, y eso es un defecto de ajuste que ninguna comparación entre modelos habría detectado. El p-valor predictivo posterior es la proporción de réplicas más extremas que lo observado.

20.7. Resumir la posteriori: teoría de la decisión, estimadores de Bayes e intervalos HPD

La posteriori es la respuesta completa, pero nadie publica una función. Hay que resumirla, y la pregunta "¿media, mediana o moda?" tiene respuesta exacta: depende de qué se pierda al equivocarse.

Teorema 20.30 (El estimador de Bayes según la pérdida)**.** *El estimador que minimiza la pérdida esperada a posteriori es*

$$\begin{aligned} L(\theta, a) = (\theta - a)^2 \quad &\Longrightarrow \quad a^* = E[\theta \mid x] \quad \textit{(media)}, \\ L(\theta, a) = |\theta - a| \quad &\Longrightarrow \quad a^* = \textit{mediana de } \pi(\theta \mid x), \\ L(\theta, a) = \mathbf{1}\{|\theta - a| > \epsilon\} \quad &\Longrightarrow \quad a^* \to \textit{moda (MAP) cuando } \epsilon \to 0. \end{aligned}$$

La demostración del primer caso es la de siempre: derivar $E[(\theta - a)^2 \mid x]$ respecto de a e igualar a cero da $a = E[\theta \mid x]$. El segundo sale de que la derivada de $E|\theta - a|$ es $P(\theta < a) - P(\theta > a)$, nula en la mediana.

Definición 20.31 (Intervalo HPD)**.** Una región de **máxima densidad posterior** al nivel $1 - \alpha$ es el conjunto $\{\theta : \pi(\theta \mid x) \geq c\}$ con c elegido para que acumule probabilidad $1 - \alpha$.

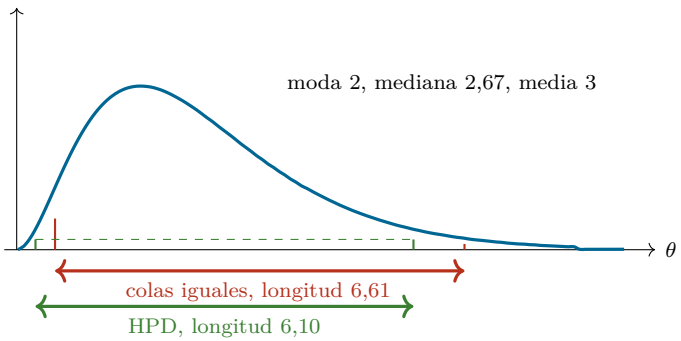


Figura 20.3: Una posteriori Gamma(3, 1) resumida de dos maneras. El HPD corta la densidad a una altura constante —la línea verde discontinua— y por eso es el más corto de todos los intervalos con esa probabilidad. El de colas iguales incluye puntos menos creíbles que otros que deja fuera, pero es invariante frente a reparametrizaciones.

Entre todas las regiones de esa probabilidad, es la de menor longitud, y todo punto de dentro tiene más densidad que cualquier punto de fuera.

Ejemplo 20.32 (Una posteriori asimétrica resumida de cinco maneras)**.** Sea la posteriori Gamma(3, 1), con densidad $\propto \theta^2 e^{-\theta}$.

(a) Los tres centros.

$$\text{media} = \frac{\alpha}{\beta} = 3, \qquad \text{mediana} = 2{,}674\,1, \qquad \text{moda} = \frac{\alpha - 1}{\beta} = 2.$$

Difieren en un 50 por ciento entre los extremos. Con una posteriori simétrica coincidirían las tres y la pregunta no tendría interés; con una asimétrica, elegir es inevitable.

(b) Los dos intervalos al 95 por ciento.

$$\text{colas iguales: } [0{,}618\,7;\ 7{,}224\,7], \quad \text{longitud } 6{,}606\,0;$$

$$\text{HPD: } [0{,}303\,5;\ 6{,}401\,2], \quad \text{longitud } 6{,}097\,7.$$

El HPD es un 8 por ciento más corto. Y se puede comprobar que es el correcto: la densidad en sus dos extremos vale 0,033 99 y 0,034 00, es decir, la misma, que es justo la condición que define el HPD. En el de colas iguales, en cambio, la densidad en el extremo izquierdo (0,103 1) es cinco veces y media la del derecho (0,019 0): se están dejando fuera puntos más creíbles que otros que se incluyen.

Cuidado

El HPD tiene un defecto serio: no es invariante frente a reparametrizaciones. El HPD de θ transformado no es el HPD de $\log\theta$, porque la transformación cambia la densidad por el jacobiano y con ella el orden de credibilidad. El intervalo de colas iguales sí es invariante, porque los cuantiles se transforman sin más. En la práctica: HPD si la parametrización es la natural del problema, colas iguales si va a haber transformaciones por medio.

Idea clave

La teoría de la decisión enseña además cómo se comparan los dos enfoques. Un estimador es *admisible* si ningún otro tiene riesgo menor o igual para todo θ y estrictamente menor para alguno. El resultado clave es que todo estimador de Bayes con priori propia es admisible, y que —bajo condiciones amplias— todo estimador admisible es de Bayes o límite de estimadores de Bayes. Es decir: la clase de los procedimientos razonables desde el punto de vista frecuentista coincide esencialmente con la de los procedimientos bayesianos. Y de ahí que la inadmisibilidad de la media muestral en dimensión tres o más —el fenómeno de Stein— se explique de forma tan limpia con una priori jerárquica: el estimador de James–Stein *es* un Bayes empírico disfrazado.

20.8. Modelos jerárquicos: aprender de los vecinos

Aquí es donde el enfoque bayesiano gana con claridad. El problema es viejo: hay que estimar muchos parámetros parecidos —el efecto de un programa en ocho colegios, la tasa de éxito de ocho hospitales, la media de bateo de doscientos jugadores— y para cada uno hay pocos datos.

Idea clave

Hay dos respuestas obvias y las dos son malas. Estimar cada grupo por separado *ignora* que los grupos se parecen, y produce estimaciones ruidosísimas: el máximo observado casi siempre es una casualidad afortunada. Agrupar todo y dar una sola estimación *niega* que los grupos difieran. La respuesta correcta está en medio, y el modelo jerárquico la calcula sola.

Definición 20.33 (Modelo de dos niveles)**.**

$$y_j \mid \theta_j \sim N(\theta_j,\ s_j^2), \qquad \theta_j \mid \mu, \tau \sim N(\mu,\ \tau^2),$$

con μ y τ desconocidos. La posteriori de cada θ_j es

$$E[\theta_j \mid y] = B_j\,\mu + (1 - B_j)\,y_j, \qquad B_j = \frac{s_j^2}{s_j^2 + \tau^2}.$$

El coeficiente B_j es el **factor de contracción**, y su lectura es transparente: si el grupo j está medido con mucho ruido (s_j grande), $B_j \to 1$ y la estimación se acerca a la media global; si está medido con precisión, $B_j \to 0$ y se queda donde está. La cantidad de préstamo que cada grupo pide a sus vecinos la decide su propia incertidumbre.

Ejemplo 20.34 (Ocho programas escolares)**.** El caso clásico. Ocho colegios prueban un programa de preparación; cada uno informa del efecto estimado y_j con su error típico s_j. Con

$\tau = 10$:

j	y_j	s_j	B_j	estimación	ee	
1	28	15	0,692	14,24	8,32	
2	8	10	0,500	8,06	7,07	
3	−3	16	0,719	5,00	8,48	
4	7	11	0,548	7,62	7,40	$\mu = 8{,}13$
5	−1	9	0,448	3,08	6,69	
6	1	11	0,548	4,90	7,40	
7	18	10	0,500	13,06	7,07	
8	12	18	0,764	9,04	8,74	

Tres lecturas.

Primera. El colegio 1 informó de un efecto de 28, el mayor con diferencia. La estimación jerárquica lo baja a 14,24: la mitad. No porque se desconfíe de ese colegio en particular, sino porque *el máximo de ocho observaciones ruidosas está sesgado hacia arriba* casi por definición. Y el colegio 3, que informó de -3, sube a $+5$.

Segunda. La contracción no es la misma para todos. El colegio 5, el mejor medido ($s = 9$), conserva más de la mitad de su valor propio; el 8, el peor medido ($s = 18$), cede casi tres cuartas partes.

Tercera, y la decisiva. La suma de las varianzas de los ocho errores típicos originales es 1328; la de las estimaciones contraídas, 472. El error total esperado se ha reducido a poco más de un tercio. No se ha añadido ni un solo dato: solo se ha reconocido que los ocho colegios pertenecen a la misma población.

Observación 20.35 (La conexión que cierra tres capítulos). Este modelo es *el mismo* que los efectos aleatorios del análisis de la varianza, escrito en lenguaje bayesiano: τ^2 es la componente de varianza entre grupos y s_j^2 la de dentro. Y el estimador de James–Stein, que demostró que la media muestral es inadmisible en dimensión tres o más, es exactamente esta fórmula con τ estimado de los propios datos: un **Bayes empírico**. El famoso "paradójico" resultado de Stein deja de ser paradójico en cuanto se ve como lo que es: contraer hacia una media común es óptimo porque los parámetros *de verdad* se parecen entre sí.

Sobre la priori de τ hay un aviso. La elección tradicional $\pi(\tau) \propto 1/\tau$ produce una posteriori impropia en este modelo: no es una priori difusa, es una priori inválida. La recomendación actual es una media-Cauchy o una media-normal sobre τ, que es débilmente informativa y siempre da posteriori propia.

20.9. Prioris de Jeffreys y cómo se calcula de verdad la evidencia

Quedan dos cabos sueltos: qué priori usar cuando uno no quiere aportar información, y cómo se calcula la integral que define la evidencia.

Definición 20.36 (Priori de Jeffreys)**.**

$$\pi_J(\theta) \;\propto\; \sqrt{I(\theta)},$$

con I la información de Fisher.

Teorema 20.37 (Invariancia)**.** *Si $\phi = g(\theta)$ es una reparametrización, la priori de Jeffreys de ϕ es la transformada de la de θ.*

Demostración. $I(\phi) = I(\theta)\,(\mathrm{d}\theta/\mathrm{d}\phi)^2$, luego $\sqrt{I(\phi)} = \sqrt{I(\theta)}\,|\mathrm{d}\theta/\mathrm{d}\phi|$, que es precisamente la regla de cambio de variable para densidades. □

Una línea, y resuelve el problema que el capítulo planteó: la uniforme en p no es uniforme en el logit, pero la de Jeffreys sí es la misma priori escrita en las dos escalas. Casos concretos:

$$\text{binomial: } \pi_J = \mathrm{Beta}(\tfrac{1}{2},\tfrac{1}{2}), \qquad \text{Poisson: } \pi_J \propto \lambda^{-1/2}, \qquad \text{escala: } \pi_J \propto \sigma^{-1}.$$

Ejemplo 20.38 (Dos éxitos de diez, con tres prioris)**.**

priori	posteriori	media	IC 95 %
uniforme $\mathrm{Beta}(1,1)$	$\mathrm{Beta}(3,9)$	0,2500	[0,060; 0,518]
Jeffreys $\mathrm{Beta}(\frac{1}{2},\frac{1}{2})$	$\mathrm{Beta}(2{,}5;\ 8{,}5)$	0,2273	[0,044; 0,503]
Haldane $\mathrm{Beta}(0,0)$	$\mathrm{Beta}(2,8)$	0,2000	[0,028; 0,483]

Las tres son "no informativas" y dan medias que difieren en un 25 por ciento relativo. Con diez datos, la elección de la priori objetiva *importa*; con mil, las tres coincidirían en la tercera cifra. La moraleja no es que haya que elegir bien, sino que hay que *declarar* cuál se ha usado y comprobar la sensibilidad.

Aviso técnico: la de Haldane es impropia ($\int \pi = \infty$) y produce posteriori impropia si se observan 0 éxitos o n éxitos. Una priori impropia es admisible solo si la posteriori sale propia, y eso hay que verificarlo, no suponerlo.

La evidencia: de la integral imposible a tres aproximaciones

Definición 20.39 (Aproximación de Laplace)**.** Desarrollando log de la integranda alrededor de su máximo $\hat\theta$ y quedándose en el término cuadrático,

$$p(y) = \int L(\theta)\pi(\theta)\mathrm{d}\theta \;\approx\; L(\hat\theta)\,\pi(\hat\theta)\,(2\pi)^{d/2}\,|H|^{-1/2},$$

con H el hessiano de $-\log$ de la integranda en el máximo.

Ejemplo 20.40 (Exacta, Laplace y BIC, lado a lado)**.** Modelo beta-binomial con priori uniforme, observando una proporción de 0,3:

n	$\log p(y)$ exacta	Laplace	$-\mathrm{BIC}/2$
20	−3,044 5	−3,011 4	−3,150 0
200	−5,303 3	−5,299 9	−5,438 5
2000	−7,601 4	−7,601 1	−7,739 7

Léase la columna de Laplace: el error pasa de 0,033 a 0,003 y a 0,0003, es decir, es $O(1/n)$. Y el BIC, que es Laplace tras tirar todos los términos que no crecen con n, se queda a una distancia *constante* de 0,14: es $O(1)$, no converge al valor exacto. Sirve para comparar modelos, donde la constante se cancela en parte, y no para estimar la evidencia.

Regla: qué usar en la práctica

- **Laplace** cuando la posteriori es aproximadamente normal y la dimensión es baja. Barato y sorprendentemente bueno.
- **Muestreo por importancia puente** (*bridge sampling*) cuando hace falta precisión y se tienen muestras de la posteriori. Es el estándar actual para factores de Bayes.
- **WAIC** y **validación cruzada dejando uno fuera con pesos de importancia** cuando el objetivo no es la evidencia sino la *capacidad predictiva*. Son criterios distintos y responden a preguntas distintas: la evidencia premia al modelo que generó los datos, la predictiva al que predice mejor los próximos.
- Para obtener las muestras: **Monte Carlo hamiltoniano** —que usa el gradiente de la log-posteriori para proponer saltos largos y aceptados casi siempre— en problemas de dimensión moderada, e **inferencia variacional** —que sustituye el muestreo por una optimización sobre una familia de aproximaciones— cuando el modelo es enorme y basta con una aproximación sesgada pero rápida.

20.10. Por qué funciona Metropolis–Hastings, y de dónde sale la conjugación normal

Dos piezas se han usado como recetas: el algoritmo que genera muestras de la posteriori, y las fórmulas del modelo normal–normal. Ambas se deducen en pocas líneas.

Definición 20.41 (Balance detallado)**.** Una cadena de Markov con núcleo de transición $q(x \to y)$ satisface el *balance detallado* respecto de una distribución π si

$$\pi(x)\, q(x \to y) = \pi(y)\, q(y \to x) \qquad \text{para todos } x, y.$$

Proposición 20.42 (Balance detallado implica estacionariedad)**.** *Si una cadena satisface el balance detallado respecto de π, entonces π es una distribución estacionaria de la cadena.*

Demostración. Súmese (o intégrese) la igualdad del balance sobre x:

$$\sum_x \pi(x) q(x \to y) = \sum_x \pi(y) q(y \to x) = \pi(y) \sum_x q(y \to x) = \pi(y),$$

porque las probabilidades de transición desde y suman 1. El miembro izquierdo es exactamente $(\pi Q)(y)$, luego $\pi Q = \pi$. □

Teorema 20.43 (Metropolis–Hastings tiene la posteriori como estacionaria)**.** *Sea π la distribución objetivo, conocida sólo salvo constante multiplicativa, y $g(x \to y)$ una densidad de propuesta. El algoritmo que, estando en x, propone $y \sim g(x \to \cdot)$ y lo acepta con probabilidad*

$$\alpha(x \to y) = \min\left\{1,\ \frac{\pi(y)\, g(y \to x)}{\pi(x)\, g(x \to y)}\right\}$$

—quedándose en x si lo rechaza— define una cadena que satisface el balance detallado respecto de π, y por tanto tiene a π por distribución estacionaria.

Demostración. El núcleo efectivo para $y \neq x$ es $q(x \to y) = g(x \to y)\,\alpha(x \to y)$. Hay que ver $\pi(x)q(x \to y) = \pi(y)q(y \to x)$. Supóngase, sin pérdida de generalidad, que

$$\frac{\pi(y)g(y \to x)}{\pi(x)g(x \to y)} \leq 1,$$

de modo que $\alpha(x \to y) = \frac{\pi(y)g(y \to x)}{\pi(x)g(x \to y)}$ y, por ser el cociente inverso ≥ 1, $\alpha(y \to x) = 1$. Entonces

$$\pi(x)\, q(x \to y) = \pi(x)\, g(x \to y) \cdot \frac{\pi(y)g(y \to x)}{\pi(x)g(x \to y)} = \pi(y)\, g(y \to x) = \pi(y)\, q(y \to x),$$

que es lo buscado. El caso contrario es simétrico, y para $x = y$ la igualdad es trivial. Por la proposición 20.42, π es estacionaria; y si la cadena es además irreducible y aperiódica —lo cual se cumple con propuestas de soporte pleno—, el teorema ergódico (capítulo 22) garantiza que los promedios a lo largo de la trayectoria convergen a las esperanzas bajo π. □

Observación 20.44 (Por qué basta conocer π salvo constante)*.* En el cociente de aceptación, π aparece dos veces, arriba y abajo: cualquier constante multiplicativa se cancela. Esa es toda la razón por la que el método es útil en estadística bayesiana, donde la posteriori se conoce como $\pi(\theta) \propto L(\theta)\, p(\theta)$ y la constante de normalización $\int L\, p\, \mathrm{d}\theta$ es justamente la integral que no se sabe calcular. Metropolis–Hastings esquiva la integral imposible aprovechando que sólo importan los *cocientes* de la posteriori en dos puntos.

Teorema 20.45 (Modelo normal–normal, deducido)**.** *Sea $\theta \sim N(\mu_0, \tau_0^2)$ la distribución a priori y $x_1, \ldots, x_n$ una muestra $N(\theta, \sigma^2)$ con σ^2 conocida. Entonces la posteriori es normal,*

$$\theta \mid x \ \sim\ N(\mu_1, \tau_1^2), \qquad \frac{1}{\tau_1^2} = \frac{1}{\tau_0^2} + \frac{n}{\sigma^2}, \qquad \mu_1 = \tau_1^2\left(\frac{\mu_0}{\tau_0^2} + \frac{n\bar{x}}{\sigma^2}\right).$$

Demostración. Escríbase la posteriori salvo constantes que no dependan de θ y tómense logaritmos:

$$\log \pi(\theta \mid x) = -\frac{(\theta - \mu_0)^2}{2\tau_0^2} - \sum_{i=1}^{n} \frac{(x_i - \theta)^2}{2\sigma^2} + \text{cte.}$$

En el segundo término se desarrolla $\sum (x_i - \theta)^2 = \sum (x_i - \bar{x})^2 + n(\bar{x} - \theta)^2$ —la descomposición habitual—, y el primer sumando no depende de θ: pasa a la constante. Queda una función cuadrática en θ:

$$\log \pi = -\frac{1}{2}\left[\left(\frac{1}{\tau_0^2} + \frac{n}{\sigma^2}\right)\theta^2 - 2\left(\frac{\mu_0}{\tau_0^2} + \frac{n\bar{x}}{\sigma^2}\right)\theta\right] + \text{cte.}$$

Una densidad cuyo logaritmo es cuadrático *es* normal; completando el cuadrado, $-\frac{1}{2\tau_1^2}(\theta-\mu_1)^2$ con precisión $1/\tau_1^2$ igual al coeficiente de θ^2 y media μ_1 igual al coeficiente lineal dividido entre aquél. Eso es exactamente el enunciado. □

Observación 20.46 (Las precisiones se suman, y la media es un promedio ponderado)**.** La fórmula se recuerda sin memorizarla si se piensa en *precisión* $= 1/$varianza:

- la precisión posterior es la suma de la precisión a priori y la de los datos;
- la media posterior es el promedio de μ_0 y $\bar{x}$ *ponderado por sus precisiones.*

De ahí se leen los dos límites importantes. Si $\tau_0 \to \infty$ —a priori no informativa— la posteriori tiende a $N(\bar{x}, \sigma^2/n)$, que es exactamente el resultado frecuentista: con información previa nula, ambas escuelas coinciden. Y si $n \to \infty$, la posteriori se concentra en $\bar{x}$ sea cual sea la a priori: con datos suficientes, la opinión inicial se vuelve irrelevante. Ese segundo hecho —la consistencia bayesiana, formalizada en el teorema de Bernstein–von Mises— es la mejor respuesta a la objeción de que las a priori son arbitrarias.

Teorema 20.47 (El estimador de Bayes depende de la función de pérdida)**.** *Sea $\pi(\theta \mid x)$ la posteriori. El estimador que minimiza la pérdida esperada* $\mathrm{E}[L(\theta,\hat\theta) \mid x]$ *es:*

1. *la* media *posterior, si $L(\theta,\hat\theta) = (\theta-\hat\theta)^2$;*
2. *la* mediana *posterior, si $L = |\theta-\hat\theta|$;*
3. *la* moda *posterior, si L es la pérdida* 0–1 *(límite de pérdidas que penalizan por igual todo error mayor que ε).*

Demostración. (1) Derívese $R(\hat\theta) = \mathrm{E}[(\theta-\hat\theta)^2 \mid x]$ respecto de $\hat\theta$: $R' = -2\,\mathrm{E}[\theta-\hat\theta \mid x]$, que se anula en $\hat\theta = \mathrm{E}[\theta \mid x]$; y $R'' = 2 > 0$.

(2) $R(\hat\theta) = \int |\theta-\hat\theta|\pi\,\mathrm{d}\theta = \int_{-\infty}^{\hat\theta}(\hat\theta-\theta)\pi + \int_{\hat\theta}^{\infty}(\theta-\hat\theta)\pi$. Derivando bajo el signo integral (los términos de frontera se cancelan porque el integrando se anula en $\theta=\hat\theta$),

$$R'(\hat\theta) = \Pi(\hat\theta) - (1-\Pi(\hat\theta)) = 2\Pi(\hat\theta) - 1,$$

con Π la función de distribución posterior. Se anula en $\Pi(\hat\theta) = \frac{1}{2}$: la mediana.

(3) Con la pérdida que vale 0 si $|\theta-\hat\theta| \le \varepsilon$ y 1 en otro caso, la pérdida esperada es $1 - P(|\theta-\hat\theta| \le \varepsilon \mid x)$, que se minimiza maximizando la masa posterior en un intervalo de anchura 2ε alrededor de $\hat\theta$. Cuando $\varepsilon \to 0$, ese máximo se alcanza donde la densidad es máxima: la moda. □

Observación 20.48 (Una consecuencia incómoda y saludable)**.** No existe "el" estimador puntual bayesiano: existe uno por cada función de pérdida, y elegirla es una decisión del analista, no del método. Si la posteriori es simétrica y unimodal —como la normal del teorema anterior— los tres coinciden y la cuestión es invisible; pero en posterioris asimétricas —una gamma, una beta cerca del borde— pueden diferir sustancialmente. Presentar sólo un número sin decir qué pérdida se ha supuesto es, estrictamente, información incompleta, y es la razón por la que la práctica bayesiana recomienda dar el intervalo de credibilidad entero antes que un resumen.

20.11. Gibbs, prioris impropias y diagnóstico de convergencia

Definición 20.49 (Condicional completa y muestreador de Gibbs)**.** Sea una posteriori sobre un vector de parámetros $\theta = (\theta_1, \dots, \theta_k)$. La *condicional completa* de θ_j es su distribución dado *todo* lo demás,

$$\pi(\theta_j \mid \theta_{-j}, x), \qquad \theta_{-j} = (\theta_1, \dots, \theta_{j-1}, \theta_{j+1}, \dots, \theta_k).$$

El *muestreador de Gibbs* genera la cadena actualizando una coordenada cada vez: partiendo de $\theta^{(0)}$, en el paso $t+1$ se extrae

$$\theta_1^{(t+1)} \sim \pi(\theta_1 \mid \theta_2^{(t)}, \dots, \theta_k^{(t)}, x), \quad \theta_2^{(t+1)} \sim \pi(\theta_2 \mid \theta_1^{(t+1)}, \theta_3^{(t)}, \dots, x), \ \dots$$

usando siempre los valores más recientes de las coordenadas ya actualizadas.

Observación 20.50 (Gibbs es Metropolis–Hastings con aceptación 1). Si en Metropolis–Hastings se propone según la condicional completa, $g(\theta \to \theta') = \pi(\theta'_j \mid \theta_{-j})$ con $\theta'_{-j} = \theta_{-j}$, el cociente de aceptación vale exactamente 1: la propuesta *nunca* se rechaza. Por eso Gibbs no tiene parámetros que ajustar y es tan cómodo cuando las condicionales completas son conocidas —lo que ocurre siempre en modelos jerárquicos con prioris conjugadas—. El precio es que sólo se mueve paralelamente a los ejes, de modo que si los parámetros están muy correlacionados la cadena avanza a pasitos diagonales y mezcla muy despacio.

Definición 20.51 (Priori impropia y posteriori propia)**.** Una *priori impropia* es una función no negativa $\pi(\theta)$ que se usa como si fuera una densidad pero cuya integral diverge: $\int \pi(\theta)\,\mathrm{d}\theta = \infty$. Ejemplos habituales: $\pi(\mu) \propto 1$ para una media, o $\pi(\sigma) \propto 1/\sigma$ para una escala.

Una priori impropia es *admisible* sólo si la posteriori resultante es *propia*, es decir, si

$$\int L(\theta; x)\,\pi(\theta)\,\mathrm{d}\theta \;<\; \infty.$$

Cuidado

[Esa integral hay que comprobarla, y a veces diverge] Con una priori impropia no hay ninguna garantía automática: la posteriori puede no existir, y en ese caso todo lo que se calcule con ella carece de sentido —incluidas las salidas de un programa de MCMC, que seguirá produciendo números tranquilamente, porque la cadena no "sabe" que su distribución objetivo no es integrable—. El caso patológico clásico es el modelo jerárquico normal con $\pi(\tau) \propto 1/\tau$ sobre la desviación típica entre grupos: la posteriori es impropia y la cadena deriva hacia $\tau \to 0$ sin estabilizarse nunca. La comprobación es analítica y no la puede hacer el ordenador por uno.

Definición 20.52 ($\hat{R}$ y tamaño efectivo de muestra)**.** Ejecútense m cadenas independientes de longitud n (tras descartar el calentamiento), y para un escalar de interés sean

$$B = \frac{n}{m-1}\sum_{j=1}^{m}(\bar\theta_{\cdot j} - \bar\theta_{\cdot\cdot})^2 \quad \text{(varianza } \textit{entre} \text{ cadenas)}, \qquad W = \frac{1}{m}\sum_{j=1}^{m} s_j^2 \quad \text{(varianza } \textit{dentro}\text{)}.$$

El *estadístico de Gelman–Rubin* es

$$\widehat{R} = \sqrt{\frac{\widehat{\mathrm{Var}}^{+}}{W}}, \qquad \widehat{\mathrm{Var}}^{+} = \frac{n-1}{n} W + \frac{1}{n} B.$$

El *tamaño efectivo de muestra* es

$$n_{\mathrm{ef}} = \frac{mn}{1 + 2\sum_{k \geq 1} \rho_k},$$

con ρ_k las autocorrelaciones de la cadena.

Observación 20.53 (Cómo se leen los dos números). $\widehat{R}$ compara la dispersión entre cadenas con la dispersión dentro de cada una: si las cadenas han olvidado su punto de partida y exploran la misma región, ambas coinciden y $\widehat{R} \to 1$. El umbral tradicional era $\widehat{R} < 1{,}1$; la recomendación actual es más exigente, $\widehat{R} < 1{,}01$, porque valores de 1,05 ya delatan cadenas que no se han mezclado. Y n_{ef} dice a cuántas muestras *independientes* equivale la cadena: si se han generado 10 000 iteraciones con $n_{\mathrm{ef}} = 300$, el error de Monte Carlo es el de 300 datos, no el de diez mil. La regla práctica es exigir $n_{\mathrm{ef}} > 400$ por cantidad de interés. Conviene subrayar lo obvio: ninguno de los dos demuestra convergencia, sólo detectan la *no* convergencia.

Regla: Elegir y justificar una distribución a priori

1. **¿Hay información previa real?** Estudios anteriores, límites físicos, rangos plausibles. Si la hay, tradúcela a una distribución y *dilo*: es información, no trampa.
2. **¿Existe conjugada?** Si el modelo la tiene (beta–binomial, gamma–Poisson, normal–normal), úsala: la posteriori sale en forma cerrada y los hiperparámetros se interpretan como "datos equivalentes". Una Beta$(2, 2)$ equivale a haber visto un éxito y un fracaso.
3. **Si no hay información, prefiere una priori débilmente informativa** —una normal ancha, una t con pocos grados de libertad, una semi-Cauchy para escalas— antes que una impropia. Descartan los valores absurdos sin imponer nada, y garantizan que la posteriori es propia.
4. **Si se usa una impropia, comprueba analíticamente** que $\int L\pi \, \mathrm{d}\theta < \infty$. La de Jeffreys, $\pi(\theta) \propto \sqrt{I(\theta)}$, tiene la virtud de ser invariante frente a reparametrizaciones, pero no es automáticamente admisible en varias dimensiones.
5. **Análisis de sensibilidad, siempre.** Repite el ajuste con dos o tres prioris razonablemente distintas. Si las conclusiones no cambian, la elección era irrelevante y se puede decir así; si cambian, *esa* es la conclusión que hay que reportar: los datos no bastan y el resultado depende de lo que se supusiera.
6. **Nunca elijas la priori mirando los datos** y luego la presentes como previa. Eso no es bayesiano ni frecuentista: es hacer trampa dos veces.

20.12. Ejercicios resueltos

Tres problemas que muestran de dónde sale la posteriori y qué se hace con ella.

Ejercicio 20.1. Con priori uniforme Beta$(1,1)$ sobre la probabilidad θ de una moneda, se observan 7 caras en 10 tiradas. Halle la posteriori, su media, su moda y un intervalo creíble al 95 por ciento.

Solución. La verosimilitud binomial es $\theta^7(1-\theta)^3$ y la priori es constante, luego la posteriori es proporcional a $\theta^7(1-\theta)^3$, es decir, Beta$(8,4)$.

Su media es $8/12 = 0{,}667$ y su moda es $(8-1)/(8+4-2) = 0{,}700$, que coincide con el estimador de máxima verosimilitud $7/10$. El intervalo creíble central al 95 por ciento es (0,390; 0,891).

Aquí sí se puede decir lo que con un intervalo de confianza no se podía: dada la priori elegida y los datos observados, hay un 95 por ciento de probabilidad de que θ esté en ese intervalo. La diferencia no es retórica, es de objeto: para el bayesiano θ es una variable aleatoria con distribución, y esa frase tiene sentido. ■

Ejercicio 20.2. Con priori $\theta \sim N(100, 15^2)$ y una muestra de $n = 9$ con $\bar{x} = 112$ y $\sigma = 18$ conocida, halle la distribución posterior.

Solución. Para el modelo normal-normal, las precisiones se suman:

$$\frac{1}{\tau^2_{\text{post}}} = \frac{1}{15^2} + \frac{9}{18^2}, \qquad \mu_{\text{post}} = \tau^2_{\text{post}}\Big(\frac{100}{15^2} + \frac{9\cdot 112}{18^2}\Big).$$

Sale $\tau^2_{\text{post}} = 31{,}03$, es decir, $\tau_{\text{post}} = 5{,}57$, y $\mu_{\text{post}} = 110{,}34$.

La media posterior queda entre la priori, 100, y la muestral, 112, y mucho más cerca de la segunda: la priori equivale a $\sigma^2/\tau^2 = 324/225 = 1{,}44$ observaciones, frente a las nueve reales. Ésa es la lectura útil de una priori conjugada: se traduce a un número de observaciones imaginarias, y así se ve de un vistazo cuánto pesa. ■

Ejercicio 20.3. Una enfermedad afecta al 1 por mil. Un test tiene sensibilidad 0,99 y especificidad 0,95. Alguien da positivo. Calcule la probabilidad de que esté enfermo y explique el resultado.

Solución. Sobre 100 000 personas: 100 enfermas, de las que 99 dan positivo; y 99 900 sanas, de las que un 5 por ciento, es decir 4995, dan positivo también. El total de positivos es 5094, y $P(\text{enfermo} \mid +) = 99/5094 = 0{,}0194$.

Menos de un 2 por ciento. El resultado sorprende a casi todo el mundo, incluidos muchos médicos, y la causa es que la intuición ignora la prevalencia: hay tantísima gente sana que incluso un 5 por ciento de falsos positivos produce cincuenta veces más positivos erróneos que correctos. El test no es malo —multiplica por veinte la probabilidad de estar enfermo, de 0,001 a 0,019—, simplemente parte de muy abajo. De ahí que no se hagan cribados masivos de enfermedades raras sin un segundo test confirmatorio. ■

20.13. Ejercicios propuestos

20.1. Con priori Beta$(2,8)$ y 3 éxitos en 10 intentos, halle la posteriori y su media.

20.2. ¿Qué priori conjugada corresponde a una verosimilitud de Poisson? ¿Y a una normal con varianza conocida?

20.3. Interprete Beta(a,b) como "$a-1$ éxitos y $b-1$ fracasos imaginarios".

20.4. ¿En qué se diferencia un intervalo creíble de uno de confianza?

20.5. ¿Qué es el factor de Bayes y cómo se interpreta un valor de 10?

20.6. Con la enfermedad del ejercicio resuelto, ¿cuánto vale la probabilidad posterior si la prevalencia fuese del 10 por ciento en lugar del 1 por mil?

20.7. Explique por qué la posteriori se concentra al aumentar n y qué le pasa a la influencia de la priori.

20.8. ¿Qué es una priori impropia y qué hay que comprobar antes de usarla?

20.9. ¿Para qué sirve el muestreo de Gibbs y cuándo es preferible a Metropolis–Hastings?

20.10. ¿Qué mide el estadístico $\hat{R}$ y qué valor indica problemas?

20.11. ¿Puede un análisis bayesiano y uno frecuentista dar el mismo intervalo numérico? ¿Cuándo?

20.12. Critique la afirmación: "el bayesiano es subjetivo y el frecuentista objetivo".

Capítulo 21

Métodos no paramétricos y remuestreo

Lo que necesitas traer

Del capítulo 18: contrastes de hipótesis, nivel y potencia. Del capítulo 16: la distribución χ^2 y la aproximación normal. Del capítulo 5: conteo de permutaciones, que es lo que hay detrás de los contrastes exactos.

Casi todo lo visto hasta aquí supone que los datos vienen de una distribución concreta, típicamente normal. Cuando ese supuesto es dudoso (muestras pequeñas, distribuciones asimétricas, presencia de atípicos) hacen falta técnicas que no lo necesiten. Hay dos familias: los contrastes basados en rangos, que sustituyen los valores por su posición en el orden, y los métodos de remuestreo, que usan el ordenador para generar la distribución del estadístico a partir de la propia muestra.

21.1. Contrastes basados en rangos

Definición 21.1. El *contraste de Mann–Whitney* compara dos muestras independientes ordenando conjuntamente todas las observaciones y sumando los rangos de cada grupo. El *contraste de Wilcoxon de los signos* hace lo análogo con datos apareados, usando los rangos de los valores absolutos de las diferencias con el signo de cada una. El *contraste de Kolmogórov–Smirnov* mide la máxima distancia vertical entre la función de distribución empírica y la teórica.

Idea clave

Qué dice esto en cristiano. La idea es tirar los valores y quedarse solo con el orden. Si la observación más grande es la misma sea cual sea la escala, el contraste se vuelve inmune a la forma de la distribución y a los valores extremos: cambiar un dato de 64 a 6400 no altera su rango. Se pierde algo de potencia cuando la normalidad sí se cumple (los rangos ignoran información), pero se gana robustez cuando no. La regla práctica es sencilla: con muestras grandes y sin atípicos, los métodos clásicos; con muestras

pequeñas, ordinales o contaminadas, los rangos.

Ejemplo 21.2 (Los mismos datos, con y sin supuestos)**. (a) Dos grupos independientes.** Sean $A = \{12, 15, 17, 19, 22, 25, 28\}$ y $B = \{14, 18, 21, 26, 30, 33, 35\}$. La t de Student da $t = -1{,}516542$ con $p = 0{,}155275$; el contraste de Mann–Whitney da $U = 14$ con $p = 0{,}208625$. Ambos coinciden en no rechazar, pero el p del segundo es algo mayor: ese es el precio de la robustez cuando la normalidad se cumple.

(b) Datos apareados. Con las ocho diferencias 2,1; −0,4; 3,3; 1,8; −1,1; 2,5; 0,9; 4,0, el contraste de Wilcoxon da $p = 0{,}054688$, justo en el límite. Es el sustituto no paramétrico de la t para muestras apareadas.

(c) El contraste del signo, todavía más burdo. Ignorando incluso la magnitud y contando solo cuántas diferencias son positivas, salen 6 de 8, con $p = 0{,}289062$ por la binomial. Es mucho menos potente porque descarta más información, pero no exige absolutamente nada sobre la distribución.

(d) Ajuste a una distribución. Para comprobar si el grupo A es compatible con una normal de media 20 y desviación 6, Kolmogórov–Smirnov da un estadístico de 0,137612 con $p = 0{,}996412$: perfectamente compatible. Este contraste tiene la ventaja de comparar distribuciones enteras, no solo medias.

Cuidado

[No paramétrico no significa sin supuestos] Mann–Whitney no contrasta la igualdad de medias sino la de *distribuciones*, y su interpretación habitual como comparación de medianas solo es válida si ambas distribuciones tienen la misma forma. Con formas distintas puede salir significativo aunque las medianas coincidan. Segundo aviso: todos estos contrastes siguen exigiendo *independencia* de las observaciones, que es el supuesto más importante y el que ninguno de ellos relaja. Y tercero: cuando la normalidad sí se cumple, usar rangos cuesta potencia; con muestras grandes la eficiencia relativa de Mann–Whitney frente a la t es $\frac{3}{\pi} \approx 0{,}955$, es decir, un 4,5 % de pérdida, que es un precio bajo pero no nulo.

21.2. Qué significa exactamente no paramétrico: estadísticos de orden, distribución empírica y el contraste del signo

El capítulo se llama “métodos no paramétricos” y ha empezado usando esa etiqueta sin definirla, igual que ha usado F_n, ha escrito $x_{(k)}$ y ha calculado un p-valor del contraste del signo sin decir en ningún sitio qué son esas tres cosas. Son las piezas más elementales del edificio, y además las más reutilizables: el estadístico de orden y la función de distribución empírica aparecen en todos los contrastes del capítulo, y el contraste del signo es el ejemplo más simple posible de un procedimiento que funciona sin suponer absolutamente nada.

Paramétrico, no paramétrico, libre de distribución

Definición 21.3 (Modelo paramétrico y modelo no paramétrico)**.** Un *modelo paramétrico* es una familia de distribuciones $\mathcal{P} = \{P_\theta : \theta \in \Theta\}$ con $\Theta \subset \mathbb{R}^d$ para algún d finito: cada distribución de la familia queda determinada por un número finito de números. Un *modelo no paramétrico* es aquel cuya familia $\mathcal{P}$ no admite tal descripción, por ejemplo $\mathcal{P} = \{$todas las distribuciones continuas sobre $\mathbb{R}\}$ o $\mathcal{P} = \{$todas las distribuciones continuas y simétricas respecto de un punto$\}$. Un *método no paramétrico* es el que es válido bajo un modelo no paramétrico.

Definición 21.4 (Estadístico libre de distribución)**.** Un estadístico T es *libre de distribución* bajo una hipótesis nula H_0 si su distribución cuando H_0 es cierta es la misma para todas las distribuciones F compatibles con H_0. En ese caso, sus valores críticos se pueden tabular de una vez y sirven para siempre.

Idea clave

Qué dice esto en cristiano. "Paramétrico" no quiere decir "que tiene parámetros": casi todo tiene parámetros. Quiere decir que la forma de la distribución está fijada de antemano y solo quedan por concretar unos pocos números. Suponer $N(\mu, \sigma^2)$ es paramétrico porque, dados dos números, la distribución está completamente determinada. Suponer únicamente "es continua" es no paramétrico, porque para determinar una función de distribución continua cualquiera harían falta infinitos números. Y "libre de distribución" es una propiedad distinta y complementaria: dice que la tabla de valores críticos no depende de cuál sea esa forma desconocida. Casi todos los métodos de este capítulo son las dos cosas a la vez, y por eso se confunden, pero no son lo mismo: el bootstrap es no paramétrico y *no* es libre de distribución.

Ejemplo 21.5 (Cuatro problemas clasificados)**. (a) Contrastar $\mu = 0$ con la t de Student** suponiendo datos normales: paramétrico. La distribución del estadístico bajo H_0 es una t_{n-1} *porque* los datos son normales; si no lo son, la tabla miente (poco con n grande, mucho con n pequeño y colas gruesas).

(b) Contrastar que la mediana vale 0 con el contraste del signo: no paramétrico y libre de distribución. Como se demuestra más abajo, la distribución nula es $\mathrm{Bin}(n, 1/2)$ sea cual sea F, con tal de que sea continua.

(c) Estimar la densidad con núcleos: no paramétrico. El objeto que se estima, la función f entera, no se deja describir con un número finito de parámetros; de ahí la lentitud de convergencia que se discute más adelante en el capítulo.

(d) Estimar el error típico de la mediana por bootstrap: no paramétrico, pero no libre de distribución. El resultado depende de la muestra concreta y no hay ninguna tabla universal: hay que remuestrear cada vez.

Los estadísticos de orden

Definición 21.6 (Estadísticos de orden)**.** Dada una muestra $x_1, \dots, x_n$, sus *estadísticos de orden* son los mismos valores reordenados de menor a mayor, y se escriben

$$x_{(1)} \le x_{(2)} \le \cdots \le x_{(n)}.$$

Así $x_{(1)} = \min_i x_i$ y $x_{(n)} = \max_i x_i$. El *recorrido* es $x_{(n)} - x_{(1)}$ y la *mediana muestral* es $x_{((n+1)/2)}$ si n es impar y el promedio de $x_{(n/2)}$ y $x_{(n/2+1)}$ si es par.

Ejemplo 21.7 (Ocho datos ordenados)**.** Para la muestra 7, 3, 11, 5, 3, 9, 14, 6 los estadísticos de orden son

$$x_{(1)} = 3, \quad x_{(2)} = 3, \quad x_{(3)} = 5, \quad x_{(4)} = 6, \quad x_{(5)} = 7, \quad x_{(6)} = 9, \quad x_{(7)} = 11, \quad x_{(8)} = 14.$$

El recorrido es $14 - 3 = 11$ y la mediana muestral, al ser $n = 8$ par, vale $(x_{(4)} + x_{(5)})/2 = (6+7)/2 = 6{,}5$. Obsérvese que los empates no estorban: dos datos iguales ocupan dos posiciones consecutivas y la definición sigue teniendo sentido, aunque la asignación de cuál es cuál sea arbitraria.

Proposición 21.8 (Distribución del máximo y del mínimo)**.** *Sean $X_1, \dots, X_n$ independientes con la misma función de distribución F. Entonces*

$$P\big(X_{(n)} \le t\big) = F(t)^n, \qquad P\big(X_{(1)} > t\big) = \big(1 - F(t)\big)^n.$$

Demostración. El máximo no pasa de t exactamente cuando *todos* los datos no pasan de t, y por independencia esa probabilidad es el producto de las n probabilidades individuales, todas iguales a $F(t)$. El mínimo supera a t exactamente cuando todos lo superan, con probabilidad $(1 - F(t))^n$. □

Ejemplo 21.9 (Para qué sirve eso)**.** Si el 10 por ciento de las piezas de un lote son defectuosas y se toman 8 al azar, la probabilidad de que *ninguna* de las ocho lo sea es $0{,}9^8 = 0{,}430\,467$, de modo que la probabilidad de detectar al menos una es $0{,}569\,533$: poco más de la mitad. Con inspecciones de ocho piezas se le escapa a uno casi la mitad de los lotes con un 10 por ciento de defectuosas, y la cuenta que lo dice es la proposición anterior aplicada al mínimo. Para las uniformes hay además una fórmula muy útil: si $U_1, \dots, U_n$ son uniformes en $[0, 1]$, entonces $\mathrm{E}[U_{(k)}] = k/(n+1)$, es decir, los n datos ordenados reparten el intervalo en $n+1$ trozos iguales en promedio. Con $n = 8$, el tercero ocupa en media la posición $3/9 = 0{,}333\,3$. De ahí salen las posiciones que se usan al dibujar un gráfico de cuantiles.

La función de distribución empírica

Definición 21.10 (Función de distribución empírica)**.** Dada la muestra $x_1, \dots, x_n$, su *función de distribución empírica* es

$$F_n(t) = \frac{1}{n} \#\{i : \ x_i \le t\} = \frac{1}{n} \sum_{i=1}^{n} \mathbf{1}_{\{x_i \le t\}}, \qquad t \in \mathbb{R}.$$

Es una función escalonada, no decreciente, que vale 0 a la izquierda de $x_{(1)}$, vale 1 a partir de $x_{(n)}$, y sube un escalón de altura $1/n$ en cada dato observado (de altura r/n si hay r datos repetidos en ese punto).

Ejemplo 21.11 (La empírica de la muestra anterior, escalón a escalón)**.** Con los ocho datos del ejemplo 21.7:

t	3	5	6	7	9	11	14
$F_n(t)$	0,25	0,375	0,5	0,625	0,75	0,875	1

El primer escalón mide $2/8 = 0{,}25$ porque el valor 3 está repetido; los demás miden $1/8$. Entre dos datos consecutivos la función es constante: $F_n(6{,}5) = F_n(6) = 0{,}5$, y $F_n(2) = 0$. Esa constancia a trozos es exactamente la razón, señalada al estudiar el contraste de Kolmogórov–Smirnov, de que el supremo $\sup_t |F_n(t) - F(t)|$ se alcance siempre en un dato y baste examinar $2n$ números.

Proposición 21.12 (La empírica es un estimador insesgado y consistente)**.** *Sea F la función de distribución de la que provienen los datos, independientes. Para cada t fijo,*

$$n\,F_n(t) \sim \mathrm{Bin}\big(n,\ F(t)\big), \qquad \mathrm{E}[F_n(t)] = F(t), \qquad \mathrm{Var}\big(F_n(t)\big) = \frac{F(t)\big(1-F(t)\big)}{n}.$$

En consecuencia $F_n(t) \to F(t)$ casi seguramente para cada t.

Demostración. Fijado t, cada indicador $\mathbf{1}_{\{x_i \le t\}}$ vale 1 con probabilidad $F(t)$ y 0 con probabilidad $1 - F(t)$, y los n indicadores son independientes; su suma es, por definición, una binomial de parámetros n y $F(t)$. Dividiendo por n: la esperanza de la binomial es $nF(t)$, luego $\mathrm{E}[F_n(t)] = F(t)$; y su varianza es $nF(t)(1-F(t))$, que dividida por n^2 da lo anunciado. La convergencia casi segura es la ley fuerte de los grandes números aplicada a esos indicadores. □

Observación 21.13 (De la convergencia en un punto a la convergencia uniforme)*.* La proposición anterior da la convergencia *para cada t por separado*. El teorema de Glivenko–Cantelli dice bastante más: la convergencia es uniforme,

$$\sup_{t \in \mathbb{R}} \big|F_n(t) - F(t)\big| \ \longrightarrow\ 0 \qquad \text{casi seguramente,}$$

y por eso se le llama a veces el teorema fundamental de la estadística: la muestra reproduce la población entera, no solo punto a punto. La desigualdad de Dvoretzky, Kiefer y Wolfowitz cuantifica la velocidad: con probabilidad al menos $1 - \alpha$, la función de distribución verdadera queda dentro de la banda $F_n(t) \pm \varepsilon$ con $\varepsilon = \sqrt{\log(2/\alpha)/(2n)}$. Con $n = 100$ y $\alpha = 0{,}05$ eso son $\pm 0{,}136$: una banda ancha, que recuerda que con cien datos no se conoce una distribución con mucha precisión. Nótese además que la varianza $F(1-F)/n$ es máxima en la mediana y se anula en las colas: la empírica estima muy bien lo que pasa en los extremos y peor lo que pasa en el centro.

Regla: Usar la función de distribución empírica

1. **Ordena los datos** y dibuja los saltos en $x_{(1)}, \dots, x_{(n)}$, de altura $1/n$ cada uno, acumulando los repetidos en un solo salto mayor.
2. **Léela como una tabla de proporciones**: $F_n(t)$ es la fracción de la muestra que no supera t, y $1 - F_n(t)$ la que lo supera.
3. **Para comparar con un modelo**, superpón F y mide la máxima separación vertical: ese es el estadístico de Kolmogórov–Smirnov.
4. **Para comparar dos muestras**, dibuja las dos empíricas: el mismo criterio da el contraste de Kolmogórov–Smirnov de dos muestras.
5. **No la suavices para que quede bonita.** Si lo que se quiere es una curva suave, lo que hay que estimar es la densidad, y ese es otro problema, con otro compromiso entre sesgo y varianza.

El contraste del signo

Definición 21.14 (Contraste del signo)**.** Sean $d_1, \dots, d_n$ observaciones independientes de una distribución continua —típicamente las diferencias de n pares antes-después— y sea M su mediana. Para contrastar

$$H_0 \colon M = 0 \qquad \text{frente a} \qquad H_1 \colon M \neq 0,$$

se descartan las diferencias nulas, se llama m al número de las que quedan y se toma como estadístico

$$S^+ = \#\{i : \ d_i > 0\}.$$

Se rechaza H_0 cuando S^+ está demasiado lejos de $m/2$, en cualquiera de los dos sentidos.

Teorema 21.15 (Distribución nula del contraste del signo)**.** *Bajo H_0, y siendo la distribución continua,*

$$S^+ \sim \mathrm{Bin}(m, \tfrac{1}{2}), \qquad \mathrm{E}[S^+] = \frac{m}{2}, \qquad \mathrm{Var}(S^+) = \frac{m}{4},$$

sea cual sea la distribución de los datos. El p-valor bilateral exacto de un valor observado $s \geq m/2$ es

$$p = 2\, P(S^+ \geq s) = \frac{2}{2^m} \sum_{j=s}^{m} \binom{m}{j}.$$

Demostración. Si la mediana de d_i es 0, entonces $P(d_i > 0) = \frac{1}{2}$ por la propia definición de mediana, y $P(d_i = 0) = 0$ por ser la distribución continua. Los signos de las m diferencias son, por tanto, m ensayos independientes con probabilidad $\frac{1}{2}$ de salir positivos, y su suma es una binomial de parámetros m y $\frac{1}{2}$; la media y la varianza son las de esa binomial. Obsérvese qué se ha usado y qué no: se ha usado la independencia y la continuidad, y no se ha usado absolutamente nada sobre la forma de la distribución. Por eso el estadístico es libre de distribución en el sentido de la definición 21.4 y el p-valor sale de una tabla binomial universal. □

Ejemplo 21.16 (Diez pacientes, un contraste hecho entero)**.** Se mide la tensión arterial de diez pacientes antes y después de un tratamiento:

antes	72	68	75	80	66	74	79	71	77	69
después	70	69	71	74	64	73	72	70	73	70
d_i	2	-1	4	6	2	1	7	1	4	-1

(a) Contar signos. Hay 8 diferencias positivas y 2 negativas; ninguna nula, así que $m = 10$ y $S^+ = 8$.

(b) El p-valor exacto. Por el teorema 21.15,

$$P(S^+ \geq 8) = \frac{\binom{10}{8} + \binom{10}{9} + \binom{10}{10}}{2^{10}} = \frac{45 + 10 + 1}{1024} = \frac{56}{1024} = 0{,}054\,688,$$

y el p-valor bilateral es $2 \cdot 0{,}054\,688 = 0{,}109\,375$. No se rechaza al 5 %.

(c) Comparación con un contraste que sí usa las magnitudes. Sobre estas mismas diez diferencias, el contraste de Wilcoxon de los rangos con signo —que se estudia con detalle más adelante en el capítulo— da $p = 0{,}021\,5$ y sí rechaza. La razón salta a la vista mirando los datos: entre las ocho subidas hay varias grandes (4, 6, 7), mientras que las dos bajadas valen ambas -1 y son, en magnitud, de las más pequeñas de la muestra. El contraste del signo tira esa información y solo ve "ocho contra dos"; Wilcoxon ve además que los dos votos en contra son de los más flojos y los ocho a favor incluyen los más rotundos.

(d) Cuánto cuesta esa pérdida. La eficiencia relativa asintótica del contraste del signo frente a la t con datos normales es $2/\pi = 0{,}636\,6$: hace falta un 57 por ciento más de datos para conseguir la misma potencia. Frente a Wilcoxon, cuya eficiencia es $3/\pi = 0{,}955$, el del signo sale claramente peor parado. Su virtud no es la potencia, sino que no exige ni siquiera simetría: vale para datos ordinales en los que la única operación permitida sea comparar.

Ejemplo 21.17 (La tabla de valores críticos, obtenida contando)**.** El valor crítico bilateral al 5 % es el menor s tal que $2P(S^+ \geq s) \leq 0{,}05$; se rechaza cuando $S^+ \geq s$ o cuando $S^+ \leq m - s$. Enumerando coeficientes binomiales:

m	6	7	8	9	10	11	12	13	14	15
s crítico	6	7	8	8	9	10	10	11	12	12
nivel real	0,031 3	0,015 6	0,007 8	0,039 1	0,021 5	0,011 7	0,038 6	0,022 5	0,012 9	0,035 2

Dos lecturas. La primera: con $m = 5$ la tabla ni siquiera empieza, porque el p-valor bilateral más pequeño posible es $2/2^5 = 0{,}0625$, mayor que 0,05; con cinco pares es imposible rechazar al 5 % aunque las cinco diferencias tengan el mismo signo. La segunda: el nivel real oscila entre el 0,8 y el 3,9 por ciento y nunca es 0,05, porque la distribución es discreta y los p-valores solo pueden tomar los valores de la tabla. Esa granularidad no es un defecto del método: es lo que ocurre siempre que se hace inferencia exacta con pocos datos.

Regla: Aplicar el contraste del signo

1. **Comprueba que los datos están apareados** —dos medidas del mismo individuo, o dos individuos emparejados— y calcula las diferencias en un orden fijado de antemano.
2. **Descarta las diferencias exactamente nulas** y reduce m en consecuencia. Con muchos ceros el método pierde casi toda su potencia, y conviene decir cuántos hubo.
3. **Cuenta los positivos**, que es el estadístico, y compara con $m/2$.
4. **Con $m \leq 20$ usa la binomial exacta**; el p-valor bilateral es el doble de la cola, y no hace falta más que una tabla de coeficientes binomiales.
5. **Con $m > 20$ tipifica**: $Z = (S^+ - m/2)/\sqrt{m/4}$, con corrección de continuidad de 0,5 en la dirección de la media.
6. **Di qué has contrastado**: la hipótesis nula es que la *mediana* de las diferencias es cero, no que su media lo sea. Con distribuciones asimétricas media y mediana no coinciden y las dos afirmaciones son distintas.
7. **Si las magnitudes son fiables, no uses este contraste**: usa Wilcoxon, que aprovecha también el tamaño de las diferencias, o la t si te crees la normalidad.

Cuidado

[Lo que el contraste del signo no dice] Es habitual leer un contraste del signo significativo como "el tratamiento mejora en promedio". No dice eso. Dice que más de la mitad de las diferencias son positivas, que es una afirmación sobre la mediana y no sobre la media. Un tratamiento que mejore ligerísimamente a nueve pacientes de cada diez y hunda catastróficamente al décimo dará un contraste del signo rotundamente significativo y podría tener media negativa. Cuando lo que importa es el efecto total —coste económico, años de vida ganados— la mediana no es la magnitud relevante y el contraste del signo, por muy válido que sea, contesta a otra pregunta.

Tablas de contingencia: de dónde salen las frecuencias esperadas

Definición 21.18 (Tabla de contingencia)**.** Sean dos variables categóricas observadas sobre los mismos n individuos, la primera con f categorías y la segunda con c. La *tabla de contingencia* $f \times c$ es la matriz de recuentos $O_{ij} = \#\{\text{individuos de la categoría } i \text{ de la primera y } j \text{ de la segunda}\}$. Sus *marginales* son

$$O_{i\cdot} = \sum_{j=1}^{c} O_{ij}, \qquad O_{\cdot j} = \sum_{i=1}^{f} O_{ij}, \qquad \sum_{i,j} O_{ij} = n.$$

Proposición 21.19 (Las frecuencias esperadas bajo independencia)**.** *Si las dos variables son independientes, con $P(\text{fila } i) = p_i$ y $P(\text{columna } j) = q_j$, entonces $\mathrm{E}[O_{ij}] = n\,p_i q_j$. Sustituyendo*

los parámetros desconocidos por sus estimadores de máxima verosimilitud $\hat{p}_i = O_{i\cdot}/n$ *y* $\hat{q}_j = O_{\cdot j}/n$ *se obtiene la frecuencia esperada que se usa en el estadístico de Pearson:*

$$E_{ij} = n\,\hat{p}_i\hat{q}_j = \frac{O_{i\cdot}\,O_{\cdot j}}{n}.$$

Los grados de libertad del contraste son $(f-1)(c-1)$.

Demostración. Bajo independencia, la probabilidad de caer en la casilla (i,j) es el producto p_iq_j, y el número de individuos que caen en ella es una binomial de parámetros n y p_iq_j, cuya media es np_iq_j. La fórmula "fila por columna partido por el total" es solo eso con las probabilidades estimadas. Para los grados de libertad se cuenta: hay fc casillas, sujetas a una restricción —que sumen n—, lo que deja $fc-1$; y se han estimado $f-1$ probabilidades de fila (la última sale de que sumen uno) y $c-1$ de columna. Restando,

$$fc - 1 - (f-1) - (c-1) = fc - f - c + 1 = (f-1)(c-1). \qquad \square$$

Ejemplo 21.20 (Una tabla 2×2, con las esperadas calculadas a mano)**.** Se pregunta a 120 personas de dos ciudades si usan el transporte público:

	sí	no	total
ciudad A	42	18	60
ciudad B	28	32	60
total	70	50	120

Las esperadas salen de la proposición 21.19: $E_{11} = 60\cdot 70/120 = 35$, $E_{12} = 60\cdot 50/120 = 25$, y por simetría de los marginales $E_{21} = 35$ y $E_{22} = 25$. Entonces

$$\chi^2 = \frac{(42-35)^2}{35} + \frac{(18-25)^2}{25} + \frac{(28-35)^2}{35} + \frac{(32-25)^2}{25} = 1{,}4 + 1{,}96 + 1{,}4 + 1{,}96 = 6{,}72,$$

con $(2-1)(2-1) = 1$ grado de libertad y $p = 0{,}009\,5$: se rechaza la independencia. La corrección de continuidad de Yates da $\chi^2_Y = 5{,}794\,3$ con $p = 0{,}016\,1$, y el contraste exacto de Fisher $p = 0{,}015\,7$; los tres coinciden en el veredicto, y aquí las esperadas son todas mayores que 25, de modo que la aproximación asintótica es de fiar. La fuerza de la asociación la mide la razón de momios, $\mathrm{OR} = (42\cdot 32)/(18\cdot 28) = 2{,}666\,7$. El resto del procedimiento —cuándo vale la aproximación, cómo mirar los residuos— está en la regla y los ejemplos de la sección de bondad de ajuste; lo que aquí se ha añadido es de dónde salen las E_{ij} y por qué los grados de libertad son los que son.

21.3. Bootstrap y remuestreo

Definición 21.21. El *bootstrap* estima la distribución muestral de un estadístico $\hat{\theta}$ generando B muestras del mismo tamaño extraídas *con reemplazamiento* de la muestra original, calculando $\hat{\theta}$ en cada una y usando la dispersión de esos B valores como estimación del error estándar. El *intervalo percentil* al 95 % toma los cuantiles 2,5 y 97,5 de esa colección.

Idea clave

Qué dice esto en cristiano. La idea es de una audacia notable: si no se conoce la población, se usa la muestra *como si fuera* la población y se remuestrea de ella. La lógica es que la relación entre la muestra y la población se parece a la relación entre una remuestra y la muestra. Lo que hace revolucionaria a la técnica es que funciona para *cualquier* estadístico, incluidos aquellos para los que no existe fórmula del error estándar: la mediana, un cuantil, una correlación, la razón de dos medias. Donde antes hacía falta un teorema, ahora basta con un bucle.

Ejemplo 21.22 (Bootstrap sobre una muestra con un atípico)**. (a) La media.** Con la muestra $12, 15, 15, 17, 18, 18, 19, 21, 22, 24, 25, 28, 31, 35, 64$ y $B = 20\,000$ remuestras, el error estándar bootstrap de la media es 3,157220, frente al valor clásico $\frac{s}{\sqrt{n}} = 3{,}273947$. Coinciden razonablemente, como debe ocurrir en un caso donde sí existe fórmula.

(b) El intervalo percentil. Los cuantiles 2,5 y 97,5 de las medias remuestreadas dan [19,07; 31,27]. Nótese que no es simétrico respecto de la media muestral 24,27: el bootstrap recoge la asimetría que el intervalo clásico, simétrico por construcción, ignora.

(c) La mediana, donde no hay fórmula. El error estándar bootstrap de la mediana es 2,584129 y su intervalo [18; 28]. No existe ninguna expresión cerrada sencilla para el error estándar de la mediana, y aquí se obtiene con el mismo bucle que para la media. Ese es el argumento decisivo a favor del método.

(d) Cuántas remuestras hacen falta. Para estimar un error estándar bastan unas 200; para intervalos de confianza conviene $B \geq 2000$, y para percentiles extremos, más. El coste es lineal en B y hoy es despreciable, lo que explica que el bootstrap fuera imposible antes de 1979 y rutinario después.

Cuidado

[El bootstrap no crea información] Remuestrear no añade datos: si la muestra original es pequeña o no representativa, el bootstrap reproducirá fielmente ese defecto. Con $n = 5$ solo hay 126 remuestras distintas, y el método da resultados poco fiables por muchas iteraciones que se hagan. Segundo aviso: el bootstrap ordinario falla para estadísticos que dependen de las colas, como el máximo muestral, porque la remuestra nunca puede superar el máximo observado. Y tercero: exige independencia; con series temporales o datos agrupados hay que usar variantes por bloques que preserven la estructura de dependencia.

Pausa para pensar

El bootstrap lo introdujo Bradley Efron en 1979, y el nombre viene de la expresión inglesa sobre levantarse tirando de los propios cordones, es decir, hacer algo aparentemente imposible sin ayuda externa. La reacción inicial fue de escepticismo, porque parecía sacar información de la nada; lo que en realidad hace es sustituir un teorema asintótico por fuerza bruta computacional, y su justificación teórica se completó en la década

siguiente. Es probablemente el mejor ejemplo de cómo el ordenador ha cambiado no solo la práctica sino la propia estructura de la estadística: hay métodos que ahora son elementales y que antes eran, sencillamente, impensables.

21.4. Bondad de ajuste y tablas de contingencia

Los contrastes de rangos de la sección anterior comparan *muestras*. Los de esta sección comparan una muestra con un *modelo*, o dos clasificaciones entre sí. Son los contrastes no paramétricos más usados en la práctica, y también los peor usados: casi todo el mundo sabe calcular un χ^2 y casi nadie recuerda cuándo deja de valer.

Definición 21.23. Sea una muestra clasificada en k categorías con frecuencias observadas $O_1, \dots, O_k$ y frecuencias esperadas bajo la hipótesis nula $E_1, \dots, E_k$ (con $\sum O_i = \sum E_i = n$). El *estadístico chi-cuadrado de Pearson* es

$$\chi^2 = \sum_{i=1}^{k} \frac{(O_i - E_i)^2}{E_i},$$

y bajo H_0 tiende en distribución a una χ^2 con $k - 1 - r$ grados de libertad, donde r es el número de parámetros estimados a partir de la propia muestra.

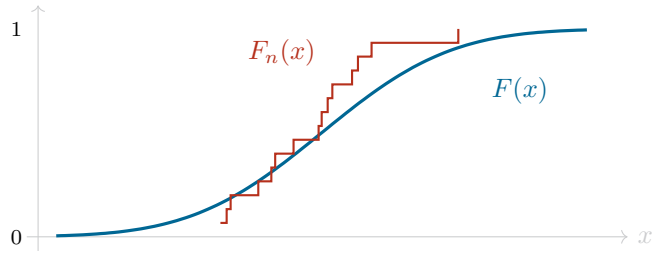


Figura 21.1: Función de distribución empírica de una muestra de tamaño 15 frente a la teórica. El estadístico de Kolmogorov–Smirnov es la mayor separación vertical entre ambas, aquí $D = 0{,}246$.

Idea clave

Qué dice esto en cristiano. Cada sumando pregunta "¿cuánto se desvía esta casilla de lo previsto, medido en unidades de lo que cabe esperar que se desvíe?". La división por E_i es la clave: una discrepancia de 5 casos es escandalosa si esperábamos 3 y es trivial si esperábamos 500. Al sumar todas las casillas se obtiene una medida global de desacuerdo, que es pequeña cuando el modelo funciona y crece cuando falla. Los grados de libertad descuentan una unidad por la restricción $\sum O_i = n$ y una más por cada parámetro que se haya estimado espiando los datos.

Regla: Aplicar un contraste χ^2 sin meter la pata

1. **Trabajar con frecuencias, nunca con porcentajes.** El estadístico depende del tamaño de muestra; alimentarlo con porcentajes equivale a declarar $n = 100$.
2. **Calcular las E_i bajo H_0**: en bondad de ajuste, $E_i = n\,p_i$ con las p_i del modelo; en independencia, $E_{ij} = (\text{total fila } i)(\text{total columna } j)/n$.
3. **Comprobar que las E_i son suficientemente grandes**: la regla habitual pide $E_i \geq 5$ en todas las casillas, o al menos en el 80 % de ellas con ninguna por debajo de 1. Si no, agrupar categorías o usar un contraste exacto.
4. **Contar los grados de libertad**: $k-1$ en bondad de ajuste sin parámetros estimados, $k-1-r$ si se estimaron r, y $(f-1)(c-1)$ en una tabla $f \times c$.
5. **Comparar con el valor crítico $\chi^2_{1-\alpha}$**, siempre por la cola derecha: un χ^2 *pequeño* no es evidencia contra H_0, es acuerdo.
6. **Si se rechaza, mirar los residuos** $(O_i - E_i)/\sqrt{E_i}$ para saber *qué* casillas causaron el rechazo. El χ^2 global no lo dice.
7. **No confundir significación con asociación fuerte**: con n grande cualquier desviación mínima resulta significativa. Para medir la fuerza, usar la V de Cramér o la razón de momios.

Ejemplo 21.24 (Cuatro contrastes de bondad de ajuste e independencia)**. (a) ¿Está cargado el dado?** En 60 tiradas salen las frecuencias $8, 12, 9, 14, 7, 10$. Bajo el modelo de dado equilibrado, $E_i = 10$ para las seis caras, y

$$\chi^2 = \frac{4+4+1+16+9+0}{10} = 3{,}4.$$

Con 5 grados de libertad el valor crítico al 5 % es 11,070498 y el p-valor es 0,638570. No hay ninguna evidencia de carga: la irregularidad observada es exactamente la que produce el azar.

(b) Una tabla 2×2. Cien personas clasificadas por tratamiento y resultado:

	éxito	fracaso	total
A	30	20	50
B	15	35	50
total	45	55	100

Las esperadas son 22,5 y 27,5 en las dos filas, y $\chi^2 = 9{,}090909$ con 1 grado de libertad, $p = 0{,}002569$. Se rechaza la independencia. La razón de momios mide la fuerza: $\text{OR} = \frac{30\cdot 35}{20\cdot 15} = 3{,}5$, es decir, los momios de éxito con el tratamiento A son tres veces y media los de B.

(c) La corrección de Yates y el exacto de Fisher. Sobre la misma tabla, la corrección de continuidad $\sum(|O-E| - 0{,}5)^2/E$ da $\chi^2_Y = 7{,}919192$ con $p = 0{,}004891$, y el contraste exacto de Fisher, que enumera todas las tablas con los mismos márgenes, da $p = 0{,}004635$. Los tres p-valores coinciden en el veredicto; el de Fisher es el único exacto y es el que hay que usar cuando alguna esperada baja de 5.

(d) Una tabla 3×2. Con

	sí	no
I	20	15
II	25	10
III	10	20

las esperadas son 19,25 y 15,75 en las dos primeras filas y 16,5 y 13,5 en la tercera, dando $\chi^2 = 9{,}571910$ con $(3-1)(2-1) = 2$ grados de libertad, valor crítico 5,991465 y $p = 0{,}008346$. Los residuos estandarizados muestran que el grupo III es el culpable: aporta $(10-16{,}5)/\sqrt{16{,}5} = -1{,}600$ y $(20-13{,}5)/\sqrt{13{,}5} = 1{,}769$, mientras que las otras dos filas se quedan por debajo de 1,5 en valor absoluto (0,171 y $-0{,}189$ la primera, 1,311 y $-1{,}449$ la segunda).

Definición 21.25. Dada una muestra $x_1 \le \cdots \le x_n$ con función de distribución empírica F_n, el *estadístico de Kolmogórov–Smirnov* para el ajuste a una distribución continua F es

$$D_n = \sup_x |F_n(x) - F(x)| = \max_{1 \le i \le n} \max\left\{ \tfrac{i}{n} - F(x_i),\ F(x_i) - \tfrac{i-1}{n} \right\}.$$

Idea clave

Qué dice esto en cristiano. En vez de trocear el recorrido en categorías, como hace el χ^2, se compara la curva acumulada de los datos con la curva acumulada del modelo y se mide la mayor separación vertical entre ambas. Es un contraste más fino para distribuciones continuas porque no pierde información al agrupar, y tiene una propiedad notable: bajo H_0 la distribución de D_n *no depende de* F, porque la transformación $u = F(x)$ convierte cualquier problema en el caso uniforme.

Ejemplo 21.26 (Kolmogórov–Smirnov contra la uniforme)**.** Con la muestra ordenada 0,12; 0,31; 0,44; 0,58; 0,62; 0,79; 0,85; 0,91 y $F(x) = x$ en $[0,1]$:

(a) El salto por arriba. $D^+ = \max_i(i/n - x_i)$. Los valores $i/8$ son 0,125; 0,25; 0,375; 0,5; 0,625; 0,75; 0,875; 1 y las diferencias 0,005; $-0{,}06$; $-0{,}065$; $-0{,}08$; 0,005; $-0{,}04$; 0,025; 0,09, de donde $D^+ = 0{,}09$ en la última observación.

(b) El salto por abajo. $D^- = \max_i(x_i - (i-1)/8)$ da 0,12; 0,185; 0,19; 0,205; 0,12; 0,165; 0,1; 0,035, luego $D^- = 0{,}205$, alcanzado en $x_4 = 0{,}58$.

(c) El veredicto. $D_8 = \max(0{,}09;\ 0{,}205) = 0{,}205$. El valor crítico aproximado al 5 % es $1{,}358/\sqrt{8} = 0{,}480126$, muy por encima; el p-valor exacto es 0,827482. La muestra es perfectamente compatible con la uniforme.

(d) Dónde está el máximo. Conviene notar que D_n se alcanza siempre en un dato, nunca entre datos, porque F_n es constante a trozos y F es creciente: entre dos saltos consecutivos la diferencia $F_n - F$ sólo puede decrecer. Por eso el supremo, que en principio recorre toda la recta, se reduce a un máximo sobre $2n$ números.

Cuidado

[Contraejemplo: el χ^2 con esperadas pequeñas miente] Considérese una tabla 2×2 con observadas $\left(\begin{smallmatrix} 4 & 0 \\ 1 & 5 \end{smallmatrix}\right)$. Las esperadas son 2 y 2 en la primera fila y 3 y 3 en la segunda, todas por debajo de 5. El χ^2 de Pearson vale 6,667 con $p = 0{,}0098$, aparentemente concluyente; el contraste exacto de Fisher da $p = 0{,}0476$, justo en el filo. La aproximación asintótica exagera sistemáticamente la significación cuando las casillas están vacías o casi, porque una distribución discreta muy gruesa se está aproximando por una continua. Con n pequeño, Fisher siempre; el χ^2 es un atajo asintótico, no una definición.

Cuidado

[Contraejemplo: Kolmogórov–Smirnov con parámetros estimados] Si se estima la media y la varianza a partir de la propia muestra y luego se contrasta la normalidad con las tablas usuales de D_n, el contraste se vuelve gravemente conservador: rechaza mucho menos de lo que debería, porque la curva ajustada persigue a los datos y las separaciones se encogen artificialmente. Los valores críticos correctos en ese caso son los de Lilliefors, sensiblemente menores. Un D_n que no llega al valor crítico estándar puede perfectamente superar el de Lilliefors y llevar al rechazo.

Pausa para pensar

El χ^2 de Pearson data de 1900 y su distribución asintótica fue de las primeras demostraciones serias de la estadística moderna; Pearson, sin embargo, se equivocó en los grados de libertad cuando había parámetros estimados, y fue Fisher quien lo corrigió en 1922 estableciendo la resta de r. La polémica entre ambos duró décadas y fue tan áspera como fértil. La moraleja técnica es sencilla y conviene recordarla: cada parámetro que se estima mirando los datos hace que el modelo se acerque a ellos gratis, y ese acercamiento hay que pagarlo con un grado de libertad menos.

21.5. Regresión sin modelo: suavizado y splines

Ajustar una recta supone que la relación es una recta. Cuando no lo es, y no se sabe qué es, hay que dejar que los datos decidan la forma. El precio vuelve a ser el mismo compromiso entre sesgo y varianza, ahora con un parámetro de suavizado en el papel protagonista.

Definición 21.27. El *estimador de Nadaraya–Watson* de la función de regresión $m(x) = E[Y \mid X = x]$ es el promedio ponderado

$$\hat{m}_h(x) = \frac{\sum_i K\left(\frac{x - x_i}{h}\right) y_i}{\sum_i K\left(\frac{x - x_i}{h}\right)}.$$

Un *spline de suavizado* es, en cambio, la función que minimiza

$$\sum_i \big(y_i - f(x_i)\big)^2 + \lambda \int \big(f''(t)\big)^2 \mathrm{d}t,$$

y resulta ser siempre un spline cúbico natural con nudos en los datos.

Idea clave

Qué dice esto en cristiano. El primer método dice: para predecir en x, promedia las respuestas de los puntos cercanos, dando más peso a los más próximos. El segundo dice: busca la curva que pase razonablemente cerca de los datos sin retorcerse demasiado, y castiga el retorcimiento con la integral de la segunda derivada, que es exactamente la curvatura acumulada. Los dos parámetros, la ventana h y la penalización λ, hacen el mismo papel: con $\lambda \to 0$ el spline interpola todos los puntos y tiembla; con $\lambda \to \infty$ la penalización obliga a $f'' = 0$ y sale la recta de mínimos cuadrados de toda la vida. La regresión lineal es el caso límite de infinito suavizado.

Ejemplo 21.28 (Los tres métodos y sus rarezas)**. (a) El sesgo de frontera.** Nadaraya–Watson se porta mal en los extremos del rango, porque allí sólo hay vecinos por un lado y el promedio se sesga hacia el interior. La regresión lineal local —ajustar una recta ponderada en vez de una constante— corrige ese sesgo automáticamente, y por eso el método LOWESS de Cleveland, de 1979, se convirtió en el estándar.

(b) Grados de libertad efectivos. Un spline de suavizado no tiene un número entero de parámetros, pero sí un análogo: la traza de la matriz de suavizado. Con ella se pueden comparar modelos y aplicar criterios como la validación cruzada generalizada, que elige λ automáticamente.

(c) Por qué salen cúbicos. No se supone que la solución sea un spline: se demuestra. Entre todas las funciones dos veces derivables que produzcan un valor dado del ajuste, la que minimiza $\int (f'')^2$ es necesariamente un spline cúbico con nudos en los datos y lineal fuera del rango. La forma funcional cae del problema variacional, no de una elección.

(d) El compromiso, otra vez. La velocidad de convergencia óptima vuelve a ser $n^{-4/5}$ para funciones dos veces derivables, y $n^{-2p/(2p+1)}$ si se supone derivabilidad de orden p. Cuanta más suavidad se esté dispuesto a suponer, más rápido se aprende. No suponer nada tiene un precio, y es exactamente ese exponente.

Cuidado

[Contraejemplo: suavizar puede fabricar estructura] Un suavizador aplicado a datos puramente aleatorios produce siempre una curva con ondulaciones que parecen tendencias. Con $n = 50$ puntos de ruido blanco y una ventana pequeña, casi siempre aparecen dos o tres jorobas convincentes, y el ojo humano las interpreta como ciclos. El remedio no es dejar de suavizar, sino acompañar la curva de una banda de confianza —que en esos casos resulta abarcar cómodamente la recta horizontal— y desconfiar de toda estructura que dependa de la elección de h.

21.6. Estimar una densidad sin suponer su forma

Un histograma es una estimación de densidad, y bastante mala: depende de dónde se pongan las barras y produce una función escalonada cuando la verdadera casi nunca lo es. Sustituirlo por algo mejor obliga a enfrentarse al compromiso más característico de la estadística no paramétrica, el que hay entre sesgo y varianza.

Definición 21.29. Dada una muestra $x_1, \dots, x_n$, el *estimador de densidad por núcleos* es

$$\hat{f}_h(x) = \frac{1}{nh} \sum_{i=1}^{n} K\Big(\frac{x - x_i}{h}\Big),$$

donde el *núcleo* K es una densidad simétrica y $h > 0$ es la *ventana* o parámetro de suavizado.

Idea clave

Qué dice esto en cristiano. En vez de contar cuántos datos caen en cada caja, se pone una campanita encima de cada dato y se suman todas. El resultado es suave, no depende de dónde empiece la primera caja y se parece mucho más a lo que uno dibujaría a mano. La ventana h dice cuán anchas son las campanitas, y en ella está todo el problema: si es pequeña, cada dato hace su propio pico y la estimación es fiel pero temblorosa; si es grande, todo se funde en una loma suave que ya no distingue nada. La ventana no es un detalle técnico, es la decisión que determina el resultado, y la elección del núcleo apenas importa en comparación.

Ejemplo 21.30 (Elegir la ventana, con números)**. (a) La regla práctica.** La regla de Silverman toma $h = 1{,}06\,\hat{\sigma}\,n^{-1/5}$. Con $n = 100$ y $\hat{\sigma} = 1$ da $h = 0{,}422$; con $n = 1000$, $h = 0{,}266$. La ventana se estrecha al crecer la muestra, pero muy despacio: hace falta multiplicar n por 32 para reducirla a la mitad.

(b) El compromiso. El sesgo del estimador es del orden de h^2 y su varianza del orden de $1/(nh)$. Minimizar la suma da $h \asymp n^{-1/5}$, con error óptimo del orden de $n^{-4/5}$.

(c) El precio de no suponer nada. Si uno acepta que la densidad es normal, estimar μ y σ da un error del orden de n^{-1}. La estimación no paramétrica converge estrictamente más despacio, y ésa es la factura por no comprometerse con ninguna forma. A cambio, si la suposición de normalidad era falsa, el estimador paramétrico converge deprisa hacia la respuesta equivocada.

(d) La maldición de la dimensión. En d dimensiones el error óptimo pasa a ser $n^{-4/(4+d)}$. Con $d = 1$ es $n^{-0{,}8}$; con $d = 10$, $n^{-0{,}286}$. Para conseguir en diez dimensiones la precisión que en una se logra con cien datos harían falta del orden de 10^7. Ésa es la razón técnica de que en dimensión alta no se estimen densidades.

Cuidado

[Contraejemplo: la validación cruzada puede colapsar] El método honesto para elegir h es la validación cruzada por verosimilitud, que maximiza $\sum_i \log \hat{f}_{h,-i}(x_i)$ dejando fuera

cada dato. Con datos que tengan valores repetidos —o con colas pesadas— ese criterio empuja h hacia cero, porque un pico infinitamente estrecho sobre un dato duplicado da verosimilitud infinita. El procedimiento no diverge: converge alegremente a una respuesta absurda, un peine de deltas.

21.7. Contrastes de permutación: inferencia exacta sin ninguna distribución

Hay un contraste que no supone normalidad, ni simetría, ni tamaño muestral grande, ni siquiera que exista la varianza. Solo supone una cosa, y es la propia hipótesis nula.

Idea clave

Si el tratamiento no hace nada, entonces las etiquetas "grupo A" y "grupo B" están pegadas a los datos al azar. Cualquier otra asignación de etiquetas habría sido igual de probable. Luego, para saber si la diferencia observada es grande, basta con calcular la diferencia que habría salido con *todas* las reasignaciones posibles y ver en qué percentil cae la real. No hace falta ninguna distribución teórica: la distribución de referencia se fabrica con los propios datos.

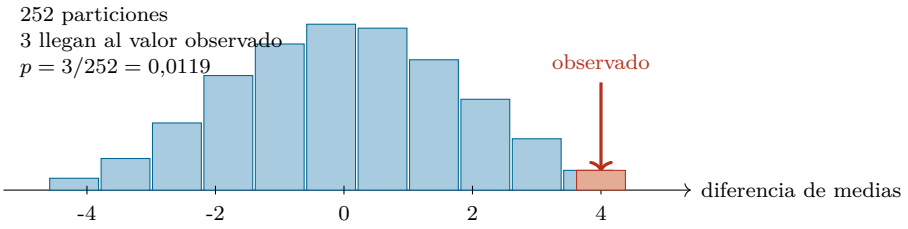


Figura 21.2: Las $\binom{10}{5} = 252$ maneras de repartir los diez datos en dos grupos de cinco, con la diferencia de medias que produce cada una. La distribución de referencia no viene de ninguna teoría: se fabrica con los propios datos. Solo tres particiones alcanzan la diferencia observada, y ese recuento *es* el p-valor exacto.

Definición 21.31 (Contraste de permutación)**.** Bajo H_0 de **intercambiabilidad**, se calcula el estadístico T para las $\binom{n_1+n_2}{n_1}$ particiones posibles de los datos en dos grupos de los tamaños dados. El p-valor exacto es la proporción de particiones con T tan extremo o más que el observado.

Ejemplo 21.32 (Cinco y cinco, con las 252 particiones contadas una a una)**.** Grupo A: $12, 15, 14, 18, 16$. Grupo B: $9, 11, 13, 10, 12$.

(a) El estadístico observado. $\bar{A} = 15$, $\bar{B} = 11$, luego $T_{\text{obs}} = 4$.

(b) La distribución de referencia. Hay $\binom{10}{5} = 252$ maneras de repartir los diez números en dos grupos de cinco. Calculando la diferencia de medias en las 252:

$$\#\{T \geq 4\} = 3, \qquad p_{\text{unilateral}} = \frac{3}{252} = 0{,}011\,905.$$

Bilateralmente, $\#\{|T| \geq 4\} = 6$ y $p = 6/252 = 0{,}023\,810$.

(c) Comparación con la t. La t de Student da $t = 3{,}266$ con 8 grados de libertad y $p = 0{,}011\,4$. Prácticamente el mismo número. Pero el de permutación es *exacto*: no se apoya en ninguna aproximación, y sería igual de válido si los datos vinieran de una distribución con colas monstruosas.

(d) La granularidad. Con solo 252 particiones, los p-valores posibles son múltiplos de $1/252 = 0{,}003\,97$. No se puede obtener $p = 0{,}001$ ni aunque la separación fuera total: el p-valor mínimo alcanzable unilateralmente es $1/252$. Esa granularidad es la firma de los contrastes exactos con muestras pequeñas.

Regla: permutación en la práctica

1. Con n pequeño, enumera todas las particiones: el resultado es exacto.
2. Con n grande, $\binom{2n}{n}$ es astronómico. Se toman B permutaciones al azar (mil, diez mil) y se estima p con $(1 + \#\{T^* \geq T_{\text{obs}}\})/(B+1)$. El $+1$ del numerador y del denominador no es cosmética: garantiza que el contraste siga siendo válido, y evita el absurdo de un p-valor cero.
3. El error de Monte Carlo de esa estimación es $\sqrt{p(1-p)/B}$: con $B = 1000$ y $p \approx 0{,}05$, unos 0,007. Si el p-valor está cerca del umbral, sube B.
4. El estadístico puede ser *cualquiera*: diferencia de medias, de medianas, de recortadas, t, incluso el coeficiente de una regresión. No hace falta que sea pivotal ni que su distribución se conozca.

Cuidado

Permutar y remuestrear no son lo mismo, y confundirlos es un error de fondo. **Permutar impone la hipótesis nula** (baraja las etiquetas, con lo que fabrica un mundo donde el tratamiento no hace nada) y sirve para contrastar. **El bootstrap no impone nada** (remuestrea con reemplazamiento del mundo tal como es) y sirve para estimar la variabilidad y construir intervalos. Usar el bootstrap para contrastar sin recentrar los datos bajo H_0 es uno de los errores más frecuentes en la literatura aplicada.
La otra advertencia es sobre la hipótesis: la intercambiabilidad es más fuerte que "las medias son iguales". Si los dos grupos tienen la misma media pero varianzas muy distintas, las etiquetas no son intercambiables y el contraste de permutación puede rechazar por esa razón, no por la media. Para el problema de Behrens–Fisher, permutar no es la respuesta.

Observación 21.33. El mismo mecanismo sirve para contrastar independencia entre dos variables: se dejan quietos los x y se permutan los y, se recalcula la correlación en cada permutación, y se compara. Y sirve en regresión, permutando los residuos del modelo reducido. Fisher lo propuso en 1935 como la justificación *lógica* de la t: la t es una aproximación cómoda al contraste de permutación, no al revés. Que durante medio siglo se usara la aproximación en lugar del original fue una cuestión de aritmética, no de teoría; con ordenadores, esa razón desapareció.

21.8. Comparar k grupos sin normalidad: Kruskal–Wallis y Friedman

Wilcoxon y Mann–Whitney resuelven el caso de dos grupos. Para k grupos, el ANOVA tiene su equivalente por rangos.

Definición 21.34 (Kruskal–Wallis)**.** Se ordenan conjuntamente las N observaciones de los k grupos y se sustituyen por sus rangos. Con R_i la suma de rangos del grupo i, de tamaño n_i:

$$H = \frac{12}{N(N+1)} \sum_{i=1}^{k} \frac{R_i^2}{n_i} - 3(N+1) \ \xrightarrow{d}\ \chi^2_{k-1}.$$

Con empates se divide H por $1 - \dfrac{\sum (t_j^3 - t_j)}{N^3 - N}$, donde t_j son los tamaños de los grupos de empatados.

Ejemplo 21.35 (Los mismos tres tratamientos, por rangos)**.** Con los datos de los tres tratamientos (A: $55, 58, 52, 60, 55$; B: $64, 68, 61, 70, 67$; C: $72, 75, 70, 78, 75$), el ANOVA paramétrico daba $F = 38{,}73$ con $p = 0{,}000\,006$. Por rangos:

$$H = 12{,}321\,0, \qquad 2 \text{ grados de libertad}, \qquad p = 0{,}002\,1.$$

La conclusión es la misma —los tres grupos difieren— pero el p-valor es tres órdenes de magnitud mayor. No es que Kruskal–Wallis sea peor: es que con cinco datos por grupo la aproximación χ^2 es tosca y, sobre todo, los rangos tiran información. Con $N = 15$ el rango máximo es 15, así que por muy separados que estén los grupos, H tiene un techo: $H_{\text{máx}} = 12{,}5$ con este diseño. Estamos en 12,32, es decir, prácticamente en el máximo posible. El contraste está saturado.

Ahí está la moraleja: los métodos de rangos pagan un precio en resolución que solo se nota cuando la separación es enorme o la muestra minúscula. Su eficiencia relativa asintótica frente al F bajo normalidad es $3/\pi = 0{,}955$ —pierden un 4,5 por ciento—, y frente a distribuciones de colas pesadas puede ser mayor que 1: entonces ganan.

Definición 21.36 (Friedman, para bloques)**.** Con b bloques y k tratamientos, se ordena *dentro de cada bloque* (rangos de 1 a k) y, con R_j la suma de rangos del tratamiento j:

$$Q = \frac{12}{bk(k+1)} \sum_{j=1}^{k} R_j^2 - 3b(k+1) \ \xrightarrow{d}\ \chi^2_{k-1}.$$

Friedman es a las medidas repetidas lo que Kruskal–Wallis es al ANOVA de un factor, y su idea es la misma que la del análisis apareado: comparar dentro de cada bloque anula la variabilidad entre bloques. Si tres jueces puntúan cinco vinos, lo que importa es el orden que cada juez establece, no la escala personal de cada uno.

Regla: después de rechazar

Igual que con el F, un H significativo no dice qué grupos difieren. La comparación posterior estándar es la de **Dunn**: para cada par,

$$z_{ij} = \frac{\bar{R}_i - \bar{R}_j}{\sqrt{\dfrac{N(N+1)}{12}\left(\dfrac{1}{n_i} + \dfrac{1}{n_j}\right)}},$$

con el nivel corregido por Bonferroni sobre las $\binom{k}{2}$ comparaciones. Nótese que se usan los rangos *del análisis conjunto*, no rangos recalculados en cada pareja: recalcularlos es un error frecuente y cambia el resultado.

21.9. Intervalos bootstrap que sí funcionan, y la validación cruzada

El intervalo percentil es el primero que se aprende y el que peor se comporta. Conviene conocer los otros tres.

Regla: los cuatro intervalos bootstrap

Sea $\hat{\theta}$ el estimador, $\hat{\theta}^*_{(\alpha)}$ el cuantil α de las réplicas bootstrap.

- **Percentil**: $[\hat{\theta}^*_{(0{,}025)},\ \hat{\theta}^*_{(0{,}975)}]$. Intuitivo y sesgado.
- **Básico o de pivote**: $[2\hat{\theta} - \hat{\theta}^*_{(0{,}975)},\ 2\hat{\theta} - \hat{\theta}^*_{(0{,}025)}]$. Refleja el intervalo alrededor de $\hat{\theta}$; corrige el sesgo de primer orden. Puede salirse del rango admisible del parámetro.
- **Bootstrap-t**: se studentiza cada réplica, $t^* = (\hat{\theta}^* - \hat{\theta})/\mathrm{ee}^*$, y se usan los cuantiles de t^*. Es el de mejor cobertura, pero exige un error típico para cada réplica (a veces un bootstrap anidado).
- **BCa**: corrige sesgo y asimetría con dos constantes,

$$\hat{z}_0 = \Phi^{-1}\left(\frac{\#\{\hat{\theta}^* < \hat{\theta}\}}{B}\right), \qquad \hat{a} = \frac{\sum(\bar{\theta}_{(\cdot)} - \hat{\theta}_{(i)})^3}{6\left[\sum(\bar{\theta}_{(\cdot)} - \hat{\theta}_{(i)})^2\right]^{3/2}},$$

  donde $\hat{\theta}_{(i)}$ son las estimaciones jackknife (dejando fuera la observación i). Es el que recomiendan por defecto Efron y Tibshirani.

Idea clave

Las dos constantes del BCa tienen lectura directa. $\hat{z}_0$ mide cuántas réplicas caen por debajo de la estimación original: si es exactamente la mitad, $\hat{z}_0 = 0$ y no hay sesgo que corregir. $\hat{a}$ mide la asimetría a través del jackknife: si es cero, la varianza del estimador no cambia con el valor del parámetro. Cuando ambas son nulas, el BCa *es* el percentil.

Es decir, el percentil no está mal: está incompleto, y el BCa le añade justo lo que le falta.

Regla: bootstrap en regresión: dos maneras, y no dan lo mismo

- **Remuestrear residuos**: se fija X, se ajusta el modelo, y se generan $y_i^* = \hat{y}_i + e_i^*$ con los residuos remuestreados. Supone el modelo correcto y homocedástico, y a cambio es más eficiente.
- **Remuestrear pares** (x_i, y_i): no supone forma funcional ni varianza constante. Es lo que hay que usar cuando el modelo puede estar mal especificado o hay heterocedasticidad.

Con datos dependientes —series temporales— ninguno de los dos vale, porque remuestrear destruye la estructura de dependencia. Ahí se usa el **bootstrap por bloques**: se remuestrean tramos consecutivos de longitud ℓ, con ℓ creciendo como $n^{1/3}$, de modo que dentro de cada bloque la dependencia se conserva y entre bloques se supone despreciable.

Validación cruzada

Definición 21.37 (Validación cruzada de k particiones)**.** Se parten los datos en k trozos. Para cada j, se ajusta el modelo con los $k-1$ trozos restantes y se evalúa el error sobre el trozo j. El error de validación es el promedio de los k errores. Con $k = n$ se llama *dejar uno fuera*.

El problema que resuelve es que el error de entrenamiento es optimista por construcción: el modelo ya ha visto esos datos, y con suficientes parámetros los reproduce exactamente. Ese optimismo crece con la complejidad, y por eso minimizar el error de entrenamiento siempre elige el modelo más complicado.

Observación 21.38 (Un atajo exacto en regresión lineal)*.* Para modelos lineales no hace falta reajustar n veces. Con h_{ii} el elemento diagonal de la matriz sombrero,

$$\mathrm{CV}_{(n)} = \frac{1}{n}\sum_{i=1}^{n}\Big(\frac{e_i}{1-h_{ii}}\Big)^2,$$

que se calcula con un solo ajuste. Sustituir cada h_{ii} por su media p/n da la validación cruzada generalizada, y de ahí sale el criterio que usan los suavizadores.

Regla: elegir k

- $k = n$ (dejar uno fuera): casi insesgado para el error de predicción, pero de *alta varianza*, porque los n modelos son casi idénticos entre sí y sus errores están muy correlacionados.
- $k = 5$ o $k = 10$: algo sesgado hacia arriba (cada modelo se ajusta con menos datos de los que se tendrán al final) y de mucha menor varianza. Es el estándar, y es lo

que hay que usar salvo razón en contra.

Cuidado

El error más caro de toda la estadística aplicada moderna cabe en una frase: **seleccionar variables mirando todos los datos y luego validar cruzadamente**. Si se eligen, por ejemplo, los veinte predictores más correlacionados con la respuesta usando la muestra entera, y *después* se hace validación cruzada sobre el modelo con esos veinte, el error estimado puede salir espectacularmente bajo aunque los predictores sean ruido puro: la selección ya ha filtrado la información del conjunto de validación. La regla es inflexible: *todo* paso que mire a la variable respuesta —selección, transformación, ajuste de hiperparámetros— tiene que quedar dentro del bucle de validación.

21.10. Intervalos libres de distribución para la mediana y el estimador de Hodges–Lehmann

Hay un intervalo de confianza para la mediana que no supone absolutamente nada sobre la distribución —ni simetría, ni varianza finita, ni continuidad más allá de no tener átomos— y que se construye contando.

Idea clave

Si M es la mediana poblacional, cada observación cae por debajo de M con probabilidad $1/2$, independientemente de las demás. Luego el número de observaciones por debajo de M es $\mathrm{Bin}(n, 1/2)$. El suceso "$M < x_{(k)}$" equivale a "al menos k observaciones están por debajo de M"... y esa probabilidad se lee en una tabla binomial. Todo el intervalo sale de ahí.

Teorema 21.39 (Intervalo para la mediana)**.** *Para* $1 \le k < n + 1 - k$,

$$P\big(x_{(k)} \le M \le x_{(n+1-k)}\big) = 1 - 2\sum_{j=0}^{k-1} \binom{n}{j} 2^{-n}.$$

Ejemplo 21.40 (Doce datos y un 96,1 por ciento que no se puede afinar)**.** Con $n = 12$ y $k = 3$, el intervalo es $[x_{(3)}, x_{(10)}]$ y su confianza exacta es

$$1 - 2\big[\tbinom{12}{0} + \tbinom{12}{1} + \tbinom{12}{2}\big]2^{-12} = 1 - 2 \cdot \frac{1 + 12 + 66}{4096} = 1 - \frac{158}{4096} = 0{,}961\,4.$$

Con $k = 4$ sería $[x_{(4)}, x_{(9)}]$, con confianza $1 - 2 \cdot 299/4096 = 0{,}854\,0$.

No hay nada entre el 96,1 y el 85,4 por ciento: los niveles alcanzables son discretos, porque los extremos tienen que ser datos observados. Esa es la contrapartida de no suponer nada. Se informa el nivel real, no un 95 por ciento inventado.

Definición 21.41 (Estimador de Hodges–Lehmann)**.** Para una muestra $x_1, \dots, x_n$, se forman las $\binom{n+1}{2}$ **medias de Walsh** $(x_i + x_j)/2$ con $i \le j$, y se toma su mediana. Para dos muestras, se toma la mediana de las $n_1 n_2$ diferencias $x_i - y_j$.

Ejemplo 21.42 (Ocho diferencias apareadas)**.** Con las diferencias $d = 2, 2, 3, 1, 3, 2, 2, 2$ de un estudio antes-después, las $\binom{9}{2} = 36$ medias de Walsh son los promedios $(d_i + d_j)/2$ con $i \le j$. Ordenadas, su mediana vale 2: el estimador de Hodges–Lehmann de la mediana de las diferencias es 2 kilos, casi idéntico a la media de 2,125 y a la mediana muestral de 2.

La virtud no es el número, que aquí coincide con todo lo demás, sino el punto de ruptura: $1 - 1/\sqrt{2} \approx 0{,}29$, es decir, resiste que casi un tercio de los datos esté corrompido. La media resiste 0 y la mediana 0,5; Hodges–Lehmann se sitúa en medio con una eficiencia del 95,5 por ciento bajo normalidad, que es casi la de la media.

Idea clave

Hodges–Lehmann no es un estimador cualquiera: es *el* estimador que corresponde al contraste de rangos con signo de Wilcoxon, en el mismo sentido en que $\bar{x}$ corresponde a la t. El intervalo obtenido invirtiendo Wilcoxon —tomando las medias de Walsh entre dos posiciones concretas— es su intervalo de confianza natural. Esa correspondencia entre un contraste y un estimador, vía inversión, es un patrón que recorre toda la estadística: cada contraste esconde un estimador y un intervalo, y conviene saber cuál.

21.11. De dónde salen los p-valores: la distribución nula, construida

Los contrastes del capítulo han producido p-valores exactos —0,208 625, 0,054 688— sin explicar de dónde venían. La respuesta es que en los métodos no paramétricos la distribución nula se *construye contando*, y esa construcción es todo el contenido del método.

Teorema 21.43 (Distribución nula de Mann–Whitney)**.** *Sean dos muestras independientes de tamaños n_1 y n_2, sin empates, y sea*

$$U = \#\{(i, j) : \; x_i < y_j\}$$

el número de pares en que una observación del primer grupo es menor que una del segundo. Bajo la hipótesis nula de que las $n_1 + n_2$ observaciones son intercambiables, todas las $\binom{n_1+n_2}{n_1}$ asignaciones de rangos son igualmente probables, y

$$P(U = u) = \frac{N(u; n_1, n_2)}{\binom{n_1+n_2}{n_1}},$$

donde $N(u; n_1, n_2)$ es el número de asignaciones con ese valor de U. Además

$$\mathrm{E}[U] = \frac{n_1 n_2}{2}, \qquad \mathrm{Var}(U) = \frac{n_1 n_2 (n_1 + n_2 + 1)}{12}.$$

Demostración. La distribución. Bajo H_0 todas las ordenaciones conjuntas son equiprobables, y U es función sólo de la ordenación; contar las que dan cada valor y dividir entre el total es la definición de probabilidad en un espacio equiprobable.

La media. Escríbase $U = \sum_{i,j} I_{ij}$ con $I_{ij} = \mathbf{1}_{\{x_i<y_j\}}$. Por simetría, $P(I_{ij}=1) = \frac{1}{2}$ para cada par, luego $\mathrm{E}[U] = n_1 n_2 \cdot \frac{1}{2}$.

La varianza. $\mathrm{Var}(U) = \sum_{i,j} \mathrm{Var}(I_{ij}) + \sum \mathrm{Cov}$. Cada $\mathrm{Var}(I_{ij}) = \frac{1}{4}$, y hay $n_1 n_2$ términos. Las covarianzas: dos indicadores que no comparten ningún índice son independientes y aportan 0; los que comparten uno, por ejemplo I_{ij} e $I_{ij'}$, cumplen $P(I_{ij} = I_{ij'} = 1) = P(x_i < \min(y_j, y_{j'})) = \frac{1}{3}$, luego su covarianza vale $\frac{1}{3} - \frac{1}{4} = \frac{1}{12}$. Hay $n_1 n_2 (n_2 - 1)$ pares que comparten el índice i y $n_1 n_2 (n_1 - 1)$ que comparten el índice j. Sumando,

$$\mathrm{Var}(U) = \frac{n_1 n_2}{4} + \frac{n_1 n_2 (n_1 + n_2 - 2)}{12} = \frac{n_1 n_2 (3 + n_1 + n_2 - 2)}{12} = \frac{n_1 n_2 (n_1 + n_2 + 1)}{12}. \quad \square$$

Ejemplo 21.44 (Un p-valor calculado a mano)**.** Con $n_1 = n_2 = 4$ hay $\binom{8}{4} = 70$ asignaciones. Un recuento directo del número de asignaciones con $U \le u$ da la tabla completa de la cola izquierda:

u	0	1	2	3	4	5	6
$\#\{U \le u\}$	1	2	4	7	12	17	24
$P(U \le u)$	0,0143	0,0286	0,0571	0,1000	0,1714	0,2429	0,3429

De aquí se lee todo: el valor crítico al 5 % de una cola es $U = 1$ (porque $P(U \le 1) = 0{,}0286 \le 0{,}05$ pero $P(U \le 2) = 0{,}0571 > 0{,}05$), y si el estadístico observado fuese $U = 5$, el p-valor a una cola sería 0,2429 y a dos colas 0,4857. Nótese que con muestras tan pequeñas el nivel 0,05 *no se puede alcanzar exactamente*: los niveles disponibles son los de la tabla, y ese carácter discreto es una propiedad intrínseca de los métodos de permutación, no un defecto.

Observación 21.45 (La aproximación normal, y cuándo usarla). Contar $\binom{n_1+n_2}{n_1}$ asignaciones es inviable a partir de muestras moderadas ($\binom{40}{20} \approx 1{,}4 \cdot 10^{11}$). Por eso, para $n_1, n_2 \ge 8$, se usa

$$Z = \frac{U - n_1 n_2/2}{\sqrt{n_1 n_2 (n_1 + n_2 + 1)/12}} \;\approx\; N(0,1),$$

con corrección de continuidad. La justificación es un teorema central del límite para estadísticos de rango —el de Hájek—, y la aproximación es notablemente buena: para $n_1 = n_2 = 10$ el error máximo en la función de distribución es inferior a 0,01. Con empates hay que corregir la varianza restando el término $\frac{n_1 n_2 \sum (t_k^3 - t_k)}{12(n_1+n_2)(n_1+n_2-1)}$, donde t_k son los tamaños de los grupos de valores repetidos.

Teorema 21.46 (Distribución nula de Kolmogórov–Smirnov)**.** *Sea F continua y F_n la función de distribución empírica de una muestra de tamaño n de F, y*

$$D_n = \sup_x \big| F_n(x) - F(x) \big|.$$

Entonces la distribución de D_n no depende de F*, y*

$$P(\sqrt{n}\, D_n \le t) \;\xrightarrow[n\to\infty]{}\; K(t) = 1 - 2\sum_{k=1}^{\infty} (-1)^{k-1} \mathrm{e}^{-2k^2 t^2}.$$

Idea clave

Por qué no depende de F. Es la transformación integral de probabilidad: si $U_i = F(X_i)$, entonces los U_i son uniformes en $(0,1)$, y como F es creciente el supremo se conserva. Así D_n es siempre el mismo estadístico calculado sobre una muestra uniforme, sea cual sea la F de partida. Esta invariancia —el carácter *libre de distribución*— es la razón de ser del contraste, y la que permite tabular unos valores críticos universales. El límite $K(t)$ es la distribución del supremo del *puente browniano*, y su serie converge muy deprisa.

Ejemplo 21.47 (Los valores críticos, y una regla práctica)**.** Resolviendo $K(t) = 1-\alpha$ se obtienen los valores críticos asintóticos:

α	0,10	0,05	0,01
t_α	1,2238	1,3581	1,6276
criterio: rechazar si $D_n >$	$1{,}2238/\sqrt{n}$	$1{,}3581/\sqrt{n}$	$1{,}6276/\sqrt{n}$

Con $n = 25$ y $\alpha = 0{,}05$, el umbral es $1{,}3581/5 = 0{,}2716$. La aproximación asintótica es aceptable a partir de $n \approx 35$; por debajo conviene la corrección de Stephens, que sustituye $\sqrt{n}$ por $\sqrt{n} + 0{,}12 + 0{,}11/\sqrt{n}$. Nótese además la lentitud: para detectar una discrepancia de 0,05 en la función de distribución hacen falta unas 740 observaciones. El contraste de Kolmogórov–Smirnov es honesto y universal, pero no es potente.

Cuidado

[El error que invalida la mitad de los usos] El teorema exige que F esté *completamente especificada de antemano*. Si se estiman parámetros a partir de los mismos datos —"contrastar normalidad estimando μ y σ con $\bar{x}$ y s"— la distribución nula deja de ser K y los valores críticos anteriores son demasiado grandes: el contraste se vuelve conservador y pierde casi toda su potencia. Para ese caso hay que usar la tabla de Lilliefors, cuyos valores críticos son notablemente menores ($0{,}886/\sqrt{n}$ en lugar de $1{,}3581/\sqrt{n}$ al 5 %), o directamente el contraste de Shapiro–Wilk, que es más potente contra alternativas de no normalidad.

21.12. Rangos, empates y las correlaciones de Spearman y Kendall

Todo este capítulo se apoya en una palabra que nunca se ha definido: *rango*. Los contrastes de Mann–Whitney, de Wilcoxon, de Kruskal–Wallis y de Friedman se enuncian sustituyendo los datos por sus rangos, la corrección por empates aparece como un factor sin explicar de dónde sale, y los ejercicios propuestos piden calcular un coeficiente de Spearman que en ningún sitio se ha definido. Esta sección pone las definiciones y las cuentas.

Definición 21.48 (Rango y rango medio)**.** Sea $x_1, \dots, x_n$ una muestra numérica. Si todos los valores son distintos, el *rango* de x_i es

$$R_i = \#\{j : \ x_j \le x_i\},$$

es decir, la posición que ocupa x_i al ordenar la muestra de menor a mayor; los rangos son entonces una permutación de $1, 2, \dots, n$.

Si hay valores repetidos, a todos los elementos de un grupo de t empatados se les asigna el mismo *rango medio*: la media aritmética de las t posiciones que habrían ocupado. Los rangos dejan entonces de ser enteros, pero siguen sumando lo mismo.

Proposición 21.49 (La suma de rangos no cambia con los empates)**.** *Con o sin empates,* $\displaystyle\sum_{i=1}^{n} R_i = \frac{n(n+1)}{2}$.

Demostración. Sin empates los rangos son $1, \dots, n$ y la suma es la de la progresión aritmética. Con empates, cada grupo de t elementos que habrían ocupado las posiciones $p+1, \dots, p+t$ recibe t veces el rango medio $\frac{1}{t}\sum_{j=1}^{t}(p+j)$, cuya suma total es $t \cdot \frac{1}{t}\sum_{j=1}^{t}(p+j) = \sum_{j=1}^{t}(p+j)$: exactamente lo mismo que aportaban las posiciones originales. Como los grupos particionan la muestra, la suma global no varía. □

Regla: asignar rangos, con empates incluidos

1. **Decide primero qué se ordena conjuntamente.** En Mann–Whitney y en Kruskal–Wallis se juntan *todos* los grupos y se ordena la mezcla; en Wilcoxon se ordenan los valores absolutos de las diferencias; en Friedman se ordena dentro de cada bloque por separado. Equivocarse aquí invalida el contraste entero, y es el error más común.
2. **Ordena de menor a mayor y numera las posiciones** $1, 2, \dots, n$**.**
3. **Localiza los bloques de valores idénticos y sustituye sus posiciones por la media.** Si tres datos empatados ocupan las posiciones 4, 5 y 6, los tres reciben el rango 5; si dos empatados ocupan 2 y 3, los dos reciben 2,5.
4. **Comprueba la suma.** Debe dar $n(n+1)/2$ por la proposición 21.49. Es una verificación de diez segundos que detecta casi cualquier error de asignación.
5. **Anota los tamaños** $t_1, t_2, \dots$ **de los grupos de empatados.** Son los que entran en las correcciones por empates del estadístico, como el factor $1 - \sum(t_j^3 - t_j)/(N^3 - N)$ de Kruskal–Wallis. Si no hay empates, todos los t_j valen 1, cada $t_j^3 - t_j$ es cero y la corrección es el factor 1: por eso solo se menciona cuando hace falta.
6. **Con muchos empates, plantéate cambiar de método.** Si más de la cuarta parte de los datos está empatada, la aproximación normal de los estadísticos de rango se deteriora y conviene usar directamente un contraste de permutación sobre los rangos, que es exacto por construcción.

Ejemplo 21.50 (Los rangos de tres muestras, con la comprobación)**. (a) Sin empates.** Para $3,\ 7,\ 10,\ 15$ los rangos son $1,2,3,4$ y suman $10 = 4\cdot 5/2$ ✓.

(b) Con un empate doble. Para $3,\ 7,\ 7,\ 10,\ 15$, los dos sietes ocuparían las posiciones 2 y 3, luego reciben ambos el rango 2,5. Los rangos son $1;\ 2{,}5;\ 2{,}5;\ 4;\ 5$ y suman $15 = 5\cdot 6/2$ ✓. Este es exactamente el ejercicio propuesto número 3 del capítulo.

(c) Con dos empates dobles. Para $12,\ 15,\ 15,\ 17,\ 18,\ 18$ los rangos son $1;\ 2{,}5;\ 2{,}5;\ 4;\ 5{,}5;\ 5{,}5$, que suman $21 = 6\cdot 7/2$ ✓. Los tamaños de los grupos de empatados son $t_1 = t_2 = 2$, de modo que $\sum(t_j^3 - t_j) = 6 + 6 = 12$.

Definición 21.51 (Coeficiente de correlación de Spearman)**.** Dadas n parejas (x_i, y_i), sean R_i los rangos de las x y S_i los de las y, calculados por separado con la definición 21.48. El *coeficiente de Spearman* r_s es el coeficiente de correlación de Pearson de las parejas de rangos:

$$r_s = \frac{\sum_i (R_i - \bar R)(S_i - \bar S)}{\sqrt{\sum_i (R_i - \bar R)^2}\,\sqrt{\sum_i (S_i - \bar S)^2}}, \qquad \bar R = \bar S = \frac{n+1}{2}.$$

Teorema 21.52 (La fórmula abreviada de Spearman)**.** *Si no hay empates en ninguna de las dos variables y $d_i = R_i - S_i$, entonces*

$$r_s = 1 - \frac{6\sum_{i=1}^n d_i^2}{n(n^2-1)}.$$

Demostración. Sin empates, tanto los R_i como los S_i son permutaciones de $1,\dots,n$, así que ambas tienen la misma media $\bar R = \bar S = (n+1)/2$ y la misma suma de cuadrados centrados,

$$\sum_i (R_i - \bar R)^2 = \sum_{k=1}^n k^2 - n\bar R^2 = \frac{n(n+1)(2n+1)}{6} - \frac{n(n+1)^2}{4} = \frac{n(n^2-1)}{12} \equiv V.$$

Los dos radicandos del denominador de la definición valen V, de modo que el denominador es V. Para el numerador, obsérvese que

$$\sum_i d_i^2 = \sum_i \big[(R_i - \bar R) - (S_i - \bar S)\big]^2 = V + V - 2\sum_i (R_i - \bar R)(S_i - \bar S),$$

porque $\bar R = \bar S$ hace que $d_i = (R_i - \bar R) - (S_i - \bar S)$. Despejando la covarianza y dividiendo entre V,

$$r_s = \frac{V - \frac12 \sum d_i^2}{V} = 1 - \frac{\sum d_i^2}{2V} = 1 - \frac{6\sum d_i^2}{n(n^2-1)}. \qquad \square$$

Definición 21.53 (Tau de Kendall)**.** Con las mismas n parejas y sin empates, una pareja de índices $i < j$ es *concordante* si $(x_j - x_i)(y_j - y_i) > 0$ y *discordante* si el producto es negativo. Llamando C y D a sus números,

$$\tau = \frac{C - D}{\binom{n}{2}}.$$

Ejemplo 21.54 (El ejercicio propuesto número 6, resuelto entero y con su p-valor exacto)**.** Se piden Spearman entre $x = (1,2,3,4,5)$ y $y = (2,1,4,3,5)$.

(a) Los rangos. No hay empates, y los propios valores ya son rangos: $R = (1, 2, 3, 4, 5)$ y $S = (2, 1, 4, 3, 5)$.

(b) Las diferencias. $d = (-1,\ 1,\ -1,\ 1,\ 0)$, de donde $\sum d_i^2 = 4$. Por el teorema 21.52,

$$r_s = 1 - \frac{6 \cdot 4}{5 \cdot 24} = 1 - \frac{24}{120} = 0{,}8.$$

Calculando en cambio la correlación de Pearson sobre los rangos, como manda la definición 21.51, sale también 0,8: las dos vías coinciden, como debe ocurrir sin empates.

(c) Kendall, sobre los mismos datos. Hay $\binom{5}{2} = 10$ parejas. Discordantes son solamente $(1, 2)$ —porque x sube de 1 a 2 mientras y baja de 2 a 1— y $(3, 4)$, por la misma razón. Luego $C = 8$, $D = 2$ y

$$\tau = \frac{8 - 2}{10} = 0{,}6.$$

Es sistemáticamente menor que r_s, y no es casualidad: τ mide una proporción de parejas y r_s una correlación, y las dos escalas no coinciden. Si los datos vinieran de una normal bivariante de correlación ρ, se demuestra que $\tau = \frac{2}{\pi}\arcsin\rho$ y $r_s = \frac{6}{\pi}\arcsin\frac{\rho}{2}$; despejando de la segunda con $r_s = 0{,}8$ sale $\rho = 0{,}813\,473$, y llevándolo a la primera, $\tau = 0{,}604\,852$, prácticamente el 0,6 observado.

(d) ¿Es significativo? Aquí está lo interesante. Bajo la hipótesis nula de independencia, cualquiera de las $5! = 120$ permutaciones de S es igual de probable, así que la distribución exacta de r_s se obtiene *enumerando las 120*. Haciéndolo, resultan 8 permutaciones con $r_s \geq 0{,}8$, de modo que

$$p_{\text{unilateral}} = \frac{8}{120} = 0{,}066\,667, \qquad p_{\text{bilateral}} = \frac{16}{120} = 0{,}133\,333.$$

No se rechaza la independencia. Nótese que la aproximación asintótica habitual devuelve $p = 0{,}104\,088$, un valor bastante distinto del exacto: con $n = 5$ ninguna aproximación es de fiar, y la enumeración completa cuesta un segundo de ordenador.

Ejemplo 21.55 (Valores críticos exactos de Spearman, calculados enumerando)**.** Enumerando las $n!$ permutaciones se obtiene, para cada n, el menor valor de r_s cuya probabilidad de ser superado no llega al 5 % por una cola:

n	5	6	7	8	9	10
r_s crítico	0,9000	0,8286	0,7143	0,6429	0,6000	0,5636
$P(r_s \geq \text{crítico})$	0,0417	0,0292	0,0440	0,0481	0,0484	0,0481

Los valores coinciden con los de las tablas clásicas, y la última fila enseña la lección de siempre en los contrastes exactos: el nivel real nunca es 0,05 exactamente, porque la distribución es discreta. Con $n = 6$ el contraste al 5 % opera de hecho al 2,92 %, es decir, es apreciablemente conservador.

Cuidado

[Con empates, la fórmula abreviada miente] La expresión $1 - 6\sum d^2/(n(n^2-1))$ del teorema 21.52 exige *ausencia de empates*: su demostración usa dos veces que los rangos son una permutación de $1, \dots, n$. Con empates hay que volver a la definición 21.51 y calcular la correlación de Pearson sobre los rangos medios. Un ejemplo mínimo: para $x = (10, 12, 12, 15, 20)$ e $y = (3, 5, 4, 4, 9)$ los rangos son $R = (1;\ 2{,}5;\ 2{,}5;\ 4;\ 5)$ y $S = (1;\ 4;\ 2{,}5;\ 2{,}5;\ 5)$; la definición correcta da $r_s = 0{,}763\,158$ y la fórmula abreviada aplicada a ciegas da 0,775. La diferencia parece pequeña, pero crece con el número de empates y siempre en la misma dirección: la fórmula abreviada *exagera* la correlación.

Observación 21.56 (El puente entre rangos y Mann–Whitney). La suma de rangos y el estadístico U de Mann–Whitney no son dos cosas distintas. Si R_1 es la suma de los rangos del primer grupo en la ordenación conjunta, entonces

$$U = R_1 - \frac{n_1(n_1+1)}{2},$$

porque R_1 cuenta, para cada observación del primer grupo, cuántos elementos de la muestra conjunta la preceden o la igualan, y hay que descontar los que pertenecen al propio grupo, que son exactamente $1+2+\cdots+n_1$. Compruébese con el ejemplo con el que se abre el capítulo: para $A = \{12, 15, 17, 19, 22, 25, 28\}$ y $B = \{14, 18, 21, 26, 30, 33, 35\}$ los rangos de A en la ordenación conjunta son $1, 3, 4, 6, 8, 9, 11$, que suman $R_A = 42$, y en efecto $42 - \frac{7\cdot 8}{2} = 42 - 28 = 14 = U$ ✓, el mismo valor que allí se daba. Esa identidad es la que permite manejar el contraste con una tabla de sumas de rangos, que era como se hacía antes de los ordenadores.

21.13. El contraste de Wilcoxon de los rangos con signo, con su distribución exacta

El capítulo enunció el contraste de Wilcoxon en una frase de una definición compartida con Mann–Whitney y Kolmogórov–Smirnov, lo aplicó a ocho diferencias y produjo el número $p = 0{,}054\,688$ sin decir cómo. Aquí está el estadístico, el procedimiento paso a paso, la distribución nula y ese p-valor recalculado a mano.

Definición 21.57 (Estadístico de los rangos con signo)**.** Sean $d_1, \dots, d_n$ las diferencias de n pares, todas no nulas. Se ordenan los valores absolutos $|d_i|$ y se les asigna su rango R_i de 1 a n. Entonces

$$W^+ = \sum_{i:\ d_i>0} R_i, \qquad W^- = \sum_{i:\ d_i<0} R_i,$$

y el estadístico del contraste es $W = \min(W^+, W^-)$. Las diferencias nulas se descartan y n se reduce en consecuencia.

Regla: aplicar el contraste de los rangos con signo

1. **Calcula las n diferencias y elimina las nulas**, ajustando n. Descartar un empate exacto es la convención habitual; existen variantes que los reparten, pero conviene

declarar cuál se usa.

2. **Ordena los valores absolutos y asigna rangos**, con rangos medios si hay empates entre magnitudes, tal como manda la regla de la sección anterior.
3. **Devuelve a cada rango el signo de su diferencia** y suma por separado los rangos positivos y los negativos. Comprueba que $W^+ + W^- = n(n+1)/2$: esa igualdad es la proposición de la suma de rangos y es la verificación que hay que hacer siempre.
4. **Con $n \le 20$, usa la distribución exacta** del teorema 21.58 o la tabla de valores críticos de más abajo. No hace falta ordenador: es un recuento de subconjuntos.
5. **Con $n > 20$, tipifica** usando $\mathrm{E}[W^+] = n(n+1)/4$ y $\mathrm{Var}(W^+) = n(n+1)(2n+1)/24$, y compara con la normal, con corrección de continuidad de 0,5.
6. **Interpreta lo que has contrastado.** La hipótesis nula es que la distribución de las diferencias es *simétrica respecto de cero*, no que su media valga cero. Si la distribución de las diferencias es asimétrica, un rechazo puede deberse a la asimetría y no al desplazamiento.

Teorema 21.58 (Distribución nula de los rangos con signo)**.** *Supóngase que las diferencias $d_1, \dots, d_n$ son independientes y con distribución continua simétrica respecto de cero. Entonces los signos son independientes de los rangos y cada uno de los 2^n patrones de signos es igualmente probable, de modo que*

$$P(W^+ = w) = \frac{c_n(w)}{2^n},$$

donde $c_n(w)$ es el número de subconjuntos de $\{1, 2, \dots, n\}$ cuya suma vale w. Esos números son los coeficientes del polinomio

$$\prod_{k=1}^{n} (1 + z^k) = \sum_{w=0}^{n(n+1)/2} c_n(w)\, z^w,$$

y además

$$\mathrm{E}[W^+] = \frac{n(n+1)}{4}, \qquad \mathrm{Var}(W^+) = \frac{n(n+1)(2n+1)}{24}.$$

Demostración. Por la simetría de cada d_i respecto de cero, el signo de d_i es independiente de $|d_i|$ y vale $+$ o $-$ con probabilidad $\frac{1}{2}$ cada uno; y como los rangos dependen solo de los valores absolutos, los signos son independientes del vector de rangos. Fijados los rangos, W^+ es la suma de los rangos que reciben signo positivo, es decir, la suma de un subconjunto aleatorio uniforme de $\{1, \dots, n\}$: de ahí el recuento $c_n(w)/2^n$. La función generatriz es inmediata, porque al desarrollar $\prod_k (1 + z^k)$ cada término del producto corresponde a elegir o no el rango k, y el exponente resultante es la suma elegida.

Para los momentos, escríbase $W^+ = \sum_{k=1}^{n} k\,\varepsilon_k$ con ε_k independientes que valen 1 o 0 con probabilidad $\tfrac12$. Entonces $\mathrm{E}[\varepsilon_k] = \tfrac12$ y $\mathrm{Var}(\varepsilon_k) = \tfrac14$, de donde

$$\mathrm{E}[W^+] = \frac{1}{2}\sum_{k=1}^{n} k = \frac{n(n+1)}{4}, \qquad \mathrm{Var}(W^+) = \frac{1}{4}\sum_{k=1}^{n} k^2 = \frac{1}{4}\cdot\frac{n(n+1)(2n+1)}{6} = \frac{n(n+1)(2n+1)}{24}.$$

□

Ejemplo 21.59 (El p-valor de 0,054688 del capítulo, recalculado contando)**.** Las ocho diferencias del ejemplo de apertura eran 2,1; −0,4; 3,3; 1,8; −1,1; 2,5; 0,9; 4,0.

(a) Rangos de los valores absolutos. Ordenando $0{,}4 < 0{,}9 < 1{,}1 < 1{,}8 < 2{,}1 < 2{,}5 < 3{,}3 < 4{,}0$, los rangos son

d_i	2,1	−0,4	3,3	1,8	−1,1	2,5	0,9	4,0
R_i	5	1	7	4	3	6	2	8

No hay empates, así que todos los rangos son enteros.

(b) Las dos sumas. Los negativos son −0,4 (rango 1) y −1,1 (rango 3), luego $W^- = 1 + 3 = 4$ y $W^+ = 36 - 4 = 32$. Comprobación: $W^+ + W^- = 36 = \frac{8\cdot 9}{2}$ ✓. El estadístico es $W = \min(32, 4) = 4$.

(c) El recuento. Por el teorema 21.58 hay que contar los subconjuntos de $\{1, \dots, 8\}$ cuya suma no pasa de 4. Se pueden escribir todos:

$$\varnothing, \quad \{1\}, \quad \{2\}, \quad \{3\}, \quad \{4\}, \quad \{1,2\}, \quad \{1,3\},$$

siete en total, de sumas $0, 1, 2, 3, 4, 3, 4$. Como $2^8 = 256$,

$$P(W^- \le 4) = \frac{7}{256} = 0{,}027\,344.$$

(d) El veredicto. Por simetría, la cola derecha aporta lo mismo, de modo que

$$p_{\text{bilateral}} = 2 \cdot \frac{7}{256} = \frac{14}{256} = 0{,}054\,688,$$

que es exactamente el número que el capítulo daba sin explicar. Está justo por encima de 0,05: con estos ocho datos no se rechaza al 5 %, aunque por muy poco. Y obsérvese la granularidad, que es la firma de los contrastes exactos: con $n = 8$ los p-valores bilaterales posibles son múltiplos de $2/256 = 0{,}007\,813$, así que no existe ningún resultado que dé exactamente 0,05.

Ejemplo 21.60 (La tabla de valores críticos, obtenida por enumeración exacta)**.** Contando subconjuntos para cada n se obtiene el mayor w tal que $2\,P(W \le w) \le 0{,}05$, que es el valor crítico bilateral al 5 % (se rechaza si el estadístico $W = \min(W^+, W^-)$ es *menor o igual* que él):

n	6	7	8	9	10	11	12	13	14
W crítico	0	2	3	5	8	10	13	17	21
nivel real	0,0313	0,0469	0,0391	0,0391	0,0488	0,0420	0,0425	0,0479	0,0494

Dos lecturas. La primera: con $n = 5$ la tabla se corta, porque el p-valor bilateral más pequeño alcanzable es $2/2^5 = 0{,}0625 > 0{,}05$; con cinco pares y este contraste es *imposible* rechazar al 5 %, aunque las cinco diferencias tengan el mismo signo. La segunda: el nivel real oscila entre el 3 y el 5 % y solo se acerca a 0,05 cuando n crece, otra manifestación de la discreción del método. Los valores coinciden con los de las tablas publicadas de Wilcoxon, que se construyeron exactamente así.

Definición 21.61 (Eficiencia relativa asintótica)**.** Sean dos contrastes válidos para el mismo problema. La *eficiencia relativa asintótica* del primero respecto del segundo es el límite del cociente n_2/n_1 de los tamaños muestrales que hacen falta para alcanzar la misma potencia frente a una sucesión de alternativas que se acercan a la nula, cuando esa potencia se mantiene fija y $n \to \infty$. Un valor de 0,95 significa que el primer contraste necesita un 5 % más de datos que el segundo para hacer el mismo trabajo.

Observación 21.62 (De dónde sale el famoso 3 partido por pi). La cifra $3/\pi = 0{,}954\,930$ que el capítulo cita tres veces es la eficiencia relativa asintótica de los contrastes de rangos frente a los contrastes normales *cuando los datos son realmente normales*. Es decir, es el peor caso posible para los rangos: es el terreno de juego del rival. Perder un 4,5 % de eficiencia en el escenario más favorable al competidor, y ganarle sin límite en cuanto las colas se ensanchan, es un intercambio extraordinariamente ventajoso; de hecho se demuestra que la eficiencia de Wilcoxon frente a la t nunca baja de $108/125 = 0{,}864$, sea cual sea la distribución continua y simétrica de partida (es la cota de Hodges y Lehmann), mientras que la de la t frente a Wilcoxon puede ser arbitrariamente próxima a cero. Esa asimetría, y no la robustez genérica, es el argumento técnico decisivo a favor de los rangos.

Cuidado

Tres confusiones que conviene deshacer. Primera: el contraste de los rangos con signo *no* es el contraste del signo. El del signo solo cuenta cuántas diferencias son positivas y usa una binomial; con las ocho diferencias del ejemplo daría 6 positivas de 8 y $p = 0{,}289\,062$, cinco veces mayor. Wilcoxon usa además las magnitudes, ordenadas, y por eso es más potente. Segunda: la hipótesis nula de Wilcoxon es la *simetría* de las diferencias respecto de cero, más fuerte que "la mediana es cero", y con diferencias muy asimétricas el contraste puede rechazar sin que haya ningún desplazamiento. Tercera: W^+ y W^- no son dos estadísticos distintos, sino complementarios, porque su suma es siempre $n(n+1)/2$; usar uno u otro cambia la cola en que hay que mirar, y ese es el origen del error clásico de dar un p-valor que no baja de 0,5 cuando el efecto es enorme.

21.14. Ejercicios resueltos

Tres problemas donde la normalidad no se puede suponer y hay que trabajar con rangos o remuestreo.

Ejercicio 21.1. De 12 pares antes-después, 9 muestran mejora y 3 empeoramiento. Aplique el test de los signos y compare con lo que haría una t pareada.

Solución. Bajo H_0, cada par mejora con probabilidad $1/2$, así que el número de mejoras sigue una $B(12; 0{,}5)$. El p-valor bilateral es $2\,P(X \geq 9) = 0{,}146$, que no permite rechazar al 5 por ciento.

La t pareada usaría las magnitudes de las diferencias y no sólo sus signos, de modo que si las nueve mejoras fuesen grandes y los tres empeoramientos minúsculos, la t detectaría el efecto y el test de los signos no. Ése es el precio de no suponer nada: el test de los signos es válido sin hipótesis de distribución, pero tira información y por tanto tiene menos potencia. La elección correcta depende de si uno se cree la normalidad; si se la cree, la t; si no, el test de Wilcoxon, que es el término medio porque usa los rangos de las magnitudes. ■

Ejercicio 21.2. Compare los grupos $A = \{12, 15, 9, 14, 11\}$ y $B = \{18, 20, 16, 19, 17\}$ con la prueba de Mann–Whitney. Interprete el resultado.

Solución. Ordenando conjuntamente las diez observaciones, todas las de A quedan por debajo de todas las de B: los rangos de A son $1, 2, 3, 4, 5$ y los de B son $6, 7, 8, 9, 10$. Por tanto $U_A = 0$, el valor mínimo posible, y el p-valor bilateral exacto es $0{,}0079$.

Se rechaza la igualdad de distribuciones. Es el caso extremo: separación total. Nótese que la prueba no ha usado en ningún momento los valores numéricos, sólo el orden, y aun así distingue los grupos con contundencia. Nótese también qué contrasta exactamente: no la igualdad de medias, sino la de distribuciones, y bajo el supuesto adicional de igual forma se interpreta como desplazamiento de la mediana. ■

Ejercicio 21.3. Explique cómo se construye un intervalo de confianza por bootstrap para la mediana y por qué funciona.

Solución. De la muestra original de tamaño n se extraen B remuestras de tamaño n *con reemplazo*, se calcula la mediana de cada una y se toman los percentiles 2,5 y 97,5 de esas B medianas. Con $B = 10\,000$ basta para casi todo.

La idea es que la muestra es la mejor aproximación disponible a la población, de modo que remuestrear de la muestra imita el proceso de muestrear de la población. Lo que se estima no es la mediana, que ya se tiene, sino su *variabilidad*. Para la mediana no hay una fórmula sencilla del error típico como la hay para la media, y ahí es donde el bootstrap resulta insustituible.

Un detalle bonito: cada remuestra deja fuera, en promedio, la fracción $(1 - 1/n)^n \to 1/\mathrm{e} = 0{,}368$ de los datos originales. Ese 36,8 por ciento excluido es lo que en aprendizaje automático se usa como conjunto de validación en los bosques aleatorios. ■

21.15. Ejercicios propuestos

21.1. ¿Cuándo es preferible una prueba no paramétrica a una paramétrica? Cite dos situaciones.

21.2. ¿Qué contrasta exactamente la prueba de Mann–Whitney y bajo qué supuesto se lee como diferencia de medianas?

21.3. Calcule los rangos de $3, 7, 7, 10, 15$ y explique cómo se tratan los empates.

21.4. Aplique el test de los signos a 7 mejoras de 10 pares.

21.5. ¿Qué añade Wilcoxon respecto del test de los signos?

21.6. Calcule el coeficiente de Spearman entre $(1, 2, 3, 4, 5)$ y $(2, 1, 4, 3, 5)$.

21.7. ¿Qué mide el estadístico de Kolmogorov–Smirnov y con qué se compara?

21.8. ¿Cuál es la eficiencia relativa asintótica de Wilcoxon frente a la t bajo normalidad?

21.9. Explique la diferencia entre bootstrap y validación cruzada.

21.10. ¿Qué fracción de los datos originales queda fuera de una remuestra bootstrap típica?

21.11. ¿Se puede usar bootstrap con datos dependientes? ¿Qué variante hace falta?

21.12. ¿Qué es un test de permutaciones y por qué su validez no depende de ninguna distribución?

Capítulo 22

Procesos estocásticos

Lo que necesitas traer

Del capítulo 7: matrices, potencias de matrices y valores propios, para las cadenas de Markov. Del capítulo 16: variables aleatorias, esperanza, independencia y la distribución normal. Del capítulo 6: límites y series.

Una variable aleatoria describe un experimento; un proceso estocástico describe un sistema que evoluciona al azar en el tiempo. Los ejemplos son omnipresentes: la cola de un servidor, el precio de un activo, la posición de una partícula en un fluido, el número de individuos de una población. Este capítulo presenta las dos familias fundamentales, las cadenas de Markov en tiempo discreto y el movimiento browniano en tiempo continuo, y la propiedad que las une: el futuro depende del presente, no del camino recorrido.

22.1. Cadenas de Markov

Definición 22.1. Una *cadena de Markov* es una sucesión $X_0, X_1, \ldots$ de variables con valores en un conjunto de estados que cumple la *propiedad de Markov*: $P(X_{n+1} = j \mid X_n = i, \ldots, X_0) = P(X_{n+1} = j \mid X_n = i) = p_{ij}$. La matriz $P = (p_{ij})$ es *estocástica* (filas que suman 1), y $P(X_n = j \mid X_0 = i) = (P^n)_{ij}$. Una distribución $\boldsymbol{\pi}$ es *estacionaria* si $\boldsymbol{\pi} P = \boldsymbol{\pi}$.

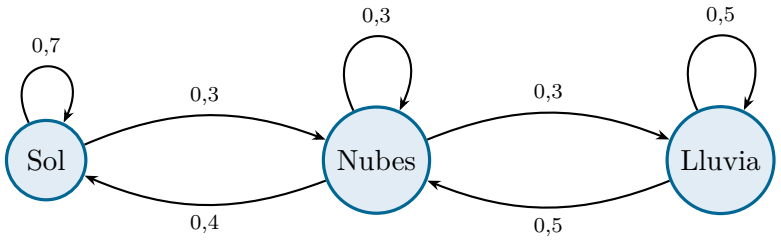


Figura 22.1: Una cadena de Markov de tres estados. Las flechas que salen de cada nodo suman uno. La cadena es irreducible y aperiódica, de modo que existe una única distribución estacionaria y el tiempo de mañana acaba olvidando el de hoy.

Idea clave

Qué dice esto en cristiano. La propiedad de Markov es una hipótesis de *amnesia*: para predecir el siguiente paso basta saber dónde estás, no cómo llegaste. Eso reduce un objeto potencialmente infinito (toda la historia) a un solo número (el estado actual), y es lo que hace manejable la teoría. La consecuencia operativa es que todo se calcula con potencias de una matriz: la distribución tras n pasos es la inicial multiplicada por P^n. Y la distribución estacionaria es el autovector por la izquierda asociado al autovalor 1, que toda matriz estocástica tiene.

Ejemplo 22.2 (Una cadena de dos y tres estados)**. (a) El tiempo meteorológico.** Con estados soleado y lluvioso y matriz $P = \begin{pmatrix} 0{,}8 & 0{,}2 \\ 0{,}4 & 0{,}6 \end{pmatrix}$, la distribución estacionaria resuelve $\pi_1 = 0{,}8\pi_1 + 0{,}4\pi_2$ con $\pi_1 + \pi_2 = 1$, y da $\boldsymbol{\pi} = \left(\frac{2}{3}, \frac{1}{3}\right)$. A largo plazo dos de cada tres días son soleados, sea cual sea el tiempo de hoy.

(b) La convergencia. Partiendo de un día lluvioso, $(0, 1)$, las distribuciones sucesivas son (0,4; 0,6), luego (0,56; 0,44), (0,624; 0,376), (0,6496; 0,3504)… tendiendo a (0,6667; 0,3333). La velocidad la marca el segundo autovalor, que aquí vale 0,4: el error se multiplica por 0,4 en cada paso.

(c) Tiempo medio de retorno. Para una cadena irreducible, el tiempo medio de volver al estado i es $\frac{1}{\pi_i}$. Aquí, volver a un día lluvioso tarda de media 3 días y a uno soleado 1,5. Es una de las fórmulas más útiles y menos conocidas de la teoría.

(d) La ruina del jugador. Con capital i entre 0 y N y probabilidad p de ganar cada apuesta, la probabilidad de alcanzar N antes de arruinarse es $\frac{1-(q/p)^i}{1-(q/p)^N}$ si $p \neq q$, y $\frac{i}{N}$ si el juego es justo. Con $i = 10$, $N = 20$ y $p = \frac{18}{37}$ (la ruleta) sale 0,368031, frente al 0,5 que daría un juego equitativo. Una desventaja del 2,7 % por tirada cuesta trece puntos de probabilidad.

Cuidado

[La estacionaria existe casi siempre, pero converger a ella exige más] Toda cadena finita tiene al menos una distribución estacionaria, pero para que la cadena *converja* a ella desde cualquier inicio hace falta que sea irreducible (se puede ir de cualquier estado a cualquier otro) y *aperiódica*. La cadena $\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$ tiene estacionaria $\left(\frac{1}{2}, \frac{1}{2}\right)$ y nunca converge: oscila para siempre entre los dos estados. Segundo aviso: la propiedad de Markov es una hipótesis fuerte y con frecuencia falsa. Si el sistema tiene memoria (un cliente que lleva mucho esperando se comporta distinto), hay que ampliar el estado para incluirla, y eso puede hacer explotar el número de estados.

22.2. Qué es exactamente un proceso estocástico, y una cadena de Markov resuelta de principio a fin

La sección anterior ha definido la cadena de Markov, la matriz estocástica y la distribución estacionaria en un solo párrafo, y ha calculado con ellas. Falta debajo el suelo: qué es un proceso estocástico, qué es una trayectoria, qué significa que el tiempo sea discreto o continuo, de dónde sale la distribución inicial y por qué las probabilidades a n pasos se leen en una potencia de matriz. Esta sección pone ese suelo y luego resuelve una cadena concreta entera —matriz, distribución estacionaria por el sistema de ecuaciones, y comprobación numérica de que las potencias convergen— para que quede un modelo de cómo se hacen estas cuentas.

Procesos, trayectorias y espacio de estados

Definición 22.3 (Proceso estocástico)**.** Un *proceso estocástico* es una familia $\{X_t\}_{t\in T}$ de variables aleatorias definidas todas sobre el mismo espacio de probabilidad y con valores en un mismo conjunto E. El conjunto T es el *conjunto de índices* o *tiempos*, y E es el *espacio de estados*. Se habla de:

- *tiempo discreto* si $T = \{0, 1, 2, \dots\}$ —o cualquier conjunto numerable—, y entonces se escribe $X_0, X_1, X_2, \dots$;
- *tiempo continuo* si $T = [0, \infty)$ o un intervalo;
- *espacio de estados discreto* si E es finito o numerable, y *continuo* si es $\mathbb{R}$, $\mathbb{R}^d$ o similar.

Definición 22.4 (Trayectoria)**.** Fijado un resultado ω del experimento aleatorio, la función

$$t \longmapsto X_t(\omega)$$

se llama *trayectoria*, *camino* o *realización* del proceso. Un proceso estocástico se puede mirar, por tanto, de dos maneras: como una colección de variables aleatorias indexada por el tiempo, o como una única variable aleatoria *cuyo valor es una función entera del tiempo*.

Idea clave

Qué dice esto en cristiano. Una variable aleatoria sortea un número. Un proceso estocástico sortea una *película*: al azar, de una vez, sale una función completa del tiempo. Cuando se dibuja la evolución de un precio, o la posición de una partícula, o la longitud de una cola a lo largo de la tarde, lo que se está viendo es un único punto del espacio muestral, no una sucesión de sorteos independientes. Esta doble lectura explica por qué hay dos tipos de preguntas completamente distintas: las que se refieren a un instante fijo (¿cómo se distribuye X_{10}?) y las que se refieren a la trayectoria entera (¿es continua? ¿cuál es su máximo? ¿cuándo cruza por primera vez el nivel a?). Las primeras se contestan con distribuciones de una variable; las segundas necesitan toda la maquinaria del capítulo.

Ejemplo 22.5 (Cuatro procesos y sus trayectorias)**. (a) Lanzamientos de moneda.** X_n vale 1 si el lanzamiento n-ésimo sale cara y 0 si sale cruz. Tiempo discreto, espacio de estados $E = \{0, 1\}$. Una trayectoria es una sucesión de ceros y unos, por ejemplo $1, 0, 0, 1, 1, 0, \ldots$. Es un proceso, pero de los aburridos: las variables son independientes y el pasado no dice nada del futuro.

(b) Paseo aleatorio. $S_n = X_1 + \cdots + X_n$ con X_i valiendo $+1$ o -1 con probabilidad $\frac{1}{2}$. Tiempo discreto, espacio de estados $E = \mathbb{Z}$. Con los lanzamientos $+, -, -, +, +, +, -$ la trayectoria es $1, 0, -1, 0, 1, 2, 1$. Aquí ya hay dependencia: S_{n+1} está pegado a S_n, a distancia exactamente 1.

(c) Proceso de recuento. $N(t)$ es el número de llamadas recibidas hasta el instante t. Tiempo continuo, espacio de estados $E = \{0, 1, 2, \ldots\}$. Las trayectorias son funciones escalonadas, constantes a trozos, que suben de uno en uno.

(d) Movimiento browniano. W_t es la posición de una partícula. Tiempo continuo y espacio de estados continuo, $E = \mathbb{R}$. Sus trayectorias son continuas y, como se verá, no derivables en ningún punto: es el caso en que la mirada "el azar sortea una función" resulta imprescindible, porque toda la dificultad está en las propiedades de esa función.

Observación 22.6 (Qué significa conocer un proceso)*.* Un proceso no queda determinado por las distribuciones de cada X_t por separado. Hacen falta las *distribuciones finito-dimensionales*: las de todos los vectores $(X_{t_1}, \ldots, X_{t_k})$ para cualesquiera instantes. Dos procesos pueden tener $X_t \sim N(0, t)$ para todo t y ser completamente distintos, según cómo estén correlacionados los valores en instantes diferentes. El teorema de extensión de Kolmogórov garantiza el recíproco: dada una familia coherente de distribuciones finito-dimensionales, existe un proceso que las tiene. Esa es la razón por la que, en la práctica, un proceso se especifica diciendo cómo depende cada valor de los anteriores, que es justo lo que hace la propiedad de Markov.

Cadenas de Markov: los datos del problema

Definición 22.7 (Cadena homogénea, matriz de transición y distribución inicial)**.** Sea $\{X_n\}_{n\geq 0}$ un proceso en tiempo discreto con espacio de estados finito $E = \{1, 2, \ldots, S\}$ que cumple la propiedad de Markov. La cadena es *homogénea* si las probabilidades de transición no dependen del instante:

$$P(X_{n+1} = j \mid X_n = i) = p_{ij} \qquad \text{para todo } n.$$

Entonces quedan definidos:

- la *matriz de transición* $P = (p_{ij})_{i,j\in E}$, que es *estocástica*: sus entradas son no negativas y cada *fila* suma 1, porque desde el estado i hay que ir a alguna parte;
- la *distribución inicial* $\mu^{(0)} = (\mu_1^{(0)}, \ldots, \mu_S^{(0)})$, dada por $\mu_i^{(0)} = P(X_0 = i)$, que es un vector fila de componentes no negativas que suman 1.

El par $(\mu^{(0)}, P)$ determina por completo la cadena.

Proposición 22.8 (Todo se calcula con potencias de P)**.** *Con la notación anterior, para todos los estados y todo $n \geq 1$:*

1. *la probabilidad de una trayectoria concreta es*
$$P(X_0 = i_0,\ X_1 = i_1,\ \ldots,\ X_n = i_n) = \mu^{(0)}_{i_0}\, p_{i_0 i_1} p_{i_1 i_2} \cdots p_{i_{n-1} i_n};$$
2. *la distribución en el instante n es $\mu^{(n)} = \mu^{(0)} P^n$, entendido como producto de un vector fila por una matriz.*

Demostración. (1) Aplíquese repetidamente la regla del producto, $P(A \cap B) = P(A)P(B \mid A)$, empezando por el final:

$$P(X_0 = i_0, \ldots, X_n = i_n) = P(X_n = i_n \mid X_0 = i_0, \ldots, X_{n-1} = i_{n-1})\, P(X_0 = i_0, \ldots, X_{n-1} = i_{n-1}),$$

y por la propiedad de Markov el primer factor es $p_{i_{n-1} i_n}$. Repitiendo n veces se llega a $P(X_0 = i_0) = \mu^{(0)}_{i_0}$ multiplicado por los n factores de transición.

(2) Por la ley de la probabilidad total, partiendo según dónde estaba la cadena un paso antes,

$$\mu^{(n)}_j = P(X_n = j) = \sum_{i \in E} P(X_{n-1} = i)\, P(X_n = j \mid X_{n-1} = i) = \sum_{i \in E} \mu^{(n-1)}_i\, p_{ij},$$

que es exactamente la componente j-ésima del producto $\mu^{(n-1)} P$. Iterando desde $n = 0$ se obtiene $\mu^{(n)} = \mu^{(0)} P^n$. □

Observación 22.9 (Vectores fila, no columna). Que la distribución se escriba a la *izquierda* de la matriz no es un capricho de notación: es consecuencia de que las filas de P sumen 1. Con este convenio, la distribución estacionaria es un vector propio *por la izquierda* de valor propio 1, es decir $\pi P = \pi$. Hay libros —sobre todo de física— que usan el convenio traspuesto, con columnas que suman uno y $P\pi = \pi$. Ambos son correctos y describen lo mismo; mezclarlos a media página es una de las fuentes de error más frecuentes en este tema. Conviene comprobar siempre en qué sentido suman uno las líneas de la matriz antes de multiplicar nada.

Regla: Escribir la matriz de transición a partir de un enunciado

1. **Decide el espacio de estados** y ponle nombre y orden. Los estados deben ser exhaustivos —el sistema siempre está en alguno— y excluyentes —nunca en dos a la vez—.
2. **Comprueba que la propiedad de Markov es plausible** con esos estados: ¿basta con saber dónde está el sistema hoy para predecir mañana? Si la respuesta es no, el problema no es que falle el modelo, sino que el estado está mal elegido; véase la regla siguiente.
3. **Rellena la matriz fila a fila**, leyendo cada fila como "estando en i, ¿a dónde voy y con qué probabilidad?".
4. **Comprueba que cada fila suma 1.** Es el error aritmético más común y se detecta en diez segundos.
5. **Escribe aparte la distribución inicial.** Si el enunciado dice que hoy hace sol, es $(1, 0, 0)$; si dice que el estado inicial se sortea, es el vector correspondiente.

6. **Dibuja el grafo** con un nodo por estado y una flecha rotulada por cada probabilidad positiva. La irreducibilidad y la periodicidad se leen ahí de un vistazo, mucho mejor que en la matriz.

Regla: Qué hacer cuando la propiedad de Markov falla

1. **Diagnostica**: si la probabilidad de lo que ocurre mañana depende de cuánto tiempo lleva el sistema como está, o de cómo llegó ahí, la cadena con ese espacio de estados no es markoviana.
2. **Amplía el estado incluyendo la memoria que haga falta.** Si el futuro depende de los dos últimos días, tómese como estado el *par* (ayer, hoy): el proceso de pares sí es una cadena de Markov, con S^2 estados en vez de S.
3. **Cuenta el coste.** Recordar k pasos multiplica el número de estados por S^{k-1}, y la matriz crece con el cuadrado de eso. La propiedad de Markov siempre se puede conseguir; lo que no siempre se puede es pagarla.
4. **Alternativa: cambia de tiempo.** A veces basta con incluir en el estado una variable de "tiempo transcurrido en el estado actual", que es lo que hacen los procesos semimarkovianos.
5. **Comprueba la hipótesis con datos si los hay**: se puede contrastar la markovianidad comparando las frecuencias de transición condicionadas a uno y a dos pasos de historia con un contraste χ^2 de independencia.

Una cadena concreta, de principio a fin

Ejemplo 22.10 (Tres estados de tiempo meteorológico, resuelta entera)**.** En cierta comarca se clasifica cada día como despejado (1), nublado (2) o lluvioso (3). Del registro histórico se deduce que tras un día despejado el siguiente es despejado la mitad de las veces, nublado tres de cada diez y lluvioso dos de cada diez; tras un día nublado, despejado dos de cada diez, nublado seis y lluvioso dos; y tras un día lluvioso, despejado tres de cada diez, nublado tres y lluvioso cuatro.

(a) La matriz. Ordenando los estados como arriba,

$$P = \begin{pmatrix} 0{,}5 & 0{,}3 & 0{,}2 \\ 0{,}2 & 0{,}6 & 0{,}2 \\ 0{,}3 & 0{,}3 & 0{,}4 \end{pmatrix},$$

cuyas tres filas suman 1 ✓. Todas las entradas son positivas, luego la cadena es irreducible (desde cualquier estado se llega a cualquier otro en un paso) y aperiódica (hay probabilidad positiva de quedarse), de modo que el teorema ergódico se le aplica.

(b) Predicción a pocos días vista. Si hoy está despejado, $\mu^{(0)} = (1, 0, 0)$ y la

proposición 22.8 da $\mu^{(n)} = \mu^{(0)} P^n$, que es simplemente la primera fila de P^n:

n	despejado	nublado	lluvioso
0	1	0	0
1	0,500 0	0,300 0	0,200 0
2	0,370 0	0,390 0	0,240 0
3	0,335 0	0,417 0	0,248 0
4	0,325 3	0,425 1	0,249 6
6	0,321 8	0,428 3	0,250 0

La cuenta del paso 2, hecha a mano para que se vea: $\mu_1^{(2)} = 0{,}5 \cdot 0{,}5 + 0{,}3 \cdot 0{,}2 + 0{,}2 \cdot 0{,}3 = 0{,}25 + 0{,}06 + 0{,}06 = 0{,}37$, y análogamente las otras dos componentes.

(c) La distribución estacionaria, resolviendo el sistema. Se busca $\pi = (\pi_1, \pi_2, \pi_3)$ con $\pi P = \pi$ y $\pi_1 + \pi_2 + \pi_3 = 1$. Escribiendo las tres ecuaciones columna a columna:

$$\begin{aligned} 0{,}5\pi_1 + 0{,}2\pi_2 + 0{,}3\pi_3 &= \pi_1, \\ 0{,}3\pi_1 + 0{,}6\pi_2 + 0{,}3\pi_3 &= \pi_2, \\ 0{,}2\pi_1 + 0{,}2\pi_2 + 0{,}4\pi_3 &= \pi_3. \end{aligned}$$

La tercera ecuación es redundante —siempre lo es una de ellas, porque las filas de P suman uno—, así que se descarta y se sustituye por la normalización. Reordenando las dos primeras:

$$-0{,}5\pi_1 + 0{,}2\pi_2 + 0{,}3\pi_3 = 0, \qquad 0{,}3\pi_1 - 0{,}4\pi_2 + 0{,}3\pi_3 = 0.$$

De la primera, $\pi_3 = (0{,}5\pi_1 - 0{,}2\pi_2)/0{,}3$. Sustituyendo en la segunda y multiplicando todo por 0,3:

$$0{,}09\pi_1 - 0{,}12\pi_2 + 0{,}3\big(0{,}5\pi_1 - 0{,}2\pi_2\big) = 0 \implies 0{,}24\pi_1 = 0{,}18\pi_2 \implies \pi_2 = \frac{4}{3}\pi_1.$$

Y entonces $\pi_3 = (0{,}5\pi_1 - 0{,}2 \cdot \tfrac{4}{3}\pi_1)/0{,}3 = (0{,}5 - 0{,}266\,667)\pi_1/0{,}3 = \tfrac{7}{9}\pi_1$. Imponiendo que sumen uno:

$$\pi_1\Big(1 + \frac{4}{3} + \frac{7}{9}\Big) = \pi_1 \cdot \frac{28}{9} = 1 \quad \implies \quad \pi_1 = \frac{9}{28}, \qquad \pi_2 = \frac{3}{7}, \qquad \pi_3 = \frac{1}{4}.$$

En decimales, $\pi = (0{,}321\,429;\ 0{,}428\,571;\ 0{,}250\,000)$.

(d) Comprobación. Se multiplica y se mira:

$$\pi_1' = 0{,}321\,429 \cdot 0{,}5 + 0{,}428\,571 \cdot 0{,}2 + 0{,}25 \cdot 0{,}3 = 0{,}160\,714 + 0{,}085\,714 + 0{,}075 = 0{,}321\,429 \ \checkmark,$$

y las otras dos componentes salen igual. Las tres suman uno y son positivas ✓.

(e) Tiempos medios de retorno. Por el teorema ergódico, $\mu_i = 1/\pi_i$:

$$\mu_1 = \frac{28}{9} = 3{,}11 \text{ días}, \qquad \mu_2 = \frac{7}{3} = 2{,}33 \text{ días}, \qquad \mu_3 = 4 \text{ días}.$$

Entre dos días lluviosos consecutivos pasan cuatro días de media, y entre dos nublados, dos y un tercio: los estados más probables se visitan más a menudo, exactamente en proporción inversa.

Ejemplo 22.11 (Las potencias de la matriz convergen, y a qué velocidad)**.** Sigue la misma cadena. Elevando la matriz:

$$P^2 = \begin{pmatrix} 0{,}370 & 0{,}390 & 0{,}240 \\ 0{,}280 & 0{,}480 & 0{,}240 \\ 0{,}330 & 0{,}390 & 0{,}280 \end{pmatrix}, \qquad P^4 = \begin{pmatrix} 0{,}325\,3 & 0{,}425\,1 & 0{,}249\,6 \\ 0{,}317\,2 & 0{,}433\,2 & 0{,}249\,6 \\ 0{,}323\,7 & 0{,}425\,1 & 0{,}251\,2 \end{pmatrix},$$

$$P^8 = \begin{pmatrix} 0{,}321\,457 & 0{,}428\,543 & 0{,}249\,999 \\ 0{,}321\,392 & 0{,}428\,609 & 0{,}249\,999 \\ 0{,}321\,455 & 0{,}428\,543 & 0{,}250\,002 \end{pmatrix}, \qquad P^{16} \approx \begin{pmatrix} 0{,}321\,429 & 0{,}428\,571 & 0{,}250\,000 \\ 0{,}321\,429 & 0{,}428\,571 & 0{,}250\,000 \\ 0{,}321\,429 & 0{,}428\,571 & 0{,}250\,000 \end{pmatrix}.$$

Las tres filas se van pareciendo entre sí y todas convergen al mismo vector, que es exactamente el π calculado a mano en el ejemplo anterior. Léase lo que eso significa: en P^{16} la primera fila y la tercera coinciden en seis decimales, es decir, *tras dieciséis días el tiempo que hacía el primer día es indiferente*. La cadena ha olvidado su origen.

La velocidad. Midiendo la discrepancia con la distancia en variación total, $\|\mu^{(n)} - \pi\| = \frac{1}{2}\sum_i |\mu_i^{(n)} - \pi_i|$, y partiendo de un día despejado:

n	1	2	3	4	5	6
$\|\mu^{(n)} - \pi\|$	0,178 6	0,048 6	0,013 6	0,003 9	0,001 1	0,000 3

Cada término es aproximadamente 0,29 veces el anterior: el error decae geométricamente. La razón está en los valores propios de P, que son 1, 0,3 y 0,2. El primero es el que corresponde a π y no decae; el segundo mayor en módulo, 0,3, gobierna la velocidad de convergencia, y por eso el cociente observado ronda esa cifra. Esta es la regla general: en una cadena finita, irreducible y aperiódica, la distancia a la estacionaria decae como $|\lambda_2|^n$, donde λ_2 es el segundo valor propio en módulo. Cuanto más cerca de 1 esté λ_2, más lenta es la mezcla, y esa es la cantidad que determina cuántas iteraciones hay que descartar en los métodos de Monte Carlo que se estudian al final del capítulo.

Cuidado

[Tres maneras de estropear este cálculo] *Primera*, resolver $P\pi = \pi$ en vez de $\pi P = \pi$: eso da el vector propio por la derecha, que en una matriz estocástica es siempre $(1, 1, \ldots, 1)$ y no dice absolutamente nada. *Segunda*, olvidar la normalización: el sistema $\pi(P - I) = 0$ tiene infinitas soluciones —todos los múltiplos de π— y la condición $\sum_i \pi_i = 1$ es la que selecciona una. *Tercera*, dar por hecho que P^n converge. Existir, la estacionaria existe siempre en el caso finito; converger es otra cosa y exige aperiodicidad. Con $P = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$, la estacionaria es $(\frac{1}{2}, \frac{1}{2})$ y sin embargo P^n alterna entre la identidad y la propia P para siempre. La comprobación práctica es inmediata: si algún $p_{ii} > 0$ en una cadena irreducible, hay aperiodicidad garantizada y no hay nada que temer.

22.3. Movimiento browniano

Definición 22.12. El *movimiento browniano* o proceso de Wiener es el proceso W_t con $W_0 = 0$, incrementos independientes, $W_t - W_s \sim N(0, t-s)$ y trayectorias continuas. Es el límite del paseo aleatorio al hacer los pasos infinitamente pequeños y frecuentes.

Idea clave

Qué dice esto en cristiano. Es el análogo continuo del paseo aleatorio: la partícula recibe empujones minúsculos e independientes en todas direcciones, y su posición acumulada resulta ser gaussiana con varianza proporcional al tiempo. De ahí sale la firma inconfundible: la distancia típica recorrida crece como $\sqrt{t}$, no como t. Es la misma raíz cuadrada que aparecía en la ecuación del calor, y no es coincidencia: la densidad de probabilidad del browniano *satisface* la ecuación del calor.

Ejemplo 22.13 (Propiedades del browniano)**. (a) El escalado.** Como $\operatorname{Var}(W_t) = t$, la desviación típica es $\sqrt{t}$: para $t = 1, 4, 9, 100$ vale $1, 2, 3, 10$. Recorrer el doble de distancia típica exige cuatro veces más tiempo, igual que en la difusión.

(b) Autosemejanza. El proceso $\frac{1}{\sqrt{c}} W_{ct}$ tiene la misma ley que W_t: al ampliar el tiempo por c y el espacio por $\sqrt{c}$, la trayectoria es estadísticamente idéntica. Es un fractal, con dimensión de Hausdorff $\frac{3}{2}$ para su gráfica.

(c) Trayectorias continuas y no derivables. Casi toda trayectoria browniana es continua en todo punto y derivable en *ninguno*. Es decir, el ejemplo de Weierstrass, que en su día se consideró una patología artificial, resulta ser lo que hace genéricamente una partícula real suspendida en un fluido.

(d) La variación cuadrática. La suma $\sum (W_{t_{k+1}} - W_{t_k})^2$ converge a t cuando la partición se afina, mientras que para una función derivable convergería a cero. Ese hecho es lo que obliga a inventar un cálculo nuevo: en el cálculo de Itô aparece el término extra $\frac{1}{2} f'' \, \mathrm{d}t$, y por eso $\mathrm{d}(W_t^2) = 2W_t \, \mathrm{d}W_t + \mathrm{d}t$ en vez de solo el primer sumando.

Cuidado

[No se puede integrar respecto de W con las reglas de siempre] La regla de la cadena clásica es falsa para trayectorias brownianas, precisamente porque su variación cuadrática no es nula. El lema de Itô la sustituye y añade el término de segundo orden; olvidarlo es el error que invalida cualquier cálculo estocástico. Segundo aviso: la fórmula de Black–Scholes para valorar opciones se deduce de ahí, y su supuesto central (volatilidad constante y rendimientos gaussianos) es conocidamente falso en los mercados reales, donde las colas son mucho más gruesas. Usar el modelo sin recordar sus hipótesis ha costado cantidades notables de dinero.

Pausa para pensar

El movimiento lo describió el botánico Robert Brown en 1827 observando granos de polen en agua, y durante ochenta años nadie supo explicarlo. Einstein lo hizo en 1905, en uno de sus cuatro artículos de aquel año, y su explicación fue además el argumento decisivo a favor de la existencia de los átomos: predijo cuánto se desplazaría una partícula en función del tamaño molecular, y Perrin lo midió en 1908 obteniendo el número de Avogadro. La formalización matemática la dio Wiener en 1923, y Itô construyó el cálculo estocástico en los años cuarenta. Un fenómeno visto por un botánico acabó dando a la vez la prueba de los átomos y la maquinaria de los mercados financieros.

22.4. El proceso de Poisson y las colas de espera

El movimiento browniano modela un desplazamiento continuo; el proceso de Poisson modela lo contrario: sucesos aislados que ocurren de golpe en instantes impredecibles. Llamadas a una centralita, desintegraciones radiactivas, clientes que entran en un banco, mutaciones en una cadena de ADN. Es, con diferencia, el proceso estocástico más útil fuera de las matemáticas, y la puerta de entrada a la teoría de colas.

Definición 22.14. Un *proceso de Poisson* de intensidad $\lambda > 0$ es una familia $\{N(t)\}_{t\geq 0}$ de variables que cuentan sucesos, con $N(0) = 0$, incrementos independientes, incrementos estacionarios, y

$$P\big(N(t+s) - N(s) = k\big) = \mathrm{e}^{-\lambda t}\frac{(\lambda t)^k}{k!}, \qquad k = 0, 1, 2, \dots$$

Los tiempos entre sucesos consecutivos $T_1, T_2, \dots$ son independientes y siguen una distribución exponencial de parámetro λ, y el instante del n-ésimo suceso $S_n = T_1 + \cdots + T_n$ sigue una distribución Gamma de parámetros (n, λ).

Idea clave

Qué dice esto en cristiano. Imagínese que en cada instante hay una probabilidad minúscula de que ocurra algo, siempre la misma, y que lo ocurrido hasta ahora no influye en lo que viene. Al pasar al límite continuo, ese "lanzar una moneda muy sesgada muchísimas veces" produce exactamente el proceso de Poisson. Las tres hipótesis se leen así: incrementos independientes es *el pasado no informa del futuro*; incrementos estacionarios es *la tasa no cambia con la hora*; y la fórmula de Poisson es lo que sale obligatoriamente de las dos anteriores más la condición de que no ocurran dos sucesos exactamente a la vez. No hay libertad de elección: si se aceptan las hipótesis, la distribución es esa y no otra.

Regla: Resolver un problema de Poisson en siete pasos

1. **Identificar la unidad de tiempo** y expresar λ en sucesos por unidad.
2. **Si el intervalo mide t unidades**, el número de sucesos es Poisson de parámetro

λt, no λ. Éste es el error más frecuente.

3. **Distinguir "cuántos" de "cuándo"**: preguntas sobre conteos usan la Poisson; preguntas sobre esperas usan la exponencial o la Gamma.
4. **Para "al menos uno"**, calcular $1 - P(N = 0) = 1 - \mathrm{e}^{-\lambda t}$ en lugar de sumar la cola.
5. **Aprovechar la falta de memoria**: $P(T > s + t \mid T > s) = P(T > t)$. Lo que ya se ha esperado no cuenta.
6. **Para superposición**, sumar intensidades: la unión de procesos independientes de intensidades λ_1, λ_2 es Poisson de intensidad $\lambda_1 + \lambda_2$.
7. **Para adelgazamiento**, multiplicar: si cada suceso se conserva con probabilidad p de forma independiente, lo que queda es Poisson de intensidad $p\lambda$.

Ejemplo 22.15 (Una centralita con tres llamadas por hora)**.** Sea $\lambda = 3$ llamadas por hora.

(a) Conteos. En una hora, $P(N = 0) = \mathrm{e}^{-3} = 0{,}049787$, $P(N = 1) = 0{,}149361$, $P(N = 2) = P(N = 3) = 0{,}224042$, $P(N = 4) = 0{,}168031$, $P(N = 5) = 0{,}100819$. Los valores $k = 2$ y $k = 3$ empatan, cosa que ocurre siempre que λ es entero: $P(k)/P(k-1) = \lambda/k$ vale 1 justo en $k = \lambda$. La probabilidad de que suene el teléfono al menos una vez es $1 - \mathrm{e}^{-3} = 0{,}950213$.

(b) Un intervalo distinto. En *dos* horas el parámetro es $\lambda t = 6$, de modo que $P(N(2) = 5) = \mathrm{e}^{-6}6^5/5! = 0{,}160623$. Usar 3 en lugar de 6 daría 0,100819, un error del 37 %.

(c) Esperas. El tiempo hasta la primera llamada es exponencial de media $1/3$ de hora, es decir 20 minutos. La probabilidad de esperar más de media hora es $\mathrm{e}^{-1{,}5} = 0{,}223130$. El tiempo hasta la tercera llamada tiene media $3/3 = 1$ hora y varianza $3/9 = 0{,}333333$.

(d) La falta de memoria en acción. Si ya han pasado 15 minutos sin llamadas, la probabilidad de esperar media hora más sigue siendo $\mathrm{e}^{-1{,}5} = 0{,}223130$, idéntica a la del apartado anterior. El teléfono no "debe" una llamada por haber estado callado. Ésta es la propiedad que hace de la exponencial la única distribución continua sin memoria, y es también la razón de que el modelo sea a menudo demasiado optimista en contextos donde la espera sí desgasta.

Definición 22.16. Una *cola* $M/M/1$ tiene llegadas de Poisson de intensidad λ, tiempos de servicio exponenciales de tasa μ, un solo servidor y disciplina de primero en llegar, primero en ser servido. Si $\rho = \lambda/\mu < 1$, existe distribución estacionaria

$$\pi_n = (1 - \rho)\rho^n, \qquad n = 0, 1, 2, \dots$$

y se cumplen $L = \dfrac{\rho}{1-\rho}$, $L_q = \dfrac{\rho^2}{1-\rho}$, $W = \dfrac{1}{\mu - \lambda}$, $W_q = \dfrac{\lambda}{\mu(\mu - \lambda)}$, ligadas por la *fórmula de Little* $L = \lambda W$.

Idea clave

Qué dice esto en cristiano. ρ es la fracción de tiempo que el servidor está ocupado. La distribución estacionaria es geométrica: la probabilidad de encontrar n personas en el sistema decae multiplicándose por ρ cada vez. La fórmula de Little dice algo casi obvio y sin embargo enormemente general: la cantidad media de gente dentro es la velocidad de entrada por el tiempo medio que cada uno pasa dentro. Vale para cualquier sistema en régimen estacionario, con cualquier distribución y cualquier número de servidores; no usa ninguna hipótesis de Poisson.

Ejemplo 22.17 (Una ventanilla con $\lambda = 3$ y $\mu = 5$ por hora)**. (a) Ocupación y distribución.** $\rho = 3/5 = 0{,}6$, luego el empleado está libre el 40 % del tiempo: $\pi_0 = 0{,}4$, $\pi_1 = 0{,}24$, $\pi_2 = 0{,}144$, $\pi_3 = 0{,}0864$, $\pi_4 = 0{,}05184$. La probabilidad de encontrar tres o más personas es $\rho^3 = 0{,}216$.

(b) Medias. $L = 0{,}6/0{,}4 = 1{,}5$ personas en el sistema y $L_q = 0{,}36/0{,}4 = 0{,}9$ en la cola; la diferencia 0,6 es justo ρ, la persona que está siendo atendida en promedio. El tiempo total en el sistema es $W = 1/(5-3) = 0{,}5$ horas $= 30$ minutos, y el tiempo de espera puro es $W_q = 3/(5 \cdot 2) = 0{,}3$ horas $= 18$ minutos. Comprobación con Little: $\lambda W = 3 \cdot 0{,}5 = 1{,}5 = L$.

(c) La no linealidad de la congestión. Si la demanda sube de 3 a 4 manteniendo $\mu = 5$, entonces $\rho = 0{,}8$ y L salta de 1,5 a $0{,}8/0{,}2 = 4$: un aumento del 33 % en la carga casi triplica la cola. Con $\lambda = 4{,}5$ sale $L = 9$. La función $\rho/(1-\rho)$ explota cerca de 1, y ése es el motivo matemático de que los sistemas reales se derrumben de golpe en vez de degradarse suavemente.

(d) Añadir un servidor en vez de acelerar. Con dos ventanillas de la misma tasa ($M/M/2$ con $\lambda = 3$, $\mu = 5$) sale $\pi_0 = 0{,}538462$ y $L_q = 0{,}059341$, de donde $W_q = 1{,}186813$ minutos frente a los 18 de antes. Duplicar servidores redujo la espera en cola por un factor de quince, mucho más de lo que sugeriría la intuición lineal.

Cuidado

[Contraejemplo: con $\rho \geq 1$ no hay régimen estacionario] Si $\lambda = \mu = 5$, cada cliente se atiende exactamente a la velocidad a la que llegan, y la tentación es concluir que la cola se mantiene estable. Es falso: la cadena es recurrente nula, la cola crece sin límite en probabilidad y $L = \rho/(1-\rho)$ se va a infinito. La razón es la variabilidad, no la media: cuando un servicio se alarga por azar se acumula gente, y no hay ningún mecanismo que recupere el tiempo perdido, porque el servidor no puede ir más rápido que μ ni siquiera cuando hay atasco. Un sistema diseñado con $\rho = 1$ no está ajustado al límite: está roto.

Cuidado

[Contraejemplo: sucesos raros no siempre son Poisson] La Poisson exige independencia. Los terremotos no la cumplen: cada uno dispara réplicas, de modo que los sucesos vienen en racimos y la varianza del conteo supera con creces a la media (sobredispersión). Otro tanto ocurre con los accidentes laborales, donde un mismo trabajador propenso concentra varios. El diagnóstico rápido es comparar media y varianza muestrales: en

una Poisson auténtica coinciden, porque $\mathbb{E}[N] = \mathrm{Var}(N) = \lambda t$. Si la varianza dobla a la media, el modelo correcto es una binomial negativa, no una Poisson.

Pausa para pensar

El proceso de Poisson tiene una caracterización sorprendente: si se condiciona a que en $[0, t]$ hayan ocurrido exactamente n sucesos, entonces los instantes en que ocurrieron se distribuyen como n puntos uniformes independientes en $[0, t]$, reordenados. Es decir, sabiendo cuántos, el "cuándo" es puro azar uniforme sin ninguna estructura. Esto convierte muchos cálculos difíciles en cálculos de estadísticos de orden de uniformes, y explica por qué al mirar un registro de sucesos de Poisson uno cree ver racimos donde no los hay: los puntos uniformes se agrupan mucho más de lo que la intuición espera.

22.5. Tiempos de parada, martingalas y el muestreo opcional

Muchas preguntas sobre procesos no son "¿dónde está en el instante t?" sino "¿cuándo llega a tal sitio y qué valor tiene entonces?". El aparato que responde a eso es el de las martingalas, y su teorema central convierte problemas que parecen requerir cálculos largos en una línea.

Definición 22.18. Un *tiempo de parada* respecto de una filtración $(\mathcal{F}_t)$ es una variable aleatoria τ tal que el suceso $\{\tau \le t\}$ pertenece a $\mathcal{F}_t$ para todo t: para saber si ya ha ocurrido basta con lo observado hasta ese momento. Un proceso (X_t) es una *martingala* si $E[|X_t|] < \infty$ y $E[X_{t+s} \mid \mathcal{F}_t] = X_t$.

Idea clave

Qué dice esto en cristiano. Una martingala es un juego justo: la mejor previsión del capital futuro es el capital de ahora, por mucho que uno sepa del pasado. Un tiempo de parada es una regla de abandono *honesta*: "me retiro cuando pierda mil euros" vale, porque en cada instante se sabe si ya ha ocurrido; "me retiro justo antes de la peor racha" no vale, porque exige mirar el futuro. Toda la teoría vive de esa distinción, y el teorema del muestreo opcional dice lo que uno sospecha y no sabe demostrar: que ninguna regla de abandono honesta convierte un juego justo en ventajoso.

Teorema 22.19 (Muestreo opcional de Doob)**.** *Si (X_t) es una martingala y τ un tiempo de parada acotado, entonces $E[X_\tau] = E[X_0]$. La conclusión se mantiene si τ es finito casi seguramente y la martingala está uniformemente integrable, o si τ tiene esperanza finita y los incrementos están acotados.*

Regla: Resolver con martingalas en siete pasos

1. **Identificar el juego justo**: buscar una cantidad cuya esperanza no cambie. El capital en un juego equitativo, $B(t)$, o $B(t)^2 - t$.

2. **Comprobar que es martingala** verificando $E[X_{t+1} \mid \mathcal{F}_t] = X_t$ paso a paso.
3. **Definir el tiempo de parada** como el primer instante en que se alcanza una frontera.
4. **Verificar las hipótesis** del teorema. Éste es el paso que se salta todo el mundo y donde están los errores.
5. **Escribir** $E[X_\tau] = E[X_0]$ y desarrollar la esperanza sobre los casos posibles.
6. **Despejar** la probabilidad o el tiempo buscados.
7. **Comprobar el resultado** en un caso extremo donde la respuesta sea obvia.

Ejemplo 22.20 (La ruina del jugador, resuelta en cuatro líneas)**. (a) La probabilidad de arruinarse.** Con juego justo, capital inicial i y objetivo N, el capital X_n es una martingala. Sea τ el primer instante en que se alcanza 0 o N. El muestreo opcional da $i = E[X_0] = E[X_\tau] = N \cdot P(\text{éxito})$, luego $P(\text{éxito}) = i/N$. Con $i = 3$ y $N = 10$, la ruina tiene probabilidad 0,7.

(b) El tiempo medio de juego. Ahora se usa la martingala $X_n^2 - n$. Aplicando el mismo teorema, $i^2 = N^2 \cdot \frac{i}{N} - E[\tau]$, de donde $E[\tau] = i(N - i)$. Con $i = 3$ y $N = 10$ salen 21 jugadas de media: una partida justa entre alguien con 3 y alguien con 7 dura veintiuna manos en promedio.

(c) El caso sesgado. Si la probabilidad de ganar es $p \neq 1/2$, el capital ya no es martingala, pero $((1-p)/p)^{X_n}$ sí lo es. El mismo argumento da $P(\text{éxito}) = \frac{1-r^i}{1-r^N}$ con $r = (1-p)/p$. Con $p = 0{,}49$, $i = 50$ y $N = 100$, la probabilidad de duplicar el capital cae a 0,119: una desventaja del uno por ciento por mano basta para hundirla.

(d) Por qué el casino gana siempre. La conclusión de (c) llevada al límite: contra una banca de capital infinito y con $p < 1/2$, la ruina tiene probabilidad 1 para cualquier estrategia y cualquier tiempo de parada honesto. Ninguna martingala en el sentido coloquial —doblar la apuesta tras cada pérdida— lo evita, y de hecho el nombre técnico viene de ahí.

Cuidado

[Contraejemplo: sin hipótesis, el muestreo opcional es falso] Tómese un paseo aleatorio simétrico desde 0 y sea τ el primer instante en que vale 1. Ese τ es finito con probabilidad 1, y sin embargo $E[X_\tau] = 1 \neq 0 = E[X_0]$. El teorema falla porque $E[\tau] = \infty$ y la martingala no está uniformemente integrable. La estrategia "juego hasta ganar un euro y me retiro" funciona con probabilidad 1, pero exige un tiempo de espera de esperanza infinita y, por el camino, pérdidas arbitrariamente grandes. Es el ejemplo que hay que tener en la cabeza cada vez que alguien presenta un sistema infalible.

22.6. Cálculo de Itô y la ecuación de Black–Scholes

El movimiento browniano no es derivable en ningún punto, de modo que $\mathrm{d}B$ no significa nada en el sentido habitual. Itô encontró la manera de definir $\int f \,\mathrm{d}B$ de todos modos, y el

precio es una regla de la cadena con un término de más. Ese término es toda la diferencia.

Definición 22.21. La *integral de Itô* de un proceso adaptado H es el límite en media cuadrática de las sumas $\sum_i H_{t_i}(B_{t_{i+1}} - B_{t_i})$, evaluando siempre H en el *extremo izquierdo* de cada subintervalo. Si $\mathrm{d}X = \mu\,\mathrm{d}t + \sigma\,\mathrm{d}B$ y f es de clase C^2, la *fórmula de Itô* afirma

$$\mathrm{d}f(X) = f'(X)\,\mathrm{d}X + \tfrac{1}{2}f''(X)\,\sigma^2\,\mathrm{d}t.$$

Idea clave

Qué dice esto en cristiano. Comparado con la regla de la cadena de siempre sobra el segundo término. Está ahí porque $(\mathrm{d}B)^2$ no es despreciable: la variación cuadrática del browniano en un intervalo de longitud t vale exactamente t, no cero. La insistencia en evaluar en el extremo izquierdo tampoco es un tecnicismo: significa que la decisión de cuánto invertir se toma *antes* de ver el movimiento del precio, y es lo que hace que la integral de Itô sea una martingala. Elegir el punto medio da otra integral, la de Stratonovich, que respeta la regla de la cadena clásica y no es martingala. Itô para finanzas y probabilidad, Stratonovich para física.

Ejemplo 22.22 (La fórmula de Itô en acción)**. (a) El cuadrado del browniano.** Con $f(x) = x^2$ y $X = B$ resulta $\mathrm{d}(B^2) = 2B\,\mathrm{d}B + \mathrm{d}t$, de donde $\int_0^t B\,\mathrm{d}B = \frac{1}{2}(B_t^2 - t)$. Ese $-t$ es exactamente lo que hace que $B_t^2 - t$ sea una martingala, como se usó en la ruina del jugador.

(b) El movimiento browniano geométrico. Si $\mathrm{d}S = \mu S\,\mathrm{d}t + \sigma S\,\mathrm{d}B$, aplicar Itô a $\log S$ da $S_t = S_0 \exp((\mu - \sigma^2/2)t + \sigma B_t)$. Ese $-\sigma^2/2$ es la corrección de Itô y explica por qué el rendimiento medio geométrico de un activo es menor que el aritmético: la volatilidad se come rentabilidad.

(c) La ecuación de Black–Scholes. Construyendo una cartera que combine la opción y el activo de modo que el término aleatorio se cancele, la ausencia de arbitraje obliga a que el valor $V(S,t)$ cumpla $V_t + \frac{1}{2}\sigma^2 S^2 V_{SS} + rSV_S - rV = 0$. Es una ecuación en derivadas parciales parabólica: con el cambio adecuado se convierte en la ecuación del calor del capítulo 13.

(d) La fórmula, con números. Para una opción de compra europea con $S = K = 100$, $r = 0{,}05$, $\sigma = 0{,}20$ y $T = 1$ año se obtiene $d_1 = 0{,}350$, $d_2 = 0{,}150$ y un precio de 10,45. La opción de venta vale 5,57, y la diferencia $10{,}45 - 5{,}57 = 4{,}88$ coincide con $S - K\mathrm{e}^{-rT} = 4{,}88$, que es la paridad entre compra y venta. La delta, $N(d_1) = 0{,}637$, dice cuántas acciones hay que tener para cubrirla.

Cuidado

[Contraejemplo: el modelo supone lo que la realidad no cumple] Black–Scholes supone volatilidad constante, rendimientos normales y negociación continua sin costes. Los tres supuestos son falsos. Si los rendimientos fuesen normales con $\sigma = 0{,}20$, una caída diaria del 20 por ciento como la del 19 de octubre de 1987 tendría una probabilidad de aproximadamente 10^{-50}, es decir, no habría ocurrido nunca en la historia del universo. Ocurrió. El mercado lo sabe y lo compensa cotizando volatilidades implícitas mayores

para los precios de ejercicio alejados, lo que produce la *sonrisa de volatilidad*: una deformación sistemática que es, literalmente, la huella del error del modelo dibujada por los precios.

Pausa para pensar

Itô desarrolló su cálculo durante la guerra, en un Japón aislado, y publicó los resultados principales entre 1942 y 1951; tardaron años en llegar a Occidente. Black, Scholes y Merton publicaron la fórmula en 1973, el mismo año en que abrió el mercado de opciones de Chicago, y en pocos meses los operadores llevaban tablas impresas en el bolsillo. En 1998 el fondo que Scholes y Merton asesoraban, *Long-Term Capital Management*, quebró de forma espectacular por subestimar precisamente los sucesos extremos que su modelo declaraba imposibles.

22.7. Recurrencia, reversibilidad y las cadenas que se diseñan a propósito

Hemos usado "irreducible", "aperiódica" y "teorema ergódico" como si estuvieran definidos. Toca definirlos, porque de ahí sale la aplicación más importante que tienen hoy las cadenas de Markov.

Definición 22.23 (Clasificación de estados)**.** Un estado i es **recurrente** si la cadena vuelve a él con probabilidad 1, y **transitorio** en otro caso. Es **recurrente positivo** si además el tiempo medio de retorno $\mu_i = E[T_i \mid X_0 = i]$ es finito, y **recurrente nulo** si vuelve con probabilidad 1 pero tarda infinito en media. El **periodo** de i es el máximo común divisor de los tiempos en que el retorno es posible; si vale 1, el estado es **aperiódico**.

Teorema 22.24 (Ergódico)**.** *Si una cadena es irreducible, aperiódica y recurrente positiva, entonces existe una única distribución estacionaria π, se cumple $P^n_{ij} \to \pi_j$ para todo par (i,j), y además*

$$\frac{1}{n}\sum_{k=1}^{n} f(X_k) \;\longrightarrow\; \sum_{j} \pi_j f(j) \qquad \text{casi seguramente,}$$

sea cual sea el estado inicial. Además $\pi_i = 1/\mu_i$.

La segunda parte es la que importa en la práctica: los promedios *a lo largo de una trayectoria* convergen a los promedios *sobre la distribución estacionaria*. Es una ley de los grandes números para datos dependientes.

Ejemplo 22.25 (Paseo aleatorio en $\mathbb{Z}^d$: el teorema de Pólya)**.** El paseo aleatorio simple —en cada paso, un vecino al azar— es recurrente en dimensiones 1 y 2, y transitorio a partir de 3. La probabilidad de volver alguna vez al origen es

$$d=1:\ 1, \qquad d=2:\ 1, \qquad d=3:\ 0{,}340\,5,$$

es decir, en el espacio el paseante se pierde definitivamente casi dos veces de cada tres. Shizuo Kakutani lo resumió así: un borracho siempre vuelve a casa, pero un pájaro borracho puede perderse para siempre. En dimensiones 1 y 2 la recurrencia es además *nula*: vuelve seguro, pero el tiempo medio de retorno es infinito.

Definición 22.26 (Balance detallado y reversibilidad)**.** Una distribución π cumple **balance detallado** respecto de P si

$$\pi_i P_{ij} = \pi_j P_{ji} \qquad \text{para todo } i, j.$$

Sumando en i se obtiene $\pi P = \pi$: el balance detallado implica estacionariedad, pero no al revés.

Idea clave

Y aquí ocurre el giro que cambió la estadística computacional. Durante décadas se planteaba el problema en un sentido: dada una cadena, ¿cuál es su π? En 1953 Metropolis y sus colaboradores lo plantearon al revés: *dada* una π que uno quiere muestrear pero no sabe cómo, ¿se puede fabricar una cadena cuya estacionaria sea exactamente esa π? El balance detallado dice que sí, y da la receta.

Regla: Metropolis–Hastings

Se elige una distribución de propuesta $q(\cdot \mid x)$ cualquiera. Desde el estado x:

1. Propón $y \sim q(\cdot \mid x)$.
2. Calcula $\alpha = \min\Big\{1,\ \dfrac{\pi(y)\, q(x \mid y)}{\pi(x)\, q(y \mid x)}\Big\}$.
3. Acepta y con probabilidad α; en caso contrario quédate en x.

La cadena resultante cumple balance detallado respecto de π por construcción.

Ejemplo 22.27 (Cinco estados, una cadena hecha a medida)**.** Se quiere muestrear

$$\pi = (0{,}1;\ 0{,}2;\ 0{,}4;\ 0{,}2;\ 0{,}1)$$

sobre $\{1, \dots, 5\}$. Como propuesta, un paseo simétrico al vecino de la izquierda o de la derecha con probabilidad $1/2$ (y rebote en los extremos). Como q es simétrica, $\alpha = \min\{1, \pi(y)/\pi(x)\}$.

Desde el estado 2: proponer el 3 se acepta con probabilidad $\min\{1,\ 0{,}4/0{,}2\} = 1$ (siempre, porque el 3 es más probable); proponer el 1 se acepta con probabilidad $\min\{1,\ 0{,}1/0{,}2\} = 0{,}5$. Luego

$$P_{23} = \tfrac{1}{2}, \qquad P_{21} = \tfrac{1}{2} \cdot \tfrac{1}{2} = \tfrac{1}{4}, \qquad P_{22} = \tfrac{1}{4}.$$

Comprobación del balance detallado entre 2 y 3: por el otro lado, desde 3 proponer el 2 se acepta con probabilidad $0{,}2/0{,}4 = 0{,}5$, luego $P_{32} = \frac{1}{2} \cdot \frac{1}{2} = \frac{1}{4}$, y

$$\pi_2 P_{23} = 0{,}2 \cdot 0{,}5 = 0{,}1, \qquad \pi_3 P_{32} = 0{,}4 \cdot 0{,}25 = 0{,}1. \ \checkmark$$

Obsérvese lo esencial: en ningún momento se ha necesitado la constante de normalización de π, porque solo aparecen cocientes. Eso es lo que hace el método utilizable en estadística bayesiana, donde la posteriori se conoce salvo una constante que es una integral imposible.

Observación 22.28. El muestreador de Gibbs es el caso particular en que se propone actualizar una sola coordenada, sorteándola de su condicional completa; entonces $\alpha = 1$ siempre y no hay rechazo. Es lo que hace practicables los modelos jerárquicos con cientos de parámetros.

Las dos advertencias de rigor: la cadena tarda en olvidar el punto de partida (periodo de calentamiento, que se descarta) y las muestras sucesivas están correlacionadas, de modo que n iteraciones no equivalen a n datos independientes. El *tamaño muestral efectivo* $n/(1+2\sum_k \rho_k)$ mide cuántas valen de verdad, y puede ser cien veces menor.

22.8. Propiedad fuerte de Markov, principio de reflexión y tiempos de llegada

Teorema 22.29 (Propiedad fuerte de Markov)**.** *Si τ es un tiempo de parada finito casi seguramente, entonces $(B_{\tau+t} - B_\tau)_{t\geq 0}$ es un movimiento browniano independiente de todo lo ocurrido hasta τ.*

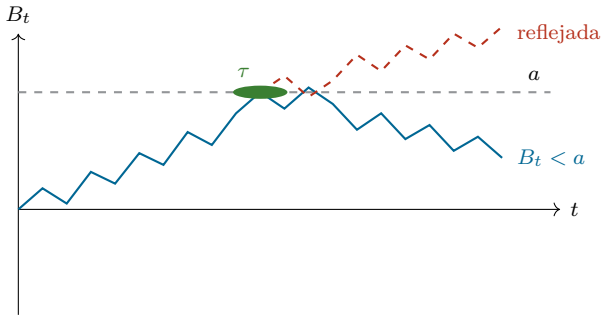


Figura 22.2: Al llegar al nivel a en el instante τ, la propiedad fuerte de Markov dice que el proceso empieza de nuevo. Reflejar el tramo posterior empareja biyectivamente las trayectorias que tocan a y acaban debajo con las que acaban encima, y de ahí $P(\max_{s\leq t} B_s \geq a) = 2P(B_t \geq a)$: exactamente el doble.

Es decir: el browniano no solo empieza de nuevo en cada instante fijo, sino también en instantes elegidos por una regla que no mire al futuro. De ahí sale el argumento más elegante del tema.

Teorema 22.30 (Principio de reflexión)**.** *Para $a > 0$,*

$$P\Big(\max_{0\leq s\leq t} B_s \geq a\Big) = 2\,P(B_t \geq a) = 2\Big(1 - \Phi(a/\sqrt{t})\Big).$$

Demostración. Sea τ el primer instante en que la trayectoria toca a. Si $\tau \leq t$, la propiedad fuerte de Markov dice que a partir de ahí el proceso arranca de nuevo desde a, y por simetría tiene la misma probabilidad de acabar por encima que por debajo de a en el instante t.

Reflejando el tramo posterior a τ se emparejan biyectivamente las trayectorias que tocan a y acaban debajo con las que acaban encima. Luego $P(\tau \le t) = 2P(B_t \ge a)$, salvo el conjunto de medida nula donde $B_t = a$ exactamente. □

Ejemplo 22.31 (El máximo dobla al valor final)**.** Con $t = 1$ y $a = 1$:

$$P(B_1 \ge 1) = 1 - \Phi(1) = 0{,}158\,655, \qquad P\Big(\max_{s\le 1} B_s \ge 1\Big) = 2 \cdot 0{,}158\,655 = 0{,}317\,311.$$

La probabilidad de *haber pasado* por el nivel 1 es exactamente el doble de la de *estar* por encima al final. Uno de cada dos caminos que tocan el uno vuelve a bajar.

De aquí se obtiene además la distribución del tiempo de llegada $\tau_a = \inf\{t : B_t = a\}$:

$$P(\tau_a \le t) = 2\big(1 - \Phi(a/\sqrt{t})\big), \qquad f_{\tau_a}(t) = \frac{a}{\sqrt{2\pi t^3}} e^{-a^2/2t}.$$

La mediana se obtiene resolviendo $2(1-\Phi(1/\sqrt{t})) = 1/2$, es decir $1/\sqrt{t} = \Phi^{-1}(0{,}75) = 0{,}674\,49$, luego

$$\text{mediana}(\tau_1) = 1/0{,}674\,49^2 = 2{,}198\,4.$$

Y sin embargo $E[\tau_a] = +\infty$: la densidad decae como $t^{-3/2}$, cuya integral por t diverge. Se llega con probabilidad 1, en un tiempo mediano de dos unidades… y con esperanza infinita. Esa disonancia es la firma de las leyes de cola pesada, y τ_a es de hecho una distribución estable de índice $1/2$.

Idea clave

La misma tensión aparece en la ruina del jugador con juego justo: se acaba tocando cualquier nivel prefijado con probabilidad 1, pero el tiempo esperado para hacerlo es infinito. Una estrategia que "siempre acaba ganando" y necesita capital y tiempo ilimitados no es una estrategia. Es la respuesta matemática a la martingala del casino, y la razón de que las casas pongan límite de apuesta.

Observación 22.32 (Por qué el browniano y no otro)*.* El principio de invariancia de Donsker explica por qué este proceso aparece en todas partes. Tómese cualquier paseo aleatorio con pasos independientes de media 0 y varianza σ^2 finita; reescálese el tiempo por n y el espacio por $\sigma\sqrt{n}$. La trayectoria reescalada converge en distribución al movimiento browniano, sea cual sea la ley de los pasos. Es el teorema central del límite elevado de variables a *trayectorias*: igual que la normal es el límite universal de las sumas, el browniano es el límite universal de los caminos. Por eso modela el precio de un activo, la posición de un grano de polen y el error acumulado de un reloj, sin que ninguno de los tres tenga nada que ver con los otros.

22.9. Procesos de ramificación: extinción, apellidos y epidemias

Definición 22.33 (Galton–Watson)**.** Cada individuo de la generación n produce, independientemente de los demás, un número aleatorio de descendientes con la misma distribución $p_k = P(X = k)$. Con $Z_0 = 1$,

$$Z_{n+1} = \sum_{i=1}^{Z_n} X_i^{(n)}.$$

Sea $f(s) = \sum_k p_k s^k$ la función generatriz y $m = E[X] = f'(1)$ la descendencia media.

Teorema 22.34 (Momentos y extinción)**.** *$E[Z_n] = m^n$, y la generatriz de Z_n es la n-ésima iterada $f^{(n)} = f \circ \cdots \circ f$. La probabilidad de extinción $q = \lim P(Z_n = 0)$ es la* menor *raíz no negativa de*

$$s = f(s),$$

y cumple:

$$q = 1 \text{ si } m \le 1, \qquad q < 1 \text{ si } m > 1.$$

El caso $m = 1$ es el que sorprende: la población *no* decrece en media —cada generación tiene el mismo tamaño esperado que la anterior— y sin embargo se extingue con probabilidad 1. Es un martingala no negativa, y toda martingala no negativa converge; el único límite compatible con valores enteros y media constante es el cero, alcanzado con probabilidad total. La media se mantiene porque las poquísimas trayectorias que sobreviven crecen sin control.

Ejemplo 22.35 (Una descendencia concreta)**.** Sea $p_0 = 0{,}25$, $p_1 = 0{,}25$, $p_2 = 0{,}5$. Entonces

$$m = 0 \cdot 0{,}25 + 1 \cdot 0{,}25 + 2 \cdot 0{,}5 = 1{,}25 > 1,$$

así que hay esperanza de supervivencia. La ecuación de punto fijo es

$$s = 0{,}25 + 0{,}25\, s + 0{,}5\, s^2 \iff 2s^2 - 3s + 1 = 0 \iff (2s-1)(s-1) = 0,$$

con raíces $s = \frac{1}{2}$ y $s = 1$. La menor es

$$q = 0{,}5.$$

El linaje se extingue en la mitad de los casos pese a que en media crece un 25 por ciento por generación. Y en las mitades que sobreviven crece exponencialmente: $E[Z_{10}] = 1{,}25^{10} = 9{,}31$, que es la media de un 50 por ciento de ceros y un 50 por ciento de valores en torno a 18.

Idea clave

Este modelo es el esqueleto de tres cosas distintas.
Apellidos. Es el problema original de Galton y Watson (1875): por qué se extinguen los apellidos de las familias aristocráticas. La respuesta es que con m apenas por encima de 1 la extinción sigue siendo muy probable.
Epidemias. El parámetro m es el número reproductivo básico R_0. El umbral $m = 1$ es el umbral epidémico: con $R_0 \le 1$ el brote se apaga con certeza; con $R_0 > 1$ hay una probabilidad positiva —exactamente $1 - q$— de epidemia. Con descendencia de Poisson de media R_0, la ecuación es $q = e^{-R_0(1-q)}$, y para $R_0 = 2$ da $q = 0{,}2032$: cuatro de cada cinco introducciones del patógeno prenden.
Reacciones en cadena. La masa crítica de un material fisible es exactamente la geometría que hace pasar m de menos de uno a más de uno. La transición en el umbral es abrupta, y esa abruptez es el fenómeno físico.

22.10. Cadenas en tiempo continuo y ecuaciones diferenciales estocásticas

Las fórmulas de la cola $M/M/1$ se enunciaron sin deducirse. La maquinaria que falta es la de las cadenas en tiempo continuo, y es breve.

Definición 22.36 (Generador infinitesimal)**.** Una cadena de Markov en tiempo continuo queda descrita por las tasas $q_{ij} \geq 0$ de salto de i a j, con $q_{ii} = -\sum_{j \neq i} q_{ij}$. La matriz $Q = (q_{ij})$ es el **generador**, y

$$P(t) = e^{tQ}, \qquad \underbrace{P'(t) = P(t)Q}_{\text{progresiva}}, \qquad \underbrace{P'(t) = QP(t)}_{\text{regresiva}}.$$

La distribución estacionaria cumple $\pi Q = 0$.

Ejemplo 22.37 (Dos estados, resuelto entero)**.** Sea $Q = \begin{pmatrix} -2 & 2 \\ 3 & -3 \end{pmatrix}$: del estado 1 se sale a tasa 2 y del 2 a tasa 3. Los autovalores de Q son 0 y -5, y

$$P(t) = e^{tQ} = \frac{1}{5}\begin{pmatrix} 3 & 2 \\ 3 & 2 \end{pmatrix} + \frac{e^{-5t}}{5}\begin{pmatrix} 2 & -2 \\ -3 & 3 \end{pmatrix}.$$

Con $t = 0$ da la identidad, y con $t \to \infty$ todas las filas convergen a $\pi = (3/5,\ 2/5)$, que es lo que resuelve $\pi Q = 0$. El tiempo de relajación es $1/5$: tras $t = 1$ la memoria del estado inicial vale $e^{-5} = 0{,}0067$, prácticamente nada.

Teorema 22.38 (Nacimiento y muerte, y de ahí la cola $M/M/1$)**.** *Si solo hay saltos a vecinos, con tasas λ_n hacia arriba y μ_n hacia abajo, el balance entre estados consecutivos da*

$$\lambda_{n-1}\pi_{n-1} = \mu_n \pi_n.$$

Con $\lambda_n = \lambda$ y $\mu_n = \mu$ constantes y $\rho = \lambda/\mu < 1$, la recursión se resuelve sola:

$$\pi_n = \rho^n \pi_0 \quad \Longrightarrow \quad \pi_n = (1-\rho)\rho^n,$$

una geométrica. De ahí $L = E[N] = \dfrac{\rho}{1-\rho}$ y, por la ley de Little, $W = L/\lambda$.

Las fórmulas de la cola dejan de ser recetas: salen de una ecuación de balance de una línea. Y la extensión a $M/M/c$ solo cambia $\mu_n = \min(n, c)\mu$.

Ecuaciones diferenciales estocásticas

Teorema 22.39 (Existencia y unicidad)**.** *La ecuación $\mathrm{d}X_t = b(X_t, t)\mathrm{d}t + \sigma(X_t, t)\mathrm{d}B_t$ con X_0 dado tiene solución única si b y σ son Lipschitz en x y de crecimiento a lo sumo lineal. Es el análogo exacto de Picard, con la misma demostración por iteración.*

Ejemplo 22.40 (Ornstein–Uhlenbeck: reversión a la media)**.**

$$\mathrm{d}X_t = -\theta X_t\,\mathrm{d}t + \sigma\,\mathrm{d}B_t.$$

El término de deriva empuja siempre hacia el cero, con fuerza proporcional a la distancia; el de ruido empuja al azar. Multiplicando por $e^{\theta t}$ y aplicando Itô se resuelve explícitamente, y resulta gaussiano con

$$E[X_t] = x_0 e^{-\theta t}, \qquad \operatorname{Var}(X_t) = \frac{\sigma^2}{2\theta}(1 - e^{-2\theta t}).$$

Con $\theta = 1$, $\sigma = 0{,}5$ y $x_0 = 3$:

t	0,5	1	5
media	1,819 6	1,103 6	0,020 2
desv. típica	0,281 1	0,328 8	0,353 5

La media decae exponencialmente con vida media $\log 2/\theta = 0{,}693$, y la varianza satura en $\sigma^2/2\theta = 0{,}125$, es decir, desviación típica 0,353 553. La ley estacionaria es $N(0,\ \sigma^2/2\theta)$ y ya está prácticamente alcanzada en $t = 5$.

Este proceso modela el tipo de interés (Vasicek), la velocidad de una partícula browniana (que es su origen físico, 1930) y cualquier magnitud que fluctúe alrededor de un nivel de equilibrio. A diferencia del browniano geométrico, no crece sin límite.

Regla: el diccionario que cierra el libro

Cada objeto estocástico tiene su gemelo determinista, y el puente tiene nombre.

- **Fokker–Planck** (o Kolmogórov progresiva): la densidad $p(x,t)$ de la solución resuelve
$$\partial_t p = -\partial_x(b\,p) + \tfrac{1}{2}\partial_x^2(\sigma^2 p),$$
una ecuación en derivadas parciales de tipo parabólico. Un proceso aleatorio y una ecuación del calor generalizada son el mismo objeto visto de dos maneras.
- **Feynman–Kac**: la solución de $\partial_t u + b\,\partial_x u + \tfrac{1}{2}\sigma^2\partial_x^2 u - ru = 0$ con dato final $u(x,T) = g(x)$ se escribe como una esperanza,
$$u(x,t) = E\big[e^{-r(T-t)} g(X_T) \mid X_t = x\big].$$
Esa fórmula es la razón de que la ecuación de Black–Scholes se pueda resolver como un valor esperado, y de que se pueda valorar por simulación cuando no hay fórmula cerrada.
- **Girsanov**: cambiar de medida de probabilidad equivale a cambiar la deriva. Es lo que permite sustituir el rendimiento real del activo por el tipo sin riesgo y valorar como si todo el mundo fuera neutral al riesgo.
- **Euler–Maruyama**: $X_{n+1} = X_n + b\,h + \sigma\sqrt{h}\,Z_n$ con Z_n normales estándar. Tiene orden fuerte $\frac{1}{2}$ —el error de una trayectoria concreta va como $\sqrt{h}$— y orden débil

1 para las esperanzas. Ese $\frac{1}{2}$ es la firma del cálculo estocástico: el ruido aporta $\sqrt{h}$, no h.

22.11. Chapman–Kolmogórov y el teorema ergódico

Se ha calculado con potencias de la matriz de transición y se ha invocado el teorema ergódico sin haber justificado ni una cosa ni la otra. Se hace ahora.

Teorema 22.41 (Ecuaciones de Chapman–Kolmogórov)**.** *Sea (X_n) una cadena de Markov homogénea con matriz de transición P, y $p_{ij}^{(n)} = P(X_n = j \mid X_0 = i)$. Entonces, para todos $m, n \geq 0$,*

$$p_{ij}^{(m+n)} = \sum_k p_{ik}^{(m)}\, p_{kj}^{(n)}, \qquad \textit{es decir} \qquad P^{(m+n)} = P^{(m)} P^{(n)}, \quad \textit{y} \quad p_{ij}^{(n)} = (P^n)_{ij}.$$

Demostración. Por la ley de la probabilidad total, partiendo según el estado en el instante intermedio m:

$$p_{ij}^{(m+n)} = P(X_{m+n} = j \mid X_0 = i) = \sum_k P\big(X_{m+n} = j \mid X_m = k,\ X_0 = i\big)\, P(X_m = k \mid X_0 = i).$$

La propiedad de Markov dice que, condicionado a $X_m = k$, el futuro es independiente del pasado, luego el primer factor es $P(X_{m+n} = j \mid X_m = k) = p_{kj}^{(n)}$ por homogeneidad. El segundo es $p_{ik}^{(m)}$. Eso es exactamente la fórmula, y en notación matricial es el producto de matrices. La última afirmación se sigue por inducción: $P^{(1)} = P$ y $P^{(n+1)} = P^{(n)}P = P^n P = P^{n+1}$. □

Observación 22.42. Toda la "magia" de calcular probabilidades a n pasos multiplicando matrices es esta cuenta de tres líneas, y su único ingrediente es la propiedad de Markov. Nótese además qué se ha usado: que condicionar al presente borra el pasado. Si la cadena tuviera memoria de dos pasos, la factorización fallaría y P^n no significaría nada.

Definición 22.43 (Irreducible, aperiódica, recurrente positiva)**.** Una cadena es *irreducible* si desde cualquier estado se puede llegar a cualquier otro en algún número de pasos. El *periodo* de un estado i es $d(i) = \gcd\{n \geq 1 : p_{ii}^{(n)} > 0\}$; la cadena es *aperiódica* si $d(i) = 1$ (en una cadena irreducible el periodo es común a todos los estados). Un estado es *recurrente* si el tiempo de retorno $\tau_i = \min\{n \geq 1 : X_n = i\}$ es finito con probabilidad 1, y *recurrente positivo* si además $\mu_i = \mathrm{E}[\tau_i \mid X_0 = i] < \infty$.

Teorema 22.44 (Ergódico para cadenas de Markov)**.** *Sea una cadena irreducible, aperiódica y con espacio de estados finito (o, más generalmente, recurrente positiva). Entonces:*

1. *existe una única distribución estacionaria π, es decir con $\pi P = \pi$ y $\sum_i \pi_i = 1$, y todos sus componentes son positivos;*
2. *$p_{ij}^{(n)} \to \pi_j$ para todos i, j, con independencia del estado inicial;*

3. $\pi_i = 1/\mu_i$*: la proporción de tiempo en un estado es el inverso de su tiempo medio de retorno;*

4. (ley de los grandes números para trayectorias.) *Para toda función acotada f,*

$$\frac{1}{N}\sum_{n=1}^{N} f(X_n) \;\longrightarrow\; \sum_i \pi_i f(i) \qquad \textit{casi seguro.}$$

Demostración de (1) y (2) en el caso finito, por acoplamiento. Paso 1: alguna potencia es estrictamente positiva. Por irreducibilidad, para cada par (i,j) hay un n_{ij} con $p_{ij}^{(n_{ij})} > 0$; y por aperiodicidad, el conjunto $\{n : p_{ii}^{(n)} > 0\}$ está cerrado por sumas y tiene máximo común divisor 1, luego contiene a todos los enteros suficientemente grandes: si un conjunto de enteros positivos cerrado por sumas tiene máximo común divisor 1, la identidad de Bézout produce una combinación entera igual a 1, y sumando múltiplos adecuados se alcanzan todos los enteros a partir de cierto umbral. Combinando ambas cosas, existe m con $(P^m)_{ij} > 0$ para todos i, j.

Paso 2: contracción. Sea $\varepsilon = \min_{i,j}(P^m)_{ij} > 0$ y S el número de estados, de modo que $S\varepsilon \le 1$. Para dos distribuciones μ, ν sobre los estados, la distancia en variación total es $\|\mu - \nu\| = \frac{1}{2}\sum_i |\mu_i - \nu_i|$. Se comprueba que

$$\|\mu P^m - \nu P^m\| \;\le\; (1 - S\varepsilon)\,\|\mu - \nu\|.$$

La razón es un acoplamiento: escríbase $P^m = S\varepsilon\,\Pi + (1 - S\varepsilon)Q$, donde Π tiene todas las filas iguales a la distribución uniforme y Q es otra matriz estocástica (los coeficientes salen no negativos por la definición de ε). Al aplicar Π, cualquier distribución se convierte en la misma, luego esa parte no aporta diferencia; queda la fracción $1 - S\varepsilon$.

Paso 3: punto fijo. La aplicación $\mu \mapsto \mu P^m$ es una contracción sobre el conjunto completo de las distribuciones de probabilidad con la distancia en variación total. Por el teorema del punto fijo de Banach (capítulo 12) tiene un único punto fijo π, y $\mu P^{mk} \to \pi$ geométricamente para toda μ inicial. Que π sea también punto fijo de P se sigue de que πP es punto fijo de P^m y de la unicidad. Aplicando lo anterior a $\mu = \delta_i$ se obtiene (2) a lo largo de los múltiplos de m, y los pasos intermedios se tratan igual componiendo con P^r. La positividad de los π_j sale de $\pi_j = \sum_i \pi_i (P^m)_{ij} \ge \varepsilon$. □

Observación 22.45 (Qué añade cada hipótesis)*.* Las tres hipótesis no son decorativas, y quitarlas rompe conclusiones distintas. *Sin irreducibilidad*, la cadena puede quedarse atrapada en una clase y no hay unicidad: la matriz identidad tiene infinitas estacionarias. *Sin aperiodicidad*, la estacionaria existe y es única, pero P^n no converge: la cadena que alterna entre dos estados tiene $\pi = (\frac{1}{2}, \frac{1}{2})$ y $p_{11}^{(n)}$ oscila entre 0 y 1 para siempre —aunque los *promedios* sí convergen, y por eso (4) sigue valiendo—. *Sin recurrencia positiva* (en espacios infinitos), puede no haber estacionaria en absoluto: el paseo aleatorio simple sobre $\mathbb{Z}$ es irreducible y recurrente, pero su tiempo medio de retorno es infinito y no admite distribución estacionaria.

Definición 22.46 (Uniformemente integrable)**.** Una familia $\{Y_\alpha\}$ de variables aleatorias es *uniformemente integrable* si

$$\lim_{K\to\infty} \sup_{\alpha} \mathrm{E}\big[|Y_\alpha|\,\mathbf{1}_{\{|Y_\alpha|>K\}}\big] = 0.$$

Observación 22.47 (Qué impide, y por qué aparece en el muestreo opcional). La condición dice que ninguna variable de la familia esconde masa en las colas lejanas, de manera uniforme. Una sola variable integrable siempre la cumple; el contenido está en el "sup". Es exactamente la hipótesis que convierte convergencia en probabilidad en convergencia en L^1, y por eso aparece en el teorema del muestreo opcional de Doob: sin ella, se puede tener $M_{\tau\wedge n}\to M_\tau$ casi seguro y aun así $\mathrm{E}[M_\tau]\neq\mathrm{E}[M_0]$.

El contraejemplo canónico es la estrategia de doblar la apuesta. Sea M_n la fortuna de un jugador que apuesta a cara o cruz doblando tras cada pérdida, y τ el instante de la primera cara. Entonces M_n es una martingala con $\mathrm{E}[M_0]=0$, y $M_\tau=1$ con probabilidad 1: parece que se ha ganado dinero de la nada. Lo que falla es la integrabilidad uniforme: $M_{\tau\wedge n}$ vale 1 con probabilidad $1-2^{-n}$ y $-(2^n-1)$ con probabilidad 2^{-n}, de modo que $\mathrm{E}[|M_{\tau\wedge n}|\mathbf{1}_{\{|M|>K\}}]\approx 1$ para todo K: la familia esconde, para siempre, una pérdida catastrófica de probabilidad diminuta. El teorema no se viola; la hipótesis no se cumple. Y esa es también la razón por la que el sistema de la martingala arruina a quien lo usa: exige capital infinito.

22.12. Filtraciones, estados absorbentes y cómo analizar una cadena

Definición 22.48 (Filtración y proceso adaptado)**.** Una *filtración* es una familia creciente de σ-álgebras $(\mathcal{F}_t)_{t\geq 0}$ sobre el espacio de probabilidad, con $\mathcal{F}_s\subset\mathcal{F}_t$ para $s\leq t$. Intuitivamente, $\mathcal{F}_t$ es *toda la información disponible en el instante* t: los sucesos de los que a esa altura ya se puede decir si han ocurrido o no.

Un proceso (X_t) es *adaptado* a la filtración si X_t es medible respecto de $\mathcal{F}_t$ para cada t; es decir, si su valor en el instante t se conoce con la información de ese instante. La filtración *natural* de un proceso es $\mathcal{F}_t=\sigma(X_s : s\leq t)$, la generada por su propio pasado.

Observación 22.49 (Para qué hace falta). Sin filtración no se pueden enunciar dos cosas del capítulo. La primera es un *tiempo de parada*: τ lo es si $\{\tau\leq t\}\in\mathcal{F}_t$, o sea, si para decidir si ya se ha parado basta con lo observado hasta t y no hace falta mirar el futuro. "Vender cuando la acción baje un 10 %" es tiempo de parada; "vender en el máximo" no lo es, y por eso nadie puede ejecutarlo. La segunda es la definición de *martingala*, $\mathrm{E}[X_t\mid\mathcal{F}_s]=X_s$, que dice literalmente "la mejor predicción del futuro dado todo lo conocido es el valor actual" y sería incomprensible sin un objeto que represente "lo conocido". Y la integral de Itô exige que el integrando sea adaptado: no se puede apostar usando información que aún no existe.

Definición 22.50 (Clasificación de estados)**.** En una cadena de Markov con matriz P:

- j es *accesible* desde i si $p_{ij}^{(n)}>0$ para algún $n\geq 0$; si además i es accesible desde j, ambos *comunican*. Comunicar es una relación de equivalencia, y sus clases son las *clases de comunicación*.
- Una clase es *cerrada* si no se puede salir de ella. La cadena es *irreducible* si hay una sola clase.
- i es *absorbente* si $p_{ii}=1$: una vez allí, no se sale nunca. Es el caso extremo de clase cerrada con un solo elemento.

- i es *recurrente* si el retorno es seguro, y *transitorio* si no. En un espacio finito, los estados de una clase cerrada son recurrentes y los de una clase no cerrada, transitorios.

Regla: Analizar una cadena de Markov concreta

1. **Dibuja el grafo de transiciones** —un nodo por estado, una flecha por cada $p_{ij} > 0$—. Casi todo se lee de ahí.
2. **Clasifica los estados:** identifica las clases de comunicación, cuáles son cerradas (recurrentes) y cuáles no (transitorias), y localiza los absorbentes.
3. **Comprueba las hipótesis antes de invocar nada.** Para el teorema ergódico hacen falta *irreducibilidad* —una sola clase— y *aperiodicidad* —basta con que algún $p_{ii} > 0$ para garantizarla—. Si hay periodo $d > 1$, la distribución estacionaria existe y es única, pero P^n no converge; sólo convergen los promedios temporales.
4. **Calcula la estacionaria** resolviendo el sistema $\pi P = \pi$ junto con $\sum_i \pi_i = 1$. En la práctica: escribe $\pi(P - I) = 0$, descarta una de las ecuaciones —siempre es redundante— y sustitúyela por la de normalización.
5. **Tiempos medios de retorno:** $\mu_i = 1/\pi_i$, directamente.
6. **Si hay estados absorbentes**, ordena la matriz por bloques $P = \left(\begin{smallmatrix} Q & R \\ 0 & I \end{smallmatrix}\right)$ con Q las transiciones entre transitorios. Entonces la *matriz fundamental* es $N = (I - Q)^{-1}$; el número medio de visitas a j partiendo de i es N_{ij}, el tiempo medio hasta la absorción es la suma de la fila i de N, y las probabilidades de absorción en cada estado son las filas de NR.
7. **Comprueba** que las filas de P suman 1, que π tiene componentes no negativas y que suman 1. Un error de aritmética se detecta casi siempre en este paso.

Ejemplo 22.51 (Una cadena con absorción, entera)**.** Un jugador con 1 o 2 euros apuesta uno a cara o cruz; se retira al arruinarse (0) o al llegar a 3. Los estados 0 y 3 son absorbentes; 1 y 2, transitorios. Con

$$Q = \begin{pmatrix} 0 & \frac{1}{2} \\ \frac{1}{2} & 0 \end{pmatrix}, \qquad R = \begin{pmatrix} \frac{1}{2} & 0 \\ 0 & \frac{1}{2} \end{pmatrix},$$

la matriz fundamental es

$$N = (I - Q)^{-1} = \begin{pmatrix} 1 & -\frac{1}{2} \\ -\frac{1}{2} & 1 \end{pmatrix}^{-1} = \frac{4}{3}\begin{pmatrix} 1 & \frac{1}{2} \\ \frac{1}{2} & 1 \end{pmatrix} = \begin{pmatrix} \frac{4}{3} & \frac{2}{3} \\ \frac{2}{3} & \frac{4}{3} \end{pmatrix}.$$

El tiempo medio hasta terminar partiendo de 1 euro es $\frac{4}{3} + \frac{2}{3} = 2$ apuestas, y las probabilidades de absorción son

$$NR = \begin{pmatrix} \frac{2}{3} & \frac{1}{3} \\ \frac{1}{3} & \frac{2}{3} \end{pmatrix}:$$

con 1 euro se arruina con probabilidad $\frac{2}{3}$ y llega a 3 con probabilidad $\frac{1}{3}$, en proporción exacta a la fortuna inicial, como debe ser porque la fortuna es una martingala.

22.13. Ejercicios resueltos

Tres problemas, uno por cada objeto central: cadena de Markov, proceso de Poisson y movimiento browniano.

Ejercicio 22.1. Una cadena de dos estados tiene matriz $P = \begin{pmatrix} 0{,}7 & 0{,}3 \\ 0{,}4 & 0{,}6 \end{pmatrix}$. Halle P^2, la distribución estacionaria y explique por qué existe y es única.

Solución. Multiplicando, $P^2 = \begin{pmatrix} 0{,}61 & 0{,}39 \\ 0{,}52 & 0{,}48 \end{pmatrix}$.

La estacionaria resuelve $\pi P = \pi$ con $\pi_1 + \pi_2 = 1$. De $0{,}7\pi_1 + 0{,}4\pi_2 = \pi_1$ sale $0{,}4\pi_2 = 0{,}3\pi_1$, luego $\pi = (4/7,\, 3/7) = (0{,}5714;\, 0{,}4286)$.

Existe y es única porque la cadena es irreducible —desde cualquier estado se llega a cualquier otro— y aperiódica, ya que hay probabilidad positiva de quedarse. El teorema ergódico garantiza entonces que P^n converge a la matriz cuyas filas son todas π, sea cual sea el estado inicial. Ya en P^2 se ve la tendencia: las dos filas se han acercado entre sí. ■

Ejercicio 22.2. Las llamadas a una centralita llegan según un proceso de Poisson con $\lambda = 3$ por minuto. Calcule la probabilidad de ninguna llamada en un minuto, la de al menos dos, y la distribución del tiempo hasta la primera.

Solución. $P(N = 0) = \mathrm{e}^{-3} = 0{,}0498$. Para al menos dos, $P(N \geq 2) = 1 - \mathrm{e}^{-3}(1+3) = 1 - 4\mathrm{e}^{-3} = 0{,}8009$. El tiempo hasta la primera llamada es exponencial de parámetro 3, con media $1/3$ de minuto, es decir, veinte segundos.

Lo notable es la falta de memoria: si han pasado dos minutos sin llamadas, el tiempo que falta hasta la siguiente sigue siendo exponencial de media veinte segundos. El proceso no "debe" nada. Esa propiedad caracteriza a la exponencial entre las distribuciones continuas, y es lo que hace que el proceso de Poisson sea el modelo natural de sucesos sin memoria. ■

Ejercicio 22.3. Sea $B(t)$ un movimiento browniano estándar. Calcule $\mathrm{Var}(B(3) - B(1))$, $P(B(4) > 2)$ y explique por qué B no es derivable en ningún punto.

Solución. Los incrementos son independientes y $B(t) - B(s) \sim N(0, t-s)$, luego $\mathrm{Var}(B(3) - B(1)) = 2$. Y $B(4) \sim N(0, 4)$, de modo que $P(B(4) > 2) = P(Z > 1) = 0{,}1587$.

En cuanto a la derivabilidad: el cociente incremental sobre un intervalo de longitud h tiene desviación típica $\sqrt{h}/h = 1/\sqrt{h}$, que tiende a infinito cuando $h \to 0$. El cociente no se estabiliza; se dispersa cada vez más. La demostración rigurosa —de Paley, Wiener y Zygmund, 1933— establece que, con probabilidad 1, la trayectoria no es derivable en *ningún* punto. Es continua en todas partes y derivable en ninguna, igual que la función de Weierstrass, sólo que aquí ese comportamiento patológico no es un ejemplo construido a propósito sino la norma: casi toda trayectoria es así. ■

22.14. Ejercicios propuestos

22.1. Para $P = \begin{pmatrix} 0{,}5 & 0{,}5 \\ 0{,}2 & 0{,}8 \end{pmatrix}$, halle la distribución estacionaria.

22.2. ¿Qué significa que una cadena sea irreducible? ¿Y aperiódica? Dé un ejemplo de cadena periódica.

22.3. ¿Qué es un estado absorbente? Dé la matriz de una cadena con uno.

22.4. En la ruina del jugador con $p = 1/2$, capital inicial 3 y objetivo 10, ¿cuál es la probabilidad de arruinarse?

22.5. Si $N(t)$ es Poisson con $\lambda = 5$ por hora, ¿cuál es la probabilidad de exactamente tres sucesos en media hora?

22.6. ¿Qué distribución tiene el tiempo entre sucesos consecutivos de un proceso de Poisson?

22.7. Explique la propiedad de falta de memoria y diga qué otra distribución la tiene en el caso discreto.

22.8. Para el movimiento browniano, calcule $\mathrm{Cov}(B(2), B(5))$.

22.9. ¿Qué es el movimiento browniano geométrico y por qué se usa en finanzas?

22.10. ¿Qué es una martingala? Diga si $B(t)$ y $B(t)^2 - t$ lo son.

22.11. Enuncie el teorema ergódico para cadenas de Markov y diga qué hipótesis necesita.

22.12. ¿Por qué la variación total de una trayectoria browniana es infinita y su variación cuadrática vale t?

Parte IV

Álgebra abstracta

Capítulo 23

Teoría de grupos desde cero: la simetría explicada con las manos

Lo que necesitas traer

Del capítulo 1: conjuntos, funciones biyectivas, relaciones de equivalencia e inducción. Del capítulo 3: divisibilidad y aritmética modular, que dan los primeros ejemplos de grupo. Del capítulo 7: matrices invertibles, que dan los ejemplos no conmutativos.

El capítulo 25 presentó los grupos como quien presenta a un adulto: de golpe, con su nombre completo y su currículum. Este capítulo hace lo contrario. Aquí empezamos con un triángulo de cartón y una mesa, y no escribimos ni un axioma hasta haber visto de dónde salen. Todo lo que se afirma se comprueba con números; todas las tablas están calculadas entera y explícitamente; ningún paso se salta con un "es evidente".

La promesa es concreta: al terminar este capítulo deberías poder tomar cualquier objeto —un cuadrado, un cubo de Rubik, un conjunto de raíces de un polinomio, una molécula— y responder por ti mismo qué grupo tiene, cuántos elementos, qué subgrupos, y qué significa cada una de esas respuestas. Y deberías poder leer el capítulo 25 completo sin sentir que te falta el suelo.

Idea clave

Intuición primero. Toda la teoría de grupos nace de una observación tonta: *las cosas que se pueden deshacer se combinan bien entre sí.* Si giro un cuadrado y luego lo giro otra vez, el resultado también es un giro. Si lo giro, puedo volver atrás. Si no hago nada, tampoco pasa nada. Un grupo es, literalmente, "la colección de todo lo que le puedo hacer a algo sin romperlo, y que puedo deshacer". Los números, las matrices, las permutaciones y las simetrías son todos ejemplos de esa misma idea, y por eso un solo teorema sobre grupos vale simultáneamente para el reloj, para el cubo de Rubik y para las ecuaciones de quinto grado.

23.1. El experimento del triángulo

Toma un triángulo equilátero de cartón. Marca sus vértices con 1, 2, 3 *por los dos lados*, y dibuja en la mesa una silueta del triángulo con las posiciones marcadas también $1, 2, 3$. La pregunta es:

¿de cuántas maneras puedo levantar el triángulo y volver a ponerlo en su silueta?

Cuéntalas con el cartón en la mano antes de seguir leyendo. Encontrarás seis:

(1) e: no hacer nada.

(2) r: girarlo $120°$ en sentido antihorario.

(3) r^2: girarlo $240°$ (o sea, aplicar r dos veces).

(4) s_1: voltearlo sobre el eje que pasa por la posición 1.

(5) s_2: voltearlo sobre el eje que pasa por la posición 2.

(6) s_3: voltearlo sobre el eje que pasa por la posición 3.

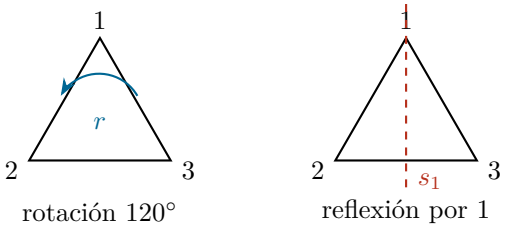


Figura 23.1: El experimento del triángulo: tres giros y tres volteos agotan todos los movimientos que devuelven el cartón a su sitio. Componer dos cualesquiera nunca produce nada nuevo, y ese cierre es exactamente lo que significa ser un grupo.

Seis movimientos, ni uno más. ¿Por qué exactamente seis? Porque el vértice que acaba en la posición 1 puede ser cualquiera de los tres (3 opciones), y una vez decidido eso, los otros dos pueden quedar en dos órdenes (2 opciones): $3 \times 2 = 6$. Este razonamiento —"cuento las opciones en cadena"— reaparecerá formalizado como el teorema órbita–estabilizador.

23.1.1. Componer movimientos

Lo interesante no es la lista, sino que los movimientos *se pueden encadenar*. Escribimos ab para "primero b, después a" (el mismo convenio que en la composición de funciones, $f \circ g$ significa aplicar g primero; ver la advertencia más abajo).

Ejemplo 23.1 (Calcular $r\,s_1$ paso a paso)**.** Partimos de la posición inicial, con el vértice i en la posición i.

Paso 1. Aplicamos s_1: el vértice 1 se queda, los vértices 2 y 3 intercambian sitios. Ahora las posiciones $(1, 2, 3)$ contienen a los vértices $(1, 3, 2)$.

Paso 2. Aplicamos r (rotación antihoraria): el contenido de la posición 1 pasa a la 2, el de la 2 a la 3, el de la 3 a la 1. El contenido era $(1,3,2)$, así que ahora es: posición 1 recibe lo que había en 3, es decir 2; posición 2 recibe lo que había en 1, es decir 1; posición 3 recibe lo que había en 2, es decir 3. Resultado: $(2,1,3)$.

Paso 3. Identificamos el resultado. Las posiciones contienen $(2,1,3)$: el vértice 3 se quedó quieto y los otros dos se intercambiaron. Eso es exactamente s_3.

Conclusión. $r\,s_1 = s_3$.

Ejemplo 23.2 (Y ahora al revés: $s_1\,r$)**. Paso 1.** Aplicamos r: las posiciones pasan a contener $(3,1,2)$.

Paso 2. Aplicamos s_1: intercambia lo que hay en las posiciones 2 y 3. De $(3,1,2)$ pasamos a $(3,2,1)$.

Paso 3. En $(3,2,1)$ el vértice 2 se quedó y los otros se intercambiaron: es s_2.

Conclusión. $s_1\,r = s_2 \neq s_3 = r\,s_1$.

Cuidado

El orden importa, y el convenio también. Acabamos de ver $rs_1 \neq s_1 r$: este grupo *no es conmutativo*. Es la primera vez en todo el libro que una operación falla al conmutar de forma esencial, y no es un accidente: si todo conmutara, la teoría de Galois no existiría y las ecuaciones de quinto grado tendrían fórmula.

Además, distintos libros usan distintos convenios: unos leen ab como "primero a", otros como "primero b". Ninguno está mal; lo que está mal es mezclarlos. En este libro **siempre** ab significa *primero b, después a*, igual que $f \circ g$. Si comparas una tabla tuya con la de otro texto y sale "transpuesta", casi seguro es esto.

23.1.2. La tabla completa

Repitiendo el cálculo anterior para las 36 parejas se obtiene la *tabla de Cayley* de este grupo, llamado D_3 (o S_3: son el mismo). La casilla en la fila a y la columna b contiene ab:

$\cdot$	e	r	r^2	s_1	s_2	s_3
e	e	r	r^2	s_1	s_2	s_3
r	r	r^2	e	s_3	s_1	s_2
r^2	r^2	e	r	s_2	s_3	s_1
s_1	s_1	s_2	s_3	e	r	r^2
s_2	s_2	s_3	s_1	r^2	e	r
s_3	s_3	s_1	s_2	r	r^2	e

Mira la tabla un minuto largo antes de continuar. Hay tres cosas que saltar a la vista, y las tres son teoremas disfrazados:

(a) **En cada fila y en cada columna aparece cada elemento exactamente una vez.** Es un cuadrado latino. Razón: si $ab = ab'$ entonces multiplicando por a^{-1} sale $b = b'$. La *cancelación* es visible a simple vista.

(b) **El bloque de arriba a la izquierda (e, r, r^2) es cerrado**: rotación por rotación da rotación. Ese bloque es un subgrupo, y encima ocupa exactamente la mitad de la tabla. Eso es el teorema de Lagrange asomándose.

(c) **Reflexión por reflexión da rotación.** Voltear dos veces devuelve el triángulo a su cara original, y lo que queda es un giro. Por eso "par × par = par": el signo del capítulo 25 vive aquí.

Pausa para pensar

Comprueba tú mismo, con el cartón o con la tabla, que $s_1 s_2 = r$ pero $s_2 s_1 = r^2$. Después contesta: ¿cuánto vale $(s_1 s_2)^{-1}$? Si dijiste r^2, muy bien; observa que también es $s_2^{-1} s_1^{-1} = s_2 s_1 = r^2$. Acabas de verificar a mano la regla $(ab)^{-1} = b^{-1} a^{-1}$: para deshacer "ponerse calcetines y luego zapatos" hay que quitarse primero los zapatos.

23.2. Cimientos: qué es exactamente una operación, qué es exactamente un grupo, y qué se deduce de ello sin añadir nada

La sección anterior contó seis movimientos de un triángulo y comprobó, con el cartón en la mano, que encadenarlos nunca produce nada nuevo. Esa experiencia es el suelo de todo el capítulo, pero una experiencia no es una definición: para poder demostrar algo hay que escribir con precisión qué se supone y qué no. Esta sección hace ese trabajo de cimentación. Aquí se define, una por una y con ejemplo resuelto, cada pieza que el resto del capítulo va a usar: qué es una operación, qué es un grupo, qué es el orden de un elemento, qué es un subgrupo, cómo se escriben y se multiplican las permutaciones, qué es un homomorfismo y qué dice el teorema de Lagrange.

Se puede leer de dos maneras. Si vienes de cero, léela entera y despacio: es el vocabulario completo, y no hace falta nada más que saber qué es un conjunto y qué es una aplicación. Si ya has visto grupos alguna vez, úsala como diccionario y salta directamente a la sección siguiente, que vuelve sobre los axiomas comentando los errores típicos al verificarlos.

23.2.1. La pieza más pequeña: la operación

Antes que el grupo está la operación. Todo lo que sigue —los axiomas, las tablas, los teoremas— es palabrería si no se ha fijado primero qué significa "combinar dos elementos". Y el detalle que casi todos los principiantes pasan por alto es que una operación tiene dos exigencias, no una: tiene que estar definida para *todas* las parejas, y el resultado tiene que caer *dentro* del conjunto.

Definición 23.3 (Operación binaria interna)**.** Sea G un conjunto no vacío. Una **operación binaria interna** en G es una aplicación

$$* : \ G \times G \longrightarrow G, \qquad (a, b) \longmapsto a * b.$$

Que sea una aplicación con dominio $G \times G$ y codominio G significa exactamente dos cosas:

(O1) Totalidad. Para *cada* pareja (a, b) de elementos de G el resultado $a * b$ está definido, sin excepciones.

(O2) Cierre. Para cada pareja, $a * b$ es de nuevo un elemento de G.

Se dice entonces que el par $(G, *)$ es un *conjunto con operación*.

Ejemplo 23.4 (Cuatro operaciones legítimas y dos impostoras)**.** Se comprueba (O1) y (O2) en cada caso; nada más.

(a) La suma y (b) el producto en $\mathbb{Z}$. Dados dos enteros cualesquiera, $a+b$ y $a \cdot b$ están siempre definidos y son enteros. Operaciones binarias internas. ✓ **(c) La resta en $\mathbb{Z}$.** $a - b$ existe siempre y es entero: *sí* es operación binaria interna. ✓ (Que luego no sirva para formar un grupo es otro asunto: fallará la asociatividad, no el cierre.)

(d) La composición de biyecciones de un conjunto X. Si f y g son biyecciones de X en X, la composición $f \circ g$ está definida y vuelve a ser una biyección. ✓ **(e) La resta en $\mathbb{N}$. NO.** Falla el cierre: $3 - 5 = -2$, y $-2 \notin \mathbb{N}$. La "operación" se sale del conjunto, así que ni siquiera es una aplicación $\mathbb{N} \times \mathbb{N} \to \mathbb{N}$. Un solo contraejemplo basta para tumbarla.

(f) La división en $\mathbb{Z}$. NO, y por dos motivos independientes. Falla el cierre: $3 : 2 = 1{,}5 \notin \mathbb{Z}$. Y falla la totalidad: $3 : 0$ no está definido, luego la pareja $(3, 0)$ no tiene imagen. Cualquiera de los dos fallos la descarta.

Cuidado

Cuando en un ejercicio te pidan comprobar que algo es un grupo, la primera comprobación no es ningún axioma sofisticado: es que la operación no te expulse del conjunto. Los ejemplos (e) y (f) son el tipo de fallo que hace que todo lo demás sobre. Y ojo con los conjuntos definidos por una propiedad: los números impares con la suma no son cerrados ($3 + 5 = 8$), y las matrices 2×2 con determinante 2 tampoco lo son con el producto (el producto de dos de ellas tiene determinante 4).

Una operación sobre un conjunto finito se puede escribir entera, y esa escritura es la herramienta de cálculo más útil que hay en teoría de grupos elemental.

Definición 23.5 (Tabla de Cayley)**.** Si $G = \{g_1, \ldots, g_n\}$ es finito, su **tabla de Cayley** es la tabla de n filas y n columnas cuya casilla situada en la fila g_i y la columna g_j contiene el elemento $g_i * g_j$. El orden de los factores es el que dicta este libro: *fila primero, columna después*.

Ejemplo 23.6 (La tabla de $\{1, 2, 3, 4\}$ con el producto módulo 5)**.** Se multiplica y se toma resto al dividir entre 5. Por ejemplo $2 \cdot 3 = 6 = 5 + 1$, luego la casilla es 1; y $4 \cdot 4 = 16 = 3 \cdot 5 + 1$, luego es 1.

$\cdot$	1	2	3	4
1	1	2	3	4
2	2	4	1	3
3	3	1	4	2
4	4	3	2	1

Todos los resultados están en el conjunto: la operación es interna. ✓ La tabla es simétrica respecto de la diagonal principal, señal de que la operación conmuta. Y en cada fila y en cada columna aparece cada elemento exactamente una vez; volveremos sobre eso al demostrar la propiedad cancelativa.

Con la operación fijada, se le pueden pedir propiedades. Estas cuatro son el vocabulario mínimo.

Definición 23.7 (Asociativa, conmutativa, neutro, simétrico)**.** Sea $*$ una operación binaria interna en G.

- $*$ es **asociativa** si $(a*b)*c = a*(b*c)$ para todos $a, b, c \in G$.
- $*$ es **conmutativa** si $a*b = b*a$ para todos $a, b \in G$.
- Un elemento $e \in G$ es **neutro** si $e*a = a*e = a$ para todo $a \in G$.
- Si hay un neutro e, un elemento b es **simétrico** (o **inverso**) de a cuando $a*b = b*a = e$.

Ejemplo 23.8 (La resta no es asociativa, la composición sí)**. Resta en $\mathbb{Z}$.** $(8-3)-2 = 5-2 = 3$, mientras que $8-(3-2) = 8-1 = 7$. Como $3 \neq 7$, la resta no es asociativa. Tampoco es conmutativa ($8-3 \neq 3-8$) y no tiene neutro por los dos lados: $a-0 = a$, pero $0-a = -a \neq a$ salvo si $a = 0$. **Composición de aplicaciones.** Siempre es asociativa, y la razón es que no hay nada que demostrar más allá de leer las definiciones: para cada x,

$$((f\circ g)\circ h)(x) = (f\circ g)(h(x)) = f\Big(g(h(x))\Big) = f\big((g\circ h)(x)\big) = \big(f\circ(g\circ h)\big)(x).$$

Como coinciden en todo punto, las dos aplicaciones son iguales. Este cálculo de dos líneas es la razón de fondo de que *casi todos* los grupos que aparecen en la práctica sean asociativos: casi todos son grupos de transformaciones disfrazados.

23.2.2. La definición de grupo, axioma por axioma

Ahora sí. Un grupo es un conjunto con una operación que cumple exactamente cuatro exigencias, ni una más. Conviene ver de dónde sale cada una antes de escribirlas: las cuatro están ya en el experimento del triángulo. *Cierre*: encadenar dos movimientos da un movimiento de la lista. *Asociatividad*: encadenar tres movimientos no depende de cómo se agrupen, porque son composiciones de aplicaciones. *Neutro*: existe el movimiento "no hacer nada". *Inverso*: todo movimiento se puede deshacer. Un grupo es la abstracción exacta de esas cuatro observaciones.

Definición 23.9 (Grupo)**.** Un **grupo** es un par $(G, *)$ formado por un conjunto no vacío G y una operación binaria interna $*$ en G que cumple:

G1. Cierre. $a * b \in G$ para todos $a, b \in G$. (Va incluido ya en la definición 23.3 de operación interna, y se enuncia aparte porque en la práctica es lo primero que falla.)

G2. Asociatividad. $(a * b) * c = a * (b * c)$ para todos $a, b, c \in G$.

G3. Elemento neutro. Existe $e \in G$ tal que $e * a = a * e = a$ para todo $a \in G$.

G4. Elemento inverso. Para cada $a \in G$ existe $a' \in G$ tal que $a * a' = a' * a = e$.

Cuando no haya ambigüedad se escribe G en lugar de $(G, *)$, y ab en lugar de $a * b$.

Definición 23.10 (Grupo abeliano)**.** Un grupo $(G, *)$ es **abeliano** (o **conmutativo**) si además cumple

$$\textbf{G5.} \quad a * b = b * a \quad \text{para todos } a, b \in G.$$

El nombre honra a Niels Henrik Abel. En la tabla de Cayley la conmutatividad se ve de un vistazo: la tabla es simétrica respecto de la diagonal principal.

Observación 23.11 (Notación multiplicativa y notación aditiva)*.* El mismo concepto se escribe de dos maneras según el grupo, y confundirlas es fuente constante de errores de lectura. En un grupo cualquiera se usa la *notación multiplicativa*; en los grupos abelianos de números se usa la *aditiva*. Es la misma teoría con otra ropa:

Concepto	**Multiplicativa**	**Aditiva**
operación	ab	$a + b$
neutro	e (o 1)	0
inverso de a	a^{-1}	$-a$
repetir n veces	a^n	na
subgrupo generado	$\langle a \rangle = \{a^n\}$	$\langle a \rangle = \{na\}$

Regla práctica: si ves un $+$, traduce mentalmente "0 hace de e y $-a$ hace de a^{-1}", y todos los teoremas siguen valiendo palabra por palabra.

Ejemplo 23.12 ($(\mathbb{Z}, +)$ verificado axioma por axioma)**.** No se da nada por supuesto; se comprueban los cuatro, uno a uno. **G1.** La suma de dos enteros es un entero. ✓ **G2.** $(a + b) + c = a + (b + c)$, propiedad conocida de la suma de enteros. ✓ **G3.** El neutro es 0, porque $0 + a = a + 0 = a$. ✓ **G4.** El inverso de a es $-a$, que también es entero, y $a + (-a) = (-a) + a = 0$. ✓ **G5.** $a + b = b + a$. ✓ Conclusión: $(\mathbb{Z}, +)$ es un grupo abeliano infinito. Obsérvese el cambio de ropa: aquí "a^{-1}" se escribe $-a$ y "e" se escribe 0.

23.2.3. Las primeras demostraciones: lo que sale gratis de los axiomas

Estas cuatro proposiciones son, probablemente, las primeras demostraciones algebraicas que lees en tu vida. Son cortas y merecen leerse con lápiz: cada igualdad se justifica con *un*

axioma, y no se usa absolutamente nada más. En particular, no se supone que el grupo sea conmutativo ni finito.

Proposición 23.13 (Unicidad del neutro y del inverso)**.** *En todo grupo G:*

(a) *hay un único elemento neutro;*

(b) *cada $a \in G$ tiene un único inverso.*

Demostración. **(a)** Supongamos que e y e' son ambos neutros. Como e es neutro, $e * e' = e'$. Como e' es neutro, $e * e' = e$. Los dos miembros izquierdos son el mismo elemento, luego

$$e' = e * e' = e.$$

Es decir, $e = e'$: no había dos, había uno escrito de dos maneras. **(b)** Supongamos que b y c son ambos inversos de a, es decir $a * b = b * a = e$ y $a * c = c * a = e$. Entonces

$$b \overset{\text{G3}}{=} b * e \overset{\text{G4}}{=} b * (a * c) \overset{\text{G2}}{=} (b * a) * c \overset{\text{G4}}{=} e * c \overset{\text{G3}}{=} c.$$

Cinco pasos, cinco axiomas. Luego $b = c$. □

Observación 23.14. La unicidad no es un adorno: es lo que permite *escribir* e y a^{-1}. Sin ella, la notación a^{-1} sería ambigua, porque no estaría claro a cuál de los posibles inversos se refiere. A partir de aquí, y para siempre, e denota el neutro y a^{-1} el inverso de a.

Proposición 23.15 (Propiedad cancelativa)**.** *En todo grupo G, para todos $a, b, c \in G$:*

$$ab = ac \implies b = c, \qquad ba = ca \implies b = c.$$

Demostración. Supongamos $ab = ac$. Multiplicamos ambos miembros por a^{-1} *por la izquierda* —y solo por la izquierda, que es lo delicado si el grupo no conmuta:

$$a^{-1}(ab) = a^{-1}(ac).$$

Aplicando G2 a los dos lados, $(a^{-1}a)b = (a^{-1}a)c$; por G4, $eb = ec$; por G3, $b = c$. La segunda implicación es idéntica multiplicando por la derecha. □

Idea clave

Por qué la tabla de Cayley es un cuadrado latino. La cancelación tiene una consecuencia visual inmediata: en la fila de a no puede repetirse ningún elemento, porque dos casillas iguales significarían $ab = ab'$ con $b \neq b'$, y eso contradice la proposición 23.15. Como la fila tiene tantas casillas como elementos, aparecen *todos* exactamente una vez. Lo mismo con las columnas. Por eso una tabla con un elemento repetido en una fila no es la tabla de ningún grupo, y esa es la comprobación más rápida que existe.

Proposición 23.16 (Inversos de inversos y de productos)**.** *En todo grupo G y para todos $a, b \in G$:*

(a) $(a^{-1})^{-1} = a$, ***(b)*** $(ab)^{-1} = b^{-1}a^{-1}$ *(¡en ese orden!).*

Demostración. **(a)** Por definición, $(a^{-1})^{-1}$ es el único elemento que multiplicado por a^{-1} da e. Pero a ya cumple $a^{-1}a = aa^{-1} = e$, luego a *es* un inverso de a^{-1}; por la unicidad (proposición 23.13) es *el* inverso. **(b)** Basta comprobar que $b^{-1}a^{-1}$ hace el trabajo, y de nuevo la unicidad se encarga del resto:

$$(ab)(b^{-1}a^{-1}) = a(bb^{-1})a^{-1} = a\,e\,a^{-1} = aa^{-1} = e,$$
$$(b^{-1}a^{-1})(ab) = b^{-1}(a^{-1}a)b = b^{-1}\,e\,b = b^{-1}b = e.$$

Los dos productos dan e, luego $b^{-1}a^{-1}$ es el inverso de ab. □

Cuidado

El orden invertido de (b) no es una manía de notación: es geometría. Para deshacer "ponerse los calcetines y después los zapatos" hay que quitarse *primero* los zapatos. Si escribieras $(ab)^{-1} = a^{-1}b^{-1}$ y multiplicaras, obtendrías $ab\,a^{-1}b^{-1}$, que no se simplifica: b y a^{-1} no conmutan en general. En un grupo abeliano las dos fórmulas coinciden, y precisamente por eso conviene aprenderla en el caso no abeliano.

Regla: comprobar si $(G, *)$ es un grupo

Cinco pasos, en este orden. Parar en cuanto uno falle.

1. **¿Hay operación?** Comprueba que $a * b$ está definido para toda pareja y que el resultado está en G. Si G es finito, escribe la tabla: el cierre se ve de un vistazo. Si G es infinito, busca un contraejemplo antes de intentar demostrar nada.
2. **¿Es asociativa?** Si la operación es composición de aplicaciones, o suma o producto de números o de matrices, lo es de oficio y basta decirlo. Si es una regla inventada, hay que comprobarla o encontrar un contraejemplo con tres elementos concretos.
3. **¿Hay neutro?** Búscalo. En la tabla es la fila que reproduce la cabecera y la columna que reproduce la primera columna. Debe funcionar *por los dos lados*.
4. **¿Hay inversos?** Para cada elemento busca otro que multiplicado por él dé e. En la tabla: comprueba que e aparece en cada fila.
5. **(Opcional) ¿Es abeliano?** Mira si la tabla es simétrica respecto de la diagonal, o si encuentras un par que no conmute.

Truco de descarte rápido: si en alguna fila de la tabla hay un elemento repetido, ya no es grupo (cuadrado latino), y no hace falta mirar nada más.

23.2.4. Un catálogo mínimo de grupos, todos verificados

Definición 23.17 (Los enteros módulo n)**.** Sea $n \geq 1$. En $\mathbb{Z}$ se declara $a \equiv b$ cuando n divide a $a - b$. Es una relación de equivalencia, y su clase $\bar{a} = a + n\mathbb{Z} = \{a + kn : k \in \mathbb{Z}\}$ agrupa a todos los enteros con el mismo resto al dividir entre n. El conjunto de las n clases se denota

$$\mathbb{Z}/n\mathbb{Z} = \{\bar{0}, \bar{1}, \ldots, \overline{n-1}\},$$

y se le da la operación $\bar{a}+\bar{b}=\overline{a+b}$. Esa operación *está bien definida*: si $\bar{a}=\overline{a'}$ y $\bar{b}=\overline{b'}$, entonces $n \mid a-a'$ y $n \mid b-b'$, luego n divide a $(a+b)-(a'+b')$ y por tanto $\overline{a+b}=\overline{a'+b'}$; el resultado no depende de los representantes elegidos. Con ella, $(\mathbb{Z}/n\mathbb{Z},+)$ es un grupo abeliano de n elementos: el neutro es $\bar{0}$ y el inverso de $\bar{a}$ es $\overline{n-a}$. En el resto del capítulo se abrevia $\mathbb{Z}_n$, y se escribe a en vez de $\bar{a}$ cuando no hay riesgo de confusión.

Ejemplo 23.18 (El grupo de Klein aparece dos veces con dos disfraces)**. Primer disfraz: $\mathbb{Z}/2\mathbb{Z}\times\mathbb{Z}/2\mathbb{Z}$.** Los cuatro pares $(0,0),(1,0),(0,1),(1,1)$ con la suma coordenada a coordenada. Llamando $e=(0,0)$, $a=(1,0)$, $b=(0,1)$, $c=(1,1)$, la tabla es

$\cdot$	e	a	b	c
e	e	a	b	c
a	a	e	c	b
b	b	c	e	a
c	c	b	a	e

Segundo disfraz: $\{1,3,5,7\}$ con el producto módulo 8. Se calcula: $3\cdot 3=9\equiv 1$, $5\cdot 5=25\equiv 1$, $7\cdot 7=49\equiv 1$, $3\cdot 5=15\equiv 7$, $3\cdot 7=21\equiv 5$, $5\cdot 7=35\equiv 3$. Escribiendo la tabla con $e=1$, $a=3$, $b=5$, $c=7$ se obtiene *exactamente* la de arriba. **Lectura.** Los dos conjuntos no tienen nada que ver —pares de bits contra restos módulo 8—, pero su tabla es la misma salvo renombrar. Eso es lo que más adelante llamaremos ser isomorfos, y es la primera prueba de que en teoría de grupos lo que importa no son los elementos sino cómo se combinan. Todo elemento cumple $x^2=e$, dato que basta para distinguir este grupo de $\mathbb{Z}/4\mathbb{Z}$.

Ejemplo 23.19 (S_3 y su tabla de Cayley completa)**.** Las seis simetrías del triángulo de la sección anterior se pueden mirar como las seis maneras de reordenar tres símbolos. El diccionario es exacto: el giro r manda el contenido de la posición 1 a la 2, el de la 2 a la 3 y el de la 3 a la 1, o sea $r=(1\,2\,3)$; y s_i es la reflexión que deja fijo el vértice i e intercambia los otros dos, luego $s_1=(2\,3)$, $s_2=(1\,3)$, $s_3=(1\,2)$.

Con el convenio del libro —fg significa *primero g, después f*— la tabla completa de S_3, con los seis elementos escritos como permutaciones, es

$\circ$	id	$(1\,2\,3)$	$(1\,3\,2)$	$(2\,3)$	$(1\,3)$	$(1\,2)$
id	id	$(1\,2\,3)$	$(1\,3\,2)$	$(2\,3)$	$(1\,3)$	$(1\,2)$
$(1\,2\,3)$	$(1\,2\,3)$	$(1\,3\,2)$	id	$(1\,2)$	$(2\,3)$	$(1\,3)$
$(1\,3\,2)$	$(1\,3\,2)$	id	$(1\,2\,3)$	$(1\,3)$	$(1\,2)$	$(2\,3)$
$(2\,3)$	$(2\,3)$	$(1\,3)$	$(1\,2)$	id	$(1\,2\,3)$	$(1\,3\,2)$
$(1\,3)$	$(1\,3)$	$(1\,2)$	$(2\,3)$	$(1\,3\,2)$	id	$(1\,2\,3)$
$(1\,2)$	$(1\,2)$	$(2\,3)$	$(1\,3)$	$(1\,2\,3)$	$(1\,3\,2)$	id

Traducida con el diccionario $r=(1\,2\,3)$, $r^2=(1\,3\,2)$, $s_1=(2\,3)$, $s_2=(1\,3)$, $s_3=(1\,2)$, esta tabla es *casilla por casilla* la misma que la del triángulo. Comprobación de una casilla, para creérselo: fila $(1\,2\,3)$, columna $(2\,3)$ significa aplicar primero $(2\,3)$ y después $(1\,2\,3)$; el 1 queda

fijo y luego va al 2, el 2 va al 3 y el 3 queda, el 3 va al 2 y luego al 3. Resultado: $1 \mapsto 2$, $2 \mapsto 1$, $3 \mapsto 3$, que es $(1\,2)$. ✓ En el lenguaje del triángulo: $r\, s_1 = s_3$.

Cada fila y cada columna contienen los seis elementos una sola vez (cuadrado latino ✓), y la tabla *no* es simétrica: S_3 no es abeliano. Es el grupo no abeliano más pequeño que existe.

Ejemplo 23.20 (D_4, las ocho simetrías del cuadrado)**.** Numera los vértices de un cuadrado $1, 2, 3, 4$ en orden cíclico. Un movimiento rígido que devuelva el cuadrado a su sitio queda determinado por cómo permuta los vértices, y hay ocho: cuatro *rotaciones* —e, $r = (1\,2\,3\,4)$ de $90°$, $r^2 = (1\,3)(2\,4)$ de $180°$ y $r^3 = (1\,4\,3\,2)$ de $270°$—, dos *reflexiones por ejes que pasan por vértices opuestos* —$(2\,4)$, que deja fijos 1 y 3, y $(1\,3)$, que deja fijos 2 y 4— y dos *reflexiones por ejes que pasan por los puntos medios de lados opuestos*, $(1\,2)(3\,4)$ y $(1\,4)(2\,3)$, que no dejan fijo ningún vértice.

Llamando $s = (2\,4)$, las ocho se escriben $e, r, r^2, r^3, s, rs, r^2s, r^3s$, y se comprueba que $rs = (1\,2)(3\,4)$, $r^2s = (1\,3)$ y $r^3s = (1\,4)(2\,3)$: las cuatro reflexiones, ni una repetida. **Órdenes** (cuántas veces hay que repetir cada movimiento para volver al punto de partida): $\mathrm{ord}(e) = 1$; $\mathrm{ord}(r) = \mathrm{ord}(r^3) = 4$; $\mathrm{ord}(r^2) = 2$; y las cuatro reflexiones tienen orden 2, porque reflejar dos veces por el mismo eje no hace nada. **No es abeliano**: sr y rs no coinciden. En efecto, la relación que gobierna todo el grupo es $srs = r^{-1}$, y la sección dedicada a D_4 más adelante en el capítulo la explota para calcular cualquier producto sin volver a tocar el cartón. Aquí basta con saber que D_4 existe, tiene ocho elementos y es un grupo.

Ejemplo 23.21 (Tres no-ejemplos, con el axioma culpable señalado)**. (a)** $(\mathbb{N}, +)$ **no es grupo.** Cierre ✓, asociatividad ✓, neutro 0 ✓. Falla **G4**: el inverso de 3 tendría que ser un x con $3 + x = 0$, es decir $x = -3$, que no está en $\mathbb{N}$. Ningún elemento salvo el 0 tiene inverso. **(b)** $(\mathbb{Z}, \cdot)$ **no es grupo.** Cierre ✓, asociatividad ✓, neutro 1 ✓. Falla **G4**: el inverso de 2 sería $1/2 \notin \mathbb{Z}$. Nótese lo fino de la distinción: el inverso *existe* en $\mathbb{Q}$, pero no en el conjunto que estamos considerando, y eso basta para arruinarlo. Los únicos enteros con inverso entero son 1 y -1. **(c)** $(\mathbb{Q}, \cdot)$ **tampoco.** Ahora casi todo tiene inverso, pero el 0 no: no hay ningún racional x con $0 \cdot x = 1$. Basta un elemento sin inverso para que falle G4. El remedio es quitarlo: $(\mathbb{Q}^*, \cdot)$ con $\mathbb{Q}^* = \mathbb{Q} \setminus \{0\}$ *sí* es un grupo abeliano —cierre ✓ (el producto de dos racionales no nulos no es nulo), asociatividad ✓, neutro 1 ✓, inverso de p/q es q/p ✓—. Es el ejemplo estándar de grupo multiplicativo infinito.

23.2.5. Orden de un grupo y orden de un elemento

Definición 23.22 (Orden de un grupo)**.** El **orden** de un grupo G es su número de elementos, y se escribe $|G|$. Si G tiene infinitos elementos se dice que es de orden infinito. Un grupo es *finito* cuando $|G| < \infty$.

Definición 23.23 (Potencias y orden de un elemento)**.** Sea G un grupo y $a \in G$. Se define $a^0 = e$, $a^{n+1} = a^n * a$ para $n \geq 0$ y $a^{-n} = (a^{-1})^n$. El **orden de** a, escrito $\mathrm{ord}(a)$, es el menor entero $k > 0$ tal que $a^k = e$; si no existe ningún k con esa propiedad, se dice que a tiene orden infinito.

En notación aditiva, a^k se escribe ka y la condición es $ka = 0$.

Proposición 23.24 (En un grupo finito todo elemento tiene orden finito)**.** *Si* $|G| < \infty$ *y* $a \in G$*, existe* $k > 0$ *con* $a^k = e$*.*

Demostración. La lista $a^1, a^2, a^3, \dots$ toma valores en un conjunto finito, así que dos de ellos coinciden: $a^i = a^j$ con $i < j$. Multiplicando por $(a^{-1})^i$ y usando la cancelación (proposición 23.15) queda $a^{j-i} = e$ con $j - i > 0$. □

Regla: calcular el orden de un elemento y el subgrupo que genera

1. Escribe la lista $a^1, a^2, a^3, \dots$ multiplicando cada vez por a, y **para en cuanto aparezca** e. El número de pasos dados es $\operatorname{ord}(a)$.
2. La lista que has escrito, $\{e, a, a^2, \dots, a^{k-1}\}$ con $k = \operatorname{ord}(a)$, *es* el subgrupo generado $\langle a \rangle$, y tiene exactamente k elementos.
3. En notación aditiva la lista es $a, 2a, 3a, \dots$ y paras al llegar a 0.
4. **Atajo en** $\mathbb{Z}/n\mathbb{Z}$: $\operatorname{ord}(a) = \dfrac{n}{\gcd(a, n)}$, porque $ka \equiv 0 \pmod{n}$ equivale a que n divida a ka.
5. **Atajo en** S_n: el orden de una permutación es el mínimo común múltiplo de las longitudes de sus ciclos disjuntos.
6. **Comprobación de sentido**: por el corolario de Lagrange, $\operatorname{ord}(a)$ debe dividir a $|G|$. Si te sale un número que no divide, hay un error de cálculo.

Ejemplo 23.25 (Órdenes calculados en tres grupos)**. (a) En** $\mathbb{Z}/12\mathbb{Z}$, con el atajo $\operatorname{ord}(a) = 12/\gcd(a, 12)$:

a	0	1	2	3	4	5	6	7	8	9	10	11
$\operatorname{ord}(a)$	1	12	6	4	3	12	2	12	3	4	6	12

Verificación a mano de uno de ellos, sin fórmula: ord(8) se calcula sumando 8 repetidamente módulo 12: 8, $16 \equiv 4$, $24 \equiv 0$. Tres pasos, luego $\operatorname{ord}(8) = 3$. ✓ Todos los órdenes de la tabla dividen a 12, como debe ser. **(b) En** $\{1, 2, 3, 4\}$ **con el producto módulo 5.** $\operatorname{ord}(1) = 1$. Para el 2: 2, 4, $8 \equiv 3$, $16 \equiv 1$; cuatro pasos, $\operatorname{ord}(2) = 4$. Para el 3: 3, $9 \equiv 4$, $27 \equiv 2$, $81 \equiv 1$; $\operatorname{ord}(3) = 4$. Para el 4: 4, $16 \equiv 1$; $\operatorname{ord}(4) = 2$. Órdenes $1, 4, 4, 2$, todos divisores de 4. ✓ **(c) En** S_3**.** $\operatorname{ord}(\mathrm{id}) = 1$; los dos 3-ciclos $(1\,2\,3)$ y $(1\,3\,2)$ tienen orden 3; las tres trasposiciones tienen orden 2. Órdenes $1, 3, 3, 2, 2, 2$: todos dividen a 6. ✓

23.2.6. Subgrupos, subgrupo generado y grupos cíclicos

Definición 23.26 (Subgrupo)**.** Sea $(G, *)$ un grupo. Un subconjunto $H \subseteq G$ es un **subgrupo** de G, y se escribe $H \leq G$, si H con la operación heredada de G es a su vez un grupo: es decir, si $e \in H$, si $ab \in H$ para todos $a, b \in H$, y si $a^{-1} \in H$ para todo $a \in H$.

Todo grupo tiene al menos dos subgrupos, $\{e\}$ y G, llamados *triviales*; los demás se llaman *propios*.

Proposición 23.27 (Criterio práctico de subgrupo)**.** *Un subconjunto $H \subseteq G$ es subgrupo si y solo si*

(a) *$H \neq \varnothing$, y*

(b) *$ab^{-1} \in H$ para todos $a, b \in H$.*

Demostración. Si H es subgrupo, (a) se cumple porque $e \in H$, y (b) porque $b^{-1} \in H$ y el producto no sale de H.

Recíprocamente, supongamos (a) y (b). Como $H \neq \varnothing$, elegimos $a \in H$; tomando $b = a$ en (b) obtenemos $aa^{-1} = e \in H$. Ahora, para cualquier $b \in H$, tomando el par (e, b) en (b) resulta $eb^{-1} = b^{-1} \in H$: los inversos están. Por último, dados $a, b \in H$ sabemos ya que $b^{-1} \in H$, y aplicando (b) al par (a, b^{-1}) sale $a(b^{-1})^{-1} = ab \in H$: hay cierre. La asociatividad se hereda de G sin más, porque una igualdad que vale para todos los elementos de G vale en particular para los de H. Luego H es un grupo con la operación heredada. □

Observación 23.28 (Atajo para grupos finitos)*.* Si G es finito, basta comprobar que $H \neq \varnothing$ y que H es cerrado para el producto: los inversos vienen gratis. En efecto, si $a \in H$, por la proposición 23.24 existe k con $a^k = e$; el cierre da $a^{k-1} \in H$, y $a^{k-1} = a^{-1}$. Este atajo *no* vale en grupos infinitos: $\mathbb{N}$ es un subconjunto no vacío y cerrado de $(\mathbb{Z}, +)$ y no es subgrupo.

Ejemplo 23.29 (El criterio aplicado dos veces)**. (a) Los pares forman subgrupo de $(\mathbb{Z}, +)$.** Sea $H = 2\mathbb{Z}$. No es vacío ($0 \in H$). Si $a = 2m$ y $b = 2k$, entonces $a - b = 2(m - k) \in H$ —recuérdese que en notación aditiva "ab^{-1}" se escribe $a - b$—. Por el criterio, $2\mathbb{Z} \leq \mathbb{Z}$. ✓ **(b) Los impares no.** Sea $H = \{\dots, -1, 1, 3, 5, \dots\}$. Falla ya el primer requisito de la definición: $0 \notin H$. Y también el cierre: $1 + 3 = 4$ no es impar. No es subgrupo.

Definición 23.30 (Subgrupo generado)**.** Dado $a \in G$, el **subgrupo generado por** a es

$$\langle a \rangle = \{a^k : k \in \mathbb{Z}\} = \{\dots, a^{-2}, a^{-1}, e, a, a^2, \dots\}.$$

Es un subgrupo: no es vacío y $a^i(a^j)^{-1} = a^{i-j}$ vuelve a ser una potencia. Si $\operatorname{ord}(a) = k$ es finito, la lista se cierra sobre sí misma y $\langle a \rangle = \{e, a, \dots, a^{k-1}\}$ tiene exactamente k elementos; es decir, $|\langle a \rangle| = \operatorname{ord}(a)$. Más en general, dado un subconjunto $S \subseteq G$, $\langle S \rangle$ denota el menor subgrupo que contiene a S, que es el conjunto de todos los productos finitos de elementos de S y de sus inversos.

Definición 23.31 (Grupo cíclico y generador)**.** Un grupo G es **cíclico** si existe $a \in G$ con $G = \langle a \rangle$. Ese elemento a se llama **generador** de G. Un grupo cíclico puede tener varios generadores distintos, y todo grupo cíclico es abeliano, porque $a^i a^j = a^{i+j} = a^j a^i$.

Ejemplo 23.32 (Todos los subgrupos de $\mathbb{Z}/12\mathbb{Z}$, calculados enteros)**.** Se hace por fuerza bruta y sin trucos: se calcula $\langle a \rangle$ para los doce elementos y se ve cuántos conjuntos distintos

salen. **Paso 1. Generar.** Sumando repetidamente y parando al llegar a 0:

$$\begin{array}{llll} \langle 0\rangle = \{0\}, & \langle 1\rangle = \mathbb{Z}/12\mathbb{Z}, & \langle 2\rangle = \{0,2,4,6,8,10\}, & \langle 3\rangle = \{0,3,6,9\}, \\ \langle 4\rangle = \{0,4,8\}, & \langle 5\rangle = \mathbb{Z}/12\mathbb{Z}, & \langle 6\rangle = \{0,6\}, & \langle 7\rangle = \mathbb{Z}/12\mathbb{Z}, \\ \langle 8\rangle = \{0,8,4\}, & \langle 9\rangle = \{0,9,6,3\}, & \langle 10\rangle = \{0,10,8,6,4,2\}, & \langle 11\rangle = \mathbb{Z}/12\mathbb{Z}. \end{array}$$

Por ejemplo $\langle 5\rangle$ se obtiene sumando cincos: $5, 10, 3, 8, 1, 6, 11, 4, 9, 2, 7, 0$, los doce elementos. Y se ve que $\langle 8\rangle = \langle 4\rangle$, $\langle 9\rangle = \langle 3\rangle$ y $\langle 10\rangle = \langle 2\rangle$. **Paso 2. Contar los distintos.** Salen seis conjuntos:

$$\{0\}, \quad \{0,6\}, \quad \{0,4,8\}, \quad \{0,3,6,9\}, \quad \{0,2,4,6,8,10\}, \quad \mathbb{Z}/12\mathbb{Z},$$

de órdenes $1, 2, 3, 4, 6, 12$: exactamente los divisores de 12, uno por divisor y sin repetir. **Paso 3. ¿Hay más subgrupos, aparte de estos seis?** No. Cualquier subgrupo H de $\mathbb{Z}/12\mathbb{Z}$ contiene el subgrupo generado por cada uno de sus elementos, y se comprueba que si H contiene a a contiene a $\langle a\rangle$; tomando en H el elemento no nulo d con menor valor, la división euclídea de cualquier otro elemento entre d obliga a que el resto esté en H y sea menor que d, luego nulo: $H = \langle d\rangle$. Así que todo subgrupo es de la forma $\langle d\rangle$, y ya los hemos listado todos. **Paso 4. Generadores.** Los cuatro elementos que generan el grupo entero son $1, 5, 7, 11$, es decir, exactamente los coprimos con 12. Son $\varphi(12) = 4$. **Paso 5. Cómo se contienen.** $\langle a\rangle \subseteq \langle b\rangle$ exactamente cuando $\gcd(b, 12)$ divide a $\gcd(a, 12)$; por ejemplo $\{0,6\} \subset \{0,3,6,9\}$ y $\{0,6\} \subset \{0,2,4,6,8,10\}$, mientras que $\{0,4,8\}$ y $\{0,3,6,9\}$ solo comparten el 0. El dibujo del retículo completo está en la figura 23.2, en el taller de ejemplos resueltos.

23.2.7. Permutaciones: dos notaciones y un convenio

Definición 23.33 (Grupo simétrico)**.** Una **permutación** de $\{1, 2, \ldots, n\}$ es una aplicación biyectiva $\sigma : \{1, \ldots, n\} \to \{1, \ldots, n\}$. El conjunto de todas ellas, con la composición de aplicaciones como operación, es el **grupo simétrico** S_n. Es un grupo: la composición de biyecciones es biyectiva (cierre), la composición es asociativa, el neutro es la identidad id y el inverso de σ es su aplicación inversa. Su orden es $|S_n| = n!$, porque una permutación se construye eligiendo la imagen del 1 (n opciones), la del 2 entre las restantes ($n-1$), y así sucesivamente. Para $n \geq 3$ no es abeliano.

Definición 23.34 (Notación de dos filas y notación de ciclos)**.** La **notación de dos filas** escribe arriba los elementos y debajo sus imágenes:

$$\sigma = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 3 & 5 & 4 & 1 & 6 & 2 \end{pmatrix} \quad \text{significa } \sigma(1) = 3,\ \sigma(2) = 5,\ \ldots,\ \sigma(6) = 2.$$

Un **ciclo** de longitud k, escrito $(a_1\, a_2\, \ldots\, a_k)$, es la permutación que manda $a_1 \mapsto a_2$, $a_2 \mapsto a_3$, $\ldots$, $a_k \mapsto a_1$ y deja fijo todo lo demás. Dos ciclos son **disjuntos** si no comparten ningún símbolo; los ciclos disjuntos conmutan entre sí, porque actúan sobre símbolos distintos. Un ciclo de longitud 2 se llama **trasposición**: intercambia dos símbolos y deja quieto el resto.

Cuidado

El convenio, que es la fuente número uno de errores. En este libro las permutaciones son aplicaciones y se componen como aplicaciones: en el producto $\sigma\tau$ actúa **primero** τ (la de la derecha) y **después** σ. Se lee, por tanto, **de derecha a izquierda**, igual que $f \circ g$. Muchos textos —y bastantes programas de ordenador, entre ellos algunas convenciones de sympy— usan el convenio contrario, de izquierda a derecha. Ninguno de los dos está mal; lo que está mal es mezclarlos. Si comparas tu resultado con el de otra fuente y obtienes la permutación inversa o la tabla traspuesta, casi con seguridad es esto. Antes de calcular nada, di en voz alta cuál estás usando.

Ejemplo 23.35 (De la notación de dos filas a los ciclos disjuntos, paso a paso)**.** Sea

$$\sigma = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 3 & 5 & 4 & 1 & 6 & 2 \end{pmatrix}.$$

Paso 1. Empezamos por el 1 y seguimos la cadena hasta volver: $1 \mapsto 3$, $3 \mapsto 4$, $4 \mapsto 1$. Cerramos el paréntesis: $(1\,3\,4)$. **Paso 2.** El menor símbolo no usado es el 2: $2 \mapsto 5$, $5 \mapsto 6$, $6 \mapsto 2$. Cerramos: $(2\,5\,6)$. **Paso 3.** No quedan símbolos, así que $\sigma = (1\,3\,4)(2\,5\,6)$, producto de dos ciclos disjuntos. **Paso 4. Orden.** Cada ciclo de longitud 3 vuelve a su sitio cada tres aplicaciones, y los dos vuelven a la vez en $\operatorname{mcm}(3,3) = 3$: $\operatorname{ord}(\sigma) = 3$. **Paso 5. Como producto de trasposiciones.** Un ciclo se rompe así: $(a\,b\,c) = (a\,c)(a\,b)$. Comprobación con el convenio de derecha a izquierda para $(1\,3\,4) = (1\,4)(1\,3)$: el 1 va a 3 y el 3 no lo toca $(1\,4)$, luego $1 \mapsto 3$ ✓; el 3 va a 1 y luego a 4, luego $3 \mapsto 4$ ✓; el 4 no lo toca $(1\,3)$ y luego va a 1, luego $4 \mapsto 1$ ✓. Así

$$\sigma = (1\,4)(1\,3)(2\,6)(2\,5),$$

cuatro trasposiciones.

Ejemplo 23.36 (Dos composiciones resueltas, y el orden importa)**.** Trabajamos en S_5 con $\sigma = (1\,3\,5)(2\,4)$ y $\tau = (1\,2\,4)$. **Primera composición:** $\sigma\tau$, es decir, primero τ y después σ. Se sigue cada símbolo, siempre en dos etapas: $1 \xrightarrow{\tau} 2 \xrightarrow{\sigma} 4$, luego $1 \mapsto 4$; $2 \xrightarrow{\tau} 4 \xrightarrow{\sigma} 2$, luego el 2 queda fijo; $3 \xrightarrow{\tau} 3 \xrightarrow{\sigma} 5$, luego $3 \mapsto 5$; $4 \xrightarrow{\tau} 1 \xrightarrow{\sigma} 3$, luego $4 \mapsto 3$; y $5 \xrightarrow{\tau} 5 \xrightarrow{\sigma} 1$, luego $5 \mapsto 1$. Armando los ciclos desde el 1: $1 \to 4 \to 3 \to 5 \to 1$, y el 2 solo. Resultado:

$$\sigma\tau = (1\,4\,3\,5), \qquad \operatorname{ord}(\sigma\tau) = 4.$$

Segunda composición: $\tau\sigma$, primero σ y después τ: $1 \xrightarrow{\sigma} 3 \xrightarrow{\tau} 3$, luego $1 \mapsto 3$; $2 \xrightarrow{\sigma} 4 \xrightarrow{\tau} 1$, luego $2 \mapsto 1$; $3 \xrightarrow{\sigma} 5 \xrightarrow{\tau} 5$, luego $3 \mapsto 5$; $4 \xrightarrow{\sigma} 2 \xrightarrow{\tau} 4$, luego el 4 queda fijo; y $5 \xrightarrow{\sigma} 1 \xrightarrow{\tau} 2$, luego $5 \mapsto 2$. Desde el 1: $1 \to 3 \to 5 \to 2 \to 1$. Resultado:

$$\tau\sigma = (1\,3\,5\,2) \;\neq\; (1\,4\,3\,5) = \sigma\tau.$$

Los dos son 4-ciclos, pero no el mismo: S_5 no es abeliano, y la única diferencia entre los dos cálculos ha sido el orden de aplicación. Si al hacerlo tú te sale $\sigma\tau = (1\,3\,5\,2)$, no te has equivocado de cuentas: has usado el convenio contrario.

Teorema 23.37 (Descomposición en trasposiciones y paridad)**.** *Toda permutación de S_n con $n \geq 2$ se puede escribir como producto de trasposiciones. La descomposición no es única, pero* la paridad del número de trasposiciones sí lo es*: si una permutación se escribe con un número par de trasposiciones, cualquier otra descomposición suya tendrá también un número par.*

Esquema de la demostración. La existencia es el paso 5 del ejemplo anterior aplicado a cada ciclo: $(a_1\, a_2\, \ldots\, a_k) = (a_1\, a_k)(a_1\, a_{k-1}) \cdots (a_1\, a_2)$, y toda permutación es producto de ciclos disjuntos. Para la unicidad de la paridad se considera el polinomio $\Delta = \prod_{i<j}(x_i - x_j)$ y se hace actuar una permutación sobre él permutando los índices; cada trasposición cambia el signo de Δ exactamente una vez, luego un producto de m trasposiciones lo multiplica por $(-1)^m$. Como el efecto sobre Δ depende solo de la permutación y no de cómo se haya escrito, $(-1)^m$ está determinado, y por tanto la paridad de m también. □

Definición 23.38 (Signo de una permutación y grupo alternado)**.** El **signo** de $\sigma \in S_n$ es $\operatorname{sgn}(\sigma) = (-1)^m$, donde m es el número de trasposiciones de cualquier descomposición de σ; el teorema 23.37 garantiza que no depende de cuál. Una permutación es **par** si su signo es $+1$ e **impar** si es -1. Como un ciclo de longitud k se escribe con $k-1$ trasposiciones, se tiene sgn de un k-ciclo igual a $(-1)^{k-1}$.

El **grupo alternado** A_n es el conjunto de las permutaciones pares de S_n. Es un subgrupo, porque el producto de dos pares es par y el inverso de una par es par, y para $n \geq 2$ tiene $n!/2$ elementos.

Ejemplo 23.39 (Signos calculados)**. (a)** $(1\,3\,4)(2\,5\,6)$ se escribió arriba con cuatro trasposiciones, luego su signo es $(-1)^4 = +1$: es *par*, y está en A_6. Sale también del atajo: dos ciclos de longitud 3, cada uno de signo $(-1)^2 = +1$, y el signo es multiplicativo. **(b) Coherencia con las composiciones de antes.** $\sigma = (1\,3\,5)(2\,4)$ tiene signo $(+1)\cdot(-1) = -1$ (impar) y $\tau = (1\,2\,4)$ tiene signo $+1$ (par); su producto debe ser impar, y en efecto $\sigma\tau = (1\,4\,3\,5)$ es un 4-ciclo, de signo $(-1)^3 = -1$. ✓ **(c) En** S_3: la identidad y los dos 3-ciclos son pares; las tres trasposiciones son impares. Luego $A_3 = \{\mathrm{id}, (1\,2\,3), (1\,3\,2)\}$ tiene $3 = 6/2$ elementos. ✓ **Cuidado**, el signo no es el orden: $(1\,2)(3\,4)$ tiene orden 2 y es *par*; $(1\,2)$ tiene orden 2 y es *impar*.

Definición 23.40 (Grupo diédrico)**.** Para $n \geq 3$, el **grupo diédrico** D_n es el grupo de las simetrías de un polígono regular de n lados: las n rotaciones de ángulo múltiplo de $360°/n$ y las n reflexiones respecto de sus ejes de simetría. Tiene $|D_n| = 2n$ elementos, y llamando r a la rotación mínima y s a una reflexión cualquiera se escribe

$$D_n = \{e, r, r^2, \ldots, r^{n-1},\ s, rs, r^2s, \ldots, r^{n-1}s\}, \qquad r^n = e, \quad s^2 = e, \quad srs = r^{-1}.$$

Los casos D_3 (el triángulo de la sección anterior, que coincide con S_3) y D_4 (el cuadrado) son los que se usan como banco de pruebas en todo el capítulo.

23.2.8. Homomorfismos, núcleo e imagen

Si un grupo es "un conjunto con una manera de combinar", las aplicaciones interesantes entre grupos son las que *respetan* esa manera de combinar. Da igual si primero combinas y

luego traduces o si primero traduces y luego combinas: eso es todo lo que dice la definición siguiente.

Definición 23.41 (Homomorfismo e isomorfismo)**.** Sean $(G, *)$ y $(H, \cdot)$ grupos. Una aplicación $\varphi : G \to H$ es un **homomorfismo** si

$$\varphi(a * b) = \varphi(a) \cdot \varphi(b) \qquad \text{para todos } a, b \in G.$$

Un homomorfismo biyectivo se llama **isomorfismo**, y entonces se dice que G y H son **isomorfos**, escrito $G \cong H$. Dos grupos isomorfos son el mismo grupo con los elementos renombrados: tienen el mismo orden, los mismos órdenes de elementos, los mismos subgrupos y la misma tabla salvo renombrar.

Definición 23.42 (Núcleo e imagen)**.** Sea $\varphi : G \to H$ un homomorfismo. Su **núcleo** y su **imagen** son

$$\ker \varphi = \{g \in G : \varphi(g) = e_H\}, \qquad \operatorname{im} \varphi = \{\varphi(g) : g \in G\} \subseteq H.$$

El núcleo mide qué información destruye φ; la imagen, qué parte de H alcanza.

Proposición 23.43 (Propiedades básicas)**.** *Si $\varphi : G \to H$ es homomorfismo, entonces $\varphi(e_G) = e_H$, $\varphi(a^{-1}) = \varphi(a)^{-1}$, $\operatorname{im} \varphi \leq H$ y $\ker \varphi \leq G$.*

Demostración. **Neutro.** $\varphi(e_G) = \varphi(e_G e_G) = \varphi(e_G)\varphi(e_G)$; cancelando $\varphi(e_G)$ en ambos miembros (proposición 23.15, aplicada en H) queda $e_H = \varphi(e_G)$. **Inversos.** $\varphi(a)\varphi(a^{-1}) = \varphi(aa^{-1}) = \varphi(e_G) = e_H$, y lo mismo por el otro lado; luego $\varphi(a^{-1})$ es el inverso de $\varphi(a)$. **La imagen es subgrupo**: no es vacía porque contiene a e_H, y dados $\varphi(a), \varphi(b)$ en ella, $\varphi(a)\varphi(b)^{-1} = \varphi(a)\varphi(b^{-1}) = \varphi(ab^{-1})$ vuelve a estar en ella; criterio 23.27. ✓ **El núcleo es subgrupo.** No es vacío porque $e_G \in \ker \varphi$. Y si $a, b \in \ker \varphi$,

$$\varphi(ab^{-1}) = \varphi(a)\varphi(b)^{-1} = e_H e_H^{-1} = e_H,$$

luego $ab^{-1} \in \ker \varphi$. Criterio 23.27. ✓ □

Proposición 23.44 (Inyectividad y núcleo trivial)**.** *Un homomorfismo $\varphi : G \to H$ es inyectivo si y solo si $\ker \varphi = \{e_G\}$.*

Demostración. (⇒) Supongamos φ inyectiva y sea $k \in \ker \varphi$. Entonces $\varphi(k) = e_H = \varphi(e_G)$, y la inyectividad obliga a $k = e_G$. Luego el núcleo solo contiene al neutro. (⇐) Supongamos $\ker \varphi = \{e_G\}$ y sean a, b con $\varphi(a) = \varphi(b)$. Multiplicando por $\varphi(b)^{-1}$,

$$\varphi(ab^{-1}) = \varphi(a)\varphi(b)^{-1} = e_H,$$

luego $ab^{-1} \in \ker \varphi = \{e_G\}$, es decir $ab^{-1} = e_G$ y por tanto $a = b$. Luego φ es inyectiva. □

Esta proposición es una máquina de ahorrar trabajo: comprobar inyectividad "a mano" exige mirar todos los pares de elementos, mientras que comprobar que el núcleo es trivial exige resolver *una sola* ecuación, $\varphi(g) = e$. Para homomorfismos siempre se hace lo segundo.

Ejemplo 23.45 (Tres homomorfismos, con núcleo e imagen calculados)**. (a) El signo,** $\operatorname{sgn} : S_3 \to \{+1,-1\}$, donde el segundo grupo lleva el producto de números. Es homomorfismo porque el signo es multiplicativo. Su núcleo son las permutaciones de signo $+1$, es decir $\ker(\operatorname{sgn}) = A_3 = \{\mathrm{id}, (1\,2\,3), (1\,3\,2)\}$, de orden 3; su imagen es todo $\{+1,-1\}$, de orden 2. No es inyectivo, y se ve al instante: el núcleo tiene tres elementos, no uno. **(b) La reducción módulo** n**,** $\pi : \mathbb{Z} \to \mathbb{Z}/n\mathbb{Z}$, $\pi(k) = \bar{k}$. Es homomorfismo porque $\overline{k+l} = \bar{k} + \bar{l}$ por definición de la suma de clases. Su imagen es todo $\mathbb{Z}/n\mathbb{Z}$ (sobreyectivo) y su núcleo es $\{k : n \mid k\} = n\mathbb{Z}$. **(c) Multiplicar por** 4 **en** $\mathbb{Z}/12\mathbb{Z}$, $\varphi(x) = 4x$. Es homomorfismo: $4(x+y) = 4x + 4y$. Núcleo: $4x \equiv 0 \pmod{12}$ equivale a $3 \mid x$, luego $\ker\varphi = \{0,3,6,9\}$, de orden 4. Imagen: $\{0,4,8\}$, de orden 3. Comprobación: $12/4 = 3$, el orden de la imagen. ✓ Esa coincidencia no es casual y es exactamente lo que afirmará el primer teorema de isomorfía.

23.2.9. Clases laterales, índice y el teorema de Lagrange

Definición 23.46 (Clase lateral e índice)**.** Sea $H \leq G$ y $g \in G$. La **clase lateral izquierda** de g es

$$gH = \{gh : h \in H\},$$

y análogamente la **clase lateral derecha** es $Hg = \{hg : h \in H\}$. El número de clases laterales izquierdas distintas se llama **índice** de H en G y se escribe $[G : H]$. En notación aditiva la clase se escribe $g + H$.

Lema 23.47 (Las clases laterales embaldosan el grupo)**.** *Sea $H \leq G$. Entonces:*

- ***(a)*** *todo $g \in G$ pertenece a alguna clase, concretamente a gH;*
- ***(b)*** *dos clases laterales o son idénticas o son disjuntas;*
- ***(c)*** *todas las clases tienen el mismo número de elementos que H.*

Demostración. **(a)** Como $e \in H$, se tiene $g = ge \in gH$. **(b)** Supongamos que aH y bH comparten un elemento x: existen $h_1, h_2 \in H$ con $x = ah_1 = bh_2$. Entonces $a = bh_2h_1^{-1}$. Ahora, cualquier elemento de aH es de la forma $ah = bh_2h_1^{-1}h$, y $h_2h_1^{-1}h \in H$ por ser H subgrupo; luego $aH \subseteq bH$. Por simetría $bH \subseteq aH$, y las dos clases coinciden. **(c)** La aplicación $H \to gH$ dada por $h \mapsto gh$ es sobreyectiva por definición de gH, e inyectiva porque $gh = gh'$ implica $h = h'$ por cancelación (proposición 23.15). Es una biyección, luego $|gH| = |H|$. □

Teorema 23.48 (Lagrange)**.** *Sea G un grupo finito y $H \leq G$. Entonces*

$$|G| = [G : H] \cdot |H|.$$

En particular, el orden de todo subgrupo divide al orden del grupo.

Demostración. Por el lema 23.47, las clases laterales izquierdas de H cubren todo G (apartado (a)) y no se solapan (apartado (b)): forman una partición de G. Sea $m = [G : H]$ el número de clases distintas, digamos $g_1H, \dots, g_mH$. Contando los elementos de G clase por clase y usando que cada una tiene $|H|$ elementos (apartado (c)):

$$|G| = |g_1H| + |g_2H| + \cdots + |g_mH| = \underbrace{|H| + \cdots + |H|}_{m \text{ veces}} = m\,|H| = [G : H]\,|H|.$$

Como m es un entero, $|H|$ divide a $|G|$. □

Ejemplo 23.49 (Las clases laterales de $H=\{0,4,8\}$ en $\mathbb{Z}/12\mathbb{Z}$)**.** $H=\langle 4\rangle$ tiene 3 elementos, así que Lagrange predice $[G:H]=12/3=4$ clases. Se calculan sumando:

$$0+H=\{0,4,8\},\quad 1+H=\{1,5,9\},\quad 2+H=\{2,6,10\},\quad 3+H=\{3,7,11\}.$$

Al llegar a 4 se repite: $4+H=\{4,8,0\}=0+H$. Y de hecho $5+H=1+H$, $6+H=2+H$, etcétera. **Verificación.** Cuatro clases, tres elementos cada una, disjuntas dos a dos, y su unión es $\mathbb{Z}/12\mathbb{Z}$ entero: $4\cdot 3=12$. ✓ Obsérvese que las clases son "los que dejan el mismo resto al dividir entre 4": el embaldosado no es una metáfora vaga, es literalmente una clasificación por restos.

Corolario 23.50 (El orden de un elemento divide al del grupo)**.** *Si G es finito y $a\in G$, entonces* $\operatorname{ord}(a)$ *divide a $|G|$. En consecuencia, $a^{|G|}=e$ para todo $a\in G$.*

Demostración. El subgrupo $\langle a\rangle$ tiene exactamente $\operatorname{ord}(a)$ elementos (definición 23.30), y por Lagrange su orden divide a $|G|$: pongamos $|G|=\operatorname{ord}(a)\cdot t$. Entonces

$$a^{|G|}=a^{\operatorname{ord}(a)\cdot t}=\left(a^{\operatorname{ord}(a)}\right)^{t}=e^{t}=e.$$ □

Corolario 23.51 (Todo grupo de orden primo es cíclico)**.** *Si $|G|=p$ con p primo, entonces $G\cong\mathbb{Z}/p\mathbb{Z}$, y además cualquier elemento distinto del neutro es generador.*

Demostración. Tomemos $a\in G$ con $a\neq e$. Por el corolario 23.50, $\operatorname{ord}(a)$ divide a p, luego vale 1 o p. No puede ser 1, porque eso significaría $a=e$. Luego $\operatorname{ord}(a)=p$ y el subgrupo $\langle a\rangle$ tiene p elementos, es decir, es todo G. Así que G es cíclico, generado por a, y la aplicación $\mathbb{Z}/p\mathbb{Z}\to G$ dada por $\bar{k}\mapsto a^{k}$ es un isomorfismo. □

Observación 23.52. Este corolario es sorprendentemente fuerte: dice que *no hay ninguna libertad* en los grupos de orden primo. Hay un único grupo de orden 2, uno de orden 3, uno de orden 5, uno de orden 1 000 003. Toda la variedad de la teoría de grupos vive en los órdenes compuestos. Y cuidado con el recíproco de Lagrange: que d divida a $|G|$ no garantiza que haya un subgrupo de orden d; el contraejemplo estándar, A_4 sin subgrupos de orden 6, se estudia más adelante en el capítulo.

23.2.10. Subgrupos normales, cociente y primer teorema de isomorfía

Definición 23.53 (Subgrupo normal)**.** Un subgrupo $N\leq G$ es **normal**, escrito $N\trianglelefteq G$, si

$$gN=Ng\quad\text{para todo } g\in G,$$

o equivalentemente si $gng^{-1}\in N$ para todos $g\in G$, $n\in N$. En un grupo abeliano todo subgrupo es normal, porque $gn=ng$ sin más. Todo subgrupo de índice 2 es normal, porque solo hay dos clases y una de ellas es N, luego la otra es $G\setminus N$ tanto por la izquierda como por la derecha.

Proposición 23.54 (Los núcleos son normales)**.** *Si $\varphi : G \to H$ es homomorfismo, entonces* $\ker\varphi \trianglelefteq G$.

Demostración. Ya sabemos que es subgrupo (proposición 23.43). Sean $k \in \ker\varphi$ y $g \in G$. Entonces

$$\varphi(gkg^{-1}) = \varphi(g)\,\varphi(k)\,\varphi(g)^{-1} = \varphi(g)\,e_H\,\varphi(g)^{-1} = e_H,$$

luego $gkg^{-1} \in \ker\varphi$. □

Teorema 23.55 (El grupo cociente)**.** *Sea $N \trianglelefteq G$ y sea G/N el conjunto de las clases laterales de N. La regla*

$$(aN)(bN) = (ab)N$$

está bien definida y convierte G/N en un grupo, llamado ***grupo cociente****, con neutro $N = eN$ e inverso $(aN)^{-1} = a^{-1}N$. Si G es finito, $|G/N| = |G|/|N| = [G : N]$.*

Demostración. Lo único que hay que comprobar es que el resultado no depende de los representantes elegidos. Supongamos $aN = a'N$ y $bN = b'N$, es decir $a' = an_1$ y $b' = bn_2$ con $n_1, n_2 \in N$. Entonces

$$a'b' = an_1bn_2 = ab\,(b^{-1}n_1b)\,n_2,$$

y $b^{-1}n_1b \in N$ precisamente porque N es normal. Luego $a'b' \in (ab)N$, de donde $(a'b')N = (ab)N$: el producto de clases no cambia. La asociatividad se hereda de G, la clase $eN = N$ es el neutro y $a^{-1}N$ es el inverso de aN. El recuento de elementos es la definición de índice más el teorema de Lagrange. □

Teorema 23.56 (Primer teorema de isomorfía)**.** *Sea $\varphi : G \to H$ un homomorfismo. Entonces*

$$G/\ker\varphi \;\cong\; \operatorname{im}\varphi.$$

Demostración. Escribamos $K = \ker\varphi$, que es normal por la proposición 23.54, de modo que G/K es un grupo. Definimos

$$\bar\varphi : G/K \longrightarrow \operatorname{im}\varphi, \qquad \bar\varphi(gK) = \varphi(g).$$

Está bien definida. Si $gK = g'K$, entonces $g^{-1}g' \in K$, luego $\varphi(g^{-1}g') = e$, es decir $\varphi(g)^{-1}\varphi(g') = e$ y por tanto $\varphi(g) = \varphi(g')$: el valor no depende del representante. **Es homomorfismo**, porque $\bar\varphi((aK)(bK)) = \bar\varphi((ab)K) = \varphi(ab) = \varphi(a)\varphi(b) = \bar\varphi(aK)\bar\varphi(bK)$. **Es sobreyectiva** sobre $\operatorname{im}\varphi$, porque todo elemento de la imagen es $\varphi(g) = \bar\varphi(gK)$ para algún g. **Y es inyectiva**: por la proposición 23.44 basta ver que su núcleo es trivial, y si $\bar\varphi(gK) = e$ entonces $\varphi(g) = e$, luego $g \in K$ y por tanto $gK = K$, que es el neutro de G/K. Biyectiva y homomorfismo: isomorfismo. □

Ejemplo 23.57 (El primer teorema de isomorfía en acción)**.** Retomamos $\varphi : \mathbb{Z}/12\mathbb{Z} \to \mathbb{Z}/12\mathbb{Z}$, $\varphi(x) = 4x$, con $\ker\varphi = \{0, 3, 6, 9\}$ e $\operatorname{im}\varphi = \{0, 4, 8\}$.

El teorema afirma $(\mathbb{Z}/12\mathbb{Z})/\{0, 3, 6, 9\} \cong \{0, 4, 8\}$. Comprobémoslo a mano. Las clases del núcleo son

$$\{0, 3, 6, 9\}, \qquad \{1, 4, 7, 10\}, \qquad \{2, 5, 8, 11\},$$

tres clases, coherente con $12/4 = 3$. ✓ Sumar clases es sumar sus representantes módulo 3, así que el cociente es cíclico de orden 3. Y la imagen $\{0, 4, 8\} = \langle 4 \rangle$ también es cíclica de orden 3. Los dos son $\mathbb{Z}/3\mathbb{Z}$ con otro nombre. **Cómo se usa en la práctica**: para identificar un cociente G/N que no reconoces, busca un homomorfismo que salga de G y cuyo núcleo sea exactamente N; el teorema te dice que el cociente es su imagen, que suele ser un grupo familiar.

23.2.11. Productos directos, y qué queda hecho

Definición 23.58 (Producto directo)**.** Dados dos grupos $(G, *)$ y $(H, \cdot)$, su **producto directo** es el conjunto de pares

$$G \times H = \{(g, h) : g \in G,\ h \in H\}$$

con la operación coordenada a coordenada, $(g_1, h_1)(g_2, h_2) = (g_1 * g_2,\ h_1 \cdot h_2)$. Es un grupo: el neutro es (e_G, e_H), el inverso es (g^{-1}, h^{-1}), y $|G \times H| = |G|\,|H|$. Es abeliano si y solo si lo son los dos factores, y el orden de un par es $\operatorname{ord}(g, h) = \operatorname{mcm}(\operatorname{ord}(g), \operatorname{ord}(h))$.

Ejemplo 23.59 ($\mathbb{Z}/2\mathbb{Z} \times \mathbb{Z}/3\mathbb{Z}$ es cíclico y $\mathbb{Z}/2\mathbb{Z} \times \mathbb{Z}/2\mathbb{Z}$ no)**.** En $\mathbb{Z}/2\mathbb{Z} \times \mathbb{Z}/3\mathbb{Z}$, el elemento $(1, 1)$ tiene orden $\operatorname{mcm}(2, 3) = 6$, luego genera los seis elementos del grupo: es cíclico y por tanto isomorfo a $\mathbb{Z}/6\mathbb{Z}$. En cambio, en $\mathbb{Z}/2\mathbb{Z} \times \mathbb{Z}/2\mathbb{Z}$ todo elemento cumple $\operatorname{mcm}(2, 2) = 2$, así que no hay ninguno de orden 4 y el grupo no es cíclico: es el de Klein. La diferencia entre los dos casos es que 2 y 3 son coprimos y 2 y 2 no; la sección sobre productos directos, más adelante, convierte esa observación en un criterio general.

Idea clave

Qué queda establecido. A partir de aquí el capítulo puede usar sin más explicación: operación binaria (23.3), grupo y grupo abeliano (23.9, 23.10), las reglas de cálculo con inversos (23.13–23.16), $\mathbb{Z}/n\mathbb{Z}$ (23.17), orden de un grupo y de un elemento (23.22, 23.23), subgrupo y su criterio (23.26, 23.27), subgrupo generado y grupo cíclico (23.30, 23.31), permutaciones, signo y grupo alternado (23.33–23.38), diédricos (23.40), homomorfismos, núcleo e imagen (23.41–23.44), clases laterales y Lagrange (23.46–23.51), normalidad, cociente y primer teorema de isomorfía (23.53–23.56) y producto directo (23.58). Todo lo que viene después son consecuencias o ejemplos de estas piezas: si te pierdes, vuelve aquí y localiza la definición exacta del objeto que te ha frenado, porque casi siempre el atasco no es de razonamiento sino de vocabulario.

23.3. Ahora sí: la definición, regla por regla

Todo lo que hicimos con el triángulo usó solo cuatro propiedades. Las aislamos y les damos nombre; a partir de aquí, cualquier cosa que las cumpla heredará automáticamente todo lo que demostremos.

Regla: los cuatro axiomas de grupo

Un conjunto G con una operación $*$ es un **grupo** si:

- **G1. Cierre.** Para todos $a, b \in G$, $a * b \in G$. (Operar no te saca del conjunto.)
- **G2. Asociatividad.** $(a * b) * c = a * (b * c)$. (Los paréntesis sobran.)
- **G3. Neutro.** Existe $e \in G$ con $e * a = a * e = a$ para todo a.
- **G4. Inverso.** Para cada a existe a^{-1} con $a * a^{-1} = a^{-1} * a = e$.

Si además $a * b = b * a$ siempre, el grupo se llama **abeliano** (o conmutativo).

Cuidado

Los tres errores clásicos al verificar que algo es un grupo.
Olvidar el cierre. Los números impares con la suma no forman grupo: $3 + 5 = 8$ no es impar. Antes de mirar nada más, comprueba que la operación no te expulsa del conjunto.
Confundir "tiene inverso" con "tiene inverso dentro del conjunto". $(\mathbb{Z}, \cdot)$ no es grupo: el inverso de 2 es $1/2$, que existe en $\mathbb{Q}$ pero no en $\mathbb{Z}$.
Dar por hecha la asociatividad. Casi siempre se cumple, pero no siempre: la resta no es asociativa, $(8 - 3) - 2 = 3$ mientras $8 - (3 - 2) = 7$. Por eso $(\mathbb{Z}, -)$ no es grupo aunque "se pueda deshacer".

Regla: consecuencias inmediatas de los axiomas

De G1–G4, sin ninguna hipótesis extra, se deducen:

- **(a)** El neutro es **único**. Si e y e' son neutros, $e = e * e' = e'$.
- **(b)** El inverso es **único**. Si b y c son inversos de a: $b = b*e = b*(a*c) = (b*a)*c = e*c = c$.
- **(c) Cancelación**: $ab = ac \Rightarrow b = c$ y $ba = ca \Rightarrow b = c$.
- **(d)** $(a^{-1})^{-1} = a$.
- **(e)** $(ab)^{-1} = b^{-1}a^{-1}$ **(¡en ese orden!)**.
- **(f)** La ecuación $ax = b$ tiene **solución única** $x = a^{-1}b$.

Observa cómo se demuestra (b): cada igualdad usa un axioma y solo uno. Esa es la textura de todas las demostraciones de este capítulo, y merece verse una vez despacio:

$$b \overset{\text{G3}}{=} b * e \overset{\text{G4}}{=} b * (a * c) \overset{\text{G2}}{=} (b * a) * c \overset{\text{G4}}{=} e * c \overset{\text{G3}}{=} c.$$

Pausa para pensar

¿Por qué en (e) el orden se invierte? Compruébalo: $(ab)(b^{-1}a^{-1}) = a(bb^{-1})a^{-1} = aea^{-1} = aa^{-1} = e$. Si hubiéramos escrito $(ab)(a^{-1}b^{-1})$ no habría forma de cancelar nada, porque b y a^{-1} no conmutan en general. En un grupo abeliano da igual; en D_3, no.

23.4. Un catálogo de grupos, verificado uno por uno

23.4.1. Los enteros con la suma

$(\mathbb{Z}, +)$: cierre ✓(suma de enteros es entera), asociatividad ✓, neutro 0 ✓, inverso de n es $-n$ ✓, y conmuta. Es abeliano e infinito. Ojo con el cambio de notación: aquí el "producto" es la suma, el "neutro" es 0 y "a^{-1}" se escribe $-a$. Es el mismo concepto con otra ropa.

23.4.2. El reloj: los enteros módulo n

$\mathbb{Z}_n = \{0, 1, \ldots, n-1\}$ con la suma *y luego tomar resto al dividir entre* n. Para $n = 4$:

$+$	0	1	2	3
0	0	1	2	3
1	1	2	3	0
2	2	3	0	1
3	3	0	1	2

Es el reloj de 4 horas: al pasar de 3 se vuelve a 0. Neutro 0; el inverso de 1 es 3 (porque $1 + 3 = 4 \equiv 0$), el de 2 es 2, el de 3 es 1. Abeliano, de orden 4.

23.4.3. El grupo de Klein

Toma $V = \{e, a, b, c\}$ con la regla "cada elemento es su propio inverso, y el producto de dos distintos no triviales es el tercero":

$\cdot$	e	a	b	c
e	e	a	b	c
a	a	e	c	b
b	b	c	e	a
c	c	b	a	e

Cuadrado latino ✓, neutro ✓, inversos ✓, y es abeliano (la tabla es simétrica respecto de la diagonal). ¿De dónde sale? De un rectángulo no cuadrado: e = no hacer nada, a = voltear horizontalmente, b = voltear verticalmente, c = rotar 180°. Hazlo con una hoja de papel: dos volteos seguidos dan el giro.

Idea clave

Dos grupos de orden 4, y son distintos. $\mathbb{Z}_4$ y V tienen ambos cuatro elementos, pero no son el mismo grupo. La forma más rápida de verlo: en $\mathbb{Z}_4$ hay un elemento (el 1) que al repetirse genera todo, y cumple $1+1=2\neq 0$; en V, en cambio, *todo* elemento cumple $x^2=e$. Como esa propiedad no depende de los nombres de los elementos, ningún renombramiento puede convertir uno en el otro. Este es el primer ejemplo de la pregunta que organiza toda la teoría: *¿cuántos grupos distintos hay de cada tamaño?* Y estos dos son, de hecho, los únicos de orden 4.

23.4.4. Ejemplos que *no* son grupos

Candidato	¿Grupo?	Por qué
$(\mathbb{N},+)$	No	no hay inversos ($-3\notin\mathbb{N}$)
$(\mathbb{Z},\cdot)$	No	solo ±1 tienen inverso
$(\mathbb{Q},\cdot)$	No	el 0 no tiene inverso
$(\mathbb{Q}^*,\cdot)$	**Sí**	quitando el 0, todo se arregla
$(\mathbb{Z},-)$	No	la resta no es asociativa
matrices 2×2 con $\cdot$	No	las singulares no tienen inversa
$GL_2(\mathbb{R})$ ($\det\neq 0$)	**Sí**	ese es justo el remedio

Pausa para pensar

¿Es $(\mathbb{Z}_4\setminus\{0\},\cdot)=\{1,2,3\}$ con el producto módulo 4 un grupo? Calcula $2\cdot 2=4\equiv 0$. El 0 no está en el conjunto: falla el cierre. En cambio $(\mathbb{Z}_5\setminus\{0\},\cdot)=\{1,2,3,4\}$ *sí* lo es (compruébalo: $2\cdot 3=6\equiv 1$, así que 2 y 3 son inversos; $4\cdot 4=16\equiv 1$). La diferencia es que 5 es primo y 4 no. Ese detalle es la semilla de los cuerpos finitos.

23.5. El orden de un elemento

Idea clave

Intuición primero. Toma un elemento y aplícalo una y otra vez: a, $a\cdot a$, $a\cdot a\cdot a$, … En un grupo finito esa lista no puede ser infinita sin repetirse, y en cuanto se repite se puede cancelar, así que tarde o temprano vuelves al punto de partida e. El **orden** de a es simplemente *cuántos pasos tarda en volver a casa.*

Es la misma idea de un reloj: girar la manecilla horaria de tres en tres horas te devuelve a las 12 al cabo de cuatro giros, porque $4\cdot 3=12$. Por eso en $\mathbb{Z}_n$ el orden de a es $n/\gcd(a,n)$: es el primer múltiplo de a que también es múltiplo de n, dividido entre a. Y por eso mismo el orden de un elemento siempre divide al tamaño del grupo: el paseo cierra un ciclo, y los ciclos embaldosan.

Regla: orden de un elemento

El **orden** de $a \in G$ es el menor entero $k > 0$ tal que $a^k = e$. Si no existe tal k, el orden es infinito. Se escribe $\text{ord}(a)$.
Método para calcularlo: ve multiplicando $a, a^2, a^3, \ldots$ hasta llegar a e por primera vez, y cuenta.

Ejemplo 23.60 (Órdenes en D_3)**.** $\text{ord}(e) = 1$, siempre, en cualquier grupo.
$\text{ord}(r)$: $r^1 = r \neq e$; $r^2 \neq e$; $r^3 = e$. Orden 3. (Tres giros de 120° son una vuelta completa.)
$\text{ord}(r^2)$: $(r^2)^1 = r^2$, $(r^2)^2 = r^4 = r$, $(r^2)^3 = r^6 = e$. Orden 3.
$\text{ord}(s_1)$: $s_1^2 = e$ según la tabla. Orden 2. Igual s_2, s_3.
Resumen: órdenes $1, 3, 3, 2, 2, 2$. Suman en la lista seis elementos, y *todos los órdenes dividen a* $6 = |D_3|$. No es casualidad: es el corolario de Lagrange.

Ejemplo 23.61 (Órdenes en $\mathbb{Z}_{12}$, calculados con una fórmula)**.** En $\mathbb{Z}_n$ el "elevar a la k" es sumar k veces, así que $\text{ord}(a)$ es el menor k con $ka \equiv 0 \pmod{n}$, es decir

$$\boxed{\text{ord}(a) = \frac{n}{\gcd(a, n)}}$$

Para $n = 12$: $\text{ord}(0) = 1$, $\text{ord}(1) = 12$, $\text{ord}(2) = 12/2 = 6$, $\text{ord}(3) = 4$, $\text{ord}(4) = 3$, $\text{ord}(5) = 12/1 = 12$, $\text{ord}(6) = 2$, $\text{ord}(7) = 12$, $\text{ord}(8) = 3$, $\text{ord}(9) = 4$, $\text{ord}(10) = 6$, $\text{ord}(11) = 12$.

Comprobación de sentido: los elementos de orden 12 son $1, 5, 7, 11$, exactamente los coprimos con 12, y hay $\varphi(12) = 4$ de ellos. Todos los órdenes $(1, 2, 3, 4, 6, 12)$ son divisores de 12, y aparecen *todos* los divisores: eso es característico de los grupos cíclicos.

23.6. Subgrupos: mundos dentro del mundo

Idea clave

Intuición primero. Un subgrupo es un trozo del grupo que *se basta a sí mismo*: si operas dentro de él nunca te sales, y todo lo que hagas se puede deshacer sin salir. Es como un idioma dentro de otro más grande, con vocabulario propio y cerrado.
La forma práctica de encontrarlos todos es al revés de lo que uno esperaría: no se buscan a ciegas, se *filtra* con Lagrange. Como el tamaño de un subgrupo tiene que dividir al del grupo, en un grupo de orden 6 solo pueden existir subgrupos de tamaño 1, 2, 3 y 6, y para cada tamaño posible basta preguntarse qué elementos tienen el orden adecuado. La búsqueda pasa de infinita a una lista corta.

Regla: criterio de subgrupo

$H \subseteq G$ es **subgrupo** ($H \leq G$) si:

S1. $H \neq \varnothing$ (basta comprobar $e \in H$);

S2. si $a, b \in H$ entonces $ab \in H$;

S3. si $a \in H$ entonces $a^{-1} \in H$.

Atajo: $H \neq \varnothing$ y "$a, b \in H \Rightarrow ab^{-1} \in H$" equivale a las tres.
Atajo para grupos finitos: si G es finito, basta S1 y S2. (Si $a \in H$ y H es finito y cerrado, las potencias $a, a^2, \ldots$ se repiten, de donde $a^k = e$ para algún k y $a^{-1} = a^{k-1} \in H$.)

Ejemplo 23.62 (Todos los subgrupos de D_3, encontrados sistemáticamente)**. Paso 1. Usar Lagrange como filtro.** Un subgrupo tiene orden divisor de 6: solo puede ser 1, 2, 3 o 6.

Paso 2. Orden 1 y orden 6. Son $\{e\}$ y D_3 entero. Siempre existen; se llaman *triviales*.

Paso 3. Orden 2. Un subgrupo de orden 2 es $\{e, x\}$ con $x^2 = e$. Los elementos de orden 2 son s_1, s_2, s_3: tres subgrupos, $\{e, s_1\}$, $\{e, s_2\}$, $\{e, s_3\}$.

Paso 4. Orden 3. Necesita un elemento de orden 3: r o r^2. Ambos generan el mismo: $\{e, r, r^2\}$. Un solo subgrupo.

Total: 6 subgrupos. Y ninguno de orden 4 ni 5, porque no dividen a 6.

Cuidado

Lagrange no tiene recíproco. Que d divida a $|G|$ *no* garantiza que exista un subgrupo de orden d. El contraejemplo estándar es A_4, de orden 12: no tiene ningún subgrupo de orden 6, aunque $6 \mid 12$. (Sus subgrupos tienen órdenes $1, 2, 3, 4, 12$.) Lo que sí es cierto siempre es el recíproco parcial de Cauchy —si un *primo* p divide a $|G|$, hay un elemento, y por tanto un subgrupo, de orden p— y el de Sylow para potencias de primo.

23.7. El grupo del cuadrado: D_4 con todo detalle

Repetimos el experimento con un cuadrado. Ahora hay 8 simetrías: cuatro rotaciones $(0°, 90°, 180°, 270°)$ y cuatro reflexiones (dos por las diagonales, dos por las mediatrices). Llamemos r a la rotación de 90° y s a la reflexión por el eje vertical. Entonces

$$D_4 = \{e,\ r,\ r^2,\ r^3,\ s,\ rs,\ r^2s,\ r^3s\}, \qquad |D_4| = 8.$$

Regla: la relación que define a los diédricos

En D_n se cumplen tres identidades, y con ellas se calcula *cualquier* producto sin volver a tocar el cartón:

$$r^n = e, \qquad s^2 = e, \qquad \boxed{srs = r^{-1}} \quad \text{(equivalentemente } sr = r^{-1}s\text{)}.$$

La tercera es la importante: dice que *pasar una reflexión al otro lado de una rotación invierte el giro*. Con ella, todo producto se reduce a la forma $r^i s^j$ con $0 \leq i < n$, $j \in \{0, 1\}$.

Ejemplo 23.63 (Reducir un producto usando solo las relaciones)**.** Calculemos $(r^2s)(r^3s)$ en D_4.

Paso 1. Escribimos todo seguido: $r^2\, s\, r^3\, s$.

Paso 2. Movemos la s del medio hacia la derecha atravesando r^3. Cada vez que cruza una r la invierte, así que $sr^3 = r^{-3}s$:

$$r^2\,(s\,r^3)\,s = r^2\,r^{-3}\,s\,s.$$

Paso 3. $ss = s^2 = e$:

$$= r^{2-3} = r^{-1}.$$

Paso 4. Como $r^4 = e$, $r^{-1} = r^3$.

Resultado: $(r^2s)(r^3s) = r^3$. Fíjate en lo que pasó: reflexión por reflexión dio rotación, igual que en el triángulo.

Ejemplo 23.64 (Los 10 subgrupos de D_4)**.** Lagrange permite órdenes $1, 2, 4, 8$. Vamos uno a uno.

Orden 1: $\{e\}$. (1 subgrupo)

Orden 2: uno por cada elemento de orden 2. Los órdenes en D_4 son: $\mathrm{ord}(e) = 1$; $\mathrm{ord}(r) = \mathrm{ord}(r^3) = 4$; $\mathrm{ord}(r^2) = 2$; y las cuatro reflexiones s, rs, r^2s, r^3s tienen orden 2. Total: cinco elementos de orden 2, luego **5** subgrupos de orden 2. (5)

Orden 4: hay tres. El cíclico $\langle r\rangle = \{e, r, r^2, r^3\}$; y dos copias del grupo de Klein: $\{e, r^2, s, r^2s\}$ (las reflexiones por las mediatrices más el giro de media vuelta) y $\{e, r^2, rs, r^3s\}$ (las dos diagonales más el giro). (3)

Orden 8: D_4 entero. (1)

Total: $1 + 5 + 3 + 1 = 10$ **subgrupos.**

Pausa para pensar

En D_4 hay dos subgrupos de orden 4 isomorfos al grupo de Klein y uno isomorfo a $\mathbb{Z}_4$. ¿Puedes decir, sin calcular, por qué $\{e, s, rs, r^2s\}$ *no* es subgrupo? Pista: $s \cdot rs = ?$ Usando $sr = r^{-1}s$: $s(rs) = (sr)s = r^{-1}s\,s = r^{-1} = r^3$, que no está en el conjunto. Falla el cierre.

23.8. Permutaciones con las manos

Idea clave

Intuición primero. Una permutación no es una lista de números: es una *instrucción de reordenar*. La notación de ciclos $(1\,2\,3)$ se lee como un carrusel: el 1 va al puesto del 2, el 2 al del 3, el 3 al del 1, y lo que no aparece no se mueve. Descomponer en ciclos es descubrir en qué carruseles independientes se rompe el barajado.

Dos costumbres que evitan casi todos los errores. La primera: se multiplica **de derecha a izquierda**, porque son funciones y fg significa "primero g, luego f". La segunda: ciclos que no comparten números conmutan entre sí, porque están trabajando en mesas distintas; los que comparten números, casi nunca. Y el signo de una permutación es la

paridad del número de trasposiciones: eso es lo que después distingue A_n dentro de S_n y lo que acaba explicando la quíntica.

S_n es el grupo de todas las biyecciones de $\{1, \dots, n\}$, y por el teorema de Cayley *todo* grupo finito vive dentro de algún S_n. Merece calcularse despacio.

Regla: notación de ciclos

$(a_1\, a_2\, \dots\, a_k)$ significa: $a_1 \mapsto a_2$, $a_2 \mapsto a_3$, $\dots$, $a_k \mapsto a_1$, y todo lo demás se queda quieto.
Para multiplicar: se lee **de derecha a izquierda**. Se toma un número, se ve qué le hace el ciclo de la derecha, luego qué le hace al resultado el siguiente, y así.
Para descomponer una permutación en ciclos: empieza por el 1, sigue la cadena hasta volver, cierra el paréntesis, empieza por el menor número no usado, repite.

Ejemplo 23.65 (Multiplicar $(1\,2\,3)(2\,4)$ paso a paso)**.** Recordatorio: primero actúa $(2\,4)$, después $(1\,2\,3)$.

¿A dónde va el 1? $(2\,4)$ lo deja quieto; $(1\,2\,3)$ lo manda a 2. Total: $1 \mapsto 2$.

¿A dónde va el 2? $(2\,4)$ lo manda a 4; $(1\,2\,3)$ deja quieto al 4. Total: $2 \mapsto 4$.

¿A dónde va el 4? $(2\,4)$ lo manda a 2; $(1\,2\,3)$ manda 2 a 3. Total: $4 \mapsto 3$.

¿A dónde va el 3? $(2\,4)$ lo deja; $(1\,2\,3)$ manda 3 a 1. Total: $3 \mapsto 1$.

Armamos los ciclos: empezamos en 1: $1 \to 2 \to 4 \to 3 \to 1$. Un solo ciclo de longitud 4.

$$(1\,2\,3)(2\,4) = (1\,2\,4\,3).$$

Ejemplo 23.66 (El orden de una permutación)**.** Sea $\sigma = (1\,2\,3)(4\,5)$, ciclos disjuntos.

Paso 1. El ciclo de longitud 3 vuelve a su sitio cada 3 aplicaciones; el de longitud 2, cada 2.

Paso 2. Los dos vuelven *a la vez* en el mínimo común múltiplo: $\operatorname{mcm}(3, 2) = 6$.

Paso 3. Verificación: $\sigma^2 = (1\,3\,2)$, $\sigma^3 = (4\,5)$, $\sigma^6 = e$. ✓

Regla general: el orden de una permutación es el mcm de las longitudes de sus ciclos disjuntos.

Ejemplo 23.67 (Inventario completo de S_4)**.** $|S_4| = 4! = 24$. Clasificamos por tipo cíclico:

Tipo	Cuántas	Orden	Signo
identidad	1	1	par
$(a\,b)$	6	2	impar
$(a\,b)(c\,d)$	3	2	par
$(a\,b\,c)$	8	3	par
$(a\,b\,c\,d)$	6	4	impar
	24		

Cuenta de la fila de los 3-ciclos: hay $\binom{4}{3} = 4$ maneras de elegir los tres símbolos y 2 ciclos distintos con cada terna ($(a\,b\,c)$ y $(a\,c\,b)$): $4 \cdot 2 = 8$. ✓ De los 4-ciclos: $3! = 6$ (fijamos el 1 al

principio y ordenamos los otros tres). ✓ De las dobles trasposiciones: emparejar 4 objetos en 2 parejas se hace de 3 maneras. ✓

Los pares suman $1+3+8=12=|A_4|$, la mitad. ✓

Cuidado

El signo *no* es lo mismo que el orden. Una doble trasposición como $(1\,2)(3\,4)$ tiene orden 2 pero es *par* (dos trasposiciones). Una trasposición sola tiene orden 2 y es *impar*. Para el signo cuenta cuántas trasposiciones necesitas; para el orden, cuántas veces hay que repetir la permutación entera.

23.9. Lagrange con las cuentas hechas

Idea clave

Intuición primero. Las clases laterales gH son copias trasladadas de H dentro de G. Todas tienen exactamente el mismo tamaño, ninguna se solapa con otra, y entre todas cubren G por completo: son baldosas del mismo formato cubriendo el suelo sin huecos ni superposiciones. Contar las baldosas y multiplicar por el tamaño de una da el suelo entero, y eso es literalmente el teorema de Lagrange, $|G| = [G:H]\cdot|H|$.
De ahí sale gratis casi todo lo elemental: el orden de un elemento divide al del grupo (mira el subgrupo que genera), todo grupo de orden primo es cíclico (no hay subgrupos intermedios donde esconderse) y $a^{|G|}=e$ siempre. El pequeño teorema de Fermat del capítulo 3 no es más que este enunciado aplicado al grupo de unidades módulo p.

Regla: clases laterales

Dado $H \leq G$ y $g \in G$, la **clase lateral izquierda** es $gH = \{gh : h \in H\}$.
Tres hechos, y con ellos se demuestra Lagrange:

(a) Todas las clases tienen el mismo tamaño $|H|$ (la aplicación $h \mapsto gh$ es biyectiva por cancelación).

(b) Dos clases o son iguales o son disjuntas (si comparten un elemento, coinciden).

(c) Todo g está en alguna clase (concretamente en gH, porque $e \in H$).

Conclusión: las clases **embaldosan** G en trozos iguales, luego $|G| = [G:H]\cdot|H|$.

Ejemplo 23.68 (Las clases de $H=\{e,s_1\}$ en D_3, todas calculadas)**.** $|H|=2$, así que esperamos $6/2=3$ clases de 2 elementos.

Clase de e: $eH=\{e,s_1\}$.

Clase de r: $rH=\{re,\ rs_1\}=\{r,\ s_3\}$ (usando la tabla).

Clase de r^2: $r^2H=\{r^2,\ r^2s_1\}=\{r^2,\ s_2\}$.

Verificación: $\{e,s_1\}\cup\{r,s_3\}\cup\{r^2,s_2\}=D_3$ y las tres son disjuntas. ✓ Tres clases de dos elementos, $3\cdot 2=6$. ✓

Un detalle crucial: calculemos ahora la clase *derecha* $Hr = \{r,\ s_1r\} = \{r,\ s_2\}$. ¡No coincide con $rH = \{r, s_3\}$! Este subgrupo **no es normal**, y esa es exactamente la razón por la que no se puede formar D_3/H.

Ejemplo 23.69 (Las clases de $A_3 = \{e, r, r^2\}$ en D_3)**.** $|A_3| = 3$, esperamos $6/3 = 2$ clases.

$eA_3 = \{e, r, r^2\}$ (las rotaciones). $s_1A_3 = \{s_1,\ s_1r,\ s_1r^2\} = \{s_1, s_2, s_3\}$ (las reflexiones).

Y por la derecha: $A_3s_1 = \{s_1,\ rs_1,\ r^2s_1\} = \{s_1, s_3, s_2\}$: *el mismo conjunto*. Aquí izquierda y derecha coinciden, así que A_3 **sí** es normal, y podemos formar el cociente D_3/A_3, que tiene 2 elementos: "rotaciones" y "reflexiones". Ese cociente es $\mathbb{Z}_2$, y su significado es el signo: par por par es par, par por impar es impar. El cociente *es* la regla de los signos.

Idea clave

Por qué todo subgrupo de índice 2 es normal. Si $[G : H] = 2$ solo hay dos clases: H y "el resto". Por la izquierda: H y $G \setminus H$. Por la derecha: H y $G \setminus H$. Son las mismas dos piezas, sin más opción. Por eso $A_n \trianglelefteq S_n$ siempre, sin calcular nada.

23.10. Cocientes: mirar con lentes borrosos

Idea clave

Intuición primero. Un cociente es lo que ves cuando decides *dejar de distinguir* ciertas cosas. Al pasar a G/N acuerdas que dos elementos que difieren en algo de N cuentan como el mismo; es desenfocar la imagen a propósito para que resalte lo que queda.

El truco, y la única dificultad real, es que el desenfoque debe ser compatible con la operación: si borras detalles antes de multiplicar tienes que obtener lo mismo que si multiplicas y borras después. Eso es lo que falla cuando N no es normal, y el contraejemplo de abajo lo muestra con números delante. La normalidad ($gNg^{-1} = N$) no es un capricho: es exactamente la condición de que la borrosidad se vea igual desde cualquier punto del grupo. El caso familiar es $\mathbb{Z}/n\mathbb{Z}$: olvidar todo salvo el resto módulo n es un cociente, y funciona porque en $\mathbb{Z}$, abeliano, todo subgrupo es normal.

Regla: cuándo existe el grupo cociente

$N \leq G$ es **normal** ($N \trianglelefteq G$) si $gN = Ng$ para todo g (equivalentemente $gng^{-1} \in N$ para todos $g \in G$, $n \in N$).

Solo entonces la regla $(aN)(bN) = (ab)N$ está **bien definida**, y el conjunto de clases con esa operación es el **grupo cociente** G/N, de orden $|G|/|N|$.

Cuidado

"Bien definida" significa: si eliges otro representante de la misma clase, el resultado no cambia. Es la única dificultad real del concepto de cociente, y conviene verla fallar con números delante.

En D_3 tomamos $H = \{e, s_1\}$, que ya vimos que *no* es normal. Sus clases son $H = \{e, s_1\}$, $rH = \{r, s_3\}$ y $r^2H = \{r^2, s_2\}$. Intentemos calcular el "producto" $(rH)(rH)$ de dos maneras:
Con los representantes r y r: $r \cdot r = r^2$, cuya clase es $r^2H = \{r^2, s_2\}$.
Con los representantes s_3 y s_3 (y s_3 está en rH, es igual de legítimo): $s_3 \cdot s_3 = e$, cuya clase es $H = \{e, s_1\}$.
Dos respuestas distintas para la misma pareja de clases: la operación **no existe**. Cuando el subgrupo no es normal no hay grupo cociente, y no por falta de ingenio, sino porque no hay nada que definir.

Ejemplo 23.70 ($\mathbb{Z}/3\mathbb{Z}$: el cociente más famoso, hecho a mano)**.** $G = \mathbb{Z}$, $N = 3\mathbb{Z} = \{\dots, -3, 0, 3, 6, \dots\}$. Como $\mathbb{Z}$ es abeliano, N es normal automáticamente.

Las clases: $0 + N = \{\dots, -3, 0, 3, \dots\}$; $1 + N = \{\dots, -2, 1, 4, \dots\}$; $2 + N = \{\dots, -1, 2, 5, \dots\}$. Tres clases: los que dejan resto 0, 1 y 2 al dividir entre 3.

La operación: $(1 + N) + (2 + N) = 3 + N = 0 + N$. Es decir, $1 + 2 = 0$ en el cociente.

Identificación: $\mathbb{Z}/3\mathbb{Z} \cong \mathbb{Z}_3$. El grupo cociente *es* la aritmética modular, y "tomar cociente" es exactamente "decidir ignorar los múltiplos de 3".

Ejemplo 23.71 ($D_4/\{e, r^2\}$: un cociente no obvio)**.** Sea $N = \{e, r^2\}$, el centro de D_4. Es normal porque r^2 conmuta con todo (compruébalo con $sr^2 = r^{-2}s = r^2s$, ya que $r^{-2} = r^2$).

Las cuatro clases $(8/2 = 4)$: $N = \{e, r^2\}$, $rN = \{r, r^3\}$, $sN = \{s, r^2s\}$, $rsN = \{rs, r^3s\}$.

La tabla: llamemos $\bar{e}, \bar{r}, \bar{s}, \overline{rs}$ a las cuatro. Como $r^2 \in N$, en el cociente $\bar{r}^2 = \bar{e}$; también $\bar{s}^2 = \bar{e}$ y $\overline{rs}^2 = \bar{e}$. Todos los elementos son su propio inverso.

Conclusión: $D_4/Z(D_4) \cong V$, el grupo de Klein. Y observa el mensaje: al "desenfocar" hasta borrar el giro de media vuelta, el grupo no conmutativo D_4 se vuelve conmutativo. Los cocientes simplifican; lo que se pierde es lo que estaba en N.

23.11. Isomorfismos: cuándo dos grupos son el mismo

Idea clave

Intuición primero. Dos grupos son isomorfos cuando son el mismo grupo con los nombres cambiados: existe un diccionario que traduce elementos de uno en elementos del otro y que respeta la tabla de multiplicar. Las simetrías de un triángulo y las permutaciones de tres letras *son* el mismo objeto; solo cambia si lo cuentas mirando un cartón o mirando una lista.
Demostrar que dos grupos sí son isomorfos exige construir el diccionario. Demostrar que *no* lo son es más fácil y más astuto: basta exhibir una propiedad que sobreviva a cualquier cambio de nombres y que uno tenga y el otro no. El multiconjunto de órdenes es el mejor detector barato, porque un diccionario que respeta productos respeta también cuántos pasos tarda cada elemento en volver a e: $\mathbb{Z}_4$ tiene un elemento de orden 4 y el grupo de Klein no tiene ninguno, luego no hay diccionario posible.

Regla: cómo demostrar que $G \cong H$ y cómo demostrar que no

Para probar que SÍ: construye una biyección $\varphi : G \to H$ con $\varphi(ab) = \varphi(a)\varphi(b)$. Normalmente basta decir a dónde van los generadores y comprobar que las relaciones se respetan.
Para probar que NO: encuentra una propiedad que solo dependa de la estructura y que uno tenga y el otro no. Las más útiles:

- el orden del grupo;
- si es abeliano o no;
- el *multiconjunto de órdenes de los elementos* (¡el más potente!);
- el número de subgrupos de cada orden;
- el orden del centro.

Ejemplo 23.72 ($\mathbb{Z}_4 \ncong V$, con rigor)**.** Ambos tienen orden 4 y ambos son abelianos, así que los dos primeros criterios no bastan.

Órdenes en $\mathbb{Z}_4$: $1, 4, 2, 4$ (los elementos $0, 1, 2, 3$). Hay un elemento de orden 4.

Órdenes en V: $1, 2, 2, 2$. No hay ningún elemento de orden 4.

Argumento: un isomorfismo preserva órdenes, pues $\varphi(a)^k = \varphi(a^k)$ y φ es inyectiva con $\varphi(e) = e$. Si existiera $\varphi : \mathbb{Z}_4 \to V$, el elemento $\varphi(1)$ tendría orden 4 en V: imposible. Luego no existe. ■

Ejemplo 23.73 ($\mathbb{Z}_6 \cong \mathbb{Z}_2 \times \mathbb{Z}_3$: un isomorfismo que sí existe)**. Paso 1.** Ambos tienen orden 6.

Paso 2. En $\mathbb{Z}_2 \times \mathbb{Z}_3$, el elemento $(1, 1)$ tiene orden $\operatorname{mcm}(2, 3) = 6$: genera todo.

Paso 3. Definimos $\varphi : \mathbb{Z}_6 \to \mathbb{Z}_2 \times \mathbb{Z}_3$ por $\varphi(k) = (k \bmod 2,\ k \bmod 3)$.

Paso 4. Es homomorfismo (reducir módulo 2 y módulo 3 respeta sumas) e inyectivo: si $k \equiv 0$ mód 2 y mód 3, entonces $6 \mid k$, luego $k = 0$ en $\mathbb{Z}_6$. Inyectivo entre conjuntos del mismo tamaño finito $\Rightarrow$ biyectivo. ■

Nota: esto es el teorema chino del resto del capítulo 3, disfrazado de grupos. Y falla para $\mathbb{Z}_4$ vs $\mathbb{Z}_2 \times \mathbb{Z}_2$ precisamente porque 2 y 2 no son coprimos.

Regla: el primer teorema de isomorfía, como herramienta de cálculo

Si $\varphi : G \to H$ es homomorfismo, entonces

$$G/\ker\varphi \ \cong\ \operatorname{im}\varphi.$$

Cómo usarlo en la práctica: para identificar un cociente G/N, busca un homomorfismo que salga de G y cuyo núcleo sea justo N; entonces G/N es su imagen, que suele ser un grupo conocido.

Ejemplo 23.74 (Tres cocientes identificados de un golpe)**. (i)** $\operatorname{sgn} : S_n \to \{\pm 1\}$ es sobreyectivo con núcleo A_n, luego $S_n/A_n \cong \mathbb{Z}_2$.

(ii) $\det : GL_n(\mathbb{R}) \to \mathbb{R}^*$ es sobreyectivo con núcleo $SL_n(\mathbb{R})$, luego $GL_n/SL_n \cong \mathbb{R}^*$.

(iii) $\varphi : \mathbb{R} \to S^1$, $\varphi(t) = e^{2\pi i t}$, es sobreyectivo con núcleo $\mathbb{Z}$, luego $\mathbb{R}/\mathbb{Z} \cong S^1$: la recta enrollada en una circunferencia. Y como S^1 es el grupo de las rotaciones del plano, acabas de ver por qué "medir ángulos módulo una vuelta" es un cociente de grupos.

23.12. Acciones: los grupos puestos a trabajar

Idea clave

Intuición primero. Hasta aquí un grupo era una tabla; una *acción* lo pone a hacer algo: cada elemento del grupo mueve las piezas de un conjunto. La órbita de una pieza es el conjunto de sitios a los que puede llegar; su estabilizador, las maniobras que la dejan exactamente donde estaba.

La fórmula órbita–estabilizador, $|Gx| \cdot |G_x| = |G|$, es Lagrange disfrazado y es una máquina de contar: si quieres saber cuántas simetrías tiene un objeto, no las enumeres, elige una pieza, cuenta a cuántos sitios va y cuántas maniobras la fijan, y multiplica. El cubo se resuelve así en dos líneas, y da igual si eliges caras, vértices o aristas: las tres cuentas tienen que dar 24, lo que sirve además como comprobación.

Regla: órbita, estabilizador y la fórmula que lo cuenta todo

Si G actúa sobre X:

- **órbita** de x: $Gx = \{g \cdot x : g \in G\}$ — "a dónde puede llegar x";
- **estabilizador** de x: $G_x = \{g : g \cdot x = x\}$ — "qué lo deja quieto"; es subgrupo.

$$\boxed{|Gx| \cdot |G_x| = |G|} \qquad \text{(órbita–estabilizador)}$$

Ejemplo 23.75 (Cuántas rotaciones tiene un cubo, sin enumerarlas)**. Paso 1.** G = rotaciones del cubo, X = las 6 caras.

Paso 2. La órbita de una cara son las 6 caras (cualquier cara se puede llevar a cualquier otra): $|Gx| = 6$.

Paso 3. El estabilizador de una cara: las rotaciones que la dejan en su sitio son los giros alrededor del eje perpendicular a ella, $0°, 90°, 180°, 270°$: $|G_x| = 4$.

Paso 4. $|G| = 6 \cdot 4 = 24$.

Y si hubiéramos usado los 8 vértices: órbita 8, estabilizador 3 (giros por la diagonal), $8 \cdot 3 = 24$. ✓ Con las 12 aristas: $12 \cdot 2 = 24$. ✓ Tres caminos, una respuesta.

Regla: lema de Burnside (contar sin repetir)

El número de órbitas es el **promedio de puntos fijos**:

$$\#\text{órbitas} = \frac{1}{|G|}\sum_{g\in G} |\operatorname{Fij}(g)|,$$

donde $\operatorname{Fij}(g) = \{x \in X : g \cdot x = x\}$.
Para qué sirve: contar configuraciones "distintas salvo simetría" —collares, coloraciones, moléculas, tableros— sin listarlas.

Ejemplo 23.76 (¿Cuántos collares de 6 cuentas con 2 colores hay?)**.** Dos collares son iguales si uno se obtiene del otro *girándolo*. $G = \mathbb{Z}_6$ (las 6 rotaciones), $X =$ las $2^6 = 64$ pulseras pintadas.

Paso 1. Puntos fijos de cada rotación. Una rotación de k posiciones fija una coloración si esta es constante en cada ciclo; el número de ciclos es $\gcd(k, 6)$, así que $|\operatorname{Fij}(k)| = 2^{\gcd(k,6)}$:

k	$\gcd(k,6)$	$2^{\gcd}$
0	6	64
1	1	2
2	2	4
3	3	8
4	2	4
5	1	2

Paso 2. Sumar: $64 + 2 + 4 + 8 + 4 + 2 = 84$.

Paso 3. Promediar: $84/6 = 14$.

Respuesta: 14 collares distintos. (Si además se permitiera darles la vuelta, el grupo sería D_6 con 12 elementos y saldrían 13.)

Ejemplo 23.77 (Cubos pintados con 3 colores)**.** $G =$ las 24 rotaciones del cubo, $X =$ coloraciones de las 6 caras con 3 colores.

Paso 1. Clasificar las 24 rotaciones y sus ciclos sobre las caras.

Tipo	**Cuántas**	**Ciclos en caras**	**Fijos**
identidad	1	6	$3^6 = 729$
caras $\pm 90°$	6	3	$3^3 = 27$
caras $180°$	3	4	$3^4 = 81$
vértices $\pm 120°$	8	2	$3^2 = 9$
aristas $180°$	6	3	$3^3 = 27$
	24		

Paso 2. Sumar con multiplicidad: $1(729) + 6(27) + 3(81) + 8(9) + 6(27) = 729 + 162 + 243 + 72 + 162 = 1368$.

Paso 3. Dividir: $1368/24 = 57$.

Respuesta: 57 cubos esencialmente distintos. Intenta contarlos a mano y entenderás por qué este lema es un milagro práctico.

23.13. ¿Cuántos grupos hay de cada tamaño?

Idea clave

Intuición primero. La pregunta "¿cuántos grupos hay de orden n?" es la versión algebraica de "¿cuántas figuras distintas se pueden armar con n piezas?". La respuesta depende sorprendentemente poco de n y muchísimo de *cómo se factoriza* n.
Dos reglas explican casi toda la tabla. Si n es primo, hay exactamente un grupo: por Lagrange no hay subgrupos intermedios, cualquier elemento distinto de e genera todo, y sale $\mathbb{Z}_n$. Cuanto más divisible es n —más potencias de primos y más primos que se dividen entre sí menos uno—, más margen hay para pegar piezas de formas no conmutativas y más grupos aparecen: por eso el orden 8 ya tiene cinco y el orden 16 tiene catorce. La clasificación completa es un programa gigantesco, y su cumbre es la de los grupos simples finitos.

Esta es la pregunta que organiza la teoría. La respuesta para los órdenes pequeños, que ya podemos justificar casi entera:

Orden	Cuántos	Cuáles
1	1	trivial
2	1	$\mathbb{Z}_2$
3	1	$\mathbb{Z}_3$
4	2	$\mathbb{Z}_4$, $V = \mathbb{Z}_2 \times \mathbb{Z}_2$
5	1	$\mathbb{Z}_5$
6	2	$\mathbb{Z}_6$, $S_3 = D_3$
7	1	$\mathbb{Z}_7$
8	5	$\mathbb{Z}_8$, $\mathbb{Z}_4 \times \mathbb{Z}_2$, $\mathbb{Z}_2^3$, D_4, Q_8
9	2	$\mathbb{Z}_9$, $\mathbb{Z}_3 \times \mathbb{Z}_3$
10	2	$\mathbb{Z}_{10}$, D_5
11	1	$\mathbb{Z}_{11}$
12	5	$\mathbb{Z}_{12}$, $\mathbb{Z}_6 \times \mathbb{Z}_2$, D_6, A_4, $\mathbb{Z}_3 \rtimes \mathbb{Z}_4$

Regla: los criterios que explican la tabla

(a) Orden primo ⇒ un solo grupo, cíclico. Por Lagrange, todo elemento $\neq e$ tiene orden p, luego genera.

(b) Orden p^2 ⇒ abeliano, y solo hay dos: $\mathbb{Z}_{p^2}$ y $\mathbb{Z}_p \times \mathbb{Z}_p$.

(c) Orden pq con $p < q$ y $p \nmid q-1$ ⇒ cíclico. Por eso solo hay un grupo de orden 15, 33 ó 35.

(d) Orden $2q$ con q primo impar ⇒ dos grupos: $\mathbb{Z}_{2q}$ y el diédrico D_q. Por eso hay dos de orden 6 y dos de orden 10.

Ejemplo 23.78 (Por qué solo hay dos grupos de orden 6)**. Paso 1.** Por Cauchy hay un elemento a de orden 3 y uno b de orden 2.

Paso 2. $\langle a\rangle$ tiene índice 2, luego es normal. Entonces $bab^{-1} \in \langle a\rangle$ y tiene orden 3: es a o a^2.

Paso 3, caso I: $bab^{-1} = a$, o sea $ab = ba$. Todo conmuta, ab tiene orden 6, el grupo es $\mathbb{Z}_6$.

Paso 4, caso II: $bab^{-1} = a^2 = a^{-1}$. Esa es *exactamente* la relación diédrica $srs = r^{-1}$: el grupo es $D_3 = S_3$.

Paso 5. No hay más casos. Dos grupos, ni uno más. ■

Fíjate en la estructura del argumento: Cauchy da los ladrillos, la normalidad da la relación de conmutación, y la relación solo admite dos valores. Es el esquema con el que se clasifican todas las familias pequeñas.

Idea clave

El cuaternio Q_8, el ejemplo que rompe las intuiciones. $Q_8 = \{\pm 1, \pm i, \pm j, \pm k\}$ con $i^2 = j^2 = k^2 = ijk = -1$. Tiene orden 8, no es abeliano, y sin embargo *todos* sus subgrupos son normales —cosa que en D_4 no ocurre—. Es el testigo mínimo de que "todos los subgrupos normales" no implica "abeliano", y aparece de verdad en la física: los cuaterniones son la herramienta estándar para representar rotaciones en gráficos por computadora y en control de satélites, justo porque su no conmutatividad copia la de las rotaciones del espacio.

23.14. Productos directos: fabricar grupos nuevos

Idea clave

Intuición primero. El producto directo es poner dos grupos a funcionar *en paralelo y sin hablarse*: la primera coordenada evoluciona con su regla, la segunda con la suya, y nadie interfiere. Como dos relojes independientes en la misma pared.

De ahí sale la única fórmula que hay que recordar: el orden de un par es el mínimo común múltiplo de los órdenes, porque el par vuelve al inicio cuando *ambos* relojes coinciden en su salida. Y de ahí sale también por qué $\mathbb{Z}_m \times \mathbb{Z}_n \cong \mathbb{Z}_{mn}$ exactamente cuando $\gcd(m, n) = 1$: si son coprimos, el mcm es el producto, hay un elemento que recorre todo y el grupo es cíclico. Ese enunciado es, palabra por palabra, el teorema chino del resto del capítulo 3 traducido al lenguaje de grupos.

Regla: producto directo

$G \times H = \{(g,h)\}$ con la operación componente a componente. Entonces:

- $|G \times H| = |G|\,|H|$;
- es abeliano si y solo si ambos lo son;
- $\operatorname{ord}(g,h) = \operatorname{mcm}(\operatorname{ord} g, \operatorname{ord} h)$;
- $\mathbb{Z}_m \times \mathbb{Z}_n \cong \mathbb{Z}_{mn}$ **si y solo si** $\gcd(m,n) = 1$.

Ejemplo 23.79 (Clasificar los abelianos de orden 8)**.** Por el teorema de estructura, todo abeliano finito es producto de cíclicos de orden potencia de primo. Para $8 = 2^3$ las particiones de 3 son 3, $2+1$, $1+1+1$:

$$\mathbb{Z}_8, \qquad \mathbb{Z}_4 \times \mathbb{Z}_2, \qquad \mathbb{Z}_2 \times \mathbb{Z}_2 \times \mathbb{Z}_2.$$

Tres grupos, y se distinguen por los órdenes de sus elementos: el primero tiene elementos de orden 8, el segundo llega hasta 4, el tercero solo hasta 2. Añadiendo los dos no abelianos (D_4 y Q_8) se obtienen los cinco de la tabla.

23.15. Dónde aparecen los grupos de verdad

- **Cubo de Rubik.** Sus movimientos forman un grupo de orden 43 252 003 274 489 856 000. Los "algoritmos" que memorizan los speedcubers son conmutadores $aba^{-1}b^{-1}$: precisamente la medida de cuánto falla la conmutatividad.
- **Cristalografía.** Las simetrías de un cristal en el espacio forman uno de exactamente **230** grupos espaciales. Ese número es un teorema de teoría de grupos, y es el que organiza la difracción de rayos X.
- **Química cuántica.** Las moléculas se clasifican por su grupo puntual (C_{2v} para el agua, D_{6h} para el benceno), y la tabla de caracteres de ese grupo predice qué vibraciones son visibles en el infrarrojo.
- **Códigos correctores.** Los códigos lineales son subgrupos de $\mathbb{F}_2^n$; la corrección de errores de un CD o de una sonda espacial es aritmética de clases laterales.
- **Criptografía.** RSA vive en $(\mathbb{Z}_n)^\times$ y la criptografía de curva elíptica, en el grupo de puntos de una curva. El teorema de Lagrange es lo que hace funcionar el sistema (capítulo 3).
- **Física de partículas.** $SU(3) \times SU(2) \times U(1)$ *es* el modelo estándar; las partículas son las representaciones irreducibles de ese grupo.
- **Y el motivo original.** La solubilidad de S_n decide si una ecuación tiene fórmula con radicales. Eso es el capítulo 26.

23.16. Taller de ejemplos resueltos

Ejercicio 23.1. Demuestra que si todo elemento de G cumple $x^2 = e$, entonces G es abeliano.

Solución. **Paso 1.** La hipótesis dice que cada elemento es su propio inverso: $x^{-1} = x$.

Paso 2. Toma $a, b \in G$. Entonces ab también cumple $(ab)^2 = e$, o sea $(ab)^{-1} = ab$.

Paso 3. Pero siempre $(ab)^{-1} = b^{-1}a^{-1} = ba$ (por el paso 1).

Paso 4. Igualando: $ab = ba$. Ejemplo mínimo: el grupo de Klein. ■

Ejercicio 23.2. Si $|G|$ es par, demuestra que G tiene al menos un elemento de orden 2.

Solución. **Paso 1.** Empareja cada $x \neq e$ con su inverso x^{-1}.

Paso 2. Si ningún elemento fuera su propio inverso, los $|G|-1$ elementos distintos de e se agruparían en parejas: $|G| - 1$ sería par, luego $|G|$ impar.

Paso 3. Como $|G|$ es par, hay algún $x \neq e$ con $x = x^{-1}$, es decir $x^2 = e$: orden 2. (Este argumento de emparejamiento es el mismo que demuestra el teorema de Wilson en el capítulo 3.) ■

Ejercicio 23.3. Encuentra todos los subgrupos de $\mathbb{Z}_{12}$ y dibuja cómo se contienen.

Solución. **Paso 1.** $\mathbb{Z}_{12}$ es cíclico, así que tiene exactamente un subgrupo por cada divisor de 12: $1, 2, 3, 4, 6, 12$. Son seis.

Paso 2. El de orden d está generado por $12/d$:

$$\langle 0\rangle = \{0\}, \quad \langle 6\rangle = \{0, 6\}, \quad \langle 4\rangle = \{0, 4, 8\}, \quad \langle 3\rangle = \{0, 3, 6, 9\},$$

$$\langle 2\rangle = \{0, 2, 4, 6, 8, 10\}, \quad \langle 1\rangle = \mathbb{Z}_{12}.$$

Paso 3. Contenciones: $\langle d\rangle \subseteq \langle d'\rangle \iff d' \mid d$. El retículo es una copia exacta del retículo de divisores de 12:

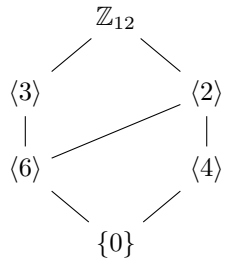


Figura 23.2: Los seis subgrupos de $\mathbb{Z}_{12}$, uno por cada divisor de 12, ordenados por inclusión. El dibujo es el retículo de divisores puesto boca abajo: $\langle d\rangle \subseteq \langle d'\rangle$ exactamente cuando d' divide a d.

El diagrama de subgrupos de un grupo cíclico es literalmente el diagrama de divisibilidad de su orden: aritmética y estructura son la misma cosa. ■

Ejercicio 23.4. Calcula el centro de D_4.

Solución. **Paso 1.** El centro es $Z(G) = \{z : zg = gz \ \forall g\}$.

Paso 2. Descartar las reflexiones. Sea $z = r^i s$ una reflexión y veamos si conmuta con r. Por un lado $zr = r^i(sr) = r^i r^{-1} s = r^{i-1}s$; por otro $rz = r^{i+1}s$. Son iguales solo si $r^{i-1} = r^{i+1}$, es decir $r^2 = e$, falso en D_4 (donde r tiene orden 4). Ninguna reflexión está en el centro.

Paso 3. Rotaciones. r^i conmuta con todas las rotaciones; ¿conmuta con s? $sr^i = r^{-i}s$, y esto es $r^i s$ si y solo si $r^{2i} = e$, es decir $4 \mid 2i$, es decir $i \in \{0, 2\}$.

Conclusión: $Z(D_4) = \{e, r^2\}$, de orden 2. En general $Z(D_n) = \{e\}$ para n impar y $\{e, r^{n/2}\}$ para n par: la media vuelta es la única simetría que conmuta con todo. ■

Ejercicio 23.5. Demuestra que un grupo de orden 4 es siempre abeliano.

Solución. **Paso 1.** Por Lagrange, los órdenes posibles de los elementos son 1, 2 y 4.

Paso 2. Caso A: hay un elemento de orden 4. Entonces genera todo, $G \cong \mathbb{Z}_4$, cíclico y por tanto abeliano.

Paso 3. Caso B: no hay ninguno de orden 4. Entonces todo $x \neq e$ tiene orden 2, o sea $x^2 = e$ para todo x; por el primer ejercicio de este taller, G es abeliano (y es V).

Conclusión: en ambos casos abeliano. El mismo argumento, con la ecuación de clases, prueba el enunciado general: todo grupo de orden p^2 es abeliano. ■

Ejercicio 23.6. En S_5, ¿cuántos elementos hay de orden 6?

Solución. **Paso 1.** El orden es el mcm de las longitudes de los ciclos. Para que dé 6 con 5 símbolos, el único tipo posible es un 3-ciclo junto a una trasposición disjunta: $\operatorname{mcm}(3, 2) = 6$.

Paso 2. Contar. Elegimos los 3 símbolos del 3-ciclo: $\binom{5}{3} = 10$ formas. Con cada terna hay 2 ciclos distintos. Los 2 símbolos restantes forman la trasposición de forma forzada: 1 manera.

Paso 3. $10 \times 2 \times 1 = 20$ elementos de orden 6.

Comprobación de coherencia: $|S_5| = 120$ y 20 divide bien dentro; además estos 20 elementos son impares (un 3-ciclo es par, una trasposición es impar), así que ninguno está en A_5 —coherente con que A_5 no tenga elementos de orden 6, ya que sus órdenes son $1, 2, 3, 5$. ■

Ejercicio 23.7. Sea $\varphi : \mathbb{Z}_{12} \to \mathbb{Z}_{12}$ dada por $\varphi(x) = 3x$. Halla su núcleo, su imagen, y comprueba el primer teorema de isomorfía.

Solución. **Paso 1. Homomorfismo:** $\varphi(x + y) = 3(x + y) = 3x + 3y$. ✓

Paso 2. Núcleo: $3x \equiv 0 \pmod{12} \iff 12 \mid 3x \iff 4 \mid x$. Luego $\ker \varphi = \{0, 4, 8\}$, de orden 3.

Paso 3. Imagen: $\{0, 3, 6, 9\} = \langle 3 \rangle$, de orden 4.

Paso 4. Verificación: $|\mathbb{Z}_{12}|/|\ker| = 12/3 = 4 = |\operatorname{im}|$. ✓ Y en efecto $\mathbb{Z}_{12}/\{0, 4, 8\} \cong \mathbb{Z}_4$: las clases son $\{0, 4, 8\}, \{1, 5, 9\}, \{2, 6, 10\}, \{3, 7, 11\}$, y sumar clases es sumar módulo 4. ■

Ejercicio 23.8. ¿Cuántos collares de 6 cuentas con 3 colores hay, si se permite girarlos pero no darles la vuelta?

Solución. Mismo esquema que el ejemplo de 2 colores, con 3 en lugar de 2:

$$\frac{1}{6}\sum_{k=0}^{5} 3^{\gcd(k,6)} = \frac{3^6 + 3^1 + 3^2 + 3^3 + 3^2 + 3^1}{6} = \frac{729 + 3 + 9 + 27 + 9 + 3}{6} = \frac{780}{6} = 130.$$

130 collares. Compara con los 14 de dos colores: el número crece rapidísimo, y sin Burnside habría que enumerar $3^6 = 729$ pulseras y descartar repeticiones a mano. ■

Ejercicio 23.9. Demuestra que la intersección de dos subgrupos es un subgrupo, y da un ejemplo de que la unión normalmente no lo es.

Solución. **Intersección.** $e \in H \cap K$ porque está en ambos. Si $a, b \in H \cap K$, entonces $ab^{-1} \in H$ (porque H es subgrupo) y $ab^{-1} \in K$ (igual), luego $ab^{-1} \in H \cap K$. Por el criterio, es subgrupo. ✓

Unión. En $\mathbb{Z}$, toma $H = 2\mathbb{Z}$ y $K = 3\mathbb{Z}$. Entonces $2 \in H \cup K$ y $3 \in H \cup K$, pero $2 + 3 = 5$ no es múltiplo de 2 ni de 3: la unión no es cerrada.

Dato fino: $H \cup K$ es subgrupo *si y solo si* uno de los dos contiene al otro. ■

Ejercicio 23.10. Comprueba con D_4 que el conjunto de reflexiones no es un subgrupo, y explica qué es lo que sí lo es.

Solución. **Paso 1.** Las reflexiones son s, rs, r^2s, r^3s. Falta el neutro e: ya ahí falla.

Paso 2. Y el producto de dos reflexiones es una rotación: $(rs)(s) = r$. El conjunto no es cerrado.

Paso 3. Lo que sí es subgrupo es el de las *rotaciones* $\{e, r, r^2, r^3\}$, de índice 2 y por tanto normal, con cociente $\mathbb{Z}_2$. Ese cociente es la respuesta a "¿le di la vuelta al cuadrado o no?", y es la versión geométrica del signo de una permutación. ■

23.17. Cierre: qué llevarte de este capítulo

Si tuvieras que quedarte con cinco frases:

1. Un grupo es el conjunto de las maneras de mover algo sin romperlo, y toda la teoría sale de que esos movimientos se encadenan y se deshacen.
2. El teorema de Lagrange es un embaldosado: los trasladados de un subgrupo cubren el grupo en piezas iguales, y por eso el orden divide.
3. Un subgrupo es *normal* cuando el embaldosado respeta la multiplicación; solo entonces existe el cociente, que es el grupo visto con lentes borrosos.
4. Órbita–estabilizador ($|Gx|\,|G_x| = |G|$) convierte preguntas de conteo imposibles en multiplicaciones de dos números pequeños.
5. Los grupos simples son los átomos, y que A_5 sea simple y no abeliano es, literalmente, la razón por la que las ecuaciones de quinto grado no tienen fórmula.

Con esto, el capítulo 25 debería leerse como una segunda vuelta más rápida sobre terreno conocido, y el capítulo 26 como el desenlace de la historia que aquí empezó con un triángulo de cartón.

23.18. Las tres piezas que faltaban: grupos cíclicos, homomorfismos y clases de conjugación

Hemos usado tres cosas continuamente sin haberlas establecido: que los subgrupos de un cíclico son uno por divisor, que el núcleo de un homomorfismo es normal, y que las clases de conjugación de S_n se cuentan con una fórmula. Van las tres.

Grupos cíclicos

Teorema 23.80 (Clasificación)**.** *Sea $G = \langle a \rangle$. Si a tiene orden infinito, $G \cong \mathbb{Z}$; si tiene orden n, $G \cong \mathbb{Z}_n$. Además:*

1. *Todo subgrupo de un cíclico es cíclico.*
2. *Si $|G| = n$, hay* exactamente un *subgrupo de orden d por cada divisor $d \mid n$, a saber $\langle a^{n/d} \rangle$.*

3. $\langle a^k\rangle = \langle a^{\gcd(k,n)}\rangle$*, y* a^k *genera* G *si y solo si* $\gcd(k,n)=1$*: hay* $\varphi(n)$ *generadores.*

Demostración. Para (1), sea $H \le \langle a\rangle$ no trivial y m el menor positivo con $a^m \in H$; dividiendo cualquier otro exponente entre m, el resto obliga a que sea nulo, luego $H = \langle a^m\rangle$. Para (2), a^k tiene orden $n/\gcd(k,n)$, lo que da (3) y fuerza la unicidad de (2). □

Ejemplo 23.81 ($\mathbb{Z}_{12}$, entero)**.** Los seis subgrupos son $\langle d\rangle$ con $d \in \{1,2,3,4,6,12\}$, de órdenes $12/d$. Y $\langle 8\rangle = \langle 4\rangle$ porque $\gcd(8,12)=4$. Los generadores son $1,5,7,11$: exactamente $\varphi(12)=4$. El retículo de subgrupos es el de divisores de 12 puesto boca abajo, tal como afirmaba el capítulo.

Homomorfismos y el primer teorema de isomorfía

Definición 23.82. $\varphi : G \to H$ es **homomorfismo** si $\varphi(ab)=\varphi(a)\varphi(b)$. Entonces $\varphi(e)=e$, $\varphi(a^{-1})=\varphi(a)^{-1}$, la imagen es subgrupo de H y el **núcleo** $\ker\varphi = \{g : \varphi(g)=e\}$ es subgrupo de G.

Teorema 23.83. $\ker\varphi$ *es* normal*, y* φ *es inyectivo si y solo si* $\ker\varphi=\{e\}$.

Demostración. Si $k\in\ker\varphi$ y $g\in G$, entonces $\varphi(gkg^{-1})=\varphi(g)e\varphi(g)^{-1}=e$. Para la segunda parte, $\varphi(a)=\varphi(b) \iff \varphi(ab^{-1})=e \iff ab^{-1}\in\ker\varphi$. □

Ahí está la motivación conceptual de la normalidad, que hasta ahora era una definición caída del cielo: *los subgrupos normales son exactamente los núcleos.* Uno es normal si y solo si es el núcleo de algún homomorfismo, a saber, la proyección al cociente.

Teorema 23.84 (Primer teorema de isomorfía)**.**

$$G/\ker\varphi \;\cong\; \operatorname{im}\varphi.$$

Demostración. Defínase $\bar\varphi(g\ker\varphi)=\varphi(g)$. Está *bien definida*: si $gK=g'K$ entonces $g^{-1}g'\in K$ y $\varphi(g)=\varphi(g')$. Es homomorfismo por serlo φ, sobreyectiva sobre la imagen por construcción, e inyectiva porque $\bar\varphi(gK)=e$ obliga a $g\in K$, es decir, a la clase trivial. □

Ejemplo 23.85. Sea $\varphi:\mathbb{Z}_{12}\to\mathbb{Z}_{12}$ con $\varphi(x)=3x$. Entonces $\ker\varphi=\{0,4,8\}$ y $\operatorname{im}\varphi=\{0,3,6,9\}$, y el teorema dice $\mathbb{Z}_{12}/\{0,4,8\}\cong\{0,3,6,9\}\cong\mathbb{Z}_4$. Comprobación de órdenes: $12/3=4$. ✓

Conjugación, centro y ecuación de clases

Definición 23.86. $a\sim b$ si $b=gag^{-1}$ para algún g. Las clases de esa relación son las **clases de conjugación**. El **centralizador** es $C(a)=\{g : ga=ag\}$ y el **centro** es $Z(G)=\bigcap_a C(a)$.

Teorema 23.87 (Ecuación de clases)**.** *Como la conjugación es una acción de G sobre sí mismo, órbita-estabilizador da $|\text{clase}(a)| = [G : C(a)]$, y separando las clases unitarias —que son exactamente los elementos del centro—:*

$$|G| = |Z(G)| + \sum_i [G : C(g_i)],$$

donde los g_i recorren representantes de las clases con más de un elemento.

Corolario 23.88. *Todo p-grupo tiene centro no trivial. En efecto, cada término de la suma es una potencia positiva de p, y $|G|$ también, luego p divide a $|Z(G)|$, que no puede ser* 1*.*

Ejemplo 23.89 (Dos ecuaciones de clases)**. (a)** S_4**.** El tamaño de la clase del tipo $1^{m_1}2^{m_2}\cdots$ es $n!/\prod_k k^{m_k}m_k!$ —el denominador cuenta las simetrías de la notación cíclica— y da

$$24 = \underbrace{1}_{1^4} + \underbrace{6}_{2\cdot 1^2} + \underbrace{3}_{2^2} + \underbrace{8}_{3\cdot 1} + \underbrace{6}_{4}.$$

El centro es trivial: solo la clase de tamaño 1.

(b) D_4**.** $8 = 2 + 2 + 2 + 2$, con $Z(D_4) = \{e, r^2\}$ de tamaño 2 y tres clases de tamaño 2. Coherente con el corolario: al ser un 2-grupo, el centro no podía ser trivial.

Teorema 23.90 (Cayley)**.** *Todo grupo se sumerge en un grupo de permutaciones: $G \hookrightarrow S_{|G|}$.*

Demostración. La acción por traslación izquierda $g \cdot x = gx$ da un homomorfismo $G \to S_G$. Su núcleo son los g con $gx = x$ para todo x, es decir, solo e. Por el primer teorema de isomorfía, $G \cong$ su imagen. □

Teorema 23.91 (Cauchy)**.** *Si p divide a $|G|$, hay un elemento de orden p.*

Argumento de McKay. Considérese

$$X = \{(g_1, \ldots, g_p) \in G^p : g_1 g_2 \cdots g_p = e\},$$

que tiene $|G|^{p-1}$ elementos, porque las $p-1$ primeras coordenadas son libres y la última queda determinada. El grupo $\mathbb{Z}_p$ actúa rotando cíclicamente —si el producto es e, cualquier rotación cíclica también lo es—, y cada órbita tiene tamaño 1 o p. Las de tamaño 1 son las tuplas constantes $(g, \ldots, g)$ con $g^p = e$. Como p divide a $|X|$ y a la suma de las órbitas grandes, divide al número de órbitas unitarias; y hay al menos una, la de $g = e$. Luego hay al menos p, y en particular algún $g \neq e$ con $g^p = e$. □

Ejemplo 23.92. Cayley para $\mathbb{Z}_3$: $0 \mapsto e$, $1 \mapsto (1\,2\,3)$, $2 \mapsto (1\,3\,2)$, es decir $\mathbb{Z}_3 \hookrightarrow S_3$. Y Cauchy en A_4, de orden 12: garantiza elementos de orden 2 y de orden 3, y en efecto los hay; pero *no* garantiza uno de orden 6, porque 6 no es primo, y de hecho A_4 no tiene ninguno —ni subgrupo de orden 6, como vimos—.

Teorema 23.93 (Clasificación de los abelianos finitos)**.** *Todo grupo abeliano finito es producto directo de cíclicos de orden potencia de primo, de manera única salvo orden (*divisores elementales*); equivalentemente, es $\mathbb{Z}_{d_1} \times \cdots \times \mathbb{Z}_{d_k}$ con $d_1 \mid d_2 \mid \cdots \mid d_k$ (*factores invariantes*).*

Ejemplo 23.94 (Contar grupos abelianos)**.** Para orden $8 = 2^3$, las particiones de 3 son 3, $2+1$ y $1+1+1$: tres grupos, $\mathbb{Z}_8$, $\mathbb{Z}_4 \times \mathbb{Z}_2$ y $\mathbb{Z}_2^3$.

Para orden $36 = 2^2 \cdot 3^2$, hay 2 particiones de 2 para cada primo, luego $2 \times 2 = 4$ grupos. En factores invariantes:

$$\mathbb{Z}_{36}, \qquad \mathbb{Z}_2 \times \mathbb{Z}_{18}, \qquad \mathbb{Z}_3 \times \mathbb{Z}_{12}, \qquad \mathbb{Z}_6 \times \mathbb{Z}_6.$$

El número de grupos abelianos de orden $n = \prod p_i^{a_i}$ es $\prod P(a_i)$ con P la función de particiones, y no depende de cuáles sean los primos: solo de los exponentes.

23.19. Los teoremas de Sylow, enunciados y demostrados

Los ejercicios del capítulo piden "demostrar con Sylow" sin que en ningún sitio se haya dicho qué afirma Sylow. Se remedia aquí, porque son la herramienta con la que se clasifican de verdad los grupos finitos pequeños.

Definición 23.95 (p-subgrupo de Sylow)**.** Sea G finito con $|G| = p^k m$, donde p es primo y $p \nmid m$. Un *p-subgrupo de Sylow* de G es un subgrupo de orden exactamente p^k: el mayor orden de p-grupo que el teorema de Lagrange permite.

Teorema 23.96 (Sylow)**.** *Sea $|G| = p^k m$ con $p \nmid m$, y sea n_p el número de p-subgrupos de Sylow. Entonces:*

1. (Existencia.) $n_p \geq 1$*: siempre hay al menos un p-subgrupo de Sylow.*
2. (Conjugación.) *Dos p-subgrupos de Sylow cualesquiera son conjugados, y todo p-subgrupo está contenido en alguno de ellos.*
3. (Recuento.) $n_p \equiv 1 \pmod p$ *y* $n_p \mid m$.

Demostración. (1) Existencia, por el argumento de Wielandt. Sea Ω la familia de *todos* los subconjuntos de G de cardinal p^k, sobre la que G actúa por traslación a izquierda, $g \cdot S = gS$. Su cardinal es

$$|\Omega| = \binom{p^k m}{p^k},$$

y un cálculo con el teorema de Kummer —o simplemente observando que en $\binom{p^k m}{p^k} = \prod_{j=0}^{p^k-1} \frac{p^k m - j}{p^k - j}$ cada factor tiene la misma potencia de p arriba y abajo, pues para $0 < j < p^k$ la mayor potencia de p que divide a $p^k m - j$ es la que divide a j— muestra que $p \nmid |\Omega|$ salvo el factor m; con precisión, $v_p(|\Omega|) = v_p(m) = 0$. Como las órbitas parten Ω y su cardinal total no es divisible por p, alguna órbita $\mathcal{O}$ tiene cardinal no divisible por p. Sea $S \in \mathcal{O}$ y $H = \{g \in G : gS = S\}$ su estabilizador. Por el teorema órbita–estabilizador, $|\mathcal{O}| = [G : H]$, y como $p \nmid |\mathcal{O}|$, se tiene $p^k \mid |H|$. Por otro lado, fijado $s \in S$, la aplicación $h \mapsto hs$ es inyectiva de H en S, luego $|H| \leq |S| = p^k$. Ambas cosas fuerzan $|H| = p^k$: H es un p-subgrupo de Sylow.

(2) y (3). Sea P un Sylow (existe por el paso 1), sea $\mathcal{S} = \{P_1, \dots, P_n\}$ el conjunto de los conjugados de P, y hágase actuar P sobre $\mathcal{S}$ por conjugación. Las órbitas tienen cardinal potencia de p, luego

$$n \equiv \#\{\text{puntos fijos}\} \pmod{p}.$$

Un punto fijo es un P_i con $PP_i = P_iP$, es decir con $P \subset N_G(P_i)$; pero entonces P y P_i son ambos Sylow del grupo $N_G(P_i)$, donde P_i es normal, y el paso siguiente muestra que eso obliga a $P = P_i$. En efecto, si Q es un p-subgrupo de $N_G(P_i)$ con $P_i \trianglelefteq N_G(P_i)$, entonces QP_i es un subgrupo de orden $|Q||P_i|/|Q \cap P_i|$, potencia de p y $\geq |P_i| = p^k$: por maximalidad $QP_i = P_i$, o sea $Q \subset P_i$. Así el único punto fijo es P mismo, y $n \equiv 1 \pmod{p}$.

Falta ver que $\mathcal{S}$ contiene a *todos* los Sylow. Sea Q cualquier p-subgrupo y hágasele actuar sobre $\mathcal{S}$. Como $|\mathcal{S}| = n \equiv 1 \pmod{p}$, no todas las órbitas pueden tener cardinal divisible por p: hay un punto fijo P_i, y por el argumento anterior $Q \subset P_i$. Si además Q es Sylow, $|Q| = |P_i|$ fuerza $Q = P_i \in \mathcal{S}$. Esto prueba (2), y de paso que $n_p = n$. Finalmente, $n = [G : N_G(P)]$ divide a $[G : P] = m$, que es (3). □

Regla: Usar Sylow para clasificar grupos de orden pequeño

1. Factorizar $|G| = p_1^{k_1} \cdots p_r^{k_r}$.
2. Para cada primo, listar los valores de n_p compatibles con $n_p \equiv 1 \pmod{p}$ y $n_p \mid m$.
3. Si para algún primo la única posibilidad es $n_p = 1$, ese Sylow es *normal* (es su propio único conjugado), y el grupo no es simple.
4. Contar elementos: n_p Sylow de orden p aportan $n_p(p-1)$ elementos de orden p, distintos entre sí porque dos Sylow de orden primo se cortan trivialmente. Si las cuentas se pasan de $|G|$, la configuración es imposible.
5. Con un Sylow normal N y otro H con $N \cap H = \{e\}$ y $NH = G$, se tiene $G \cong N \rtimes H$: producto semidirecto, y sólo queda clasificar los homomorfismos $H \to \mathrm{Aut}(N)$.

Ejemplo 23.97 (Ningún grupo de orden 15 es no cíclico)**.** Sea $|G| = 15 = 3 \cdot 5$. Para $p = 5$: $n_5 \equiv 1 \pmod{5}$ y $n_5 \mid 3$, luego $n_5 = 1$. Para $p = 3$: $n_3 \equiv 1 \pmod{3}$ y $n_3 \mid 5$, y las opciones $1, 5$ dejan sólo $n_3 = 1$ (pues $5 \not\equiv 1 \bmod 3$). Ambos Sylow, P_5 y P_3, son normales, se cortan trivialmente por Lagrange, y su producto tiene 15 elementos: $G \cong \mathbb{Z}_3 \times \mathbb{Z}_5 \cong \mathbb{Z}_{15}$. Con exactamente el mismo argumento se prueba que todo grupo de orden pq con $p < q$ y $p \nmid q - 1$ es cíclico.

Ejemplo 23.98 (Ningún grupo de orden 30 es simple)**.** $|G| = 30 = 2 \cdot 3 \cdot 5$. Se tiene $n_5 \in \{1, 6\}$ y $n_3 \in \{1, 10\}$. Si fuesen $n_5 = 6$ y $n_3 = 10$, habría $6 \cdot 4 = 24$ elementos de orden 5 y $10 \cdot 2 = 20$ de orden 3, en total $44 > 30$: imposible. Luego alguno vale 1 y G tiene un subgrupo normal propio. El recuento de elementos es la técnica que hay que interiorizar; casi todos los ejercicios de no simplicidad salen así.

23.20. Acción, simple y resoluble: los términos que los ejercicios usan

Tres términos aparecen en los ejercicios propuestos de este capítulo y están definidos en el capítulo 25, que viene *después*. Un estudiante que llegue aquí no tiene dónde mirar, así que se dejan escritos.

Definición 23.99 (Acción de un grupo)**.** Una *acción* (por la izquierda) de un grupo G sobre un conjunto X es una aplicación $G \times X \to X$, escrita $(g, x) \mapsto g \cdot x$, que cumple

$$e \cdot x = x \qquad \text{y} \qquad (gh) \cdot x = g \cdot (h \cdot x)$$

para todos $g, h \in G$ y $x \in X$. Equivalentemente, es un homomorfismo $G \to \mathrm{Sym}(X)$ al grupo de biyecciones de X.

Dado $x \in X$, su *órbita* es $G{\cdot}x = \{g{\cdot}x : g \in G\}$ y su *estabilizador* es $G_x = \{g \in G : g{\cdot}x = x\}$, que es un subgrupo.

Ejemplo 23.100 (Las tres acciones que se usan en las demostraciones)**.** Comprobar los dos axiomas es siempre inmediato, y conviene hacerlo una vez para cada una:

- *Traslación a izquierda:* G actúa sobre sí mismo por $g \cdot x = gx$. Es la que usa el teorema de Cayley y la que aparece en la demostración de Sylow por el argumento de Wielandt.
- *Conjugación:* G actúa sobre sí mismo por $g \cdot x = gxg^{-1}$. Las órbitas son las clases de conjugación y los estabilizadores, los centralizadores; de ahí sale la ecuación de clases.
- *Sobre las clases laterales:* G actúa sobre G/H por $g \cdot (aH) = (ga)H$. Es la que demuestra que un subgrupo de índice p mínimo es normal.

En los tres casos, $e \cdot x = x$ es evidente y la asociatividad de la acción es la asociatividad del producto en G.

Definición 23.101 (Grupo simple)**.** Un grupo G es *simple* si $G \neq \{e\}$ y sus únicos subgrupos normales son $\{e\}$ y el propio G.

Observación 23.102 (Por qué importa)*.* Los grupos simples son los "átomos" de la teoría: toda serie de composición de un grupo finito tiene factores simples, y por el teorema de Jordan–Hölder esa lista de factores es única salvo el orden. Los simples abelianos son exactamente los $\mathbb{Z}/p$ con p primo. Los no abelianos empiezan en A_5, de orden 60, y su clasificación completa —terminada en 2004 y repartida en más de diez mil páginas— es uno de los mayores esfuerzos colectivos de la historia de las matemáticas.

Definición 23.103 (Conmutador, serie derivada y grupo resoluble)**.** El *conmutador* de dos elementos es $[a, b] = aba^{-1}b^{-1}$, y el *subgrupo derivado* $[G, G]$ —también escrito G'— es el generado por todos los conmutadores. Es normal, y el cociente $G/[G, G]$ es el mayor cociente abeliano de G: el *abelianizado*.

La *serie derivada* es

$$G^{(0)} = G, \qquad G^{(n+1)} = \big[G^{(n)}, G^{(n)}\big].$$

El grupo G es *resoluble* si esa serie llega a la identidad en un número finito de pasos: $G^{(k)} = \{e\}$ para algún k.

Ejemplo 23.104 (Las series derivadas que hay que conocer)**.**

- Todo grupo abeliano tiene $G' = \{e\}$: resoluble en un paso.
- S_3: su derivado es $A_3 = \mathbb{Z}/3$, cuyo derivado es trivial. Serie $S_3 \supset A_3 \supset \{e\}$: resoluble.
- S_4: $S_4' = A_4$, $A_4' = V_4$ (el grupo de Klein), $V_4' = \{e\}$. Resoluble en tres pasos, y por eso la ecuación de cuarto grado tiene fórmula.
- S_5: $S_5' = A_5$, y $A_5' = A_5$ porque A_5 es simple no abeliano. La serie se estanca y *no* es resoluble. Ése es exactamente el motivo por el que la quíntica general no se resuelve por radicales, y la conexión se demuestra en el capítulo 26.

Un grupo simple no abeliano nunca es resoluble, porque su derivado —siendo normal y distinto del trivial, ya que el grupo no es abeliano— tiene que ser él mismo.

23.21. Ejercicios resueltos

Ejercicio 23.11. Escribe la ecuación de clases de S_4 y úsala para localizar todos sus subgrupos normales.

Solución. En S_n dos permutaciones son conjugadas exactamente cuando tienen el mismo tipo de ciclos, así que las clases se leen de las particiones de 4. El tamaño de la clase del tipo $1^{m_1}2^{m_2}\cdots$ es $n!/\prod_k k^{m_k} m_k!$:

tipo	1^4	$2\cdot 1^2$	2^2	$3\cdot 1$	4
tamaño	1	6	3	8	6

$$1+6+3+8+6 = 24.$$

Un subgrupo normal es unión de clases, contiene al neutro y su orden divide a 24. Las sumas que empiezan por 1 y dan un divisor de 24 son 1, $1+3=4$, $1+3+8=12$ y 24. Cada una corresponde a un subgrupo de verdad: $\{e\}$, el grupo de Klein $V = \{e, (12)(34), (13)(24), (14)(23)\}$, $A_4 = \{e\} \cup 2^2 \cup 3\cdot 1$ y S_4. Y ninguna otra combinación funciona: $1+6=7$ no divide a 24, $1+8=9$ tampoco, $1+6+3=10$ tampoco. Conclusión: los únicos normales son $\{e\} \subset V \subset A_4 \subset S_4$, que es justo la serie que hace resoluble a S_4.

Obsérvese la potencia del método: hemos clasificado los subgrupos normales sin examinar ni un solo subgrupo, solo sumando enteros. ∎

Ejercicio 23.12. ¿De cuántas maneras esencialmente distintas se pueden pintar las seis caras de un cubo con tres colores, si dos coloraciones que difieren en una rotación cuentan como la misma?

Solución. Es el lema de Burnside: el número de órbitas es el promedio de los puntos fijos. El grupo de rotaciones del cubo tiene 24 elementos repartidos en cinco tipos, y para cada uno contamos en

cuántos ciclos parte las seis caras (una coloración es fija si y solo si es constante en cada ciclo, luego hay $3^{\#\text{ciclos}}$ fijas):

tipo de rotación	cuántas	ciclos de caras	fijas
identidad	1	6	$3^6 = 729$
caras $\pm 90°$	6	3	$3^3 = 27$
caras $180°$	3	4	$3^4 = 81$
vértices $\pm 120°$	8	2	$3^2 = 9$
aristas $180°$	6	3	$3^3 = 27$

El total de pares (rotación, coloración fija) es $729+6\cdot 27+3\cdot 81+8\cdot 9+6\cdot 27 = 729+162+243+72+162 = 1368$, y

$$\text{órbitas} = \frac{1368}{24} = 57.$$

Con dos colores el mismo cálculo da $(64 + 6\cdot 8 + 3\cdot 16 + 8\cdot 4 + 6\cdot 8)/24 = 240/24 = 10$, y ese sí se puede comprobar a mano enumerando: $0, 1, 2, \ldots, 6$ caras negras dan $1, 1, 2, 2, 2, 1, 1$ configuraciones, y en efecto suman 10. Esa comprobación independiente es la que da confianza en el 57. ■

Ejercicio 23.13. Demuestra que ningún grupo de orden 12 es simple, y encuentra el orden máximo de un elemento de S_7.

Solución. **(a) Orden 12.** Sea $|G| = 12 = 2^2\cdot 3$. Por Sylow, $n_3 \equiv 1 \pmod 3$ y $n_3 \mid 4$, luego $n_3 \in \{1, 4\}$. Si $n_3 = 1$ el 3-Sylow es normal y hemos terminado. Si $n_3 = 4$, los cuatro subgrupos de orden 3 se cortan solo en el neutro, luego aportan $4\cdot 2 = 8$ elementos de orden 3. Quedan $12 - 8 = 4$ elementos, y un 2-Sylow tiene exactamente 4: no cabe más que uno, luego $n_2 = 1$ y ese es normal. En ambos casos hay un normal propio.

(b) Orden máximo en S_7. El orden de una permutación es el mínimo común múltiplo de las longitudes de sus ciclos, así que hay que maximizar mcm sobre las particiones de 7. Las candidatas serias son $7 \to 7$, $6+1 \to 6$, $5+2 \to 10$ y $4+3 \to 12$; el resto queda por debajo. El máximo es $\boxed{12}$, alcanzado solo por el tipo $4+3$, por ejemplo $(1\,2\,3\,4)(5\,6\,7)$.

Curiosidad que conviene retener: el orden máximo *no* crece de forma monótona con la estructura obvia. En S_7 vale 12, y no 7, pese a que 7 es el mayor ciclo posible; la función que da este máximo es la función de Landau. ■

Ejercicio 23.14. Determina la estructura de $(\mathbb{Z}/16\mathbb{Z})^\times$ y compárala con $(\mathbb{Z}/15\mathbb{Z})^\times$.

Solución. $(\mathbb{Z}/16\mathbb{Z})^\times = \{1, 3, 5, 7, 9, 11, 13, 15\}$ tiene $\varphi(16) = 8$ elementos. Calculamos los órdenes:

a	1	3	5	7	9	11	13	15
$\operatorname{ord}(a)$	1	4	4	2	2	4	4	2

No hay ningún elemento de orden 8, luego no es cíclico. Hay exactamente tres de orden 2 (7, 9 y 15), lo que descarta $\mathbb{Z}/2\times\mathbb{Z}/2\times\mathbb{Z}/2$ (ese tendría siete) y deja $\mathbb{Z}/4\times\mathbb{Z}/2$, con generadores 3 (de orden 4) y $15 = -1$ (de orden 2): en efecto $-1 \notin \langle 3\rangle = \{1, 3, 9, 11\}$.

$(\mathbb{Z}/15\mathbb{Z})^\times$ tiene también orden $\varphi(15) = 8$ y, por el teorema chino, es $(\mathbb{Z}/3)^\times \times (\mathbb{Z}/5)^\times \cong \mathbb{Z}/2\times\mathbb{Z}/4$: los dos grupos son isomorfos. Es el ejemplo más pequeño de dos módulos distintos con el mismo grupo de unidades, y explica por qué 2^n con $n \geq 3$ nunca tiene raíz primitiva. ■

Ejercicio 23.15. Demuestra que A_4 no tiene ningún subgrupo de orden 6, y explica por qué eso no contradice el teorema de Lagrange.

Solución. Supongamos que $H \leq A_4$ con $|H| = 6$. El índice es $[A_4 : H] = 2$, y todo subgrupo de índice 2 es normal (las clases por la izquierda y por la derecha coinciden porque solo hay dos y una es H). Entonces A_4/H tiene orden 2, luego $g^2 \in H$ para todo $g \in A_4$.

Ahora bien, A_4 tiene ocho 3-ciclos, y si g es un 3-ciclo entonces $g = g^4 = (g^2)^2$ también es un cuadrado. Así que los ocho 3-ciclos estarían en H, que solo tiene seis elementos. Contradicción.

Esto no contradice a Lagrange: Lagrange dice que el orden de un subgrupo divide al del grupo, no que exista un subgrupo por cada divisor. El recíproco de Lagrange es falso, y A_4 con $6 \mid 12$ es el contraejemplo mínimo. Lo que sí es cierto siempre es el teorema de Cauchy (hay elementos de orden p para cada primo p que divide) y el de Sylow (hay subgrupos del orden de la máxima potencia de cada primo). ■

Ejercicio 23.16. Comprueba que $D_4/Z(D_4)$ es el grupo de Klein, y deduce que D_4 no es abeliano pero sí nilpotente.

Solución. Escribimos $D_4 = \langle r, s \mid r^4 = s^2 = e,\ srs = r^{-1}\rangle$, con r el giro de 90°. El centro es $Z = \{e, r^2\}$: el giro de 180° conmuta con todo (conmuta con r obviamente, y $sr^2s = r^{-2} = r^2$), mientras que r no conmuta con s y s no conmuta con r.

El cociente tiene orden $8/2 = 4$, y en él las clases de r y de s cumplen $\bar{r}^2 = \overline{r^2} = \bar{e}$ y $\bar{s}^{\,2} = \bar{e}$: los cuatro elementos $\bar{e}, \bar{r}, \bar{s}, \overline{rs}$ tienen todos orden ≤ 2. Un grupo de orden 4 sin elementos de orden 4 es el de Klein, luego $D_4/Z \cong \mathbb{Z}/2 \times \mathbb{Z}/2$.

De paso queda un aviso: el cociente por el centro es abeliano, pero eso *no* hace abeliano al grupo. Lo que sí se sigue es que la serie central $\{e\} \subset Z \subset D_4$ termina, es decir, D_4 es nilpotente de clase 2; y hay un teorema clásico que dice que si $G/Z(G)$ es *cíclico* entonces G sí es abeliano. Aquí el cociente es de Klein, no cíclico, y por eso no aplica. ■

Ejercicio 23.17. En S_5, ¿cuántos elementos tiene la clase de conjugación de un 5-ciclo? Úsalo para calcular el centralizador de $(1\,2\,3\,4\,5)$.

Solución. Hay $5!/5 = 24$ ciclos de longitud 5 (cada ciclo se puede escribir de 5 maneras rotando el punto de partida). Como todos los 5-ciclos son conjugados entre sí, la clase tiene 24 elementos.

La ecuación órbita-estabilizador aplicada a la acción por conjugación dice

$$|\text{clase}| = \frac{|S_5|}{|C(g)|} \quad \Longrightarrow \quad |C((1\,2\,3\,4\,5))| = \frac{120}{24} = 5.$$

Un centralizador de orden 5 que contiene a las cinco potencias del propio ciclo tiene que ser exactamente $\langle(1\,2\,3\,4\,5)\rangle$. Dicho de otro modo: lo único que conmuta con un 5-ciclo en S_5 son sus propias potencias.

Este cálculo es el motor de la demostración de que A_5 es simple: los 5-ciclos se parten en A_5 en dos clases de 12, y $1 + 12 + 12 + 15 + 20 = 60$ no admite ninguna suma parcial que contenga al 1 y divida a 60 salvo 1 y 60. ■

23.22. Ejercicios propuestos

Sin solución desarrollada: solo enunciado y, al final, la respuesta. Aquí muchas respuestas son un nombre ("el grupo de Klein", "sí, por Eisenstein") más que un número: comprueba que coincida el fondo del argumento, no la redacción.

23.1. ¿Cuántos elementos tiene S_4? ¿Y A_4? ¿Y D_4 (simetrías del cuadrado)?

23.2. Escribe $(1\,2\,3)(3\,4)$ como producto de ciclos disjuntos y di si es par o impar.

23.3. ¿Cuál es el orden de la permutación $(1\,2\,3\,4\,5)(6\,7)$ en S_7?

23.4. ¿Cuántos subgrupos tiene $\mathbb{Z}/12\mathbb{Z}$? Enuméralos.

23.5. ¿Es $\mathbb{Z}/2\mathbb{Z} \times \mathbb{Z}/3\mathbb{Z}$ isomorfo a $\mathbb{Z}/6\mathbb{Z}$? ¿Y $\mathbb{Z}/2\mathbb{Z} \times \mathbb{Z}/2\mathbb{Z}$ a $\mathbb{Z}/4\mathbb{Z}$?

23.6. ¿Cuántos elementos de orden 2 tiene D_4?

23.7. Sea G de orden 15. Demuestra con Sylow que G es cíclico.

23.8. Para $|G| = 12$, ¿qué valores puede tomar n_3 (número de 3-Sylow)?

23.9. ¿Es A_4 simple? ¿Y A_5? ¿Cuál es el subgrupo normal propio de A_4?

23.10. ¿Es S_4 resoluble? ¿Y S_5? Da la serie derivada de S_4.

23.11. Calcula el orden de $(\mathbb{Z}/15\mathbb{Z})^\times$ y di a qué grupo abeliano conocido es isomorfo.

23.12. Clasifica todos los grupos abelianos de orden 8 salvo isomorfismo.

Respuestas

1. $|S_4| = 24$, $|A_4| = 12$, $|D_4| = 8$. **2.** $(1\,2\,3)(3\,4) = (1\,2\,3\,4)$ leyendo de derecha a izquierda; es un 4-ciclo, luego impar. **3.** $\operatorname{mcm}(5, 2) = 10$. **4.** Seis, uno por cada divisor de 12: los generados por $0, 6, 4, 3, 2, 1$, es decir de órdenes $1, 2, 3, 4, 6, 12$. **5.** Sí lo primero ($\gcd(2, 3) = 1$); no lo segundo: el de Klein no tiene elementos de orden 4. **6.** Cinco: las dos reflexiones en diagonales, las dos en mediatrices y el giro de 180°. **7.** $n_5 \equiv 1\ (5)$ y $n_5 \mid 3$ fuerzan $n_5 = 1$; igual $n_3 = 1$. Ambos Sylow son normales, se intersecan trivialmente, luego $G \cong \mathbb{Z}/5\mathbb{Z} \times \mathbb{Z}/3\mathbb{Z} \cong \mathbb{Z}/15\mathbb{Z}$. **8.** $n_3 \equiv 1$ (mód 3) y $n_3 \mid 4$: solo $n_3 \in \{1, 4\}$. **9.** A_4 no es simple: contiene el grupo de Klein $V = \{e, (12)(34), (13)(24), (14)(23)\}$ como normal. A_5 sí es simple. **10.** S_4 sí: $S_4 \rhd A_4 \rhd V \rhd \{e\}$ con cocientes $\mathbb{Z}/2, \mathbb{Z}/3, \mathbb{Z}/2 \times \mathbb{Z}/2$. S_5 no, porque A_5 es simple no abeliano. **11.** Orden $\varphi(15) = 8$; es isomorfo a $\mathbb{Z}/2\mathbb{Z} \times \mathbb{Z}/4\mathbb{Z}$ (no es cíclico). **12.** Tres: $\mathbb{Z}/8\mathbb{Z}$, $\mathbb{Z}/4\mathbb{Z} \times \mathbb{Z}/2\mathbb{Z}$ y $\mathbb{Z}/2\mathbb{Z} \times \mathbb{Z}/2\mathbb{Z} \times \mathbb{Z}/2\mathbb{Z}$.

Capítulo 24

Anillos, cuerpos y polinomios desde cero

Lo que necesitas traer

Del capítulo 23: grupos, subgrupos, subgrupos normales, cocientes y homomorfismos. Del capítulo 3: la división euclídea y la factorización única en $\mathbb{Z}$, que es el modelo del que se copia toda la teoría. Del capítulo 4: los complejos como ejemplo de cuerpo.

El capítulo 23 construyó la teoría de grupos empezando por un triángulo de cartón. Este hace lo mismo con la estructura que sigue en la escalera: el **anillo**. La diferencia de partida es de una sola palabra. Un grupo tiene *una* operación; un anillo tiene *dos*, y están puestas a imitar el comportamiento de los números enteros.

No hace falta nada del capítulo 25 para leer este: aquí se demuestra todo desde el principio y con números delante. Al terminar sabrás construir cuerpos finitos con tus manos, dividir polinomios como se dividen enteros, decidir si un polinomio es irreducible, y entender por qué la factorización única —que en $\mathbb{Z}$ parece obvia— es en realidad un teorema frágil que se rompe en cuanto sales de terreno conocido.

Idea clave

Intuición primero. Toda la teoría de anillos sale de una observación sobre $\mathbb{Z}$: en los enteros puedes sumar, restar y multiplicar sin salirte, pero *no* puedes dividir. La resta siempre se deshace; la multiplicación, casi nunca.

Esa asimetría es la clave. Un anillo es cualquier conjunto donde ocurre exactamente eso, y toda la teoría consiste en preguntar, una y otra vez, qué propiedades de los enteros sobreviven cuando se cambia el conjunto: ¿se puede dividir con resto?, ¿hay factorización única en primos?, ¿puede un producto dar cero sin que ningún factor lo sea? Cada respuesta afirmativa define un tipo de anillo, y la jerarquía completa es la lista ordenada de "cuánto se parece esto a $\mathbb{Z}$".

24.1. El experimento: dos operaciones en vez de una

Antes de definir nada, hagamos cuentas en un sitio pequeño. Tomemos las horas de un reloj de seis horas, es decir $\mathbb{Z}_6 = \{0, 1, 2, 3, 4, 5\}$, con la suma y el producto módulo 6.

La suma. $4 + 5 = 9 = 6 + 3$, así que $4 + 5 = 3$. Nada raro: la suma módulo 6 se comporta exactamente como esperamos. Todo elemento tiene su opuesto ($4 + 2 = 0$, luego $-4 = 2$), y el 0 no molesta a nadie.

El producto. Aquí empiezan las sorpresas. Calculemos la tabla completa:

$\cdot$	0	1	2	3	4	5
0	0	0	0	0	0	0
1	0	1	2	3	4	5
2	0	2	4	0	2	4
3	0	3	0	3	0	3
4	0	4	2	0	4	2
5	0	5	4	3	2	1

Mira la fila del 2 y la columna del 3: $2 \cdot 3 = 6 = 0$. **Dos números distintos de cero cuyo producto es cero.** En $\mathbb{Z}$ eso es imposible, y de hecho lo usamos constantemente: cuando resolvemos $(x - 2)(x + 5) = 0$ concluimos que $x = 2$ o $x = -5$ precisamente porque un producto solo es cero si un factor lo es. En $\mathbb{Z}_6$ ese razonamiento falla.

Mira ahora quién tiene inverso multiplicativo, es decir, en qué filas aparece un 1. Solo en las filas del 1 y del 5: $1 \cdot 1 = 1$ y $5 \cdot 5 = 25 = 24 + 1 = 1$. Los demás —2, 3 y 4— *no se pueden invertir*, por mucho que busques.

Pausa para pensar

Compara $\mathbb{Z}_6$ con $\mathbb{Z}_5 = \{0, 1, 2, 3, 4\}$. Calcula la tabla del producto módulo 5 y comprueba que ahí sí aparece un 1 en cada fila salvo la del 0: $1 \cdot 1 = 1$, $2 \cdot 3 = 6 = 1$, $4 \cdot 4 = 16 = 1$. Ningún par de elementos no nulos multiplica a cero.
¿Qué tienen de distinto el 5 y el 6? Uno es primo y el otro no. Esa sola diferencia es lo que separa un cuerpo de un anillo cualquiera, y lo demostraremos en la sección 24.7.

Ese es el experimento entero. Ahora la definición.

24.2. El vocabulario estructural completo: subanillos, ideales generados, anillos locales y localización

Este capítulo promete en su apertura que no hace falta nada de ningún otro para leerlo, y esa promesa hay que sostenerla palabra por palabra. Al recorrerlo aparecen varios términos que se emplean como si estuvieran definidos y no lo están: la palabra *subanillo* sale al hablar de ideales y nunca se dice qué significa; la notación $(2, x)$ se usa sin haber explicado qué es el ideal generado por dos elementos; se habla de que $\mathbb{Z}_n$ se descompone en *anillos locales* y de que $K[[x]]$ es un anillo local sin que aparezca en ningún sitio la definición de anillo local; y no se menciona ni una vez la localización, que es la construcción de la que el cuerpo de

fracciones es solo un caso particular. Esta sección cierra esa lista entera, desde cero y con las demostraciones hechas. Lo único que hace falta para leerla es lo que la sección siguiente escribe con todo detalle —que un anillo es un conjunto con dos operaciones, la suma con estructura de grupo abeliano y el producto asociativo y distributivo respecto de ella—; quien lo prefiera puede leer primero la sección siguiente y volver aquí.

Idea clave

Intuición primero. Dentro de un anillo hay dos clases de subconjuntos interesantes, y son casi opuestas. Los **subanillos** son las copias pequeñas del anillo: se parecen al ambiente, tienen su misma unidad y su misma aritmética, como $\mathbb{Z}$ dentro de $\mathbb{Q}$. Los **ideales** son agujeros negros: absorben todo lo que se les multiplique, como los múltiplos de 5 dentro de $\mathbb{Z}$. Un subanillo propio nunca es ideal y un ideal propio nunca es subanillo, porque contener al 1 y absorber son incompatibles. La tercera idea de la sección es la contraria de un cociente: en vez de tirar elementos, *añadir inversos* a los que uno elija. Eso es localizar, y es lo que convierte $\mathbb{Z}$ en $\mathbb{Q}$ si se invierte todo, y en algo mucho más fino si se invierte solo una parte.

24.2.1. Subanillos

Definición 24.1. Sea R un anillo con unidad. Un subconjunto $S \subseteq R$ es un *subanillo* si contiene al 1 de R, es cerrado para restas y es cerrado para productos:

$$1 \in S, \qquad a, b \in S \Rightarrow a - b \in S, \qquad a, b \in S \Rightarrow ab \in S.$$

Con las operaciones heredadas, S es entonces un anillo con unidad por derecho propio.

Observación 24.2. Se pide la resta y no por separado la suma y el opuesto porque una sola condición cubre las dos: de $a, b \in S$ sale $a - a = 0 \in S$, luego $0 - b = -b \in S$, y por tanto $a - (-b) = a + b \in S$. Es el mismo truco de economía que en el criterio de subgrupo.

Ejemplo 24.3 (Subanillos sí y subanillos no, uno por uno)**. (a)** $\mathbb{Z} \subset \mathbb{Q} \subset \mathbb{R} \subset \mathbb{C}$**.** Los tres son subanillos, con la misma unidad.

(b) $\mathbb{Z}[\mathrm{i}] \subset \mathbb{C}$**.** Subanillo: la resta es obvia y $(a + b\mathrm{i})(c + d\mathrm{i}) = (ac - bd) + (ad + bc)\mathrm{i}$ tiene coordenadas enteras. Igualmente $\mathbb{Z}[\sqrt{-5}\,] \subset \mathbb{C}$, porque $(a + b\sqrt{-5})(c + d\sqrt{-5}) = (ac - 5bd) + (ad + bc)\sqrt{-5}$.

(c) $2\mathbb{Z} \subset \mathbb{Z}$**.** *No* es subanillo: es cerrado para restas y productos, pero $1 \notin 2\mathbb{Z}$. En cambio sí es un ideal. Este par de hechos es el contraste esencial: un ideal propio de un anillo con unidad nunca es subanillo, porque si contuviera al 1 absorbería todo el anillo.

(d) Los únicos subanillos de $\mathbb{Z}$**.** Un subgrupo aditivo de $\mathbb{Z}$ es de la forma $n\mathbb{Z}$, y contiene al 1 solo si $n = 1$. Luego $\mathbb{Z}$ tiene exactamente un subanillo: él mismo. Es un anillo notablemente rígido.

(e) Las matrices diagonales dentro de $M_2(\mathbb{R})$**.** Subanillo: restas y productos de diagonales son diagonales y la identidad es diagonal. En cambio las matrices con segunda columna nula *no* son subanillo —no contienen a I_2— y sí son un ideal por la izquierda, como se comprueba viendo que BA actúa columna a columna.

(f) Un caso engañoso en $\mathbb{Z}_{12}$. El subconjunto $\{0,4,8\}$ es cerrado para restas y productos ($4\cdot 4 = 16 = 4$, $4\cdot 8 = 32 = 8$, $8\cdot 8 = 64 = 4$) y no contiene al 1: no es subanillo. Ahora bien, tiene una unidad *propia*, el 4, y con ella es isomorfo a $\mathbb{Z}_3$. Lo mismo pasa con $\{0,3,6,9\}$, cuya unidad propia es el 9 y que resulta isomorfo a $\mathbb{Z}_4$. Ambos son ideales de $\mathbb{Z}_{12}$, y los elementos 4 y 9 son sus idempotentes no triviales: la descomposición $\mathbb{Z}_{12} \cong \mathbb{Z}_3 \times \mathbb{Z}_4$ está escrita en ellos. ✓

24.2.2. Ideales generados y operaciones con ideales

Definición 24.4. Sea R un anillo conmutativo con unidad. El *ideal generado* por $a_1,\dots,a_n \in R$ es

$$(a_1,\dots,a_n) = \Big\{ r_1a_1 + \cdots + r_na_n \ : \ r_i \in R \Big\},$$

el menor ideal que los contiene a todos. Dados dos ideales I y J se definen

$$I+J = \{a+b : a\in I,\ b\in J\}, \qquad I\cap J, \qquad IJ = \Big\{\textstyle\sum_k a_kb_k :\ a_k\in I,\ b_k\in J\Big\},$$

que son de nuevo ideales, y se dice que I y J son *comaximales* si $I+J=R$.

Demostración. El conjunto de combinaciones $\sum r_ia_i$ es cerrado para restas y absorbe, luego es ideal; contiene a cada a_j tomando $r_j=1$ y el resto nulo; y cualquier ideal que contenga a los a_i contiene a todas esas combinaciones. Para IJ hay que tomar sumas finitas de productos y no productos sueltos, porque el conjunto de productos ab no es cerrado para la suma. □

Observación 24.5. La palabra *comaximal* es la traducción abstracta de "coprimos". En $\mathbb{Z}$, la igualdad $(m)+(n)=\mathbb{Z}$ equivale a que exista una identidad de Bézout $am+bn=1$, y eso equivale a $\gcd(m,n)=1$. Con ese vocabulario el teorema chino del resto se enuncia de una vez para todos los anillos: si I y J son comaximales, $R/(I\cap J)\cong R/I\times R/J$.

Regla: describir el ideal generado por un conjunto y operar con ideales

1. **Escribe el ideal como conjunto de combinaciones.** $(a_1,\dots,a_n)$ es $\{\sum r_ia_i\}$ y nada más. Todas las preguntas sobre pertenencia son preguntas sobre si un elemento se escribe así.

2. **En $\mathbb{Z}$ y en $K[x]$, funde los generadores en uno.** Ambos son euclídeos, luego todo ideal es principal y $(a,b)=(\gcd(a,b))$, calculado con el algoritmo de Euclides. Si te piden (a,b,c), haz gcd dos veces.

3. **Para probar que un elemento *no* está, busca un invariante.** Encuentra una propiedad que cumplan todos los generadores, que se conserve al multiplicar por cualquier cosa y al sumar, y que falle en el candidato. Paridad del término independiente, valor en un punto, grado mínimo: casi siempre sirve una de las tres.

4. **Para probar que un ideal no es principal, razona sobre el hipotético generador.** Si $(a,b)=(d)$ entonces $d\mid a$ y $d\mid b$; suele bastar con mirar grados o normas para llegar a que d es una unidad y luego contradecir que el ideal es propio.

5. **Recuerda las tres relaciones de tamaño:** $IJ\subseteq I\cap J\subseteq I\subseteq I+J$. La suma es el ideal más grande, el producto el más pequeño, y $IJ=I\cap J$ solo cuando I y J

son comaximales.

6. **Traduce a divisibilidad cuando puedas.** En $\mathbb{Z}$: la suma es el máximo común divisor, la intersección es el mínimo común múltiplo, el producto es el producto. Contener es dividir, y cuanto mayor es el ideal, menor es su generador.

Ejemplo 24.6 (Cinco cálculos con ideales)**. (a) En** $\mathbb{Z}$**.** $(4,6) = \{4r + 6s\}$. Como $2 = 4 \cdot (-1) + 6 \cdot 1$, el 2 está dentro, y todo elemento es par: $(4,6) = (2) = (\gcd(4,6))$. Además $(4)(6) = (24)$ y $(4) \cap (6) = (12) = (\operatorname{mcm}(4,6))$, de modo que aquí $IJ \subsetneq I \cap J$: la inclusión del paso 5 es estricta porque 4 y 6 no son comaximales. ✓

(b) Dos ideales comaximales. $(4) + (9) = \mathbb{Z}$, porque $4 \cdot (-2) + 9 \cdot 1 = 1$. Y en efecto $(4)(9) = (36) = (4) \cap (9)$, ya que $\operatorname{mcm}(4,9) = 36$. Ese es el caso en que el teorema chino se aplica: $\mathbb{Z}_{36} \cong \mathbb{Z}_4 \times \mathbb{Z}_9$.

(c) En $\mathbb{Q}[x]$**.** $(x^2-1,\, x^3-1)$. Euclides: $x^3-1 = x(x^2-1)+(x-1)$ y $x^2-1 = (x+1)(x-1)$, luego el máximo común divisor es $x-1$ y el ideal es $(x-1)$. ✓

(d) En $\mathbb{Z}[x]$**, el ideal** $(2,x)$**.** Por el paso 1 es $\{2f+xg\}$, es decir, exactamente los polinomios con término independiente par: ése es el invariante del paso 3, y el 1 no lo cumple, luego el ideal es propio. El capítulo ya demuestra en otro lugar que no es principal; lo que aquí se añade es el significado exacto de la notación, que hasta ahora se usaba sin definir.

(e) Un maximal con dos generadores, distinto del anterior. Considérese $(3, x^2+1)$ en $\mathbb{Z}[x]$. Reducir coeficientes módulo 3 y luego módulo x^2+1 define un homomorfismo sobreyectivo $\mathbb{Z}[x] \to \mathbb{F}_3[x]/(x^2+1)$ cuyo núcleo es ese ideal. Ahora bien, x^2+1 no tiene raíces en $\mathbb{F}_3$ —vale 1, 2 y 2 en 0, 1 y 2— y es de grado 2, luego es irreducible allí y el cociente es el cuerpo $\mathbb{F}_9$. Por tanto $(3, x^2+1)$ es maximal. ✓ Igual de concreto: $(2, x^2+x+1)$ tiene cociente $\mathbb{F}_2[x]/(x^2+x+1) = \mathbb{F}_4$, otro cuerpo, luego también es maximal.

Definición 24.7 (Ideales por la izquierda, por la derecha y bilateros)**.** Si R no es conmutativo, un subgrupo aditivo $I \subseteq R$ es *ideal por la izquierda* si $ra \in I$ para todos $r \in R$, $a \in I$; *ideal por la derecha* si $ar \in I$ en las mismas condiciones; y *bilátero* si cumple las dos. Solo los biláteros permiten formar el anillo cociente.

Ejemplo 24.8 (Los tres tipos, vistos en $M_2(K)$)**.** Sea L el conjunto de las matrices con la segunda columna nula. Multiplicar por la izquierda actúa columna a columna, luego BA conserva la columna nula: L es ideal por la izquierda. No lo es por la derecha, pues $\left(\begin{smallmatrix}1&0\\0&0\end{smallmatrix}\right)\left(\begin{smallmatrix}0&1\\0&0\end{smallmatrix}\right) = \left(\begin{smallmatrix}0&1\\0&0\end{smallmatrix}\right)$ se sale. Simétricamente, las matrices con la segunda fila nula forman un ideal por la derecha y no por la izquierda. Sobre $K = \mathbb{F}_2$ cada uno de esos conjuntos tiene 4 elementos de los 16 del anillo, y la comprobación exhaustiva por ordenador confirma las dos afirmaciones. ✓

Y sin embargo no hay ideales biláteros propios. Si $I \neq \{0\}$ es bilátero y $A \in I$ tiene una entrada $a_{k\ell} \neq 0$, entonces $E_{ik}AE_{\ell j} = a_{k\ell}E_{ij}$ está en I para todos i, j, donde E_{ij} es la matriz con un 1 en la posición (i,j). Dividiendo por el escalar y sumando las E_{ii} se obtiene $I_2 \in I$, luego $I = M_2(K)$. Se dice que $M_n(K)$ es un anillo *simple*: lleno de estructura lateral y sin ninguna bilátera.

24.2.3. El teorema de correspondencia y los otros dos de isomorfía

Teorema 24.9 (Correspondencia, segundo y tercero)**.** *Sea R un anillo conmutativo con unidad e I un ideal, y $\pi : R \to R/I$ la proyección.*

(i) $J \mapsto J/I$ es una biyección que respeta inclusiones entre los ideales de R que contienen a I y los ideales de R/I, con inversa $\bar{J} \mapsto \pi^{-1}(\bar{J})$.

(ii) Si S es un subanillo de R, entonces $(S+I)/I \cong S/(S \cap I)$.

(iii) Si $I \subseteq J$ son ideales, $(R/I)/(J/I) \cong R/J$.

Demostración. (i) Si $J \supseteq I$ es ideal, $\pi(J)$ lo es porque π es sobreyectiva; si $\bar{J}$ es ideal de R/I, su preimagen es un ideal que contiene a $\ker \pi = I$. Las dos construcciones se deshacen: $\pi^{-1}(\pi(J)) = J + I = J$ al ser $I \subseteq J$, y $\pi(\pi^{-1}(\bar{J})) = \bar{J}$ por sobreyectividad. Ambas son monótonas, luego respetan inclusiones.

(ii) Compóngase la inclusión $S \hookrightarrow R$ con π. Sale un homomorfismo $S \to R/I$ de imagen $(S+I)/I$ y núcleo $S \cap I$; el primer teorema de isomorfía, ya demostrado en el capítulo, concluye.

(iii) Compóngase $R \to R/I \to (R/I)/(J/I)$. Es sobreyectiva y su núcleo es $\{a : a + I \in J/I\} = J$; otra vez el primer teorema. □

Ejemplo 24.10 (La correspondencia en $\mathbb{Z}_{12}$, con la lista completa)**.** Los ideales de $\mathbb{Z}$ que contienen a (12) son los (d) con $d \mid 12$: seis en total. Por el teorema, $\mathbb{Z}_{12}$ tiene exactamente seis ideales, y son

$$\mathbb{Z}_{12}, \quad \{0,2,4,6,8,10\}, \quad \{0,3,6,9\}, \quad \{0,4,8\}, \quad \{0,6\}, \quad \{0\},$$

correspondientes a $d = 1, 2, 3, 4, 6, 12$. El retículo de ideales de $\mathbb{Z}_{12}$ es el de divisores de 12 puesto del revés: contener es dividir. ✓

Y el tercer teorema, con $I = (12) \subseteq J = (4)$, da $\mathbb{Z}_{12}/\{0,4,8\} \cong \mathbb{Z}/(4) = \mathbb{Z}_4$, lo que cuadra con el recuento $12/3 = 4$.

24.2.4. Anillos locales: la palabra que faltaba

Definición 24.11. Un anillo conmutativo con unidad R, con $1 \neq 0$, es *local* si tiene un único ideal maximal. Ese ideal se suele denotar $\mathfrak{m}$, y el cuerpo $R/\mathfrak{m}$ se llama su *cuerpo residual*.

Proposición 24.12. *R es local si y solo si el conjunto de sus elementos no invertibles es cerrado para la suma; y en ese caso ese conjunto* es *el único ideal maximal.*

Demostración. Supóngase que $\mathfrak{n}$, el conjunto de los no invertibles, es cerrado para la suma. Absorbe, porque si a no es unidad y ra lo fuera, digamos $(ra)s = 1$, entonces $a(rs) = 1$ y a sería unidad. Con la suma cerrada, $\mathfrak{n}$ es por tanto un ideal, y es propio porque $1 \notin \mathfrak{n}$. Además todo ideal propio está contenido en él, ya que un ideal que contenga una unidad es el anillo entero. Luego $\mathfrak{n}$ es el único maximal.

Recíprocamente, si R es local con maximal $\mathfrak{m}$, todo elemento no invertible a genera un ideal propio (a), que está contenido en algún maximal, necesariamente $\mathfrak{m}$; luego los no invertibles

están todos en $\mathfrak{m}$. Y ningún elemento de $\mathfrak{m}$ es unidad, porque $\mathfrak{m} \neq R$. Así $\mathfrak{m}$ coincide con el conjunto de no invertibles, que por ser ideal es cerrado para la suma. □

Ejemplo 24.13 (Cuatro anillos locales y uno que no lo es)**. (a) Todo cuerpo** es local, con $\mathfrak{m} = \{0\}$ y cuerpo residual él mismo.

(b) $\mathbb{Z}_{p^k}$ con p primo. Los no invertibles son los múltiplos de p, que son cerrados para la suma; luego es local con $\mathfrak{m} = (p)$ y cuerpo residual $\mathbb{Z}_{p^k}/(p) \cong \mathbb{F}_p$. Por ejemplo $\mathbb{Z}_8$ tiene unidades $\{1, 3, 5, 7\}$ y no unidades $\{0, 2, 4, 6\}$, que forman ideal. ✓

(c) $K[[x]]$, las series formales. Una serie es invertible exactamente cuando su término constante no lo es; las no invertibles son las de término constante nulo, es decir los múltiplos de x, que son cerrados para la suma. Luego $K[[x]]$ es local con $\mathfrak{m} = (x)$ y cuerpo residual K. Esta es la afirmación que el capítulo hacía sin haber definido la palabra.

(d) $\mathbb{Z}_{12}$ no es local. Sus ideales maximales son (2) y (3): los cocientes son $\mathbb{F}_2$ y $\mathbb{F}_3$. Y en efecto los no invertibles no son cerrados para la suma: 3 y 4 no son unidades y $3 + 4 = 7$ sí lo es. La razón profunda es la descomposición $\mathbb{Z}_{12} \cong \mathbb{Z}_4 \times \mathbb{Z}_3$, y en general $\mathbb{Z}_n$ es local si y solo si n es potencia de un primo. Eso es exactamente lo que quiere decir la frase, que el capítulo usa al hablar de idempotentes, de que el teorema chino "descompone $\mathbb{Z}_n$ en un producto de anillos locales".

24.2.5. Localizar: añadir inversos a la carta

Definición 24.14. Sea R conmutativo con unidad y $S \subseteq R$ un subconjunto *multiplicativo*, es decir, con $1 \in S$ y cerrado para el producto. El *localizado* $S^{-1}R$ es el conjunto de las fracciones formales $\frac{r}{s}$ con $r \in R$ y $s \in S$, identificando

$$\frac{r}{s} = \frac{r'}{s'} \quad \Longleftrightarrow \quad t\,(rs' - r's) = 0 \ \text{ para algún } t \in S,$$

con la suma y el producto habituales de fracciones. La aplicación $r \mapsto \frac{r}{1}$ es un homomorfismo de anillos $R \to S^{-1}R$, y todo elemento de S se vuelve invertible en el localizado.

Observación 24.15. El factor t de la definición parece un capricho y no lo es: sin él la relación no sería transitiva cuando R tiene divisores de cero. Si R es un dominio se puede tomar siempre $t = 1$ y la definición se simplifica a $rs' = r's$, que es la igualdad de fracciones del colegio. Con $S = R \setminus \{0\}$ en un dominio se obtiene exactamente el cuerpo de fracciones $\mathrm{Frac}(R)$ que este capítulo ya construye: la localización es esa misma construcción, ejecutada solo parcialmente.

Ejemplo 24.16 (Tres lupas de aumento distinto)**. (a) Invertirlo todo.** Con $R = \mathbb{Z}$ y $S = \mathbb{Z} \setminus \{0\}$ sale $S^{-1}\mathbb{Z} = \mathbb{Q}$.

(b) Invertir un solo elemento. Con $S = \{1, 2, 4, 8, \dots\}$ sale $\mathbb{Z}[\frac{1}{2}] = \{\frac{a}{2^k}\}$, los racionales diádicos. Sus unidades son $\pm 2^k$, y no es un cuerpo: el 3 sigue sin invertirse. Es el anillo de las fracciones que un ordenador maneja de forma exacta en binario.

(c) Invertir todo menos un primo. Ésta es la importante. Con $\mathfrak{p} = (3)$ y $S = \mathbb{Z} \setminus (3)$ —los enteros no divisibles por 3, que forman conjunto multiplicativo precisamente porque (3) es primo— sale

$$\mathbb{Z}_{(3)} = \left\{ \tfrac{a}{b} \in \mathbb{Q} \ : \ 3 \nmid b \right\}.$$

Es un anillo *local*: una fracción $\frac{a}{b}$ con $3 \nmid b$ es invertible si y solo si $3 \nmid a$, y las no invertibles —las de numerador múltiplo de 3— son cerradas para la suma. Su único maximal es $3\mathbb{Z}_{(3)}$ y su cuerpo residual es $\mathbb{F}_3$. Así, $\frac{2}{5}$ y $\frac{7}{4}$ son unidades, mientras que $\frac{3}{5}$ y $\frac{9}{2}$ no lo son. ✓

Que la construcción del apartado (c) se pueda hacer con cualquier ideal primo, y que siempre produzca un anillo local, es la razón de que los primos —y no los irreducibles— sean la noción correcta en álgebra conmutativa: la condición "$ab \in \mathfrak{p} \Rightarrow a \in \mathfrak{p}$ o $b \in \mathfrak{p}$" dice literalmente que el complementario es multiplicativo.

24.2.6. Módulos y elementos enteros, en dos definiciones

Definición 24.17. Sea R un anillo conmutativo con unidad. Un *R-módulo* es un grupo abeliano $(M, +)$ junto con una operación $R \times M \to M$ que cumple, para todos $r, s \in R$ y $m, n \in M$,

$$r(m+n) = rm + rn, \quad (r+s)m = rm + sm, \quad (rs)m = r(sm), \quad 1 \cdot m = m.$$

Son los axiomas de espacio vectorial con la única diferencia de que R no tiene por qué ser un cuerpo.

Ejemplo 24.18. Un K-módulo con K cuerpo es exactamente un espacio vectorial. Un $\mathbb{Z}$-módulo es exactamente un grupo abeliano, porque multiplicar por un entero ya estaba definido como suma repetida; en particular $\mathbb{Z}_6$ es un $\mathbb{Z}$-módulo con *torsión*, pues $6 \cdot \bar{1} = \bar{0}$ con $\bar{1} \neq 0$, fenómeno imposible en un espacio vectorial. Todo ideal $I \subseteq R$ es un R-módulo, y el cociente R/I también: por eso los ideales son la herramienta natural para estudiar anillos.

Definición 24.19. Sea $R \subseteq T$ un subanillo. Un elemento $\alpha \in T$ es *entero sobre* R si es raíz de algún polinomio *mónico* con coeficientes en R, es decir, si existe

$$\alpha^n + c_{n-1}\alpha^{n-1} + \cdots + c_1\alpha + c_0 = 0, \qquad c_i \in R.$$

La extensión $R \subseteq T$ es *entera* si todo elemento de T lo es.

Ejemplo 24.20 (Enteros y no enteros sobre $\mathbb{Z}$)**. (a) $\sqrt{2}$ es entero sobre** $\mathbb{Z}$, por ser raíz de $x^2 - 2$, que es mónico. Igual i con $x^2 + 1$, y $\sqrt{-5}$ con $x^2 + 5$. Por eso $\mathbb{Z}[\mathrm{i}]$, $\mathbb{Z}[\sqrt{2}]$ y $\mathbb{Z}[\sqrt{-5}\,]$ son extensiones enteras de $\mathbb{Z}$.

(b) $\frac{1+\sqrt{5}}{2}$ es entero sobre $\mathbb{Z}$, aunque tenga un 2 en el denominador: cumple $x^2 - x - 1 = 0$, como se comprueba sustituyendo. Éste es el motivo de que el anillo de enteros de $\mathbb{Q}(\sqrt{5})$ sea $\mathbb{Z}\big[\frac{1+\sqrt{5}}{2}\big]$ y no $\mathbb{Z}[\sqrt{5}]$: la palabra "entero" se define por la ecuación mónica, no por el aspecto del número.

(c) $\frac{1}{2}$ *no* es entero sobre $\mathbb{Z}$**.** Si lo fuera, se tendría $\big(\frac{1}{2}\big)^n + c_{n-1}\big(\frac{1}{2}\big)^{n-1} + \cdots + c_0 = 0$; multiplicando por 2^n queda $1 + 2\big(c_{n-1} + 2c_{n-2} + \cdots + 2^{n-1}c_0\big) = 0$, es decir, un número impar

igual a un número par. Imposible. En general, los únicos racionales enteros sobre $\mathbb{Z}$ son los enteros: eso es exactamente el teorema de la raíz racional, y se resume diciendo que $\mathbb{Z}$ es *íntegramente cerrado* en $\mathbb{Q}$. ✓

Observación 24.21 (Autosuficiencia del capítulo). Conviene dejar constancia de que la cadena

$$\text{cuerpo} \Rightarrow \text{euclídeo} \Rightarrow \text{DIP} \Rightarrow \text{DFU}$$

está demostrada íntegramente dentro de este capítulo, en los teoremas 24.90 y 24.93, y que cualquier uso de ese resultado —por ejemplo el que hace la demostración del teorema de Gauss en su versión fuerte al afirmar que $K[x]$ es un dominio de factorización única por ser euclídeo— debe apoyarse en esos dos enunciados y no en ningún otro capítulo. Con eso, y con el vocabulario que esta sección acaba de fijar, el capítulo no delega nada fuera de sí mismo.

Idea clave

Lo que esta sección añade se resume en cuatro pares de opuestos. Subanillo frente a ideal: parecerse frente a absorber. Ideal por un lado frente a bilátero: la conmutatividad decidiendo cuántas nociones hay. Cociente frente a localización: quitar elementos frente a añadir inversos. Y anillo local frente a producto de anillos: mirar un solo primo frente a mirarlos todos a la vez. Casi todo el álgebra conmutativa consiste en moverse entre esos cuatro pares con soltura.

24.3. Ahora sí: la definición, regla por regla

Regla: qué es un anillo

Un **anillo** es un conjunto R con dos operaciones, $+$ y $\cdot$, que cumplen:

A1. $(R,+)$ **es un grupo abeliano**: la suma es asociativa y conmutativa, hay un neutro 0, y todo a tiene opuesto $-a$.

A2. El producto es asociativo: $(ab)c = a(bc)$.

A3. Distributiva: $a(b+c) = ab + ac$ y $(b+c)a = ba + ca$.

Dos apellidos que casi siempre añadiremos:

- **con unidad**: existe $1 \in R$ con $1 \cdot a = a \cdot 1 = a$;
- **conmutativo**: $ab = ba$ para todos.

En este capítulo, *anillo* significará siempre **anillo conmutativo con unidad**, salvo aviso.

Fíjate en lo que *no* se pide: no se pide que exista a^{-1}. Esa ausencia es deliberada y es toda la gracia del asunto. Fíjate también en que la distributiva es la única regla que conecta las dos operaciones; sin ella tendrías dos estructuras independientes puestas en el mismo conjunto, no un anillo.

Regla: las consecuencias inmediatas, demostradas

De los axiomas salen, sin más herramientas, estos hechos:
(1) $0 \cdot a = 0$ **siempre.** Demostración: $0 \cdot a = (0+0) \cdot a = 0 \cdot a + 0 \cdot a$ por distributiva. Restando $0 \cdot a$ a ambos lados (se puede, porque la suma es grupo): $0 = 0 \cdot a$. *Nota que esto no se supone: se demuestra.*
(2) $(-a)b = -(ab)$**.** Demostración: $ab + (-a)b = (a + (-a))b = 0 \cdot b = 0$, luego $(-a)b$ es el opuesto de ab.
(3) $(-a)(-b) = ab$**.** "Menos por menos es más" no es una convención arbitraria: aplica (2) dos veces.
(4) Si $1 = 0$**, el anillo tiene un solo elemento.** Porque $a = 1 \cdot a = 0 \cdot a = 0$ para todo a. Por eso en la definición de cuerpo se exige $1 \neq 0$: para excluir ese caso degenerado.

Observación 24.22. El punto (3) merece un segundo. En la escuela "menos por menos da más" se aprende de memoria o con analogías dudosas ("el enemigo de mi enemigo"). Aquí queda claro de dónde sale: es la *única* posibilidad compatible con la distributiva. Si quisieras que $(-1)(-1) = -1$, entonces $0 = (-1)\cdot 0 = (-1)(1+(-1)) = (-1)+(-1)(-1) = (-1)+(-1) = -2$, y el anillo colapsaría.

24.4. Un catálogo de anillos, verificado uno por uno

Ejemplo 24.23 (Los que sí son anillos)**.** $\mathbb{Z}$, con la suma y el producto usuales. Conmutativo, con unidad. Es el anillo de referencia.

$\mathbb{Q}$**,** $\mathbb{R}$**,** $\mathbb{C}$: anillos, y además cuerpos, porque todo elemento no nulo se invierte.

$\mathbb{Z}_n$ para cualquier $n \geq 2$: anillo conmutativo con unidad. La comprobación de que las operaciones están bien definidas (que el resultado no depende del representante elegido) es exactamente la del capítulo 3.

$K[x]$, los polinomios con coeficientes en un cuerpo K: anillo conmutativo con unidad (el polinomio constante 1). Será el protagonista de la segunda mitad del capítulo.

$M_n(\mathbb{R})$, las matrices $n \times n$: anillo con unidad (la identidad) pero **no conmutativo**, porque $AB \neq BA$ en general. Es el ejemplo estándar de anillo no conmutativo, y muestra que la no conmutatividad es la norma, no la excepción.

$C[0,1]$, las funciones continuas de $[0,1]$ en $\mathbb{R}$ con suma y producto punto a punto: anillo conmutativo con unidad (la función constante 1).

$\mathbb{Z}[i] = \{a + bi : a, b \in \mathbb{Z}\}$, los *enteros de Gauss*: cerrado para suma y producto, porque $(a+bi)(c+di) = (ac-bd) + (ad+bc)i$ tiene otra vez componentes enteras.

Ejemplo 24.24 (Los que NO son anillos, y por qué exactamente)**.** $\mathbb{N} = \{0, 1, 2, \dots\}$: falla A1. No hay opuestos; 3 no tiene un -3 dentro. Es un *semianillo*, no un anillo.

Los enteros impares: falla la suma. $3 + 5 = 8$ no es impar. Ni siquiera es cerrado.

Los enteros pares $2\mathbb{Z}$: es anillo *sin unidad*, porque $1 \notin 2\mathbb{Z}$. Sirve para ver que la unidad no es gratis.

$\mathbb{R}^3$ **con el producto vectorial**: falla A2. El producto cruz no es asociativo, como muestra $(\mathbf{i} \times \mathbf{i}) \times \mathbf{j} = \mathbf{0}$ frente a $\mathbf{i} \times (\mathbf{i} \times \mathbf{j}) = \mathbf{i} \times \mathbf{k} = -\mathbf{j}$.

Cuidado

Un error frecuente al empezar: creer que "anillo" significa "números". $C[0,1]$ es un anillo cuyos elementos son *funciones*, y $M_n(\mathbb{R})$ uno cuyos elementos son *tablas de números*. Lo que hace anillo a un conjunto no es la naturaleza de sus elementos, sino el comportamiento de sus dos operaciones. Esta es la misma lección abstracta del capítulo 23, y conviene interiorizarla ya: en álgebra abstracta las cosas se definen por lo que hacen, no por lo que son.

24.5. Unidades y divisores de cero: los dos defectos de $\mathbb{Z}_6$

Volvamos a lo que descubrimos en la tabla de $\mathbb{Z}_6$ y démosle nombre.

Regla: unidad y divisor de cero

Sea R un anillo conmutativo con unidad.

- $u \in R$ es una **unidad** si existe v con $uv = 1$. El conjunto de unidades se escribe $R^\times$ o $U(R)$.
- $a \neq 0$ es un **divisor de cero** si existe $b \neq 0$ con $ab = 0$.

Un elemento nunca puede ser las dos cosas. Si u es unidad y $ub = 0$, multiplica por v: $b = 1 \cdot b = (vu)b = v(ub) = v \cdot 0 = 0$. Así que b tenía que ser cero, y u no es divisor de cero.

Ejemplo 24.25 (Las unidades de varios anillos, calculadas)**.** $U(\mathbb{Z}) = \{1, -1\}$**.** Si $ab = 1$ con enteros, entonces $|a||b| = 1$ y ambos valen 1.

$U(\mathbb{Z}_n) = \{a : \gcd(a, n) = 1\}$**.** Razón: a tiene inverso módulo n si y solo si existen u, v con $au + nv = 1$, y por Bézout (capítulo 3) eso ocurre exactamente cuando $\gcd(a, n) = 1$. Casos concretos, comprobados por ordenador:

$$U(\mathbb{Z}_8) = \{1,3,5,7\}, \quad U(\mathbb{Z}_9) = \{1,2,4,5,7,8\}, \quad U(\mathbb{Z}_{10}) = \{1,3,7,9\}, \quad U(\mathbb{Z}_{12}) = \{1,5,7,11\}.$$

El número de unidades es $\varphi(n)$: 4, 6, 4 y 4 respectivamente. ✓

$U(K[x]) = K \setminus \{0\}$, las constantes no nulas. Razón: los grados se suman, $\deg(fg) = \deg f + \deg g$, así que $fg = 1$ obliga a $\deg f = \deg g = 0$. *Este hecho —que las unidades son solo las constantes— es lo que hace que "irreducible" signifique algo en $K[x]$.*

$U(\mathbb{Z}[i]) = \{1, -1, i, -i\}$**.** Con la norma $N(a + bi) = a^2 + b^2$, que es multiplicativa, $\alpha\beta = 1$ fuerza $N(\alpha)N(\beta) = 1$, luego $N(\alpha) = 1$ y $a^2 + b^2 = 1$: solo cuatro soluciones enteras.

Ejemplo 24.26 (Los divisores de cero de $\mathbb{Z}_6$, todos)**.** Buscando en la tabla los pares (a, b) con $a, b \neq 0$ y $ab = 0$:

$$(2,3), \quad (3,2), \quad (3,4), \quad (4,3).$$

Es decir, los divisores de cero son $\{2, 3, 4\}$, y las unidades $\{1, 5\}$. Entre ambos conjuntos cubren todo $\mathbb{Z}_6$ salvo el 0: no queda ningún elemento "neutro". Eso no es casualidad en $\mathbb{Z}_n$

—todo elemento no nulo es unidad o divisor de cero—, pero sí lo es en general: en $\mathbb{Z}$, el elemento 2 no es unidad ni divisor de cero.

Cuidado

Sin divisores de cero no se puede cancelar. La regla escolar "si $ab = ac$ y $a \neq 0$ entonces $b = c$" es *falsa* en un anillo cualquiera. En $\mathbb{Z}_6$: $2 \cdot 1 = 2$ y $2 \cdot 4 = 8 = 2$, luego $2 \cdot 1 = 2 \cdot 4$ pero $1 \neq 4$.
Lo que sí es cierto: cancelar es legítimo exactamente cuando a no es divisor de cero, porque $ab = ac \Rightarrow a(b-c) = 0 \Rightarrow b - c = 0$. Por eso los dominios de integridad, donde no hay divisores de cero, son el terreno donde el álgebra escolar sigue valiendo.

24.6. Dominios de integridad y cuerpos

Regla: la escalera de tres peldaños

- **Dominio de integridad**: anillo conmutativo con $1 \neq 0$ y *sin divisores de cero*. Ejemplos: $\mathbb{Z}$, $K[x]$, $\mathbb{Z}[i]$, todo cuerpo.
- **Cuerpo**: anillo conmutativo con $1 \neq 0$ donde *todo* elemento no nulo es unidad. Ejemplos: $\mathbb{Q}$, $\mathbb{R}$, $\mathbb{C}$, $\mathbb{Z}_p$ con p primo.

Todo cuerpo es dominio (una unidad nunca es divisor de cero), pero no al revés: $\mathbb{Z}$ es dominio y no cuerpo.

Teorema 24.27 (Todo dominio finito es un cuerpo)**.** *Si R es un dominio de integridad con un número finito de elementos, entonces R es un cuerpo.*

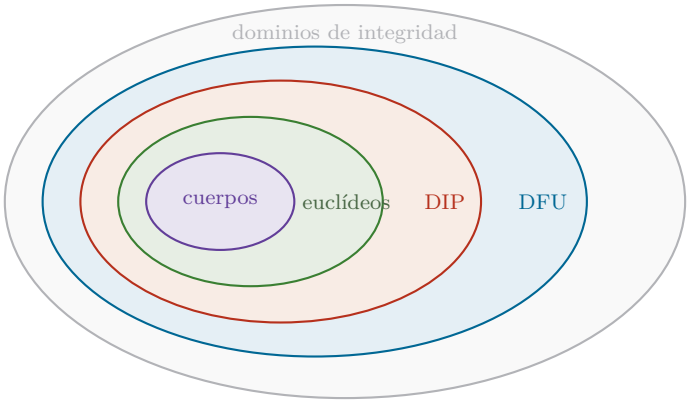


Figura 24.1: La jerarquía de dominios. Cada inclusión es estricta, y hay un testigo concreto de ello: $\mathbb{Z}$ es euclídeo y no es cuerpo; $\mathbb{Z}\left[\frac{1+\sqrt{-19}}{2}\right]$ es de ideales principales sin ser euclídeo; $\mathbb{Z}[x]$ es de factorización única sin ser de ideales principales, porque $(2, x)$ no es principal; y en $\mathbb{Z}[\sqrt{-5}]$ la factorización deja de ser única.

Demostración. Sea $a \neq 0$. Considera la aplicación $\mu_a : R \to R$ dada por $\mu_a(x) = ax$. Es inyectiva: si $ax = ay$ entonces $a(x-y) = 0$, y como $a \neq 0$ y no hay divisores de cero, $x = y$. Una aplicación inyectiva de un conjunto finito en sí mismo es sobreyectiva, así que existe x con $ax = 1$. Luego a es unidad. □

Observación 24.28. La demostración usa la finitud dos veces en la misma frase —"inyectiva de un conjunto finito en sí mismo es sobreyectiva"— y ese es todo el truco. Es el principio del palomar disfrazado. En $\mathbb{Z}$, que es infinito, μ_2 es inyectiva pero no sobreyectiva: su imagen son los pares. Por eso $\mathbb{Z}$ se queda en dominio y no llega a cuerpo.

24.7. Los anillos $\mathbb{Z}_n$ con todo detalle

Teorema 24.29 (El criterio de primalidad, en versión algebraica)**.** *$\mathbb{Z}_n$ es un cuerpo $\iff$ $\mathbb{Z}_n$ es un dominio $\iff$ n es primo.*

Demostración. Si $n = ab$ con $1 < a, b < n$, entonces en $\mathbb{Z}_n$ se tiene $a \neq 0$, $b \neq 0$ y $ab = n = 0$: hay divisores de cero, luego no es dominio.

Si $n = p$ es primo y $ab \equiv 0 \pmod p$, entonces $p \mid ab$, y por el lema de Euclides (capítulo 3) $p \mid a$ o $p \mid b$, es decir $a = 0$ o $b = 0$ en $\mathbb{Z}_p$: es dominio. Y como es finito, el teorema 24.27 lo asciende a cuerpo. □

Ejemplo 24.30 (La tabla de inversos de $\mathbb{Z}_7$ y de $\mathbb{Z}_{11}$)**.** Buscando en cada fila dónde aparece el 1:

$$\mathbb{Z}_7: \quad 1^{-1}=1,\ 2^{-1}=4,\ 3^{-1}=5,\ 4^{-1}=2,\ 5^{-1}=3,\ 6^{-1}=6.$$

$$\mathbb{Z}_{11}: \quad 1^{-1}=1,\ 2^{-1}=6,\ 3^{-1}=4,\ 4^{-1}=3,\ 5^{-1}=9,\ 6^{-1}=2,\ 7^{-1}=8,\ 8^{-1}=7,\ 9^{-1}=5,\ 10^{-1}=10.$$

Dos comprobaciones de sentido que conviene hacer siempre: la lista es simétrica (si $a^{-1} = b$ entonces $b^{-1} = a$), y los únicos elementos que son su propio inverso son 1 y -1 —es decir 6 en $\mathbb{Z}_7$ y 10 en $\mathbb{Z}_{11}$—, porque $a^2 = 1$ significa $(a-1)(a+1) = 0$ y en un cuerpo eso obliga a $a = \pm 1$.

Cómo calcular un inverso sin tabla. Para 3^{-1} en $\mathbb{Z}_{11}$: aplica Euclides extendido a 3 y 11. Como $11 = 3 \cdot 3 + 2$, $3 = 2 + 1$, se retrocede: $1 = 3 - 2 = 3 - (11 - 3 \cdot 3) = 4 \cdot 3 - 11$. Luego $4 \cdot 3 \equiv 1 \pmod{11}$ y $3^{-1} = 4$. ✓ Este método funciona con módulos de cientos de dígitos, que es exactamente lo que hace falta en RSA.

24.8. Homomorfismos: traducir entre anillos

Regla: homomorfismo de anillos

$\varphi : R \to S$ es un **homomorfismo** si respeta las dos operaciones y la unidad:

$$\varphi(a+b) = \varphi(a) + \varphi(b), \qquad \varphi(ab) = \varphi(a)\varphi(b), \qquad \varphi(1_R) = 1_S.$$

Su **núcleo** es $\ker \varphi = \{a \in R : \varphi(a) = 0\}$, y su **imagen** $\varphi(R)$.
φ es **inyectivo** $\iff$ $\ker \varphi = \{0\}$, exactamente por la misma razón que en grupos.

Ejemplo 24.31 (Tres homomorfismos que ya usabas sin saberlo)**. Reducción módulo** n: $\pi : \mathbb{Z} \to \mathbb{Z}_n$, $\pi(a) = a$ mód n. Que sea homomorfismo es precisamente la afirmación "el resto de una suma es la suma de los restos, y lo mismo con productos", que es lo que hace útil la aritmética modular. Su núcleo es $n\mathbb{Z}$.

Evaluación: fijado $c \in K$, la aplicación $\mathrm{ev}_c : K[x] \to K$ que manda $f \mapsto f(c)$. Que respete sumas y productos es la afirmación de que "evaluar y luego operar da lo mismo que operar y luego evaluar". Su núcleo son los polinomios que se anulan en c, es decir los múltiplos de $x - c$. *Ese hecho, leído al revés, es el teorema del factor.*

Conjugación: $\mathbb{C} \to \mathbb{C}$, $z \mapsto \bar{z}$. Respeta sumas y productos, es biyectiva, y por tanto es un *automorfismo* de $\mathbb{C}$. Es el ejemplo más pequeño posible de grupo de Galois no trivial, y reaparecerá en el capítulo 26.

24.9. Ideales: los subconjuntos que sirven para dividir

Idea clave

Intuición primero. En grupos, los subgrupos por los que se puede cocientar son los normales. En anillos, el papel lo hacen los **ideales**, y la condición que los define es más fuerte que ser subanillo: un ideal absorbe.
Piensa en los múltiplos de 5 dentro de $\mathbb{Z}$. Si sumas dos múltiplos de 5 te queda un múltiplo de 5; pero además, si multiplicas un múltiplo de 5 por *cualquier* entero —no necesariamente múltiplo de 5— sigues dentro. Esa segunda propiedad, la absorción, es la que permite después decir "trabajemos módulo 5 y olvidemos todo lo demás". Un ideal es un agujero negro para la multiplicación.

Regla: ideal

$I \subseteq R$ es un **ideal** si:

- **I1.** I es subgrupo aditivo: $0 \in I$ y si $a, b \in I$ entonces $a - b \in I$;
- **I2. absorbe**: si $a \in I$ y $r \in R$ (cualquiera), entonces $ra \in I$.

El **ideal generado** por a es $(a) = \{ra : r \in R\}$; los ideales de esta forma se llaman **principales**.

Teorema 24.32 (Todos los ideales de $\mathbb{Z}$)**.** *Todo ideal de* $\mathbb{Z}$ *es* $(n) = n\mathbb{Z}$ *para algún* $n \geq 0$. *Es decir,* $\mathbb{Z}$ *es un* dominio de ideales principales.

Demostración. Sea $I \neq \{0\}$ un ideal y sea n el menor entero positivo de I (existe por buena ordenación). Claramente $n\mathbb{Z} \subseteq I$ por absorción. Recíprocamente, si $a \in I$, divide con resto: $a = qn + r$ con $0 \leq r < n$. Entonces $r = a - qn \in I$, y como n era el menor positivo de I, forzosamente $r = 0$. Luego $a \in n\mathbb{Z}$. □

Observación 24.33. La demostración anterior usa *solo* la existencia de división con resto. Palabra por palabra, la misma prueba funciona en $K[x]$ cambiando "el menor entero positivo"

por "el polinomio no nulo de menor grado". Esa es la razón profunda de que $\mathbb{Z}$ y $K[x]$ se parezcan tanto: ambos son euclídeos, y de ahí sale todo lo demás en cascada.

Ejemplo 24.34 (Un ideal que no es principal)**.** En $\mathbb{Z}[x]$ (coeficientes enteros, no de un cuerpo) considera $I = (2, x)$, los polinomios con término independiente par. Si fuera $I = (f)$, entonces f dividiría a 2, luego f sería constante ± 1 o ± 2; pero ± 1 genera todo $\mathbb{Z}[x]$ (y $1 \notin I$), y ± 2 no divide a x. Contradicción.

Moraleja: $\mathbb{Z}[x]$ es dominio de factorización única pero *no* de ideales principales. La cadena de implicaciones va en un solo sentido, y este ejemplo lo prueba.

24.10. Anillos cociente: fabricar $\mathbb{C}$ a partir de $\mathbb{R}$

Regla: el anillo cociente

Dado un ideal $I \subseteq R$, el **cociente** R/I es el conjunto de clases $a + I$ con las operaciones

$$(a + I) + (b + I) = (a + b) + I, \qquad (a + I)(b + I) = ab + I.$$

La absorción (I2) es exactamente lo que hace que el producto esté **bien definido**: si $a' = a + i$ y $b' = b + j$ con $i, j \in I$, entonces $a'b' = ab + (aj + ib + ij)$, y los tres sumandos del paréntesis están en I por absorción.

Ejemplo 24.35 ($\mathbb{Z}/n\mathbb{Z} = \mathbb{Z}_n$)**.** El caso que ya conocías. Cocientar $\mathbb{Z}$ por el ideal $n\mathbb{Z}$ es exactamente "trabajar módulo n". Todo el capítulo 3 era, sin decirlo, un estudio de anillos cociente.

Ejemplo 24.36 ($\mathbb{R}[x]/(x^2 + 1) \cong \mathbb{C}$: los complejos construidos, no postulados)**.** Cocientar $\mathbb{R}[x]$ por el ideal generado por $x^2 + 1$ significa: trabajar con polinomios reales, pero declarando que $x^2 + 1 = 0$, es decir, $x^2 = -1$.

Paso 1. Quiénes son las clases. Dividiendo con resto por $x^2 + 1$, todo polinomio es equivalente a uno de grado ≤ 1. Así que las clases son exactamente $a + bx$ con $a, b \in \mathbb{R}$.

Paso 2. Cómo se multiplican.

$$(a + bx)(c + dx) = ac + (ad + bc)x + bd\,x^2 = (ac - bd) + (ad + bc)x,$$

usando $x^2 = -1$ en el último paso.

Paso 3. Reconocerlo. Esa es, letra por letra, la regla del producto de números complejos. Llamando i a la clase de x, hemos *construido* $\mathbb{C}$ sin postular nada: no hemos dicho "sea i un número cuyo cuadrado es -1" —afirmación que a cualquiera con criterio le suena a trampa—, sino que hemos fabricado un anillo perfectamente legítimo donde ese elemento existe.

Idea clave

Ese ejemplo es la receta universal para construir cuerpos, y merece leerse dos veces. **Si un polinomio irreducible no tiene raíces donde estás, cocienta por él y**

aparecerá una. x^2+1 no tiene raíz en $\mathbb{R}$; en $\mathbb{R}[x]/(x^2+1)$ la tiene. El mismo gesto, aplicado a $\mathbb{Q}$ y x^3-2, construye $\mathbb{Q}(\sqrt[3]{2})$; aplicado a $\mathbb{F}_2$ y x^2+x+1, construye el cuerpo de cuatro elementos. Todo el capítulo 26 vive de esta única idea.

Regla: ideales primos y maximales

Sea R conmutativo con unidad e $I \neq R$ un ideal.

- I es **primo** si $ab \in I$ implica $a \in I$ o $b \in I$. $\iff$ R/I es dominio.
- I es **maximal** si no hay ideales estrictamente entre I y R. $\iff$ R/I es cuerpo.

Como todo cuerpo es dominio, **maximal** $\Rightarrow$ **primo**. El recíproco falla: en $\mathbb{Z}$, el ideal (0) es primo (porque $\mathbb{Z}$ es dominio) pero no maximal.

Ejemplo 24.37 (Los dos diccionarios, comprobados)**.** En $\mathbb{Z}$: el ideal (6) no es primo, porque $2\cdot 3 \in (6)$ pero ni 2 ni 3 lo están —y en efecto $\mathbb{Z}_6$ tiene divisores de cero—. El ideal (5) sí es maximal, y $\mathbb{Z}_5$ es cuerpo. ✓

En $\mathbb{R}[x]$: (x^2+1) es maximal porque x^2+1 es irreducible, y en efecto el cociente es $\mathbb{C}$, un cuerpo. En cambio (x^2-1) no lo es, porque $x^2-1=(x-1)(x+1)$; el cociente tiene divisores de cero, ya que la clase de $x-1$ por la de $x+1$ da cero sin que ninguna sea nula. ✓

24.11. Polinomios: la aritmética paralela a la de $\mathbb{Z}$

Idea clave

Intuición primero. A partir de aquí conviene tener el capítulo 3 al lado, porque vamos a repetir su contenido cambiando "entero" por "polinomio" y "valor absoluto" por "grado". La correspondencia es esta:

En $\mathbb{Z}$	**En** $K[x]$
$\lvert a\rvert$ mide el tamaño	$\deg f$ mide el tamaño
división con resto $\lvert r\rvert < \lvert b\rvert$	división con resto $\deg r < \deg g$
unidades ± 1	unidades $=$ constantes no nulas
números primos	polinomios irreducibles
factorización única	factorización única
algoritmo de Euclides	algoritmo de Euclides
Bézout $au+bv=\gcd$	Bézout $fu+gv=\gcd$

No es una analogía poética: es el mismo teorema demostrado dos veces, y por eso ambos son *dominios euclídeos*.

Regla: grado y sus reglas

Para $f \neq 0$, $\deg f$ es el mayor exponente con coeficiente no nulo. Convenio: $\deg 0 = -\infty$.

- $\deg(f+g) \leq \max(\deg f, \deg g)$, con posible caída si los términos líderes se cancelan;
- $\deg(fg) = \deg f + \deg g$ **siempre que los coeficientes vivan en un dominio**.

La segunda regla es la que se usa todo el tiempo, y la condición del dominio es imprescindible: en $\mathbb{Z}_4[x]$, $(2x)(2x) = 4x^2 = 0$, cuyo grado no es 2.

Teorema 24.38 (División con resto en $K[x]$)**.** *Si K es un cuerpo y $f, g \in K[x]$ con $g \neq 0$, existen* únicos *$q, r \in K[x]$ tales que*

$$f = qg + r, \qquad \deg r < \deg g.$$

Demostración. Existencia, por inducción sobre $\deg f$. Si $\deg f < \deg g$, toma $q = 0$, $r = f$. Si $\deg f = n \geq m = \deg g$, sean a y b los coeficientes líderes; el polinomio $f_1 = f - \frac{a}{b}x^{n-m}g$ tiene grado menor que n (el término líder se cancela por construcción), y por hipótesis de inducción $f_1 = q_1 g + r$. Entonces $f = (q_1 + \frac{a}{b}x^{n-m})g + r$.

Unicidad: si $q_1 g + r_1 = q_2 g + r_2$, entonces $(q_1 - q_2)g = r_2 - r_1$. El lado derecho tiene grado $< \deg g$; si $q_1 \neq q_2$, el izquierdo tendría grado $\geq \deg g$. Luego $q_1 = q_2$ y $r_1 = r_2$. □

Cuidado

Que K sea **cuerpo** es esencial: en el paso de la inducción se divide entre b, el coeficiente líder de g. En $\mathbb{Z}[x]$ la división falla: no se puede dividir x^2 entre $2x + 1$ con coeficientes enteros. Por eso $\mathbb{Z}[x]$ no es euclídeo, aunque sí conserve la factorización única.

Ejemplo 24.39 (Una división completa, término a término)**.** Dividamos $f = x^4 - 3x^3 + 2x - 5$ entre $g = x^2 - 1$.

Paso 1. Término líder de f entre el de g: $x^4/x^2 = x^2$. Multiplicamos: $x^2(x^2 - 1) = x^4 - x^2$. Restamos:

$$f - (x^4 - x^2) = -3x^3 + x^2 + 2x - 5.$$

Paso 2. $-3x^3/x^2 = -3x$. Multiplicamos: $-3x(x^2 - 1) = -3x^3 + 3x$. Restamos:

$$(-3x^3 + x^2 + 2x - 5) - (-3x^3 + 3x) = x^2 - x - 5.$$

Paso 3. $x^2/x^2 = 1$. Multiplicamos: $x^2 - 1$. Restamos:

$$(x^2 - x - 5) - (x^2 - 1) = -x - 4.$$

Grado 1 < grado 2: paramos. Resultado:

$$x^4 - 3x^3 + 2x - 5 = (x^2 - 3x + 1)(x^2 - 1) + (-x - 4).$$

Comprobación: $(x^2 - 3x + 1)(x^2 - 1) = x^4 - 3x^3 + x^2 - x^2 + 3x - 1 = x^4 - 3x^3 + 3x - 1$, y sumando $-x - 4$ queda $x^4 - 3x^3 + 2x - 5$. ✓ (Verificado también con álgebra computacional.)

Ejemplo 24.40 (Euclides con polinomios: $\gcd(x^4-1,\, x^3-1)$)**. Paso 1.** $x^4-1 = x(x^3-1)+(x-1)$.

Paso 2. $x^3-1 = (x^2+x+1)(x-1)+0$.

Resto cero: el gcd es el último resto no nulo, $\boxed{x-1}$. Tiene sentido: la única raíz común de $x^4=1$ y $x^3=1$ en $\mathbb{C}$ es $x=1$.

Bézout, retrocediendo: de la primera igualdad, $x-1 = (x^4-1) - x(x^3-1)$, es decir

$$1\cdot(x^4-1) + (-x)\cdot(x^3-1) = x-1.$$

Coeficientes $u=1$, $v=-x$. ✓

24.12. Raíces, factores y el teorema del resto

Teorema 24.41 (Resto, factor y cota de raíces)**.** *Sea K un cuerpo y $f \in K[x]$.*

(i) ***Teorema del resto****: el resto de dividir f entre $x-c$ es la constante $f(c)$.*

(ii) ***Teorema del factor****: c es raíz de $f \iff (x-c) \mid f$.*

(iii) Un polinomio no nulo de grado n tiene ***a lo sumo*** *n* ***raíces*** *en K.*

Demostración. (i) Divide: $f = q\cdot(x-c)+r$ con $\deg r < 1$, luego r es constante. Evaluando en c: $f(c) = q(c)\cdot 0 + r = r$.

(ii) Inmediato de (i): $f(c)=0$ equivale a resto cero.

(iii) Por inducción. Si f no tiene raíces, listo. Si c es raíz, $f=(x-c)g$ con $\deg g = n-1$; cualquier otra raíz $d \neq c$ cumple $0 = f(d) = (d-c)g(d)$, y como K es dominio y $d-c\neq 0$, es raíz de g. Por inducción g tiene a lo sumo $n-1$ raíces. □

Cuidado

El punto (iii) **falla estrepitosamente** si el anillo tiene divisores de cero. En $\mathbb{Z}_8$, el polinomio x^2-1 tiene grado 2 y *cuatro* raíces: $1, 3, 5, 7$ (comprueba: $3^2 = 9 = 1$, $5^2 = 25 = 1$, $7^2 = 49 = 1$). El culpable es que $\mathbb{Z}_8$ no es dominio, así que el paso "$(d-c)g(d)=0$ implica $g(d)=0$" se cae.

Este ejemplo, además, tiene consecuencias prácticas: es la razón de que ciertos ataques a sistemas criptográficos funcionen módulo compuestos y no módulo primos.

Ejemplo 24.42 (Raíces racionales: el criterio y su uso)**.** Si $f = a_n x^n + \cdots + a_0$ tiene coeficientes enteros y p/q es raíz racional en forma reducida, entonces $p \mid a_0$ y $q \mid a_n$. (Sustituye, multiplica por q^n y observa qué divide a qué.)

Aplicación: ¿tiene $f = 2x^3 - 3x^2 - 3x + 2$ raíces racionales? Candidatos: $p \in \{\pm 1, \pm 2\}$, $q \in \{\pm 1, \pm 2\}$, o sea $\pm 1, \pm 2, \pm\frac{1}{2}$. Probando: $f(1) = 2-3-3+2 = -2$; $f(-1) = -2-3+3+2 = 0$. ✓ Dividiendo, $f = (x+1)(2x^2-5x+2) = (x+1)(2x-1)(x-2)$. Las tres raíces son -1, $\frac{1}{2}$ y 2, todas racionales.

Uso negativo, que es el más potente: para $f = x^3 - x - 1$ los candidatos son ± 1, y $f(1) = -1$, $f(-1) = -1$. Sin raíces racionales; como el grado es 3, eso basta para concluir que *es irreducible sobre* $\mathbb{Q}$.

24.13. Irreducibilidad: los criterios que de verdad se usan

Regla: irreducible

$f \in K[x]$ no constante es **irreducible sobre** K si no se puede escribir $f = gh$ con $\deg g, \deg h \geq 1$ y coeficientes en K.
La advertencia número uno: la irreducibilidad *depende del cuerpo*. x^2-2 es irreducible sobre $\mathbb{Q}$ y reducible sobre $\mathbb{R}$; x^2+1 es irreducible sobre $\mathbb{R}$ y reducible sobre $\mathbb{C}$. Nunca digas "este polinomio es irreducible" sin decir dónde.

Regla: la caja de herramientas, en orden de uso

1. Grados 2 y 3: basta buscar raíces. Si un polinomio de grado 2 o 3 se factoriza, uno de los factores tiene grado 1, o sea hay una raíz en K. *Sin raíces* $\Rightarrow$ *irreducible.* **Cuidado: esto es falso desde grado 4**, donde x^4+1 podría partirse en dos cuadráticos sin tener ninguna raíz.
2. Criterio de las raíces racionales (coeficientes enteros): descrito arriba.
3. Criterio de Eisenstein. Si existe un primo p tal que $p \nmid a_n$, $p \mid a_i$ para todo $i < n$, y $p^2 \nmid a_0$, entonces f es irreducible sobre $\mathbb{Q}$.
4. Reducción módulo p. Si al reducir los coeficientes módulo un primo p (que no divida al coeficiente líder) el resultado es irreducible en $\mathbb{F}_p[x]$, entonces f era irreducible sobre $\mathbb{Q}$. Es potentísimo porque en $\mathbb{F}_p[x]$ solo hay finitos polinomios que probar.
5. Cambio de variable. Sustituir $x \to x+1$ no cambia la irreducibilidad, y a veces convierte un polinomio intratable en uno donde Eisenstein sí funciona.

Ejemplo 24.43 (Cinco polinomios decididos, uno por método)**.** x^3-2 **sobre** $\mathbb{Q}$: Eisenstein con $p=2$ ($2 \nmid 1$, $2 \mid 0, 0, -2$, $4 \nmid -2$). Irreducible. También por raíces: $\sqrt[3]{2}$ no es racional.

x^3-x-1 **sobre** $\mathbb{Q}$: grado 3 sin raíces racionales (visto arriba). Irreducible.

$\Phi_5 = x^4+x^3+x^2+x+1$ **sobre** $\mathbb{Q}$: Eisenstein no aplica directamente. Sustituye $x \to x+1$:

$$\Phi_5(x+1) = \frac{(x+1)^5-1}{x} = x^4+5x^3+10x^2+10x+5,$$

y ahora Eisenstein con $p=5$ funciona ($5 \nmid 1$, 5 divide al resto, $25 \nmid 5$). Irreducible.

x^4+1 **sobre** $\mathbb{Q}$: no tiene raíces reales, y tampoco se parte en cuadráticos racionales (habría que resolver $x^4+1 = (x^2+ax+b)(x^2-ax+c)$ y sale $a^2 = \pm 2$, irracional). Irreducible sobre $\mathbb{Q}$. *Pero es reducible módulo todo primo*, lo que lo convierte en el contraejemplo favorito para mostrar que el criterio 4 solo funciona en un sentido.

x^4-10x^2+1 **sobre** $\mathbb{Q}$: irreducible (es el polinomio mínimo de $\sqrt{2}+\sqrt{3}$, como se verifica desarrollando $((\sqrt{2}+\sqrt{3})^2-5)^2 = 24$). Un caso donde la ausencia de raíces no bastaría y hace falta el argumento de grado 4 completo.

Todas estas factorizaciones fueron comprobadas con álgebra computacional.

24.14. Construir cuerpos finitos con las manos

Ya tenemos la receta: cocientar por un irreducible. Vamos a usarla dos veces y a escribir las tablas completas.

Ejemplo 24.44 ($\mathbb{F}_4$: el cuerpo de cuatro elementos)**.** Sobre $\mathbb{F}_2 = \{0, 1\}$, el polinomio x^2+x+1 no tiene raíces ($0^2 + 0 + 1 = 1$, $1^2 + 1 + 1 = 1$), y al ser de grado 2 eso basta: es irreducible. Cocientamos:

$$\mathbb{F}_4 = \mathbb{F}_2[x]/(x^2 + x + 1) = \{0,\ 1,\ a,\ a + 1\}, \qquad a^2 = a + 1,$$

donde a es la clase de x y la relación sale de $x^2 + x + 1 = 0$, recordando que en característica 2 restar es sumar.

Tabla del producto (calculada reduciendo con $a^2 = a + 1$ y verificada por ordenador):

$\cdot$	0	1	a	$a+1$
0	0	0	0	0
1	0	1	a	$a+1$
a	0	a	$a+1$	1
$a+1$	0	$a+1$	1	a

Comprobaciones. Cada fila no nula contiene un 1: es cuerpo. ✓ Además $a^2 = a + 1$, $a^3 = a \cdot a^2 = a(a + 1) = a^2 + a = (a + 1) + a = 1$: el elemento a tiene orden 3 en el grupo multiplicativo, que tiene $4 - 1 = 3$ elementos. Luego $\mathbb{F}_4^\times$ es cíclico generado por a. ✓

Y un aviso importante: $\mathbb{F}_4$ *no es* $\mathbb{Z}_4$. En $\mathbb{Z}_4$, el 2 es divisor de cero ($2 \cdot 2 = 0$) y no hay cuerpo posible. Los cuerpos de p^n elementos con $n > 1$ nunca se obtienen como $\mathbb{Z}_{p^n}$; hay que construirlos así.

Ejemplo 24.45 ($\mathbb{F}_8$ y sus potencias, listadas)**.** Sobre $\mathbb{F}_2$, el polinomio $x^3 + x + 1$ tampoco tiene raíces, y al ser de grado 3 es irreducible. Con b la clase de x (relación $b^3 = b + 1$), las potencias sucesivas son:

$$b^0 = 1, \quad b^1 = b, \quad b^2 = b^2, \quad b^3 = b+1, \quad b^4 = b^2+b, \quad b^5 = b^2+b+1, \quad b^6 = b^2+1, \quad b^7 = 1.$$

Siete potencias distintas antes de volver al 1: b genera los $8 - 1 = 7$ elementos no nulos, así que $\mathbb{F}_8^\times \cong \mathbb{Z}_7$. ✓ (Cálculo verificado por ordenador.)

Esta tabla de "logaritmos discretos" es cómo se implementan de verdad los cuerpos finitos en los códigos correctores: multiplicar se reduce a sumar exponentes módulo 7.

Teorema 24.46 (Estructura de los cuerpos finitos)**.** *Todo cuerpo finito tiene p^n elementos con p primo. Para cada p^n existe uno y* solo *uno salvo isomorfismo, denotado $\mathbb{F}_{p^n}$. Su grupo multiplicativo $\mathbb{F}_q^\times$ es siempre cíclico. Y en característica p vale $(u + v)^p = u^p + v^p$, porque todos los coeficientes binomiales intermedios $\binom{p}{k}$ con $0 < k < p$ son múltiplos de p.*

Observación 24.47. La última identidad, que un estudiante distraído escribiría por error, aquí es un teorema y define el *automorfismo de Frobenius* $u \mapsto u^p$. Es la pieza que hace funcionar toda la teoría de cuerpos finitos, y en el capítulo 26 resultará ser el generador del grupo de Galois de $\mathbb{F}_{p^n}$ sobre $\mathbb{F}_p$.

24.15. Cuando la factorización única se rompe

Idea clave

Intuición primero. Que todo entero se factorice en primos *de una sola manera* nos parece evidente porque llevamos toda la vida con ello. No lo es. Es un teorema, depende de tener división con resto, y en cuanto pisas un anillo un poco más grande deja de valer. Descubrirlo fue lo que hundió varias "demostraciones" del último teorema de Fermat en el siglo XIX, y lo que obligó a inventar la teoría de ideales.

Ejemplo 24.48 ($\mathbb{Z}[\sqrt{-5}]$: dos factorizaciones del mismo número)**.** En el anillo $\mathbb{Z}[\sqrt{-5}] = \{a + b\sqrt{-5} : a, b \in \mathbb{Z}\}$, considera

$$6 = 2 \cdot 3 = (1 + \sqrt{-5})(1 - \sqrt{-5}).$$

La segunda igualdad se comprueba directamente: $(1 + \sqrt{-5})(1 - \sqrt{-5}) = 1 - (-5) = 6$. ✓

¿Son los cuatro factores irreducibles? Usamos la norma $N(a + b\sqrt{-5}) = a^2 + 5b^2$, que es multiplicativa. Los valores relevantes son $N(2) = 4$, $N(3) = 9$, $N(1 \pm \sqrt{-5}) = 6$.

Si $2 = \alpha\beta$ con ambos no unidades, entonces $N(\alpha)N(\beta) = 4$ con ambos factores mayores que 1, luego $N(\alpha) = 2$. Pero $a^2 + 5b^2 = 2$ no tiene soluciones enteras ($b = 0$ da $a^2 = 2$; $|b| \geq 1$ da al menos 5). Luego 2 es irreducible. Igual con 3: haría falta norma 3, y $a^2 + 5b^2 = 3$ tampoco tiene solución. Y con $1 \pm \sqrt{-5}$: harían falta normas 2 y 3.

¿Son asociados entre sí? No: las unidades de este anillo son solo ± 1 (norma 1), y 2 no es $\pm(1 + \sqrt{-5})$.

Conclusión: el 6 tiene dos factorizaciones en irreducibles genuinamente distintas. La factorización única *falla*.

Observación 24.49. Lo que se rompe exactamente es la equivalencia entre "irreducible" y "primo". En $\mathbb{Z}$, un número es irreducible si y solo si es primo en el sentido fuerte ($p \mid ab \Rightarrow p \mid a$ o $p \mid b$). En $\mathbb{Z}[\sqrt{-5}]$, el 2 es irreducible pero no primo: divide a $6 = (1 + \sqrt{-5})(1 - \sqrt{-5})$ sin dividir a ninguno de los dos factores.

La salida, hallada por Kummer y Dedekind, fue genial: si los *elementos* no se factorizan de forma única, quizá los *ideales* sí. Y así es —en los anillos de enteros de cuerpos de números, todo ideal se factoriza de manera única en ideales primos—. La palabra "ideal" viene literalmente de "número ideal", el objeto que Kummer inventó para reparar este desperfecto.

24.16. La jerarquía completa

Regla: la escalera de anillos, de menos a más

anillo conmutativo ⊃ dominio ⊃ DFU ⊃ DIP ⊃ dominio euclídeo ⊃ cuerpo

- **Dominio**: sin divisores de cero. Se puede cancelar.
- **DFU** (factorización única): todo elemento no nulo ni unidad se factoriza en irreducibles de forma única salvo orden y unidades.

- **DIP** (ideales principales): todo ideal está generado por un elemento.
- **Euclídeo**: existe una función "tamaño" que permite dividir con resto.
- **Cuerpo**: todo no nulo se invierte.

Cada inclusión es estricta, y hay un contraejemplo estándar para cada escalón:

Anillo	Es	No es
$\mathbb{Z}_6$	anillo	dominio (por $2 \cdot 3 = 0$)
$\mathbb{Z}[\sqrt{-5}]$	dominio	DFU (por el 6)
$\mathbb{Z}[x]$	DFU	DIP (por el ideal $(2, x)$)
$\mathbb{Z}$	euclídeo	cuerpo (el 2 no se invierte)
$\mathbb{Q}$, $\mathbb{F}_p$, $\mathbb{F}_4$	cuerpo	—

Idea clave

Los dos ejemplos que hay que tener siempre en la cabeza son $\mathbb{Z}$ y $K[x]$: ambos son euclídeos, y de esa sola propiedad se deduce en cascada todo lo demás —Bézout, ideales principales, factorización única—. Si alguna vez olvidas por qué un resultado es cierto en polinomios, tradúcelo a enteros y recuerda la demostración de allí; casi siempre es la misma.

24.17. Dónde aparecen los anillos de verdad

Criptografía. RSA vive en $\mathbb{Z}_n$ con $n = pq$; toda su seguridad depende de que ese anillo *no* sea cuerpo y de que descomponer n sea difícil. Las curvas elípticas trabajan sobre $\mathbb{F}_p$ y $\mathbb{F}_{2^n}$.

Códigos correctores. Los códigos Reed–Solomon —los de los CD, los DVD, los códigos QR y las sondas espaciales— son ideales dentro de $\mathbb{F}_q[x]/(x^n - 1)$. Que el disco rayado siga sonando es un teorema sobre raíces de polinomios en cuerpos finitos.

Geometría algebraica. A cada sistema de ecuaciones polinómicas se le asocia el anillo cociente $K[x_1, \ldots, x_n]/I$, y las propiedades geométricas de la figura se leen como propiedades algebraicas del anillo. Es la traducción de Descartes llevada al límite.

Teoría de números. Los anillos de enteros de cuerpos de números, con su factorización única de ideales, son el marco donde se demuestran resultados sobre $\mathbb{Z}$ que en $\mathbb{Z}$ solo no se ven.

Análisis funcional. $C[0, 1]$ con la norma del supremo es un anillo y un espacio de Banach a la vez; sus ideales maximales corresponden exactamente a los puntos del intervalo, resultado que abre la teoría de las álgebras de Banach.

24.18. Taller de ejemplos resueltos

Ejercicio 24.1. Encuentra todas las unidades y todos los divisores de cero de $\mathbb{Z}_{12}$.

Solución. Unidades: los coprimos con 12, es decir $\{1,5,7,11\}$; hay $\varphi(12)=4$. Sus inversos son cada uno sí mismo, porque $5^2=25=1$, $7^2=49=1$, $11^2=121=1$ módulo 12.

Divisores de cero: los no nulos que no son unidades, o sea $\{2,3,4,6,8,9,10\}$. Verificación de uno: $4\cdot 3=12=0$ con ambos no nulos. ✓ En $\mathbb{Z}_n$ todo elemento no nulo es unidad o divisor de cero, porque si $d=\gcd(a,n)>1$ entonces $a\cdot(n/d)\equiv 0$ con $n/d\neq 0$. ■

Ejercicio 24.2. Calcula 3^{-1} en $\mathbb{Z}_{17}$ usando el algoritmo de Euclides extendido.

Solución. $17=5\cdot 3+2$, $3=1\cdot 2+1$. Retrocediendo: $1=3-2=3-(17-5\cdot 3)=6\cdot 3-17$. Luego $6\cdot 3\equiv 1 \pmod{17}$ y $3^{-1}=6$. Comprobación: $18=17+1$. ✓ ■

Ejercicio 24.3. Divide x^5+x^3-2 entre x^2+1 y comprueba el resultado.

Solución. $x^5/x^2=x^3$; $x^3(x^2+1)=x^5+x^3$; restando queda -2, de grado $0<2$. Paramos ya.

$$x^5+x^3-2=x^3(x^2+1)-2.$$

Comprobación: $x^3\cdot x^2+x^3=x^5+x^3$, y menos 2. ✓ (El cociente es x^3 y el resto -2; verificado simbólicamente.) ■

Ejercicio 24.4. Decide si x^2+x+1 y x^2+1 son irreducibles sobre $\mathbb{F}_2$ y sobre $\mathbb{F}_3$.

Solución. Grado 2: basta buscar raíces.

x^2+x+1 sobre $\mathbb{F}_2$: $f(0)=1$, $f(1)=1+1+1=1$. Sin raíces, **irreducible**. Sobre $\mathbb{F}_3$: $f(1)=3=0$, luego **reducible**, y de hecho $x^2+x+1=(x-1)^2$ en $\mathbb{F}_3$.

x^2+1 sobre $\mathbb{F}_2$: $f(1)=1+1=0$, **reducible**, $x^2+1=(x+1)^2$. Sobre $\mathbb{F}_3$: $f(0)=1$, $f(1)=2$, $f(2)=5=2$. Sin raíces, **irreducible**.

Moraleja: los mismos dos polinomios intercambian su comportamiento al cambiar de cuerpo. Es exactamente el aviso de la sección de irreducibilidad, con números delante. ■

Ejercicio 24.5. Construye un cuerpo de nueve elementos y da un elemento que genere su grupo multiplicativo.

Solución. Como $9=3^2$, tomamos $\mathbb{F}_3$ y un irreducible de grado 2. Por el ejercicio anterior, x^2+1 sirve. Entonces

$$\mathbb{F}_9=\mathbb{F}_3[x]/(x^2+1)=\{a+bi : a,b\in\mathbb{F}_3\},\qquad i^2=-1=2.$$

Nueve elementos: $0,1,2,i,1+i,2+i,2i,1+2i,2+2i$. ✓

El grupo multiplicativo tiene 8 elementos. Probemos $g=1+i$: $g^2=1+2i+i^2=1+2i-1=2i$; $g^4=(2i)^2=4i^2=1\cdot(-1)=-1=2$; $g^8=2^2=4=1$. Como $g^4=2\neq 1$ y $g^2=2i\neq 1$, el orden no divide a 4, luego es 8: $1+i$ genera $\mathbb{F}_9^\times$. ✓ ■

Ejercicio 24.6. Demuestra que en $\mathbb{Z}[i]$ el número 5 no es irreducible, pero el 3 sí.

Solución. $5=(2+i)(2-i)$, y ninguno de los dos factores es unidad porque sus normas valen $4+1=5\neq 1$. Luego 5 es **reducible**.

Para el 3: si $3=\alpha\beta$ con ambos no unidades, tomando normas $9=N(\alpha)N(\beta)$ con ambas mayores que 1, luego $N(\alpha)=3$, es decir $a^2+b^2=3$. Pero las sumas de dos cuadrados menores que 3 son $0,1,2$: no hay solución. Luego 3 es **irreducible**.

Este es el germen del teorema de Fermat de los dos cuadrados: un primo impar p deja de ser irreducible en $\mathbb{Z}[i]$ exactamente cuando $p\equiv 1 \pmod 4$, que es cuando p es suma de dos cuadrados. Y en efecto $5\equiv 1$ y $3\equiv 3$. ✓ ■

Ejercicio 24.7. Muestra que $\mathbb{Z}_2\times\mathbb{Z}_3\cong\mathbb{Z}_6$ como anillos, y que $\mathbb{Z}_2\times\mathbb{Z}_2\not\cong\mathbb{Z}_4$.

Solución. **Primera parte.** Define $\varphi : \mathbb{Z}_6 \to \mathbb{Z}_2 \times \mathbb{Z}_3$ por $\varphi(a) = (a \bmod 2,\ a \bmod 3)$. Respeta sumas y productos porque cada componente lo hace, y manda $1 \mapsto (1,1)$. Es inyectiva: si $a \equiv 0$ módulo 2 y módulo 3, entonces $6 \mid a$. Como ambos conjuntos tienen 6 elementos, es biyección. Este es el teorema chino del resto en versión de anillos.

Segunda parte. En $\mathbb{Z}_2 \times \mathbb{Z}_2$ todo elemento cumple $u + u = 0$ y además $(1,0)(0,1) = (0,0)$: hay divisores de cero. En $\mathbb{Z}_4$ también hay divisores de cero ($2 \cdot 2 = 0$), así que ese criterio no basta; pero en $\mathbb{Z}_4$ el elemento 1 tiene orden aditivo 4, mientras que en $\mathbb{Z}_2 \times \mathbb{Z}_2$ todo elemento tiene orden aditivo ≤ 2. Luego ni siquiera son isomorfos como grupos aditivos.

Moraleja: $\mathbb{Z}_m \times \mathbb{Z}_n \cong \mathbb{Z}_{mn}$ exactamente cuando $\gcd(m,n) = 1$, igual que en el capítulo 23. ■

Ejercicio 24.8. Verifica que $\frac{1}{1+\sqrt{2}} = \sqrt{2} - 1$ y calcula $\frac{1}{1+\sqrt[3]{2}}$ en $\mathbb{Q}(\sqrt[3]{2})$.

Solución. **Primera:** $(1+\sqrt{2})(\sqrt{2}-1) = \sqrt{2} - 1 + 2 - \sqrt{2} = 1$. ✓ El truco escolar de "racionalizar multiplicando por el conjugado" es, en realidad, calcular un inverso en el cuerpo $\mathbb{Q}(\sqrt{2})$.

Segunda: con $t = \sqrt[3]{2}$ se tiene $t^3 = 2$, luego $t^3 + 1 = 3$, y factorizando la suma de cubos,

$$(1+t)(1-t+t^2) = 1 + t^3 = 3 \quad \Longrightarrow \quad \frac{1}{1+t} = \frac{1-t+t^2}{3} = \frac{1 - \sqrt[3]{2} + \sqrt[3]{4}}{3}.$$

Comprobación numérica: $\sqrt[3]{2} \approx 1{,}259921$, así que $1/(1+\sqrt[3]{2}) \approx 0{,}4424933$; y $(1 - 1{,}259921 + 1{,}587401)/3 \approx 0{,}4424933$. ✓

Fíjate en lo que acaba de pasar: el "conjugado" que hay que usar no se adivina, se *calcula* con Bézout aplicado a $1 + x$ y al polinomio mínimo $x^3 - 2$. Ese es el método general para invertir en cualquier extensión. ■

Ejercicio 24.9. Encuentra el polinomio mínimo de $\sqrt{2} + \sqrt{3}$ sobre $\mathbb{Q}$.

Solución. Sea $\alpha = \sqrt{2} + \sqrt{3}$. Entonces $\alpha^2 = 2 + 2\sqrt{6} + 3 = 5 + 2\sqrt{6}$, luego $\alpha^2 - 5 = 2\sqrt{6}$ y elevando otra vez: $(\alpha^2 - 5)^2 = 24$, es decir

$$\alpha^4 - 10\alpha^2 + 1 = 0.$$

El polinomio $x^4 - 10x^2 + 1$ es irreducible sobre $\mathbb{Q}$ (verificado con álgebra computacional; a mano: no tiene raíces racionales por el criterio ± 1, y no se parte en cuadráticos racionales). Luego es el polinomio mínimo y $[\mathbb{Q}(\sqrt{2}+\sqrt{3}) : \mathbb{Q}] = 4$.

Observación que se usará en el capítulo 26: como $\sqrt{2}$ y $\sqrt{3}$ se recuperan a partir de α —basta calcular $\alpha^3 = 11\sqrt{2} + 9\sqrt{3}$ y combinarlo con α—, resulta que $\mathbb{Q}(\sqrt{2},\sqrt{3}) = \mathbb{Q}(\sqrt{2}+\sqrt{3})$: dos generadores se pueden fundir en uno. Eso es el teorema del elemento primitivo en miniatura. ■

24.19. Cierre: qué llevarte de este capítulo

Si de todo lo anterior te quedas con cinco frases, que sean estas.

Un anillo es $\mathbb{Z}$ sin la garantía de dividir. Todo lo demás son preguntas sobre cuánto de $\mathbb{Z}$ sobrevive.

Los divisores de cero son el enemigo. Su presencia rompe la cancelación, rompe la cota de raíces de un polinomio y rompe el razonamiento "producto cero implica factor cero" que usas desde la secundaria. Los dominios de integridad son, precisamente, los sitios donde el álgebra escolar sigue siendo válida.

$\mathbb{Z}$ y $K[x]$ son el mismo teorema dos veces. División con resto, Euclides, Bézout, factorización única: si lo recuerdas en uno, lo tienes en el otro.

Cocientar por un irreducible fabrica cuerpos. Es la receta que produce $\mathbb{C}$ desde $\mathbb{R}$, $\mathbb{F}_4$ desde $\mathbb{F}_2$ y toda la teoría de extensiones del capítulo 26.

La factorización única es un privilegio, no un derecho. Se pierde en $\mathbb{Z}[\sqrt{-5}]$, y repararla —pasando de números a ideales— es lo que dio origen al álgebra moderna.

Con esto estás en condiciones de leer el capítulo 25, que retoma todo lo anterior con notación compacta y añade la maquinaria pesada, y después el capítulo 26, donde grupos y cuerpos se encuentran.

24.20. Característica, el primer teorema de isomorfía y el cuerpo de fracciones

Característica y cuerpo primo

Definición 24.50. La **característica** de un anillo con unidad es el menor $n > 0$ con $n \cdot 1 = 0$, y 0 si no existe.

Teorema 24.51. *En un dominio de integridad, la característica es* 0 *o un primo. Y todo cuerpo contiene un* ***cuerpo primo****: una copia de* $\mathbb{Q}$ *si la característica es* 0*, y de* $\mathbb{F}_p$ *si es* p*.*

Demostración. Si $\operatorname{car} R = n = ab$ con $1 < a, b < n$, entonces $(a \cdot 1)(b \cdot 1) = n \cdot 1 = 0$ con ambos factores no nulos, contra la hipótesis de dominio. Para lo segundo, el homomorfismo $\mathbb{Z} \to K$ que manda $1 \mapsto 1$ tiene núcleo (0) o (p); en el primer caso su imagen genera una copia de $\mathbb{Q}$, en el segundo de $\mathbb{F}_p$. □

Teorema 24.52 (El sueño del estudiante)**.** *Si* $\operatorname{car} R = p$ *y* R *es conmutativo, entonces*

$$(u + v)^p = u^p + v^p,$$

y la aplicación $F(u) = u^p$ *—el* ***endomorfismo de Frobenius****— es un homomorfismo de anillos.*

Demostración. En el binomio, $\binom{p}{k} = \frac{p!}{k!(p-k)!}$ tiene un p en el numerador que ningún factor del denominador cancela cuando $0 < k < p$, porque p es primo y todos los factores son menores que p. Luego todos los términos intermedios son múltiplos de p, es decir, nulos. □

Ejemplo 24.53. En $\mathbb{F}_2[x]$: $(x+1)^2 = x^2 + 2x + 1 = x^2 + 1$. En $\mathbb{F}_4$ la característica es 2 aunque el cuerpo tenga cuatro elementos, y su cuerpo primo es $\mathbb{F}_2$: característica y cardinal no son lo mismo, y el segundo es siempre una potencia del primero. Y $\operatorname{car}(\mathbb{Z}_6) = 6$ no es primo precisamente porque $\mathbb{Z}_6$ no es dominio.

El Frobenius es además el generador del grupo de Galois de $\mathbb{F}_{p^n}$ sobre $\mathbb{F}_p$: la razón última de que los cuerpos finitos sean tan manejables.

El primer teorema de isomorfía para anillos

Teorema 24.54. *Si $\varphi : R \to S$ es homomorfismo de anillos, entonces* $\ker\varphi$ *es un ideal, todo ideal es núcleo de la proyección al cociente, y*

$$R/\ker\varphi \;\cong\; \operatorname{im}\varphi.$$

Esto convierte en una línea todas las identificaciones que el capítulo hacía a mano.

Ejemplo 24.55 (Tres cocientes, de un golpe)**. (a)** $\mathbb{Z}/(n) \cong \mathbb{Z}_n$, con φ la reducción módulo n.

(b) $\mathbb{R}[x]/(x^2+1) \cong \mathbb{C}$**.** Tómese la evaluación $\mathrm{ev}_i : \mathbb{R}[x] \to \mathbb{C}$, $f \mapsto f(i)$. Es sobreyectiva —$a+bx \mapsto a+bi$— y su núcleo son los polinomios con raíz i, que por tener coeficientes reales tienen también raíz $-i$, luego son múltiplos de x^2+1. El teorema hace el resto.

(c) $\mathbb{F}_2[x]/(x^3+x+1) \cong \mathbb{F}_8$, con la evaluación en un generador. La construcción "a mano" de las ocho clases queda sustituida por un argumento estructural.

El cuerpo de fracciones

Teorema 24.56. *Todo dominio de integridad R se sumerge en un cuerpo* $\mathrm{Frac}(R)$*, construido como el conjunto de pares (a,b) con $b \neq 0$ módulo la relación $(a,b) \sim (c,d) \iff ad = bc$, con las operaciones habituales. Es el menor: todo homomorfismo inyectivo de R en un cuerpo se factoriza por él.*

La comprobación de que la relación es de equivalencia usa *exactamente* que R es dominio: la transitividad requiere cancelar. Sin esa hipótesis la construcción falla, y hay que sustituirla por la localización general.

Ejemplo 24.57.

$$\mathrm{Frac}(\mathbb{Z}) = \mathbb{Q}, \qquad \mathrm{Frac}(K[x]) = K(x), \qquad \mathrm{Frac}(\mathbb{Z}[i]) = \mathbb{Q}(i).$$

El primero es la construcción *oficial* de los racionales; el segundo es el cuerpo de funciones racionales, que el capítulo de Galois usa sin más; y en el tercero, $\dfrac{1}{2+i} = \dfrac{2-i}{5}$, es decir, todo elemento de $\mathbb{Q}(i)$ es un cociente de enteros de Gauss.

24.21. El lema de Gauss, y polinomios en varias variables

Hemos aplicado Eisenstein y el criterio de la raíz racional a polinomios de coeficientes *enteros* para concluir irreducibilidad en $\mathbb{Q}[x]$. Ese salto de $\mathbb{Z}[x]$ a $\mathbb{Q}[x]$ es exactamente el contenido del lema de Gauss, y hay que justificarlo.

Definición 24.58. El **contenido** $c(f)$ de $f \in \mathbb{Z}[x]$ es el máximo común divisor de sus coeficientes. f es **primitivo** si $c(f) = 1$.

Teorema 24.59 (Lema de Gauss)**.** *El producto de primitivos es primitivo. En consecuencia* $c(fg) = c(f)c(g)$*, y un polinomio de* $\mathbb{Z}[x]$ *de grado* ≥ 1 *es irreducible en* $\mathbb{Z}[x]$ *si y solo si es primitivo e irreducible en* $\mathbb{Q}[x]$*.*

Demostración. Sean f, g primitivos y supóngase que un primo p divide a todos los coeficientes de fg. Reduciendo módulo p, se tendría $\bar{f}\bar{g} = 0$ en $\mathbb{F}_p[x]$, que es dominio por serlo $\mathbb{F}_p$; luego $\bar{f} = 0$ o $\bar{g} = 0$, es decir, p divide a todos los coeficientes de f o de g, contra la primitividad. □

Ejemplo 24.60. $2x^2+4x+2$ tiene contenido 2: es reducible en $\mathbb{Z}[x]$ —se factoriza como $2\cdot(x+1)^2$— aunque su parte primitiva x^2+2x+1 sea la que decide en $\mathbb{Q}[x]$. Y la multiplicatividad se ve en $(3x+3)(2x+4) = 6(x+1)(x+2)$, con $c = 3$, $c = 2$ y $c = 6$. ✓

La consecuencia práctica es la que usábamos: si $f \in \mathbb{Z}[x]$ es primitivo y no factoriza en $\mathbb{Z}[x]$, tampoco factoriza en $\mathbb{Q}[x]$. Sin el lema, un polinomio entero podría en principio partirse en factores de coeficientes fraccionarios, y todos los criterios del capítulo quedarían en el aire.

Teorema 24.61 (Gauss, versión fuerte)**.** *Si* R *es DFU, entonces* $R[x]$ *también lo es. Por inducción,* $R[x_1, \ldots, x_n]$ *es DFU, y en particular* $\mathbb{Z}[x]$ *y* $K[x_1, \ldots, x_n]$ *lo son.*

Varias variables y series formales

Ejemplo 24.62 ($K[x, y]$ es DFU y no es DIP)**.** El ideal (x, y) no es principal. Si $(x, y) = (f)$, entonces $f \mid x$ y $f \mid y$; por grados en cada variable, f tendría que ser constante, luego una unidad, luego (f) sería todo el anillo. Pero $1 \notin (x, y)$, porque todo elemento de ese ideal se anula en el origen. Contradicción.

Es el gemelo exacto de $(2, x) \subset \mathbb{Z}[x]$: en ambos casos el anillo es DFU por Gauss y falla el ser DIP por el mismo motivo, que hace falta más de un generador para describir "anularse en un punto". Y el cociente $K[x, y]/(x, y) \cong K$ muestra que sí es maximal. Toda la geometría algebraica vive en esta distinción.

Definición 24.63 (Series formales)**.** $K[[x]]$ es el anillo de las expresiones $\sum_{n\geq 0} a_n x^n$ sin ninguna condición de convergencia, con la suma y el producto de Cauchy.

Teorema 24.64. $f = \sum a_n x^n$ *es unidad de* $K[[x]]$ *si y solo si* $a_0 \neq 0$*. Además,* $K[[x]]$ *es un anillo local, con único ideal maximal* (x)*, y es DIP: todo ideal es* (x^k)*.*

Ejemplo 24.65. En $K[x]$, el polinomio $1 - x$ no es invertible —su inverso tendría que ser de grado negativo—. En $K[[x]]$ sí lo es:

$$\frac{1}{1-x} = 1 + x + x^2 + x^3 + \cdots,$$

y la igualdad es *formal*: se comprueba multiplicando y viendo que todos los coeficientes salvo el primero se cancelan. No hay ninguna cuestión de convergencia porque no hay ningún x que sustituir.

Ese es el fundamento del método de las funciones generatrices: manipular series como objetos algebraicos, resolver, y solo al final —si se quiere— preguntarse por la convergencia. La sucesión de Fibonacci se obtiene así de $\frac{x}{1-x-x^2}$ sin justificar ni un límite.

24.22. Cadenas ascendentes: anillos noetherianos y la base de Hilbert

La jerarquía cuerpo $\subset$ DIP $\subset$ DFU $\subset$ dominio se cerró sin mencionar la condición que en realidad gobierna todo lo que viene después —geometría algebraica, álgebra conmutativa, teoría de módulos—: que las cadenas de ideales no crezcan indefinidamente.

Definición 24.66 (Anillo noetheriano)**.** Un anillo conmutativo R es *noetheriano* si toda cadena ascendente de ideales

$$I_1 \subset I_2 \subset I_3 \subset \cdots$$

se estabiliza, es decir, existe N con $I_N = I_{N+1} = \cdots$

Proposición 24.67 (Tres formulaciones equivalentes)**.** *Para R conmutativo son equivalentes:*

1. *R es noetheriano;*
2. *todo ideal de R es finitamente generado;*
3. *toda familia no vacía de ideales tiene un elemento maximal para la inclusión.*

Demostración. (1)⇒(2). Sea I un ideal y supóngase que no es finitamente generado. Elíjase $a_1 \in I$, y recursivamente $a_{n+1} \in I \setminus (a_1, \ldots, a_n)$, lo cual es posible porque ningún $(a_1, \ldots, a_n)$ agota I. La cadena $(a_1) \subsetneq (a_1, a_2) \subsetneq \cdots$ no se estabiliza: contradicción.

(2)⇒(1). Dada una cadena, su unión $I = \bigcup I_n$ es un ideal (la unión de una cadena sí lo es), luego $I = (x_1, \ldots, x_s)$ con cada x_i en algún I_{n_i}. Tomando $N = \max n_i$, todos están en I_N, luego $I = I_N$ y la cadena se para ahí.

(1)⇔(3). Si una familia no tuviera maximal, se construiría una cadena estrictamente creciente infinita; recíprocamente, aplicando (3) a la familia de los términos de una cadena se obtiene su estabilización. □

Ejemplo 24.68 (Quién es y quién no es noetheriano)**.** Todo DIP lo es (los ideales están generados por un elemento), y en particular $\mathbb{Z}$ y $K[x]$. En cambio, el anillo de polinomios en infinitas variables $K[x_1, x_2, x_3, \ldots]$ no lo es: la cadena $(x_1) \subsetneq (x_1, x_2) \subsetneq \cdots$ crece para siempre. Tampoco lo es el anillo de las funciones continuas en $[0, 1]$, ni el de los enteros algebraicos $\overline{\mathbb{Z}}$ (donde $(2) \subsetneq (\sqrt{2}) \subsetneq (\sqrt[4]{2}) \subsetneq \cdots$). La condición noetheriana es exactamente lo que separa el álgebra manejable de la que no lo es.

Teorema 24.69 (Base de Hilbert)**.** *Si R es noetheriano, también lo es $R[x]$. En consecuencia, $R[x_1, \ldots, x_n]$ es noetheriano para todo n, y en particular $K[x_1, \ldots, x_n]$ y $\mathbb{Z}[x_1, \ldots, x_n]$ lo son.*

Demostración. Sea $I \subset R[x]$ un ideal; hay que probar que es finitamente generado. Para cada $d \geq 0$ sea

$$L_d = \{0\} \cup \{\text{coeficiente director de } f : \ f \in I, \ \deg f = d\}.$$

Cada L_d es un ideal de R (si $f, g \in I$ tienen grado d y coeficientes directores a, b, o bien $a+b=0$ o bien $f+g$ tiene grado d y coeficiente director $a+b$; y rf da ra), y $L_d \subset L_{d+1}$ porque multiplicar por x sube el grado sin cambiar el coeficiente director. Como R es noetheriano, la cadena $L_0 \subset L_1 \subset \cdots$ se estabiliza en algún L_D.

Para cada $d \leq D$, escríbase $L_d = (a_{d,1}, \ldots, a_{d,k_d})$ —finitamente generado por la proposición 24.67— y elíjase $f_{d,j} \in I$ de grado d con coeficiente director $a_{d,j}$. Afirmación: esa colección finita $\{f_{d,j}\}_{d \leq D,\ j \leq k_d}$ genera I.

Sea J el ideal que generan, contenido en I, y supóngase $J \neq I$. Entre los elementos de $I \setminus J$ tómese uno, f, de *grado mínimo* n, con coeficiente director $a \in L_n$.

Si $n \leq D$: se escribe $a = \sum_j r_j a_{n,j}$, y entonces $g = \sum_j r_j f_{n,j} \in J$ tiene grado n y el mismo coeficiente director, luego $f - g \in I$ tiene grado $< n$. Por minimalidad de n, $f - g \in J$, y por tanto $f \in J$: contradicción.

Si $n > D$: como $L_n = L_D$, se escribe $a = \sum_j r_j a_{D,j}$ y se toma $g = x^{\,n-D} \sum_j r_j f_{D,j} \in J$, de grado n y coeficiente director a. Otra vez $f - g \in I$ tiene grado $< n$, luego está en J, luego $f \in J$: contradicción.

Así $I = J$ es finitamente generado. □

Observación 24.70 (Por qué esto importa tanto). El teorema dice que *todo* sistema de ecuaciones polinómicas, por infinito que sea el conjunto de ecuaciones que uno escriba, equivale a un sistema *finito*: el ideal que generan está generado por finitos polinomios. Esa es la razón por la que la geometría algebraica es posible, y la razón por la que el algoritmo de Buchberger para bases de Gröbner termina. Cuando Hilbert lo demostró en 1888, sin exhibir los generadores, Gordan protestó con la frase célebre "esto no es matemática, es teología"; la posteridad le ha dado la razón a Hilbert, aunque también a Gordan en un sentido práctico: saber que la base existe no dice cómo calcularla, y calcularla es computacionalmente costoso.

24.23. Cuerpos finitos: existencia y unicidad

Se afirmó que para cada potencia de primo $q = p^n$ existe un cuerpo de q elementos y que es único salvo isomorfismo, y esa unicidad se usó al identificar $\mathbb{F}_2[x]/(x^3 + x + 1)$ con $\mathbb{F}_8$. Se demuestra ahora.

Teorema 24.71 (Existencia y unicidad de $\mathbb{F}_q$)**.** *Sea p primo y $n \geq 1$, y $q = p^n$.*

1. *Existe un cuerpo con exactamente q elementos.*
2. *Dos cuerpos con q elementos son isomorfos.*
3. *No existe ningún cuerpo finito cuyo cardinal no sea potencia de un primo.*

Demostración. (3) Si F es un cuerpo finito, su característica es un primo p (si fuera ab compuesto, habría divisores de cero), luego contiene a $\mathbb{F}_p$ y es un $\mathbb{F}_p$-espacio vectorial de dimensión finita n; por tanto $|F| = p^n$.

(1) Considérese el polinomio $f(x) = x^q - x$ sobre $\mathbb{F}_p$ y sea L un cuerpo de descomposición suyo (existe, por la construcción de Kronecker del capítulo 26). Sea

$$F = \{\alpha \in L : \ \alpha^q = \alpha\},$$

el conjunto de sus raíces. Es un *subcuerpo*: la aplicación $\phi(\alpha) = \alpha^p$ es un homomorfismo de anillos —el "sueño del estudiante" $(\alpha+\beta)^p = \alpha^p + \beta^p$, válido porque $p \mid \binom{p}{k}$ para $0 < k < p$—, luego $\phi^n(\alpha) = \alpha^q$ también lo es, y F es exactamente el conjunto de puntos fijos de un homomorfismo: cerrado por suma, producto e inversos. Y tiene exactamente q elementos porque $f'(x) = qx^{q-1} - 1 = -1$ (ya que $q = 0$ en característica p), que no tiene raíces comunes con f: las q raíces son distintas.

(2) Sean F y F' dos cuerpos con q elementos. El grupo multiplicativo $F^\times$ tiene $q-1$ elementos, luego $\alpha^{q-1} = 1$ para todo $\alpha \neq 0$ por el teorema de Lagrange, y por tanto $\alpha^q = \alpha$ para *todo* $\alpha \in F$: los elementos de F son exactamente las q raíces de $x^q - x$, y F es un cuerpo de descomposición de ese polinomio sobre $\mathbb{F}_p$. Lo mismo vale para F'. Y dos cuerpos de descomposición del mismo polinomio sobre el mismo cuerpo base son isomorfos —resultado ya demostrado en el capítulo 26—, luego $F \cong F'$. □

Corolario 24.72 (Subcuerpos y retículo)**.** *$\mathbb{F}_{p^m}$ es subcuerpo de $\mathbb{F}_{p^n}$ si y sólo si $m \mid n$. En particular el retículo de subcuerpos de $\mathbb{F}_{p^n}$ es isomorfo al de divisores de n.*

Demostración. Si $\mathbb{F}_{p^m} \subset \mathbb{F}_{p^n}$, el mayor es espacio vectorial sobre el menor de alguna dimensión k, luego $p^n = (p^m)^k$ y $m \mid n$. Recíprocamente, si $m \mid n$ entonces $p^m - 1 \mid p^n - 1$ (pues $a-1 \mid a^k - 1$ con $a = p^m$), y de ahí $x^{p^m} - x \mid x^{p^n} - x$: las raíces del primero están entre las del segundo, y forman un subcuerpo con p^m elementos. □

Ejemplo 24.73 (La unicidad, en acción)**.** Los polinomios $x^3 + x + 1$ y $x^3 + x^2 + 1$ son ambos irreducibles sobre $\mathbb{F}_2$, y dan dos construcciones aparentemente distintas de un cuerpo de 8 elementos. El teorema garantiza que son isomorfos, y el isomorfismo se encuentra buscando en el segundo cuerpo una raíz del primer polinomio: si β genera $\mathbb{F}_2[x]/(x^3+x^2+1)$, se comprueba que $\alpha = \beta^2 + \beta$ satisface $\alpha^3 + \alpha + 1 = 0$, y $x \mapsto \beta^2 + \beta$ es el isomorfismo buscado. Sin el teorema de unicidad no habría ninguna razón para esperar que existiera.

24.24. El teorema de Gauss en su versión fuerte

Se ha usado que $K[x, y]$ y $\mathbb{Z}[x]$ son dominios de factorización única, apoyándose en un enunciado que se dejó sin demostrar. El lema débil —el producto de primitivos es primitivo— ya está probado; falta el salto.

Lema 24.74 (Irreducibilidad: en $R[x]$ y en $K[x]$ es lo mismo)**.** *Sea R un DFU con cuerpo de fracciones K, y sea $f \in R[x]$ primitivo de grado ≥ 1. Entonces f es irreducible en $R[x]$ si y sólo si lo es en $K[x]$.*

Demostración. Si $f = gh$ en $R[x]$ con ambos de grado ≥ 1, la misma igualdad vale en $K[x]$.

Recíprocamente, supóngase $f = GH$ en $K[x]$ con $\deg G, \deg H \geq 1$. Multiplicando por un denominador común y sacando contenidos, existen $c, d \in K^\times$ tales que $g = cG$ y $h = dH$ están en $R[x]$ y son *primitivos*. Entonces

$$cd\, f = g\, h.$$

Escríbase $cd = u/v$ con $u, v \in R$ sin factores comunes. De $u\, f = v\, gh$ se sigue, tomando contenidos y usando el lema de Gauss ($c(gh) = c(g)c(h) = 1$ y $c(f) = 1$), que u y v son asociados; luego cd es una unidad de R y f es, salvo unidad, el producto gh en $R[x]$, con ambos factores de grado ≥ 1: f no es irreducible en $R[x]$. □

Teorema 24.75 (Gauss, versión fuerte)**.** *Si R es un dominio de factorización única, también lo es $R[x]$. En consecuencia, $R[x_1, \ldots, x_n]$ es DFU para todo n; en particular lo son $\mathbb{Z}[x]$, $K[x, y]$ y $K[x_1, \ldots, x_n]$.*

Demostración. Sea K el cuerpo de fracciones de R. Recuérdese que $K[x]$ *sí* es DFU, por ser euclídeo y por el teorema 25.149.

Existencia de la factorización. Sea $f \in R[x]$ no nulo y no unidad. Escríbase $f = c(f)\, f_1$ con f_1 primitivo. El contenido $c(f) \in R$ se factoriza en irreducibles de R porque R es DFU, y los irreducibles de R siguen siéndolo en $R[x]$ (un factor de grado positivo haría subir el grado). Para f_1: factorícese en $K[x]$, que es DFU, como producto de irreducibles $F_1 \cdots F_s$; por el argumento del lema 24.74, cada F_j se puede sustituir por un múltiplo escalar primitivo $g_j \in R[x]$, y entonces f_1 y $g_1 \cdots g_s$ son ambos primitivos y proporcionales sobre K, luego iguales salvo unidad de R. Cada g_j es irreducible en $R[x]$ por el lema. Así f queda factorizado.

Unicidad. Supóngase dos factorizaciones de f en irreducibles de $R[x]$. Sepárense en cada una los factores de grado 0 —que son irreducibles de R— y los de grado ≥ 1 —que son necesariamente primitivos, pues un factor irreducible de grado positivo con contenido no unidad se partiría—. Por el lema de Gauss, el producto de los factores primitivos es primitivo en ambas descomposiciones, de modo que los productos de los factores constantes coinciden salvo unidad; la unicidad en R empareja esos factores. Y los factores de grado positivo son, por el lema 24.74, irreducibles en $K[x]$; la unicidad en $K[x]$ los empareja salvo constantes de K, y siendo todos primitivos, esas constantes son unidades de R. □

Observación 24.76 (Un DFU que no es DIP, y por qué importa). $\mathbb{Z}[x]$ es ahora, con todo rigor, un DFU; y *no* es un DIP, porque el ideal $(2, x)$ no es principal. La misma situación se da en $K[x, y]$ con el ideal (x, y). Esto explica por qué la jerarquía del capítulo tiene un escalón entre DIP y DFU que a primera vista parecía cosmético: en cuanto se pasa de una variable a dos, o de un cuerpo a $\mathbb{Z}$, se sale de los DIP y se cae exactamente en los DFU. Toda la geometría algebraica del capítulo 35 vive en $K[x_1, \ldots, x_n]$, es decir, en el primer sitio donde el teorema de Gauss fuerte es imprescindible.

24.25. Divisibilidad en un dominio: asociados, irreducibles, primos y máximo común divisor

La sección de la jerarquía acaba de definir el dominio de factorización única como aquel en el que "todo elemento se factoriza en irreducibles de forma única *salvo orden y unidades*".

Esa expresión esconde una palabra que el capítulo lleva usando desde el ejemplo de $\mathbb{Z}[\sqrt{-5}\,]$ y que nunca ha definido: *asociado*. Y hay otra pareja de palabras en la misma situación: el capítulo dice que en $\mathbb{Z}[\sqrt{-5}\,]$ el 2 es "irreducible pero no primo", pero la única definición de irreducible que ha dado es para polinomios sobre un cuerpo, y de *primo* en un anillo general no ha dado ninguna. Esta sección pone el vocabulario completo de la divisibilidad, con las demostraciones de las relaciones que hay entre esos conceptos.

Definición 24.77 (Divisibilidad y elementos asociados)**.** Sea R un dominio de integridad y $a, b \in R$. Se dice que a *divide* a b, y se escribe $a \mid b$, si existe $c \in R$ con $b = ac$. Dos elementos a y b son *asociados* si existe una unidad $u \in R^\times$ con $b = ua$.

Proposición 24.78 (Tres maneras de decir lo mismo)**.** *En un dominio de integridad R, y para a, b no nulos, son equivalentes:*

(i) *a y b son asociados;*

(ii) *$a \mid b$ y $b \mid a$;*

(iii) *$(a) = (b)$ como ideales.*

Demostración. $(i) \Rightarrow (ii)$: si $b = ua$ con u unidad, entonces $a \mid b$, y de $a = u^{-1}b$ sale $b \mid a$.

$(ii) \Rightarrow (i)$: escríbanse $b = ac$ y $a = bd$. Sustituyendo, $a = acd$, es decir $a(1 - cd) = 0$. Como $a \neq 0$ y R es dominio, $1 - cd = 0$, luego $cd = 1$ y c es una unidad; por tanto $b = ac$ con c unidad. *Aquí es donde se usa la hipótesis de dominio*, y sin ella la equivalencia es falsa.

$(ii) \Leftrightarrow (iii)$: $a \mid b$ equivale a $b \in (a)$, es decir a $(b) \subseteq (a)$; y $b \mid a$ equivale a $(a) \subseteq (b)$. Las dos inclusiones juntas son la igualdad. □

Definición 24.79 (Irreducible y primo, en un anillo cualquiera)**.** Sea R un dominio y $p \in R$ no nulo y no unidad.

- p es *irreducible* si de $p = ab$ se sigue que a o b es una unidad. Dicho de otro modo: sus únicos divisores son las unidades y sus propios asociados.
- p es *primo* si de $p \mid ab$ se sigue que $p \mid a$ o $p \mid b$.

Observación 24.80. Conviene ver que estas dos definiciones son la traducción exacta de las dos caracterizaciones escolares del número primo. "No se puede descomponer" es irreducible; "si divide a un producto, divide a un factor" es primo. En $\mathbb{Z}$ ambas describen el mismo conjunto de números y por eso en la escuela se usan indistintamente. Que sean cosas distintas —y que la equivalencia sea un teorema y no una tautología— es precisamente lo que el ejemplo de $\mathbb{Z}[\sqrt{-5}\,]$ enseñaba.

Teorema 24.81 (Primo implica irreducible)**.** *En todo dominio de integridad, todo elemento primo es irreducible.*

Demostración. Sea p primo y supóngase $p = ab$. En particular $p \mid ab$, luego por primalidad $p \mid a$ o $p \mid b$; sin pérdida de generalidad $p \mid a$, digamos $a = pc$. Sustituyendo, $p = ab = pcb$,

es decir $p(1-cb)=0$, y como $p\neq 0$ en un dominio queda $cb=1$: b es una unidad. Por la definición 24.79, p es irreducible. □

Cuidado

El recíproco es *falso*, y el capítulo ya exhibió el contraejemplo: en $\mathbb{Z}[\sqrt{-5}\,]$ el elemento 2 es irreducible —porque la ecuación de normas $a^2+5b^2=2$ no tiene soluciones enteras— pero no es primo, ya que divide a $6=(1+\sqrt{-5})(1-\sqrt{-5})$ sin dividir a ninguno de los dos factores. Lo que ahora se puede decir con precisión es que ese anillo falla la *implicación que sí se cumple en los DIP y en los DFU*, y que por eso deja de ser DFU.

Definición 24.82 (Máximo común divisor en un dominio)**.** Sean $a,b\in R$ no ambos nulos. Un elemento d es un *máximo común divisor* de a y b si $d\mid a$, $d\mid b$, y además todo c que divida a los dos divide también a d. Si existe, es único salvo asociados, por la proposición 24.78.

Ejemplo 24.83 (Dos elementos sin máximo común divisor)**.** Que la definición hable de "si existe" no es una cautela retórica. En $\mathbb{Z}[\sqrt{-5}\,]$ considérense $\alpha=6$ y $\beta=2+2\sqrt{-5}$. Un recorrido exhaustivo de los elementos de norma pequeña —comprobado por ordenador— muestra que sus divisores comunes son exactamente

$$\pm 1,\qquad \pm 2,\qquad \pm(1+\sqrt{-5}),$$

de normas 1, 4 y 6 respectivamente. En efecto, $2\mid 6$ y $2\mid 2+2\sqrt{-5}$, y $(1+\sqrt{-5})(1-\sqrt{-5})=6$ mientras que $2+2\sqrt{-5}=2(1+\sqrt{-5})$.

Ahora bien, *ninguno de los dos candidatos serios divide al otro*: si 2 dividiera a $1+\sqrt{-5}$ las normas darían $4\mid 6$, falso; y si $1+\sqrt{-5}$ dividiera a 2 darían $6\mid 4$, también falso. Como tampoco hay ningún divisor común de norma mayor, no existe ningún d divisible por todos los divisores comunes: **el máximo común divisor de 6 y $2+2\sqrt{-5}$ no existe**. En $\mathbb{Z}$ o en $K[x]$ ese fenómeno es impensable, y su ausencia es justamente lo que garantiza el algoritmo de Euclides.

Definición 24.84 (Nilpotentes, idempotentes y nilradical)**.** Sea R un anillo conmutativo con unidad. Un elemento a es *nilpotente* si $a^k=0$ para algún $k\geq 1$, e *idempotente* si $a^2=a$. El conjunto de los nilpotentes se llama *nilradical* y se denota $\operatorname{nil}(R)$; el anillo se dice *reducido* si $\operatorname{nil}(R)=\{0\}$.

Proposición 24.85 (El nilradical es un ideal)**.** *Si R es conmutativo, $\operatorname{nil}(R)$ es un ideal, y todo dominio de integridad es reducido.*

Demostración. Si $a^m=0$ y $r\in R$, entonces $(ra)^m=r^m a^m=0$ por conmutatividad: hay absorción. Si además $b^n=0$, el binomio de Newton da

$$(a-b)^{m+n-1}=\sum_{k=0}^{m+n-1}\binom{m+n-1}{k}a^k(-b)^{m+n-1-k},$$

y en cada sumando o bien $k\geq m$, con lo que $a^k=0$, o bien $k<m$ y entonces $m+n-1-k>n-1$, es decir $m+n-1-k\geq n$ y $b^{m+n-1-k}=0$. Todos los términos se anulan, luego $a-b$

es nilpotente. En un dominio, $a^k = 0$ con $a \neq 0$ es imposible, porque un producto de factores no nulos no puede ser nulo. □

Ejemplo 24.86 (Nilpotentes e idempotentes de varios $\mathbb{Z}_n$, calculados)**.** Los cálculos siguientes están comprobados por ordenador recorriendo todos los elementos.

anillo	nilpotentes	idempotentes
$\mathbb{Z}_{12}$	$\{0, 6\}$	$\{0, 1, 4, 9\}$
$\mathbb{Z}_8$	$\{0, 2, 4, 6\}$	$\{0, 1\}$
$\mathbb{Z}_{20}$	$\{0, 10\}$	$\{0, 1, 5, 16\}$
$\mathbb{Z}_{30}$	$\{0\}$	$\{0, 1, 6, 10, 15, 16, 21, 25\}$
$\mathbb{Z}_{36}$	$\{0, 6, 12, 18, 24, 30\}$	$\{0, 1, 9, 28\}$

Las dos columnas obedecen a reglas distintas y ambas se leen en la factorización de n. **Los nilpotentes** de $\mathbb{Z}_n$ son los múltiplos del *radical* de n, es decir del producto de sus primos distintos: para $n = 12$ el radical es $2 \cdot 3 = 6$ y los múltiplos de 6 son $\{0, 6\}$; para $n = 36$ el radical vuelve a ser 6 y hay seis múltiplos; para $n = 30$, que es libre de cuadrados, el radical es 30 y solo queda el 0, de modo que $\mathbb{Z}_{30}$ es reducido. **Los idempotentes** son $2^{\omega(n)}$, con $\omega(n)$ el número de primos distintos: cuatro para 12, 20 y 36; dos para 8; ocho para 30. La razón es el teorema chino del resto, que descompone $\mathbb{Z}_n$ en un producto de anillos locales y hace que un idempotente sea simplemente una elección de 0 o 1 en cada factor. Compruébese en $\mathbb{Z}_{12} \cong \mathbb{Z}_4 \times \mathbb{Z}_3$: el idempotente 9 corresponde a $(1, 0)$ y el 4 a $(0, 1)$, con $4 + 9 = 13 = 1$ y $4 \cdot 9 = 36 = 0$.

24.26. Dominios euclídeos: la definición formal y la cadena euclídeo ⇒ DIP ⇒ DFU

La escalera de la sección anterior describe el peldaño euclídeo así: "existe una función tamaño que permite dividir con resto". Eso es una paráfrasis, no una definición, y las tres implicaciones de la cadena se afirman sin prueba, con una remisión a otro capítulo que este promete no necesitar. Se completa aquí.

Definición 24.87 (Función euclídea y dominio euclídeo)**.** Un dominio de integridad R es *euclídeo* si existe una función $N : R \setminus \{0\} \to \mathbb{N}$, llamada *función euclídea* o *norma*, tal que para cualesquiera $a, b \in R$ con $b \neq 0$ existen $q, r \in R$ con

$$a = qb + r, \qquad \text{y o bien } r = 0 \text{ o bien } N(r) < N(b).$$

No se exige que q y r sean únicos.

Ejemplo 24.88 (Los tres dominios euclídeos del capítulo, verificados)**. (a) $\mathbb{Z}$ con $N(a) = |a|$.** Es la división entera con resto: dados a y $b \neq 0$ existen q y r con $a = qb + r$ y $0 \leq r < |b|$, de modo que $N(r) = r < |b| = N(b)$.

(b) $K[x]$ con $N(f) = \deg f$, para K cuerpo. Es el teorema 24.38, ya demostrado: existen q, r únicos con $f = qg + r$ y $\deg r < \deg g$. Aquí el cociente y el resto son además únicos, cosa que la definición no exige y que $\mathbb{Z}$ tampoco cumple si se admiten restos negativos.

(c) $\mathbb{Z}[i]$ con $N(a+bi)=a^2+b^2$. Esta merece la comprobación completa, porque es la única que no se ha hecho en el capítulo. Dados α y $\beta\neq 0$, calcúlese el cociente en $\mathbb{Q}(i)$:

$$\frac{\alpha}{\beta}=u+vi, \qquad u,v\in\mathbb{Q},$$

y tómese $q=m+ni$ con m y n los enteros *más próximos* a u y a v, de modo que $|u-m|\leq\frac{1}{2}$ y $|v-n|\leq\frac{1}{2}$. Con $r=\alpha-\beta q$ resulta

$$N(r)=N(\beta)\,N\!\left(\frac{\alpha}{\beta}-q\right)=N(\beta)\big[(u-m)^2+(v-n)^2\big]\leq N(\beta)\left(\frac{1}{4}+\frac{1}{4}\right)=\frac{N(\beta)}{2}<N(\beta),$$

usando que la norma es multiplicativa. Obsérvese que la desigualdad obtenida es *mucho* mejor que la exigida: cada paso del algoritmo al menos divide la norma por dos, lo que garantiza que Euclides en $\mathbb{Z}[i]$ termina muy deprisa.

Regla: dividir y hallar el máximo común divisor en $\mathbb{Z}[i]$

1. **Multiplica por el conjugado para dividir.** Para calcular α/β, escribe $\alpha\bar{\beta}/N(\beta)$; el numerador es un entero de Gauss y el denominador un entero positivo, así que las dos coordenadas racionales se leen de inmediato.
2. **Redondea cada coordenada al entero más próximo** para obtener q. Si alguna cae en un punto medio, cualquiera de los dos redondeos vale: el cociente no es único y no importa.
3. **Calcula el resto $r=\alpha-\beta q$ y verifica $N(r)<N(\beta)$.** Si no se cumple, te has equivocado al redondear; de hecho debe salir $N(r)\leq N(\beta)/2$.
4. **Itera con el par** (β, r) hasta que el resto sea cero.
5. **El último resto no nulo es el máximo común divisor**, definido salvo las cuatro unidades $\pm1,\pm i$; puedes elegir el representante que prefieras.
6. **Comprueba con normas.** La norma del máximo común divisor debe dividir a las normas de los dos elementos de partida. Es la verificación más rápida que existe.

Ejemplo 24.89 (Euclides con enteros de Gauss, con todos los pasos)**.** Calculemos $\mathrm{mcd}(17+7i,\ 5+3i)$. Las normas son $N(17+7i)=289+49=338$ y $N(5+3i)=25+9=34$.

Paso 1. $\dfrac{17+7i}{5+3i}=\dfrac{(17+7i)(5-3i)}{34}=\dfrac{106-16i}{34}=3{,}1176-0{,}4706\,i$. Redondeando, $q=3$, y

$$17+7i-(5+3i)\cdot 3=17+7i-15-9i=2-2i, \qquad N(2-2i)=8<34.\checkmark$$

Paso 2. $\dfrac{5+3i}{2-2i}=\dfrac{(5+3i)(2+2i)}{8}=\dfrac{4+16i}{8}=0{,}5+2i$. Redondeando (el 0,5 puede ir a 0 o a 1; tomamos 0), $q=2i$, y

$$5+3i-(2-2i)\cdot 2i=5+3i-(4i+4)=1-i, \qquad N(1-i)=2<8.\checkmark$$

Paso 3. $\dfrac{2-2i}{1-i} = 2$ exactamente, luego el resto es 0 y el algoritmo termina.

Conclusión. $\mathrm{mcd}(17+7i,\ 5+3i) = 1-i$, o cualquiera de sus asociados $-1+i$, $1+i$, $-1-i$. Comprobación por normas: $N(1-i)=2$, que divide a 338 y a 34 ✓. Y comprobación por construcción: $17+7i = (1-i)(5+12i)$ y $5+3i = (1-i)(1+4i)$, con $N(5+12i)=169$ y $N(1+4i)=17$, dos números sin factores comunes; los cofactores no comparten nada y $1-i$ es efectivamente el máximo común divisor. ✓

Teorema 24.90 (Todo dominio euclídeo es de ideales principales)**.** *Si R es euclídeo con función N, entonces todo ideal de R es principal.*

Demostración. Sea I un ideal. Si $I=\{0\}$ entonces $I=(0)$ y no hay nada que probar. Si no, el conjunto $\{N(b) : b \in I,\ b \neq 0\}$ es un subconjunto no vacío de $\mathbb{N}$ y por buena ordenación tiene un mínimo; sea $b \in I$ un elemento no nulo que lo alcanza.

Claramente $(b) \subseteq I$ por absorción. Recíprocamente, sea $a \in I$ y divídase: $a = qb + r$ con $r = 0$ o $N(r) < N(b)$. Ahora bien, $r = a - qb$ está en I, porque $a \in I$ y $qb \in I$. Si fuese $r \neq 0$ tendría $N(r) < N(b)$, contra la minimalidad de $N(b)$. Luego $r = 0$ y $a = qb \in (b)$. □

Observación 24.91. Esta es, palabra por palabra, la demostración que el capítulo dio para $\mathbb{Z}$: "sea n el menor entero positivo de I". Lo que se ha hecho es sustituir "menor entero positivo" por "menor valor de N", y así la misma prueba cubre de golpe $\mathbb{Z}$, $K[x]$ y $\mathbb{Z}[i]$. Ese es el sentido de abstraer: no se gana generalidad por gusto, se ahorra escribir tres veces lo mismo.

Teorema 24.92 (En un DIP, irreducible y primo coinciden)**.** *Sea R un dominio de ideales principales y $p \in R$ no nulo ni unidad. Entonces p es irreducible si y solo si es primo.*

Demostración. Una implicación es el teorema 24.81 y vale en cualquier dominio. Para la otra, sea p irreducible y supóngase $p \mid ab$ con $p \nmid a$; hay que probar $p \mid b$.

Considérese el ideal $(p,a) = \{xp + ya : x, y \in R\}$. Como R es DIP, $(p,a) = (d)$ para algún d. En particular $d \mid p$, y al ser p irreducible, d es una unidad o un asociado de p. Si d fuese asociado de p, de $d \mid a$ se seguiría $p \mid a$, contra lo supuesto. Luego d es una unidad y $(p,a) = R$, es decir, existen x, y con

$$xp + ya = 1.$$

Esa es la identidad de Bézout, aquí obtenida sin ningún algoritmo. Multiplicando por b,

$$b = xpb + yab.$$

El primer sumando es múltiplo de p y el segundo también, porque $p \mid ab$. Luego $p \mid b$. □

Teorema 24.93 (Todo DIP es de factorización única)**.** *Si R es un dominio de ideales principales, todo elemento no nulo y no unidad se escribe como producto finito de irreducibles, y la escritura es única salvo el orden de los factores y sustitución de cada uno por un asociado.*

Demostración. Paso 1: las cadenas de ideales principales se estabilizan. Sea $(a_1) \subseteq (a_2) \subseteq \cdots$ una cadena ascendente. Su unión $I = \bigcup_n (a_n)$ es un ideal —la unión de una cadena de ideales lo es, porque dos elementos cualesquiera están en un mismo eslabón—, luego $I = (a)$ por

hipótesis, y ese a está en algún (a_N). Entonces $I = (a) \subseteq (a_N) \subseteq I$, y la cadena es constante desde N.

Paso 2: existencia de la factorización. Supóngase que algún a no nulo ni unidad no admite factorización en irreducibles. Entonces a no es irreducible, así que $a = bc$ con b y c no unidades; y al menos uno de los dos, digamos b, tampoco admite factorización (si los dos la admitieran, concatenándolas se obtendría una de a). Además $(a) \subsetneq (b)$: la inclusión es clara, y es estricta porque si fueran iguales b y a serían asociados por la proposición 24.78 y c sería una unidad. Repitiendo el argumento con b se obtiene una cadena estrictamente creciente e infinita $(a) \subsetneq (b) \subsetneq (b_2) \subsetneq \cdots$, imposible por el paso 1.

Paso 3: unicidad. Sean $p_1 \cdots p_s = q_1 \cdots q_t$ dos factorizaciones en irreducibles, y procédase por inducción sobre s. Si $s = 0$ el producto es una unidad y $t = 0$. Si $s \geq 1$, el irreducible p_1 es primo por el teorema 24.92, y como divide al producto $q_1 \cdots q_t$, divide a algún q_j; reordenando, $p_1 \mid q_1$. Pero q_1 es irreducible y p_1 no es unidad, así que $q_1 = up_1$ con u unidad: son asociados. Cancelando p_1 —lícito en un dominio— queda $p_2 \cdots p_s = (uq_2)q_3 \cdots q_t$, y la hipótesis de inducción empareja el resto. □

Observación 24.94 (La cadena, ahora completa y sin remisiones). Con estos tres teoremas la escalera del capítulo queda demostrada de arriba abajo dentro del propio capítulo:

$$\text{cuerpo} \;\Rightarrow\; \text{euclídeo} \;\Rightarrow\; \text{DIP} \;\Rightarrow\; \text{DFU} \;\Rightarrow\; \text{dominio},$$

donde la primera implicación es trivial (en un cuerpo se puede tomar $N \equiv 0$ y el resto siempre nulo) y las tres siguientes son los teoremas 24.90, 24.93 y la definición misma de DFU. Y las cuatro implicaciones son *estrictas*, con los testigos que el capítulo ya listó: $\mathbb{Z}$ es euclídeo y no cuerpo; $\mathbb{Z}\left[\frac{1+\sqrt{-19}}{2}\right]$ es DIP y no euclídeo —para *ninguna* función N, lo cual es un teorema fino y no una simple ausencia de candidatos evidentes—; $\mathbb{Z}[x]$ es DFU y no DIP, por el ideal $(2, x)$; y $\mathbb{Z}[\sqrt{-5}\,]$ es dominio y no DFU, por las dos factorizaciones del 6.

Idea clave

Merece la pena quedarse con la estructura lógica, porque se repite en todo el álgebra. La propiedad *operativa* es ser euclídeo: es la única de las cuatro que dice cómo *hacer* algo, a saber, dividir. De ella sale una propiedad *estructural*, que los ideales sean principales, y de esta una propiedad *aritmética*, la factorización única. Las tres flechas van en un solo sentido y ninguna se puede invertir. Cuando en un anillo nuevo uno quiere saber cuánto álgebra escolar sobrevive, la pregunta útil no es "¿se parece a $\mathbb{Z}$?", sino "¿en qué peldaño de esta escalera está?"; y la manera más rápida de averiguarlo es intentar construir una función euclídea y, si falla, buscar dos factorizaciones distintas de algún número pequeño.

24.27. Ejercicios resueltos

Ejercicio 24.10. Construye $\mathbb{F}_8$ como $\mathbb{F}_2[x]/(x^3 + x + 1)$, tabula las potencias de $\alpha = \bar{x}$ y úsalas para calcular $(\alpha^2 + 1)^{-1}$.

Solución. Primero hay que ver que $x^3 + x + 1$ es irreducible sobre $\mathbb{F}_2$: al ser de grado 3, bastaría con que tuviera una raíz, y $0^3 + 0 + 1 = 1$, $1^3 + 1 + 1 = 1$; ninguna de las dos raíces posibles sirve. Luego el cociente es un cuerpo con $2^3 = 8$ elementos.

En él $\alpha^3 = \alpha + 1$ (el polinomio dice justo eso, recordando que $-1 = 1$), y de ahí sale todo lo demás multiplicando por α y reduciendo:

k	1	2	3	4	5	6	7
α^k	α	α^2	$\alpha+1$	$\alpha^2+\alpha$	$\alpha^2+\alpha+1$	α^2+1	1

Las siete potencias recorren los siete elementos no nulos: α es un generador del grupo multiplicativo $\mathbb{F}_8^\times \cong \mathbb{Z}/7$. Como 7 es primo, de hecho *cualquier* elemento distinto de 0 y 1 es generador.

Para el inverso ya no hace falta Euclides: $\alpha^2 + 1 = \alpha^6$, luego

$$(\alpha^2+1)^{-1} = \alpha^{-6} = \alpha^{7-6} = \alpha.$$

Comprobación: $\alpha(\alpha^2+1) = \alpha^3 + \alpha = (\alpha+1) + \alpha = 1$. ✓

La moraleja práctica es que una tabla de logaritmos discretos convierte cada división en una resta de exponentes; así se implementan de verdad los cuerpos finitos en los códigos correctores y en AES. ■

Ejercicio 24.11. Calcula $\gcd(4+7i,\ -2+9i)$ en $\mathbb{Z}[i]$ mediante el algoritmo de Euclides.

Solución. La norma $N(a+bi) = a^2+b^2$ es multiplicativa y hace de $\mathbb{Z}[i]$ un dominio euclídeo: dado $\beta \neq 0$ siempre hay q con $N(\alpha - \beta q) < N(\beta)$, y se encuentra dividiendo en $\mathbb{Q}(i)$ y redondeando cada coordenada al entero más próximo. Aquí $N(4+7i) = 65$ y $N(-2+9i) = 85$; empezamos por el de norma mayor.

$$\begin{aligned}
4+7i &= (-2+9i)(1-i) + (-3-4i), && N(-3-4i) = 25 < 85,\\
-2+9i &= (-3-4i)(-1-i) + (-1+2i), && N(-1+2i) = 5 < 25,\\
-3-4i &= (-1+2i)(-1+2i) + 0.
\end{aligned}$$

El último resto no nulo es $-1+2i$, luego $\gcd = -1+2i$, o cualquiera de sus asociados $1-2i$, $2+i$, $-2-i$ (el máximo común divisor solo está definido salvo unidades, y las unidades de $\mathbb{Z}[i]$ son $\pm 1, \pm i$).

Comprobación por construcción: $4+7i = (2+i)(3+2i)$ y $-2+9i = (2+i)(1+4i)$, con $N(3+2i) = 13$ y $N(1+4i) = 17$ primos distintos, así que los cofactores no comparten nada y el máximo común divisor es exactamente $2+i$. ✓ ■

Ejercicio 24.12. Demuestra que $(2, x)$ no es un ideal principal de $\mathbb{Z}[x]$, y concluye que $\mathbb{Z}[x]$ es DFU pero no DIP.

Solución. Supongamos $(2,x) = (f)$. Entonces $f \mid 2$, y como 2 tiene grado 0, también f tiene grado 0: f es un entero, y al dividir a 2 solo puede ser ± 1 o ± 2. Pero $f \mid x$ obliga a $f = \pm 1$ (un 2 no divide a x en $\mathbb{Z}[x]$, porque $x/2$ no tiene coeficientes enteros). Luego $(f) = \mathbb{Z}[x]$, el ideal total.

Y eso es falso: $(2,x)$ es propio, porque todo elemento suyo tiene término independiente par y el 1 no lo tiene. Contradicción.

Que $\mathbb{Z}[x]$ sí sea DFU es el teorema de Gauss: si R es DFU, $R[x]$ también. Así que la cadena de inclusiones

$$\text{cuerpos} \subsetneq \text{DIP} \subsetneq \text{DFU} \subsetneq \text{dominios}$$

es estricta en todos los eslabones, y $\mathbb{Z}[x]$ es el testigo del segundo. Del tercero es testigo $\mathbb{Z}[\sqrt{-5}\,]$, y del primero, $\mathbb{Z}$ mismo.

De paso, $(2,x)$ es maximal: $\mathbb{Z}[x]/(2,x) \cong \mathbb{F}_2$, que es un cuerpo. Es un ejemplo de ideal maximal que necesita dos generadores, algo imposible en un DIP. ■

Ejercicio 24.13. Describe por completo $\mathbb{Z}/12\mathbb{Z}$: unidades, divisores de cero, nilpotentes e idempotentes. Interpreta el resultado con el teorema chino del resto.

Solución. El teorema chino da $\mathbb{Z}/12 \cong \mathbb{Z}/4 \times \mathbb{Z}/3$, y casi todo se lee ahí.

Unidades: $\{1,5,7,11\}$, los cuatro coprimos con 12; en efecto $\varphi(12)=\varphi(4)\varphi(3)=2\cdot 2=4$. Como grupo es $\mathbb{Z}/2\times\mathbb{Z}/2$ (todos cumplen $a^2=1$: $5^2=25=1$, $7^2=49=1$, $11^2=121=1$).

Divisores de cero: todos los demás salvo el 0, es decir $\{2,3,4,6,8,9,10\}$; por ejemplo $3\cdot 4=0$.

Nilpotentes: solo 0 y 6, porque $6^2=36=0$. En general el nilradical de $\mathbb{Z}/n$ está generado por el radical de n, y $\mathrm{rad}(12)=2\cdot 3=6$. Como hay nilpotentes no nulos, $\mathbb{Z}/12$ no es reducido.

Idempotentes: $\{0,1,4,9\}$, cuatro en total. Y esto es exactamente lo que predice el teorema chino: en $\mathbb{Z}/4\times\mathbb{Z}/3$ los idempotentes son los pares de idempotentes, $(0,0),(1,0),(0,1),(1,1)$, es decir $2^2=4$. La correspondencia es $9\leftrightarrow(1,0)$ y $4\leftrightarrow(0,1)$, y en efecto $4+9=13=1$ y $4\cdot 9=36=0$: son idempotentes ortogonales complementarios, la descomposición del 1 que *es* la descomposición del anillo. ■

Ejercicio 24.14. Demuestra que x^4+1 es irreducible sobre $\mathbb{Q}$ y sin embargo factoriza módulo todo primo.

Solución. **Irreducible sobre $\mathbb{Q}$.** Eisenstein no se aplica directamente (ningún primo divide a los coeficientes intermedios, que son cero, y a la vez cumple lo demás), pero sí tras trasladar. Con $x\mapsto x+1$:

$$(x+1)^4+1=x^4+4x^3+6x^2+4x+2,$$

y aquí Eisenstein con $p=2$ funciona: 2 divide a $4,6,4,2$, no divide al coeficiente principal 1, y $4=2^2$ no divide al término independiente 2. Luego el trasladado es irreducible, y una traslación es un automorfismo de $\mathbb{Q}[x]$: el original también lo es. (Es el octavo polinomio ciclotómico, Φ_8.)

Factoriza módulo todo primo. Módulo 2, $x^4+1=(x+1)^4$. Para p impar, $p^2\equiv 1 \pmod 8$ siempre, luego $\mathbb{F}_{p^2}^{\times}$ tiene orden p^2-1 divisible por 8 y contiene las raíces octavas de la unidad: las raíces de x^4+1 viven en $\mathbb{F}_{p^2}$, así que ningún factor irreducible puede tener grado mayor que 2. Muestras:

$$p=3:(x^2+x+2)(x^2+2x+2),\qquad p=5:(x^2+2)(x^2+3),$$
$$p=7:(x^2+3x+1)(x^2+4x+1),\qquad p=13:(x^2+5)(x^2+8).$$

La lección es que la reducción módulo p es una prueba de irreducibilidad *de un solo sentido*: si el reducido es irreducible, el original también; pero que el reducido factorice para todo p no dice nada. x^4+1 es el contraejemplo canónico, y la razón profunda es que su grupo de Galois es $\mathbb{Z}/2\times\mathbb{Z}/2$, que no tiene elementos de orden 4: ningún Frobenius puede actuar transitivamente sobre las cuatro raíces. ■

Ejercicio 24.15. En $\mathbb{Z}[\sqrt{-5}\,]$ demuestra que 2 es irreducible pero no primo.

Solución. Usamos la norma $N(a+b\sqrt{-5})=a^2+5b^2$, que es multiplicativa y toma valores enteros no negativos. Sus unidades son los elementos de norma 1: solo ± 1.

Irreducible. Si $2=\alpha\beta$ entonces $4=N(\alpha)N(\beta)$. Si ninguno fuera unidad, ambas normas valdrían 2; pero $a^2+5b^2=2$ no tiene soluciones enteras ($b=0$ pide $a^2=2$, y $|b|\geq 1$ ya pasa de 2). Luego uno de los dos es unidad: 2 es irreducible.

No primo. De $6=2\cdot 3=(1+\sqrt{-5})(1-\sqrt{-5})$ se sigue que 2 divide al producto $(1+\sqrt{-5})(1-\sqrt{-5})$. ¿Divide a alguno de los factores? Si $1+\sqrt{-5}=2\gamma$, tomando normas $6=4N(\gamma)$, imposible con $N(\gamma)$ entero. Lo mismo con el otro. Así que 2 divide al producto sin dividir a ningún factor: no es primo.

Ahí está la grieta exacta por la que $\mathbb{Z}[\sqrt{-5}\,]$ deja de ser DFU. En un DFU irreducible y primo son lo mismo; aquí no, y por eso 6 admite dos factorizaciones esencialmente distintas. La reparación histórica de Kummer y Dedekind fue subir de nivel: los *ideales* sí factorizan de forma única, y en efecto $(2)=(2,1+\sqrt{-5})^2$, con lo que la unicidad se restaura un piso más arriba. ■

24.28. Ejercicios propuestos

Sin solución desarrollada: solo enunciado y, al final, la respuesta. Aquí muchas respuestas son un nombre ("el grupo de Klein", "sí, por Eisenstein") más que un número: comprueba que coincida el fondo del argumento, no la redacción.

24.1. ¿Es $\mathbb{Z}/6\mathbb{Z}$ un dominio de integridad? Da un par de divisores de cero.

24.2. ¿Para qué n es $\mathbb{Z}/n\mathbb{Z}$ un cuerpo?

24.3. ¿Es (x) un ideal primo de $\mathbb{Z}[x]$? ¿Es maximal?

24.4. Describe $\mathbb{Z}[x]/(x^2+1)$: ¿a qué anillo conocido es isomorfo?

24.5. ¿Es x^3-2 irreducible sobre $\mathbb{Q}$? Cita el criterio.

24.6. Factoriza x^4+1 sobre $\mathbb{F}_2$.

24.7. Factoriza x^2+1 sobre $\mathbb{F}_5$.

24.8. ¿Es $\mathbb{Z}[\sqrt{-5}]$ un DFU? Da la factorización doble de 6.

24.9. Ordena por inclusión: cuerpo $\subset$ DIP $\subset$ DFU $\subset$ dominio. ¿Dónde encaja $\mathbb{Z}[x]$?

24.10. ¿Cuántos elementos tiene $\mathbb{F}_2[x]/(x^3+x+1)$? ¿Es un cuerpo?

24.11. Enuncia el primer teorema de isomorfía para anillos y aplícalo a $\mathbb{Z} \to \mathbb{Z}/n\mathbb{Z}$.

24.12. ¿Es todo ideal maximal primo? ¿Y al revés? Da un contraejemplo en $\mathbb{Z}[x]$.

Respuestas

1. No: $2\cdot 3 = 0$ en $\mathbb{Z}/6\mathbb{Z}$. **2.** Exactamente cuando n es primo. **3.** Primo sí, porque $\mathbb{Z}[x]/(x) \cong \mathbb{Z}$ es dominio; maximal no, porque $\mathbb{Z}$ no es cuerpo. **4.** $\mathbb{Z}[x]/(x^2+1) \cong \mathbb{Z}[i]$, los enteros de Gauss. **5.** Sí, por Eisenstein con $p=2$ (o porque no tiene raíces racionales y es de grado 3). **6.** $x^4+1 = (x+1)^4$ en $\mathbb{F}_2[x]$. **7.** $x^2+1 = (x+2)(x-2) = (x+2)(x+3)$ en $\mathbb{F}_5[x]$. **8.** No es DFU: $6 = 2\cdot 3 = (1+\sqrt{-5})(1-\sqrt{-5})$, dos factorizaciones en irreducibles no asociados. **9.** Todo cuerpo es DIP, todo DIP es DFU, todo DFU es dominio, y ninguna implicación se invierte. $\mathbb{Z}[x]$ es DFU pero no DIP: el ideal $(2,x)$ no es principal. **10.** Ocho elementos, y sí es un cuerpo: $\mathbb{F}_8$, porque x^3+x+1 es irreducible sobre $\mathbb{F}_2$. **11.** Si $\varphi : R \to S$ es homomorfismo, $R/\ker\varphi \cong \operatorname{im}\varphi$. Con la reducción $\mathbb{Z} \to \mathbb{Z}/n\mathbb{Z}$ el núcleo es (n), luego $\mathbb{Z}/(n) \cong \mathbb{Z}/n\mathbb{Z}$. **12.** Maximal $\Rightarrow$ primo siempre (en anillos conmutativos con unidad); el recíproco falla: $(x) \subset \mathbb{Z}[x]$ es primo y no maximal.

Capítulo 25

Estructuras algebraicas: grupos, anillos y cuerpos

Lo que necesitas traer

De los capítulos 23 y 24: todo. Éste es el capítulo que recoge y sistematiza las dos teorías, y añade lo que hace falta para Galois: acciones de grupo, teoremas de Sylow, módulos y extensiones de cuerpos. Del capítulo 7: espacios vectoriales, base y dimensión, que reaparecen como caso particular de módulo.

Cambiamos de continente. El álgebra abstracta no estudia números concretos sino *estructuras*: conjuntos con operaciones que cumplen reglas. Este capítulo construye con detalle las tres estructuras que la teoría de Galois necesita —grupos, anillos y cuerpos— con especial atención a los grupos de permutaciones y a los polinomios, los dos protagonistas del capítulo final.

25.1. Grupos: definición y ejemplos

Idea clave

Intuición primero. Un grupo es la formalización de la palabra *simetría*. Toma un objeto —un triángulo, una ecuación, una molécula, un cubo de Rubik— y considera todas las maneras de manipularlo *sin cambiarlo esencialmente*. Esas manipulaciones siempre cumplen cuatro cosas obvias: hacer una y luego otra vuelve a ser una manipulación (cierre), da igual cómo agrupes tres seguidas (asociatividad), existe la de "no hacer nada" (neutro) y toda manipulación se puede *deshacer* (inverso). Eso es todo lo que dice la definición. Fíjate en lo que *no* se pide: no se pide conmutar. Ponerte los calcetines y luego los zapatos no es lo mismo que el orden inverso, y esa asimetría tan cotidiana es exactamente la que hará que S_5 sea complicado y que las ecuaciones de quinto grado no tengan fórmula. Un grupo, entonces, es un "sistema de acciones reversibles"; los números son solo uno de sus muchos disfraces.

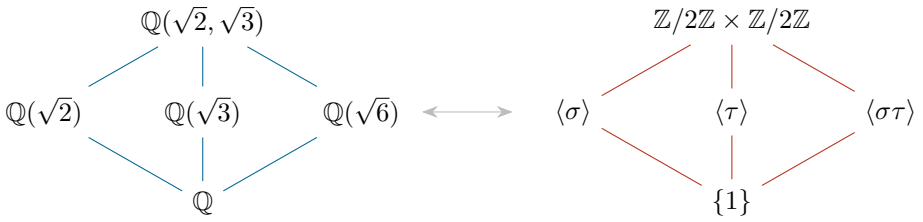


Figura 25.1: La correspondencia de Galois para $\mathbb{Q}(\sqrt{2},\sqrt{3})/\mathbb{Q}$. A la izquierda, los cinco cuerpos intermedios; a la derecha, los cinco subgrupos del grupo de Klein. La biyección invierte las inclusiones: cuanto mayor es el cuerpo, menor es el grupo que lo fija.

Definición 25.1. Un *grupo* $(G,*)$ es un conjunto con una operación asociativa, con elemento neutro e, y donde todo elemento tiene inverso. Si la operación conmuta, el grupo es *abeliano*. El *orden* $|G|$ es su cardinal.

De los axiomas se deducen: el neutro es único, el inverso es único, valen las cancelaciones ($ax=ay\Rightarrow x=y$), y $(ab)^{-1}=b^{-1}a^{-1}$.

Ejemplo 25.2 (El repertorio básico)**.** *Numéricos*: $(\mathbb{Z},+)$, $(\mathbb{Q}^*,\cdot)$, $(\mathbb{Z}_n,+)$ (enteros módulo n), $(\mathbb{Z}_p^*,\cdot)$ para p primo. *Geométricos*: el grupo *diédrico* D_n de las $2n$ simetrías del polígono regular de n lados (n rotaciones y n reflexiones); las rotaciones del espacio $SO(3)$. *De matrices*: $GL_n(K)$ (invertibles con el producto), SL_n (determinante 1), $O(n)$ (ortogonales). *De permutaciones*: el *grupo simétrico* S_n de las biyecciones de $\{1,\dots,n\}$ con la composición, $|S_n|=n!$, no abeliano para $n\geq 3$. Es el ejemplo decisivo para Galois.

25.1.1. Permutaciones en detalle

Idea clave

Intuición primero. Una permutación de $\{1,\dots,n\}$ es un *barajado* de n fichas. Si sigues una ficha y ves a dónde va, y luego a dónde va esa, tarde o temprano vuelves al inicio: has recorrido un *ciclo*, como un carrusel. Todo barajado se descompone en carruseles que giran a la vez y no comparten fichas: por eso los ciclos disjuntos conmutan (dos carruseles independientes no se estorban) y por eso el orden de la permutación es el mínimo común múltiplo de sus longitudes (el momento en que todos los carruseles coinciden simultáneamente en su posición de partida). El *signo* cuenta si el barajado se logra con un número par o impar de intercambios de dos fichas; que la paridad no dependa del camino elegido es un hecho sutil, y es el mismo hecho que hace que el determinante del capítulo 7 esté bien definido.

Toda permutación se factoriza de forma única (salvo orden) en *ciclos disjuntos*; los ciclos disjuntos conmutan. Un ciclo de longitud k tiene orden k, y el orden de una permutación es el mcm de las longitudes de sus ciclos. Todo ciclo es producto de *trasposiciones*: $(a_1a_2\cdots a_k)=(a_1a_k)\cdots(a_1a_3)(a_1a_2)$.

Teorema 25.3 (Signo)**.** *La paridad del número de trasposiciones en cualquier factorización*

de σ es invariante. Esto define el homomorfismo $\operatorname{sgn} : S_n \to \{\pm 1\}$*, cuyo núcleo es el* grupo alternante A_n *de las permutaciones pares, con* $|A_n| = n!/2$*. (El signo es el mismo que gobierna la fórmula de Leibniz del determinante.)*

Idea clave

Qué dice esto en cristiano. Una permutación se puede descomponer en intercambios de dos elementos de infinitas maneras distintas, y con distinto número de intercambios cada vez. Lo que el teorema afirma es que ese número, aunque cambie, *siempre tiene la misma paridad*: si una descomposición usa un número par, todas lo usarán. Es como una moneda pegada a cada permutación: par o impar, y no hay forma de cambiarla. Gracias a eso se puede definir el signo, y el signo es lo que separa el mundo en dos mitades exactas y da lugar al grupo alternado.

Ejemplo 25.4 (La paridad, comprobada de varias maneras)**. (a) Contar por ciclos.** Un ciclo de longitud L se escribe con $L-1$ trasposiciones. Así, $(0\,1)$ usa 1 y es impar; $(0\,1\,2)$ usa 2 y es par; $(0\,1\,2\,3)$ usa 3 y es impar; y $(0\,1)(2\,3)$ usa $1+1=2$ y es par. La regla práctica es: un ciclo de longitud par es una permutación impar, y viceversa.

(b) Dos descomposiciones de la misma permutación. El ciclo $(0\,1\,2)$ se escribe como $(0\,1)(1\,2)$, con dos trasposiciones. Pero también como $(0\,2)(0\,1)(0\,2)(0\,1)$, con cuatro. Dos y cuatro son distintos, pero ambos son pares: eso es exactamente lo que el teorema garantiza. Lo que no puede ocurrir jamás es escribirla con tres.

(c) La mitad exacta. En S_4, que tiene 24 elementos, hay 12 pares y 12 impares. No es casualidad: multiplicar por una trasposición fija empareja cada permutación par con una impar, luego hay las mismas de cada clase. Por eso $|A_n| = n!/2$ siempre, y $|A_4| = 12$.

(d) El signo es un determinante. Si a cada permutación le asociamos su matriz (un 1 en la posición $(\sigma(i), i)$ y ceros en el resto), el signo resulta ser exactamente el determinante de esa matriz. Para $(0\,1)$ el determinante vale -1; para $(0\,1\,2)$ vale $+1$. Esta es la razón profunda de que el determinante cambie de signo al intercambiar dos filas: el determinante *es* el signo, extendido a matrices cualesquiera.

Cuidado

[El número de trasposiciones no está determinado, solo su paridad] Es un error frecuente decir que "toda permutación se escribe de forma única como producto de trasposiciones". No es cierto ni en número ni en elección: el apartado (b) da dos escrituras distintas de la misma permutación. Lo único invariante es la paridad. Segundo aviso: la descomposición en *ciclos disjuntos* sí es esencialmente única, pero esa es otra descomposición y no debe confundirse con la de trasposiciones, que nunca son disjuntas entre sí.

Dos permutaciones son *conjugadas* en S_n ($\tau = \rho\sigma\rho^{-1}$) si y solo si tienen el mismo *tipo cíclico* (las mismas longitudes de ciclos): conjugar es simplemente "renombrar los objetos permutados".

Regla: descomponer una permutación y leer en ella el orden y el signo

Una permutación de $\{1,\dots,n\}$ suele venir dada en *notación de dos líneas* (arriba las entradas, abajo sus imágenes). Casi nada se ve en esa forma; todo se ve en ciclos. El procedimiento es mecánico.

1. **Persigue órbitas.** Empieza en el 1, escribe su imagen, luego la imagen de la imagen, y sigue hasta volver al 1: ese es el primer ciclo, y se cierra siempre (el conjunto es finito y σ es inyectiva). Toma el menor número no usado y repite. Al terminar tienes σ escrita como producto de ciclos *disjuntos*.

2. **Los ciclos disjuntos conmutan.** Por eso el orden en que los escribas da igual, y por eso todo lo que sigue funciona ciclo a ciclo, sin interferencias.

3. **El orden es el mínimo común múltiplo de las longitudes.** Un ciclo de longitud ℓ tiene orden ℓ; como los ciclos conmutan, $\sigma^m = \mathrm{id}$ exige que cada ciclo vuelva a su sitio, es decir que $\ell_i \mid m$ para todo i. Luego
$$\operatorname{ord}(\sigma) = \operatorname{mcm}(\ell_1,\dots,\ell_r).$$

4. **El signo se cuenta con las longitudes.** Un ciclo de longitud ℓ es producto de $\ell-1$ transposiciones, así que su signo es $(-1)^{\ell-1}$: los ciclos de longitud *impar* son pares y los de longitud *par* son impares. Multiplicando,
$$\operatorname{sgn}(\sigma) = (-1)^{\sum(\ell_i-1)} = (-1)^{n-r},$$
contando también los puntos fijos como ciclos de longitud 1. Regla de bolsillo: *cuenta cuántos ciclos de longitud par hay; si son un número par, σ es par.*

5. **Para elevar a una potencia, eleva cada ciclo por separado.** Un ciclo de longitud ℓ elevado a m se rompe en $d = \gcd(\ell, m)$ ciclos de longitud ℓ/d. En particular sobrevive entero si y solo si $\gcd(\ell,m)=1$, y se desintegra en puntos fijos si $\ell \mid m$.

6. **Para conjugar, renombra.** $\tau\sigma\tau^{-1}$ se obtiene aplicando τ a cada símbolo dentro de los ciclos de σ, sin tocar la estructura. Consecuencia grande: *dos permutaciones son conjugadas en S_n si y solo si tienen el mismo tipo de ciclos*, y por tanto las clases de conjugación de S_n están en biyección con las particiones de n.

Ejemplo 25.5 (Una permutación de S_9, analizada del todo)**.** Sea $\sigma \in S_9$ dada por
$$\sigma = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 \\ 3 & 1 & 5 & 7 & 2 & 8 & 4 & 9 & 6 \end{pmatrix}.$$

Ciclos. Empezamos en 1: $1 \mapsto 3 \mapsto 5 \mapsto 2 \mapsto 1$, ciclo cerrado $(1\,3\,5\,2)$. El menor no usado es 4: $4 \mapsto 7 \mapsto 4$, ciclo $(4\,7)$. El siguiente es 6: $6 \mapsto 8 \mapsto 9 \mapsto 6$, ciclo $(6\,8\,9)$. Todos los símbolos están usados, así que
$$\sigma = (1\,3\,5\,2)(4\,7)(6\,8\,9).$$

Orden. Longitudes $4, 2, 3$, luego $\operatorname{ord}(\sigma) = \operatorname{mcm}(4, 2, 3) = 12$. Es decir: hay que aplicar σ doce veces para volver al principio, aunque ningún ciclo individual necesite más de cuatro.

Signo. Hay dos ciclos de longitud par (el de 4 y el de 2), y dos es par, luego σ es **par**: $\operatorname{sgn}(\sigma) = +1$. Con la otra fórmula: $(4-1)+(2-1)+(3-1) = 6$ transposiciones, y $(-1)^6 = +1$. Coincide. En particular $\sigma \in A_9$.

Potencias. Aplicando la regla del paso 5:

$$\sigma^2 = (1\,5)(2\,3)(6\,9\,8), \qquad \sigma^3 = (1\,2\,5\,3)(4\,7), \qquad \sigma^6 = (1\,5)(2\,3).$$

Léase con cuidado lo que ha pasado en cada caso. En σ^2 el 4-ciclo se partió en $\gcd(4,2) = 2$ ciclos de longitud 2, el 2-ciclo murió $(2 \mid 2)$ y el 3-ciclo sobrevivió pero invertido. Los órdenes son $\operatorname{mcm}(2,3) = 6$ y $\operatorname{mcm}(4,2) = 4$ y 2, respectivamente, y en efecto $12/\gcd(12, m)$ da $6, 4, 2$ para $m = 2, 3, 6$: la fórmula del orden de una potencia en un grupo cíclico, verificada sobre el terreno.

Conjugación. Si $\tau = (1\,6)(3\,4)$, entonces $\tau\sigma\tau^{-1}$ se obtiene sustituyendo cada símbolo por su imagen bajo τ:

$$\tau\sigma\tau^{-1} = (6\,4\,5\,2)(3\,7)(1\,8\,9).$$

Mismo tipo de ciclos $4+2+3$, como tenía que ser. Y como $9 = 4+2+3$ es una partición de 9, esta clase de conjugación es una de las 30 clases de S_9 — tantas como particiones de 9.

25.1.2. Subgrupos y grupos cíclicos

Idea clave

Intuición primero. Un subgrupo es un "mundo cerrado dentro del mundo": si te quedas dentro de H y operas, nunca sales. Un grupo *cíclico* es el caso más simple posible: todo se obtiene repitiendo *una* acción, como el cuentakilómetros o el reloj. Y aquí aparece un patrón que se repetirá todo el capítulo: los grupos cíclicos finitos son *exactamente* el reloj de n horas, sus subgrupos corresponden uno a uno con los divisores de n, y sus generadores son los k coprimos con n (avanzar de 5 en 5 en un reloj de 12 recorre las 12 horas; avanzar de 4 en 4 solo visita $0, 4, 8$). Toda la teoría de números del capítulo 3 está escondida en esta frase.

Definición 25.6. $H \leq G$ es *subgrupo* si es no vacío y cerrado bajo la operación y los inversos. El subgrupo *generado* por a es $\langle a \rangle = \{a^k : k \in \mathbb{Z}\}$; su orden es el *orden del elemento* a (el menor $n > 0$ con $a^n = e$, si existe). Un grupo es *cíclico* si lo genera un solo elemento.

Idea clave

Qué dice esto en cristiano. Un subgrupo es un trozo del grupo que *sigue siendo un grupo con la misma operación*: si multiplicas dos elementos del trozo no te sales del trozo, y el inverso de cada uno también está dentro. El subgrupo *generado* por a es lo mínimo que te ves obligado a incluir si metes a: hay que meter a, entonces $a \cdot a$, entonces $a \cdot a \cdot a \ldots$ y también a^{-1}. El *orden de* a es cuándo esa lista se cierra y vuelve al neutro. Y un grupo es cíclico cuando un solo elemento, dando vueltas, produce el grupo entero.

Ejemplo 25.7 (Todos los subgrupos de $\mathbb{Z}_{12}$, uno por uno)**. (a) El orden de cada elemento.** En $\mathbb{Z}_{12}$ (los restos módulo 12 con la suma) el orden de k es el menor $n > 0$ con $nk \equiv 0$, que resulta ser $12/\gcd(k, 12)$:

k	0	1	2	3	4	5	6	7	8	9	10	11
orden	1	12	6	4	3	12	2	12	3	4	6	12

(b) Los subgrupos generados. $\langle 3\rangle = \{0, 3, 6, 9\}$ tiene cuatro elementos, y en efecto el orden de 3 es 4. $\langle 4\rangle = \{0, 4, 8\}$ tiene tres. $\langle 6\rangle = \{0, 6\}$ tiene dos. Y $\langle 2\rangle = \{0, 2, 4, 6, 8, 10\}$ tiene seis. Obsérvese que $\langle 8\rangle$ da otra vez $\{0, 4, 8\}$: elementos distintos pueden generar el mismo subgrupo.

(c) La lista completa. Los divisores de 12 son $1, 2, 3, 4, 6, 12$, y hay exactamente un subgrupo de cada uno de esos órdenes: $\{0\}$, $\{0, 6\}$, $\{0, 4, 8\}$, $\{0, 3, 6, 9\}$, $\{0, 2, \ldots, 10\}$ y el total. No hay ningún otro, ni siquiera uno de orden 5, porque 5 no divide a 12.

(d) Los generadores. Los k que generan todo $\mathbb{Z}_{12}$ son $1, 5, 7, 11$, exactamente los coprimos con 12; son cuatro, y $\varphi(12) = 4$. Compárese con $\mathbb{Z}_7$, donde $\varphi(7) = 6$ y *todos* los elementos no nulos son generadores, o con $\mathbb{Z}_{36}$, donde solo lo son $\varphi(36) = 12$ de los 36.

Cuidado

[Contraejemplo: cerrado no basta] El subconjunto $\{0, 1, 2\} \subseteq \mathbb{Z}_{12}$ no es subgrupo, porque $1 + 2 = 3$ se sale. Y $\{1, 2, 3, \ldots\} \subseteq \mathbb{Z}$ con la suma es cerrado bajo la operación pero *no* contiene inversos ni el neutro, así que tampoco lo es: en un grupo infinito hace falta pedir las tres cosas por separado. En cambio, en un grupo *finito* basta con la clausura, porque las potencias de cualquier elemento acaban dando la vuelta y produciendo el neutro y el inverso gratis. Por último, D_4 tiene ocho elementos con órdenes $1, 2, 2, 2, 2, 2, 4, 4$: ninguno de orden 8, luego D_4 *no* es cíclico aunque todos sus elementos generen subgrupos.

Teorema 25.8 (Clasificación de los cíclicos)**.** *Todo grupo cíclico infinito es isomorfo a $(\mathbb{Z}, +)$; todo cíclico de orden n es isomorfo a $(\mathbb{Z}_n, +)$. Todo subgrupo de un cíclico es cíclico, y $\mathbb{Z}_n$ tiene exactamente un subgrupo de orden d por cada divisor $d \mid n$. El elemento a^k genera $\mathbb{Z}_n$ $\iff$ $\gcd(k, n) = 1$; hay $\varphi(n)$ generadores (función de Euler).*

Idea clave

Qué dice esto en cristiano. Solo hay un grupo cíclico de cada tamaño, salvo cambio de nombres. Si el grupo es infinito, es $\mathbb{Z}$ disfrazado; si tiene n elementos, es el reloj de n horas disfrazado. Por tanto, en cuanto compruebas que un grupo lo genera un solo elemento, ya sabes absolutamente todo sobre él. La segunda parte dice que los cíclicos son además los grupos más ordenados que existen: tienen *exactamente* un subgrupo por cada divisor del orden, ni uno más ni uno menos, de modo que su retículo de subgrupos es una copia del retículo de divisores.

Ejemplo 25.9 (El diccionario entre subgrupos y divisores)**. (a) El isomorfismo explícito.** El grupo de las rotaciones de un pentágono regular tiene cinco elementos y lo genera el giro

de 72°; la correspondencia "girar k veces $\leftrightarrow$ el número k" es un isomorfismo con $\mathbb{Z}_5$. El grupo $\{1, -1, \mathrm{i}, -\mathrm{i}\}$ con el producto lo genera i, luego es $\mathbb{Z}_4$: la correspondencia es $\mathrm{i}^k \leftrightarrow k$.

(b) Un subgrupo por divisor, con la fórmula. En $\mathbb{Z}_n$ el subgrupo de orden d es el generado por n/d. Para $n = 12$ y $d = 3$, eso da $\langle 4 \rangle = \{0, 4, 8\}$; para $d = 4$, da $\langle 3 \rangle = \{0, 3, 6, 9\}$. La receta funciona siempre y da todos los subgrupos.

(c) Contar generadores. Los generadores de $\mathbb{Z}_{12}$ son $1, 5, 7, 11$: los k con $\gcd(k, 12) = 1$. Comprobación directa con $k = 5$: los múltiplos de 5 módulo 12 son $5, 10, 3, 8, 1, 6, 11, 4, 9, 2, 7, 0$, que son los doce. Con $k = 8$, en cambio, salen solo $8, 4, 0$: como $\gcd(8, 12) = 4$, genera un subgrupo de orden $12/4 = 3$.

(d) La consecuencia práctica. Como todo subgrupo de un cíclico es cíclico, en $\mathbb{Z}$ los únicos subgrupos son $n\mathbb{Z}$ para cada $n \geq 0$. Eso es exactamente el hecho, que ya usamos en aritmética, de que todo conjunto de enteros cerrado bajo suma y resta consiste en los múltiplos de un solo número: su generador es el máximo común divisor de todos ellos.

Cuidado

[Contraejemplo: fuera de los cíclicos esto se rompe] La frase "un subgrupo por cada divisor" es una propiedad exclusiva de los cíclicos, no un hecho general. El grupo A_4 tiene doce elementos y $6 \mid 12$, y sin embargo *no tiene ningún subgrupo de orden* 6. También falla la unicidad: $\mathbb{Z}_2 \times \mathbb{Z}_2$ tiene orden 4 y contiene *tres* subgrupos distintos de orden 2, mientras que $\mathbb{Z}_4$ tiene uno solo. Precisamente por eso $\mathbb{Z}_4$ y $\mathbb{Z}_2 \times \mathbb{Z}_2$ no son isomorfos, aunque ambos tengan cuatro elementos.

25.2. Anillos desde cero: axiomas, subanillos, ideales por los dos lados y cocientes

Este capítulo va a usar anillos sin descanso: los polinomios, los cuerpos, los módulos y la teoría de Galois entera son anillos vistos desde ángulos distintos. Y sin embargo hay una capa elemental que se emplea a partir de aquí como si estuviera dicha y no lo está: qué es exactamente un anillo cuando no se le exige conmutar, qué es un *subanillo*, cómo se describe el ideal generado por dos elementos —la notación $(2, x)$ aparece varias veces más adelante—, qué es un homomorfismo de anillos y por qué su núcleo es un ideal, qué dice el teorema de correspondencia para ideales, y qué significa el símbolo $\deg f$ cuando los coeficientes no viven en un cuerpo. Esta sección monta todo eso desde el principio, con las demostraciones hechas y con ejemplos resueltos enteros. Quien la lea no necesita nada previo salvo saber qué es un grupo abeliano.

Idea clave

Intuición primero. Un grupo tiene una operación. Un anillo tiene dos, y la segunda es deliberadamente peor que la primera: sumar siempre se deshace, multiplicar casi nunca. Todo el resto del capítulo consiste en preguntar qué privilegios de $\mathbb{Z}$ sobreviven cuando se cambia el conjunto. Y hay una segunda decisión de diseño que conviene ver de entrada: aquí *no* vamos a suponer que el producto conmuta. Las matrices son el

ejemplo cotidiano de anillo donde $AB \neq BA$, y en cuanto se pierde la conmutatividad la palabra "ideal" se desdobla en tres —por la izquierda, por la derecha y por los dos lados—, distinción invisible en $\mathbb{Z}$ y decisiva en cuanto uno mira una matriz.

25.2.1. Los axiomas, y lo que se deduce de ellos

Definición 25.10. Un *anillo* es una terna $(R, +, \cdot)$ formada por un conjunto R y dos operaciones internas que cumplen:

- **A1.** $(R, +)$ es un grupo abeliano: la suma es asociativa y conmutativa, hay un neutro 0 y todo a tiene opuesto $-a$;
- **A2.** el producto es asociativo: $(ab)c = a(bc)$;
- **A3.** valen las dos distributivas: $a(b+c) = ab + ac$ y $(b+c)a = ba + ca$.

El anillo es *unitario* (o *con unidad*) si existe $1 \in R$ con $1a = a1 = a$ para todo a, y *conmutativo* si $ab = ba$ para todos. Un *anillo de división* es un anillo unitario con $1 \neq 0$ en el que todo elemento no nulo tiene inverso multiplicativo; un anillo de división conmutativo es un *cuerpo*.

Observación 25.11. Las dos distributivas son enunciados distintos cuando el producto no conmuta, y hay que exigir las dos. Obsérvese también que no se pide inverso multiplicativo: esa ausencia no es una carencia de la definición, es su contenido. Y que hagan falta *dos* distributivas y ninguna otra ley mixta es lo que impide que un anillo se rompa en dos estructuras independientes puestas encima del mismo conjunto.

Proposición 25.12 (Lo que sale gratis de los axiomas)**.** *En todo anillo R y para todos $a, b \in R$:*

$$0 \cdot a = a \cdot 0 = 0, \qquad (-a)b = a(-b) = -(ab), \qquad (-a)(-b) = ab.$$

Si además R es unitario y $1 = 0$, entonces $R = \{0\}$.

Demostración. Por distributiva, $0 \cdot a = (0+0)a = 0 \cdot a + 0 \cdot a$; restando $0 \cdot a$ en el grupo aditivo queda $0 = 0 \cdot a$, y con la otra distributiva sale $a \cdot 0 = 0$. Para la segunda, $ab + (-a)b = (a + (-a))b = 0 \cdot b = 0$, luego $(-a)b$ es el opuesto de ab; el mismo cálculo por el otro lado da $a(-b) = -(ab)$. La tercera es aplicar la segunda dos veces: $(-a)(-b) = -(a(-b)) = -(-(ab)) = ab$. Por último, si $1 = 0$ entonces $a = 1 \cdot a = 0 \cdot a = 0$ para todo a. □

Idea clave

Qué dice esto en cristiano. "Menos por menos da más" no es una convención pactada en secundaria: es el único valor de $(-1)(-1)$ compatible con la distributiva. Si uno decidiera que $(-1)(-1) = -1$, entonces $0 = (-1) \cdot 0 = (-1)(1 + (-1)) = (-1) + (-1)(-1) = -1 - 1 = -2$, y el anillo se vendría abajo. Lo mismo con $0 \cdot a = 0$: no es un axioma, es un teorema de dos líneas.

Ejemplo 25.13 (Un catálogo, con el detalle de qué cumple cada uno)**. (a)** $\mathbb{Z}$: conmutativo y unitario; las únicas unidades son ± 1.

(b) $2\mathbb{Z}$, los enteros pares: es anillo, y *no* es unitario, porque $1 \notin 2\mathbb{Z}$. Sirve para ver que la unidad no viene regalada con los axiomas A1–A3.

(c) $M_n(K)$, las matrices $n \times n$ sobre un cuerpo K con $n \geq 2$: unitario (la identidad I_n) y **no conmutativo**. Además tiene divisores de cero: con $E = \left(\begin{smallmatrix} 1 & 0 \\ 0 & 0 \end{smallmatrix}\right)$ y $F = \left(\begin{smallmatrix} 0 & 0 \\ 0 & 1 \end{smallmatrix}\right)$ se tiene $EF = 0$ con $E \neq 0 \neq F$.

(d) El anillo producto $R \times S$, con las operaciones componente a componente y unidad $(1_R, 1_S)$. Siempre tiene divisores de cero si ninguno de los dos factores es nulo: $(1,0)\cdot(0,1) = (0,0)$. Esta construcción es la que aparece al otro lado del teorema chino del resto.

(e) Las funciones continuas $C[0,1]$ con suma y producto punto a punto: conmutativo y unitario (la función constante 1), y con divisores de cero abundantes —dos funciones continuas no nulas con soportes disjuntos—.

(f) Los cuaterniones de Hamilton $\mathbb{H} = \{a + b\mathrm{i} + c\mathrm{j} + d\mathrm{k}\}$ con $\mathrm{i}^2 = \mathrm{j}^2 = \mathrm{k}^2 = \mathrm{ijk} = -1$: anillo de división *no* conmutativo, pues $\mathrm{ij} = \mathrm{k}$ y $\mathrm{ji} = -\mathrm{k}$. Se puede dividir y aun así no es un cuerpo.

Definición 25.14. La *característica* de un anillo unitario R, escrita $\operatorname{car} R$, es el menor entero $n > 0$ tal que $n \cdot 1 = \underbrace{1 + \cdots + 1}_{n} = 0$, y es 0 si ningún múltiplo positivo de 1 se anula.

Ejemplo 25.15. $\operatorname{car}\mathbb{Z} = 0$ y $\operatorname{car}\mathbb{Q} = \operatorname{car}\mathbb{R} = \operatorname{car}\mathbb{C} = 0$. $\operatorname{car}\mathbb{Z}_n = n$, sea n primo o no: $\operatorname{car}\mathbb{Z}_6 = 6$. En un producto, $\operatorname{car}(\mathbb{Z}_4 \times \mathbb{Z}_6) = \operatorname{mcm}(4,6) = 12$, porque $m\cdot(1,1) = (0,0)$ exige que m sea múltiplo de 4 y de 6 a la vez. La característica de $\mathbb{F}_4$ es 2 aunque el cuerpo tenga cuatro elementos: característica y cardinal son cosas distintas. Y en un *dominio* la característica solo puede ser 0 o prima, porque si $n = ab$ con $1 < a, b < n$ entonces $(a\cdot 1)(b\cdot 1) = n \cdot 1 = 0$ con ambos factores no nulos.

25.2.2. Subanillos: la palabra que el capítulo usa sin decirla

Definición 25.16. Sea R un anillo unitario. Un subconjunto $S \subseteq R$ es un *subanillo* si es un subgrupo aditivo de R, es cerrado para el producto y contiene a la unidad: es decir,

$$0 \in S, \qquad a, b \in S \Rightarrow a - b \in S, \qquad a, b \in S \Rightarrow ab \in S, \qquad 1 \in S.$$

Con esas condiciones S es por sí mismo un anillo unitario con las operaciones heredadas.

Cuidado

Hay dos convenios en circulación: algunos autores no exigen $1 \in S$. La diferencia no es cosmética. Con el convenio de arriba, $2\mathbb{Z}$ *no* es subanillo de $\mathbb{Z}$; sin él, sí. Y hay un tercer fenómeno que despista aún más: un subconjunto puede tener unidad propia distinta de la del anillo grande, como se verá en el ejemplo siguiente. En este libro "subanillo" significa siempre lo de la definición 25.16, con la unidad del anillo ambiente.

Ejemplo 25.17 (Subanillos y falsos subanillos, comprobados)**. (a) La cadena numérica.** $\mathbb{Z} \subset \mathbb{Q} \subset \mathbb{R} \subset \mathbb{C}$: cada uno es subanillo del siguiente, porque las cuatro condiciones son inmediatas y el 1 es el mismo en todos.

(b) Los enteros de Gauss. $\mathbb{Z}[\mathrm{i}] = \{a + b\mathrm{i} : a, b \in \mathbb{Z}\}$ es subanillo de $\mathbb{C}$: la resta es obvia, el producto lo es porque $(a + b\mathrm{i})(c + d\mathrm{i}) = (ac - bd) + (ad + bc)\mathrm{i}$ vuelve a tener coordenadas enteras, y $1 = 1 + 0\mathrm{i}$ está dentro. Lo mismo con $\mathbb{Z}[\sqrt{2}] = \{a + b\sqrt{2}\}$, porque $(a + b\sqrt{2})(c + d\sqrt{2}) = (ac + 2bd) + (ad + bc)\sqrt{2}$.

(c) Los únicos subanillos de $\mathbb{Z}$. Un subgrupo aditivo de $\mathbb{Z}$ es $n\mathbb{Z}$; para contener al 1 hace falta $n = 1$. Luego $\mathbb{Z}$ tiene un único subanillo, él mismo. Los $n\mathbb{Z}$ con $n \geq 2$ son cerrados para el producto pero se quedan sin unidad.

(d) Las matrices diagonales. Dentro de $M_2(\mathbb{R})$, el conjunto de las diagonales es subanillo: la resta y el producto de diagonales son diagonales, e I_2 es diagonal. En cambio las matrices *triangulares estrictas* $\left(\begin{smallmatrix} 0 & b \\ 0 & 0 \end{smallmatrix}\right)$ no lo son: no contienen a I_2, y de hecho su producto es siempre cero.

(e) Un subconjunto con unidad propia. En $R = \mathbb{Z}_{12}$ tómese $S = \{0, 4, 8\}$. Es subgrupo aditivo y es cerrado para el producto ($4 \cdot 4 = 16 = 4$, $4 \cdot 8 = 32 = 8$, $8 \cdot 8 = 64 = 4$). Pero $1 \notin S$: no es subanillo en nuestro sentido. Lo curioso es que S *tiene* unidad propia, a saber el 4, porque $4 \cdot 4 = 4$ y $4 \cdot 8 = 8$; con ella, $S \cong \mathbb{Z}_3$. Lo mismo ocurre con $T = \{0, 3, 6, 9\}$, cuya unidad propia es el 9 ($9 \cdot 9 = 81 = 9$, $9 \cdot 3 = 27 = 3$, $9 \cdot 6 = 54 = 6$) y que resulta ser $\cong \mathbb{Z}_4$. Ninguno de los dos es subanillo de $\mathbb{Z}_{12}$, y ambos son ideales. Los elementos 4 y 9 son precisamente los idempotentes no triviales de $\mathbb{Z}_{12}$: la descomposición $\mathbb{Z}_{12} \cong \mathbb{Z}_3 \times \mathbb{Z}_4$ está escrita ahí. ✓

25.2.3. Ideales: uno, dos o tres tipos según conmute o no

Definición 25.18. Sea R un anillo. Un subgrupo aditivo $I \subseteq R$ es

- un *ideal por la izquierda* si $ra \in I$ para todos $r \in R$, $a \in I$;
- un *ideal por la derecha* si $ar \in I$ para todos $r \in R$, $a \in I$;
- un *ideal bilátero* (o simplemente *ideal*) si es las dos cosas a la vez.

Si R es conmutativo las tres nociones coinciden y se dice sencillamente *ideal*. Un ideal I es *propio* si $I \neq R$, y un anillo unitario con $1 \neq 0$ cuyos únicos ideales biláteros son $\{0\}$ y él mismo se llama *simple*.

Observación 25.19. El criterio rápido para "propio" es: un ideal I de un anillo unitario es todo el anillo si y solo si $1 \in I$, y también si y solo si I contiene alguna unidad. En efecto, si $u \in I$ es unidad entonces $1 = u^{-1}u \in I$ por absorción, y de $1 \in I$ sale $r = r \cdot 1 \in I$ para todo r. Esa observación de una línea es la que se usa continuamente para comprobar que un ideal es propio.

Definición 25.20 (Ideal generado)**.** En un anillo conmutativo unitario R, el *ideal generado* por $a_1, \ldots, a_n$ es

$$(a_1, \ldots, a_n) \;=\; \Big\{ r_1 a_1 + r_2 a_2 + \cdots + r_n a_n \;:\; r_i \in R \Big\},$$

el menor ideal que los contiene a todos. Si $n = 1$ el ideal $(a) = aR$ se llama *principal*. Dados dos ideales I, J se definen además

$$I + J = \{a + b : a \in I,\ b \in J\}, \qquad I \cap J, \qquad IJ = \left\{\sum_{k=1}^{m} a_k b_k :\ a_k \in I,\ b_k \in J\right\},$$

y los tres son ideales de R.

Demostración. Que el conjunto descrito es un ideal es inmediato: la diferencia de dos combinaciones $\sum r_i a_i$ es otra, y multiplicar por $s \in R$ da $\sum (sr_i)a_i$. Contiene a cada a_j (tómese $r_j = 1$ y los demás nulos), y cualquier ideal que contenga a los a_i contiene, por absorción y por ser subgrupo, a todas esas combinaciones: es por tanto el menor. En IJ hay que tomar *sumas* de productos y no solo productos sueltos, precisamente para que el conjunto sea cerrado para la suma. □

Regla: comprobar si un subconjunto es subanillo, ideal, o ninguna de las dos cosas

Es la comprobación más frecuente del tema y se hace siempre en el mismo orden, del test más barato al más caro. Nótese que subanillo e ideal son condiciones *incompatibles* en un anillo unitario salvo en el caso extremo $I = R$: un ideal propio nunca contiene al 1.

1. **Comprueba que no es vacío y que** 0 **está dentro.** Si el 0 falla, ya está: ni subanillo ni ideal. Es el descarte más rápido que existe.
2. **Comprueba el subgrupo aditivo con el criterio de la resta:** $a, b \in S \Rightarrow a - b \in S$. Una sola condición cubre a la vez el cierre para la suma y la existencia de opuestos. Si esto falla no hay nada más que mirar. Ejemplo de fallo: los enteros impares, donde $3 + 5 = 8$ se sale.
3. **Decide qué estás buscando.** Si buscas *subanillo*, comprueba $a, b \in S \Rightarrow ab \in S$ y, sobre todo, $1 \in S$. Si buscas *ideal*, comprueba la absorción $r \in R,\ a \in I \Rightarrow ra \in I$, que es mucho más fuerte: el multiplicador r recorre *todo* el anillo, no solo I.
4. **Si el anillo no es conmutativo, comprueba la absorción por los dos lados** separadamente. Es habitual que se cumpla una sola. Trabajar en $M_n(K)$ y decir "ideal" a secas es el error clásico del tema.
5. **Para descartar que un ideal sea todo el anillo, busca el** 1 **o una unidad.** Si $1 \in I$ o si I contiene una unidad, entonces $I = R$. Si no, ya sabes que es propio.
6. **Si el subconjunto viene descrito por generadores,** usa la definición 25.20: $(a_1, \ldots, a_n)$ es el conjunto de combinaciones $\sum r_i a_i$. Para comprobar que un elemento concreto está o no está, la vía práctica es buscar un invariante que todos los generadores compartan —una paridad, un valor en un punto, un grado, un resto módulo algo— y verlo fallar en el candidato.
7. **Control final.** Un ideal es núcleo de algún homomorfismo y un subanillo es imagen de alguno. Si ves de qué homomorfismo se trata, has terminado sin hacer cuentas.

Ejemplo 25.21 (Siete subconjuntos clasificados)**. (a)** $n\mathbb{Z} \subset \mathbb{Z}$**.** Ideal, y principal: $n\mathbb{Z} = (n)$. No es subanillo si $n \neq \pm 1$, porque no contiene al 1. Y son *todos*: si $I \neq \{0\}$ es un ideal de $\mathbb{Z}$ y n es su menor elemento positivo, la división $a = qn + r$ con $0 \leq r < n$ da $r = a - qn \in I$, luego $r = 0$ por minimalidad y $I = (n)$.

(b) $\mathbb{Z} \subset \mathbb{Q}$**.** Subanillo, y *no* ideal: $\frac{1}{2} \cdot 1 = \frac{1}{2} \notin \mathbb{Z}$, así que falla la absorción. Es el contraste que conviene memorizar: los subanillos se parecen al ambiente, los ideales son absorbentes.

(c) $(4, 6)$ **en** $\mathbb{Z}$**.** Por la definición, es $\{4r + 6s\}$, y por Bézout ese conjunto es exactamente el de los múltiplos de $\gcd(4, 6) = 2$, con $2 = 4 \cdot (-1) + 6 \cdot 1$. Luego $(4, 6) = (2)$. En general, en $\mathbb{Z}$ se cumple $(a, b) = (\gcd(a, b))$: la suma de ideales *es* el máximo común divisor. Y el producto es $(4)(6) = (24)$, la intersección es $(4) \cap (6) = (12) = (\operatorname{mcm}(4, 6))$. Obsérvese que $IJ \subseteq I \cap J$ y que la inclusión suele ser estricta: $(24) \subsetneq (12)$.

(d) $(x^2 - 1,\ x^3 - 1)$ **en** $\mathbb{Q}[x]$**.** El algoritmo de Euclides da $\gcd(x^2 - 1, x^3 - 1) = x - 1$ —de hecho $x^3 - 1 = x(x^2 - 1) + (x - 1)$ y $x^2 - 1 = (x + 1)(x - 1)$—, luego el ideal es $(x - 1)$, otra vez principal. En $K[x]$, igual que en $\mathbb{Z}$, dos generadores siempre se funden en uno. ✓

(e) $(2, x)$ **en** $\mathbb{Z}[x]$**.** Aquí no se funden. Por la definición, $(2, x) = \{2f + xg\}$, que es exactamente el conjunto de polinomios con término independiente par —el invariante del paso 6 de la regla—. Si fuese (h), entonces $h \mid 2$ obligaría a h constante, $h \in \{\pm 1, \pm 2\}$; pero ± 1 genera todo el anillo y $1 \notin (2, x)$, y ± 2 no divide a x en $\mathbb{Z}[x]$. No es principal, y ése es el testigo de que $\mathbb{Z}[x]$ no es un dominio de ideales principales.

(f) $(3,\ x^2 + 1)$ **en** $\mathbb{Z}[x]$**.** Reducir módulo 3 y luego módulo $x^2 + 1$ da un homomorfismo sobreyectivo $\mathbb{Z}[x] \to \mathbb{F}_3[x]/(x^2 + 1)$ cuyo núcleo es justamente este ideal. Como $x^2 + 1$ no tiene raíces en $\mathbb{F}_3$ —vale 1, 2 y 2 en $0, 1, 2$— y tiene grado 2, es irreducible allí y el cociente es el cuerpo $\mathbb{F}_9$. Conclusión: $(3, x^2 + 1)$ es un ideal *maximal* de $\mathbb{Z}[x]$ que necesita dos generadores. ✓

(g) Las matrices con segunda columna nula en $M_2(K)$**.** Sea $L = \left\{ \left(\begin{smallmatrix} a & 0 \\ c & 0 \end{smallmatrix}\right) \right\}$. Multiplicar por la izquierda por cualquier B actúa columna a columna, luego BA vuelve a tener la segunda columna nula: L es ideal por la izquierda. **No** es ideal por la derecha: $\left(\begin{smallmatrix} 1 & 0 \\ 0 & 0 \end{smallmatrix}\right)\left(\begin{smallmatrix} 0 & 1 \\ 0 & 0 \end{smallmatrix}\right) = \left(\begin{smallmatrix} 0 & 1 \\ 0 & 0 \end{smallmatrix}\right)$, que se sale. Simétricamente, las matrices con segunda fila nula forman un ideal por la derecha y no por la izquierda. Sobre $K = \mathbb{F}_2$ ambos conjuntos tienen 4 elementos de los 16 del anillo, y la comprobación exhaustiva confirma las dos afirmaciones. ✓

Teorema 25.22 ($M_n(K)$ es simple)**.** *Si K es un cuerpo y $n \geq 1$, los únicos ideales bilateros de $M_n(K)$ son $\{0\}$ y $M_n(K)$.*

Demostración. Sea $I \neq \{0\}$ un ideal bilátero y tómese $A \in I$ con alguna entrada $a_{k\ell} \neq 0$. Sean E_{ij} las matrices que valen 1 en la posición (i, j) y 0 en las demás; se comprueba directamente que $E_{ik} A E_{\ell j} = a_{k\ell} E_{ij}$. Como I absorbe por los dos lados, esa matriz está en I, y multiplicando por el escalar $a_{k\ell}^{-1}$ —lícito porque K es cuerpo— queda $E_{ij} \in I$ para todos i, j. Sumando las E_{ii} se obtiene $I_n \in I$, luego I contiene una unidad y es todo el anillo. □

Observación 25.23. Este teorema es la razón de que la distinción entre los tres tipos de ideal sea tan importante: $M_2(K)$ tiene ideales por la izquierda y por la derecha en abundancia —uno por cada subespacio de K^2, de hecho— y ninguno bilátero. Un anillo puede estar lleno de estructura lateral y ser simple visto por los dos lados a la vez.

25.2.4. Homomorfismos, cocientes y los teoremas de isomorfía

Definición 25.24. Sean R y S anillos unitarios. Una aplicación $\varphi : R \to S$ es un *homomorfismo de anillos* si para todos $a, b \in R$

$$\varphi(a+b) = \varphi(a) + \varphi(b), \qquad \varphi(ab) = \varphi(a)\varphi(b), \qquad \varphi(1_R) = 1_S.$$

Su *núcleo* es $\ker\varphi = \{a \in R : \varphi(a) = 0\}$ y su *imagen* es $\operatorname{im}\varphi = \varphi(R)$. Si φ es biyectiva se llama *isomorfismo* y se escribe $R \cong S$; un isomorfismo de R en sí mismo es un *automorfismo*.

Proposición 25.25. *Si $\varphi : R \to S$ es homomorfismo de anillos, entonces $\ker\varphi$ es un ideal bilátero de R, $\operatorname{im}\varphi$ es un subanillo de S, y φ es inyectiva si y solo si $\ker\varphi = \{0\}$.*

Demostración. Si $a, b \in \ker\varphi$ entonces $\varphi(a-b) = \varphi(a) - \varphi(b) = 0$, luego el núcleo es subgrupo aditivo. Si $a \in \ker\varphi$ y $r \in R$, entonces $\varphi(ra) = \varphi(r)\varphi(a) = \varphi(r)\cdot 0 = 0$ y análogamente $\varphi(ar) = 0$: absorbe por los dos lados. La imagen es cerrada para resta y producto porque φ los respeta, y contiene a $1_S = \varphi(1_R)$. Para lo último, si $\varphi(a) = \varphi(b)$ entonces $a - b \in \ker\varphi$. □

Cuidado

El núcleo es un ideal y *no* un subanillo: no contiene al 1 salvo que S sea el anillo nulo. La imagen es un subanillo y *no* un ideal en general: la inclusión $\mathbb{Z} \hookrightarrow \mathbb{Q}$ tiene imagen $\mathbb{Z}$, que no absorbe. Confundir los dos papeles es el error más común al empezar.

Teorema 25.26 (Los tres teoremas de isomorfía y la correspondencia, para anillos)**.** *Sea R un anillo unitario e I un ideal bilátero de R. El cociente R/I, formado por las clases $a+I$ con las operaciones $(a+I)+(b+I) = (a+b)+I$ y $(a+I)(b+I) = ab+I$, es un anillo, y la proyección $\pi : R \to R/I$ es un homomorfismo sobreyectivo de núcleo I. Además:*

(i) Primer teorema. *Para todo homomorfismo $\varphi : R \to S$, $R/\ker\varphi \cong \operatorname{im}\varphi$.*

(ii) Correspondencia. *La aplicación $J \mapsto J/I$ es una biyección, que respeta las inclusiones, entre los ideales de R que contienen a I y los ideales de R/I; la inversa es $\bar{J} \mapsto \pi^{-1}(\bar{J})$.*

(iii) Segundo teorema. *Si S es un subanillo de R e I un ideal, entonces $S+I$ es un subanillo, $S \cap I$ es un ideal de S, y $(S+I)/I \cong S/(S\cap I)$.*

(iv) Tercer teorema. *Si $I \subseteq J$ son ideales de R, entonces J/I es un ideal de R/I y $(R/I)/(J/I) \cong R/J$: las fracciones se simplifican.*

Demostración. Que el producto de clases está bien definido es donde se usa la absorción: si $a' = a+i$ y $b' = b+j$ con $i, j \in I$, entonces $a'b' = ab + (aj + ib + ij)$ y los tres sumandos del paréntesis están en I.

(i) La aplicación $\bar{\varphi}(a + \ker\varphi) = \varphi(a)$ está bien definida —dos representantes difieren en el núcleo—, es homomorfismo porque φ lo es, es sobreyectiva sobre la imagen por construcción, y es inyectiva porque $\bar{\varphi}(a+\ker\varphi) = 0$ significa $a \in \ker\varphi$.

(ii) Si $J \supseteq I$ es ideal, $\pi(J) = J/I$ es ideal de R/I por sobreyectividad de π. Si $\bar{J}$ es ideal de R/I, su preimagen $\pi^{-1}(\bar{J})$ es ideal de R y contiene a $\ker\pi = I$. Las dos construcciones

se deshacen mutuamente: $\pi^{-1}(\pi(J)) = J + I = J$ porque $I \subseteq J$, y $\pi(\pi^{-1}(\bar{J})) = \bar{J}$ por sobreyectividad. Las inclusiones se conservan en ambos sentidos porque ambas aplicaciones son monótonas.

(iii) Compóngase la inclusión $S \hookrightarrow R$ con $\pi : R \to R/I$. El resultado es un homomorfismo $S \to R/I$ de imagen $(S+I)/I$ y de núcleo $S \cap I$; aplíquese (i).

(iv) Compóngase $\pi_I : R \to R/I$ con $\pi_{J/I} : R/I \to (R/I)/(J/I)$. La composición es sobreyectiva y su núcleo es $\{a : a + I \in J/I\} = J$; aplíquese (i) otra vez. □

Ejemplo 25.27 (Los cuatro enunciados con $\mathbb{Z}$ y con $\mathbb{Z}_{12}$)**. (a) Primer teorema.** La reducción $\mathbb{Z} \to \mathbb{Z}_n$ tiene núcleo (n) e imagen todo $\mathbb{Z}_n$, luego $\mathbb{Z}/(n) \cong \mathbb{Z}_n$. La evaluación $\mathbb{Z}[x] \to \mathbb{Z}$, $f \mapsto f(0)$, tiene núcleo (x), luego $\mathbb{Z}[x]/(x) \cong \mathbb{Z}$: como $\mathbb{Z}$ es dominio y no cuerpo, (x) es primo y no maximal.

(b) Correspondencia en $\mathbb{Z}_{12}$. Los ideales de $\mathbb{Z} = \mathbb{Z}$ que contienen a (12) son exactamente los (d) con $d \mid 12$, es decir seis: $(1), (2), (3), (4), (6), (12)$. La correspondencia dice que $\mathbb{Z}_{12}$ tiene exactamente seis ideales, y en efecto son

$$\mathbb{Z}_{12}, \quad \{0,2,4,6,8,10\}, \quad \{0,3,6,9\}, \quad \{0,4,8\}, \quad \{0,6\}, \quad \{0\}.$$

El retículo de ideales de $\mathbb{Z}_{12}$ es, letra por letra, el retículo de divisores de 12 puesto del revés. ✓

(c) Tercer teorema. Con $I = (12) \subseteq J = (4)$ dentro de $\mathbb{Z}$, el enunciado da

$$\mathbb{Z}_{12}/\{0,4,8\} \;=\; (\mathbb{Z}/(12))/((4)/(12)) \;\cong\; \mathbb{Z}/(4) = \mathbb{Z}_4,$$

y las cuentas cuadran: el de la izquierda tiene $12/3 = 4$ elementos. Es la misma simplificación de fracciones que en grupos.

(d) Segundo teorema. En $R = \mathbb{Z}$, con el subanillo $S = \mathbb{Z}$ y el ideal $I = (6)$ no se ve nada; el enunciado luce cuando el subanillo es pequeño. Tómese $R = \mathbb{Q}[x]$, $S = \mathbb{Q}$ (las constantes) e $I = (x^2+1)$. Entonces $S \cap I = \{0\}$ —una constante no nula no es múltiplo de x^2+1 por grados— y el teorema da $(\mathbb{Q} + I)/I \cong \mathbb{Q}$: la copia de $\mathbb{Q}$ dentro de $\mathbb{Q}[x]/(x^2+1) = \mathbb{Q}(\mathrm{i})$ es fiel, ningún racional se pierde al cocientar. Ese es el mecanismo por el que un cociente por un irreducible siempre *contiene* al cuerpo de partida.

Regla: construir un anillo cociente y operar dentro de él

El cociente es la máquina de fabricar anillos nuevos, y en la práctica todo se reduce a cinco gestos mecánicos.

1. **Asegúrate de que cocientas por un ideal bilátero.** Si el anillo es conmutativo no hay nada que comprobar más allá de la absorción. Si no lo es, un ideal por un solo lado *no* sirve: el producto de clases no queda bien definido.
2. **Elige representantes canónicos.** En $\mathbb{Z}/(n)$ son los restos $0, \ldots, n-1$. En $K[x]/(f)$ con $\deg f = m$ son los polinomios de grado $< m$, por el teorema de la división: cada clase tiene uno y solo uno. Esa elección convierte el cociente en un objeto concreto y contable.

3. **Cuenta los elementos antes de operar.** Si K es finito con q elementos, $K[x]/(f)$ tiene $q^{\deg f}$ elementos, porque hay $\deg f$ coeficientes libres. Es la primera verificación de que no te has equivocado de conjunto.

4. **Opera como siempre y reduce al final.** Suma y multiplica los representantes en el anillo grande y luego toma resto. La relación $\bar{f} = 0$ es la única regla nueva: en $K[x]/(f)$ con f mónico de grado m, sirve para reescribir $\bar{x}^{\,m}$ en términos de potencias menores, y aplicándola repetidamente se baja cualquier grado.

5. **Identifica el resultado.** Si f factoriza en factores coprimos, $f = gh$, el teorema chino da $K[x]/(f) \cong K[x]/(g) \times K[x]/(h)$ y aparecen divisores de cero e idempotentes. Si f es irreducible, el cociente es un cuerpo. Si f tiene factores repetidos, aparecen nilpotentes. Esos tres casos agotan las posibilidades y se distinguen mirando $\gcd(f, f')$.

Ejemplo 25.28 (Un cociente entero sobre el cuerpo de dos elementos)**.** Tomemos $K = \mathbb{F}_2$ y $f = x^3 + 1$. Nótese que f *no* es irreducible: en característica 2,

$$x^3 + 1 = (x + 1)(x^2 + x + 1),$$

como se comprueba multiplicando, y los dos factores son coprimos porque $x^2 + x + 1$ no se anula en $x = 1$. Llamemos t a la clase de x, de modo que la regla de reducción es $t^3 = 1$. Por el paso 3 hay $2^3 = 8$ elementos, que son $c_0 + c_1 t + c_2 t^2$ con $c_i \in \mathbb{F}_2$. La tabla completa del producto, obtenida reduciendo con $t^3 = 1$ y verificada por ordenador elemento a elemento, es:

$\cdot$	0	1	t	t^2	$1{+}t$	$1{+}t^2$	$t{+}t^2$	$1{+}t{+}t^2$
0	0	0	0	0	0	0	0	0
1	0	1	t	t^2	$1{+}t$	$1{+}t^2$	$t{+}t^2$	$1{+}t{+}t^2$
t	0	t	t^2	1	$t{+}t^2$	$1{+}t$	$1{+}t^2$	$1{+}t{+}t^2$
t^2	0	t^2	1	t	$1{+}t^2$	$t{+}t^2$	$1{+}t$	$1{+}t{+}t^2$
$1{+}t$	0	$1{+}t$	$t{+}t^2$	$1{+}t^2$	$1{+}t^2$	$t{+}t^2$	$1{+}t$	0
$1{+}t^2$	0	$1{+}t^2$	$1{+}t$	$t{+}t^2$	$t{+}t^2$	$1{+}t$	$1{+}t^2$	0
$t{+}t^2$	0	$t{+}t^2$	$1{+}t^2$	$1{+}t$	$1{+}t$	$1{+}t^2$	$t{+}t^2$	0
$1{+}t{+}t^2$	0	$1{+}t{+}t^2$	$1{+}t{+}t^2$	$1{+}t{+}t^2$	0	0	0	$1{+}t{+}t^2$

Lectura de la tabla. Solo tres filas contienen un 1: las de 1, t y t^2. Luego las unidades son exactamente $\{1, t, t^2\}$, con $t^{-1} = t^2$. Los otros cuatro elementos no nulos —$1 + t$, $1 + t^2$, $t + t^2$ y $1 + t + t^2$— son divisores de cero, y la última fila lo enseña de golpe: $(1 + t + t^2)(1 + t) = 0$. No hay ningún nilpotente distinto de 0, y los idempotentes son cuatro: 0, 1, $t + t^2$ y $1 + t + t^2$.

Por qué sale eso. Por el paso 5 y el teorema chino,

$$\mathbb{F}_2[x]/(x^3 + 1) \;\cong\; \mathbb{F}_2[x]/(x + 1) \times \mathbb{F}_2[x]/(x^2 + x + 1) \;\cong\; \mathbb{F}_2 \times \mathbb{F}_4,$$

y todo cuadra: las unidades del producto son $|\mathbb{F}_2^\times| \cdot |\mathbb{F}_4^\times| = 1 \cdot 3 = 3$; los idempotentes son $2^2 = 4$, uno por cada elección de 0 o 1 en cada factor; y no hay nilpotentes porque $x^3 + 1$ no tiene factores repetidos. El diccionario explícito es

$$1 + t + t^2 \;\longleftrightarrow\; (1, 0), \qquad t + t^2 \;\longleftrightarrow\; (0, 1),$$

que se comprueba evaluando: en el primer factor $t \mapsto 1$, luego $1+t+t^2 \mapsto 1+1+1=1$ y $t+t^2 \mapsto 1+1=0$; en el segundo $t \mapsto a$ con $a^2=a+1$, luego $1+t+t^2 \mapsto 1+a+a+1=0$ y $t+t^2 \mapsto a+a+1=1$. Y en efecto los dos son ortogonales y complementarios: $(1+t+t^2)+(t+t^2)=1$ y $(1+t+t^2)(t+t^2)=0$. ✓

Cuidado

Compárese con $\mathbb{F}_2[x]/(x^3+x+1)$, donde el polinomio *sí* es irreducible: allí los ocho elementos forman el cuerpo $\mathbb{F}_8$ y las siete clases no nulas son unidades. El mismo número de elementos y estructuras completamente distintas. La lección es que el cociente no depende del tamaño del ideal sino de cómo factoriza su generador, y por eso decidir irreducibilidad es la pregunta central del capítulo.

25.2.5. El anillo de polinomios, definido de verdad

Definición 25.29. Sea R un anillo conmutativo unitario. El *anillo de polinomios* $R[x]$ es el conjunto de las sucesiones $(a_0, a_1, a_2, \dots)$ de elementos de R con solo finitos términos no nulos, con la suma componente a componente y el producto de convolución

$$(ab)_k = \sum_{i+j=k} a_i b_j.$$

Escribiendo $x=(0,1,0,\dots)$, cada sucesión se escribe de forma única como $a_0+a_1x+\cdots+a_nx^n$. Para $f \neq 0$, el *grado* $\deg f$ es el mayor índice con $a_i \neq 0$, el *coeficiente principal* es ese $a_{\deg f}$, y f es *mónico* si su coeficiente principal vale 1. Se conviene $\deg 0 = -\infty$.

Observación 25.30. Definir un polinomio como una sucesión de coeficientes y no como "una función" no es pedantería. Sobre $\mathbb{F}_2$, los polinomios x y x^2 definen la misma función $\mathbb{F}_2 \to \mathbb{F}_2$ y sin embargo son elementos distintos de $\mathbb{F}_2[x]$, con grados distintos. En cuanto el cuerpo es finito hay más polinomios que funciones, y toda la teoría —grado, división, irreducibilidad— se refiere a los polinomios, no a las funciones que inducen.

Proposición 25.31. *Si R es un dominio de integridad, entonces para $f, g \neq 0$:*

$$\deg(f+g) \leq \max(\deg f, \deg g), \qquad \deg(fg) = \deg f + \deg g.$$

En consecuencia $R[x]$ es un dominio de integridad y sus unidades son exactamente las unidades de R, vistas como polinomios constantes.

Demostración. La desigualdad de la suma es clara, con posible caída si los términos principales se cancelan. Para el producto, si a y b son los coeficientes principales de f y g, el coeficiente de $x^{\deg f+\deg g}$ en fg es exactamente ab, que es no nulo por ser R dominio; los grados superiores son cero. De ahí $fg \neq 0$, luego $R[x]$ es dominio. Y si $fg=1$ entonces $\deg f + \deg g = 0$, obligando a $\deg f = \deg g = 0$: ambos son constantes, y constantes inversas entre sí. □

Cuidado

La hipótesis de dominio es imprescindible. En $\mathbb{Z}_4[x]$ se tiene $(2x)(2x) = 4x^2 = 0$, de grado $-\infty$ y no 2; y $(1+2x)(1+2x) = 1+4x+4x^2 = 1$, de modo que $1+2x$ es una unidad de $\mathbb{Z}_4[x]$ que no es constante. Con divisores de cero en los coeficientes se cae todo el capítulo de polinomios, empezando por la cota "un polinomio de grado n tiene a lo sumo n raíces".

Definición 25.32 (Raíz, multiplicidad y derivada formal)**.** Sea K un cuerpo y $f \in K[x]$ no nulo. Un elemento $c \in K$ es *raíz* de f si $f(c) = 0$; su *multiplicidad* es el mayor $m \geq 1$ tal que $(x-c)^m \mid f$. La *derivada formal* de $f = \sum a_i x^i$ es $f' = \sum_{i \geq 1} i\, a_i x^{i-1}$, definida sin ningún límite: es una regla puramente algebraica que cumple $(fg)' = f'g + fg'$.

Proposición 25.33. *c es raíz múltiple de f —es decir, de multiplicidad ≥ 2— si y solo si $f(c) = f'(c) = 0$. En consecuencia, f tiene algún factor repetido si y solo si* $\gcd(f, f') \neq 1$.

Demostración. Escríbase $f = (x-c)^m g$ con $g(c) \neq 0$. Derivando con la regla del producto, $f' = m(x-c)^{m-1} g + (x-c)^m g'$. Si $m \geq 2$, ambos sumandos se anulan en c. Si $m = 1$, queda $f'(c) = g(c) \neq 0$. La segunda afirmación es la misma cuenta aplicada a un factor irreducible cualquiera en lugar de a $x - c$. □

Ejemplo 25.34 (Multiplicidades detectadas con la derivada)**.** Sea $f = x^4 - 2x^3 + 2x - 1$. Su derivada es $f' = 4x^3 - 6x^2 + 2$, y el algoritmo de Euclides da

$$\gcd(f, f') = x^2 - 2x + 1 = (x-1)^2.$$

Como el máximo común divisor no es 1, hay factores repetidos, y como es $(x-1)^2$, la raíz 1 tiene multiplicidad 3 en f. Dividiendo, $f / \gcd(f, f') = (x-1)(x+1)$, lo que confirma

$$f = (x-1)^3 (x+1),$$

igualdad que se verifica desarrollando y que se ha comprobado con álgebra computacional. ✓ El interés práctico es que este método detecta y localiza raíces múltiples *sin factorizar*: solo con divisiones sucesivas. Es lo que hace por dentro cualquier programa de cálculo simbólico antes de intentar una factorización cara.

Idea clave

Repasando lo montado: un anillo son dos operaciones con una sola ley que las conecta; los subanillos se le parecen y los ideales lo absorben; los ideales biláteros son exactamente los núcleos, y cocientar por ellos produce anillos nuevos cuya estructura se lee en la factorización del generador. A partir de aquí, todo lo que el capítulo hace con $K[x]$, con $\mathbb{Z}_n$ y con las extensiones de cuerpos es este mismo esquema aplicado una y otra vez.

25.3. Lagrange, clases laterales y subgrupos normales

Idea clave

Intuición primero. Las clases laterales son un *embaldosado*: el subgrupo H es una baldosa, y sus trasladados gH cubren el grupo entero sin traslaparse y todos del mismo tamaño: de ahí Lagrange ($|H|$ divide a $|G|$) sin ninguna cuenta. "Normal" es la condición de que el embaldosado respete la multiplicación: se puede operar con baldosas enteras sin que importe qué representante se toma de cada una. El cociente G/N es entonces el grupo visto *con lentes borrosos* que funden cada baldosa en un punto: se pierde el detalle interno de N y sobrevive la estructura entre baldosas. Toda la teoría de cocientes, aquí y en anillos y espacios vectoriales, es esta única imagen.

Definición 25.35. Las *clases laterales* de $H \leq G$ son $gH = \{gh : h \in H\}$. Dos clases o coinciden o son disjuntas; todas tienen $|H|$ elementos; el número de clases es el *índice* $[G : H]$.

Teorema 25.36 (Lagrange)**.** *Si G es finito, $|G| = [G : H] \cdot |H|$: el orden de todo subgrupo divide al del grupo.*

Idea clave

Qué dice esto en cristiano. Las clases laterales trocean el grupo en pedazos del mismo tamaño, como cortar una tarta en porciones idénticas: cada porción tiene tantos elementos como H, no se solapan, y entre todas cubren G. De ahí el teorema sale solo contando: si hay $[G : H]$ porciones y cada una tiene $|H|$ elementos, el total es el producto. La consecuencia práctica es enorme: en un grupo de 12 elementos, ningún subgrupo puede tener 5 ni 7 ni 8 elementos, sin necesidad de mirar el grupo. Lagrange no construye nada, prohíbe.

Ejemplo 25.37 (Contar antes de mirar)**. (a) Las porciones, explícitas.** En $\mathbb{Z}_{12}$ con $H = \{0, 4, 8\}$ las clases son $H = \{0, 4, 8\}$, $1 + H = \{1, 5, 9\}$, $2 + H = \{2, 6, 10\}$ y $3 + H = \{3, 7, 11\}$. Son cuatro porciones de tres elementos: $4 \cdot 3 = 12$. Y $4 + H$ no es una quinta porción, sino otra vez H.

(b) Lagrange prohibiendo. S_3 tiene seis elementos, luego sus subgrupos solo pueden tener orden $1, 2, 3$ o 6; en efecto esos cuatro tamaños se dan y ningún otro. Y como los órdenes de los elementos de S_3 son $1, 2, 2, 2, 3, 3$, todos dividen a 6, tal como exige el corolario.

(c) El pequeño teorema de Fermat, gratis. Aplicando el corolario $a^{|G|} = e$ al grupo $\mathbb{F}_7^*$, que tiene seis elementos, sale $a^6 \equiv 1 \pmod 7$ para todo a no divisible por 7: por ejemplo $2^6 = 64 = 63 + 1 \equiv 1$. En $\mathbb{F}_{11}^*$, con diez elementos, $3^{10} \equiv 1 \pmod{11}$. El teorema más famoso de la aritmética elemental resulta ser un caso particular de contar porciones de tarta.

(d) Grupos de orden primo. Si $|G| = p$ con p primo, los únicos divisores son 1 y p, luego los únicos subgrupos son el trivial y el total. Tomando cualquier $a \neq e$, el subgrupo $\langle a \rangle$ no puede ser trivial, así que es todo G: *todo grupo de orden primo es cíclico*, y por tanto isomorfo a $\mathbb{Z}_p$. Hay un único grupo de orden 5, uno de orden 7 y uno de orden 13, y esto se

ha demostrado sin construir ninguno.

Cuidado

[Contraejemplo: el recíproco de Lagrange es falso] Lagrange dice que el orden de un subgrupo divide al del grupo. Lo que *no* dice, y mucha gente asume, es lo contrario: que por cada divisor haya un subgrupo. El contraejemplo estándar es A_4, de orden 12: el número 6 divide a 12 y sin embargo A_4 no tiene ningún subgrupo de seis elementos. La razón es que un tal subgrupo tendría índice 2, luego sería normal, y su cociente tendría orden 2; pero los ocho elementos de orden 3 de A_4 tendrían que ir a parar al neutro de ese cociente, y ocho es demasiado para un subgrupo de seis. Lo que sí es cierto es el recíproco parcial de Cauchy: si un primo p divide a $|G|$, existe un elemento de orden p.

Corolario 25.38. *El orden de todo elemento divide a $|G|$; en consecuencia $a^{|G|} = e$ para todo a. Todo grupo de orden primo es cíclico (generado por cualquier elemento no trivial). El pequeño teorema de Fermat, $a^{p-1} \equiv 1 \pmod p$, es Lagrange aplicado a $\mathbb{Z}_p^*$.*

Definición 25.39 (Subgrupo normal y cociente)**.** $N \leq G$ es *normal* ($N \trianglelefteq G$) si $gNg^{-1} = N$ para todo g; equivalentemente, si las clases izquierdas y derechas coinciden. Solo entonces la operación $(aN)(bN) = abN$ está bien definida y el conjunto de clases forma el *grupo cociente* G/N, de orden $[G : N]$.

En un grupo abeliano todo subgrupo es normal. Criterio útil: todo subgrupo de índice 2 es normal (las dos clases son H y su complemento, por ambos lados). Así, $A_n \trianglelefteq S_n$.

Regla: clasificar los grupos de un orden dado

Dado n, ¿cuántos grupos de orden n hay, salvo isomorfismo? La respuesta se ataca casi siempre con la misma batería, en este orden.

1. **Factoriza n y aplica Lagrange.** El orden de todo subgrupo y de todo elemento divide a n. Si $n = p$ es primo, cualquier elemento distinto del neutro genera todo: *hay un único grupo de orden p*, el cíclico $\mathbb{Z}/p$. Este caso ya está cerrado y no hay más que decir.

2. **Cuenta los p-Sylow.** Para cada primo p con $p^a \| n$, escribe $n = p^a m$ y recuerda que el número n_p de p-subgrupos de Sylow cumple

$$n_p \equiv 1 \pmod p \qquad \text{y} \qquad n_p \mid m.$$

   Lista los divisores de m y tacha los que no son $\equiv 1$ módulo p. Con suerte solo queda $n_p = 1$.

3. **Traduce $n_p = 1$ como normalidad.** Si solo hay un p-Sylow, es normal (los conjugados de un Sylow son Sylow). Esta es la fábrica de subgrupos normales, y por tanto la fábrica de demostraciones de *no simplicidad*.

4. **Si dos Sylow son normales y se cortan trivialmente, el grupo es su producto.** Si $P \trianglelefteq G$, $Q \trianglelefteq G$, $P \cap Q = \{e\}$ y $|P||Q| = |G|$, entonces $G \cong P \times Q$. Con esto muchos órdenes quedan resueltos de golpe.

5. **Usa los teoremas de estructura de los p-grupos.** Todo grupo de orden p^2 es abeliano, luego $\mathbb{Z}/p^2$ o $\mathbb{Z}/p \times \mathbb{Z}/p$. Todo p-grupo tiene centro no trivial. Estos dos hechos resuelven la parte local.

6. **Cuando queden varios candidatos, distínguelos por invariantes.** Abeliano o no; número de elementos de cada orden; orden del centro; si es cíclico. Dos grupos con distinto perfil de órdenes no pueden ser isomorfos.

7. **Control.** La suma de los cuadrados de los grados de las representaciones, o más elementalmente: los elementos de orden p vienen en paquetes de $p-1$ por cada subgrupo de orden p. Si tu conteo no cuadra con $|G|$, hay un error.

Ejemplo 25.40 (Tres órdenes resueltos: 15, 12 y el aviso de A_4)**. Orden 15: hay exactamente un grupo.** Sea $|G| = 15 = 3 \cdot 5$. Para $p = 5$: $n_5 \mid 3$ y $n_5 \equiv 1 \pmod 5$, luego $n_5 \in \{1, 3\}$ y $3 \not\equiv 1$, así que $n_5 = 1$. Para $p = 3$: $n_3 \mid 5$ y $n_3 \equiv 1 \pmod 3$, luego $n_3 \in \{1, 5\}$ y $5 \equiv 2$, así que $n_3 = 1$. Ambos Sylow son normales, se cortan trivialmente (órdenes coprimos) y $3 \cdot 5 = 15$, luego

$$G \cong \mathbb{Z}/3 \times \mathbb{Z}/5 \cong \mathbb{Z}/15.$$

Todo grupo de orden 15 es cíclico. El mismo argumento funciona siempre que $n = pq$ con $p < q$ y $p \nmid q - 1$.

Orden 12: hay exactamente cinco. Ahora $12 = 2^2 \cdot 3$. Para $p = 3$: $n_3 \mid 4$ y $n_3 \equiv 1 \pmod 3$, luego $n_3 \in \{1, 4\}$: *no* se decide, y esa indecisión es justo la razón de que haya varios grupos. Si $n_3 = 4$, esos cuatro subgrupos de orden 3 aportan $4 \cdot 2 = 8$ elementos de orden 3; quedan $12 - 8 = 4$ elementos, que forman forzosamente el único 2-Sylow, luego $n_2 = 1$. Ese es el caso de A_4. Si $n_3 = 1$, el 3-Sylow es normal y se obtienen los otros cuatro. La lista completa es

$$\mathbb{Z}/12, \qquad \mathbb{Z}/2 \times \mathbb{Z}/6, \qquad D_6, \qquad A_4, \qquad \mathrm{Dic}_3 = Q_{12}.$$

Los dos primeros son abelianos y se distinguen porque uno tiene elemento de orden 12 y el otro no; los tres últimos no son abelianos y se distinguen por el número de elementos de orden 2: D_6 tiene siete, A_4 tiene tres y Dic_3 tiene uno solo.

Cuidado

El recíproco del teorema de Lagrange es **falso**. Que $d \mid |G|$ no garantiza que exista un subgrupo de orden d. El contraejemplo mínimo es A_4, de orden 12: no tiene ningún subgrupo de orden 6. La razón es corta y merece leerse. Si $H \leq A_4$ tuviera orden 6, su índice sería 2, luego H sería normal y $A_4/H \cong \mathbb{Z}/2$; entonces el cuadrado de todo elemento estaría en H. Pero en A_4 hay ocho 3-ciclos, y el cuadrado de un 3-ciclo es otro 3-ciclo, de modo que H contendría los ocho: imposible en un subgrupo de seis elementos. Lo que sí es cierto en general es el recíproco para potencias de primos, y eso

es exactamente el primer teorema de Sylow.

25.4. Homomorfismos y teoremas de isomorfía

Idea clave

Intuición primero. Un homomorfismo es una *traducción que respeta la gramática*: da igual operar primero y traducir después, o traducir primero y operar después. Como toda traducción imperfecta, pierde información: el *núcleo* es exactamente "lo que la traducción decidió olvidar", el conjunto de elementos que acaban confundidos con el neutro. El primer teorema de isomorfía dice algo casi tautológico una vez visto así: *si borras lo que se olvida, lo que queda es lo que se ve*, $G/\ker\varphi \cong \operatorname{im}\varphi$. El logaritmo es el ejemplo cotidiano: traduce $(\mathbb{R}^{>0},\cdot)$ a $(\mathbb{R},+)$ sin perder nada, y por eso multiplicar se convierte en sumar. El determinante traduce matrices a números y olvida todo salvo el factor de área; lo olvidado es SL_n.

Definición 25.41. $\varphi : G \to G'$ es *homomorfismo* si $\varphi(ab) = \varphi(a)\varphi(b)$; entonces $\varphi(e) = e'$ y $\varphi(a^{-1}) = \varphi(a)^{-1}$. Su núcleo $\ker\varphi = \varphi^{-1}(e')$ es un subgrupo normal, y φ es inyectivo $\iff \ker\varphi = \{e\}$. Un *isomorfismo* es un homomorfismo biyectivo: $G \cong G'$ significa "el mismo grupo con otros nombres".

Teorema 25.42 (Primer teorema de isomorfía)**.** $G/\ker\varphi \cong \operatorname{im}\varphi$, *mediante* $g\ker\varphi \mapsto \varphi(g)$.

Ejemplo 25.43. $\det : GL_n(\mathbb{R}) \to \mathbb{R}^*$ es homomorfismo con núcleo SL_n: entonces $GL_n/SL_n \cong \mathbb{R}^*$. El signo $\operatorname{sgn} : S_n \to \{\pm 1\}$ tiene núcleo A_n: $S_n/A_n \cong \mathbb{Z}_2$. La exponencial $t \mapsto e^{it}$ de $(\mathbb{R},+)$ al círculo unitario tiene núcleo $2\pi\mathbb{Z}$: el círculo es $\mathbb{R}/2\pi\mathbb{Z}$, "la recta enrollada".

Regla: calcular un grupo cociente G/N y reconocerlo

Un cociente no es un objeto nuevo y misterioso: es el grupo original visto "con gafas borrosas", donde dos elementos que difieren por algo de N pasan a ser el mismo.

1. **Comprueba primero que N es normal.** Sin $gNg^{-1} = N$ el cociente ni siquiera existe como grupo. Atajos: todo subgrupo de un grupo abeliano es normal; todo subgrupo de índice 2 es normal; el núcleo de cualquier homomorfismo es normal; el centro $Z(G)$ es normal.
2. **Calcula el orden del cociente:** $|G/N| = |G|/|N|$ (Lagrange). Este número solo ya suele decidir casi todo, porque hay muy pocos grupos de orden pequeño.
3. **Escribe las clases.** Los elementos de G/N son los conjuntos gN. Dos representantes g y h dan la misma clase exactamente cuando $g^{-1}h \in N$.
4. **Multiplica con representantes:** $(gN)(hN) = (gh)N$. La normalidad es justo lo que hace que este resultado no dependa del representante elegido.

5. **Calcula el orden de una clase.** El orden de gN en G/N es el menor $n \geq 1$ con $g^n \in N$. Ojo: puede ser mucho menor que el orden de g en G.
6. **Reconoce el grupo.** Si encuentras una clase de orden $|G/N|$, el cociente es cíclico $\mathbb{Z}/|G/N|$. Si no, y el orden es 4, es el grupo de Klein. Si el orden es p primo, es forzosamente $\mathbb{Z}/p$.
7. **Control:** la proyección $\pi : G \to G/N$, $\pi(g) = gN$, es un homomorfismo sobreyectivo con núcleo exactamente N. Si tu cociente no cumple esto, algo se torció.

Ejemplo 25.44 (Cuatro cocientes calculados a mano)**. (a)** $\mathbb{Z}/12$ **módulo** $N = \langle 4 \rangle$**.** Aquí $N = \{0, 4, 8\}$, de orden 3, y como el grupo es abeliano N es normal. El cociente tiene orden $12/3 = 4$ y sus clases son

$$\bar{0} = \{0, 4, 8\}, \quad \bar{1} = \{1, 5, 9\}, \quad \bar{2} = \{2, 6, 10\}, \quad \bar{3} = \{3, 7, 11\}.$$

La clase $\bar{1}$ tiene orden 4, porque $1, 2, 3 \notin N$ pero $4 \in N$. Al haber un elemento de orden igual al del grupo, el cociente es cíclico:

$$(\mathbb{Z}/12)/\langle 4 \rangle \;\cong\; \mathbb{Z}/4.$$

(b) D_4 **módulo su centro.** En el grupo diédrico $D_4 = \{1, r, r^2, r^3, s, rs, r^2s, r^3s\}$ de orden 8, el centro es $Z(D_4) = \{1, r^2\}$, de orden 2. El cociente tiene orden 4. Sus clases son

$$\bar{1} = \{1, r^2\}, \quad \bar{r} = \{r, r^3\}, \quad \bar{s} = \{s, r^2s\}, \quad \overline{rs} = \{rs, r^3s\}.$$

Ahora bien: $r^2 \in Z$, luego $\bar{r}^{\,2} = \bar{1}$; y $s^2 = 1$, luego $\bar{s}^{\,2} = \bar{1}$. *Toda* clase no trivial tiene orden 2, así que no hay elemento de orden 4 y el cociente no es cíclico:

$$D_4/Z(D_4) \;\cong\; \mathbb{Z}/2 \times \mathbb{Z}/2.$$

(c) S_4 **módulo** A_4**.** El subgrupo alternado tiene índice 2, luego es normal y el cociente tiene orden 2: hay solo dos clases, las permutaciones pares y las impares. El isomorfismo $S_4/A_4 \cong \mathbb{Z}/2$ no es más que la función signo ε.

(d) S_4 **módulo el grupo de Klein** $V = \{e, (12)(34), (13)(24), (14)(23)\}$**.** Este V sí es normal en S_4 (la conjugación conserva el tipo de ciclo, y esas tres son *todas* las permutaciones de tipo $2 + 2$). El cociente tiene orden $24/4 = 6$. La clase de (12) tiene orden 2 y la de (123) tiene orden 3; además el cociente no es abeliano, porque $\overline{(12)} \cdot \overline{(123)} \neq \overline{(123)} \cdot \overline{(12)}$. El único grupo no abeliano de orden 6 es S_3:

$$S_4/V \;\cong\; S_3.$$

Este isomorfismo tiene un significado geométrico: S_4 actúa sobre los tres pares de aristas opuestas del tetraedro, y V es exactamente el conjunto de simetrías que no mueve ninguno de los tres pares.

Cuidado

El cociente *no* es un subgrupo de G. En (b), $D_4/Z \cong \mathbb{Z}/2 \times \mathbb{Z}/2$, pero el grupo de Klein no aparece como subgrupo de D_4 conteniendo a Z de manera que "complete" el grupo: D_4 no es isomorfo a $Z \times (D_4/Z) = \mathbb{Z}/2 \times \mathbb{Z}/2 \times \mathbb{Z}/2$, que es abeliano. Conocer N y G/N no determina G.

Pausa para pensar

En (b) todo elemento del cociente cumple $\bar{g}^{\,2} = \bar{1}$, es decir $g^2 \in Z(G)$ para todo $g \in D_4$. Comprueba esta afirmación directamente en la lista de los ocho elementos: los cuatro giros tienen $g^2 \in \{1, r^2\}$ y las cuatro reflexiones tienen $g^2 = 1$. Que $G/Z(G)$ sea cíclico obligaría a G a ser abeliano; aquí no lo es, y en efecto el cociente resultó ser Klein y no $\mathbb{Z}/4$.

25.5. Presentaciones: describir un grupo con generadores y relaciones

Hasta aquí los grupos han llegado ya construidos: permutaciones, matrices, clases de restos. Pero muchas veces lo que se conoce de un grupo no es una lista de elementos sino un puñado de *movimientos básicos* y las *ecuaciones* que satisfacen. Girar un cubo de Rubik, componer simetrías de un polígono, trenzar hilos: en todos esos casos uno sabe qué hacen los generadores y qué combinaciones vuelven al punto de partida. Una presentación convierte esa información en una definición completa del grupo.

Definición 25.45 (Grupo libre y presentación)**.** El *grupo libre* $F(S)$ sobre un conjunto $S = \{a, b, \dots\}$ tiene por elementos las *palabras reducidas* en las letras de S y sus inversas —cadenas finitas en las que nunca aparecen juntas x y x^{-1}— con la concatenación seguida de reducción como operación, y la palabra vacía como neutro. Dado un subconjunto $R \subseteq F(S)$ de *relatores*, la *presentación* $\langle S \mid R \rangle$ designa el cociente

$$\langle S \mid R \rangle \;=\; F(S)/N(R),$$

donde $N(R)$ es el menor subgrupo normal que contiene a R, llamado *clausura normal* de R. Escribir una relación $w = v$ equivale a incluir el relator wv^{-1}.

Idea clave

Qué dice esto en cristiano. Una presentación es la descripción más económica posible de un grupo: se dan unas letras y unas reglas, y el grupo es todo lo que se puede escribir con esas letras respetando esas reglas y nada más. El grupo libre es el caso sin reglas, donde no se cancela nunca nada salvo lo obligatorio ($aa^{-1} = 1$); es el grupo más grande que esas letras pueden generar. Al imponer relaciones se van pegando elementos que antes eran distintos, y el grupo se encoge. Es exactamente como definir un objeto por sus propiedades en vez de listarlo.

Ejemplo 25.46 (Presentaciones que ya conocemos)**. (a) Los cíclicos.** La presentación $\langle a \mid a^n = 1\rangle$ describe a $\mathbb{Z}_n$: la única regla es que dar n pasos te devuelve al principio. Sin ninguna regla, $\langle a \mid \rangle$ es $\mathbb{Z}$, infinito. Y $\langle a, b \mid ab = ba\rangle$ es $\mathbb{Z} \times \mathbb{Z}$: dos direcciones independientes que conmutan.

(b) D_4 en una línea. La presentación $\langle a, b \mid a^4 = 1,\ b^2 = 1,\ bab^{-1} = a^{-1}\rangle$ describe el grupo de simetrías del cuadrado: a es el giro de 90°, b una reflexión, y la tercera relación dice que reflejar invierte el sentido del giro. Toda palabra se reduce a la forma $a^i b^j$ con $0 \le i < 4$ y $0 \le j < 2$, es decir ocho elementos, que son exactamente los de D_4.

(c) Cuánto crece el grupo libre. En $F(\{a, b\})$, las palabras reducidas de longitud 1 son cuatro (a, a^{-1}, b, b^{-1}); las de longitud 2 son doce, porque tras cada letra hay tres continuaciones posibles (todas salvo su propia inversa); las de longitud 3 son 36, y en general $4 \cdot 3^{n-1}$. El crecimiento es exponencial, y ese es el sentido preciso de que el grupo libre sea "lo más grande posible".

(d) Una presentación sorprendente. El grupo $\langle a, b \mid a^2 = b^3 = (ab)^3 = 1\rangle$ tiene exactamente doce elementos y es A_4. Ni el número 12 ni la identificación son evidentes mirando la presentación, y eso ilustra a la vez su potencia y su peligro: describe grupos muy complicados con muy pocos símbolos, pero averiguar cuál es puede ser durísimo.

Cuidado

[Añadir relaciones nunca agranda un grupo, y a veces lo aniquila] Cada relación nueva identifica elementos, así que el grupo solo puede encogerse. Es fácil pasarse: $\langle a \mid a = 1\rangle$ es el grupo trivial, y hay presentaciones de aspecto inofensivo que resultan ser triviales sin que se note. De hecho el *problema de la palabra* (decidir si dos palabras representan el mismo elemento) es indecidible en general: Novikov y Boone demostraron en los años cincuenta que no existe ningún algoritmo que funcione para toda presentación. Ni siquiera la pregunta "¿es este grupo trivial?" tiene respuesta algorítmica. Una notación tan compacta esconde una dificultad ilimitada.

Idea clave

El grupo libre es el grupo que *no obedece ninguna regla salvo las obligatorias*: las únicas simplificaciones permitidas son cancelar xx^{-1}, porque eso lo exigen los axiomas. Todo lo demás queda distinto de todo lo demás. Después, imponer relaciones es ir pegando: cada relator declara que cierta palabra es el neutro, y ese decreto arrastra consigo todos sus conjugados y productos — de ahí que haya que tomar la clausura *normal* y no simplemente el subgrupo generado. El resultado es que $\langle S \mid R\rangle$ es el grupo más grande que satisface las relaciones pedidas: cualquier otro grupo con generadores que las cumplan es un cociente suyo. Esa es la propiedad universal, y es la que hace útiles a las presentaciones.

Regla: leer una presentación y averiguar qué grupo es

Una presentación no dice de entrada ni el orden del grupo ni si es trivial; hay que trabajarla.

1. **Búsquese una forma normal para las palabras.** Úsense las relaciones como reglas de reescritura orientadas: si $sr = r^{-1}s$, toda palabra puede llevarse a la forma $r^i s^j$ moviendo las s a la derecha. La forma normal es el objetivo, porque contar formas normales es contar elementos.

2. **Acótese el orden por arriba contando formas normales.** Si toda palabra se reduce a $r^i s^j$ con $0 \le i < n$ y $0 \le j < 2$, entonces $|G| \le 2n$. Nótese que esto es solo una cota: las relaciones podrían colapsar más de lo que se ve.

3. **Acótese por abajo exhibiendo un grupo concreto que cumpla las relaciones.** Si se encuentra un grupo H de orden $2n$ generado por dos elementos que satisfacen exactamente esos relatores, la propiedad universal da un epimorfismo $G \twoheadrightarrow H$, luego $|G| \ge 2n$. Con la cota anterior se concluye $G \cong H$.

4. **Abelianícese para obtener información barata.** El cociente $G/[G,G]$ se calcula tratando los generadores como conmutativos: las relaciones pasan a ser un sistema lineal sobre $\mathbb{Z}$, resoluble con la forma normal de Smith. Un abelianizado no trivial demuestra que G no es trivial.

5. **Búsquense homomorfismos hacia grupos pequeños.** Enviar los generadores a permutaciones concretas de S_3, S_4 o a matrices, y comprobar que las relaciones se cumplen, produce cocientes conocidos y por tanto cotas inferiores del orden.

6. **Sospéchese de las presentaciones que colapsan.** Una relación aparentemente inofensiva puede matar el grupo entero: conviene siempre intentar deducir $a = 1$ manipulando los relatores antes de dar por buena la intuición.

7. **Acéptese que en general no hay algoritmo.** Decidir si dos palabras representan el mismo elemento — el *problema de la palabra* — es indecidible para presentaciones finitas arbitrarias (Novikov–Boone), y también lo es decidir si el grupo es trivial. Las herramientas de los pasos anteriores funcionan en los casos concretos, no en general.

Ejemplo 25.47 (Cuatro presentaciones desmontadas)**. (a) El cíclico y el diédrico.** La presentación $\langle a \mid a^n \rangle$ da $\mathbb{Z}/n\mathbb{Z}$: toda palabra es a^i y la relación identifica a^n con 1, luego hay exactamente n formas normales. Para el diédrico, $D_n = \langle r, s \mid r^n,\ s^2,\ srsr \rangle$. La tercera relación se lee mejor como $srs^{-1} = r^{-1}$, y es la que permite mover toda s a la derecha: cada vez que una s se cruza con una r, la r cambia de signo. Así toda palabra se reduce a $r^i s^j$ con $0 \le i < n$, $0 \le j < 2$, de donde $|D_n| \le 2n$; y como las simetrías del n-gono regular forman un grupo de orden $2n$ que cumple las tres relaciones, se tiene la igualdad.

(b) Los cuaterniones. Sea $Q_8 = \langle i, j \mid i^4,\ i^2 j^{-2},\ jij^{-1}i \rangle$. Las relaciones dicen $i^4 = 1$, $j^2 = i^2$ y $jij^{-1} = i^{-1}$. Póngase $-1 := i^2$; entonces $(-1)^2 = i^4 = 1$, y -1 es central. La forma

normal es $i^a j^b$ con $a < 4$, $b < 2$: ocho palabras. Que el orden es exactamente 8 se comprueba con la representación matricial

$$i \mapsto \begin{pmatrix} \mathrm{i} & 0 \\ 0 & -\mathrm{i} \end{pmatrix}, \qquad j \mapsto \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix},$$

donde i es la unidad imaginaria: se verifica $i^2 = j^2 = -I$ y $ij = -ji$, y la imagen tiene ocho matrices distintas. Nótese la diferencia con D_4: en Q_8 hay un solo elemento de orden 2 (a saber, -1), mientras que D_4 tiene cinco.

(c) Dos presentaciones con dos letras y ninguna relación seria. El grupo libre $F_2 = \langle a, b \mid \rangle$ es infinito y muy grande: el número de palabras reducidas de longitud exactamente n es $4 \cdot 3^{n-1}$ (cuatro elecciones para la primera letra, tres para cada siguiente, porque solo se prohíbe deshacer la anterior), es decir $4, 12, 36, 108, 324, \ldots$ Añadiendo el único relator $aba^{-1}b^{-1}$ se obtiene $\mathbb{Z}^2$, donde ahora hay una sola palabra de cada forma $a^m b^n$. Añadiendo en cambio a^2 y b^2 se obtiene el *diédrico infinito*: ab tiene orden infinito y el grupo es una copia de $\mathbb{Z} \rtimes \mathbb{Z}/2\mathbb{Z}$.

(d) Una presentación de A_4. Considérese $\langle a, b \mid a^2,\ b^3,\ (ab)^3 \rangle$. Enviando $a \mapsto (12)(34)$ y $b \mapsto (123)$ se comprueba que las tres relaciones se cumplen, y las imágenes generan A_4, de orden 12. Un análisis de formas normales da la cota superior 12, luego el grupo presentado es exactamente A_4. La familia $\langle a, b \mid a^2, b^3, (ab)^n \rangle$ da los grupos de triángulo: $n = 3$ produce A_4, $n = 4$ produce S_4, $n = 5$ produce A_5, y a partir de $n = 6$ el grupo se vuelve infinito, reflejando el paso de la geometría esférica a la euclídea y a la hiperbólica.

Ejemplo 25.48 (Presentaciones que engañan)**. (a) Un grupo trivial disfrazado.** La presentación $\langle a, b \mid aba^{-1}b^{-2},\ bab^{-1}a^{-2} \rangle$ parece definir un grupo infinito y no abeliano. En realidad es el grupo trivial: las relaciones dicen $aba^{-1} = b^2$ y $bab^{-1} = a^2$, y combinándolas se llega a $a = b = 1$. Ninguna inspección superficial lo delata; hay que manipular.

(b) El mismo grupo con dos caras. Las presentaciones $\langle a, b \mid a^2, b^2, (ab)^3 \rangle$ y $\langle x, y \mid x^3, y^2, yxyx \rangle$ describen ambas a S_3, pese a no parecerse en nada. La traducción es $x = ab$, $y = a$. Moraleja: la presentación no es un invariante del grupo, es solo una descripción entre muchas, y decidir si dos presentaciones dan grupos isomorfos es, de nuevo, indecidible en general.

(c) Menos relaciones no significa grupo mayor. Quitar un relator siempre agranda el grupo o lo deja igual, eso es cierto; pero añadir un *generador* junto con una relación que lo define no cambia nada. Así $\langle a, b \mid a^n,\ b = a \rangle \cong \mathbb{Z}/n\mathbb{Z}$, con dos generadores y dos relaciones para un grupo cíclico. Esta operación, llamada *transformación de Tietze*, junto con su inversa, genera todas las presentaciones equivalentes de un mismo grupo.

(d) El abelianizado como detector. Para $G = \langle a, b \mid a^2 b^{-3},\ b^3 a^{-1} b^{-1} a \rangle$, abelianizar convierte las relaciones en el sistema $2a - 3b = 0$, $3b - b = 0$ sobre $\mathbb{Z}$, es decir $2a = 3b$ y $2b = 0$. De ahí $4a = 6b = 0$ y el abelianizado es finito no trivial, luego $G \neq 1$. Sin resolver el problema de la palabra, y con solo álgebra lineal sobre $\mathbb{Z}$, se ha demostrado algo sustancial sobre G.

Cuidado

Una presentación finita no garantiza un grupo finito, ni siquiera un grupo del que se pueda decidir nada. $\langle a, b \mid \ \rangle$ tiene dos generadores, cero relaciones y es infinito; $\langle a, b \mid a^2, b^2 \rangle$ tiene dos relaciones y sigue siendo infinito. Y el error más frecuente en la práctica es olvidar que la clausura normal contiene a los *conjugados* de los relatores: imponer $r^n = 1$ obliga también a $wr^n w^{-1} = 1$ para toda palabra w, lo cual puede colapsar el grupo mucho más de lo previsto.

Pausa para pensar

Toda presentación tiene una lectura geométrica: constrúyase un grafo cuyos vértices sean los elementos del grupo y cuyas aristas etiquetadas conecten g con gs para cada generador s. Es el *grafo de Cayley*, y en él los relatores son exactamente los ciclos cerrados. Para $\mathbb{Z}^2$ sale la cuadrícula infinita; para D_4, un prisma octogonal; para F_2, un árbol infinito de valencia 4 sin ningún ciclo — que es la manera visual de decir que no hay relaciones. La teoría geométrica de grupos, disciplina entera del siglo XX, consiste en estudiar el grupo mirando la forma de ese grafo a gran escala.

25.6. Acciones de grupos

Idea clave

Intuición primero. Un grupo abstracto es una lista de simetrías desempleadas; una *acción* las pone a trabajar sobre un conjunto concreto. La órbita de un punto es "todos los lugares a los que puedes viajar" usando las simetrías; el estabilizador es "cuántas simetrías te dejan quieto". El teorema órbita–estabilizador dice que viaje y lealtad se compensan exactamente: $|\text{órbita}| \times |\text{estabilizador}| = |G|$ (los elementos que te llevan al mismo sitio forman una clase lateral del estabilizador: ¡el embaldosado otra vez!). Casi todos los teoremas de conteo del capítulo —ecuación de clases, Sylow, Burnside y los collares— son esta igualdad disfrazada.

Definición 25.49. Una *acción* de G sobre un conjunto X asigna a cada $g \in G$ una biyección de X, compatiblemente: $(gh) \cdot x = g \cdot (h \cdot x)$, $e \cdot x = x$. La *órbita* de x es $Gx = \{g \cdot x\}$; el *estabilizador* es $G_x = \{g : g \cdot x = x\} \leq G$.

Teorema 25.50 (Órbita–estabilizador)**.** $|Gx| = [G : G_x]$*; en particular, el tamaño de toda órbita divide a* $|G|$.

Ejemplo 25.51. El grupo de rotaciones del cubo actúa sobre las 6 caras: la acción es transitiva (una órbita) y el estabilizador de una cara son sus 4 rotaciones: $|G| = 6 \cdot 4 = 24$. Sin enumerar una sola rotación, sabemos que hay exactamente 24. Las acciones cuentan lo incontable; el mismo principio (con el lema de Burnside) cuenta collares, coloraciones y moléculas.

Dos consecuencias estructurales profundas:

Teorema 25.52 (Cayley)**.** *Todo grupo es isomorfo a un subgrupo de un grupo de permutaciones (G actúa sobre sí mismo por multiplicación). Los grupos abstractos* son *grupos de simetrías.*

Teorema 25.53 (Cauchy y Sylow, enunciados)**.** *Si un primo p divide a $|G|$, existe un elemento de orden p (Cauchy). Más fino (Sylow): si p^k es la mayor potencia de p que divide a $|G|$, existen subgrupos de orden p^k, todos conjugados entre sí, y su número es $\equiv 1 \pmod p$ y divide a $|G|$. Estos teoremas son el punto de partida de la clasificación de grupos finitos.*

Regla: aplicar los teoremas de Sylow para clasificar grupos

Los teoremas de Sylow son la herramienta más potente para deducir la estructura de un grupo finito a partir únicamente de su orden.

1. **Factoriza el orden.** Escribe $|G| = p^k m$ con $p \nmid m$, para cada primo p que divida a $|G|$. Un p-subgrupo de Sylow es un subgrupo de orden exactamente p^k.
2. **Recuerda los tres enunciados.** Primero: existe al menos un p-Sylow. Segundo: todos los p-Sylow son conjugados entre sí. Tercero: su número n_p cumple
$$n_p \equiv 1 \pmod p \qquad \text{y} \qquad n_p \mid m.$$
3. **Lista los candidatos.** Para cada p, escribe los divisores de m y quédate con los congruentes con 1 módulo p. Suele quedar una lista muy corta, y a veces solo el 1.
4. **Traduce $n_p = 1$ a normalidad.** Si solo hay un p-Sylow, es invariante por conjugación, luego *normal*. Esta es la conexión clave: contar Sylow es buscar subgrupos normales, y por tanto decidir si el grupo es simple.
5. **Cuenta elementos si hay varias posibilidades.** Dos p-Sylow distintos con p primo se cortan solo en la identidad cuando $k = 1$, así que n_p subgrupos aportan $n_p(p-1)$ elementos de orden p. Si al sumar sobre varios primos te pasas de $|G|$, has descartado esa combinación.
6. **Combina para obtener la estructura.** Si dos Sylow P y Q son ambos normales, se cortan trivialmente y sus órdenes multiplican al de G, entonces $G \cong P \times Q$. Si solo uno es normal, se obtiene un producto semidirecto $P \rtimes Q$.
7. **Contrasta con la lista conocida.** Verifica que el grupo obtenido tiene el orden correcto y que sus elementos tienen los órdenes esperados. Recuerda que $|G| = p$ primo obliga a $G \cong \mathbb{Z}/p$, y que todo grupo de orden p^2 es abeliano.

Ejemplo 25.54 (Cuatro órdenes analizados con Sylow)**. (a) Orden 15: solo hay un grupo.** Con $|G| = 15 = 3 \cdot 5$. Para $p = 3$: $m = 5$, y los divisores de 5 congruentes con 1 módulo 3 son solo $\{1\}$, porque $5 \equiv 2$. Luego $n_3 = 1$. Para $p = 5$: $m = 3$, y los divisores $\{1, 3\}$ con $3 \equiv 3 \not\equiv 1$ módulo 5, así que $n_5 = 1$. Ambos Sylow son normales, se cortan trivialmente por

tener órdenes coprimos, y $3 \cdot 5 = 15$. Por tanto

$$G \cong \mathbb{Z}/3 \times \mathbb{Z}/5 \cong \mathbb{Z}/15.$$

Conclusión: todo grupo de orden 15 es cíclico. El mismo argumento funciona siempre que $|G| = pq$ con $p < q$ y $p \nmid q - 1$.

(b) Orden 30: no hay grupos simples. Aquí $30 = 2 \cdot 3 \cdot 5$. Para $p = 5$: $m = 6$, candidatos $\{1, 6\}$. Para $p = 3$: $m = 10$, candidatos $\{1, 10\}$. Supongamos que ninguno vale 1, es decir $n_5 = 6$ y $n_3 = 10$. Entonces los seis 5-Sylow aportan $6 \cdot (5 - 1) = 24$ elementos de orden 5, y los diez 3-Sylow aportan $10 \cdot (3 - 1) = 20$ elementos de orden 3. En total $24 + 20 = 44$ elementos distintos, más la identidad. Pero $|G| = 30 < 45$: contradicción. Luego $n_3 = 1$ o $n_5 = 1$, y en cualquier caso G tiene un subgrupo normal propio: *no hay grupos simples de orden* 30.

(c) Orden 56: el mismo truco, con desenlace distinto. Con $56 = 2^3 \cdot 7$. Para $p = 7$: $m = 8$, candidatos $\{1, 8\}$. Si $n_7 = 8$, los ocho 7-Sylow son cíclicos de orden 7 y se cortan trivialmente, aportando $8 \cdot 6 = 48$ elementos de orden 7. Quedan $56 - 48 = 8$ elementos, que deben albergar a todos los 2-Sylow, cada uno de orden 8. Pero eso solo cabe si hay exactamente uno, luego $n_2 = 1$ y el 2-Sylow es normal. Y si en cambio $n_7 = 1$, el 7-Sylow es normal. En ambas ramas hay un subgrupo normal propio: tampoco hay grupos simples de orden 56.

(d) Orden 255: cíclico obligatorio. Con $255 = 3 \cdot 5 \cdot 17$. Para $p = 17$: $m = 15$, y de sus divisores $\{1, 3, 5, 15\}$ ninguno salvo 1 es congruente con 1 módulo 17. Luego $n_{17} = 1$ y el 17-Sylow P es normal. Considerando el cociente G/P de orden 15, el apartado (a) dice que es cíclico. A partir de ahí se demuestra que G mismo es cíclico, es decir $G \cong \mathbb{Z}/255$. Con solo contar divisores hemos clasificado por completo un orden de tres cifras.

Cuidado

La condición $n_p \equiv 1$ módulo p y $n_p \mid m$ es necesaria, no suficiente: que un valor sea candidato no significa que se realice. En el orden 12, por ejemplo, los candidatos son $n_2 \in \{1, 3\}$ y $n_3 \in \{1, 4\}$, y de hecho existen grupos que realizan varias combinaciones: $\mathbb{Z}/12$ tiene $n_2 = n_3 = 1$, mientras que A_4 tiene $n_3 = 4$. Sylow acota las posibilidades; no las determina.

Idea clave

La búsqueda de grupos simples finitos ocupó a la comunidad matemática durante más de un siglo y culminó en 2004 con una demostración repartida en unas diez mil páginas y cientos de artículos. Los argumentos de conteo de este ejemplo son la versión elemental de esa maquinaria: descartar órdenes uno a uno hasta que solo quedan las familias infinitas conocidas (cíclicos de orden primo, alternados A_n con $n \geq 5$, grupos de tipo Lie) y los veintiséis grupos esporádicos, entre ellos el Monstruo, de orden aproximadamente 8×10^{53}.

25.7. El producto semidirecto y la clasificación de los grupos de orden pq

El producto directo $A \times B$ es la manera aburrida de juntar dos grupos: cada uno vive por su lado y los elementos de uno conmutan con los del otro. Pero la mayoría de los grupos interesantes no son productos directos. S_3 contiene un $\mathbb{Z}/3\mathbb{Z}$ normal y un $\mathbb{Z}/2\mathbb{Z}$, sus órdenes multiplican a 6 y se cortan trivialmente, y sin embargo $S_3 \neq \mathbb{Z}/3\mathbb{Z} \times \mathbb{Z}/2\mathbb{Z}$, porque S_3 no es abeliano. Lo que falta es permitir que la segunda pieza *actúe* sobre la primera. Esa es exactamente la construcción del producto semidirecto.

Definición 25.55 (Producto semidirecto)**.** Sean N y H grupos y $\varphi\colon H \to \mathrm{Aut}(N)$ un homomorfismo. El *producto semidirecto* $N \rtimes_\varphi H$ es el conjunto $N \times H$ con la operación

$$(n_1, h_1)(n_2, h_2) = \big(n_1\, \varphi(h_1)(n_2),\ h_1 h_2\big).$$

Es un grupo de orden $|N|\,|H|$, con neutro $(1,1)$ e inverso $(n,h)^{-1} = (\varphi(h^{-1})(n^{-1}), h^{-1})$. Identificando N con $N \times \{1\}$ y H con $\{1\} \times H$, se tiene $N \trianglelefteq N \rtimes_\varphi H$, $N \cap H = 1$, NH el grupo entero, y la conjugación reproduce la acción: $hnh^{-1} = \varphi(h)(n)$. Si φ es trivial se recupera el producto directo.

Idea clave

Qué dice esto en cristiano. El producto directo $N \times H$ pega dos grupos sin que se enteren el uno del otro: cada componente se opera por su lado. El semidirecto pega lo mismo pero dejando que H *revuelva* a N al pasar por encima: la fórmula $(n_1, h_1)(n_2, h_2) = (n_1\, \varphi(h_1)(n_2), h_1 h_2)$ dice que, antes de multiplicar los dos trozos de N, hay que aplicar a n_2 la deformación que h_1 ordena. Si φ es trivial (nadie revuelve nada) se recupera el producto directo. Todos los grupos no abelianos pequeños son de esta forma: es la manera más barata de fabricar no conmutatividad.

Ejemplo 25.56 (Fabricar D_4 y S_3 desde cero)**. (a) La receta para D_4.** Tomemos $N = \mathbb{Z}_4$ (las rotaciones), $H = \mathbb{Z}_2$ (reflejar o no) y $\varphi(1) =$ "cambiar de signo", es decir el automorfismo $n \mapsto -n$ de $\mathbb{Z}_4$. La operación queda $(a,0)(b,\epsilon) = (a+b, \epsilon)$ y $(a,1)(b,\epsilon) = (a-b, 1+\epsilon)$. El grupo resultante tiene ocho elementos con órdenes $1, 2, 2, 2, 2, 2, 4, 4$, exactamente los de D_4.

(b) Comprobar que no conmuta. $(1,0)(0,1) = (1,1)$, mientras que $(0,1)(1,0) = (0-1, 1) = (3,1)$. Son distintos, luego el grupo no es abeliano; en el lenguaje de las simetrías del cuadrado, girar y luego reflejar no es lo mismo que reflejar y luego girar.

(c) Qué pasa si φ es trivial. Con $\varphi(1) = \mathrm{id}$ la fórmula se reduce a sumar componente a componente: sale $\mathbb{Z}_4 \times \mathbb{Z}_2$, cuyos órdenes son $1, 2, 2, 2, 4, 4, 4, 4$. Compárese con la lista de (a): son grupos distintos, ambos de orden 8, y lo único que los separa es la elección de φ.

(d) S_3 por el mismo procedimiento. Con $N = \mathbb{Z}_3$, $H = \mathbb{Z}_2$ y $\varphi(1)(n) = -n$ sale un grupo de seis elementos con órdenes $1, 2, 2, 2, 3, 3$: es S_3. Aquí N son las rotaciones del triángulo y H es reflejar. Y en general, para $p < q$ primos, existe un grupo no abeliano de orden pq si y solo si p divide a $q-1$: para $p = 3$, $q = 7$ sí $(3 \mid 6)$ y existe uno de orden 21; para $p = 3$, $q = 5$ no $(3 \nmid 4)$ y el único grupo de orden 15 es $\mathbb{Z}_{15}$.

Cuidado

[Cuidado con el orden de los factores] En $N \rtimes_\varphi H$ el factor de la izquierda, N, es el *normal*, y el de la derecha actúa sobre él; escribir $H \rtimes N$ es otra cosa y normalmente ni siquiera tiene sentido. Segundo aviso: el mismo par (N, H) puede dar grupos distintos según qué φ se elija, como muestran los apartados (a) y (c), así que *no* se puede hablar de "el" producto semidirecto de $\mathbb{Z}_4$ y $\mathbb{Z}_2$ sin decir cuál. Y tercero: no todo grupo es semidirecto de dos trozos suyos. El grupo de los cuaterniones $Q_8 = \{\pm 1, \pm \mathrm{i}, \pm \mathrm{j}, \pm \mathrm{k}\}$ tiene orden 8 y todos sus subgrupos no triviales contienen a -1, de modo que no existen N y H con $N \cap H = \{1\}$; Q_8 no se descompone.

Teorema 25.57 (Reconocimiento interno)**.** *Sea G un grupo con subgrupos $N \trianglelefteq G$ y $H \leq G$ tales que $N \cap H = \{1\}$ y $NH = G$. Entonces $G \cong N \rtimes_\varphi H$, donde $\varphi(h)$ es la conjugación por h restringida a N. Se dice que H es un* complemento *de N en G.*

Idea clave

Qué dice esto en cristiano. El producto semidirecto, tal como se definió, fabrica un grupo nuevo a partir de dos piezas sueltas. Este teorema hace el camino inverso: si te dan un grupo ya hecho y encuentras dentro de él dos subgrupos que encajan bien, puedes afirmar que el grupo *era* un semidirecto disfrazado. Las tres condiciones que hay que comprobar son sencillas de recordar: uno de los dos debe ser normal, deben cortarse solo en el neutro, y entre los dos deben generar todo. Es el criterio que se usa una y otra vez para clasificar grupos pequeños.

Ejemplo 25.58 (Reconocer D_4 y S_3 por dentro)**. (a) Las tres comprobaciones en D_4.** Tomemos $N = \langle r \rangle$, el subgrupo de las cuatro rotaciones, y $H = \langle s \rangle$ con s una reflexión. Primero, N tiene índice $8/4 = 2$, y todo subgrupo de índice 2 es normal. Segundo, $N \cap H = \{1\}$, porque la única rotación que es también reflexión sería la identidad. Tercero, $|N||H| = 4 \cdot 2 = 8 = |D_4|$, luego NH es todo el grupo.

(b) Quién es φ. La conjugación da $srs^{-1} = r^{-1}$: reflejar, girar y volver a reflejar equivale a girar al revés. Por tanto $\varphi(s)$ es el automorfismo "invertir" de $\mathbb{Z}_4$, y queda $D_4 \cong \mathbb{Z}_4 \rtimes \mathbb{Z}_2$ con ese φ concreto, exactamente el que ya habíamos usado para construirlo.

(c) S_3 con el mismo argumento. Con $N = A_3$ (de índice $6/3 = 2$, luego normal) y $H = \langle (1\,2) \rangle$, se tiene $N \cap H = \{e\}$ porque A_3 solo contiene permutaciones pares y $(1\,2)$ es impar, y $3 \cdot 2 = 6$. Conclusión inmediata: $S_3 \cong \mathbb{Z}_3 \rtimes \mathbb{Z}_2$.

(d) La utilidad real. Este criterio es el motor de la clasificación de los grupos de orden pq. Los teoremas de Sylow garantizan que existe un subgrupo normal de orden q; se toma un subgrupo de orden p, se comprueba que la intersección es trivial por Lagrange (los órdenes p y q son primos distintos) y que el producto de órdenes es pq. El teorema hace el resto, y solo queda decidir cuántos φ posibles hay.

Cuidado

[Contraejemplo: hay grupos que no se descomponen] No todo grupo admite esta descomposición. En el grupo de los cuaterniones $Q_8 = \{\pm 1, \pm \mathrm{i}, \pm \mathrm{j}, \pm \mathrm{k}\}$, *todo* subgrupo no trivial contiene al elemento -1, de modo que dos subgrupos cualesquiera no triviales se cortan al menos en $\{1, -1\}$ y nunca solo en $\{1\}$. Lo mismo ocurre en $\mathbb{Z}_4$, cuyos únicos subgrupos son $\{0\}$, $\{0, 2\}$ y el total: no hay manera de partirlo. Segundo aviso: la condición de normalidad se le pide a N, no a H; si ambos fueran normales el producto sería *directo*, no semidirecto, y el grupo sería una simple yuxtaposición.

Idea clave

Piénsese en el grupo de movimientos rígidos del plano que fijan un polígono, o mejor, en las isometrías del plano entero: una traslación seguida de un giro no es lo mismo que el giro seguido de la traslación, pero cualquier isometría se escribe de forma única como traslación $\circ$ giro. Las traslaciones forman la parte normal N — conjugar una traslación por un giro da otra traslación, girada — y los giros forman el complemento H. La fórmula del producto no es más que la contabilidad de ese cruce: para juntar (n_1, h_1) con (n_2, h_2) hay que pasar h_1 por delante de n_2, y al hacerlo n_2 se transforma en $\varphi(h_1)(n_2)$. El producto directo es el caso en que nada se transforma al cruzar.

Regla: clasificar los grupos de orden pq con $p > q$ primos

El caso $|G| = pq$ es el ejemplo modelo donde Sylow y el producto semidirecto se combinan para dar una respuesta completa.

1. **Cuéntense los p-subgrupos de Sylow.** n_p divide a q y es $\equiv 1 \pmod p$. Como $q < p$, la única posibilidad es $n_p = 1$: hay un único subgrupo P de orden p, y por unicidad es normal.

2. **Cuéntense los q-subgrupos de Sylow.** n_q divide a p y es $\equiv 1 \pmod q$, luego $n_q \in \{1, p\}$, y $n_q = p$ solo es posible si $q \mid p - 1$. Este es el punto donde se bifurca todo.

3. **Si $q \nmid p - 1$, concluir que G es cíclico.** Entonces $n_q = 1$, ambos Sylow son normales, se cortan trivialmente y $G \cong \mathbb{Z}/p\mathbb{Z} \times \mathbb{Z}/q\mathbb{Z} \cong \mathbb{Z}/pq\mathbb{Z}$ por el teorema chino del resto. No hay nada más que decir.

4. **Si $q \mid p - 1$, montar el producto semidirecto.** Tómese $P \cong \mathbb{Z}/p\mathbb{Z}$ normal y $Q \cong \mathbb{Z}/q\mathbb{Z}$ un q-Sylow. Como $P \cap Q = 1$ por órdenes coprimos y $|PQ| = pq$, se tiene $G \cong \mathbb{Z}/p\mathbb{Z} \rtimes_\varphi \mathbb{Z}/q\mathbb{Z}$.

5. **Enumérense los homomorfismos $\varphi\colon \mathbb{Z}/q\mathbb{Z} \to \mathrm{Aut}(\mathbb{Z}/p\mathbb{Z})$.** El grupo de automorfismos es cíclico de orden $p - 1$, así que un tal φ queda determinado por la imagen del generador, que debe ser un elemento de orden divisor de q. Hay exactamente q elecciones: la trivial y $q - 1$ de orden q.

6. **Agrúpense las elecciones no triviales, que dan todas el mismo grupo.** Las $q-1$ acciones fieles se obtienen unas de otras reparametrizando el generador de Q, luego producen grupos isomorfos. Conclusión: exactamente *dos* grupos de orden pq, el cíclico y uno no abeliano.

7. **Descríbase el no abeliano con una presentación.** Es $\langle a, b \mid a^p,\ b^q,\ bab^{-1} = a^k\rangle$, donde k es un entero cuyo orden multiplicativo módulo p es exactamente q. Distintos k válidos dan grupos isomorfos, en coherencia con el paso anterior.

Ejemplo 25.59 (Cuatro órdenes clasificados por completo)**. (a) Orden** $6 = 3 \cdot 2$**.** Aquí $p = 3$, $q = 2$ y $q \mid p - 1 = 2$, luego hay dos grupos: $\mathbb{Z}/6\mathbb{Z}$ y el no abeliano. Éste último es $\mathbb{Z}/3\mathbb{Z} \rtimes \mathbb{Z}/2\mathbb{Z}$ con la acción $a \mapsto a^{-1}$, es decir $\langle a, b \mid a^3, b^2, bab^{-1} = a^2\rangle$, que es S_3, que a su vez es D_3. Tres nombres para el mismo grupo.

(b) Orden $15 = 5 \cdot 3$**.** Ahora $p = 5$, $q = 3$ y $p - 1 = 4$, que no es divisible por 3. El paso 3 se aplica y la conclusión es inmediata: *todo* grupo de orden 15 es cíclico. Este es el argumento entero, y merece subrayarse lo fuerte que es la conclusión comparada con lo poco que costó.

(c) Orden $21 = 7 \cdot 3$**.** Aquí $p = 7$, $q = 3$ y $3 \mid 6$, luego hay exactamente dos grupos de orden 21. El no abeliano se construye buscando un k de orden 3 módulo 7: como $2^1 = 2$, $2^2 = 4$ y $2^3 = 8 \equiv 1$, sirve $k = 2$. Queda $\langle a, b \mid a^7,\ b^3,\ bab^{-1} = a^2\rangle$. El otro k posible es 4, y da un grupo isomorfo (basta sustituir b por b^2).

(d) Orden 20**: más allá de** pq**.** Con $|G| = 20 = 2^2 \cdot 5$ el 5-Sylow es normal y hay cinco grupos en total. Uno de ellos es $\mathbb{Z}/5\mathbb{Z} \rtimes \mathbb{Z}/4\mathbb{Z}$ con acción *fiel*, el llamado grupo de Frobenius F_{20}: el generador de $\mathbb{Z}/4\mathbb{Z}$ actúa como $a \mapsto a^2$, y en efecto 2 es raíz primitiva módulo 5, pues $2, 4, 3, 1$ recorre todas las unidades. Ese grupo aparece de forma natural como el grupo de Galois de ciertas quínticas resolubles.

Ejemplo 25.60 (Semidirectos por todas partes)**. (a) Diédricos.** $D_n = \mathbb{Z}/n\mathbb{Z} \rtimes \mathbb{Z}/2\mathbb{Z}$ con la acción $\varphi(s)(r) = r^{-1}$. La comprobación es la relación $srs^{-1} = r^{-1}$ de la presentación. Aquí $N = \langle r\rangle$ es normal por tener índice 2, y $H = \langle s\rangle$ es un complemento porque $s \notin \langle r\rangle$.

(b) Alternado y simétrico. $S_n = A_n \rtimes \mathbb{Z}/2\mathbb{Z}$ para todo $n \geq 2$: A_n es normal de índice 2, y cualquier transposición genera un complemento. Un escalón más abajo, $A_4 = V_4 \rtimes \mathbb{Z}/3\mathbb{Z}$, donde $V_4 = \{1, (12)(34), (13)(24), (14)(23)\}$ es normal y el 3-ciclo (123) permuta cíclicamente los tres elementos no triviales de V_4. Encadenando, $S_4 = V_4 \rtimes S_3$, que es la razón estructural de que la cuártica sea resoluble por radicales (capítulo 26).

(c) Matrices afines. El grupo de las aplicaciones $x \mapsto ax + b$ con $a \neq 0$, sobre un cuerpo K, es $K \rtimes K^\times$: las traslaciones b forman el subgrupo normal y las homotecias a el complemento, con la acción $a \cdot b = ab$. En forma matricial son las matrices $\begin{pmatrix} a & b \\ 0 & 1 \end{pmatrix}$, y el subgrupo normal es el de las que tienen $a = 1$. Sobre $\mathbb{F}_5$ este grupo es precisamente el F_{20} del ejemplo anterior.

(d) Cuándo NO hay complemento. $\mathbb{Z}/4\mathbb{Z}$ contiene $N = \{0, 2\} \cong \mathbb{Z}/2\mathbb{Z}$ como subgrupo normal, pero no existe ningún H de orden 2 con $N \cap H = 1$, porque N es el *único* subgrupo de orden 2. Luego $\mathbb{Z}/4\mathbb{Z}$ no se descompone como semidirecto no trivial de sus piezas: es una

extensión de $\mathbb{Z}/2\mathbb{Z}$ por $\mathbb{Z}/2\mathbb{Z}$ que *no escinde*. Lo mismo le pasa a Q_8 respecto de su centro. Medir cuántas extensiones no escinden existe es el objeto de la cohomología de grupos.

Cuidado

Tres cautelas. Primera: en $N \rtimes H$ el subgrupo normal es N, el de la izquierda; escribir $H \rtimes N$ es un error de tipo, no una variante notacional. Segunda: distintas φ pueden dar grupos isomorfos, así que contar homomorfismos $H \to \mathrm{Aut}(N)$ sobreestima el número de grupos — hay que agrupar las acciones que difieren en un automorfismo de H o de N. Tercera: que N sea normal no basta para que exista complemento, como muestra el caso (d); el teorema de Schur–Zassenhaus garantiza que sí existe cuando $|N|$ y $|G/N|$ son coprimos, que es justamente lo que hace funcionar todo el análisis de los grupos de orden pq.

Pausa para pensar

Con la regla anterior el lector puede clasificar sin ayuda todos los grupos de orden pq para $pq < 100$. Pruébese con $33 = 11 \cdot 3$ ($3 \nmid 10$, luego solo el cíclico), con $55 = 11 \cdot 5$ ($5 \mid 10$, luego dos grupos) y con $65 = 13 \cdot 5$ ($5 \nmid 12$, solo el cíclico). ¿Y p^2? Ahí el argumento cambia por completo: todo grupo de orden p^2 es abeliano, y por tanto es $\mathbb{Z}/p^2\mathbb{Z}$ o $\mathbb{Z}/p\mathbb{Z} \times \mathbb{Z}/p\mathbb{Z}$. La demostración usa que el centro de un p-grupo no es trivial, y es uno de los argumentos más limpios de toda la teoría.

25.8. Tablas de caracteres: la huella dactilar de un grupo finito

Un grupo abstracto es difícil de manipular; una matriz se calcula. Una representación traduce el grupo a matrices, y quedarse solo con la traza de cada matriz — un único número por clase de conjugación — parece tirar casi toda la información. La sorpresa, y es de las mayores del álgebra, es que esa sombra basta: la tabla de caracteres determina la representación salvo isomorfismo, y con ella se leen el orden del grupo, sus subgrupos normales, su carácter abeliano o soluble. Esta sección enseña a construirla a mano.

Definición 25.61 (Representación, carácter y tabla)**.** Una *representación* de un grupo finito G sobre $\mathbb{C}$ es un homomorfismo $\rho\colon G \to \mathrm{GL}(V)$ con V un espacio vectorial complejo de dimensión finita d, llamada el *grado*. Su *carácter* es la función $\chi_\rho(g) = \operatorname{tr} \rho(g)$, que es constante en cada clase de conjugación y cumple $\chi_\rho(1) = d$. La representación es *irreducible* si V no tiene subespacios propios invariantes. La *tabla de caracteres* de G es la matriz cuadrada cuyas filas son los caracteres irreducibles y cuyas columnas son las clases de conjugación.

Teorema 25.62 (Relaciones de ortogonalidad)**.** *Con el producto interno* $\langle \chi, \psi \rangle = \frac{1}{|G|} \sum_{g \in G} \chi(g)\overline{\psi(g)}$*, los caracteres irreducibles* $\chi_1, \dots, \chi_r$ *forman una base ortonormal del espacio de funciones de clase. En particular* $\langle \chi_i, \chi_j \rangle = \delta_{ij}$*, el número* r *de irreducibles coincide con el número de clases de conjugación, los grados cumplen* $\sum_i \chi_i(1)^2 = |G|$ *y cada grado divide a* $|G|$*. Además, para las columnas,* $\sum_i \chi_i(g)\overline{\chi_i(g)} = |C_G(g)| = |G|/|\mathrm{cl}(g)|$.

Idea clave

Qué dice esto en cristiano. Los caracteres irreducibles se comportan como una base ortonormal, igual que los vectores $\hat{\imath}, \hat{\jmath}, \hat{k}$ del espacio: cada uno tiene "longitud" 1 y son perpendiculares dos a dos, con el producto interno que promedia sobre el grupo. La consecuencia práctica es enorme, porque convierte preguntas difíciles en cuentas de sumar y dividir: para saber si un carácter es irreducible basta comprobar que $\langle \chi, \chi \rangle = 1$, y para saber cuántas veces aparece χ_i dentro de una representación cualquiera basta calcular un producto interno. Toda la teoría de representaciones se reduce a hacer geometría en un espacio de funciones.

Ejemplo 25.63 (La tabla de S_3, comprobada casilla a casilla)**. (a) Cuántos hay y de qué tamaño.** S_3 tiene tres clases de conjugación: la identidad (1 elemento), las trasposiciones (3) y los tres-ciclos (2). Luego hay exactamente tres caracteres irreducibles. Sus dimensiones d_i cumplen $\sum d_i^2 = |G| = 6$, y la única posibilidad con tres sumandos es $1+1+4$, es decir $d = 1, 1, 2$.

(b) La tabla. Escribiendo los valores en el orden (identidad, trasposiciones, tres-ciclos):

	1	3	2
χ_{triv}	1	1	1
χ_{sgn}	1	-1	1
χ_{std}	2	0	-1

La primera columna da las dimensiones, y en efecto $1^2 + 1^2 + 2^2 = 6$.

(c) Comprobar la ortonormalidad. Para χ_{std}: $\langle \chi, \chi \rangle = \frac{1}{6}\left(1 \cdot 2^2 + 3 \cdot 0^2 + 2 \cdot (-1)^2\right) = \frac{4+0+2}{6} = 1$, luego es irreducible. Y el cruce con el trivial: $\frac{1}{6}\left(1 \cdot 2 \cdot 1 + 3 \cdot 0 \cdot 1 + 2 \cdot (-1) \cdot 1\right) = \frac{2-2}{6} = 0$. Los tres productos cruzados dan cero y las tres normas dan uno.

(d) Un grupo abeliano. En $\mathbb{Z}_3$ cada elemento es su propia clase, luego hay tres caracteres, todos de dimensión 1: son $\chi_k(j) = \omega^{jk}$ con $\omega = \mathrm{e}^{2\pi \mathrm{i}/3}$. La ortogonalidad se convierte aquí en la identidad $1 + \omega + \omega^2 = 0$, que ya conocíamos de las raíces de la unidad. La transformada de Fourier discreta no es más que el cambio de base a estos caracteres.

Cuidado

[Qué determina y qué no determina la tabla] El número de irreducibles es el número de *clases de conjugación*, no el de elementos ni el de subgrupos: S_4 tiene 24 elementos y solo 5 irreducibles, con dimensiones $1, 1, 2, 3, 3$ y $1+1+4+9+9 = 24$. Segundo aviso, y es el famoso: la tabla de caracteres *no* determina el grupo. D_4 y Q_8 tienen ambos orden 8, ambos cinco clases de conjugación y *exactamente la misma tabla de caracteres*, y sin embargo no son isomorfos. La huella dactilar es muy buena, pero no es única.

Idea clave

La ortogonalidad convierte la teoría de representaciones en geometría euclídea. Cada carácter es un vector; los irreducibles son ejes perpendiculares de longitud uno; y descomponer una representación cualquiera en irreducibles es, literalmente, proyectar un vector sobre esos ejes: el número de veces que χ_i aparece en χ es el producto escalar $\langle \chi, \chi_i \rangle$. Por eso el criterio de irreducibilidad es tan cómodo: χ es irreducible si y solo si $\langle \chi, \chi \rangle = 1$, es decir si el vector tiene norma uno. Y por eso la tabla es cuadrada: hay tantos ejes como dimensiones tiene el espacio de funciones de clase, que es el número de clases.

Regla: construir la tabla de caracteres de un grupo pequeño

Se construye por eliminación: se ponen las filas fáciles, se usan las restricciones numéricas y lo que queda se despeja.

1. **Calcúlense las clases de conjugación y sus tamaños.** El número de clases es el número de filas y de columnas. Compruébese que los tamaños suman $|G|$ y que cada uno divide a $|G|$, pues son índices de centralizadores.

2. **Determínense los grados con la ecuación $\sum d_i^2 = |G|$.** Se sabe que $d_1 = 1$ (el carácter trivial), que cada d_i divide a $|G|$ y que el número de grados iguales a 1 es $|G/[G,G]|$. Con esas tres restricciones, la partición de $|G|$ en cuadrados suele quedar determinada.

3. **Escríbanse las filas de grado 1.** Son los homomorfismos $G \to \mathbb{C}^\times$, es decir los caracteres del abelianizado $G/[G,G]$. Si ese cociente es $\mathbb{Z}/n\mathbb{Z}$, hay n filas de unos y raíces n-ésimas de la unidad.

4. **Constrúyanse representaciones concretas para las filas de grado alto.** Las fuentes habituales son la representación por permutación de una acción de G y los productos de caracteres ya obtenidos. La representación por permutación sobre un conjunto X tiene carácter $\chi(g) = \#\{x \in X : gx = x\}$: se cuentan puntos fijos, nada más.

5. **Descompóngase lo obtenido restando lo ya conocido.** El carácter de permutación siempre contiene al trivial. Réstese: $\chi - \chi_{\text{triv}}$, y compruébese con $\langle \cdot, \cdot \rangle = 1$ si lo que queda ya es irreducible.

6. **Complétese la última fila con la ortogonalidad de columnas.** Si falta un solo carácter, cada entrada suya se despeja de $\sum_i |\chi_i(g)|^2 = |G|/|\mathrm{cl}(g)|$, o más cómodamente de la ortogonalidad de la columna de g contra la del neutro: $\sum_i \chi_i(1)\overline{\chi_i(g)} = 0$ para $g \neq 1$.

7. **Verifíquese la tabla entera.** Toda fila debe tener norma 1, todo par de filas distintas producto escalar 0, la primera columna debe contener los grados y las columnas deben cumplir la segunda relación de ortogonalidad. Si algo falla, el error está casi siempre en un tamaño de clase.

Ejemplo 25.64 (Tres tablas construidas paso a paso)**. (a) El grupo cíclico $\mathbb{Z}/3\mathbb{Z}$.** Es abeliano, luego cada elemento es su propia clase: tres clases, tres irreducibles, y $1^2+1^2+1^2=3$ obliga a que todos los grados sean 1. Con $\omega=\mathrm{e}^{2\pi\mathrm{i}/3}$ los tres caracteres son $\chi_k(a^j)=\omega^{kj}$:

	1	a	a^2
χ_0	1	1	1
χ_1	1	ω	ω^2
χ_2	1	ω^2	ω

La ortogonalidad de las filas χ_0 y χ_1 es exactamente la identidad $1+\omega+\omega^2=0$. En general, para un grupo abeliano todos los grados valen 1 y la tabla es la matriz de Fourier discreta: la transformada de Fourier finita *es* la teoría de caracteres de $\mathbb{Z}/n\mathbb{Z}$.

(b) El grupo simétrico S_3. Las clases son $\{1\}$, las tres transposiciones y los dos 3-ciclos, de tamaños $1,3,2$ que suman 6. Tres clases, tres irreducibles; como $[S_3,S_3]=A_3$ y el abelianizado es $\mathbb{Z}/2\mathbb{Z}$, hay exactamente dos filas de grado 1, y $1+1+d^2=6$ da $d=2$. Las dos primeras filas son el carácter trivial y el signo. Para la tercera, tómese la acción sobre $\{1,2,3\}$: su carácter cuenta puntos fijos, o sea $(3,1,0)$. Restándole el trivial queda $(2,0,-1)$, cuya norma es $\frac{1}{6}(1\cdot4+3\cdot0+2\cdot1)=1$: irreducible.

	1	$3\cdot(12)$	$2\cdot(123)$
trivial	1	1	1
signo	1	-1	1
estándar	2	0	-1

Las columnas también cuadran: la del neutro da $1+1+4=6=|G|$, la de las transposiciones $1+1+0=2=6/3$ y la de los 3-ciclos $1+1+1=3=6/2$.

(c) D_4 y Q_8, o por qué la tabla no lo dice *todo*. Ambos tienen orden 8 y cinco clases de conjugación, ambos tienen abelianizado $\mathbb{Z}/2\mathbb{Z}\times\mathbb{Z}/2\mathbb{Z}$ — luego cuatro filas de grado 1 — y $1+1+1+1+d^2=8$ fuerza $d=2$. Al completar la tabla siguiendo los pasos anteriores se obtiene *exactamente la misma matriz* para los dos grupos. Y sin embargo no son isomorfos: D_4 tiene cinco elementos de orden 2 y Q_8 tiene uno solo. Es el contraejemplo estándar: la tabla de caracteres es un invariante muy fino, pero no completo.

Ejemplo 25.65 (Lo que la tabla revela de un vistazo)**. (a) El orden del grupo y sus grados.** La primera columna son los grados, y la suma de sus cuadrados es $|G|$. Para A_4 los grados son $1,1,1,3$ y en efecto $1+1+1+9=12$; para S_4 son $1,1,2,3,3$ con $1+1+4+9+9=24$; para A_5 son $1,3,3,4,5$ con $1+9+9+16+25=60$. Además cada grado divide al orden, lo que en el caso de A_5 descarta de entrada particiones como $1+1+\ldots$

(b) Si el grupo es abeliano. Ocurre si y solo si todos los grados valen 1, es decir si la tabla tiene tantas filas como elementos. Más aún, el número de filas de grado 1 es siempre $|G/[G,G]|$, de modo que la tabla mide de un golpe cuánto le falta al grupo para ser conmutativo.

(c) Los subgrupos normales, sin buscarlos. Para cada carácter irreducible χ, el conjunto $\ker\chi=\{g:\chi(g)=\chi(1)\}$ es un subgrupo normal, y *todo* subgrupo normal de G es

intersección de núcleos de ese tipo. Así, un grupo es simple exactamente cuando ninguna fila salvo la trivial tiene una entrada igual a su grado. Mirando la tabla de A_5 se comprueba en cinco segundos que es simple, sin manipular ni una permutación.

(d) Solubilidad, vía Burnside. El teorema $p^a q^b$ de Burnside — todo grupo cuyo orden solo tiene dos divisores primos es soluble — se demostró en 1904 con caracteres, y durante más de medio siglo no se conoció ninguna demostración que los evitase. Fue la primera gran victoria de las representaciones sobre la teoría de grupos pura, y el antecedente directo del teorema de Feit–Thompson y de la clasificación de los grupos simples finitos (capítulo 26 para lo que esto significa en las ecuaciones).

Cuidado

El carácter no es un homomorfismo: en general $\chi(gh) \neq \chi(g)\chi(h)$, porque la traza de un producto no es el producto de las trazas. Lo que sí vale es $\chi(hgh^{-1}) = \chi(g)$, que es la constancia en clases, y $\chi(g^{-1}) = \overline{\chi(g)}$. Otro error frecuente es contar mal las clases de conjugación: en S_n coinciden con los tipos de ciclo, pero en A_n una clase de S_n puede partirse en dos — ocurre por ejemplo con los 5-ciclos en A_5, y por eso A_5 tiene cinco clases y no cuatro. Y todo lo anterior se apoya en que el cuerpo sea $\mathbb{C}$; sobre un cuerpo de característica que divida a $|G|$ la teoría cambia radicalmente.

Pausa para pensar

La tabla de caracteres de un grupo cabe en una servilleta y contiene una cantidad asombrosa de su estructura. Por eso la química cuántica clasifica los modos de vibración de una molécula con la tabla de caracteres de su grupo de simetría, y por eso la física de partículas describe familias enteras de partículas como representaciones irreducibles de SU(3). La pregunta natural para el lector es la inversa: dada una matriz cuadrada de números complejos, ¿cómo saber si es la tabla de caracteres de algún grupo? Se conocen muchas condiciones necesarias — ortogonalidad de filas y columnas, grados que dividen al orden, entradas que son enteros algebraicos — pero ninguna caracterización completa. Sigue siendo un problema abierto.

25.9. Grupos solubles y grupos simples

Idea clave

Intuición primero. Piensa en números: los primos son los que no se pueden partir, y todo número se desarma en primos. Con grupos ocurre lo mismo, y los "primos" son los grupos *simples*: los que no admiten cociente propio, los que no se pueden desmontar. Un grupo es *soluble* cuando se puede desmontar por completo en piezas *abelianas*, es decir, en piezas conmutativas y por tanto dóciles. La palabra "soluble" no es casual: resolver una ecuación con radicales significa ir despejando por capas, y cada radical añade una capa conmutativa. Si el grupo de la ecuación se desmonta en capas abelianas, hay fórmula; si en el desmontaje aparece una pieza simple no abeliana —un "átomo

indivisible y no conmutativo", como A_5— no la hay. Esa es toda la teoría de Galois en cuatro líneas; el capítulo 26 solo la hace precisa.

Definición 25.66. G es *soluble* si existe una cadena $\{e\} = G_0 \trianglelefteq G_1 \trianglelefteq \cdots \trianglelefteq G_k = G$ con cada cociente G_{i+1}/G_i abeliano. G es *simple* si no tiene subgrupos normales propios no triviales.

Idea clave

Qué dice esto en cristiano. Un grupo es soluble si se puede desmontar en capas y todas las capas son abelianas. Piénsese en una cebolla: se va quitando un subgrupo normal, se mira lo que queda (G_{i+1}/G_i) y se exige que eso sea conmutativo; si al final se llega al grupo trivial, el grupo era soluble. Un grupo simple es el extremo opuesto: no se puede quitar nada, es un átomo, no hay ningún subgrupo normal por donde cortar. Y aquí está la conexión que da nombre a todo el capítulo siguiente: una ecuación se resuelve por radicales exactamente cuando cierto grupo asociado a ella es soluble, y los radicales sucesivos son precisamente esas capas abelianas.

Ejemplo 25.67 (Desmontar grupos en capas)**. (a) S_3 es soluble.** La cadena es $\{e\} \trianglelefteq A_3 \trianglelefteq S_3$. El primer cociente es $A_3/\{e\} = A_3 \cong \mathbb{Z}_3$, abeliano; el segundo es $S_3/A_3 \cong \mathbb{Z}_2$, también abeliano. Dos capas y hemos terminado.

(b) S_4 es soluble, con tres capas. La cadena es $\{e\} \trianglelefteq V_4 \trianglelefteq A_4 \trianglelefteq S_4$, donde V_4 son la identidad y los tres productos de dos trasposiciones disjuntas. Los cocientes tienen órdenes 4, $\frac{12}{4} = 3$ y $\frac{24}{12} = 2$, y son respectivamente $\mathbb{Z}_2 \times \mathbb{Z}_2$, $\mathbb{Z}_3$ y $\mathbb{Z}_2$: los tres abelianos. Por eso las ecuaciones de grado 3 y 4 tienen fórmula por radicales, con sus raíces cuadradas y cúbicas encadenadas.

(c) A_5 no es soluble. Su subgrupo derivado (el generado por todos los conmutadores) es *el propio* A_5, de orden 60. Como la serie derivada se queda ahí atascada y nunca llega al trivial, no hay manera de desmontarlo en capas abelianas. Compárese con A_4, cuyo derivado tiene orden 4 y desde ahí baja al trivial sin problema.

(d) Y además A_5 es simple. Sus clases de conjugación tienen tamaños $1, 20, 15, 12, 12$, que suman 60. Un subgrupo normal es unión de clases de conjugación, tiene que contener al 1 y su orden debe dividir a 60. Pero $1 + 15 = 16$, $1 + 20 = 21$, $1 + 12 = 13$, $1 + 12 + 12 = 25$ y $1 + 15 + 20 = 36$: ninguna suma parcial que contenga al 1 da un divisor de 60, salvo el 1 solo y el total. Luego no hay subgrupos normales propios. Esta cuenta de cinco números es la demostración entera.

Cuidado

[Contraejemplo: soluble no significa fácil, ni abeliano] Soluble no quiere decir abeliano: S_3 y S_4 son solubles y ninguno es conmutativo. Segundo aviso: la relación $\trianglelefteq$ de la cadena solo exige que cada G_i sea normal *en el siguiente*, no en todo G; en la cadena de S_4, el subgrupo V_4 resulta ser normal en S_4, pero eso es un regalo y no forma parte de la definición. Y tercero: que S_5 no sea soluble no significa que la quíntica no tenga

soluciones, sino que no hay una *fórmula con radicales* para ellas; $x^5 - 2 = 0$ tiene la solución perfectamente expresable $\sqrt[5]{2}$, porque su grupo concreto sí es soluble aunque el grupo genérico de grado 5 no lo sea.

Hechos: todo abeliano es soluble; subgrupos y cocientes de solubles son solubles; si N y G/N son solubles, G lo es. S_3 es soluble ($\{e\} \trianglelefteq A_3 \trianglelefteq S_3$, cocientes $\mathbb{Z}_3, \mathbb{Z}_2$); S_4 es soluble (cadena $\{e\} \trianglelefteq V \trianglelefteq A_4 \trianglelefteq S_4$, donde $V = \{e, (12)(34), (13)(24), (14)(23)\}$ es el grupo de Klein, con cocientes $V, \mathbb{Z}_3, \mathbb{Z}_2$). Un grupo simple no abeliano *no* es soluble.

Teorema 25.68. *A_5 es simple. En consecuencia, S_n no es soluble para $n \geq 5$.*

Idea de la demostración de la simplicidad. $|A_5| = 60$, y sus clases de conjugación tienen tamaños $1, 15, 20, 12, 12$ (identidad, dobles trasposiciones, 3-ciclos, y los 5-ciclos partidos en dos clases). Un subgrupo normal es unión de clases de conjugación que contiene a la identidad, y su orden debe dividir a 60. Pero ninguna suma $1 + \{15, 20, 12, 12\}$ da un divisor propio de 60 (probar: $1 + 15 = 16$, $1 + 20 = 21$, $1 + 12 = 13$, $1 + 15 + 20 = 36$, ...): no hay subgrupos normales propios. Para $n > 5$, A_n también es simple, y la cadena $\{e\} \trianglelefteq A_n \trianglelefteq S_n$ no puede refinarse con cocientes abelianos: S_n no es soluble. □

Idea clave

Qué dice esto en cristiano. Este es el teorema que rompe la simetría entre los grados bajos y los altos. Hasta grado 4 los grupos asociados a las ecuaciones se pueden desmontar en capas abelianas, y por eso existen las fórmulas de Cardano y Ferrari; a partir de grado 5 aparece A_5, que es un átomo indivisible, y la cadena de capas se atasca antes de empezar. No es que nadie haya encontrado la fórmula: es que la fórmula no puede existir, y este teorema de una línea es la razón.

Ejemplo 25.69 (Por qué A_5 no se puede partir)**. (a) Las clases de conjugación.** A_5 tiene 60 elementos repartidos en cinco clases de tamaños 1, 15, 20, 12 y 12: la identidad, los productos de dos trasposiciones disjuntas, los tres-ciclos y dos familias distintas de cinco-ciclos. La suma es $1 + 15 + 20 + 12 + 12 = 60$.

(b) El argumento completo. Un subgrupo normal es unión de clases de conjugación, debe contener a la identidad y su orden debe dividir a 60. Las sumas posibles que incluyen el 1 son $1+15 = 16$, $1+20 = 21$, $1+12 = 13$, $1+12+12 = 25$, $1+15+20 = 36$, $1+15+12 = 28$, $1+20+12 = 33$... y ninguna divide a 60 salvo el 1 solo y el total 60. Luego no hay subgrupos normales propios.

(c) La consecuencia sobre S_n. Para $n \geq 5$, el grupo A_n es simple y no abeliano, de modo que aparece como una capa en cualquier intento de descomponer S_n y no se deja partir más. Numéricamente: el subgrupo derivado de A_5 es el propio A_5, de orden 60, mientras que el de A_4 tiene orden 4 y desde ahí la serie derivada baja hasta el trivial sin problema.

(d) El contraste con los grados pequeños. S_3 se desmonta con la cadena $1 \trianglelefteq A_3 \trianglelefteq S_3$, de cocientes $\mathbb{Z}_3$ y $\mathbb{Z}_2$. S_4 con $1 \trianglelefteq V_4 \trianglelefteq A_4 \trianglelefteq S_4$, de cocientes de órdenes 4, 3 y 2. Las raíces

cuadradas y cúbicas de las fórmulas clásicas se corresponden exactamente con esos cocientes. En grado 5 no hay ninguna cadena que ofrecer, y por tanto tampoco ninguna torre de radicales.

Cuidado

[Lo que el teorema no dice] No dice que las ecuaciones de grado 5 carezcan de soluciones: toda ecuación polinómica compleja tiene sus raíces, garantizadas por el teorema fundamental del álgebra. Tampoco dice que ninguna quíntica concreta sea resoluble por radicales: $x^5 - 2 = 0$ tiene la solución $\sqrt[5]{2}$, escrita con un radical, porque su grupo particular sí es soluble. Lo que afirma es que no existe una *fórmula general* válida para todas, hecha con las cuatro operaciones y radicales. Y tampoco significa que no haya soluciones explícitas de ningún tipo: Hermite dio en 1858 una fórmula para la quíntica usando funciones elípticas, que no son radicales pero son perfectamente explícitas.

Idea clave

Guarda este teorema: la frontera entre S_4 (soluble) y S_5 (no soluble) será *exactamente* la frontera entre las ecuaciones de grado ≤ 4 (con fórmula) y las de grado ≥ 5 (sin fórmula). Todo el capítulo 26 desemboca ahí.

Regla: contar órbitas con la ecuación de clases y el lema de Burnside

Cuando hay que contar configuraciones «salvo simetría», la respuesta casi nunca es dividir por el número de simetrías. La herramienta correcta es Burnside.

1. **Identifica el conjunto y el grupo.** Escribe explícitamente el conjunto X de configuraciones *etiquetadas* (sin identificar nada) y el grupo G de simetrías que actúa sobre él. Es el paso donde se cometen los errores.

2. **Recuerda órbita–estabilizador.** Para cada $x \in X$,

$$|G| = |\mathrm{Orb}(x)| \cdot |\mathrm{Est}(x)|.$$

   De aquí sale que el tamaño de toda órbita divide al orden del grupo.

3. **Escribe la ecuación de clases.** Si G actúa sobre sí mismo por conjugación, las órbitas son las clases de conjugación y los estabilizadores los centralizadores:

$$|G| = |Z(G)| + \sum_i [G : C_G(x_i)],$$

   sumando sobre representantes de las clases con más de un elemento. Es la herramienta para demostrar que los p-grupos tienen centro no trivial.

4. **Para contar órbitas, aplica Burnside.** El número de órbitas es

$$\frac{1}{|G|} \sum_{g \in G} |\mathrm{Fij}(g)|,$$

donde $\mathrm{Fij}(g) = \{x \in X : g \cdot x = x\}$. Es un promedio de puntos fijos, no un cociente.

5. **Agrupa por tipo de simetría.** No recorras los elementos uno a uno: agrúpalos por clases de conjugación, que tienen el mismo número de puntos fijos. Basta calcular un representante por clase y multiplicar por el tamaño.

6. **Cuenta puntos fijos por ciclos.** Si g permuta las posiciones formando $c(g)$ ciclos y hay n colores disponibles, entonces $|\mathrm{Fij}(g)| = n^{c(g)}$: cada ciclo debe ser monocromo. Esta observación resuelve casi todos los problemas de coloreado.

7. **Verifica en casos extremos.** Con un solo color el resultado debe ser 1. Con muchísimos colores, el resultado debe acercarse a $|X|/|G|$, porque casi ninguna configuración tiene simetrías.

Ejemplo 25.70 (Cuatro conteos con simetría)**. (a) Collares de seis cuentas y dos colores.** El grupo es C_6, las rotaciones del collar, y X tiene $2^6 = 64$ elementos. La rotación por k posiciones descompone las seis posiciones en $\gcd(k, 6)$ ciclos, luego $|\mathrm{Fij}(r^k)| = 2^{\gcd(k,6)}$. Sumando sobre $k = 0, \ldots, 5$:

$$2^6 + 2^1 + 2^2 + 2^3 + 2^2 + 2^1 = 64 + 2 + 4 + 8 + 4 + 2 = 84,$$

y el número de collares distintos es $84/6 = 14$. Obsérvese que $64/6$ no es ni siquiera entero: dividir a lo bruto habría sido imposible, precisamente porque algunas configuraciones (como la alternada) tienen simetrías propias.

(b) Caras de un cubo con dos colores. El grupo de rotaciones del cubo tiene 24 elementos, repartidos en cinco clases: la identidad (1 elemento, 6 ciclos sobre las caras), los giros de 90° y 270° respecto a ejes de caras (6 elementos, 3 ciclos), los de 180° de cara (3 elementos, 4 ciclos), los de 120° de vértice (8 elementos, 2 ciclos) y los de 180° de arista (6 elementos, 3 ciclos). Con $n = 2$:

$$\frac{1}{24}\left(2^6 + 6 \cdot 2^3 + 3 \cdot 2^4 + 8 \cdot 2^2 + 6 \cdot 2^3\right) = \frac{64 + 48 + 48 + 32 + 48}{24} = \frac{240}{24} = 10.$$

Hay exactamente 10 dados de dos colores esencialmente distintos.

(c) El mismo cubo con tres colores. Cambiando n:

$$\frac{1}{24}\left(3^6 + 6 \cdot 3^3 + 3 \cdot 3^4 + 8 \cdot 3^2 + 6 \cdot 3^3\right) = \frac{729 + 162 + 243 + 72 + 162}{24} = \frac{1368}{24} = 57.$$

La fórmula es un polinomio en n, y ese polinomio es el *índice de ciclos* del grupo. Es la base del teorema de enumeración de Pólya, que con la misma maquinaria cuenta también cuántas configuraciones usan cada color un número prefijado de veces.

(d) La ecuación de clases de S_4 y A_5. En S_4 las clases de conjugación son los tipos de ciclo: la identidad (1 elemento), las trasposiciones (6), los 3-ciclos (8), los 4-ciclos (6) y los productos de dos trasposiciones (3). Comprobación: $1 + 6 + 8 + 6 + 3 = 24$. En A_5 las clases son la identidad (1), los 3-ciclos (20), los productos de dos trasposiciones (15) y los 5-ciclos, que en A_5 se *parten* en dos clases de 12 cada una. Total: $1 + 20 + 15 + 12 + 12 = 60$. Esta descomposición es la que demuestra que A_5 es simple: cualquier subgrupo normal es unión

de clases enteras incluyendo la identidad, y ninguna combinación de $\{1, 20, 15, 12, 12\}$ que contenga el 1 suma un divisor propio de 60.

Cuidado

El error más frecuente es dividir el total entre el orden del grupo. En (a) daría $64/6 \approx 10{,}67$, que ni siquiera es entero. La división correcta solo funciona cuando la acción es *libre*, es decir, cuando ningún elemento no trivial fija ninguna configuración. Burnside corrige exactamente ese defecto promediando los puntos fijos en vez de contar solo la identidad.

Pausa para pensar

La ecuación de clases tiene un corolario que parece magia: todo grupo de orden p^n con p primo tiene centro no trivial. La razón es aritmética pura. Cada término $[G : C_G(x_i)]$ de la suma es un divisor de p^n mayor que 1, luego múltiplo de p; y $|G| = p^n$ también lo es. Despejando, $|Z(G)|$ debe ser múltiplo de p, y como contiene al menos la identidad, es al menos p. De ahí se sigue en tres líneas que todo grupo de orden p^2 es abeliano.

25.10. Anillos

Idea clave

Intuición primero. Un grupo tiene una operación; un anillo tiene dos, y están pensadas para imitar a $\mathbb{Z}$: sumar siempre se puede deshacer, multiplicar no. Todo lo demás del capítulo consiste en preguntar qué propiedades de los enteros sobreviven cuando se cambia el conjunto: ¿se puede dividir con resto?, ¿hay factorización única en primos?, ¿un producto puede dar cero sin que ningún factor lo sea? Cada respuesta afirmativa define un tipo de anillo, y la jerarquía que sigue (dominio, DFU, DIP, euclídeo, cuerpo) es literalmente la lista ordenada de "cuánto se parece esto a los enteros".

Definición 25.71. Un *anillo* (conmutativo con unidad) es $(R, +, \cdot)$ donde $(R, +)$ es grupo abeliano, el producto es asociativo y conmutativo con unidad 1, y distribuye sobre la suma. Una *unidad* es un elemento con inverso multiplicativo (R^* denota el grupo de unidades). Un *divisor de cero* es $a \neq 0$ con $ab = 0$ para algún $b \neq 0$; un *dominio de integridad* es un anillo sin divisores de cero (donde vale cancelar).

Ejemplo 25.72. $\mathbb{Z}$ (unidades ± 1); $\mathbb{Z}_n$ (dominio $\iff$ n primo; las unidades son los coprimos con n); $K[x]$ para K cuerpo (dominio; unidades: constantes no nulas); los enteros de Gauss $\mathbb{Z}[i] = \{a + bi\}$; matrices $M_n(K)$ (no conmutativo, con divisores de cero).

Definición 25.73 (Ideales y cocientes)**.** $I \subseteq R$ es un *ideal* si es subgrupo aditivo y absorbe: $r \in R, a \in I \Rightarrow ra \in I$. El cociente R/I (clases $a + I$ con las operaciones heredadas) es un anillo. El ideal *principal* generado por a es $(a) = aR$. Los ideales son los núcleos de los homomorfismos de anillos, y vale el análogo del primer teorema de isomorfía: $R/\ker\varphi \cong \operatorname{im}\varphi$.

Teorema 25.74. *I es* maximal *(sin ideales estrictamente intermedios)* $\iff$ *R/I es un* ***cuerpo****. I es* primo *($ab \in I \Rightarrow a \in I$ o $b \in I$)* $\iff$ *R/I es dominio.*

Ejemplo 25.75. En $\mathbb{Z}$: los ideales son (n), y (p) es maximal $\iff$ p primo: $\mathbb{Z}_p = \mathbb{Z}/(p)$ es cuerpo. Esta es la plantilla que copiaremos con polinomios.

25.10.1. La jerarquía aritmética

Definición 25.76. Un dominio es *euclídeo* si tiene división con resto (con una noción de "tamaño"); es *de ideales principales* (DIP) si todo ideal es principal; es *de factorización única* (DFU) si todo elemento se factoriza en irreducibles de forma única salvo orden y unidades. Se cumple: euclídeo $\Rightarrow$ DIP $\Rightarrow$ DFU, y ninguna flecha se invierte.

$\mathbb{Z}$, $K[x]$ y $\mathbb{Z}[i]$ son euclídeos (tamaños: valor absoluto, grado, norma a^2+b^2). En cambio $\mathbb{Z}[\sqrt{-5}]$ no es DFU: $6 = 2\cdot 3 = (1+\sqrt{-5})(1-\sqrt{-5})$ son dos factorizaciones genuinamente distintas —el fenómeno cuyo estudio (Kummer, Dedekind) creó la teoría de ideales—. En un DFU existen mcd y vale el lema de Euclides ($p \mid ab \Rightarrow p \mid a$ o $p \mid b$ para p irreducible).

Regla: decidir si un ideal es primo o maximal y describir el cociente

Toda la cuestión se reduce a una sola frase que conviene memorizar: *primo significa que el cociente es dominio; maximal significa que el cociente es cuerpo*. Todo lo demás es aritmética.

1. **Trabaja siempre en un anillo conmutativo con unidad** A. Fuera de ese contexto las definiciones cambian y las equivalencias de abajo dejan de valer.
2. **Prueba la definición directa de primo:** I es primo si $I \neq A$ y de $ab \in I$ se sigue $a \in I$ o $b \in I$. Es literalmente la propiedad del "primo" de toda la vida: si $p \mid ab$ entonces $p \mid a$ o $p \mid b$.
3. **O calcula el cociente,** que suele ser más rápido: I primo $\iff$ A/I es dominio de integridad; I maximal $\iff$ A/I es cuerpo.
4. **Recuerda la jerarquía:** todo ideal maximal es primo (los cuerpos son dominios), pero no al revés. El contraejemplo estándar es (0) en $\mathbb{Z}$, o (x) en $\mathbb{Z}[x]$.
5. **En un dominio de ideales principales** (como $\mathbb{Z}$ o $K[x]$ con K cuerpo) los ideales no nulos primos y los maximales coinciden, y son exactamente los (p) con p irreducible. Aquí "decidir si el ideal es maximal" se convierte en "decidir si el generador es irreducible".
6. **Si el generador se factoriza, usa el teorema chino del resto:** de $f = gh$ con g, h coprimos sale $K[x]/(f) \cong K[x]/(g) \times K[x]/(h)$, un producto que siempre tiene divisores de cero, luego (f) no es ni siquiera primo.
7. **Control final:** cuenta. Si A/I es finito y su cardinal es primo, es $\mathbb{Z}/p$ y por tanto cuerpo. En un anillo finito, además, *todo* ideal primo es maximal.

Ejemplo 25.77 (Seis ideales clasificados)**. (a) (0) y (5) en $\mathbb{Z}$.** El cociente $\mathbb{Z}/(0) = \mathbb{Z}$ es dominio pero no cuerpo: (0) es primo y no maximal. El cociente $\mathbb{Z}/(5) = \mathbb{F}_5$ es cuerpo: (5) es maximal (y primo). En cambio $\mathbb{Z}/(6)$ tiene divisores de cero, $\bar{2}\cdot\bar{3} = \bar{0}$, así que (6) no es primo.

(b) (x^2+1) en $\mathbb{R}[x]$. El polinomio no tiene raíces reales y tiene grado 2, luego es irreducible sobre $\mathbb{R}$; el ideal es maximal y

$$\mathbb{R}[x]/(x^2+1) \;\cong\; \mathbb{C},$$

donde la clase $\bar{x}$ hace de i. Ésta es, literalmente, la construcción algebraica de los números complejos.

(c) El mismo polinomio en $\mathbb{Q}[x]$. Sigue siendo irreducible, así que el ideal sigue siendo maximal, pero el cuerpo que sale es más pequeño:

$$\mathbb{Q}[x]/(x^2+1) \;\cong\; \mathbb{Q}(i) = \{a+bi : \; a,b\in\mathbb{Q}\},$$

que no es $\mathbb{C}$. El cociente depende del anillo de partida, no solo del polinomio.

(d) (x^2-1) en $\mathbb{Q}[x]$. Ahora $x^2-1 = (x-1)(x+1)$ con factores coprimos, y ninguno de los dos está en el ideal. Por el paso 6,

$$\mathbb{Q}[x]/(x^2-1) \;\cong\; \mathbb{Q}[x]/(x-1)\times\mathbb{Q}[x]/(x+1) \;\cong\; \mathbb{Q}\times\mathbb{Q},$$

que tiene divisores de cero: $(1,0)\cdot(0,1) = (0,0)$. El ideal no es primo.

(e) (x) en $\mathbb{Z}[x]$. Evaluar en 0 da un homomorfismo sobreyectivo $\mathbb{Z}[x]\to\mathbb{Z}$ de núcleo (x), luego

$$\mathbb{Z}[x]/(x) \;\cong\; \mathbb{Z}.$$

Como $\mathbb{Z}$ es dominio pero no cuerpo, (x) es primo y *no* maximal. Y no lo es porque cabe algo en medio: $(x)\subsetneq(2,x)\subsetneq\mathbb{Z}[x]$.

(f) $(2,x)$ en $\mathbb{Z}[x]$. Este ideal ya no es principal. Reducir los coeficientes módulo 2 y evaluar en 0 da $\mathbb{Z}[x]\to\mathbb{F}_2$, sobreyectivo y con núcleo exactamente $(2,x)$, de modo que

$$\mathbb{Z}[x]/(2,x) \;\cong\; \mathbb{F}_2,$$

un cuerpo: el ideal es maximal. Nótese la moraleja: en $\mathbb{Z}[x]$, que no es un dominio de ideales principales, los ideales maximales necesitan dos generadores.

Cuidado

No confundas "f no tiene raíces en K" con "f es irreducible sobre K". Vale para grados 2 y 3, y solo ahí. El polinomio $x^4+2x^2+1 = (x^2+1)^2$ no tiene raíces reales y sin embargo es reducible sobre $\mathbb{R}$; el ideal (x^4+2x^2+1) en $\mathbb{R}[x]$ no es ni primo, pues $x^2+1\notin(x^4+2x^2+1)$ por grados, pero su cuadrado sí está.

Idea clave

Los tres cocientes $\mathbb{R}[x]/(x^2+1)\cong\mathbb{C}$, $\mathbb{Q}[x]/(x^2+1)\cong\mathbb{Q}(i)$ y $\mathbb{F}_2[x]/(x^2+x+1)\cong\mathbb{F}_4$ se construyen exactamente igual: se toma un polinomio irreducible y se declara que su clase es una raíz. Ésta es la máquina universal de fabricar cuerpos, y es la que en el capítulo de Galois permite "añadir" raíces a voluntad.

Regla: usar el teorema chino del resto en $\mathbb{Z}$ y en $K[x]$

El teorema chino del resto es la herramienta que convierte *un* problema grande en *varios* problemas pequeños e independientes. En $\mathbb{Z}$ se lee como aritmética modular; en $K[x]$ se lee como interpolación. Es literalmente el mismo teorema.

1. **Comprueba la hipótesis de coprimalidad.** El enunciado

$$\mathbb{Z}/m_1 m_2 \cdots m_r \;\cong\; \mathbb{Z}/m_1 \times \mathbb{Z}/m_2 \times \cdots \times \mathbb{Z}/m_r$$

   como isomorfismo de *anillos* exige $\gcd(m_i, m_j) = 1$ para $i \neq j$, dos a dos, no solo que el conjunto entero tenga máximo común divisor 1. Si falla, el sistema puede no tener solución o tenerla no única.

2. **Escribe el sistema.** Un sistema de congruencias $x \equiv a_i \pmod{m_i}$, $i = 1, \ldots, r$, es exactamente pedir la preimagen del elemento $(a_1, \ldots, a_r)$ bajo el isomorfismo. Como el isomorfismo es biyectivo, la solución existe y es *única módulo* $M = m_1 \cdots m_r$.

3. **Construye la solución con los complementarios.** Pon $M_i = M/m_i$. Entonces $\gcd(M_i, m_i) = 1$, así que existe y_i con $M_i y_i \equiv 1 \pmod{m_i}$ (calcúlalo con el algoritmo de Euclides extendido). La solución es

$$x \;=\; \sum_{i=1}^{r} a_i \, M_i \, y_i \pmod{M}.$$

   El truco es que $M_i \equiv 0$ módulo todos los m_j con $j \neq i$: cada sumando "habla" en una sola congruencia y es mudo en las demás.

4. **Reduce al final, no al principio.** Suma los tres o cuatro productos completos y reduce una sola vez módulo M. Reducir a mitad de camino es donde se cometen los errores de signo.

5. **Úsalo para calcular potencias.** Para evaluar $a^N \bmod m$ con $m = \prod p_i^{k_i}$: calcula a^N módulo cada $p_i^{k_i}$ por separado (allí el teorema de Euler reduce el exponente módulo $\varphi(p_i^{k_i})$) y recombina. Cada pieza es diminuta comparada con el original.

6. **Traduce a $K[x]$.** Los módulos m_i se vuelven polinomios coprimos f_i, y $\gcd(f_i, f_j) = 1$. Con $f_i = x - c_i$ y puntos c_i distintos:

$$K[x]/((x-c_1)\cdots(x-c_r)) \;\cong\; K \times \cdots \times K, \qquad p \longmapsto (p(c_1), \ldots, p(c_r)),$$

   porque $K[x]/(x-c) \cong K$ por evaluación. La fórmula del paso 3 se convierte, término a término, en la base de Lagrange.

7. **Verifica siempre por sustitución directa.** Reduce tu x (o evalúa tu p) en cada módulo y comprueba que salen los a_i de partida. Es una comprobación de dos líneas y detecta cualquier inversa mal calculada.

Ejemplo 25.78 (Cinco cálculos con el teorema chino)**. (a) Un sistema de tres congruencias.** Resolvamos

$$x \equiv 3 \pmod 5, \qquad x \equiv 2 \pmod 7, \qquad x \equiv 6 \pmod{11}.$$

Los módulos $5, 7, 11$ son primos distintos, luego coprimos dos a dos: la hipótesis se cumple. Aquí $M = 5 \cdot 7 \cdot 11 = 385$ y los complementarios son $M_1 = 77$, $M_2 = 55$, $M_3 = 35$.

Las inversas: $77 \equiv 2 \pmod 5$ y $2 \cdot 3 = 6 \equiv 1$, luego $y_1 = 3$. $55 \equiv 6 \pmod 7$ y $6 \cdot 6 = 36 \equiv 1$, luego $y_2 = 6$. $35 \equiv 2 \pmod{11}$ y $2 \cdot 6 = 12 \equiv 1$, luego $y_3 = 6$. Por tanto

$$x = 3 \cdot 77 \cdot 3 + 2 \cdot 55 \cdot 6 + 6 \cdot 35 \cdot 6 = 693 + 660 + 1260 = 2613 \equiv 303 \pmod{385},$$

puesto que $2613 - 6 \cdot 385 = 2613 - 2310 = 303$. Verificación: $303 = 60 \cdot 5 + 3$, $303 = 43 \cdot 7 + 2$, $303 = 27 \cdot 11 + 6$. Las tres salen.

(b) El isomorfismo, con idempotentes. El isomorfismo $\mathbb{Z}/385 \cong \mathbb{Z}/5 \times \mathbb{Z}/7 \times \mathbb{Z}/11$ lleva 1 a $(1,1,1)$, y los tres idempotentes ortogonales de la derecha, $(1,0,0)$, $(0,1,0)$, $(0,0,1)$, provienen respectivamente de

$$e_1 = M_1 y_1 = 231, \qquad e_2 = M_2 y_2 = 330, \qquad e_3 = M_3 y_3 = 210 \pmod{385}.$$

Se comprueba que $231^2 \equiv 231$, $330^2 \equiv 330$ y $210^2 \equiv 210$ módulo 385: son idempotentes de verdad. Además $231 + 330 + 210 = 771 \equiv 1 \pmod{385}$, y sus residuos son $(1,0,0)$, $(0,1,0)$, $(0,0,1)$ como debía ser. La fórmula del paso 3 no es más que $x = \sum a_i e_i$.

Como el isomorfismo es de anillos, restringe a las unidades: $(\mathbb{Z}/385)^\times \cong (\mathbb{Z}/5)^\times \times (\mathbb{Z}/7)^\times \times (\mathbb{Z}/11)^\times$, y contando, $\varphi(385) = 4 \cdot 6 \cdot 10 = 240$, que es justamente lo que da la fórmula multiplicativa de la función de Euler. Esa multiplicatividad *es* el teorema chino.

(c) Una potencia enorme sin calcular la potencia. Queremos 7^{222} mód 11. Como 11 es primo y $11 \nmid 7$, el pequeño teorema de Fermat da $7^{10} \equiv 1$, así que el exponente se reduce módulo 10: $222 = 22 \cdot 10 + 2$, y

$$7^{222} \equiv 7^2 = 49 \equiv 5 \pmod{11}.$$

Un caso con módulo compuesto: 2^{1000} mód 1000. Factorizamos $1000 = 8 \cdot 125$ con $\gcd(8, 125) = 1$. Módulo 8: $2^{1000} \equiv 0$, porque $2^3 = 8$ ya es cero y $1000 \geq 3$. Módulo 125: $\varphi(125) = 100$ y $\gcd(2, 125) = 1$, luego $2^{100} \equiv 1$ y $2^{1000} = (2^{100})^{10} \equiv 1$. Recombinando, buscamos $x \equiv 0 \pmod 8$ y $x \equiv 1 \pmod{125}$: los múltiplos de 8 son $8, 16, \ldots$ y $376 = 47 \cdot 8$ cumple $376 = 3 \cdot 125 + 1$. Así

$$2^{1000} \equiv 376 \pmod{1000},$$

es decir, las tres últimas cifras de 2^{1000} son 376. Nótese que no hemos escrito en ningún momento el número 2^{1000}, que tiene 302 cifras.

(d) Las inversas, con Euclides extendido. Para invertir 17 módulo 101: $101 = 5 \cdot 17 + 16$, $17 = 1 \cdot 16 + 1$, luego $1 = 17 - 16 = 17 - (101 - 5 \cdot 17) = 6 \cdot 17 - 101$. Por tanto $17^{-1} \equiv 6 \pmod{101}$, y en efecto $17 \cdot 6 = 102 = 101 + 1$. El mismo algoritmo da $3^{-1} \equiv 67 \pmod{100}$, pues $3 \cdot 67 = 201 = 2 \cdot 100 + 1$. Y para $\gcd(1234, 5678)$ el algoritmo devuelve $1234 \cdot 704 + 5678 \cdot (-153) = 868\,736 - 868\,734 = 2$: el máximo común divisor es 2, no 1, de modo que 1234 *no* es invertible módulo 5678. La coprimalidad del paso 1 no es decorativa.

(e) El mismo teorema en $K[x]$: interpolación. Busquemos $p \in \mathbb{Q}[x]$ con

$$p \equiv 1 \pmod{x-1}, \qquad p \equiv 3 \pmod{x-2}, \qquad p \equiv 2 \pmod{x-4},$$

que por la identificación $K[x]/(x-c) \cong K$ significa simplemente $p(1)=1$, $p(2)=3$, $p(4)=2$. Los módulos $x-1$, $x-2$, $x-4$ son coprimos dos a dos porque las raíces son distintas. Los complementarios y sus inversas producen la base de Lagrange

$$L_1 = \tfrac{x^2}{3} - 2x + \tfrac{8}{3}, \qquad L_2 = -\tfrac{x^2}{2} + \tfrac{5x}{2} - 2, \qquad L_3 = \tfrac{x^2}{6} - \tfrac{x}{2} + \tfrac{1}{3},$$

y la solución es $p = 1\cdot L_1 + 3\cdot L_2 + 2\cdot L_3$, es decir

$$p(x) = -\tfrac{5}{6}x^2 + \tfrac{9}{2}x - \tfrac{8}{3}.$$

Verificación: $p(1) = -\frac{5}{6}+\frac{9}{2}-\frac{8}{3} = \frac{-5+27-16}{6} = \frac{6}{6} = 1$; $p(2) = -\frac{10}{3}+9-\frac{8}{3} = \frac{-20+54-16}{6} = \frac{18}{6} = 3$; $p(4) = -\frac{40}{3}+18-\frac{8}{3} = \frac{-80+108-16}{6} = \frac{12}{6} = 2$. Y la unicidad "módulo M" del paso 2 se lee aquí como: p es único módulo $(x-1)(x-2)(x-4)$, o sea, único entre los polinomios de grado ≤ 2. Ese es el teorema de existencia y unicidad de la interpolación polinómica, demostrado sin tocar un determinante de Vandermonde.

Cuidado

La coprimalidad tiene que ser *dos a dos*. El sistema $x \equiv 1 \pmod 4$, $x \equiv 3 \pmod 6$ tiene módulos con $\gcd(4,6) = 2$: la primera congruencia obliga a x impar y la segunda también, y de hecho aquí sí hay solución ($x \equiv 9 \pmod{12}$), pero $x \equiv 1 \pmod 4$ junto con $x \equiv 2 \pmod 6$ es incompatible, porque la primera pide x impar y la segunda x par. Cuando los módulos no son coprimos hay que comprobar a mano la compatibilidad en cada primo común, y si hay solución ya no es única módulo el producto sino módulo el mínimo común múltiplo. En $K[x]$ el fenómeno análogo es pedir valores en un punto repetido: $(x-1)$ y $(x-1)$ no son coprimos, y lo que aparece entonces es la interpolación de Hermite, con derivadas.

Idea clave

Un anillo se "rompe" en un producto exactamente cuando tiene idempotentes no triviales, y el teorema chino es la forma canónica de fabricarlos. En $\mathbb{Z}/385$ encontramos tres; en $\mathbb{Z}/p^k$ no hay ninguno distinto de 0 y 1, y por eso los $\mathbb{Z}/p^k$ son las piezas atómicas: no se descomponen más. La misma dicotomía reaparece por todo el álgebra: un espacio se descompone en suma directa de subespacios invariantes cuando el polinomio mínimo factoriza en factores coprimos —y la descomposición primaria de un módulo, el teorema de estructura de los grupos abelianos finitos y la forma de Jordan del capítulo 7 son, todos, el teorema chino del resto aplicado en anillos distintos.

25.11. Polinomios

Idea clave

Intuición primero. $K[x]$ es la copia más fiel de $\mathbb{Z}$ que existe fuera de $\mathbb{Z}$. Donde en los enteros se divide dejando un resto *más pequeño*, en polinomios se divide dejando un resto *de menor grado*; el grado hace de valor absoluto. Y en cuanto hay división con resto, todo lo demás se hereda en cadena: algoritmo de Euclides, identidad de Bézout, ideales generados por un solo elemento, factorización única. Por eso $K[x]$ es euclídeo, y por eso conviene leer este apartado con el capítulo 3 al lado: son el mismo argumento dos veces.
El diccionario completo es este. Los primos de $\mathbb{Z}$ son los polinomios irreducibles de $K[x]$; las unidades (± 1) son ahora las constantes no nulas; "a divide a b" se comprueba igual; y el teorema del resto —que $f(a)$ es el resto de dividir por $x-a$— dice que **buscar raíces es buscar factores lineales**. Un aviso que la práctica olvida: ser irreducible depende del cuerpo. x^2+1 es irreducible sobre $\mathbb{R}$ y se parte sobre $\mathbb{C}$; x^2-2 es irreducible sobre $\mathbb{Q}$ y se parte sobre $\mathbb{R}$. Esa dependencia del terreno es exactamente lo que el capítulo 26 convertirá en una teoría.

$K[x]$ (K cuerpo) es el anillo protagonista. Su aritmética es paralela a la de $\mathbb{Z}$:

Teorema 25.79 (División y consecuencias)**.** *Para $f, g \in K[x]$, $g \neq 0$, existen únicos q, r con $f = qg + r$, $\deg r < \deg g$. De ahí:*

(i) Resto*: $f(a)$ es el resto de dividir por $(x-a)$; a es raíz $\iff$ $(x-a) \mid f$.*

(ii) Un polinomio de grado n tiene a lo sumo n raíces en K.

(iii) $K[x]$ es DIP: todo ideal es (d) con d el polinomio mónico de menor grado; existen gcd *y la identidad de Bézout* $\gcd(f,g) = uf + vg$ *(algoritmo de Euclides).*

(iv) $K[x]$ es DFU: factorización única en irreducibles.

Idea clave

Qué dice esto en cristiano. Dividir polinomios funciona igual que dividir números enteros: hay un cociente y un resto, el resto es más pequeño que el divisor (aquí "más pequeño" significa de grado menor) y ambos son únicos. Toda la aritmética de $\mathbb{Z}$ se traslada mecánicamente: máximo común divisor por Euclides, identidad de Bézout, factorización única. Y la consecuencia más usada de todas es el teorema del resto, que convierte la pregunta "¿es a raíz de f?" en la cuenta trivial de evaluar $f(a)$.

Ejemplo 25.80 (Dividir, evaluar y factorizar)**. (a) Una división completa.** Con $f = x^4 - 3x^3 + 2x - 5$ y $g = x^2 - 1$ se obtiene $q = x^2 - 3x + 1$ y $r = -x - 4$. Comprobación: $q\,g + r = (x^2 - 3x + 1)(x^2 - 1) + (-x - 4)$, que desarrollado da exactamente f. Y el grado del resto, 1, es menor que el del divisor, 2, como exige el teorema.

(b) El teorema del resto. Dividiendo el mismo f entre $x-a$, el resto es siempre el número $f(a)$: para $a=1$ da -5, para $a=2$ da -9 y para $a=-1$ da -3. Los tres coinciden con evaluar directamente. En particular, a es raíz exactamente cuando el resto es cero, que es el fundamento de la regla de Ruffini.

(c) Raíces y factores. Para $h=x^3-6x^2+11x-6$, las raíces son 1, 2 y 3, y en consecuencia $h=(x-1)(x-2)(x-3)$: cada raíz aporta un factor lineal, y como el grado es 3 no puede haber una cuarta. Ese tope es justamente la consecuencia (iii) del teorema.

(d) Euclides y Bézout con polinomios. Con $u=x^4-1$ y $v=x^3-x$, el algoritmo de Euclides da $\gcd(u,v)=x^2-1$, y la identidad de Bézout resulta ser $1\cdot u+(-x)\cdot v=x^2-1$. Comprobación: $x^4-1-x(x^3-x)=x^4-1-x^4+x^2=x^2-1$. Es literalmente el mismo algoritmo que con números, cambiando "menor valor absoluto" por "menor grado".

Cuidado

[Contraejemplo: hace falta que K sea un cuerpo] La división con resto necesita poder dividir entre el coeficiente principal de g. En $\mathbb{Z}[x]$ eso falla: no se puede dividir x^2 entre $2x+1$ dentro de los enteros. Por eso el teorema se enuncia sobre $K[x]$ con K cuerpo, y por eso $\mathbb{Z}[x]$ no es un dominio de ideales principales aunque $\mathbb{Z}$ y $K[x]$ lo sean por separado. Segundo aviso: la cota "un polinomio de grado n tiene a lo sumo n raíces" también exige cuerpo. En $\mathbb{Z}_8$, el polinomio x^2-1 tiene *cuatro* raíces, $1,3,5$ y 7, porque hay divisores de cero y la factorización $(x-1)(x+1)=0$ deja de implicar que uno de los factores se anule.

25.11.1. Irreducibilidad

Cuidado

"Irreducible" depende del cuerpo: x^2-2 es irreducible en $\mathbb{Q}[x]$, se parte en $\mathbb{R}[x]$; x^2+1 es irreducible en $\mathbb{R}[x]$, se parte en $\mathbb{C}[x]$. Sobre $\mathbb{C}$ los irreducibles son los lineales (teorema fundamental del álgebra); sobre $\mathbb{R}$, lineales y cuadráticos sin raíces reales; sobre $\mathbb{Q}$, de cualquier grado.

Herramientas para $\mathbb{Q}[x]$ (y $\mathbb{Z}[x]$):

(1) *Raíces racionales*: si $\frac{p}{q}$ (irreducible) es raíz de $c_nx^n+\cdots+c_0\in\mathbb{Z}[x]$, entonces $p\mid c_0$ y $q\mid c_n$. Para grados 2 y 3, no tener raíces racionales equivale a ser irreducible sobre $\mathbb{Q}$.

(2) *Lema de Gauss*: un polinomio con coeficientes enteros que factoriza sobre $\mathbb{Q}$ factoriza ya sobre $\mathbb{Z}$. (Permite trabajar módulo primos.)

(3) *Reducción módulo p*: si $f \bmod p$ es irreducible en $\mathbb{F}_p[x]$ (y el grado no cae), f es irreducible sobre $\mathbb{Q}$.

(4) *Criterio de Eisenstein*: si un primo p divide a todos los coeficientes menos el líder, y p^2 no divide al término independiente, f es irreducible sobre $\mathbb{Q}$. Ejemplo estrella: x^n-2 es irreducible para todo n (con $p=2$), lo que da extensiones de $\mathbb{Q}$ de cualquier grado.

(5) El truco del *desplazamiento*: $f(x)$ es irreducible $\iff$ $f(x+c)$ lo es. Con él, el *polinomio ciclotómico* $\Phi_p(x) = x^{p-1}+\cdots+x+1 = \frac{x^p-1}{x-1}$ (p primo) es irreducible: $\Phi_p(x+1) = x^{p-1} + \binom{p}{1}x^{p-2} + \cdots + \binom{p}{p-1}$ es Eisenstein en p.

Teorema 25.81. *Si $p(x)$ es irreducible en $K[x]$, el cociente $K[x]/(p(x))$ es un **cuerpo** que contiene a K, y la clase $\alpha = \overline{x}$ satisface $p(\alpha) = 0$. Sus elementos se representan por polinomios de grado $< \deg p$, con la aritmética módulo $p(x)$.*

Demostración. (p) es maximal en el DIP $K[x]$ (un ideal $(d) \supseteq (p)$ obliga $d \mid p$: d es unidad o asociado de p); por el Teorema 25.74, el cociente es cuerpo. Concretamente, el inverso de $\overline{f} \neq 0$ se calcula con Bézout: $\gcd(f,p) = 1 = uf + vp \Rightarrow \overline{u} = \overline{f}^{\,-1}$. □

Idea clave

Qué dice esto en cristiano. Esta es la máquina de fabricar cuerpos, y es la misma idea que se usó para inventar los números complejos. Si un polinomio no tiene raíz donde estás, te inventas una: declaras que α es una raíz, trabajas con expresiones en α y cada vez que aparece una potencia demasiado alta la reduces usando la relación $p(\alpha) = 0$. El cociente $K[x]/(p(x))$ es exactamente eso escrito con rigor, y el teorema garantiza que el resultado es un cuerpo de verdad, con divisiones y todo, siempre que p sea irreducible. Toda la teoría de Galois se construye repitiendo esta operación.

Ejemplo 25.82 (Tres cuerpos fabricados con la misma receta)**. (a) Los números complejos.** Tomando $K = \mathbb{R}$ y $p(x) = x^2+1$, que no tiene raíces reales, el cociente $\mathbb{R}[x]/(x^2+1)$ es $\mathbb{C}$, y la clase de x es i. Multiplicar $(1+2x)(3+4x)$ da $8x^2+10x+3$; reduciendo módulo x^2+1, es decir sustituyendo x^2 por -1, queda $10x-5$. Y en efecto $(1+2\mathrm{i})(3+4\mathrm{i}) = -5+10\mathrm{i}$.

(b) Un cuerpo con cuatro elementos. Con $K = \mathbb{F}_2$ y $p(t) = t^2+t+1$, que no tiene raíces en $\mathbb{F}_2$ porque $p(0) = p(1) = 1$, el cociente tiene los cuatro elementos $0, 1, t, t+1$. La regla de reducción es $t^2 = t+1$, de donde $t^3 = t\cdot t^2 = t^2+t = (t+1)+t = 1$. Ese es el $\mathbb{F}_4$ del apartado anterior, construido ahora explícitamente.

(c) Dividir en $\mathbb{Q}(\sqrt{2})$. Con $K = \mathbb{Q}$ y $p(x) = x^2-2$, el cociente es $\mathbb{Q}(\sqrt{2})$. El inverso de $1+x$ es $x-1$, porque $(1+x)(x-1) = x^2-1 = 2-1 = 1$. Traducido a números: $\frac{1}{1+\sqrt{2}} = \sqrt{2}-1$, que es justo lo que sale al racionalizar el denominador. La racionalización del colegio y el cálculo de inversos en un cociente son la misma operación.

(d) El tamaño del cuerpo nuevo. Si $\deg p = n$, todo elemento se escribe de forma única como $c_0 + c_1\alpha + \cdots + c_{n-1}\alpha^{n-1}$: hay n coeficientes libres, luego la extensión tiene grado n sobre K. Con $K = \mathbb{F}_2$ y $n = 2$ salen $2^2 = 4$ elementos, y con $n = 3$ saldrían 8. Esa es la razón de que todo cuerpo finito tenga p^n elementos.

Cuidado

[Contraejemplo: si p no es irreducible, no sale un cuerpo] La hipótesis de irreducibilidad no es un adorno. En $\mathbb{F}_2[t]$ el polinomio t^2+1 *no* es irreducible, porque $t^2+1 = (t+1)^2$ (recuérdese que $2 = 0$). En el cociente $\mathbb{F}_2[t]/(t^2+1)$, la clase de $t+1$ es no nula y sin

embargo su cuadrado es cero: hay un divisor de cero, y por tanto ese elemento no tiene inverso. El cociente es un anillo perfectamente respetable, pero no un cuerpo. La regla general es: $K[x]/(p)$ es cuerpo si y solo si p es irreducible, y es un dominio si y solo si p es irreducible también, porque en $K[x]$ los ideales primos no nulos y los maximales coinciden.

Idea clave

Este teorema es la *máquina de fabricar cuerpos*, el corazón técnico del capítulo 26. $\mathbb{R}[x]/(x^2+1)$ es un cuerpo donde $\overline{x}^2 = -1$: acabamos de *construir* $\mathbb{C}$ sin "imaginar" nada. $\mathbb{Q}[x]/(x^2-2) \cong \mathbb{Q}(\sqrt{2})$. $\mathbb{F}_2[x]/(x^2+x+1)$ es el cuerpo de 4 elementos. Adjuntar una raíz = cocientar por el irreducible que la define.

Regla: decidir si un polinomio es irreducible sobre $\mathbb{Q}$

No hay un algoritmo único que se aplique siempre con elegancia; hay una batería de criterios, y el oficio consiste en probarlos en el orden correcto, del más barato al más caro.

1. **Limpia primero.** Saca el contenido (el máximo común divisor de los coeficientes) y quítate los denominadores multiplicando por el mínimo común múltiplo. Por el lema de Gauss, un polinomio de $\mathbb{Z}[x]$ primitivo es irreducible en $\mathbb{Q}[x]$ si y solo si lo es en $\mathbb{Z}[x]$: se puede trabajar entero, que es mucho más cómodo.

2. **Si el grado es 2 o 3, basta buscar raíces.** Un polinomio de grado ≤ 3 que factoriza tiene forzosamente un factor de grado 1, luego una raíz. Por el teorema de la raíz racional, los candidatos de $a_n x^n + \cdots + a_0$ son p/q con $p \mid a_0$ y $q \mid a_n$; son cuatro o seis números, se prueban a mano y se acabó.

3. **Ojo: en grado ≥ 4 eso ya no vale.** Puede no haber ninguna raíz racional y aun así factorizar en dos cuadráticos. Este es el error más frecuente del tema.

4. **Prueba Eisenstein.** Si existe un primo p con $p \nmid a_n$, $p \mid a_i$ para todo $i < n$ y $p^2 \nmid a_0$, el polinomio es irreducible. Es el criterio más potente cuando se aplica.

5. **Si Eisenstein falla, desplaza.** La irreducibilidad no cambia al sustituir $x \mapsto x+c$ (el cambio es un automorfismo de $\mathbb{Q}[x]$). Prueba $x+1$ y $x-1$ antes de rendirte: muchos polinomios famosos solo revelan su Eisenstein tras el desplazamiento.

6. **Reduce módulo un primo pequeño.** Si $p \nmid a_n$ y la reducción $\bar{f} \in \mathbb{F}_p[x]$ es irreducible, entonces f es irreducible sobre $\mathbb{Q}$. Prueba $p = 2, 3, 5$. Es barato porque en $\mathbb{F}_p[x]$ solo hay finitos polinomios que ensayar.

7. **Cuidado con la dirección de este último criterio.** Solo funciona en un sentido. Que $\bar{f}$ factorice *no* dice nada: hay polinomios irreducibles sobre $\mathbb{Q}$ que se rompen módulo todos los primos. Lo que sí se puede aprovechar es la información de grados: si $\bar{f}$ factoriza en irreducibles de grados 2 y 3 módulo un primo y en grados 1

y 4 módulo otro, ninguna factorización entera es compatible con ambas y f es irreducible.

Ejemplo 25.83 (Cinco polinomios, cinco criterios distintos)**. (a)** $f = x^3 + x + 1$**: raíz racional.** Grado 3, coeficientes $a_0 = a_3 = 1$, luego los únicos candidatos a raíz son ± 1. Se prueban: $f(1) = 3$ y $f(-1) = -1$. Ninguna raíz, y en grado 3 eso basta. *Irreducible.*

(b) $f = x^4 + x^3 + x^2 + x + 1$**: Eisenstein tras desplazar.** Directamente Eisenstein no se aplica: no hay primo que divida a $1, 1, 1, 1$. Pero sustituyendo $x \mapsto x + 1$, y usando que $f(x) = (x^5 - 1)/(x - 1)$,

$$f(x+1) = \frac{(x+1)^5 - 1}{x} = x^4 + 5x^3 + 10x^2 + 10x + 5,$$

y ahora $p = 5$ divide a $5, 10, 10, 5$, no divide al coeficiente principal 1, y $25 \nmid 5$. Eisenstein se aplica. *Irreducible.* (Es el polinomio ciclotómico Φ_5, y este truco funciona igual para todo Φ_p.)

(c) $f = x^4 + 1$**: el mismo truco, con** $p = 2$**.** Otra vez Eisenstein falla en crudo, pero

$$f(x+1) = (x+1)^4 + 1 = x^4 + 4x^3 + 6x^2 + 4x + 2,$$

y $p = 2$ divide a $4, 6, 4, 2$ mientras que $4 \nmid 2$. *Irreducible sobre* $\mathbb{Q}$. Este ejemplo es además la ilustración clásica del aviso del paso 7: módulo 2 se tiene $f \equiv (x+1)^4$, módulo 3 se tiene $f \equiv (x^2 + x - 1)(x^2 - x - 1)$ y módulo 5 se tiene $f \equiv (x^2 + 2)(x^2 - 2)$. De hecho $x^4 + 1$ factoriza módulo *todo* primo, y sin embargo es irreducible sobre $\mathbb{Q}$. El criterio de reducción, mal leído, aquí te haría concluir lo contrario.

(d) $f = x^5 - x - 1$**: reducción módulo 5.** Los candidatos racionales ± 1 no son raíces ($f(1) = -1$, $f(-1) = -1$), pero en grado 5 eso no cierra nada. Eisenstein no se aplica ni tras desplazar. Reducimos: módulo 2 el polinomio factoriza como $(x^2 + x + 1)(x^3 + x^2 + 1)$, así que ese primo no sirve; pero *módulo 5 es irreducible*, y con eso basta. *Irreducible sobre* $\mathbb{Q}$.

(e) $f = x^4 + 4$**: la trampa.** No tiene raíces racionales: $f(\pm 1) = 5$, $f(\pm 2) = 20$, y ningún divisor de 4 anula. Un lector apresurado concluiría irreducibilidad. Pero

$$x^4 + 4 = (x^2 - 2x + 2)(x^2 + 2x + 2),$$

como se comprueba multiplicando. *Reducible.* Es un caso de la identidad de Sophie Germain $a^4 + 4b^4 = (a^2 - 2ab + 2b^2)(a^2 + 2ab + 2b^2)$. Guárdese como recordatorio permanente de que "sin raíces" y "irreducible" son cosas distintas en cuanto el grado pasa de 3.

25.12. Cuerpos

Idea clave

Intuición primero. Un cuerpo es el sitio donde por fin se puede dividir entre cualquier cosa que no sea cero: es un anillo sin excusas. $\mathbb{Q}$, $\mathbb{R}$, $\mathbb{C}$ y $\mathbb{F}_p$ son los ejemplos de cabecera, y toda el álgebra lineal del capítulo 7 se hizo, sin decirlo, sobre un cuerpo — porque despejar coeficientes exige dividir.

La *característica* responde a una pregunta ingenua: ¿qué pasa si sumo 1 consigo mismo muchas veces? En $\mathbb{Q}$ nunca llegas a cero y la característica es 0; en $\mathbb{F}_5$ llegas a los cinco pasos. Y solo puede ser cero o un primo, porque si $n = ab$ diera cero, $a \cdot 1$ y $b \cdot 1$ serían divisores de cero, y en un cuerpo no los hay. De esa observación tan pequeña sale el "sueño del estudiante" $(a+b)^p = a^p + b^p$ en característica p: no es un error de principiante ahí, es un teorema, porque los coeficientes binomiales de en medio son múltiplos de p y se anulan. El automorfismo de Frobenius que eso define es la pieza que hace funcionar los cuerpos finitos, y con ellos los códigos correctores de tu tarjeta y de los códigos QR.

Definición 25.84. Un *cuerpo* es un anillo conmutativo con $1 \neq 0$ donde todo elemento no nulo es unidad. La *característica* $\operatorname{car} K$ es el menor n con $n \cdot 1 = 0$ (o 0 si no existe); siempre es 0 o un primo (si $n = ab$, $(a \cdot 1)(b \cdot 1) = 0$ y no hay divisores de cero). Todo cuerpo contiene un *cuerpo primo*: $\mathbb{Q}$ si $\operatorname{car} = 0$, $\mathbb{F}_p$ si $\operatorname{car} = p$.

Idea clave

Qué dice esto en cristiano. Un cuerpo es un sitio donde se puede sumar, restar, multiplicar *y dividir* entre cualquier cosa que no sea cero: es decir, donde valen todas las reglas del colegio. La característica responde a una pregunta rara pero decisiva: ¿si voy sumando $1 + 1 + 1 + \cdots$, llego alguna vez a cero? En $\mathbb{Q}$ o $\mathbb{R}$ nunca, y se dice que la característica es 0. En $\mathbb{F}_5$ sí, a los cinco sumandos, y la característica es 5. Que siempre salga 0 o un primo es un teorema pequeño con una consecuencia grande: cada cuerpo lleva dentro una copia de $\mathbb{Q}$ o una de $\mathbb{F}_p$, y ya no hay más opciones de arranque.

Ejemplo 25.85 (Cuerpos y no cuerpos, comprobados)**. (a) $\mathbb{F}_5$ sí lo es.** Los inversos son $1^{-1} = 1$, $2^{-1} = 3$ (porque $2 \cdot 3 = 6 \equiv 1$), $3^{-1} = 2$ y $4^{-1} = 4$ (porque $16 \equiv 1$). Todo elemento no nulo tiene inverso, luego se puede dividir siempre: $\frac{3}{4}$ en $\mathbb{F}_5$ significa $3 \cdot 4^{-1} = 3 \cdot 4 = 12 \equiv 2$.

(b) $\mathbb{Z}_6$ no lo es. Aquí $2 \cdot 3 = 6 \equiv 0$ con ambos factores no nulos: hay divisores de cero. Las únicas unidades son 1 y 5, así que 2, 3 y 4 no tienen inverso y no se puede dividir entre ellos. La regla general es que $\mathbb{Z}_n$ es cuerpo exactamente cuando n es primo.

(c) La característica en acción. En $\mathbb{F}_5$, $1 + 1 + 1 + 1 + 1 = 5 \equiv 0$, luego $\operatorname{car} \mathbb{F}_5 = 5$. En característica 2 ocurre el llamado sueño del principiante: $(a+b)^2 = a^2 + 2ab + b^2 = a^2 + b^2$, porque el 2 del término central *es* cero. Análogamente en característica 3 se tiene $(a + b)^3 = a^3 + b^3$, ya que los coeficientes 3 de $3a^2b + 3ab^2$ se anulan.

(d) Característica y cardinal no son lo mismo. El cuerpo $\mathbb{F}_4 = \mathbb{F}_2[t]/(t^2 + t + 1)$ tiene cuatro elementos $\{0, 1, t, t + 1\}$, pero su característica es 2, no 4: sumar $1 + 1$ ya da cero. Sus reglas son $t^2 = t + 1$ y $t^3 = t \cdot t^2 = t^2 + t = 1$, de modo que $\mathbb{F}_4^* = \{1, t, t^2\}$ es cíclico de orden 3. Nótese además que $\mathbb{F}_4$ *no* es $\mathbb{Z}_4$: en $\mathbb{Z}_4$ el elemento 2 no tiene inverso, mientras que en $\mathbb{F}_4$ todo elemento no nulo lo tiene.

Cuidado

[Contraejemplo: qué se rompe si se quita una condición] Si se quita la conmutatividad quedan los *anillos de división*, como los cuaterniones de Hamilton, donde se puede dividir pero $ij = -ji$; no son cuerpos. Si se quita la existencia de inversos queda un dominio como $\mathbb{Z}$, donde 2 no tiene inverso. Y si se admite $1 = 0$ queda el anillo nulo, con un solo elemento, que se excluye expresamente porque haría trivial toda la teoría. Por último, cuidado con la característica 0: no significa que 1 sea cero, sino todo lo contrario, que ningún múltiplo de 1 llega nunca a cero.

Teorema 25.86 (Cuerpos finitos)**.** *Todo cuerpo finito tiene p^n elementos ($p = \mathrm{car}$), y para cada p^n existe exactamente uno salvo isomorfismo, $\mathbb{F}_{p^n}$, construible como cociente de $\mathbb{F}_p[x]$ por un irreducible de grado n. Su grupo multiplicativo $\mathbb{F}_q^*$ es* cíclico*. En característica p vale el "sueño del estudiante": $(a+b)^p = a^p + b^p$ (los binomiales intermedios son divisibles por p), y $x \mapsto x^p$ es un automorfismo (*Frobenius*).*

Los cuerpos finitos no son curiosidades: sostienen la criptografía (RSA, curvas elípticas), los códigos correctores (Reed–Solomon, los de los QR y los CD) y buena parte de la combinatoria.

Regla: construir y calcular en un cuerpo finito $\mathbb{F}_{p^n}$

Los cuerpos finitos no se "descubren" uno a uno: se fabrican con una receta fija, y para cada tamaño posible hay uno y solo uno salvo isomorfismo.

1. **Comprueba primero que el tamaño es legal.** Existe un cuerpo con q elementos si y solo si $q = p^n$ con p primo y $n \geq 1$, y en ese caso es único salvo isomorfismo. No hay cuerpos de 6, 10 ni 12 elementos.

2. **Elige un polinomio irreducible** $f \in \mathbb{F}_p[x]$ de grado exactamente n. Para grados 2 y 3 basta comprobar que f no tiene raíces en $\mathbb{F}_p$: se prueban los p valores y ya está.

3. **Define** $\mathbb{F}_{p^n} = \mathbb{F}_p[x]/(f)$**.** Sus elementos son los restos, es decir los polinomios de grado $< n$: hay p opciones para cada uno de los n coeficientes, y de ahí salen los p^n elementos.

4. **Suma coeficiente a coeficiente módulo** p**.** La suma es la de un espacio vectorial $\mathbb{F}_p^n$: nada más.

5. **Multiplica como polinomios y reduce con** $f(\alpha) = 0$**,** donde $\alpha = \bar{x}$. Esa relación es la única regla nueva: convierte α^n en una combinación de potencias menores.

6. **Construye la tabla de potencias de un generador.** El grupo multiplicativo $\mathbb{F}_{p^n}^\times$ es cíclico de orden $p^n - 1$; si α lo genera se le llama *primitivo*, y entonces multiplicar es sumar exponentes e invertir es cambiar α^k por α^{p^n-1-k}.

7. **Controles:** todo elemento cumple $a^{p^n} = a$; la aplicación de Frobenius $a \mapsto a^p$ es un automorfismo; la característica es p, de modo que $p \cdot a = 0$ para todo a; y $\mathbb{F}_{p^m}$ es subcuerpo de $\mathbb{F}_{p^n}$ exactamente cuando $m \mid n$.

Ejemplo 25.87 (El cuerpo de ocho elementos, en detalle)**.** Sobre $\mathbb{F}_2$ los polinomios mónicos irreducibles de grado 3 son exactamente dos, x^3+x+1 y x^3+x^2+1 (el recuento general da $(2^3-2)/3 = 2$). Tomamos $f = x^3+x+1$: no tiene raíces, pues $f(0) = 1$ y $f(1) = 1+1+1 = 1$.

Ponemos $\mathbb{F}_8 = \mathbb{F}_2[x]/(x^3+x+1)$ y $\alpha = \bar{x}$, de forma que la regla de reducción es

$$\alpha^3 = \alpha + 1$$

(recuérdese que en característica 2 restar es sumar). Los ocho elementos son $c_0 + c_1\alpha + c_2\alpha^2$ con $c_i \in \{0,1\}$. Calculando potencias sucesivas y reduciendo cada vez:

potencia	forma reducida	(c_0, c_1, c_2)
α^1	α	$(0,1,0)$
α^2	α^2	$(0,0,1)$
α^3	$1+\alpha$	$(1,1,0)$
α^4	$\alpha+\alpha^2$	$(0,1,1)$
α^5	$1+\alpha+\alpha^2$	$(1,1,1)$
α^6	$1+\alpha^2$	$(1,0,1)$
α^7	1	$(1,0,0)$

Las siete potencias recorren los siete elementos no nulos, así que α es primitivo y $\mathbb{F}_8^\times \cong \mathbb{Z}/7$. Con la tabla, toda la aritmética se vuelve mecánica:

$$(1+\alpha)(\alpha+\alpha^2) = \alpha^3 \cdot \alpha^4 = \alpha^7 = 1 \quad \Longrightarrow \quad (1+\alpha)^{-1} = \alpha + \alpha^2.$$

Y en efecto, desarrollando a mano: $(1+\alpha)(\alpha+\alpha^2) = \alpha+\alpha^2+\alpha^2+\alpha^3 = \alpha+\alpha^3 = \alpha+\alpha+1 = 1$.

Obsérvese también que 7 es primo: *todo* elemento distinto de 0 y 1 es primitivo. Y como $2 \nmid 3$, el cuerpo $\mathbb{F}_4$ *no* es subcuerpo de $\mathbb{F}_8$; los únicos subcuerpos de $\mathbb{F}_8$ son $\mathbb{F}_2$ y él mismo.

Ejemplo 25.88 (El cuerpo de nueve elementos)**.** Sobre $\mathbb{F}_3$ hay tres polinomios mónicos irreducibles de grado 2, a saber x^2+1, x^2+x+2 y x^2+2x+2; el recuento general da $(3^2-3)/2 = 3$. Elegimos $f = x^2+1$, que no tiene raíces porque $0^2+1 = 1$, $1^2+1 = 2$ y $2^2+1 = 5 \equiv 2$.

Con $\mathbb{F}_9 = \mathbb{F}_3[x]/(x^2+1)$ y $\alpha = \bar{x}$ la regla de reducción es $\alpha^2 = -1 = 2$. Los nueve elementos son $a + b\alpha$ con $a, b \in \{0,1,2\}$: son los "complejos módulo 3".

Aquí α *no* es primitivo: $\alpha^2 = 2$, $\alpha^3 = 2\alpha$ y $\alpha^4 = 4 = 1$, luego su orden es 4 y no 8. Probamos $g = 1+\alpha$:

potencia	valor	potencia	valor
g^1	$1+\alpha$	g^5	$2+2\alpha$
g^2	2α	g^6	α
g^3	$1+2\alpha$	g^7	$2+\alpha$
g^4	2	g^8	1

Las ocho potencias son distintas, luego $g = 1 + \alpha$ es primitivo y $\mathbb{F}_9^\times \cong \mathbb{Z}/8$. La primera fila se comprueba de un vistazo: $(1+\alpha)^2 = 1 + 2\alpha + \alpha^2 = 1 + 2\alpha - 1 = 2\alpha$, y $g^4 = (2\alpha)^2 = 4\alpha^2 = 1 \cdot (-1) = 2$, de donde $g^8 = 2^2 = 4 = 1$.

Como control del paso 7, el Frobenius $\varphi(a) = a^3$ actúa aquí como una "conjugación": $\varphi(\alpha) = \alpha^3 = 2\alpha = -\alpha$, exactamente igual que $i \mapsto -i$ en $\mathbb{C}$. Y φ deja fijos precisamente los elementos de $\mathbb{F}_3$, que son las raíces de $x^3 - x$.

Cuidado

$\mathbb{F}_9$ *no* es $\mathbb{Z}/9$. El anillo $\mathbb{Z}/9$ tiene divisores de cero ($\bar{3} \cdot \bar{3} = \bar{0}$) y por tanto no es un cuerpo. En general $\mathbb{Z}/p^n$ solo es cuerpo cuando $n = 1$: para $n \geq 2$ hay que pasar por el cociente de polinomios. La confusión entre ambos es el error más frecuente del tema.

Pausa para pensar

¿Por qué el resultado no depende del polinomio irreducible elegido? Si se toma $f = x^2 + x + 2$ en vez de $x^2 + 1$ se obtiene otro cuerpo de nueve elementos, con otras tablas; pero ambos son isomorfos, porque el cuerpo de p^n elementos es único salvo isomorfismo: es el cuerpo de descomposición de $x^{p^n} - x$ sobre $\mathbb{F}_p$, y los cuerpos de descomposición son únicos. Cambiar de f es como cambiar de base en un espacio vectorial: cambian las coordenadas, no el objeto.

25.13. Fundamentos: las demostraciones estructurales del álgebra

25.13.1. La ecuación de clases y sus frutos

Definición 25.89. G actúa sobre sí mismo por *conjugación*: $g \cdot x = gxg^{-1}$. Las órbitas son las *clases de conjugación*; el estabilizador de x es su *centralizador* $C_G(x) = \{g : gx = xg\}$; los puntos fijos de todos los elementos forman el *centro* $Z(G)$.

Teorema 25.90 (Ecuación de clases)**.** *Para G finito, sumando los tamaños de las órbitas (órbita–estabilizador, Teorema 25.50):*

$$|G| = |Z(G)| + \sum_i [G : C_G(x_i)],$$

donde los x_i recorren representantes de las clases no centrales (cada índice > 1 y divisor de $|G|$).

Corolario 25.91. *Todo p-*grupo *(orden p^k) tiene centro no trivial: en la ecuación de clases, $|G|$ y cada índice $[G : C_G(x_i)] > 1$ son divisibles por p, luego $|Z(G)|$ también, y $|Z(G)| \geq p$. De aquí, por inducción, todo p-grupo es soluble (el cociente G/Z es un p-grupo menor), y todo grupo de orden p^2 es abeliano.*

Teorema 25.92 (Cauchy, demostración de McKay)**.** *Si $p \mid |G|$, existe un elemento de orden p.*

Demostración. Sea $X = \{(g_1, \dots, g_p) : g_1 g_2 \cdots g_p = e\}$: tiene $|G|^{p-1}$ elementos (los primeros $p-1$ son libres, el último queda determinado). El grupo cíclico $\mathbb{Z}_p$ actúa sobre X rotando las tuplas (si el producto es e, el producto rotado también: $g_2 \cdots g_p g_1 = g_1^{-1}(g_1 \cdots g_p)g_1 = e$). Las órbitas tienen tamaño 1 o p (divisores de p); las de tamaño 1 son las tuplas constantes $(g, \dots, g)$, es decir, los g con $g^p = e$. Contando módulo p: $|G|^{p-1} \equiv \#\{g : g^p = e\} \pmod p$. El lado izquierdo es $\equiv 0$ (pues $p \mid |G|$), y el conjunto de la derecha contiene a e: para que su cardinal sea múltiplo de p, debe contener al menos p elementos, y cualquiera distinto de e tiene orden exactamente p. □

Teorema 25.93 (Sylow, con idea de demostración)**.** *Sea $|G| = p^k m$ con $p \nmid m$.* (I) *Existen subgrupos de orden p^k (p-Sylow).* (II) *Todos los p-Sylow son conjugados.* (III) *Su número n_p cumple $n_p \equiv 1 \pmod p$ y $n_p \mid m$.*

Idea de (I). G actúa por traslación sobre la familia Ω de subconjuntos de tamaño p^k. El cardinal $|\Omega| = \binom{p^k m}{p^k}$ no es divisible por p (teorema de Kummer/Lucas sobre binomiales), así que alguna órbita tiene tamaño coprimo con p; el estabilizador de un elemento de esa órbita resulta tener orden exactamente p^k. Las partes (II) y (III) se obtienen haciendo actuar un Sylow sobre el conjunto de los demás por conjugación y contando puntos fijos módulo p. □

Ejemplo 25.94 (Sylow en acción: no hay grupos simples de orden 15)**.** $n_5 \equiv 1 \pmod 5$ y $n_5 \mid 3$: forzosamente $n_5 = 1$; el único 5-Sylow es normal. Igualmente $n_3 = 1$. Con ambos normales y de intersección trivial, $G \cong \mathbb{Z}_3 \times \mathbb{Z}_5 \cong \mathbb{Z}_{15}$: *todo* grupo de orden 15 es cíclico. Este estilo de conteo estructural es la puerta a la clasificación de grupos finitos.

Teorema 25.95 (Estructura de los grupos abelianos finitos)**.** *Todo grupo abeliano finito es producto directo de cíclicos de órdenes de potencias de primos, de forma única salvo orden. (Idea: se separa primero en p-componentes con Bézout —como en la descomposición primaria del capítulo 7, que es el mismo teorema para el anillo $K[\lambda]$ actuando en lugar de $\mathbb{Z}$—, y dentro de cada p-grupo se extraen cíclicos maximales inductivamente.) Junto con Lagrange, esto resuelve por completo la clasificación abeliana: por ejemplo, hay exactamente dos grupos abelianos de orden 12: $\mathbb{Z}_4 \times \mathbb{Z}_3$ y $\mathbb{Z}_2 \times \mathbb{Z}_2 \times \mathbb{Z}_3$.*

Idea clave

Qué dice esto en cristiano. Los grupos abelianos finitos están completamente clasificados: no hay sorpresas, no hay ejemplares exóticos esperando a ser descubiertos. Cada uno es un montoncito de relojes $\mathbb{Z}_{p^k}$ colocados en paralelo, y dos montoncitos distintos dan grupos distintos. Esto convierte la pregunta "¿cuántos grupos abelianos hay de orden n?" en un problema de contar particiones, que es aritmética pura y se resuelve con lápiz y papel. Ningún otro trozo del álgebra tiene una clasificación tan limpia.

Ejemplo 25.96 (Contar todos los grupos abelianos de un orden dado)**. (a) La receta.** Se factoriza $n = p_1^{e_1} \cdots p_r^{e_r}$ y, para cada primo, se cuentan las *particiones* del exponente. El total es el producto. Para $n = 8 = 2^3$, las particiones de 3 son 3, $2+1$ y $1+1+1$: tres grupos, que

son $\mathbb{Z}_8$, $\mathbb{Z}_4 \times \mathbb{Z}_2$ y $\mathbb{Z}_2 \times \mathbb{Z}_2 \times \mathbb{Z}_2$.

(b) Varios órdenes de un vistazo.

n	factorización	número de grupos abelianos
4	2^2	2
8	2^3	3
12	$2^2 \cdot 3$	$2 \cdot 1 = 2$
16	2^4	5
36	$2^2 \cdot 3^2$	$2 \cdot 2 = 4$
100	$2^2 \cdot 5^2$	$2 \cdot 2 = 4$

Los cuatro de orden 36 son $\mathbb{Z}_4 \times \mathbb{Z}_9$, $\mathbb{Z}_4 \times \mathbb{Z}_3 \times \mathbb{Z}_3$, $\mathbb{Z}_2 \times \mathbb{Z}_2 \times \mathbb{Z}_9$ y $\mathbb{Z}_2 \times \mathbb{Z}_2 \times \mathbb{Z}_3 \times \mathbb{Z}_3$.

(c) Las dos formas normales. El mismo grupo se puede escribir de dos maneras canónicas. Como potencias de primos: $\mathbb{Z}_2 \times \mathbb{Z}_2 \times \mathbb{Z}_9$. O como *factores invariantes*, cada uno dividiendo al siguiente: $\mathbb{Z}_2 \times \mathbb{Z}_{18}$, ya que $2 \mid 18$ y $2 \cdot 18 = 36$. Se pasa de una a otra agrupando el mayor de cada primo, luego el segundo mayor, y así sucesivamente. Análogamente $\mathbb{Z}_4 \times \mathbb{Z}_3 \times \mathbb{Z}_3$ se reescribe como $\mathbb{Z}_3 \times \mathbb{Z}_{12}$.

(d) Cuándo el producto se colapsa. $\mathbb{Z}_4 \times \mathbb{Z}_9$ es cíclico, porque $\gcd(4, 9) = 1$ y el elemento $(1, 1)$ tiene orden 36: es simplemente $\mathbb{Z}_{36}$. En cambio $\mathbb{Z}_2 \times \mathbb{Z}_2$ no es cíclico, porque todos sus elementos tienen orden 1 o 2. La regla es la del resto chino: el producto se funde en un solo reloj exactamente cuando los órdenes son coprimos.

Cuidado

[Contraejemplo: fuera de lo abeliano no hay nada parecido] Esta clasificación es exclusiva de los grupos abelianos. De orden 8 hay tres abelianos, pero también D_4 y Q_8, que no lo son: cinco grupos en total. Y para órdenes mayores la cuenta se descontrola: hay 267 grupos de orden 64 y más de diez millones de orden 1024. La clasificación de *todos* los grupos finitos simples ocupó a cientos de matemáticos durante treinta años y llena más de diez mil páginas. Segundo aviso: la unicidad es salvo el orden de los factores, así que $\mathbb{Z}_2 \times \mathbb{Z}_9$ y $\mathbb{Z}_9 \times \mathbb{Z}_2$ son el mismo grupo escrito al revés, no dos grupos distintos.

25.13.2. La simplicidad de A_5, con las cuentas completas

Demostración detallada del Teorema 25.68. Paso 1: las clases de conjugación de A_5. Los tipos cíclicos pares en S_5 son: identidad (1), dobles trasposiciones (15), 3-ciclos (20) y 5-ciclos (24). En A_n una clase de S_n se conserva o se parte en dos mitades, según si algún elemento impar centraliza a un representante. Las dobles trasposiciones y los 3-ciclos tienen centralizadores con elementos impares (por ejemplo, una trasposición disjunta): clases de 15 y 20. Los 5-ciclos no: su centralizador en S_5 es el cíclico de orden 5 que generan, todo par; la clase de 24 se parte en dos de 12. Total: $60 = 1 + 15 + 20 + 12 + 12$.

Paso 2: ningún subconjunto sirve. Un subgrupo normal es unión de clases completas, contiene a la clase $\{e\}$, y su orden divide a 60. Los candidatos son sumas de $\{1\}$ con algunas de $\{15, 20, 12, 12\}$:

$$1 + 15 = 16, \quad 1 + 20 = 21, \quad 1 + 12 = 13, \quad 1 + 15 + 20 = 36, \quad 1 + 15 + 12 = 28,$$

$1{+}20{+}12 = 33$, $1{+}12{+}12 = 25$, $1{+}15{+}12{+}12 = 40$, $1{+}20{+}12{+}12 = 45$, $1{+}15{+}20{+}12 = 48$.

Ninguno divide a 60 (los divisores propios no triviales son 2,3,4,5,6,10,12,15,20,30). Solo quedan 1 y 60: A_5 es simple. □ □

Para $n > 5$ la simplicidad de A_n se demuestra por inducción (todo normal no trivial contiene un 3-ciclo, y los 3-ciclos generan A_n). La consecuencia estructural, que el capítulo 26 explota: la única cadena normal de S_n ($n \geq 5$) es $\{e\} \trianglelefteq A_n \trianglelefteq S_n$, y el cociente A_n no es abeliano: no hay refinamiento soluble posible.

25.13.3. Aritmética de anillos: los teoremas que faltaban

Teorema 25.97 (Isomorfía para anillos)**.** *Si $\varphi : R \to S$ es homomorfismo de anillos,* $\ker\varphi$ *es un ideal y* $R/\ker\varphi \cong \operatorname{im}\varphi$*. (Misma demostración que en grupos, respetando además el producto.)*

Teorema 25.98 (Resto chino)**.** *Si I, J son ideales* comaximales *($I + J = R$), entonces*

$$R/(I \cap J) \cong R/I \times R/J.$$

En $\mathbb{Z}$: si $\gcd(m,n) = 1$, $\mathbb{Z}_{mn} \cong \mathbb{Z}_m \times \mathbb{Z}_n$.

Demostración. El homomorfismo $\varphi(r) = (r + I,\ r + J)$ tiene núcleo $I \cap J$. Sobreyectividad: por comaximalidad existen $i \in I$, $j \in J$ con $i + j = 1$; entonces $\varphi(aj + bi) = (a + I,\ b + J)$ para cualesquiera a, b, pues $aj = a(1 - i) \equiv a \bmod I$ y $bi \equiv b \bmod J$. El primer teorema de isomorfía concluye. □

Idea clave

Qué dice esto en cristiano. Si dos módulos no comparten factores, trabajar módulo el producto es exactamente lo mismo que llevar dos contabilidades separadas. Un reloj de 12 horas equivale a llevar a la vez un reloj de 4 y otro de 3: cada hora del grande corresponde a un par de horas de los pequeños, y ese emparejamiento no se repite nunca. La palabra *comaximal* es la versión abstracta de "coprimos": que entre los dos ideales se genere todo el anillo, igual que $\gcd(m,n) = 1$ permite escribir $1 = am + bn$. El teorema es la razón de que tantos cálculos aritméticos se puedan partir en trozos independientes y luego recomponer.

Ejemplo 25.99 (El diccionario $\mathbb{Z}_{12} \cong \mathbb{Z}_4 \times \mathbb{Z}_3$, entero)**. (a) La tabla completa.** A cada $a \in \mathbb{Z}_{12}$ le asignamos el par (a mód 4, a mód 3):

a	0	1	2	3	4	5	6	7	8	9	10	11
(a_4, a_3)	00	11	22	30	01	12	20	31	02	10	21	32

Los doce pares son distintos, luego la correspondencia es biyectiva, y hay exactamente $4 \cdot 3 = 12$ pares posibles: no sobra ni falta ninguno.

(b) Respeta la suma. $5+9=14\equiv 2$ en $\mathbb{Z}_{12}$. Por el otro camino, $5\mapsto(1,2)$ y $9\mapsto(1,0)$, cuya suma componente a componente es $(1+1,2+0)=(2,2)$; y el único a con $a\equiv 2$ (4) y $a\equiv 2$ (3) es $a=2$. Coincide.

(c) Resolver un sistema. Buscar x con $x\equiv 3$ (5), $x\equiv 4$ (7) y $x\equiv 2$ (9) es pedir el elemento de $\mathbb{Z}_{315}$ correspondiente al triple $(3,4,2)$; la solución es $x=263$, única módulo $315=5\cdot 7\cdot 9$. Comprobación: $263=52\cdot 5+3$, $263=37\cdot 7+4$ y $263=29\cdot 9+2$.

(d) Por qué $\mathbb{Z}_4\times\mathbb{Z}_3$ es cíclico. El elemento $(1,1)$ tiene orden $\operatorname{mcm}(4,3)=12$, luego genera el grupo entero; por eso el producto vuelve a ser un $\mathbb{Z}_{12}$. En general, $\mathbb{Z}_m\times\mathbb{Z}_n$ es cíclico exactamente cuando $\gcd(m,n)=1$, que es la misma condición del teorema vista desde el lado de los órdenes.

Cuidado

[Contraejemplo: sin coprimalidad falla] Con $m=n=2$ los módulos no son coprimos, y en efecto $\mathbb{Z}_4$ *no* es isomorfo a $\mathbb{Z}_2\times\mathbb{Z}_2$: en $\mathbb{Z}_4$ los órdenes de los elementos son $1,4,2,4$, mientras que en $\mathbb{Z}_2\times\mathbb{Z}_2$ son $1,2,2,2$. El primero tiene un elemento de orden 4 y el segundo no, así que ningún cambio de nombres puede identificarlos. La correspondencia $a\mapsto(a \bmod 2, a \bmod 2)$ ni siquiera es inyectiva: 0 y 2 van al mismo par. En términos de ideales, (2) y (2) no son comaximales porque $(2)+(2)=(2)\neq\mathbb{Z}$, y ahí es exactamente donde se cae la demostración.

El teorema chino es el motor de la aritmética computacional (trabajar módulo primos y recombinar), de la criptografía RSA (descifrado rápido módulo p y q), y de la descomposición de $(\mathbb{Z}_n)^\times$ que se usará con los cuerpos ciclotómicos en el capítulo 26.

Teorema 25.100 (Todo dominio euclídeo es DIP; todo DIP es DFU)**.** (Euclídeo $\Rightarrow$ DIP)*: dado un ideal $I\neq 0$, sea $d\in I$ de tamaño mínimo; para $a\in I$, la división $a=qd+r$ con r menor que d fuerza $r=a-qd\in I$, luego $r=0$: $I=(d)$.* (DIP $\Rightarrow$ DFU, idea)*: en un DIP, todo irreducible es primo (Bézout da el lema de Euclides), las cadenas crecientes de ideales estabilizan (la unión de la cadena es un ideal, principal, generado por un elemento que ya vive en algún eslabón), y con esas dos piezas la factorización existe (por la condición de cadenas) y es única (por el lema de Euclides, cancelando irreducible a irreducible).*

Teorema 25.101 (Lema de Gauss, demostrado)**.** *El producto de polinomios* primitivos *(mcd de coeficientes 1) en $\mathbb{Z}[x]$ es primitivo. En consecuencia, si $f\in\mathbb{Z}[x]$ factoriza en $\mathbb{Q}[x]$, factoriza ya en $\mathbb{Z}[x]$ con factores de los mismos grados.*

Demostración. Si fg no fuera primitivo, un primo p dividiría todos sus coeficientes. Reduciendo módulo p (homomorfismo $\mathbb{Z}[x]\to\mathbb{F}_p[x]$): $\bar f\bar g=0$ en $\mathbb{F}_p[x]$, que es dominio de integridad; luego $\bar f=0$ o $\bar g=0$, es decir, p divide todos los coeficientes de f o de g: contradice su primitividad. Para la consecuencia: de una factorización racional se sacan denominadores y contenidos, y el lema garantiza que las constantes se ajustan para aterrizar en $\mathbb{Z}[x]$. □

Idea clave

Qué dice esto en cristiano. El lema de Gauss resuelve una preocupación muy razonable: al factorizar un polinomio con coeficientes enteros, ¿puede aparecer una factorización que use fracciones y que no exista con enteros? La respuesta es no. Si f se parte en $\mathbb{Q}[x]$, se parte también en $\mathbb{Z}[x]$ con los mismos grados. Esto es lo que legitima buscar factorizaciones probando solo con divisores enteros, y lo que hace que criterios como el de Eisenstein sirvan para demostrar irreducibilidad sobre $\mathbb{Q}$ examinando únicamente números enteros.

Ejemplo 25.102 (Contenidos, productos y factorizaciones)**. (a) El lema en acción.** Los polinomios $2x+1$ y $3x+2$ son primitivos, porque el máximo común divisor de sus coeficientes es 1 en ambos casos. Su producto es $6x^2+7x+2$, cuyo contenido es $\gcd(6,7,2)=1$: primitivo también, tal como afirma el lema. Nótese que el 6 y el 2 comparten factor, pero el 7 de en medio impide que el contenido crezca.

(b) Qué pasa con los contenidos generales. $(2x+2)(3x+3)=6x^2+12x+6$, cuyo contenido es $6=2\cdot 3$: exactamente el producto de los contenidos de los factores. Esa es la forma general del lema: el contenido de un producto es el producto de los contenidos, y el caso primitivo es simplemente $1\cdot 1=1$.

(c) La consecuencia práctica. El polinomio $2x^2-x-1$ factoriza sobre $\mathbb{Q}$; el lema garantiza que hay una factorización entera, y en efecto es $(x-1)(2x+1)$. Sin el lema habría que temer factorizaciones como $\left(x-\frac{1}{2}\right)(2x+2)$, que existen pero se reducen a la entera sacando y metiendo constantes.

(d) Eisenstein, que depende de esto. Para $f=x^3+2x^2+2x+2$, el primo $p=2$ divide a todos los coeficientes salvo el principal, y $p^2=4$ no divide al término independiente 2. El criterio de Eisenstein concluye que f es irreducible sobre $\mathbb{Q}$, y esa conclusión solo es legítima porque el lema de Gauss traduce "no factoriza en $\mathbb{Z}[x]$" a "no factoriza en $\mathbb{Q}[x]$". También son irreducibles x^2-2 y el ciclotómico $x^4+x^3+x^2+x+1$.

Cuidado

[Irreducible depende del cuerpo] Un polinomio no es irreducible en abstracto, sino *sobre* un cuerpo concreto. El polinomio x^4+1 es irreducible sobre $\mathbb{Q}$, y sin embargo sobre $\mathbb{F}_2$ se descompone del todo: $x^4+1=(x+1)^4$. Sobre $\mathbb{R}$ tampoco es irreducible, pues es producto de dos cuadráticas. Segundo aviso: el lema de Gauss habla de *factorizaciones no triviales*, es decir en factores de grado positivo; sacar la constante 2 de $2x+2$ no cuenta como factorizar, porque las constantes no nulas son unidades en $\mathbb{Q}[x]$ aunque no lo sean en $\mathbb{Z}[x]$. Esa diferencia entre unidades es justamente la que el lema se encarga de controlar.

Demostración del criterio de Eisenstein. Sea $f=c_nx^n+\cdots+c_0$ con $p\mid c_i$ $(i<n)$, $p\nmid c_n$, $p^2\nmid c_0$. Si $f=gh$ en $\mathbb{Z}[x]$ (basta por Gauss) con g,h no constantes, reducimos módulo p: $\bar f=\bar c_nx^n$, así que $\bar g\bar h=\bar c_nx^n$ en $\mathbb{F}_p[x]$. Por factorización única en $\mathbb{F}_p[x]$, ambos $\bar g,\bar h$ son monomios: $\bar g=ax^r$, $\bar h=bx^s$ con $r,s\geq 1$ (grados positivos). Entonces p divide los términos independientes de g y de h, y $p^2\mid g(0)h(0)=c_0$: contradicción. □

Obsérvese la arquitectura de ambas demostraciones: reducir módulo p convierte una pregunta sobre $\mathbb{Z}[x]$ en una sobre el dominio $\mathbb{F}_p[x]$, donde la factorización única ya está disponible. "Cambiar de anillo por un homomorfismo bien elegido" es quizá la técnica más productiva de toda el álgebra.

25.13.4. El grupo multiplicativo de un cuerpo finito es cíclico

Teorema 25.103. *Todo subgrupo finito G del grupo multiplicativo de un cuerpo es cíclico. En particular $\mathbb{F}_q^* \cong \mathbb{Z}_{q-1}$.*

Demostración. Sea $n = |G|$ y d el máximo de los órdenes de los elementos (el *exponente*). Por la estructura de los abelianos finitos, todo orden divide a d, así que los n elementos de G satisfacen $x^d = 1$. Pero en un cuerpo, el polinomio $x^d - 1$ tiene a lo sumo d raíces (Teorema 25.79): $n \leq d$. Como siempre $d \leq n$ (Lagrange), $d = n$: hay un elemento de orden n, que genera. □

Idea clave

Qué dice esto en cristiano. En un cuerpo, los elementos no nulos con la multiplicación siempre se organizan en un solo círculo: existe un elemento cuyas potencias sucesivas barren todos los demás. A ese elemento se le llama *raíz primitiva* o generador. La razón profunda es que en un cuerpo la ecuación $x^d = 1$ no puede tener más de d soluciones, y eso impide que los órdenes se repartan de forma desordenada. Es un resultado con consecuencias muy prácticas: convierte multiplicar en sumar exponentes, que es la base de los logaritmos discretos y de buena parte de la criptografía.

Ejemplo 25.104 (Encontrar generadores a mano)**. (a) $\mathbb{F}_7^*$ tiene seis elementos.** Las potencias de 3 son $3, 2, 6, 4, 5, 1$: aparecen los seis, luego 3 es generador. Las de 2, en cambio, son $2, 4, 1, 2, 4, 1$: solo produce tres elementos, porque el orden de 2 es 3. Los generadores de $\mathbb{F}_7^*$ son exactamente 3 y 5, y son $\varphi(6) = 2$, como predice la teoría.

(b) Cuántos generadores hay. En $\mathbb{F}_{11}^*$, de orden 10, hay $\varphi(10) = 4$ generadores, que son $2, 6, 7, 8$. En $\mathbb{F}_{13}^*$, de orden 12, hay $\varphi(12) = 4$, que son $2, 6, 7, 11$. La regla es siempre la misma: una vez fijado un generador g, el elemento g^k vuelve a ser generador exactamente cuando $\gcd(k, q-1) = 1$.

(c) La utilidad: multiplicar es sumar. Con $g = 3$ en $\mathbb{F}_7$, la tabla de potencias de (a) da el diccionario $3^1 = 3$, $3^2 = 2$, $3^3 = 6$, $3^4 = 4$, $3^5 = 5$, $3^6 = 1$. Para multiplicar $2 \cdot 6$ basta sumar exponentes: $2 = 3^2$ y $6 = 3^3$, luego $2 \cdot 6 = 3^5 = 5$. Y en efecto $12 \equiv 5 \pmod 7$. Es un logaritmo, con la misma mecánica que el del capítulo escolar pero en un mundo finito.

(d) Por qué la demostración necesita el cuerpo. La clave es que $x^2 = 1$ solo puede tener dos soluciones. En $\mathbb{F}_7$ las tiene: $x = 1$ y $x = 6$. Ese control sobre el número de raíces es lo que obliga a que exista un elemento de orden máximo $q - 1$, y es exactamente lo que se pierde en cuanto el anillo no es un cuerpo.

Cuidado

[Contraejemplo: fuera de un cuerpo esto es falso] El grupo $\mathbb{Z}_8^* = \{1,3,5,7\}$ tiene cuatro elementos y *no* es cíclico: los órdenes son $1,2,2,2$, ninguno vale 4. La razón salta a la vista con el criterio del apartado (d): la ecuación $x^2 = 1$ tiene en $\mathbb{Z}_8$ *cuatro* soluciones, porque $3^2 = 9 \equiv 1$, $5^2 = 25 \equiv 1$ y $7^2 = 49 \equiv 1$. Eso solo puede ocurrir si hay divisores de cero, y en efecto $\mathbb{Z}_8$ no es un cuerpo. Moraleja: la hipótesis "dentro de un cuerpo" no es decorativa; es toda la demostración.

La demostración es una miniatura del libro entero: teoría de grupos (exponente, Lagrange) y teoría de polinomios (cota de raíces) colaborando. Este teorema explica los "logaritmos discretos" del ejercicio de $\mathbb{F}_8$, la existencia de *raíces primitivas* módulo p, y es la base de la criptografía de Diffie–Hellman.

25.14. Complementos: productos de grupos, automorfismos y simétricos

25.14.1. Productos directos y semidirectos

$G \cong N \times H$ si ambos normales, intersección trivial, $NH = G$; si solo N es normal: *semidirecto* $N \rtimes H$ (H actúa por conjugación). $S_3 \cong \mathbb{Z}_3 \rtimes \mathbb{Z}_2$, $D_n \cong \mathbb{Z}_n \rtimes \mathbb{Z}_2$, $F_{20} = \mathbb{Z}_5 \rtimes \mathbb{Z}_4$: la fábrica de no abelianos con piezas abelianas. Grupos pequeños: hasta orden 10 son cíclicos, $\mathbb{Z}_2^2$, $\mathbb{Z}_4$, S_3, los cinco de orden 8 ($\mathbb{Z}_8, \mathbb{Z}_4 \times \mathbb{Z}_2, \mathbb{Z}_2^3, D_4, Q_8$), $\mathbb{Z}_9, \mathbb{Z}_3^2, \mathbb{Z}_{10}, D_5$. Q_8: los cuaterniones de Hamilton, hoy el estándar de rotaciones 3D (sin gimbal lock).

25.14.2. Automorfismos

$\mathrm{Aut}(\mathbb{Z}_n) \cong (\mathbb{Z}_n)^\times$: y $\mathrm{Gal}(\mathbb{Q}(\zeta_n)/\mathbb{Q}) \cong \mathrm{Aut}(\mathbb{Z}_n)$ —el grupo ciclotómico es el grupo de automorfismos de las raíces de la unidad—; los datos de un semidirecto son un homomorfismo $H \to \mathrm{Aut}(N)$.

25.14.3. Polinomios simétricos

Teorema 25.105. *Todo simétrico es polinomio único en los elementales $e_1, \dots, e_n$ (algoritmo del monomio líder). Newton: $p_1 = e_1$, $p_2 = e_1 p_1 - 2e_2$, $p_3 = e_1 p_2 - e_2 p_1 + 3e_3$,...*

Para $x^3 + px + q$: $p_2 = -2p$, $p_3 = -3q$ sin conocer raíces. La semilla histórica de Galois: lo simétrico se calcula desde los coeficientes; lo invariante bajo un subgrupo genera el cuerpo intermedio correspondiente (el discriminante y la resolvente cúbica son los ejemplos operativos).

25.15. Cierre del temario: los subtemas restantes

25.15.1. Segundo y tercer teoremas de isomorfía

Teorema 25.106. $\frac{HN}{N} \cong \frac{H}{H \cap N}$ *(con $N \trianglelefteq G$); y si $N \subseteq M$ ambos normales: $\frac{G/N}{M/N} \cong \frac{G}{M}$ (las fracciones se simplifican). Demostraciones: componer con la proyección y aplicar el primero.*

Teorema 25.107 (Correspondencia)**.** *Subgrupos de G/N $\leftrightarrow$ subgrupos de G que contienen a N, respetando normalidad e índices: la Galois interna de los grupos.*

Idea clave

Qué dice esto en cristiano. El teorema de correspondencia dice que al pasar al cociente G/N no se pierde información sobre lo que había *por encima* de N: los subgrupos del cociente son exactamente los subgrupos de G que contenían a N, vistos desde arriba. Es como mirar un edificio y quedarse solo con las plantas por encima de la quinta: la lista de plantas visibles se conserva íntegra, con su orden y sus alturas relativas. Junto con los dos primeros teoremas de isomorfía, esto convierte los cocientes en una herramienta manejable en lugar de un objeto abstracto: se puede razonar sobre G/N mirando a G.

Ejemplo 25.108 (Los tres teoremas con números)**. (a) Segundo teorema de isomorfía, en $\mathbb{Z}$.** Tomemos $G = \mathbb{Z}$, $N = 6\mathbb{Z}$ y $H = 4\mathbb{Z}$. Entonces $H + N = \gcd(4,6)\mathbb{Z} = 2\mathbb{Z}$ y $H \cap N = \operatorname{mcm}(4,6)\mathbb{Z} = 12\mathbb{Z}$. El teorema afirma que $(H+N)/N \cong H/(H\cap N)$: el primero es $2\mathbb{Z}/6\mathbb{Z}$, con $6/2 = 3$ elementos, y el segundo es $4\mathbb{Z}/12\mathbb{Z}$, con $12/4 = 3$. Coinciden.

(b) Tercer teorema, en $\mathbb{Z}_{12}$. Con $N = \langle 6\rangle = \{0,6\}$ y $M = \langle 2\rangle$, se tiene $|G/N| = 12/2 = 6$ y $|M/N| = 6/2 = 3$, luego el cociente doble $(G/N)/(M/N)$ tiene $6/3 = 2$ elementos. Por otro lado $|G/M| = 12/6 = 2$. Iguales, tal como dice la regla de que "las fracciones se simplifican".

(c) Correspondencia, con la lista completa. En $\mathbb{Z}_{12}$ con $N = \langle 4\rangle = \{0,4,8\}$, los subgrupos que contienen a N son exactamente tres: el propio $\{0,4,8\}$, el subgrupo $\{0,2,4,6,8,10\}$ y todo $\mathbb{Z}_{12}$. Por otro lado $\mathbb{Z}_{12}/N$ tiene $12/3 = 4$ elementos y es $\mathbb{Z}_4$, cuyos subgrupos son también tres, de órdenes 1, 2 y 4. La correspondencia es biyectiva y respeta los índices: $12/3 = 4$, $12/6 = 2$, $12/12 = 1$.

(d) Para qué se usa. Este es el mecanismo que permite hacer inducción sobre el orden del grupo: se busca un subgrupo normal, se pasa al cociente, que es más pequeño, se aplica la hipótesis de inducción y se sube el resultado. Casi todas las demostraciones sobre grupos finitos solubles o nilpotentes tienen esta forma, y la correspondencia es la que garantiza que subir no pierde nada.

Cuidado

[La correspondencia solo ve hacia arriba] Los subgrupos de G/N corresponden a los subgrupos de G que *contienen* a N, no a todos los subgrupos de G. En el ejemplo (c), $\mathbb{Z}_{12}$ tiene seis subgrupos en total, pero solo tres sobreviven al cociente; los otros tres ($\{0\}$, $\{0,6\}$ y $\{0,3,6,9\}$) desaparecen o se funden. Pasar al cociente sí destruye información, y la destruye toda por debajo de N. Segundo aviso: en el segundo teorema, la expresión HN es un subgrupo únicamente porque N es normal; con dos subgrupos cualesquiera, el conjunto de productos HK no tiene por qué ser subgrupo, y ese es el error más común al aplicarlo.

25.15.2. Jordan–Hölder

Serie de composición: cadena maximal con cocientes *simples*; los factores son únicos salvo orden: *los primos de un grupo.* $\mathbb{Z}_{12}$: factores $\mathbb{Z}_2, \mathbb{Z}_2, \mathbb{Z}_3$ en cualquier orden ($12 = 2 \cdot 2 \cdot 3$); S_5: $\mathbb{Z}_2$ y A_5 —la obstrucción con nombre—. Soluble $\iff$ factores cíclicos de orden primo.

25.15.3. Presentaciones

$G = \langle \text{generadores} \mid \text{relaciones} \rangle$: $\mathbb{Z}_n = \langle a \mid a^n \rangle$, $D_n = \langle r, s \mid r^n, s^2, srsr \rangle$, $Q_8 = \langle i, j \mid i^4, i^2 j^{-2}, jij^{-1}i \rangle$; formalmente, cociente del grupo libre. Decidir si una presentación da el grupo trivial es *indecidible* (Novikov–Boone): el álgebra toca los límites de la computación.

25.15.4. El cuerpo de fracciones

Todo dominio R se sumerge en $\mathrm{Frac}(R)$: pares $\frac{a}{b}$ módulo $ad = bc$, con propiedad universal. $\mathrm{Frac}(\mathbb{Z}) = \mathbb{Q}$ (¡la construcción oficial!), $\mathrm{Frac}(K[x]) = K(x)$ —el cuerpo de la ecuación general—.

25.15.5. Factorización efectiva

Cómo factoriza el software en $\mathbb{Q}[x]$: libre de cuadrados con $\gcd(f, f')$; Berlekamp mod p (núcleo del Frobenius: álgebra lineal sobre $\mathbb{F}_p$); levantamiento de Hensel (Newton p-ádico); recombinación con LLL. Frobenius, Newton y retículos: capítulos 6, 7 y 25 corriendo dentro de cada CAS.

25.16. Profundización final: Wilson, dos cuadrados, resultantes y caracteres

25.16.1. Wilson

Teorema 25.109. *p primo $\iff$ $(p-1)! \equiv -1 \pmod p$. (Emparejar inversos; solo ± 1 son autoinversos: $x^2 = 1$ tiene ≤ 2 raíces en un cuerpo.)*

Corolario: para $p \equiv 1 \pmod 4$, $x = (\frac{p-1}{2})!$ cumple $x^2 \equiv -1$: la llave del siguiente teorema.

25.16.2. Suma de dos cuadrados

Teorema 25.110 (Fermat)**.** *p impar es $a^2 + b^2 \iff p \equiv 1 \pmod 4$.*

Demostración. ($\Rightarrow$) módulo 4. ($\Leftarrow$) con $x^2 \equiv -1$: en $\mathbb{Z}[i]$, $p \mid (x+i)(x-i)$ sin dividir factores: p no es primo gaussiano, $p = \alpha\beta$, y las normas dan $N(\alpha) = p = a^2 + b^2$. □

Aritmética clásica demostrada *factorizando en otro anillo*: el manifiesto de la teoría algebraica de números, con el vocabulario partido/inerte/ramificado que el Frobenius sistematiza.

25.16.3. Resultantes

$\operatorname{Res}(f,g) = \det(\text{Sylvester}) = 0 \iff$ raíz común; $\Delta = \pm\frac{1}{a}\operatorname{Res}(f,f')$: el discriminante como determinante (así se calcula $-4p^3 - 27q^2$). Elimina variables: $\operatorname{Res}_y(f,g)$ da las abscisas de intersección de curvas (Bézout: $\leq mn$ puntos); ancestro de las bases de Gröbner.

25.16.4. Representaciones y caracteres

Idea clave

Intuición primero. Un grupo abstracto es difícil de tocar; una matriz se puede calcular. Una *representación* es una forma de hacer que el grupo actúe por matrices, es decir, de convertir simetrías abstractas en álgebra lineal del capítulo 7. El *carácter* es la traza de esas matrices: una sombra numérica bastante pobre a primera vista —un número por clase de conjugación— y sin embargo, milagrosamente, *esa sombra determina la representación*. La tabla de caracteres es la huella dactilar del grupo: cabe en una servilleta y contiene su estructura. Por eso la química cuántica y la física de partículas trabajan con tablas de caracteres y no con los grupos mismos.

$\rho : G \to GL_n(\mathbb{C})$; carácter $\chi = \operatorname{tr}\rho$ (constante en clases). Maschke y Schur: todo se parte en irreducibles; tantos como clases; $\sum d_i^2 = |G|$; caracteres ortonormales.

Ejemplo 25.111 (Tabla de S_3).

	e	(12) [3]	(123) [2]
trivial	1	1	1
signo	1	-1	1
estándar	2	0	-1

$1+1+4=6$ ✓; $\langle\chi_{\text{est}},\chi_{\text{est}}\rangle = \frac{1}{6}(4+0+2) = 1$ ✓. Con tablas así, la química predice espectros moleculares y la física clasifica partículas: el matrimonio de los capítulos 7 y 25.

25.16.5. La base de Hilbert

R noetheriano $\Rightarrow R[x]$ noetheriano: todo sistema polinomial equivale a uno finito. Fundó el álgebra conmutativa ("teología" protestó Gordan); sobre él, la geometría algebraica y Gröbner.

25.17. Última capa: Sylow demostrado, módulos sobre DIP y Dedekind

25.17.1. Sylow, con demostración

Idea clave

Intuición primero. Lagrange dice que el orden de un subgrupo *divide* al del grupo, pero no promete que exista uno para cada divisor (y de hecho A_4, de orden 12, no tiene subgrupo de orden 6). Sylow rescata lo máximo que se puede rescatar: para cada primo, la *potencia máxima* sí aparece siempre como subgrupo. Es como decir que aunque no

todo divisor tenga representante, el "esqueleto primo" del grupo está garantizado. Y añade dos regalos de conteo: todos esos subgrupos son el mismo visto desde distintos ángulos (conjugados), y su cantidad está atrapada entre dos condiciones aritméticas tan estrictas que muchas veces solo queda la posibilidad $n_p = 1$ —o sea, normal— y con eso el grupo se desmonta solo. Así se clasifican familias enteras de grupos sin enumerar un solo elemento.

Teorema 25.112. *$|G| = p^k m$, $p \nmid m$: existen subgrupos de orden p^k; todos conjugados; su número $n_p \equiv 1 \pmod p$ y $n_p \mid m$.*

Demostración de existencia (Wielandt). G actúa por traslación sobre $\Omega = \{$subconjuntos de tamaño $p^k\}$, con $|\Omega| = \binom{p^k m}{p^k} \not\equiv 0 \pmod p$ (teorema de Lucas: los dígitos p-ádicos de $p^k m$ dominan a los de p^k). Alguna órbita tiene tamaño coprimo con p; su estabilizador H cumple $p^k \mid |H|$ (la órbita de tamaño $\frac{|G|}{|H|}$ coprimo con p obliga) y $|H| \le p^k$ (un estabilizador de un conjunto S actúa libremente en S por traslación: $|H|$ divide... más directo: cada clase $Hs \subseteq S$ da $|H| \le |S| = p^k$). Luego $|H| = p^k$. Conjugación y conteo módulo p (haciendo actuar un Sylow sobre los demás) dan (II) y (III). □

Ejemplo 25.113 (Aplicación total: grupos de orden pq)**.** Sean $p < q$ primos. $n_q \mid p$ y $n_q \equiv 1 \pmod q$ fuerzan $n_q = 1$: el q-Sylow es normal, y $G \cong \mathbb{Z}_q \rtimes \mathbb{Z}_p$. Si $p \nmid q-1$, el homomorfismo $\mathbb{Z}_p \to \mathrm{Aut}(\mathbb{Z}_q) \cong \mathbb{Z}_{q-1}$ es trivial: $G \cong \mathbb{Z}_{pq}$ cíclico (así, todo grupo de orden 15, 33, 35 es cíclico). Si $p \mid q-1$, hay exactamente un no abeliano (S_3 para 6, D_5 no—: el de orden $2q$ es diédrico, y para pq impar el de Frobenius). Clasificación completa de una familia infinita con dos teoremas.

25.17.2. Módulos sobre un DIP: el teorema que unifica

Un *módulo* es un "espacio vectorial sobre un anillo". El teorema de estructura: *todo módulo finitamente generado sobre un DIP es suma directa de cíclicos* $R/(d_1) \oplus \cdots \oplus R/(d_k) \oplus R^r$ con $d_1 \mid d_2 \mid \cdots$ (factores invariantes, únicos). Dos cosechas de un solo teorema:

- $R = \mathbb{Z}$: **la clasificación de los grupos abelianos finitamente generados** ($\mathbb{Z}_{d_1} \oplus \cdots \oplus \mathbb{Z}^r$), ahora demostrada y no solo enunciada.
- $R = K[\lambda]$, con V visto como módulo vía $p(\lambda) \cdot v = p(A)v$: **las formas canónicas racional y de Jordan** —los d_i son los factores invariantes de A, y los cíclicos primarios son los bloques—.

Que los grupos abelianos y las formas de Jordan sean *el mismo teorema* con dos anillos distintos es una de las grandes unificaciones del álgebra: el argumento (diagonalizar matrices sobre el DIP con la forma normal de Smith: Gauss permitiendo divisiones con resto) es el mismo algoritmo de eliminación del capítulo 7, corriendo sobre $\mathbb{Z}$ o $K[\lambda]$.

25.17.3. Símbolo de Jacobi

Extensión multiplicativa de Legendre a impares compuestos: $\left(\frac{a}{n}\right) = \prod \left(\frac{a}{p_i}\right)^{e_i}$; satisface la misma reciprocidad, lo que da un algoritmo de tipo Euclides para calcular cualquier símbolo sin factorizar:

$$\left(\tfrac{219}{383}\right) \overset{(1)}{=} -\left(\tfrac{383}{219}\right) = -\left(\tfrac{164}{219}\right) \overset{(2)}{=} -\left(\tfrac{41}{219}\right) \overset{(3)}{=} -\left(\tfrac{219}{41}\right) = -\left(\tfrac{14}{41}\right) = -\left(\tfrac{2}{41}\right)\left(\tfrac{7}{41}\right) \overset{(4)}{=} -(+1)(-1) = +1 :$$

(1) reciprocidad con signo ($219 \equiv 383 \equiv 3$ mód 4); (2) extraer el cuadrado $164 = 4 \cdot 41$; (3) reciprocidad sin signo ($41 \equiv 1$ mód 4); (4) $\left(\frac{2}{41}\right) = +1$ pues $41 \equiv 1$ mód 8, y $\left(\frac{7}{41}\right) = \left(\frac{41}{7}\right) = \left(\frac{6}{7}\right) = \left(\frac{2}{7}\right)\left(\frac{3}{7}\right) = (+1)(-1) = -1$. Todo por volteos y reducciones, sin factorizar 383. Es la subrutina de los tests de primalidad de Solovay–Strassen y de los algoritmos de raíces cuadradas modulares (Tonelli–Shanks) que usa la criptografía de curvas elípticas.

25.17.4. Enteros algebraicos y dominios de Dedekind

Un *entero algebraico* es raíz de un mónico entero; en $\mathbb{Q}(\sqrt{d})$ forman $\mathbb{Z}[\sqrt{d}]$ o $\mathbb{Z}[\frac{1+\sqrt{d}}{2}]$ (según d mod 4). En estos anillos la factorización de *elementos* puede fallar ($\mathbb{Z}[\sqrt{-5}]$), pero Dedekind restauró el orden un piso arriba:

Teorema 25.114. *En el anillo de enteros de un cuerpo de números,* ***todo ideal factoriza de forma única en ideales primos****.*

Idea clave

Qué dice esto en cristiano. En los enteros ordinarios, cada número se descompone en primos de una sola manera. En los anillos de enteros de otros cuerpos eso se rompe: un mismo número admite dos factorizaciones distintas en irreducibles, lo que durante el siglo XIX estropeó varios intentos de demostrar el último teorema de Fermat. La salida de Kummer y Dedekind fue genial: si los *números* no factorizan bien, súbase un nivel y factorícense los *ideales*, que sí lo hacen. Los ideales primos son los "números ideales" que faltaban y que reparan la unicidad.

Ejemplo 25.115 (El ejemplo que obligó a inventar los ideales)**. (a) La doble factorización.** En $\mathbb{Z}[\sqrt{-5}]$ se tiene

$$6 = 2 \cdot 3 = (1 + \sqrt{-5})(1 - \sqrt{-5}),$$

y las dos son factorizaciones en irreducibles genuinamente distintas: los factores de una no son asociados de los de la otra.

(b) Comprobar que los cuatro son irreducibles. La norma es $N(a+b\sqrt{-5}) = a^2+5b^2$ y es multiplicativa. Aquí $N(2) = 4$, $N(3) = 9$ y $N(1 \pm \sqrt{-5}) = 1 + 5 = 6$; el producto cuadra por ambos lados, $4 \cdot 9 = 36 = 6 \cdot 6$. Si 2 factorizara, sus factores tendrían normas 2 y 2; pero $a^2 + 5b^2 = 2$ no tiene solución entera (con $b = 0$ haría falta $a^2 = 2$, y con $|b| \geq 1$ ya se pasa de 5). Lo mismo con $a^2 + 5b^2 = 3$. Luego ninguno de los cuatro se puede partir.

(c) La reparación con ideales. Existen ideales primos P, Q, R tales que $(2) = P^2$, $(3) = QR$, $(1+\sqrt{-5}) = PQ$ y $(1-\sqrt{-5}) = PR$. Entonces por un camino $(6) = (2)(3) = P^2QR$,

y por el otro $(6) = (1+\sqrt{-5})(1-\sqrt{-5}) = PQ \cdot PR = P^2QR$. Las dos factorizaciones de antes eran dos maneras de agrupar los mismos cuatro ideales primos.

(d) Cuándo no hace falta. En $\mathbb{Z}[\mathrm{i}]$, los enteros de Gauss, la factorización única sí vale y no hay que subir a los ideales: allí $5 = (2+\mathrm{i})(2-\mathrm{i})$ y esa es la descomposición, con normas 5 y 5 cuyo producto es $N(5) = 25$. La diferencia entre los dos anillos se mide con un número, el *número de clases*: vale 1 para $\mathbb{Z}[\mathrm{i}]$ y 2 para $\mathbb{Z}[\sqrt{-5}]$, y el 1 es precisamente la condición de que haya factorización única.

Cuidado

[Irreducible y primo dejan de ser lo mismo] En $\mathbb{Z}$ se usan como sinónimos, pero son dos propiedades distintas: irreducible significa "no se puede partir" y primo significa "si divide a un producto, divide a un factor". Primo siempre implica irreducible; el recíproco solo vale en dominios de factorización única. En $\mathbb{Z}[\sqrt{-5}]$, el número 2 es irreducible pero *no* primo: divide a $6 = (1+\sqrt{-5})(1-\sqrt{-5})$ sin dividir a ninguno de los dos factores, porque los cocientes no son enteros del anillo. Es exactamente ahí donde se rompe la unicidad, y por eso Dedekind trasladó el juego a los ideales, donde el ideal primo se define directamente por la segunda propiedad.

El contraejemplo se disuelve: en $\mathbb{Z}[\sqrt{-5}]$, $(6) = \mathfrak{p}_2^2\,\mathfrak{p}_3\,\bar{\mathfrak{p}}_3$ con $\mathfrak{p}_2 = (2, 1+\sqrt{-5})$, etc.: las dos factorizaciones $2 \cdot 3 = (1+\sqrt{-5})(1-\sqrt{-5})$ son dos maneras de agrupar los *mismos* cuatro primos ideales. La medida del fallo elemento/ideal es el *grupo de clases* (finito: teorema de Minkowski), cuya trivialidad equivale a DFU —para $\mathbb{Q}(\sqrt{-d})$ ocurre solo para nueve valores de d (el teorema de Heegner–Baker–Stark que explica el casi-entero $e^{\pi\sqrt{163}}$)—. Con Dedekind, la aritmética del capítulo culmina donde empieza la teoría de números moderna.

25.18. Ronda de temarios: nilpotencia y automorfismos

Del temario del curso básico de álgebra (Cinvestav) tomamos los dos temas de grupos que aún no habíamos desarrollado: grupos nilpotentes y el grupo de automorfismos.

25.18.1. Grupos nilpotentes

Idea clave

Intuición primero. Si "abeliano" significa que todo conmuta y "soluble" que se puede desmontar en capas conmutativas, *nilpotente* es el escalón intermedio: no solo se desmonta en capas abelianas, sino que cada capa está en el *centro* del piso siguiente, es decir, conmuta con todo lo que tiene encima. Es una abelianidad que se propaga hacia arriba de forma ordenada, en vez de aparecer a trozos. La imagen final del capítulo es la mejor: un grupo nilpotente finito *es* el producto de sus piezas de cada primo, igual que un entero es el producto de sus potencias de primos. Nilpotente = "se factoriza como un número".

Definición 25.116. La *serie central ascendente* de G es $Z_0 = \{e\} \subseteq Z_1 = Z(G) \subseteq Z_2 \subseteq \cdots$, donde Z_{i+1} es la preimagen en G del centro de G/Z_i. G es *nilpotente* si $Z_c = G$ para algún c (el menor, su *clase*).

Abeliano = nilpotente de clase 1. La nilpotencia es "abelianidad por capas": cada piso es central en el cociente por el anterior. Todo nilpotente es soluble (la serie central es en particular una serie con cocientes abelianos), pero no al revés: S_3 es soluble y $Z(S_3) = \{e\}$, así que su serie central no arranca.

Lema 25.117. *Si $G/Z(G)$ es cíclico, G es abeliano.*

Demostración. Si $G/Z = \langle gZ \rangle$, todo elemento es $g^a z$ con z central; dos tales conmutan: $(g^a z)(g^b w) = g^{a+b} zw = (g^b w)(g^a z)$. □ □

Idea clave

Qué dice esto en cristiano. El centro $Z(G)$ recoge a los elementos que conmutan con todos. El cociente $G/Z(G)$ mide, por tanto, "cuánta no conmutatividad queda" una vez retirado lo que ya conmutaba. El lema dice que esa cantidad restante nunca puede ser un simple círculo: si lo fuera, el grupo entero habría sido abeliano desde el principio, y entonces el centro habría sido todo G y el cociente trivial. Dicho de otro modo: $G/Z(G)$ nunca es cíclico no trivial. Es un lema minúsculo pero de los más rentables del álgebra, porque prohíbe de un plumazo montones de configuraciones imposibles.

Ejemplo 25.118 (El lema en funcionamiento y su mejor consecuencia)**. (a) D_4 lo cumple.** El centro de D_4 tiene dos elementos (la identidad y el giro de media vuelta), luego $|D_4/Z| = 8/2 = 4$. El lema exige que ese cociente de orden 4 no sea cíclico, y en efecto es $\mathbb{Z}_2 \times \mathbb{Z}_2$, no $\mathbb{Z}_4$. Si hubiera salido $\mathbb{Z}_4$ tendríamos una contradicción, porque D_4 no es abeliano.

(b) Los casos con centro trivial. En S_3 el centro es $\{e\}$, así que $S_3/Z \cong S_3$, de orden 6 y desde luego no cíclico. Lo mismo en A_4: centro trivial y cociente de orden 12 no cíclico. En el otro extremo, $\mathbb{Z}_6$ es abeliano, su centro es todo el grupo y el cociente es trivial, que sí es cíclico; coherente, porque el lema solo prohíbe el caso cíclico *no trivial*.

(c) La consecuencia estrella: todo grupo de orden p^2 es abeliano. Sea $|G| = p^2$. Todo p-grupo tiene centro no trivial, luego $|Z|$ vale p o p^2. Si vale p^2, el grupo es abeliano y hemos terminado. Si valiera p, el cociente G/Z tendría orden p, que es primo y por tanto cíclico; el lema obligaría entonces a que G fuese abeliano, con lo que $Z = G$ y $|Z| = p^2$, contradicción. Solo queda la primera opción.

(d) Qué se acaba de demostrar. Todos los grupos de orden 4, 9, 25, 49… son abelianos, y de hecho solo hay dos de cada orden p^2: $\mathbb{Z}_{p^2}$ y $\mathbb{Z}_p \times \mathbb{Z}_p$. La clasificación completa de un caso infinito de tamaños ha salido de un lema de una línea. Nótese que para orden p^3 ya falla: D_4 y Q_8 tienen orden $8 = 2^3$ y no son abelianos.

Cuidado

[El lema no dice que el centro sea grande] Es fácil leer el lema como si garantizara que $Z(G)$ tiene muchos elementos, y no es así: S_3, A_4 y S_n para $n \geq 3$ tienen centro trivial. Lo único que se afirma es sobre la *forma* del cociente, no sobre su tamaño. Segundo aviso: el enunciado recíproco es falso en el sentido obvio; que $G/Z(G)$ no sea cíclico no implica nada sobre si G es abeliano o no, porque si G es abeliano el cociente es trivial, que *sí* es cíclico. El lema se usa siempre en la dirección contrarrecíproca: si G no es abeliano, entonces $G/Z(G)$ no es cíclico.

Teorema 25.119. *Todo p-grupo finito es nilpotente.*

Demostración. Ya probamos con la ecuación de clases que $Z(G) \neq \{e\}$ para todo p-grupo no trivial. Entonces G/Z_1 es un p-grupo menor: su centro tampoco es trivial, así que $Z_2 \supsetneq Z_1$; la serie central crece estrictamente hasta agotar G (finitud). El lema entra en el último paso: si G/Z_{c-1} fuera cíclico no trivial no habría problema, pero garantiza además que un p-grupo de orden p^2 es abeliano (G/Z tendría orden p, cíclico). □ □

Teorema 25.120 (Caracterización de los nilpotentes finitos)**.** *Para G finito son equivalentes: (i) G es nilpotente; (ii) todo subgrupo de Sylow es normal; (iii) G es el producto directo de sus subgrupos de Sylow.*

Esquema. (iii)⇒(i): producto de nilpotentes es nilpotente (los centros se multiplican). (ii)⇒(iii): Sylows normales de órdenes coprimos se intersecan trivialmente, conmutan entre sí ($[P,Q] \subseteq P \cap Q = \{e\}$) y sus órdenes multiplican al de G. (i)⇒(ii) usa el *argumento del normalizador*: en un nilpotente todo subgrupo propio está estrictamente contenido en su normalizador (subir por la serie central); si un Sylow P no fuera normal, $N(P)$ sería propio, pero $N(N(P)) = N(P)$ (un Sylow normal en $N(P)$ es característico en él), contradicción. □ □

Idea clave

Qué dice esto en cristiano. Nilpotente es una condición más fuerte que soluble. Soluble decía "se puede desmontar en capas abelianas"; nilpotente dice además que esas capas se pueden tomar *centrales*, es decir, que en cada paso lo que se quita conmuta con todo lo que queda. La caracterización da la versión útil: un grupo finito es nilpotente exactamente cuando sus subgrupos de Sylow no se estorban entre sí, y entonces el grupo es simplemente el producto de ellos. Dicho de otro modo, un grupo nilpotente es un montón de p-grupos apilados sin interacción.

Ejemplo 25.121 (Quién es nilpotente y quién no)**. (a) Los p-grupos siempre lo son.** Todo grupo de orden 4, 8, 9, 16 o 27 es nilpotente. La razón es que un p-grupo tiene siempre centro no trivial, luego la serie central ascendente arranca de verdad y no se queda parada en el trivial. Para D_4, de orden 8: el centro tiene dos elementos, el cociente D_4/Z tiene orden 4 y es abeliano, luego $Z_2 = D_4$ y la serie termina en dos pasos.

(b) S_3 es soluble pero no nilpotente. Su centro es trivial, así que la serie central

ascendente se queda atascada en $\{e\}$ desde el primer paso y nunca llega a S_3. Mirándolo por la caracterización de Sylow: el 3-Sylow es A_3, que sí es normal, pero hay *tres* subgrupos de Sylow de orden 2 (uno por cada trasposición), y tres es más de uno, luego no son normales.

(c) A_4 tampoco. Tiene ocho elementos de orden 3, que se reparten en cuatro subgrupos de orden 3 distintos. Al haber cuatro 3-Sylow en lugar de uno, ninguno es normal y el grupo no es nilpotente, aunque sí es soluble. La cadena de implicaciones es estricta: cíclico $\Rightarrow$ abeliano $\Rightarrow$ nilpotente $\Rightarrow$ soluble, y ninguna flecha se puede invertir.

(d) La descomposición cuando sí lo es. $\mathbb{Z}_{12}$ es abeliano, luego nilpotente, y en efecto es el producto directo de sus Sylow: el 2-Sylow tiene orden 4, el 3-Sylow orden 3, y $\mathbb{Z}_{12} \cong \mathbb{Z}_4 \times \mathbb{Z}_3$. Esa es la forma general que garantiza la caracterización, y es la razón de que estudiar grupos nilpotentes se reduzca a estudiar p-grupos.

Cuidado

[Los contraejemplos que delimitan cada implicación] D_4 es nilpotente y no abeliano: nilpotente no significa conmutativo. S_3 es soluble y no nilpotente. Y A_5 no es ni lo uno ni lo otro. Segundo aviso: la condición "todo Sylow es normal" hay que comprobarla para *todos* los primos que dividen al orden, no solo para uno; en S_3 el 3-Sylow es normal y aun así el grupo falla, porque el 2-Sylow no lo es. Y tercero: la longitud de la serie central (la *clase de nilpotencia*) no es lo mismo que el número de primos; D_4 tiene clase 2 y un solo primo.

Observación 25.122. La jerarquía completa, con testigos mínimos: abelianos $\subsetneq$ nilpotentes $\subsetneq$ solubles $\subsetneq$ todos; Q_8 es nilpotente no abeliano, S_3 soluble no nilpotente, A_5 no soluble. Los nilpotentes finitos son "productos de p-grupos": su aritmética se estudia un primo a la vez, como los módulos sobre DIP del capítulo se estudiaban un primo a la vez. El *subgrupo de Frattini* $\Phi(G)$ (intersección de los maximales) es siempre nilpotente y consiste en los "no generadores": elementos que pueden borrarse de cualquier conjunto generador; para p-grupos, $G/\Phi(G)$ es un espacio vectorial sobre $\mathbb{F}_p$ y el número mínimo de generadores es su dimensión (Burnside): el álgebra lineal del capítulo 7 gobernando la teoría de grupos.

25.18.2. Automorfismos interiores y exteriores

Idea clave

Intuición primero. Un automorfismo es una simetría *del grupo mismo*: una forma de renombrar sus elementos sin que la tabla de multiplicar cambie. Siempre hay unas cuantas gratis: conjugar por un elemento fijo ($x \mapsto gxg^{-1}$) es cambiar de punto de vista, como describir las rotaciones de un cubo desde otra esquina. Esas son las *interiores*, y solo son interesantes en la medida en que el centro sea pequeño: $\mathrm{Inn}(G) \cong G/Z(G)$ dice que las simetrías gratis son el grupo menos lo que ya conmutaba con todo. Las *exteriores* son las sorpresas: renombramientos que no provienen de mirar desde otro sitio. Que S_6 tenga una, y sea el único S_n que la tiene, es una de las rarezas más bonitas del álgebra.

Teorema 25.123. *La aplicación $g \mapsto (\varphi_g : x \mapsto gxg^{-1})$ es un homomorfismo $G \to \mathrm{Aut}(G)$ con núcleo $Z(G)$ e imagen $\mathrm{Inn}(G)$; por el primer teorema de isomorfismo,*

$$\mathrm{Inn}(G) \cong G/Z(G).$$

Además $\mathrm{Inn}(G) \trianglelefteq \mathrm{Aut}(G)$, y el cociente $\mathrm{Out}(G) = \mathrm{Aut}(G)/\mathrm{Inn}(G)$ mide las simetrías "inesperadas" de G.

Demostración. $\varphi_g \circ \varphi_h = \varphi_{gh}$ es un cálculo directo; el núcleo es $\{g : gxg^{-1} = x\ \forall x\} = Z(G)$. La normalidad: para $\alpha \in \mathrm{Aut}(G)$, $\alpha\varphi_g\alpha^{-1} = \varphi_{\alpha(g)}$ (verificar sobre un x arbitrario). □ □

Ejemplo 25.124. $\mathrm{Aut}(\mathbb{Z}_n) \cong (\mathbb{Z}_n)^\times$ ya nos acompañaba desde los ciclotómicos. Para S_n: toda permutación de las trasposiciones inducida por un automorfismo debe preservar el tipo cíclico, y un análisis de conteo muestra $\mathrm{Aut}(S_n) = \mathrm{Inn}(S_n) \cong S_n$ para todo $n \neq 2, 6$. La excepción $n = 6$ es real: S_6 tiene un automorfismo exterior (único caso en toda la familia), descubierto vía las dos acciones no equivalentes de S_6 sobre 6 objetos; está emparentado con las estructuras excepcionales que salpican la matemática (los 6 pentágonos sintéticos, la combinatoria del icosaedro). Que la excepción ocurra exactamente una vez, y en $n = 6$, es de esas rarezas que el formalismo no hace esperar y la matemática entrega.

25.19. Cierre definitivo: noetherianidad, Zorn y localización

Tres herramientas de anillos que faltaban para que el capítulo alcance el nivel de entrada de la geometría algebraica y del álgebra conmutativa.

25.19.1. Anillos noetherianos y el teorema de la base de Hilbert

Teorema 25.125 (Caracterización)**.** *Para un anillo conmutativo R son equivalentes: (i) todo ideal es finitamente generado; (ii) toda cadena ascendente de ideales $I_1 \subseteq I_2 \subseteq \cdots$ se estabiliza; (iii) toda familia no vacía de ideales tiene un elemento maximal. Tal R es* noetheriano*.*

Demostración. (i)⇒(ii): la unión de la cadena es un ideal, generado por finitos elementos, todos presentes en algún I_N: la cadena para ahí. (ii)⇒(iii): si no hubiera maximal, se construye una cadena estrictamente creciente infinita (elección recursiva). (iii)⇒(i): dado I, el conjunto de subideales finitamente generados de I tiene un maximal J; si $J \neq I$, añadir un elemento faltante contradice la maximalidad. □ □

Teorema 25.126 (Base de Hilbert)**.** *Si R es noetheriano, $R[x]$ es noetheriano. En consecuencia $K[x_1, \ldots, x_n]$ y $\mathbb{Z}[x_1, \ldots, x_n]$ lo son.*

Idea de la demostración. Dado un ideal $I \subseteq R[x]$, mírense los *coeficientes líderes* de sus elementos de cada grado: forman una cadena ascendente de ideales de R, que se estabiliza; tomando generadores de esos finitos ideales de líderes y polinomios que los realicen, todo elemento de I se reduce por división (restar múltiplos para matar el término líder, como en el algoritmo de Euclides) a combinación de los elegidos. *Controlar los líderes controla todo*: la misma idea que la división con resto, industrializada. □ □

Idea clave

Qué dice esto en cristiano. Un anillo noetheriano es uno donde no se puede crecer para siempre: si vas construyendo ideales cada vez más grandes, antes o después te atascas y dejas de avanzar. Las tres condiciones equivalentes son tres maneras de decir lo mismo (todo ideal se genera con finitos elementos, las cadenas crecientes se estabilizan, toda familia tiene un maximal), y en la práctica se usa la que resulte más cómoda. El teorema de la base de Hilbert añade la noticia decisiva: esta propiedad sobrevive al añadir variables, de modo que los anillos de polinomios con los que se trabaja de verdad la tienen todos.

Ejemplo 25.127 (Cadenas que se paran y anillos que aguantan)**. (a) $\mathbb{Z}$ es noetheriano.** Todo ideal de $\mathbb{Z}$ es de la forma (n), generado por un solo elemento. Una cadena creciente como $(12) \subseteq (6) \subseteq (3) \subseteq (1)$ corresponde a divisores cada vez menores del 12, y como 12 tiene un número finito de divisores, la cadena se para necesariamente: aquí, en cuatro pasos. Crecer en la relación de contenido es decrecer en la de divisibilidad.

(b) Finitamente generado no es lo mismo que principal. En $K[x, y]$ el ideal (x, y) (los polinomios sin término independiente) *no* es principal: ningún polinomio solo genera a la vez x e y. Pero sí es finitamente generado, con dos generadores. Por eso $K[x, y]$ es noetheriano sin ser dominio de ideales principales: la noetherianidad es una condición mucho más débil y mucho más frecuente.

(c) La base de Hilbert en cadena. K es un cuerpo, luego trivialmente noetheriano (sus únicos ideales son 0 y K). Aplicando el teorema, $K[x]$ es noetheriano; aplicándolo otra vez, $K[x, y]$; y repitiendo n veces, $K[x_1, \ldots, x_n]$. Lo mismo partiendo de $\mathbb{Z}$ da $\mathbb{Z}[x_1, \ldots, x_n]$. Un teorema demostrado una vez cubre así infinitos casos.

(d) Qué significa geométricamente. Una curva o superficie definida por un sistema de ecuaciones polinómicas, aunque el sistema tenga infinitas ecuaciones, se puede definir con un número *finito* de ellas: el ideal que generan está finitamente generado. Por ejemplo, todos los polinomios que se anulan sobre la parábola $y = x^2$ forman el ideal $(y - x^2)$, con un único generador; el polinomio $y^2 - x^4 = (y - x^2)(y + x^2)$ pertenece a él, como se comprueba dividiendo.

Cuidado

[Contraejemplo: con infinitas variables se rompe] El anillo $K[x_1, x_2, x_3, \ldots]$ con infinitas variables *no* es noetheriano: la cadena $(x_1) \subsetneq (x_1, x_2) \subsetneq (x_1, x_2, x_3) \subsetneq \cdots$ crece indefinidamente y nunca se estabiliza, y el ideal formado por todos los x_i no admite un sistema finito de generadores. Obsérvese que la base de Hilbert añade *una* variable cada vez, luego solo cubre los casos con finitas. Segundo aviso: noetheriano no implica factorización única. El anillo $\mathbb{Z}[\sqrt{-5}]$ es noetheriano y en él $6 = 2 \cdot 3 = (1+\sqrt{-5})(1-\sqrt{-5})$ son dos factorizaciones distintas en irreducibles. Lo que sí se recupera, y es el teorema de Dedekind, es la factorización única en *ideales* primos.

Observación 25.128. El teorema (1888) escandalizó a su época: prueba que todo sistema de ecuaciones polinomiales en n variables equivale a uno *finito*, sin exhibirlo ("¡esto no es

matemática, es teología!", protestó Gordan). La versión efectiva llegó un siglo después: las *bases de Gröbner* (Buchberger, 1965) eligen los generadores con líderes ordenados de modo que la división multivariable tenga resto único, y convierten en algoritmo lo que Hilbert dejó en existencia: hoy son el motor de todo sistema de álgebra computacional que resuelva sistemas polinomiales (el análogo multivariable exacto de la escalonada de Gauss del capítulo 7).

25.19.2. Zorn y los ideales maximales

Teorema 25.129. *Todo anillo conmutativo con* $1 \neq 0$ *tiene un ideal maximal; más aún, todo ideal propio está contenido en uno maximal.*

Demostración. La familia de ideales propios que contienen al dado, ordenada por inclusión: toda cadena tiene cota superior (la unión: es ideal, y es *propia* porque el 1 no aparece en ningún eslabón: aquí se usa que ser propio es "no contener al 1", una condición finitista). El lema de Zorn —equivalente al axioma de elección, el mismo que fabricó el conjunto de Vitali y las bases de trascendencia— entrega el maximal. □ □

Observación 25.130. Consecuencias inmediatas con lo ya demostrado: como $R/\mathfrak{m}$ es cuerpo para $\mathfrak{m}$ maximal, *todo anillo tiene cocientes que son cuerpos*: así se fabrican los cuerpos residuales de la teoría de números ($\mathbb{Z} \to \mathbb{F}_p$) y de la geometría ($K[x_1, \dots, x_n] \to K$, evaluar en un punto). El diccionario del Nullstellensatz de Hilbert, que el capítulo dejó señalado, empieza exactamente aquí: en $\mathbb{C}[x_1, \dots, x_n]$ los ideales maximales *son* los puntos de $\mathbb{C}^n$: el álgebra de los ideales es la geometría de los lugares.

25.19.3. Localización: mirar con lupa

Definición 25.131. Si $S \subseteq R$ es multiplicativo ($1 \in S$, cerrado bajo producto), el *localizado* $S^{-1}R$ consta de fracciones $\frac{r}{s}$ con $r \in R$, $s \in S$, identificando $\frac{r}{s} = \frac{r'}{s'}$ si $t(rs' - r's) = 0$ para algún $t \in S$: la construcción del cuerpo de fracciones, ejecutada parcialmente.

Ejemplo 25.132. Tres lupas: (a) $S = R \setminus \{0\}$ en un dominio: el cuerpo de fracciones ya construido. (b) $S = \{1, f, f^2, \dots\}$: se invierte solo f: el anillo de "funciones definidas donde $f \neq 0$". (c) La más importante: $S = R \setminus \mathfrak{p}$ con $\mathfrak{p}$ primo (multiplicativo *por definición* de primo: he aquí por qué los primos son la noción correcta): el localizado $R_{\mathfrak{p}}$ tiene un *único* ideal maximal, $\mathfrak{p}R_{\mathfrak{p}}$: es un *anillo local*, la lupa que solo ve lo que pasa "cerca de $\mathfrak{p}$". En $\mathbb{Z}_{(p)}$ (fracciones con denominador ajeno a p) la aritmética entera se reduce a la de un solo primo: es el paso intermedio conceptual hacia los p-ádicos del capítulo 26, y la razón del eslogan que organiza el álgebra moderna: *propiedad global = propiedad en cada localización*: verificar algo de $\mathbb{Z}$ es verificarlo primo por primo, tal como los módulos sobre DIP se estudiaron primo a primo.

25.20. Segunda ronda: Pólya, idempotentes y $\mathbb{F}_{16}$

25.20.1. El cubo coloreado

¿Cuántos cubos con 6 caras y 3 colores, salvo rotación? Por Burnside sobre las 24 rotaciones (identidad 3^6; caras $\pm 90°$: $6 \cdot 3^3$; caras $180°$: $3 \cdot 3^4$; vértices: $8 \cdot 3^2$; aristas: $6 \cdot 3^3$):

$$\frac{729 + 162 + 243 + 72 + 162}{24} = \frac{1368}{24} = 57.$$

Imposible a ojo, rutinario con la acción; el teorema de Pólya lo empaqueta en una generatriz y cuenta isómeros químicos desde 1937.

25.20.2. Idempotentes: el CRT hecho elementos

En $\mathbb{Z}_6$: $e_1 = 3$, $e_2 = 4$ cumplen $e_i^2 = e_i$, $e_1 e_2 = 0$, $e_1 + e_2 = 1$: las coordenadas de $\mathbb{Z}_6 \cong \mathbb{Z}_2 \times \mathbb{Z}_3$ (y los coeficientes de Bézout del CRT). Toda descomposición de un anillo = sistema de idempotentes ortogonales: la versión en anillos de las proyecciones espectrales del capítulo 7 —el mismo Bézout demuestra CRT y descomposición primaria—.

25.20.3. $\mathbb{F}_{16}$ completo

$x^4 + x + 1$ irreducible sobre $\mathbb{F}_2$: $\mathbb{F}_{16} = \mathbb{F}_2(\alpha)$, $\alpha^4 = \alpha + 1$; $\mathbb{F}_{16}^* \cong \mathbb{Z}_{15}$ con α generador. Subcuerpos: $\mathbb{F}_2 \subset \mathbb{F}_4 \subset \mathbb{F}_{16}$ y *no* hay $\mathbb{F}_8$ ($3 \nmid 4$); Frobenius de orden 4, $\mathbb{F}_4$ = fijo de su cuadrado. El capítulo 26 en dieciséis elementos —y el alfabeto de los Reed–Solomon cortos—.

25.21. Coda: orden 12 y los teoremas de solubilidad masiva

Con Sylow y semidirectos, la clasificación de orden 12: $n_3 \in \{1, 4\}$ y $n_2 \in \{1, 3\}$; el análisis de casos produce exactamente cinco grupos: $\mathbb{Z}_{12}$, $\mathbb{Z}_2 \times \mathbb{Z}_6$, A_4 (el de las rotaciones del tetraedro: $n_3 = 4$), D_6 y el dicíclico Dic_3. Nótese que A_4, de orden 12, *no* tiene subgrupo de orden 6: el recíproco de Lagrange es falso, y este es el contraejemplo mínimo. Dos hitos de solubilidad que enmarcan el capítulo: el teorema $p^a q^b$ de *Burnside* (1904): todo grupo de orden $p^a q^b$ es soluble —demostrado con caracteres: la primera gran victoria de las representaciones—; y *Feit–Thompson* (1963): **todo grupo de orden impar es soluble** (255 páginas: el disparo de salida de la clasificación). Corolario conjunto: todo simple no abeliano tiene orden par y divisible por al menos tres primos: $60 = 2^2 \cdot 3 \cdot 5$ es el mínimo posible, y es A_5: nuestro viejo conocido no es un accidente sino el primer permitido por la aritmética.

25.22. Temas extra y ejemplos desarrollados

25.22.1. Burnside

#órbitas $= \frac{1}{|G|} \sum_g |\operatorname{Fix}(g)|$ (doble conteo + órbita–estabilizador). Collares de 6 cuentas y 2 colores: $\frac{84}{6} = 14$ con rotaciones; $\frac{84+72}{12} = 13$ con volteos. Pólya lo industrializó para isómeros.

25.22.2. RSA completo

$n = pq$, e coprimo con $\varphi(n)$, $d = e^{-1}$; $c = m^e$, $m = c^d$ por Euler (=Lagrange en $(\mathbb{Z}_n)^\times$). Miniatura: $n = 55$, $e = 3$, $d = 27$: $7 \mapsto 13 \mapsto 7$ ✓ (con CRT). La banca en línea apuesta a que factorizar es duro.

25.22.3. El retículo de $\mathbb{Z}_{12}$

Subgrupos $\langle d \rangle$, $d \mid 12$: el retículo de divisores (mcd/mcm); modelo de la correspondencia de Galois y de la teoría musical (el ciclo de quintas genera: $\gcd(7, 12) = 1$).

25.22.4. Los grupos simples finitos

Jordan–Hölder hace átomos; la clasificación (el teorema más largo de la historia): cíclicos primos, A_n, tipo Lie, y 26 esporádicos hasta el Monstruo ($\sim 8 \times 10^{53}$, "moonshine"). Nuestro A_5 es el átomo exótico más pequeño: el responsable de la quíntica.

25.23. Módulos sobre un DIP: la forma normal de Smith y el teorema de estructura

El teorema de estructura de los módulos finitamente generados sobre un dominio de ideales principales es el resultado más reutilizado del libro: de él salen la clasificación de los grupos abelianos finitos y las formas canónicas racional y de Jordan. Se enunció en ocho líneas y se remitió a "la forma normal de Smith" sin definirla. Toca cerrarlo.

Definición 25.133 (Módulo)**.** Un **R-módulo** es un grupo abeliano M con una multiplicación $R \times M \to M$ que cumple los axiomas de espacio vectorial. La única diferencia con un espacio vectorial es que R no tiene por qué ser cuerpo, y esa diferencia lo cambia todo: puede haber **torsión**, elementos $m \neq 0$ con $rm = 0$.

Idea clave

Los módulos unifican tres cosas que parecían distintas. Un $\mathbb{Z}$-módulo *es* un grupo abeliano, ni más ni menos: la multiplicación por enteros ya estaba definida. Un K-módulo es un espacio vectorial. Y un $K[x]$-módulo es un espacio vectorial con un endomorfismo distinguido —el que "multiplica por x"—. Un solo teorema de clasificación sobre estos tres casos produce la clasificación de los abelianos finitos, la de los espacios vectoriales por dimensión, y la forma de Jordan.

Definición 25.134 (Forma normal de Smith)**.** Toda matriz A con entradas en un DIP R admite U, V invertibles con

$$UAV = \operatorname{diag}(d_1, d_2, \ldots, d_r, 0, \ldots, 0), \qquad d_1 \mid d_2 \mid \cdots \mid d_r,$$

y los d_i —los **factores invariantes**— son únicos salvo unidades.

Regla: el algoritmo

1. Con operaciones de fila y columna, lleva a la posición $(1,1)$ un elemento de norma mínima no nula.
2. Usa la división euclídea para anular el resto de la primera fila y columna; si algún resto no es cero, ponlo en $(1,1)$ y repite. El proceso termina porque la norma decrece estrictamente.
3. Si d_1 no divide a alguna entrada del bloque restante, súmale esa fila a la primera y vuelve al paso 1. Al terminar, d_1 divide a todo.
4. Recurre sobre el bloque $(n-1)\times(n-1)$.

Los $d_1 \cdots d_k$ son los máximos comunes divisores de los menores de orden k, lo que demuestra la unicidad de golpe.

Teorema 25.135 (Estructura)**.** *Todo módulo finitamente generado sobre un DIP es*

$$M \;\cong\; R^s \;\oplus\; R/(d_1) \;\oplus \cdots \oplus\; R/(d_r), \qquad d_1 \mid \cdots \mid d_r,$$

y la descomposición es única. El s es el rango *y los d_i los factores invariantes de la torsión.*

Cómo Smith lo demuestra. M generado por n elementos es cociente R^n/K, y K —submódulo de un libre sobre un DIP— es libre de rango $m \le n$. Escríbase la matriz A de la inclusión $R^m \hookrightarrow R^n$ en unas bases; cambiar de base equivale a multiplicar por matrices invertibles, así que se puede tomar A en forma de Smith. En esas bases el cociente se parte en sumandos $R/(d_i)$ y copias libres. □

Ejemplo 25.136 (Un cociente de $\mathbb{Z}^2$, identificado)**.** Sea

$$M = \mathbb{Z}^2/\langle (2,4),\ (6,8)\rangle.$$

La matriz de relaciones es $A = \left(\begin{smallmatrix} 2 & 6 \\ 4 & 8 \end{smallmatrix}\right)$. Su forma de Smith:

$$\begin{pmatrix} 2 & 6 \\ 4 & 8 \end{pmatrix} \rightsquigarrow \begin{pmatrix} 2 & 0 \\ 4 & -4 \end{pmatrix} \rightsquigarrow \begin{pmatrix} 2 & 0 \\ 0 & -4 \end{pmatrix} \rightsquigarrow \operatorname{diag}(2,4),$$

restando tres veces la primera columna de la segunda, luego dos veces la primera fila de la segunda, y ajustando el signo. Luego

$$M \cong \mathbb{Z}_2 \oplus \mathbb{Z}_4.$$

Comprobación de orden: $|\det A| = |16-24| = 8 = 2\cdot 4$. ✓ Y la unicidad de Smith descarta las otras dos posibilidades de orden 8: no es $\mathbb{Z}_8$ ni $\mathbb{Z}_2^3$. Obsérvese que $d_1 = 2$ es el máximo común divisor de las entradas y $d_1 d_2 = 8$ es el del único menor de orden 2, tal como afirma la regla.

Observación 25.137 (Los dos corolarios que el libro usaba). **(a) Grupos abelianos finitos.** Con $R = \mathbb{Z}$ y M finito, el rango es 0 y queda $M \cong \mathbb{Z}_{d_1} \oplus \cdots \oplus \mathbb{Z}_{d_r}$: exactamente la clasificación por factores invariantes.

(b) Formas canónicas. Con $R = K[x]$ y $M = V$ visto como $K[x]$-módulo vía $x \cdot v = T(v)$, la torsión lo es todo (porque V tiene dimensión finita) y el teorema da

$$V \cong K[x]/(p_1) \oplus \cdots \oplus K[x]/(p_r), \qquad p_1 \mid \cdots \mid p_r.$$

Escribiendo cada sumando en la base $\{1, x, \ldots, x^{\deg p_i - 1}\}$ sale la matriz compañera: es la *forma racional*. Y si K es algebraicamente cerrado se refina a potencias de $(x - \lambda)$, y cada sumando da un bloque de Jordan. Las dos formas canónicas del capítulo de álgebra lineal son el mismo teorema leído dos veces.

25.24. Jordan–Hölder demostrado, álgebras de división y el teorema de Wedderburn

Los primos de un grupo

Definición 25.138. Una **serie de composición** de G es una cadena $\{e\} = G_0 \lhd G_1 \lhd \cdots \lhd G_n = G$ con todos los cocientes G_{i+1}/G_i *simples*.

Teorema 25.139 (Jordan–Hölder)**.** *Dos series de composición de un mismo grupo finito tienen la misma longitud y sus factores coinciden salvo isomorfismo y reordenación.*

Lema 25.140 (Mariposa de Zassenhaus)**.** *Si $A \lhd A^*$ y $B \lhd B^*$ son subgrupos de G, entonces*

$$\frac{A(A^* \cap B^*)}{A(A^* \cap B)} \;\cong\; \frac{B(A^* \cap B^*)}{B(A \cap B^*)}.$$

De la mariposa sale el **teorema de refinamiento de Schreier** —dos series cualesquiera admiten refinamientos equivalentes— y de él Jordan–Hölder de inmediato: una serie de composición no admite refinamientos propios, así que los refinamientos equivalentes tienen que ser las series originales.

Ejemplo 25.141 (La misma factorización por dos caminos)**.** En $\mathbb{Z}_{12}$:

$$\begin{aligned} 0 \lhd \langle 6\rangle \lhd \langle 2\rangle \lhd \mathbb{Z}_{12} &: \quad \text{factores } \mathbb{Z}_2,\ \mathbb{Z}_3,\ \mathbb{Z}_2, \\ 0 \lhd \langle 4\rangle \lhd \langle 2\rangle \lhd \mathbb{Z}_{12} &: \quad \text{factores } \mathbb{Z}_3,\ \mathbb{Z}_2,\ \mathbb{Z}_2. \end{aligned}$$

Mismo multiconjunto $\{\mathbb{Z}_2, \mathbb{Z}_2, \mathbb{Z}_3\}$, distinto orden. Exactamente lo que el teorema predice, y exactamente el comportamiento de la factorización en primos de $12 = 2^2 \cdot 3$.

Idea clave

Jordan–Hölder es lo que hace que la palabra *resoluble* esté bien definida: si una serie de composición tiene todos los factores abelianos, todas la tienen, porque los factores son los mismos. Sin este teorema, la afirmación central de la teoría de Galois —"la ecuación se resuelve por radicales si y solo si su grupo es resoluble"— dependería de qué serie eligiera cada uno.

Y explica el sentido de la clasificación de los grupos simples finitos, terminada en 2004 tras diez mil páginas: los simples son los "primos" de la teoría de grupos, y Jordan–Hölder garantiza que todo grupo finito se factoriza en ellos de manera única. Lo que *no* está resuelto es el problema de la extensión: saber los factores no determina el grupo, como muestran $\mathbb{Z}_4$ y $\mathbb{Z}_2 \times \mathbb{Z}_2$, que tienen los mismos.

Álgebras y el teorema pequeño de Wedderburn

Definición 25.142. Una **K-álgebra** es un anillo A que además es K-espacio vectorial, con $\lambda(ab) = (\lambda a)b = a(\lambda b)$. Un **anillo de división** es un anillo con unidad en el que todo elemento no nulo es invertible, sin exigir conmutatividad.

Ejemplo 25.143 (Los cuaterniones)**.** $\mathbb{H} = \{a + bi + cj + dk\}$ sobre $\mathbb{R}$, con $i^2 = j^2 = k^2 = ijk = -1$. Es un álgebra de dimensión 4, asociativa y *no conmutativa* ($ij = k$ pero $ji = -k$). La norma $N(q) = a^2 + b^2 + c^2 + d^2$ es multiplicativa, y

$$q^{-1} = \frac{\bar{q}}{N(q)},$$

así que todo elemento no nulo es invertible: $\mathbb{H}$ es un anillo de división. Es el único, junto a $\mathbb{R}$ y $\mathbb{C}$, de dimensión finita sobre $\mathbb{R}$ y asociativo (teorema de Frobenius, 1877), y de él salen las rotaciones del espacio sin las singularidades de los ángulos de Euler.

Teorema 25.144 (Wedderburn, 1905)**.** *Todo anillo de división* finito *es conmutativo, es decir, es un cuerpo.*

Esquema. Sea D finito de división y Z su centro, un cuerpo con q elementos. Entonces D es un Z-espacio de dimensión n, luego $|D| = q^n$, y el centralizador de cada x tiene orden q^d con $d \mid n$. La ecuación de clases del grupo multiplicativo da

$$q^n - 1 = (q-1) + \sum_i \frac{q^n - 1}{q^{d_i} - 1}.$$

El polinomio ciclotómico $\Phi_n(q)$ divide a $q^n - 1$ y a cada sumando de la suma, luego divide a $q - 1$. Pero si $n > 1$, $|\Phi_n(q)| = \prod |q - \zeta| > q - 1$ para toda raíz primitiva ζ, porque $q \geq 2$ y las raíces no son reales positivas. Contradicción, luego $n = 1$ y $D = Z$. □

Observación 25.145. El teorema explica una ausencia llamativa: hay cuerpos finitos de todos los órdenes p^n, y no hay *ningún* "cuerpo finito no conmutativo". La no conmutatividad y la finitud son incompatibles en presencia de división. Y con eso queda también justificada la descomposición $\mathbb{C}[G] \cong \bigoplus M_{d_i}(\mathbb{C})$ del capítulo de representaciones: es el teorema *grande* de Wedderburn —toda álgebra semisimple es producto de álgebras de matrices sobre anillos de división— combinado con que sobre $\mathbb{C}$ el único anillo de división de dimensión finita es $\mathbb{C}$ mismo.

25.25. Tres demostraciones que se habían dejado como idea

Lagrange, la buena definición del signo y la cadena euclídeo ⇒ DIP ⇒ DFU aparecieron como esbozos. Los tres son cortos y los tres se usan a diario, aquí y en los capítulos posteriores.

Teorema 25.146 (Lagrange)**.** *Sea G finito y $H \leq G$. Entonces $|G| = [G : H]\,|H|$; en particular $|H|$ divide a $|G|$, y el orden de todo elemento divide al orden del grupo.*

Demostración. Paso 1: las clases laterales son disjuntas o iguales. Si $aH \cap bH \neq \emptyset$, sea $x = ah_1 = bh_2$. Entonces $a = bh_2h_1^{-1}$, luego para todo $h \in H$ se tiene $ah = b(h_2h_1^{-1}h) \in bH$: así $aH \subset bH$, y por simetría $bH \subset aH$.

Paso 2: todas tienen el mismo cardinal. La aplicación $H \to aH$ dada por $h \mapsto ah$ es biyectiva: es sobreyectiva por definición, e inyectiva porque $ah_1 = ah_2$ implica $h_1 = h_2$ al multiplicar por a^{-1}. Luego $|aH| = |H|$.

Paso 3: recuento. Todo $g \in G$ está en la clase gH (pues $g = ge$), de modo que las clases recubren G; por el paso 1 lo hacen sin solaparse, y por el paso 2 todas tienen $|H|$ elementos. Si hay $[G : H]$ clases distintas, $|G| = [G : H]|H|$.

Consecuencia. El orden de g es el cardinal del subgrupo cíclico $\langle g\rangle$, que divide a $|G|$. De ahí, $g^{|G|} = e$ para todo g, y con $G = (\mathbb{Z}/n)^\times$ eso es exactamente el teorema de Euler–Fermat del capítulo 3: un mismo enunciado con dos nombres. □

Teorema 25.147 (El signo está bien definido)**.** *Toda permutación $\sigma \in S_n$ se escribe como producto de transposiciones, y aunque la descomposición no es única, la* paridad *del número de transposiciones sí lo es. En consecuencia* $\operatorname{sgn}(\sigma) = (-1)^k$ *está bien definido, es un homomorfismo $S_n \to \{\pm 1\}$, y $|A_n| = n!/2$ para $n \geq 2$.*

Demostración. Existencia de la descomposición. Todo ciclo se escribe $(a_1a_2\cdots a_m) = (a_1a_m)(a_1a_{m-1})\cdots(a_1a_2)$, y toda permutación es producto de ciclos disjuntos.

Invariancia de la paridad. Considérese el polinomio en n variables

$$\Delta(x_1,\dots,x_n) = \prod_{1\leq i<j\leq n} (x_i - x_j),$$

y hágase actuar σ permutando las variables: $(\sigma\cdot\Delta)(x) = \prod_{i<j}(x_{\sigma(i)} - x_{\sigma(j)})$. Cada factor del producto original reaparece exactamente una vez, posiblemente con el signo cambiado si σ invierte el orden del par, luego $\sigma\cdot\Delta = \pm\Delta$; defínase $\varepsilon(\sigma)$ como ese signo. La definición es intrínseca —no involucra ninguna descomposición—, y es un homomorfismo porque $(\sigma\tau)\cdot\Delta = \sigma\cdot(\tau\cdot\Delta)$.

Falta calcular ε sobre una transposición. Para $\tau = (1\,2)$, el único factor que cambia de signo es $(x_1 - x_2)$: los factores $(x_1 - x_j)$ y $(x_2 - x_j)$ con $j \geq 3$ simplemente se intercambian entre sí, y los que no involucran a 1 ni a 2 quedan fijos. Luego $\varepsilon(\tau) = -1$, y por conjugación lo mismo vale para toda transposición. Por tanto, si σ es producto de k transposiciones, $\varepsilon(\sigma) = (-1)^k$: la paridad de k está determinada por $\varepsilon(\sigma)$, que no depende de la descomposición.

Cardinal de A_n. Para $n \geq 2$, $\varepsilon : S_n \to \{\pm 1\}$ es sobreyectivo (una transposición va a -1), luego $A_n = \ker\varepsilon$ tiene índice 2 y $|A_n| = n!/2$. □

Observación 25.148. Este es el resultado que el capítulo 7 invoca para la unicidad del determinante: la fórmula $\det A = \sum_\sigma \operatorname{sgn}(\sigma) \prod_i a_{i\sigma(i)}$ sólo tiene sentido si el signo está bien definido, y ahora lo está. La demostración con Δ tiene la virtud de que es constructiva y no circular: no presupone el determinante, que es el error en el que se cae al intentar definir el signo como "el determinante de la matriz de permutación".

Teorema 25.149 (La cadena aritmética, completa)**.** *Todo dominio euclídeo es un dominio de ideales principales, y todo dominio de ideales principales es un dominio de factorización única.*

Demostración. Euclídeo ⇒ DIP. Sea $I \neq 0$ un ideal y $b \in I$ de norma euclídea mínima entre los no nulos. Para $a \in I$, la división da $a = qb + r$ con $r = 0$ o $N(r) < N(b)$. Como $r = a - qb \in I$, la minimalidad fuerza $r = 0$: luego $a \in (b)$ e $I = (b)$.

DIP ⇒ las cadenas ascendentes se estabilizan. Sea $I_1 \subset I_2 \subset \cdots$ una cadena de ideales. Su unión I es un ideal, luego $I = (c)$ para algún c; y c está en algún I_N, de donde $I = (c) \subset I_N \subset I$, y la cadena se para.

DIP ⇒ existe factorización. Supóngase que algún a no se factoriza en irreducibles. Entonces a no es irreducible ni unidad, así que $a = a_1 b_1$ con ambos factores no unidades, y al menos uno de ellos —digamos a_1— tampoco se factoriza. Repitiendo, se obtiene una cadena estrictamente creciente $(a) \subsetneq (a_1) \subsetneq (a_2) \subsetneq \cdots$ (estricta porque b_i no es unidad), contradiciendo el paso anterior.

DIP ⇒ irreducible implica primo. Sea p irreducible y supóngase $p \mid ab$ con $p \nmid a$. El ideal (p, a) es principal, digamos (d); entonces $d \mid p$, y como p es irreducible, d es unidad o asociado de p. No puede ser asociado de p, porque $d \mid a$ y $p \nmid a$. Luego d es unidad y $(p, a) = (1)$: existen x, y con $px + ay = 1$. Multiplicando por b: $b = pbx + aby$, y p divide a ambos sumandos (al segundo porque $p \mid ab$), luego $p \mid b$.

Unicidad. Con "irreducible = primo" disponible, la unicidad es el argumento clásico: si $p_1 \cdots p_r = q_1 \cdots q_s$, entonces p_1 divide al producto de la derecha, luego divide a algún q_j, y siendo ambos irreducibles son asociados; se cancelan y se induce sobre r. □

Cuidado

[Ninguna de las dos implicaciones se invierte] El anillo $\mathbb{Z}[\frac{1+\sqrt{-19}}{2}]$ es un DIP que *no* es euclídeo para ninguna norma: es el contraejemplo estándar, y su demostración es delicada. Y $\mathbb{Z}[x]$ es un DFU que no es DIP: el ideal $(2, x)$ no es principal, porque un generador tendría que dividir a 2 y a x, luego ser ± 1, y $(2, x) \neq (1)$ ya que todos sus elementos tienen término independiente par. La jerarquía es, por tanto, estrictamente creciente, y cada escalón se paga con un teorema menos.

25.26. Cuerpos de descomposición, extensiones normales y módulos libres: los términos que el capítulo usa sin definir

Este capítulo deja tres deudas de vocabulario. Dos de ellas están en los ejercicios propuestos: el ejercicio que pregunta si $\mathbb{Q}(\sqrt[3]{2})/\mathbb{Q}$ es una extensión *normal* y el que pide el *cuerpo de descomposición* de x^4-2 usan palabras que solo se definen en el capítulo 26, es decir, *después*. La tercera está en la demostración del teorema de estructura sobre un dominio de ideales principales, donde se afirma sin más que un submódulo de un módulo *libre* es libre de *rango* menor o igual. Un lector que avance en orden no tiene dónde mirar. Aquí quedan las tres definiciones, con lo mínimo imprescindible para que los ejercicios se puedan hacer sin saltar de capítulo.

Descomponer un polinomio: el cuerpo mínimo donde se parte del todo

Definición 25.150 (Cuerpo de descomposición)**.** Sea K un cuerpo y $f \in K[x]$ no constante, de grado n. Una extensión E/K es un *cuerpo de descomposición* de f sobre K si se cumplen las dos condiciones a la vez:

1. f factoriza en $E[x]$ como producto de factores lineales, $f = c\,(x-\alpha_1)(x-\alpha_2)\cdots(x-\alpha_n)$ con $c \in K$ y $\alpha_i \in E$;
2. $E = K(\alpha_1, \ldots, \alpha_n)$, es decir, E está generado sobre K por esas raíces y por nada más.

La segunda condición es la que convierte a E en el cuerpo *mínimo* con la propiedad: ningún subcuerpo propio de E que contenga a K parte a f en lineales.

Teorema 25.151 (Existencia por la construcción de Kronecker, y unicidad)**.** *Todo $f \in K[x]$ no constante tiene un cuerpo de descomposición sobre K, y $[E : K] \leq n!$ con $n = \deg f$. Además dos cuerpos de descomposición cualesquiera de f sobre K son isomorfos mediante un isomorfismo que deja fijo a K punto a punto.*

Cómo se construye, en dos frases. El paso de Kronecker es el único ingrediente: si $p \in K[x]$ es irreducible, el cociente $K_1 = K[x]/(p)$ *es un cuerpo* —porque (p) es maximal en un dominio de ideales principales— y en él la clase $\bar{x}$ satisface $p(\bar{x}) = 0$. Es decir, se fabrica una raíz de la nada, declarando que existe. Ahora se itera: se toma un factor irreducible p_1 de f, se construye K_1, en $K_1[x]$ el polinomio f pierde al menos un factor lineal, y se repite con el cociente. Tras a lo sumo n pasos todos los factores son lineales, y quedarse con el subcuerpo generado por las raíces da E. La cota $n!$ sale de la torre: el primer paso multiplica el grado por a lo sumo n, el segundo por a lo sumo $n-1$, y así sucesivamente.

La unicidad se demuestra por inducción sobre $[E : K]$, extendiendo isomorfismos paso a paso: un isomorfismo $\sigma : K \to K'$ transforma f en σf y lleva factores irreducibles a factores irreducibles, de modo que $K(\alpha) \cong K'(\alpha')$ para cualquier raíz α' de σp; adjuntando raíces una a una se levanta σ hasta todo E. □

Observación 25.152 (Por qué la unicidad es la parte útil). La existencia es una construcción y podría parecer lo importante, pero en la práctica se usa mil veces más la unicidad. Es ella la

que permite decir *el* cuerpo de descomposición en vez de *un* cuerpo de descomposición, y con ello hablar de $\mathrm{Gal}(f)$ como de un objeto bien definido en vez de uno por cada construcción. Es también la que demuestra que hay un único cuerpo de p^n elementos salvo isomorfismo: $\mathbb{F}_{p^n}$ es el cuerpo de descomposición de $x^{p^n} - x$ sobre $\mathbb{F}_p$, y dos cuerpos de ese tamaño son ambos cuerpos de descomposición del mismo polinomio. Obsérvese que la unicidad es *salvo isomorfismo* y no canónica: el isomorfismo entre dos copias no es único, y contar de cuántas maneras se puede elegir es exactamente contar los elementos del grupo de Galois.

Ejemplo 25.153 (Tres descomposiciones calculadas)**. (a)** x^2+1 **sobre** $\mathbb{R}$**.** Ya es Kronecker puro: $\mathbb{C} = \mathbb{R}[x]/(x^2+1)$, de grado 2, y la factorización es $(x-\mathrm{i})(x+\mathrm{i})$. Adjuntar una raíz bastó, porque la otra es su opuesta y ya estaba.

(b) x^3-5 **sobre** $\mathbb{Q}$**.** Las raíces son $\sqrt[3]{5}$, $\omega\sqrt[3]{5}$ y $\omega^2\sqrt[3]{5}$ con $\omega = \mathrm{e}^{2\pi\mathrm{i}/3}$. Adjuntar solo $\sqrt[3]{5}$ no basta: $\mathbb{Q}(\sqrt[3]{5})$ está contenido en $\mathbb{R}$ y las otras dos raíces no son reales. Hace falta también ω, cuyo polinomio mínimo sobre $\mathbb{Q}(\sqrt[3]{5})$ sigue siendo x^2+x+1, así que

$$E = \mathbb{Q}(\sqrt[3]{5},\ \omega), \qquad [E:\mathbb{Q}] = 3\cdot 2 = 6.$$

La cota de Kronecker daba $3! = 6$: aquí se alcanza con igualdad.

(c) x^3+x+1 **sobre** $\mathbb{F}_2$**.** El polinomio no tiene raíces en $\mathbb{F}_2$ ($f(0) = f(1) = 1$), luego es irreducible, y $\mathbb{F}_8 = \mathbb{F}_2[x]/(x^3+x+1)$ contiene una raíz α. Aquí ocurre algo que sobre $\mathbb{Q}$ no pasa nunca: las otras dos raíces son α^2 y α^4, y ambas *ya viven* en $\mathbb{F}_8$, porque el Frobenius $a \mapsto a^2$ permuta las raíces. El cuerpo de descomposición es $\mathbb{F}_8$ mismo, de grado 3 y no 6. Adjuntar una raíz trajo las tres gratis; ese es exactamente el fenómeno que la siguiente definición bautiza.

Extensiones normales: o vienen todas las raíces, o no viene ninguna

Definición 25.154 (Extensión normal)**.** Una extensión algebraica L/K es *normal* si todo polinomio irreducible de $K[x]$ que tiene *alguna* raíz en L tiene *todas* sus raíces en L; equivalentemente, si el polinomio mínimo sobre K de cada elemento de L factoriza en lineales dentro de $L[x]$.

Teorema 25.155 (Caracterización en el caso finito)**.** *Para una extensión* finita L/K *son equivalentes:*

$$L/K \text{ es normal} \qquad \Longleftrightarrow \qquad L \text{ es el cuerpo de descomposición sobre } K \text{ de algún } f \in K[x].$$

Las dos implicaciones, en resumen. ($\Rightarrow$) Escríbase $L = K(\beta_1, \ldots, \beta_r)$, cosa posible por ser finita, y sea m_i el polinomio mínimo de β_i sobre K. Cada m_i es irreducible y tiene la raíz β_i en L, luego por normalidad tiene todas sus raíces en L. Entonces $f = m_1 m_2 \cdots m_r$ parte en lineales dentro de L, y el subcuerpo generado por sus raíces contiene a todos los β_i: es L entero.

($\Leftarrow$) Sea L el cuerpo de descomposición de f y sea $p \in K[x]$ irreducible con una raíz $\alpha \in L$; sea α' otra raíz de p en un cuerpo de descomposición mayor. El argumento de extensión de isomorfismos da un K-isomorfismo $K(\alpha) \to K(\alpha')$, y como L y $L(\alpha')$ son ambos cuerpos de descomposición de f sobre esos dos cuerpos isomorfos, la parte de unicidad del teorema 25.151

lo levanta a un isomorfismo $L \to L(\alpha')$ que fija K. Grados iguales fuerzan $L = L(\alpha')$, es decir $\alpha' \in L$. □

Ejemplo 25.156 (Normal, no normal, y la trampa de las torres)**. (a) Toda extensión de grado 2 es normal.** Si $[L : K] = 2$ y $\alpha \in L \setminus K$, su mínimo tiene grado 2 y una raíz en L; dividiendo, el otro factor es lineal con coeficientes en L, luego la segunda raíz también está en L. Sin excepciones y sin cuentas.

(b) $\mathbb{Q}(\sqrt[3]{2})/\mathbb{Q}$ no es normal. El irreducible $x^3 - 2$ tiene ahí la raíz $\sqrt[3]{2}$ y no tiene las otras dos, que son complejas mientras que $\mathbb{Q}(\sqrt[3]{2}) \subset \mathbb{R}$. Esta es la respuesta al ejercicio propuesto, y conviene ver que el motivo no es aritmético sino de tamaño: falta sitio para ω.

(c) La normalidad no es transitiva. Considérese la torre

$$\mathbb{Q} \;\subset\; \mathbb{Q}(\sqrt{2}) \;\subset\; \mathbb{Q}(\sqrt[4]{2}).$$

Cada peldaño tiene grado 2, luego cada peldaño es normal por el apartado (a). Y sin embargo la extensión total $\mathbb{Q}(\sqrt[4]{2})/\mathbb{Q}$ *no* es normal: contiene la raíz $\sqrt[4]{2}$ de $x^4 - 2$ pero no contiene $\mathrm{i}\sqrt[4]{2}$, que no es real. Este es el contraejemplo que hay que llevar en el bolsillo: ser normal es una propiedad de la extensión completa y no se hereda por composición de peldaños. Lo que sí es cierto es la mitad de arriba: si L/K es normal y $K \subseteq F \subseteq L$, entonces L/F sigue siendo normal.

Cuidado

[Normal no es lo mismo que de Galois] Ser normal no basta para que la correspondencia de Galois funcione: hace falta además que la extensión sea *separable*, cosa que en característica 0 y sobre cuerpos finitos es automática, pero no en general. Y cuidado con el falso amigo: el adjetivo "normal" aplicado a extensiones no tiene ninguna relación de definición con los subgrupos normales de la primera parte del capítulo. La relación existe, pero es un teorema —los subgrupos normales del grupo de Galois corresponden a los subcuerpos intermedios normales— y no un juego de palabras.

Módulos libres y rango: lo que el teorema de estructura da por sabido

Definición 25.157 (Módulo libre, base y rango)**.** Sea R un anillo conmutativo con 1 y M un R-módulo. Una familia $(e_i)_{i \in I}$ de elementos de M es una *base* si genera M y es *libre*, es decir, si toda relación $\sum_i r_i e_i = 0$ con casi todos los r_i nulos obliga a que *todos* los r_i sean 0. El módulo M es *libre* si posee alguna base; en tal caso todo elemento se escribe de forma única como $\sum_i r_i e_i$ y $M \cong R^{(I)}$. Cuando I es finito con $|I| = n$ se dice que M es libre de *rango* n y se escribe $\operatorname{rango} M = n$.

Ejemplo 25.158 (El ejemplo canónico y tres que no lo son)**. (a) R^n.** Las n-uplas con las operaciones coordenada a coordenada forman el módulo libre por antonomasia, con la base canónica $e_1, \ldots, e_n$. Es la definición de espacio vectorial con la palabra "cuerpo" rebajada a "anillo", y todo lo que se demuestra sobre R^n usando solo sumas y productos por escalares vale igual. Sobre un cuerpo, *todo* módulo es libre —eso es exactamente el teorema de existencia de bases— y por eso el álgebra lineal es fácil.

(b) $\mathbb{Z}/6$ como $\mathbb{Z}$-módulo no es libre. Ningún elemento m puede formar parte de una base, porque $6m = 0$ es una relación no trivial. En general, un módulo con torsión nunca es libre.

(c) $\mathbb{Q}$ como $\mathbb{Z}$-módulo no es libre, y no tiene torsión. Dos racionales cualesquiera a/b y c/d satisfacen la relación entera $(cb)\cdot\frac{a}{b} - (ad)\cdot\frac{c}{d} = 0$, luego una base tendría que tener un solo elemento; pero $\mathbb{Z}\cdot\frac{a}{b}$ no es todo $\mathbb{Q}$, porque no contiene $\frac{a}{2b}$. Sin torsión y sin ser libre: la implicación "libre $\Rightarrow$ sin torsión" no se puede invertir.

(d) Un ideal que no es libre. En $R = \mathbb{Z}[x]$, que es un dominio, el ideal $I = (2, x)$ cumple $x\cdot 2 - 2\cdot x = 0$, luego $\{2, x\}$ no es base; y ningún elemento solo genera I, porque I no es principal. Así que I no es libre de ningún rango. Aquí está la razón profunda de que el teorema de estructura exija un *dominio de ideales principales* y no un dominio cualquiera: sobre un DIP todo ideal es (d), que sí es libre de rango 1.

Cuidado

[El rango puede no estar bien definido, pero no sobre un dominio] Que $R^m \cong R^n$ implique $m = n$ es un enunciado que hay que demostrar, y sobre un anillo arbitrario es *falso*: existen anillos no conmutativos —el ejemplo clásico es el de los endomorfismos de un espacio vectorial de dimensión infinita, donde $V \cong V \oplus V$ obliga a $R \cong R^2$ como R-módulos— en los que un mismo módulo libre tiene bases de dos, tres o cualquier número de elementos. Un anillo en el que eso no ocurre se dice que tiene *número de base invariante*.
La buena noticia es que todo anillo conmutativo no nulo lo tiene, y sobre un *dominio* D el argumento es de una línea: sea F su cuerpo de fracciones; un isomorfismo $D^m \cong D^n$ induce por paso a fracciones un isomorfismo de F-espacios vectoriales $F^m \cong F^n$, y entonces $m = n$ por la invariancia de la dimensión, que sí está demostrada en el capítulo 7.
En otras palabras: el rango de un módulo libre sobre un dominio es la dimensión del espacio vectorial que se obtiene al permitir dividir, y por eso hereda toda la buena conducta del álgebra lineal.

Observación 25.159 (Por qué el teorema de estructura necesita esta palabra). La demostración del teorema de estructura sobre un DIP tiene dos usos de la libertad, y ninguno es prescindible. El primero es de arranque: un módulo generado por n elementos es un cociente R^n/N, y para poder escribir una *matriz* de relaciones y llevarla a forma normal de Smith hace falta que N sea también libre, con una base que sirva de columnas. Ese es el lema —todo submódulo de un módulo libre de rango finito sobre un DIP es libre de rango menor o igual— que el capítulo usaba sin nombre. Sobre $\mathbb{Z}[x]$ el lema es falso, como muestra el ideal $(2, x)$ del ejemplo anterior, que es un submódulo de $\mathbb{Z}[x]$ y no es libre; de ahí que el teorema de estructura no valga sobre dominios cualesquiera.

El segundo uso es de enunciado: en la descomposición $M \cong R^s \oplus R/(d_1) \oplus \cdots \oplus R/(d_r)$, el número s *es* el rango de la parte libre, y afirmar que la descomposición es única incluye afirmar que s está determinado por M. Sin la advertencia anterior esa unicidad ni siquiera tendría sentido. Con ella, s admite una descripción intrínseca y cómoda de calcular: $s = \dim_F(M \otimes_R F)$ con F el cuerpo de fracciones, o dicho sin producto tensorial, s es el número máximo de elementos de M linealmente independientes sobre R. Para grupos abelianos finitamente

generados, s es el número de copias de $\mathbb{Z}$ y r mide la parte finita: $\mathbb{Z}^s \oplus$ (torsión), con $s = 0$ exactamente cuando el grupo es finito.

25.27. Ejercicios resueltos

Ejercicio 25.1. Escribe la tabla del grupo $D_3 \cong S_3$ con generadores r (rotación de 120°) y s (reflexión), y encuentra todos sus subgrupos, señalando cuáles son normales.

Solución. $S_3 = \{e, r, r^2, s, rs, r^2s\}$ con $r^3 = s^2 = e$, $sr = r^2s$. De la última relación, $rs \neq sr$: no abeliano (el más pequeño posible). Subgrupos: triviales; $\langle r\rangle = \{e, r, r^2\} \cong \mathbb{Z}_3$; y $\langle s\rangle, \langle rs\rangle, \langle r^2s\rangle \cong \mathbb{Z}_2$. Lagrange prohíbe órdenes 4 y 5. Normalidad: $\langle r\rangle$ tiene índice 2, normal; los tres de orden 2 no lo son (conjugar una reflexión da otra reflexión: $rsr^{-1} = r^2s \notin \langle s\rangle$). Cadena soluble: $\{e\} \trianglelefteq \langle r\rangle \trianglelefteq S_3$ con cocientes $\mathbb{Z}_3$ y $\mathbb{Z}_2$. ■

Ejercicio 25.2. Sea $\sigma = (1\,3\,4)(2\,5) \in S_5$ y $\tau = (1\,2)$. Calcula $\sigma\tau$, $\tau\sigma$, σ^{-1}, el orden y el signo de σ, y su clase de conjugación.

Solución. Componiendo de derecha a izquierda. $\sigma\tau$: $1 \mapsto 2 \mapsto 5$, $2 \mapsto 1 \mapsto 3$, $3 \mapsto 4$, $4 \mapsto 1$, $5 \mapsto 2$: $\sigma\tau = (1\,5\,2\,3\,4)$. $\tau\sigma$: $1 \mapsto 3$, $3 \mapsto 4$, $4 \mapsto 1 \mapsto 2$, $2 \mapsto 5$, $5 \mapsto 2 \mapsto 1$: $\tau\sigma = (1\,3\,4\,2\,5)$. Distintos: no conmutan. $\sigma^{-1} = (1\,4\,3)(2\,5)$ (invertir cada ciclo). Orden: $\operatorname{mcm}(3, 2) = 6$. Signo: 3-ciclo par, trasposición impar: $\operatorname{sgn}\sigma = -1$, impar. Su clase de conjugación: todas las permutaciones de tipo $3 + 2$, que son $\frac{5!}{3\cdot 2} = 20$. ■

Ejercicio 25.3. Demuestra el primer teorema de isomorfía y aplícalo para identificar $\mathbb{Z}_{12}/\langle 4\rangle$.

Solución. Sea $\Phi : G/\ker\varphi \to \operatorname{im}\varphi$, $\Phi(gK) = \varphi(g)$ (con $K = \ker\varphi$). *Bien definida e inyectiva*: $gK = hK \iff h^{-1}g \in K \iff \varphi(h^{-1}g) = e' \iff \varphi(g) = \varphi(h)$. *Homomorfismo*: $\Phi(gK \cdot hK) = \varphi(gh) = \varphi(g)\varphi(h)$. *Sobre*: evidente sobre la imagen. Aplicación: $\varphi : \mathbb{Z}_{12} \to \mathbb{Z}_4$, $x \mapsto x \bmod 4$, es sobreyectivo con núcleo $\{0, 4, 8\} = \langle 4\rangle$; por tanto $\mathbb{Z}_{12}/\langle 4\rangle \cong \mathbb{Z}_4$. ■

Ejercicio 25.4. ¿Cuántas rotaciones tiene el grupo del icosaedro (20 caras triangulares)? Úsalo para explicar la afirmación "el grupo de rotaciones del icosaedro es A_5".

Solución. Órbita–estabilizador sobre las 20 caras: la acción es transitiva y el estabilizador de una cara son sus 3 rotaciones: $|G| = 20 \times 3 = 60$. Se demuestra que G actúa fielmente sobre 5 objetos geométricos (los 5 cubos inscritos), dando una inyección $G \hookrightarrow S_5$; la imagen, de orden 60 y formada por rotaciones (orientación positiva: permutaciones pares), es A_5. Así, el grupo simple A_5 está tallado en piedra desde la antigüedad: es la simetría del icosaedro. (Y por eso la insolubilidad de la quíntica está relacionada con los sólidos platónicos, como explotó Klein.) ■

Ejercicio 25.5. Encuentra las unidades y los divisores de cero de $\mathbb{Z}_{12}$, y todos sus ideales.

Solución. Unidades: clases coprimas con 12: $\{1, 5, 7, 11\}$ (cada una es su propio inverso: $5^2 = 25 \equiv 1$, etc.: $\mathbb{Z}_{12}^* \cong \mathbb{Z}_2 \times \mathbb{Z}_2$). Divisores de cero: el resto de los no nulos: $\{2, 3, 4, 6, 8, 9, 10\}$ (p.ej. $3 \cdot 4 = 0$). Ideales: los subgrupos aditivos (d) con $d \mid 12$: $(1) = \mathbb{Z}_{12}$, $(2), (3), (4), (6), (0)$. Maximales: (2) y (3), con cocientes los cuerpos $\mathbb{Z}_2$ y $\mathbb{Z}_3$; en cambio $\mathbb{Z}_{12}/(4) \cong \mathbb{Z}_4$ no es cuerpo: (4) no es maximal (cabe en (2)).■

Ejercicio 25.6. Estudia la irreducibilidad sobre $\mathbb{Q}$ de: (a) $x^3 + 3x + 3$; (b) $x^4 + x^2 + 1$; (c) $x^4 + 1$; (d) $x^5 + x + 1$.

Solución. (a) Eisenstein con $p = 3$: irreducible. *(b)* Sin raíces racionales, pero ¡cuidado!: puede factorizar en cuadráticos. En efecto: $x^4 + x^2 + 1 = (x^2 + x + 1)(x^2 - x + 1)$. **Reducible**. (Moral: para grado ≥ 4,

no tener raíces no basta.) *(c)* Sin raíces reales; probamos cuadráticos: $x^4+1=(x^2+ax+b)(x^2-ax+c)$ fuerza $b+c=a^2$, $a(c-b)=0$, $bc=1$; con $a\neq 0$: $b=c=\pm 1$ y $a^2=\pm 2$: sin solución racional. **Irreducible sobre** $\mathbb{Q}$ (aunque, curiosamente, es reducible módulo *todo* primo: la reducción mod p solo funciona en una dirección). *(d)* $x^5+x+1=(x^2+x+1)(x^3-x^2+1)$ (se descubre notando que las raíces cúbicas primitivas de la unidad lo anulan: si $\omega^3=1, \omega\neq 1$: $\omega^5+\omega+1=\omega^2+\omega+1=0$). **Reducible**. ■

Ejercicio 25.7. En $\mathbb{F}_2[x]/(x^3+x+1)$: comprueba que es un cuerpo de 8 elementos, y que la clase α de x genera el grupo multiplicativo.

Solución. x^3+x+1 no tiene raíces en $\mathbb{F}_2$ ($f(0)=f(1)=1$) y es cúbico: irreducible. El cociente es un cuerpo con los 8 polinomios de grado ≤ 2: $\mathbb{F}_8$. Con $\alpha^3=\alpha+1$ (característica 2):

$$\alpha^1=\alpha,\ \alpha^2,\ \alpha^3=\alpha+1,\ \alpha^4=\alpha^2+\alpha,\ \alpha^5=\alpha^2+\alpha+1,\ \alpha^6=\alpha^2+1,\ \alpha^7=\alpha^3+\alpha=1.$$

Las siete potencias recorren todos los no nulos: $\mathbb{F}_8^*=\langle\alpha\rangle\cong\mathbb{Z}_7$, cíclico como promete la teoría. (Esta estructura de "logaritmo discreto" es la base de la criptografía en cuerpos finitos.) ■

Ejercicio 25.8. Demuestra con Bézout que $\overline{x}+1$ es invertible en $\mathbb{Q}[x]/(x^3-2)$ y calcula su inverso. Interpreta el resultado en $\mathbb{Q}(\sqrt[3]{2})$.

Solución. Euclides con $f=x+1$ y $p=x^3-2$: $x^3-2=(x+1)(x^2-x+1)-3$, así que $3=(x+1)(x^2-x+1)-(x^3-2)$ y

$$(x+1)^{-1}\equiv\frac{x^2-x+1}{3}\pmod{x^3-2}.$$

En $\mathbb{Q}(\sqrt[3]{2})$ con $\alpha=\sqrt[3]{2}$: $\frac{1}{1+\alpha}=\frac{\alpha^2-\alpha+1}{3}$. Verificación: $(1+\alpha)(\alpha^2-\alpha+1)=\alpha^3+1=3$ ✓. Racionalizar denominadores cúbicos —que en secundaria parece un truco de magia— es simplemente el algoritmo de Euclides en $\mathbb{Q}[x]$. ■

25.28. Ejercicios adicionales resueltos

Ejercicio 25.9. Todo grupo de orden 6 es $\mathbb{Z}_6$ o S_3.

Solución. Cauchy da a (orden 3), b (orden 2); $\langle a\rangle$ normal (índice 2); $bab^{-1}=a$ da $\mathbb{Z}_6$, $=a^{-1}$ da S_3. ■

Ejercicio 25.10. $\mathrm{Aut}(\mathbb{Z}_8)$.

Solución. $(\mathbb{Z}_8)^\times=\{1,3,5,7\}$ con todos los cuadrados $\equiv 1$: Klein, no cíclico —no hay raíces primitivas módulo 2^k, $k\geq 3$—. ■

Ejercicio 25.11. Expresa $x^2y+xy^2+y^2z+yz^2+z^2x+zx^2$ en elementales.

Solución. $e_1e_2-3e_3$; verificación en $(1,1,1)$: $6=9-3$ ✓. ■

Ejercicio 25.12. 5^{-1} en $\mathbb{Z}_{37}$ y el sistema $x\equiv 2\ (5)$, $x\equiv 3\ (7)$.

Solución. Euclides: $1=15\cdot 5-2\cdot 37$: $5^{-1}=15$. CRT: $x\equiv 17\pmod{35}$. (RSA con 300 cifras: lo mismo.) ■

Ejercicio 25.13. Construye $\mathbb{F}_9$ y un generador de $\mathbb{F}_9^*$.

Solución. $\mathbb{F}_3[i]$ con $i^2 = -1$ (x^2+1 irreducible). $1+i$: $(1+i)^2 = 2i$, $(1+i)^4 = 2$, orden 8: generador ✓. ■

Ejercicio 25.14. $\mathbb{Z}[x]$ no es DIP.

Solución. $(2, x)$: un generador dividiría a 2 y a x: solo ± 1, pero $1 \notin (2, x)$ (términos independientes pares). La cadena euclídeo $\subset$ DIP $\subset$ DFU es estricta. ■

Ejercicio 25.15. Orden $p^2 \Rightarrow$ abeliano; clasifícalos.

Solución. $Z \neq \{e\}$ (ecuación de clases); $|Z| = p$ daría G/Z cíclico y G abeliano (contradicción): $Z = G$. Estructura: $\mathbb{Z}_{p^2}$ o $\mathbb{Z}_p^2$. ■

25.29. Taller de ejemplos explicados: grupos con números concretos

Idea clave

Intuición primero. Un grupo suena abstracto, pero es solo un conjunto donde puedes *combinar* elementos, *deshacer* (inversos) y hay un *neutro*. El reloj de 6 horas ($\mathbb{Z}_6$) es un grupo; las simetrías de un cuadrado son un grupo. Al trabajar con $\mathbb{Z}_6$ concreto vas a *ver* el teorema de Lagrange como un simple embaldosado, y el "orden" de un elemento como cuántos pasos das antes de volver al inicio.

Un *grupo* es un conjunto con una operación que se puede combinar, deshacer (inversos) y tiene un elemento neutro. Los ejemplos concretos hacen la idea tangible.

25.29.1. El grupo $\mathbb{Z}_6$ y el orden de un elemento

$\mathbb{Z}_6 = \{0, 1, 2, 3, 4, 5\}$ con la suma módulo 6 (el reloj de 6 horas). El *orden* de un elemento es cuántas veces hay que sumarlo consigo mismo para volver al 0.

Ejemplo 25.160. Hallar el orden de 2 en $\mathbb{Z}_6$.

Vamos sumando 2 y reduciendo módulo 6:

$$2, \quad 2+2=4, \quad 4+2=6\equiv 0.$$

Llegamos a 0 tras *tres* pasos, así que el orden de 2 es 3. Fíjate que 3 divide a 6: eso no es casualidad, es el teorema de Lagrange en acción.

25.29.2. Clases laterales y el teorema de Lagrange

Ejemplo 25.161. Considera el subgrupo $H = \{0, 3\}$ dentro de $\mathbb{Z}_6$. Vamos a "embaldosar" $\mathbb{Z}_6$ con copias trasladadas de H (las clases laterales).

Paso 1. La baldosa base es $H = \{0, 3\}$.

Paso 2. Le sumamos 1 a cada elemento: $1 + H = \{1, 4\}$.

Paso 3. Le sumamos 2: $2 + H = \{2, 5\}$.

Ya cubrimos todo $\mathbb{Z}_6 = \{0,3\} \cup \{1,4\} \cup \{2,5\}$ con *tres* baldosas de *dos* elementos cada una. Y en efecto $3 \times 2 = 6$: el número de baldosas por el tamaño de cada una da el total. Eso es **Lagrange**: el tamaño de un subgrupo siempre divide al del grupo.

25.29.3. Comprobar que algo es un grupo

Ejemplo 25.162. ¿Es $G = \{1, -1, i, -i\}$ un grupo con la multiplicación (donde $i^2 = -1$)?

Comprobamos las propiedades: el neutro es 1 (multiplicar por 1 no cambia nada). Cada elemento tiene inverso: el de i es $-i$ (porque $i \cdot (-i) = -i^2 = 1$), el de -1 es -1. Y el producto de dos cualesquiera vuelve a estar en el conjunto (por ejemplo $i \cdot i = -1$, que está). Sí es un grupo, y de hecho es *cíclico*: todo sale de multiplicar i por sí mismo: $i, i^2 = -1, i^3 = -i, i^4 = 1$.

25.30. Ejercicios propuestos

Sin solución desarrollada: solo enunciado y, al final, la respuesta. Aquí muchas respuestas son un nombre ("el grupo de Klein", "sí, por Eisenstein") más que un número: comprueba que coincida el fondo del argumento, no la redacción.

25.1. Calcula $[\mathbb{Q}(\sqrt{2}) : \mathbb{Q}]$ y da una base como espacio vectorial sobre $\mathbb{Q}$.

25.2. Calcula $[\mathbb{Q}(\sqrt{2}, \sqrt{3}) : \mathbb{Q}]$ y da una base.

25.3. Halla el polinomio mínimo de $\sqrt{2} + \sqrt{3}$ sobre $\mathbb{Q}$.

25.4. Halla el polinomio mínimo de $\sqrt[3]{2}$ sobre $\mathbb{Q}$ y sobre $\mathbb{Q}(\sqrt[3]{2})$.

25.5. ¿Es $\mathbb{Q}(\sqrt[3]{2})/\mathbb{Q}$ una extensión normal? Justifica.

25.6. ¿Cuántos elementos tiene $\mathbb{F}_3[x]/(x^2+1)$? ¿Es cuerpo?

25.7. ¿Existe un cuerpo finito de orden 6? ¿Y de orden 27?

25.8. Da el grado de $\mathbb{Q}(\zeta_5)/\mathbb{Q}$, donde $\zeta_5 = e^{2\pi i/5}$.

25.9. ¿Es π algebraico sobre $\mathbb{Q}$? ¿Y sobre $\mathbb{Q}(\pi^2)$?

25.10. Si $[K : \mathbb{Q}] = 6$ y $\mathbb{Q} \subset F \subset K$, ¿qué grados puede tener F sobre $\mathbb{Q}$?

25.11. Da el cuerpo de descomposición de $x^4 - 2$ sobre $\mathbb{Q}$ y su grado.

25.12. ¿Es toda extensión finita algebraica? ¿Y al revés?

Respuestas

1. Grado 2; base $\{1, \sqrt{2}\}$. **2.** Grado 4; base $\{1, \sqrt{2}, \sqrt{3}, \sqrt{6}\}$. **3.** $x^4 - 10x^2 + 1$. **4.** Sobre $\mathbb{Q}$: $x^3 - 2$. Sobre $\mathbb{Q}(\sqrt[3]{2})$: $x - \sqrt[3]{2}$, de grado 1. **5.** No: contiene una raíz de $x^3 - 2$ pero no las otras dos, que son complejas. **6.** Nueve elementos, y sí es cuerpo ($x^2 + 1$ es irreducible sobre $\mathbb{F}_3$): es $\mathbb{F}_9$. **7.** De orden 6 no (los órdenes son potencias de primos); de orden $27 = 3^3$ sí, y es único salvo isomorfismo. **8.** $[\mathbb{Q}(\zeta_5) : \mathbb{Q}] = \varphi(5) = 4$. **9.** Sobre $\mathbb{Q}$ no: es trascendente. Sobre $\mathbb{Q}(\pi^2)$ sí, con polinomio mínimo $x^2 - \pi^2$. **10.** 1, 2, 3 o 6: los divisores de 6, por la fórmula de la torre. **11.** $\mathbb{Q}(\sqrt[4]{2}, i)$, de grado 8 sobre $\mathbb{Q}$. **12.** Finita $\Rightarrow$ algebraica sí; el recíproco no: la clausura algebraica $\overline{\mathbb{Q}}$ sobre $\mathbb{Q}$ es algebraica de grado infinito.

Capítulo 26

Teoría de Galois

Lo que necesitas traer

Del capítulo 25: grupos solubles, extensiones de cuerpos, polinomios irreducibles y grado de una extensión. Del capítulo 7: dimensión de un espacio vectorial, porque el grado de una extensión es una dimensión. Del capítulo 4: raíces de la unidad.

Llegamos a la cima. La teoría de Galois conecta dos mundos: las *extensiones de cuerpos* (dónde viven las raíces de un polinomio) y los *grupos* (las simetrías de esas raíces). El diccionario entre ambos —la correspondencia de Galois— responde preguntas que resistieron siglos: por qué la cuadrática, la cúbica y la cuártica tienen fórmula y la quíntica no, y por qué los tres problemas griegos de la regla y el compás son imposibles.

26.1. Extensiones de cuerpos

Idea clave

Intuición primero. Una extensión de cuerpos es lo que ocurre cuando a un cuerpo le falta una raíz y decides *agregarla*. A $\mathbb{Q}$ le falta $\sqrt{2}$; se la añades y aparece $\mathbb{Q}(\sqrt{2}) = \{a + b\sqrt{2}\}$, un mundo nuevo que contiene al viejo. A $\mathbb{R}$ le falta $\sqrt{-1}$; se la añades y sale $\mathbb{C}$. Siempre es el mismo gesto.

La idea decisiva —y es de Dedekind, no de Galois— es que el cuerpo grande L es un **espacio vectorial** sobre el pequeño K, así que tiene dimensión. Esa dimensión es el *grado* $[L : K]$, y de golpe todo el capítulo 7 queda disponible aquí: $\mathbb{C}$ sobre $\mathbb{R}$ tiene base $\{1, i\}$ y grado 2, y $\mathbb{Q}(\sqrt{2})$ sobre $\mathbb{Q}$ tiene base $\{1, \sqrt{2}\}$ y grado 2. El grado mide *cuánto* agrandaste, y coincide con el grado del polinomio mínimo del elemento añadido.

La multiplicatividad del grado en una torre, $[M : K] = [M : L][L : K]$, es entonces la versión de cuerpos del teorema de Lagrange: los grados se multiplican como se multiplican los índices de subgrupos. Es la primera señal de que hay una correspondencia escondida entre torres de cuerpos y cadenas de subgrupos, y toda esta teoría consiste en hacerla explícita.

Definición 26.1. Si $K \subseteq L$ son cuerpos, L/K es una *extensión*, y L es un espacio vectorial sobre K; su dimensión es el *grado* $[L : K]$. Extensión *finita* = grado finito.

Ejemplo 26.2. $[\mathbb{C} : \mathbb{R}] = 2$ (base $\{1, i\}$); $[\mathbb{Q}(\sqrt{2}) : \mathbb{Q}] = 2$; $[\mathbb{R} : \mathbb{Q}] = \infty$ (incluso no numerable).

Teorema 26.3 (Multiplicatividad del grado)**.** *Si $K \subseteq L \subseteq M$: $[M : K] = [M : L]\,[L : K]$.*

Demostración. Si $\{u_i\}$ es base de L/K y $\{v_j\}$ de M/L, entonces $\{u_i v_j\}$ es base de M/K: todo $m \in M$ es $m = \sum_j \ell_j v_j$ con $\ell_j \in L$, y cada $\ell_j = \sum_i c_{ij} u_i$ con $c_{ij} \in K$, luego generan; la independencia se comprueba invirtiendo los dos pasos (una combinación nula agrupa en coeficientes de los v_j que deben anularse en L, y luego cada uno se anula en K). □

Este teorema —una pura afirmación de álgebra lineal— es el motor de incontables demostraciones de imposibilidad: si el grado de lo que quieres construir no divide al grado de donde vives, no se puede.

26.1.1. Elementos algebraicos y polinomio mínimo

Definición 26.4. $\alpha \in L$ es *algebraico* sobre K si anula algún polinomio no nulo de $K[x]$; si no, *trascendente*. Para α algebraico, el *polinomio mínimo* m_α es el mónico irreducible de menor grado con $m_\alpha(\alpha) = 0$; divide a todo polinomio que anule a α. Además

$$K(\alpha) \cong K[x]/(m_\alpha), \qquad [K(\alpha) : K] = \deg m_\alpha,$$

con base $\{1, \alpha, \ldots, \alpha^{d-1}\}$: la máquina de cuerpos del capítulo 25 en acción.

Ejemplo 26.5. $m_{\sqrt{2}} = x^2 - 2$; $m_{\sqrt[3]{2}} = x^3 - 2$ (Eisenstein, $p = 2$), así que $[\mathbb{Q}(\sqrt[3]{2}) : \mathbb{Q}] = 3$; $m_i = x^2 + 1$. Los números e y π son trascendentes (Hermite 1873, Lindemann 1882): no viven en *ninguna* extensión finita de $\mathbb{Q}$.

Proposición 26.6. *Toda extensión finita es algebraica (cada α genera una subextensión de grado finito, y las potencias $1, \alpha, \alpha^2, \ldots$ deben ser linealmente dependientes). El conjunto de los elementos de L algebraicos sobre K es un subcuerpo: si α, β son algebraicos, $\alpha \pm \beta$, $\alpha\beta$, α/β viven en $K(\alpha, \beta)$, de grado finito $\leq [K(\alpha) : K] \cdot [K(\beta) : K]$.*

Regla: hallar el polinomio mínimo y el grado de una extensión

Todo el capítulo descansa en saber calcular $[K(\alpha) : K]$. El procedimiento es siempre el mismo.

1. **Fabrica un polinomio que anule a α.** Aísla los radicales y eleva a potencias hasta que desaparezcan todos. Cada elevación al cuadrado suele duplicar el grado, así que conviene aislar primero el radical más incómodo.
2. **Comprueba que es irreducible sobre K.** Aquí están los criterios, en orden de comodidad:

- *Grados 2 y 3:* un polinomio de grado 2 o 3 es reducible si y solo si tiene una raíz en K. Basta buscar raíces.
- *Raíces racionales:* para coeficientes enteros y polinomio mónico, las únicas raíces racionales posibles son divisores enteros del término independiente.
- *Eisenstein:* si existe un primo p que divide a todos los coeficientes salvo al principal, y p^2 no divide al término independiente, el polinomio es irreducible. A veces hay que sustituir $x \mapsto x+1$ para que Eisenstein se aplique.
- *Reducción módulo p:* si al reducir los coeficientes módulo un primo p que no divide al coeficiente principal el polinomio resulta irreducible en $\mathbb{F}_p[x]$, entonces lo era en $\mathbb{Q}[x]$. El recíproco es falso.

3. **Hazlo mónico.** El polinomio mínimo es, por convenio, mónico: divide por el coeficiente principal.
4. **El grado del polinomio mínimo *es* el grado de la extensión,** y $\{1, \alpha, \ldots, \alpha^{d-1}\}$ es una base de $K(\alpha)$ sobre K.
5. **Para extensiones con varios generadores, usa la torre:** $[M : K] = [M : L]\,[L : K]$. Calcula un piso, luego el siguiente *sobre el anterior* (no sobre K), y multiplica. Que $\sqrt{3} \notin \mathbb{Q}(\sqrt{2})$ hay que justificarlo, no darlo por hecho.
6. **Control por divisibilidad.** El grado de cualquier subextensión divide al grado total (es Lagrange disfrazado). Si te sale un grado 5 dentro de una extensión de grado 6, hay un error.

Ejemplo 26.7 (El polinomio mínimo de $\sqrt{2}+\sqrt{3}$)**.** Sea $\alpha = \sqrt{2}+\sqrt{3}$.

Paso 1. Elevamos al cuadrado: $\alpha^2 = 2 + 2\sqrt{6} + 3 = 5 + 2\sqrt{6}$. Aislamos y volvemos a elevar:

$$\alpha^2 - 5 = 2\sqrt{6} \implies (\alpha^2-5)^2 = 24 \implies \alpha^4 - 10\alpha^2 + 25 = 24,$$

es decir $\alpha^4 - 10\alpha^2 + 1 = 0$.

Paso 2. ¿Es $x^4 - 10x^2 + 1$ irreducible sobre $\mathbb{Q}$? Raíces racionales: los únicos candidatos son ± 1, y $1 - 10 + 1 = -8 \neq 0$. No hay factores lineales. ¿Y un producto de dos cuadráticos? Si fuera $(x^2 + ax + b)(x^2 - ax + c)$ con a, b, c racionales (el coeficiente de x^3 es cero, luego los términos en x son opuestos), desarrollando se obtiene $b + c - a^2 = -10$, $a(c-b) = 0$ y $bc = 1$. Si $a = 0$, entonces $b + c = -10$ y $bc = 1$, cuyas soluciones son $-5 \pm 2\sqrt{6}$: irracionales. Si $c = b$, entonces $b^2 = 1$, luego $b = \pm 1$ y $2b - a^2 = -10$, es decir $a^2 = 12$ o $a^2 = 8$: tampoco racionales. No hay factorización, así que el polinomio es irreducible.

Pasos 3 y 4. $m_\alpha(x) = x^4 - 10x^2 + 1$ y $[\mathbb{Q}(\alpha) : \mathbb{Q}] = 4$.

Paso 5, y la sorpresa. Por otro lado, $[\mathbb{Q}(\sqrt{2}, \sqrt{3}) : \mathbb{Q}] = 2 \cdot 2 = 4$: el primer piso es $\mathbb{Q}(\sqrt{2})$, de grado 2, y $\sqrt{3} \notin \mathbb{Q}(\sqrt{2})$ porque de $\sqrt{3} = a + b\sqrt{2}$ se seguiría, elevando al cuadrado, $3 = a^2 + 2b^2 + 2ab\sqrt{2}$, lo que obliga a $ab = 0$ y lleva a que $\sqrt{3}$ o $\sqrt{3/2}$ sean racionales; ninguno lo es. Ahora bien, $\mathbb{Q}(\alpha) \subseteq \mathbb{Q}(\sqrt{2}, \sqrt{3})$ y ambos tienen grado 4, luego *son el mismo cuerpo*:

$$\mathbb{Q}(\sqrt{2}, \sqrt{3}) = \mathbb{Q}(\sqrt{2} + \sqrt{3}).$$

Este es el teorema del elemento primitivo en un caso concreto: dos generadores se han fundido en uno solo. Para verlo a mano, obsérvese que $\alpha^3 = 11\sqrt{2}+9\sqrt{3}$, de modo que $\frac{1}{2}(\alpha^3-9\alpha) = \sqrt{2}$ y por tanto $\sqrt{2}$, y con él $\sqrt{3} = \alpha - \sqrt{2}$, están en $\mathbb{Q}(\alpha)$.

Ejemplo 26.8 (Dos grados calculados con Eisenstein y con la torre)**.** *(a)* $\alpha = \sqrt[4]{2}$. Anula a $x^4 - 2$, irreducible por Eisenstein con $p = 2$ (2 divide a $0, 0, 0, -2$ y 4 no divide a -2). Luego $[\mathbb{Q}(\sqrt[4]{2}) : \mathbb{Q}] = 4$. Nótese la torre interna: $\mathbb{Q} \subset \mathbb{Q}(\sqrt{2}) \subset \mathbb{Q}(\sqrt[4]{2})$, con grados 2 y 2, que multiplican a 4. ✓

(b) $\zeta_7 = \mathrm{e}^{2\pi i/7}$. Anula a $x^7 - 1$, que no es irreducible ($x - 1$ lo divide). El cociente es

$$\Phi_7(x) = x^6 + x^5 + x^4 + x^3 + x^2 + x + 1,$$

y sustituyendo $x \mapsto x + 1$ se obtiene un polinomio cuyos coeficientes son los números combinatorios $\binom{7}{k}$ con $k \geq 1$, todos divisibles por 7, y cuyo término independiente es exactamente 7, no divisible por 49: Eisenstein con $p = 7$ se aplica. Por tanto Φ_7 es irreducible y $[\mathbb{Q}(\zeta_7) : \mathbb{Q}] = 6 = \varphi(7)$.

(c) Un control que falla. Supongamos que alguien afirma que $\sqrt[3]{2} \in \mathbb{Q}(\sqrt{2}, \sqrt{3})$. Ese cuerpo tiene grado 4 sobre $\mathbb{Q}$, y $\mathbb{Q}(\sqrt[3]{2})$ tiene grado 3; pero 3 no divide a 4. Imposible, sin necesidad de ningún cálculo. El criterio de divisibilidad del paso 6 resuelve en un segundo preguntas que a mano serían laboriosas.

Regla: calcular el polinomio mínimo de un elemento algebraico

El polinomio mínimo es el "carné de identidad" de un número algebraico: el mónico de menor grado que lo anula. Todo lo demás en teoría de Galois se apoya en saber calcularlo.

1. **Fija el cuerpo base.** "El polinomio mínimo de $\sqrt{2}$" no significa nada por sí solo: sobre $\mathbb{Q}$ es $x^2 - 2$, pero sobre $\mathbb{R}$ es $x - \sqrt{2}$. Escribe siempre $m_{\alpha,K}$.

2. **Fabrica una relación algebraica.** Llama α al número, aísla el radical más incómodo y eleva a la potencia adecuada. Repite hasta que desaparezcan todos los radicales.

3. **Comprueba que el candidato es irreducible** sobre K. Con Eisenstein (posiblemente tras trasladar $x \mapsto x + 1$), con reducción módulo un primo, o buscando raíces racionales si el grado es 2 o 3.

4. **Si no logras probar irreducibilidad, acota el grado.** Se tiene $[K(\alpha) : K] = \deg m_{\alpha,K}$, y la torre $[K(\alpha, \beta) : K] = [K(\alpha, \beta) : K(\alpha)] \cdot [K(\alpha) : K]$ suele fijar el grado exacto por multiplicatividad.

5. **Cuenta los conjugados.** Si sabes que α tiene exactamente n conjugados distintos $\alpha_1, \ldots, \alpha_n$ (las imágenes de α por los encajes en una clausura), entonces $m_\alpha = \prod(x - \alpha_i)$ y su grado es n. Es la vía más rápida cuando los signos de los radicales son evidentes.

6. **Verifica numéricamente.** Sustituye el valor decimal de α en el candidato: debe salir cero hasta la precisión con que trabajes. Es un control barato que detecta casi todos los errores de álgebra.

7. **Control estructural:** m_α divide a *cualquier* polinomio de $K[x]$ que anule a α. Si tu candidato tiene grado mayor que otro polinomio anulador, te has equivocado.

Ejemplo 26.9 (Cuatro polinomios mínimos sobre $\mathbb{Q}$)**. (a)** $\alpha = \sqrt{2}+\sqrt{3}$**.** Elevamos: $\alpha^2 = 5+2\sqrt{6}$, luego $\alpha^2-5 = 2\sqrt{6}$ y elevando otra vez $\alpha^4 - 10\alpha^2 + 25 = 24$. Por tanto

$$m_\alpha = x^4 - 10x^2 + 1.$$

¿Es irreducible? El grado 4 es correcto por la vía del paso 5: los conjugados son $\pm\sqrt{2}\pm\sqrt{3}$, cuatro números distintos, y su producto $\prod(x \mp \sqrt{2} \mp \sqrt{3})$ da exactamente $x^4 - 10x^2 + 1$. Control numérico: $\alpha \approx 3{,}1463$, $\alpha^2 \approx 9{,}8990$, y $9{,}8990^2 - 10\cdot 9{,}8990 + 1 \approx 98{,}0 - 99{,}0 + 1 \approx 0$.

De paso queda demostrado que $\mathbb{Q}(\sqrt{2},\sqrt{3}) = \mathbb{Q}(\sqrt{2}+\sqrt{3})$ y que su grado sobre $\mathbb{Q}$ es 4: un ejemplo concreto del teorema del elemento primitivo.

(b) $\alpha = \sqrt[3]{2}+\sqrt[3]{4}$**.** Aquí $\sqrt[3]{4} = (\sqrt[3]{2})^2$, así que $\alpha \in \mathbb{Q}(\sqrt[3]{2})$ y su grado divide a 3; como α es irracional, el grado es exactamente 3. Escribiendo $\beta = \sqrt[3]{2}$ y usando $\beta^3 = 2$:

$$\alpha^3 = (\beta+\beta^2)^3 = \beta^3 + 3\beta^4 + 3\beta^5 + \beta^6 = 2 + 6\beta + 6\beta^2 + 4 = 6 + 6\alpha,$$

de donde

$$m_\alpha = x^3 - 6x - 6,$$

irreducible por Eisenstein en $p=2$ (o en $p=3$). Numéricamente $\alpha \approx 1{,}2599 + 1{,}5874 = 2{,}8473$ y $2{,}8473^3 - 6\cdot 2{,}8473 - 6 \approx 23{,}08 - 17{,}08 - 6 \approx 0$.

(c) $\alpha = \zeta_7 = e^{2\pi i/7}$**.** Anula a $x^7-1 = (x-1)(x^6+x^5+\cdots+1)$, y como $\alpha \neq 1$ el candidato es el segundo factor,

$$\Phi_7 = x^6 + x^5 + x^4 + x^3 + x^2 + x + 1,$$

irreducible porque $\Phi_7(x+1)$ cumple Eisenstein en $p=7$. Luego $[\mathbb{Q}(\zeta_7):\mathbb{Q}] = 6 = \varphi(7)$.

(d) $\alpha = \cos(2\pi/7)$**.** Este vive dentro del anterior: $\alpha = \tfrac{1}{2}(\zeta_7 + \zeta_7^{-1})$, el subcuerpo real, de índice 2 dentro de $\mathbb{Q}(\zeta_7)$. Su grado es entonces $6/2 = 3$ y el polinomio mínimo es

$$8x^3 + 4x^2 - 4x - 1,$$

que en forma mónica es $x^3 + \tfrac{1}{2}x^2 - \tfrac{1}{2}x - \tfrac{1}{8}$. Control: $\cos(2\pi/7) \approx 0{,}62349$, y $8(0{,}62349)^3 + 4(0{,}62349)^2 - 4(0{,}62349) - 1 \approx 1{,}9397 + 1{,}5550 - 2{,}4940 - 1 \approx 0{,}0007$, cero salvo redondeo.

Cuidado

Que un polinomio anule a α no lo convierte en el mínimo. En (b), $\alpha = \sqrt[3]{2}+\sqrt[3]{4}$ también es raíz de $(x^3-6x-6)(x^2+1)$, de grado 5; y en (c), ζ_7 es raíz de x^7-1, de grado 7. El mínimo es el de *menor* grado, y por el paso 7 divide a todos los demás.

Pausa para pensar

En (d) el grado bajó de 6 a 3 al pasar de ζ_7 a $\cos(2\pi/7)$. La razón es que la conjugación compleja $\zeta \mapsto \bar{\zeta}$ es un elemento de orden 2 del grupo de Galois, y el subcuerpo real es exactamente su cuerpo fijo. Éste es el mecanismo de la correspondencia de Galois funcionando en un caso que puedes tocar con las manos: subgrupo de orden 2 $\leftrightarrow$ subcuerpo de índice 2.

26.1.2. Cuerpos de descomposición y normalidad

Definición 26.10. El *cuerpo de descomposición* de $f \in K[x]$ es la menor extensión donde f factoriza en lineales. Existe (se adjuntan raíces iterando el cociente por irreducibles) y es único salvo isomorfismo. Una extensión finita L/K es *normal* si es cuerpo de descomposición de algún polinomio; equivalentemente, si todo irreducible de $K[x]$ con *una* raíz en L tiene *todas* sus raíces en L.

Ejemplo 26.11. $\mathbb{Q}(\sqrt[3]{2})/\mathbb{Q}$ **no** es normal: contiene una raíz de x^3-2 pero no las complejas. El cuerpo de descomposición de x^3-2 es $E=\mathbb{Q}(\sqrt[3]{2},\omega)$ con $\omega=e^{2\pi i/3}$ (raíces: $\sqrt[3]{2}, \omega\sqrt[3]{2}, \omega^2\sqrt[3]{2}$). Grado: $[\mathbb{Q}(\sqrt[3]{2}):\mathbb{Q}]=3$, y $\omega \notin \mathbb{Q}(\sqrt[3]{2}) \subset \mathbb{R}$ añade grado 2 (su mínimo es x^2+x+1): por la torre, $[E:\mathbb{Q}]=6$.

26.1.3. Separabilidad

Un polinomio es *separable* si no tiene raíces repetidas (en su cuerpo de descomposición); criterio: $\gcd(f,f')=1$. En característica 0 *todo* irreducible es separable (si f irreducible tuviera raíz doble, $f'\neq 0$ sería múltiplo común de grado menor: absurdo), así que sobre $\mathbb{Q}$ nunca habrá problema; las patologías inseparables solo viven en característica p (por ejemplo x^p-t sobre $\mathbb{F}_p(t)$). Una extensión es *separable* si el mínimo de cada elemento lo es.

Teorema 26.12 (Elemento primitivo)**.** *Toda extensión finita separable es* simple*: $L=K(\gamma)$ para algún γ. Por ejemplo, $\mathbb{Q}(\sqrt{2},\sqrt{3})=\mathbb{Q}(\sqrt{2}+\sqrt{3})$, como veremos en los ejercicios.*

26.2. Cimientos: qué es un cuerpo, qué es una extensión y qué mide exactamente el grado

Idea clave

Antes de nada. Todo lo que viene después en este capítulo —la correspondencia de Galois, la imposibilidad de la quíntica, los polígonos que se pueden dibujar y los que no— se apoya en cuatro objetos y en nada más: un *cuerpo*, una *extensión* de cuerpos, el *grado* de esa extensión y los *automorfismos* que la respetan. Esta sección construye los cuatro desde cero, con sus demostraciones y con las cuentas hechas enteras, para que quien llegue aquí sin haber visto nunca la palabra "cuerpo" pueda seguir el resto sin salir del capítulo.

La única idea que hay que llevarse en la mochila es esta: si K está dentro de L, entonces L **es un espacio vectorial sobre** K, y por tanto tiene dimensión. Esa dimensión se llama grado, se escribe $[L:K]$, y convierte preguntas sobre números en preguntas sobre dimensiones, que sabemos contar. Todo el capítulo es la explotación sistemática de esa frase.

26.2.1. Cuerpo, subcuerpo y característica

Un cuerpo es, sin adornos, un sitio donde se puede sumar, restar, multiplicar y dividir con todas las reglas de la aritmética escolar. Se define con axiomas porque queremos poder reconocer sitios así aunque sus elementos no sean números en el sentido corriente: los restos módulo 7 forman uno, y las funciones racionales también.

Definición 26.13 (Cuerpo)**.** Un *cuerpo* es un conjunto K con dos operaciones, suma $+$ y producto $\cdot$, y dos elementos distinguidos $0 \neq 1$, tales que para todos $a, b, c \in K$:

(C1) $a+b=b+a$ y $(a+b)+c=a+(b+c)$ (la suma es conmutativa y asociativa);

(C2) $a+0=a$, y existe $-a \in K$ con $a+(-a)=0$ (neutro y opuesto);

(C3) $ab=ba$ y $(ab)c=a(bc)$ (el producto es conmutativo y asociativo);

(C4) $a\cdot 1=a$, y si $a\neq 0$ existe $a^{-1}\in K$ con $a\,a^{-1}=1$ (neutro e inverso);

(C5) $a(b+c)=ab+ac$ (distributiva).

Restar y dividir se definen a partir de ahí: $a-b=a+(-b)$ y $a/b=a\,b^{-1}$ cuando $b\neq 0$.

El axioma que de verdad manda es (C4): la existencia de inversos. Es lo que separa a los cuerpos de los anillos, y es exactamente lo que hace falta para despejar en una ecuación lineal.

Ejemplo 26.14 (Cuatro cuerpos y un impostor)**. (a)** $\mathbb{Q}$**.** Los racionales. El inverso de p/q (con $p,q\neq 0$) es q/p. Es el cuerpo más pequeño que contiene a 1, y será el suelo de casi todo el capítulo.

(b) $\mathbb{R}$ **y** $\mathbb{C}$**.** Los reales cumplen los cinco axiomas, y los complejos también: el inverso de $a+b\mathrm{i}$ con $(a,b)\neq(0,0)$ es

$$(a+b\mathrm{i})^{-1}=\frac{a-b\mathrm{i}}{a^2+b^2},$$

porque $(a+b\mathrm{i})(a-b\mathrm{i})=a^2+b^2$, un real no nulo.

(c) $\mathbb{F}_p$ **con** p **primo.** Los restos $\{0,1,\dots,p-1\}$ con suma y producto módulo p. Todos los axiomas son evidentes salvo (C4). Veámoslo en $\mathbb{F}_5$ a mano, que es la forma honesta de creérselo:

$$1^{-1}=1,\qquad 2^{-1}=3\ \ (2\cdot 3=6\equiv 1),\qquad 3^{-1}=2,\qquad 4^{-1}=4\ \ (4\cdot 4=16\equiv 1).$$

Todos los no nulos tienen inverso, luego $\mathbb{F}_5$ es un cuerpo, y en él "dividir" tiene sentido literal: $3/4=3\cdot 4^{-1}=3\cdot 4=12\equiv 2$. En general, si $0<a<p$ entonces $\gcd(a,p)=1$ y la identidad

de Bézout da enteros u, v con $au + pv = 1$; reduciendo módulo p queda $a\,u \equiv 1$, así que u es el inverso buscado. La primalidad es imprescindible.

(d) El impostor: $\mathbb{Z}$. Los enteros cumplen (C1), (C2), (C3) y (C5), pero fallan (C4) estrepitosamente: el único entero x con $2x = 1$ sería $x = 1/2$, que no es entero. Así que $\mathbb{Z}$ *no* es un cuerpo. Tampoco lo es $\mathbb{Z}_6$: allí $2 \cdot 3 = 6 \equiv 0$ con ambos factores no nulos, y un cuerpo no puede tener *divisores de cero*, porque si $ab = 0$ con $a \neq 0$, multiplicando por a^{-1} sale $b = 0$.

Definición 26.15 (Subcuerpo)**.** Un subconjunto $K \subseteq L$ de un cuerpo L es un *subcuerpo* si contiene a 0 y a 1 y es cerrado para las cuatro operaciones: si $a, b \in K$ entonces $a + b$, $-a$, ab y (si $a \neq 0$) a^{-1} están en K. Un subcuerpo es, con las operaciones heredadas, un cuerpo por derecho propio.

Ejemplo 26.16 (Subcuerpos que ya conoces y uno que hay que comprobar)**.** $\mathbb{Q} \subseteq \mathbb{R} \subseteq \mathbb{C}$ es una cadena de subcuerpos. Más interesante:

$$\mathbb{Q}(\sqrt{2}) = \{a + b\sqrt{2} : \ a, b \in \mathbb{Q}\}$$

es un subcuerpo de $\mathbb{R}$. Suma y producto no dan problema, pues $(a + b\sqrt{2})(c + d\sqrt{2}) = (ac + 2bd) + (ad + bc)\sqrt{2}$. El inverso exige una cuenta: si $(a, b) \neq (0, 0)$,

$$\frac{1}{a + b\sqrt{2}} = \frac{a - b\sqrt{2}}{(a + b\sqrt{2})(a - b\sqrt{2})} = \frac{a - b\sqrt{2}}{a^2 - 2b^2} = \frac{a}{a^2 - 2b^2} - \frac{b}{a^2 - 2b^2}\sqrt{2},$$

que vuelve a tener la forma $a' + b'\sqrt{2}$ con $a', b' \in \mathbb{Q}$. Falta justificar que el denominador no se anula: si $a^2 = 2b^2$ con $b \neq 0$ entonces $(a/b)^2 = 2$ y $\sqrt{2}$ sería racional, que no lo es; y si $b = 0$ entonces $a \neq 0$ y el denominador es $a^2 \neq 0$. Truco a recordar: **para invertir, multiplica y divide por el conjugado**.

Hay una pregunta ingenua que resulta ser decisiva: si voy sumando $1 + 1 + 1 + \cdots$, ¿llego alguna vez a 0? En $\mathbb{Q}$ nunca; en $\mathbb{F}_5$ sí, a los cinco sumandos. Esa alternativa se llama característica y no admite término medio.

Definición 26.17 (Característica)**.** Escribamos $n \cdot 1 = 1 + 1 + \cdots + 1$ (n sumandos). La *característica* de K, denotada $\operatorname{car} K$, es el menor entero $n \geq 1$ con $n \cdot 1 = 0$ si tal n existe, y 0 si no existe ninguno.

Proposición 26.18 (La característica es 0 o un primo)**.** *Si* $\operatorname{car} K = n > 0$*, entonces* n *es primo.*

Demostración. Supongamos $n > 0$. Primero, $n \neq 1$, porque $1 \cdot 1 = 1 \neq 0$ por el axioma $0 \neq 1$. Supongamos ahora que n no fuera primo: $n = ab$ con $1 < a < n$ y $1 < b < n$. La distributiva da, sin más que desarrollar el producto de dos sumas de unos,

$$(a \cdot 1)(b \cdot 1) = (ab) \cdot 1 = n \cdot 1 = 0.$$

Como en un cuerpo no hay divisores de cero (ejemplo 26.14(d)), uno de los dos factores es cero: $a \cdot 1 = 0$ o $b \cdot 1 = 0$. Pero $a < n$ y $b < n$, y n era el *menor* entero positivo que anula al 1: contradicción. Luego n es primo. □

Observación 26.19. De ahí sale que todo cuerpo lleva dentro una copia de $\mathbb{Q}$ (si $\operatorname{car} K = 0$) o una de $\mathbb{F}_p$ (si $\operatorname{car} K = p$): basta mirar el subcuerpo generado por el 1. En este capítulo trabajaremos casi siempre en característica 0 —$\mathbb{Q}$, $\mathbb{R}$, $\mathbb{C}$ y sus extensiones—, y por eso muchas patologías no aparecerán nunca. El capítulo 25 trató la característica con más detalle, y el apartado sobre inseparabilidad, más adelante en este mismo capítulo, es justo el catálogo de lo que se rompe cuando $\operatorname{car} K = p$.

26.2.2. Extensiones de cuerpos y el grado

Definición 26.20 (Extensión y la notación L/K)**.** Si K es un subcuerpo de L, se dice que L es una *extensión* de K y se escribe L/K. La barra **no es un cociente**: es solo la manera universal de decir "L encima de K". Se lee "L sobre K". A K se le llama cuerpo base y a L cuerpo grande.

Cuidado

El símbolo L/K se parece a un cociente de grupos o de anillos, pero no lo es en absoluto: aquí no se está identificando nada. Es notación heredada, y coexiste en este mismo capítulo con cocientes de verdad como $K[x]/(m_\alpha)$. Distínguelos por lo que hay a los lados: cuerpos a ambos lados de la barra significa extensión.

Proposición 26.21 (La observación clave)**.** *Si L/K es una extensión, entonces L es un espacio vectorial sobre K, tomando como suma de vectores la suma de L y como producto por escalares la restricción del producto de L a $K \times L$.*

Demostración. Hay que comprobar los axiomas de espacio vectorial del capítulo 7, y todos son casos particulares de axiomas de cuerpo. $(L, +)$ es grupo abeliano por (C1) y (C2). Para $\lambda, \mu \in K$ y $u, v \in L$: $\lambda(u+v) = \lambda u + \lambda v$ es la distributiva (C5); $(\lambda+\mu)u = \lambda u + \mu u$ es la distributiva otra vez; $(\lambda\mu)u = \lambda(\mu u)$ es la asociativa del producto (C3); y $1 \cdot u = u$ es (C4). No hay nada más que hacer: el enunciado es una relectura, no un teorema profundo. Su valor está en lo que autoriza, no en su dificultad. □

Definición 26.22 (Grado de una extensión)**.** El *grado* de L/K es la dimensión de L como K-espacio vectorial:

$$[L : K] := \dim_K L.$$

La extensión es *finita* si $[L : K] < \infty$, e *infinita* en caso contrario.

Ejemplo 26.23 (Tres grados calculados de principio a fin)**. (a)** $[\mathbb{C} : \mathbb{R}] = 2$**.** Todo complejo se escribe $z = a + b\mathrm{i}$ con $a, b \in \mathbb{R}$ de forma única, así que $\{1, \mathrm{i}\}$ genera $\mathbb{C}$ sobre $\mathbb{R}$. Es libre: si $a \cdot 1 + b \cdot \mathrm{i} = 0$ con a, b reales, entonces $a = b = 0$ por definición de igualdad de complejos. Base de dos elementos, luego $[\mathbb{C} : \mathbb{R}] = 2$.

(b) $[\mathbb{Q}(\sqrt{2}) : \mathbb{Q}] = 2$**.** Por el ejemplo 26.16, todo elemento es $a + b\sqrt{2}$ con $a, b \in \mathbb{Q}$: $\{1, \sqrt{2}\}$ genera. ¿Es libre? Si $a + b\sqrt{2} = 0$ con $b \neq 0$, entonces $\sqrt{2} = -a/b \in \mathbb{Q}$, falso. Luego $b = 0$, y entonces $a = 0$. Base de dos elementos: $[\mathbb{Q}(\sqrt{2}) : \mathbb{Q}] = 2$.

(c) $[\mathbb{R}:\mathbb{Q}]$ **es infinito.** Si fuera finito, digamos n, entonces $\mathbb{R}$ tendría una base $\{v_1,\dots,v_n\}$ sobre $\mathbb{Q}$ y todo real sería $\sum q_i v_i$ con $q_i \in \mathbb{Q}$. Pero las n-tuplas de racionales forman un conjunto numerable, y $\mathbb{R}$ no lo es (capítulo 6). Contradicción. De hecho ni siquiera es numerable: $[\mathbb{R}:\mathbb{Q}]$ es infinito no numerable.

Pausa para pensar

Fíjate en lo que acaba de pasar en (c): una propiedad de *cardinalidad* ha decidido una cuestión de álgebra sin tocar un solo polinomio. El grado es un puente de doble sentido, y en este capítulo lo cruzaremos casi siempre en la dirección contraria: contando dimensiones para demostrar que ciertos números no existen.

26.2.3. Torres de extensiones: el teorema del grado

Si $K \subseteq L \subseteq M$ son cuerpos, se habla de una *torre* de extensiones. El resultado que sigue es el caballo de batalla del capítulo entero: los grados se multiplican al subir pisos.

Teorema 26.24 (Multiplicatividad del grado)**.** *Sea $K \subseteq L \subseteq M$ una torre de cuerpos. Entonces*

$$[M:K] = [M:L]\cdot[L:K],$$

en el sentido de que si los dos factores de la derecha son finitos, el de la izquierda también, y vale la igualdad. Además, si $\{u_1,\dots,u_m\}$ es base de L sobre K y $\{v_1,\dots,v_n\}$ es base de M sobre L, entonces los mn productos $\{u_i v_j\}$ forman una base de M sobre K.

Demostración. Probamos directamente la afirmación sobre la base producto; la igualdad de grados es su consecuencia inmediata, porque una base de mn elementos da dimensión mn.

Generan. Sea $z \in M$. Como $\{v_j\}$ es base de M sobre L, existen $\ell_1,\dots,\ell_n \in L$ con

$$z = \sum_{j=1}^{n} \ell_j v_j.$$

Como $\{u_i\}$ es base de L sobre K, cada coeficiente se desarrolla a su vez: $\ell_j = \sum_{i=1}^{m} c_{ij} u_i$ con $c_{ij} \in K$. Sustituyendo,

$$z = \sum_{j=1}^{n} \Big(\sum_{i=1}^{m} c_{ij} u_i \Big) v_j = \sum_{i=1}^{m}\sum_{j=1}^{n} c_{ij}\,(u_i v_j),$$

que es una combinación K-lineal de los $u_i v_j$. Luego generan M sobre K.

Son libres. Supongamos $\sum_{i,j} c_{ij} u_i v_j = 0$ con todos los $c_{ij} \in K$. Agrupamos por j:

$$\sum_{j=1}^{n} \Big(\underbrace{\sum_{i=1}^{m} c_{ij} u_i}_{=:\ \ell_j \ \in\ L} \Big) v_j = 0.$$

Cada ℓ_j está en L, porque es combinación de elementos $u_i \in L$ con coeficientes $c_{ij} \in K \subseteq L$. Tenemos entonces una combinación L-lineal nula de los v_j, que son libres sobre L: por tanto

$\ell_j = 0$ para todo j. Pero $\ell_j = \sum_i c_{ij} u_i = 0$ es una combinación K-lineal nula de los u_i, que son libres sobre K: por tanto $c_{ij} = 0$ para todos i, j. Los mn productos son libres.

Como generan y son libres, forman base, y hay exactamente mn de ellos (son distintos entre sí precisamente porque son libres). Luego $[M : K] = mn = [M : L][L : K]$. □

Corolario 26.25 (Control por divisibilidad)**.** *Si M/K es finita y L es un cuerpo intermedio, $K \subseteq L \subseteq M$, entonces $[L : K]$ divide a $[M : K]$.*

Demostración. Inmediato del teorema 26.24: $[M : K] = [M : L][L : K]$ exhibe a $[L : K]$ como factor. □

Este corolario es, de lejos, la herramienta de imposibilidad más barata del capítulo: si lo que quieres construir tiene un grado que no divide al grado de donde vives, no se puede, y punto. Así caerán la duplicación del cubo y la trisección del ángulo en la sección de regla y compás.

Ejemplo 26.26 (La torre $\mathbb{Q} \subset \mathbb{Q}(\sqrt{2}) \subset \mathbb{Q}(\sqrt{2}, \sqrt{3})$, entera)**.** Queremos $[\mathbb{Q}(\sqrt{2}, \sqrt{3}) : \mathbb{Q}]$, donde $\mathbb{Q}(\sqrt{2}, \sqrt{3})$ denota el menor subcuerpo de $\mathbb{R}$ que contiene a $\mathbb{Q}$, a $\sqrt{2}$ y a $\sqrt{3}$.

Primer piso. $[\mathbb{Q}(\sqrt{2}) : \mathbb{Q}] = 2$ con base $\{1, \sqrt{2}\}$, ya calculado en el ejemplo 26.23(b).

Segundo piso. Afirmamos que $[\mathbb{Q}(\sqrt{2}, \sqrt{3}) : \mathbb{Q}(\sqrt{2})] = 2$ con base $\{1, \sqrt{3}\}$. Que genera es claro (todo se escribe $c + d\sqrt{3}$ con $c, d \in \mathbb{Q}(\sqrt{2})$, y ese conjunto ya es cuerpo por el mismo truco del conjugado). Lo que hay que demostrar de verdad es que $\sqrt{3} \notin \mathbb{Q}(\sqrt{2})$, porque si estuviera el grado sería 1 y no 2. Supongamos que $\sqrt{3} = a + b\sqrt{2}$ con $a, b \in \mathbb{Q}$. Elevando al cuadrado:

$$3 = a^2 + 2b^2 + 2ab\sqrt{2}.$$

Si $ab \neq 0$, despejamos $\sqrt{2} = (3 - a^2 - 2b^2)/(2ab) \in \mathbb{Q}$, falso. Luego $ab = 0$. Si $b = 0$, queda $\sqrt{3} = a \in \mathbb{Q}$, falso. Si $a = 0$, queda $3 = 2b^2$, o sea $(2b)^2 = 6$, o sea $\sqrt{6} = |2b| \in \mathbb{Q}$, también falso. En los tres casos hay contradicción, así que $\sqrt{3} \notin \mathbb{Q}(\sqrt{2})$ y el segundo piso mide 2.

Multiplicamos. Por el teorema 26.24,

$$[\mathbb{Q}(\sqrt{2}, \sqrt{3}) : \mathbb{Q}] = [\mathbb{Q}(\sqrt{2}, \sqrt{3}) : \mathbb{Q}(\sqrt{2})] \cdot [\mathbb{Q}(\sqrt{2}) : \mathbb{Q}] = 2 \cdot 2 = 4.$$

La base, escrita. La demostración del teorema no solo da el número: da la base como producto de las dos bases parciales, $\{1, \sqrt{2}\} \times \{1, \sqrt{3}\}$, es decir

$$\{\, 1,\ \sqrt{2},\ \sqrt{3},\ \sqrt{6}\,\},$$

porque $\sqrt{2} \cdot \sqrt{3} = \sqrt{6}$. Todo elemento de $\mathbb{Q}(\sqrt{2}, \sqrt{3})$ se escribe de forma *única* como

$$a + b\sqrt{2} + c\sqrt{3} + d\sqrt{6}, \qquad a, b, c, d \in \mathbb{Q}.$$

Esta expresión con cuatro coeficientes es la que usaremos una y otra vez: es la que permite calcular cuerpos fijos resolviendo un sistema lineal en a, b, c, d.

Regla: calcular el grado de una extensión

El grado es el número que decide casi todo. Este es el procedimiento básico; los criterios finos de irreducibilidad se detallan en la regla de la página siguiente, "hallar el polinomio mínimo y el grado de una extensión".

1. **Escribe la extensión como una torre.** Si el cuerpo grande es $K(\alpha_1, \ldots, \alpha_r)$, adjunta los generadores *de uno en uno*:
$$K \subseteq K(\alpha_1) \subseteq K(\alpha_1, \alpha_2) \subseteq \cdots \subseteq K(\alpha_1, \ldots, \alpha_r).$$
2. **Mide cada piso sobre el piso inmediatamente anterior**, nunca sobre K. El grado del piso i-ésimo es el grado del polinomio mínimo de α_i *sobre* $K(\alpha_1, \ldots, \alpha_{i-1})$, que puede ser menor que su grado sobre K: adjuntar cosas puede volver reducible un polinomio que era irreducible.
3. **Multiplica** los grados de los pisos (teorema 26.24).
4. **Justifica cada bajada de grado.** Que un piso mida 1 significa que el elemento ya estaba, y eso hay que demostrarlo, no suponerlo. Que mida lo máximo también: el paso delicado del ejemplo 26.26 fue probar $\sqrt{3} \notin \mathbb{Q}(\sqrt{2})$.
5. **Controla con la divisibilidad.** Por el corolario 26.25, el grado de cualquier subextensión divide al total. Si te sale un subcuerpo de grado 3 dentro de una extensión de grado 4, hay un error garantizado en algún paso anterior.
6. **Controla con una cota superior.** Siempre $[K(\alpha, \beta) : K] \leq [K(\alpha) : K] \cdot [K(\beta) : K]$. Si tu resultado supera ese producto, revisa.

26.2.4. Elementos algebraicos, trascendentes y el polinomio mínimo

Definición 26.27 (Algebraico y trascendente)**.** Sea L/K una extensión y $\alpha \in L$. Se dice que α es *algebraico sobre* K si existe un polinomio no nulo $f \in K[x]$ con $f(\alpha) = 0$. Si no existe ninguno, α es *trascendente sobre* K.

Obsérvese que ser algebraico depende del cuerpo base: $\sqrt{2}$ es algebraico sobre $\mathbb{Q}$ (anula a $x^2 - 2$) y también sobre $\mathbb{R}$ (anula a $x - \sqrt{2}$), pero π es trascendente sobre $\mathbb{Q}$ y algebraico sobre $\mathbb{R}$. Nunca digas "α es algebraico" a secas.

Ejemplo 26.28 (Algebraicos y trascendentes)**.** $\sqrt{2}$, $\sqrt[3]{2}$, i y $\sqrt{2} + \sqrt{3}$ son algebraicos sobre $\mathbb{Q}$: anulan respectivamente a $x^2 - 2$, $x^3 - 2$, $x^2 + 1$ y $x^4 - 10x^2 + 1$ (esto último se comprueba abajo). En cambio e (Hermite, 1873) y π (Lindemann, 1882) son trascendentes sobre $\mathbb{Q}$: no anulan a ningún polinomio racional, y por tanto no viven dentro de ninguna extensión finita de $\mathbb{Q}$. La trascendencia de π es la que hace imposible la cuadratura del círculo, y se demuestra al final de este capítulo.

Teorema 26.29 (Polinomio mínimo: existencia, unicidad e irreducibilidad)**.** *Sea $\alpha \in L$ algebraico sobre K. Entonces existe un único polinomio $m_{\alpha,K} \in K[x]$ que es mónico, anula a α y tiene grado mínimo entre los que lo anulan. Además:*

(i) $m_{\alpha,K}$ divide a todo *$f \in K[x]$ con $f(\alpha) = 0$;*

(ii) $m_{\alpha,K}$ es irreducible en $K[x]$;

(iii) recíprocamente, si $g \in K[x]$ es mónico, irreducible y $g(\alpha) = 0$, entonces $g = m_{\alpha,K}$.

Demostración. Existencia. El conjunto de grados de polinomios no nulos de $K[x]$ que anulan a α no es vacío (porque α es algebraico) y es un conjunto de enteros ≥ 0; tiene mínimo. Tomamos f de ese grado mínimo y lo dividimos por su coeficiente principal, que es invertible por ser K cuerpo: sale mónico, sigue anulando a α y conserva el grado. Lo llamamos $m = m_{\alpha,K}$.

(i) Divide a todos. Sea $f \in K[x]$ con $f(\alpha) = 0$. La división euclídea en $K[x]$ da $f = qm + r$ con $r = 0$ o $\deg r < \deg m$. Evaluando en α: $0 = f(\alpha) = q(\alpha)m(\alpha) + r(\alpha) = r(\alpha)$. Si fuese $r \neq 0$, tendríamos un polinomio no nulo de grado menor que $\deg m$ anulando a α, contra la minimalidad. Luego $r = 0$ y $m \mid f$.

Unicidad. Si m' es otro mónico de grado mínimo que anula a α, por (i) se tiene $m \mid m'$ y, por el mismo argumento con los papeles cambiados, $m' \mid m$. Dos polinomios que se dividen mutuamente difieren en una constante multiplicativa; siendo ambos mónicos, esa constante es 1 y $m = m'$.

(ii) Irreducibilidad. Supongamos $m = gh$ con $\deg g < \deg m$ y $\deg h < \deg m$. Evaluando, $0 = m(\alpha) = g(\alpha)h(\alpha)$, y como L es un cuerpo (sin divisores de cero), $g(\alpha) = 0$ o $h(\alpha) = 0$. Cualquiera de las dos posibilidades da un polinomio no nulo de grado estrictamente menor que $\deg m$ que anula a α: contradice la minimalidad. Luego m no factoriza y es irreducible.

(iii) Recíproco. Si g es mónico, irreducible y $g(\alpha) = 0$, por (i) $m \mid g$. Como g es irreducible, sus únicos divisores son las constantes y los múltiplos constantes de g; siendo m no constante (pues α anula a m y m es mónico no nulo de grado ≥ 1), debe ser $m = cg$ con c constante, y siendo ambos mónicos, $c = 1$. □

La parte (iii) es la que se usa en la práctica: para identificar el polinomio mínimo basta fabricar *cualquier* polinomio mónico irreducible que anule al elemento. No hay que buscar el de menor grado a ciegas.

Ejemplo 26.30 (El polinomio mínimo de $\sqrt{2} + \sqrt{3}$, hecho entero)**.** Sea $\alpha = \sqrt{2} + \sqrt{3}$ y $K = \mathbb{Q}$.

Paso 1: fabricar un anulador. Elevamos al cuadrado, con cuidado con el término cruzado:

$$\alpha^2 = (\sqrt{2})^2 + 2\sqrt{2}\sqrt{3} + (\sqrt{3})^2 = 2 + 2\sqrt{6} + 3 = 5 + 2\sqrt{6}.$$

Aislamos el radical superviviente y volvemos a elevar:

$$\alpha^2 - 5 = 2\sqrt{6} \implies (\alpha^2 - 5)^2 = 4 \cdot 6 = 24 \implies \alpha^4 - 10\alpha^2 + 25 = 24,$$

es decir $\alpha^4 - 10\alpha^2 + 1 = 0$. Ya tenemos el candidato $f(x) = x^4 - 10x^2 + 1$.

Paso 2: irreducibilidad sobre $\mathbb{Q}$. Hay que descartar dos cosas.

Factores lineales. Una raíz racional de un mónico con coeficientes enteros es un divisor entero del término independiente, aquí ± 1. Y $f(1) = 1 - 10 + 1 = -8 \neq 0$, $f(-1) = -8 \neq 0$. No hay raíces racionales, luego no hay factores de grado 1 (ni, por tanto, de grado 3, pues un factor cúbico obligaría a otro lineal).

Producto de dos cuadráticos. Si $f = (x^2 + ax + b)(x^2 + a'x + b')$ con todo en $\mathbb{Q}$, comparando el coeficiente de x^3 sale $a + a' = 0$, o sea $a' = -a$. Desarrollando el producto $(x^2 + ax + b)(x^2 - ax + b')$:

$$x^4 + (b + b' - a^2)x^2 + a(b' - b)x + bb'.$$

Igualando coeficientes con $x^4 - 10x^2 + 0 \cdot x + 1$:

$$b + b' - a^2 = -10, \qquad a(b' - b) = 0, \qquad bb' = 1.$$

De la segunda, $a = 0$ o $b = b'$.

- Si $a = 0$: entonces $b + b' = -10$ y $bb' = 1$, luego b, b' son las raíces de $t^2 + 10t + 1 = 0$, es decir $t = -5 \pm 2\sqrt{6}$, irracionales. No sirven.
- Si $b = b'$: entonces $b^2 = 1$, o sea $b = \pm 1$, y $2b - a^2 = -10$ da $a^2 = 12$ (si $b = 1$) o $a^2 = 8$ (si $b = -1$). Ni $\sqrt{12}$ ni $\sqrt{8}$ son racionales. Tampoco sirven.

No hay factorización posible en $\mathbb{Q}[x]$: f es irreducible.

Paso 3: conclusión. f es mónico, irreducible y anula a α; por el teorema 26.29(iii),

$$m_{\alpha,\mathbb{Q}}(x) = x^4 - 10x^2 + 1, \qquad [\mathbb{Q}(\alpha) : \mathbb{Q}] = 4.$$

Paso 4: control numérico. $\alpha \approx 1{,}41421 + 1{,}73205 = 3{,}14626$, luego $\alpha^2 \approx 9{,}89898$ y $\alpha^4 \approx 97{,}9898$; entonces $97{,}9898 - 10 \cdot 9{,}89898 + 1 \approx 97{,}9898 - 98{,}9898 + 1 = 0$. ✓

Paso 5: la sorpresa. Tenemos $\mathbb{Q}(\alpha) \subseteq \mathbb{Q}(\sqrt{2}, \sqrt{3})$ y ambos cuerpos tienen grado 4 sobre $\mathbb{Q}$ (ejemplo 26.26), así que *son el mismo*. Se puede ver a mano: desarrollando,

$$\alpha^3 = 11\sqrt{2} + 9\sqrt{3} \implies \tfrac{1}{2}(\alpha^3 - 9\alpha) = \tfrac{1}{2}(11\sqrt{2} + 9\sqrt{3} - 9\sqrt{2} - 9\sqrt{3}) = \sqrt{2},$$

de modo que $\sqrt{2} \in \mathbb{Q}(\alpha)$, y con él $\sqrt{3} = \alpha - \sqrt{2}$. Dos generadores fundidos en uno solo: es el teorema del elemento primitivo, que se demuestra en general más adelante en este capítulo.

Definición 26.31 (Extensión algebraica)**.** Una extensión L/K es *algebraica* si todo elemento de L es algebraico sobre K.

Proposición 26.32 (Finita implica algebraica)**.** *Si $[L : K] = n < \infty$, entonces L/K es algebraica, y el polinomio mínimo de cada $\alpha \in L$ tiene grado a lo sumo n.*

Demostración. Sea $\alpha \in L$. Los $n + 1$ elementos

$$1,\ \alpha,\ \alpha^2,\ \ldots,\ \alpha^n$$

viven en L, que es un K-espacio vectorial de dimensión n. Cualquier familia de $n + 1$ vectores en un espacio de dimensión n es linealmente dependiente (capítulo 7). Existen, pues, $c_0, \ldots, c_n \in K$ no todos nulos con

$$c_0 + c_1\alpha + \cdots + c_n\alpha^n = 0.$$

El polinomio $f(x) = c_0 + c_1 x + \cdots + c_n x^n$ es no nulo (algún $c_i \neq 0$), está en $K[x]$ y anula a α: α es algebraico. Y su polinomio mínimo divide a f, luego tiene grado $\leq n$. □

Cuidado

El recíproco es falso: hay extensiones algebraicas infinitas. El cuerpo $\overline{\mathbb{Q}}$ de todos los números algebraicos sobre $\mathbb{Q}$ es algebraico sobre $\mathbb{Q}$ y de grado infinito, porque contiene $\sqrt[n]{2}$ para todo n, con grados n arbitrariamente grandes. "Algebraica" y "finita" no son sinónimos: finita es estrictamente más fuerte.

26.2.5. $K[\alpha]$ frente a $K(\alpha)$: el pequeño milagro

Cuando se adjunta un elemento $\alpha \in L$ a un cuerpo K hay dos candidatos naturales al resultado, y a primera vista son distintos.

Definición 26.33 (Los dos conjuntos, y extensión simple)**.** Sea L/K y $\alpha \in L$. Se define

$$K[\alpha] := \{f(\alpha) : \ f \in K[x]\} = \{c_0 + c_1\alpha + \cdots + c_r\alpha^r : \ r \geq 0, \ c_i \in K\},$$

el conjunto de las *expresiones polinómicas* en α, y $K(\alpha)$ como el menor subcuerpo de L que contiene a K y a α; equivalentemente,

$$K(\alpha) = \left\{ \tfrac{f(\alpha)}{g(\alpha)} : \ f, g \in K[x], \ g(\alpha) \neq 0 \right\},$$

el conjunto de las *expresiones racionales* en α. Una extensión L/K se llama *simple* si $L = K(\alpha)$ para algún α, llamado entonces *elemento primitivo*. Con varios generadores se escribe $K(\alpha_1, \ldots, \alpha_r)$, definido igual como el menor subcuerpo que los contiene a todos.

Siempre $K[\alpha] \subseteq K(\alpha)$, porque $f(\alpha) = f(\alpha)/1$. La inclusión contraria es el problema: un cociente no tiene por qué ser un polinomio. Y sin embargo, cuando α es algebraico, lo es.

Teorema 26.34 (Estructura de $K(\alpha)$ para α algebraico)**.** *Sea α algebraico sobre K con polinomio mínimo $m = m_{\alpha,K}$ de grado d. Entonces:*

(i) $K[\alpha] = K(\alpha)$*;*

(ii) $\{1, \alpha, \alpha^2, \ldots, \alpha^{d-1}\}$ *es una base de $K(\alpha)$ sobre K;*

(iii) $[K(\alpha) : K] = d = \deg m_{\alpha,K}$.

Demostración. (ii) primero, la parte de generación. Sea $z \in K[\alpha]$, digamos $z = f(\alpha)$ con $f \in K[x]$. Dividimos f entre m: $f = qm + r$ con $r = 0$ o $\deg r < d$. Evaluando en α y usando $m(\alpha) = 0$:

$$z = f(\alpha) = q(\alpha) \cdot 0 + r(\alpha) = r(\alpha).$$

Como $\deg r < d$, z es combinación K-lineal de $1, \alpha, \ldots, \alpha^{d-1}$. Luego esos d elementos generan $K[\alpha]$ sobre K.

Independencia. Si $c_0 + c_1\alpha + \cdots + c_{d-1}\alpha^{d-1} = 0$ con los $c_i \in K$ no todos nulos, el polinomio $c_0 + c_1 x + \cdots + c_{d-1}x^{d-1}$ sería no nulo, de grado $< d$ y anularía a α: contradice la minimalidad de m. Luego todos los c_i son cero y la familia es libre.

(i) $K[\alpha]$ es un cuerpo. Es claro que $K[\alpha]$ es cerrado para suma, resta y producto (suma y producto de polinomios son polinomios). Falta el inverso. Sea $z = f(\alpha) \neq 0$ con $f \in K[x]$; podemos suponer $\deg f < d$ por el argumento anterior. Como m es irreducible y f es no nulo de grado menor que $\deg m$, el único divisor común de f y m es una constante: $\gcd(f, m) = 1$. La identidad de Bézout en $K[x]$ da polinomios $u, v \in K[x]$ con

$$u\,f + v\,m = 1.$$

Evaluando en α: $u(\alpha)f(\alpha) + v(\alpha)\cdot 0 = 1$, es decir $u(\alpha)\,z = 1$. Así $z^{-1} = u(\alpha) \in K[\alpha]$. Por tanto $K[\alpha]$ es un subcuerpo que contiene a K y a α, luego contiene a $K(\alpha)$ por minimalidad de este; y como también $K[\alpha] \subseteq K(\alpha)$, son iguales.

(iii) Es (ii): una base de d elementos da dimensión d. □

Observación 26.35. La demostración anterior dice, en el fondo, que la aplicación $K[x] \to K[\alpha]$ que manda $f \mapsto f(\alpha)$ es sobreyectiva con núcleo el ideal (m), de donde $K(\alpha) \cong K[x]/(m_{\alpha,K})$. Es la "máquina de fabricar cuerpos" del capítulo 25: cocientar el anillo de polinomios por un irreducible produce un cuerpo. Aquí solo hemos necesitado la versión concreta, con Bézout a la vista.

Ejemplo 26.36 (Invertir $1+\beta$, con β la raíz cúbica real de dos, paso a paso)**.** Sea $\beta = \sqrt[3]{2}$. Su polinomio mínimo sobre $\mathbb{Q}$ es $m(x) = x^3 - 2$: anula a β, es mónico y es irreducible por Eisenstein con $p = 2$ (el primo 2 divide a los coeficientes $0, 0, -2$ pero 4 no divide a -2). Luego $[\mathbb{Q}(\beta) : \mathbb{Q}] = 3$ y la base es $\{1, \beta, \beta^2\} = \{1, \sqrt[3]{2}, \sqrt[3]{4}\}$. La pregunta: ¿cómo se escribe

$$z = \frac{1}{1 + \sqrt[3]{2}}$$

en esa base, es decir, con la forma $a + b\sqrt[3]{2} + c\sqrt[3]{4}$ y $a, b, c \in \mathbb{Q}$?

Paso 1: identificar el polinomio. Aquí $z = 1/f(\beta)$ con $f(x) = x + 1$.

Paso 2: comprobar que $\gcd(f, m) = 1$. Como m es irreducible de grado 3 y $\deg f = 1 < 3$, el único divisor común posible es una constante. (Concretamente, $f = x + 1$ dividiría a m solo si $m(-1) = 0$, y $m(-1) = (-1)^3 - 2 = -3 \neq 0$.)

Paso 3: Bézout, con la división euclídea explícita. Dividimos m entre f:

$$x^3 - 2 = (x + 1)(x^2 - x + 1) - 3,$$

lo que se comprueba desarrollando $(x+1)(x^2 - x + 1) = x^3 + 1$ y restando 3. Despejamos el 1:

$$3 = (x+1)(x^2 - x + 1) - (x^3 - 2) \implies 1 = (x+1)\cdot \tfrac{1}{3}(x^2 - x + 1) - \tfrac{1}{3}(x^3 - 2).$$

Es decir, $u(x) = \frac{1}{3}(x^2 - x + 1)$ y $v(x) = -\frac{1}{3}$ cumplen $uf + vm = 1$.

Paso 4: evaluar en β. Como $m(\beta) = 0$, el segundo sumando desaparece:

$$1 = (\beta+1)\cdot \tfrac{1}{3}(\beta^2 - \beta + 1) \implies \frac{1}{1 + \sqrt[3]{2}} = \frac{1 - \sqrt[3]{2} + \sqrt[3]{4}}{3} = \frac{1}{3} - \frac{1}{3}\sqrt[3]{2} + \frac{1}{3}\sqrt[3]{4}.$$

Ya está escrito en la base: las coordenadas respecto de $\{1, \beta, \beta^2\}$ son $a = \frac{1}{3}$, $b = -\frac{1}{3}$ y $c = \frac{1}{3}$.

Paso 5: dos controles. Multiplicando directamente,

$$(1+\beta)(1-\beta+\beta^2) = 1-\beta+\beta^2+\beta-\beta^2+\beta^3 = 1+\beta^3 = 1+2 = 3, \ \checkmark$$

que es exactamente lo que hacía falta. Y numéricamente: $\sqrt[3]{2} \approx 1{,}259921$, $\sqrt[3]{4} \approx 1{,}587401$, luego el miembro derecho vale $(1 - 1{,}259921 + 1{,}587401)/3 = 1{,}327480/3 \approx 0{,}442493$; y el izquierdo, $1/(1 + 1{,}259921) = 1/2{,}259921 \approx 0{,}442493$. ✓

Cuidado

La hipótesis "α algebraico" del teorema 26.34 no se puede quitar. Si t es trascendente sobre K, entonces $K[t]$ es un anillo de polinomios de verdad —con infinitos elementos linealmente independientes $1, t, t^2, \ldots$— y $1/t$ *no* es un polinomio en t: allí $K[t] \subsetneq K(t)$ y la extensión $K(t)/K$ es infinita. El milagro de que "los cocientes se conviertan en polinomios" es exactamente el privilegio de lo algebraico.

26.2.6. Automorfismos: las simetrías de una extensión

Llegamos al segundo protagonista. Un automorfismo de un cuerpo es una biyección que no nota las operaciones: si la aplicas antes o después de sumar y multiplicar, da lo mismo.

Definición 26.37 (Automorfismo de cuerpos y K-automorfismo)**.** Sea L un cuerpo. Un *automorfismo* de L es una biyección $\sigma\colon L \to L$ tal que, para todos $a, b \in L$,

$$\sigma(a+b) = \sigma(a) + \sigma(b), \qquad \sigma(ab) = \sigma(a)\sigma(b).$$

Si además L/K es una extensión y $\sigma(a) = a$ para todo $a \in K$, se dice que σ *fija* K *punto a punto*, o que es un K*-automorfismo*, y se escribe $\sigma|_K = \mathrm{id}$.

De los dos axiomas se sigue automáticamente que $\sigma(0) = 0$, $\sigma(1) = 1$, $\sigma(-a) = -\sigma(a)$ y $\sigma(a^{-1}) = \sigma(a)^{-1}$. Para $\sigma(1) = 1$: de $\sigma(1) = \sigma(1 \cdot 1) = \sigma(1)^2$ y $\sigma(1) \neq 0$ (por inyectividad, ya que $\sigma(0) = 0$) se despeja $\sigma(1) = 1$. En particular todo automorfismo fija al cuerpo primo: sobre $\mathbb{Q}$, *ningún* automorfismo mueve un racional.

Definición 26.38 (El grupo de Galois)**.** Para una extensión L/K se define

$$\mathrm{Gal}(L/K) := \{\sigma\colon L \to L \text{ automorfismo con } \sigma|_K = \mathrm{id}\},$$

con la composición de aplicaciones como operación.

Proposición 26.39 ($\mathrm{Gal}(L/K)$ es un grupo)**.** *El conjunto* $\mathrm{Gal}(L/K)$ *con la composición es un grupo.*

Demostración. Operación interna. Si $\sigma, \tau \in \mathrm{Gal}(L/K)$, la composición $\sigma \circ \tau$ es biyectiva (composición de biyecciones) y respeta las operaciones: $\sigma(\tau(a+b)) = \sigma(\tau(a) + \tau(b)) = \sigma(\tau(a)) + \sigma(\tau(b))$, e igual para el producto. Y fija K: si $a \in K$, $\tau(a) = a$ y luego $\sigma(a) = a$.

Asociatividad. La composición de aplicaciones siempre es asociativa.

Neutro. La identidad id_L es un automorfismo y fija K trivialmente.

Inversos. Si σ es biyectiva, existe σ^{-1} como aplicación. Respeta la suma: dados $a, b \in L$, escribamos $a = \sigma(a')$, $b = \sigma(b')$; entonces $\sigma^{-1}(a+b) = \sigma^{-1}(\sigma(a') + \sigma(b')) = \sigma^{-1}(\sigma(a'+b')) = a' + b' = \sigma^{-1}(a) + \sigma^{-1}(b)$, y lo mismo para el producto. Y fija K: si $a \in K$ entonces $\sigma(a) = a$, luego $\sigma^{-1}(a) = a$. □

Teorema 26.40 (Los automorfismos permutan las raíces)**.** *Sea L/K una extensión, $\sigma \in \mathrm{Gal}(L/K)$ y $f \in K[x]$. Si $\alpha \in L$ cumple $f(\alpha) = 0$, entonces $\sigma(\alpha)$ también es raíz de f. En particular, si α es algebraico sobre K, entonces $\sigma(\alpha)$ es raíz del polinomio mínimo $m_{\alpha,K}$, y σ restringe a una permutación del conjunto* finito *de raíces de $m_{\alpha,K}$ que estén en L.*

Demostración. Escribamos $f(x) = c_0 + c_1 x + \cdots + c_n x^n$ con todos los $c_i \in K$. Aplicamos σ a la igualdad $f(\alpha) = 0$ y usamos que respeta sumas y productos, y que fija los coeficientes por estar en K:

$$0 = \sigma(0) = \sigma(c_0 + c_1\alpha + \cdots + c_n\alpha^n) = \sigma(c_0) + \sigma(c_1)\sigma(\alpha) + \cdots + \sigma(c_n)\sigma(\alpha)^n$$
$$= c_0 + c_1\sigma(\alpha) + \cdots + c_n\sigma(\alpha)^n = f(\sigma(\alpha)).$$

(La igualdad $\sigma(\alpha^k) = \sigma(\alpha)^k$ es la multiplicatividad aplicada $k-1$ veces.) Luego $\sigma(\alpha)$ es raíz de f. Tomando $f = m_{\alpha,K}$ se obtiene la segunda afirmación; y como σ es inyectiva y el conjunto de raíces de $m_{\alpha,K}$ en L es finito (a lo sumo $\deg m$), σ lo permuta. □

Idea clave

Esto es la teoría entera en una línea. Un automorfismo no puede mandar $\sqrt{2}$ a 7, ni a $\sqrt{3}$: solo puede mandarlo a una raíz de $x^2 - 2$, es decir, a $\sqrt{2}$ o a $-\sqrt{2}$. Como además un automorfismo queda completamente determinado por las imágenes de los generadores —por el teorema 26.34, todo elemento es un polinomio en ellos—, **contar automorfismos se reduce a contar elecciones coherentes de raíces**. De ahí sale, de golpe, la cota $|\mathrm{Gal}(L/K)| \le [L:K]$ y el método práctico de la regla siguiente. Todo lo demás del capítulo consiste en explotar esta observación.

Ejemplo 26.41 ($\mathrm{Gal}(\mathbb{Q}(\sqrt{2},\sqrt{3})/\mathbb{Q})$, calculado entero)**.** Sea $L = \mathbb{Q}(\sqrt{2},\sqrt{3})$, de grado 4 sobre $\mathbb{Q}$ con base $\{1, \sqrt{2}, \sqrt{3}, \sqrt{6}\}$ (ejemplo 26.26).

Paso 1: qué puede hacer un automorfismo. Sea $\sigma \in \mathrm{Gal}(L/\mathbb{Q})$. Por el teorema 26.40, $\sigma(\sqrt{2})$ ha de ser raíz de $x^2 - 2$, luego $\sigma(\sqrt{2}) = \pm\sqrt{2}$; y $\sigma(\sqrt{3})$ ha de ser raíz de $x^2 - 3$, luego $\sigma(\sqrt{3}) = \pm\sqrt{3}$. Como todo elemento de L es $a + b\sqrt{2} + c\sqrt{3} + d\sqrt{6}$ con $a, b, c, d \in \mathbb{Q}$ y σ fija los racionales y respeta productos,

$$\sigma(a + b\sqrt{2} + c\sqrt{3} + d\sqrt{6}) = a + b\,\sigma(\sqrt{2}) + c\,\sigma(\sqrt{3}) + d\,\sigma(\sqrt{2})\sigma(\sqrt{3}),$$

así que σ queda *determinado* por los dos signos. A lo sumo hay $2 \cdot 2 = 4$ automorfismos.

Paso 2: los cuatro existen de verdad. Definimos

$$\mathrm{id}, \qquad \sigma : \begin{cases} \sqrt{2} \mapsto -\sqrt{2} \\ \sqrt{3} \mapsto \sqrt{3} \end{cases} \qquad \tau : \begin{cases} \sqrt{2} \mapsto \sqrt{2} \\ \sqrt{3} \mapsto -\sqrt{3} \end{cases} \qquad \sigma\tau : \begin{cases} \sqrt{2} \mapsto -\sqrt{2} \\ \sqrt{3} \mapsto -\sqrt{3}. \end{cases}$$

Que cada uno es realmente un automorfismo se comprueba con la fórmula del paso 1: es $\mathbb{Q}$-lineal por construcción, biyectiva porque su cuadrado es la identidad, y multiplicativa porque los signos se propagan bien al producto. Por ejemplo, para σ: $\sigma(\sqrt{6}) = \sigma(\sqrt{2})\sigma(\sqrt{3}) = (-\sqrt{2})(\sqrt{3}) = -\sqrt{6}$, y basta verificar la multiplicatividad sobre la base, que es un cálculo finito. Nótese que hay exactamente $4 = [L : \mathbb{Q}]$ automorfismos: la extensión es de Galois, en el sentido que se define en la sección siguiente.

Paso 3: la tabla del grupo. Cada uno de los tres no triviales cambia signos de forma independiente, y aplicar dos veces el mismo cambio de signo devuelve el original:

$$\sigma^2 = \tau^2 = (\sigma\tau)^2 = \mathrm{id}, \qquad \sigma\tau = \tau\sigma.$$

Un grupo de orden 4 en el que todo elemento distinto del neutro tiene orden 2 es el *grupo de Klein* $V_4 \cong \mathbb{Z}/2 \times \mathbb{Z}/2$, no el cíclico $\mathbb{Z}/4$ (que sí tiene elementos de orden 4). Por tanto

$$\mathrm{Gal}\left(\mathbb{Q}(\sqrt{2}, \sqrt{3})/\mathbb{Q}\right) \cong \mathbb{Z}/2 \times \mathbb{Z}/2.$$

Paso 4: comprobación con un elemento concreto. Tomemos $z = 1 + \sqrt{2} + 2\sqrt{3} - \sqrt{6}$. Entonces $\sigma(z) = 1 - \sqrt{2} + 2\sqrt{3} + \sqrt{6}$ y $\tau(z) = 1 + \sqrt{2} - 2\sqrt{3} + \sqrt{6}$, y se verifica que $\sigma(\tau(z)) = 1 - \sqrt{2} - 2\sqrt{3} - \sqrt{6} = \sigma\tau(z)$, coherente con la tabla.

Regla: calcular a mano el grupo de Galois de una extensión concreta

Este es el procedimiento elemental, que solo usa lo definido hasta aquí. Más adelante, la regla "calcular el grupo de Galois de un polinomio concreto" con discriminante y resolvente, y el método de Dedekind, lo afinan para casos donde contar a mano se hace pesado.

1. **Halla todas las raíces de** f (simbólica o descriptivamente: "las cuatro $\mathrm{i}^k \sqrt[4]{2}$", no hace falta valor decimal).
2. **Escribe el cuerpo de descomposición** L adjuntando *todas* las raíces a K. No basta con una: la pregunta operativa es "adjuntada α_1, ¿están ya las demás dentro?".
3. **Calcula** $[L : K]$ **con la torre** (regla anterior). Ese número es el candidato al orden del grupo.
4. **Lista las imágenes posibles de cada generador.** Por el teorema 26.40, cada generador solo puede ir a una raíz de *su* polinomio mínimo. Multiplica el número de opciones: es una cota superior de $|\,\mathrm{Gal}(L/K)|$.
5. **Descarta las combinaciones incompatibles.** Si entre los generadores hay relaciones (por ejemplo $\sqrt{6} = \sqrt{2}\sqrt{3}$, o que un cuerpo sea real), algunas elecciones de signos no dan automorfismo.
6. **Compara con el grado.** Si el recuento coincide con $[L : K]$, has encontrado el grupo entero. Siempre $|\,\mathrm{Gal}(L/K)| \leq [L : K]$.
7. **Identifica el grupo abstracto** calculando los órdenes de sus elementos: orden 4 con un elemento de orden 4 es $\mathbb{Z}/4$; orden 4 con todos de orden 2 es el de Klein; orden 6 no conmutativo es S_3.

8. **Control.** $|\mathrm{Gal}(L/K)|$ debe dividir a $n!$ (actúa por permutaciones de las n raíces) y, si f es irreducible, debe ser múltiplo de n.

26.2.7. Cuerpo de descomposición y cuerpo fijo

Definición 26.42 (Cuerpo de descomposición)**.** Sea $f \in K[x]$ no constante. Una extensión L/K es un *cuerpo de descomposición* de f sobre K si se cumplen las dos condiciones a la vez:

(a) f factoriza en factores lineales dentro de $L[x]$, es decir $f = c(x-\alpha_1)\cdots(x-\alpha_n)$ con $\alpha_i \in L$;

(b) $L = K(\alpha_1, \ldots, \alpha_n)$, o sea, L está generado por esas raíces: no sobra nada.

Teorema 26.43 (Existencia)**.** *Todo $f \in K[x]$ no constante tiene un cuerpo de descomposición sobre K, y su grado sobre K es a lo sumo* $(\deg f)!$*. (La unicidad salvo isomorfismo se demuestra más adelante, con el teorema de extensión.)*

Demostración. Por inducción sobre $n = \deg f$. Si $n = 1$, f ya es lineal y $L = K$ sirve. Si $n > 1$, sea g un factor irreducible de f en $K[x]$. El cociente $K_1 := K[x]/(g)$ es un cuerpo (el cociente de un anillo de polinomios por un irreducible lo es, capítulo 25), contiene una copia de K, y la clase α_1 de x cumple $g(\alpha_1) = 0$, luego $f(\alpha_1) = 0$; además $[K_1 : K] = \deg g \le n$. En $K_1[x]$ podemos escribir $f = (x-\alpha_1)f_1$ con $\deg f_1 = n-1$. Por hipótesis de inducción, f_1 tiene un cuerpo de descomposición L sobre K_1, con $[L : K_1] \le (n-1)!$, y en él f factoriza del todo. Tomando el subcuerpo generado por K y todas las raíces se cumplen (a) y (b), y por la torre $[L : K] \le n \cdot (n-1)! = n!$. □

Ejemplo 26.44 (Tres cuerpos de descomposición sobre $\mathbb{Q}$)**. (a)** $f = x^2 - 2$**.** Sus raíces son $\pm\sqrt{2}$. Adjuntando una ya está la otra (es su opuesta), así que $L = \mathbb{Q}(\sqrt{2})$ y $[L : \mathbb{Q}] = 2$. Aquí adjuntar *una* raíz bastó.

(b) $f = x^3 - 2$**.** Sus raíces son $\sqrt[3]{2}$, $\omega\sqrt[3]{2}$ y $\omega^2\sqrt[3]{2}$, con $\omega = \mathrm{e}^{2\pi\mathrm{i}/3} = -\frac{1}{2} + \frac{\sqrt{3}}{2}\mathrm{i}$, que cumple $\omega^3 = 1$ y $\omega^2 + \omega + 1 = 0$. Aquí adjuntar una raíz *no* basta: $\mathbb{Q}(\sqrt[3]{2})$ está contenido en $\mathbb{R}$ y las otras dos raíces no son reales. El cuerpo de descomposición es

$$L = \mathbb{Q}(\sqrt[3]{2},\ \omega).$$

Grado: el primer piso mide 3 (Eisenstein con $p = 2$ sobre $x^3 - 2$); el segundo mide 2, porque el polinomio mínimo de ω sobre $\mathbb{Q}$ es $x^2 + x + 1$ (irreducible: sus raíces no son reales, luego no son racionales) y sigue siendo irreducible sobre $\mathbb{Q}(\sqrt[3]{2})$ por ser este un cuerpo real. Por la torre,

$$[L : \mathbb{Q}] = 2 \cdot 3 = 6.$$

Ese 6 es el número que después identificará el grupo como S_3.

(c) $f = x^4 + 1$**.** Sus raíces son los cuatro complejos $\frac{\pm 1 \pm \mathrm{i}}{\sqrt{2}}$, es decir, las raíces primitivas octavas de la unidad ζ_8^k con $k = 1, 3, 5, 7$. Adjuntando $\zeta_8 = \frac{1+\mathrm{i}}{\sqrt{2}}$ se obtienen todas (son sus potencias impares), así que $L = \mathbb{Q}(\zeta_8)$. Y como

$$\zeta_8 + \zeta_8^{-1} = 2\cos\frac{\pi}{4} = \sqrt{2}, \qquad \zeta_8^2 = \mathrm{i},$$

resulta $L = \mathbb{Q}(\sqrt{2}, \mathrm{i})$, con $[L:\mathbb{Q}] = 2 \cdot 2 = 4$ (el segundo piso mide 2 porque x^2+1 no tiene raíces en el cuerpo real $\mathbb{Q}(\sqrt{2})$). Obsérvese la sorpresa: un polinomio de grado 4 cuyo cuerpo de descomposición tiene grado 4 y no 24.

Definición 26.45 (Cuerpo fijo de un subgrupo)**.** Sea L un cuerpo y H un subgrupo del grupo de automorfismos de L. El *cuerpo fijo* de H es

$$L^H := \{x \in L : \ \sigma(x) = x \text{ para todo } \sigma \in H\}.$$

Es un subcuerpo de L: si $\sigma(x) = x$ y $\sigma(y) = y$, entonces $\sigma(x+y) = x+y$, $\sigma(xy) = xy$, $\sigma(-x) = -x$ y $\sigma(x^{-1}) = x^{-1}$, y además $0, 1 \in L^H$ siempre.

Ejemplo 26.46 (Un cuerpo fijo calculado)**.** En $L = \mathbb{Q}(\sqrt{2}, \sqrt{3})$ con los automorfismos del ejemplo 26.41, tomemos $H = \{\mathrm{id}, \sigma\}$, donde σ cambia el signo de $\sqrt{2}$ y deja $\sqrt{3}$. Un elemento genérico es $z = a + b\sqrt{2} + c\sqrt{3} + d\sqrt{6}$, y

$$\sigma(z) = a - b\sqrt{2} + c\sqrt{3} - d\sqrt{6}$$

(recuérdese que $\sigma(\sqrt{6}) = -\sqrt{6}$). Imponer $\sigma(z) = z$ y usar la unicidad de coordenadas en la base $\{1, \sqrt{2}, \sqrt{3}, \sqrt{6}\}$ da $b = -b$ y $d = -d$, es decir $b = d = 0$. Luego

$$L^H = \{a + c\sqrt{3} : \ a, c \in \mathbb{Q}\} = \mathbb{Q}(\sqrt{3}),$$

de grado 2 sobre $\mathbb{Q}$. Nótese el patrón: $|H| = 2$ y $[L : L^H] = 2$, mientras que $[L^H : \mathbb{Q}] = 2$ es el índice de H en el grupo total. Esa coincidencia no es casual y es, literalmente, el enunciado de la correspondencia de Galois, que se demuestra en la sección correspondiente de este capítulo.

26.2.8. El vocabulario que queda, en una página

Los términos siguientes se usan a partir de aquí sin más aviso. Se recogen juntos para que el lector tenga todas las definiciones a mano antes de entrar en materia; cada uno se desarrolla, con ejemplos y demostraciones, en la sección que se indica.

Definición 26.47 (Glosario operativo del capítulo)**.** Sea L/K una extensión finita.

- *Extensión normal*: L es cuerpo de descomposición de algún polinomio de $K[x]$; equivalentemente, todo irreducible de $K[x]$ con una raíz en L tiene todas sus raíces en L. Ejemplo de lo contrario: $\mathbb{Q}(\sqrt[3]{2})/\mathbb{Q}$ no es normal.
- *Extensión separable*: el polinomio mínimo de cada elemento de L no tiene raíces repetidas. En característica 0 esto es automático, así que sobre $\mathbb{Q}$ nunca estorba.
- *Extensión de Galois*: normal y separable; equivale a $|\operatorname{Gal}(L/K)| = [L:K]$, y equivale a $L^{\operatorname{Gal}(L/K)} = K$. Es la hipótesis bajo la cual la correspondencia funciona.
- *Raíz primitiva n-ésima de la unidad*: un ζ con $\zeta^n = 1$ y $\zeta^m \neq 1$ para $0 < m < n$; sobre $\mathbb{C}$, $\zeta_n = \mathrm{e}^{2\pi \mathrm{i}/n}$.
- *Extensión ciclotómica*: la de la forma $K(\zeta_n)/K$. Sobre $\mathbb{Q}$ tiene grado $\varphi(n)$ y grupo abeliano $(\mathbb{Z}/n)^\times$.

- *Torre radical*: una cadena $K = F_0 \subset F_1 \subset \cdots \subset F_m$ con cada $F_{i+1} = F_i(\sqrt[n_i]{a_i})$ y $a_i \in F_i$; un polinomio es *resoluble por radicales* si su cuerpo de descomposición cabe en una torre así.
- *Número construible*: un real que se obtiene, partiendo de un segmento unidad, mediante un número finito de intersecciones de rectas y circunferencias trazadas con regla sin marcas y compás.

Pausa para pensar

Merece la pena releer la lista anterior después de terminar el capítulo. Ahora mismo son siete etiquetas; al final serán siete piezas de una única máquina, y la sensación de que "normal", "separable" y "de Galois" son tres tecnicismos arbitrarios habrá desaparecido: son exactamente las tres condiciones mínimas para que contar automorfismos equivalga a contar dimensiones.

26.3. El grupo de Galois

Definición 26.48. $\mathrm{Gal}(L/K) = \{\sigma : L \to L \text{ automorfismo de cuerpos con } \sigma|_K = \mathrm{id}\}$, con la composición. Siempre $|\mathrm{Gal}(L/K)| \leq [L : K]$; la extensión es *de Galois* cuando hay igualdad, lo que (en característica 0) equivale a ser **normal**: cuerpo de descomposición de algún polinomio.

Proposición 26.49. *Si $\sigma \in \mathrm{Gal}(L/K)$ y α es raíz de $f \in K[x]$, entonces $\sigma(\alpha)$ también: $f(\sigma(\alpha)) = \sigma(f(\alpha)) = 0$, pues σ respeta las operaciones y fija los coeficientes.*

Idea clave

Esta proposición es el alma de la teoría: **los automorfismos permutan las raíces**. El grupo de Galois de (el cuerpo de descomposición de) un polinomio separable de grado n es un subgrupo de S_n, y es *transitivo* sobre las raíces cuando el polinomio es irreducible. Las "simetrías ocultas" de una ecuación son las permutaciones de sus raíces que respetan *todas* las relaciones algebraicas entre ellas: cuantas más relaciones, más chico el grupo. Un automorfismo queda determinado por dónde manda los generadores, y solo puede mandarlos a raíces de sus polinomios mínimos: contar automorfismos es contar esas elecciones coherentes.

Ejemplo 26.50 (Repertorio de grupos de Galois)**.** *(a)* $\mathrm{Gal}(\mathbb{Q}(\sqrt{2})/\mathbb{Q}) = \{\mathrm{id}, \sigma\} \cong \mathbb{Z}_2$, con $\sigma(a + b\sqrt{2}) = a - b\sqrt{2}$. *(b)* $\mathrm{Gal}(\mathbb{C}/\mathbb{R}) = \{\mathrm{id}, \text{conjugación}\} \cong \mathbb{Z}_2$. *(c)* $E = \mathbb{Q}(\sqrt[3]{2}, \omega)$: σ puede mandar $\sqrt[3]{2}$ a cualquiera de las 3 raíces de $x^3 - 2$ y ω a cualquiera de las 2 raíces de $x^2 + x + 1$: a lo sumo 6, y como $[E : \mathbb{Q}] = 6$ es de Galois, exactamente 6: $\mathrm{Gal}(E/\mathbb{Q}) \cong S_3$, *todas* las permutaciones de las tres raíces. No abeliano. *(d) Ciclotómico*: $\mathbb{Q}(\zeta_n)$ con $\zeta_n = e^{2\pi i/n}$: cada automorfismo es $\zeta_n \mapsto \zeta_n^k$ con $\gcd(k, n) = 1$, y $\mathrm{Gal}(\mathbb{Q}(\zeta_n)/\mathbb{Q}) \cong (\mathbb{Z}_n)^\times$, *abeliano* de orden $\varphi(n)$. *(e) Cuerpos finitos*: $\mathrm{Gal}(\mathbb{F}_{p^n}/\mathbb{F}_p)$ es cíclico de orden n, generado por el *Frobenius* $x \mapsto x^p$.

Regla: calcular el grupo de Galois de un polinomio concreto

Sea $f \in \mathbb{Q}[x]$ de grado n. El grupo $\mathrm{Gal}(f)$ es, por definición, el grupo de Galois de su cuerpo de descomposición sobre $\mathbb{Q}$. Estos son los pasos que hay que dar, en este orden.

1. **Factoriza f primero.** Todo lo que sigue supone que f es irreducible. Si f se parte, el cuerpo de descomposición es el compuesto de los cuerpos de descomposición de los factores, y el grupo es un subgrupo del producto de los grupos de cada trozo. Nunca apliques los criterios de abajo a un polinomio reducible: darán basura.

2. **Construye el cuerpo de descomposición L explícitamente.** No basta con adjuntar una raíz: hay que adjuntarlas *todas*. La pregunta operativa es siempre la misma: adjuntada una raíz α, ¿están ya las demás en $\mathbb{Q}(\alpha)$, o falta algo? Si falta, adjúntalo y repite.

3. **Calcula $[L:\mathbb{Q}]$ con la torre.** Ese número *es* el orden del grupo: $|\mathrm{Gal}(L/\mathbb{Q})| = [L:\mathbb{Q}]$. Este es el paso que hace todo el trabajo pesado; identificar el grupo después suele ser rutina.

4. **Sitúa el grupo dentro de S_n.** Todo automorfismo permuta las n raíces de f y queda determinado por esa permutación (porque las raíces generan L). Así $\mathrm{Gal}(f)$ se ve como un subgrupo de S_n, y como f es irreducible, es un subgrupo *transitivo*: puede llevar cualquier raíz a cualquier otra. Esto reduce muchísimo las posibilidades. Para $n=3$ los subgrupos transitivos de S_3 son solo dos: $A_3 \cong \mathbb{Z}/3$ y S_3. Para $n=4$ son cinco: $\mathbb{Z}/4$, V, D_4, A_4 y S_4.

5. **Usa el discriminante para decidir entre A_n y S_n.** El discriminante $\Delta = \prod_{i<j}(\alpha_i - \alpha_j)^2$ es un número racional, y $\sqrt{\Delta}$ cambia de signo exactamente con las permutaciones impares. Por tanto:

$$\mathrm{Gal}(f) \subseteq A_n \iff \Delta \text{ es un cuadrado en } \mathbb{Q}.$$

   Para el cúbico reducido $x^3 + px + q$ la fórmula es $\Delta = -4p^3 - 27q^2$.

6. **Cierra el caso con el orden.** Combinando el paso 3 (que da $|G|$) con el paso 4 (que da la lista de candidatos) casi siempre queda un único grupo posible. Si quedan dos con el mismo orden, distínguelos buscando un elemento del orden adecuado o mirando si el grupo es abeliano.

7. **Control final.** $|G| = [L:\mathbb{Q}]$ debe ser divisible por n (por transitividad) y debe dividir a $n!$. Si tu respuesta falla cualquiera de las dos, hay un error antes.

Ejemplo 26.51 (Los dos cúbicos que hay que conocer)**.** Comparamos $f = x^3 - 2$ y $g = x^3 - 3x - 1$. Ambos son irreducibles sobre $\mathbb{Q}$: son de grado 3 y no tienen raíces racionales (para f, ningún $\pm 1, \pm 2$ sirve; para g, ± 1 dan -3 y 1). Salen grupos distintos.

El caso $f = x^3 - 2$. Sus raíces son $\sqrt[3]{2}$, $\omega\sqrt[3]{2}$ y $\omega^2\sqrt[3]{2}$ con $\omega = \mathrm{e}^{2\pi i/3}$. La primera es real y las otras dos no, de modo que $\mathbb{Q}(\sqrt[3]{2}) \subset \mathbb{R}$ no puede contenerlas: el cuerpo de descomposición

es $L = \mathbb{Q}(\sqrt[3]{2}, \omega)$, estrictamente mayor. La torre da

$$[L : \mathbb{Q}] = [L : \mathbb{Q}(\sqrt[3]{2})] \cdot [\mathbb{Q}(\sqrt[3]{2}) : \mathbb{Q}] = 2 \cdot 3 = 6,$$

donde el 2 es el grado del polinomio mínimo $x^2 + x + 1$ de ω, que sigue siendo irreducible sobre $\mathbb{Q}(\sqrt[3]{2})$ por ser este real. Luego $|G| = 6$, y el único subgrupo transitivo de S_3 de orden 6 es S_3 entero. El discriminante confirma: $\Delta = -4 \cdot 0^3 - 27 \cdot (-2)^2 = -108$, que ni siquiera es positivo, así que desde luego no es un cuadrado racional. Conclusión: $\mathrm{Gal}(x^3 - 2) \cong S_3$.

El caso $g = x^3 - 3x - 1$**.** Aquí $p = -3$, $q = -1$ y

$$\Delta = -4(-3)^3 - 27(-1)^2 = 108 - 27 = 81 = 9^2.$$

Es un cuadrado, luego el grupo está contenido en $A_3 \cong \mathbb{Z}/3$, y por transitividad no puede ser trivial: $\mathrm{Gal}(g) \cong \mathbb{Z}/3$, de orden 3. Eso obliga a que $[L : \mathbb{Q}] = 3$, es decir, *adjuntar una sola raíz ya produce las otras dos.* Y así es, de forma completamente explícita: si α es raíz, también lo es $2 - \alpha^2$. Numéricamente las tres raíces son

$$\alpha_1 = 1{,}879385\ldots, \qquad \alpha_2 = -1{,}532089\ldots, \qquad \alpha_3 = -0{,}347296\ldots$$

y el lector puede comprobar que $2 - \alpha_1^2 = \alpha_2$, $2 - \alpha_2^2 = \alpha_3$ y $2 - \alpha_3^2 = \alpha_1$: un 3-ciclo perfecto, que es literalmente el generador del grupo escrito como fórmula. El origen de esa magia es trigonométrico: las raíces son $2\cos\theta$ con $\theta = 20°, 140°, 260°$, y de la identidad del ángulo doble $\cos 2\theta = 2\cos^2\theta - 1$ sale

$$2 - (2\cos\theta)^2 = -2\cos 2\theta = 2\cos(180° - 2\theta).$$

Es decir, la aplicación $\alpha \mapsto 2 - \alpha^2$ actúa sobre el ángulo como $\theta \mapsto 180° - 2\theta$, y esa transformación manda $20 \mapsto 140 \mapsto -100 \equiv 260 \mapsto -340 \equiv 20$ módulo 360°: el 3-ciclo.

Moraleja: dos cúbicos de aspecto casi idéntico, uno con grupo de orden 6 y otro de orden 3. Lo que los separa es un solo número, el discriminante.

Ejemplo 26.52 (Un cuártico con grupo diédrico: $x^4 - 2$)**.** Sea $f = x^4 - 2$, irreducible por Eisenstein con $p = 2$. Sus raíces son

$$\sqrt[4]{2}, \quad i\sqrt[4]{2}, \quad -\sqrt[4]{2}, \quad -i\sqrt[4]{2},$$

es decir, $i^k \sqrt[4]{2}$ para $k = 0, 1, 2, 3$. El cuerpo de descomposición es entonces $L = \mathbb{Q}(\sqrt[4]{2},\, i)$: hace falta $\sqrt[4]{2}$ para tener una raíz y hace falta i para pasar de una a otra.

El grado. $[\mathbb{Q}(\sqrt[4]{2}) : \mathbb{Q}] = 4$ por Eisenstein, y ese cuerpo es real, luego $x^2 + 1$ sigue siendo irreducible sobre él y $[L : \mathbb{Q}(\sqrt[4]{2})] = 2$. Por la torre,

$$[L : \mathbb{Q}] = 2 \cdot 4 = 8.$$

Así que $|G| = 8$. (Comprobación independiente: el polinomio mínimo de $\sqrt[4]{2} + i$ sobre $\mathbb{Q}$ es $x^8 + 4x^6 + 2x^4 + 28x^2 + 1$, de grado 8; el elemento primitivo confirma el cálculo.)

El grupo. De los cinco subgrupos transitivos de S_4, solo uno tiene orden 8: el diédrico D_4. Ya está identificado, pero vale la pena verlo en acción. Un automorfismo queda fijado

por dónde manda $\sqrt[4]{2}$ (cuatro opciones) y dónde manda i (dos opciones): $4 \cdot 2 = 8$, coherente. Llamemos

$$\sigma : \ \sqrt[4]{2} \mapsto i\sqrt[4]{2}, \ i \mapsto i, \qquad \tau : \ \sqrt[4]{2} \mapsto \sqrt[4]{2}, \ i \mapsto -i$$

(τ es la conjugación compleja). Entonces σ tiene orden 4, τ tiene orden 2, y se comprueba que $\tau\sigma\tau^{-1} = \sigma^{-1}$: exactamente la relación que define D_4, el grupo de simetrías del cuadrado. Y la imagen geométrica es perfecta, porque las cuatro raíces *son* los vértices de un cuadrado en el plano complejo, inscrito en la circunferencia de radio $\sqrt[4]{2}$: σ lo gira un cuarto de vuelta y τ lo refleja en el eje real.

Control. $|G| = 8$ es divisible por $n = 4$ y divide a $4! = 24$. Ambas condiciones se cumplen.

Regla: calcular el grupo de Galois de un polinomio concreto

Casi siempre el grupo se determina sin escribir un solo automorfismo: basta con contar grados y mirar el discriminante.

1. **Comprueba que f es separable** (en característica 0 basta con que sea irreducible, o más generalmente con $\gcd(f, f') = 1$). Si no lo es, factoriza primero.
2. **Construye mentalmente el cuerpo de descomposición L**: adjunta las raíces una a una y lleva la cuenta de los grados con la torre.
3. **Calcula $|G| = [L : \mathbb{Q}]$.** Éste es el dato central: el grupo tiene exactamente ese orden, ni uno más.
4. **Encájalo en S_n.** El grupo actúa fielmente permutando las n raíces, y si f es irreducible la acción es transitiva, lo que descarta de golpe casi todos los subgrupos.
5. **Mira el discriminante.** Si Δ es un cuadrado en $\mathbb{Q}$, entonces $G \subseteq A_n$; si no lo es, $G \not\subseteq A_n$ y $\mathbb{Q}(\sqrt{\Delta})$ es un subcuerpo cuadrático concreto de L.
6. **Cruza los datos.** Para $n = 3$: transitivo y de orden 3 o 6, luego $G = A_3 \cong \mathbb{Z}/3$ si Δ es cuadrado y $G = S_3$ si no. Para $n = 4$ hay cinco posibilidades ($\mathbb{Z}/4$, V, D_4, A_4, S_4) y se separan con el orden y con la resolvente cúbica.
7. **Control:** el número de cuerpos intermedios debe coincidir con el número de subgrupos de G. Si tu grupo tiene tres subgrupos propios y tú ves cuatro subcuerpos, revisa.

Ejemplo 26.53 (Cinco grupos de Galois sobre $\mathbb{Q}$)**. (a)** $f = x^2 - 2$**.** El cuerpo de descomposición es $\mathbb{Q}(\sqrt{2})$, de grado 2. Luego $G \cong \mathbb{Z}/2$, generado por $\sqrt{2} \mapsto -\sqrt{2}$. Aquí $\Delta = 8$, que no es cuadrado, coherente con $G \not\subseteq A_2 = \{e\}$.

(b) $f = x^3 - 2$**.** Las raíces son $\sqrt[3]{2}$, $\omega\sqrt[3]{2}$, $\omega^2\sqrt[3]{2}$ con $\omega = e^{2\pi i/3}$. El cuerpo de descomposición es $L = \mathbb{Q}(\sqrt[3]{2}, \omega)$ y la torre da

$$[L : \mathbb{Q}] = [L : \mathbb{Q}(\sqrt[3]{2})] \cdot [\mathbb{Q}(\sqrt[3]{2}) : \mathbb{Q}] = 2 \cdot 3 = 6.$$

Un grupo transitivo de orden 6 dentro de S_3 solo puede ser S_3 entero. El discriminante confirma: $\Delta = -108$, que no es cuadrado en $\mathbb{Q}$, y en efecto $\sqrt{-108} = 6\sqrt{-3}$ genera el subcuerpo $\mathbb{Q}(\omega)$.

(c) $f = x^3 - 3x - 1$**.** También irreducible (no tiene raíces racionales: ± 1 dan -3 y 1). Pero ahora

$$\Delta = -4(-3)^3 - 27(-1)^2 = 108 - 27 = 81 = 9^2,$$

¡un cuadrado! Luego $G \subseteq A_3$, y siendo transitivo, $G = A_3 \cong \mathbb{Z}/3$. El cuerpo de descomposición tiene grado 3, no 6: las tres raíces reales $2\cos(\pi/9) \approx 1{,}8794$, $2\cos(5\pi/9) \approx -0{,}3473$ y $2\cos(7\pi/9) \approx -1{,}5321$ se expresan unas en función de otras mediante el generador $\beta \mapsto 2 - \beta^2$, que las permuta cíclicamente (compruébalo: $2 - 1{,}8794^2 \approx -1{,}5321$).

(d) $f = x^4 - 2$**.** Las raíces son $\pm\sqrt[4]{2}$, $\pm i\sqrt[4]{2}$, así que $L = \mathbb{Q}(\sqrt[4]{2}, i)$ y

$$[L : \mathbb{Q}] = 2 \cdot 4 = 8.$$

Un subgrupo transitivo de orden 8 de S_4 es un 2-Sylow, es decir D_4. Se ve directamente: $\sigma : \sqrt[4]{2} \mapsto i\sqrt[4]{2}$ tiene orden 4, la conjugación τ tiene orden 2, y $\tau\sigma\tau = \sigma^{-1}$: la relación exacta del diédrico.

(e) $f = x^4 + 1$**.** Sus raíces son las ocho raíces primitivas de la unidad de orden 8, así que $L = \mathbb{Q}(\zeta_8)$ y $[L : \mathbb{Q}] = \varphi(8) = 4$. El grupo es $(\mathbb{Z}/8)^\times = \{1, 3, 5, 7\}$, en el que todo elemento eleva al cuadrado a 1: es el grupo de Klein $\mathbb{Z}/2 \times \mathbb{Z}/2$, no $\mathbb{Z}/4$. Los tres subcuerpos cuadráticos correspondientes son $\mathbb{Q}(i)$, $\mathbb{Q}(\sqrt{2})$ y $\mathbb{Q}(\sqrt{-2})$, uno por cada subgrupo de orden 2.

Idea clave

Compara (b) y (c): dos cúbicas irreducibles con el mismo aspecto y grupos distintos, S_3 frente a $\mathbb{Z}/3$. Lo único que las separa es si Δ es o no un cuadrado. Esa es, en miniatura, toda la potencia del método: una cuenta puramente aritmética con los coeficientes decide una propiedad de simetría del cuerpo de descomposición.

Cuidado

El orden del grupo es $[L : \mathbb{Q}]$ con L el cuerpo de descomposición *completo*, no $[\mathbb{Q}(\alpha) : \mathbb{Q}]$ con una sola raíz. En (b), $\mathbb{Q}(\sqrt[3]{2})$ tiene grado 3 pero su grupo de automorfismos sobre $\mathbb{Q}$ es trivial, porque los otros dos conjugados no son reales y no caben dentro. Confundir ambos grados es el error clásico del tema.

26.4. La correspondencia de Galois

Idea clave

Intuición primero. La idea de Galois: para entender un objeto, estudia sus *simetrías*. Las raíces de un polinomio se pueden permutar entre sí sin que ninguna ecuación racional lo note; esas permutaciones indistinguibles forman el grupo de Galois. La correspondencia es un diccionario perfecto y *al revés*: cuanto más grande el subcuerpo (más cantidades conoces), más pequeño el subgrupo (menos permutaciones pasan desapercibidas). Conocer es romper simetría: resolver una ecuación por pasos es demoler el grupo piso por piso, y por eso la estructura interna del grupo (¿tiene pisos abelianos hasta abajo?) decide si la ecuación es resoluble por radicales. El álgebra de cuerpos, traducida a teoría finita de grupos: ese es el milagro operativo.

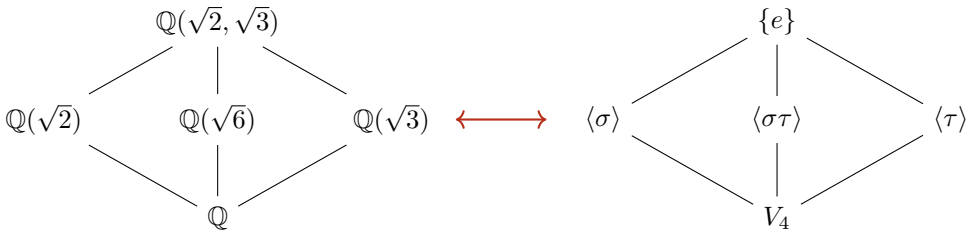


Figura 26.1: La correspondencia de Galois para $\mathbb{Q}(\sqrt{2},\sqrt{3})$: el retículo de subcuerpos (izquierda) y el de subgrupos de $V_4 = \{e, \sigma, \tau, \sigma\tau\}$ (derecha) son imágenes en espejo, con las inclusiones invertidas. Aquí σ manda $\sqrt{2} \mapsto -\sqrt{2}$ y τ manda $\sqrt{3} \mapsto -\sqrt{3}$: cada subcuerpo es lo que su subgrupo no puede mover.

Teorema 26.54 (Teorema fundamental de la teoría de Galois)**.** *Sea L/K de Galois finita con grupo $G = \mathrm{Gal}(L/K)$. Las aplicaciones*

$$F \longmapsto \mathrm{Gal}(L/F), \qquad H \longmapsto L^H = \{x \in L : \sigma(x) = x \ \forall \sigma \in H\}$$

son biyecciones mutuamente inversas entre cuerpos intermedios $K \subseteq F \subseteq L$ y subgrupos $H \leq G$, y ***invierten inclusiones****. Además:*

(i) $[L : F] = |\mathrm{Gal}(L/F)|$ *y* $[F : K] = [G : \mathrm{Gal}(L/F)]$ *(grados $\leftrightarrow$ índices);*

(ii) *F/K es normal (de Galois)* $\iff$ $\mathrm{Gal}(L/F) \trianglelefteq G$*, y entonces* $\mathrm{Gal}(F/K) \cong G/\mathrm{Gal}(L/F)$.

Idea clave

El teorema convierte preguntas sobre cuerpos (infinitos, difíciles de enumerar) en preguntas sobre un grupo finito (finito, computable). "¿Cuántos cuerpos intermedios hay?" = "¿cuántos subgrupos tiene G?". "¿Cuáles son normales sobre K?" = "¿qué subgrupos son normales?". Es uno de los diccionarios más perfectos de la matemática, y el modelo de todas las "correspondencias de Galois" posteriores (revestimientos y grupo fundamental en topología, por ejemplo).

Ejemplo 26.55 (El retículo completo de $x^3 - 2$)**.** $E = \mathbb{Q}(\sqrt[3]{2}, \omega)$, $G \cong S_3$ con subgrupos $\{e\}$, A_3, tres de orden 2, y S_3 (ejercicios del capítulo 25). El diccionario, invirtiendo inclusiones:

subgrupo H	orden	cuerpo fijo E^H	grado sobre $\mathbb{Q}$
$\{e\}$	1	E	6
$\langle(2\,3)\rangle$	2	$\mathbb{Q}(\sqrt[3]{2})$	3
$\langle(1\,3)\rangle$	2	$\mathbb{Q}(\omega\sqrt[3]{2})$	3
$\langle(1\,2)\rangle$	2	$\mathbb{Q}(\omega^2\sqrt[3]{2})$	3
A_3	3	$\mathbb{Q}(\omega)$	2
S_3	6	$\mathbb{Q}$	1

(numerando las raíces $1 = \sqrt[3]{2}$, $2 = \omega\sqrt[3]{2}$, $3 = \omega^2\sqrt[3]{2}$; el estabilizador de la raíz 1 fija $\mathbb{Q}(\sqrt[3]{2})$, etc.). Solo A_3 es normal, y su cuerpo $\mathbb{Q}(\omega)$ es el único intermedio normal sobre $\mathbb{Q}$, con

$\mathrm{Gal}(\mathbb{Q}(\omega)/\mathbb{Q}) \cong S_3/A_3 \cong \mathbb{Z}_2$. No existe ningún otro cuerpo intermedio: el diccionario es exhaustivo.

Regla: usar la correspondencia de Galois en un caso concreto

La correspondencia convierte una pregunta sobre cuerpos (difícil, infinitos objetos) en una pregunta sobre subgrupos de un grupo finito (fácil, se listan todos). El procedimiento es siempre el mismo.

1. **Verifica que L/K es de Galois:** normal y separable, o equivalentemente $|\mathrm{Aut}(L/K)| = [L : K]$, o también que L es cuerpo de descomposición de un polinomio separable. Sin esto la correspondencia falla.
2. **Lista *todos* los subgrupos de $G = \mathrm{Gal}(L/K)$.** Son finitos y pocos; para grupos de orden ≤ 12 se hacen a mano usando Lagrange (el orden de un subgrupo divide al del grupo).
3. **Invierte el orden.** A subgrupo más grande le corresponde cuerpo fijo más pequeño: $H_1 \subseteq H_2 \iff L^{H_2} \subseteq L^{H_1}$. Los extremos son $G \leftrightarrow K$ y $\{e\} \leftrightarrow L$.
4. **Calcula grados con índices:** $[L : L^H] = |H|$ y $[L^H : K] = [G : H]$. Esto identifica cada subcuerpo antes de escribirlo.
5. **Encuentra cada cuerpo fijo explícitamente.** Toma un elemento genérico de L en la base sobre K, impón que sea fijo por los generadores de H y resuelve el sistema lineal.
6. **Detecta las extensiones normales:** L^H/K es normal (de Galois) exactamente cuando $H \trianglelefteq G$, y en ese caso $\mathrm{Gal}(L^H/K) \cong G/H$.
7. **Control:** cuenta. El número de subcuerpos intermedios es *exactamente* el número de subgrupos. Ni uno más, ni uno menos.

Ejemplo 26.56 (El retículo completo de $\mathbb{Q}(\sqrt{2},\sqrt{3})$)**.** Sea $L = \mathbb{Q}(\sqrt{2},\sqrt{3})$. Es el cuerpo de descomposición de $(x^2-2)(x^2-3)$, luego $L/\mathbb{Q}$ es de Galois, y ya sabemos que $[L : \mathbb{Q}] = 4$.

Los automorfismos quedan determinados por lo que hacen con $\sqrt{2}$ y $\sqrt{3}$, y cada uno puede cambiar de signo independientemente:

$$e, \qquad \sigma : \begin{cases} \sqrt{2} \mapsto -\sqrt{2} \\ \sqrt{3} \mapsto \sqrt{3} \end{cases} \qquad \tau : \begin{cases} \sqrt{2} \mapsto \sqrt{2} \\ \sqrt{3} \mapsto -\sqrt{3} \end{cases} \qquad \sigma\tau : \begin{cases} \sqrt{2} \mapsto -\sqrt{2} \\ \sqrt{3} \mapsto -\sqrt{3}. \end{cases}$$

Los cuatro cumplen $g^2 = e$, así que $G \cong \mathbb{Z}/2 \times \mathbb{Z}/2$, el grupo de Klein.

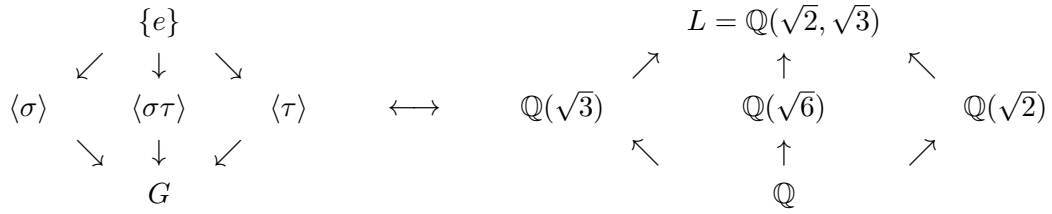
Subgrupos. Hay exactamente cinco: el trivial, el total, y tres de orden 2, generados por σ, por τ y por $\sigma\tau$.

Cuerpos fijos. Un elemento general de L es $a + b\sqrt{2} + c\sqrt{3} + d\sqrt{6}$ con $a, b, c, d \in \mathbb{Q}$.

- Fijo por σ: como $\sigma(\sqrt{6}) = \sigma(\sqrt{2}\sqrt{3}) = -\sqrt{6}$, hace falta $b = d = 0$, luego $L^{\langle\sigma\rangle} = \mathbb{Q}(\sqrt{3})$.
- Fijo por τ: análogamente $c = d = 0$ y $L^{\langle\tau\rangle} = \mathbb{Q}(\sqrt{2})$.

- Fijo por $\sigma\tau$: ahora $\sigma\tau(\sqrt{6}) = (-\sqrt{2})(-\sqrt{3}) = \sqrt{6}$, así que $\sqrt{6}$ *sí* sobrevive; hacen falta $b = c = 0$ y $L^{\langle\sigma\tau\rangle} = \mathbb{Q}(\sqrt{6})$.

El retículo, con el orden invertido:



Como G es abeliano, *todos* sus subgrupos son normales, y en efecto los tres subcuerpos cuadráticos son extensiones de Galois de $\mathbb{Q}$, con grupo $G/H \cong \mathbb{Z}/2$ en cada caso.

Control del paso 7: cinco subgrupos, cinco subcuerpos (contando $\mathbb{Q}$ y L). En particular, $\mathbb{Q}(\sqrt{2}, \sqrt{3})$ *no* contiene ningún otro subcuerpo cuadrático: por ejemplo $\sqrt{5} \notin L$, porque si estuviera habría un cuarto subgrupo de orden 2, y no lo hay.

Ejemplo 26.57 (Un retículo no abeliano: $x^3 - 2$)**.** Sea $L = \mathbb{Q}(\sqrt[3]{2}, \omega)$, de grado 6 sobre $\mathbb{Q}$, con $G \cong S_3$ actuando sobre las raíces $\alpha_1 = \sqrt[3]{2}$, $\alpha_2 = \omega\sqrt[3]{2}$, $\alpha_3 = \omega^2\sqrt[3]{2}$.

S_3 tiene seis subgrupos: $\{e\}$, tres de orden 2 (generados por las transposiciones), uno de orden 3 (A_3) y el total. Luego hay seis subcuerpos:

subgrupo	orden	cuerpo fijo	grado sobre $\mathbb{Q}$
$\{e\}$	1	L	6
$\langle(23)\rangle$	2	$\mathbb{Q}(\sqrt[3]{2})$	3
$\langle(13)\rangle$	2	$\mathbb{Q}(\omega\sqrt[3]{2})$	3
$\langle(12)\rangle$	2	$\mathbb{Q}(\omega^2\sqrt[3]{2})$	3
A_3	3	$\mathbb{Q}(\omega)$	2
S_3	6	$\mathbb{Q}$	1

La transposición (23) es la que deja quieta a α_1 y permuta las otras dos: es la conjugación compleja, y su cuerpo fijo es la parte real $\mathbb{Q}(\sqrt[3]{2})$.

Aquí sí hay diferencia entre normal y no normal. $A_3 \trianglelefteq S_3$, y en efecto $\mathbb{Q}(\omega)/\mathbb{Q}$ es de Galois con grupo $S_3/A_3 \cong \mathbb{Z}/2$. En cambio $\langle(23)\rangle$ *no* es normal, y en efecto $\mathbb{Q}(\sqrt[3]{2})/\mathbb{Q}$ no es normal: contiene una raíz de $x^3 - 2$ pero no las otras dos.

Cuidado

La expresión "$\mathbb{Q}(\sqrt[3]{2})$ tiene grado 3, luego su grupo tiene 3 elementos" es falsa. Esa extensión no es de Galois y su grupo de automorfismos sobre $\mathbb{Q}$ es trivial: cualquier automorfismo manda $\sqrt[3]{2}$ a una raíz de $x^3 - 2$, y las otras dos no son reales, así que no caben dentro. La correspondencia solo funciona con la extensión de Galois completa arriba del todo.

Pausa para pensar

Que los tres subgrupos de orden 2 sean conjugados entre sí en S_3 se traduce en que los tres cuerpos $\mathbb{Q}(\alpha_i)$ son isomorfos entre sí, pero distintos como subconjuntos de $\mathbb{C}$: solo el primero es real. Conjugación de subgrupos $\leftrightarrow$ isomorfía de subcuerpos: la correspondencia respeta también esa capa de información.

Regla: calcular el grupo de Galois y los subcuerpos de una extensión ciclotómica

Las extensiones ciclotómicas $\mathbb{Q}(\zeta_n)$ son el laboratorio ideal de la teoría: su grupo de Galois es abeliano, se conoce de antemano, y todo el retículo de subcuerpos se lee sin resolver una sola ecuación. Conviene dominarlas antes de atacar casos generales.

1. **Fija la raíz y el grado.** Sea $\zeta_n = e^{2\pi i/n}$ una raíz primitiva n-ésima de la unidad. El polinomio mínimo de ζ_n sobre $\mathbb{Q}$ es el polinomio ciclotómico Φ_n, que es irreducible sobre $\mathbb{Q}$ (teorema de Gauss), y por tanto
$$[\mathbb{Q}(\zeta_n) : \mathbb{Q}] = \deg \Phi_n = \varphi(n).$$

2. **Identifica el grupo.** $\mathbb{Q}(\zeta_n)$ es el cuerpo de descomposición de $x^n - 1$, luego es normal, y en característica 0 todo es separable: la extensión es de Galois. Cada $\sigma \in \mathrm{Gal}(\mathbb{Q}(\zeta_n)/\mathbb{Q})$ manda $\zeta_n \mapsto \zeta_n^k$ con $\gcd(k, n) = 1$, y componer dos de ellos multiplica los exponentes. Por tanto
$$\mathrm{Gal}(\mathbb{Q}(\zeta_n)/\mathbb{Q}) \;\cong\; (\mathbb{Z}/n)^\times, \qquad \sigma_k \longleftrightarrow k.$$
Es *siempre* abeliano. Nunca busques un grupo no conmutativo aquí.

3. **Determina la estructura de $(\mathbb{Z}/n)^\times$.** Descompón $n = \prod p_i^{k_i}$ y usa el teorema chino: $(\mathbb{Z}/n)^\times \cong \prod (\mathbb{Z}/p_i^{k_i})^\times$. Cada factor con p impar es cíclico de orden $p^{k-1}(p-1)$; para $p = 2$ se tiene $(\mathbb{Z}/2)^\times = 1$, $(\mathbb{Z}/4)^\times \cong \mathbb{Z}/2$ y $(\mathbb{Z}/2^k)^\times \cong \mathbb{Z}/2 \times \mathbb{Z}/2^{k-2}$ si $k \geq 3$.

4. **Lista los subgrupos.** Como el grupo es abeliano, *todos* sus subgrupos son normales, así que *todos* los subcuerpos intermedios son de Galois sobre $\mathbb{Q}$ y el número de subcuerpos es exactamente el número de subgrupos. Para $(\mathbb{Z}/n)^\times$ cíclico de orden m hay un subgrupo por cada divisor de m: ni uno más, ni uno menos.

5. **Traduce cada subgrupo en un subcuerpo con períodos de Gauss.** Si $H \leq (\mathbb{Z}/n)^\times$ tiene índice d, su cuerpo fijo tiene grado d sobre $\mathbb{Q}$ y está generado por el *período*
$$\eta = \sum_{k \in H} \zeta_n^k,$$
junto con sus $d - 1$ conjugados. El polinomio mínimo de η se obtiene multiplicando $(x - \eta_j)$ sobre los conjugados; sale con coeficientes racionales precisamente porque el grupo permuta los períodos.

6. **Reconoce el subcuerpo cuadrático.** Si p es primo impar, $(\mathbb{Z}/p)^\times$ es cíclico de orden $p-1$ y tiene un único subgrupo de índice 2 (los residuos cuadráticos), luego $\mathbb{Q}(\zeta_p)$ contiene un único subcuerpo cuadrático, y es
$$\mathbb{Q}(\sqrt{p^*}), \qquad p^* = (-1)^{(p-1)/2}p,$$
es decir $\mathbb{Q}(\sqrt{p}\,)$ si $p \equiv 1 \pmod 4$ y $\mathbb{Q}(\sqrt{-p}\,)$ si $p \equiv 3 \pmod 4$.
7. **Contrasta con la construcción con regla y compás.** El polígono regular de n lados es construible exactamente cuando $[\mathbb{Q}(\zeta_n):\mathbb{Q}] = \varphi(n)$ es una potencia de 2, porque entonces el grupo es un 2-grupo y admite una torre de subgrupos de índice 2, que es justo una torre de extensiones cuadráticas. Eso equivale a $n = 2^a p_1 \cdots p_s$ con primos de Fermat distintos.

Ejemplo 26.58 (Cuatro ciclotómicas desmontadas pieza a pieza)**. (a) $\mathbb{Q}(\zeta_7)$: cíclica de orden 6.** Aquí $\varphi(7) = 6$ y $\Phi_7(x) = x^6+x^5+x^4+x^3+x^2+x+1$, de grado 6. El grupo $(\mathbb{Z}/7)^\times = \{1,2,3,4,5,6\}$ es cíclico generado por 3: las potencias $3,2,6,4,5,1$ recorren todo. Luego Gal $\cong \mathbb{Z}/6$ y hay exactamente cuatro subgrupos ($1,2,3,6$ como órdenes), es decir cuatro subcuerpos intermedios contando los extremos:
$$\mathbb{Q} \;\subset\; \mathbb{Q}(\sqrt{-7}\,) \;\subset\; \mathbb{Q}(\zeta_7), \qquad \mathbb{Q} \;\subset\; K_3 \;\subset\; \mathbb{Q}(\zeta_7),$$
con $[K_3:\mathbb{Q}] = 3$. El cuadrático es $\mathbb{Q}(\sqrt{-7}\,)$ porque $7 \equiv 3 \pmod 4$, conforme al paso 6.

El cúbico K_3 se obtiene con el período del subgrupo $H = \{1,6\}$ (índice 3): $\eta_0 = \zeta_7 + \zeta_7^6 = 2\cos\frac{2\pi}{7}$, y sus conjugados $\eta_1 = 2\cos\frac{4\pi}{7}$, $\eta_2 = 2\cos\frac{6\pi}{7}$. Numéricamente $\eta_0 = 1.24698$, $\eta_1 = -0.44504$, $\eta_2 = -1.80194$. Su suma es -1 (la suma de todas las raíces de Φ_7), su producto es 1 y la suma de productos dos a dos es -2, de modo que
$$m_{\eta_0}(x) = x^3 + x^2 - 2x - 1,$$
que efectivamente tiene esas tres raíces. Así $K_3 = \mathbb{Q}(\cos\frac{2\pi}{7})$ es un cuerpo cúbico *real y cíclico*, y $\mathbb{Q}(\zeta_7) = K_3(\sqrt{-7}\,)$ salvo detalles de composición.

(b) $\mathbb{Q}(\zeta_5)$ y el pentágono. $\varphi(5) = 4$, potencia de 2: el pentágono regular es construible, como sabían los griegos. El período de $H = \{1,4\}$ es $2\cos\frac{2\pi}{5}$, que satisface $x^2 + x - 1 = 0$ y vale $\frac{\sqrt{5}-1}{2} = 0.618034$: la razón áurea, obtenida sin geometría, solo contando subgrupos. El subcuerpo cuadrático es $\mathbb{Q}(\sqrt{5}\,)$, coherente con $5 \equiv 1 \pmod 4$.

(c) $\mathbb{Q}(\zeta_{15})$: el primero que no es cíclico. $\varphi(15) = \varphi(3)\varphi(5) = 2\cdot 4 = 8$, y por el teorema chino
$$(\mathbb{Z}/15)^\times \cong (\mathbb{Z}/3)^\times \times (\mathbb{Z}/5)^\times \cong \mathbb{Z}/2 \times \mathbb{Z}/4.$$
Este grupo *no* es cíclico: no hay ningún elemento de orden 8. Comprobación directa sobre $\{1,2,4,7,8,11,13,14\}$: el 2 tiene orden 4 ($2,4,8,16 \equiv 1$), el 7 también ($7,4,13,1$), y los elementos de orden 2 son 4, 11 y 14. Tres involuciones significan tres subgrupos de índice 2, es decir *tres* subcuerpos cuadráticos:
$$\mathbb{Q}(\sqrt{-3}\,) \subset \mathbb{Q}(\zeta_3), \qquad \mathbb{Q}(\sqrt{5}\,) \subset \mathbb{Q}(\zeta_5), \qquad \mathbb{Q}(\sqrt{-15}\,),$$

el tercero siendo el "producto" de los otros dos. En total $\mathbb{Z}/4 \times \mathbb{Z}/2$ tiene 8 subgrupos (uno trivial, tres de orden 2, tres de orden 4 y el total), luego $\mathbb{Q}(\zeta_{15})$ tiene exactamente 8 subcuerpos. Compárese con (a), donde $\mathbb{Z}/6$ solo tenía 4.

(d) $\mathbb{Q}(\zeta_{16})$ y el heptadecágono. $\varphi(16) = 8$ y $(\mathbb{Z}/16)^\times \cong \mathbb{Z}/2 \times \mathbb{Z}/4$, generado por -1 y por 3 (cuyo orden es 4: $3, 9, 11, 1$). De nuevo abeliano no cíclico, pero lo importante es que $8 = 2^3$ es potencia de 2: el polígono de 16 lados es construible, como era de esperar bisecando el cuadrado. En cambio $\varphi(7) = 6$ y $\varphi(9) = 6$ no son potencias de 2, y por eso el heptágono y el eneágono regulares *no* son construibles con regla y compás. El caso célebre es $n = 17$: $\varphi(17) = 16 = 2^4$, y de ahí el heptadecágono de Gauss, cuya construcción no es más que la torre de cuatro subgrupos de índice 2 dentro de $\mathbb{Z}/16$.

Cuidado

El isomorfismo $\mathrm{Gal}(\mathbb{Q}(\zeta_n)/\mathbb{Q}) \cong (\mathbb{Z}/n)^\times$ es con el grupo *multiplicativo* de unidades, no con $\mathbb{Z}/n$ aditivo. Confundirlos lleva a decir que el grupo tiene n elementos en vez de $\varphi(n)$, y a creer que siempre es cíclico. Solo lo es cuando n es 1, 2, 4, p^k o $2p^k$ con p primo impar; el ejemplo (c) con $n = 15$ ya lo rompe.

Idea clave

El teorema de Kronecker–Weber afirma que *toda* extensión abeliana de $\mathbb{Q}$ está contenida en alguna ciclotómica. Dicho de otro modo, los ejemplos de arriba no son una familia entre muchas: son *todas* las extensiones de $\mathbb{Q}$ con grupo de Galois conmutativo. Por eso $\sqrt{2}$ vive dentro de $\mathbb{Q}(\zeta_8)$, $\sqrt{-3}$ dentro de $\mathbb{Q}(\zeta_3)$ y $\sqrt{5}$ dentro de $\mathbb{Q}(\zeta_5)$: no es casualidad, es un teorema. Encontrar el análogo para cuerpos base distintos de $\mathbb{Q}$ es el programa de la teoría de cuerpos de clases, y su versión no abeliana sigue abierta.

26.5. Las fórmulas clásicas: grados 2, 3 y 4

Idea clave

Intuición primero. La pregunta "¿por qué hay fórmula hasta grado 4 y no más?" se responde mejor al revés: mira las fórmulas que sí existen y pregúntate qué tienen dentro. Todas son la misma receta —sumar, restar, multiplicar, dividir y *sacar raíces*— aplicada por etapas, y cada raíz que sacas es un escalón.

Lo notable es que los escalones de la fórmula están dictados, uno a uno, por la cadena de subgrupos del grupo de permutaciones de las raíces. En grado 2 el grupo es $\mathbb{Z}_2$ y hay que sacar una raíz cuadrada: la del discriminante. En grado 3 la cadena $\{e\} \trianglelefteq A_3 \trianglelefteq S_3$ tiene cocientes $\mathbb{Z}_3$ y $\mathbb{Z}_2$, y la fórmula de Cardano saca exactamente una raíz cúbica y una cuadrada, en ese orden. En grado 4 la cadena pasa por el grupo de Klein y por eso Ferrari necesita resolver antes una *cúbica* auxiliar.

Ahí está el enunciado completo del teorema antes de demostrarlo: **hay fórmula por radicales si y solo si el grupo se puede desmontar en pasos abelianos**. Es la traducción literal de "la resolución en escalones". Y en grado 5 esa cadena no existe, porque A_5 es simple y no se deja desmontar. La imposibilidad de la quíntica no es que

nadie haya sido lo bastante listo: es que el grupo no tiene por dónde romperse.

Antes de demostrar que la quíntica no tiene fórmula, veamos *por qué* las de grados menores sí, y cómo el grupo lo explica.

Grado 2: x^2+bx+c: completar el cuadrado da $x=\frac{-b\pm\sqrt{b^2-4c}}{2}$. El grupo es $S_2\cong\mathbb{Z}_2$: una sola raíz cuadrada, la del discriminante.

Grado 3 (Cardano, 1545): la cúbica reducida $x^3+px+q=0$ (a la que se llega con $x\to x-\frac{a}{3}$) se resuelve con la sustitución $x=u+v$ imponiendo $3uv=-p$:

$$u^3+v^3=-q,\qquad u^3v^3=-\frac{p^3}{27},$$

así que u^3, v^3 son raíces de la *resolvente cuadrática* $t^2+qt-\frac{p^3}{27}=0$:

$$x=\sqrt[3]{-\frac{q}{2}+\sqrt{\frac{q^2}{4}+\frac{p^3}{27}}}+\sqrt[3]{-\frac{q}{2}-\sqrt{\frac{q^2}{4}+\frac{p^3}{27}}}.$$

Estructuralmente: S_3 es soluble con cadena $\{e\}\trianglelefteq A_3\trianglelefteq S_3$ de cocientes $\mathbb{Z}_3$ y $\mathbb{Z}_2$; la fórmula toma *una raíz cuadrada* (el paso $S_3\to A_3$, cuyo objeto es el discriminante) y *una raíz cúbica* (el paso $A_3\to\{e\}$). La cadena del grupo *es* el esqueleto de la fórmula. Curiosidad histórica: cuando las tres raíces son reales, la fórmula pasa obligatoriamente por complejos (*casus irreducibilis*): los números complejos entraron a la matemática por la cúbica, no por la cuadrática.

Grado 4 (Ferrari): la resolvente es una *cúbica*, que corresponde al cociente $S_4/V\cong S_3$ (con V el grupo de Klein): se resuelve la cúbica y se rematan dos raíces cuadradas. De nuevo, la cadena soluble $\{e\}\trianglelefteq\mathbb{Z}_2\trianglelefteq V\trianglelefteq A_4\trianglelefteq S_4$ dicta la fórmula.

Grado 5: haría falta una cadena análoga en $S_5\ldots$ que no existe.

Regla: determinar el grupo de Galois de una cuártica con el discriminante y la resolvente cúbica

Para grados 2 y 3 el grupo de Galois se decide con el discriminante y poco más. En grado 4 hay cinco grupos transitivos posibles —$\mathbb{Z}/4$, V_4, D_4, A_4 y S_4— y el procedimiento que sigue los separa a todos con dos objetos calculables: el discriminante y un polinomio cúbico auxiliar.

1. **Reduce y comprueba la irreducibilidad.** Parte de f mónica de grado 4. Si tiene término cúbico, la sustitución $x\mapsto x-\frac{a_3}{4}$ lo elimina y deja $f(x)=x^4+px^2+qx+r$ sin cambiar el grupo. Comprueba que f es irreducible sobre $\mathbb{Q}$ (Eisenstein, reducción módulo un primo, o buscar raíces racionales y factorizaciones en dos cuadráticas). Si f se factoriza, el grupo *no* es transitivo y hay que analizar los factores por separado.

2. **Calcula el discriminante.** Con $\Delta=\prod_{i<j}(\alpha_i-\alpha_j)^2$. Para la forma reducida,

$$\Delta=16p^4r-4p^3q^2-128p^2r^2+144pq^2r-27q^4+256r^3.$$

3. **Decide si Δ es un cuadrado en $\mathbb{Q}$.** Si Δ es cuadrado perfecto, $G\subseteq A_4$, y como

G es transitivo solo quedan V_4 y A_4. Si no lo es, $G \not\subseteq A_4$ y quedan $\mathbb{Z}/4$, D_4 y S_4. Este único bit ya parte el problema en dos.

4. **Escribe la resolvente cúbica.** Es

$$R(y) = y^3 - p\,y^2 - 4r\,y + (4pr - q^2),$$

cuyas raíces son $\alpha_1\alpha_2 + \alpha_3\alpha_4$, $\alpha_1\alpha_3 + \alpha_2\alpha_4$, $\alpha_1\alpha_4 + \alpha_2\alpha_3$. Estas tres cantidades son permutadas por S_4, y el núcleo de esa acción es exactamente V_4: ahí está toda la información.

5. **Cuenta las raíces racionales de R.** Si R es irreducible sobre $\mathbb{Q}$, el grupo es A_4 (caso Δ cuadrado) o S_4 (caso Δ no cuadrado). Si R se descompone completamente en $\mathbb{Q}$, el grupo es V_4. Si R tiene exactamente una raíz racional, el grupo es $\mathbb{Z}/4$ o D_4.

6. **Separa $\mathbb{Z}/4$ de D_4.** Es el único caso ambiguo. El criterio general: $G \cong \mathbb{Z}/4$ si y solo si f se vuelve reducible sobre $\mathbb{Q}(\sqrt{\Delta}\,)$. Para la subfamilia bicuadrática $f = x^4 + px^2 + r$ (con $q = 0$) hay un atajo aritmético: $G \cong V_4$ si r es un cuadrado en $\mathbb{Q}$; $G \cong \mathbb{Z}/4$ si $r(p^2 - 4r)$ es un cuadrado; y $G \cong D_4$ en cualquier otro caso.

7. **Contrasta el orden con el grado del cuerpo de descomposición.** El orden del grupo debe coincidir con $[L : \mathbb{Q}]$, donde L es el cuerpo de descomposición: 4 para $\mathbb{Z}/4$ y V_4, 8 para D_4, 12 para A_4, 24 para S_4. Si has identificado a mano dos generadores y el subcuerpo que fijan, esta cuenta es la comprobación final.

Ejemplo 26.59 (Las cinco cuárticas, una por grupo)**. (a) $f = x^4 + 1$, grupo V_4.** Aquí $p = q = 0$, $r = 1$. El discriminante es $\Delta = 256 \cdot 1 = 256 = 16^2$: cuadrado, luego $G \subseteq A_4$. La resolvente es $R(y) = y^3 - 4y = y(y-2)(y+2)$, que se descompone completamente sobre $\mathbb{Q}$: por el paso 5, $G \cong V_4$. Coincide con lo esperado, pues $x^4 + 1 = \Phi_8$ y $\mathbb{Q}(\zeta_8)$ tiene grupo $(\mathbb{Z}/8)^\times \cong \mathbb{Z}/2 \times \mathbb{Z}/2$. También lo predice el atajo del paso 6: $r = 1$ es un cuadrado.

(b) $f = x^4 - 2$, grupo D_4. Eisenstein en 2 da la irreducibilidad. $\Delta = -2048$, que ni siquiera es positivo, luego no es un cuadrado: $G \not\subseteq A_4$. La resolvente es $R(y) = y^3 + 8y = y(y^2 + 8)$, con *exactamente* una raíz racional ($y = 0$), así que estamos en el caso ambiguo. Atajo del paso 6 con $p = 0$, $r = -2$: $r = -2$ no es cuadrado, y $r(p^2 - 4r) = -2 \cdot 8 = -16$ tampoco. Por tanto $G \cong D_4$, de orden 8. Se confirma directamente: el cuerpo de descomposición es $\mathbb{Q}(\sqrt[4]{2}, i)$, de grado $4 \cdot 2 = 8$ sobre $\mathbb{Q}$.

(c) $f = x^4 - 4x^2 + 2$, grupo $\mathbb{Z}/4$. Eisenstein en 2: irreducible. $\Delta = 2048$, y $\sqrt{2048} = 45.25\ldots$ no es entero, luego no es cuadrado. La resolvente es $R(y) = y^3 + 4y^2 - 8y - 32 = (y+4)(y^2 - 8)$, con la única raíz racional $y = -4$: caso ambiguo otra vez. Atajo con $p = -4$, $r = 2$: $r = 2$ no es cuadrado, pero $r(p^2 - 4r) = 2(16 - 8) = 16 = 4^2$ *sí* lo es. Por tanto $G \cong \mathbb{Z}/4$, cíclico de orden 4. Nótese que (b) y (c) tienen resolventes con el mismo aspecto —una raíz racional— y sin embargo grupos de órdenes 8 y 4: sin el paso 6 no se distinguen.

(d) $f = x^4 + 8x + 12$, grupo A_4. Aquí $p = 0$, $q = 8$, $r = 12$. El discriminante es $\Delta = 331\,776 = 576^2$: cuadrado, luego $G \subseteq A_4$. La resolvente es $R(y) = y^3 - 48y - 64$; sus posibles raíces racionales son divisores de 64, y ninguno de $\pm1, \pm2, \pm4, \pm8, \pm16, \pm32, \pm64$

la anula, de modo que R es irreducible. Por el paso 5, $G \cong A_4$, de orden 12. Este es el ejemplo estándar: los grupos A_4 son raros y hay que buscarlos, porque exigen simultáneamente discriminante cuadrado y resolvente irreducible.

(e) $f = x^4+x+1$, grupo S_4: el caso genérico. $p = 0$, $q = 1$, $r = 1$, y $\Delta = -27+256 = 229$, que es primo y por tanto no es cuadrado. La resolvente es $R(y) = y^3 - 4y - 1$, sin raíces racionales (solo cabrían ± 1, y $R(1) = -4$, $R(-1) = 2$), luego irreducible. Conclusión: $G \cong S_4$, de orden 24, y el cuerpo de descomposición tiene grado 24 sobre $\mathbb{Q}$. Una cuártica "tomada al azar" cae aquí: los otros cuatro grupos corresponden a coincidencias aritméticas.

(f) Un chequeo con un caso conocido. Para $f = x^4 - 10x^2 + 1$, cuyas raíces son $\pm\sqrt{2} \pm \sqrt{3}$, el discriminante es $147\,456 = 384^2$ —cuadrado— y $r = 1$ es un cuadrado, así que el atajo predice V_4. En efecto, el cuerpo de descomposición es $\mathbb{Q}(\sqrt{2}, \sqrt{3}\,)$, de grado 4, con grupo $\mathbb{Z}/2 \times \mathbb{Z}/2$ generado por los dos cambios de signo independientes. Todo encaja.

Cuidado

La resolvente cúbica hay que escribirla *después* de eliminar el término en x^3. Si se aplica la fórmula $R(y) = y^3 - py^2 - 4ry + (4pr - q^2)$ a un polinomio que todavía tiene término cúbico, los coeficientes p, q, r no significan lo que la fórmula supone y el resultado es basura. Del mismo modo, todo el procedimiento presupone que f es *irreducible*: si $f = (x^2 + 1)(x^2 - 2)$, el grupo sigue siendo V_4, pero no por ser una cuártica transitiva, sino por ser el compuesto de dos cuadráticas, y la resolvente no dice nada útil.

Pausa para pensar

¿Por qué la resolvente tiene grado 3 y no otro? Porque S_4 tiene un subgrupo normal V_4 con $S_4/V_4 \cong S_3$, y ese cociente es el que actúa sobre las tres cantidades $\alpha_1\alpha_2 + \alpha_3\alpha_4$ y compañía. La resolvente cúbica *es* el polinomio de ese cociente. Y aquí está el motivo profundo de que la cuártica se resuelva por radicales y la quíntica no: S_4 tiene esa escalera $S_4 \trianglerighteq A_4 \trianglerighteq V_4 \trianglerighteq \mathbb{Z}/2 \trianglerighteq 1$ con cocientes abelianos, mientras que S_5 contiene a A_5, que es simple y no abeliano y no admite ningún escalón. La resolvente cúbica no es un truco de Ferrari: es la escalera hecha polinomio.

26.6. El método de Dedekind: leer el grupo de Galois en las factorizaciones módulo p

Calcular el grupo de Galois de un polinomio concreto puede ser trabajoso: hay que localizar las raíces, decidir qué permutaciones las respetan y descartar las que no. Dedekind observó que hay un atajo asombroso: basta factorizar el polinomio módulo unos cuantos primos y anotar los grados de los factores. Esos grados son, literalmente, los tipos de ciclo de ciertos elementos del grupo de Galois. En quince minutos de aritmética modular, y sin tocar un solo número real, queda determinado el grupo de la mayoría de los polinomios que uno se encuentra.

Teorema 26.60 (Dedekind)**.** *Sea $f \in \mathbb{Z}[x]$ mónico, separable, de grado n, y sea p un primo que no divide al discriminante de f. Si módulo p el polinomio factoriza como producto de*

irreducibles de grados $d_1, d_2, \ldots, d_r$ (necesariamente distintos entre sí en el sentido de que no se repiten factores), entonces el grupo de Galois $G = \mathrm{Gal}(f/\mathbb{Q})$, visto como subgrupo de S_n por su acción en las raíces, contiene un elemento cuyo tipo de ciclo es exactamente $(d_1, d_2, \ldots, d_r)$.

Idea clave

Qué dice esto en cristiano. Calcular un grupo de Galois parece exigir entender la extensión entera. El teorema de Dedekind da un atajo asombroso: se reduce el polinomio módulo un primo, se factoriza (cosa que un ordenador hace en un instante) y los grados de los factores dan directamente la forma cíclica de un elemento del grupo. Probando con unos pocos primos se acumulan elementos hasta reconocer el grupo. Es la técnica que convierte un problema de teoría de cuerpos en un ejercicio de aritmética modular.

Ejemplo 26.61 (Determinar el grupo de Galois con dos primos)**.** Sea $f(x) = x^5 - x - 1$, irreducible sobre $\mathbb{Q}$, con discriminante $2869 = 19 \cdot 151$.

(a) Módulo 2. La factorización es en un factor de grado 2 y otro de grado 3, luego el grupo contiene un elemento de tipo cíclico $(2, 3)$. Su cubo es una transposición.

(b) Módulo 3. El polinomio es irreducible módulo 3, de modo que el grupo contiene un 5-ciclo. Lo mismo ocurre módulo 5, 11 y 13.

(c) La conclusión. Un subgrupo transitivo de S_5 que contenga un 5-ciclo y una transposición es todo S_5. Luego $\mathrm{Gal}(f) = S_5$, el polinomio no es soluble por radicales y sus raíces no se pueden escribir con radicales por más ingenio que se ponga. Dos factorizaciones han bastado.

(d) Los primos prohibidos. Módulo 19 la factorización tiene grados $(1, 3)$, que suman 4 y no 5: el polinomio deja de ser separable porque 19 divide al discriminante. Ésa es exactamente la hipótesis del teorema, y saltársela produce basura.

Idea clave

Reducir módulo p es mirar el polinomio a través de una lente que colapsa los enteros en $\mathbb{F}_p$. El grupo de Galois de la reducción es siempre cíclico, generado por el Frobenius $x \mapsto x^p$, y ese generador permuta las raíces exactamente según los grados de los factores irreducibles: un factor de grado d agrupa d raíces que el Frobenius rota entre sí, es decir un d-ciclo. Lo que dice Dedekind es que esa permutación no es un accidente de la reducción, sino la sombra fiel de un elemento que ya estaba en el grupo de arriba. Cada primo entrega, gratis, una permutación del grupo verdadero. Y si se consiguen suficientes permutaciones de tipos distintos, el grupo queda acorralado.

Regla: determinar un grupo de Galois factorizando módulo primos

El método es de acorralamiento: cada primo aporta una restricción por abajo, y el discriminante una por arriba.

1. **Compruébese que f es irreducible sobre $\mathbb{Q}$.** Si lo es, el grupo actúa transitiva-

mente sobre las n raíces, lo que ya descarta muchísimos subgrupos de S_n. Sirven el criterio de Eisenstein o la propia reducción módulo un primo pequeño.

2. **Calcúlese el discriminante Δ y decídase si es un cuadrado.** Si Δ es un cuadrado perfecto en $\mathbb{Q}$, el grupo está contenido en A_n; si no, contiene permutaciones impares. Este dato solo ya reduce a la mitad las posibilidades.

3. **Elíjanse primos que no dividan a Δ.** Un primo que divide al discriminante da una factorización con factores repetidos y el teorema no se aplica; hay que descartarlo y pasar al siguiente.

4. **Factorícese f módulo cada primo y anótese la lista de grados.** Basta con probar la divisibilidad por los irreducibles de grado bajo sobre $\mathbb{F}_p$; para $n \leq 5$ es cuestión de minutos a mano.

5. **Tradúzcase cada lista en un tipo de ciclo.** Grados $(1,4)$ significan un 4-ciclo; grados $(2,3)$ significan un elemento de orden 6 que es producto de una transposición y un 3-ciclo; grados $(1,1,\dots,1)$ significan la identidad, es decir un primo que se descompone del todo.

6. **Elévense los elementos a potencias para obtener más.** Si el grupo contiene un elemento de tipo $(2,3)$, su cubo es una transposición. Es el truco decisivo: rara vez aparece una transposición directamente, pero se fabrica elevando.

7. **Concluir con el teorema de generación.** Un subgrupo transitivo de S_p (p primo) que contenga una transposición es todo S_p. Más en general, si los tipos de ciclo encontrados no caben en ningún subgrupo propio, el grupo es S_n. Compruébese siempre la conclusión contra el paso 2: si Δ es cuadrado, S_n es imposible.

Ejemplo 26.62 (Cuatro polinomios diagnosticados)**. (a) $f(x) = x^5 - 4x + 2$ tiene grupo S_5.** Es irreducible por Eisenstein en $p = 2$. Su discriminante es -212144, que no es cuadrado, así que el grupo no cabe en A_5. Factorizando:

p	3	5	7	11	13	17	19
tipo	(5)	$(1,2,2)$	$(2,3)$	$(1,4)$	$(1,1,3)$	(5)	$(1,4)$

De $p = 3$ sale un 5-ciclo, luego el grupo es transitivo y su orden es múltiplo de 5. De $p = 7$ sale un elemento de tipo $(2,3)$; al cubo, una transposición. Un subgrupo de S_5 transitivo con una transposición es S_5 entero. Conclusión: Gal $= S_5$, luego la quíntica *no* es resoluble por radicales — y todo se ha decidido con aritmética módulo 3 y módulo 7.

(b) $f(x) = x^5 - 2$ tiene grupo F_{20}, de orden 20. El discriminante es $50000 = 2^4 \cdot 5^5$, así que se descarta $p = 5$. Los tipos que aparecen son $(1,4)$ para $p = 3, 7, 13, 17, 23$, luego (5) para $p = 11$ y $(1,2,2)$ para $p = 19$. Nunca aparece $(2,3)$ ni $(1,1,3)$, y en efecto no pueden: el grupo es $\mathbb{Z}/5\mathbb{Z} \rtimes \mathbb{Z}/4\mathbb{Z}$, que tiene elementos de órdenes $1, 2, 4, 5$ pero ninguno de orden 3 ni 6. La ausencia de ciertos tipos es tan informativa como la presencia de otros.

(c) $f(x) = x^4 + 1$ tiene grupo de Klein. Este es el ejemplo clásico de polinomio irreducible sobre $\mathbb{Q}$ que es *reducible módulo todo primo*. En efecto, para $p = 3, 5, 7, 11, 13, 19, 23$ el tipo es

siempre $(2,2)$, y para $p=17$ es $(1,1,1,1)$. Jamás aparece un 4-ciclo, y no puede aparecer: el grupo es $\mathbb{Z}/2\mathbb{Z}\times\mathbb{Z}/2\mathbb{Z}$, donde todo elemento no trivial tiene orden 2. Como $f=\Phi_8$, el cuerpo de descomposición es $\mathbb{Q}(\zeta_8)$ y el grupo es $(\mathbb{Z}/8)^\times\cong\mathbb{Z}/2\times\mathbb{Z}/2$. Un primo se descompone del todo exactamente cuando $p\equiv 1 \pmod 8$, y 17 es el primero.

(d) $f(x)=x^3-2$ tiene grupo S_3. El discriminante es -108, no cuadrado, luego el grupo contiene permutaciones impares y es S_3 y no A_3. Los tipos confirman: (3) para $p=7,13,19$ y $(1,2)$ para $p=5,11,17,23$. La proporción es reveladora — aparecen 3-ciclos y transposiciones, y ninguna otra cosa, que es exactamente el catálogo de S_3. El primo 3 se descarta porque divide al discriminante.

Ejemplo 26.63 (Chebotarev: las proporciones también hablan)**. (a) El enunciado, en versión utilizable.** El teorema de densidad de Chebotarev afirma que, al recorrer los primos, la proporción de aquellos cuyo tipo de ciclo es τ tiende a la proporción de elementos de tipo τ en el grupo de Galois. En otras palabras: el grupo es *estadísticamente visible* en la tabla de factorizaciones, no solo lógicamente deducible de ella.

(b) Las densidades de S_5. En S_5, de los 120 elementos, 1 es la identidad, 10 son transposiciones, 20 son 3-ciclos, 15 son productos de dos transposiciones, 30 son 4-ciclos, 24 son 5-ciclos y 20 son de tipo $(2,3)$. Así, para x^5-4x+2 la proporción de primos que se descomponen totalmente debe tender a $1/120$, la de los que quedan irreducibles a $24/120=1/5$, y la de tipo $(2,3)$ a $20/120=1/6$.

(c) El caso ciclotómico, que es el más viejo. Para $f=\Phi_5=x^4+x^3+x^2+x+1$ el grupo es $(\mathbb{Z}/5)^\times\cong\mathbb{Z}/4$, cíclico. Los tipos observados son (4) para $p=3,7,13,17,23$, $(2,2)$ para $p=19$ y $(1,1,1,1)$ para $p=11$. La regla exacta es que el tipo de p depende solo de p mód 5: se descompone del todo si $p\equiv 1$, da $(2,2)$ si $p\equiv 4$ y da (4) si $p\equiv 2,3$. Las densidades son $1/4$, $1/4$ y $1/2$: exactamente las proporciones de elementos de orden 1, 2 y 4 en $\mathbb{Z}/4$. Este caso particular es el teorema de Dirichlet sobre primos en progresiones aritméticas.

(d) El valor práctico y su límite. El método de Dedekind es *completamente algorítmico*, y es en esencia lo que implementan los sistemas de álgebra computacional para calcular grupos de Galois. Su límite es que nunca da una demostración de que el grupo sea *pequeño*: cada primo solo aporta cotas por abajo. Para acotar por arriba hacen falta el discriminante, las resolventes, o el conocimiento previo de que la extensión es abeliana.

Cuidado

Tres precauciones. Primera: hay que descartar los primos que dividen al discriminante; usarlos produce factores repetidos y conclusiones falsas. Segunda: el teorema da elementos *contenidos* en el grupo, nunca la afirmación de que el grupo se agote con ellos; concluir "el grupo es $\mathbb{Z}/4$ porque solo he visto 4-ciclos" es un error de razonamiento, aunque a veces acierte por casualidad. Y tercera: el caso (c) del primer ejemplo enseña que un polinomio puede ser irreducible sobre $\mathbb{Q}$ y reducible módulo todos los primos a la vez; eso ocurre exactamente cuando el grupo de Galois no tiene ningún elemento de orden n, es decir cuando no es cíclico generado por un n-ciclo.

Pausa para pensar

Aplíquese el método a x^4-2. Los tipos que aparecen son $(2,2)$ para $p=3,11,17,19$, un 4-ciclo para $p=5,13$ y $(1,1,2)$ para $p=7,23$. Hay un 4-ciclo, luego el grupo tiene orden múltiplo de 4 y contiene un elemento de orden 4; hay un elemento de tipo $(1,1,2)$, que es una transposición. Un subgrupo transitivo de S_4 con un 4-ciclo y una transposición no adyacente es D_4, de orden 8 — y en efecto el cuerpo de descomposición es $\mathbb{Q}(\sqrt[4]{2},\mathrm{i})$, de grado 8 sobre $\mathbb{Q}$. ¿Sabría el lector predecir, sin calcular, qué tipos *no* pueden aparecer nunca para este polinomio?

26.7. Sumas de Gauss, reciprocidad cuadrática y el teorema de Kronecker–Weber

Hemos visto que $\mathbb{Q}(\zeta_p)$ tiene grupo de Galois cíclico de orden $p-1$. Un grupo cíclico de orden par tiene un único subgrupo de índice 2, luego $\mathbb{Q}(\zeta_p)$ contiene un único subcuerpo cuadrático. ¿Cuál es? La respuesta es sorprendentemente limpia, se obtiene con una suma finita de raíces de la unidad, y de ella cae como fruta madura la ley de reciprocidad cuadrática, el resultado que Gauss llamó *theorema aureum* y demostró de ocho maneras distintas a lo largo de su vida.

Definición 26.64 (Símbolo de Legendre y suma de Gauss)**.** Sea p un primo impar y a un entero. El *símbolo de Legendre* $\left(\frac{a}{p}\right)$ vale 0 si $p \mid a$, vale $+1$ si a es un resto cuadrático módulo p (es decir, si $x^2 \equiv a$ tiene solución) y vale -1 en caso contrario. El criterio de Euler lo calcula:

$$\left(\frac{a}{p}\right) \equiv a^{(p-1)/2} \pmod{p}.$$

La *suma de Gauss* asociada a p es

$$g_p \;=\; \sum_{a=1}^{p-1} \left(\frac{a}{p}\right) \zeta_p^{\,a}, \qquad \zeta_p = \mathrm{e}^{2\pi i/p}.$$

Teorema 26.65 (El cuadrado de la suma de Gauss)**.** *Con la notación anterior,*

$$g_p^2 \;=\; (-1)^{(p-1)/2}\, p \;=\; p^*,$$

de modo que $\sqrt{p^*} \in \mathbb{Q}(\zeta_p)$*. En consecuencia el único subcuerpo cuadrático de* $\mathbb{Q}(\zeta_p)$ *es* $\mathbb{Q}(\sqrt{p^*})$*, con* $p^* = p$ *si* $p \equiv 1 \pmod 4$ *y* $p^* = -p$ *si* $p \equiv 3 \pmod 4$*.*

Idea clave

Qué dice esto en cristiano. El símbolo de Legendre es un semáforo: vale 1 si a es un cuadrado módulo p, -1 si no lo es y 0 si p divide a a. La suma de Gauss lo empaqueta todo en un solo número complejo, y el teorema dice que su cuadrado es $\pm p$, con el signo decidido por el resto de p módulo 4. La consecuencia es espectacular: dentro del cuerpo ciclotómico $\mathbb{Q}(\zeta_p)$ vive $\sqrt{\pm p}$, es decir, todo cuerpo cuadrático está metido dentro de uno

ciclotómico. Ese es el germen del teorema de Kronecker–Weber.

Ejemplo 26.66 (Cuadrados módulo un primo)**. (a) La lista completa módulo 7.** Los cuadrados no nulos son $1,4,2,2,4,1$, es decir el conjunto $\{1,2,4\}$: tres residuos cuadráticos y tres no. Siempre son exactamente $\frac{p-1}{2}$ de cada, porque a y $-a$ dan el mismo cuadrado.

(b) Los símbolos. Con $p=7$: $\left(\frac{1}{7}\right)=\left(\frac{2}{7}\right)=\left(\frac{4}{7}\right)=1$ y $\left(\frac{3}{7}\right)=\left(\frac{5}{7}\right)=\left(\frac{6}{7}\right)=-1$, coherente con la lista anterior.

(c) El signo de la suma de Gauss. Como $7\equiv 3 \pmod 4$, se tiene $g_7^2=(-1)^3\cdot 7=-7$, luego $\sqrt{-7}\in\mathbb{Q}(\zeta_7)$. Para $p=5$, que es $1 \pmod 4$, sale $g_5^2=5$ y por tanto $\sqrt{5}\in\mathbb{Q}(\zeta_5)$: esa es la razón de que el pentágono regular sea construible con regla y compás.

(d) Reciprocidad cuadrática en acción. Se cumple $\left(\frac{3}{7}\right)\left(\frac{7}{3}\right)=-1$, y la ley predice $(-1)^{\frac{3-1}{2}\cdot\frac{7-1}{2}}=(-1)^{1\cdot 3}=-1$. Coincide. La ley suplementaria dice que $\left(\frac{2}{p}\right)=1$ exactamente cuando $p\equiv\pm 1 \pmod 8$, y en efecto vale 1 para $p=7,17,23$ (restos $7,1,7$) y -1 para $p=11,19$ (resto 3).

Cuidado

[El símbolo de Legendre no distingue cuántas raíces hay] $\left(\frac{a}{p}\right)=1$ significa que hay *dos* raíces cuadradas de a módulo p, no una; son $\pm r$. Y $\left(\frac{a}{p}\right)=-1$ significa cero raíces. Segundo aviso: el símbolo es multiplicativo, $\left(\frac{ab}{p}\right)=\left(\frac{a}{p}\right)\left(\frac{b}{p}\right)$, de modo que el producto de dos no residuos es un residuo; "no cuadrado por no cuadrado da cuadrado" suena raro y es correcto. Y tercero: el símbolo de Jacobi generaliza el de Legendre a módulos compuestos, pero allí un valor 1 ya *no* garantiza que haya raíz cuadrada.

Idea clave

El símbolo de Legendre es un homomorfismo $(\mathbb{Z}/p)^\times\to\{\pm 1\}$, y es exactamente el carácter que corta el grupo cíclico $\mathrm{Gal}(\mathbb{Q}(\zeta_p)/\mathbb{Q})\cong(\mathbb{Z}/p)^\times$ por la mitad. La suma de Gauss es lo que se obtiene al promediar las raíces de la unidad *con signos*: los restos cuadráticos suman una cosa, los no restos suman otra, y la diferencia entre ambas mitades es precisamente la raíz cuadrada que buscamos. Es la misma maniobra que hicimos con los períodos de Gauss del heptadecágono: $\eta_0-\eta_1$ era $\sqrt{17}$. Aquí la maniobra se hace de golpe para todo primo. Y una vez que se sabe que $\sqrt{p^*}$ vive dentro de $\mathbb{Q}(\zeta_p)$, comparar dos primos p y q se convierte en un ejercicio de teoría de Galois: hay que ver cómo actúa el Frobenius en q sobre ese $\sqrt{p^*}$, y la respuesta se puede leer de dos maneras. Igualar las dos lecturas *es* la ley de reciprocidad.

Regla: localizar el subcuerpo cuadrático y aplicar la reciprocidad

Este es el procedimiento que convierte una pregunta aritmética (¿es a un cuadrado módulo p?) en un cálculo mecánico de pocas líneas. Se usa constantemente al decidir si

un polinomio tiene raíces módulo un primo.

1. **Fija el primo y calcula p^*.** Si $p \equiv 1 \pmod 4$ toma $p^* = p$; si $p \equiv 3 \pmod 4$ toma $p^* = -p$. Ese entero es el discriminante del subcuerpo cuadrático de $\mathbb{Q}(\zeta_p)$.
2. **Escribe la torre.** $\mathbb{Q} \subset \mathbb{Q}(\sqrt{p^*}) \subset \mathbb{Q}(\zeta_p)$, con grados 2 y $(p-1)/2$. El subgrupo que fija $\mathbb{Q}(\sqrt{p^*})$ es el de los restos cuadráticos, de índice 2.
3. **Reduce el símbolo a factores primos.** $\left(\frac{a}{p}\right)$ es multiplicativo en a, así que factoriza a y trata por separado -1, 2 y cada primo impar.
4. **Usa los dos suplementos.** $\left(\frac{-1}{p}\right) = (-1)^{(p-1)/2}$, que vale $+1$ si $p \equiv 1 \pmod 4$; y $\left(\frac{2}{p}\right) = (-1)^{(p^2-1)/8}$, que vale $+1$ si $p \equiv \pm 1 \pmod 8$.
5. **Da la vuelta con la reciprocidad.** Para p, q primos impares distintos, $\left(\frac{p}{q}\right)\left(\frac{q}{p}\right) = (-1)^{\frac{p-1}{2}\cdot\frac{q-1}{2}}$. Es decir: los dos símbolos coinciden salvo que p y q sean *ambos* $\equiv 3 \pmod 4$, en cuyo caso son opuestos.
6. **Reduce el numerador módulo el nuevo denominador.** Tras cada vuelta el numerador se hace más pequeño; repite los pasos 3 a 5 hasta llegar a símbolos triviales. El algoritmo termina siempre, y en un número de pasos comparable al del algoritmo de Euclides.
7. **Traduce el resultado.** Un $+1$ dice que el polinomio $x^2 - a$ se parte módulo p; un -1 dice que es irreducible módulo p, y entonces el Frobenius en p actúa como la transposición sobre las dos raíces de $\sqrt{a}$.

Ejemplo 26.67 (Sumas de Gauss calculadas y la ley leída en cuatro casos)**. (a) Los primeros valores de g_p.** Para $p = 3$ los restos cuadráticos son $\{1\}$ y los no restos $\{2\}$, luego $g_3 = \zeta_3 - \zeta_3^2 = i\sqrt{3}$, numéricamente $1{,}73205\,i$, y $g_3^2 = -3$. Aquí $3 \equiv 3 \pmod 4$, así que $p^* = -3$: el subcuerpo cuadrático de $\mathbb{Q}(\zeta_3)$ es $\mathbb{Q}(\sqrt{-3})$, que de hecho *es* $\mathbb{Q}(\zeta_3)$ entero, pues $\zeta_3 = (-1+\sqrt{-3})/2$. Para $p = 5$ los restos son $\{1, 4\}$ y los no restos $\{2, 3\}$, de modo que $g_5 = (\zeta_5 + \zeta_5^4) - (\zeta_5^2 + \zeta_5^3) = 2{,}23607 = \sqrt{5}$ y $g_5^2 = 5$; en efecto $5 \equiv 1 \pmod 4$. Para $p = 7$ se obtiene $g_7 = 2{,}64575\,i = i\sqrt{7}$ y $g_7^2 = -7$; para $p = 11$, $g_{11} = 3{,}31662\,i$ y $g_{11}^2 = -11$; para $p = 13$, $g_{13} = 3{,}60555 = \sqrt{13}$; y para $p = 17$, $g_{17} = 4{,}12311 = \sqrt{17}$, que es exactamente el $\eta_0 - \eta_1$ del heptadecágono.

(b) La ley verificada a mano en siete pares. Recordemos el enunciado: el producto $\left(\frac{p}{q}\right)\left(\frac{q}{p}\right)$ vale $(-1)^{\frac{p-1}{2}\frac{q-1}{2}}$.

p	q	$\left(\frac{p}{q}\right)$	$\left(\frac{q}{p}\right)$	producto	$(-1)^{\frac{p-1}{2}\frac{q-1}{2}}$
3	5	-1	-1	$+1$	$+1$
3	7	-1	$+1$	-1	-1
5	7	-1	-1	$+1$	$+1$
7	11	-1	$+1$	-1	-1
13	17	$+1$	$+1$	$+1$	$+1$
11	19	$+1$	-1	-1	-1
5	13	-1	-1	$+1$	$+1$

Los únicos casos con producto -1 son $(3,7)$, $(7,11)$ y $(11,19)$, y en los tres ambos primos son $\equiv 3 \pmod 4$, tal como anuncia la fórmula.

(c) Un cálculo completo con el algoritmo. ¿Es 30 un resto cuadrático módulo 53? Factorizamos: $\left(\frac{30}{53}\right) = \left(\frac{2}{53}\right)\left(\frac{3}{53}\right)\left(\frac{5}{53}\right)$. Como $53 \equiv 5 \pmod 8$, el suplemento da $\left(\frac{2}{53}\right) = -1$. Como $53 \equiv 1 \pmod 4$, la reciprocidad no cambia signos: $\left(\frac{3}{53}\right) = \left(\frac{53}{3}\right) = \left(\frac{2}{3}\right) = -1$ (porque $3 \equiv 3 \pmod 8$), y $\left(\frac{5}{53}\right) = \left(\frac{53}{5}\right) = \left(\frac{3}{5}\right) = -1$ (los restos módulo 5 son 1 y 4). El producto de los tres es $(-1)^3 = -1$: no lo es. Comprobación directa: los cuadrados módulo 53 son $1, 4, 9, 16, 25, 36, 49, 11, 28, 47, 15, 38, 10, 37, 13, 44, 24, 6, 43, 29, 17, 7, 52, 46, 42, 40$, y 30 no aparece.

(d) Por qué esto decide factorizaciones. El polinomio $x^2 - a$ se parte módulo p exactamente cuando $\left(\frac{a}{p}\right) = +1$. Por ejemplo x^2+1 se parte módulo $5, 13, 17, 29$ y es irreducible módulo $3, 7, 11, 19, 23$: la frontera es $p \equiv 1$ frente a $p \equiv 3 \pmod 4$, que es el suplemento del paso 4. Con el método de Dedekind eso significa que el Frobenius en p es la identidad en el primer caso y la transposición en el segundo, y la densidad de cada comportamiento es $1/2$, como manda el teorema de Chebotarev.

Ejemplo 26.68 (Kronecker–Weber: todo lo abeliano es ciclotómico)**. (a) El enunciado.** *Toda* extensión finita $K/\mathbb{Q}$ con grupo de Galois abeliano está contenida en algún $\mathbb{Q}(\zeta_m)$. El menor m que sirve se llama el *conductor* de K. Es un teorema enorme: dice que los cuerpos ciclotómicos, que parecen una familia muy particular, agotan de un plumazo todo el mundo abeliano sobre $\mathbb{Q}$. Enunciado por Kronecker en 1853 con una demostración incompleta, completado por Weber en 1886 (con una laguna) y por Hilbert en 1896.

(b) El caso cuadrático, comprobado. Todo $\mathbb{Q}(\sqrt{d})$ es abeliano (grupo $\mathbb{Z}/2$), así que debe caber en algún ciclotómico. Y cabe:

d	vive en	$\varphi(m)$	conductor
-1	$\mathbb{Q}(\zeta_4)$	2	4
2	$\mathbb{Q}(\zeta_8)$	4	8
-2	$\mathbb{Q}(\zeta_8)$	4	8
3	$\mathbb{Q}(\zeta_{12})$	4	12
-3	$\mathbb{Q}(\zeta_3)$	2	3
5	$\mathbb{Q}(\zeta_5)$	4	5
17	$\mathbb{Q}(\zeta_{17})$	16	17
6	$\mathbb{Q}(\zeta_{24})$	8	24

La regla general: el discriminante de $\mathbb{Q}(\sqrt{d})$ (con d libre de cuadrados) es d si $d \equiv 1 \pmod 4$ y $4d$ en caso contrario, y el conductor es su valor absoluto. Así $d=-1$ da disc -4 y conductor 4; $d=2$ da disc 8; $d=-3$ da disc -3 y conductor 3; $d=5$ da disc 5; $d=17$ da disc 17. Los casos compuestos se arman multiplicando: $\sqrt{6}=\sqrt{2}\cdot\sqrt{3}$ necesita a la vez el 8 y el 12, y $\operatorname{mcm}(8,12)=24$.

(c) De dónde salen los $\sqrt{p^*}$. El teorema de la suma de Gauss ya nos da la mitad del trabajo: $\sqrt{p^*}\in\mathbb{Q}(\zeta_p)$ para todo primo impar p. Para $\sqrt{-1}$ usamos $\zeta_4=i$; para $\sqrt{2}$ usamos $\zeta_8+\zeta_8^{-1}=2\cos(\pi/4)=\sqrt{2}$. Con esas tres piezas y multiplicatividad se alcanza cualquier $\sqrt{d}$: por ejemplo $\sqrt{3}=\sqrt{-3}/\sqrt{-1}$ vive en $\mathbb{Q}(\zeta_4,\zeta_3)=\mathbb{Q}(\zeta_{12})$, y $\sqrt{-2}=\sqrt{-1}\sqrt{2}$ vive en $\mathbb{Q}(\zeta_8)$ porque $\zeta_8^2=i$.

(d) Qué queda fuera. La palabra *abeliano* es esencial. $\mathbb{Q}(\sqrt[3]{2})$ no es normal y ni siquiera se plantea; $\mathbb{Q}(\sqrt[3]{2},\zeta_3)$ sí es normal pero su grupo es S_3, no abeliano, y en efecto no cabe en ningún ciclotómico. El intento de repetir el teorema con $\mathbb{Q}$ sustituido por otro cuerpo base es el *duodécimo problema de Hilbert*: para cuerpos cuadráticos imaginarios lo resuelve la teoría de la multiplicación compleja, con valores especiales de funciones modulares en lugar de raíces de la unidad; en general sigue abierto. La teoría de cuerpos de clases da la descripción abstracta del grupo de Galois abeliano máximo, pero no siempre los generadores explícitos.

Cuidado

Tres tropiezos habituales. Primero: el símbolo de Legendre $\left(\frac{a}{p}\right)=+1$ dice que a es un cuadrado *módulo* p, no que sea un cuadrado en $\mathbb{Z}$; 2 es un resto cuadrático módulo 7 porque $3^2=9\equiv 2$. Segundo: la reciprocidad exige que *ambos* sean primos impares distintos; para el 2 y para el -1 hay que usar los suplementos, y con un numerador compuesto hay que factorizar antes (o pasar al símbolo de Jacobi, que se calcula igual pero cuya interpretación como "es un cuadrado" ya no es válida). Tercero: en Kronecker–Weber la contención es en *algún* $\mathbb{Q}(\zeta_m)$, no en $\mathbb{Q}(\zeta_n)$ con n el grado; el conductor puede ser mucho mayor que el grado, como en $\mathbb{Q}(\sqrt{6})$, de grado 2 y conductor 24.

Pausa para pensar

La suma de Gauss g_p tiene módulo $\sqrt{p}$, pero determinar su *signo* (o su argumento) costó a Gauss cuatro años: sabía desde 1801 que $g_p = \sqrt{p}$ o $-\sqrt{p}$ para $p \equiv 1 \pmod 4$, y no consiguió decidir cuál hasta 1805, cuando la respuesta —siempre el signo positivo— se le ocurrió, según escribió, "como un relámpago". Compruébelo en los datos de arriba: $g_5 = +\sqrt{5}$, $g_{13} = +\sqrt{13}$, $g_{17} = +\sqrt{17}$, y para $p \equiv 3 \pmod 4$ siempre $g_p = +i\sqrt{p}$. Que un número esté determinado hasta el signo por álgebra pura y que el signo exija un argumento analítico es un fenómeno recurrente; la demostración moderna más corta usa la fórmula de sumación de Poisson.

26.8. Solubilidad por radicales

Idea clave

Intuición primero. "Resolver por radicales" significa construir las raíces con una torre de cuerpos en la que cada piso se obtiene del anterior añadiendo una raíz n-ésima. Cada piso de la torre es un paso elemental de la fórmula.
El puente entre las dos mitades del argumento es este: añadir $\sqrt[n]{a}$ produce una extensión cuyo grupo es *cíclico*, es decir, lo más simple que hay después de trivial. Como la correspondencia de Galois traduce torres de cuerpos en cadenas de subgrupos invirtiendo el orden, una torre de radicales se convierte automáticamente en una cadena de subgrupos con cocientes abelianos: exactamente la definición de *grupo soluble*. Por eso el adjetivo "soluble", que suena a jerga, es literalmente "que se puede resolver".
De modo que la demostración final tiene solo dos ingredientes, y ninguno es un cálculo: la correspondencia —que ya está probada— y el hecho de que A_5 es simple, así que S_5 no admite ninguna cadena de esas. Toda la maquinaria del capítulo existe para poder decir esas dos frases con derecho.

Definición 26.69. $f \in K[x]$ es *soluble por radicales* si sus raíces se expresan desde K con operaciones de cuerpo y raíces n-ésimas; formalmente, si su cuerpo de descomposición cabe en una torre $K = F_0 \subset F_1 \subset \cdots \subset F_m$ con cada $F_{i+1} = F_i(\sqrt[n_i]{a_i})$, $a_i \in F_i$.

Teorema 26.70 (Criterio de Galois)**.** *f es soluble por radicales* $\iff$ Gal *(de su cuerpo de descomposición sobre K) es un grupo* soluble*.*

Idea de la demostración. ($\Leftarrow$ de la solubilidad del grupo a los radicales, y viceversa; característica 0.) El ingrediente técnico: si F contiene las raíces n-ésimas de la unidad, la extensión $F(\sqrt[n]{a})/F$ es de Galois con grupo *cíclico* (un automorfismo manda $\sqrt[n]{a} \mapsto \zeta^k \sqrt[n]{a}$); recíprocamente, toda extensión cíclica de grado n es de esa forma (teorema 90 de Hilbert / resolventes de Lagrange). Entonces: una torre de radicales produce, vía la correspondencia de Galois, una cadena de subgrupos con cocientes cíclicos (y adjuntar de antemano las raíces de la unidad solo agrega trozos abelianos): el grupo es soluble. A la inversa, una cadena con cocientes abelianos se refina a cocientes cíclicos, y cada peldaño cíclico se realiza como extensión por un radical: la fórmula existe. □

Idea clave

Qué dice esto en cristiano. Este es el corazón del capítulo y probablemente del libro. Resolver por radicales significa llegar a las raíces con una torre de extensiones, cada piso añadiendo una raíz n-ésima. Cada piso de esa torre se traduce, en el lado de los grupos, en un cociente abeliano. Por tanto "existe una fórmula con radicales" y "el grupo se desmonta en capas abelianas" son literalmente la misma frase escrita en dos idiomas. El diccionario de Galois convierte una pregunta imposible sobre fórmulas en una pregunta finita sobre un grupo pequeño.

Ejemplo 26.71 (Grupos de Galois y solubilidad, caso por caso)**. (a) Los grados pequeños siempre se resuelven.** Para x^2-2 el grupo es $\mathbb{Z}_2$; para x^3-2 es S_3; para x^4-2 es el diedral de orden 8. Los tres son solubles, y de ahí salen la fórmula cuadrática, la de Cardano y la de Ferrari.

(b) Una quíntica que sí es soluble. Para x^5-2 el grupo tiene orden 20: es $\mathbb{Z}_5 \rtimes \mathbb{Z}_4$, el grupo de Frobenius F_{20}. Es soluble, con la cadena $1 \trianglelefteq \mathbb{Z}_5 \trianglelefteq F_{20}$, y en efecto sus raíces se escriben explícitamente como $\sqrt[5]{2}\,\zeta_5^k$ para $k=0,\dots,4$.

(c) Una quíntica que no. Para x^5-4x+2 el grupo es S_5 entero. Es irreducible por Eisenstein con $p=2$, y S_5 no es soluble porque contiene a A_5, que es simple no abeliano. Por tanto esta ecuación concreta *no* tiene solución por radicales. No es que nadie la haya encontrado: es que no existe.

(d) El discriminante como detector. Para x^3-3x-1 el discriminante vale $81=9^2$, un cuadrado perfecto, luego $\sqrt{\Delta}$ es racional y el grupo está dentro de $A_3=\mathbb{Z}_3$: es cíclico de orden 3. Para x^3-2 el discriminante es -108, que no es cuadrado, y el grupo es S_3 completo. Un solo cálculo distingue los dos casos posibles en grado 3.

Cuidado

[Abel–Ruffini no dice que no haya soluciones] Toda ecuación polinómica compleja tiene sus raíces, garantizadas por el teorema fundamental del álgebra; lo que no existe es una *fórmula general* con radicales para grado 5 o más. Segundo aviso: hay infinitas quínticas concretas que sí se resuelven, como la del apartado (b); la imposibilidad es sobre el caso genérico. Y tercero: sí existen fórmulas explícitas para la quíntica usando otras funciones. Hermite dio una en 1858 con funciones elípticas modulares. La prohibición es sobre los radicales, no sobre la explicitud.

Teorema 26.72 (Abel–Ruffini, versión de Galois)**.** *Para cada $n \geq 5$ existen polinomios de grado n sobre $\mathbb{Q}$ no solubles por radicales; en particular no existe fórmula general. Ejemplo concreto: $f(x)=x^5-4x+2$.*

Demostración completa para x^5-4x+2. Paso 1: irreducible por Eisenstein con $p=2$. En consecuencia $G=\mathrm{Gal}$ actúa *transitivamente* sobre las 5 raíces, y por órbita–estabilizador $5 \mid |G|$. *Paso 2: exactamente 3 raíces reales.* $f'=5x^4-4$ tiene dos ceros reales $\pm(4/5)^{1/4} \approx \pm 0.946$; $f(-0.946) \approx 4.8 > 0$ y $f(0.946) \approx -0.8 < 0$, con $f(\pm\infty)=\pm\infty$: tres cruces con el eje, y las

otras 2 raíces son complejas conjugadas. *Paso 3: G contiene un 5-ciclo y una trasposición.* Por Cauchy (capítulo 25), $5 \mid |G|$ da un elemento de orden 5, que en S_5 es un 5-ciclo. La restricción de la conjugación compleja al cuerpo de descomposición es un automorfismo que intercambia las dos raíces complejas y fija las tres reales: una trasposición. *Paso 4: 5-ciclo + trasposición generan S_5.* Renumerando, sea $\tau = (1\,2)$ y σ un 5-ciclo; alguna potencia de σ manda $1 \mapsto 2$, y sigue siendo 5-ciclo, digamos $\sigma = (1\,2\,a\,b\,c)$. Conjugando: $\sigma\tau\sigma^{-1} = (2\,a)$, $\sigma(2\,a)\sigma^{-1} = (a\,b)$, etc.: obtenemos las trasposiciones consecutivas $(1\,2), (2\,a), (a\,b), (b\,c)$, que generan todas las trasposiciones, y estas generan S_5. *Paso 5*: $G = S_5$ no es soluble (Teorema 25.68 del capítulo 25: A_5 es simple no abeliano). Por el criterio de Galois, f no es soluble por radicales. □ □

Ejemplo 26.73 (Solubles y no solubles)**. (a) Grados bajos.** Para $n \leq 4$ todos los polinomios son solubles, porque S_2, S_3, S_4 son grupos solubles: las cadenas son $S_3 \trianglerighteq A_3 \trianglerighteq 1$ y $S_4 \trianglerighteq A_4 \trianglerighteq V_4 \trianglerighteq 1$, con cocientes cíclicos. Las fórmulas de Cardano y Ferrari son la traducción explícita de esas cadenas.

(b) El primer no soluble. $x^5 - x - 1$ tiene grupo S_5, y S_5 no es soluble porque A_5 es simple y no abeliano. Otro ejemplo clásico es $x^5 - 6x + 3$, cuyo grupo también es S_5 por el criterio de Eisenstein más un conteo de raíces reales.

(c) Quintas sí solubles. $x^5 - 2$ tiene grupo de orden 20, el producto semidirecto $\mathbb{Z}/5 \rtimes \mathbb{Z}/4$, que es soluble; sus raíces son $\sqrt[5]{2}\,\zeta_5^k$, y se escriben con radicales sin ningún problema. El grado no decide: el grupo decide.

(d) Qué significa exactamente. El teorema no dice que las raíces no existan ni que no se puedan calcular: $x^5 - x - 1$ tiene una raíz real aproximadamente igual a 1,167304, calculable con la precisión que se quiera. Dice que no hay ninguna *fórmula* con sumas, productos y radicales que la exprese. Es una afirmación sobre el lenguaje disponible, no sobre la existencia.

Idea clave

Uno de los grandes momentos de la matemática. Las fórmulas de grados 2, 3, 4 existen porque S_2, S_3, S_4 son solubles; la de grado 5 no existe porque S_5 no lo es. La pregunta "¿hay fórmula?" —aparentemente sobre manipulaciones algebraicas— tiene respuesta puramente *estructural*: sobre simetrías. Nótese la precisión del criterio: no dice que ninguna quíntica se resuelva ($x^5 - 1$ sí, ver ejercicios), sino que la resolubilidad de cada una la decide su grupo. Galois escribió esta teoría antes de morir en duelo a los 20 años, la noche anterior, en una carta que terminaba: "no tengo tiempo".

Regla: decidir si un número es construible con regla y compás

La regla trazada entre dos puntos y el compás centrado en un punto que pasa por otro solo permiten cortar rectas con rectas, rectas con circunferencias y circunferencias con circunferencias. Traducido al álgebra: las coordenadas de cada punto nuevo salen de resolver sistemas lineales o ecuaciones de segundo grado sobre el cuerpo generado por lo ya construido. De ahí el procedimiento.

1. **Convierte la construcción en un número.** Casi todo problema clásico se reduce

a: ¿es construible cierto número real α? Duplicar el cubo es construir $\sqrt[3]{2}$; trisecar un ángulo 3θ dado es construir $\cos\theta$ a partir de $\cos 3\theta$; el polígono regular de n lados es construir $\cos(2\pi/n)$.

2. **Recuerda el criterio necesario.** Si α es construible, entonces
$$[\mathbb{Q}(\alpha):\mathbb{Q}] = 2^k \quad \text{para algún } k \geq 0.$$
La razón: cada paso de la construcción sube el grado en 1 o en 2, así que α vive en una torre $\mathbb{Q} = K_0 \subset K_1 \subset \cdots \subset K_m$ con $[K_{j+1} : K_j] = 2$; por la torre $[K_m : \mathbb{Q}] = 2^m$, y $[\mathbb{Q}(\alpha):\mathbb{Q}]$ divide a ese número.

3. **Calcula el polinomio mínimo de α y mide su grado.** Este es todo el trabajo. Si el grado no es una potencia de 2, la construcción es *imposible*, y no hay ingenio que valga: no es que nadie la haya encontrado, es que no existe.

4. **Cuidado con la dirección del criterio.** Grado potencia de 2 es necesario pero *no* suficiente. La condición exacta es que el cuerpo de descomposición del polinomio mínimo tenga grado potencia de 2 sobre $\mathbb{Q}$, equivalentemente que Gal sea un 2-grupo. Existen polinomios irreducibles de grado 4 con grupo S_4 o A_4 cuyas raíces no son construibles.

5. **Para construir de verdad, exhibe la torre.** Si el número sí es construible, la demostración útil es la cadena de raíces cuadradas anidadas: escribir α con sumas, productos, cocientes y $\sqrt{\ }$ es exactamente dar la receta de regla y compás, porque $\sqrt{a}$ se construye con el teorema de la altura.

6. **Para polígonos, usa Gauss–Wantzel.** El n-ágono regular es construible si y solo si
$$n = 2^a p_1 p_2 \cdots p_r$$
con los p_i primos de Fermat distintos, es decir primos de la forma $2^{2^s}+1$: $3, 5, 17, 257, 65537$. El motivo es que $[\mathbb{Q}(\zeta_n):\mathbb{Q}] = \varphi(n)$ y $[\mathbb{Q}(\cos(2\pi/n)):\mathbb{Q}] = \varphi(n)/2$, y hay que exigir que eso sea potencia de 2.

Ejemplo 26.74 (Las tres imposibilidades clásicas, liquidadas en tres líneas cada una)**.** Los griegos dejaron tres problemas abiertos que resistieron dos mil años. Con el criterio del grado caen los tres.

Duplicar el cubo. Se pide un cubo de volumen doble que el dado, es decir el lado $\alpha = \sqrt[3]{2}$. Su polinomio mínimo es $x^3 - 2$, irreducible por Eisenstein con $p = 2$, luego $[\mathbb{Q}(\sqrt[3]{2}):\mathbb{Q}] = 3$. Y 3 no es potencia de 2 (dicho con precisión: $3 \nmid 2^k$ para todo k). Imposible.

Trisecar el ángulo. No se trata de trisecar *todos* los ángulos: basta exhibir uno que resista. Tomemos $60°$, que es construible sin esfuerzo (triángulo equilátero). Trisecarlo significa construir $c = \cos 20°$. De la identidad $\cos 3\theta = 4\cos^3\theta - 3\cos\theta$ con $3\theta = 60°$ y $\cos 60° = \frac{1}{2}$ sale
$$4c^3 - 3c = \tfrac{1}{2}, \qquad \text{es decir} \qquad 8c^3 - 6c - 1 = 0.$$

Este cúbico es irreducible sobre $\mathbb{Q}$: por el teorema de la raíz racional los únicos candidatos son $\pm 1, \pm\frac{1}{2}, \pm\frac{1}{4}, \pm\frac{1}{8}$, y ninguno anula. Luego $[\mathbb{Q}(\cos 20^\circ) : \mathbb{Q}] = 3$. Imposible. (Equivalentemente, $2\cos 20^\circ$ tiene polinomio mínimo $x^3 - 3x - 1$, el mismo del ejemplo anterior; su grupo de Galois es $\mathbb{Z}/3$, que no es un 2-grupo.)

Cuadrar el círculo. Se pide un cuadrado del área del círculo unidad, es decir el lado $\sqrt{\pi}$. Aquí el criterio del grado no basta por sí solo: hace falta el teorema de Lindemann (1882) de que π es trascendente. Si $\sqrt{\pi}$ fuera construible sería algebraico, luego $\pi = (\sqrt{\pi})^2$ también lo sería. Imposible. Nótese la diferencia de naturaleza: las dos primeras son imposibilidades de grado, esta es una imposibilidad de trascendencia, y por eso tardó cincuenta años más en resolverse.

Pausa para pensar

Hay una cuarta imposibilidad clásica que suele olvidarse y que se resuelve igual: la *rectificación* de la circunferencia, construir un segmento de longitud 2π. Misma razón, misma frase de Lindemann.

Ejemplo 26.75 (Qué polígonos regulares se pueden dibujar)**.** Aplicamos Gauss–Wantzel a los primeros casos. El número clave es siempre $[\mathbb{Q}(\cos(2\pi/n)) : \mathbb{Q}] = \varphi(n)/2$ para $n \geq 3$.

n	$\varphi(n)$	$\varphi(n)/2$	¿potencia de 2?	veredicto
3	2	1	sí	construible (equilátero)
4	2	1	sí	construible (cuadrado)
5	4	2	sí	construible
6	2	1	sí	construible
7	6	3	**no**	**imposible**
8	4	2	sí	construible
9	6	3	**no**	**imposible**
10	4	2	sí	construible
11	10	5	**no**	**imposible**
15	8	4	sí	construible ($15 = 3 \cdot 5$)
17	16	8	sí	construible (Gauss, 1796)

El pentágono, hecho explícito. Para $n = 5$, el número $\beta = 2\cos(2\pi/5)$ tiene polinomio mínimo $x^2 + x - 1$: en efecto, si $\zeta = \mathrm{e}^{2\pi i/5}$ entonces $1 + \zeta + \zeta^2 + \zeta^3 + \zeta^4 = 0$; dividiendo entre ζ^2 y agrupando $\zeta + \zeta^{-1} = \beta$ queda $\beta^2 + \beta - 1 = 0$. Resolviendo,

$$\beta = \frac{-1 + \sqrt{5}}{2} = 0{,}6180339887\ldots, \qquad \cos 72^\circ = \frac{\sqrt{5} - 1}{4}.$$

Grado 2: una sola raíz cuadrada, un solo paso de compás. Ahí está la construcción del pentágono, y de paso la aparición de la razón áurea, que no es coincidencia sino la misma ecuación.

El heptágono, hecho explícito. Para $n = 7$, el número $2\cos(2\pi/7) = 1{,}2469796\ldots$ tiene polinomio mínimo $x^3 + x^2 - 2x - 1$, de grado 3. Fin de la discusión: el heptágono regular no

se construye con regla y compás. Compárese con el eneágono ($n = 9$), que falla por la misma razón — y no por casualidad, porque construir el eneágono es trisecar 120°.

Cuidado

El criterio dice qué es imposible con *regla y compás*, no qué es imposible en absoluto. Con una regla graduada (construcción llamada *neusis*), o doblando papel (origami), o con una espiral auxiliar, la trisección del ángulo y la duplicación del cubo sí se logran: el origami alcanza grado 3 porque sus axiomas resuelven cúbicas. Lo que Wantzel demostró en 1837 es una afirmación sobre unas herramientas concretas, y su fuerza está justamente en ser tan precisa.

Regla: demostrar que una ecuación concreta no es resoluble por radicales

El teorema de Abel–Ruffini dice que *la* quíntica general no se resuelve por radicales. Eso es un enunciado sobre coeficientes indeterminados. Lo que sigue es la versión que de verdad se usa: dado un polinomio *concreto* con coeficientes enteros, probar que *esa* ecuación no tiene solución por radicales.

1. **Fija el criterio.** La ecuación $f(x) = 0$ es resoluble por radicales sobre $\mathbb{Q}$ si y solo si $\mathrm{Gal}(f)$ es un grupo soluble, es decir, admite una cadena $1 = G_0 \trianglelefteq G_1 \trianglelefteq \cdots \trianglelefteq G_m = G$ con todos los cocientes G_{i+1}/G_i abelianos. Todo el trabajo consiste, pues, en identificar el grupo lo bastante como para decidir esa palabra.

2. **Demuestra que f es irreducible sobre $\mathbb{Q}$.** Es imprescindible: solo entonces el grupo actúa *transitivamente* sobre las raíces. Eisenstein es la vía más rápida cuando existe un primo p que divide todos los coeficientes salvo el principal y con $p^2 \nmid a_0$; si no, prueba la reducción módulo un primo pequeño o el criterio de las raíces racionales combinado con factorizaciones tentativas.

3. **Deduce que 5 divide el orden del grupo.** Si $\deg f = 5$ y f es irreducible, la órbita de una raíz tiene 5 elementos, luego $5 \mid |G|$ por el teorema de órbita–estabilizador. Por Cauchy, G contiene un elemento de orden 5, y en S_5 los únicos elementos de orden 5 son los 5-ciclos.

4. **Cuenta las raíces reales.** Estudia f', localiza los puntos críticos, evalúa f en ellos y aplica Bolzano en cada intervalo. Si salen exactamente 3 raíces reales y 2 complejas conjugadas, la conjugación compleja restringida al cuerpo de descomposición es un elemento de G que fija tres raíces e intercambia dos: es una *transposición*.

5. **Concluye que $G = S_5$.** Un subgrupo de S_5 que sea transitivo y contenga a la vez un 5-ciclo y una transposición es todo S_5: conjugando el 5-ciclo, la transposición genera sucesivamente todas las transposiciones adyacentes, y estas generan S_5.

6. **Invoca la simplicidad de A_5.** S_5 no es soluble, porque su única cadena normal es $1 \trianglelefteq A_5 \trianglelefteq S_5$ y el cociente $A_5/1 = A_5$ es simple y no abeliano. Por el paso 1, la ecuación no es resoluble por radicales.

7. **Comprueba que el argumento no prueba de más.** Repite el conteo de raíces reales y el orden del grupo sobre un ejemplo *resoluble* de grado 5 para verificar que tu razonamiento distingue los casos. Si tu argumento "demuestra" que x^5-2 no es resoluble, algo has hecho mal: esa sí lo es.

Ejemplo 26.76 (Cuatro quínticas: dos imposibles y dos resolubles)**. (a)** $f(x)=x^5-6x+3$**: el ejemplo canónico.** *Irreducibilidad:* Eisenstein con $p=3$, pues $3\mid 6$, $3\mid 3$ y $9\nmid 3$. Luego $5\mid |G|$ y hay un 5-ciclo.

Raíces reales: $f'(x)=5x^4-6$ se anula en $x=\pm(6/5)^{1/4}=\pm 1.04664$. Los valores críticos son $f(1.04664)=-2.02385$ y $f(-1.04664)=+8.02385$: signos opuestos, luego hay exactamente tres raíces reales. Se ve también con Bolzano puro: $f(-2)=-17$, $f(-1)=8$, $f(0)=3$, $f(1)=-2$, $f(2)=23$, con cambios de signo en $(-2,-1)$, $(0,1)$ y $(1,2)$. Numéricamente las raíces son

$$-1.67094,\quad 0.505501,\quad 1.40164,\quad -0.118104\pm 1.58746\,i,$$

tres reales y un par conjugado. La conjugación compleja es entonces una transposición de G.

Conclusión: G es transitivo, contiene un 5-ciclo y una transposición, luego $G=S_5$, de orden 120. Como S_5 no es soluble, **la ecuación $x^5-6x+3=0$ no se resuelve por radicales**. Y obsérvese lo que esto *no* dice: las cinco raíces existen, están ahí arriba con seis cifras. Lo que no existe es una fórmula que las produzca a partir de -6 y 3 usando solo $+,-,\times,\div$ y radicales.

(b) $f(x)=x^5-4x+2$**: el mismo diagnóstico.** Eisenstein con $p=2$ ($2\mid 4$, $2\mid 2$, $4\nmid 2$). Las raíces son -1.51851, 0.508500, 1.24360 y $-0.116792\pm 1.43845\,i$: de nuevo tres reales y dos complejas. Mismo argumento, mismo veredicto: $G=S_5$, no resoluble por radicales.

(c) $f(x)=x^5-2$**: resoluble, y se ve la fórmula.** Eisenstein en 2, luego irreducible; pero aquí solo hay *una* raíz real ($\sqrt[5]{2}$) y cuatro complejas, así que el paso 4 no produce ninguna transposición y el argumento anterior no arranca. De hecho el cuerpo de descomposición es $\mathbb{Q}(\sqrt[5]{2},\zeta_5)$, de grado $5\cdot 4=20$, y

$$G\cong \mathbb{Z}/5\rtimes\mathbb{Z}/4,\qquad |G|=20,$$

el llamado grupo de Frobenius F_{20}. Es soluble: la cadena $1\trianglelefteq\mathbb{Z}/5\trianglelefteq F_{20}$ tiene cocientes $\mathbb{Z}/5$ y $\mathbb{Z}/4$, ambos abelianos. Y en efecto la fórmula es explícita: las raíces son $\sqrt[5]{2}\,\zeta_5^{k}$ con $k=0,1,2,3,4$. La solubilidad del grupo se traduce, literalmente, en la torre "primero adjunta ζ_5, después una raíz quinta".

(d) $f(x)=x^5-5x+12$**: resoluble aunque no lo parezca.** Este polinomio es irreducible sobre $\mathbb{Q}$ y tiene una sola raíz real. Su discriminante es $64\,000\,000=8000^2$, un cuadrado perfecto, de modo que $G\subseteq A_5$. Un cálculo del grupo da $|G|=10$, es decir $G\cong D_5$, el diédrico de orden 10, que es soluble ($1\trianglelefteq\mathbb{Z}/5\trianglelefteq D_5$ con cocientes $\mathbb{Z}/5$ y $\mathbb{Z}/2$). Por tanto *esta* quíntica sí se resuelve por radicales, pese a no tener ninguna forma especial a simple vista. Es la prueba de que "grado 5" no significa "imposible": significa "imposible en general".

Cuidado

El criterio de las tres raíces reales es *suficiente*, no necesario. El polinomio $x^5 - x - 1$ es irreducible y tiene una sola raíz real, así que el paso 4 no da transposición alguna; sin embargo su grupo también es S_5, de orden 120, y tampoco es resoluble. Cuando el conteo de raíces no cierra el caso hay que recurrir a otras herramientas: el discriminante (2869 para este polinomio, que no es cuadrado, luego $G \not\subseteq A_5$), la factorización módulo varios primos para leer los tipos de ciclo del Frobenius, o la resolvente sextica de Cayley. Nunca concluyas "es resoluble" solo porque el argumento rápido haya fallado.

Idea clave

Merece la pena separar tres afirmaciones que suelen confundirse. Primera: las raíces *existen* —eso lo garantiza el teorema fundamental del álgebra, y no tiene nada que ver con Galois—. Segunda: las raíces se pueden *calcular* con la precisión que se quiera —eso lo hace cualquier método numérico, y el apartado (a) lo hizo con seis cifras—. Tercera: las raíces se pueden *expresar* con radicales —eso es lo único que Galois prohíbe—. La imposibilidad es sobre un lenguaje, no sobre los números. Si se amplía el lenguaje, la quíntica vuelve a ser resoluble: Hermite y Kronecker la resolvieron con funciones elípticas modulares en 1858. Lo que Galois demostró es exactamente dónde está la frontera del lenguaje de los radicales, y esa frontera es la palabra "soluble" en teoría de grupos.

26.9. Construcciones con regla y compás

Idea clave

Intuición primero. La regla traza rectas y el compás traza circunferencias; cortar dos de esas figuras lleva, en coordenadas, a ecuaciones de grado uno o dos y nunca más. Cada paso de una construcción, por tanto, a lo sumo *duplica* el grado del cuerpo donde viven tus puntos. Empiezas en $\mathbb{Q}$ (grado 1) y multiplicas por 2 tantas veces como quieras: todo número construible tiene grado potencia de 2 sobre $\mathbb{Q}$.

Con esa sola frase caen dos milenios de intentos. Duplicar el cubo pide $\sqrt[3]{2}$, de grado 3, y 3 no es potencia de 2. Trisecar el ángulo pide $\cos 20°$, raíz de un cúbico irreducible: mismo motivo. Cuadrar el círculo pide $\sqrt{\pi}$, que ni siquiera es algebraico. Y el heptágono regular es imposible mientras que el de 17 lados sí se puede, porque $\varphi(7) = 6$ no es potencia de 2 y $\varphi(17) = 16$ sí.

Merece la pena detenerse en lo que acaba de pasar: preguntas de geometría con dibujo, planteadas por los griegos, resueltas contando *dimensiones de espacios vectoriales*. Ese cambio de terreno —de la figura al álgebra que la describe— es la lección más transferible de todo el capítulo.

Un real α es *construible* si, partiendo de un segmento unidad, se obtiene con regla (sin marcas) y compás. Álgebra de por medio: intersecar rectas y circunferencias solo resuelve ecuaciones de grado ≤ 2 sobre el cuerpo de lo ya construido, así que cada paso a lo sumo duplica el grado.

Teorema 26.77. *Todo número construible es algebraico sobre $\mathbb{Q}$ con $[\mathbb{Q}(\alpha):\mathbb{Q}]$ una potencia de 2. (El recíproco exacto involucra la clausura normal: α es construible $\iff$ el grupo de Galois de la clausura normal de $\mathbb{Q}(\alpha)$ tiene orden potencia de 2.)*

Idea clave

Qué dice esto en cristiano. Con regla y compás solo se pueden hacer dos cosas: cortar rectas con rectas y con circunferencias. Algebraicamente, eso significa resolver ecuaciones de grado uno o dos, es decir, extraer raíces cuadradas. Por tanto cada paso de una construcción duplica el grado, y el grado total de cualquier número construible tiene que ser una potencia de 2. Con esa sola frase caen de golpe tres problemas que resistieron dos mil años, y caen no por falta de ingenio, sino porque son imposibles.

Ejemplo 26.78 (Los tres problemas clásicos, liquidados)**. (a) Duplicar el cubo.** Exige construir $\sqrt[3]{2}$. Su polinomio mínimo es x^3-2, que es irreducible sobre $\mathbb{Q}$ por Eisenstein con $p=2$, luego $[\mathbb{Q}(\sqrt[3]{2}):\mathbb{Q}]=3$. Como 3 no es potencia de 2, es imposible.

(b) Trisecar un ángulo. Trisecar 60° exige construir $\cos 20^\circ$, que es raíz de $8t^3-6t-1=0$ (la identidad $\cos 3\theta = 4\cos^3\theta - 3\cos\theta$ con $\cos 60^\circ = \frac{1}{2}$). Ese polinomio es irreducible sobre $\mathbb{Q}$, luego el grado es 3 y otra vez es imposible. Nótese que *algunos* ángulos sí se trisecan, como el de 90°; lo imposible es un método general.

(c) Cuadrar el círculo. Exige construir $\sqrt{\pi}$, y por tanto π. Pero π ni siquiera es algebraico: Lindemann demostró en 1882 que es trascendente. Un número trascendente no tiene grado finito, así que la construcción es imposible por una razón todavía más radical.

(d) El criterio es necesario, no suficiente. Que el grado sea potencia de 2 no basta para ser construible: hace falta además que el cuerpo de descomposición tenga grado potencia de 2. Existen polinomios de grado 4 irreducibles cuyas raíces no son construibles, porque su grupo de Galois es S_4 o A_4 y no una torre de 2-extensiones.

Cuidado

[Con otras herramientas los tres problemas sí se resuelven] La imposibilidad es de la regla y el compás, no de la geometría. Con una regla marcada (construcciones *neusis*) Arquímedes ya trisecaba ángulos, y con papiroflexia se resuelven ecuaciones cúbicas, de modo que duplicar el cubo se hace doblando papel. Segundo aviso: los antiguos no fracasaron por torpeza; el enunciado correcto del problema necesitaba una teoría, la de cuerpos, que tardó veintidós siglos en aparecer. Wantzel cerró los dos primeros en 1837, con veintitrés años, y su trabajo pasó casi desapercibido.

Corolario 26.79 (Los tres problemas griegos)**.** Duplicar el cubo*: exige $\sqrt[3]{2}$, de grado 3: imposible.* Trisecar el ángulo general*: trisecar 60° exige $\cos 20^\circ$; de $\cos 3\theta = 4\cos^3\theta - 3\cos\theta$ con $\theta = 20^\circ$: $8x^3-6x-1=0$, irreducible sobre $\mathbb{Q}$ (sin raíces racionales), grado 3: imposible.* Cuadrar el círculo*: exige $\sqrt{\pi}$, trascendente (Lindemann): imposible con ganas.*

Teorema 26.80 (Gauss–Wantzel)**.** *El polígono regular de n lados es construible $\iff$ $n=$*

$2^k\, p_1 \cdots p_r$ *con los* p_i primos de Fermat *distintos* ($p = 2^{2^m} + 1$*: 3, 5, 17, 257, 65537). Razón: construir el* n*-ágono es construir* ζ_n*, y* $[\mathbb{Q}(\zeta_n) : \mathbb{Q}] = \varphi(n)$ *debe ser potencia de 2. Así, el heptágono* ($\varphi(7) = 6$) *no es construible, pero el de 17 lados sí* ($\varphi(17) = 16$)*: el descubrimiento que decidió a Gauss, a los 19 años, a dedicarse a la matemática.*

Idea clave

Qué dice esto en cristiano. Gauss–Wantzel resuelve por completo una pregunta que parecía inabordable: ¿qué polígonos regulares se pueden dibujar con regla y compás? La respuesta depende solo de la factorización de n: potencias de dos multiplicadas por primos de Fermat distintos, y nada más. La razón es que construir el n-gono equivale a construir ζ_n, cuyo grado sobre $\mathbb{Q}$ es $\varphi(n)$, y hay que exigir que $\varphi(n)$ sea potencia de 2. Esa condición sobre la función de Euler se traduce exactamente en la factorización del enunciado.

Ejemplo 26.81 (Qué polígonos se pueden y cuáles no)**. (a) Los primos de Fermat conocidos.** Son 3, 5, 17, 257 y 65 537, es decir $2^{2^k} + 1$ para $k = 0, 1, 2, 3, 4$. El siguiente candidato, $2^{32} + 1 = 4\,294\,967\,297$, *no* es primo: Euler encontró en 1732 que es $641 \times 6\,700\,417$, desmintiendo la conjetura de Fermat de que todos lo fueran. No se conoce ningún otro primo de Fermat.

(b) La lista de construibles hasta 30**.** Son $3, 4, 5, 6, 8, 10, 12, 15, 16, 17, 20, 24, 30$. Faltan el 7, el 9, el 11, el 13, el 14, el 18, el 19, el 21… El 9 falla porque 3 aparece al cuadrado, y la condición exige primos de Fermat *distintos* y con exponente uno.

(c) El heptadecágono de Gauss. Como $17 = 2^4 + 1$ es primo de Fermat, $\varphi(17) = 16 = 2^4$ es potencia de dos y el polígono de 17 lados es construible. Gauss lo descubrió en 1796, con diecinueve años, y fue lo que le decidió a dedicarse a las matemáticas; pidió que grabaran un heptadecágono en su tumba.

(d) Por qué el 7 **no.** Se tiene $\varphi(7) = 6 = 2 \cdot 3$, que no es potencia de 2 por el factor 3. Ese 3 es exactamente el mismo obstáculo que impide duplicar el cubo. Y el 9 falla porque $\varphi(9) = 6$ también.

Cuidado

[Construible el polígono no es lo mismo que construible el ángulo] El 15-gono es construible porque $15 = 3 \cdot 5$ son dos primos de Fermat distintos, y sin embargo el ángulo de $24°$ no se puede trisecar. Son preguntas diferentes. Segundo aviso: la condición es sobre n entero, así que multiplicar por potencias de 2 nunca estropea nada; bisecar ángulos siempre se puede, y por eso si el n-gono es construible también lo es el $2n$-gono. Y tercero: el teorema dice qué es posible, no da la construcción; la del 257-gono la publicó Richelot en 1832 y ocupa ochenta páginas.

Dos milenios de intentos con regla y compás, cerrados por un cómputo de dimensiones de espacios vectoriales.

26.10. Los períodos de Gauss y la construcción del heptadecágono

En 1796, con diecinueve años, Gauss demostró que el polígono regular de 17 lados es construible con regla y compás, y el hallazgo lo decidió a dedicarse a las matemáticas. Lo notable es que su argumento no es geométrico: es puro álgebra de la que hoy llamamos de Galois, cincuenta años antes de que Galois la escribiera. La idea consiste en agrupar las diecisiete raíces de la unidad en bloques cada vez más pequeños, de forma que cada bloque sea raíz de una ecuación de segundo grado con coeficientes en el bloque anterior. Cuatro ecuaciones de segundo grado encadenadas, y el ángulo queda construido.

Definición 26.82 (Períodos de Gauss)**.** Sea p primo, $\zeta = \mathrm{e}^{2\pi \mathrm{i}/p}$ y g una raíz primitiva módulo p, de modo que $\mathrm{Gal}(\mathbb{Q}(\zeta)/\mathbb{Q}) \cong (\mathbb{Z}/p)^\times$ es cíclico de orden $p-1$ generado por $\sigma\colon \zeta \mapsto \zeta^g$. Si $ef = p-1$, el *período de Gauss* de f términos correspondiente a la clase j es

$$\eta_j^{(f)} = \sum_{k \equiv j \pmod{e}} \zeta^{g^k}, \qquad j = 0, 1, \dots, e-1.$$

Es decir: se recorren las potencias de g y se reparten cíclicamente en e montones de f sumandos cada uno. Cada $\eta_j^{(f)}$ queda fijo por el subgrupo $\langle \sigma^e \rangle$, luego vive en el único subcuerpo de grado e sobre $\mathbb{Q}$.

Ejemplo 26.83 (Los períodos de Gauss para $p = 7$ y $p = 17$)**. (a) El caso $p = 7$.** El grupo $(\mathbb{Z}/7)^*$ es cíclico de orden 6 generado por 3, con potencias $3, 2, 6, 4, 5, 1$. Agrupando de dos en dos se obtienen los tres períodos $\eta_0 = \zeta + \zeta^6$, $\eta_1 = \zeta^3 + \zeta^4$, $\eta_2 = \zeta^2 + \zeta^5$, que satisfacen $\eta^3 + \eta^2 - 2\eta - 1 = 0$. Cada período es $2\cos(2\pi k/7)$ para un k.

(b) La torre. $\mathbb{Q}(\zeta_7)$ tiene grado 6 sobre $\mathbb{Q}$, y como $6 = 2 \cdot 3$, el grupo cíclico $\mathbb{Z}/6$ tiene una única cadena de subgrupos: $\mathbb{Q} \subset \mathbb{Q}(\sqrt{-7}) \subset \mathbb{Q}(\zeta_7)$ y $\mathbb{Q} \subset \mathbb{Q}(\eta) \subset \mathbb{Q}(\zeta_7)$. Los períodos generan los cuerpos intermedios.

(c) El heptadecágono. Para $p = 17$ se tiene $17 - 1 = 16 = 2^4$, luego la torre es de cuatro extensiones cuadráticas seguidas y todo se resuelve con raíces cuadradas: el polígono regular de 17 lados es construible con regla y compás. Gauss lo descubrió a los diecinueve años y pidió que grabaran uno en su tumba.

(d) Por qué $p = 7$ no sirve. Aquí $7 - 1 = 6 = 2 \cdot 3$, y ese factor 3 obliga a una extensión cúbica que no se alcanza con raíces cuadradas. El heptágono regular no es construible. La condición de Gauss–Wantzel es exactamente que $p - 1$ sea potencia de dos, es decir, que p sea un primo de Fermat: $3, 5, 17, 257, 65537$.

Idea clave

La correspondencia de Galois dice que los subcuerpos de $\mathbb{Q}(\zeta_{17})$ están en biyección con los subgrupos de $\mathbb{Z}/16$, y como $16 = 2^4$ hay una torre de subgrupos de índice 2 en cada peldaño:

$$1 \subset \mathbb{Z}/2 \subset \mathbb{Z}/4 \subset \mathbb{Z}/8 \subset \mathbb{Z}/16.$$

Los períodos son simplemente los generadores explícitos de esa torre leída al revés: los dos períodos de ocho términos generan el cuerpo de grado 2, los cuatro de cuatro términos el de grado 4, y así hasta llegar a ζ mismo. Que cada peldaño tenga grado 2 es exactamente lo que permite resolverlo con una raíz cuadrada, y una raíz cuadrada es lo único que la regla y el compás saben hacer. Toda la construcción del 17-gono es esa observación puesta en números.

Regla: construir el polígono regular de p lados con períodos de Gauss

El método funciona siempre que $p-1$ sea potencia de 2, es decir para los primos de Fermat.

1. **Compruébese que $\varphi(p) = p-1$ es potencia de 2.** Si no lo es, la construcción es imposible por el criterio de Gauss–Wantzel y no hay nada que hacer. Para $p = 17$ se tiene $16 = 2^4$: cuatro peldaños.

2. **Búsquese una raíz primitiva g módulo p y escríbase la lista de potencias.** Para $p = 17$ la menor es $g = 3$, y sus potencias $3^0, 3^1, \ldots, 3^{15}$ son $1, 3, 9, 10, 13, 5, 15, 11, 16, 14, 8, 7, 4, 12, 2, 6$: los dieciséis restos no nulos, cada uno una vez.

3. **Pártase la lista en dos montones alternos.** Los de índice par dan η_0, los de índice impar dan η_1. Los índices pares corresponden a los restos cuadráticos módulo p.

4. **Calcúlense $\eta_0 + \eta_1$ y $\eta_0\eta_1$, que son racionales.** La suma es -1 siempre, porque suma todas las raíces salvo la unidad. El producto se calcula desarrollando y agrupando; para $p = 17$ vale -4. Resuélvase la cuadrática resultante.

5. **Repítase el reparto dentro de cada montón.** Cada η_j se parte en dos períodos con la mitad de términos; su suma es η_j y su producto vuelve a ser expresable en el nivel anterior. Otra cuadrática.

6. **Continúese hasta llegar a períodos de dos términos.** Un período de dos términos es $\zeta^a + \zeta^{-a} = 2\cos(2\pi a/p)$: el coseno buscado, ya expresado con raíces cuadradas encadenadas.

7. **Tradúzcase a construcción geométrica.** Cada raíz cuadrada se materializa con el teorema de la altura sobre una semicircunferencia, y cada operación racional con semejanza de triángulos. Compruébese el resultado numéricamente antes de dibujar: un signo equivocado en el paso 4 produce el coseno del ángulo equivocado.

Ejemplo 26.84 (El heptadecágono, nivel por nivel)**. (a) Primer nivel: dos períodos de ocho.** Con $g = 3$ y la lista del paso 2, los índices pares dan $\{1, 9, 13, 15, 16, 8, 4, 2\}$, es decir, ordenados, $\{1, 2, 4, 8, 9, 13, 15, 16\}$ — los restos cuadráticos módulo 17 — y los impares

$\{3,5,6,7,10,11,12,14\}$. Los períodos

$$\eta_0 = \sum_{a\in\{1,2,4,8,9,13,15,16\}} \zeta^a, \qquad \eta_1 = \sum_{a\in\{3,5,6,7,10,11,12,14\}} \zeta^a$$

cumplen $\eta_0+\eta_1=-1$ y $\eta_0\eta_1=-4$, luego son las raíces de $t^2+t-4=0$:

$$\eta_0 = \frac{-1+\sqrt{17}}{2} = 1{,}5615528, \qquad \eta_1 = \frac{-1-\sqrt{17}}{2} = -2{,}5615528.$$

Ya ha aparecido el $\sqrt{17}$, y con él el cuerpo cuadrático $\mathbb{Q}(\sqrt{17})$, que es el subcuerpo de grado 2.

(b) Segundo nivel: cuatro períodos de cuatro. Repartiendo ahora cada montón en dos, se obtienen

$$q_0 = \zeta+\zeta^4+\zeta^{13}+\zeta^{16} = 2{,}0494812, \qquad q_2 = \zeta^2+\zeta^8+\zeta^9+\zeta^{15} = -0{,}4879284,$$

que suman η_0 y multiplican -1; y análogamente $q_1 = 0{,}3441507$ y $q_3 = -2{,}9057035$, que suman η_1 y multiplican -1. Dos cuadráticas más: $t^2-\eta_0 t-1=0$ y $t^2-\eta_1 t-1=0$. El cuerpo generado es de grado 4 sobre $\mathbb{Q}$.

(c) Tercer nivel: los períodos de dos y el coseno. El período $o_0 = \zeta+\zeta^{16} = 2\cos(2\pi/17) = 1{,}8649445$ y su compañero $o_4 = \zeta^4+\zeta^{13} = 0{,}1845367$ cumplen $o_0+o_4=q_0$ y $o_0o_4=q_1$. Una última cuadrática, $t^2-q_0t+q_1=0$, y $\cos(2\pi/17)$ queda despejado. Agrupando las cuatro raíces cuadradas encadenadas se llega a la fórmula que Gauss publicó:

$$16\cos\frac{2\pi}{17} = -1+\sqrt{17}+\sqrt{34-2\sqrt{17}}+2\sqrt{17+3\sqrt{17}-\sqrt{34-2\sqrt{17}}-2\sqrt{34+2\sqrt{17}}}.$$

Numéricamente el miembro derecho vale 14,9195557, y dividido entre 16 da 0,9324722294, que es en efecto $\cos(2\pi/17)$ con doce cifras exactas.

(d) Por qué hacía falta que 16 **fuese potencia de** 2**.** En cada nivel se pasó de e períodos a $2e$, es decir, se subió un peldaño de grado 2 en la torre. Si $p-1$ tuviese un factor primo impar ℓ, en algún momento habría un peldaño de grado ℓ, y ese peldaño exigiría resolver una ecuación de grado ℓ que en general no se reduce a raíces cuadradas. Por eso el heptágono ($\varphi(7)=6=2\cdot 3$) y el eneágono ($\varphi(9)=6$) no son construibles, mientras que el pentágono ($\varphi(5)=4$) y el heptadecágono sí lo son.

Ejemplo 26.85 (Qué polígonos son construibles)**. (a) El criterio completo.** El polígono regular de n lados es construible con regla y compás si y solo si $n = 2^k p_1\cdots p_r$ con los p_i primos de Fermat distintos, es decir de la forma $2^{2^m}+1$. Los únicos conocidos son 3, 5, 17, 257 y 65537, y no se sabe si hay más. La razón es que $\varphi(n)$ debe ser potencia de 2, y φ es multiplicativa.

(b) La lista pequeña. Son construibles

$$\begin{aligned} n = {} & 3,4,5,6,8,10,12,15,16,17,20,24,30,32,34,40,48,51,60,64,\\ & 68,80,85,96,102,120,128,136,160,170,192,204,240,255,256,257, \end{aligned}$$

y por debajo de 300 sólo falta 272. No lo son 7, 9, 11, 13, 14, 18, 19, 21, 22, 23, 25, 26, 27, 28, 29. Compruébese: $\varphi(15)=8$ sí, $\varphi(18)=6$ no, $\varphi(51)=32$ sí, $\varphi(255)=128$ sí.

(c) El 257-gono y el 65537-gono. Ambos son construibles, y ambos han sido construidos explícitamente: Richelot publicó el 257-gono en 1832 en un trabajo de ciento noventa y cuatro páginas, y Hermes dedicó diez años al 65537-gono, dejando un manuscrito que se conserva en Gotinga. Con períodos de Gauss el procedimiento es idéntico al del 17-gono, solo que con ocho y con dieciséis niveles de cuadráticas respectivamente.

(d) Los tres problemas clásicos, de un plumazo. La duplicación del cubo pide $\sqrt[3]{2}$, de grado 3 sobre $\mathbb{Q}$, y 3 no es potencia de 2. La trisección del ángulo general lleva a la cúbica $4t^3 - 3t = \cos\theta$, irreducible por ejemplo para $\theta = 60°$. La cuadratura del círculo pide $\sqrt{\pi}$, y π ni siquiera es algebraico (Lindemann, 1882). Los tres problemas resistieron dos mil años y cayeron todos con el mismo argumento: todo número construible tiene grado potencia de 2 sobre $\mathbb{Q}$.

Cuidado

El criterio de Gauss–Wantzel es una condición sobre el *grado*, y el error más común es aplicarlo al revés. Que un número tenga grado potencia de 2 sobre $\mathbb{Q}$ no basta para que sea construible: hace falta además que su clausura normal tenga grado potencia de 2, o equivalentemente que el grupo de Galois sea un 2-grupo. Existen polinomios irreducibles de grado 4 con grupo S_4 o A_4, cuyas raíces no son construibles pese a tener grado 4. Para las raíces de la unidad el problema no se plantea porque la extensión ciclotómica ya es abeliana, y ahí sí grado potencia de 2 equivale a construibilidad.

Pausa para pensar

Los períodos de Gauss no son una curiosidad del 17-gono: son el instrumento con el que se estudian las extensiones abelianas de $\mathbb{Q}$ en general. El período de f términos genera exactamente el subcuerpo de índice f, y toda la aritmética del cuerpo ciclotómico se refleja en cómo se combinan. Cuando $e = 2$ y $f = (p-1)/2$, la diferencia $\eta_0 - \eta_1$ es esencialmente la *suma de Gauss* que aparece en la sección siguiente, y su cuadrado vale $\pm p$: de ahí sale $\sqrt{\pm p}$ y, con él, la ley de reciprocidad cuadrática. Merece la pena que el lector calcule los períodos de $p = 5$ y compruebe que reproduce la fórmula clásica $\cos(2\pi/5) = (\sqrt{5}-1)/4$ del pentágono regular.

26.11. Fundamentos: la demostración de la correspondencia

Esta sección demuestra el teorema fundamental. La cadena lógica tiene tres eslabones: independencia de caracteres, el teorema de Artin sobre cuerpos fijos, y el ensamblaje final.

26.11.1. Independencia de caracteres

Teorema 26.86 (Dedekind)**.** *Automorfismos distintos $\sigma_1, \dots, \sigma_n$ de un cuerpo L son* linealmente independientes*: si $c_1\sigma_1(x) + \cdots + c_n\sigma_n(x) = 0$ para todo $x \in L$, entonces todos los $c_i = 0$.*

Demostración. Inducción en n; para $n = 1$, evaluar en $x = 1$. Supongamos una relación mínima $\sum_{i\le n} c_i\sigma_i = 0$ con todos los $c_i \neq 0$ y $n \geq 2$. Como $\sigma_1 \neq \sigma_n$, existe $a \in L$ con $\sigma_1(a) \neq \sigma_n(a)$. Evaluamos la relación en ax:

$$\sum_i c_i\sigma_i(a)\sigma_i(x) = 0 \quad \text{para todo } x,$$

y le restamos $\sigma_n(a)$ veces la relación original:

$$\sum_{i<n} c_i(\sigma_i(a) - \sigma_n(a))\sigma_i(x) = 0.$$

Es una relación más corta con el coeficiente de σ_1 no nulo ($c_1(\sigma_1(a) - \sigma_n(a)) \neq 0$): contradice la minimalidad. □

Ejemplo 26.87 (Independencia de caracteres y para qué sirve)**. (a) Un caso concreto.** En $\mathbb{C}/\mathbb{R}$, la identidad y la conjugación son dos automorfismos distintos. Que sean linealmente independientes significa que no existen $a, b \in \mathbb{C}$ no ambos nulos con $az + b\bar{z} = 0$ para todo z. En efecto, tomando $z = 1$ sale $a + b = 0$ y tomando $z = \mathrm{i}$ sale $a - b = 0$, luego $a = b = 0$.

(b) La consecuencia sobre grados. De aquí se deduce que $|\operatorname{Gal}(L/K)| \le [L : K]$, con igualdad si y sólo si la extensión es de Galois. Sin este lema la correspondencia de Galois no arranca: es literalmente el primer ladrillo.

(c) En cuerpos finitos. Los automorfismos de $\mathbb{F}_{p^n}$ son $x \mapsto x^{p^k}$ para $k = 0, \dots, n-1$: exactamente n, y el grupo es cíclico generado por el Frobenius. La independencia lineal dice que ninguna combinación no trivial de esas n funciones se anula idénticamente, hecho nada evidente porque el cuerpo es finito y las funciones son polinómicas.

(d) La versión de Artin. El mismo argumento sirve para caracteres de grupos: los homomorfismos distintos de un grupo en el grupo multiplicativo de un cuerpo son linealmente independientes. De ahí sale la ortogonalidad de caracteres de la teoría de representaciones, y con ella la transformada discreta de Fourier.

26.11.2. El teorema de Artin

Teorema 26.88 (Artin)**.** *Sea G un grupo finito de automorfismos de L y $F = L^G$ su cuerpo fijo. Entonces*

$$[L : F] = |G|.$$

Demostración de la desigualdad sustancial, $[L : F] \le |G|$. Sea $|G| = n$ con $G = \{\sigma_1, \dots, \sigma_n\}$; veamos que cualesquiera $n + 1$ elementos $\alpha_1, \dots, \alpha_{n+1} \in L$ son dependientes sobre F. El sistema lineal homogéneo

$$\sum_{j=1}^{n+1} x_j\,\sigma_i(\alpha_j) = 0 \qquad (i = 1, \dots, n)$$

tiene n ecuaciones y $n + 1$ incógnitas: por Rouché–Frobenius (¡capítulo 7!) hay solución no trivial en L. Elegimos una con el *menor número* de entradas no nulas, normalizada para que una de ellas sea 1, digamos $x_1 = 1$. Afirmamos que todas las x_j están en F: si no, algún $\sigma \in G$

mueve alguna, y aplicando σ a todo el sistema (que solo permuta las ecuaciones, pues $\sigma\sigma_i$ recorre G) obtenemos otra solución $(\sigma(x_j))$; la diferencia de ambas es una solución con *menos* entradas no nulas (la posición 1 se cancela: $\sigma(1) = 1$) y no nula: contradice la minimalidad. Con todas las $x_j \in F$, la ecuación correspondiente a $\sigma_i = \mathrm{id}$ dice $\sum x_j\alpha_j = 0$: dependencia sobre F. La otra mitad, $[L : F] \geq |G|$, sale de la independencia de caracteres de Dedekind con un argumento dual. □

Idea clave

Qué dice esto en cristiano. El teorema de Artin es la mitad difícil de la correspondencia de Galois, y dice algo sorprendente: si se toma cualquier grupo finito de automorfismos de un cuerpo y se mira qué elementos deja quietos, el resultado es un subcuerpo cuya extensión tiene grado *exactamente* el orden del grupo. Ni más ni menos. Es lo que garantiza que el diccionario entre subgrupos y subcuerpos sea una biyección perfecta y no una correspondencia parcial llena de agujeros.

Ejemplo 26.89 (Contar grados contando automorfismos)**. (a) El caso más simple.** Sea $L = \mathbb{C}$ y $G = \{\mathrm{id}, \text{conjugación}\}$, de orden 2. El cuerpo fijo es $\mathbb{R}$, porque $\overline{z} = z$ equivale a que z sea real. Y en efecto $[\mathbb{C} : \mathbb{R}] = 2$, igual al orden de G.

(b) Una extensión cuadrática. Sea $L = \mathbb{Q}(\sqrt{2})$ y G generado por $\sigma(a+b\sqrt{2}) = a-b\sqrt{2}$, de orden 2. El cuerpo fijo son los elementos con $b = 0$, es decir $\mathbb{Q}$. Y $[\mathbb{Q}(\sqrt{2}) : \mathbb{Q}] = 2$. Otra vez coincide.

(c) Un caso de grado seis. Sea L el cuerpo de descomposición de x^3-2, que es $\mathbb{Q}(\sqrt[3]{2}, \zeta_3)$, con $[L : \mathbb{Q}] = 6$. Su grupo de Galois es S_3, de orden 6: coinciden. Los subgrupos de S_3 tienen órdenes $1, 2, 3, 6$, y en efecto hay subcuerpos intermedios de grados $6, 3, 2, 1$ sobre $\mathbb{Q}$, con los índices invertidos.

(d) El diccionario, con los índices al revés. Al subgrupo A_3 de orden 3 le corresponde el cuerpo fijo $\mathbb{Q}(\zeta_3)$, de grado $\frac{6}{3} = 2$ sobre $\mathbb{Q}$. A los tres subgrupos de orden 2 les corresponden los tres cuerpos $\mathbb{Q}(\sqrt[3]{2})$, $\mathbb{Q}(\omega\sqrt[3]{2})$ y $\mathbb{Q}(\omega^2\sqrt[3]{2})$, cada uno de grado $\frac{6}{2} = 3$. Subgrupo grande, cuerpo pequeño.

Cuidado

[La correspondencia invierte el orden y solo los normales dan extensiones de Galois] Cuanto mayor es el subgrupo, menor es el cuerpo fijo: es un diccionario que da la vuelta a las inclusiones. Segundo aviso: un subcuerpo intermedio F da una extensión F/K que es de Galois *solo* si el subgrupo correspondiente es normal. En el ejemplo (d), $\mathbb{Q}(\sqrt[3]{2})/\mathbb{Q}$ no es de Galois, porque el subgrupo de orden 2 no es normal en S_3, y en efecto ese cuerpo contiene una raíz de $x^3 - 2$ pero no las otras dos. En cambio $\mathbb{Q}(\zeta_3)/\mathbb{Q}$ sí lo es, porque A_3 es normal.

26.11.3. Ensamblaje del teorema fundamental

Demostración del Teorema 26.54. Sea L/K de Galois con grupo G (así $|G| = [L : K]$, y por Artin el cuerpo fijo de G es exactamente K).

Las dos aplicaciones son inversas. Dado $H \leq G$, sea $F = L^H$: por Artin, $[L : F] = |H|$, y claramente $H \subseteq \mathrm{Gal}(L/F)$ con $|\mathrm{Gal}(L/F)| \leq [L : F] = |H|$: luego $\mathrm{Gal}(L/L^H) = H$ exactamente. Recíprocamente, dado un cuerpo intermedio F, la extensión L/F es de Galois (un cuerpo de descomposición sobre K lo sigue siendo sobre el cuerpo más grande F), así que $|\mathrm{Gal}(L/F)| = [L : F]$ y su cuerpo fijo es F por Artin de nuevo: $L^{\mathrm{Gal}(L/F)} = F$. Las composiciones en ambos órdenes son la identidad, y que invierten inclusiones es inmediato de las definiciones.

Grados e índices. $[L : F] = |\mathrm{Gal}(L/F)|$ acaba de probarse; $[F : K] = [L : K]/[L : F] = |G|/|\mathrm{Gal}(L/F)| = [G : \mathrm{Gal}(L/F)]$ por la torre.

Normalidad. Para $\sigma \in G$, el subgrupo que fija $\sigma(F)$ es $\sigma\,\mathrm{Gal}(L/F)\sigma^{-1}$ (verificación directa). Entonces: $\mathrm{Gal}(L/F)$ es normal en $G \iff \sigma(F) = F$ para todo $\sigma \in G \iff$ toda raíz en L de un polinomio mínimo de un elemento de F queda en F (las imágenes $\sigma(\alpha)$ recorren esas raíces) $\iff F/K$ es normal. En ese caso, restringir automorfismos ($\sigma \mapsto \sigma|_F$) es un homomorfismo $G \to \mathrm{Gal}(F/K)$ sobreyectivo con núcleo $\mathrm{Gal}(L/F)$, y el primer teorema de isomorfía (¡capítulo 25!) da $\mathrm{Gal}(F/K) \cong G/\mathrm{Gal}(L/F)$. □ □

Observación 26.90. Mirando la demostración se ve el reparto de papeles: *álgebra lineal* (sistemas y dimensiones) para Artin, *teoría de grupos* (isomorfía, normalidad) para el diccionario, y *polinomios* (raíces y factorización) como pegamento. La teoría de Galois no es una rama más: es el punto donde las tres convergen.

26.11.4. El teorema del elemento primitivo, demostrado

Demostración (característica 0). Basta el caso $L = K(\alpha, \beta)$ (e iterar). Consideremos los elementos $\gamma_c = \alpha + c\beta$ con $c \in K$. Hay infinitos c (característica 0) y solo finitas parejas (raíz de m_α, raíz de m_β); se puede elegir c tal que los valores $\alpha_i + c\beta_j$ sean todos distintos para las distintas raíces α_i, β_j (evitar finitas ecuaciones lineales en c). Con tal c, sea $\gamma = \gamma_c$: el polinomio $h(x) = m_\alpha(\gamma - cx) \in K(\gamma)[x]$ tiene a β como raíz, y *ninguna otra* raíz común con m_β (si $m_\beta(\beta_j) = 0$ y $h(\beta_j) = 0$ con $\beta_j \neq \beta$, entonces $\gamma - c\beta_j$ sería alguna α_i, es decir $\alpha + c\beta = \alpha_i + c\beta_j$: excluido por la elección de c). Entonces $\gcd(h, m_\beta) = x - \beta$ en $K(\gamma)[x]$ —el mcd se calcula con Euclides *dentro* de $K(\gamma)$—, luego $\beta \in K(\gamma)$, y $\alpha = \gamma - c\beta$ también: $K(\alpha, \beta) = K(\gamma)$. □

26.11.5. El teorema fundamental del álgebra, vía Galois

Como espectacular aplicación cruzada, demostremos con teoría de grupos el teorema que el análisis suele atribuirse:

Teorema 26.91. $\mathbb{C}$ *es algebraicamente cerrado.*

Demostración. Solo dos datos analíticos, ambos del capítulo 6: (a) todo polinomio real de grado *impar* tiene raíz real (Bolzano: los límites en $\pm\infty$ tienen signos opuestos), y (b) todo complejo tiene raíz cuadrada compleja (fórmula explícita con módulos y argumentos). Traducción algebraica: (a) $\mathbb{R}$ no tiene extensiones propias de grado impar (el polinomio mínimo de un elemento primitivo sería impar e irreducible, imposible con raíz real); (b) $\mathbb{C}$ no tiene extensiones de grado 2 (la cuadrática se resuelve).

Sea ahora $L/\mathbb{C}$ una extensión finita; agrandándola, podemos suponer $L/\mathbb{R}$ de Galois, con grupo G de orden $2^a m$ (m impar). Un 2-Sylow $H \leq G$ (¡capítulo 25!) tiene índice m impar: su cuerpo fijo L^H tiene grado m impar sobre $\mathbb{R}$, y por (a), $m = 1$: G es un 2-grupo. Entonces $\mathrm{Gal}(L/\mathbb{C})$ es un 2-grupo; si no fuera trivial, tendría (por la ecuación de clases y la teoría de p-grupos) un subgrupo de índice 2, cuyo cuerpo fijo sería una extensión de grado 2 de $\mathbb{C}$: imposible por (b). Luego $L = \mathbb{C}$. □ □

Idea clave

Qué dice esto en cristiano. El teorema fundamental del álgebra dice que $\mathbb{C}$ es el final del camino: al buscar raíces de polinomios complejos nunca hace falta salir de $\mathbb{C}$. Lo notable es que este enunciado, que parece puramente algebraico, no se puede demostrar solo con álgebra: en algún punto hay que usar una propiedad analítica de $\mathbb{R}$, típicamente el teorema del valor intermedio o la compacidad. La demostración vía Galois lo reduce a lo mínimo: basta con que todo polinomio real de grado impar tenga raíz real, y todo lo demás es teoría de grupos.

Ejemplo 26.92 (Dónde entra el análisis y dónde el álgebra)**. (a) El único ingrediente analítico.** Todo polinomio real de grado impar tiene una raíz real, porque sus dos brazos van a lados opuestos del eje y se aplica Bolzano. Por ejemplo $x^5 + x + 1$ vale -1 en $x = -1$ y 1 en $x = 0$, luego tiene raíz entre ambos: es $-0{,}7548777$.

(b) El segundo ingrediente, trivial. Todo número complejo tiene raíz cuadrada compleja, lo cual se ve en forma polar: la raíz de $r\,\mathrm{e}^{\mathrm{i}\theta}$ es $\sqrt{r}\,\mathrm{e}^{\mathrm{i}\theta/2}$. Por tanto $\mathbb{C}$ no admite extensiones de grado 2.

(c) El argumento de grupos. Si $\mathbb{C}$ tuviera una extensión propia, su grupo de Galois sobre $\mathbb{R}$ tendría orden par; por Sylow habría un 2-subgrupo, y el cuerpo fijo correspondiente tendría grado impar sobre $\mathbb{R}$, lo que por (a) obliga a que ese grado sea 1. Entonces el grupo sería un 2-grupo, que tiene subgrupos de todos los índices, y aparecería una extensión de grado 2 de $\mathbb{C}$, imposible por (b).

(d) Consecuencias inmediatas. Todo polinomio complejo de grado n tiene exactamente n raíces contadas con multiplicidad, y todo polinomio real factoriza sobre $\mathbb{R}$ en factores de grado 1 y 2. Por ejemplo $x^4 + 1$ es irreducible sobre $\mathbb{Q}$ y sobre $\mathbb{R}$ se parte como $(x^2 + \sqrt{2}x + 1)(x^2 - \sqrt{2}x + 1)$.

Cuidado

[El nombre es engañoso: no es un teorema de álgebra pura] Ninguna demostración es puramente algebraica, y no puede serlo: el cuerpo de los números algebraicos también es algebraicamente cerrado, así que la propiedad no distingue a $\mathbb{C}$ por razones algebraicas. El ingrediente analítico es imprescindible. Segundo aviso: el teorema afirma que las raíces *existen*, no que se puedan calcular ni escribir con radicales; Abel–Ruffini demuestra justamente lo contrario a partir de grado 5. Y tercero: Gauss dio cuatro demostraciones a lo largo de su vida y la primera, de 1799, tenía un hueco topológico que no se cerró del todo hasta el siglo XX.

Sylow y la ecuación de clases demostrando que los polinomios complejos tienen raíces: pocas veces se ve tan claro que las fronteras entre "análisis" y "álgebra" son administrativas.

26.11.6. El discriminante y el cálculo de grupos de Galois

¿Cómo se calcula, en la práctica, el grupo de Galois de un polinomio concreto? La primera herramienta es el *discriminante*

$$\Delta = \prod_{i<j} (\alpha_i - \alpha_j)^2,$$

simétrico en las raíces, luego expresable en los coeficientes (para la cúbica $x^3 + px + q$: $\Delta = -4p^3 - 27q^2$; para la cuadrática, el $b^2 - 4ac$ de siempre).

Teorema 26.93. *Sea f irreducible separable con grupo $G \leq S_n$. Entonces $\sqrt{\Delta} = \prod_{i<j}(\alpha_i - \alpha_j)$ es fijado exactamente por las permutaciones* pares *de las raíces, y por tanto:*

$$G \subseteq A_n \iff \Delta \text{ es un cuadrado en } K.$$

Idea clave

Qué dice esto en cristiano. El discriminante es un detector de paridad. Su raíz cuadrada, $\sqrt{\Delta} = \prod_{i<j}(\alpha_i - \alpha_j)$, cambia de signo al intercambiar dos raíces, es decir, al aplicar una permutación impar. Por tanto queda fija exactamente por las permutaciones pares. La conclusión es un criterio de una línea: el grupo de Galois está contenido en A_n si y solo si Δ es un cuadrado perfecto en el cuerpo base. Con un único cálculo se corta a la mitad la lista de grupos posibles.

Ejemplo 26.94 (Distinguir grupos con el discriminante)**. (a) Grado tres, caso S_3.** Para $x^3 - 2$ el discriminante vale -108, que no es cuadrado en $\mathbb{Q}$. Luego el grupo no está en A_3 y, siendo transitivo de grado 3, tiene que ser S_3 entero, de orden 6. Coincide con que el cuerpo de descomposición $\mathbb{Q}(\sqrt[3]{2}, \zeta_3)$ tenga grado 6.

(b) Grado tres, caso cíclico. Para $x^3 - 3x - 1$ el discriminante es $81 = 9^2$, un cuadrado perfecto. El grupo está dentro de $A_3 = \mathbb{Z}_3$, y como el polinomio es irreducible el grupo es transitivo, luego es $\mathbb{Z}_3$ exactamente. Su cuerpo de descomposición tiene grado 3 y es una extensión cíclica de $\mathbb{Q}$: las tres raíces se expresan racionalmente unas en función de otras.

(c) La cuenta en grado dos. Para $x^2 - 5$ el discriminante es 20, que no es cuadrado en $\mathbb{Q}$, y el grupo es $\mathbb{Z}_2$. Aquí A_2 es trivial, así que el criterio distingue simplemente entre que el polinomio tenga o no raíces racionales.

(d) Para qué sirve en la práctica. En grado 4 los grupos transitivos posibles son cinco: $\mathbb{Z}_4$, V_4, D_4, A_4 y S_4. El discriminante separa de un golpe $\{V_4, A_4\}$ (cuadrado) de $\{\mathbb{Z}_4, D_4, S_4\}$ (no cuadrado). Combinado con la resolvente cúbica, eso basta para determinar el grupo por completo.

Cuidado

[Cuadrado *en el cuerpo base*, no en $\mathbb{C}$] En $\mathbb{C}$ todo número es un cuadrado, así que la condición sería vacía. Lo que se pregunta es si Δ es cuadrado en K, típicamente $\mathbb{Q}$. Segundo aviso: el criterio solo vale si el polinomio es separable, es decir, si $\Delta \neq 0$; un discriminante nulo indica raíces repetidas y ahí no hay nada que decidir. Y tercero: el criterio dice si el grupo está *dentro* de A_n, no cuál es exactamente; para identificarlo hacen falta más datos, como la factorización módulo varios primos que da el teorema de Dedekind.

Corolario 26.95 (Cúbicas irreducibles sobre $\mathbb{Q}$)**.** *$G \cong A_3 = \mathbb{Z}_3$ si Δ es un cuadrado racional; $G \cong S_3$ si no. Ejemplos: $x^3 - 3x + 1$ tiene $\Delta = 81 = 9^2$: grupo $\mathbb{Z}_3$ (sus tres raíces reales generan cada una a las otras dos: es el "caso cíclico"); $x^3 - 2$ tiene $\Delta = -108$, no cuadrado: grupo S_3, como calculamos a mano.*

Para la cuártica, el papel auxiliar lo juega la *resolvente cúbica*: sus raíces son las combinaciones $\alpha_1\alpha_2 + \alpha_3\alpha_4$ y análogas, permutadas por S_4 a través del cociente $S_4/V \cong S_3$; según factorice la resolvente y sea o no cuadrado el discriminante, el grupo es S_4, A_4, D_4, V o $\mathbb{Z}_4$. Este árbol de decisión —completamente algorítmico— es lo que implementan los sistemas de álgebra computacional para calcular grupos de Galois.

26.11.7. Extensiones ciclotómicas y de Kummer: los peldaños de las torres

Los dos tipos de extensión que arman toda torre de radicales, con su teoría exacta:

Teorema 26.96 (Extensiones ciclotómicas)**.** *$\Phi_n(x)$ (el polinomio mónico cuyas raíces son las raíces* primitivas *n-ésimas de la unidad, de grado $\varphi(n)$) es irreducible sobre $\mathbb{Q}$, y por tanto $\mathrm{Gal}(\mathbb{Q}(\zeta_n)/\mathbb{Q}) \cong (\mathbb{Z}_n)^\times$ exactamente. (La irreducibilidad para n primo es el Eisenstein desplazado del capítulo 25; el caso general usa reducción módulo primos $p \nmid n$ y el Frobenius.)*

Idea clave

Qué dice esto en cristiano. Las extensiones ciclotómicas son las más amables de todas: adjuntar una raíz n-ésima de la unidad produce siempre una extensión de Galois cuyo grupo es $(\mathbb{Z}/n)^*$, un grupo abeliano perfectamente conocido. El grado es $\varphi(n)$. Esto convierte cualquier pregunta sobre $\mathbb{Q}(\zeta_n)$ en una pregunta de aritmética modular elemental, y es la razón de que estas extensiones sean el laboratorio donde se prueba toda la teoría.

Ejemplo 26.97 (Los polinomios ciclotómicos y su grupo)**. (a) Los primeros.** $\Phi_1 = x - 1$, $\Phi_2 = x + 1$, $\Phi_3 = x^2 + x + 1$, $\Phi_4 = x^2 + 1$, $\Phi_5 = x^4 + x^3 + x^2 + x + 1$, $\Phi_6 = x^2 - x + 1$, $\Phi_8 = x^4 + 1$ y $\Phi_{12} = x^4 - x^2 + 1$. En cada caso el grado coincide con $\varphi(n)$.

(b) La factorización de $x^n - 1$. Se tiene $x^n - 1 = \prod_{d|n} \Phi_d$. Para $n = 6$: $x^6 - 1 = (x - 1)(x + 1)(x^2 + x + 1)(x^2 - x + 1)$, con grados $1 + 1 + 2 + 2 = 6$, que es exactamente $\sum_{d|6} \varphi(d) = 1 + 1 + 2 + 2$.

(c) Los grupos. Para $n=5$, el grupo es $(\mathbb{Z}/5)^*$, cíclico de orden 4; por tanto $\mathbb{Q}(\zeta_5)$ tiene un único subcuerpo intermedio, de grado 2, que resulta ser $\mathbb{Q}(\sqrt{5})$. Para $n=8$, el grupo es $(\mathbb{Z}/8)^* = \{1,3,5,7\} \cong \mathbb{Z}_2 \times \mathbb{Z}_2$, también de orden 4 pero no cíclico, de modo que hay *tres* subcuerpos cuadráticos: $\mathbb{Q}(\mathrm{i})$, $\mathbb{Q}(\sqrt{2})$ y $\mathbb{Q}(\sqrt{-2})$.

(d) La conexión con la construibilidad. $\mathbb{Q}(\zeta_{17})$ tiene grado $\varphi(17) = 16 = 2^4$, y su grupo es cíclico de orden 16, luego admite una cadena de subgrupos de índices 2 en cada paso. Esa cadena es exactamente la torre de raíces cuadradas que Gauss usó para construir el heptadecágono.

Cuidado

[Que Φ_n sea irreducible sobre $\mathbb{Q}$ no es evidente] Parece obvio y no lo es: la demostración de que Φ_n es irreducible sobre $\mathbb{Q}$ para todo n la dio Gauss para n primo y Dedekind en general, y es delicada. Sobre otros cuerpos falla: Φ_5 factoriza sobre $\mathbb{F}_{11}$ en cuatro factores lineales, porque $11 \equiv 1 \pmod 5$. Segundo aviso: los coeficientes de Φ_n no son siempre 0 y ± 1; el primer contraejemplo aparece en Φ_{105}, que tiene un coeficiente -2. Y tercero: el teorema de Kronecker–Weber afirma que *toda* extensión abeliana de $\mathbb{Q}$ está dentro de una ciclotómica, lo que convierte a esta familia en el catálogo completo de lo abeliano sobre $\mathbb{Q}$.

Teorema 26.98 (Extensiones de Kummer)**.** *Sea K un cuerpo que contiene las raíces n-ésimas de la unidad, y $a \in K$ que no es potencia d-ésima para ningún divisor $d>1$ de n. Entonces $x^n - a$ es irreducible, $K(\sqrt[n]{a})/K$ es de Galois, y*

$$\mathrm{Gal}\left(K(\sqrt[n]{a})/K\right) \cong \mathbb{Z}_n, \qquad \sigma_k : \sqrt[n]{a} \mapsto \zeta_n^k \sqrt[n]{a}.$$

Recíprocamente, toda *extensión cíclica de grado n de tal K es de esta forma. (La dirección recíproca usa las* resolventes de Lagrange *$\rho = \sum_k \zeta_n^{-k}\sigma^k(\theta)$, cuya potencia n-ésima cae al cuerpo base: son la generalización exacta del truco u^3, v^3 de Cardano.)*

Ejemplo 26.99 (Extensiones cíclicas por radicales)**. (a) Un caso pequeño.** Sobre $\mathbb{Q}(\mathrm{i})$, que contiene las raíces cuartas de la unidad, la extensión $\mathbb{Q}(\mathrm{i}, \sqrt[4]{2})$ es cíclica de grado 4; el generador manda $\sqrt[4]{2} \mapsto \mathrm{i}\sqrt[4]{2}$, y sus potencias multiplican por -1, $-\mathrm{i}$ y 1.

(b) Sin las raíces de la unidad, falla. Sobre $\mathbb{Q}$ a secas, $\mathbb{Q}(\sqrt[3]{2})$ tiene grado 3 y *no* es de Galois: sólo contiene una de las tres raíces cúbicas de 2, y su grupo de automorfismos es trivial. Hay que adjuntar primero ζ_3 para que la teoría funcione, y entonces el grado sube a 6.

(c) El recíproco. Toda extensión cíclica de grado n de un cuerpo que contenga las raíces n-ésimas de la unidad es de la forma $K(\sqrt[n]{a})$. La demostración construye el radical como una *resolvente de Lagrange* $\alpha + \zeta\sigma(\alpha) + \zeta^2\sigma^2(\alpha) + \cdots$, cuyo n-ésimo poder resulta estar en K. Ése es el mecanismo que produce las fórmulas de Cardano.

(d) En característica p. Si $\mathrm{car}\, K = p$ y $n = p$, no hay raíces p-ésimas de la unidad distintas de 1 y la teoría de Kummer no sirve. La sustituye la teoría de Artin–Schreier, donde las extensiones cíclicas de grado p vienen de $x^p - x = a$ en lugar de $x^p = a$. El cambio de ecuación es exactamente el que hace falta.

Estos dos teoremas son las piezas con las que la demostración del criterio de Galois (Teorema 26.70) arma el puente "torre de radicales ↔ cadena con cocientes cíclicos": cada peldaño de radical es una extensión de Kummer (cíclica), previa adjunción ciclotómica (abeliana); y cada cociente cíclico de una cadena soluble se realiza como radical vía las resolventes de Lagrange. Con las dos direcciones cerradas, la teoría queda completa: **la estructura del grupo es exactamente la estructura de las fórmulas posibles**.

26.12. Complementos: norma y traza, cuerpos finitos y la ecuación general

26.12.1. Norma y traza

$N_{L/K}(\alpha) = \det(m_\alpha)$, $T_{L/K}(\alpha) = \operatorname{tr}(m_\alpha)$ (multiplicación como operador lineal: ¡capítulo 7!); en Galois, producto y suma de conjugados. En $\mathbb{Q}(\sqrt{2})$: $N = a^2 - 2b^2$: la ecuación de Pell, y el puente a la teoría de números.

26.12.2. Cuerpos finitos: teoría completa

$\mathbb{F}_{p^n}$ = descomposición de $x^{p^n} - x$; $\operatorname{Gal}(\mathbb{F}_{p^n}/\mathbb{F}_p) = \langle\varphi\rangle \cong \mathbb{Z}_n$ con $\varphi(x) = x^p$ (Frobenius; orden exacto n por conteo de raíces); subcuerpos $\mathbb{F}_{p^d} \iff d \mid n$: la correspondencia en su versión más transparente. El número de irreducibles de grado n: $\frac{1}{n}\sum_{d|n}\mu(d)p^{n/d}$ (Möbius): códigos y criptografía.

26.12.3. La ecuación general

Sobre $K = \mathbb{Q}(e_1, \ldots, e_n)$, el polinomio general tiene grupo *exactamente* S_n (permutaciones de las variables fijan los simétricos; polinomios simétricos + Artin dan la cota). Con S_n no soluble para $n \geq 5$: no hay fórmula universal —la versión complementaria del ejemplo concreto—.

26.12.4. Cardano numérico completo

$x^3 - 6x + 4$: $\frac{q^2}{4} + \frac{p^3}{27} = -4 < 0$ (casus irreducibilis); $u^3 = -2 + 2i = \sqrt{8}e^{3i\pi/4}$: $u = 1 + i$, $v = 1 - i$: $x_1 = 2$; $x_{2,3} = 2\sqrt{2}\cos(\frac{\pi}{4} \pm \frac{2\pi}{3}) = -1 \mp \sqrt{3}$. Verificación: $(x-2)(x^2+2x-2)$ ✓. Con tres raíces reales, el camino es forzosamente complejo (teorema del casus irreducibilis): los complejos entraron a la matemática por la cúbica.

26.12.5. Panorama

Problema inverso (abierto; solubles: Shafarevich), Galois infinito ($\operatorname{Gal}(\overline{\mathbb{Q}}/\mathbb{Q})$, Langlands, Fermat–Wiles), Galois diferencial (Picard–Vessiot: por qué $\int e^{-x^2}$ no es elemental: la misma imposibilidad que la quíntica, en otra categoría).

26.13. Cierre del temario: los subtemas restantes

26.13.1. Clausura algebraica

Existe $\overline{K}$ (adjuntar raíces de todo con Zorn), única salvo isomorfismo. $\overline{\mathbb{Q}}$ es *numerable*: casi todo número es trascendente (Cantor: por conteo, sin exhibir ninguno).

26.13.2. Liouville y los primeros trascendentes

Teorema 26.100. *α algebraico de grado d: $|\alpha - \frac{p}{q}| > \frac{c}{q^d}$. (TVM sobre el polinomio mínimo: $\frac{1}{q^d} \leq |m(\frac{p}{q})| \leq M|\alpha - \frac{p}{q}|$.)*

$L = \sum 10^{-n!}$ viola toda cota: trascendente (el primero de la historia, 1844). Luego e (Hermite), π (Lindemann), e^π (Gelfond–Schneider, problema 7 de Hilbert); de $e+\pi$ y $e\pi$ ni se sabe si son irracionales —pero no ambos racionales: son raíces de $x^2 - (e+\pi)x + e\pi$—. Grado de trascendencia: $\mathbb{Q}(\pi) \cong \mathbb{Q}(t)$: π es "una variable".

26.13.3. Ciclotómicos completos

$x^n - 1 = \prod_{d|n} \Phi_d$; $\Phi_n \in \mathbb{Z}[x]$ e irreducible: $[\mathbb{Q}(\zeta_n) : \mathbb{Q}] = \varphi(n)$ exacto. Recursión: $\Phi_{12} = \frac{x^{12}-1}{\Phi_1\Phi_2\Phi_3\Phi_4\Phi_6} = x^4 - x^2 + 1$; $\Phi_8 = x^4 + 1$. Coeficientes $0, \pm 1$ hasta $n = 104$; Φ_{105} trae un -2: los patrones exigen demostración.

26.13.4. La cuártica por su resolvente

Para $x^4 + px^2 + qx + r$: resolvente $t^3 - pt^2 - 4rt + (4pr - q^2)$; el grupo se lee de su factorización y del discriminante: irreducible + Δ no cuadrado: S_4; cuadrado: A_4; una raíz racional: D_4 o $\mathbb{Z}_4$; tres: V.

Ejemplo 26.101. $x^4 - 10x^2 + 1$: resolvente $t^3 + 10t^2 - 4t - 40 = (t+10)(t-2)(t+2)$: tres raíces racionales $\Rightarrow V$: el Klein calculado a mano, ahora por algoritmo.

26.13.5. Base normal e independencia de radicales

Toda Galois finita tiene θ cuya órbita es base (L libre de rango 1 sobre $K[G]$). Y $[\mathbb{Q}(\sqrt{p_1}, \ldots, \sqrt{p_k}) : \mathbb{Q}] = 2^k$ con grupo $(\mathbb{Z}_2)^k$: la familia infinita con grupo calculado para siempre.

26.14. Profundización final: Hilbert 90, reciprocidad y un diédrico grande

26.14.1. Hilbert 90

Teorema 26.102. *L/K cíclica con $\langle\sigma\rangle$: $N(\alpha) = 1 \iff \alpha = \frac{\beta}{\sigma(\beta)}$.*

Ejemplo 26.103 (El teorema 90 de Hilbert en acción)**. (a) El caso complejo.** Para $\mathbb{C}/\mathbb{R}$ con σ la conjugación, la norma es $N(z) = z\bar{z} = |z|^2$. Los elementos de norma 1 son los del círculo unidad, y el teorema dice que todos son de la forma $\beta/\bar{\beta}$. Con $z = \mathrm{i}$: tomando $\beta = 1+\mathrm{i}$ sale $\beta/\bar{\beta} = (1+\mathrm{i})/(1-\mathrm{i}) = \mathrm{i}$. Cuadra.

(b) Ternas pitagóricas. Aplicando lo mismo con $\beta = m + n\mathrm{i}$ racional se obtienen todos los puntos racionales del círculo unidad: $(\frac{m^2-n^2}{m^2+n^2}, \frac{2mn}{m^2+n^2})$. Con $m = 2, n = 1$ sale $(3/5, 4/5)$, es decir, la terna $(3, 4, 5)$. La parametrización clásica es un corolario de Hilbert 90.

(c) En cuerpos finitos. Para $\mathbb{F}_{p^n}/\mathbb{F}_p$ la norma es sobreyectiva sobre $\mathbb{F}_p^*$, y el núcleo tiene $\frac{p^n-1}{p-1}$ elementos, todos de la forma β^{1-p}. Con $p = 2$, $n = 2$: el núcleo de la norma en $\mathbb{F}_4$ tiene $3/1 = 3$ elementos, que son todos los no nulos.

(d) La lectura moderna. El enunciado equivale a $H^1(\mathrm{Gal}(L/K), L^*) = 0$: el primer grupo de cohomología se anula. Ésa es la formulación que se generaliza, y de ella salen la teoría de cuerpos de clases y buena parte de la geometría aritmética contemporánea. El número 90 viene de que era el teorema número noventa del informe de Hilbert sobre números de 1897.

En $\mathbb{Q}(i)/\mathbb{Q}$: los puntos racionales de $x^2 + y^2 = 1$ son $\frac{m+in}{m-in}$: ¡la parametrización de las ternas pitagóricas! $((m^2 - n^2, 2mn, m^2 + n^2)$.) El mismo teorema que construye extensiones de Kummer resuelve una diofántica de cuatro milenios.

26.14.2. Sumas de Gauss y reciprocidad cuadrática

$g = \sum_a (\frac{a}{p})\zeta_p^a$ cumple $g^2 = (-1)^{(p-1)/2}p$: $\mathbb{Q}(\sqrt{\pm p}) \subset \mathbb{Q}(\zeta_p)$ (Kronecker–Weber cuadrático, explícito). Comparando el Frobenius en q de dos maneras:

$$\Big(\frac{p}{q}\Big)\Big(\frac{q}{p}\Big) = (-1)^{\frac{p-1}{2}\frac{q-1}{2}},$$

el *theorema aureum*. Su descendencia: cuerpos de clases y Langlands ("reciprocidad no abeliana").

26.14.3. $x^6 - 2$: el diédrico D_6

$[E : \mathbb{Q}] = 12$ con $E = \mathbb{Q}(2^{1/6}, \zeta_6)$; $\sigma : \alpha \mapsto \zeta_6\alpha$ (orden 6), $\tau =$ conjugación, $\tau\sigma\tau^{-1} = \sigma^{-1}$: D_6, las simetrías del hexágono de raíces. Tres subgrupos de índice $2 \to$ tres cuadráticos: $\mathbb{Q}(\sqrt{-3})$, $\mathbb{Q}(\sqrt{2})$, $\mathbb{Q}(\sqrt{-6})$ (pues $D_6^{\mathrm{ab}} = \mathbb{Z}_2^2$). Soluble, como debe: es un radical.

26.14.4. El grupo absoluto y Krull

$\mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q})$ = límite profinito de los grupos finitos, con la topología de Krull (cercanía = coincidir en subextensiones grandes); el teorema fundamental sobrevive con "subgrupos *cerrados*". Sus representaciones son el lenguaje de la teoría de números del siglo XXI.

26.14.5. Especialización e irreducibilidad de Hilbert

Grupo G sobre $\mathbb{Q}(t) \Rightarrow$ infinitas especializaciones racionales con grupo G (las excepciones son delgadas): con la ecuación general, infinitos polinomios concretos con grupo S_n; y la

proporción con grupo menor tiende a 0: la genericidad de S_n es el estado natural de las ecuaciones.

26.15. Última capa: solubilidad en detalle, p-ádicos y Lüroth

26.15.1. El criterio de solubilidad, con las dos direcciones

Radicales $\Rightarrow$ *soluble*: dada la torre $K = F_0 \subset \cdots \subset F_m$ con $F_{i+1} = F_i(\sqrt[n_i]{a_i})$, se adjuntan primero las raíces de la unidad necesarias (ζ_N, $N = \operatorname{mcm}(n_i)$): eso agrega un trozo abeliano $\operatorname{Gal}(\mathbb{Q}(\zeta_N)/\mathbb{Q}) \cong (\mathbb{Z}_N)^\times$; después cada peldaño es Kummer: cíclico. Tomando la clausura normal y aplicando la correspondencia, el grupo del cuerpo de descomposición queda con una cadena de subgrupos normales con cocientes abelianos: soluble (y subgrupos/cocientes de solubles son solubles: el grupo de f, cociente del de la torre, hereda).

Soluble $\Rightarrow$ *radicales*: refinar la cadena a cocientes *cíclicos* de orden primo (posible en grupos finitos solubles); adjuntar ζ_N de entrada (abeliano, no estorba); y realizar cada peldaño cíclico de grado p como extensión por un radical mediante la *resolvente de Lagrange*: si $\operatorname{Gal}(M/F) = \langle\sigma\rangle$ de orden p y $\zeta = \zeta_p \in F$, para θ generador se forma

$$\rho = \theta + \zeta^{-1}\sigma(\theta) + \zeta^{-2}\sigma^2(\theta) + \cdots + \zeta^{-(p-1)}\sigma^{p-1}(\theta)\,:$$

entonces $\sigma(\rho) = \zeta\rho$, luego $\sigma(\rho^p) = \rho^p$: ¡$\rho^p \in F$!, y $M = F(\rho)$ con $\rho = \sqrt[p]{\rho^p}$ (algún $\rho \neq 0$ por independencia de caracteres). Cada paso cíclico *es* una raíz p-ésima: la torre queda construida. □ (Las fórmulas de Cardano son exactamente estas resolventes para S_3: la demostración abstracta y la fórmula del siglo XVI son el mismo cálculo.)

26.15.2. Números p-ádicos

Otra manera de "completar" $\mathbb{Q}$ (capítulo 6): el *valor absoluto p-ádico* $|x|_p = p^{-v_p(x)}$ (grande cuando x es muy divisible por... pequeño cuando es muy divisible por p) satisface la desigualdad ultramétrica $|x+y|_p \leq$ máx; completando se obtiene $\mathbb{Q}_p$, donde las series convergen $\iff$ el término general $\to 0$ (¡el sueño del estudiante, verdadero!). **Ostrowski**: los únicos valores absolutos de $\mathbb{Q}$ son el usual y los p-ádicos: $\mathbb{R}$ y los $\mathbb{Q}_p$ son *todas* las sombras analíticas de $\mathbb{Q}$.

Teorema 26.104 (Lema de Hensel)**.** *Si $f \in \mathbb{Z}_p[x]$ y a cumple $f(a) \equiv 0 \pmod p$ con $f'(a) \not\equiv 0$, existe una única raíz $\alpha \in \mathbb{Z}_p$ con $\alpha \equiv a$:* las raíces módulo p se levantan. *(La iteración es la de Newton, convergente en la métrica p-ádica.)*

Idea clave

Qué dice esto en cristiano. El lema de Hensel es el método de Newton trasplantado a la aritmética modular. Si encuentras una raíz módulo p y la derivada no se anula ahí, esa raíz se puede *levantar* a módulo p^2, luego a p^3, y así indefinidamente, con una única manera de hacerlo en cada paso. La condición $f'(a) \not\equiv 0$ es la misma que en Newton: hace falta que la tangente no sea horizontal para poder corregir. Es la herramienta básica de los números p-ádicos y de casi todo el análisis local en teoría de números.

Ejemplo 26.105 (Levantar una raíz paso a paso)**. (a) Raíz cuadrada de dos módulo siete.** Las soluciones de $x^2 \equiv 2 \pmod 7$ son $x = 3$ y $x = 4$, pues $9 = 7 + 2$ y $16 = 14 + 2$. La derivada es $2x$, que en $x = 3$ vale 6, no divisible por 7: se puede levantar.

(b) Los levantamientos sucesivos. Aplicando la corrección de Newton $a \mapsto a - \frac{f(a)}{f'(a)}$ módulo la potencia siguiente, se obtiene $a = 10$ módulo 49 (en efecto $100 - 2 = 98 = 2 \cdot 49$), luego $a = 108$ módulo 343 ($108^2 - 2 = 11662 = 34 \cdot 343$), y luego $a = 2166$ módulo 2401. En cada paso el número de dígitos 7-ádicos correctos se duplica, igual que en Newton se duplican las cifras decimales.

(c) El resultado en el límite. La sucesión $3, 10, 108, 2166, \ldots$ converge en los enteros 7-ádicos a un número cuyo cuadrado es exactamente 2. Es decir, $\sqrt{2}$ *existe* en $\mathbb{Z}_7$, aunque no exista en $\mathbb{Q}$. Los cuerpos p-ádicos tienen raíces que los racionales no tienen, y viceversa.

(d) Cuándo no arranca. Para $p = 5$, la congruencia $x^2 \equiv 2 \pmod 5$ no tiene ninguna solución, porque los cuadrados módulo 5 son $\{0, 1, 4\}$. Sin raíz inicial no hay nada que levantar: $\sqrt{2}$ no existe en $\mathbb{Z}_5$. La existencia depende del primo, y el símbolo de Legendre $\left(\frac{2}{p}\right)$ es precisamente el que decide.

Cuidado

[La hipótesis sobre la derivada es esencial] Si $f'(a) \equiv 0 \pmod p$ el lema no dice nada, y de hecho puede fallar. Para $f(x) = x^2 - p$ la raíz $a = 0$ cumple $f(0) = -p \equiv 0$, pero $f'(0) = 0$: y en efecto $\sqrt{p}$ no está en $\mathbb{Z}_p$. Existe una versión refinada del lema que cubre estos casos comparando $|f(a)|$ con $|f'(a)|^2$, pero la versión simple no basta. Segundo aviso: el lema garantiza una raíz *única* que levanta la dada; puede haber otras raíces que levanten otras soluciones módulo p, como el $x = 4$ del apartado (a), que da la otra raíz cuadrada.

Ejemplo: $x^2 = 2$ en $\mathbb{Q}_7$: $3^2 = 9 \equiv 2 \pmod 7$; Hensel levanta: $x = 3 + 7t$ con $9 + 42t \equiv 2 \pmod{49}$: $t = 1$: $x = 10$ ($100 = 2 \cdot 49 + 2$ ✓), y así sucesivamente: $\sqrt{2} = 3 + 1 \cdot 7 + 2 \cdot 7^2 + \cdots \in \mathbb{Q}_7$. La filosofía *local–global* (estudiar en $\mathbb{R}$ y todos los $\mathbb{Q}_p$, y pegar: Hasse–Minkowski para formas cuadráticas) es el sistema operativo de la teoría de números moderna, y los grupos de Galois locales $\mathrm{Gal}(\overline{\mathbb{Q}}_p/\mathbb{Q}_p)$ son las piezas con que se analiza el global.

26.15.3. El teorema de Lüroth

Todo cuerpo intermedio $\mathbb{Q} \subsetneq F \subseteq \mathbb{Q}(t)$ es de nuevo una función racional pura: $F = \mathbb{Q}(g(t))$ para alguna g. Consecuencia geométrica: toda curva parametrizable racionalmente admite parametrización "sin redundancia"; en la correspondencia de Galois del cuerpo $\mathbb{Q}(t)$ (con grupos infinitos como PGL_2), Lüroth es el análogo del teorema del elemento primitivo. Su fallo en dos variables (superficies unirracionales no racionales, resuelto en los 70) es un teorema profundo de geometría algebraica: la frontera exacta donde la intuición de una variable se rompe.

26.15.4. Cohomología de Galois, en una cápsula

Hilbert 90 moderno: $H^1(\mathrm{Gal}(L/K), L^\times) = 1$ —los "1-cociclos" multiplicativos son todos triviales—, del cual la versión cíclica clásica es el caso computable. El lenguaje cohomológico organiza la teoría entera: las extensiones cíclicas de grado n de un cuerpo con ζ_n están clasificadas por $K^\times/(K^\times)^n$ (Kummer, releído), las álgebras centrales simples por el grupo de Brauer H^2, y la teoría de cuerpos de clases es un cálculo cohomológico gigante. Para este libro basta el mensaje: *los teoremas 90, de Kummer y de las resolventes de Lagrange son la misma agua en tres pozos*, y el idioma que los unifica es la cohomología —la puerta por donde el lector continuará, si quiere, hacia la teoría de números del presente—.

26.16. Ronda de temarios: base normal, inseparabilidad y trascendencia

Tres temas de los programas de álgebra de posgrado que completan el panorama de las extensiones de cuerpos por sus tres flancos: el módulo galoisiano, la característica p y lo no algebraico.

26.16.1. El teorema de la base normal

Teorema 26.106 (Base normal)**.** *Toda extensión de Galois finita L/F admite $\alpha \in L$ tal que $\{\sigma(\alpha) : \sigma \in \mathrm{Gal}(L/F)\}$ es una base de L como F-espacio vectorial.*

Demostración para el caso cíclico. Sea $G = \langle\sigma\rangle$ de orden n. Mírese $\sigma : L \to L$ como operador F-lineal. Satisface $\sigma^n = \mathrm{id}$, así que su polinomio mínimo divide a $x^n - 1$; pero por la independencia lineal de caracteres (Artin–Dedekind, ya demostrada) los operadores $\mathrm{id}, \sigma, \ldots, \sigma^{n-1}$ son linealmente independientes sobre L (luego sobre F), de modo que ningún polinomio de grado $< n$ anula a σ: el polinomio mínimo es *exactamente* $x^n - 1$, de grado $n = \dim_F L$: mínimo = característico. Por la teoría del capítulo 7 (forma racional), σ admite un *vector cíclico* α: los iterados $\alpha, \sigma\alpha, \ldots, \sigma^{n-1}\alpha$ forman base. Eso es la base normal. □ □

Ejemplo 26.107. $\mathbb{Q}(\sqrt{2})/\mathbb{Q}$: el candidato ingenuo $\alpha = \sqrt{2}$ *falla* ($\{\sqrt{2}, -\sqrt{2}\}$ es dependiente), pero $\alpha = 1 + \sqrt{2}$ funciona: $\{1+\sqrt{2},\ 1-\sqrt{2}\}$ es base. Para $\mathbb{Q}(\zeta_p)/\mathbb{Q}$: $\alpha = \zeta_p$ da la base normal $\{\zeta_p, \zeta_p^2, \ldots, \zeta_p^{p-1}\}$ (las potencias primitivas: base porque el polinomio ciclotómico tiene grado $p-1$). La base normal convierte a L en *módulo libre de rango 1* sobre el anillo de grupo $F[G]$: la extensión entera es "un solo elemento movido por el grupo", el punto de partida de la teoría de módulos galoisianos y de aplicaciones tan concretas como la aritmética rápida en $\mathbb{F}_{2^n}$ para criptografía (elevar al cuadrado en base normal es *rotar* coordenadas: el Frobenius sale gratis en hardware).

26.16.2. Inseparabilidad: lo que la característica p rompe

Ejemplo 26.108 (El ejemplo obligatorio)**.** Sea $K = \mathbb{F}_p(t)$ y $f(x) = x^p - t$. Es irreducible (Eisenstein con el primo t del DFU $\mathbb{F}_p[t]$), pero en $L = K(\alpha)$ con $\alpha^p = t$:

$$x^p - t = x^p - \alpha^p = (x - \alpha)^p$$

(¡el sueño del novato, teorema en característica p!): una sola raíz con multiplicidad p. La extensión L/K es *puramente inseparable*: ningún automorfismo puede mover α (¿a qué otra raíz?), así que $\mathrm{Aut}(L/K) = \{\mathrm{id}\}$ pese a $[L:K] = p$: la correspondencia de Galois colapsa por completo. El criterio general: f irreducible es inseparable $\iff f' = 0 \iff f(x) = g(x^p)$: la derivada formal (¡sin límites!) detecta el fenómeno.

Definición 26.109. Para L/K finita, el *grado de separabilidad* $[L:K]_s$ es el número de K-inmersiones de L en una cerradura algebraica. Siempre $[L:K] = [L:K]_s \cdot p^m$ para algún $m \geq 0$, y la extensión se factoriza en una torre $K \subseteq L_s \subseteq L$ con L_s/K separable (el "cuerpo de los separables") y L/L_s puramente inseparable.

Ejemplo 26.110 (Separable y no separable)**. (a) En característica cero, todo es separable.** Si $\operatorname{car} K = 0$, todo polinomio irreducible tiene derivada no nula y por tanto raíces distintas, luego $[L:K]_s = [L:K]$ siempre. Las patologías de esta definición sólo aparecen en característica positiva.

(b) El ejemplo canónico. Sea $K = \mathbb{F}_p(t)$ y $L = K(t^{1/p})$. El polinomio mínimo es $x^p - t = (x - t^{1/p})^p$: una única raíz con multiplicidad p. Por tanto $[L:K] = p$ mientras que $[L:K]_s = 1$. La extensión es puramente inseparable.

(c) La factorización del grado. Siempre se cumple $[L:K] = [L:K]_s \cdot [L:K]_i$, donde el segundo factor es una potencia de p. En el ejemplo anterior, $p = 1 \cdot p$. Toda extensión se descompone en una parte separable seguida de una puramente inseparable.

(d) Cuándo hay que preocuparse. Los cuerpos *perfectos* (los de característica cero y los finitos) no tienen extensiones inseparables, y por eso en la mayoría de los cursos la palabra "separable" se puede ignorar. Deja de poderse ignorar en geometría algebraica sobre cuerpos de funciones en característica p, que es justo donde más se usa la teoría de Galois moderna.

Teorema 26.111. *Un cuerpo K de característica p es* perfecto *(toda extensión algebraica es separable) si y solo si el Frobenius $x \mapsto x^p$ es sobreyectivo. En particular, todo cuerpo finito es perfecto, y toda extensión de $\mathbb{Q}$ es separable.*

Demostración. Si Frobenius es sobre y $f(x) = g(x^p) = \sum a_i x^{pi}$ fuera irreducible, escribiendo $a_i = b_i^p$: $f = \left(\sum b_i x^i\right)^p$, no irreducible: no hay inseparables. Recíproco: si algún t no es potencia p-ésima, $x^p - t$ repite el ejemplo de arriba. Los finitos: Frobenius es inyectivo (es homomorfismo de cuerpos) y un endomorfismo inyectivo de un conjunto finito es sobre. Característica 0: $f' \neq 0$ automáticamente. □ □

Observación 26.112. Por esto pudimos desarrollar toda la teoría de Galois sobre $\mathbb{Q}$ y sobre $\mathbb{F}_q$ sin mencionar la separabilidad más que como hipótesis técnica: nuestros dos mundos favoritos son perfectos. El fenómeno inseparable vive en los cuerpos de funciones $\mathbb{F}_p(t)$ —el hábitat de la geometría algebraica en característica p— y allí no es patología sino estructura: el Frobenius no sobreyectivo es la herramienta central (la que Weil y Grothendieck convirtieron en la maquinaria que cuenta puntos de curvas, y con la que se probó la "hipótesis de Riemann sobre cuerpos finitos").

26.16.3. Bases de trascendencia

Definición 26.113. $S \subseteq L$ es *algebraicamente independiente* sobre K si ningún polinomio no nulo con coeficientes en K se anula en elementos distintos de S. Una *base de trascendencia* de L/K es un conjunto algebraicamente independiente maximal; entonces L es algebraico sobre $K(S)$.

Teorema 26.114 (Steinitz)**.** *Dos bases de trascendencia de L/K tienen el mismo cardinal: el* grado de trascendencia $\operatorname{trdeg}(L/K)$.

Idea. Es el *lema de intercambio* de Steinitz, calcado del capítulo 7: si $\{x_1, \dots, x_m\}$ y $\{y_1, \dots, y_n\}$ son bases de trascendencia con $m < n$, cada y_j es algebraico sobre $K(x_1, \dots, x_m)$, y uno a uno se intercambian los x_i por y_j manteniendo la independencia (el polinomio que relaciona a y_1 con los x debe involucrar a algún x_i: ese se despeja y se sustituye); al agotar los x sobran y algebraicos entre sí: contradicción. La independencia algebraica satisface los mismos axiomas de "dependencia" que la lineal (los *matroides*): por eso la prueba es la misma palabra por palabra. □ □

Ejemplo 26.115. $\operatorname{trdeg} \mathbb{Q}(t_1, \dots, t_n)/\mathbb{Q} = n$: por fin es un *teorema* que el cuerpo de funciones racionales en n variables "tiene n variables". $\operatorname{trdeg}(\mathbb{C}/\mathbb{Q})$ es la cardinalidad del continuo (una base de trascendencia de $\mathbb{C}$ es enorme e invisible: nadie ha exhibido una, Zorn garantiza que existe). Y la geometría lee el grado de trascendencia como *dimensión*: el cuerpo de funciones de una curva tiene grado 1, el de una superficie 2; Lüroth (ya probado) dice en este lenguaje que "subcuerpo de dimensión 1 de una recta es una recta". El diccionario cuerpos↔geometría con que cerramos el capítulo es la puerta de entrada a la geometría algebraica.

26.17. Cierre definitivo: Galois infinita y teoría de Kummer

Las dos últimas piezas del edificio: qué pasa cuando el grupo es infinito, y la descripción exacta de las extensiones abelianas cuando hay raíces de la unidad en casa.

26.17.1. Teoría de Galois infinita: la topología de Krull en serio

Para extensiones infinitas como $\overline{\mathbb{F}}_p/\mathbb{F}_p$ o $\mathbb{Q}(\zeta_{p^\infty})/\mathbb{Q}$, el grupo $\operatorname{Gal}(L/F)$ es infinito y la correspondencia ingenua *falla*: hay más subgrupos que cuerpos intermedios. El arreglo: el grupo de Galois no es un grupo desnudo sino un *grupo topológico*.

Definición 26.116. La *topología de Krull* en $G = \operatorname{Gal}(L/F)$: dos automorfismos están cerca si coinciden en una subextensión finita grande. Formalmente, $G = \varprojlim \operatorname{Gal}(E/F)$ sobre las subextensiones finitas de Galois E (*límite inverso*: un elemento de G es una elección compatible de un automorfismo por cada piso finito), con la topología heredada del producto de los grupos finitos discretos. Así G es compacto y totalmente disconexo: un grupo *profinito*.

Teorema 26.117 (Correspondencia de Galois infinita)**.** *Para L/F de Galois (algebraica, normal, separable) la correspondencia*

$$H \mapsto L^H, \qquad E \mapsto \mathrm{Gal}(L/E)$$

es una biyección entre los subgrupos cerrados *de G y los cuerpos intermedios; los subgrupos* abiertos *corresponden a las subextensiones* finitas*.*

Idea. Cada piso finito usa la teoría ya demostrada; la compacidad pega los pisos. Un subgrupo no cerrado y su clausura fijan el mismo cuerpo (coincidir en cada piso finito es la definición de cercanía): por eso los no cerrados sobran, y esa es exactamente la corrección que la topología aporta. □ □

Ejemplo 26.118 (El ejemplo canónico: $\widehat{\mathbb{Z}}$)**.** $\mathrm{Gal}(\overline{\mathbb{F}}_p/\mathbb{F}_p)$: cada piso finito $\mathbb{F}_{p^n}/\mathbb{F}_p$ tiene grupo cíclico $\mathbb{Z}_n$ generado por Frobenius (ya demostrado); el límite inverso es

$$\widehat{\mathbb{Z}} = \varprojlim \mathbb{Z}_n \cong \prod_q \mathbb{Z}_q \quad \text{(producto de los enteros } q\text{-ádicos del capítulo, por resto chino),}$$

el "cierre profinito" de $\mathbb{Z}$. Frobenius genera un subgrupo denso pero *no* todo: hay automorfismos de $\overline{\mathbb{F}}_p$ que no son potencias de Frobenius (¡solo límites de potencias!), y los subgrupos no cerrados asociados son la razón visible de que la correspondencia necesite la topología. La misma maquinaria describe $\mathrm{Gal}(\mathbb{Q}(\zeta_{p^\infty})/\mathbb{Q}) \cong \mathbb{Z}_p^\times$: los grupos p-ádicos que construimos por Hensel son grupos de Galois de torres ciclotómicas: dos capítulos del libro resultaron ser el mismo objeto.

26.17.2. Teoría de Kummer: las abelianas, clasificadas

Teorema 26.119 (Kummer)**.** *Sea F con $\zeta_n \in F$ (y característica ajena a n). Las extensiones abelianas de exponente n de F están en biyección con los subgrupos Δ de $F^\times/(F^\times)^n$:*

$$\Delta \;\longmapsto\; F(\sqrt[n]{\Delta}), \qquad y \qquad \mathrm{Gal}\left(F(\sqrt[n]{\Delta})/F\right) \cong \mathrm{Hom}\left(\Delta,\ \langle\zeta_n\rangle\right).$$

En particular $[F(\sqrt[n]{a}) : F]$ es el orden de a en $F^\times/(F^\times)^n$.

Idea, con las herramientas del capítulo. Cada radical $\sqrt[n]{a}$ da una extensión cíclica: el emparejamiento $\sigma \mapsto \frac{\sigma(\sqrt[n]{a})}{\sqrt[n]{a}} \in \langle\zeta_n\rangle$ es un homomorfismo inyectivo (¡la cocadena de Hilbert 90 del capítulo!). Recíprocamente, toda cíclica de grado n viene de un radical: esa es exactamente la mitad "resolvente de Lagrange" del teorema de solubilidad, ya demostrada ($\sigma(\rho) = \zeta\rho \Rightarrow \rho^n \in F$). Kummer solo organiza ambas mitades en una biyección funtorial, y la independencia de caracteres (Artin–Dedekind otra vez) separa radicales independientes. □ □

Ejemplo 26.120. Sobre $\mathbb{Q}(\zeta_3)$: las extensiones cúbicas abelianas son exactamente $\mathbb{Q}(\zeta_3, \sqrt[3]{a})$ con a no cubo, y $\sqrt[3]{2}, \sqrt[3]{3}, \sqrt[3]{6}, \ldots$ generan extensiones distintas porque $2, 3, 6$ son independientes en $\mathbb{Q}(\zeta_3)^\times/(\text{cubos})$. Sin ζ_n en casa el teorema es falso ($\mathbb{Q}(\sqrt[3]{2})/\mathbb{Q}$ ni siquiera es normal: la obstrucción con que abrimos el capítulo), y *clasificar las abelianas sobre $\mathbb{Q}$ a secas* exige juntar

todos los ciclotómicos: eso es el teorema de Kronecker–Weber (toda abeliana sobre $\mathbb{Q}$ vive en un $\mathbb{Q}(\zeta_m)$), la puerta de la teoría de cuerpos de clases con la que el capítulo se despidió. Kummer es, además, el motor del criterio con que atacó el último teorema de Fermat: la aritmética de $\mathbb{Z}[\zeta_p]$ y sus primos regulares.

26.18. Segunda ronda: extensión de isomorfismos y el Frobenius en acción

26.18.1. El teorema de extensión

Teorema 26.121. *Un isomorfismo $\sigma : K \to K'$ se extiende a $K(\alpha) \to K'(\alpha')$ mandando α a cualquier raíz α' de $\sigma(p)$ (p = mínimo de α); el número de extensiones a un cuerpo de descomposición es el grado (caso separable).* Idea*: $K(\alpha) \cong K[x]/(p) \to K'[x]/(\sigma p) \cong K'(\alpha')$.*

Ejemplo 26.122 (Contar extensiones de un isomorfismo)**. (a) La conjugación compleja.** Con $K = K' = \mathbb{Q}$ y $\alpha = \mathrm{i}$, el polinomio mínimo x^2+1 tiene dos raíces, $\pm\mathrm{i}$, luego la identidad de $\mathbb{Q}$ se extiende de exactamente dos maneras a $\mathbb{Q}(\mathrm{i})$: la identidad y la conjugación. Eso dice que $|\operatorname{Gal}(\mathbb{Q}(\mathrm{i})/\mathbb{Q})| = 2$.

(b) Una raíz cúbica real. Con $\alpha = \sqrt[3]{2}$, el polinomio x^3-2 tiene una sola raíz *dentro de* $\mathbb{R}$, así que la identidad de $\mathbb{Q}$ se extiende de una única manera a $\mathbb{Q}(\sqrt[3]{2}) \subset \mathbb{R}$: el grupo de automorfismos es trivial pese a que el grado es 3. La extensión no es normal.

(c) Completando la torre. Si se adjunta también ζ_3, el cuerpo $\mathbb{Q}(\sqrt[3]{2}, \zeta_3)$ tiene grado 6 y ahora sí hay seis extensiones: se puede mandar $\sqrt[3]{2}$ a cualquiera de las tres raíces y ζ_3 a cualquiera de las dos primitivas. El grupo es S_3.

(d) La demostración de la existencia del cuerpo de descomposición. El teorema se aplica repetidamente, adjuntando una raíz cada vez, y por inducción se obtiene que dos cuerpos de descomposición de un mismo polinomio son isomorfos. La unicidad salvo isomorfismo, que es lo que permite hablar de "el" cuerpo de descomposición, sale íntegramente de este enunciado.

El andamio invisible del capítulo: da la unicidad del cuerpo de descomposición, la cuenta $|\operatorname{Gal}| = [L:K]$, y el mecanismo para subir la conjugación o el Frobenius a las extensiones.

26.18.2. El Frobenius, visto factorizar

$x^4+1 = \Phi_8$, grupo $(\mathbb{Z}_8)^\times$: el Frobenius en p es $\zeta \mapsto \zeta^p$, y su clase mod 8 decide:

p mód 8	orden	factorización mod p
1	1	cuatro lineales ($p = 17$: raíces $2, 8, 9, 15$)
$3, 5, 7$	2	dos cuadráticos

Verificado: mod 3: $(x^2+x+2)(x^2+2x+2)$; mod 7: $(x^2+3x+1)(x^2+4x+1)$. Nunca un factor lineal con $p \not\equiv 1$: *por eso* x^4+1 es reducible módulo todo primo pero irreducible sobre $\mathbb{Q}$: el grupo de Klein no tiene 4-ciclos que el Frobenius pudiera realizar. Chebotarev: densidades $\frac{1}{4}$ y $\frac{3}{4}$.

26.18.3. Un mínimo con estrategia: $\sqrt{2}+i$

Conjugados $\pm\sqrt{2}\pm i$ (Klein actuando):

$$m(x)=((x-\sqrt{2})^2+1)((x+\sqrt{2})^2+1)=x^4-2x^2+9,$$

irreducible, de grado $4=[\mathbb{Q}(\sqrt{2},i):\mathbb{Q}]$: α primitivo, con el algoritmo universal (órbita, producto, expandir) que la resultante mecaniza.

Idea clave

Intuición de cierre. ¿Por qué las fórmulas mueren en grado 5? Cada radical $\sqrt[p]{\ }$ que uno adjunta es un piso de simetría *cíclica*: las torres de radicales solo construyen grupos hechos de capas abelianas (solubles). Los grados 2, 3 y 4 tienen grupos S_2, S_3, S_4 desmontables en esas capas —por eso existen la fórmula cuadrática, Cardano y la cuártica—. Pero dentro de S_5 vive A_5: un engranaje *simple*, sin ninguna capa intermedia que extraer, y ninguna torre de radicales puede fabricar un piso indivisible no abeliano. No es que no hayamos encontrado la fórmula de la quíntica: es que el edificio que tendría que construir tiene un piso imposible. La sección siguiente muestra el reverso: quínticas particulares con grupo pequeño *sí* se resuelven: la obstrucción es del grupo, no del grado.

26.19. Coda: la quíntica soluble y Chebotarev cuantitativo

¿Cómo se decide si una quíntica *concreta* es soluble? El criterio práctico (Cayley): f de grado 5 irreducible es soluble por radicales $\iff$ su grupo cabe en F_{20} $\iff$ la *resolvente séxtica* (un polinomio auxiliar de grado 6 en los coeficientes) tiene una raíz racional. Para x^5-2: la tiene (grupo F_{20}, soluble, como vimos); para x^5-4x+2: no (grupo S_5). Y el teorema de *densidad de Chebotarev* vuelve cuantitativo el método de Dedekind: para x^5-4x+2 (grupo S_5), las proporciones de primos según la factorización mod p convergen a las de tipos de ciclo en S_5: $\frac{24}{120}=20\,\%$ irreducible (5-ciclos), $\frac{30}{120}=25\,\%$ tipo (4)(1), $\frac{20}{120}$ tipo (3)(2), $\ldots$, $\frac{1}{120}$ totalmente partido. Contar factorizaciones es *medir* el grupo: con unos cientos de primos, el histograma delata a S_5 con confianza estadística —la teoría de Galois convertida en experimento—. Este mismo teorema, aplicado a los ciclotómicos, contiene a Dirichlet (infinitos primos en cada progresión $a+nq$ con $\gcd(a,q)=1$, equidistribuidos): el capítulo se despide tocando la teoría analítica de números.

26.20. Temas extra y ejemplos desarrollados

26.20.1. El pentágono con regla y compás

$M=(\frac{1}{2},0)$; circunferencia de centro M y radio $\frac{\sqrt{5}}{2}$ corta el eje en $\frac{\sqrt{5}-1}{2}=2\cos 72^\circ$: bisecar, levantar perpendicular, transportar cuerda: la torre $\mathbb{Q}\subset\mathbb{Q}(\sqrt{5})$ hecha con las manos, con la razón áurea (diagonal/lado del pentágono) emergiendo.

26.20.2. Cuerpos cuadráticos y Kronecker–Weber

Toda cuadrática es $\mathbb{Q}(\sqrt{d})$; en la familia viven $\mathbb{Z}[i]$ (dos cuadrados), Pell, y el fallo de factorización que motivó los ideales. Kronecker–Weber: toda abeliana sobre $\mathbb{Q}$ vive en un ciclotómico ($\sqrt{5} \in \mathbb{Q}(\zeta_5)$ por los períodos de Gauss); su generalización es el problema 12 de Hilbert.

26.20.3. Un reducible instructivo

$x^4 - 5x^2 + 6 = (x^2 - 2)(x^2 - 3)$: grupo Klein no transitivo: *las órbitas del grupo son los factores irreducibles*; y el grupo de un producto es el producto exactamente cuando los cuerpos se cortan en $\mathbb{Q}$.

26.20.4. Galois computacional

Dedekind: la factorización de f mod p copia el tipo de ciclos del Frobenius. $x^5 - 4x + 2$: mod 3 irreducible (5-ciclo), mod 7 tipo (2)(3) (su cubo: trasposición): S_5 confirmado sin análisis real. Chebotarév: las frecuencias de factorización convergen a las proporciones de tipos de ciclo: el grupo es estadísticamente visible.

26.20.5. AES y los cuerpos de Galois

AES opera en $\mathbb{F}_{256} = \mathbb{F}_2[x]/(x^8 + x^4 + x^3 + x + 1)$: bytes como polinomios, XOR como suma, S-box esencialmente $b \mapsto b^{-1}$ (Bézout). Reed–Solomon (QR, discos) y curvas elípticas también viven en cuerpos finitos: la teoría de Galois procesa cada segundo de internet.

26.21. Tres deudas del capítulo: irreducibilidad de Φ_n, el casus irreducibilis y la trascendencia de π

Φ_n es irreducible sobre $\mathbb{Q}$

Solo se demostró el caso n primo, y sin embargo de este resultado dependen $[\mathbb{Q}(\zeta_n) : \mathbb{Q}] = \varphi(n)$, la identificación $\mathrm{Gal}(\mathbb{Q}(\zeta_n)/\mathbb{Q}) \cong (\mathbb{Z}/n)^\times$, los períodos de Gauss y el criterio de constructibilidad. Va la demostración general.

Teorema 26.123. *Φ_n es irreducible en $\mathbb{Q}[x]$ para todo n.*

Demostración. Sea ζ una raíz primitiva n-ésima y f su polinomio mínimo, que divide a $x^n - 1$; escríbase $x^n - 1 = fg$ con $f, g \in \mathbb{Z}[x]$ mónicos (por el lema de Gauss). Sea p un primo que no divide a n; queremos ver que ζ^p también es raíz de f.

Si no lo fuera, sería raíz de g, luego ζ sería raíz de $g(x^p)$, y por minimalidad $f \mid g(x^p)$. Redúzcase módulo p: en $\mathbb{F}_p[x]$ se tiene $g(x^p) = g(x)^p$ por Frobenius, luego $\bar{f} \mid \bar{g}^p$. Así que $\bar{f}$ y $\bar{g}$ comparten algún factor irreducible, y por tanto $x^n - 1 = \bar{f}\bar{g}$ tendría una raíz *múltiple* en la clausura de $\mathbb{F}_p$. Pero su derivada es nx^{n-1}, que no se anula en ninguna raíz porque $p \nmid n$: el polinomio es separable. Contradicción.

Luego f tiene por raíces todas las ζ^p con $p \nmid n$, y componiendo esos pasos, todas las ζ^k con $\gcd(k,n)=1$: es decir, todas las raíces primitivas. Luego $f=\Phi_n$ y Φ_n es irreducible. □

Ejemplo 26.124. $\Phi_{12}=x^4-x^2+1$. Eisenstein no se le aplica en ninguna traslación, y sin embargo el teorema lo declara irreducible. Luego $[\mathbb{Q}(\zeta_{12}):\mathbb{Q}]=\varphi(12)=4$ y

$$\mathrm{Gal}(\mathbb{Q}(\zeta_{12})/\mathbb{Q}) \cong (\mathbb{Z}/12)^\times = \{1,5,7,11\},$$

donde *todos* los elementos tienen orden 2: es el grupo de Klein. Por la correspondencia, hay tres subcuerpos cuadráticos, y son

$$\mathbb{Q}(i), \qquad \mathbb{Q}(\sqrt{3}), \qquad \mathbb{Q}(\sqrt{-3}).$$

Todo ello sin calcular una sola raíz.

Casus irreducibilis

Teorema 26.125. *Sea $f\in\mathbb{Q}[x]$ cúbica irreducible con las* tres *raíces reales. Entonces ninguna de ellas pertenece a una torre de radicales* reales*: cualquier expresión por radicales tiene que pasar por los complejos.*

Idea. Con tres raíces reales el discriminante es positivo, y en la fórmula de Cardano aparece $\sqrt{q^2/4+p^3/27}<0$. Que eso sea inevitable —y no un defecto de la fórmula— se prueba con Galois: sea L el cuerpo de descomposición, de grado 3 o 6. Si alguna raíz viviera en una torre real $\mathbb{Q}=F_0\subset\cdots\subset F_m\subset\mathbb{R}$ con cada paso $F_{i+1}=F_i(\sqrt[n_i]{a_i})$, se puede suponer los n_i primos. Tómese el primer nivel que contenga una raíz. Un paso de grado 3 real obliga a que F_i contenga las raíces cúbicas de la unidad, que no son reales; y un paso de grado 2 no puede generar una extensión de grado 3. Contradicción. □

Ejemplo 26.126 (Tres raíces reales y un rodeo obligado por $\mathbb{C}$)**.** Sea

$$f(x)=x^3-3x+1.$$

No tiene raíces racionales ($f(1)=-1$, $f(-1)=3$), luego es irreducible. Su discriminante es

$$\Delta=-4p^3-27q^2=-4(-27)-27=81>0,$$

luego las tres raíces son reales. De hecho son

$$2\cos\frac{2\pi}{9}=1{,}532\,088\,9, \qquad 2\cos\frac{8\pi}{9}=-1{,}879\,385\,2, \qquad 2\cos\frac{14\pi}{9}=0{,}347\,296\,4,$$

como se comprueba con la identidad $(2\cos\theta)^3-3(2\cos\theta)=2\cos 3\theta$ y $3\theta=2\pi/3$.

Y sin embargo Cardano exige calcular

$$\sqrt[3]{-\tfrac12+\sqrt{\tfrac14-1}}=\sqrt[3]{-\tfrac12+\tfrac{\sqrt3}{2}i},$$

es decir, pasar por $\sqrt{-3}$. Tres números reales cuya única expresión por radicales viaja por el plano complejo y vuelve. Ese fue, históricamente, el argumento que obligó a tomarse en serio los números imaginarios: no eran un capricho para ecuaciones sin solución, sino un rodeo *necesario* para ecuaciones con tres soluciones reales.

Trascendencia

Teorema 26.127 (Hermite, 1873)**.** *e es trascendente.*

Esquema. Supóngase $a_0 + a_1 e + \cdots + a_n e^n = 0$ con $a_i \in \mathbb{Z}$, $a_0 \neq 0$. Para un polinomio f auxiliar, defínase $I(t) = \int_0^t e^{t-x} f(x)\,\mathrm{d}x$, que integrando por partes repetidamente vale $e^t \sum f^{(j)}(0) - \sum f^{(j)}(t)$. Aplicando la relación supuesta se obtiene un entero $J = \sum_k a_k I(k)$. Con la elección

$$f(x) = x^{p-1}(x-1)^p \cdots (x-n)^p$$

y p primo grande, se prueba que J es un entero *no divisible* por p —luego $|J| \geq 1$— y a la vez que $|J| < 1$ por acotación directa de la integral. Contradicción. □

Teorema 26.128 (Lindemann–Weierstrass, 1882)**.** *Si $\alpha_1, \ldots, \alpha_n$ son algebraicos distintos, entonces $e^{\alpha_1}, \ldots, e^{\alpha_n}$ son linealmente independientes sobre $\overline{\mathbb{Q}}$.*

Corolario 26.129 (π es trascendente)**.** *Si π fuera algebraico, también lo sería $i\pi$. Aplicando el teorema a $\{0,\ i\pi\}$, los números $e^0 = 1$ y $e^{i\pi} = -1$ serían linealmente independientes sobre $\overline{\mathbb{Q}}$. Pero*

$$1 \cdot e^{i\pi} + 1 \cdot e^0 = -1 + 1 = 0$$

es una relación no trivial. Contradicción.

Observación 26.130. Y con eso se cierra la cuadratura del círculo, que el capítulo dejaba pendiente de una cita. Todo número construible con regla y compás es algebraico —de hecho, de grado potencia de 2—; $\sqrt{\pi}$ no lo es, porque si lo fuera π también; luego no se puede construir un cuadrado de área π. Dos mil años de intentos quedaron zanjados por una relación de una línea entre e, i, π y -1.

Nótese la diferencia de nivel: Liouville, demostrado antes en el capítulo, construye trascendentes *a medida* con desarrollos de fracción continua muy rápidos. Hermite y Lindemann demuestran la trascendencia de números *dados de antemano*, que es incomparablemente más difícil. Y sigue habiendo preguntas abiertas del todo elementales: no se sabe si $e + \pi$ o $e\pi$ son trascendentes, aunque sí que al menos uno de los dos lo es —porque si ambos fueran algebraicos, e y π serían raíces de un mismo polinomio de grado 2 con coeficientes algebraicos—.

26.22. Gauss–Wantzel: la implicación que faltaba, y el teorema 90 de Hilbert

Del criterio de constructibilidad del n-ágono sólo se justificó una dirección —la que dice qué es *imposible*—, y el propio capítulo advertía de que "grado potencia de dos" no basta en general. Aquí está la vuelta, con la explicación de por qué en este caso sí basta.

Teorema 26.131 (Constructibilidad de las torres cuadráticas)**.** *Un número complejo α es construible con regla y compás a partir de $\{0,1\}$ si y sólo si existe una torre de cuerpos*

$$\mathbb{Q} = K_0 \subset K_1 \subset \cdots \subset K_m \ni \alpha \qquad \textit{con} \qquad [K_{i+1} : K_i] = 2.$$

Demostración. ($\Rightarrow$) Cada paso de una construcción es la intersección de dos rectas, de una recta y una circunferencia, o de dos circunferencias, cuyas coordenadas resuelven un sistema lineal o una ecuación de segundo grado sobre el cuerpo generado por los datos: se sube de grado 1 o 2.

($\Leftarrow$) Una extensión de grado 2 es $K(\sqrt{d})$ para algún $d \in K$, y la raíz cuadrada de un número dado es construible con regla y compás —la construcción clásica de la media geométrica: un semicírculo de diámetro $1+d$ y la altura sobre el punto que lo divide en 1 y d—. Encadenando las m construcciones se obtiene α. □

Teorema 26.132 (Gauss–Wantzel, ambas direcciones)**.** *El polígono regular de n lados es construible con regla y compás si y sólo si*

$$n = 2^k\, p_1 p_2 \cdots p_r,$$

donde los p_i son primos de Fermat distintos (primos de la forma $2^{2^s}+1$).

Demostración. Construir el n-ágono equivale a construir $\zeta_n = \mathrm{e}^{2\pi \mathrm{i}/n}$, y $[\mathbb{Q}(\zeta_n):\mathbb{Q}] = \varphi(n)$ porque el polinomio ciclotómico Φ_n es irreducible —resultado ya demostrado en este capítulo—.

($\Rightarrow$) Si ζ_n es construible, el teorema 26.131 da una torre cuadrática, luego $[\mathbb{Q}(\zeta_n):\mathbb{Q}] = \varphi(n)$ divide a 2^m: es una potencia de 2. Escribiendo $n = 2^k \prod p_i^{a_i}$ con los p_i primos impares, se tiene $\varphi(n) = 2^{k-1}\prod p_i^{a_i-1}(p_i-1)$, y para que sea potencia de 2 hace falta $a_i = 1$ y que cada $p_i - 1$ sea potencia de 2, es decir, que cada p_i sea primo de Fermat.

($\Leftarrow$) Aquí está lo que faltaba. Supóngase $\varphi(n) = 2^m$. La extensión $\mathbb{Q}(\zeta_n)/\mathbb{Q}$ es de Galois —es el cuerpo de descomposición de $x^n - 1$— con grupo

$$G = \mathrm{Gal}(\mathbb{Q}(\zeta_n)/\mathbb{Q}) \cong (\mathbb{Z}/n)^\times,$$

que es *abeliano* de orden 2^m. Todo grupo de orden 2^m tiene una cadena de subgrupos

$$\{e\} = H_0 \subset H_1 \subset \cdots \subset H_m = G, \qquad [H_{i+1}:H_i] = 2,$$

porque un 2-grupo tiene centro no trivial y se procede por inducción tomando un elemento central de orden 2 (en el caso abeliano es aún más directo: por el teorema de estructura, G es producto de cíclicos de orden potencia de 2, y la cadena se lee de ahí). Aplicando la correspondencia de Galois —demostrada en este capítulo— se obtiene una torre de cuerpos fijos

$$\mathbb{Q} = K_0 \subset K_1 \subset \cdots \subset K_m = \mathbb{Q}(\zeta_n), \qquad [K_{i+1}:K_i] = 2,$$

y por el teorema 26.131, ζ_n es construible. □

Observación 26.133 (Por qué aquí sí basta el grado, y en general no). La advertencia del capítulo es correcta: que $[\mathbb{Q}(\alpha):\mathbb{Q}]$ sea potencia de 2 *no* implica que α sea construible. Tómese una raíz α de un polinomio irreducible de grado 4 cuyo grupo de Galois sea S_4 —por ejemplo $x^4 - x - 1$—. Entonces $[\mathbb{Q}(\alpha):\mathbb{Q}] = 4$ es potencia de 2, pero el cierre normal tiene grado $|S_4| = 24$, que no lo es, y se demuestra que no existe ninguna torre cuadrática que contenga a α. Lo que hace falta de verdad no es que el grado sea potencia de 2, sino que lo sea el grado del *cierre de Galois*.

Y ahí está la razón por la que el caso ciclotómico funciona: $\mathbb{Q}(\zeta_n)/\mathbb{Q}$ ya es de Galois y su grupo es abeliano, de modo que el grado del cierre normal coincide con el grado de la extensión. Es un caso feliz, no la regla, y esa distinción es exactamente lo que separa el enunciado correcto del incorrecto.

Ejemplo 26.134 (Los polígonos construibles hasta 100)**.** Los primos de Fermat conocidos son $3, 5, 17, 257, 65537$, y no se conoce ningún otro (se sabe que $2^{2^s}+1$ es compuesto para $5 \le s \le 32$). Los $n \le 100$ construibles son, por tanto,

$$3, 4, 5, 6, 8, 10, 12, 15, 16, 17, 20, 24, 30, 32, 34, 40, 48, 51, 60, 64, 68, 80, 85, 96,$$

y los primeros no construibles son 7, 9, 11, 13, 14, 18, 19. Obsérvese que 9 no lo es —$\varphi(9)=6$— aunque 3 sí: elevar un primo de Fermat al cuadrado destruye la propiedad, y por eso el enunciado exige primos *distintos*. El 65 537-ágono es construible, y en 1894 Johann Gustav Hermes dedicó diez años a escribir la construcción; el manuscrito, de más de doscientas páginas, se conserva en Gotinga.

Teorema 26.135 (Teorema 90 de Hilbert)**.** *Sea L/K una extensión cíclica finita de grado n, con $\mathrm{Gal}(L/K)=\langle\sigma\rangle$. Si $\alpha \in L$ cumple*

$$N_{L/K}(\alpha) = \alpha\,\sigma(\alpha)\,\sigma^2(\alpha)\cdots\sigma^{n-1}(\alpha) = 1,$$

entonces existe $\beta \in L^\times$ con

$$\alpha = \frac{\beta}{\sigma(\beta)}.$$

Demostración. Defínanse los coeficientes

$$b_0 = 1, \qquad b_i = \alpha\,\sigma(\alpha)\cdots\sigma^{i-1}(\alpha) \quad (1 \le i \le n-1),$$

de modo que $b_n = N_{L/K}(\alpha) = 1$. Considérese, para $\theta \in L$, la expresión

$$\gamma(\theta) = \sum_{i=0}^{n-1} b_i\,\sigma^i(\theta).$$

Por el teorema de independencia lineal de caracteres de Dedekind–Artin, los automorfismos $\mathrm{id}, \sigma, \dots, \sigma^{n-1}$ son linealmente independientes sobre L, luego esta combinación no es idénticamente nula: existe θ con $\beta = \gamma(\theta) \neq 0$.

Ahora obsérvese que $\alpha\,\sigma(b_i) = \alpha\,\sigma(\alpha)\cdots\sigma^{i}(\alpha) = b_{i+1}$. Por tanto

$$\alpha\,\sigma(\beta) = \alpha\sum_{i=0}^{n-1}\sigma(b_i)\,\sigma^{i+1}(\theta) = \sum_{i=0}^{n-1} b_{i+1}\,\sigma^{i+1}(\theta) = \sum_{k=1}^{n} b_k\,\sigma^k(\theta).$$

El término $k=n$ vale $b_n\sigma^n(\theta) = 1\cdot\theta = b_0\sigma^0(\theta)$, porque $\sigma^n = \mathrm{id}$ y $b_n = 1$. Así que la última suma es exactamente $\sum_{k=0}^{n-1} b_k\sigma^k(\theta) = \beta$. Es decir,

$$\alpha\,\sigma(\beta) = \beta \quad\Longleftrightarrow\quad \alpha = \frac{\beta}{\sigma(\beta)}. \qquad \square$$

Corolario 26.136 (Teoría de Kummer: las extensiones cíclicas son radicales)**.** *Sea K un cuerpo que contiene una raíz primitiva n-ésima de la unidad ζ, con $\operatorname{car} K \nmid n$. Entonces toda extensión cíclica L/K de grado n es de la forma $L = K(\sqrt[n]{a})$ para algún $a \in K$.*

Demostración. Sea σ un generador del grupo de Galois y aplíquese el teorema 26.135 a $\alpha = \zeta^{-1}$, cuya norma es $(\zeta^{-1})^n = 1$ porque $\zeta \in K$ queda fijo por σ. Se obtiene $\beta \in L^\times$ con $\zeta^{-1} = \beta/\sigma(\beta)$, es decir

$$\sigma(\beta) = \zeta\,\beta.$$

Entonces $\sigma(\beta^n) = \zeta^n \beta^n = \beta^n$, luego $a = \beta^n$ es fijo por σ y por tanto está en K. Además β no es fijo por ninguna potencia σ^k con $0 < k < n$ (pues $\sigma^k(\beta) = \zeta^k \beta \neq \beta$), de modo que β genera L sobre K: $L = K(\beta) = K(\sqrt[n]{a})$. □

Observación 26.137 (Esta es la pieza que hace funcionar Galois). El teorema fundamental dice que "resoluble por radicales" equivale a "grupo de Galois resoluble". La dirección difícil —de grupo resoluble a fórmula con radicales— consiste en descomponer el grupo en una torre de cocientes cíclicos y *convertir cada piso cíclico en una extracción de raíz*. Ese último paso es exactamente el corolario anterior, y su demostración es exactamente el teorema 90 de Hilbert. Sin él, la teoría de Galois demostraría que ciertas ecuaciones no se pueden resolver, pero no que las otras sí.

El nombre viene de que es el enunciado número 90 del *Zahlbericht* que Hilbert escribió en 1897; la versión moderna lo enuncia como $H^1(\mathrm{Gal}(L/K), L^\times) = 0$, la anulación del primer grupo de cohomología, y desde ahí se generaliza a la teoría de cuerpos de clases.

26.23. Separabilidad, clausura algebraica y un criterio para reconocer $\mathrm{Gal}(f) \cong S_p$

La palabra *separable* ha aparecido ya varias veces como hipótesis técnica, y la *cerradura algebraica* se ha usado para definir el grado de separabilidad sin haberla construido. Ambas cosas se enderezan aquí. Después viene el pago de una deuda mayor: el capítulo demostró caso por caso que ciertas quínticas concretas tienen grupo S_5, repitiendo cada vez el mismo razonamiento. Ese razonamiento es en realidad un teorema general sobre polinomios de grado primo, y demostrado una sola vez sirve para grados 5, 7, 11 y cualquier otro primo, donde ya no hay resolventes ni tablas que consultar.

Separabilidad: raíces distintas, y cómo verlo sin calcularlas

Definición 26.138 (Polinomio separable y extensión separable)**.** Un polinomio no constante $f \in K[x]$ es *separable* si en un cuerpo de descomposición suyo tiene $\deg f$ raíces *distintas*, es decir, si ninguna raíz es múltiple. Un elemento α algebraico sobre K es *separable sobre K* si su polinomio mínimo lo es, y una extensión algebraica L/K es *separable* si todo elemento de L es separable sobre K.

Proposición 26.139 (El criterio operativo)**.** *Sea $f \in K[x]$ no constante y f' su derivada formal. Entonces*

$$f \text{ es separable} \iff \gcd(f, f') = 1.$$

Si además f es irreducible, esto se simplifica: f es separable si y solo si $f' \neq 0$.

Demostración. Si α es raíz múltiple de f en un cuerpo de descomposición, entonces $f = (x-\alpha)^2 g$ y derivando, $f' = 2(x-\alpha)g + (x-\alpha)^2 g'$, de modo que $(x-\alpha)$ divide a f y a f': el máximo común divisor no es 1. Recíprocamente, si $f = (x-\alpha)h$ con $h(\alpha) \neq 0$, entonces $f'(\alpha) = h(\alpha) \neq 0$ y $(x-\alpha)$ no divide a f'. Como el máximo común divisor se calcula por el algoritmo de Euclides y este no cambia al agrandar el cuerpo, $\gcd(f, f')$ es el mismo en $K[x]$ que en el cuerpo de descomposición: por eso el criterio se puede aplicar *sin calcular ninguna raíz*, que es lo que lo hace útil.

Para f irreducible, $\gcd(f, f')$ solo puede ser 1 o un asociado de f; lo segundo obliga a $f \mid f'$, imposible por grados salvo que $f' = 0$. □

Ejemplo 26.140 (Dónde todo es separable y dónde no)**. (a) Característica 0.** Si $\operatorname{car} K = 0$ y f es irreducible de grado $n \geq 1$, entonces f' tiene grado $n-1$ y no es el polinomio nulo, porque el coeficiente principal $n\,a_n$ no se anula al ser $n \neq 0$ en K. Por la proposición, f es separable. Conclusión: sobre $\mathbb{Q}$, sobre $\mathbb{R}$ y sobre $\mathbb{C}$ la palabra "separable" no aporta ninguna restricción, y todo el capítulo se pudo escribir sin ella.

(b) Cuerpos finitos. Si K es finito de característica p, el Frobenius $a \mapsto a^p$ es inyectivo por ser homomorfismo de cuerpos, y una aplicación inyectiva de un conjunto finito en sí mismo es sobreyectiva: todo elemento es potencia p-ésima. Eso impide el mecanismo del apartado siguiente y hace que todo irreducible sea separable. Junto con (a), esto cubre los dos únicos mundos en que trabaja este libro.

(c) El contraejemplo estándar, en característica p. Sea $K = \mathbb{F}_p(t)$, el cuerpo de funciones racionales en una indeterminada t sobre $\mathbb{F}_p$, y sea

$$f(x) = x^p - t \in K[x].$$

Su derivada formal es $f' = p\,x^{p-1} = 0$, así que el criterio dice de inmediato que f no es separable. Y se ve directamente: si θ es una raíz en un cuerpo de descomposición, entonces $\theta^p = t$ y por el sueño del estudiante

$$x^p - t = x^p - \theta^p = (x-\theta)^p.$$

Una sola raíz, con multiplicidad p. Además f es irreducible sobre K —t no es potencia p-ésima en $\mathbb{F}_p(t)$, porque las potencias p-ésimas son las funciones de t^p—, de modo que se trata de un irreducible con todas sus raíces iguales. La extensión $K(\theta)/K$ tiene grado p y un único K-automorfismo, el trivial: no hay ningún otro sitio adonde mandar θ. La correspondencia de Galois se queda sin nada que corresponder.

La clausura algebraica: el sitio donde caben todas las raíces a la vez

Definición 26.141 (Cuerpo algebraicamente cerrado y clausura algebraica)**.** Un cuerpo Ω es *algebraicamente cerrado* si todo polinomio no constante de $\Omega[x]$ tiene alguna raíz en Ω, y por tanto factoriza en lineales. Una *clausura algebraica* (o *cierre algebraico*) de K es una extensión $\overline{K}/K$ que es algebraica y con $\overline{K}$ algebraicamente cerrado.

Teorema 26.142 (Existencia y unicidad, enunciadas)**.** *Todo cuerpo K admite una clausura algebraica, y dos clausuras algebraicas cualesquiera de K son isomorfas por un isomorfismo que fija K punto a punto.*

Idea clave

La existencia se demuestra con el **lema de Zorn**, y conviene saber por qué hace falta. Iterando la construcción de Kronecker se adjuntan las raíces de un polinomio, luego las de otro, y así sucesivamente; pero los polinomios de $K[x]$ pueden ser infinitos y cada extensión trae consigo polinomios nuevos, de modo que el proceso no termina en un número finito de pasos ni en una sucesión numerable de ellos. Zorn ordena esa torre transfinita: se considera el conjunto de las extensiones algebraicas de K dentro de un conjunto suficientemente grande fijado de antemano —ese detalle técnico, que evita hablar de la "clase de todas las extensiones", es de Artin—, se ordena por inclusión, se observa que toda cadena tiene cota superior (la unión) y se toma un elemento maximal. Un maximal no admite extensiones algebraicas propias, y eso es exactamente ser algebraicamente cerrado.

La unicidad, en cambio, es el mismo argumento de extensión de isomorfismos que da la unicidad del cuerpo de descomposición, aplicado con Zorn a las extensiones parciales. Y de nuevo: es única salvo isomorfismo, pero el isomorfismo no es único, y ahí está toda la teoría de Galois.

Observación 26.143 (Para qué se necesita: el grado de separabilidad)*.* El grado de separabilidad $[L:K]_s$ se definió como el número de K-inmersiones de L en una cerradura algebraica, y esa definición no se sostiene sin el teorema anterior. Hacen falta las dos mitades. La existencia garantiza que hay *algún* sitio donde contar las inmersiones: un recipiente que contiene todas las raíces de todos los polinomios mínimos a la vez, de modo que ninguna imagen posible se quede fuera por falta de espacio. La unicidad garantiza que el número no depende del recipiente elegido: si $\overline{K}$ y $\overline{K}'$ son dos clausuras, el isomorfismo entre ellas que fija K pone en biyección las inmersiones $L \to \overline{K}$ con las $L \to \overline{K}'$. Solo entonces $[L:K]_s$ es un invariante de la extensión y tiene sentido la factorización $[L:K] = [L:K]_s \cdot p^m$.

Un aviso de tamaño. $\overline{\mathbb{Q}}$, la clausura algebraica de $\mathbb{Q}$, es *numerable* y por tanto mucho más pequeña que $\mathbb{C}$: los números algebraicos son numerables porque los polinomios con coeficientes racionales lo son y cada uno aporta finitas raíces. Así que $\mathbb{C}$ es una clausura algebraica de $\mathbb{R}$ pero no de $\mathbb{Q}$, aunque contenga una: casi todo número complejo es trascendente.

Un criterio de grado primo

Teorema 26.144 (Grupo de Galois de un irreducible de grado primo con exactamente dos raíces no reales)**.** *Sea p un número primo y sea $f \in \mathbb{Q}[x]$ irreducible de grado p. Si f tiene exactamente $p-2$ raíces reales —equivalentemente, exactamente dos raíces complejas no reales, que serán conjugadas entre sí—, entonces*

$$\mathrm{Gal}(f) \;\cong\; S_p.$$

Demostración. Sea $L \subset \mathbb{C}$ el cuerpo de descomposición de f sobre $\mathbb{Q}$ y $G = \mathrm{Gal}(L/\mathbb{Q})$. Como

$\operatorname{car}\mathbb{Q}=0$, f es separable y sus p raíces $\alpha_1,\dots,\alpha_p$ son distintas; cada $\sigma\in G$ las permuta, y σ queda determinado por esa permutación porque las raíces generan L. Eso identifica G con un subgrupo de S_p, y todo el trabajo consiste en demostrar que es el subgrupo total. Se hace en cuatro pasos.

Paso 1: G es transitivo, luego p divide a $|G|$. Sean α_i y α_j dos raíces. Como f es irreducible, es el polinomio mínimo de ambas, y por la unicidad del cuerpo de descomposición existe un $\mathbb{Q}$-isomorfismo $\mathbb{Q}(\alpha_i)\to\mathbb{Q}(\alpha_j)$ con $\alpha_i\mapsto\alpha_j$, que se levanta a un automorfismo de L. Por tanto la acción de G sobre las raíces tiene una sola órbita, de tamaño p. Por el teorema de órbita–estabilizador (capítulo 25), $p=|G|/|G_{\alpha_1}|$ divide a $|G|$.

Paso 2: G contiene un p-ciclo. Por el teorema de Cauchy, un grupo cuyo orden es divisible por el primo p contiene un elemento σ de orden exactamente p. Ahora bien, en S_p un elemento de orden p tiene que ser un p-ciclo: su descomposición en ciclos disjuntos tiene longitudes cuyo mínimo común múltiplo es p, luego alguna longitud es múltiplo de p; y como las longitudes suman a lo sumo p, hay un único ciclo y ocupa todos los símbolos.

Paso 3: la conjugación compleja es una transposición. Por hipótesis hay exactamente dos raíces no reales; al ser f de coeficientes reales, son conjugadas la una de la otra, digamos α_{p-1} y $\alpha_p=\overline{\alpha_{p-1}}$. La conjugación compleja $c:\mathbb{C}\to\mathbb{C}$ es un automorfismo de cuerpos que fija $\mathbb{Q}$ y permuta el conjunto de raíces de f; en consecuencia aplica $L=\mathbb{Q}(\alpha_1,\dots,\alpha_p)$ en sí mismo, y su restricción $\tau=c|_L$ pertenece a G. Esa restricción fija las $p-2$ raíces reales e intercambia las otras dos: es una transposición. Nótese dónde se usa la hipótesis con precisión —si hubiera cuatro raíces no reales, τ sería un producto de dos transposiciones y el argumento se vendría abajo—.

Paso 4: un p-ciclo y una transposición generan S_p cuando p es primo. Sea $\tau=(a\ b)$ la transposición del paso 3 y σ el p-ciclo del paso 2. Como p es primo, *toda* potencia σ^k con $1\le k\le p-1$ tiene también orden p y es por tanto otro p-ciclo; entre esas potencias hay una que lleva a a b, pues σ recorre todos los símbolos. Sustituyendo σ por esa potencia y renombrando los símbolos como $a=1,\ b=2,\ \sigma(2)=3,\ \dots$, se puede suponer

$$\sigma=(1\ 2\ 3\ \cdots\ p),\qquad \tau=(1\ 2).$$

Conjugando repetidamente, $\sigma^i\tau\sigma^{-i}=(1+i\quad 2+i)$ con los índices leídos módulo p, de modo que el subgrupo generado contiene todas las transposiciones de consecutivos $(1\ 2),(2\ 3),\dots,(p-1\ \ p)$. Estas generan S_p, porque cualquier transposición $(i\ j)$ se obtiene componiendo consecutivas y toda permutación es producto de transposiciones. Luego $\langle\sigma,\tau\rangle=S_p$, y como $\langle\sigma,\tau\rangle\subseteq G\subseteq S_p$, se concluye $G=S_p$. □

Cuidado

[Dónde se usa que p sea primo, y qué pasa si no lo es] La primalidad interviene dos veces, y ninguna es cosmética. En el paso 2, para que "orden p" implique "p-ciclo"; en el paso 4, para que alguna potencia del ciclo lleve a a b sin dejar de ser un ciclo de longitud máxima. En grado compuesto el enunciado es sencillamente falso: en S_4, el 4-ciclo $\sigma=(1\,2\,3\,4)$ y la transposición $\tau=(1\,3)$ generan un grupo de orden 8 —el diédrico D_4, las simetrías del cuadrado con los vértices numerados en orden— y no S_4, que tiene 24 elementos. La transposición "salta" un vértice y por eso no rompe la simetría. En cambio, si se toma $\tau=(1\,2)$, el grupo generado sí es todo S_4. Con p primo

esa distinción desaparece: cualquier transposición sirve.

Ejemplo 26.145 (Una quíntica y un polinomio de grado siete, verificados)**. (a)** $f(x) = x^5 - 10x + 5$**.** *Irreducibilidad:* Eisenstein con $p = 5$, pues 5 divide a los coeficientes $0, 0, 0, -10, 5$ y 25 no divide al término independiente 5. *Conteo de raíces reales:* $f'(x) = 5x^4 - 10$ se anula exactamente en $x = \pm\sqrt[4]{2}$, y los valores críticos son

$$f(\sqrt[4]{2}) = 5 - 8\sqrt[4]{2} = -4{,}5137\ldots < 0, \qquad f(-\sqrt[4]{2}) = 5 + 8\sqrt[4]{2} = 14{,}5137\ldots > 0.$$

Un polinomio de grado impar con dos puntos críticos y valores críticos de signos opuestos corta al eje exactamente tres veces. Se confirma con Bolzano puro y aritmética entera: $f(-2) = -7$, $f(-1) = 14$, $f(0) = 5$, $f(1) = -4$, $f(2) = 17$, con cambios de signo en $(-2, -1)$, $(0, 1)$ y $(1, 2)$ y sin ningún cambio más. Numéricamente las cinco raíces son

$$-1{,}885961, \quad 0{,}503227, \quad 1{,}621725, \quad -0{,}119496 \pm 1{,}798421\,\mathrm{i},$$

tres reales y un par conjugado, es decir $p - 2 = 3$ reales. El teorema 26.144 da $\mathrm{Gal}(f) \cong S_5$, de orden 120; y como S_5 no es soluble, la ecuación $x^5 - 10x + 5 = 0$ no se resuelve por radicales.

(b) $g(x) = x^7 - 16x^5 - 16x^3 + 256x - 2$**, de grado 7.** Aquí el interés es que ninguna tabla de subgrupos transitivos ni ninguna resolvente resulta practicable, y el teorema decide igual. *Irreducibilidad:* Eisenstein con $p = 2$, pues 2 divide a 16, 16, 256 y 2, mientras que 4 no divide al término independiente -2. *Conteo:* el polinomio es una perturbación de $x(x^2 - 4)(x^2 - 16)(x^2 + 4) = x^7 - 16x^5 - 16x^3 + 256x$, cuyas raíces reales son $0, \pm 2, \pm 4$; restar 2 las mueve muy poco y no destruye ninguna, y en efecto las raíces de g son

$$-3{,}999739, \quad -2{,}002598, \quad 0{,}007813, \quad 1{,}997390, \quad 4{,}000260, \quad -0{,}001562 \pm 2{,}000004\,\mathrm{i},$$

esto es, exactamente $5 = 7 - 2$ raíces reales y un par conjugado. Por el teorema, $\mathrm{Gal}(g) \cong S_7$, de orden 5040. Obsérvese la economía: dos hechos verificables a mano —un criterio de irreducibilidad y un recuento de cortes con el eje— determinan por completo un grupo de cinco mil elementos.

Cuidado

[El criterio es suficiente, nunca necesario] Que el recuento de raíces reales no salga $p - 2$ no permite concluir nada. El ejemplo canónico ya apareció en el capítulo: $x^5 - x - 1$ es irreducible y tiene una sola raíz real, de modo que el paso 3 no produce transposición alguna; y sin embargo su grupo también es S_5. Cuando el criterio no se aplica hay que recurrir al discriminante, a la factorización módulo varios primos (el método de Dedekind) o a las resolventes. Y en la otra dirección, tampoco vale invertir la lectura del teorema de Abel–Ruffini: la mayoría de las quínticas no solubles no cumplen esta hipótesis, simplemente porque tener exactamente dos raíces no reales es una condición bastante particular.

26.24. Ejercicios resueltos

Ejercicio 26.1. Calcula $[\mathbb{Q}(\sqrt{2},\sqrt{3}) : \mathbb{Q}]$, da una base, y demuestra que $\mathbb{Q}(\sqrt{2},\sqrt{3}) = \mathbb{Q}(\sqrt{2}+\sqrt{3})$ (elemento primitivo explícito).

Solución. $\sqrt{3} \notin \mathbb{Q}(\sqrt{2})$: si $\sqrt{3} = a + b\sqrt{2}$, al cuadrar, $3 = a^2 + 2b^2 + 2ab\sqrt{2}$ fuerza $ab = 0$, y ambos casos ($\sqrt{3} \in \mathbb{Q}$ o $\sqrt{3/2} \in \mathbb{Q}$) son absurdos. Así $x^2 - 3$ sigue irreducible sobre $\mathbb{Q}(\sqrt{2})$ y la torre da $[\mathbb{Q}(\sqrt{2},\sqrt{3}) : \mathbb{Q}] = 4$, base $\{1, \sqrt{2}, \sqrt{3}, \sqrt{6}\}$. Sea $\gamma = \sqrt{2} + \sqrt{3}$: $\gamma^3 = 11\sqrt{2} + 9\sqrt{3}$, luego $\gamma^3 - 9\gamma = 2\sqrt{2}$: $\sqrt{2} \in \mathbb{Q}(\gamma)$, y entonces $\sqrt{3} = \gamma - \sqrt{2} \in \mathbb{Q}(\gamma)$: $\mathbb{Q}(\gamma)$ contiene ambos generadores, y como está contenido en el cuerpo, coinciden. ■

Ejercicio 26.2. Determina $\mathrm{Gal}(\mathbb{Q}(\sqrt{2},\sqrt{3})/\mathbb{Q})$ y *todos* los cuerpos intermedios vía la correspondencia.

Solución. Es el cuerpo de descomposición de $(x^2-2)(x^2-3)$: Galois de grado 4. Los automorfismos eligen signos independientes: $\sigma : \sqrt{2} \mapsto -\sqrt{2}$, $\tau : \sqrt{3} \mapsto -\sqrt{3}$; $G = \{\mathrm{id}, \sigma, \tau, \sigma\tau\} \cong \mathbb{Z}_2 \times \mathbb{Z}_2$ (Klein). Subgrupos propios: $\langle\sigma\rangle, \langle\tau\rangle, \langle\sigma\tau\rangle$, todos normales (abeliano). Cuerpos fijos: σ fija $\sqrt{3}$: $E^{\langle\sigma\rangle} = \mathbb{Q}(\sqrt{3})$; τ fija $\sqrt{2}$: $\mathbb{Q}(\sqrt{2})$; $\sigma\tau$ fija $\sqrt{2}\sqrt{3} = \sqrt{6}$: $\mathbb{Q}(\sqrt{6})$. Por la biyección del Teorema 26.54, estos tres son *todos* los cuerpos intermedios propios —afirmación imposible de verificar a mano sin el teorema—. ■

Ejercicio 26.3. Halla el polinomio mínimo de $\alpha = \sqrt{2} + \sqrt{3}$ sobre $\mathbb{Q}$ y factorízalo con la órbita de Galois.

Solución. $\alpha^2 = 5 + 2\sqrt{6} \Rightarrow (\alpha^2 - 5)^2 = 24$: $m(x) = x^4 - 10x^2 + 1$. Es el mínimo porque $[\mathbb{Q}(\alpha) : \mathbb{Q}] = 4$ (ejercicio anterior). La órbita de α bajo el grupo de Klein es $\{\pm\sqrt{2} \pm \sqrt{3}\}$, y en efecto

$$x^4 - 10x^2 + 1 = \prod_{\pm,\pm} \left(x - (\pm\sqrt{2} \pm \sqrt{3})\right) :$$

el polinomio mínimo de α es siempre $\prod_{\beta \in \text{órbita}} (x - \beta)$ —sus coeficientes son simétricos bajo G, luego racionales—. Este es el método general para calcular mínimos: aplicar el grupo y multiplicar. ■

Ejercicio 26.4. Describe $\mathrm{Gal}(\mathbb{Q}(\zeta_5)/\mathbb{Q})$, su retículo de subgrupos, y encuentra el único cuerpo intermedio cuadrático. Deduce que $\cos 72^\circ = \frac{\sqrt{5}-1}{4}$.

Solución. $[\mathbb{Q}(\zeta_5) : \mathbb{Q}] = \varphi(5) = 4$ y $G \cong (\mathbb{Z}_5)^\times = \{1,2,3,4\} \cong \mathbb{Z}_4$, generado por $\sigma : \zeta \mapsto \zeta^2$ (pues 2 genera: $2, 4, 3, 1$). $\mathbb{Z}_4$ tiene un único subgrupo propio, $\langle\sigma^2\rangle$ con $\sigma^2 : \zeta \mapsto \zeta^4 = \bar{\zeta}$ (la conjugación): su cuerpo fijo es el único intermedio, de grado 2: $\mathbb{Q}(\zeta + \bar{\zeta}) = \mathbb{Q}(2\cos 72^\circ)$. Calculemos $\eta = \zeta + \zeta^4$ y $\eta' = \zeta^2 + \zeta^3$: $\eta + \eta' = -1$ (suma de todas las raíces primitivas) y $\eta\eta' = \zeta^3 + \zeta^4 + \zeta^6 + \zeta^7 = \zeta + \zeta^2 + \zeta^3 + \zeta^4 = -1$. Así η, η' son raíces de $t^2 + t - 1 = 0$: $\eta = \frac{-1+\sqrt{5}}{2}$ (la raíz positiva, pues $\eta = 2\cos 72^\circ > 0$). El cuerpo intermedio es $\mathbb{Q}(\sqrt{5})$, y

$$\cos 72^\circ = \frac{\sqrt{5}-1}{4}.$$

Estas "sumas parciales de raíces" son los *períodos de Gauss*, y el mismo método (anidando subgrupos de $\mathbb{Z}_{16}$) construye el 17-ágono. ■

Ejercicio 26.5. Para $f = x^4 - 2$ con cuerpo de descomposición $E = \mathbb{Q}(\sqrt[4]{2}, i)$: calcula $[E : \mathbb{Q}]$, demuestra que $\mathrm{Gal}(E/\mathbb{Q}) \cong D_4$, y localiza en el retículo los cuerpos $\mathbb{Q}(i)$, $\mathbb{Q}(\sqrt{2})$ y $\mathbb{Q}(i\sqrt{2})$.

Solución. $x^4 - 2$ es Eisenstein: $[\mathbb{Q}(\sqrt[4]{2}) : \mathbb{Q}] = 4$, e $i \notin \mathbb{R} \supset \mathbb{Q}(\sqrt[4]{2})$ duplica: $[E : \mathbb{Q}] = 8$. Generadores: $\sigma : \sqrt[4]{2} \mapsto i\sqrt[4]{2}$, $i \mapsto i$ (orden 4) y $\tau =$ conjugación (orden 2), con $\tau\sigma\tau = \sigma^{-1}$: el diédrico D_4, las

simetrías del cuadrado formado por las raíces $\pm\sqrt[4]{2}, \pm i\sqrt[4]{2}$ en el plano complejo. Cuerpos cuadráticos intermedios = subgrupos de índice 2 (hay tres en D_4): $\langle\sigma\rangle$ fija i: $\mathbb{Q}(i)$; $\langle\sigma^2,\tau\rangle$ fija $\sqrt{2} = (\sqrt[4]{2})^2$ (nota: $\sigma(\sqrt{2}) = (i\sqrt[4]{2})^2 = -\sqrt{2}$, así que es σ^2 quien la fija junto con τ): $\mathbb{Q}(\sqrt{2})$; y $\langle\sigma^2,\sigma\tau\rangle$ fija $i\sqrt{2}$: $\mathbb{Q}(i\sqrt{2})$. Los tres son normales (índice 2) y en efecto son cuerpos de descomposición (x^2+1, x^2-2, x^2+2).■

Ejercicio 26.6. Demuestra que el heptágono regular no es construible, y que el polígono de 15 lados sí.

Solución. Heptágono: exige $\cos\frac{2\pi}{7}$, con $[\mathbb{Q}(\cos\frac{2\pi}{7}) : \mathbb{Q}] = \frac{\varphi(7)}{2} = 3$, no potencia de 2. *15 lados*: $\varphi(15) = \varphi(3)\varphi(5) = 8 = 2^3$ ✓; explícitamente, $15 = 3\cdot 5$ con 3 y 5 primos de Fermat, y como $\gcd(3,5) = 1$, de $\frac{2}{5} - \frac{1}{3} = \frac{1}{15}$ se obtiene el ángulo $\frac{2\pi}{15}$ restando ángulos construibles: $\frac{2\pi}{15} = 2\cdot\frac{2\pi}{5} - \frac{2\pi}{3}$. Construibles el pentágono (ejercicio del $\cos 72°$) y el triángulo, construible el 15-ágono, tal como ya sabía Euclides. ■

Ejercicio 26.7. Explica por qué x^5-1 y x^5-2 *sí* son solubles por radicales, sin contradecir a Abel–Ruffini.

Solución. x^5-1: grupo $(\mathbb{Z}_5)^\times \cong \mathbb{Z}_4$, abeliano, soluble (y las raíces son explícitas vía $\cos 72°$). x^5-2: cuerpo de descomposición $\mathbb{Q}(\sqrt[5]{2}, \zeta_5)$, de grado 20; su grupo es el producto semidirecto $\mathbb{Z}_5 \rtimes \mathbb{Z}_4$ (Frobenius F_{20}), con cadena $\{e\} \trianglelefteq \mathbb{Z}_5 \trianglelefteq F_{20}$ de cocientes $\mathbb{Z}_5$ y $\mathbb{Z}_4$: soluble, y en efecto la raíz es el radical $\sqrt[5]{2}$. Abel–Ruffini no dice "ninguna quíntica se resuelve", sino "no hay fórmula universal": la ecuación *general* tiene grupo S_5, pero cada quíntica particular tiene el grupo que tiene, y si es soluble (como $\mathbb{Z}_4$ o F_{20}), hay radicales para ella. La frontera exacta la traza el grupo, no el grado. ■

Ejercicio 26.8. [Síntesis final] Traza el arco completo del libro: ¿qué papel jugó cada capítulo en la demostración de que $x^5 - 4x + 2$ no se resuelve por radicales?

Solución. Cálculo (cap. 1): el análisis de f' y el teorema de Bolzano contaron las raíces reales (paso 2). *Álgebra lineal (cap. 2)*: el grado de una extensión es una dimensión de espacio vectorial, y la multiplicatividad de la torre —pura base por base— gobierna todos los cómputos de grados. *Grupos (cap. 7)*: Lagrange, Cauchy, órbita–estabilizador dieron el 5-ciclo; la conjugación compleja, la trasposición; la simplicidad de A_5, la no solubilidad. *Anillos y polinomios (cap. 7)*: Eisenstein dio la irreducibilidad; el cociente por irreducibles construyó los cuerpos donde viven las raíces. *Galois (cap. 8)*: la correspondencia tradujo "torre de radicales" en "cadena de subgrupos con cocientes abelianos", y la traducción cerró el círculo. Un problema de 300 años (1540–1830) resuelto no calculando más fuerte, sino *entendiendo la estructura*: esa es la lección de toda esta guía. ■

26.25. Ejercicios adicionales resueltos

Ejercicio 26.9. Grupo y cuerpos intermedios de x^4+1.

Solución. $\mathbb{Q}(\zeta_8)$, grado $\varphi(8) = 4$, grupo $(\mathbb{Z}_8)^\times \cong \mathbb{Z}_2^2$: tres cuadráticos: $\mathbb{Q}(\sqrt{2}), \mathbb{Q}(i), \mathbb{Q}(\sqrt{-2})$ (de $\zeta_8 \pm \zeta_8^{-1}$ y ζ_8^2). Todo automorfismo tiene orden 2: por eso x^4+1 factoriza mod *todo* primo. ■

Ejercicio 26.10. Polinomio mínimo de $2\cos\frac{2\pi}{7}$.

Solución. De $\zeta^6 + \cdots + 1 = 0$ dividido por ζ^3, con $\alpha = \zeta + \zeta^{-1}$: $\alpha^3 + \alpha^2 - 2\alpha - 1$, irreducible: grado $3 = \varphi(7)/2$: el heptágono inconstructible; discriminante 49: grupo $\mathbb{Z}_3$. ■

Ejercicio 26.11. Retículo de subcuerpos de $\mathbb{F}_{2^{12}}$.

Solución. $\mathbb{F}_{2^d}$, $d \mid 12$; cadenas maximales con grados $(2,2,3)$, $(2,3,2)$, $(3,2,2)$; $\mathbb{F}_{16} \cap \mathbb{F}_{64} = \mathbb{F}_4$ (mcd), generan $\mathbb{F}_{2^{12}}$ (mcm): el retículo de divisores, la correspondencia hecha aritmética. ■

Ejercicio 26.12. $\mathrm{Gal}(\mathbb{Q}(\sqrt[3]{2})/\mathbb{Q})$ trivial.

Solución. Las otras raíces no son reales: $\sqrt[3]{2} \mapsto \sqrt[3]{2}$: $|\,\mathrm{Gal}\,| = 1 < 3$: no normal; el déficit de simetrías mide las raíces ausentes. ■

Ejercicio 26.13. N y T de $1+\sqrt[3]{2}$; su polinomio mínimo.

Solución. Conjugados $1+\omega^k\sqrt[3]{2}$: $T = 3$, $N = \prod(1+\omega^k\sqrt[3]{2}) = 3$; mínimo $(x-1)^3 = 2$: x^3-3x^2+3x-3 (¡Eisenstein en 3!), con T y N leídos de los coeficientes. ■

Ejercicio 26.14. Grupos de x^3-4x+2 y x^3-3x+1 por el discriminante.

Solución. $\Delta = 148$ (no cuadrado): S_3; $\Delta = 81 = 9^2$: $\mathbb{Z}_3$ —las raíces $2\cos\frac{2\pi}{9}$ etc., parientes del nonágono: cúbicas con simetría extra, detectadas sin resolver—. ■

26.26. Taller de ejemplos explicados: grupos de Galois concretos

Idea clave

Intuición primero. La idea de Galois, en una frase: para entender las raíces de un polinomio, estudia sus *simetrías* (las formas de intercambiarlas que ninguna ecuación racional puede notar). Esas simetrías forman un grupo, y la *estructura* de ese grupo (¿se puede desarmar en pasos simples?) decide si el polinomio se resuelve con radicales. En los ejemplos verás grupos pequeños ($\mathbb{Z}_2$, S_3) para tocar la idea antes de la maquinaria pesada.

La idea de Galois es medir un polinomio por las *simetrías* de sus raíces. Veámoslo en casos pequeños, sin la maquinaria pesada.

26.26.1. Un caso de grado 2

Ejemplo 26.146. El grupo de Galois de x^2-2 sobre los racionales $\mathbb{Q}$.

Paso 1. Hallar las raíces: $x^2-2=0$ da $x=\sqrt{2}$ y $x=-\sqrt{2}$.

Paso 2. Preguntarse qué simetrías hay. Una "simetría" es una forma de mezclar las raíces sin que ninguna ecuación con coeficientes racionales lo note. Hay dos: la identidad (dejar todo igual) y el intercambio σ que manda $\sqrt{2} \mapsto -\sqrt{2}$. Este intercambio es legal porque $\sqrt{2}$ y $-\sqrt{2}$ son "indistinguibles" para $\mathbb{Q}$: ambas cumplen $x^2 = 2$.

Paso 3. Identificar el grupo. Son dos simetrías, y σ aplicada dos veces vuelve al inicio ($\sigma^2 = \mathrm{id}$). Ese grupo es $\mathbb{Z}_2$, el grupo más simple no trivial.

26.26.2. Un caso de grado 3 (más rico)

Ejemplo 26.147. El grupo de Galois de x^3-2 sobre $\mathbb{Q}$.

Paso 1. Las raíces son las tres raíces cúbicas de 2: la real $\sqrt[3]{2}$, y dos complejas $\sqrt[3]{2}\,\omega$ y $\sqrt[3]{2}\,\omega^2$, donde $\omega = e^{2\pi i/3}$ es una raíz cúbica de la unidad.

Paso 2. Contar las simetrías. Las simetrías pueden *permutar* las tres raíces entre sí, y resulta que las seis permutaciones posibles son todas legales. Seis simetrías de tres objetos: ese es el grupo S_3 (el grupo de todas las reordenaciones de tres cosas).

Paso 3. La conclusión profunda. S_3 tiene una estructura "escalonada" que se puede desmontar en pasos abelianos, y por eso x^3-2 *sí* se resuelve con radicales (raíces): la fórmula de las cúbicas existe. En cambio, para grado 5 el grupo puede ser S_5, que contiene un bloque indivisible (A_5), y por eso no hay fórmula general para la quíntica. Los grupos de Galois miden exactamente esa diferencia.

26.27. Ejercicios propuestos

Sin solución desarrollada: solo enunciado y, al final, la respuesta. Aquí muchas respuestas son un nombre ("el grupo de Klein", "sí, por Eisenstein") más que un número: comprueba que coincida el fondo del argumento, no la redacción.

26.1. Calcula $\mathrm{Gal}(\mathbb{Q}(\sqrt{2})/\mathbb{Q})$ y describe explícitamente sus elementos.

26.2. Calcula $\mathrm{Gal}(\mathbb{Q}(\sqrt{2},\sqrt{3})/\mathbb{Q})$ y di a qué grupo conocido es isomorfo.

26.3. ¿Cuáles son los cuerpos intermedios de $\mathbb{Q}(\sqrt{2},\sqrt{3})/\mathbb{Q}$? Relaciónalos con los subgrupos.

26.4. Calcula el grupo de Galois de x^3-2 sobre $\mathbb{Q}$.

26.5. Calcula el grupo de Galois de x^4-2 sobre $\mathbb{Q}$ y su orden.

26.6. Calcula el grupo de Galois de x^4+1 sobre $\mathbb{Q}$.

26.7. ¿Cuál es el grupo de Galois de $\mathbb{F}_{p^n}/\mathbb{F}_p$? Da un generador.

26.8. ¿Por qué la quíntica general no es resoluble por radicales? Nombra el grupo culpable.

26.9. Da un polinomio de grado 5 sobre $\mathbb{Q}$ con grupo de Galois S_5 y explica el argumento de las raíces reales.

26.10. ¿Es constructible con regla y compás el ángulo de 20°? ¿Y el pentágono regular?

26.11. Expresa $\cos(2\pi/5)$ en radicales.

26.12. ¿Qué polígonos regulares de n lados son constructibles? Cita el teorema.

Respuestas

1. $\mathbb{Z}/2\mathbb{Z}$: la identidad y $\sqrt{2}\mapsto-\sqrt{2}$. **2.** El grupo de Klein $\mathbb{Z}/2\mathbb{Z}\times\mathbb{Z}/2\mathbb{Z}$, de orden 4: cada signo se elige por separado. **3.** Tres propios: $\mathbb{Q}(\sqrt{2})$, $\mathbb{Q}(\sqrt{3})$ y $\mathbb{Q}(\sqrt{6})$, correspondientes a los tres subgrupos de orden 2; más $\mathbb{Q}$ y el total. **4.** S_3, de orden 6 (el cuerpo de descomposición $\mathbb{Q}(\sqrt[3]{2},\zeta_3)$ tiene grado 6). **5.** D_4, el grupo diédrico de orden 8. **6.** El grupo de Klein $\mathbb{Z}/2\mathbb{Z}\times\mathbb{Z}/2\mathbb{Z}$, de orden 4. **7.** Cíclico de orden n, generado por el Frobenius $x\mapsto x^p$. **8.** Porque S_5 no es resoluble: su serie derivada se estanca en A_5, que es simple no abeliano. **9.** x^5-4x+2 es irreducible por Eisenstein con $p=2$ y tiene exactamente tres raíces reales; la conjugación compleja da una transposición y el 5-ciclo viene de la irreducibilidad, y una transposición más un 5-ciclo generan S_5. **10.** 20° no: equivale a trisecar 60° y su polinomio mínimo tiene grado 3, que no es potencia de 2. El pentágono sí. **11.** $\cos(2\pi/5)=\dfrac{\sqrt{5}-1}{4}$. **12.** Aquellos con $n=2^k p_1\cdots p_r$ donde los p_i son primos de Fermat distintos (teorema de Gauss–Wantzel).

Capítulo 27

Teoría de la representación de grupos

Lo que necesitas traer

Del capítulo 23: grupos finitos, clases de conjugación y subgrupos normales. Del capítulo 7: matrices, traza, diagonalización, producto interno y bases ortonormales. Del capítulo 4: los complejos, porque las representaciones viven sobre $\mathbb{C}$.

Un grupo abstracto es difícil de manipular. La idea de la teoría de la representación es convertirlo en matrices: si se consigue realizar el grupo como un conjunto de matrices con el producto habitual, todo el álgebra lineal queda disponible para estudiarlo. Y lo notable es que casi toda la información del grupo sobrevive a esa traducción, concentrada en unos pocos números: las trazas de esas matrices, llamadas caracteres.

27.1. Representaciones, caracteres y ortogonalidad

Definición 27.1. Una *representación* de un grupo finito G sobre $\mathbb{C}$ es un homomorfismo $\rho : G \to \mathrm{GL}(V)$, y su *grado* es $\dim V$. Es *irreducible* si V no tiene subespacios propios invariantes. El *carácter* es $\chi_\rho(g) = \operatorname{tr} \rho(g)$, y es constante en cada clase de conjugación. Con el producto interno $\langle \chi, \psi \rangle = \frac{1}{|G|} \sum_g \chi(g)\overline{\psi(g)}$, los caracteres irreducibles forman un sistema ortonormal.

Idea clave

Qué dice esto en cristiano. Una representación es una manera de ver los elementos del grupo como transformaciones de un espacio. Descomponerla en irreducibles es como descomponer un vector en su base: el teorema de Maschke garantiza que sobre $\mathbb{C}$ toda representación de un grupo finito se parte en trozos irreducibles, y esos trozos son los ladrillos. Los caracteres son la herramienta que hace todo esto computable, porque reducen matrices enteras a un solo número por clase de conjugación y aun así determinan la representación por completo. Es una compresión brutal de información

que, milagrosamente, no pierde nada.

Regla: descomponer una representación en irreducibles

El procedimiento es enteramente mecánico una vez se tiene la tabla de caracteres:

1. **Cuenta las clases de conjugación** del grupo: ese número es exactamente el número de representaciones irreducibles.
2. **Halla las dimensiones** d_i usando que $\sum d_i^2 = |G|$ y que cada d_i divide a $|G|$; con esas dos restricciones casi siempre queda una única posibilidad.
3. **Construye la tabla** empezando por la trivial (todo unos) y la del signo si el grupo es de permutaciones.
4. **Comprueba la ortonormalidad**: cada fila debe tener norma 1 y ser ortogonal a las demás, lo que valida la tabla.
5. **Calcula el carácter de tu representación** evaluando trazas, que en las de permutación es simplemente contar puntos fijos.
6. **Proyecta**: la multiplicidad de la irreducible χ_i es $\langle \chi, \chi_i \rangle$, un número entero no negativo.
7. **Verifica** que $\sum m_i d_i$ reproduce el grado de la representación original.

Ejemplo 27.2 (La tabla de S_3 y una descomposición)**. (a) Las dimensiones.** S_3 tiene tres clases (identidad, tres trasposiciones, dos tres-ciclos), luego tres irreducibles. Sus dimensiones cumplen $\sum d_i^2 = 6$, y la única posibilidad con tres sumandos es $1+1+4$, es decir $d = 1, 1, 2$.

(b) La tabla. En el orden (identidad, trasposiciones, tres-ciclos), con tamaños de clase $1, 3, 2$:

	1	3	2
χ_{triv}	1	1	1
χ_{sgn}	1	-1	1
χ_{std}	2	0	-1

Las tres normas valen 1 y los tres productos cruzados 0, luego la tabla es correcta.

(c) Descomponer la representación de permutación. Su carácter cuenta puntos fijos: $\chi = (3, 1, 0)$. Proyectando, $\langle \chi, \chi_{\text{triv}} \rangle = 1$, $\langle \chi, \chi_{\text{sgn}} \rangle = 0$ y $\langle \chi, \chi_{\text{std}} \rangle = 1$. Luego la representación natural de S_3 sobre $\mathbb{C}^3$ se parte como trivial más estándar, con $1 \cdot 1 + 1 \cdot 2 = 3$: cuadra.

(d) Grupos mayores. S_4 tiene cinco clases y dimensiones $1, 1, 2, 3, 3$, con $1+1+4+9+9 = 24$. A_5 tiene cinco clases y dimensiones $1, 3, 3, 4, 5$, con $1 + 9 + 9 + 16 + 25 = 60$. En ambos casos la restricción $\sum d_i^2 = |G|$ deja muy poco margen y basta casi para adivinar la tabla.

Cuidado

[La tabla de caracteres no determina el grupo] D_4 y Q_8 tienen ambos orden 8, cinco clases de conjugación y *exactamente la misma tabla de caracteres*, con dimensiones $1,1,1,1,2$, y sin embargo no son isomorfos: D_4 tiene cinco elementos de orden 2 y Q_8 solo uno. La tabla es una huella dactilar muy buena y no es única. Segundo aviso: todo esto vale sobre $\mathbb{C}$ y para grupos *finitos*. Sobre cuerpos de característica p que divida a $|G|$, el teorema de Maschke falla y aparecen representaciones que no se descomponen, terreno de la llamada teoría modular, mucho más difícil.

Pausa para pensar

La teoría la creó Frobenius en 1896 respondiendo a una carta de Dedekind sobre un problema aparentemente ajeno: factorizar cierto determinante asociado a un grupo. En pocos años Burnside la usó para demostrar que todo grupo de orden $p^a q^b$ es soluble, un teorema puramente algebraico cuya única demostración conocida durante setenta años pasaba por los caracteres. Más tarde resultó ser el lenguaje natural de la mecánica cuántica: las partículas elementales se clasifican por representaciones irreducibles del grupo de simetría, y el espín no es más que el índice que etiqueta cuál. Wigner llamó a esto la irrazonable eficacia de la teoría de grupos en física.

27.2. Representaciones, matrices y equivalencia: la tabla de caracteres de S_3 construida desde las matrices

La definición de representación con la que abre el capítulo es correcta y comprimida: dice "homomorfismo $\rho : G \to \mathrm{GL}(V)$" y sigue adelante. Conviene desplegarla, porque cada palabra que aparece después —subrepresentación, equivalencia, suma directa, irreducible— se apoya en ella y ninguna se ha definido. Y conviene sobre todo construir una tabla de caracteres partiendo de matrices explícitas, y no de una fila que "sale de restar la trivial": la tabla de S_3 se usa ocho veces en el capítulo y aquí se levanta entera desde cero, con todas las comprobaciones hechas.

Las definiciones desplegadas

Definición 27.3 (Grupo lineal general y representación)**.** Sea V un espacio vectorial complejo de dimensión finita. El *grupo lineal general* $\mathrm{GL}(V)$ es el conjunto de las aplicaciones lineales biyectivas $V \to V$, que con la composición es un grupo, con neutro id_V. Si $\dim V = d$ y se fija una base, $\mathrm{GL}(V)$ se identifica con $\mathrm{GL}_d(\mathbb{C})$, el grupo de las matrices $d \times d$ complejas de determinante no nulo.

Una *representación* de un grupo G sobre V es un homomorfismo de grupos $\rho : G \to \mathrm{GL}(V)$, es decir, una asignación $g \mapsto \rho(g)$ que cumple

$$\rho(gh) = \rho(g)\,\rho(h) \quad \text{para todos } g, h \in G.$$

El *grado* —o *dimensión*— de la representación es $d = \dim V$.

Observación 27.4 (Consecuencias inmediatas de la definición). De $\rho(gh)=\rho(g)\rho(h)$ se sigue todo lo demás sin hipótesis adicionales. Tomando $g=h=e$ se obtiene $\rho(e)=\rho(e)^2$, y como $\rho(e)$ es invertible, $\rho(e)=\mathrm{id}_V$. Y entonces $\rho(g)\rho(g^{-1})=\rho(e)=\mathrm{id}_V$, luego $\rho(g^{-1})=\rho(g)^{-1}$: no hay que exigirlo, sale gratis. Además, si G es finito de orden n, cada $\rho(g)$ cumple $\rho(g)^{|g|}=\mathrm{id}_V$, de modo que su polinomio mínimo divide a $x^{|g|}-1$, que tiene raíces simples: *toda* matriz de una representación de un grupo finito es diagonalizable y sus autovalores son raíces de la unidad. Ese hecho se usa constantemente en el capítulo, y es una línea de álgebra lineal.

Definición 27.5 (Representación matricial)**.** Fijada una base $\mathcal{B}=(v_1,\dots,v_d)$ de V, cada $\rho(g)$ tiene una matriz $R(g)=(r_{ij}(g))\in \mathrm{GL}_d(\mathbb{C})$ determinada por

$$\rho(g)\,v_j=\sum_{i=1}^{d} r_{ij}(g)\,v_i,$$

es decir, la columna j de $R(g)$ recoge las coordenadas de la imagen de v_j. La aplicación $R:G\to\mathrm{GL}_d(\mathbb{C})$ es un homomorfismo y se llama la *representación matricial* asociada a ρ y a $\mathcal{B}$. Cambiar de base sustituye $R(g)$ por $P^{-1}R(g)P$ con P la matriz de cambio, *la misma P para todos los g a la vez*.

Definición 27.6 (Morfismo y equivalencia de representaciones)**.** Sean (ρ_V,V) y (ρ_W,W) representaciones de G. Una aplicación lineal $f:V\to W$ es un *morfismo de representaciones* —o *entrelazador*— si

$$f\circ\rho_V(g)=\rho_W(g)\circ f\qquad\text{para todo } g\in G,$$

es decir, si conmuta con la acción. Las dos representaciones son *equivalentes*, o *isomorfas*, y se escribe $V\cong W$, si existe un morfismo biyectivo. En términos matriciales: R y S son equivalentes si existe una única matriz invertible P con

$$S(g)=P^{-1}R(g)P\qquad\text{para todo } g\in G.$$

Cuidado

El cuantificador es la clave: se exige *una sola* P que valga simultáneamente para los $|G|$ elementos. Dos representaciones cuyas matrices sean conjugadas elemento a elemento, con una P distinta para cada g, no tienen por qué ser equivalentes. De hecho, dos representaciones son equivalentes si y sólo si tienen el mismo carácter, y eso es un teorema con contenido, no una trivialidad.

Definición 27.7 (Subrepresentación y suma directa)**.** Un subespacio $W\subseteq V$ es *invariante* si $\rho(g)(W)\subseteq W$ para todo $g\in G$; entonces la restricción $\rho|_W(g)=\rho(g)|_W$ es una representación de G sobre W, llamada *subrepresentación*. Siempre hay dos, 0 y V, llamadas triviales; la representación es *irreducible* si $V\neq 0$ y no hay ninguna otra.

Dadas (ρ_V,V) y (ρ_W,W), su *suma directa* es la representación sobre $V\oplus W$ definida por $(\rho_V\oplus\rho_W)(g)(v,w)=(\rho_V(g)v,\rho_W(g)w)$. En matrices, con una base formada por una de V

seguida de una de W, es la matriz por bloques

$$\begin{pmatrix} R_V(g) & 0 \\ 0 & R_W(g) \end{pmatrix},$$

y su carácter es $\chi_{V\oplus W} = \chi_V + \chi_W$, porque la traza de una diagonal por bloques es la suma de las trazas.

Idea clave

Con estas palabras, el teorema de Maschke dice algo muy concreto: para toda representación de un grupo finito sobre $\mathbb{C}$ existe *una* base en la que *todas* las matrices $R(g)$ son diagonales por bloques a la vez, y los bloques ya no se pueden partir más. Descomponer una representación es literalmente encontrar esa base. Los caracteres permiten averiguar cuántos bloques hay y de qué tamaño sin llegar a encontrarla nunca, y por eso son útiles: dan la respuesta saltándose el cálculo.

La tabla de caracteres de S_3, levantada desde las matrices

Ejemplo 27.8 (Las tres representaciones irreducibles de S_3, explícitas)**.** El grupo S_3 tiene seis elementos: e, las tres transposiciones (12), (13), (23) y los dos tres-ciclos (123), (132). Se construyen las tres irreducibles una a una.

(a) La trivial. $\rho_{\text{triv}}(\sigma) = 1$ para todo σ, sobre $V = \mathbb{C}$. Es homomorfismo trivialmente, y es irreducible por tener dimensión 1: un espacio de dimensión 1 no tiene subespacios propios no nulos. Su carácter es $\chi_{\text{triv}} = (1, 1, 1)$ en las clases e, transposiciones, tres-ciclos.

(b) El signo. $\rho_{\text{sgn}}(\sigma) = \operatorname{sgn}(\sigma) \in \{\pm 1\}$, también sobre $\mathbb{C}$. Es homomorfismo porque la signatura lo es. Vale 1 en e, -1 en cada transposición y $+1$ en los tres-ciclos, que son producto de dos transposiciones. Carácter: $\chi_{\text{sgn}} = (1, -1, 1)$.

(c) La estándar, construida. Sea $\mathbb{C}^3$ con base e_1, e_2, e_3 y sea σ actuando por $\sigma \cdot e_i = e_{\sigma(i)}$: eso es la *representación por permutación*, de grado 3. El subespacio

$$V_{\text{est}} = \{(x_1, x_2, x_3) \in \mathbb{C}^3 : x_1 + x_2 + x_3 = 0\}$$

es invariante, porque permutar coordenadas no altera su suma, y tiene dimensión 2. Tómese la base

$$f_1 = e_1 - e_2, \qquad f_2 = e_2 - e_3.$$

Se calculan las matrices. Para $\sigma = (12)$: $f_1 = e_1 - e_2 \mapsto e_2 - e_1 = -f_1$, y $f_2 = e_2 - e_3 \mapsto e_1 - e_3 = f_1 + f_2$. Para $\sigma = (123)$, que manda $1 \mapsto 2$, $2 \mapsto 3$, $3 \mapsto 1$: $f_1 \mapsto e_2 - e_3 = f_2$, y $f_2 \mapsto e_3 - e_1 = -(f_1 + f_2)$. Poniendo cada imagen en columna:

$$R(e) = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad R((12)) = \begin{pmatrix} -1 & 1 \\ 0 & 1 \end{pmatrix}, \quad R((13)) = \begin{pmatrix} 0 & -1 \\ -1 & 0 \end{pmatrix},$$

$$R((23)) = \begin{pmatrix} 1 & 0 \\ 1 & -1 \end{pmatrix}, \quad R((123)) = \begin{pmatrix} 0 & -1 \\ 1 & -1 \end{pmatrix}, \quad R((132)) = \begin{pmatrix} -1 & 1 \\ -1 & 0 \end{pmatrix}.$$

Son matrices enteras, y se comprueba que R es homomorfismo: por ejemplo $R((12))R((23)) = \begin{pmatrix}-1 & 1\\ 0 & 1\end{pmatrix}\begin{pmatrix}1 & 0\\ 1 & -1\end{pmatrix} = \begin{pmatrix}0 & -1\\ 1 & -1\end{pmatrix} = R((123))$, y en efecto $(12)\circ(23) = (123)$ con la convención de componer de derecha a izquierda. Los órdenes cuadran: R de una transposición tiene cuadrado la identidad, y R de un tres-ciclo tiene cubo la identidad. Además $\det R(\sigma) = \operatorname{sgn}(\sigma)$ en los seis casos.

Proposición 27.9 (La estándar es irreducible, sin usar caracteres)**.** *La representación de S_3 sobre V_{est} es irreducible.*

Demostración. Un subespacio invariante propio no nulo tendría dimensión 1, luego sería una recta $\mathbb{C}v$ con v autovector de *todas* las matrices a la vez. Ahora bien, los autovalores de $R((12)) = \begin{pmatrix}-1 & 1\\ 0 & 1\end{pmatrix}$ son -1 y 1, con autovectores $(1,0)$ y $(1,2)$ salvo escalares, ambos reales. Y el polinomio característico de $R((123)) = \begin{pmatrix}0 & -1\\ 1 & -1\end{pmatrix}$ es $\lambda^2+\lambda+1$, cuyas raíces son las dos raíces cúbicas primitivas de la unidad, con autovectores de coordenadas no reales. Ninguna de las dos rectas de $R((12))$ es autovector de $R((123))$, luego no hay recta invariante común. □

Ejemplo 27.10 (La tabla completa, con todas las comprobaciones)**. (a) Las clases.** Dos permutaciones de S_n son conjugadas si y sólo si tienen el mismo tipo cíclico, así que las clases son $\{e\}$, las tres transposiciones y los dos tres-ciclos, de tamaños 1, 3 y 2. Hay tres clases, luego tres irreducibles, y la única solución de $d_1^2+d_2^2+d_3^2 = 6$ en enteros positivos es $1+1+4$: dimensiones 1, 1, 2. Las tres construidas en el ejemplo 27.8 las agotan.

(b) Las trazas. Tomando la traza de las matrices anteriores: $\operatorname{tr} R(e) = 1+1 = 2$; $\operatorname{tr} R((12)) = -1+1 = 0$, y lo mismo para (13) y (23), como debe ser por ser función de clase; $\operatorname{tr} R((123)) = 0+(-1) = -1$, igual que para (132). Luego $\chi_{\text{est}} = (2,0,-1)$. La tabla queda

	e	(12)	(123)
tamaño de la clase	1	3	2
χ_{triv}	1	1	1
χ_{sgn}	1	-1	1
χ_{est}	2	0	-1

(c) Ortogonalidad por filas. Con $\langle\chi,\psi\rangle = \frac{1}{6}\sum_\sigma \chi(\sigma)\overline{\psi(\sigma)}$, agrupando por clases y pesando con su tamaño:

$$\langle\chi_{\text{triv}},\chi_{\text{triv}}\rangle = \tfrac{1\cdot1+3\cdot1+2\cdot1}{6} = 1, \qquad \langle\chi_{\text{sgn}},\chi_{\text{sgn}}\rangle = \tfrac{1+3+2}{6} = 1,$$

$$\langle\chi_{\text{est}},\chi_{\text{est}}\rangle = \tfrac{1\cdot4+3\cdot0+2\cdot1}{6} = 1, \qquad \langle\chi_{\text{triv}},\chi_{\text{sgn}}\rangle = \tfrac{1-3+2}{6} = 0,$$

$$\langle\chi_{\text{triv}},\chi_{\text{est}}\rangle = \tfrac{2+0-2}{6} = 0, \qquad \langle\chi_{\text{sgn}},\chi_{\text{est}}\rangle = \tfrac{1\cdot2+3\cdot0+2\cdot(-1)}{6} = 0.$$

Las tres normas valen 1 —lo que reprueba la irreducibilidad de la estándar, ahora por caracteres— y los tres productos cruzados se anulan.

(d) Ortogonalidad por columnas. La suma $\sum_i \chi_i(g)\overline{\chi_i(h)}$ debe valer el orden del centralizador $|C_G(g)| = |G|/|\text{clase de } g|$ si $g \sim h$, y cero si no. En la columna de e: $1+1+4 = 6 = 6/1$. En la de las transposiciones: $1+1+0 = 2 = 6/3$. En la de los tres-ciclos: $1+1+1 = 3 = 6/2$. Y los cruces: e con (12) da $1-1+0=0$; e con (123) da $1+1-2=0$; (12) con (123) da $1-1+0=0$. Los seis valores son los previstos.

(e) Cierre de la cuenta. $\sum d_i^2 = 1+1+4 = 6 = |S_3|$. La tabla es correcta y está completa.

Ejemplo 27.11 (Tres descomposiciones hechas con esta tabla)**. (a) La representación por permutación.** Su carácter cuenta puntos fijos, porque la matriz de permutación tiene un 1 en la diagonal exactamente en las posiciones fijadas: $\chi = (3,1,0)$. Proyectando,

$$\langle \chi, \chi_{\text{triv}} \rangle = \tfrac{3+3+0}{6} = 1, \quad \langle \chi, \chi_{\text{sgn}} \rangle = \tfrac{3-3+0}{6} = 0, \quad \langle \chi, \chi_{\text{est}} \rangle = \tfrac{6+0+0}{6} = 1,$$

luego $\mathbb{C}^3 \cong \text{triv} \oplus \text{est}$, con $1+2=3$. Geométricamente: la recta de los vectores con las tres coordenadas iguales, más el plano de suma cero con el que se construyó la estándar.

(b) La regular. Su carácter es $(6,0,0)$: al trasladar por $\sigma \neq e$ ningún elemento queda fijo. Proyectando, $\langle \chi_{\text{reg}}, \chi_i \rangle = \frac{6\,d_i}{6} = d_i$, es decir 1, 1 y 2. Luego $\mathbb{C}[S_3] \cong \text{triv} \oplus \text{sgn} \oplus \text{est}^{\oplus 2}$, con $1+1+2\cdot 2 = 6$.

(c) El cuadrado tensorial. $\chi^2_{\text{est}} = (4,0,1)$, y $\langle (4,0,1), \chi_i \rangle$ vale $\frac{4+0+2}{6} = 1$, $\frac{4-0+2}{6} = 1$ y $\frac{8+0-2}{6} = 1$: la descomposición es $\text{triv} \oplus \text{sgn} \oplus \text{est}$, con $1+1+2 = 4$. Coincide con lo que se obtiene en la sección sobre producto tensorial, ahora con la tabla justificada desde las matrices.

Regla: comprobar que una tabla de caracteres no tiene errores

Toda tabla admite seis verificaciones baratas, y conviene hacerlas siempre porque un error de signo se propaga a todas las descomposiciones posteriores.

1. **Filas y columnas coinciden en número** con el número de clases de conjugación.
2. **La primera columna son las dimensiones** y sus cuadrados suman $|G|$.
3. **La primera fila es todo unos**, la del carácter trivial.
4. **Cada fila tiene norma** 1 y dos filas distintas son ortogonales, con el peso $|C_j|$ de cada clase dentro de la suma.
5. **Cada columna al cuadrado da** $|C_G(g)| = |G|/|C|$, y dos columnas distintas son ortogonales. Esta comprobación es independiente de la anterior y suele delatar errores que las filas no ven.
6. **Los valores son enteros algebraicos** y, más concretamente, sumas de d_i raíces de la unidad: si aparece un $\frac{3}{2}$ o un valor de módulo mayor que d_i, hay un error seguro, porque $|\chi_i(g)| \leq d_i$ siempre.

Una última comprobación estructural: el número de caracteres de grado 1 ha de ser exactamente $|G/[G,G]|$. Para S_3, el conmutador es A_3 y el cociente tiene orden 2: hay dos, la trivial y el signo, y en efecto no hay más.

27.3. La representación regular y las aplicaciones

Teorema 27.12 (Descomposición de la regular)**.** *La* representación regular*, que hace actuar G sobre el espacio de funciones $G \to \mathbb{C}$ por traslación, tiene carácter $\chi_{reg}(e) = |G|$ y $\chi_{reg}(g) = 0$ para $g \neq e$. Descomponiéndola, cada irreducible aparece exactamente d_i veces, de donde $\sum_i d_i^2 = |G|$.*

Idea clave

Qué dice esto en cristiano. La representación regular es el grupo actuando sobre sí mismo, y su carácter es tan simple que se calcula de un vistazo: al trasladar por $g \neq e$ ningún elemento queda fijo, luego la traza es cero. Proyectar ese carácter sobre cada irreducible da directamente d_i, y sumar reproduce el orden del grupo. Es decir, la regular contiene a *todas* las irreducibles y en la proporción exacta de su dimensión: es un catálogo completo. De ahí sale la identidad $\sum d_i^2 = |G|$, que es la herramienta con la que se adivinan las tablas.

Ejemplo 27.13 (La regular y para qué sirven los caracteres)**. (a) La cuenta en S_3.** La regular tiene dimensión 6 y se descompone como $1 \cdot \chi_{\text{triv}} + 1 \cdot \chi_{\text{sgn}} + 2 \cdot \chi_{\text{std}}$, con dimensión total $1 + 1 + 2 \cdot 2 = 6$. Cada irreducible aparece tantas veces como su dimensión.

(b) Grupos abelianos. Si G es abeliano, cada elemento es su propia clase, luego hay $|G|$ irreducibles y $\sum d_i^2 = |G|$ obliga a que todas tengan dimensión 1. Para $\mathbb{Z}_n$, los caracteres son $\chi_k(j) = \mathrm{e}^{2\pi \mathrm{i} jk/n}$, y descomponer una función en esa base *es* la transformada de Fourier discreta. El análisis armónico clásico es la teoría de representación de los grupos abelianos.

(c) Ortogonalidad por columnas. Además de las filas, las columnas de la tabla también son ortogonales: $\sum_i \chi_i(g)\overline{\chi_i(h)}$ vale $\frac{|G|}{|\text{clase de } g|}$ si g y h son conjugados y cero si no. Esa segunda relación permite completar tablas cuando falta una fila y detectar errores de cálculo.

(d) Burnside, un teorema puramente algebraico. Todo grupo de orden $p^a q^b$ con p, q primos es soluble. La única demostración conocida durante setenta años usaba caracteres, y consiste en probar que un grupo simple no abeliano no puede tener una clase de conjugación de tamaño p^k. Es el ejemplo canónico de un enunciado sin análisis a la vista cuya prueba lo necesita.

Cuidado

[Sobre otros cuerpos casi nada de esto vale] Todo el capítulo supone $\mathbb{C}$ y G finito. Si el cuerpo tiene característica p que divide a $|G|$, el teorema de Maschke falla: en $\mathbb{F}_2$ la representación de $\mathbb{Z}_2$ sobre $\mathbb{F}_2^2$ dada por $\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$ tiene un subespacio invariante sin complemento invariante, y no se descompone. Esa es la teoría modular, mucho más difícil y todavía en desarrollo. Segundo aviso: para grupos infinitos hay que exigir continuidad y trabajar con grupos compactos, donde la suma sobre G se sustituye por la integral de Haar; el resultado, el teorema de Peter–Weyl, es el que conecta esta teoría con las series de Fourier.

27.4. Los grupos simétricos y los grupos continuos

La teoría general dice que un grupo finito tiene tantas representaciones irreducibles como clases de conjugación y que la suma de los cuadrados de sus dimensiones es el orden del grupo. Lo que no dice es cómo encontrarlas. Para dos familias concretas —los grupos simétricos y los grupos de Lie compactos— hay respuestas completas y sorprendentemente combinatorias.

Definición 27.14. Una *partición* de n es una sucesión decreciente $\lambda_1 \geq \lambda_2 \geq \cdots$ de enteros positivos con suma n; se dibuja como un *diagrama de Young*, una fila de λ_i casillas por cada índice. Las representaciones irreducibles de S_n están en biyección con las particiones de n, y la dimensión de la asociada a λ viene dada por la *fórmula del gancho*

$$\dim V_\lambda = \frac{n!}{\prod_c h(c)},$$

donde $h(c)$, el *gancho* de la casilla c, cuenta las casillas a su derecha, las de debajo y la propia.

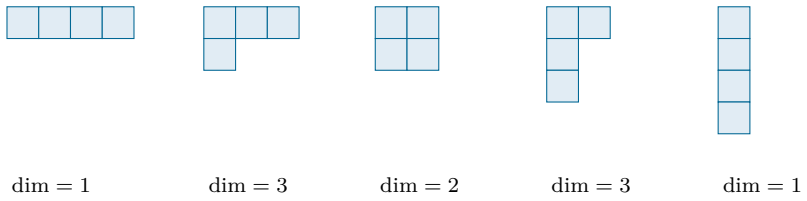


Figura 27.1: Las cinco particiones de 4 y, por tanto, las cinco representaciones irreducibles de S_4. Las dimensiones $1, 3, 2, 3, 1$ salen de contar ganchos, y sus cuadrados suman $1 + 9 + 4 + 9 + 1 = 24$, el orden del grupo.

Idea clave

Qué dice esto en cristiano. Las clases de conjugación de S_n son los tipos cíclicos, y un tipo cíclico es exactamente una partición de n: la permutación que es un 3-ciclo y un 2-ciclo en S_5 corresponde a la partición $3 + 2$. Así que el número de irreducibles ya está contado. Lo asombroso es que la correspondencia es explícita y que la dimensión se calcula dibujando un diagrama y contando casillas, sin construir ninguna matriz ni resolver ningún sistema. Un problema de álgebra lineal en dimensión desconocida se resuelve con una cuadrícula y aritmética elemental.

Ejemplo 27.15 (Las tablas de S_4 y S_5, calculadas)**. (a)** S_4**, de orden** 24**.** Las cinco particiones de 4 y sus dimensiones:

λ	4	3+1	2+2	2+1+1
dim	1	3	2	3

más 1+1+1+1 con dim = 1.

La comprobación: $1^2 + 3^2 + 2^2 + 3^2 + 1^2 = 1 + 9 + 4 + 9 + 1 = 24$, el orden del grupo. Los dos unos son la trivial y el signo; el 2 es la representación que factoriza por $S_4/V_4 \cong S_3$; los dos treses son la representación estándar y su producto por el signo.

(b) S_5, de orden 120. Siete particiones, con dimensiones 1, 4, 5, 6, 5, 4, 1, y en efecto $1+16+25+36+25+16+1=120$. Nótese la simetría de la lista: transponer el diagrama equivale a multiplicar por el signo, y por eso las dimensiones se leen igual en los dos sentidos.

(c) Un gancho calculado a mano. Para $\lambda=3+1$ en S_4, el diagrama tiene tres casillas arriba y una abajo a la izquierda. Los ganchos valen $4,2,1$ en la fila superior y 1 en la inferior, luego $\dim = 24/(4\cdot 2\cdot 1\cdot 1)=3$. Cuadra.

(d) Para qué sirve. La descomposición de $\mathbb{C}[S_n]$ en bloques por particiones es lo que hay detrás de la simetrización de tensores: los tensores completamente simétricos corresponden a $\lambda=n$, los antisimétricos a $\lambda=1+1+\cdots+1$, y todo lo demás son las simetrías mixtas que aparecen en la teoría de invariantes y en la clasificación de las partículas.

Definición 27.16. Para el grupo $SU(2)$, las representaciones irreducibles complejas están indexadas por un semientero $j\in\{0,\tfrac12,1,\tfrac32,\dots\}$, llamado *espín*, y la de índice j tiene dimensión $2j+1$. Se realizan sobre los polinomios homogéneos de grado $2j$ en dos variables. La descomposición del producto obedece la *regla de Clebsch–Gordan*

$$V_{j_1}\otimes V_{j_2}\cong V_{|j_1-j_2|}\oplus V_{|j_1-j_2|+1}\oplus\cdots\oplus V_{j_1+j_2}.$$

Ejemplo 27.17 (Grupos continuos y sus representaciones)**. (a) Las dimensiones.** Para $j=0,\tfrac12,1,\tfrac32,2$ salen $1,2,3,4,5$. La de dimensión 1 es la trivial; la de dimensión 2 es $SU(2)$ actuando sobre $\mathbb{C}^2$, es decir, sobre sí mismo; la de dimensión 3 es la que factoriza por $SO(3)$ y describe los vectores del espacio ordinario.

(b) Clebsch–Gordan comprobado. Con $j_1=j_2=\tfrac12$: el producto tiene dimensión $2\times 2=4$ y se descompone en $V_0\oplus V_1$, de dimensiones $1+3=4$. En física es el singlete y el triplete de dos electrones. Con $j_1=1$ y $j_2=\tfrac12$: dimensión $3\times 2=6$, que se parte en $V_{1/2}\oplus V_{3/2}$, es decir, $2+4=6$.

(c) Por qué semienteros. $SU(2)$ recubre a $SO(3)$ dos a uno, de modo que las representaciones de $SU(2)$ con j entero descienden a $SO(3)$ y las de j semientero no: al girar 360 grados cambian de signo y hace falta girar 720 para volver al estado inicial. Ése es el contenido matemático del espín del electrón, y no es una excentricidad cuántica sino un hecho topológico sobre $SO(3)$, cuyo grupo fundamental es $\mathbb{Z}/2\mathbb{Z}$.

(d) Ortogonalidad en el caso continuo. Todo el aparato de caracteres sobrevive sustituyendo la suma sobre el grupo por una integral respecto de la medida de Haar, que existe y es única salvo constante en todo grupo compacto. El carácter de V_j evaluado en una rotación de ángulo θ vale $\sin((2j+1)\theta/2)/\sin(\theta/2)$, y con esa fórmula las relaciones de ortogonalidad se comprueban con una integral trigonométrica elemental.

Cuidado

[Contraejemplo: sin compacidad, todo cambia] El teorema de Maschke y la ortogonalidad de caracteres descansan en poder promediar sobre el grupo, y eso exige una medida total finita. Para $SL(2,\mathbb{R})$, que no es compacto, no hay ninguna representación unitaria irreducible de dimensión finita salvo la trivial, y las interesantes son todas de dimensión infinita. La teoría existe —es la de Harish-Chandra— pero es incomparablemente más

difícil y no se parece a nada de este capítulo.

Pausa para pensar

La fórmula del gancho la publicaron Frame, Robinson y Thrall en 1954, y su demostración original era complicada. Hoy existe una prueba probabilística de Greene, Nijenhuis y Wilf que consiste en pasearse aleatoriamente por el diagrama saltando de gancho en gancho, y que cabe en una página. Es un ejemplo excelente de algo que ocurre a menudo: un resultado combinatorio limpio suele acabar teniendo una demostración limpia, aunque tarde décadas en aparecer.

27.5. Inducción, restricción y el teorema de Burnside

Las representaciones de un grupo y las de sus subgrupos están enlazadas por dos operaciones que van en direcciones opuestas y que resultan ser adjuntas la una de la otra. De ese enlace sale uno de los teoremas más impresionantes del álgebra finita: que ningún grupo de orden $p^a q^b$ puede ser simple.

Definición 27.18. Sea $H \leq G$ de índice n. La *restricción* $\operatorname{Res}_H^G V$ es la representación de H que se obtiene mirando sólo los elementos de H. La *inducción* $\operatorname{Ind}_H^G W$ es la representación de G de dimensión $n \dim W$ construida como $\mathbb{C}[G] \otimes_{\mathbb{C}[H]} W$. Se cumple la *reciprocidad de Frobenius*

$$\langle \operatorname{Ind}_H^G W,\ V \rangle_G = \langle W,\ \operatorname{Res}_H^G V \rangle_H.$$

Idea clave

Qué dice esto en cristiano. Restringir es fácil y pierde información: se olvida cómo actúan los elementos que no están en H. Inducir es lo contrario: se parte de una acción pequeña y se fabrica la mayor acción de G compatible con ella, colocando una copia de W en cada clase lateral. La reciprocidad dice que las dos operaciones son adjuntas, y su utilidad práctica es enorme: preguntar cuántas veces aparece una irreducible de G dentro de algo inducido es lo mismo que preguntar cuántas veces aparece una irreducible de H dentro de algo restringido, y una de las dos preguntas suele ser mucho más fácil.

Ejemplo 27.19 (Inducir y restringir en S_3 y S_4)**. (a) Inducir la trivial desde A_3.** El índice es 2, luego la inducida tiene dimensión 2. Por Frobenius se comprueba que $\operatorname{Ind}_{A_3}^{S_3} \mathbf{1} = \mathbf{1} \oplus \operatorname{sgn}$.

(b) Inducir desde una transposición. Con $H = \langle (12) \rangle$, de índice 3, la inducida de la trivial tiene dimensión 3 y es la representación por permutación de S_3 sobre las tres clases laterales, que vale $\mathbf{1} \oplus$ estándar.

(c) La regla general. Inducir la representación trivial de H da siempre la representación por permutación de G sobre G/H. Toda la teoría de acciones de grupos sobre conjuntos es un caso particular de inducción.

(d) Restringir de S_4 a S_3. La representación estándar de S_4, de dimensión 3, restringe a $\mathbf{1} \oplus$ estándar de S_3. La regla de ramificación para grupos simétricos dice que $\mathrm{Res}^{S_n}_{S_{n-1}} V_\lambda$ es la suma de los V_μ con μ obtenido quitando una casilla al diagrama de λ: otra vez, combinatoria pura donde uno esperaba álgebra lineal.

Teorema 27.20 (Burnside)**.** *Todo grupo finito de orden $p^a q^b$, con p y q primos, es resoluble. En particular, ningún grupo simple no abeliano tiene orden divisible por sólo dos primos.*

Idea clave

Qué dice esto en cristiano. El enunciado no menciona representaciones y la demostración las necesita: durante setenta años no se conoció ninguna prueba que no pasara por la teoría de caracteres, hasta que Bender y Goldschmidt encontraron una en los años setenta, mucho más larga. El argumento clave es que si χ es un carácter irreducible de grado d y una clase de conjugación tiene tamaño coprimo con d, entonces $\chi(g) = 0$ o bien g actúa como un escalar. De ahí se extrae un subgrupo normal propio, y la inducción sobre el orden hace el resto. La moraleja metodológica es la que interesa: para demostrar algo sobre la *estructura* de un grupo se pasa por sus *acciones lineales*, se hace álgebra lineal allí y se vuelve.

Cuidado

[Contraejemplo: tres primos ya bastan para romperlo] A_5 tiene orden $60 = 2^2 \cdot 3 \cdot 5$ y es simple no abeliano, luego no resoluble. Añadir un tercer primo destruye la conclusión de Burnside por completo, y no de forma marginal: A_5 es la razón de que la quíntica no se resuelva por radicales. El teorema es agudo en el sentido más fuerte, y el salto de dos a tres primos es uno de esos escalones donde las matemáticas cambian de régimen sin previo aviso.

Pausa para pensar

William Burnside publicó el teorema en 1904 y también formuló el problema que lleva su nombre: si un grupo finitamente generado en el que todo elemento tiene orden acotado ha de ser finito. La respuesta, hallada por Golod y Shafarevich en 1964, es que no. Burnside escribió además el primer tratado en inglés sobre grupos finitos, y en la primera edición, de 1897, omitió deliberadamente la teoría de caracteres por considerarla poco útil. En la segunda, de 1911, la incluyó y reconoció el error en el prólogo.

27.6. Los dos teoremas que sostienen todo: Maschke y el lema de Schur

Hemos usado las relaciones de ortogonalidad, el recuento de irreducibles y la descomposición en irreducibles como si fueran datos. Los tres salen de dos teoremas, y ambos se demuestran con el mismo truco: promediar sobre el grupo.

Teorema 27.21 (Maschke)**.** *Sea G finito y K un cuerpo cuya característica no divide a $|G|$. Entonces toda representación de G sobre K es suma directa de irreducibles.*

Demostración. Basta ver que todo subespacio invariante $W \subset V$ tiene un complemento invariante. Tómese cualquier proyección lineal $\pi : V \to W$ y promédiese:

$$\tilde{\pi} = \frac{1}{|G|} \sum_{g \in G} \rho(g)\, \pi\, \rho(g)^{-1}.$$

Sigue siendo proyección sobre W (si $w \in W$, cada sumando devuelve w), y ahora conmuta con la acción, luego $\ker \tilde{\pi}$ es invariante y $V = W \oplus \ker \tilde{\pi}$. Inducción sobre la dimensión. □

Cuidado

El $|G|$ del denominador es todo el teorema, y por eso la hipótesis sobre la característica no es técnica. Tómese $G = \mathbb{Z}/2$ actuando sobre $\mathbb{F}_2^2$ por

$$\rho(1) = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}.$$

El subespacio generado por e_1 es invariante, y no tiene ningún complemento invariante: el único otro vector fijo es él mismo. La demostración se rompe exactamente donde tiene que romperse: $|G| = 2 = 0$ en $\mathbb{F}_2$ y no se puede dividir. Toda la teoría de representaciones modulares existe para lidiar con este caso, y es incomparablemente más difícil.

Teorema 27.22 (Lema de Schur)**.** *Sean V, W irreducibles y $f : V \to W$ un morfismo de representaciones.*

1. *Si $V \not\cong W$, entonces $f = 0$.*
2. *Si $V = W$ y el cuerpo es algebraicamente cerrado, entonces $f = \lambda\,\mathrm{id}$ para algún escalar λ.*

Demostración. $\ker f$ e $\operatorname{im} f$ son invariantes; por irreducibilidad, o $f = 0$ o f es isomorfismo, lo que da (1). Para (2), sobre $\mathbb{C}$ el endomorfismo f tiene algún autovalor λ, y entonces $f - \lambda\,\mathrm{id}$ es un morfismo con núcleo no trivial: por (1) es nulo. □

Ejemplo 27.23 (El operador promediado, calculado)**.** Tómese la irreducible de dimensión 2 de S_3 y la matriz $A = E_{11}$, que no es morfismo de nada. Promediando,

$$\tilde{A} = \frac{1}{6} \sum_{g \in S_3} \rho(g)\, E_{11}\, \rho(g)^{-1},$$

el resultado sí conmuta con toda la acción, luego por Schur es un escalar. ¿Cuál? La traza se conserva al conjugar, así que $\operatorname{tr} \tilde{A} = \operatorname{tr} E_{11} = 1$, y un escalar λI en dimensión 2 tiene traza 2λ. Luego

$$\tilde{A} = \tfrac{1}{2} I.$$

En general, el promedio de A sobre una irreducible de dimensión d es $\frac{\operatorname{tr} A}{d} I$: el promedio olvida absolutamente todo de A salvo su traza. Esa es la maquinaria que produce las relaciones de ortogonalidad, tomando A igual a cada matriz elemental.

Teorema 27.24 (Relaciones de ortogonalidad y recuento)**.** *Los caracteres irreducibles forman una base ortonormal del espacio de* ***funciones de clase*** *(las constantes en cada clase de conjugación), con el producto*

$$\langle \chi, \psi \rangle = \frac{1}{|G|} \sum_{g \in G} \chi(g) \overline{\psi(g)}.$$

En particular, el número de irreducibles es exactamente el número de clases de conjugación.

Ahora sí está justificado el paso 1 de la receta. Y el paso 2 —que cada d_i divide a $|G|$— necesita un ingrediente distinto: los enteros algebraicos. La clave es que $\frac{|C|\chi(g)}{\chi(1)}$ es entero algebraico para toda clase C, cosa que se prueba viendo que los elementos de clase generan un subanillo finitamente generado del centro del álgebra de grupo. De ahí sale $d_i \mid |G|$, y de ahí sale el teorema $p^a q^b$ de Burnside.

Ejemplo 27.25 (La divisibilidad como herramienta de descarte)**.** Para $|G| = 8$ con cinco clases, $\sum d_i^2 = 8$ con cinco sumandos deja solo $1+1+1+1+4$. Y para A_5, con $|G| = 60$ y cinco clases, la ecuación $\sum d_i^2 = 60$ con $d_1 = 1$ tiene varias soluciones enteras, pero la divisibilidad $d_i \mid 60$ y el hecho de que A_5 es simple (luego no tiene más caracteres de grado 1) dejan únicamente

$$1 + 3^2 + 3^2 + 4^2 + 5^2 = 1 + 9 + 9 + 16 + 25 = 60.$$

Las dimensiones de las irreducibles de A_5 son $1, 3, 3, 4, 5$, y se han obtenido sin construir ni una sola matriz.

27.7. Producto tensorial, dual, simétrica y exterior

Las representaciones no solo se suman: se multiplican. Y la aritmética de caracteres correspondiente es de una sencillez casi injusta.

Definición 27.26. Dadas V y W con caracteres χ_V, χ_W:

$$\chi_{V \oplus W} = \chi_V + \chi_W, \qquad \chi_{V \otimes W} = \chi_V \cdot \chi_W, \qquad \chi_{V^*} = \overline{\chi_V},$$

y el espacio $V \otimes V$ se parte en la parte simétrica y la antisimétrica, con

$$\chi_{\mathrm{Sym}^2 V}(g) = \tfrac{1}{2}(\chi(g)^2 + \chi(g^2)), \qquad \chi_{\Lambda^2 V}(g) = \tfrac{1}{2}(\chi(g)^2 - \chi(g^2)).$$

La razón de la primera fórmula es que los autovalores de g en $V \otimes W$ son los productos $\lambda_i \mu_j$, y $\sum_{i,j} \lambda_i \mu_j = (\sum \lambda_i)(\sum \mu_j)$. La de Sym^2 sale de que sus autovalores son los $\lambda_i \lambda_j$ con $i \le j$, y $\sum_{i \le j} \lambda_i \lambda_j = \frac{1}{2}[(\sum \lambda_i)^2 + \sum \lambda_i^2]$.

Ejemplo 27.27 (S_3, con toda la aritmética explícita)**.** Clases: e (tamaño 1), transposiciones (3), 3-ciclos (2). Caracteres irreducibles:

$$\chi_{\text{triv}} = (1,1,1), \qquad \chi_{\text{sgn}} = (1,-1,1), \qquad \chi_{\text{std}} = (2,0,-1).$$

(a) El cuadrado. $\chi^2_{\text{std}} = (4,0,1)$, de dimensión 4. Sus productos con las tres irreducibles:

$$\langle \chi^2, \chi_{\text{triv}} \rangle = \frac{1\cdot 4 + 3\cdot 0 + 2\cdot 1}{6} = 1, \quad \langle \chi^2, \chi_{\text{sgn}} \rangle = \frac{4+0+2}{6} = 1, \quad \langle \chi^2, \chi_{\text{std}} \rangle = \frac{8+0-2}{6} = 1.$$

Luego

$$\text{std} \otimes \text{std} = \mathbf{1} \oplus \text{sgn} \oplus \text{std},$$

y las dimensiones cuadran: $4 = 1+1+2$.

(b) Separando simétrica y exterior. Hace falta $\chi(g^2)$: el cuadrado de la identidad es la identidad ($\chi = 2$), el de una transposición es la identidad ($\chi = 2$) y el de un 3-ciclo es otro 3-ciclo ($\chi = -1$). Luego $\chi(g^2) = (2,2,-1)$ y

$$\chi_{\text{Sym}^2} = \tfrac{1}{2}\big[(4,0,1) + (2,2,-1)\big] = (3,1,0), \qquad \chi_{\Lambda^2} = \tfrac{1}{2}\big[(4,0,1) - (2,2,-1)\big] = (1,-1,1).$$

(c) Identificación. Λ^2 es exactamente χ_{sgn}: la potencia exterior máxima de la representación estándar es el signo, que es el hecho que da nombre a la signatura de una permutación. Y $\text{Sym}^2 = (3,1,0)$ se descompone como $\mathbf{1} \oplus \text{std}$, porque $\langle \text{Sym}^2, \chi_{\text{triv}} \rangle = 1$ y $\langle \text{Sym}^2, \chi_{\text{std}} \rangle = 1$. Todo cuadra: $4 = 3+1$ y $\mathbf{1} \oplus \text{std} \oplus \text{sgn}$.

Idea clave

Estas fórmulas convierten la teoría de representaciones en *aritmética de vectores cortos*. Para descomponer un producto tensorial no hace falta ver ni una matriz: se multiplican dos listas de números y se proyecta sobre la tabla de caracteres. Ese es el cálculo que en física se llama Clebsch–Gordan y que dice cómo se combinan dos momentos angulares; para $SU(2)$, $V_a \otimes V_b = V_{a+b} \oplus V_{a+b-2} \oplus \cdots \oplus V_{|a-b|}$, con la misma lógica pero con caracteres que son funciones y no listas finitas.

Observación 27.28 (El álgebra de grupo). Todo lo anterior tiene una formulación estructural en una sola línea. El **álgebra de grupo** $\mathbb{C}[G]$ —combinaciones lineales formales de elementos de G, con el producto extendido por linealidad— se descompone como

$$\mathbb{C}[G] \;\cong\; \bigoplus_i M_{d_i}(\mathbb{C}),$$

que es el teorema de Artin–Wedderburn aplicado a un álgebra semisimple (semisimple por Maschke). Contando dimensiones se obtiene $\sum d_i^2 = |G|$, que deja de ser una casualidad numérica. Para S_3:

$$\mathbb{C}[S_3] \cong \mathbb{C} \oplus \mathbb{C} \oplus M_2(\mathbb{C}), \qquad 1+1+4 = 6.$$

Los proyectores sobre cada bloque son los idempotentes centrales

$$e_i = \frac{d_i}{|G|} \sum_g \overline{\chi_i(g)}\, g,$$

y por ejemplo $e_{\text{sgn}} = \frac{1}{6}\sum_g \text{sgn}(g)\, g$ es el antisimetrizador de toda la vida, el que aparece en el determinante y en los estados fermiónicos. El centro de $\mathbb{C}[G]$ es, dimensión a dimensión, el espacio de funciones de clase.

27.8. Dimensión infinita: grupos compactos, medida de Haar y el teorema de Peter–Weyl

Todo lo anterior supone G finito, y esa hipótesis se usó exactamente dos veces: para poder *promediar* sobre el grupo (Maschke) y para que la suma $\sum_g$ tuviera sentido (ortogonalidad). Si G es infinito pero **compacto**, ambas cosas se recuperan sustituyendo la suma por una integral.

Teorema 27.29 (Medida de Haar)**.** *Todo grupo topológico compacto admite una única medida de probabilidad μ invariante por traslaciones:*

$$\int_G f(hg)\,\mathrm{d}\mu(g) = \int_G f(g)\,\mathrm{d}\mu(g) = \int_G f(gh)\,\mathrm{d}\mu(g) \qquad \text{para todo } h \in G.$$

Con ella, la receta $\frac{1}{|G|}\sum_g$ se convierte en $\int_G \cdots \mathrm{d}\mu$ y *todo* el capítulo se traslada palabra por palabra: Maschke sigue valiendo (toda representación unitaria de dimensión finita es suma de irreducibles), Schur sigue valiendo, y los caracteres siguen siendo ortonormales:

$$\langle \chi, \psi \rangle = \int_G \chi(g)\overline{\psi(g)}\,\mathrm{d}\mu(g).$$

Teorema 27.30 (Peter–Weyl, 1927)**.** *Sea G compacto. Entonces:*

1. *Toda representación unitaria irreducible de G tiene dimensión* finita*.*
2. *Los coeficientes matriciales de las irreducibles forman una base ortogonal completa de $L^2(G,\mu)$:*
$$L^2(G) \;\cong\; \widehat{\bigoplus_{\rho \in \hat{G}}} V_\rho^* \otimes V_\rho,$$
donde cada irreducible ρ aparece $\dim V_\rho$ *veces.*
3. *Las funciones de clase de L^2 tienen a los caracteres irreducibles por base ortonormal.*

Idea clave

La primera afirmación es la sorpresa: aunque G sea infinito, sus piezas irreducibles siguen siendo matrices finitas. La segunda es la que da el nombre a todo el tema, porque *contiene* el análisis de Fourier. Tómese $G = \mathbb{T}$, el círculo, que es compacto y abeliano. Sus irreducibles son unidimensionales —los caracteres $e^{2\pi i n x}$, uno por cada $n \in \mathbb{Z}$— y Peter–Weyl dice que

$$L^2(\mathbb{T}) = \widehat{\bigoplus_{n \in \mathbb{Z}}} \mathbb{C}\, e^{2\pi i n x},$$

que es exactamente la completitud del sistema trigonométrico. La serie de Fourier *es* la descomposición de la representación regular del círculo.

Y ahí está la respuesta a la pregunta que quedaba pendiente al final del capítulo de análisis armónico: el análisis de Fourier es teoría de representaciones del círculo, y esta es la frase que lo demuestra.

Ejemplo 27.31 ($SU(2)$, entero)**.** $SU(2)$ es compacto y no abeliano: el ejemplo mínimo donde la teoría dice algo nuevo.

(a) Las irreducibles. Una por cada dimensión: V_n = polinomios homogéneos de grado n en dos variables, con $\dim V_n = n+1$. No hay más.

(b) Los caracteres. Toda matriz de $SU(2)$ es conjugada a $\operatorname{diag}(e^{i\theta}, e^{-i\theta})$, luego un carácter es una función de θ:

$$\chi_n(\theta) = \frac{\sin\big((n+1)\theta\big)}{\sin\theta} = e^{in\theta} + e^{i(n-2)\theta} + \cdots + e^{-in\theta},$$

y en efecto $\chi_n(0) = n+1$, la dimensión. La medida de Haar sobre las clases es $\frac{2}{\pi}\sin^2\theta\,\mathrm{d}\theta$ en $[0,\pi]$, y respecto de ella los χ_n son ortonormales.

(c) Clebsch–Gordan. Multiplicando dos caracteres y reagrupando exponenciales sale

$$V_a \otimes V_b \;\cong\; V_{a+b} \oplus V_{a+b-2} \oplus \cdots \oplus V_{|a-b|}.$$

Compruébense las dimensiones: $V_1 \otimes V_1 = V_2 \oplus V_0$ da $2\cdot 2 = 3+1$; y $V_2 \otimes V_1 = V_3 \oplus V_1$ da $3\cdot 2 = 4+2$. ✓

(d) Qué significa en física. V_n es el estado de espín $j = n/2$, y la regla de Clebsch–Gordan es la de composición de momentos angulares: dos partículas de espín $\frac{1}{2}$ (V_1 cada una) dan un triplete (V_2, espín 1) y un singlete (V_0, espín 0). La distinción entre representaciones de $SU(2)$ y de $SO(3) = SU(2)/\{\pm I\}$ —solo las V_n con n par bajan al cociente— es exactamente la distinción entre bosones y fermiones: el espín semientero solo existe porque $SU(2)$ es el recubrimiento doble de $SO(3)$.

Observación 27.32 (Y cuando el grupo no es compacto)*.* Todo se rompe a la vez. Para $G = \mathbb{R}$ los caracteres son $e^{i\xi x}$ con ξ *continuo*: no hay suma directa sino integral, y aparece la transformada de Fourier en lugar de la serie. Para $SL_2(\mathbb{R})$ las irreducibles unitarias son de dimensión infinita y se organizan en familias (serie principal, serie discreta, complementaria) cuya clasificación ocupó a Bargmann, Harish-Chandra y varias décadas. Y el programa de Langlands es, en una frase demasiado corta, la conjetura de que las representaciones automorfas de esos grupos se corresponden con las representaciones de Galois del capítulo correspondiente: los dos hilos del libro, el analítico y el aritmético, atados por el mismo nudo.

27.9. Restricción, inducción y la reciprocidad de Frobenius

Dado un subgrupo $H \le G$, hay dos maneras obvias de mover representaciones entre los dos, y resultan ser adjuntas.

Definición 27.33. Restricción: si V es representación de G, $\operatorname{Res}^G_H V$ es la misma V mirando solo la acción de H; su carácter es la restricción del carácter.

Inducción: si W es representación de H, se define

$$\operatorname{Ind}^G_H W = \bigoplus_{gH \in G/H} g \otimes W,$$

de dimensión $[G:H]\dim W$, con G permutando los sumandos. Su carácter es

$$\chi_{\operatorname{Ind} W}(g) = \frac{1}{|H|} \sum_{\substack{x\in G \\ x^{-1}gx\in H}} \chi_W(x^{-1}gx).$$

Teorema 27.34 (Reciprocidad de Frobenius)**.**

$$\langle \operatorname{Ind}_H^G W,\ V\rangle_G \ =\ \langle W,\ \operatorname{Res}_H^G V\rangle_H.$$

En lenguaje categórico, inducción y restricción son funtores adjuntos, y esta igualdad es la biyección de sus conjuntos de morfismos. En lenguaje práctico: para saber cuántas veces aparece V en algo inducido, basta restringir V y contar allí, que es un cálculo en un grupo más pequeño.

Ejemplo 27.35 (S_3 desde su subgrupo cíclico)**.** Sea $H=\langle(1\,2\,3)\rangle\cong\mathbb{Z}/3$, de índice 2 en S_3. Sus tres irreducibles son unidimensionales: $\psi_k(c)=\omega^k$ con $\omega=e^{2\pi i/3}$.

(a) Inducir la trivial. $\operatorname{Ind}\psi_0$ tiene dimensión $2\cdot 1=2$, y su carácter en las tres clases de S_3 —identidad, transposiciones, 3-ciclos— vale $(2,0,2)$. Por Frobenius,

$$\langle \operatorname{Ind}\psi_0, \mathbf{1}\rangle = \langle \psi_0, \operatorname{Res}\mathbf{1}\rangle = 1, \qquad \langle \operatorname{Ind}\psi_0, \operatorname{sgn}\rangle = \langle \psi_0, \operatorname{Res}\operatorname{sgn}\rangle = 1,$$

porque el signo restringido a los 3-ciclos es trivial. Luego $\operatorname{Ind}\psi_0 = \mathbf{1}\oplus\operatorname{sgn}$, y en efecto $(2,0,2)=(1,1,1)+(1,-1,1)$. ✓

(b) Inducir una no trivial. $\operatorname{Ind}\psi_1$ tiene también dimensión 2 y carácter $(2,0,-1)$: es *irreducible*, y es la representación estándar. Ahí está construida sin escribir ni una matriz.

(c) La comprobación cruzada. Res de la estándar a $\mathbb{Z}/3$ da el carácter $(2,-1,-1)$ sobre $\{e,c,c^2\}$, que es $\psi_1+\psi_2$. Y Frobenius predice $\langle\operatorname{Ind}\psi_1,\operatorname{std}\rangle=\langle\psi_1,\psi_1+\psi_2\rangle=1$: coherente. ✓

Observación 27.36. La inducción es la herramienta con la que se construyen de verdad las tablas de caracteres de los grupos grandes: se inducen caracteres de subgrupos manejables y se descomponen con Frobenius. El teorema de Brauer garantiza que *todo* carácter de un grupo finito es combinación entera de caracteres inducidos desde subgrupos elementales, lo que convierte el método en universal. Y en el caso de $\mathbb{Z}/n \le$ nada, la inducción desde el trivial da la representación regular, cuya descomposición $\mathbb{C}[G]=\bigoplus V_i^{\oplus d_i}$ es el teorema de Artin–Wedderburn de la sección anterior.

27.10. Las relaciones de ortogonalidad, demostradas

Todo el capítulo —la construcción de tablas de caracteres, el recuento de irreducibles, la descomposición de representaciones— descansa en un teorema que se enunció sin prueba. Se demuestra aquí, y con él la divisibilidad de las dimensiones, que se usó de inmediato para A_5.

En todo lo que sigue G es finito, el cuerpo es $\mathbb{C}$ y las representaciones son de dimensión finita. Recuérdese el producto escalar de funciones de clase:

$$\langle f_1, f_2\rangle = \frac{1}{|G|}\sum_{g\in G} f_1(g)\,\overline{f_2(g)}.$$

Lema 27.37 (Promediar produce entrelazadores)**.** *Sean (ρ_1, V_1) y (ρ_2, V_2) representaciones y $\varphi : V_1 \to V_2$ lineal cualquiera. Entonces*

$$\widetilde{\varphi} = \frac{1}{|G|} \sum_{g \in G} \rho_2(g)\, \varphi\, \rho_1(g)^{-1}$$

es G-equivariante, y $\operatorname{tr} \widetilde{\varphi} = \operatorname{tr} \varphi$.

Demostración. Para $h \in G$, $\rho_2(h)\widetilde{\varphi}\rho_1(h)^{-1} = \frac{1}{|G|}\sum_g \rho_2(hg)\varphi\rho_1(hg)^{-1} = \widetilde{\varphi}$ (reindexando $g \mapsto h^{-1}g$). La traza se conserva porque $\operatorname{tr}(\rho_2(g)\varphi\rho_1(g)^{-1}) = \operatorname{tr}\varphi$ cuando $V_1 = V_2$ y $\rho_1 = \rho_2$, y se promedian $|G|$ copias. □

Teorema 27.38 (Primera relación de ortogonalidad)**.** *Si χ_1, χ_2 son los caracteres de dos representaciones irreducibles, entonces*

$$\langle \chi_1, \chi_2 \rangle = \begin{cases} 1, & \text{si son isomorfas,} \\ 0, & \text{si no lo son.} \end{cases}$$

Demostración. Paso 1: coeficientes matriciales. Fíjense bases y escríbanse $\rho_1(g) = (a_{ij}(g))$, $\rho_2(g) = (b_{kl}(g))$ en forma unitaria (siempre se puede, promediando un producto escalar). Sea E_{ji} la matriz elemental y aplíquese el lema 27.37 a $\varphi = E_{ji}$.

Si las representaciones *no* son isomorfas, el lema de Schur dice que todo entrelazador es nulo, luego $\widetilde{\varphi} = 0$ para toda φ; leyendo la entrada (k,l) de esa igualdad,

$$\frac{1}{|G|}\sum_g b_{kj}(g)\, a_{il}(g^{-1}) = 0 \qquad \text{para todos } i,j,k,l.$$

Si son isomorfas —tómense iguales, $\rho_1 = \rho_2 = \rho$ de dimensión d—, Schur dice que $\widetilde{\varphi}$ es un múltiplo de la identidad, y por el lema su traza es $\operatorname{tr} E_{ji} = \delta_{ij}$; por tanto $\widetilde{\varphi} = \frac{\delta_{ij}}{d} I$, es decir

$$\frac{1}{|G|}\sum_g a_{kj}(g)\, a_{il}(g^{-1}) = \frac{\delta_{ij}\delta_{kl}}{d}.$$

Paso 2: pasar a caracteres. Como $\chi(g) = \sum_i a_{ii}(g)$ y, por unitariedad, $a_{il}(g^{-1}) = \overline{a_{li}(g)}$, se obtiene en el caso no isomorfo

$$\langle \chi_1, \chi_2 \rangle = \frac{1}{|G|}\sum_g \Big(\sum_k b_{kk}(g)\Big)\Big(\sum_i \overline{a_{ii}(g)}\Big) = \sum_{k,i} \frac{1}{|G|}\sum_g b_{kk}(g) a_{ii}(g^{-1}) = 0,$$

y en el caso isomorfo

$$\langle \chi, \chi \rangle = \sum_{k,i} \frac{\delta_{ik}\delta_{ki}}{d} = \sum_k \frac{1}{d} = \frac{d}{d} = 1.$$

□

Corolario 27.39 (Criterio de irreducibilidad y descomposición)**.** *Una representación con carácter χ es irreducible si y sólo si $\langle \chi, \chi \rangle = 1$. Y si $\chi = \sum_i m_i \chi_i$ con los χ_i irreducibles distintos, entonces $m_i = \langle \chi, \chi_i \rangle$ y $\langle \chi, \chi \rangle = \sum_i m_i^2$.*

Teorema 27.40 (El número de irreducibles es el número de clases)**.** *Los caracteres irreducibles forman una* base ortonormal *del espacio de funciones de clase de G. En particular, el número de representaciones irreducibles no isomorfas es igual al número de clases de conjugación.*

Demostración. Son ortonormales por el teorema 27.38, luego linealmente independientes; falta que generen. Sea f una función de clase ortogonal a todos los χ_i, y para cada representación irreducible (ρ, V) defínase

$$T_f = \sum_{g\in G} \overline{f(g)}\,\rho(g).$$

Como f es de clase, T_f conmuta con todos los $\rho(h)$ —reindexando la suma por conjugación—, luego por el lema de Schur $T_f = \lambda I$ con

$$\lambda = \frac{\operatorname{tr} T_f}{d} = \frac{1}{d}\sum_g \overline{f(g)}\chi(g) = \frac{|G|}{d}\,\overline{\langle f, \chi\rangle} = 0.$$

Así $T_f = 0$ en toda irreducible, y por tanto también en la representación regular, que es suma de irreducibles. Pero en la regular, aplicando T_f al vector básico e_e se obtiene $\sum_g \overline{f(g)}\,e_g$, combinación de vectores de una base: que sea nula obliga a $f(g) = 0$ para todo g. Luego el complemento ortogonal de los caracteres es trivial, y como la dimensión del espacio de funciones de clase es el número de clases de conjugación, se concluye. □

Teorema 27.41 (Las dimensiones dividen al orden)**.** *Si $d_1, \dots, d_r$ son las dimensiones de las representaciones irreducibles de G, entonces $d_i \mid |G|$ para todo i. Además $\sum_i d_i^2 = |G|$.*

Demostración. La segunda afirmación sale de descomponer la representación regular, cuyo carácter vale $|G|$ en la identidad y 0 en lo demás: $\langle \chi_{\text{reg}}, \chi_i\rangle = d_i$, luego $\chi_{\text{reg}} = \sum d_i\chi_i$ y evaluando en la identidad, $|G| = \sum d_i^2$.

Para la divisibilidad hace falta una noción del capítulo 26: un número complejo es un *entero algebraico* si es raíz de un polinomio mónico con coeficientes enteros. Los enteros algebraicos forman un anillo, y los *racionales* que además son enteros algebraicos son exactamente los enteros (ese es el lema de la raíz racional).

Sean $C_1, \dots, C_r$ las clases de conjugación y $z_j = \sum_{g\in C_j} g \in \mathbb{C}[G]$ las *sumas de clase*. Cada z_j es central en $\mathbb{C}[G]$, luego actúa en una irreducible V_i como un escalar ω_{ij}, y tomando trazas,

$$\omega_{ij} = \frac{|C_j|\,\chi_i(g_j)}{d_i}, \qquad g_j \in C_j.$$

El punto clave es que los ω_{ij} son enteros algebraicos: los productos $z_j z_k = \sum_l c_{jkl} z_l$ tienen coeficientes c_{jkl} *enteros no negativos* (cuentan factorizaciones), de modo que el $\mathbb{Z}$-módulo generado por las sumas de clase es finitamente generado y estable por multiplicación; el criterio del módulo fiel da que cada ω_{ij} es entero algebraico. También lo es cada $\chi_i(g)$, por ser suma de raíces de la unidad (los autovalores de $\rho_i(g)$ satisfacen $\lambda^{|G|} = 1$).

Ahora, por la relación de ortogonalidad agrupada por clases,

$$|G| = |G|\,\langle \chi_i, \chi_i\rangle = \sum_{j=1}^{r} |C_j|\,\chi_i(g_j)\overline{\chi_i(g_j)},$$

y dividiendo entre d_i:

$$\frac{|G|}{d_i} = \sum_{j=1}^{r} \underbrace{\frac{|C_j|\chi_i(g_j)}{d_i}}_{\omega_{ij}} \cdot \underbrace{\overline{\chi_i(g_j)}}_{\text{entero alg.}},$$

que es una suma de productos de enteros algebraicos, luego entero algebraico. Y es racional. Por tanto es un entero: $d_i \mid |G|$. □

Ejemplo 27.42 (Las dimensiones de A_5, ahora justificadas)**.** $|A_5| = 60$ y hay 5 clases de conjugación, luego 5 irreducibles con $\sum_{i=1}^{5} d_i^2 = 60$ y cada $d_i \mid 60$. Una es la trivial, $d_1 = 1$; y como A_5 es simple, no hay más de dimensión 1 (un carácter lineal tendría núcleo normal propio salvo que sea trivial). Quedan cuatro dimensiones ≥ 2 con $\sum_{i\geq 2} d_i^2 = 59$ y cada una dividiendo a 60. Las opciones son escasas: $59 = 9 + 9 + 16 + 25$, es decir $\{3, 3, 4, 5\}$, y es la única descomposición de 59 en cuatro cuadrados de divisores de 60 mayores que 1. Sin el teorema 27.41 el argumento no cerraría, porque $59 = 4 + 9 + 21, \ldots$ admitiría otras combinaciones de cuadrados.

27.11. El teorema de Burnside, demostrado

El teorema que da título a la sección se enunció como cita. Con las relaciones de ortogonalidad y la divisibilidad de las dimensiones ya demostradas, su prueba está al alcance: es el ejemplo canónico de un enunciado puramente sobre grupos que sólo se sabe demostrar con caracteres.

Lema 27.43 (Lema clave)**.** *Sea χ un carácter irreducible de grado d, C una clase de conjugación con $\gcd(|C|, d) = 1$, y $g \in C$ con $\chi(g) \neq 0$. Entonces $\rho(g)$ es una matriz* escalar*.*

Demostración. En la demostración del teorema 27.41 se vio que $\omega = \dfrac{|C|\,\chi(g)}{d}$ es un entero algebraico, y que $\chi(g)$ también lo es. Como $\gcd(|C|, d) = 1$, la identidad de Bézout da enteros u, v con $u|C| + vd = 1$, luego

$$\frac{\chi(g)}{d} = \frac{(u|C| + vd)\chi(g)}{d} = u\,\frac{|C|\chi(g)}{d} + v\,\chi(g) = u\,\omega + v\,\chi(g),$$

que es un entero algebraico por ser combinación entera de enteros algebraicos.

Ahora, $\chi(g)$ es la suma de los d autovalores de $\rho(g)$, todos raíces de la unidad (pues $\rho(g)^{|G|} = I$). Por la desigualdad triangular,

$$\left|\frac{\chi(g)}{d}\right| \leq \frac{1}{d}\sum_{i=1}^{d} |\lambda_i| = 1,$$

con igualdad si y sólo si todos los λ_i son iguales. Supóngase que no lo son, de modo que $|\chi(g)/d| < 1$ estrictamente. Todo conjugado algebraico de $\chi(g)/d$ es de la misma forma —una media de d raíces de la unidad, porque los conjugados de Galois permutan las raíces de la unidad—, luego también tiene módulo ≤ 1. El producto de todos los conjugados es la norma,

un número *racional* que además es entero algebraico, o sea un entero de $\mathbb{Z}$; y su módulo es < 1. El único entero de módulo menor que 1 es 0, luego $\chi(g)/d = 0$, contra la hipótesis $\chi(g) \neq 0$.

Por tanto todos los autovalores coinciden, y siendo $\rho(g)$ diagonalizable (tiene orden finito), es la matriz escalar λI. □

Lema 27.44 (Ninguna clase de un grupo simple tiene tamaño potencia de primo)**.** *Sea G un grupo simple no abeliano. Entonces ninguna clase de conjugación de G tiene cardinal p^k con p primo y $k \geq 1$.*

Demostración. Supóngase $|C| = p^k$ con $k \geq 1$ y sea $g \in C$; nótese que $g \neq e$, porque la clase de la identidad tiene un solo elemento. La segunda relación de ortogonalidad (por columnas) aplicada a las columnas de e y de g, que son distintas, da

$$\sum_{\chi \in \mathrm{Irr}(G)} \chi(e)\,\overline{\chi(g)} = 0, \qquad \text{es decir} \qquad 1 + \sum_{\chi \neq \mathbf{1}} d_\chi\,\overline{\chi(g)} = 0.$$

Supóngase, buscando contradicción, que para todo carácter no trivial χ con $\chi(g) \neq 0$ se tiene $p \mid d_\chi$. Entonces

$$-\frac{1}{p} = \sum_{\chi \neq \mathbf{1},\ \chi(g) \neq 0} \frac{d_\chi}{p}\,\overline{\chi(g)}$$

sería un entero algebraico —suma de productos de enteros (d_χ/p) por enteros algebraicos $(\overline{\chi(g)})$— y a la vez un racional no entero: imposible.

Existe pues un carácter irreducible no trivial χ con $\chi(g) \neq 0$ y $p \nmid d_\chi$. Como $|C| = p^k$, se tiene $\gcd(|C|, d_\chi) = 1$, y el lema 27.43 dice que $\rho(g)$ es escalar.

Considérese ahora $Z = \{x \in G : \rho(x) \text{ es escalar}\}$. Es un subgrupo normal de G —las matrices escalares son centrales, y la condición se conserva por conjugación—, y contiene a $g \neq e$, luego $Z \neq \{e\}$. Por simplicidad, $Z = G$: toda la imagen de ρ consiste en escalares y es, por tanto, abeliana. Pero $\ker \rho$ es normal y no es todo G (pues χ no es trivial), luego por simplicidad $\ker \rho = \{e\}$ y ρ es fiel: G sería isomorfo a un grupo abeliano, contra la hipótesis de que G es no abeliano. □

Teorema 27.45 (Burnside)**.** *Todo grupo de orden $p^a q^b$, con p, q primos, es* resoluble.

Demostración. Por inducción fuerte sobre $|G|$. Si $|G| = 1$ no hay nada que probar.

Caso $a = 0$ o $b = 0$. G es un p-grupo. Todo p-grupo tiene centro no trivial —por la ecuación de clases, $|Z(G)| \equiv |G| \equiv 0 \pmod p$—, y $G/Z(G)$ es un p-grupo menor, resoluble por inducción; una extensión de un grupo resoluble por un grupo abeliano es resoluble.

Caso $a, b \geq 1$. Sea Q un q-subgrupo de Sylow, no trivial, y sea $g \in Z(Q)$ con $g \neq e$ (existe, otra vez por la ecuación de clases). Como Q centraliza a g, se tiene $Q \subset C_G(g)$, luego

$$|C| = [G : C_G(g)] \ \text{ divide a } \ [G : Q] = p^a,$$

es decir, la clase de conjugación de g tiene cardinal p^k para algún $k \geq 0$.

Si $k = 0$, entonces $g \in Z(G)$ y el centro es no trivial; $G/Z(G)$ tiene orden de la misma forma y es resoluble por inducción, luego G lo es.

Si $k \geq 1$, el lema 27.44 dice que G no puede ser simple no abeliano. Si G es abeliano, es resoluble. Si no, tiene un subgrupo normal propio no trivial N; tanto N como G/N tienen orden de la forma $p^i q^j$ y son estrictamente menores, luego resolubles por inducción, y una extensión de resoluble por resoluble es resoluble. □

Observación 27.46 (Sesenta años sin demostración algebraica). Burnside publicó este teorema en 1904 usando exactamente el argumento anterior, y durante décadas fue el ejemplo predilecto de un enunciado elemental —habla sólo de órdenes de grupos— cuya única prueba conocida pasaba por un territorio ajeno: matrices complejas, caracteres y enteros algebraicos. La búsqueda de una demostración "puramente algebraica" fue un programa explícito, y sólo se completó en los años ochenta, con trabajos de Goldschmidt, Matsuyama y Bender que son considerablemente más largos y difíciles que ésta. El episodio es una lección sobre qué significa que una demostración sea "natural": la de Burnside lo es, aunque cruce tres áreas.

Conviene además fijar el alcance. Con *tres* primos el teorema es falso: $|A_5| = 60 = 2^2 \cdot 3 \cdot 5$ y A_5 es simple no abeliano, luego no resoluble. No es un fallo marginal del enunciado —60 es el orden no resoluble más pequeño que existe—, sino la frontera exacta.

27.12. Funciones de clase y el producto tensorial de representaciones

Definición 27.47 (Función de clase)**.** Una *función de clase* sobre un grupo finito G es una aplicación $f : G \to \mathbb{C}$ constante sobre las clases de conjugación, es decir, con

$$f(hgh^{-1}) = f(g) \qquad \text{para todos } g, h \in G.$$

El conjunto de todas ellas, $\mathcal{Z}(G)$, es un espacio vectorial complejo de dimensión igual al número de clases de conjugación —una base evidente son los indicadores de cada clase—, dotado del producto escalar

$$\langle f_1, f_2 \rangle = \frac{1}{|G|} \sum_{g \in G} f_1(g)\, \overline{f_2(g)}.$$

Observación 27.48. Todo carácter es función de clase, porque la traza es invariante por conjugación: $\chi(hgh^{-1}) = \operatorname{tr}\big(\rho(h)\rho(g)\rho(h)^{-1}\big) = \operatorname{tr}\rho(g)$. Ése es el motivo de que una tabla de caracteres tenga una columna por clase y no una por elemento, y de que el teorema 27.40 —los caracteres irreducibles son base ortonormal de $\mathcal{Z}(G)$— cierre el recuento: hay tantas representaciones irreducibles como clases porque ambas cosas son la dimensión del mismo espacio.

Definición 27.49 (Producto tensorial de representaciones)**.** Sean (ρ_V, V) y (ρ_W, W) representaciones de G. El espacio $V \otimes W$ se convierte en representación mediante la acción *diagonal*

$$g \cdot (v \otimes w) = (\rho_V(g)v) \otimes (\rho_W(g)w),$$

extendida por linealidad. Su carácter es el producto de los caracteres:

$$\chi_{V \otimes W}(g) = \chi_V(g)\, \chi_W(g).$$

Cuando $V = W$, el espacio se descompone en dos subrepresentaciones,

$$V \otimes V = \operatorname{Sym}^2 V \ \oplus \ \Lambda^2 V,$$

generadas respectivamente por los tensores simétricos $v \otimes w + w \otimes v$ y los antisimétricos $v \otimes w - w \otimes v$, con caracteres

$$\chi_{\operatorname{Sym}^2}(g) = \tfrac{1}{2}(\chi(g)^2 + \chi(g^2)), \qquad \chi_{\Lambda^2}(g) = \tfrac{1}{2}(\chi(g)^2 - \chi(g^2)).$$

Justificación de las fórmulas. Que la acción diagonal es una representación se comprueba en una línea: $g \cdot (h \cdot (v \otimes w)) = (gh) \cdot (v \otimes w)$. Para el carácter, diagonalícese $\rho_V(g)$ con autovalores λ_i y $\rho_W(g)$ con μ_j; entonces los $v_i \otimes w_j$ son autovectores de g con autovalores $\lambda_i \mu_j$, y la traza es $\sum_{i,j} \lambda_i \mu_j = (\sum \lambda_i)(\sum \mu_j)$. Para la parte simétrica, una base son los $v_i \otimes v_j + v_j \otimes v_i$ con $i \le j$, de autovalores $\lambda_i \lambda_j$, y $\sum_{i \le j} \lambda_i \lambda_j = \frac{1}{2}[(\sum \lambda_i)^2 + \sum \lambda_i^2]$; el segundo sumando es $\chi(g^2)$ porque los autovalores de $\rho(g^2)$ son los λ_i^2. La antisimétrica sale igual con $i < j$ y el signo cambiado. □

Regla: Construir una tabla de caracteres desde cero

1. **Clases de conjugación.** Halla las clases y sus cardinales. El número de clases es el número de columnas y el de filas.
2. **Dimensiones.** Escribe $\sum_i d_i^2 = |G|$ con $d_1 = 1$ (el carácter trivial) y cada $d_i \mid |G|$ por el teorema 27.41. Suele quedar una sola posibilidad.
3. **Caracteres lineales.** Son los homomorfismos $G \to \mathbb{C}^\times$, y coinciden con los del abelianizado $G/[G,G]$: hay exactamente $|G/[G,G]|$ de ellos, todos de dimensión 1.
4. **Representaciones evidentes.** La de permutación (si G actúa sobre algo), la regular, o inducidas desde subgrupos. Resta el trivial de la de permutación para obtener la *estándar*.
5. **Fabrica más.** Multiplica caracteres ya obtenidos —el producto tensorial de la definición 27.49— y descompón con $\langle \chi, \chi_i \rangle$; toma Sym^2 y Λ^2 del estándar.
6. **Cierra por ortogonalidad.** Si falta una sola fila, se despeja imponiendo que las columnas sean ortogonales: $\sum_i \chi_i(g)\overline{\chi_i(h)} = 0$ si $g \not\sim h$, y $= |C_G(g)|$ si $g \sim h$.
7. **Comprueba.** Cada fila debe tener norma 1; filas distintas, producto 0; la primera columna, las dimensiones; y $\sum d_i^2 = |G|$.

27.13. Ejercicios resueltos

Tres problemas que construyen una tabla de caracteres desde cero y la usan para algo.

Ejercicio 27.1. Construya la tabla de caracteres de S_3 y compruebe las relaciones de ortogonalidad.

Solución. S_3 tiene tres clases de conjugación —la identidad, las tres transposiciones y los dos 3-ciclos—, luego tres irreducibles. La suma de los cuadrados de sus dimensiones ha de valer 6, y la única posibilidad es $1+1+4$.

	e	(12)	(123)
tamaño	1	3	2
χ_{triv}	1	1	1
χ_{sgn}	1	-1	1
χ_{est}	2	0	-1

La fila de la estándar sale restando la trivial a la representación por permutación, cuyo carácter cuenta puntos fijos: $3, 1, 0$.

Ortogonalidad: $\langle \chi_{\text{est}}, \chi_{\text{est}} \rangle = \frac{1}{6}(1\cdot 4 + 3\cdot 0 + 2\cdot 1) = 1$, luego es irreducible. Y $\langle \chi_{\text{triv}}, \chi_{\text{sgn}} \rangle = \frac{1}{6}(1-3+2) = 0$. Nótese que la comprobación de irreducibilidad se reduce a una suma de tres términos: eso es exactamente lo que la teoría de caracteres compra. ■

Ejercicio 27.2. Descomponga en irreducibles la representación por permutación de S_3 sobre $\mathbb{C}^3$ y la representación regular.

Solución. El carácter de la representación por permutación es $(3, 1, 0)$. Sus productos con las tres irreducibles valen 1, 0 y 1 respectivamente, luego $\mathbb{C}^3 \cong \text{triv} \oplus \text{est}$. Geométricamente: la recta de los vectores con todas las coordenadas iguales, que queda fija, y el plano de suma cero.

La regular tiene carácter $(6, 0, 0)$, y cada irreducible aparece tantas veces como su dimensión: $\text{triv} \oplus \text{sgn} \oplus 2\cdot\text{est}$, con $1+1+4=6$. Ése es el contenido del teorema que dice $|G| = \sum d_i^2$: no es una coincidencia numérica, es la descomposición de la regular contada en dimensiones. ■

Ejercicio 27.3. Demuestre que un grupo finito es abeliano si y sólo si todas sus representaciones irreducibles complejas tienen dimensión 1.

Solución. Si G es abeliano, cada clase de conjugación es un punto, luego hay $|G|$ clases y por tanto $|G|$ irreducibles; como $\sum d_i^2 = |G|$ con $|G|$ sumandos positivos, todos los d_i valen 1.

Recíprocamente, si todas las irreducibles tienen dimensión 1, entonces hay $|G|$ de ellas y por tanto $|G|$ clases de conjugación, lo que obliga a que cada clase sea un solo elemento, es decir, a que todo elemento conmute con todos. ■

27.14. Ejercicios propuestos

27.1. ¿Cuántas representaciones irreducibles complejas tiene $\mathbb{Z}/6\mathbb{Z}$ y de qué dimensiones?

27.2. Escriba los cuatro caracteres de $\mathbb{Z}/4\mathbb{Z}$.

27.3. ¿Cuántas irreducibles tiene D_4 y qué dimensiones? Compruebe la suma de cuadrados.

27.4. Halle el carácter de la representación por permutación de S_4 sobre $\mathbb{C}^4$ y descompóngala.

27.5. ¿Qué afirma el teorema de Maschke y dónde falla si la característica divide a $|G|$?

27.6. Demuestre que el carácter es constante en cada clase de conjugación.

27.7. ¿Cómo se detecta con caracteres si una representación es irreducible?

27.8. ¿Cuál es la dimensión de la representación de S_5 asociada a la partición $3+2$?

27.9. ¿Cuántas irreducibles tiene S_6? (No hace falta calcularlas.)

27.10. Descomponga $V_1 \otimes V_{1/2}$ para $SU(2)$ y compruebe las dimensiones.

27.11. ¿Por qué los caracteres de un grupo abeliano finito forman a su vez un grupo? ¿Cuál?

27.12. Explique en una frase qué tiene que ver el análisis de Fourier clásico con este capítulo.

Parte V

Análisis

Capítulo 28

Topología: la geometría de lo que no se rompe

Lo que necesitas traer

Del capítulo 6: límites, continuidad y la idea de sucesión convergente. Del capítulo 1: conjuntos, uniones e intersecciones arbitrarias. Del capítulo 23: qué es un grupo, imprescindible para la sección sobre el grupo fundamental.

La topología estudia las propiedades de las formas que sobreviven al estirarlas y doblarlas, pero no al romperlas o pegarlas. Es la "geometría de goma": para un topólogo, una taza y una dona son *la misma cosa* (ambas tienen exactamente un agujero, y una se deforma en la otra sin cortar). Este capítulo desarrolla las ideas con rigor —definiciones, teoremas, demostraciones— pero, como el resto del libro, con cajas de intuición que explican de qué va cada cosa antes de formalizarla, y un taller de ejemplos al final. Se apoya en el capítulo 6 (límites, sucesiones, compacidad) pero es en gran parte autocontenido.

28.1. La idea: geometría sin distancias

Idea clave

Intuición primero. Imagina que todas las figuras están hechas de goma infinitamente elástica. Puedes estirarlas, encogerlas y doblarlas cuanto quieras, pero no rasgarlas ni pegar partes que estaban separadas. Dos figuras son "la misma" para la topología si una se transforma en la otra con esas reglas. Un cuadrado y un círculo son iguales (uno se infla hasta el otro); una dona y una taza son iguales (el agujero de la dona es el asa de la taza); pero una esfera y una dona son *distintas*, porque no puedes crear un agujero sin romper. La topología es lo que queda de la geometría cuando olvidas las distancias y los ángulos, y solo recuerdas qué está "cerca" de qué y cuántos agujeros hay.

Para hacer esto preciso necesitamos capturar la noción de "cercanía" sin usar números concretos de distancia. El primer paso son los espacios métricos (que sí tienen distancia); el segundo, los espacios topológicos (que se quedan solo con lo esencial).

28.2. El vocabulario que el capítulo da por sabido: interior, clausura, acumulación, subespacio y subbase

Este capítulo construye la topología desde el principio, y las secciones que siguen dan la intuición de cada idea con calma: primero las distancias, luego los abiertos, luego los espacios topológicos abstractos. Pero hay media docena de palabras que el capítulo utiliza constantemente —*interior*, *clausura*, *frontera*, *punto de acumulación*, *subespacio*, *subbase*, *topología más fina*— y que nunca se detiene a definir formalmente. Esta sección las fija de una vez, en el orden lógico, para poder remitir a ellas después. Basta con saber, y las secciones siguientes lo explican con detalle, que una *topología* sobre un conjunto X es una familia τ de subconjuntos de X —llamados *abiertos*— que contiene a $\varnothing$ y a X y es estable por uniones arbitrarias y por intersecciones finitas; que un conjunto es *cerrado* si su complemento es abierto; y que en un espacio métrico los abiertos son los conjuntos que contienen una bola alrededor de cada uno de sus puntos.

Idea clave

Intuición primero. Las cinco palabras de esta sección responden a cinco preguntas muy concretas. *Interior*: ¿qué puntos de un conjunto tienen sitio a su alrededor? *Clausura*: ¿qué puntos están tan pegados al conjunto que no se los puede separar de él? *Frontera*: ¿cuál es la piel que separa el conjunto de su exterior? *Punto de acumulación*: ¿a qué puntos se puede llegar aproximándose desde el conjunto sin llegar nunca? Y *subespacio*: si me encierro dentro de un trozo del espacio y solo veo lo que hay ahí, ¿qué me parece abierto? Esta última pregunta tiene una respuesta que sorprende a todo el mundo la primera vez, y por eso lleva su propia advertencia.

Definición 28.1 (Interior, clausura, frontera y exterior)**.** Sea X un espacio topológico y $A \subseteq X$.

- El *interior* de A, denotado $\operatorname{int} A$ o $\mathring{A}$, es la unión de todos los abiertos contenidos en A. Por tanto es el mayor abierto contenido en A, y $x \in \operatorname{int} A$ si y solo si existe un abierto U con $x \in U \subseteq A$.
- La *clausura* de A, denotada $\overline{A}$, es la intersección de todos los cerrados que contienen a A. Por tanto es el menor cerrado que contiene a A, y $x \in \overline{A}$ si y solo si todo abierto que contiene a x corta a A.
- La *frontera* de A es $\partial A = \overline{A} \setminus \operatorname{int} A = \overline{A} \cap \overline{X \setminus A}$.
- El *exterior* de A es $\operatorname{ext} A = \operatorname{int}(X \setminus A) = X \setminus \overline{A}$.

Se dice que A es *denso* en X si $\overline{A} = X$. Los tres conjuntos $\operatorname{int} A$, ∂A y $\operatorname{ext} A$ son disjuntos dos a dos y su unión es todo X: cada punto del espacio está en exactamente uno de ellos, y esa partición es el control de sensatez de cualquier cálculo.

Proposición 28.2. *A es abierto si y solo si $A = \operatorname{int} A$; y A es cerrado si y solo si $A = \overline{A}$. Además $\operatorname{int}(X \setminus A) = X \setminus \overline{A}$ y $\overline{X \setminus A} = X \setminus \operatorname{int} A$.*

Demostración. Si A es abierto, A es uno de los abiertos contenidos en A, luego $A \subseteq \operatorname{int} A$; y la otra inclusión vale siempre. Recíprocamente, $\operatorname{int} A$ es unión de abiertos, luego abierto, y si coincide con A entonces A lo es. El caso cerrado es idéntico intercambiando uniones por intersecciones. Para la dualidad: $x \notin \overline{A}$ significa que hay un abierto U con $x \in U$ y $U \cap A = \varnothing$, es decir $x \in U \subseteq X \setminus A$, que es exactamente $x \in \operatorname{int}(X \setminus A)$. La segunda identidad es la primera aplicada a $X \setminus A$. □

Definición 28.3 (Punto de acumulación, punto aislado y conjunto derivado)**.** Sea $A \subseteq X$ y $x \in X$. Se dice que x es un *punto de acumulación* (o *punto límite*) de A si todo abierto que contiene a x corta a A *en algún punto distinto del propio* x:

$$\text{para todo abierto } U \ni x, \qquad (U \setminus \{x\}) \cap A \neq \varnothing.$$

El conjunto de todos ellos se llama *conjunto derivado* de A y se denota A'. Un punto $x \in A$ que *no* es de acumulación de A se llama *punto aislado* de A: existe entonces un abierto U con $U \cap A = \{x\}$.

La cláusula "distinto del propio x" es la única diferencia con la clausura, y es toda la diferencia: un punto de A pertenece siempre a $\overline{A}$, pero no tiene por qué pertenecer a A'.

Proposición 28.4. *Para todo $A \subseteq X$ se cumple $\overline{A} = A \cup A'$. En consecuencia, A es cerrado si y solo si $A' \subseteq A$:* un conjunto es cerrado exactamente cuando contiene a todos sus puntos de acumulación.

Demostración. Si $x \in A$, entonces $x \in \overline{A}$ trivialmente; y si $x \in A'$, todo abierto que contiene a x corta a A, luego $x \in \overline{A}$. Eso da $A \cup A' \subseteq \overline{A}$. Recíprocamente, sea $x \in \overline{A}$ y supongamos $x \notin A$. Todo abierto U con $x \in U$ corta a A; y como $x \notin A$, ese punto de corte es automáticamente distinto de x, de modo que $x \in A'$. Para la consecuencia: A es cerrado $\iff A = \overline{A} = A \cup A' \iff A' \subseteq A$. □

Ejemplo 28.5 (Cinco conjuntos derivados)**.** Todos en $\mathbb{R}$ con la topología usual.

(a) Un intervalo. $A = (0,1)$. Entonces $A' = [0,1]$: los puntos 0 y 1 son de acumulación sin pertenecer a A. Ningún punto de A está aislado, y $\overline{A} = A \cup A' = [0,1]$. ✓

(b) Una sucesión. $A = \{1, \frac{1}{2}, \frac{1}{3}, \frac{1}{4}, \dots\}$. Cada punto $1/k$ está *aislado*: el intervalo abierto $(\frac{1}{k+1}, \frac{1}{k-1})$ corta a A solo en $1/k$; por ejemplo $(\frac{1}{4}, \frac{1}{2}) \cap A = \{\frac{1}{3}\}$. En cambio 0 sí es de acumulación, porque todo intervalo $(-\varepsilon, \varepsilon)$ contiene infinitos términos $1/n$ (todos los $n > 1/\varepsilon$). Por tanto $A' = \{0\}$ y $\overline{A} = A \cup \{0\}$. Nótese que $0 \notin A$: el conjunto derivado no tiene por qué estar contenido en el conjunto.

(c) Los enteros. $A = \mathbb{Z}$. Cada entero está aislado, porque $(n - \frac{1}{2}, n + \frac{1}{2}) \cap \mathbb{Z} = \{n\}$, y ningún real es de acumulación de $\mathbb{Z}$. Luego $A' = \varnothing$ y $\mathbb{Z}$ es cerrado (contiene trivialmente a su derivado vacío). Un conjunto infinito, cerrado y sin ningún punto de acumulación.

(d) Los racionales. $A = \mathbb{Q}$. Todo intervalo abierto contiene infinitos racionales, luego $\mathbb{Q}' = \mathbb{R}$ y $\overline{\mathbb{Q}} = \mathbb{R}$: $\mathbb{Q}$ es denso. No tiene ningún punto aislado.

(e) Una mezcla. $A = (0,1) \cup \{2\}$. Aquí $A' = [0,1]$ —el punto 2 no es de acumulación, porque (1,5; 2,5) corta a A solo en el propio 2— y 2 es el único punto aislado. Entonces

$\overline{A} = [0,1] \cup \{2\}$, $\operatorname{int} A = (0,1)$ y $\partial A = \{0,1,2\}$. Comprobación de la partición: $\mathbb{R} = (0,1) \cup \{0,1,2\} \cup ((-\infty,0) \cup (1,2) \cup (2,\infty))$, tres trozos disjuntos que lo cubren todo. ✓

Definición 28.6 (Topología de subespacio)**.** Sea (X,τ) un espacio topológico y $Y \subseteq X$ un subconjunto cualquiera. La *topología de subespacio* (también llamada *relativa* o *inducida*) sobre Y es

$$\tau_Y = \{\, U \cap Y \;:\; U \in \tau \,\},$$

es decir: los abiertos de Y son exactamente las trazas sobre Y de los abiertos de X. Con ella, Y se llama *subespacio* de X. Cuando haga falta distinguir, se dirá que un conjunto es *abierto en* Y o *abierto en* X, nunca "abierto" a secas.

Proposición 28.7. *τ_Y es efectivamente una topología sobre Y.*

Demostración. $\varnothing = \varnothing \cap Y$ e $Y = X \cap Y$ están en τ_Y. Para las uniones, si $V_\alpha = U_\alpha \cap Y$ con $U_\alpha \in \tau$, entonces

$$\bigcup_\alpha V_\alpha = \bigcup_\alpha (U_\alpha \cap Y) = \Big(\bigcup_\alpha U_\alpha\Big) \cap Y,$$

y la unión de los U_α es abierta en X. Para las intersecciones finitas, el mismo cálculo con $\cap$ en lugar de $\cup$: $(U_1 \cap Y) \cap (U_2 \cap Y) = (U_1 \cap U_2) \cap Y$. Todo se reduce a que la intersección con un conjunto fijo conmuta con uniones e intersecciones. □

Cuidado

Aquí está la trampa que hay que interiorizar de una vez: **un mismo conjunto puede ser abierto en un subespacio y no serlo en el espacio total.** "Abierto" no es una propiedad de un conjunto; es una propiedad de la pareja formada por el conjunto y el ambiente donde se le mira. Tres testigos, de menor a mayor tamaño:

(i) El ejemplo mínimo, con tres puntos. Sea $X = \{a,b,c\}$ con la topología $\tau = \{\varnothing, \{a\}, \{a,b\}, X\}$ —compruébese que es topología: las uniones e intersecciones de esos cuatro conjuntos vuelven a estar en la lista— y sea $Y = \{b,c\}$. Intersecando cada abierto con Y se obtiene

$$\varnothing \cap Y = \varnothing, \quad \{a\} \cap Y = \varnothing, \quad \{a,b\} \cap Y = \{b\}, \quad X \cap Y = Y,$$

de modo que $\tau_Y = \{\varnothing, \{b\}, Y\}$. Así pues $\{b\}$ *es abierto en* Y y *no es abierto en* X, porque $\{b\} \notin \tau$. Todo el fenómeno cabe en cuatro conjuntos.

(ii) El ejemplo de la recta. En $\mathbb{R}$, el intervalo $[0,1)$ no es abierto: ninguna bola centrada en 0 cabe dentro, porque siempre asoma hacia los negativos. Pero en el subespacio $Y = [0,2]$ sí lo es, porque

$$[0,1) = (-1,1) \cap [0,2],$$

traza de un abierto de $\mathbb{R}$. El punto 0 tiene sitio a su alrededor *dentro de* Y: a su izquierda, en Y, no hay nada.

(iii) Un conjunto abierto y cerrado a la vez. Sea $Y = [0,1] \cup [2,3]$ con la topología inducida de $\mathbb{R}$. El trozo $[0,1]$ es cerrado en $\mathbb{R}$, y sin embargo es *abierto en* Y, porque $[0,1] = (-1;\,1{,}5) \cap Y$. También es cerrado en Y, por ser cerrado en $\mathbb{R}$. Que exista un

subconjunto propio no vacío abierto y cerrado a la vez es precisamente lo que significa que Y no sea conexo, y así es como se demuestran las desconexiones en la práctica.

Moraleja operativa: ante la pregunta "¿es A abierto?", la respuesta correcta empieza siempre por otra pregunta, *¿abierto en qué?*

Proposición 28.8 (Cerrados y clausuras relativas)**.** *Sea $Y \subseteq X$ un subespacio y $A \subseteq Y$.*

1. *$F \subseteq Y$ es cerrado en Y si y solo si $F = C \cap Y$ para algún cerrado C de X.*
2. *La clausura de A en Y es $\overline{A}^{Y} = \overline{A} \cap Y$, donde $\overline{A}$ es la clausura en X.*
3. *Si Y es abierto en X, todo abierto de Y es abierto en X. Si Y es cerrado en X, todo cerrado de Y es cerrado en X.*

Demostración. (1) F es cerrado en Y si y solo si $Y \setminus F$ es abierto en Y, o sea $Y \setminus F = U \cap Y$ con U abierto en X; tomando complementos dentro de Y queda $F = (X \setminus U) \cap Y$, y $C = X \setminus U$ es cerrado. El recíproco se recorre al revés.

(2) Por (1), $\overline{A} \cap Y$ es un cerrado de Y que contiene a A, luego $\overline{A}^{Y} \subseteq \overline{A} \cap Y$. Para la otra inclusión, $\overline{A}^{Y} = C \cap Y$ con C cerrado en X y $A \subseteq C$; entonces $\overline{A} \subseteq C$, y por tanto $\overline{A} \cap Y \subseteq C \cap Y = \overline{A}^{Y}$.

(3) Si Y es abierto y $V = U \cap Y$ con U abierto en X, entonces V es intersección de dos abiertos de X, luego abierto en X. El caso cerrado es el mismo argumento con cerrados. □

Ejemplo 28.9 (Un subespacio donde casi todo está aislado)**.** Sea $Y = \{0\} \cup \{\frac{1}{n} : n \geq 1\} \subseteq \mathbb{R}$ con la topología inducida. Para cada $k \geq 2$, el conjunto $\{\frac{1}{k}\}$ *es abierto en* Y: en efecto,

$$\left(\tfrac{1}{k+1}, \tfrac{1}{k-1}\right) \cap Y = \left\{\tfrac{1}{k}\right\},$$

por ejemplo $(\frac{1}{4}, \frac{1}{2}) \cap Y = \{\frac{1}{3}\}$; y $\{1\} = (\frac{1}{2}, 2) \cap Y$. En cambio $\{0\}$ *no es abierto en* Y: cualquier intervalo $(-\varepsilon, \varepsilon)$ contiene infinitos puntos $1/n$, así que su traza sobre Y nunca se reduce a $\{0\}$. Y $\{0\}$ tampoco puede escribirse como unión de los abiertos anteriores.

La consecuencia es bonita: Y es "casi discreto" —todos sus puntos salvo uno están aislados— y sin embargo no es discreto, y esa única excepción es lo que lo hace compacto. Es el ejemplo mínimo de espacio compacto infinito, y todo su interés vive en el punto 0.

Regla: responder a la pregunta "¿es abierto?" sin equivocarse de ambiente

1. **Identifica el espacio ambiente antes que nada.** Escríbelo: "trabajo dentro de Y". Si el enunciado no lo dice, el ambiente por defecto es el espacio total del que se venía hablando.
2. **Si el ambiente es un subespacio Y, busca un abierto U del total con $A = U \cap Y$.** Encontrar ese U *es* la demostración completa, y suele ser un intervalo o una bola muy sencilla.
3. **Para negarlo, exhibe un punto sin sitio.** A no es abierto en Y si hay un $x \in A$

tal que todo abierto de Y que lo contiene se sale de A. En espacios métricos basta comprobar que ninguna bola $B(x,r)\cap Y$ cabe dentro de A.

4. **Usa el atajo de la proposición 28.8(3).** Si Y es abierto en X, no hay diferencia entre las dos preguntas para los abiertos; si Y es cerrado, no la hay para los cerrados. Toda la sutileza aparece cuando Y no es ni una cosa ni la otra, como $[0,2)$ en $\mathbb{R}$.

5. **Para clausuras, interseca y no recalcules.** Por la proposición 28.8(2), la clausura relativa es la clausura de siempre cortada con Y. No hay que rehacer el trabajo.

6. **Recuerda que abierto y cerrado no son excluyentes ni exhaustivos.** Un conjunto puede ser ambos (y entonces delata una desconexión) o ninguno. Son dos preguntas independientes y hay que responderlas por separado.

7. **Control.** $\varnothing$ e Y son siempre abiertos en Y. Si tu lista de abiertos relativos no los incluye, has calculado mal.

Definición 28.10 (Subbase y topología generada)**.** Sea X un conjunto y $\mathcal{S}$ una familia cualquiera de subconjuntos de X cuya unión sea X. La *topología generada* por $\mathcal{S}$ es la menor topología sobre X que contiene a $\mathcal{S}$, y se describe explícitamente en dos pasos:

1. se forman todas las intersecciones *finitas* de miembros de $\mathcal{S}$, obteniendo una familia $\mathcal{B}$;

2. se forman todas las uniones *arbitrarias* de miembros de $\mathcal{B}$.

La familia $\mathcal{B}$ resultante es una base de esa topología, y se dice entonces que $\mathcal{S}$ es una *subbase*. Dicho de otro modo: una base genera los abiertos con uniones, y una subbase los genera con intersecciones finitas primero y uniones después.

Proposición 28.11. *Con la notación anterior, la familia $\mathcal{B}$ de las intersecciones finitas de miembros de $\mathcal{S}$ cumple las dos condiciones de base: cubre X, y la intersección de dos miembros de $\mathcal{B}$ es de nuevo un miembro de $\mathcal{B}$. En consecuencia,* toda *familia $\mathcal{S}$ que cubra X es subbase de exactamente una topología.*

Demostración. $\mathcal{B}$ cubre X porque $\mathcal{S}$ ya lo cubre y $\mathcal{S}\subseteq\mathcal{B}$ (una intersección de un solo elemento). Y si $B_1 = S_1\cap\cdots\cap S_p$ y $B_2 = S'_1\cap\cdots\cap S'_q$ son intersecciones finitas de miembros de $\mathcal{S}$, entonces $B_1\cap B_2$ es la intersección de los $p+q$ conjuntos juntos, que sigue siendo finita: pertenece a $\mathcal{B}$. La condición de base se cumple entonces de la forma más fuerte posible, con $B_3 = B_1\cap B_2$. Que la topología resultante es la menor que contiene a $\mathcal{S}$ es inmediato: cualquier topología que contenga a $\mathcal{S}$ debe contener sus intersecciones finitas y las uniones de estas. □

Ejemplo 28.12 (Dos semirrectas generan toda la recta)**.** Sobre $X=\mathbb{R}$ tómese

$$\mathcal{S} = \{(a,\infty) : a\in\mathbb{R}\} \;\cup\; \{(-\infty,b) : b\in\mathbb{R}\}.$$

Cubre $\mathbb{R}$ (de hecho cada semirrecta ya cubre casi todo). Las intersecciones finitas producen exactamente tres tipos de conjunto: las propias semirrectas, los intervalos acotados $(a,b) =$

$(a,\infty)\cap(-\infty,b)$, y el vacío cuando $a\geq b$. Intersecar más semirrectas no da nada nuevo, porque $(a_1,\infty)\cap(a_2,\infty)=(\max(a_1,a_2),\infty)$. Así que $\mathcal{B}$ es la familia de los intervalos abiertos y las semirrectas, y sus uniones son precisamente los abiertos usuales de $\mathbb{R}$: $\mathcal{S}$ es una subbase de la topología usual.

La utilidad práctica es inmediata y aparece varias veces en este capítulo: para comprobar que una función $f\colon X\to\mathbb{R}$ es continua basta verificar que $f^{-1}((a,\infty))$ y $f^{-1}((-\infty,b))$ son abiertos, porque la preimagen respeta uniones e intersecciones y, por tanto, lo que vale sobre una subbase vale sobre toda la topología. Dos familias de comprobaciones en lugar de infinitas.

Regla: trabajar con una base o una subbase en lugar de con todos los abiertos

1. **Comprueba primero que tu familia cubre el espacio.** Sin eso no genera ninguna topología, porque el total tiene que ser abierto.
2. **Para una base, verifica la condición de encaje:** si $x\in B_1\cap B_2$, debe existir B_3 de la familia con $x\in B_3\subseteq B_1\cap B_2$. Para una subbase no hay nada que verificar: cualquier familia que cubra sirve.
3. **Al probar continuidad, comprueba solo sobre la subbase.** f es continua si y solo si $f^{-1}(S)$ es abierto para todo S de una subbase del destino. Es el ahorro más grande que ofrece este vocabulario.
4. **Al probar que un conjunto es abierto, basta encontrar un básico dentro.** U es abierto si y solo si cada $x\in U$ tiene un miembro de la base con $x\in B\subseteq U$.
5. **Para comparar dos topologías, compara sus bases.** τ_1 es más fina que τ_2 si y solo si cada básico de τ_2 contiene, alrededor de cada uno de sus puntos, un básico de τ_1.
6. **Cuenta el tamaño de la base para los axiomas de numerabilidad.** Que exista una base numerable es una propiedad fuerte y se comprueba exhibiendo una: en $\mathbb{R}$, los intervalos de extremos racionales.
7. **Control.** Los abiertos generados nunca son menos que la familia de partida ni más que todos los subconjuntos. Si tu topología generada no contiene a $\mathcal{S}$, hay un error en el paso de las intersecciones.

Definición 28.13 (Comparación de topologías)**.** Si τ_1 y τ_2 son dos topologías sobre el mismo conjunto X y $\tau_2\subseteq\tau_1$, se dice que τ_1 es *más fina* que τ_2, y que τ_2 es *más gruesa* que τ_1. Más fina significa *más abiertos*: más conjuntos declarados abiertos, más funciones continuas saliendo de X y menos entrando en él. La topología discreta es la más fina de todas y la indiscreta la más gruesa; dos topologías pueden además no ser comparables en ninguno de los dos sentidos.

Pausa para pensar

Junta ahora las dos ideas de esta sección y verás que son la misma vista desde dos sitios. La topología de subespacio es la *más gruesa* que hace continua la inclusión $Y \hookrightarrow X$: se declaran abiertos los mínimos conjuntos necesarios para que meter Y dentro de X no rompa nada. La topología generada por una subbase es la *más gruesa* que contiene a la familia dada. Y la topología producto, que aparece más adelante, es la más gruesa que hace continuas todas las proyecciones. Tres construcciones, un solo patrón: *tómese la topología más económica que hace continuo lo que uno quiere que sea continuo*. Cuando ese patrón se abstrae del todo se llama topología inicial, y es el motivo de que las tres definiciones se parezcan tanto.

28.3. Espacios métricos

Idea clave

Intuición primero. Un espacio métrico es un conjunto donde se ha decidido qué significa "distancia", y lo interesante es lo poco que hay que exigirle: que nunca sea negativa, que valga cero solo entre un punto y sí mismo, que sea simétrica, y la desigualdad triangular (el atajo directo nunca es más largo que pasar por un tercer punto). Nada más. Con esas cuatro reglas ya se puede hablar de convergencia, de continuidad y de todo lo demás.
Lo llamativo es cuántas distancias distintas cumplen esas reglas sobre el mismo conjunto. En el plano puedes medir en línea recta, o al modo de un taxi que solo circula por calles perpendiculares, o quedarte con la mayor de las dos diferencias de coordenadas: las tres son métricas legítimas, dan bolas de formas distintas (un círculo, un rombo, un cuadrado) y, sin embargo, dan exactamente las mismas sucesiones convergentes y las mismas funciones continuas. Eso sugiere que la distancia concreta era un andamio y que la información esencial estaba en otro sitio: en qué conjuntos son abiertos. Esa sospecha es lo que lleva, dos secciones más adelante, a tirar la distancia y quedarse solo con los abiertos.

Un *espacio métrico* es un conjunto donde sabemos medir distancias, con las propiedades mínimas que uno esperaría de una distancia.

Definición 28.14. Un *espacio métrico* es un conjunto X con una función *distancia* $d(x, y)$ que a cada par de puntos le asigna un número ≥ 0, cumpliendo:

1. $d(x, y) \geq 0$, y $d(x, y) = 0$ solo si $x = y$ (la distancia es cero solo de un punto a sí mismo);
2. $d(x, y) = d(y, x)$ (ir y volver mide lo mismo);
3. $d(x, z) \leq d(x, y) + d(y, z)$ (la *desigualdad triangular*: el camino directo no es más largo que el que pasa por un tercer punto).

Ejemplo 28.15 (Los ejemplos que hay que tener en mente)**.** El más importante: $\mathbb{R}^n$ con la distancia usual $d(x, y) = \sqrt{\sum (x_i - y_i)^2}$ (la del capítulo anterior, Pitágoras). Otro,

sorprendente: la *distancia discreta*, donde $d(x,y) = 1$ si $x \neq y$ y 0 si $x = y$ —como si todos los puntos estuvieran a la misma distancia unos de otros, sin ninguno más cerca que otro—. Los espacios métricos abarcan desde la recta real hasta espacios de funciones.

Definición 28.16. La *bola abierta* de centro x y radio r es el conjunto de puntos a distancia menor que r:

$$B(x,r) = \{\, y \in X : d(x,y) < r \,\}.$$

En $\mathbb{R}$ es un intervalo abierto; en $\mathbb{R}^2$, un disco sin su borde; en $\mathbb{R}^3$, una bola sólida sin cáscara.

28.4. Conjuntos abiertos y cerrados

Estas dos palabras son el vocabulario básico de la topología.

Definición 28.17. Un conjunto U es *abierto* si alrededor de cada uno de sus puntos cabe una bolita enteramente dentro de U: para todo $x \in U$ existe $r > 0$ con $B(x,r) \subseteq U$. Un conjunto es *cerrado* si su complemento es abierto.

Idea clave

Intuición primero. "Abierto" significa *sin frontera incluida*: desde cualquier punto puedes moverte un poquito en cualquier dirección sin salirte. El intervalo $(0,1)$ es abierto: estés donde estés dentro, siempre hay algo de espacio a ambos lados. En cambio $[0,1]$ *no* es abierto, porque en el punto 0 (que sí pertenece) cualquier bolita se sale hacia los negativos: el extremo te tiene atrapado contra el borde. "Cerrado" es lo contrario: incluye toda su frontera. Cuidado: abierto y cerrado no son opuestos como en las puertas; hay conjuntos que no son ni lo uno ni lo otro (como $[0,1)$), y conjuntos que son ambos a la vez (como el propio X y el vacío).

Ejemplo 28.18. $(0,1)$ es abierto en $\mathbb{R}$; $[0,1]$ es cerrado (su complemento $(-\infty,0) \cup (1,\infty)$ es abierto); $[0,1)$ no es ni abierto (falla en 0) ni cerrado (su complemento falla en 1).

Regla: demostrar que un conjunto es abierto o que es cerrado

Casi todos los ejercicios iniciales de topología son de este tipo, y casi todos se resuelven sin tocar la definición, usando en su lugar preimágenes por funciones continuas. Conviene tener las cinco vías ordenadas de la más cómoda a la menos.

1. **Vía preimagen (la que hay que probar primero).** Si f es continua, f^{-1}(abierto) es abierto y f^{-1}(cerrado) es cerrado. Traduce el conjunto a una condición sobre funciones continuas y mira el signo de la desigualdad: *estricta* da abierto, *no estricta* da cerrado. Así, $\{(x,y) : x^2 + y^2 < 1\}$ es abierto y $\{(x,y) : x^2 + y^2 \leq 1\}$ es cerrado, sin más argumento que la continuidad de $x^2 + y^2$.
2. **Vía operaciones.** Unión arbitraria e intersección finita de abiertos es abierta; intersección arbitraria y unión finita de cerrados es cerrada. Descompón el conjunto

y aplica esto. *Ojo con las palabras "arbitraria" y "finita": están cruzadas, y ese cruce es el corazón de la definición de topología.*

3. **Vía complementario.** Cerrado significa, por definición, complementario de abierto. Muchas veces el complementario es visiblemente más fácil.

4. **Vía bolas (la definición directa).** Para probar que U es abierto en un espacio métrico: toma $x \in U$ cualquiera y *construye explícitamente* un radio $r > 0$ con $B(x, r) \subset U$. El radio casi siempre se lee como "la distancia de x al borde", y hay que escribirlo, no invocarlo.

5. **Vía sucesiones, solo para cerrados y solo en espacios métricos.** F es cerrado si y solo si toda sucesión de puntos de F que converge tiene su límite en F. Es la vía más intuitiva, pero cuidado: en espacios topológicos generales las sucesiones no bastan y hay que usar redes.

6. **Recuerda que "abierto" y "cerrado" no son opuestos.** Un conjunto puede ser ambas cosas ($\emptyset$ y el total siempre lo son; en un espacio conexo son los únicos), y puede no ser ninguna ($[0, 1)$ en $\mathbb{R}$). La pregunta "¿es abierto o cerrado?" está mal planteada; las preguntas correctas son dos, y se responden por separado.

7. **Y que todo depende del ambiente.** $[0, 1)$ no es abierto en $\mathbb{R}$, pero *sí* es abierto en el subespacio $[0, 2]$, porque $[0, 1) = (-1, 1) \cap [0, 2]$. Antes de responder, pregúntate siempre: ¿abierto *en* qué?

Ejemplo 28.19 (Cinco conjuntos, cinco veredictos)**. (a)** $U = \{(x, y) \in \mathbb{R}^2 : y > x^2\}$ **es abierto.** La función $g(x, y) = y - x^2$ es continua y $U = g^{-1}((0, \infty))$, preimagen de un abierto. Abierto. (Vía 1, en una línea.)

(b) $F = \{(x, y) : xy = 1\}$ **es cerrado.** Es $h^{-1}(\{1\})$ con $h(x, y) = xy$ continua y $\{1\}$ cerrado en $\mathbb{R}$. Cerrado. Y obsérvese que es cerrado *pero no acotado*, luego no compacto: la hipérbola es el recordatorio de que cerrado no implica compacto.

(c) $A = [0, 1)$ **en** $\mathbb{R}$ **no es ni abierto ni cerrado.** No es abierto porque ninguna bola centrada en 0 cabe dentro ($B(0, r)$ contiene puntos negativos para todo $r > 0$). No es cerrado porque la sucesión $x_n = 1 - 1/n$ está en A, converge a 1, y $1 \notin A$. Ambos fallos, cada uno con su testigo.

(d) $\mathbb{Q}$ **en** $\mathbb{R}$ **no es ni abierto ni cerrado, y su interior y su clausura son extremos.** Ninguna bola cabe dentro de $\mathbb{Q}$ (todo intervalo contiene irracionales), luego $\operatorname{int}(\mathbb{Q}) = \emptyset$. Y toda bola corta a $\mathbb{Q}$, luego $\overline{\mathbb{Q}} = \mathbb{R}$. Un conjunto de interior vacío y clausura total: la frontera es todo $\mathbb{R}$.

(e) La vía 4, escrita con el radio explícito. Probemos directamente que $U = \{x \in \mathbb{R} : |x - 3| < 2\} = (1, 5)$ es abierto. Dado $x \in U$, tómese

$$r = 2 - |x - 3| > 0.$$

Si $|y - x| < r$ entonces $|y - 3| \leq |y - x| + |x - 3| < r + |x - 3| = 2$, luego $y \in U$. Es decir $B(x, r) \subset U$. El radio no se sacó de la nada: es exactamente lo que le falta a x para llegar al

borde.

Cuidado

La intersección *infinita* de abiertos puede no ser abierta. El ejemplo obligatorio:

$$\bigcap_{n\geq 1}\left(-\frac{1}{n},\frac{1}{n}\right)=\{0\},$$

intersección de abiertos que da un conjunto cerrado de interior vacío. Por eso la definición de topología exige intersecciones *finitas* y no más. Dualmente, $\bigcup_{n\geq 1}[\frac{1}{n},1]=(0,1]$, unión infinita de cerrados que no es cerrada.

Regla: calcular el interior, la clausura y la frontera de un conjunto

Estas tres operaciones son el pan de cada día de la topología, y se calculan siempre de la misma manera: preguntando, punto por punto, qué ven los entornos.

1. **Interior: los puntos que tienen sitio a su alrededor.** $x\in \operatorname{int} A$ si existe un abierto U con $x\in U\subseteq A$. Equivalentemente, $\operatorname{int} A$ es el mayor abierto contenido en A, es decir la unión de todos los abiertos contenidos en A.

2. **Clausura: los puntos que no se pueden separar de A.** $x\in\overline{A}$ si *todo* entorno de x corta a A. Equivalentemente, $\overline{A}$ es el menor cerrado que contiene a A. En un espacio métrico hay un criterio más cómodo: $x\in\overline{A}$ si y solo si existe una sucesión de puntos de A que converge a x.

3. **Frontera: lo que sobra.** $\partial A=\overline{A}\setminus\operatorname{int} A$. Un punto está en la frontera cuando todo entorno suyo corta a A *y* a su complemento. La frontera es siempre cerrada, porque $\partial A=\overline{A}\cap\overline{A^c}$.

4. **Usa la dualidad para no calcular dos veces.** Siempre se cumple $\operatorname{int}(A^c)=(\overline{A})^c$ y $\overline{A^c}=(\operatorname{int} A)^c$. Si ya tienes la clausura de A, el interior del complemento sale gratis.

5. **Comprueba las caracterizaciones.** A es abierto $\iff A=\operatorname{int} A$. A es cerrado $\iff A=\overline{A}$. A es denso $\iff \overline{A}=X$. A es denso en ninguna parte $\iff \operatorname{int}(\overline{A})=\varnothing$.

6. **No supongas que estas operaciones respetan uniones e intersecciones.** Se cumple $\overline{A\cup B}=\overline{A}\cup\overline{B}$ y $\operatorname{int}(A\cap B)=\operatorname{int} A\cap\operatorname{int} B$, pero solo hay *contenidos* en los otros dos casos, y solo para familias finitas. Con familias infinitas incluso la unión falla.

7. **Verifica con la partición del espacio.** Todo punto del espacio está exactamente en uno de tres lugares: $\operatorname{int} A$, ∂A o $\operatorname{int}(A^c)$. Si tus tres conjuntos no son disjuntos o no cubren todo X, hay un error.

Ejemplo 28.20 (Seis conjuntos, tres topologías, y el conjunto cuya frontera es todo)**.** Aplicamos el procedimiento en orden creciente de rareza. Salvo aviso, la topología es la usual.

(a) Un intervalo semiabierto en $\mathbb{R}$. Sea $A = [0, 1)$. Ningún entorno de 0 cabe dentro de A (siempre asoma por la izquierda), así que $\operatorname{int} A = (0, 1)$. El punto 1 es límite de $1 - \frac{1}{n} \in A$, luego $\overline{A} = [0, 1]$. Y por tanto $\partial A = \{0, 1\}$. Comprobación del paso 7: $\mathbb{R}$ se parte en $(0, 1)$, $\{0, 1\}$ y $(-\infty, 0) \cup (1, \infty)$. ✓

(b) Los racionales. Sea $A = \mathbb{Q} \subseteq \mathbb{R}$. Todo intervalo contiene irracionales, luego $\operatorname{int} \mathbb{Q} = \varnothing$; todo intervalo contiene racionales, luego $\overline{\mathbb{Q}} = \mathbb{R}$. Así que $\partial\mathbb{Q} = \mathbb{R}$: la frontera de $\mathbb{Q}$ es *todo el espacio*. Y aplicando el paso 4 al complemento, los irracionales también tienen interior vacío y clausura $\mathbb{R}$. Dos conjuntos disjuntos, ambos densos.

Cuidado

$\mathbb{Q}$ es denso, luego *no* es denso en ninguna parte: $\operatorname{int}(\overline{\mathbb{Q}}) = \operatorname{int}(\mathbb{R}) = \mathbb{R} \neq \varnothing$. Las dos expresiones se parecen y significan cosas casi opuestas. «Denso en ninguna parte» pide que la clausura tenga interior vacío, no que el conjunto lo tenga.

(c) La sucesión $\{1/n\}$. Sea $A = \{1, \frac{1}{2}, \frac{1}{3}, \ldots\} \subseteq \mathbb{R}$. Cada punto está aislado, así que $\operatorname{int} A = \varnothing$; pero 0 es límite de la sucesión, luego $\overline{A} = A \cup \{0\}$ y $\partial A = A \cup \{0\}$. Este conjunto sí es denso en ninguna parte, y además $\overline{A}$ es compacto: es el ejemplo mínimo de espacio compacto infinito.

(d) Un disco en el plano, y el mismo disco visto desde $\mathbb{R}^3$. En $\mathbb{R}^2$, para $A = \{x^2 + y^2 < 1\}$ se tiene $\operatorname{int} A = A$, $\overline{A} = \{x^2 + y^2 \leq 1\}$ y ∂A es la circunferencia. Pero si metemos ese mismo disco dentro de $\mathbb{R}^3$ (en el plano $z = 0$), entonces $\operatorname{int} A = \varnothing$, porque ninguna bola tridimensional cabe en un objeto plano, y ∂A pasa a ser el disco cerrado entero.

Pausa para pensar

Este es el punto que más cuesta interiorizar: el interior, la clausura y la frontera *no* son propiedades del conjunto, sino de la pareja (conjunto, espacio ambiente). El mismo disco tiene frontera de dimensión 1 o de dimensión 2 según dónde lo mires. Cuando un enunciado dice «la frontera de A», siempre hay que preguntar: ¿frontera dentro de qué?

(e) Cambiando la topología: discreta, indiscreta y cofinita. Tomemos $X = \mathbb{R}$ y $A = [0, 1)$ otra vez.

Topología	$\operatorname{int} A$	$\overline{A}$	∂A
usual	$(0, 1)$	$[0, 1]$	$\{0, 1\}$
discreta (todo es abierto)	A	A	$\varnothing$
indiscreta (solo $\varnothing$ y X)	$\varnothing$	$\mathbb{R}$	$\mathbb{R}$
cofinita (cerrados = finitos y X)	$\varnothing$	$\mathbb{R}$	$\mathbb{R}$

En la discreta todo conjunto es abierto y cerrado a la vez, así que ningún conjunto tiene frontera. En la cofinita, como A es infinito y el único cerrado que lo contiene es $\mathbb{R}$, resulta

$\overline{A} = \mathbb{R}$; y como A no contiene ningún abierto no vacío (los abiertos son cofinitos y A^c es infinito), $\operatorname{int} A = \varnothing$.

(f) La unión infinita rompe el paso 6. Enumeremos $\mathbb{Q} \cap [0,1] = \{q_1, q_2, \ldots\}$ y tomemos $A_n = \{q_n\}$. Cada A_n es cerrado, luego $\overline{A_n} = A_n$ y $\bigcup_n \overline{A_n} = \mathbb{Q} \cap [0,1]$. Pero

$$\overline{\bigcup_n A_n} = \overline{\mathbb{Q} \cap [0,1]} = [0,1] \supsetneq \mathbb{Q} \cap [0,1].$$

La igualdad $\overline{A \cup B} = \overline{A} \cup \overline{B}$ vale para dos conjuntos, y por inducción para cualquier cantidad finita, pero se rompe en el paso al límite. Del mismo modo, $\operatorname{int} \bigcap_n A_n$ puede ser estrictamente menor que $\bigcap_n \operatorname{int} A_n$: con $A_n = (-\frac{1}{n}, \frac{1}{n})$, cada interior es A_n y la intersección de los interiores es $\{0\}$, mientras que el interior de la intersección es $\varnothing$.

Idea clave

Un ejercicio clásico, el «problema de Kuratowski»: partiendo de un conjunto y aplicando repetidamente clausura y complemento, ¿cuántos conjuntos distintos se pueden obtener? La respuesta es a lo sumo 14, y el máximo se alcanza en $\mathbb{R}$ con el conjunto

$$A = (0,1) \cup (1,2) \cup \{3\} \cup (\mathbb{Q} \cap [4,5]).$$

Que el número sea finito no es evidente en absoluto: se sigue de la identidad $\overline{\operatorname{int}(\overline{\operatorname{int} A})} = \overline{\operatorname{int} A}$, que dice que tras cuatro operaciones el proceso se estabiliza.

28.5. Espacios topológicos: quedarse solo con lo esencial

Aquí está el salto conceptual. Resulta que casi todo en topología —continuidad, compacidad, conexión— se puede definir usando *solo* la lista de conjuntos abiertos, sin mencionar jamás la distancia. Así que damos el paso audaz: olvidamos la distancia y declaramos, por decreto, cuáles conjuntos son abiertos.

Definición 28.21. Un *espacio topológico* es un conjunto X junto con una colección τ de subconjuntos (llamados *abiertos*) tal que:

1. El vacío $\emptyset$ y el total X son abiertos;
2. La unión de cualquier cantidad de abiertos es abierta;
3. La intersección de una cantidad *finita* de abiertos es abierta.

Idea clave

Intuición primero. Un espacio topológico es lo que queda cuando destilas la esencia de "cercanía". No dices "estos puntos están a distancia 0.3"; solo dices "estos conjuntos cuentan como abiertos" (como zonas de vecindad). Con eso basta para hablar de límites y continuidad. La condición de que solo las intersecciones *finitas* sean abiertas es importante y sutil: en $\mathbb{R}$, los intervalos $(-\frac{1}{n}, \frac{1}{n})$ son todos abiertos, pero su intersección infinita es solo el punto $\{0\}$, que *no* es abierto. Por eso la regla pide "finita".

Ejemplo 28.22 (Topologías extremas)**.** Sobre cualquier conjunto X hay dos topologías extremas. La *discreta*: *todos* los subconjuntos son abiertos (máxima separación, cada punto aislado). La *indiscreta*: solo $\emptyset$ y X son abiertos (mínima información, nada distingue a un punto de otro). Entre estos dos extremos viven todas las topologías interesantes, como la usual de $\mathbb{R}$ (donde los abiertos son las uniones de intervalos abiertos).

28.6. Continuidad, redefinida

En el capítulo 6, una función continua era una "sin saltos", definida con ε y δ. La topología da una definición equivalente pero mucho más limpia, que ni siquiera menciona distancias.

Definición 28.23. Una función $f : X \to Y$ entre espacios topológicos es *continua* si la preimagen de todo abierto es abierta: para cada abierto $V \subseteq Y$, el conjunto $f^{-1}(V) = \{x : f(x) \in V\}$ es abierto en X.

Idea clave

Qué dice esto en cristiano. La definición topológica de continuidad (*la preimagen de todo abierto es abierta*) parece no tener nada que ver con la de ε–δ, y es exactamente la misma. La traducción es directa: "para todo ε existe δ" dice que la preimagen de la bola de radio ε contiene una bola de radio δ, es decir, que la preimagen es abierta. La ventaja de la versión topológica es que no menciona distancias, y por tanto tiene sentido en espacios donde no hay ninguna.

Ejemplo 28.24 (La traducción entre los dos lenguajes)**. (a) Preimagen sí, imagen no.** La condición es sobre *preimágenes*. La función $f(x) = x^2$ es continua, y sin embargo la imagen del abierto $(-1, 1)$ es $[0, 1)$, que no es abierto. Definir continuidad con imágenes daría una noción distinta, la de aplicación abierta, y no coincide con la continuidad.

(b) Comprobar con un abierto concreto. Para $f(x) = x^2$ y el abierto $(1, 4)$, la preimagen es $(-2, -1) \cup (1, 2)$: abierto, como debe. Para $f(x) = \lfloor x \rfloor$ y el abierto $\left(-\frac{1}{2}, \frac{1}{2}\right)$, la preimagen es $[0, 1)$, que *no* es abierto; por eso la parte entera no es continua.

(c) La topología decide. La identidad de $\mathbb{R}$ con la topología discreta en $\mathbb{R}$ con la usual es continua, porque en la discreta todo conjunto es abierto. En sentido contrario no lo es. La misma fórmula puede ser continua o no según qué topologías se pongan a los lados, cosa que la definición con distancias oculta.

(d) Un caso sin distancias. En un espacio con la topología cofinita (los abiertos son los complementarios de conjuntos finitos), la continuidad sigue teniendo sentido y se comprueba igual, aunque no exista ninguna noción de "estar cerca". Esa es la ganancia de la definición topológica.

Cuidado

[Continua no es abierta ni cerrada] Las tres nociones son independientes: $f(x) = x^2$ es continua y no abierta; la proyección $\mathbb{R}^2 \to \mathbb{R}$ es continua y abierta pero no cerrada (la imagen de la hipérbola $xy = 1$, que es cerrada, es $\mathbb{R} \setminus \{0\}$); y una inclusión puede ser cerrada sin ser abierta. Segundo aviso: basta comprobar la condición sobre una *base* de abiertos del espacio de llegada, no sobre todos; en $\mathbb{R}$ basta con los intervalos, y eso es lo que hace practicable la comprobación.

Idea clave

Intuición primero. La idea es: una función continua no "desgarra" el espacio. Si miras hacia atrás (la preimagen) una zona de vecindad del destino, lo que ves en el origen también es una zona de vecindad, no un conjunto roto. Comparado con los ε–δ del capítulo 6, esta definición dice lo mismo ("puntos cercanos van a puntos cercanos") pero empaquetado en una sola frase sobre abiertos, y funciona incluso donde no hay distancias. Es uno de esos momentos en que subir de abstracción *simplifica* en vez de complicar.

Proposición 28.25. *Para funciones $f : \mathbb{R} \to \mathbb{R}$, esta definición topológica coincide con la de ε–δ del capítulo 6.*

Idea. Si f es continua con ε–δ y V es abierto, toma x con $f(x) \in V$; como V es abierto hay una bolita $B(f(x), \varepsilon) \subseteq V$, y la continuidad ε–δ da un δ con $f(B(x, \delta)) \subseteq B(f(x), \varepsilon) \subseteq V$, así que $B(x, \delta) \subseteq f^{-1}(V)$: la preimagen es abierta. El recíproco desanda el mismo camino. Las dos definiciones son dos idiomas para la misma idea. □

Idea clave

Qué dice esto en cristiano. La definición de ε–δ dice "si me acerco lo bastante en el origen, llego tan cerca como quiera en el destino". La topológica dice "la preimagen de todo abierto es abierta". Son la misma frase leída al revés: un abierto del destino es precisamente un conjunto donde cabe una bolita ε alrededor de cada punto, y pedir que su preimagen sea abierta es pedir que alrededor de cada punto de partida quepa una bolita δ. La versión topológica gana porque no menciona distancias, y por eso sigue teniendo sentido en espacios donde no hay manera de medir.

Ejemplo 28.26 (Comprobar continuidad con preimágenes)**. (a) La parábola.** Para $f(x) = x^2$, la preimagen de $(1, 4)$ es $(-2, -1) \cup (1, 2)$, unión de dos abiertos y por tanto abierta; la de $(-1, 4)$ es $(-2, 2)$, también abierta. En general $f^{-1}((a, b))$ con $0 \le a < b$ es $(-\sqrt{b}, -\sqrt{a}) \cup (\sqrt{a}, \sqrt{b})$, siempre abierto: f es continua.

(b) La función escalón. Sea $H(x) = 0$ si $x < 0$ y $H(x) = 1$ si $x \ge 0$. La preimagen del abierto $(0{,}5;\ 1{,}5)$ es $[0, \infty)$, que *no* es abierto: H no es continua, y el punto de fallo es exactamente $x = 0$, donde el intervalo cerrado empieza.

(c) Cuidado con las imágenes. La condición es sobre preimágenes, no sobre imágenes. Para $f(x) = x^2$ la imagen del abierto $(-1, 1)$ es $[0, 1)$, que no es abierto. Una función continua puede perfectamente mandar abiertos en no abiertos; lo que no puede es que la preimagen de un abierto falle.

(d) Basta con una subbase. No hace falta comprobar todos los abiertos: basta con los de una subbase. En $\mathbb{R}$ los intervalos (a, ∞) y $(-\infty, b)$ generan la topología, así que f es continua si y sólo si $\{x : f(x) > a\}$ y $\{x : f(x) < b\}$ son abiertos para todos a, b. Con $f(x) = x^2$ eso da $\{|x| > \sqrt{a}\}$ y $\{|x| < \sqrt{b}\}$, abiertos ambos, y el trabajo se acaba en dos líneas.

28.7. Taller: comprobar la continuidad en la práctica

Regla: demostrar que una función es continua

La definición topológica —la preimagen de todo abierto es abierta— parece pedir que se examinen infinitos conjuntos. En la práctica nunca se hace eso: se usan bases, teoremas de construcción y el lema del pegado. La definición se reserva para los contraejemplos.

1. **Comprueba solo sobre una base o una subbase.** Si $\mathcal{B}$ es base de la topología del destino, basta que $f^{-1}(B)$ sea abierto para todo $B \in \mathcal{B}$; el resto sale gratis, porque la preimagen respeta uniones. Para $f : X \to \mathbb{R}$ basta mirar los intervalos (a, b), e incluso solo las semirrectas (a, ∞) y $(-\infty, b)$, que forman subbase.

2. **Construye a partir de continuas conocidas.** La composición de continuas es continua; sumas, productos y cocientes (con denominador no nulo) de funciones reales continuas son continuos; las proyecciones $\pi_X : X \times Y \to X$ son continuas; toda función constante lo es, y también la identidad y las inclusiones de subespacios.

3. **Hacia un producto, comprueba coordenada a coordenada.** Una aplicación $f : Z \to X \times Y$ es continua si y solo si lo son sus dos componentes $\pi_X \circ f$ y $\pi_Y \circ f$. Esta es la propiedad que *define* la topología producto, y convierte casi todo cálculo de curvas y superficies en dos comprobaciones separadas.

4. **Usa que la continuidad es local.** Si cada punto tiene un entorno abierto en el que f es continua, entonces f es continua. Consecuencia: si $X = \bigcup U_i$ con los U_i *abiertos* y f restringida a cada U_i es continua, f lo es.

5. **Para funciones definidas a trozos, usa el lema del pegado.** Si $X = A \cup B$ con A y B *ambos cerrados* (o ambos abiertos), $f|_A$ y $f|_B$ continuas y *coincidentes en* $A \cap B$, entonces f es continua. Las tres hipótesis son necesarias, y la que más se olvida es la última.

6. **Criterios equivalentes que a veces salen más baratos.** f es continua si y solo si la preimagen de todo *cerrado* es cerrada; si y solo si $f(\overline{A}) \subseteq \overline{f(A)}$ para todo $A \subseteq X$. En espacios métricos, además, si y solo si $x_n \to x$ implica $f(x_n) \to f(x)$.

7. **Para probar que NO es continua, exhibe un abierto testigo.** Basta *un* abierto V del destino cuya preimagen no sea abierta, o —en espacios métricos—

una sucesión $x_n \to x$ con $f(x_n) \not\to f(x)$. Un solo testigo concreto vale más que cualquier argumento general.

Ejemplo 28.27 (Siete funciones interrogadas)**.** Aplicamos los pasos por orden de comodidad, no de aparición.

(a) $f(x) = x^2$ es continua, comprobado sobre una base. Por el paso 1 basta ver que $f^{-1}((a,b))$ es abierto. Si $b \le 0$ la preimagen es vacía; si $a < 0 \le b$ es $(-\sqrt{b}, \sqrt{b})$; si $0 \le a < b$ es $(-\sqrt{b}, -\sqrt{a}) \cup (\sqrt{a}, \sqrt{b})$. Los tres casos son abiertos. Por ejemplo, $f^{-1}((1,4)) = (-2,-1) \cup (1,2)$. Nótese que la preimagen de un intervalo *no* tiene por qué ser un intervalo: continua no significa que conserve la forma de los abiertos, solo que la preimagen sea abierta.

(b) La curva $\gamma(t) = (\cos t, \sin t)$ es continua, coordenada a coordenada. Por el paso 3, como $t \mapsto \cos t$ y $t \mapsto \sin t$ son continuas de $\mathbb{R}$ en $\mathbb{R}$, la aplicación $\gamma : \mathbb{R} \to \mathbb{R}^2$ es continua. No hace falta tocar ningún abierto del plano. El mismo argumento vale para cualquier curva o superficie parametrizada.

(c) El valor absoluto, pegado correctamente. Sea $f(x) = -x$ en $A = (-\infty, 0]$ y $f(x) = x$ en $B = [0, \infty)$. Ambos trozos son cerrados, ambas restricciones son continuas, y en $A \cap B = \{0\}$ las dos fórmulas dan 0. El lema del pegado (paso 5) concluye que $f = |x|$ es continua, sin ningún ε ni ningún δ.

(d) El mismo pegado, mal hecho. Sea ahora $g(x) = x$ en $(-\infty, 0)$ y $g(x) = x + 1$ en $[0, \infty)$. Las restricciones son continuas y los trozos cubren $\mathbb{R}$, pero *no* se cumplen las hipótesis del paso 5: uno de los trozos es abierto y el otro cerrado. Y en efecto g no es continua. El testigo del paso 7 es la semirrecta $V = (\frac{1}{2}, \infty)$, que es abierta y cumple

$$g^{-1}(V) = [0, \infty),$$

porque ningún $x < 0$ tiene imagen mayor que $\frac{1}{2}$ y todo $x \ge 0$ la tiene. Preimagen de un abierto que no es abierta: no es continua.

(e) La función de Dirichlet no es continua en ningún punto. Sea $\chi(x) = 1$ si $x \in \mathbb{Q}$ y $\chi(x) = 0$ si no. El abierto $V = (\frac{1}{2}, \frac{3}{2})$ tiene preimagen $\chi^{-1}(V) = \mathbb{Q}$, que no es abierto (ningún intervalo cabe en $\mathbb{Q}$). Un solo testigo basta. Con sucesiones se ve igual de rápido: si x es racional, tómese $x_n \to x$ con x_n irracionales; entonces $\chi(x_n) = 0 \not\to 1 = \chi(x)$.

(f) La continuidad depende de las dos topologías, no de la fórmula. Considérese la identidad $\mathrm{id} : \mathbb{R} \to \mathbb{R}$. De la topología discreta a la usual es continua (la preimagen de cualquier cosa es abierta, porque *todo* es abierto en el origen). De la usual a la discreta *no* lo es: $\{0\}$ es abierto en el destino y su preimagen $\{0\}$ no es abierta en el origen. La misma fórmula, dos veredictos opuestos. Cuando alguien pregunta «¿es continua?», la pregunta está incompleta hasta que se dicen las dos topologías.

(g) $f(x) = 1/x$ es continua, y no hay ninguna discontinuidad en 0. Su dominio es $\mathbb{R} \setminus \{0\}$, que es abierto; en él, f es cociente de continuas con denominador no nulo, luego es continua por el paso 2. La costumbre escolar de decir «tiene una discontinuidad en $x = 0$» es un abuso de lenguaje: en 0 la función no está definida, y una función no puede ser discontinua donde no existe.

Cuidado

La continuidad se comprueba con *preimágenes*, nunca con imágenes. La imagen de un abierto por una función continua puede no ser abierta: $f(x) = x^2$ manda el abierto $(-1, 1)$ en $[0, 1)$, que no lo es; y $f(x) = 0$ manda cualquier abierto en el punto $\{0\}$. Las funciones que sí mandan abiertos en abiertos se llaman *aplicaciones abiertas*, y son otra cosa: un homeomorfismo es exactamente una biyección continua y abierta.

Pausa para pensar

Vuelve al paso 5 y pregúntate por qué el lema del pegado exige que los dos trozos sean cerrados. La razón está en el paso 6: la preimagen de un cerrado por f es $(f|_A)^{-1}(C) \cup (f|_B)^{-1}(C)$, y cada uno de esos dos conjuntos es cerrado *en su trozo*; para que la unión sea cerrada en X hace falta que los trozos lo sean. Con infinitos trozos cerrados el lema falla: cada punto $\{q\}$ de $\mathbb{Q}$ es cerrado y la función de Dirichlet es constante en cada uno, y ya hemos visto que no es continua.

Idea clave

Los pasos 1 y 3 dicen, entre líneas, algo más profundo: una topología queda determinada por qué funciones son continuas hacia ella o desde ella. La topología producto es la más gruesa que hace continuas a las proyecciones; la cociente, la más fina que hace continua a la proyección al cociente; la inicial y la final generalizan ambas. Esta forma de definir objetos por su comportamiento frente a aplicaciones, y no por sus elementos, es la manera moderna de hacer matemáticas, y reaparece en el capítulo 24 al caracterizar el anillo cociente por su propiedad universal.

28.8. Compacidad

La compacidad es quizá la propiedad más importante de toda la topología. En el capítulo 6 la vimos como "cerrado y acotado"; aquí está su forma verdadera.

Definición 28.28. Un *cubrimiento abierto* de X es una colección de abiertos cuya unión es todo X. El espacio X es *compacto* si de *todo* cubrimiento abierto se puede extraer un *subcubrimiento finito* (una cantidad finita de esos abiertos que ya cubren todo).

Idea clave

Intuición primero. Compacto significa, más o menos, "finito para efectos prácticos, aunque tenga infinitos puntos". Imagina que cubres un país con infinitas manchas de luz de linternas; si el país es compacto, con un número *finito* de esas linternas ya lo iluminas entero. Un intervalo cerrado $[0, 1]$ es compacto: por más ingeniosamente que lo cubras con abiertos, siempre bastan unos pocos. Un intervalo abierto $(0, 1)$ *no* lo es: se puede cubrir de forma que necesites infinitos y ninguna subcolección finita alcance

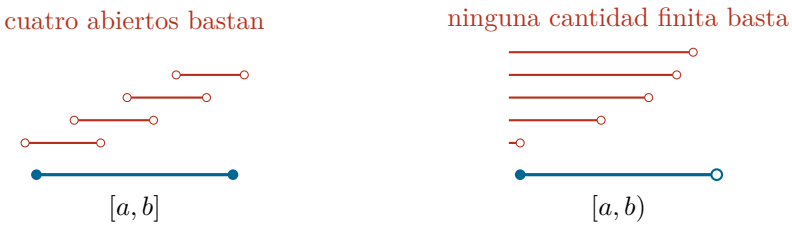


Figura 28.1: Compacidad por recubrimientos. A la izquierda, un intervalo cerrado y acotado: de cualquier familia de abiertos que lo cubra se puede extraer una subfamilia finita. A la derecha, el mismo intervalo sin el extremo derecho: los abiertos se acercan al punto que falta sin llegar nunca, y por muchos que se tomen queda siempre un trozo sin cubrir.

(lo veremos en el taller). La compacidad es lo que hace que las funciones continuas alcancen su máximo, que las sucesiones tengan subsucesiones convergentes, y que muchas demostraciones "de infinito" se reduzcan a "finito".

Teorema 28.29 (Heine–Borel)**.** *Un subconjunto de* $\mathbb{R}^n$ *es compacto si y solo si es* cerrado y acotado*.*

Idea clave

Qué dice esto en cristiano. Compacto significa que de cualquier manera de cubrir el espacio con abiertos se puede escoger una cantidad finita que ya lo cubra. Es una forma de decir "se comporta como si fuera finito": todo lo que se pueda hacer con finitos abiertos vale. Heine–Borel traduce eso en $\mathbb{R}^n$ a algo comprobable de un vistazo, cerrado y acotado, y el teorema del máximo es la razón por la que a nadie le importaría la compacidad si no fuera por ella: una función continua sobre un compacto alcanza sus extremos, y eso es la base de toda la optimización.

Ejemplo 28.30 (Compactos, no compactos y por qué)**. (a) Un cubrimiento sin subcubrimiento finito.** El intervalo $(0,1)$ está cubierto por los abiertos $U_n = \left(\frac{1}{n}, 1\right)$ para $n \geq 2$. Cualquier subcolección finita tiene un n máximo y deja fuera todo $\left(0, \frac{1}{n}\right]$. Luego $(0,1)$ no es compacto: le falta ser cerrado.

(b) Otro que falla por no acotado. La recta $\mathbb{R}$ está cubierta por los intervalos $(-n, n)$, y ninguna subcolección finita la cubre. Es cerrada y no acotada: falla la otra mitad de Heine–Borel.

(c) El teorema del máximo, con y sin hipótesis. Sobre $[0,1]$, la función $f(x) = x$ alcanza máximo 1 y mínimo 0. Sobre $(0,1)$ no alcanza ninguno de los dos: el supremo es 1 y el ínfimo 0, y ninguno pertenece. Y sobre $(0,1]$, la función $\frac{1}{x}$ ni siquiera está acotada: vale 10, 100 y 1000 en 0,1, 0,01 y 0,001.

(d) Dónde Heine–Borel deja de valer. La equivalencia "compacto = cerrado y acotado" es exclusiva de $\mathbb{R}^n$. En un espacio de dimensión infinita falla: la bola unidad cerrada de ℓ^2 es cerrada y acotada y *no* es compacta, porque la sucesión de vectores de la base canónica está

en ella y no tiene ninguna subsucesión convergente (dos de ellos distan siempre $\sqrt{2}$).

Cuidado

[Compacto es una propiedad del espacio, no del ambiente] Un conjunto es compacto en sí mismo, con la topología inducida, y eso no cambia al meterlo en otro espacio. En cambio "cerrado" sí depende del ambiente: $(0,1)$ es cerrado dentro de sí mismo y no lo es dentro de $\mathbb{R}$. Segundo aviso: en espacios métricos, compacto equivale a que toda sucesión tenga subsucesión convergente (compacidad secuencial), pero en topología general esas dos nociones *no* coinciden. Y tercero: la imagen continua de un compacto es siempre compacta, y de ahí sale el teorema del máximo en una línea: la imagen es un compacto de $\mathbb{R}$, luego cerrado y acotado, luego contiene a su supremo.

Este teorema (demostrado en el capítulo 6 para $\mathbb{R}$) es la razón de que en cálculo baste con "cerrado y acotado": en $\mathbb{R}^n$ eso equivale a la compacidad de verdad. Pero ojo: la equivalencia es especial de $\mathbb{R}^n$; en espacios más generales, compacto es estrictamente más fuerte que cerrado y acotado.

Teorema 28.31 (Por qué importa: el máximo se alcanza)**.** *Si X es compacto y $f : X \to \mathbb{R}$ es continua, entonces f alcanza su valor máximo y su mínimo.*

Idea. La imagen $f(X)$ es compacta (la continuidad manda compactos a compactos: un cubrimiento de la imagen se tira hacia atrás a uno del dominio, se toma el subcubrimiento finito y se empuja de vuelta). Un compacto de $\mathbb{R}$ es cerrado y acotado por Heine–Borel, así que tiene un supremo *que le pertenece*: ese es el máximo. Este es exactamente el teorema del valor extremo del capítulo 6, ahora entendido en su raíz: *el máximo existe porque el dominio es compacto.* □

Ejemplo 28.32 (Dónde falla el teorema si se quita una hipótesis)**. (a) Con compacidad, todo bien.** En $X = [0,3]$ la función $f(x) = x^3 - 3x$ es continua; de sus dos puntos críticos $x = \pm 1$ sólo $x = 1$ cae dentro del intervalo, y los valores candidatos son $f(0) = 0$, $f(1) = -2$ y $f(3) = 18$. El mínimo es -2 en $x = 1$ y el máximo 18 en $x = 3$; los dos se alcanzan, como promete el teorema.

(b) Sin cerrado. En $X = (0,1)$, la función $f(x) = x$ es continua y acotada, su supremo es 1 y su ínfimo 0, y *ninguno* se alcanza. El intervalo abierto no es compacto porque el recubrimiento $\{(1/n, 1-1/n)\}_{n\geq 3}$ no tiene subrecubrimiento finito.

(c) Sin acotado. En $X = [0,\infty)$, la función $f(x) = x$ no alcanza máximo porque ni siquiera está acotada. En $X = \mathbb{R}$, la función $f(x) = \arctan x$ sí está acotada, entre $-\pi/2$ y $\pi/2$, y tampoco alcanza ninguno de los dos extremos.

(d) Sin continuidad. En $X = [0,1]$, que sí es compacto, sea $f(x) = x$ para $x < 1$ y $f(1) = 0$. El supremo es 1 y no se alcanza. El teorema necesita las dos hipótesis: la compacidad del dominio y la continuidad de la función. Ninguna sobra.

Cuidado

[Contraejemplo: en dimensión infinita cerrado y acotado no basta] En $\mathbb{R}^n$, compacto equivale a cerrado y acotado (Heine–Borel), y de ahí viene la costumbre de tratarlos como sinónimos. En el espacio ℓ^2 de sucesiones de cuadrado sumable, la bola unidad cerrada es cerrada y acotada y *no* es compacta: la sucesión de vectores $e_1, e_2, e_3, \ldots$ de la base canónica vive en ella y cumple $\|e_i - e_j\| = \sqrt{2}$ para $i \neq j$, de modo que no tiene ninguna subsucesión de Cauchy. Sobre esa bola, el funcional continuo $f(x) = \sum_n x_n/n$ tiene supremo que no se alcanza.

Regla: demostrar que un espacio es compacto, o que no lo es

Compacidad es la propiedad más útil y la peor entendida del tema, porque su definición (todo recubrimiento abierto admite subrecubrimiento finito) parece diseñada para no poder verificarse. La clave es no usar nunca la definición para lo positivo y usarla siempre para lo negativo.

1. **Para demostrar que SÍ es compacto, no uses la definición.** Usa teoremas. En $\mathbb{R}^n$, Heine–Borel: compacto $\iff$ cerrado y acotado. En general: la imagen continua de un compacto es compacta, un producto de compactos es compacto (Tychonoff), y un subconjunto cerrado de un compacto es compacto. Con estas cuatro herramientas se resuelve casi todo.

2. **Para demostrar que NO es compacto, exhibe un recubrimiento rebelde.** Basta *uno* sin subrecubrimiento finito. Los dos modelos que sirven casi siempre: para conjuntos no acotados, $\{B(0,n)\}_{n\geq 1}$; para conjuntos a los que les falta un punto de su borde, bolas que se acercan al punto que falta sin llegar.

3. **Alternativa para lo negativo: una sucesión sin subsucesión convergente.** En espacios métricos, compacto equivale a secuencialmente compacto, y a menudo es más rápido exhibir la sucesión. Los modelos: $x_n = n$ (se escapa al infinito) y $x_n = 1/n$ en $(0,1]$ (converge fuera del conjunto).

4. **Fuera de $\mathbb{R}^n$, olvídate de Heine–Borel.** "Cerrado y acotado" *no* implica compacto en dimensión infinita. La bola unidad cerrada de un espacio normado es compacta si y solo si la dimensión es finita (teorema de Riesz). Aplicar Heine–Borel en ℓ^2 o en $C[0,1]$ es el error más caro del capítulo.

5. **Aprovecha lo que da la compacidad.** Una función continua sobre un compacto es acotada, alcanza su máximo y su mínimo, y es uniformemente continua. En cuanto veas "compacto" en una hipótesis, estas tres consecuencias están disponibles gratis.

6. **Úsala también como invariante.** Si X es compacto e Y no lo es, no hay homeomorfismo entre ellos, porque la imagen continua de un compacto es compacta. Es la forma más barata de distinguir espacios.

Ejemplo 28.33 (Cuatro veredictos de compacidad)**. (a)** $[0,1]$ **es compacto;** $(0,1)$ **no.** El primero, por Heine–Borel: cerrado y acotado. El segundo, con el recubrimiento rebelde

$$\mathcal{U} = \left\{\left(\tfrac{1}{n}, 1\right)\right\}_{n\geq 2}.$$

Cubre $(0,1)$, porque todo $x \in (0,1)$ cumple $x > 1/n$ para n grande. Pero cualquier subfamilia finita tiene un n máximo, digamos N, y su unión es $(\frac{1}{N}, 1)$, que no contiene al punto $\frac{1}{2N}$. No hay subrecubrimiento finito. Alternativamente, la sucesión $x_n = 1/n$ vive en $(0,1)$ y su único límite posible es 0, que no está: ninguna subsucesión converge *dentro* del espacio.

(b) La esfera S^2 es compacta; el plano no. $S^2 = f^{-1}(\{1\})$ con $f(x,y,z) = x^2+y^2+z^2$ continua, luego es cerrada; y está acotada. Heine–Borel. El plano $\mathbb{R}^2$ no lo está: el recubrimiento $\{B(0,n)\}_{n\geq 1}$ no admite subrecubrimiento finito, porque una subfamilia finita se queda dentro de la bola de radio máximo.

(c) La bola unidad cerrada de ℓ^2 NO es compacta. Es cerrada y acotada, y sin embargo falla: considérese la sucesión de vectores de la base $e_1, e_2, e_3, \ldots$, todos de norma 1. Para $n \neq m$,

$$\|e_n - e_m\|_2 = \sqrt{1+1} = \sqrt{2},$$

así que *ninguna* pareja se acerca nunca: ninguna subsucesión es de Cauchy y por tanto ninguna converge. Este único ejemplo desmonta Heine–Borel fuera de dimensión finita, y es exactamente por eso que en análisis funcional hubo que inventar la convergencia débil: para recuperar algún tipo de compacidad.

(d) $\mathbb{R}$ **y** $[0,1]$ **no son homeomorfos.** Uno es compacto y el otro no; por el paso 6, ya está. (Compárese con el argumento alternativo de conexión: quitando un punto interior, $[0,1]$ puede quedar conexo — quitando un extremo — pero $\mathbb{R}$ nunca. Ambos funcionan; el de compacidad es más corto.)

Pausa para pensar

Hay una forma de recordar por qué la compacidad es tan potente: es la versión topológica de "finito". Todo lo que en un conjunto finito es evidente — que una función alcanza su máximo, que hay un elemento más cercano, que las cosas no se escapan — sigue siendo cierto en un compacto. La definición por recubrimientos es fea, pero lo que hace es precisamente comprar esa finitud sin exigir que el conjunto sea finito.

28.9. Conexión

Definición 28.34. Un espacio es *conexo* si no se puede partir en dos abiertos no vacíos y disjuntos; es decir, si no está hecho de "dos pedazos separados".

Ejemplo 28.35 (Conexos y no conexos)**. (a) El intervalo.** $[0,1]$ es conexo. La demostración clásica supone $[0,1] = A \cup B$ con A, B abiertos disjuntos no vacíos, toma $a \in A$, $b \in B$ con $a < b$ y considera $c = \sup\{x \in A : x < b\}$: ese c no puede estar ni en A ni en B sin contradicción, y sin embargo está en $[0,1]$.

(b) Dos trozos. $[0,1]\cup[2,3]$ no es conexo: los conjuntos $(-1;\ 1{,}5)$ y $(1{,}5;\ 4)$ intersecados con el espacio dan dos abiertos disjuntos que lo recubren. Tampoco lo es $\mathbb{Q}$ con la topología usual, y aquí la partición es más sorprendente: $\{x\in\mathbb{Q} : x<\sqrt{2}\}$ y $\{x\in\mathbb{Q} : x>\sqrt{2}\}$ son abiertos, disjuntos y recubren $\mathbb{Q}$, porque el punto que los separaría no es racional. De hecho las componentes conexas de $\mathbb{Q}$ son los puntos sueltos.

(c) Conexo pero no por caminos. La *curva del topólogo*, unión de la gráfica de $\sin(1/x)$ para $x\in(0,1]$ con el segmento $\{0\}\times[-1,1]$, es conexa (la adherencia de un conexo es conexa) y no es conexa por caminos: ningún camino continuo llega del segmento a la gráfica, porque tendría que oscilar infinitas veces en tiempo finito.

(d) Por qué importa. El teorema del valor intermedio es exactamente el enunciado "la imagen continua de un conexo es conexa" más "los conexos de $\mathbb{R}$ son los intervalos". Con ese par de frases se demuestra en dos renglones lo que en cálculo cuesta una página de biseccciones.

Idea clave

Intuición primero. Conexo significa "de una sola pieza". La recta $\mathbb{R}$ es conexa; el conjunto $(0,1)\cup(2,3)$ no lo es (son dos trozos con un hueco en medio). La conexión es lo que hay detrás del *teorema del valor intermedio* del capítulo 6: si una función continua vale -1 en un punto y $+1$ en otro, tiene que valer 0 en medio, porque no puede "saltarse" el cero sin romper la conexión de su recorrido.

Teorema 28.36. *La imagen continua de un conexo es conexa. En particular, si $f:[a,b]\to\mathbb{R}$ es continua, su imagen es un intervalo: eso es el teorema del valor intermedio.*

Regla: demostrar que un espacio es conexo, o que no lo es

La conexión es la formalización de «está hecho de una sola pieza». Las dos direcciones de la demostración son completamente distintas: una se hace exhibiendo algo, la otra invocando teoremas.

1. **Escribe la definición correcta.** X es *disconexo* si existen abiertos U,V no vacíos, disjuntos, con $U\cup V=X$. Es *conexo* si no existe tal partición. Fíjate en que U y V son entonces también cerrados: los conjuntos que son abiertos y cerrados a la vez son los que detectan la disconexión.

2. **Para probar que NO es conexo, exhibe la partición.** Basta encontrar un conjunto no vacío, distinto de X, que sea abierto y cerrado. Alternativamente, construye una función continua y sobreyectiva $f:X\to\{0,1\}$ con $\{0,1\}$ discreto: eso *es* una partición disfrazada.

3. **Para probar que SÍ es conexo, usa la caja de herramientas.** Los ladrillos: (i) los intervalos de $\mathbb{R}$ son conexos; (ii) la imagen continua de un conexo es conexa; (iii) si $\{A_i\}$ son conexos con un punto en común, $\bigcup A_i$ es conexo; (iv) si A es conexo y $A\subseteq B\subseteq\overline{A}$, entonces B es conexo.

4. **El atajo habitual: conexión por caminos.** Si dos puntos cualesquiera se unen por un camino continuo, el espacio es conexo por (i) y (iii). Conexo por caminos $\Rightarrow$ conexo, siempre. El recíproco es falso en general, pero *sí* vale para abiertos de $\mathbb{R}^n$.

5. **Descompón en componentes.** La componente conexa de x es la unión de todos los conexos que contienen a x; es el mayor conexo que lo contiene, y es cerrada. Las componentes parten el espacio, y su número es un invariante topológico.

6. **Para distinguir dos espacios, quita puntos.** Si $X \cong Y$ mediante h, entonces $X \setminus \{p\} \cong Y \setminus \{h(p)\}$. Contando componentes tras quitar un punto se distinguen espacios que a simple vista se parecen.

7. **Verifica con el teorema del valor intermedio.** Si has decidido que X es conexo, toda $f : X \to \mathbb{R}$ continua debe tomar todos los valores intermedios. Si encuentras una que salta, tu conclusión era falsa.

Ejemplo 28.37 (Seis espacios interrogados, y una curva conexa por la que no se puede caminar)**.** Aplicamos los dos procedimientos alternándolos.

(a) $\mathbb{R} \setminus \{0\}$ frente a $\mathbb{R}^2 \setminus \{0\}$. El primero se parte en $U = (-\infty, 0)$ y $V = (0, \infty)$, ambos abiertos, no vacíos y disjuntos: es disconexo, con dos componentes. El segundo es conexo por caminos: dados dos puntos no nulos del plano, si el segmento que los une no pasa por el origen ya está; y si pasa, se rodea el origen con un arco. Luego es conexo.

Consecuencia inmediata por el paso 6: $\mathbb{R}$ *no es homeomorfo a* $\mathbb{R}^2$. Si lo fuera, quitar el punto correspondiente al origen produciría espacios homeomorfos, y uno tiene dos componentes y el otro una. Este es el argumento de dimensión más barato que existe, y funciona solo para $n = 1$ frente a $n \geq 2$; distinguir $\mathbb{R}^2$ de $\mathbb{R}^3$ requiere ya el grupo fundamental.

(b) $\mathbb{Q}$ es totalmente disconexo. Dados dos racionales $p < q$, elige un irracional α entre ellos (por ejemplo un múltiplo adecuado de $\sqrt{2} \approx 1{,}41421$). Entonces

$$U = \mathbb{Q} \cap (-\infty, \alpha), \qquad V = \mathbb{Q} \cap (\alpha, \infty)$$

son abiertos en $\mathbb{Q}$, disjuntos, no vacíos y cubren $\mathbb{Q}$ porque $\alpha \notin \mathbb{Q}$. Como esto se puede hacer alrededor de cualquier punto, las componentes conexas de $\mathbb{Q}$ son los puntos sueltos. Y sin embargo $\mathbb{Q}$ no es discreto: ningún punto es abierto. «Totalmente disconexo» y «discreto» no son lo mismo.

(c) Las matrices invertibles reales, y las complejas. El determinante $\det : \mathrm{GL}_n(\mathbb{R}) \to \mathbb{R} \setminus \{0\}$ es continuo y sobreyectivo. Componiendo con el signo se obtiene una función continua sobre $\{0, 1\}$, luego $\mathrm{GL}_n(\mathbb{R})$ es disconexo: se parte en $\det > 0$ y $\det < 0$. De hecho tiene exactamente dos componentes (cada una es conexa por caminos, vía descomposición en operaciones elementales). En cambio $\mathrm{GL}_n(\mathbb{C})$ *sí* es conexo, porque $\mathbb{C} \setminus \{0\}$ lo es y ya no hay ningún signo que separar. Véase el capítulo 7 para el determinante y el capítulo 4 para $\mathbb{C} \setminus \{0\}$.

(d) La curva seno del topólogo: conexa sin ser conexa por caminos. Sea

$$S = \left\{ \left(x, \sin \tfrac{1}{x}\right) : 0 < x \leq 1 \right\} \cup \left(\{0\} \times [-1, 1]\right).$$

Llamemos A a la primera parte. A es la imagen continua del intervalo $(0, 1]$, luego es conexa por el paso 3(ii). Su clausura es exactamente S, porque al acercarse a $x = 0$ el seno oscila y sus valores se acumulan en todo $[-1, 1]$. Por el paso 3(iv), S es conexa.

Pero no hay ningún camino continuo de $(0, 0)$ a un punto de A: un camino así tendría que recorrer infinitas oscilaciones completas en un tiempo finito, y su segunda coordenada no podría ser continua en el instante en que llega a $x = 0$. Formalmente, el conjunto de tiempos en que el camino está sobre el eje $\{x = 0\}$ es cerrado y también abierto, lo que fuerza al camino a quedarse ahí.

Cuidado

Aquí muere la tentación de usar «conexo» y «conexo por caminos» como sinónimos. La implicación va en un solo sentido. Lo que sí es cierto: en un *abierto* de $\mathbb{R}^n$ ambas nociones coinciden, y por eso en cálculo de varias variables (capítulo 9) uno puede permitirse la confusión sin consecuencias.

(e) El intervalo y la circunferencia. $[0, 1]$ y S^1 son ambos compactos y conexos, así que esas dos propiedades no los distinguen. Pero quita un punto: $[0, 1] \setminus \{\frac{1}{2}\}$ tiene dos componentes, mientras que S^1 menos cualquier punto es homeomorfo a un intervalo abierto, y por tanto conexo. Luego *no* son homeomorfos. El mismo truco separa la letra Y de la letra I: quitando el punto de ramificación de la Y quedan tres componentes, y ningún punto de la I hace eso.

(f) Un cerrado que es conexo y cuyo interior no lo es. En $\mathbb{R}^2$, sea B la unión de los dos discos cerrados de radio 1 centrados en $(\pm 1, 0)$. Se tocan en el origen, luego B es conexo por el paso 3(iii). Pero $\operatorname{int} B$ es la unión de los dos discos *abiertos*, que son disjuntos: disconexo. La conexión no pasa al interior. (Sí pasa a la clausura, por el paso 3(iv).)

Pausa para pensar

Vuelve al paso 1 y fíjate en la reformulación: X es conexo si y solo si sus únicos subconjuntos abiertos y cerrados a la vez son $\varnothing$ y X. Esta versión es la que se generaliza mejor: en teoría de anillos, un anillo no tiene idempotentes no triviales exactamente cuando su espectro es conexo (capítulo 24). El diccionario entre álgebra y geometría empieza justo aquí.

Idea clave

El teorema del valor intermedio, que en el capítulo las funciones elementales parecía un hecho sobre gráficas, es en realidad el enunciado «la imagen continua de un conexo es conexa» aplicado a un intervalo. Toda la fuerza del teorema está en un solo hecho: los intervalos de $\mathbb{R}$ son conexos, y son los únicos subconjuntos conexos de $\mathbb{R}$.

28.10. Homeomorfismos: cuándo dos espacios son "el mismo"

Definición 28.38. Un *homeomorfismo* es una función $f : X \to Y$ que es biyectiva (empareja los puntos uno a uno), continua, y con inversa también continua. Si existe uno, se dice que X e Y son *homeomorfos*: topológicamente idénticos.

Idea clave

Intuición primero. Un homeomorfismo es la "deformación de goma" hecha precisa: una correspondencia que no rompe ni pega (continua en ambos sentidos). Que dos espacios sean homeomorfos significa que, para la topología, son indistinguibles. El intervalo $(0,1)$ es homeomorfo a toda la recta $\mathbb{R}$ (¡un trozo finito es topológicamente igual que la recta infinita!). La taza y la dona son homeomorfas. La esfera y la dona *no* lo son, y demostrarlo requiere un "invariante" que las distinga: el número de agujeros.

Ejemplo 28.39 (Un trozo finito igual que toda la recta)**.** La función $f:(-1,1)\to\mathbb{R}$ dada por $f(x)=\tan(\frac{\pi}{2}x)$ es un homeomorfismo: es continua, biyectiva (estira el intervalo hasta cubrir toda la recta), y su inversa $\frac{2}{\pi}\arctan$ también es continua. Por eso $(-1,1)$ y $\mathbb{R}$ son "el mismo espacio" topológicamente, aunque uno mida 2 y el otro sea infinito: la topología no ve longitudes.

Observación 28.40 (Invariantes topológicos)*.* ¿Cómo se demuestra que dos espacios *no* son homeomorfos? Con *invariantes*: propiedades que se conservan bajo homeomorfismo. Si una figura tiene una y la otra no, no pueden ser la misma. El número de "agujeros" es uno (formalizado por el *grupo fundamental* y los grupos de homología); la compacidad y la conexión son otros. Por ejemplo, $[0,1]$ no es homeomorfo a $(0,1)$ porque el primero es compacto y el segundo no. Cazar invariantes que distingan espacios es el corazón de la *topología algebraica*, una de las grandes ramas de la matemática del siglo XX.

Regla: demostrar que dos espacios NO son homeomorfos

Exhibir un homeomorfismo es cuestión de ingenio; demostrar que no existe requiere método. El método es siempre el mismo: encontrar una propiedad que se conserve por homeomorfismos y que uno tenga y el otro no. A esas propiedades se las llama *invariantes topológicos*.

1. **Prueba los invariantes baratos primero.** Cardinal, compacidad, conexión, conexión por caminos, número de componentes conexas, separabilidad, metrizabilidad. Uno de estos resuelve la mayoría de los casos, y cada uno se comprueba en dos líneas.

2. **Si eso no basta, quita puntos.** Este es el truco central del capítulo. Si $h: X\to Y$ es homeomorfismo y $p\in X$, entonces h restringido da un homeomorfismo $X\setminus\{p\}\to Y\setminus\{h(p)\}$. Así que si al quitar *cualquier* punto de X el resultado tiene una propiedad que Y no puede lograr quitando *ningún* punto, no hay homeomorfismo.

3. **Cuenta puntos de corte.** Un punto p es *de corte* si $X\setminus\{p\}$ es disconexo. El número de puntos de corte, y el número de componentes en que cada uno parte el espacio, son invariantes. Con esto se distinguen todas las letras del alfabeto vistas como espacios topológicos.

4. **Si el espacio es de dimensión ≥ 2, quitar un punto no basta: quita curvas.** En el plano, quitar un punto deja algo conexo. Lo que distingue $\mathbb{R}^2$ de $\mathbb{R}$ no es

eso, sino que $\mathbb{R} \setminus \{p\}$ es disconexo mientras que $\mathbb{R}^2 \setminus \{p\}$ no lo es — es decir, el argumento se usa al revés.

5. **Cuando todo lo anterior falla, sube al grupo fundamental.** π_1 es un invariante mucho más fino. $\pi_1(\mathbb{R}^2) = \{1\}$ pero $\pi_1(S^1) = \mathbb{Z}$: eso demuestra que el disco y la circunferencia no son homeomorfos, y de paso que la circunferencia no es contráctil.

6. **Recuerda que un homeomorfismo es más que una biyección continua.** Hace falta que la inversa también sea continua. El ejemplo obligatorio es $h : [0, 2\pi) \to S^1$, $h(t) = (\cos t, \sin t)$: es biyectiva y continua, pero su inversa no lo es (los puntos cercanos a $(1, 0)$ por debajo del eje se aplican en valores cercanos a 2π, lejos de 0). Y en efecto no puede haber homeomorfismo, porque S^1 es compacta y $[0, 2\pi)$ no.

Ejemplo 28.41 (Distinguir espacios con invariantes: cuatro casos)**. (a) El intervalo** $(0, 1)$ **y la circunferencia** S^1**.** Compacidad ya los separa (S^1 lo es, $(0, 1)$ no). Pero el argumento instructivo es el de los puntos de corte: quitando cualquier punto de S^1 queda un arco, que es conexo; quitando cualquier punto interior de $(0, 1)$ queda algo disconexo. Todo punto de $(0, 1)$ es de corte, ninguno de S^1 lo es. No son homeomorfos.

(b) La letra "Y" y el segmento "I". Ambos son compactos, conexos, conexos por caminos y con el cardinal del continuo: los invariantes baratos no deciden. Pero la "Y" tiene un punto (el nudo central) cuya eliminación deja *tres* componentes, mientras que en el segmento la eliminación de cualquier punto deja como mucho dos. Ese número — tres — es un invariante y no coincide. No son homeomorfos. El mismo argumento clasifica las letras: $\{C, I, L, M, N, S, U, V, W, Z\}$ son todas homeomorfas entre sí (arcos), $\{E, F, T, Y\}$ forman otra clase, y $\{O, D\}$ otra (circunferencias).

(c) $\mathbb{R}$ **y** $\mathbb{R}^2$**.** Cardinal igual, ambos conexos, ambos no compactos, ambos separables. Los invariantes baratos fallan todos. La solución es el paso 2: quítese un punto. $\mathbb{R} \setminus \{0\}$ tiene dos componentes conexas; $\mathbb{R}^2 \setminus \{p\}$ tiene una sola (dos puntos cualesquiera se unen rodeando el agujero). Como el número de componentes es invariante y un homeomorfismo llevaría un caso al otro, no existe. Este argumento elemental es el caso $n = 1$ de la invariancia de la dimensión, que en general es un teorema difícil de Brouwer.

(d) El disco cerrado y la circunferencia. Ambos son compactos y conexos por caminos. Quitar un punto no basta: quitando el centro, el disco sigue conexo. Aquí hace falta el paso 5: $\pi_1(D^2)$ es trivial porque todo lazo se contrae al centro, mientras que $\pi_1(S^1) \cong \mathbb{Z}$, con el entero midiendo cuántas vueltas da el lazo. Grupos no isomorfos, espacios no homeomorfos.

Cuidado

Un invariante solo sirve en una dirección. Que dos espacios compartan todos los invariantes que sabes comprobar *no* demuestra que sean homeomorfos; solo significa que tus herramientas no han bastado. Para lo afirmativo no hay atajo: hay que construir la aplicación, comprobar que es biyectiva, y comprobar la continuidad en ambos sentidos.

28.11. Taller: fabricar un homeomorfismo con todas sus letras

Regla: construir un homeomorfismo explícito

El apartado anterior enseñaba a demostrar que dos espacios *no* son homeomorfos. Aquí va la dirección contraria, que no tiene atajos teóricos pero sí un repertorio de trucos que cubre casi todos los casos de un primer curso.

1. **Escribe la lista de comprobaciones y no la abrevies.** Hay que dar $h : X \to Y$ y verificar cuatro cosas: que está bien definida, que es biyectiva, que es continua y que h^{-1} es continua. La cuarta es la que se olvida, y es la única que puede fallar cuando las otras tres se cumplen.

2. **Si X es compacto e Y es Hausdorff, la cuarta es automática.** Toda biyección continua de un compacto en un Hausdorff es un homeomorfismo, porque manda cerrados (compactos) en compactos (cerrados), es decir, es cerrada. Este teorema ahorra la mitad del trabajo siempre que se pueda invocar.

3. **Para trozos de recta, usa aplicaciones afines y monótonas.** $t \mapsto a + (b-a)t$ lleva $[0,1]$ a $[a,b]$, y $(0,1)$ a (a,b), respetando qué extremos se incluyen. Cualquier biyección continua y *estrictamente monótona* entre intervalos es homeomorfismo, porque su inversa vuelve a ser monótona y continua.

4. **Para pasar de acotado a no acotado, usa una función que se dispare en los extremos.** Los modelos: tan, $\log \frac{x}{1-x}$, $\frac{x}{1-x^2}$. Todos convierten un intervalo abierto en toda la recta. La clave es que sean estrictamente crecientes y que sus límites en los extremos sean $\pm\infty$.

5. **Para figuras del plano o del espacio, usa la aplicación radial.** $v \mapsto v/\|v\|$ lleva cualquier borde estrellado respecto del origen a la esfera; $v \mapsto v/(1+\|v\|)$ lleva $\mathbb{R}^n$ a la bola abierta. La receta general es: escribir el punto en coordenadas «dirección más tamaño» y cambiar solo el tamaño.

6. **Para superficies, describe el pegado y compáralo.** Dos cocientes de polígonos con el mismo esquema de identificación son homeomorfos. Así se ve que el cilindro es $S^1 \times [0,1]$, que la corona circular es lo mismo, y que la esfera es el cociente del disco por su borde.

7. **Verifica en puntos concretos y en los invariantes.** Evalúa h y h^{-1} en dos o tres puntos y comprueba que se deshacen mutuamente; y antes de empezar, comprueba que los invariantes baratos (compacidad, conexión, número de componentes) coinciden. Si no coinciden, no pierdas el tiempo buscando la fórmula.

Ejemplo 28.42 (Seis homeomorfismos escritos con todas sus letras)**.** Cada apartado da la fórmula, la inversa y la razón de que ambas sean continuas.

(a) Todos los intervalos abiertos acotados son el mismo. La aplicación $h(t) = a + (b-a)t$ lleva $(0,1)$ en (a,b) y su inversa es $h^{-1}(x) = (x-a)/(b-a)$. Ambas son afines y

por tanto continuas. Lo mismo con corchetes: $[0,1] \cong [a,b]$. Pero $[0,1] \ncong (0,1)$, porque uno es compacto y el otro no; el paso 7 lo habría avisado antes de escribir nada.

(b) Un intervalo acotado y toda la recta. La aplicación

$$h(x) = \log \frac{x}{1-x}, \qquad h : (0,1) \to \mathbb{R},$$

es estrictamente creciente, con $h(x) \to -\infty$ cuando $x \to 0^+$ y $h(x) \to +\infty$ cuando $x \to 1^-$. Su inversa es la función logística

$$h^{-1}(y) = \frac{e^y}{e^y + 1},$$

continua en todo $\mathbb{R}$. Comprobación del paso 7: $h(\frac{1}{2}) = \log 1 = 0$ y $h^{-1}(0) = \frac{1}{2}$; también $h(\frac{1}{4}) = \log \frac{1}{3} \approx -1{,}0986$ y $h^{-1}(-1{,}0986) \approx \frac{1}{4}$. Luego $(0,1) \cong \mathbb{R}$: la topología no distingue lo acotado de lo ilimitado.

(c) La circunferencia menos un punto es una recta: la proyección estereográfica. Sea $N = (0,1)$ el polo norte de S^1 y

$$h(a,b) = \frac{a}{1-b}, \qquad h : S^1 \setminus \{N\} \to \mathbb{R}.$$

Está bien definida porque $b \neq 1$. Su inversa es

$$h^{-1}(t) = \Big(\frac{2t}{t^2+1},\ \frac{t^2-1}{t^2+1}\Big),$$

que efectivamente cae en la circunferencia, pues

$$\Big(\frac{2t}{t^2+1}\Big)^2 + \Big(\frac{t^2-1}{t^2+1}\Big)^2 = \frac{4t^2 + t^4 - 2t^2 + 1}{(t^2+1)^2} = 1.$$

Las dos fórmulas son cocientes de polinomios con denominador no nulo, luego continuas. Comprobación: $h(1,0) = 1$, $h(0,-1) = 0$, $h(-1,0) = -1$, y h^{-1} los devuelve. La misma fórmula, con más coordenadas, da $S^n \setminus \{N\} \cong \mathbb{R}^n$.

(d) Todo el espacio es una bola. En $\mathbb{R}^n$, sea

$$h(v) = \frac{v}{1 + \|v\|}.$$

Entonces $\|h(v)\| = \|v\|/(1 + \|v\|) < 1$, así que h cae dentro de la bola abierta unidad. Despejando $\|v\| = \|h(v)\|/(1 - \|h(v)\|)$ se obtiene la inversa

$$h^{-1}(w) = \frac{w}{1 - \|w\|},$$

continua porque $\|w\| < 1$. Luego $\mathbb{R}^n$ es homeomorfo a una bola de radio 1. Es el mismo fenómeno de (b) en dimensión cualquiera, y explica por qué la topología no puede ver el tamaño.

(e) El cuadrado y el círculo, con la aplicación radial. Sea Q el borde del cuadrado $[-1,1]^2$ y $h(v) = v/\|v\|$ con la norma euclídea. Como el origen no está en Q, h está bien definida y es continua; y es biyectiva sobre S^1 porque cada semirrecta que sale del origen

corta a Q exactamente una vez. La inversa también es continua, pero aquí no hace falta comprobarlo: Q es compacto y S^1 es Hausdorff, así que el paso 2 lo regala. Este es el sentido preciso de la frase «un cuadrado y un círculo son la misma cosa».

(f) El cilindro y la corona circular. Sea $C = S^1 \times [0,1]$ y $A = \{v \in \mathbb{R}^2 : 1 \le \|v\| \le 2\}$. Defínase

$$h(u,t) = (1+t)\,u.$$

Es continua (producto de continuas), y biyectiva: dado $v \in A$, el par que le corresponde es $u = v/\|v\|$ y $t = \|v\| - 1$. De nuevo C es compacto y A es Hausdorff, luego h es homeomorfismo por el paso 2. Comprobación: $h(u,0) = u$ recorre el borde interior y $h(u,1) = 2u$ el exterior, como debe ser.

Cuidado

El paso 2 exige que el *dominio* sea compacto, no el codominio. La biyección continua $t \mapsto (\cos t, \sin t)$ de $[0, 2\pi)$ en S^1 tiene codominio compacto y no es homeomorfismo: su dominio no lo es. Y fuera de esa hipótesis no hay ningún atajo; la continuidad de la inversa hay que probarla, normalmente exhibiendo su fórmula como en los apartados (b), (c) y (d).

Pausa para pensar

Compárese (b) con el hecho de que $[0,1]$ y $(0,1)$ no son homeomorfos. Un intervalo abierto acotado es topológicamente idéntico a la recta entera, pero difiere de sí mismo con los extremos añadidos. Lo que la topología ve no es la longitud, sino si el espacio «se acaba» en algún sitio incluyendo el final. Añadir dos puntos cambia todo; multiplicar la longitud por infinito no cambia nada.

Idea clave

Un homeomorfismo es un isomorfismo en la categoría de los espacios topológicos, igual que un isomorfismo de grupos lo es en la de los grupos (capítulo 23) y una aplicación lineal biyectiva en la de los espacios vectoriales (capítulo 7). La diferencia incómoda es esta: en grupos y en espacios vectoriales, la inversa de un morfismo biyectivo es automáticamente un morfismo, mientras que en topología hay que exigirlo aparte. Esa asimetría, que parece un tecnicismo, es la razón de que existan biyecciones continuas que deforman irreparablemente el espacio.

28.12. Bases y subbases: describir una topología sin listarla

Una topología en un conjunto infinito tiene demasiados abiertos para enumerarlos. La solución es dar un catálogo pequeño a partir del cual se generan todos.

Definición 28.43. Una familia $\mathcal{B}$ de subconjuntos de X es *base* de una topología si todo abierto es unión de miembros de $\mathcal{B}$. Para que $\mathcal{B}$ genere una topología basta con que cubra X y que, para $B_1, B_2 \in \mathcal{B}$ y $x \in B_1 \cap B_2$, exista $B_3 \in \mathcal{B}$ con $x \in B_3 \subset B_1 \cap B_2$.

Ejemplo 28.44. En $\mathbb{R}$, los intervalos abiertos (a, b) con $a < b$ forman una base de la topología usual. De hecho basta con los de extremos racionales, y eso demuestra que $\mathbb{R}$ tiene una base *numerable*: es lo que se llama segundo axioma de numerabilidad.

Ejemplo 28.45 (la recta de Sorgenfrey)**.** Toma como base los intervalos semiabiertos $[a, b)$. La topología resultante es estrictamente más fina que la usual: $[0, 1)$ es abierto aquí y no allá. Este espacio es separable y de Lindelöf, pero no es segundo numerable ni metrizable, y su cuadrado no es normal. Es la fuente inagotable de contraejemplos de la topología general.

Idea clave

Una base es a una topología lo que una base vectorial es a un espacio vectorial: un conjunto pequeño que lo genera todo. La analogía no es perfecta —aquí no hay unicidad ni dimensión— pero la actitud es la misma: describir un objeto infinito con datos finitos o numerables.

28.13. Topología producto y topología cociente

Definición 28.46. En $X \times Y$, la *topología producto* tiene como base los rectángulos $U \times V$ con U abierto en X y V abierto en Y. Para productos infinitos $\prod_\alpha X_\alpha$ la base la forman los productos $\prod U_\alpha$ donde $U_\alpha = X_\alpha$ salvo para *un número finito* de índices.

Cuidado

Esa restricción a "salvo finitos" parece arbitraria y es esencial. Con la definición ingenua (la topología de cajas) el teorema de Tychonoff es falso y la proyección deja de ser abierta. La topología producto está definida para que sea la más gruesa que hace continuas todas las proyecciones, y esa es su verdadera razón de ser.

Teorema 28.47 (Tychonoff)**.** *Un producto arbitrario de espacios compactos es compacto en la topología producto.*

Observación 28.48. Tychonoff es equivalente al axioma de elección. Es de los pocos teoremas de matemática ordinaria de los que se sabe con exactitud cuánta elección consumen; el caso de productos numerables de métricos, en cambio, es mucho más inocente.

Definición 28.49. Dada una relación de equivalencia $\sim$ en X, la *topología cociente* en $X/\!\sim$ declara abierto a U si y solo si su preimagen por la proyección $\pi : X \to X/\!\sim$ es abierta en X.

Ejemplo 28.50. Pegando los extremos de $[0, 1]$ (es decir, $0 \sim 1$) se obtiene la circunferencia S^1. Pegando los lados opuestos de un cuadrado con la misma orientación se obtiene el toro; pegando un par con orientación invertida, la banda de Möbius; los dos pares invertidos, la botella de Klein.

Idea clave

El cociente es la formalización de "pegar". Toda la topología de superficies consiste en tomar un polígono, decidir qué lados se identifican con qué orientación, y estudiar el objeto resultante. Es exactamente el mismo mecanismo que en el capítulo 23 produce los grupos cociente G/N: partir por una relación de equivalencia y heredar la estructura.

28.14. Axiomas de numerabilidad y compactificación por un punto

Los axiomas de separación miden cuánto se pueden distinguir los puntos unos de otros. Hay una segunda familia de axiomas, independiente de aquella, que mide algo distinto: cuánta información hace falta para describir la topología. Un espacio con *pocos* abiertos básicos es manejable, porque todo argumento se reduce a un proceso de sucesiones y de listas numerables; un espacio que necesita una base descomunal se resiste a casi todas las herramientas del análisis. Y al final de la sección, un truco de una sola línea que convierte cualquier espacio en compacto añadiendo un único punto.

Definición 28.51 (Primer y segundo axioma de numerabilidad)**.** Un espacio topológico X satisface el *primer axioma de numerabilidad* (se dice 1AN) si todo punto tiene una base numerable de entornos, es decir, una familia numerable $\{B_n\}$ de entornos suyos tal que todo entorno del punto contiene algún B_n. Satisface el *segundo axioma* (2AN) si la topología entera admite una base numerable. Se dice que X es *separable* si contiene un subconjunto numerable denso, y *de Lindelöf* si todo recubrimiento abierto admite un subrecubrimiento numerable.

Idea clave

Qué dice esto en cristiano. Los axiomas de numerabilidad preguntan cuántos abiertos hacen falta para describir el espacio. El primero, 1AN, pide que *alrededor de cada punto* baste una lista numerable de entornos cada vez más pequeños; eso es lo que permite razonar con sucesiones, porque se puede elegir un punto en el entorno n-ésimo y formar una sucesión que converja. El segundo, 2AN, es más fuerte: pide una lista numerable de abiertos que sirva para *todo el espacio* a la vez. La recta real cumple los dos gracias a las bolas de radio racional centradas en puntos racionales, y por eso todo lo que uno aprende en cálculo con sucesiones funciona; en espacios que no son 1AN, las sucesiones dejan de detectar la topología y hay que sustituirlas por redes.

Ejemplo 28.52 (Los cuatro axiomas en cuatro espacios)**. (a) La recta usual.** Base numerable: los intervalos (p, q) con $p, q \in \mathbb{Q}$. Es 2AN, y por tanto 1AN, separable ($\mathbb{Q}$ es denso y numerable) y Lindelöf.

(b) Un discreto no numerable. En $\mathbb{R}$ con la topología discreta, cada punto $\{x\}$ es abierto, así que $\{\{x\}\}$ es una base local numerable: es 1AN. Pero cualquier base debe contener todos los $\{x\}$, luego no es 2AN; el único denso es todo $\mathbb{R}$, luego no es separable; y el recubrimiento por puntos no admite subrecubrimiento numerable, luego no es Lindelöf.

(c) La recta de Sorgenfrey. $\mathbb{R}$ con la base $\{[a,b)\}$ es 1AN (tómense $[x, x+1/n)$), separable (los racionales siguen siendo densos) y Lindelöf, pero *no* es 2AN: cualquier base debe contener, para cada x, un abierto cuyo mínimo sea x, y hay una cantidad no numerable de valores de x. Es el contraejemplo estándar de que separable más Lindelöf no implica 2AN.

(d) La topología conumerable. En $\mathbb{R}$, declárense abiertos el vacío y los complementarios de conjuntos numerables. No es 1AN en ningún punto, y en él las sucesiones convergentes son exactamente las que acaban siendo constantes, de modo que la convergencia de sucesiones no distingue nada. Es el aviso de que "sucesión" y "topología" no son sinónimos.

Teorema 28.53 (Relaciones entre los cuatro)**.** *Siempre se cumple* $2AN \Rightarrow 1AN$, $2AN \Rightarrow$ *separable y* $2AN \Rightarrow$ *Lindelöf. Todo espacio métrico es 1AN. Y para espacios métricos las tres condiciones 2AN, separable y Lindelöf son equivalentes entre sí. Ninguna de las implicaciones restantes es cierta en general.*

Demostración de las tres implicaciones. Sea $\mathcal{B} = \{B_n\}$ una base numerable. Para 1AN, la subfamilia de los B_n que contienen a un punto dado es numerable y es base de entornos suyos. Para separabilidad, elíjase un punto x_n en cada B_n no vacío; el conjunto $\{x_n\}$ es numerable, y es denso porque todo abierto no vacío contiene algún B_n y por tanto algún x_n. Para Lindelöf, dado un recubrimiento abierto $\mathcal{U}$, sea N el conjunto de índices n tales que $B_n \subseteq U$ para algún $U \in \mathcal{U}$; escójase para cada $n \in N$ uno de esos U, llámese U_n. La familia $\{U_n\}_{n\in N}$ es numerable, y recubre: si $x \in U \in \mathcal{U}$, hay un B_n con $x \in B_n \subseteq U$, luego $n \in N$ y $x \in B_n \subseteq U_n$. Que todo métrico es 1AN se ve tomando las bolas $B(x, 1/n)$. □

Ejemplo 28.54 (Las implicaciones y sus recíprocos)**. (a) 2AN implica los otros tres.** Si $\mathcal{B} = \{B_n\}$ es base numerable: los B_n que contienen a x forman base local numerable (1AN); eligiendo un punto en cada B_n no vacío se obtiene un denso numerable (separable); y de todo recubrimiento se extrae uno numerable quedándose con un abierto por cada B_n contenido en alguno (Lindelöf).

(b) 1AN no implica 2AN. El discreto no numerable del ejemplo anterior. La base local es tan pequeña como se quiera, pero hay demasiados puntos.

(c) Separable no implica 2AN. La recta de Sorgenfrey otra vez. Y en sentido contrario, un producto no numerable de copias de $[0,1]$ es separable (por un teorema de Hewitt–Marczewski–Pondiczery) y ni siquiera es 1AN.

(d) En espacios métricos todo se colapsa. Para espacios métricos, 2AN, separable y Lindelöf son *equivalentes*, y 1AN se cumple siempre. Por eso en cálculo estas distinciones no aparecen nunca: la métrica las borra. Sólo se vuelven visibles al salir de los espacios métricos, y ahí es donde la topología general empieza a pagar por sí misma.

Idea clave

Ser 2AN significa que la topología cabe en una lista. Y cuando algo cabe en una lista, se le puede aplicar inducción, diagonalización y todos los trucos del análisis clásico: por eso $\mathbb{R}^n$ es tan cómodo y los espacios de dimensión infinita tan incómodos. El primer axioma es más débil pero ya basta para lo que más se usa a diario: que las sucesiones detecten

la clausura y la continuidad. Fuera de 1AN hay que sustituir las sucesiones por redes, y la intuición se resiente.

Regla: decidir qué axiomas de numerabilidad cumple un espacio

1. **Búsquese primero una base explícita.** Si se encuentra una base indexada por un conjunto numerable, el espacio es 2AN y por el teorema anterior cumple también los otros tres. En $\mathbb{R}^n$ sirven las bolas de centro de coordenadas racionales y radio racional, que son tantas como $\mathbb{Q}^n \times \mathbb{Q}$, es decir numerables.

2. **Si no, pruébese el primer axioma punto a punto.** Basta exhibir, para cada x, una sucesión decreciente de entornos que atrape a todos los demás. En cualquier espacio métrico las bolas $B(x, 1/n)$ hacen el trabajo, así que todo métrico es 1AN sin más discusión.

3. **Para negar el segundo axioma, cuéntese.** La técnica estándar: si se logra asociar a cada punto de un conjunto no numerable un elemento *distinto* de cualquier base, la base no puede ser numerable. Suele bastar con encontrar, para cada x, un abierto mínimo que lo contenga.

4. **Compruébese la separabilidad buscando un denso concreto.** Los candidatos habituales son $\mathbb{Q}$, $\mathbb{Q}^n$, los polinomios de coeficientes racionales, o las sucesiones finitas de racionales. Hay que verificar que cortan a todo abierto no vacío.

5. **Para Lindelöf, atáquese el recubrimiento por partes.** La estrategia que funciona casi siempre: primero se cubre con una subfamilia numerable una parte grande del espacio, y luego se comprueba que lo que queda fuera es numerable, de modo que un abierto por punto ya basta.

6. **Recuérdese que en métricos las tres coinciden.** Si el espacio es metrizable, comprobar cualquiera de 2AN, separable o Lindelöf demuestra las otras dos. Este atajo ahorra la mitad del trabajo y conviene usarlo antes que ningún otro.

7. **Ténganse a mano los contraejemplos.** La recta de Sorgenfrey separa 1AN de 2AN; un discreto no numerable no es separable ni Lindelöf pero sí es 1AN; y el cuadrado de Sorgenfrey es el contraejemplo clásico de que Lindelöf no se hereda a productos.

Ejemplo 28.55 (Cuatro espacios y su ficha de numerabilidad)**. (a) La recta usual $\mathbb{R}$.** Es 2AN: la familia de intervalos $(p - 1/n,\ p + 1/n)$ con $p \in \mathbb{Q}$ y $n \in \mathbb{N}$ es numerable, porque está indexada por $\mathbb{Q} \times \mathbb{N}$, y es base, porque dado x y un abierto $U \ni x$ hay un $\varepsilon > 0$ con $(x - \varepsilon, x + \varepsilon) \subseteq U$; tómese n con $2/n < \varepsilon$ y un racional p a distancia menor que $1/n$ de x. Entonces $x \in (p - 1/n, p + 1/n) \subseteq U$. De 2AN se siguen los otros tres. El mismo argumento sirve en $\mathbb{R}^d$ con bolas de centro racional.

(b) La recta de Sorgenfrey, con base $[a, b)$. Es 1AN, porque los entornos $[x, x + 1/n)$ forman base de entornos de x. Es separable, porque $\mathbb{Q}$ corta a todo $[a, b)$ no vacío. Y sin

embargo *no* es 2AN: sea $\mathcal{B}$ una base cualquiera; para cada $x \in \mathbb{R}$ el abierto $[x, x+1)$ contiene a x, luego hay $B_x \in \mathcal{B}$ con $x \in B_x \subseteq [x, x+1)$, y de ahí $x = \min B_x$. Puntos distintos dan conjuntos distintos, porque el mínimo los determina, así que $x \mapsto B_x$ es inyectiva de $\mathbb{R}$ en $\mathcal{B}$ y $\mathcal{B}$ no puede ser numerable. Corolario inmediato y muy útil: la recta de Sorgenfrey *no es metrizable*, pues en un métrico separable implicaría 2AN.

(c) Un conjunto no numerable con la topología discreta. Es 1AN — el propio $\{x\}$ es una base de entornos de un solo elemento — pero no es separable, porque el único denso es todo el espacio, ni es Lindelöf, porque el recubrimiento por singletons no admite subrecubrimiento numerable, ni por tanto es 2AN. Obsérvese que es metrizable (con la distancia discreta) y separable falla: esto confirma que en la equivalencia del teorema el que manda es el conjunto, no la métrica.

(d) El cuadrado de Sorgenfrey. El producto de la recta de Sorgenfrey consigo misma es separable — $\mathbb{Q} \times \mathbb{Q}$ sigue siendo denso — pero *no* es de Lindelöf. La razón es la antidiagonal $L = \{(x, -x)\}$, que es un subespacio cerrado no numerable en el que la topología inducida es la discreta, porque el abierto básico $[x, x+1) \times [-x, -x+1)$ corta a L solo en el punto $(x, -x)$. Un cerrado discreto no numerable impide Lindelöf. Como la recta de Sorgenfrey sí es de Lindelöf, este ejemplo demuestra que la propiedad no pasa a productos, ni siquiera a productos de dos factores idénticos.

Ejemplo 28.56 (La compactificación de Alexandroff, en cuatro casos)**.** La construcción es esta: dado X, añádase un punto nuevo, llamado ∞, y declárese abierto en $X^* = X \cup \{\infty\}$ todo abierto de X, más todo conjunto de la forma $\{\infty\} \cup (X \setminus K)$ con $K \subseteq X$ compacto y cerrado. El espacio X^* resulta ser siempre compacto, y es de Hausdorff exactamente cuando X es de Hausdorff y localmente compacto.

(a) La recta y el círculo. $\mathbb{R}^* \cong S^1$. La biyección es la proyección estereográfica desde el polo norte del círculo sobre la recta tangente en el polo sur, extendida enviando el polo norte a ∞. Los entornos de ∞ son los complementarios de los compactos de $\mathbb{R}$, es decir, los conjuntos que contienen a $(-\infty, -M) \cup (M, +\infty)$: exactamente los arcos alrededor del polo norte. Los dos extremos de la recta se pegan en un solo punto, y por eso el círculo tiene un agujero que la recta no tenía.

(b) El plano y la esfera. $(\mathbb{R}^2)^* \cong S^2$, de nuevo por proyección estereográfica; es la esfera de Riemann del capítulo 30, y explica por qué en variable compleja se habla de "el punto del infinito" en singular. En general $(\mathbb{R}^n)^* \cong S^n$. Obsérvese que $(0,1)^* \cong S^1$ también, porque $(0,1)$ es homeomorfo a $\mathbb{R}$: la compactificación por un punto pega los *dos* extremos en uno, a diferencia de la compactificación $[0,1]$, que añade dos puntos distintos. Son compactificaciones distintas del mismo espacio.

(c) Los naturales y una sucesión convergente. Sea $\mathbb{N}$ con la topología discreta. Sus compactos son exactamente los finitos, luego los entornos de ∞ son los cofinitos, y $\mathbb{N}^*$ es homeomorfo al subespacio de $\mathbb{R}$

$$\Big\{0\Big\} \cup \Big\{1, \tfrac{1}{2}, \tfrac{1}{3}, \tfrac{1}{4}, \tfrac{1}{5}, \dots\Big\},$$

con ∞ haciendo de 0. Es el ejemplo más sencillo de espacio compacto infinito, y es el modelo mental correcto de "sucesión que converge": un espacio en el que un punto distinguido es el límite de todos los demás.

(d) Los racionales, donde falla Hausdorff. $\mathbb{Q}$ es de Hausdorff pero no localmente compacto: ningún punto tiene un entorno de clausura compacta, porque toda bola cerrada de $\mathbb{Q}$ tiene agujeros irracionales y admite recubrimientos sin subrecubrimiento finito. Por tanto $\mathbb{Q}^*$ es compacto y *no* es de Hausdorff: el punto ∞ y cualquier racional no se pueden separar. Sirve de recordatorio de que compacidad y separación son propiedades independientes, y de que la hipótesis de compacidad local en el criterio no es un adorno técnico.

Cuidado

[Lindelöf no implica el segundo axioma] La tentación es pensar que "todo recubrimiento se reduce a uno numerable" equivale a "hay una base numerable". La recta de Sorgenfrey desmiente esto: es de Lindelöf — dado un recubrimiento por básicos $[a_x, b_x)$, el conjunto de los x que no quedan en el interior de la unión resulta ser numerable, y se cubren de uno en uno — y no es 2AN, según se acaba de demostrar en el apartado (b). La flecha 2AN $\Rightarrow$ Lindelöf solo va en un sentido, y únicamente en presencia de una métrica se puede recorrer al revés. Del mismo modo, separable no implica 2AN, ni Lindelöf implica separable, ni 1AN implica ninguna de las tres.

Pausa para pensar

La compactificación de Alexandroff es la más barata posible: añade un solo punto. En el otro extremo está la compactificación de Stone–Čech, que añade tantos puntos como haga falta para que *toda* función continua y acotada $X \to \mathbb{R}$ se extienda con continuidad. Para $\mathbb{N}$, la primera añade un punto y la segunda añade una cantidad de puntos mayor que la del continuo. Entre ambas hay un retículo entero de compactificaciones, ordenado por refinamiento, con la de Alexandroff como mínimo y la de Stone–Čech como máximo. Que un problema tan concreto — "hacer compacto un espacio" — tenga una respuesta con estructura de retículo es una muestra más de que en topología las construcciones rara vez son únicas, y de que la pregunta correcta no es "cuál" sino "cuál es la mejor para lo que quiero hacer".

28.15. Axiomas de separación

No todos los espacios topológicos permiten distinguir puntos con abiertos. La jerarquía estándar mide cuánto se puede separar.

Definición 28.57. T_1: los puntos son cerrados.
T_2 (*Hausdorff*): dos puntos distintos tienen entornos disjuntos.
T_3 (*regular*): un punto y un cerrado que no lo contiene tienen entornos disjuntos.
T_4 (*normal*): dos cerrados disjuntos tienen entornos disjuntos.

Ejemplo 28.58. La topología de complementos finitos en un conjunto infinito es T_1 pero no Hausdorff: dos abiertos no vacíos cualesquiera siempre se cortan, porque sus complementos son finitos.

Teorema 28.59. *Todo espacio métrico es normal, y por tanto cumple todos los axiomas anteriores.*

Ejemplo 28.60 (Espacios que no separan)**. (a) La demostración en métricos.** Dados cerrados disjuntos A, B, la función $f(x) = d(x, A)/(d(x, A) + d(x, B))$ es continua, vale 0 en A y 1 en B; los conjuntos $f^{-1}([0;\ 0{,}5))$ y $f^{-1}((0{,}5;\ 1])$ son abiertos disjuntos que separan A de B. El denominador no se anula porque A y B son cerrados disjuntos.

(b) La topología trivial. En un conjunto con dos puntos y sólo los abiertos $\emptyset$ y X, ningún par de puntos se separa: no es ni T_1 ni Hausdorff. En él, toda sucesión converge a *todos* los puntos a la vez, lo que muestra por qué la unicidad del límite es exactamente la propiedad de Hausdorff.

(c) T_1 sin Hausdorff. En un conjunto infinito con la topología cofinita, los puntos son cerrados (T_1) pero dos abiertos no vacíos cualesquiera se cortan, porque sus complementarios son finitos: no es Hausdorff. La sucesión $1, 2, 3, \ldots$ converge a todos los puntos.

(d) Hausdorff sin normal. El plano de Sorgenfrey, producto de dos rectas de Sorgenfrey, es Hausdorff, regular y *no* normal: la antidiagonal $\{(x, -x)\}$ es un cerrado discreto no numerable que no se puede separar por abiertos de sus subconjuntos. La cadena $T_1 \subset T_2 \subset T_3 \subset T_4$ es estrictamente creciente, y cada eslabón tiene su contraejemplo clásico.

Idea clave

Hausdorff es la condición que hace que los límites sean únicos. Sin ella, una sucesión puede converger a dos puntos distintos a la vez, y toda la intuición analítica se desmorona. Por eso casi toda la geometría y el análisis trabajan solo con espacios Hausdorff, y por eso conviene saber que la hipótesis está ahí, aunque casi nunca se mencione.

Teorema 28.61 (Lema de Urysohn)**.** *Si X es normal y A, B son cerrados disjuntos, existe $f : X \to [0, 1]$ continua con $f \equiv 0$ en A y $f \equiv 1$ en B.*

Idea clave

Qué dice esto en cristiano. La normalidad se define con abiertos: dos cerrados disjuntos se pueden meter en abiertos disjuntos. El lema de Urysohn dice que eso, que parece poca cosa, basta para construir una *función continua* que valga cero en uno y uno en el otro. Es decir: separar con abiertos y separar con funciones es lo mismo. La importancia práctica es enorme, porque las funciones se pueden sumar, multiplicar y combinar, mientras que los abiertos no; de aquí salen las particiones de la unidad, y de ellas casi toda la geometría diferencial.

Ejemplo 28.62 (Funciones de Urysohn concretas)**. (a) En la recta.** Con $A = (-\infty, 0]$ y $B = [1, \infty)$, la función $f(x) = \min(\max(x, 0), 1)$ vale 0 en A, 1 en B y es continua. Aquí no hace falta ninguna maquinaria: la fórmula de la distancia normalizada del ejemplo anterior da lo mismo.

(b) Suave en vez de continua. Si además se quiere que f sea de clase C^∞, se usa

$\varphi(t) = \mathrm{e}^{-1/t}$ para $t > 0$ y 0 para $t \le 0$, y se construye $f(x) = \varphi(x)/(\varphi(x) + \varphi(1-x))$. Es infinitamente derivable, vale 0 para $x \le 0$ y 1 para $x \ge 1$, y todas sus derivadas se anulan en los dos puntos de empalme. Ésta es la versión que se usa en variedades.

(c) La construcción del lema. La demostración general asigna un abierto U_q a cada racional $q \in [0,1]$ de modo que $q < q'$ implique $\overline{U_q} \subset U_{q'}$, y define $f(x) = \inf\{q : x \in U_q\}$. Es un encaje de infinitas capas entre A y B, como las curvas de nivel de un mapa topográfico.

(d) Lo que el lema no da. No garantiza que $f^{-1}(0)$ sea exactamente A: sólo que f vale 0 *en* A. Para conseguir la igualdad hace falta que A sea un conjunto G_δ, y ésa es la versión fuerte del lema. En espacios métricos siempre ocurre, porque $A = \bigcap_n \{x : d(x,A) < 1/n\}$.

Observación 28.63. El lema de Urysohn es sorprendente: la normalidad es una condición sobre abiertos, puramente cualitativa, y de ella se extrae una función continua a valores reales. Es la primera vez que la topología *fabrica* análisis. Su consecuencia inmediata, el teorema de extensión de Tietze, permite extender funciones continuas desde un cerrado a todo el espacio.

28.16. Hausdorff: el axioma que se usa a diario

Regla: demostrar que un espacio es de Hausdorff, o que no lo es

De todos los axiomas de separación, el único que se usa a diario es T_2. Merece un procedimiento propio, porque casi todos los teoremas que uno querrá aplicar más adelante lo llevan escondido en las hipótesis.

1. **Escribe lo que hay que construir.** Dados $x \neq y$, hacen falta abiertos $U \ni x$ y $V \ni y$ con $U \cap V = \varnothing$. Son dos conjuntos y una condición; nada más. Si el espacio es métrico, tómese $r = d(x,y)/2$ y las bolas $B(x,r)$, $B(y,r)$: la desigualdad triangular garantiza que no se cortan.

2. **Hereda en vez de demostrar.** Todo subespacio de un Hausdorff es Hausdorff, y todo producto (finito o infinito) de Hausdorff lo es. Como los espacios métricos son Hausdorff, cualquier subconjunto de $\mathbb{R}^n$ con la topología usual lo es sin más argumento.

3. **Desconfía de los cocientes.** La propiedad de Hausdorff *no* se hereda al pasar al cociente. Es la excepción de la lista anterior, y la fuente de casi todos los ejemplos no separados que aparecen en geometría.

4. **Usa el criterio de la diagonal cuando los puntos estorben.** X es Hausdorff si y solo si $\Delta = \{(x,x)\}$ es cerrada en $X \times X$. Convierte una condición con dos cuantificadores existenciales en la clausura de un solo conjunto, y es la vía habitual en las demostraciones abstractas.

5. **Cosecha las consecuencias en cuanto la tengas.** En un Hausdorff, los límites son únicos, todo compacto es cerrado, los puntos son cerrados (luego $T_2 \Rightarrow T_1$), y dos funciones continuas que coinciden en un denso coinciden en todo el espacio. Estas cuatro afirmaciones son falsas sin T_2.

6. **Para probar que NO lo es, exhibe dos puntos inseparables.** Basta encontrar $x \neq y$ tales que todo entorno de x corte a todo entorno de y. Alternativamente, exhibe una sucesión con dos límites distintos, o un compacto que no sea cerrado: cualquiera de las tres cosas delata la falta de T_2.

7. **Comprueba dónde queda el espacio en la jerarquía.** T_2 implica T_1, e implica que los puntos son cerrados; pero T_1 *no* implica T_2. Antes de dar por concluido el análisis, sitúa el ejemplo: ¿es T_1? ¿es T_2? ¿es normal? Los tres veredictos son independientes y hay que darlos por separado.

Ejemplo 28.64 (Cinco espacios ante el espejo de Hausdorff)**.** Los dos primeros son fáciles; los tres últimos son los contraejemplos que conviene tener guardados.

(a) Todo espacio métrico, con el radio explícito. Sean $x \neq y$ y $r = d(x,y)/2 > 0$. Si existiera $z \in B(x,r) \cap B(y,r)$, la desigualdad triangular daría

$$d(x,y) \leq d(x,z) + d(z,y) < r + r = d(x,y),$$

lo cual es absurdo. Luego las dos bolas son disjuntas y el espacio es Hausdorff. Este cálculo de tres líneas cubre de golpe $\mathbb{R}^n$, todos sus subconjuntos, los espacios normados y los espacios de funciones con la norma del supremo.

(b) La topología indiscreta, el extremo opuesto. En $X = \{a, b\}$ con $\tau = \{\varnothing, X\}$, el único abierto que contiene a a es X, y el único que contiene a b también. Nunca son disjuntos: no es Hausdorff, ni siquiera T_1. Y la consecuencia del paso 5 se ve en carne viva: la sucesión constante $a, a, a, \ldots$ converge a a *y* a b, porque todo entorno de b —a saber, X— contiene a todos los términos. Sin separación, «el límite» deja de ser una expresión legítima.

(c) La recta con dos orígenes. Tómense dos copias de $\mathbb{R}$ y péguese cada punto $x \neq 0$ de una con el mismo x de la otra, dejando sin pegar los dos ceros, que llamaremos 0_1 y 0_2. El resultado es T_1 —cada punto es cerrado— y localmente indistinguible de $\mathbb{R}$: todo punto tiene un entorno homeomorfo a un intervalo. Y sin embargo no es Hausdorff: cualquier entorno de 0_1 contiene un intervalo $(-\varepsilon, \varepsilon)$ sin el cero, y lo mismo cualquier entorno de 0_2; esos dos intervalos siempre se cortan. Aquí se ve el paso 3 en acción: el espacio es un cociente de dos rectas, que son Hausdorff, y la propiedad se perdió al pegar. Por eso la definición de variedad incluye «Hausdorff» como hipótesis expresa: no se sigue de ser localmente euclídeo.

(d) Un compacto que no es cerrado. En la recta con dos orígenes, el conjunto $K = [-1, 0_1] \cup [0_1, 1]$ —es decir, el intervalo $[-1,1]$ que pasa por el primer origen— es compacto, por ser imagen continua de $[-1,1]$. Pero no es cerrado: el punto 0_2 está en su clausura, ya que todo entorno de 0_2 corta a K, y sin embargo $0_2 \notin K$. El teorema «compacto $\Rightarrow$ cerrado» del paso 5 necesita T_2, y este es el testigo de que la hipótesis no sobra.

(e) La topología del punto particular. En $X = \mathbb{R}$, declárese abierto el vacío y todo conjunto que contenga al 0. Es una topología legítima: el vacío y X están, la unión de conjuntos que contienen al 0 contiene al 0, y la intersección de dos de ellos también. Los cerrados son, por tanto, X y los conjuntos que *no* contienen al 0.

No es Hausdorff, y por la razón más brusca posible: dos abiertos no vacíos cualesquiera contienen ambos al 0, luego jamás son disjuntos. Tampoco es T_1, pero conviene mirar de cerca

por qué. El conjunto $\{1\}$ no contiene al 0, así que es cerrado; en cambio $\{0\}$ sí lo contiene, así que no es cerrado. Hay entonces puntos cerrados y puntos no cerrados conviviendo en el mismo espacio, y basta ese único punto rebelde para que falle T_1: la condición se exige sobre *todos* los puntos, no sobre algunos.

Cuidado

No se confunda «no es Hausdorff» con «es raro y no importa». Los espacios no separados son el pan de cada día en geometría algebraica: la topología de Zariski sobre una variedad casi nunca es Hausdorff, porque sus cerrados son los conjuntos de soluciones de sistemas de ecuaciones y hay muy pocos. Sobre una recta, la topología de Zariski es exactamente la cofinita. Y sin embargo toda la teoría funciona, porque se sustituye la separación por otras hipótesis.

Pausa para pensar

Los apartados (c) y (d) están conectados por un hilo fino. Sin T_2 se pierde la unicidad de los límites; y perdida esa unicidad, se cae el argumento estándar de que un compacto es cerrado, que consiste precisamente en separar el punto de fuera de todo el compacto con abiertos disjuntos. Un axioma de separación no es un capricho de clasificación: es la hipótesis mínima que hace que las palabras «el límite» y «el punto más cercano» signifiquen algo.

Idea clave

La jerarquía $T_1 \subset T_2 \subset T_3 \subset T_4$ mide, literalmente, cuántos abiertos tiene el espacio: cuanto más fina es la topología, más fácil es separar. En la discreta se separa todo; en la indiscreta, nada. Los espacios métricos están cómodamente arriba —son normales, y por tanto cumplen los cuatro—, y el resultado recíproco, los *teoremas de metrización*, dice cuándo un espacio topológico abstracto puede venir de una distancia: el de Urysohn afirma que basta con ser regular, T_1 y segundo numerable. Es la respuesta a la pregunta que abría el capítulo: ¿cuánta información de la distancia queda realmente en los abiertos?

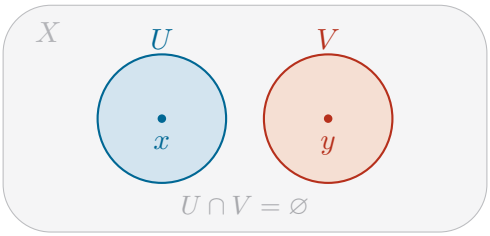


Figura 28.2: La condición de Hausdorff: dos puntos distintos siempre pueden encerrarse en abiertos que no se tocan. Es lo que garantiza que los límites, cuando existen, sean únicos.

28.17. Compacidad: las tres caras

Teorema 28.65. *Para un espacio métrico X son equivalentes:* (i) *todo cubrimiento abierto admite subcubrimiento finito;* (ii) *toda sucesión tiene una subsucesión convergente (compacidad secuencial);* (iii) *X es completo y totalmente acotado.*

Cuidado

Esa equivalencia *solo* vale en espacios métricos. En topología general, compacidad y compacidad secuencial son independientes: hay compactos no secuencialmente compactos (el cubo $[0,1]^{[0,1]}$) y al revés (el primer ordinal no numerable con la topología del orden).

Teorema 28.66 (Heine–Borel)**.** *En $\mathbb{R}^n$, un conjunto es compacto si y solo si es cerrado y acotado.*

Cuidado

Heine–Borel es una propiedad de $\mathbb{R}^n$, no una definición de compacidad. En dimensión infinita falla estrepitosamente: la bola unidad cerrada de ℓ^2 es cerrada y acotada y no es compacta, como se ve en el capítulo 32. Confundir "cerrado y acotado" con "compacto" es el error que separa a quien entendió la topología de quien memorizó un teorema.

28.18. Conexión por caminos

Definición 28.67. X es *conexo por caminos* si para todos $x, y \in X$ existe una función continua $\gamma : [0,1] \to X$ con $\gamma(0) = x$ y $\gamma(1) = y$.

Proposición 28.68. *Conexo por caminos implica conexo. El recíproco es falso.*

Ejemplo 28.69 (la curva del topólogo)**.** Sea $S = \{(x, \sin(1/x)) : 0 < x \le 1\} \cup (\{0\} \times [-1,1])$. Es conexo, porque es la clausura de un conexo, pero no conexo por caminos: ningún camino continuo llega desde el segmento vertical hasta la curva oscilante, porque tendría que recorrer infinitas oscilaciones en tiempo finito.

Idea clave

Este ejemplo explica por qué hay dos nociones y no una. La conexión es la condición "no puedo partirte en dos abiertos"; la conexión por caminos es "puedo caminar de aquí a allá". La segunda es más fuerte, más intuitiva y más útil, y es la que se usa para definir el grupo fundamental.

28.19. Una mirada al grupo fundamental

Definición 28.70. Fijado un punto base $x_0 \in X$, el *grupo fundamental* $\pi_1(X, x_0)$ es el conjunto de clases de homotopía de lazos que empiezan y terminan en x_0, con la operación de

recorrer un lazo tras otro.

Ejemplo 28.71. $\pi_1(\mathbb{R}^n) = \{e\}$: todo lazo se contrae al punto. $\pi_1(S^1) = \mathbb{Z}$: la clase de un lazo queda determinada por cuántas vueltas da y en qué sentido. $\pi_1(T^2) = \mathbb{Z} \times \mathbb{Z}$: dos vueltas independientes, una por cada agujero del toro.

Teorema 28.72. *π_1 es un invariante topológico: espacios homeomorfos tienen grupos fundamentales isomorfos.*

Idea clave

Qué dice esto en cristiano. El grupo fundamental cuenta los agujeros de un espacio mirando qué lazos se pueden encoger hasta un punto y cuáles no. Si en un espacio todo lazo se puede contraer, el grupo es trivial y el espacio es simplemente conexo. Si hay un agujero, un lazo que lo rodee no se puede contraer, y el número de vueltas que da se convierte en un número entero. Que sea un invariante topológico significa que dos espacios con grupos distintos no pueden ser homeomorfos, y eso convierte una pregunta geométrica difícil en una comparación de grupos.

Ejemplo 28.73 (Contar agujeros con lazos)**. (a) Los casos triviales.** En $\mathbb{R}^n$, en una bola o en cualquier convexo, todo lazo se contrae tirando de él hacia el centro: $\pi_1 = \{1\}$. Lo mismo en la esfera S^2, donde un lazo se puede deslizar hasta un polo. Ninguno de estos espacios tiene agujeros que atrapen lazos.

(b) La circunferencia. En S^1, un lazo queda caracterizado por cuántas vueltas da y en qué sentido, y ese número entero no cambia al deformar: $\pi_1(S^1) = \mathbb{Z}$. Es el primer cálculo no trivial de la topología algebraica y de él salen el teorema del punto fijo de Brouwer en dimensión dos y el teorema fundamental del álgebra.

(c) Distinguir espacios. El plano $\mathbb{R}^2$ tiene π_1 trivial y el plano perforado $\mathbb{R}^2 \setminus \{0\}$ tiene $\pi_1 = \mathbb{Z}$, luego no son homeomorfos. Ese es exactamente el mismo agujero que hacía fallar el lema de Poincaré en cálculo vectorial, y el "número de vueltas" es la integral $\frac{1}{2\pi} \oint \frac{-y\,\mathrm{d}x + x\,\mathrm{d}y}{x^2+y^2}$.

(d) Casos con más estructura. El toro tiene $\pi_1 = \mathbb{Z} \times \mathbb{Z}$, con un generador por cada uno de sus dos agujeros, y es abeliano. La figura de ocho tiene el grupo libre de dos generadores, que *no* es abeliano: dar la vuelta a un lazo y luego al otro no es lo mismo que hacerlo al revés. La topología distingue esos dos espacios sin necesidad de mirarlos.

Cuidado

[El grupo fundamental no lo ve todo] π_1 trivial no significa que el espacio no tenga agujeros: la esfera S^2 tiene $\pi_1 = \{1\}$ y sin embargo tiene un hueco de dimensión dos, que detecta $\pi_2 = \mathbb{Z}$. Hace falta toda la familia de grupos de homotopía, y aun así no bastan para clasificar. Segundo aviso: el grupo depende del punto base elegido, aunque en un espacio conexo por caminos todos los puntos dan grupos isomorfos. Y tercero: invariante significa que grupos distintos implican espacios distintos; el recíproco es falso, y hay espacios no homeomorfos con el mismo grupo fundamental.

Idea clave

Aquí ocurre la fusión que da nombre a la topología algebraica: a cada espacio se le asocia un grupo, y las preguntas geométricas —¿son homeomorfos? ¿tiene agujeros?— se convierten en preguntas algebraicas sobre grupos, que sabemos responder con las herramientas del capítulo 23. Que $\pi_1(S^1) = \mathbb{Z}$ y $\pi_1(\mathbb{R}^2)$ sea trivial demuestra de un golpe que el círculo no es contráctil, algo que ninguna cuenta con abiertos habría dado tan limpiamente.

Observación 28.74. De $\pi_1(S^1) = \mathbb{Z}$ se deduce en pocas líneas el teorema del punto fijo de Brouwer en dimensión dos: toda función continua del disco en sí mismo tiene un punto fijo. Y de Brouwer se deduce el teorema fundamental del álgebra, el teorema de Borsuk–Ulam y buena parte de la teoría de juegos —el teorema de Nash usa un pariente cercano, el punto fijo de Kakutani—. La topología, que empezó preguntándose qué es un agujero, termina demostrando que existe un equilibrio económico.

Regla: calcular el grupo fundamental y usarlo para distinguir espacios

El grupo fundamental convierte «cuántos agujeros tiene» en un objeto algebraico. Es el primer invariante que distingue espacios que la conexión y la compacidad no logran separar.

1. **Elige un punto base y olvídalo enseguida.** $\pi_1(X, x_0)$ es el conjunto de lazos que empiezan y terminan en x_0, módulo homotopía, con la concatenación como operación. Si X es conexo por caminos, cambiar x_0 da grupos isomorfos, así que se escribe simplemente $\pi_1(X)$.

2. **Reduce el espacio antes de calcular.** Si X se retrae por deformación a A, entonces $\pi_1(X) \cong \pi_1(A)$. Este es el paso que más trabajo ahorra: casi todo espacio razonable se contrae a algo de la lista siguiente.

3. **Memoriza la lista básica.** $\pi_1(\mathbb{R}^n) = 1$ y en general todo convexo o contráctil tiene grupo trivial. $\pi_1(S^1) = \mathbb{Z}$. $\pi_1(S^n) = 1$ para $n \geq 2$. $\pi_1(T^2) = \mathbb{Z}^2$. $\pi_1(S^1 \vee S^1) = F_2$, el grupo libre en dos generadores. $\pi_1(\mathbb{RP}^2) = \mathbb{Z}/2\mathbb{Z}$.

4. **Usa el producto.** $\pi_1(X \times Y) \cong \pi_1(X) \times \pi_1(Y)$. De aquí sale $\pi_1(T^2) = \pi_1(S^1 \times S^1) = \mathbb{Z} \times \mathbb{Z}$ sin ningún trabajo.

5. **Para pegar espacios, usa Van Kampen.** Si $X = U \cup V$ con U, V abiertos y $U \cap V$ conexo por caminos, $\pi_1(X)$ es el producto libre de $\pi_1(U)$ y $\pi_1(V)$ amalgamado sobre $\pi_1(U \cap V)$. Es la herramienta de cálculo general.

6. **Aplícalo: espacios homeomorfos tienen grupos isomorfos.** Por contrarrecíproco, si los grupos no son isomorfos los espacios no son homeomorfos. Y como el grupo es invariante por homotopía, distingue incluso cosas que no son homeomorfas por razones más finas.

7. **Verifica el resultado con un lazo concreto.** Dibuja el generador y comprueba que no se puede contraer. Si el grupo te sale trivial pero ves un lazo que rodea un agujero, has usado una retracción que no era válida.

Ejemplo 28.75 (Seis espacios, sus grupos, y dos teoremas famosos como corolarios)**.** Calculamos con el procedimiento y luego cosechamos.

(a) El plano agujereado. $\mathbb{R}^2 \setminus \{0\}$ se retrae por deformación a la circunferencia unidad mediante $H_t(v) = (1-t)v + t\frac{v}{\|v\|}$, que está bien definida porque $v \neq 0$. Por los pasos 2 y 3,

$$\pi_1(\mathbb{R}^2 \setminus \{0\}) \cong \pi_1(S^1) = \mathbb{Z}.$$

El isomorfismo es explícito: a cada lazo γ le asocia su *número de vueltas*

$$n(\gamma) = \frac{1}{2\pi} \oint_\gamma \frac{x\,\mathrm{d}y - y\,\mathrm{d}x}{x^2+y^2},$$

la misma integral que aparece en el capítulo 11 al estudiar campos cerrados no exactos, y la misma que en el capítulo 30 se escribe $\frac{1}{2\pi i}\oint \frac{\mathrm{d}z}{z}$. Tres capítulos distintos, un solo entero.

(b) El espacio agujereado, y por qué $\mathbb{R}^2 \not\cong \mathbb{R}^3$. En cambio $\mathbb{R}^3 \setminus \{0\}$ se retrae a S^2, cuyo grupo es trivial: en dimensión tres se puede deslizar cualquier lazo alrededor del agujero hasta contraerlo. Luego

$$\pi_1(\mathbb{R}^2 \setminus \{0\}) = \mathbb{Z} \;\neq\; 1 = \pi_1(\mathbb{R}^3 \setminus \{0\}),$$

y por el paso 6, $\mathbb{R}^2$ y $\mathbb{R}^3$ no son homeomorfos. Compara con el argumento de conexión del apartado anterior: aquel separaba $\mathbb{R}$ de $\mathbb{R}^2$ contando componentes; este separa $\mathbb{R}^2$ de $\mathbb{R}^3$ contando lazos. Cada dimensión necesita un invariante de un grado más.

(c) El toro y el ocho: dos agujeros, dos grupos distintos. El toro $T^2 = S^1 \times S^1$ tiene, por el paso 4, $\pi_1(T^2) = \mathbb{Z}^2$, generado por los dos lazos a (meridiano) y b (paralelo), que *conmutan*: el lazo $aba^{-1}b^{-1}$ se contrae, deslizándolo por la superficie. El ocho $S^1 \vee S^1$ (dos circunferencias pegadas por un punto) tiene $\pi_1 = F_2$, el grupo libre, donde $aba^{-1}b^{-1} \neq 1$. Como $\mathbb{Z}^2$ es abeliano y F_2 no lo es, los dos espacios no son homeomorfos, pese a que ambos «tienen dos agujeros».

Cuidado

«Contar agujeros» no es un invariante: hay que contar *cómo se relacionan*. El toro y el ocho tienen la misma homología en dimensión uno ($\mathbb{Z}^2$ en ambos casos, porque la homología es la abelianización del grupo fundamental), y sin embargo se distinguen por π_1. El grupo fundamental ve más que la homología, al precio de ser mucho más difícil de calcular.

(d) La banda de Möbius y el cilindro. Ambos se retraen a su circunferencia central, luego ambos tienen $\pi_1 = \mathbb{Z}$: el grupo fundamental *no* los distingue. Lo que los separa es la orientabilidad, o de forma más elemental la frontera: la del cilindro tiene dos componentes, la de la banda de Möbius solo una. Un invariante no lo resuelve todo, y saber cuál falla es parte del oficio.

(e) Primer corolario: no hay retracción del disco a su borde. Supón que existiera $r : D^2 \to S^1$ continua con $r(x) = x$ para $x \in S^1$. Aplicando π_1 a la composición $S^1 \hookrightarrow D^2 \xrightarrow{r} S^1$, que es la identidad, se obtendría una factorización de $\mathrm{id}_{\mathbb{Z}}$ a través de $\pi_1(D^2) = 1$:

$$\mathbb{Z} \longrightarrow 1 \longrightarrow \mathbb{Z}.$$

Imposible, porque la composición sería la aplicación nula y no la identidad. Luego no existe tal r.

(f) Segundo corolario: el teorema del punto fijo de Brouwer en dimensión dos. Si $f : D^2 \to D^2$ fuera continua y sin puntos fijos, para cada x el rayo que sale de $f(x)$ y pasa por x cortaría a S^1 en un punto $r(x)$; esa r sería continua y fijaría el borde, es decir, una retracción. Contradice (e). Luego *toda* función continua del disco en sí mismo tiene un punto fijo. Un teorema sobre existencia de soluciones, demostrado contando vueltas.

Pausa para pensar

Observa la estructura del argumento (e)–(f): se traduce un problema geométrico a un problema de grupos, se ve que la traducción es imposible, y se concluye que el problema geométrico no tenía solución. Ese movimiento — funtorialidad — es la idea central de toda la topología algebraica, y el mismo esquema demuestra el teorema fundamental del álgebra: un polinomio sin raíces daría una homotopía entre un lazo de grado n y uno constante.

Idea clave

π_1 es solo el primer piso. Existen $\pi_n(X)$ para todo n, definidos con esferas S^n en lugar de lazos, y todos son abelianos para $n \geq 2$. Son extraordinariamente difíciles: nadie conoce todos los $\pi_n(S^2)$, y $\pi_3(S^2) = \mathbb{Z}$ (la fibración de Hopf) fue una de las sorpresas mayores del siglo, porque muestra que una esfera de dimensión baja puede tener «agujeros» detectables solo desde dimensiones más altas.

28.20. Espacios recubridores: desenrollar un espacio

El grupo fundamental mide los agujeros contando lazos que no se pueden deshacer. Hay una segunda manera de mirar exactamente lo mismo, más geométrica y en muchos casos más fácil de dibujar: en lugar de estudiar los lazos dentro del espacio, se construye *otro* espacio encima del primero en el que esos lazos se abren y dejan de ser lazos. La imagen mental es una rampa de aparcamiento en espiral: vista desde arriba parece un círculo, pero al recorrerla uno no vuelve al mismo sitio, sino un piso más arriba. Ese es el ejemplo fundador de toda la teoría.

Definición 28.76 (Espacio recubridor)**.** Sean $\widetilde{X}$ y X espacios topológicos y $p\colon \widetilde{X} \to X$ una aplicación continua y sobreyectiva. Se dice que p es un *recubrimiento* (y que $\widetilde{X}$ es un *espacio recubridor* de X) si todo punto $x \in X$ tiene un entorno abierto U tal que

$$p^{-1}(U) = \bigsqcup_{i \in I} V_i,$$

unión disjunta de abiertos $V_i \subseteq \widetilde{X}$, cada uno de los cuales p aplica homeomórficamente sobre U. Un tal U se llama *entorno uniformemente recubierto* y los V_i, las *hojas* sobre U. Si X es conexo, el cardinal de la fibra $p^{-1}(x)$ no depende de x; se llama el *número de hojas* del recubrimiento.

Idea clave

Qué dice esto en cristiano. Un recubrimiento es un espacio que se enrolla sobre otro de forma localmente indistinguible. Mirando de cerca, $\widetilde{X}$ y X son idénticos; mirando de lejos, $\widetilde{X}$ puede tener varias copias apiladas encima de cada punto de X. La imagen mental es una escalera de caracol proyectándose sobre un círculo: cada punto del círculo tiene infinitos escalones encima, y sin embargo cada tramo de escalera se ve como un tramo de círculo. Esa combinación de igualdad local con diferencia global es lo que hace del recubrimiento la herramienta para detectar agujeros.

Ejemplo 28.77 (Cuatro recubrimientos)**. (a) La escalera de caracol.** $p\colon \mathbb{R} \to S^1$, $t \mapsto (\cos 2\pi t, \sin 2\pi t)$. Cada punto de S^1 tiene una fibra $p^{-1}(x)$ en biyección con $\mathbb{Z}$: infinitas hojas. Es el recubrimiento universal de la circunferencia y es el que demuestra que $\pi_1(S^1) = \mathbb{Z}$.

(b) Envolver n veces. $p\colon S^1 \to S^1$, $z \mapsto z^n$, tiene fibras de exactamente n puntos: las n raíces n-ésimas de z. Con $n = 3$, sobre el punto 1 están 1, $\mathrm{e}^{2\pi i/3}$ y $\mathrm{e}^{4\pi i/3}$.

(c) La esfera sobre el proyectivo. $p\colon S^2 \to \mathbb{R}P^2$, que identifica cada punto con su antípoda, tiene fibras de dos puntos. Como S^2 es simplemente conexa, es el recubrimiento universal, y de ahí se lee $\pi_1(\mathbb{R}P^2) = \mathbb{Z}/2\mathbb{Z}$.

(d) El plano sobre el toro. $p\colon \mathbb{R}^2 \to T^2$, $(s,t) \mapsto (\mathrm{e}^{2\pi i s}, \mathrm{e}^{2\pi i t})$, con fibras en biyección con $\mathbb{Z}^2$. Y un no ejemplo: la aplicación $p\colon (0,3) \to S^1$, $t \mapsto \mathrm{e}^{2\pi i t}$, *no* es recubrimiento, porque sobre el punto 1 hay tres preimágenes pero sobre los puntos con argumento pequeño hay cuatro, y ningún entorno se parte en hojas homeomorfas.

Teorema 28.78 (Levantamiento de caminos y de homotopías)**.** *Sea $p\colon \widetilde{X} \to X$ un recubrimiento, $\gamma\colon [0,1] \to X$ un camino y $\widetilde{x}_0 \in p^{-1}(\gamma(0))$. Entonces existe un único camino $\widetilde{\gamma}\colon [0,1] \to \widetilde{X}$ con $p \circ \widetilde{\gamma} = \gamma$ y $\widetilde{\gamma}(0) = \widetilde{x}_0$. Además, si dos caminos γ_0, γ_1 con los mismos extremos son homótopos en X, sus levantamientos desde el mismo $\widetilde{x}_0$ terminan en el mismo punto y son homótopos en $\widetilde{X}$.*

Idea de la demostración. Cúbrase X por entornos uniformemente recubiertos; sus preimágenes por γ forman un recubrimiento abierto de $[0,1]$, que es compacto, luego admite un número de Lebesgue. Pártase $[0,1]$ en trozos $0 = t_0 < t_1 < \cdots < t_n = 1$ tan finos que $\gamma([t_k, t_{k+1}])$ quepa dentro de un entorno uniformemente recubierto U_k. Sobre el primer trozo hay una única hoja de $p^{-1}(U_0)$ que contiene a $\widetilde{x}_0$, y en ella se levanta usando el homeomorfismo inverso; el extremo obtenido sirve de punto de partida para el trozo siguiente. La unicidad en cada paso es la unicidad del homeomorfismo local, y por inducción se obtiene la unicidad global. Para homotopías se repite el argumento en el cuadrado $[0,1]^2$, que también es compacto. □

Ejemplo 28.79 (Levantar caminos y contar vueltas)**. (a) Una vuelta a la circunferencia.**

Con $p\colon \mathbb{R} \to S^1$ y $\gamma(s) = (\cos 2\pi s, \sin 2\pi s)$, el levantamiento que arranca en 0 es $\widetilde{\gamma}(s) = s$, y termina en 1. El levantamiento que arranca en 5 es $\widetilde{\gamma}(s) = 5 + s$ y termina en 6: el punto inicial cambia, pero la *diferencia* entre los extremos vale 1 siempre.

(b) El grado. Para el lazo que da n vueltas, $\gamma(s) = \mathrm{e}^{2\pi i n s}$, el levantamiento es $\widetilde{\gamma}(s) = ns$ y la diferencia final es n. Ese entero es el grado, y el teorema de levantamiento de homotopías garantiza que no cambia al deformar el lazo: es un invariante genuino, y es el isomorfismo $\pi_1(S^1) \cong \mathbb{Z}$.

(c) La unicidad es esencial. Fijado el punto inicial $\widetilde{x}_0$, el levantamiento es único. Sin fijarlo hay tantos como puntos en la fibra: sobre S^1, infinitos, uno por cada entero. Por eso todos los enunciados sobre recubrimientos llevan punto base.

(d) Una aplicación inmediata. No existe ninguna retracción continua $D^2 \to S^1$: si la hubiera, la identidad de S^1 factorizaría por el disco, y al contar vueltas saldría que el lazo de grado 1 es homótopo al constante, de grado 0. De ahí sale el teorema del punto fijo de Brouwer, y de él, por ejemplo, el teorema fundamental del álgebra.

Idea clave

Un recubrimiento es una fotocopiadora topológica: encima de cada trocito pequeño de X hay varias copias idénticas, apiladas y sin tocarse. Localmente $\widetilde{X}$ y X son indistinguibles — p es un homeomorfismo local —, pero globalmente pueden ser radicalmente distintos: $\mathbb{R}$ y S^1 son localmente el mismo espacio, y sin embargo uno tiene un agujero y el otro no. Toda la información sobre el agujero está en cómo se pegan las copias, no en las copias mismas.

Regla: leer un recubrimiento y extraer información de él

1. **Compruébese que p es un recubrimiento.** Hay que exhibir, para cada punto de la base, un entorno cuya preimagen se parta en abiertos disjuntos aplicados homeomórficamente. No basta con que p sea un homeomorfismo local: la inclusión $(0, 2) \to S^1$ dada por $t \mapsto e^{2\pi i t}$ lo es y no es recubrimiento, porque falla la sobreyectividad y falla el reparto en hojas cerca del punto 1.

2. **Cuéntense las hojas.** Tómese un punto cómodo x_0 de la base y determínese el cardinal de $p^{-1}(x_0)$. Si X es conexo, ese número vale para todos los puntos y es un invariante del recubrimiento.

3. **Levántese un lazo concreto.** Elíjase un lazo γ en x_0, levántese desde un punto $\widetilde{x}_0$ de la fibra y mírese dónde termina. Si termina en un punto distinto del de partida, el lazo no era trivial: se ha demostrado que $\pi_1(X, x_0) \neq 1$ sin manipular ninguna homotopía.

4. **Léase el morfismo inducido.** La aplicación $p_*\colon \pi_1(\widetilde{X}, \widetilde{x}_0) \to \pi_1(X, x_0)$ es siempre inyectiva, por el teorema de levantamiento de homotopías. Su imagen es un subgrupo $H \leq \pi_1(X, x_0)$, y el número de hojas coincide con el índice $[\pi_1(X, x_0) : H]$.

5. **Si $\widetilde{X}$ es simplemente conexo, cámbiese de estrategia.** Entonces $H = 1$, el índice es el orden del grupo y se concluye directamente $|\pi_1(X, x_0)|$ = número de hojas. Este es el atajo que calcula $\pi_1(S^1)$ y $\pi_1(\mathbb{RP}^2)$ en dos líneas.

6. **Identifíquense las transformaciones de cubierta.** Son los homeomorfismos φ de $\widetilde{X}$ con $p \circ \varphi = p$: mueven cada hoja a otra hoja sin alterar la base. Forman un grupo, y cuando el recubrimiento es *normal* ese grupo es isomorfo al cociente $\pi_1(X, x_0)/H$.

7. **Úsese el resultado para distinguir espacios.** Dos espacios con recubrimientos universales de distinto número de hojas sobre ellos tienen grupos fundamentales de distinto orden, luego no son homeomorfos. Es el mismo criterio del apartado anterior, escrito en el idioma de las hojas.

Ejemplo 28.80 (Cuatro recubrimientos, contados a mano)**. (a) La recta sobre el círculo:** $p(t) = (\cos 2\pi t,\ \sin 2\pi t)$**.** Es el ejemplo del que sale todo lo demás. La fibra sobre el punto $(1, 0)$ es $p^{-1}((1, 0)) = \mathbb{Z}$: los enteros, y solo ellos, dan ángulo múltiplo de 360°. Un entorno uniformemente recubierto del punto $(1, 0)$ es el arco de ángulo menor que 180° alrededor suyo, cuya preimagen es la unión disjunta $\bigsqcup_{n\in\mathbb{Z}}(n - \frac{1}{4}, n + \frac{1}{4})$. El lazo $w_k(s) = (\cos 2\pi ks, \sin 2\pi ks)$, que da k vueltas, se levanta a $\widetilde{w}_k(s) = ks$, que empieza en 0 y termina en k. Como $\mathbb{R}$ es simplemente conexo y el punto final del levantamiento es un invariante de homotopía, la asignación $[w_k] \mapsto k$ está bien definida y es un isomorfismo: $\pi_1(S^1) \cong \mathbb{Z}$.

(b) El círculo sobre sí mismo: $q(z) = z^n$**.** Aquí $\widetilde{X} = X = S^1$ visto dentro de $\mathbb{C}$ (véase el capítulo 4). Cada punto tiene exactamente n preimágenes, las n raíces n-ésimas, separadas por ángulos de $360°/n$: para $n = 2$, medio giro; para $n = 3$, 120°; para $n = 4$, 90°; para $n = 5$, 72°. El subgrupo imagen es $q_*(\mathbb{Z}) = n\mathbb{Z}$, de índice n en $\mathbb{Z}$: exactamente el número de hojas, como anuncia el paso 4. Y como los únicos subgrupos de $\mathbb{Z}$ de índice n son $n\mathbb{Z}$ y solo $n\mathbb{Z}$, *este es el único* recubrimiento conexo de n hojas del círculo, salvo isomorfismo.

(c) La esfera sobre el plano proyectivo. Sea $p\colon S^2 \to \mathbb{RP}^2$ la aplicación que identifica cada punto con su antípoda. Cada punto de $\mathbb{RP}^2$ tiene exactamente 2 preimágenes, $\{x, -x\}$, y un casquete abierto que no contenga puntos antipodales es uniformemente recubierto. Ahora bien, S^2 es simplemente conexa, luego el paso 5 se aplica sin más: $|\pi_1(\mathbb{RP}^2)| = 2$, es decir $\pi_1(\mathbb{RP}^2) \cong \mathbb{Z}/2\mathbb{Z}$. Consecuencia visual, y famosa: un lazo que recorre una vez la banda de identificación no se puede contraer, pero recorrido dos veces sí. Las transformaciones de cubierta son la identidad y la antipodal, y forman un grupo de orden 2, coherente con el paso 6.

(d) El plano sobre el toro. La aplicación $P(s, t) = \big((\cos 2\pi s, \sin 2\pi s), (\cos 2\pi t, \sin 2\pi t)\big)$ recubre $T^2 = S^1 \times S^1$ con fibra $\mathbb{Z} \times \mathbb{Z}$: infinitas hojas, una por cada par de enteros. Como $\mathbb{R}^2$ es simplemente conexo, $\pi_1(T^2) \cong \mathbb{Z}^2$, y el par de enteros que se lee en el punto final del levantamiento es literalmente "cuántas vueltas al agujero horizontal y cuántas al vertical" ha dado el lazo. Obsérvese cómo el enrejado de cuadrados unidad del plano se convierte, al pegar, en el cuadrado con lados identificados que ya se usó para construir el toro.

Ejemplo 28.81 (El recubrimiento universal y para qué sirve)**. (a) Qué es.** Un recubrimiento

$p\colon \widetilde{X} \to X$ con $\widetilde{X}$ conexo por caminos y simplemente conexo se llama *universal*. Existe siempre que X sea conexo por caminos, localmente conexo por caminos y semilocalmente simplemente conexo, y en ese caso es único salvo isomorfismo. Los cuatro ejemplos del recuadro anterior son universales excepto $z \mapsto z^n$, cuyo espacio de arriba es el círculo, que no es simplemente conexo.

(b) Por qué se llama universal. Porque domina a todos los demás: si $q\colon Y \to X$ es cualquier otro recubrimiento con Y conexo, existe un recubrimiento $r\colon \widetilde{X} \to Y$ tal que $q \circ r = p$. El universal está, literalmente, encima de todos. En el caso del círculo esto se ve de golpe: $\mathbb{R} \to S^1 \to S^1$, donde la primera flecha es $t \mapsto e^{2\pi i t}$ y la segunda $z \mapsto z^n$, se compone en $t \mapsto e^{2\pi i n t}$, que vuelve a ser el recubrimiento universal reparametrizado.

(c) La correspondencia de Galois topológica. Para X razonable y conexo, hay una biyección entre las clases de isomorfismo de recubrimientos conexos de X y las clases de conjugación de subgrupos de $\pi_1(X, x_0)$. Bajo ella, el número de hojas es el índice, el recubrimiento universal corresponde al subgrupo trivial, el propio X corresponde al grupo entero, y los recubrimientos normales corresponden a los subgrupos normales, con grupo de transformaciones de cubierta $\pi_1(X, x_0)/H$. La analogía con la teoría de Galois de cuerpos no es una casualidad de nombres: es la misma estructura reticular, y quien haya leído el capítulo 23 reconocerá el diagrama.

(d) Un cálculo que sale gratis. ¿Cuántos recubrimientos conexos de 3 hojas tiene el círculo? Tantos como subgrupos de índice 3 en $\mathbb{Z}$, es decir, exactamente uno: $3\mathbb{Z}$. ¿Y el toro? Tantos como subgrupos de índice 3 en $\mathbb{Z}^2$; hay cuatro, a saber los generados por las columnas de

$$\begin{pmatrix} 3 & 0 \\ 0 & 1 \end{pmatrix}, \quad \begin{pmatrix} 1 & 0 \\ 0 & 3 \end{pmatrix}, \quad \begin{pmatrix} 3 & 1 \\ 0 & 1 \end{pmatrix}, \quad \begin{pmatrix} 3 & 2 \\ 0 & 1 \end{pmatrix},$$

y como $\mathbb{Z}^2$ es abeliano todos son normales, así que hay cuatro recubrimientos triples del toro esencialmente distintos. Toda esta información combinatoria se ha obtenido sin dibujar ni un solo toro.

Cuidado

[Homeomorfismo local no es recubrimiento] La confusión más frecuente. La aplicación $f\colon (0,3) \to S^1$, $f(t) = e^{2\pi i t}$, es un homeomorfismo local y es sobreyectiva, y sin embargo *no* es un recubrimiento: la fibra sobre el punto 1 es $\{1, 2\}$, con dos elementos, mientras que la fibra sobre un punto de ángulo pequeño positivo tiene tres. Como la base es conexa, un recubrimiento debe tener fibras del mismo cardinal, luego f no puede serlo. Geométricamente el fallo se ve mejor: cerca del punto 1 de la base, la preimagen contiene un trocito de $(0,3)$ que *no* se aplica homeomórficamente, porque le falta el extremo. La moraleja: siempre hay que exhibir el entorno uniformemente recubierto, no basta con mirar la derivada o el carácter local de f.

Pausa para pensar

Obsérvese lo que se ha conseguido. El grupo fundamental es difícil de calcular directamente, porque exige razonar sobre clases de homotopía de lazos, que son objetos infinitos y escurridizos. Los recubrimientos convierten ese problema en un problema de *contar*: cuántas hojas hay encima de un punto. Y contar es algo que se sabe hacer. Este es un patrón que reaparece en toda la matemática avanzada: cuando una estructura es difícil de manejar, se busca un objeto auxiliar sobre el que actúe, y se estudia la acción en lugar de la estructura. En álgebra son las representaciones; aquí, los recubrimientos.

28.21. La característica de Euler y la clasificación de las superficies

Hay un número entero, asociado a cada superficie, que se calcula contando trozos y que no cambia aunque la superficie se deforme, se estire o se vuelva a triangular de otra manera. Ese número es la *característica de Euler*, y con él — más el dato de si la superficie tiene o no dos caras — queda determinada por completo cualquier superficie compacta sin borde. Es uno de los teoremas de clasificación más satisfactorios de toda la matemática: una lista finita de piezas y una regla para pegarlas agota todas las posibilidades.

Definición 28.82 (Triangulación y característica de Euler)**.** Una *triangulación* de una superficie compacta S es una descomposición de S en un número finito de triángulos curvilíneos que se cortan dos a dos, o bien en un vértice común, o bien en una arista común, o bien en nada. Si la triangulación tiene V vértices, A aristas y C caras, la *característica de Euler* de S es

$$\chi(S) = V - A + C.$$

El teorema básico afirma que este número no depende de la triangulación elegida y que es invariante por homeomorfismos: si $S \cong S'$, entonces $\chi(S) = \chi(S')$.

Ejemplo 28.83 (Triangulaciones y el número que no cambia)**. (a) El tetraedro.** Tiene $V = 4$, $A = 6$, $C = 4$, luego $\chi = 4 - 6 + 4 = 2$. Como es homeomorfo a la esfera, ese 2 es la característica de S^2.

(b) El cubo y el octaedro. El cubo da $8 - 12 + 6 = 2$ y el octaedro $6 - 12 + 8 = 2$: los tres poliedros tienen triangulaciones (o descomposiciones celulares) muy distintas y el mismo número. Ésa es la afirmación con contenido: χ no depende de la triangulación elegida, sólo del espacio.

(c) El toro. Su triangulación mínima usa $V = 7$, $A = 21$, $C = 14$, de donde $\chi = 7 - 21 + 14 = 0$. Y en general, para la superficie orientable de género g se cumple $\chi = 2 - 2g$: la esfera ($g = 0$) da 2, el toro ($g = 1$) da 0, el toro doble ($g = 2$) da -2.

(d) Una consecuencia elemental. De $\chi = 2$ y de que cada cara de un poliedro convexo tiene al menos tres aristas se deduce, con un par de desigualdades, que todo poliedro convexo tiene alguna cara con menos de seis lados. Ése es el ingrediente combinatorio del que arrancan todas las demostraciones del teorema de los cinco colores.

Teorema 28.84 (Clasificación de las superficies compactas)**.** *Toda superficie compacta, conexa y sin borde es homeomorfa exactamente a una de las siguientes:*

1. *la esfera S^2, con $\chi = 2$;*
2. *la suma conexa de $g \geq 1$ toros, con $\chi = 2 - 2g$ (superficies orientables de* género *g);*
3. *la suma conexa de $k \geq 1$ planos proyectivos, con $\chi = 2 - k$ (superficies no orientables).*

En consecuencia, el par formado por la característica de Euler y la orientabilidad determina la superficie salvo homeomorfismo.

Idea clave

La característica de Euler es lo que queda de una superficie cuando se le quita toda la geometría. No importa el tamaño, ni la curvatura, ni cómo esté sentada en el espacio: solo importa el balance entre trozos de dimensión par y trozos de dimensión impar. Que ese balance no cambie al refinar la triangulación es el milagro; una vez aceptado, contar vértices, aristas y caras se convierte en una manera de *medir agujeros* con los dedos.

Regla: calcular la característica de Euler y clasificar una superficie

1. **Constrúyase una descomposición cómoda.** No hace falta que sean triángulos: sirve cualquier descomposición en celdas, es decir, cualquier reparto de la superficie en trozos homeomorfos a discos, pegados por sus bordes. Cuantas menos celdas, menos cuentas.
2. **Cuéntense V, A y C con cuidado en las identificaciones.** Este es el paso donde se cometen los errores. Si la superficie viene dada como un polígono con lados identificados, dos lados que se pegan cuentan como *una* sola arista, y varios vértices del polígono que acaban en el mismo punto cuentan como *un* solo vértice.
3. **Aplíquese $\chi = V - A + C$.** El resultado debe ser un entero menor o igual que 2; si sale mayor que 2, hay un error de conteo seguro.
4. **Decídase la orientabilidad.** Recórrase la superficie con un pequeño reloj dibujado en ella. Si existe un camino cerrado que lo devuelve girando al revés, la superficie es no orientable; equivalentemente, si contiene una banda de Möbius, no es orientable.
5. **Léase la superficie en la tabla.** Si es orientable, el género es $g = (2 - \chi)/2$, que ha de salir entero. Si no lo es, el número de planos proyectivos es $k = 2 - \chi$.
6. **Para sumas conexas, úsese la fórmula aditiva.** Al pegar dos superficies quitando un disco a cada una,

$$\chi(S \# S') = \chi(S) + \chi(S') - 2,$$

porque se eliminan dos caras y se identifican los dos bordes circulares, lo cual no cambia el balance de vértices y aristas.

7. **Compruébese con un invariante independiente.** El grupo fundamental sirve de control: para el género g orientable, π_1 tiene $2g$ generadores; para la esfera es trivial; para el plano proyectivo es $\mathbb{Z}/2\mathbb{Z}$. Si el π_1 y la χ discrepan, revísese el conteo del paso 2.

Ejemplo 28.85 (Los cinco sólidos platónicos, y por qué todos dan 2)**. (a) La cuenta, uno por uno.** Todos los poliedros convexos son homeomorfos a la esfera, así que todos deben dar $\chi = 2$. Compruébese:

Sólido	V	A	C	$V - A + C$
Tetraedro	4	6	4	2
Cubo	8	12	6	2
Octaedro	6	12	8	2
Dodecaedro	20	30	12	2
Icosaedro	12	30	20	2

(b) De dónde salen esos números sin contar a mano. Si cada cara tiene m lados y en cada vértice concurren n aristas, entonces $mC = 2A$ (cada arista pertenece a dos caras) y $nV = 2A$ (cada arista tiene dos extremos). Para el dodecaedro, $m = 5$ y $n = 3$: de $5 \cdot 12 = 2A$ sale $A = 30$, y de $3V = 2 \cdot 30$ sale $V = 20$. Para el icosaedro, $m = 3$ y $n = 5$: $3 \cdot 20 = 2A$ da $A = 30$, y $5V = 60$ da $V = 12$. Obsérvese la dualidad, que intercambia V con C y deja A fija.

(c) Por qué solo hay cinco. Sustituyendo $C = 2A/m$ y $V = 2A/n$ en $V - A + C = 2$ se obtiene

$$\frac{2A}{n} - A + \frac{2A}{m} = 2, \qquad \text{es decir} \qquad \frac{1}{n} + \frac{1}{m} = \frac{1}{2} + \frac{1}{A} > \frac{1}{2}.$$

Con $m, n \geq 3$ enteros, la desigualdad $1/m + 1/n > 1/2$ solo admite las parejas $(3,3)$, $(3,4)$, $(4,3)$, $(3,5)$ y $(5,3)$. Cinco parejas, cinco sólidos. Este es un teorema de geometría clásica demostrado enteramente con una cuenta topológica.

(d) Un poliedro no convexo con $\chi \neq 2$. Tómese un cubo y perfórese de lado a lado con un agujero de sección cuadrada. El resultado es homeomorfo al toro, y en efecto cualquier descomposición en celdas de esa figura da $\chi = 0$. El teorema de Euler "$V - A + C = 2$", tal como se enuncia en la escuela, es un teorema sobre la *esfera*, no sobre los poliedros en general; el enunciado correcto es que χ depende solo del tipo topológico.

Ejemplo 28.86 (Cuatro superficies a partir del cuadrado)**.** Todas las superficies compactas se obtienen de un polígono identificando sus lados por parejas. Cuatro casos, con sus cuentas completas.

(a) La esfera. El polígono más cómodo aquí no es el cuadrado sino el biángulo: dos vértices, dos lados curvos, y el patrón aa^{-1}, que pega un lado con el otro tal cual. Es la construcción del globo terráqueo a partir de un gajo. Quedan $V = 2$ (los dos polos), $A = 1$ (el meridiano, pues los dos lados se han identificado) y $C = 1$, luego

$$\chi(S^2) = 2 - 1 + 1 = 2.$$

Todavía más económica es la descomposición en celdas mínima: un solo punto y un disco pegado por *todo* su borde a ese punto, es decir $V = 1$, $A = 0$, $C = 1$, que vuelve a dar $\chi = 2$. Las dos cuentas coinciden, como debe ser. Orientable, género 0.

(b) El toro. Patrón $aba^{-1}b^{-1}$: lados opuestos identificados conservando la orientación. Los *cuatro* vértices del cuadrado se pegan en un único punto, los cuatro lados en dos aristas a y b, y hay una cara. Así $V = 1$, $A = 2$, $C = 1$ y

$$\chi(T^2) = 1 - 2 + 1 = 0,$$

coherente con $2 - 2g$ para $g = 1$. Es orientable: el reloj vuelve siempre girando en el mismo sentido, porque las identificaciones no invierten ningún lado.

(c) La botella de Klein. Patrón $abab^{-1}$: como el toro, pero uno de los pares se pega con un giro. El conteo es idéntico — $V = 1$, $A = 2$, $C = 1$ —, luego también $\chi = 0$. Y sin embargo la botella de Klein *no* es el toro: no es orientable, contiene una banda de Möbius, y por el teorema de clasificación es la suma conexa de dos planos proyectivos, con $\chi = 2-k = 2-2 = 0$. Este par es la prueba de que la característica de Euler sola no basta: hace falta el dato de la orientabilidad.

(d) El plano proyectivo, y las sumas conexas. Patrón $abab$, que identifica puntos antipodales del borde: queda $V = 1$, $A = 1$, $C = 1$ y $\chi(\mathbb{RP}^2) = 1 - 1 + 1 = 1$. Con la fórmula aditiva del paso 6 ya se puede rellenar la tabla entera: $\chi(T\#T) = 0+0-2 = -2$, que es $2-2g$ con $g = 2$; $\chi(\mathbb{RP}^2\#\mathbb{RP}^2) = 1 + 1 - 2 = 0$, la botella de Klein; y $\chi(T\#\mathbb{RP}^2) = 0 + 1 - 2 = -1$, que es $2-k$ con $k = 3$, es decir la suma conexa de *tres* planos proyectivos. Esa última igualdad — toro más proyectivo igual a tres proyectivos — no es intuitiva en absoluto, y es una de las sorpresas del teorema de clasificación: en cuanto una superficie deja de ser orientable, los asas se convierten en gorros cruzados.

Cuidado

[La característica no distingue por sí sola] Se acaba de ver: toro y botella de Klein tienen ambos $\chi = 0$, y no son homeomorfos. Del mismo modo, $\chi = 2$ lo tiene la esfera pero también — si se admiten superficies con borde — ningún otro caso cerrado, mientras que $\chi = 1$ corresponde tanto al plano proyectivo como al disco cerrado, que sí tiene borde. Los tres datos que hay que fijar antes de aplicar el teorema de clasificación son: compacta, conexa y *sin borde*. Faltando cualquiera de ellos, la lista cambia. Con borde, por ejemplo, entran el cilindro ($\chi = 0$, orientable, dos componentes de borde) y la banda de Möbius ($\chi = 0$, no orientable, una sola componente de borde), que comparten característica y difieren en todo lo demás.

Pausa para pensar

Una consecuencia inesperada, para terminar. El grafo completo K_7 — siete vértices, todos unidos entre sí, es decir $7 \cdot 6/2 = 21$ aristas — no cabe en el plano ni en la esfera, pero sí en el toro, y al hacerlo divide el toro en 14 caras triangulares. Compruébese: $7 - 21 + 14 = 0 = \chi(T^2)$. Como los siete vértices son vecinos dos a dos, colorear ese mapa exige siete colores, y de ahí que el análogo del teorema de los cuatro colores en el

toro sea un teorema de *siete* colores. La característica de Euler, que empezó siendo un conteo de aristas de poliedros, acaba fijando cuántos colores hacen falta para pintar un mapa.

28.22. Taller de ejemplos explicados

28.22.1. Demostrar que un conjunto es abierto

Ejemplo 28.87. Demostrar que $(0,1)$ es abierto en $\mathbb{R}$.

Paso 1. Recordar qué hay que probar: que alrededor de cada punto de $(0,1)$ cabe un intervalito dentro de $(0,1)$.

Paso 2. Tomar un punto cualquiera $x \in (0,1)$. Está a cierta distancia positiva de ambos extremos; llamemos r a la menor de esas dos distancias, $r = \min(x, 1-x)$, que es positiva porque $0 < x < 1$.

Paso 3. La bolita $B(x,r) = (x-r, x+r)$ **cabe dentro:** por como elegimos r, este intervalo no llega ni al 0 ni al 1. Como todo punto tiene su bolita dentro, $(0,1)$ es abierto.

28.22.2. Demostrar que $(0,1)$ no es compacto

Ejemplo 28.88. Vamos a exhibir un cubrimiento abierto de $(0,1)$ del que *no* se puede sacar un subcubrimiento finito.

Paso 1. Proponer el cubrimiento. Consideremos los intervalos $U_n = \left(\frac{1}{n}, 1\right)$ para $n = 2, 3, 4, \ldots$

Paso 2. Verificar que cubren $(0,1)$**.** Cualquier punto $x \in (0,1)$ es positivo, así que para n suficientemente grande $\frac{1}{n} < x$, y entonces $x \in U_n$. Entre todos, los U_n cubren $(0,1)$.

Paso 3. Ver que no basta una cantidad finita. Si tomáramos solo finitos U_n, habría uno con el n más grande, digamos $U_N = (\frac{1}{N}, 1)$, y esa colección finita solo cubriría hasta $\frac{1}{N}$: los puntos entre 0 y $\frac{1}{N}$ quedarían fuera. Ningún subcubrimiento finito funciona.

Conclusión: $(0,1)$ no es compacto. El problema es el extremo 0, que "falta": por eso $[0,1]$ (que sí lo incluye) sí es compacto.

28.22.3. Un homeomorfismo explícito

Ejemplo 28.89. Comprobar que $f(x) = 2x$ es un homeomorfismo de $\mathbb{R}$ en $\mathbb{R}$.

Es biyectiva (cada y viene de un único $x = y/2$), continua (es una recta), y su inversa $f^{-1}(y) = y/2$ también es continua. Cumple las tres condiciones, así que es un homeomorfismo: estirar la recta al doble no cambia su topología, como era de esperar. (Un ejemplo trivial, pero muestra la lista de comprobaciones que se hace siempre: biyectiva, continua, inversa continua.)

28.22.4. Distinguir dos espacios con un invariante

Ejemplo 28.90. Demostrar que $[0,1]$ y $(0,1)$ *no* son homeomorfos.

Usamos un invariante: la *compacidad*. Ya vimos que $[0,1]$ es compacto (Heine–Borel: cerrado y acotado) y que $(0,1)$ no lo es. Como la compacidad se conserva bajo homeomorfismo (la

imagen continua de un compacto es compacta), si hubiera un homeomorfismo entre ellos, ambos serían compactos o ninguno lo sería. Pero uno lo es y el otro no: contradicción. Por lo tanto no son homeomorfos. Así es como la topología prueba que dos formas son "esencialmente distintas": encontrando una propiedad que una tiene y la otra no.

28.23. Entornos, redes y homotopía: los términos que faltaban

Tres palabras sostienen buena parte del capítulo y ninguna estaba definida. La primera es, de hecho, la más usada de todas.

Definición 28.91 (Entorno)**.** Sea X un espacio topológico y $x \in X$. Un conjunto $N \subset X$ es un *entorno* de x si existe un abierto U con

$$x \in U \subset N.$$

Si además N es abierto, se habla de *entorno abierto*. Una *base de entornos* de x es una familia $\mathcal{B}_x$ de entornos de x tal que todo entorno de x contiene a alguno de $\mathcal{B}_x$.

Cuidado

[El convenio importa, y no es universal] Con la definición de arriba —la de Bourbaki, y la que usa este libro— un entorno *no* tiene por qué ser abierto: en $\mathbb{R}$, el intervalo $[-1, 1]$ es entorno de 0. Otros textos reservan la palabra para los abiertos que contienen al punto. La diferencia parece cosmética y no lo es: con el convenio amplio, la familia de entornos de un punto es cerrada por superconjuntos y por intersecciones finitas —es un *filtro*—, lo que permite enunciar la convergencia con elegancia; con el estrecho, hay que añadir la palabra "contiene" en cada enunciado. Lo esencial es que las dos formulaciones dan la misma topología, porque *U es abierto si y sólo si es entorno de todos sus puntos*, y ésa es la caracterización que conviene retener.

Proposición 28.92 (Todo se puede decir con entornos)**.** *En un espacio topológico:*

1. *U es abierto $\iff$ U es entorno de cada uno de sus puntos;*
2. *$x \in \overline{A}$ $\iff$ todo entorno de x corta a A;*
3. *x es punto interior de A $\iff$ A es entorno de x;*
4. *f es continua en x $\iff$ la preimagen de todo entorno de $f(x)$ es entorno de x.*

Demostración. (1) Si U es abierto, sirve él mismo como el abierto intermedio. Recíprocamente, si para cada $x \in U$ hay un abierto U_x con $x \in U_x \subset U$, entonces $U = \bigcup_{x \in U} U_x$ es unión de abiertos. *(2)* $x \notin \overline{A}$ significa que hay un abierto que contiene a x y no corta a A, o sea un entorno disjunto de A. *(3)* y *(4)* son reformulaciones directas de las definiciones. □

Definición 28.93 (Conjunto dirigido y red)**.** Un *conjunto dirigido* es un conjunto D con una relación $\preceq$ reflexiva y transitiva tal que para todos $\alpha, \beta \in D$ existe γ con $\alpha \preceq \gamma$ y $\beta \preceq \gamma$ (dos elementos siempre tienen una cota superior).

Una *red* en X es una aplicación $D \to X$, escrita $(x_\alpha)_{\alpha \in D}$, con D dirigido. Se dice que *converge* a x, y se escribe $x_\alpha \to x$, si para todo entorno N de x existe $\alpha_0 \in D$ tal que $x_\alpha \in N$ para todo $\alpha \succeq \alpha_0$.

Observación 28.94 (Por qué las sucesiones no bastan)*.* Una sucesión es una red con $D = \mathbb{N}$. La diferencia es que $\mathbb{N}$ es *numerable*, y esa limitación se nota en cuanto un espacio no tiene bases numerables de entornos: en tales espacios puede haber un punto en $\overline{A}$ al que *ninguna* sucesión de A converge. Con redes eso no ocurre nunca, y los tres enunciados que en espacios métricos se dan con sucesiones vuelven a ser ciertos en general:

$$x \in \overline{A} \iff \exists\, \text{red en } A \text{ que converge a } x,$$

f es continua si y sólo si conserva límites de redes, y K es compacto si y sólo si toda red en K tiene una subred convergente en K. El conjunto dirigido que sirve en el caso general es, precisamente, la familia de entornos de x ordenada por inclusión inversa: se elige un punto de A en cada entorno, y esa red converge a x por construcción. Ahí está el papel de la definición 28.91.

Definición 28.95 (Homotopía y homotopía de caminos)**.** Dos aplicaciones continuas $f, g : X \to Y$ son *homótopas*, y se escribe $f \simeq g$, si existe una aplicación continua

$$H : X \times [0,1] \longrightarrow Y \qquad \text{con} \qquad H(x,0) = f(x), \quad H(x,1) = g(x)$$

para todo x. La aplicación H se llama *homotopía*, y conviene pensarla como una familia continua $h_t = H(\cdot, t)$ que deforma f en g.

Si f, g son *caminos* —es decir, $X = [0,1]$— con los mismos extremos, se exige además que la deformación mantenga los extremos fijos: $H(0,t) = f(0)$ y $H(1,t) = f(1)$ para todo t. Esa relación se llama *homotopía de caminos* y es la que se usa para definir el grupo fundamental.

Proposición 28.96. *La homotopía es una relación de equivalencia.*

Demostración. Reflexiva: $H(x,t) = f(x)$. *Simétrica*: si H va de f a g, entonces $H'(x,t) = H(x, 1-t)$ va de g a f. *Transitiva*: si H_1 va de f a g y H_2 de g a h, se concatenan,

$$H(x,t) = \begin{cases} H_1(x, 2t), & 0 \le t \le \tfrac{1}{2}, \\ H_2(x, 2t-1), & \tfrac{1}{2} \le t \le 1, \end{cases}$$

que es continua por el lema del pegado —los dos trozos son cerrados y coinciden en la intersección, donde ambos valen $g(x)$—. Con extremos fijos, las tres construcciones los respetan. □

Observación 28.97 (Y ahora el grupo fundamental tiene sentido)*.* Con esta definición, $\pi_1(X, x_0)$ es el conjunto de clases de homotopía de lazos basados en x_0, con el producto inducido por la concatenación. La proposición anterior es lo que garantiza que "clase de homotopía" signifique

algo; la buena definición del producto —que la clase del producto sólo depende de las clases de los factores— se comprueba concatenando las homotopías de cada mitad. Y el enunciado "X es simplemente conexo si todo lazo es homótopo a un punto" pasa a ser verificable, porque ya se sabe qué hay que exhibir: una H concreta.

28.24. Ejercicios resueltos

Los seis problemas siguientes recorren, en orden, los conceptos del capítulo: métrica, compacidad, homeomorfismos, separación y conexión. Conviene intentarlos antes de leer la solución.

Ejercicio 28.1. Sea X un conjunto con al menos dos puntos y d la métrica discreta: $d(x,y)=1$ si $x\neq y$ y $d(x,x)=0$. Describa las bolas abiertas, las bolas cerradas y la clausura de una bola abierta. Concluya que $\overline{B(x,r)}$ puede ser estrictamente menor que la bola cerrada $\bar B(x,r)$.

Solución. Si $r\le 1$, entonces $B(x,r)=\{y : d(x,y)<r\}=\{x\}$, porque la única distancia menor que 1 es 0. Si $r>1$, $B(x,r)=X$. Para las bolas cerradas: $\bar B(x,r)=\{x\}$ si $r<1$, y $\bar B(x,r)=X$ si $r\ge 1$.

Tomemos ahora $r=1$. La bola abierta es $B(x,1)=\{x\}$, que es cerrada (su complemento es unión de bolas $\{y\}$, todas abiertas), luego $\overline{B(x,1)}=\{x\}$. Pero la bola cerrada es $\bar B(x,1)=X$, que tiene al menos dos puntos. Por tanto

$$\overline{B(x,1)}=\{x\}\subsetneq X=\bar B(x,1).$$

La notación es traicionera: la barra de la clausura y la barra de la bola cerrada no significan lo mismo. En $\mathbb{R}^n$ con la métrica usual coinciden, y esa coincidencia es una propiedad de los espacios normados, no de los espacios métricos. ■

Ejercicio 28.2. Exhiba un recubrimiento abierto de $(0,1)$ sin subrecubrimiento finito, y otro de $[0,\infty)$. Explique en cada caso cuál de las dos hipótesis de Heine–Borel falla.

Solución. Para $(0,1)$ tomemos $U_n=\left(\frac{1}{n},1\right)$ con $n\ge 2$. Es un recubrimiento: dado $x\in(0,1)$, basta $n>1/x$. Pero cualquier subfamilia finita $U_{n_1},\dots,U_{n_k}$ tiene unión $\left(\frac{1}{N},1\right)$ con $N=\max n_i$, que no contiene a $\frac{1}{2N}$. Aquí falla que el conjunto sea *cerrado* en $\mathbb{R}$: le faltan los extremos, y el recubrimiento se cuela por el que falta.

Para $[0,\infty)$ tomemos $V_n=(-1,n)$. Cubre, y ninguna subfamilia finita cubre, por el mismo argumento con $\max n_i$. Aquí falla que el conjunto sea *acotado*.

Los dos ejemplos juntos muestran que ninguna de las dos hipótesis de Heine–Borel sobra. Y conviene recordar que la equivalencia «compacto $\iff$ cerrado y acotado» es una propiedad de $\mathbb{R}^n$, no de los espacios métricos en general: en un espacio de Hilbert de dimensión infinita la bola unidad cerrada es cerrada y acotada, y no es compacta. ■

Ejercicio 28.3. Sea $f:[0,2\pi)\to S^1$ definida por $f(t)=(\cos t,\sin t)$. Demuestre que f es una biyección continua pero no un homeomorfismo. ¿Por qué no contradice el teorema «biyección continua de un compacto en un Hausdorff es homeomorfismo»?

Solución. Que f es biyectiva y continua es inmediato. Para ver que f^{-1} no es continua, considere la sucesión $z_n=f\big(2\pi-\frac{1}{n}\big)$. Geométricamente $z_n\to(1,0)=f(0)$, pero

$$f^{-1}(z_n)=2\pi-\tfrac{1}{n}\longrightarrow 2\pi\neq 0=f^{-1}\big((1,0)\big).$$

Luego f^{-1} no conserva límites y no es continua.

Otra demostración, más topológica: $[0,2\pi)$ menos el punto 0 es conexo, mientras que S^1 menos el punto $f(0)$ también lo es — eso no sirve. Use mejor la compacidad: S^1 es compacto y $[0,2\pi)$ no lo

es (el recubrimiento $\left[0, 2\pi - \frac{1}{n}\right)$ no admite subrecubrimiento finito). Un homeomorfismo preserva la compacidad, luego no puede haberlo.

Y ahí está la respuesta a la pregunta: el teorema exige que el *dominio* sea compacto, y $[0, 2\pi)$ no lo es. Si se cambia el dominio a $[0, 2\pi]$, la aplicación deja de ser inyectiva. No hay forma de arreglarlo, y esa imposibilidad es precisamente lo que hace del círculo un objeto distinto de un segmento. ■

Ejercicio 28.4. Demuestre que un espacio topológico X es de Hausdorff si y solo si la diagonal $\Delta = \{(x, x) : x \in X\}$ es cerrada en $X \times X$ con la topología producto.

Solución. ($\Rightarrow$) Supongamos X Hausdorff y tomemos $(x, y) \notin \Delta$, es decir $x \neq y$. Hay abiertos disjuntos $U \ni x$ y $V \ni y$. Entonces $U \times V$ es un abierto básico del producto que contiene a (x, y) y que no corta a Δ: si $(z, z) \in U \times V$, entonces $z \in U \cap V = \varnothing$, absurdo. Luego Δ^c es abierto.

($\Leftarrow$) Supongamos Δ cerrada y sean $x \neq y$. Como $(x, y) \in \Delta^c$, que es abierto, hay un abierto básico $U \times V$ con $(x, y) \in U \times V \subseteq \Delta^c$. Entonces $x \in U$, $y \in V$, y $U \cap V = \varnothing$: si hubiera z en la intersección, (z, z) estaría en $U \times V$ y en Δ. Luego X es Hausdorff.

Este criterio es más útil de lo que parece: convierte una propiedad de separación (un enunciado con dos cuantificadores existenciales) en una propiedad de cierre de un solo conjunto. Es la vía habitual para demostrar que el conjunto donde dos funciones continuas coinciden es cerrado cuando el codominio es Hausdorff. ■

Ejercicio 28.5. Sea X un conjunto infinito con la topología cofinita: los abiertos son $\varnothing$ y los complementos de conjuntos finitos. Demuestre que X es compacto, conexo y T_1, pero no es de Hausdorff.

Solución. Compacto. Sea $\{U_i\}$ un recubrimiento abierto. Elija cualquier $U_{i_0} \neq \varnothing$. Su complemento es finito, digamos $\{x_1, \dots, x_k\}$, y cada x_j está en algún U_{i_j}. Entonces $U_{i_0}, U_{i_1}, \dots, U_{i_k}$ cubren X: un subrecubrimiento de a lo sumo $k+1$ elementos. Nótese que ni siquiera hizo falta mirar el recubrimiento.

Conexo. Dos abiertos no vacíos tienen complementos finitos, luego su intersección tiene complemento finito y en particular no es vacía (porque X es infinito). Así que nunca hay dos abiertos no vacíos disjuntos, y menos aún una partición.

T_1. Cada punto $\{x\}$ es finito, luego cerrado, luego su complemento es abierto. Dados $x \neq y$, el abierto $X \setminus \{y\}$ contiene a x y no a y: eso es exactamente T_1.

No Hausdorff. Es lo mismo que ya se probó al ver la conexión: dos abiertos no vacíos siempre se cortan, así que jamás se pueden separar dos puntos. Este espacio muestra que T_1 es estrictamente más débil que T_2, y también que un compacto no tiene por qué ser Hausdorff —motivo por el cual muchos textos incluyen «Hausdorff» dentro de la definición de compacto. ■

Ejercicio 28.6. Sea $f : S^1 \to \mathbb{R}$ continua. Demuestre que existe $x \in S^1$ con $f(x) = f(-x)$, donde $-x$ denota el punto antipodal.

Solución. Defina $g : S^1 \to \mathbb{R}$ por $g(x) = f(x) - f(-x)$, que es continua por serlo f y la aplicación antipodal. Observe la antisimetría:

$$g(-x) = f(-x) - f(x) = -g(x).$$

Fije cualquier $x_0 \in S^1$. Si $g(x_0) = 0$ hemos terminado. Si no, $g(x_0)$ y $g(-x_0) = -g(x_0)$ tienen signos opuestos. Tome un camino continuo $\gamma : [0, 1] \to S^1$ con $\gamma(0) = x_0$ y $\gamma(1) = -x_0$ (existe, porque S^1 es conexo por caminos: basta recorrer media circunferencia). Entonces $g \circ \gamma : [0, 1] \to \mathbb{R}$ es continua y cambia de signo entre sus extremos, así que por el teorema del valor intermedio existe t^* con $g(\gamma(t^*)) = 0$, es decir

$$f(\gamma(t^*)) = f(-\gamma(t^*)).$$

Este es el teorema de Borsuk–Ulam en dimensión uno. Su versión en dimensión dos dice que en todo instante hay dos puntos antipodales de la Tierra con la misma temperatura y la misma presión simultáneamente; su demostración ya no cabe en un párrafo y necesita el grupo fundamental de $\mathbb{RP}^2$.■

28.25. Ejercicios propuestos

Sin solución desarrollada: solo enunciado y, al final, la respuesta. En esta parte del libro muchas respuestas son una frase ("sí, y la razón es…") más que un número; comprueba que la tuya coincide en el fondo, no en la letra.

28.1. En $\mathbb{R}$ con la métrica usual, ¿es $A = (0, 1]$ abierto, cerrado, ambos o ninguno? Da su interior, su clausura y su frontera.

28.2. Demuestra que en cualquier espacio métrico la bola abierta $B(x, r)$ es un conjunto abierto.

28.3. Sea X un conjunto con la topología discreta. ¿Qué subconjuntos son abiertos? ¿Cuáles son compactos?

28.4. ¿Es $\mathbb{Z} \subset \mathbb{R}$ cerrado? ¿Es compacto? Justifica ambas respuestas por separado.

28.5. Da un cubrimiento abierto de $(0, 1)$ que no admita subcubrimiento finito.

28.6. Sean τ_1 la topología usual de $\mathbb{R}$ y τ_2 la de los complementos finitos. ¿Cuál es más fina? ¿Es la identidad $(\mathbb{R}, \tau_1) \to (\mathbb{R}, \tau_2)$ continua?

28.7. Prueba que la imagen continua de un conjunto conexo es conexa.

28.8. ¿Son homeomorfos $[0, 1]$ y $(0, 1)$? Da un invariante topológico que los distinga.

28.9. ¿Son homeomorfos la circunferencia S^1 y el intervalo $[0, 1]$? Usa el argumento del punto que se quita.

28.10. Sea $f : X \to \mathbb{R}$ continua con X compacto y no vacío. Demuestra que f alcanza su máximo.

28.11. ¿Es $\mathbb{Q} \subset \mathbb{R}$ conexo? Describe sus componentes conexas.

28.12. En $\mathbb{R}^2$, ¿es el conjunto $\{(x, y) : xy = 1\}$ conexo? ¿Y cerrado? ¿Y compacto?

Respuestas

1. Ni abierto ni cerrado; interior $(0, 1)$, clausura $[0, 1]$, frontera $\{0, 1\}$. **2.** Si $y \in B(x, r)$, toma $s = r - d(x, y) > 0$; la desigualdad triangular da $B(y, s) \subset B(x, r)$. **3.** Todos los subconjuntos son abiertos; los compactos son exactamente los finitos. **4.** Cerrado sí (su complemento es unión de intervalos abiertos); compacto no (no es acotado). **5.** $\{(1/n, 1)\}_{n\geq 2}$: cualquier subfamilia finita deja fuera puntos cercanos a 0. **6.** τ_1 es más fina. Sí, la identidad $(\mathbb{R}, \tau_1) \to (\mathbb{R}, \tau_2)$ es continua (ir hacia la topología más gruesa siempre lo es); la inversa no. **7.** Si $f(C) = U \sqcup V$ con U, V abiertos disjuntos no vacíos, entonces $f^{-1}(U), f^{-1}(V)$ desconectan C. **8.** No. $[0, 1]$ es compacto y $(0, 1)$ no; la compacidad es un invariante topológico. **9.** No. Quitar un punto interior a $[0, 1]$ lo desconecta; quitar cualquier punto a S^1 lo deja conexo. **10.** $f(X)$ es compacto en $\mathbb{R}$, luego cerrado y acotado; su supremo pertenece a $f(X)$. **11.** No es conexo. Sus componentes conexas son los puntos: $\mathbb{Q}$ es totalmente disconexo. **12.** Cerrado sí; compacto no (no acotado); conexo no: son dos ramas de hipérbola en cuadrantes opuestos, dos componentes.

Capítulo 29

Análisis real: los cimientos y las patologías

Lo que necesitas traer

Del capítulo 6: la construcción de $\mathbb{R}$, el supremo, sucesiones, límites, continuidad y derivada. Del capítulo 2: numerabilidad y el argumento diagonal. Del capítulo 28: abiertos, cerrados y compacidad, que aquí se usan como lenguaje.

29.1. Qué falta todavía

El capítulo 6 demostró los teoremas del cálculo en una variable, y el capítulo 28 extrajo de ellos la noción abstracta de cercanía. Entre los dos queda un hueco: las preguntas que uno se hace *después* de creer que ya entendió el cálculo, y cuya respuesta cambia la manera de pensar.

¿Cuántos números reales hay, comparados con cuántos racionales? ¿Existe una función continua en todo punto y derivable en ninguno? ¿Se puede decir con exactitud qué funciones son integrables Riemann, no con una condición suficiente sino con una equivalencia? ¿Puede una función ser continua exactamente en los irracionales? ¿Y exactamente en los racionales? ¿Si reordeno los términos de una serie convergente, cambia la suma?

Todas tienen respuesta, y todas las respuestas son sorprendentes. Este capítulo las demuestra. Es el capítulo de las patologías: los ejemplos que rompieron la intuición del siglo XIX y obligaron a construir el análisis moderno — y, con él, la integral de Lebesgue del capítulo 31.

Idea clave

Intuición primero. Hasta ahora las funciones que has visto se portan bien: son continuas salvo en unos pocos puntos, derivables salvo en unas pocas esquinas, integrables sin discusión. Eso no es porque las funciones sean así, sino porque son las que uno *escribe*. Las funciones que uno escribe con fórmulas cortas son una minoría diminuta y educada dentro del universo de todas las funciones. Este capítulo sale de esa minoría a propósito. La moraleja no es que el análisis esté roto, sino al revés: los teoremas que sobreviven

al contacto con las patologías son los teoremas de verdad, y los que no sobreviven necesitaban hipótesis que nadie había notado.

29.2. Los cimientos elementales del análisis real, definidos desde cero

Este capítulo va a hablar de patologías: funciones continuas sin derivada, conjuntos no numerables de longitud nula, series que suman lo que uno quiera según cómo se ordenen. Nada de eso tiene sentido si antes no está escrito con precisión qué es un supremo, qué es converger, qué es ser continuo y qué es integrar. Esta sección escribe esas definiciones una por una, con un ejemplo resuelto entero detrás de cada una y con las demostraciones asequibles hechas. Quien venga de otro capítulo puede saltársela; quien llegue de cero debe leerla primero, porque todo lo que sigue se apoya en ella y en ningún sitio más.

Idea clave

Intuición primero. El análisis real es el cálculo hecho con cuidado. En el cálculo se dice "x se acerca a 3" y se dibuja una flecha; en el análisis eso se sustituye por una frase con dos números, ε y δ, que se pueden verificar. La razón no es el gusto por la pedantería: es que la flecha miente en cuanto las funciones dejan de ser amables, y este capítulo está dedicado precisamente a las funciones que no lo son. La herramienta que hace posible todo el edificio es una sola: el axioma del supremo, que dice que en la recta real no hay agujeros. De él salen, uno detrás de otro, el teorema del valor intermedio, el de Bolzano–Weierstrass, la continuidad uniforme en compactos y la existencia de la integral.

Cotas, supremo e ínfimo: el axioma que distingue $\mathbb{R}$ de $\mathbb{Q}$

Definición 29.1 (Cota, supremo, máximo)**.** Sea $A \subseteq \mathbb{R}$ no vacío. Un número M es *cota superior* de A si $a \le M$ para todo $a \in A$; el conjunto se llama entonces *acotado superiormente*. Un número s es el *supremo* de A, y se escribe $s = \sup A$, si cumple las dos condiciones siguientes:

(i) $a \le s$ para todo $a \in A$ —es decir, s es cota superior—;

(ii) para todo $\varepsilon > 0$ existe $a \in A$ con $a > s - \varepsilon$ —es decir, ninguna cota superior es menor que s—.

Si además $s \in A$, se llama *máximo* de A. Simétricamente, m es *cota inferior* si $m \le a$ para todo $a \in A$, y el *ínfimo* $\inf A$ es la mayor de las cotas inferiores: cumple $\inf A \le a$ para todo a, y para todo $\varepsilon > 0$ hay algún $a \in A$ con $a < \inf A + \varepsilon$. Si $\inf A \in A$, se llama *mínimo*.

Definición 29.2 (Axioma del supremo)**.** Todo subconjunto de $\mathbb{R}$ no vacío y acotado superiormente tiene supremo *en* $\mathbb{R}$. Esta propiedad, llamada *completitud* de la recta, es el único axioma que separa $\mathbb{R}$ de $\mathbb{Q}$: todo lo demás (suma, producto, orden) lo cumplen los dos.

Ejemplo 29.3 (Dos supremos calculados con la definición, paso a paso)**. (a)** $A = \left\{ \dfrac{n}{n+1} : n \in \mathbb{N} \right\} = \left\{ \frac{1}{2}, \frac{2}{3}, \frac{3}{4}, \dots \right\}$. Afirmamos que $\sup A = 1$ y que $1 \notin A$.

Condición (i). Para todo $n \geq 1$ se tiene $\frac{n}{n+1} = 1 - \frac{1}{n+1} < 1$, luego 1 es cota superior.

Condición (ii). Sea $\varepsilon > 0$. Queremos n con $1 - \frac{1}{n+1} > 1 - \varepsilon$, es decir $\frac{1}{n+1} < \varepsilon$, es decir $n > \frac{1}{\varepsilon} - 1$. Como los naturales no están acotados, ese n existe. Con $\varepsilon = 0{,}01$ sirve $n = 100$, que da $\frac{100}{101} = 0{,}990099 > 0{,}99$.

Las dos condiciones se cumplen, luego $\sup A = 1$. Y $1 \notin A$ porque $\frac{n}{n+1} = 1$ exigiría $n = n+1$. Hay supremo pero no máximo. En cambio $\inf A = \frac{1}{2}$, que sí pertenece a A: es mínimo.

(b) $C = \{q \in \mathbb{Q} : q > 0,\ q^2 < 2\}$. Dentro de $\mathbb{R}$ este conjunto tiene supremo, y vale $\sqrt{2} = 1{,}414214\ldots$ Dentro de $\mathbb{Q}$, en cambio, *no* tiene supremo: si $s \in \mathbb{Q}$ fuera la menor cota superior racional, se comprueba que ni $s^2 < 2$ (se podría subir un poco), ni $s^2 > 2$ (se podría bajar un poco), ni $s^2 = 2$ (no hay racional con cuadrado 2) son posibles. Este ejemplo es la razón entera de que $\mathbb{R}$ exista, y la sección de reglas sobre supremos que viene más abajo en este mismo capítulo lo explota sistemáticamente.

Cuidado

Supremo y máximo no son lo mismo, y confundirlos es el error más frecuente de todo el capítulo. La condición (ii) pide un elemento *mayor que* $s - \varepsilon$, no un elemento *igual a* s. El supremo de $(0, 1)$ es 1 y ese conjunto no tiene máximo.

Sucesiones, límites y el criterio de Cauchy

Definición 29.4 (Sucesión y convergencia)**.** Una *sucesión* de números reales es una función $a : \mathbb{N} \to \mathbb{R}$, que se escribe $(a_n)_{n \geq 1}$. Se dice que (a_n) *converge* a $L \in \mathbb{R}$, y se escribe $a_n \to L$ o $\lim_{n \to \infty} a_n = L$, si

$$\text{para todo } \varepsilon > 0 \text{ existe } N \in \mathbb{N} \text{ tal que} \quad n \geq N \implies |a_n - L| < \varepsilon.$$

Una sucesión que no converge a ningún real se llama *divergente*. Y (a_n) está *acotada* si existe M con $|a_n| \leq M$ para todo n.

Ejemplo 29.5 (Un límite demostrado entero)**.** Sea $a_n = \dfrac{3n+1}{n+2}$. Afirmamos que $a_n \to 3$.

Paso 1: el candidato y el error. Se calcula la diferencia exacta,

$$3 - a_n = 3 - \frac{3n+1}{n+2} = \frac{3(n+2) - (3n+1)}{n+2} = \frac{5}{n+2},$$

que es positiva, así que $|a_n - 3| = \frac{5}{n+2}$.

Paso 2: resolver la desigualdad. Queremos $\frac{5}{n+2} < \varepsilon$, es decir $n > \frac{5}{\varepsilon} - 2$.

Paso 3: elegir N y cerrar. Tomamos cualquier natural $N > \frac{5}{\varepsilon} - 2$. Para $n \geq N$ se tiene $|a_n - 3| = \frac{5}{n+2} < \varepsilon$, que es exactamente lo que pedía la definición.

Paso 4: comprobación numérica. Con $\varepsilon = 0{,}01$ la cuenta da $n > 498$, luego $N = 499$. En efecto $a_{499} = 2{,}990020$, con error $0{,}009980 < 0{,}01$; y $a_{498} = 2{,}990000$, con error exactamente $0{,}01$, que ya no es *menor* que ε. La frontera está donde el álgebra decía.

Regla: demostrar un límite con la definición de épsilon y ene

Casi todas las demostraciones de límite tienen la misma forma, y conviene ejecutarlas siempre en este orden.

1. **Adivina el candidato L.** Con las reglas usuales del cálculo: divide por la potencia dominante, simplifica, mira quién manda. Este paso es heurístico y no forma parte de la demostración.
2. **Calcula $|a_n - L|$ exactamente.** No lo estimes todavía: haz el álgebra y deja una expresión cerrada. Aquí es donde se ganan o se pierden las demostraciones.
3. **Acota por arriba con algo sencillo y decreciente.** Sustituye la expresión por otra mayor pero manejable, del tipo $\frac{C}{n}$. Toda acotación intermedia vale con tal de que vaya en la dirección correcta.
4. **Despeja n en la desigualdad "cota $< \varepsilon$".** Obtienes una condición del tipo $n > g(\varepsilon)$.
5. **Declara N y escribe la implicación.** "Sea N cualquier natural mayor que $g(\varepsilon)$; entonces $n \geq N$ implica $|a_n - L| < \varepsilon$." Ahí termina la demostración.
6. **No inviertas los cuantificadores.** Primero se da ε, después se elige N. Un N que dependa de n, o un ε que dependa de N, es una demostración sin valor.
7. **Control numérico.** Sustituye un ε concreto y comprueba con una calculadora que el N que has obtenido funciona y que el anterior no. Cuesta un minuto y detecta casi todos los errores de álgebra.

Teorema 29.6 (Unicidad del límite)**.** *Una sucesión no puede converger a dos números distintos.*

Demostración. Supongamos $a_n \to L$ y $a_n \to L'$ con $L \neq L'$, y llamemos $d = |L - L'| > 0$. Aplicamos la definición con $\varepsilon = d/2$, que es un número positivo legítimo. Existe N_1 tal que $|a_n - L| < d/2$ para $n \geq N_1$, y existe N_2 tal que $|a_n - L'| < d/2$ para $n \geq N_2$. Tomando $n \geq \max(N_1, N_2)$, la desigualdad triangular da

$$d = |L - L'| \leq |L - a_n| + |a_n - L'| < \frac{d}{2} + \frac{d}{2} = d,$$

es decir $d < d$, que es absurdo. Luego $L = L'$. □

Definición 29.7 (Sucesión de Cauchy)**.** Una sucesión (a_n) es *de Cauchy* si

$$\text{para todo } \varepsilon > 0 \text{ existe } N \text{ tal que} \quad n, m \geq N \implies |a_n - a_m| < \varepsilon.$$

La diferencia con la definición 29.4 es que aquí no aparece el límite: la condición se comprueba mirando solo los términos de la sucesión entre sí.

Teorema 29.8 (Toda sucesión convergente es de Cauchy)**.** *Si $a_n \to L$, entonces (a_n) es de Cauchy. Además toda sucesión de Cauchy está acotada.*

Demostración. Sea $\varepsilon > 0$. Por convergencia existe N con $|a_n - L| < \varepsilon/2$ para todo $n \geq N$. Entonces, para $n, m \geq N$,

$$|a_n - a_m| \leq |a_n - L| + |L - a_m| < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon.$$

El truco de partir ε en dos mitades y pasar por el punto intermedio L es el argumento más usado de todo el análisis.

Para la acotación: tomando $\varepsilon = 1$ en la definición de Cauchy hay un N con $|a_n - a_N| < 1$ para $n \geq N$, luego $|a_n| < |a_N| + 1$ a partir de ese índice; los términos anteriores son finitos en número, así que $M = \max\{|a_1|, \ldots, |a_{N-1}|, |a_N| + 1\}$ acota a toda la sucesión. □

Observación 29.9. El recíproco —toda sucesión de Cauchy converge— es *verdadero en $\mathbb{R}$ y falso en $\mathbb{Q}$*, y es exactamente otra forma de enunciar el axioma del supremo. En $\mathbb{Q}$, las truncaciones decimales $1;\ 1{,}4;\ 1{,}41;\ 1{,}414;\ldots$ de $\sqrt{2}$ forman una sucesión de Cauchy sin límite racional. Por eso el criterio de Cauchy es tan valioso: permite demostrar que algo converge *sin conocer el límite*, y esa es la única vía cuando el límite es justamente lo que se quiere construir.

Ejemplo 29.10 (Una sucesión de Cauchy sin fórmula para el límite)**.** Sea $a_n = \sum_{k=1}^{n} \frac{1}{k^2}$. Para $m > n$,

$$|a_m - a_n| = \sum_{k=n+1}^{m} \frac{1}{k^2} < \sum_{k=n+1}^{m} \frac{1}{k(k-1)} = \sum_{k=n+1}^{m} \Big(\frac{1}{k-1} - \frac{1}{k}\Big) = \frac{1}{n} - \frac{1}{m} < \frac{1}{n},$$

así que basta tomar $N > 1/\varepsilon$. Es de Cauchy, luego converge. Numéricamente, la cola a partir de $n = 100$ vale $0{,}009940 < 0{,}01$, coherente con la cota $1/n$. El límite es $\pi^2/6 = 1{,}644934$, pero eso es un teorema muy posterior (Euler, 1735): la convergencia se sabe antes y con menos.

Subsucesiones, Bolzano–Weierstrass y los límites superior e inferior

Definición 29.11 (Subsucesión)**.** Dada (a_n) y una aplicación estrictamente creciente $k \mapsto n_k$ de $\mathbb{N}$ en $\mathbb{N}$, la sucesión $(a_{n_k})_{k \geq 1}$ es una *subsucesión* de (a_n). Es decir: se conserva el orden y se tachan términos, pero quedan infinitos. Si $a_n \to L$, toda subsucesión converge también a L; el recíproco es falso.

Teorema 29.12 (Bolzano–Weierstrass)**.** *Toda sucesión acotada de números reales tiene una subsucesión convergente.*

Demostración. Sea $|a_n| \leq M$ para todo n, de modo que todos los términos viven en $I_0 = [-M, M]$. Partimos I_0 por la mitad: al menos una de las dos mitades contiene a_n para infinitos índices n (si ambas contuviesen solo finitos, la sucesión tendría finitos términos). Llamamos I_1 a una mitad con infinitos índices y repetimos: obtenemos intervalos encajados

$$I_0 \supseteq I_1 \supseteq I_2 \supseteq \cdots, \qquad \text{longitud}(I_j) = \frac{2M}{2^j},$$

cada uno con infinitos índices dentro. Ahora construimos la subsucesión: elegimos n_1 con $a_{n_1} \in I_1$; elegido n_k, como I_{k+1} contiene infinitos índices, existe $n_{k+1} > n_k$ con $a_{n_{k+1}} \in I_{k+1}$. La aplicación $k \mapsto n_k$ es estrictamente creciente por construcción.

Falta ver que (a_{n_k}) converge. Sea c_j el extremo izquierdo de I_j: la sucesión (c_j) es creciente y acotada por M, luego por el axioma del supremo tiene supremo $L = \sup_j c_j$. Como a_{n_k} y L están ambos en I_k, cuya longitud es $2M \cdot 2^{-k}$, se tiene $|a_{n_k} - L| \le 2M \cdot 2^{-k}$, que tiende a 0. Luego $a_{n_k} \to L$. □

Ejemplo 29.13 (Una sucesión que no converge y sí tiene subsucesiones que sí)**.** Sea $a_n = (-1)^n \left(1 - \frac{1}{n}\right)$. No converge: los términos pares valen $1 - \frac{1}{n}$ y se acercan a 1 (0,9 para $n = 10$, 0,99 para $n = 100$, 0,999 para $n = 1000$), mientras que los impares se acercan a -1. Pero está acotada por 1, y Bolzano–Weierstrass promete al menos una subsucesión convergente: hay dos evidentes, la de índices pares (que tiende a 1) y la de índices impares (que tiende a -1).

Definición 29.14 (Límite superior e inferior)**.** Para (a_n) acotada se definen

$$\limsup_{n\to\infty} a_n = \lim_{n\to\infty} \Big(\sup_{k\ge n} a_k\Big), \qquad \liminf_{n\to\infty} a_n = \lim_{n\to\infty} \Big(\inf_{k\ge n} a_k\Big).$$

Los dos límites existen siempre: $s_n = \sup_{k\ge n} a_k$ es decreciente y acotada, e $i_n = \inf_{k\ge n} a_k$ es creciente y acotada. Siempre vale $\liminf a_n \le \limsup a_n$, y la sucesión converge *si y solo si* ambos coinciden, en cuyo caso su valor común es el límite.

Ejemplo 29.15 (Los dos límites, calculados)**.** Con $a_n = (-1)^n + \frac{1}{n}$: para n grande, $\sup_{k\ge n} a_k$ vale poco más de 1 (lo alcanza el primer índice par $\ge n$) y $\inf_{k\ge n} a_k$ vale poco menos de -1. Numéricamente, mirando la cola desde $n = 1000$ se obtiene supremo 1,000998 e ínfimo $-0{,}999500$. En el límite, $\limsup a_n = 1$ y $\liminf a_n = -1$: como no coinciden, la sucesión no converge, y esos dos números son exactamente los límites de las dos subsucesiones que Bolzano–Weierstrass garantizaba.

Series

Definición 29.16 (Serie y convergencia de series)**.** Dada una sucesión (a_n), su *serie* asociada es la sucesión de *sumas parciales* $S_N = \sum_{n=1}^{N} a_n$. Se dice que la serie $\sum_{n\ge1} a_n$ *converge* con suma S si $S_N \to S$ en el sentido de la definición 29.4; en caso contrario *diverge*. Una serie no es una suma infinita: es el límite de una sucesión de sumas finitas, y por eso el orden de los términos puede importar, como demuestra el teorema de Riemann más adelante en este capítulo.

Ejemplo 29.17 (Tres series decididas con la definición)**. (a) Geométrica.** Para $|r| < 1$, $S_N = \sum_{n=0}^{N} r^n = \frac{1-r^{N+1}}{1-r} \to \frac{1}{1-r}$, porque $r^{N+1} \to 0$. Con $r = \frac{1}{3}$ la suma vale $\frac{3}{2}$, y con $r = \frac{1}{2}$ empezando en $n = 1$ vale 1. Si $|r| \ge 1$ el término general no tiende a 0 y la serie diverge.

(b) Telescópica. $\sum_{n\ge1} \frac{1}{n(n+1)}$ tiene $S_N = \sum_{n=1}^{N} \left(\frac{1}{n} - \frac{1}{n+1}\right) = 1 - \frac{1}{N+1} \to 1$. La suma es exactamente 1, y aquí se ve la definición funcionando: la serie converge porque *la sucesión de sumas parciales* converge.

(c) Armónica. $\sum_{n\geq 1}\frac{1}{n}$ diverge, aunque $\frac{1}{n}\to 0$. La demostración agrupa: $\frac{1}{3}+\frac{1}{4}>\frac{1}{2}$, $\frac{1}{5}+\cdots+\frac{1}{8}>\frac{1}{2}$, y en general el bloque de 2^{k-1} términos que termina en 2^k suma más de $\frac{1}{2}$; luego $S_{2^k}>1+\frac{k}{2}$. Numéricamente $S_{10}=2{,}928968$, $S_{100}=5{,}187378$, $S_{1000}=7{,}485471$: crece como $\ln N$, despacio pero sin techo. Que el término general tienda a cero es necesario y *no* suficiente.

Continuidad, valor intermedio y continuidad uniforme

Definición 29.18 (Continuidad en un punto)**.** Sea $f:A\to\mathbb{R}$ con $A\subseteq\mathbb{R}$ y $x_0\in A$. La función f es *continua en* x_0 si

$$\text{para todo } \varepsilon>0 \text{ existe } \delta>0 \text{ tal que}\quad x\in A,\ |x-x_0|<\delta \implies |f(x)-f(x_0)|<\varepsilon.$$

Es *continua en* A si lo es en todos los puntos de A. Obsérvese el orden: primero se recibe ε (la tolerancia en la imagen) y después se busca δ (la tolerancia en el dominio), que puede depender de ε *y de* x_0.

Ejemplo 29.19 (Una continuidad demostrada con el delta explícito)**.** Sea $f(x)=x^2$ y $x_0=3$; veamos que es continua ahí.

Paso 1: factorizar el error. $|f(x)-f(3)|=|x^2-9|=|x-3|\,|x+3|$.

Paso 2: domar el factor molesto. El factor $|x+3|$ depende de x, así que lo acotamos restringiendo de antemano el radio: si $|x-3|<1$ entonces $2<x<4$ y por tanto $|x+3|<7$.

Paso 3: juntar. Bajo esa restricción, $|f(x)-f(3)|<7|x-3|$. Para que esto sea menor que ε basta $|x-3|<\varepsilon/7$.

Paso 4: declarar el delta. Tomamos $\delta=\min(1,\frac{\varepsilon}{7})$: el 1 protege el paso 2 y el $\varepsilon/7$ produce la conclusión.

Paso 5: comprobar. Con $\varepsilon=0{,}1$ sale $\delta=\frac{1}{70}=0{,}014286$. En efecto, $(3+\delta)^2-9=0{,}085918<0{,}1$ y $9-(3-\delta)^2=0{,}085510<0{,}1$. Funciona, y con margen.

Teorema 29.20 (Bolzano y el teorema del valor intermedio)**.** *Sea $f:[a,b]\to\mathbb{R}$ continua con $f(a)<0<f(b)$. Entonces existe $c\in(a,b)$ con $f(c)=0$. En consecuencia, si f es continua en $[a,b]$ y y es cualquier valor entre $f(a)$ y $f(b)$, existe $c\in[a,b]$ con $f(c)=y$: una función continua no puede pasar de un valor a otro sin tomar todos los intermedios.*

Demostración. Sea $S=\{x\in[a,b]:f(x)<0\}$. Es no vacío ($a\in S$) y está acotado superiormente por b, así que por el axioma del supremo existe $c=\sup S$, y claramente $c\in[a,b]$. Descartamos las dos alternativas a $f(c)=0$.

Si fuera $f(c)<0$, la continuidad con $\varepsilon=|f(c)|/2$ daría un $\delta>0$ tal que $f<0$ en todo el intervalo $(c-\delta,c+\delta)\cap[a,b]$. Como $f(b)>0$, se tiene $c<b$, y entonces habría puntos de S mayores que c, contra la definición de supremo.

Si fuera $f(c)>0$, el mismo argumento con $\varepsilon=f(c)/2$ daría un entorno donde $f>0$; como $f(a)<0$, se tiene $c>a$, y ningún punto de $(c-\delta,c]$ estaría en S, de modo que $c-\delta$ sería una cota superior de S menor que c: otra contradicción.

Luego $f(c)=0$. El caso general se obtiene aplicando lo anterior a $g(x)=f(x)-y$. □

Ejemplo 29.21 (El teorema en acción, con números)**.** Sea $f(x) = x^3 + x - 1$ en $[0,1]$. Es continua, $f(0) = -1 < 0$ y $f(1) = 1 > 0$, luego hay una raíz. La bisección la localiza: $f(0{,}6) = -0{,}184$ y $f(0{,}7) = 0{,}043$, así que la raíz está en $(0{,}6\,;\,0{,}7)$; el valor exacto es $c = 0{,}682328$. Obsérvese lo que el teorema *no* hace: no la calcula. Garantiza que existe, y esa garantía es lo que autoriza a buscarla.

Definición 29.22 (Continuidad uniforme)**.** $f : A \to \mathbb{R}$ es *uniformemente continua* en A si

$$\text{para todo } \varepsilon > 0 \text{ existe } \delta > 0 \text{ tal que} \quad x, y \in A,\ |x - y| < \delta \implies |f(x) - f(y)| < \varepsilon.$$

La única diferencia con la definición 29.18 es que aquí δ debe servir *a la vez para todos los puntos*: no puede depender de dónde se esté mirando. Toda función uniformemente continua es continua; el recíproco es falso.

Ejemplo 29.23 (Dos funciones continuas que no son uniformemente continuas)**. (a)** $f(x) = 1/x$ **en** $(0,1]$**.** Tomemos $x_n = \frac{1}{n}$ e $y_n = \frac{1}{n+1}$. La distancia $|x_n - y_n| = \frac{1}{n(n+1)}$ tiende a cero —vale 0,009091 para $n = 10$ y 0,000099 para $n = 100$—, pero $|f(x_n) - f(y_n)| = |n - (n+1)| = 1$ siempre. Ningún δ puede servir para $\varepsilon = \frac{1}{2}$. El culpable es que el dominio no es cerrado: la función se dispara cerca del borde que falta.

(b) $f(x) = x^2$ **en** $\mathbb{R}$**.** Con $x_n = n$ e $y_n = n + \frac{1}{n}$ la distancia tiende a cero, y sin embargo $f(y_n) - f(x_n) = 2 + \frac{1}{n^2}$, que vale 2,01 para $n = 10$ y 2,0001 para $n = 100$: nunca baja de 2. Aquí el culpable es que el dominio no está acotado. Los dos ejemplos juntos anuncian el teorema que sigue.

Teorema 29.24 (Heine–Cantor: continuidad uniforme en compactos)**.** *Toda función continua en un intervalo cerrado y acotado $[a,b]$ es uniformemente continua en él.*

Demostración. Por reducción al absurdo. Si f no fuese uniformemente continua, existiría un $\varepsilon_0 > 0$ tal que ningún δ sirve; en particular, tomando $\delta = \frac{1}{n}$ para cada n, habría puntos $x_n, y_n \in [a,b]$ con

$$|x_n - y_n| < \frac{1}{n} \qquad \text{y} \qquad |f(x_n) - f(y_n)| \geq \varepsilon_0.$$

La sucesión (x_n) está acotada, así que por Bolzano–Weierstrass (teorema 29.12) tiene una subsucesión $x_{n_k} \to c$, y $c \in [a,b]$ porque el intervalo es cerrado. Como $|x_{n_k} - y_{n_k}| < \frac{1}{n_k} \to 0$, también $y_{n_k} \to c$. Ahora la continuidad de f *en el punto* c da que $f(x_{n_k}) \to f(c)$ y $f(y_{n_k}) \to f(c)$, luego $|f(x_{n_k}) - f(y_{n_k})| \to 0$, en contradicción con que esa cantidad sea siempre $\geq \varepsilon_0 > 0$. □

Observación 29.25. Las dos hipótesis se usan y ninguna sobra: cerrado, para que el límite c no se escape del dominio (véase $1/x$ en $(0,1]$); acotado, para que Bolzano–Weierstrass se aplique (véase x^2 en $\mathbb{R}$). Este teorema es el que sostiene, más adelante en el capítulo, que toda función continua en $[a,b]$ es integrable Riemann y que los polinomios de Bernstein convergen uniformemente.

Derivada

Definición 29.26 (Derivada)**.** Sea f definida en un entorno de x_0. La *derivada* de f en x_0 es

$$f'(x_0) = \lim_{h\to 0} \frac{f(x_0+h)-f(x_0)}{h},$$

cuando ese límite existe y es finito; entonces f es *derivable* en x_0. Toda función derivable en un punto es continua en él, porque $f(x_0+h)-f(x_0) = h\cdot\frac{f(x_0+h)-f(x_0)}{h} \to 0\cdot f'(x_0) = 0$. El recíproco es falso, y la sección de este capítulo dedicada a la función de Weierstrass lo lleva al extremo.

Ejemplo 29.27 (Una derivada y una no derivada)**. (a)** $f(x)=x^3$ **en** $x_0=2$**.** El cociente incremental es

$$\frac{(2+h)^3-8}{h} = \frac{12h+6h^2+h^3}{h} = 12+6h+h^2 \longrightarrow 12.$$

Numéricamente, con $h=10^{-3}$ da 12,006001 y con $h=10^{-6}$ da 12,000006: converge a $12 = 3\cdot 2^2$, como manda la regla de la potencia.

(b) $f(x)=|x|$ **en** $x_0=0$**.** El cociente vale $\frac{|h|}{h}$, que es $+1$ para $h>0$ y -1 para $h<0$. Los dos límites laterales existen y son distintos, luego el límite no existe: la función es continua en 0 y no derivable ahí.

La integral de Riemann: particiones y sumas superiores e inferiores

Definición 29.28 (Partición, sumas de Darboux e integral de Riemann)**.** Sea $f:[a,b]\to\mathbb{R}$ *acotada*. Una *partición* de $[a,b]$ es un conjunto finito $P=\{a=x_0<x_1<\cdots<x_n=b\}$; el subintervalo i-ésimo es $I_i=[x_{i-1},x_i]$ y su longitud $\Delta x_i = x_i - x_{i-1}$. Se ponen

$$M_i = \sup_{I_i} f, \qquad m_i = \inf_{I_i} f,$$

que existen por el axioma del supremo, y se definen la *suma superior* y la *suma inferior*

$$U(f,P) = \sum_{i=1}^{n} M_i\,\Delta x_i, \qquad L(f,P) = \sum_{i=1}^{n} m_i\,\Delta x_i.$$

La *integral superior* es $\overline{\int_a^b} f = \inf_P U(f,P)$ y la *integral inferior* es $\underline{\int_a^b} f = \sup_P L(f,P)$. Cuando ambas coinciden, f se llama *integrable Riemann* y el valor común es $\int_a^b f$. Criterio equivalente y mucho más práctico: f es integrable si y solo si para todo $\varepsilon>0$ existe una partición P con $U(f,P)-L(f,P)<\varepsilon$.

Ejemplo 29.29 (Dos integrales calculadas con la definición)**. (a)** $\int_0^1 x\,\mathrm{d}x$**.** Con la partición uniforme de n trozos, $x_i = i/n$, la función es creciente, así que $M_i = i/n$ y $m_i = (i-1)/n$. Entonces

$$U(f,P_n) = \sum_{i=1}^{n} \frac{i}{n}\cdot\frac{1}{n} = \frac{n+1}{2n}, \qquad L(f,P_n) = \frac{n-1}{2n}, \qquad U-L = \frac{1}{n}.$$

Para $n = 10$: $U = 0{,}55$ y $L = 0{,}45$; para $n = 100$: $U = 0{,}505$ y $L = 0{,}495$. La diferencia $1/n$ se hace tan pequeña como se quiera, luego f es integrable, y como ambas sumas atrapan a $\frac{1}{2}$, se concluye $\int_0^1 x\,\mathrm{d}x = \frac{1}{2}$.

(b) $\int_0^1 x^2\,\mathrm{d}x$**.** Igual, con $U(f, P_n) = \frac{(n+1)(2n+1)}{6n^2}$ y $L(f, P_n) = \frac{(n-1)(2n-1)}{6n^2}$. Para $n = 10$ dan 0,385 y 0,285, con diferencia $0{,}1 = 1/n$; ambas tienden a $\frac{1}{3}$, que es el valor de la integral. Nótese el patrón: para una función monótona, $U - L = \frac{(b-a)(f(b)-f(a))}{n}$, y de ahí sale de un golpe que *toda función monótona acotada es integrable.*

Regla: decidir integrabilidad con sumas superiores e inferiores

1. **Comprueba primero que f está acotada.** Es hipótesis de la definición 29.28, no conclusión. Sin ella los supremos M_i pueden ser infinitos y no hay nada que sumar.
2. **Elige la partición uniforme salvo que la función pida otra cosa.** Con n trozos iguales, $\Delta x_i = \frac{b-a}{n}$ sale factor común y las sumas se convierten en sumas aritméticas conocidas.
3. **Identifica M_i y m_i mirando la monotonía.** En un tramo creciente, el supremo está a la derecha y el ínfimo a la izquierda; en uno decreciente, al revés. Si la función cambia de sentido, parte el intervalo por los puntos críticos.
4. **Calcula la diferencia $U - L$ y no las sumas por separado.** Para funciones monótonas la suma telescopa y queda $\frac{(b-a)\lvert f(b)-f(a)\rvert}{n}$; ese solo cálculo demuestra la integrabilidad sin conocer el valor.
5. **Para demostrar que NO es integrable, exhibe una cota inferior de $U - L$.** Basta ver que $U(f, P) - L(f, P) \geq c > 0$ para *toda* partición. Es lo que ocurre con la función de Dirichlet: cada subintervalo contiene racionales e irracionales, luego $M_i = 1$ y $m_i = 0$ siempre, y $U - L = b - a$ para cualquier P.
6. **Usa los atajos antes de sufrir.** Continua en $[a, b]$ implica integrable (vía Heine–Cantor, teorema 29.24); monótona implica integrable; continua salvo en finitos puntos, también. El criterio definitivo, con un "si y solo si", es el de Lebesgue, que se demuestra más adelante en este mismo capítulo.
7. **Control.** Refinar una partición nunca empeora: añadir puntos hace bajar U y subir L. Si tus cuentas dan lo contrario, hay un error de signo.

Convergencia puntual frente a uniforme

Definición 29.30 (Las dos convergencias de funciones)**.** Sea $f_n : A \to \mathbb{R}$ una sucesión de funciones. Se dice que $f_n \to f$ *puntualmente* si para cada $x \in A$ fijo la sucesión numérica $f_n(x)$ converge a $f(x)$; y *uniformemente* si

$$\sup_{x \in A} |f_n(x) - f(x)| \longrightarrow 0,$$

es decir, si para todo $\varepsilon > 0$ existe un N —*el mismo para todos los puntos*— tal que $n \geq N$ implica $|f_n(x) - f(x)| < \varepsilon$ para todo $x \in A$. La única diferencia está en el orden de los

cuantificadores, y es toda la diferencia: el límite uniforme de funciones continuas es continuo, el puntual no tiene por qué serlo. El ejemplo canónico es $f_n(x) = x^n$ en $[0,1]$, cuyo límite puntual salta en $x = 1$, y la regla operativa completa para distinguirlas está más adelante en este capítulo.

Abiertos, cerrados, compactos, densos y numerables

Definición 29.31 (El vocabulario topológico de la recta)**.** Sea $A \subseteq \mathbb{R}$.

- A es *abierto* si para todo $x \in A$ existe $r > 0$ con $(x - r, x + r) \subseteq A$.
- A es *cerrado* si su complementario es abierto; equivalentemente, si el límite de toda sucesión de puntos de A que converja en $\mathbb{R}$ pertenece a A.
- A es *compacto* si todo recubrimiento suyo por abiertos admite un subrecubrimiento finito. El teorema de Heine–Borel afirma que en $\mathbb{R}$ esto equivale a ser cerrado y acotado, y también a que toda sucesión de A tenga una subsucesión convergente a un punto de A.
- A es *denso* en $\mathbb{R}$ si todo intervalo abierto no vacío contiene algún punto de A.
- A es *numerable* si existe una biyección entre A y $\mathbb{N}$, y *a lo sumo numerable* si es finito o numerable. Esta noción se desarrolla entera en la sección siguiente.

Ejemplo 29.32 (El vocabulario, con casos concretos)**. (a) Abiertos y cerrados no son opuestos.** $(0,1)$ es abierto y no cerrado; $[0,1]$ es cerrado y no abierto; $[0,1)$ no es ninguna de las dos cosas; y $\varnothing$ y $\mathbb{R}$ son a la vez abiertos y cerrados. "No abierto" no significa "cerrado".

(b) Compacto de verdad. $[0,1]$ es compacto. En cambio $(0,1)$ no lo es: el recubrimiento por los abiertos $U_n = \left(\frac{1}{n}, 1\right)$ cubre todo el intervalo, y ninguna subfamilia finita lo hace, porque una cantidad finita de ellos está contenida en el mayor, que deja fuera los puntos próximos a 0. Tampoco $[0,\infty)$ es compacto, por no ser acotado.

(c) Denso. $\mathbb{Q}$ es denso en $\mathbb{R}$: entre dos reales cualesquiera hay siempre un racional. También lo son los irracionales. Y sin embargo $\mathbb{Q}$ es numerable y los irracionales no, como demuestra la sección siguiente: densidad y cardinal son propiedades independientes.

(d) Cerrado y numerable. $\{0\} \cup \left\{\frac{1}{n} : n \in \mathbb{N}\right\}$ es cerrado (contiene el límite 0 de la sucesión) y numerable, y no contiene ningún intervalo. Quitarle el 0 lo convierte en un conjunto que ya no es cerrado.

Pausa para pensar

Merece la pena ver el esqueleto lógico de lo que se acaba de construir, porque todo cuelga de un único clavo. El axioma del supremo (definición 29.2) da el teorema del valor intermedio (teorema 29.20) y da Bolzano–Weierstrass (teorema 29.12); Bolzano–Weierstrass da Heine–Cantor (teorema 29.24); y Heine–Cantor da la integrabilidad de las funciones continuas. Si se quita el axioma —es decir, si se trabaja en $\mathbb{Q}$— caen los cuatro a la vez: $f(x) = x^2 - 2$ pasa de negativa a positiva sin anularse, la sucesión de truncaciones de $\sqrt{2}$ no tiene subsucesión convergente en $\mathbb{Q}$, y el edificio entero se

desploma. Que la recta no tenga agujeros no es una observación geométrica ociosa: es la hipótesis de trabajo de todo el capítulo.

29.3. Contar lo infinito: cardinalidad

Definición 29.33 (Equipotencia)**.** Dos conjuntos A y B son *equipotentes*, y se escribe $|A| = |B|$, si existe una biyección $f : A \to B$. Un conjunto es *numerable* si es equipotente a $\mathbb{N}$, y *a lo sumo numerable* si es finito o numerable.

Idea clave

Intuición primero. Contar es emparejar. Si en una sala hay sillas y personas y cada persona tiene exactamente una silla y cada silla exactamente una persona, hay tantas sillas como personas, y no hizo falta contar ninguna de las dos cosas. Cantor tomó esa definición —la única que no usa números— y la aplicó a conjuntos infinitos. El resultado es que "infinito" deja de ser una palabra y pasa a ser una familia de tamaños distintos.

Teorema 29.34 (La unión numerable de numerables es numerable)**.** *Si $A_1, A_2, \ldots$ son a lo sumo numerables, entonces $\bigcup_{n\geq 1} A_n$ es a lo sumo numerable. En particular $\mathbb{N} \times \mathbb{N}$ es numerable.*

Demostración. Escribe los elementos de A_n como $a_{n1}, a_{n2}, \ldots$ (repitiendo si A_n es finito). Coloca todo en una tabla infinita cuya fila n es A_n y recórrela por diagonales:

$$a_{11}; \quad a_{12}, a_{21}; \quad a_{13}, a_{22}, a_{31}; \quad a_{14}, a_{23}, a_{32}, a_{41}; \ \ldots$$

Cada diagonal es finita, así que el recorrido llega a todo elemento en un número finito de pasos: eso es una sobreyección $\mathbb{N} \to \bigcup A_n$. Suprimiendo repeticiones se obtiene una biyección con $\mathbb{N}$ o con un subconjunto finito. Para $\mathbb{N} \times \mathbb{N}$ basta tomar $A_n = \{n\} \times \mathbb{N}$. La biyección explícita es el *emparejamiento de Cantor*

$$\pi(m, n) = \frac{(m+n)(m+n+1)}{2} + n. \qquad \square$$

Idea clave

Qué dice esto en cristiano. Contar es emparejar. Dos conjuntos tienen el mismo tamaño si se pueden emparejar uno a uno sin que sobre ni falte nadie, y esa definición, que es la única razonable, produce resultados que chocan de frente con la intuición: hay tantos números pares como naturales, y tantos racionales como naturales, aunque en ambos casos uno sea parte propia del otro. Esa es precisamente la característica de lo infinito. Y sin embargo hay infinitos de tamaños distintos: $\mathbb{R}$ no se puede emparejar con $\mathbb{N}$, y eso lo demostró Cantor con el argumento diagonal.

Ejemplo 29.35 (Emparejamientos que sorprenden y uno que es imposible)**. (a) Los pares son tantos como los naturales.** El emparejamiento $n \mapsto 2n$ es una biyección de $\mathbb{N}$ en los

pares. Quitar la mitad de un conjunto infinito puede dejar un conjunto del mismo tamaño; es el hotel de Hilbert.

(b) $\mathbb{N}\times\mathbb{N}$ es numerable. Se recorren los pares por diagonales: primero los que suman 1, luego los que suman 2, y así. La fórmula de emparejamiento de Cantor es $(a,b)\mapsto \frac{(a+b)(a+b+1)}{2}+b$, que a $(1,2)$ le asigna 8 y a $(2,1)$ le asigna 7. De aquí sale que $\mathbb{Q}$ es numerable, escribiendo cada fracción como un par.

(c) La unión numerable de numerables es numerable. Se colocan los conjuntos en filas y se recorre por diagonales, exactamente igual que en (b). Es la razón de que los números algebraicos sean numerables: hay numerables polinomios con coeficientes enteros y cada uno tiene finitas raíces.

(d) El argumento diagonal. Supongamos una lista de todos los números de $[0,1]$. Se construye uno nuevo cambiando la primera cifra del primero, la segunda del segundo, y así. El número resultante difiere de cada elemento de la lista en al menos una cifra, luego no estaba: la lista no era completa. Por tanto $[0,1]$ no es numerable, y como consecuencia *casi todos* los reales son trascendentes, aunque exhibir uno concreto costó hasta Liouville en 1844.

Cuidado

[Infinito no es un número, y "más grande" hay que definirlo] Decir que hay "más" reales que naturales significa exactamente que no existe biyección, no que uno contenga al otro. Los pares están contenidos en $\mathbb{N}$ y tienen el mismo cardinal. Segundo aviso: el argumento diagonal exige cuidado con los desarrollos decimales dobles, como $0{,}4999\ldots = 0{,}5$; se evita eligiendo cifras que no sean 0 ni 9. Y tercero: el teorema de Cantor sobre $\mathcal{P}(A)$ garantiza que no hay un infinito máximo, porque siempre se puede construir uno mayor tomando partes; la torre de cardinales no tiene techo.

Corolario 29.36. *$\mathbb{Z}$ y $\mathbb{Q}$ son numerables.*

Demostración. $\mathbb{Z}=\{0,1,-1,2,-2,\ldots\}$ se enumera directamente. Para $\mathbb{Q}$, cada racional se escribe de forma única como p/q con $q\geq 1$ y $\gcd(p,q)=1$; la aplicación $p/q\mapsto(p,q)$ es inyectiva de $\mathbb{Q}$ en $\mathbb{Z}\times\mathbb{N}$, que es numerable, y un subconjunto infinito de un numerable es numerable. □

Cuidado

Que $\mathbb{Q}$ sea numerable choca con la intuición geométrica: entre dos racionales cualesquiera hay infinitos racionales, y sin embargo se pueden poner en fila india indexados por $\mathbb{N}$. La densidad y el cardinal son propiedades *independientes*. El conjunto de Cantor, más abajo, tiene la propiedad opuesta: no contiene ningún intervalo y aun así no es numerable.

Teorema 29.37 (Cantor, 1891: el argumento diagonal)**.** *$\mathbb{R}$ no es numerable. De hecho $[0,1]$ no lo es.*

Demostración. Supón que existiera una enumeración $x_1, x_2, x_3, \dots$ de todos los elementos de $[0,1]$. Escribe cada uno en base 10,

$$x_n = 0{,}d_{n1}d_{n2}d_{n3}\dots,$$

eligiendo, para los que admiten dos desarrollos, el que no termina en infinitos nueves. Define un número $y = 0{,}e_1e_2e_3\dots$ con

$$e_n = \begin{cases} 5, & \text{si } d_{nn} \neq 5, \\ 6, & \text{si } d_{nn} = 5. \end{cases}$$

Entonces $y \in [0,1]$, y y no usa nunca los dígitos 0 ni 9, así que su desarrollo es único. Pero $y \neq x_n$ para todo n, porque difieren en el dígito n-ésimo. Contradicción: la enumeración no era completa. □

Ejemplo 29.38 (La diagonal y sus parientes)**. (a) La construcción explícita.** Si alguien presenta una lista $x_1 = 0{,}\mathbf{1}357\dots$, $x_2 = 0{,}2\mathbf{4}68\dots$, $x_3 = 0{,}71\mathbf{9}2\dots$, se fabrica y cambiando la cifra n-ésima de x_n: por ejemplo, poniendo 5 donde había otra cosa y 6 donde había un 5. Con los tres primeros datos, $y = 0{,}555\dots$ difiere de x_1 en la primera cifra, de x_2 en la segunda y de x_3 en la tercera. Sea cual sea la lista, y no está en ella.

(b) El detalle técnico. Hay que evitar las colas de nueves, porque $0{,}4999\dots = 0{,}5$ y dos desarrollos distintos podrían representar el mismo número. Elegir las cifras dentro de $\{5,6\}$ resuelve el problema de golpe: ningún número tiene dos desarrollos que usen sólo esas dos cifras.

(c) Los que sí son numerables. $\mathbb{N}$, $\mathbb{Z}$, $\mathbb{Q}$ y el conjunto de los números algebraicos son todos numerables. Como $\mathbb{R}$ no lo es, existen números trascendentes, y de hecho "casi todos" lo son. Cantor demostró así su existencia en 1874 sin exhibir ninguno, treinta años después de que Liouville construyera el primero a mano.

(d) El conjunto de Cantor. Tiene medida cero y sin embargo tiene el mismo cardinal que $\mathbb{R}$: cada uno de sus puntos se codifica por una sucesión de ceros y unos (izquierda o derecha en cada paso), y hay tantas sucesiones como reales. Cardinal grande y medida nula son propiedades independientes.

Idea clave

Por qué funciona. La lista propuesta es un cuadrado infinito de dígitos. El argumento camina por la diagonal y en cada paso *desobedece* al número que está en esa fila. El número construido se diferencia de todos los de la lista por diseño, uno a la vez. No es un truco sobre decimales: es el mismo mecanismo que produce la paradoja del barbero, el teorema de incompletitud de Gödel y el problema de la parada. Siempre que un objeto pueda "leer" su propia descripción, la diagonal produce algo que la lista no previó.

Corolario 29.39 (Existen números trascendentes)**.** *El conjunto de los números algebraicos —las raíces de polinomios con coeficientes racionales— es numerable. Como $\mathbb{R}$ no lo es, existen (muchísimos) reales trascendentes.*

Demostración. Los polinomios de grado n con coeficientes en $\mathbb{Q}$ forman un conjunto equipotente a $\mathbb{Q}^{n+1}$, numerable; la unión sobre n sigue siendo numerable; cada polinomio no nulo tiene a lo sumo n raíces (teorema demostrado en el capítulo 24). Así, el conjunto de algebraicos es unión numerable de conjuntos finitos. □

Observación 29.40. Esta demostración, de 1874, es célebre por lo que *no* hace: no exhibe ni un solo número trascendente. Liouville había tenido que construir uno a mano en 1844, y la prueba de que e y π son trascendentes (Hermite 1873, Lindemann 1882) es dura. Cantor demuestra que casi todos los números lo son sin señalar ninguno. Es el primer argumento de existencia por conteo de la historia de la matemática.

Teorema 29.41 (Cantor: el conjunto de partes es estrictamente mayor)**.** *Para todo conjunto A no existe ninguna sobreyección $A \to \mathcal{P}(A)$. En particular $|A| < |\mathcal{P}(A)|$.*

Demostración. Sea $f : A \to \mathcal{P}(A)$ cualquiera y define

$$D = \{a \in A : a \notin f(a)\} \in \mathcal{P}(A).$$

Si $D = f(b)$ para algún b, pregúntate si $b \in D$. Si $b \in D$, entonces por definición de D se tiene $b \notin f(b) = D$; si $b \notin D = f(b)$, entonces b cumple la condición y $b \in D$. Ambas ramas son contradictorias, luego D no está en la imagen y f no es sobreyectiva. □

Idea clave

Qué dice esto en cristiano. El conjunto de todos los subconjuntos de A es siempre estrictamente mayor que A, por grande que sea A. No hay un infinito máximo: dado cualquiera, tomando partes se obtiene otro mayor. La demostración es la misma diagonal de antes en versión abstracta: si $f\colon A \to \mathcal{P}(A)$ fuese sobreyectiva, el conjunto $D = \{a : a \notin f(a)\}$ tendría que ser $f(d)$ para algún d, y entonces $d \in D \iff d \notin D$.

Ejemplo 29.42 (Cardinales que crecen sin techo)**. (a) Casos finitos.** Si $|A| = 3$, entonces $|\mathcal{P}(A)| = 8$. En general $2^n > n$ para todo n, cosa evidente; lo notable es que la misma demostración funciona sin cambiar una coma para conjuntos infinitos.

(b) La escalera infinita. $|\mathbb{N}| = \aleph_0$, $|\mathcal{P}(\mathbb{N})| = 2^{\aleph_0} = |\mathbb{R}|$, $|\mathcal{P}(\mathbb{R})| = 2^{2^{\aleph_0}}$, y así indefinidamente. El último es el cardinal del conjunto de todas las funciones de $\mathbb{R}$ en $\{0,1\}$, o el de todas las topologías sobre $\mathbb{R}$.

(c) La paradoja de Russell es el mismo argumento. Aplicando el razonamiento al "conjunto de todos los conjuntos" se obtiene el conjunto de los que no se contienen a sí mismos, y la contradicción de siempre. La respuesta de la teoría axiomática es negar que exista ese conjunto: los axiomas de Zermelo–Fraenkel sólo permiten separar subconjuntos de conjuntos ya existentes.

(d) Lo que queda abierto. ¿Hay algún cardinal estrictamente entre $\aleph_0$ y $2^{\aleph_0}$? Ésa es la hipótesis del continuo, y Gödel (1940) y Cohen (1963) demostraron que ni ella ni su negación se pueden deducir de los axiomas usuales. No es que no se sepa: es que la pregunta no tiene respuesta dentro del sistema.

Pausa para pensar

Aplica el teorema a $A = \mathbb{N}$: hay más subconjuntos de $\mathbb{N}$ que números naturales. Aplícalo otra vez a $\mathcal{P}(\mathbb{N})$, y otra vez: obtienes una torre infinita de infinitos, cada uno estrictamente mayor que el anterior. La pregunta de si hay algún cardinal *entre* $|\mathbb{N}|$ y $|\mathbb{R}| = |\mathcal{P}(\mathbb{N})|$ es la *hipótesis del continuo*; Gödel (1940) y Cohen (1963) demostraron que los axiomas usuales de la teoría de conjuntos no la deciden. Es indecidible, no desconocida.

Regla: decidir si un conjunto es numerable

La pregunta "¿cuántos elementos tiene este conjunto infinito?" se responde casi siempre con una de estas cinco maniobras, y conviene probarlas en este orden.

1. **¿Es un subconjunto de uno numerable?** Todo subconjunto de un conjunto numerable es numerable (o finito). Si tu conjunto vive dentro de $\mathbb{Z}$, de $\mathbb{Q}$ o del conjunto de las sucesiones finitas de enteros, ya has terminado.

2. **¿Es una unión numerable de conjuntos numerables?** Entonces es numerable. Este es el argumento que hace numerable a $\mathbb{Q} = \bigcup_{q\geq 1}\{p/q : p \in \mathbb{Z}\}$ y al conjunto de los polinomios con coeficientes enteros (uno por cada grado).

3. **¿Es un producto finito de numerables?** También numerable: $\mathbb{Z}\times\mathbb{Z}$ se enumera en zigzag por diagonales. Cuidado: el producto *infinito* no lo es. $\{0,1\}^{\mathbb{N}}$ —las sucesiones infinitas de ceros y unos— es no numerable.

4. **¿Puedo describir cada elemento con una cantidad finita de información tomada de un alfabeto numerable?** Si sí, es numerable. Con este criterio se ve de un golpe que los números algebraicos son numerables (cada uno queda descrito por un polinomio de coeficientes enteros más un índice que dice cuál de sus raíces es) y que hay solo una cantidad numerable de programas de ordenador, de demostraciones y de frases en español.

5. **Si sospechas que *no* es numerable, busca la diagonal.** Supón una enumeración completa $x_1, x_2, x_3, \ldots$ y construye un elemento que difiera del n-ésimo justo en la posición n. Ese elemento pertenece al conjunto pero no está en la lista, y la contradicción cierra el argumento.

6. **Control final.** Numerable y no numerable no son "grande" y "pequeño" en el sentido de la medida: el conjunto de Cantor es no numerable y tiene medida cero, mientras que $\mathbb{Q}$ es numerable y denso en $\mathbb{R}$. Densidad, medida y cardinalidad son tres nociones independientes de tamaño, y mezclarlas es el error conceptual más frecuente de este capítulo.

Ejemplo 29.43 (Cuatro conjuntos, cuatro respuestas)**.** *(a) Los racionales de* $[0,1]$. Subconjunto de $\mathbb{Q}$, que es numerable: numerable. Y sin embargo es denso: entre dos reales cualesquiera hay infinitos. Denso no implica grande en cardinal.

(b) Los números algebraicos. Para cada n y cada elección de $n+1$ coeficientes enteros hay

un polinomio, y cada polinomio tiene a lo sumo n raíces. El conjunto de polinomios de grado n con coeficientes enteros es $\mathbb{Z}^{n+1}$, numerable; la unión sobre n sigue siendo numerable; y sustituir cada polinomio por su lista finita de raíces conserva la numerabilidad. Luego los algebraicos son numerables.

(c) Los trascendentes. Si fueran numerables, $\mathbb{R} = \{\text{algebraicos}\} \cup \{\text{trascendentes}\}$ sería unión de dos numerables y por tanto numerable, contra el argumento diagonal. Luego los trascendentes son no numerables. Este es el célebre argumento de Cantor: demuestra que "casi todos" los reales son trascendentes *sin exhibir ni uno solo*. Hicieron falta décadas más para probar que e y π lo son, y todavía hoy no se sabe si $e + \pi$ lo es.

(d) Las funciones $f : \mathbb{N} \to \{0, 1\}$. Es exactamente $\{0, 1\}^{\mathbb{N}}$: no numerable, por diagonalización. Como consecuencia, hay más funciones de $\mathbb{N}$ en $\{0, 1\}$ que programas capaces de calcularlas: existen funciones no computables, y existen por razones de puro conteo, antes de cualquier discusión sobre qué es un algoritmo.

Regla: trabajar con supremos e ínfimos usando la completitud

El axioma del supremo es lo único que distingue $\mathbb{R}$ de $\mathbb{Q}$, y casi todo teorema de análisis lo usa en algún punto. Manejarlo bien es una destreza mecánica.

1. **Comprueba que el supremo existe.** El axioma dice: todo conjunto $A \subseteq \mathbb{R}$ *no vacío* y *acotado superiormente* tiene supremo. Las dos hipótesis son necesarias; sin ellas no hay nada que discutir. Por convenio, $\sup \varnothing = -\infty$ y $\sup A = +\infty$ si A no está acotado.

2. **Usa la caracterización en dos partes.** Que $s = \sup A$ significa exactamente dos cosas: (i) $a \leq s$ para todo $a \in A$, es decir s es cota superior; (ii) para todo $\varepsilon > 0$ existe $a \in A$ con $a > s - \varepsilon$, es decir ninguna cota es menor. Toda demostración de un supremo consta de estas dos verificaciones, ni una más.

3. **El supremo no tiene por qué pertenecer al conjunto.** Si pertenece, se llama máximo. Recuerda que $\sup(0, 1) = 1 \notin (0, 1)$. Confundir supremo con máximo es el error más frecuente del capítulo.

4. **Convierte el supremo en una sucesión.** De la parte (ii) con $\varepsilon = 1/n$ se obtiene $a_n \in A$ con $s - \frac{1}{n} < a_n \leq s$, luego $a_n \to s$. Este truco convierte un enunciado sobre supremos en uno sobre límites, que suele ser más fácil de manipular.

5. **Pasa al ínfimo por reflexión.** Siempre vale $\inf A = -\sup(-A)$, donde $-A = \{-a : a \in A\}$. Nunca hace falta demostrar dos veces la misma cosa.

6. **Recuerda las reglas de combinación.** Se tiene $\sup(A + B) = \sup A + \sup B$, y $\sup(cA) = c \sup A$ para $c > 0$ pero $\sup(cA) = c \inf A$ para $c < 0$. En cambio $\sup(A \cap B)$ no admite fórmula: puede ser cualquier cosa, incluso $-\infty$.

7. **Verifica con el caso racional.** Si tu argumento funcionaría igual en $\mathbb{Q}$, es que no has usado la completitud y probablemente el resultado es falso. El conjunto $\{q \in \mathbb{Q} : q^2 < 2\}$ está acotado y no tiene supremo *en* $\mathbb{Q}$: ese es el hueco que $\mathbb{R}$ rellena.

Ejemplo 29.44 (Cinco supremos calculados con la definición)**. (a)** $A = \left\{1 - \frac{1}{n} : n \in \mathbb{N}\right\}$**.** Afirmamos $\sup A = 1$. Cota superior: $1 - \frac{1}{n} < 1$ para todo n, hecho. Y ninguna cota menor: dado $\varepsilon > 0$, tomamos $n > 1/\varepsilon$, con lo que $1 - \frac{1}{n} > 1 - \varepsilon$. Ambas partes verificadas. Nótese que $1 \notin A$: hay supremo pero no máximo. En cambio $\inf A = 0 = 1 - \frac{1}{1}$, que sí es mínimo.

(b) $B = \left\{\frac{(-1)^n n}{n+1} : n \in \mathbb{N}\right\}$**.** Los términos pares suben hacia 1 y los impares bajan hacia -1. Entonces $\sup B = 1$ e $\inf B = -1$, y ninguno se alcanza. Comprobación de (ii) para el supremo: dado $\varepsilon > 0$, con n par y $n > \frac{1}{\varepsilon}$ se tiene $\frac{n}{n+1} = 1 - \frac{1}{n+1} > 1 - \varepsilon$.

(c) $C = \{q \in \mathbb{Q} : q^2 < 2\}$**.** Dentro de $\mathbb{R}$, $\sup C = \sqrt{2} \approx 1{,}41421$. La demostración de que $\sqrt{2}$ es cota superior es directa; la de que ninguna cota es menor usa que entre dos reales siempre hay un racional. Pero *dentro de* $\mathbb{Q}$ este conjunto no tiene supremo: si $s \in \mathbb{Q}$ fuera cota superior mínima, se demuestra que ni $s^2 < 2$, ni $s^2 > 2$, ni $s^2 = 2$ son posibles. Este ejemplo es el motivo de que $\mathbb{R}$ exista.

(d) Suma de conjuntos. Con $A = (0, 1)$ y $B = (2, 5)$ se tiene $A + B = (2, 6)$, y en efecto $\sup(A + B) = 6 = 1 + 5 = \sup A + \sup B$. Esta identidad se demuestra en dos direcciones: $\leq$ es inmediato porque $a + b \leq \sup A + \sup B$; y $\geq$ se obtiene eligiendo $a > \sup A - \varepsilon/2$ y $b > \sup B - \varepsilon/2$, con lo que $a + b > \sup A + \sup B - \varepsilon$.

(e) Un supremo que se convierte en teorema. Sea $f : [a, b] \to \mathbb{R}$ continua con $f(a) < 0 < f(b)$, y sea

$$S = \{x \in [a, b] : f(x) < 0\}.$$

Este conjunto es no vacío (contiene a) y está acotado por b, luego existe $c = \sup S$. Se demuestra entonces que $f(c) = 0$: si fuera $f(c) < 0$, la continuidad daría un intervalo alrededor de c con $f < 0$, y habría puntos de S mayores que c; si fuera $f(c) > 0$, habría un entorno con $f > 0$, y $c - \delta$ sería una cota superior menor. Ambas contradicen que c sea el supremo. Esto es el teorema de Bolzano, y toda su fuerza reside en el axioma del supremo. En $\mathbb{Q}$ es falso: $f(x) = x^2 - 2$ pasa de negativa a positiva sin anularse nunca.

Cuidado

La propiedad (ii) exige la existencia de un elemento de A mayor que $s - \varepsilon$, no de un elemento mayor o igual que s. Escribir «existe $a \in A$ con $a = s$» equivale a afirmar que hay máximo, que es una propiedad mucho más fuerte y en general falsa. La diferencia entre $> s - \varepsilon$ y $= s$ es exactamente la diferencia entre supremo y máximo.

Idea clave

Todas las formulaciones de la completitud de $\mathbb{R}$ son equivalentes entre sí: el axioma del supremo, que toda sucesión de Cauchy converja, que toda sucesión monótona acotada converja, el teorema de los intervalos encajados, y el de Bolzano–Weierstrass. Elegir una como axioma y deducir las demás es una cuestión de gusto expositivo. Lo que importa es la idea común: en $\mathbb{R}$ no hay huecos, y por eso los procesos límite terminan *dentro* del conjunto en vez de escaparse fuera.

29.4. El conjunto de Cantor y la escalera del diablo

Definición 29.45 (Conjunto ternario de Cantor)**.** Parte de $C_0 = [0,1]$ y define C_{n+1} quitando de cada intervalo de C_n su tercio central abierto. Así

$$C_1 = [0,\tfrac{1}{3}] \cup [\tfrac{2}{3},1], \qquad C_2 = [0,\tfrac{1}{9}] \cup [\tfrac{2}{9},\tfrac{1}{3}] \cup [\tfrac{2}{3},\tfrac{7}{9}] \cup [\tfrac{8}{9},1], \ \ldots$$

El *conjunto de Cantor* es $C = \bigcap_{n\geq 0} C_n$.



Figura 29.1: Las seis primeras etapas de la construcción del conjunto de Cantor: en cada paso se quita el tercio central de cada intervalo. La longitud total tiende a cero y sin embargo lo que queda no es numerable.

Idea clave

Qué dice esto en cristiano. El conjunto de Cantor es el ejemplo que rompe todas las intuiciones a la vez. Se construye quitando tercios centrales infinitas veces, y lo que queda tiene longitud cero (se ha quitado todo) y sin embargo tiene tantos puntos como la recta entera. Es cerrado, no tiene ningún intervalo dentro, no tiene puntos aislados y es homeomorfo a su propio cuadrado. Sirve como contraejemplo estándar cada vez que alguien afirma que "longitud cero significa pocos puntos" o que "muchos puntos significa algo de longitud".

Ejemplo 29.46 (Longitud cero con tantos puntos como $\mathbb{R}$)**. (a) Lo que queda en cada paso.** C_n tiene 2^n intervalos de longitud 3^{-n}, luego mide $\left(\frac{2}{3}\right)^n$: vale 0,666667 para $n=1$, 0,444444 para $n=2$, 0,296296 para $n=3$ y 0,017342 para $n=10$. El límite es cero.

(b) Lo que se ha quitado suma uno. En el paso k se quitan 2^k intervalos de longitud $3^{-(k+1)}$, y $\sum_{k\geq 0} \frac{2^k}{3^{k+1}} = \frac{1/3}{1-2/3} = 1$. Se ha quitado longitud 1 de un intervalo de longitud 1, y aun así queda algo.

(c) Cuántos puntos quedan. Un número está en C exactamente cuando admite un desarrollo en base 3 sin ninguna cifra 1. Cambiando cada 2 por un 1 se obtiene una biyección con los desarrollos binarios, es decir con $[0,1]$: el cardinal es el mismo que el de $\mathbb{R}$. Por ejemplo $\frac{1}{3} = 0{,}0222\ldots_3$ y $\frac{2}{3} = 0{,}2_3$ están en C; $\frac{1}{2} = 0{,}111\ldots_3$ no; y $\frac{1}{4} = 0{,}020202\ldots_3$ sí, aunque no sea extremo de ningún intervalo.

(d) Las propiedades topológicas. C es compacto (cerrado y acotado), no contiene ningún intervalo (su interior es vacío), y no tiene puntos aislados: cada punto es límite de otros de C. Un conjunto con esas tres propiedades se llama *perfecto y totalmente disconexo*,

y resulta que todos ellos son homeomorfos entre sí: el conjunto de Cantor es único salvo deformación.

Cuidado

[Medida cero y "pequeño" no son lo mismo] C tiene medida cero y cardinal $2^{\aleph_0}$: es enorme en cardinal y nulo en medida. Al revés también ocurre: $\mathbb{Q}$ tiene medida cero y es numerable, pero es *denso* en $\mathbb{R}$, mientras que C no lo es. Hay al menos tres nociones distintas de "pequeño" (numerable, medida nula, primera categoría) y no coinciden. Segundo aviso: existen conjuntos de Cantor "gordos", que se construyen quitando trozos más pequeños y tienen medida positiva sin contener ningún intervalo; así que la medida cero es un accidente de esta construcción concreta, no una consecuencia de su forma.

Teorema 29.47 (Propiedades de C)**.** *(a) C es compacto y no vacío.*

(b) C tiene medida (longitud) cero.

(c) C no contiene ningún intervalo: su interior es vacío.

(d) C no es numerable; de hecho $|C| = |\mathbb{R}|$.

(e) C es perfecto: es cerrado y todo punto suyo es límite de otros puntos de C.

Demostración. (a) Cada C_n es unión finita de cerrados acotados, luego compacto; la intersección de una familia decreciente de compactos no vacíos es compacta y no vacía. Además $0, 1 \in C_n$ para todo n.

(b) C_n consta de 2^n intervalos de longitud 3^{-n}, así que su longitud total es $(2/3)^n \to 0$. Como $C \subseteq C_n$ para todo n, su medida exterior es $\leq (2/3)^n$ para todo n, luego es cero. Equivalentemente, la longitud total de lo eliminado es

$$\sum_{n\geq 0} \frac{2^n}{3^{n+1}} = \frac{1}{3} \cdot \frac{1}{1 - 2/3} = 1.$$

(c) Si C contuviese un intervalo de longitud $\ell > 0$, ese intervalo estaría dentro de C_n para todo n; pero los intervalos de C_n miden $3^{-n} < \ell$ para n grande. Imposible.

(d) Caracterización en base 3: $x \in C$ si y solo si admite un desarrollo $x = \sum_{k\geq 1} a_k 3^{-k}$ con todos los $a_k \in \{0, 2\}$. (Quitar el tercio central en el paso k es exactamente prohibir $a_k = 1$.) La aplicación

$$\varphi\Big(\sum_{k\geq 1} a_k 3^{-k}\Big) = \sum_{k\geq 1} \frac{a_k}{2}\, 2^{-k}$$

manda C sobre $[0, 1]$ leyendo la sucesión de ceros y doses como binario. Es sobreyectiva, luego $|C| \geq |[0, 1]|$; y $C \subseteq [0, 1]$ da la otra desigualdad.

(e) C es cerrado por (a). Sea $x \in C$ y $\varepsilon > 0$. Toma n con $3^{-n} < \varepsilon$: el punto x está en uno de los 2^n intervalos de C_n, cuyos dos extremos pertenecen a C y distan de x menos que ε; al menos uno de ellos es distinto de x. □

Ejemplo 29.48 (El conjunto de Cantor, medido y contado)**. (a) La medida.** En el paso n quedan 2^n intervalos de longitud 3^{-n}, con longitud total $(2/3)^n$: vale 0,666667; 0,444444; 0,296296; 0,197531 para $n = 1, 2, 3, 4$, y tiende a cero. Lo que se quita suma $\sum_{n\geq 0} 2^n/3^{n+1} = \frac{1}{3} \cdot \frac{1}{1-2/3} = 1$: se ha quitado todo el intervalo y sin embargo queda algo.

(b) Qué queda. Los puntos de C son exactamente los que admiten desarrollo en base 3 sin ningún 1. Así, $1/3 = 0{,}0222\ldots_3$ y $1/4 = 0{,}020202\ldots_3$ pertenecen a C, mientras que $1/2 = 0{,}111\ldots_3$ no. Hay tantos como sucesiones de ceros y doses, es decir, tantos como reales.

(c) La dimensión. La dimensión de Hausdorff de C es $\log 2/\log 3 = 0{,}630930$: estrictamente entre 0 (un punto) y 1 (un segmento). El conjunto de Cantor es el fractal más sencillo que existe y el primero que se estudió, medio siglo antes de que la palabra "fractal" se inventara.

(d) Para qué sirve. C es el contraejemplo universal del análisis real: compacto, no numerable, de medida nula, totalmente disconexo, sin puntos aislados, homeomorfo a $\{0,1\}^{\mathbb{N}}$ y homeomorfo a su propio producto por sí mismo. Casi cualquier afirmación plausible que uno se invente sobre conjuntos de la recta se refuta con él.

Idea clave

Qué acaba de pasar. C tiene la cantidad de puntos de toda la recta y la longitud del conjunto vacío. Esto separa dos ideas que la intuición confunde: *cuántos puntos hay* (cardinal) y *cuánto espacio ocupan* (medida). $\mathbb{Q}$ es infinito numerable y denso, de medida cero. C es no numerable y de medida cero. La medida no cuenta puntos: mide cómo están repartidos. Sin este ejemplo, la teoría de la medida del capítulo 31 parecería una formalidad; con él, es una necesidad.

Definición 29.49 (Función de Cantor o escalera del diablo)**.** Define $c : [0,1] \to [0,1]$ así: si $x \in C$ tiene desarrollo ternario $\sum a_k 3^{-k}$ con $a_k \in \{0,2\}$, pon $c(x) = \sum_{k\geq 1} (a_k/2)2^{-k}$; si $x \notin C$, entonces x cae en uno de los tercios centrales eliminados, y se define c como constante en ese intervalo, con el valor común de sus extremos.

Teorema 29.50. *La función c es continua, creciente (no estrictamente), $c(0) = 0$, $c(1) = 1$, y $c'(x) = 0$ en todo punto fuera de C, es decir, en un conjunto de medida 1. Sin embargo, c sube de 0 a 1.*

Demostración. La monotonía es inmediata del orden lexicográfico de los desarrollos. Para la continuidad: c es creciente y su imagen es todo $[0,1]$ (por la sobreyectividad de φ), y una función monótona cuya imagen es un intervalo no tiene saltos, luego es continua. Fuera de C, la función es localmente constante, así que su derivada es 0; y $[0,1] \setminus C$ tiene medida $1 - 0 = 1$. □

Ejemplo 29.51 (Valores de la escalera)**.** $c(1/3) = c(2/3) = 1/2$ (todo el tercio central $[1/3, 2/3]$ tiene imagen $1/2$); $c(1/9) = c(2/9) = 1/4$; $c(7/9) = c(8/9) = 3/4$. Y $1/4 = 0{,}020202\ldots_3$ está *en* C, con $c(1/4) = 0{,}010101\ldots_2 = 1/3$.

Cuidado

La escalera del diablo demuele el teorema fundamental del cálculo si uno lo enuncia con descuido. Es continua en $[0,1]$, derivable con derivada nula en casi todo punto, y sin embargo

$$\int_0^1 c'(x)\,\mathrm{d}x = 0 \neq 1 = c(1) - c(0).$$

No hay error: la hipótesis que falla es la *continuidad absoluta*, definida al final de este capítulo. Toda la sección de diferenciación del capítulo 31 existe para arreglar exactamente esto.

29.5. Oscilación: exactamente cuándo existe la integral de Riemann

Definición 29.52 (Oscilación)**.** Sea $f : A \to \mathbb{R}$ acotada. La oscilación de f en un conjunto $S \subseteq A$ es

$$\operatorname{osc}(f,S) = \sup_{x,y\in S} |f(x) - f(y)| = \sup_S f - \inf_S f,$$

y la oscilación en un punto x_0 es

$$\omega_f(x_0) = \lim_{\delta\to 0^+} \operatorname{osc}\left(f, (x_0-\delta, x_0+\delta)\cap A\right).$$

El límite existe porque la cantidad decrece al decrecer δ y está acotada inferiormente por 0.

Ejemplo 29.53 (Medir cuánto salta una función)**. (a) Un salto limpio.** Sea $f = \mathbf{1}_{[0,\infty)}$. En cualquier entorno de 0, la función toma los valores 0 y 1, luego $\omega_f(0) = 1$: la oscilación mide exactamente el tamaño del salto. En cualquier otro punto vale 0.

(b) Una oscilación infinita. Para $f(x) = \sin(1/x)$ con $f(0) = 0$, en todo entorno del origen la función recorre $[-1,1]$ entera, luego $\omega_f(0) = 2$. Para $f(x) = 1/x$ con $f(0) = 0$, la oscilación en el origen es infinita.

(c) Una que se amansa. Para $f(x) = x\sin(1/x)$ con $f(0) = 0$, en el entorno $(-\delta,\delta)$ los valores están en $[-\delta,\delta]$, luego $\omega_f(0) \le 2\delta$ y el ínfimo es 0: la función es continua en el origen pese a oscilar infinitas veces.

(d) La función de Thomae. Vale $1/q$ en el racional p/q irreducible y 0 en los irracionales. Su oscilación es $1/q$ en cada racional p/q y 0 en cada irracional, de donde resulta continua exactamente en los irracionales y discontinua exactamente en los racionales. Es el ejemplo que hace falta tener presente en el resto de la sección.

Proposición 29.54. *f es continua en x_0 si y solo si $\omega_f(x_0) = 0$. Además, para cada $\varepsilon > 0$ el conjunto $\{x : \omega_f(x) \ge \varepsilon\}$ es cerrado.*

Demostración. La equivalencia es la definición de continuidad traducida: $\omega_f(x_0) < \varepsilon$ significa que hay un entorno donde f varía menos de ε. Para lo segundo, sea x un punto con $\omega_f(x) < \varepsilon$; entonces existe δ con $\operatorname{osc}(f,(x-\delta,x+\delta)) < \varepsilon$, y para todo y en ese intervalo se tiene $\omega_f(y) \le \operatorname{osc}(f,(x-\delta,x+\delta)) < \varepsilon$. Luego $\{\omega_f < \varepsilon\}$ es abierto. □

Ejemplo 29.55 (Los conjuntos $\{\omega_f \geq \varepsilon\}$)**. (a) Son cerrados.** El conjunto $F_\varepsilon = \{x : \omega_f(x) \geq \varepsilon\}$ es siempre cerrado: si $\omega_f(x) < \varepsilon$, hay un entorno donde la variación de f es menor que ε, y en todo ese entorno la oscilación es también menor. El complementario es abierto.

(b) El conjunto de discontinuidades. Como f es discontinua en x exactamente cuando $\omega_f(x) > 0$, el conjunto de discontinuidades es $\bigcup_{n\geq 1} F_{1/n}$: una unión numerable de cerrados, es decir, un F_σ. Ésta es la observación que decide todo lo que sigue.

(c) Con la función de Thomae. $F_{1/n}$ es el conjunto de racionales con denominador $\leq n$, que es discreto y cerrado; la unión de todos ellos son los racionales, que efectivamente son un F_σ y no son cerrados.

(d) Con la de Dirichlet. $\mathbf{1}_{\mathbb{Q}}$ tiene oscilación 1 en *todo* punto, luego $F_1 = \mathbb{R}$ y el conjunto de discontinuidades es toda la recta. Es el caso extremo: una función puede ser discontinua en absolutamente todos los puntos.

Idea clave

Intuición primero. La oscilación mide "cuánto salta f si me acerco todo lo que quiera a este punto". Una discontinuidad de salto tiene oscilación igual al tamaño del salto; una discontinuidad esencial como $\sin(1/x)$ en 0 tiene oscilación 2. Convertir la continuidad en un número, y no en una frase con cuantificadores, permite *medir* el conjunto de discontinuidades. Ese es el paso que faltaba para responder "¿qué funciones son integrables?" con un si y solo si.

Teorema 29.56 (Criterio de Lebesgue de integrabilidad Riemann)**.** *Sea $f : [a, b] \to \mathbb{R}$ acotada y sea D su conjunto de puntos de discontinuidad. Entonces f es integrable Riemann si y solo si D tiene medida cero.*

Demostración. Escribe $D = \bigcup_{k\geq 1} D_{1/k}$ con $D_\varepsilon = \{x : \omega_f(x) \geq \varepsilon\}$, que es cerrado y acotado, luego compacto.

($\Leftarrow$) Supón $|D| = 0$ y sea $\varepsilon > 0$. Entonces $D_\varepsilon \subseteq D$ tiene medida cero y es compacto, así que se cubre con un número *finito* de intervalos abiertos de longitud total $< \varepsilon$. Sea K el complemento de esa unión dentro de $[a, b]$: es compacto y en cada punto suyo $\omega_f < \varepsilon$, luego (por compacidad, tomando un subrecubrimiento finito de los entornos correspondientes) existe $\delta > 0$ tal que todo subintervalo de K de longitud $< \delta$ tiene oscilación $< 2\varepsilon$. Toma una partición P de paso menor que δ que además refine los extremos de los intervalos del recubrimiento. Separa los subintervalos de P en dos grupos. Los que están en K contribuyen a $U(f, P) - L(f, P)$ a lo sumo $2\varepsilon(b - a)$; los que cortan al recubrimiento contribuyen a lo sumo $2M\varepsilon$, con $M = \sup |f|$. Así $U - L \leq 2\varepsilon(b - a) + 2M\varepsilon$, que es arbitrariamente pequeño: criterio de Riemann satisfecho.

($\Rightarrow$) Supón f integrable y fija $\varepsilon > 0$. Basta ver que D_ε tiene medida cero para todo ε, pues D es la unión numerable de los $D_{1/k}$. Dado $\eta > 0$, elige P con $U(f, P) - L(f, P) < \varepsilon\eta$. Los subintervalos de P cuyo interior corta a D_ε tienen oscilación $\geq \varepsilon$, luego si L es la suma de sus longitudes,

$$\varepsilon L \leq \sum_{\text{esos } i} \operatorname{osc}(f, I_i)\, |I_i| \leq U(f, P) - L(f, P) < \varepsilon\eta,$$

de donde $L < \eta$. Los finitos puntos de la partición se cubren con longitud total tan pequeña como se quiera. Luego $|D_\varepsilon| = 0$. □

Ejemplo 29.57 (La función de Thomae: continua exactamente en los irracionales)**.** Define $t : [0,1] \to \mathbb{R}$ por

$$t(x) = \begin{cases} 1/q, & \text{si } x = p/q \text{ en forma irreducible, } q \geq 1, \\ 0, & \text{si } x \text{ es irracional.} \end{cases}$$

Entonces t es discontinua en todo racional y continua en todo irracional; en consecuencia es integrable Riemann con $\int_0^1 t = 0$.

Demostración. En un racional p/q el valor es $1/q > 0$, pero hay irracionales arbitrariamente cerca donde t vale 0: la oscilación es $\geq 1/q$, y hay discontinuidad. En un irracional x_0 y dado $\varepsilon > 0$, solo hay *finitos* racionales $p/q \in [0,1]$ con $q \leq 1/\varepsilon$; sea δ la distancia mínima de x_0 a ese conjunto finito (positiva, porque x_0 no está en él). Si $|x - x_0| < \delta$ entonces o bien x es irracional y $t(x) = 0$, o bien $x = p/q$ con $q > 1/\varepsilon$ y $t(x) = 1/q < \varepsilon$. Luego t es continua en x_0. El conjunto de discontinuidades es $\mathbb{Q} \cap [0,1]$, numerable, de medida cero: t es integrable. Y como toda suma inferior es 0, la integral vale 0. □

Cuidado

La función de Dirichlet $\mathbf{1}_{\mathbb{Q}}$ es el ejemplo hermano y el contraste exacto: es discontinua en *todos* los puntos, su conjunto de discontinuidades es $[0,1]$, de medida 1, y no es integrable Riemann ($U = 1$, $L = 0$ para toda partición). Sin embargo sí es integrable Lebesgue, con integral 0. La diferencia entre las dos funciones no está en "cuántos" racionales hay —son los mismos— sino en si la función se aplasta o no al acercarse a ellos.

Regla: decidir si una función acotada es Riemann-integrable

El criterio de Lebesgue lo resuelve todo: *una función acotada en $[a,b]$ es Riemann-integrable si y solo si su conjunto de discontinuidades tiene medida cero*. Convertirlo en procedimiento:

1. **Comprobar primero que está acotada.** Es hipótesis, no conclusión. Una función no acotada en $[a,b]$ nunca es Riemann-integrable en sentido propio, por muy pocas discontinuidades que tenga: $1/\sqrt{x}$ en $(0,1]$ ya queda fuera y hay que tratarla como impropia.

2. **Localizar el conjunto D de discontinuidades.** No basta con decir "es discontinua"; hay que describir *dónde*. Este es el paso que realmente exige pensar.

3. **Medir D.** Si D es finito, numerable, o de medida cero por cualquier otra razón: integrable. Si D contiene un intervalo, o tiene medida positiva: no integrable. Los casos frontera —conjuntos no numerables de medida cero, como el de Cantor— caen del lado bueno.

4. **Atajos que evitan todo el análisis anterior.** Continua en $[a,b] \Rightarrow$ integrable ($D = \varnothing$). Monótona $\Rightarrow$ integrable (sus discontinuidades son a lo sumo numerables). Continua salvo en un número finito de saltos $\Rightarrow$ integrable. Estos tres cubren la inmensa mayoría de los casos prácticos.

5. **Si hace falta la definición, usar el criterio de las sumas.** Integrable $\iff$ para todo $\varepsilon > 0$ existe una partición P con $U(f,P) - L(f,P) < \varepsilon$. En términos de oscilación, $\sum_i \operatorname{osc}(f, I_i)\,|I_i| < \varepsilon$: los intervalos donde f oscila mucho han de poder cubrirse con longitud total pequeña.

6. **Cerrar con el valor.** Demostrar que la integral existe y calcularla son cosas distintas. Si f coincide con una función conocida salvo en un conjunto de medida cero, la integral es la misma, porque cambiar f en medida cero no altera ni la integrabilidad ni el valor.

7. **Control.** Si $f \geq 0$ y no es idénticamente nula en ningún intervalo, la integral puede perfectamente ser 0; pero si f es *continua*, $f \geq 0$ y $\int f = 0$, entonces $f \equiv 0$. Confundir estas dos afirmaciones es un error clásico.

Ejemplo 29.58 (Cuatro funciones al borde de la integrabilidad)**. (a) Dirichlet: no integrable.** $f = \mathbf{1}_{\mathbb{Q}}$ en $[0,1]$. Es discontinua en *todos* los puntos, y $[0,1]$ no tiene medida cero: no es Riemann-integrable. Directamente con sumas: toda partición tiene $U(f,P) = 1$ y $L(f,P) = 0$, porque cada subintervalo contiene racionales e irracionales. La diferencia es 1 para toda partición, y no hay refinamiento que la baje. (En el sentido de Lebesgue sí es integrable, con integral 0; véase el capítulo 31.)

(b) Thomae: integrable, y sorprendentemente. $f(x) = 1/q$ si $x = p/q$ en forma irreducible, $f(x) = 0$ si x es irracional. Esta función es discontinua exactamente en los racionales y *continua* en todos los irracionales, y $\mathbb{Q}$ es numerable, luego de medida cero: es integrable. Su integral vale 0, porque para cada ε solo hay finitos puntos con $f(x) \geq \varepsilon$ y se pueden encerrar en intervalos de longitud total tan pequeña como se quiera. Es el contraejemplo perfecto a la idea de que "discontinua en un conjunto denso" implique "no integrable".

(c) Oscilación salvaje, pero en un punto. $f(x) = \sin(1/x)$ para $x \in (0,1]$ y $f(0) = 0$. Oscila infinitas veces cerca del origen, pero está acotada por 1 y es continua en todo punto salvo el 0. Un solo punto tiene medida cero: es integrable. Lo que arruina la integrabilidad de Riemann no es oscilar mucho, sino oscilar mucho *en un conjunto grande*.

(d) El Cantor gordo: no integrable. Sea C el conjunto de Cantor obtenido retirando en el paso m los intervalos centrales de longitud 4^{-m}, cuya medida es $1/2$. Entonces $\mathbf{1}_C$ es discontinua en todo punto de C (que no tiene interior pero sí medida $1/2$): no es Riemann-integrable. Compárese con el conjunto de Cantor usual, de medida cero, para el cual $\mathbf{1}_C$ *sí* es integrable con integral 0 pese a que C es no numerable. Dos conjuntos con la misma topología y respuestas opuestas: lo que decide es la medida, no el cardinal.

Cuidado

El teorema fundamental del cálculo tiene una hipótesis que suele darse por supuesta. Existen funciones F derivables en todo $[0,1]$ cuya derivada F' es acotada pero *no* Riemann-integrable; el ejemplo estándar es una variante de $x^2\sin(1/x^2)$ construida sobre un Cantor gordo. En esos casos la fórmula $\int_0^1 F' = F(1)-F(0)$ ni siquiera tiene sentido, porque el lado izquierdo no existe. Es una de las razones históricas por las que se necesitó la integral de Lebesgue: allí, con la hipótesis de continuidad absoluta, el teorema fundamental vuelve a ser cierto sin excepciones.

29.6. El teorema de Baire

Definición 29.59. En un espacio métrico (X,d), un conjunto A es *denso en ninguna parte* si su clausura tiene interior vacío, y es *magro* (o de primera categoría) si es unión numerable de conjuntos densos en ninguna parte.

Teorema 29.60 (Baire, 1899)**.** *Sea (X,d) un espacio métrico completo y no vacío. Si $U_1, U_2, \dots$ son abiertos densos, entonces $\bigcap_{n\ge 1} U_n$ es denso en X. Equivalentemente: X no es magro en sí mismo.*

Demostración. Sea $B_0 = B(x_0, r_0)$ una bola cualquiera; hay que ver que corta a la intersección. Como U_1 es abierto y denso, $U_1 \cap B_0$ es abierto no vacío, y contiene una bola cerrada $\overline{B(x_1,r_1)} \subseteq U_1 \cap B_0$ con $0 < r_1 < r_0/2$. Repitiendo, se construyen bolas encajadas

$$\overline{B(x_n,r_n)} \subseteq U_n \cap B(x_{n-1},r_{n-1}), \qquad 0 < r_n < r_{n-1}/2.$$

Como $r_n < 2^{-n} r_0$, la sucesión (x_n) es de Cauchy: para $m > n$, $x_m \in B(x_n, r_n)$, luego $d(x_n,x_m) < r_n \to 0$. Por completitud converge a un x, que está en todas las bolas cerradas (cada una es cerrada y contiene la cola de la sucesión). Entonces $x \in U_n$ para todo n y $x \in B_0$. □

Idea clave

Qué dice esto en cristiano. El teorema de Baire dice que un espacio completo no se puede construir amontonando una cantidad numerable de conjuntos "delgados". Es un principio de abundancia: si algo se puede escribir como intersección numerable de abiertos densos, entonces ese algo sigue siendo denso, es decir, sigue habiendo mucho. De ahí salen demostraciones de existencia que no construyen nada: se prueba que el conjunto de objetos con la propiedad buscada es residual, y por tanto no vacío. Casi toda la existencia en análisis funcional funciona así.

Ejemplo 29.61 (Existencia sin construcción)**. (a) $\mathbb{Q}$ no es intersección numerable de abiertos.** Si lo fuera, como también es unión numerable de puntos (cerrados con interior vacío), se llegaría a que $\mathbb{R}$ es unión numerable de conjuntos sin interior, contradiciendo Baire. Los irracionales, en cambio, sí lo son: son $\bigcap_{q\in\mathbb{Q}} (\mathbb{R}\setminus\{q\})$, intersección numerable de abiertos densos.

(b) La consecuencia sorprendente. No existe ninguna función $f : \mathbb{R} \to \mathbb{R}$ continua exactamente en los racionales. El conjunto de puntos de continuidad de cualquier función es siempre un G_δ, y $\mathbb{Q}$ no lo es. Es un resultado de imposibilidad demostrado sin examinar ninguna función concreta.

(c) El caso que sí existe. La función de Thomae, que vale $\frac{1}{q}$ en la fracción irreducible $\frac{p}{q}$ y 0 en los irracionales, es continua exactamente en los irracionales. Y los irracionales sí son un G_δ, como se vio en (a): no hay contradicción, todo encaja.

(d) El uso típico en análisis. Con Baire se demuestra que "casi toda" función continua en $[0, 1]$ no es derivable en ningún punto, en el sentido de que las derivables en algún punto forman un conjunto de primera categoría. Es decir: las funciones patológicas como la de Weierstrass no son excepciones raras, son la norma, y las funciones amables son las excepciones.

Cuidado

[La completitud es imprescindible] Baire falla en espacios no completos: $\mathbb{Q}$ con la distancia usual es unión numerable de puntos, y cada punto es cerrado sin interior, de modo que $\mathbb{Q}$ es de primera categoría en sí mismo. La hipótesis no es decorativa. Segundo aviso: el teorema también vale en espacios localmente compactos de Hausdorff, no solo métricos completos, y ambas versiones se usan. Y tercero: "residual" o "de segunda categoría" es una noción de abundancia *topológica* que no coincide con la de medida; existen conjuntos residuales de medida cero, así que "casi todo" significa cosas distintas según el contexto.

Idea clave

Intuición primero. Un conjunto denso en ninguna parte es "flaco": por muy repartido que esté, no llena ningún trocito de espacio. Baire dice que un espacio completo no se puede armar apilando numerablemente muchos conjuntos flacos: la completitud es una reserva de sustancia que ninguna colección numerable de flacos agota. Es un principio de existencia disfrazado de topología: para probar que existe un objeto con cierta propiedad, se demuestra que los objetos *sin* esa propiedad forman un conjunto magro, y Baire garantiza que quedan otros. Así se demuestra, sin construir ninguna, que casi toda función continua no es derivable en ningún punto.

Corolario 29.62. *$\mathbb{R}$ no es unión numerable de conjuntos densos en ninguna parte. En particular, $\mathbb{R}$ no es numerable (los puntos son densos en ninguna parte), y $\mathbb{Q}$ no es intersección numerable de abiertos.*

Demostración. Lo primero es Baire aplicado a $\mathbb{R}$, que es completo. Para lo último: si $\mathbb{Q} = \bigcap_n U_n$ con U_n abiertos, cada U_n contiene a $\mathbb{Q}$, luego es denso. Sea $\mathbb{Q} = \{q_1, q_2, \dots\}$; los abiertos $V_n = \mathbb{R} \setminus \{q_n\}$ también son densos. Por Baire, $\bigcap_n U_n \cap \bigcap_n V_n = \mathbb{Q} \setminus \mathbb{Q} = \emptyset$ debería ser denso, absurdo. □

Teorema 29.63 (No existe función continua exactamente en los racionales)**.** *Para toda $f : \mathbb{R} \to \mathbb{R}$, el conjunto de sus puntos de continuidad es un G_δ, es decir, una intersección numerable de abiertos. Como $\mathbb{Q}$ no es un G_δ, ninguna función es continua exactamente en $\mathbb{Q}$.*

Demostración. El conjunto de continuidad es $\{x : \omega_f(x) = 0\} = \bigcap_{k\geq 1}\{x : \omega_f(x) < 1/k\}$, y cada uno de esos conjuntos es abierto por la proposición de la sección anterior. □

Idea clave

Qué dice esto en cristiano. Existe una función continua exactamente en los irracionales (la de Thomae) y *no* existe ninguna continua exactamente en los racionales. La asimetría no es un capricho: el conjunto de puntos de continuidad tiene que ser intersección numerable de abiertos, y los racionales no lo son, mientras que los irracionales sí. Es un ejemplo perfecto de cómo una restricción topológica decide una pregunta que parecía puramente de construcción de ejemplos.

Ejemplo 29.64 (Qué conjuntos pueden ser de continuidad)**. (a) Los irracionales sí.** $\mathbb{R}\setminus\mathbb{Q} = \bigcap_{q\in\mathbb{Q}}(\mathbb{R}\setminus\{q\})$, una intersección numerable de abiertos: es un G_δ. Y la función de Thomae lo realiza.

(b) Los racionales no. Si $\mathbb{Q}$ fuese G_δ, entonces $\mathbb{R} = \mathbb{Q}\cup(\mathbb{R}\setminus\mathbb{Q})$ sería unión de dos conjuntos magros (de primera categoría), lo que contradice el teorema de Baire. Por tanto ninguna función real es continua exactamente en los racionales, por más ingenio que se le ponga.

(c) Un cerrado cualquiera. Dado un cerrado F, la función que vale 0 en F y $d(x,F)$ multiplicada por la indicatriz de $\mathbb{Q}$ fuera de F es continua exactamente en F y en los irracionales de fuera. Ajustando la construcción se puede realizar cualquier G_δ, de modo que el teorema es óptimo: la condición no sólo es necesaria, es suficiente.

(d) Baire en una frase. En un espacio métrico completo, una intersección numerable de abiertos densos es densa. De ahí salen éste, el principio de acotación uniforme, la existencia de funciones continuas sin derivada en ningún punto y la de series de Fourier divergentes. Es el teorema más rentable del análisis por unidad de longitud de enunciado.

Pausa para pensar

Ponlo junto: la función de Thomae *sí* es continua exactamente en los irracionales, porque $\mathbb{R}\setminus\mathbb{Q} = \bigcap_{q\in\mathbb{Q}}(\mathbb{R}\setminus\{q\})$ es un G_δ. Pero ninguna función puede ser continua exactamente en los racionales. La asimetría no viene de la aritmética ni del tamaño de los conjuntos —ambos son densos, uno numerable y el otro no— sino de la forma topológica que puede tener un conjunto de continuidad.

Regla: usar el teorema de Baire para demostrar que algo existe

Baire es la herramienta que demuestra existencia sin construir. Su uso sigue casi siempre el mismo guion.

1. **Verifica el escenario.** El teorema vale en espacios métricos *completos* (o en espacios localmente compactos de Hausdorff). En un espacio incompleto es falso: $\mathbb{Q}$ con la métrica usual es unión numerable de puntos, todos cerrados de interior vacío.

2. **Recuerda los dos enunciados equivalentes.** En un espacio completo no vacío: (i) la intersección de una familia numerable de abiertos densos es densa; (ii) el espacio no es unión numerable de cerrados de interior vacío. La segunda es la forma que se usa para las demostraciones por contradicción.

3. **Traduce «raro» a «magro».** Un conjunto es *magro* (o de primera categoría) si es unión numerable de conjuntos cuya clausura tiene interior vacío. Baire dice que un espacio completo nunca es magro en sí mismo, es decir: lo magro es genuinamente pequeño.

4. **Para demostrar existencia, prueba que el complementario es magro.** Si quieres probar que existe un objeto con la propiedad P, escribe $\{\text{no } P\}$ como unión numerable $\bigcup F_n$ y demuestra que cada F_n es cerrado con interior vacío. Baire concluye que el complementario es denso, luego el objeto no solo existe: es *típico*.

5. **Para demostrar imposibilidad, aplica la forma (ii).** Si un objeto se escribiera como unión numerable de piezas cerradas de interior vacío, Baire lo prohíbe. Así se demuestra que $\mathbb{R}$ no es numerable y que no existe función continua en los racionales y discontinua en los irracionales.

6. **Cada F_n debe ser cerrado.** Es el punto donde fallan las demostraciones. Si tu F_n no es cerrado, sustitúyelo por su clausura y comprueba que sigue teniendo interior vacío; si la clausura ya llena un abierto, hay que redefinir la descomposición.

7. **Interpreta el resultado como genericidad.** Baire no da ejemplos: da abundancia. La conclusión típica es «el conjunto de objetos con la propiedad deseada es residual», que significa «casi todos, en sentido topológico». No confundir con «casi todos» en sentido de medida: son nociones distintas y hasta incompatibles.

Ejemplo 29.65 (Cuatro existencias regaladas por Baire)**. (a) Los irracionales no son numerables.** Supongamos que $\mathbb{R} \setminus \mathbb{Q} = \{y_1, y_2, \dots\}$. Entonces

$$\mathbb{R} = \mathbb{Q} \cup (\mathbb{R} \setminus \mathbb{Q}) = \bigcup_{q \in \mathbb{Q}} \{q\} \ \cup \ \bigcup_{n} \{y_n\},$$

una unión numerable de conjuntos unipuntuales. Cada punto es cerrado y tiene interior vacío en $\mathbb{R}$. Pero $\mathbb{R}$ es completo, así que Baire lo prohíbe: contradicción. El mismo argumento demuestra de un golpe que $\mathbb{R}$ entero no es numerable, sin diagonal de Cantor.

(b) No hay función continua exactamente en $\mathbb{Q}$. Para cualquier $f : \mathbb{R} \to \mathbb{R}$, el conjunto de puntos de continuidad es siempre un G_δ, es decir, intersección numerable de abiertos. Si f fuera continua exactamente en $\mathbb{Q}$, entonces $\mathbb{Q}$ sería un G_δ, digamos $\mathbb{Q} = \bigcap_n U_n$ con U_n abierto y denso. Pero entonces

$$\varnothing = \Big(\bigcap_n U_n\Big) \cap \Big(\bigcap_{q \in \mathbb{Q}} (\mathbb{R} \setminus \{q\})\Big)$$

sería intersección numerable de abiertos densos, que por Baire debería ser densa y en particular no vacía. Contradicción. En cambio la recíproca sí ocurre: la función de Thomae, que vale $1/q$

en p/q irreducible y 0 en los irracionales, es continua exactamente en los irracionales, y los irracionales sí son un G_δ.

(c) Casi toda función continua no es derivable en ningún punto. En el espacio $C[0,1]$ con la norma del supremo, que es completo, sea

$$F_n = \Big\{ f : \exists\, x \in [0,1] \text{ tal que } |f(x+h) - f(x)| \le n|h| \text{ para todo } h \text{ admisible} \Big\}.$$

Se demuestra que cada F_n es cerrado (por compacidad de $[0,1]$ y convergencia uniforme) y de interior vacío (toda función se perturba uniformemente poco añadiéndole una sierra de pendiente enorme). Como toda función derivable en algún punto pertenece a algún F_n, el conjunto de funciones derivables en al menos un punto es magro. Por Baire, su complementario es residual: las funciones continuas sin derivada en ningún punto no son una rareza construida a mano, son la regla. La función de Weierstrass $\sum 2^{-n}\cos(3^n \pi x)$ es solo el ejemplo explícito más famoso; obsérvese que en $x = 0$ vale $\sum 2^{-n} = 2$, y en $x = 1/2$ todos los términos se anulan.

(d) Principio de acotación uniforme. Sea $\{T_\alpha\}$ una familia de operadores lineales continuos de un Banach X en un normado Y, con $\sup_\alpha \|T_\alpha x\| < \infty$ para cada x fijo. Definiendo $F_n = \{x : \|T_\alpha x\| \le n \ \forall \alpha\}$, cada F_n es cerrado por continuidad, y su unión es todo X por hipótesis. Baire fuerza que algún F_{n_0} tenga interior no vacío, y de ahí sale $\sup_\alpha \|T_\alpha\| < \infty$: la acotación puntual implica acotación uniforme. Es el teorema de Banach–Steinhaus, y su demostración es literalmente Baire aplicado con una descomposición bien elegida.

Cuidado

Magro y de medida nula no son lo mismo, y ni siquiera van juntos. El intervalo $[0,1]$ se descompone en la unión de un conjunto magro y uno de medida nula: existe un conjunto residual (topológicamente enorme) que tiene medida cero, y un conjunto magro de medida uno. Cuando un texto dice «casi todo», hay que mirar en qué sentido lo dice, porque las dos nociones de genericidad pueden apuntar en direcciones opuestas.

Pausa para pensar

Baire demuestra que existen funciones continuas sin derivada en ningún punto, pero no exhibe ninguna. ¿Es eso satisfactorio? Depende de para qué. Si buscas una fórmula, no: hay que ir a Weierstrass o a Takagi. Si buscas entender el paisaje, Baire dice algo que ningún ejemplo concreto puede decir: que la derivabilidad es la excepción, y que nuestra intuición de que «las funciones continuas son suaves salvo en unos pocos picos» está exactamente al revés. El ejemplo explícito y el argumento de categoría responden a preguntas distintas.

29.7. Espacios completos y el principio de contracción

Teorema 29.66 (Punto fijo de Banach)**.** *Sea (X,d) métrico completo no vacío y $T : X \to X$ con*

$$d(T(x),T(y)) \le k\, d(x,y) \qquad \text{para todo } x, y, \ \text{con } 0 \le k < 1.$$

Entonces T tiene un único punto fijo x^*, y para todo $x_0 \in X$ la sucesión $x_{n+1} = T(x_n)$ converge a x^* con la cota de error

$$d(x_n, x^*) \le \frac{k^n}{1-k}\, d(x_1, x_0).$$

Demostración. De $d(x_{n+1}, x_n) \le k\, d(x_n, x_{n-1})$ sale $d(x_{n+1}, x_n) \le k^n d(x_1, x_0)$. Para $m > n$, la desigualdad triangular da

$$d(x_m, x_n) \le \sum_{j=n}^{m-1} k^j d(x_1, x_0) \le \frac{k^n}{1-k} d(x_1, x_0) \xrightarrow[n\to\infty]{} 0,$$

luego (x_n) es de Cauchy y converge a cierto x^*. T es continua (es lipschitziana), así que $T(x^*) = \lim T(x_n) = \lim x_{n+1} = x^*$. Si y^* fuera otro punto fijo, $d(x^*, y^*) = d(Tx^*, Ty^*) \le k\, d(x^*, y^*)$ obliga a $d = 0$. Haciendo $m \to \infty$ en la desigualdad se obtiene la cota. □

Ejemplo 29.67 (Resolver $\cos x = x$ sin fórmula)**.** Sea $X = [0,1]$, completo por ser cerrado en $\mathbb{R}$, y $T(x) = \cos x$. Como cos es decreciente en $[0,1]$ y $\cos 1 \approx 0{,}5403$, $\cos 0 = 1$, se tiene $T([0,1]) \subseteq [0{,}54, 1] \subseteq [0,1]$. Además $|T'(x)| = |\sin x| \le \sin 1 \approx 0{,}8415 < 1$, y por el teorema del valor medio T es contracción con $k = \sin 1$. Luego la ecuación $\cos x = x$ tiene exactamente una solución. Iterando desde $x_0 = 1$:

$$0{,}540302,\ 0{,}857553,\ 0{,}654290,\ 0{,}793480,\ 0{,}701369,\ 0{,}763960,\ \ldots$$

que oscila y converge lentamente (k está cerca de 1) al valor $x^* = 0{,}7390851332\ldots$

Observación 29.68. Este es el mismo teorema que demuestra la existencia y unicidad de soluciones de ecuaciones diferenciales en el capítulo 12 (Picard–Lindelöf): allí X es un espacio de funciones continuas con la norma del supremo, completo, y T es el operador integral asociado a la ecuación. También es el motor del método de Newton y de la convergencia de los métodos iterativos del álgebra lineal numérica. Un solo argumento de tres líneas, tres teoremas grandes.

Regla: aplicar el principio de contracción

El teorema del punto fijo de Banach no solo demuestra que la solución existe: la *calcula*, y además dice de antemano cuántos pasos harán falta. Para usarlo hay que verificar tres cosas y solo tres.

1. **Escribe el problema como $x = f(x)$.** Casi cualquier ecuación admite muchas reescrituras de esta forma, y no todas sirven: la elección de f decide si el método converge. Si una no funciona, prueba otra.
2. **Elige un conjunto cerrado X y comprueba que f lo respeta: $f(X) \subseteq X$.** Este paso se olvida constantemente y es imprescindible: si las iteraciones se escapan del conjunto, el teorema no dice nada.
3. **Comprueba la contracción:** $|f(x) - f(y)| \le k|x - y|$ con $k < 1$ *fijo*, el mismo para todos los puntos. Con f derivable basta acotar $|f'| \le k < 1$ en todo X, por

el teorema del valor medio. Que $|f'(x)| < 1$ en cada punto no basta si el supremo vale 1.

4. **Concluye y calcula.** Existe un único punto fijo, y la sucesión $x_{n+1} = f(x_n)$ converge a él desde *cualquier* punto de partida en X.

5. **Usa la cota de error a priori:**

$$|x_n - x^*| \le \frac{k^n}{1-k}\,|x_1 - x_0|.$$

Con ella se decide cuántas iteraciones hacen falta *antes* de empezar a iterar, lo cual es justo lo que un método numérico necesita.

6. **Interpreta k.** Es el factor por el que se reduce el error en cada paso: $k = 0{,}5$ gana un bit por iteración (unas tres iteraciones por cifra decimal); $k = 0{,}9$ necesita más de veinte iteraciones por cifra. Un k cercano a 1 no invalida el teorema, pero lo vuelve inútil en la práctica.

Ejemplo 29.69 (La ecuación $x = \cos x$)**.** No tiene solución en términos elementales, pero el principio de contracción la resuelve.

Paso 2. Tomamos $X = [0,1]$. Como cos es decreciente ahí, $\cos([0,1]) = [\cos 1,\ 1] = [0{,}540,\ 1] \subset [0,1]$. ✓

Paso 3. $f'(x) = -\sin x$, y en $[0,1]$ se tiene $|f'(x)| \le \sin 1 \approx 0{,}8415 < 1$. Es una contracción con $k = 0{,}8415$.

Paso 4. Existe un único x^* con $\cos x^* = x^*$. Iterando desde $x_0 = 0{,}5$:

n	1	3	5	12
x_n	0,8776	0,8027	0,7682	0,7372

El valor exacto es $x^* = 0{,}739085\ldots$ Las iteraciones oscilan alrededor de él —porque $f' < 0$— y se acercan despacio.

Paso 6: por qué despacio. Con $k = 0{,}8415$ y $|x_1 - x_0| = 0{,}3776$, la cota a priori tras 10 pasos es

$$\frac{0{,}8415^{10}}{1 - 0{,}8415} \cdot 0{,}3776 \approx 0{,}42,$$

una cota malísima. Y en efecto tras doce iteraciones apenas tenemos dos cifras correctas. El teorema garantiza la convergencia, pero un k tan próximo a 1 la hace lenta: para diez cifras harían falta más de cien iteraciones. Esta es exactamente la razón por la que existen métodos como el de Newton (capítulo 14).

Ejemplo 29.70 (Una contracción rápida: la raíz cuadrada)**.** Reescribamos $x^2 = 2$ como $x = f(x)$ con $f(x) = \frac{1}{2}(x + \frac{2}{x})$, sobre $X = [1,2]$.

Paso 2. $f(1) = 1{,}5$, $f(2) = 1{,}5$ y el mínimo de f en $[1,2]$ es $\sqrt{2} \approx 1{,}414$; así que $f([1,2]) = [1{,}414,\ 1{,}5] \subset [1,2]$. ✓

Paso 3. $f'(x) = \frac{1}{2} - \frac{1}{x^2}$, que en $[1,2]$ varía entre $-\frac{1}{2}$ y $\frac{1}{4}$: contracción con $k = \frac{1}{2}$.

Paso 4 y 5. Desde $x_0 = 1{,}5$:

$$x_1 = 1{,}41\overline{6}, \qquad x_2 = 1{,}4142156862\ldots, \qquad x_3 = 1{,}4142135623746\ldots, \qquad x_4 = 1{,}414213562373095,$$

frente a $\sqrt{2} = 1{,}414213562373095\ldots$ Cuéntense las cifras correctas: x_1 acierta dos decimales, x_2 acierta cinco, x_3 acierta once y x_4 acierta las quince que caben en la aritmética de doble precisión. El número de cifras buenas se duplica en cada paso.

La lección. La cota de la contracción con $k = \frac{1}{2}$ predecía ganar una cifra binaria por paso, es decir unas cuatro cifras decimales tras trece iteraciones; la realidad ha sido muchísimo mejor. El motivo es que $f'(\sqrt{2}) = \frac{1}{2} - \frac{1}{2} = 0$: cerca del punto fijo la contracción no es lineal sino *cuadrática*, y el número de cifras correctas se duplica en cada paso. El principio de contracción da una garantía *suficiente*; cuando la derivada se anula en el punto fijo, la convergencia real es mucho más rápida que la garantizada. Esta observación es, ni más ni menos, el método de Newton.

29.8. Una función continua sin derivada en ningún punto

Teorema 29.71 (Weierstrass, 1872)**.** *Sean $0 < a < 1$ y b un entero impar con $ab > 1 + \frac{3\pi}{2}$. Entonces*

$$W(x) = \sum_{n=0}^{\infty} a^n \cos(b^n \pi x)$$

converge uniformemente en $\mathbb{R}$, es continua en todo punto, y no es derivable en ningún punto.

Demostración de la continuidad y esquema de la no derivabilidad. La convergencia uniforme es el criterio M de Weierstrass: $|a^n \cos(b^n \pi x)| \le a^n$ y $\sum a^n < \infty$. Un límite uniforme de continuas es continuo, luego W es continua.

Para la no derivabilidad, fija x_0 y $m \in \mathbb{N}$. Escribe $b^m x_0 = \alpha_m + \xi_m$ con α_m entero y $\xi_m \in (-\frac{1}{2}, \frac{1}{2}]$, y toma

$$h_m = \frac{1 - \xi_m}{b^m},$$

que tiende a 0 y tiene el signo de $1 - \xi_m > 0$. Separa la serie del cociente incremental en los términos $n < m$ y $n \ge m$. Para $n < m$, el teorema del valor medio da

$$\Big| \sum_{n<m} a^n \frac{\cos(b^n \pi (x_0 + h_m)) - \cos(b^n \pi x_0)}{h_m} \Big| \le \sum_{n<m} a^n b^n \pi = \pi \frac{(ab)^m - 1}{ab - 1} < \frac{\pi (ab)^m}{ab - 1}.$$

Para $n \ge m$, la elección de h_m y la paridad de b hacen que *todos* los términos tengan el mismo signo, y el primero de ellos ya aporta

$$\Big| \sum_{n \ge m} a^n \frac{\cos(b^n \pi (x_0 + h_m)) - \cos(b^n \pi x_0)}{h_m} \Big| \ge \frac{2}{3} (ab)^m.$$

Por tanto, el cociente incremental es en valor absoluto mayor que

$$\left(\frac{2}{3} - \frac{\pi}{ab - 1} \right) (ab)^m,$$

y la hipótesis $ab > 1 + \frac{3\pi}{2}$ es exactamente la que hace positivo ese paréntesis. Como $(ab)^m \to \infty$, el cociente incremental no tiene límite finito: W no es derivable en x_0, y x_0 era arbitrario. $\square$

Idea clave

Qué dice esto en cristiano. Durante siglo y medio todo el mundo dio por supuesto que una función continua tenía que ser derivable salvo en unos pocos puntos malos: la intuición del trazo suave lo exigía. Weierstrass demostró en 1872 que eso es falso, exhibiendo una función continua en toda la recta y derivable en *ningún* punto. La receta es sumar cosenos con frecuencias que crecen geométricamente y amplitudes que decrecen geométricamente: las amplitudes bajan lo bastante para que la suma converja y sea continua, pero las frecuencias suben más deprisa aún, de modo que la pendiente nunca se estabiliza.

Ejemplo 29.72 (La función monstruo, con números)**. (a) Los parámetros.** Con $a = \frac{1}{2}$ y $b = 13$ se cumple la condición $ab = 6{,}5 > 1 + \frac{3\pi}{2} = 5{,}712389$, luego $W(x) = \sum_{n\geq 0} 2^{-n}\cos(13^n \pi x)$ es continua y no derivable en ningún punto.

(b) Por qué es continua. Cada término está acotado por $M_n = 2^{-n}$ y $\sum 2^{-n} = 2$ converge. Por el criterio M de Weierstrass, la serie converge uniformemente en toda la recta, y un límite uniforme de funciones continuas es continuo. La continuidad es la parte fácil y sale gratis.

(c) Por qué no es derivable. Los cocientes incrementales en el origen, con $h = 10^{-2}$, 10^{-3}, 10^{-4}, 10^{-5}, 10^{-6}, valen aproximadamente $-27{,}95$, $-229{,}00$, $-2570{,}07$, $-6168{,}31$ y $-66\,761{,}52$. No se estabilizan en ningún número: crecen sin control. Cada vez que se mira más de cerca aparecen oscilaciones nuevas de la misma pendiente relativa.

(d) La intuición geométrica. La gráfica es autosemejante: al ampliar cualquier trozo se vuelve a ver la misma clase de dentado. Nunca se llega a un tramo que parezca recto, que es exactamente lo que se necesitaría para tener tangente. Su dimensión fractal es $2 + \frac{\ln a}{\ln b}$, un número entre 1 y 2: es más que una curva y menos que una superficie.

Cuidado

[No es una rareza aislada, es lo normal] La reacción de la época fue de rechazo; Hermite habló de "esta plaga lamentable de funciones sin derivada". Pero el teorema de Baire demuestra que las funciones continuas no derivables en ningún punto forman un conjunto residual en $C[0,1]$: son la abrumadora mayoría, y las funciones suaves son las excepciones. Segundo aviso: la condición sobre ab es la que Weierstrass necesitaba para su demostración; Hardy la rebajó después a $ab \geq 1$, así que el rango de parámetros válidos es mayor de lo que el enunciado sugiere. Y tercero: continua no derivable en ningún punto no es lo mismo que no monótona en ningún intervalo, aunque esta función cumpla también eso.

Idea clave

Qué está pasando geométricamente. Cada sumando es un coseno con amplitud a^n (que se encoge) y frecuencia b^n (que crece más rápido de lo que la amplitud se encoge). La gráfica tiene arrugas dentro de arrugas dentro de arrugas, sin escala en la que se vea

recta. Derivar es preguntar "¿a qué se parece esto de cerca?", y la respuesta es: a sí mismo, con más arrugas. La condición $ab > 1 + 3\pi/2$ es la contabilidad exacta de la carrera entre amplitud y frecuencia. Con $a = 1/2$ y $b = 13$ se tiene $ab = 6{,}5 > 5{,}713$: sirve.

Observación 29.73. En 1872 esto fue un escándalo. Ampère había "demostrado" que toda función continua es derivable salvo en puntos aislados, y la intuición geométrica de la época daba por hecho que una curva sin cortes tiene tangente casi siempre. Hoy se sabe algo más fuerte, y por Baire: en el espacio $C[0,1]$ con la norma del supremo —que es completo—, el conjunto de funciones que son derivables en *algún* punto es magro. Es decir, las funciones derivables en algún punto son la excepción, y las de Weierstrass son la regla. La única razón por la que nos parecen raras es que son difíciles de escribir.

29.9. Familias de funciones: Dini y Arzelà–Ascoli

Teorema 29.74 (Dini)**.** *Sea K compacto y $f_n : K \to \mathbb{R}$ continuas con $f_n \to f$ puntualmente, f continua, y la convergencia monótona ($f_n \le f_{n+1}$ para todo n, o bien $\ge$). Entonces la convergencia es uniforme.*

Demostración. Supón $f_n \uparrow f$ y sea $g_n = f - f_n \ge 0$, continua, decreciente a 0 puntualmente. Dado $\varepsilon > 0$, los conjuntos $U_n = \{x : g_n(x) < \varepsilon\}$ son abiertos (continuidad), crecientes (monotonía) y cubren K (convergencia puntual). Por compacidad, un número finito los cubre, y por ser crecientes uno solo: $K = U_N$. Entonces $0 \le g_n < \varepsilon$ en todo K para $n \ge N$. □

Ejemplo 29.75 (Dini y sus cuatro hipótesis)**. (a) Funciona.** $f_n(x) = x^n$ en $[0, \frac{1}{2}]$: continuas, decrecientes en n, límite 0 continuo, dominio compacto. La convergencia es uniforme, y en efecto $\sup_{[0,1/2]} x^n = 2^{-n} \to 0$.

(b) Sin límite continuo, falla. Las mismas $f_n(x) = x^n$ en $[0,1]$ cumplen todas las hipótesis salvo que el límite es discontinuo en 1. Y la convergencia no es uniforme: $\sup_{[0,1]} |x^n - f(x)| = 1$ para todo n.

(c) Sin compacidad, falla. $f_n(x) = 1/(1+nx)$ en $(0,1]$ decrece a 0, que es continuo, y sin embargo $\sup f_n = 1$ para todo n: no hay convergencia uniforme. El intervalo abierto es lo que rompe el argumento, porque el punto problemático se escapa.

(d) Sin monotonía, falla. La joroba viajera f_n que vale 1 en un pico estrecho alrededor de $1/n$ y 0 fuera converge puntualmente a 0 en $[0,1]$ compacto, con límite continuo, y $\sup f_n = 1$ siempre. Las cuatro hipótesis son indispensables y cada una tiene su contraejemplo.

Cuidado

Las tres hipótesis son necesarias. Sin compacidad: $f_n(x) = x^n$ en $[0,1)$ crece a $0\ldots$ no, crece hacia 0 solo si se reordena; el ejemplo estándar es $f_n(x) = \min(nx, 1)$ en $(0,1]$, monótona y con límite continuo 1, pero no uniforme. Sin continuidad del límite: $f_n(x) = x^n$ en $[0,1]$ es monótona decreciente con límite discontinuo, y la convergencia

no es uniforme. Sin monotonía: la "joroba viajera" f_n = triángulo de altura 1 y base $[1/(n+1), 1/n]$ tiende puntualmente a 0 en $[0,1]$ y no uniformemente.

Definición 29.76 (Equicontinuidad)**.** Una familia $\mathcal{F}$ de funciones $K \to \mathbb{R}$ es *equicontinua* si para cada $\varepsilon > 0$ existe $\delta > 0$, *el mismo para todas las funciones de la familia*, tal que $|x-y| < \delta \Rightarrow |f(x) - f(y)| < \varepsilon$ para toda $f \in \mathcal{F}$. Es *uniformemente acotada* si existe M con $\|f\|_\infty \le M$ para toda $f \in \mathcal{F}$.

Teorema 29.77 (Arzelà–Ascoli)**.** *Sea $K \subseteq \mathbb{R}$ compacto. Una sucesión (f_n) en $C(K)$ uniformemente acotada y equicontinua tiene una subsucesión uniformemente convergente. (Es decir: en $C(K)$, cerrado + acotado + equicontinuo = compacto.)*

Demostración. Sea $\{q_1, q_2, \dots\}$ un subconjunto numerable denso de K. La sucesión $(f_n(q_1))$ está acotada en $\mathbb{R}$, luego por Bolzano–Weierstrass tiene una subsucesión convergente $f_n^{(1)}$. De esta se extrae otra $f_n^{(2)}$ que converge también en q_2, y así sucesivamente. La *diagonal* $g_n = f_n^{(n)}$ es, a partir del índice k, subsucesión de $f^{(k)}$, luego $(g_n(q_k))$ converge para todo k.

Falta pasar de la convergencia en un denso a la uniforme. Dado $\varepsilon > 0$, toma δ de la equicontinuidad para $\varepsilon/3$. Por compacidad, finitos puntos $q_{i_1}, \dots, q_{i_r}$ del denso cumplen que todo $x \in K$ dista menos de δ de alguno. Elige N tal que $|g_n(q_{i_j}) - g_m(q_{i_j})| < \varepsilon/3$ para $n, m \ge N$ y todo $j \le r$. Entonces, para x cualquiera y q el punto cercano correspondiente,

$$|g_n(x) - g_m(x)| \le |g_n(x) - g_n(q)| + |g_n(q) - g_m(q)| + |g_m(q) - g_m(x)| < \varepsilon.$$

Luego (g_n) es de Cauchy en la norma del supremo, y $C(K)$ es completo. □

Idea clave

Qué dice esto en cristiano. Dini y Arzelà–Ascoli son dos maneras de conseguir gratis lo que normalmente cuesta: uniformidad y compacidad. Dini dice que si la convergencia es monótona, sobre un compacto y con límite continuo, entonces automáticamente es uniforme; se ahorra comprobarlo. Arzelà–Ascoli dice cuándo una familia de funciones tiene subsucesiones convergentes: hace falta que estén uniformemente acotadas y que sean equicontinuas, es decir, que ninguna sea más empinada que las demás. Es el Bolzano–Weierstrass de los espacios de funciones.

Ejemplo 29.78 (Cuándo se aplican y cuándo no)**. (a) Dini funcionando.** Sea $f_n(x) = \frac{x}{n}$ en $[0,1]$: es monótona decreciente en n, el compacto es $[0,1]$ y el límite 0 es continuo. Dini garantiza uniformidad, y en efecto $\sup |f_n| = \frac{1}{n} \to 0$.

(b) Dini fallando por el límite. Sea $f_n(x) = x^n$ en $[0,1]$: también es monótona y el compacto es el mismo, pero el límite es discontinuo (vale 0 para $x < 1$ y 1 en $x = 1$). Dini no se aplica, y de hecho la convergencia no es uniforme. En $[0;\, 0{,}9]$, en cambio, el límite es continuo y sí hay uniformidad: $\sup = 0{,}9^n$, que vale 0,348678 para $n = 10$ y 0,005154 para $n = 50$.

(c) Una familia equicontinua. Todas las funciones con $|f'| \leq 1$ (Lipschitz de constante 1) forman una familia equicontinua: dado ε, el mismo $\delta = \varepsilon$ sirve para todas a la vez. Si además están acotadas, Arzelà–Ascoli da subsucesiones uniformemente convergentes.

(d) Una que no lo es. La familia $\{\sin(nx)\}$ está uniformemente acotada por 1, pero no es equicontinua: las derivadas valen $n\cos(nx)$ y crecen sin límite, así que el δ que sirve para $\sin x$ no sirve para $\sin(100x)$. Y en efecto ninguna subsucesión de $\sin(nx)$ converge uniformemente. La acotación sola no basta.

Cuidado

[Las hipótesis de Dini son cuatro y todas hacen falta] Monotonía, compacidad, continuidad de cada f_n y continuidad del límite. Quitar la última rompe el teorema, como en (b); quitar la compacidad también, pues $f_n(x) = \frac{x}{n}$ en $[0, \infty)$ es monótona con límite continuo y no converge uniformemente. Segundo aviso: en Arzelà–Ascoli, la equicontinuidad es lo que sustituye a la acotación de las derivadas, y es más débil: hay familias equicontinuas sin derivada. Y tercero: la conclusión es que hay una *subsucesión* convergente, no que converja la sucesión entera; es exactamente el mismo matiz que en Bolzano–Weierstrass.

Idea clave

Intuición primero. En $\mathbb{R}^n$, acotado más cerrado da compacto (Heine–Borel). En espacios de funciones eso es falso: la joroba viajera está acotada y no tiene subsucesión convergente. Lo que falta es una condición que impida que las funciones "se escapen hacia lo infinitamente empinado". Equicontinuidad es exactamente eso: una cota común a la rapidez con que todas pueden cambiar. Con ella, la dimensión infinita se comporta como la finita. Es el teorema que hace funcionar los métodos de compacidad en ecuaciones diferenciales (Peano) y en cálculo de variaciones.

Regla: demostrar convergencia uniforme, o demostrar que no la hay

La distinción entre convergencia puntual y uniforme es la que decide si el límite hereda las buenas propiedades. El procedimiento es corto pero hay que seguirlo entero.

1. **Hallar primero el límite puntual.** Se fija x, se hace $n \to \infty$ y se obtiene $f(x)$. Sin este paso no hay nada que discutir: la convergencia uniforme es convergencia puntual *más* una condición extra.

2. **Calcular $M_n = \sup_x |f_n(x) - f(x)|$.** Toda la cuestión se reduce a este número. Si $M_n \to 0$, hay convergencia uniforme; si $M_n \not\to 0$, no la hay. Es un criterio, no una heurística: sirve igual para probar que sí como para probar que no.

3. **Calcular el supremo derivando.** Para cada n fijo, se busca el máximo de $|f_n - f|$ como función de x con la derivada de siempre. El punto crítico suele depender de n y *migrar*: ver hacia dónde huye ese máximo es ver por qué falla la uniformidad.

4. **Atajo para demostrar que NO es uniforme: buscar una sucesión x_n.** Basta

exhibir puntos x_n con $|f_n(x_n) - f(x_n)| \geq c > 0$ para todo n. No hace falta el supremo exacto, solo una cota inferior que no se vaya a cero.

5. **Atajo estructural para demostrar que no es uniforme.** Si cada f_n es continua y el límite f *no* lo es, la convergencia no puede ser uniforme, porque el límite uniforme de continuas es continuo. Este argumento resuelve el caso más famoso en una línea.

6. **Para series de funciones, el test M de Weierstrass.** Si $|u_n(x)| \leq M_n$ para todo x y $\sum M_n < \infty$, la serie $\sum u_n$ converge uniforme y absolutamente. Es la herramienta de trabajo diaria; su limitación es que solo detecta convergencia *absoluta*, y hay series uniformemente convergentes que no lo son.

7. **Recordar qué se compra con la uniformidad.** El límite de continuas es continua; se puede intercambiar lím con $\int$; y, con la hipótesis adicional de que f_n' converja uniformemente, también con la derivada. Sin uniformidad, los tres intercambios pueden fallar.

8. **Recordar el atajo de Dini.** Si f_n son continuas en un *compacto*, la convergencia es monótona y el límite es continuo, entonces es automáticamente uniforme. Las tres hipótesis son necesarias: quitar cualquiera de ellas produce un contraejemplo.

Ejemplo 29.79 (Cuatro sucesiones de funciones)**. (a) El caso clásico: $f_n(x) = x^n$ en $[0,1]$.** El límite puntual es $f(x) = 0$ para $x < 1$ y $f(1) = 1$: discontinuo. Por el paso 5, la convergencia no es uniforme, y se acabó. Con el supremo se ve igual: $M_n = \sup_{[0,1)} x^n = 1$ para todo n. En cambio, en $[0, a]$ con $a < 1$ se tiene $M_n = a^n \to 0$: *sí* es uniforme. La uniformidad depende del dominio, no solo de las funciones.

(b) El supremo que huye: $f_n(x) = nxe^{-nx^2}$ en $[0,1]$. Puntualmente, para $x > 0$ fijo, la exponencial gana y $f_n(x) \to 0$; y $f_n(0) = 0$. El límite es $f \equiv 0$, continuo, así que el atajo del paso 5 no sirve. Derivando, $f_n'(x) = ne^{-nx^2}(1 - 2nx^2) = 0$ da $x_n = \frac{1}{\sqrt{2n}}$, y

$$M_n = f_n(x_n) = n \cdot \frac{1}{\sqrt{2n}}\, e^{-1/2} = \sqrt{\frac{n}{2}}\, e^{-1/2} \longrightarrow \infty.$$

No hay convergencia uniforme, y de la manera más espectacular: el máximo no solo no baja, sino que crece sin cota mientras se refugia cada vez más cerca del origen. Consecuencia inmediata en las integrales:

$$\int_0^1 f_n = \frac{1 - e^{-n}}{2} \longrightarrow \frac{1}{2}, \qquad \text{pero} \qquad \int_0^1 f = 0.$$

Ahí está el intercambio lím $\leftrightarrow \int$ fallando, con toda la masa escapándose hacia el origen sin desaparecer nunca.

(c) Uniforme sin esfuerzo: $f_n(x) = \frac{\sin(nx)}{n}$ en $\mathbb{R}$. El límite puntual es 0 y $M_n = \sup |f_n| = \frac{1}{n} \to 0$: convergencia uniforme en toda la recta. Pero obsérvese que $f_n'(x) = \cos(nx)$ no converge a nada: la convergencia uniforme de las funciones no dice absolutamente nada sobre las derivadas. Es la asimetría fundamental entre integrar (que suaviza) y derivar (que amplifica).

(d) Una serie por el test M. $\sum_{n\geq 1} \frac{\sin(nx)}{n^2}$. Como $\left|\frac{\sin nx}{n^2}\right| \leq \frac{1}{n^2}$ para todo x y $\sum \frac{1}{n^2} = \frac{\pi^2}{6} < \infty$, la serie converge uniformemente en $\mathbb{R}$ y su suma es continua. La serie derivada $\sum \frac{\cos nx}{n}$, en cambio, ya no admite el test M porque $\sum \frac{1}{n}$ diverge, y de hecho su suma no es continua: tiene un salto en $x = 0$.

Pausa para pensar

Hay una manera de entender la diferencia sin fórmulas. En la convergencia puntual, cada x tiene derecho a su propio N: uno puede esperar tanto como quiera, pero el tiempo de espera depende del punto. En la convergencia uniforme, existe un N que sirve *a la vez* para todos los puntos. La frase "para todo ε existe N tal que para todo x" frente a "para todo ε y para todo x existe N" resume el capítulo entero: el orden de los cuantificadores es la única diferencia, y es toda la diferencia.

Regla: decidir si una familia de funciones es relativamente compacta (Arzelà–Ascoli)

En dimensión finita, "cerrado y acotado" basta para tener compacidad. En espacios de funciones eso es falso, y hace falta una hipótesis más: la equicontinuidad. Saber verificarla es lo que convierte muchos teoremas de existencia en dos líneas.

1. **Sitúa el problema en $C(K)$.** La familia $\mathcal{F}$ debe vivir en $C(K)$ con K compacto, dotado de la norma del supremo $\|f\|_\infty = \max_{x\in K} |f(x)|$. Convergencia en esa norma es convergencia uniforme: eso es lo que el teorema va a producir.

2. **Comprueba la acotación uniforme.** Necesitas una sola constante M con $|f(x)| \leq M$ para *toda* $f \in \mathcal{F}$ y todo $x \in K$. Que cada f esté acotada por separado no sirve de nada; la constante tiene que ser común.

3. **Comprueba la equicontinuidad.** Para cada $\varepsilon > 0$ debe existir un $\delta > 0$, *independiente de* f, tal que $|x - y| < \delta$ implica $|f(x) - f(y)| < \varepsilon$ para toda $f \in \mathcal{F}$. La palabra clave es "independiente de f": ese es todo el contenido del teorema.

4. **Usa la cota de la derivada como atajo.** Si todas las funciones de $\mathcal{F}$ son derivables y existe L con $|f'| \leq L$ para toda $f \in \mathcal{F}$, entonces por el teorema del valor medio $|f(x) - f(y)| \leq L|x - y|$, y basta tomar $\delta = \varepsilon/L$. Una cota uniforme de la derivada implica equicontinuidad, y es la forma más rápida de verificarla.

5. **Aplica el teorema.** Arzelà–Ascoli: $\mathcal{F} \subseteq C(K)$ es relativamente compacta (es decir, toda sucesión de $\mathcal{F}$ tiene una subsucesión uniformemente convergente) si y solo si $\mathcal{F}$ es uniformemente acotada y equicontinua. Si además $\mathcal{F}$ es cerrada, es compacta.

6. **Para negar la compacidad, exhibe el fallo en un punto.** Basta encontrar un x_0 y una sucesión $f_n \in \mathcal{F}$ tal que las oscilaciones de f_n cerca de x_0 no se controlen con un δ común. En la práctica: si los gráficos se ponen cada vez más verticales en algún punto, no hay equicontinuidad.

7. **Cobra el teorema de existencia.** El uso típico es: construir una sucesión aproximada, comprobar que es uniformemente acotada y equicontinua, extraer una subsucesión convergente, y demostrar que el límite resuelve el problema. Así se demuestran Peano para ecuaciones diferenciales y la existencia de geodésicas minimizantes.

Ejemplo 29.80 (Cuatro familias examinadas, dos compactas y dos no)**. (a)** $f_n(x) = x^n$ **en** $[0,1]$**: acotada pero no equicontinua.** La acotación uniforme es inmediata: $0 \le x^n \le 1$, luego $M = 1$ sirve para todas. Pero la equicontinuidad falla en $x_0 = 1$. Tomemos $\varepsilon = \frac{1}{2}$ y veamos que ningún δ funciona: dado $\delta > 0$, elegimos $y = 1 - \delta/2$ y observamos que $f_n(1) - f_n(y) = 1 - (1-\delta/2)^n \to 1$ cuando $n \to \infty$, de modo que para n grande esa diferencia supera $\frac{1}{2}$ aunque $|1-y| < \delta$. No hay δ común. Coherentemente, x^n converge puntualmente a la función discontinua que vale 0 en $[0,1)$ y 1 en 1, y *ninguna* subsucesión converge uniformemente, porque el límite uniforme de continuas sería continuo.

(b) $g_n(x) = \sin(nx)$ **en** $[0,\pi]$**: el mismo fallo, en todas partes.** De nuevo $M = 1$. Pero $g_n'(x) = n\cos(nx)$ tiene norma n, que no está acotada uniformemente, y el paso 4 no da nada. De hecho la equicontinuidad falla en *cada* punto: entre x y $x + \pi/n$ la función recorre todo el intervalo $[-1,1]$, y $\pi/n \to 0$. Ninguna subsucesión de $\sin(nx)$ converge uniformemente; ni siquiera puntualmente en casi ningún punto.

(c) $h_n(x) = \dfrac{x}{1+nx^2}$ **en** $[0,1]$**: sí es relativamente compacta.** Acotación: el máximo de h_n se alcanza en $x = 1/\sqrt{n}$ y vale $\dfrac{1}{2\sqrt{n}}$, que para $n = 1, 4, 25, 100$ da 0,5; 0,25; 0,1; 0,05. Luego $|h_n| \le \frac{1}{2}$ para toda n: la constante $M = \frac{1}{2}$ es común. Equicontinuidad por el paso 4:

$$h_n'(x) = \frac{1 - nx^2}{(1+nx^2)^2},$$

y poniendo $u = nx^2 \ge 0$ resulta $|h_n'| = \dfrac{|1-u|}{(1+u)^2} \le 1$ para todo $u \ge 0$. Una cota $L = 1$ válida para toda n: la familia es equicontinua con $\delta = \varepsilon$. Por Arzelà–Ascoli hay subsucesiones uniformemente convergentes, y en este caso la sucesión entera converge uniformemente a 0, pues $\|h_n\|_\infty = \frac{1}{2\sqrt{n}} \to 0$.

(d) La bola unidad de $C[0,1]$ **no es compacta.** Es el ejemplo que separa dimensión finita de dimensión infinita. La bola $\{f : \|f\|_\infty \le 1\}$ es cerrada y acotada, pero contiene a las g_n del apartado (b) divididas por nada —ya están acotadas por 1—, y por tanto no es equicontinua ni relativamente compacta. Este es el teorema de Riesz en acción: la bola unidad cerrada de un espacio normado es compacta *si y solo si* la dimensión es finita. Arzelà–Ascoli dice exactamente qué hay que añadir para recuperar la compacidad.

(e) El teorema de Peano, en cuatro líneas. Para $y' = F(x,y)$ con F continua y acotada por M en un rectángulo, las poligonales de Euler y_h satisfacen $|y_h(x) - y_h(z)| \le M|x-z|$ para todo paso h: son uniformemente acotadas y equicontinuas con la *misma* constante de Lipschitz M, por el paso 4. Arzelà–Ascoli entrega una subsucesión uniformemente convergente, y pasando al límite en la forma integral $y(x) = y_0 + \int_{x_0}^{x} F(t, y(t))\,\mathrm{d}t$ se obtiene una solución. Nótese que aquí no se ha supuesto que F sea Lipschitz en y: por eso el resultado da existencia pero *no* unicidad, a diferencia de Picard–Lindelöf.

Cuidado

Equicontinuidad no es lo mismo que continuidad uniforme de cada miembro. Cada x^n es uniformemente continua en $[0,1]$ —lo es toda función continua en un compacto—, y sin embargo la familia no es equicontinua. La diferencia está en el orden de los cuantificadores: continuidad uniforme de cada f permite que δ dependa de f; equicontinuidad no. Cuando dudes, escribe los cuantificadores en orden y mira dónde aparece f.

Pausa para pensar

Compárese este teorema con el de Bolzano–Weierstrass en $\mathbb{R}^n$: toda sucesión acotada tiene subsucesión convergente. En $C(K)$ eso es falso, y Arzelà–Ascoli identifica la hipótesis exacta que hay que añadir. La moraleja general es que en dimensión infinita la compacidad es un bien escaso y hay que ganársela: cada teorema importante del análisis funcional —Ascoli, Rellich–Kondrachov, Banach–Alaoglu— es una receta distinta para fabricar compacidad, y casi siempre el precio es o bien una regularidad extra (una derivada acotada) o bien una topología más débil. Los teoremas de existencia del análisis moderno son, casi sin excepción, teoremas de compacidad disfrazados.

29.10. Aproximación: Bernstein, Weierstrass y Stone

Teorema 29.81 (Weierstrass, demostración de Bernstein)**.** *Toda $f \in C[0,1]$ es límite uniforme de polinomios. En concreto, los polinomios de Bernstein*

$$B_n f(x) = \sum_{k=0}^{n} f\left(\frac{k}{n}\right) \binom{n}{k} x^k (1-x)^{n-k}$$

convergen a f uniformemente en $[0,1]$.

Demostración. Sea S_n una variable aleatoria binomial de parámetros n y x; entonces $B_n f(x) = \mathbb{E}[f(S_n/n)]$ y $\mathbb{E}[S_n/n] = x$, $\operatorname{Var}(S_n/n) = x(1-x)/n \le 1/(4n)$. Dado $\varepsilon > 0$, la continuidad uniforme de f en el compacto $[0,1]$ da δ con $|u-v| < \delta \Rightarrow |f(u)-f(v)| < \varepsilon/2$. Separa según $|S_n/n - x| < \delta$ o no:

$$|B_n f(x) - f(x)| \le \mathbb{E}|f(S_n/n) - f(x)| \le \frac{\varepsilon}{2} + 2\|f\|_\infty \Pr\left(|S_n/n - x| \ge \delta\right).$$

Por la desigualdad de Chebyshev, esa probabilidad es $\le \frac{x(1-x)}{n\delta^2} \le \frac{1}{4n\delta^2}$, cota independiente de x. Tomando $n > \|f\|_\infty/(\varepsilon\delta^2)$ se consigue $|B_n f - f| < \varepsilon$ uniformemente. □

Ejemplo 29.82 (Bernstein sobre $f(x) = x^2$)**.** Aquí el cálculo se hace exacto. Usando $\sum k\binom{n}{k}x^k(1-x)^{n-k} = nx$ y $\sum k^2\binom{n}{k}x^k(1-x)^{n-k} = nx(1-x) + n^2x^2$,

$$B_n(x^2) = x^2 + \frac{x(1-x)}{n}.$$

Para $n=2$ da $\frac{x^2}{2} + \frac{x}{2}$; para $n=3$, $\frac{2x^2}{3} + \frac{x}{3}$; para $n=5$, $\frac{4x^2}{5} + \frac{x}{5}$. El error máximo es $\max_{[0,1]} x(1-x)/n = 1/(4n)$: la convergencia es uniforme pero lenta, del orden $1/n$. Bernstein aproxima bien y *robustamente*, no rápido.

Teorema 29.83 (Stone–Weierstrass)**.** *Sea K un espacio métrico compacto y $\mathcal{A} \subseteq C(K,\mathbb{R})$ una subálgebra (cerrada bajo sumas, productos y múltiplos escalares) que contiene a las constantes y* separa puntos*: para $x \neq y$ existe $f \in \mathcal{A}$ con $f(x) \neq f(y)$. Entonces $\mathcal{A}$ es densa en $C(K,\mathbb{R})$ con la norma del supremo.*

Idea de la demostración. Primero se ve que la clausura $\overline{\mathcal{A}}$ es cerrada bajo $|\cdot|$: basta aproximar $t \mapsto |t|$ en $[-\|f\|, \|f\|]$ por polinomios (que es el teorema de Weierstrass anterior, o una sucesión explícita $p_{n+1}(t) = p_n(t) + \frac{1}{2}(t^2 - p_n(t)^2)$ que crece a $|t|$, con Dini para la uniformidad). De $|\cdot|$ se obtienen máx y mín vía $\max(f,g) = \frac{f+g+|f-g|}{2}$. Luego, dados $x \neq y$ y valores α, β, la separación de puntos más las constantes permiten fabricar $h \in \mathcal{A}$ con $h(x) = \alpha$, $h(y) = \beta$. Fijado $\varepsilon > 0$ y f objetivo, para cada par de puntos se construye una función que aproxima bien cerca de ellos, y dos argumentos de compacidad (uno con mín, otro con máx) las pegan en una función de $\overline{\mathcal{A}}$ a distancia $< \varepsilon$ de f. □

Ejemplo 29.84 (Tres consecuencias inmediatas)**.** (i) Los polinomios son densos en $C[a,b]$: el álgebra generada por 1 y x separa puntos. (ii) Los polinomios en dos variables son densos en $C(K)$ para $K \subseteq \mathbb{R}^2$ compacto. (iii) Los polinomios trigonométricos son densos en las funciones continuas 2π-periódicas: aquí K es la circunferencia, y $\cos x, \sin x$ separan sus puntos. Este último es el teorema que sostiene toda la teoría de series de Fourier del capítulo 31.

Cuidado

Sobre $\mathbb{C}$ el teorema es *falso* tal cual: el álgebra de polinomios en z separa puntos del disco cerrado y contiene las constantes, pero su clausura son las funciones holomorfas en el interior y continuas hasta el borde, no todas las continuas ($\bar{z}$ no se aproxima). La versión compleja correcta exige además que $\mathcal{A}$ sea cerrada bajo conjugación. Esa asimetría es la misma rigidez de la holomorfía que estudia el capítulo 30.

29.11. Reordenar series: el teorema de Riemann

Definición 29.85. $\sum a_n$ converge *absolutamente* si $\sum |a_n| < \infty$, y *condicionalmente* si converge pero $\sum |a_n| = \infty$.

Teorema 29.86 (Las absolutamente convergentes se pueden reordenar)**.** *Si $\sum a_n$ converge absolutamente y $\sigma : \mathbb{N} \to \mathbb{N}$ es una biyección, entonces $\sum a_{\sigma(n)}$ converge al mismo valor.*

Demostración. Sea $S = \sum a_n$ y $\varepsilon > 0$. Elige N con $\sum_{n>N} |a_n| < \varepsilon$. Sea M tan grande que $\{1,\dots,N\} \subseteq \sigma(\{1,\dots,M\})$. Para $m \geq M$, la diferencia entre $\sum_{n\leq m} a_{\sigma(n)}$ y $\sum_{n\leq N} a_n$ solo contiene términos de índice $> N$, luego es $< \varepsilon$ en valor absoluto. Así $|\sum_{n\leq m} a_{\sigma(n)} - S| < 2\varepsilon$. □

Teorema 29.87 (Riemann, 1854)**.** *Si $\sum a_n$ converge condicionalmente, entonces para todo $L \in \mathbb{R} \cup \{\pm\infty\}$ existe un reordenamiento cuya suma es L. También existen reordenamientos que oscilan sin converger.*

Demostración. Sean $p_n = \max(a_n, 0)$ y $q_n = \max(-a_n, 0)$, de modo que $a_n = p_n - q_n$ y

$|a_n| = p_n + q_n$. Si ambas series $\sum p_n$ y $\sum q_n$ convergieran, $\sum |a_n|$ convergería; si solo una divergiera, $\sum a_n$ divergería. Luego $\sum p_n = \sum q_n = +\infty$. Además $a_n \to 0$, luego $p_n, q_n \to 0$.

Dado L finito, procede así: suma términos positivos, en su orden original, hasta pasar por encima de L (posible porque $\sum p_n = \infty$); después suma términos negativos hasta bajar por debajo de L (posible porque $\sum q_n = \infty$); después positivos otra vez, y así indefinidamente. Cada cambio de dirección ocurre a distancia de L menor que el último término usado, y esos términos tienden a 0. Luego las sumas parciales convergen a L. Para $L = +\infty$, se pasa por encima de 1, luego de 2, etc., intercalando un solo negativo cada vez. □

Ejemplo 29.88 (La armónica alternada, reordenada)**.** Se sabe que $\sum_{n\geq 1} \frac{(-1)^{n+1}}{n} = \ln 2 \approx 0{,}693147$. Reordena tomando dos positivos por cada negativo:

$$1 + \frac{1}{3} - \frac{1}{2} + \frac{1}{5} + \frac{1}{7} - \frac{1}{4} + \frac{1}{9} + \frac{1}{11} - \frac{1}{6} + \cdots$$

Esta serie converge a $\frac{3}{2}\ln 2 \approx 1{,}039721$. (La suma de los primeros 200 000 bloques da 1,0397195, coherente con $\frac{3}{2}\ln 2 = 1{,}0397208$.)

Por qué sale $\frac{3}{2}\ln 2$. Sea $H_n = \sum_{k\leq n} 1/k$ y $H_n = \ln n + \gamma + o(1)$. La suma parcial tras n bloques es

$$\Big(\sum_{k=1}^{2n} \frac{1}{2k-1}\Big) - \frac{1}{2}H_n = \Big(H_{4n} - \tfrac{1}{2}H_{2n}\Big) - \tfrac{1}{2}H_n.$$

Sustituyendo la asintótica, los γ se cancelan $(1 - \frac{1}{2} - \frac{1}{2} = 0)$ y queda

$$\ln(4n) - \tfrac{1}{2}\ln(2n) - \tfrac{1}{2}\ln n = \ln 4 - \tfrac{1}{2}\ln 2 + (1 - \tfrac{1}{2} - \tfrac{1}{2})\ln n = 2\ln 2 - \tfrac{1}{2}\ln 2 = \tfrac{3}{2}\ln 2. \quad \square$$

Idea clave

La moraleja. La suma de infinitos términos *no* es una operación conmutativa: es un límite de sumas parciales, y el orden decide qué sumas parciales se están mirando. La convergencia absoluta es exactamente la hipótesis bajo la cual el infinito se comporta como una suma finita. Por eso, en el capítulo 31, todos los teoremas de intercambio de límite e integral piden dominación: son la versión continua de esta misma exigencia.

Regla: decidir si una serie converge, y si lo hace absoluta o condicionalmente

Un orden de ataque que resuelve prácticamente cualquier serie de examen.

1. **Mirar el término general.** Si $a_n \not\to 0$, la serie diverge y se acabó en un renglón. Es lo primero que hay que comprobar siempre. El recíproco es falso —$\sum 1/n$ diverge con $a_n \to 0$—, así que este paso solo sirve para descartar.

2. **Pasar a $|a_n|$ y decidir la convergencia absoluta.** Aquí sirven todos los criterios de términos positivos, en este orden de coste: comparación con las series patrón, cociente, raíz, integral.

   - Patrones que hay que saber de memoria: $\sum n^{-p}$ converge $\iff$ $p > 1$; $\sum r^n$

converge $\iff |r|<1$; $\sum \frac{1}{n(\ln n)^p}$ converge $\iff p>1$.

- Cociente y raíz: si el límite es <1 converge, si es >1 diverge, y si es exactamente 1 *no dicen nada* y hay que cambiar de criterio.
- Comparación por paso al límite: si $a_n/b_n \to L \in (0,\infty)$, las dos series hacen lo mismo. Es la herramienta más rápida: se compara con la serie patrón que tenga el mismo comportamiento asintótico.

3. **Si $\sum |a_n|$ converge, terminado:** la serie converge absolutamente, y por tanto converge. Además se la puede reordenar libremente sin que cambie la suma.

4. **Si $\sum |a_n|$ diverge, todavía puede converger.** Ahora toca buscar convergencia condicional, y hay dos herramientas.

   - *Leibniz*, para series alternadas $\sum (-1)^n b_n$: si b_n es decreciente y $b_n \to 0$, converge. Y de regalo, el error al truncar es menor que el primer término omitido.
   - *Dirichlet*, más general: si las sumas parciales de $\sum c_n$ están acotadas y b_n decrece a cero, entonces $\sum b_n c_n$ converge. Es lo que salva a las series con $\sin n$ o $\cos n$, cuyas sumas parciales están acotadas.

5. **Nombrar el resultado con precisión.** Absolutamente convergente, condicionalmente convergente o divergente son tres respuestas distintas, y la del medio es la interesante.

6. **Recordar lo que está en juego.** Si la convergencia es condicional, el teorema de reordenación de Riemann dice que reordenando los términos se puede obtener *cualquier* suma que uno quiera, o $\pm\infty$. La suma de una serie condicionalmente convergente no es una propiedad del conjunto de sus términos, sino del orden en que se escriben.

7. **Controles.** Si se ha usado el criterio del cociente y ha dado 1, la respuesta obtenida por ese camino no vale nada. Y una comprobación numérica de las primeras veinte sumas parciales cuesta un minuto y detecta errores de signo.

Ejemplo 29.89 (Cinco series decididas)**. (a) Condicional pura.** $\sum_{n\ge 1} \frac{(-1)^{n+1}}{n}$. En valor absoluto es la armónica, que diverge. Alternada con $b_n = 1/n$ decreciente a cero: por Leibniz converge, y su suma es $\ln 2 \approx 0{,}6931$. Es *condicionalmente* convergente, y por tanto el material bruto del teorema de Riemann: reordenándola se puede llegar a π, a -17 o a $+\infty$.

(b) La frontera logarítmica. $\sum_{n\ge 2} \frac{1}{n\ln n}$ frente a $\sum_{n\ge 2} \frac{1}{n(\ln n)^2}$. Los términos se parecen muchísimo, y sin embargo el criterio integral las separa: $\int_2^\infty \frac{dx}{x\ln x} = [\ln\ln x]_2^\infty = \infty$, mientras que $\int_2^\infty \frac{dx}{x(\ln x)^2} = \big[-\frac{1}{\ln x}\big]_2^\infty = \frac{1}{\ln 2} < \infty$. La primera diverge y la segunda converge, con suma $\approx 1{,}92297$. Ni el cociente ni la raíz habrían servido: ambos dan exactamente 1 en los dos casos.

(c) El cociente en su terreno. $\sum_{n\ge 1} \frac{n!}{n^n}$. El cociente es

$$\frac{a_{n+1}}{a_n} = \frac{(n+1)!}{(n+1)^{n+1}} \cdot \frac{n^n}{n!} = \Big(\frac{n}{n+1}\Big)^n = \frac{1}{\big(1+\frac{1}{n}\big)^n} \longrightarrow \frac{1}{e} \approx 0{,}368 < 1.$$

Converge absolutamente; su suma vale $\approx 1{,}87985$. Este es el caso ideal para el cociente: factoriales y potencias n-ésimas se simplifican de golpe.

(d) Dirichlet en acción. $\sum_{n\geq 1} \frac{\sin n}{n}$. En valor absoluto diverge, porque $|\sin n|$ no es pequeño en promedio y la comparación acaba en la armónica. Pero las sumas parciales de $\sum \sin n$ están acotadas (se suman como una progresión geométrica compleja), y $1/n$ decrece a cero: por Dirichlet converge. Su suma es exactamente $\frac{\pi-1}{2} \approx 1{,}07080$. Nótese que no es alternada, así que Leibniz no habría bastado.

(e) Un caso que engaña. $\sum_{n\geq 1}(-1)^n \frac{n}{n+1}$. Es alternada y podría parecer un Leibniz, pero $b_n = \frac{n}{n+1} \to 1 \neq 0$: el paso 1 la mata antes de empezar. Las sumas parciales oscilan aproximadamente entre $-\frac{1}{2}$ y 0 sin acercarse a nada. Comprobar $a_n \to 0$ antes que ninguna otra cosa ahorra páginas de trabajo inútil.

Cuidado

El criterio de Leibniz exige que b_n sea *decreciente*, no solo que tienda a cero, y la hipótesis no es decorativa. Tomando $b_n = \frac{1}{n}$ para n par y $b_n = \frac{1}{n^2}$ para n impar se obtiene una sucesión que tiende a cero pero no es monótona, y la serie alternada correspondiente diverge: al separar los términos pares se recupera media serie armónica, que se lleva la suma a $-\infty$. Cuando la monotonía falle, hay que recurrir a Dirichlet o a un análisis directo.

Regla: reordenar una serie condicionalmente convergente para que sume lo que se quiera

El teorema de Riemann sobre reordenaciones no es una curiosidad: es la razón por la que "suma infinita" y "suma" no son la misma operación. Conviene saber ejecutarlo, no solo enunciarlo.

1. **Comprueba que la convergencia es condicional.** Necesitas $\sum a_n$ convergente pero $\sum |a_n| = +\infty$. Si la convergencia fuese absoluta, *toda* reordenación daría la misma suma (teorema de Dirichlet) y no habría nada que hacer.

2. **Separa las partes positiva y negativa.** Define $a_n^+ = \max(a_n, 0)$ y $a_n^- = \max(-a_n, 0)$, de modo que $a_n = a_n^+ - a_n^-$ y $|a_n| = a_n^+ + a_n^-$. La condicionalidad obliga a que *ambas* series $\sum a_n^+$ y $\sum a_n^-$ diverjan a $+\infty$: si solo divergiera una, la serie original no convergería, y si ninguna divergiera, la convergencia sería absoluta.

3. **Fija el objetivo S.** Puede ser cualquier número real, o $+\infty$, o $-\infty$, o incluso ningún límite (oscilación entre dos valores prefijados). El algoritmo es el mismo.

4. **Ejecuta el algoritmo voraz.** Suma términos positivos, en su orden original, hasta que la suma parcial *sobrepase* S por primera vez. Entonces suma términos negativos, en su orden, hasta *bajar* de S. Vuelve a positivos, y así indefinidamente. Los pasos 2 garantizan que siempre queda "combustible" para cruzar la línea.

5. **Justifica la convergencia con $a_n \to 0$.** En cada cambio de dirección, la suma parcial se ha pasado de S como mucho en el valor del último término usado. Como

$a_n \to 0$ —que es condición necesaria de convergencia— ese exceso tiende a 0, y las sumas parciales convergen a S exactamente.

6. **Comprueba que es una reordenación de verdad.** Hay que verificar que la biyección $\sigma : \mathbb{N} \to \mathbb{N}$ construida usa cada índice *una vez*: se consumen los positivos en orden y los negativos en orden, y como ambos grupos se agotan, ningún término queda fuera. Sin esta comprobación, lo que se ha hecho es cambiar la serie, no reordenarla.

7. **Contrasta con el caso absoluto.** Aplica el mismo intento a una serie absolutamente convergente y observa dónde falla: en el paso 2, porque una de las dos partes converge y el algoritmo se queda sin combustible antes de la primera vuelta. Esa es la demostración de que la reordenación libre es exclusiva de lo condicional.

Ejemplo 29.90 (La serie armónica alternada, reordenada tres veces)**.** La protagonista es

$$\sum_{n=1}^{\infty} \frac{(-1)^{n+1}}{n} = 1 - \frac{1}{2} + \frac{1}{3} - \frac{1}{4} + \cdots = \ln 2 = 0{,}693147.$$

Es condicionalmente convergente: converge por el criterio de Leibniz, pero $\sum 1/n$ diverge. Las partes positiva y negativa son $\sum \frac{1}{2k-1}$ y $\sum \frac{1}{2k}$, y ambas divergen, conforme al paso 2. Ya tenemos permiso para jugar.

(a) Dos positivos por cada negativo: la suma crece a $\frac{3}{2}\ln 2$. Reordenemos como

$$1 + \frac{1}{3} - \frac{1}{2} + \frac{1}{5} + \frac{1}{7} - \frac{1}{4} + \frac{1}{9} + \frac{1}{11} - \frac{1}{6} + \cdots$$

Agrupando de tres en tres, el bloque k-ésimo es $\frac{1}{4k-3} + \frac{1}{4k-1} - \frac{1}{2k}$. Sumando numéricamente 200 000 bloques se obtiene 1,0397195, frente al valor teórico

$$\frac{3}{2}\ln 2 = 1{,}0397208.$$

Los mismos términos, cada uno usado exactamente una vez, y la suma ha aumentado en un 50 %. La razón heurística: los positivos son ahora "el doble de frecuentes", y como ambas familias divergen, el desequilibrio de frecuencia se acumula sin límite hasta estabilizarse en el nuevo valor.

(b) El algoritmo voraz hacia $S = 1$. Apliquemos el paso 4 literalmente con objetivo $S = 1$. Empezamos sumando positivos hasta pasarnos, luego negativos hasta bajar, y así:

término añadido	suma parcial	posición respecto a $S=1$
$+1$	1,000000	alcanza
$-\frac{1}{2}$	0,500000	por debajo
$+\frac{1}{3}$	0,833333	por debajo
$+\frac{1}{5}$	1,033333	por encima
$-\frac{1}{4}$	0,783333	por debajo
$+\frac{1}{7}$	0,926190	por debajo
$+\frac{1}{9}$	1,037302	por encima
$-\frac{1}{6}$	0,870635	por debajo
$+\frac{1}{11}$	0,961544	por debajo
$+\frac{1}{13}$	1,038467	por encima
$-\frac{1}{8}$	0,913467	por debajo

Las oscilaciones alrededor de 1 tienen amplitud decreciente —0,5; 0,25; 0,167; 0,125—, exactamente los términos negativos que se van usando, y por el paso 5 el límite es 1. La reordenación es explícita: $1, -\frac{1}{2}, \frac{1}{3}, \frac{1}{5}, -\frac{1}{4}, \frac{1}{7}, \frac{1}{9}, -\frac{1}{6}, \dots$

(c) Hacia $+\infty$, y hacia ninguna parte. Para que la reordenación diverja a $+\infty$ basta cambiar el objetivo en cada vuelta: sumar positivos hasta pasar de 1, un negativo, positivos hasta pasar de 2, un negativo, positivos hasta pasar de 3, y así. Cada negativo se usa, luego es una reordenación legítima, y las sumas parciales superan cualquier cota. Y para que *no* converja ni a un número ni a infinito, alterna dos objetivos: cruza por encima de 1, luego por debajo de 0, luego por encima de 1, y repite. El límite superior es 1 y el inferior 0: la serie reordenada oscila para siempre.

(d) Por qué con series absolutamente convergentes no se puede. Tomemos $\sum \frac{(-1)^{n+1}}{n^2}$, que converge absolutamente a $\frac{\pi^2}{12}$. Aquí $\sum a_n^+ = \sum \frac{1}{(2k-1)^2} = \frac{\pi^2}{8}$ y $\sum a_n^- = \sum \frac{1}{4k^2} = \frac{\pi^2}{24}$, ambas *finitas*. El algoritmo voraz con objetivo $S=5$ agota todos los positivos sin llegar nunca a 5, y se queda sin combustible: el paso 2 falla. Y en efecto $\frac{\pi^2}{8} - \frac{\pi^2}{24} = \frac{\pi^2}{12}$, que es la suma original: no hay reordenación que la cambie.

Cuidado

Reordenar no es lo mismo que agrupar. Insertar paréntesis en una serie convergente nunca cambia la suma (las sumas parciales del agrupado son una subsucesión de las originales), y quitarlos sí puede romperlo todo: $(1-1)+(1-1)+\cdots = 0$, pero $1-1+1-1+\cdots$ no converge. Y ninguna de las dos operaciones es una reordenación. Cuando manipules una serie, di siempre cuál de las tres cosas estás haciendo —reordenar, agrupar o insertar términos—, porque las reglas son distintas para cada una.

Idea clave

La lección de fondo es que la convergencia absoluta es lo que permite tratar una serie como una *suma de un conjunto* en vez de como el límite de una sucesión de sumas parciales. Cuando $\sum |a_n| < \infty$, la expresión $\sum_{n\in A} a_n$ tiene sentido para cualquier

subconjunto A y no depende de ningún orden; ese es exactamente el contenido del teorema de Fubini–Tonelli para series dobles, y es la razón por la que en teoría de la medida se exige integrabilidad de $|f|$ y no solo de f. El teorema de Riemann es el aviso de qué se pierde cuando esa hipótesis no está: sin ella, "sumar" deja de ser una operación sobre conjuntos y pasa a ser una operación sobre listas.

29.12. Variación acotada y continuidad absoluta

Definición 29.91 (Variación total)**.** Para $f:[a,b]\to\mathbb{R}$, la variación total es

$$V_a^b(f)=\sup_P\sum_{i=1}^{n}|f(x_i)-f(x_{i-1})|,$$

donde el supremo recorre todas las particiones $a=x_0<\cdots<x_n=b$. Si $V_a^b(f)<\infty$, se dice que f es de *variación acotada*, y se escribe $f\in BV[a,b]$.

Idea clave

Intuición primero. $V_a^b(f)$ es la distancia total recorrida por un lápiz que traza la gráfica midiendo solo el movimiento vertical: sube 3, baja 2, sube 1, y la variación es 6, aunque el desplazamiento neto sea 2. Es el cuentakilómetros, no el desplazamiento. Una función es de variación acotada cuando ese cuentakilómetros no se dispara al infinito, es decir, cuando no oscila infinitas veces con amplitud que no decrece lo suficiente.

Ejemplo 29.92 (Cálculo de variaciones)**.** Si f es monótona, $V_a^b(f)=|f(b)-f(a)|$ (toda suma telescopa). Así $V_0^2(x^2)=4$. Si f es C^1, entonces $V_a^b(f)=\int_a^b|f'|$; para $\sin x$ en $[0,2\pi]$ eso da $\int_0^{2\pi}|\cos x|\mathrm{d}x=4$, que coincide con leerlo a mano: sube 1, baja 2, sube 1.

Ejemplo 29.93 (Continua pero no de variación acotada)**.** Sea

$$f(x)=\begin{cases}x\sin(1/x), & x\in(0,1],\\ 0, & x=0.\end{cases}$$

Es continua en $[0,1]$. Tomando los puntos $x_k=\frac{2}{(2k+1)\pi}$, donde $|f(x_k)|=x_k$ con signos alternos, la suma de $|f(x_k)-f(x_{k+1})|$ es del orden de $\sum 1/k$, que diverge. Luego $V_0^1(f)=\infty$. En cambio $g(x)=x^2\sin(1/x)$ sí es de variación acotada, porque g' es acotada e integrable.

Teorema 29.94 (Descomposición de Jordan)**.** *$f\in BV[a,b]$ si y solo si f es diferencia de dos funciones crecientes. En consecuencia (por el teorema de Lebesgue sobre monótonas), toda $f\in BV$ es derivable en casi todo punto y tiene a lo sumo una cantidad numerable de discontinuidades, todas de salto.*

Demostración. Si $f = g - h$ con g, h crecientes, toda suma de la definición está acotada por $(g(b) - g(a)) + (h(b) - h(a))$. Recíprocamente, define $v(x) = V_a^x(f)$, que es creciente, y observa que $v - f$ también lo es: para $x < y$,

$$(v(y) - f(y)) - (v(x) - f(x)) = V_x^y(f) - (f(y) - f(x)) \geq 0,$$

porque la variación en $[x, y]$ es al menos $|f(y) - f(x)|$. Entonces $f = v - (v - f)$. □

Idea clave

Qué dice esto en cristiano. Variación acotada significa que la función no sube y baja infinitamente: sumando todas sus subidas y todas sus bajadas se obtiene un número finito. La descomposición de Jordan dice que eso equivale a ser diferencia de dos funciones crecientes, una que recoge las subidas y otra las bajadas. La continuidad absoluta es una condición más fuerte: no solo la variación total es finita, sino que se reparte de forma controlada, sin concentrarse en conjuntos diminutos. Y esa es exactamente la condición que hace verdadero el teorema fundamental del cálculo en su forma más general.

Ejemplo 29.95 (Las tres clases, con sus fronteras)**. (a) Una continua que no es de variación acotada.** Sea $f(x) = x \sin \frac{1}{x}$ en $[0, 1]$, con $f(0) = 0$. Es continua, pero en los puntos $x = \frac{2}{(2k+1)\pi}$ alcanza alternativamente valores $\pm\frac{2}{(2k+1)\pi}$, de modo que la variación acumulada es del orden de $\sum \frac{2}{k\pi}$, que diverge. Sumando cincuenta términos ya va por 2,85, y sigue creciendo sin tope.

(b) La descomposición de Jordan. Toda función monótona es de variación acotada, con variación $|f(b) - f(a)|$. Y toda función de variación acotada se escribe como diferencia de dos crecientes: se toma $V(x)$ la variación en $[a, x]$, que es creciente, y entonces $f = V - (V - f)$, con ambos sumandos crecientes. De ahí sale que una función de variación acotada es derivable en casi todo punto, porque las monótonas lo son.

(c) La escalera de Cantor. Es continua, creciente (luego de variación acotada) y sube de 0 a 1; sin embargo su derivada vale 0 en casi todo punto, porque solo crece sobre el conjunto de Cantor, que tiene medida nula. Por tanto $\int_0^1 f' = 0 \neq 1 = f(1) - f(0)$: el teorema fundamental del cálculo *falla*.

(d) Dónde vuelve a valer. La escalera no es absolutamente continua, y ese es exactamente el motivo del fallo. El teorema fundamental en su versión definitiva dice que $f(b) - f(a) = \int_a^b f'$ ocurre si y solo si f es absolutamente continua. Toda función Lipschitz lo es, y toda absolutamente continua es de variación acotada y uniformemente continua, pero no al revés en ninguno de los dos casos.

Cuidado

[La cadena de implicaciones es estricta en cada paso] Lipschitz ⇒ absolutamente continua ⇒ variación acotada y uniformemente continua ⇒ continua. Ninguna flecha se invierte: $\sqrt{x}$ es absolutamente continua sin ser Lipschitz, la escalera de Cantor es de variación acotada sin ser absolutamente continua, y $x \sin \frac{1}{x}$ es continua sin ser de variación acotada. Segundo aviso: derivable en todo punto no implica absolutamente continua; la función

$x^2 \sin \frac{1}{x^2}$ es derivable en todas partes y su derivada no es integrable, de modo que el teorema fundamental tampoco vale para ella.

Definición 29.96 (Continuidad absoluta)**.** $f : [a, b] \to \mathbb{R}$ es *absolutamente continua* si para cada $\varepsilon > 0$ existe $\delta > 0$ tal que, para toda familia finita de intervalos disjuntos (a_i, b_i) con $\sum (b_i - a_i) < \delta$, se tiene $\sum |f(b_i) - f(a_i)| < \varepsilon$.

Idea clave

Qué dice esto en cristiano. La continuidad uniforme controla *un* intervalo pequeño; la continuidad absoluta controla *muchos a la vez*: si la suma de las longitudes de una colección de intervalos disjuntos es pequeña, la suma de los saltos de f sobre ellos también lo es. Esa condición aparentemente técnica es exactamente la que impide que la función acumule todo su crecimiento en un conjunto de medida cero, y por eso es la hipótesis correcta del teorema fundamental del cálculo.

Ejemplo 29.97 (Continuidad absoluta y sus vecinas)**. (a) Las de Lipschitz.** Si $|f(x) - f(y)| \leq M|x - y|$, entonces la suma de saltos está acotada por M por la suma de longitudes: basta tomar $\delta = \varepsilon/M$. Toda función de Lipschitz es absolutamente continua, y en particular toda función de clase C^1 en un compacto.

(b) Absolutamente continua sin ser de Lipschitz. $f(x) = \sqrt{x}$ en $[0, 1]$ es absolutamente continua (es la integral de $1/(2\sqrt{x})$, que es integrable con integral 1) y no es de Lipschitz, porque su derivada explota en el origen.

(c) La escalera del diablo. La función de Cantor es uniformemente continua, creciente y de variación acotada, y *no* es absolutamente continua: se puede recubrir el conjunto de Cantor con intervalos de longitud total tan pequeña como se quiera, y sobre ellos la función sube 1 entero. Éste es el contraejemplo que justifica toda la definición.

(d) La jerarquía. $C^1 \subset$ Lipschitz $\subset$ absolutamente continua $\subset$ variación acotada $\subset$ diferenciable en casi todo punto, y todas las inclusiones son estrictas. Situar una función en esa escalera es lo primero que hay que hacer antes de derivar o integrar con confianza.

Regla: la jerarquía de regularidad

Para funciones en $[a, b]$ valen las inclusiones *estrictas*

$$C^1 \subset \text{Lipschitz} \subset AC \subset BV \cap C \subset C.$$

Los contraejemplos que separan cada nivel:

- $|x|$ es Lipschitz pero no C^1.
- $\sqrt{x}$ en $[0, 1]$ es absolutamente continua pero no Lipschitz (pendiente infinita en 0).

- La escalera de Cantor c es continua, creciente (luego BV, con $V_0^1(c) = 1$) pero *no* absolutamente continua: el conjunto C se cubre con intervalos de longitud total $< \delta$ y aun así las imágenes suman 1.
- $x \sin(1/x)$ es continua pero no de variación acotada.

Teorema 29.98 (Teorema fundamental del cálculo, versión definitiva)**.** *$f : [a, b] \to \mathbb{R}$ es absolutamente continua si y solo si f es derivable en casi todo punto, $f' \in L^1[a, b]$, y*

$$f(x) - f(a) = \int_a^x f'(t)\,\mathrm{d}t \qquad \text{para todo } x.$$

Ejemplo 29.99 (El teorema fundamental, versión definitiva)**. (a) Lo que sí vale.** Si f es absolutamente continua, entonces f' existe en casi todo punto, es integrable Lebesgue y $f(x) - f(a) = \int_a^x f'$. Para $f(x) = \sqrt{x}$ en $[0, 1]$: $f' = 1/(2\sqrt{x})$ salvo en el origen, y $\int_0^1 \frac{\mathrm{d}t}{2\sqrt{t}} = 1 = f(1) - f(0)$. Cuadra, aunque la derivada no esté acotada.

(b) Lo que no vale sin la hipótesis. Para la función de Cantor, $F' = 0$ en casi todo punto, luego $\int_0^1 F' = 0$, mientras que $F(1) - F(0) = 1$. El teorema fundamental falla por completo, y el culpable es un conjunto de medida nula donde la derivada no existe.

(c) La otra mitad. Si $g \in L^1$, la función $G(x) = \int_a^x g$ es absolutamente continua y $G' = g$ en casi todo punto. Las dos mitades juntas dicen que "absolutamente continua" e "integral indefinida de una función L^1" son la misma clase de funciones, sin holgura.

(d) La comparación con Riemann. La versión clásica exige f' continua, o al menos integrable Riemann. La de Lebesgue no exige nada sobre la derivada salvo su existencia en casi todo punto, que viene regalada, y desplaza toda la hipótesis a la función: es más fácil de comprobar y cubre muchos más casos.

Observación 29.100. Este enunciado, demostrado en el capítulo 31, cierra el círculo que abrió la escalera del diablo. El teorema fundamental del cálculo no es válido para toda función derivable en casi todo punto: hace falta continuidad absoluta, ni más ni menos. La escalera de Cantor es el ejemplo mínimo de lo que sale mal, y la continuidad absoluta es exactamente la hipótesis que lo prohíbe.

Regla: calcular la variación total de una función

La variación total mide cuánto sube y baja una función en total, sin que las subidas y bajadas se cancelen. Es el concepto que hace funcionar la integral de Riemann–Stieltjes y la continuidad absoluta.

1. **Escribe la definición.** Para $f : [a, b] \to \mathbb{R}$,

$$V_a^b(f) = \sup_P \sum_{i=1}^{n} |f(x_i) - f(x_{i-1})|,$$

donde el supremo recorre todas las particiones $a = x_0 < \cdots < x_n = b$. Si el supremo es finito, f es de variación acotada, y se escribe $f \in BV[a, b]$.

2. **Si f es monótona, la respuesta es inmediata.** Todas las diferencias tienen el mismo signo y la suma telescopea: $V_a^b(f) = |f(b) - f(a)|$. Ninguna partición mejora esto.

3. **Si f es monótona a trozos, suma los tramos.** Localiza los puntos donde f cambia de sentido (los críticos, si f es derivable), y suma los saltos $|f(c_k) - f(c_{k-1})|$ entre cambios consecutivos. La variación es aditiva: $V_a^c + V_c^b = V_a^b$.

4. **Si f es de clase C^1, integra el valor absoluto de la derivada.**

$$V_a^b(f) = \int_a^b |f'(x)|\,\mathrm{d}x.$$

   Esta es la fórmula práctica más útil, y explica por qué la variación total es «la longitud recorrida», no «el desplazamiento neto».

5. **Si f es lipschitziana, ya está acotada.** De $|f(x) - f(y)| \leq K|x - y|$ se sigue $V_a^b(f) \leq K(b - a)$ sumando sobre cualquier partición. Toda función con derivada acotada es lipschitziana y, por tanto, BV.

6. **Para demostrar que *no* es de variación acotada, exhibe particiones.** Construye una sucesión explícita de particiones cuyas sumas diverjan. El patrón habitual: elige los puntos donde la función alcanza sus máximos y mínimos locales y comprueba que la serie de las diferencias es divergente.

7. **Usa la descomposición de Jordan como control.** Toda $f \in BV$ se escribe como diferencia de dos funciones crecientes, concretamente $f = \frac{V+f}{2} - \frac{V-f}{2}$ con $V(x) = V_a^x(f)$. Si tu función no admite tal descomposición, no puede ser de variación acotada.

Ejemplo 29.101 (Cinco variaciones totales)**. (a) Una función monótona.** Para $f(x) = x^3$ en $[-1, 2]$, que es creciente, $V_{-1}^2(f) = f(2) - f(-1) = 8 - (-1) = 9$. Verificación por la fórmula integral: $\int_{-1}^2 3x^2\mathrm{d}x = [x^3]_{-1}^2 = 9$. Coincide.

(b) Una función con dos cambios de sentido. Sea $f(x) = x^3 - 3x$ en $[-2, 2]$. Su derivada $3x^2 - 3$ se anula en $x = \pm 1$, y los valores relevantes son

$$f(-2) = -2, \quad f(-1) = 2, \quad f(1) = -2, \quad f(2) = 2.$$

La función sube de -2 a 2, baja de 2 a -2, y vuelve a subir de -2 a 2. Por tanto

$$V_{-2}^2(f) = |2 - (-2)| + |-2 - 2| + |2 - (-2)| = 4 + 4 + 4 = 12.$$

Contraste: el desplazamiento neto es $f(2) - f(-2) = 4$, muy inferior. La variación total mide el camino recorrido, no la distancia entre extremos, igual que el cuentakilómetros frente a la distancia en línea recta.

(c) Una función continua que no es BV. Sea

$$f(x) = \begin{cases} x\sin(1/x), & 0 < x \leq 1, \\ 0, & x = 0. \end{cases}$$

Es continua en todo $[0,1]$, incluido el origen, porque $|f(x)| \le x$. Pero tomemos los puntos $x_k = \dfrac{2}{(2k+1)\pi}$, donde $\sin(1/x_k) = \pm 1$ alternadamente, de modo que $f(x_k) = \pm x_k$. Entre dos consecutivos, la diferencia en valor absoluto es al menos x_k, y

$$\sum_{k} x_k = \sum_{k} \frac{2}{(2k+1)\pi} = \infty$$

por comparación con la serie armónica. Sumando los primeros 50 términos ya se supera 3,74, y con 5000 términos se pasa de 6,67; el crecimiento es logarítmico, lento pero imparable. Luego $V_0^1(f) = \infty$: continua pero no de variación acotada.

(d) El mismo ejemplo, arreglado. Sea ahora $g(x) = x^2 \sin(1/x)$ con $g(0) = 0$. Su derivada es $g'(x) = 2x\sin(1/x) - \cos(1/x)$, que está acotada por $|g'| \le 2|x| + 1 \le 3$ en $[0,1]$. Luego g es lipschitziana con constante 3 y, por el paso 5, $V_0^1(g) \le 3$. El cálculo numérico de $\int_0^1 |g'|$ da aproximadamente 1,042. El exponente 2 en vez de 1 marca toda la diferencia: amortigua las oscilaciones lo justo para que la suma converja.

(e) La escalera del diablo. La función de Cantor $c : [0,1] \to [0,1]$ es continua y creciente, con $c(0) = 0$ y $c(1) = 1$. Por el paso 2, $V_0^1(c) = 1$. Pero su derivada existe y vale 0 en casi todo punto, luego

$$\int_0^1 |c'(x)|\,\mathrm{d}x = 0 \neq 1 = V_0^1(c).$$

La fórmula integral del paso 4 *falla*. Este es precisamente el fenómeno que motiva la noción de continuidad absoluta: la igualdad $V = \int |f'|$ vale exactamente para las funciones absolutamente continuas, y la escalera del diablo es el ejemplo canónico de función continua y BV que no lo es.

Cuidado

Ser BV no implica ser continua, y ser continua no implica ser BV. Los dos ejemplos que lo demuestran son: una función escalonada con un salto, que es BV con variación igual al salto pero discontinua; y el apartado (c), continua con variación infinita. Ambas propiedades son independientes, y solo su combinación con la continuidad absoluta da el teorema fundamental del cálculo en su versión más general.

Pausa para pensar

La descomposición de Jordan dice que toda función de variación acotada es diferencia de dos crecientes. Combinada con el teorema de Lebesgue de que toda función monótona es derivable en casi todo punto, se obtiene de inmediato: *toda función de variación acotada es derivable en casi todo punto*. Compáralo con el apartado (c) de la regla de Baire, donde casi ninguna función continua era derivable en ningún punto. La moraleja: la variación acotada, no la continuidad, es la hipótesis que garantiza derivabilidad. El control sobre cuánto oscila una función es mucho más fuerte que el control sobre si da saltos.

Regla: decidir si una función es absolutamente continua y aplicar el teorema fundamental

La continuidad absoluta es la hipótesis exacta bajo la cual "derivar e integrar se deshacen". Ni la continuidad ni la derivabilidad en casi todo punto bastan, y conviene tener un procedimiento para distinguir los casos.

1. **Escribe la definición con cuidado.** f es absolutamente continua en $[a,b]$ si para cada $\varepsilon>0$ existe $\delta>0$ tal que, para *cualquier* colección finita de intervalos disjuntos (x_k,y_k) contenidos en $[a,b]$,
$$\sum_k (y_k-x_k)<\delta \quad\Longrightarrow\quad \sum_k |f(y_k)-f(x_k)|<\varepsilon.$$
La diferencia con la continuidad uniforme es que aquí se controlan *muchos* intervalos a la vez, no uno solo.

2. **Busca primero un atajo por Lipschitz.** Si $|f(x)-f(y)|\le M|x-y|$, entonces $\sum|f(y_k)-f(x_k)|\le M\sum(y_k-x_k)$ y basta $\delta=\varepsilon/M$. Toda función Lipschitz es absolutamente continua; en particular toda f de clase C^1 en un compacto.

3. **Usa la caracterización integral.** f es absolutamente continua en $[a,b]$ si y solo si existe $g\in L^1[a,b]$ con
$$f(x)=f(a)+\int_a^x g(t)\,\mathrm{d}t \quad \text{para todo } x\in[a,b],$$
y en tal caso $g=f'$ en casi todo punto. Este es el criterio operativo: si sabes exhibir la primitiva, has terminado.

4. **Cobra el teorema fundamental en su versión de Lebesgue.** Si f es absolutamente continua, entonces f es derivable en casi todo punto, $f'\in L^1$ y
$$f(b)-f(a)=\int_a^b f'(t)\,\mathrm{d}t.$$
Sin continuidad absoluta esta igualdad puede fallar, y falla.

5. **Recuerda la cadena de implicaciones.** $C^1\Rightarrow$ Lipschitz $\Rightarrow$ absolutamente continua $\Rightarrow$ de variación acotada y continua $\Rightarrow$ derivable en casi todo punto. Ninguna de las flechas se puede invertir, y para cada una hay un contraejemplo estándar.

6. **Para negar la continuidad absoluta, ataca por la variación o por la integral.** Dos vías: o bien demuestras que f no es de variación acotada (y entonces no puede ser absolutamente continua, por el paso 5), o bien compruebas que $\int_a^b f'\neq f(b)-f(a)$ (y entonces el paso 4 lo prohíbe).

7. **Usa Banach–Zarecki si nada de lo anterior decide.** f es absolutamente continua si y solo si es continua, de variación acotada y cumple la propiedad N de Luzin: la imagen de todo conjunto de medida nula tiene medida nula. Es la caracterización que separa los casos difíciles.

Ejemplo 29.102 (Cinco funciones en la escalera de la regularidad)**. (a) $f(x)=\sqrt{x}$ en $[0,1]$: absolutamente continua sin ser Lipschitz.** La derivada $f'(x)=\dfrac{1}{2\sqrt{x}}$ no está acotada cerca de 0, así que el atajo del paso 2 no sirve. Pero sí está en L^1:

$$\int_0^1 \frac{\mathrm{d}t}{2\sqrt{t}} = \big[\sqrt{t}\,\big]_0^1 = 1 < \infty,$$

y en efecto $\sqrt{x}=\int_0^x \frac{\mathrm{d}t}{2\sqrt{t}}$ para todo $x\in[0,1]$. Por el paso 3, f es absolutamente continua, y el teorema fundamental se cumple: $f(1)-f(0)=1=\int_0^1 f'$. Moraleja: una derivada no acotada no impide nada, siempre que sea integrable.

(b) $f(x)=x^{1/3}$ en $[-1,1]$: lo mismo, con un pico peor. Aquí $f'(x)=\frac{1}{3}x^{-2/3}$, que explota en 0 como $x^{-2/3}$. Sigue siendo integrable, pues $\int_0^1 \frac{1}{3}t^{-2/3}\mathrm{d}t=\big[t^{1/3}\big]_0^1=1$, y la variación total es $\int_{-1}^1|f'|=1+1=2=f(1)-f(-1)$, coherente con que f sea creciente. Absolutamente continua, pese a tener tangente vertical.

(c) $g(x)=x^2\sin(1/x^2)$ con $g(0)=0$: derivable en *todo* punto y no absolutamente continua. Que es derivable en 0 se ve directamente: $\big|\frac{g(h)-g(0)}{h}\big| = |h\sin(1/h^2)|\le |h|\to 0$, luego $g'(0)=0$. Fuera del origen,

$$g'(x) = 2x\sin\frac{1}{x^2} - \frac{2}{x}\cos\frac{1}{x^2},$$

cuyo segundo sumando explota como $2/x$. Los extremos locales están en $1/x_k^2=\frac{\pi}{2}+k\pi$, donde $|g(x_k)|=x_k^2=\dfrac{1}{\pi/2+k\pi}$, y esas alturas se alternan de signo. La variación entre extremos consecutivos es al menos $\sum_k \dfrac{2}{\pi/2+k\pi}$, que diverge como la serie armónica: con 50 términos ya pasa de 3,74 y con 5000 de 6,67, sin techo. Luego g no es de variación acotada, y por el paso 5 no es absolutamente continua. Aquí falla el paso 4: g' existe *en todos* los puntos y aun así no es integrable, de modo que $\int_0^1 g'$ ni siquiera tiene sentido como integral de Lebesgue.

(d) La escalera de Cantor: el contraejemplo canónico. La función $c:[0,1]\to[0,1]$ es continua, creciente (luego de variación acotada, con $V=1$) y satisface $c'(x)=0$ en casi todo punto, porque es constante en cada uno de los intervalos que se quitan al construir el conjunto de Cantor, y esos intervalos ocupan medida total 1. Por tanto

$$\int_0^1 c'(t)\,\mathrm{d}t = 0 \qquad \text{mientras que} \qquad c(1)-c(0)=1.$$

El teorema fundamental falla estrepitosamente, y por el paso 6 concluimos que c *no* es absolutamente continua. Es continua y de variación acotada: las dos primeras condiciones de Banach–Zarecki se cumplen. Lo que falla es la tercera.

(e) Dónde falla exactamente la escalera: la propiedad N. El conjunto de Cantor $\mathcal{C}$ tiene medida nula, y sin embargo $c(\mathcal{C})=[0,1]$, que tiene medida 1. La función manda un conjunto despreciable sobre un conjunto de medida total: viola la propiedad N de Luzin del paso 7. Toda la variación de c está concentrada sobre $\mathcal{C}$, exactamente donde la derivada no existe, y por eso c' no puede ver nada. Compárese con el apartado (a): allí la derivada explotaba, pero lo hacía sobre un conjunto suficientemente "gordo" como para que la integral lo registrase.

Cuidado

Ser derivable en casi todo punto no dice nada sobre el teorema fundamental. La escalera de Cantor es derivable en casi todo punto —con derivada 0— y la fórmula falla. Y la función del apartado (c) es derivable en *todo* punto, sin excepción, y la fórmula también falla. La hipótesis correcta no es sobre dónde existe f', sino sobre si f' es integrable y recupera a f al integrarla; eso es justo lo que dice la continuidad absoluta y no lo dice ninguna hipótesis de derivabilidad puntual.

Pausa para pensar

El teorema de descomposición de Lebesgue ordena todo esto: toda función continua de variación acotada se escribe de manera única como suma de una parte absolutamente continua y una parte *singular*, cuya derivada es cero en casi todo punto. La escalera de Cantor es el ejemplo puro de la segunda; $\sqrt{x}$ lo es de la primera. Y hay una traducción exacta al lenguaje de medidas: la derivada de una función creciente define una medida, la parte absolutamente continua corresponde a las medidas con densidad respecto de Lebesgue —las que admiten un $\mathrm{d}\mu = g\,\mathrm{d}x$— y la singular a las que viven sobre conjuntos de medida nula. El teorema de Radon–Nikodym del capítulo 31 es este mismo enunciado, escrito en el lenguaje que le corresponde.

29.13. Taller de ejercicios resueltos

Ejercicio 29.1. Demuestra que el conjunto de sucesiones de ceros y unos no es numerable, y deduce de nuevo que $\mathbb{R}$ no lo es.

Solución. Sea $S = \{0,1\}^{\mathbb{N}}$ y supón una enumeración $s_1, s_2, \dots$ con $s_n = (s_{n1}, s_{n2}, \dots)$. Define $t = (t_k)$ con $t_k = 1 - s_{kk}$. Entonces $t \in S$ y $t \neq s_n$ para todo n (difieren en la posición n): contradicción. Para $\mathbb{R}$: la aplicación $S \to C \subseteq \mathbb{R}$ que manda (a_k) a $\sum 2a_k 3^{-k}$ es inyectiva con imagen el conjunto de Cantor, luego $|\mathbb{R}| \geq |S| > |\mathbb{N}|$. ∎

Ejercicio 29.2. Calcula $\omega_f(0)$ para $f(x) = \sin(1/x)$ $(x \neq 0)$, $f(0) = 0$, y para $g(x) = x\sin(1/x)$, $g(0) = 0$. ¿Cuál es continua en 0?

Solución. En todo entorno de 0, $\sin(1/x)$ toma los valores 1 y -1 (en $x = \frac{2}{(4k+1)\pi}$ y $x = \frac{2}{(4k+3)\pi}$), luego la oscilación en $(-\delta,\delta)$ vale 2 para todo δ y $\omega_f(0) = 2 \neq 0$: discontinua. Para g, $|g(x)| \leq |x|$, luego $\operatorname{osc}(g,(-\delta,\delta)) \leq 2\delta \to 0$ y $\omega_g(0) = 0$: continua. Ambas son integrables Riemann en $[-1,1]$, pues su único punto de discontinuidad posible es 0, de medida cero. ∎

Ejercicio 29.3. Sea $f = \mathbf{1}_C$ la indicadora del conjunto de Cantor. ¿Es integrable Riemann? ¿Cuánto vale la integral?

Solución. El conjunto de discontinuidades de $\mathbf{1}_C$ es la frontera de C, que es C mismo (pues C es cerrado con interior vacío, luego $\partial C = \overline{C} \setminus C^\circ = C$). Como $|C| = 0$, el criterio de Lebesgue da integrabilidad. Toda suma inferior es 0 (todo subintervalo contiene puntos fuera de C, porque C no contiene intervalos), luego $\int_0^1 \mathbf{1}_C = 0$. Compara con $\mathbf{1}_{\mathbb{Q}}$, cuyo conjunto de discontinuidades es todo $[0,1]$ y que no es integrable: la diferencia la marca la medida del conjunto, no su cardinal. ∎

Ejercicio 29.4. Demuestra que $\mathbb{R}$ no es unión numerable de rectas… es decir, que $[0,1]$ no es unión numerable de conjuntos cerrados disjuntos no triviales. (Versión débil: demuestra que $[0,1]$ no es unión numerable de puntos.)

Solución. Los puntos son cerrados de interior vacío, luego densos en ninguna parte. Si $[0,1]=\bigcup_n\{x_n\}$, entonces $[0,1]$ sería magro en sí mismo, contradiciendo Baire ($[0,1]$ es completo por ser cerrado en $\mathbb{R}$). Esta es una tercera demostración de la no numerabilidad de $\mathbb{R}$, distinta de la diagonal y de la del conjunto de Cantor. ■

Ejercicio 29.5. Verifica que $T(x)=\frac{x}{2}+\frac{1}{x}$ es una contracción en $[1,2]$ y encuentra su punto fijo.

Solución. Primero, T manda $[1,2]$ en sí mismo: T tiene mínimo en $x=\sqrt{2}$ con $T(\sqrt{2})=\sqrt{2}\approx 1{,}414$, y $T(1)=1{,}5$, $T(2)=1{,}5$; luego $T([1,2])=[\sqrt{2},1{,}5]\subseteq[1,2]$. Segundo, $T'(x)=\frac{1}{2}-\frac{1}{x^2}$ cumple $|T'(x)|\le\frac{1}{2}$ en $[1,2]$ (en $x=1$ vale $-1/2$, en $x=2$ vale $1/4$), luego $k=1/2$. El punto fijo resuelve $x=\frac{x}{2}+\frac{1}{x}$, o sea $x^2=2$: es $\sqrt{2}$. Este es exactamente el método babilónico, y la cota de Banach explica su convergencia; de hecho converge más rápido que k^n porque $T'(\sqrt{2})=0$ (convergencia cuadrática, como en Newton). ■

Ejercicio 29.6. Sea $f_n(x)=\frac{nx}{1+n^2x^2}$ en $[0,1]$. ¿Converge puntualmente? ¿Uniformemente? ¿Es equicontinua la familia?

Solución. Puntualmente: para $x>0$ fijo, $f_n(x)\approx\frac{1}{nx}\to 0$; y $f_n(0)=0$. Luego $f_n\to 0$ puntualmente. Uniformemente no: $f_n(1/n)=\frac{1}{2}$ para todo n, luego $\|f_n\|_\infty\ge\frac{1}{2}$. La familia está uniformemente acotada ($|f_n|\le\frac{1}{2}$ por la desigualdad $2|u|\le 1+u^2$ con $u=nx$) pero no es equicontinua en 0: la pendiente $f_n'(0)=n$ crece sin cota, así que no hay δ común. Arzelà–Ascoli no se aplica, y en efecto no hay subsucesión uniformemente convergente (cualquier límite uniforme tendría que ser el límite puntual 0). ■

Ejercicio 29.7. Calcula $B_3 f$ para $f(x)=x^3$ y compara con f.

Solución.

$$B_3(x^3)=\sum_{k=0}^{3}\Big(\frac{k}{3}\Big)^3\binom{3}{k}x^k(1-x)^{3-k}=\frac{1}{27}\Big[3x(1-x)^2+8\cdot 3x^2(1-x)+27x^3\Big].$$

Desarrollando: $3x(1-2x+x^2)+24x^2-24x^3+27x^3=3x-6x^2+3x^3+24x^2+3x^3=3x+18x^2+6x^3$, dividido entre 27 da

$$B_3(x^3)=\frac{x}{9}+\frac{2x^2}{3}+\frac{2x^3}{9}.$$

En $x=1$ vale $\frac{1}{9}+\frac{2}{3}+\frac{2}{9}=1$, como debe ser (los polinomios de Bernstein siempre son exactos en los extremos). En $x=1/2$ da $\frac{1}{18}+\frac{1}{6}+\frac{1}{36}=\frac{2+6+1}{36}=\frac{9}{36}=\frac{1}{4}$, frente a $f(1/2)=1/8$: error de $1/8$, coherente con la lentitud $O(1/n)$ del método. ■

Ejercicio 29.8. Encuentra un reordenamiento de la serie armónica alternada que sume 0, y otro que diverja a $+\infty$.

Solución. Para sumar 0, aplica el algoritmo de Riemann con $L=0$: empieza con 1 (por encima de 0), luego resta negativos hasta bajar de 0: $1-\frac{1}{2}-\frac{1}{4}-\frac{1}{6}-\frac{1}{8}=1-1{,}0417<0$; luego suma positivos hasta pasar de 0: $+\frac{1}{3}$; y así. Como los términos tienden a 0, las sumas parciales quedan atrapadas cada vez más cerca de 0. Para $+\infty$: suma positivos hasta pasar de 1, resta $\frac{1}{2}$, suma positivos hasta pasar de 2, resta $\frac{1}{4}$, suma hasta pasar de 3, resta $\frac{1}{6}$… Cada tramo positivo es posible porque $\sum\frac{1}{2k-1}=\infty$, y los negativos aportan en total solo $-\frac{1}{2}\sum\frac{1}{k}$ a lo largo de infinitos pasos, pero se consumen de uno en uno, de modo que las sumas parciales superan cualquier cota. ■

Ejercicio 29.9. Demuestra que si $f \in BV[a,b]$ y $g \in BV[a,b]$, entonces $fg \in BV[a,b]$, y da la cota de la variación.

Solución. Ambas son acotadas: $|f| \le |f(a)| + V_a^b(f) =: M_f$, e igual para g. Entonces

$$|f(x_i)g(x_i) - f(x_{i-1})g(x_{i-1})| \le |f(x_i)||g(x_i) - g(x_{i-1})| + |g(x_{i-1})||f(x_i) - f(x_{i-1})|,$$

y sumando sobre la partición, $V_a^b(fg) \le M_f V_a^b(g) + M_g V_a^b(f) < \infty$. Nota que BV es entonces un álgebra; con la norma $\|f\|_{BV} = |f(a)| + V_a^b(f)$ es incluso un espacio de Banach. ∎

Ejercicio 29.10. Sea c la escalera de Cantor. Calcula $V_0^1(c)$ y explica por qué eso no contradice que $c' = 0$ en casi todo punto.

Solución. c es creciente, luego $V_0^1(c) = c(1) - c(0) = 1$. No hay contradicción porque la igualdad $V_a^b(f) = \int_a^b |f'|$ vale para funciones absolutamente continuas, no para toda función de variación acotada. La descomposición de Lebesgue de una función BV tiene una parte absolutamente continua, una parte de saltos y una *parte singular continua*; la escalera de Cantor es puro término singular: toda su variación se acumula sobre un conjunto de medida cero, donde la derivada no existe. Es el ejemplo canónico de que "derivada nula en casi todo punto" no implica "constante". ∎

29.14. Cierre: qué llevarte de este capítulo

Cinco frases para retener.

Uno. Hay más de un infinito, y $\mathbb{R}$ es estrictamente mayor que $\mathbb{Q}$; la demostración es una diagonal que desobedece a cada fila de cualquier lista propuesta.

Dos. Cardinal y medida son cosas distintas: el conjunto de Cantor tiene tantos puntos como la recta y longitud cero.

Tres. La integral de Riemann existe exactamente cuando el conjunto de discontinuidades tiene medida cero; ni una hipótesis más, ni una menos.

Cuatro. La completitud, vía Baire y vía el principio de contracción, es una máquina de demostrar que las cosas existen sin construirlas.

Cinco. Las hipótesis de convergencia absoluta y de continuidad absoluta son la misma idea en dos disfraces: sin ellas, reordenar una serie cambia su suma y el teorema fundamental del cálculo se rompe. Los teoremas de convergencia del capítulo 31 y la teoría de espacios de funciones que sigue existen para administrar exactamente esa frontera.

29.15. Medida cero, sin esperar al capítulo de Lebesgue

El criterio que decide qué funciones son integrables Riemann —el teorema de Lebesgue— se enuncia y se usa aquí, pero su concepto central se remitía a un capítulo posterior. No hace falta: la definición es elemental y sólo usa longitudes de intervalos.

Definición 29.103 (Conjunto de medida cero)**.** Un conjunto $A \subset \mathbb{R}$ tiene *medida cero* —o es *nulo*— si para todo $\varepsilon > 0$ existe una familia *numerable* de intervalos abiertos $(I_k)_{k\ge1}$ tal que

$$A \subset \bigcup_{k\ge1} I_k \qquad \text{y} \qquad \sum_{k\ge1} |I_k| < \varepsilon,$$

donde $|I_k|$ denota la longitud. En $\mathbb{R}^n$ se pide lo mismo con rectángulos y volúmenes.

Proposición 29.104 (Las tres propiedades que se usan)**.** *1. Todo conjunto* numerable *tiene medida cero.*

2. Una unión numerable *de conjuntos de medida cero tiene medida cero.*

3. Un intervalo $[a,b]$ *con* $a<b$ no *tiene medida cero.*

Demostración. (1) Sea $A=\{a_1,a_2,\dots\}$ y $\varepsilon>0$. Tómese $I_k=(a_k-\frac{\varepsilon}{2^{k+2}},\,a_k+\frac{\varepsilon}{2^{k+2}})$, de longitud $\frac{\varepsilon}{2^{k+1}}$. Entonces $A\subset\bigcup I_k$ y $\sum_k|I_k|=\varepsilon/2<\varepsilon$.

(2) Sea $A=\bigcup_n A_n$ con cada A_n nulo. Recúbrase A_n con intervalos de longitud total $<\varepsilon/2^n$. La unión de todas esas familias es numerable —unión numerable de familias numerables— y su longitud total es $<\sum_n \varepsilon/2^n=\varepsilon$.

(3) Si $[a,b]\subset\bigcup_k I_k$, por compacidad basta un subrecubrimiento finito $I_{k_1},\dots,I_{k_m}$; y una inducción elemental sobre m muestra que $\sum_j|I_{k_j}|\geq b-a$. Luego ninguna familia puede tener longitud total menor que $b-a>0$. □

Ejemplo 29.105 (Los tres conjuntos que hay que tener en la cabeza)**.**

- $\mathbb{Q}\cap[0,1]$ es numerable, luego *nulo*. Por eso la función de Dirichlet —que vale 1 en los racionales y 0 en los irracionales— tiene un conjunto de discontinuidades de medida *total*, no nula: es discontinua en todo punto, y por eso no es integrable Riemann.
- La función de Thomae —que vale $1/q$ en p/q irreducible y 0 en los irracionales— es discontinua exactamente en $\mathbb{Q}$, un conjunto nulo. Por el teorema de Lebesgue *sí* es integrable Riemann, con integral 0, pese a tener infinitas discontinuidades. Ése es el ejemplo que separa "muchas discontinuidades" de "discontinuidades gordas".
- El conjunto de Cantor es *no numerable* —tiene el cardinal del continuo— y sin embargo es nulo: en la etapa n está cubierto por 2^n intervalos de longitud 3^{-n}, con longitud total $(2/3)^n\to 0$. Y el *Cantor gordo*, construido quitando en cada etapa intervalos más cortos, es homeomorfo al anterior pero tiene medida positiva. Nulo y numerable no son lo mismo, en ninguna de las dos direcciones.

Observación 29.106 (Por qué la numerabilidad es esencial)*.* Si en la definición se permitieran recubrimientos *arbitrarios*, todo conjunto tendría medida cero: cada punto se recubre con un intervalo tan corto como se quiera, y $[0,1]$ es unión de sus puntos. Lo que hace que la definición diga algo es que sólo se admiten familias numerables, de modo que las longitudes se suman en una serie. Ésa es exactamente la razón por la que la teoría de la medida es *numerablemente* aditiva y no más, y por la que en el capítulo 31 la σ-aditividad aparece como axioma en vez de deducirse.

29.16. Ejercicios resueltos

Ejercicio 29.11. Estudia la iteración $x_{n+1}=\cos x_n$ desde $x_0=0$: demuestra que converge, identifica el límite y estima la velocidad.

Solución. **(a) El intervalo invariante.** Para cualquier x real, $\cos x \in [-1,1]$, y aplicando el coseno otra vez, $\cos x \in [\cos 1, 1] \subset [0{,}540, 1]$. Así que a partir del segundo paso la sucesión vive en $I = [\cos 1, 1]$, y $\cos(I) \subset I$.

(b) Es contractiva ahí. En I se tiene $|\cos' x| = |\sin x| \le \sin 1 = 0{,}841471 < 1$, porque $I \subset [0,1]$ y el seno crece. Por el teorema del valor medio, $|\cos a - \cos b| \le 0{,}8415\,|a-b|$: es una contracción con constante $L = \sin 1$. Como I es completo (cerrado en $\mathbb{R}$), el principio de contracción da un punto fijo único y convergencia desde cualquier arranque.

(c) El límite. Es la única solución de $x = \cos x$, llamada a veces el *número de Dottie*:

$$x^* = 0{,}739085133215\ldots$$

No tiene forma cerrada elemental, lo que resulta chocante para una ecuación tan inocente.

(d) La velocidad. La cota teórica $|x_n - x^*| \le L^n |x_0 - x^*|$ con $L = 0{,}8415$ predice una convergencia lenta: hacen falta unas 27 iteraciones para ganar dos cifras ($0{,}8415^{27} \approx 0{,}0094$). Los primeros términos desde $x_0 = 0$ lo confirman:

$$0,\quad 1,\quad 0{,}540302,\quad 0{,}857553,\quad 0{,}654290,\quad 0{,}793480,\quad 0{,}701369,\ \ldots$$

oscilando alrededor de x^* y estrechándose por un factor $\approx 0{,}67 = |\sin x^*|$ cada vez. Ese $|\sin x^*| = |f'(x^*)|$ es la constante que gobierna *de verdad* el régimen asintótico; la L global solo da la garantía. ■

Ejercicio 29.12. Demuestra que el conjunto de Cantor tiene medida nula y sin embargo el cardinal del continuo.

Solución. **(a) Medida nula.** En el paso n quedan 2^n intervalos de longitud 3^{-n}, luego la longitud total superviviente es $(2/3)^n \to 0$. Alternativamente, sumando lo que se quita: en el paso n se retiran 2^{n-1} intervalos de longitud 3^{-n}, y

$$\sum_{n=1}^{\infty} \frac{2^{n-1}}{3^n} = \frac{1}{3}\sum_{n\ge 0}\Big(\frac{2}{3}\Big)^n = \frac{1}{3}\cdot\frac{1}{1-2/3} = 1.$$

Se quita longitud total 1 de un intervalo de longitud 1: lo que queda mide 0.

(b) Cardinal $\mathfrak{c}$. Un punto de $[0,1]$ está en el conjunto de Cantor si y solo si admite un desarrollo en base 3 que solo usa las cifras 0 y 2 (el 1 es precisamente la cifra que se elimina en cada paso). La aplicación que cambia cada 2 por un 1 y reinterpreta el resultado en base 2 es sobreyectiva sobre $[0,1]$, luego el Cantor tiene al menos el cardinal del continuo; y está contenido en $[0,1]$, luego lo tiene exactamente.

Aquí está la lección: *medida* y *cardinal* son dos nociones de tamaño independientes. Hay conjuntos numerables densos (los racionales) y conjuntos no numerables de medida cero (el Cantor). Y afinando la construcción — quitando en el paso n intervalos de longitud $\varepsilon\, 4^{-n}$ en lugar de 3^{-n} — se obtiene un "Cantor gordo" que sigue siendo cerrado, sin puntos interiores y totalmente disconexo, pero de medida $1-\varepsilon > 0$. Su función indicadora es acotada y no Riemann-integrable, con lo que se resuelve de paso el ejercicio propuesto 4 por una vía más fina que la de Dirichlet. ■

Ejercicio 29.13. Demuestra que la función de Thomae es continua en los irracionales, discontinua en los racionales, y calcula su integral de Riemann en $[0,1]$.

Solución. Definimos $f(x) = 1/q$ si $x = p/q$ en forma irreducible con $q \ge 1$, y $f(x) = 0$ si x es irracional.

(a) Discontinua en cada racional. Si $x = p/q$, entonces $f(x) = 1/q > 0$, pero hay irracionales tan cerca como se quiera, donde f vale 0. No hay límite igual al valor.

(b) Continua en cada irracional. Sea x irracional y $\varepsilon > 0$. Solo hay una cantidad *finita* de racionales en $[0,1]$ con denominador $q < 1/\varepsilon$ (a lo sumo $\sum_{q<1/\varepsilon} q$ de ellos). Tomando δ menor que la

distancia de x al más cercano de esa lista finita, todo y con $|y-x|<\delta$ cumple $f(y)<\varepsilon$: es continua en x, y con límite 0.

(c) La integral. El criterio de Lebesgue dice que una función acotada es Riemann-integrable si y solo si su conjunto de discontinuidades tiene medida nula. Aquí ese conjunto es $\mathbb{Q}\cap[0,1]$, numerable, luego de medida 0: es integrable. Y como es 0 salvo en un conjunto de medida nula, $\int_0^1 f = 0$.

Directamente, sin invocar a Lebesgue: dado ε, los puntos con $f \geq \varepsilon/2$ son finitos, digamos N; cúbranse con intervalos de longitud total $<\varepsilon/2$. La suma superior asociada es menor que $1\cdot(\varepsilon/2)+(\varepsilon/2)\cdot 1=\varepsilon$, y la inferior es 0. Luego $\overline{\int} = \underline{\int} = 0$.

Contraste con Dirichlet (1 en $\mathbb{Q}$, 0 fuera), que es discontinua *en todas partes* y no integrable: la diferencia entre las dos es que Thomae "se apaga" al crecer el denominador. Ambas son iguales en casi todo punto a la función nula, y ambas tienen integral de Lebesgue 0; solo una tiene integral de Riemann. ■

Ejercicio 29.14. Decide si $f(x)=x\sin(1/x)$ y $g(x)=x^2\sin(1/x)$ (ambas con valor 0 en 0) son de variación acotada en $[0,1]$.

Solución. **(a) f no lo es.** En los puntos $t_n = \dfrac{1}{(n+\frac12)\pi}$ el seno vale ±1, luego $|f(t_n)| = t_n$ y los signos alternan. Tomando la partición formada por $t_1 > t_2 > \cdots > t_N$ y el 0, la variación de f es al menos

$$\sum_{n=1}^{N} \big|f(t_n) - f(t_{n+1})\big| \;\geq\; \sum_{n=1}^{N} \frac{1}{(n+\frac12)\pi},$$

que es una serie armónica disfrazada y por tanto diverge. (Con $N = 20\,000$ la suma parcial vale 3,1408, y sigue creciendo como $\log N/\pi$; en efecto $\log 20\,000/\pi = 3{,}152$.) Luego $V_0^1(f)=\infty$.

(b) g sí lo es. Ahora $g'(x) = 2x\sin(1/x) - \cos(1/x)$ para $x>0$, y $g'(0)=0$; en todo caso $|g'| \leq 2x+1 \leq 3$ en $[0,1]$. Una función con derivada acotada es Lipschitz, y toda función Lipschitz de constante K tiene variación $\leq K(b-a)$. Luego $V_0^1(g)\leq 3$.

El factor x extra es todo lo que separa ambos casos: las oscilaciones tienen amplitud $\sim x$ en f y $\sim x^2$ en g, mientras que su número por unidad de longitud crece como $1/x$ en las dos. El producto amplitud×número es ~ 1 en el primer caso (serie divergente) y $\sim x$ en el segundo (serie convergente).

Nota fina: g es de variación acotada y derivable en todo punto, pero g' no es Riemann-integrable en el sentido impropio absoluto ni g es de clase C^1; y la función de Cantor es continua, creciente y de variación acotada, y sin embargo *no* absolutamente continua. Los tres escalones —continua, variación acotada, absolutamente continua— son estrictamente distintos. ■

Ejercicio 29.15. Demuestra que $\mathbb{R}$ no es unión numerable de conjuntos cerrados con interior vacío, y saca una consecuencia sobre $\mathbb{Q}$.

Solución. Es el teorema de Baire: en un espacio métrico completo, la unión numerable de cerrados con interior vacío tiene interior vacío; en particular no puede ser todo el espacio. $\mathbb{R}$ es completo, luego la respuesta es no.

Consecuencia 1: $\mathbb{Q}$ no es un G_δ. Si lo fuera, $\mathbb{Q}=\bigcap_n U_n$ con U_n abiertos, y cada U_n contiene a $\mathbb{Q}$, luego es denso. Entonces

$$\varnothing = \mathbb{Q}\cap(\mathbb{R}\setminus\mathbb{Q}) = \Big(\bigcap_n U_n\Big) \cap \Big(\bigcap_{q\in\mathbb{Q}} (\mathbb{R}\setminus\{q\})\Big)$$

sería una intersección numerable de abiertos densos, que por Baire tiene que ser densa. Contradicción.

Consecuencia 2: no hay ninguna función continua exactamente en $\mathbb{Q}$. El conjunto de puntos de continuidad de cualquier función $\mathbb{R}\to\mathbb{R}$ es siempre un G_δ (es $\bigcap_n \{x : \operatorname{osc} f(x) < 1/n\}$, y

cada uno de esos conjuntos es abierto). Como $\mathbb{Q}$ no es G_δ, no hay función que lo tenga por conjunto de continuidad.

Y sin embargo sí existe una función continua exactamente en los irracionales: la de Thomae del ejercicio anterior. La asimetría es real, no un defecto de nuestra imaginación: los irracionales sí son un G_δ, y los racionales no. ■

Ejercicio 29.16. Sea $f_n(x) = x^n$ en $[0,1]$. Estudia la convergencia puntual y uniforme, y explica por qué Dini no se aplica pero Arzelà–Ascoli tampoco.

Solución. **(a) Puntual.** $f_n(x) \to 0$ para $x < 1$ y $f_n(1) = 1$: el límite es la función f que vale 0 en $[0,1)$ y 1 en 1, discontinua.

(b) No uniforme. $\sup_{[0,1]} |f_n - f| = \sup_{x<1} x^n = 1$ para todo n; la distancia uniforme ni siquiera tiende a cero. Y no podía ser de otra manera: un límite uniforme de continuas es continuo, y este límite no lo es.

(c) Dini. El teorema de Dini pide cuatro cosas: dominio compacto, funciones continuas, convergencia monótona y *límite continuo*. Aquí se cumplen las tres primeras (x^n decrece en n para cada x fijo) y falla la cuarta. Dini no se aplica, y la conclusión es efectivamente falsa: la continuidad del límite no es un adorno de la hipótesis, es lo que hace funcionar la demostración.

(d) Arzelà–Ascoli. La familia está uniformemente acotada ($\|f_n\|_\infty = 1$), pero no es equicontinua en $x = 1$: para $\varepsilon = 1/2$ y cualquier δ, tomando n grande se tiene $|f_n(1) - f_n(1-\delta)| = 1 - (1-\delta)^n \to 1 > \varepsilon$. Las pendientes $f_n'(1) = n$ explotan. Sin equicontinuidad no hay subsucesión uniformemente convergente, y en efecto no la hay: cualquier subsucesión converge puntualmente a la misma función discontinua.

Compárese con $g_n(x) = x^n$ restringida a $[0, 1/2]$: ahí $\|g_n\|_\infty = 2^{-n} \to 0$, la convergencia es uniforme y la familia sí es equicontinua. Todo el problema está concentrado en un único punto. ■

29.17. Ejercicios propuestos

Sin solución desarrollada: solo enunciado y, al final, la respuesta. En esta parte del libro muchas respuestas son una frase ("sí, y la razón es…") más que un número; comprueba que la tuya coincide en el fondo, no en la letra.

29.1. ¿Es numerable el conjunto de los polinomios con coeficientes racionales? ¿Y el de los números algebraicos?

29.2. Demuestra que $[0,1]$ y $(0,1)$ tienen el mismo cardinal exhibiendo una biyección.

29.3. ¿Cuál es la medida (longitud total) del conjunto de Cantor? ¿Y su cardinal?

29.4. Da una función acotada en $[0,1]$ que no sea Riemann-integrable y di dónde falla el criterio de Lebesgue.

29.5. ¿Es Riemann-integrable la función de Thomae? ¿Cuánto vale su integral?

29.6. ¿Puede $\mathbb{R}$ escribirse como unión numerable de conjuntos cerrados con interior vacío? Cita el teorema.

29.7. Aplica el principio de contracción a $f(x) = \cos x$ en $[0,1]$: ¿converge la iteración? ¿A qué?

29.8. Si $f_n \to f$ puntualmente en $[0,1]$, con todas continuas, f continua y la convergencia monótona, ¿es uniforme? Cita el teorema.

29.9. Enuncia qué hipótesis le falta a $\{f_n(x) = x^n\}$ en $[0,1]$ para ser equicontinua y por qué Arzelà–Ascoli no aplica.

29.10. ¿A qué puede converger un reordenamiento de $\sum(-1)^{n+1}/n$? Da el rango de valores posibles.

29.11. ¿Es $f(x) = x\sin(1/x)$ (con $f(0) = 0$) de variación acotada en $[0, 1]$?

29.12. ¿Toda función continua y de variación acotada es absolutamente continua? Da el contraejemplo estándar.

Respuestas

1. Sí a ambos: unión numerable de conjuntos numerables; los algebraicos son numerables, luego "casi todo" real es trascendente. **2.** Trata los extremos con un truco de Hilbert: manda $0 \mapsto 1/2$, $1 \mapsto 1/3$, $1/n \mapsto 1/(n+2)$ y deja fijo el resto. **3.** Medida 0; cardinal $\mathfrak{c}$ (el mismo que $\mathbb{R}$). **4.** La de Dirichlet (1 en $\mathbb{Q}$, 0 fuera): su conjunto de discontinuidades es todo $[0, 1]$, de medida positiva. **5.** Sí es integrable; la integral vale 0 (sus discontinuidades son solo los racionales, de medida nula). **6.** No: Baire lo prohíbe. $\mathbb{R}$ es completo, luego no es de primera categoría en sí mismo. **7.** Sí: $|\cos'| = |\sin x| \le \sin 1 < 1$ en $[0, 1]$. Converge al punto fijo $x^* \approx 0{,}739085$. **8.** Sí: es exactamente el teorema de Dini. **9.** No es equicontinua en $x = 1$: las pendientes nx^{n-1} explotan. Falla la hipótesis de equicontinuidad, no la de acotación. **10.** A cualquier número real, y también a $+\infty$ o $-\infty$: teorema de reordenamiento de Riemann (la serie es condicionalmente convergente). **11.** No: la variación total diverge como $\sum 1/n$. **12.** No. La escalera del diablo (función de Cantor) es continua, monótona (luego de variación acotada) y no absolutamente continua: $f' = 0$ c.t.p. pero sube de 0 a 1.

Capítulo 30

Variable compleja: el puente entre los mundos

Lo que necesitas traer

Del capítulo 4: aritmética compleja, forma polar y raíces de la unidad. Del capítulo 6: derivadas, integrales y series de potencias. Del capítulo 11: integrales de línea y campos conservativos, que son lo mismo que la integral de contorno.

Veinte veces la rozamos: Euler en el capítulo 6, los valores propios complejos en el 5, el casus irreducibilis en el 8, la deformación de contornos en el 4. El análisis complejo merece cerrar el libro: es el punto donde el cálculo, la geometría y el álgebra de todos los capítulos anteriores se funden en una sola teoría, más rígida y más bella que cualquiera de sus partes.

30.1. Funciones holomorfas y las ecuaciones de Cauchy–Riemann

Idea clave

Intuición primero. Derivar en $\mathbb{R}^2$ solo pide un buen plano tangente; derivar en $\mathbb{C}$ pide muchísimo más: que el cociente incremental dé *el mismo número complejo* sin importar la dirección de acercamiento. Geométricamente, la diferencial debe ser "multiplicar por un complejo", y multiplicar por un complejo es *girar y escalar*: una función holomorfa es, bajo el microscopio, una rotación con zoom en cada punto (por eso preserva ángulos: conforme). Ese contrato draconiano explica toda la rigidez milagrosa del capítulo: quien firma "giro y escala en todas partes" queda determinado por sus valores en cualquier arco diminuto, es automáticamente infinitamente derivable, y no puede tener máximos interiores. Las ecuaciones de Cauchy–Riemann son simplemente la letra pequeña del contrato.

Definición 30.1. $f : \mathbb{C} \to \mathbb{C}$ es *holomorfa* en z_0 si existe

$$f'(z_0) = \lim_{h \to 0} \frac{f(z_0 + h) - f(z_0)}{h} \qquad (h \in \mathbb{C}) :$$

el *mismo* límite por todas las direcciones complejas. Esta exigencia —invisible en apariencia— es brutalmente restrictiva.

Teorema 30.2 (Cauchy–Riemann)**.** *Escribiendo $f = u + iv$ con $z = x + iy$: f es holomorfa $\iff$ u, v son diferenciables y*

$$u_x = v_y, \qquad u_y = -v_x.$$

Demostración. Igualar el límite por h real ($f' = u_x + iv_x$) con el límite por $h = ik$ imaginario ($f' = \frac{u_y + iv_y}{i} = v_y - iu_y$): partes reales e imaginarias dan las dos ecuaciones; el recíproco es la diferenciabilidad real más el álgebra de la multiplicación compleja. □

Lectura geométrica: la jacobiana $\left(\begin{smallmatrix} u_x & u_y \\ v_x & v_y \end{smallmatrix}\right)$ con CR es $\left(\begin{smallmatrix} a & -b \\ b & a \end{smallmatrix}\right)$: *rotación por estiramiento* (capítulo 7): las holomorfas preservan ángulos donde $f' \neq 0$ (son *conformes*). Y derivando CR una vez más:

$$u_{xx} + u_{yy} = v_{yx} - v_{xy} = 0 :$$

las partes real e imaginaria de una holomorfa son armónicas (¡capítulos 11 y 12!): toda la teoría del potencial plano es análisis complejo disfrazado, y las curvas de nivel de u y v forman las redes ortogonales de los mapas de flujo.

Ejemplo 30.3. e^z, polinomios, $\sin z$, $\cos z$ (por sus series del capítulo 6, que convergen en todo $\mathbb{C}$), $\frac{1}{z}$ fuera del origen. *No* holomorfa: $\bar{z}$ (CR falla: $u_x = 1 \neq -1 = v_y$): conjugar invierte ángulos. $\ln z$ es holomorfa pero *multivaluada* (cada vuelta al origen suma $2\pi i$: ¡el campo del índice del capítulo 11 era exactamente $\operatorname{Im} \frac{dz}{z}$!): se elige una *rama* cortando el plano.

Regla: comprobar si una función es holomorfa

La derivabilidad compleja es mucho más exigente que la real, y el procedimiento para verificarla es siempre el mismo.

1. **Separa parte real e imaginaria.** Escribe $z = x + iy$ y desarrolla hasta dejar $f(z) = u(x, y) + i\, v(x, y)$ con u y v reales. Este paso, puramente algebraico, es donde se cometen la mayoría de los errores: hay que recordar que $i^2 = -1$ y que $\bar{z} = x - iy$.
2. **Calcula las cuatro parciales** u_x, u_y, v_x, v_y.
3. **Impón Cauchy–Riemann:** $u_x = v_y$ y $u_y = -v_x$. Fíjate en el signo de la segunda: es el que distingue una función holomorfa de su conjugada.
4. **Mira *dónde* se cumplen.** Si las dos igualdades se cumplen solo en un punto o sobre una curva, la función **no** es holomorfa en ningún sitio: la holomorfía exige derivabilidad en todo un entorno abierto, no en puntos sueltos.
5. **Comprueba la continuidad de las parciales** (basta que sean continuas para que Cauchy–Riemann sea suficiente, no solo necesario). Con funciones elementales esto es automático.

6. **Si es holomorfa, la derivada sale gratis:** $f'(z) = u_x + i\,v_x$. Y como control, el resultado debería poder reescribirse en términos de z sola, sin que aparezcan x ni y por separado.

Ejemplo 30.4 (Tres funciones y tres desenlaces distintos)**.** *(a)* $f(z) = z^2$. Con $z = x + iy$, $z^2 = (x^2 - y^2) + i(2xy)$, luego $u = x^2 - y^2$ y $v = 2xy$. Las parciales:

$$u_x = 2x, \quad u_y = -2y, \quad v_x = 2y, \quad v_y = 2x.$$

Se cumple $u_x = v_y$ $(2x = 2x)$ y $u_y = -v_x$ $(-2y = -2y)$ en *todo* el plano. Es entera, y $f' = u_x + iv_x = 2x + 2iy = 2z$, como esperábamos.

(b) $f(z) = \overline{z}$. Ahora $u = x$, $v = -y$, así que $u_x = 1$ pero $v_y = -1$. La primera ecuación falla en todos los puntos: la conjugación no es derivable en ningún sitio. Geométricamente es razonable: conjugar es reflejar respecto del eje real, y una reflexión invierte los ángulos en lugar de conservarlos.

(c) $f(z) = |z|^2$. Aquí $u = x^2 + y^2$ y $v = 0$. Las condiciones piden $2x = 0$ y $2y = -0$, es decir, se cumplen *únicamente* en el origen. Como no hay ningún entorno abierto donde valgan, $|z|^2$ no es holomorfa en ninguna parte, ni siquiera en el origen (allí es derivable, pero eso no basta). Este es el ejemplo que hay que tener en la cabeza: una función puede ser infinitamente derivable como función de dos variables reales y aun así no ser holomorfa.

Regla: construir la función holomorfa a partir de su parte real

Si te dan $u(x, y)$ y te piden la f holomorfa con $\operatorname{Re} f = u$, el camino es el mismo que el de los potenciales del capítulo 11, con Cauchy–Riemann haciendo de gradiente.

1. **Verifica que u es armónica:** $u_{xx} + u_{yy} = 0$. Si no lo es, no existe tal f y el problema no tiene solución.
2. **Usa la primera ecuación:** $v_y = u_x$. Integra respecto de y para obtener $v = \int u_x \,\mathrm{d}y + \varphi(x)$.
3. **Usa la segunda para fijar φ:** impón $v_x = -u_y$ y despeja $\varphi'(x)$; intégrala. Queda una constante real arbitraria, que refleja que f está determinada salvo una constante imaginaria pura.
4. **Reescribe en términos de z.** Sustituye $x = \frac{z+\bar{z}}{2}$, o más rápido: evalúa f sobre el eje real $(y = 0)$ y sustituye x por z, truco que funciona porque dos funciones holomorfas que coinciden en un segmento coinciden en todo el dominio (principio de identidad).

Ejemplo 30.5 (De una función armónica a su función holomorfa)**.** Sea $u(x, y) = x^3 - 3xy^2$. Primero, ¿es armónica?

$$u_{xx} = 6x, \qquad u_{yy} = -6x, \qquad u_{xx} + u_{yy} = 0. \ \checkmark$$

Ahora $u_x = 3x^2 - 3y^2$, y por Cauchy–Riemann $v_y = 3x^2 - 3y^2$, de donde

$$v = 3x^2y - y^3 + \varphi(x).$$

La segunda condición pide $v_x = -u_y$. Calculamos $u_y = -6xy$ y $v_x = 6xy + \varphi'(x)$, luego

$$6xy + \varphi'(x) = 6xy \implies \varphi'(x) = 0 \implies \varphi = c.$$

Tomando $c = 0$,

$$f = (x^3 - 3xy^2) + i(3x^2y - y^3).$$

Para reconocerla, evaluamos en el eje real ($y = 0$): queda x^3. Sustituyendo x por z conjeturamos $f(z) = z^3$, y en efecto $(x+iy)^3 = x^3 + 3ix^2y - 3xy^2 - iy^3$ reagrupa exactamente en lo anterior.

La moraleja es que la parte real de una función holomorfa determina la parte imaginaria salvo constante: hay muchísima menos libertad de la que uno esperaría. Dos funciones armónicas relacionadas de este modo se llaman *conjugadas armónicas*, y sus curvas de nivel se cortan siempre en ángulo recto, propiedad que se usa para dibujar a la vez las líneas de campo y las equipotenciales de un problema físico.

30.2. Los cimientos, uno a uno: funciones, límites, derivada, armónicas, funciones elementales y la integral de contorno

La sección anterior ha ido deprisa. Ha dicho qué son las ecuaciones de Cauchy–Riemann, ha enseñado a comprobarlas y ha exhibido las tres funciones de rigor, pero apoyándose en media docena de objetos que nunca se definieron: qué es exactamente una función de variable compleja, qué significa que un límite exista en el plano, qué es un camino, qué es esa integral $\oint$ que a partir de ahora aparecerá en cada página. Esta sección paga la deuda y está escrita para quien nunca ha visto variable compleja: no se supone nada más que la aritmética compleja del capítulo 4 y el cálculo de una variable real.

Conviene decir de entrada por qué merece la pena tanta ceremonia. En análisis real uno puede ser descuidado con las definiciones porque la intuición del gráfico le rescata: se ve la curva, se ve la pendiente, se ve el área. En $\mathbb{C}$ no hay gráfico que ver —el dominio ya tiene dos dimensiones y la imagen otras dos, así que una gráfica viviría en cuatro—, de modo que la única brújula fiable son las definiciones. La buena noticia es que son pocas y que casi todas se copian del caso real cambiando el valor absoluto por el módulo. La sorprendente es que ese cambio inocente convierte la derivada en una condición durísima, y ahí empieza toda la teoría.

Idea clave

El plan. Primero decimos qué objeto estudiamos: una función que come puntos del plano y devuelve puntos del plano, es decir, *dos* funciones reales de dos variables a la vez. Después copiamos del análisis real las nociones de límite y continuidad, que no dan ninguna sorpresa. Luego copiamos la derivada, que sí la da: exigir que un cociente incremental tenga el mismo límite por infinitas direcciones no es lo mismo que exigirlo por dos. De esa única exigencia salen Cauchy–Riemann y las funciones armónicas.

Cerramos con las funciones elementales complejas y con la integral de contorno, que es la herramienta de trabajo de todo lo que viene después.

30.2.1. Funciones de variable compleja

Definición 30.6 (Función de variable compleja)**.** Sea $\Omega \subseteq \mathbb{C}$ un conjunto abierto. Una *función de variable compleja* definida en Ω es una aplicación $f : \Omega \to \mathbb{C}$ que a cada $z \in \Omega$ le asigna un único número complejo $f(z)$. A Ω se le llama *dominio* de f.

Escribiendo $z = x+iy$ con $x, y \in \mathbb{R}$ y separando el valor $f(z)$ en sus partes real e imaginaria quedan definidas dos funciones reales de dos variables reales,

$$u(x,y) = \operatorname{Re} f(x+iy), \qquad v(x,y) = \operatorname{Im} f(x+iy),$$

llamadas *parte real* y *parte imaginaria* de f, de modo que $f(z) = u(x,y) + i\,v(x,y)$.

Esta descomposición no es un adorno notacional: es el diccionario que traduce todo enunciado de variable compleja a un enunciado sobre dos funciones de $\mathbb{R}^2$ en $\mathbb{R}$, y al revés. Cada vez que algo parezca misterioso, el primer reflejo debe ser escribir u y v y mirar qué está pasando en el plano real.

Observación 30.7 (Por qué no hay gráfica que dibujar). La gráfica de una función real $f : \mathbb{R} \to \mathbb{R}$ vive en $\mathbb{R}^2$ y cabe en un papel. La de $f : \Omega \to \mathbb{C}$ necesitaría dos coordenadas para la entrada y dos para la salida: vive en $\mathbb{R}^4$. Por eso en variable compleja **no se dibujan gráficas: se dibujan imágenes**. Se toma en el plano de partida una malla de referencia —las rectas $x = c$ y $y = k$, o bien las circunferencias $|z| = r$ y los rayos $\arg z = \theta$— y se dibuja en el plano de llegada en qué se convierte. Dos dibujos lado a lado, el "antes" y el "después", es la manera estándar de ver una función compleja. Y la propiedad que se lee de inmediato en ese par de dibujos es la que dará nombre al final del capítulo: si las curvas de la malla se cortaban en ángulo recto y sus imágenes se siguen cortando en ángulo recto, la función es *conforme*, que es tanto como decir holomorfa con derivada no nula.

Ejemplo 30.8 (La función $f(z) = z^2$, con sus dos partes y su malla)**.** Tomemos $\Omega = \mathbb{C}$ y $f(z) = z^2$. Desarrollamos con $z = x + iy$, recordando que $i^2 = -1$:

$$z^2 = (x+iy)^2 = x^2 + 2ixy + i^2y^2 = (x^2 - y^2) + i\,(2xy),$$

de donde

$$u(x,y) = x^2 - y^2, \qquad v(x,y) = 2xy.$$

Comprobación en un punto: en $z = 1+i$ es $x = y = 1$, luego $u = 1 - 1 = 0$ y $v = 2$, es decir $f(1+i) = 2i$; y directamente $(1+i)^2 = 1 + 2i - 1 = 2i$. Coincide.

Qué le hace a la malla cartesiana. Fijemos $x = c$ con $c \neq 0$ y dejemos correr y: entonces $u = c^2 - y^2$ y $v = 2cy$, de donde $y = v/(2c)$ y

$$u = c^2 - \frac{v^2}{4c^2},$$

una parábola abierta hacia la izquierda con vértice en $u = c^2$. Si fijamos en cambio $y = k$ con $k \neq 0$, queda $u = x^2 - k^2$ y $v = 2kx$, es decir $x = v/(2k)$ y $u = \dfrac{v^2}{4k^2} - k^2$, otra parábola, abierta hacia la derecha. Las dos familias se cortan perpendicularmente, igual que se cortaban las rectas de partida: la conformidad, visible a simple vista.

Qué le hace a la malla polar. Si $z = re^{i\theta}$ entonces $z^2 = r^2 e^{2i\theta}$: los módulos se elevan al cuadrado y los argumentos se *duplican*. La circunferencia de radio r va a la de radio r^2, el rayo de ángulo θ va al de ángulo 2θ, y el primer cuadrante, que ocupa ángulos de 0 a $\pi/2$, se abre hasta ocupar de 0 a π: todo el semiplano superior. Esa última frase es todo lo que hay detrás del ejercicio de potencial resuelto al final del capítulo.

30.2.2. Límite y continuidad en el plano complejo

La definición de límite es palabra por palabra la del caso real, con $|\cdot|$ significando ahora el módulo complejo. Lo único que cambia —y lo cambia todo— es la forma del conjunto por el que uno puede acercarse.

Definición 30.9 (Límite)**.** Sean f definida en un entorno de z_0 salvo quizá en z_0 mismo, y $L \in \mathbb{C}$. Se dice que $\lim_{z\to z_0} f(z) = L$ si para todo $\varepsilon > 0$ existe $\delta > 0$ tal que

$$0 < |z - z_0| < \delta \implies |f(z) - L| < \varepsilon.$$

Definición 30.10 (Continuidad)**.** f es *continua* en $z_0 \in \Omega$ si $\lim_{z\to z_0} f(z) = f(z_0)$; es continua en Ω si lo es en todos sus puntos.

Observación 30.11 (La observación clave de toda la sección)*.* En $\mathbb{R}$, acercarse a un punto solo se puede hacer de dos maneras, por la izquierda y por la derecha, y por eso los límites laterales lo deciden todo. En $\mathbb{C}$, la condición $0 < |z - z_0| < \delta$ describe un *disco* perforado, y a un punto del plano se llega por infinitas direcciones distintas —y ni siquiera hace falta ir en línea recta: se puede llegar en espiral—. Que el límite exista significa que **todas** esas maneras de acercarse dan el mismo número, y basta que dos discrepen para que no exista.

Para la continuidad esto no tiene consecuencias dramáticas: f es continua si y solo si u y v lo son como funciones de $\mathbb{R}^2$ en $\mathbb{R}$, porque $\max(|u-u_0|, |v-v_0|) \le |f - f_0| \le |u-u_0| + |v-v_0|$. Pero para la *derivada*, donde el cociente incremental depende de la dirección de acercamiento, la exigencia de unanimidad será brutal. Recuérdese esta observación: es la única razón por la que el análisis complejo se parece tan poco al real.

Ejemplo 30.12 (Un límite que no existe, y por qué)**.** Estudiemos $\lim\limits_{z\to 0} \dfrac{\bar z}{z}$. La función está definida en $\mathbb{C} \setminus \{0\}$ y su módulo vale $|\bar z|/|z| = 1$ en todas partes, así que no se dispara: si el límite falla, no será por tamaño. Escribamos $z = re^{i\theta}$ y acerquémonos al origen por el rayo de ángulo θ fijo, haciendo $r \to 0^+$. Como $\bar z = re^{-i\theta}$,

$$\frac{\bar z}{z} = \frac{re^{-i\theta}}{re^{i\theta}} = e^{-2i\theta},$$

que *no depende de* r: sobre cada rayo la función es constante. Por el eje real positivo ($\theta = 0$) el valor es 1; por el eje imaginario positivo ($\theta = \pi/2$) es $e^{-i\pi} = -1$; por la bisectriz ($\theta = \pi/4$)

es $e^{-i\pi/2} = -i$. Tres direcciones, tres valores distintos: el límite no existe. De hecho, eligiendo el rayo se obtiene como valor límite *cualquier* número de módulo 1.

En cambio $\lim_{z\to z_0} z^2 = z_0^2$ sin misterio: si $|z - z_0| < \delta \le 1$ entonces $|z^2 - z_0^2| = |z - z_0||z + z_0| \le \delta(2|z_0| + 1)$, menor que ε tomando $\delta = \min(1, \varepsilon/(2|z_0| + 1))$. Los polinomios son continuos en todo $\mathbb{C}$ y, por las reglas de suma, producto y cociente de límites —que se demuestran exactamente igual que en $\mathbb{R}$—, también lo son los cocientes de polinomios allí donde el denominador no se anula.

30.2.3. La derivada compleja

Definición 30.13 (Derivada compleja)**.** Sea $f : \Omega \to \mathbb{C}$ con Ω abierto y $z_0 \in \Omega$. Se dice que f es *derivable en sentido complejo* en z_0 si existe el límite

$$f'(z_0) = \lim_{h\to 0} \frac{f(z_0 + h) - f(z_0)}{h},$$

donde h recorre los números *complejos* no nulos suficientemente pequeños. El número $f'(z_0) \in \mathbb{C}$ es la *derivada de f en z_0*.

Escrito así parece la definición de siempre, y en cierto sentido lo es: las reglas de derivación del cálculo de una variable —suma, producto, cociente, cadena, la derivada de z^n— se demuestran con la misma cuenta, porque en esas demostraciones solo se usa que se está en un cuerpo y que el límite respeta las operaciones. Lo que cambia es el conjunto por el que se mueve h: en el caso real recorre un intervalito de la recta; aquí, un disquito del plano.

Cuidado

[Por qué la derivada compleja es muchísimo más exigente que la real] Merece la pena detenerse, porque es el punto que todo el mundo cruza de largo la primera vez y luego no entiende por qué los teoremas del capítulo son tan potentes.

Una función $f : \Omega \to \mathbb{C}$ es, por el diccionario de la definición 30.6, lo mismo que una aplicación $F = (u, v)$ de un abierto de $\mathbb{R}^2$ en $\mathbb{R}^2$. Que F sea diferenciable en el sentido de varias variables significa que se aproxima cerca del punto por una aplicación *lineal* cualquiera, es decir, por una matriz 2×2: cuatro números libres. Que f sea derivable en sentido complejo significa muchísimo más: que esa aproximación lineal es *multiplicar por el número complejo* $f'(z_0)$, y las multiplicaciones por complejos tienen solo dos números libres, módulo y argumento, o sea un escalado y un giro. De las cuatro dimensiones disponibles, la derivabilidad compleja obliga a caer en un subespacio de dimensión dos: eso son dos ecuaciones, y esas dos ecuaciones son Cauchy–Riemann.

La forma operativa de ver la exigencia es mirar el cociente $(f(z_0 + h) - f(z_0))/h$ como función de la dirección por la que h tiende a 0: la definición pide que *todas* las direcciones den el mismo número. Hay un ejemplo que lo enseña de golpe.

La conjugación es continua en todo $\mathbb{C}$ y no es derivable en ningún punto. Sea $f(z) = \bar{z}$. Es continua en todas partes, porque $|\bar{z} - \bar{z}_0| = |\overline{z - z_0}| = |z - z_0|$ y basta tomar $\delta = \varepsilon$. Formemos el cociente incremental en un punto cualquiera z_0:

$$\frac{f(z_0 + h) - f(z_0)}{h} = \frac{\overline{z_0 + h} - \bar{z}_0}{h} = \frac{\bar{z}_0 + \bar{h} - \bar{z}_0}{h} = \frac{\bar{h}}{h}.$$

Nótese que z_0 ha desaparecido: el cociente vale $\bar{h}/h$ sea cual sea el punto. Ahora los dos límites direccionales:

- *por el eje real*, $h = t$ con $t \in \mathbb{R}$, $t \to 0$: entonces $\bar{h} = t$ y el cociente vale $t/t = 1$, luego el límite por esta dirección es $\mathbf{1}$;
- *por el eje imaginario*, $h = it$ con $t \in \mathbb{R}$, $t \to 0$: entonces $\bar{h} = -it$ y el cociente vale $-it/(it) = -1$, luego el límite por esta dirección es $-\mathbf{1}$.

Como $1 \neq -1$, el límite de la definición 30.13 no existe, y esto ocurre **en todos y cada uno de los puntos del plano**. Es decir: $\bar{z}$ es continua en todo $\mathbb{C}$, y tan suave como se quiera vista en $\mathbb{R}^2$ —sus partes $u = x$ y $v = -y$ son polinomios de primer grado, luego de clase C^∞—, y aun así no tiene derivada compleja en ningún punto. Nada remotamente parecido puede ocurrir en $\mathbb{R}$.
Con la lectura geométrica es aún más claro: $\bar{z}$ es la reflexión respecto del eje real, y una reflexión invierte el sentido de los ángulos; como la derivada compleja obliga a que la aproximación lineal sea un giro con escalado —que *conserva* el sentido de los ángulos—, ninguna reflexión puede ser derivable. Compárese con $f(z) = z^2$, cuyo cociente incremental es $((z_0+h)^2 - z_0^2)/h = 2z_0 + h \to 2z_0$ sin que la dirección de h intervenga en absoluto.

30.2.4. Las ecuaciones de Cauchy–Riemann, deducidas

Las dos ecuaciones que gobiernan el capítulo salen de hacer con una f genérica lo que acabamos de hacer con $\bar{z}$: comparar el límite por el eje real con el límite por el eje imaginario. Supongamos que $f = u + iv$ es derivable en $z_0 = x_0 + iy_0$.

Por el eje real. Con $h = t$ real, $z_0 + h = (x_0 + t) + iy_0$ y

$$\frac{f(z_0+t) - f(z_0)}{t} = \frac{u(x_0+t, y_0) - u(x_0, y_0)}{t} + i\,\frac{v(x_0+t, y_0) - v(x_0, y_0)}{t} \xrightarrow[t\to 0]{} \frac{\partial u}{\partial x} + i\,\frac{\partial v}{\partial x},$$

con las parciales evaluadas en (x_0, y_0). Es decir, $f'(z_0) = u_x + i\,v_x$.

Por el eje imaginario. Con $h = it$ y t real, $z_0 + h = x_0 + i(y_0 + t)$ y, usando $1/i = -i$,

$$\frac{f(z_0+it) - f(z_0)}{it} = \frac{1}{i}\left(\frac{u(x_0, y_0+t) - u(x_0, y_0)}{t} + i\,\frac{v(x_0, y_0+t) - v(x_0, y_0)}{t}\right)$$

y el límite cuando $t \to 0$ vale

$$-i\left(\frac{\partial u}{\partial y} + i\,\frac{\partial v}{\partial y}\right) = \frac{\partial v}{\partial y} - i\,\frac{\partial u}{\partial y}.$$

Es decir, $f'(z_0) = v_y - i\,u_y$. Los dos límites tienen que coincidir, porque la derivada es un único número; igualando partes reales por un lado y partes imaginarias por otro salen las dos ecuaciones.

Teorema 30.14 (Cauchy–Riemann, condición necesaria)**.** *Si $f = u + iv$ es derivable en sentido complejo en $z_0 = x_0 + iy_0$, entonces existen las cuatro derivadas parciales de u y v en (x_0, y_0) y satisfacen*

$$\frac{\partial u}{\partial x} = \frac{\partial v}{\partial y}, \qquad \frac{\partial u}{\partial y} = -\frac{\partial v}{\partial x},$$

y además $f'(z_0) = \frac{\partial u}{\partial x} + i\,\frac{\partial v}{\partial x} = \frac{\partial v}{\partial y} - i\,\frac{\partial u}{\partial y}$.

Demostración. Es la cuenta de los dos párrafos anteriores: la existencia del límite completo garantiza la de los dos límites direccionales, y su coincidencia da las dos igualdades al separar parte real y parte imaginaria. □

El recíproco es lo que uno querría usar en la práctica —comprobar dos ecuaciones y concluir que la función es derivable—, pero *no es cierto tal cual*: hay funciones cuyas cuatro parciales existen y cumplen Cauchy–Riemann en un punto y que aun así no son derivables allí, porque tener parciales no equivale a ser diferenciable en $\mathbb{R}^2$. Hace falta una hipótesis más, y la más cómoda de verificar es la continuidad de las parciales.

Teorema 30.15 (Cauchy–Riemann, condición suficiente)**.** *Sea $f = u + iv$ definida en un abierto Ω. Supongamos que en Ω las cuatro parciales u_x, u_y, v_x, v_y* existen y son continuas*, y que satisfacen las ecuaciones de Cauchy–Riemann. Entonces f es derivable en sentido complejo en todo punto de Ω, es decir, holomorfa.*

Demostración. La continuidad de las parciales implica, por el teorema correspondiente del cálculo en varias variables, que u y v son *diferenciables* en el sentido real: escribiendo $h = s+it$,

$$u(x_0+s, y_0+t) - u(x_0, y_0) = u_x s + u_y t + \varepsilon_1(h), \qquad v(x_0+s, y_0+t) - v(x_0, y_0) = v_x s + v_y t + \varepsilon_2(h),$$

con $\varepsilon_j(h)/|h| \to 0$. Sumando la segunda multiplicada por i y usando $u_y = -v_x$ y $v_y = u_x$ para eliminar dos de las cuatro parciales,

$$f(z_0 + h) - f(z_0) = (u_x + iv_x)(s + it) + \varepsilon(h) = (u_x + iv_x)\, h + \varepsilon(h),$$

con $\varepsilon(h)/|h| \to 0$. Dividiendo entre h y pasando al límite queda $f'(z_0) = u_x + iv_x$. □

Cuidado

[Las dos hipótesis que se olvidan] Primera: hace falta la *continuidad* de las parciales (o, en su defecto, la diferenciabilidad real de u y v); solo con que las parciales existan y cumplan las ecuaciones el teorema es falso. Con funciones elementales la hipótesis se cumple automáticamente y no da guerra, pero al enunciar el resultado hay que decirla. Segunda: comprobar Cauchy–Riemann *en un punto* no sirve de nada por sí solo. Como se ve enseguida, la propiedad interesante —ser holomorfa— exige derivabilidad en todo un entorno abierto, y hay funciones que cumplen las ecuaciones únicamente sobre una recta o en un punto aislado.

Ejemplo 30.16 (Cuatro funciones puestas a prueba)**.** *(a)* $f(z) = z^2$. Con $u = x^2 - y^2$ y $v = 2xy$, las cuatro parciales son $u_x = 2x$, $u_y = -2y$, $v_x = 2y$, $v_y = 2x$, todas continuas. Se cumple $u_x = v_y$ y $u_y = -v_x$ en todo el plano, luego por el teorema 30.15 f es entera, con $f' = u_x + iv_x = 2x + 2iy = 2z$: el mismo valor que dio el cociente incremental.

(b) $f(z) = e^z$. Anticipando la definición de más abajo, e^{x+iy} tiene partes $u = e^x \cos y$ y $v = e^x \sin y$, con

$$u_x = e^x \cos y, \quad u_y = -e^x \sin y, \quad v_x = e^x \sin y, \quad v_y = e^x \cos y.$$

Se ve de un vistazo que $u_x = v_y$ y $u_y = -v_x$, en todo el plano y con todas las parciales continuas: e^z es entera. Y su derivada es $f' = u_x + iv_x = e^x \cos y + ie^x \sin y = e^x(\cos y + i \sin y) = e^z$. La exponencial compleja es, igual que la real, su propia derivada.

(c) $f(z) = \bar{z}$. Aquí $u = x$ y $v = -y$, luego $u_x = 1$ y $v_y = -1$: la primera ecuación pide $1 = -1$, falso en todo punto. Coincide con el cálculo de los dos límites direccionales de la advertencia anterior: la conjugación no es derivable en ningún sitio.

(d) $f(z) = |z|^2$. Como $|z|^2 = x^2 + y^2$ es real, $u = x^2 + y^2$ y $v = 0$. Las ecuaciones piden $2x = 0$ y $2y = 0$: se cumplen *únicamente* en $z = 0$. Por tanto $|z|^2$ es derivable en el origen —con $f'(0) = u_x + iv_x = 0$— y en ningún otro punto, luego no es holomorfa en ninguna parte, ni siquiera en el origen, porque no hay ningún abierto donde sea derivable. Es el contraejemplo que conviene tener a mano, porque $|z|^2$ es de clase C^∞ como función de dos variables reales. Los apartados (a), (c) y (d) se retoman, con la lectura geométrica, en el ejemplo de las tres funciones de la sección anterior.

30.2.5. Holomorfa, entera, analítica: tres palabras y su relación

Definición 30.17 (Función holomorfa)**.** Sea $\Omega \subseteq \mathbb{C}$ abierto. Se dice que $f : \Omega \to \mathbb{C}$ es *holomorfa* en Ω si es derivable en sentido complejo en *todos* los puntos de Ω. Se dice que f es holomorfa *en un punto* z_0 si lo es en algún abierto que contenga a z_0; es decir, si es derivable en todo un entorno de z_0, y no meramente en z_0.

Definición 30.18 (Función entera)**.** f es *entera* si es holomorfa en todo el plano, es decir, si $\Omega = \mathbb{C}$. Volveremos sobre este término, con ejemplos y contraejemplos, en la definición 30.99.

Definición 30.19 (Función analítica)**.** $f : \Omega \to \mathbb{C}$ es *analítica* en Ω si para cada $z_0 \in \Omega$ existen un radio $\rho > 0$ y una serie de potencias $\sum_{n \geq 0} a_n (z - z_0)^n$ que converge a $f(z)$ para todo z con $|z - z_0| < \rho$. Dicho brevemente: si localmente es la suma de su serie de potencias.

Observación 30.20 (Qué relación hay entre las tres palabras en este libro)*.* Las definiciones de holomorfa y de analítica son *a priori* muy distintas: una habla de un límite, la otra de una serie infinita. En análisis real las dos nociones análogas no coinciden en absoluto: la función $g(x) = e^{-1/x^2}$, extendida con $g(0) = 0$, es infinitamente derivable en toda la recta y todas sus derivadas se anulan en el origen, así que su serie de Taylor en 0 es la serie nula y no representa a g en ningún entorno; g es C^∞ y no es analítica.

En $\mathbb{C}$, en cambio, **las dos nociones son exactamente la misma**, y esa equivalencia es uno de los resultados centrales del capítulo: se demuestra con la fórmula integral de Cauchy en la sección siguiente. Toda función holomorfa resulta ser automáticamente infinitamente derivable y coincidir con su serie de Taylor en cada disco contenido en el dominio. Por eso, en este libro y en la literatura en general, *holomorfa* y *analítica* se usan como sinónimos, y a partir de la próxima sección lo haremos sin más aviso; pero conviene tener claro que eso es un **teorema** y no un convenio: hasta que no se demuestre, las dos palabras significan cosas distintas y no se pueden intercambiar. *Entera* no es sinónimo de ninguna de las dos: es holomorfa *más* la condición geométrica de que el dominio sea todo el plano.

Definición 30.21 (Serie de potencias y radio de convergencia)**.** Una *serie de potencias* centrada en $z_0 \in \mathbb{C}$ es una expresión $\sum_{n\geq 0} a_n (z-z_0)^n$ con $a_n \in \mathbb{C}$. Su *radio de convergencia* es

$$R = \frac{1}{\limsup_{n\to\infty} |a_n|^{1/n}} \in [0,\infty],$$

con los convenios $1/0 = \infty$ y $1/\infty = 0$. La serie converge absolutamente si $|z-z_0| < R$ y diverge si $|z-z_0| > R$; sobre la circunferencia $|z-z_0| = R$ puede ocurrir cualquier cosa. Al conjunto $\{|z-z_0| < R\}$ se le llama *disco de convergencia*.

Ejemplo 30.22 (La serie geométrica, que es el modelo de todas)**.** Para $\sum_{n\geq 0} z^n$ todos los coeficientes valen 1, luego $|a_n|^{1/n} = 1$ y $R = 1$. Dentro del disco unidad la suma es $\dfrac{1}{1-z}$, por el mismo cálculo telescópico que en el caso real: $(1-z)(1+z+\cdots+z^N) = 1 - z^{N+1}$, y $z^{N+1} \to 0$ cuando $|z| < 1$. Control numérico en $z = i/2$: la suma exacta es $\dfrac{1}{1-i/2} = \dfrac{2}{2-i} = \dfrac{2(2+i)}{5} = 0{,}8 + 0{,}4\,i$, y sumando los veinte primeros términos se obtiene $0{,}799999 + 0{,}400000\,i$. Para $|z| > 1$ el término general no tiende a cero y la serie diverge sin remedio. Obsérvese que $1/(1-z)$ deja de ser holomorfa exactamente en $z = 1$, a distancia 1 del centro: el radio de convergencia es la distancia a la singularidad más cercana. Esa idea, con seis ejemplos más y los criterios prácticos para calcular R, está desarrollada en la regla sobre radios de convergencia de este mismo capítulo.

30.2.6. Funciones armónicas y armónicas conjugadas

Definición 30.23 (Función armónica)**.** Una función real $u : \Omega \to \mathbb{R}$, con $\Omega \subseteq \mathbb{R}^2$ abierto, de clase C^2, es *armónica* si satisface la *ecuación de Laplace*

$$\Delta u := u_{xx} + u_{yy} = 0 \qquad \text{en todo } \Omega.$$

Definición 30.24 (Armónica conjugada)**.** Dadas dos funciones armónicas u, v en Ω, se dice que v es una *armónica conjugada* de u si juntas satisfacen las ecuaciones de Cauchy–Riemann, esto es, si $u_x = v_y$ y $u_y = -v_x$; equivalentemente, si $f = u + iv$ es holomorfa en Ω.

Proposición 30.25 (Las dos partes de una holomorfa son armónicas)**.** *Si $f = u + iv$ es holomorfa en Ω, entonces u y v son armónicas en Ω.*

Demostración. Admitimos aquí el hecho —que se demostrará con la fórmula integral de Cauchy— de que una función holomorfa es infinitamente derivable, de modo que u y v son de clase C^∞ y sus parciales cruzadas coinciden. Derivando la primera ecuación de Cauchy–Riemann respecto de x y la segunda respecto de y se obtiene $u_{xx} = v_{yx}$ y $u_{yy} = -v_{xy}$; sumando y usando $v_{yx} = v_{xy}$ queda $u_{xx} + u_{yy} = 0$. Para v se procede al revés: derivando la primera respecto de y y la segunda respecto de x salen $u_{xy} = v_{yy}$ y $u_{yx} = -v_{xx}$, luego $v_{xx} + v_{yy} = 0$. □

Este resultado es la razón por la que todo problema plano de potencial —temperaturas estacionarias, campos electrostáticos, flujos incompresibles e irrotacionales— es análisis

complejo disfrazado. Y tiene un recíproco local muy útil: toda función armónica en un disco es la parte real de alguna función holomorfa, y encontrarla es puramente mecánico.

Regla: hallar la armónica conjugada de una función armónica dada

Se parte de $u(x,y)$ y se busca v tal que $f = u + iv$ sea holomorfa. Es la contrapartida compleja del cálculo de un potencial escalar a partir de un campo conservativo, con Cauchy–Riemann en el papel del gradiente.

1. **Comprueba que u es armónica.** Calcula u_{xx} y u_{yy} y verifica que suman cero. Si no lo hacen, no existe ninguna v: es un filtro, no un trámite.
2. **Integra la primera ecuación.** De $v_y = u_x$ se obtiene $v(x,y) = \int u_x(x,y)\,\mathrm{d}y + \varphi(x)$, integrando respecto de y con x como parámetro. La "constante" de integración es una función arbitraria de x, no un número.
3. **Determina φ con la segunda ecuación.** Impón $v_x = -u_y$; al derivar la expresión del paso 2 respecto de x, todo se cancela salvo $\varphi'(x)$, que queda despejada. Intégrala: ahora sí, lo que sobra es una constante real c.
4. **Escribe $f = u + iv$ y comprueba.** Verifica las dos ecuaciones sobre el resultado final; es una comprobación de treinta segundos que atrapa casi todos los errores de signo.
5. **Identifica f en función de z.** El atajo estándar: evalúa f sobre el eje real, es decir, pon $y = 0$, y sustituye x por z. Funciona por el principio de identidad, que se demostrará más adelante.
6. **Recuerda la ambigüedad y el límite del método.** La armónica conjugada está determinada *salvo una constante real aditiva*, o sea, f está determinada salvo sumarle una constante imaginaria pura. Y en dominios con agujeros puede no existir globalmente aunque exista en cada disco: el ejemplo canónico es $u = \ln|z|$ en $\mathbb{C}\setminus\{0\}$, cuya conjugada querría ser $\arg z$, que no es uniforme.

Ejemplo 30.26 (De $u = e^x\cos y$ a la función holomorfa, entero)**.** *Paso 1.* $u_x = e^x\cos y$ y $u_{xx} = e^x\cos y$; $u_y = -e^x\sin y$ y $u_{yy} = -e^x\cos y$. Sumando, $u_{xx} + u_{yy} = e^x\cos y - e^x\cos y = 0$: armónica en todo el plano.

Paso 2. La primera ecuación dice $v_y = u_x = e^x\cos y$. Integrando respecto de y con x fijo,

$$v(x,y) = \int e^x\cos y\,\mathrm{d}y + \varphi(x) = e^x\sin y + \varphi(x).$$

Paso 3. La segunda dice $v_x = -u_y$. Por un lado $v_x = e^x\sin y + \varphi'(x)$; por otro $-u_y = e^x\sin y$. Igualando,

$$e^x\sin y + \varphi'(x) = e^x\sin y \implies \varphi'(x) = 0 \implies \varphi(x) = c,$$

y tomamos $c = 0$.

Paso 4. Queda $f = e^x\cos y + i\,e^x\sin y$. Comprobación: $u_x = e^x\cos y$ y $v_y = e^x\cos y$, iguales; $u_y = -e^x\sin y$ y $v_x = e^x\sin y$, opuestas. Correcto.

Paso 5. Sobre el eje real, $y=0$, se tiene $u=e^x$ y $v=0$, luego $f=e^x$; sustituyendo x por z conjeturamos $f(z)=e^z$, y en efecto la definición de la exponencial compleja que damos a continuación es exactamente $e^x(\cos y + i\sin y)$. Hemos *reconstruido* la exponencial compleja partiendo únicamente de su parte real.

Paso 6. La solución general es $f(z)=e^z+ic$ con $c\in\mathbb{R}$: sumar una constante imaginaria pura no cambia la parte real, y no hay más libertad que esa. Un segundo ejemplo, el de $u=x^3-3xy^2$ que conduce a z^3, está resuelto paso a paso más arriba en este mismo capítulo.

30.2.7. Las funciones elementales complejas

Las series de e^x, $\sin x$ y $\cos x$ del cálculo real convergen para todo valor de la variable y sus términos son polinomios: nada impide sustituir x por un número complejo. Eso es lo que se hace, y todo funciona. Damos las definiciones en la forma más operativa.

Definición 30.27 (Exponencial compleja)**.** Para $z=x+iy$ se define

$$e^z := e^x\big(\cos y + i\sin y\big),$$

donde e^x, $\cos y$ y $\sin y$ son las funciones reales de siempre. Equivalentemente, $e^z=\sum_{n\ge 0} z^n/n!$, serie que converge absolutamente para todo $z\in\mathbb{C}$.

Proposición 30.28 (Propiedades de la exponencial)**.** *Para todos $z,w\in\mathbb{C}$: (1) $e^{z+w}=e^z e^w$; (2) $e^z\neq 0$ siempre, y $|e^z|=e^{\operatorname{Re} z}$; (3) e^z es entera y $\frac{\mathrm{d}}{\mathrm{d}z}e^z=e^z$; (4) $e^{z+2\pi i}=e^z$, es decir, la exponencial compleja es* periódica de periodo $2\pi i$*, y sus únicos periodos son los múltiplos enteros de $2\pi i$.*

Demostración. (1) es la fórmula de adición del seno y el coseno: si $z=x+iy$ y $w=s+it$, entonces $e^z e^w = e^{x+s}(\cos y+i\sin y)(\cos t+i\sin t) = e^{x+s}\big(\cos(y+t)+i\sin(y+t)\big)=e^{z+w}$. (2) sale de $|e^z|=e^x\sqrt{\cos^2 y+\sin^2 y}=e^x>0$. (3) se comprobó en el apartado (b) del ejemplo anterior. Para (4), $e^{2\pi i}=\cos 2\pi + i\sin 2\pi = 1$, luego $e^{z+2\pi i}=e^z e^{2\pi i}=e^z$; recíprocamente, si $e^w=1$ entonces $e^{\operatorname{Re} w}=1$ obliga a $\operatorname{Re} w=0$, y $\cos(\operatorname{Im} w)=1$ obliga a $\operatorname{Im} w\in 2\pi\mathbb{Z}$. □

La periodicidad es la diferencia estructural con el caso real y la fuente de todo lo que viene después: e^x es inyectiva sobre $\mathbb{R}$, pero e^z toma cada valor no nulo *infinitas* veces, en puntos separados verticalmente por $2\pi i$. Por eso su inversa, el logaritmo, no podrá ser una función.

Definición 30.29 (Seno y coseno complejos)**.** Para $z\in\mathbb{C}$ se define, invirtiendo las fórmulas de Euler,

$$\cos z := \frac{e^{iz}+e^{-iz}}{2}, \qquad \sin z := \frac{e^{iz}-e^{-iz}}{2i}.$$

Sobre el eje real estas fórmulas devuelven el seno y el coseno de siempre, así que la definición *extiende* las funciones reales y no las contradice. Y todas las identidades algebraicas habituales siguen valiendo, porque son consecuencias formales de $e^{z+w}=e^z e^w$: en particular $\sin^2 z+\cos^2 z=1$ para todo $z\in\mathbb{C}$, ambas son enteras, $(\sin z)'=\cos z$ y $(\cos z)'=-\sin z$.

Cuidado

[El seno y el coseno complejos NO están acotados] En la recta real, $|\sin x| \le 1$ y $|\cos x| \le 1$ son de las primeras cosas que uno aprende. En el plano complejo son **rotundamente falsas**, y confiar en ellas es uno de los errores más caros del capítulo, porque es el que arruina la elección del contorno al calcular integrales reales.
Hagamos la cuenta con $z = i$. Por la definición,

$$\cos(i) = \frac{e^{i\cdot i} + e^{-i\cdot i}}{2} = \frac{e^{-1} + e^{1}}{2} = \cosh(1) = \frac{0{,}367879 + 2{,}718282}{2} = 1{,}543081,$$

que ya es mayor que 1. Con $z = 10i$ la cosa se dispara: $\cos(10i) = \cosh(10) = 11\,013{,}23$. En general $\cos(iy) = \cosh y$ y $\sin(iy) = i\sinh y$, de modo que sobre el eje imaginario el seno y el coseno crecen exponencialmente. La razón es transparente en la definición: $\cos z$ es una combinación de e^{iz} y e^{-iz}, y $|e^{iz}| = e^{-\operatorname{Im} z}$, que explota cuando $\operatorname{Im} z \to -\infty$ y muere cuando $\operatorname{Im} z \to +\infty$; al alejarse del eje real, uno de los dos sumandos siempre gana y arrastra al conjunto.
Consecuencia práctica, que se usará varias veces: para calcular una integral real con $\cos ax$ o $\sin ax$ en el integrando **nunca** se sustituye por $\cos az$ o $\sin az$ en el plano complejo, porque el arco de cierre no moriría. Se sustituye por e^{iaz}, que sí decae en el semiplano superior, y al final se toma la parte real o la imaginaria.

Definición 30.30 (Logaritmo complejo, rama principal y corte de rama)**.** Dado $z \in \mathbb{C} \setminus \{0\}$, se llama *logaritmo* de z a cualquier $w \in \mathbb{C}$ con $e^w = z$. Escribiendo $z = re^{i\theta}$ con $r = |z| > 0$, las soluciones son exactamente

$$\log z = \ln r + i\theta + 2\pi i k, \qquad k \in \mathbb{Z},$$

donde ln es el logaritmo real. Así que log es una función *multivaluada*: cada $z \neq 0$ tiene infinitos logaritmos, separados entre sí por múltiplos de $2\pi i$.

Una *rama* del logaritmo es una elección continua de uno de esos valores sobre un abierto. La *rama principal* Log es la que toma el argumento en $(-\pi, \pi]$:

$$\operatorname{Log} z = \ln|z| + i \operatorname{Arg} z, \qquad \operatorname{Arg} z \in (-\pi, \pi].$$

El rayo donde esa elección deja de ser continua —para la rama principal, el semieje real negativo junto con el origen— se llama *corte de rama*. El origen es el *punto de ramificación*: ninguna rama puede definirse continuamente en un entorno perforado suyo.

Definición 30.31 (Potencias complejas)**.** Para $z \neq 0$ y $a \in \mathbb{C}$ se define $z^a := e^{a \log z}$, expresión multivaluada salvo que se fije una rama del logaritmo; con la rama principal se obtiene el *valor principal* de z^a. Si $a \in \mathbb{Z}$ el resultado es unívoco, porque cambiar de rama lo multiplica por $e^{2\pi i a} = 1$; si $a = 1/n$ hay exactamente n valores, las n raíces n-ésimas.

Ejemplo 30.32 (Tres cálculos numéricos con logaritmos y potencias)**.** *(a)* $\log(-1)$. El módulo de -1 es 1 y su argumento principal es π, o sea $-1 = 1 \cdot e^{i\pi}$. Entonces

$$\log(-1) = \ln 1 + i\pi + 2\pi i k = i\pi + 2\pi i k, \qquad k \in \mathbb{Z},$$

y el valor principal es $\operatorname{Log}(-1) = i\pi \approx 3{,}141593\,i$. El logaritmo de un número negativo existe perfectamente en $\mathbb{C}$: lo que no existe es un *único* logaritmo.

(b) i^i. Como $i = e^{i\pi/2}$, sus logaritmos son $\log i = i\pi/2 + 2\pi ik$, y por la definición de potencia

$$i^i = e^{i\log i} = e^{i(i\pi/2+2\pi ik)} = e^{-\pi/2-2\pi k},$$

que es un número *real positivo* para cada k. El valor principal ($k = 0$) es $i^i = e^{-\pi/2} = 0{,}207880$. Que una potencia imaginaria de un número imaginario salga real deja de ser paradójico en cuanto se ve que solo estamos evaluando la exponencial real en un exponente que resultó ser real.

(c) $\sqrt{i}$. Por definición $\sqrt{i} = i^{1/2} = e^{\frac{1}{2}\log i}$, y con la rama principal $\log i = i\pi/2$, luego

$$\sqrt{i} = e^{i\pi/4} = \cos\frac{\pi}{4} + i\sin\frac{\pi}{4} = \frac{\sqrt{2}}{2} + \frac{\sqrt{2}}{2}\,i \approx 0{,}707107 + 0{,}707107\,i.$$

Verificación: $\left(\frac{\sqrt{2}}{2} + \frac{\sqrt{2}}{2}i\right)^2 = \frac{1}{2} + 2\cdot\frac{1}{2}i + \frac{1}{2}i^2 = \frac{1}{2} - \frac{1}{2} + i = i$. Correcto. El otro valor, con $k = 1$, es el opuesto: las dos raíces cuadradas de siempre.

Las trampas de este terreno —que $\log(zw)$ no es $\log z + \log w$ como igualdad de ramas, que $\sqrt{z}\sqrt{w}$ no es $\sqrt{zw}$, que $(z^a)^b$ no es z^{ab}— están tratadas con detalle en la regla sobre ramas del logaritmo y su ejemplo, más adelante en el capítulo.

30.2.8. Caminos e integral de contorno

Todo lo que queda del capítulo se apoya en integrar funciones complejas a lo largo de curvas. Hay que decir con precisión qué es una curva y qué es esa integral.

Definición 30.33 (Camino y camino cerrado simple)**.** Un *camino* (o *arco*, o *curva*) en $\mathbb{C}$ es una aplicación continua $\gamma : [a, b] \to \mathbb{C}$ de clase C^1 a trozos, es decir, derivable con derivada continua salvo quizá en un número finito de puntos, donde puede tener esquinas. Los puntos $\gamma(a)$ y $\gamma(b)$ son su *origen* y su *extremo*, y su *longitud* es

$$L(\gamma) = \int_a^b |\gamma'(t)|\,\mathrm{d}t.$$

El camino es *cerrado* si $\gamma(a) = \gamma(b)$, y *simple* si no se corta a sí mismo, esto es, si γ es inyectiva en $[a, b)$. Un *camino cerrado simple* es un camino cerrado y simple: sale de un punto, da una vuelta sin cruzarse y regresa. Por convenio, y salvo aviso, los caminos cerrados se recorren en *sentido positivo* o antihorario.

El camino más importante del capítulo es la circunferencia de centro z_0 y radio $r > 0$ recorrida una vez en sentido positivo, que se parametriza como $\gamma(t) = z_0 + re^{it}$ con $t \in [0, 2\pi]$ y $\gamma'(t) = ire^{it}$. Es cerrado y simple, y su longitud es $\int_0^{2\pi} |ire^{it}|\mathrm{d}t = \int_0^{2\pi} r\,\mathrm{d}t = 2\pi r$, como debe ser.

Definición 30.34 (Integral de contorno)**.** Sea $\gamma : [a, b] \to \mathbb{C}$ un camino y f una función continua sobre la imagen de γ. La *integral de contorno* de f a lo largo de γ es

$$\int_\gamma f(z)\,\mathrm{d}z := \int_a^b f\big(\gamma(t)\big)\,\gamma'(t)\,\mathrm{d}t,$$

donde el miembro derecho es la integral de una función continua a trozos de variable *real* t con valores complejos, que se calcula integrando por separado su parte real y su parte imaginaria:

$$\int_a^b (P(t)+iQ(t))\,\mathrm{d}t := \int_a^b P(t)\,\mathrm{d}t + i\int_a^b Q(t)\,\mathrm{d}t.$$

Cuando el camino es cerrado se escribe $\oint_\gamma$.

Es decir: la integral compleja no es un objeto nuevo, es una integral de toda la vida sobre un intervalo real, con el integrando fabricado sustituyendo z por $\gamma(t)$ y $\mathrm{d}z$ por $\gamma'(t)\,\mathrm{d}t$. Todo el arte consistirá en elegir bien el camino.

Regla: calcular una integral de contorno directamente desde la definición

1. **Parametriza el camino.** Escribe $\gamma(t)$ con t recorriendo un intervalo $[a,b]$ y comprueba que el sentido es el que pide el problema. Las dos parametrizaciones que cubren casi todo: el segmento de p a q es $\gamma(t)=p+t(q-p)$ con $t\in[0,1]$; la circunferencia de centro z_0 y radio r es $\gamma(t)=z_0+re^{it}$ con $t\in[0,2\pi]$.
2. **Calcula $\gamma'(t)$.** Es una derivada ordinaria respecto de la variable real t: para el segmento sale la constante $q-p$; para la circunferencia, ire^{it}.
3. **Sustituye y simplifica antes de integrar.** Forma el integrando $f(\gamma(t))\,\gamma'(t)$. En caminos circulares conviene dejarlo todo en exponenciales e^{ikt}: casi siempre se cancela más de lo que uno espera.
4. **Integra en t.** Si quedó en la forma $\int_a^b ce^{ikt}\mathrm{d}t$, úsese la primitiva $ce^{ikt}/(ik)$ para $k\neq 0$ —que da 0 sobre un periodo completo— y $c(b-a)$ para $k=0$. Si no, sepárense parte real y parte imaginaria y hágase cada una por su lado.
5. **Si el camino tiene esquinas, trocéalo.** Una poligonal se integra sumando las integrales de sus lados, cada uno con su propia parametrización.
6. **Comprueba orientación y orden de magnitud.** Invertir el sentido cambia el signo. Y la estimación ML del lema 30.38 da en dos líneas una cota superior del módulo: si el resultado la supera, hay un error.

Ejemplo 30.35 (La integral fundamental del capítulo entero)**.** Calculemos $\displaystyle\oint_{|z|=1}\frac{\mathrm{d}z}{z}$ sobre la circunferencia unidad recorrida una vez en sentido positivo.

Paso 1. Parametrizamos $\gamma(t)=e^{it}$ con $t\in[0,2\pi]$, que recorre la circunferencia unidad una vez en sentido antihorario.

Paso 2. $\gamma'(t)=ie^{it}$.

Paso 3. El integrando es

$$f(\gamma(t))\,\gamma'(t)=\frac{1}{e^{it}}\cdot ie^{it}=i.$$

Toda la dependencia en t se ha cancelado: el integrando es la constante i. Esa cancelación es el corazón del asunto, y ocurre porque $1/z$ decrece exactamente al mismo ritmo al que crece el factor γ'.

Paso 4. Integramos:

$$\oint_{|z|=1} \frac{\mathrm{d}z}{z} = \int_0^{2\pi} i\,\mathrm{d}t = i\,(2\pi - 0) = 2\pi i \approx 6{,}283185\,i.$$

Comprobación. Sobre la circunferencia de centro el origen y radio arbitrario r el integrando es $\frac{1}{re^{it}} \cdot ire^{it} = i$ otra vez: el resultado es $2\pi i$ *independientemente del radio.* Y recorriendo en sentido horario, con $\gamma(t) = e^{-it}$, sale $-2\pi i$.

Por qué importa. Este $2\pi i$ es, literalmente, todo el análisis complejo en miniatura. Es lo que impide que $1/z$ tenga primitiva en el plano perforado —una primitiva querría ser $\log z$, que salta $2\pi i$ al dar una vuelta al origen—; es el único término de una serie de Laurent que sobrevive a la integración, y por eso el residuo se define como el coeficiente de $1/(z - z_0)$; y es el factor que aparece en la fórmula integral de Cauchy y en el teorema de los residuos. Todo el resto del capítulo es contabilidad alrededor de este cálculo de cuatro líneas.

Ejemplo 30.36 (Dos caminos, dos desenlaces: z^2 frente a $\bar{z}$)**.** Integremos del punto 0 al punto $1 + i$ por dos caminos. *Camino A*: el segmento recto, $\alpha(t) = (1 + i)t$ con $t \in [0, 1]$ y $\alpha'(t) = 1 + i$. *Camino B*: la poligonal que va de 0 a 1 y luego de 1 a $1 + i$, en dos trozos, $\beta_1(t) = t$ con $\beta_1' = 1$ y $\beta_2(t) = 1 + it$ con $\beta_2' = i$, ambos con $t \in [0, 1]$.

(a) Con $f(z) = z^2$, holomorfa. Por el camino A,

$$\int_A z^2\,\mathrm{d}z = \int_0^1 \big((1+i)t\big)^2 (1+i)\,\mathrm{d}t = (1+i)^3 \int_0^1 t^2\,\mathrm{d}t = \frac{(1+i)^3}{3},$$

y como $(1+i)^3 = (1+i)(2i) = 2i-2 = -2+2i$, la integral vale $\dfrac{-2+2i}{3} \approx -0{,}666667+0{,}666667\,i$. Por el camino B, el primer trozo da $\int_0^1 t^2\,\mathrm{d}t = \frac{1}{3}$ y el segundo

$$\int_0^1 (1+it)^2\, i\,\mathrm{d}t = i\int_0^1 (1 + 2it - t^2)\,\mathrm{d}t = i\Big(1 + i - \tfrac{1}{3}\Big) = i\Big(\tfrac{2}{3} + i\Big) = -1 + \tfrac{2}{3}i.$$

Sumando, $\frac{1}{3} - 1 + \frac{2}{3}i = -\frac{2}{3} + \frac{2}{3}i$: exactamente lo mismo que por el camino A. **La integral no ha dependido del camino**, y no es casualidad: z^2 tiene primitiva $z^3/3$, y evaluándola en los extremos sale $\frac{(1+i)^3}{3} - 0 = \frac{-2+2i}{3}$. Esa es la esencia del teorema de Cauchy de la sección siguiente.

(b) Con $f(z) = \bar{z}$, que no es holomorfa. Por el camino A, usando $\overline{(1+i)t} = (1-i)t$,

$$\int_A \bar{z}\,\mathrm{d}z = \int_0^1 (1-i)t\,(1+i)\,\mathrm{d}t = (1-i)(1+i)\int_0^1 t\,\mathrm{d}t = 2 \cdot \tfrac{1}{2} = 1.$$

Por el camino B, el primer trozo da $\int_0^1 t\,\mathrm{d}t = \frac{1}{2}$ y el segundo, con $\overline{1+it} = 1 - it$,

$$\int_0^1 (1-it)\,i\,\mathrm{d}t = \int_0^1 (i + t)\,\mathrm{d}t = i + \tfrac{1}{2},$$

de modo que el total es $\frac{1}{2} + i + \frac{1}{2} = 1 + i$. Los dos caminos dan resultados **distintos**: 1 frente a $1 + i$. La integral de contorno de una función continua depende, en general, del camino; que no dependa es una propiedad especial, y la sección siguiente demostrará que la holomorfía es exactamente lo que la garantiza. De paso: sobre el camino cerrado "ir por A y volver por B" se obtiene $\oint \bar{z}\,\mathrm{d}z = 1 - (1 + i) = -i \neq 0$.

Proposición 30.37 (Independencia de la parametrización). *Sea $\gamma : [a,b] \to \mathbb{C}$ un camino y $\sigma : [c,d] \to [a,b]$ una biyección de clase C^1 con $\sigma' > 0$, $\sigma(c) = a$ y $\sigma(d) = b$. Entonces $\tilde\gamma = \gamma \circ \sigma$ es un camino con la misma imagen y el mismo sentido, y para toda f continua*

$$\int_{\tilde\gamma} f(z)\,\mathrm{d}z = \int_{\gamma} f(z)\,\mathrm{d}z.$$

Si en cambio $\sigma' < 0$, con lo que se invierte el sentido, la integral cambia de signo.

Demostración. Por la regla de la cadena, $\tilde\gamma'(s) = \gamma'(\sigma(s))\,\sigma'(s)$, luego

$$\int_{\tilde\gamma} f = \int_c^d f\big(\gamma(\sigma(s))\big)\,\gamma'(\sigma(s))\,\sigma'(s)\,\mathrm{d}s = \int_a^b f(\gamma(t))\gamma'(t)\,\mathrm{d}t = \int_\gamma f,$$

donde el paso central es el cambio de variable $t = \sigma(s)$ en una integral real ordinaria. Si $\sigma' < 0$ los límites se intercambian y aparece el signo menos. □

Esto es lo que permite hablar de "la integral sobre la circunferencia unidad" sin decir con qué fórmula se ha parametrizado: solo importan el recorrido y el sentido. Por ejemplo, $\tilde\gamma(s) = e^{2\pi i s}$ con $s \in [0,1]$ tiene $\tilde\gamma'(s) = 2\pi i e^{2\pi i s}$, el integrando de $1/z$ es la constante $2\pi i$ y la integral sobre $[0,1]$ vuelve a valer $2\pi i$.

Lema 30.38 (Estimación ML). *Sea γ un camino de longitud L y f continua sobre su imagen, con $|f(z)| \le M$ en todos los puntos de esa imagen. Entonces*

$$\left|\int_\gamma f(z)\,\mathrm{d}z\right| \le M\,L.$$

Demostración. Hace falta primero la desigualdad $\left|\int_a^b g(t)\,\mathrm{d}t\right| \le \int_a^b |g(t)|\,\mathrm{d}t$ para g continua a trozos con valores complejos, que no es del todo obvia porque el módulo no es monótono. Se prueba así: sea $I = \int_a^b g(t)\mathrm{d}t$; si $I = 0$ no hay nada que demostrar, y si no, escríbase $I = |I|e^{i\alpha}$, de modo que

$$|I| = e^{-i\alpha} I = \int_a^b e^{-i\alpha} g(t)\,\mathrm{d}t.$$

El miembro izquierdo es real, luego el derecho también y coincide con su parte real:

$$|I| = \int_a^b \mathrm{Re}\big(e^{-i\alpha}g(t)\big)\mathrm{d}t \le \int_a^b \big|e^{-i\alpha}g(t)\big|\,\mathrm{d}t = \int_a^b |g(t)|\,\mathrm{d}t,$$

usando que $\mathrm{Re}\, w \le |w|$ y que $|e^{-i\alpha}| = 1$. Aplicándolo con $g(t) = f(\gamma(t))\gamma'(t)$,

$$\left|\int_\gamma f\,\mathrm{d}z\right| = \left|\int_a^b f(\gamma(t))\gamma'(t)\,\mathrm{d}t\right| \le \int_a^b |f(\gamma(t))|\,|\gamma'(t)|\,\mathrm{d}t \le M\int_a^b |\gamma'(t)|\,\mathrm{d}t = M\,L. \qquad \square$$

Ejemplo 30.39 (La estimación ML en acción, y cuánto se pierde con ella). *(a) Un caso en que es exacta.* Para $f(z) = 1/z$ sobre $|z| = 1$ se tiene $M = 1$ —el módulo vale exactamente 1 en toda la curva— y $L = 2\pi$, luego la cota es 2π; y el valor real de la integral es $2\pi i$, de módulo 2π. La desigualdad es una igualdad, porque el integrando resultó ser constante.

(b) Un caso realista. Acotemos $\displaystyle\oint_{|z|=2} \frac{e^z}{z^2+1}\,\mathrm{d}z$ sin calcularla. Sobre $|z|=2$, el numerador cumple $|e^z| = e^{\operatorname{Re} z} \le e^2 = 7{,}389056$, con el máximo alcanzado en $z=2$; y el denominador cumple, por la desigualdad triangular invertida, $|z^2+1| \ge |z|^2 - 1 = 4-1 = 3$. Por tanto $M = e^2/3 = 2{,}463019$, y como $L = 2\pi\cdot 2 = 4\pi = 12{,}566371$,

$$\left|\oint_{|z|=2} \frac{e^z}{z^2+1}\,\mathrm{d}z\right| \le \frac{e^2}{3}\cdot 4\pi = 30{,}951206.$$

El valor exacto, que sabremos calcular con el teorema de los residuos, es $2\pi i \sin(1)$, de módulo $5{,}287118$: la cota es correcta pero generosa, se ha perdido casi un factor 6. Eso es lo normal con ML, y no es un defecto, porque casi siempre se usa para demostrar que algo *tiende a cero*, no para estimar con precisión. El ejemplo típico: sobre el arco de radio R en el semiplano superior, un integrando que decae como R^{-2} tiene $M \sim R^{-2}$ y $L = \pi R$, de modo que $ML \sim \pi/R \to 0$ y el arco desaparece del cálculo. Todo el método de integrales reales por residuos descansa en argumentos de esta forma.

Definición 30.40 (Singularidad aislada y polo)**.** Sea f holomorfa en un disco perforado $\{z : 0 < |z-z_0| < R\}$ para algún $R>0$, y no holomorfa (o no definida) en z_0. Se dice entonces que z_0 es una *singularidad aislada* de f: aislada porque hay todo un disco a su alrededor libre de otras singularidades.

Se dice que z_0 es un *polo de orden* $m \ge 1$ si existe un entero m tal que el límite

$$\lim_{z\to z_0} (z-z_0)^m f(z)$$

existe, es finito y no vale cero. Un polo de orden 1 se llama *simple* y uno de orden 2, *doble*. Estas nociones se reformularán en términos de la serie de Laurent más adelante en el capítulo, donde se define también el *residuo* de f en z_0; las dos descripciones son equivalentes.

Ejemplo 30.41 (Reconocer una singularidad aislada y el orden de un polo)**.** Para $f(z) = \dfrac{z+3}{z^2(z-1)}$ hay dos singularidades aisladas, $z=0$ y $z=1$, porque el denominador se anula exactamente ahí y el numerador no. En $z=1$ multiplicamos por $(z-1)^1$ y tomamos límite: $\lim_{z\to 1} \dfrac{z+3}{z^2} = \dfrac{4}{1} = 4$, finito y no nulo, luego es un polo simple. En $z=0$, con $m=1$ el límite $\lim_{z\to 0} \dfrac{z+3}{z(z-1)}$ es infinito, así que $m=1$ no basta; con $m=2$,

$$\lim_{z\to 0} z^2\cdot\frac{z+3}{z^2(z-1)} = \lim_{z\to 0}\frac{z+3}{z-1} = \frac{3}{-1} = -3,$$

finito y no nulo: polo de orden 2, es decir, doble. La receta general es la del enunciado: subir m de uno en uno hasta que el límite se estabilice en un número finito distinto de cero.

Dos contraejemplos instructivos. El origen *no* es una singularidad aislada de $1/\sin(1/z)$, porque los puntos $z = 1/(k\pi)$ donde se anula el denominador se acumulan en él y no dejan ningún disco perforado libre. Y para $\sqrt{z}$ o $\log z$ el origen tampoco lo es: es un punto de ramificación, y ninguna rama es holomorfa en un disco perforado completo a su alrededor.

Pausa para pensar

Repasa la lista de lo que ha hecho falta para llegar hasta aquí: la definición de límite, la de derivada, dos ecuaciones en derivadas parciales y la definición de una integral sobre un intervalo real. Nada más. Y sin embargo, con estos materiales, la sección siguiente demostrará que una función holomorfa está determinada en todo el interior de una curva por sus valores sobre la curva, y que derivar una sola vez implica poder derivar infinitas. Merece la pena preguntarse dónde entra exactamente esa fuerza. La respuesta, en una frase: en el "para todas las direcciones" de la definición 30.13. Todo lo demás son consecuencias.

30.3. El teorema de Cauchy y la fórmula integral

Idea clave

Intuición primero. Este es el resultado que hace que la variable compleja sea otro mundo. El teorema de Cauchy dice que la integral de una función holomorfa a lo largo de una curva cerrada, dentro de una región sin agujeros, vale cero. En términos del capítulo 11 eso es decir que las funciones holomorfas son conservativas: la integral no depende del camino, solo de los extremos, y si vuelves al punto de partida no has acumulado nada.
La fórmula integral es la consecuencia asombrosa: el valor de la función en un punto interior queda determinado por sus valores en la curva que lo rodea. Es decir, *lo que pasa dentro está completamente dictado por lo que pasa en el borde*, aunque el borde esté lejos. Ninguna función real tiene una propiedad ni remotamente parecida; puedes modificar una función real en el interior de un intervalo sin tocar los extremos. La rigidez de las holomorfas es lo que hace que de aquí salgan, casi gratis, que sean infinitamente derivables, que sean analíticas, el teorema de Liouville y el teorema fundamental del álgebra.

La integral compleja $\oint_\gamma f\,\mathrm{d}z = \oint (u+iv)(\mathrm{d}x + i\,\mathrm{d}y)$ es un par de integrales de línea del capítulo 11.

Teorema 30.42 (Cauchy–Goursat)**.** *Si f es holomorfa en un dominio simplemente conexo, $\oint_\gamma f\,\mathrm{d}z = 0$ para toda curva cerrada.*

Demostración vía Green. Separando partes reales e imaginarias, las dos integrales de línea tienen, por Green, integrandos $-u_y - v_x$ y $u_x - v_y$: *exactamente las ecuaciones de Cauchy–Riemann*: ambos cero. (Goursat lo demostró sin suponer continuidad de las derivadas, por bisección de triángulos: el refinamiento técnico que hace a la teoría autosuficiente.) □

Idea clave

Qué dice esto en cristiano. Cauchy–Goursat dice que la integral de una función holomorfa a lo largo de cualquier curva cerrada vale cero, siempre que no haya agujeros

dentro. Es el resultado más asombroso del análisis complejo, porque en el caso real nada parecido es cierto: la integral de una función derivable a lo largo de un camino cerrado no tiene por qué anularse. La razón es que ser holomorfa es muchísimo más fuerte que ser derivable: obliga a cumplir las ecuaciones de Cauchy–Riemann, y esas ecuaciones dicen exactamente que el campo asociado es conservativo y sin fuentes.

Ejemplo 30.43 (Integrales que se anulan y la que no)**. (a) El caso general.** Para $f(z) = z^2$, holomorfa en todo $\mathbb{C}$, la integral sobre cualquier curva cerrada vale 0. Lo mismo para e^z, para $\sin z$ y para cualquier polinomio. No hace falta calcular nada.

(b) La excepción que lo explica todo. Sobre el círculo unidad recorrido en sentido antihorario, $\oint \frac{\mathrm{d}z}{z} = 2\pi\mathrm{i}$. Parametrizando con $z = \mathrm{e}^{\mathrm{i}\theta}$ queda $\int_0^{2\pi} \mathrm{i}\,\mathrm{d}\theta = 2\pi\mathrm{i}$. No hay contradicción: $\frac{1}{z}$ no es holomorfa en el origen, que está dentro de la curva.

(c) Las potencias superiores sí se anulan. Para $n \neq -1$, $\oint z^n\,\mathrm{d}z = 0$ sobre el círculo unidad, incluso con n negativo. Por ejemplo $\oint \frac{\mathrm{d}z}{z^2} = 0$. Solo el exponente -1 da algo distinto de cero, y por eso el residuo se define como el coeficiente de $\frac{1}{z-z_0}$ y de ningún otro término.

(d) La fórmula integral de Cauchy. De ahí sale que $f(z_0) = \frac{1}{2\pi\mathrm{i}} \oint \frac{f(z)}{z-z_0}\mathrm{d}z$: el valor en un punto interior queda determinado por los valores en el borde. Con $f = 1$ y $z_0 = 0$ se recupera el apartado (b). Esta fórmula no tiene ningún análogo real, y de ella se deduce que toda función holomorfa es infinitamente derivable y analítica.

Cuidado

[La hipótesis de simple conexión es la que se olvida] El apartado (b) es el contraejemplo: en el plano perforado, que no es simplemente conexo, la integral cerrada no se anula. Ese $2\pi\mathrm{i}$ es el mismo agujero que hacía fallar el lema de Poincaré y que da $\pi_1(S^1) = \mathbb{Z}$. Segundo aviso: holomorfa significa derivable en un *entorno* de cada punto, no solo en el punto; una función derivable en un solo punto no cumple nada de esto. Y tercero: el teorema exige holomorfía en todo el interior de la curva, no solo sobre ella; basta un polo dentro para que todo cambie.

Corolario inmediato: en dominios con agujeros vale la *deformación de contornos* del capítulo 11: la integral solo ve las singularidades encerradas.

Teorema 30.44 (Fórmula integral de Cauchy)**.** *Si f es holomorfa dentro y sobre γ (simple, antihoraria) y z_0 está dentro:*

$$f(z_0) = \frac{1}{2\pi i} \oint_\gamma \frac{f(z)}{z - z_0}\,\mathrm{d}z.$$

Demostración. Deformar γ a un círculo diminuto alrededor de z_0; ahí $f(z) \approx f(z_0)$ y $\oint \frac{\mathrm{d}z}{z-z_0} = 2\pi i$ (parametrizar: es el índice), con el error controlado por la continuidad. □

Ejemplo 30.45 (La fórmula integral de Cauchy en acción)**. (a) Un cálculo directo.**

$\oint_{|z|=2} \dfrac{e^z}{z-1}\,dz = 2\pi i\, e^1 = 2\pi i \cdot 2{,}718282$, porque 1 está dentro de la circunferencia. Si el polo estuviera fuera, por ejemplo en $z = 5$, la integral valdría 0.

(b) Derivadas. La versión para derivadas da $\oint_{|z|=2} \dfrac{e^z}{(z-1)^3}\,dz = \dfrac{2\pi i}{2!}e^1 = \pi i\, e$. De aquí se sigue que toda función holomorfa es infinitamente derivable, cosa que en el caso real es rotundamente falsa.

(c) La desigualdad de Cauchy. Acotando la integral se obtiene $|f^{(n)}(z_0)| \le n!\, M/R^n$. Con M fijo y $R \to \infty$ todas las derivadas de una entera acotada se anulan, luego es constante: eso es el teorema de Liouville, y de él sale el teorema fundamental del álgebra en tres líneas.

(d) La propiedad del valor medio. Tomando γ una circunferencia de centro z_0 y radio r, la fórmula dice que $f(z_0)$ es el promedio de f sobre esa circunferencia. De ahí sale el principio del módulo máximo, y de él, que una función holomorfa no constante nunca alcanza el máximo de $|f|$ en el interior.

Idea clave

La fórmula es asombrosa: *los valores de una holomorfa en el borde determinan todos los del interior* —compárese con la fórmula de representación de las armónicas del capítulo 11: es el mismo fenómeno, y no es casual—. Derivando bajo la integral (¡capítulo 6!) infinitas veces:

$$f^{(n)}(z_0) = \frac{n!}{2\pi i}\oint \frac{f(z)}{(z-z_0)^{n+1}}\,dz\,:$$

una holomorfa es automáticamente C^∞ y, acotando esta integral, igual a su serie de Taylor: **holomorfa = analítica**. El contraste con el caso real (e^{-1/x^2}, capítulo 6) no puede ser mayor: en $\mathbb{C}$, derivar una vez regala derivar infinitas.

Dos consecuencias con historia. *Liouville*: una entera y acotada es constante (la estimación de Cauchy $|f'(z_0)| \le \frac{\text{máx}\,|f|}{R}$ con $R \to \infty$); de ahí el *teorema fundamental del álgebra* en cuatro líneas: si p no tuviera raíces, $\frac{1}{p}$ sería entera y acotada, luego constante: absurdo —la tercera demostración del libro, tras Galois+Sylow: cada gran teoría reclama este teorema como suyo—. Y el *principio del módulo máximo*: $|f|$ no tiene máximos interiores (hermano del principio del máximo armónico del capítulo 11).

Regla: determinar el radio y el disco de convergencia de una serie de potencias

Una serie $\sum_{n\ge 0} a_n (z-z_0)^n$ converge absolutamente dentro de un disco, diverge fuera, y sobre el borde puede hacer cualquier cosa. Todo el trabajo es encontrar el radio R de ese disco.

1. **Identifica el centro z_0.** Es el punto donde la serie está escrita; todas las distancias se miden desde ahí. Si la serie viene en potencias de $z-1$, el centro es 1, no 0.
2. **Prueba primero el cociente.** Si todos los a_n son no nulos a partir de cierto punto, $R = \lim_{n\to\infty} \left|\dfrac{a_n}{a_{n+1}}\right|$, siempre que ese límite exista.

3. **Si hay huecos, usa Cauchy–Hadamard.** La fórmula que *siempre* funciona es $\dfrac{1}{R} = \limsup_{n\to\infty} |a_n|^{1/n}$, con los convenios $1/0 = \infty$ y $1/\infty = 0$. Es obligatoria cuando infinitos a_n valen cero.

4. **O reconoce una serie geométrica disfrazada.** Si la serie es $\sum c^n z^{kn} = \sum (cz^k)^n$, converge exactamente cuando $|cz^k| < 1$, es decir $|z| < |c|^{-1/k}$. Más rápido que cualquier fórmula.

5. **Atajo geométrico: R es la distancia a la singularidad más cercana.** Si f es holomorfa en un abierto y la desarrollas en z_0, entonces $R = \operatorname{dist}(z_0, S)$ donde S es el conjunto de puntos donde f deja de ser holomorfa. Esto convierte casi todos los ejercicios en un dibujo.

6. **Estudia el borde aparte.** En $|z - z_0| = R$ no hay teorema general: hay que analizar punto por punto, o usar el criterio de Dirichlet. Dentro la convergencia es absoluta y uniforme en compactos; en el borde puede ser condicional, absoluta o nula.

7. **Verifica.** Derivar o integrar término a término no cambia R. Y evalúa numéricamente en un punto interior cercano al borde: las sumas parciales deben estabilizarse, lentamente pero sin explotar.

Ejemplo 30.46 (Siete radios de convergencia)**.** (a) **Los dos extremos.** Para $\sum n!\, z^n$ el cociente da $|a_n/a_{n+1}| = 1/(n+1) \to 0$, luego $R = 0$: la serie solo converge en $z = 0$. Para $\sum z^n/n!$ el mismo cociente da $n + 1 \to \infty$, luego $R = \infty$: es e^z, entera. Entre estos dos extremos vive todo lo demás.

(b) **El disco unidad, con tres comportamientos distintos en el borde.** Las tres series

$$\sum_{n\geq 1} z^n, \qquad \sum_{n\geq 1} \frac{z^n}{n}, \qquad \sum_{n\geq 1} \frac{z^n}{n^2}$$

tienen las tres $R = 1$ (el cociente da 1 en los tres casos). Pero sobre $|z| = 1$: la primera diverge en *todos* los puntos (el término general no tiende a 0); la segunda diverge en $z = 1$ (serie armónica) y converge en todos los demás puntos del círculo, por ejemplo en $z = -1$ con suma $-\ln 2 \approx -0{,}693147$; la tercera converge absolutamente en *todo* el círculo, porque $\sum 1/n^2 = \pi^2/6$. Mismo radio, tres bordes.

(c) **Una serie con huecos.** Para $\sum_{n\geq 0} 2^n z^{3n}$ el criterio del cociente ni se puede aplicar: los coeficientes de z^1, z^2, z^4, z^5, … son cero. Cauchy–Hadamard: los coeficientes no nulos son $a_{3n} = 2^n$, luego $|a_{3n}|^{1/(3n)} = 2^{1/3}$ y $R = 2^{-1/3} \approx 0{,}793701$. El atajo del paso 4 lo confirma: $\sum (2z^3)^n$ converge si $|2z^3| < 1$, o sea $|z| < 2^{-1/3}$.

(d) **El misterio de $1/(1+x^2)$.** La función real $g(x) = \dfrac{1}{1+x^2}$ es infinitamente derivable en toda la recta, no se dispara en ningún sitio, y sin embargo su serie de Taylor en 0,

$$1 - x^2 + x^4 - x^6 + \cdots,$$

solo converge para $|x| < 1$: en $x = 2$ da $1 - 4 + 16 - 64 + \cdots$, un disparate. Desde la recta real no hay ninguna explicación. Desde $\mathbb{C}$ la hay, y es de una línea: $g(z) = 1/(1+z^2)$ tiene polos en $z = \pm i$, que están a distancia 1 del origen. Por el paso 5, $R = 1$. Los polos son invisibles sobre el eje real, pero el radio de convergencia los ve.

(e) **La misma función, otro centro.** Desarrollemos g alrededor de $z_0 = 1$. La singularidad más cercana sigue siendo $\pm i$, y ahora la distancia es $|1 - i| = \sqrt{2} \approx 1{,}414214$. Luego $R = \sqrt{2}$: cambiando el centro se gana alcance. Repitiendo esta idea —desarrollar, avanzar el centro, volver a desarrollar— se llega a la continuación analítica.

(f) **La tangente.** $\tan z$ es holomorfa en 0 y sus polos más cercanos son $\pm\pi/2$. Por tanto su serie de Taylor $z + \frac{z^3}{3} + \frac{2z^5}{15} + \cdots$ tiene $R = \pi/2 \approx 1{,}570796$, sin necesidad de mirar un solo coeficiente. Compáralo con el trabajo que costaría estimar $\limsup |a_n|^{1/n}$ a partir de los números de Bernoulli.

(g) **Un radio que no es una distancia obvia.** $h(z) = \dfrac{1}{e^z - 1} - \dfrac{1}{z}$ es holomorfa en 0 (la singularidad es evitable) y sus polos son $z = 2\pi i k$ con $k \neq 0$. El más cercano está a distancia $2\pi \approx 6{,}283185$, luego $R = 2\pi$. Es el radio que gobierna la serie de los números de Bernoulli.

Cuidado

El criterio del cociente *no* es equivalente a Cauchy–Hadamard. En (c) el cociente $|a_n/a_{n+1}|$ ni siquiera está definido para casi todo n. Y hay series sin huecos donde el cociente oscila y no existe el límite —por ejemplo $a_n = 2^n$ si n es par, $a_n = 3^n$ si es impar— mientras que $\limsup |a_n|^{1/n} = 3$ da $R = 1/3$ sin ninguna dificultad. Cuando dudes, usa el lím sup.

Pausa para pensar

Fíjate en lo que dice el apartado (d): un hecho sobre funciones de variable real (¿hasta dónde converge esta serie?) cuya respuesta correcta *no se puede formular* sin salir a $\mathbb{C}$. No es que el argumento complejo sea más elegante; es que en la recta el enunciado «hay una singularidad a distancia 1» no significa nada. Esta es la razón principal por la que el análisis complejo se estudia aunque uno solo quiera hacer análisis real.

Idea clave

El paso 5 tiene un recíproco fuerte: sobre la circunferencia $|z - z_0| = R$ hay *siempre* al menos un punto singular; si no lo hubiera, la función se extendería a un disco mayor y el radio sería más grande. A veces *todos* los puntos del borde son singulares: la serie lacunar $\sum_{n \geq 0} z^{2^n}$ tiene $R = 1$ y no admite continuación analítica a través de ningún arco de $|z| = 1$; ese círculo es una *frontera natural*. La función existe dentro y no existe fuera, y no hay nada que hacer al respecto.

30.4. Series de Laurent, singularidades y residuos

Idea clave

Intuición primero. Integra z^n en un círculo alrededor del origen: para todo $n \neq -1$ existe primitiva ($\frac{z^{n+1}}{n+1}$) y la integral cerrada es cero. El *único* término sin primitiva es $\frac{1}{z}$ —su primitiva querría ser $\ln z$, que no cierra el circuito: da $2\pi i$ de desfase por vuelta—. Así que al integrar una serie de Laurent término a término, el contorno es sordo a todo excepto al coeficiente de $\frac{1}{z}$: ese superviviente es el *residuo*. Todo el método de residuos es esta observación más contabilidad: una integral real endiablada se cierra por el plano complejo y su valor es $2\pi i$ por la suma de los pocos coeficientes que el contorno alcanza a oír.

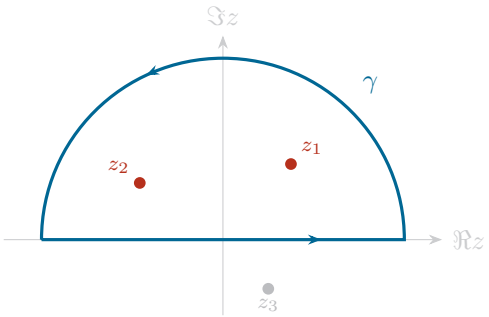


Figura 30.1: El contorno semicircular del teorema de los residuos: al hacer crecer el radio, el arco aporta cada vez menos y la integral sobre el eje real queda igualada a 2πi por la suma de los residuos de los polos encerrados. El polo z_3 no cuenta.

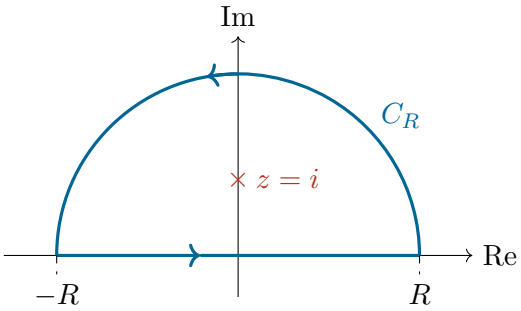


Figura 30.2: El contorno semicircular: la integral real deseada es el diámetro, el arco C_R muere cuando $R \to \infty$, y el valor queda en manos del único polo encerrado: $2\pi i$ $\mathrm{Res}_{z=i}$.

Alrededor de una singularidad aislada, f se expande en *serie de Laurent* $\sum_{n=-\infty}^{\infty} c_n (z-z_0)^n$; según la cola negativa, la singularidad es *evitable* (sin términos negativos), *polo de orden m* (finitos) o *esencial* (infinitos: cerca de ella f toma todos los valores salvo quizá uno —Picard—, como $e^{1/z}$).

Definición 30.47. El *residuo* es $\mathrm{Res}(f, z_0) = c_{-1}$: el único coeficiente que sobrevive a la integración ($\oint (z-z_0)^n \mathrm{d}z = 0$ salvo $n = -1$, que da $2\pi i$).

Teorema 30.48 (De los residuos)**.** *Para f holomorfa salvo singularidades aisladas dentro de γ:*

$$\oint_\gamma f\,\mathrm{d}z = 2\pi i \sum_{dentro} \operatorname{Res}(f, z_k).$$

Cálculo práctico: en un polo simple, Res $= \lim (z-z_0)f(z)$, y si $f = \frac{g}{h}$ con $h(z_0)=0$ simple: Res $= \frac{g(z_0)}{h'(z_0)}$; en polos de orden m, derivar $m-1$ veces.

Ejemplo 30.49 (Integrales reales por residuos)**.** *(a)* $\displaystyle\int_{-\infty}^{\infty} \frac{\mathrm{d}x}{(1+x^2)^2}$: cerramos con un semicírculo superior gigante (su aporte muere como $\frac{1}{R^3}\cdot R$). Dentro: polo doble en $z=i$ de $\frac{1}{(z-i)^2(z+i)^2}$; residuo $= \frac{\mathrm{d}}{\mathrm{d}z}(z+i)^{-2}\big|_{z=i} = -\frac{2}{(2i)^3} = \frac{1}{4i}$; total $2\pi i\cdot\frac{1}{4i} = \frac{\pi}{2}$. *(b)* $\displaystyle\int_{-\infty}^{\infty} \frac{\cos ax}{1+x^2}\mathrm{d}x = \operatorname{Re}\oint \frac{e^{iaz}}{1+z^2}$: polo en i con residuo $\frac{e^{-a}}{2i}$: resultado πe^{-a} —¡nuestra transformada de la gaussiana del capítulo 6 tiene ahora una prima exacta, y la integral de Dirichlet del capítulo 9 se hace igual con un rodeo por el origen!—. *(c)* Sumas de series: la función $\pi\cot\pi z$ tiene polos en todos los enteros con residuo 1; integrándola contra $\frac{1}{z^2}$ en cuadrados crecientes:

$$\sum_{n\geq 1}\frac{1}{n^2} = \frac{\pi^2}{6},$$

la *cuarta* demostración de Basilea del libro (Parseval, producto de Euler, y ahora residuos): cuando un resultado cae por tantos caminos, es que está en el centro de la red.

Regla: clasificar una singularidad y calcular su residuo

Casi todo el uso práctico del análisis complejo pasa por estos dos cálculos encadenados.

1. **Localiza las singularidades.** Para un cociente $f = g/h$ con g, h holomorfas, están donde h se anula. Comprueba antes si g se anula también ahí: si lo hace, quizá la singularidad se cancele y sea evitable.

2. **Determina el orden del polo.** Es el orden del cero de h menos el orden del cero de g en ese punto. En la práctica: multiplica f por $(z-z_0)^m$ con $m = 1, 2, 3, \ldots$ hasta que el límite en z_0 salga finito y no nulo; ese m es el orden.

3. **Polo simple ($m=1$):** $\operatorname{Res}(f, z_0) = \lim_{z\to z_0}(z-z_0)f(z)$. Y si $f = g/h$ con h con cero simple, hay un atajo muy cómodo:

$$\operatorname{Res}(f, z_0) = \frac{g(z_0)}{h'(z_0)},$$

   que evita factorizar h y es el que conviene usar cuando h es un polinomio de grado alto o una exponencial.

4. **Polo de orden m:** deriva $m-1$ veces,

$$\operatorname{Res}(f, z_0) = \frac{1}{(m-1)!}\lim_{z\to z_0}\frac{\mathrm{d}^{m-1}}{\mathrm{d}z^{m-1}}\Big[(z-z_0)^m f(z)\Big].$$

Cuidado con el factorial y con derivar *después* de multiplicar, no antes.

5. **Singularidad esencial:** no hay fórmula. Hay que desarrollar la serie de Laurent y leer directamente el coeficiente de $(z-z_0)^{-1}$. Con exponenciales y senos esto suele ser fácil, porque sus series se conocen de memoria.

6. **Control.** Si f es un cociente de polinomios y el grado del denominador supera al del numerador en dos o más unidades, la suma de *todos* los residuos es cero. Es una verificación excelente y muy barata.

Ejemplo 30.50 (Cuatro residuos de cuatro tipos)**.** *(a) Polo simple con la fórmula del cociente.* Sea $f(z)=\dfrac{1}{z^2+1}$. El denominador se anula en $\pm i$ con ceros simples. Con $g=1$ y $h=z^2+1$, $h'=2z$:

$$\operatorname{Res}(f,i)=\frac{1}{2i}=-\frac{i}{2},\qquad \operatorname{Res}(f,-i)=\frac{1}{-2i}=\frac{i}{2}.$$

La suma es cero, como anticipaba el control del paso 6 (grado 2 frente a grado 0).

(b) Polo doble. Sea $f(z)=\dfrac{z+1}{z^2(z-2)}$. En $z=0$ el denominador tiene un cero de orden 2 y el numerador no se anula, así que es un polo doble. Aplicamos la fórmula con $m=2$:

$$\operatorname{Res}(f,0)=\frac{1}{1!}\lim_{z\to 0}\frac{\mathrm{d}}{\mathrm{d}z}\Big[\frac{z+1}{z-2}\Big]=\lim_{z\to 0}\frac{(z-2)-(z+1)}{(z-2)^2}=\frac{-3}{4}.$$

En $z=2$ el polo es simple: $\operatorname{Res}(f,2)=\dfrac{2+1}{2^2}=\dfrac{3}{4}$. Otra vez la suma es cero (grados 3 y 1), lo que confirma los dos cálculos a la vez.

Consecuencia inmediata: $\displaystyle\oint_{|z|=3}\frac{z+1}{z^2(z-2)}\,\mathrm{d}z=2\pi i\Big(-\frac{3}{4}+\frac{3}{4}\Big)=0$, mientras que $\displaystyle\oint_{|z|=1}$ del mismo integrando vale $2\pi i\cdot\big(-\frac{3}{4}\big)=-\frac{3\pi i}{2}$, porque ese contorno más pequeño solo encierra el polo del origen. *El contorno decide qué residuos cuentan*: esa es toda la técnica.

(c) Singularidad esencial. Para $f(z)=\mathrm{e}^{1/z}$ usamos la serie de la exponencial con $1/z$ en el argumento:

$$\mathrm{e}^{1/z}=\sum_{n\ge 0}\frac{1}{n!\,z^n}=1+\frac{1}{z}+\frac{1}{2z^2}+\frac{1}{6z^3}+\cdots$$

El coeficiente de $1/z$ es 1, luego Res $=1$. Si en cambio tomamos $f(z)=z^2\mathrm{e}^{1/z}$, el término genérico es $z^{2-n}/n!$ y hay que resolver $2-n=-1$, es decir $n=3$: el residuo es $1/3!=1/6$.

(d) Singularidad evitable. $f(z)=\dfrac{\sin z}{z}$ parece singular en 0, pero $\sin z=z-\frac{z^3}{6}+\cdots$, de modo que $f=1-\frac{z^2}{6}+\cdots$: no hay ningún término negativo. La singularidad es evitable, el residuo es 0, y toda integral cerrada alrededor del origen se anula.

Cuidado

Una misma función tiene *varias* series de Laurent, una por cada corona donde se la desarrolle, y solo la de la corona más interior ($0<|z-z_0|<r$) da el residuo en z_0. Por

ejemplo, $f(z) = \dfrac{1}{z(z-1)}$ admite en $0 < |z| < 1$ el desarrollo

$$\frac{1}{z(z-1)} = -\frac{1}{z}\cdot\frac{1}{1-z} = -\frac{1}{z}(1+z+z^2+\cdots) = -\frac{1}{z} - 1 - z - z^2 - \cdots,$$

con residuo -1 en el origen. Pero en la región exterior $|z| > 1$ el mismo cociente se desarrolla como $\dfrac{1}{z^2}\cdot\dfrac{1}{1-1/z} = \dfrac{1}{z^2} + \dfrac{1}{z^3} + \cdots$, cuyo coeficiente de $1/z$ es 0. No hay contradicción: ese segundo desarrollo no es válido cerca del origen y no calcula ningún residuo. Antes de leer un coeficiente, hay que saber en qué corona se está.

Regla: elegir el contorno para una integral real

El método de residuos resuelve integrales reales que no tienen primitiva elemental, pero exige escoger bien el contorno. Hay tres plantillas que cubren casi todos los casos de un curso.

1. **Función racional en $(-\infty,\infty)$, con denominador sin raíces reales y grado al menos dos unidades mayor que el numerador:** usa el *semicírculo superior*. El arco tiende a cero porque su longitud crece como R mientras el integrando decrece como R^{-2} o más rápido, y el valor queda en manos de los polos del semiplano superior.

2. **Integrando con $\cos ax$ o $\sin ax$ (con $a > 0$):** nunca lo pases al plano complejo tal cual, porque $\cos az$ explota en el semiplano superior. Sustitúyelo por e^{iaz}, integra, y toma al final la parte real (para el coseno) o la imaginaria (para el seno). El lema de Jordan garantiza entonces que el arco muere.

3. **Integrando trigonométrico sobre $[0, 2\pi]$:** sustituye $z = \mathrm{e}^{i\theta}$, de modo que

$$\cos\theta = \frac{z + z^{-1}}{2}, \qquad \sin\theta = \frac{z - z^{-1}}{2i}, \qquad \mathrm{d}\theta = \frac{\mathrm{d}z}{iz},$$

   y la integral se convierte en una integral sobre la circunferencia unitaria: solo cuentan los polos con módulo menor que 1.

4. **Si hay un polo sobre el eje real**, hay que esquivarlo con un semicírculo pequeño; ese rodeo aporta $\pm i\pi$ veces el residuo (la mitad de una vuelta completa), no $2\pi i$.

5. **Verificación final.** Si la integral de partida era real, el resultado debe salir real; si el integrando era positivo, el resultado debe salir positivo. Cualquier i superviviente delata un error de signo o un polo mal contado.

Ejemplo 30.51 (Una racional: $\int_{-\infty}^{\infty} \frac{\mathrm{d}x}{x^4+1}$)**.** El denominador no se anula sobre el eje real y su grado supera al del numerador en cuatro unidades: la plantilla del semicírculo superior es aplicable.

Los polos son las cuatro raíces de $z^4 = -1$, es decir $\mathrm{e}^{i\pi/4}$, $\mathrm{e}^{3i\pi/4}$, $\mathrm{e}^{5i\pi/4}$ y $\mathrm{e}^{7i\pi/4}$. En el

semiplano superior quedan las dos primeras. Como los ceros son simples y el numerador es 1, usamos el atajo del cociente con $h' = 4z^3$:

$$\operatorname{Res} = \frac{1}{4z^3} = \frac{z}{4z^4} = \frac{z}{4(-1)} = -\frac{z}{4},$$

donde hemos usado $z^4 = -1$ para simplificar. Sumando los dos polos relevantes,

$$-\frac{1}{4}\left(\mathrm{e}^{i\pi/4} + \mathrm{e}^{3i\pi/4}\right) = -\frac{1}{4}\left(\frac{\sqrt{2}}{2} + i\frac{\sqrt{2}}{2} - \frac{\sqrt{2}}{2} + i\frac{\sqrt{2}}{2}\right) = -\frac{i\sqrt{2}}{4}.$$

Por el teorema de los residuos,

$$\int_{-\infty}^{\infty} \frac{\mathrm{d}x}{x^4+1} = 2\pi i \cdot \Big(-\frac{i\sqrt{2}}{4}\Big) = \frac{2\pi\sqrt{2}}{4} = \frac{\pi}{\sqrt{2}} \approx 2{,}221.$$

El resultado es real y positivo, como exige un integrando positivo. ✓ Y el orden de magnitud es creíble: el integrando vale 1 en el origen y decae rápido, así que un área algo mayor que 2 es lo esperable.

Ejemplo 30.52 (Una trigonométrica: $\int_0^{2\pi} \frac{\mathrm{d}\theta}{2+\cos\theta}$)**.** Aplicamos la tercera plantilla. Con $z = \mathrm{e}^{i\theta}$ y $\cos\theta = \frac{1}{2}(z + z^{-1})$:

$$\int_0^{2\pi} \frac{\mathrm{d}\theta}{2+\cos\theta} = \oint_{|z|=1} \frac{1}{2 + \frac{z+z^{-1}}{2}} \cdot \frac{\mathrm{d}z}{iz} = \oint_{|z|=1} \frac{2\,\mathrm{d}z}{i\,(z^2+4z+1)}.$$

Las raíces de z^2+4z+1 son $z = -2 \pm \sqrt{3}$. Solo $z_0 = -2+\sqrt{3} \approx -0{,}268$ tiene módulo menor que 1 y está dentro del contorno; la otra, $\approx -3{,}73$, queda fuera y no cuenta.

El residuo, con la fórmula del cociente ($h = z^2+4z+1$, $h' = 2z+4$):

$$\operatorname{Res} = \frac{2/i}{2z_0+4} = \frac{2/i}{2\sqrt{3}} = \frac{1}{i\sqrt{3}}.$$

Y por tanto

$$\int_0^{2\pi} \frac{\mathrm{d}\theta}{2+\cos\theta} = 2\pi i \cdot \frac{1}{i\sqrt{3}} = \frac{2\pi}{\sqrt{3}} \approx 3{,}628.$$

Comprobación de sensatez. El integrando oscila entre $\frac{1}{3}$ (cuando $\cos\theta = 1$) y 1 (cuando $\cos\theta = -1$), así que la integral sobre un intervalo de longitud $2\pi \approx 6{,}28$ tiene que estar entre 2,09 y 6,28. El valor 3,63 encaja. Nótese además que el resultado no depende de haber sabido ninguna primitiva: la sustitución de Weierstrass daría lo mismo, pero con bastante más trabajo.

Regla: Cómo desarrollar una función en serie de Laurent en cada anillo

Una misma función racional no tiene *una* serie: tiene una por cada anillo donde es holomorfa. El procedimiento es siempre el mismo.

1. **Localiza las singularidades.** Si f tiene polos en $z_1, \dots, z_k$, ordena sus módulos $|z_1| \le \cdots \le |z_k|$. Estos números parten el plano en $k+1$ anillos centrados en

0: $|z| < |z_1|$, luego $|z_1| < |z| < |z_2|$, y así hasta $|z| > |z_k|$. En cada uno hay un desarrollo distinto y en cada uno es único.

2. **Descompón en fracciones simples.** Escribe f como suma de términos $\dfrac{A}{z-a}$ (y potencias superiores si hay polos múltiples). Cada término se desarrolla por separado y luego se suman las series.

3. **Elige la dirección de cada término.** Para el término $\dfrac{1}{z-a}$ hay dos maneras de forzar una serie geométrica, y la elección depende de si estás dentro o fuera del círculo $|z| = |a|$:

$$\underbrace{\frac{1}{z-a} = -\frac{1}{a}\cdot\frac{1}{1-z/a} = -\sum_{n\geq 0}\frac{z^n}{a^{n+1}}}_{\text{válido si } |z|<|a|}, \qquad \underbrace{\frac{1}{z-a} = \frac{1}{z}\cdot\frac{1}{1-a/z} = \sum_{n\geq 0}\frac{a^n}{z^{n+1}}}_{\text{válido si } |z|>|a|}.$$

   La regla mnemotécnica: *saca factor común el que tenga mayor módulo*, para que el cociente que queda sea menor que 1.

4. **Aplica la serie geométrica.** Usa $\dfrac{1}{1-w} = \sum_{n\geq 0} w^n$ solo cuando $|w| < 1$. Todo el trabajo del paso anterior fue precisamente construir un w con $|w| < 1$ en el anillo que te interesa.

5. **Suma y reordena.** Junta las dos series en una sola escrita como $\sum_{n=-\infty}^{\infty} c_n z^n$. Las potencias negativas forman la *parte principal*; las no negativas, la *parte regular*.

6. **Lee el residuo en el anillo correcto.** El residuo en $z = 0$ es c_{-1} del desarrollo válido en el anillo *más interior* (el que rodea al origen sin encerrar otros polos). En los anillos exteriores, c_{-1} ya no es un residuo: es la suma de todos los residuos encerrados.

7. **Comprueba numéricamente.** Elige un punto del anillo, suma unos cuantos términos y compáralo con el valor exacto de f. Si no coinciden, casi siempre te equivocaste de dirección en el paso 3.

Ejemplo 30.53 (Una función, tres series distintas)**.** Desarrollemos $f(z) = \dfrac{1}{(z-1)(z-2)}$ en todos sus anillos. Los polos están en 1 y 2, así que hay tres regiones: $|z| < 1$, $1 < |z| < 2$ y $|z| > 2$.

Primero, fracciones simples: buscamos A, B con $\frac{A}{z-1} + \frac{B}{z-2} = f$. Multiplicando y evaluando en $z = 1$ da $A = \frac{1}{1-2} = -1$; en $z = 2$ da $B = \frac{1}{2-1} = 1$. Por tanto

$$f(z) = \frac{1}{z-2} - \frac{1}{z-1}.$$

(a) El disco $|z| < 1$. Aquí $|z| < 1$ y $|z| < 2$, así que en ambos términos el polo tiene mayor

módulo y sacamos factor común el polo:

$$\frac{1}{z-2} = -\sum_{n\geq 0} \frac{z^n}{2^{n+1}}, \qquad -\frac{1}{z-1} = +\sum_{n\geq 0} z^n.$$

Sumando:

$$f(z) = \sum_{n\geq 0} \left(1 - \frac{1}{2^{n+1}}\right) z^n = \frac{1}{2} + \frac{3}{4}z + \frac{7}{8}z^2 + \frac{15}{16}z^3 + \cdots$$

No hay potencias negativas, como debe ser: f es holomorfa en el disco. Control numérico en $z = \frac{1}{2}$: la serie da 1,33333 y $f(\frac{1}{2}) = \frac{1}{(-\frac{1}{2})(-\frac{3}{2})} = \frac{4}{3}$. Coincide.

(b) El anillo $1 < |z| < 2$. Ahora $|z| > 1$ pero $|z| < 2$, así que los dos términos van en direcciones opuestas:

$$-\frac{1}{z-1} = -\sum_{n\geq 0} \frac{1}{z^{n+1}} = -\sum_{n\geq 1} z^{-n}, \qquad \frac{1}{z-2} = -\sum_{n\geq 0} \frac{z^n}{2^{n+1}}.$$

Es decir

$$f(z) = -\sum_{n\geq 1} \frac{1}{z^n} - \sum_{n\geq 0} \frac{z^n}{2^{n+1}} = \cdots - \frac{1}{z^2} - \frac{1}{z} - \frac{1}{2} - \frac{z}{4} - \frac{z^2}{8} - \cdots$$

Aquí sí hay parte principal, y es infinita. Control en $z = \frac{3}{2}$: sumando doscientos términos de cada serie se obtiene $-4{,}000$, y en efecto $f(\frac{3}{2}) = \frac{1}{(\frac{1}{2})(-\frac{1}{2})} = -4$.

(c) El exterior $|z| > 2$. Ambos polos quedan dentro, así que en los dos términos sacamos factor común z:

$$\frac{1}{z-2} = \sum_{n\geq 0} \frac{2^n}{z^{n+1}}, \qquad -\frac{1}{z-1} = -\sum_{n\geq 0} \frac{1}{z^{n+1}},$$

$$f(z) = \sum_{n\geq 0} \frac{2^n - 1}{z^{n+1}} = \frac{0}{z} + \frac{1}{z^2} + \frac{3}{z^3} + \frac{7}{z^4} + \cdots$$

Control en $z = 3$: la suma da 0,5 y $f(3) = \frac{1}{2\cdot 1} = \frac{1}{2}$. Coincide.

(d) Lectura de los coeficientes. Obsérvese el coeficiente de z^{-1} en cada caso: vale 0 en (a), -1 en (b) y 0 en (c). Ninguno de los tres es el residuo de f en un polo, porque 0 no es polo de f. Lo que sí miden es la integral $\frac{1}{2\pi i}\oint f$ sobre una circunferencia del anillo correspondiente, es decir, la suma de residuos encerrados. En (a) no se encierra nada: 0. En (b) se encierra solo $z = 1$, cuyo residuo es -1. En (c) se encierran ambos, y $\operatorname{Res}_{z=1} f + \operatorname{Res}_{z=2} f = -1 + 1 = 0$. Las tres series conocen la geometría.

Cuidado

Es un error clásico decir «la serie de Laurent de f». No existe: existe la serie de Laurent de f *en un anillo dado*. Las tres series de arriba representan la misma función analítica, pero ninguna converge fuera de su anillo, y en la intersección vacía de sus regiones no hay nada que comparar. Antes de escribir un desarrollo, di siempre dónde vive.

Pausa para pensar

En el desarrollo (c) todos los coeficientes de potencias positivas son cero y el de z^{-1} también. ¿Qué te dice eso sobre el comportamiento de f cuando $z \to \infty$? Respuesta: $f(z) \sim z^{-2}$, o sea f tiene un cero de orden 2 en el infinito, que es exactamente lo que se espera de una racional con grado del denominador dos unidades mayor que el del numerador.

Regla: elegir una rama del logaritmo (o de la raíz) y trabajar con el corte

En $\mathbb{C}$ el logaritmo no es una función: es una familia infinita de funciones. Trabajar bien con él consiste en elegir explícitamente cuál, y no olvidarse de la elección a mitad del cálculo.

1. **Escribe z en forma polar.** $z = re^{i\theta}$ con $r = |z| > 0$. Entonces cualquier logaritmo de z es $\log z = \ln r + i\theta$, y como θ está determinado solo salvo múltiplos de 2π, hay infinitos valores separados por $2\pi i$.

2. **Elige un intervalo para el argumento.** Fijar una *rama* es escoger un intervalo semiabierto de longitud 2π donde vive θ. La *rama principal* usa $\arg z \in (-\pi, \pi]$ y se escribe $\operatorname{Log} z$.

3. **Identifica el corte.** El rayo desde el origen que corresponde al extremo del intervalo queda excluido: para la rama principal es el semieje real negativo, $(-\infty, 0]$. Sobre el corte la rama salta $2\pi i$; el origen es un punto de ramificación y nunca puede estar dentro del dominio.

4. **Define las potencias a partir del logaritmo.** $z^a := e^{a \log z}$. Cambiar de rama cambia z^a en un factor $e^{2\pi i a}$; solo cuando $a \in \mathbb{Z}$ ese factor es 1 y la potencia es unívoca. Para $a = 1/n$ hay exactamente n valores: las n raíces.

5. **No uses las identidades reales sin comprobarlas.** $\log(zw) = \log z + \log w$ y $\sqrt{z}\sqrt{w} = \sqrt{zw}$ son válidas *módulo* $2\pi i$ (respectivamente, salvo signo), no como igualdades de ramas fijas.

6. **Dentro del dominio, deriva con tranquilidad.** En el abierto donde la rama es continua, es holomorfa y $\dfrac{\mathrm{d}}{\mathrm{d}z} \log z = \dfrac{1}{z}$. Todo el análisis funciona; lo único prohibido es cruzar el corte.

7. **Verifica dando una vuelta.** Sigue el valor de log continuamente a lo largo de una curva cerrada que rodee el origen: si al volver al punto de partida el valor cambió, la elección de rama era necesaria y acabas de medir la monodromía.

Ejemplo 30.54 (Seis logaritmos y dos identidades que fallan)**.** (a) **Valores principales.** Con $\arg \in (-\pi, \pi]$: $\operatorname{Log}(-1) = \ln 1 + i\pi = i\pi$; $\operatorname{Log}(i) = i\pi/2$; $\operatorname{Log}(1-i) = \ln\sqrt{2} - i\pi/4 = \frac{1}{2}\ln 2 - i\pi/4 \approx 0{,}346574 - 0{,}785398\,i$. Los demás logaritmos de cada uno de estos números se obtienen sumando $2\pi i k$.

(b) i^i **es real.** Por el paso 4, $i^i = e^{i \log i} = e^{i(i\pi/2 + 2\pi i k)} = e^{-\pi/2 - 2\pi k}$. Todos los valores son *reales positivos*. El principal es

$$i^i = e^{-\pi/2} \approx 0{,}207880,$$

y los demás forman la sucesión $\ldots, e^{-\pi/2+2\pi}, e^{-\pi/2}, e^{-\pi/2-2\pi}, \ldots$ Que una potencia imaginaria de un número imaginario salga real no es una casualidad: es lo que dice $e^{i\theta}$ cuando el exponente vuelve a girar.

(c) **La suma de logaritmos falla.** Sea $z = w = e^{2\pi i/3}$. Entonces $\operatorname{Log} z + \operatorname{Log} w = \frac{2\pi i}{3} + \frac{2\pi i}{3} = \frac{4\pi i}{3}$. Pero $zw = e^{4\pi i/3} = e^{-2\pi i/3}$, cuyo argumento principal es $-2\pi/3$, así que $\operatorname{Log}(zw) = -\frac{2\pi i}{3}$. La diferencia es exactamente $2\pi i$: la suma se salió del intervalo $(-\pi, \pi]$ y la rama principal la devolvió al redil.

(d) **La raíz del producto falla.** Con la rama principal, $\sqrt{-1} = e^{\frac{1}{2}\operatorname{Log}(-1)} = e^{i\pi/2} = i$. Entonces

$$\sqrt{-1} \cdot \sqrt{-1} = i \cdot i = -1, \qquad \text{pero} \qquad \sqrt{(-1)(-1)} = \sqrt{1} = 1.$$

Esta es la fuente de la «demostración» de que $1 = -1$ que circula por internet. El error no está en el álgebra: está en suponer que $\sqrt{\ }$ es una función multiplicativa, cuando es una rama.

(e) **Una vuelta al origen.** Empieza en $z = 1$ con $\log 1 = 0$ y recorre el círculo unidad en sentido positivo, siguiendo el valor de forma continua: $\log(e^{i\theta}) = i\theta$ con θ creciendo de 0 a 2π. Al volver a $z = 1$ el valor es $2\pi i$, no 0. La misma cuenta escrita como integral es

$$\oint_{|z|=1} \frac{\mathrm{d}z}{z} = 2\pi i,$$

o sea: la integral de Cauchy más básica *es* el salto de rama del logaritmo. Si en cambio recorres un círculo que no rodee al origen, la vuelta cierra y la integral da 0.

(f) **Ramas distintas para un mismo problema.** Para integrar $\int_0^\infty \frac{x^{a-1}}{1+x}\,\mathrm{d}x$ con $0 < a < 1$ no sirve la rama principal: el contorno estándar (el «ojo de cerradura») necesita $\arg z \in (0, 2\pi)$, con el corte sobre el semieje real *positivo*, para que los dos bordes del corte den valores que difieran en el factor $e^{2\pi i(a-1)}$. La elección de rama no es cosmética: es lo que hace que el contorno funcione, y el resultado es $\pi / \sin(\pi a)$.

Cuidado

$(z^a)^b \neq z^{ab}$ en general, por la misma razón. Con la rama principal, $((-1)^2)^{1/2} = 1^{1/2} = 1$, mientras que $(-1)^{2\cdot\frac{1}{2}} = (-1)^1 = -1$. Cada vez que aparece un exponente no entero hay un logaritmo escondido, y por tanto una rama que alguien tuvo que elegir.

Pausa para pensar

El corte no es una propiedad del logaritmo: es una propiedad de la *rama*. Puedes poner el corte en el semieje negativo, en el positivo, o en una espiral: lo único obligatorio es que sea una curva que vaya del origen a infinito, para impedir que ninguna curva cerrada

del dominio rodee al punto de ramificación. La función no tiene ninguna singularidad en, digamos, $z = -1$; somos nosotros los que hemos decidido cortar ahí.

Idea clave

La forma limpia de decir todo esto es que el logaritmo es una función unívoca sobre otra superficie: la *superficie de Riemann* del logaritmo, una espiral infinita de hojas donde cada vuelta alrededor del origen te sube un piso y el valor aumenta en $2\pi i$. La de $\sqrt{z}$ tiene solo dos hojas y se cierra tras dos vueltas; la de $z^{1/n}$, n hojas. Los cortes desaparecen: eran el precio de aplastar la superficie sobre el plano. Es el mismo cambio de punto de vista que convierte a los espacios recubridores en la herramienta natural, capítulo 28.

30.5. Mapas conformes y el teorema de Riemann

Idea clave

Intuición primero. Una función holomorfa con derivada no nula, vista como transformación del plano en el plano, *conserva los ángulos*: dos curvas que se cortan formando cierto ángulo van a parar a dos curvas que se cortan formando el mismo ángulo. Deforma las longitudes, deforma las áreas, pero no los ángulos. La razón es la lectura geométrica de la derivada compleja: multiplicar por un número complejo es rotar y escalar, y a escala infinitesimal eso es exactamente lo que hace la función, la misma rotación y el mismo escalado en todas las direcciones.
De ahí sale su utilidad práctica: si sabes resolver un problema de potencial, de temperatura estacionaria o de flujo en una región sencilla (un disco, un semiplano), puedes transportar la solución a una región complicada mediante un mapa conforme, porque las ecuaciones implicadas son invariantes bajo esas transformaciones. El teorema del mapeo de Riemann dice cuán general es esto, y es difícil de creer la primera vez: *cualquier* región simplemente conexa del plano que no sea todo el plano es conformemente equivalente al disco unitario. Da igual lo retorcida que sea su frontera.

Como las holomorfas preservan ángulos, transforman problemas de potencial entre dominios: resolver Laplace en un dominio feo = mapearlo conformemente a un disco y resolver ahí (donde la serie del capítulo 12 lo hace). El *teorema del mapa de Riemann* garantiza que **todo** dominio simplemente conexo propio de $\mathbb{C}$ se mapea conformemente al disco unitario: uno de los teoremas de existencia más potentes de la matemática. Ejemplos operativos: $z \mapsto z^2$ abre el primer cuadrante en el semiplano; la transformación de Yukovski $z + \frac{1}{z}$ manda círculos a perfiles de ala (y con ella y el teorema de Kutta del capítulo 11 se calculó la sustentación de los primeros aviones); las transformaciones de Möbius $\frac{az+b}{cz+d}$ (¡las isometrías de $\mathbb{H}$ del capítulo 36!) permutan discos y semiplanos. La cartografía conforme (Mercator preserva ángulos: por eso sirve para navegar) es el caso esférico.

Regla: Cómo construir un mapa conforme entre dominios estándar

El teorema de Riemann garantiza que existe un biholomorfismo entre casi cualquier dominio simplemente conexo y el disco, pero no lo construye. En la práctica se arma por composición, encadenando piezas de un catálogo corto.

1. **Identifica origen y destino.** Escribe explícitamente el dominio Ω de partida y el de llegada. Si el destino no es el disco unidad $\mathbb{D} = \{|w| < 1\}$, casi siempre conviene pasar por él: construye $\Omega \to \mathbb{D}$ y luego $\mathbb{D} \to$ destino, componiendo la inversa de un mapa conocido.

2. **Reduce a un semiplano o a un sector.** Los dominios se abaratan en este orden: sector $\to$ semiplano $\to$ disco. Casi todo ejercicio se resuelve llevando Ω primero al semiplano superior $\mathbb{H} = \{\operatorname{Im} z > 0\}$.

3. **Usa potencias para abrir o cerrar ángulos.** El mapa $z \mapsto z^\alpha$ (con la rama principal) multiplica por α el ángulo en el vértice. Un sector de abertura θ con vértice en 0 se convierte en un semiplano tomando $\alpha = \pi/\theta$. Ojo: z^α es conforme salvo en el vértice mismo, donde el ángulo cambia y por tanto la conformidad se pierde.

4. **Usa la exponencial para bandas.** La franja horizontal $\{0 < \operatorname{Im} z < \pi\}$ se transforma por $z \mapsto e^z$ exactamente en $\mathbb{H}$: en efecto e^{x+iy} tiene argumento y, que recorre $(0, \pi)$. Para una franja de altura h se compone antes con $z \mapsto \pi z/h$.

5. **Usa la transformación de Cayley para pasar de semiplano a disco.** El mapa
$$w = \frac{z - i}{z + i}$$
lleva $\mathbb{H}$ biyectivamente sobre $\mathbb{D}$, con $i \mapsto 0$ y el eje real sobre la circunferencia unidad. Su inversa es $z = i\dfrac{1 + w}{1 - w}$.

6. **Ajusta con automorfismos del disco.** Si necesitas que un punto concreto $a \in \mathbb{D}$ vaya al centro, compón con
$$\varphi_a(z) = e^{i\theta}\,\frac{z - a}{1 - \overline{a}z},$$
que es biholomorfismo de $\mathbb{D}$ en sí mismo. Todo automorfismo del disco tiene esta forma, así que dos condiciones (un punto y una dirección) fijan el mapa por completo.

7. **Verifica en tres puntos frontera.** Un mapa conforme queda determinado por su acción sobre tres puntos del borde. Evalúa tu composición en tres puntos concretos y comprueba que caen donde deben y en el orden correcto: si el orden se invierte, has intercambiado el interior por el exterior.

Ejemplo 30.55 (Cinco mapas del catálogo, verificados punto por punto)**. (a) Semiplano**

superior al disco: la transformación de Cayley. Sea $T(z) = \dfrac{z-i}{z+i}$. Si $z = x$ es real, $|x-i| = |x+i| = \sqrt{x^2+1}$, luego $|T(x)| = 1$: el eje real va a la circunferencia. Y $T(i) = 0$, el centro. Comprobaciones numéricas: $T(2i) = \frac{i}{3i} = \frac{1}{3}$, de módulo 0,3333; $T(1+i) = \frac{1}{1+2i} = \frac{1-2i}{5} = 0{,}2 - 0{,}4i$, de módulo 0,4472; $T(-3+\frac{1}{2}i)$ da $0{,}7333 + 0{,}5333i$, de módulo 0,9068. Los tres puntos están en $\mathbb{H}$ y sus imágenes en $\mathbb{D}$, cada vez más cerca del borde cuanto más cerca del eje real está el original.

(b) Automorfismos del disco. Tomemos $a = \frac{1}{2}$ y $\varphi(z) = \dfrac{z-1/2}{1-z/2}$. Entonces $\varphi(\frac{1}{2}) = 0$ y $\varphi(1) = \frac{1/2}{1/2} = 1$. Numéricamente: $\varphi(0{,}9) = 0{,}7273$; $\varphi(0{,}99i)$ tiene módulo 0,99399; $\varphi(\frac{1}{2}+\frac{1}{2}i) = -0{,}2+0{,}6i$, de módulo 0,6325. Los puntos interiores siguen siendo interiores y los muy cercanos al borde siguen muy cercanos al borde, como debe ocurrir con un automorfismo.

(c) Cuadrante a semiplano: la potencia. El primer cuadrante $\{0 < \arg z < \pi/2\}$ es un sector de abertura $\theta = \pi/2$. Con $\alpha = \pi/\theta = 2$, el mapa $z \mapsto z^2$ duplica argumentos y lo lleva sobre $\{0 < \arg w < \pi\} = \mathbb{H}$. Componiendo con (a) obtenemos el mapa del cuadrante al disco:

$$w = \frac{z^2 - i}{z^2 + i}.$$

Comprobación: $z = 1+i$ está en el cuadrante, $z^2 = 2i$, y ya vimos que $T(2i) = \frac{1}{3} \in \mathbb{D}$.

(d) Franja a semiplano: la exponencial. En $\Omega = \{0 < \operatorname{Im} z < \pi\}$ tomemos $E(z) = e^z$. Como $\operatorname{Im}(e^{x+iy}) = e^x \sin y$ y $\sin y > 0$ para $0 < y < \pi$, la imagen cae en $\mathbb{H}$. Verificación: $E(\frac{1}{2}i) = 0{,}87758 + 0{,}47943i$, con parte imaginaria positiva; $E(1+i) = 1{,}46869 + 2{,}28736i$; $E(-2+3i) = -0{,}13398 + 0{,}01910i$, todavía en $\mathbb{H}$ aunque muy cerca del eje real, porque 3 está cerca de π. El mapa completo franja $\to$ disco es entonces $w = \dfrac{e^z - i}{e^z + i}$.

(e) Joukowski. El mapa $J(z) = z + \dfrac{1}{z}$ transforma el exterior del disco unidad $\{|z| > 1\}$ en el complemento del segmento $[-2, 2]$. En efecto, sobre $|z| = 1$ se tiene $z = e^{it}$ y $J = 2\cos t \in [-2, 2]$: la circunferencia se aplasta sobre el segmento. Sobre $|z| = 2$ los puntos ya no se aplastan: $t = 0{,}3$ da $J \approx 2{,}3883 + 0{,}4433i$; $t = 1$ da $1{,}3508 + 1{,}2622i$; $t = 2$ da $-1{,}0404 + 1{,}3639i$. Es el mapa con el que Joukowski convirtió circunferencias en perfiles de ala y fundó la aerodinámica del siglo XX.

Idea clave

Todo el catálogo se resume en tres verbos. Las *potencias* abren y cierran ángulos en un vértice; la *exponencial* convierte bandas en sectores (porque convierte sumas en productos, y por tanto traslaciones verticales en rotaciones); las *transformaciones de Möbius* $\frac{az+b}{cz+d}$ intercambian rectas y circunferencias sin distinguirlas. Con esos tres verbos y la composición se resuelve la inmensa mayoría de los problemas concretos, y el teorema de Riemann cubre teóricamente todo lo demás.

Cuidado

El teorema de Riemann exige que Ω sea simplemente conexo *y distinto de todo* $\mathbb{C}$. La segunda condición no es un tecnicismo: no existe ninguna biyección holomorfa $\mathbb{C} \to \mathbb{D}$, porque tal cosa sería una función entera y acotada, y Liouville la obligaría a ser constante.

Por la misma razón un anillo no es conformemente equivalente a un disco perforado, y dos anillos $\{r < |z| < R\}$ solo son equivalentes entre sí cuando coincide el cociente R/r: la conformidad recuerda ese número.

Regla: construir y usar una transformación de Möbius

Las transformaciones $T(z) = \dfrac{az+b}{cz+d}$ con $ad-bc \neq 0$ son las únicas biyecciones holomorfas de la esfera de Riemann en sí misma. Son la caja de herramientas básica de los mapas conformes.

1. **Comprueba el determinante.** Si $ad-bc=0$ la aplicación es constante y no sirve. Con $ad-bc \neq 0$ siempre hay inversa, y también es de Möbius: $T^{-1}(w) = \dfrac{dw-b}{-cw+a}$.
2. **Piensa en matrices.** Asocia a T la matriz $\left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right)$. Componer transformaciones es *multiplicar* matrices, y la inversa es la matriz inversa. Toda la teoría es álgebra lineal 2×2 disfrazada.
3. **Recuerda qué hace en el infinito.** Si $c \neq 0$: $T(-d/c) = \infty$ y $T(\infty) = a/c$. Si $c = 0$ la transformación es afín, $T(z) = (az+b)/d$, y deja fijo el infinito.
4. **Usa que manda circunferencias y rectas en circunferencias y rectas.** No conserva «circunferencia» ni «recta» por separado, pero sí la familia conjunta. Para saber en qué se transforma una curva, basta con seguir tres de sus puntos.
5. **Constrúyela con tres puntos.** Hay una única T que manda z_1, z_2, z_3 a $0, 1, \infty$, y es
$$T(z) = \frac{z-z_1}{z-z_3} \cdot \frac{z_2-z_3}{z_2-z_1}.$$
Para mandar z_1, z_2, z_3 a w_1, w_2, w_3, compón la de la izquierda con la inversa de la de la derecha: $S^{-1} \circ T$.
6. **Localiza los puntos fijos.** Resuelve $T(z) = z$, es decir $cz^2 + (d-a)z - b = 0$. Genéricamente hay dos; si hay uno solo (raíz doble) la transformación es *parabólica*. Los puntos fijos son los autovectores de la matriz vistos en la esfera.
7. **Verifica con puntos de prueba.** Evalúa T en tres puntos conocidos y comprueba que caen donde deben, y que un punto interior de la región de salida cae dentro de la región de llegada (esto último decide la orientación, y se olvida constantemente).

Ejemplo 30.56 (Cinco transformaciones de Möbius, paso a paso)**.** (a) **La transformación de Cayley: semiplano superior $\to$ disco.** Sea $T(z) = \dfrac{z-i}{z+i}$. Comprobemos con el paso 7: $T(i) = 0$ (el centro del disco), $T(0) = -1$, $T(1) = (1-i)/(1+i) = -i$, $T(\infty) = 1$. Los tres últimos tienen módulo 1: la recta real (una «circunferencia» que pasa por ∞) va a la circunferencia unidad, como anuncia el paso 4. Y el punto interior $z = i$

va al interior. Además $T(2i) = \dfrac{i}{3i} = \dfrac{1}{3}$, otra confirmación. Luego T manda $\{\operatorname{Im} z > 0\}$ biyectivamente sobre $\{|w| < 1\}$.

(b) **El diccionario inverso: disco → semiplano derecho.** Sea $S(z) = \dfrac{1+z}{1-z}$. Entonces $S(0) = 1$, $S(i) = \dfrac{1+i}{1-i} = i$, $S(-1) = 0$, $S(1) = \infty$ y $S(1/2) = 3$. La circunferencia unidad va al eje imaginario y el disco al semiplano $\{\operatorname{Re} w > 0\}$. Entre (a) y (b) tienes ya los tres dominios estándar —semiplano superior, disco, semiplano derecho— conectados por transformaciones explícitas; casi todo problema de mapas conformes empieza reduciéndose a uno de ellos.

(c) **Los automorfismos del disco.** Para $|a| < 1$, la transformación de Blaschke $\varphi_a(z) = \dfrac{z-a}{1-\bar{a}z}$ manda el disco en sí mismo y a en 0. Con $a = \frac{1}{2}$: $\varphi(1/2) = 0$, $\varphi(0) = -\frac{1}{2}$, y sobre el borde $\varphi(1) = 1$, $\varphi(-1) = -1$, $\varphi(i) = \dfrac{i - \frac{1}{2}}{1 - \frac{i}{2}} = -0{,}8 + 0{,}6\,i$, de módulo $\sqrt{0{,}64 + 0{,}36} = 1$. Todo el borde va al borde. Estas, compuestas con rotaciones $z \mapsto e^{i\alpha}z$, son *todos* los automorfismos holomorfos del disco.

(d) **Construir una a medida.** Busquemos la que manda $1, i, -1$ a $0, 1, \infty$. Por el paso 5 con $z_1 = 1$, $z_2 = i$, $z_3 = -1$:

$$T(z) = \frac{z-1}{z+1} \cdot \frac{i+1}{i-1} = -i\,\frac{z-1}{z+1},$$

porque $\dfrac{i+1}{i-1} = -i$. Verificamos: $T(1) = 0$ ✓, $T(-1) = \infty$ ✓, y $T(i) = -i \cdot \dfrac{i-1}{i+1} = -i \cdot i = 1$ ✓. De paso, $T(0) = -i \cdot (-1) = i$ y $T(\infty) = -i$: la circunferencia unidad (que pasa por $1, i, -1$) va a la recta que pasa por $0, 1, \infty$, es decir el eje real, y el interior del disco va a uno de los dos semiplanos.

(e) **Puntos fijos: dos, y uno.** Para $T(z) = \dfrac{z+2}{z+1}$ la ecuación $T(z) = z$ es $z^2 + z = z + 2$, o sea $z^2 = 2$: hay dos puntos fijos, $z = \pm\sqrt{2}$. En cambio para $U(z) = \dfrac{z}{z+1}$ la ecuación es $z^2 + z = z$, o sea $z^2 = 0$: un único punto fijo $z = 0$, doble. U es *parabólica*, y en efecto su matriz $\left(\begin{smallmatrix} 1 & 0 \\ 1 & 1 \end{smallmatrix}\right)$ no es diagonalizable —tiene el autovalor 1 repetido con un solo autovector—. La clasificación de las Möbius en elípticas, parabólicas, hiperbólicas y loxodrómicas es literalmente la clasificación de las formas de Jordan de sus matrices, capítulo 7.

Cuidado

La matriz de T está determinada solo salvo un factor escalar: $\left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right)$ y $\left(\begin{smallmatrix} 2a & 2b \\ 2c & 2d \end{smallmatrix}\right)$ dan la *misma* transformación. Por eso el grupo correcto no es $\mathrm{GL}_2(\mathbb{C})$ sino el cociente $\mathrm{PGL}_2(\mathbb{C})$, y por eso siempre se puede normalizar a $ad - bc = 1$ —lo cual todavía deja la ambigüedad de signo—.

Pausa para pensar

En el plano, «recta» y «circunferencia» parecen objetos distintos y el paso 4 suena a coincidencia afortunada. En la esfera de Riemann dejan de serlo: una recta *es* una circunferencia, la que pasa por el polo norte. Bajo la proyección estereográfica las dos familias son una sola, y el enunciado del paso 4 se vuelve la afirmación obvia de que una biyección conforme de la esfera manda circunferencias en circunferencias. Cambiar de modelo no demuestra nada nuevo, pero explica por qué el enunciado tenía que ser así.

Idea clave

Restringidas al disco (o al semiplano superior), las Möbius con coeficientes reales y determinante positivo son exactamente los *movimientos rígidos* de la geometría hiperbólica: conservan la métrica $\mathrm{d}s = \frac{2|\mathrm{d}z|}{1-|z|^2}$ igual que las rotaciones y traslaciones conservan la distancia euclídea. El lema de Schwarz–Pick dice que toda función holomorfa del disco en sí mismo contrae esa distancia, con igualdad solo para las Möbius. Y el subgrupo con coeficientes enteros, $\mathrm{PSL}_2(\mathbb{Z})$, es el grupo modular: el punto de partida de las formas modulares y, por un camino largo, de la demostración del último teorema de Fermat.

30.6. El panorama desde arriba

El análisis complejo re-teje todos los hilos del libro. Las *funciones elípticas* (inversas de las integrales del péndulo, capítulo 12) son las funciones doblemente periódicas de $\mathbb{C}$: viven en toros, y sus toros son las curvas elípticas de la criptografía y de Fermat–Wiles (capítulo 26). La función $\zeta(s) = \prod_p (1 - p^{-s})^{-1}$ (capítulo 6) se continúa analíticamente a $\mathbb{C}$ y sus ceros gobiernan la distribución de los primos: la hipótesis de Riemann, el problema abierto central de la matemática, es una afirmación de variable compleja sobre teoría de números. Las formas modulares —holomorfas en $\mathbb{H}$ (capítulo 36) con simetrías de Möbius aritméticas— son el lenguaje en que Wiles demostró Fermat vía las representaciones de Galois (capítulo 26). Y la maquinaria de contornos y residuos es la caja de herramientas diaria de la física teórica, del procesamiento de señales (la transformada de Laplace del capítulo 12 se invierte con un contorno de Bromwich) y de la combinatoria analítica. El estudiante que ha llegado hasta aquí tiene todos los prerrequisitos para un curso completo de variable compleja —Ahlfors o el visual de Needham—, y con él, las puertas de la matemática del siglo XX abiertas de par en par.

30.7. Ronda de temarios: Schwarz, Poisson, reflexión y la función Gamma

Los programas de variable compleja (Ahlfors, Stein–Shakarchi) cierran con cuatro piezas que aún no habíamos tocado y que rematan este capítulo.

30.7.1. El lema de Schwarz y los automorfismos del disco

Teorema 30.57 (Lema de Schwarz)**.** *Si $f : \mathbb{D} \to \mathbb{D}$ es holomorfa en el disco unitario con $f(0) = 0$, entonces $|f(z)| \leq |z|$ para todo z y $|f'(0)| \leq 1$; si hay igualdad en un punto interior (o $|f'(0)| = 1$), entonces $f(z) = e^{i\theta} z$: una rotación.*

Demostración. $g(z) = \frac{f(z)}{z}$ tiene singularidad evitable en 0 (pues $f(0) = 0$: el desarrollo de f empieza en z). En el círculo $|z| = r < 1$: $|g| \leq \frac{1}{r}$; por el principio del módulo máximo la cota vale en todo el disco de radio r, y haciendo $r \to 1$: $|g| \leq 1$, que es la desigualdad. Si $|g(z_0)| = 1$ en un punto interior, g alcanza su máximo dentro: g es constante de módulo 1. □ □

Idea clave

Qué dice esto en cristiano. Los tres teoremas de este bloque dicen hasta qué punto una función holomorfa está encorsetada. El lema de Schwarz dice que una función del disco en sí mismo que fija el origen no puede expandir: $|f(z)| \leq |z|$ siempre, y si en algún punto se da la igualdad, entonces f es una rotación y nada más. Picard va mucho más lejos: una función entera no constante toma *todos* los valores complejos salvo, como mucho, uno. Ser holomorfa es una condición tan rígida que deja poquísimo margen de maniobra.

Ejemplo 30.58 (La rigidez, en casos concretos)**. (a) Schwarz con igualdad.** La función $f(z) = \mathrm{e}^{\mathrm{i}\alpha} z$ cumple $|f(z)| = |z|$ en todo punto y $|f'(0)| = 1$: es el caso de igualdad, y en efecto es una rotación. La función $f(z) = z^2$ cumple $|f(z)| = |z|^2 < |z|$ para $0 < |z| < 1$ y $f'(0) = 0$: desigualdad estricta, como manda el lema.

(b) Picard con exactamente un valor omitido. La exponencial e^z es entera y no constante, y omite exactamente el valor 0: no hay ningún z con $\mathrm{e}^z = 0$. Toma todos los demás, incluso los negativos, pues $\mathrm{e}^{\mathrm{i}\pi} = -1$. Es el ejemplo extremal del teorema.

(c) Sin ningún valor omitido. Un polinomio no constante toma *todos* los valores complejos, que es el teorema fundamental del álgebra. Así que los polinomios están en el otro extremo: no omiten nada. Entre ambos casos no hay nada más, y esa es la fuerza del enunciado.

(d) Liouville como corolario barato. Si f es entera y acotada, entonces es constante. Se deduce de la fórmula integral de Cauchy: la cota sobre $|f|$ da una cota sobre $|f'(z_0)|$ que tiende a cero al agrandar el radio, luego $f' \equiv 0$. De aquí sale, en tres líneas, el teorema fundamental del álgebra: si un polinomio no tuviera raíces, $\frac{1}{p}$ sería entera y acotada, luego constante.

Cuidado

[Nada de esto vale en el caso real] La función real e^{-x^2} es infinitamente derivable, acotada y no constante: Liouville falla estrepitosamente en $\mathbb{R}$. Y $\arctan x$ es derivable en toda la recta y omite infinitos valores. La rigidez es una propiedad exclusivamente compleja, y viene de que la derivabilidad compleja impone dos ecuaciones (Cauchy–Riemann) donde la real solo impone un límite. Segundo aviso: el lema de Schwarz exige las tres

hipótesis, disco en disco, holomorfía y $f(0)=0$; sin la última se sustituye por el lema de Schwarz–Pick, que es la versión con la métrica hiperbólica.

Corolario 30.59 (Automorfismos del disco)**.** *Toda biyección holomorfa $f:\mathbb{D}\to\mathbb{D}$ es de la forma*

$$f(z)=e^{i\theta}\,\frac{z-a}{1-\bar a z},\qquad a\in\mathbb{D}.$$

Demostración. Los $\varphi_a(z)=\frac{z-a}{1-\bar a z}$ son automorfismos (envían el círculo unitario en sí mismo —comprobar $|\varphi_a(e^{it})|=1$— y $\varphi_a^{-1}=\varphi_{-a}$). Dado f con $f(a)=0$, la composición $h=f\circ\varphi_a^{-1}$ fija el 0; aplicando Schwarz a h y a h^{-1}: $|h(z)|\le|z|$ y $|z|\le|h(z)|$: igualdad, luego h es rotación. □ □

Observación 30.60. Compárese con Liouville: la rigidez holomorfa una vez más. Un difeomorfismo real del disco tiene infinitos grados de libertad; uno holomorfo, exactamente tres (a y θ). Pick observó que Schwarz dice que toda $f:\mathbb{D}\to\mathbb{D}$ holomorfa *contrae la métrica hiperbólica* del capítulo 36: la geometría de curvatura -1 es la geometría intrínseca del análisis complejo, y el lema de Schwarz es su desigualdad fundamental. Ese puente (Schwarz–Pick) une los capítulos 36 y 30 en una sola frase.

30.7.2. El núcleo de Poisson y el problema de Dirichlet en el disco

Teorema 30.61. *Si u es armónica en $\overline{\mathbb{D}}$ con valores de borde $u(e^{it})$, entonces*

$$u(re^{i\theta})=\frac{1}{2\pi}\int_0^{2\pi}P_r(\theta-t)\,u(e^{it})\,\mathrm{d}t,\qquad P_r(\psi)=\frac{1-r^2}{1-2r\cos\psi+r^2}.$$

Recíprocamente, esta fórmula resuelve *el problema de Dirichlet: dado cualquier dato de borde continuo, produce la única armónica que lo extiende.*

De dónde sale. $u=\operatorname{Re}f$ con f holomorfa (existencia de la armónica conjugada en el disco, capítulo al inicio). La fórmula de Cauchy da $f(z)$ integrando sobre el borde; el truco es restar la integral respecto del punto reflejado $1/\bar z$ (que está *fuera*, así que su integral de Cauchy es 0):

$$f(z)=\frac{1}{2\pi i}\oint f(w)\Big(\frac{1}{w-z}-\frac{1}{w-1/\bar z}\Big)\mathrm{d}w,$$

y el paréntesis, sobre $|w|=1$, se simplifica exactamente a $\frac{1-|z|^2}{|w-z|^2}\cdot\frac{1}{w}$: real y positivo. Tomando partes reales queda Poisson. Las tres virtudes del núcleo: $P_r>0$, $\frac{1}{2\pi}\int P_r=1$ (probar con la serie geométrica $P_r(\psi)=\sum_{n\in\mathbb{Z}}r^{|n|}e^{in\psi}$), y P_r se concentra en $\psi=0$ cuando $r\to1$: es una *identidad aproximada*, y por eso los valores de borde se recuperan continuamente. □ □

Ejemplo 30.62 (El núcleo de Poisson)**. (a) Qué hace.** La fórmula $u(r\mathrm{e}^{\mathrm{i}\theta})=\frac{1}{2\pi}\int_0^{2\pi}\frac{1-r^2}{1-2r\cos(\theta-t)+r^2}\,u(\mathrm{e}^{\mathrm{i}t})\,\mathrm{d}t$ reconstruye la función armónica en todo el disco a partir de sus valores en el borde. Es la solución explícita del problema de Dirichlet en el disco.

(b) En el centro. Con $r=0$ el núcleo vale 1 y la fórmula se reduce al promedio: $u(0)=\frac{1}{2\pi}\int u$. Recupera la propiedad del valor medio como caso particular.

(c) Un ejemplo. Con dato de borde $u(\mathrm{e}^{it}) = \cos t$, la solución es $u(x, y) = x$, que es armónica y coincide en el borde. La fórmula lo confirma sin necesidad de adivinar, integrando el núcleo contra $\cos t$.

(d) El comportamiento en el borde. El núcleo es positivo, tiene integral 1 y se concentra alrededor de θ cuando $r \to 1$: es una aproximación de la identidad. Por eso u recupera continuamente sus valores de borde, y por eso el mismo núcleo aparece en el estudio de las series de Fourier bajo el nombre de sumación de Abel.

Observación 30.63. En $r = 0$: $u(0) = \frac{1}{2\pi}\int u(e^{it})\,\mathrm{d}t$: la propiedad del valor medio que en el capítulo 12 enunciamos sin demostrar, ahora con demostración y con fórmula para *todos* los puntos, no solo el centro. El principio del máximo armónico es corolario inmediato (un promedio ponderado con pesos positivos no puede superar a su máximo). Y la serie $P_r = \sum r^{|n|}e^{in\psi}$ enseña qué hace el disco con los datos de borde: amortigua el armónico n-ésimo por $r^{|n|}$: Fourier del capítulo 12, sumabilidad de Abel del capítulo 6 y armónicas del capítulo 11, en una sola fórmula.

30.7.3. Prolongación analítica: el principio de reflexión

Teorema 30.64 (Reflexión de Schwarz)**.** *Sea f holomorfa en la mitad superior de un dominio simétrico respecto de $\mathbb{R}$, continua hasta el segmento real, y* real sobre el segmento*. Entonces*

$$F(z) = \begin{cases} f(z) & \text{arriba,} \\ \overline{f(\bar{z})} & \text{abajo} \end{cases}$$

es holomorfa en todo el dominio simétrico.

Demostración. $\overline{f(\bar{z})}$ es holomorfa donde está definida (sus parciales heredan Cauchy–Riemann: conjugar dos veces restaura las ecuaciones), y en el segmento ambas definiciones coinciden (ahí f es real). La holomorfía a través del segmento se verifica con *Morera* (ya probado): la integral de F sobre cualquier triangulito que cruce el eje se parte en dos trozos que tienden a cero, pues F es continua y holomorfa a cada lado. □ □

Ejemplo 30.65 (Reflexión de Schwarz)**. (a) La receta.** Si f es holomorfa arriba, continua hasta el eje real y toma valores reales allí, entonces $F(z) = \overline{f(\bar{z})}$ para z abajo pega holomorfamente con f. Con $f(z) = z^2$: abajo queda $\overline{\bar{z}^2} = z^2$, la misma función, como debe ser.

(b) Un uso. Una función holomorfa en el semidisco superior que sea real en el diámetro se extiende al disco completo. Eso duplica el dominio gratis y es la herramienta estándar para estudiar regularidad en la frontera.

(c) La hipótesis es esencial. Con $f(z) = \mathrm{i}z$, los valores en el eje real son imaginarios puros y no reales; la reflexión que corresponde entonces es respecto de la recta imaginaria, no de la real. Si los valores de borde están en una circunferencia en vez de en una recta, la reflexión es la inversión respecto de esa circunferencia.

(d) La versión general. El principio se extiende a bordes analíticos cualesquiera y es la base del método de las imágenes en electrostática: una carga frente a un plano conductor se

sustituye por la carga más su imagen especular. La física y la variable compleja están diciendo literalmente lo mismo.

Observación 30.66. Es el primer teorema de *prolongación analítica*: la función no tenía definición abajo y el teorema la *crea*, única por el principio de identidad. Así se extienden la función Γ y la zeta de Riemann más allá de sus dominios originales; así se construyen los mapeos conformes de polígonos (reflejar por cada lado, una y otra vez: el patrón de espejos de un caleidoscopio); y la versión para circunferencias (conjugar = invertir respecto del círculo) explica por qué la simetría $z \mapsto 1/\bar{z}$ apareció en el núcleo de Poisson: Poisson *es* una reflexión de Schwarz aplicada a la fórmula de Cauchy.

30.7.4. La función Gamma y la fórmula de reflexión

Del capítulo 6: $\Gamma(s) = \int_0^\infty t^{s-1}e^{-t}\mathrm{d}t$ con $\Gamma(s+1) = s\Gamma(s)$. La ecuación funcional permite prolongarla: $\Gamma(s) = \frac{\Gamma(s+1)}{s}$ define Γ en $-1 < s < 0$, luego en $-2 < s < -1$, y así: una función meromorfa en todo $\mathbb{C}$ con polos simples en $0, -1, -2, \ldots$ de residuos $\frac{(-1)^n}{n!}$ (iterar la ecuación funcional y evaluar).

Teorema 30.67 (Fórmula de reflexión de Euler)**.**

$$\Gamma(s)\,\Gamma(1-s) = \frac{\pi}{\sin \pi s}.$$

Demostración desde el producto del seno. Del producto de Weierstrass–Euler del capítulo 6, $\sin \pi s = \pi s \prod_{n\geq 1}\left(1 - \frac{s^2}{n^2}\right)$, y de la definición por límite de Gauss $\Gamma(s) = \lim_N \frac{N!\,N^s}{s(s+1)\cdots(s+N)}$ (equivalente a la integral: se prueba con convergencia dominada sobre $\left(1 - \frac{t}{N}\right)^N \to e^{-t}$), se calcula directamente

$$\frac{1}{\Gamma(s)\Gamma(1-s)} = \frac{1}{\Gamma(s)\cdot(-s)\Gamma(-s)} = s\prod_{n\geq 1}\Big(1+\frac{s}{n}\Big)e^{-s/n}\cdot\Big(1-\frac{s}{n}\Big)e^{s/n} = s\prod\Big(1-\frac{s^2}{n^2}\Big) = \frac{\sin \pi s}{\pi}.$$

Los factores exponenciales de convergencia se cancelan por parejas: el producto del seno y la función Gamma eran la misma infraestructura. □ □

Ejemplo 30.68. $s = \frac{1}{2}$: $\Gamma(\frac{1}{2})^2 = \frac{\pi}{\sin(\pi/2)} = \pi$, es decir $\Gamma(\frac{1}{2}) = \sqrt{\pi}$: la integral gaussiana por *tercera* vía (capítulo 9: coordenadas polares; capítulo 11: por volúmenes; ahora: por reflexión). Además la fórmula muestra que Γ *no tiene ceros* (si $\Gamma(s_0) = 0$, el lado izquierdo se anularía donde el derecho no): $\frac{1}{\Gamma}$ es entera, la campeona de las funciones enteras con los ceros prescritos $0, -1, -2, \ldots$, y su producto de Weierstrass $\frac{1}{\Gamma(s)} = se^{\gamma s}\prod(1 + \frac{s}{n})e^{-s/n}$ es donde la constante γ de Euler (capítulo 6) encuentra su papel estructural: no era una curiosidad de la serie armónica sino el "coeficiente de normalización" de la función especial más importante después de las elementales.

30.8. Cierre definitivo: mapeo de Riemann, Picard y funciones elípticas

Las tres cumbres restantes de la variable compleja clásica, con las ideas completas de sus demostraciones.

30.8.1. El teorema del mapeo de Riemann, demostrado

Teorema 30.69. *Todo dominio simplemente conexo $\Omega \subsetneq \mathbb{C}$ es biholomorfo al disco $\mathbb{D}$.*

Esquema de la demostración (Riesz–Fejér). Paso 0. Hay que usar $\Omega \neq \mathbb{C}$ (Liouville prohíbe mapear $\mathbb{C}$ al disco): con $a \notin \Omega$, la raíz $\sqrt{z-a}$ admite rama holomorfa en Ω (simple conexidad: el logaritmo existe, capítulo anterior); esta rama es inyectiva y esquiva un disco entero de valores (si toma w, no toma $-w$), y componiendo con una inversión de Möbius se logra $\Omega \hookrightarrow \mathbb{D}$: *podemos suponer* $\Omega \subseteq \mathbb{D}$ *con* $0 \in \Omega$.

Paso 1 (el problema extremal). Considérese la familia $\mathcal{F}$ de inyecciones holomorfas $f : \Omega \to \mathbb{D}$ con $f(0) = 0$, y búsquese la que *maximiza* $|f'(0)|$. La familia es no vacía (la identidad) y uniformemente acotada; el teorema de *Montel* —toda familia uniformemente acotada de holomorfas tiene subsucesiones que convergen uniformemente en compactos; se demuestra con Arzelà–Ascoli (¡capítulo 12!) más la estimación de Cauchy que da equicontinuidad gratis— entrega una f óptima (la inyectividad sobrevive al límite por Hurwitz/Rouché).

Paso 2 (el óptimo es sobreyectivo). Si f omitiera un valor $w \in \mathbb{D}$, se fabrica una competidora mejor: componer con el automorfismo φ_w (que mueve w a 0), sacar raíz cuadrada (legal: el dominio imagen es simplemente conexo y esquiva 0), y recentrar con otro automorfismo. La raíz cuadrada es *expansiva* en el sentido de Schwarz (su inversa $z \mapsto z^2$ contrae estrictamente, lema de Schwarz), y el cómputo da $|g'(0)| > |f'(0)|$: contradicción con la maximalidad. Luego f es sobre: biholomorfismo. □ □

Idea clave

Qué dice esto en cristiano. El teorema de la aplicación de Riemann dice algo casi increíble: cualquier región del plano que no tenga agujeros y no sea el plano entero se puede deformar holomorfamente hasta convertirla en un disco. Da igual lo retorcida que sea su forma: un triángulo, una banda, el interior de una espiral, todos son "el mismo" desde el punto de vista del análisis complejo. Eso permite resolver un problema en la región complicada trasladándolo al disco, donde todo se sabe, y volviendo. Es la base de la cartografía conforme y de la aerodinámica clásica.

Ejemplo 30.70 (Deformaciones conformes concretas)**. (a) El semiplano y el disco.** La transformación de Cayley $f(z) = \frac{z-\mathrm{i}}{z+\mathrm{i}}$ manda el semiplano superior al disco unidad. Comprobación: para $z = 0$ da -1, que está en el borde; para $z = \mathrm{i}$ da 0, el centro; y para $z = 2\mathrm{i}$ da $\frac{\mathrm{i}}{3\mathrm{i}} = \frac{1}{3}$, dentro del disco.

(b) La banda y el semiplano. La exponencial manda la banda $\{0 < \operatorname{Im} z < \pi\}$ al semiplano superior: si $z = x + \mathrm{i}y$ con $0 < y < \pi$, entonces e^z tiene argumento y, que recorre exactamente $(0, \pi)$. Componiendo con (a) se llega de la banda al disco.

(c) Los automorfismos del disco. Las únicas biyecciones holomorfas del disco en sí mismo son $f(z) = e^{i\alpha}\frac{z-a}{1-\bar{a}z}$ con $|a| < 1$. Son las que hacen que la aplicación de Riemann no sea única: se puede componer con cualquiera de ellas. La unicidad se recupera fijando adónde va un punto y en qué dirección.

(d) Lo que se conserva y lo que no. Una transformación conforme conserva los *ángulos* entre curvas, pero no las longitudes ni las áreas. Por eso las funciones armónicas se transportan bien de una región a otra (el laplaciano se conserva salvo un factor) y los problemas de Dirichlet se pueden mover al disco, donde la fórmula de Poisson los resuelve de golpe.

Cuidado

[Las dos exclusiones del enunciado son imprescindibles] El plano entero $\mathbb{C}$ *no* es biholomorfo al disco: si lo fuera, la inversa sería una función entera y acotada, luego constante por Liouville. Y un dominio con un agujero, como una corona, no es simplemente conexo y tampoco puede serlo, porque el disco sí lo es y esa propiedad se conserva. Segundo aviso: el teorema es de existencia pura y no da la fórmula; encontrar explícitamente la aplicación para una región concreta puede ser durísimo, y para polígonos existe la fórmula de Schwarz–Christoffel, que se escribe como una integral que rara vez es elemental.

Observación 30.71. Obsérvese la arquitectura, que es la del libro entero: un problema de existencia convertido en *problema variacional* (maximizar $|f'(0)|$), compacidad para obtener el óptimo (Montel = Bolzano–Weierstrass funcional), y el lema de Schwarz como desigualdad que castiga a los no-óptimos. La unicidad (salvo rotación, fijando $f(0) = 0$, $f'(0) > 0$) es puro Schwarz. Que la frontera se extienda continuamente (Carathéodory) y las fórmulas explícitas (Schwarz–Christoffel para polígonos) son los siguientes escalones del mismo edificio.

30.8.2. Los teoremas de Picard

Teorema 30.72 (Picard pequeño)**.** *Una función entera no constante omite* a lo sumo un *valor complejo. (Ejemplo límite: e^z omite exactamente el* 0*.)*

Idea. Si f omite 0 y 1, su imagen vive en $\mathbb{C} \setminus \{0, 1\}$, cuyo *recubridor universal es el disco* (la función modular λ: el plano rayado por un mosaico hiperbólico de triángulos ideales, la geometría $K = -1$ del capítulo 36 trabajando para el análisis). Levantar f al recubridor produce una entera con valores en $\mathbb{D}$: constante por Liouville, luego f constante. Todo el teorema es: *el plano doblemente pinchado es hiperbólico, y Liouville mata los mapas de lo parabólico a lo hiperbólico.* □ □

Ejemplo 30.73 (Cuánto omite una función entera)**. (a) El caso extremo.** e^z omite exactamente un valor, el 0: para todo $w \neq 0$ la ecuación $e^z = w$ tiene infinitas soluciones. La cota del teorema se alcanza.

(b) Los polinomios no omiten nada. Por el teorema fundamental del álgebra, $p(z) = w$ siempre tiene solución. Un polinomio no constante es sobreyectivo sobre $\mathbb{C}$.

(c) La escalera de resultados. Liouville dice que una entera acotada es constante; Casorati–Weierstrass dice que cerca de una singularidad esencial la imagen es densa; Picard pequeño dice que omite a lo sumo un valor; Picard grande dice que en cualquier entorno de una singularidad esencial toma *todos* los valores salvo a lo sumo uno, e infinitas veces cada uno. Cada peldaño es sustancialmente más fuerte y más difícil.

(d) Un ejemplo de Picard grande. $\mathrm{e}^{1/z}$ tiene una singularidad esencial en el origen y, en cualquier disco perforado alrededor de él, toma todos los valores complejos salvo el 0, cada uno infinitas veces: las soluciones de $\mathrm{e}^{1/z} = w$ son $z = 1/(\ln w + 2\pi \mathrm{i} k)$, que se acumulan en el origen.

Observación 30.74. El *gran* teorema de Picard afina: cerca de una singularidad esencial, la función toma *todos* los valores salvo a lo sumo uno, infinitas veces cada uno (Casorati–Weierstrass, que ya probamos, era la sombra tibia de esto). La jerarquía final de singularidades queda así: evitable (acotada), polo (explota ordenadamente), esencial (visita todo el plano en cada vecindad: caos total).

30.8.3. Funciones elípticas: $\wp$ de Weierstrass

¿Existen funciones *doblemente* periódicas ($f(z+\omega_1) = f(z+\omega_2) = f(z)$ con periodos independientes)? Holomorfas no (serían acotadas en el paralelogramo compacto: Liouville: constantes); con polos, sí:

Definición 30.75. Para el retículo $\Lambda = \mathbb{Z}\omega_1 + \mathbb{Z}\omega_2$:

$$\wp(z) = \frac{1}{z^2} + \sum_{\omega \in \Lambda \setminus 0} \Big(\frac{1}{(z-\omega)^2} - \frac{1}{\omega^2} \Big)$$

(el contratérmino $-\frac{1}{\omega^2}$ fuerza la convergencia absoluta: comparar con $\sum |\omega|^{-3}$, que converge: es la serie de Eisenstein del truco $\varepsilon/2^k$ en dos dimensiones).

Teorema 30.76. *$\wp$ es par, Λ-periódica, y satisface la ecuación diferencial*

$$(\wp')^2 = 4\wp^3 - g_2\,\wp - g_3, \qquad g_2 = 60\sum_{\omega \neq 0} \omega^{-4}, \quad g_3 = 140\sum_{\omega \neq 0} \omega^{-6}.$$

Idea. Desarrollar $\wp$ en Laurent cerca de 0 ($\wp = z^{-2} + \frac{g_2}{20} z^2 + \frac{g_3}{28} z^4 + \cdots$), formar la combinación $(\wp')^2 - 4\wp^3 + g_2\wp + g_3$ y comprobar que su parte principal se cancela por diseño: queda una función elíptica *sin polos* que vale 0 en 0: constante cero por Liouville en el toro. La demostración entera es "Liouville más álgebra de coeficientes": el método del capítulo, elevado a fábrica de identidades. □ □

Observación 30.77 (El puente final del libro). La ecuación dice que $z \mapsto (\wp(z), \wp'(z))$ parametriza la *curva cúbica* $y^2 = 4x^3 - g_2 x - g_3$: todo toro complejo es una curva elíptica y viceversa. Aquí confluyen de golpe: las integrales $\int \frac{\mathrm{d}x}{\sqrt{\text{cúbica}}}$ que el capítulo 6 declaró no elementales (son la inversa de $\wp$: por eso no eran elementales), el péndulo exacto del capítulo 12 (su solución *es* elíptica), los retículos del capítulo 7, la suma geométrica de puntos en la cúbica

que es un *grupo* abeliano (capítulo 25) —el grupo con que trabaja la criptografía de curvas elípticas y con el que Wiles demostró Fermat—, y la función theta de Poisson del capítulo 31, que es el ladrillo con que se construyen todas las funciones elípticas. Ningún objeto del libro anuda más hilos: es el punto de fuga donde convergen las perspectivas de todo el libro.

30.9. Segunda ronda: Rouché y el principio del argumento

30.9.1. El principio del argumento

Teorema 30.78. *Para f meromorfa dentro de γ: $\frac{1}{2\pi i}\oint \frac{f'}{f}\,\mathrm{d}z = Z - P$ (ceros menos polos, con multiplicidad): cerca de un cero de orden m, $\frac{f'}{f} \approx \frac{m}{z}$. Geométricamente: las vueltas de $f(\gamma)$ alrededor del origen —el índice del capítulo 11 aplicado a la imagen—.*

Teorema 30.79 (Rouché)**.** *$|g| < |f|$ sobre $\gamma \Rightarrow$ f y $f+g$ tienen los mismos ceros dentro (el perro atado corto no cambia las vueltas del amo alrededor del árbol).*

Ejemplo 30.80 (Contar raíces sin resolver)**.** $z^5 + 3z + 1$: en $|z| = 1$, $|3z| = 3 > 2 \geq |z^5 + 1|$: **una** raíz en el disco unitario (las de $3z$). En $|z| = 2$: $|z^5| = 32 > 7$: cinco. Conclusión sin cálculo: una dentro del disco, cuatro en el anillo $1 < |z| < 2$. Rouché da también la demostración más corta del TFA (en un círculo grande z^n domina: n raíces) —la cuarta del libro— y el criterio de Nyquist del control es el principio del argumento con casco de ingeniero.

Regla: Cómo contar ceros y polos con el principio del argumento y Rouché

Contar raíces dentro de una región no requiere calcularlas. Se cuenta integrando, o comparando con una función más simple.

1. **Verifica las hipótesis.** Toma γ una curva cerrada simple, positivamente orientada, y f meromorfa dentro y holomorfa sobre γ, *sin ceros ni polos sobre* γ. Si un cero cae justo encima de la curva, el método no dice nada: mueve el radio.

2. **Aplica el principio del argumento.** Entonces

$$\frac{1}{2\pi i}\oint_\gamma \frac{f'(z)}{f(z)}\,\mathrm{d}z = Z - P,$$

   donde Z y P son el número de ceros y de polos dentro de γ, contados con multiplicidad. Si f es holomorfa, $P = 0$ y la integral cuenta ceros directamente.

3. **Interpreta geométricamente.** Esa integral es $\dfrac{1}{2\pi}\Delta_\gamma \arg f$: el número de vueltas que da la curva imagen $f(\gamma)$ alrededor del origen. Contar ceros es contar vueltas.

4. **Para polinomios, usa Rouché en vez de integrar.** Si $|f(z) - g(z)| < |g(z)|$ para todo $z \in \gamma$, entonces f y g tienen el mismo número de ceros dentro. La estrategia: elige como g el término *dominante* sobre esa curva concreta y mete todo lo demás en la diferencia.

5. **Elige bien el término dominante según el radio.** Sobre círculos pequeños domina el término de menor grado; sobre círculos grandes, el de mayor grado. Por eso el mismo polinomio se compara con g distintas en $|z| = 1$ y en $|z| = 2$.

6. **Acota la diferencia con la desigualdad triangular.** Sobre $|z| = R$ se usa $|z^k| = R^k$ y $|a + b| \le |a| + |b|$. La desigualdad de Rouché debe ser *estricta*; si sale una igualdad, prueba con otro radio.

7. **Cuenta por diferencia para anillos.** El número de ceros en $r < |z| < R$ es (ceros en $|z| < R$) menos (ceros en $|z| < r$). Y como control global, el teorema fundamental del álgebra: la suma sobre todos los anillos debe dar el grado.

Ejemplo 30.81 (Localizar raíces sin resolver la ecuación)**. (a)** $f(z) = z^4 - 6z + 3$**: ¿cuántas raíces en el disco unidad?** Sobre $|z| = 1$ comparamos con $g(z) = -6z$, que ahí vale $|g| = 6$. La diferencia es $f - g = z^4 + 3$, y

$$|z^4 + 3| \le |z|^4 + 3 = 1 + 3 = 4 < 6 = |-6z|.$$

Rouché aplica: f tiene tantos ceros en $|z| < 1$ como $g(z) = -6z$, es decir, exactamente 1.

(b) El mismo f **en** $|z| < 2$**.** Ahora domina el grado máximo: tomamos $g(z) = z^4$, que sobre $|z| = 2$ vale 16. La diferencia es $-6z + 3$, y

$$|-6z + 3| \le 6 \cdot 2 + 3 = 15 < 16.$$

Luego f tiene 4 ceros en $|z| < 2$, que son todos (el grado es 4). Restando, en el anillo $1 < |z| < 2$ hay $4 - 1 = 3$ raíces.

(c) Contraste con el cálculo directo. Las cuatro raíces de $z^4 - 6z + 3$ son, numéricamente, 0,51140, 1,60444 y el par conjugado $-1{,}05792 \pm 1{,}59282i$, de módulos 0,51140, 1,60444, 1,91214 y 1,91214. Exactamente una tiene módulo menor que 1 y las otras tres están entre 1 y 2, tal y como predijo Rouché sin resolver nada.

(d) Verificación por el principio del argumento. Integrando numéricamente

$$\frac{1}{2\pi i} \oint_{|z|=1} \frac{4z^3 - 6}{z^4 - 6z + 3}\, \mathrm{d}z$$

se obtiene 1,0000, y sobre $|z| = 2$ se obtiene 4,00. Los dos caminos, el analítico y el numérico, dan la misma cuenta.

(e) Otro polinomio: $h(z) = z^5 + 3z + 1$**.** Sobre $|z| = 1$ tomamos $g(z) = 3z$ con $|g| = 3$, y $|h - g| = |z^5 + 1| \le 1 + 1 = 2 < 3$. Luego h tiene exactamente 1 raíz en el disco unidad. Numéricamente sus raíces son $-0{,}33199$, el par $1{,}00507 \pm 0{,}93726i$ de módulo 1,26289, y el par $-0{,}83907 \pm 0{,}94385i$ de módulo 1,37427: en efecto una sola dentro. La integral del principio del argumento sobre $|z| = 1$ da 1,000.

Cuidado

La desigualdad de Rouché debe verificarse *en toda* la curva, no en un punto conveniente. Y el sentido importa: la conclusión es que f y g tienen el mismo número de ceros, no que estén cerca unos de otros. En (a) el cero de $g(z) = -6z$ está en el origen y el de f en 0,5114; Rouché cuenta, no localiza.

Pausa para pensar

Aplica Rouché a $f(z) = z^n$ y $g(z) = z^n - \varepsilon$ sobre $|z| = 1$, con ε pequeño. Obtienes que $z^n - \varepsilon$ tiene n raíces en el disco, lo cual es cierto: son las n raíces n-ésimas de ε, todas de módulo $\varepsilon^{1/n}$. Ahora observa qué pasa cuando $\varepsilon \to 0$: las n raíces convergen todas al origen, donde z^n tiene un cero de orden n. Esta es la forma correcta de entender la multiplicidad: un cero de orden n es el límite de n ceros simples que se juntan.

30.9.2. La inversión de Laplace: Bromwich

$$f(t) = \frac{1}{2\pi i}\int_{c-i\infty}^{c+i\infty} F(s)e^{st}\,\mathrm{d}s\,:$$

cerrando por la izquierda, *la antitransformada es la suma de residuos de* $F(s)e^{st}$: para $F = \frac{1}{s^2+1}$, los residuos en $\pm i$ dan $\frac{e^{it}-e^{-it}}{2i} = \sin t$ ✓. Toda la tabla del capítulo 12 es contabilidad de residuos: la variable compleja no compite con lo aprendido: lo explica.

30.10. El índice de una curva y el teorema de los residuos, demostrado

El teorema de los residuos se enunció hablando de lo que hay "dentro de γ". Esa frase no es una definición: una curva cerrada puede dar tres vueltas alrededor de un punto, o dar una en un sentido y otra en el contrario. La noción correcta es el *índice*, y con ella el teorema admite un enunciado limpio y una demostración completa.

Definición 30.82 (Índice o número de vueltas)**.** Sea $\gamma : [0,1] \to \mathbb{C}$ una curva cerrada de clase C^1 a trozos y $a \notin \gamma([0,1])$. El *índice de* γ *respecto de* a es

$$n(\gamma, a) = \frac{1}{2\pi \mathrm{i}} \oint_\gamma \frac{\mathrm{d}z}{z-a}.$$

Teorema 30.83 (Propiedades del índice)**.** *Con las hipótesis anteriores:*

1. $n(\gamma, a)$ *es un número* entero*;*
2. $a \mapsto n(\gamma, a)$ *es constante en cada componente conexa de* $\mathbb{C} \setminus \gamma([0,1])$*;*
3. *vale* 0 *en la componente no acotada.*

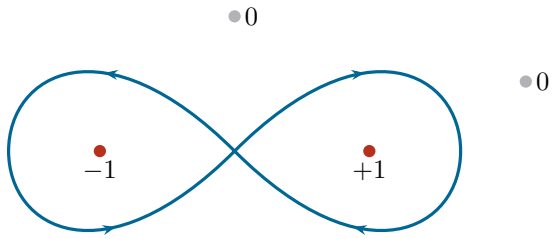


Figura 30.3: Una curva en forma de ocho, recorrida en el sentido que indican las flechas. El índice $n(\gamma,a)=\frac{1}{2\pi\mathrm{i}}\oint\frac{\mathrm{d}z}{z-a}$ vale $+1$ en el lóbulo derecho, -1 en el izquierdo —porque allí la curva gira en sentido contrario— y 0 fuera. Es la razón por la que la frase "los polos que hay dentro de γ" no basta: los dos lóbulos encierran polos, pero contribuyen con signos opuestos.

Demostración. (1) Defínase $h(t)=\displaystyle\int_0^t\frac{\gamma'(s)}{\gamma(s)-a}\,\mathrm{d}s$, de modo que $h(0)=0$ y $h(1)=2\pi\mathrm{i}\,n(\gamma,a)$. Considérese $\varphi(t)=\mathrm{e}^{-h(t)}(\gamma(t)-a)$. Donde γ es derivable,

$$\varphi'(t)=\mathrm{e}^{-h(t)}\Big(\gamma'(t)-h'(t)(\gamma(t)-a)\Big)=\mathrm{e}^{-h(t)}\Big(\gamma'(t)-\frac{\gamma'(t)}{\gamma(t)-a}(\gamma(t)-a)\Big)=0.$$

Como φ es continua y de derivada nula salvo en un número finito de puntos, es constante. En particular $\varphi(1)=\varphi(0)$, es decir $\mathrm{e}^{-h(1)}(\gamma(1)-a)=\gamma(0)-a$. Al ser la curva cerrada, $\gamma(1)=\gamma(0)$, y como $\gamma(0)\neq a$ se puede cancelar: $\mathrm{e}^{-h(1)}=1$, luego $h(1)\in 2\pi\mathrm{i}\mathbb{Z}$ y $n(\gamma,a)\in\mathbb{Z}$.

(2) La función $a\mapsto n(\gamma,a)$ es continua en el complementario de la curva (el integrando depende continuamente de a y el dominio de integración es compacto). Una función continua con valores enteros sobre un conjunto conexo es constante.

(3) Sea L la longitud de γ y $R=\max_t|\gamma(t)|$. Si $|a|>R$, la estimación ML da

$$|n(\gamma,a)|\le\frac{1}{2\pi}\cdot\frac{L}{|a|-R}\xrightarrow[|a|\to\infty]{}0,$$

y siendo entero, vale 0 para $|a|$ grande. Por (2) vale 0 en toda la componente no acotada. □

Ejemplo 30.84 (Contar vueltas)**.** Para $\gamma(t)=\mathrm{e}^{2\pi\mathrm{i}mt}$, $t\in[0,1]$, con $m\in\mathbb{Z}$, se tiene $n(\gamma,0)=m$: la circunferencia recorrida m veces. Para una figura de ocho, el índice vale $+1$ en un lóbulo, -1 en el otro y 0 fuera. Y ese es exactamente el motivo por el que hace falta el índice: una integral sobre el ocho no "rodea" los dos polos por igual, sino que uno cuenta con signo $+$ y el otro con signo $-$.

Definición 30.85 (Ciclo homólogo a cero)**.** Un *ciclo* en Ω es una suma formal finita de curvas cerradas contenidas en Ω; el índice se extiende por linealidad. Un ciclo Γ es *homólogo a cero en* Ω, y se escribe $\Gamma\sim 0$, si $n(\Gamma,a)=0$ para todo $a\notin\Omega$.

Intuitivamente: el ciclo no rodea ningún agujero del dominio. Con esto, la versión definitiva del teorema de Cauchy es la siguiente.

Teorema 30.86 (Cauchy, versión homológica)**.** *Si f es holomorfa en un abierto Ω y Γ es un ciclo homólogo a cero en Ω, entonces*

$$\oint_\Gamma f(z)\,\mathrm{d}z = 0, \qquad \text{y además} \qquad n(\Gamma,a)\,f(a) = \frac{1}{2\pi\mathrm{i}}\oint_\Gamma \frac{f(z)}{z-a}\,\mathrm{d}z \quad \text{para } a\in\Omega\setminus\Gamma.$$

Idea clave

La demostración de Dixon, en un párrafo. Se define $g(z,w)=\frac{f(w)-f(z)}{w-z}$ si $w\neq z$ y $g(z,z)=f'(z)$; la clave es que g es continua en $\Omega\times\Omega$ y holomorfa en cada variable. Se pone $h(z)=\frac{1}{2\pi\mathrm{i}}\oint_\Gamma g(z,w)\mathrm{d}w$ para $z\in\Omega$, y $h(z)=\frac{1}{2\pi\mathrm{i}}\oint_\Gamma \frac{f(w)}{w-z}\mathrm{d}w$ para z en el abierto $\Omega_0=\{z\notin\Gamma : n(\Gamma,z)=0\}$; la hipótesis $\Gamma\sim 0$ garantiza que $\Omega\cup\Omega_0=\mathbb{C}$, y las dos definiciones coinciden en la intersección, de modo que h es entera. Fuera de un disco grande, Ω_0 contiene todo, y allí $h\to 0$; por Liouville $h\equiv 0$, que es exactamente la segunda fórmula. La primera sale aplicando la segunda a $z\mapsto (z-a)f(z)$.

Teorema 30.87 (Teorema de los residuos)**.** *Sea f holomorfa en $\Omega\setminus\{a_1,\dots,a_k\}$, con singularidades aisladas en los a_j, y sea Γ un ciclo homólogo a cero en Ω que no pasa por ninguna a_j. Entonces*

$$\frac{1}{2\pi\mathrm{i}}\oint_\Gamma f(z)\,\mathrm{d}z = \sum_{j=1}^{k} n(\Gamma,a_j)\ \mathrm{Res}(f,a_j).$$

Demostración. Alrededor de cada a_j desarróllese f en serie de Laurent (teorema 30.89 de la sección siguiente) y llámese *parte principal* a

$$P_j(z)=\sum_{n\geq 1}\frac{c_{-n}^{(j)}}{(z-a_j)^n},$$

que es una serie convergente en $\mathbb{C}\setminus\{a_j\}$ y define allí una función holomorfa. La función

$$F = f - \sum_{j=1}^{k} P_j$$

es holomorfa en $\Omega\setminus\{a_1,\dots,a_k\}$ y, en cada a_j, tiene una singularidad evitable: al restar la parte principal, lo que queda es la parte regular de la serie de Laurent, que es holomorfa. Extendida a esos puntos, F es holomorfa en todo Ω, y el teorema 30.86 da $\oint_\Gamma F=0$, es decir

$$\oint_\Gamma f = \sum_{j=1}^{k}\oint_\Gamma P_j.$$

Queda calcular cada $\oint_\Gamma P_j$. Para $n\geq 2$, la función $(z-a_j)^{-n}$ tiene la primitiva $\frac{(z-a_j)^{1-n}}{1-n}$ en $\mathbb{C}\setminus\{a_j\}$, y la integral de una derivada sobre una curva cerrada es nula. Sólo sobrevive el término $n=1$:

$$\oint_\Gamma P_j = c_{-1}^{(j)}\oint_\Gamma \frac{\mathrm{d}z}{z-a_j} = 2\pi\mathrm{i}\,n(\Gamma,a_j)\,c_{-1}^{(j)},$$

por la definición 30.82. Como $c_{-1}^{(j)} = \operatorname{Res}(f, a_j)$, se obtiene el enunciado. (El intercambio de suma e integral está justificado porque la serie de la parte principal converge uniformemente sobre el compacto Γ, que está a distancia positiva de a_j.) □

Observación 30.88 (Qué se ha ganado). El enunciado con índices no es un adorno. Sin él no se puede integrar sobre una curva que se cruce a sí misma, ni sobre una que dé varias vueltas, ni justificar por qué al recorrer un contorno en sentido horario cambia el signo; y el "dentro de γ" informal esconde el teorema de la curva de Jordan, que es difícil y que aquí no hace ninguna falta. La versión homológica sustituye una intuición topológica delicada por un número que se calcula con una integral.

30.11. El teorema de Laurent, y el cálculo sistemático de residuos

El residuo se definió como el coeficiente c_{-1} de un desarrollo cuya existencia no se había probado. Se prueba ahora.

Teorema 30.89 (Laurent). *Sea f holomorfa en la corona $A = \{z : r < |z - z_0| < R\}$ con $0 \le r < R \le \infty$. Entonces existe una única familia de coeficientes $(c_n)_{n\in\mathbb{Z}}$ tal que*

$$f(z) = \sum_{n=-\infty}^{\infty} c_n\,(z - z_0)^n \qquad \text{para todo } z \in A,$$

con convergencia absoluta y uniforme sobre compactos de A. Además, para cualquier $\rho \in (r, R)$,

$$c_n = \frac{1}{2\pi \mathrm{i}} \oint_{|w-z_0|=\rho} \frac{f(w)}{(w - z_0)^{n+1}}\,\mathrm{d}w.$$

Demostración. Existencia. Fíjese $z \in A$ y elíjanse ρ_1, ρ_2 con $r < \rho_1 < |z - z_0| < \rho_2 < R$. El ciclo $\Gamma = C_{\rho_2} - C_{\rho_1}$ (las dos circunferencias, la exterior en sentido positivo y la interior en sentido negativo) es homólogo a cero en A y tiene índice 1 respecto de z, de modo que el teorema 30.86 da

$$f(z) = \frac{1}{2\pi \mathrm{i}} \oint_{C_{\rho_2}} \frac{f(w)}{w - z}\mathrm{d}w - \frac{1}{2\pi \mathrm{i}} \oint_{C_{\rho_1}} \frac{f(w)}{w - z}\mathrm{d}w.$$

En la primera integral $|z - z_0| < |w - z_0| = \rho_2$, y se desarrolla

$$\frac{1}{w - z} = \frac{1}{(w - z_0) - (z - z_0)} = \frac{1}{w - z_0} \cdot \frac{1}{1 - \frac{z - z_0}{w - z_0}} = \sum_{n\ge 0} \frac{(z - z_0)^n}{(w - z_0)^{n+1}},$$

serie geométrica de razón de módulo $|z - z_0|/\rho_2 < 1$, luego uniformemente convergente en w; se integra término a término y salen los c_n con $n \ge 0$.

En la segunda, en cambio, $|w - z_0| = \rho_1 < |z - z_0|$, y se desarrolla al revés:

$$-\frac{1}{w - z} = \frac{1}{(z - z_0) - (w - z_0)} = \frac{1}{z - z_0} \cdot \frac{1}{1 - \frac{w - z_0}{z - z_0}} = \sum_{m\ge 1} \frac{(w - z_0)^{m-1}}{(z - z_0)^m},$$

con razón de módulo $\rho_1/|z-z_0|<1$. Integrando término a término salen los coeficientes con índice negativo, $c_{-m}=\frac{1}{2\pi\mathrm{i}}\oint_{C_{\rho_1}} f(w)(w-z_0)^{m-1}\mathrm{d}w$, que es la fórmula del enunciado con $n=-m$. Que el radio ρ pueda tomarse cualquiera en (r,R) se sigue de que el integrando es holomorfo en la corona y dos circunferencias concéntricas son homólogas allí.

Unicidad. Supóngase $f(z)=\sum_m b_m(z-z_0)^m$ en A con convergencia uniforme en compactos. Multiplíquese por $(z-z_0)^{-n-1}$ e intégrese sobre $|z-z_0|=\rho$. La convergencia uniforme permite intercambiar, y

$$\oint_{|z-z_0|=\rho}(z-z_0)^{m-n-1}\mathrm{d}z=\begin{cases}2\pi\mathrm{i}, & m=n,\\ 0, & m\neq n,\end{cases}$$

(pues para $m\neq n$ el integrando tiene primitiva), de donde $b_n=c_n$. □

Definición 30.90 (Clasificación de singularidades por la serie)**.** Si f es holomorfa en $0<|z-z_0|<R$ y $\sum c_n(z-z_0)^n$ es su serie de Laurent, la singularidad en z_0 es

- *evitable* si $c_n=0$ para todo $n<0$;
- un *polo de orden* m si $c_{-m}\neq 0$ y $c_n=0$ para $n<-m$;
- *esencial* si hay infinitos $n<0$ con $c_n\neq 0$.

El *residuo* es $\operatorname{Res}(f,z_0)=c_{-1}$.

Regla: Calcular residuos sin desarrollar la serie

1. **Polo simple:** $\operatorname{Res}(f,a)=\lim_{z\to a}(z-a)f(z)$.
2. **Cociente con denominador de cero simple:** si $f=p/q$ con $p(a)\neq 0$, $q(a)=0$ y $q'(a)\neq 0$, entonces $\operatorname{Res}(f,a)=\dfrac{p(a)}{q'(a)}$. Es la versión que más se usa, y evita factorizar q.
3. **Polo de orden m:** $\operatorname{Res}(f,a)=\dfrac{1}{(m-1)!}\lim_{z\to a}\dfrac{\mathrm{d}^{m-1}}{\mathrm{d}z^{m-1}}\big[(z-a)^m f(z)\big]$.
4. **Singularidad esencial:** no hay atajo; hay que desarrollar. Por ejemplo, $\mathrm{e}^{1/z}=\sum_{n\geq 0}\frac{1}{n!z^n}$ tiene residuo 1 en 0.
5. **Residuo en el infinito:** $\operatorname{Res}(f,\infty)=-\operatorname{Res}\big(\frac{1}{w^2}f(\frac{1}{w}),0\big)$, y la suma de *todos* los residuos, incluido el del infinito, de una función racional es cero: un truco que ahorra medio examen.

Justificación de (2) y (3). Para (2), $\lim_{z\to a}(z-a)\frac{p(z)}{q(z)}=p(a)\lim_{z\to a}\frac{z-a}{q(z)-q(a)}=\frac{p(a)}{q'(a)}$. Para (3), si f tiene polo de orden m entonces $(z-a)^m f(z)=c_{-m}+c_{-m+1}(z-a)+\cdots+c_{-1}(z-a)^{m-1}+\cdots$ es holomorfa cerca de a; derivando $m-1$ veces y evaluando en a, el único término que sobrevive es $(m-1)!\,c_{-1}$. □

Ejemplo 30.91 (Los tres tipos, en un vistazo)**.** En $z_0=0$: la función $\frac{\sin z}{z}=1-\frac{z^2}{6}+\cdots$ tiene singularidad evitable; $\frac{\cos z}{z^3}=\frac{1}{z^3}-\frac{1}{2z}+\cdots$ tiene un polo de orden 3 con residuo $-\frac{1}{2}$;

y $\mathrm{e}^{1/z}$ tiene singularidad esencial. La diferencia no es estética: cerca de una singularidad esencial, el teorema de Casorati–Weierstrass dice que f toma valores arbitrariamente próximos a *cualquier* número complejo, y el gran teorema de Picard, que los toma *todos* salvo a lo sumo uno, infinitas veces. Ninguna función con polo hace eso.

30.12. Integrales reales por residuos: lema de Jordan, valor principal y el semicírculo pequeño

Los ejemplos del capítulo usaban dos pasos sin justificar: que el arco grande "muere" cuando hay un factor $\mathrm{e}^{\mathrm{i}az}$, y que un polo *sobre* el eje real aporta media vuelta. Los dos son lemas cortos.

Lema 30.92 (Estimación ML y su límite)**.** *Si $|f(z)| \le M_R$ sobre el arco $C_R = \{R\mathrm{e}^{\mathrm{i}\theta} : 0 \le \theta \le \pi\}$, entonces $\left|\int_{C_R} f\right| \le \pi R\, M_R$. En particular, si $R\,M_R \to 0$ la integral tiende a cero.*

Esto basta para $f = p/q$ con $\deg q \ge \deg p + 2$, pero *no* basta para $\frac{\mathrm{e}^{\mathrm{i}az}}{z}$, donde $RM_R \to 1$. Ahí hace falta algo mejor.

Lema 30.93 (Jordan)**.** *Sea $a > 0$ y f continua sobre los arcos C_R del semiplano superior, con $M_R = \max_{C_R}|f| \to 0$ cuando $R \to \infty$. Entonces*

$$\int_{C_R} f(z)\,\mathrm{e}^{\mathrm{i}az}\,\mathrm{d}z \xrightarrow[R\to\infty]{} 0.$$

Demostración. Sobre $z = R\mathrm{e}^{\mathrm{i}\theta}$ se tiene $|\mathrm{e}^{\mathrm{i}az}| = \mathrm{e}^{-aR\sin\theta}$ y $|\mathrm{d}z| = R\,\mathrm{d}\theta$, luego

$$\left|\int_{C_R} f\,\mathrm{e}^{\mathrm{i}az}\mathrm{d}z\right| \le M_R\,R\int_0^{\pi} \mathrm{e}^{-aR\sin\theta}\,\mathrm{d}\theta = 2M_R\,R\int_0^{\pi/2} \mathrm{e}^{-aR\sin\theta}\,\mathrm{d}\theta,$$

por simetría respecto de $\theta = \pi/2$. Ahora se usa la *desigualdad de Jordan*: la función $\theta \mapsto \frac{\sin\theta}{\theta}$ es decreciente en $(0, \pi/2]$, luego $\frac{\sin\theta}{\theta} \ge \frac{\sin(\pi/2)}{\pi/2} = \frac{2}{\pi}$, es decir $\sin\theta \ge \frac{2\theta}{\pi}$ en $[0, \pi/2]$. Por tanto $\mathrm{e}^{-aR\sin\theta} \le \mathrm{e}^{-2aR\theta/\pi}$ y

$$2M_R R\int_0^{\pi/2} \mathrm{e}^{-2aR\theta/\pi}\mathrm{d}\theta = 2M_R R\cdot\frac{\pi}{2aR}\left(1-\mathrm{e}^{-aR}\right) \le \frac{\pi M_R}{a} \xrightarrow[R\to\infty]{} 0. \qquad \square$$

Observación 30.94. Nótese lo que hace el lema: la cota π/a *no depende de R*. Toda la exponencial oscilante del semiplano superior aporta, como mucho, un factor acotado, y basta con que f tienda a cero, por lentamente que sea. Con la estimación ML cruda haría falta que f decayera más rápido que $1/R$; con Jordan basta $f \to 0$. Esa es la diferencia entre poder calcular $\int_0^{\infty} \frac{\sin x}{x}\mathrm{d}x$ y no poder.

Definición 30.95 (Valor principal de Cauchy)**.** Si f es integrable fuera de un entorno de x_0 pero no en un entorno suyo, se define

$$\mathrm{v.p.}\int_a^b f(x)\,\mathrm{d}x = \lim_{\varepsilon\to 0^+}\left[\int_a^{x_0-\varepsilon} f + \int_{x_0+\varepsilon}^{b} f\right],$$

cuando el límite existe. Análogamente en el infinito, v.p. $\int_{-\infty}^{\infty} f = \lim_{R\to\infty} \int_{-R}^{R} f$.

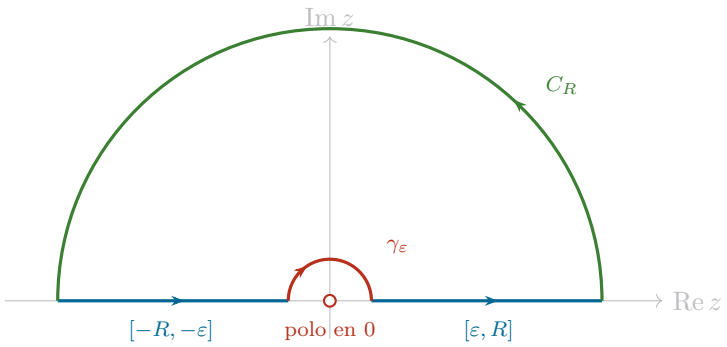


Figura 30.4: El contorno con el que se calcula $\int_0^\infty \frac{\sin x}{x}\mathrm{d}x = \frac{\pi}{2}$. El polo del integrando $\mathrm{e}^{\mathrm{i}z}/z$ está *sobre* el camino, así que se esquiva con un semicírculo pequeño (rojo) que aporta $-\mathrm{i}\pi$ veces el residuo; el arco grande (verde) muere por el lema de Jordan, y los dos segmentos se combinan en $2\mathrm{i}\int \sin x/x$. Dentro del contorno no queda ninguna singularidad, de modo que la suma vale cero y la integral se despeja.

Cuidado

El valor principal *no* es la integral: es un límite simétrico que puede existir cuando la integral no. Por ejemplo v.p. $\int_{-1}^{1} \frac{\mathrm{d}x}{x} = 0$ por simetría, mientras que $\int_{-1}^{1} \frac{\mathrm{d}x}{x}$ no existe ni como integral impropia ni como integral de Lebesgue. Al escribir el resultado hay que decir cuál de las dos cosas se ha calculado.

Lema 30.96 (Semicírculo pequeño)**.** *Sea f con un polo* simple *en x_0 y sea $\gamma_\varepsilon(\theta) = x_0 + \varepsilon \mathrm{e}^{\mathrm{i}\theta}$ con $\theta \in [\theta_1, \theta_2]$. Entonces*

$$\lim_{\varepsilon\to 0^+} \int_{\gamma_\varepsilon} f(z)\,\mathrm{d}z = \mathrm{i}\,(\theta_2 - \theta_1)\ \operatorname{Res}(f, x_0).$$

En particular, un semicírculo recorrido en sentido positivo (θ de 0 a π) aporta $\mathrm{i}\pi \operatorname{Res}$, y recorrido en sentido negativo, $-\mathrm{i}\pi \operatorname{Res}$.

Demostración. Escríbase $f(z) = \frac{c_{-1}}{z-x_0} + g(z)$, donde g es holomorfa en un entorno de x_0 (es la serie de Laurent menos su único término singular) y por tanto acotada, digamos $|g| \le K$, en un disco pequeño. Para el primer sumando, con $z = x_0 + \varepsilon \mathrm{e}^{\mathrm{i}\theta}$ y $\mathrm{d}z = \mathrm{i}\varepsilon \mathrm{e}^{\mathrm{i}\theta}\mathrm{d}\theta$,

$$\int_{\gamma_\varepsilon} \frac{c_{-1}}{z - x_0}\mathrm{d}z = c_{-1}\int_{\theta_1}^{\theta_2} \frac{\mathrm{i}\varepsilon \mathrm{e}^{\mathrm{i}\theta}}{\varepsilon \mathrm{e}^{\mathrm{i}\theta}}\mathrm{d}\theta = \mathrm{i}\,c_{-1}(\theta_2 - \theta_1),$$

independientemente de ε. Para el segundo, $\left|\int_{\gamma_\varepsilon} g\right| \le K \cdot (\theta_2 - \theta_1)\varepsilon \to 0$. □

Cuidado

[Sólo vale para polos simples] Con un polo doble el lema es falso: el término $c_{-2}(z-x_0)^{-2}$ integrado sobre el arco produce un factor $1/\varepsilon$ que diverge. Si al indentar el contorno aparece un polo de orden ≥ 2 sobre el eje real, la integral simplemente no existe, ni siquiera en valor principal.

Ejemplo 30.97 (La integral de Dirichlet, entera)**.** Se calcula $\displaystyle\int_0^\infty \frac{\sin x}{x}\,\mathrm{d}x$. Tómese $f(z)=\frac{\mathrm{e}^{\mathrm{i}z}}{z}$ y el contorno indentado: el segmento $[-R,-\varepsilon]$, el semicírculo pequeño γ_ε recorrido *en sentido negativo* (de π a 0, para esquivar el origen por arriba), el segmento $[\varepsilon,R]$ y el arco grande C_R. Dentro del contorno no hay ninguna singularidad —el único polo, $z=0$, ha quedado fuera por la indentación—, así que por el teorema de Cauchy la integral total vale 0:

$$\int_{-R}^{-\varepsilon}\frac{\mathrm{e}^{\mathrm{i}x}}{x}\mathrm{d}x+\int_{\gamma_\varepsilon^-}f+\int_{\varepsilon}^{R}\frac{\mathrm{e}^{\mathrm{i}x}}{x}\mathrm{d}x+\int_{C_R}f=0.$$

Por el lema 30.93 con $f(z)=1/z$ y $a=1$, el último término tiende a 0. Por el lema 30.96, y siendo $\operatorname{Res}(f,0)=1$, el segundo tiende a $-\mathrm{i}\pi$. Y los dos segmentos se combinan sustituyendo $x\mapsto -x$ en el primero:

$$\int_{\varepsilon}^{R}\frac{\mathrm{e}^{\mathrm{i}x}-\mathrm{e}^{-\mathrm{i}x}}{x}\mathrm{d}x=2\mathrm{i}\int_{\varepsilon}^{R}\frac{\sin x}{x}\mathrm{d}x.$$

Pasando al límite: $2\mathrm{i}\int_0^\infty \frac{\sin x}{x}\mathrm{d}x-\mathrm{i}\pi=0$, es decir

$$\int_0^\infty \frac{\sin x}{x}\,\mathrm{d}x=\frac{\pi}{2}.$$

Numéricamente, $\int_0^{2000}\frac{\sin x}{x}\mathrm{d}x=1{,}570\,98$ frente a $\pi/2=1{,}570\,80$: la convergencia es lenta —la integral es condicionalmente convergente— pero el valor es el correcto.

Ejemplo 30.98 (Un contorno de ojo de cerradura)**.** Para $0<a<1$ se calcula $\displaystyle\int_0^\infty \frac{x^{a-1}}{1+x}\mathrm{d}x$. Se toma $f(z)=\frac{z^{a-1}}{1+z}$ con la rama de $z^{a-1}=\mathrm{e}^{(a-1)\log z}$ y el logaritmo definido con argumento en $(0,2\pi)$, es decir, cortado a lo largo del semieje real positivo. El contorno es un *ojo de cerradura*: rodea el corte por arriba de ε a R, da la vuelta grande, vuelve por debajo del corte de R a ε y cierra con el círculo pequeño. Al volver por debajo, el argumento ha aumentado en 2π, de modo que allí $z^{a-1}=x^{a-1}\mathrm{e}^{2\pi\mathrm{i}(a-1)}$. Los dos círculos aportan 0 (por ML, usando $0<a<1$ en los dos extremos), y el único polo encerrado es $z=-1=\mathrm{e}^{\mathrm{i}\pi}$, con residuo $\mathrm{e}^{\mathrm{i}\pi(a-1)}$. Así:

$$\left(1-\mathrm{e}^{2\pi\mathrm{i}(a-1)}\right)\int_0^\infty\frac{x^{a-1}}{1+x}\mathrm{d}x=2\pi\mathrm{i}\,\mathrm{e}^{\mathrm{i}\pi(a-1)},$$

y dividiendo entre $\mathrm{e}^{\mathrm{i}\pi(a-1)}$ el miembro izquierdo se convierte en $-2\mathrm{i}\sin(\pi(a-1))\int = 2\mathrm{i}\sin(\pi a)\int$, de donde

$$\int_0^\infty \frac{x^{a-1}}{1+x}\,\mathrm{d}x=\frac{\pi}{\sin(\pi a)}.$$

Comprobación numérica: para $a=\frac12$ da $3{,}141\,59=\pi$; para $a=0{,}3$ da $3{,}883\,22$; para $a=0{,}75$ da $4{,}442\,88$, y los tres coinciden con $\pi/\sin(\pi a)$ hasta la sexta cifra. De paso, esta fórmula *es* la

de reflexión de Euler $\Gamma(a)\Gamma(1-a) = \pi/\sin(\pi a)$ disfrazada, porque la integral de la izquierda vale $B(a, 1-a) = \Gamma(a)\Gamma(1-a)$.

Regla: Elegir el contorno según el integrando

1. **Racional par,** $\deg q \geq \deg p + 2$**:** semicírculo superior; el arco muere por ML.
2. **Con** $\sin(ax)$ **o** $\cos(ax)$**:** escríbase como parte imaginaria o real de $\mathrm{e}^{\mathrm{i}az}$ y úsese el lema de Jordan. *Nunca* se pone $\sin z$ en el integrando complejo: $|\sin z|$ crece exponencialmente fuera del eje real.
3. **Con un polo sobre el eje real:** indéntese y aplíquese el lema del semicírculo pequeño; el resultado es un valor principal.
4. **Con** x^a **o** $\log x$**:** ojo de cerradura con corte en el semieje positivo.
5. **Racional en** $\sin\theta, \cos\theta$ **sobre** $[0, 2\pi]$**:** sustitúyase $z = \mathrm{e}^{\mathrm{i}\theta}$, con $\cos\theta = \frac{z+z^{-1}}{2}$, $\sin\theta = \frac{z-z^{-1}}{2\mathrm{i}}$ y $\mathrm{d}\theta = \frac{\mathrm{d}z}{\mathrm{i}z}$: la integral se convierte en una sobre la circunferencia unidad.
6. **Con** e^x **periódico:** rectángulo de altura $2\pi\mathrm{i}$ o $\pi\mathrm{i}$, aprovechando que los lados horizontales dan el mismo integrando salvo constante.

30.13. Entera, meromorfa y familia normal: las tres palabras que faltaban

Definición 30.99 (Función entera y función meromorfa)**.** Una función f es *entera* si es holomorfa en *todo* el plano complejo $\mathbb{C}$.

Es *meromorfa* en un abierto Ω si es holomorfa en Ω salvo en un conjunto de puntos aislados, y en cada uno de ellos tiene un *polo* —no una singularidad esencial—. Equivalentemente, si localmente es cociente de dos holomorfas con denominador no idénticamente nulo.

Ejemplo 30.100 (Quién es qué)**.** Son enteras: los polinomios, e^z, $\sin z$, $\cos z$ y la función de Bessel J_0. No lo es $1/z$, que falla en el origen; ni $\sqrt{z}$ ni $\log z$, que ni siquiera son funciones uniformes.

Son meromorfas en $\mathbb{C}$: toda función racional p/q, con polos en los ceros de q; también $\tan z = \sin z/\cos z$, con polos simples en $\frac{\pi}{2} + k\pi$; y $1/\sin z$, con polos en $k\pi$. La función Γ es meromorfa con polos simples en $0, -1, -2, \ldots$ En cambio $\mathrm{e}^{1/z}$ *no* es meromorfa en $\mathbb{C}$: su singularidad en 0 es esencial. Y $1/\sin(1/z)$ tampoco, porque sus polos se acumulan en el origen y dejan de estar aislados.

Observación 30.101 (Las meromorfas son un cuerpo, y eso lo explica todo). Las funciones meromorfas en un dominio conexo forman un *cuerpo*: se pueden sumar, multiplicar y dividir (por una no idénticamente nula) sin salirse de la clase. Las holomorfas, en cambio, sólo forman un anillo, porque $1/f$ deja de ser holomorfa donde f se anula. Esa es la razón por la que el principio del argumento y el teorema de Rouché se enuncian para meromorfas: la cantidad

que cuentan, $\frac{1}{2\pi \mathrm{i}} \oint \frac{f'}{f} = Z - P$, sólo tiene sentido si se admiten polos, y su naturaleza es la de un cociente.

Definición 30.102 (Familia normal)**.** Sea $\mathcal{F}$ una familia de funciones holomorfas en un abierto Ω. Se dice que $\mathcal{F}$ es *normal* si toda sucesión de elementos de $\mathcal{F}$ admite una subsucesión que converge *uniformemente sobre los compactos* de Ω (se permite que el límite sea la constante ∞ en la versión de Montel para meromorfas).

Teorema 30.103 (Montel)**.** *Toda familia de funciones holomorfas* uniformemente acotada *sobre los compactos de Ω es normal.*

Demostración. Sea $K \subset \Omega$ compacto y K' un compacto ligeramente mayor con K en su interior, a distancia $r > 0$ de $\partial K'$. Por hipótesis $|f| \leq M$ en K' para toda $f \in \mathcal{F}$, y la estimación de Cauchy da, para $z \in K$,

$$|f'(z)| = \left| \frac{1}{2\pi \mathrm{i}} \oint_{|w-z|=r} \frac{f(w)}{(w-z)^2} \mathrm{d}w \right| \leq \frac{M}{r}.$$

Luego la familia es *equicontinua* sobre K: por el teorema del valor medio, $|f(z_1) - f(z_2)| \leq \frac{M}{r}|z_1 - z_2|$ con constante independiente de f. Equicontinua y uniformemente acotada, el teorema de Arzelà–Ascoli (capítulo 12) da una subsucesión uniformemente convergente en K; un argumento diagonal sobre una sucesión de compactos que agote Ω produce una subsucesión que sirve en todos a la vez. □

Observación 30.104. Ésta es la compacidad que hace funcionar la demostración del teorema de la aplicación de Riemann: permite extraer, de una sucesión maximizante de $|f'(0)|$, una función límite que sigue siendo holomorfa e inyectiva. Es el análogo de Bolzano–Weierstrass en un espacio de funciones, y la razón de que se le exija a la familia estar acotada: sin esa hipótesis, la sucesión $f_n(z) = nz$ no tiene ninguna subsucesión convergente.

30.14. Ejercicios resueltos

Ejercicio 30.1. Verifica Cauchy–Riemann para $f(z) = z^2$ y calcula f' de tres maneras.

Solución. $z^2 = (x^2 - y^2) + i(2xy)$: $u_x = 2x = v_y$ ✓, $u_y = -2y = -v_x$ ✓. Derivada: por la definición, $\lim \frac{(z+h)^2 - z^2}{h} = 2z$; por $f' = u_x + iv_x = 2x + 2iy = 2z$; por la regla formal $\frac{\mathrm{d}}{\mathrm{d}z} z^2 = 2z$: consistente —las reglas del capítulo 6 valen verbatim, esa es la gracia—. ■

Ejercicio 30.2. Calcula $\oint_{|z|=2} \frac{e^z}{z(z-1)}\, \mathrm{d}z$.

Solución. Dos polos simples dentro. $\mathrm{Res}_0 = \frac{e^0}{0-1} = -1$; $\mathrm{Res}_1 = \frac{e^1}{1} = e$. Total: $2\pi i(e-1)$. ■

Ejercicio 30.3. Calcula $\displaystyle\int_0^{2\pi} \frac{\mathrm{d}\theta}{5 + 4\cos\theta}$ por residuos.

Solución. Con $z = e^{i\theta}$: $\cos\theta = \frac{z+z^{-1}}{2}$, $\mathrm{d}\theta = \frac{\mathrm{d}z}{iz}$:

$$\oint_{|z|=1} \frac{\mathrm{d}z}{iz\big(5 + 2z + 2z^{-1}\big)} = \oint \frac{\mathrm{d}z}{i(2z^2 + 5z + 2)} = \oint \frac{\mathrm{d}z}{2i(z+2)(z+\frac{1}{2})}.$$

Solo $z=-\frac{1}{2}$ está dentro: residuo $\frac{1}{2i(-\frac{1}{2}+2)}=\frac{1}{3i}$. Total: $2\pi i\cdot\frac{1}{3i}=\frac{2\pi}{3}$. (Toda integral trigonométrica racional se rinde ante este cambio: es el $t=\tan\frac{\theta}{2}$ del capítulo 6, profesionalizado.) ■

Ejercicio 30.4. Demuestra con Liouville que una entera con $|f(z)|\le C|z|^n$ es un polinomio de grado $\le n$.

Solución. Por las estimaciones de Cauchy en el círculo de radio R: $|f^{(n+1)}(z_0)|\le\frac{(n+1)!\,CR^n(1+|z_0|/R)^n}{R^{n+1}}\to 0$ cuando $R\to\infty$: $f^{(n+1)}\equiv 0$, y una función con derivada $(n+1)$-ésima nula es un polinomio de grado $\le n$ (Taylor). Con $n=0$ es Liouville; el crecimiento controla el grado: la rigidez holomorfa en estado puro. ■

Ejercicio 30.5. Usa el mapa $w=z^2$ para resolver el problema de potencial en el primer cuadrante con bordes a potenciales 0 (eje x) y 1 (eje y).

Solución. z^2 manda el cuadrante al semiplano superior, con el eje x positivo del semiplano viniendo del borde a potencial 0 y el negativo del borde a potencial 1. En el semiplano, la solución es el ángulo normalizado: $u=\frac{1}{\pi}\arg w$. Tirando atrás: $\arg(z^2)=2\arg z$:

$$u(z)=\frac{2}{\pi}\arg z=\frac{2}{\pi}\arctan\frac{y}{x}\,:$$

armónica (parte imaginaria de $\frac{2}{\pi}\ln z$), con los valores de borde correctos. Mapear, resolver en el dominio fácil, regresar: el método conforme completo en miniatura. ■

30.15. Taller de ejemplos explicados: números complejos y un residuo

Idea clave

Intuición primero. Un número complejo $a+bi$ es un punto del plano, y operar con ellos es geometría disfrazada de álgebra (multiplicar rota y escala). El *residuo* es el truco estrella: al integrar alrededor de un polo, todo se cancela salvo el coeficiente de $\frac{1}{z}$, así que una integral real endiablada se reduce a leer ese único número y multiplicar por $2\pi i$. En el ejemplo lo verás resolver en tres líneas una integral que a mano sería una pesadilla.

30.15.1. Aritmética compleja

Un número complejo es $a+bi$ con $i^2=-1$. Se opera como con binomios, recordando reemplazar i^2 por -1.

Ejemplo 30.105 (Multiplicar)**.** Calcular $(3+2i)(1-i)$.

Paso 1. Multiplicar como binomios: $(3)(1)+(3)(-i)+(2i)(1)+(2i)(-i)=3-3i+2i-2i^2$.

Paso 2. Reemplazar $i^2=-1$: el término $-2i^2$ se vuelve $-2(-1)=+2$. Entonces

$$3-3i+2i+2=(3+2)+(-3i+2i)=5-i.$$

Ejemplo 30.106 (Módulo y división)**.** El *módulo* (tamaño) de $3+4i$ es $|3+4i|=\sqrt{3^2+4^2}=\sqrt{25}=5$.

Para dividir $\dfrac{1}{1+i}$, multiplicamos arriba y abajo por el conjugado $1-i$:

$$\frac{1}{1+i}\cdot\frac{1-i}{1-i}=\frac{1-i}{1^2-i^2}=\frac{1-i}{1+1}=\frac{1-i}{2}=\tfrac{1}{2}-\tfrac{1}{2}i.$$

30.15.2. Una integral real resuelta por residuos

Uno de los milagros de la variable compleja: integrales reales difíciles se resuelven cerrando un contorno en el plano complejo y mirando los "residuos" en los polos.

Ejemplo 30.107. Calcular $\displaystyle\int_{-\infty}^{\infty}\frac{1}{1+x^2}\,\mathrm{d}x$.

Paso 1. Localizar los polos. La función $\dfrac{1}{1+z^2}=\dfrac{1}{(z-i)(z+i)}$ tiene polos en $z=i$ y $z=-i$. Al cerrar el contorno por la mitad superior del plano, solo encerramos $z=i$.

Paso 2. Calcular el residuo en $z=i$. Para un polo simple, el residuo es

$$\operatorname{Res}_{z=i}\frac{1}{(z-i)(z+i)}=\frac{1}{z+i}\Big|_{z=i}=\frac{1}{i+i}=\frac{1}{2i}.$$

Paso 3. Aplicar el teorema de los residuos: la integral vale $2\pi i$ por la suma de los residuos encerrados:

$$\int_{-\infty}^{\infty}\frac{\mathrm{d}x}{1+x^2}=2\pi i\cdot\frac{1}{2i}=\pi.$$

El resultado es π (que también sale con arctan, pero por residuos fue casi automático).

30.16. Ejercicios propuestos

Sin solución desarrollada: solo enunciado y, al final, la respuesta. En esta parte del libro muchas respuestas son una frase ("sí, y la razón es…") más que un número; comprueba que la tuya coincide en el fondo, no en la letra.

30.1. ¿Es $f(z)=\bar{z}$ holomorfa en algún punto? Verifica con Cauchy–Riemann.

30.2. Halla u_y y v_x para $f(z)=z^2$ y comprueba las ecuaciones de Cauchy–Riemann.

30.3. Calcula $\displaystyle\oint_{|z|=2}\frac{\mathrm{d}z}{z-1}$.

30.4. Calcula $\operatorname{Res}\left(\frac{1}{z^2+1},\, z=i\right)$.

30.5. Calcula $\operatorname{Res}\left(\frac{e^z}{z^3},\, z=0\right)$.

30.6. Clasifica la singularidad de $\frac{\sin z}{z}$ en $z=0$ y la de $e^{1/z}$ en $z=0$.

30.7. Da los tres primeros términos de la serie de Laurent de $\frac{1}{z(z-1)}$ alrededor de $z=0$ en $0<|z|<1$.

30.8. Calcula $\displaystyle\int_{-\infty}^{\infty}\frac{\mathrm{d}x}{1+x^4}$ por residuos.

30.9. Calcula $\displaystyle\int_{-\infty}^{\infty}\frac{\cos x}{1+x^2}\,\mathrm{d}x$.

30.10. Calcula $\displaystyle\int_0^{2\pi} \frac{\mathrm{d}\theta}{2+\cos\theta}$ con la sustitución $z = e^{i\theta}$.

30.11. ¿Cuántos ceros tiene $z^4 + 5z + 1$ dentro de $|z| < 1$? Usa Rouché.

30.12. Si f es entera y acotada, ¿qué puedes concluir? ¿Y si $|f|$ alcanza un máximo interior en un dominio?

Respuestas

1. En ninguno: $u_x = 1$ pero $v_y = -1$. **2.** $u = x^2 - y^2$, $v = 2xy$; $u_x = 2x = v_y$ y $u_y = -2y = -v_x$. Se cumplen en todo $\mathbb{C}$. **3.** $2\pi i$. **4.** $-i/2$ (equivalentemente $1/(2i)$). **5.** $1/2$. **6.** $\sin z/z$: singularidad evitable (límite 1). $e^{1/z}$: singularidad esencial. **7.** $-\dfrac{1}{z} - 1 - z - \cdots$ **8.** $\dfrac{\pi\sqrt{2}}{2} = \dfrac{\pi}{\sqrt{2}}$. **9.** π/e. **10.** $\dfrac{2\pi}{\sqrt{3}} = \dfrac{2\sqrt{3}\,\pi}{3}$. **11.** Uno. Con $g(z) = 5z$ en $|z| = 1$: $|z^4 + 1| \le 2 < 5 = |g|$, y g tiene un cero dentro. **12.** Liouville: f es constante. Y por el principio del módulo máximo, si $|f|$ alcanza un máximo interior, f es constante en la componente.

Capítulo 31

La integral de Lebesgue y la transformada de Fourier

Lo que necesitas traer

Del capítulo 29: supremos, sucesiones de funciones, convergencia uniforme y las patologías de la integral de Riemann. Del capítulo 28: abiertos y σ-álgebras. Del capítulo 16: nociones de probabilidad, que aquí se reinterpretan como teoría de la medida.

Este libro prometió varias veces que "la primera parada del análisis posterior" sería la integral de Lebesgue. Cumplimos la promesa: este capítulo construye la medida y la integral de Lebesgue, demuestra los tres teoremas de convergencia que hacen funcionar todo el análisis moderno, y con esa maquinaria desarrolla la transformada de Fourier *con demostraciones*: la inversión, Plancherel y el principio de incertidumbre. Es el punto exacto donde el cálculo del capítulo 6 se convierte en análisis.

31.1. Por qué Riemann no basta

Idea clave

Intuición primero. La diferencia entre Riemann y Lebesgue se explica con una imagen que se atribuye al propio Lebesgue: para contar el dinero de un montón de monedas, Riemann las va tomando en el orden en que aparecen y suma; Lebesgue primero las agrupa por valor —todas las de un quetzal juntas, todas las de cinco juntas— y luego multiplica valor por cantidad. Riemann corta el *dominio* en trozos; Lebesgue corta el *recorrido* y pregunta cuánto mide el conjunto donde la función toma cada valor.

El cambio parece cosmético y no lo es. Al cortar por el recorrido, ya no importa lo desordenada que esté la función en el eje horizontal, solo cuánto "mide" cada conjunto de nivel; por eso la función de Dirichlet, que a Riemann le resulta imposible, para Lebesgue es trivial. Pero el precio es que hay que construir antes una teoría de la medida capaz de asignarle un tamaño a conjuntos muy feos, y a eso se dedica la sección siguiente. La ganancia real, sin embargo, no son las funciones patológicas: son los teoremas de

convergencia, que permiten intercambiar límite e integral con hipótesis suaves, y sin los cuales el análisis moderno no existiría.

La integral de Riemann nos llevó lejos, pero tiene tres grietas estructurales.

Ejemplo 31.1 (La función de Dirichlet)**.** $D(x) = 1$ si $x \in \mathbb{Q}$, 0 si no. En cualquier subintervalo hay racionales e irracionales: toda suma superior vale 1 y toda inferior 0: *no es integrable Riemann*. Y sin embargo es un límite puntual de funciones integrables: enumerando $\mathbb{Q} \cap [0,1] = \{r_1, r_2, \dots\}$, las $D_n = \mathbf{1}_{\{r_1,\dots,r_n\}}$ son escalonadas con integral 0 y $D_n \to D$ puntualmente. *La clase de Riemann no es cerrada bajo límites puntuales*: el análisis, que vive de tomar límites, trabaja en una casa sin techo.

Las otras dos grietas: los teoremas de intercambio $\lim \int = \int \lim$ de Riemann exigen convergencia *uniforme* (capítulo 6), una hipótesis frecuentemente falsa e incómoda de verificar; y el espacio de funciones integrables con la norma $\int |f - g|$ *no es completo*: hay sucesiones de Cauchy sin límite integrable, como $\mathbb{Q}$ visto desde los racionales. La solución de Lebesgue (1902) a las tres grietas es la misma idea: en vez de partir el *dominio* en intervalitos (Riemann), partir el *rango* en franjas horizontales y preguntar "¿cuánto mide el conjunto donde f está en esta franja?". Todo se reduce entonces a saber *medir conjuntos*.

31.2. Qué es medir: sigma-álgebras, medidas, funciones medibles y la integral en tres pasos

La sección anterior ha explicado por qué Riemann no basta. Antes de construir la medida de Lebesgue conviene tener escrito, con precisión y desde cero, el vocabulario que se va a usar sin descanso durante todo el capítulo: qué es una familia de conjuntos medibles, qué es una medida, qué significa "en casi todo punto", qué es una función medible y cómo se define la integral en los tres pasos que la construyen. Nada de esto exige conocer todavía la medida de Lebesgue: son definiciones abstractas que después se instancian en ella, y tenerlas primero evita tener que adivinar el significado de las palabras a mitad de una demostración.

Idea clave

Intuición primero. Medir es asignar un número a un conjunto. La primera sorpresa es que no se puede medir todo: hay conjuntos tan retorcidos que cualquier asignación razonable de tamaños entra en contradicción consigo misma. La solución no es rendirse, sino decidir de antemano *qué colección de conjuntos* se va a medir, y exigirle a esa colección que sea estable bajo las operaciones que uno hace en análisis: complementar y unir numerablemente muchas veces. Esa colección es una σ-álgebra, y una medida es una función que le asigna a cada uno de sus miembros un número no negativo, de forma que trocear un conjunto en una cantidad *numerable* de pedazos disjuntos y sumar dé el mismo resultado. Todo el capítulo es la historia de esa palabra: numerable.

Sigma-álgebras

Definición 31.2 (σ-álgebra)**.** Sea X un conjunto. Una familia $\mathcal{M}$ de subconjuntos de X es una *σ-álgebra* sobre X si cumple tres condiciones:

(i) $X \in \mathcal{M}$;

(ii) si $E \in \mathcal{M}$, entonces $E^c = X \setminus E \in \mathcal{M}$ (estable por complementos);

(iii) si $E_1, E_2, E_3, \dots \in \mathcal{M}$, entonces $\bigcup_{n\geq 1} E_n \in \mathcal{M}$ (estable por uniones *numerables*).

Los elementos de $\mathcal{M}$ se llaman *conjuntos medibles*, y el par $(X, \mathcal{M})$ es un *espacio medible*.

Proposición 31.3 (Lo que sale gratis de los tres axiomas)**.** *Toda σ-álgebra contiene $\varnothing$ y es estable por intersecciones numerables y por diferencias.*

Demostración. $\varnothing = X^c \in \mathcal{M}$ por (i) y (ii). Para intersecciones se usa la ley de De Morgan: si todos los E_n son medibles, entonces

$$\bigcap_{n\geq 1} E_n = \Big(\bigcup_{n\geq 1} E_n^c \Big)^c,$$

y el segundo miembro es medible por (ii), (iii) y (ii) otra vez. Para la diferencia, $E \setminus F = E \cap F^c$, que es intersección de dos medibles. □

Ejemplo 31.4 (Cuatro σ-álgebras, y una familia que no lo es)**. (a) La trivial y la total.** $\{\varnothing, X\}$ es la menor σ-álgebra sobre X, y el conjunto de partes $\mathcal{P}(X)$ es la mayor. Cualquier otra está entre ambas.

(b) La generada por un conjunto. Sobre $X = \{1, 2, 3\}$, la menor σ-álgebra que contiene a $\{1\}$ es $\{\varnothing, \{1\}, \{2, 3\}, X\}$: exactamente cuatro elementos, y se comprueba a mano que es estable por complementos y uniones.

(c) Los borelianos. Sobre $\mathbb{R}$, la menor σ-álgebra que contiene a todos los intervalos abiertos se llama σ-álgebra de Borel, y sus miembros son los *borelianos*. Contiene todos los abiertos (unión numerable de intervalos con extremos racionales), todos los cerrados, todas las intersecciones numerables de abiertos y así indefinidamente. Que exista la "menor" es un hecho general: la intersección de cualquier familia de σ-álgebras vuelve a serlo.

(d) Una familia que no es σ-álgebra. La familia de las uniones *finitas* de intervalos de $\mathbb{R}$ es estable por complementos y uniones finitas, pero no por uniones numerables: uniendo los intervalos $\left(q - \varepsilon 2^{-k}, q + \varepsilon 2^{-k}\right)$ centrados en los racionales se obtiene un abierto que no es unión finita de intervalos. Esa familia es un *álgebra*, no una σ-álgebra, y la diferencia entre ambas palabras es la letra sigma, es decir, lo numerable.

Regla: comprobar que una familia de conjuntos es una sigma-álgebra

1. **Verifica los tres axiomas literalmente y en orden.** El total, los complementos, las uniones numerables. Son tres comprobaciones y ninguna se puede saltar.

2. **No demuestres por separado lo que ya es consecuencia.** El vacío, las intersecciones numerables y las diferencias salen gratis por la proposición 31.3; incluirlas en la lista de axiomas es redundante.

3. **Si la familia se describe por una propiedad, comprueba que la propiedad sobrevive a las tres operaciones.** Ejemplo típico: "los E tales que E o E^c es a lo sumo numerable" es una σ-álgebra sobre cualquier X, porque una unión numerable de numerables es numerable.

4. **Si solo necesitas la menor que contiene a cierta familia $\mathcal{E}$, no la construyas.** Defínela como la intersección de todas las σ-álgebras que contienen a $\mathcal{E}$; esa intersección es una σ-álgebra y por definición es la menor. Intentar describir sus elementos uno a uno es, en general, imposible.

5. **Sospecha en cuanto veas la palabra "finito".** Estable por uniones finitas no es estable por uniones numerables, y ahí es donde fallan casi todos los candidatos.

6. **Control.** Si tu familia contiene a los intervalos y es una σ-álgebra, entonces contiene a todos los borelianos. Si crees haber construido una que contiene los intervalos y no contiene, por ejemplo, al conjunto de Cantor, hay un error.

Medidas

Definición 31.5 (Medida y espacio de medida)**.** Sea $(X, \mathcal{M})$ un espacio medible. Una *medida* sobre él es una función $\mu : \mathcal{M} \to [0, +\infty]$ tal que

(I) $\mu(\varnothing) = 0$;

(II) (σ-*aditividad*) si $E_1, E_2, \dots \in \mathcal{M}$ son disjuntos dos a dos, entonces

$$\mu\Big(\bigsqcup_{n\geq 1} E_n\Big) = \sum_{n\geq 1} \mu(E_n).$$

La terna $(X, \mathcal{M}, \mu)$ se llama *espacio de medida*. Se admite el valor $+\infty$, y la serie del miembro derecho es de términos no negativos, así que converge o vale $+\infty$, sin ambigüedad de orden.

Proposición 31.6 (Las tres propiedades que se usan a diario)**.** *Toda medida es* monótona *($E \subseteq F$ implica $\mu(E) \leq \mu(F)$),* numerablemente subaditiva *($\mu(\bigcup E_n) \leq \sum \mu(E_n)$ sin hipótesis de disjunción) y* continua por abajo *(si $E_1 \subseteq E_2 \subseteq \cdots$ entonces $\mu(\bigcup E_n) = \lim \mu(E_n)$).*

Demostración. Monotonía: $F = E \sqcup (F \setminus E)$ con los dos trozos medibles y disjuntos, luego $\mu(F) = \mu(E) + \mu(F \setminus E) \geq \mu(E)$. Subaditividad: se *disjuntan* los conjuntos poniendo $F_1 = E_1$ y $F_n = E_n \setminus (E_1 \cup \cdots \cup E_{n-1})$, que son medibles por la proposición 31.3, disjuntos, con la misma unión y con $F_n \subseteq E_n$; entonces $\mu(\bigcup E_n) = \sum \mu(F_n) \leq \sum \mu(E_n)$. Continuidad por abajo: con los mismos F_n, $\bigcup_{k\leq n} E_k = E_n$ y $\mu(\bigcup E_n) = \sum_{k\geq 1} \mu(F_k) = \lim_n \sum_{k\leq n} \mu(F_k) = \lim_n \mu(E_n)$. □

Ejemplo 31.7 (Tres medidas que conviene tener en la cabeza)**. (a) La medida de contar.** Sobre cualquier X con $\mathcal{M} = \mathcal{P}(X)$, se define $\mu(E)$ como el número de elementos de E, y $+\infty$ si

es infinito. Es una medida, y con ella la integral se convierte en una suma: $\int f\,\mathrm{d}\mu = \sum_{x\in X} f(x)$. Toda la teoría de series es un caso particular de este capítulo.

(b) La delta de Dirac. Fijado $x_0 \in X$, se pone $\delta_{x_0}(E) = 1$ si $x_0 \in E$ y 0 si no. Es una medida, y $\int f\,\mathrm{d}\delta_{x_0} = f(x_0)$: la "función delta" de los físicos es, correctamente entendida, una medida y no una función.

(c) La medida de Lebesgue. Es la que la sección siguiente construye: sobre los medibles de $\mathbb{R}$, asigna a cada intervalo su longitud, $m((a,b)) = b-a$. La construcción ocupa una sección entera precisamente porque no es evidente que exista.

Conjuntos nulos y "en casi todo punto"

Definición 31.8 (Conjunto de medida nula)**.** En un espacio de medida $(X, \mathcal{M}, \mu)$, un conjunto $N \in \mathcal{M}$ es *nulo*, o *de medida cero*, si $\mu(N) = 0$. En el caso concreto de la recta con la medida de Lebesgue, esto equivale a la condición elemental siguiente: para todo $\varepsilon > 0$ existe una familia *numerable* de intervalos abiertos $(I_k)_{k\geq 1}$ con $N \subseteq \bigcup_k I_k$ y $\sum_k |I_k| < \varepsilon$.

Teorema 31.9 (Todo conjunto numerable tiene medida nula)**.** *Si $A = \{a_1, a_2, a_3, \dots\} \subseteq \mathbb{R}$ es a lo sumo numerable, entonces A tiene medida de Lebesgue cero. En particular $m(\mathbb{Q}) = 0$.*

Demostración. Sea $\varepsilon > 0$. La idea es repartir el ε entre los infinitos puntos usando una serie geométrica, de modo que la suma total quede controlada. Cubrimos el punto a_k con el intervalo abierto

$$I_k = \Big(a_k - \frac{\varepsilon}{2^{k+2}},\ a_k + \frac{\varepsilon}{2^{k+2}}\Big), \qquad \text{de longitud} \quad |I_k| = \frac{\varepsilon}{2^{k+1}}.$$

Evidentemente $A \subseteq \bigcup_{k\geq 1} I_k$, porque cada a_k está en su propio intervalo. Y la longitud total es

$$\sum_{k\geq 1} |I_k| = \sum_{k\geq 1} \frac{\varepsilon}{2^{k+1}} = \frac{\varepsilon}{2}\sum_{k\geq 1}\frac{1}{2^k} = \frac{\varepsilon}{2}\cdot 1 = \frac{\varepsilon}{2} < \varepsilon.$$

Como ε era arbitrario, la medida exterior de A es cero, y un conjunto de medida exterior nula es medible con medida nula. Con $\varepsilon = 0{,}1$, por ejemplo, los cuatro primeros intervalos miden 0,05; 0,025; 0,0125 y 0,00625, y la serie entera suma 0,1. Cada punto recibe su ración y la suma sigue siendo pequeña: eso es todo el argumento. □

Cuidado

La numerabilidad es esencial y no un tecnicismo. Si se permitiesen recubrimientos arbitrarios, *todo* conjunto tendría medida cero, porque cada punto se cubre con un intervalo tan corto como se quiera y $[0,1]$ es la unión de sus puntos. Lo que hace que la definición diga algo es que las longitudes se suman en una *serie*. Y el recíproco del teorema 31.9 es falso: el conjunto de Cantor es no numerable y tiene medida cero, como se comprueba más abajo en este capítulo.

Definición 31.10 (En casi todo punto)**.** Una propiedad se cumple *en casi todo punto* —abreviado c.t.p., y en probabilidad "casi seguramente"— si el conjunto de puntos donde falla está contenido en un conjunto de medida nula. Así, "$f = g$ c.t.p." significa $\mu(\{x : f(x) \neq$

$g(x)\}\big) = 0$, y "$f_n \to f$ c.t.p." significa que el conjunto de puntos donde la convergencia falla es nulo. La integral no distingue funciones iguales c.t.p., y por eso esta expresión aparece en casi todos los enunciados del capítulo.

El criterio de Carathéodory: la unión numerable de medibles es medible

La sección siguiente introduce la medida exterior m^* y declara medible a todo conjunto E que "corte limpio": $m^*(A) = m^*(A\cap E) + m^*(A\setminus E)$ para todo $A \subseteq \mathbb{R}$. Allí se da la idea de la demostración de que esos conjuntos forman una σ-álgebra; aquí está la demostración entera, porque es el único punto técnico serio de toda la construcción y merece verse una vez.

Teorema 31.11 (Los medibles de Carathéodory forman una σ-álgebra)**.** *Sea m^* una medida exterior, es decir, una función monótona y numerablemente subaditiva con $m^*(\varnothing) = 0$. La familia $\mathcal{M}$ de los conjuntos E que cumplen el criterio de Carathéodory es una σ-álgebra, y m^* restringida a ella es σ-aditiva.*

Demostración. Paso 1: complementos. La condición $m^*(A) = m^*(A\cap E) + m^*(A\setminus E)$ es simétrica al cambiar E por E^c, porque $A\cap E^c = A\setminus E$ y $A\setminus E^c = A\cap E$. Luego $E\in\mathcal{M}$ implica $E^c\in\mathcal{M}$. Y $X\in\mathcal{M}$ trivialmente.

Paso 2: unión de dos. Sean $E_1, E_2 \in \mathcal{M}$ y A cualquiera. Aplicando el criterio con E_1 y luego, al trozo $A\setminus E_1$, con E_2:

$$m^*(A) = m^*(A\cap E_1) + m^*\big((A\setminus E_1)\cap E_2\big) + m^*\big(A\setminus E_1\setminus E_2\big).$$

Ahora $A\cap(E_1\cup E_2) = (A\cap E_1)\cup\big((A\setminus E_1)\cap E_2\big)$, así que por subaditividad la suma de los dos primeros términos es $\geq m^*\big(A\cap(E_1\cup E_2)\big)$; y el tercero es exactamente $m^*\big(A\setminus(E_1\cup E_2)\big)$. Por tanto $m^*(A)\geq m^*(A\cap(E_1\cup E_2)) + m^*(A\setminus(E_1\cup E_2))$, y la desigualdad contraria es subaditividad. Luego $E_1\cup E_2\in\mathcal{M}$, y por inducción también toda unión finita.

Paso 3: aditividad sobre un conjunto de prueba. Si $E_1,\dots,E_n\in\mathcal{M}$ son disjuntos, entonces para todo A

$$m^*\Big(A\cap \bigsqcup_{k\leq n} E_k\Big) = \sum_{k\leq n} m^*(A\cap E_k).$$

Se demuestra por inducción: aplicando el criterio de E_n al conjunto $A\cap\bigsqcup_{k\leq n}E_k$, la parte dentro de E_n es $A\cap E_n$ y la parte fuera es $A\cap\bigsqcup_{k\leq n-1}E_k$, por ser disjuntos.

Paso 4: unión numerable. Sean $E_1, E_2, \dots\in\mathcal{M}$ y llamemos $E = \bigcup_n E_n$. Podemos suponerlos disjuntos, sustituyendo E_n por $E_n\setminus(E_1\cup\dots\cup E_{n-1})$, que sigue en $\mathcal{M}$ por los pasos 1 y 2. Sea $F_n = \bigsqcup_{k\leq n}E_k\in\mathcal{M}$. Para todo A,

$$m^*(A) = m^*(A\cap F_n) + m^*(A\setminus F_n) \;\geq\; \sum_{k\leq n} m^*(A\cap E_k) + m^*(A\setminus E),$$

donde se ha usado el paso 3 en el primer sumando y la monotonía en el segundo, ya que $A\setminus F_n\supseteq A\setminus E$. La desigualdad vale para todo n, así que se puede pasar al límite —es una serie de términos no negativos—:

$$m^*(A) \;\geq\; \sum_{k\geq 1} m^*(A\cap E_k) + m^*(A\setminus E) \;\geq\; m^*(A\cap E) + m^*(A\setminus E) \;\geq\; m^*(A),$$

donde la segunda desigualdad es subaditividad aplicada a $A \cap E = \bigsqcup_k (A \cap E_k)$ y la tercera lo es otra vez. Como los extremos coinciden, *todas* las desigualdades son igualdades: $E \in \mathcal{M}$.

Paso 5: σ-aditividad. Tomando $A = E$ en la cadena de igualdades recién obtenida, $m^*(E) = \sum_{k\geq 1} m^*(E \cap E_k) = \sum_{k\geq 1} m^*(E_k)$, que es exactamente la condición (ii) de la definición 31.5. □

Observación 31.12. Nótese dónde se ha usado que la unión es numerable: en el paso 4, al pasar al límite en una *serie*. Con uniones arbitrarias no hay serie que pasar al límite y el argumento se detiene; por eso la teoría es numerablemente aditiva y no más, y por eso el conjunto de Vitali, que se construye con una familia numerable de traslaciones, no consigue ser medible.

Funciones medibles y funciones simples

Definición 31.13 (Función medible)**.** Sea $(X, \mathcal{M}, \mu)$ un espacio de medida. Una función $f : X \to [-\infty, +\infty]$ es *medible* si

$$\{x \in X : f(x) > a\} \in \mathcal{M} \qquad \text{para todo } a \in \mathbb{R}.$$

Basta comprobarlo con $> a$: los conjuntos $\{f \geq a\} = \bigcap_n \{f > a - \frac{1}{n}\}$, $\{f < a\}$ y $\{f \leq a\}$ se obtienen por intersecciones numerables y complementos, que la proposición 31.3 garantiza dentro de $\mathcal{M}$.

Definición 31.14 (Función simple)**.** Una función $\varphi : X \to \mathbb{R}$ es *simple* si es medible y toma solo una cantidad *finita* de valores. Toda función simple se escribe de forma única en su *representación canónica*

$$\varphi = \sum_{i=1}^{n} c_i \, \mathbf{1}_{E_i},$$

donde $c_1, \ldots, c_n$ son sus valores distintos y $E_i = \{x : \varphi(x) = c_i\}$ son medibles y disjuntos. Aquí $\mathbf{1}_E$ denota la función indicadora, que vale 1 en E y 0 fuera. Las funciones simples son a la integral de Lebesgue lo que las funciones escalonadas son a la de Riemann, con una diferencia decisiva: los E_i no tienen por qué ser intervalos, sino cualquier conjunto medible.

Teorema 31.15 (Aproximación por simples)**.** *Si $f : X \to [0, +\infty]$ es medible, existe una sucesión creciente de funciones simples no negativas $0 \leq \varphi_1 \leq \varphi_2 \leq \cdots$ con $\varphi_n(x) \to f(x)$ en todo punto.*

Demostración. Se corta el *recorrido*, que es la idea entera de Lebesgue. Para cada n se define

$$\varphi_n = \sum_{k=0}^{n2^n - 1} \frac{k}{2^n} \, \mathbf{1}_{E_{n,k}} + n \, \mathbf{1}_{\{f \geq n\}}, \qquad E_{n,k} = \Big\{ x : \ \frac{k}{2^n} \leq f(x) < \frac{k+1}{2^n} \Big\}.$$

Cada $E_{n,k}$ es medible por la definición 31.13, y φ_n toma finitos valores: es simple. Es creciente en n porque al pasar de n a $n+1$ cada franja se parte en dos y el valor asignado solo puede subir. Y converge: si $f(x) < \infty$, en cuanto $n > f(x)$ se tiene $0 \leq f(x) - \varphi_n(x) \leq 2^{-n}$; y si $f(x) = \infty$, entonces $\varphi_n(x) = n \to \infty$. □

Ejemplo 31.16 (La aproximación, con números)**.** Tomemos $f(x) = x$ en $[0,1]$. Aquí $\varphi_n(x) = \dfrac{\lfloor 2^n x \rfloor}{2^n}$, es decir, la escalera que redondea hacia abajo en múltiplos de 2^{-n}. Su integral es

$$\int_0^1 \varphi_n = \sum_{k=0}^{2^n-1} \frac{k}{2^n}\cdot\frac{1}{2^n} = \frac{1}{2} - \frac{1}{2^{n+1}},$$

que vale 0,25 para $n=1$; 0,375 para $n=2$; 0,4375 para $n=3$ y 0,499512 para $n=10$. Sube hacia $\frac{1}{2} = \int_0^1 x\,\mathrm{d}x$, exactamente como promete el teorema, y el error es $2^{-(n+1)}$.

La integral, en los tres pasos que la construyen

Definición 31.17 (La integral de Lebesgue en tres pasos)**.** Sea $(X, \mathcal{M}, \mu)$ un espacio de medida.

Paso 1: funciones simples no negativas. Si $\varphi = \sum_{i=1}^{n} c_i \mathbf{1}_{E_i}$ es simple con $c_i \geq 0$ y los E_i medibles y disjuntos, se define

$$\int_X \varphi\,\mathrm{d}\mu = \sum_{i=1}^{n} c_i\,\mu(E_i),$$

con el convenio $0\cdot\infty = 0$. El valor no depende de la representación elegida.

Paso 2: funciones medibles no negativas. Si $f \geq 0$ es medible,

$$\int_X f\,\mathrm{d}\mu = \sup\Big\{ \int_X \varphi\,\mathrm{d}\mu \;:\; \varphi \text{ simple},\ 0 \leq \varphi \leq f \Big\},$$

un número de $[0,+\infty]$ que siempre existe. Por el teorema 31.15 basta calcularlo con *cualquier* sucesión creciente de simples que tienda a f.

Paso 3: funciones de signo cualquiera. Se escribe $f = f^+ - f^-$ con $f^+ = \max(f,0)$ y $f^- = \max(-f,0)$, ambas medibles y no negativas. Se dice que f es *integrable* si $\int |f|\,\mathrm{d}\mu < \infty$, lo que equivale a que $\int f^+$ y $\int f^-$ sean ambas finitas, y entonces

$$\int_X f\,\mathrm{d}\mu = \int_X f^+\,\mathrm{d}\mu - \int_X f^-\,\mathrm{d}\mu.$$

El conjunto de las funciones integrables se denota $L^1(\mu)$.

Ejemplo 31.18 (Una integral por los tres pasos)**. (a) Paso 1, una simple.** Sea $\varphi = 3\cdot\mathbf{1}_{[0,1)} + 5\cdot\mathbf{1}_{[1,4)} + 2\cdot\mathbf{1}_{[4,6)}$ sobre $\mathbb{R}$ con la medida de Lebesgue. Los tres conjuntos son disjuntos y miden 1, 3 y 2, luego

$$\int_{\mathbb{R}} \varphi\,\mathrm{d}m = 3\cdot 1 + 5\cdot 3 + 2\cdot 2 = 3 + 15 + 4 = 22.$$

Ni particiones ni sumas superiores: se multiplica valor por medida y se suma.

(b) Paso 2, una no negativa. Para $f(x) = x$ en $[0,1]$, el ejemplo anterior exhibe simples $\varphi_n \uparrow f$ con $\int \varphi_n = \frac{1}{2} - 2^{-(n+1)}$, cuyo supremo es $\frac{1}{2}$. Luego $\int_{[0,1]} x\,\mathrm{d}m = \frac{1}{2}$, que coincide con la integral de Riemann, como debe ser.

(c) Paso 3, con signo. Para $f(x) = x$ en $[-1,1]$ se tiene $f^+(x) = \max(x,0)$ y $f^-(x) = \max(-x,0)$, con $\int f^+ = \int f^- = \frac{1}{2}$. Ambas son finitas, luego $f \in L^1$ y $\int_{[-1,1]} x\,\mathrm{d}m = \frac{1}{2} - \frac{1}{2} = 0$.

(d) Medible pero no integrable. $f(x) = \frac{1}{x}$ en $(0,1]$ es continua, luego medible; pero $\int_0^1 \frac{\mathrm{d}x}{x} = +\infty$, de modo que no está en L^1. Medibilidad es el permiso para *plantear* la integral; la integrabilidad es la finitud de $\int |f|$, y son dos cuestiones distintas.

Regla: calcular una integral de Lebesgue desde la definición

1. **Comprueba primero la medibilidad.** Sin ella no hay nada que integrar. Continuas, monótonas, indicadoras de medibles, y todo lo que se obtenga de ellas con sumas, productos, máximos y *límites puntuales* es medible.
2. **Si la función es simple, termina en una línea.** Multiplica cada valor por la medida de su conjunto de nivel y suma. Toda la dificultad se ha trasladado a saber medir conjuntos, que es justo lo que la teoría construye.
3. **Si es no negativa, busca una sucesión creciente de simples.** No hace falta la canónica del teorema 31.15: cualquiera vale, y la más cómoda suele ser truncar por arriba y discretizar el recorrido.
4. **Si tiene signo, parte en f^+ y f^- y comprueba que ambas integrales son finitas.** Si una de las dos es infinita y la otra no, la integral vale $\pm\infty$; si las dos lo son, la integral *no está definida*. Lebesgue es una teoría absoluta: no hay cancelaciones condicionales.
5. **Cambia la función en un conjunto nulo sin miedo.** Si $f = g$ c.t.p. en el sentido de la definición 31.10, entonces $\int f = \int g$. Es la libertad que Riemann no da y la que hace tratable la función de Dirichlet.
6. **Para intercambiar límite e integral, no improvises.** Usa los tres teoremas de convergencia de la sección siguiente —monótona, Fatou y dominada— y verifica sus hipótesis literalmente.
7. **Control.** Si $f \geq 0$ y $\int f = 0$, entonces $f = 0$ c.t.p. Si tu cuenta da integral nula para una función que es positiva en un conjunto de medida positiva, hay un error.

Dirichlet: el ejemplo que separa las dos integrales

Ejemplo 31.19 (La función de Dirichlet, calculada de las dos maneras)**.** Sea $D = \mathbf{1}_{\mathbb{Q}}$ en $[0,1]$, es decir, $D(x) = 1$ si x es racional y $D(x) = 0$ si es irracional.

(a) Riemann: no es integrable. Sea $P = \{0 = x_0 < \cdots < x_n = 1\}$ una partición *cualquiera*. Todo subintervalo $I_i = [x_{i-1}, x_i]$ tiene longitud positiva y por tanto contiene racionales e irracionales, porque ambos conjuntos son densos. Luego $M_i = \sup_{I_i} D = 1$ y $m_i = \inf_{I_i} D = 0$ para todo i, de donde

$$U(D,P) = \sum_{i=1}^{n} 1 \cdot \Delta x_i = 1, \qquad L(D,P) = \sum_{i=1}^{n} 0 \cdot \Delta x_i = 0.$$

Esto vale para toda partición, así que la integral superior es 1, la inferior es 0, y $U(D,P) - L(D,P) = 1$ no baja de 1 por mucho que se refine. Las dos integrales no coinciden: D *no* es

integrable Riemann.

(b) Lebesgue: es integrable, con integral cero. La función D es simple: toma exactamente dos valores, 1 sobre $E_1 = \mathbb{Q} \cap [0,1]$ y 0 sobre $E_2 = [0,1] \setminus \mathbb{Q}$. Ambos conjuntos son medibles —E_1 es numerable y por tanto boreliano; E_2 es su complementario en $[0,1]$—, y por el teorema 31.9 se tiene $m(E_1) = 0$, luego $m(E_2) = 1 - 0 = 1$. Aplicando el paso 1 de la definición 31.17:

$$\int_{[0,1]} D \, \mathrm{d}m = 1 \cdot m(E_1) + 0 \cdot m(E_2) = 1 \cdot 0 + 0 \cdot 1 = 0.$$

Una línea. La integral existe y vale 0.

(c) Dónde está la diferencia. Riemann exige que las oscilaciones se puedan aplastar partiendo el *dominio*, y aquí no se puede: cada trocito de recta, por pequeño que sea, ve a la función saltar entre 0 y 1. Lebesgue parte el *recorrido*: solo hay dos valores, y lo único que pregunta es cuánto miden los dos conjuntos donde se toman. Como uno de ellos es nulo, no aporta nada. Dicho de otro modo, $D = 0$ en casi todo punto, y la integral no ve la diferencia.

(d) La consecuencia estructural. Enumerando $\mathbb{Q} \cap [0,1] = \{r_1, r_2, \dots\}$, las funciones $D_n = \mathbf{1}_{\{r_1,\dots,r_n\}}$ son escalonadas, integrables Riemann con integral 0, y $D_n \to D$ puntualmente. La clase de las funciones integrables Riemann no es estable bajo límites puntuales; la de las medibles sí. Ese es el motivo real de todo el capítulo, muy por encima de la anécdota de una función rara.

Teorema 31.20 (Compatibilidad con Riemann)**.** *Si $f : [a,b] \to \mathbb{R}$ es integrable Riemann, entonces es integrable Lebesgue y las dos integrales coinciden. El recíproco es falso, como muestra la función de Dirichlet. Además, una función acotada en $[a,b]$ es integrable Riemann* si y solo si *su conjunto de discontinuidades tiene medida cero, criterio que se demuestra en el capítulo 29.*

Observación 31.21. La teoría nueva no invalida nada de la vieja: la contiene. Todo lo que se calculó con primitivas y con la regla de Barrow sigue siendo correcto y sigue dando el mismo número. Lo que se gana es que la clase de funciones admisibles es mucho mayor, que es estable bajo límites y que los teoremas de intercambio $\lim \int = \int \lim$ dejan de exigir convergencia uniforme. Como contrapartida, se pierde la integración condicional: $\int_0^\infty \frac{\sin x}{x} \mathrm{d}x$ existe como impropia de Riemann y no como integral de Lebesgue, porque $\int_0^\infty \frac{|\sin x|}{x}$ diverge. La sección siguiente construye la medida concreta y las que le siguen explotan la maquinaria.

Pausa para pensar

Una última observación sobre los espacios L^p, que este capítulo define más adelante como las funciones medibles con $\int |f|^p < \infty$ y norma $\|f\|_p = \left(\int |f|^p\right)^{1/p}$. Conviene tener un ejemplo numérico en la cabeza desde ya: para $f(x) = x$ en $[0,1]$ se obtiene $\|f\|_1 = \frac{1}{2} = 0{,}5$, $\|f\|_2 = \frac{1}{\sqrt{3}} = 0{,}577350$, $\|f\|_3 = \frac{1}{\sqrt[3]{4}} = 0{,}629961$ y $\|f\|_\infty = 1$. Las normas crecen con p porque el dominio tiene medida 1: elevar a potencias altas premia los valores grandes. Y el hecho de que identifiquemos funciones iguales en casi todo punto —definición 31.10— es lo único que convierte a $\|\cdot\|_p$ en una norma de verdad, porque sin esa identificación habría funciones no nulas de norma cero, como la propia función de

Dirichlet.

31.3. La medida de Lebesgue

Idea clave

Intuición primero. Medir es asignarle un número a un conjunto de forma que se comporte como uno espera: la medida de un intervalo debe ser su longitud, la de una unión de trozos separados debe ser la suma, y trasladar un conjunto no debe cambiar su medida. La estrategia para lograrlo es directa: cubre el conjunto con intervalos, suma sus longitudes, y toma el mejor cubrimiento posible. Eso es la medida exterior.
El problema —y es un problema real, no un tecnicismo— es que la medida exterior no es aditiva sobre *todos* los conjuntos: existen conjuntos tan patológicos que romperlos en dos y sumar las medidas da algo distinto del total. El conjunto de Vitali es el ejemplo, y su construcción usa el axioma de elección. La salida de Carathéodory es elegante y algo inesperada: en vez de arreglar la medida, se restringe la clientela. Se declaran "medibles" solo aquellos conjuntos que parten limpiamente a cualquier otro, y sobre esa familia —que es enorme y contiene todo lo que usarás jamás— la teoría funciona perfectamente.
Merece la pena quedarse con dos ejemplos opuestos: los racionales, que son infinitos y densos y sin embargo miden cero; y el conjunto de Cantor, que es no numerable y también mide cero.

Definición 31.22. La *medida exterior* de $A \subseteq \mathbb{R}$ es

$$m^*(A) = \inf\Big\{ \sum_{k=1}^{\infty} \ell(I_k) : A \subseteq \bigcup_k I_k,\ I_k \text{ intervalos abiertos}\Big\},$$

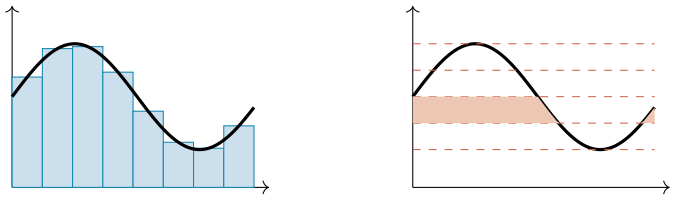
el costo más barato de cubrir A con intervalos. Es monótona, invariante por traslaciones, y *numerablemente subaditiva*: $m^*(\bigcup A_k) \leq \sum m^*(A_k)$ (unir los cubrimientos casi óptimos, con margen $\varepsilon/2^k$ al k-ésimo: el truco de la serie geométrica que ya usamos para contar $\mathbb{Q}$).

Ejemplo 31.23. m^* de un punto es 0; de un conjunto numerable, 0 (subaditividad numerable con el truco $\varepsilon/2^k$): en particular $m^*(\mathbb{Q}) = 0$: *los racionales, densos en todas partes, miden nada.* Y $m^*([a,b]) = b - a$: la desigualdad $\leq$ es obvia; la $\geq$ usa compacidad (Heine–Borel, capítulo 6: un cubrimiento abierto de $[a,b]$ tiene subcubrimiento finito, y para finitos intervalos la cuenta es inducción). La compacidad que demostramos por deporte resulta ser el cimiento de la teoría de la medida.

Idea clave

Intuición primero. ¿Qué conjunto merece tener medida? El criterio de Carathéodory responde con una prueba de *cuchillo*: E es medible si corta limpio, es decir, si al partir cualquier conjunto A en "la parte dentro de E" y "la parte fuera", las medidas exteriores suman exacto, sin pérdida ni doble conteo. Los conjuntos razonables (intervalos, abiertos, todo lo construible) son cuchillos afilados; el conjunto de Vitali es un serrucho tan

patológico que *ningún* corte suyo suma bien. La genialidad de la definición es que no dice cómo es E por dentro: lo juzga por cómo trata a los demás.



Riemann: partir el dominio Lebesgue: partir el rango

Figura 31.1: Las dos integrales: Riemann rebana el dominio en franjas verticales; Lebesgue rebana el *rango* en bandas horizontales y pregunta cuánto mide el conjunto donde f cae en cada banda: por eso solo necesita saber *medir conjuntos*.

Definición 31.24 (Carathéodory)**.** $E \subseteq \mathbb{R}$ es *medible* si parte bien a todo el mundo:

$$m^*(A) = m^*(A \cap E) + m^*(A \setminus E) \qquad \text{para todo } A \subseteq \mathbb{R}.$$

Para medibles escribimos $m(E) = m^*(E)$.

Teorema 31.25 (Estructura de los medibles)**.** *Los conjuntos medibles forman una* σ-álgebra *(cerrada bajo complementos y uniones numerables) que contiene a todos los intervalos, y sobre ella* m *es* numerablemente aditiva*: para medibles disjuntos,*

$$m\Big(\bigsqcup_{k=1}^{\infty} E_k\Big) = \sum_{k=1}^{\infty} m(E_k).$$

Idea de la demostración. Complementos: la definición es simétrica en E y E^c. Uniones finitas: aplicar la condición dos veces y reagrupar. La aditividad finita sale probando $m^*(A \cap (E_1 \sqcup E_2)) = m^*(A \cap E_1) + m^*(A \cap E_2)$ con A arbitrario. El salto a lo numerable: con $F_n = \bigsqcup_{k \le n} E_k$, la monotonía da $m^*(A) \ge \sum_{k \le n} m^*(A \cap E_k) + m^*(A \setminus \bigcup_k E_k)$ para todo n; se pasa al límite (¡una serie de términos no negativos, capítulo 6!) y la subaditividad cierra la pinza de desigualdades. Que un intervalo es medible es un cálculo directo con la definición de m^*. □ □

Todo abierto es unión numerable de intervalos, así que abiertos, cerrados, y cualquier cosa construible con operaciones numerables a partir de ellos (los *borelianos*) son medibles. La pregunta honesta es si existe algo *no* medible.

Ejemplo 31.26 (El conjunto de Vitali: sí existe, pero cuesta el axioma de elección)**.** En $[0,1)$, declárese $x \sim y$ si $x - y \in \mathbb{Q}$. Elíjase (¡axioma de elección!) un representante de cada clase: el conjunto V. Sus trasladados racionales $\{V + q \text{ mód } 1 : q \in \mathbb{Q} \cap [0,1)\}$ son disjuntos (dos elementos de V nunca difieren en un racional) y cubren $[0,1)$ (todo x difiere de su representante en un racional). Si V fuera medible, por aditividad numerable e invariancia por traslación:

$$1 = m[0,1) = \sum_q m(V+q) = \sum_q m(V) = \begin{cases} 0 & \text{si } m(V) = 0 \\ \infty & \text{si } m(V) > 0 \end{cases}:$$

absurdo en ambos casos. *No todo conjunto es medible*: la aditividad numerable, la invariancia por traslación y "medir todo" son tres deseos incompatibles, y Lebesgue sacrifica el tercero. (Sin el axioma de elección, es consistente que todo conjunto sea medible: Solovay 1970. La no medibilidad no se construye: se invoca.)

Ejemplo 31.27 (El conjunto de Cantor: grande y nulo a la vez)**.** Quítese a $[0,1]$ el tercio central abierto, a cada pedazo restante su tercio central, y así por siempre. Lo que queda, C, es cerrado, y lo removido mide

$$\frac{1}{3}+\frac{2}{9}+\frac{4}{27}+\cdots=\frac{1/3}{1-2/3}=1:$$

así $m(C)=0$. Pero C es *no numerable*: sus elementos son exactamente los números con expansión en base 3 sin unos, en biyección con las sucesiones de ceros y doses, es decir con $\{0,1\}^{\infty}$: la diagonal de Cantor del capítulo 6 aplica. Moraleja doble: medida cero no es pequeñez de cardinal, y la recta esconde fractales; la *dimensión* correcta de C es $\log 2/\log 3\approx 0.6309$, el primer ejemplo de dimensión fraccionaria.

Regla: demostrar que un conjunto tiene medida cero

Que $E\subset\mathbb{R}$ tenga medida de Lebesgue nula significa, por definición, que para todo $\varepsilon>0$ existe un recubrimiento numerable de E por intervalos cuya longitud total es menor que ε. Toda demostración de medida nula es, en el fondo, un reparto de ε.

1. **Si el conjunto es numerable, ya está.** Enumera $E=\{x_1,x_2,\dots\}$ y cubre x_k con el intervalo de longitud $\varepsilon/2^k$ centrado en él. La suma es $\sum_{k\geq 1}\varepsilon/2^k=\varepsilon$. Este truco — *repartir el ε en una serie geométrica* — es el que hay que tener siempre a mano.

2. **Unión numerable de nulos es nulo.** Mismo reparto: al conjunto E_k dale $\varepsilon/2^k$. Es la propiedad que hace que la familia de conjuntos nulos sea cómoda de manejar.

3. **Subconjunto de nulo es nulo.** Trivial, pero es lo que permite decir "basta cubrir algo mayor".

4. **Si el conjunto no es numerable, construye el recubrimiento por etapas.** La técnica estándar: describe E como intersección decreciente $E=\bigcap_m E_m$ donde cada E_m es unión finita de intervalos de longitud total λ_m, y comprueba que $\lambda_m\to 0$. Entonces $\mu(E)\leq\lambda_m$ para todo m, luego $\mu(E)=0$.

5. **Para conjuntos en $\mathbb{R}^n$, busca una dimensión que falte.** Un subconjunto de un hiperplano, o el grafo de una función continua $\{(x,f(x))\}\subset\mathbb{R}^2$, tiene medida nula: se cubre con rectángulos tan bajos como se quiera usando la continuidad uniforme sobre compactos.

6. **No confundas nulo con pequeño en otros sentidos.** Un conjunto nulo puede ser denso (los racionales), puede ser no numerable (el de Cantor) y puede tener el cardinal del continuo. Y al revés, un conjunto puede ser magro y de interior vacío

y aun así tener medida positiva. *Medida, cardinal y categoría son tres nociones de tamaño independientes.*

Ejemplo 31.28 (Cuatro conjuntos y sus medidas)**. (a) $\mathbb{Q}\cap[0,1]$ tiene medida cero.** Es numerable: aplíquese el paso 1. La consecuencia es célebre: la función de Dirichlet, que vale 1 en los racionales y 0 en los irracionales, es igual *en casi todo punto* a la función constante 0, luego es Lebesgue-integrable con $\int_0^1 \mathbf{1}_{\mathbb{Q}} = 0$. Riemann, en cambio, no la integra: sus sumas superiores valen 1 y las inferiores 0 sea cual sea la partición. Aquí se ve de golpe por qué hizo falta una teoría nueva.

(b) El conjunto de Cantor tiene medida cero y no es numerable. En la etapa m de la construcción quedan 2^m intervalos de longitud 3^{-m}, luego $\lambda_m = (2/3)^m \to 0$. Por el paso 4, $\mu(C) = 0$. Y sin embargo C está en biyección con $\{0,1\}^{\mathbb{N}}$ mediante los desarrollos en base 3 sin cifra 1, luego tiene el cardinal del continuo. Es el contraejemplo que hay que recordar: *tan grande como $\mathbb{R}$ en cardinal, invisible en medida.*

(c) El grafo de $f(x)=x^2$ en $[0,1]$ tiene medida cero en $\mathbb{R}^2$. Partiendo $[0,1]$ en N trozos, sobre cada uno la oscilación de f es a lo sumo $2/N$ (porque $|f'|\le 2$), luego el grafo cabe en N rectángulos de área $\frac{1}{N}\cdot\frac{2}{N}$. Área total $2/N\to 0$. El mismo argumento vale para cualquier función continua sobre un compacto.

(d) El conjunto de Cantor "gordo" tiene medida positiva. Repítase la construcción de Cantor pero quitando en la etapa m un intervalo central de longitud 4^{-m} de cada uno de los 2^{m-1} intervalos presentes. Lo eliminado suma

$$\sum_{m\ge 1} 2^{m-1}\cdot 4^{-m} = \frac{1}{2}\sum_{m\ge 1} 2^{-m} = \frac{1}{2},$$

así que lo que queda mide $1-\frac{1}{2}=\frac{1}{2}>0$. Y sin embargo el resultado sigue siendo cerrado, sin puntos interiores y totalmente disconexo, exactamente como el Cantor clásico. Compárese con (b) y véase el paso 6 en acción: la misma forma topológica, dos medidas incompatibles.

Regla: decidir si una función es medible

Antes de integrar hay que saber que se puede. La buena noticia es que la medibilidad es una propiedad extraordinariamente estable: casi todo lo que uno escribe con lápiz es medible, y la dificultad real es construir un ejemplo que *no* lo sea.

1. **Usa la definición operativa:** $f: X\to\mathbb{R}$ es medible si $\{x : f(x) > a\}$ es medible para todo $a\in\mathbb{R}$. Basta comprobarlo con $>a$; los $\ge a$, $<a$ y $\le a$ salen gratis por complementos e intersecciones numerables.

2. **Reconoce los ladrillos.** Son medibles: las funciones continuas, las monótonas, las indicadoras χ_E de conjuntos medibles y las funciones simples $\sum c_i\chi_{E_i}$.

3. **Aplica las operaciones que la conservan:** suma, producto, cociente (donde el denominador no se anula), múltiplos, máx, mín, $|f|$, f^+, f^-, y la composición $g\circ f$ con g *continua* por fuera.

4. **Explota el límite, que es el arma fuerte.** Si f_n son medibles, también lo son $\sup f_n$, $\inf f_n$, $\limsup f_n$, $\liminf f_n$ y, cuando existe, $\lim f_n$. Ésta es la ventaja decisiva sobre la integrabilidad de Riemann, que no sobrevive a los límites puntuales.

5. **Descompón en trozos.** Si $X = \bigcup_n X_n$ con los X_n medibles y f es medible en cada X_n, lo es en todo X. Sirve para funciones definidas a trozos, aunque sean infinitos trozos.

6. **Cuidado con el orden de la composición:** "continua $\circ$ medible" es medible; "medible $\circ$ continua" puede fallar. Y para que valga hay que trabajar con la σ-álgebra correcta: f Lebesgue-medible y g Lebesgue-medible no garantizan que $g \circ f$ lo sea.

7. **Control:** si te piden decidir sobre un conjunto *no* medible, la respuesta siempre pasa por el axioma de elección. No hay ningún ejemplo explícito, así que si tu función se puede escribir con una fórmula, es medible.

Ejemplo 31.29 (Cinco funciones ante el tribunal de la medibilidad)**. (a) La función de Dirichlet** $\chi_{\mathbb{Q}}$ vale 1 en los racionales y 0 fuera. Es medible por el paso 2, porque $\mathbb{Q}$ es numerable y por tanto medible con $\lambda(\mathbb{Q}) = 0$. Como consecuencia $\int_0^1 \chi_{\mathbb{Q}}\, d\lambda = 1 \cdot 0 + 0 \cdot 1 = 0$: la función que Riemann no puede integrar, Lebesgue la despacha en una línea.

(b) $f(x) = \sin(1/x)$ **para** $x \neq 0$**,** $f(0) = 0$**.** Es continua en $\mathbb{R} \setminus \{0\}$, luego medible ahí; y $\{0\}$ tiene medida cero. Por el paso 5, f es medible en todo $\mathbb{R}$. Nótese que la salvaje oscilación cerca del origen es irrelevante: la medibilidad no ve nada de eso.

(c) El límite puntual $f_n(x) = \big(\cos(n!\,\pi x)\big)^{2n}$ **en** $[0,1]$**.** Cada f_n es continua, luego medible. Su límite puntual existe y vale exactamente $\chi_{\mathbb{Q}}$: si $x = p/q$ es racional, para $n \geq q$ el número $n!\,x$ es entero y $\cos(n!\pi x) = \pm 1$, de modo que $f_n(x) = 1$; si x es irracional, $|\cos(n!\pi x)| < 1$ y la potencia $2n$ lo aplasta a 0. Por el paso 4, el límite es medible, y así reobtenemos (a).

(d) $f(x) = \chi_C(x)$ **con** C **el conjunto de Cantor.** C es cerrado, luego medible, y $\lambda(C) = 0$; así que f es medible con integral 0. Y esto pese a que C es no numerable, con la misma cardinalidad que $\mathbb{R}$: medida cero y "ser pequeño en cardinal" son cosas distintas.

(e) χ_V **con** V **un conjunto de Vitali.** Se construye eligiendo un representante de cada clase de $\mathbb{R}/\mathbb{Q}$ dentro de $[0,1]$ (aquí entra el axioma de elección). Si V fuera medible, las traslaciones $V + q$ con $q \in \mathbb{Q} \cap [-1,1]$ serían disjuntas, todas de la misma medida m, y su unión estaría entre $[0,1]$ y $[-1,2]$; entonces $1 \leq \sum_q m \leq 3$. Pero una suma numerable de copias del mismo número solo puede valer 0 (si $m = 0$) o $+\infty$ (si $m > 0$). Contradicción: V no es medible, y χ_V tampoco.

Idea clave

El paso 4 es toda la razón de ser de este capítulo. En Riemann, un límite puntual de funciones integrables puede no ser integrable —el ejemplo (c) lo exhibe—; en Lebesgue, un límite puntual de funciones medibles es siempre medible. Por eso los teoremas de convergencia son enunciables aquí y no allí.

Cuidado

Medible no implica integrable. La función $f(x) = 1/x$ en $(0, 1]$ es continua y por tanto medible, pero $\int_0^1 f \, \mathrm{d}\lambda = +\infty$. Medibilidad es el permiso para *plantear* la integral; la finitud es una cuestión aparte, que se resuelve con $\int |f| < \infty$.

Regla: construir un conjunto de Cantor y calcular su medida

Los conjuntos de Cantor son la fábrica de contraejemplos de toda la teoría de la medida. Casi cualquier pregunta del tipo «¿implica esto aquello?» se responde con uno de ellos. Vale la pena saber fabricarlos a voluntad.

1. **Fija el patrón de borrado.** Partes de $[0, 1]$. En el paso n tendrás 2^{n-1} intervalos y de cada uno quitarás un trozo abierto centrado. Lo único que decides es la longitud ℓ_n de cada trozo borrado.

2. **Suma lo que borras.** En el paso n borras 2^{n-1} trozos de longitud ℓ_n, o sea $2^{n-1}\ell_n$ de medida. Como los trozos borrados son disjuntos dos a dos, la medida total borrada es exactamente
$$\mu(\text{borrado}) = \sum_{n=1}^{\infty} 2^{n-1}\ell_n.$$

3. **Resta.** El conjunto que queda, $C = [0, 1] \setminus (\text{borrado})$, es cerrado (complemento de una unión de abiertos) y su medida es $\mu(C) = 1 - \sum_{n\geq 1} 2^{n-1}\ell_n$. Aquí no hay ninguna sutileza: la aditividad numerable hace todo el trabajo.

4. **Elige ℓ_n según lo que quieras probar.** Si tomas $\ell_n = 3^{-n}$ la suma da 1 y obtienes medida cero. Si tomas $\ell_n = 4^{-n}$ la suma da $1/2$ y obtienes medida $1/2$. El patrón es siempre el mismo objeto topológico; solo cambia la medida.

5. **Comprueba que es denso en ninguna parte.** En el paso n los intervalos que sobreviven tienen longitud $< 2^{-n}$, así que C no contiene ningún intervalo. Al ser además cerrado, su interior es vacío: es denso en ninguna parte, *sea cual sea su medida*.

6. **Cuenta sus puntos.** Codifica cada punto superviviente por la sucesión de decisiones «izquierda/derecha» en cada paso. Esa codificación es una biyección con $\{0, 1\}^{\mathbb{N}}$, luego C tiene la cardinalidad del continuo. Nunca es numerable.

7. **Verifica con las longitudes que sobreviven.** Tras el paso n quedan 2^n intervalos; multiplica su número por su longitud y comprueba que coincide con $1 - \sum_{k\leq n} 2^{k-1}\ell_k$. Si no cuadra, has borrado trozos que se solapaban o que se salían.

Ejemplo 31.30 (Dos conjuntos de Cantor, una escalera del diablo y cuatro mitos derribados)**.** Construimos los dos casos extremos y los usamos para separar nociones que la intuición confunde.

(a) El conjunto de Cantor clásico: medida cero, cardinalidad del continuo.

Tomamos $\ell_n = 3^{-n}$: en el paso 1 borramos el tercio central $(1/3, 2/3)$, en el paso 2 los dos tercios centrales de lo que queda, y así sucesivamente. La medida borrada es

$$\sum_{n=1}^{\infty} 2^{n-1}3^{-n} = \frac{1}{3}\sum_{n=1}^{\infty}\left(\frac{2}{3}\right)^{n-1} = \frac{1}{3}\cdot\frac{1}{1-2/3} = 1,$$

luego $\mu(C) = 0$. Por el paso 6, C tiene $2^{\aleph_0}$ puntos. En términos de expansiones ternarias, C es exactamente el conjunto de los $x \in [0,1]$ que admiten una escritura en base 3 sin la cifra 1.

Cuidado

Aquí muere el mito más común: *«medida cero» no significa «pocos puntos»*. C tiene tantos puntos como $[0,1]$ entero y sin embargo mide 0. Los conjuntos numerables tienen medida cero, pero el recíproco es rotundamente falso.

(b) El conjunto gordo de Cantor: denso en ninguna parte, medida $1/2$. Ahora $\ell_n = 4^{-n}$. Paso a paso: borramos primero un intervalo central de longitud $1/4$ y quedan dos intervalos de longitud $3/8$ cada uno; borramos de cada uno un trozo central de longitud $1/16$ y quedan cuatro intervalos de longitud $5/32$; después ocho de longitud $9/128$. Las comprobaciones del paso 7 cuadran: $2\cdot\frac{3}{8} = \frac{3}{4}$, $4\cdot\frac{5}{32} = \frac{5}{8}$, $8\cdot\frac{9}{128} = \frac{9}{16}$. En el límite,

$$\mu(\text{borrado}) = \sum_{n=1}^{\infty} 2^{n-1}4^{-n} = \frac{1}{2}\sum_{n=1}^{\infty}\left(\frac{2}{4}\right)^{n} = \frac{1}{2}\cdot 1 = \frac{1}{2}, \qquad \mu(F) = \frac{1}{2}.$$

F es cerrado, sin puntos interiores, y mide la mitad del intervalo.

Cuidado

Segundo mito derribado: *«denso en ninguna parte» no implica «medida cero»*. La topología y la medida son dos formas distintas de decir «pequeño», y no se implican en ninguna dirección. (En la otra dirección: $\mathbb{Q}\cap[0,1]$ es denso, luego no es denso en ninguna parte, y sin embargo mide 0.)

(c) Una función acotada que Riemann no puede integrar, pero Lebesgue sí. Sea $g = \mathbf{1}_F$ la indicadora del conjunto gordo. Su conjunto de discontinuidades es precisamente F, que mide $1/2 > 0$; por el criterio de Lebesgue para la integrabilidad de Riemann, g *no* es Riemann integrable. Como Lebesgue no le pide continuidad sino medibilidad, y F es cerrado (luego boreliano), la integral de Lebesgue existe sin ningún problema:

$$\int_{[0,1]} \mathbf{1}_F \, d\mu = \mu(F) = \frac{1}{2}.$$

Compara con la indicadora del Cantor clásico, $\mathbf{1}_C$: esa *sí* es Riemann integrable (su conjunto de discontinuidades C mide cero) y su integral vale 0.

(d) La escalera del diablo. Sobre el Cantor clásico se define $c : [0,1] \to [0,1]$ así: si $x = \sum a_k 3^{-k}$ con $a_k \in \{0,2\}$, entonces $c(x) = \sum (a_k/2)2^{-k}$; en cada intervalo borrado c se declara constante e igual al valor común de sus extremos. La función resultante es continua y

creciente, con $c(0) = 0$ y $c(1) = 1$. Dos valores concretos: $1/3 = 0{,}0222\ldots_3$ da $c(1/3) = 1/2$, y $1/4 = 0{,}\overline{02}_3$ está en C con

$$c(1/4) = \sum_{k\geq 1} 4^{-k} = \frac{1/4}{1 - 1/4} = \frac{1}{3}.$$

(e) Por qué la escalera rompe el teorema fundamental del cálculo. La función c es constante en cada intervalo borrado, y esos intervalos ocupan medida total 1. Luego $c'(x) = 0$ para casi todo x. Por tanto

$$\int_0^1 c'(x)\,dx = 0 \neq 1 = c(1) - c(0).$$

No hay error: la regla de Barrow en su forma $\int_a^b f' = f(b) - f(a)$ exige *continuidad absoluta*, no solo continuidad. c es continua, incluso uniformemente continua, y falla de todos modos. El «salto» de altura 1 está repartido sobre un conjunto de medida cero, invisible para la integral.

Pausa para pensar

¿Cuánto mide la gráfica de c, es decir, su longitud como curva en el plano? Las partes horizontales suman longitud 1 (recorren todo el eje x salvo medida cero) y las subidas suman altura 1; la longitud total es exactamente 2, la misma que la de la escalera de dos peldaños. Una curva continua que sube 1 sin tener nunca pendiente positiva.

Idea clave

El Cantor clásico tiene dimensión de Hausdorff $\log 2/\log 3 \approx 0{,}6309$: ni punto (0) ni segmento (1), sino algo intermedio. La medida de Lebesgue solo sabe decir «cero» sobre él; hace falta una familia más fina de medidas, las de Hausdorff $\mathcal{H}^s$, para distinguir un Cantor de otro. Ese es el punto de partida de la geometría fractal. Y un dato que suele sorprender: $C + C = \{x + y : x, y \in C\} = [0, 2]$, es decir, dos copias de un conjunto de medida cero se suman para llenar un intervalo entero.

31.4. La integral de Lebesgue y los tres teoremas de convergencia

Idea clave

Intuición primero. Aquí está el pago de todo el trabajo anterior. La pregunta práctica es: si $f_n \to f$, ¿es cierto que $\int f_n \to \int f$? Con Riemann la respuesta honesta es "casi nunca sin hipótesis fuertes", y eso obliga a comprobar convergencia uniforme a cada paso, cosa que rara vez se tiene. Los tres teoremas de esta sección dan condiciones mucho más laxas y son, en la práctica, la razón por la que se usa la integral de Lebesgue.
La convergencia monótona dice que si las funciones solo suben, se puede pasar al límite sin más. El lema de Fatou da una desigualdad que vale *siempre*, sin hipótesis, y por eso es la herramienta de emergencia. Y la convergencia dominada es la que se usa a

diario: basta encontrar una única función integrable que acote a todas las f_n a la vez, y el intercambio queda justificado. Esa función dominante es una especie de techo común; la intuición de por qué hace falta es que sin techo la masa puede escaparse al infinito, y los contraejemplos clásicos (la joroba que se aleja, el pico que se estrecha) muestran exactamente eso: la masa no desaparece, se va.

Definición 31.31. $f:\mathbb{R}\to[-\infty,\infty]$ es *medible* si $\{x : f(x) > a\}$ es medible para todo a. Para f medible no negativa,

$$\int f\,\mathrm{d}m = \sup\Big\{\sum_i c_i\, m(E_i) : \text{funciones simples } \sum c_i \mathbf{1}_{E_i} \le f\Big\}$$

(aproximar por debajo con "histogramas horizontales"); en general, $\int f = \int f^+ - \int f^-$ cuando $\int |f| < \infty$ (f *integrable*, $f \in L^1$).

La función de Dirichlet, intratable para Riemann, es trivial ahora: $D = \mathbf{1}_{\mathbb{Q}}$ es simple y $\int D = 1 \cdot m(\mathbb{Q}) = 0$. Más aún: si f es integrable Riemann, es integrable Lebesgue con la misma integral (las sumas de Riemann son integrales de funciones simples que emparedan), y el criterio definitivo es de Lebesgue: *f acotada es integrable Riemann si y solo si su conjunto de discontinuidades tiene medida cero*. La teoría nueva no rompe nada de lo viejo: lo explica y lo extiende.

Teorema 31.32 (Convergencia monótona, Beppo Levi)**.** *Si $0 \le f_1 \le f_2 \le \cdots$ y $f_n \to f$ puntualmente, entonces $\int f_n \to \int f$.*

Idea. $\le$ es monotonía. Para $\ge$: fijada una simple $\varphi \le f$ y $c < 1$, los conjuntos $E_n = \{f_n \ge c\varphi\}$ crecen hasta llenar todo; la aditividad numerable da $\int_{E_n} \varphi \to \int \varphi$, así que $\lim \int f_n \ge c \int \varphi$ para todo $c < 1$ y toda φ: tomar supremos. Toda la fuerza viene de la aditividad *numerable*: el teorema es falso para la integral de Riemann precisamente porque las sumas finitas no soportan este paso. □ □

Idea clave

Qué dice esto en cristiano. Los tres teoremas de convergencia son la razón de ser de la integral de Lebesgue: dicen cuándo se puede meter el límite dentro de la integral. Con Riemann eso casi nunca se puede sin exigir convergencia uniforme, que es una condición fuerte y rara. La monótona dice que si las funciones solo crecen, el límite pasa siempre. Fatou da una desigualdad que vale sin ninguna hipótesis extra. Y la dominada, la más usada, dice que basta con encontrar una función integrable que las tape a todas: entonces el límite pasa aunque la convergencia sea solo puntual.

Ejemplo 31.33 (Los tres teoremas y sus fallos)**. (a) La monótona en acción.** Sea $f_n = \mathbf{1}_{[0,n]} \cdot \frac{1}{1+x^2}$, creciente en n. Sus integrales son $2\arctan n$, que crecen hasta π, y el límite es $\int_{\mathbb{R}} \frac{\mathrm{d}x}{1+x^2} = \pi$. Coinciden, como promete el teorema, y no hizo falta ninguna uniformidad.

(b) La joroba viajera: Fatou con desigualdad estricta. Sea $f_n = n\,\mathbf{1}_{[0,1/n]}$. Puntualmente $f_n \to 0$ en todo punto, luego $\int \liminf f_n = 0$. Pero $\int f_n = 1$ para todo n, así que $\liminf \int f_n = 1$. Se cumple $0 \leq 1$: Fatou vale, y con desigualdad estricta. La masa se escapa hacia arriba sin llegar al límite.

(c) Por qué falla la dominada ahí. En el ejemplo (b) no existe ninguna $g \in L^1$ que domine a todas las f_n: haría falta $g(x) \geq n$ en $[0, 1/n]$ para todo n, es decir $g(x) \geq \frac{1}{x}$, cuya integral cerca del origen diverge. Sin dominante no hay teorema, y en efecto la conclusión es falsa.

(d) La dominada funcionando. Sea $f_n(x) = \frac{\sin(nx)}{n}$ en $[0, \pi]$. Converge puntualmente a 0 y está dominada por $g = 1$, que es integrable en un intervalo finito. Por tanto $\int_0^\pi f_n \to 0$. Comprobación directa: $\int_0^\pi \frac{\sin(nx)}{n} \mathrm{d}x = \frac{1-\cos(n\pi)}{n^2}$, que vale $\frac{2}{n^2}$ para n impar y 0 para n par; en ambos casos tiende a cero.

Cuidado

[La dominante debe ser una sola función, integrable y fija] No vale una dominante distinta para cada n, ni una que dependa de n: tiene que ser una única $g \in L^1$ que sirva para todas a la vez. Ese es exactamente el punto donde falla el ejemplo (b). Segundo aviso: la monótona exige que las funciones sean *no negativas* y crecientes; con sucesiones decrecientes puede fallar, como $f_n = \mathbf{1}_{[n,\infty)}$, cuyas integrales valen ∞ para todo n mientras que el límite puntual es 0. Y tercero: Fatou solo da una desigualdad, nunca una igualdad, y su valor está justo en no pedir nada a cambio.

Lema 31.34 (Fatou)**.** *Para $f_n \geq 0$ medibles:* $\int \liminf f_n \leq \liminf \int f_n$.

Demostración. $g_n = \inf_{k \geq n} f_k$ es monótona creciente hacia $\liminf f_n$ y $g_n \leq f_n$: convergencia monótona a la izquierda, monotonía a la derecha. □ □

Idea clave

Qué dice esto en cristiano. Al pasar al límite, la integral puede *perder* masa pero nunca ganarla. La masa se puede escapar de tres maneras: hacia arriba (una joroba cada vez más alta y estrecha), hacia los lados (una joroba que se va al infinito) o diluyéndose (una joroba cada vez más baja y ancha). Fatou no impide ninguna de las tres; sólo garantiza que lo que quede en el límite no sea más de lo que había. Por eso la desigualdad es en ese sentido y no en el otro.

Ejemplo 31.35 (Las tres fugas de masa)**. (a) La joroba alta.** Sea $f_n = n\,\mathbf{1}_{(0,1/n)}$ en $[0, 1]$. Entonces $\int f_n = 1$ para todo n, mientras que $f_n(x) \to 0$ en todo punto. Fatou da $0 \leq 1$: desigualdad estricta. La masa se escapó por arriba.

(b) La joroba viajera. Sea $f_n = \mathbf{1}_{[n,n+1]}$ en $\mathbb{R}$. De nuevo $\int f_n = 1$ y $f_n \to 0$ puntualmente. La masa no desapareció: se fue al infinito.

(c) La joroba que se aplana. Sea $f_n = \frac{1}{n}\mathbf{1}_{[0,n]}$. Otra vez $\int f_n = 1$ y $f_n \to 0$. Aquí la masa se diluyó sobre un dominio cada vez mayor.

(d) Sin la hipótesis de positividad, falla. Con $f_n = -n\,\mathbf{1}_{(0,1/n)}$ se tiene $\int f_n = -1$ y $\liminf f_n = 0$, de modo que la desigualdad de Fatou diría $0 \le -1$, que es falsa. La condición $f_n \ge 0$ (o, más generalmente, que estén acotadas por debajo por una función integrable) es imprescindible.

Teorema 31.36 (Convergencia dominada, Lebesgue)**.** *Si $f_n \to f$ puntualmente y existe $g \in L^1$ con $|f_n| \le g$ para todo n, entonces*

$$\lim_{n\to\infty} \int f_n = \int f.$$

Demostración. Fatou aplicado a $g + f_n \ge 0$ da $\int g + \int f \le \int g + \liminf \int f_n$; aplicado a $g - f_n \ge 0$ da $\int g - \int f \le \int g - \limsup \int f_n$. Juntas: $\limsup \int f_n \le \int f \le \liminf \int f_n$. □ □

Ejemplo 31.37 (La convergencia dominada y su hipótesis)**. (a) Un caso donde funciona.** Sea $f_n(x) = \dfrac{\sin(nx)}{1+x^2}$ en $\mathbb{R}$. Está dominada por $g(x) = 1/(1+x^2)$, que es integrable con $\int g = \pi$. Aunque f_n no converge puntualmente en ningún sitio interesante, la dominación permite manipular integrales de sucesiones parecidas con total libertad.

(b) Un caso limpio. $f_n(x) = x^n$ en $[0,1]$ está dominada por $g = 1$ y converge puntualmente a la función que vale 0 en $[0,1)$ y 1 en $x = 1$. El teorema da $\int_0^1 x^n\,dx \to 0$, y en efecto $1/(n+1) \to 0$. Nótese que la convergencia no es uniforme; a la convergencia dominada eso le da igual.

(c) Sin dominante, falla. La joroba alta $f_n = n\mathbf{1}_{(0,1/n)}$ converge puntualmente a 0 y sus integrales valen 1. No hay contradicción: la menor función que domina a todas es $g(x) \ge 1/x$ cerca del origen, que no es integrable. El teorema no se puede aplicar y el resultado falla.

(d) Frente a Riemann. La sucesión $f_n = \mathbf{1}_{\{q_1,\dots,q_n\}}$, indicatriz de los n primeros racionales de $[0,1]$, converge puntualmente a la función de Dirichlet, que no es integrable Riemann. Con Lebesgue todo va bien: $\int f_n = 0$ para todo n y la integral del límite también vale 0. Ésta es exactamente la ventaja por la que se construyó la teoría.

Observación 31.38. Este es el teorema que el capítulo 6 pedía a gritos. La convergencia uniforme queda jubilada como hipótesis de intercambio: basta una *dominante integrable*. La derivación bajo la integral (el truco de Feynman) es un corolario limpio: el cociente incremental se domina con el valor medio. Y la hipótesis no es decorativa:

$$f_n = n\,\mathbf{1}_{(0,1/n)} \to 0 \text{ puntualmente, pero } \int f_n = 1 \not\to 0 :$$

la masa se escapa por una chimenea cada vez más alta y flaca; ninguna g integrable domina a todas las f_n (tendría que superar a $\sup_n f_n \ge \frac{1}{x}$ cerca de 0). Las tres formas de perder masa —chimenea, meseta que se aplana hacia ∞, joroba que se fuga a ∞— son las tres únicas patologías, y Fatou dice que la masa solo puede *perderse* en el límite, nunca aparecer.

Cuidado

La frontera exacta entre lo viejo y lo nuevo: $\int_0^\infty \frac{\sin x}{x}\,\mathrm{d}x = \frac{\pi}{2}$ existe como integral *impropia* de Riemann (cancelación de signos, capítulo 30 por residuos), pero *no* es integrable Lebesgue: $\int_0^\infty \frac{|\sin x|}{x} = \infty$ (comparar con la serie armónica sobre los bloques $[k\pi, (k+1)\pi]$). Lebesgue es una teoría *absoluta*: integra $|f|$ o no integra; las cancelaciones condicionales quedan para los límites de integrales propias, igual que las series condicionalmente convergentes quedaron fuera de la reordenación libre (Riemann otra vez, capítulo 6: es la misma dicotomía absoluto/condicional).

Regla: elegir cuál de los tres teoremas de convergencia aplicar

El problema es siempre el mismo: se tiene $f_n \to f$ puntualmente y se quiere concluir $\int f_n \to \int f$. Eso *no* es automático, y los tres teoremas son las tres licencias que permiten hacerlo. La elección se decide en dos preguntas.

1. **Primero: ¿la sucesión es creciente y no negativa?** Si $0 \le f_1 \le f_2 \le \cdots$ puntualmente, usa **convergencia monótona**. No necesitas nada más: ni cota, ni integrabilidad previa, ni siquiera que el límite sea finito. Es el teorema más barato de aplicar y por eso hay que preguntarlo primero.

2. **Corolario gratis de la monótona: se pueden intercambiar $\sum$ e $\int$ si los términos son ≥ 0.** Las sumas parciales son crecientes. Este es, con diferencia, el uso más frecuente del teorema.

3. **Si no es monótona: ¿hay una cota integrable uniforme?** Busca $g \ge 0$ con $\int g < \infty$ y $|f_n| \le g$ para todo n (la misma g para todos, ese es el punto). Si la encuentras, usa **convergencia dominada**, y de propina obtienes también $\int |f_n - f| \to 0$.

4. **Cómo se busca la dominante.** No la adivines: acótala. Usa $|\sin t| \le 1$, $|\sin t| \le |t|$, $(1+t)^{-1} \le 1$, o la propia cota del enunciado. La dominante suele ser el "peor caso" de la familia, y muchas veces es simplemente $|f_1|$ o una función explícita como e^{-x} o $1/(1+x^2)$.

5. **Si no hay dominante, todavía queda Fatou.** Para $f_n \ge 0$ siempre vale
$$\int \lim\inf f_n \;\le\; \lim\inf \int f_n.$$
No da igualdad, pero da una desigualdad en la dirección útil, y basta para muchísimas cosas. Recuerda la dirección con la imagen: *la masa se puede escapar al pasar al límite, pero no puede aparecer de la nada.*

6. **Si el límite de las integrales no es el que esperas, busca la fuga de masa.** Cuando falla, falla siempre por lo mismo: una joroba que se estrecha y se alza (masa que se va al infinito en altura) o que se desplaza (masa que se va al infinito en posición). Identificarla es la forma de convencerse de que no hay dominante posible.

Ejemplo 31.39 (Los tres teoremas, uno tras otro, sobre ejemplos concretos)**. (a) Monótona, para intercambiar suma e integral.** Calculemos

$$\int_0^1 \frac{-\ln(1-x)}{x}\,\mathrm{d}x.$$

Desarrollando $-\ln(1-x)=\sum_{n\geq 1} x^n/n$ y dividiendo por x, el integrando es $\sum_{n\geq 1} x^{n-1}/n$, una serie de términos *no negativos* en $(0,1)$. Las sumas parciales crecen, luego la convergencia monótona autoriza el intercambio:

$$\int_0^1 \sum_{n\geq 1} \frac{x^{n-1}}{n}\,\mathrm{d}x = \sum_{n\geq 1} \frac{1}{n}\int_0^1 x^{n-1}\,\mathrm{d}x = \sum_{n\geq 1}\frac{1}{n^2} = \frac{\pi^2}{6}.$$

Numéricamente la integral vale $1{,}6449340668\ldots$ y $\pi^2/6 = 1{,}6449340668\ldots$: coinciden. Obsérvese que no hizo falta ninguna cota; la monotonía lo hizo todo.

(b) Dominada, con la cota escrita a mano. Sea

$$I_n = \int_0^n \Big(1-\frac{x}{n}\Big)^n \mathrm{e}^{x/2}\,\mathrm{d}x.$$

Extendemos el integrando por 0 fuera de $[0,n]$ y llamamos f_n al resultado. Puntualmente $(1-x/n)^n \to \mathrm{e}^{-x}$, luego $f_n(x)\to \mathrm{e}^{-x}\mathrm{e}^{x/2}=\mathrm{e}^{-x/2}$. La dominante sale de la desigualdad elemental $(1-x/n)^n \leq \mathrm{e}^{-x}$ válida para $0\leq x\leq n$: entonces $0\leq f_n(x)\leq \mathrm{e}^{-x/2}$, que es integrable en $(0,\infty)$ con integral 2. Convergencia dominada:

$$\lim_{n\to\infty} I_n = \int_0^\infty \mathrm{e}^{-x/2}\,\mathrm{d}x = 2.$$

La aproximación numérica lo confirma: $I_5=1{,}258$, $I_{20}=1{,}702$, $I_{100}=1{,}9257$, $I_{1000}=1{,}99206$, $I_{10000}=1{,}99920$. Converge, despacio pero sin duda, a 2.

(c) Fatou, y por qué la desigualdad es estricta. Sea $f_n = n\cdot \mathbf{1}_{(0,1/n)}$ en $(0,1)$: un rectángulo de base $1/n$ y altura n. Para cada $x>0$ fijo se tiene $f_n(x)=0$ en cuanto $n>1/x$, luego $f_n\to 0$ puntualmente. Pero $\int_0^1 f_n = 1$ para todo n. Así que

$$0 = \int \lim f_n \;<\; \lim \int f_n = 1 :$$

Fatou se cumple, con desigualdad *estricta*. Y no hay dominante posible, porque cualquier $g\geq f_n$ para todo n cumpliría $g(x)\geq n$ para $x<1/n$, es decir $g(x)\gtrsim 1/x$ cerca de 0, que no es integrable. Aquí está la fuga de masa del paso 6: el área 1 no desaparece, se escapa hacia arriba mientras la base se encoge.

Cuidado

La hipótesis de la convergencia dominada es una *única* g que sirva para *todos* los n. El error típico es acotar $|f_n|\leq g_n$ con una cota que depende de n — eso no dice nada. En el ejemplo (c) cada f_n está acotada por la constante n, que es integrable en $(0,1)$; y sin embargo el teorema no se aplica, precisamente porque esa constante cambia con n.

31.5. El cambio de variable y las coordenadas polares

Hay una fórmula que se usa en cada página de un curso de cálculo en varias variables y que nadie demuestra allí, porque su demostración honesta necesita exactamente la teoría de la medida de este capítulo: el teorema del cambio de variable. Merece la pena enunciarlo bien, porque la versión de Lebesgue es notablemente más limpia que la de Riemann —no hay que preocuparse por fronteras, ni por regiones raras, ni por la orientación— y porque de ella salen, casi de regalo, el valor de la integral de Gauss y el volumen de la bola en cualquier dimensión.

Teorema 31.40 (Cambio de variable)**.** *Sean $U, V \subseteq \mathbb{R}^n$ abiertos y $\Phi\colon U \to V$ un difeomorfismo de clase C^1. Para toda $f\colon V \to [0,\infty]$ medible,*

$$\int_V f(y)\,\mathrm{d}y = \int_U f(\Phi(u))\,\bigl|\det D\Phi(u)\bigr|\,\mathrm{d}u,$$

y la misma identidad vale para $f \in L^1(V)$. El factor $|\det D\Phi|$ se llama jacobiano de la transformación.

Idea clave

El jacobiano es un factor de dilatación local de volumen, y esto no es una metáfora: la diferencial $D\Phi(u)$ es la mejor aproximación lineal de Φ cerca de u, y una aplicación lineal T multiplica la medida de Lebesgue de cualquier conjunto por $|\det T|$. Ese hecho lineal, que se comprueba primero para cubos y se extiende por la unicidad de la medida, es todo el contenido del teorema; el resto es cubrir U con trozos pequeños donde Φ casi coincida con su diferencial. El valor absoluto está porque la medida no distingue orientaciones: a diferencia de la integral de una forma, aquí voltear el espacio no cambia el signo.

Regla: aplicar un cambio de variable en varias dimensiones

1. **Elige el cambio mirando el dominio, no el integrando.** Si la región es un disco, un anillo o un sector, las polares casi siempre ganan aunque la función parezca fea; si es un paralelogramo, el cambio correcto es lineal.
2. **Escribe Φ del nuevo espacio al viejo.** La fórmula pide $y = \Phi(u)$ con u la variable nueva; escribirla al revés es el error más común y produce el jacobiano invertido.
3. **Calcula la matriz de derivadas parciales y su determinante.** En polares $\Phi(r,\theta) = (r\cos\theta, r\sin\theta)$ y $\det D\Phi = r$; en esféricas $\det D\Phi = \rho^2 \sin\varphi$; en un cambio lineal $y = Au$, el jacobiano es la constante $|\det A|$.
4. **Toma valor absoluto y no lo olvides nunca.** Un jacobiano negativo indica que la transformación invierte la orientación, cosa que a la medida no le importa en absoluto.
5. **Traduce también la región de integración.** Los nuevos límites describen

$U = \Phi^{-1}(V)$; dibujar la región en las coordenadas nuevas suele ser más rápido que despejar.

6. **No te asustes si el difeomorfismo falla en un conjunto pequeño.** Las polares no son inyectivas sobre la semirrecta $\theta = 0$ ni están definidas en el origen, pero eso es un conjunto de medida cero y la integral de Lebesgue lo ignora: aquí la teoría es estrictamente más cómoda que la de Riemann.

7. **Comprueba el resultado con un caso conocido.** Integrar la función constante 1 debe devolver el volumen de la región; si no lo hace, el jacobiano está mal.

Ejemplo 31.41 (Las polares, y la integral que Riemann no sabe calcular)**. (a) Un anillo.** Para $\int\int_{1\le x^2+y^2\le 4}(x^2+y^2)\,\mathrm{d}A$, las polares convierten el integrando en r^2 y el elemento de área en $r\,\mathrm{d}r\,\mathrm{d}\theta$, de modo que la integral vale $\int_0^{2\pi}\int_1^2 r^3\,\mathrm{d}r\,\mathrm{d}\theta = 2\pi\cdot\frac{2^4-1}{4} = \frac{15\pi}{2}$. En cartesianas habría que partir la región en cuatro trozos y aparecerían raíces por todas partes.

(b) La integral de Gauss. El truco clásico consiste en elevar al cuadrado. Si $I = \int_{-\infty}^{\infty} \mathrm{e}^{-x^2}\mathrm{d}x$, entonces por Tonelli —que es lícito porque el integrando es no negativo—

$$I^2 = \int_{\mathbb{R}^2} \mathrm{e}^{-(x^2+y^2)}\,\mathrm{d}A = \int_0^{2\pi}\int_0^{\infty} \mathrm{e}^{-r^2} r\,\mathrm{d}r\,\mathrm{d}\theta = 2\pi\cdot\frac{1}{2} = \pi,$$

luego $I = \sqrt{\pi}$. La aparición del factor r es lo único que hace la integral elemental: sin él, e^{-r^2} no tiene primitiva en términos de funciones usuales.

(c) Una gaussiana con términos cruzados. Para $\int_{\mathbb{R}^2} \mathrm{e}^{-(x^2+xy+y^2)}\mathrm{d}A$, la forma cuadrática tiene matriz $\begin{pmatrix} 1 & 1/2 \\ 1/2 & 1 \end{pmatrix}$ de determinante $3/4$. Diagonalizándola con un cambio lineal de jacobiano constante, la integral resulta $\pi/\sqrt{3/4} = 2\pi\sqrt{3}/3 \approx 3{,}6275987$, y una cuadratura numérica sobre $[-15,15]^2$ devuelve 3,62759873. El caso general es $\int_{\mathbb{R}^n} \mathrm{e}^{-\langle Ax,x\rangle}\mathrm{d}x = \pi^{n/2}/\sqrt{\det A}$ para A definida positiva.

(d) Una elipse. El cambio lineal $x = au$, $y = bv$ tiene jacobiano constante ab y transforma el disco unidad en la elipse $x^2/a^2 + y^2/b^2 \le 1$; por tanto su área es $ab\cdot\pi = \pi ab$, sin integrar nada. Este es el patrón general: un cambio lineal nunca cuesta trabajo, sólo multiplica por un número.

Ejemplo 31.42 (La bola en dimensión n, y una sorpresa)**. (a) La fórmula.** Combinando el cambio a coordenadas esféricas con la definición de la función Γ se obtiene que el volumen de la bola unidad de $\mathbb{R}^n$ es

$$\omega_n = \frac{\pi^{n/2}}{\Gamma\left(\frac{n}{2}+1\right)}.$$

Para $n = 1$ da 2, la longitud del segmento $[-1,1]$; para $n = 2$ da π; para $n = 3$ da $4\pi/3$, como debía.

(b) La sorpresa. Los valores siguientes son $\omega_4 = \pi^2/2 \approx 4{,}9348$, $\omega_5 = 8\pi^2/15 \approx 5{,}2638$, $\omega_6 = \pi^3/6 \approx 5{,}1677$, $\omega_7 = 16\pi^3/105 \approx 4{,}7248$ y $\omega_8 = \pi^4/24 \approx 4{,}0587$. El volumen *crece hasta*

la dimensión cinco y luego decrece a cero. La razón es que $\pi^{n/2}$ crece exponencialmente pero $\Gamma(n/2+1)$ crece como un factorial, y el factorial siempre gana: en dimensión alta, casi todo el cubo unidad está fuera de la bola inscrita.

(c) La superficie sale gratis. Integrando en cáscaras, $\omega_n = \int_0^1 \sigma_{n-1} r^{n-1} \mathrm{d}r = \sigma_{n-1}/n$, donde σ_{n-1} es la medida de la esfera unidad. Para $n = 3$ eso da $\sigma_2 = 3\omega_3 = 4\pi$, la fórmula del área de la esfera. El mismo cálculo en cáscaras da $\int_{|x|\le 1} |x|^2 \mathrm{d}x = \int_0^1 r^2 \cdot 4\pi r^2 \, \mathrm{d}r = 4\pi/5$ en dimensión tres.

(d) Por qué esto exige Lebesgue. Las coordenadas esféricas degeneran en los polos y en el origen, y la esfera unidad tiene medida nula en $\mathbb{R}^3$. Con Riemann habría que discutir cada una de esas patologías; con Lebesgue basta decir "salvo un conjunto de medida cero" y seguir. Es el mismo argumento que permite escribir $\int_0^\infty$ sin distinguir integrales impropias: para Lebesgue no hay integrales impropias, sólo funciones que están o no están en L^1.

Cuidado

La hipótesis de que Φ sea un difeomorfismo no se puede rebajar a "derivable". Si Φ no es inyectiva, la fórmula cuenta varias veces la misma región: para $\Phi(u) = u^2$ sobre $(-1, 1)$, el miembro derecho da $\int_{-1}^1 f(u^2)\, 2|u| \, \mathrm{d}u$, que es el doble de $\int_0^1 f(y) \, \mathrm{d}y$. La versión correcta en ese caso incluye un factor de multiplicidad, la llamada fórmula del área. Y si $\det D\Phi$ se anula en un conjunto de medida positiva, la imagen puede colapsar y perderse información sin remedio.

Pausa para pensar

El teorema dice que la medida de Lebesgue se transforma por $|\det|$ bajo aplicaciones lineales. Tómese T ortogonal: entonces $|\det T| = 1$ y la medida no cambia, lo que demuestra que la medida de Lebesgue es invariante por rotaciones sin haberlo pedido nunca en su construcción, que sólo hablaba de rectángulos con lados paralelos a los ejes. ¿Es casualidad? No: la medida de Lebesgue es, salvo constante multiplicativa, la única medida sobre los borelianos de $\mathbb{R}^n$ que es invariante por traslaciones y finita en compactos, y esa unicidad fuerza la invariancia por cualquier isometría.

31.6. Los espacios L^p y la completitud

Idea clave

Intuición primero. Los espacios L^p son espacios vectoriales cuyos elementos son funciones, con una norma que mide "el tamaño" de una función. Son la continuación natural del capítulo 7 en dimensión infinita, y L^2 es especialmente cómodo porque su norma viene de un producto interno: en L^2 se puede hablar de funciones ortogonales, de proyecciones y de bases, exactamente igual que con vectores.
El teorema de Riesz–Fischer es el que justifica todo el rodeo por la teoría de la medida: dice que estos espacios son *completos*, es decir, que toda sucesión de Cauchy de funciones converge a una función del espacio. Con la integral de Riemann esto sería falso, y no por

poco: se pueden construir sucesiones de funciones Riemann-integrables que se apiñan pero cuyo límite natural no es Riemann-integrable. Completitud es lo que permite resolver ecuaciones por aproximaciones sucesivas y estar seguro de que el límite existe, y por eso es la propiedad sobre la que descansan el análisis funcional, las ecuaciones en derivadas parciales y la mecánica cuántica.

Definición 31.43. $L^p = \{f \text{ medibles} : \int |f|^p < \infty\}$ con la norma $\|f\|_p = (\int |f|^p)^{1/p}$, identificando funciones iguales *casi en todas partes* (fuera de un conjunto de medida cero: la igualdad que la integral puede ver).

Teorema 31.44 (Riesz–Fischer)**.** *L^p es* completo*: toda sucesión de Cauchy converge en la norma a un elemento de L^p.*

Idea. De una sucesión de Cauchy se extrae una subsucesión con $\|f_{n_{k+1}} - f_{n_k}\| < 2^{-k}$; la serie telescópica $\sum |f_{n_{k+1}} - f_{n_k}|$ tiene integral finita (convergencia monótona), luego converge casi en todas partes: el límite puntual f existe c.t.p., y la convergencia dominada remata $\|f_n - f\| \to 0$. Los tres teoremas de convergencia trabajando en equipo: esta demostración es su hora estelar. □ □

Idea clave

Qué dice esto en cristiano. Riesz–Fischer dice que los espacios L^p están completos: no tienen agujeros, toda sucesión de Cauchy converge dentro. Eso es lo que los convierte en espacios de Banach y permite hacer análisis en ellos, igual que la completitud de $\mathbb{R}$ permite hacer cálculo. Sin ese teorema, L^2 no sería un espacio de Hilbert y no habría series de Fourier ni mecánica cuántica. Es la razón profunda de que la integral de Lebesgue sustituyera a la de Riemann: con Riemann, el espacio análogo *no* es completo.

Ejemplo 31.45 (Completitud, y el agujero que deja Riemann)**. (a) El agujero de Riemann.** Sea $\{q_1, q_2, \dots\}$ una enumeración de los racionales de $[0,1]$ y $f_n = \mathbf{1}_{\{q_1,\dots,q_n\}}$. Cada f_n es Riemann-integrable (finitas discontinuidades) y la sucesión es de Cauchy en la norma L^1, porque $\|f_n - f_m\|_1 = 0$. Pero su límite puntual es la función de Dirichlet, que *no* es Riemann-integrable. Con Riemann el espacio tiene un agujero justo ahí; con Lebesgue, el límite está dentro y vale 0 casi por doquier.

(b) Las normas p son distintas. Sea $f(x) = x^{-1/2}$ en $(0,1]$. Está en L^1, con $\int_0^1 x^{-1/2} = 2$, pero no en L^2, porque $\int_0^1 x^{-1}$ diverge. En un intervalo acotado, $L^2 \subset L^1$; en $\mathbb{R}$ entero no hay inclusión en ninguno de los dos sentidos.

(c) El caso contrario. La función $f(x) = \frac{1}{x}$ en $[1,\infty)$ está en L^2, pues $\int_1^\infty x^{-2} = 1$, y no está en L^1, pues $\int_1^\infty x^{-1}$ diverge. Junto con (b), esto muestra que ninguno de los dos espacios contiene al otro sobre dominios no acotados.

(d) L^2 es el único que es de Hilbert. Solo con $p = 2$ la norma proviene de un producto interno, $\langle f, g\rangle = \int f\bar{g}$, y solo entonces hay ortogonalidad, proyecciones y bases ortonormales. Por eso las series de Fourier viven en L^2 y no en L^1 ni en L^3: la identidad de Parseval, $\|f\|_2^2 = \sum |c_n|^2$, no tiene análogo en los otros.

Cuidado

[Los elementos de L^p son clases, no funciones] Dos funciones que difieren en un conjunto de medida nula tienen distancia cero, así que técnicamente $\|\cdot\|_p$ no es una norma sobre funciones sino sobre clases de equivalencia. Por eso "la función de L^2" no tiene un valor bien definido en un punto concreto: preguntar cuánto vale $f(3)$ no tiene sentido. Segundo aviso: convergencia en L^p no implica convergencia puntual. La sucesión de las jorobas viajeras que barren $[0,1]$ converge a 0 en L^1 y no converge en ningún punto; lo que sí se puede extraer es una *subsucesión* que converja casi por doquier.

Observación 31.46. L^2, con el producto $\langle f,g\rangle = \int f\bar{g}$, es entonces un espacio de *Hilbert*: el espacio euclidiano del capítulo 7 en dimensión infinita, ahora *completo*. Aquí es donde las series de Fourier del capítulo 12 encuentran su casa definitiva: para *cualquier* $f \in L^2[-\pi,\pi]$, la serie de Fourier converge a f en norma L^2 (la base trigonométrica es ortonormal completa), la identidad de Parseval es el teorema de Pitágoras en dimensión infinita, y no hacen falta las hipótesis de suavidad que arrastrábamos. La convergencia *puntual* c.t.p. para L^2 es un teorema monumental (Carleson, 1966); para L^1 es *falsa* (Kolmogórov construyó una $f \in L^1$ con serie de Fourier divergente en todas partes): la frontera fina del análisis armónico pasa exactamente entre estos dos espacios.

Regla: decidir a qué espacios L^p pertenece una función

La pregunta "¿está f en L^p?" es la pregunta "¿converge $\int |f|^p$?". Y esa integral solo puede fallar en dos sitios: donde f se dispara y donde el dominio se va al infinito. El método es separar los dos y tratarlos aparte.

1. **Localiza los puntos conflictivos.** Marca las singularidades de f (donde $|f|\to\infty$) y, si el dominio no está acotado, marca también el infinito. Parte el dominio en trozos, uno por conflicto, más un resto donde f es acotada y el trozo tiene medida finita — ese resto nunca da problemas.

2. **Usa el patrón potencia, que resuelve casi todo.** Estos dos hechos hay que saberlos de memoria:

$$\int_0^1 x^{-a}\,\mathrm{d}x < \infty \iff a<1, \qquad \int_1^\infty x^{-a}\,\mathrm{d}x < \infty \iff a>1.$$

   Nótese que las condiciones son *opuestas*: cerca de cero conviene que el exponente sea suave, cerca del infinito conviene que sea agresivo. Ninguna potencia sirve en ambos sitios a la vez, y de ahí sale casi toda la teoría.

3. **Aplica el patrón al integrando elevado a p.** Si $|f(x)|\sim x^{-a}$ cerca de un punto, entonces $|f|^p \sim x^{-ap}$, y la condición se convierte en $ap<1$ (cerca de 0) o $ap>1$ (cerca de ∞). Despejando p se obtiene el rango exacto.

4. **Recuerda la asimetría entre medida finita e infinita.** Si $\mu(E)<\infty$ y $p<q$, entonces $L^q(E)\subset L^p(E)$: los espacios están *encajados*, y el grande es el de exponente pequeño. Si $\mu(E)=\infty$, no hay inclusión en ninguna de las dos direcciones.

5. **Para acotar productos, usa Hölder.** Si $\frac{1}{p}+\frac{1}{q}=1$ entonces $\|fg\|_1 \le \|f\|_p\|g\|_q$. El caso $p=q=2$ es Cauchy–Schwarz. El truco más útil es el aparentemente tonto $g\equiv 1$: sobre un conjunto de medida finita da
$$\|f\|_1 \le \|f\|_p\,\mu(E)^{1-1/p},$$
que es exactamente la inclusión del paso 4, demostrada en una línea.
6. **Control.** Antes de dar el rango de p, comprueba los extremos por separado sustituyendo el valor crítico: casi siempre el extremo *no* pertenece, porque ahí el exponente cae justo en -1 y el logaritmo diverge.

Ejemplo 31.47 (Tres funciones y sus rangos de p)**. (a) $f(x)=x^{-1/2}$ en $(0,1)$.** Solo hay un conflicto, en $x=0$. Aquí $|f|^p = x^{-p/2}$, integrable en $(0,1)$ si y solo si $p/2<1$, es decir $p<2$. Luego
$$f\in L^p(0,1) \iff 1\le p<2.$$
En particular $f\in L^1$ (y de hecho $\int_0^1 x^{-1/2}\mathrm{d}x = 2$) pero $f\notin L^2$, porque $\int_0^1 x^{-1}\mathrm{d}x$ diverge. Este es el ejemplo mínimo de que la inclusión $L^2\subset L^1$ sobre medida finita es *estricta*.

(b) $f(x)=\dfrac{x^{-1/2}}{1+x}$ en $(0,\infty)$. Ahora hay dos conflictos. Cerca de 0 el denominador tiende a 1, luego $f(x)\sim x^{-1/2}$ y hace falta $p/2<1$, o sea $p<2$. Cerca de ∞ el denominador manda y $f(x)\sim x^{-3/2}$, luego hace falta $3p/2>1$, o sea $p>2/3$. Ambas a la vez:
$$f\in L^p(0,\infty) \iff \tfrac{2}{3}<p<2.$$
Un rango *acotado por los dos lados*: fuera de él no está, ni por arriba ni por abajo. Esto solo puede pasar en medida infinita, y es la ilustración limpia del paso 4. De propina, para $p=1$ la integral tiene valor cerrado:
$$\int_0^\infty \frac{\mathrm{d}x}{\sqrt{x}\,(1+x)} = \pi,$$
como se ve con la sustitución $x=t^2$, que la convierte en $2\int_0^\infty \mathrm{d}t/(1+t^2)$.

(c) Cauchy–Schwarz en acción. Acotemos $\int_0^1 \sqrt{x}\,\mathrm{e}^x\,\mathrm{d}x$ sin calcularla. Con $p=q=2$,
$$\int_0^1 \sqrt{x}\,\mathrm{e}^x\,\mathrm{d}x \le \Big(\int_0^1 x\,\mathrm{d}x\Big)^{1/2}\Big(\int_0^1 \mathrm{e}^{2x}\,\mathrm{d}x\Big)^{1/2} = \Big(\tfrac12\Big)^{1/2}\Big(\tfrac{\mathrm{e}^2-1}{2}\Big)^{1/2}.$$
Numéricamente el segundo factor es $\sqrt{3{,}194528}$ y la cota vale 1,26383, mientras que el valor verdadero de la integral es 1,25563. Un error del orden del 0,7 por ciento: la desigualdad es casi una igualdad aquí, porque $\sqrt{x}$ y e^x son ambas crecientes y por tanto "casi proporcionales" en el sentido que mide Cauchy–Schwarz.

Regla: usar Hölder, Minkowski y Jensen, y elegir bien los exponentes

Las tres desigualdades básicas de los espacios L^p se enuncian en un renglón, pero casi todo el arte está en elegir los exponentes. Este es el procedimiento.

1. **Identifica qué tipo de desigualdad necesitas.** Si tienes un *producto* dentro de la integral, es Hölder. Si tienes una *suma* dentro de una norma, es Minkowski. Si tienes una *función convexa* aplicada a un promedio, es Jensen.
2. **Hölder: emparéjalos.** Si $\frac{1}{p}+\frac{1}{q}=1$ con $1\le p\le\infty$, entonces $\int|fg|\,d\mu\le\|f\|_p\|g\|_q$. Los exponentes p y q se llaman conjugados; $p=q=2$ da Cauchy–Schwarz, y $p=1$, $q=\infty$ da la cota trivial $\int|fg|\le\|g\|_\infty\int|f|$.
3. **Para que Hölder sea una igualdad, iguala las potencias.** Hay igualdad exactamente cuando $|f|^p$ y $|g|^q$ son proporcionales. Si estás buscando la mejor cota posible, resuelve esa proporcionalidad: te dirá qué p elegir.
4. **Minkowski: $\|f+g\|_p\le\|f\|_p+\|g\|_p$, y solo si $p\ge1$.** Es lo que convierte $\|\cdot\|_p$ en una norma. Para $p<1$ la desigualdad se invierte y L^p deja de ser un espacio normado.
5. **Jensen: exige medida de probabilidad.** Si $\mu(X)=1$ y φ es convexa, entonces $\varphi(\int f\,d\mu)\le\int\varphi\circ f\,d\mu$. Si tu medida no es de probabilidad, normalízala dividiendo por $\mu(X)$ antes de aplicarla, o la desigualdad será falsa.
6. **Inclusión en medida finita.** Si $\mu(X)<\infty$ y $r<p$, aplica Hölder a $|f|^r\cdot1$ con exponentes p/r y su conjugado: obtienes $\|f\|_r\le\|f\|_p\,\mu(X)^{1/r-1/p}$, es decir $L^p\subseteq L^r$. En medida infinita esto es falso y no hay ninguna inclusión entre los L^p.
7. **Verifica con un caso extremo.** Prueba tu desigualdad con f constante, o con una indicadora, o con $f=g$. Si falla ahí, has intercambiado los exponentes o has olvidado normalizar.

Ejemplo 31.48 (Seis desigualdades, y la elección de exponentes que las hace exactas)**.** Trabajamos casi siempre en $(0,1)$ con la medida de Lebesgue, donde las potencias x^{-a} dan todos los ejemplos que hacen falta.

(a) La familia de referencia. Para $f_a(x)=x^{-a}$ en $(0,1)$ y $p\ge1$,

$$\|f_a\|_p^p=\int_0^1 x^{-ap}\,dx=\frac{1}{1-ap}\quad\text{si } ap<1,$$

y la integral diverge si $ap\ge1$. Es decir, $x^{-a}\in L^p(0,1)$ exactamente cuando $p<1/a$. Con $a=1/2$: $x^{-1/2}\in L^p$ si y solo si $p<2$; en particular está en L^1 (con $\|f\|_1=2$) pero no en L^2.

(b) Cauchy–Schwarz, y por qué no es óptima aquí. Estimemos $I=\int_0^1 x^{-1/3}x^{-1/4}\,dx$. El valor exacto es

$$I=\int_0^1 x^{-7/12}\,dx=\frac{1}{1-7/12}=\frac{12}{5}=2{,}4.$$

Con $p=q=2$: $\|x^{-1/3}\|_2=(1/(1-2/3))^{1/2}=\sqrt{3}\approx1{,}732$ y $\|x^{-1/4}\|_2=(1/(1-1/2))^{1/2}=\sqrt{2}\approx1{,}414$, luego Hölder da $I\le\sqrt{6}\approx2{,}449$. Es correcta, pero no exacta.

(c) Los exponentes que hacen exacta la desigualdad. Aplicamos el paso 3: queremos $|f|^p\propto|g|^q$, es decir $x^{-p/3}\propto x^{-q/4}$, o sea $p/3=q/4$. Junto con $\frac{1}{p}+\frac{1}{q}=1$ esto da $p=\frac{7}{4}$ y

$q = \frac{7}{3}$. Entonces ambas normas involucran la misma integral:

$$\|f\|_{7/4}^{7/4} = \int_0^1 x^{-7/12} = \frac{12}{5}, \qquad \|g\|_{7/3}^{7/3} = \int_0^1 x^{-7/12} = \frac{12}{5},$$

y por tanto $\|f\|_{7/4}\|g\|_{7/3} = (12/5)^{4/7}(12/5)^{3/7} = (12/5)^1 = \frac{12}{5} = 2{,}4$. Con la elección correcta de exponentes, Hölder no pierde absolutamente nada.

(d) Interpolación: $L^p \cap L^q$ está contenido en todo lo intermedio. Si $\frac{1}{r} = \frac{\theta}{p} + \frac{1-\theta}{q}$ con $\theta \in [0,1]$, Hölder aplicado a $|f|^r = |f|^{\theta r}|f|^{(1-\theta)r}$ da

$$\|f\|_r \le \|f\|_p^{\theta}\, \|f\|_q^{1-\theta}.$$

Con $f(x) = x^{-1/3}$ en $(0,1)$, $p = 1$, $q = 2$ y $r = 3/2$ resulta $\theta = 1/3$. Los números: $\|f\|_1 = 3/2$, $\|f\|_2 = \sqrt{3}$, $\|f\|_{3/2} = 2^{2/3} \approx 1{,}587$, y la cota es $(3/2)^{1/3}(\sqrt{3})^{2/3} \approx 1{,}651$. Se cumple, con holgura moderada.

(e) Jensen en acción. En $(0,1)$ la medida de Lebesgue *es* de probabilidad, así que con $\varphi(t) = e^t$, que es convexa, y $f(x) = x$:

$$e^{\int_0^1 x\,dx} = e^{1/2} \approx 1{,}6487 \;\le\; \int_0^1 e^x\,dx = e - 1 \approx 1{,}7183.$$

El mismo argumento con $\varphi(t) = -\log t$ sobre $f > 0$ da la desigualdad entre medias geométrica y aritmética en su forma integral: $\exp\big(\int \log f\big) \le \int f$.

(f) Minkowski falla para $p < 1$. Sean $f = \mathbf{1}_{[0,1/2]}$ y $g = \mathbf{1}_{(1/2,1]}$ en $[0,1]$, y tomemos $p = 1/2$. Entonces $\|f\|_{1/2} = \big(\int_0^{1/2} 1\big)^2 = \frac{1}{4}$, e igual para g, mientras que $f + g = \mathbf{1}_{[0,1]}$ tiene $\|f+g\|_{1/2} = 1$. Es decir

$$1 = \|f+g\|_{1/2} \;>\; \|f\|_{1/2} + \|g\|_{1/2} = \tfrac{1}{2}.$$

La «norma» $\|\cdot\|_{1/2}$ no satisface la desigualdad triangular: $L^{1/2}$ es un espacio métrico completo con la distancia $\int |f-g|^{1/2}$, pero no es normado.

Cuidado

La inclusión $L^p \subseteq L^r$ para $r < p$ vale *solo en medida finita*. En $(0,\infty)$ no hay ninguna inclusión en ninguna dirección: $x^{-1/2}\mathbf{1}_{(0,1)}$ está en L^1 pero no en L^2, y $x^{-1}\mathbf{1}_{(1,\infty)}$ está en L^2 pero no en L^1. El primero falla por la singularidad en 0, el segundo por la cola en el infinito, y los exponentes que arreglan una cosa estropean la otra.

Idea clave

Hölder no es un truco: es la afirmación de que L^q es el espacio dual de L^p cuando $1 \le p < \infty$. Cada $g \in L^q$ define el funcional lineal $f \mapsto \int fg$, la desigualdad dice que su norma es a lo sumo $\|g\|_q$, y el caso de igualdad del paso 3 dice que es exactamente $\|g\|_q$. Ese es el puente hacia el análisis funcional (véase el capítulo 32), donde L^2 ocupa el lugar privilegiado por ser el único autodual, con $p = q = 2$.

31.7. La transformada de Fourier, con demostraciones

Idea clave

Intuición primero. La transformada de Fourier es un *prisma*: entra luz blanca $f(x)$ y sale su arcoíris $\hat{f}(\xi)$, que dice cuánta onda pura $e^{i\xi x}$ de cada frecuencia lleva dentro la señal. La fórmula de $\hat{f}$ es el análisis (descomponer en colores) y el teorema de inversión es la síntesis (recombinar el arcoíris y recuperar la luz blanca): que la ida y vuelta no pierda nada es exactamente lo que hay que demostrar, y Plancherel añade que tampoco se pierde *energía*. Con ese lenguaje, los teoremas del capítulo se leen solos: derivar acentúa los colores agudos (multiplicar por $i\xi$), suavizar los apaga, y el principio de incertidumbre dice que ninguna señal puede ser a la vez un destello brevísimo y una nota purísima.

Del círculo a la recta: para $f \in L^1(\mathbb{R})$ definimos

$$\hat{f}(\xi) = \int_{-\infty}^{\infty} f(x)\, e^{-i\xi x}\, \mathrm{d}x.$$

Las series descomponen lo periódico en armónicos discretos; la transformada descompone lo no periódico en un *continuo* de frecuencias.

Ejemplo 31.49 (Los dos ejemplos generadores)**.** *(a) La gaussiana es su propia transformada*: si $f(x) = e^{-x^2/2}$, derivando bajo la integral (¡ahora legal por convergencia dominada!) e integrando por partes:

$$\frac{\mathrm{d}\hat{f}}{\mathrm{d}\xi} = \int (-ix) e^{-x^2/2} e^{-i\xi x} \mathrm{d}x = i \int (e^{-x^2/2})' e^{-i\xi x} \mathrm{d}x = -\xi\, \hat{f}(\xi).$$

La transformada satisface la *misma* EDO $\hat{f}' = -\xi \hat{f}$ que la gaussiana (capítulo 12: separable), así que $\hat{f}(\xi) = \hat{f}(0) e^{-\xi^2/2}$, y $\hat{f}(0) = \int e^{-x^2/2} = \sqrt{2\pi}$ (¡la integral de los capítulos 9, 11 y 30!):

$$\widehat{e^{-x^2/2}} = \sqrt{2\pi}\, e^{-\xi^2/2}.$$

(b) Para $f(x) = e^{-|x|}$: dos integrales elementales dan $\hat{f}(\xi) = \frac{2}{1+\xi^2}$: la lorentziana. Obsérvese el patrón que gobierna todo el capítulo: f con esquina (no derivable en 0) $\to$ $\hat{f}$ decae como ξ^{-2}; la gaussiana suave $\to$ decaimiento supergaussiano. *La suavidad de f es el decaimiento de $\hat{f}$*, y viceversa: derivar en x es multiplicar por $i\xi$ en frecuencia ($\widehat{f'} = i\xi \hat{f}$, por partes), así que cada derivada disponible compra un factor de decaimiento.

Lema 31.50 (Riemann–Lebesgue)**.** *Si $f \in L^1$, entonces $\hat{f}$ es continua y $\hat{f}(\xi) \to 0$ cuando $|\xi| \to \infty$.*

Demostración. Continuidad: convergencia dominada (el integrando se domina por $|f|$). Decaimiento: para el indicador de un intervalo es un cálculo ($\widehat{\mathbf{1}_{[a,b]}} = \frac{e^{-ia\xi} - e^{-ib\xi}}{i\xi}$, que decae como $\frac{1}{|\xi|}$); las combinaciones de indicadores (funciones escalonadas) son densas en L^1 (así se construyó la integral), y la estimación uniforme $|\hat{f} - \hat{g}| \le \|f - g\|_1$ transporta el límite. Densidad más estimación uniforme: el argumento estándar del análisis, usado aquí por primera vez en el libro con todas sus piezas justificadas. □ □

Teorema 31.51 (Inversión)**.** *Si $f \in L^1$ y también $\hat f \in L^1$, entonces en casi todo punto (y en todo punto de continuidad)*

$$f(x) = \frac{1}{2\pi}\int_{-\infty}^{\infty} \hat f(\xi)\, e^{i\xi x}\,\mathrm{d}\xi.$$

Demostración. El obstáculo: sustituir la definición de $\hat f$ y "cambiar el orden" produce $\int e^{i\xi(x-y)}\mathrm{d}\xi$, que no converge. El remedio: amortiguar con la gaussiana. Para $\varepsilon > 0$,

$$I_\varepsilon(x) = \frac{1}{2\pi}\int \hat f(\xi)\, e^{-\varepsilon^2\xi^2/2}\, e^{i\xi x}\,\mathrm{d}\xi.$$

Ahora sí es legal intercambiar (Fubini–Tonelli: el integrando doble es absolutamente integrable gracias al amortiguador), y la integral interior es la transformada de la gaussiana, ya calculada:

$$I_\varepsilon(x) = \int f(y)\,\frac{1}{\sqrt{2\pi}\,\varepsilon}\, e^{-(x-y)^2/2\varepsilon^2}\,\mathrm{d}y = (f * G_\varepsilon)(x),$$

la convolución de f con una gaussiana de masa 1 concentrándose en el origen: una *identidad aproximada* (el núcleo de Poisson del capítulo 30 era otra). Cuando $\varepsilon \to 0$: el lado izquierdo tiende a $\frac{1}{2\pi}\int \hat f e^{i\xi x}$ por convergencia dominada (dominante $|\hat f| \in L^1$: aquí se usa la hipótesis), y el derecho tiende a f en L^1 y en los puntos de continuidad (la gaussiana concentra el promedio). Los dos límites de la misma cantidad son iguales. □ □

Ejemplo 31.52 (La inversión en acción, cerrando un círculo con el capítulo 30)**.** Invirtiendo la lorentziana: $\frac{1}{2\pi}\int \frac{2}{1+\xi^2} e^{i\xi x}\mathrm{d}\xi = e^{-|x|}$, es decir

$$\int_{-\infty}^{\infty} \frac{\cos(\xi x)}{1+\xi^2}\,\mathrm{d}\xi = \pi e^{-|x|}:$$

exactamente la integral que calculamos por residuos en el capítulo 30. Dos maquinarias independientes —el teorema de inversión y el contorno semicircular— obligadas a coincidir: cuando eso pasa, uno empieza a creerse la matemática.

Teorema 31.53 (Plancherel)**.** *Para $f \in L^1 \cap L^2$: $\int |f|^2 = \frac{1}{2\pi}\int |\hat f|^2$, y la transformada se extiende a una isometría (salvo $\sqrt{2\pi}$) de L^2 en sí mismo.*

Idea. Aplicar inversión a $g = f * \tilde f$ con $\tilde f(x) = \overline{f(-x)}$, cuyo valor en 0 es $\int |f|^2$ y cuya transformada es $|\hat f|^2 \geq 0$; la fórmula de inversión en $x = 0$ es exactamente la identidad. La extensión a todo L^2 es densidad más Riesz–Fischer: la completitud, otra vez, haciendo el trabajo silencioso. □ □

Ejemplo 31.54. Con $f = \mathbf{1}_{[-1,1]}$: $\hat f(\xi) = \frac{2\sin\xi}{\xi}$, y Plancherel da

$$\int_{-\infty}^{\infty} \frac{\sin^2\xi}{\xi^2}\,\mathrm{d}\xi = \frac{2\pi\cdot\int_{-1}^{1} 1}{4} = \pi.$$

(Y de paso, con inversión en $x = 0$ y un límite: $\int \frac{\sin\xi}{\xi} = \pi$, la integral de Dirichlet del capítulo 30, tercera derivación en el libro.) El par indicador↔sinc es el corazón del *teorema*

del muestreo de Shannon: una señal con frecuencias en $[-W, W]$ se reconstruye exactamente de sus muestras cada $\frac{\pi}{W}$: la razón matemática de que el audio digital funcione, y el motivo de que el CD muestree a 44.1 kHz (el doble del límite audible más margen).

Teorema 31.55 (Principio de incertidumbre de Heisenberg)**.** *Si* $\|f\|_2 = 1$*, entonces*

$$\Big(\int x^2 |f(x)|^2 \mathrm{d}x\Big)\Big(\frac{1}{2\pi}\int \xi^2 |\hat f(\xi)|^2 \mathrm{d}\xi\Big) \geq \frac{1}{4},$$

con igualdad exactamente para las gaussianas.

Demostración. Para f suave que decae (el caso general por densidad): integrando por partes,

$$1 = \int |f|^2 = -\int x\,\frac{\mathrm{d}}{\mathrm{d}x}|f|^2 = -2\,\mathrm{Re}\int x\, f'\,\bar f,$$

y Cauchy–Schwarz (¡capítulo 7!) da $1 \leq 2\|xf\|_2\,\|f'\|_2$. Pero $\|f'\|_2^2 = \frac{1}{2\pi}\|\widehat{f'}\|_2^2 = \frac{1}{2\pi}\|\xi\hat f\|_2^2$ (Plancherel y $\widehat{f'} = i\xi\hat f$): elevar al cuadrado. La igualdad en Cauchy–Schwarz exige $f' \propto xf$: la EDO de la gaussiana una vez más. □ □

Observación 31.56. Una función no puede estar simultáneamente concentrada en posición y en frecuencia: no es un defecto de los aparatos sino un teorema sobre L^2. En mecánica cuántica ($\hat f$ = amplitud en momento) es *el* principio de incertidumbre, con la gaussiana como estado de mínima incertidumbre (el estado fundamental del oscilador armónico del capítulo 12: todo cierra); en procesamiento de señales dice que no existe una señal breve y de banda angosta, la disyuntiva que gobierna desde el radar hasta los ataques de compresión de audio.

Ejemplo 31.57 (La ecuación del calor en la recta, resuelta del todo)**.** Para $u_t = k u_{xx}$ con dato inicial $u(x,0) = f(x)$: transformando en x, la EDP se vuelve la EDO $\partial_t \hat u = -k\xi^2 \hat u$ (derivar es multiplicar: la transformada *diagonaliza* ∂_x, como prometía el capítulo 12), así que $\hat u = \hat f(\xi) e^{-k\xi^2 t}$, y anti-transformando (producto → convolución, gaussiana → gaussiana):

$$u(x,t) = \int_{-\infty}^{\infty} f(y)\,\frac{1}{\sqrt{4\pi k t}}\,e^{-(x-y)^2/4kt}\,\mathrm{d}y.$$

El *núcleo del calor* en persona: en el círculo la separación de variables daba una serie; en la recta, la transformada da esta fórmula cerrada. Comprobación directa: el núcleo satisface la EDP (derivar y agrupar) y es una identidad aproximada cuando $t \to 0^+$ (recupera f: el mismo lema que usamos en la inversión, tercera aparición). La suavización instantánea, la irreversibilidad, la velocidad infinita de propagación: todo se lee de la fórmula.

Regla: calcular una transformada de Fourier sin integrar

Con la convención de este capítulo, $\hat f(\xi) = \int_{\mathbb{R}} f(x) e^{-i\xi x}\mathrm{d}x$. Casi ninguna transformada se calcula por fuerza bruta: se parte de cuatro transformadas conocidas y se llega al resto con seis propiedades.

1. **Comprueba que** $f \in L^1$**.** Si $\int |f| < \infty$, la integral converge absolutamente, $\hat f$ es

continua y acotada por $\|f\|_1$, y tiende a 0 en el infinito (Riemann–Lebesgue).

2. **Memoriza el diccionario base:**

$$\widehat{\chi_{[-a,a]}}(\xi) = \frac{2\sin(a\xi)}{\xi}, \qquad \widehat{e^{-|x|}}(\xi) = \frac{2}{1+\xi^2}, \qquad \widehat{e^{-x^2/2}}(\xi) = \sqrt{2\pi}\, e^{-\xi^2/2}.$$

3. **Aplica traslación y modulación:** $\widehat{f(x-a)} = e^{-ia\xi}\hat{f}(\xi)$ y $\widehat{e^{ibx}f(x)} = \hat{f}(\xi-b)$. Trasladar en el tiempo es girar la fase; girar la fase es trasladar en frecuencia.

4. **Aplica el escalado:** $\widehat{f(ax)} = \frac{1}{|a|}\hat{f}(\xi/a)$. Comprimir la señal ensancha el espectro: es el principio de incertidumbre en su forma más elemental.

5. **Convierte cálculo en álgebra:** $\widehat{f'} = i\xi\hat{f}$ y $\widehat{f*g} = \hat{f}\cdot\hat{g}$. Derivar se vuelve multiplicar, y convolucionar se vuelve multiplicar: por eso la transformada resuelve ecuaciones diferenciales y ecuaciones de convolución.

6. **Usa la simetría:** si f es real y par, $\hat{f}$ es real y par; si f es real e impar, $\hat{f}$ es imaginaria pura e impar. Comprobar la paridad del resultado es el control más barato que existe.

7. **Control por Plancherel:** $\int |f|^2\,\mathrm{d}x = \frac{1}{2\pi}\int |\hat{f}|^2\,\mathrm{d}\xi$. Si tu candidato no cumple esta igualdad, hay un factor mal puesto.

Ejemplo 31.58 (Cuatro transformadas obtenidas del diccionario)**. (a) El pulso rectangular y el control de Plancherel.** Con $f = \chi_{[-1,1]}$,

$$\hat{f}(\xi) = \int_{-1}^{1} e^{-i\xi x}\mathrm{d}x = \left[\frac{e^{-i\xi x}}{-i\xi}\right]_{-1}^{1} = \frac{2\sin\xi}{\xi}.$$

Es real y par, como exige el paso 6, y en $\xi = 0$ vale $2 = \int f$, que es lo que debe. Verifiquemos el paso 7: $\int |f|^2 = 2$, y por otro lado $\frac{1}{2\pi}\int \frac{4\sin^2\xi}{\xi^2}\mathrm{d}\xi = \frac{4\pi}{2\pi} = 2$, usando $\int_{\mathbb{R}} \frac{\sin^2\xi}{\xi^2}\mathrm{d}\xi = \pi$. Coinciden. Y de paso hemos *deducido* esa integral clásica sin calcularla.

(b) Un pulso desplazado y comprimido. Sea $g(x) = \chi_{[1,3]}(x)$. Es el pulso $\chi_{[-1,1]}$ trasladado dos unidades a la derecha, así que por el paso 3

$$\hat{g}(\xi) = e^{-2i\xi}\cdot\frac{2\sin\xi}{\xi}.$$

El módulo no cambió: $|\hat{g}| = |\hat{f}|$. Trasladar una señal no altera su contenido de frecuencias, solo sus fases; ésta es la razón por la que el oído no distingue el instante exacto en que empieza una nota a partir de su timbre.

(c) El triángulo, sin integrar. La convolución de $\chi_{[-1,1]}$ consigo misma es la función triangular

$$\Lambda(x) = (2-|x|)^+ \quad \text{en } [-2,2],$$

así que por el paso 5

$$\hat{\Lambda}(\xi) = \left(\frac{2\sin\xi}{\xi}\right)^2 = \frac{4\sin^2\xi}{\xi^2} \ \geq 0.$$

Obsérvese el patrón anunciado en la sección: el pulso tiene saltos y su transformada decae como ξ^{-1}; el triángulo es continuo (una derivada más) y decae como ξ^{-2}. Cada grado de suavidad compra un factor de decaimiento.

(d) Una integral difícil, gratis. La lorentziana $\widehat{e^{-|x|}} = \frac{2}{1+\xi^2}$ combinada con Plancherel da

$$\int_{\mathbb{R}} e^{-2|x|}\mathrm{d}x = \frac{1}{2\pi}\int_{\mathbb{R}} \frac{4}{(1+\xi^2)^2}\mathrm{d}\xi \quad\Longrightarrow\quad 1 = \frac{2}{\pi}\int_{\mathbb{R}} \frac{\mathrm{d}\xi}{(1+\xi^2)^2},$$

es decir $\int_{\mathbb{R}} \frac{\mathrm{d}\xi}{(1+\xi^2)^2} = \frac{\pi}{2}$. Comprobación numérica: la cuadratura da 1,5707963, y $\pi/2 =$ 1,5707963. La integral se ha obtenido sin sustituciones trigonométricas ni residuos, solo emparejando un par transformada–función.

Idea clave

Las tres transformadas del diccionario no son ejemplos sueltos: son los tres arquetipos de decaimiento. El pulso (discontinuo) decae como $1/\xi$; la exponencial $e^{-|x|}$ (continua con esquina) como $1/\xi^2$; la gaussiana (C^∞) más rápido que cualquier potencia. Sabiendo dónde cae tu función en esa escala ya sabes cómo será su espectro antes de calcular nada.

Cuidado

Los factores 2π dependen de la convención. Con $\hat f(\xi) = \int f(x)e^{-2\pi i\xi x}\mathrm{d}x$ la fórmula de Plancherel queda sin constante, $\|f\|_2 = \|\hat f\|_2$, y la gaussiana $e^{-\pi x^2}$ es exactamente su propia transformada. Antes de usar una tabla, comprueba con qué convención está escrita; el paso 7 detecta cualquier discrepancia inmediatamente.

Regla: distinguir los cinco modos de convergencia y saber cuál implica cuál

Una sucesión de funciones puede «converger» en al menos cinco sentidos distintos, y confundirlos es la fuente número uno de errores en teoría de la medida. Este es el mapa completo.

1. **Escribe las cinco definiciones y no las mezcles.** Convergencia *puntual*: $f_n(x) \to f(x)$ para todo x. *En casi todo punto* (c.t.p.): lo mismo salvo en un conjunto de medida cero. *Uniforme*: $\sup_x |f_n(x) - f(x)| \to 0$. *En L^p*: $\|f_n - f\|_p \to 0$. *En medida*: $\mu\{x : |f_n(x) - f(x)| > \varepsilon\} \to 0$ para todo $\varepsilon > 0$.

2. **Memoriza las implicaciones que sí valen siempre.** Uniforme $\Rightarrow$ puntual $\Rightarrow$ c.t.p. Y en cualquier espacio: $L^p \Rightarrow$ en medida (por Chebyshev: $\mu\{|f_n - f| > \varepsilon\} \le \varepsilon^{-p}\|f_n - f\|_p^p$).

3. **Recuerda las dos que necesitan medida finita.** Si $\mu(X) < \infty$: c.t.p. $\Rightarrow$ en medida, y uniforme $\Rightarrow L^p$. En medida infinita ambas fallan.

4. **Usa la recíproca parcial de la convergencia en medida.** Si $f_n \to f$ en medida, existe una *subsucesión* $f_{n_k} \to f$ c.t.p. Nunca toda la sucesión, solo una subsucesión: este es el único puente de vuelta que hay.

5. **Para pasar de c.t.p. a L^p necesitas un control extra.** O bien una dominante integrable (convergencia dominada), o bien la convergencia de las normas (teorema de Scheffé: si $f_n \to f$ c.t.p. y $\|f_n\|_1 \to \|f\|_1$, entonces $\|f_n - f\|_1 \to 0$). Sin uno de los dos, la implicación es falsa.

6. **Aplica Egorov cuando quieras uniformidad.** Si $\mu(X) < \infty$ y $f_n \to f$ c.t.p., para todo $\delta > 0$ hay un conjunto E con $\mu(E) < \delta$ tal que la convergencia es uniforme fuera de E. La convergencia c.t.p. es «uniforme salvo un conjunto tan pequeño como quieras», pero nunca salvo medida cero.

7. **Cuando dudes, prueba los tres contraejemplos canónicos.** La máquina de escribir, la joroba que se escapa y el pico que crece. Entre los tres separan todas las flechas que no valen.

Ejemplo 31.59 (Los tres contraejemplos que separan todos los modos de convergencia)**.** Cada uno de estos ejemplos está diseñado para romper exactamente una flecha. Vale la pena aprenderlos de memoria.

(a) La máquina de escribir: converge en L^p y en medida, pero en ningún punto. En $[0,1]$, escribe cada $n \geq 1$ como $n = 2^k + j$ con $0 \leq j < 2^k$, y define

$$f_n = \mathbf{1}_{[\,j/2^k,\,(j+1)/2^k\,]}.$$

Los primeros términos son $\mathbf{1}_{[0,1]}$; luego $\mathbf{1}_{[0,1/2]}, \mathbf{1}_{[1/2,1]}$; luego los cuatro cuartos $\mathbf{1}_{[0,1/4]}, \dots, \mathbf{1}_{[3/4,1]}$; y así sucesivamente. La franja barre el intervalo una y otra vez, cada vez más estrecha. Entonces

$$\|f_n\|_p^p = 2^{-k} \longrightarrow 0,$$

luego $f_n \to 0$ en todo L^p con $p < \infty$, y por el paso 2 también en medida. Pero fijado cualquier $x \in [0,1]$, la sucesión $f_n(x)$ vale 1 infinitas veces (una vez por cada nivel k) y 0 infinitas veces: no converge *en ningún punto*.

Cuidado

Este ejemplo mata la implicación «$L^p \Rightarrow$ c.t.p.». Y confirma el paso 4: la subsucesión $f_{2^k} = \mathbf{1}_{[0,2^{-k}]}$ sí converge a 0 en todo punto salvo $x = 0$. Una subsucesión, no la sucesión.

(b) La joroba que se escapa: converge puntualmente en todas partes, pero en ningún L^p. En $\mathbb{R}$ con la medida de Lebesgue, sea $g_n = \mathbf{1}_{[n,\,n+1]}$. Para cada x fijo, en cuanto $n > x$ se tiene $g_n(x) = 0$; luego $g_n \to 0$ puntualmente en todo $\mathbb{R}$. Sin embargo

$$\|g_n\|_p = 1 \quad \text{para todo } n \text{ y todo } p, \qquad \mu\{|g_n| > 1/2\} = 1,$$

así que no hay convergencia ni en L^p ni en medida. La masa no desaparece: se marcha al infinito. Este ejemplo muestra por qué el paso 3 exige $\mu(X) < \infty$, y por qué la convergencia dominada necesita una dominante: aquí la menor dominante posible es $\sup_n g_n = \mathbf{1}_{[1,\infty)}$, que no es integrable.

(c) El pico que crece: c.t.p. y en medida, pero no en L^1. En $(0,1]$, sea $h_n = n\,\mathbf{1}_{(0,1/n]}$. Para cada $x > 0$, en cuanto $n > 1/x$ se tiene $h_n(x) = 0$: convergencia puntual a 0 en todo $(0,1]$. Y $\mu\{h_n > \varepsilon\} \le 1/n \to 0$, luego también en medida. Pero

$$\int_0^1 h_n \, d\mu = n \cdot \frac{1}{n} = 1 \quad \text{para todo } n.$$

La masa tampoco desaparece: se concentra. Aquí se ve el lema de Fatou con desigualdad *estricta*:

$$\int \liminf h_n = \int 0 = 0 \;<\; 1 = \liminf \int h_n.$$

(d) Afinando el pico: convergencia en L^1 pero no en L^2. Sea ahora $h_n = \sqrt{n}\,\mathbf{1}_{(0,1/n]}$. Un cálculo directo:

$$\|h_n\|_1 = \frac{\sqrt{n}}{n} = n^{-1/2} \to 0, \qquad \|h_n\|_2^2 = \frac{n}{n} = 1.$$

Converge a 0 en L^1 y no en L^2. En general, con $h_n = n^{\alpha}\mathbf{1}_{(0,1/n]}$ se tiene $\|h_n\|_p^p = n^{\alpha p - 1}$, de modo que hay convergencia en L^p exactamente cuando $\alpha p < 1$. Ajustando α se separa cualquier par de exponentes que uno quiera.

(e) Uniforme sin L^1: hace falta medida finita. En $\mathbb{R}$, sea $u_n = \frac{1}{n}\mathbf{1}_{[0,n]}$. Entonces $\sup_x |u_n(x)| = 1/n \to 0$, es decir convergencia uniforme a 0; y sin embargo $\|u_n\|_1 = \frac{1}{n} \cdot n = 1$ para todo n. La altura baja, la base crece, y el área no se mueve. En un espacio de medida finita esto no puede pasar, porque $\|u_n\|_1 \le \mu(X) \sup |u_n|$.

(f) Egorov, con números. En $[0,1]$, $f_n(x) = x^n$ converge a 0 en casi todo punto (falla solo en $x = 1$), pero no uniformemente: $\sup_{[0,1]} x^n = 1$ para todo n. Egorov dice que basta retirar un conjunto pequeño. Retirando $E = (0{,}99,\, 1]$, de medida 0,01, sobre $[0,\, 0{,}99]$ se tiene

$$\sup_{[0,\,0{,}99]} x^n = (0{,}99)^n,$$

que baja de 10^{-3} a partir de $n = 688$, ya que $(0{,}99)^{687} \approx 1{,}003 \cdot 10^{-3}$ y $(0{,}99)^{688} \approx 9{,}93 \cdot 10^{-4}$. Cuanto menor sea δ, más lenta la convergencia uniforme, pero uniforme al fin.

Pausa para pensar

Compara (b) y (c) con cuidado: en ambos la integral vale constantemente 1 mientras el límite puntual es 0. La diferencia es *adónde* se va la masa. En (b) escapa horizontalmente hacia el infinito; en (c) escapa verticalmente concentrándose en un punto. Son las dos únicas formas de perder masa en el límite, y la hipótesis de dominación de Lebesgue está diseñada precisamente para prohibir las dos a la vez.

Idea clave

Este mapa es la razón de ser de L^p como *espacio*: la convergencia en norma es una noción de convergencia que respeta la estructura métrica, y el teorema de Riesz–Fischer garantiza que L^p es completo con ella. La convergencia c.t.p., en cambio, no proviene de ninguna métrica sobre el espacio de funciones medibles cuando la medida es infinita. La

convergencia en medida sí es metrizable en medida finita, con la distancia $\int \frac{|f-g|}{1+|f-g|}\,d\mu$, y bajo ella el espacio también resulta completo.

31.8. Cierre definitivo: Fubini, Radon–Nikodym y diferenciación

Tres teoremas que usamos y debíamos, con sus ideas completas.

31.8.1. Medidas producto y Fubini–Tonelli

La medida producto en $\mathbb{R}^2$ se construye repitiendo Carathéodory con rectángulos $A \times B$ de "área" $m(A)m(B)$ como cubrimientos básicos.

Teorema 31.60 (Tonelli y Fubini)**.** *Para f medible en $\mathbb{R}^2$: si $f \geq 0$ (Tonelli) o si $\iint |f| < \infty$ (Fubini), entonces*

$$\iint f \,\mathrm{d}(m \times m) = \int \Big(\int f(x,y)\,\mathrm{d}y \Big)\mathrm{d}x = \int \Big(\int f(x,y)\,\mathrm{d}x \Big)\mathrm{d}y$$

(en Fubini, las rebanadas están definidas para casi todo x).

Idea. La identidad es trivial para indicadores de rectángulos; el *teorema de la clase monótona* la propaga a todos los conjuntos medibles (la familia donde vale es cerrada bajo límites monótonos, por convergencia monótona a cada lado de la igualdad, y contiene a los rectángulos); funciones simples por linealidad, no negativas por convergencia monótona otra vez, e integrables partiendo $f = f^+ - f^-$. La maquinaria de convergencia del capítulo no es un accesorio de Fubini: *es* la demostración. □ □

Idea clave

Qué dice esto en cristiano. Tonelli y Fubini dicen cuándo una integral doble se puede calcular integrando primero en una variable y luego en la otra, y cuándo se puede cambiar el orden. La diferencia entre los dos es la hipótesis: Tonelli solo exige que la función sea no negativa, y entonces todo vale aunque el resultado sea infinito; Fubini exige que la integral del valor absoluto sea finita. En la práctica se usan en pareja: primero Tonelli sobre $|f|$ para comprobar que la integral converge, y luego Fubini sobre f para hacer el cálculo en el orden que convenga.

Ejemplo 31.61 (Cambiar el orden, y cuándo no se puede)**. (a) Un cambio de orden que salva el cálculo.** La integral $\int_0^1 \int_0^1 \mathrm{e}^{-xy}\,\mathrm{d}y\,\mathrm{d}x$ tiene integrando positivo y acotado, así que Tonelli permite integrar en cualquier orden. Integrando primero en y sale $\int_0^1 \frac{1-\mathrm{e}^{-x}}{x}\mathrm{d}x$, que no tiene primitiva elemental pero sí valor numérico 0,796600.

(b) La estrategia de las dos etapas. Para $\iint_{\mathbb{R}^2} \mathrm{e}^{-(x^2+y^2)}$, primero Tonelli sobre la función (que es positiva) da $\left(\int_{\mathbb{R}} \mathrm{e}^{-x^2}\mathrm{d}x\right)^2$. Como ese producto vale π en coordenadas polares,

se deduce el valor de la integral de Gauss: $\int_{\mathbb{R}} \mathrm{e}^{-x^2} = \sqrt{\pi} = 1{,}772454$. Es el truco clásico, y depende enteramente de Tonelli.

(c) El contraejemplo estándar. Sea $f(x,y) = \frac{x^2-y^2}{(x^2+y^2)^2}$ en $(0,1] \times (0,1]$. Integrando primero en y y luego en x sale $\frac{\pi}{4}$; en el orden contrario sale $-\frac{\pi}{4}$. Los dos órdenes dan resultados distintos, y ninguno de los dos es "el" valor de la integral doble.

(d) Por qué no hay contradicción. Porque $\iint |f| = \infty$: la hipótesis de Fubini falla. La función toma valores positivos y negativos que se cancelan de manera distinta según el orden, y sin integrabilidad absoluta el resultado depende del camino. Es el mismo fenómeno que la reordenación de series condicionalmente convergentes.

Cuidado

[Comprobar $\iint |f| < \infty$ no es opcional] El apartado (c) es la advertencia entera: aplicar Fubini sin comprobar la hipótesis produce resultados sin sentido, y no hay ningún aviso durante el cálculo. La comprobación es barata, porque para $|f|$ vale Tonelli y se puede integrar en el orden que sea. Segundo aviso: ambos teoremas exigen medidas σ-finitas; con medidas exóticas fallan. Y tercero: que las dos integrales iteradas coincidan no demuestra que la integral doble exista; puede haber casos donde coincidan por casualidad y la función no sea integrable.

Cuidado

Sin integrabilidad absoluta, los dos órdenes pueden diferir: $f(x,y) = \frac{x^2-y^2}{(x^2+y^2)^2}$ en $(0,1)^2$ da $\frac{\pi}{4}$ en un orden y $-\frac{\pi}{4}$ en el otro (cálculo directo con arctan). No es paradoja: $\iint |f| = \infty$, y las dos iteradas son dos maneras condicionales de sumar una serie no absolutamente convergente: la reordenación de Riemann del capítulo 6, en versión continua. El mismo fenómeno exacto que ya vimos con $\int \frac{\sin x}{x}$: condicional $\neq$ absoluto, ahora en dos dimensiones.

Regla: intercambiar el orden de integración con Fubini y Tonelli

Cambiar $\iint$ por $\iint$ es uno de los trucos más rentables del análisis, pero es también donde más se equivoca la gente. La secuencia correcta tiene un orden fijo: primero Tonelli, después Fubini.

1. **Comprueba que las medidas son σ-finitas** y que f es medible en el producto. En la práctica, con la medida de Lebesgue en $\mathbb{R}^n$ esto siempre se cumple.
2. **Si $f \geq 0$, aplica Tonelli sin más trámite.** Para funciones no negativas los tres valores —la integral doble y las dos iteradas— coinciden siempre, aunque valgan $+\infty$. No hay nada que verificar.
3. **Si f cambia de signo, verifica primero la integrabilidad.** Calcula $\iint |f|$, y hazlo con Tonelli (que $|f| \geq 0$ es gratis). Solo si sale finito puedes usar Fubini.

4. **Ahora sí, aplica Fubini:** si $\iint |f| < \infty$, entonces f es integrable, ambas iteradas existen y coinciden con la doble.
5. **El truco estándar: fabrica una segunda variable.** Escribe una constante incómoda como una integral —típicamente $\frac{1}{x} = \int_0^\infty e^{-xt}\,\mathrm{d}t$ para $x > 0$, o $\frac{1}{1-u} = \sum u^n$— y usa Tonelli para intercambiar.
6. **Dibuja el recinto al cambiar el orden.** Los límites de integración se recalculan a partir del dibujo, no a partir de la fórmula anterior: es aquí donde se pierden la mitad de los puntos en un examen.
7. **Control:** si las dos iteradas te salen distintas, no has encontrado una paradoja; has encontrado que $\iint |f| = +\infty$. Compruébalo y quedará explicado.

Ejemplo 31.62 (Tres intercambios legales y uno ilegal)**. (a) Una serie que sale de una integral doble.** En $[0,1]^2$ el integrando $\frac{1}{1-xy}$ es no negativo, así que Tonelli permite hacer lo que se quiera. Desarrollando en serie geométrica e intercambiando suma e integral (Tonelli para series, todo positivo):

$$\int_0^1\!\!\int_0^1 \frac{\mathrm{d}x\,\mathrm{d}y}{1-xy} = \sum_{n\ge 0} \int_0^1\!\!\int_0^1 x^n y^n\,\mathrm{d}x\,\mathrm{d}y = \sum_{n\ge 0} \frac{1}{(n+1)^2} = \frac{\pi^2}{6}.$$

Control numérico: $\pi^2/6 \approx 1{,}644934$, y la cuadratura directa da 1,644934. La integral es impropia en la esquina $(1,1)$ y aun así el resultado es finito; con Tonelli eso no exige ninguna justificación adicional.

(b) La integral de Dirichlet. Para calcular $\int_0^\infty \frac{\sin x}{x}\,\mathrm{d}x$ se aplica el paso 5 con un amortiguador e^{-sx} que garantiza la integrabilidad absoluta:

$$I(s) = \int_0^\infty e^{-sx}\frac{\sin x}{x}\,\mathrm{d}x = \int_0^\infty\!\!\int_s^\infty e^{-tx}\sin x\,\mathrm{d}t\,\mathrm{d}x = \int_s^\infty \frac{\mathrm{d}t}{1+t^2} = \frac{\pi}{2} - \arctan s.$$

Haciendo $s \to 0^+$ se obtiene $\int_0^\infty \frac{\sin x}{x}\,\mathrm{d}x = \frac{\pi}{2}$. El amortiguador no es un adorno: sin él $\int_0^\infty |\sin x|/x\,\mathrm{d}x = +\infty$ y Fubini no sería aplicable.

(c) La gaussiana, otra vez. Con $I = \int_0^\infty e^{-x^2}\mathrm{d}x$, el integrando de I^2 es no negativo y Tonelli da directamente

$$I^2 = \int_0^\infty\!\!\int_0^\infty e^{-(x^2+y^2)}\mathrm{d}x\,\mathrm{d}y = \int_0^{\pi/2}\!\!\int_0^\infty e^{-r^2} r\,\mathrm{d}r\,\mathrm{d}\theta = \frac{\pi}{2}\cdot\frac{1}{2} = \frac{\pi}{4},$$

luego $I = \sqrt{\pi}/2$. Obsérvese que aquí no hubo que comprobar absolutamente nada: ésa es la ventaja de Tonelli sobre Fubini.

(d) El intercambio ilegal. Sea

$$f(x,y) = \frac{x^2-y^2}{(x^2+y^2)^2} \qquad \text{en } [0,1]^2.$$

Una primitiva respecto de x es $-\frac{x}{x^2+y^2}$, de modo que

$$\int_0^1\!\left(\int_0^1 f\,\mathrm{d}x\right)\mathrm{d}y = \int_0^1 \frac{-1}{1+y^2}\,\mathrm{d}y = -\frac{\pi}{4},$$

mientras que por la antisimetría $f(y,x) = -f(x,y)$ el otro orden da $+\frac{\pi}{4}$. Las dos iteradas existen, son finitas y son *distintas*. No hay contradicción con Fubini: lo que ocurre es que $\iint |f| = +\infty$, porque cerca del origen $|f| \sim r^{-2}$ y en el plano $\int_0 r^{-2} \cdot r \,\mathrm{d}r = \int_0 r^{-1}\mathrm{d}r$ diverge. El paso 3 lo habría detectado antes de escribir la primera línea.

Cuidado

El error típico es aplicar Fubini y *después* preguntarse si valía. Los pasos 2–3 no son burocracia: en (d) el resultado difiere en $\pi/2$ según el orden. Con integrandos que cambian de signo, la comprobación $\iint |f| < \infty$ es obligatoria, y se hace siempre con Tonelli.

Pausa para pensar

¿Qué le pasa exactamente a (d) en el origen? Las masas positivas (donde $|x| > |y|$) y negativas (donde $|x| < |y|$) son ambas infinitas, y cada orden de integración las cancela en una proporción diferente. Es el análogo bidimensional del teorema de reordenación de Riemann: una serie condicionalmente convergente puede sumar lo que uno quiera si se reordenan sus términos.

31.8.2. Radon–Nikodym: toda medida razonable tiene densidad

Teorema 31.63 (Radon–Nikodym)**.** *Si ν, μ son medidas σ-finitas y ν es* absolutamente continua *respecto de μ (todo conjunto μ-nulo es ν-nulo), entonces existe una densidad $h \geq 0$ medible, única c.t.p., con*

$$\nu(E) = \int_E h \,\mathrm{d}\mu \quad \textit{para todo } E; \qquad h = \frac{\mathrm{d}\nu}{\mathrm{d}\mu}.$$

Idea (von Neumann). En el Hilbert $L^2(\mu+\nu)$, el funcional $f \mapsto \int f \,\mathrm{d}\nu$ es lineal y continuo; el teorema de representación de Riesz (proyección ortogonal del capítulo 7, en dimensión infinita: el punto más cercano de un subespacio cerrado existe por completitud) da g con $\int f \,\mathrm{d}\nu = \int f g \,\mathrm{d}(\mu+\nu)$ para toda f. Despejar algebraicamente esta identidad produce $h = \frac{g}{1-g}$ con las propiedades pedidas (la continuidad absoluta garantiza $g < 1$ c.t.p.). El álgebra lineal euclidiana, madurada en L^2, demuestra el teorema central de la teoría de la medida. □ □

Idea clave

Qué dice esto en cristiano. Si cada vez que μ dice "este conjunto no pesa nada" también ν lo dice, entonces ν se puede escribir como "μ multiplicada por una densidad". Esa densidad es la derivada de Radon–Nikodým. Es la generalización exacta de la idea de función de densidad de probabilidad, y es lo que permite convertir preguntas sobre medidas en preguntas sobre funciones, que son mucho más manejables.

Ejemplo 31.64 (Densidades y cuándo no existen)**. (a) Una densidad familiar.** Si $\nu(A) = \int_A \frac{1}{\sqrt{2\pi}} \mathrm{e}^{-x^2/2} \mathrm{d}x$ y μ es la medida de Lebesgue, entonces $\mathrm{d}\nu/\mathrm{d}\mu$ es la campana de Gauss. Toda

función de densidad de probabilidad del cálculo elemental es una derivada de Radon–Nikodým disfrazada.

(b) Sin absoluta continuidad, no hay densidad. Sea $\nu = \delta_0$, la delta de Dirac, y μ la de Lebesgue. El conjunto $\{0\}$ tiene medida de Lebesgue cero y $\delta_0(\{0\}) = 1$, luego ν no es absolutamente continua respecto de μ y no existe ninguna función f con $\delta_0(A) = \int_A f \, \mathrm{d}\mu$. La delta no es una función, por más que la notación de la física insista.

(c) La medida de Cantor. La medida asociada a la función de Cantor está concentrada en un conjunto de medida de Lebesgue cero y no tiene átomos: no es absolutamente continua ni discreta. Es el tercer tipo, el *singular continuo*, y su existencia es lo que hace que el teorema de descomposición de Lebesgue tenga tres piezas y no dos.

(d) Cambio de medida. En finanzas, pasar de la probabilidad real a la de riesgo neutro es exactamente multiplicar por una derivada de Radon–Nikodým; el teorema de Girsanov da su forma explícita para el movimiento browniano. Sin la hipótesis de absoluta continuidad en los dos sentidos (equivalencia de medidas) el cambio no sería reversible y el modelo permitiría arbitraje.

Observación 31.65. El nombre $\frac{\mathrm{d}\nu}{\mathrm{d}\mu}$ no es decorativo: se comporta como derivada (regla de la cadena $\frac{\mathrm{d}\nu}{\mathrm{d}\lambda} = \frac{\mathrm{d}\nu}{\mathrm{d}\mu}\frac{\mathrm{d}\mu}{\mathrm{d}\lambda}$). Este teorema es la licencia oficial de dos gestos cotidianos: en probabilidad, "X tiene densidad f" significa exactamente $\frac{\mathrm{d}P_X}{\mathrm{d}m} = f$; y la *esperanza condicional* —el corazón de las martingalas y de todo el aprendizaje automático probabilista— se *define* como una derivada de Radon–Nikodym. La contraparte es el teorema de descomposición de Lebesgue: toda medida se parte en (absolutamente continua) + (singular), como la función de Cantor ilustra.

31.8.3. El teorema de diferenciación de Lebesgue

Teorema 31.66. *Si $f \in L^1$, entonces para casi todo x:*

$$\lim_{r\to 0} \frac{1}{2r} \int_{x-r}^{x+r} f(t)\,\mathrm{d}t = f(x).$$

En consecuencia, $F(x) = \int_a^x f$ es derivable c.t.p. con $F' = f$: el teorema fundamental del cálculo, en su forma final.

Idea. Para f continua es el teorema fundamental del capítulo 6. El caso general se controla con la *función maximal* de Hardy–Littlewood $Mf(x) = \sup_r \frac{1}{2r}\int_{x-r}^{x+r} |f|$ y su desigualdad $m\{Mf > \lambda\} \le \frac{3}{\lambda}\|f\|_1$, que se demuestra con un *lema de cubrimiento de Vitali*: de cualquier colección de intervalos se extraen disjuntos que, triplicados, cubren la unión (ordenar por tamaño y ser voraz). Densidad de las continuas en L^1 más la maximal transportan el resultado: el patrón "denso + estimación uniforme" por tercera vez en el capítulo. □ □

Ejemplo 31.67 (Derivar la integral)**. (a) El caso continuo.** Si f es continua en x, el promedio $\frac{1}{2r}\int_{x-r}^{x+r} f$ tiende a $f(x)$ sin ninguna dificultad: es el teorema fundamental del cálculo de toda la vida. Lo notable del enunciado de Lebesgue es que vale para funciones que no son continuas en ningún punto.

(b) Un escalón. Sea $f = \mathbf{1}_{[0,\infty)}$. En $x = 0$ el promedio simétrico vale $1/2$ para todo r, y no converge a $f(0) = 1$. No hay contradicción: el teorema dice "para casi todo x", y $\{0\}$ tiene medida cero.

(c) La función de Dirichlet. $f = \mathbf{1}_{\mathbb{Q}}$ es igual a 0 en casi todo punto, luego sus promedios valen 0 y el teorema se cumple trivialmente en todos los irracionales. Con la integral de Riemann esta función ni siquiera es integrable.

(d) Puntos de Lebesgue. El enunciado fuerte dice que $\frac{1}{2r}\int_{x-r}^{x+r} |f(y) - f(x)|\,\mathrm{d}y \to 0$ en casi todo x. Esos puntos se llaman de Lebesgue, y la diferencia con el enunciado débil importa: garantizan que f es "aproximadamente continua" allí, lo que se usa constantemente en teoría de operadores maximales y en análisis armónico.

31.9. Variación acotada, continuidad absoluta y el teorema fundamental

El teorema de diferenciación de Lebesgue dice que una función integrable se recupera casi en todo punto promediando en intervalos pequeños. Falta la mitad simétrica de la historia: dada una función F, ¿cuándo es cierto que $F(b) - F(a) = \int_a^b F'$? La respuesta separa tres clases de funciones que la intuición del cálculo elemental confunde alegremente, y el contraejemplo que las separa ya apareció en este capítulo: la escalera del diablo.

Definición 31.68 (Variación total)**.** Sea $F\colon [a,b] \to \mathbb{R}$. Su *variación total* es

$$V_a^b(F) = \sup\Big\{\sum_{i=1}^{n} |F(x_i) - F(x_{i-1})| \;:\; a = x_0 < x_1 < \cdots < x_n = b\Big\}.$$

Si $V_a^b(F) < \infty$ se dice que F es de *variación acotada*. La función F es *absolutamente continua* si para todo $\varepsilon > 0$ existe $\delta > 0$ tal que, para cualquier familia finita de intervalos disjuntos (a_i, b_i) con $\sum (b_i - a_i) < \delta$, se tiene $\sum |F(b_i) - F(a_i)| < \varepsilon$.

Ejemplo 31.69 (Variación total de funciones concretas)**. (a) Monótonas.** Si F es creciente, $V_a^b(F) = F(b) - F(a)$, porque todas las diferencias son positivas y la suma telescópica. Para $F(x) = x^2$ en $[0,2]$ sale 4.

(b) Oscilantes. Para $F(x) = \sin x$ en $[0, 2\pi]$, la función sube 1, baja 2 y sube 1: $V = 4$. En general, para una función de clase C^1 se cumple $V_a^b(F) = \int_a^b |F'|$, y aquí $\int_0^{2\pi} |\cos x|\,\mathrm{d}x = 4$.

(c) Variación infinita. Sea $F(x) = x\sin(1/x)$ para $x \neq 0$ y $F(0) = 0$. Es continua en $[0,1]$ y su variación total es infinita: evaluando en los puntos $x_k = 2/((2k+1)\pi)$, las oscilaciones suman una serie del tipo $\sum 1/k$, que diverge. Continua no implica de variación acotada.

(d) La escalera del diablo. La función de Cantor es continua, creciente y va de 0 a 1, luego su variación total es exactamente 1. Y sin embargo su derivada vale 0 en casi todo punto, de modo que $\int_0^1 F' = 0 \neq 1 = F(1) - F(0)$. La fórmula $V = \int |F'|$ falla, y eso es precisamente lo que la continuidad absoluta viene a arreglar.

Teorema 31.70 (Jordan, y el teorema fundamental del cálculo de Lebesgue)**.** *Una función es de variación acotada si y sólo si es diferencia de dos funciones crecientes; en tal caso*

es derivable en casi todo punto y $F' \in L^1[a,b]$. *Y son equivalentes:* (i) F *es absolutamente continua;* (ii) F *es derivable c.t.p.,* $F' \in L^1$ *y* $F(x) = F(a) + \int_a^x F'(t)\,\mathrm{d}t$ *para todo* x*;* (iii) F *es la integral indefinida de alguna función de* L^1.

Idea clave

Hay tres condiciones encajadas, cada una estrictamente más fuerte que la anterior: continua, de variación acotada, absolutamente continua. Ser de variación acotada garantiza que la derivada existe casi siempre y es integrable, pero no que la integral la reconstruya: puede perderse masa en un conjunto de medida cero. La continuidad absoluta es exactamente la condición que impide esa fuga, y por eso es la hipótesis correcta —no la continuidad, no la derivabilidad c.t.p.— en el teorema fundamental del cálculo, en la integración por partes y en el cambio de variable unidimensional.

Regla: decidir a cuál de las tres clases pertenece una función

1. **Si es monótona, ya es de variación acotada.** Su variación total es simplemente $|F(b)-F(a)|$, porque todos los sumandos tienen el mismo signo y la suma telescopa.
2. **Si es C^1 en un intervalo cerrado, calcula $V = \int_a^b |F'|$.** Esa fórmula vale para toda función absolutamente continua y reduce el supremo sobre particiones a una integral corriente.
3. **Si es lipschitziana, es absolutamente continua.** Basta tomar $\delta = \varepsilon/L$; el recíproco es falso, y $\sqrt{x}$ es el contraejemplo estándar.
4. **Sospecha de las oscilaciones infinitas.** Una función continua puede no ser de variación acotada si oscila cada vez más rápido cerca de un punto; el criterio es sumar las alturas de las oscilaciones y ver si la serie converge.
5. **Sospecha de las funciones singulares.** Si F es continua y creciente pero $F' = 0$ en casi todo punto, es de variación acotada y *no* absolutamente continua: toda su variación se concentra en un conjunto nulo.
6. **Aplica el criterio decisivo.** Comprueba si $F(b) - F(a) = \int_a^b F'$. Si la igualdad falla —y siempre falla por defecto, nunca por exceso, en el caso creciente— la función no es absolutamente continua.
7. **Usa la descomposición.** Toda función de variación acotada se escribe como suma de una absolutamente continua, una de saltos y una singular continua; identificar las tres piezas responde la pregunta de una vez.

Ejemplo 31.71 (Cuatro funciones y sus variaciones)**. (a) Una parábola.** Para $F(x) = x^2 - x$ en $[0,1]$, la derivada se anula en $x = \frac{1}{2}$ con $F(\frac{1}{2}) = -\frac{1}{4}$, y como $F(0) = F(1) = 0$ la variación total es $\frac{1}{4} + \frac{1}{4} = \frac{1}{2}$. La fórmula integral confirma: $\int_0^1 |2x-1|\,\mathrm{d}x = \frac{1}{2}$. Es C^1, luego lipschitziana, luego absolutamente continua.

(b) Un seno completo. Para $F(x) = \operatorname{sin} x$ en $[0, 2\pi]$, la función sube de 0 a 1, baja de 1 a

-1 y vuelve a subir de -1 a 0: la variación total es $1+2+1=4$, y en efecto $\int_0^{2\pi} |\cos x|\,\mathrm{d}x = 4$. Contar los tramos de monotonía es siempre el camino más corto.

(c) Continua pero no de variación acotada. Sea $F(x) = x\sin(1/x)$ con $F(0)=0$, que es continua en $[0,1]$. En los puntos $x_k = \frac{2}{(2k+1)\pi}$ el seno vale ± 1, así que $|F(x_k)| = \frac{2}{(2k+1)\pi}$, y la suma de esas alturas se comporta como la serie armónica: las sumas parciales son 0,2122, 0,3395, 0,4305, 0,5012, 0,5591, 0,6081, 0,6505 para $k \le 1,\dots,7$, y crecen sin cota. La función es continua en un compacto —luego uniformemente continua— y aun así su variación es infinita.

(d) Absolutamente continua con derivada no acotada. Para $F(x)=\sqrt{x}$ en $[0,1]$, la derivada $\frac{1}{2\sqrt{x}}$ explota en el origen y F no es lipschitziana. Pero $\int_0^1 \frac{\mathrm{d}x}{2\sqrt{x}} = 1 = F(1)-F(0)$, así que F sí es absolutamente continua. Moraleja: la continuidad absoluta no exige derivada acotada, sólo derivada integrable.

Ejemplo 31.72 (La escalera del diablo como frontera exacta)**. (a) Lo que sí cumple.** La función de Cantor $c\colon [0,1]\to[0,1]$ es continua y creciente, luego de variación acotada con $V_0^1(c) = c(1)-c(0) = 1$. Por el teorema de Jordan es derivable en casi todo punto, y de hecho $c'(x)=0$ para todo x fuera del conjunto de Cantor, es decir, en un conjunto de medida 1.

(b) Lo que falla. Entonces $\int_0^1 c'(t)\,\mathrm{d}t = 0$, mientras que $c(1)-c(0)=1$. La igualdad del teorema fundamental se incumple por el máximo margen posible, y la conclusión es que c no es absolutamente continua. Toda la subida ocurre sobre el conjunto de Cantor, que tiene medida nula: la función transporta masa desde un conjunto invisible.

(c) La lectura en términos de medidas. La medida de Lebesgue–Stieltjes asociada a c es una medida de probabilidad sobre $[0,1]$ sin átomos y concentrada en un conjunto de medida de Lebesgue nula: es *singular*. Por el teorema de Radon–Nikodym no puede tener densidad, y esa es exactamente la razón analítica de que c' no la reconstruya.

(d) La descomposición general. Cualquier F creciente se escribe como $F_{\mathrm{ac}} + F_{\mathrm{saltos}} + F_{\mathrm{sing}}$. La primera pieza es una integral indefinida, la segunda una suma de escalones, la tercera un objeto tipo Cantor. La función $F(x)=x$ tiene sólo la primera; la función escalón de Heaviside sólo la segunda; la escalera del diablo sólo la tercera. Cualquier función real es una mezcla de las tres.

Cuidado

Es tentador creer que una función derivable en *todo* punto cumple el teorema fundamental. No es cierto. La función $F(x) = x^2\sin(1/x^2)$ con $F(0)=0$ es derivable en todos los puntos de $[0,1]$, incluido el origen, y sin embargo su derivada no es integrable y F no es de variación acotada. Derivabilidad puntual en todas partes es más fuerte que derivabilidad c.t.p. y aun así no basta: lo que hace falta es controlar el *tamaño* de la derivada, no su existencia.

Pausa para pensar

La integración por partes $\int_a^b FG' = [FG]_a^b - \int_a^b F'G$ se demuestra en los cursos elementales suponiendo que ambas funciones son C^1. El enunciado correcto sólo pide que F y G sean absolutamente continuas, y esa mejora no es cosmética: en probabilidad, F suele ser una función de distribución que no es derivable en ningún sentido clásico. ¿Y si ni siquiera es continua? Entonces hay que sustituir $\int F'G$ por una integral de Stieltjes respecto de la medida que F genera, y el resultado sigue siendo válido; la sección siguiente construye esa medida.

31.10. Medidas de Lebesgue–Stieltjes y medidas signadas

La medida de Lebesgue mide longitudes, y eso es una elección entre muchas. Si se admite que la recta pueda tener densidad variable, o incluso masas puntuales, se obtiene toda una familia de medidas parametrizada por funciones crecientes; es la construcción que hace de la probabilidad un caso particular de este capítulo. Y si además se permite que la "masa" sea negativa, aparecen las medidas signadas, cuya teoría es sorprendentemente rígida: siempre se parten en una parte positiva y otra negativa que viven en trozos disjuntos del espacio.

Definición 31.73 (Medida de Lebesgue–Stieltjes y medida signada)**.** Sea $F\colon \mathbb{R} \to \mathbb{R}$ creciente y continua por la derecha. Existe una única medida boreliana μ_F con

$$\mu_F((a,b]) = F(b) - F(a) \qquad \text{para todos } a < b,$$

llamada medida de Lebesgue–Stieltjes de F. Por otra parte, una *medida signada* sobre una σ-álgebra es una función ν numerablemente aditiva con valores en $\mathbb{R}$ (o en $[-\infty, \infty)$) y $\nu(\varnothing) = 0$.

Ejemplo 31.74 (Medidas a partir de funciones)**. (a) La medida de Lebesgue.** Con $F(x) = x$ se obtiene $\mu((a,b]) = b - a$: la longitud usual. Todo el caso clásico es un caso particular de esta construcción.

(b) Una masa puntual. Con $F = \mathbf{1}_{[0,\infty)}$ se obtiene $\mu = \delta_0$, que asigna 1 al conjunto $\{0\}$ y 0 a todo lo que no lo contenga. El salto de la función se convierte en un átomo de la medida, y el tamaño del salto es la masa del átomo.

(c) Una mezcla. Con $F(x) = \frac{1}{2}x + \frac{1}{2}\mathbf{1}_{[0,\infty)}(x)$ en $[0,1]$ se obtiene media unidad repartida uniformemente y media concentrada en el origen. Es el modelo de una variable aleatoria que con probabilidad $1/2$ vale exactamente cero y con probabilidad $1/2$ es uniforme.

(d) Por qué continua por la derecha. Si F no lo fuera, la medida no quedaría determinada: dos funciones que difieren sólo en el valor que toman en sus puntos de salto darían la misma medida. Exigir continuidad por la derecha elige un representante canónico, y es el mismo convenio que hace que las funciones de distribución de probabilidad se definan como $P(X \le x)$ y no como $P(X < x)$.

Teorema 31.75 (Hahn–Jordan)**.** *Si ν es una medida signada sobre $(X, \mathcal{A})$, existe una partición $X = P \sqcup N$ con P positivo y N negativo, es decir $\nu(E) \ge 0$ para todo medible $E \subseteq P$*

y $\nu(E) \le 0$ para todo medible $E \subseteq N$. Definiendo $\nu^+(E) = \nu(E \cap P)$ y $\nu^-(E) = -\nu(E \cap N)$ se obtienen dos medidas positivas mutuamente singulares con

$$\nu = \nu^+ - \nu^-, \qquad |\nu| = \nu^+ + \nu^-,$$

y esta descomposición es la única con esa propiedad de mutua singularidad.

Idea clave

Una función creciente es un "repartidor de masa": la cantidad de masa que hay en el intervalo $(a, b]$ es cuánto sube F al recorrerlo. Si F sube de forma continua y suave, la masa está esparcida y hay densidad; si F da un salto, hay un átomo, una masa puntual; si F sube continuamente pero sólo sobre un conjunto nulo, la masa es singular sin ser puntual. La continuidad por la derecha es una convención de bordes, elegida para que $(a, b]$ sea el intervalo natural. Y la descomposición de Hahn dice que una medida con signo no puede mezclar signos de forma inextricable: siempre hay una frontera que separa la zona donde suma de la zona donde resta.

Regla: trabajar con una medida de Stieltjes o con una medida signada

1. **Para una F dada, mide primero los intervalos semiabiertos.** La receta $\mu_F((a, b]) = F(b) - F(a)$ determina todo lo demás por la unicidad de la extensión de Carathéodory.
2. **Localiza los saltos: son los átomos.** $\mu_F(\{c\}) = F(c) - F(c^-)$, la altura del salto en c. Si F es continua, no hay átomos y todo punto tiene medida nula.
3. **Donde F sea C^1, la densidad es F'.** Ahí $\mathrm{d}\mu_F = F'(x)\,\mathrm{d}x$ y toda integral respecto de μ_F se calcula como una integral corriente con un peso.
4. **Desconfía si F es continua pero singular.** La escalera del diablo no tiene átomos ni densidad, y su medida es un objeto legítimo que ninguna de las dos descripciones anteriores captura.
5. **Para una medida signada, busca primero el conjunto positivo.** Si $\nu(E) = \int_E g\,\mathrm{d}\mu$, entonces $P = \{g \ge 0\}$ y $N = \{g < 0\}$; el caso general se demuestra tomando supremos, pero el ejemplo con densidad es el que hay que tener en la cabeza.
6. **Calcula ν^+, ν^- y $|\nu|$ por separado.** Son $\int_{E\cap P} g$, $-\int_{E\cap N} g$ y $\int_E |g|$ respectivamente. Nunca confundas $|\nu|(E)$ con $|\nu(E)|$: la primera es siempre mayor o igual que la segunda.
7. **Recuerda que $|\nu|(X)$ es una norma.** Con ella el conjunto de medidas signadas finitas es un espacio de Banach, y resulta ser el dual de $C_0(X)$: es el teorema de representación de Riesz para medidas.

Ejemplo 31.76 (Cuatro funciones de distribución y las medidas que generan)**. (a) La**

identidad. Con $F(x) = x$ se recupera $\mu_F((a,b]) = b - a$: la medida de Lebesgue. Es el caso puramente absolutamente continuo, con densidad constante 1.

(b) Una densidad más un átomo. Sea $F(x) = x$ para $x < 1$ y $F(x) = x + 1$ para $x \geq 1$. Entonces $\mu_F((0,2]) = F(2) - F(0) = 3$, que se descompone en longitud 2 más una masa 1 situada exactamente en el punto 1. En efecto $\mu_F(\{1\}) = F(1) - F(1^-) = 2 - 1 = 1$, mientras que $\mu_F((0,1)) = 1$ y $\mu_F((0,1]) = 2$. La medida es $\lambda + \delta_1$, suma de Lebesgue y una delta de Dirac.

(c) Sólo saltos. Si F es la función escalón que vale 0 para $x < 0$ y 1 para $x \geq 0$, entonces $\mu_F = \delta_0$: toda la masa está en el origen y cualquier conjunto que no lo contenga mide cero. Poniendo escalones de altura 2^{-k} en los puntos $1/k$ se obtiene una medida de probabilidad puramente atómica con infinitos átomos.

(d) Continua y singular. Si F es la escalera del diablo, μ_F es una medida de probabilidad sobre $[0,1]$ sin ningún átomo —porque F es continua— y concentrada en el conjunto de Cantor, que tiene medida de Lebesgue nula. No tiene densidad, y sin embargo no es discreta. Los tres ejemplos anteriores más este agotan las piezas: toda medida sobre $\mathbb{R}$ es suma de una parte con densidad, una parte atómica y una parte continua singular.

Ejemplo 31.77 (Una medida signada, medida con lupa)**. (a) La definición.** Sobre $[0,1]$ con la medida de Lebesgue, sea $\nu(E) = \int_E (x - \frac{1}{2})\,\mathrm{d}x$. Es numerablemente aditiva porque la integral lo es, y toma valores de ambos signos: $\nu([0, \frac{1}{4}]) = -\frac{3}{32}$ mientras que $\nu([\frac{1}{2}, 1]) = \frac{1}{8}$.

(b) La descomposición de Hahn. El conjunto positivo es $P = [\frac{1}{2}, 1]$ y el negativo $N = [0, \frac{1}{2})$, porque ahí es donde la densidad $x - \frac{1}{2}$ cambia de signo. Entonces $\nu^+([0,1]) = \int_{1/2}^1 (x - \frac{1}{2})\mathrm{d}x = \frac{1}{8}$ y $\nu^-([0,1]) = \int_0^{1/2} (\frac{1}{2} - x)\mathrm{d}x = \frac{1}{8}$.

(c) La diferencia entre $|\nu(E)|$ y $|\nu|(E)$. Se tiene $\nu([0,1]) = \frac{1}{8} - \frac{1}{8} = 0$, de modo que $|\nu([0,1])| = 0$; en cambio $|\nu|([0,1]) = \frac{1}{8} + \frac{1}{8} = \frac{1}{4}$. La variación total mide cuánta masa hay en juego, no cuánta sobrevive tras la cancelación. Obsérvese también que $\nu([\frac{1}{4}, \frac{3}{4}]) = 0$ sin que el conjunto sea despreciable: hay cancelación exacta.

(d) Otro ejemplo con la misma estructura. Con $\nu(E) = \int_E \sin x\,\mathrm{d}x$ sobre $[0, 2\pi]$, el conjunto positivo es $[0, \pi]$ y el negativo $[\pi, 2\pi]$, con $\nu^+ = \nu^- = 2$, luego $\nu([0,2\pi]) = 0$ y $|\nu|([0,2\pi]) = 4$. Ese número 4 es también la variación total de la función $-\cos x$ sobre el mismo intervalo, y no es coincidencia: la variación total de la función es la masa total de la medida que genera.

Cuidado

La descomposición de Hahn no es única como partición: el conjunto P puede modificarse en cualquier conjunto de $|\nu|$-medida nula sin dejar de ser positivo. Lo único que sí es único es el par de medidas (ν^+, ν^-). Y conviene no confundir dos descomposiciones distintas que aparecen juntas: la de Hahn–Jordan separa signos dentro de una misma medida, mientras que la de Lebesgue separa una medida positiva en parte absolutamente continua y parte singular respecto de otra. La segunda es la que precede a Radon–Nikodym; la primera es la que permite definir la integral respecto de una medida con signo.

Pausa para pensar

Una función de distribución de probabilidad es exactamente una F creciente, continua por la derecha, con límites 0 en $-\infty$ y 1 en $+\infty$; su medida de Stieltjes es la ley de la variable aleatoria. La esperanza $\int x\,\mathrm{d}\mu_F$ engloba en una sola fórmula el caso discreto, con su sumatorio, y el caso continuo, con su densidad, y además cubre las mezclas de los dos, que aparecen en cuanto una variable tiene un valor con probabilidad positiva y por lo demás se reparte con densidad. ¿Cuántas de las fórmulas que se aprenden por separado en un curso elemental de probabilidad son en realidad la misma fórmula escrita en dos casos particulares?

31.11. El broche: la probabilidad es teoría de la medida

El regalo final de Lebesgue (Kolmogórov, 1933): la probabilidad entera es un caso particular de este capítulo, y con ese diccionario los teoremas límite quedan a un paso.

Definición 31.78 (El diccionario)**.** Un *espacio de probabilidad* es un espacio de medida $(\Omega, \mathcal{F}, P)$ con $P(\Omega) = 1$. Una *variable aleatoria* es una función medible $X : \Omega \to \mathbb{R}$; su *distribución* es la medida imagen $P_X(B) = P(X \in B)$; la *esperanza* es la integral $\mathbb{E}[X] = \int X\,\mathrm{d}P$; "casi seguramente" es "c.t.p."; la *independencia* de X, Y es que $P_{(X,Y)}$ sea la medida *producto* $P_X \times P_Y$: cada concepto probabilista es un teorema de este capítulo con otro nombre. En particular $\mathbb{E}[XY] = \mathbb{E}[X]\mathbb{E}[Y]$ para independientes *es* Fubini.

Idea clave

Qué dice esto en cristiano. La probabilidad no es una teoría aparte: es teoría de la medida con la masa total normalizada a uno y con otro vocabulario. Un conjunto medible es un "suceso"; una función medible es una "variable aleatoria"; la integral es la "esperanza"; la convergencia en medida es la "convergencia en probabilidad"; la convergencia en casi todo punto es la "convergencia casi segura". Aprender el diccionario ahorra volver a demostrar cada teorema dos veces.

Ejemplo 31.79 (El diccionario en acción)**. (a) Un dado.** $\Omega = \{1, \dots, 6\}$, $\mathcal{F}$ todas las partes, $P(\{k\}) = 1/6$. La variable $X(\omega) = \omega$ tiene $\mathbb{E}[X] = \int X\,\mathrm{d}P = 3{,}5$ y $\mathrm{Var}(X) = 35/12 = 2{,}916667$.

(b) La uniforme continua. $\Omega = [0, 1]$ con la medida de Lebesgue. La variable $X(\omega) = \omega$ tiene $\mathbb{E}[X] = 1/2$ y $\mathrm{Var}(X) = 1/12$. El suceso "X es racional" tiene probabilidad 0 pese a no ser imposible: la distinción entre "probabilidad cero" e "imposible" es una consecuencia directa de trabajar con medidas.

(c) Independencia. Que dos sucesos sean independientes es que $P(A \cap B) = P(A)P(B)$. Es una condición sobre la medida, no sobre la lógica: los sucesos "el dado da par" y "el dado da menos de 3" cumplen $P = 1/2$, $P = 1/3$ y $P(\text{ambos}) = 1/6 = 1/2 \cdot 1/3$, luego son independientes, aunque intuitivamente parezcan relacionados.

(d) Los tres tipos de convergencia. Casi segura implica en probabilidad, que implica en distribución, y ninguna recíproca es cierta. La sucesión de indicatrices de intervalos que barren $[0,1]$ una y otra vez con longitudes $1/n$ converge en probabilidad a 0 y no converge en casi ningún punto: es el contraejemplo estándar, y es exactamente el mismo que en teoría de la medida se usa para separar convergencia en medida de convergencia puntual.

Teorema 31.80 (Ley débil de los grandes números, demostrada)**.** *Si $X_1, X_2, \dots$ son independientes con media μ y varianza $\sigma^2 < \infty$, entonces para todo $\varepsilon > 0$:*

$$P\Big(\Big|\frac{X_1 + \cdots + X_n}{n} - \mu\Big| > \varepsilon\Big) \le \frac{\sigma^2}{n\varepsilon^2} \xrightarrow{n\to\infty} 0.$$

Demostración. Dos líneas con lo ya construido. *Chebyshev*: $P(|Y - \mathbb{E}Y| > \varepsilon) \le \frac{\operatorname{Var} Y}{\varepsilon^2}$ (integrar la desigualdad $\mathbf{1}_{\{|Y-\mathbb{E}Y|>\varepsilon\}} \le \frac{(Y-\mathbb{E}Y)^2}{\varepsilon^2}$: monotonía de la integral). *Aditividad de la varianza*: para independientes, $\operatorname{Var}(\sum X_i) = \sum \operatorname{Var}(X_i) = n\sigma^2$ (los términos cruzados mueren por Fubini). Con $Y = \frac{\sum X_i}{n}$: $\operatorname{Var} Y = \frac{\sigma^2}{n}$. □ □

Ejemplo 31.81 (La ley de los grandes números, cuantificada)**. (a) La cota de Chebyshev.** Con $\sigma = 2$, $\varepsilon = 0{,}5$ y $n = 100$, la desigualdad da $P(|\bar X - \mu| \ge 0{,}5) \le \dfrac{4}{100 \cdot 0{,}25} = 0{,}16$. Si además las variables fuesen normales, la probabilidad real sería 0,012419: Chebyshev es correcta pero muy generosa, porque no usa nada sobre la forma de la distribución.

(b) Cuántos datos hacen falta. Para garantizar $P(|\bar X - \mu| \ge 0{,}1) \le 0{,}05$ con $\sigma = 2$ hace falta $n \ge \sigma^2/(\varepsilon^2 \cdot 0{,}05) = 4/(0{,}01 \cdot 0{,}05) = 8000$. Con la aproximación normal bastarían unas 1537 observaciones. La diferencia mide el precio de no suponer nada.

(c) Débil frente a fuerte. La ley débil dice que $\bar X_n \to \mu$ en probabilidad; la fuerte, que converge en casi todo punto. La primera se demuestra en dos líneas con Chebyshev; la segunda necesita el lema de Borel–Cantelli y es genuinamente más profunda, porque afirma algo sobre *cada* trayectoria, no sobre cada instante.

(d) Sin varianza finita. Para la distribución de Cauchy, la media muestral de n observaciones tiene *exactamente* la misma distribución que una sola observación: promediar no mejora nada. La ley de los grandes números falla porque no hay media, y es el recordatorio de que las hipótesis de momentos no son un adorno técnico.

Observación 31.82. La ley *fuerte* ($\frac{\sum X_i}{n} \to \mu$ casi seguramente, solo con $\mathbb{E}|X| < \infty$) requiere Borel–Cantelli (si $\sum P(A_n) < \infty$, casi seguro ocurren solo finitos A_n: convergencia monótona aplicada a $\sum \mathbf{1}_{A_n}$) y es la justificación matemática de la frecuencia como probabilidad: el casino gana *siempre*, no en promedio. Nótese qué compró la σ-aditividad de Kolmogórov: sin medida numerablemente aditiva no hay "casi seguramente", y la frase "la moneda justa saca 50 % de caras a la larga" ni siquiera puede enunciarse.

Teorema 31.83 (Límite central, con la transformada de Fourier)**.** *Con las hipótesis de arriba y $\mu = 0$, $\sigma = 1$:*

$$\frac{X_1 + \cdots + X_n}{\sqrt{n}} \xrightarrow{\ distr\ } N(0,1) \qquad (\textit{la gaussiana } \tfrac{1}{\sqrt{2\pi}} e^{-x^2/2}).$$

Idea completa. La *función característica* de una variable es exactamente la transformada de Fourier de su distribución: $\varphi_X(t) = \mathbb{E}[e^{itX}]$. Tres hechos del capítulo 31 la convierten en arma: (1) independencia $\Rightarrow$ la característica de la suma es el *producto* (Fubini); (2) Taylor con resto (¡capítulo 6!): $\varphi_X(t) = 1 - \frac{t^2}{2} + o(t^2)$; (3) por tanto

$$\varphi_{\frac{\sum X_i}{\sqrt{n}}}(t) = \left(\varphi_X\left(\tfrac{t}{\sqrt{n}}\right)\right)^n = \left(1 - \frac{t^2}{2n} + o\left(\tfrac{1}{n}\right)\right)^n \longrightarrow e^{-t^2/2},$$

el límite fundamental del capítulo 6, y $e^{-t^2/2}$ es la característica de la gaussiana (¡la gaussiana es su propia transformada!). El teorema de continuidad de Lévy (convergencia de características $\Rightarrow$ convergencia de distribuciones: se demuestra con la inversión de Fourier de este capítulo) concluye. □ □

Ejemplo 31.84 (El límite central en números)**. (a) Cien monedas.** Sea S el número de caras en 100 lanzamientos, con $\mu = 50$ y $\sigma = 5$. El valor exacto de $P(S \geq 60)$ es 0,028444. La aproximación normal cruda da $1 - \Phi(2) = 0{,}022750$, un error del veinte por ciento; con la corrección de continuidad, $1 - \Phi(1{,}9) = 0{,}028717$, error de menos del uno por ciento. La corrección merece la pena siempre.

(b) Mil monedas. Con $n = 1000$, $P(S \geq 550) = 0{,}000865$ y la aproximación normal con corrección da prácticamente lo mismo. Cuanto mayor es n, mejor va, y el teorema de Berry–Esseen cuantifica el error: es del orden de $1/\sqrt{n}$.

(c) Qué NO dice. El teorema no afirma que la suma *sea* normal, ni que las colas se aproximen bien. Para sucesos muy raros la aproximación normal puede errar en varios órdenes de magnitud, y por eso existen las desigualdades de grandes desviaciones. Usar la normal para estimar la probabilidad de un colapso financiero es el error clásico.

(d) Sin varianza finita, otro límite. Para variables con colas de índice $\alpha < 2$, las sumas normalizadas convergen a distribuciones estables de Lévy, no a la normal. La normal es simplemente el caso $\alpha = 2$ de una familia más amplia, y es atractora sólo dentro de su dominio.

Observación 31.85 (El final del libro)*.* Mírese la lista de ingredientes del teorema más famoso de la probabilidad: el límite $(1 + \frac{x}{n})^n \to e^x$ y Taylor (capítulo 6), la independencia como producto y Fubini (este capítulo), la transformada de Fourier con su inversión y la gaussiana autotransformada (este capítulo). El teorema del límite central no es un resultado *más*: es el punto donde el análisis del libro entero se dobla sobre sí mismo para explicar por qué la campana aparece en las estaturas, en los errores de medición y en el ruido de fondo del universo: porque *sumar muchas causas independientes pequeñas ejecuta, físicamente, una convolución iterada, y la convolución iterada es una multiplicación iterada en frecuencia cuyo único punto fijo estable es la gaussiana.* No hay mejor última página para un libro que empezó con ε y δ.

31.12. L^∞, el supremo esencial, y las desigualdades demostradas

Los espacios L^p se presentaron con la desigualdad de Hölder enunciada para $1 \le p \le \infty$, pero el caso $p = \infty$ requiere una norma que aún no se ha definido, y ni Hölder ni Minkowski se han demostrado. Se cierra aquí el hueco.

Definición 31.86 (Supremo esencial)**.** Sea $(X, \mathcal{M}, \mu)$ un espacio de medida y f medible. El *supremo esencial* de $|f|$ es

$$\|f\|_\infty = \operatorname*{ess\,sup}_{x \in X} |f(x)| = \inf\{M \ge 0 : \ \mu(\{x : |f(x)| > M\}) = 0\},$$

con el convenio $\inf \emptyset = +\infty$. El espacio $L^\infty(X, \mu)$ es el de las funciones medibles con $\|f\|_\infty < \infty$, identificando las que coinciden en casi todo punto.

Proposición 31.87. *El ínfimo de la definición se alcanza: $|f(x)| \le \|f\|_\infty$ en casi todo punto. Además $\|\cdot\|_\infty$ es una norma y L^∞ es completo.*

Demostración. Sea $M_0 = \|f\|_\infty < \infty$. Para cada n existe $M_n < M_0 + \frac{1}{n}$ con $\mu(\{|f| > M_n\}) = 0$, luego $\mu(\{|f| > M_0 + \frac{1}{n}\}) = 0$. Como $\{|f| > M_0\} = \bigcup_n \{|f| > M_0 + \frac{1}{n}\}$ es unión numerable de conjuntos nulos, es nulo. La desigualdad triangular se sigue de que $|f + g| \le \|f\|_\infty + \|g\|_\infty$ fuera de la unión de dos conjuntos nulos. Para la completitud, si (f_n) es de Cauchy, el conjunto donde alguna de las desigualdades $|f_n - f_m| \le \|f_n - f_m\|_\infty$ falla es una unión numerable de nulos; fuera de él (f_n) es uniformemente de Cauchy y converge uniformemente a una f acotada. □

Cuidado

[Supremo esencial no es supremo] La función que vale 0 en los irracionales y 10^6 en los racionales tiene supremo 10^6 y supremo esencial 0: la medida de Lebesgue de $\mathbb{Q}$ es nula. Es exactamente lo que se quiere, porque en teoría de la medida las funciones que difieren en un conjunto nulo son la misma función, y el supremo ordinario no respeta esa identificación. Nótese también que L^∞ es el único L^p donde la convergencia en norma es la convergencia uniforme (fuera de un nulo), y por eso su comportamiento es distinto: no es separable, y su dual no es L^1.

Lema 31.88 (Young)**.** *Sean $p, q > 1$ con $\frac{1}{p} + \frac{1}{q} = 1$ (*exponentes conjugados*). Entonces, para $a, b \ge 0$,*

$$ab \ \le \ \frac{a^p}{p} + \frac{b^q}{q},$$

con igualdad si y sólo si $a^p = b^q$.

Demostración. Si $a = 0$ o $b = 0$ es obvio. En otro caso, tómense logaritmos: como log es cóncava y $\frac{1}{p} + \frac{1}{q} = 1$,

$$\log\left(\frac{a^p}{p} + \frac{b^q}{q}\right) \ \ge \ \frac{1}{p}\log(a^p) + \frac{1}{q}\log(b^q) = \log a + \log b = \log(ab),$$

y log es creciente. La igualdad en la desigualdad de Jensen para una función estrictamente cóncava exige que los dos puntos coincidan: $a^p = b^q$. □

Teorema 31.89 (Hölder)**.** *Sean $1 \le p \le \infty$ y q su conjugado ($\frac{1}{p} + \frac{1}{q} = 1$, con $q = \infty$ si $p = 1$). Si $f \in L^p$ y $g \in L^q$, entonces $fg \in L^1$ y*

$$\|fg\|_1 \;\le\; \|f\|_p \, \|g\|_q.$$

Demostración. Casos extremos. Si $p = 1$, $q = \infty$: $|fg| \le |f| \, \|g\|_\infty$ en casi todo punto, e integrando sale la desigualdad. El caso $p = \infty$ es simétrico.

Caso $1 < p < \infty$. Si $\|f\|_p = 0$ entonces $f = 0$ c.t.p. y no hay nada que probar; igual con g. Supóngase pues que ambas normas son positivas y, dividiendo, que $\|f\|_p = \|g\|_q = 1$: hay que probar $\int |fg| \le 1$. Aplíquese el lema 31.88 puntualmente con $a = |f(x)|$ y $b = |g(x)|$:

$$|f(x)g(x)| \;\le\; \frac{|f(x)|^p}{p} + \frac{|g(x)|^q}{q}.$$

Integrando,

$$\int |fg| \;\le\; \frac{1}{p}\int |f|^p + \frac{1}{q}\int |g|^q = \frac{1}{p} + \frac{1}{q} = 1. \qquad \square$$

Teorema 31.90 (Minkowski)**.** *Para $1 \le p \le \infty$ y $f, g \in L^p$ se tiene $\|f + g\|_p \le \|f\|_p + \|g\|_p$. Es decir, $\|\cdot\|_p$ es una norma.*

Demostración. Los casos $p = 1$ y $p = \infty$ son inmediatos. Sea $1 < p < \infty$ y supóngase $\|f + g\|_p > 0$. Entonces

$$\|f+g\|_p^p = \int |f+g|^p \le \int |f| \, |f+g|^{p-1} + \int |g| \, |f+g|^{p-1},$$

usando $|f + g| \le |f| + |g|$. A cada sumando se le aplica Hölder con exponentes p y $q = \frac{p}{p-1}$; obsérvese que $\left(|f+g|^{p-1}\right)^q = |f+g|^p$, luego $\left\||f+g|^{p-1}\right\|_q = \|f+g\|_p^{p/q} = \|f+g\|_p^{p-1}$. Así

$$\|f+g\|_p^p \;\le\; \left(\|f\|_p + \|g\|_p\right) \|f+g\|_p^{p-1},$$

y dividiendo entre $\|f+g\|_p^{p-1}$, que es finito y no nulo, se obtiene el resultado. (Que $\|f+g\|_p < \infty$ se sigue de $|f+g|^p \le 2^p(|f|^p + |g|^p)$.) □

Corolario 31.91 (Desigualdad de interpolación)**.** *Si $1 \le p < r < q \le \infty$ y $f \in L^p \cap L^q$, entonces $f \in L^r$ y $\|f\|_r \le \|f\|_p^\theta \|f\|_q^{1-\theta}$, donde $\frac{1}{r} = \frac{\theta}{p} + \frac{1-\theta}{q}$.*

Demostración. Escríbase $|f|^r = |f|^{\theta r}|f|^{(1-\theta)r}$ y aplíquese Hölder con exponentes $\frac{p}{\theta r}$ y $\frac{q}{(1-\theta)r}$, que son conjugados por la relación entre los exponentes. □

Observación 31.92 (Por qué L^q es el dual de L^p). La desigualdad de Hölder dice que cada $g \in L^q$ define un funcional lineal continuo sobre L^p, a saber $\Lambda_g(f) = \int fg$, con $\|\Lambda_g\| \le \|g\|_q$. Que la norma sea exactamente $\|g\|_q$ se ve tomando $f = |g|^{q-1}\operatorname{sgn}(g)$, con la que Hölder es una igualdad. Y que *todo* funcional continuo sea de esa forma —el teorema de representación de

Riesz para L^p, válido para $1 \le p < \infty$ y σ-finitud— es una consecuencia del teorema de Radon–Nikodym: dado Λ, se define la medida $\nu(E) = \Lambda(\mathbf{1}_E)$, se comprueba que es absolutamente continua respecto de μ y se toma $g = \mathrm{d}\nu/\mathrm{d}\mu$. Nótese la exclusión: $(L^\infty)^*$ es *estrictamente* mayor que L^1.

31.13. Egorov y Lusin: casi uniforme, casi continua

Dos teoremas que el capítulo ha usado sin enunciar, y que responden a la misma pregunta: ¿cuánto se parece una función medible a una función buena, si se permite descartar un conjunto pequeño?

Teorema 31.93 (Egorov)**.** *Sea $\mu(X) < \infty$ y $f_n \to f$ en casi todo punto, con todas las funciones medibles y finitas c.t.p. Entonces, para cada $\varepsilon > 0$, existe $E \subset X$ medible con $\mu(X \setminus E) < \varepsilon$ tal que $f_n \to f$* uniformemente *sobre E.*

Demostración. Para $n, k \ge 1$ defínase

$$E_{n,k} = \bigcup_{m \ge n} \left\{ x : |f_m(x) - f(x)| \ge \tfrac{1}{k} \right\}.$$

Para k fijo, $(E_{n,k})_n$ es decreciente en n, y su intersección está contenida en el conjunto de puntos donde $f_n \not\to f$, que es nulo. Como $\mu(X) < \infty$, la continuidad de la medida por arriba da $\mu(E_{n,k}) \to 0$ cuando $n \to \infty$. Elíjase entonces, para cada k, un índice n_k con

$$\mu(E_{n_k,k}) < \frac{\varepsilon}{2^k},$$

y póngase $E = X \setminus \bigcup_k E_{n_k,k}$. Por subaditividad, $\mu(X \setminus E) \le \sum_k \varepsilon 2^{-k} = \varepsilon$. Y sobre E la convergencia es uniforme: dado $\delta > 0$, tómese k con $\frac{1}{k} < \delta$; como $E \cap E_{n_k,k} = \emptyset$, para todo $x \in E$ y todo $m \ge n_k$ se tiene $|f_m(x) - f(x)| < \frac{1}{k} < \delta$, y el índice n_k no depende de x. □

Cuidado

[La hipótesis $\mu(X) < \infty$ es indispensable] En $\mathbb{R}$ con la medida de Lebesgue, sea $f_n = \mathbf{1}_{[n,n+1]}$. Entonces $f_n \to 0$ en todo punto, pero sobre cualquier conjunto E de complemento de medida finita la convergencia no es uniforme: para todo n hay puntos de $[n, n+1] \cap E$ donde $f_n = 1$. El teorema falla, y su fallo señala qué está haciendo realmente la hipótesis: impedir que la "mala convergencia" se escape hacia el infinito.

Teorema 31.94 (Lusin)**.** *Sea $f : \mathbb{R}^n \to \mathbb{C}$ medible y finita c.t.p., y sea A de medida finita con $f = 0$ fuera de A. Para cada $\varepsilon > 0$ existe una función* continua *g de soporte compacto tal que*

$$\lambda(\{x : f(x) \ne g(x)\}) < \varepsilon,$$

y además puede elegirse $\|g\|_\infty \le \|f\|_\infty$.

Idea clave

Cómo se demuestra. Primero para $f = \mathbf{1}_E$: por regularidad de la medida de Lebesgue existen K compacto y U abierto con $K \subset E \subset U$ y $\lambda(U \setminus K) < \varepsilon$; el lema de Urysohn produce una g continua que vale 1 en K, 0 fuera de U y está entre ambos. Luego, por linealidad, para funciones simples. Finalmente, una f medible es límite puntual de simples, y Egorov convierte esa convergencia puntual en uniforme fuera de un conjunto de medida $< \varepsilon/2$; el límite uniforme de continuas es continuo.

Observación 31.95 (El eslogan de Littlewood). Los tres teoremas se resumen en los "tres principios" de Littlewood, que conviene memorizar porque orientan casi toda la técnica del capítulo:

1. *Todo conjunto medible es casi una unión finita de intervalos* (regularidad).
2. *Toda función medible es casi continua* (Lusin).
3. *Toda sucesión convergente c.t.p. es casi uniformemente convergente* (Egorov).

"Casi" significa siempre lo mismo: salvo un conjunto de medida arbitrariamente pequeña. La teoría de la medida no es más difícil que la de Riemann; sólo hay que aceptar que las igualdades se cumplen fuera de conjuntos despreciables, y estos tres teoremas dicen cuánto se puede recuperar de la intuición clásica pagando ese precio.

31.14. El lema de Vitali y la derivación de funciones monótonas

Se ha invocado varias veces —aquí y en el capítulo 29— el teorema de Lebesgue según el cual toda función monótona es derivable en casi todo punto, y sobre él descansa toda la teoría de la variación acotada. Es la última pieza que faltaba, y se demuestra con un lema de recubrimiento que además tiene interés propio.

Lema 31.96 (Recubrimiento de Vitali, versión finita)**.** *Sea $\mathcal{B} = \{B_1, \dots, B_N\}$ una colección finita de bolas en $\mathbb{R}^n$. Existe una subcolección* disjunta *$B_{i_1}, \dots, B_{i_k}$ tal que*

$$\lambda\Big(\bigcup_{j=1}^{N} B_j\Big) \;\le\; 3^n \sum_{m=1}^{k} \lambda(B_{i_m}).$$

Demostración. Algoritmo voraz: elíjase B_{i_1} de radio máximo; descártense todas las bolas que la cortan; entre las restantes, elíjase la de radio máximo, y así sucesivamente hasta agotar la colección. Las elegidas son disjuntas por construcción. Y si B fue descartada, es porque corta a alguna B_{i_m} elegida *antes* y por tanto de radio mayor o igual; si r es el radio de B y $r_m \ge r$ el de B_{i_m}, y se cortan, entonces B está contenida en la bola concéntrica con B_{i_m} de radio $3r_m$ (hay que recorrer, desde el centro de B_{i_m}, a lo sumo $r_m + 2r \le 3r_m$). Por tanto $\bigcup_j B_j \subset \bigcup_m 3B_{i_m}$, y como $\lambda(3B) = 3^n\lambda(B)$, se concluye. □

Definición 31.97 (Números derivados de Dini)**.** Para $f:[a,b]\to\mathbb{R}$ y x interior, se definen

$$D^+f(x)=\limsup_{h\to 0^+}\frac{f(x+h)-f(x)}{h},\qquad D_+f(x)=\liminf_{h\to 0^+}\frac{f(x+h)-f(x)}{h},$$

y análogamente D^-f, D_-f por la izquierda. La función es derivable en x exactamente cuando los cuatro coinciden y son finitos.

Teorema 31.98 (Lebesgue)**.** *Toda función monótona $f:[a,b]\to\mathbb{R}$ es derivable en casi todo punto, f' es medible, y*

$$\int_a^b f'(x)\,\mathrm{d}x\ \le\ f(b)-f(a),$$

con desigualdad posiblemente estricta.

Idea clave

La estructura del argumento. Supóngase f creciente. Para $u<v$ racionales, sea

$$E_{u,v}=\{x: D_+f(x)<u<v<D^+f(x)\}.$$

El conjunto donde los números derivados no coinciden es la unión numerable de estos $E_{u,v}$, así que basta probar que cada uno es nulo. Fíjese $\varepsilon>0$ y tómese un abierto $U\supset E_{u,v}$ con $\lambda(U)<\lambda^*(E_{u,v})+\varepsilon$. Como $D_+f<u$ en cada punto de $E_{u,v}$, hay intervalos arbitrariamente pequeños $[x,x+h]\subset U$ con $f(x+h)-f(x)<uh$; el lema de Vitali permite extraer una familia disjunta de ellos que cubre casi todo $E_{u,v}$, y sumando: el incremento total de f sobre esa familia es $<u\lambda(U)$. Ahora se repite el razonamiento *dentro* de esos intervalos con la desigualdad $D^+f>v$, obteniendo una segunda familia disjunta cuyo incremento total es $>v\,\lambda^*(E_{u,v})$. Como el incremento de la segunda familia no puede superar al de la primera (está contenida en ella y f crece),

$$v\,\lambda^*(E_{u,v})\ \le\ u\,\lambda(U)\ \le\ u(\lambda^*(E_{u,v})+\varepsilon).$$

Haciendo $\varepsilon\to 0$ queda $v\,\lambda^*(E_{u,v})\le u\,\lambda^*(E_{u,v})$ con $u<v$, lo cual fuerza $\lambda^*(E_{u,v})=0$. Que la derivada sea finita c.t.p. se obtiene igual, comparando con $D^+f=+\infty$. Finalmente, poniendo $g_n(x)=n\big(f(x+\tfrac1n)-f(x)\big)\ge 0$, se tiene $g_n\to f'$ c.t.p., y el lema de Fatou da

$$\int_a^b f'\le\liminf\int_a^b g_n=\liminf\ n\int_b^{b+1/n}f-n\int_a^{a+1/n}f\ \le\ f(b)-f(a),$$

usando que f es creciente.

Ejemplo 31.99 (La desigualdad es estricta: la escalera de Cantor)**.** Sea $c:[0,1]\to[0,1]$ la función de Cantor: continua, creciente, $c(0)=0$, $c(1)=1$, y constante en cada uno de los intervalos que se van quitando al construir el conjunto de Cantor. Como esos intervalos tienen medida total 1, se tiene $c'=0$ en casi todo punto, luego

$$\int_0^1 c'(x)\,\mathrm{d}x=0\ <\ 1=c(1)-c(0).$$

La función sube de 0 a 1 sin tener nunca derivada positiva, salvo en un conjunto de medida nula. Esto muestra que "derivable c.t.p. con derivada integrable" *no* basta para recuperar la función integrando: hace falta la continuidad absoluta, y ese es exactamente el contenido del teorema fundamental del cálculo en su versión de Lebesgue.

Corolario 31.100 (Jordan y la variación acotada)**.** *Toda función de variación acotada es diferencia de dos crecientes (descomposición de Jordan), luego es derivable en casi todo punto. En particular, toda función absolutamente continua lo es.*

Teorema 31.101 (Teorema fundamental del cálculo, versión de Lebesgue)**.** *Sea $F:[a,b]\to\mathbb{R}$. Son equivalentes:*

1. *F es absolutamente continua;*
2. *existe $g\in L^1(a,b)$ con $F(x)=F(a)+\int_a^x g$ para todo x;*
3. *F es derivable c.t.p., $F'\in L^1$ y $F(x)=F(a)+\int_a^x F'$.*

Demostración. (2)⇒(1). Por la continuidad absoluta de la integral: para $\varepsilon>0$ existe $\delta>0$ tal que $\lambda(E)<\delta$ implica $\int_E|g|<\varepsilon$ (consecuencia del teorema de convergencia dominada aplicado a $|g|\mathbf{1}_{\{|g|>n\}}$). Aplicado a una unión finita de intervalos disjuntos de longitud total $<\delta$, da exactamente la definición de continuidad absoluta de F.

(1)⇒(3). Una función absolutamente continua es de variación acotada (partiendo $[a,b]$ en un número finito de trozos de longitud $<\delta$, la variación de cada uno es ≤ 1), luego por el corolario anterior F' existe c.t.p. y es integrable. Sea $G(x)=\int_a^x F'$. Por la implicación ya probada, G es absolutamente continua, y por el teorema de diferenciación de Lebesgue $G'=F'$ c.t.p. Entonces $H=F-G$ es absolutamente continua con $H'=0$ c.t.p., y el lema siguiente fuerza H constante; evaluando en a, $H\equiv F(a)$.

Lema. Si H es absolutamente continua y $H'=0$ c.t.p., entonces H es constante. En efecto, sea E el conjunto donde $H'=0$, de medida total, y sean $c<d$ en $[a,b]$ y $\varepsilon>0$. Cada $x\in E\cap(c,d)$ está en intervalos arbitrariamente cortos con $|H(x+h)-H(x)|<\varepsilon h$; por el lema de Vitali se extrae una familia disjunta que cubre $E\cap(c,d)$ salvo medida $<\delta$, donde δ es el de la continuidad absoluta para ε. Sumando sobre la familia se obtiene $|H(d)-H(c)|\le\varepsilon(d-c)+\varepsilon$, y haciendo $\varepsilon\to 0$, $H(d)=H(c)$.

(3)⇒(2) es trivial con $g=F'$. □

Observación 31.102 (La referencia circular, cerrada). El capítulo 29 enunciaba este teorema y remitía aquí; aquí se enunciaba y se usaba el teorema de Lebesgue sobre monótonas sin demostrar. Con el lema de Vitali, el teorema 31.98 y el teorema 31.101, la cadena queda cerrada por completo: continuidad absoluta ⇒ variación acotada ⇒ (Jordan) diferencia de monótonas ⇒ (Lebesgue–Vitali) derivable c.t.p.; y la continuidad absoluta es lo que convierte esa derivada c.t.p. en suficiente para reconstruir la función. La escalera de Cantor marca exactamente dónde falla si se quita esa hipótesis.

31.15. σ-finitud y medida producto: las hipótesis que las reglas mandan comprobar

Las reglas del capítulo ordenan verificar hipótesis que no estaban definidas en ninguna parte.

Definición 31.103 (Medida finita y σ-finita)**.** Una medida μ sobre $(X, \mathcal{M})$ es *finita* si $\mu(X) < \infty$, y *σ-finita* si X se puede escribir como unión numerable de conjuntos de medida finita:

$$X = \bigcup_{n \geq 1} X_n \qquad \text{con} \qquad \mu(X_n) < \infty.$$

Ejemplo 31.104 (Quién lo es y quién no)**.** La medida de Lebesgue en $\mathbb{R}^n$ es σ-finita —basta tomar las bolas de radio n— pero no finita. Toda medida de probabilidad es finita. La medida de *contar* sobre $\mathbb{N}$ es σ-finita; sobre $\mathbb{R}$ *no* lo es, porque los conjuntos de medida finita son los finitos y no hay unión numerable de finitos que dé $\mathbb{R}$. Y la medida de Hausdorff $\mathcal{H}^s$ en $\mathbb{R}^n$ con $s < n$ tampoco es σ-finita respecto de la de Lebesgue.

Ahí está el ejemplo que muestra por qué Fubini exige la hipótesis: sobre $[0,1]^2$ con μ la medida de Lebesgue y ν la de contar, la función indicadora de la diagonal da $\int\!\!\int f \, \mathrm{d}\mu \, \mathrm{d}\nu = 0$ y $\int\!\!\int f \, \mathrm{d}\nu \, \mathrm{d}\mu = 1$. Los dos órdenes de integración dan resultados distintos, y lo único que falla es la σ-finitud de ν.

Definición 31.105 (σ-álgebra y medida producto)**.** Dados $(X, \mathcal{M}, \mu)$ e $(Y, \mathcal{N}, \nu)$, la *σ-álgebra producto* $\mathcal{M} \otimes \mathcal{N}$ es la generada por los *rectángulos medibles* $A \times B$ con $A \in \mathcal{M}$, $B \in \mathcal{N}$. La *medida producto* $\mu \times \nu$ es la única medida sobre ella que cumple

$$(\mu \times \nu)(A \times B) = \mu(A)\, \nu(B),$$

con el convenio $0 \cdot \infty = 0$. La unicidad está garantizada si ambas son σ-finitas, por el teorema de extensión de Carathéodory aplicado al álgebra de uniones finitas de rectángulos.

Cuidado

[La σ-álgebra producto no es la de Lebesgue del producto] $\mathcal{L}(\mathbb{R}) \otimes \mathcal{L}(\mathbb{R})$ está *estrictamente contenida* en $\mathcal{L}(\mathbb{R}^2)$: hay conjuntos medibles-Lebesgue del plano que no están en la σ-álgebra producto. El motivo es que la medida producto no es completa aunque los factores lo sean, y $\mathcal{L}(\mathbb{R}^2)$ es su completación. En la práctica no molesta —se trabaja con la completación— pero explica por qué los enunciados cuidadosos de Fubini distinguen entre "medible respecto del producto" y "medible-Lebesgue", y por qué la versión de Tonelli para funciones no negativas se enuncia primero en la σ-álgebra producto y luego se extiende.

31.16. Ejercicios resueltos

Ejercicio 31.1. Demuestre que $\lim_{n\to\infty} \int_0^\infty \frac{\mathrm{d}x}{(1+x/n)^n\, x^{1/n}}$ existe y calcúlelo.

Solución. El integrando tiende puntualmente a e^{-x} (límite fundamental del capítulo 6; $x^{1/n}\to 1$ para $x>0$). Dominante: para $n\geq 2$, $(1+x/n)^n \geq (1+x/2)^2$ (la sucesión $(1+x/n)^n$ es creciente en n: desigualdad AM–GM, capítulo 6), y $x^{-1/n}\leq \max(1,x^{-1/2})$, así que el integrando se domina por $g(x)=\frac{\max(1,x^{-1/2})}{(1+x/2)^2}\in L^1(0,\infty)$ (integrable en 0 por comparación con $x^{-1/2}$, en ∞ por x^{-2}). Convergencia dominada:

$$\lim_n \int_0^\infty = \int_0^\infty e^{-x}\mathrm{d}x = 1.$$

Sin el teorema, este límite exigiría estimaciones artesanales; con él, es rellenar una lista de verificación. ■

Ejercicio 31.2. Sea $f_n(x)=\frac{n}{x}\,\mathbf{1}_{[n,2n]}(x)$. Calcule $\lim f_n$ puntual, $\lim \int f_n$ y explique por qué no coinciden sin contradecir ningún teorema.

Solución. Para cada x fijo, $f_n(x)=0$ en cuanto $n>x$ (el soporte $[n,2n]$ se marcha): el límite puntual es 0. Pero $\int f_n = \int_n^{2n}\frac{n}{x}\,\mathrm{d}x = n\ln 2\to\infty$. No hay contradicción: la sucesión no está dominada —cualquier g con $g\geq f_n$ para todo n cumple $\int_{\mathbb{R}} g \geq \int_n^{2n} g \geq n\ln 2$ para cada n, así que $\int g=\infty$— y Fatou solo promete $\int \liminf \leq \liminf \int$: $0\leq\infty$, verdadero y estricto. Es la "joroba que se fuga al infinito": la tercera patología del catálogo, en estado puro. ■

Ejercicio 31.3. Calcule $\widehat{\mathbf{1}_{[-a,a]}}$ y deduzca $\int_{-\infty}^{\infty}\frac{\sin^2(a\xi)}{\xi^2}\mathrm{d}\xi=\pi a$.

Solución. $\hat f(\xi)=\int_{-a}^{a}e^{-i\xi x}\mathrm{d}x=\frac{e^{ia\xi}-e^{-ia\xi}}{i\xi}=\frac{2\sin(a\xi)}{\xi}$. Plancherel: $\frac{1}{2\pi}\int\frac{4\sin^2(a\xi)}{\xi^2}\mathrm{d}\xi=\int_{-a}^{a}1=2a$, y despejando, $\int\frac{\sin^2(a\xi)}{\xi^2}\mathrm{d}\xi=\pi a$. (Comprobación de cordura: en el capítulo 30 la misma integral con $a=1$ salió por el contorno con muesca: π. Coincide.) Nótese la linealidad en a: dilatar el intervalo comprime la sinc, y las áreas se compensan exactamente como dicta el escalamiento $\widehat{f(x/\lambda)}=\lambda\hat f(\lambda\xi)$. ■

Ejercicio 31.4. Demuestre que no existe $f\in L^1(\mathbb{R})$ que sea *elemento identidad* para la convolución: $f*g=g$ para toda $g\in L^1$.

Solución. Transformando: la convolución se vuelve producto ($\widehat{f*g}=\hat f\hat g$: Fubini directo), así que haría falta $\hat f(\xi)\hat g(\xi)=\hat g(\xi)$ para toda g; eligiendo g gaussiana ($\hat g$ nunca se anula), $\hat f\equiv 1$. Pero Riemann–Lebesgue exige $\hat f(\xi)\to 0$: contradicción. La identidad de la convolución "quiere ser" la delta de Dirac —la chimenea $n\mathbf{1}_{(0,1/n)}$ en el límite—, que no es una función: exactamente el objeto que la teoría de *distribuciones* de Schwartz (años 40) construye para legalizar lo que físicos e ingenieros ya usaban. Este ejercicio es la puerta señalizada hacia ese siguiente curso. ■

Ejercicio 31.5. (Síntesis final.) Usando el núcleo del calor, demuestre la *fórmula de sumación de Poisson* en su caso más famoso: para $t>0$,

$$\sum_{n\in\mathbb{Z}} e^{-\pi n^2 t} = \frac{1}{\sqrt{t}}\sum_{n\in\mathbb{Z}} e^{-\pi n^2/t}.$$

Solución. Sea $\theta(t)$ el lado izquierdo. La función $u(x,t)=\sum_n e^{-\pi n^2 t}e^{2\pi i n x}$ resuelve el calor en el *círculo* (serie de Fourier, capítulo 12) con dato inicial el "peine" de deltas; pero la misma evolución se obtiene enrollando la solución de la *recta*: sumando el núcleo gaussiano sobre todas las copias enteras, $u(x,t)=\sum_n \frac{1}{\sqrt{t}}e^{-\pi(x-n)^2/t}$ (con la normalización k, tiempos reescalados para que el núcleo sea $t^{-1/2}e^{-\pi x^2/t}$). Evaluar ambas expresiones en $x=0$ e igualar da la fórmula. La identidad es la dualidad círculo↔recta hecha ecuación (periodizar↔muestrear), y no es un juguete: es el ingrediente exacto con que Riemann demostró la ecuación funcional de su función ζ, y la razón profunda del teorema del muestreo. El libro cierra donde empieza la siguiente aventura. ■

31.17. Taller de ejemplos explicados: convergencia y medida cero

Idea clave

Intuición primero. El teorema de convergencia dominada dice, en cristiano: si tus funciones se acercan a un límite y todas están "bajo un mismo techo" integrable, entonces el límite de las integrales es la integral del límite (puedes meter el límite dentro). Y "medida cero" formaliza la idea de un conjunto tan flaco que no aporta área, aunque tenga infinitos puntos, como los racionales. Los dos ejemplos hacen tangibles ambas ideas.

31.17.1. Un límite de integrales con el teorema de convergencia dominada

El teorema de convergencia dominada dice, en cristiano: si una sucesión de funciones f_n se acerca a un límite f, y todas están "techadas" por una misma función integrable, entonces el límite de las integrales es la integral del límite. Veámoslo en un caso.

Ejemplo 31.106. Calcular $\lim_{n\to\infty} \int_0^1 x^n \, \mathrm{d}x$.

Paso 1. Ver a qué tiende x^n punto por punto. Para $0 \le x < 1$, elevar a potencias grandes lleva el número hacia 0 (por ejemplo 0.9^{100} es minúsculo). Solo en $x = 1$ se queda en 1. Así que $x^n \to 0$ en casi todo el intervalo.

Paso 2. Verificar la "dominación". Todas las x^n cumplen $0 \le x^n \le 1$ en $[0, 1]$, y la función constante 1 es integrable ahí. Hay techo común: se puede aplicar el teorema.

Paso 3. Concluir. El teorema permite meter el límite dentro de la integral:

$$\lim_{n\to\infty} \int_0^1 x^n \, \mathrm{d}x = \int_0^1 \left(\lim_{n\to\infty} x^n \right) \mathrm{d}x = \int_0^1 0 \, \mathrm{d}x = 0.$$

Comprobación directa: $\displaystyle\int_0^1 x^n \mathrm{d}x = \frac{1}{n+1}$, que en efecto tiende a 0 cuando $n \to \infty$. El teorema acertó, y nos ahorró calcular la integral cada vez.

31.17.2. Un conjunto de medida cero

Ejemplo 31.107. Los racionales dentro de $[0, 1]$ forman un conjunto de *medida cero*: aunque hay infinitos y están por todas partes, "ocupan longitud cero". La razón intuitiva: son *numerables* (se pueden listar $r_1, r_2, r_3, \dots$), y podemos cubrir el k-ésimo con un intervalito de largo $\dfrac{\varepsilon}{2^k}$; la suma total de las longitudes es

$$\frac{\varepsilon}{2} + \frac{\varepsilon}{4} + \frac{\varepsilon}{8} + \cdots = \varepsilon,$$

que se puede hacer tan pequeña como queramos. Por eso la función que vale 1 en los racionales y 0 en los irracionales tiene integral de Lebesgue 0: el "1" vive en un conjunto de medida nula y no aporta área.

31.18. Ejercicios propuestos

Sin solución desarrollada: solo enunciado y, al final, la respuesta. En esta parte del libro muchas respuestas son una frase ("sí, y la razón es…") más que un número; comprueba que la tuya coincide en el fondo, no en la letra.

31.1. ¿Cuál es la medida de Lebesgue de $\mathbb{Q}\cap[0,1]$? ¿Y la de los irracionales de $[0,1]$?

31.2. Construye un abierto denso en $[0,1]$ de medida menor que $1/10$.

31.3. ¿Es Lebesgue-integrable la función de Dirichlet en $[0,1]$? ¿Cuánto vale la integral?

31.4. Sea $f_n = n\chi_{(0,1/n)}$. Calcula $\lim \int f_n$ y $\int \lim f_n$. ¿Qué teorema falla y por qué?

31.5. Aplica convergencia dominada a $f_n(x) = \frac{\sin(nx)}{n}$ en $[0,1]$: ¿cuál es el límite de $\int f_n$?

31.6. Enuncia la desigualdad del lema de Fatou y di en qué sentido puede ser estricta.

31.7. ¿Es $1/\sqrt{x}$ integrable en $(0,1)$? ¿Y $1/x$?

31.8. ¿Es $1/x^2$ integrable en $(1,\infty)$? ¿Y $1/x$?

31.9. Calcula $\int_{\mathbb{R}} e^{-x^2}\mathrm{d}x$ y di si $\frac{\sin x}{x}$ es Lebesgue-integrable en $(0,\infty)$.

31.10. ¿Toda función medible es continua en casi todo punto? Cita el teorema correcto (Lusin) y di qué afirma realmente.

31.11. ¿Existe un subconjunto de $\mathbb{R}$ no medible? ¿Qué axioma se usa en la construcción?

31.12. Si $f \geq 0$ es medible y $\int f = 0$, ¿qué puedes concluir sobre f?

Respuestas

1. 0 y 1 respectivamente: los racionales son numerables, luego de medida nula. **2.** Enumera $\mathbb{Q}\cap[0,1] = \{q_k\}$ y toma $\bigcup_k (q_k - \varepsilon 2^{-k-1}, q_k + \varepsilon 2^{-k-1})$ con $\varepsilon = 1/20$: es denso y de medida $\leq \varepsilon < 1/10$. **3.** Sí, y vale 0: es 0 salvo en un conjunto de medida nula. **4.** $\lim \int f_n = 1$ pero $\int \lim f_n = 0$. Falla la convergencia dominada: no hay dominante integrable (la mínima sería $\sim 1/x$). **5.** $|f_n| \leq 1$, dominante constante integrable en $[0,1]$; $f_n \to 0$, luego $\int f_n \to 0$. **6.** $\int \liminf f_n \leq \liminf \int f_n$. Es estricta, por ejemplo, con las jorobas viajeras del ejercicio 4. **7.** $1/\sqrt{x}$ sí (la integral vale 2); $1/x$ no (diverge logarítmicamente en 0). **8.** $1/x^2$ sí (vale 1); $1/x$ no (diverge en ∞). **9.** $\sqrt{\pi}$. Y no: $\int_0^\infty |\sin x/x| = \infty$, así que es integrable impropia de Riemann pero *no* Lebesgue. **10.** No. Lusin dice que para todo $\varepsilon > 0$ existe un cerrado F con $\mu(F^c) < \varepsilon$ tal que $f|_F$ es continua: continuidad tras quitar poco, no continuidad casi en todo punto. **11.** Sí: el conjunto de Vitali. Usa el axioma de elección. **12.** Que $f = 0$ en casi todo punto.

Capítulo 32

Análisis funcional: álgebra lineal en dimensión infinita

Lo que necesitas traer

Del capítulo 7: espacios vectoriales, aplicaciones lineales, producto interno y bases ortonormales. Del capítulo 29: completitud, sucesiones de Cauchy y convergencia uniforme. Del capítulo 31: los espacios L^p y la integral de Lebesgue.

32.1. La idea: cuando los vectores son funciones

Idea clave

Intuición primero. El capítulo 7 enseñó que un vector no es una flecha: es un elemento de un conjunto donde se puede sumar y multiplicar por escalares. Bajo esa definición, las funciones continuas en $[0, 1]$ son vectores: se suman, se escalan, hay un cero. El espacio que forman tiene dimensión infinita, y ahí es donde la historia se pone interesante, porque en dimensión infinita fallan cosas que en $\mathbb{R}^n$ se daban por sentadas: hay sucesiones acotadas sin subsucesión convergente, hay subespacios que no son cerrados, hay operadores lineales sin autovalores, y la bola unidad deja de ser compacta.
El análisis funcional es lo que queda del álgebra lineal cuando se le añade una topología y se le quita la dimensión finita. Su método es siempre el mismo: tratar una función como un punto, una ecuación diferencial o integral como un operador lineal actuando sobre ese punto, y resolverla con argumentos geométricos —proyecciones, ortogonalidad, bases— en vez de con manipulaciones analíticas.

Este capítulo usa el álgebra lineal del capítulo 7, la completitud y los teoremas de Baire y Arzelà–Ascoli del capítulo 29, y los espacios L^p del capítulo 31. A cambio, explica *por qué* funcionan las series de Fourier, qué es de verdad un desarrollo en autofunciones, y de dónde salen los espacios de Hilbert de la mecánica cuántica.

32.2. Aplicaciones lineales, funcionales y operadores acotados: el vocabulario que sostiene el capítulo

Todo lo que hace este capítulo consiste en tratar funciones como puntos y en estudiar aplicaciones que mueven esos puntos. Las secciones siguientes definen con calma la norma, la completitud y el producto interno; esta se ocupa de la otra mitad del vocabulario, la que se refiere a las *aplicaciones*: qué es un operador lineal, qué significa que sea acotado, qué es la norma de un operador, por qué acotado y continuo son sinónimos aquí y solo aquí, y qué es un funcional. Al final se añaden dos definiciones que el capítulo utiliza sin haberlas escrito: la ortogonalidad con su complemento $M^{\perp}$, y los coeficientes de Fourier abstractos.

Para leerla basta con saber lo que la sección siguiente formaliza: una *norma* $\|\cdot\|$ sobre un espacio vectorial mide el tamaño de un vector, se anula solo en el cero, cumple $\|\lambda x\| = |\lambda|\,\|x\|$ y satisface la desigualdad triangular; y la distancia asociada es $d(x,y) = \|x-y\|$.

Idea clave

Intuición primero. En $\mathbb{R}^n$ toda aplicación lineal es una matriz, y toda matriz es continua: no hay nada que discutir. En dimensión infinita la situación cambia por completo. La derivada, por ejemplo, es lineal —la derivada de una suma es la suma de las derivadas— y sin embargo puede convertir funciones diminutas en funciones enormes: $\sin(nx)/n$ mide $1/n$ y su derivada mide 1. Una aplicación así es lineal y *discontinua*, algo imposible en dimensión finita.
De ahí que en este capítulo la palabra "operador" venga casi siempre acompañada del adjetivo *acotado*. Y de ahí también el teorema que abre esta sección, y que es la razón de que todo funcione: para una aplicación lineal, ser continua y estar acotada son *exactamente* lo mismo. Un solo número, la norma del operador, decide la cuestión.

Definición 32.1 (Aplicación lineal, operador y funcional)**.** Sean X e Y espacios vectoriales sobre el mismo cuerpo $\mathbb{K}$ (que será $\mathbb{R}$ o $\mathbb{C}$). Una aplicación $T\colon X \to Y$ es *lineal* si

$$T(x_1+x_2) = T(x_1)+T(x_2) \quad \text{y} \quad T(\lambda x) = \lambda\, T(x)$$

para todos $x_1, x_2, x \in X$ y $\lambda \in \mathbb{K}$; las dos condiciones se resumen en $T(\lambda x+\mu y) = \lambda Tx + \mu Ty$. Es costumbre escribir Tx en lugar de $T(x)$. Cuando X e Y son espacios normados se dice *operador lineal*; y si el espacio de llegada es el propio cuerpo, $Y=\mathbb{K}$, se dice *funcional lineal*, y se suele denotar con letras griegas minúsculas, φ, ψ.

De la linealidad se sigue de inmediato $T0 = 0$, y también que el *núcleo* $\ker T = \{x : Tx = 0\}$ y la *imagen* $\operatorname{im} T = T(X)$ son subespacios vectoriales.

Ejemplo 32.2 (Cinco aplicaciones lineales del capítulo)**. (a) Una matriz.** $T\colon \mathbb{R}^2 \to \mathbb{R}^2$, $T(x,y) = (x+y,\, x-y)$. Lineal, como toda matriz.

(b) Un funcional integral. $\varphi\colon C[0,1] \to \mathbb{R}$, $\varphi(f) = \int_0^1 f(t)\,\mathrm{d}t$. Es lineal porque la integral lo es.

(c) La evaluación. $\delta_{t_0}\colon C[0,1] \to \mathbb{R}$, $\delta_{t_0}(f) = f(t_0)$. Lineal: $(f+g)(t_0) = f(t_0)+g(t_0)$.

(d) La derivada. $D\colon C^1[0,1] \to C[0,1]$, $Df = f'$. Lineal, y protagonista del ejemplo que sigue.

(e) El desplazamiento. $S\colon \ell^2 \to \ell^2$, $S(x_1, x_2, x_3, \dots) = (0, x_1, x_2, \dots)$. Lineal, inyectivo y no sobreyectivo a la vez, cosa imposible en dimensión finita.

Definición 32.3 (Operador acotado y norma de operador)**.** Sean X e Y espacios normados y $T\colon X \to Y$ lineal. Se dice que T es *acotado* si existe una constante $C \geq 0$ tal que

$$\|Tx\|_Y \leq C\,\|x\|_X \qquad \text{para todo } x \in X.$$

En ese caso, la *norma del operador* es el menor valor admisible de C, es decir

$$\|T\| = \inf\{C \geq 0 : \|Tx\| \leq C\|x\|\ \forall x\} = \sup_{x\neq 0} \frac{\|Tx\|}{\|x\|} = \sup_{\|x\|\leq 1} \|Tx\| = \sup_{\|x\|=1} \|Tx\|.$$

(Las cuatro expresiones coinciden por homogeneidad: $\|T(x/\|x\|)\| = \|Tx\|/\|x\|$; la última exige $X \neq \{0\}$.) El conjunto de los operadores acotados de X en Y se denota $B(X,Y)$, y $B(X) = B(X,X)$.

Aviso terminológico. "Acotado" aquí *no* significa que la imagen de T sea un conjunto acotado: salvo que $T = 0$, la imagen nunca lo es, porque $\|T(\lambda x)\| = |\lambda|\,\|Tx\|$ crece sin tope. Significa "acotado sobre la bola unidad", que por homogeneidad es lo mismo que "de crecimiento como mucho lineal".

Teorema 32.4 (Para una aplicación lineal, continuo y acotado son lo mismo)**.** *Sean X e Y normados y $T\colon X \to Y$ lineal. Las cinco afirmaciones siguientes son equivalentes:*

(i) T es continua en algún punto $x_0 \in X$;

(ii) T es continua en 0;

(iii) T es acotada, es decir, existe $C \geq 0$ con $\|Tx\| \leq C\|x\|$ para todo x;

(iv) T es lipschitziana: $\|Tx - Tz\| \leq \|T\|\,\|x - z\|$ para todos x, z;

(v) T es continua en todo punto de X.

Demostración. Demostramos el ciclo (i) ⇒ (ii) ⇒ (iii) ⇒ (iv) ⇒ (v) ⇒ (i).

(i) ⇒ (ii). Supongamos T continua en x_0 y sea $\varepsilon > 0$. Existe $\delta > 0$ tal que $\|z - x_0\| \leq \delta$ implica $\|Tz - Tx_0\| \leq \varepsilon$. Dado ahora u con $\|u\| \leq \delta$, aplíquese lo anterior a $z = x_0 + u$, que cumple $\|z - x_0\| = \|u\| \leq \delta$; por linealidad, $Tz - Tx_0 = T(x_0 + u) - Tx_0 = Tu$, de modo que $\|Tu\| \leq \varepsilon$. Como además $T0 = 0$, esto dice exactamente que T es continua en 0. *Este paso es toda la magia*: la linealidad traslada la continuidad de cualquier punto al origen, cosa que una función no lineal no puede hacer.

(ii) ⇒ (iii). Tomando $\varepsilon = 1$ en la continuidad en 0, hay $\delta > 0$ con $\|Tu\| \leq 1$ siempre que $\|u\| \leq \delta$. Sea ahora $x \neq 0$ cualquiera y póngase $u = \dfrac{\delta}{\|x\|}\,x$, que tiene $\|u\| = \delta$. Entonces $\|Tu\| \leq 1$, y por homogeneidad

$$\|Tu\| = \frac{\delta}{\|x\|}\,\|Tx\| \leq 1 \qquad \Longrightarrow \qquad \|Tx\| \leq \frac{1}{\delta}\,\|x\|.$$

Para $x = 0$ la desigualdad es trivial. Luego T es acotada con $C = 1/\delta$.

(iii) ⇒ (iv). Si $\|Tx\| \le C\|x\|$ para todo x, entonces el supremo que define $\|T\|$ es finito y $\|T\| \le C$; además $\|Tw\| \le \|T\|\,\|w\|$ para todo w, por definición de supremo. Aplicando esto a $w = x - z$ y usando la linealidad,

$$\|Tx - Tz\| = \|T(x - z)\| \le \|T\|\,\|x - z\|.$$

(iv) ⇒ (v). Una aplicación lipschitziana es continua: dado $\varepsilon > 0$, basta tomar $\delta = \varepsilon/(\|T\|+1)$ para que $\|x - z\| < \delta$ implique $\|Tx - Tz\| < \varepsilon$. Y esto vale en todo punto a la vez, es decir, la continuidad es incluso uniforme.

(v) ⇒ (i). Inmediato: continua en todo punto lo es en particular en uno. □

Observación 32.5. Merece la pena subrayar lo que el teorema dice y lo que no. Dice que para una aplicación *lineal* la continuidad es un fenómeno *global*, gobernado por un solo número $\|T\|$: o es continua en todas partes y uniformemente, o no lo es en ningún punto. No hay término medio, no hay operadores lineales continuos en unos puntos y discontinuos en otros. Nada de esto vale para funciones no lineales, donde la continuidad es local y hay funciones continuas en un único punto. Por eso en análisis funcional "operador acotado" y "operador continuo" se usan como sinónimos sin que nadie se moleste.

Ejemplo 32.6 (Un funcional lineal discontinuo, calculado entero)**.** Que existan aplicaciones lineales discontinuas es la razón de ser del teorema anterior. Sea $X = C^1[0,1]$, el espacio de las funciones con derivada continua, con la norma del supremo $\|f\|_\infty = \max_{[0,1]} |f|$, y considérese el funcional

$$\varphi\colon X \to \mathbb{R}, \qquad \varphi(f) = f'(0).$$

Es lineal, porque la derivada lo es y la evaluación también. Veamos que *no* es acotado exhibiendo la sucesión de testigos

$$f_n(x) = \frac{\sin(nx)}{n}, \qquad n = 1, 2, 3, \ldots$$

Cada f_n es de clase C^∞, y su tamaño es minúsculo:

$$\|f_n\|_\infty = \max_{[0,1]} \Big|\frac{\sin(nx)}{n}\Big| \le \frac{1}{n} \xrightarrow[n\to\infty]{} 0.$$

Pero su derivada es $f_n'(x) = \cos(nx)$, de modo que

$$\varphi(f_n) = f_n'(0) = \cos 0 = 1 \qquad \text{para todo } n.$$

Si φ fuera acotado con constante C, tendríamos $1 = |\varphi(f_n)| \le C\|f_n\|_\infty \le C/n$ para todo n, es decir $C \ge n$ para todo n: imposible. Luego φ no es acotado y, por el teorema 32.4, tampoco es continuo. La lectura directa: $f_n \to 0$ en X y sin embargo $\varphi(f_n) = 1 \not\to \varphi(0) = 0$. La continuidad falla en el origen y, por el teorema, en todos los demás puntos.

Obsérvese qué es lo que se rompe: la norma del supremo controla el *tamaño* de la función y no dice nada de su *pendiente*. Una función puede ser diminuta y oscilar violentamente. Con la norma $\|f\|_\infty + \|f'\|_\infty$ el mismo funcional pasaría a ser acotado, con $\|\varphi\| = 1$. *La acotación no es propiedad de la aplicación sola, sino de la aplicación y las dos normas elegidas.*

Ejemplo 32.7 (La misma matriz, tres normas distintas)**.** Sea $T(x,y)=(x+y,\,x-y)$, de matriz $A=\left(\begin{smallmatrix}1&1\\1&-1\end{smallmatrix}\right)$. Su norma de operador depende de con qué norma se mida $\mathbb{R}^2$.

(a) Norma euclídea. Aquí $\|T(x,y)\|_2^2=(x+y)^2+(x-y)^2=2x^2+2y^2=2\|(x,y)\|_2^2$, luego $\|T(x,y)\|_2=\sqrt{2}\,\|(x,y)\|_2$ para *todo* vector. Así que $\|T\|_2=\sqrt{2}\approx 1{,}41421$, y cualquier vector unitario sirve de testigo: T es $\sqrt{2}$ veces una isometría.

(b) Norma del máximo. Con $\|(x,y)\|_\infty=\max(|x|,|y|)$ se tiene $|x\pm y|\le 2\max(|x|,|y|)$, luego $\|T\|_\infty\le 2$; y el testigo es $(1,1)$, de norma 1, con $T(1,1)=(2,0)$ de norma 2. Por tanto $\|T\|_\infty=2$, el máximo de las sumas por filas de $|A|$.

(c) Norma de la suma. Con $\|(x,y)\|_1=|x|+|y|$ sale $\|T(x,y)\|_1=|x+y|+|x-y|\le 2(|x|+|y|)$, y el testigo $(1,0)$ da $T(1,0)=(1,1)$ de norma 1... que no alcanza el 2. El testigo bueno es de nuevo $(1,1)$, pero ahora tiene norma 2 y su imagen $(2,0)$ tiene norma 2: cociente 1. En realidad $\|T\|_1=2$, el máximo de las sumas por columnas, alcanzado en $(1,0)$ cuando se mide bien: $\|T(1,0)\|_1=2$ y $\|(1,0)\|_1=1$. ✓

Tres números distintos —$\sqrt{2}$, 2 y 2— para la misma aplicación lineal. La norma de un operador nunca es un dato del operador solo.

Proposición 32.8 ($B(X,Y)$ es un espacio normado)**.** *Sean X,Y,Z normados. Entonces:*

1. *$B(X,Y)$ es un espacio vectorial y $\|\cdot\|$ es una norma sobre él;*
2. *si $T\in B(X,Y)$ y $S\in B(Y,Z)$, entonces $ST\in B(X,Z)$ y $\|ST\|\le\|S\|\,\|T\|$;*
3. *si Y es completo, $B(X,Y)$ también lo es.*

Demostración. (1) Si $\|Tx\|\le\|T\|\|x\|$ y $\|Sx\|\le\|S\|\|x\|$, entonces $\|(T+S)x\|\le(\|T\|+\|S\|)\|x\|$, lo que da a la vez que $T+S$ es acotado y la desigualdad triangular $\|T+S\|\le\|T\|+\|S\|$. La homogeneidad $\|\lambda T\|=|\lambda|\|T\|$ sale de sacar $|\lambda|$ del supremo. Y $\|T\|=0$ obliga a $\|Tx\|\le 0$ para todo x, o sea $T=0$.

(2) $\|S(Tx)\|\le\|S\|\,\|Tx\|\le\|S\|\,\|T\|\,\|x\|$, y se toma el supremo sobre $\|x\|\le 1$.

(3) Sea (T_n) de Cauchy en $B(X,Y)$. Para cada x fijo, $\|T_nx-T_mx\|\le\|T_n-T_m\|\,\|x\|$, luego (T_nx) es de Cauchy en Y y converge a un vector que llamamos Tx. La aplicación T así definida es lineal por paso al límite en $T_n(\lambda x+\mu y)=\lambda T_nx+\mu T_ny$. Es acotada porque $(\|T_n\|)$ es de Cauchy en $\mathbb{R}$ y por tanto acotada por un M, de donde $\|Tx\|=\lim\|T_nx\|\le M\|x\|$. Y $T_n\to T$ en norma de operador: dado $\varepsilon>0$, para $n,m\ge N$ se tiene $\|T_nx-T_mx\|\le\varepsilon\|x\|$, y haciendo $m\to\infty$ queda $\|T_nx-Tx\|\le\varepsilon\|x\|$ para todo x, es decir $\|T_n-T\|\le\varepsilon$. □

Definición 32.9 (Dual topológico)**.** El *dual topológico* de un espacio normado X es el espacio $X^*=B(X,\mathbb{K})$ de los funcionales lineales *acotados* sobre X, con la norma

$$\|\varphi\|=\sup_{\|x\|\le 1}|\varphi(x)|.$$

Por la proposición 32.8(3), y como $\mathbb{K}$ es completo, X^* es siempre un espacio de Banach, *aunque X no lo sea*. El adjetivo "topológico" distingue este espacio del dual algebraico, formado por *todos* los funcionales lineales, acotados o no; el ejemplo anterior muestra que ambos pueden ser distintos.

Definición 32.10 (Ortogonalidad y complemento ortogonal)**.** Sea H un espacio con producto interno $\langle\cdot,\cdot\rangle$. Dos vectores x, y son *ortogonales*, y se escribe $x \perp y$, si $\langle x, y\rangle = 0$. Dado un subconjunto $A \subseteq H$, su *complemento ortogonal* es

$$A^{\perp} = \{\, x \in H : \ \langle x, a\rangle = 0 \text{ para todo } a \in A \,\}.$$

Una familia (e_i) es *ortogonal* si sus elementos son ortogonales dos a dos, y *ortonormal* si además $\|e_i\| = 1$ para todo i.

Proposición 32.11 (Pitágoras y las propiedades de $A^{\perp}$)**.** *Sea H con producto interno y $A \subseteq H$.*

1. *Si $x \perp y$, entonces $\|x+y\|^2 = \|x\|^2+\|y\|^2$; y más en general, si $x_1, \ldots, x_n$ son ortogonales dos a dos, $\left\|\sum x_k\right\|^2 = \sum \|x_k\|^2$.*
2. *$A^{\perp}$ es siempre un subespacio vectorial* cerrado *de H, sea A lo que sea.*
3. *$A \cap A^{\perp} \subseteq \{0\}$, y $A \subseteq (A^{\perp})^{\perp}$.*

Demostración. (1) $\|x+y\|^2 = \langle x+y, x+y\rangle = \|x\|^2 + \langle x, y\rangle + \langle y, x\rangle + \|y\|^2$, y los dos términos centrales son nulos. El caso general sale por inducción.

(2) Es subespacio porque la condición $\langle x, a\rangle = 0$ es lineal en x. Es cerrado porque

$$A^{\perp} = \bigcap_{a \in A} \{x : \langle x, a\rangle = 0\} = \bigcap_{a \in A} \varphi_a^{-1}(\{0\}),$$

donde $\varphi_a(x) = \langle x, a\rangle$ es continuo por Cauchy–Schwarz ($|\varphi_a(x)| \le \|a\|\,\|x\|$, luego φ_a es acotado con $\|\varphi_a\| \le \|a\|$). Cada conjunto de la intersección es la preimagen de un cerrado por una función continua, y una intersección arbitraria de cerrados es cerrada.

(3) Si $x \in A \cap A^{\perp}$, entonces $\langle x, x\rangle = 0$, luego $x = 0$. La segunda inclusión es inmediata de la simetría de la condición. □

Ejemplo 32.12 (Proyección y complemento ortogonal en $\mathbb{R}^3$)**.** Sea $H = \mathbb{R}^3$ con el producto escalar usual y $M = \operatorname{span}\{(1,1,0)\}$. Un vector unitario de M es $e = \frac{1}{\sqrt{2}}(1,1,0)$. Proyectemos $v = (3,1,4)$:

$$\langle v, e\rangle = \frac{3+1+0}{\sqrt{2}} = \frac{4}{\sqrt{2}} = 2\sqrt{2}, \qquad P_M v = \langle v, e\rangle\, e = 2\sqrt{2}\cdot\frac{1}{\sqrt{2}}(1,1,0) = (2,2,0).$$

El residuo es $v - P_M v = (1,-1,4)$, y en efecto $\langle (1,-1,4), (1,1,0)\rangle = 1-1+0 = 0$: es ortogonal a M, tal como exige la caracterización de la proyección. La distancia de v a M es

$$\|v - P_M v\| = \sqrt{1+1+16} = \sqrt{18} = 3\sqrt{2} \approx 4{,}24264.$$

Y Pitágoras cuadra: $\|v\|^2 = 9+1+16 = 26$, mientras que $\|P_M v\|^2 + \|v - P_M v\|^2 = 8+18 = 26$. ✓

El complemento ortogonal es $M^{\perp} = \{(a,b,c) : a+b=0\} = \{(a,-a,c)\}$, un plano de dimensión 2, generado por $(1,-1,0)$ y $(0,0,1)$. Como debe ser, $\dim M + \dim M^{\perp} = 1+2 = 3$ y $\mathbb{R}^3 = M \oplus M^{\perp}$: cada vector se escribe de forma única como suma de su proyección y su residuo. En dimensión infinita esta descomposición sigue siendo cierta, pero *solo si M es cerrado*, y esa hipótesis es una de las que más se olvidan.

Definición 32.13 (Coeficientes de Fourier abstractos)**.** Sea H un espacio con producto interno y $(e_n)_{n\ge 1}$ una familia ortonormal. Para cada $x \in H$, los números

$$\hat{x}_n = \langle x, e_n \rangle \qquad (n = 1, 2, \dots)$$

se llaman *coeficientes de Fourier* de x respecto del sistema (e_n), y la serie formal $\sum_n \hat{x}_n\, e_n$ es la *serie de Fourier abstracta* de x. Que esa serie converja y que su suma sea x son dos afirmaciones distintas y ambas requieren demostración: la primera es consecuencia de la desigualdad de Bessel y de la completitud de H; la segunda es equivalente a que el sistema sea una base ortonormal, y es el contenido de la identidad de Parseval.

Se llama *serie* de un espacio normado a la sucesión de sumas parciales $S_N = \sum_{n\le N} x_n$; decir que $\sum x_n$ converge a s significa exactamente que $\|S_N - s\| \to 0$, ni más ni menos.

Ejemplo 32.14 (Coeficientes de Fourier de la función constante)**.** En $H = L^2[0,1]$ con $\langle f, g\rangle = \int_0^1 fg$, el sistema $e_n(x) = \sqrt{2}\,\sin(n\pi x)$ es ortonormal: un cálculo directo da $\int_0^1 2\sin^2(n\pi x)\,\mathrm{d}x = 1$ y $\int_0^1 2\sin(n\pi x)\sin(m\pi x)\,\mathrm{d}x = 0$ para $n \neq m$. Tomemos la función constante $f \equiv 1$, cuyos coeficientes son

$$\hat{f}_n = \int_0^1 \sqrt{2}\,\sin(n\pi x)\,\mathrm{d}x = \sqrt{2}\left[\frac{-\cos(n\pi x)}{n\pi}\right]_0^1 = \frac{\sqrt{2}\,(1-\cos(n\pi))}{n\pi} = \begin{cases} \dfrac{2\sqrt{2}}{n\pi}, & n \text{ impar},\\[2mm] 0, & n \text{ par}. \end{cases}$$

Los primeros valen $\hat{f}_1 = \frac{2\sqrt{2}}{\pi} \approx 0{,}90032$, $\hat{f}_2 = 0$, $\hat{f}_3 = \frac{2\sqrt{2}}{3\pi} \approx 0{,}30011$, $\hat{f}_4 = 0$, $\hat{f}_5 \approx 0{,}18006$. La suma de sus cuadrados es

$$\sum_{n \text{ impar}} \frac{8}{n^2\pi^2} = \frac{8}{\pi^2}\sum_{k\ge 1}\frac{1}{(2k-1)^2} = \frac{8}{\pi^2}\cdot\frac{\pi^2}{8} = 1,$$

que coincide exactamente con $\|f\|^2 = \int_0^1 1\,\mathrm{d}x = 1$. La igualdad —y no solo la desigualdad de Bessel— indica que el sistema es completo en $L^2[0,1]$, que es justamente el desarrollo en serie de senos de las ecuaciones del calor con extremos fijos. Nótese de paso que la serie converge a f *en norma* L^2, y desde luego no puntualmente en los extremos, donde cada e_n se anula y f vale 1.

Regla: decidir si una aplicación lineal es continua

La pregunta se resuelve siempre con el teorema 32.4, en una dirección o en la otra.

1. **Comprueba primero que es lineal.** Si no lo es, el teorema no aplica y hay que estudiar la continuidad punto a punto, como en cálculo.

2. **Para probar que *sí* es continua, busca la desigualdad $\|Tx\| \le C\|x\|$.** Se manipula la expresión de Tx con las herramientas disponibles —triangular, Cauchy–Schwarz, Hölder, $|\int| \le \int|\cdot|$— hasta que aparezca $\|x\|$ multiplicado por una constante. Esa constante ya es una cota de $\|T\|$.

3. **Para probar que *no* lo es, construye una sucesión testigo.** Hace falta (x_n) con $\|x_n\|$ acotada (o incluso tendiendo a 0) y $\|Tx_n\| \to \infty$. La forma canónica es

tomar vectores de norma 1 cuya imagen crezca; si además se dividen por algo que tienda a infinito, se obtiene el contraste más vistoso, como con $\sin(nx)/n$.

4. **Recuerda que basta mirar el origen.** Por el paso (i)⇒(ii) del teorema, comprobar la continuidad en 0 resuelve el problema entero. No hay que estudiar ningún otro punto.

5. **Sospecha siempre de la derivada y de la evaluación en normas integrales.** Los dos ejemplos clásicos de discontinuidad son $f \mapsto f'$ con la norma del supremo y $f \mapsto f(t_0)$ con una norma L^p: ambas normas miden promedios o tamaños globales y no controlan el comportamiento local.

6. **En dimensión finita, no te molestes.** Toda aplicación lineal entre espacios normados de dimensión finita es continua, porque todas las normas son equivalentes y en coordenadas T es una matriz. Los ejemplos interesantes viven solo en dimensión infinita.

7. **Control.** $\|T\| \geq \|Tx\|/\|x\|$ para cualquier x concreto que se te ocurra. Evalúa en dos o tres vectores sencillos: si alguno da un cociente mayor que la $\|T\|$ que has calculado, la cuenta está mal.

Regla: usar la ortogonalidad para calcular en un espacio con producto interno

1. **Normaliza antes de proyectar.** Todas las fórmulas de esta sección suponen $\|e_k\| = 1$. Si tus generadores no son unitarios, divídelos por su norma; y si no son ortogonales, aplica Gram–Schmidt primero.

2. **Calcula cada coeficiente por separado.** $\hat{x}_k = \langle x, e_k \rangle$ es una integral o una suma independiente de las demás. Ese es todo el ahorro de trabajar con bases ortonormales: no hay ningún sistema que resolver.

3. **Suma para proyectar.** $P_M x = \sum_k \hat{x}_k \, e_k$ sobre los índices que generan M.

4. **Verifica con la ortogonalidad del residuo.** $x - P_M x$ debe ser ortogonal a todos los e_k. Es el control más rápido y detecta casi cualquier error aritmético.

5. **Calcula el error con Pitágoras, no desarrollando cuadrados.** $\|x - P_M x\|^2 = \|x\|^2 - \sum_k |\hat{x}_k|^2$, por la proposición 32.11(1).

6. **Distingue Bessel de Parseval.** La desigualdad $\sum |\hat{x}_k|^2 \leq \|x\|^2$ vale siempre; la igualdad solo si el sistema es completo. Si en un cálculo te sale una igualdad con un sistema del que no has comprobado la completitud, has supuesto lo que había que demostrar.

7. **Recuerda que $A^\perp$ es cerrado aunque A no lo sea.** Por eso $(A^\perp)^\perp$ es en general más grande que A: es, de hecho, la clausura del subespacio generado por A.

Pausa para pensar

Compara las dos mitades de esta sección. La primera dice que un operador lineal queda descrito, en lo tocante a continuidad, por un único número $\|T\|$; la segunda, que en un espacio con producto interno la geometría —ángulos, perpendicularidad, teorema de Pitágoras— sobrevive intacta a la dimensión infinita. Todo el capítulo vive de la tensión entre esas dos afirmaciones. Los espacios de Banach solo tienen la primera y por eso sus teoremas son más duros de demostrar y menos constructivos; los de Hilbert tienen las dos, y por eso en ellos casi todo se reduce a proyectar sobre algo y mirar qué queda perpendicular.

32.3. Espacios normados y espacios de Banach

Definición 32.15 (Norma)**.** Una *norma* sobre un espacio vectorial X sobre $\mathbb{R}$ o $\mathbb{C}$ es una función $\|\cdot\| : X \to [0,\infty)$ con

(N1) $\|x\| = 0 \iff x = 0$;

(N2) $\|\lambda x\| = |\lambda|\,\|x\|$;

(N3) $\|x+y\| \le \|x\| + \|y\|$ (desigualdad triangular).

El par $(X, \|\cdot\|)$ es un *espacio normado*; $d(x,y) = \|x-y\|$ lo convierte en espacio métrico. Si esa métrica es completa, X es un *espacio de Banach*.

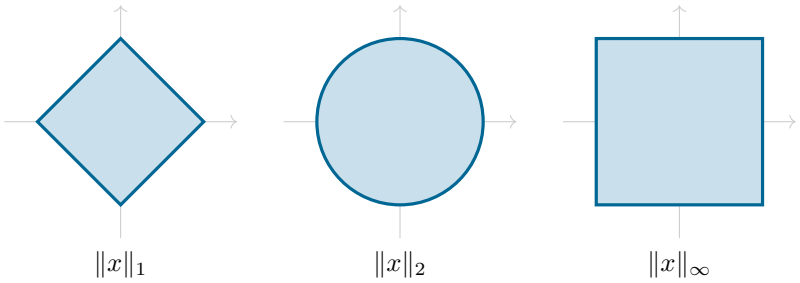


Figura 32.1: La bola unidad de $\mathbb{R}^2$ según se mida con la suma de valores absolutos, con la distancia euclídea o con el máximo. Las tres normas son equivalentes en dimensión finita, y esa equivalencia es precisamente lo que se pierde en dimensión infinita.

Ejemplo 32.16 (El catálogo básico)**.**

- $\mathbb{R}^n$ con $\|x\|_p = \left(\sum |x_i|^p\right)^{1/p}$ para $1 \le p < \infty$, y $\|x\|_\infty = \max |x_i|$. Todos son Banach.
- $\ell^p = \{(x_n) : \sum |x_n|^p < \infty\}$ con la norma análoga; ℓ^∞ = sucesiones acotadas con el supremo. Todos son Banach.
- c_0 = sucesiones que tienden a 0, con la norma del supremo: es Banach, subespacio cerrado de ℓ^∞.

- $C[a,b]$ con $\|f\|_\infty = \max|f|$: es Banach (un límite uniforme de continuas es continua).
- $L^p(\Omega)$ con $\|f\|_p = (\int |f|^p)^{1/p}$: Banach por el teorema de Riesz–Fischer del capítulo 31.

Cuidado

$C[0,1]$ con la norma $\|f\|_1 = \int_0^1 |f|$ *no* es completo. La sucesión de funciones continuas que suben de 0 a 1 en el intervalo $[\frac{1}{2}, \frac{1}{2}+\frac{1}{n}]$ es de Cauchy en $\|\cdot\|_1$, pero su límite es la función escalón, que no es continua. La compleción de ese espacio es precisamente $L^1[0,1]$: la integral de Lebesgue no se inventó por gusto, sino para tener un espacio donde los límites existan. **La misma colección de funciones con dos normas distintas puede ser Banach o no serlo.**

Teorema 32.17 (En dimensión finita todas las normas son equivalentes)**.** *Si* $\dim X = n < \infty$ *y* $\|\cdot\|_a$, $\|\cdot\|_b$ *son normas sobre* X*, existen* $c, C > 0$ *con* $c\|x\|_a \le \|x\|_b \le C\|x\|_a$ *para todo* x*. En consecuencia todo espacio normado de dimensión finita es Banach y todos sus subespacios son cerrados.*

Demostración. Fija una base y compara toda norma con $\|\cdot\|_\infty$ en coordenadas. Por la desigualdad triangular, $\|x\|_a \le (\sum \|e_i\|_a)\|x\|_\infty =: C_a\|x\|_\infty$, lo que además muestra que $\|\cdot\|_a$ es continua respecto de $\|\cdot\|_\infty$. La esfera $\{\|x\|_\infty = 1\}$ es compacta (Heine–Borel en coordenadas), y $\|\cdot\|_a$ es continua y positiva en ella, luego alcanza un mínimo $c_a > 0$. Por homogeneidad, $c_a\|x\|_\infty \le \|x\|_a \le C_a\|x\|_\infty$ para todo x, y componiendo las dos comparaciones se obtiene el resultado. □

Idea clave

Qué dice esto en cristiano. En dimensión finita da igual con qué norma se trabaje: todas producen los mismos abiertos, las mismas sucesiones convergentes y las mismas funciones continuas. En dimensión infinita eso se rompe por completo, y elegir la norma pasa a ser una decisión con consecuencias. El teorema de Riesz remata la idea: la bola unidad cerrada es compacta *exactamente* en dimensión finita. Por eso todos los argumentos de análisis que se apoyan en compacidad hay que reinventarlos al pasar a espacios de funciones.

Ejemplo 32.18 (Normas equivalentes y normas que no lo son)**. (a) En $\mathbb{R}^2$ todas dan lo mismo.** Para $v = (3,4)$: $\|v\|_1 = 7$, $\|v\|_2 = 5$ y $\|v\|_\infty = 4$. Son números distintos, pero se acotan mutuamente: $\|v\|_\infty \le \|v\|_2 \le \|v\|_1 \le 2\|v\|_\infty$. Una sucesión converge en una norma si y solo si converge en las otras.

(b) En dimensión infinita no. En el espacio de funciones continuas sobre $[0,1]$, sea $f_n(x) = x^n$. Entonces $\|f_n\|_1 = \frac{1}{n+1} \to 0$, mientras que $\|f_n\|_\infty = 1$ para todo n. La sucesión converge a cero en la norma L^1 y no converge en la del supremo: las dos normas definen topologías distintas.

(c) La bola unidad no es compacta. En ℓ^2, la sucesión de vectores de la base canónica $e_1, e_2, e_3, \ldots$ vive toda en la bola unidad, y la distancia entre dos cualesquiera es $\|e_i - e_j\| = \sqrt{2}$.

Ninguna subsucesión puede ser de Cauchy, luego ninguna converge. La bola es cerrada y acotada y no es compacta.

(d) La consecuencia práctica. En dimensión finita, "cerrado y acotado" basta para garantizar que una función continua alcance su máximo. En dimensión infinita no, y hay que buscar sustitutos: la compacidad débil, la convexidad o la semicontinuidad inferior. Media disciplina del análisis funcional consiste en encontrar reemplazos de la compacidad perdida.

Cuidado

[Equivalente no significa igual] Que dos normas sean equivalentes no significa que den los mismos números, sino que se acotan mutuamente por constantes, y por tanto tienen las mismas sucesiones convergentes. En (a) los tres valores son 7, 5 y 4, y las normas son equivalentes. Segundo aviso: el teorema de Riesz es una equivalencia, así que también funciona al revés como criterio de dimensión infinita; si encuentras una sucesión acotada sin subsucesión convergente, ya sabes que el espacio no es de dimensión finita.

Teorema 32.19 (Riesz: la bola unidad delata la dimensión)**.** *En un espacio normado X, la bola cerrada unidad es compacta si y solo si* $\dim X < \infty$.

Demostración. Si la dimensión es finita, es Heine–Borel vía el teorema anterior. Recíprocamente, supón $\dim X = \infty$ y construye por inducción vectores unitarios $x_1, x_2, \dots$ con $\|x_n - x_m\| \geq \frac{1}{2}$ para $n \neq m$. Para el paso inductivo se usa el *lema de Riesz*: si $Y \subsetneq X$ es un subespacio cerrado propio, existe x unitario con $\operatorname{dist}(x, Y) \geq \frac{1}{2}$. (Prueba del lema: toma $z \notin Y$, sea $d = \operatorname{dist}(z, Y) > 0$ por ser Y cerrado, elige $y_0 \in Y$ con $\|z - y_0\| \leq 2d$ y pon $x = \frac{z-y_0}{\|z-y_0\|}$; entonces para todo $y \in Y$, $\|x - y\| = \frac{\|z-y_0 - \|z-y_0\| y\|}{\|z-y_0\|} \geq \frac{d}{2d} = \frac{1}{2}$.) Aplicando el lema a $Y_n = \operatorname{span}(x_1, \dots, x_n)$, que es cerrado por tener dimensión finita, se obtiene la sucesión. Esa sucesión vive en la bola unidad y no tiene ninguna subsucesión de Cauchy: la bola no es compacta. □

Idea clave

Qué dice esto en cristiano. En dimensión finita, "cerrado y acotado" equivale a compacto y de toda sucesión acotada se extrae una convergente. El teorema de Riesz dice que esa comodidad es un privilegio exclusivo de la dimensión finita: en cuanto el espacio es infinito-dimensional, la bola unidad deja de ser compacta, sin excepciones. Es una frontera limpia, y es la razón de que el análisis funcional exista como disciplina separada: hay que inventar sustitutos de la compacidad (convergencia débil, operadores compactos, reflexividad) porque la original ya no está disponible.

Ejemplo 32.20 (La bola unidad en cuatro espacios)**. (a) En $\mathbb{R}^n$.** La bola unidad es compacta por Heine–Borel, y por eso toda función continua sobre ella alcanza su máximo. Todo el cálculo de optimización clásico descansa en este hecho.

(b) En ℓ^2. Los vectores $e_1, e_2, e_3, \dots$ de la base canónica están todos en la bola unidad y cumplen $\|e_i - e_j\| = \sqrt{2}$ para $i \neq j$. Ninguna subsucesión es de Cauchy, luego ninguna

converge: la bola no es compacta. Sin embargo $e_n \to 0$ *débilmente*, porque $\langle e_n, x\rangle = x_n \to 0$ para todo $x \in \ell^2$. La convergencia débil es el sustituto que sí funciona.

(c) En $C[0,1]$. Las funciones $f_n(x) = x^n$ tienen todas norma 1 y convergen puntualmente a una función discontinua, luego no hay subsucesión uniformemente convergente. Aquí el sustituto es el teorema de Arzelà–Ascoli, que recupera la compacidad a cambio de exigir equicontinuidad.

(d) El lema que lo demuestra. El lema de Riesz dice que si $Y \subsetneq X$ es un subespacio cerrado propio, para cada $\theta < 1$ existe x de norma 1 con $d(x, Y) \geq \theta$. Iterándolo se construye una sucesión de vectores unitarios separados entre sí más de $1/2$, que no puede tener subsucesión convergente. En dimensión finita el proceso se detiene tras n pasos; en dimensión infinita no se detiene nunca.

Idea clave

Qué significa esto. En $\mathbb{R}^n$, la estrategia estándar del análisis es: toma una sucesión acotada, extrae una subsucesión convergente, pasa al límite. En dimensión infinita esa estrategia está prohibida por el teorema de Riesz, y hay que sustituirla por algo. Las tres sustituciones históricas son: (i) pedir más a la familia —equicontinuidad— y usar Arzelà–Ascoli; (ii) cambiar de topología y usar convergencia débil (sección final de este capítulo); (iii) trabajar con operadores *compactos*, que por definición devuelven la compacidad perdida. Casi todo el análisis funcional aplicado es una de esas tres.

Regla: decidir si una norma viene de un producto interno

Todo espacio de Hilbert es de Banach, pero no al revés, y la diferencia es exactamente si la norma "recuerda ángulos". Hay un test limpio que lo decide sin ambigüedad.

1. **Prueba la ley del paralelogramo.** Una norma procede de un producto interno si y solo si

$$\|x+y\|^2 + \|x-y\|^2 = 2\|x\|^2 + 2\|y\|^2 \qquad \text{para todos } x, y.$$

   Este es el teorema de Jordan–von Neumann, y es un si y solo si: no hay excepciones, no hay hipótesis extra.

2. **Para refutar, basta un par de vectores.** Si sospechas que no, no busques una demostración general: elige dos vectores sencillos con soportes disjuntos —típicamente e_1 y e_2 — y comprueba que la identidad falla. Una sola pareja tumba la norma entera.

3. **Si se cumple, el producto se recupera por polarización.** En el caso real,

$$\langle x, y\rangle = \tfrac{1}{4}\big(\|x+y\|^2 - \|x-y\|^2\big),$$

   y en el complejo hay que añadir los términos con i. Es decir: el producto interno no es un dato adicional, está codificado en la norma.

4. **Aparte, comprueba la completitud.** Que la norma sea euclídea no dice nada sobre si el espacio es completo. Banach quiere decir "normado y completo"; Hilbert quiere decir "con producto interno y completo". Son dos propiedades independientes que se verifican por separado.

5. **Para probar completitud, el método estándar tiene tres pasos.** Dada una sucesión de Cauchy: (i) encuentra un candidato a límite mediante una convergencia más débil (puntual, o coeficiente a coeficiente); (ii) comprueba que el candidato está en el espacio; (iii) comprueba que la convergencia es en la norma del espacio, no solo puntual. Saltarse (ii) es el error clásico.

6. **Ten a mano la lista de casos.** ℓ^p y L^p son de Banach para todo $1 \le p \le \infty$, y de Hilbert *solo* para $p = 2$. $C[a,b]$ con la norma del supremo es de Banach pero no de Hilbert. $C[a,b]$ con la norma L^2 tiene producto interno pero *no* es completo.

Ejemplo 32.21 (El test aplicado a tres normas de $\mathbb{R}^2$, y un espacio incompleto)**. (a) La norma del supremo no es euclídea.** En $\mathbb{R}^2$ con $\|(u,v)\|_\infty = \max(|u|,|v|)$, tomemos $x = (1,0)$ e $y = (0,1)$. Entonces $x+y = (1,1)$ y $x-y = (1,-1)$, luego

$$\|x+y\|_\infty^2 + \|x-y\|_\infty^2 = 1+1 = 2, \qquad 2\|x\|_\infty^2 + 2\|y\|_\infty^2 = 2+2 = 4.$$

$2 \neq 4$: la ley del paralelogramo falla y no existe ningún producto interno que induzca $\|\cdot\|_\infty$. Fin.

(b) La norma de la suma tampoco. Con $\|(u,v)\|_1 = |u| + |v|$ y los mismos vectores: $\|x+y\|_1 = 2$, $\|x-y\|_1 = 2$, luego el miembro izquierdo vale $4+4 = 8$ frente a $2+2 = 4$ del derecho. Falla otra vez, y en la dirección contraria a la de (a).

(c) La norma euclídea, por supuesto, sí. Con $\|\cdot\|_2$ los dos miembros valen 4. Y no es casualidad: la identidad

$$\|x+y\|^2 + \|x-y\|^2 = (\|x\|^2 + 2\langle x,y\rangle + \|y\|^2) + (\|x\|^2 - 2\langle x,y\rangle + \|y\|^2)$$

muestra que los términos cruzados se cancelan siempre que haya un producto interno detrás. La ley del paralelogramo *es* esa cancelación, y su nombre viene de la geometría elemental: en un paralelogramo, la suma de los cuadrados de las diagonales iguala a la suma de los cuadrados de los cuatro lados.

(d) Producto interno sin completitud. Sea $X = C[0,1]$ con $\langle f,g\rangle = \int_0^1 fg$. Hay producto interno, luego se cumple el paralelogramo. Pero X no es completo: la sucesión de funciones continuas

$$f_n(x) = \begin{cases} 0, & x \le \tfrac{1}{2}, \\ n(x - \tfrac{1}{2}), & \tfrac{1}{2} < x < \tfrac{1}{2} + \tfrac{1}{n}, \\ 1, & x \ge \tfrac{1}{2} + \tfrac{1}{n}, \end{cases}$$

es de Cauchy en la norma L^2 (dos de ellas difieren solo en un intervalo de longitud $\le 1/n$, luego $\|f_n - f_m\|_2^2 \le 1/\min(n,m)$), y converge en L^2 al escalón de Heaviside, que *no es continuo*. El límite existe, pero se ha salido del espacio: exactamente el fallo del paso (ii) de la receta

de completitud. Completar X da $L^2[0,1]$, y esa es una de las motivaciones históricas de la integral de Lebesgue.

Regla: decidir si un espacio normado es completo

Ser de Banach no es un adorno: es lo que permite construir objetos como límites de aproximaciones sucesivas. Comprobarlo tiene una rutina.

1. **Recuerda qué hay que probar:** toda sucesión de Cauchy converge *dentro del espacio*. Las dos maneras de fallar son que el límite no exista, o que exista pero se salga del espacio.
2. **Reconoce los completos de catálogo:** $\mathbb{R}^n$ y $\mathbb{C}^n$ con cualquier norma; ℓ^p para $1 \le p \le \infty$; $L^p(\mu)$ para $1 \le p \le \infty$ (teorema de Riesz–Fischer); $C(K)$ con la norma del supremo y K compacto; y todo espacio de operadores $\mathcal{B}(X,Y)$ con Y de Banach.
3. **Usa el criterio del subespacio:** un subespacio de un espacio de Banach es completo si y solo si es *cerrado*. Éste es el atajo más útil de todos: en vez de manejar sucesiones de Cauchy, basta con demostrar que el subespacio es cerrado.
4. **Cuidado con lo de dimensión infinita:** en dimensión finita todas las normas son equivalentes y todo subespacio es cerrado. En dimensión infinita ambas cosas fallan, y es justo ahí donde aparecen los ejemplos incompletos.
5. **Para construir un contraejemplo, busca un límite que se escape.** Exhibe una sucesión del espacio que sea de Cauchy pero cuyo límite natural viva en un espacio mayor: funciones continuas que convergen a una discontinua, sucesiones finitas que convergen a una infinita, polinomios que convergen a algo que no lo es.
6. **Comprueba las series:** un normado es completo si y solo si toda serie absolutamente convergente converge, es decir, $\sum \|x_n\| < \infty$ implica que $\sum x_n$ converge. A veces es más cómodo que trabajar con sucesiones de Cauchy.
7. **Control:** si el espacio es completo, tienes derecho a usar Baire y con él los cuatro teoremas grandes. Si no lo es, esos teoremas no se aplican, y es exactamente ahí donde suelen aparecer los contraejemplos famosos.

Ejemplo 32.22 (Cuatro espacios examinados)**. (a)** $C[0,1]$ **con la norma del supremo: completo.** Si (f_n) es de Cauchy en $\|\cdot\|_\infty$, entonces para cada x la sucesión numérica $(f_n(x))$ es de Cauchy en $\mathbb{R}$ y converge a un $f(x)$. La convergencia es uniforme (basta pasar al límite en $|f_n(x) - f_m(x)| \le \varepsilon$ dejando $m \to \infty$), y un límite uniforme de continuas es continuo. El límite existe y sigue dentro: es de Banach.

(b) El mismo $C[0,1]$ **con la norma** $\|f\|_1 = \int_0^1 |f|$**: *no* completo.** Tomemos la rampa

$$f_n(x) = \begin{cases} 0, & x \le \frac{1}{2}, \\ n\left(x - \frac{1}{2}\right), & \frac{1}{2} < x < \frac{1}{2} + \frac{1}{n}, \\ 1, & x \ge \frac{1}{2} + \frac{1}{n}. \end{cases}$$

Cada f_n es continua y para $m > n$ se tiene $\|f_n - f_m\|_1 \le \frac{1}{2n} \to 0$, luego la sucesión es de Cauchy. Pero en L^1 converge a la función escalón $\chi_{[1/2,1]}$, que no es continua: el límite se escapa. La compleción de $(C[0,1], \|\cdot\|_1)$ es precisamente $L^1[0,1]$.

(c) Las sucesiones de soporte finito c_{00} con la norma del supremo: *no* completo. Sea $x^{(n)} = (1, \frac{1}{2}, \frac{1}{3}, \dots, \frac{1}{n}, 0, 0, \dots)$. Para $m > n$, $\|x^{(n)} - x^{(m)}\|_\infty = \frac{1}{n+1} \to 0$, así que es de Cauchy; pero el límite es $(1, \frac{1}{2}, \frac{1}{3}, \dots)$, que tiene infinitas entradas no nulas. Por el paso 3 la explicación estructural es que c_{00} es un subespacio *denso* y no cerrado de c_0, y un subespacio denso propio nunca es cerrado.

(d) Los polinomios dentro de $C[0,1]$: *no* completo. Por el teorema de Weierstrass los polinomios son densos en $C[0,1]$, y como no son todo $C[0,1]$ (por ejemplo e^x no es un polinomio), no son cerrados. Paso 3 otra vez: no completo. Un testigo explícito es $p_n(x) = \sum_{k=0}^{n} x^k/k!$, de Cauchy en $\|\cdot\|_\infty$ con límite e^x.

Idea clave

Los tres contraejemplos tienen la misma forma: un subespacio denso y propio. Ése es el patrón universal de la incompletitud, y también la receta para arreglarla: todo normado tiene una compleción, única salvo isometría, en la que él queda como subespacio denso. Así nacen L^1 a partir de las continuas y $\mathbb{R}$ a partir de $\mathbb{Q}$.

Cuidado

Cambiar la norma cambia la respuesta. El conjunto de (a) y el de (b) es *el mismo*, $C[0,1]$; solo cambia la manera de medir, y uno es de Banach y el otro no. Por eso en dimensión infinita nunca se dice "el espacio $C[0,1]$" sin decir con qué norma: la equivalencia de normas de dimensión finita ya no está disponible.

32.4. Espacios de Hilbert: la geometría regresa

Definición 32.23 (Producto interno)**.** Un *producto interno* sobre un espacio vectorial complejo H es $\langle\cdot,\cdot\rangle : H \times H \to \mathbb{C}$ lineal en la primera variable, con $\langle y, x\rangle = \overline{\langle x, y\rangle}$ y $\langle x, x\rangle > 0$ para $x \neq 0$. Induce la norma $\|x\| = \sqrt{\langle x, x\rangle}$. Si con ella H es completo, es un *espacio de Hilbert*.

Proposición 32.24 (Cauchy–Schwarz y ley del paralelogramo)**.** $|\langle x, y\rangle| \le \|x\|\,\|y\|$, *con igualdad si y solo si x, y son proporcionales. Además*

$$\|x + y\|^2 + \|x - y\|^2 = 2\|x\|^2 + 2\|y\|^2.$$

Recíprocamente (Jordan–von Neumann), una norma que cumple la ley del paralelogramo proviene de un producto interno.

Demostración. Para Cauchy–Schwarz, si $y \neq 0$ toma $\lambda = \langle x, y\rangle/\|y\|^2$ y desarrolla $0 \le \|x - \lambda y\|^2 = \|x\|^2 - |\langle x, y\rangle|^2/\|y\|^2$. La ley del paralelogramo sale de desarrollar ambos cuadrados y sumar: los términos cruzados se cancelan. □

Ejemplo 32.25 (Cuáles son de Hilbert y cuáles no)**.** ℓ^2 con $\langle x,y\rangle=\sum x_n\overline{y_n}$ y $L^2(\Omega)$ con $\langle f,g\rangle=\int f\bar g$ son de Hilbert. En cambio ℓ^p y L^p con $p\neq 2$ *no* lo son: en ℓ^p, tomando $x=(1,0,0,\dots)$ e $y=(0,1,0,\dots)$, la ley del paralelogramo exigiría $2\cdot 2^{2/p}=4$, o sea $p=2$. Y $C[0,1]$ con la norma del supremo tampoco: con $f=1$ y $g=x$ se obtiene $\|f+g\|_\infty^2+\|f-g\|_\infty^2=4+1=5\neq 2+2=4$.

Teorema 32.26 (Proyección sobre un convexo cerrado)**.** *Sea H de Hilbert, $K\subseteq H$ convexo, cerrado y no vacío, y $x\in H$. Entonces existe un único $p\in K$ con $\|x-p\|=\operatorname{dist}(x,K)$. Si además $K=M$ es un subespacio cerrado, p queda caracterizado por $x-p\perp M$, y la aplicación $P_M:x\mapsto p$ es lineal, idempotente y de norma 1 (si $M\neq\{0\}$).*

Demostración. Sea $d=\operatorname{dist}(x,K)$ y $(y_n)\subseteq K$ con $\|x-y_n\|\to d$. La ley del paralelogramo aplicada a $x-y_n$ y $x-y_m$ da

$$\|y_n-y_m\|^2=2\|x-y_n\|^2+2\|x-y_m\|^2-4\Big\|x-\frac{y_n+y_m}{2}\Big\|^2\le 2\|x-y_n\|^2+2\|x-y_m\|^2-4d^2,$$

donde se usó que $\frac{y_n+y_m}{2}\in K$ por convexidad. El lado derecho tiende a 0, luego (y_n) es de Cauchy; por completitud converge a un $p\in K$ (cerrado) con $\|x-p\|=d$. La unicidad sale de la misma desigualdad aplicada a dos minimizantes. Si $K=M$ es subespacio, para todo $m\in M$ y $t\in\mathbb{R}$ la función $t\mapsto\|x-p-tm\|^2=\|x-p\|^2-2t\,\mathrm{Re}\langle x-p,m\rangle+t^2\|m\|^2$ tiene mínimo en $t=0$, lo que fuerza $\langle x-p,m\rangle=0$. □

Ejemplo 32.27 (Proyectar sobre convexos)**. (a) Sobre un subespacio.** Si K es el subespacio generado por $(1,0,0)$ en $\mathbb{R}^3$, la proyección de $(2,3,4)$ es $(2,0,0)$ y el error $(0,3,4)$ es ortogonal a K. Ésa es la caracterización: el punto más cercano es aquel en que el residuo se vuelve perpendicular.

(b) Sobre un cerrado no convexo, falla. Tómese $K=\{x\in\mathbb{R}^2:\|x\|=1\}$, la circunferencia, que es cerrada y no convexa. La proyección del origen no existe de forma única: *todos* los puntos de K están a distancia 1. La convexidad es lo que garantiza la unicidad, vía la identidad del paralelogramo.

(c) Sobre un convexo con restricciones. Con $K=\{x\in\mathbb{R}^2:x_1\ge 0,\ x_2\ge 0\}$, la proyección de $(-3,5)$ es $(0,5)$: se anulan las coordenadas negativas. Este cálculo es el corazón de los algoritmos de gradiente proyectado.

(d) Mínimos cuadrados. Resolver $\min_\beta\|y-X\beta\|^2$ es proyectar y sobre el subespacio imagen de X. La condición de ortogonalidad $X^\top(y-X\hat\beta)=0$ son exactamente las ecuaciones normales. Toda la regresión lineal es este teorema aplicado en $\mathbb{R}^n$.

Idea clave

La imagen mental. Este es el teorema del "pie de la perpendicular" de la geometría de secundaria, elevado a dimensión infinita: desde un punto exterior a un plano, la distancia mínima se alcanza en un único punto, y el segmento que los une es perpendicular al plano. Todo lo que sigue —series de Fourier, mínimos cuadrados, esperanza condicional en probabilidad, aproximación óptima por polinomios— es este único teorema en distintos disfraces. Y su demostración solo usa dos cosas: convexidad y la ley del paralelogramo.

Por eso los espacios de Hilbert son tan cómodos y los de Banach tanto menos.

Corolario 32.28 (Descomposición ortogonal)**.** *Si $M \subseteq H$ es un subespacio cerrado, entonces $H = M \oplus M^{\perp}$, y $(M^{\perp})^{\perp} = M$. Un subespacio M es denso si y solo si $M^{\perp} = \{0\}$.*

32.5. Bases ortonormales, Bessel y Parseval

Definición 32.29. Una familia $(e_i)_{i\in I}$ en H es *ortonormal* si $\langle e_i, e_j\rangle = \delta_{ij}$, y es una *base ortonormal* (o sistema completo) si además su envolvente lineal es densa.

Teorema 32.30 (Bessel, Parseval y Riesz–Fischer)**.** *Sea $(e_n)_{n\ge 1}$ ortonormal en H y $\hat{x}_n = \langle x, e_n\rangle$. Entonces*

(a) ***(Bessel)*** *$\sum_n |\hat{x}_n|^2 \le \|x\|^2$, y $\sum_n \hat{x}_n e_n$ converge, con suma $P_M x$, donde $M = \overline{\operatorname{span}}(e_n)$.*

(b) ***(Parseval)*** *(e_n) es base ortonormal $\iff$ $\|x\|^2 = \sum_n |\hat{x}_n|^2$ para todo x $\iff$ $x = \sum_n \hat{x}_n e_n$ para todo x.*

(c) ***(Riesz–Fischer)*** *Para toda $(c_n) \in \ell^2$, la serie $\sum c_n e_n$ converge en H.*

En consecuencia, todo espacio de Hilbert separable de dimensión infinita es isométricamente isomorfo a ℓ^2.

Demostración. (a) Sea $S_N = \sum_{n\le N} \hat{x}_n e_n$. Un cálculo directo da $\|x - S_N\|^2 = \|x\|^2 - \sum_{n\le N} |\hat{x}_n|^2 \ge 0$, lo que prueba Bessel; las sumas parciales son de Cauchy porque $\|S_N - S_M\|^2 = \sum_{M<n\le N} |\hat{x}_n|^2$ es la cola de una serie convergente, y la completitud hace el resto. (b) Sale de la misma identidad $\|x - S_N\|^2 = \|x\|^2 - \sum_{n\le N} |\hat{x}_n|^2$: el resto tiende a 0 exactamente cuando vale Parseval, y eso ocurre para todo x exactamente cuando el sistema es completo. (c) Es de nuevo el argumento de Cauchy de (a). El isomorfismo con ℓ^2 es $x \mapsto (\hat{x}_n)$, que es lineal, biyectivo por (b) y (c), e isométrico por Parseval. □

Ejemplo 32.31 (Fourier es exactamente esto)**.** En $H = L^2[-\pi, \pi]$, el sistema $\left\{\frac{1}{\sqrt{2\pi}} e^{inx}\right\}_{n\in\mathbb{Z}}$ es ortonormal, y es completo por el teorema de Stone–Weierstrass del capítulo 29 (los polinomios trigonométricos son densos en las continuas periódicas, y estas lo son en L^2). Por tanto toda $f \in L^2$ es la suma de su serie de Fourier *en el sentido de L^2*, y Parseval dice

$$\frac{1}{2\pi}\int_{-\pi}^{\pi} |f|^2 = \sum_{n\in\mathbb{Z}} |c_n|^2.$$

Con $f(x) = x$ se obtiene $b_n = \frac{2(-1)^{n+1}}{n}$, y Parseval en forma real da

$$\int_{-\pi}^{\pi} x^2\,\mathrm{d}x = \frac{2\pi^3}{3} = \pi \sum_{n\ge 1} \frac{4}{n^2} \implies \sum_{n\ge 1} \frac{1}{n^2} = \frac{\pi^2}{6}.$$

El problema de Basilea cae como un subproducto de haber elegido bien la base.

Cuidado

La convergencia que da este teorema es en L^2, *no* puntual. Existen funciones continuas cuya serie de Fourier diverge en un punto (du Bois-Reymond, 1873), y ese fenómeno se demuestra con el principio de acotación uniforme, más abajo. La convergencia puntual en casi todo punto para $f \in L^2$ es cierta, pero es el teorema de Carleson (1966), uno de los más difíciles del análisis del siglo XX. No confundas los tres modos de convergencia.

Ejemplo 32.32 (Mínimos cuadrados como proyección)**.** Aproximar $f(x) = x^2$ en $L^2[0,1]$ por una recta $a+bx$ es proyectar f sobre $M = \operatorname{span}\{1, x\}$. Las ecuaciones normales $\langle f-a-bx, 1\rangle = \langle f-a-bx, x\rangle = 0$ dan

$$a + \tfrac{b}{2} = \tfrac{1}{3}, \qquad \tfrac{a}{2} + \tfrac{b}{3} = \tfrac{1}{4} \implies a = -\tfrac{1}{6},\ b = 1.$$

La mejor recta es $x - \frac{1}{6}$, con error cuadrático $\int_0^1 (x^2 - x + \frac{1}{6})^2 \mathrm{d}x = \frac{1}{180}$, es decir $\|f - P_M f\|_2 = \frac{1}{6\sqrt{5}} \approx 0{,}0745$. Compárese con la mejor aproximación en norma del supremo, que es otro problema (Chebyshev) y da otra recta: la geometría depende de la norma.

Regla: proyectar sobre un subespacio y usar Bessel y Parseval

En un espacio de Hilbert, "la mejor aproximación" tiene siempre la misma forma: se proyecta ortogonalmente. Y proyectar, cuando hay una base ortonormal, es simplemente quedarse con unos cuantos coeficientes.

1. **Ortonormaliza primero.** Si el subespacio M viene dado por generadores cualesquiera, aplícales Gram–Schmidt para obtener una base ortonormal $\{e_1, \dots, e_n\}$ de M. Todo lo que sigue supone ortonormalidad; sin ella las fórmulas son falsas.

2. **Proyecta coeficiente a coeficiente.** La proyección ortogonal de x sobre M es

$$P_M x = \sum_{k=1}^{n} \langle x, e_k\rangle\, e_k.$$

   Cada coeficiente se calcula por separado: no hay que resolver ningún sistema. Esa es toda la ventaja de haber ortonormalizado.

3. **Reconoce que esa es la mejor aproximación.** $P_M x$ es el *único* punto de M que minimiza $\|x - m\|$, y la caracterización es geométrica: $x - P_M x \perp M$. Si alguna vez dudas de un cálculo, comprueba esa ortogonalidad; es el control más rápido que existe.

4. **Calcula el error con Pitágoras.** Como $x = P_M x + (x - P_M x)$ con los dos trozos ortogonales,

$$\|x\|^2 = \|P_M x\|^2 + \|x - P_M x\|^2, \qquad \text{luego} \qquad \|x - P_M x\|^2 = \|x\|^2 - \sum_{k=1}^{n} |\langle x, e_k\rangle|^2.$$

   Nunca hace falta desarrollar el cuadrado a mano.

5. **De ahí sale Bessel, gratis.** Como el error es ≥ 0, la fórmula anterior da $\sum_k |\langle x, e_k\rangle|^2 \leq \|x\|^2$ para *cualquier* familia ortonormal, finita o no. Bessel no es un teorema aparte: es Pitágoras más la observación de que una longitud no es negativa.

6. **Parseval es Bessel con igualdad, y equivale a completitud.** La desigualdad se convierte en
$$\|x\|^2 = \sum_k |\langle x, e_k\rangle|^2$$
si y solo si la familia ortonormal es *completa*, es decir, si su envolvente cerrada es todo el espacio. Dicho al revés: el defecto en Bessel mide exactamente cuánto de x no ve la familia.

7. **Control.** El error debe bajar (nunca subir) al añadir vectores a la base, y debe ser cero si y solo si $x \in M$. Y todo coeficiente cumple $|\langle x, e_k\rangle| \leq \|x\|$.

Ejemplo 32.33 (La mejor recta que aproxima a x^2, y la serie $\sum 1/n^2$ vía Parseval)**. (a) Aproximación en $L^2(0,1)$.** Busquemos el polinomio de grado ≤ 1 que mejor aproxima a $f(x) = x^2$ en la norma de $L^2(0,1)$; es decir, la proyección de x^2 sobre $M = \operatorname{gen}\{1, x\}$. Ortonormalizando $\{1, x\}$ por Gram–Schmidt sale
$$e_1(x) = 1, \qquad e_2(x) = \sqrt{12}\,(x - \tfrac{1}{2}),$$
(los primeros polinomios de Legendre trasladados a $[0,1]$). Los coeficientes son
$$\langle f, e_1\rangle = \int_0^1 x^2\,\mathrm{d}x = \tfrac{1}{3}, \qquad \langle f, e_2\rangle = \sqrt{12}\int_0^1 x^2(x - \tfrac{1}{2})\mathrm{d}x = \sqrt{12}\Big(\tfrac{1}{4} - \tfrac{1}{6}\Big) = \frac{\sqrt{12}}{12} = \frac{1}{2\sqrt{3}}.$$
Reconstruyendo,
$$P_M f(x) = \tfrac{1}{3}\cdot 1 + \tfrac{1}{2\sqrt{3}}\cdot\sqrt{12}(x - \tfrac{1}{2}) = \tfrac{1}{3} + x - \tfrac{1}{2} = x - \tfrac{1}{6}.$$
La mejor recta es $y = x - \frac{1}{6}$. Por el paso 4 el error al cuadrado es
$$\|f\|^2 - \tfrac{1}{9} - \tfrac{1}{12} = \tfrac{1}{5} - \tfrac{1}{9} - \tfrac{1}{12} = \frac{36 - 20 - 15}{180} = \frac{1}{180},$$
luego $\|x^2 - (x - \frac{1}{6})\|_2 = 1/\sqrt{180} = 0{,}0745356\ldots$ Control del paso 3: $\int_0^1 (x^2 - x + \frac{1}{6})\mathrm{d}x = \frac{1}{3} - \frac{1}{2} + \frac{1}{6} = 0$ y $\int_0^1 (x^2 - x + \frac{1}{6})x\,\mathrm{d}x = \frac{1}{4} - \frac{1}{3} + \frac{1}{12} = 0$. El error es ortogonal a 1 y a x, como tenía que ser.

Pausa para pensar

Merece la pena notar que $y = x - \frac{1}{6}$ no pasa por ninguno de los puntos que uno elegiría a ojo: no interpola a x^2 en los extremos ni en el punto medio. La mejor aproximación en media cuadrática y la mejor interpolación son objetivos distintos, y dan respuestas distintas.

(b) Parseval, y una serie famosa como subproducto. En $L^2(-\pi,\pi)$ la familia $\{\frac{1}{\sqrt{2\pi}}\}\cup\{\frac{\cos nx}{\sqrt{\pi}},\frac{\sin nx}{\sqrt{\pi}}\}_{n\geq 1}$ es una base ortonormal completa. Tomemos $f(x)=x$. Por imparidad todos los coeficientes en coseno se anulan, y

$$\frac{1}{\sqrt{\pi}}\int_{-\pi}^{\pi} x\sin nx\,\mathrm{d}x=\frac{1}{\sqrt{\pi}}\cdot\frac{2\pi(-1)^{n+1}}{n}=\frac{2\sqrt{\pi}\,(-1)^{n+1}}{n}.$$

Parseval dice que la suma de los cuadrados de estos coeficientes es $\|f\|^2=\int_{-\pi}^{\pi}x^2\mathrm{d}x=\frac{2\pi^3}{3}$:

$$\sum_{n\geq 1}\frac{4\pi}{n^2}=\frac{2\pi^3}{3}\qquad\Longrightarrow\qquad\sum_{n\geq 1}\frac{1}{n^2}=\frac{\pi^2}{6}.$$

El problema de Basilea, resuelto como corolario de una identidad de longitudes. Que la igualdad valga — y no solo la desigualdad de Bessel — es precisamente la afirmación de que el sistema trigonométrico es completo en L^2.

Regla: ortonormalizar por Gram–Schmidt y desarrollar una función en la base

En un espacio de Hilbert, «tener una base» significa poder escribir cualquier vector como una serie de coeficientes calculados por productos internos. Fabricar esa base y usarla es un procedimiento mecánico.

1. **Fija el producto interno antes de nada.** En $L^2(a,b)$ es $\langle f,g\rangle=\int_a^b f\bar{g}\,dx$, pero puede llevar un peso $w(x)>0$. Cambiar el peso cambia por completo la base que sale; los polinomios de Legendre, Hermite y Chebyshev son el mismo algoritmo con tres pesos distintos.

2. **Aplica Gram–Schmidt a la familia de partida.** Con $v_1,v_2,\ldots$ linealmente independientes, define recursivamente

$$u_k=v_k-\sum_{j<k}\frac{\langle v_k,u_j\rangle}{\langle u_j,u_j\rangle}\,u_j.$$

   Cada u_k es ortogonal a todos los anteriores por construcción; compruébalo con un producto interno antes de seguir.

3. **Normaliza al final, no en cada paso.** Es más limpio arrastrar los u_k sin normalizar y dividir por $\|u_k\|$ solo cuando escribas la base. Así los coeficientes salen en fracciones y no en raíces.

4. **Calcula los coeficientes de Fourier generalizados.** Si $\{e_k\}$ es ortonormal, $\hat{f}(k)=\langle f,e_k\rangle$. Si prefieres trabajar con los u_k sin normalizar, el coeficiente es $c_k=\langle f,u_k\rangle/\langle u_k,u_k\rangle$ y el desarrollo es $f\approx\sum c_k u_k$.

5. **Usa la paridad y las simetrías para anular coeficientes.** Si el intervalo es simétrico y f es par, todos los coeficientes de índice impar se anulan sin calcular nada. Esta es la mitad del trabajo ahorrada en la mayoría de los ejercicios.

6. **Controla el error con Bessel y Parseval.** La suma parcial $S_N = \sum_{k\le N} c_k u_k$ es la *mejor* aproximación en L^2 dentro del subespacio generado, y el error cumple $\|f - S_N\|^2 = \|f\|^2 - \sum_{k\le N} |c_k|^2 \|u_k\|^2$. Si esa resta te sale negativa, has calculado mal un coeficiente.

7. **Verifica la completitud antes de escribir una igualdad.** Bessel ($\le$) vale siempre; Parseval ($=$) solo si la base es completa. Si el sistema no es completo, la serie converge, pero no a f.

Ejemplo 32.34 (De Gram–Schmidt a Legendre, y de Parseval a $\zeta(2)$)**.** Recorremos el procedimiento completo dos veces: una con polinomios y otra con senos.

(a) Gram–Schmidt sobre $1, x, x^2, x^3$ en $L^2(-1,1)$. Empezamos con $u_0 = 1$, cuyo cuadrado de norma es $\int_{-1}^{1} 1\,dx = 2$. Luego

$$u_1 = x - \frac{\langle x, 1\rangle}{2}\cdot 1 = x,$$

porque $\int_{-1}^{1} x\,dx = 0$ por paridad (paso 5). Para u_2: $\langle x^2, 1\rangle = \frac{2}{3}$ y $\langle x^2, x\rangle = 0$, así que

$$u_2 = x^2 - \frac{2/3}{2} = x^2 - \frac{1}{3}.$$

Para u_3: $\langle x^3, 1\rangle = 0$, $\langle x^3, x\rangle = \frac{2}{5}$, $\langle x^3, x^2 - \frac{1}{3}\rangle = 0$, luego

$$u_3 = x^3 - \frac{2/5}{2/3}\,x = x^3 - \frac{3}{5}x.$$

Los cuadrados de norma son $\|u_0\|^2 = 2$, $\|u_1\|^2 = \frac{2}{3}$, $\|u_2\|^2 = \frac{8}{45}$, $\|u_3\|^2 = \frac{8}{175}$. Estos son exactamente los polinomios de Legendre salvo normalización: $P_2(x) = \frac{3}{2}x^2 - \frac{1}{2} = \frac{3}{2}u_2$ y $P_3(x) = \frac{5}{2}x^3 - \frac{3}{2}x = \frac{5}{2}u_3$, con la convención $P_k(1) = 1$.

(b) Desarrollar $f(x) = |x|$ en esa base. Como f es par, $c_1 = c_3 = 0$ sin calcular nada. Los otros dos:

$$c_0 = \frac{\int_{-1}^{1} |x|\,dx}{2} = \frac{1}{2}, \qquad c_2 = \frac{\int_{-1}^{1} |x|(x^2 - \frac{1}{3})\,dx}{8/45} = \frac{1/6}{8/45} = \frac{15}{16}.$$

Por tanto la mejor aproximación cuadrática es

$$S_2(x) = \frac{1}{2} + \frac{15}{16}\Big(x^2 - \frac{1}{3}\Big) = \frac{3}{16} + \frac{15}{16}x^2.$$

En $x = 0$ vale $3/16 = 0{,}1875$ en lugar de 0, y en $x = \pm 1$ vale $9/8 = 1{,}125$ en lugar de 1: la parábola no interpola, y no tiene por qué hacerlo. Lo que hace es minimizar el error *en media cuadrática*.

(c) El error, medido con el paso 6. Tenemos $\|f\|^2 = \int_{-1}^{1} x^2 dx = \frac{2}{3}$ y

$$\sum_{k\le 2} c_k^2 \|u_k\|^2 = \frac{1}{4}\cdot 2 + \Big(\frac{15}{16}\Big)^2 \cdot \frac{8}{45} = \frac{1}{2} + \frac{5}{32} = \frac{21}{32}.$$

Luego $\|f - S_2\|^2 = \frac{2}{3} - \frac{21}{32} = \frac{1}{96}$, es decir $\|f - S_2\| = 1/\sqrt{96} \approx 0{,}1021$. La resta es positiva, como exige Bessel: si te hubiera salido negativa, habría un error en c_2.

Cuidado

S_2 no es el polinomio de Taylor de grado 2 (que no existe, porque $|x|$ no es derivable en 0) ni el interpolador en tres puntos. Es la *proyección ortogonal* sobre el subespacio de polinomios de grado ≤ 2. Cada noción de «mejor aproximación» tiene su propio ganador, y solo la de mínimos cuadrados se calcula con productos internos.

(d) La base trigonométrica y una identidad clásica. En $L^2(-\pi,\pi)$ el sistema $\{\frac{1}{\sqrt{2\pi}}\}\cup\{\frac{1}{\sqrt{\pi}}\cos kx\}\cup\{\frac{1}{\sqrt{\pi}}\sin kx\}$ es ortonormal y completo. Para $f(x)=x$, que es impar, solo sobreviven los senos:

$$b_k=\frac{1}{\pi}\int_{-\pi}^{\pi} x\sin(kx)\,dx=\frac{2(-1)^{k+1}}{k},\qquad\text{luego}\qquad x=2\sum_{k=1}^{\infty}\frac{(-1)^{k+1}}{k}\sin(kx).$$

(e) Parseval y $\sum 1/k^2$. Aplicando la identidad de Parseval a ese desarrollo:

$$\frac{1}{\pi}\int_{-\pi}^{\pi}x^2\,dx=\sum_{k=1}^{\infty}b_k^2\quad\Longrightarrow\quad\frac{1}{\pi}\cdot\frac{2\pi^3}{3}=\sum_{k=1}^{\infty}\frac{4}{k^2},$$

o sea $\frac{2\pi^2}{3}=4\sum k^{-2}$, y por tanto

$$\sum_{k=1}^{\infty}\frac{1}{k^2}=\frac{\pi^2}{6}\approx 1{,}644934.$$

El problema de Basilea resuelto como un corolario de que $\|f\|^2$ se puede calcular de dos maneras. Repitiendo con $f(x)=x^2$ se obtiene $\sum k^{-4}=\pi^4/90$.

Pausa para pensar

Fíjate en lo que acaba de pasar en (e): una igualdad entre un número transcendente y una serie de racionales salió de que el teorema de Pitágoras vale en dimensión infinita. No hubo ningún truco analítico; hubo geometría. Esa es toda la razón de ser de los espacios de Hilbert.

Idea clave

Legendre, Hermite, Laguerre y Chebyshev son el *mismo* Gram–Schmidt aplicado a $1,x,x^2,\dots$ con cuatro productos internos distintos: peso 1 en $(-1,1)$, peso e^{-x^2} en $\mathbb{R}$, peso e^{-x} en $(0,\infty)$, y peso $(1-x^2)^{-1/2}$ en $(-1,1)$. Cada uno resuelve una ecuación diferencial de Sturm–Liouville distinta (véase el capítulo 12), y esa es la razón profunda de que aparezcan una y otra vez en física.

32.6. Operadores acotados

Teorema 32.35 (Continuidad equivale a acotación)**.** *Sea $T:X\to Y$ lineal entre espacios normados. Son equivalentes: (i) T es continua en algún punto; (ii) T es continua; (iii) existe $C\ge 0$ con $\|Tx\|\le C\|x\|$ para todo x. El ínfimo de tales C es*

$$\|T\|=\sup_{\|x\|\le 1}\|Tx\|.$$

Demostración. (i)⇒(iii): si T es continua en x_0, existe δ con $\|Tz - Tx_0\| \le 1$ para $\|z - x_0\| \le \delta$; por linealidad, $\|Tu\| \le 1$ para $\|u\| \le \delta$, luego $\|Tx\| \le \delta^{-1}\|x\|$. (iii)⇒(ii): T es lipschitziana. (ii)⇒(i) es trivial. □

Observación 32.36. Esto es una particularidad de la linealidad y no tiene análogo para funciones generales: para un operador lineal, la continuidad es un fenómeno global determinado por un solo número. El conjunto $B(X,Y)$ de operadores acotados, con esta norma, es un espacio normado; y si Y es Banach, $B(X,Y)$ también lo es.

Ejemplo 32.37 (Normas de operadores calculadas)**.**

- Matriz $A = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}$: en ℓ^1 la norma es el máximo de las sumas por columnas, 6; en ℓ^∞, el máximo de las sumas por filas, 7; en ℓ^2 es el mayor valor singular, $\sqrt{15+\sqrt{221}} \approx 5{,}46499$.
- Operador de multiplicación $(M_g f)(x) = g(x)f(x)$ en L^2: $\|M_g\| = \|g\|_\infty$.
- Operador de Volterra $(Vf)(x) = \int_0^x f(t)\,\mathrm{d}t$ en $L^2[0,1]$: sus valores singulares son $\frac{2}{(2k-1)\pi}$, luego $\|V\| = \frac{2}{\pi} \approx 0{,}63662$. Es el ejemplo canónico de operador acotado *sin ningún autovalor*: $Vf = \lambda f$ con $\lambda \neq 0$ obliga, derivando, a $f' = f/\lambda$ con $f(0) = 0$, cuya única solución es $f = 0$.
- La derivada $D = \mathrm{d}/\mathrm{d}x$ en $C^1[0,1] \subseteq C[0,1]$ con la norma del supremo *no* es acotada: $\|\sin(nx)\|_\infty = 1$ pero $\|D\sin(nx)\|_\infty = n$.

Cuidado

Que la derivada no sea acotada no es un accidente técnico: es la razón por la que los operadores de la mecánica cuántica (posición, momento, energía) se definen solo en un subespacio denso y se estudian como *operadores no acotados con dominio*. Todo el aparato de operadores autoadjuntos no acotados existe porque D no cabe en $B(H)$.

Regla: calcular la norma de un operador acotado

Por definición $\|T\| = \sup_{\|x\|\le 1} \|Tx\|$, pero ese supremo casi nunca se ataca de frente. La técnica es siempre la misma: una desigualdad y un testigo.

1. **Cota superior: acota $\|Tx\|$ por una constante por $\|x\|$.** Manipula la expresión de Tx con las desigualdades disponibles (triangular, Hölder, Cauchy–Schwarz, $|\sup| \le \sup|\ |$) hasta obtener $\|Tx\| \le C\|x\|$. Eso demuestra a la vez que T es acotado y que $\|T\| \le C$.
2. **Cota inferior: exhibe un testigo.** Encuentra un vector concreto x_0 con $\|x_0\| = 1$ y $\|Tx_0\| = C$. Entonces $\|T\| \ge C$, y junto con el paso 1 queda $\|T\| = C$. *Sin testigo no hay igualdad*, solo una cota.
3. **Si el supremo no se alcanza, usa una sucesión de casi-testigos.** A veces no hay ningún x_0 con $\|Tx_0\| = C$, pero sí una sucesión x_k de norma 1 con $\|Tx_k\| \to C$.

Basta igual: el supremo es C aunque no se alcance. Esta situación es la norma, no la excepción, en dimensión infinita.

4. **Acotado equivale a continuo.** Para aplicaciones lineales las dos palabras significan lo mismo, y la demostración es corta: si T es continuo en 0 hay δ con $\|Tx\| \le 1$ para $\|x\| \le \delta$, y por linealidad $\|T\| \le 1/\delta$. Por eso "acotado" es sinónimo de "bueno" en este contexto, y no de "de imagen acotada".

5. **Aprende de memoria los cuatro modelos.** El operador de multiplicación por g tiene norma $\|g\|_\infty$. El operador diagonal en ℓ^2 tiene norma $\sup_n |\lambda_n|$. El desplazamiento (a izquierda o derecha) tiene norma 1. Un funcional dado por $\langle \cdot, y\rangle$ tiene norma $\|y\|$ — eso es el teorema de representación de Riesz.

6. **Control.** $\|T\| = 0$ solo si $T = 0$; $\|ST\| \le \|S\|\|T\|$; y si T es una proyección ortogonal no nula, $\|T\| = 1$ exactamente. Si tu cuenta contradice alguna de estas, revísala.

Ejemplo 32.38 (Cuatro operadores, cuatro normas)**. (a) Un funcional en $L^2(0,1)$.** Sea $\varphi(f) = \int_0^1 x\, f(x)\,\mathrm{d}x$. Por Cauchy–Schwarz,

$$|\varphi(f)| \le \Big(\int_0^1 x^2 \mathrm{d}x\Big)^{1/2} \|f\|_2 = \frac{1}{\sqrt{3}}\|f\|_2,$$

luego $\|\varphi\| \le 1/\sqrt{3}$. El testigo es el vector que hace de Cauchy–Schwarz una igualdad, es decir el proporcional a x: tomando $f_0(x) = \sqrt{3}\,x$, que tiene $\|f_0\|_2 = 1$, sale $\varphi(f_0) = \sqrt{3}\int_0^1 x^2\mathrm{d}x = 1/\sqrt{3}$. Por tanto

$$\|\varphi\| = \frac{1}{\sqrt{3}} = 0{,}5773502692\ldots$$

Y obsérvese que esto es justo Riesz: $\varphi = \langle \cdot, y\rangle$ con $y(x) = x$, y en efecto $\|y\|_2 = 1/\sqrt{3}$.

(b) Multiplicación en $C[0,1]$. Sea $(Tf)(x) = g(x)f(x)$ con g continua fija. Entonces $\|Tf\|_\infty \le \|g\|_\infty \|f\|_\infty$, luego $\|T\| \le \|g\|_\infty$; y el testigo es la función constante $f \equiv 1$, de norma 1, que da $\|Tf\|_\infty = \|g\|_\infty$. Luego $\|T\| = \|g\|_\infty$ exactamente. Nótese que el supremo *sí* se alcanza aquí.

(c) Operador diagonal en ℓ^2. Sea $T(x_1, x_2, \ldots) = (\lambda_1 x_1, \lambda_2 x_2, \ldots)$ con (λ_n) acotada. Entonces $\|Tx\|_2^2 = \sum |\lambda_n|^2 |x_n|^2 \le (\sup|\lambda_n|)^2 \|x\|_2^2$, y los testigos son los vectores de la base, $\|Te_n\| = |\lambda_n|$. Luego $\|T\| = \sup_n |\lambda_n|$. Aquí el supremo se alcanza si y solo si el supremo de la sucesión es un máximo: con $\lambda_n = 1 - 1/n$ se tiene $\|T\| = 1$ pero *ningún* vector de norma 1 cumple $\|Tx\| = 1$. Es el paso 3 en estado puro.

(d) El desplazamiento a la derecha. Sea $S(x_1, x_2, \ldots) = (0, x_1, x_2, \ldots)$ en ℓ^2. Es una isometría: $\|Sx\| = \|x\|$ para todo x, luego $\|S\| = 1$ y todo vector es testigo. Pero no es sobreyectivo (nada se aplica en e_1), lo cual en dimensión finita sería imposible para una isometría. Este ejemplo, de una simplicidad engañosa, es el que rompe la intuición de dimensión finita y el que hay que recordar cuando se afirme algo demasiado optimista sobre operadores.

Regla: calcular el adjunto de un operador y clasificarlo

El adjunto es la generalización de la traspuesta conjugada, y clasificar un operador por su relación con su adjunto es lo que decide qué teorema espectral se le puede aplicar.

1. **Parte siempre de la definición, no de una fórmula.** El adjunto T^* es el único operador que cumple $\langle Tx, y\rangle = \langle x, T^*y\rangle$ para todos x, y. Escribe el miembro izquierdo, manipúlalo hasta que x quede solo en el primer hueco, y lo que quede en el segundo *es* T^*y.

2. **En integrales, la manipulación es cambiar el orden.** Si $Tf(x) = \int K(x,y)f(y)\,dy$, entonces al aplicar Fubini sale $T^*g(y) = \int \overline{K(x,y)}\,g(x)\,dx$; es decir, el núcleo del adjunto es $\overline{K(y,x)}$: se traspone y se conjuga, igual que en matrices.

3. **Comprueba las propiedades básicas.** $(T^*)^* = T$, $(ST)^* = T^*S^*$ (se invierte el orden), $\|T^*\| = \|T\|$ y $\|T^*T\| = \|T\|^2$. Esta última, la identidad C*, es la que hace que toda la teoría funcione.

4. **Clasifica.** T es *autoadjunto* si $T = T^*$; *unitario* si $T^*T = TT^* = I$; *normal* si $T^*T = TT^*$; *positivo* si $\langle Tx, x\rangle \geq 0$ para todo x; *isometría* si $T^*T = I$ (sin exigir $TT^* = I$).

5. **No confundas isometría con unitario en dimensión infinita.** En dimensión finita $T^*T = I$ obliga a $TT^* = I$; en dimensión infinita no. Esa diferencia es la fuente de los ejemplos más instructivos del capítulo.

6. **Traduce la clasificación al espectro.** Autoadjunto ⇒ espectro real. Unitario ⇒ espectro en la circunferencia unidad. Positivo ⇒ espectro en $[0, \infty)$. Normal ⇒ hay teorema espectral. Si el operador no es normal, ninguna de estas garantías vale.

7. **Verifica con los cuatro subespacios.** Siempre se cumple $\ker T^* = (\operatorname{im} T)^\perp$ y $\overline{\operatorname{im} T^*} = (\ker T)^\perp$. Si tu candidato a T^* no respeta esas dos igualdades, está mal.

Ejemplo 32.39 (Cinco operadores, sus adjuntos, y qué se rompe en cada caso)**.** Cada ejemplo ilustra exactamente una patología que en dimensión finita no existe.

(a) El desplazamiento a la derecha: una isometría que no es unitaria. En ℓ^2, sea $S(x_1, x_2, x_3, \ldots) = (0, x_1, x_2, \ldots)$. Aplicando el paso 1:

$$\langle Sx, y\rangle = \sum_{k\geq 2} x_{k-1}\bar{y}_k = \sum_{j\geq 1} x_j \bar{y}_{j+1} = \langle x, S^*y\rangle \quad \text{con} \quad S^*(y_1, y_2, \ldots) = (y_2, y_3, \ldots),$$

el desplazamiento a la izquierda. Entonces

$$S^*S = I, \qquad SS^*(y_1, y_2, \ldots) = (0, y_2, y_3, \ldots) \neq I.$$

S conserva normas ($\|Sx\| = \|x\|$) pero no es sobreyectivo: su imagen es el subespacio $\{x_1 = 0\}$. Es una isometría, no un operador unitario, y por tanto *no es normal*.

Cuidado

En dimensión finita, «inyectivo» e «invertible» coinciden para endomorfismos. S es la demostración de que en dimensión infinita no: es inyectivo, isométrico, y sin embargo no invertible. Su inversa por la izquierda existe (S^*) pero no es inversa por la derecha.

(b) Los espectros de S y S^* no se parecen en nada. S no tiene ningún autovalor: si $Sx = \lambda x$ con $\lambda \neq 0$, comparando componentes sale $x = 0$. En cambio S^* tiene *todo el disco abierto* como autovalores: para $|\lambda| < 1$, el vector $v_\lambda = (1, \lambda, \lambda^2, \dots)$ está en ℓ^2 (su norma al cuadrado es $1/(1-|\lambda|^2)$) y cumple $S^* v_\lambda = \lambda v_\lambda$. Ambos tienen el mismo espectro, el disco cerrado $\{|\lambda| \leq 1\}$, pero uno lo alcanza con autovectores y el otro no. Esto no puede pasar si el operador es normal.

(c) Multiplicación: autoadjunto, con espectro y sin autovalores. En $L^2(0,1)$ define $(Mf)(x) = x\,f(x)$. Como el multiplicador es real, $\langle Mf, g\rangle = \int x f \bar{g} = \langle f, Mg\rangle$: es autoadjunto. Su norma es $\|M\| = \sup_{(0,1)} |x| = 1$, y su espectro es $\sigma(M) = [0,1]$. Y sin embargo no tiene ningún autovector: $xf(x) = \lambda f(x)$ obliga a f a anularse fuera del punto $\{\lambda\}$, que tiene medida cero, luego $f = 0$ en L^2. El espectro continuo es un fenómeno genuinamente infinito-dimensional. En general $M_\varphi^* = M_{\bar\varphi}$, con M_φ autoadjunto si φ es real y unitario si $|\varphi| = 1$ en casi todo punto.

(d) El operador de Volterra: adjunto explícito y suma de rango uno. En $L^2(0,1)$, sea $Vf(x) = \int_0^x f(t)\,dt$, es decir el operador integral de núcleo $K(x,y) = \mathbf{1}_{\{y<x\}}$. Por el paso 2 el núcleo del adjunto es $\mathbf{1}_{\{x<y\}}$, o sea

$$V^* g(y) = \int_y^1 g(x)\,dx.$$

Lo notable es la suma:

$$(V + V^*)f(x) = \int_0^x f + \int_x^1 f = \int_0^1 f = \langle f, \mathbf{1}\rangle \cdot \mathbf{1}(x),$$

un operador de rango uno. Es decir, V es «autoadjunto salvo rango uno». Su norma es $\|V\| = 2/\pi \approx 0{,}6366$ (los valores singulares son $\frac{2}{(2n-1)\pi}$, es decir 0,6366, 0,2122, 0,1273, …), su espectro es $\sigma(V) = \{0\}$ y no tiene autovalores: de $Vf = \lambda f$ con $\lambda \neq 0$ se sigue $f = \lambda f'$ y $f(0) = 0$, luego $f \equiv 0$. Un operador no nulo cuyo espectro es un solo punto se llama *cuasinilpotente*, y $\|V^n\|^{1/n} \to 0$.

(e) Un operador sin adjunto que le sea conmutable, en dimensión dos. No hace falta ir al infinito para ver que la normalidad es una condición fuerte. Sea $A = \left(\begin{smallmatrix} 0 & 1 \\ 0 & 0 \end{smallmatrix}\right)$ en $\mathbb{C}^2$. Entonces

$$A^*A = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}, \qquad AA^* = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix},$$

que no conmutan: A no es normal. Su espectro es $\{0\}$ (doble) mientras que $\|A\| = 1$: el radio espectral es estrictamente menor que la norma, cosa imposible en un operador normal, donde siempre $r(A) = \|A\|$. En cambio la rotación $R = \left(\begin{smallmatrix} 0 & -1 \\ 1 & 0 \end{smallmatrix}\right)$ cumple $R^*R = RR^* = I$: es unitaria, sus autovalores son $\pm i$ (en la circunferencia unidad, como anuncia el paso 6) y sí es diagonalizable en una base ortonormal.

Pausa para pensar

Compara (d) y (e): en ambos el espectro es $\{0\}$ y el operador no es cero. En dimensión finita eso significa nilpotente, y en efecto $A^2 = 0$. En dimensión infinita $V^n \neq 0$ para todo n, pero $\|V^n\|^{1/n} \to 0$: el operador no llega nunca a anularse, aunque sus potencias decaigan más rápido que cualquier geométrica. ¿Cuánto vale exactamente V^n? El teorema fundamental del cálculo iterado da $V^n f(x) = \frac{1}{(n-1)!}\int_0^x (x-t)^{n-1} f(t)\,dt$, y el factorial del denominador explica el decaimiento.

Idea clave

La identidad $\|T^*T\| = \|T\|^2$ del paso 3 parece una curiosidad y es en realidad el axioma que define un C^*-álgebra. A partir de ella sola se demuestra que el espectro de un elemento autoadjunto es real, que el radio espectral de un normal coincide con su norma, y que toda C^*-álgebra conmutativa es un álgebra de funciones continuas (teorema de Gelfand–Naimark). Es decir: toda la teoría espectral está codificada en una igualdad entre dos números.

32.7. Los cuatro teoremas grandes

Teorema 32.40 (Hahn–Banach)**.** *Sea X un espacio vectorial real, $p : X \to \mathbb{R}$ sublineal ($p(x+y) \le p(x)+p(y)$, $p(tx) = tp(x)$ para $t \ge 0$), $M \subseteq X$ subespacio y $f : M \to \mathbb{R}$ lineal con $f \le p$ en M. Entonces existe una extensión lineal $F : X \to \mathbb{R}$ con $F|_M = f$ y $F \le p$ en todo X. En particular, si X es normado y $x_0 \neq 0$, existe $F \in X^*$ con $\|F\| = 1$ y $F(x_0) = \|x_0\|$: el dual* separa puntos.

Esquema. El paso clave es extender de M a $M \oplus \mathbb{R}z$ para un solo $z \notin M$. Hay que elegir $\alpha = F(z)$ compatible: la condición $f(m) + t\alpha \le p(m+tz)$ para todo m, t se reduce, dividiendo por $|t|$, a

$$\sup_{m \in M} \big(f(m) - p(m-z)\big) \;\le\; \alpha \;\le\; \inf_{m' \in M} \big(p(m'+z) - f(m')\big),$$

y ese intervalo es no vacío porque $f(m) + f(m') \le p(m+m') \le p(m-z) + p(m'+z)$. Un argumento de Zorn sobre la familia de extensiones parciales, ordenada por prolongación, da la extensión a todo X. Para el corolario, aplica lo anterior a $M = \mathbb{R}x_0$, $f(tx_0) = t\|x_0\|$ y $p = \|\cdot\|$. □

Idea clave

Para qué sirve. Hahn–Banach es el teorema que garantiza que hay *suficientes* funcionales lineales continuos. Sin él, el dual X^* podría ser trivial y toda la teoría de dualidad se caería. De él salen: la separación de convexos por hiperplanos (y con ella la dualidad de la optimización convexa y el teorema minimax), la existencia de límites de Banach, y el hecho de que $\|x\| = \sup_{\|F\| \le 1} |F(x)|$, que permite razonar sobre vectores midiéndolos con funcionales. A diferencia de los otros tres grandes, no necesita completitud: necesita el axioma de elección.

Teorema 32.41 (Banach–Steinhaus, o principio de acotación uniforme)**.** *Sea X de Banach, Y normado y $\mathcal{F} \subseteq B(X,Y)$ una familia tal que $\sup_{T\in\mathcal{F}} \|Tx\| < \infty$ para cada $x \in X$ por separado. Entonces $\sup_{T\in\mathcal{F}} \|T\| < \infty$.*

Demostración. Sea $E_n = \{x \in X : \ \|Tx\| \le n \text{ para toda } T \in \mathcal{F}\}$. Cada E_n es cerrado (intersección de preimágenes de cerrados por funciones continuas) y $\bigcup_n E_n = X$ por hipótesis. Como X es completo, el teorema de Baire (capítulo 29) impide que todos los E_n tengan interior vacío: algún E_N contiene una bola $B(x_0, r)$. Entonces, para $\|u\| \le r$ y toda $T \in \mathcal{F}$,

$$\|Tu\| \le \|T(x_0 + u)\| + \|Tx_0\| \le 2N,$$

de donde $\|T\| \le 2N/r$ para toda T. □

Ejemplo 32.42 (Existe una función continua cuya serie de Fourier diverge en 0)**.** Sea $S_n f(0) = \frac{1}{2\pi}\int_{-\pi}^{\pi} f(t) D_n(t)\,\mathrm{d}t$ con el núcleo de Dirichlet D_n. Cada $S_n(\cdot)(0)$ es un funcional continuo sobre $C_{2\pi}$ con norma exactamente $\frac{1}{2\pi}\|D_n\|_1$, las *constantes de Lebesgue*, que crecen como $\frac{4}{\pi^2}\ln n \to \infty$. Si la serie de Fourier de toda función continua convergiera en 0, la familia $\{S_n(\cdot)(0)\}$ estaría puntualmente acotada, y Banach–Steinhaus forzaría $\sup_n \|S_n\| < \infty$, contra el crecimiento logarítmico. Luego existe f continua con $\sup_n |S_n f(0)| = \infty$.

Idea clave

Cómo leer esa demostración. No construye ninguna función: demuestra que *tiene* que existir, porque lo contrario contradice un principio general. Es el mismo estilo de argumento que el conteo de Cantor para los trascendentes, y es característico del análisis funcional: los teoremas grandes son máquinas de existencia. Construir explícitamente una función continua con serie de Fourier divergente es posible pero laborioso; el argumento de Baire lo hace en cinco líneas.

Teorema 32.43 (Aplicación abierta y aplicación inversa)**.** *Sean X, Y de Banach y $T \in B(X,Y)$ sobreyectiva. Entonces T es abierta: manda abiertos en abiertos. En consecuencia, si además T es inyectiva, T^{-1} es continua.*

Esquema. Como T es sobreyectiva, $Y = \bigcup_n \overline{T(B(0,n))}$, y Baire da que algún $\overline{T(B(0,n))}$ tiene interior; por linealidad y simetría, $\overline{T(B_X)} \supseteq B(0,c)$ para algún $c > 0$. El paso técnico es quitar la clausura: dado y con $\|y\| < c$, se elige x_1 con $\|y - Tx_1\| < c/2$, luego x_2 con $\|y - Tx_1 - Tx_2\| < c/4$, etc.; la serie $\sum x_k$ converge absolutamente y, por completitud de X, a un x con $Tx = y$ y $\|x\| < 2$. Eso da $T(B_X) \supseteq B(0, c/2)$, que es la apertura. □

Ejemplo 32.44 (Cuándo la inversa es continua)**. (a) El caso bueno.** Si $T\colon X \to Y$ es lineal, continua y biyectiva entre espacios de Banach, su inversa es automáticamente continua. Esto *no* es evidente: la continuidad de T^{-1} no se sigue del álgebra, hace falta la completitud de ambos espacios y el teorema de Baire.

(b) Sin completitud, falla. Sea X el espacio de sucesiones con soporte finito con la norma del supremo (no completo) y $T(x_1, x_2, \dots) = (x_1, x_2/2, x_3/3, \dots)$. Es lineal, continua

y biyectiva sobre su imagen, y su inversa multiplica la coordenada n por n: no es acotada. La hipótesis de Banach no es decorativa.

(c) Consecuencia sobre normas. Si un mismo espacio vectorial es completo con dos normas y una domina a la otra, entonces las dos son equivalentes. Aplicado a $C^1[0,1]$ con $\|f\|_\infty$ y $\|f\|_\infty + \|f'\|_\infty$: la segunda domina, pero con la primera el espacio no es completo, y en efecto las normas no son equivalentes.

(d) Por qué se llama "abierta". La conclusión literal es que T manda abiertos en abiertos, es decir, que $T(B(0,1))$ contiene una bola alrededor del origen. Eso significa que se puede resolver $Tx = y$ con control sobre el tamaño de x: existe C tal que para todo y hay x con $Tx = y$ y $\|x\| \le C\|y\|$. Es una estimación a priori, y por eso el teorema es tan útil en ecuaciones en derivadas parciales.

Teorema 32.45 (Gráfica cerrada)**.** *Sean X, Y de Banach y $T : X \to Y$ lineal. Si su gráfica $\{(x, Tx)\}$ es cerrada en $X \times Y$, entonces T es continua.*

Demostración. La gráfica G es un subespacio cerrado de $X \times Y$, luego es Banach. La proyección $\pi_1 : G \to X$ es lineal, continua y biyectiva, así que por el teorema de la aplicación inversa π_1^{-1} es continua; y $T = \pi_2 \circ \pi_1^{-1}$ es composición de continuas. □

Idea clave

Qué dice esto en cristiano. Para probar que un operador es continuo hay que comprobar: si $x_n \to x$ entonces $Tx_n \to Tx$. Eso son dos cosas a la vez: que Tx_n converge, y que converge al sitio correcto. El teorema de la gráfica cerrada dice que basta con la segunda: si se *supone* que Tx_n converge a algo, y se demuestra que ese algo es Tx, la continuidad está probada. Es la mitad del trabajo, y en la práctica la mitad fácil.

Ejemplo 32.46 (Usar la gráfica cerrada)**. (a) Un operador que resulta continuo.** Sea $T\colon L^2 \to L^2$ definido por $Tf = gf$ con g medible tal que $gf \in L^2$ para toda $f \in L^2$. Si $f_n \to f$ en L^2 y $gf_n \to h$, pasando a subsucesiones que convergen en casi todo punto se ve que $h = gf$. Gráfica cerrada, luego T es continuo, y de ahí se deduce $g \in L^\infty$.

(b) Un operador con gráfica no cerrada. La derivación $D\colon C^1[0,1] \to C[0,1]$ vista como operador definido en un subespacio denso de $C[0,1]$: tomando $f_n(x) = \sin(n^2 x)/n$, se tiene $f_n \to 0$ uniformemente mientras que $f_n'(x) = n\cos(n^2 x)$ no converge a nada. La gráfica no es cerrada y el operador no es acotado.

(c) La derivación sí es cerrada en otro sentido. Como operador *no acotado* con dominio C^1, la derivada es un operador cerrado: si $f_n \to f$ y $f_n' \to g$ uniformemente, entonces f es derivable y $f' = g$. Cerrado no implica acotado cuando el dominio no es todo el espacio, y esa distinción es la base de la teoría de operadores no acotados de la mecánica cuántica.

(d) La relación con la aplicación abierta. Los dos teoremas son equivalentes entre sí y ambos se deducen del de Baire, junto con el principio de acotación uniforme. Se les llama "los tres pilares" del análisis funcional, y los tres dicen variantes de lo mismo: en espacios completos, la información algebraica fuerza información topológica.

Idea clave

Por qué la gráfica cerrada es útil. Para probar continuidad hay que ver que $x_n \to x$ implica $Tx_n \to Tx$: dos cosas a la vez, que (Tx_n) converja y que converja al sitio correcto. La gráfica cerrada regala la primera: basta demostrar que *si* $x_n \to x$ *y* $Tx_n \to y$, entonces $y = Tx$. Es estrictamente más fácil, y es la razón de que este teorema aparezca constantemente en ecuaciones en derivadas parciales.

Cuidado

Los tres últimos teoremas —Steinhaus, aplicación abierta, gráfica cerrada— descansan sobre Baire, y por tanto sobre la *completitud*. En espacios normados incompletos son falsos. Ejemplo: en el espacio de sucesiones con finitos términos no nulos, con la norma del supremo (que no es completo), el funcional $x \mapsto \sum nx_n$ tiene gráfica cerrada y no es continuo.

Regla: elegir cuál de los cuatro teoremas grandes aplicar

Hahn–Banach, Banach–Steinhaus, la aplicación abierta y la gráfica cerrada resuelven cuatro problemas distintos. Identificar cuál toca es cuestión de mirar qué te dan y qué te piden.

1. **¿Necesitas *fabricar* un funcional?** Hahn–Banach. Extiende un funcional definido en un subespacio a todo el espacio sin aumentar la norma. Sus corolarios son los que se usan de verdad: para cada $x \neq 0$ existe φ con $\|\varphi\| = 1$ y $\varphi(x) = \|x\|$; y si $\varphi(x) = 0$ para todo $\varphi \in X^*$, entonces $x = 0$.

2. **¿Tienes una familia de operadores acotada punto a punto?** Banach–Steinhaus (principio de acotación uniforme). Si $\sup_n \|T_n x\| < \infty$ para cada x y X es de Banach, entonces $\sup_n \|T_n\| < \infty$: lo puntual se vuelve uniforme gratis.

3. **¿Tienes un operador biyectivo y quieres invertirlo con control?** Aplicación abierta, o su corolario, el teorema del isomorfismo de Banach: si $T : X \to Y$ es lineal, continuo y biyectivo entre espacios de Banach, entonces T^{-1} es automáticamente continuo. No hay que probar nada más.

4. **¿Quieres demostrar que un operador es continuo y te cuesta?** Gráfica cerrada. Basta con probar la implicación más débil: si $x_n \to x$ *y además* $Tx_n \to y$, entonces $y = Tx$. Es mucho menos que probar continuidad directamente, porque te regalan la convergencia de la imagen.

5. **Verifica siempre la hipótesis de completitud.** Los cuatro descansan sobre el teorema de Baire, y los cuatro son falsos sin ella. Antes de invocarlos, comprueba que estás en espacios de Banach.

6. **Recuerda la dirección de cada uno.** Hahn–Banach construye; los otros tres transforman una hipótesis débil en una conclusión fuerte de manera automática.

7. **Control:** si tu demostración no usa la completitud en ninguna parte, o no necesitabas el teorema, o la demostración está mal.

Ejemplo 32.47 (Cada teorema en su problema)**. (a) Hahn–Banach: el dual separa puntos.** ¿Puede haber dos vectores distintos $x \neq y$ que ningún funcional continuo distinga? No. Aplicando el corolario a $z = x - y \neq 0$ se obtiene φ con $\varphi(z) = \|z\| \neq 0$, luego $\varphi(x) \neq \varphi(y)$. Este hecho, aparentemente inocente, es el que hace que la convergencia débil sea una noción razonable: si $x_n \rightharpoonup x$ y $x_n \rightharpoonup y$, entonces $\varphi(x) = \varphi(y)$ para todo φ y por tanto $x = y$. El límite débil es único gracias a Hahn–Banach.

(b) Banach–Steinhaus: toda sucesión débilmente convergente es acotada. Si $x_n \rightharpoonup x$ en X de Banach, cada $\varphi \in X^*$ cumple que $(\varphi(x_n))_n$ converge y por tanto está acotada. Mirando los x_n como funcionales sobre X^* (que es de Banach), tenemos una familia acotada punto a punto; el principio de acotación uniforme da $\sup_n \|x_n\| < \infty$. Nótese lo que se ha ganado: de infinitas cotas separadas, una sola cota común.

(c) Aplicación abierta: normas comparables son equivalentes. Supón que X es completo con dos normas $\|\cdot\|_a$ y $\|\cdot\|_b$, y que existe C con $\|x\|_b \leq C\|x\|_a$ para todo x. Entonces la identidad $(X, \|\cdot\|_a) \to (X, \|\cdot\|_b)$ es continua y biyectiva entre espacios de Banach; por el paso 3 su inversa es continua, es decir, existe $c > 0$ con $\|x\|_a \leq c\|x\|_b$. *Una* desigualdad implica la otra. Sin completitud esto es rotundamente falso: en $C[0,1]$ se tiene $\|f\|_1 \leq \|f\|_\infty$ y las normas no son equivalentes, porque $\|\cdot\|_1$ no es completa.

(d) Gráfica cerrada: un operador simétrico definido en todo el espacio es acotado (Hellinger–Toeplitz). Sea $T : H \to H$ lineal en un Hilbert con $\langle Tx, y\rangle = \langle x, Ty\rangle$ para todos x, y. Para aplicar el paso 4, supongamos $x_n \to x$ y $Tx_n \to z$; hay que ver que $z = Tx$. Para cualquier y,

$$\langle z, y\rangle = \lim \langle Tx_n, y\rangle = \lim \langle x_n, Ty\rangle = \langle x, Ty\rangle = \langle Tx, y\rangle,$$

y como esto vale para todo y, resulta $z = Tx$. La gráfica es cerrada, luego T es continuo. Es un resultado sorprendente: la simetría, que es una condición algebraica, obliga a la continuidad, que es analítica.

Cuidado

El paso 5 no es una formalidad. El funcional $\varphi(p) = p'(1)$ sobre el espacio de los polinomios con la norma del supremo en $[0,1]$ es lineal y no acotado: con $p_n(x) = x^n$ se tiene $\|p_n\|_\infty = 1$ y $\varphi(p_n) = n \to \infty$. La familia $\{\varphi_n(p) = p^{(n)}(1)/n!\,\}$ está acotada punto a punto sobre los polinomios y sin embargo no uniformemente. No hay contradicción con Banach–Steinhaus: los polinomios con esa norma no forman un espacio de Banach.

Pausa para pensar

Los cuatro teoremas se dividen en dos familias por su demostración. Hahn–Banach se prueba con el lema de Zorn y vale en cualquier normado, incluso incompleto. Los otros tres se prueban con el teorema de Baire y exigen completitud. Si alguna vez dudas de si puedes usar uno, pregúntate cuál de las dos herramientas hay debajo.

Regla: calcular la norma de un funcional e identificar el dual de un espacio

El dual X^* es el espacio de los funcionales lineales continuos $\varphi : X \to \mathbb{C}$, con la norma $\|\varphi\| = \sup_{\|x\|\le 1} |\varphi(x)|$. Calcular esa norma es un ejercicio de dos desigualdades.

1. **Acota por arriba con la desigualdad que corresponda.** Escribe $|\varphi(x)|$ y estímalo por algo de la forma $C\|x\|$. En ℓ^p y L^p la herramienta es Hölder; en $C[a,b]$ basta sacar el supremo fuera de la integral. Esto da $\|\varphi\| \le C$.

2. **Acota por abajo exhibiendo un vector.** Busca x_0 con $\|x_0\| \le 1$ y $|\varphi(x_0)|$ tan cerca de C como quieras. Si lo encuentras con $|\varphi(x_0)| = C$, la norma *se alcanza*; si solo consigues una sucesión que se aproxima, la norma es C igualmente pero no se alcanza.

3. **Recuerda cuándo la cota superior es exacta.** En Hölder hay igualdad cuando $|f|^p$ y $|g|^q$ son proporcionales; ese es el vector que buscas en el paso 2. Si no puede existir dentro del espacio (por ejemplo porque tendría que ser discontinuo), la norma no se alcanza.

4. **Consulta la tabla de duales.** $(\ell^p)^* = \ell^q$ y $(L^p)^* = L^q$ para $1 \le p < \infty$ con $\frac{1}{p} + \frac{1}{q} = 1$; $c_0^* = \ell^1$; y $H^* \cong H$ para todo espacio de Hilbert (Riesz). En cambio $(\ell^\infty)^*$ es estrictamente mayor que ℓ^1.

5. **En un Hilbert, usa Riesz y ahórrate los dos pasos anteriores.** Todo funcional continuo es $\varphi(x) = \langle x, y\rangle$ para un único y, y entonces $\|\varphi\| = \|y\|$ automáticamente. El problema se reduce a identificar y.

6. **Usa Hahn–Banach cuando necesites que el funcional exista.** Para todo $x_0 \neq 0$ hay $\varphi \in X^*$ con $\|\varphi\| = 1$ y $\varphi(x_0) = \|x_0\|$. De aquí sale que el dual separa puntos y que la inmersión canónica $X \hookrightarrow X^{**}$ es isométrica.

7. **Verifica con la fórmula de la distancia.** Para un subespacio cerrado M, se cumple $\operatorname{dist}(x, M) = \sup\{|\varphi(x)| : \varphi \in M^\perp,\ \|\varphi\| \le 1\}$. Si tu funcional da un valor mayor que la distancia real, has calculado mal la norma.

Ejemplo 32.48 (Cinco funcionales, y el que no alcanza su norma)**.** Recorremos los espacios más habituales, uno por uno.

(a) Un funcional en ℓ^2 vía Riesz. Sea $\varphi(x) = \sum_{k\ge 1} \frac{x_k}{k}$. El vector representante es $a = (1, \frac{1}{2}, \frac{1}{3}, \ldots)$, que está en ℓ^2 porque $\sum k^{-2}$ converge. Por el paso 5,

$$\|\varphi\| = \|a\|_2 = \Big(\sum_{k\ge 1} \frac{1}{k^2}\Big)^{1/2} = \frac{\pi}{\sqrt{6}} \approx 1{,}28255,$$

usando el valor de Basilea. La norma se alcanza en $x = a/\|a\|_2$, y ahí $\varphi(x) = \|a\|_2$ exactamente.

(b) El mismo funcional en ℓ^1 y en ℓ^∞. En ℓ^1 la desigualdad de Hölder con $p = 1$, $q = \infty$ da $\|\varphi\| = \|a\|_\infty = 1$, alcanzada en $x = e_1$. Pero en ℓ^∞ el funcional $\varphi(x) = \sum x_k/k$ ni siquiera está definido para todo x: con $x = (1, 1, 1, \ldots)$ la serie es la armónica y diverge. Cambiar el espacio de partida cambia el funcional, no solo su norma.

(c) Un funcional en $L^2(0,1)$. Sea $\varphi(f) = \int_0^1 x\, f(x)\, dx$. Su representante de Riesz es $y(x) = x$, luego

$$\|\varphi\| = \|y\|_2 = \Big(\int_0^1 x^2 dx\Big)^{1/2} = \frac{1}{\sqrt{3}} \approx 0{,}57735,$$

alcanzada en $f(x) = \sqrt{3}\, x$. Compruébalo: $\int_0^1 x \cdot \sqrt{3} x\, dx = \sqrt{3}/3 = 1/\sqrt{3}$. ✓

(d) Dos funcionales en $C[0,1]$: uno alcanza su norma y el otro no. Con la norma del supremo, sea primero $\varphi_1(f) = f(0) - f(1)$. Claramente $|\varphi_1(f)| \le 2\|f\|_\infty$, y tomando cualquier f continua con $f(0) = 1$, $f(1) = -1$ y $|f| \le 1$ (por ejemplo $f(x) = 1 - 2x$) se obtiene $\varphi_1(f) = 2$: la norma es 2 y se alcanza. Ahora

$$\varphi_2(f) = \int_0^{1/2} f(x)\, dx - \int_{1/2}^1 f(x)\, dx.$$

La cota superior es inmediata: $|\varphi_2(f)| \le \int_0^1 |f| \le \|f\|_\infty$, luego $\|\varphi_2\| \le 1$. Y se aproxima: tomando f_n continua, igual a 1 en $[0, \frac12 - \frac1n]$, a -1 en $[\frac12 + \frac1n, 1]$ y lineal en medio, se tiene $\varphi_2(f_n) = 1 - \frac1n \to 1$, luego $\|\varphi_2\| = 1$. Pero la igualdad exigiría $f = 1$ en $[0, \frac12)$ y $f = -1$ en $(\frac12, 1]$, y esa función no es continua. *La norma no se alcanza.*

Cuidado

Que un funcional de norma 1 no alcance su norma es imposible en dimensión finita (la bola es compacta) y también en todo espacio reflexivo. El teorema de James dice justamente eso al revés: un espacio de Banach es reflexivo si y solo si *todo* funcional alcanza su norma. El ejemplo (d) es, por tanto, una demostración de que $C[0,1]$ no es reflexivo.

(e) La fórmula de la distancia, con un número conocido. En $L^2(-1,1)$, sea M el subespacio de los polinomios de grado ≤ 2 y $f(x) = |x|$. Ya calculamos la proyección: $\operatorname{dist}(f, M) = \|f - S_2\| = 1/\sqrt{96} \approx 0{,}1021$. El funcional que realiza ese supremo en el paso 7 es

$$\varphi(g) = \frac{\langle g,\, f - S_2\rangle}{\|f - S_2\|},$$

que se anula sobre M (porque $f - S_2 \perp M$), tiene norma 1, y cumple $\varphi(f) = \|f - S_2\|$. En un Hilbert la distancia siempre se alcanza y siempre hay un funcional óptimo; en un Banach general Hahn–Banach garantiza que existe, pero no da ninguna fórmula.

(f) Un dual que no se puede escribir. $(\ell^1)^* = \ell^\infty$ es concreto: cada $a \in \ell^\infty$ da $\varphi(x) = \sum a_k x_k$ y todos son de esa forma. Pero $(\ell^\infty)^*$ contiene funcionales que no provienen de ningún elemento de ℓ^1: los *límites de Banach*, que asignan a cada sucesión acotada un «límite» consistente con el habitual, invariante por desplazamiento, y definido incluso para $(1,0,1,0,\ldots)$, donde vale $\frac12$. Su existencia se demuestra con Hahn–Banach y no admite ninguna construcción explícita: son objetos que sabemos que hay y que nadie puede escribir.

Pausa para pensar

Repasa (b) y (f) juntos. En (b) el mismo símbolo $\sum x_k/k$ define un funcional acotado sobre ℓ^1 y sobre ℓ^2, pero ni siquiera está definido sobre ℓ^∞. En (f) sobre ℓ^∞ hay funcionales

que sobre ℓ^1 serían imposibles. La moraleja es que un funcional no es una fórmula: es una fórmula *más* un dominio, y el dominio manda.

Idea clave

El teorema de Banach–Alaoglu dice que la bola unidad de X^* es compacta en la topología débil-$*$, siempre, en cualquier espacio de Banach. Es la compacidad que se perdió al pasar a dimensión infinita, recuperada al precio de debilitar la topología. Combinado con Hahn–Banach, es la herramienta que permite extraer límites en problemas variacionales: existen mínimos porque hay compacidad débil, aunque no haya compacidad en norma.

32.8. El principio de contracción de Banach

Los cuatro teoremas grandes de la sección anterior son teoremas de estructura: dicen cómo se comportan los operadores acotados entre espacios de Banach. Hay un quinto resultado, más humilde y muchísimo más usado, que no necesita linealidad para nada y que sin embargo es el motivo por el que la completitud vale la pena. Se llama principio de contracción o teorema del punto fijo de Banach, y viene con algo que casi ningún teorema de existencia regala: un algoritmo que produce la solución y una cota del error en cada paso.

Definición 32.49 (Contracción)**.** Sea (X, d) un espacio métrico. Una aplicación $T\colon X \to X$ es una *contracción* si existe una constante $k < 1$, la misma para todos los puntos, tal que

$$d\big(T(x), T(y)\big) \le k\, d(x, y) \qquad \text{para todos } x, y \in X.$$

La constante k se llama factor de contracción. Obsérvese que se exige $k < 1$ estrictamente: la desigualdad $d(Tx, Ty) < d(x, y)$ punto a punto *no* basta.

Idea clave

Qué dice esto en cristiano. El teorema del punto fijo de Banach es el resultado más rentable del análisis: dice que si una aplicación acerca los puntos entre sí en una proporción fija menor que uno, entonces tiene un único punto fijo, y además se encuentra iterando desde donde sea. Es a la vez un teorema de existencia, uno de unicidad y un algoritmo con cota de error. De él salen el teorema de existencia de soluciones de ecuaciones diferenciales, el de la función implícita y la mitad de los métodos numéricos.

Ejemplo 32.50 (Iterar hasta el punto fijo)**. (a) Resolver** $\cos x = x$**.** La aplicación $T(x) = \cos x$ sobre $[0, 1]$ cumple $|T'(x)| = |\sin x| \le \sin 1 = 0{,}841471 < 1$: es una contracción. Iterando desde $x_0 = 0$ se obtiene 1; 0,540302; 0,857553; 0,654290; 0,793480… y converge a 0,7390851, el punto fijo.

(b) El método babilónico. Para $\sqrt{2}$, la aplicación $T(x) = \frac{1}{2}\left(x + \frac{2}{x}\right)$ es contractiva cerca de $\sqrt{2}$. Desde $x_0 = 1$ da 1,5; 1,416667; 1,414216; 1,414214. Cuatro pasos y seis cifras correctas: aquí la convergencia es cuadrática porque además la derivada se anula en el punto fijo.

(c) La cota de error, gratis. El teorema da $d(x_n, x^*) \le \frac{k^n}{1-k} d(x_1, x_0)$, con k la constante de contracción. En (a), con $k = 0{,}841471$ y $d(x_1, x_0) = 1$, tras diez iteraciones la cota es $\frac{0{,}841471^{10}}{0{,}158529} = 1{,}1227$: floja, pero calculable sin conocer la solución. La cota siempre existe, aunque a veces sea pesimista.

(d) Existencia de soluciones de ecuaciones diferenciales. El teorema de Picard–Lindelöf convierte $y' = f(x, y)$ con $y(x_0) = y_0$ en la ecuación integral $y(x) = y_0 + \int_{x_0}^{x} f(t, y(t))\,\mathrm{d}t$, y demuestra que el operador del lado derecho es una contracción en un intervalo pequeño. El punto fijo es la solución, y las iteraciones son las aproximaciones sucesivas de Picard.

Cuidado

[Contracción exige $k < 1$ estricto y espacio completo] Con $k = 1$ el teorema falla: la aplicación $T(x) = x + \frac{1}{x}$ sobre $[1, \infty)$ cumple $|T(x) - T(y)| < |x - y|$ para todos los pares distintos y no tiene ningún punto fijo, porque $T(x) > x$ siempre. Acercar no basta; hay que acercar en una proporción *uniforme*. Segundo aviso: el espacio debe ser completo. En $(0, 1]$ con la distancia usual, $T(x) = \frac{x}{2}$ es una contracción con $k = \frac{1}{2}$ y su punto fijo sería 0, que no está en el espacio. Y tercero: T debe llevar el espacio dentro de sí mismo; si las iteraciones se escapan, no hay nada que iterar.

Teorema 32.51 (Principio de contracción de Banach)**.** *Sea X un espacio métrico* completo *y no vacío, y sea $T\colon X \to X$ una contracción de factor $k < 1$. Entonces T tiene exactamente un punto fijo x^*, y para cualquier punto inicial $x_0 \in X$ la sucesión $x_{n+1} = T(x_n)$ converge a x^* con las cotas*

$$d(x_n, x^*) \le \frac{k^n}{1-k}\, d(x_1, x_0), \qquad d(x_n, x^*) \le \frac{k}{1-k}\, d(x_n, x_{n-1}).$$

Demostración. Por inducción $d(x_{n+1}, x_n) \le k^n d(x_1, x_0)$. Para $m > n$, la desigualdad triangular y la suma geométrica dan

$$d(x_m, x_n) \le \sum_{j=n}^{m-1} k^j d(x_1, x_0) \le \frac{k^n}{1-k} d(x_1, x_0) \xrightarrow[n\to\infty]{} 0,$$

luego (x_n) es de Cauchy. Aquí es donde entra la completitud: sin ella la sucesión existiría igual pero podría no converger a nada dentro de X. Sea x^* su límite; como toda contracción es continua, $T(x^*) = \lim T(x_n) = \lim x_{n+1} = x^*$. Si y^* fuera otro punto fijo, $d(x^*, y^*) = d(Tx^*, Ty^*) \le k\, d(x^*, y^*)$, y como $k < 1$ esto obliga a $d(x^*, y^*) = 0$. Haciendo $m \to \infty$ en la estimación anterior se obtiene la primera cota; la segunda sale de aplicarla con x_{n-1} como punto inicial. □

Ejemplo 32.52 (El punto fijo, calculado)**. (a) El coseno.** Sea $T(x) = \cos x$ sobre $[0, 1]$, con $|T'| = |\sin x| \le \sin 1 = 0{,}841471 < 1$. Iterando desde $x_0 = 0{,}5$ salen 0,877583; 0,639012; 0,802685; luego 0,768196 en el quinto paso y 0,730081 en el octavo, oscilando alrededor del punto fijo 0,739085. La convergencia es lenta porque k está cerca de 1.

(b) Newton como contracción. Para $\sqrt{2}$, la iteración $T(x) = (x + 2/x)/2$ desde $x_0 = 1{,}5$ da 1,4166666667; 1,4142156863; 1,4142135624, ya exacto en diez cifras al tercer paso. Aquí la

convergencia es cuadrática porque T' se anula en el punto fijo: es mejor que una contracción cualquiera.

(c) Sin completitud, falla. En $X = (0, 1]$ con la distancia usual, $T(x) = x/2$ es una contracción de factor $1/2$ y *no* tiene punto fijo, porque el único candidato, el 0, no está en el espacio. La completitud es exactamente lo que garantiza que el límite de la sucesión iterada exista dentro de X.

(d) Sin $k < 1$ estricto, falla. En $X = [1, \infty)$, la aplicación $T(x) = x + 1/x$ cumple $|T(x) - T(y)| < |x - y|$ para $x \neq y$, es decir, contrae estrictamente distancias, y sin embargo no tiene punto fijo: $T(x) > x$ siempre. Hace falta una constante $k < 1$ *uniforme*, no basta con la desigualdad estricta punto a punto.

Idea clave

El teorema convierte una pregunta de existencia en una receta de cocina. En vez de demostrar que la ecuación $T(x) = x$ tiene solución, uno empieza en cualquier sitio y aplica T una y otra vez; el factor k mide cuánto se encoge el error en cada vuelta, así que el número de decimales correctos crece linealmente con el número de iteraciones. La completitud del espacio es lo que garantiza que el punto al que se dirige la sucesión *existe*: por eso este resultado es de análisis funcional y no de álgebra.

Regla: resolver una ecuación por iteración de punto fijo

Casi cualquier ecuación $F(x) = 0$ puede reescribirse como $x = g(x)$, pero no todas las reescrituras sirven. Estos pasos separan las que funcionan de las que no.

1. **Despeja para obtener la forma $x = g(x)$.** Hay muchas maneras de hacerlo y el teorema no se aplica a todas por igual: la elección del despeje es la decisión importante.
2. **Elige un conjunto cerrado C y comprueba que $g(C) \subseteq C$.** Un cerrado dentro de un espacio completo vuelve a ser completo, y sin invariancia la iteración puede escaparse del dominio donde controlas g.
3. **Acota la derivada en C.** Si $|g'(x)| \leq k$ en todo C con $k < 1$, el teorema del valor medio da $|g(x) - g(y)| \leq k|x - y|$, es decir, g es contracción con factor k.
4. **Si $k \geq 1$, cambia de despeje.** No hay nada que salvar: prueba otro reagrupamiento, o aplica el método de Newton, cuya función de iteración tiene derivada nula en la raíz.
5. **Itera desde cualquier punto de C.** El punto inicial afecta a la velocidad pero nunca al límite, porque el punto fijo es único.
6. **Detente con la cota a posteriori.** La diferencia entre dos iterados consecutivos, multiplicada por $k/(1-k)$, acota el error verdadero; es el criterio de parada honesto.

7. **Traduce el resultado.** En espacios de funciones el "punto" es una función entera, así que el punto fijo es la solución de una ecuación diferencial o integral, no un número.

Ejemplo 32.53 (Cuatro iteraciones, dos que convergen y una que huye)**.** (a) **La ecuación $x = \cos x$.** Aquí $g = \cos$ manda $[0,1]$ dentro de $[0{,}5403,\,1] \subseteq [0,1]$ y $|g'(x)| = |\sin x| \le \sin 1 \approx 0{,}84147$, luego es contracción con $k \approx 0{,}841$. Desde $x_0 = 0{,}5$ los iterados son 0,87758, 0,63901, 0,80269, 0,69478, 0,76820, oscilando alrededor del punto fijo $x^* \approx 0{,}7390851332$. Converge, sí, pero con k tan cerca de 1 hacen falta unas cuarenta vueltas para seis decimales.

(b) **La ecuación $x^3+x-1=0$, bien despejada.** Escribiéndola como $x = 1/(1+x^2)$ se tiene $g'(x) = -2x/(1+x^2)^2$, cuyo máximo en valor absoluto sobre $[\tfrac12, 1]$ se alcanza en $x = 1/\sqrt{3}$ y vale $3\sqrt{3}/8 \approx 0{,}64952$. Además $g([\tfrac12,1]) = [\tfrac12,\,0{,}8] \subseteq [\tfrac12, 1]$. Desde $y_0 = 0{,}5$: 0,8, 0,60976, 0,72897, 0,65300, 0,70106, y ya $y_{10} = 0{,}680394$, $y_{15} = 0{,}682528$, $y_{20} = 0{,}6823071$, contra la raíz 0,682327803828. La cota a priori exige $n = 32$ para garantizar 10^{-6}; el error real en $n = 20$ ya es $2{,}1 \cdot 10^{-5}$, lo que recuerda que la cota es pesimista.

(c) **La misma ecuación, mal despejada.** Si en cambio se escribe $x = 1 - x^3$, entonces $g'(x) = -3x^2$ y en la raíz $|g'| \approx 1{,}40 > 1$: no es contracción cerca del punto fijo y la iteración se desestabiliza. El caso extremo es $g(x) = 2x - 1$, con punto fijo $x^* = 1$ y $k = 2$; desde 1,1 los iterados son 1,2, 1,4, 1,8, 2,6, 4,2, 7,4, alejándose sin remedio.

(d) **Newton como punto fijo privilegiado.** Para $F(x) = x^3 + x - 1$, la iteración de Newton es $g(x) = x - \frac{x^3+x-1}{3x^2+1}$, y desde 0,7 da 0,682591093, 0,682327863, 0,682327803828: tres pasos para doce decimales. La razón es que $g'(x^*) = 0$, así que localmente k es tan pequeño como se quiera y la convergencia deja de ser lineal para ser cuadrática.

Ejemplo 32.54 (El punto fijo vive en un espacio de funciones)**.** (a) **Picard y las ecuaciones diferenciales.** El problema $u' = u$, $u(0) = 1$ equivale por integración a $u(x) = 1 + \int_0^x u(t)\,dt$, es decir, al punto fijo del operador $Tu(x) = 1 + \int_0^x u$. En $C[0, \tfrac12]$ con la norma del supremo, $\|Tu - Tv\|_\infty \le \tfrac12 \|u - v\|_\infty$, luego $k = \tfrac12$. Iterando desde $u_0 \equiv 1$ salen $1 + x$, $1 + x + \frac{x^2}{2}$, $1 + x + \frac{x^2}{2} + \frac{x^3}{6}$, que en $x = \tfrac12$ valen 1,5, 1,625, 1,6458333, 1,6484375, camino de $\mathrm{e}^{1/2} = 1{,}6487212707$.

(b) **Por qué eso es el teorema de existencia.** El mismo argumento con $u' = f(x,u)$ y f lipschitziana en la segunda variable con constante L da una contracción en $C[0,a]$ siempre que $La < 1$, y su punto fijo es la solución del problema de valor inicial. Eso es exactamente el teorema de Picard–Lindelöf del capítulo 12: existencia y unicidad local no son un milagro, son un punto fijo.

(c) **La serie de Neumann.** Si A es un operador acotado con $\|A\| < 1$, la aplicación $Tx = Ax + b$ es contracción de factor $\|A\|$, y su único punto fijo es $x = (I - A)^{-1}b = \sum_{n\ge0} A^n b$. Aquí la iteración *es* la suma parcial de la serie geométrica de operadores, y el teorema explica de una vez por qué la serie converge y por qué su suma es la inversa.

(d) **El truco de la norma equivalente.** Para el operador de Volterra $Vf(x) = \int_0^x f$ en $C[0,1]$ se tiene $\|V\| = 1$, así que la serie de Neumann parece no aplicarse. Pero $\|V^n f\|_\infty \le \|f\|_\infty / n!$, luego $\|\lambda V\| < 1$ tras iterar unas cuantas veces para cualquier λ, y $I - \lambda V$ resulta invertible siempre. Ese es el motivo estructural de que el espectro de V sea $\{0\}$.

Cuidado

La condición $k < 1$ debe valer con una constante *uniforme*. La aplicación $T(x) = x + \frac{1}{x}$ sobre $[1,\infty)$ cumple $|T(x)-T(y)| < |x-y|$ para todos los pares de puntos distintos, y sin embargo no tiene ningún punto fijo: los supuestos candidatos se escapan al infinito. La otra hipótesis que no se puede quitar es la completitud: sobre $\mathbb{Q}$, la iteración $x \mapsto \frac{x}{2} + \frac{1}{x}$ es contractiva en $[1,2]\cap\mathbb{Q}$ y sus iterados son todos racionales, pero convergen a $\sqrt{2}$, que no está en el espacio.

Pausa para pensar

El teorema del punto fijo no usa en ningún momento que X sea un espacio vectorial: sólo necesita una distancia y completitud. ¿Qué se gana entonces enunciándolo dentro del análisis funcional? La respuesta es que los espacios donde uno quiere aplicarlo —$C[a,b]$, L^2, los espacios de sucesiones— son precisamente los que este capítulo demuestra completos, y la norma es la herramienta con la que se calcula el factor k. Existen versiones que renuncian a la contracción y piden convexidad y compacidad en su lugar; la más famosa es el teorema de Schauder, que da existencia pero pierde tanto la unicidad como el algoritmo.

32.9. Dualidad y el teorema de representación de Riesz

Definición 32.55. El *dual* de un espacio normado X es $X^* = B(X,\mathbb{R})$ (o $B(X,\mathbb{C})$), siempre un espacio de Banach.

Teorema 32.56 (Riesz–Fréchet: en Hilbert, el dual es el propio espacio)**.** *Sea H de Hilbert y $\varphi \in H^*$. Entonces existe un único $y \in H$ con*

$$\varphi(x) = \langle x, y\rangle \quad \text{para todo } x, \qquad y \qquad \|\varphi\| = \|y\|.$$

Demostración. Si $\varphi = 0$ toma $y = 0$. Si no, $M = \ker\varphi$ es un subespacio cerrado propio, luego $M^\perp \neq \{0\}$; toma $z \in M^\perp$ unitario. Para todo x, el vector $\varphi(x)z - \varphi(z)x$ está en M, luego es ortogonal a z:

$$0 = \langle \varphi(x)z - \varphi(z)x,\ z\rangle = \varphi(x) - \varphi(z)\langle x, z\rangle,$$

de donde $\varphi(x) = \langle x, \overline{\varphi(z)}z\rangle$ y basta tomar $y = \overline{\varphi(z)}z$. La unicidad y la igualdad de normas son inmediatas por Cauchy–Schwarz, con igualdad en $x = y$. □

Observación 32.57. Este teorema es el que da sentido a la notación de Dirac de la física: el *bra* $\langle\psi|$ es un funcional continuo, el *ket* $|\psi\rangle$ es el vector que lo representa, y el teorema de

Riesz dice que la correspondencia entre ambos es biyectiva e isométrica. También es la base del método variacional para ecuaciones en derivadas parciales: el teorema de Lax–Milgram, que resuelve $a(u,v)=\varphi(v)$, es Riesz más una perturbación acotada.

Ejemplo 32.58 (Duales conocidos)**.** Para $1<p<\infty$ y $\frac{1}{p}+\frac{1}{q}=1$, se tiene $(L^p)^*\cong L^q$ y $(\ell^p)^*\cong\ell^q$, mediante $g\mapsto(f\mapsto\int fg)$. Además $(\ell^1)^*\cong\ell^\infty$ y $c_0^*\cong\ell^1$. En cambio $(\ell^\infty)^*$ es estrictamente mayor que ℓ^1: contiene funcionales exóticos (límites de Banach) cuya existencia solo se demuestra con Hahn–Banach y que no se pueden escribir. Un espacio es *reflexivo* si $X^{**}=X$ canónicamente: L^p con $1<p<\infty$ lo es, L^1, L^∞, c_0 y $C[0,1]$ no.

32.10. Operadores compactos y el teorema espectral

Definición 32.59. $T\in B(X,Y)$ es *compacto* si la imagen de la bola unidad tiene clausura compacta. Equivalentemente, toda sucesión acotada (x_n) tiene una subsucesión con (Tx_{n_k}) convergente.

Idea clave

Intuición primero. El teorema de Riesz decía que en dimensión infinita la bola unidad no es compacta. Un operador compacto es el que la vuelve a comprimir hasta hacerla casi de dimensión finita: aplasta la bola en algo pequeño. De hecho, en espacios de Hilbert, los operadores compactos son exactamente los límites en norma de operadores de rango finito. Por eso se comportan como matrices grandes, y por eso admiten un teorema espectral casi idéntico al de matrices simétricas del capítulo 7.

Ejemplo 32.60. El operador integral $(Kf)(x)=\int_0^1 k(x,t)f(t)\,\mathrm{d}t$ con núcleo $k\in L^2([0,1]^2)$ es compacto en $L^2[0,1]$ (operador de Hilbert–Schmidt); el de Volterra es el caso $k=\mathbf{1}_{t\le x}$. La identidad en un espacio de dimensión infinita *no* es compacta —eso es literalmente el teorema de Riesz—, y por tanto un operador compacto nunca es invertible en dimensión infinita.

Teorema 32.61 (Espectral para operadores compactos autoadjuntos)**.** *Sea H de Hilbert separable y $T\in B(H)$ compacto y autoadjunto ($\langle Tx,y\rangle=\langle x,Ty\rangle$). Entonces existe una base ortonormal de $\overline{\operatorname{im}T}$ formada por autovectores de T, con autovalores reales $\lambda_n\to0$, y*

$$Tx=\sum_n \lambda_n\langle x,e_n\rangle e_n.$$

Además $\|T\|=\max_n|\lambda_n|$, y ese máximo se alcanza.

Esquema. El paso decisivo es que $\|T\|$ o $-\|T\|$ es autovalor. Para verlo, se usa que para T autoadjunto $\|T\|=\sup_{\|x\|=1}|\langle Tx,x\rangle|$; se toma una sucesión maximizante x_n, y por compacidad se extrae $Tx_{n_k}\to y$; un cálculo con $\|Tx_n-\lambda x_n\|^2=\|Tx_n\|^2-2\lambda\langle Tx_n,x_n\rangle+\lambda^2\to0$ muestra que x_{n_k} converge a un autovector. Después se restringe T al complemento ortogonal del autovector encontrado —que es invariante por autoadjunción— y se repite. Los autovalores obtenidos tienden a 0 porque, si no, una sucesión de autovectores ortonormales con $|\lambda_n|\ge\varepsilon$ daría $\|Te_n-Te_m\|^2=|\lambda_n|^2+|\lambda_m|^2\ge2\varepsilon^2$, contra la compacidad. □

Ejemplo 32.62 (Diagonalizar en dimensión infinita)**. (a) El caso finito.** En $\mathbb{R}^n$, el teorema espectral para matrices simétricas es el caso particular: base ortonormal de vectores propios, valores propios reales. La versión de Hilbert dice que lo mismo vale para operadores compactos autoadjuntos, con la salvedad de que ahora hay infinitos valores propios y su única acumulación posible es el cero.

(b) Un operador integral. Sea $(Tf)(x) = \int_0^1 \min(x,y) f(y)\,dy$ en $L^2[0,1]$. Es compacto (núcleo continuo) y autoadjunto (el núcleo es simétrico). Sus funciones propias son $\sqrt{2}\sin((n-\frac{1}{2})\pi x)$ con valores propios $\lambda_n = 1/((n-\frac{1}{2})\pi)^2$, que valen 0,405285; 0,045032; 0,016211 para $n = 1, 2, 3$ y tienden a cero como n^{-2}.

(c) Por qué el cero es especial. Como $\lambda_n \to 0$, el operador T no puede tener inversa acotada: la ecuación $Tf = g$ es un problema mal planteado, y por eso invertir un operador integral (deconvolución, tomografía) exige regularización. En dimensión finita esto no pasa nunca.

(d) Sin compacidad, falla. El operador de multiplicación $(Tf)(x) = xf(x)$ en $L^2[0,1]$ es autoadjunto y acotado, y *no* tiene ningún valor propio: $xf(x) = \lambda f(x)$ obliga a $f = 0$ en casi todo punto. Su espectro es todo $[0,1]$ y es puramente continuo. La compacidad es lo que garantiza que el espectro sea una sucesión de puntos.

Observación 32.63 (Para qué se inventó todo esto)*.* Aplicado al operador de Green de un problema de Sturm–Liouville $-(pu')' + qu = \lambda w u$ con condiciones de contorno, este teorema dice que las autofunciones forman una base ortonormal de L^2. Es decir: *todo* desarrollo en autofunciones que aparece al resolver la ecuación del calor, la de ondas o la de Schrödinger por separación de variables está justificado por este único enunciado. Las series de Fourier son el caso particular $p = w = 1$, $q = 0$. Ese fue el objetivo original de Hilbert, y de ahí viene el nombre del espacio.

Regla: calcular el espectro de un operador

En dimensión infinita "autovalor" deja de ser suficiente: hay operadores sin ningún autovalor y con espectro enorme. El espectro es la noción correcta, y se calcula por partes.

1. **Parte de la definición correcta:** $\lambda \in \sigma(T)$ si $T - \lambda I$ *no* es invertible en $\mathcal{B}(X)$. Fallar puede ocurrir por tres motivos: no ser inyectivo, no ser sobreyectivo, o serlo con inversa no acotada (esto último no pasa si X es de Banach, por la aplicación abierta).

2. **Separa las tres piezas.** Espectro puntual: $T - \lambda I$ no es inyectivo, y entonces λ sí es autovalor. Espectro continuo: es inyectivo con imagen densa pero no cerrada. Espectro residual: es inyectivo pero la imagen ni siquiera es densa.

3. **Acota primero con el radio espectral:** $\sigma(T)$ es un compacto no vacío de $\mathbb{C}$ contenido en el disco $|\lambda| \le \|T\|$, y su radio exacto es $r(T) = \lim_n \|T^n\|^{1/n}$. Calcular $\|T^n\|$ suele ser mucho más informativo que calcular $\|T\|$.

4. **Busca los autovalores resolviendo** $Tx = \lambda x$ explícitamente. Si el sistema fuerza $x = 0$, no hay autovalores y hay que seguir con los pasos siguientes.
5. **Aprovecha las estructuras especiales.** Si T es autoadjunto, $\sigma(T) \subset \mathbb{R}$ y $r(T) = \|T\|$. Si T es compacto en dimensión infinita, $0 \in \sigma(T)$ y el resto del espectro son autovalores aislados de multiplicidad finita que solo pueden acumularse en 0. Si T es unitario, $\sigma(T)$ vive en la circunferencia unidad.
6. **Pasa al adjunto cuando te atasques:** $\sigma(T^*) = \overline{\sigma(T)}$ (los conjugados), y el espectro residual de T está contenido en el puntual de T^*. Muchos espectros se calculan por este atajo.
7. **Control:** el espectro nunca es vacío y siempre es cerrado y acotado. Si te sale vacío, o abierto, o no acotado, revisa.

Ejemplo 32.64 (Cuatro espectros, cuatro comportamientos)**. (a) Multiplicación por x en $L^2[0,1]$.** Sea $(Mf)(x) = x\,f(x)$. ¿Autovalores? De $xf(x) = \lambda f(x)$ se sigue $(x-\lambda)f(x) = 0$ en casi todo punto, y como $x - \lambda$ solo se anula en un punto (medida cero), forzosamente $f = 0$ en casi todo punto. *No hay ningún autovalor.* Sin embargo, si $\lambda \in [0,1]$ el operador $M - \lambda I$ tiene inversa formal $f \mapsto f/(x-\lambda)$, que no es acotada porque $1/(x-\lambda)$ no está acotada cerca de $x = \lambda$. Y si $\lambda \notin [0,1]$ esa misma inversa sí es acotada. Conclusión:

$$\sigma(M) = [0,1], \qquad \sigma_p(M) = \varnothing.$$

Aquí M es autoadjunto, el espectro es real, y $\|M\| = 1 = r(M)$, coherente con el paso 5.

(b) El desplazamiento a la derecha en ℓ^2. Sea $S(x_1, x_2, \dots) = (0, x_1, x_2, \dots)$. Es una isometría, $\|Sx\| = \|x\|$, luego $\|S\| = 1$ y también $\|S^n\| = 1$, así que $r(S) = 1$ y $\sigma(S)$ cabe en el disco cerrado. Autovalores no hay: $Sx = \lambda x$ da $0 = \lambda x_1$, $x_1 = \lambda x_2$, …, que obliga a $x = 0$ tanto si $\lambda = 0$ como si no. Pero el adjunto $S^*(x_1, x_2, \dots) = (x_2, x_3, \dots)$ sí tiene muchísimos: para $|\lambda| < 1$ el vector $(1, \lambda, \lambda^2, \dots)$ está en ℓ^2 y cumple $S^*x = \lambda x$. Por el paso 6 y por ser σ cerrado,

$$\sigma(S) = \{\lambda : |\lambda| \le 1\}, \qquad \sigma_p(S) = \varnothing, \qquad \sigma_p(S^*) = \{|\lambda| < 1\}.$$

Un operador sin un solo autovalor y con un disco entero de espectro: el mejor argumento de por qué en dimensión infinita no basta con el polinomio característico.

(c) Un operador compacto diagonal. En ℓ^2, sea $D(x_1, x_2, \dots) = (x_1, \frac{x_2}{2}, \frac{x_3}{3}, \dots)$. Cada e_n es autovector con autovalor $1/n$. El operador es compacto (límite en norma de los truncamientos de rango finito, con $\|D - D_N\| = \frac{1}{N+1} \to 0$), así que por el paso 5 el 0 tiene que estar en el espectro:

$$\sigma(D) = \{0\} \cup \left\{\tfrac{1}{n} : n \ge 1\right\}, \qquad \sigma_p(D) = \left\{\tfrac{1}{n}\right\}.$$

El 0 está en el espectro pero *no* es autovalor: $Dx = 0$ obliga a $x = 0$. Es el punto de acumulación que la teoría de Riesz–Schauder predice, y la imagen de D es densa pero no cerrada.

(d) El operador de Volterra. En $L^2[0,1]$, $(Vf)(x) = \int_0^x f(t)\,\mathrm{d}t$. Aquí conviene el paso 3. Iterando, V^n tiene núcleo $\frac{(x-t)^{n-1}}{(n-1)!}$, de donde $\|V^n\| \le \frac{1}{(n-1)!}$ y

$$r(V) = \lim_n \|V^n\|^{1/n} \le \lim_n \left(\tfrac{1}{(n-1)!}\right)^{1/n} = 0.$$

Como el espectro no puede ser vacío, $\sigma(V) = \{0\}$, y sin embargo $V \neq 0$: es un operador *cuasinilpotente*. Tampoco 0 es autovalor, pues $\int_0^x f = 0$ para todo x implica $f = 0$ derivando. Un operador no nulo cuyo espectro es un solo punto y sin ningún autovector: imposible en dimensión finita, donde espectro $\{0\}$ significa nilpotente.

Cuidado

Que $\|T\|$ sea grande no dice nada sobre el espectro. En (d), $\|V\| = 2/\pi \approx 0{,}6366$ y $r(V) = 0$: la desigualdad $r(T) \le \|T\|$ puede ser estrictísima. Solo cuando T es autoadjunto (o, más generalmente, normal) se tiene la igualdad $r(T) = \|T\|$.

Idea clave

Los cuatro ejemplos ilustran las cuatro maneras de ser espectro. En dimensión finita el espectro siempre es un conjunto finito de autovalores; en dimensión infinita puede ser un intervalo sin autovalores (a), un disco sin autovalores (b), una sucesión de autovalores más su límite (c), o un único punto que tampoco es autovalor (d). El polinomio característico se acabó; lo que queda es la teoría espectral.

32.11. Ecuaciones integrales: el caso del núcleo degenerado

La teoría de operadores compactos se inventó para resolver ecuaciones integrales, y conviene ver cómo se aplica en el caso más concreto posible: aquel en el que el núcleo es una suma finita de productos de una función de x por una función de t. Entonces toda la maquinaria infinito-dimensional se derrumba sobre un sistema lineal de tamaño finito, que se resuelve con las técnicas del capítulo 7. Es el mejor laboratorio para entender qué dice de verdad la alternativa de Fredholm.

Definición 32.65 (Núcleo degenerado)**.** Un núcleo $k(x,t)$ es *degenerado* (o separable, o de rango finito) si se escribe como

$$k(x,t) = \sum_{i=1}^{n} a_i(x)\, b_i(t)$$

con $a_i, b_i \in L^2[0,1]$. El operador asociado $Kf(x) = \int_0^1 k(x,t) f(t)\,\mathrm{d}t$ tiene entonces imagen contenida en el subespacio $\operatorname{span}\{a_1, \dots, a_n\}$, de dimensión a lo sumo n, y en particular es compacto.

Idea clave

Si el operador sólo puede producir combinaciones de n funciones fijas, la incógnita de una ecuación como $u - \lambda K u = f$ ya no es una función arbitraria: es f más una combinación de esas n funciones. Las verdaderas incógnitas son los n coeficientes, y se determinan sustituyendo el ansatz dentro de la propia ecuación. Un problema en dimensión infinita se ha convertido en un sistema de n ecuaciones con n incógnitas, y todo lo que diga la teoría abstracta —cuándo hay solución, cuándo es única, qué valores de λ son excepcionales— se lee en un determinante.

Regla: resolver una ecuación de Fredholm de núcleo degenerado

La ecuación de segunda especie es $u(x) - \lambda \int_0^1 k(x,t)u(t)\,\mathrm{d}t = f(x)$, con λ un parámetro y u la incógnita.

1. **Separa el núcleo.** Escribe $k(x,t) = \sum_i a_i(x)b_i(t)$ agrupando todo lo que depende de x a la izquierda; el número n de sumandos es el tamaño del sistema que vas a obtener.
2. **Bautiza las integrales desconocidas.** Pon $c_i = \int_0^1 b_i(t)u(t)\,\mathrm{d}t$. Son números, no funciones, y son las únicas incógnitas verdaderas del problema.
3. **Despeja la incógnita en función de los c_i.** La ecuación se lee $u(x) = f(x) + \lambda \sum_i c_i a_i(x)$: la solución ya está escrita, salvo por n constantes.
4. **Sustituye ese ansatz en la definición de los c_i.** Multiplicando por b_j e integrando sale un sistema lineal $(I - \lambda A)c = \varphi$, con $A_{ji} = \int_0^1 b_j a_i$ y $\varphi_j = \int_0^1 b_j f$.
5. **Estudia el determinante $\det(I - \lambda A)$.** Es un polinomio en λ de grado a lo sumo n; sus raíces son exactamente los valores excepcionales, y son los recíprocos de los autovalores no nulos de K.
6. **Si el determinante no se anula, resuelve y sustituye.** Hay solución única para toda f, y se obtiene volviendo al paso 3 con los c_i ya calculados.
7. **Si el determinante se anula, aplica la alternativa.** Entonces la ecuación homogénea tiene soluciones no triviales, y la inhomogénea sólo es resoluble si f es ortogonal a todas las soluciones del problema adjunto; en ese caso hay infinitas.

Ejemplo 32.66 (El núcleo $k(x,t) = x + t$, resuelto entero)**.** (a) **Montaje.** Se busca u con $u(x) - \lambda \int_0^1 (x+t)u(t)\,\mathrm{d}t = x$. Como $x + t = x \cdot 1 + 1 \cdot t$, el núcleo es degenerado con $n = 2$. Poniendo $c_1 = \int_0^1 u$ y $c_2 = \int_0^1 tu(t)\,\mathrm{d}t$, el ansatz es $u(x) = x + \lambda(c_1 x + c_2)$.

(b) **El sistema.** Integrando ese ansatz y luego multiplicándolo por t e integrando otra vez se obtiene

$$\left(1 - \tfrac{\lambda}{2}\right)c_1 - \lambda c_2 = \tfrac{1}{2}, \qquad -\tfrac{\lambda}{3}c_1 + \left(1 - \tfrac{\lambda}{2}\right)c_2 = \tfrac{1}{3},$$

cuyo determinante es $\left(1 - \tfrac{\lambda}{2}\right)^2 - \tfrac{\lambda^2}{3} = -\tfrac{1}{12}(\lambda^2 + 12\lambda - 12)$.

(c) **La solución general.** Resolviendo, $c_1 = \dfrac{-\lambda - 6}{\lambda^2 + 12\lambda - 12}$ y $c_2 = \dfrac{-4}{\lambda^2 + 12\lambda - 12}$, y por tanto

$$u(x) = \frac{2(3\lambda x - 2\lambda - 6x)}{\lambda^2 + 12\lambda - 12}.$$

Para $\lambda = 1$ esto da $u(x) = -6x - 4$, y en efecto $-6x - 4 - \int_0^1 (x+t)(-6t-4)\,\mathrm{d}t = x$. Para $\lambda = \frac{1}{2}$ da $u(x) = \frac{36x+8}{23}$, que también se comprueba en una línea.

(d) **Los valores excepcionales.** El determinante se anula cuando $\lambda^2 + 12\lambda - 12 = 0$, es decir en $\lambda = -6+4\sqrt{3} \approx 0{,}9282$ y $\lambda = -6-4\sqrt{3} \approx -12{,}9282$. Y no es casualidad: K deja invariante $\operatorname{span}\{1, x\}$, donde su matriz en la base $\{1, x\}$ es $\begin{pmatrix} 1/2 & 1/3 \\ 1 & 1/2 \end{pmatrix}$, con polinomio característico $12\mu^2 - 12\mu - 1$ y autovalores $\mu = \frac{1}{2} \pm \frac{\sqrt{3}}{3}$, o sea 1,07735 y $-0{,}07735$. Sus recíprocos son 0,9282 y $-12{,}9282$, los mismos números de antes.

Ejemplo 32.67 (Rango uno, y la alternativa de Fredholm en carne y hueso)**.** (a) **Un núcleo de rango uno.** Sea $Kf(x) = \int_0^1 xt\,f(t)\,\mathrm{d}t$ y considérese $u - \lambda K u = 1$. El ansatz es $u(x) = 1 + \lambda c x$ con $c = \int_0^1 t u(t)\,\mathrm{d}t$; sustituyendo, $c = \frac{1}{2} + \frac{\lambda c}{3}$, de donde $c = \dfrac{3}{2(3-\lambda)}$ y $u(x) = 1 + \dfrac{3\lambda x}{2(3-\lambda)}$. Para $\lambda = 1$ queda $u(x) = 1 + \frac{3}{4}x$, y la comprobación es inmediata porque $\int_0^1 xt(1 + \frac{3}{4}t)\,\mathrm{d}t = \frac{x}{2} + \frac{x}{4} = \frac{3x}{4}$.

(b) **El valor excepcional.** La fórmula estalla en $\lambda = 3$, y ahí la ecuación homogénea $u = 3Ku$ tiene la solución no trivial $u(x) = x$, pues $\int_0^1 xt \cdot t\,\mathrm{d}t = \frac{x}{3}$. El núcleo es simétrico, así que $K^* = K$ y el problema adjunto tiene la misma solución.

(c) **Cuándo no hay solución.** Con $\lambda = 3$ y $f \equiv 1$, la condición de compatibilidad pide $\int_0^1 1 \cdot x\,\mathrm{d}x = 0$, pero esa integral vale $\frac{1}{2}$. Conclusión: la ecuación $u - 3Ku = 1$ *no tiene ninguna solución*, por muy inocente que parezca el miembro derecho.

(d) **Cuándo hay infinitas.** Con $\lambda = 3$ y $f(x) = x - \frac{2}{3}$ sí se cumple la condición, porque $\int_0^1 (x - \frac{2}{3})x\,\mathrm{d}x = \frac{1}{3} - \frac{1}{3} = 0$. Y efectivamente $u(x) = x - \frac{2}{3} + \gamma x$ resuelve la ecuación para todo $\gamma \in \mathbb{R}$: el conjunto de soluciones es una recta afín, paralela al núcleo de $I - 3K$. Es la dicotomía de siempre: solución única para casi todo λ, ninguna o infinitas en los valores excepcionales.

Cuidado

La alternativa de Fredholm es un enunciado sobre ecuaciones de *segunda* especie, del tipo $u - \lambda K u = f$, donde la identidad aparece explícitamente. Para las de primera especie, $Ku = f$, todo cambia: si K es compacto e inyectivo en dimensión infinita, su inversa nunca es acotada, y el problema es mal planteado en el sentido de Hadamard. Perturbaciones minúsculas de f producen cambios enormes en u, y por eso resolver $Ku = f$ numéricamente exige regularización en vez de eliminación gaussiana. La diferencia entre las dos especies no es un tecnicismo de notación.

Pausa para pensar

Un núcleo cualquiera de $L^2([0,1]^2)$ se aproxima en norma por núcleos degenerados: basta truncar su desarrollo en una base ortonormal del producto. Ese es el argumento por el que todo operador de Hilbert–Schmidt es límite en norma de operadores de rango finito, y por tanto compacto. ¿Qué dice esto al revés? Que la teoría de Fredholm general no es más que el paso al límite de lo que acabas de hacer con dos incógnitas, y que la razón por la que el espectro de un operador compacto sólo puede acumularse en 0 es que cada aproximación de rango finito aporta un número finito de autovalores grandes.

32.12. Convergencia débil y compacidad recuperada

Definición 32.68. $x_n \rightharpoonup x$ *débilmente* si $\varphi(x_n) \to \varphi(x)$ para todo $\varphi \in X^*$. En un Hilbert: $\langle x_n, y\rangle \to \langle x, y\rangle$ para todo y.

Ejemplo 32.69. En ℓ^2, la base canónica e_n cumple $e_n \rightharpoonup 0$ (para todo $y \in \ell^2$, $\langle e_n, y\rangle = y_n \to 0$ por ser $\sum |y_n|^2 < \infty$), pero $\|e_n\| = 1$ para todo n, así que *no* converge en norma. La convergencia débil no conserva la norma: solo garantiza $\|x\| \le \liminf \|x_n\|$ (semicontinuidad inferior débil). En $L^2[0, 2\pi]$, el mismo fenómeno es $\sin(nx) \rightharpoonup 0$: las oscilaciones se cancelan contra cualquier función fija, que es el lema de Riemann–Lebesgue.

Teorema 32.70 (Banach–Alaoglu; caso separable reflexivo)**.** *En un espacio de Banach reflexivo, toda sucesión acotada tiene una subsucesión débilmente convergente. En particular esto vale en todo espacio de Hilbert y en L^p con $1 < p < \infty$.*

Ejemplo 32.71 (Compacidad débil y el método directo)**. (a) Qué se gana.** En ℓ^2 la sucesión e_n está acotada y no tiene subsucesión convergente en norma; sí tiene una que converge débilmente, a 0. El teorema recupera para las sucesiones acotadas la propiedad que la compacidad les había quitado, a cambio de debilitar la noción de convergencia.

(b) El método directo del cálculo de variaciones. Para minimizar un funcional J, se toma una sucesión minimizante, se extrae una subsucesión débilmente convergente (aquí entra el teorema) y se usa que J es débilmente semicontinuo inferiormente para concluir que el límite es el mínimo. Así se demuestra la existencia de soluciones de muchas ecuaciones en derivadas parciales sin construirlas.

(c) La norma no es débilmente continua. Con $e_n \rightharpoonup 0$ se tiene $\|e_n\| = 1$ para todo n mientras que $\|0\| = 0$: la norma no pasa al límite. Lo que sí se conserva es la desigualdad $\|x\| \le \liminf \|x_n\|$, que es precisamente la semicontinuidad inferior que el método directo necesita.

(d) Sin reflexividad, falla. En c_0 (sucesiones que tienden a cero, con la norma del supremo), que no es reflexivo, la sucesión acotada $x_n = (1, 1, \dots, 1, 0, 0, \dots)$ con n unos no tiene ninguna subsucesión débilmente convergente en c_0: el candidato natural, la sucesión constante 1, no pertenece al espacio. En L^1 ocurre lo mismo, y por eso los problemas variacionales se plantean en L^p con $p > 1$.

Idea clave

El cierre del círculo. Riesz nos quitó la compacidad de la bola unidad; Alaoglu nos la devuelve, a cambio de debilitar el significado de "converger". Esa es la transacción central del análisis funcional moderno y el motor del *método directo del cálculo de variaciones*: para minimizar un funcional J, se toma una sucesión minimizante, se usa la acotación para extraer una subsucesión débilmente convergente, y se concluye con la semicontinuidad inferior débil de J que el límite es el mínimo. Así se demuestra la existencia de soluciones de infinidad de problemas de física e ingeniería sin resolverlos.

32.13. Formas bilineales, Lax–Milgram y el método de Galerkin

El teorema de representación de Riesz dice que todo funcional continuo sobre un Hilbert es un producto interno con un vector fijo. Si en vez del producto interno se admite una forma bilineal cualquiera que se le parezca lo suficiente, el resultado sobrevive; ese es el teorema de Lax–Milgram, y es la base matemática del método de elementos finitos con el que hoy se calculan puentes, alas y circuitos. Lo notable es que la demostración abstracta viene acompañada de un procedimiento numérico casi trivial de programar.

Definición 32.72 (Forma bilineal acotada y coerciva)**.** Sea H un espacio de Hilbert real. Una forma bilineal $a\colon H\times H\to\mathbb{R}$ es *acotada* si existe M con $|a(u,v)|\le M\|u\|\,\|v\|$ para todos u,v, y *coerciva* (o elíptica) si existe $\alpha>0$ con

$$a(u,u)\ge \alpha\|u\|^2 \qquad \text{para todo } u\in H.$$

No se exige simetría. Cuando a sí es simétrica, a es un producto interno equivalente al de H y la constante α mide cuánto se deforma la geometría.

Idea clave

Qué dice esto en cristiano. Una forma bilineal acotada es una que no explota: $|a(u,v)|\le M\|u\|\|v\|$, igual que un producto escalar. Coerciva significa que no se aplana: $a(u,u)\ge\alpha\|u\|^2$, es decir, que sobre la diagonal crece al menos como la norma al cuadrado. Juntas, las dos condiciones dicen que a se comporta "como un producto escalar deformado", sin llegar a ser simétrica. Y eso, según se verá enseguida, basta para resolver ecuaciones.

Ejemplo 32.73 (Formas acotadas y coercivas)**. (a) El producto escalar.** $a(u,v)=\langle u,v\rangle$ es acotada con $M=1$ (por Cauchy–Schwarz) y coerciva con $\alpha=1$. Es el caso trivial y el patrón de todos los demás.

(b) La forma del laplaciano. En $H^1_0(\Omega)$, la forma $a(u,v)=\int_\Omega \nabla u\cdot\nabla v$ es acotada, y es coerciva gracias a la desigualdad de Poincaré, que acota $\|u\|_{L^2}$ por $\|\nabla u\|_{L^2}$ cuando u se anula en la frontera. Sin la condición de frontera la coercividad se pierde, porque las constantes tienen gradiente nulo.

(c) Una no simétrica. $a(u,v)=\int(\nabla u\cdot\nabla v+b\cdot\nabla u\,v)$, con un término de transporte, no es simétrica y sigue siendo acotada y coerciva si $|b|$ es pequeño. Éste es el caso interesante: no hay producto escalar asociado y sin embargo todo funciona.

(d) Una que no es coerciva. En $\mathbb{R}^2$, la forma $a(u,v)=u_1v_1-u_2v_2$ es acotada y cumple $a(u,u)=u_1^2-u_2^2$, que es negativo para $u=(0,1)$. No hay ningún $\alpha>0$ posible, y en efecto la ecuación asociada (de tipo hiperbólico) no se resuelve con estas técnicas.

Teorema 32.74 (Lax–Milgram)**.** *Si a es bilineal, acotada y coerciva sobre un Hilbert H, entonces para cada funcional $\varphi\in H^*$ existe un único $u\in H$ con*

$$a(u,v)=\varphi(v)\qquad \text{para todo } v\in H,$$

y además $\|u\|\le\|\varphi\|/\alpha$.

Demostración. Por Riesz, $a(u,v)=\langle Au,v\rangle$ para un operador lineal acotado A con $\|A\|\le M$, y $\varphi(v)=\langle w,v\rangle$ para un único w. Hay que resolver $Au=w$. La coercividad da $\alpha\|u\|^2\le a(u,u)=\langle Au,u\rangle\le\|Au\|\,\|u\|$, luego $\|Au\|\ge\alpha\|u\|$: A es inyectivo y su imagen es cerrada. Si la imagen no fuera todo H, habría $v_0\neq0$ ortogonal a ella, y entonces $\alpha\|v_0\|^2\le a(v_0,v_0)=\langle Av_0,v_0\rangle=0$, contradicción. Así A es biyectivo y $u=A^{-1}w$ cumple $\|u\|\le\|w\|/\alpha=\|\varphi\|/\alpha$. □

Ejemplo 32.75 (Lax–Milgram resolviendo ecuaciones)**. (a) El caso simétrico.** Si además a es simétrica, el teorema se reduce al de representación de Riesz aplicado con el producto escalar $a(\cdot,\cdot)$, y la solución es el minimizador de la energía $\frac12 a(u,u)-\varphi(u)$. Lax–Milgram es la versión que sobrevive cuando la simetría desaparece y ya no hay energía que minimizar.

(b) El problema de Dirichlet. Para $-\Delta u=f$ en Ω con $u=0$ en el borde, la formulación débil es $\int\nabla u\cdot\nabla v=\int fv$ para toda $v\in H_0^1$. Con la forma del apartado anterior y $\varphi(v)=\int fv$, el teorema da existencia y unicidad de golpe, sin construir ninguna solución explícita.

(c) La estimación que viene de regalo. El teorema no sólo da la solución, sino la cota $\|u\|\le\|\varphi\|/\alpha$: la respuesta depende continuamente del dato. El problema está bien planteado en el sentido de Hadamard, y ésa es la parte que hace utilizable la teoría.

(d) El método de los elementos finitos. Aplicando Lax–Milgram en un subespacio de dimensión finita $V_h\subset H$ se obtiene un sistema lineal cuadrado con solución única, y el lema de Céa acota el error por $\frac{M}{\alpha}$ veces la mejor aproximación posible en V_h. Es decir: el método numérico es casi óptimo, y el cociente M/α mide cuánto se pierde. Toda la ingeniería de estructuras por ordenador descansa en este párrafo.

Idea clave

La ecuación $a(u,v)=\varphi(v)$ *para todo* v se llama formulación débil o variacional. Su gracia es que sólo pide una derivada de u en vez de dos, y que la restricción a un subespacio de dimensión finita produce automáticamente un sistema lineal cuadrado: si $V_h=\operatorname{span}\{\phi_1,\dots,\phi_n\}$, imponer la igualdad únicamente para $v=\phi_1,\dots,\phi_n$ da n ecuaciones con las n coordenadas de u_h como incógnitas. Eso es el método de Galerkin, y cuando las ϕ_i son funciones a trozos con soporte pequeño se llama método de los elementos finitos.

Regla: aproximar por Galerkin una ecuación en forma débil

1. **Multiplica por una función de prueba e integra por partes.** La ecuación fuerte $-(pu')' + qu = f$ con u nula en los extremos se convierte en $\int pu'v' + quv = \int fv$; el término de frontera desaparece porque v también se anula allí.
2. **Identifica a y φ, y comprueba las dos constantes.** Acotación con Cauchy–Schwarz, coercividad con la desigualdad de Poincaré si hace falta. Sin $\alpha > 0$ no hay teorema.
3. **Elige un subespacio de dimensión finita que respete la condición de contorno.** Cada función de base debe anularse donde lo hace la solución; si no, el método aproxima otro problema.
4. **Calcula la matriz de rigidez $A_{ij} = a(\phi_j, \phi_i)$ y el vector $F_i = \varphi(\phi_i)$.** Son integrales explícitas; con bases polinómicas se hacen a mano.
5. **Resuelve $Ac = F$.** Si a es coerciva, A es definida positiva y por tanto invertible, así que el sistema discreto *siempre* tiene solución única.
6. **Reconstruye $u_h = \sum_j c_j \phi_j$ y compárala con la solución exacta si la conoces.** El error se mide en la norma de energía, no puntualmente.
7. **Enriquece la base y observa que la energía baja.** Cuando a es simétrica, Galerkin minimiza $J(v) = \frac{1}{2}a(v,v) - \varphi(v)$ sobre V_h, así que ampliar el subespacio nunca empeora el resultado.

Ejemplo 32.76 (Un problema de contorno y su formulación débil)**.** (a) **El problema.** Sea $-u'' + u = f$ en $(0,1)$ con $u(0) = u(1) = 0$. Multiplicando por v con $v(0) = v(1) = 0$ e integrando por partes se llega a $a(u,v) = \int_0^1 (u'v' + uv)\,\mathrm{d}x = \int_0^1 fv\,\mathrm{d}x$.

(b) **Las constantes son inmejorables.** Aquí $a(u,u) = \int (u')^2 + u^2 = \|u\|^2$ exactamente, con la norma natural que mide a la vez la función y su derivada. Es decir $\alpha = M = 1$, y Lax–Milgram se reduce a Riesz puro: la forma *es* el producto interno.

(c) **Las soluciones exactas, para poder comparar.** Con $f \equiv 1$ la solución es $u(x) = 1 - \dfrac{\mathrm{e}^x + \mathrm{e}^{1-x}}{1+\mathrm{e}}$, que en $x = \frac{1}{2}$ vale 0,11318112. Con $f(x) = x$ la solución es $u(x) = x - \dfrac{\sinh x}{\sinh 1}$, que en $x = \frac{1}{2}$ vale 0,05659056, en $x = \frac{1}{4}$ vale 0,03504760 y en $x = \frac{3}{4}$ vale 0,05027579.

(d) **Qué se ha ganado.** La formulación débil tiene sentido para $f \in L^2$ aunque u no sea dos veces derivable en el sentido clásico, y la existencia queda garantizada por un teorema de tres líneas en vez de por una fórmula explícita. Cuando p y q son discontinuos —capas de materiales distintos— la formulación fuerte ni siquiera se puede escribir, y la débil sigue funcionando.

Ejemplo 32.77 (Galerkin con dos funciones de base, hecho a mano)**.** (a) **La base.** Tómense $\phi_1(x) = x(1-x)$ y $\phi_2(x) = x^2(1-x)$, ambas nulas en 0 y en 1. Las integrales dan $a(\phi_1,\phi_1) = \frac{11}{30}$, $a(\phi_1,\phi_2) = \frac{11}{60}$ y $a(\phi_2,\phi_2) = \frac{1}{7}$, de modo que la matriz de rigidez

es $A = \begin{pmatrix} 11/30 & 11/60 \\ 11/60 & 1/7 \end{pmatrix}$, con $\det A = \frac{473}{25200} > 0$: definida positiva, como anunciaba la coercividad.

(b) **El caso $f \equiv 1$ y una lección de simetría.** Aquí $F = (\frac{1}{6}, \frac{1}{12})$ y el sistema da $c = (\frac{5}{11}, 0)$: el segundo coeficiente se anula. No es un accidente, es que el problema es simétrico respecto de $x = \frac{1}{2}$ y ϕ_2 no lo es. Queda $u_h(x) = \frac{5}{11}x(1-x)$, con $u_h(\frac{1}{2}) = \frac{5}{44} = 0{,}11363636$ frente al valor exacto 0,11318112: un error de $4{,}6 \cdot 10^{-4}$ con dos funciones de base.

(c) **El caso $f(x) = x$, donde sí hacen falta las dos.** Ahora $F = (\frac{1}{12}, \frac{1}{20})$ y el sistema da $c_1 = \frac{69}{473} \approx 0{,}145877$, $c_2 = \frac{7}{43} \approx 0{,}162791$, luego $u_h(x) = \frac{69}{473}x + \frac{8}{473}x^2 - \frac{7}{43}x^3$. En $x = \frac{1}{2}$ vale $\frac{5}{88} = 0{,}05681818$ contra el exacto 0,05659056; en $x = \frac{1}{4}$ da 0,03498282 contra 0,03504760, y en $x = \frac{3}{4}$ da 0,05024445 contra 0,05027579.

(d) **La energía baja al enriquecer.** Con la base entera, $J(u_h) = -\frac{24}{2365} = -0{,}01014799$; usando sólo ϕ_1 el coeficiente óptimo es $\frac{5}{22}$ y la energía sube a $-\frac{5}{528} = -0{,}00946970$. Como J se está minimizando, el valor más negativo es el mejor, y eso confirma numéricamente el paso 7 de la regla: añadir una función de base sólo puede ayudar.

Cuidado

Galerkin no interpola: u_h no coincide con u en ningún punto prefijado, y comparar valores puntuales, como se ha hecho arriba, es una comodidad expositiva, no la manera correcta de medir el error. El enunciado honesto es el lema de Céa: $\|u - u_h\| \le \frac{M}{\alpha} \inf_{v \in V_h} \|u - v\|$, es decir, la solución de Galerkin es tan buena como la mejor aproximación disponible en el subespacio, salvo el factor M/α. Si la forma es simétrica ese factor mejora a $\sqrt{M/\alpha}$, y en el ejemplo anterior, con $M = \alpha = 1$, u_h *es* exactamente la proyección ortogonal de u sobre V_h en la norma de energía.

Pausa para pensar

Si se pierde la coercividad el teorema desaparece, y con él la unicidad. Piénsese en $-u'' - \pi^2 u = f$ con las mismas condiciones de contorno: la forma $a(u,v) = \int u'v' - \pi^2 uv$ se anula en $u = v = \sin(\pi x)$, así que no puede ser coerciva, y en efecto el problema homogéneo tiene soluciones no triviales. ¿Reconoces la situación? Es la alternativa de Fredholm de la sección anterior, vista desde el lado variacional: el operador es identidad más compacto, y en los valores excepcionales del parámetro la existencia se paga con una condición de ortogonalidad sobre f.

32.14. Taller de ejercicios resueltos

Ejercicio 32.1. Demuestra que $\ell^1 \subsetneq \ell^2 \subsetneq \ell^\infty$ y que las inclusiones son continuas con norma 1.

Solución. Si $\sum |x_n| < \infty$ entonces $|x_n| \le \|x\|_1$ para todo n, luego $\sum |x_n|^2 \le \|x\|_1 \sum |x_n| = \|x\|_1^2$ y $\|x\|_2 \le \|x\|_1$. Si $\sum |x_n|^2 < \infty$ entonces $|x_n|^2 \le \|x\|_2^2$, luego $\|x\|_\infty \le \|x\|_2$. Ambas normas de inclusión valen 1 (basta evaluar en e_1). Las inclusiones son estrictas: $x_n = 1/n$ está en ℓ^2 y no en ℓ^1; $x_n = 1$

está en ℓ^∞ y no en ℓ^2. Nota que para *medidas finitas* ocurre lo contrario: $L^2[0,1] \subseteq L^1[0,1]$, por Cauchy–Schwarz contra la función 1. ■

Ejercicio 32.2. Sea $\varphi(f) = \int_0^1 f(x)\,\mathrm{d}x$ sobre $C[0,1]$ con la norma del supremo. Calcula $\|\varphi\|$. Ídem para $\psi(f) = f(1/2)$.

Solución. $|\varphi(f)| \le \int_0^1 |f| \le \|f\|_\infty$, luego $\|\varphi\| \le 1$; y con $f \equiv 1$ se alcanza la igualdad, así que $\|\varphi\| = 1$. Igual $\|\psi\| = 1$, alcanzado también en $f \equiv 1$. Ambos son continuos. En cambio, sobre $C[0,1]$ con la norma $\|\cdot\|_1$, el funcional ψ *no* es continuo: las funciones triangulares de altura 1, base $2/n$ y pico en $1/2$ tienen $\|f_n\|_1 = 1/n \to 0$ pero $\psi(f_n) = 1$. ■

Ejercicio 32.3. En $L^2[-1,1]$, ortonormaliza $\{1, x, x^2\}$ por Gram–Schmidt.

Solución. $\|1\|^2 = 2$, luego $e_0 = 1/\sqrt{2}$. Como $\langle x, 1\rangle = 0$, x ya es ortogonal; $\|x\|^2 = 2/3$, luego $e_1 = \sqrt{3/2}\,x$. Para x^2: $\langle x^2, e_0\rangle = \frac{1}{\sqrt{2}} \cdot \frac{2}{3} = \frac{\sqrt{2}}{3}$ y $\langle x^2, e_1\rangle = 0$ por paridad, así que $u = x^2 - \frac{1}{3}$, con $\|u\|^2 = \int_{-1}^1 (x^2 - \frac{1}{3})^2 = \frac{8}{45}$ y $e_2 = \sqrt{\frac{45}{8}}\,(x^2 - \frac{1}{3})$. Estos son los polinomios de Legendre normalizados; su existencia como base ortonormal de $L^2[-1,1]$ es Stone–Weierstrass otra vez. ■

Ejercicio 32.4. Demuestra que si T es compacto y S acotado, entonces TS y ST son compactos. Deduce que la identidad en dimensión infinita no es compacta.

Solución. Si (x_n) es acotada, (Sx_n) también lo es (por continuidad de S), luego $T(Sx_n)$ tiene subsucesión convergente: TS es compacto. Para ST: extrae primero $Tx_{n_k} \to y$ y aplica S, que es continua, para obtener $STx_{n_k} \to Sy$. Si la identidad fuera compacta, la bola unidad tendría clausura compacta, contra el teorema de Riesz. Corolario: los operadores compactos forman un ideal bilátero cerrado de $B(H)$, y en dimensión infinita ningún compacto es invertible —si T compacto y T^{-1} acotado existieran, $I = T^{-1}T$ sería compacto. ■

Ejercicio 32.5. Sea H de Hilbert y M un subespacio *no* cerrado. Da un ejemplo y muestra que el teorema de proyección falla.

Solución. En ℓ^2, sea M el conjunto de sucesiones con finitos términos no nulos: es subespacio, es denso (luego $M^\perp = \{0\}$) y no es cerrado ($M \neq \ell^2$). Toma $x = (1, \frac{1}{2}, \frac{1}{3}, \dots) \in \ell^2 \setminus M$. Entonces $\operatorname{dist}(x, M) = 0$ por densidad, pero no existe $p \in M$ con $\|x - p\| = 0$, porque eso obligaría a $p = x \notin M$. El ínfimo no se alcanza: la hipótesis "cerrado" del teorema de proyección no es decorativa. Además, aquí $M \oplus M^\perp = M \neq \ell^2$: la descomposición ortogonal también falla. ■

Ejercicio 32.6. Usa Parseval con la función $f(x) = |x|$ en $[-\pi, \pi]$ para calcular $\sum_{k\ge 0} \frac{1}{(2k+1)^4}$.

Solución. f es par, así que solo hay cosenos: $a_0 = \frac{1}{\pi}\int_{-\pi}^{\pi} |x|\mathrm{d}x = \pi$, y para $n \ge 1$,

$$a_n = \frac{1}{\pi}\int_{-\pi}^{\pi} |x|\cos(nx)\mathrm{d}x = \frac{2}{\pi}\int_0^{\pi} x\cos(nx)\mathrm{d}x = \frac{2(\cos(n\pi) - 1)}{\pi n^2} = \begin{cases} -\frac{4}{\pi n^2}, & n \text{ impar}, \\ 0, & n \text{ par}. \end{cases}$$

Parseval en forma real, $\frac{1}{\pi}\int_{-\pi}^{\pi} f^2 = \frac{a_0^2}{2} + \sum_{n\ge 1} a_n^2$, con $\frac{1}{\pi}\int_{-\pi}^{\pi} x^2 = \frac{2\pi^2}{3}$, da

$$\frac{2\pi^2}{3} = \frac{\pi^2}{2} + \frac{16}{\pi^2}\sum_{k\ge 0}\frac{1}{(2k+1)^4} \implies \sum_{k\ge 0}\frac{1}{(2k+1)^4} = \frac{\pi^2}{16}\Big(\frac{2\pi^2}{3} - \frac{\pi^2}{2}\Big) = \frac{\pi^4}{96}.$$

Comprobación cruzada: $\sum_{n\ge 1} n^{-4} = \pi^4/90$ y los pares aportan $\frac{1}{16}\cdot\frac{\pi^4}{90}$, así que los impares suman $\frac{\pi^4}{90}(1 - \frac{1}{16}) = \frac{\pi^4}{90}\cdot\frac{15}{16} = \frac{\pi^4}{96}$. Coincide. ■

Ejercicio 32.7. Demuestra que en un Hilbert, si $x_n \rightharpoonup x$ y $\|x_n\| \to \|x\|$, entonces $x_n \to x$ en norma.

Solución.

$$\|x_n - x\|^2 = \|x_n\|^2 - 2\,\mathrm{Re}\langle x_n, x\rangle + \|x\|^2.$$

El primer término tiende a $\|x\|^2$ por hipótesis, y $\langle x_n, x\rangle \to \langle x, x\rangle = \|x\|^2$ por convergencia débil. Luego el total tiende a $\|x\|^2 - 2\|x\|^2 + \|x\|^2 = 0$. (Esta propiedad se llama *propiedad de Radon–Riesz* y es una particularidad de los Hilbert y de los L^p con $1 < p < \infty$; falla en L^1.) ■

Ejercicio 32.8. Sea $T : \ell^2 \to \ell^2$ el operador *desplazamiento* $T(x_1, x_2, \dots) = (0, x_1, x_2, \dots)$. Calcula $\|T\|$, su adjunto, y muestra que T no tiene autovalores mientras que T^* tiene infinitos.

Solución. $\|Tx\| = \|x\|$, luego $\|T\| = 1$ y T es una isometría, pero no es sobreyectiva: su imagen es el subespacio cerrado $\{x_1 = 0\}$. El adjunto es el desplazamiento a la izquierda $T^*(x_1, x_2, \dots) = (x_2, x_3, \dots)$, pues $\langle Tx, y\rangle = \sum_{n\ge 1} x_n \overline{y_{n+1}} = \langle x, T^* y\rangle$. Si $Tx = \lambda x$, comparando coordenadas se obtiene $\lambda x_1 = 0$ y $x_n = \lambda x_{n+1}$, lo que fuerza $x = 0$: sin autovalores. En cambio, para todo $|\lambda| < 1$ el vector $x = (1, \lambda, \lambda^2, \dots) \in \ell^2$ cumple $T^* x = \lambda x$: todo el disco abierto es espectro puntual de T^*. Un operador y su adjunto tienen la misma norma y espectros muy distintos; en dimensión finita esto sería imposible, porque A y $A^{\mathsf T}$ tienen el mismo polinomio característico. ■

Ejercicio 32.9. Sea $(Vf)(x) = \int_0^x f$. Demuestra que $\|V^n\| \le \frac{1}{n!}$ en $C[0,1]$ y deduce que $I - V$ es invertible.

Solución. Por inducción, $|V^n f(x)| \le \|f\|_\infty \frac{x^n}{n!}$: para $n = 1$ es directo, y

$$|V^{n+1} f(x)| \le \int_0^x |V^n f(t)|\mathrm{d}t \le \|f\|_\infty \int_0^x \frac{t^n}{n!}\mathrm{d}t = \|f\|_\infty \frac{x^{n+1}}{(n+1)!}.$$

Luego $\|V^n\| \le 1/n!$ y la serie de Neumann $\sum_{n\ge 0} V^n$ converge absolutamente en $B(C[0,1])$, que es Banach. Su suma es la inversa de $I - V$, porque $(I - V)\sum_{n\le N} V^n = I - V^{N+1} \to I$. Esto resuelve la ecuación integral $u(x) - \int_0^x u = g(x)$ para toda g continua, y es exactamente el argumento de Picard del capítulo 12 escrito en lenguaje de operadores: la serie de Neumann *es* la iteración de Picard. ■

32.15. Cierre: qué llevarte de este capítulo

Cinco frases para retener.

Uno. Un espacio de funciones es un espacio vectorial con una norma; que sea completo o no depende de la norma elegida, y la completitud es lo que permite construir límites.

Dos. En dimensión infinita la bola unidad no es compacta (Riesz), y casi toda la técnica del área consiste en recuperar la compacidad perdida: por equicontinuidad, por operadores compactos, o por convergencia débil.

Tres. En espacios de Hilbert vuelve la geometría: hay perpendiculares, proyecciones y bases ortonormales, y con ellas Fourier, mínimos cuadrados y el desarrollo en autofunciones son un solo teorema.

Cuatro. Los cuatro teoremas grandes son máquinas de existencia: Hahn–Banach garantiza que hay funcionales suficientes, y Steinhaus, la aplicación abierta y la gráfica cerrada convierten completitud en continuidad, todos vía Baire.

Cinco. El teorema espectral para operadores compactos autoadjuntos es el motivo por el que separar variables funciona: convierte una ecuación diferencial en una suma de problemas unidimensionales, igual que diagonalizar una matriz simétrica en el capítulo 7.

32.16. Espectro y resolvente: las definiciones y la fórmula de Gelfand

El capítulo ha calculado espectros con reglas enunciadas sin justificar. Se dan aquí las definiciones y las demostraciones.

Definición 32.78 (Resolvente y espectro)**.** Sea X un espacio de Banach complejo y $T \in \mathcal{B}(X)$ un operador lineal acotado. El *conjunto resolvente* es

$$\rho(T) = \{\lambda \in \mathbb{C} : \ \lambda I - T \text{ es biyectivo con inverso acotado}\},$$

y $R(\lambda, T) = (\lambda I - T)^{-1}$ se llama *resolvente*. El *espectro* es $\sigma(T) = \mathbb{C} \setminus \rho(T)$, y se descompone en:

- *espectro puntual* σ_p: los λ para los que $\lambda I - T$ no es inyectivo (autovalores);
- *espectro continuo* σ_c: inyectivo, imagen densa pero no cerrada;
- *espectro residual* σ_r: inyectivo con imagen no densa.

Observación 32.79. En dimensión finita sólo existe σ_p: un operador inyectivo entre espacios de la misma dimensión finita es sobreyectivo. Toda la riqueza del espectro es un fenómeno de dimensión infinita, y por eso la teoría espectral de operadores no es álgebra lineal repetida.

Lema 32.80 (Serie de Neumann)**.** *Si* $\|S\| < 1$ *entonces* $I - S$ *es invertible, con* $(I-S)^{-1} = \sum_{n\geq 0} S^n$ *y* $\|(I-S)^{-1}\| \leq \frac{1}{1-\|S\|}$.

Demostración. La serie $\sum \|S^n\| \leq \sum \|S\|^n$ converge, y $\mathcal{B}(X)$ es completo por serlo X, luego la serie converge absolutamente a un operador A. Multiplicando término a término, $(I - S)\sum_{n=0}^{N} S^n = I - S^{N+1} \to I$, y lo mismo por el otro lado. □

Teorema 32.81 (El espectro es compacto y no vacío)**.** *Para todo* $T \in \mathcal{B}(X)$ *con* $X \neq \{0\}$ *complejo,* $\sigma(T)$ *es un subconjunto compacto y no vacío de* $\{|\lambda| \leq \|T\|\}$.

Demostración. Acotado. Si $|\lambda| > \|T\|$, entonces $\lambda I - T = \lambda(I - \frac{T}{\lambda})$ con $\|T/\lambda\| < 1$, invertible por el lema 32.80. Luego $\sigma(T) \subset \{|\lambda| \leq \|T\|\}$.

Cerrado. Sea $\lambda_0 \in \rho(T)$ y $|\lambda - \lambda_0| < \|R(\lambda_0, T)\|^{-1}$. Entonces

$$\lambda I - T = (\lambda_0 I - T)\Big(I + (\lambda - \lambda_0)R(\lambda_0, T)\Big),$$

y el segundo factor es invertible por Neumann. Así $\rho(T)$ es abierto.

No vacío. Supóngase $\sigma(T) = \emptyset$, de modo que $\lambda \mapsto R(\lambda, T)$ está definida en todo $\mathbb{C}$. La identidad del resolvente $R(\lambda) - R(\mu) = (\mu - \lambda)R(\lambda)R(\mu)$ muestra que es holomorfa (derivable en norma). Para $|\lambda| > \|T\|$, la serie de Neumann da $\|R(\lambda, T)\| \leq \frac{1}{|\lambda| - \|T\|} \to 0$. Sea ahora φ un funcional continuo cualquiera sobre $\mathcal{B}(X)$: la función escalar $\lambda \mapsto \varphi(R(\lambda, T))$ es entera y tiende a 0 en el infinito, luego es idénticamente nula por Liouville (capítulo 30). Como esto vale para todo φ, el teorema de Hahn–Banach fuerza $R(\lambda, T) = 0$ para todo λ, lo cual es absurdo porque un inverso no puede ser el operador nulo. □

Teorema 32.82 (Fórmula del radio espectral, Gelfand)**.** *Sea* $r(T) = \max\{|\lambda| : \lambda \in \sigma(T)\}$. *Entonces*

$$r(T) = \lim_{n\to\infty} \|T^n\|^{1/n} = \inf_{n\geq 1} \|T^n\|^{1/n} \;\leq\; \|T\|.$$

Idea clave

Las dos desigualdades. Para $\geq$: si $\lambda \in \sigma(T)$, la factorización $\lambda^n I - T^n = (\lambda I - T)(\lambda^{n-1} I + \cdots + T^{n-1})$ muestra que $\lambda^n \in \sigma(T^n)$, luego $|\lambda|^n \leq \|T^n\|$ por la parte ya probada, es decir $|\lambda| \leq \|T^n\|^{1/n}$ para todo n. Para $\leq$: la serie de Laurent $R(\lambda, T) = \sum_{n\geq 0} \lambda^{-n-1} T^n$ converge para $|\lambda| > r(T)$, porque el resolvente es holomorfo fuera del espectro y el radio de convergencia de una serie de Laurent llega hasta la singularidad más próxima; por la fórmula de Cauchy–Hadamard, $\limsup \|T^n\|^{1/n} \leq r(T)$. El lema de subaditividad de Fekete aplicado a $a_n = \log \|T^n\|$ (que cumple $a_{n+m} \leq a_n + a_m$) garantiza que el límite existe y coincide con el ínfimo.

Ejemplo 32.83 (Radio espectral cero con norma grande)**.** Sea V el operador de Volterra en $L^2(0,1)$, $(Vf)(x) = \int_0^x f$. Se comprueba por inducción que $\|V^n\| \leq \frac{1}{(n-1)!}$, de donde $\|V^n\|^{1/n} \to 0$ y $r(V) = 0$: el espectro es $\{0\}$ y V no tiene ningún autovalor. Sin embargo $\|V\| = 2/\pi \approx 0{,}6366$, muy lejos de cero. Un operador así se llama *cuasinilpotente*, y no existe en dimensión finita, donde radio espectral cero implica nilpotencia. Es el ejemplo que muestra por qué Gelfand no es una trivialidad: la sucesión $\|T^n\|^{1/n}$ puede decrecer indefinidamente aunque $\|T\|$ no sea pequeño.

Proposición 32.84 (Espectro del adjunto)**.** *Para* $T \in \mathcal{B}(H)$ *en un Hilbert,* $\sigma(T^*) = \overline{\sigma(T)} = \{\bar\lambda : \lambda \in \sigma(T)\}$.

Demostración. $(\lambda I - T)^* = \bar\lambda I - T^*$, y un operador es invertible si y sólo si lo es su adjunto (el inverso del adjunto es el adjunto del inverso). Por tanto $\lambda \in \rho(T)$ equivale a $\bar\lambda \in \rho(T^*)$. □

Corolario 32.85 (Autoadjuntos y unitarios)**.** *Si* $T = T^*$*, entonces* $\sigma(T) \subset \mathbb{R}$ *y* $r(T) = \|T\|$. *Si* U *es unitario, entonces* $\sigma(U)$ *está contenido en la circunferencia unidad.*

Demostración. Para el primero, si $\lambda = a + b\mathrm{i}$ con $b \neq 0$, entonces para todo x

$$\|(\lambda I - T)x\|\,\|x\| \;\geq\; \big|\langle (\lambda I - T)x, x\rangle\big| \;\geq\; \big|\mathrm{Im}\langle (\lambda I - T)x, x\rangle\big| = |b|\,\|x\|^2,$$

usando que $\langle Tx, x\rangle$ es real por ser T autoadjunto; luego $\lambda I - T$ es acotado inferiormente, con imagen cerrada, y lo mismo vale para $\bar\lambda I - T$, cuya imagen es densa por la proposición anterior: es biyectivo. Que $r(T) = \|T\|$ sale de $\|T^2\| = \|T\|^2$ para autoadjuntos (pues $\|T^2\| \geq \|T^2 x\| \geq \langle T^2 x, x\rangle = \|Tx\|^2$ para $\|x\| = 1$), que iterada da $\|T^{2^k}\| = \|T\|^{2^k}$ y por Gelfand $r(T) = \|T\|$. Para U unitario, $\|U\| = \|U^{-1}\| = 1$, luego $\sigma(U)$ y $\sigma(U^{-1}) = \{\lambda^{-1}\}$ están ambos en el disco unidad cerrado, lo que obliga a $|\lambda| = 1$. □

32.17. Espacios de Sobolev, derivada débil y la topología débil-∗

Dos nociones se han usado en los ejemplos de Lax–Milgram y de Galerkin, y en el enunciado de Banach–Alaoglu, sin haberse definido. Aquí están.

Definición 32.86 (Derivada débil)**.** Sea $\Omega \subset \mathbb{R}^n$ abierto y $u \in L^1_{\mathrm{loc}}(\Omega)$. Se dice que $v \in L^1_{\mathrm{loc}}(\Omega)$ es la *derivada débil* $\partial_i u$ si

$$\int_\Omega u\, \partial_i \varphi \,\mathrm{d}x = -\int_\Omega v\, \varphi \,\mathrm{d}x \qquad \text{para toda } \varphi \in C_c^\infty(\Omega).$$

La definición no es un capricho: es la fórmula de integración por partes, con el término de frontera desaparecido porque φ tiene soporte compacto. Si u es de clase C^1, su derivada débil coincide con la clásica; la ventaja es que ahora hay funciones no derivables que sí tienen derivada débil.

Proposición 32.87 (Unicidad)**.** *La derivada débil, si existe, es única en casi todo punto.*

Demostración. Si v_1, v_2 sirven, entonces $\int (v_1 - v_2)\varphi = 0$ para toda $\varphi \in C_c^\infty$. El lema fundamental del cálculo de variaciones —que se prueba regularizando por convolución con una aproximación de la identidad— da $v_1 = v_2$ c.t.p. □

Ejemplo 32.88 (Una función sin derivada, con derivada débil)**.** Sea $u(x) = |x|$ en $(-1,1)$. No es derivable en 0, pero para toda $\varphi \in C_c^\infty(-1,1)$,

$$\int_{-1}^{1} |x|\varphi'(x)\mathrm{d}x = \int_{-1}^{0} (-x)\varphi' + \int_0^1 x\varphi' = \int_{-1}^0 \varphi - \int_0^1 \varphi = -\int_{-1}^1 \operatorname{sgn}(x)\varphi(x)\mathrm{d}x,$$

integrando por partes en cada trozo (los términos de frontera se cancelan). Luego $u' = \operatorname{sgn}$ en sentido débil. En cambio sgn *no* tiene derivada débil en L^1_{loc}: su "derivada" es la delta de Dirac, que no es una función. Ahí está la frontera exacta de la teoría.

Definición 32.89 (Espacios de Sobolev)**.** Para $k \in \mathbb{N}$ y $1 \le p \le \infty$,

$$W^{k,p}(\Omega) = \{u \in L^p(\Omega) :\ \partial^\alpha u \in L^p(\Omega) \text{ en sentido débil para todo } |\alpha| \le k\},$$

con la norma $\|u\|_{W^{k,p}} = \left(\sum_{|\alpha|\le k} \|\partial^\alpha u\|_p^p\right)^{1/p}$. Se escribe $H^k = W^{k,2}$, que es un espacio de Hilbert con $\langle u, v\rangle_{H^k} = \sum_{|\alpha|\le k} \int \partial^\alpha u\, \partial^\alpha v$. Y $H^1_0(\Omega)$ es el cierre de $C_c^\infty(\Omega)$ en $H^1(\Omega)$: las funciones que "se anulan en la frontera" en sentido débil.

Teorema 32.90 ($W^{k,p}$ es completo)**.** *$W^{k,p}(\Omega)$ es un espacio de Banach, y H^k un Hilbert.*

Demostración. Sea (u_m) de Cauchy en $W^{k,p}$. Entonces cada $(\partial^\alpha u_m)$ es de Cauchy en L^p, que es completo, luego converge a algún $u_\alpha \in L^p$; en particular $u_m \to u_0 =: u$. Falta ver que u_α es la derivada débil $\partial^\alpha u$. Para $\varphi \in C_c^\infty$,

$$\int u\, \partial^\alpha \varphi = \lim_m \int u_m\, \partial^\alpha \varphi = \lim_m (-1)^{|\alpha|} \int \partial^\alpha u_m\, \varphi = (-1)^{|\alpha|} \int u_\alpha \varphi,$$

donde los pasos al límite se justifican por Hölder, ya que $u_m \to u$ y $\partial^\alpha u_m \to u_\alpha$ en L^p y $\varphi \in L^q$ por tener soporte compacto. □

Observación 32.91 (Qué se gana). Con estas definiciones, la formulación variacional del capítulo 13 deja de ser un formalismo. La ecuación $-\Delta u = f$ con $u = 0$ en la frontera se escribe como: hallar $u \in H^1_0(\Omega)$ con

$$\int_\Omega \nabla u \cdot \nabla v \, dx = \int_\Omega f v \, dx \qquad \text{para toda } v \in H^1_0,$$

que es exactamente $a(u,v) = L(v)$. La forma a es acotada por Cauchy–Schwarz, y coerciva por la desigualdad de Poincaré; Lax–Milgram da existencia y unicidad. Nótese lo que ha ocurrido: no se ha demostrado que exista una solución *clásica*, es decir dos veces derivable. Se ha demostrado que existe una solución débil, y la cuestión de si es de clase C^2 —la *regularidad elíptica*— es un problema aparte, con su propia teoría. Separar existencia de regularidad es el movimiento conceptual que hizo posible la teoría moderna de EDP.

Definición 32.92 (Topología débil y débil-$*$)**.** En un Banach X, la *topología débil* es la menos fina que hace continuos a todos los $\varphi \in X^*$; se escribe $x_n \rightharpoonup x$ si $\varphi(x_n) \to \varphi(x)$ para todo $\varphi \in X^*$. En el dual X^*, la *topología débil-$*$* es la menos fina que hace continuas a todas las evaluaciones $\varphi \mapsto \varphi(x)$ con $x \in X$; se escribe $\varphi_n \overset{*}{\rightharpoonup} \varphi$ si $\varphi_n(x) \to \varphi(x)$ para todo $x \in X$.

Cuidado

[Tres convergencias, no dos] En X^* conviven la convergencia en norma, la débil (contra elementos de X^{**}) y la débil-$*$ (contra elementos de X). Norma implica débil implica débil-$*$, y ninguna implicación se invierte en general; las dos débiles coinciden si y sólo si X es reflexivo. En $\ell^\infty = (\ell^1)^*$, la sucesión e_n de vectores de la base converge débil-$*$ a 0 pero no débilmente.

Teorema 32.93 (Banach–Alaoglu)**.** *La bola unidad cerrada de X^* es compacta en la topología débil-$*$.*

Demostración. Para cada $x \in X$ sea $D_x = \{z \in \mathbb{C} : |z| \le \|x\|\}$, un compacto. El producto $P = \prod_{x \in X} D_x$ es compacto en la topología producto, por el teorema de Tíjonov. La aplicación que envía φ a la familia $(\varphi(x))_{x \in X}$ identifica la bola unidad de X^* con un subconjunto de P, y la topología producto restringida a esa imagen es exactamente la topología débil-$*$ (ambas son la convergencia puntual en x). Falta ver que la imagen es cerrada en P: si (f_i) es una red en la imagen que converge puntualmente a $f \in P$, entonces f hereda la linealidad —porque $f(\alpha x + \beta y) = \lim f_i(\alpha x + \beta y) = \lim[\alpha f_i(x) + \beta f_i(y)] = \alpha f(x) + \beta f(y)$— y la cota $|f(x)| \le \|x\|$, luego es un funcional de norma ≤ 1. Un cerrado dentro de un compacto es compacto. □

Corolario 32.94 (Versión sucesional)**.** *Si X es separable, la bola unidad de X^* es* metrizable *en la topología débil-$*$, y por tanto secuencialmente compacta: toda sucesión acotada en X^* tiene una subsucesión convergente débil-$*$. Si X es reflexivo, toda sucesión acotada en X tiene una subsucesión débilmente convergente.*

Demostración. Sea $\{x_k\}$ denso en la bola de X. La métrica $d(\varphi,\psi)=\sum_k 2^{-k}\,\mín(1,|\varphi(x_k)-\psi(x_k)|)$ induce sobre la bola de X^* la topología débil-$*$: la convergencia en d equivale a la convergencia en cada x_k, y esa, por densidad y equiacotación, a la convergencia en todo x. Compacto más metrizable es secuencialmente compacto. La segunda afirmación es la primera aplicada a X^* con $X^{**}=X$. □

Idea clave

Para qué sirve todo esto. La compacidad es lo que permite demostrar que un problema de minimización tiene solución: se toma una sucesión minimizante, se extrae una subsucesión convergente y se pasa al límite. Pero en dimensión infinita las bolas *nunca* son compactas en norma —ese es el teorema de Riesz—, de modo que sin una topología más débil el método no arranca. Banach–Alaoglu proporciona la compacidad que falta, y el precio es que la convergencia es más débil: al pasar al límite hace falta que el funcional sea semicontinuo inferiormente para esa topología, lo cual, para funcionales convexos, es automático. Ahí está, en tres frases, el *método directo del cálculo de variaciones*, y con él se resuelven desde la existencia de geodésicas hasta la de soluciones de EDP no lineales.

32.18. El adjunto y las clases de operadores, definidos

Toda la teoría espectral del capítulo cuantifica sobre operadores autoadjuntos, unitarios y normales, y esas palabras sólo aparecían como pasos de una receta. Aquí están, con la existencia del adjunto demostrada.

Teorema 32.95 (Existencia y unicidad del adjunto)**.** *Sea H un espacio de Hilbert y $T\in\mathcal{B}(H)$. Existe un único operador $T^*\in\mathcal{B}(H)$ tal que*

$$\langle Tx,\,y\rangle=\langle x,\,T^*y\rangle\qquad \textit{para todos } x,y\in H,$$

y además $\|T^*\|=\|T\|$ y $\|T^*T\|=\|T\|^2$.

Demostración. Fíjese $y\in H$. La aplicación $x\mapsto\langle Tx,y\rangle$ es lineal y continua, con norma $\le\|T\|\,\|y\|$ por Cauchy–Schwarz. Por el teorema de representación de Riesz–Fréchet existe un único vector, que se llama T^*y, con

$$\langle Tx,y\rangle=\langle x,T^*y\rangle\quad\text{para todo } x,\qquad \|T^*y\|\le\|T\|\,\|y\|.$$

La unicidad en Riesz garantiza que $y\mapsto T^*y$ está bien definida, y su linealidad se sigue de la unicidad aplicada a combinaciones. La desigualdad anterior da $\|T^*\|\le\|T\|$; aplicándola a T^* y usando $T^{**}=T$ —que sale de tomar conjugados en la identidad definitoria— se obtiene la igualdad. Finalmente,

$$\|Tx\|^2=\langle Tx,Tx\rangle=\langle T^*Tx,x\rangle\le\|T^*T\|\,\|x\|^2,$$

luego $\|T\|^2\le\|T^*T\|$; y la desigualdad contraria es $\|T^*T\|\le\|T^*\|\|T\|=\|T\|^2$. □

Definición 32.96 (Las cinco clases de operadores)**.** Sea $T\in\mathcal{B}(H)$. Se dice que T es:

- *autoadjunto* (o hermítico) si $T^* = T$;
- *normal* si $T^*T = TT^*$;
- *unitario* si $T^*T = TT^* = I$, es decir, si es biyectivo con $T^{-1} = T^*$;
- una *isometría* si $T^*T = I$, equivalentemente si $\|Tx\| = \|x\|$ para todo x;
- *positivo*, y se escribe $T \geq 0$, si $\langle Tx, x\rangle \geq 0$ para todo x.

Observación 32.97 (Cómo encajan entre sí, y con los números). Autoadjuntos y unitarios son normales; positivo implica autoadjunto (en un Hilbert complejo, la condición $\langle Tx, x\rangle \in \mathbb{R}$ para todo x ya fuerza $T = T^*$, cosa que *no* ocurre en el caso real). Toda isometría sobreyectiva es unitaria, pero hay isometrías no unitarias en dimensión infinita: el desplazamiento $S(x_1, x_2, \dots) = (0, x_1, x_2, \dots)$ en ℓ^2 cumple $S^*S = I$ y $SS^* \neq I$, y ése es exactamente el fenómeno que no existe en dimensión finita, donde inyectivo implica sobreyectivo.

La analogía que organiza todo: $\mathcal{B}(H)$ es a $\mathbb{C}$ lo que la conjugación compleja es al adjunto. Los autoadjuntos son los "reales" (espectro en $\mathbb{R}$), los unitarios los de "módulo uno" (espectro en la circunferencia), los positivos los "≥ 0" (espectro en $[0, \infty)$), y los normales son aquellos para los que esa analogía funciona del todo: son precisamente los diagonalizables por el teorema espectral. Los no normales son los que no tienen análogo numérico, y por eso su teoría es incomparablemente más difícil.

32.19. Ejercicios resueltos

Ejercicio 32.10. Calcula la norma del funcional $\varphi(f) = \int_0^{1/2} f - \int_{1/2}^1 f$ sobre $(C[0,1], \|\cdot\|_\infty)$ y decide si se alcanza.

Solución. **(a) Cota superior.** $|\varphi(f)| \leq \int_0^1 |f| \leq \|f\|_\infty$, luego $\|\varphi\| \leq 1$.

(b) La cota es óptima. Tomemos la función continua f_ε que vale 1 en $[0, \frac{1}{2} - \varepsilon]$, -1 en $[\frac{1}{2} + \varepsilon, 1]$ y es lineal en medio. Entonces $\|f_\varepsilon\|_\infty = 1$ y

$$\varphi(f_\varepsilon) = \left(\tfrac{1}{2} - \varepsilon\right) + \left(\tfrac{1}{2} - \varepsilon\right) + 0 = 1 - 2\varepsilon \xrightarrow[\varepsilon \to 0]{} 1,$$

(el trozo lineal aporta 0 por simetría). Luego $\|\varphi\| = 1$.

(c) No se alcanza. Si $\varphi(f) = 1$ con $\|f\|_\infty = 1$, haría falta $f \equiv 1$ en $[0, \frac{1}{2}]$ y $f \equiv -1$ en $[\frac{1}{2}, 1]$ salvo en medida nula; por continuidad, $f(1/2)$ tendría que valer 1 y -1 a la vez. Imposible.

Este es el fenómeno que distingue $C[0,1]$ de un espacio reflexivo: hay funcionales de norma 1 que no alcanzan su norma en la bola unidad cerrada. El teorema de James lo convierte en criterio: un Banach es reflexivo si y solo si *todo* funcional continuo alcanza su norma. Aquí falla uno, luego $C[0,1]$ no es reflexivo. Compárese con un Hilbert, donde Riesz garantiza que todo funcional es $\langle \cdot, y\rangle$ y la norma se alcanza en $y/\|y\|$. ∎

Ejercicio 32.11. Sea $S : \ell^2 \to \ell^2$ el desplazamiento a la derecha, $S(x_1, x_2, \dots) = (0, x_1, x_2, \dots)$. Estudia isometría, sobreyectividad, adjunto, autovalores y espectro. Haz lo mismo con S^*.

Solución. **(a) Isometría, no sobreyectiva.** $\|Sx\|^2 = \sum |x_n|^2 = \|x\|^2$, así que S es isométrico e inyectivo. Pero su imagen es $\{y : y_1 = 0\}$, un subespacio cerrado propio: no es sobreyectivo. En

dimensión finita esto es imposible (inyectivo $\Rightarrow$ sobreyectivo); en dimensión infinita, la "habitación de Hilbert" siempre tiene sitio para uno más.

(b) Adjunto. De $\langle Sx, y\rangle = \sum_{n\geq 1} x_n\overline{y_{n+1}} = \langle x, S^*y\rangle$ sale $S^*(y_1, y_2, \dots) = (y_2, y_3, \dots)$: el desplazamiento a la izquierda. Es sobreyectivo pero no inyectivo (mata a e_1); exactamente el reverso de S.

(c) Autovalores. $Sx = \lambda x$ compara $(0, x_1, x_2, \dots)$ con $(\lambda x_1, \lambda x_2, \dots)$: la primera coordenada da $\lambda x_1 = 0$, y de ahí en cascada $x = 0$ si $\lambda \neq 0$; y si $\lambda = 0$ también. *S no tiene ningún autovalor.*

Pero S^* tiene muchísimos: para $|\lambda| < 1$, el vector $v_\lambda = (1, \lambda, \lambda^2, \dots)$ está en ℓ^2 (serie geométrica) y cumple $S^*v_\lambda = \lambda v_\lambda$. El disco abierto entero es espectro puntual de S^*.

(d) Espectro. El espectro es cerrado, y $\sigma(S) = \sigma(S^*)^-$ conjugado; como $\|S\| = 1$, está contenido en el disco unidad cerrado, y como contiene al disco abierto (por lo anterior y la relación $\sigma(S) = \overline{\sigma(S^*)}$), resulta

$$\sigma(S) = \sigma(S^*) = \{\lambda \in \mathbb{C} : |\lambda| \leq 1\}.$$

Todo el disco, sin un solo autovalor para S.

Esta es la lección central del capítulo: en dimensión infinita el espectro es mucho más rico que el conjunto de autovalores. Aquí S tiene espectro residual (el disco abierto) y S^* espectro puntual en el mismo sitio. Un operador *compacto* no puede hacer esto: su espectro no nulo es una sucesión de autovalores que tiende a 0. Y en efecto S no es compacto: manda la base ortonormal (e_n) en (e_{n+1}), que no tiene ninguna subsucesión convergente. ■

Ejercicio 32.12. Calcula $\|A\|$ como operador de $(\mathbb{R}^2, \|\cdot\|_2)$ en sí mismo para $A = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}$, y compárala con $\|A\|_1$, $\|A\|_\infty$ y la norma de Frobenius.

Solución. La norma de operador inducida por la euclídea es el mayor valor singular, es decir $\sqrt{\lambda_{\max}(A^\top A)}$. Aquí

$$A^\top A = \begin{pmatrix} 10 & 14 \\ 14 & 20 \end{pmatrix}, \qquad \lambda = \frac{30 \pm \sqrt{100 + 4\cdot 196}}{2} = 15 \pm \sqrt{221},$$

luego $\lambda_{\max} = 15 + \sqrt{221} = 29{,}866069$ y

$$\|A\|_2 = \sqrt{29{,}866069} = 5{,}464986.$$

El otro valor singular es $\sqrt{15 - \sqrt{221}} = 0{,}365966$, y el producto de ambos es $|\det A| = 2$: en efecto $5{,}464986 \cdot 0{,}365966 = 2{,}000000$. ✓

Las otras normas, todas fáciles de calcular a mano:

$$\|A\|_1 = \max_j \sum_i |a_{ij}| = 6 \text{ (columna 2)}, \qquad \|A\|_\infty = \max_i \sum_j |a_{ij}| = 7 \text{ (fila 2)},$$

$$\|A\|_F = \sqrt{1 + 4 + 9 + 16} = \sqrt{30} = 5{,}477226.$$

Obsérvese el encaje $\|A\|_2 \leq \|A\|_F$, que se cumple siempre porque $\|A\|_F^2$ es la suma de *todos* los cuadrados de valores singulares; aquí la diferencia es minúscula (5,465 frente a 5,477) porque el segundo valor singular es casi nulo: la matriz está cerca de ser de rango 1. Ese cociente $\sigma_1/\sigma_2 = 14{,}93$ es su número de condición, y avisa de que resolver $Ax = b$ con ella amplifica los errores relativos por un factor de hasta 15. ■

Ejercicio 32.13. En $L^2(-\pi, \pi)$ halla la mejor aproximación de $f(x) = x$ por elementos de $V = \operatorname{span}\{1, \cos x, \sin x\}$ y calcula el error.

Solución. V es un subespacio cerrado de dimensión 3 y la base $\{1, \cos x, \sin x\}$ es ortogonal, con $\|1\|^2 = 2\pi$ y $\|\cos\|^2 = \|\sin\|^2 = \pi$. La mejor aproximación es la proyección ortogonal, y sus coeficientes son cocientes de productos escalares:

$$\frac{\langle x, 1\rangle}{2\pi} = 0, \qquad \frac{\langle x, \cos x\rangle}{\pi} = 0, \qquad \frac{\langle x, \sin x\rangle}{\pi} = \frac{2\pi}{\pi} = 2.$$

Los dos primeros se anulan por paridad (x es impar, 1 y $\cos x$ son pares) sin necesidad de integrar; el tercero usa $\int_{-\pi}^{\pi} x \sin x \,\mathrm{d}x = 2\pi$. Luego

$$P_V f = 2\sin x.$$

Por Pitágoras, el error al cuadrado es $\|f\|^2 - \|P_V f\|^2$:

$$\|x\|^2 = \int_{-\pi}^{\pi} x^2 \mathrm{d}x = \frac{2\pi^3}{3} = 20{,}670851, \qquad \|2\sin x\|^2 = 4\pi = 12{,}566371,$$

$$\|f - P_V f\|^2 = 8{,}104481, \qquad \|f - P_V f\| = 2{,}846837.$$

Dos observaciones. Primera: $2\sin x$ es justo el primer término de la serie de Fourier de la onda en diente de sierra, cuyo desarrollo completo es $x = 2\sum_{n\ge 1}(-1)^{n+1}\sin(nx)/n$; la proyección sobre un subespacio de la base es simplemente truncar. Segunda: el error no tiende a cero con *este* V fijo, pero sí al aumentar los armónicos, porque el sistema trigonométrico es completo; y de la identidad de Parseval $\frac{2\pi^3}{3} = \pi\sum 4/n^2$ se recupera $\sum 1/n^2 = \pi^2/6$, lo cual sirve de comprobación independiente de todos los números anteriores. ■

Ejercicio 32.14. Demuestra que la base ortonormal (e_n) de ℓ^2 converge débilmente a 0 pero no en norma, y que por tanto la bola unidad cerrada no es compacta en norma.

Solución. **(a) Débilmente a cero.** Por el teorema de Riesz, todo funcional continuo en ℓ^2 es $\varphi(x) = \langle x, y\rangle$ para algún $y \in \ell^2$. Entonces $\varphi(e_n) = \overline{y_n}$, y como $\sum |y_n|^2 < \infty$, necesariamente $y_n \to 0$. Luego $\varphi(e_n) \to 0$ para todo φ: eso es exactamente $e_n \rightharpoonup 0$.

(b) No en norma. $\|e_n - 0\| = 1$ para todo n. Es más: $\|e_n - e_m\| = \sqrt{2}$ si $n \neq m$, así que la sucesión no tiene ninguna subsucesión de Cauchy.

(c) La bola no es compacta. (e_n) vive en la bola unidad cerrada y no tiene subsucesión convergente. Luego la bola no es secuencialmente compacta, y en un métrico eso equivale a no ser compacta.

Esto ilustra el teorema de Riesz: la bola unidad cerrada de un espacio normado es compacta si y solo si la dimensión es finita. La compacidad que se pierde se recupera en parte cambiando de topología: por Banach–Alaoglu la bola del *dual* siempre es compacta en la topología débil-$*$; y como ℓ^2 es reflexivo, su propia bola es débilmente compacta. En efecto, (e_n) sí converge en esa topología —a 0, como acabamos de ver—, aunque el límite tenga norma 0 y todos los términos norma 1: la norma no es débilmente continua, solo débilmente semicontinua inferiormente, y $0 = \|\lim\| \le \liminf \|e_n\| = 1$ lo confirma. ■

Ejercicio 32.15. Usa el principio de acotación uniforme para demostrar que si $\sum_n a_n b_n$ converge para todo $b \in \ell^2$, entonces $a \in \ell^2$.

Solución. Para cada N definimos $\varphi_N : \ell^2 \to \mathbb{C}$ por $\varphi_N(b) = \sum_{n\le N} a_n b_n$. Cada φ_N es lineal y continuo, con norma exacta $\|\varphi_N\| = \big(\sum_{n\le N} |a_n|^2\big)^{1/2}$ (Cauchy–Schwarz da la cota, y $b = (\bar a_1, \dots, \bar a_N, 0, \dots)$ la alcanza).

La hipótesis dice que para cada b fijo la sucesión numérica $\varphi_N(b)$ converge, luego está acotada: $\sup_N |\varphi_N(b)| < \infty$ para todo b. Eso es acotación puntual. Como ℓ^2 es completo, el principio de acotación uniforme (Banach–Steinhaus) da acotación en norma:

$$\sup_N \|\varphi_N\| = M < \infty \quad\Longrightarrow\quad \sum_{n\le N} |a_n|^2 \le M^2 \ \text{ para todo } N,$$

y haciendo $N \to \infty$, $\sum |a_n|^2 \le M^2$: $a \in \ell^2$.

Merece la pena ver dónde entra la completitud, porque es lo único que no es formal: sin ella el teorema es falso. Tómese el subespacio c_{00} de las sucesiones con soporte finito, con la norma de ℓ^2; los mismos φ_N están puntualmente acotados sobre él (cada b tiene soporte finito, así que $\varphi_N(b)$ es constante en N a partir de un momento) y sin embargo $\|\varphi_N\|$ puede tender a infinito eligiendo $a_n = 1$ para todo n. La categoría de Baire es la que separa un caso del otro. ■

32.20. Ejercicios propuestos

Sin solución desarrollada: solo enunciado y, al final, la respuesta. Varias respuestas son un nombre de teorema; si tu argumento llega al mismo sitio por otro camino, está bien igual.

32.1. ¿Es $\|x\|_\infty = \max_i |x_i|$ una norma en $\mathbb{R}^n$? ¿Proviene de un producto interno?

32.2. ¿Son equivalentes todas las normas en $\mathbb{R}^n$? ¿Y en un espacio de dimensión infinita?

32.3. ¿Es $C[0,1]$ con la norma del supremo un espacio de Banach? ¿Y con la norma L^1?

32.4. ¿Cuál de los espacios ℓ^p es de Hilbert? ¿Por qué solo ese?

32.5. Verifica la ley del paralelogramo en ℓ^2 y muestra que falla en ℓ^1 con $x = (1, 0, \dots)$, $y = (0, 1, 0, \dots)$.

32.6. Calcula la norma del funcional $\varphi(f) = \int_0^1 f$ sobre $(C[0,1], \|\cdot\|_\infty)$.

32.7. Enuncia Hahn–Banach y di qué garantiza sobre la abundancia de funcionales.

32.8. ¿Qué afirma el principio de acotación uniforme y sobre qué hipótesis descansa?

32.9. ¿Es sobreyectiva toda aplicación lineal continua e inyectiva entre Banach? ¿Qué teorema controla el caso sobreyectivo?

32.10. ¿Es la bola unidad cerrada de un espacio normado de dimensión infinita compacta en norma? ¿Y en la topología débil-$*$ del dual?

32.11. Sea $T : \ell^2 \to \ell^2$ el desplazamiento $T(x_1, x_2, \dots) = (0, x_1, x_2, \dots)$. ¿Es isométrico? ¿Es sobreyectivo? ¿Tiene autovalores?

32.12. ¿Qué dice el teorema espectral para operadores compactos autoadjuntos en un Hilbert?

Respuestas

1. Es norma sí; de un producto interno no (falla la ley del paralelogramo con $x = (1,0)$, $y = (0,1)$). **2.** En dimensión finita sí, siempre. En dimensión infinita no: en $C[0,1]$, $\|\cdot\|_1$ y $\|\cdot\|_\infty$ no lo son. **3.** Con el supremo sí es Banach; con la norma L^1 no (las sucesiones de Cauchy pueden salirse hacia funciones discontinuas). **4.** Solo ℓ^2: es el único p para el que se cumple la ley del paralelogramo, y por tanto el único con producto interno. **5.** En ℓ^1: $\|x+y\| = \|x-y\| = 2$, luego el lado izquierdo da 8, mientras que $2(\|x\|^2 + \|y\|^2) = 4$. No hay producto interno. **6.** $\|\varphi\| = 1$, alcanzada en la función constante 1. **7.** Todo funcional lineal acotado en un subespacio se extiende a todo el espacio con la misma norma. Consecuencia: el dual separa puntos, hay funcionales de sobra. **8.** Una familia de operadores acotados puntualmente en un Banach está acotada en norma. Descansa en la completitud, vía Baire. **9.** No: el desplazamiento del ejercicio 11 es inyectivo, continuo y no sobreyectivo. El caso sobreyectivo lo controla el teorema de la aplicación abierta (y, con ello, el de la inversa acotada). **10.** En norma no: la compacidad de la bola caracteriza la dimensión finita (Riesz). En débil-$*$ sí, por Banach–Alaoglu. **11.** Isométrico

sí ($\|Tx\| = \|x\|$); sobreyectivo no (nada llega a e_1); autovalores no tiene ninguno, aunque su espectro es todo el disco unidad cerrado. **12.** Que existe una base ortonormal de autovectores con autovalores reales que tienden a 0, y T se escribe $Tx = \sum \lambda_n \langle x, e_n\rangle e_n$.

Capítulo 33

Análisis armónico

Lo que necesitas traer

Del capítulo 31: la integral de Lebesgue y la transformada de Fourier. Del capítulo 32: espacios de Hilbert, bases ortonormales y operadores acotados. Del capítulo 4: la exponencial compleja, que es el ladrillo de todo el capítulo.

La idea central es una sola y es de las más fértiles de la matemática: cualquier función razonable se puede escribir como superposición de ondas puras. En un intervalo, las ondas forman una familia discreta y la superposición es una serie; en toda la recta, forman un continuo y la superposición es una integral. Ese cambio de punto de vista convierte derivadas en multiplicaciones y convoluciones en productos, y por eso el análisis armónico es la herramienta que resuelve las ecuaciones en derivadas parciales, procesa las señales y sostiene la mecánica cuántica.

33.1. Series de Fourier

Definición 33.1. Para f de cuadrado integrable en $[-\pi,\pi]$, su *serie de Fourier* es

$$f(x)\sim \frac{a_0}{2}+\sum_{n\geq 1}\left(a_n\cos nx+b_n\sin nx\right),\qquad a_n=\frac{1}{\pi}\int_{-\pi}^{\pi} f\cos nx\,\mathrm{d}x,\quad b_n=\frac{1}{\pi}\int_{-\pi}^{\pi} f\sin nx\,\mathrm{d}x.$$

La *identidad de Parseval* afirma que $\frac{1}{\pi}\int|f|^2=\frac{a_0^2}{2}+\sum(a_n^2+b_n^2)$.

Idea clave

Qué dice esto en cristiano. Los senos y cosenos forman una base ortonormal de L^2, y las fórmulas de los coeficientes no son más que el producto escalar con cada elemento de la base: exactamente lo mismo que proyectar un vector sobre los ejes coordenados, con la integral haciendo de producto escalar. Parseval es entonces el teorema de Pitágoras en dimensión infinita: la norma al cuadrado es la suma de los cuadrados de las coordenadas. Todo el capítulo es geometría en un espacio de Hilbert, y esa es la razón de que las series de Fourier funcionen tan bien y las de Taylor tan mal.

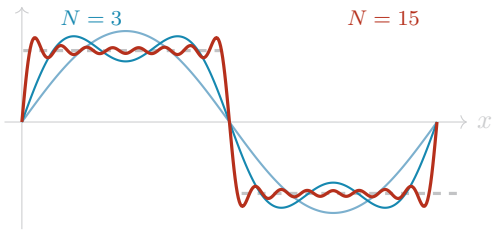


Figura 33.1: Sumas parciales de la serie de Fourier de la onda cuadrada. Al aumentar N la aproximación mejora en todo el interior, pero junto al salto aparece un pico que no se reduce: sobrepasa el valor correcto en un 8,9 por ciento por muchos términos que se sumen. Es el fenómeno de Gibbs, y la razón de que la convergencia sea puntual pero no uniforme.

Ejemplo 33.2 (La onda cuadrada y sus consecuencias)**. (a) Los coeficientes.** Para $f=1$ en $(0,\pi)$ y -1 en $(-\pi,0)$, los cosenos se anulan por ser f impar y $b_n=\frac{4}{n\pi}$ para n impar, cero para n par. Numéricamente: $b_1=1{,}273240$, $b_3=0{,}424413$, $b_5=0{,}254648$, $b_7=0{,}181891$.

(b) Parseval da una suma célebre. Como $|f|^2=1$, se tiene $\frac{1}{\pi}\int_{-\pi}^{\pi}1=2$, luego $\sum b_n^2=2$, es decir $\frac{16}{\pi^2}\sum_{n\text{ impar}}\frac{1}{n^2}=2$, de donde $\sum_{n\text{ impar}}\frac{1}{n^2}=\frac{\pi^2}{8}=1{,}233701$. Sumando los diez mil primeros impares se obtiene 1,233676: coincide.

(c) Otra suma, con otra función. Para $f(x)=x$ en $(-\pi,\pi)$ salen $b_n=\frac{2(-1)^{n+1}}{n}$, y Parseval da $\sum\frac{1}{n^2}=\frac{\pi^2}{6}=1{,}644934$, que es el problema de Basilea. Cada función produce una identidad numérica distinta, y ese es uno de los usos más elegantes de la teoría.

(d) El fenómeno de Gibbs. Cerca del salto, las sumas parciales sobrepasan el valor de la función en aproximadamente un 8,9 %, y ese exceso *no* desaparece al añadir términos: solo se estrecha. El límite del sobreimpulso es $\frac{2}{\pi}\int_0^{\pi}\frac{\sin t}{t}\,dt=1{,}178980$ veces la altura del salto. Es la razón de que comprimir imágenes con transformadas produzca halos alrededor de los bordes.

Cuidado

[Convergencia puntual, uniforme y en media son tres cosas distintas] La serie converge a f en la norma de L^2 siempre que f sea de cuadrado integrable, pero eso no garantiza convergencia puntual. Du Bois-Reymond construyó en 1873 una función continua cuya serie de Fourier diverge en un punto, y Kolmogórov en 1926 una función integrable cuya serie diverge en *todos* los puntos. Lo que sí es cierto, y costó hasta 1966 (teorema de Carleson), es que para $f\in L^2$ la serie converge en casi todo punto. Segundo aviso: en un salto la serie converge al promedio de los dos límites laterales, no a ninguno de los dos, así que redefinir f en el punto de salto no cambia nada.

33.2. El vocabulario de las ondas: periodicidad, paridad y ortogonalidad

La sección anterior ha soltado la fórmula de los coeficientes de Fourier como quien reparte cartas: ahí están a_n y b_n, con su $\frac{1}{\pi}$ delante y sus integrales entre $-\pi$ y π. Pero ¿de dónde salen? ¿Por qué el $\frac{a_0}{2}$ lleva ese dos dividiendo y los demás no? ¿Por qué las funciones pares no tienen senos? Nada de eso es un convenio arbitrario: todo se deduce de una única propiedad, la

ortogonalidad del sistema trigonométrico, que a su vez no es más que el teorema de Pitágoras trasplantado a un espacio de funciones. Esta sección construye ese vocabulario desde cero —periodo, paridad, producto interno, ortogonalidad—, deduce los coeficientes en lugar de postularlos y calcula dos series completas de principio a fin.

Funciones periódicas

Definición 33.3 (Función periódica y periodo)**.** Una función $f\colon \mathbb{R} \to \mathbb{C}$ es *periódica* de periodo $T > 0$ si

$$f(x+T) = f(x) \qquad \text{para todo } x \in \mathbb{R}.$$

Si existe el menor de todos los periodos positivos, se le llama *periodo fundamental*. La *frecuencia* es $1/T$ y la *frecuencia angular* es $\omega = 2\pi/T$.

Observación 33.4. De la definición se sigue de inmediato que si T es un periodo también lo son $2T$, $3T$ y en general kT con k entero positivo: basta aplicar la igualdad k veces. Por eso hablar de «el» periodo sin más es ambiguo, y por eso se distingue el fundamental. Una consecuencia práctica: para conocer una función periódica basta conocerla en *un* intervalo de longitud T, por ejemplo $[-T/2, T/2]$ o $[0, T]$; lo demás se obtiene repitiendo. Ese es el motivo de que todas las integrales de este capítulo se tomen sobre un intervalo de longitud un periodo, y de que dé exactamente igual cuál se elija.

Ejemplo 33.5 (Periodos que conviene tener contados)**. (a) Las funciones básicas.** $\sin x$ y $\cos x$ tienen periodo fundamental 2π. La función $\cos(nx)$ con n entero positivo tiene periodo fundamental $2\pi/n$, porque $\cos(n(x+2\pi/n)) = \cos(nx + 2\pi) = \cos(nx)$; pero *también* tiene a 2π por periodo, ya que $2\pi = n \cdot (2\pi/n)$. Este detalle es la clave de todo: *todas* las funciones $\cos(nx)$ y $\sin(nx)$, para todo $n \geq 1$, comparten el periodo 2π, y por eso cualquier combinación suya vuelve a ser 2π-periódica.

(b) Una constante. La función constante es periódica con cualquier periodo, y no tiene periodo fundamental. Es el término $n = 0$ de la familia, el de frecuencia cero.

(c) Una suma que no es periódica. $\sin x + \sin(\pi x)$ no es periódica: haría falta T múltiplo entero de 2π y de 2 a la vez, es decir, π racional. La periodicidad de una suma exige que los periodos sean conmensurables.

(d) Extensión periódica. Si f solo está dada en $[-\pi, \pi)$, se define su *extensión periódica* declarando $f(x + 2k\pi) = f(x)$. Así $f(x) = x$ en $[-\pi, \pi)$ se convierte en el diente de sierra, con saltos de altura 2π en los puntos $\pi + 2k\pi$. Toda serie de Fourier converge a la extensión periódica, nunca al trozo original: por eso aparecen saltos en los extremos aunque la función de partida fuese continua allí dentro.

Paridad

Definición 33.6 (Función par e impar)**.** Una función f es *par* si $f(-x) = f(x)$ para todo x, e *impar* si $f(-x) = -f(x)$.

Proposición 33.7 (Reglas de paridad)**.** *Sean f y g funciones integrables en $[-L, L]$. Entonces:*

(i) si f es impar, $\int_{-L}^{L} f = 0$; si f es par, $\int_{-L}^{L} f = 2\int_{0}^{L} f$;

(ii) par por par = par, impar por impar = par, par por impar = impar;

(iii) toda función se descompone de manera única como suma de una par y una impar,

$$f(x) = \underbrace{\frac{f(x)+f(-x)}{2}}_{\text{parte par}} + \underbrace{\frac{f(x)-f(-x)}{2}}_{\text{parte impar}}.$$

Demostración. (i) Pártase la integral y hágase el cambio $u = -x$ en el primer trozo: $\int_{-L}^{0} f(x)\mathrm{d}x = \int_{0}^{L} f(-u)\mathrm{d}u$. Si f es impar, eso vale $-\int_{0}^{L} f$ y cancela el segundo trozo; si es par, vale $+\int_{0}^{L} f$ y lo duplica.

(ii) Evalúese el producto en $-x$: por ejemplo, con f par y g impar, $f(-x)g(-x) = f(x)\cdot(-g(x)) = -f(x)g(x)$, que es la definición de impar. Los otros dos casos son idénticos.

(iii) La primera fracción es par y la segunda impar, como se comprueba sustituyendo x por $-x$, y su suma es $f(x)$. Para la unicidad, si $f = p_1 + i_1 = p_2 + i_2$ con p_j pares e i_j impares, entonces $p_1 - p_2 = i_2 - i_1$ es a la vez par e impar, luego $h(-x) = h(x) = -h(x)$ y por tanto $h = 0$. □

Idea clave

El apartado (i) es la razón de que la paridad ahorre la mitad del trabajo: convierte integrales que hay que calcular en ceros que se escriben directamente. Y el apartado (iii) explica por qué una serie de Fourier general necesita senos *y* cosenos: los cosenos construyen la parte par de la función y los senos la impar. Si la función ya es par, la mitad impar de la maquinaria sobra y todos los b_n salen nulos. No hay ningún milagro: hay una simetría.

Producto interno y ortogonalidad

Aquí está la idea que organiza el capítulo entero. En el plano, dos vectores son perpendiculares cuando su producto escalar es cero, y las coordenadas de un vector en una base ortonormal se calculan proyectando, es decir, haciendo el producto escalar con cada elemento de la base. Pues bien: las funciones admiten un producto escalar, los senos y cosenos son perpendiculares entre sí, y los coeficientes de Fourier son las coordenadas. Literalmente, no por analogía.

Definición 33.8 (Producto interno de funciones y sistema ortogonal)**.** Para funciones f, g de cuadrado integrable en un intervalo $[a, b]$ se define el *producto interno*

$$\langle f, g\rangle = \int_{a}^{b} f(x)\,\overline{g(x)}\,\mathrm{d}x,$$

(la barra es la conjugación compleja, que sobra si las funciones son reales) y la *norma* $\|f\| = \sqrt{\langle f, f\rangle} = \left(\int_a^b |f|^2\right)^{1/2}$. Dos funciones son *ortogonales* si $\langle f, g\rangle = 0$. Una familia $\{\varphi_0, \varphi_1, \varphi_2, \ldots\}$ es un *sistema ortogonal* si $\langle \varphi_n, \varphi_m\rangle = 0$ siempre que $n \neq m$ y ninguna φ_n es nula, y es *ortonormal* si además $\|\varphi_n\| = 1$ para todo n.

Regla: comprobar que un sistema de funciones es ortogonal y normalizarlo

1. **Fija el intervalo y el peso.** La ortogonalidad no es una propiedad de las funciones sino de la pareja «funciones más intervalo»: $\sin x$ y $\cos x$ son ortogonales en $[-\pi,\pi]$ y no lo son en $[0,\pi/2]$, donde $\int_0^{\pi/2}\sin x\cos x=\frac12$.
2. **Usa la paridad antes que la primitiva.** Si el intervalo es simétrico y el producto $\varphi_n\overline{\varphi_m}$ es impar, la integral es cero por la proposición 33.7 y no hay nada que calcular.
3. **Convierte productos en sumas.** Para senos y cosenos, las identidades $2\cos A\cos B=\cos(A-B)+\cos(A+B)$, $2\sin A\sin B=\cos(A-B)-\cos(A+B)$ y $2\sin A\cos B=\sin(A+B)+\sin(A-B)$ transforman cada producto en dos términos cuya integral sobre un periodo completo es visiblemente nula.
4. **Trata aparte el caso** $n=m$, que nunca da cero: es el que produce la norma.
5. **Calcula** $\|\varphi_n\|^2=\langle\varphi_n,\varphi_n\rangle$ y normaliza dividiendo: $\tilde\varphi_n=\varphi_n/\|\varphi_n\|$. El sistema ortonormal es el mismo sistema con las escalas ajustadas.
6. **Comprueba el resultado en un caso pequeño**, integrando numéricamente con $n=1$ y $m=2$; un signo mal puesto se detecta al instante.
7. **No confundas ortogonal con independiente ni con completo.** Ortogonal implica linealmente independiente, pero que un sistema ortogonal permita reconstruir *toda* función —que sea *completo*— es un teorema aparte, y para el sistema trigonométrico es justo lo que garantiza el teorema de Fejér de la sección correspondiente.

Teorema 33.9 (Relaciones de ortogonalidad del sistema trigonométrico)**.** *Para enteros $n,m\ge1$ se cumple, sobre el intervalo $[-\pi,\pi]$:*

$$\int_{-\pi}^{\pi}\cos nx\,\cos mx\,\mathrm{d}x=\begin{cases}0, & n\neq m\\ \pi, & n=m\end{cases}\qquad \int_{-\pi}^{\pi}\sin nx\,\sin mx\,\mathrm{d}x=\begin{cases}0, & n\neq m\\ \pi, & n=m\end{cases}$$

$$\int_{-\pi}^{\pi}\sin nx\,\cos mx\,\mathrm{d}x=0\quad\text{para todos } n,m\ge1,\qquad \int_{-\pi}^{\pi}\cos nx\,\mathrm{d}x=\int_{-\pi}^{\pi}\sin nx\,\mathrm{d}x=0,$$

y $\int_{-\pi}^{\pi}1\cdot1\,\mathrm{d}x=2\pi$. Es decir, la familia $\{1,\cos x,\sin x,\cos 2x,\sin 2x,\dots\}$ es un sistema ortogonal en $[-\pi,\pi]$, y al dividir cada elemento por su norma se obtiene el sistema ortonormal

$$\left\{\frac{1}{\sqrt{2\pi}},\ \frac{\cos nx}{\sqrt{\pi}},\ \frac{\sin nx}{\sqrt{\pi}}\right\}_{n\ge1}.$$

Demostración. La última fila es inmediata: una primitiva de $\cos nx$ es $\sin(nx)/n$, que vale lo mismo en $-\pi$ y en π por ser sin impar y $n\pi$ múltiplo de π; y $\sin nx$ es impar, luego su integral sobre un intervalo simétrico es nula por la proposición 33.7.

Para la tercera, obsérvese que $\sin nx\,\cos mx$ es impar (impar por par), así que su integral sobre $[-\pi,\pi]$ es cero sin necesidad de calcular nada.

Para la primera, aplíquese $2\cos A\cos B=\cos(A-B)+\cos(A+B)$ con $A=nx$ y $B=mx$:

$$\int_{-\pi}^{\pi}\cos nx\cos mx\,\mathrm{d}x=\frac{1}{2}\int_{-\pi}^{\pi}\cos((n-m)x)\mathrm{d}x+\frac{1}{2}\int_{-\pi}^{\pi}\cos((n+m)x)\mathrm{d}x.$$

Si $n\neq m$, ambos enteros $n-m$ y $n+m$ son no nulos y las dos integrales se anulan por la última fila; si $n=m$, el primer integrando es $\cos 0=1$ y aporta $\frac{1}{2}\cdot 2\pi=\pi$, mientras que el segundo sigue siendo cero. La segunda relación es idéntica con $2\sin A\sin B=\cos(A-B)-\cos(A+B)$: para $n=m$ queda $\frac{1}{2}\int_{-\pi}^{\pi}1=\pi$.

Finalmente, las normas son $\|1\|^2=2\pi$ y $\|\cos nx\|^2=\|\sin nx\|^2=\pi$, de donde salen los divisores $\sqrt{2\pi}$ y $\sqrt{\pi}$ del enunciado. □

Teorema 33.10 (Deducción de los coeficientes de Fourier)**.** *Supóngase que una función f, 2π-periódica, admite un desarrollo*

$$f(x)=\frac{a_0}{2}+\sum_{n\geq 1}(a_n\cos nx+b_n\sin nx)$$

en el que la serie se puede integrar término a término. Entonces los coeficientes están obligados a ser

$$a_n=\frac{1}{\pi}\int_{-\pi}^{\pi}f(x)\cos nx\,\mathrm{d}x\quad(n\geq 0),\qquad b_n=\frac{1}{\pi}\int_{-\pi}^{\pi}f(x)\sin nx\,\mathrm{d}x\quad(n\geq 1).$$

Demostración. Multiplíquense los dos lados por $\cos mx$ con $m\geq 1$ fijo e intégrese entre $-\pi$ y π:

$$\int_{-\pi}^{\pi}f\cos mx\,\mathrm{d}x=\frac{a_0}{2}\int_{-\pi}^{\pi}\cos mx\,\mathrm{d}x+\sum_{n\geq 1}a_n\int_{-\pi}^{\pi}\cos nx\cos mx\,\mathrm{d}x+\sum_{n\geq 1}b_n\int_{-\pi}^{\pi}\sin nx\cos mx\,\mathrm{d}x.$$

Por el teorema 33.9, la primera integral es cero, la tercera suma es idénticamente cero y de la segunda suma *sobrevive un único término*, el de $n=m$, que vale π. Queda $\int_{-\pi}^{\pi}f\cos mx=a_m\pi$, que es la fórmula. Multiplicando en cambio por $\sin mx$ se obtiene igual la de b_m.

Falta el término constante, y aquí se ve de dónde sale el dos. Integrando la serie sin multiplicar por nada, todos los senos y cosenos se anulan y queda $\int_{-\pi}^{\pi}f=\frac{a_0}{2}\cdot 2\pi=a_0\pi$, es decir $a_0=\frac{1}{\pi}\int_{-\pi}^{\pi}f$. Si el término constante se hubiera escrito como a_0 a secas, su fórmula llevaría un $\frac{1}{2\pi}$ y sería la única distinta de todas las demás. Escribirlo como $a_0/2$ es exactamente el precio que se paga para que *una sola* fórmula sirva para todos los a_n, incluido $n=0$. No hay más misterio en ese dos. □

Corolario 33.11 (Efecto de la paridad sobre los coeficientes)**.** *Si f es par, entonces $b_n=0$ para todo n y $a_n=\frac{2}{\pi}\int_0^{\pi}f\cos nx\,\mathrm{d}x$. Si f es impar, entonces $a_n=0$ para todo n y $b_n=\frac{2}{\pi}\int_0^{\pi}f\sin nx\,\mathrm{d}x$.*

Demostración. Si f es par, el producto $f\sin nx$ es par por impar, luego impar, y su integral sobre $[-\pi,\pi]$ es nula: $b_n=0$. En cambio $f\cos nx$ es par por par, luego par, y su integral vale el doble de la de $[0,\pi]$. El caso impar es el mismo argumento con los papeles cambiados. □

Proposición 33.12 (Bessel, mejor aproximación y Parseval)**.** *Sea f de cuadrado integrable en $[-\pi,\pi]$ con coeficientes a_n, b_n, y sea $S_N f = \frac{a_0}{2} + \sum_{n=1}^{N}(a_n \cos nx + b_n \sin nx)$ su suma parcial. Entonces:*

(i) entre todos los polinomios trigonométricos de grado $\le N$, el que minimiza $\int_{-\pi}^{\pi} |f-T|^2$ es exactamente $T = S_N f$;

(ii) desigualdad de Bessel*: $\dfrac{a_0^2}{2} + \displaystyle\sum_{n=1}^{N}(a_n^2+b_n^2) \le \frac{1}{\pi}\int_{-\pi}^{\pi} |f|^2$ para todo N; en particular la serie de la izquierda converge y $a_n, b_n \to 0$;*

(iii) identidad de Parseval*: como el sistema trigonométrico es completo, la desigualdad es en realidad una igualdad,*

$$\frac{a_0^2}{2} + \sum_{n\ge 1}(a_n^2+b_n^2) = \frac{1}{\pi}\int_{-\pi}^{\pi} |f(x)|^2\,\mathrm{d}x,$$

y esto equivale a que $\|f - S_N f\| \to 0$, es decir, a la convergencia en media cuadrática*.*

Demostración. Escríbase $T = \frac{c_0}{2} + \sum_{n\le N}(c_n \cos nx + d_n \sin nx)$ y desarróllese, usando la ortogonalidad del teorema 33.9 para que todos los productos cruzados desaparezcan:

$$\frac{1}{\pi}\int_{-\pi}^{\pi} |f-T|^2 = \frac{1}{\pi}\int_{-\pi}^{\pi} |f|^2 - \Big[a_0 c_0 + 2\sum_{n\le N}(a_n c_n + b_n d_n)\Big] + \Big[\frac{c_0^2}{2} + \sum_{n\le N}(c_n^2+d_n^2)\Big].$$

Completando cuadrados, esto es

$$\frac{1}{\pi}\int |f|^2 - \Big[\frac{a_0^2}{2} + \sum_{n\le N}(a_n^2+b_n^2)\Big] + \Big[\frac{(c_0-a_0)^2}{2} + \sum_{n\le N}\big((c_n-a_n)^2 + (d_n-b_n)^2\big)\Big].$$

El último corchete es una suma de cuadrados y es lo único que depende de T: se minimiza, anulándose, tomando $c_n = a_n$ y $d_n = b_n$, es decir $T = S_N f$. Eso prueba (i). Y como el error resultante es no negativo, queda $\frac{a_0^2}{2} + \sum_{n\le N}(a_n^2+b_n^2) \le \frac{1}{\pi}\int |f|^2$, que es (ii); las sumas parciales de una serie de términos no negativos acotadas superiormente convergen, y su término general tiende a cero. Para (iii), el error cometido es $\frac{1}{\pi}\int |f - S_N f|^2 = \frac{1}{\pi}\int |f|^2 - [\frac{a_0^2}{2} + \sum_{n\le N}(a_n^2+b_n^2)]$, de modo que Parseval y la convergencia en media cuadrática son literalmente el mismo enunciado leído de dos maneras. □

Idea clave

Conviene separar bien dos cosas que suenan parecidas. La *convergencia en media cuadrática* dice que el área del error al cuadrado tiende a cero, y vale *siempre* que f sea de cuadrado integrable, sin ninguna hipótesis adicional: es geometría de Hilbert pura, la que acaba de demostrarse. La *convergencia puntual* dice que en cada x concreto el número $S_N f(x)$ se acerca a $f(x)$, y eso es mucho más delicado: puede fallar incluso para funciones continuas. Una serie puede converger en media cuadrática y divergir en un punto sin contradicción alguna, porque un punto no tiene área. Los teoremas de Dirichlet, de Fejér y de Carleson, que se demuestran más adelante en este mismo

capítulo, se ocupan precisamente de esa segunda pregunta; aquí basta con retener que la respuesta por defecto —y la más útil— es la primera.

Ejemplo 33.13 (Serie de Fourier de $f(x)=x^2$, entera)**.** Sea $f(x)=x^2$ en $[-\pi,\pi]$, extendida por periodicidad.

(a) Paridad. f es par, luego por el corolario 33.11 todos los b_n valen 0 y no hay que calcular ni una sola integral de seno.

(b) El término constante. $a_0=\frac{1}{\pi}\int_{-\pi}^{\pi}x^2\mathrm{d}x=\frac{1}{\pi}\cdot\frac{2\pi^3}{3}=\frac{2\pi^2}{3}=6{,}579736$, de modo que en la serie aparece $\frac{a_0}{2}=\frac{\pi^2}{3}=3{,}289868$, que es el valor medio de x^2 sobre el periodo.

(c) Los cosenos, integrando dos veces por partes.

$$a_n=\frac{2}{\pi}\int_0^{\pi}x^2\cos nx\,\mathrm{d}x=\frac{2}{\pi}\left[\frac{x^2\sin nx}{n}+\frac{2x\cos nx}{n^2}-\frac{2\sin nx}{n^3}\right]_0^{\pi}=\frac{2}{\pi}\cdot\frac{2\pi(-1)^n}{n^2}=\frac{4(-1)^n}{n^2}.$$

En la evaluación, $\sin n\pi=0$ mata el primer y el tercer sumando, y $\cos n\pi=(-1)^n$ deja el del medio. Numéricamente $a_1=-4$, $a_2=1$, $a_3=-0{,}444444$, $a_4=0{,}25$, $a_5=-0{,}16$.

(d) La serie.

$$x^2=\frac{\pi^2}{3}+4\sum_{n\ge1}\frac{(-1)^n}{n^2}\cos nx,\qquad -\pi\le x\le\pi.$$

Como x^2 es continua y su extensión periódica también lo es (los dos extremos valen π^2), y además es derivable a trozos, la convergencia es uniforme: los coeficientes decaen como $1/n^2$ y $\sum|a_n|<\infty$. Aquí no hay fenómeno de Gibbs porque no hay salto.

(e) Comprobación numérica. En $x=1{,}5$, donde $f=2{,}25$, las sumas parciales valen $3{,}289868$ con $N=0$, y luego $3{,}006919$; $2{,}016927$; $2{,}110614$; $2{,}350656$; $2{,}295195$; $2{,}193958$ para $N=1,\dots,6$: oscilan alrededor del valor correcto y se van cerrando sobre él. El error máximo sobre todo el intervalo pasa de $1{,}135$ con $N=3$ a $0{,}381$ con $N=10$ y a $0{,}079$ con $N=50$.

(f) Dos identidades numéricas de regalo. Evaluando en $x=\pi$, donde $\cos n\pi=(-1)^n$ y por tanto $(-1)^n\cos n\pi=1$:

$$\pi^2=\frac{\pi^2}{3}+4\sum_{n\ge1}\frac{1}{n^2}\quad\Longrightarrow\quad\sum_{n\ge1}\frac{1}{n^2}=\frac{\pi^2}{6}=1{,}644934,$$

que es el problema de Basilea. Y evaluando en $x=0$, $0=\frac{\pi^2}{3}+4\sum(-1)^n/n^2$, de donde $\sum_{n\ge1}\frac{(-1)^{n-1}}{n^2}=\frac{\pi^2}{12}=0{,}822467$.

(g) Y Parseval da la tercera. El lado izquierdo es $\frac{1}{\pi}\int_{-\pi}^{\pi}x^4\mathrm{d}x=\frac{2\pi^4}{5}=38{,}963636$, y el derecho $\frac{a_0^2}{2}+\sum a_n^2=\frac{2\pi^4}{9}+16\sum\frac{1}{n^4}$. Igualando, $16\sum n^{-4}=\frac{2\pi^4}{5}-\frac{2\pi^4}{9}=\frac{8\pi^4}{45}$, es decir

$$\sum_{n\ge1}\frac{1}{n^4}=\frac{\pi^4}{90}=1{,}082323,$$

que es el valor de $\zeta(4)$. La suma numérica de los cien mil primeros términos da $1{,}0823232337$, coincidente hasta la décima cifra.

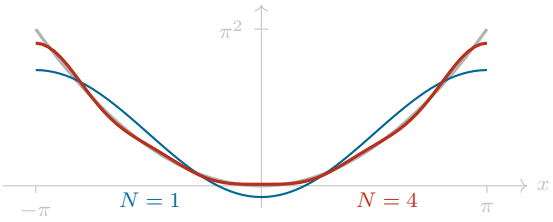


Figura 33.2: La parábola x^2 (gris) y dos sumas parciales de su serie de Fourier. Con un solo coseno ya se reconoce la forma, y con cuatro la aproximación es visualmente indistinguible salvo en los extremos. Al no haber saltos, la convergencia es uniforme y no hay sobreimpulso.

Ejemplo 33.14 (Una función sin simetría: la media rampa)**.** Sea ahora $f(x) = 0$ en $(-\pi, 0)$ y $f(x) = x$ en $(0, \pi)$, extendida por periodicidad. No es par ni impar, así que habrá cosenos y senos.

(a) El término constante. $a_0 = \frac{1}{\pi}\int_0^\pi x\,\mathrm{d}x = \frac{1}{\pi}\cdot\frac{\pi^2}{2} = \frac{\pi}{2} = 1{,}570796$, luego el término independiente de la serie es $\frac{a_0}{2} = \frac{\pi}{4} = 0{,}785398$, que es el valor medio de la función sobre el periodo.

(b) Los cosenos. Integrando por partes,

$$a_n = \frac{1}{\pi}\int_0^\pi x\cos nx\,\mathrm{d}x = \frac{1}{\pi}\left[\frac{x\sin nx}{n} + \frac{\cos nx}{n^2}\right]_0^\pi = \frac{(-1)^n - 1}{\pi n^2},$$

que vale 0 si n es par y $-\frac{2}{\pi n^2}$ si es impar. Numéricamente $a_1 = -0{,}636620$, $a_3 = -0{,}070736$, $a_5 = -0{,}025465$.

(c) Los senos. Con la misma técnica,

$$b_n = \frac{1}{\pi}\int_0^\pi x\sin nx\,\mathrm{d}x = \frac{1}{\pi}\left[-\frac{x\cos nx}{n} + \frac{\sin nx}{n^2}\right]_0^\pi = \frac{(-1)^{n+1}}{n},$$

es decir $b_1 = 1$, $b_2 = -0{,}5$, $b_3 = 0{,}333333$, $b_4 = -0{,}25$. Estos decaen como $1/n$, mucho más despacio que los cosenos: es la huella del salto.

(d) La serie completa.

$$f(x) = \frac{\pi}{4} - \frac{2}{\pi}\sum_{k\ge 0}\frac{\cos\big((2k+1)x\big)}{(2k+1)^2} + \sum_{n\ge 1}\frac{(-1)^{n+1}}{n}\sin nx.$$

(e) Qué pasa en los dos puntos delicados. En $x = 0$ la función no salta: sus dos límites laterales valen 0, y en efecto la serie da $\frac{\pi}{4} - \frac{2}{\pi}\sum(2k+1)^{-2} = \frac{\pi}{4} - \frac{2}{\pi}\cdot\frac{\pi^2}{8} = 0$ ✓. En $x = \pi$, en cambio, la extensión periódica salta de $f(\pi^-) = \pi$ a $f(\pi^+) = f(-\pi^+) = 0$, y el teorema de Dirichlet predice el promedio $\frac{\pi+0}{2} = \frac{\pi}{2} = 1{,}570796$; la serie truncada con 500 términos da $1{,}570160$. ✓

(f) Parseval. El lado izquierdo vale $\frac{1}{\pi}\int_0^\pi x^2\mathrm{d}x = \frac{\pi^2}{3} = 3{,}289868$. El derecho es

$$\frac{a_0^2}{2} + \sum_{n\ge 1}(a_n^2 + b_n^2) = \frac{\pi^2}{8} + \frac{4}{\pi^2}\sum_{k\ge 0}\frac{1}{(2k+1)^4} + \sum_{n\ge 1}\frac{1}{n^2} = \frac{\pi^2}{8} + \frac{4}{\pi^2}\cdot\frac{\pi^4}{96} + \frac{\pi^2}{6} = \frac{\pi^2}{8} + \frac{\pi^2}{24} + \frac{\pi^2}{6} = \frac{\pi^2}{3}.$$

Coinciden exactamente, y la comprobación numérica sumando doscientos mil términos da 3,289863 frente a 3,289868: la diferencia es la cola que falta.

(g) La lectura gráfica. La suma parcial con $N = 3$ ya reproduce la rampa y el tramo nulo, pero junto a $x = \pi$ y a $x = -\pi$ aparece el sobreimpulso característico. Con $N = 15$ el pico no baja: solo se estrecha y se pega más al salto. Es el fenómeno de Gibbs, que se estudia a fondo más adelante, y aquí se ve nacer.

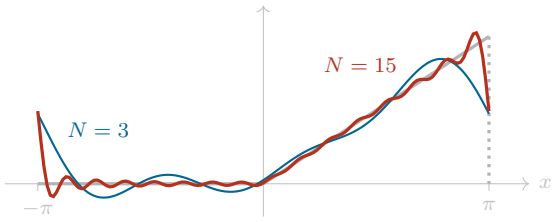


Figura 33.3: La función que vale 0 a la izquierda y x a la derecha (gris), con dos sumas parciales de su serie. En el salto de $x = \pi$ la serie apunta al promedio $\pi/2$, y a ambos lados aparece el sobreimpulso de Gibbs, que al aumentar N se estrecha pero no se reduce.

La forma exponencial compleja

Definición 33.15 (Serie de Fourier en forma exponencial)**.** Los *coeficientes exponenciales* de una función 2π-periódica son

$$c_n = \frac{1}{2\pi}\int_{-\pi}^{\pi} f(x)\,\mathrm{e}^{-\mathrm{i}nx}\,\mathrm{d}x, \qquad n \in \mathbb{Z},$$

y la serie se escribe $f(x) \sim \sum_{n\in\mathbb{Z}} c_n\,\mathrm{e}^{\mathrm{i}nx}$. La familia $\{\mathrm{e}^{\mathrm{i}nx}\}_{n\in\mathbb{Z}}$ es ortogonal en $[-\pi,\pi]$ con el producto interno complejo, pues

$$\int_{-\pi}^{\pi} \mathrm{e}^{\mathrm{i}nx}\overline{\mathrm{e}^{\mathrm{i}mx}}\,\mathrm{d}x = \int_{-\pi}^{\pi} \mathrm{e}^{\mathrm{i}(n-m)x}\,\mathrm{d}x = \begin{cases} 0, & n \neq m \\ 2\pi, & n = m. \end{cases}$$

Proposición 33.16 (Traducción entre las dos formas)**.** *Para una función real,*

$$c_0 = \frac{a_0}{2}, \qquad c_n = \frac{a_n - \mathrm{i}b_n}{2}, \qquad c_{-n} = \frac{a_n + \mathrm{i}b_n}{2} = \overline{c_n} \qquad (n \geq 1),$$

y recíprocamente $a_n = c_n + c_{-n}$ *y* $b_n = \mathrm{i}(c_n - c_{-n})$*. La identidad de Parseval toma la forma*

$$\sum_{n\in\mathbb{Z}} |c_n|^2 = \frac{1}{2\pi}\int_{-\pi}^{\pi} |f(x)|^2\,\mathrm{d}x.$$

Demostración. Por la fórmula de Euler, $\mathrm{e}^{-\mathrm{i}nx} = \cos nx - \mathrm{i}\sin nx$, de modo que

$$c_n = \frac{1}{2\pi}\int_{-\pi}^{\pi} f\cos nx\,\mathrm{d}x - \frac{\mathrm{i}}{2\pi}\int_{-\pi}^{\pi} f\sin nx\,\mathrm{d}x = \frac{a_n}{2} - \mathrm{i}\frac{b_n}{2},$$

usando las fórmulas del teorema 33.10. Cambiando n por $-n$ y usando que el coseno es par y el seno impar sale $c_{-n} = (a_n + \mathrm{i}b_n)/2$, que es el conjugado si f es real. Sumando y restando las dos expresiones se despejan a_n y b_n. Para Parseval, obsérvese que $|c_n|^2 + |c_{-n}|^2 = \frac{a_n^2+b_n^2}{2}$ y $|c_0|^2 = \frac{a_0^2}{4}$; multiplicando la identidad de la proposición 33.12 por $\frac{1}{2}$ se obtiene exactamente la versión compleja. □

Regla: pasar de la forma trigonométrica a la exponencial y al revés

1. **Recuerda que hay el doble de índices.** La forma trigonométrica usa $n \geq 0$; la exponencial usa todos los enteros. Cada pareja (a_n, b_n) con $n \geq 1$ se reparte entre c_n y c_{-n}, y por eso los módulos salen mitad.
2. **Si f es real, comprueba la simetría** $c_{-n} = \overline{c_n}$. Si no se cumple, hay un error de signo: es la comprobación más barata que existe.
3. **Si f es par, los c_n son reales** (porque $b_n = 0$); **si es impar, son imaginarios puros**.
4. **Para calcular, usa la forma exponencial**: una sola integral en lugar de dos, y las integraciones por partes son inmediatas porque la primitiva de $\mathrm{e}^{-\mathrm{i}nx}$ es ella misma dividida por $-\mathrm{i}n$.
5. **Para interpretar, usa la trigonométrica**: a_n y b_n son amplitudes reales medibles, y $\sqrt{a_n^2 + b_n^2}$ es la amplitud del armónico n-ésimo.
6. **Al aplicar Parseval, mira qué constante lleva delante**: es $\frac{1}{\pi}\int |f|^2$ en la forma trigonométrica y $\frac{1}{2\pi}\int |f|^2$ en la exponencial. Mezclar las dos es el error de factor dos más repetido del tema.

Ejemplo 33.17 (Los dos ejemplos anteriores, en forma exponencial)**. (a) La parábola.** Con $a_n = 4(-1)^n/n^2$ y $b_n = 0$ resulta $c_0 = \pi^2/3$ y $c_n = 2(-1)^n/n^2$ para todo $n \neq 0$, reales y simétricos como corresponde a una función par y real. Parseval compleja: $\frac{\pi^4}{9} + 2\sum_{n\geq 1}\frac{4}{n^4} = \frac{\pi^4}{9} + \frac{8\pi^4}{90} = \frac{\pi^4}{5}$, que en efecto es $\frac{1}{2\pi}\int_{-\pi}^{\pi} x^4 \mathrm{d}x$.

(b) La media rampa. Con $a_n = \frac{(-1)^n-1}{\pi n^2}$ y $b_n = \frac{(-1)^{n+1}}{n}$,

$$c_n = \frac{a_n - \mathrm{i}b_n}{2} = \frac{(-1)^n - 1}{2\pi n^2} + \mathrm{i}\,\frac{(-1)^n}{2n}, \qquad c_0 = \frac{\pi}{4}.$$

Para $n = 1$: $c_1 = -\frac{1}{\pi} - \frac{\mathrm{i}}{2} = -0{,}318310 - 0{,}5\,\mathrm{i}$, y $c_{-1} = -0{,}318310 + 0{,}5\,\mathrm{i}$, conjugados ✓. El módulo $|c_1| = 0{,}592700$ mide la amplitud del primer armónico.

De la serie a la integral: por qué existe la transformada

Todo lo anterior vive en un intervalo. ¿Y si la función no es periódica, sino que está definida en toda la recta y se desvanece a lo lejos? El truco es tomarla periódica de periodo $2L$ y dejar que L crezca. En $[-L, L]$ el sistema ortogonal es $\{\mathrm{e}^{\mathrm{i}n\pi x/L}\}$ y la serie es

$$f(x) = \sum_{n\in\mathbb{Z}} c_n\,\mathrm{e}^{\mathrm{i}\omega_n x}, \qquad \omega_n = \frac{n\pi}{L}, \qquad c_n = \frac{1}{2L}\int_{-L}^{L} f(y)\,\mathrm{e}^{-\mathrm{i}\omega_n y}\,\mathrm{d}y.$$

Llámese $\hat f(\omega)=\int_{-L}^{L} f(y)\mathrm{e}^{-\mathrm{i}\omega y}\mathrm{d}y$, de modo que $c_n=\hat f(\omega_n)/(2L)$. Las frecuencias ω_n están separadas por $\Delta\omega=\pi/L$, y como $\frac{1}{2L}=\frac{\Delta\omega}{2\pi}$, la serie se lee

$$f(x)=\frac{1}{2\pi}\sum_{n\in\mathbb{Z}}\hat f(\omega_n)\,\mathrm{e}^{\mathrm{i}\omega_n x}\,\Delta\omega,$$

que es una suma de Riemann. Cuando $L\to\infty$ el paso $\Delta\omega$ tiende a cero, las frecuencias dejan de ser un conjunto discreto y se convierten en un continuo, y la suma se vuelve integral. Ese paso al límite es la transformada de Fourier, y la sección siguiente la define y la explota; lo que viene aquí son las reglas de manejo, deducidas de una vez.

Definición 33.18 (Transformada de Fourier y su inversa)**.** Para f integrable en $\mathbb{R}$,

$$\hat f(\omega)=\int_{-\infty}^{\infty} f(x)\,\mathrm{e}^{-\mathrm{i}\omega x}\,\mathrm{d}x,\qquad\text{y, cuando }\hat f\text{ también es integrable,}\qquad f(x)=\frac{1}{2\pi}\int_{-\infty}^{\infty}\hat f(\omega)\,\mathrm{e}^{\mathrm{i}\omega x}\,\mathrm{d}\omega.$$

El factor $\frac{1}{2\pi}$ de la fórmula de inversión es el que sobrevive del $\frac{1}{2L}$ anterior; otros libros lo reparten de otro modo, y por eso hay que comprobar siempre el convenio antes de copiar una fórmula, como avisa la advertencia de la sección siguiente.

Proposición 33.19 (Reglas operativas)**.** *Sean f,g integrables, $c\in\mathbb{R}$ y $s>0$. Entonces*

(i) linealidad*: $\widehat{\alpha f+\beta g}=\alpha\hat f+\beta\hat g$;*

(ii) traslación*: si $g(x)=f(x-c)$, entonces $\hat g(\omega)=\mathrm{e}^{-\mathrm{i}c\omega}\hat f(\omega)$: trasladar no cambia los módulos, solo las fases;*

(iii) modulación*: si $g(x)=\mathrm{e}^{\mathrm{i}cx}f(x)$, entonces $\hat g(\omega)=\hat f(\omega-c)$;*

(iv) escalado*: si $g(x)=f(sx)$, entonces $\hat g(\omega)=\frac{1}{s}\hat f(\omega/s)$: estrechar en x es ensanchar en ω;*

(v) derivación*: si f es derivable y f,f' son integrables con $f(x)\to 0$ en el infinito, entonces $\widehat{f'}(\omega)=\mathrm{i}\omega\hat f(\omega)$;*

(vi) convolución*: $\widehat{f*g}=\hat f\cdot\hat g$.*

Demostración. (i) La integral es lineal. (ii) Cámbiese $u=x-c$: $\int f(x-c)\mathrm{e}^{-\mathrm{i}\omega x}\mathrm{d}x=\int f(u)\mathrm{e}^{-\mathrm{i}\omega(u+c)}\mathrm{d}u=\mathrm{e}^{-\mathrm{i}c\omega}\hat f(\omega)$. (iii) Agrúpense las exponenciales: $\mathrm{e}^{\mathrm{i}cx}\mathrm{e}^{-\mathrm{i}\omega x}=\mathrm{e}^{-\mathrm{i}(\omega-c)x}$. (iv) Cámbiese $u=sx$, con $\mathrm{d}x=\mathrm{d}u/s$. (v) Intégrese por partes:

$$\int f'(x)\mathrm{e}^{-\mathrm{i}\omega x}\mathrm{d}x=\big[f(x)\mathrm{e}^{-\mathrm{i}\omega x}\big]_{-\infty}^{\infty}+\mathrm{i}\omega\int f(x)\mathrm{e}^{-\mathrm{i}\omega x}\mathrm{d}x,$$

y el corchete se anula porque f tiende a cero y la exponencial tiene módulo 1. (vi) Escríbase la doble integral, cámbiese el orden e introdúzcase $u=x-y$:

$$\widehat{f*g}(\omega)=\iint f(y)g(x-y)\mathrm{e}^{-\mathrm{i}\omega x}\mathrm{d}y\,\mathrm{d}x=\int f(y)\mathrm{e}^{-\mathrm{i}\omega y}\Big(\int g(u)\mathrm{e}^{-\mathrm{i}\omega u}\mathrm{d}u\Big)\mathrm{d}y=\hat f(\omega)\hat g(\omega).\qquad\square$$

Ejemplo 33.20 (Cuatro transformadas calculadas y comprobadas)**. (a) El pulso rectangular.** Sea $f=1$ en $[-a,a]$ y 0 fuera. Entonces

$$\hat f(\omega)=\int_{-a}^{a}\mathrm{e}^{-\mathrm{i}\omega x}\mathrm{d}x=\left[\frac{\mathrm{e}^{-\mathrm{i}\omega x}}{-\mathrm{i}\omega}\right]_{-a}^{a}=\frac{2\sin(a\omega)}{\omega},$$

usando $\mathrm{e}^{\mathrm{i}\theta}-\mathrm{e}^{-\mathrm{i}\theta}=2\mathrm{i}\sin\theta$. Con $a=1$: $\hat f(0)=2$ (que es la integral de f, como debe), $\hat f(1)=1{,}682942$ y $\hat f(2)=0{,}909297$; la integración numérica da los mismos valores. Obsérvese que f tiene soporte compacto y $\hat f$ decae solo como $1/\omega$: la discontinuidad se paga en la cola.

(b) La exponencial de dos lados. Para $f(x)=\mathrm{e}^{-a|x|}$ con $a>0$,

$$\hat f(\omega)=\int_0^\infty \mathrm{e}^{-ax}\big(\mathrm{e}^{-\mathrm{i}\omega x}+\mathrm{e}^{\mathrm{i}\omega x}\big)\mathrm{d}x=\frac{1}{a+\mathrm{i}\omega}+\frac{1}{a-\mathrm{i}\omega}=\frac{2a}{a^2+\omega^2}.$$

Con $a=1$ sale $\hat f(0)=2$, $\hat f(1)=1$ y $\hat f(2)=0{,}4$, confirmados numéricamente. Es la *lorentziana*, y su decaimiento en $1/\omega^2$ corresponde al pico anguloso de f en el origen.

(c) La gaussiana. $\widehat{\mathrm{e}^{-x^2/2}}(\omega)=\sqrt{2\pi}\,\mathrm{e}^{-\omega^2/2}$; en general $\widehat{\mathrm{e}^{-ax^2}}(\omega)=\sqrt{\pi/a}\,\mathrm{e}^{-\omega^2/(4a)}$, que es la fórmula que la sección siguiente utiliza. Numéricamente, para $a=\frac12$: $\hat f(0)=2{,}506628=\sqrt{2\pi}$, $\hat f(1)=1{,}520347$ y $\hat f(2)=0{,}339235$. La gaussiana es la única función que reproduce su propia forma, y por eso es el caso de igualdad del principio de incertidumbre.

(d) El triángulo, gratis por convolución. Sea g el pulso de anchura 1 centrado, es decir $g=\mathbf{1}_{[-1/2,1/2]}$, cuya transformada es $2\sin(\omega/2)/\omega$ por el apartado (a). Entonces $g*g$ es la función triangular $(1-|x|)^+$, y sin integrar nada la regla (vi) da

$$\widehat{(1-|x|)^+}(\omega)=\Big(\frac{2\sin(\omega/2)}{\omega}\Big)^2=\frac{4\sin^2(\omega/2)}{\omega^2}.$$

Valores: $0{,}919395$ en $\omega=1$, $0{,}708073$ en $\omega=2$ y $0{,}442221$ en $\omega=3$, que coinciden con la integración numérica directa. Y ahora la transformada decae como $1/\omega^2$, no como $1/\omega$: convolucionar ha suavizado la función y ha acelerado el decaimiento, que es siempre la misma historia contada dos veces.

(e) Comprobación de la inversión. Con $f(x)=\mathrm{e}^{-|x|}$ y $x=0{,}7$, la fórmula de la definición 33.18 da $\frac{1}{2\pi}\int\frac{2}{1+\omega^2}\cos(0{,}7\,\omega)\mathrm{d}\omega=0{,}496589$, y en efecto $\mathrm{e}^{-0{,}7}=0{,}496585$. Con este convenio, la identidad de Plancherel es $\int|f|^2=\frac{1}{2\pi}\int|\hat f|^2$: aquí ambos lados valen 1.

Observación 33.21. Falta un tercer escenario, el que de verdad ejecuta un ordenador. Si en lugar de una función se tienen N muestras, el sistema ortogonal es finito —las N raíces N-ésimas de la unidad— y las integrales se convierten en sumas: eso es la *transformada discreta de Fourier*, que se define y se calcula en la sección correspondiente de este capítulo. Los tres casos son la misma construcción sobre tres grupos distintos: la serie vive en el círculo, la transformada en la recta, y la discreta en $\mathbb{Z}/N$. La ortogonalidad, que es lo único que se ha usado en toda esta sección, es la misma en los tres.

Pausa para pensar

Conviene medir lo que se ha ganado al escribir $\langle f,g\rangle=\int f\bar g$. Con esa única definición, calcular un coeficiente de Fourier es proyectar sobre un eje, la desigualdad de Bessel es la afirmación de que una proyección no alarga un vector, Parseval es el teorema de Pitágoras y la mejor aproximación por polinomios trigonométricos es el pie de la perpendicular. Todo el álgebra lineal de dimensión finita del capítulo 7 se traslada palabra por palabra a un espacio de funciones, y las únicas dificultades genuinamente nuevas son las de convergencia: si la proyección sobre infinitos ejes reconstruye el vector,

y en qué sentido. Esa pregunta es la que ocupa el resto del capítulo, y ahora se sabe exactamente por qué es la pregunta correcta.

33.3. La transformada de Fourier

Definición 33.22. Para f integrable en $\mathbb{R}$, su *transformada de Fourier* es $\hat f(\omega) = \int_{\mathbb{R}} f(x)\,\mathrm{e}^{-\mathrm{i}\omega x}\,\mathrm{d}x$. Convierte derivación en multiplicación, $\widehat{f'}(\omega) = \mathrm{i}\omega\hat f(\omega)$, y convolución en producto, $\widehat{f * g} = \hat f \cdot \hat g$. El *principio de incertidumbre* afirma que $\sigma_x\,\sigma_\omega \geq \frac{1}{2}$: una función y su transformada no pueden estar ambas muy concentradas.

Idea clave

Qué dice esto en cristiano. La transformada descompone una señal en sus frecuencias puras, diciendo cuánto hay de cada una. Su gran virtud operativa es que convierte los problemas difíciles en fáciles: derivar pasa a ser multiplicar por $\mathrm{i}\omega$, de modo que una ecuación diferencial se convierte en una ecuación algebraica; y convolucionar, que es una integral pesada, pasa a ser multiplicar punto a punto. La estrategia general es siempre la misma: transformar, resolver el problema fácil, y deshacer la transformada.

Ejemplo 33.23 (Transformadas y sus lecturas)**. (a) La gaussiana es su propia forma.** Se cumple $\widehat{\mathrm{e}^{-ax^2}}(\omega) = \sqrt{\frac{\pi}{a}}\,\mathrm{e}^{-\omega^2/(4a)}$. Para $a = 1$ da $\sqrt{\pi}\,\mathrm{e}^{-\omega^2/4} = 1{,}772454\,\mathrm{e}^{-\omega^2/4}$. La gaussiana es la única función que es esencialmente su propia transformada, y también la única que alcanza la igualdad en el principio de incertidumbre.

(b) Estrechar en el tiempo es ensanchar en frecuencia. Si f se comprime por un factor a, su transformada se dilata por el mismo factor. Por eso un pulso muy corto contiene muchas frecuencias: un chasquido suena a todo el espectro a la vez, y una nota pura y sostenida ocupa apenas una frecuencia. Ese es todo el contenido del principio de incertidumbre, que en mecánica cuántica se convierte en la relación de Heisenberg sin más que interpretar ω como momento.

(c) Resolver el calor transformando. La ecuación $u_t = k\,u_{xx}$ se convierte en $\hat u_t = -k\omega^2\hat u$, una ecuación ordinaria cuya solución es $\hat u(\omega,t) = \hat f(\omega)\mathrm{e}^{-k\omega^2 t}$. Deshaciendo la transformada aparece la convolución con la gaussiana, que es exactamente el núcleo del calor. El factor $\mathrm{e}^{-k\omega^2 t}$ explica de un vistazo por qué la difusión mata primero las frecuencias altas.

(d) Muestreo y transformada rápida. El teorema de Nyquist dice que una señal sin frecuencias por encima de $f_{\text{máx}}$ se reconstruye exactamente muestreándola a más de $2f_{\text{máx}}$: por eso el disco compacto usa 44 100 muestras por segundo para cubrir los 20 kHz audibles. Y la transformada rápida reduce el coste de n^2 a $n\log n$ operaciones: para $n = 2^{20} \approx 10^6$ el ahorro es un factor de unas 52 000 veces, que es lo que hace practicable todo el procesamiento digital moderno.

Cuidado

[Los convenios difieren y las fórmulas cambian] Hay al menos tres convenios habituales para la transformada, según dónde se coloque el 2π: en el exponente, delante de la integral o repartido entre ida y vuelta. Las fórmulas de Parseval y de la convolución cambian de constante según cuál se use, así que copiar una fórmula de otra fuente sin comprobar el convenio es una fuente segura de errores de factor 2π. Segundo aviso: la transformada exige integrabilidad, y funciones tan básicas como la constante 1 o $\sin x$ no la tienen en el sentido clásico; sus transformadas existen solo como *distribuciones*, y de ahí sale la delta de Dirac, que no es una función sino un funcional.

Pausa para pensar

Fourier presentó sus series en 1807 para resolver la ecuación del calor, y la Academia de París le rechazó el trabajo: Lagrange encabezó la oposición porque no aceptaba que una función discontinua pudiera escribirse como suma de senos, todos ellos continuos. Tardó quince años en publicarlo. La disputa era legítima y fue extraordinariamente fecunda: para responderla hubo que definir con precisión qué es una función, qué es la convergencia y qué es una integral, lo que llevó directamente a Dirichlet, a Riemann, a Cantor (que llegó a la teoría de conjuntos estudiando los conjuntos de unicidad de series trigonométricas) y a Lebesgue. Es difícil encontrar una sola pregunta que haya generado más matemática.

33.4. Convolución, núcleos y el teorema de muestreo

La transformada de Fourier no sería tan útil si no fuese por una identidad casi mágica: convierte la convolución, que es una operación complicada, en un producto ordinario. De esa identidad cuelgan el filtrado de señales, la teoría de la aproximación por núcleos y el teorema que hace posible el audio digital.

Definición 33.24. La *convolución* de dos funciones integrables $f, g\colon \mathbb{R} \to \mathbb{C}$ es

$$(f * g)(x) = \int_{-\infty}^{\infty} f(y)\, g(x-y)\, \mathrm{d}y.$$

Con el convenio $\hat{f}(\xi) = \int f(x) \mathrm{e}^{-2\pi i x \xi} \mathrm{d}x$ se cumple el *teorema de convolución*

$$\widehat{f * g} = \hat{f} \cdot \hat{g}, \qquad \text{y recíprocamente} \qquad \widehat{f \cdot g} = \hat{f} * \hat{g}.$$

Idea clave

Qué dice esto en cristiano. Convolucionar es emborronar: se sustituye cada valor de f por un promedio ponderado de sus vecinos, con pesos dados por g. Desenfocar una fotografía, suavizar una serie temporal con una media móvil o repartir una carga puntual en una nube difusa son todos convoluciones. El teorema dice que ese emborronado, que en el dominio original obliga a integrar en cada punto, se convierte en el dominio de

frecuencias en una simple multiplicación número a número. Como la transformada de un núcleo suavizante vale casi 1 en frecuencias bajas y casi 0 en las altas, multiplicar por ella *apaga los detalles finos y conserva las tendencias*: eso es exactamente lo que hace un filtro paso bajo, y ahora se ve por qué.

Regla: Manejar convoluciones sin equivocarse

1. **Determinar el soporte antes de integrar**: si f vive en $[a,b]$ y g en $[c,d]$, entonces $f*g$ vive en $[a+c,b+d]$. Los soportes se suman.
2. **Reflejar y desplazar**: en $(f*g)(x)$ la función g aparece como $g(x-y)$, es decir, dada la vuelta y trasladada. Olvidar la reflexión es el error clásico.
3. **Usar la conmutatividad** $f*g=g*f$ para reflejar la más sencilla de las dos.
4. **Trocear en casos** según el solapamiento de los soportes: entra, solapa totalmente, sale. Cada caso da un trozo de la fórmula.
5. **Comprobar la masa**: $\int(f*g)=(\int f)(\int g)$, lo que en frecuencia es $\widehat{f*g}(0)=\hat f(0)\hat g(0)$.
6. **Comprobar la regularidad**: $f*g$ es al menos tan suave como la más suave de las dos. Convolucionar sólo puede mejorar la regularidad, nunca empeorarla.
7. **Si aparece un producto en vez de una convolución**, aplicar el teorema al revés: multiplicar por una ventana en el tiempo equivale a convolucionar con su transformada en frecuencia, y de ahí vienen las fugas espectrales.

Ejemplo 33.25 (Cuatro convoluciones calculadas)**. (a) Caja con caja.** Sea $f=g=\mathbf{1}_{[0,1]}$. El solapamiento de $[0,1]$ con $[x-1,x]$ tiene longitud x si $0\le x\le 1$ y $2-x$ si $1\le x\le 2$, de modo que $f*f$ es el triángulo

$$(f*f)(0)=0,\quad (f*f)(0{,}5)=0{,}5,\quad (f*f)(1)=1,\quad (f*f)(1{,}5)=0{,}5,\quad (f*f)(2)=0.$$

El soporte pasó de $[0,1]$ a $[0,2]$ (se sumaron) y la función pasó de discontinua a continua: la convolución suavizó un escalón.

(b) La gaussiana autorrecíproca. Con $f(x)=\mathrm{e}^{-\pi x^2}$ resulta $\hat f=f$. Numéricamente, $\hat f(0)=1$, $\hat f(0{,}5)=0{,}455938$ y $\hat f(1)=0{,}043214$, que coinciden con $\mathrm{e}^{-\pi\xi^2}$ en esos puntos. Consecuencia inmediata del teorema de convolución: la convolución de dos gaussianas es otra gaussiana, y las varianzas se suman, que es precisamente por qué la suma de normales independientes es normal.

(c) Parseval sobre la caja. Para $f=\mathbf{1}_{[-1/2,1/2]}$ se tiene $\hat f(\xi)=\operatorname{sinc}(\xi)=\sin(\pi\xi)/(\pi\xi)$. Entonces $\int|f|^2=1$ y $\int|\hat f|^2=1$ también, como exige Plancherel; la integración numérica sobre $[-400,400]$ da $0{,}999747$, y lo que falta está en las colas, que decaen sólo como ξ^{-2}.

(d) La aproximación de la identidad. Si $\varphi_\varepsilon(x)=\varepsilon^{-1}\varphi(x/\varepsilon)$ con $\int\varphi=1$, entonces $f*\varphi_\varepsilon\to f$ cuando $\varepsilon\to 0$ en casi todo sentido razonable. No existe una función δ que sea neutro de la convolución, pero sí sucesiones que se le acercan tanto como se quiera: la delta

de Dirac es exactamente el límite que falta, y por eso hace falta la teoría de distribuciones para darle estatus legal.

Teorema 33.26 (Muestreo de Nyquist–Shannon)**.** *Si $f \in L^2(\mathbb{R})$ tiene transformada soportada en $[-B, B]$ (es de* banda limitada*), entonces f queda determinada por sus valores en los instantes $n/(2B)$, $n \in \mathbb{Z}$, y se reconstruye exactamente mediante*

$$f(t) = \sum_{n \in \mathbb{Z}} f\Big(\frac{n}{2B}\Big)\, \operatorname{sinc}(2Bt - n).$$

Idea clave

Qué dice esto en cristiano. Una señal sin frecuencias por encima de B no puede hacer nada sorprendente entre dos muestras tomadas suficientemente juntas: para colarse entre ellas tendría que oscilar deprisa, y eso es justo lo que la limitación de banda le prohíbe. Por tanto, muestreando al doble de la frecuencia máxima no se pierde absolutamente nada, y la señal continua original se puede recuperar íntegra. Éste es el teorema que justifica que un disco compacto muestree a 44 100 hercios: el oído humano llega a unos 20 000, y $2 \times 20\,000 = 40\,000$, con un margen para el filtro.

Ejemplo 33.27 (Aliasing: qué pasa cuando se muestrea despacio)**.** Tómese $f(t) = \sin(2\pi \cdot 3\, t)$, de frecuencia 3 hercios, y muéstrese a 4 hercios, es decir, por debajo del mínimo exigido de 6.

(a) Las muestras. En $t = n/4$ los valores de $\sin(6\pi t)$ son $0;\ -1;\ 0;\ 1;\ 0;\ -1$ para $n = 0, \dots, 5$.

(b) El impostor. La señal $-\sin(2\pi t)$, de frecuencia 1 hercio, produce en esos mismos instantes $0;\ -1;\ 0;\ 1;\ 0;\ -1$: exactamente las mismas muestras. Ninguna cantidad de procesamiento posterior puede distinguirlas, porque los datos son idénticos.

(c) La aritmética del alias. Al muestrear a f_s, toda frecuencia f se confunde con $|f - k f_s|$ para cualquier entero k. Aquí $|3 - 4| = 1$, de donde el 1 hercio. Ésta es la razón de que las ruedas de los carros parezcan girar hacia atrás en el cine: a 24 imágenes por segundo, una rueda que gira a 25 vueltas por segundo aparenta girar a -1.

(d) El remedio. Filtrar *antes* de muestrear, nunca después. Un filtro antialias elimina las frecuencias por encima de $f_s/2$ mientras la señal es todavía continua; una vez tomadas las muestras, la información se ha perdido de forma irreversible y no hay filtro que la devuelva.

Cuidado

[Contraejemplo: el núcleo de Dirichlet no aproxima la identidad] Las sumas parciales de una serie de Fourier son $S_n f = f * D_n$, donde D_n es el núcleo de Dirichlet. Uno esperaría que D_n fuese una aproximación de la identidad y que $S_n f \to f$ siempre; falla porque D_n *cambia de signo* y su norma L^1 no está acotada: las constantes de Lebesgue $L_n = \frac{1}{2\pi} \int |D_n|$ valen 1,435991 para $n = 1$, 1,961361 para $n = 5$ y 2,494493 para $n = 20$, creciendo como $\frac{4}{\pi^2} \ln n$. Ese crecimiento logarítmico es exactamente lo que permite construir funciones continuas cuya serie de Fourier diverge en un punto. El núcleo de

Fejér, en cambio, es no negativo y con integral 1: sí es aproximación de la identidad, y por eso las medias de Cesàro convergen uniformemente para toda función continua allí donde las sumas parciales fracasan.

Cuidado

[Contraejemplo: nada de banda limitada tiene soporte compacto] El teorema de muestreo se enuncia para señales de banda limitada, y ninguna señal real lo es: si $\hat f$ se anula fuera de $[-B,B]$, entonces f es la restricción de una función entera, y una función entera que se anule en un intervalo se anula en todas partes. Es decir, una señal de banda limitada *dura desde* $-\infty$ *hasta* $+\infty$. Toda grabación tiene principio y fin, luego tiene energía en frecuencias arbitrariamente altas, y el filtro antialias sólo puede dejarla pequeña, nunca nula. El teorema es una idealización excelente, no una descripción.

Pausa para pensar

El principio de incertidumbre de Heisenberg es, matemáticamente, un enunciado sobre transformadas de Fourier y no sobre física: para toda f de norma 1,

$$\Big(\int x^2|f(x)|^2\mathrm{d}x\Big)\Big(\int \xi^2|\hat f(\xi)|^2\mathrm{d}\xi\Big)\ \geq\ \frac{1}{16\pi^2},$$

con igualdad exactamente para las gaussianas. Con $f(x)=\sqrt[4]{2}\,\mathrm{e}^{-\pi x^2}$ las dos dispersiones valen $1/(4\pi)=0{,}079577$ y su producto $0{,}00633257$, que coincide con $1/(16\pi^2)$ hasta el último decimal calculado. Una función no puede estar concentrada a la vez en el tiempo y en la frecuencia; la mecánica cuántica hereda el teorema porque posición y momento son un par de Fourier, no porque la naturaleza sea caprichosa.

33.5. Ondículas: mirar la frecuencia sin perder el tiempo

La transformada de Fourier dice qué frecuencias contiene una señal y no dice cuándo. Para una nota sostenida da igual; para una señal que cambia —una grabación, una imagen, un electrocardiograma— es una pérdida fatal. Las ondículas resuelven el problema cambiando la base.

Definición 33.28. Una *ondícula* madre $\psi\in L^2(\mathbb{R})$ con media nula genera, por dilataciones y traslaciones diádicas, la familia

$$\psi_{j,k}(x)=2^{j/2}\psi(2^jx-k),\qquad j,k\in\mathbb{Z},$$

y para elecciones adecuadas de ψ esa familia es una base ortonormal de $L^2(\mathbb{R})$. La *ondícula de Haar* es $\psi=\mathbf{1}_{[0,1/2)}-\mathbf{1}_{[1/2,1)}$.

Idea clave

Qué dice esto en cristiano. Una exponencial $e^{i\omega x}$ oscila con la misma amplitud desde menos infinito hasta más infinito: sabe perfectamente qué frecuencia es y no tiene ni idea de dónde está. Una ondícula es un pulso corto: sabe bastante bien dónde está y algo menos qué frecuencia tiene. Al dilatarla se cubren todas las escalas —las versiones estrechas detectan detalles finos y las anchas, tendencias lentas— y al trasladarla se cubre todo el eje. El resultado es un diccionario que localiza en tiempo y en frecuencia a la vez, dentro de lo que el principio de incertidumbre permite, que no es todo pero es mucho más que Fourier.

Ejemplo 33.29 (Qué se gana al cambiar de base)**. (a) Haar, a mano.** Con la señal $(4, 6, 10, 12)$, el primer nivel da medias $(5, 11)$ y detalles $(-1, -1)$; el segundo, media 8 y detalle -3. La señal entera queda codificada como $(8, -3, -1, -1)$, y el proceso es reversible. Cada coeficiente dice "cuánto difiere esta mitad de la otra", a una escala distinta.

(b) Por qué comprime. En una señal suave por tramos, casi todos los detalles finos son minúsculos y se pueden poner a cero sin que se note. Una imagen de un millón de píxeles se reconstruye de forma aceptable con unos pocos miles de coeficientes. Con Fourier no funciona igual: un solo borde en la imagen —una discontinuidad— excita *todas* las frecuencias, y por eso las compresiones basadas en cosenos producen los artefactos en bloques que se ven alrededor del texto en una imagen muy comprimida.

(c) El coste. La transformada rápida de ondículas es $O(n)$, incluso mejor que el $O(n \log n)$ de la transformada rápida de Fourier, porque cada nivel trabaja sobre la mitad de los datos del anterior y la serie geométrica suma.

(d) Dónde se usa. El formato JPEG2000 y el estándar del FBI para huellas dactilares usan ondículas; el JPEG clásico usa la transformada del coseno. Y en 2015, la detección de ondas gravitacionales por LIGO empleó métodos de tiempo-frecuencia emparentados para aislar una señal de menos de un segundo enterrada en ruido.

Cuidado

[Contraejemplo: Haar no sirve para casi nada práctico] La ondícula de Haar es discontinua, de modo que representa mal las señales suaves: hacen falta muchísimos coeficientes para aproximar una recta inclinada, porque cada escalón hay que construirlo. Es magnífica para explicar la idea y mala como herramienta. Las que se usan de verdad son las de Daubechies, que tienen soporte compacto y varios momentos nulos —la de orden N anula los polinomios de grado menor que N—, y no admiten fórmula cerrada: se definen por una recursión y sólo se conocen numéricamente. Esa fue la aportación de 1988 que convirtió las ondículas en tecnología.

Pausa para pensar

La historia de las ondículas es un caso de reinvención múltiple. Haar publicó su base en 1909 sin sospechar para qué serviría. En los setenta, un geofísico llamado Morlet las redescubrió analizando ecos sísmicos para prospección petrolera, y sus colegas le dijeron que aquello no podía funcionar; se fue a hablar con el físico Grossmann, que le dio el marco teórico. Meyer encontró la primera base ortonormal suave en 1985 y contó que se enteró del tema por un artículo que le pasaron en la fotocopiadora. Daubechies cerró el asunto en 1988. En menos de quince años, una curiosidad de la prospección petrolera se convirtió en un estándar internacional de compresión de imágenes.

33.6. Inversión, Plancherel, clase de Schwartz y distribuciones temperadas

Hemos calculado transformadas sin decir nunca cómo se vuelve, ni en qué espacio vive la operación, ni qué significa transformar una función que no es integrable. Toca cerrar esos tres huecos, que son el mismo hueco.

Teorema 33.30 (Riemann–Lebesgue)**.** *Si $f \in L^1(\mathbb{R})$, entonces $\hat{f}$ es continua, acotada por $\|f\|_1$, y $\hat{f}(\xi) \to 0$ cuando $|\xi| \to \infty$.*

Esto ya avisa del problema: la transformada de una función de L^1 está en C_0, no necesariamente en L^1. Así que la fórmula de inversión no se puede aplicar sin más.

Teorema 33.31 (Inversión)**.** *Si $f \in L^1$ y además $\hat{f} \in L^1$, entonces*

$$f(x) = \int_{\mathbb{R}} \hat{f}(\xi) e^{2\pi i x \xi} \, \mathrm{d}\xi$$

en casi todo punto, y en todo punto si f es continua.

Definición 33.32 (Clase de Schwartz)**.** $\mathcal{S}(\mathbb{R}^n)$ es el espacio de las funciones infinitamente derivables que decrecen, junto con todas sus derivadas, más deprisa que cualquier potencia:

$$\sup_x |x^\alpha \partial^\beta f(x)| < \infty \quad \text{para todos los multiíndices } \alpha, \beta.$$

Ahí la teoría es perfecta: la transformada es una *biyección* de $\mathcal{S}$ en $\mathcal{S}$, con inversa la fórmula de arriba. La razón es la dualidad derivada-multiplicación: derivar en x es multiplicar por ξ y viceversa, luego decrecimiento rápido y suavidad se intercambian, y $\mathcal{S}$ es el espacio donde ambas cosas se dan a la vez. La gaussiana $e^{-\pi x^2}$, que es su propia transformada, es el habitante emblemático.

Teorema 33.33 (Plancherel)**.** *La transformada de Fourier, definida en $\mathcal{S}$, se extiende de manera única a un operador unitario de $L^2(\mathbb{R}^n)$ en sí mismo:*

$$\int |f|^2 = \int |\hat{f}|^2, \qquad \langle f, g \rangle = \langle \hat{f}, \hat{g} \rangle.$$

Esto es lo que legitima toda la intuición geométrica: Fourier es una *rotación* del espacio de Hilbert L^2, que conserva longitudes y ángulos. La energía de una señal es la misma medida en el tiempo o en la frecuencia.

Ejemplo 33.34 (Caja y sinc, con la energía contada por los dos lados)**.** Sea $f = \mathbf{1}_{[-1/2,1/2]}$. Entonces

$$\hat f(\xi) = \int_{-1/2}^{1/2} e^{-2\pi i x\xi}\mathrm{d}x = \frac{\sin(\pi\xi)}{\pi\xi} = \operatorname{sinc}\xi.$$

(a) La inversión no es inmediata. $\hat f$ decae como $1/|\xi|$, luego *no* está en L^1: el teorema de inversión no se aplica literalmente. La integral hay que entenderla como valor principal, y en $x = \pm 1/2$ converge a $1/2$, la media de los dos límites laterales, exactamente como en las series de Fourier.

(b) Plancherel sí funciona. f está en L^2 con $\|f\|_2^2 = 1$, y en efecto

$$\int_{-\infty}^{\infty} \frac{\sin^2(\pi\xi)}{(\pi\xi)^2}\,\mathrm{d}\xi = 1.$$

De paso, esa integral es el valor de $\int \operatorname{sinc}^2$, que muchos textos deducen por residuos y que aquí sale de contar energía.

Distribuciones temperadas

Definición 33.35. $\mathcal{S}'$ es el dual topológico de $\mathcal{S}$: los funcionales lineales continuos sobre la clase de Schwartz. La transformada de una distribución se define por traspaso,

$$\langle \hat T, \varphi\rangle = \langle T, \hat\varphi\rangle,$$

que tiene sentido porque $\hat\varphi \in \mathcal{S}$ siempre.

Con esa definición todo lo que antes "no tenía transformada" la tiene:

$$\hat\delta = 1, \qquad \hat 1 = \delta, \qquad \widehat{e^{2\pi i a x}} = \delta_a, \qquad \widehat{\cos(2\pi a x)} = \tfrac12(\delta_a + \delta_{-a}),$$

$$\widehat{\operatorname{sgn}}(\xi) = \frac{1}{i\pi}\,\mathrm{v.p.}\frac{1}{\xi}, \qquad \widehat{\mathrm{v.p.}\tfrac{1}{x}}(\xi) = -i\pi\operatorname{sgn}\xi.$$

Cada una de estas líneas justifica una práctica de ingeniería que llevaba un siglo funcionando sin permiso. Que la transformada de una constante sea una delta dice que una señal continua "tiene toda su energía en frecuencia cero"; que la de un coseno sean dos deltas dice que un tono puro es una raya en el espectro. Y el peine de Dirac $\sum_{n\in\mathbb{Z}}\delta_n$ es su propia transformada, hecho que *es* la fórmula de sumación de Poisson y del que sale el teorema del muestreo.

Cuidado

Las distribuciones no se pueden multiplicar entre sí. No existe δ^2, ni $\delta\cdot\mathrm{v.p.}\frac{1}{x}$; el producto solo está definido contra funciones suaves. Ese es el obstáculo que hace que las EDP no lineales no se puedan tratar con distribuciones sin más, y el origen de toda la teoría posterior de espacios de Sobolev, paraproductos y estructuras de regularidad. En el

terreno lineal, en cambio, la teoría de distribuciones es un instrumento sin fisuras.

33.7. La transformada discreta de Fourier y el algoritmo rápido

Todo lo anterior es continuo. Lo que se calcula en un ordenador es otra cosa, y conviene decir exactamente cuál.

Definición 33.36 (DFT)**.** Para $x = (x_0, \dots, x_{N-1}) \in \mathbb{C}^N$ y $\omega = e^{-2\pi i/N}$,

$$X_k = \sum_{n=0}^{N-1} x_n\, \omega^{nk}, \qquad k = 0, \dots, N-1.$$

La matriz $F = (\omega^{nk})$ cumple $F^*F = NI$: es $\sqrt{N}$ veces una matriz unitaria. La inversa es la misma fórmula con ω^{-1} y un factor $1/N$.

La DFT es, pues, un cambio de base ortogonal en $\mathbb{C}^N$: pasar de la base canónica (una coordenada por instante) a la base de exponenciales discretas (una coordenada por frecuencia). Y hereda las propiedades del caso continuo con una salvedad importante: la convolución que diagonaliza es la *circular*, porque los índices se leen módulo N. Ahí es donde vive el error clásico de la convolución rápida, que se corrige rellenando con ceros hasta que la longitud sea al menos $n_1 + n_2 - 1$.

Ejemplo 33.37 (Ocho puntos, calculados y contados)**.** Sea $x = (1,1,1,1,0,0,0,0)$: media onda cuadrada.

k	X_k	$\lvert X_k\rvert$
0	4	4
1	$1 - 2{,}414\,2\,i$	$2{,}613\,1$
2	0	0
3	$1 - 0{,}414\,2\,i$	$1{,}082\,4$
4	0	0
5	$1 + 0{,}414\,2\,i$	$1{,}082\,4$
6	0	0
7	$1 + 2{,}414\,2\,i$	$2{,}613\,1$

Tres lecturas inmediatas. Primera: $X_0 = 4$ es la suma de los datos, es decir N veces la media. Segunda: $X_{N-k} = \overline{X_k}$, porque los datos son reales; solo hay que calcular la mitad. Tercera: los coeficientes de índice par, salvo el de índice cero, se anulan: es la firma de una señal con simetría de media onda —exactamente la razón por la que la serie de Fourier de la onda cuadrada solo tiene armónicos impares—.

El recuento de operaciones. Directamente son $N^2 = 64$ productos complejos. La recursión de Cooley–Tukey parte la suma en índices pares e impares,

$$X_k = \underbrace{\sum_m x_{2m}\omega_{N/2}^{mk}}_{\text{DFT de los pares}} + \omega_N^k \underbrace{\sum_m x_{2m+1}\omega_{N/2}^{mk}}_{\text{DFT de los impares}},$$

y como el segundo sumando solo cambia de signo al pasar de k a $k+N/2$, cada par de salidas cuesta *una* multiplicación. Son $\frac{N}{2}\log_2 N = 12$ mariposas en total, frente a 64: cinco veces menos con $N=8$.

Idea clave

El ahorro no impresiona con ocho puntos; con un millón, sí. $N^2 = 10^{12}$ frente a $\frac{N}{2}\log_2 N \approx 10^7$: cien mil veces menos trabajo. Un cálculo que llevaría once días pasa a ser instantáneo. Gauss conocía el algoritmo en 1805 y lo dejó en un cuaderno sin publicar; Cooley y Tukey lo redescubrieron en 1965 y cambiaron la ingeniería del siglo XX. El tratamiento digital de señales, la tomografía, la compresión de audio y la multiplicación rápida de enteros son todos consecuencias directas de esas doce mariposas.

Cuidado

[Aliasing y fugas espectrales] Dos trampas que la fórmula no advierte. La primera: si la señal contiene frecuencias por encima de la mitad de la de muestreo, aparecen *dobladas* sobre las bajas y son indistinguibles de ellas. Es el aliasing, y no hay procesado posterior que lo repare; hay que filtrar antes de muestrear. La segunda: si la señal no es periódica en la ventana observada, sus discontinuidades en los extremos reparten energía por todo el espectro. Es la fuga espectral, y se atenúa multiplicando por una ventana suave (Hann, Hamming) antes de transformar, a costa de perder resolución en frecuencia. Ninguna de las dos es un defecto del algoritmo: son consecuencias de haber discretizado.

33.8. Resultados positivos: Fejér, Poisson y la traducción suavidad–decaimiento

Hasta aquí las noticias han sido malas: hay continuas cuya serie diverge en un punto, integrables cuya serie diverge en todas partes, y el fenómeno de Gibbs no desaparece nunca. Conviene equilibrar la balanza, porque los resultados positivos son igual de contundentes.

Definición 33.38 (Medias de Cesàro y núcleo de Fejér)**.** $\sigma_N f = \dfrac{S_0 f + \cdots + S_{N-1} f}{N}$, y se cumple $\sigma_N f = f * F_N$ con

$$F_N(t) = \frac{1}{N}\Big(\frac{\sin(Nt/2)}{\sin(t/2)}\Big)^2 \geq 0.$$

Todo está en ese ≥ 0. El núcleo de Dirichlet, que produce las sumas parciales, cambia de signo y tiene norma L^1 que crece como $\log N$; de ahí todas las patologías. El de Fejér es no negativo, tiene integral 1 y se concentra en el origen: es una aproximación de la identidad de manual.

Teorema 33.39 (Fejér, 1900)**.** *Si f es continua y 2π-periódica, $\sigma_N f \to f$* uniformemente*. Si f es solo integrable, $\sigma_N f \to f$ en L^1 y en todo punto de continuidad.*

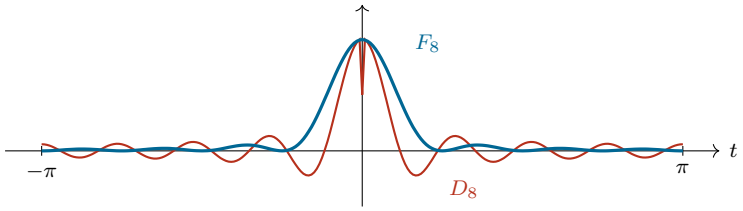


Figura 33.4: Los dos núcleos con $N = 8$, reescalados a la misma altura. El de Dirichlet, que produce las sumas parciales, oscila y cambia de signo: su norma L^1 crece como $\log N$ y de ahí salen todas las patologías, incluido Gibbs. El de Fejér, que produce las medias de Cesàro, es *no negativo* y por eso $\sigma_N f$ nunca se sale del rango de f: promediar mata el sobreimpulso.

Observación 33.40. El teorema de aproximación de Weierstrass sale de aquí en dos líneas: $\sigma_N f$ es un polinomio trigonométrico, y aproxima uniformemente a cualquier continua. Componiendo con el cambio $x = \cos\theta$ se obtiene la versión polinómica. Fejér tenía diecinueve años.

Ejemplo 33.41 (Gibbs desaparece al promediar)**.** Onda cuadrada de altura 1, con salto en el origen.

(a) Sumas parciales. S_{31} alcanza un máximo de 1,179 3 cerca del salto, y ese sobreimpulso —un 8,95 por ciento de la altura del salto— *no* disminuye al aumentar N: solo se estrecha. El valor límite es la célebre constante de Gibbs $\frac{1}{\pi}\int_0^\pi \frac{\sin t}{t}\mathrm{d}t \cdot 2 = 1{,}178\,979\,7$.

(b) Medias de Cesàro. σ_{31} alcanza un máximo de 0,995 5: *no* sobrepasa. Nunca lo hace, y la razón es que el núcleo de Fejér es no negativo, luego $\sigma_N f$ está siempre acotada entre el ínfimo y el supremo de f. Promediar las sumas parciales elimina el fenómeno de Gibbs por completo.

(c) El precio. En $x = 0{,}05$, donde el valor exacto es 1, se tiene $S_{15} = 0{,}491\,6$ y $\sigma_{15} = 0{,}266\,8$. Cesàro converge, pero *más despacio*. Es el compromiso de siempre: se gana en estabilidad lo que se pierde en velocidad, y por eso en la práctica se usan ventanas intermedias.

Teorema 33.42 (Suavidad y decaimiento son la misma cosa)**.** *Para f periódica:*

$$
\begin{aligned}
&f \in C^k \implies \hat f(n) = o(|n|^{-k}) && \textit{(integrando por partes } k \textit{ veces)}\\
&\textstyle\sum |n|^k |\hat f(n)| < \infty \implies f \in C^k && \textit{(derivando la serie)}\\
&f \textit{ analítica} \iff |\hat f(n)| \le Ce^{-a|n|} && \textit{(decaimiento exponencial)}\\
&f \textit{ con un salto} \implies \hat f(n) \sim c/n && \textit{(decaimiento mínimo)}
\end{aligned}
$$

Este diccionario es probablemente lo más útil de todo el capítulo. Explica de un golpe: por qué la onda cuadrada, que salta, tiene coeficientes $\sim 1/n$ y converge tan mal; por qué la onda triangular, continua con derivada a saltos, los tiene $\sim 1/n^2$ y converge de maravilla; por qué el fenómeno de Gibbs solo aparece en presencia de discontinuidades; y por qué en compresión de imágenes conviene detectar bordes antes de transformar. También explica que la regularidad de una función se pueda *medir* contando el decaimiento de su espectro, que es la idea de la que salen los espacios de Sobolev definidos por Fourier y toda la teoría moderna de regularidad.

33.9. Integrales singulares: función maximal, transformada de Hilbert y Calderón–Zygmund

Todo lo anterior es análisis *de* Fourier: la transformada como herramienta. El análisis armónico como disciplina empieza aquí, con operadores cuyos núcleos no son integrables y que sin embargo están acotados.

Definición 33.43 (Función maximal de Hardy–Littlewood)**.**

$$Mf(x) = \sup_{r>0} \frac{1}{|B(x,r)|} \int_{B(x,r)} |f(y)|\,\mathrm{d}y.$$

Mide el mayor promedio de $|f|$ alrededor de x. Es un operador brutal —toma un supremo sobre todas las escalas— y aun así se controla.

Teorema 33.44 (Desigualdad maximal, tipo débil $(1,1)$)**.** *Existe C tal que para todo $\lambda > 0$*

$$\big|\{x : Mf(x) > \lambda\}\big| \le \frac{C}{\lambda}\|f\|_{L^1}.$$

Además M es acotado en L^p para todo $p > 1$.

Idea. Cada punto del conjunto de nivel está en alguna bola donde el promedio supera λ, es decir $|B| < \frac{1}{\lambda}\int_B |f|$. El lema de recubrimiento de Vitali permite extraer una subfamilia disjunta cuyas dilataciones por 5 cubren el todo, y sumando sobre ella sale la desigualdad con $C = 5^n$. □

Observación 33.45. El resultado *no* vale con $p = 1$ en el sentido fuerte: Mf nunca está en L^1 salvo si $f = 0$, porque decae como $1/|x|$ en el infinito. Esa es la razón de ser del "tipo débil". Y el teorema de diferenciación de Lebesgue —que $\frac{1}{|B_r|}\int_{B_r} f \to f(x)$ en casi todo punto— es un corolario directo: se prueba para funciones continuas, donde es trivial, y la desigualdad maximal permite pasar al límite en L^1 por densidad. Toda la teoría de la derivación en medida cuelga de esa cota.

Definición 33.46 (Transformada de Hilbert)**.**

$$Hf(x) = \frac{1}{\pi}\,\text{v.p.}\!\int_{\mathbb{R}} \frac{f(y)}{x-y}\,\mathrm{d}y,$$

donde el valor principal es imprescindible porque $1/(x-y)$ no es integrable en el origen. En Fourier es un multiplicador de módulo constante:

$$\widehat{Hf}(\xi) = -i\,\mathrm{sgn}(\xi)\,\hat{f}(\xi).$$

Esa fórmula lo dice todo: H no cambia el tamaño de ninguna frecuencia, solo su fase, en $\pm 90°$ según el signo. Es un operador *unitario* en L^2 y cumple $H^2 = -\mathrm{id}$: una raíz cuadrada de -1 actuando sobre funciones.

Ejemplo 33.47 (La transformada de un escalón)**.** Para $f = \mathbf{1}_{[-1,1]}$ la integral se calcula directamente:

$$Hf(x) = \frac{1}{\pi}\log\Big|\frac{x+1}{x-1}\Big|.$$

Valores:

x	0,5	0,9	0,99	2	5
$Hf(x)$	0,349 7	0,937 2	1,684 9	0,349 7	0,129 1

Tres observaciones. Primera: Hf *explota logarítmicamente* en $x = \pm 1$, justo donde f salta. Segunda: Hf no tiene soporte compacto aunque f sí lo tenga, y decae solo como $2/(\pi x)$: no es integrable. Luego H manda una función perfectamente buena de L^1 fuera de L^1. Tercera: y sin embargo $Hf \in L^p$ para todo $p > 1$.

Teorema 33.48 (Acotación L^p)**.** *H es acotado en $L^p(\mathbb{R})$ para $1 < p < \infty$, y no lo es en $p = 1$ ni en $p = \infty$.*

La demostración moderna es el argumento de **Calderón–Zygmund**, y su estructura es un patrón que conviene conocer porque se reutiliza en todo el análisis contemporáneo:

Regla: el esquema de Calderón–Zygmund

1. Se prueba la acotación en L^2, que suele ser fácil porque en Fourier el operador es una multiplicación por una función acotada.
2. Se descompone $f = g + b$ a la altura λ: la parte *buena* g está acotada por $C\lambda$ y la *mala* b es una suma de trozos con media cero, cada uno soportado en un cubo, con los cubos disjuntos y de medida total $\le \frac{1}{\lambda}\|f\|_1$. La descomposición se obtiene subdividiendo diádicamente y parando en cuanto el promedio supera λ.
3. En g se usa el paso 1; en b se usa la media cero y la condición de Hörmander sobre el núcleo, $\int_{|x|>2|y|} |K(x-y) - K(x)|\,\mathrm{d}x \le C$, que es lo que hace que las cancelaciones sobrevivan.
4. Sale el tipo débil $(1,1)$, y por interpolación de Marcinkiewicz con el caso L^2 se obtiene L^p para $1 < p < 2$. Dualidad da el resto.

Idea clave

La lección general es que estos operadores son acotados *por cancelación*, no por tamaño. El núcleo $1/x$ tiene integral divergente, y lo único que salva la situación es que sus partes positiva y negativa se compensan; por eso el valor principal es esencial y por eso las demostraciones tienen que usar la media cero de b.
En dimensión superior el análogo son las **transformadas de Riesz** R_j, con multiplicador $-i\xi_j/|\xi|$, que satisfacen $\sum R_j^2 = -\mathrm{id}$ y permiten escribir $\partial_i\partial_j u = R_i R_j \Delta u$: de ahí sale la estimación elíptica $\|D^2 u\|_{L^p} \le C\|\Delta u\|_{L^p}$, que es el resultado que hace funcionar la teoría de regularidad de las ecuaciones en derivadas parciales. La sección de soluciones débiles del capítulo de EDP descansa, sin decirlo, en este teorema.

33.10. La fórmula de sumación de Poisson, y Fourier en varias variables

Teorema 33.49 (Sumación de Poisson)**.** *Para f suficientemente buena (por ejemplo, de la clase de Schwartz),*

$$\sum_{n\in\mathbb{Z}} f(n) = \sum_{k\in\mathbb{Z}} \hat{f}(k).$$

Demostración. Periodícese f: sea $F(x)=\sum_n f(x+n)$, que es 1-periódica. Sus coeficientes de Fourier son

$$c_k = \int_0^1 F(x)e^{-2\pi ikx}\mathrm{d}x = \sum_n \int_0^1 f(x+n)e^{-2\pi ikx}\mathrm{d}x = \int_{\mathbb{R}} f(x)e^{-2\pi ikx}\mathrm{d}x = \hat{f}(k).$$

Evaluando la serie de Fourier de F en $x=0$ se obtiene el enunciado. □

Tres líneas, y de ellas cuelga media matemática. Es la afirmación de que el peine de Dirac es su propia transformada, escrita sin distribuciones.

Ejemplo 33.50 (La identidad theta, comprobada)**.** Aplicando Poisson a la gaussiana $f(x)=e^{-\pi x^2 t}$, cuya transformada es $t^{-1/2}e^{-\pi\xi^2/t}$:

$$\vartheta(t) = \sum_{n\in\mathbb{Z}} e^{-\pi n^2 t} = \frac{1}{\sqrt{t}} \sum_{k\in\mathbb{Z}} e^{-\pi k^2/t} = \frac{1}{\sqrt{t}}\,\vartheta(1/t).$$

Con $t=2$ y cuatro términos por lado:

$$\text{izquierda} = 1{,}003\,734\,885\,5, \qquad \text{derecha} = 1{,}003\,734\,885\,5.$$

Coinciden hasta la décima cifra. Y esa identidad es exactamente la que, pasada por la transformada de Mellin, produce la ecuación funcional de la función zeta: la simetría $t\leftrightarrow 1/t$ del calor se convierte en la simetría $s\leftrightarrow 1-s$ de ζ.

Observación 33.51 (El teorema del muestreo, otra vez)*.* Aplicando Poisson a $f(x)e^{-2\pi ix\xi}$ con paso T se obtiene

$$\sum_k \hat{f}\Big(\xi - \frac{k}{T}\Big) = T\sum_n f(nT)e^{-2\pi inT\xi}\,:$$

muestrear en el tiempo *periodifica* en frecuencia. Si $\hat{f}$ se anula fuera de $[-B,B]$ y $T\le 1/(2B)$, las copias no se solapan y se puede recuperar f quedándose con una de ellas. Ese es el teorema de Nyquist–Shannon, y el solapamiento cuando T es demasiado grande es el aliasing. Todo el tratamiento digital de señales cabe en un corolario de tres líneas.

Fourier como teoría de representaciones

Definición 33.52 (Caracteres y dualidad de Pontriaguin)**.** Un **carácter** de un grupo abeliano localmente compacto G es un homomorfismo continuo $\chi: G\to S^1$. El conjunto $\hat{G}$ de caracteres es a su vez un grupo localmente compacto, y la transformada de Fourier es

$$\hat{f}(\chi) = \int_G f(x)\,\overline{\chi(x)}\,\mathrm{d}x$$

respecto de la medida de Haar. El teorema de Pontriaguin dice que $\hat{\hat{G}} \cong G$.

Y el cuadro que unifica todo el capítulo cabe en cuatro renglones:

G	$\hat{G}$	qué es la transformada
$\mathbb{T} = \mathbb{R}/\mathbb{Z}$	$\mathbb{Z}$	serie de Fourier
$\mathbb{Z}$	$\mathbb{T}$	serie de Fourier del dual
$\mathbb{R}$	$\mathbb{R}$	transformada integral
$\mathbb{Z}/N$	$\mathbb{Z}/N$	DFT

Los caracteres de $\mathbb{T}$ son $x \mapsto e^{2\pi i n x}$, indexados por $\mathbb{Z}$: por eso los coeficientes de Fourier son una sucesión. Los de $\mathbb{Z}/N$ son $n \mapsto \omega^{nk}$: por eso la DFT es un cambio de base finito. Y los caracteres de Dirichlet del capítulo de teoría analítica son los caracteres del grupo $(\mathbb{Z}/q)^\times$: la ortogonalidad que allí se usaba para aislar progresiones aritméticas *es* la ortogonalidad de caracteres de este capítulo.

Observación 33.53 (Varias variables). En $\mathbb{R}^n$ todo lo anterior vale con $\xi \cdot x$ en el exponente. Dos hechos nuevos merecen mención. Primero, la transformada conmuta con las rotaciones, luego manda funciones radiales en funciones radiales, y el resultado se expresa con funciones de Bessel: la transformada de la indicatriz del disco unidad del plano es $J_1(2\pi|\xi|)/|\xi|$, el análogo bidimensional del sinc, con sus ceros en los ceros de J_1. Ese patrón es el disco de Airy que limita la resolución de todo telescopio.

Segundo, la **transformada de Radon**: integrar f a lo largo de todas las rectas del plano. El *teorema de la sección central* dice que la transformada de Fourier unidimensional de la proyección en una dirección es la restricción de $\hat{f}$ a la recta correspondiente. Es decir: midiendo proyecciones desde muchos ángulos se llena $\hat{f}$ y se invierte. Eso es exactamente lo que hace un escáner de tomografía, y por eso Cormack y Hounsfield recibieron el Nobel de Medicina en 1979 por un teorema de análisis armónico.

33.11. El núcleo de Dirichlet y la constante de Lebesgue

Se ha hablado varias veces del núcleo de Dirichlet sin escribirlo. Es hora, porque de su fórmula sale tanto el teorema de convergencia puntual como la razón profunda de por qué la convergencia puntual falla para funciones continuas.

Definición 33.54 (Núcleo de Dirichlet)**.** Para $n \geq 0$,

$$D_n(t) = \sum_{k=-n}^{n} \mathrm{e}^{\mathrm{i}kt}.$$

Proposición 33.55 (Fórmula cerrada)**.** *Para $t \notin 2\pi\mathbb{Z}$,*

$$D_n(t) = \frac{\sin\big((n+\tfrac{1}{2})t\big)}{\sin\big(\tfrac{t}{2}\big)}, \qquad y \qquad D_n(0) = 2n+1.$$

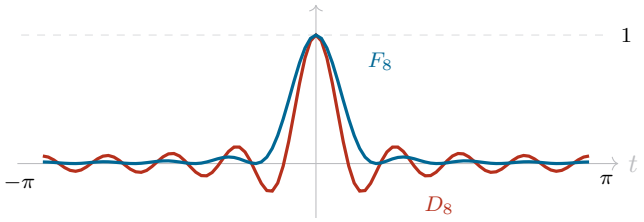


Figura 33.5: Los dos núcleos con $n = 8$, escalados a altura 1 en el origen. El de Dirichlet (rojo) cambia de signo una y otra vez, y son sus lóbulos negativos los que hacen divergir a $\int |D_n|$. El de Fejér (azul) nunca baja de cero, y por eso su norma L^1 es constante. Toda la diferencia entre que la serie diverja y que sus medias converjan está en ese cambio de signo.

Además D_n es par, real, y $\dfrac{1}{2\pi}\displaystyle\int_{-\pi}^{\pi} D_n(t)\,\mathrm{d}t = 1.$

Demostración. Es una suma geométrica de razón $\mathrm{e}^{\mathrm{i}t} \neq 1$:

$$D_n(t) = \mathrm{e}^{-\mathrm{i}nt}\sum_{j=0}^{2n}\mathrm{e}^{\mathrm{i}jt} = \mathrm{e}^{-\mathrm{i}nt}\,\frac{\mathrm{e}^{\mathrm{i}(2n+1)t}-1}{\mathrm{e}^{\mathrm{i}t}-1} = \frac{\mathrm{e}^{\mathrm{i}(n+1)t}-\mathrm{e}^{-\mathrm{i}nt}}{\mathrm{e}^{\mathrm{i}t}-1}.$$

Multiplíquense numerador y denominador por $\mathrm{e}^{-\mathrm{i}t/2}$:

$$D_n(t) = \frac{\mathrm{e}^{\mathrm{i}(n+\frac12)t}-\mathrm{e}^{-\mathrm{i}(n+\frac12)t}}{\mathrm{e}^{\mathrm{i}t/2}-\mathrm{e}^{-\mathrm{i}t/2}} = \frac{2\mathrm{i}\sin\!\left((n+\frac12)t\right)}{2\mathrm{i}\sin(t/2)}.$$

El valor en 0 es la suma de $2n+1$ unos. La integral se calcula término a término: sólo el término $k=0$ contribuye, y aporta 2π. □

Proposición 33.56 (Las sumas parciales son una convolución)**.** *Si* $\widehat{f}(k) = \frac{1}{2\pi}\int_{-\pi}^{\pi} f(t)\mathrm{e}^{-\mathrm{i}kt}\mathrm{d}t$ *y* $S_n f(x) = \sum_{|k|\le n}\widehat{f}(k)\mathrm{e}^{\mathrm{i}kx}$*, entonces*

$$S_n f(x) = \frac{1}{2\pi}\int_{-\pi}^{\pi} f(t)\,D_n(x-t)\,\mathrm{d}t = \frac{1}{2\pi}\int_{-\pi}^{\pi} f(x-t)\,D_n(t)\,\mathrm{d}t.$$

Demostración. Sustitúyase la definición de $\widehat{f}(k)$ e intercámbiense suma finita e integral: $S_n f(x) = \frac{1}{2\pi}\int f(t)\sum_{|k|\le n}\mathrm{e}^{\mathrm{i}k(x-t)}\mathrm{d}t$, y la suma es $D_n(x-t)$. La segunda igualdad es el cambio $t \mapsto x-t$ y la periodicidad. □

Todo el problema de la convergencia puntual está, pues, en cómo se comporta D_n. Y se comporta mal.

Teorema 33.57 (Las constantes de Lebesgue divergen)**.** *Sea* $L_n = \dfrac{1}{2\pi}\displaystyle\int_{-\pi}^{\pi}|D_n(t)|\,\mathrm{d}t$*. Entonces*

$$L_n \;\ge\; \frac{4}{\pi^2}\sum_{k=1}^{n}\frac{1}{k} \;\ge\; \frac{4}{\pi^2}\log n \;\xrightarrow[n\to\infty]{}\; \infty.$$

Demostración. Por paridad, $L_n = \frac{1}{\pi}\int_0^\pi \frac{|\sin((n+\frac12)t)|}{|\sin(t/2)|}\mathrm{d}t$. Como $|\sin(t/2)| \le t/2$ para $t \ge 0$,

$$L_n \;\ge\; \frac{1}{\pi}\int_0^\pi \frac{|\sin((n+\frac12)t)|}{t/2}\,\mathrm{d}t = \frac{2}{\pi}\int_0^\pi \frac{|\sin((n+\frac12)t)|}{t}\,\mathrm{d}t.$$

Con la sustitución $u = (n+\frac12)t$ el integrando se vuelve $|\sin u|/u$ y el intervalo pasa a ser $[0,(n+\frac12)\pi]$, que contiene a $[0,n\pi]$:

$$L_n \;\ge\; \frac{2}{\pi}\int_0^{n\pi} \frac{|\sin u|}{u}\,\mathrm{d}u = \frac{2}{\pi}\sum_{k=1}^{n}\int_{(k-1)\pi}^{k\pi} \frac{|\sin u|}{u}\,\mathrm{d}u.$$

En el k-ésimo intervalo, $u \le k\pi$, luego $\frac{1}{u} \ge \frac{1}{k\pi}$, y $\int_{(k-1)\pi}^{k\pi}|\sin u|\mathrm{d}u = 2$. Por tanto

$$L_n \;\ge\; \frac{2}{\pi}\sum_{k=1}^{n}\frac{2}{k\pi} = \frac{4}{\pi^2}\sum_{k=1}^{n}\frac{1}{k},$$

y la serie armónica supera a $\log n$. □

Ejemplo 33.58 (Los valores, calculados)**.** La cota es fina: el comportamiento verdadero es $L_n = \frac{4}{\pi^2}\log n + c + O(1/n)$ con $c \approx 1{,}270$. Numéricamente:

n	1	10	50	100	500	1000
L_n	1,436	2,223	2,860	3,139	3,790	4,071

Divergen, sí, pero de forma exasperantemente lenta: hacen falta $n \approx 10^{10}$ para llegar a $L_n = 10$. Esa lentitud explica por qué el fenómeno pasó desapercibido durante décadas y por qué en la práctica numérica las sumas parciales suelen portarse bien.

Teorema 33.59 (Existe una función continua con serie de Fourier divergente)**.** *Existe f continua y 2π-periódica tal que* $\sup_n |S_n f(0)| = \infty$.

Demostración. Considérense los funcionales $\Lambda_n : C(\mathbb{T}) \to \mathbb{C}$ dados por $\Lambda_n(f) = S_n f(0) = \frac{1}{2\pi}\int_{-\pi}^{\pi} f(-t)D_n(t)\mathrm{d}t$. Cada uno es lineal y continuo con $\|\Lambda_n\| = L_n$: la cota $\le$ es inmediata, y la igualdad se obtiene aproximando uniformemente el signo de D_n —que es una función escalonada de ±1— por funciones continuas de norma 1. Por el teorema 33.57, $\sup_n \|\Lambda_n\| = \infty$. El principio de acotación uniforme (Banach–Steinhaus, capítulo 32) dice entonces que no puede ocurrir que $\sup_n |\Lambda_n(f)| < \infty$ para toda f: existe alguna f continua con $\sup_n |S_n f(0)| = \infty$. □

Observación 33.60 (Qué se ha demostrado y qué no)*.* Esto es una demostración de existencia, no una construcción: no exhibe ninguna función concreta. Se pueden construir explícitamente (du Bois-Reymond lo hizo en 1873, sumando bloques de polinomios trigonométricos con amplitudes cuidadosamente elegidas), pero el argumento funcional es más corto y más informativo, porque señala la causa: $\|D_n\|_{L^1} \to \infty$. Y señala también el remedio: sustituir D_n por un núcleo cuya norma L^1 esté acotada. Ese núcleo existe y es el de Fejér.

33.12. Convergencia puntual: Riemann–Lebesgue, Dini y el teorema de Dirichlet

El teorema de Dirichlet se ha citado varias veces sin enunciarlo. Se enuncia y se demuestra ahora, y de paso queda claro qué hipótesis hace falta y por qué el límite es el promedio de los límites laterales.

Lema 33.61 (Riemann–Lebesgue)**.** *Si $g \in L^1(a,b)$, entonces $\displaystyle\int_a^b g(t)\sin(\lambda t)\,\mathrm{d}t \to 0$ cuando $\lambda \to \infty$, e igual con* cos.

Demostración. Si $g = \mathbf{1}_{[c,d]}$, entonces $\int_c^d \sin(\lambda t)\mathrm{d}t = \frac{\cos(\lambda c)-\cos(\lambda d)}{\lambda}$, cuyo módulo es $\le 2/\lambda \to 0$. Por linealidad el resultado vale para funciones escalonadas. Para $g \in L^1$ general y $\varepsilon > 0$, tómese una escalonada s con $\|g-s\|_{L^1} < \varepsilon/2$ (las escalonadas son densas en L^1, capítulo 31); entonces

$$\left|\int g\sin(\lambda t)\right| \le \left|\int (g-s)\sin(\lambda t)\right| + \left|\int s\sin(\lambda t)\right| \le \|g-s\|_{L^1} + \left|\int s\sin(\lambda t)\right| < \varepsilon$$

para λ suficientemente grande. □

Teorema 33.62 (Criterio de Dini)**.** *Sea $f \in L^1(\mathbb{T})$, x un punto y $s \in \mathbb{C}$. Escríbase*

$$\psi_x(t) = f(x+t) + f(x-t) - 2s.$$

Si $\displaystyle\int_0^\pi \frac{|\psi_x(t)|}{t}\,\mathrm{d}t < \infty$, entonces $S_n f(x) \to s$.

Demostración. Como $\frac{1}{2\pi}\int_{-\pi}^{\pi} D_n = 1$,

$$S_n f(x) - s = \frac{1}{2\pi}\int_{-\pi}^{\pi} \big[f(x-t)-s\big]D_n(t)\,\mathrm{d}t.$$

El núcleo es par; agrupando t con $-t$,

$$S_n f(x) - s = \frac{1}{2\pi}\int_0^\pi \big[f(x+t)+f(x-t)-2s\big]D_n(t)\,\mathrm{d}t = \frac{1}{2\pi}\int_0^\pi \frac{\psi_x(t)}{\sin(t/2)}\,\sin\Big((n+\tfrac12)t\Big)\mathrm{d}t.$$

Llámese $g(t) = \psi_x(t)/\sin(t/2)$. Cerca de 0 se tiene $\sin(t/2) \sim t/2$, de modo que $|g(t)| \le \frac{\pi|\psi_x(t)|}{t}$ en $(0,\pi]$ (pues $\sin(t/2) \ge t/\pi$ allí), y la hipótesis de Dini dice justamente que ese mayorante es integrable; lejos de 0 el denominador está acotado inferiormente y g es integrable porque f lo es. Así $g \in L^1(0,\pi)$, y el lema 33.61 con $\lambda = n + \frac12$ da que la integral tiende a 0. □

Teorema 33.63 (Dirichlet)**.** *Sea f 2π-periódica e integrable, y supóngase que en el punto x existen los límites laterales $f(x^\pm)$ y las derivadas laterales*

$$\lim_{t\to 0^+}\frac{f(x+t)-f(x^+)}{t}, \qquad \lim_{t\to 0^+}\frac{f(x-t)-f(x^-)}{t}.$$

Entonces

$$S_n f(x) \;\longrightarrow\; \frac{f(x^+)+f(x^-)}{2}.$$

En particular, si f es continua y derivable a trozos, la serie converge a $f(x)$ en todo punto de continuidad, y al promedio de los saltos en las discontinuidades.

Demostración. Tómese $s = \frac{f(x^+)+f(x^-)}{2}$. Entonces

$$\psi_x(t) = [f(x+t) - f(x^+)] + [f(x-t) - f(x^-)],$$

y por hipótesis cada corchete dividido entre t tiene límite finito cuando $t \to 0^+$. Por tanto $|\psi_x(t)|/t$ está acotada en un entorno $(0,\delta)$ del origen, luego es integrable allí; en $[\delta,\pi]$ es integrable porque f lo es y $1/t \le 1/\delta$. Se cumple la hipótesis de Dini y el teorema 33.62 concluye. □

Ejemplo 33.64 (La onda cuadrada, y por qué vale 0 en el salto)**.** Sea $f(x) = 1$ en $(0,\pi)$ y $f(x) = -1$ en $(-\pi,0)$. Su serie es $\frac{4}{\pi}\sum_{m\ge 0}\frac{\sin((2m+1)x)}{2m+1}$. En $x = 0$ todos los términos valen 0, luego la serie converge a 0; y en efecto $\frac{f(0^+)+f(0^-)}{2} = \frac{1+(-1)}{2} = 0$. El teorema no está haciendo una elección arbitraria en los saltos: la simetría del núcleo de Dirichlet, que es par, obliga a que la suma parcial "vea" los dos lados por igual. Evaluando en $x = \pi/2$ se obtiene $1 = \frac{4}{\pi}(1 - \frac{1}{3} + \frac{1}{5} - \cdots)$, es decir la serie de Leibniz para $\pi/4$: un subproducto gratuito.

Cuidado

[El fenómeno de Gibbs sobrevive al teorema] El teorema de Dirichlet garantiza convergencia *puntual*, y nada más. En cada punto fijo la suma parcial se acerca al valor correcto; pero cerca del salto, las sumas parciales sobrepasan el valor en aproximadamente un 8,95 % de la altura del salto, y ese exceso *no* disminuye al crecer n: sólo se estrecha, desplazándose hacia el punto de discontinuidad. Convergencia puntual sin convergencia uniforme, en estado puro. El promotor del exceso es, otra vez, la oscilación de D_n: el término $\int_0^\pi \frac{\sin u}{u}\mathrm{d}u = 1{,}8519$, frente a $\frac{\pi}{2} = 1{,}5708$, y su cociente 1,0895 es el 8,95 % del enunciado.

33.13. Fejér, demostrado, y sus consecuencias

La cura para la divergencia del apartado anterior es no sumar la serie, sino promediar sus sumas parciales.

Definición 33.65 (Medias de Cesàro y núcleo de Fejér)**.** $\sigma_n f = \dfrac{S_0 f + S_1 f + \cdots + S_n f}{n+1}$, y el *núcleo de Fejér* es $F_n = \frac{D_0+\cdots+D_n}{n+1}$, de modo que $\sigma_n f(x) = \frac{1}{2\pi}\int_{-\pi}^{\pi} f(x-t)F_n(t)\mathrm{d}t$.

Proposición 33.66 (Fórmula cerrada y positividad)**.** *Para $t \notin 2\pi\mathbb{Z}$,*

$$F_n(t) = \frac{1}{n+1}\left(\frac{\sin\left(\frac{(n+1)t}{2}\right)}{\sin\left(\frac{t}{2}\right)}\right)^2 \;\ge\; 0,$$

con $\frac{1}{2\pi}\int_{-\pi}^{\pi} F_n = 1$ y, para cada $\delta > 0$, $\max\limits_{\delta\le|t|\le\pi} F_n(t) \le \dfrac{1}{(n+1)\sin^2(\delta/2)} \to 0$.

Demostración. Multiplíquese $\sum_{k=0}^{n}\sin\big((k+\frac12)t\big)$ por $2\sin(t/2)$ y úsese $2\sin A\sin B=\cos(A-B)-\cos(A+B)$:

$$2\sin\tfrac{t}{2}\sum_{k=0}^{n}\sin\big((k+\tfrac12)t\big)=\sum_{k=0}^{n}\big[\cos(kt)-\cos((k+1)t)\big]=1-\cos((n+1)t),$$

suma telescópica. Como $1-\cos\theta=2\sin^2(\theta/2)$, queda $\sum_{k=0}^{n}\sin\big((k+\frac12)t\big)=\dfrac{\sin^2\big(\frac{(n+1)t}{2}\big)}{\sin(t/2)}$. Dividiendo entre $(n+1)\sin(t/2)$ y usando la proposición 33.55 se obtiene la fórmula. La positividad es evidente por ser un cuadrado; la integral vale 1 porque cada D_k tiene integral 1; y la última cota sale de acotar el numerador por 1 y el denominador por $\sin^2(\delta/2)$. □

Observación 33.67 (Aquí está toda la diferencia). $\int|D_n|\to\infty$, pero $\int|F_n|=\int F_n=2\pi$ para todo n, porque $F_n\ge0$. Un núcleo que no cambia de signo no puede tener norma L^1 mayor que su integral, y por eso el promedio de Cesàro se comporta bien donde la suma parcial se comporta mal. Toda la teoría de núcleos sumatorios —Abel, Poisson, Gauss–Weierstrass— gira alrededor de esta observación.

Teorema 33.68 (Fejér)**.** *Si f es continua y 2π-periódica, entonces $\sigma_n f\to f$ uniformemente. Si $f\in L^p(\mathbb{T})$ con $1\le p<\infty$, entonces $\sigma_n f\to f$ en L^p. Y si f es integrable y en el punto x existen los límites laterales, entonces $\sigma_n f(x)\to\frac{f(x^+)+f(x^-)}{2}$.*

Demostración. (Caso continuo.) Como $\frac{1}{2\pi}\int F_n=1$,

$$\sigma_n f(x)-f(x)=\frac{1}{2\pi}\int_{-\pi}^{\pi}\big[f(x-t)-f(x)\big]F_n(t)\,\mathrm{d}t.$$

Sea $\varepsilon>0$. Por ser f continua sobre un compacto, es uniformemente continua: existe $\delta>0$ con $|f(x-t)-f(x)|<\varepsilon/2$ para $|t|<\delta$ y todo x. Pártase la integral:

$$|\sigma_n f(x)-f(x)|\le\underbrace{\frac{1}{2\pi}\int_{|t|<\delta}\frac{\varepsilon}{2}F_n(t)\mathrm{d}t}_{\le\varepsilon/2}+\underbrace{\frac{1}{2\pi}\int_{\delta\le|t|\le\pi}2\|f\|_\infty F_n(t)\mathrm{d}t}_{\le\,2\|f\|_\infty\cdot\frac{1}{(n+1)\sin^2(\delta/2)}}.$$

En el primer trozo se ha usado $F_n\ge0$ y $\frac{1}{2\pi}\int F_n=1$; en el segundo, la cota uniforme de la proposición 33.66. El segundo sumando es menor que $\varepsilon/2$ para n grande, y la cota no depende de x: la convergencia es uniforme.

(Caso L^p.) Idéntico, sustituyendo la continuidad uniforme por la continuidad de las traslaciones en L^p: $\|f(\cdot-t)-f\|_{L^p}\to0$ cuando $t\to0$, y aplicando la desigualdad integral de Minkowski. (Caso puntual.) Idéntico al de Dini, usando que F_n es par y concentra su masa cerca de 0. □

Corolario 33.69 (Weierstrass trigonométrico)**.** *Los polinomios trigonométricos son densos en $C(\mathbb{T})$ con la norma del supremo.*

Demostración. $\sigma_n f$ es un polinomio trigonométrico (combinación finita de $\mathrm{e}^{\mathrm{i}kx}$ con $|k|\le n$) y converge uniformemente a f. □

Corolario 33.70 (Unicidad de los coeficientes)**.** *Si f es integrable y $\widehat{f}(k) = 0$ para todo $k \in \mathbb{Z}$, entonces $f = 0$ en casi todo punto. En consecuencia, dos funciones integrables con los mismos coeficientes de Fourier coinciden c.t.p.*

Demostración. Si todos los coeficientes son nulos, todas las sumas parciales son nulas, luego todas las medias $\sigma_n f$ son nulas; y $\sigma_n f \to f$ en L^1 por el teorema 33.68, así que $\|f\|_{L^1} = 0$. □

Corolario 33.71 (Si converge, converge a lo que debe)**.** *Si f es continua en x y la sucesión $S_n f(x)$ converge, su límite es $f(x)$.*

Demostración. Si una sucesión converge, sus medias de Cesàro convergen al mismo límite; y las medias convergen a $f(x)$ por Fejér. □

Idea clave

El mapa completo de la convergencia de series de Fourier. Conviene tenerlo junto, porque es fácil confundir los enunciados.

- *f continua:* $S_n f$ puede divergir en un punto (sección anterior); $\sigma_n f$ converge uniformemente (Fejér).
- *f continua y derivable a trozos:* $S_n f \to f$ puntualmente (Dirichlet), y uniformemente si además f es continua en todo punto.
- *$f \in L^2$:* $S_n f \to f$ en L^2, siempre, sin hipótesis extra: es Parseval y la teoría de Hilbert, y no necesita ningún núcleo.
- *$f \in L^p$, $1 < p < \infty$:* $S_n f \to f$ en L^p (teorema de M. Riesz, vía la transformada de Hilbert). Para $p = 1$ es *falso*.
- *$f \in L^2$, convergencia puntual:* $S_n f \to f$ en casi todo punto. Es el teorema de Carleson (1966), uno de los resultados más difíciles del análisis del siglo XX, y responde afirmativamente a una conjetura que Luzin planteó en 1915.
- *$f \in L^1$:* puede divergir en *todo* punto (Kolmogórov, 1926).

La moraleja operativa: si se necesita convergencia en norma, úsese L^2; si se necesita convergencia uniforme y sólo hay continuidad, úsese Fejér; y si se necesita convergencia puntual, compruébese Dini.

33.14. Calcular una serie de Fourier y decidir cómo converge

Los ingredientes están repartidos por todo el capítulo —paridad al principio, Dirichlet y Fejér casi mil líneas después— y los ejercicios propuestos piden exactamente el proceso completo. Aquí está reunido.

Regla: De la función a la serie, y de la serie a la convergencia

Primera parte: calcular los coeficientes.

1. **Fija el periodo y el intervalo.** Si la función se da en $[-L, L]$, el periodo es $2L$ y las frecuencias son $\frac{n\pi}{L}$. Si se da en $[0, L]$ y se quiere una serie *de senos*, extiéndela de forma impar; para una de *cosenos*, de forma par.

2. **Explota la paridad antes de integrar.** Si f es par, todos los b_n son cero y basta calcular $a_n = \frac{2}{L}\int_0^L f\cos\frac{n\pi x}{L}$. Si es impar, todos los a_n son cero. Esto ahorra la mitad del trabajo y evita la mitad de los errores.

3. **Integra por partes cuando haya polinomios**, y usa que $\int_{-L}^{L}\operatorname{sin}\frac{n\pi x}{L}\cos\frac{m\pi x}{L} = 0$ siempre.

4. **Trata a_0 aparte:** es el doble del valor medio, $a_0 = \frac{1}{L}\int_{-L}^{L} f$, y en la serie aparece como $a_0/2$.

Segunda parte: decidir en qué sentido converge.

5. **¿$f \in L^2$?** Entonces la serie converge a f *en media cuadrática*, siempre, sin ninguna hipótesis más, y vale la identidad de Parseval $\frac{a_0^2}{2} + \sum(a_n^2 + b_n^2) = \frac{1}{L}\int_{-L}^{L}|f|^2$. Ésta es la respuesta por defecto.

6. **¿f es derivable a trozos?** Entonces converge *puntualmente* en cada x al promedio $\frac{f(x^+)+f(x^-)}{2}$, por el teorema de Dirichlet. En los puntos de continuidad, a $f(x)$.

7. **¿f es continua y periódica, pero no se sabe más?** Entonces las sumas parciales pueden divergir, pero las *medias de Cesàro* convergen *uniformemente*, por Fejér.

8. **¿f es continua, periódica y derivable a trozos?** Entonces la convergencia es *uniforme* en todo el intervalo, y además $\sum(|a_n| + |b_n|) < \infty$.

9. **Si hay un salto, espera Gibbs.** Cerca de la discontinuidad las sumas parciales sobrepasan el salto en un 8,95 % de su altura, y ese exceso no desaparece al aumentar n: sólo se estrecha. Por eso no hay convergencia uniforme en un entorno del salto, por muy buena que sea la función en el resto.

10. **Comprueba con un valor.** Evaluar la serie en un punto cómodo suele producir una identidad numérica conocida —$\pi^2/6$, $\pi/4$— y es la mejor verificación de que los coeficientes están bien.

Ejemplo 33.72 (La serie de $f(x) = |x|$, de principio a fin)**.** *Paso 2:* la función es *par*, luego todos los $b_n = 0$ y sólo hay cosenos. *Paso 4:* $a_0 = \frac{1}{\pi}\int_{-\pi}^{\pi}|x|\mathrm{d}x = \frac{2}{\pi}\cdot\frac{\pi^2}{2} = \pi$. *Paso 3:* integrando por partes,

$$a_n = \frac{2}{\pi}\int_0^{\pi} x\cos(nx)\,\mathrm{d}x = \frac{2}{\pi}\left[\frac{x\sin nx}{n} + \frac{\cos nx}{n^2}\right]_0^{\pi} = \frac{2\big((-1)^n - 1\big)}{\pi n^2},$$

que vale 0 si n es par y $-\frac{4}{\pi n^2}$ si es impar. Por tanto

$$|x| = \frac{\pi}{2} - \frac{4}{\pi}\sum_{k\geq 0}\frac{\cos\big((2k+1)x\big)}{(2k+1)^2}.$$

Pasos 5 a 8: $|x|$ es continua, periódica al extenderla y derivable a trozos, luego la convergencia es *uniforme*; y en efecto $\sum 1/(2k+1)^2$ converge, lo que confirma el criterio del paso 9. No hay saltos, así que no hay Gibbs. *Paso 11:* evaluando en $x=0$,

$$0 = \frac{\pi}{2} - \frac{4}{\pi}\sum_{k\geq 0}\frac{1}{(2k+1)^2} \quad\Longrightarrow\quad \sum_{k\geq 0}\frac{1}{(2k+1)^2} = \frac{\pi^2}{8},$$

identidad correcta, de la que se deduce $\sum_{n\geq 1} 1/n^2 = \pi^2/6$ separando pares e impares.

Definición 33.73 (Fenómeno de Gibbs y aliasing)**.** El *fenómeno de Gibbs* es el sobreimpulso que presentan las sumas parciales de Fourier junto a una discontinuidad de salto: para toda función derivable a trozos con un salto de altura h en x_0,

$$\lim_{n\to\infty}\Big(\max_{x \text{ cerca de } x_0} S_n f(x)\Big) - f(x_0^+) = h\left(\frac{1}{\pi}\int_0^{\pi}\frac{\sin u}{u}\,\mathrm{d}u - \frac{1}{2}\right) \approx 0{,}0895\,h,$$

un 8,95 % del salto, *independiente de n y de la función*. Lo que sí depende de n es la anchura de la zona afectada, que decrece como $1/n$.

El *aliasing* es el fenómeno por el cual, al muestrear una señal a intervalos T, dos frecuencias f_1 y f_2 con $f_1 - f_2 = k/T$ producen *exactamente* las mismas muestras y se vuelven indistinguibles. El teorema de muestreo de Nyquist–Shannon dice que se evita si y sólo si la señal no contiene frecuencias por encima de $\frac{1}{2T}$, la *frecuencia de Nyquist*.

Observación 33.74 (Los dos son la misma advertencia con distinto disfraz)*.* Gibbs dice que una suma *finita* de senos no puede reproducir un salto; aliasing, que un muestreo *finito* no puede distinguir frecuencias arbitrariamente altas. En ambos casos el problema es que se pretende capturar información de frecuencia ilimitada con recursos limitados, y en ambos casos el remedio es el mismo: suavizar antes. Contra Gibbs se multiplican los coeficientes por una ventana que decae —lo que equivale a sumar con Fejér en vez de con Dirichlet—; contra el aliasing se aplica un filtro paso bajo *antes* de muestrear, no después, porque una vez que las frecuencias se han confundido ya no hay forma de separarlas.

33.15. Ejercicios resueltos

Tres problemas que van del cálculo de coeficientes a lo que la serie sabe y la función no.

Ejercicio 33.1. Halle la serie de Fourier de $f(x) = x$ en $(-\pi, \pi)$ extendida por periodicidad, y úsela para sumar una serie numérica.

Solución. La función es impar, luego todos los a_n se anulan y

$$b_n = \frac{2}{\pi}\int_0^{\pi} x\sin nx\,\mathrm{d}x = \frac{2(-1)^{n+1}}{n},$$

de modo que $f(x) \sim 2\sum_{n\geq 1} \frac{(-1)^{n+1}}{n} \sin nx$.

Evaluando en $x = \pi/2$, donde f es continua y vale $\pi/2$, queda $\frac{\pi}{2} = 2\left(1 - \frac{1}{3} + \frac{1}{5} - \cdots\right)$, es decir, la serie de Leibniz. Aplicando Parseval en cambio, $\frac{1}{\pi}\int_{-\pi}^{\pi} x^2 \mathrm{d}x = \sum b_n^2$ da $\frac{2\pi^2}{3} = 4\sum \frac{1}{n^2}$, y de ahí $\sum \frac{1}{n^2} = \frac{\pi^2}{6}$: el problema de Basilea resuelto de pasada. ■

Ejercicio 33.2. Explique por qué la serie de Fourier de la onda cuadrada converge puntualmente pero no uniformemente, y qué tiene que ver con el fenómeno de Gibbs.

Solución. En todo punto de continuidad la serie converge al valor de la función, y en el salto converge al promedio de los límites laterales: eso es el teorema de Dirichlet.

La convergencia no puede ser uniforme por una razón elemental: cada suma parcial es una función continua, y el límite uniforme de funciones continuas es continuo. Como el límite tiene un salto, la convergencia no es uniforme. El fenómeno de Gibbs es la manifestación cuantitativa de ese fallo: junto al salto, las sumas parciales sobrepasan el valor correcto en aproximadamente un 8,9 por ciento de la altura del salto, y ese exceso *no disminuye* al aumentar N; sólo se estrecha el intervalo donde ocurre. La cifra exacta es $\frac{2}{\pi}\int_0^{\pi} \frac{\sin t}{t}\mathrm{d}t - 1 = 0{,}0894$. ■

Ejercicio 33.3. Calcule la transformada de Fourier de la gaussiana $f(x) = \mathrm{e}^{-\pi x^2}$ y comente el principio de incertidumbre.

Solución. Con la normalización $\hat{f}(\xi) = \int f(x)\mathrm{e}^{-2\pi i x\xi}\mathrm{d}x$, completar el cuadrado en el exponente da $\hat{f}(\xi) = \mathrm{e}^{-\pi\xi^2}$: la gaussiana es su propia transformada.

El principio de incertidumbre dice que una función y su transformada no pueden estar ambas muy concentradas: en la formulación de Hardy, si $f(x) = O(\mathrm{e}^{-\pi a x^2})$ y $\hat{f}(\xi) = O(\mathrm{e}^{-\pi\xi^2/a})$ con $a > 1$, entonces $f \equiv 0$. Comprimir en tiempo es ensanchar en frecuencia, y la gaussiana es el caso de igualdad, el compromiso óptimo. En física cuántica ése es el enunciado de Heisenberg; aquí es un teorema de análisis, y el hecho de que sean el mismo enunciado no es casualidad, porque posición y momento son variables conjugadas por Fourier. ■

33.16. Ejercicios propuestos

33.1. Halle los coeficientes de Fourier de $f(x) = |x|$ en $(-\pi, \pi)$.

33.2. ¿Por qué una función par sólo tiene cosenos y una impar sólo senos?

33.3. Enuncie la identidad de Parseval y úsela para sumar $\sum 1/(2k+1)^2$.

33.4. ¿Qué relación hay entre la suavidad de f y la velocidad de decaimiento de sus coeficientes?

33.5. ¿Converge la serie de Fourier de toda función continua en todo punto? Cite a Kolmogorov y a Carleson.

33.6. Calcule la transformada de Fourier de la indicatriz de $[-1, 1]$.

33.7. ¿Qué le ocurre a la transformada al trasladar la función? ¿Y al dilatarla?

33.8. Enuncie el teorema de convolución y diga para qué se usa en la práctica.

33.9. ¿Qué dice el teorema de muestreo de Nyquist y Shannon, y qué es el aliasing?

33.10. ¿Por qué la transformada rápida de Fourier cuesta $O(n \log n)$ y no $O(n^2)$?

33.11. ¿Qué es el núcleo de Dirichlet y por qué su norma L^1 crece como $\log n$?

33.12. Explique en qué sentido el análisis de Fourier es teoría de representaciones del círculo.

Capítulo 34

Teoría analítica de números

Lo que necesitas traer

Del capítulo 3: primos, divisibilidad y funciones aritméticas. Del capítulo 30: funciones holomorfas, continuación analítica, polos y el teorema de los residuos. Del capítulo 6: series, productos infinitos y criterios de convergencia.

Los números primos son objetos discretos y su distribución parece caótica; sin embargo, las herramientas que mejor la describen son continuas: integrales, series y funciones de variable compleja. Esa aparente incongruencia es la marca de la teoría analítica de números, y su resultado central, el teorema de los números primos, es un enunciado sobre enteros cuya demostración pasa inevitablemente por el plano complejo.

34.1. La función zeta y la distribución de los primos

Definición 34.1. La *función zeta de Riemann* es $\zeta(s) = \sum_{n\geq 1} n^{-s}$ para $\operatorname{Re}(s) > 1$, y admite el *producto de Euler*

$$\zeta(s) = \prod_{p \text{ primo}} \frac{1}{1-p^{-s}}.$$

Si $\pi(x)$ cuenta los primos menores o iguales que x, el *teorema de los números primos* afirma que $\pi(x) \sim \dfrac{x}{\ln x}$.

Idea clave

Qué dice esto en cristiano. El producto de Euler es la factorización única escrita como identidad analítica: desarrollar cada factor en serie geométrica y multiplicar reproduce cada n^{-s} exactamente una vez, porque cada entero se factoriza de una sola manera. De ahí sale la conexión: propiedades analíticas de ζ se traducen en propiedades aritméticas de los primos. El teorema de los números primos dice que la densidad de primos alrededor de x es $\frac{1}{\ln x}$: cerca de un millón, aproximadamente uno de cada catorce números es primo, y esa densidad decrece muy despacio.

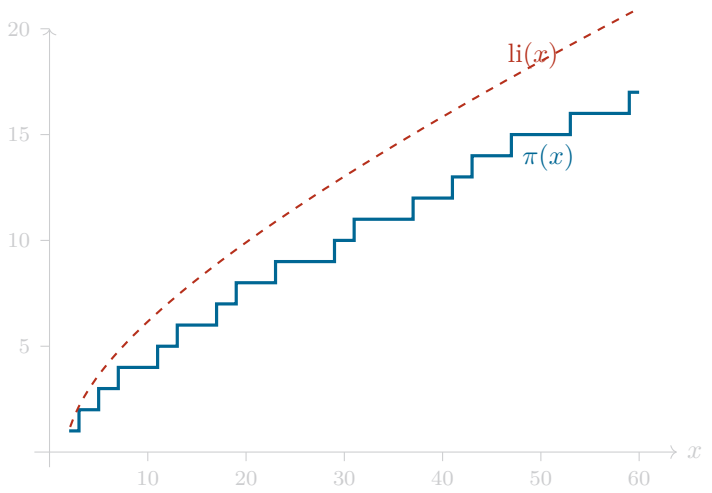


Figura 34.1: La función contadora de primos hasta 60, una escalera que sube justo en cada primo, comparada con el logaritmo integral. Aquí $\pi(60) = 17$ y $\mathrm{li}(60) = 20{,}97$.

Ejemplo 34.2 (Contar primos y medir el error)**. (a) La aproximación básica.** Los valores de $\pi(x)$ frente a $\frac{x}{\ln x}$:

x	$\pi(x)$	$x/\ln x$	cociente
10^3	168	144,8	1,161
10^6	78 498	72 382,4	1,084
10^9	50 847 534	48 254 942,4	1,054

El cociente tiende a 1, pero muy despacio: es un teorema asintótico y la convergencia es logarítmica.

(b) Una aproximación mucho mejor. El logaritmo integral $\mathrm{Li}(x) = \int_2^x \frac{\mathrm{d}t}{\ln t}$ da 178 para $x = 10^3$ y 78 628 para 10^6, con errores de 10 y 130 frente a los 23 y 6116 de la aproximación anterior. Gauss conjeturó esta versión a los quince años, en 1792, examinando tablas de primos.

(c) La divergencia de $\sum 1/p$. Euler demostró que la suma de los inversos de los primos diverge, lo que da una prueba analítica de que hay infinitos y además dice que son "más abundantes" que los cuadrados, cuya suma de inversos converge. La suma crece como $\ln\ln x$, es decir, con una lentitud absurda: para superar el valor 4 hay que llegar a primos del orden de 10^{18}.

(d) La hipótesis de Riemann. Todos los ceros no triviales de ζ tendrían parte real $\frac{1}{2}$. Es equivalente a que el error del teorema de los números primos sea del orden de $\sqrt{x}\ln x$, es decir, lo más pequeño posible. Se han verificado más de diez billones de ceros y todos cumplen la conjetura; sigue sin demostrarse desde 1859 y es uno de los problemas del milenio.

Cuidado

[Asintótico no significa buena aproximación para valores concretos] Que $\pi(x) \sim \frac{x}{\ln x}$ solo dice que el cociente tiende a 1; para $x = 10^9$ el error relativo sigue siendo del 5 %.

Segundo aviso, y es espectacular: durante mucho tiempo se creyó que $\mathrm{Li}(x) > \pi(x)$ siempre, porque así ocurre en todo el rango calculable. Littlewood demostró en 1914 que la desigualdad cambia de signo infinitas veces, y la primera vez que lo hace está en algún punto astronómicamente alto, más allá de 10^{19}. Es la advertencia definitiva contra conjeturar a partir de datos numéricos por muchos que sean.

Pausa para pensar

Riemann publicó en 1859 un artículo de ocho páginas, el único que escribió sobre teoría de números, en el que extendió ζ al plano complejo, relacionó sus ceros con la distribución de los primos y enunció de pasada la hipótesis que lleva su nombre, diciendo que le parecía probable pero que no había conseguido demostrarla y que no era esencial para su objetivo inmediato. Ese comentario al margen se ha convertido en el problema abierto más importante de la matemática. El teorema de los números primos lo demostraron independientemente Hadamard y de la Vallée Poussin en 1896, siguiendo el programa de Riemann; una demostración "elemental", sin variable compleja, no llegó hasta Erdős y Selberg en 1949, y provocó una amarga disputa de prioridad entre ambos.

34.2. Qué significa n^{-s}: series de Dirichlet, productos de Euler y prolongación analítica

La sección anterior ha escrito $\zeta(s) = \sum n^{-s}$ y ha hablado de ceros con parte real $\frac{1}{2}$ sin haber dicho qué es elevar un número natural a un exponente complejo, por qué esa serie converge en algún sitio, qué justifica el producto de Euler ni qué significa «extender» una función más allá de donde su fórmula tiene sentido. Todo eso se construye aquí, con las demostraciones hechas y con las comprobaciones numéricas al lado.

Elevar a un exponente complejo

Definición 34.3 (La potencia n^{-s})**.** Para un entero $n \geq 1$ y un número complejo $s = \sigma + it$ con $\sigma, t \in \mathbb{R}$ se define

$$n^{-s} = \mathrm{e}^{-s \ln n},$$

donde $\ln n$ es el logaritmo real, de modo que no hay ninguna ambigüedad de rama. Separando partes,

$$n^{-s} = n^{-\sigma}\big(\cos(t \ln n) - \mathrm{i}\,\sin(t \ln n)\big), \qquad \text{luego} \qquad |n^{-s}| = n^{-\sigma}.$$

Es costumbre en esta materia escribir $\sigma = \mathrm{Re}(s)$ y $t = \mathrm{Im}(s)$, y así se hará a partir de ahora.

Ejemplo 34.4 (Una potencia compleja, calculada entera)**.** Sea $s = 2 + 3\mathrm{i}$ y $n = 2$. Entonces $\ln 2 = 0{,}693\,147$ y $3 \ln 2 = 2{,}079\,442$, luego

$$2^{-s} = 2^{-2}(\cos 2{,}079\,442 - \mathrm{i}\,\sin 2{,}079\,442) = 0{,}25\,(-0{,}486\,994 - 0{,}873\,405\,\mathrm{i}) = -0{,}121\,749 - 0{,}218\,351\,\mathrm{i}.$$

Su módulo es $|2^{-s}| = 0{,}25 = 2^{-2}$, exactamente como predice la fórmula. ✓

La lectura es la que gobierna todo el capítulo: *la parte real de s controla el tamaño y la parte imaginaria sólo hace girar*. El factor $n^{-\sigma}$ decide si la serie converge; el factor $\cos(t\ln n) - \mathrm{i}\sin(t\ln n)$, de módulo 1, hace que los términos oscilen y se cancelen parcialmente. Toda la teoría analítica de números vive de esa tensión entre tamaño y cancelación.

Dónde converge una serie de Dirichlet

Teorema 34.5 (Convergencia de la serie de ζ)**.** *La serie $\sum_{n\geq 1} n^{-s}$ converge absolutamente si $\sigma > 1$, y uniformemente en cada semiplano $\sigma \geq 1+\delta$ con $\delta > 0$; en particular define una función holomorfa en $\sigma > 1$. Para $s = \sigma$ real y $\sigma > 1$ se tiene además la cota explícita*

$$\frac{1}{\sigma - 1} \;<\; \zeta(\sigma) \;\leq\; \frac{\sigma}{\sigma - 1}.$$

En $s = 1$ la serie diverge.

Demostración. Como $|n^{-s}| = n^{-\sigma}$, basta estudiar $\sum n^{-\sigma}$ con σ real. La función $t \mapsto t^{-\sigma}$ es decreciente en $[1,\infty)$, luego para cada $n \geq 2$

$$\int_{n}^{n+1} t^{-\sigma}\,\mathrm{d}t \;\leq\; n^{-\sigma} \;\leq\; \int_{n-1}^{n} t^{-\sigma}\,\mathrm{d}t.$$

Sumando en $n \geq 2$ y añadiendo el término $n = 1$,

$$\int_{1}^{\infty} t^{-\sigma}\mathrm{d}t \;\leq\; \sum_{n\geq 1} n^{-\sigma} \;\leq\; 1 + \int_{1}^{\infty} t^{-\sigma}\mathrm{d}t, \qquad \text{es decir} \qquad \frac{1}{\sigma-1} \leq \zeta(\sigma) \leq 1 + \frac{1}{\sigma-1} = \frac{\sigma}{\sigma-1},$$

y la integral es finita exactamente cuando $\sigma > 1$. Si $\sigma \geq 1+\delta$ entonces $|n^{-s}| \leq n^{-1-\delta}$, que es una serie convergente de constantes: por el criterio M de Weierstrass la convergencia es uniforme, y el límite uniforme de funciones holomorfas es holomorfo. Para $s = 1$ queda la serie armónica, que diverge. □

Ejemplo 34.6 (La cota, comprobada)**.** Con $\sigma = 1{,}5$ la cota superior es $\sigma/(\sigma-1) = 1{,}5/0{,}5 = 3$ y la inferior $1/(\sigma-1) = 2$. El valor verdadero es $\zeta(1{,}5) = 2{,}612\,375$: encajado entre ambas. ✓Con $\sigma = 1{,}001$ las cotas son 1000 y 1001, y el valor real es $\zeta(1{,}001) = 1000{,}577\,288$. Se ve nacer el polo: al acercarse σ a 1 la función crece como $1/(\sigma-1)$, y la diferencia $\zeta(\sigma) - 1/(\sigma-1)$ tiende a la constante de Euler–Mascheroni $\gamma = 0{,}577\,216$. Compárese con el valor calculado: $1000{,}577\,288$ frente a $1000 + 0{,}577\,216$. ✓

Definición 34.7 (Serie de Dirichlet y abscisa de convergencia)**.** Dada una función aritmética $f : \mathbb{N} \to \mathbb{C}$, su *serie de Dirichlet* es $D_f(s) = \sum_{n\geq 1} f(n) n^{-s}$. El mismo argumento de comparación de la demostración anterior muestra que si converge absolutamente en s_0, converge absolutamente en todo s con $\sigma \geq \sigma_0$: el dominio de convergencia absoluta es un *semiplano*, y su borde

$$\sigma_a(f) = \inf\{\sigma \;:\; \textstyle\sum_n |f(n)| n^{-\sigma} < \infty\}$$

es la *abscisa de convergencia absoluta*. Dos series de Dirichlet que coinciden en un semiplano tienen los mismos coeficientes, así que la correspondencia $f \leftrightarrow D_f$ no pierde información.

Ejemplo 34.8 (Cuatro abscisas)**.** Si $|f(n)| \le C\, n^k$ entonces $\sigma_a \le k+1$, por comparación con $\zeta(\sigma - k)$. Así: μ y $\mathbf{1}$ tienen $\sigma_a = 1$; $\tau(n)$ también, porque $\tau(n) \le n$ y sin embargo $\sum \tau(n)/n$ diverge; $\varphi(n)$, que es del tamaño de n, tiene $\sigma_a = 2$; y la función de von Mangoldt Λ, acotada por $\ln n$, tiene $\sigma_a = 1$. Todas las identidades de esta sección son válidas en el semiplano donde ambos lados convergen absolutamente, y sólo ahí; extender su validez es justamente el trabajo de la prolongación analítica.

Multiplicativa y completamente multiplicativa

Definición 34.9 (Las dos multiplicatividades)**.** Una función aritmética f no idénticamente nula es

- *multiplicativa* si $f(mn) = f(m)f(n)$ siempre que $\gcd(m,n) = 1$;
- *completamente multiplicativa* si $f(mn) = f(m)f(n)$ para *todos* m, n, coprimos o no.

Toda función multiplicativa cumple $f(1) = 1$: tomando $m = n = 1$ resulta $f(1) = f(1)^2$, luego $f(1)$ es 0 o 1; si fuera 0, entonces $f(n) = f(n)f(1) = 0$ para todo n y la función sería nula.

Proposición 34.10 (Una multiplicativa está determinada en las potencias de primos)**.** *Si f es multiplicativa y $n = p_1^{a_1} \cdots p_k^{a_k}$ es la factorización de n, entonces*

$$f(n) = f\big(p_1^{a_1}\big) \cdots f\big(p_k^{a_k}\big).$$

Si además f es completamente multiplicativa, $f(p^a) = f(p)^a$ y basta conocerla en los primos.

Demostración. Las potencias $p_i^{a_i}$ son coprimas dos a dos, así que se aplica la definición repetidamente por inducción en k. La segunda afirmación es la definición aplicada $a-1$ veces. □

Ejemplo 34.11 (Quién es quién, con contraejemplos numéricos)**. (a) Multiplicativas y no completamente.** τ, σ, φ y μ. Los contraejemplos son de una línea: $\tau(4) = 3$ pero $\tau(2)^2 = 4$; $\varphi(4) = 2$ pero $\varphi(2)^2 = 1$; $\mu(4) = 0$ pero $\mu(2)^2 = 1$; $\sigma(4) = 7$ pero $\sigma(2)^2 = 9$. En los cuatro casos el argumento $4 = 2 \cdot 2$ no es un producto de coprimos, y por eso no hay contradicción con la multiplicatividad.

(b) Completamente multiplicativas. Las potencias $\mathrm{Id}_k(n) = n^k$, la constante $\mathbf{1}(n) = 1$, la función δ que vale 1 en $n = 1$ y 0 después, los caracteres de Dirichlet χ —esa es exactamente la condición que se les exige— y la función de Liouville $\lambda(n) = (-1)^{\Omega(n)}$, con $\Omega(n)$ el número de factores primos contados con multiplicidad.

(c) Ni siquiera multiplicativa. La función de von Mangoldt Λ. Falla ya en el primer requisito, $\Lambda(1) = 0 \neq 1$; y en concreto $\Lambda(6) = 0$ mientras que $\Lambda(2)\Lambda(3) = \ln 2 \cdot \ln 3 = 0{,}761\,500 \neq 0$. Que Λ no sea multiplicativa no le impide tener una serie de Dirichlet utilísima —es $-\zeta'/\zeta$—; simplemente esa serie no tiene producto de Euler.

El producto de Euler, en general

Teorema 34.12 (Producto de Euler)**.** *Sea f multiplicativa con $\sum_n |f(n)|n^{-\sigma} < \infty$. Entonces*

$$D_f(s) = \sum_{n\geq 1} \frac{f(n)}{n^s} = \prod_{p \ primo} \left(1 + \frac{f(p)}{p^s} + \frac{f(p^2)}{p^{2s}} + \cdots\right),$$

y si f es completamente multiplicativa el factor de cada primo es una serie geométrica y queda

$$D_f(s) = \prod_p \left(1 - \frac{f(p)}{p^s}\right)^{-1}.$$

Demostración. Sea P_N el producto de los factores correspondientes a los primos $p \leq N$. Cada factor es una serie absolutamente convergente, así que se puede multiplicar término a término y reordenar; al hacerlo, cada sumando del desarrollo es

$$\frac{f(p_1^{a_1}) \cdots f(p_k^{a_k})}{(p_1^{a_1} \cdots p_k^{a_k})^s} = \frac{f(n)}{n^s}$$

por la proposición 34.10, y por el teorema fundamental de la aritmética cada entero n cuyos factores primos son todos $\leq N$ aparece *exactamente una vez*. Por tanto

$$P_N = \sum_{n \in A_N} \frac{f(n)}{n^s}, \qquad A_N = \{n \ : \ \text{todo primo que divide a } n \text{ es } \leq N\},$$

y la diferencia con la serie completa está acotada por $\left|D_f(s) - P_N\right| \leq \sum_{n>N} |f(n)|n^{-\sigma}$, que tiende a 0 porque es la cola de una serie convergente. Haciendo $N \to \infty$ se obtiene la identidad. □

Corolario 34.13. *Tomando $f = \mathbf{1}$ resulta $\zeta(s) = \prod_p (1 - p^{-s})^{-1}$ para $\sigma > 1$; tomando $f = \mu$, que es multiplicativa con $\mu(p) = -1$ y $\mu(p^k) = 0$ para $k \geq 2$, resulta*

$$\sum_{n\geq 1} \frac{\mu(n)}{n^s} = \prod_p \Big(1 - \frac{1}{p^s}\Big) = \frac{1}{\zeta(s)}.$$

En particular ζ no se anula en $\sigma > 1$: su inversa está dada por una serie convergente.

Ejemplo 34.14 (Cuatro series de Dirichlet, con comprobación numérica)**.** La regla de oro es que la convolución de Dirichlet se convierte en producto: si $h = f * g$ entonces $D_h = D_f\, D_g$, porque al multiplicar $\sum f(a)a^{-s}$ por $\sum g(b)b^{-s}$ el coeficiente de n^{-s} recoge todos los pares con $ab = n$.

(a) $\tau = \mathbf{1} * \mathbf{1}$, luego $\sum \tau(n)n^{-s} = \zeta(s)^2$. Comprobación en $s = 2$: $\zeta(2)^2 = 1{,}644\,934^2 = 2{,}705\,808$, y la suma parcial hasta $n = 20\,000$ vale $2{,}705\,205$, que se acerca por debajo como corresponde a una serie de términos positivos. ✓

(b) $\mu * \mathbf{1} = \delta$, luego $\sum \mu(n)n^{-s} = 1/\zeta(s)$. En $s = 2$ eso es $1/1{,}644\,934 = 0{,}607\,927$, y la suma parcial hasta $50\,000$ da $0{,}607\,927$. ✓

(c) $\sigma = \mathbf{1} * \mathrm{Id}$, luego $\sum \sigma(n)n^{-s} = \zeta(s)\zeta(s-1)$, válido para $\sigma > 2$.

(d) $\varphi * \mathbf{1} = \mathrm{Id}$, luego $\sum \varphi(n)n^{-s} = \zeta(s-1)/\zeta(s)$, válido para $\sigma > 2$. Comprobación en $s = 3$: $\zeta(2)/\zeta(3) = 1{,}644\,934/1{,}202\,057 = 1{,}368\,433$, y la suma parcial hasta $20\,000$ vale $1{,}368\,402$. ✓

Regla: De una identidad aritmética a una identidad analítica

1. **Escribe la identidad como convolución.** «$\sum_{d\mid n}\varphi(d)=n$» es $\varphi * \mathbf{1} = \mathrm{Id}$.
2. **Traduce cada factor a su serie.** $\mathbf{1}\to\zeta(s)$, $\mathrm{Id}\to\zeta(s-1)$, $\mu\to 1/\zeta(s)$, $\delta\to 1$.
3. **Convolución pasa a producto**: la identidad se convierte en $D_\varphi(s)\,\zeta(s)=\zeta(s-1)$, y se despeja.
4. **Anota el semiplano de validez**: el mayor de las abscisas de convergencia absoluta de las series implicadas. Aquí $\sigma>2$.
5. **Comprueba con un valor concreto** antes de fiarte: una suma parcial larga frente al valor cerrado. Todos los ejemplos de esta sección se han verificado así.

Prolongación analítica: qué se está afirmando exactamente

Definición 34.15 (Prolongación analítica)**.** Sea F holomorfa en un abierto conexo $U\subseteq\mathbb{C}$ y $\Omega\supseteq U$ otro abierto conexo. Una *prolongación analítica* de F a Ω es una función G holomorfa en Ω tal que $G=F$ en U. Por el principio de identidad (capítulo 30), si dos funciones holomorfas en un abierto conexo coinciden en un conjunto con un punto de acumulación, coinciden en todo el abierto; en consecuencia *la prolongación, si existe, es única.*

Cuidado

[Prolongar no es sumar] Éste es el punto que más confusión causa. La serie $\sum n^{-s}$ *no* converge en $s=-1$ ni en $s=0$: allí sus términos crecen sin límite. Lo que existe es una única función holomorfa en $\mathbb{C}\setminus\{1\}$ que *coincide con esa suma donde la suma tiene sentido*, y es esa función la que vale $-\frac{1}{12}$ en $s=-1$. Decir «$1+2+3+\cdots=-\frac{1}{12}$» es un abuso de lenguaje inofensivo entre quien sabe qué significa y una fuente de disparates fuera de ese contexto.

Teorema 34.16 (La identidad de la serie alternada)**.** *Para $\sigma>1$,*

$$\eta(s)=\sum_{n\geq 1}\frac{(-1)^{n-1}}{n^s}=(1-2^{1-s})\zeta(s),$$

y la serie de la izquierda converge —condicionalmente— en todo el semiplano $\sigma>0$.

Demostración. En el rango de convergencia absoluta se puede reordenar. Los términos pares son $\sum_{n\text{ par}} n^{-s}=\sum_{m\geq 1}(2m)^{-s}=2^{-s}\zeta(s)$. Sumar todos con signo $+$ da $\zeta(s)$; para cambiar el signo a los pares hay que restarlos dos veces:

$$\eta(s)=\zeta(s)-2\cdot 2^{-s}\zeta(s)=(1-2^{1-s})\zeta(s).$$

Para la convergencia en $\sigma>0$ basta el criterio de Dirichlet: las sumas parciales de $(-1)^{n-1}$ están acotadas por 1 y los factores n^{-s} tienen variación total acotada y tienden a 0 cuando $\sigma>0$. □

Observación 34.17 (Por qué sólo sobrevive el polo de $s=1$). Despejando, $\zeta(s)=\eta(s)/(1-2^{1-s})$, y esa expresión parece tener polos en todos los ceros del denominador, es decir, en

$$s_k = 1 + \frac{2\pi \mathrm{i}\, k}{\ln 2}, \qquad k \in \mathbb{Z}.$$

No los tiene: en cada uno de esos puntos con $k \neq 0$ el numerador η también se anula, y el cociente tiene una singularidad *evitable*. Con $k=1$ resulta $s_1 = 1 + 9{,}064\,720\,\mathrm{i}$, y el cálculo numérico confirma que $\eta(s_1)=0$ mientras que $\zeta(s_1) = 1{,}346\,579 + 0{,}109\,882\,\mathrm{i}$, un valor perfectamente finito. El único polo verdadero es $s=1$, y allí $\eta(1)=\ln 2 \neq 0$ mientras el denominador se anula simplemente.

Teorema 34.18 (Prolongación a $\sigma>0$ y el polo, con residuo)**.** *Para $\sigma>1$ se cumple*

$$\zeta(s) = \frac{s}{s-1} - s\int_1^\infty \frac{\{t\}}{t^{s+1}}\,\mathrm{d}t,$$

donde $\{t\}=t-\lfloor t\rfloor \in [0,1)$ es la parte fraccionaria. La integral converge absolutamente para $\sigma>0$ y define ahí una función holomorfa, luego el segundo miembro es una prolongación analítica de ζ al semiplano $\sigma>0$, con un único polo, simple, en $s=1$ y de residuo 1.

Demostración. Separando $\{t\}=t-\lfloor t\rfloor$,

$$\int_1^\infty \frac{\{t\}}{t^{s+1}}\mathrm{d}t = \int_1^\infty t^{-s}\mathrm{d}t - \int_1^\infty \frac{\lfloor t\rfloor}{t^{s+1}}\mathrm{d}t = \frac{1}{s-1} - \sum_{n\ge 1} n\int_n^{n+1} t^{-s-1}\mathrm{d}t.$$

La integral interior vale $\big(n^{-s}-(n+1)^{-s}\big)/s$, de modo que la suma es

$$\frac{1}{s}\sum_{n\ge1}\Big(n\cdot n^{-s} - n\,(n+1)^{-s}\Big) = \frac{1}{s}\Big(\sum_{n\ge1} n^{1-s} - \sum_{m\ge2}(m-1)m^{-s}\Big) = \frac{1}{s}\Big(\zeta(s-1)-\zeta(s-1)+1-1+\zeta(s)\Big) = \frac{\zeta(s)}{s},$$

donde se ha reindexado $m=n+1$ y se han cancelado los términos $\sum m^{1-s}$. Por tanto

$$\int_1^\infty \frac{\{t\}}{t^{s+1}}\mathrm{d}t = \frac{1}{s-1} - \frac{\zeta(s)}{s},$$

que despejada es la fórmula del enunciado. Como $0\le\{t\}<1$, la integral está dominada por $\int_1^\infty t^{-\sigma-1}\mathrm{d}t = 1/\sigma$, finita para $\sigma>0$, y depende holomorfamente de s por derivación bajo el signo integral. Finalmente, $s/(s-1)$ tiene un polo simple en $s=1$ con residuo $\lim_{s\to1}(s-1)\cdot\frac{s}{s-1}=1$. □

Ejemplo 34.19 (Comprobación en $s=2$ y lectura del polo)**.** Con $s=2$ la fórmula dice $\zeta(2) = 2 - 2\int_1^\infty \{t\}t^{-3}\mathrm{d}t$, es decir $\int_1^\infty \{t\}t^{-3}\mathrm{d}t = 1-\zeta(2)/2 = 1-0{,}822\,467 = 0{,}177\,533$; y en efecto $2-2\cdot 0{,}177\,533 = 1{,}644\,934 = \zeta(2)$. ✓

Haciendo $s\to1$ en la fórmula se obtiene además

$$\lim_{s\to1}\Big(\zeta(s)-\frac{1}{s-1}\Big) = 1 - \int_1^\infty \frac{\{t\}}{t^2}\mathrm{d}t = \gamma = 0{,}577\,216,$$

la constante de Euler–Mascheroni, que aparece así como el término constante del desarrollo de ζ en su polo. Es la misma constante que gobierna $\sum_{n\le x} 1/n = \ln x + \gamma + o(1)$, y no es

casualidad: el polo de ζ en $s=1$ *es* la divergencia de la serie armónica vista desde el plano complejo. Y como el producto de Euler convierte esa divergencia en $\sum 1/p=\infty$, el polo es también la razón analítica de que haya infinitos primos.

Observación 34.20 (Qué falta y dónde está). Con esto, ζ está definida en $\sigma>0$ salvo el polo. Para llegar al resto del plano —y poder hablar de los ceros triviales en $-2,-4,-6,\dots$— hace falta la ecuación funcional, que la sección correspondiente de este capítulo enuncia y sitúa históricamente. Lo que aquí queda establecido es lo previo: qué significa la serie, dónde converge, por qué el producto de Euler es legítimo, qué quiere decir prolongar y por qué la prolongación es única.

34.3. Funciones aritméticas y primos en progresiones

Definición 34.21. Una *función aritmética* es cualquier $f:\mathbb{N}\to\mathbb{C}$; es *multiplicativa* si $f(mn)=f(m)f(n)$ cuando $\gcd(m,n)=1$. Las tres básicas son $\tau(n)$ (número de divisores), $\sigma(n)$ (suma de divisores) y la indicatriz $\varphi(n)$. La *serie de Dirichlet* asociada es $D_f(s)=\sum_n f(n)n^{-s}$, y el *teorema de Dirichlet* afirma que toda progresión $an+b$ con $\gcd(a,b)=1$ contiene infinitos primos.

Idea clave

Qué dice esto en cristiano. Las series de Dirichlet son a las funciones aritméticas lo que las series de potencias a las funciones ordinarias: convierten convoluciones en productos. Como $\zeta(s)^2=\sum\tau(n)n^{-s}$ y $\frac{\zeta(s-1)}{\zeta(s)}=\sum\varphi(n)n^{-s}$, cada identidad analítica encierra una identidad aritmética. El teorema de Dirichlet es el ejemplo mayor de este método: para demostrar un enunciado sobre primos en progresiones se introducen unos caracteres, se construyen unas series y se estudia dónde no se anulan. Aritmética pura demostrada con análisis complejo.

Ejemplo 34.22 (Funciones aritméticas y progresiones)**. (a) Los tres valores básicos.** Para $n=12$: $\tau=6$, $\sigma=28$ y $\varphi=4$. Para $n=36$: 9, 91 y 12. Para $n=100$: 9, 217 y 40. Las tres son multiplicativas, así que basta conocerlas en potencias de primos: $\tau(p^k)=k+1$, $\sigma(p^k)=\frac{p^{k+1}-1}{p-1}$ y $\varphi(p^k)=p^k-p^{k-1}$.

(b) Un número perfecto. $\sigma(28)=1+2+4+7+14+28=56=2\cdot 28$: el 28 es perfecto. Euclides demostró que $2^{p-1}(2^p-1)$ lo es siempre que 2^p-1 sea primo, y Euler que todos los perfectos pares son de esa forma. Si existe alguno impar sigue sin saberse.

(c) Primos en progresiones. Hasta 100,

$$p\equiv 1 \pmod 4 : 5,13,17,29,37,41,53,61,73,89,97,$$
$$p\equiv 3 \pmod 4 : 3,7,11,19,23,31,43,47,59,67,71,79,83.$$

Ambas listas son infinitas, y además tienen la misma densidad asintótica, $\frac12$ cada una, aunque la segunda va casi siempre por delante en las cuentas parciales (el llamado sesgo de Chebyshev).

(d) Valores especiales de zeta. Euler calculó $\zeta(2)=\frac{\pi^2}{6}=1{,}644934$ y $\zeta(4)=\frac{\pi^4}{90}=1{,}082323$, y en general $\zeta(2k)$ es un múltiplo racional de π^{2k}. Para los impares no se conoce

nada parecido: que $\zeta(3) = 1{,}202057$ es irracional lo demostró Apéry en 1978 y causó estupor; de $\zeta(5)$ ni siquiera se sabe eso.

Cuidado

[Multiplicativa no es lo mismo que completamente multiplicativa] φ es multiplicativa pero $\varphi(4) = 2 \neq \varphi(2)^2 = 1$: la propiedad solo vale para argumentos coprimos. Solo unas pocas funciones, como $f(n) = n^k$, cumplen la versión fuerte. Segundo aviso: la convergencia de $\sum \frac{1}{p}$ hacia infinito y la de $\sum \frac{1}{p}$ restringida a primos gemelos son cosas distintas. Los gemelos hasta 100 son $(3,5), (5,7), (11,13), (17,19), (29,31), (41,43), (59,61), (71,73)$, y Brun demostró en 1919 que la suma de sus inversos *converge*, a un valor cercano a 1,902160. Eso no prueba que sean finitos: podrían ser infinitos y aun así sumar poco. La conjetura sigue abierta.

34.4. El teorema de los números primos y la hipótesis de Riemann

La pregunta "¿cuántos primos hay por debajo de x?" no tiene respuesta exacta, y probablemente no la tendrá nunca. Tiene, en cambio, una respuesta asintótica de una precisión asombrosa, y la historia de cómo se afinó esa respuesta es la historia de la teoría analítica de números.

Teorema 34.23 (De los números primos)**.** *Sea $\pi(x)$ el número de primos menores o iguales que x. Entonces*

$$\pi(x) \sim \frac{x}{\ln x}, \qquad \text{y con más precisión} \qquad \pi(x) = \operatorname{li}(x) + O\big(x\,\mathrm{e}^{-c\sqrt{\ln x}}\big),$$

donde $\operatorname{li}(x) = \int_2^x \dfrac{\mathrm{d}t}{\ln t}$ es el logaritmo integral.

Idea clave

Qué dice esto en cristiano. Cerca de un número grande x, la "densidad" de primos es aproximadamente $1/\ln x$: entre los números de nueve cifras, uno de cada veintiuno es primo. El teorema dice que esa intuición es correcta y que integrarla —eso es exactamente $\operatorname{li}(x)$— da una estimación mucho mejor que la versión burda $x/\ln x$. Los primos no siguen ninguna regla, y sin embargo su cantidad obedece una ley suave, del mismo modo que las moléculas de un gas se mueven al azar y la presión es predecible.

Ejemplo 34.24 (Cuánto acierta cada aproximación)**. (a) La tabla.** Para tres tamaños:

x	$\pi(x)$	$x/\ln x$	$\operatorname{li}(x)$
10^3	168	145	178
10^6	78 498	72 382	78 628
10^9	50 847 534	48 254 942	50 849 235

(b) Quién gana. El error relativo de $x/\ln x$ es del 14 % en 10^3 y baja al 5 % en 10^9: converge, pero con una lentitud desesperante, porque el error se comporta como $1/\ln x$. El de $\mathrm{li}(x)$ es del 5,7 % en 10^3 y de sólo 0,0033 % en 10^9. Por eso las tablas serias usan el logaritmo integral y no el cociente sencillo, aunque ambos sean asintóticamente equivalentes.

(c) La densidad local. Entre 10^9 y $10^9 + 1000$ cabe esperar $1000/\ln(10^9) = 48$ primos. La predicción funciona en promedio; en un tramo concreto puede haber treinta o sesenta, y eso es exactamente lo que las fluctuaciones miden.

(d) El sesgo de Littlewood. En todos los valores tabulados $\mathrm{li}(x)$ es mayor que $\pi(x)$, y durante décadas se conjeturó que siempre lo era. Littlewood demostró en 1914 que la diferencia cambia de signo infinitas veces; el primer cambio ocurre en algún punto que todavía no se sabe localizar, y que las cotas actuales sitúan por debajo de 10^{316}. Es el aviso definitivo contra generalizar a partir de datos numéricos, por muchos que sean.

Definición 34.25. La *hipótesis de Riemann* afirma que todos los ceros no triviales de $\zeta(s)$ tienen parte real igual a $\tfrac{1}{2}$. Es equivalente a la estimación

$$\pi(x) = \mathrm{li}(x) + O(\sqrt{x}\,\ln x),$$

que es la mejor posible: ningún error menor es compatible con la existencia de ceros.

Idea clave

Qué dice esto en cristiano. La conexión entre los ceros de una función compleja y el recuento de primos parece cosa de magia, y no lo es: la fórmula explícita de Riemann expresa $\pi(x)$ como el término principal $\mathrm{li}(x)$ menos una suma de contribuciones, una por cada cero de ζ. Cada cero $\rho = \beta + \mathrm{i}\gamma$ aporta un término oscilante de tamaño x^{β}. Si todos los β valen $\tfrac{1}{2}$, todas esas ondas tienen amplitud $\sqrt{x}$ y el error total es del orden de $\sqrt{x}$: los primos están repartidos con la mayor regularidad posible. Un solo cero con $\beta = 0{,}6$ produciría una onda de amplitud $x^{0{,}6}$ que estropearía la regularidad. La hipótesis dice, en el fondo, que no hay ninguna conspiración entre los primos.

Ejemplo 34.26 (Lo que se sabe y lo que no)**. (a) Los primeros ceros.** Están en $\tfrac{1}{2} \pm 14{,}134725\,\mathrm{i}$, $\tfrac{1}{2} \pm 21{,}022040\,\mathrm{i}$ y $\tfrac{1}{2} \pm 25{,}010858\,\mathrm{i}$. Se han verificado más de diez billones de ceros y todos caen en la recta crítica.

(b) Lo demostrado. Hardy probó en 1914 que hay infinitos ceros *sobre* la recta; Selberg, que una proporción positiva lo está; y Conrey, en 1989, que al menos el 40 %. Nadie ha demostrado que no haya ninguno fuera, y esa es toda la diferencia.

(c) La región libre de ceros. Lo que sí se sabe es que no hay ceros con $\beta = 1$, y ése es exactamente el ingrediente que Hadamard y de la Vallée Poussin usaron en 1896 para demostrar el teorema de los números primos. Cuanto mayor sea la región libre de ceros que uno sepa demostrar, mejor el término de error.

(d) Qué se caería con ella. Miles de teoremas están publicados en la forma "suponiendo la hipótesis de Riemann". Uno concreto y bonito: el criterio de Robin dice que la hipótesis equivale a que $\sigma(n) < \mathrm{e}^{\gamma} n \ln\ln n$ para todo $n > 5040$, donde σ es la suma de divisores y

$\gamma = 0{,}577216$ es la constante de Euler–Mascheroni. Un solo entero que lo incumpliera derribaría la conjetura, y se puede buscar con un ordenador.

Cuidado

[Contraejemplo: la equivalencia no es con $x/\ln x$] El teorema de los números primos, en la forma $\pi(x) \sim x/\ln x$, *no* basta para deducir la hipótesis de Riemann ni nada parecido, porque el símbolo $\sim$ sólo controla el cociente y tolera errores enormes en términos absolutos: la diferencia $\pi(x) - x/\ln x$ crece como $x/(\ln x)^2$, es decir, unos 2,6 millones ya en 10^9. Toda la información fina está en el término de error, no en el principal, y por eso los enunciados equivalentes se formulan siempre con $\mathrm{li}(x)$.

Pausa para pensar

Riemann publicó su conjetura en 1859, en un artículo de ocho páginas que era el único que escribió sobre teoría de números, y la enunció de pasada, diciendo que sería deseable demostrarla pero que la había dejado de lado para no distraerse del objetivo principal. Ciento sesenta y siete años después sigue abierta, es el octavo problema de Hilbert y uno de los siete del milenio, con un millón de dólares esperando. Cuando en 1932 se examinaron los papeles privados de Riemann, se descubrió que había calculado a mano los primeros ceros con una fórmula que nadie conocía y que hoy se usa para calcularlos por ordenador. Había visto mucho más de lo que publicó.

34.5. Cribas y la conjetura de los primos gemelos

Contar primos es difícil; contar primos con condiciones adicionales lo es mucho más. La técnica que lo intenta es la criba, y su historia reciente incluye uno de los avances más inesperados de las últimas décadas, obra de un matemático que trabajaba en la oscuridad más completa.

Definición 34.27. Una *criba* estima el tamaño de un conjunto tras eliminar los múltiplos de una familia de primos. La *criba de Eratóstenes* con inclusión-exclusión da, para el número de enteros hasta x libres de factores primos menores que z,

$$\Phi(x, z) \approx x \prod_{p<z} \Big(1 - \frac{1}{p}\Big) \approx \frac{2\mathrm{e}^{-\gamma} x}{\ln z},$$

por el teorema de Mertens. La *criba de Brun* refina el proceso truncando la inclusión-exclusión en un número par o impar de términos para obtener cotas por arriba y por abajo.

Idea clave

Qué dice esto en cristiano. La idea es la de siempre: quitar los múltiplos de 2, luego los de 3, luego los de 5, y ver qué queda. El problema es que inclusión-exclusión con miles de primos tiene un número astronómico de términos, y los errores de cada uno se acumulan hasta tragarse el resultado. La aportación de Brun fue notar que si se corta

la suma en un número par de términos se obtiene una cota superior, y si se corta en impar, una inferior: se renuncia a la igualdad exacta a cambio de un control efectivo. Casi toda la teoría de cribas del siglo pasado consiste en refinar ese truco, y su limitación fundamental —el llamado problema de la paridad— explica por qué las cribas nunca han bastado por sí solas para atrapar primos gemelos.

Ejemplo 34.28 (Lo que se sabe sobre los gemelos)**. (a) La conjetura.** Se cree que hay infinitos pares $(p, p+2)$ ambos primos, y que su cantidad hasta x es asintóticamente $2C_2\, x/(\ln x)^2$, con $C_2 = 0{,}6601618$, la constante de los primos gemelos. Sigue abierta.

(b) Brun, 1919**.** Demostró que la suma de los inversos de los primos gemelos *converge*, a la constante $B_2 = 1{,}902160$. Es un resultado curioso: si hubiera "muchos" gemelos la suma divergería como diverge $\sum 1/p$; que converja indica que son escasos, y no resuelve si son finitos o infinitos.

(c) Chen, 1973**.** Hay infinitos primos p tales que $p+2$ es primo o producto de dos primos. Es lo más cerca que nadie ha estado, y se demostró durante la Revolución Cultural, en condiciones que el propio Chen describió como escribir sobre el suelo.

(d) Zhang, 2013**.** Existe una constante N tal que hay infinitos pares de primos consecutivos que difieren en menos de N. Zhang la fijó en 70 millones. Tenía cincuenta y ocho años, no tenía plaza fija y había trabajado años en un restaurante. En meses, el proyecto colaborativo Polymath y una idea independiente de Maynard bajaron la cota a 246. Para llegar a 2 haría falta una idea nueva que nadie tiene.

Cuidado

[Contraejemplo: la heurística probabilística falla si se usa mal] Tratar los primos como aleatorios con densidad $1/\ln x$ y suponer independencia da $x/(\ln x)^2$ pares de gemelos hasta x: el orden de magnitud correcto, pero con la constante equivocada, porque falta el factor $2C_2 = 1{,}320$. La razón es que los primos *no* son independientes módulo primos pequeños: si p es impar, $p+2$ también lo es, lo que duplica sus posibilidades respecto de un número al azar. La heurística de Hardy y Littlewood corrige clase por clase y acierta; la ingenua se equivoca en un 32 por ciento. Es un buen recordatorio de que "comportarse aleatoriamente" nunca significa "ser independiente".

34.6. Caracteres, funciones L y primos en progresiones

Euclides demostró que hay infinitos primos. La pregunta siguiente —¿hay infinitos que terminen en 7?— parece del mismo nivel y resultó necesitar análisis complejo. La respuesta de Dirichlet, de 1837, es el primer teorema de la teoría analítica de números.

Definición 34.29. Un *carácter de Dirichlet* módulo q es una función $\chi\colon \mathbb{Z} \to \mathbb{C}$ completamente multiplicativa, periódica de periodo q, con $\chi(n) = 0$ si $\gcd(n, q) > 1$. Hay exactamente $\varphi(q)$ de ellos, y forman un grupo isomorfo a $(\mathbb{Z}/q\mathbb{Z})^\times$. Su *función L* es $L(s, \chi) = \sum_{n\geq 1} \chi(n) n^{-s} = \prod_p (1 - \chi(p)p^{-s})^{-1}$ para $\Re s > 1$.

Idea clave

Qué dice esto en cristiano. Un carácter es un filtro. Multiplicar por él y sumar sirve para aislar una clase de restos entre todas las demás, exactamente igual que en el análisis de Fourier se multiplica por una exponencial e integra para extraer una frecuencia. De hecho es lo mismo: los caracteres de Dirichlet son los caracteres del grupo $(\mathbb{Z}/q\mathbb{Z})^\times$, y la identidad que aísla la progresión a mód q,

$$\frac{1}{\varphi(q)}\sum_{\chi}\bar{\chi}(a)\chi(n)=\begin{cases}1 & \text{si } n\equiv a\ (q),\\ 0 & \text{en otro caso,}\end{cases}$$

es la fórmula de inversión de Fourier sobre un grupo abeliano finito. El capítulo de análisis armónico y éste están hablando del mismo objeto con vocabulario distinto.

Teorema 34.30 (Dirichlet)**.** *Si* $\gcd(a,q)=1$*, la progresión* $a,\ a+q,\ a+2q,\dots$ *contiene infinitos primos. Más aún,* $\sum_{p\equiv a\,(q)}1/p$ *diverge, y los primos se reparten equitativamente entre las* $\varphi(q)$ *clases posibles.*

Regla: La estructura de la demostración, en siete pasos

1. **Tomar logaritmos** del producto de Euler para convertirlo en la suma $\log L(s,\chi)=\sum_p \chi(p)p^{-s}+O(1)$.
2. **Filtrar con caracteres** usando la ortogonalidad para quedarse sólo con los primos de la clase a.
3. **Aislar el carácter principal**, cuya L es esencialmente ζ y explota en $s=1$: ése es el término que hace divergir la suma.
4. **Demostrar que los demás no explotan**, es decir, que $L(1,\chi)$ es finito para $\chi\neq\chi_0$. Eso es fácil, porque las series alternan.
5. **Demostrar que tampoco se anulan**: $L(1,\chi)\neq 0$. Éste es el paso duro, y sin él la divergencia podría cancelarse.
6. **Separar el caso complejo del real**: para caracteres complejos, χ y $\bar{\chi}$ son distintos y un argumento de conteo basta.
7. **Rematar el caso real**, que resiste, y para el que Dirichlet acabó usando su fórmula del número de clases.

Ejemplo 34.31 (Módulo 4, con todos los números)**. (a) Los dos caracteres.** Módulo 4 hay $\varphi(4)=2$: el principal χ_0, que vale 1 en los impares y 0 en los pares, y χ_1, que vale 1 en los $\equiv 1$, -1 en los $\equiv 3$ y 0 en los pares.

(b) La función L no trivial. $L(1,\chi_1)=1-\frac{1}{3}+\frac{1}{5}-\frac{1}{7}+\cdots=\pi/4=0{,}785398$. Es la serie de Leibniz, y lo importante aquí no es su valor bonito sino que *no es cero*: eso basta para concluir que hay infinitos primos de cada tipo.

(c) El reparto real. Hasta 100 hay 11 primos de la forma $4k+1$ y 13 de la forma $4k+3$.

El teorema dice que la proporción tiende a un medio, pero converge muy despacio.

(d) El sesgo de Chebyshev. Los primos $4k+3$ van por delante casi siempre. Chebyshev lo notó en 1853 y durante mucho tiempo pareció una regla; Littlewood demostró en 1914 que la ventaja cambia de signo infinitas veces, y el primer cambio ocurre en 26 861. Es el mismo fenómeno que con $\mathrm{li}(x)$ y $\pi(x)$.

Cuidado

[Contraejemplo: la hipótesis $\gcd(a,q)=1$ no se puede quitar] La progresión $4,\ 10,\ 16,\dots$ —es decir, $a=4$ y $q=6$— no contiene ningún primo, porque todos sus términos son pares y mayores que 2. Con $\gcd(a,q)=d>1$, todo término es múltiplo de d y a lo sumo puede haber un primo. La condición de coprimalidad es exactamente la que hace que la clase pueda contener primos, y una vez que puede, contiene infinitos.

Pausa para pensar

Dirichlet leyó la demostración en la Academia de Berlín en 1837 y ese día nació la teoría analítica de números. El escándalo metodológico era mayúsculo: para demostrar un enunciado sobre enteros se usaban series infinitas, límites y funciones de variable compleja. Gauss, todavía vivo, desconfiaba de esa clase de argumentos. Hoy la mitad de lo que se sabe sobre los primos se demuestra así y no se conoce ninguna alternativa elemental para casi nada de ello.

34.7. Continuación analítica de ζ y la ecuación funcional

Hemos hablado de "ceros no triviales" sin haber extendido nunca ζ más allá de $\Re s>1$, que es donde converge la serie. Ahí no hay ni un solo cero: el producto de Euler lo prohíbe. Todo el interés está fuera, y hay que llegar.

Teorema 34.32 (Primera extensión)**.** *La serie alternada $\eta(s)=\sum_{n\ge1}(-1)^{n-1}n^{-s}$ converge para $\Re s>0$, y cumple $\eta(s)=(1-2^{1-s})\zeta(s)$. Luego*

$$\zeta(s)=\frac{\eta(s)}{1-2^{1-s}}$$

extiende ζ a $\Re s>0$ salvo posibles polos en $1-2^{1-s}=0$. El único que sobrevive es $s=1$, y es un polo simple con residuo 1.

Ese polo es la traducción analítica de que hay infinitos primos: si ζ fuera regular en 1, el producto de Euler daría un valor finito para $\sum 1/p$.

Teorema 34.33 (Ecuación funcional, Riemann 1859)**.** *Definiendo la función completada*

$$\xi(s)=\pi^{-s/2}\Gamma\!\left(\tfrac{s}{2}\right)\zeta(s),$$

se cumple $\xi(s)=\xi(1-s)$ para todo s. Equivalentemente,

$$\zeta(s)=2^s\pi^{s-1}\sin\!\Big(\frac{\pi s}{2}\Big)\Gamma(1-s)\,\zeta(1-s).$$

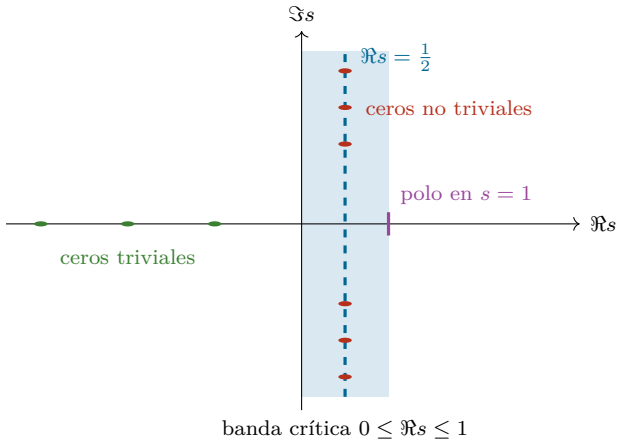


Figura 34.2: La ecuación funcional $\xi(s) = \xi(1-s)$ hace de $\Re s = \frac{1}{2}$ un eje de simetría. A la derecha de $\Re s = 1$ el producto de Euler prohíbe los ceros; a la izquierda de $\Re s = 0$ solo quedan los triviales, en los enteros pares negativos. Todo lo interesante ocurre en la banda, y la hipótesis de Riemann dice que ocurre justo en la raya.

La demostración clásica pasa por la función theta de Jacobi $\vartheta(t) = \sum_{n\in\mathbb{Z}} e^{-\pi n^2 t}$, que satisface $\vartheta(1/t) = \sqrt{t}\,\vartheta(t)$ por sumación de Poisson; aplicando la transformada de Mellin se obtiene ξ y la simetría $t \leftrightarrow 1/t$ se convierte en $s \leftrightarrow 1-s$.

Ejemplo 34.34 (Los ceros triviales y dos valores célebres)**. (a) Ceros triviales.** En $s = -2, -4, -6, \dots$ el factor $\sin(\pi s/2)$ se anula, y ni $\Gamma(1-s)$ ni $\zeta(1-s)$ lo compensan. Luego

$$\zeta(-2) = \zeta(-4) = \zeta(-6) = \cdots = 0.$$

En $s = 0$ y en los enteros negativos impares el seno no se anula, y ahí hay valores no nulos.

(b) $\zeta(-1)$**.** Poniendo $s = -1$: $\sin(-\pi/2) = -1$, $\Gamma(2) = 1$ y $\zeta(2) = \pi^2/6$, luego

$$\zeta(-1) = 2^{-1}\pi^{-2}\cdot(-1)\cdot 1\cdot\frac{\pi^2}{6} = -\frac{1}{12}.$$

Este es el número que aparece en la broma de que "$1+2+3+\cdots = -1/12$". La suma diverge, por supuesto; lo que vale $-1/12$ es la continuación analítica de la función que *donde converge* vale esa suma. Y sin embargo el valor tiene contenido físico: es el que da la energía del vacío en la cuerda bosónica y el efecto Casimir.

(c) $\zeta(0) = -\frac{1}{2}$ **y** $\zeta(-3) = \frac{1}{120}$, por el mismo cálculo. En general $\zeta(-n) = -B_{n+1}/(n+1)$ con B los números de Bernoulli, lo que enlaza esta sección con la fórmula de Euler–Maclaurin.

Idea clave

La ecuación funcional coloca todo en su sitio: la recta $\Re s = \frac{1}{2}$ es el *eje de simetría* de ξ. Los ceros no triviales viven en la banda $0 \le \Re s \le 1$ y están dispuestos simétricamente respecto de esa recta (y del eje real, por reflexión de Schwarz). La hipótesis de Riemann afirma que todos están justo encima del eje de simetría; sin la ecuación funcional, la

conjetura ni siquiera se podría formular. Y el recuento de ceros hasta altura T es

$$N(T) \sim \frac{T}{2\pi}\log\frac{T}{2\pi \mathrm{e}},$$

que es la fórmula que permite hablar de la *densidad* de los ceros y compararla con estadísticas de matrices aleatorias.

34.8. Von Mangoldt, las funciones de Chebyshev y los órdenes medios

Contar primos directamente es incómodo. Toda la teoría analítica trabaja con dos funciones auxiliares que pesan cada primo con su logaritmo, y con ellas las fórmulas salen limpias.

Definición 34.35 (Von Mangoldt y Chebyshev)**.**

$$\Lambda(n) = \begin{cases} \log p & \text{si } n = p^k, \\ 0 & \text{en otro caso,} \end{cases} \qquad \theta(x) = \sum_{p \le x} \log p, \qquad \psi(x) = \sum_{n \le x} \Lambda(n).$$

La razón de ser de Λ es esta identidad, que se obtiene derivando logarítmicamente el producto de Euler:

$$-\frac{\zeta'(s)}{\zeta(s)} = \sum_{n \ge 1} \frac{\Lambda(n)}{n^s} \qquad (\Re s > 1).$$

A la izquierda hay una función meromorfa cuyos polos son exactamente los ceros de ζ (y el polo $s = 1$); a la derecha, la información sobre los primos. Esa igualdad *es* el puente entre el análisis complejo y la aritmética, y es la razón de que los ceros de ζ gobiernen la distribución de los primos.

Teorema 34.36 (Equivalencias)**.**

$$\psi(x) \sim x \quad \Longleftrightarrow \quad \theta(x) \sim x \quad \Longleftrightarrow \quad \pi(x) \sim \frac{x}{\log x}.$$

El paso de ψ a π se hace por sumación de Abel, y la diferencia $\psi - \theta$ es $O(\sqrt{x}\log^2 x)$: las potencias de primos son despreciables.

Ejemplo 34.37 (Tres funciones, tres velocidades)**.**

x	$\pi(x)$	$\theta(x)$	$\psi(x)$	$\psi(x)/x$	$\pi(x)/\operatorname{li}(x)$
10^2	25	83,728	94,045	0,940 5	0,829 8
10^3	168	956,245	996,681	0,996 7	0,945 9
10^4	1229	9895,991	10013,397	1,001 3	0,986 2

Léase la columna $\psi(x)/x$: con $x = 1000$ ya está en el 99,7 por ciento, y con 10^4 pasa de 1 por trece milésimas. Es una convergencia excelente. Compárese con $\pi(x)/\operatorname{li}(x)$, que con 10^4 todavía va un 1,4 por ciento por debajo.

Esa es la razón de que nadie demuestre el teorema de los números primos en su forma original: se demuestra para ψ, donde la asintótica es limpia, y se traduce al final. El precio de la traducción son los factores logarítmicos, que no afectan al término principal.

Teorema 34.38 (Órdenes medios)**.** *Con $\gamma = 0{,}577\,215\,7$ la constante de Euler–Mascheroni:*

$$\sum_{n\le x}\tau(n) = x\log x + (2\gamma-1)x + O(\sqrt{x}), \qquad \sum_{n\le x}\varphi(n) = \frac{3}{\pi^2}x^2 + O(x\log x),$$

$$\sum_{n\le x}\sigma(n) = \frac{\pi^2}{12}x^2 + O(x\log x).$$

La primera se demuestra con el **método de la hipérbola** de Dirichlet: contar los pares (a,b) con $ab\le x$ sumando primero los de $a\le\sqrt{x}$, luego los de $b\le\sqrt{x}$, y restando el cuadrado contado dos veces. El exponente $1/2$ del error se ha bajado hasta 0,3149 y el problema de determinar el óptimo —se conjetura $1/4$— sigue abierto: es el *problema del divisor de Dirichlet*.

Ejemplo 34.39 (Comprobación hasta mil)**.**

	valor exacto	asintótica
$\sum_{n\le1000}\tau(n)$	7 069	7 062,19
$\sum_{n\le1000}\varphi(n)$	304 192	303 963,55
libres de cuadrados ≤ 1000	608	607,93

La primera discrepancia es de 6,81, muy por debajo del $\sqrt{1000}=31{,}6$ que permite el término de error. Y el último renglón merece un comentario aparte: la densidad de los enteros libres de cuadrados es

$$\prod_p\Big(1-\frac{1}{p^2}\Big) = \frac{1}{\zeta(2)} = \frac{6}{\pi^2} = 0{,}607\,927,$$

y hasta mil hay exactamente 608. La lectura probabilística es directa: dos enteros elegidos al azar son coprimos con probabilidad $6/\pi^2\approx 60{,}8$ por ciento, porque ser divisibles a la vez por p tiene "probabilidad" $1/p^2$ y los sucesos son independientes por el teorema chino. Que π aparezca en una pregunta sobre divisibilidad es, en una frase, de qué va este capítulo.

34.9. Sumas exponenciales, sumas de Gauss y el método del círculo

Junto a la función zeta, el otro gran motor de la teoría analítica es la estimación de *sumas exponenciales*: expresiones $\sum e^{2\pi i f(n)}$ cuyo tamaño mide cuánto se cancelan las fases.

Definición 34.40 (Suma de Gauss)**.** Para un primo impar p,

$$g_p = \sum_{n=0}^{p-1} e^{2\pi i n^2/p} = \sum_{n=0}^{p-1}\Big(\frac{n}{p}\Big)e^{2\pi i n/p},$$

donde $\left(\frac{n}{p}\right)$ es el símbolo de Legendre.

Teorema 34.41. *$|g_p| = \sqrt{p}$, y con más precisión $g_p = \sqrt{p}$ si $p \equiv 1 \pmod 4$ y $g_p = i\sqrt{p}$ si $p \equiv 3 \pmod 4$.*

Ejemplo 34.42 (Módulo 5, calculado entero)**.** Los cuadrados no nulos módulo 5 son $1, 4, 4, 1$: los residuos cuadráticos son $\{1, 4\}$ y los no residuos $\{2, 3\}$. Luego

$$g_5 = 1 + 2e^{2\pi i/5} + 2e^{8\pi i/5} = 1 + 4\cos\tfrac{2\pi}{5} = 1 + 4(0{,}309\,017) = 2{,}236\,068 = \sqrt{5}.$$

Real y positivo, como predice el teorema para $p = 5 \equiv 1 \pmod 4$.

Lo notable es el tamaño. La suma tiene p términos de módulo 1; si no hubiera cancelación valdría p, y si las fases fueran totalmente aleatorias cabría esperar $\sqrt{p}$ por el paseo aleatorio. El teorema dice que las fases se comportan *exactamente* como si fueran aleatorias, ni mejor ni peor. Esa "raíz cuadrada de cancelación" es el objetivo de casi toda estimación en la materia, y cuando se consigue se dice que la estimación es de *tipo raíz cuadrada*.

La versión con caracteres generales es la desigualdad de **Pólya–Vinogradov**, $\left|\sum_{n \le N} \chi(n)\right| \le \sqrt{q}\log q$, que dice que un carácter no principal no puede tener el mismo signo durante mucho rato; de ahí salen las cotas para el menor no residuo cuadrático.

El método del círculo

Idea clave

Hardy y Littlewood observaron que el número de representaciones de N como suma de s elementos de un conjunto A es un *coeficiente de Fourier*. Si $f(\alpha) = \sum_{a \in A,\, a \le N} e^{2\pi i a\alpha}$, entonces

$$r_s(N) = \int_0^1 f(\alpha)^s\, e^{-2\pi i N\alpha}\, \mathrm{d}\alpha.$$

La integral es exacta y no dice nada por sí sola. El truco está en partir el intervalo en dos: los **arcos mayores**, entornos minúsculos de los racionales de denominador pequeño, donde f es grande y se puede aproximar bien; y los **arcos menores**, todo lo demás, donde hay que *demostrar* que f es pequeña por cancelación.

Los arcos mayores producen el término principal, que siempre tiene la misma forma:

$$r_s(N) \;\approx\; \mathfrak{S}(N) \cdot \frac{N^{s/k-1}}{\Gamma(\cdot)},$$

donde $\mathfrak{S}(N) = \prod_p \sigma_p(N)$ es la **serie singular**: un producto sobre primos de las densidades locales, es decir, de cuántas soluciones hay módulo cada potencia de primo. Los arcos menores solo tienen que dar un error más pequeño, y ahí es donde entran las sumas exponenciales.

Regla: qué ha resuelto el método

- **Problema de Waring.** Todo entero suficientemente grande es suma de s potencias k-ésimas si s es bastante mayor que k; el método da la asintótica exacta del número de representaciones.

- **Goldbach ternario** (Vinogradov, 1937): todo impar suficientemente grande es suma de tres primos. Helfgott completó el rango bajo en 2013, cerrando el problema.
- **Goldbach binario** sigue abierto, y la razón es estructural: con solo dos sumandos los arcos menores no se pueden controlar. El método necesita "espacio de maniobra", y dos no bastan.

Ejemplo 34.43 (La constante de los primos gemelos, deducida)**.** La heurística del círculo aplicada a los gemelos da como densidad local en el primo q el factor

$$\sigma_q = \frac{q(q-2)}{(q-1)^2},$$

que compara el número real de pares admisibles módulo q con lo que darían dos eventos independientes. El producto sobre los primos impares es

$$C_2 = \prod_{q>2} \frac{q(q-2)}{(q-1)^2} = 0{,}660\,161\,8,$$

y la conjetura de Hardy–Littlewood predice

$$\pi_2(x) \sim 2C_2 \int_2^x \frac{\mathrm{d}t}{\log^2 t}, \qquad 2C_2 = 1{,}320\,323\,6.$$

Con $x = 10^6$ la fórmula da 8 248 pares de primos gemelos, y el recuento real es 8 169: un error del uno por ciento. La constante que el capítulo usaba sin explicar es, pues, el producto de las correcciones locales que el método del círculo predice.

Observación 34.44 (Cribas modernas y el problema de la paridad)*.* La criba de Brun se refina en dos direcciones. La de **Selberg** sustituye los signos alternos por unos pesos λ_d libres y minimiza una forma cuadrática en ellos: el óptimo se calcula exactamente y da cotas superiores mucho mejores. La **criba grande** y la desigualdad de **Bombieri–Vinogradov** controlan la distribución de los primos en progresiones *en promedio* sobre los módulos, y sirven como sustituto incondicional de la hipótesis de Riemann generalizada en muchísimas aplicaciones.

Pero hay un techo. El **problema de la paridad**, aislado por Selberg, dice que ningún método de criba puro puede distinguir entre números con un número par de factores primos y con un número impar. Por eso Chen (1973) solo pudo llegar a "$p+2$ es primo o producto de dos primos", y por eso los huecos acotados entre primos (Zhang 2013, Maynard y Tao 2013) necesitaron una idea nueva: pesos multidimensionales sobre tuplas admisibles, que rompen la barrera desde otro ángulo.

34.10. La criba de Selberg: optimizar en vez de alternar signos

La criba de Eratóstenes con inclusión–exclusión tiene un defecto fatal: el término de error crece como $2^{\pi(z)}$ y devora el término principal en cuanto z pasa de ser minúsculo. Brun lo remedió truncando la alternancia. Selberg tuvo una idea mejor: no truncar, sino *optimizar*.

Idea clave

El punto de partida es una desigualdad de una línea. Para cualesquiera números reales λ_d con $\lambda_1 = 1$,

$$\mathbf{1}_{n \text{ no cribado}} \;\le\; \Big(\sum_{d\mid n,\ d\le D} \lambda_d\Big)^2,$$

porque si n sobrevive a la criba su único divisor relevante es 1 y el cuadrado vale 1; y si no sobrevive, el cuadrado es ≥ 0. La desigualdad vale *sean cuales sean* los λ_d. Así que se suma sobre $n \le N$, se obtiene una forma cuadrática en los λ, y se *minimiza*. No hay que adivinar nada: el óptimo es un problema de álgebra lineal.

Teorema 34.45 (La forma y su mínimo)**.** *Con $|A_d| \approx N\,g(d)$ multiplicativa,*

$$\sum_{n\le N}\Big(\sum_{d\mid n,\, d\le D}\lambda_d\Big)^2 \;\approx\; N\sum_{d_1,d_2\le D}\frac{\lambda_{d_1}\lambda_{d_2}}{[d_1,d_2]} + O(D^2),$$

y el mínimo de esa forma cuadrática sujeto a $\lambda_1 = 1$ vale exactamente $1/G(D)$, con

$$G(D) = \sum_{\substack{d\le D\\ d\mid P(z)}} \prod_{p\mid d}\frac{1}{p-1}.$$

Ejemplo 34.46 (Cribar $\{1,\dots,N\}$ por 2, 3 y 5, con los λ explícitos)**.** La respuesta exacta se conoce —la densidad es $\frac12\cdot\frac23\cdot\frac45 = \frac{4}{15} = 0{,}266\,667$— y eso permite ver funcionar la criba con red debajo.

(a) La cota según el nivel D.

D	$G(D)$	cota $N/G(D)$
1	1	$1{,}000\,000\,N$
2	2	$0{,}500\,000\,N$
3	2,5	$0{,}400\,000\,N$
5	2,75	$0{,}363\,636\,N$
6	3,25	$0{,}307\,692\,N$
10	3,5	$0{,}285\,714\,N$
15	3,625	$0{,}275\,862\,N$
30	3,75	$0{,}266\,667\,N$

Con $D = 30$ la cota es *exacta*: $G(30) = \prod_{p\le 5}\frac{p}{p-1} = 2\cdot\frac32\cdot\frac54 = \frac{15}{4}$, y su inverso es $\frac{4}{15}$. Y a medida que D baja, la cota se degrada de forma controlada y monótona. Ese es el compromiso de toda criba: D grande da mejor término principal y peor error, porque el error es $O(D^2)$.

(b) Los pesos óptimos con $D = 6$. Resolviendo el sistema lineal que anula el gradiente,

$$\lambda_1 = 1,\quad \lambda_2 = \lambda_3 = -\tfrac{12}{13} = -0{,}923\,077,\quad \lambda_5 = -\tfrac{5}{13} = -0{,}384\,615,\quad \lambda_6 = \tfrac{12}{13} = 0{,}923\,077.$$

Compárense con los de inclusión–exclusión, que serían $\mu(d)$: $1,-1,-1,-1,+1$. Los de Selberg tienen los *mismos signos* —la estructura de Möbius sobrevive— pero magnitudes menores

que 1, y menores cuanto mayor es el primo. La criba no dice "quita todos los múltiplos de 5"; dice "quita el 38,5 por ciento de ellos", porque quitarlos todos sobrepasaría. Esa moderación cuantitativa es exactamente lo que evita que los errores se acumulen.

(c) Comprobación. El valor mínimo de la forma cuadrática con esos λ es $0{,}307\,692\,31 = 1/3{,}25 = 1/G(6)$, tal como afirma el teorema. ✓

Cuidado

Con toda su elegancia, la criba de Selberg da *cotas superiores*, y las inferiores son mucho más difíciles: hacen falta combinaciones de dos cribas de Selberg (el método de Jurkat–Richert) y siempre se pierde un factor. Ese es el motivo estructural de que las cribas demuestren "hay pocos" con facilidad y "hay muchos" casi nunca, y de que el problema de la paridad —ningún método de criba puro distingue un número con dos factores primos de uno con tres— no se pueda esquivar desde dentro de la teoría.

34.11. La notación asintótica y las funciones aritméticas, por escrito

El capítulo entero consiste en estimaciones, y el símbolo con el que se escriben no estaba definido. Tampoco lo estaban la función de Möbius ni la convolución que un ejercicio pide usar.

Definición 34.47 (Notación de Landau)**.** Sean f, g definidas para x grande, con $g(x) > 0$. Se escribe:

- $f(x) = O(g(x))$ —"f es a lo sumo del orden de g"— si existen constantes $C > 0$ y x_0 tales que $|f(x)| \le C\, g(x)$ para todo $x \ge x_0$. La constante C se llama *constante implícita*. La notación de Vinogradov $f \ll g$ significa lo mismo.
- $f(x) = o(g(x))$ si $\displaystyle\lim_{x\to\infty} \frac{f(x)}{g(x)} = 0$: es estrictamente más fuerte.
- $f(x) \sim g(x)$ —"f es asintóticamente igual a g"— si $\displaystyle\lim_{x\to\infty} \frac{f(x)}{g(x)} = 1$. Equivale a $f = g + o(g)$.
- $f(x) = \Theta(g(x))$ si $f = O(g)$ y $g = O(f)$ a la vez.

Cuidado

[Tres trampas de la notación O] *Primera: el signo igual no es una igualdad.* $O(g)$ denota una *clase* de funciones, y $f = O(g)$ debe leerse $f \in O(g)$. Por eso la relación no es simétrica: se escribe $x = O(x^2)$ pero jamás $O(x^2) = x$. Y una cadena como $O(x) + O(x) = O(x)$ es cierta, mientras que $O(x) - O(x) = 0$ es falsa.
Segunda: hay que decir en qué variable y sobre qué rango. La afirmación "$f(x, y) = O(x)$" es ambigua si y también varía: hay que especificar si la constante implícita es *uniforme*

en y o depende de y. En este capítulo, salvo aviso, todas las constantes son absolutas y el límite es $x \to \infty$.
Tercera: O acota, no describe. Que $\pi(x) = \mathrm{li}(x) + O(\sqrt{x}\ln x)$ sea el enunciado equivalente a la hipótesis de Riemann no significa que el error *sea* de ese tamaño, sino que *no lo supera*. De hecho se sabe que el error cambia de signo infinitas veces y que no es $o(\sqrt{x}\,\ln\ln\ln x/\ln x)$: la cota superior y el tamaño real son cosas distintas.

Definición 34.48 (Convolución de Dirichlet y función de Möbius)**.** Dadas dos funciones aritméticas $f, g : \mathbb{N} \to \mathbb{C}$, su *convolución de Dirichlet* es

$$(f * g)(n) = \sum_{d \mid n} f(d)\, g\!\left(\frac{n}{d}\right).$$

La *función de Möbius* es

$$\mu(n) = \begin{cases} 1, & n = 1, \\ (-1)^k, & n = p_1 p_2 \cdots p_k \text{ con los } p_i \text{ primos distintos}, \\ 0, & \text{si algún primo divide a } n \text{ con exponente } \geq 2. \end{cases}$$

Se escriben además $\mathbf{1}(n) = 1$ para todo n, y $\delta(n) = 1$ si $n = 1$ y 0 en otro caso.

Teorema 34.49 (Inversión de Möbius)**.** *La convolución de Dirichlet es conmutativa y asociativa, tiene a δ por elemento neutro, y*

$$\mu * \mathbf{1} = \delta.$$

En consecuencia, para funciones aritméticas cualesquiera,

$$F(n) = \sum_{d \mid n} f(d) \qquad \Longleftrightarrow \qquad f(n) = \sum_{d \mid n} \mu(d)\, F\!\left(\frac{n}{d}\right).$$

Demostración. La conmutatividad se ve reindexando $d \mapsto n/d$, y la asociatividad reescribiendo ambos lados como $\sum_{abc=n} f(a)g(b)h(c)$. Para la identidad clave hay que probar $\sum_{d\mid n}\mu(d) = \delta(n)$. Es evidente para $n = 1$. Para $n > 1$ con factorización $n = p_1^{a_1}\cdots p_k^{a_k}$, sólo contribuyen los divisores libres de cuadrados, que son los productos de subconjuntos de $\{p_1, \ldots, p_k\}$; el que tiene j primos aporta $(-1)^j$ y hay $\binom{k}{j}$ de ellos, luego

$$\sum_{d \mid n} \mu(d) = \sum_{j=0}^{k} \binom{k}{j} (-1)^j = (1-1)^k = 0.$$

La inversión es ahora álgebra: $F = f * \mathbf{1}$ implica $F * \mu = f * \mathbf{1} * \mu = f * \delta = f$, y recíprocamente. □

Ejemplo 34.50 (Despejar la función φ de Euler)**.** La identidad $\sum_{d\mid n}\varphi(d) = n$ —clasificar los n elementos de $\mathbb{Z}/n$ por su orden— se escribe $\varphi * \mathbf{1} = \mathrm{Id}$. Invirtiendo,

$$\varphi(n) = \sum_{d \mid n} \mu(d)\, \frac{n}{d} = n \prod_{p \mid n} \left(1 - \frac{1}{p}\right),$$

donde la última igualdad sale de que sólo cuentan los divisores libres de cuadrados y el producto se factoriza primo a primo. Para $n = 12 = 2^2 \cdot 3$: $12(1-\tfrac12)(1-\tfrac13) = 12 \cdot \tfrac12 \cdot \tfrac23 = 4$, y en efecto los coprimos con 12 menores que él son $1, 5, 7, 11$.

Regla: Pasar de una estimación de ψ a una de π

El teorema de los números primos se demuestra para $\psi(x) = \sum_{p^k \le x} \ln p$ y se enuncia para $\pi(x)$. El puente son dos pasos mecánicos.

1. **De ψ a θ.** Con $\theta(x) = \sum_{p \le x} \ln p$, la diferencia sólo recoge las potencias:
$$\psi(x) - \theta(x) = \sum_{k \ge 2} \theta(x^{1/k}) = O(\sqrt{x}\,\ln^2 x),$$
porque sólo hay $O(\ln x)$ valores de k que contribuyen y cada término es $O(\sqrt{x}\ln x)$. Luego $\psi \sim x$ equivale a $\theta \sim x$.

2. **De θ a π, por sumación de Abel.** La *fórmula de sumación parcial* dice que si $A(t) = \sum_{n \le t} a_n$ y ϕ es de clase C^1,
$$\sum_{n \le x} a_n\,\phi(n) = A(x)\phi(x) - \int_1^x A(t)\,\phi'(t)\,\mathrm{d}t.$$
Aplicada con $a_n = \ln n$ si n es primo y 0 si no, y $\phi(t) = 1/\ln t$, da
$$\pi(x) = \frac{\theta(x)}{\ln x} + \int_2^x \frac{\theta(t)}{t\,\ln^2 t}\,\mathrm{d}t.$$
Sustituyendo $\theta(t) = t + o(t)$ y evaluando la integral se obtiene $\pi(x) \sim \dfrac{x}{\ln x}$, y con el desarrollo completo, $\pi(x) = \mathrm{li}(x) + \text{error}$.

Por qué se trabaja con ψ y no con π. Porque ψ es la función cuya transformada de Mellin es $-\zeta'/\zeta$, y es ahí —y no en π— donde los ceros de ζ aparecen como polos. Contar primos con peso $\ln p$ no es un capricho: es lo que hace que la fórmula explícita salga limpia.

34.12. Ejercicios resueltos

Tres problemas que usan funciones aritméticas y la zeta para decir algo sobre los primos.

Ejercicio 34.1. Demuestre que $\sum_{d \mid n} \varphi(d) = n$ y compruébelo con $n = 12$.

Solución. Clasifique los enteros de 1 a n según su máximo común divisor con n. Los que cumplen $\gcd(k, n) = n/d$ son exactamente los de la forma $k = (n/d)j$ con $\gcd(j, d) = 1$ y $1 \le j \le d$, y hay $\varphi(d)$ de ellos. Como cada k cae en exactamente una clase, sumando sobre los divisores se obtiene n.

Con $n = 12$: los divisores son $1, 2, 3, 4, 6, 12$ y φ vale $1, 1, 2, 2, 2, 4$, que suman 12. La identidad dice, en el lenguaje de las series de Dirichlet, que $\varphi * \mathbf{1} = \mathrm{Id}$, y ése es el verdadero motivo por el que funciona. ∎

Ejercicio 34.2. Enuncie la inversión de Möbius y úsela para despejar φ de la identidad anterior.

Solución. Si $g(n)=\sum_{d\mid n} f(d)$ para todo n, entonces $f(n)=\sum_{d\mid n}\mu(n/d)\,g(d)$, donde μ vale 1 si el argumento es producto de un número par de primos distintos, -1 si es impar y 0 si tiene algún factor cuadrado.

Aplicándolo resulta $\varphi(n)=\sum_{d\mid n}\mu(n/d)\,d = n\prod_{p\mid n}\bigl(1-\tfrac{1}{p}\bigr)$, que es la fórmula habitual. La inversión de Möbius es, en el fondo, la observación de que μ es el inverso de la función constante $\mathbf{1}$ para la convolución de Dirichlet: $\mu * \mathbf{1} = \delta$. Todo el aparato es álgebra en un anillo de funciones aritméticas. ■

Ejercicio 34.3. Demuestre la fórmula del producto de Euler $\zeta(s)=\prod_p (1-p^{-s})^{-1}$ para $\Re s>1$ y explique por qué implica que hay infinitos primos.

Solución. Desarrolle cada factor como serie geométrica: $(1-p^{-s})^{-1}=1+p^{-s}+p^{-2s}+\cdots$. Al multiplicar sobre todos los primos y expandir, cada término del producto es n^{-s} para exactamente un n, por el teorema fundamental de la aritmética. La convergencia absoluta para $\Re s>1$ justifica el reordenamiento.

Para la infinitud de los primos, tome $s\to 1^{+}$. El lado izquierdo tiende a infinito, porque la serie armónica diverge. Si hubiese finitos primos, el lado derecho sería un producto finito de números finitos y por tanto acotado. Contradicción. Es la demostración de Euler, y dice bastante más que la de Euclides: refinándola se obtiene que $\sum 1/p$ diverge, de modo que los primos no sólo son infinitos sino que son abundantes, mucho más que los cuadrados. ■

34.13. Ejercicios propuestos

34.1. Calcule $\mu(12)$, $\mu(30)$ y $\mu(1)$.

34.2. Compruebe $\sum_{d\mid 30}\varphi(d)=30$.

34.3. ¿Cuánto valen $\zeta(2)$ y $\zeta(4)$?

34.4. ¿Es multiplicativa la función σ_1? Calcule $\sigma_1(12)$.

34.5. ¿Cuántos primos hay hasta 100? Compare con $100/\ln 100$ y con $\mathrm{li}(100)$.

34.6. Enumere los pares de primos gemelos menores que 100.

34.7. ¿Qué afirma el teorema de Dirichlet sobre progresiones aritméticas? ¿Cuántos primos hay de la forma $4k+3$?

34.8. ¿Dónde tiene ζ su único polo y cuál es su residuo?

34.9. ¿Qué son los ceros triviales de ζ y de dónde salen?

34.10. Enuncie la hipótesis de Riemann en términos del error de $\pi(x)$.

34.11. ¿Qué dice el postulado de Bertrand? ¿Quién lo demostró?

34.12. ¿Por qué la divergencia de $\sum 1/p$ es una afirmación más fuerte que la infinitud de los primos?

Parte VI

Geometría, topología y abstracción

Capítulo 35

Geometría algebraica

Lo que necesitas traer

Del capítulo 24: anillos conmutativos, ideales, ideales primos y maximales, y cocientes.
Del capítulo 25: cuerpos algebraicamente cerrados y el teorema de la base de Hilbert.
Del capítulo 9: derivadas parciales, para hablar de puntos singulares.

La geometría algebraica estudia los conjuntos definidos por ecuaciones polinómicas. Su idea motriz es un diccionario: a cada figura geométrica le corresponde un anillo de funciones, y a cada propiedad geométrica una propiedad algebraica. Trabajar con figuras se convierte así en trabajar con ideales, y las herramientas del álgebra conmutativa pasan a ser herramientas de geometría.

35.1. Variedades e ideales: el diccionario

Definición 35.1. Dado un conjunto S de polinomios en $K[x_1, \dots, x_n]$, su *variedad afín* es $V(S) = \{\mathbf{p} : \ f(\mathbf{p}) = 0 \ \forall f \in S\}$. Recíprocamente, dado un conjunto X de puntos, su *ideal* es $I(X) = \{f : \ f(\mathbf{p}) = 0 \ \forall \mathbf{p} \in X\}$. El *Nullstellensatz* de Hilbert afirma que, sobre un cuerpo algebraicamente cerrado, $I(V(J)) = \sqrt{J}$, el radical de J.

Idea clave

Qué dice esto en cristiano. Las operaciones V e I van en sentidos opuestos y forman un diccionario que invierte inclusiones: cuantas más ecuaciones, más pequeña la figura. El Nullstellensatz dice hasta qué punto ese diccionario es fiel, y la respuesta es: casi del todo, salvo por el detalle de las multiplicidades. Los ideales (x) y (x^2) definen el mismo conjunto de puntos (el origen), y solo se distinguen por el radical. Sobre cuerpos no algebraicamente cerrados el diccionario se rompe, y por eso la geometría algebraica trabaja sobre $\mathbb{C}$ o sobre clausuras algebraicas.

Ejemplo 35.2 (Del álgebra a la geometría y vuelta)**. (a) Casos básicos.** $V(y - x^2)$ es una parábola; $V(x^2 + y^2 - 1)$ una circunferencia; $V(xy)$ la unión de los dos ejes; $V(x, y)$ el origen

solo. Cuantos más polinomios, menos puntos.

(b) La suma y la intersección. $V(f) \cap V(g) = V(f, g)$: intersecar figuras corresponde a sumar ideales. Y $V(f) \cup V(g) = V(fg)$: unir corresponde a multiplicar. Los dos ejes son $V(xy) = V(x) \cup V(y)$, y esa factorización refleja que la figura no es irreducible.

(c) El radical, con números. El ideal (x^2) y el ideal (x) tienen la misma variedad, el eje y, pero solo el segundo es radical. El Nullstellensatz dice que $I(V((x^2))) = (x)$: al ir y volver se pierde la multiplicidad. Por eso el diccionario limpio es entre variedades e ideales *radicales*.

(d) Dónde falla sobre $\mathbb{R}$. El ideal $(x^2 + y^2)$ en $\mathbb{R}[x, y]$ tiene por variedad real solo el origen, y sin embargo su radical no es (x, y). Sobre $\mathbb{C}$ la variedad son dos rectas, $y = \pm ix$, y todo funciona. La necesidad de un cuerpo algebraicamente cerrado no es técnica: sobre $\mathbb{R}$ el diccionario es sencillamente falso.

Cuidado

[La dimensión y el grado se comportan de forma poco intuitiva] Una curva plana de grado d y otra de grado e se cortan en exactamente *de* puntos, pero solo si se cuenta en el plano *proyectivo*, sobre $\mathbb{C}$ y con multiplicidades: ese es el teorema de Bézout. En el plano afín real, dos circunferencias pueden no cortarse, dos rectas paralelas tampoco, y la cuenta falla. Segundo aviso: una variedad puede tener puntos *singulares* donde no hay tangente bien definida, como el origen de $y^2 = x^3$ (una cúspide) o de $y^2 = x^2(x + 1)$ (un cruce); la mayor parte de la teoría fina consiste en manejarlos.

Pausa para pensar

Descartes fundó la geometría analítica en 1637 al traducir figuras en ecuaciones; la geometría algebraica hace lo contrario, traducir ecuaciones en figuras, y de paso convertir el álgebra en geometría. La reformulación decisiva es de Grothendieck en los años sesenta, con los *esquemas*: en vez de partir de puntos, se parte del anillo y se *definen* los puntos como sus ideales primos. Suena a giro gratuito y resultó ser el marco donde caben a la vez la geometría sobre $\mathbb{C}$ y la aritmética sobre $\mathbb{Z}$, lo que permitió a Wiles demostrar el último teorema de Fermat en 1995 usando geometría de curvas elípticas.

35.2. El diccionario por escrito: conjuntos algebraicos, topología de Zariski y división de polinomios

La sección anterior anunció un diccionario entre ecuaciones y figuras y lo usó de inmediato. Pero casi todas las piezas de ese diccionario se han nombrado sin construirse: qué es exactamente el espacio donde viven los puntos, por qué los conjuntos de soluciones forman una *topología*, qué significa que un ideal sea radical, qué autoriza a llamar «variedad» a una figura y qué quiere decir «dividir» un polinomio cuando hay varias variables y varios divisores. Esta sección construye todo eso desde cero, con demostraciones cuando son asequibles y con cuentas hechas siempre. Quien la lea entera podrá leer el resto del capítulo sin dar nada por sabido.

El espacio afín y los conjuntos algebraicos

Definición 35.3 (Espacio afín y conjunto algebraico)**.** Sea K un cuerpo. El *espacio afín* de dimensión n sobre K es el conjunto

$$\mathbb{A}^n(K) = \{(a_1, \ldots, a_n) \ : \ a_i \in K\},$$

que como conjunto es K^n, pero del que se olvidan la suma y el producto por escalares: en geometría afín no hay un origen privilegiado ni tiene sentido sumar dos puntos. Sus elementos se llaman *puntos*. Cuando K está claro se escribe $\mathbb{A}^n$.

Dado un conjunto $S \subseteq K[x_1, \ldots, x_n]$ de polinomios, su *conjunto de ceros* es

$$V(S) = \{\mathbf{p} \in \mathbb{A}^n \ : \ f(\mathbf{p}) = 0 \ \text{ para toda } f \in S\},$$

y un subconjunto $X \subseteq \mathbb{A}^n$ se llama *conjunto algebraico afín* si $X = V(S)$ para algún S.

Observación 35.4 (Siempre bastan finitas ecuaciones). Si $I = (S)$ es el ideal generado por S, entonces $V(S) = V(I)$: toda combinación $\sum h_i f_i$ se anula donde se anulan los f_i, y recíprocamente $S \subseteq I$. Y por el teorema de la base de Hilbert (capítulo 24) ese ideal está generado por un número finito de polinomios $f_1, \ldots, f_r$, luego

$$V(S) = V(f_1, \ldots, f_r).$$

Un sistema con infinitas ecuaciones polinómicas es siempre equivalente a uno finito. Es el primer regalo del álgebra a la geometría, y se usa continuamente sin mencionarlo.

Ejemplo 35.5 (Seis conjuntos algebraicos, calculados)**. (a) Los dos extremos.** $V(0) = \mathbb{A}^n$, porque el polinomio nulo se anula en todas partes, y $V(1) = \varnothing$, porque la constante 1 no se anula en ninguna. El espacio entero y el vacío son conjuntos algebraicos.

(b) Un punto. En $\mathbb{A}^2$, $V(x, y) = \{(0, 0)\}$; y en general $V(x_1 - a_1, \ldots, x_n - a_n) = \{(a_1, \ldots, a_n)\}$. Todo punto es un conjunto algebraico.

(c) Una intersección resuelta. $V(y - x^2,\, x - 1)$: la segunda ecuación fuerza $x = 1$ y la primera da entonces $y = 1$. Luego $V(y - x^2, x - 1) = \{(1, 1)\}$, un solo punto, aunque cada ecuación por separado defina una curva.

(d) Depende del cuerpo. $V(x^2 + 1) \subset \mathbb{A}^1$ es vacío si $K = \mathbb{R}$ y es $\{\mathrm{i}, -\mathrm{i}\}$ si $K = \mathbb{C}$. El conjunto de ceros no es una propiedad del polinomio solamente: es una propiedad del polinomio y del cuerpo donde se buscan las soluciones.

(e) Una figura que no es algebraica. La gráfica del seno, $\{(x, \sin x)\} \subset \mathbb{A}^2(\mathbb{R})$, no es un conjunto algebraico: corta a la recta $y = 0$ en infinitos puntos, y un polinomio no nulo restringido a una recta tiene un número finito de raíces. Ser algebraico es una condición fuerte.

(f) Un conjunto finito cualquiera. $\{a_1, \ldots, a_m\} \subset \mathbb{A}^1$ es $V\big((x - a_1) \cdots (x - a_m)\big)$. Todo subconjunto finito de la recta es algebraico.

Proposición 35.6 (Cómo se combinan los conjuntos de ceros)**.** *Para ideales I, J, I_α de $K[x_1, \ldots, x_n]$:*

1. *si $I \subseteq J$ entonces $V(J) \subseteq V(I)$ —más ecuaciones, menos puntos—;*

2. $V(0) = \mathbb{A}^n$ *y* $V(1) = \varnothing$*;*

3. $\bigcap_\alpha V(I_\alpha) = V\big(\sum_\alpha I_\alpha\big)$*, para cualquier familia, incluso infinita;*

4. $V(I) \cup V(J) = V(I \cap J) = V(IJ)$.

Demostración. (1) y (2) son inmediatos. Para (3), un punto anula a todos los elementos de todos los I_α si y sólo si anula a todas sus sumas finitas, que es lo que hay en $\sum I_\alpha$.

La única parte con contenido es (4), y basta probar $V(IJ) \subseteq V(I) \cup V(J)$, porque las inclusiones $V(I) \cup V(J) \subseteq V(I \cap J) \subseteq V(IJ)$ salen de (1) y de $IJ \subseteq I \cap J$. Sea $\mathbf{p} \in V(IJ)$ y supóngase que $\mathbf{p} \notin V(I)$: existe $f \in I$ con $f(\mathbf{p}) \neq 0$. Para cualquier $g \in J$ se tiene $fg \in IJ$, luego $f(\mathbf{p})g(\mathbf{p}) = 0$; como K es un cuerpo y $f(\mathbf{p}) \neq 0$, resulta $g(\mathbf{p}) = 0$. Esto vale para toda $g \in J$, así que $\mathbf{p} \in V(J)$. □

Definición 35.7 (Topología de Zariski)**.** La proposición anterior dice exactamente que los conjuntos algebraicos de $\mathbb{A}^n$ cumplen los axiomas de los cerrados de una topología: el total y el vacío lo son, la intersección arbitraria de cerrados es cerrada y la unión *finita* de cerrados es cerrada. La topología que resulta —cerrados = conjuntos algebraicos, abiertos = sus complementarios— se llama *topología de Zariski* de $\mathbb{A}^n$.

Ejemplo 35.8 (La topología de Zariski en la recta y en el plano)**.** Tómese K infinito, por ejemplo $K = \mathbb{C}$.

(a) Los cerrados de $\mathbb{A}^1$. Un polinomio no nulo de grado d tiene a lo sumo d raíces, luego $V(f)$ es finito salvo si $f = 0$. Así que los cerrados de la recta afín son exactamente los conjuntos finitos y la recta entera. Los abiertos no vacíos son los complementarios de conjuntos finitos: enormes.

(b) No es de Hausdorff, ni de lejos. Dos abiertos no vacíos de $\mathbb{A}^1$ son complementarios de dos conjuntos finitos, y su intersección es el complementario de un conjunto finito, que no es vacío porque K es infinito. Dos puntos distintos *nunca* se pueden separar por abiertos disjuntos. Equivalentemente: todo abierto no vacío es denso.

(c) No es la topología producto. La diagonal $\Delta = V(x-y) \subset \mathbb{A}^2$ es cerrada para Zariski, por definición. En cambio, en la topología producto de $\mathbb{A}^1 \times \mathbb{A}^1$ los cerrados se construyen a partir de productos de cerrados de la recta, es decir, de conjuntos finitos y rectas horizontales o verticales, y una unión finita de esos no puede ser la diagonal, que es infinita y no contiene ninguna recta horizontal ni vertical. La topología de Zariski del plano es estrictamente más fina que el producto de las de la recta.

(d) Qué se conserva. Pese a todo, las nociones topológicas siguen funcionando formalmente: clausura, densidad, irreducibilidad, compacidad. Lo que no sobrevive es la intuición métrica. Conviene leer «cerrado de Zariski» como «solución de un sistema de ecuaciones» y no como «conjunto que contiene sus límites».

De los puntos a las ecuaciones: el ideal, el radical y el Nullstellensatz

Definición 35.9 (Ideal de un conjunto, radical de un ideal)**.** Dado $X \subseteq \mathbb{A}^n$, su *ideal* es

$$I(X) = \{f \in K[x_1, \ldots, x_n] \ : \ f(\mathbf{p}) = 0 \ \text{ para todo } \mathbf{p} \in X\}.$$

Dado un ideal J, su *radical* es

$$\sqrt{J} = \{f \ : \ f^m \in J \ \text{ para algún entero } m \geq 1\},$$

que es un ideal, y J se llama *radical* si $J = \sqrt{J}$.

Proposición 35.10 (Las cuatro propiedades básicas)**.** *1. $I(X)$ es un ideal radical.*

2. Si $X \subseteq Y$ entonces $I(Y) \subseteq I(X)$.

3. $X \subseteq V(I(X))$, con igualdad si y sólo si X es un conjunto algebraico.

4. $\sqrt{J} \subseteq I(V(J))$ para todo ideal J.

Demostración. (1) Que sea ideal es inmediato: si f y g se anulan en X, también $f+g$ y hf. Y si $f^m \in I(X)$, entonces $f(\mathbf{p})^m = 0$ para todo $\mathbf{p} \in X$; como K es un cuerpo, no hay nilpotentes no nulos y $f(\mathbf{p}) = 0$, luego $f \in I(X)$.

(2) es evidente. (3): todo punto de X anula por definición a todos los elementos de $I(X)$. Si $X = V(S)$ es algebraico, entonces $S \subseteq I(X)$, luego $V(I(X)) \subseteq V(S) = X$ y hay igualdad; recíprocamente, $V(I(X))$ es algebraico, así que la igualdad obliga a que X lo sea.

(4): si $f^m \in J$, en cada punto de $V(J)$ se tiene $f(\mathbf{p})^m = 0$ y por tanto $f(\mathbf{p}) = 0$. □

Ejemplo 35.11 (Tres ideales calculados a mano)**. (a) El ideal de un punto.** $I(\{(0,0)\}) = (x,y)$ en $K[x,y]$. En efecto, (x,y) está contenido en él; y si f se anula en el origen, escribiendo $f = c + (\text{términos con } x) + (\text{términos con } y \text{ solo})$ se ve que el término independiente es $c = f(0,0) = 0$, y todo lo demás es múltiplo de x o de y. Luego $f \in (x,y)$.

(b) La multiplicidad se pierde al ir y volver. $V(x^2) \subset \mathbb{A}^2$ es el eje y, y I de ese eje es (x), porque un polinomio se anula en $x = 0$ exactamente cuando no tiene término libre en x. De modo que $I(V(x^2)) = (x) \neq (x^2)$, y en cambio $\sqrt{(x^2)} = (x)$: el resultado es el radical, tal como anuncia el Nullstellensatz.

(c) El ideal del espacio total. Si K es infinito, $I(\mathbb{A}^n) = (0)$: un polinomio que se anula en todos los puntos es el polinomio nulo. Para $n = 1$ es la cuenta de las raíces; para n general se hace por inducción fijando x_n. *Si K es finito la afirmación es falsa*: en $\mathbb{F}_2[x]$ el polinomio $x^2 - x$ se anula en los dos puntos de $\mathbb{A}^1(\mathbb{F}_2)$ y no es nulo.

Teorema 35.12 (Nullstellensatz de Hilbert, forma débil y forma fuerte)**.** *Sea K algebraicamente cerrado y $J \subseteq K[x_1, \ldots, x_n]$ un ideal.*

1. Forma débil*: si $V(J) = \varnothing$, entonces $J = (1)$, es decir, $1 \in J$.*

2. Forma fuerte*: $I(V(J)) = \sqrt{J}$.*

Demostración. Se admite la forma débil, cuya demostración necesita el lema de normalización de Noether, y se deduce la fuerte con el *truco de Rabinowitsch*, que es una idea de una línea y merece verse.

Por la proposición 35.10 ya se sabe $\sqrt{J} \subseteq I(V(J))$. Para la otra inclusión, sea $f \in I(V(J))$, es decir, f se anula en todos los ceros comunes de J. Escríbase $J = (f_1, \dots, f_r)$ e introdúzcase una *variable nueva* t, considerando en $K[x_1, \dots, x_n, t]$ el ideal

$$J' = (f_1, \dots, f_r,\ 1 - t\,f).$$

Ese ideal no tiene ningún cero: en un cero habría que anular todos los f_i, luego el punto $(\mathbf{p})$ estaría en $V(J)$ y por tanto $f(\mathbf{p}) = 0$, con lo que la última ecuación daría $1 - t \cdot 0 = 1 = 0$, absurdo. Por la forma débil, $1 \in J'$, o sea

$$1 = \sum_{i=1}^{r} h_i(\mathbf{x}, t)\, f_i(\mathbf{x}) + h_0(\mathbf{x}, t)\,(1 - t f(\mathbf{x}))$$

para ciertos polinomios h_i. Ahora el truco: esta identidad vale en el cuerpo de fracciones, así que se puede *sustituir* $t = 1/f$, lo que anula el último sumando:

$$1 = \sum_{i=1}^{r} h_i\Big(\mathbf{x}, \tfrac{1}{f}\Big)\, f_i(\mathbf{x}).$$

Multiplicando por una potencia f^m suficientemente grande para eliminar todos los denominadores queda

$$f^m = \sum_{i=1}^{r} g_i(\mathbf{x})\, f_i(\mathbf{x}) \in J,$$

luego $f \in \sqrt{J}$. □

Ejemplo 35.13 (Un certificado de inconsistencia, escrito)**.** El sistema $xy = 1$, $x = 0$ no tiene solución en $\mathbb{A}^2(\mathbb{C})$: si $x = 0$ la primera ecuación dice $0 = 1$. La forma débil promete entonces que el 1 se escribe como combinación de los dos miembros izquierdos, y aquí se ve a simple vista:

$$y \cdot x \ - \ 1 \cdot (xy - 1) = 1.$$

Esa identidad es un *certificado* verificable de que el sistema es imposible: no hace falta buscar soluciones ni razonar, basta desarrollar el producto. Los algoritmos de bases de Gröbner producen certificados así de forma automática, y esa es una de las razones prácticas de todo el aparato.

Sobre $\mathbb{R}$ no hay certificado. El sistema $x^2 + 1 = 0$ tampoco tiene soluciones reales, y sin embargo $(x^2 + 1) \neq (1)$ en $\mathbb{R}[x]$, porque todo elemento de ese ideal tiene grado ≥ 2 o es nulo. La hipótesis de cierre algebraico no es un tecnicismo: sin ella el teorema es falso.

Irreducibilidad: cuándo una figura merece llamarse variedad

Definición 35.14 (Conjunto algebraico irreducible y variedad)**.** Un conjunto algebraico $X \neq \varnothing$ es *irreducible* si no se puede escribir $X = X_1 \cup X_2$ con X_1, X_2 cerrados propios de X. Una *variedad afín* es un conjunto algebraico irreducible.

Proposición 35.15. *X es irreducible si y sólo si $I(X)$ es un ideal primo.*

Demostración. Supóngase X irreducible y sean f, g con $fg \in I(X)$. Entonces $X \subseteq V(fg) = V(f) \cup V(g)$, luego $X = (X \cap V(f)) \cup (X \cap V(g))$, unión de dos cerrados de X. Por irreducibilidad uno de ellos es todo X; si es el primero, f se anula en X y $f \in I(X)$. Luego $I(X)$ es primo.

Recíprocamente, sea $I(X)$ primo y supóngase $X = X_1 \cup X_2$ con X_i cerrados propios. Como $X_1 \subsetneq X$, existe $f \in I(X_1) \setminus I(X)$, y análogamente $g \in I(X_2) \setminus I(X)$. El producto fg se anula en X_1 y en X_2, luego en X: $fg \in I(X)$ con $f \notin I(X)$ y $g \notin I(X)$, contra la primalidad. □

Ejemplo 35.16 (Una reducible y una irreducible, comprobadas)**. (a) Los dos ejes.** $X = V(xy)$ no es irreducible: $X = V(x) \cup V(y)$, dos cerrados propios. Y en efecto $I(X) = (xy)$ no es primo, porque $x \cdot y \in (xy)$ y ninguno de los dos factores está.

(b) La parábola. $X = V(y - x^2)$ sí es irreducible. La forma limpia de verlo es algebraica: el homomorfismo $K[x, y] \to K[x]$ que manda $x \mapsto x$ y $y \mapsto x^2$ es sobreyectivo y su núcleo es exactamente $(y - x^2)$ —dividir cualquier f entre $y - x^2$ como polinomio en y deja un resto sin y, que es $f(x, x^2)$—. Por tanto

$$K[x, y]/(y - x^2) \;\cong\; K[x],$$

que es un dominio de integridad, luego $(y - x^2)$ es primo y la parábola es irreducible. De paso, ese isomorfismo dice que las funciones polinómicas sobre la parábola son exactamente los polinomios en una variable: la parábola «es» una recta desde el punto de vista del álgebra.

Cuidado

[La palabra "variedad" se usa con dos sentidos] En la literatura —y en este capítulo— «variedad afín» significa unas veces *cualquier* conjunto algebraico y otras solamente los irreducibles. La diferencia importa: los dos ejes $V(xy)$ son un conjunto algebraico, pero no una variedad en sentido estricto, y su ideal es radical sin ser primo. El diccionario fino es de tres pisos: conjuntos algebraicos $\leftrightarrow$ ideales radicales, conjuntos irreducibles $\leftrightarrow$ ideales primos, puntos $\leftrightarrow$ ideales maximales. Cuando un enunciado exija irreducibilidad conviene comprobarla; el teorema de Bézout, por ejemplo, exige que las curvas no compartan componentes.

Funciones racionales, anillo local y orden en un punto

Definición 35.17 (Cuerpo de funciones, anillo local y orden)**.** Sea X una variedad afín irreducible y $K[X] = K[x_1, \ldots, x_n]/I(X)$ su anillo de coordenadas —el anillo de las funciones polinómicas sobre X, que se estudia con detalle más adelante en este mismo capítulo—. Como $I(X)$ es primo, $K[X]$ es un dominio, y se define su *cuerpo de funciones* $K(X)$ como su cuerpo de fracciones: los cocientes f/g con $g \neq 0$ en $K[X]$.

Para un punto $P \in X$, el *anillo local* es

$$\mathcal{O}_P = \{\, f/g \in K(X) \;:\; g(P) \neq 0 \,\},$$

las funciones racionales *definidas en* P. Es un anillo local: sus elementos no invertibles son los f/g con $f(P) = 0$, y forman su único ideal maximal.

Si X es una curva y P un punto liso, ese ideal maximal es principal, generado por cualquier *parámetro local* t; entonces toda $h \in K(X)^\times$ se escribe de manera única como $h = t^m u$

con $u(P) \neq 0$, y ese entero $m \in \mathbb{Z}$ es el *orden* $\operatorname{ord}_P(h)$: positivo si h tiene un cero de esa multiplicidad, negativo si tiene un polo.

Ejemplo 35.18 (Órdenes en la recta, contados uno a uno)**.** Sea $X = \mathbb{A}^1$, con $K[X] = K[x]$ y $K(X) = K(x)$, y sea

$$h(x) = \frac{(x-1)^2}{(x+2)^3}.$$

En $P = 1$ el parámetro local es $t = x - 1$ y $h = t^2 \cdot (x+2)^{-3}$ con $(1+2)^{-3} = 1/27 \neq 0$, luego $\operatorname{ord}_1(h) = 2$. En $P = -2$, con $t = x + 2$, sale $h = t^{-3}(x-1)^2$ y $\operatorname{ord}_{-2}(h) = -3$: un polo triple. En cualquier otro punto el orden es 0, porque numerador y denominador son no nulos allí.

La suma de los órdenes es $2 - 3 = -1$, y no 0. No hay contradicción con la afirmación de que los divisores principales tienen grado 0: falta contar el punto del infinito. En la recta proyectiva, con el parámetro local $u = 1/x$ en el infinito, se tiene $h = u^3(1-u)^2/(1+2u)^3 \cdot u^{-2} = u\,(\cdots)$ con $(\cdots)$ no nulo en $u = 0$, es decir $\operatorname{ord}_\infty(h) = 1$, y ahora sí $2 - 3 + 1 = 0$. Los ceros y los polos de una función racional se compensan exactamente cuando se cuenta en el espacio proyectivo, y ese es el contenido de que $\deg \operatorname{div}(h) = 0$.

Curvas planas: grado, multiplicidad de un punto y cono tangente

Definición 35.19 (Curva plana, grado y multiplicidad de un punto)**.** Una *curva plana afín* sobre K es el conjunto de ceros $V(F)$ de un polinomio no constante $F \in K[x, y]$ sin factores repetidos; su *grado* es el grado total de F. Las componentes irreducibles de la curva se corresponden con los factores irreducibles de F.

Sea $P \in V(F)$; trasladando se supone $P = (0, 0)$ y se escribe F como suma de sus partes homogéneas,

$$F = F_m + F_{m+1} + \cdots + F_d, \qquad F_m \neq 0,$$

donde F_k agrupa los monomios de grado exactamente k. El entero m se llama *multiplicidad* de P en la curva y se escribe $m_P(F)$; la forma F_m es el *cono tangente*, y sus factores lineales son las *rectas tangentes* a la curva en P. El punto es *liso* si $m_P(F) = 1$ —y entonces hay una única tangente— y *singular* si $m_P(F) \geq 2$. Que $m_P \geq 2$ equivale a que se anulen F y sus dos derivadas parciales en P, que es el criterio del gradiente que usa la sección siguiente.

Ejemplo 35.20 (Cuatro curvas, cuatro multiplicidades)**. (a) La parábola $F = y - x^2$ en el origen.** La parte de grado más bajo es $F_1 = y$, luego $m_P = 1$: punto liso, con tangente única $y = 0$. Coincide con el gradiente: $\nabla F = (-2x, 1)$, que en el origen vale $(0, 1) \neq \mathbf{0}$.

(b) El nodo $F = y^2 - x^2 - x^3$. La parte de grado más bajo es $F_2 = y^2 - x^2$, luego $m_P = 2$ y el cono tangente factoriza como $(y - x)(y + x)$: *dos* tangentes distintas. La curva se cruza consigo misma. Grado de la curva: 3.

(c) La cúspide $F = y^2 - x^3$. Ahora $F_2 = y^2$, otra vez $m_P = 2$, pero el cono tangente es un cuadrado: una *única* tangente $y = 0$ contada dos veces. Esa es exactamente la diferencia algebraica entre nodo y cúspide, y explica por qué las multiplicidades de intersección con la tangente son distintas —3 para la cúspide, 2 para una recta genérica—.

(d) El folium de Descartes $F = x^3+y^3-3xy$. La parte de grado más bajo es $F_2 = -3xy$, luego $m_P = 2$ con tangentes $x = 0$ e $y = 0$: un nodo con las dos tangentes perpendiculares. El grado de la curva es 3 y el origen es su único punto singular.

Órdenes monomiales y el algoritmo de división

Definición 35.21 (Orden monomial)**.** Escríbase $\mathbf{x}^\alpha = x_1^{\alpha_1}\cdots x_n^{\alpha_n}$ con $\alpha \in \mathbb{N}^n$. Un *orden monomial* es un orden total $\succ$ sobre los monomios tal que

1. si $\alpha \succ \beta$ entonces $\alpha+\gamma \succ \beta+\gamma$ para todo γ —compatibilidad con el producto—, y
2. $\succ$ es un buen orden: todo conjunto no vacío de monomios tiene mínimo. Equivale a que $\mathbf{x}^\alpha \succeq 1$ para todo α, y es lo que garantiza que los algoritmos terminen.

Los tres órdenes usuales son, para $x_1 \succ \cdots \succ x_n$:

- *lexicográfico* (lex): $\alpha \succ \beta$ si la primera coordenada no nula de $\alpha-\beta$ es positiva. Manda la primera variable, cueste lo que cueste el grado.
- *lexicográfico graduado* (grlex): primero se compara el grado total $|\alpha| = \alpha_1+\cdots+\alpha_n$, y en caso de empate se desempata por lex.
- *lexicográfico graduado inverso* (grevlex): primero el grado total, y en caso de empate gana el que tenga *menor* exponente en la *última* variable.

Fijado $\succ$, el mayor monomio que aparece en f con coeficiente no nulo da el *monomio principal* $\mathrm{lm}(f)$, su coeficiente es $\mathrm{lc}(f)$ y el producto de ambos es el *término principal* $\mathrm{lt}(f)$.

Ejemplo 35.22 (Los tres órdenes discrepan, con números)**.** Compárense $u = x\,y^5z^2$ y $v = x^4y\,z^3$ en $K[x,y,z]$ con $x \succ y \succ z$. Los dos tienen grado total 8.

- *lex*: se mira el exponente de x, que es 1 frente a 4; gana v.
- *grlex*: empatan en grado, y el desempate lex vuelve a dar v.
- *grevlex*: empatan en grado, y se mira el exponente de la *última* variable z: 2 frente a 3; gana el *menor*, luego gana u.

Así que grlex y grevlex no son el mismo orden, aunque ambos empiecen comparando grados; la primera discrepancia entre ellos aparece justo con tres variables. Y con $f = 4xy^5z^2 - 7x^4yz^3$ se tiene $\mathrm{lt}(f) = -7x^4yz^3$ en lex y en grlex, pero $\mathrm{lt}(f) = 4xy^5z^2$ en grevlex.

Regla: Dividir un polinomio entre una lista de polinomios

Fijado un orden monomial y una lista ordenada de divisores $(g_1,\dots,g_s)$, para dividir f:

1. **Inicializa** $r = 0$, cocientes $q_1 = \cdots = q_s = 0$ y un resto de trabajo $h = f$.
2. **Mira el término principal de h.** Recorre la lista de divisores *en orden* y busca el primer g_i cuyo $\mathrm{lt}(g_i)$ *divida* a $\mathrm{lt}(h)$.

3. **Si lo encuentras**, pon $c = \mathrm{lt}(h)/\mathrm{lt}(g_i)$, suma c a q_i y sustituye h por $h - c\,g_i$. El término principal se cancela y h baja.
4. **Si no lo encuentra ninguno**, mueve $\mathrm{lt}(h)$ al resto: súmalo a r y quítalo de h. Ese término ya no se podrá tocar nunca más.
5. **Repite** desde el paso 2 hasta que $h = 0$. Al terminar, $f = q_1 g_1 + \cdots + q_s g_s + r$, y ningún monomio de r es divisible por ningún $\mathrm{lt}(g_i)$.

El proceso acaba siempre, porque el término principal de h decrece estrictamente en cada vuelta y el orden es un buen orden.

Ejemplo 35.23 (Una división completa, y el resto que cambia de orden)**.** Divídase $f = x^2y + xy^2 + y^2$ entre $g_1 = xy - 1$ y $g_2 = y^2 - 1$ en $K[x, y]$ con lex y $x \succ y$.

(a) Con la lista (g_1, g_2)**.** El término principal de f es x^2y, divisible por $\mathrm{lt}(g_1) = xy$ con cociente x: resta $x\,g_1$ y queda $xy^2 + x + y^2$. Ahora el principal es xy^2, otra vez divisible por xy con cociente y: resta $y\,g_1$ y queda $x + y + y^2$. El principal es ahora x, que no lo divide ni xy ni y^2: pasa al resto. Queda $y^2 + y$; el principal y^2 sí lo divide $\mathrm{lt}(g_2)$, con cociente 1: resta g_2 y queda $y + 1$, que ya no es divisible por nada y pasa entero al resto. Resultado:

$$f = (x + y)\,g_1 + 1 \cdot g_2 + (x + y + 1).$$

Comprobación: $(x+y)(xy-1) = x^2y+xy^2-x-y$, más y^2-1, más $x+y+1$, da $x^2y+xy^2+y^2$. ✓

(b) Con la lista (g_2, g_1)**.** Repitiendo el proceso con los divisores intercambiados sale

$$f = (x + 1)\,g_2 + x\,g_1 + (2x + 1),$$

y en efecto $(x+1)(y^2-1)+x(xy-1)+2x+1 = xy^2+y^2-x-1+x^2y-x+2x+1 = x^2y+xy^2+y^2$. ✓

Los dos restos, $x + y + 1$ y $2x + 1$, son distintos. El algoritmo es determinista, pero su salida depende del orden en que se listen los divisores: eso es precisamente lo que en una variable no pasa nunca.

Observación 35.24 (Y por eso hacen falta las bases de Gröbner)*.* El defecto es más grave de lo que parece: como los dos restos son no nulos, la división no permite decidir si f pertenece o no al ideal (g_1, g_2). La base de Gröbner reducida de ese ideal para lex con $x \succ y$ resulta ser

$$G = \{\, x - y,\ \ y^2 - 1 \,\},$$

y al dividir f entre G —en cualquiera de los dos órdenes— el resto es el mismo, $2y + 1$. Como no es cero, $f \notin (g_1, g_2)$, y ahora sí es una conclusión y no un accidente del orden elegido. Devolver la unicidad del resto es *la* propiedad que define una base de Gröbner, y la sección correspondiente de este capítulo la enuncia y explica cómo calcularla con el algoritmo de Buchberger.

Cuidado

[Un resto no nulo no demuestra nada] Con divisores cualesquiera, resto cero implica pertenencia al ideal, pero resto no nulo *no* implica no pertenencia. Ejemplo mínimo: divídase $f = xy$ entre $g_1 = xy - y$ y $g_2 = y$ con lex; con la lista (g_1, g_2) sale $f = 1 \cdot g_1 + 1 \cdot g_2 + 0$, resto cero; pero si se prueba con divisores mal elegidos es fácil obtener restos no nulos de polinomios que sí están en el ideal. Sólo tras sustituir los generadores por una base de Gröbner el resto se convierte en una *forma normal*, y entonces «resto cero» es un criterio de pertenencia con las dos implicaciones.

35.3. Curvas planas: singularidades, género y Bézout

Definición 35.25. Un punto $\mathbf{p}$ de la curva $V(f) \subset \mathbb{C}^2$ es *singular* si $\nabla f(\mathbf{p}) = \mathbf{0}$; en caso contrario es *liso* y la curva tiene tangente bien definida. Para una curva plana *lisa* de grado d en el plano proyectivo, su *género* es

$$g = \frac{(d-1)(d-2)}{2},$$

que coincide con el número de asas de la superficie de Riemann asociada. El *teorema de Bézout* afirma que dos curvas proyectivas de grados d y e sin componentes comunes se cortan en exactamente de puntos, contados con multiplicidad.

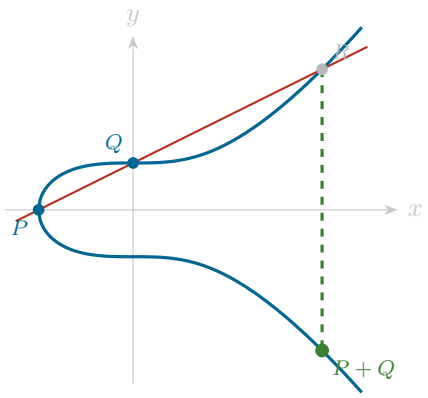


Figura 35.1: La ley de grupo sobre $y^2 = x^3 + 1$. La recta que une $P = (-1, 0)$ con $Q = (0, 1)$ corta la curva en un tercer punto $R = (2, 3)$ —existe por Bézout— y se define $P + Q$ como su reflejado $(2, -3)$. El neutro es el punto del infinito.

Idea clave

Qué dice esto en cristiano. Una curva compleja de dimensión uno es, vista como objeto real, una superficie: por eso tiene género, igual que un toro o una esfera. Y ese género queda determinado por el grado mediante una fórmula, lo que enlaza el álgebra (el grado del polinomio) con la topología (el número de asas). Bézout es la versión limpia de "dos curvas de grados d y e se cortan de veces", y para que sea exacta hacen falta

tres correcciones: trabajar sobre $\mathbb{C}$, trabajar en el plano proyectivo para no perder las intersecciones en el infinito, y contar con multiplicidad las tangencias.

Ejemplo 35.26 (Género, singularidades y cuentas de intersección)**. (a) El género según el grado.** Para $d = 1$ y $d = 2$ sale $g = 0$: rectas y cónicas son esferas, y por eso admiten parametrización racional. Para $d = 3$ sale $g = 1$: un toro, y de ahí el nombre de curva elíptica. Para $d = 4$ sale $g = 3$ y para $d = 5$, $g = 6$.

(b) La parametrización de género cero, con consecuencias. La circunferencia $x^2+y^2 = 1$ se parametriza como $\left(\frac{1-t^2}{1+t^2}, \frac{2t}{1+t^2}\right)$; para $t = 2$ da $\left(-\frac{3}{5}, \frac{4}{5}\right)$, cuya suma de cuadrados es 1. Tomando $t = \frac{p}{q}$ racional se obtienen todas las ternas pitagóricas: $(q^2 - p^2,\ 2pq,\ q^2 + p^2)$, y con $p = 1$, $q = 2$ sale $(3, 4, 5)$. El género cero es exactamente la razón de que esta receta exista.

(c) Dos singularidades clásicas. La curva $y^2 = x^3$ tiene gradiente $(3x^2, -2y)$, que se anula en el origen: es una *cúspide*, la curva llega a un punto y da la vuelta. La curva $y^2 = x^2(x+1)$ tiene un *nodo* en el origen, donde se cruza consigo misma. En cambio $y^2 = x^3 - x$ es lisa, y sus únicos puntos racionales son $(-1, 0)$, $(0, 0)$ y $(1, 0)$.

(d) Bézout con las tres correcciones. Una recta y una cónica se cortan en $1 \cdot 2 = 2$ puntos; dos cónicas en 4. Sobre $\mathbb{R}$ y en el plano afín eso falla constantemente: dos circunferencias disjuntas se cortan en cero puntos reales, y dos paralelas en ninguno. Sobre $\mathbb{C}$ y proyectivamente los puntos aparecen: las paralelas se cortan en el infinito, y las circunferencias disjuntas en dos puntos complejos.

Cuidado

[El género es de la curva lisa, y las curvas singulares mienten] La fórmula $g = \frac{(d-1)(d-2)}{2}$ exige lisura. La cúspide $y^2 = x^3$ tiene grado 3 y su género *geométrico* es 0, no 1: cada singularidad rebaja el género. Para curvas con nodos y cúspides hay que restar una contribución por cada una. Segundo aviso: que una curva de género 0 admita parametrización racional y una de género ≥ 1 no la admita es un teorema profundo, y es la razón última de que la ecuación $x^n + y^n = 1$ con $n \geq 3$ (género ≥ 1) tenga tan pocos puntos racionales, mientras que con $n = 2$ tiene infinitos.

35.4. El plano proyectivo: por qué las cuentas salen bien en el infinito

La geometría algebraica afín está llena de excepciones molestas: dos rectas se cortan en un punto *salvo si son paralelas*, dos cónicas se cortan en cuatro puntos *salvo cuando no*. Todas esas excepciones desaparecen de golpe al añadir los puntos del infinito y al permitir coordenadas complejas. El precio es un espacio nuevo; la recompensa es que los teoremas dejan de tener letra pequeña.

Definición 35.27. El *plano proyectivo* $\mathbb{P}^2(K)$ es el conjunto de las ternas $(X : Y : Z)$ con

K un cuerpo, no todas nulas, identificando $(X : Y : Z) \sim (\lambda X : \lambda Y : \lambda Z)$ para todo $\lambda \neq 0$. Los puntos con $Z \neq 0$ se identifican con el plano afín mediante $(x, y) = (X/Z,\ Y/Z)$; los que tienen $Z = 0$ forman la *recta del infinito*. Un polinomio se *homogeneiza* sustituyendo cada monomio de grado k por el mismo multiplicado por Z^{d-k}, donde d es el grado total.

Idea clave

Qué dice esto en cristiano. Un punto proyectivo no es un punto: es una *dirección* vista desde el origen del espacio de tres dimensiones, es decir, una recta que pasa por el origen. Las que no son horizontales cortan al plano $Z = 1$ y dan los puntos de siempre; las horizontales no lo cortan y son los puntos nuevos, uno por cada dirección del plano. Dos rectas paralelas van en la misma dirección, así que se encuentran en ese punto nuevo, y con eso la excepción desaparece. Homogeneizar es simplemente escribir las ecuaciones en el lenguaje de esas direcciones.

Regla: Pasar al proyectivo y volver, en siete pasos

1. **Homogeneizar**: multiplicar cada monomio por la potencia de Z que haga falta para igualar el grado. De $y = x^2$ sale $YZ = X^2$.
2. **Deshomogeneizar** poniendo $Z = 1$ para recuperar la curva afín original.
3. **Buscar los puntos del infinito** poniendo $Z = 0$ y resolviendo. En $YZ = X^2$ sale $X = 0$, luego el único punto del infinito de la parábola es $(0 : 1 : 0)$.
4. **Comprobar la homogeneidad**: todos los monomios deben tener el mismo grado, o la ecuación no está bien definida sobre clases de equivalencia.
5. **Contar con multiplicidad**: un punto de tangencia cuenta dos veces, un punto de inflexión tres.
6. **Trabajar sobre un cuerpo algebraicamente cerrado**, típicamente $\mathbb{C}$: sin eso los teoremas de conteo fallan por falta de raíces.
7. **Traducir de vuelta** el resultado al plano afín real, señalando qué se ha quedado en el infinito y qué era complejo.

Ejemplo 35.28 (Las excepciones que desaparecen)**. (a) Rectas paralelas.** $y = x$ e $y = x+1$ no se cortan en el plano afín. Homogeneizadas son $Y = X$ e $Y = X + Z$; restando sale $Z = 0$, y con $Y = X$ eso da el punto $(1 : 1 : 0)$. Se cortan exactamente en un punto, como toda pareja de rectas distintas.

(b) Dos circunferencias sin puntos comunes. $x^2 + y^2 = 1$ y $x^2 + y^2 = 4$ no se cortan en $\mathbb{R}^2$. En $\mathbb{P}^2(\mathbb{C})$ se cortan en cuatro puntos contando multiplicidad, y los cuatro están en el infinito: son $(1 : \mathrm{i} : 0)$ y $(1 : -\mathrm{i} : 0)$, cada uno doble. Se llaman *puntos cíclicos* y por ellos pasa toda circunferencia del plano, lo que explica de un plumazo por qué dos circunferencias nunca se cortan en más de dos puntos visibles.

(c) Bézout sin excepciones. Dos curvas proyectivas planas sin componentes comunes, de grados m y n, se cortan en exactamente mn puntos de $\mathbb{P}^2(\mathbb{C})$ contando multiplicidades.

Recta con cónica: $1 \cdot 2 = 2$. Dos cónicas: 4. Cónica con cúbica: 6. El enunciado afín necesitaría tres párrafos de salvedades; el proyectivo cabe en una línea.

(d) Las cónicas se reducen a tres. En el plano afín real hay elipses, parábolas e hipérbolas, tres familias con propiedades muy distintas. En $\mathbb{P}^2(\mathbb{C})$ toda cónica no degenerada es equivalente a $XY = Z^2$: hay *una sola*. La diferencia entre ellas es únicamente cómo cortan a la recta del infinito: la elipse no la toca, la parábola es tangente y la hipérbola la cruza en dos puntos.

Definición 35.29. Una *curva elíptica* sobre un cuerpo K de característica distinta de 2 y 3 es una curva proyectiva lisa de ecuación $Y^2Z = X^3 + aXZ^2 + bZ^3$ con discriminante $-16(4a^3 + 27b^2) \neq 0$. Tiene un único punto en el infinito, $O = (0 : 1 : 0)$, y sus puntos forman un *grupo abeliano* con la ley de la cuerda y la tangente, tomando O como neutro.

Ejemplo 35.30 (Contar puntos sobre cuerpos finitos)**.** Sea $E : y^2 = x^3 - x$, cuyo discriminante es $-16 \cdot 4 \cdot (-1)^3 = 64 \neq 0$.

(a) Los conteos. Incluyendo el punto del infinito, $|E(\mathbb{F}_5)| = 8$, $|E(\mathbb{F}_7)| = 8$, $|E(\mathbb{F}_{11})| = 12$ y $|E(\mathbb{F}_{13})| = 8$.

(b) La cota de Hasse. El teorema afirma que $|p + 1 - |E(\mathbb{F}_p)|| \leq 2\sqrt{p}$. Aquí las trazas $a_p = p + 1 - |E(\mathbb{F}_p)|$ valen -2, 0, 0 y 6, frente a cotas 4,472; 5,292; 6,633 y 7,211. Se cumplen holgadamente. Es el análogo de la hipótesis de Riemann para curvas sobre cuerpos finitos, y ése sí está demostrado, por Hasse en 1933 y por Weil en general.

(c) La ley de grupo. Para sumar P y Q se traza la recta que los une, se toma el tercer punto de corte con la curva —que existe por Bézout, porque una recta y una cúbica se cortan en tres puntos— y se refleja respecto del eje x. La asociatividad de esa operación no es nada evidente y se demuestra, precisamente, con geometría proyectiva.

(d) Para qué sirve. La criptografía de curva elíptica usa el problema del logaritmo discreto en ese grupo, que es mucho más duro que en $\mathbb{F}_p^*$: una clave de 256 bits sobre una curva ofrece la seguridad que RSA necesita 3072 bits para igualar. Y la demostración del último teorema de Fermat, en 1995, consistió en probar que cierta curva elíptica asociada a una hipotética solución no podía existir.

Cuidado

[Contraejemplo: sobre $\mathbb{R}$ el conteo falla] Bézout exige un cuerpo algebraicamente cerrado. Sobre $\mathbb{R}$, las cúbicas $y = x^3$ y $y = 0$ se cortan en un solo punto visible, no en tres, porque $x^3 = 0$ tiene una raíz triple; y $x^2 + y^2 + 1 = 0$ no tiene ningún punto real aunque sea una cónica perfectamente legítima sobre $\mathbb{C}$. Trabajar sobre los reales es más intuitivo y menos regular; la geometría algebraica prefiere $\mathbb{C}$ por la misma razón por la que el álgebra prefiere los complejos.

Pausa para pensar

El plano proyectivo no lo inventaron los algebristas sino los pintores. Los tratados de perspectiva del Quattrocento ya manejaban el punto de fuga, que es exactamente el punto del infinito de una dirección, y Desargues escribió en 1639 el primer tratado geométrico que lo trataba como un punto más. El libro fue tan mal recibido que se perdió, y sólo se recuperó dos siglos después, cuando Poncelet había reinventado la disciplina desde una prisión rusa en 1813, sin libros y de memoria.

35.5. Del anillo al espacio: el espectro y los esquemas

El diccionario entre ideales y variedades funciona sobre un cuerpo algebraicamente cerrado y se rompe en cuanto uno se sale de ahí. Grothendieck propuso arreglarlo cambiando la definición de espacio: en lugar de un conjunto de puntos con funciones encima, un anillo del que los puntos se deducen. El resultado es que la geometría se puede hacer sobre $\mathbb{Z}$.

Definición 35.31. El *espectro* $\operatorname{Spec} A$ de un anillo conmutativo A es el conjunto de sus ideales primos, con la *topología de Zariski*, en la que los cerrados son los $V(I) = \{\mathfrak{p} : I \subseteq \mathfrak{p}\}$. Un punto se llama *cerrado* si su clausura es él mismo, y *genérico* si su clausura es todo el espacio. Un *esquema afín* es un espectro con su haz de funciones; un *esquema* es un espacio que localmente lo es.

Idea clave

Qué dice esto en cristiano. La idea de partida es del capítulo anterior: un punto de una variedad se corresponde con un ideal maximal. Grothendieck observó que si además se admiten los primos no maximales, aparecen puntos nuevos que no son puntos en el sentido clásico sino *subvariedades irreducibles vistas como un único punto grueso*. El ideal (0) en un dominio es el punto genérico: vive en todas partes a la vez y su clausura es el espacio entero. Suena a capricho y resuelve un problema real: permite hablar de "lo que ocurre en el punto general" sin tener que decir en cuál, que es exactamente lo que uno quiere cuando dice "una curva genérica no tiene singularidades".

Ejemplo 35.32 (Espectros que se pueden dibujar)**. (a)** $\operatorname{Spec}\mathbb{Z}$**.** Un punto cerrado por cada primo —(2), (3), (5), …— más el punto genérico (0). Es un espacio de dimensión 1: una curva cuyos puntos son los primos. La topología es rarísima: los abiertos no vacíos son los complementarios de conjuntos finitos de primos, así que dos abiertos no vacíos cualesquiera se cortan. No es Hausdorff, ni de lejos.

(b) $\operatorname{Spec} k[x]$ **con** k **algebraicamente cerrado.** Un punto cerrado $(x-a)$ por cada $a \in k$, más el genérico (0). Es la recta afín de toda la vida, con un punto extra. La analogía con (a) es literal y es la base de la aritmética geométrica: $\mathbb{Z}$ y $k[x]$ son ambos dominios de ideales principales, y sus espectros son ambos curvas.

(c) $\operatorname{Spec}\mathbb{Z}[x]$**.** Tiene dimensión 2: es una superficie fibrada sobre $\operatorname{Spec}\mathbb{Z}$, cuya fibra sobre el primo p es $\operatorname{Spec}\mathbb{F}_p[x]$. Reducir una ecuación módulo p es, en este lenguaje, restringirla a

una fibra, y estudiarla para todos los p a la vez es estudiar la superficie entera.

(d) Nilpotentes y multiplicidad. $\operatorname{Spec} k[x]/(x^2)$ tiene un solo punto, igual que $\operatorname{Spec} k$. Pero sus anillos son distintos: el primero tiene un elemento nilpotente no nulo. Los esquemas distinguen "un punto" de "un punto doble", y eso permite que las intersecciones tengan multiplicidad de forma automática, sin definiciones *ad hoc*. Es la razón técnica por la que Bézout funciona limpiamente en este lenguaje.

Cuidado

[Contraejemplo: la intuición topológica no sobrevive] La topología de Zariski es tan grosera que casi todo lo aprendido en el capítulo de topología deja de servir. En $\operatorname{Spec}\mathbb{Z}$ no hay dos abiertos no vacíos disjuntos, luego el espacio no es Hausdorff; los puntos genéricos no son cerrados, luego no es ni siquiera T_1; y sin embargo es cuasicompacto. Quien intente visualizar un esquema como una figura geométrica se equivocará; el objeto que hay que tener en la cabeza es el anillo, y el espacio es un modo de organizarlo.

Pausa para pensar

Grothendieck reconstruyó la geometría algebraica entera entre 1958 y 1970, y los *Éléments de géométrie algébrique* que escribió con Dieudonné suman unas mil ochocientas páginas de un plan que quedó inacabado: estaban previstos trece capítulos y se publicaron cuatro. La comunidad tardó dos décadas en digerirlo. En 1970 Grothendieck abandonó las matemáticas institucionales por motivos políticos, a los cuarenta y dos años. Dejó escrita una frase que resume su método: prefería ahogar el problema en un mar de generalidad hasta que se disolviera solo, en vez de atacarlo con un martillo.

35.6. Anillo de coordenadas, morfismos y el diccionario completo

El diccionario que abre el capítulo traduce *objetos*: ideales por un lado, variedades por otro. Falta la mitad que lo convierte en un diccionario de verdad, que es la de las *aplicaciones*.

Definición 35.33 (Anillo de coordenadas)**.** Para una variedad afín $X \subset \mathbb{A}^n$,

$$K[X] = K[x_1, \dots, x_n]/I(X),$$

el anillo de las funciones polinómicas sobre X, donde dos polinomios se identifican cuando coinciden en todos los puntos de X.

Definición 35.34 (Morfismo)**.** Una aplicación $\varphi : X \to Y \subset \mathbb{A}^m$ es un **morfismo** (o aplicación regular) si sus m coordenadas son funciones de $K[X]$.

Teorema 35.35 (La antiequivalencia)**.** *Todo morfismo $\varphi : X \to Y$ induce un homomorfismo de K-álgebras $\varphi^{\#} : K[Y] \to K[X]$ por composición, $f \mapsto f \circ \varphi$; y la correspondencia*

$$X \longleftrightarrow K[X]$$

es una antiequivalencia entre la categoría de variedades afines sobre K algebraicamente cerrado y la de K-álgebras reducidas finitamente generadas.

Las flechas se invierten porque componer con φ tira de las funciones hacia atrás. Esa inversión es la misma que la de **FinBool** y **FinSet**, y la misma que la de Gelfand: geometría y álgebra son categorías opuestas.

Ejemplo 35.36 (La cúspide: biyectiva, regular, y no isomorfismo)**.** Sea $C = V(y^2 - x^3) \subset \mathbb{A}^2$ y $\varphi : \mathbb{A}^1 \to C$, $t \mapsto (t^2, t^3)$.

(a) Es morfismo. Sus coordenadas son polinomios. Y su imagen está en C: $(t^3)^2 = (t^2)^3$.

(b) Es biyectiva. Sobreyectiva: dado $(a, b) \in C$ con $a \neq 0$, tómese $t = b/a$, y entonces $t^2 = b^2/a^2 = a^3/a^2 = a$ y $t^3 = ta = b$. El origen viene de $t = 0$. Inyectiva: $t = y/x$ recupera t.

(c) Y sin embargo no es isomorfismo. A nivel de anillos, $\varphi^{\#} : K[C] \to K[t]$ tiene imagen $K[t^2, t^3]$, que es un subanillo *propio* de $K[t]$: le falta t. Si φ fuera isomorfismo, $\varphi^{\#}$ lo sería. La inversa $t = y/x$ no es un polinomio en C; es solo una función *racional*, indefinida en el origen.

(d) Qué significa. Una biyección regular no basta para identificar variedades, igual que una biyección continua no basta para identificar espacios topológicos. Lo que falla está localizado en un punto: la cúspide tiene ahí una singularidad, y $\mathbb{A}^1$ no. El álgebra lo detecta sin necesidad de dibujar nada, comparando dos anillos.

Definición 35.37 (Irreducibilidad, componentes y dimensión)**.** X es **irreducible** si no es unión de dos cerrados propios; equivale a que $I(X)$ sea primo, y a que $K[X]$ sea dominio. Toda variedad es unión finita y única de componentes irreducibles (por noetherianidad). La **dimensión** de X es la longitud máxima de una cadena de subvariedades irreducibles $X_0 \subsetneq \cdots \subsetneq X_d$, y coincide con el grado de trascendencia de $K(X)$ sobre K y con la dimensión de Krull de $K[X]$.

Ejemplo 35.38 (Dos descomposiciones)**. (a)** $V(xy) \subset \mathbb{A}^2$**.** El ideal (xy) no es primo ($x \cdot y \in (xy)$ y ninguno de los dos está). La descomposición es $V(xy) = V(x) \cup V(y)$: los dos ejes. Dos componentes, ambas de dimensión 1.

(b) $V(xz, yz) \subset \mathbb{A}^3$**.** Un punto está en la variedad si $z = 0$ o si $x = y = 0$. Luego

$$V(xz, yz) = \underbrace{V(z)}_{\text{plano}} \cup \underbrace{V(x, y)}_{\text{eje } z},$$

de dimensiones 2 y 1. Las componentes irreducibles *no* tienen por qué tener la misma dimensión: la variedad no es equidimensional, y su dimensión —definida como el máximo— es 2, aunque cerca de un punto del eje z distinto del origen sea localmente de dimensión 1.

35.7. Bases de Gröbner: cómo se calcula con ideales

Todo lo anterior afirma correspondencias entre ideales y variedades sin dar ningún método para operar con ideales. Preguntas tan básicas como "¿pertenece f al ideal I?" o "¿cuál es la proyección de esta variedad?" necesitan un algoritmo, y ese algoritmo existe desde 1965.

Idea clave

En una variable, la división euclídea resuelve todo: $f \in (g)$ si y solo si el resto de dividir f entre g es cero. En varias variables la división es ambigua —depende del orden en que se elijan los divisores— y el resto puede ser no nulo aunque f sí pertenezca al ideal. Una base de Gröbner es un sistema de generadores tan bueno que devuelve la unicidad del resto.

Definición 35.39 (Orden monomial y base de Gröbner)**.** Fíjese un orden total sobre los monomios compatible con el producto y con buena ordenación (lexicográfico, graduado lexicográfico, graduado inverso). Para cada f, $\mathrm{LT}(f)$ es su término principal. Un conjunto $G = \{g_1, \dots, g_s\} \subset I$ es **base de Gröbner** de I si

$$(\mathrm{LT}(g_1), \dots, \mathrm{LT}(g_s)) = (\mathrm{LT}(f) : f \in I).$$

Entonces el resto de la división por G es único, y $f \in I$ si y solo si ese resto es cero.

El algoritmo de Buchberger construye una: se toman todos los pares (g_i, g_j), se forma el S-polinomio que cancela sus términos principales, se divide por G, y si el resto no es cero se añade a G. Termina porque el ideal de términos principales crece y el anillo es noetheriano.

Teorema 35.40 (Eliminación)**.** *Si G es base de Gröbner de I en orden* lexicográfico *con $x_1 > \cdots > x_n$, entonces $G \cap K[x_{k+1}, \dots, x_n]$ es base de Gröbner del ideal de eliminación $I \cap K[x_{k+1}, \dots, x_n]$.*

Eliminar variables es exactamente proyectar la variedad, así que el orden lexicográfico convierte una operación geométrica en una lectura de la lista de generadores.

Ejemplo 35.41 (Dos eliminaciones concretas)**. (a) Proyectar una intersección.** Sea $I = (x^2 + y^2 - 1,\ x - y)$: la circunferencia unidad cortada con la bisectriz. En orden lex con $x > y$, la base de Gröbner reducida es

$$G = \{\, x - y,\quad 2y^2 - 1 \,\}.$$

El segundo generador no contiene x: es el ideal de eliminación, y describe la proyección sobre el eje y, es decir $y = \pm 1/\sqrt{2}$. Los dos puntos de corte son $(\pm\frac{1}{\sqrt{2}}, \pm\frac{1}{\sqrt{2}})$, y el sistema ha quedado triangulado: primero se resuelve en y, luego $x = y$.

(b) Implicitar una parametrización. La curva $x = t^2$, $y = t^3$ se describe por el ideal $(x - t^2,\ y - t^3)$ en $K[t, x, y]$. Con orden lex y $t > x > y$, la base de Gröbner es

$$\{\, t^2 - x,\quad tx - y,\quad ty - x^2,\quad x^3 - y^2 \,\},$$

y el único generador sin t es

$$x^3 - y^2 = 0.$$

Es la ecuación implícita de la cúspide, obtenida por un algoritmo y no por ingenio. Ese paso —de la parametrización a la ecuación— es el que se automatiza en diseño asistido por ordenador, en robótica (cinemática inversa) y en demostración automática de teoremas de geometría.

Observación 35.42. El precio es el coste. El orden lexicográfico es el que elimina, y es también el más caro: en el peor caso el número de operaciones es doblemente exponencial en el número de variables. En la práctica se calcula primero una base en orden graduado inverso, que es mucho más barata, y se convierte al lexicográfico con el algoritmo FGLM. Con eso, sistemas de veinte o treinta variables son abordables, y la geometría algebraica computacional se convirtió en una herramienta de ingeniería y no solo en teoría.

35.8. Singularidades, género y el teorema de Riemann–Roch

El género apareció como un invariante decisivo y quedó como un número sin oficio. Riemann–Roch le da el oficio: contar funciones.

Definición 35.43 (Divisores)**.** Un **divisor** sobre una curva lisa C es una suma formal finita $D = \sum n_P\, P$ con $n_P \in \mathbb{Z}$; su **grado** es $\sum n_P$. Cada función racional f no nula define el divisor principal $\operatorname{div}(f) = \sum_P \operatorname{ord}_P(f)\, P$, de grado 0. Dos divisores son *linealmente equivalentes* si difieren en uno principal. Y

$$L(D) = \{f \text{ racional} : \operatorname{div}(f) + D \geq 0\} \cup \{0\}, \qquad \ell(D) = \dim_K L(D).$$

En cristiano: $L(D)$ es el espacio de las funciones a las que se les permiten polos, pero solo en los puntos de D y con el orden que D autoriza. Cuanto más generoso el divisor, más funciones caben.

Teorema 35.44 (Riemann–Roch)**.** *Para toda curva proyectiva lisa de género g y todo divisor D,*

$$\ell(D) - \ell(K - D) = \deg D - g + 1,$$

donde K es el divisor canónico. En particular $\deg K = 2g - 2$ y $\ell(K) = g$, y si $\deg D > 2g - 2$ el término correctivo desaparece y $\ell(D) = \deg D - g + 1$ exactamente.

Ejemplo 35.45 (Deducir la forma de Weierstrass en lugar de postularla)**.** Sea C una curva de género 1 con un punto marcado O. Como $\deg K = 0$ y $\ell(K) = 1$, la fórmula con $D = nO$ da $\ell(nO) = n$ para $n \geq 1$.

n	$\ell(nO)$	base
1	1	$\{1\}$
2	2	$\{1, x\}$
3	3	$\{1, x, y\}$
4	4	$\{1, x, y, x^2\}$
5	5	$\{1, x, y, x^2, xy\}$
6	6	$\{1, x, y, x^2, xy, x^3\}$ y también y^2

Léase la primera fila: $\ell(O) = 1$, luego la única función con a lo sumo un polo simple en O es la constante. Eso ya es un teorema: *no hay* función con un solo polo simple, porque si la hubiera daría un isomorfismo con la recta proyectiva y el género sería 0.

En $n=2$ aparece una función nueva, con polo doble: llámese x. En $n=3$, otra con polo triple: y. En $n=6$ el espacio tiene dimensión 6, y hemos listado *siete* funciones con polo de orden ≤ 6: $1, x, y, x^2, xy, x^3, y^2$. Siete vectores en un espacio de dimensión seis son linealmente dependientes, luego existe una relación

$$y^2 + a_1xy + a_3y = x^3 + a_2x^2 + a_4x + a_6,$$

que tras completar cuadrados y cubos —si la característica no es 2 ni 3— se reduce a

$$y^2 = x^3 + ax + b.$$

La ecuación de Weierstrass no es una elección afortunada: es *obligatoria*, y sale de contar dimensiones.

Teorema 35.46 (Género de una curva plana, con corrección por singularidades)**.** *Una curva plana lisa de grado d tiene género*

$$g = \frac{(d-1)(d-2)}{2},$$

y si tiene singularidades ordinarias con multiplicidades m_P, hay que restar $\delta_P = \binom{m_P}{2}$ por cada una:

$$g = \frac{(d-1)(d-2)}{2} - \sum_P \delta_P.$$

Ejemplo 35.47 (Dos cúbicas, dos destinos)**. (a) $y^2 = x^2(x+1)$, cúbica nodal.** $d=3$ da $\frac{2\cdot 1}{2} = 1$, y el nodo en el origen ($m=2$) resta $\delta = 1$: el género es 0. Una curva de género 0 es racional, es decir, admite parametrización. Y en efecto

$$x = t^2 - 1, \qquad y = t(t^2-1)$$

la parametriza: al sustituir, $y^2 - x^2(x+1) = t^2(t^2-1)^2 - (t^2-1)^2t^2 = 0$. ✓

(b) $y^2 = x^3 - x$, lisa. El discriminante $-16(4a^3+27b^2)$ con $a=-1$, $b=0$ vale $64 \neq 0$: no hay singularidades, y el género es 1. Por tanto *no* admite ninguna parametrización racional. Ese es el motivo profundo de que las integrales elípticas no se puedan resolver en términos elementales: el género es una obstrucción, no una falta de ingenio.

Observación 35.48 (El género aparece tres veces con tres caras). Es el número de asas de la superficie de Riemann asociada (topología), la dimensión del espacio de formas diferenciales holomorfas (análisis), y el término correctivo de Riemann–Roch (álgebra). Que las tres coincidan es el contenido del teorema de Riemann–Roch en su forma completa, y la primera aparición del patrón que Grothendieck generalizaría hasta el teorema de Riemann–Roch–Hirzebruch y el teorema del índice de Atiyah–Singer, donde un invariante analítico y otro topológico resultan ser el mismo número.

35.9. Haces de funciones y el espectro de un anillo

Un esquema afín se definió como "un espectro con su haz de funciones". Como la palabra haz no se había explicado, esa definición estaba vacía. Se llena aquí.

Definición 35.49 (Prehaz y haz)**.** Un **prehaz** $\mathcal{F}$ sobre un espacio X asigna a cada abierto U un conjunto (o grupo, o anillo) $\mathcal{F}(U)$, y a cada inclusión $V \subseteq U$ una restricción $\mathcal{F}(U) \to \mathcal{F}(V)$, compatible con la composición. Es un **haz** si además, para todo recubrimiento $U = \bigcup U_i$:

- *Localidad*: si dos secciones de $\mathcal{F}(U)$ coinciden en cada U_i, son iguales.
- *Pegado*: si hay secciones $s_i \in \mathcal{F}(U_i)$ que coinciden en las intersecciones, existe una $s \in \mathcal{F}(U)$ que las restringe.

El **tallo** en un punto es $\mathcal{F}_x = \varinjlim_{U \ni x} \mathcal{F}(U)$: el comportamiento de las secciones "infinitamente cerca" de x.

Observación 35.50. Las dos condiciones dicen que ser una sección es una propiedad *local*: se comprueba en un entorno de cada punto y se pega. Las funciones continuas, las derivables, las holomorfas y los campos vectoriales forman haces; las funciones *acotadas* no, porque el pegado falla (cada trozo puede estar acotado y el todo no). Ese contraejemplo es el que hace ver qué se está pidiendo exactamente.

Definición 35.51 (Espectro y su haz estructural)**.** Para un anillo conmutativo A, $\operatorname{Spec} A$ es el conjunto de sus ideales *primos*, con la topología de Zariski: los cerrados son los $V(I) = \{\mathfrak{p} \supseteq I\}$. Los abiertos básicos son los $D(f) = \{\mathfrak{p} : f \notin \mathfrak{p}\}$, y el haz estructural se define sobre ellos por

$$\mathcal{O}(D(f)) = A_f$$

—localizar es invertir f—, con tallos $\mathcal{O}_{\mathfrak{p}} = A_{\mathfrak{p}}$.

Ejemplo 35.52 ($\operatorname{Spec}\mathbb{Z}$, entero)**.** Sus puntos son los ideales primos de $\mathbb{Z}$: uno por cada primo, $(2), (3), (5), \ldots$, más el ideal (0).

(a) La topología. Los cerrados propios son los conjuntos finitos de primos, porque $V((n))$ es el conjunto de los divisores primos de n. El punto (0) está en *todo* abierto no vacío: es un **punto genérico**, denso en todo el espacio. Un espacio así no es de Hausdorff, y esa es la primera cosa que hay que aceptar de la topología de Zariski.

(b) Las secciones. $\mathcal{O}(D(2)) = \mathbb{Z}[\frac{1}{2}]$: sobre el abierto que quita el punto (2), las funciones son las fracciones con denominador potencia de 2.

(c) Los tallos. $\mathcal{O}_{(3)} = \mathbb{Z}_{(3)}$, las fracciones con denominador no divisible por 3: un anillo local cuyo ideal maximal es $3\mathbb{Z}_{(3)}$. Y en el punto genérico, $\mathcal{O}_{(0)} = \mathbb{Q}$.

(d) Qué es una "función" aquí. El entero $n = 12$, visto como sección global, "vale" en cada punto $\mathfrak{p} = (p)$ su clase en el cuerpo residual $\mathbb{Z}/p$. Así 12 vale 0 en (2) y en (3), vale 2 en (5), vale 5 en (7). Los ceros de esa función son exactamente $\{(2), (3)\} = V((12))$: la factorización de 12 es su conjunto de ceros. La aritmética se ha convertido en geometría, literalmente.

Idea clave

Ahí está la revolución de Grothendieck en una frase. Hilbert había identificado los puntos de una variedad afín sobre un cuerpo algebraicamente cerrado con los ideales *maximales* de su anillo de coordenadas. Grothendieck añadió los primos no maximales

—que aportan los puntos genéricos, uno por cada subvariedad irreducible— y dejó de exigir que el anillo fuera reducido, ni finitamente generado, ni sobre un cuerpo. El resultado, $\operatorname{Spec} A$ con su haz, tiene sentido para *cualquier* anillo conmutativo: para $\mathbb{Z}$, para $\mathbb{Z}/4$ (que tiene un solo punto y elementos nilpotentes, es decir, un punto "con grosor infinitesimal"), para anillos de series formales.
Y el precio de esa generalidad se paga con el haz. Sin él, $\operatorname{Spec} A$ es un espacio topológico feo y pobre; con él, es un objeto que recuerda *todo* el anillo, porque $A = \mathcal{O}(\operatorname{Spec} A)$. Un esquema es lo que se obtiene pegando espectros como se pegan cartas en una variedad diferenciable, y esa es la definición que hoy sostiene la geometría aritmética entera, desde Weil hasta Wiles.

35.10. Multiplicidad de intersección, divisor canónico y el algoritmo de Buchberger

Tres piezas del capítulo se usan sin estar definidas: el número con el que Bézout cuenta las tangencias, el divisor K que aparece en Riemann–Roch, y el algoritmo que produce las bases de Gröbner cuyos resultados se exhiben ya calculados.

Definición 35.53 (Multiplicidad de intersección)**.** Sean F, G dos curvas planas y P un punto. La *multiplicidad de intersección* de F y G en P es

$$I_P(F,G) = \dim_K \ \mathcal{O}_P/(F,G),$$

la dimensión, como espacio vectorial sobre K, del cociente del anillo local $\mathcal{O}_P$ —las funciones racionales definidas en P— por el ideal que generan las dos ecuaciones.

Regla: Calcular $I_P(F,G)$ sin álgebra conmutativa

La definición anterior es la correcta, pero rara vez es la que se usa. Estas cinco propiedades la determinan por completo y permiten calcularla como un algoritmo euclídeo:

1. $I_P(F,G) = \infty$ si F y G comparten una componente por P; en otro caso es un entero ≥ 0.
2. $I_P(F,G) = 0$ si y sólo si P no está en ambas curvas.
3. $I_P(F,G) = I_P(G,F)$, y es invariante por cambios de coordenadas.
4. $I_P(F,GH) = I_P(F,G) + I_P(F,H)$: es aditiva en cada argumento respecto del producto.
5. $I_P(F,G) = I_P(F,\, G + AF)$ para cualquier polinomio A. Es el paso que hace de "división".

Con (4) y (5) se va bajando el grado alternativamente, igual que en el algoritmo de Euclides, hasta llegar a un caso trivial por (2). Y una cota rápida: siempre $I_P(F,G) \geq m_P(F) \cdot m_P(G)$, el producto de las multiplicidades de P como punto de cada curva, con igualdad si y sólo si las curvas no tienen *tangentes comunes* en P.

Ejemplo 35.54 (Tangencia, nodo y cúspide, contados)**.** Sean la parábola $F = y - x^2$ y las rectas o curvas siguientes, todas en $P = (0,0)$.

- $G = y$ (la tangente): por (5), $I_P(y - x^2,\, y) = I_P(-x^2,\, y) = 2I_P(x, y) = 2$. La tangencia cuenta *dos*, y eso es lo que salva el enunciado de Bézout.
- $G = x$ (una recta transversal): $I_P(y - x^2, x) = I_P(y, x) = 1$. Corte simple.
- *Nodo*: $F = y^2 - x^2 - x^3$ tiene en el origen dos ramas con tangentes $y = \pm x$, luego $m_P(F) = 2$; cortada con una recta genérica G da $I_P = 2$, y con cada una de sus dos tangentes, 3.
- *Cúspide*: $F = y^2 - x^3$ tiene $m_P(F) = 2$ con la tangente doble $y = 0$. Entonces $I_P(F, y) = I_P(-x^3, y) = 3$, mientras que con una recta genérica da 2. La diferencia entre nodo y cúspide se lee en ese número.

En los cuatro casos la suma de las multiplicidades sobre todos los puntos de intersección, contados en el plano proyectivo y sobre un cuerpo algebraicamente cerrado, da exactamente $\deg F \cdot \deg G$: eso es Bézout, y sin la definición de I_P el enunciado era una frase sin contenido verificable.

Definición 35.55 (Diferenciales y divisor canónico)**.** Sea C una curva proyectiva suave sobre un cuerpo algebraicamente cerrado, con cuerpo de funciones $K(C)$. Una *diferencial racional* es una expresión formal $\omega = f\,\mathrm{d}g$ con $f, g \in K(C)$, con las reglas $\mathrm{d}(g+h) = \mathrm{d}g + \mathrm{d}h$, $\mathrm{d}(gh) = g\,\mathrm{d}h + h\,\mathrm{d}g$ y $\mathrm{d}c = 0$ para las constantes. Forman un espacio vectorial de dimensión 1 sobre $K(C)$: fijada cualquier t con $\mathrm{d}t \neq 0$, toda ω se escribe $\omega = f\,\mathrm{d}t$ de manera única.

En cada punto P se elige un *parámetro local* t_P —una función que se anula en P a orden exactamente 1— y se define

$$\operatorname{ord}_P(\omega) = \operatorname{ord}_P(f) \qquad \text{si } \omega = f\,\mathrm{d}t_P.$$

El *divisor* de ω es $\operatorname{div}(\omega) = \sum_P \operatorname{ord}_P(\omega)\, P$, una suma finita. Cualquiera de esos divisores se llama *divisor canónico* y se denota K.

Observación 35.56 (Está bien definido salvo equivalencia, y eso basta)*.* Dos diferenciales no nulas se diferencian en un factor: $\omega' = h\,\omega$ con $h \in K(C)^\times$, luego $\operatorname{div}(\omega') = \operatorname{div}(h) + \operatorname{div}(\omega)$. Es decir, *el divisor canónico es único salvo equivalencia lineal*, que es exactamente lo que Riemann –Roch necesita: la fórmula

$$\ell(D) - \ell(K - D) = \deg D - g + 1$$

sólo depende de la clase de K, porque ℓ y deg son invariantes por equivalencia lineal. Y de ahí salen los dos hechos que el capítulo usaba sin justificar: tomando $D = 0$ se obtiene $\ell(K) = g$, y tomando $D = K$ resulta $\deg K = 2g - 2$.

Ejemplo 35.57 (El divisor canónico en los tres casos básicos)**.**

- *Recta proyectiva* $\mathbb{P}^1$, género $g = 0$: con la coordenada x, la diferencial $\mathrm{d}x$ no tiene ceros en el plano afín y tiene un polo de orden 2 en el infinito —porque con el parámetro local $u = 1/x$ se tiene $\mathrm{d}x = -u^{-2}\mathrm{d}u$—. Así $K = -2\,\infty$, con $\deg K = -2 = 2 \cdot 0 - 2$.

- *Curva elíptica* $y^2 = x^3 + ax + b$, género $g = 1$: la diferencial $\omega = \mathrm{d}x/y$ no tiene ni ceros ni polos —es la *diferencial invariante*—, luego $K = 0$ y $\deg K = 0 = 2 \cdot 1 - 2$. Que exista una diferencial sin ceros ni polos es lo que hace de la curva elíptica un grupo.
- *Curva plana suave de grado* d: su género es $g = \frac{(d-1)(d-2)}{2}$, y el divisor canónico es el que cortan las curvas de grado $d-3$; para $d = 4$ eso da $g = 3$ y K cortado por rectas, con $\deg K = 4 = 2 \cdot 3 - 2$.

Regla: Algoritmo de Buchberger: calcular una base de Gröbner

1. **Fija un orden monomial** $\prec$ —lexicográfico si se va a eliminar variables, grevlex si sólo se quiere resolver rápido— y escribe cada polinomio con su término líder $\mathrm{lt}(f)$ delante.
2. **Parte de** $G = \{f_1, \dots, f_s\}$, los generadores del ideal.
3. **Forma el** S**-polinomio** de cada par $f, g \in G$:
$$S(f,g) = \frac{L}{\mathrm{lt}(f)}\, f - \frac{L}{\mathrm{lt}(g)}\, g, \qquad L = \mathrm{mcm}(\mathrm{lm}(f), \mathrm{lm}(g)).$$
El diseño es exactamente el que cancela los dos términos líderes.
4. **Divide** $S(f,g)$ **por** G con el algoritmo de división multivariante, y quédate con el resto r.
5. **Si** $r \neq 0$**, añádelo a** G y vuelve al paso 3 con los nuevos pares. Si todos los restos son 0, G ya es una base de Gröbner.
6. **Reduce.** Quita de G todo polinomio cuyo término líder sea divisible por el de otro, y divide cada superviviente por los demás. Con el término líder normalizado a 1, la base *reducida* es **única**: es la forma canónica del ideal.

Por qué termina. Cada polinomio añadido aporta un término líder nuevo, de modo que el ideal de términos líderes crece estrictamente; y $K[x_1, \dots, x_n]$ es noetheriano por la base de Hilbert (capítulo 24), luego la cadena se estabiliza en un número finito de pasos.

Ejemplo 35.58 (Una base de Gröbner a mano, y una eliminación)**.** Sea $I = (f_1, f_2)$ con $f_1 = x^2 - y$ y $f_2 = xy - 1$, con orden lexicográfico $x \succ y$. Los términos líderes son x^2 y xy, con mínimo común múltiplo $x^2 y$. El S-polinomio es

$$S(f_1, f_2) = y\, f_1 - x\, f_2 = y(x^2 - y) - x(xy - 1) = x - y^2.$$

Dividido por $\{f_1, f_2\}$ no se reduce —su término líder x no es divisible por x^2 ni por xy—, así que se añade $f_3 = x - y^2$. Repitiendo con los nuevos pares, f_1 y f_2 se reducen a cero módulo f_3 (pues $x \equiv y^2$), y aparece $f_4 = y^3 - 1$. La base reducida es

$$G = \{\, x - y^2, \;\; y^3 - 1 \,\}.$$

Y ahí está la eliminación: el segundo polinomio *no contiene* x, de modo que $y^3 = 1$ da las tres soluciones $y \in \{1, \omega, \omega^2\}$ y la primera ecuación despeja $x = y^2$. El sistema original, que parecía enredado, resultó tener exactamente tres soluciones, y el orden lexicográfico las ha puesto en forma triangular sin que hiciera falta ninguna astucia.

35.11. Ejercicios resueltos

Tres problemas sobre el diccionario entre álgebra y geometría.

Ejercicio 35.1. Describa $V(I)$ para $I = (xy)$ y para $I = (x^2 + y^2 - 1,\, x - y)$ en $\mathbb{C}^2$, y calcule el ideal de cada variedad.

Solución. $V(xy)$ es la unión de los dos ejes. Su ideal es (xy), que es radical pero no primo, y esa falta de primalidad corresponde exactamente a que la variedad es reducible.

Para el segundo, sustituyendo $y = x$ en la circunferencia queda $2x^2 = 1$, luego $x = \pm 1/\sqrt{2}$: dos puntos. El ideal de ese conjunto es la intersección de los dos ideales maximales correspondientes.

La moraleja del diccionario: ideales radicales se corresponden con variedades, ideales primos con variedades irreducibles, e ideales maximales con puntos. Y todo ello sobre un cuerpo algebraicamente cerrado, porque sobre $\mathbb{R}$ el ideal $(x^2 + y^2)$ define un solo punto y no es radical, y el diccionario se rompe. ∎

Ejercicio 35.2. Enuncie el Nullstellensatz de Hilbert y explique por qué la hipótesis de cierre algebraico es imprescindible.

Solución. Si K es algebraicamente cerrado e $I \subseteq K[x_1, \dots, x_n]$ es un ideal, entonces $I(V(I)) = \sqrt{I}$. En particular, si $V(I) = \varnothing$ entonces $I = (1)$: un sistema de ecuaciones polinómicas sin solución tiene una combinación de sus miembros igual a 1, que es un certificado algebraico de la inconsistencia.

Sobre $\mathbb{R}$ falla estrepitosamente. El polinomio $x^2 + 1$ no tiene raíces reales, luego $V(x^2 + 1) = \varnothing$, y sin embargo $(x^2 + 1) \neq (1)$ en $\mathbb{R}[x]$. El diccionario entre ideales y variedades deja de ser una biyección, y la geometría algebraica real necesita herramientas completamente distintas —los conjuntos semialgebraicos, el teorema de Tarski–Seidenberg— para decir algo comparable. ∎

Ejercicio 35.3. Sobre la curva $y^2 = x^3 + 1$, calcule $P + Q$ con $P = (-1, 0)$ y $Q = (0, 1)$, y verifique el resultado.

Solución. La recta que une P y Q tiene pendiente 1 y ecuación $y = x + 1$. Sustituyendo,

$$(x+1)^2 = x^3 + 1 \iff x^3 - x^2 - 2x = 0 \iff x(x-2)(x+1) = 0,$$

de modo que los tres cortes están en $x = -1, 0, 2$. El tercero es $R = (2, 3)$, y por definición $P + Q = -R = (2, -3)$.

Verificación: $(-3)^2 = 9$ y $2^3 + 1 = 9$. Correcto. Obsérvese que el tercer punto de corte existe siempre gracias a Bézout —una recta y una cúbica se cortan en tres puntos del plano proyectivo sobre $\mathbb{C}$, contando multiplicidades—, y que ahí es donde entra el punto del infinito cuando la recta es vertical: ése es el neutro del grupo. ∎

35.12. Ejercicios propuestos

35.1. Describa $V(x^2 + y^2)$ en $\mathbb{C}^2$ y en $\mathbb{R}^2$, y explique la diferencia.

35.2. ¿Es primo el ideal (xy) en $\mathbb{C}[x, y]$? ¿Es radical?

35.3. Halle $I(V)$ para $V = \{(0,0)\}$ en $\mathbb{C}^2$.

35.4. ¿Cuál es la dimensión de $V(xy)$? ¿Y el número de componentes irreducibles?

35.5. Enuncie el teorema de la base de Hilbert y diga qué consecuencia tiene sobre las variedades.

35.6. ¿Cuántos puntos tiene la intersección de dos cónicas en $\mathbb{P}^2(\mathbb{C})$, contando multiplicidades?

35.7. ¿Cuál es el género de una curva plana lisa de grado 4?

35.8. ¿Por qué las cúbicas lisas no admiten parametrización racional?

35.9. Halle los puntos singulares de $y^2 = x^3$ y clasifique la singularidad.

35.10. ¿Qué es $\operatorname{Spec}\mathbb{Z}$ y qué papel juega el ideal (0)?

35.11. Calcule $|E(\mathbb{F}_5)|$ para $E : y^2 = x^3 + 1$ y compruebe la cota de Hasse.

35.12. Explique por qué el teorema de Bézout necesita el plano proyectivo y un cuerpo algebraicamente cerrado.

Capítulo 36

Cálculo tensorial

Lo que necesitas traer

Del capítulo 7: espacios vectoriales, base dual, formas bilineales y cambio de base. Del capítulo 9: derivadas parciales, regla de la cadena y jacobiano. Del capítulo 11: campos vectoriales, formas diferenciales y el teorema de Stokes. Del capítulo 25: producto tensorial y grupos de Lie.

El cálculo tensorial responde una pregunta simple: ¿cómo escribir las leyes de la geometría y de la física de modo que *no dependan del sistema de coordenadas*? La respuesta —los tensores— generaliza a la vez los vectores, las matrices y las formas cuadráticas, y desemboca en la geometría de los espacios curvos: el lenguaje de la relatividad general y de la mecánica de medios continuos.

36.1. Convenio de Einstein

Idea clave

Intuición primero. El convenio de Einstein no es un concepto: es un ahorro de tinta que además funciona como corrector ortográfico. La observación es que en física casi todos los sumatorios recorren un índice que aparece dos veces, una arriba y una abajo, así que se acuerda no escribir el signo $\sum$ y dar la suma por supuesta.
Lo valioso viene después. Como la posición del índice codifica *cómo se transforma* el objeto al cambiar de coordenadas, la gramática del convenio detecta errores sola: si un índice te queda arriba en un lado de la ecuación y abajo en el otro, o si aparece tres veces, la ecuación es incorrecta antes de calcular nada. Un índice repetido arriba y abajo desaparece del resultado —está sumado— y lo que sobrevive tiene que casar término a término. Úsalo como checksum: en este capítulo ahorra más errores que cualquier otra precaución.

De ahora en adelante, **un índice repetido una vez arriba y una vez abajo se suma automáticamente**:

$$a^i b_i \equiv \sum_{i=1}^{n} a^i b_i .$$

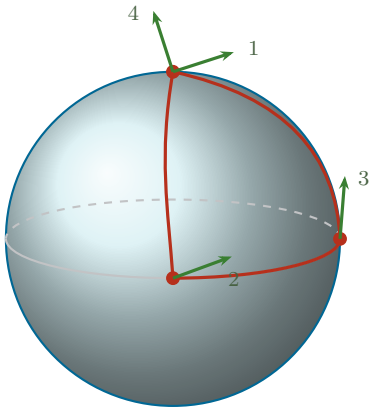


Figura 36.1: Transporte paralelo alrededor de un triángulo geodésico. El vector 1 parte del polo, baja por un meridiano hasta el ecuador sin girar respecto de la superficie (2), recorre un cuarto de ecuador (3) y vuelve al polo por el otro meridiano. Al llegar (4) ha rotado 90°. Ese desajuste *es* la curvatura.

Reglas gramaticales del convenio: un índice repetido (*mudo*) puede renombrarse a placer ($a^i b_i = a^k b_k$); un índice libre debe aparecer exactamente una vez en cada término y en la misma posición en ambos lados de una ecuación; ningún índice puede aparecer tres veces. La posición del índice (arriba/abajo) no es decorativa: codifica cómo se transforma el objeto.

36.2. Índices arriba y abajo, sumatorios y el convenio de Einstein: el alfabeto con el que está escrito todo el capítulo

La sección anterior ha anunciado el convenio de Einstein en dos líneas y ha dado por supuesto que el lector ya sabe qué es un índice, qué significa que esté arriba o abajo, y qué quiere decir la expresión $a^i b_i$. Nada de eso es evidente la primera vez, y saltárselo tiene un coste altísimo: un capítulo entero escrito en un alfabeto que no se ha aprendido. Esta sección enseña ese alfabeto desde cero, letra a letra, con todos los sumatorios escritos hasta el último término. No hay ninguna idea profunda aquí: hay contabilidad, y la contabilidad hay que aprenderla antes de usarla.

Conviene decir de entrada qué problema resuelve la notación. En física y en geometría casi todas las cantidades interesantes son *sumas* de productos: la energía cinética suma sobre las tres direcciones del espacio, el producto de matrices suma sobre un índice intermedio, la longitud de un vector suma sobre las componentes. Escribir todos esos sumatorios con su símbolo $\sum$ ocupa mucho sitio y, sobre todo, esconde la estructura detrás de un bosque de signos. La notación indicial hace lo contrario: deja a la vista qué se suma con qué, y convierte la comprobación de una fórmula larga en un recuento de letras que se hace de un vistazo.

36.2.1. Índices y componentes

Definición 36.1 (Notación indicial)**.** Sea n un entero positivo, la *dimensión*. Un *índice* es una letra que recorre los valores $1, 2, \ldots, n$. Escribir v^i significa: la lista de n números $v^1, v^2, \ldots, v^n$,

y v^i es el que ocupa el lugar i. Escribir $T^i{}_j$ significa la tabla de n^2 números $T^i{}_j$, uno por cada par (i,j). En general, un símbolo con p índices arriba y q índices abajo denota una colección de n^{p+q} números, sus *componentes*, indexadas por todas las combinaciones posibles de valores.

Los superíndices **no son potencias**: v^2 es la segunda componente de la lista v, no el cuadrado de v. Cuando de verdad haga falta un cuadrado se escribirá $(v^2)^2$ o se dirá con palabras.

Cuidado

Este choque de notaciones es real y provoca errores durante las primeras semanas. La costumbre universal en geometría es reservar los superíndices para los índices y renunciar a las potencias salvo con paréntesis explícitos. Cuando aparezca x^2 en este capítulo, significará casi siempre *la segunda coordenada*; si el contexto pidiera un cuadrado, se escribirá $(x^1)^2$, con su paréntesis.

Ejemplo 36.2 (Componentes, contadas y escritas)**. (a) Un vector en dimensión 3.** La lista v^i con $v^1 = 2$, $v^2 = -1$, $v^3 = 4$ son tres números y nada más. En la notación de bachillerato serían $v = (2, -1, 4)$.

(b) Una tabla con dos índices. En dimensión 2, el símbolo $A^i{}_j$ designa cuatro números, que es cómodo colocar en una matriz con el índice de arriba marcando la fila y el de abajo la columna:

$$A^i{}_j = \begin{pmatrix} A^1{}_1 & A^1{}_2 \\ A^2{}_1 & A^2{}_2 \end{pmatrix} = \begin{pmatrix} 2 & -1 \\ 0 & 3 \end{pmatrix}.$$

Así, $A^1{}_2 = -1$ y $A^2{}_1 = 0$: el orden importa.

(c) Cuántos números hay. Un símbolo con $p+q$ índices en dimensión n tiene n^{p+q} componentes. En dimensión 4, que es la del espaciotiempo, un símbolo con dos índices tiene 16 componentes, con tres índices 64 y con cuatro índices 256. Esa explosión es la razón de que más adelante se busquen simetrías: son lo único que hace manejables esas tablas.

(d) Los nombres de las letras dan igual. Las colecciones v^i y v^k son la misma; la letra es un hueco, no un dato. Lo único que no se puede hacer es reutilizar una letra que ya esté ocupada en la misma expresión, y a eso se dedica el apartado siguiente.

36.2.2. El convenio de sumación

Definición 36.3 (Convenio de sumación de Einstein)**.** En una expresión formada por productos de componentes, si un mismo índice aparece *exactamente dos veces en el mismo término, una vez arriba y una vez abajo*, se entiende que hay un sumatorio sobre ese índice desde 1 hasta n, aunque el signo $\sum$ no se escriba:

$$a^i b_i \;\equiv\; \sum_{i=1}^{n} a^i b_i.$$

El convenio es solo eso: un acuerdo de escritura para no repetir el símbolo $\sum$.

Definición 36.4 (Índice mudo e índice libre)**.** En un término, un índice que aparece dos veces (una arriba y otra abajo) se llama *mudo*: está sumado y *no* aparece en el resultado. Un índice que aparece una sola vez se llama *libre*: sobrevive, y el resultado tiene un índice por cada uno de ellos. Un índice mudo puede rebautizarse con cualquier letra que no esté ya en uso en ese término, sin que cambie nada.

Ejemplo 36.5 (Seis expresiones expandidas hasta el último término)**.** Todo lo que sigue es literalmente la definición 36.3 aplicada con paciencia. Merece la pena hacerlo una vez a mano.

(a) Un índice mudo en dimensión 3**.** La expresión $a^i b_i$ tiene un solo índice, y es mudo. Expandiendo,

$$a^i b_i = a^1 b_1 + a^2 b_2 + a^3 b_3.$$

Con $a^i = (2, -1, 4)$ y $b_i = (3, 0, 5)$ sale $2 \cdot 3 + (-1) \cdot 0 + 4 \cdot 5 = 6 + 0 + 20 = 26$. El resultado es *un número*, sin ningún índice: como debía ser, porque no quedaban índices libres.

(b) Un índice libre y otro mudo. La expresión $w^i = A^i{}_j v^j$ tiene la j muda y la i libre. En dimensión 2 hay que escribirla dos veces, una por cada valor de i:

$$w^1 = A^1{}_1 v^1 + A^1{}_2 v^2, \qquad w^2 = A^2{}_1 v^1 + A^2{}_2 v^2.$$

Con la matriz del ejemplo anterior y $v^j = (4, 5)$: $w^1 = 2 \cdot 4 + (-1) \cdot 5 = 3$ y $w^2 = 0 \cdot 4 + 3 \cdot 5 = 15$. Es, exactamente, el producto de una matriz por un vector de toda la vida: la notación indicial no inventa operaciones nuevas, solo las escribe de otra manera.

(c) Dos índices mudos. La expresión $g_{ij} u^i v^j$ tiene los dos índices mudos, luego es un número, y expandirla en dimensión 2 da *cuatro* sumandos:

$$g_{ij} u^i v^j = g_{11} u^1 v^1 + g_{12} u^1 v^2 + g_{21} u^2 v^1 + g_{22} u^2 v^2.$$

Tomemos las coordenadas polares del plano, donde (como se verá) $g_{11} = 1$, $g_{22} = r^2$ y $g_{12} = g_{21} = 0$, y evaluemos en $r = 2$ con $u^i = (1;\, 0{,}5)$ y $v^j = (2, -1)$:

$$1 \cdot 1 \cdot 2 + 0 + 0 + 4 \cdot 0{,}5 \cdot (-1) = 2 - 2 = 0.$$

Cero: los dos vectores son perpendiculares en esa geometría, cosa que *no* se ve mirando las componentes a secas.

(d) Una suma doble con coeficientes de tres índices. La expresión $a^k = \Gamma^k{}_{ij} v^i v^j$ tiene k libre e i, j mudos: son n ecuaciones de n^2 sumandos cada una. En dimensión 2, para $k = 1$:

$$a^1 = \Gamma^1{}_{11} v^1 v^1 + \Gamma^1{}_{12} v^1 v^2 + \Gamma^1{}_{21} v^2 v^1 + \Gamma^1{}_{22} v^2 v^2,$$

y otro tanto para $k = 2$. Ocho sumandos en total, escritos en once caracteres. Cuando más adelante los Γ de las coordenadas polares resulten ser $\Gamma^r{}_{\theta\theta} = -r$ y $\Gamma^\theta{}_{r\theta} = \Gamma^\theta{}_{\theta r} = 1/r$, con todos los demás nulos, esta expresión se reducirá a $a^r = -r(v^\theta)^2$ y $a^\theta = \frac{2}{r} v^r v^\theta$: los términos centrífugo y de Coriolis, aparecidos solos.

(e) Índices que no se suman. En T^{ij} no hay nada que sumar: los dos índices están arriba y aparecen una sola vez cada uno. Son n^2 números, no un número. Del mismo modo, $\partial_i f$ es una colección de n derivadas parciales, no una suma de ellas.

(f) Renombrar un mudo no cambia nada. Las tres expresiones $a^i b_i$, $a^k b_k$ y $a^m b_m$ son idénticas: las tres valen 26 con los números de (a). Pero en el término $A^i{}_k B^k{}_j$ no se puede renombrar k como i ni como j, porque esas letras ya están ocupadas por índices libres y la expresión resultante significaría otra cosa.

Regla: expandir a mano una expresión con índices

Un procedimiento mecánico para pasar de la notación comprimida a la lista completa de sumandos. Vale la pena aplicarlo palabra por palabra las primeras diez veces.

1. **Subraya los índices repetidos arriba y abajo.** Son los mudos, y cada uno aporta un sumatorio de n términos.
2. **Anota los que quedan.** Son los libres, y dicen cuántas ecuaciones hay que escribir: una por cada combinación de valores de los índices libres.
3. **Cuenta antes de escribir.** Con ℓ índices libres y m mudos, la expresión equivale a n^ℓ ecuaciones de n^m sumandos cada una. Saber el número de términos antes de empezar evita perder alguno por el camino.
4. **Fija los libres y recorre los mudos.** Escoge un valor concreto para cada índice libre y recorre todos los valores de los mudos, en orden, sin saltarte ninguno.
5. **Comprueba la posición de los índices en el resultado.** Los índices libres deben quedar en los dos miembros de la igualdad en el mismo número y en el mismo nivel, arriba o abajo. Si no, la expresión estaba mal escrita antes de calcular nada.
6. **Elimina de un plumazo los términos con factores nulos.** En casi todos los ejemplos del capítulo, la mayoría de las componentes son cero, y una lista previa de las no nulas reduce ocho sumandos a uno.

36.2.3. La delta de Kronecker

Hay un símbolo que aparecerá en cada página de aquí en adelante y que conviene definir de una vez, porque se usa sin avisar en la propia definición de base dual.

Definición 36.6 (Delta de Kronecker)**.** La *delta de Kronecker* es la colección de n^2 números

$$\delta^i_j = \begin{cases} 1 & \text{si } i = j, \\ 0 & \text{si } i \neq j, \end{cases}$$

es decir, la matriz identidad escrita con un índice arriba y otro abajo. En dimensión 3, por ejemplo, $\delta^1_1 = \delta^2_2 = \delta^3_3 = 1$ y las otras seis componentes valen 0.

Proposición 36.7 (Las tres propiedades que se usan siempre)**.** *Con el convenio de sumación se cumplen:*

(I) $\delta^i_j v^j = v^i$ *y* $\delta^i_j \omega_i = \omega_j$*: la delta* renombra *el índice contra el que se contrae;*

(ii) $\delta^i_i = n$, *y no* 1;

(iii) $\delta^i_k \delta^k_j = \delta^i_j$.

Demostración. (i) Al expandir $\delta^i_j v^j = \delta^i_1 v^1 + \cdots + \delta^i_n v^n$, todos los sumandos son cero salvo aquel en que el índice mudo toma el valor i, y ese vale $1 \cdot v^i$. Queda v^i. El segundo enunciado es idéntico cambiando el papel de los dos índices.

(ii) Aquí los dos índices son el mismo, luego es mudo y hay que sumar: $\delta^i_i = \delta^1_1 + \delta^2_2 + \cdots + \delta^n_n = 1 + 1 + \cdots + 1 = n$. En dimensión 4 vale 4, y confundirlo con 1 estropea cualquier cálculo en el que aparezca.

(iii) Por (i), aplicado con v^k sustituido por δ^k_j: contraer con la delta renombra el índice, y el resultado es δ^i_j. □

Ejemplo 36.8 (La delta trabajando)**. (a) Renombrar.** En dimensión 2, $\delta^i_j v^j$ para $i = 1$ es $\delta^1_1 v^1 + \delta^1_2 v^2 = 1 \cdot v^1 + 0 \cdot v^2 = v^1$. Lo mismo con $i = 2$. Efectivamente, la delta no hace nada salvo cambiar el nombre del índice.

(b) La traza mal contada. En dimensión 4, $\delta^\mu_\mu = 4$. Este número aparece constantemente en relatividad, por ejemplo al tomar la traza de la ecuación de Einstein, y escribir 1 en su lugar cambia el resultado de arriba abajo.

(c) Un producto de matrices disfrazado. La igualdad $A^i{}_k \delta^k_j = A^i{}_j$ dice que multiplicar por la identidad no cambia nada; escrita con índices es un caso particular de la propiedad (i).

(d) El error clásico de las tres apariciones. La expresión $\delta^i_i v^i$ no significa nada: el índice i aparece tres veces y el convenio no dice qué hacer con él. Si se quiere $n\, v^i$, hay que escribir $\delta^k_k v^i$, con un índice mudo distinto.

36.2.4. Arriba, abajo y el cambio de coordenadas

Queda por explicar lo único que no es notación: por qué unos índices van arriba y otros abajo. La respuesta corta es que los dos tipos de componentes *cambian de manera opuesta* cuando se cambia de coordenadas, y que la posición del índice es la etiqueta que recuerda cuál es cuál.

Definición 36.9 (Cambio de coordenadas y sus dos jacobianas)**.** Sea un cambio de coordenadas que asigna a cada punto dos juegos de nombres, x^i y $\tilde{x}^\mu$, mediante funciones derivables y con inversa derivable. Sus *matrices jacobianas* son

$$J^\mu{}_i = \frac{\partial \tilde{x}^\mu}{\partial x^i}, \qquad (J^{-1})^i{}_\mu = \frac{\partial x^i}{\partial \tilde{x}^\mu},$$

y ambas son funciones del punto: hay una matriz distinta en cada sitio.

Proposición 36.10. *Las dos jacobianas son inversas una de la otra:* $\dfrac{\partial \tilde{x}^\mu}{\partial x^i}\dfrac{\partial x^i}{\partial \tilde{x}^\nu} = \delta^\mu_\nu$.

Demostración. Basta aplicar la regla de la cadena a la función que lleva $\tilde{x}^\nu$ en $\tilde{x}^\mu$ pasando por las x^i: su derivada es la suma de la izquierda. Pero esa función es la identidad, y la derivada de $\tilde{x}^\mu$ respecto de $\tilde{x}^\nu$ vale 1 si $\mu = \nu$ y 0 en otro caso, es decir δ^μ_ν. □

Definición 36.11 (Componentes contravariantes y covariantes)**.** Una colección de n números asociada a cada sistema de coordenadas se llama *contravariante* (y se escribe con el índice **arriba**) si al cambiar de coordenadas obedece

$$\tilde{V}^{\mu} = \frac{\partial \tilde{x}^{\mu}}{\partial x^{i}}\, V^{i},$$

y se llama *covariante* (índice **abajo**) si obedece

$$\tilde{\omega}_{\mu} = \frac{\partial x^{i}}{\partial \tilde{x}^{\mu}}\, \omega_{i}.$$

Es decir: los índices de arriba se transforman con J y los de abajo con J^{-1}.

Definición 36.12 (Ley de transformación, o definición clásica de tensor)**.** Una colección de n^{p+q} números $T^{i_1\dots i_p}{}_{j_1\dots j_q}$, dada en cada sistema de coordenadas, es un *tensor de tipo* (p,q) si al cambiar de coordenadas se transforma con *un jacobiano por índice*: una copia de $\partial\tilde{x}/\partial x$ por cada índice de arriba y una copia de $\partial x/\partial\tilde{x}$ por cada índice de abajo. Para los dos casos que más se usan,

$$\tilde{T}^{\mu}{}_{\nu} = \frac{\partial \tilde{x}^{\mu}}{\partial x^{i}}\frac{\partial x^{j}}{\partial \tilde{x}^{\nu}}\, T^{i}{}_{j}, \qquad \tilde{g}_{\mu\nu} = \frac{\partial x^{i}}{\partial \tilde{x}^{\mu}}\frac{\partial x^{j}}{\partial \tilde{x}^{\nu}}\, g_{ij}.$$

Idea clave

El motivo de que la ley sea esta y no otra es puramente aritmético: es la única que hace que las contracciones no dependan de las coordenadas. Al contraer un índice de arriba con uno de abajo aparece el producto $J\,J^{-1}$, que por la proposición 36.10 es la delta y desaparece. Si los dos tipos de índice se transformaran igual, el número $\omega_i V^i$ cambiaría de valor al cambiar de coordenadas y no serviría para describir nada físico.

Ejemplo 36.13 (El cambio de cartesianas a polares, calculado entero)**.** Este es el ejemplo de referencia del capítulo, y conviene hacerlo con todos los números. Trabajamos en el plano, con coordenadas cartesianas $x^i = (x, y)$ y polares $\tilde{x}^\mu = (r, \theta)$, ligadas por $x = r\cos\theta$, $y = r\sin\theta$. Fijamos de una vez el punto

$$P = (x, y) = (\sqrt{3},\, 1), \qquad \text{es decir} \qquad r = \sqrt{3+1} = 2, \quad \theta = \frac{\pi}{6},$$

donde $\cos\theta = \frac{\sqrt{3}}{2} \approx 0{,}8660$ y $\sin\theta = \frac{1}{2} = 0{,}5$.

Paso 1: las dos jacobianas. Derivando x e y respecto de r y de θ,

$$\frac{\partial x^{i}}{\partial \tilde{x}^{\mu}} = \begin{pmatrix} \cos\theta & -r\sin\theta \\ \sin\theta & r\cos\theta \end{pmatrix} = \begin{pmatrix} 0{,}8660 & -1 \\ 0{,}5 & 1{,}7321 \end{pmatrix},$$

y despejando $r = \sqrt{x^2+y^2}$ y derivando (o invirtiendo la matriz anterior),

$$\frac{\partial \tilde{x}^{\mu}}{\partial x^{i}} = \begin{pmatrix} \cos\theta & \sin\theta \\ -\dfrac{\sin\theta}{r} & \dfrac{\cos\theta}{r} \end{pmatrix} = \begin{pmatrix} 0{,}8660 & 0{,}5 \\ -0{,}25 & 0{,}4330 \end{pmatrix}.$$

Comprobación de la proposición 36.10: el producto de las dos matrices vale $\left(\begin{smallmatrix} 0{,}8660\cdot 0{,}8660+0{,}5\cdot 0{,}5 & \cdot \\ \cdot & \cdot \end{smallmatrix}\right)$, y la entrada calculada da $0{,}75+0{,}25=1$; las demás salen 0, 0 y 1. Es la identidad, como debía.

Paso 2: un objeto con el índice arriba. Sea V con componentes cartesianas $V^x=1$, $V^y=2$. Por la definición 36.11,

$$V^r=\frac{\partial r}{\partial x}V^x+\frac{\partial r}{\partial y}V^y=0{,}8660\cdot 1+0{,}5\cdot 2=1{,}8660,$$

$$V^\theta=\frac{\partial \theta}{\partial x}V^x+\frac{\partial \theta}{\partial y}V^y=-0{,}25\cdot 1+0{,}4330\cdot 2=0{,}6160.$$

Control con un escalar. La métrica polar es $g=\operatorname{diag}(1,r^2)=\operatorname{diag}(1,4)$, luego

$$g_{\mu\nu}V^\mu V^\nu=1\cdot(1{,}8660)^2+4\cdot(0{,}6160)^2=3{,}4821+1{,}5179=5,$$

y en cartesianas $1^2+2^2=5$. Coincide exactamente: la longitud no depende del sistema de coordenadas, aunque las componentes sí.

Paso 3: un objeto con el índice abajo. Sea $f(x,y)=x^2+y^2$ y $\omega_i=\partial_i f$, es decir $\omega_x=2x=2\sqrt{3}\approx 3{,}4641$ y $\omega_y=2y=2$. Ahora hay que usar la *otra* jacobiana:

$$\omega_r=\frac{\partial x}{\partial r}\omega_x+\frac{\partial y}{\partial r}\omega_y=0{,}8660\cdot 3{,}4641+0{,}5\cdot 2=3+1=4,$$

$$\omega_\theta=\frac{\partial x}{\partial \theta}\omega_x+\frac{\partial y}{\partial \theta}\omega_y=(-1)\cdot 3{,}4641+1{,}7321\cdot 2=-3{,}4641+3{,}4641=0.$$

El resultado es el esperado y sirve de comprobación independiente: en polares la función es $f=r^2$, luego $\partial_r f=2r=4$ y $\partial_\theta f=0$. La ley de transformación de un índice abajo reproduce exactamente las derivadas parciales calculadas en las coordenadas nuevas, que es justo lo que tenía que pasar.

Paso 4: un objeto con dos índices abajo. Tomemos T_{ij} con la única componente no nula $T_{xx}=1$ (en el lenguaje que viene después, $T=\mathrm{d}x\otimes\mathrm{d}x$). Con dos jacobianas, una por índice:

$$\tilde T_{rr}=\Big(\frac{\partial x}{\partial r}\Big)^2 T_{xx}=\cos^2\theta=0{,}75,$$

$$\tilde T_{r\theta}=\tilde T_{\theta r}=\frac{\partial x}{\partial r}\frac{\partial x}{\partial \theta}T_{xx}=\cos\theta\,(-r\sin\theta)=-0{,}8660,\qquad \tilde T_{\theta\theta}=\Big(\frac{\partial x}{\partial \theta}\Big)^2 T_{xx}=r^2\sin^2\theta=1.$$

Control: contrayendo con la métrica inversa $g^{\mu\nu}=\operatorname{diag}(1;\,0{,}25)$ se obtiene $1\cdot 0{,}75+0{,}25\cdot 1=1$, que es el mismo número que $\delta^{ij}T_{ij}=T_{xx}=1$ en cartesianas. Nótese que en polares el tensor tiene un término cruzado que en cartesianas no existía: las componentes cambian de forma, el objeto no.

Paso 5: un objeto con un índice arriba y otro abajo. Sea $A^i{}_j$ la matriz $\operatorname{diag}(1,-1)$ en cartesianas. Ahora hacen falta las dos jacobianas a la vez, y el cálculo es un producto de tres matrices, $\tilde A=J^{-1}AJ$ en la notación de siempre. El resultado, simplificado con las fórmulas del ángulo doble, es

$$\tilde A^\mu{}_\nu=\begin{pmatrix} \cos 2\theta & -r\sin 2\theta \\ -\dfrac{\sin 2\theta}{r} & -\cos 2\theta \end{pmatrix}=\begin{pmatrix} 0{,}5 & -1{,}7321 \\ -0{,}4330 & -0{,}5 \end{pmatrix}\quad \text{en } P.$$

Control: la traza vale $0{,}5 - 0{,}5 = 0$, igual que la traza cartesiana $1 + (-1) = 0$. La contracción de un índice de arriba con uno de abajo produce siempre un número que no depende de las coordenadas, y comprobarlo cuesta dos sumas.

Moraleja. Cinco objetos, cinco leyes distintas, y en los cinco casos la regla es la misma: un jacobiano por índice, con $\partial\tilde{x}/\partial x$ para los de arriba y $\partial x/\partial\tilde{x}$ para los de abajo. Todo el resto del capítulo consiste en aplicar esto sin equivocarse.

Cuidado

Dos avisos que ahorran horas. Primero: las jacobianas dependen del punto, así que hay que evaluarlas *en el mismo punto* en el que están dadas las componentes; mezclar la jacobiana de un sitio con las componentes de otro produce números sin ningún significado. Segundo: que las componentes de un objeto cambien al cambiar de coordenadas no significa que el objeto cambie. En el paso 2, un vector constante en cartesianas tiene componentes polares que dependen de r y de θ; el vector no se ha movido, se han movido los ejes con los que se le mide.

36.2.5. Dos operaciones que ya se pueden definir con índices

Definición 36.14 (Producto tensorial en componentes)**.** Dadas dos colecciones de componentes, su *producto tensorial* es la colección que se obtiene multiplicando todas las componentes de una por todas las de la otra, sin sumar nada:

$$(S \otimes T)^{ik}{}_{jl} = S^{i}{}_{j}\, T^{k}{}_{l}.$$

Si S es de tipo (p, q) y T de tipo (p', q'), el producto es de tipo $(p+p', q+q')$ y tiene $n^{p+q} \cdot n^{p'+q'}$ componentes.

Definición 36.15 (Contracción)**.** *Contraer* una colección de componentes consiste en igualar un índice de arriba con uno de abajo y sumar según el convenio. De un objeto de tipo (p, q) se obtiene otro de tipo $(p-1, q-1)$. El caso más sencillo es la *traza* de un objeto de tipo $(1, 1)$:

$$T^{i}{}_{i} = T^{1}{}_{1} + T^{2}{}_{2} + \cdots + T^{n}{}_{n}.$$

Proposición 36.16. *La traza $T^{i}{}_{i}$ de un tensor de tipo $(1, 1)$ es un número que no depende del sistema de coordenadas.*

Demostración. Por la definición 36.12 y poniendo $\nu = \mu$ para contraer,

$$\tilde{T}^{\mu}{}_{\mu} = \frac{\partial\tilde{x}^{\mu}}{\partial x^{i}}\frac{\partial x^{j}}{\partial\tilde{x}^{\mu}}\, T^{i}{}_{j} = \delta^{j}_{i}\, T^{i}{}_{j} = T^{i}{}_{i},$$

donde el paso central es la proposición 36.10 (los dos jacobianas se multiplican y dan la delta) y el último es la propiedad (i) de la proposición 36.7. □

Ejemplo 36.17 (Producto y contracción con números)**.** Trabajemos en dimensión 2 con $u^{i} = (3, -1)$ y $\omega_{j} = (2, 5)$.

(a) El producto. $M^i{}_j = u^i\omega_j$ tiene cuatro componentes: $M^1{}_1 = 3\cdot 2 = 6$, $M^1{}_2 = 3\cdot 5 = 15$, $M^2{}_1 = -1\cdot 2 = -2$ y $M^2{}_2 = -1\cdot 5 = -5$. Como matriz, $\left(\begin{smallmatrix} 6 & 15 \\ -2 & -5 \end{smallmatrix}\right)$.

(b) La contracción. $M^i{}_i = 6+(-5) = 1$, que coincide con $u^i\omega_i = 3\cdot 2+(-1)\cdot 5 = 6-5 = 1$. No es casualidad: contraer el producto de un vector por un covector es exactamente evaluarlos uno en el otro.

(c) Contraer no es lo mismo que sumar. La expresión M^{ij} con los dos índices arriba no admite contracción: M^{ii} no significa nada invariante, porque el convenio pide un índice arriba y otro abajo. Para tomar su traza hay que bajar antes un índice con la métrica, y el resultado en general *no* es la suma de las entradas de la diagonal.

(d) El orden importa. $u^i\omega_j$ y $\omega_i u^j$ son objetos distintos: el primero es la matriz de (a) y el segundo su traspuesta, $\left(\begin{smallmatrix} 6 & -2 \\ 15 & -5 \end{smallmatrix}\right)$. Las dos tienen la misma traza, 1, coherente con la proposición 36.16.

Pausa para pensar

Vale la pena resumir qué se ha ganado. Un índice repetido arriba y abajo es una suma que desaparece; un índice suelto es una etiqueta que sobrevive; la posición del índice recuerda con cuál de las dos jacobianas se transforma; y la delta de Kronecker es el símbolo que convierte una suma en un renombramiento. Con estas cuatro frases, cualquier ecuación del capítulo se puede leer, expandir y comprobar antes de entender qué significa. Que la comprobación sintáctica pueda hacerse sin entender el contenido no es un defecto de la notación: es su mayor virtud.

36.3. Vectores, covectores y el espacio dual

Idea clave

Intuición primero. Hay dos clases de objetos con n números, y confundirlas es el error clásico. Un *vector* es una flecha: un desplazamiento, una velocidad. Un *covector* es una máquina de medir flechas: le das una flecha y te devuelve un número.
La prueba de que son distintos aparece al cambiar de unidades. Pasa de metros a centímetros: la base se encoge cien veces, y las componentes de un desplazamiento se multiplican por 100 —van *al revés* que la base, por eso se llaman contravariantes—. Pero las componentes de un gradiente, que se miden en "grados por unidad de longitud", se dividen entre 100: van *con* la base, son covariantes. Nada de esto se nota en coordenadas cartesianas con la métrica euclídea, y esa es justamente la razón de que en cursos anteriores nadie lo mencione; en cuanto las coordenadas son curvas o la métrica no es la identidad, la distinción se vuelve inevitable.
Regla mnemotécnica que no falla: los $\mathrm{d}x^i$ son covectores (índice arriba en el símbolo, pero miden), los $\partial/\partial x^i$ son vectores. Y el emparejamiento $\omega_i v^i$ da un número que no depende de la base: eso es lo único que de verdad existe.

Sea V un espacio vectorial real de dimensión n con base $\{e_1, \dots, e_n\}$; un vector es $v = v^i e_i$.

Definición 36.18 (Espacio dual)**.** V^* es el espacio de las aplicaciones lineales $\omega : V \to \mathbb{R}$ (*covectores* o *1-formas*). Tiene dimensión n, con *base dual* $\{e^1, \dots, e^n\}$ definida por $e^i(e_j) = \delta^i_j$. Un covector es $\omega = \omega_i e^i$ y su acción es $\omega(v) = \omega_i v^i$.

Idea clave

Qué dice esto en cristiano. Un vector es una flecha; un covector es una máquina que se come flechas y escupe números. En $\mathbb{R}^3$, la flecha (x, y, z) es un vector y la regla "dame una flecha y te devuelvo su altura z" es un covector. Parecen lo mismo porque en presencia de un producto escalar se pueden identificar, pero no lo son: al cambiar de unidades, los vectores y los covectores se transforman de maneras *opuestas*. Si se mide en centímetros en vez de metros, las componentes de un vector se multiplican por cien y las de un covector se dividen por cien. De ahí vienen los nombres "contravariante" y "covariante", y de ahí viene la costumbre de escribir unos índices arriba y otros abajo.

Ejemplo 36.19 (Covectores concretos)**. (a) La base dual.** En $\mathbb{R}^3$ con base canónica e_1, e_2, e_3, la base dual e^1, e^2, e^3 está formada por las tres proyecciones: $e^2(x, y, z) = y$. La relación que las define es $e^i(e_j) = \delta^i_j$, es decir, vale 1 si los índices coinciden y 0 si no.

(b) Un gradiente. Si $f(x, y, z) = 3x - 2y + z$, la diferencial $\mathrm{d}f = 3\,\mathrm{d}x - 2\,\mathrm{d}y + \mathrm{d}z$ es un covector: aplicada al vector $v = (1, 1, 1)$ da $3 - 2 + 1 = 2$, que es la derivada direccional de f en la dirección de v. El gradiente "natural" de f es la diferencial, no el vector $(3, -2, 1)$; ese vector sólo aparece después de usar el producto escalar para subir el índice.

(c) Cambio de base. Si se cambia e_1 por $\tilde{e}_1 = 2e_1$, entonces $\tilde{e}^1 = \frac{1}{2}e^1$, porque hay que seguir cumpliendo $\tilde{e}^1(\tilde{e}_1) = 1$. Un vector con componente 3 en e_1 pasa a tener componente 1,5 en $\tilde{e}_1$; un covector con componente 3 en e^1 pasa a tener componente 6. Se transforman al revés.

(d) Sin dimensión finita no hay isomorfismo natural. Para $\dim V = n < \infty$ se tiene $V \cong V^*$, pero el isomorfismo depende de la base elegida: no es canónico. En cambio $V \cong V^{**}$ sí es canónico, mediante $v \mapsto (\omega \mapsto \omega(v))$. Y en dimensión infinita ni siquiera eso: el dual de ℓ^1 es ℓ^∞, que es estrictamente mayor.

36.3.1. La ley de transformación

Cambiemos de base: $e'_j = A^i_j e_i$. Para que el objeto v no cambie, sus componentes deben moverse "al revés" que la base:

$$v'^j = (A^{-1})^j_i \, v^i \qquad (\textit{contravariantes}, \text{ índice arriba}),$$

mientras que las componentes de un covector se mueven *igual* que la base:

$$\omega'_j = A^i_j \, \omega_i \qquad (\textit{covariantes}, \text{ índice abajo}).$$

Verificación de consistencia: $\omega'_j v'^j = A^i_j (A^{-1})^j_k \, \omega_i v^k = \delta^i_k \omega_i v^k = \omega_i v^i$: el número $\omega(v)$ es invariante, como debe ser.

Idea clave

Ejemplo físico: al pasar de metros a centímetros la base se encoge por 100 y las componentes de un desplazamiento se multiplican por 100 (contravariantes); pero las componentes de un gradiente —variación *por unidad de longitud*— se dividen entre 100 (covariantes). El gradiente es por naturaleza un covector, no un vector: la distinción existe físicamente aunque en coordenadas cartesianas con la métrica euclídea pase desapercibida.

En cálculo en variedades, las derivadas parciales $\{\partial/\partial x^i\}$ forman la base natural de los vectores (tangentes) y los diferenciales $\{\mathrm{d}x^i\}$ la base dual de los covectores, con $\mathrm{d}x^i(\partial_j) = \delta^i_j$.

Regla: leer, escribir y verificar una ecuación con índices

Antes de calcular nada conviene aprender a *auditar* una expresión tensorial. Estas comprobaciones cuestan segundos y detectan la inmensa mayoría de los errores.

1. **Clasifica cada índice.** Un índice que aparece *dos veces* en un mismo término, una arriba y una abajo, es *mudo*: está sumado y desaparece del resultado. Un índice que aparece *una sola vez* es *libre*: sobrevive.

2. **Los índices libres deben coincidir término a término.** En $T^{ij} = A^i_k B^{kj} + C^{ij}$ los índices libres son i y j en los tres términos, arriba en los tres. Si en un sumando apareciera $T^i{}_j$ con la j abajo, la ecuación estaría mal escrita, sin necesidad de saber qué son A, B y C.

3. **Ningún índice puede aparecer tres veces.** Escribir $a^i b_i c_i$ no significa nada; si de verdad quieres una triple suma, hay que introducir un índice nuevo o un factor métrico explícito.

4. **Los índices mudos se pueden renombrar libremente,** siempre que el nombre nuevo no choque con otro ya presente en el mismo término. $a^k b_k$ y $a^m b_m$ son idénticos; pero si el término ya contiene una m libre, renombrar a m crea una colisión y rompe la expresión.

5. **Cuenta el tipo del resultado.** Si a un lado sobreviven dos índices arriba y uno abajo, el objeto es de tipo $(2,1)$ y el otro lado debe serlo también.

6. **Recuerda la asimetría de la delta.** δ^i_j actúa como un renombrador: $\delta^i_j v^j = v^i$. Y su traza no es 1 sino la dimensión: $\delta^i_i = n$. Este último es un despiste frecuentísimo.

Ejemplo 36.20 (Auditar cuatro expresiones sin calcular nada)**.** *(a)* $S_{ij} = A_{ik} B^k{}_j$. A la derecha, k aparece arriba y abajo: es mudo y se suma. Quedan libres i y j, ambos abajo, exactamente como a la izquierda. La ecuación es gramaticalmente correcta y define un tensor $(0,2)$.

(b) $v^i = A_{ij} w^j$. Está mal. A la izquierda hay un índice libre *arriba*; a la derecha, la i de A_{ij} está *abajo* y no está contraída con nada. La expresión correcta sería $v^i = A^i{}_j w^j$, o

bien $v_i = A_{ij}w^j$. Este error, muy común al principio, es precisamente lo que el convenio está diseñado para hacer visible.

(c) $T^i{}_i$. Los dos índices son mudos y el resultado es un escalar: es la *traza* del tensor. Nótese que T^{ij} no admite traza directa: hay que bajar antes un índice con la métrica, $g_{jk}T^{jk}$. Una traza solo tiene sentido entre un índice arriba y uno abajo.

(d) δ^i_i. Sumando sobre i de 1 a n se obtiene n, no 1. En relatividad, donde $n = 4$, este número aparece constantemente y confundirlo con 1 echa a perder páginas enteras de cálculo.

Ejemplo 36.21 (Por qué la posición del índice no es decorativa)**.** En coordenadas polares, un mismo objeto geométrico tiene componentes numéricamente distintas según se le mire como vector o como covector. Tomemos el campo v con componentes contravariantes $v^r = 1$, $v^\theta = 0$ (un vector unitario radial). Bajando índices con $g = \mathrm{diag}(1, r^2)$:

$$v_r = g_{rr}v^r = 1, \qquad v_\theta = g_{\theta\theta}v^\theta = r^2 \cdot 0 = 0.$$

Aquí coinciden. Pero tomemos ahora $v^r = 0$, $v^\theta = 1$:

$$v_r = 0, \qquad v_\theta = r^2.$$

El mismo objeto tiene componente 1 arriba y r^2 abajo. En cartesianas esto nunca pasa (porque $g = \mathrm{diag}(1, 1)$) y por eso en el cálculo elemental jamás se distinguen las dos versiones; en cuanto las coordenadas se curvan, la distinción es obligatoria y omitirla produce fórmulas que fallan por factores de r.

36.4. Tensores y sus operaciones

Idea clave

Intuición primero. Un tensor es una máquina multilineal: se come una cantidad fija de vectores y de covectores y devuelve un número, y es lineal en cada ranura por separado. Los objetos que ya conoces son casos de esto: un escalar es una máquina con cero ranuras, un covector tiene una, el producto interno es una máquina de dos vectores, una matriz es una máquina que come un vector y un covector.

Por qué merece la pena la definición abstracta: porque separa el *objeto* de sus *componentes*. Las componentes cambian al cambiar de base; el objeto no. Esa es toda la moraleja del capítulo, y es lo que hace que las leyes de la física escritas con tensores valgan en cualquier sistema de referencia sin reescribirlas.

Las operaciones se leen igual de fácil. El producto tensorial junta dos máquinas en una con más ranuras. La contracción tapa una ranura de entrada con una de salida y suma —la traza es exactamente eso—. Y el criterio del cociente es la herramienta práctica: si algo se come tensores y devuelve tensores, linealmente, entonces es un tensor; no hace falta verificar la ley de transformación índice por índice.

Definición 36.22. Un *tensor de tipo* (p, q) es una aplicación multilineal

$$T : \underbrace{V^* \times \cdots \times V^*}_{p} \times \underbrace{V \times \cdots \times V}_{q} \to \mathbb{R},$$

con componentes $T^{i_1\dots i_p}_{\ \ j_1\dots j_q} = T(e^{i_1},\dots,e_{j_1},\dots)$. Bajo cambio de base, cada índice superior se transforma con A^{-1} y cada inferior con A; esta *ley de transformación* es la definición clásica de tensor, equivalente a la multilineal.

Idea clave

Qué dice esto en cristiano. Un tensor no es una matriz con más índices: es una máquina multilineal. Come q vectores y p covectores, y devuelve un número, siendo lineal en cada entrada por separado. Los índices arriba y abajo llevan la contabilidad de qué come cada ranura, y esa contabilidad es lo que hace que la fórmula de cambio de base salga sola. Todo lo que ya se conoce es un tensor: un vector es de tipo $(1,0)$, un covector de tipo $(0,1)$, el producto escalar de tipo $(0,2)$ y una aplicación lineal de tipo $(1,1)$.

Ejemplo 36.23 (Objetos conocidos vistos como tensores)**. (a) El producto escalar.** Es una máquina que come dos vectores y devuelve un número, lineal en cada uno: tensor de tipo $(0,2)$. En coordenadas es g_{ij}, y en la base canónica de $\mathbb{R}^3$ vale $g_{ij}=\delta_{ij}$. Con otra base deja de ser la identidad: en la base $\{(1,0),(1,1)\}$ la matriz métrica es $\begin{pmatrix}1&1\\1&2\end{pmatrix}$.

(b) Una aplicación lineal. Come un vector y un covector y devuelve un número: $T(\omega,v)=\omega(Av)$. Tipo $(1,1)$, y en coordenadas es la matriz $A^i{}_j$ de siempre, con un índice arriba y otro abajo. Esa posición de los índices es la que explica por qué la matriz de una aplicación lineal cambia como $P^{-1}AP$ y no de otra manera.

(c) El determinante. En $\mathbb{R}^3$ es una máquina que come tres vectores y devuelve un número, lineal en cada uno y alternada: tipo $(0,3)$. Es el tensor de Levi-Civita ε_{ijk}, que vale 1, -1 o 0 según la permutación de los índices.

(d) Contar componentes. Un tensor de tipo (p,q) en dimensión n tiene n^{p+q} componentes. En dimensión 4, un tensor de tipo $(0,2)$ tiene 16; si además es simétrico, solo 10 independientes, que son exactamente las del tensor métrico de la relatividad general. El tensor de Riemann, de tipo $(1,3)$, tendría 256 componentes, y sus simetrías las reducen a 20.

Cuidado

[Un array de números no es un tensor] Lo que convierte a un conjunto de números en un tensor es *cómo cambia* al cambiar de base: cada índice arriba se transforma con P^{-1} y cada índice abajo con P. Una tabla de números que no obedezca esa ley no es un tensor, por muchos índices que tenga; los símbolos de Christoffel son el ejemplo clásico de objeto con tres índices que *no* es un tensor. Segundo aviso: la posición de los índices no es decoración tipográfica, distingue objetos genuinamente distintos, y en $\mathbb{R}^n$ con base ortonormal parece irrelevante solo porque ahí $P^{-1}=P^{\mathsf{T}}$.

Casos: $(0,0)$ escalar; $(1,0)$ vector; $(0,1)$ covector; $(1,1)$ transformación lineal; $(0,2)$ forma bilineal (como el producto interno); $(2,0)$ forma bilineal sobre covectores.

Operaciones:

- *Suma* y producto por escalares (mismo tipo).
- *Producto tensorial*: $(S \otimes T)$ concatena argumentos; componentes $S^i_j T^k_l$. El espacio de tensores (p, q) tiene dimensión n^{p+q} y base $e_{i_1} \otimes \cdots \otimes e^{j_q}$.
- *Contracción*: igualar un índice arriba con uno abajo y sumar: de (p, q) a $(p-1, q-1)$. La traza T^i_i es la contracción del $(1, 1)$.
- *Simetrización* $T_{(ij)} = \frac{1}{2}(T_{ij} + T_{ji})$ y *antisimetrización* $T_{[ij]} = \frac{1}{2}(T_{ij} - T_{ji})$. Un tensor $(0, 2)$ siempre se parte en simétrico + antisimétrico.

Observación 36.24 (Criterio del cociente). Si se sabe que $T^i_j v^j$ es vector para *todo* vector v, entonces T es un tensor $(1, 1)$. Este "criterio del cociente" es la manera práctica de reconocer tensores: los objetos que comen tensores y devuelven tensores, linealmente, son tensores.

Las *k-formas* del capítulo 11 son exactamente los tensores $(0, k)$ totalmente antisimétricos; el producto cuña es el producto tensorial antisimetrizado, y el determinante es (salvo constante) la única n-forma en dimensión n: por eso mide volúmenes.

Regla: decidir si un objeto con índices es de verdad un tensor

Escribir algo con índices no lo convierte en tensor. La pregunta siempre es una sola: *¿sus componentes se transforman con la ley correcta al cambiar de coordenadas?* En la práctica casi nunca se comprueba a mano; se usan estos atajos, en este orden.

1. **¿Está construido solo con operaciones tensoriales?** Sumar tensores del mismo tipo, multiplicarlos tensorialmente, contraer un índice arriba con uno abajo, simetrizar y antisimetrizar: todas ellas producen tensores. Si el objeto es una receta hecha únicamente con esas piezas, es tensor y no hay nada que verificar.
2. **¿Pasa el criterio del cociente?** Si $T^i{}_j v^j$ resulta ser un vector *para todo* vector v, entonces T es un tensor $(1, 1)$. Es el atajo más potente del capítulo: en vez de comprobar la ley de transformación índice por índice, se comprueba que el objeto *come tensores y devuelve tensores.*
3. **¿Aparece una derivada parcial suelta?** Bandera roja inmediata. $\partial_j v^i$ *no* es un tensor: al cambiar de coordenadas la regla de la cadena genera un término extra con segundas derivadas del cambio. La única derivada que sí es tensorial es la covariante, $\nabla_j v^i = \partial_j v^i + \Gamma^i_{jk} v^k$, y su misión es exactamente cancelar ese término extra.
4. **¿Es un objeto definido por una fórmula que menciona un sistema de coordenadas concreto?** Los símbolos de Christoffel son el ejemplo canónico: se definen a partir de las derivadas de la métrica *en unas coordenadas*, y en coordenadas cartesianas de $\mathbb{R}^n$ son todos nulos, mientras que en polares no. Un tensor que sea cero en un sistema es cero en todos; luego Γ no es un tensor.
5. **Contar los índices libres a ambos lados de la ecuación.** En toda igualdad tensorial correcta, los índices libres coinciden en nombre, número y posición (arriba o abajo) en los dos miembros, y los índices repetidos aparecen exactamente una vez arriba y una vez abajo. Un índice que aparezca tres veces, o repetido dos veces abajo, es siempre un error.

6. **Recordar la excepción útil: las diferencias sí son tensores.** Aunque Γ no sea tensor, la diferencia $\Gamma^i_{jk} - \tilde{\Gamma}^i_{jk}$ de dos conexiones *sí* lo es: el término anómalo es el mismo para las dos y se cancela al restar. La misma idea explica por qué el tensor de torsión $\Gamma^i_{jk} - \Gamma^i_{kj}$ es tensorial.

Ejemplo 36.25 (Cinco objetos, ¿tensor o no?)**. (a) La delta de Kronecker δ^i_j.** Sí, y además es el mismo en todas las coordenadas. Transformándola con la ley de un $(1,1)$:

$$\tilde{\delta}^a{}_b = \frac{\partial \tilde{x}^a}{\partial x^i}\frac{\partial x^j}{\partial \tilde{x}^b}\,\delta^i_j = \frac{\partial \tilde{x}^a}{\partial x^i}\frac{\partial x^i}{\partial \tilde{x}^b} = \frac{\partial \tilde{x}^a}{\partial \tilde{x}^b} = \delta^a_b,$$

donde el penúltimo paso es la regla de la cadena. La delta es el único tensor con esta propiedad de invariancia total, y por eso puede escribirse en cualquier fórmula sin especificar coordenadas.

(b) La derivada parcial de un campo escalar, $\partial_i f$. Sí, es un covector: $\tilde{\partial}_a f = \frac{\partial x^i}{\partial \tilde{x}^a}\partial_i f$, que es exactamente la ley de un índice abajo. Este es el caso feliz, y engaña: hace pensar que derivar siempre produce tensores.

(c) La derivada parcial de un campo vectorial, $\partial_j v^i$. *No.* Al transformar aparece

$$\tilde{\partial}_b \tilde{v}^a = \frac{\partial x^j}{\partial \tilde{x}^b}\frac{\partial \tilde{x}^a}{\partial x^i}\,\partial_j v^i \;+\; \underbrace{\frac{\partial^2 \tilde{x}^a}{\partial x^i \partial x^j}\frac{\partial x^j}{\partial \tilde{x}^b}\,v^i}_{\text{término indeseado}}.$$

El primer sumando es la ley correcta de un $(1,1)$; el segundo la estropea. Solo se anula cuando el cambio de coordenadas es afín, y por eso en cartesianas todo parece funcionar. Este término sobrante es la razón entera de que exista la derivada covariante.

(d) Los símbolos de Christoffel Γ^i_{jk}. *No*, por el paso 4: en cartesianas planas valen todos cero, y en coordenadas polares del mismo plano $\Gamma^r_{\theta\theta} = -r$ y $\Gamma^\theta_{r\theta} = 1/r$ no son cero. Un tensor nulo en una base es nulo en todas; luego Γ no lo es. Lo que mide Γ no es la curvatura del espacio, sino la curvatura de las *líneas coordenadas*.

(e) El tensor de curvatura $R^i{}_{jkl}$. Sí, y es notable: está construido con derivadas parciales de Γ y productos $\Gamma\Gamma$, ninguno de los cuales es tensor, pero la combinación concreta

$$R^i{}_{jkl} = \partial_k \Gamma^i_{lj} - \partial_l \Gamma^i_{kj} + \Gamma^i_{km}\Gamma^m_{lj} - \Gamma^i_{lm}\Gamma^m_{kj}$$

cancela todos los términos anómalos. Y esa es precisamente la razón de que la curvatura signifique algo físico: no se puede hacer desaparecer eligiendo bien las coordenadas.

Cuidado

En el plano euclídeo, con coordenadas polares, los símbolos de Christoffel son distintos de cero y sin embargo el espacio es *plano*. Confundir "$\Gamma \neq 0$" con "hay curvatura" es el malentendido más frecuente del capítulo. Lo que decide es $R^i{}_{jkl}$: en el plano vale cero en todas las coordenadas, mientras que en la esfera no vale cero en ninguna. Los Γ hablan del sistema de coordenadas; el R habla del espacio.

Regla: transformar un tensor concreto de un sistema de coordenadas a otro

La ley de transformación es la definición misma de tensor, pero cuando hay que aplicarla a números concretos conviene un procedimiento fijo: qué jacobiano va con qué índice, y en qué punto se evalúa todo.

1. **Escribe las dos jacobianas.** Si pasas de coordenadas x^i a $\tilde{x}^\mu$, necesitas ambas matrices,
$$J^\mu{}_i = \frac{\partial \tilde{x}^\mu}{\partial x^i}, \qquad (J^{-1})^i{}_\mu = \frac{\partial x^i}{\partial \tilde{x}^\mu},$$
y son inversas una de la otra. Casi todos los errores vienen de usar una donde tocaba la otra.

2. **Aplica la regla del arriba y el abajo.** Los índices *contravariantes* (arriba) se transforman con $\partial\tilde{x}/\partial x$; los *covariantes* (abajo) con $\partial x/\partial\tilde{x}$:
$$\tilde{V}^\mu = \frac{\partial \tilde{x}^\mu}{\partial x^i} V^i, \qquad \tilde{\omega}_\mu = \frac{\partial x^i}{\partial \tilde{x}^\mu}\, \omega_i.$$
La regla mnemotécnica: los índices contravariantes van "con" el cambio y los covariantes "en contra", que es de donde vienen los nombres.

3. **Para tensores de rango superior, un jacobiano por índice.** Un tensor de tipo (p, q) arrastra p copias de J y q copias de J^{-1}:
$$\tilde{T}^\mu{}_\nu = \frac{\partial \tilde{x}^\mu}{\partial x^i}\frac{\partial x^j}{\partial \tilde{x}^\nu}\, T^i{}_j, \qquad \tilde{g}_{\mu\nu} = \frac{\partial x^i}{\partial \tilde{x}^\mu}\frac{\partial x^j}{\partial \tilde{x}^\nu}\, g_{ij}.$$

4. **Evalúa todo en el mismo punto.** Las derivadas parciales son funciones del punto: hay que sustituir las coordenadas del punto concreto *antes* de multiplicar, o al menos ser coherente. Un tensor en p solo se relaciona con componentes en p.

5. **Deja los índices mudos en su sitio.** En cada término, el índice repetido aparece una vez arriba y otra abajo, y los índices libres son los mismos a ambos lados de la igualdad y en el mismo nivel. Si en $\tilde{T}^\mu{}_\nu$ te sobra o falta un índice, la fórmula está mal antes de calcular nada.

6. **Verifica con un escalar.** La contracción total no debe cambiar de valor: $\tilde{g}_{\mu\nu}\tilde{V}^\mu\tilde{V}^\nu = g_{ij}V^iV^j$, y la traza $\tilde{T}^\mu{}_\mu = T^i{}_i$. Es la comprobación más barata que existe y detecta casi cualquier error de jacobiano.

7. **Distingue componentes de objeto.** Las componentes cambian; el vector no. Escribir "$V^r = 0{,}707$" sin decir en qué base es tan incompleto como dar una longitud sin unidades. El objeto invariante es $V = V^i\partial_i = \tilde{V}^\mu\tilde{\partial}_\mu$.

Ejemplo 36.26 (Un mismo vector visto de cuatro maneras)**.** Trabajemos en el plano con coordenadas cartesianas (x, y) y polares (r, θ), ligadas por $x = r\cos\theta$, $y = r\sin\theta$. Las

jacobianas del paso 1 son

$$\frac{\partial(r,\theta)}{\partial(x,y)} = \begin{pmatrix} \cos\theta & \sin\theta \\ -\dfrac{\sin\theta}{r} & \dfrac{\cos\theta}{r} \end{pmatrix}, \qquad \frac{\partial(x,y)}{\partial(r,\theta)} = \begin{pmatrix} \cos\theta & -r\sin\theta \\ \sin\theta & r\cos\theta \end{pmatrix},$$

y en efecto su producto es la identidad. Fijemos el punto $p = (x, y) = (1, 1)$, donde $r = \sqrt{2} = 1{,}41421$ y $\theta = \pi/4$, de modo que $\cos\theta = \sin\theta = 0{,}707107$.

(a) Un vector contravariante. Sea V el vector que en cartesianas tiene componentes $V^x = 1$, $V^y = 0$ (apunta hacia el este, de longitud 1). Por el paso 2,

$$V^r = \cos\theta \cdot 1 + \sin\theta \cdot 0 = 0{,}707107, \qquad V^\theta = -\frac{\sin\theta}{r}\cdot 1 + \frac{\cos\theta}{r}\cdot 0 = -\frac{0{,}707107}{1{,}41421} = -0{,}5.$$

Verificación del paso 6 con la métrica polar $g = \mathrm{diag}(1, r^2) = \mathrm{diag}(1, 2)$:

$$g_{\mu\nu}V^\mu V^\nu = 1\cdot(0{,}707107)^2 + 2\cdot(-0{,}5)^2 = 0{,}5 + 0{,}5 = 1 = g_{ij}V^iV^j.$$

La longitud se conserva, como debía. Y el signo negativo de V^θ tiene un significado claro: moverse hacia el este desde el punto $(1, 1)$ hace *disminuir* el ángulo polar.

(b) El covector correspondiente. Sea $\omega = \mathrm{d}x$, con componentes $\omega_x = 1$, $\omega_y = 0$. Ahora hay que usar la *otra* jacobiana:

$$\omega_r = \frac{\partial x}{\partial r}\,\omega_x = \cos\theta = 0{,}707107, \qquad \omega_\theta = \frac{\partial x}{\partial\theta}\,\omega_x = -r\sin\theta = -1.$$

Nótese que $\omega_\theta = -1$ mientras que $V^\theta = -0{,}5$: aunque en cartesianas ambos objetos tenían las mismas componentes $(1, 0)$, en polares *no* coinciden. Justamente por eso hay dos tipos de índice. La comprobación con la métrica inversa $g^{\mu\nu} = \mathrm{diag}(1, 1/r^2) = \mathrm{diag}(1, 0{,}5)$:

$$g^{\mu\nu}\omega_\mu\omega_\nu = (0{,}707107)^2 + 0{,}5\cdot(-1)^2 = 0{,}5 + 0{,}5 = 1.$$

(c) La métrica misma, transformada desde cero. Partiendo de $g_{ij} = \delta_{ij}$ en cartesianas y aplicando el paso 3:

$$g_{rr} = \Big(\frac{\partial x}{\partial r}\Big)^2 + \Big(\frac{\partial y}{\partial r}\Big)^2 = \cos^2\theta + \sin^2\theta = 1,$$

$$g_{\theta\theta} = (-r\sin\theta)^2 + (r\cos\theta)^2 = r^2, \qquad g_{r\theta} = (\cos\theta)(-r\sin\theta) + (\sin\theta)(r\cos\theta) = 0.$$

De ahí $\mathrm{d}s^2 = \mathrm{d}r^2 + r^2\mathrm{d}\theta^2$, la métrica polar de siempre, obtenida sin dibujar ningún triángulo: solo aplicando la ley de transformación.

(d) Una transformación de Lorentz. En el espacio de Minkowski con $\eta = \mathrm{diag}(-1, 1)$ y un impulso de velocidad $\beta = 0{,}6$, se tiene $\gamma = (1-\beta^2)^{-1/2} = 1{,}25$ y

$$\Lambda = \begin{pmatrix} \gamma & -\gamma\beta \\ -\gamma\beta & \gamma \end{pmatrix} = \begin{pmatrix} 1{,}25 & -0{,}75 \\ -0{,}75 & 1{,}25 \end{pmatrix}.$$

El cuadrivector $U = (1, 0)$ (una partícula en reposo, con $c = 1$) se transforma en $\tilde{U} = (1{,}25,\ -0{,}75)$. Verificación por el paso 6:

$$\eta_{\mu\nu}\tilde{U}^\mu\tilde{U}^\nu = -(1{,}25)^2 + (0{,}75)^2 = -1{,}5625 + 0{,}5625 = -1 = \eta_{ij}U^iU^j.$$

El invariante -1 es la masa en reposo al cuadrado (con signo): lo mismo que veía el observador original. Toda la relatividad especial cabe en el paso 6.

Cuidado

La ley de transformación se aplica a las *componentes*, no al punto. Es muy fácil calcular V^θ con las jacobianas evaluadas en un punto y luego compararlo con un V^θ de otro punto; el resultado no significa nada. Además, las componentes polares de un campo vectorial constante en cartesianas *no* son constantes: V^r y V^θ del apartado (a) dependen de r y θ. Que las componentes varíen no significa que el campo varíe: eso es exactamente lo que la derivada covariante viene a arreglar.

Pausa para pensar

¿Por qué la ley de transformación se elige así y no de otra forma? Porque es la única que hace consistentes las contracciones. Si V^i se transformara con J y ω_i también con J, entonces $\omega_i V^i$ cambiaría de valor al cambiar de coordenadas y no serviría para nada físico. Al poner J arriba y J^{-1} abajo, cada contracción arrastra un $JJ^{-1} = \mathrm{Id}$ que se cancela, y el resultado es un número que todos los observadores comparten. Toda la maquinaria de índices existe para garantizar esa cancelación automáticamente, sin tener que comprobarla caso por caso.

36.5. La métrica

Idea clave

Intuición primero. El tensor métrico es un *campo de reglas y transportadores*: en cada punto del espacio hay un aparatito g_{ij} que dice cómo medir longitudes y ángulos de los vectores que viven *ahí*. En el plano con coordenadas cartesianas todos los aparatitos son idénticos (δ_{ij}) y uno olvida que existen; en una esfera, o en coordenadas polares, o cerca de una masa, los aparatitos varían de punto a punto, y *toda* la geometría —distancias, geodésicas, curvatura, y en relatividad el paso mismo del tiempo— es contabilidad hecha con estas reglas locales. Aprender geometría diferencial es aprender a calcular sin suponer jamás que las reglas de aquí sirven allá.

Definición 36.27. Un *tensor métrico* es un $(0,2)$ simétrico no degenerado g_{ij}. Define el producto escalar $\langle u, v\rangle = g_{ij}u^i v^j$, normas, ángulos, y la longitud de curvas mediante el *elemento de línea*

$$\mathrm{d}s^2 = g_{ij}\,\mathrm{d}x^i \mathrm{d}x^j, \qquad L = \int_a^b \sqrt{g_{ij}\dot{x}^i\dot{x}^j}\,\mathrm{d}t.$$

Si g es definida positiva la geometría es *riemanniana*; con signatura mixta (como en relatividad), *pseudo-riemanniana*.

Ejemplo 36.28. *Cartesianas en* $\mathbb{R}^2$: $g = \mathrm{diag}(1,1)$, $\mathrm{d}s^2 = \mathrm{d}x^2 + \mathrm{d}y^2$. *Polares*: $\mathrm{d}s^2 = \mathrm{d}r^2 + r^2\mathrm{d}\theta^2$, $g = \mathrm{diag}(1, r^2)$: la métrica ya no es constante, aunque el plano sea el mismo. *Esfera de radio* R (colatitud ϕ, longitud θ): $\mathrm{d}s^2 = R^2\mathrm{d}\phi^2 + R^2\sin^2\phi\,\mathrm{d}\theta^2$. *Minkowski* (relatividad especial): $\mathrm{d}s^2 = -c^2\mathrm{d}t^2 + \mathrm{d}x^2 + \mathrm{d}y^2 + \mathrm{d}z^2$; los intervalos con $\mathrm{d}s^2 < 0$ son temporales, > 0 espaciales, $= 0$ luminosos.

36.5.1. Subir y bajar índices

La métrica identifica V con V^*: a cada vector v^i le asocia el covector $v_i = g_{ij}v^j$, y la inversa g^{ij} (definida por $g^{ik}g_{kj} = \delta^i_j$) sube índices: $\omega^i = g^{ij}\omega_j$. Así, "el mismo" objeto físico tiene versiones covariante y contravariante; en cartesianas euclídeas coinciden numéricamente (por eso en cálculo elemental nunca se distinguen), pero en polares o en Minkowski no.

El *gradiente genuino* de una función es el covector $\partial_i f$; el "vector gradiente" es su versión subida $g^{ij}\partial_j f$. En polares: $(\nabla f)^r = f_r$ pero $(\nabla f)^\theta = \frac{1}{r^2} f_\theta$ —el factor que aparece "misteriosamente" en las fórmulas de análisis vectorial en coordenadas curvilíneas—.

El *elemento de volumen* invariante es $\sqrt{|\det g|}\,\mathrm{d}x^1 \cdots \mathrm{d}x^n$: en polares $\sqrt{\det g} = r$, en esféricas $\rho^2 \sin\phi$ —exactamente los jacobianos del capítulo 9, ahora explicados: son la raíz del determinante métrico—.

Regla: subir y bajar índices sin equivocarse

Una vez hay métrica, vectores y covectores dejan de ser mundos separados: g los traduce. Pero la traducción tiene reglas estrictas de contabilidad.

1. **Baja con $g_{\mu\nu}$, sube con $g^{\mu\nu}$.** Las dos operaciones básicas son

$$V_\mu = g_{\mu\nu}V^\nu, \qquad V^\mu = g^{\mu\nu}V_\nu,$$

   donde $g^{\mu\nu}$ es la matriz *inversa* de $g_{\mu\nu}$, definida por $g^{\mu\alpha}g_{\alpha\nu} = \delta^\mu_\nu$. Nunca es la matriz de las entradas invertidas una a una, salvo que g sea diagonal.

2. **Comprueba el balance de índices.** En toda ecuación correcta, cada índice libre aparece en el mismo sitio (arriba o abajo) y el mismo número de veces en los dos miembros. Un índice repetido arriba y abajo está sumado y desaparece del balance; dos índices repetidos en la misma posición es siempre un error.

3. **Usa un índice mudo distinto cada vez.** Al bajar dos índices de golpe, escribe $T_{\mu\nu} = g_{\mu\alpha}g_{\nu\beta}T^{\alpha\beta}$ con letras mudas diferentes. Reutilizar la misma letra produce una contracción que no querías.

4. **Respeta el orden de los índices.** Si T no es simétrico, $T^\mu{}_\nu$ y $T_\nu{}^\mu$ son objetos distintos. Por eso se deja el hueco vertical: marca de qué ranura vino cada índice.

5. **Recuerda que subir y bajar son inversas.** Bajar un índice y volver a subirlo devuelve el original, porque $g^{\mu\alpha}g_{\alpha\nu} = \delta^\mu_\nu$. Es un buen control de errores en cálculos largos.

6. **Contrae solo un índice arriba con uno abajo.** La contracción $T^\mu{}_\mu$ es un escalar; la expresión $T^{\mu\mu}$ no significa nada invariante. Si quieres la traza de un tensor con dos índices arriba, primero baja uno: $T^\mu{}_\mu = g_{\mu\nu}T^{\mu\nu}$.

7. **Controla con un escalar conocido.** El producto $V^\mu V_\mu = g_{\mu\nu}V^\mu V^\nu$ es la norma al cuadrado, un número que no depende de coordenadas. Calcúlalo en dos sistemas distintos: si sale distinto, hay un error de índices en algún paso.

Ejemplo 36.29 (Cuatro traducciones con métricas concretas)**. (a) Coordenadas polares en el plano.** La métrica es $\mathrm{d}s^2 = \mathrm{d}r^2 + r^2\mathrm{d}\theta^2$, es decir

$$g_{\mu\nu} = \begin{pmatrix} 1 & 0 \\ 0 & r^2 \end{pmatrix}, \qquad g^{\mu\nu} = \begin{pmatrix} 1 & 0 \\ 0 & 1/r^2 \end{pmatrix}.$$

Sea $V^\mu = (V^r, V^\theta) = (1, 1)$. Bajando índices: $V_r = g_{rr}V^r = 1$ y $V_\theta = g_{\theta\theta}V^\theta = r^2$. La componente angular cambia de valor, y esto no es un capricho: V^θ es una velocidad angular (radianes por unidad de parámetro) mientras que V_θ tiene dimensiones de longitud al cuadrado por radián. La norma es

$$V^\mu V_\mu = 1 \cdot 1 + 1 \cdot r^2 = 1 + r^2,$$

que es lo mismo que $\|V\|^2 = (\mathrm{d}r/\mathrm{d}t)^2 + r^2(\mathrm{d}\theta/\mathrm{d}t)^2$ del análisis vectorial ordinario. Coincide.

(b) La esfera de radio a. Con $\mathrm{d}s^2 = a^2\mathrm{d}\theta^2 + a^2\sin^2\theta\,\mathrm{d}\varphi^2$ se tiene $g^{\theta\theta} = 1/a^2$ y $g^{\varphi\varphi} = 1/(a^2\sin^2\theta)$. Si $W_\mu = (0, 1)$, entonces $W^\varphi = g^{\varphi\varphi}W_\varphi = \dfrac{1}{a^2\sin^2\theta}$, que *explota* cuando $\theta \to 0$. No es una singularidad de la esfera, sino de las coordenadas: en el polo, φ deja de estar definida.

(c) Minkowski y el cambio de signo. Con $\eta_{\mu\nu} = \mathrm{diag}(-1, 1, 1, 1)$ se tiene $\eta^{\mu\nu} = \mathrm{diag}(-1, 1, 1, 1)$, la misma matriz. Bajar índices aquí solo cambia el signo de la componente temporal: si $V^\mu = (V^0, \mathbf{V})$ entonces $V_\mu = (-V^0, \mathbf{V})$. Por eso

$$V^\mu V_\mu = -(V^0)^2 + |\mathbf{V}|^2,$$

y ese signo es toda la relatividad especial: los vectores con $V^\mu V_\mu < 0$ son temporales, los que dan > 0 espaciales, y los que dan exactamente 0 son los rayos de luz.

(d) Una traza mal calculada y su arreglo. Sea $T^{\mu\nu}$ un tensor simétrico en el plano polar, con $T^{rr} = 1$, $T^{\theta\theta} = 1$ y $T^{r\theta} = 0$. La tentación es escribir la traza como $T^{rr} + T^{\theta\theta} = 2$. Es incorrecto: hay que bajar un índice primero.

$$T^\mu{}_\mu = g_{\mu\nu}T^{\mu\nu} = g_{rr}T^{rr} + g_{\theta\theta}T^{\theta\theta} = 1 \cdot 1 + r^2 \cdot 1 = 1 + r^2.$$

El primer resultado dependía de la parametrización elegida; el segundo es un escalar auténtico. Para convencerse: en coordenadas cartesianas el mismo tensor tendría otras componentes, y solo la segunda cuenta daría el mismo número.

Cuidado

La expresión $g^{\mu\nu}$ es la inversa *matricial*, no el recíproco entrada por entrada. Coinciden solo cuando g es diagonal, que es el caso de casi todos los ejemplos de este capítulo y por eso el error pasa desapercibido tanto tiempo. En cuanto aparece un término cruzado, como en la métrica de Kerr, la diferencia es total.

Pausa para pensar

¿Por qué δ^μ_ν se escribe siempre con un índice arriba y otro abajo, y nunca como $\delta_{\mu\nu}$? Porque la delta con las dos posiciones mixtas es la identidad, y la identidad sí es un tensor: sus componentes valen 1 y 0 en *todo* sistema de coordenadas. En cambio $\delta_{\mu\nu}$ con

los dos índices abajo pretendería ser una métrica euclídea disfrazada, y sus componentes cambiarían al cambiar de coordenadas. La posición de los índices no es notación: es el enunciado de qué se transforma cómo.

36.6. Derivación covariante

Idea clave

Intuición primero. Comparar vectores en puntos distintos es como comparar precios en monedas distintas: antes de restar hay que *convertir*. En coordenadas curvas, parte del cambio aparente de un campo vectorial es ilusión contable (los ejes de referencia giraron bajo nuestros pies), y los símbolos de Christoffel son exactamente los *tipos de cambio* que corrigen esa ilusión. La derivada covariante es el cambio verdadero: lo que queda tras descontar el movimiento de los ejes. Prueba de fuego: en el plano, un campo constante tiene parciales no nulas en polares, pero derivada covariante cero: el campo no cambió; cambiaron los ejes.

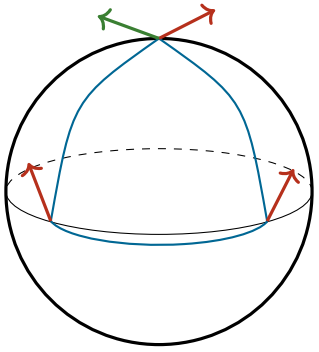


Figura 36.2: Transporte paralelo en la esfera: el vector rojo baja por un meridiano, viaja por el ecuador y regresa por otro meridiano sin girarlo jamás… y llega girado 90° (verde). Ese giro por circuito *es* la curvatura.

En coordenadas curvilíneas los vectores de la base cambian de punto a punto; derivar componentes a secas *no* da un tensor (las $\partial_i v^k$ de un vector no se transforman como tensor: aparece un término extra con derivadas de la matriz de cambio). La corrección:

Definición 36.30 (Símbolos de Christoffel)**.** Γ^k_{ij} se definen por $\partial_i e_j = \Gamma^k_{ij} e_k$ y se calculan solo con la métrica:

$$\Gamma^k_{ij} = \tfrac{1}{2} g^{kl}(\partial_i g_{jl} + \partial_j g_{il} - \partial_l g_{ij}). \tag{36.1}$$

Son simétricos en los índices inferiores ($\Gamma^k_{ij} = \Gamma^k_{ji}$) y **no** son tensores: pueden anularse en unas coordenadas y no en otras.

Definición 36.31 (Derivada covariante)**.**

$$\nabla_i v^k = \partial_i v^k + \Gamma^k_{ij} v^j, \qquad \nabla_i \omega_j = \partial_i \omega_j - \Gamma^k_{ij} \omega_k,$$

y en general un término $+\Gamma$ por cada índice arriba y $-\Gamma$ por cada índice abajo; sobre escalares, $\nabla_i f = \partial_i f$. El resultado *sí* es un tensor: los términos no tensoriales de ∂ y de Γ se cancelan exactamente. Propiedades: linealidad, regla de Leibniz, conmuta con contracciones y *compatibilidad métrica*: $\nabla_k g_{ij} = 0$ (la fórmula (36.1) está diseñada para esto; junto con la simetría, la determina de forma única: es la *conexión de Levi-Civita*).

Ejemplo 36.32 (Divergencia y laplaciano covariantes)**.** La divergencia de un campo es $\nabla_i v^i = \frac{1}{\sqrt{g}}\partial_i(\sqrt{g}\, v^i)$ (con $\sqrt{g} = \sqrt{|\det g|}$), y el laplaciano de un escalar:

$$\Delta f = \nabla_i(g^{ij}\partial_j f) = \frac{1}{\sqrt{g}}\partial_i(\sqrt{g}\, g^{ij}\partial_j f).$$

En polares ($\sqrt{g} = r$, $g^{rr} = 1$, $g^{\theta\theta} = r^{-2}$):

$$\Delta f = \frac{1}{r}\partial_r(r\, f_r) + \frac{1}{r^2} f_{\theta\theta} = f_{rr} + \frac{1}{r} f_r + \frac{1}{r^2} f_{\theta\theta},$$

la fórmula que en el capítulo 9 costó una página de regla de la cadena: el formalismo tensorial la produce en una línea, y en *cualquier* sistema de coordenadas.

Regla: calcular los símbolos de Christoffel de una métrica

Este es el cálculo central del capítulo y el que abre la puerta a geodésicas y curvatura. Parece temible por los tres índices, pero es mecánico.

1. **Escribe la métrica como matriz g_{ij} e inviértela**. Si g es diagonal —y en casi todos los ejemplos clásicos lo es— la inversa se obtiene invirtiendo cada entrada: $g^{ii} = 1/g_{ii}$, sin sumar.

2. **Anota qué componentes dependen de qué coordenadas.** Toda derivada $\partial_k g_{ij}$ con g_{ij} constante es cero, y eso elimina de un plumazo la mayoría de los términos. En polares, la única entrada no constante es $g_{\theta\theta} = r^2$, y solo depende de r: la única derivada no nula de toda la métrica es $\partial_r g_{\theta\theta} = 2r$.

3. **Aplica la fórmula $\Gamma^k_{ij} = \frac{1}{2}g^{kl}(\partial_i g_{jl} + \partial_j g_{il} - \partial_l g_{ij})$**, recorriendo k por fuera. Con métrica diagonal solo sobrevive $l = k$, así que

$$\Gamma^k_{ij} = \frac{1}{2g_{kk}}(\partial_i g_{jk} + \partial_j g_{ik} - \partial_k g_{ij}) \quad \text{(sin sumar en } k\text{)}.$$

4. **Usa la simetría para ahorrar la mitad del trabajo: $\Gamma^k_{ij} = \Gamma^k_{ji}$.** En dimensión 2 hay $2 \cdot 3 = 6$ símbolos independientes, no 8; en dimensión 4 son 40, no 64.

5. **Verifica con un caso conocido.** Si tu métrica se reduce a la euclídea en alguna región, allí todos los Γ deben anularse. Y si la métrica no depende de una coordenada x^a, ningún Γ puede contener una derivada respecto de x^a.

6. **No los llames tensores.** Los Γ pueden ser todos nulos en cartesianas y no nulos en polares *describiendo el mismo plano*. Un tensor nulo en unas coordenadas es nulo en todas; los símbolos de Christoffel no cumplen eso, y por eso su anulación no significa "espacio plano".

Ejemplo 36.33 (Christoffel del plano en coordenadas polares)**.** La métrica es $\mathrm{d}s^2 = \mathrm{d}r^2 + r^2\mathrm{d}\theta^2$, es decir $g = \mathrm{diag}(1, r^2)$ y $g^{-1} = \mathrm{diag}(1, r^{-2})$.

Paso 2. La única derivada no nula de la métrica es $\partial_r g_{\theta\theta} = 2r$. Todo símbolo que no involucre esa derivada se anula.

Paso 3, componente $k = r$. Como $g_{rr} = 1$,

$$\Gamma^r_{ij} = \tfrac{1}{2}(\partial_i g_{jr} + \partial_j g_{ir} - \partial_r g_{ij}).$$

Los dos primeros sumandos son siempre cero (la fila $g_{\cdot r}$ es constante). Queda $-\frac{1}{2}\partial_r g_{ij}$, que solo es no nulo para $i = j = \theta$:

$$\Gamma^r_{\theta\theta} = -\tfrac{1}{2}\,\partial_r g_{\theta\theta} = -\tfrac{1}{2}(2r) = -r.$$

Paso 3, componente $k = \theta$. Ahora $g_{\theta\theta} = r^2$, luego

$$\Gamma^\theta_{ij} = \frac{1}{2r^2}(\partial_i g_{j\theta} + \partial_j g_{i\theta} - \partial_\theta g_{ij}).$$

El último sumando siempre es cero (nada depende de θ). Para $i = r$, $j = \theta$:

$$\Gamma^\theta_{r\theta} = \frac{1}{2r^2}\,\partial_r g_{\theta\theta} = \frac{2r}{2r^2} = \frac{1}{r},$$

y por simetría $\Gamma^\theta_{\theta r} = \frac{1}{r}$ también. El resto se anula.

Resumen. Los únicos símbolos no nulos son

$$\Gamma^r_{\theta\theta} = -r, \qquad \Gamma^\theta_{r\theta} = \Gamma^\theta_{\theta r} = \frac{1}{r}.$$

Dos números para describir toda la "curvatura aparente" de las coordenadas polares en un plano que, recordemos, es completamente plano.

Ejemplo 36.34 (La prueba de fuego: un campo constante en polares)**.** Tomemos el campo vectorial más aburrido posible, el campo constante $v = \mathbf{i}$ que en cartesianas tiene componentes $(1, 0)$ en todo punto. Sus componentes polares no son constantes:

$$v^r = \cos\theta, \qquad v^\theta = -\frac{\sin\theta}{r}.$$

(Se obtienen imponiendo $v^r\partial_r + v^\theta\partial_\theta = \partial_x$, con $\partial_r = \cos\theta\,\partial_x + \sin\theta\,\partial_y$ y $\partial_\theta = -r\sin\theta\,\partial_x + r\cos\theta\,\partial_y$.)

Las derivadas parciales a secas **no** se anulan: por ejemplo $\partial_\theta v^r = -\sin\theta \neq 0$. Si nos quedáramos ahí concluiríamos, absurdamente, que un campo constante cambia. La derivada covariante corrige esa ilusión. Calculemos las cuatro componentes de $\nabla_i v^k = \partial_i v^k + \Gamma^k_{ij} v^j$:

$$\begin{aligned}
\nabla_r v^r &= \partial_r(\cos\theta) + 0 = 0,\\
\nabla_\theta v^r &= \partial_\theta(\cos\theta) + \Gamma^r_{\theta\theta} v^\theta = -\sin\theta + (-r)\Big(-\frac{\sin\theta}{r}\Big) = -\sin\theta + \sin\theta = 0,\\
\nabla_r v^\theta &= \partial_r\Big(-\frac{\sin\theta}{r}\Big) + \Gamma^\theta_{r\theta} v^\theta = \frac{\sin\theta}{r^2} + \frac{1}{r}\Big(-\frac{\sin\theta}{r}\Big) = 0,\\
\nabla_\theta v^\theta &= \partial_\theta\Big(-\frac{\sin\theta}{r}\Big) + \Gamma^\theta_{\theta r} v^r = -\frac{\cos\theta}{r} + \frac{1}{r}\cos\theta = 0.
\end{aligned}$$

Las cuatro se anulan. El campo no cambia: lo que cambiaba eran los ejes, y los símbolos de Christoffel han descontado exactamente ese movimiento, término a término. Este cálculo, hecho una vez con papel, convence más que cualquier explicación de qué es una derivada covariante y por qué hace falta.

Ejemplo 36.35 (Christoffel de la esfera, y una geodésica que sí y otra que no)**.** Para la esfera de radio R con colatitud ϕ y longitud θ, $g = \mathrm{diag}(R^2,\ R^2\sin^2\phi)$. La única derivada no nula es $\partial_\phi g_{\theta\theta} = 2R^2\sin\phi\cos\phi$, y repitiendo el procedimiento anterior se obtiene

$$\Gamma^{\phi}_{\theta\theta} = -\sin\phi\cos\phi, \qquad \Gamma^{\theta}_{\phi\theta} = \Gamma^{\theta}_{\theta\phi} = \cot\phi,$$

y todos los demás nulos. Obsérvese que R ha desaparecido: los símbolos no dependen del radio, aunque la curvatura sí.

¿Es el ecuador una geodésica? El ecuador es $\phi = \pi/2$ constante, con θ creciendo uniformemente: $\dot\phi = 0$, $\ddot\phi = 0$, $\dot\theta = \omega$, $\ddot\theta = 0$. La ecuación geodésica $\ddot x^k + \Gamma^k_{ij}\dot x^i\dot x^j = 0$ da, para $k = \phi$:

$$0 + \Gamma^{\phi}_{\theta\theta}\dot\theta^2 = -\sin\tfrac{\pi}{2}\cos\tfrac{\pi}{2}\,\omega^2 = -1\cdot 0\cdot\omega^2 = 0. \ \checkmark$$

Y para $k = \theta$: $0 + 2\Gamma^{\theta}_{\phi\theta}\dot\phi\dot\theta = 0$ porque $\dot\phi = 0$. Las dos ecuaciones se cumplen: el ecuador es una geodésica.

¿Y un paralelo cualquiera? Repitamos con $\phi = \phi_0$ fijo, distinto de $\pi/2$. La ecuación para $k = \phi$ pide ahora

$$-\sin\phi_0\cos\phi_0\,\omega^2 = 0,$$

lo que obliga a $\cos\phi_0 = 0$, es decir $\phi_0 = \pi/2$. Ningún otro paralelo es geodésica. Esto explica por qué las rutas aéreas entre dos ciudades de la misma latitud no siguen el paralelo sino que se desvían hacia el polo: el paralelo no es el camino más corto, el círculo máximo sí. Un vuelo Madrid–Nueva York, ambos cerca de los 40° de latitud, pasa notablemente más al norte que cualquiera de las dos ciudades.

Regla: derivar covariantemente un tensor de cualquier tipo

Una sola fórmula cubre todos los casos, y se recuerda con una frase: *una derivada parcial, más un Γ por cada índice, con signo $+$ si el índice está arriba y $-$ si está abajo.*

1. **Escribir el término parcial.** ∇_k(lo que sea) empieza siempre por ∂_k aplicado a las componentes.
2. **Añadir un Γ por índice contravariante (arriba), con signo $+$.** Para v^i: $+\Gamma^i_{km}v^m$. El índice de arriba del tensor sube al Γ; el índice de derivación k va abajo en el Γ; el índice mudo m sustituye al original en el tensor.
3. **Añadir un Γ por índice covariante (abajo), con signo $-$.** Para ω_i: $-\Gamma^m_{ki}\omega_m$. Ahora el mudo sube al Γ y el índice original baja.
4. **Ensamblar en el caso general.** Para un (p,q) hay p términos con $+$ y q con $-$. Por ejemplo, para un $(1,1)$:

$$\nabla_k T^i{}_j = \partial_k T^i{}_j + \Gamma^i_{km}T^m{}_j - \Gamma^m_{kj}T^i{}_m.$$

5. **Comprobar los índices libres.** El resultado de ∇_k sobre un (p, q) es un $(p, q+1)$: tiene un índice abajo más, el k. Cada sumando debe tener exactamente los mismos índices libres, en las mismas posiciones. Es el control que atrapa el 90 % de los errores de signo y de colocación.

6. **Usar la compatibilidad métrica como comodín.** $\nabla_k g_{ij} = 0$ y $\nabla_k g^{ij} = 0$ siempre (con la conexión de Levi-Civita). Esto significa que *subir y bajar índices conmuta con derivar covariantemente*: si $v_i = g_{ij} v^j$, entonces $\nabla_k v_i = g_{ij} \nabla_k v^j$. Es lo que permite mover los índices dentro de las ecuaciones de la física sin miedo.

7. **Atajos que ahorran mucho trabajo.** Sobre escalares, $\nabla_k f = \partial_k f$, sin ningún Γ. La divergencia tiene fórmula cerrada, $\nabla_i v^i = \frac{1}{\sqrt{|g|}} \partial_i \big(\sqrt{|g|}\, v^i\big)$, y el rotacional de un covector también, $\nabla_i \omega_j - \nabla_j \omega_i = \partial_i \omega_j - \partial_j \omega_i$: los Γ se cancelan por ser simétricos en los índices de abajo.

Ejemplo 36.36 (La derivada covariante en la esfera)**.** Tomemos la esfera de radio a con la métrica $g = \operatorname{diag}(a^2,\ a^2 \sin^2\theta)$ en coordenadas (θ, φ). Sus símbolos de Christoffel no nulos son

$$\Gamma^{\theta}_{\varphi\varphi} = -\sin\theta\cos\theta, \qquad \Gamma^{\varphi}_{\theta\varphi} = \Gamma^{\varphi}_{\varphi\theta} = \cot\theta.$$

(El radio a ha desaparecido: los Γ solo dependen de la métrica salvo factor constante global.)

(a) Un campo vectorial. Sea $v = (v^\theta, v^\varphi) = (0, 1)$, es decir, el campo que apunta siempre hacia el este con componente φ constante. Sus derivadas parciales son todas nulas, y sin embargo

$$\nabla_\varphi v^\theta = \partial_\varphi v^\theta + \Gamma^{\theta}_{\varphi\varphi} v^\varphi = 0 - \sin\theta\cos\theta = -\sin\theta\cos\theta \neq 0.$$

El campo tiene componentes constantes pero *no* es paralelo: al desplazarse hacia el este sobre un paralelo, el vector se ve arrastrado hacia el ecuador. Ese $-\sin\theta\cos\theta$ es exactamente lo que la parcial no veía.

(b) El comodín métrico, comprobado. Verifiquemos $\nabla_\theta g_{\varphi\varphi} = 0$ a mano:

$$\nabla_\theta g_{\varphi\varphi} = \partial_\theta\big(a^2 \sin^2\theta\big) - 2\,\Gamma^{m}_{\theta\varphi} g_{m\varphi} = 2a^2 \sin\theta\cos\theta - 2\cot\theta \cdot a^2 \sin^2\theta = 0.$$

Los dos términos se cancelan exactamente. Nótese el 2: hay *dos* índices abajo, luego dos términos con signo $-$, y aquí ambos coinciden por simetría.

(c) La divergencia por la fórmula corta. Aquí $\sqrt{|g|} = a^2 \sin\theta$, así que para cualquier campo

$$\nabla_i v^i = \frac{1}{a^2 \sin\theta}\Big[\partial_\theta\big(a^2 \sin\theta\, v^\theta\big) + \partial_\varphi\big(a^2 \sin\theta\, v^\varphi\big)\Big] = \partial_\theta v^\theta + \cot\theta\, v^\theta + \partial_\varphi v^\varphi.$$

El mismo resultado sale sumando $\nabla_\theta v^\theta + \nabla_\varphi v^\varphi$ con la regla larga, pero con la fórmula corta no hace falta ningún Γ. Para el campo del apartado (a), $v = (0, 1)$, la divergencia es cero: el flujo hacia el este no acumula nada.

Pausa para pensar

Merece la pena entender *por qué* el signo es + arriba y − abajo, y no al revés. La razón es que la contracción de un vector con un covector, $\omega_i v^i$, es un escalar, y sobre escalares la derivada covariante debe reducirse a la parcial. Al aplicar la regla de Leibniz,

$$\nabla_k(\omega_i v^i) = (\partial_k \omega_i) v^i + \omega_i (\partial_k v^i) + (\text{términos con } \Gamma),$$

y para que los términos con Γ se anulen entre sí es obligatorio que uno lleve $+\Gamma^i_{km}$ y el otro $-\Gamma^m_{ki}$: son el mismo término con los índices renombrados y signos opuestos. El signo no es un convenio arbitrario, sino la única elección compatible con que derivar un número sea derivar un número.

36.6.1. Transporte paralelo y geodésicas

Definición 36.37. Un vector $v^k(t)$ se *transporta paralelamente* a lo largo de la curva $x^i(t)$ si su derivada covariante a lo largo de la curva es nula:

$$\frac{Dv^k}{\mathrm{d}t} \equiv \dot{v}^k + \Gamma^k_{ij} \dot{x}^i v^j = 0.$$

Una *geodésica* es una curva que transporta paralelamente su propio tangente:

$$\ddot{x}^k + \Gamma^k_{ij} \dot{x}^i \dot{x}^j = 0.$$

Las geodésicas son también las curvas de longitud estacionaria (se obtienen minimizando $\int \mathrm{d}s$ con cálculo de variaciones): rectas en el plano, círculos máximos en la esfera, y las trayectorias de caída libre en relatividad general.

Ejemplo 36.38 (Holonomía en la esfera)**.** Transporta un vector tangente a la esfera a lo largo de este circuito: del polo norte al ecuador por un meridiano, un cuarto de vuelta por el ecuador, y de regreso al polo por otro meridiano. Al volver, el vector llega *girado* 90° respecto del inicial, ¡sin que nunca lo hayamos girado! El ángulo de giro es proporcional al área encerrada ($\frac{4\pi R^2}{8} = \frac{\pi R^2}{2}$, y el ángulo es área$/R^2 = \frac{\pi}{2}$). Ese giro —la *holonomía*— es la manifestación operativa de la curvatura: en el plano no ocurre.

Regla: escribir y resolver la ecuación de las geodésicas

Una geodésica es la curva que un objeto libre recorre: la trayectoria «lo más recta posible» que permite la métrica. Su ecuación se monta siempre igual.

1. **Parte de los Christoffel.** Con $\Gamma^\lambda_{\mu\nu}$ ya calculados a partir de la métrica, la ecuación es
$$\frac{\mathrm{d}^2 x^\lambda}{\mathrm{d}s^2} + \Gamma^\lambda_{\mu\nu} \frac{\mathrm{d}x^\mu}{\mathrm{d}s} \frac{\mathrm{d}x^\nu}{\mathrm{d}s} = 0,$$
una ecuación por cada coordenada. El parámetro s debe ser *afín*: longitud de arco, tiempo propio, o cualquier función afín de ellos.

2. **Atajo por el lagrangiano.** Suele ser mucho más rápido escribir $L = g_{\mu\nu}\dot{x}^\mu\dot{x}^\nu$ y aplicar Euler–Lagrange
$$\frac{\mathrm{d}}{\mathrm{d}s}\frac{\partial L}{\partial \dot{x}^\lambda} - \frac{\partial L}{\partial x^\lambda} = 0.$$
Sale el mismo sistema y, de paso, los Christoffel se leen de los coeficientes. Este es el método que conviene en un examen.

3. **Busca coordenadas cíclicas.** Si L no depende de una coordenada x^λ, entonces $\partial L/\partial \dot{x}^\lambda$ es constante a lo largo de la geodésica. Cada coordenada cíclica es una cantidad conservada y baja en uno el orden del sistema.

4. **Usa la normalización como primera integral.** La cantidad $g_{\mu\nu}\dot{x}^\mu\dot{x}^\nu$ es constante: vale 1 si parametrizas por longitud de arco, -1 (o $-c^2$) para una partícula masiva en signatura lorentziana, y 0 para la luz. Esta relación sustituye a una de las ecuaciones de segundo orden.

5. **Resuelve el sistema reducido.** Con las conservadas del paso 3 y la normalización del paso 4, lo que queda suele ser una sola ecuación de primer orden, integrable por separación de variables.

6. **Verifica sustituyendo.** Mete la solución candidata en todas las ecuaciones geodésicas, no solo en una. Es la comprobación más barata y la que más errores detecta.

7. **Interpreta el resultado.** Comprueba que las geodésicas obvias salen: rectas en el plano, círculos máximos en la esfera. Si una curva que *debería* ser geodésica no satisface la ecuación, revisa los Christoffel.

Ejemplo 36.39 (Geodésicas de tres superficies)**. (a) El plano en polares: las rectas disfrazadas.** Con $\mathrm{d}s^2 = \mathrm{d}r^2 + r^2\mathrm{d}\theta^2$, los únicos Christoffel no nulos son

$$\Gamma^r_{\theta\theta} = -r, \qquad \Gamma^\theta_{r\theta} = \Gamma^\theta_{\theta r} = \frac{1}{r},$$

y el sistema queda

$$\ddot{r} - r\dot{\theta}^2 = 0, \qquad \ddot{\theta} + \frac{2}{r}\dot{r}\dot{\theta} = 0.$$

La segunda ecuación dice que $r^2\dot{\theta}$ es constante: es la conservación del momento angular, y aparece aquí como consecuencia de que θ es cíclica. La primera es la ecuación radial con el término centrífugo. Su solución general es $r\cos(\theta - \theta_0) = d$, que en cartesianas es la recta $x\cos\theta_0 + y\sin\theta_0 = d$. Las geodésicas del plano son rectas, como debían ser: los Christoffel no nulos no indican curvatura, solo que las coordenadas son curvilíneas.

(b) La esfera: los círculos máximos. Con $\mathrm{d}s^2 = a^2\mathrm{d}\theta^2 + a^2\sin^2\theta\,\mathrm{d}\varphi^2$ los Christoffel son

$$\Gamma^\theta_{\varphi\varphi} = -\sin\theta\cos\theta, \qquad \Gamma^\varphi_{\theta\varphi} = \cot\theta,$$

y las ecuaciones

$$\ddot{\theta} - \sin\theta\cos\theta\,\dot{\varphi}^2 = 0, \qquad \ddot{\varphi} + 2\cot\theta\,\dot{\theta}\dot{\varphi} = 0.$$

Verifiquemos dos candidatas. Un *meridiano*, $\theta = s/a$ con φ constante: entonces $\ddot\theta = 0$ y $\dot\varphi = 0$, luego ambas ecuaciones se satisfacen. Es geodésica. El *ecuador*, $\theta = \pi/2$ y $\varphi = s/a$: aquí $\ddot\theta = 0$ y $\sin\theta\cos\theta = \cos(\pi/2) = 0$, de nuevo cero; y $\cot(\pi/2) = 0$ mata la segunda. También es geodésica.

(c) Un paralelo que no es geodésica. Tomemos ahora $\theta = \theta_0$ constante con $\theta_0 \neq \pi/2$, recorrido a rapidez unidad, o sea $\dot\varphi = 1/(a\sin\theta_0)$. La primera ecuación exige

$$0 - \sin\theta_0\cos\theta_0 \cdot \frac{1}{a^2\sin^2\theta_0} = -\frac{1}{a^2\tan\theta_0},$$

que solo es cero si $\tan\theta_0 = \infty$, es decir $\theta_0 = \pi/2$. Conclusión: de todos los paralelos, únicamente el ecuador es geodésica. Por eso las rutas aéreas entre dos ciudades de la misma latitud se desvían hacia el polo: el paralelo es más largo que el círculo máximo. La longitud del paralelo es $2\pi a\sin\theta_0$, siempre menor que la circunferencia máxima $2\pi a$ pero mayor, como camino entre dos de sus puntos, que el arco de círculo máximo correspondiente.

(d) El plano hiperbólico: semicircunferencias. En el semiplano superior con $\mathrm{d}s^2 = \dfrac{\mathrm{d}x^2 + \mathrm{d}y^2}{y^2}$, los Christoffel son

$$\Gamma^x_{xy} = -\frac{1}{y}, \qquad \Gamma^y_{xx} = \frac{1}{y}, \qquad \Gamma^y_{yy} = -\frac{1}{y},$$

y las ecuaciones geodésicas quedan

$$\ddot x - \frac{2}{y}\dot x\dot y = 0, \qquad \ddot y + \frac{\dot x^2 - \dot y^2}{y} = 0.$$

Propongamos la curva $x(s) = R\tanh s$, $y(s) = R\operatorname{sech} s$. Primero comprobemos que está parametrizada por longitud de arco: $\dot x = R\operatorname{sech}^2 s$, $\dot y = -R\operatorname{sech} s\tanh s$, y

$$\frac{\dot x^2 + \dot y^2}{y^2} = \frac{R^2\operatorname{sech}^2 s(\operatorname{sech}^2 s + \tanh^2 s)}{R^2\operatorname{sech}^2 s} = 1.$$

Sustituyendo en las dos ecuaciones geodésicas, ambas se anulan idénticamente. Y como $x^2 + y^2 = R^2(\tanh^2 s + \operatorname{sech}^2 s) = R^2$, la curva es una semicircunferencia centrada en el eje real. Las rectas verticales también lo son: x constante y $y = e^s$ satisface $\ddot y - \dot y^2/y = e^s - e^s = 0$.

Como aplicación, la distancia hiperbólica entre i y $2i$ se calcula sobre la geodésica vertical:

$$d = \int_1^2 \frac{\mathrm{d}y}{y} = \ln 2 \approx 0{,}6931.$$

Idea clave

Las tres geometrías clásicas se distinguen justo por sus geodésicas y por cuántas paralelas admiten. En el plano euclídeo, por un punto exterior a una recta pasa exactamente una paralela; en la esfera, ninguna, porque dos círculos máximos siempre se cortan; en el plano hiperbólico, infinitas. Durante dos milenios se intentó *demostrar* el quinto postulado de Euclides. La respuesta llegó al construir modelos como el del apartado (d), donde todos los demás postulados valen y el quinto falla: el postulado era independiente,

no deducible.

Cuidado

Que los Christoffel no sean nulos no significa que el espacio sea curvo. En (a) el plano es plano y sin embargo $\Gamma^r_{\theta\theta} = -r \neq 0$. Los Christoffel dependen de las coordenadas y siempre pueden anularse en un punto dado eligiendo coordenadas normales. Lo que no se puede anular, si de verdad hay curvatura, es el tensor de Riemann: ese es el criterio, y es el tema de la sección siguiente.

36.7. Curvatura

Idea clave

Intuición primero. La curvatura mide una sola cosa: si transportas un vector paralelamente alrededor de un circuito cerrado y vuelve girado, el espacio es curvo. Hazlo en una hoja de papel y el vector regresa idéntico; hazlo sobre una esfera —baja del polo al ecuador, avanza un cuarto de vuelta, sube— y regresa girado 90°. El tensor de Riemann es esa discrepancia por unidad de área.

De ahí sale su definición formal sin misterio: R mide cuánto *fallan en conmutar* dos derivadas covariantes. En un espacio plano derivar en una dirección y luego en otra da lo mismo en cualquier orden; en uno curvo, no. Por eso el teorema de abajo —$R \equiv 0$ si y solo si existen coordenadas donde la métrica es constante— dice justo lo que uno esperaría: la curvatura es la obstrucción exacta a poder aplanar.

Y ojo con la trampa habitual: los símbolos de Christoffel Γ pueden ser no nulos en un espacio perfectamente plano —el plano en coordenadas polares es el ejemplo—, porque miden cómo se tuercen las *coordenadas*, no el espacio. Solo R distingue de verdad la esfera del plano, y ninguna elección de coordenadas puede hacerlo desaparecer.

Definición 36.40 (Tensor de Riemann)**.**

$$(\nabla_i \nabla_j - \nabla_j \nabla_i)\, v^k = R^k{}_{lij}\, v^l, \qquad R^k{}_{lij} = \partial_i \Gamma^k_{jl} - \partial_j \Gamma^k_{il} + \Gamma^k_{im}\Gamma^m_{jl} - \Gamma^k_{jm}\Gamma^m_{il}.$$

Mide el fallo de conmutación de las derivadas covariantes, o equivalentemente la holonomía por unidad de área. Simetrías (con todos los índices abajo, $R_{klij} = g_{km} R^m{}_{lij}$): antisimétrico en (k,l) y en (i,j), simétrico al intercambiar los pares, y la identidad de Bianchi $R_{k[lij]} = 0$. En dimensión n tiene $\frac{n^2(n^2-1)}{12}$ componentes independientes: 1 en dimensión 2, 6 en dimensión 3, 20 en dimensión 4 (el espaciotiempo).

Definición 36.41. Contracciones: *tensor de Ricci* $R_{lj} = R^i{}_{lij}$ (simétrico) y *escalar de curvatura* $R = g^{lj} R_{lj}$. En dimensión 2, toda la curvatura se reduce al escalar: $R = 2K$, donde K es la curvatura de Gauss.

Teorema 36.42. *$R^k{}_{lij} \equiv 0$ en una región $\iff$ existen coordenadas en las que g_{ij} es constante (espacio* plano*). La curvatura es la obstrucción exacta a aplanar: el plano en polares tiene*

$\Gamma \neq 0$ *pero* $R = 0$*; la esfera tiene* $R \neq 0$ *en cualquier coordenada.*

Ejemplo 36.43 (La curvatura de la esfera, cálculo completo)**.** Métrica: $g_{\phi\phi} = R_0^2$, $g_{\theta\theta} = R_0^2 \sin^2\phi$ (radio R_0). Christoffel no nulos (de (36.1)):

$$\Gamma^{\phi}_{\theta\theta} = -\sin\phi\cos\phi, \qquad \Gamma^{\theta}_{\phi\theta} = \Gamma^{\theta}_{\theta\phi} = \cot\phi.$$

Componente clave del Riemann:

$$\begin{aligned} R^{\phi}{}_{\theta\phi\theta} &= \partial_\phi\Gamma^{\phi}_{\theta\theta} - \partial_\theta\Gamma^{\phi}_{\phi\theta} + \Gamma^{\phi}_{\phi m}\Gamma^{m}_{\theta\theta} - \Gamma^{\phi}_{\theta m}\Gamma^{m}_{\phi\theta} \\ &= \partial_\phi(-\sin\phi\cos\phi) - 0 + 0 - \Gamma^{\phi}_{\theta\theta}\Gamma^{\theta}_{\phi\theta} \\ &= (\sin^2\phi - \cos^2\phi) + \sin\phi\cos\phi\cdot\cot\phi = \sin^2\phi. \end{aligned}$$

Ricci: $R_{\theta\theta} = R^{\phi}{}_{\theta\phi\theta} = \sin^2\phi$ y $R_{\phi\phi} = 1$ (cálculo análogo). Escalar:

$$R = g^{\phi\phi}R_{\phi\phi} + g^{\theta\theta}R_{\theta\theta} = \frac{1}{R_0^2} + \frac{\sin^2\phi}{R_0^2\sin^2\phi} = \frac{2}{R_0^2}.$$

Curvatura de Gauss $K = \frac{R}{2} = \frac{1}{R_0^2}$: constante y positiva, decreciente con el radio —una esfera enorme parece localmente plana, que es exactamente nuestra experiencia en la Tierra—.

Regla: calcular la curvatura de una métrica, de principio a fin

Este es el cálculo largo del capítulo. Hacerlo con orden es la diferencia entre veinte minutos y una tarde perdida.

1. **Escribir g_{ij} y g^{ij}.** Si la métrica es diagonal —y en los ejercicios casi siempre lo es—, la inversa es la diagonal de los recíprocos, y eso solo ya ahorra la mitad del trabajo.
2. **Calcular los Γ^{i}_{jk} y anotar solo los no nulos.** Recuérdese la simetría $\Gamma^{i}_{jk} = \Gamma^{i}_{kj}$: en dimensión 2 hay a lo sumo 6 independientes, no 8; en dimensión 4, 40 y no 64. Antes de seguir conviene hacer una lista limpia de los supervivientes; todo lo demás es cero y no volverá a aparecer.
3. **Montar el Riemann con la fórmula:**
$$R^{i}{}_{jkl} = \partial_k\Gamma^{i}_{lj} - \partial_l\Gamma^{i}_{kj} + \Gamma^{i}_{km}\Gamma^{m}_{lj} - \Gamma^{i}_{lm}\Gamma^{m}_{kj}.$$
Regla mnemotécnica del orden de índices: el primer término deriva respecto del *tercer* índice y el Γ lleva el cuarto y el segundo; el segundo término es el mismo con k y l intercambiados. La antisimetría en $k \leftrightarrow l$ está incorporada en la fórmula.
4. **Aprovechar las simetrías para no calcular de más.** Con todos los índices abajo, $R_{ijkl} = -R_{jikl} = -R_{ijlk} = R_{klij}$. En dimensión 2 eso deja *una sola* componente independiente, R_{1212}; en dimensión 3, seis; en dimensión 4, veinte. Calcular las n^4 es siempre un desperdicio.
5. **Contraer para obtener Ricci:** $R_{jl} = R^{i}{}_{jil}$, sumando sobre el primer índice y el tercero. Es simétrico, $R_{jl} = R_{lj}$, lo cual sirve de control.

6. **Contraer otra vez para el escalar:** $R = g^{jl}R_{jl}$. Es un número (o una función), y en dimensión 2 se relaciona con la curvatura de Gauss por $R = 2K$.

7. **Controles.** Si el espacio es plano, todo el Riemann debe salir cero, no solo el escalar. Si la métrica tiene una simetría evidente (rotacional, por ejemplo), el resultado debe respetarla. Y en las métricas de curvatura constante debe salir $R_{ijkl} = K(g_{ik}g_{jl} - g_{il}g_{jk})$: es la comprobación más rápida de todas.

Ejemplo 36.44 (Tres superficies y sus curvaturas)**. (a) La esfera de radio** a**.** Con $g = \operatorname{diag}(a^2, a^2\sin^2\theta)$ y los Christoffel $\Gamma^{\theta}_{\varphi\varphi} = -\sin\theta\cos\theta$, $\Gamma^{\varphi}_{\theta\varphi} = \cot\theta$, la única componente independiente del Riemann es

$$R^{\theta}{}_{\varphi\theta\varphi} = \partial_\theta\Gamma^{\theta}_{\varphi\varphi} - \Gamma^{\theta}_{\varphi\varphi}\Gamma^{\varphi}_{\theta\varphi} = -\cos 2\theta + \sin\theta\cos\theta\cot\theta = \sin^2\theta.$$

De ahí $R_{\theta\theta} = 1$ y $R_{\varphi\varphi} = \sin^2\theta$, y contrayendo con g^{ij},

$$R = \frac{1}{a^2}\cdot 1 + \frac{1}{a^2\sin^2\theta}\cdot\sin^2\theta = \frac{2}{a^2}.$$

Luego $K = R/2 = 1/a^2$: constante y positiva, y tanto más pequeña cuanto mayor es la esfera. Una esfera muy grande es localmente casi plana —que es la razón por la que la Tierra parece llana— y el resultado tiene las unidades correctas, uno partido por longitud al cuadrado. Control final: $R_{\theta\varphi} = 0$, como exige la simetría de Ricci junto con la simetría rotacional.

(b) El plano hiperbólico. Con el modelo del semiplano superior, $\mathrm{d}s^2 = \frac{\mathrm{d}x^2+\mathrm{d}y^2}{y^2}$, los Christoffel no nulos son

$$\Gamma^{x}_{xy} = \Gamma^{x}_{yx} = -\frac{1}{y}, \qquad \Gamma^{y}_{xx} = \frac{1}{y}, \qquad \Gamma^{y}_{yy} = -\frac{1}{y},$$

y el cálculo da $R_{xx} = R_{yy} = -1/y^2$ (con $R_{xy} = 0$), es decir $R_{ij} = -g_{ij}$, con

$$R = g^{xx}R_{xx} + g^{yy}R_{yy} = y^2\Big(-\frac{1}{y^2}\Big) + y^2\Big(-\frac{1}{y^2}\Big) = -2.$$

Luego $K = -1$ en *todos* los puntos, aunque la métrica dependa de y: la aparente falta de homogeneidad es solo del sistema de coordenadas, no de la geometría. Este es el ejemplo que hay que tener en la cabeza cuando se dice que la curvatura es intrínseca.

(c) Una familia entera de golpe. Sea $\mathrm{d}s^2 = \mathrm{d}r^2 + f(r)^2\mathrm{d}\varphi^2$, la forma general de una superficie de revolución en coordenadas geodésicas. El cálculo da

$$R = -\frac{2f''(r)}{f(r)}, \qquad \text{es decir} \qquad K = -\frac{f''(r)}{f(r)}.$$

Una sola fórmula que contiene tres casos: con $f(r) = r$ sale $K = 0$ (el plano en polares), con $f(r) = a\sin(r/a)$ sale $K = 1/a^2$ (la esfera, midiendo r desde el polo), y con $f(r) = \sinh r$ sale $K = -1$ (el plano hiperbólico). La ecuación de Jacobi que gobierna la separación de geodésicas es exactamente $f'' + Kf = 0$; que las geodésicas se junten, sigan paralelas o se separen es el mismo hecho que el signo de K.

Cuidado

En dimensión 2 el tensor de Ricci contiene toda la información: $R_{ij} = \frac{R}{2} g_{ij}$ siempre, y no hay nada más que saber. En dimensión 3, Ricci determina todavía el Riemann completo. Pero a partir de dimensión 4 deja de ser cierto: hay curvatura que Ricci no ve, recogida en el tensor de Weyl. Por eso la ecuación de Einstein en el vacío, $R_{ij} = 0$, *no* obliga al espacio-tiempo a ser plano —si lo hiciera, no existirían ni los agujeros negros ni las ondas gravitacionales—. La solución de Schwarzschild tiene $R_{ij} = 0$ en todas partes fuera de la masa y sin embargo $R_{ijkl} \neq 0$: toda su curvatura vive en el Weyl.

Regla: medir longitudes, ángulos, áreas y volúmenes con una métrica

La métrica no es un objeto abstracto: es literalmente el instrumento de medida. Todo lo que se puede medir se obtiene de ella con cuatro fórmulas.

1. **Longitud de una curva.** Si $x^\mu(t)$ con $t \in [t_0, t_1]$, entonces
$$L = \int_{t_0}^{t_1} \sqrt{g_{\mu\nu}\, \dot{x}^\mu \dot{x}^\nu}\; \mathrm{d}t.$$
El resultado no depende de la parametrización elegida, lo cual conviene comprobar en un caso sencillo la primera vez.

2. **Ángulo entre dos direcciones.** Con U^μ y V^μ en el mismo punto,
$$\cos\alpha = \frac{g_{\mu\nu} U^\mu V^\nu}{\sqrt{g_{\alpha\beta} U^\alpha U^\beta}\,\sqrt{g_{\rho\sigma} V^\rho V^\sigma}}.$$
Si g es diagonal, las curvas coordenadas se cortan perpendicularmente, y recíprocamente.

3. **Elemento de volumen.** El factor correcto es
$$\mathrm{d}V = \sqrt{|\det g_{\mu\nu}|}\; \mathrm{d}x^1 \cdots \mathrm{d}x^n.$$
Ese $\sqrt{|\det g|}$ es exactamente el jacobiano que aparece al cambiar de variable en una integral múltiple; aquí se obtiene sin acordarse de ninguna tabla.

4. **Área de una superficie dentro del espacio.** Restringe la métrica a la superficie (sustituyendo las coordenadas por su parametrización), obtén la métrica inducida 2×2 y aplica el paso anterior con $n = 2$.

5. **Verifica dimensionalmente.** Cada término de $\mathrm{d}s^2$ debe tener dimensiones de longitud al cuadrado. Si θ es un ángulo, adimensional, su coeficiente debe llevar un a^2 o un r^2. Esta comprobación detecta la mitad de los errores de transcripción.

6. **Compara con el caso plano.** Pon la curvatura a cero, o el radio a infinito, y verifica que recuperas la fórmula euclídea conocida. Si no, hay un error de signo o de factor.

7. **Recuerda que $\mathrm{d}s^2$ es una notación, no un cuadrado.** La expresión $\mathrm{d}s^2 = g_{\mu\nu}\mathrm{d}x^\mu \mathrm{d}x^\nu$ es la forma cuadrática que asigna a cada vector tangente su norma; el símbolo $\mathrm{d}s^2$ nunca se «saca la raíz» término a término.

Ejemplo 36.45 (Medidas concretas sobre la esfera y el plano hiperbólico)**. (a) Longitud de un paralelo.** Sobre la esfera de radio a, el paralelo de colatitud θ_0 se parametriza con $\theta = \theta_0$ constante y $\varphi \in [0, 2\pi]$. Entonces $\dot\theta = 0$ y $\dot\varphi = 1$, luego

$$L = \int_0^{2\pi} \sqrt{a^2 \sin^2\theta_0}\, \mathrm{d}\varphi = 2\pi a \sin\theta_0.$$

En el ecuador ($\theta_0 = \pi/2$) da $2\pi a$, la circunferencia máxima; cerca del polo ($\theta_0 \to 0$) tiende a cero, como debe. Para la Tierra con $a \approx 6371$ km, el paralelo $60°$ norte corresponde a $\theta_0 = 30°$, y mide $2\pi \cdot 6371 \cdot \sin 30° = 2\pi \cdot 6371 \cdot 0{,}5 \approx 20\,015$ km, exactamente la mitad del ecuador.

(b) Área de un casquete esférico. El determinante de la métrica esférica es $\det g = a^2 \cdot a^2 \sin^2\theta = a^4 \sin^2\theta$, luego $\sqrt{\det g} = a^2 \sin\theta$ y

$$A = \int_0^{2\pi}\int_0^{\theta_0} a^2 \sin\theta \,\mathrm{d}\theta\,\mathrm{d}\varphi = 2\pi a^2(1 - \cos\theta_0).$$

Con $\theta_0 = \pi/2$ da $2\pi a^2$, el hemisferio; con $\theta_0 = \pi$ da $4\pi a^2$, la esfera entera. La fórmula tiene una consecuencia curiosa que ya conocía Arquímedes: el área depende solo de la *altura* $h = a(1 - \cos\theta_0)$ del casquete, pues $A = 2\pi a h$. Rebanadas de igual grosor de una naranja esférica tienen todas la misma cáscara.

(c) Ángulos y ortogonalidad. En la esfera, g es diagonal, luego los meridianos y los paralelos se cortan en ángulo recto en todo punto salvo en los polos, donde el sistema degenera. Comprobación con la fórmula del paso 2: para $U = (1, 0)$ y $V = (0, 1)$ se tiene $g_{\mu\nu}U^\mu V^\nu = g_{\theta\varphi} = 0$, luego $\cos\alpha = 0$. Nótese que la ortogonalidad se lee del término cruzado, no de los diagonales.

(d) Distancias hiperbólicas. En el semiplano superior, el determinante es $\det g = 1/y^4$, luego $\sqrt{\det g} = 1/y^2$ y el área hiperbólica es $\displaystyle\iint \frac{\mathrm{d}x\,\mathrm{d}y}{y^2}$. La longitud del segmento vertical de i a $2i$ es

$$\int_1^2 \frac{\mathrm{d}y}{y} = \ln 2 \approx 0{,}6931,$$

mientras que el mismo segmento mide 1 en la métrica euclídea. Y si se recorre el segmento de $0{,}1i$ a $0{,}2i$, euclídeamente mide 0,1 pero hiperbólicamente sigue midiendo $\ln 2$: acercarse al eje real no acorta nada, porque la métrica se dilata al mismo ritmo. Ese eje está infinitamente lejos, y por eso se llama *borde ideal*.

(e) Un área hiperbólica infinita en región acotada. La franja $0 < x < 1$, $0 < y < 1$ tiene área euclídea 1, pero área hiperbólica

$$\int_0^1\int_0^1 \frac{\mathrm{d}y\,\mathrm{d}x}{y^2} = \int_0^1 \left[-\frac{1}{y}\right]_0^1 \mathrm{d}x = \infty.$$

Una región acotada a la vista, pero de área infinita para quien vive dentro. Esta es la razón de que el plano hiperbólico no quepa isométricamente en $\mathbb{R}^3$: crece demasiado deprisa.

Cuidado

El factor $\sqrt{|\det g|}$ va *dentro* de la integral, no fuera, porque en general depende del punto. Olvidarlo es el error más común al integrar en coordenadas curvilíneas, y es exactamente el mismo error que escribir $\iiint f \,\mathrm{d}r\,\mathrm{d}\theta\,\mathrm{d}\varphi$ en vez de $\iiint f\, r^2 \sin\theta \,\mathrm{d}r\,\mathrm{d}\theta\,\mathrm{d}\varphi$ en coordenadas esféricas. El formalismo tensorial lo hace imposible de olvidar, porque el factor sale del propio determinante de la métrica.

Pausa para pensar

La fórmula del casquete da $A = 2\pi a^2(1-\cos\theta_0)$. Desarrolla el coseno en serie para θ_0 pequeño: $1-\cos\theta_0 \approx \theta_0^2/2$, luego $A \approx \pi a^2\theta_0^2 = \pi\rho^2$ con $\rho = a\theta_0$ la distancia medida *sobre la superficie*. Recuperas el área del disco plano. El siguiente término del desarrollo, negativo, es el primer indicio cuantitativo de curvatura: en una esfera, los discos pequeños tienen *menos* área que sus equivalentes planos del mismo radio, y ese déficit es proporcional a la curvatura. Esa observación, llevada hasta el final, es el Teorema Egregio.

Regla: escribir gradiente, divergencia, rotacional y laplaciano en coordenadas cualesquiera

Las fórmulas del gradiente en esféricas o de la divergencia en cilíndricas suelen memorizarse. No hace falta: todas salen de la métrica con dos expresiones, y el mismo procedimiento funciona en superficies curvas donde no hay ninguna tabla que consultar.

1. **Escribe la métrica y su determinante.** Para coordenadas ortogonales, $g = \mathrm{diag}(h_1^2,\dots,h_n^2)$ con los factores de escala h_i, y entonces $\sqrt{|g|} = h_1 h_2 \cdots h_n$. En cilíndricas (ρ,ϕ,z) es $\sqrt{|g|}=\rho$; en esféricas (r,θ,φ) es $\sqrt{|g|} = r^2\sin\theta$.

2. **El gradiente es un covector.** Sin ninguna métrica, $(\mathrm{d}f)_\mu = \partial_\mu f$: las componentes covariantes del gradiente son, literalmente, las derivadas parciales. Para obtener el *vector* gradiente hay que subir el índice:
$$(\operatorname{grad} f)^\mu = g^{\mu\nu}\partial_\nu f.$$
En ortogonales eso divide la i-ésima componente por h_i^2.

3. **La divergencia se calcula con $\sqrt{|g|}$.** La identidad clave es $\Gamma^\mu_{\mu\nu} = \partial_\nu \ln\sqrt{|g|}$, que colapsa la fórmula general $\nabla_\mu V^\mu = \partial_\mu V^\mu + \Gamma^\mu_{\mu\nu}V^\nu$ en
$$\operatorname{div} V = \frac{1}{\sqrt{|g|}}\,\partial_\mu\!\left(\sqrt{|g|}\,V^\mu\right).$$
No hay que calcular ningún símbolo de Christoffel.

4. **El laplaciano es la composición.** Encadenando los pasos 2 y 3 se obtiene el operador de Laplace–Beltrami,
$$\Delta f = \operatorname{div}(\operatorname{grad} f) = \frac{1}{\sqrt{|g|}}\,\partial_\mu\!\left(\sqrt{|g|}\,g^{\mu\nu}\partial_\nu f\right).$$

Esta única fórmula reemplaza todas las tablas de laplacianos en coordenadas especiales.

5. **El rotacional es la derivada exterior.** En cualquier dimensión, la antisimetrización $(\mathrm{d}\omega)_{\mu\nu} = \partial_\mu\omega_\nu - \partial_\nu\omega_\mu$ no necesita métrica ni conexión, porque los Christoffel se cancelan por ser simétricos en los índices bajos. Solo en dimensión 3 se puede volver a convertir en un vector, usando el tensor de Levi-Civita:

$$(\mathrm{rot}\, V)^\mu = \frac{1}{\sqrt{|g|}}\,\varepsilon^{\mu\nu\rho}\,\partial_\nu(g_{\rho\sigma}V^\sigma).$$

Por eso el rotacional "solo existe" en $\mathbb{R}^3$: es una coincidencia dimensional, no una propiedad general.

6. **Cuidado con las componentes físicas.** Los libros de física escriben las componentes *normalizadas* $V_{\hat{i}} = h_i V^i$, que tienen las unidades correctas. Al comparar tu resultado con una tabla, comprueba primero cuál de las dos convenciones usa, o los factores de h_i no cuadrarán.

7. **Verifica con funciones conocidas.** Prueba tu fórmula sobre casos cuyo resultado sabes de antemano: $\Delta(x^2+y^2+z^2) = 6$, $\Delta(1/r) = 0$ fuera del origen, $\Delta(\ln\rho) = 0$ fuera del eje. Si alguno falla, el error está en $\sqrt{|g|}$ o en $g^{\mu\nu}$.

Ejemplo 36.46 (Cinco operadores deducidos, no memorizados)**. (a) Esféricas: el laplaciano completo.** La métrica es $\mathrm{d}s^2 = \mathrm{d}r^2 + r^2\mathrm{d}\theta^2 + r^2\sin^2\theta\,\mathrm{d}\varphi^2$, luego $g^{\mu\nu} = \mathrm{diag}(1,\, r^{-2},\, r^{-2}\sin^{-2}\theta)$ y $\sqrt{|g|} = r^2\sin\theta$ (con $0 < \theta < \pi$, donde $\sin\theta > 0$). Sustituyendo en el paso 4, y separando los tres sumandos:

$$\Delta f = \frac{1}{r^2}\frac{\partial}{\partial r}\left(r^2\frac{\partial f}{\partial r}\right) + \frac{1}{r^2\sin\theta}\frac{\partial}{\partial\theta}\left(\sin\theta\frac{\partial f}{\partial\theta}\right) + \frac{1}{r^2\sin^2\theta}\frac{\partial^2 f}{\partial\varphi^2}.$$

Esa es la fórmula que aparece en todos los formularios, deducida en dos líneas a partir del determinante de la métrica.

(b) Los tres controles del paso 7. Con $f = 1/r$: el primer sumando da $\frac{1}{r^2}\partial_r(r^2\cdot(-r^{-2})) = \frac{1}{r^2}\partial_r(-1) = 0$, y los otros dos se anulan por no depender de θ ni de φ. Luego $\Delta(1/r) = 0$ para $r > 0$: el potencial de Coulomb es armónico, como debía. Con $f = r^2$: sale $\frac{1}{r^2}\partial_r(2r^3) = \frac{6r^2}{r^2} = 6$, que coincide con $\Delta(x^2+y^2+z^2) = 2+2+2 = 6$ calculado en cartesianas. Con $f = r\cos\theta$, que no es otra cosa que z: el primer término da $\frac{1}{r^2}\partial_r(r^2\cos\theta) = \frac{2\cos\theta}{r}$ y el segundo $\frac{1}{r^2\sin\theta}\partial_\theta(-r\sin^2\theta) = -\frac{2\cos\theta}{r}$; suman cero, como debía ser porque $\Delta z = 0$. Tres controles, tres aciertos.

(c) Cilíndricas, y una divergencia célebre. Con $\mathrm{d}s^2 = \mathrm{d}\rho^2 + \rho^2\mathrm{d}\phi^2 + \mathrm{d}z^2$ se tiene $\sqrt{|g|} = \rho$, y el paso 4 da

$$\Delta f = \frac{1}{\rho}\frac{\partial}{\partial\rho}\left(\rho\frac{\partial f}{\partial\rho}\right) + \frac{1}{\rho^2}\frac{\partial^2 f}{\partial\phi^2} + \frac{\partial^2 f}{\partial z^2}.$$

Control: $\Delta(\ln\rho) = \frac{1}{\rho}\partial_\rho(\rho\cdot\rho^{-1}) = \frac{1}{\rho}\partial_\rho(1) = 0$ para $\rho > 0$, que es el potencial logarítmico del hilo cargado.

Y en esféricas, para el campo radial con $V^r = 1/r^2$ (el campo eléctrico de una carga puntual, salvo constantes):

$$\operatorname{div} V = \frac{1}{r^2 \sin\theta}\,\partial_r\!\left(r^2 \sin\theta \cdot \frac{1}{r^2}\right) = \frac{1}{r^2 \sin\theta}\,\partial_r(\sin\theta) = 0.$$

La divergencia es cero en todo punto con $r > 0$, y toda la fuente está concentrada en el origen, donde la fórmula no se aplica. Ese "cero en todas partes salvo un punto" es exactamente la delta de Dirac de la ley de Gauss.

(d) El gradiente y el desliz de las unidades. En esféricas, el paso 2 da las componentes contravariantes

$$(\operatorname{grad} f)^r = \partial_r f, \qquad (\operatorname{grad} f)^\theta = \frac{1}{r^2}\partial_\theta f, \qquad (\operatorname{grad} f)^\varphi = \frac{1}{r^2 \sin^2\theta}\partial_\varphi f,$$

mientras que las componentes *físicas* del paso 6, multiplicando por $h_r = 1$, $h_\theta = r$, $h_\varphi = r\sin\theta$, son

$$\partial_r f, \qquad \frac{1}{r}\partial_\theta f, \qquad \frac{1}{r\sin\theta}\partial_\varphi f,$$

que es la versión de los libros de física. Los dos juegos de números son correctos; describen el mismo vector en bases distintas, una coordenada y otra ortonormal.

(e) Una superficie curva, donde no hay tabla. En la esfera de radio a, con $\mathrm{d}s^2 = a^2\mathrm{d}\theta^2 + a^2\sin^2\theta\,\mathrm{d}\varphi^2$, es $\sqrt{|g|} = a^2\sin\theta$ y $g^{\mu\nu} = \operatorname{diag}(a^{-2}, a^{-2}\sin^{-2}\theta)$, de donde

$$\Delta_{S^2} f = \frac{1}{a^2\sin\theta}\frac{\partial}{\partial\theta}\!\left(\sin\theta\frac{\partial f}{\partial\theta}\right) + \frac{1}{a^2\sin^2\theta}\frac{\partial^2 f}{\partial\varphi^2}.$$

Aplicado a $f = \cos\theta$ da $\frac{1}{a^2\sin\theta}\partial_\theta(-\sin^2\theta) = -\frac{2\cos\theta}{a^2}$, es decir $\Delta_{S^2} f = -\frac{2}{a^2}f$: la función $\cos\theta$ es una autofunción con autovalor $-2/a^2$. Son los armónicos esféricos, y de ahí salen los orbitales atómicos y los multipolos del campo gravitatorio terrestre. Con la fórmula del paso 4 se obtienen sin salir de la superficie.

Cuidado

El $\sqrt{|g|}$ va *dentro* de la derivada y el $1/\sqrt{|g|}$ *fuera*. Escribir $\operatorname{div} V = \partial_\mu V^\mu$ —olvidando los dos factores— es el error más común al pasar a coordenadas curvilíneas, y da resultados que ni siquiera son escalares. Compruébalo con el apartado (c): $\partial_r(1/r^2) = -2/r^3 \neq 0$, mientras que la divergencia correcta es 0. La diferencia no es un detalle numérico; es la diferencia entre haber calculado un tensor y no haber calculado nada.

Idea clave

Las cuatro fórmulas de este bloque son casos particulares de una sola: la derivada exterior d actuando sobre formas de grado 0, 1 y 2, combinada con el operador estrella de Hodge que la métrica proporciona. En ese lenguaje, gradiente, rotacional y divergencia son d aplicada tres veces seguidas, y las identidades $\operatorname{rot}(\operatorname{grad} f) = 0$ y $\operatorname{div}(\operatorname{rot} V) = 0$ son ambas la misma afirmación: $\mathrm{d}\circ\mathrm{d} = 0$. Que en $\mathbb{R}^3$ parezcan tres operadores distintos es un accidente de la dimensión; en dimensión n hay n de ellos, y su fallo en ser exactos es lo que mide la cohomología de De Rham.

36.8. Aplicaciones físicas

Idea clave

Intuición primero. Todo el aparato de índices existe por una exigencia física: las leyes de la naturaleza no pueden depender de quién las escribe ni desde dónde. Si una ecuación iguala tensores del mismo tipo, sigue siendo cierta en cualquier sistema de coordenadas automáticamente, sin recalcular. Eso es lo que significa "covariante", y es la razón de ser de este lenguaje.

Las tres aplicaciones que siguen ilustran la misma idea a distintas escalas. En electromagnetismo, empaquetar $\mathbf{E}$ y $\mathbf{B}$ en un solo tensor $F_{\mu\nu}$ revela que no son dos campos sino uno visto desde distintos observadores: el imán y la corriente son la misma cosa a distinta velocidad. En relatividad general, la ecuación de Einstein se lee en una frase —la materia le dice al espaciotiempo cómo curvarse, el espaciotiempo le dice a la materia cómo moverse—, y la gravedad deja de ser una fuerza para ser geometría. En medios continuos, el tensor de tensiones es simplemente la contabilidad honesta de una fuerza que depende de la orientación de la superficie sobre la que actúa.

Los tensores no son exóticos: son lo que hace falta en cuanto una magnitud tiene varias direcciones acopladas.

36.8.1. Electromagnetismo covariante

En el espaciotiempo de Minkowski (índices griegos $0,1,2,3$; $x^0 = ct$), el potencial es el covector $A_\mu = (-\varphi/c, \mathbf{A})$ y el *tensor electromagnético* es la 2-forma

$$F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu,$$

que empaqueta $\mathbf{E}$ y $\mathbf{B}$ ($F_{0i} = E_i/c$, $F_{ij} = \epsilon_{ijk}B_k$). Las cuatro ecuaciones de Maxwell colapsan en dos:

$$\partial_\mu F^{\mu\nu} = \mu_0 J^\nu, \qquad \partial_{[\lambda} F_{\mu\nu]} = 0,$$

y la segunda es automática ($F = \mathrm{d}A$ y $\mathrm{d}\circ\mathrm{d} = 0$). La covariancia manifiesta de estas ecuaciones bajo transformaciones de Lorentz es, históricamente, lo que llevó a Einstein a la relatividad: el campo magnético *es* el campo eléctrico visto desde otro sistema de referencia.

36.8.2. Relatividad general

Los postulados: el espaciotiempo es una variedad pseudo-riemanniana de dimensión 4; las partículas libres siguen *geodésicas*; y la métrica obedece las *ecuaciones de Einstein*

$$R_{\mu\nu} - \tfrac{1}{2} R\, g_{\mu\nu} = \frac{8\pi G}{c^4}\, T_{\mu\nu},$$

donde $T_{\mu\nu}$ es el tensor energía-momento de la materia. Lectura: *la materia dice al espaciotiempo cómo curvarse; el espaciotiempo dice a la materia cómo moverse*. La gravedad no es una fuerza: es geometría. La primera solución exacta (Schwarzschild, 1916) para el exterior de una masa esférica M:

$$\mathrm{d}s^2 = -\Big(1 - \frac{r_s}{r}\Big)c^2\mathrm{d}t^2 + \Big(1 - \frac{r_s}{r}\Big)^{-1}\mathrm{d}r^2 + r^2(\mathrm{d}\phi^2 + \sin^2\phi\,\mathrm{d}\theta^2), \quad r_s = \frac{2GM}{c^2},$$

predice la precesión del perihelio de Mercurio, la desviación de la luz, el corrimiento gravitacional al rojo (que el GPS debe corregir a diario) y, para $r < r_s$, los agujeros negros.

36.8.3. Medios continuos

El *tensor de tensiones* σ^{ij} de un sólido da la fuerza por unidad de área a través de una superficie de normal n_j: $F^i = \sigma^{ij} n_j$ (un tensor $(2,0)$ simétrico); el *tensor de deformaciones* $\varepsilon_{ij} = \frac{1}{2}(\partial_i u_j + \partial_j u_i)$ mide el estiramiento local, y la ley de Hooke generalizada $\sigma^{ij} = C^{ijkl}\varepsilon_{kl}$ los conecta mediante un tensor de cuarto orden. El *tensor de inercia* I_{ij} de un sólido rígido (simétrico: diagonalizable en ejes principales, teorema espectral en acción) relaciona velocidad angular y momento angular: $L_i = I_{ij}\omega^j$. Los tensores no son exóticos: son la contabilidad correcta de cualquier magnitud con varias direcciones acopladas.

Regla: contar y usar las simetrías del tensor de Riemann

El tensor de Riemann tiene n^4 componentes, pero casi todas están determinadas por las demás. Saber cuántas son verdaderamente independientes, y cuáles se anulan de oficio, ahorra la mayor parte del trabajo de cualquier cálculo de curvatura.

1. **Baja el índice y escribe las tres simetrías algebraicas.** Con $R_{\rho\sigma\mu\nu} = g_{\rho\lambda}R^{\lambda}{}_{\sigma\mu\nu}$ se cumplen

   $$R_{\rho\sigma\mu\nu} = -R_{\sigma\rho\mu\nu}, \qquad R_{\rho\sigma\mu\nu} = -R_{\rho\sigma\nu\mu}, \qquad R_{\rho\sigma\mu\nu} = R_{\mu\nu\rho\sigma},$$

   es decir: antisimetría en el primer par, antisimetría en el segundo par, y simetría al intercambiar los dos pares.

2. **Añade la primera identidad de Bianchi.** La suma cíclica de los tres últimos índices se anula,

   $$R_{\rho[\sigma\mu\nu]} = 0, \qquad \text{o bien} \qquad R_{\rho\sigma\mu\nu} + R_{\rho\mu\nu\sigma} + R_{\rho\nu\sigma\mu} = 0.$$

   Esta es la única relación que no se deduce de las tres anteriores.

3. **Cuenta las componentes independientes.** El resultado de imponer todo lo anterior es

   $$N(n) = \frac{n^2(n^2-1)}{12},$$

   o sea $N(2) = 1$, $N(3) = 6$, $N(4) = 20$ y $N(5) = 50$. En dimensión 2 hay *una sola* función de curvatura; en dimensión 4, veinte.

4. **Aprovecha las antisimetrías al calcular.** Como $R_{\rho\sigma\mu\nu}$ se anula si $\rho = \sigma$ o si $\mu = \nu$, en dimensión 2 solo hay que calcular R_{1212}: las demás son cero o iguales a esa salvo signo. Nunca calcules las n^4 componentes a fuerza bruta.

5. **Contrae para obtener Ricci y el escalar.** Se define $R_{\sigma\nu} = R^{\mu}{}_{\sigma\mu\nu}$, que resulta *simétrico*, y $R = g^{\sigma\nu}R_{\sigma\nu}$. Las simetrías del paso 1 garantizan que las demás contracciones posibles no dan nada nuevo: o son cero, o son $\pm R_{\sigma\nu}$.

6. **Construye el tensor de Einstein y usa Bianchi contraída.** Con $G_{\mu\nu} = R_{\mu\nu} - \frac{1}{2}R\,g_{\mu\nu}$, la segunda identidad de Bianchi $\nabla_{[\lambda}R_{\rho\sigma]\mu\nu} = 0$ implica al contraerla que $\nabla^{\mu}G_{\mu\nu} = 0$ *idénticamente*. Esa conservación automática es lo que obliga a poner $G_{\mu\nu}$ —y no $R_{\mu\nu}$— al lado izquierdo de las ecuaciones de campo.

7. **Comprueba con los casos extremos.** En dimensión 2, $G_{\mu\nu} \equiv 0$ siempre. En dimensión 3, el Riemann queda completamente determinado por el Ricci ($N(3) = 6$ y el Ricci simétrico tiene también 6 componentes). Solo a partir de dimensión 4 sobra información en el Riemann: los $20 - 10 = 10$ grados restantes forman el tensor de Weyl, la parte de la curvatura que no fija la materia local.

Ejemplo 36.47 (La curvatura de cuatro geometrías, contada y calculada)**. (a) La esfera de radio a: una sola componente.** Con $\mathrm{d}s^2 = a^2\mathrm{d}\theta^2 + a^2\sin^2\theta\,\mathrm{d}\varphi^2$ estamos en $n = 2$, así que por el paso 3 hay $N(2) = 1$ componente independiente. Los símbolos de Christoffel no nulos son

$$\Gamma^{\theta}_{\varphi\varphi} = -\sin\theta\cos\theta, \qquad \Gamma^{\varphi}_{\theta\varphi} = \Gamma^{\varphi}_{\varphi\theta} = \cot\theta,$$

y de ahí la única componente

$$R_{\theta\varphi\theta\varphi} = a^2\sin^2\theta.$$

Contrayendo: $R_{\mu\nu} = \dfrac{1}{a^2}\,g_{\mu\nu}$ y $R = \dfrac{2}{a^2}$. La curvatura de Gauss es $K = R/2 = 1/a^2$, positiva y constante, y decrece con el tamaño: una esfera grande es localmente casi plana, que es por lo que la Tierra parece llana.

(b) El plano hiperbólico: el mismo cálculo con el signo cambiado. Con $\mathrm{d}s^2 = \dfrac{\mathrm{d}x^2 + \mathrm{d}y^2}{y^2}$ en el semiplano $y > 0$, el mismo procedimiento da

$$R_{\mu\nu} = -g_{\mu\nu}, \qquad R = -2, \qquad K = -1.$$

Es el modelo de curvatura negativa constante. Nótese que R es constante pese a que la métrica depende de y: la curvatura es un invariante geométrico, no un artefacto de las coordenadas elegidas.

(c) El tensor de Einstein en dimensión 2 se anula solo. Usando (a),

$$G_{\mu\nu} = R_{\mu\nu} - \tfrac{1}{2}R\,g_{\mu\nu} = \frac{1}{a^2}g_{\mu\nu} - \frac{1}{2}\cdot\frac{2}{a^2}\,g_{\mu\nu} = 0.$$

Y con (b): $-g_{\mu\nu} - \frac{1}{2}(-2)g_{\mu\nu} = 0$ también. No es coincidencia de estos dos ejemplos: en $n = 2$ siempre $R_{\mu\nu} = \frac{1}{2}Rg_{\mu\nu}$, y por tanto $G \equiv 0$. Consecuencia física notable: *no existe* relatividad general en dos dimensiones, porque las ecuaciones de campo se convierten en $0 = 8\pi T_{\mu\nu}$, que solo admite el vacío. Se necesitan al menos tres dimensiones, y para que además haya ondas y grados de libertad locales, cuatro.

(d) El conteo en dimensión 4, y por qué el vacío no es plano. Aquí $N(4) = \dfrac{16\cdot 15}{12} = 20$ componentes independientes del Riemann, mientras que el Ricci, al ser simétrico 4×4, tiene solo 10. La diferencia, $20 - 10 = 10$, es el tensor de Weyl. Esto explica la aparente paradoja de la solución de Schwarzschild: fuera de la estrella se cumple $R_{\mu\nu} = 0$, y sin embargo

el espacio-tiempo *no* es plano, porque las diez componentes de Weyl siguen ahí. Son ellas las que producen las fuerzas de marea y las que se propagan como ondas gravitacionales. En dimensión 3, en cambio, $N(3) = 6$ coincide con las 6 componentes del Ricci simétrico: no hay Weyl, el vacío es forzosamente plano, y no hay ondas gravitacionales. La existencia misma de la gravedad a distancia depende de que vivamos en cuatro dimensiones y no en tres.

(e) El ahorro del paso 4, en números. Calcular a fuerza bruta el Riemann en $n = 4$ significaría $4^4 = 256$ componentes. Las antisimetrías reducen los pares $(\rho\sigma)$ y $(\mu\nu)$ a $\binom{4}{2} = 6$ valores cada uno, es decir 36 entradas; la simetría de pares las baja a $\binom{6}{2} + 6 = 21$; y la identidad de Bianchi del paso 2 impone una relación más, dejando 20. De 256 a 20: un factor de casi 13. En dimensión 2 el ahorro es aún más drástico: de 16 componentes a 1.

Cuidado

Las convenciones de signo del tensor de Riemann difieren entre libros, y hay tres elecciones independientes (el signo de la métrica, el de la definición de $R^{\rho}{}_{\sigma\mu\nu}$ y el de la contracción de Ricci). Con las convenciones de este capítulo, la esfera tiene $R = +2/a^2$; con otras, $-2/a^2$. Antes de comparar un resultado con otra fuente, comprueba *siempre* el signo que esa fuente asigna a la esfera: es el control más rápido que existe, y evita horas de buscar un error que no está.

Pausa para pensar

El paso 6 merece detenerse. Einstein buscó durante años el lado izquierdo correcto de sus ecuaciones, y su primer intento fue $R_{\mu\nu} = 8\pi T_{\mu\nu}$. El problema es que el lado derecho satisface $\nabla^\mu T_{\mu\nu} = 0$ por conservación de la energía, mientras que $\nabla^\mu R_{\mu\nu} = \frac{1}{2}\nabla_\nu R$ no se anula en general: la ecuación era inconsistente salvo que R fuese constante. La identidad de Bianchi contraída dice exactamente qué combinación arregla el problema, y la respuesta $G_{\mu\nu} = R_{\mu\nu} - \frac{1}{2}Rg_{\mu\nu}$ es única salvo añadir $\Lambda g_{\mu\nu}$ —la constante cosmológica, que también tiene divergencia nula porque $\nabla_\lambda g_{\mu\nu} = 0$—. Dicho de otro modo: la forma de las ecuaciones de campo no se eligió, se dedujo de una identidad geométrica, y el único parámetro libre que quedaba es justamente el que hoy medimos como energía oscura.

36.9. Fundamentos: la teoría detrás del formalismo

36.9.1. El producto tensorial, construido formalmente

Hasta ahora tratamos $\otimes$ como una notación; tiene una definición precisa que lo caracteriza.

Definición 36.48 (Propiedad universal)**.** El *producto tensorial* $V \otimes W$ es un espacio vectorial junto con una aplicación bilineal $\otimes : V \times W \to V \otimes W$ tal que **toda** aplicación bilineal $B : V \times W \to Z$ se factoriza de forma única como $B = \tilde{B} \circ \otimes$ con $\tilde{B} : V \otimes W \to Z$ *lineal*. En palabras: $V \otimes W$ es la máquina universal que convierte lo bilineal en lineal.

Idea clave

Qué dice esto en cristiano. El producto tensorial es la manera canónica de convertir algo bilineal en algo lineal. Si una operación come dos vectores y es lineal en cada uno por separado, se puede reinterpretar como una operación lineal que come un solo objeto del espacio $V \otimes W$. La propiedad universal dice exactamente eso, y además que $V \otimes W$ es el *menor* espacio con esa capacidad. Por eso se define por lo que hace y no por lo que es: la construcción concreta es fea y la propiedad es limpia.

Ejemplo 36.49 (Cómo se calcula en la práctica)**. (a) Dimensión y base.** Si $\dim V = m$ y $\dim W = n$, entonces $\dim(V \otimes W) = mn$, con base $\{e_i \otimes f_j\}$. Para $\mathbb{R}^2 \otimes \mathbb{R}^3$ la dimensión es 6. Compárese con $\mathbb{R}^2 \oplus \mathbb{R}^3$, que tiene dimensión 5: el producto tensorial multiplica, la suma directa suma.

(b) Un cálculo explícito. En $\mathbb{R}^2 \otimes \mathbb{R}^2$, con $u = (1, 2)$ y $v = (3, 4)$, el tensor $u \otimes v$ tiene componentes $u_i v_j$, es decir la matriz $\begin{pmatrix} 3 & 4 \\ 6 & 8 \end{pmatrix}$. Es el producto exterior de dos vectores, y tiene siempre rango 1.

(c) No todo tensor es un producto de dos vectores. La matriz $\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$ tiene rango 2 y no se puede escribir como $u \otimes v$ para ningún par: es $e_1 \otimes e_1 + e_2 \otimes e_2$, una suma de dos. Los tensores que sí son producto se llaman *descomponibles* o puros, y son los de rango 1; en física cuántica los demás son precisamente los estados *entrelazados*.

(d) La bilinealidad, comprobada. Se cumple $(2u) \otimes v = 2(u \otimes v)$ y $(u + u') \otimes v = u \otimes v + u' \otimes v$, pero $u \otimes v \neq v \otimes u$ en general: con los vectores de (b), $v \otimes u$ es $\begin{pmatrix} 3 & 6 \\ 4 & 8 \end{pmatrix}$, la traspuesta. El producto tensorial no es conmutativo.

Cuidado

[La dimensión se multiplica, y por eso todo explota] Con n sistemas de dimensión 2, el espacio conjunto tiene dimensión 2^n: con 50 partículas ya supera el número de átomos del planeta. Ese crecimiento es la razón de que simular sistemas cuánticos sea intratable y la motivación original de la computación cuántica. Segundo aviso: la notación $u \otimes v$ para un elemento genérico es engañosa, porque el elemento genérico de $V \otimes W$ es una *suma* de esos productos, no uno solo, como enseña el apartado (c).

Teorema 36.50. *$V \otimes W$ existe y es único salvo isomorfismo canónico; si $\{v_i\}$, $\{w_j\}$ son bases, $\{v_i \otimes w_j\}$ es base de $V \otimes W$, de dimensión $(\dim V)(\dim W)$.*

Construcción. Se toma el espacio (enorme) de combinaciones formales de pares $[v, w]$ y se cociente por el subespacio generado por las relaciones de bilinealidad ($[v+v', w] - [v, w] - [v', w]$, $[cv, w] - c[v, w]$, etc.) —el mismo patrón "objeto libre módulo relaciones" de los cocientes del capítulo 25—. La clase de $[v, w]$ es $v \otimes w$. La propiedad universal se verifica directamente, y

la unicidad es el argumento estándar: dos objetos universales se factorizan mutuamente y las composiciones son la identidad. □

Ejemplo 36.51 (El producto tensorial en la práctica)**. (a) Dimensiones.** Si $\dim V = 2$ y $\dim W = 3$, entonces $\dim(V \otimes W) = 6$, con base $\{v_1 \otimes w_1,\ v_1 \otimes w_2,\ v_1 \otimes w_3,\ v_2 \otimes w_1,\ v_2 \otimes w_2,\ v_2 \otimes w_3\}$. Compárese con la suma directa, que tendría dimensión 5: el tensorial multiplica, la suma directa suma.

(b) Como matrices. Un elemento de $\mathbb{R}^2 \otimes \mathbb{R}^3$ se puede escribir como una matriz 2×3. El tensor $(1,2) \otimes (1,0,-1)$ es la matriz $\left(\begin{smallmatrix} 1 & 0 & -1 \\ 2 & 0 & -2 \end{smallmatrix}\right)$, de rango 1. Los tensores de la forma $v \otimes w$ son exactamente las matrices de rango uno.

(c) No todo tensor es puro. La matriz identidad 2×2, es decir $e_1 \otimes e_1 + e_2 \otimes e_2$, tiene rango 2 y *no* se puede escribir como un único $v \otimes w$. Ésta es la razón matemática del entrelazamiento cuántico: los estados de un sistema compuesto que no factorizan son precisamente los tensores no puros.

(d) La propiedad universal en acción. El determinante de una matriz 2×2 es bilineal en sus dos columnas, luego por la propiedad universal existe una única aplicación lineal $\mathbb{R}^2 \otimes \mathbb{R}^2 \to \mathbb{R}$ que lo reproduce. Ese es el mecanismo general: toda construcción bilineal se convierte automáticamente en una lineal sobre el producto tensorial, y trabajar con aplicaciones lineales es infinitamente más cómodo.

Con esto, las definiciones del capítulo cierran el círculo: el espacio de tensores (p,q) es *canónicamente* $V^{\otimes p} \otimes (V^*)^{\otimes q}$, y un tensor $(1,1)$ es un elemento de $V \otimes V^* \cong \mathrm{End}(V)$: la identificación "matriz = tensor mixto" es un isomorfismo natural, no una casualidad de índices. Las formas diferenciales del capítulo 11 viven en los cocientes antisimétricos $\Lambda^k V^*$, de dimensión $\binom{n}{k}$.

36.9.2. La conexión de Levi-Civita: existencia y unicidad

En el capítulo dimos la fórmula (36.1) como definición. El teorema real es más profundo: la fórmula está *forzada*.

Teorema 36.52 (Fundamental de la geometría riemanniana)**.** *En una variedad con métrica g existe una única conexión ∇ que es (i)* compatible con la métrica*: $\partial_k(g(u,v)) = g(\nabla_k u, v) + g(u, \nabla_k v)$, y (ii)* simétrica *(sin torsión): $\Gamma^k_{ij} = \Gamma^k_{ji}$.*

Demostración: el truco de Koszul. Escribimos la compatibilidad tres veces, permutando cíclicamente los índices:

$$\begin{aligned} \partial_i g_{jl} &= \Gamma^m_{ij} g_{ml} + \Gamma^m_{il} g_{jm}, \\ \partial_j g_{il} &= \Gamma^m_{ji} g_{ml} + \Gamma^m_{jl} g_{im}, \\ \partial_l g_{ij} &= \Gamma^m_{li} g_{mj} + \Gamma^m_{lj} g_{im}. \end{aligned}$$

Sumamos las dos primeras y restamos la tercera. Usando la simetría $\Gamma^m_{ij} = \Gamma^m_{ji}$, cuatro de los seis términos se cancelan por pares y queda

$$\partial_i g_{jl} + \partial_j g_{il} - \partial_l g_{ij} = 2\,\Gamma^m_{ij} g_{ml}.$$

Contrayendo con $\frac{1}{2}g^{kl}$ se despeja exactamente la fórmula (36.1): *cualquier* conexión con las dos propiedades tiene esos coeficientes (unicidad), y la fórmula misma define una que las cumple (existencia). □

Idea clave

Qué dice esto en cristiano. En una superficie curva no hay manera obvia de comparar un vector de un punto con uno de otro: viven en planos tangentes distintos. Una conexión es una receta para transportarlos. El teorema dice que, una vez fijada la forma de medir (la métrica), hay *una sola* receta razonable: la que no cambia las longitudes al transportar y la que no introduce torsión artificial. Es un resultado de unicidad, y es lo que permite hablar de "la" derivada covariante y de "la" curvatura de un espacio, sin tener que elegir nada a mano.

Ejemplo 36.53 (La conexión de la esfera, calculada)**.** Tómese la esfera de radio R con coordenadas (θ,φ) y métrica $g=\operatorname{diag}(R^2,\ R^2\sin^2\theta)$.

(a) Los símbolos. La fórmula $\Gamma^k_{ij}=\frac{1}{2}g^{kl}(\partial_i g_{lj}+\partial_j g_{li}-\partial_l g_{ij})$ da sólo dos no nulos:

$$\Gamma^{\theta}_{\varphi\varphi}=-\sin\theta\cos\theta,\qquad \Gamma^{\varphi}_{\theta\varphi}=\Gamma^{\varphi}_{\varphi\theta}=\cot\theta.$$

Todos los demás se anulan. Nótese que no dependen de R: la conexión no ve la escala.

(b) Qué curvas son geodésicas. La ecuación $\ddot{x}^k+\Gamma^k_{ij}\dot{x}^i\dot{x}^j=0$ para un paralelo $\theta=\theta_0$ constante exige $\Gamma^{\theta}_{\varphi\varphi}=0$, es decir $\sin\theta_0\cos\theta_0=0$, que sólo se cumple en $\theta_0=\pi/2$. Conclusión: de todos los paralelos, el único que es geodésica es el ecuador. Los meridianos, en cambio, lo son todos.

(c) La curvatura. A partir de estos símbolos sale $R_{\theta\varphi\theta\varphi}=R^2\sin^2\theta$ y la curvatura de Gauss $K=1/R^2$: constante y positiva, y tanto menor cuanto mayor es la esfera. Además $\mathrm{Ric}=K\,g$, lo que dice que la esfera es un espacio de Einstein.

(d) La torsión que se descarta. Si se admitiesen conexiones con torsión, habría infinitas compatibles con la misma métrica: basta sumar a Γ^k_{ij} cualquier tensor antisimétrico en i,j. La condición $\Gamma^k_{ij}=\Gamma^k_{ji}$ es la que reduce esas infinitas a una. En relatividad general se toma la de Levi-Civita; en algunas teorías de gravedad con espín se conserva la torsión a propósito.

Por qué $\nabla_i v^k$ es un tensor. Bajo un cambio de coordenadas $x\to x'$, la derivada a secas transforma con un término espurio: $\partial'_i v'^k=$ (ley tensorial)$+v^m\,\partial'_i\big(\frac{\partial x'^k}{\partial x^m}\big)$, que involucra *segundas* derivadas del cambio. Un cálculo directo con la fórmula de Koszul muestra que Γ'^k_{ij} transforma con exactamente el mismo término espurio (con signo opuesto en la combinación $\partial v+\Gamma v$): en la suma, ambos se cancelan y $\nabla_i v^k$ transforma limpio. Los símbolos de Christoffel son, precisamente, *el error de tensorialidad de la derivada parcial*, empaquetado. □

36.9.3. Geodésicas por cálculo de variaciones

Afirmamos que las geodésicas son las curvas de longitud estacionaria; demostrémoslo, introduciendo de paso la herramienta variacional.

Teorema 36.54 (Ecuación de Euler–Lagrange)**.** *La curva $x(t)$ que hace estacionario el funcional $S[x]=\int_a^b L(x,\dot{x})\,\mathrm{d}t$ satisface*

$$\frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial L}{\partial \dot{x}^k}-\frac{\partial L}{\partial x^k}=0.$$

Demostración. Perturbamos $x\to x+\varepsilon\eta$ con $\eta(a)=\eta(b)=0$ y exigimos $\frac{\mathrm{d}}{\mathrm{d}\varepsilon}S\big|_{\varepsilon=0}=0$:

$$0=\int_a^b\Big(\frac{\partial L}{\partial x^k}\eta^k+\frac{\partial L}{\partial \dot{x}^k}\dot{\eta}^k\Big)\mathrm{d}t=\int_a^b\Big(\frac{\partial L}{\partial x^k}-\frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial L}{\partial \dot{x}^k}\Big)\eta^k\,\mathrm{d}t,$$

integrando por partes el segundo término (el borde muere por $\eta(a)=\eta(b)=0$). Como η es arbitraria, el paréntesis debe anularse (lema fundamental del cálculo de variaciones: si $\int fg=0$ para toda g de soporte compacto, $f\equiv 0$). □

Idea clave

Qué dice esto en cristiano. La ecuación de Euler–Lagrange resuelve un tipo de problema que no es de cálculo ordinario: en vez de buscar el punto donde una función es mínima, busca la *curva* donde un funcional es mínimo. El truco es el mismo de siempre, imponer que la variación primera se anule, pero aplicado a una perturbación de la curva entera. El resultado es una ecuación diferencial cuya solución es la curva óptima, y de ella salen las leyes de Newton, las geodésicas y buena parte de la física teórica.

Ejemplo 36.55 (Cuatro problemas resueltos con la misma ecuación)**. (a) El camino más corto en el plano.** Minimizar $\int\sqrt{1+y'^2}\,\mathrm{d}x$. Como el lagrangiano no depende de y, la ecuación se reduce a que $\frac{y'}{\sqrt{1+y'^2}}$ sea constante, luego y' es constante y la curva es una recta. La geometría euclídea sale de una ecuación diferencial.

(b) Las leyes de Newton. Con $L=T-V=\frac{1}{2}m\dot{x}^2-V(x)$, la ecuación $\frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial L}{\partial \dot{x}}-\frac{\partial L}{\partial x}=0$ da $m\ddot{x}=-V'(x)$, es decir $F=ma$. Toda la mecánica clásica es un principio variacional disfrazado, y esa reformulación es lo que permitió extenderla a la relatividad y a la teoría cuántica de campos.

(c) Las geodésicas. Minimizando la longitud sobre una superficie se obtienen las geodésicas: en la esfera son los círculos máximos, y por eso las rutas aéreas entre continentes parecen curvas en un mapa plano y son rectas sobre el globo.

(d) La braquistócrona. El problema que lanzó Johann Bernoulli en 1696 (hallar la curva de descenso más rápido bajo la gravedad) tiene por solución una cicloide, no una recta. Newton lo resolvió en una noche y envió la respuesta anónima; Bernoulli, al leerla, dijo que reconocía al león por su zarpa. Es el problema fundacional del cálculo de variaciones.

Cuidado

[Estacionario no es mínimo] La ecuación de Euler–Lagrange caracteriza los puntos *estacionarios* del funcional, que pueden ser mínimos, máximos o ninguna de las dos cosas. En la esfera, un arco de círculo máximo es geodésica tanto si es el camino corto como si es el largo dando la vuelta por el otro lado: ambos satisfacen la ecuación.

Segundo aviso: el teorema presupone que el mínimo existe y es suave, y esa suposición no es gratuita; el llamado principio de Dirichlet se usó durante décadas sin justificarla, hasta que Weierstrass exhibió funcionales sin mínimo y Hilbert tuvo que reconstruir el argumento.

Aplicación a las geodésicas. Tomamos como lagrangiano la *energía* $L = \frac{1}{2} g_{ij}(x)\dot{x}^i\dot{x}^j$ (más cómoda que la longitud, y con los mismos extremales parametrizados por arco). Entonces

$$\frac{\partial L}{\partial \dot{x}^k} = g_{kj}\dot{x}^j, \qquad \frac{\partial L}{\partial x^k} = \tfrac{1}{2}\partial_k g_{ij}\,\dot{x}^i\dot{x}^j,$$

y Euler–Lagrange da $g_{kj}\ddot{x}^j + \partial_i g_{kj}\dot{x}^i\dot{x}^j - \frac{1}{2}\partial_k g_{ij}\dot{x}^i\dot{x}^j = 0$. Simetrizando el término del medio en (i,j) y contrayendo con g^{lk}, los coeficientes que aparecen son exactamente los de Koszul:

$$\ddot{x}^l + \Gamma^l_{ij}\dot{x}^i\dot{x}^j = 0.$$

La ecuación geodésica *es* la ecuación de Euler–Lagrange de la energía. En relatividad general este cálculo es literalmente la ley de movimiento: "caída libre = trayectoria estacionaria del tiempo propio". Y de la mecánica lagrangiana se hereda el teorema de Noether: cada coordenada de la que g no depende regala una cantidad conservada $g_{kj}\dot{x}^j$ —en la esfera, la independencia de θ da la "componente angular" conservada, que es la clave práctica para integrar las geodésicas—. □

Idea clave

Intuición primero. El tensor de Riemann es la *memoria del transporte*: pasea un vector por un circuitito cerrado (lados ε) manteniéndolo "siempre paralelo", y al volver estará girado un ángulo proporcional a ε^2 por la curvatura del parche encerrado. $R^i{}_{jkl}$ empaqueta todos esos giros: los índices k, l eligen el plano del circuitito, y i, j describen el giro resultante. Sus simetrías, que se demuestran a continuación, no son numerología: dicen que un giro es antisimétrico (rotar es antisimétrico), que el circuito puede recorrerse al revés, y que intercambiar "plano del paseo" con "plano del giro" no cambia la memoria: la geometría del bucle es una sola.

36.9.4. Las simetrías del tensor de Riemann, demostradas

Teorema 36.56. *En un punto cualquiera existen* coordenadas normales*: $g_{ij} = \delta_{ij}$ y $\Gamma^k_{ij} = 0$* en ese punto *(no en un entorno, salvo que la curvatura sea nula). Se construyen lanzando geodésicas radiales desde el punto.*

Idea clave

Qué dice esto en cristiano. Por curvado que esté un espacio, siempre se puede elegir un sistema de coordenadas en el que, *justo en un punto*, todo parece plano: la métrica es la identidad y todos los símbolos de Christoffel se anulan. Es la versión geométrica de que una superficie lisa se confunde con su plano tangente si uno se acerca lo suficiente. Lo decisivo es la letra pequeña: se anulan en el punto, no en sus alrededores. Las *derivadas*

de los símbolos no se pueden anular, y ahí es donde se esconde la curvatura, que por eso es un objeto de segundo orden.

Ejemplo 36.57 (Aplanar localmente y por qué no basta)**. (a) La esfera en su polo.** Con coordenadas normales centradas en el polo norte, la métrica de la esfera unidad se escribe $g_{ij} = \delta_{ij} - \frac{1}{3}R_{ikjl}x^k x^l + O(|x|^3)$, es decir, la identidad más una corrección cuadrática. A primer orden es indistinguible del plano.

(b) Cuánto se despega. En la esfera unidad, la circunferencia de radio geodésico r mide $2\pi \sin r$ en lugar de $2\pi r$: para $r = 0{,}1$ da $2\pi \cdot 0{,}0998334$, un defecto del uno y medio por mil; para $r = 1$ da $2\pi \cdot 0{,}841471$, un defecto ya del 16 %. El desarrollo $2\pi r(1 - Kr^2/6 + \cdots)$ muestra que el término que delata la curvatura es de orden r^3.

(c) Por qué se usa. Muchas identidades tensoriales se demuestran eligiendo coordenadas normales en un punto, donde los Γ desaparecen y las derivadas covariantes se convierten en derivadas parciales ordinarias. Como una igualdad entre tensores que vale en un sistema vale en todos, y el punto era arbitrario, la identidad queda demostrada en general. Es el truco más rentable del cálculo tensorial.

(d) El principio de equivalencia. Físicamente, esto es la afirmación de Einstein de que un observador en caída libre no siente gravedad: en su sistema local todo es Minkowski y los Γ se anulan. Lo que *sí* siente son las fuerzas de marea, que son precisamente las derivadas de los Γ, es decir, la curvatura. La gravedad no se puede eliminar globalmente porque el tensor de Riemann no es cero, y ningún cambio de coordenadas anula un tensor no nulo.

Las coordenadas normales son el "laboratorio local": en ellas, en el punto base,

$$R_{klij} = \tfrac{1}{2}(\partial_i\partial_k g_{jl} + \partial_j\partial_l g_{ik} - \partial_i\partial_l g_{jk} - \partial_j\partial_k g_{il}),$$

(los términos $\Gamma\Gamma$ desaparecen), y de esta expresión *se leen* las simetrías:

$$R_{klij} = -R_{lkij} = -R_{klji} = R_{ijkl}, \qquad R_{k[lij]} = 0 \text{ (primera identidad de Bianchi)},$$

verificando cada una por inspección del patrón de índices en las segundas derivadas. Como son ecuaciones tensoriales, valen en *todas* las coordenadas: demostrar en el punto cómodo y concluir en general es *el* método de trabajo de la geometría diferencial. Un conteo cuidadoso de estas simetrías da las $\frac{n^2(n^2-1)}{12}$ componentes independientes que enunciamos.

Teorema 36.58 (Desviación geodésica)**.** *Sea ξ^k el vector que separa dos geodésicas vecinas. Entonces*

$$\frac{D^2\xi^k}{\mathrm{d}t^2} = -R^k{}_{lij}\,\dot{x}^l\,\xi^i\,\dot{x}^j :$$

la curvatura mide la aceleración relativa de geodésicas próximas. En el plano ($R = 0$), las rectas paralelas mantienen su separación lineal; en la esfera ($R > 0$), los meridianos vecinos —paralelos en el ecuador— convergen en el polo. En relatividad, esta ecuación son *las fuerzas de marea: la Luna estira los océanos porque el tensor de Riemann de la métrica terrestre-lunar no es cero, y la detección de ondas gravitacionales (LIGO) consiste en medir oscilaciones de ξ entre espejos en caída libre.*

36.9.5. El Teorema Egregium y qué significa "intrínseco"

Teorema 36.59 (Gauss, Theorema Egregium)**.** *La curvatura de Gauss K de una superficie, definida originalmente de forma* extrínseca *(cómo se curva dentro de $\mathbb{R}^3$: producto de las curvaturas principales), se calcula solo con la métrica intrínseca: $K = \frac{R}{2}$ con R el escalar de curvatura. En consecuencia, es invariante bajo isometrías: doblar sin estirar no la cambia.*

Idea clave

Qué dice esto en cristiano. Gauss–Bonnet relaciona algo que se mide localmente, la curvatura, con algo que se cuenta globalmente, la topología. En su versión local dice que la suma de los ángulos de un triángulo geodésico no vale π salvo en el plano: el exceso o el defecto es exactamente la integral de la curvatura sobre el triángulo. Y el Theorema Egregium de Gauss añade la sorpresa: esa curvatura, que se definió mirando la superficie desde fuera, se puede calcular sin salir de ella. Un habitante de la superficie puede medirla sin asomarse.

Ejemplo 36.60 (Medir la curvatura desde dentro)**. (a) Un triángulo con tres ángulos rectos.** En la esfera unidad, el triángulo formado por dos meridianos separados 90° y el ecuador tiene los tres ángulos rectos. Su suma es $\frac{3\pi}{2} = 4{,}712389$, con un exceso de $\frac{\pi}{2} = 1{,}570796$ sobre π. Y su área es la octava parte de la esfera, $\frac{4\pi}{8} = 1{,}570796$: exceso igual a área, porque $K = 1$.

(b) Los tres casos. En el plano, $K = 0$ y la suma es exactamente π: la geometría de Euclides. En la esfera, $K > 0$ y hay exceso. En el plano hiperbólico, $K = -1$ y hay defecto, de modo que los triángulos tienen menos de 180° y su área está *acotada* por π, cosa que no ocurre en ninguno de los otros dos.

(c) El Theorema Egregium en acción. Para la esfera de radio a la curvatura es $K = \frac{1}{a^2}$, que vale 0,25 para $a = 2$. Para el cilindro, $K = 0$, porque una de las dos curvaturas principales es nula. Por eso una hoja de papel se enrolla en un cilindro sin arrugarse y no se puede envolver una pelota: el papel tiene $K = 0$ y la esfera no, y K no cambia al doblar sin estirar.

(d) La consecuencia cartográfica. No existe ningún mapa plano de la Tierra que conserve las distancias, y no es un problema técnico: es el Theorema Egregium. La Tierra tiene $K = \frac{1}{R^2} \neq 0$ y el papel tiene $K = 0$, así que cualquier proyección deforma necesariamente algo. Mercator conserva ángulos y estropea áreas; Peters conserva áreas y estropea formas. Elegir mapa es elegir qué se sacrifica.

Cuidado

[Curvatura de Gauss y curvatura media son cosas distintas] K es el *producto* de las dos curvaturas principales y es intrínseca; la curvatura media es su promedio y *no* lo es. El cilindro tiene $K = 0$ y curvatura media no nula: por dentro es indistinguible del plano y por fuera es evidentemente curvo. Segundo aviso: la versión global del teorema, $\int K = 2\pi\chi(S)$, dice que la integral de la curvatura solo depende de la topología, así que deformar una esfera abollándola redistribuye la curvatura pero no cambia el total, que

sigue valiendo 4π.

Consecuencias tangibles. Un plano tiene $K = 0$ y una esfera $K = 1/R_0^2 > 0$: *no existe ningún mapa plano de la Tierra que preserve las distancias*, ni siquiera de una región pequeña —todos los mapas mienten en algo, y la cartografía es el arte de elegir la mentira—. La pizza se sostiene doblándola: al curvar la rebanada en una dirección ($K = 0$ debe conservarse), la otra dirección se ve obligada a mantenerse recta. Y el cilindro, que "se ve curvo", tiene $K = 0$: se desenrolla en un plano sin estirar, por eso sus geodésicas (hélices) se convierten en rectas al desenrollar. La distinción intrínseco/extrínseco que Gauss descubrió en las superficies es la que Riemann generalizó a cualquier dimensión y Einstein convirtió en física: el espaciotiempo no está curvado "dentro de" nada; su curvatura es intrínseca, medible desde adentro con relojes y geodésicas.

36.10. Complementos: Levi-Civita, Killing y el plano hiperbólico

36.10.1. El símbolo de Levi-Civita

ϵ_{ijk} = signo de la permutación (0 con repetidos): $(\mathbf{a}\times\mathbf{b})_i = \epsilon_{ijk}a_jb_k$, $(\operatorname{rot}\mathbf{F})_i = \epsilon_{ijk}\partial_jF_k$, $\det A = \epsilon_{ijk}A_{1i}A_{2j}A_{3k}$. La identidad maestra:

$$\epsilon_{ijk}\,\epsilon_{ilm} = \delta_{jl}\delta_{km} - \delta_{jm}\delta_{kl}, \tag{36.2}$$

que sustituye todos los formularios "bac–cab". En curvilíneas, el tensor de volumen es $\sqrt{g}\,\epsilon_{ijk}$ (densidad corregida).

36.10.2. Derivada de Lie y campos de Killing

$(\mathcal{L}_X g)_{ij} = \nabla_i X_j + \nabla_j X_i$; *Killing*: $\mathcal{L}_X g = 0$ (isometrías infinitesimales).

Teorema 36.61. *X Killing y $x(t)$ geodésica $\Rightarrow$ $g_{ij}X^i\dot{x}^j$ constante. (El término $\nabla_iX_j\dot{x}^i\dot{x}^j$ es antisimétrico contraído con simétrico: cero.)*

Ejemplo 36.62 (Simetrías y cantidades conservadas)**. (a) La esfera.** El campo $X = \partial_\varphi$ es de Killing porque la métrica no depende de φ. La cantidad conservada a lo largo de una geodésica es $g_{\varphi\varphi}\dot\varphi = R^2\sin^2\theta\,\dot\varphi$: es el momento angular respecto del eje polar. De ahí sale, sin resolver ninguna ecuación diferencial, que una geodésica que empieza tangente a un meridiano se queda en él para siempre, porque su constante vale cero.

(b) El plano. En $\mathbb{R}^2$ los campos ∂_x, ∂_y y $x\partial_y - y\partial_x$ son de Killing: tres simetrías, tres constantes, que son las dos componentes del momento lineal y el momento angular. Una superficie de dimensión 2 no puede tener más de tres campos de Killing independientes, y las que alcanzan el máximo son exactamente las de curvatura constante: plano, esfera y plano hiperbólico.

(c) Schwarzschild. La métrica exterior de un agujero negro no depende de t ni de φ, luego ∂_t y ∂_φ son de Killing y dan dos constantes: la energía y el momento angular de una

partícula en caída. Con ellas, el problema de las órbitas se reduce a una única ecuación de primer orden, igual que en el problema de Kepler newtoniano.

(d) Un espacio sin simetrías. Una superficie con una protuberancia irregular no admite ningún campo de Killing no trivial: no hay ninguna cantidad conservada y las geodésicas hay que integrarlas numéricamente. Éste es el caso genérico; los espacios con simetrías son la excepción, y es por eso que son los que aparecen en todos los libros.

Noether geométrico: cada simetría regala una conservada. Plano: 3 Killing (momentos lineal y angular); Schwarzschild: ∂_t y rotaciones: energía y momento angular, que integran las órbitas.

36.10.3. El plano hiperbólico

$\mathbb{H} = \{y > 0\}$ con $\mathrm{d}s^2 = \frac{\mathrm{d}x^2+\mathrm{d}y^2}{y^2}$: Christoffel $\Gamma^x_{xy} = -\frac{1}{y}$, $\Gamma^y_{xx} = \frac{1}{y}$, $\Gamma^y_{yy} = -\frac{1}{y}$, curvatura constante $K = -1$; geodésicas: verticales y semicircunferencias ortogonales al borde. Por un punto exterior pasan infinitas paralelas (falla el quinto postulado: Bolyai–Lobachevski); los triángulos suman menos de π con déficit = área; los círculos crecen exponencialmente. Escenario de las formas modulares y la geometría de grupos. Los tres mundos: $K = +1, 0, -1$; la cosmología pregunta cuál es el nuestro (medido: compatible con 0).

36.11. Cierre del temario: los subtemas restantes

36.11.1. Variedades formalmente

Definición 36.63. Variedad de dimensión n: espacio cubierto por *cartas* $\varphi_\alpha : U_\alpha \to \mathbb{R}^n$ con transiciones $\varphi_\beta \circ \varphi_\alpha^{-1}$ diferenciables; el conjunto de cartas es un *atlas*.

Idea clave

Qué dice esto en cristiano. Una variedad es un espacio que, mirado de cerca, se parece a $\mathbb{R}^n$, aunque de lejos pueda estar retorcido o cerrado sobre sí mismo. La palabra "carta" está tomada literalmente de la cartografía: no existe un mapa plano fiel de toda la Tierra, pero sí un atlas de mapas locales que encajan entre sí. La condición sobre las transiciones es exactamente eso: donde dos mapas se solapan, la traducción de uno a otro tiene que ser suave. Todo lo que se define en una variedad debe poder expresarse carta a carta y dar el mismo resultado en la zona común; ésa es la regla del juego de toda la geometría diferencial.

Ejemplo 36.64 (Variedades y no variedades)**. (a) La circunferencia.** S^1 necesita al menos dos cartas: una carta no basta, porque S^1 no es homeomorfa a ningún abierto de $\mathbb{R}$. Con las dos proyecciones estereográficas desde el polo norte y desde el polo sur se cubre todo, y la transición es $t \mapsto 1/t$ en $\mathbb{R} \setminus \{0\}$, que es suave.

(b) La esfera. S^2 también se cubre con dos cartas estereográficas, con transición $z \mapsto 1/\bar{z}$ vista en $\mathbb{C}$. Con coordenadas esféricas (θ, φ) haría falta más cuidado, porque en los polos la

carta degenera: la métrica $R^2 \sin^2\theta$ se anula y φ deja de estar definida. Eso es un defecto de la carta, no de la esfera.

(c) El toro y el grupo de rotaciones. $T^2 = S^1 \times S^1$ es una variedad de dimensión 2; $SO(3)$ lo es de dimensión 3 y resulta ser difeomorfa a $\mathbb{R}P^3$. El grupo $SU(2)$ es difeomorfo a S^3 y recubre a $SO(3)$ dos a uno, hecho que en física se traduce en que un espinor cambia de signo al girar 360 grados y vuelve a su valor a los 720.

(d) Lo que no es variedad. Dos rectas que se cruzan en el origen no forman una variedad: cerca del cruce el espacio tiene cuatro ramas y no se parece a $\mathbb{R}$. Un cono con vértice tampoco lo es en el vértice. Y el conjunto de Cantor no lo es en ningún punto. La condición de parecerse localmente a $\mathbb{R}^n$ es fuerte y excluye todas las singularidades.

La esfera: dos cartas estereográficas con transición $\mathbf{u} \mapsto \frac{\mathbf{u}}{\|\mathbf{u}\|^2}$; los "defectos" de (ϕ, θ) en los polos son de la carta, no de la esfera. Los tensores se definen carta a carta con la ley de transformación pegando consistentemente.

36.11.2. Vectores tangentes como derivaciones

$X =$ operador con Leibniz: $X(fg) = X(f)g + fX(g)$; en coordenadas $X = v^i \partial_i$: la base $\{\partial_i\}$ es la definición, y la ley contravariante es la regla de la cadena. *Pushforward* F_* (jacobiana) arrastra vectores; *pullback* F^* trae formas: $\int_M F^*\omega = \int_{F(M)} \omega$ —el cambio de variables es un pullback—.

36.11.3. Conos de luz y geodésicas nulas

$g(v,v) < 0$ temporal, > 0 espacial, $= 0$ nulo (cono de luz). Geodésicas nulas = rayos de luz. La estructura causal es la disposición de los conos: en Schwarzschild se inclinan hacia la masa y en $r = r_s$ apuntan adentro: *esa* es la definición del horizonte. La desviación de la luz (Eddington 1919, $1.75''$) es una geodésica nula con nuestras conservadas de Killing.

36.11.4. Christoffel esféricos en $\mathbb{R}^3$

$$\Gamma^{\rho}_{\phi\phi} = -\rho, \quad \Gamma^{\rho}_{\theta\theta} = -\rho \sin^2\phi, \quad \Gamma^{\phi}_{\rho\phi} = \Gamma^{\theta}_{\rho\theta} = \frac{1}{\rho}, \quad \Gamma^{\phi}_{\theta\theta} = -\sin\phi\cos\phi, \quad \Gamma^{\theta}_{\phi\theta} = \cot\phi.$$

Con ρ fijo: la esfera; con $\phi = \frac{\pi}{2}$: polares; $\sqrt{g} = \rho^2 \sin\phi$ reproduce el laplaciano esférico; y el Riemann es idénticamente cero: $\mathbb{R}^3$ plano en disfraz.

36.11.5. Isometrías y espacios homogéneos

Las isometrías forman grupo (dimensión 3 para plano, esfera y $\mathbb{H}$: el máximo $\frac{n(n+1)}{2}$ en $n = 2$); curvatura constante $\iff$ simetría máxima, y solo hay tres modelos (Killing–Hopf). El principio cosmológico fuerza la métrica FRW con un solo $a(t)$ y $k \in \{-1, 0, 1\}$: la simetría reduce las diez ecuaciones de Einstein a una EDO.

36.12. Profundización final: curvatura con significado y ondas gravitacionales

36.12.1. Qué mide cada curvatura

Seccional $K(\Pi)$: Gauss de la superficie de geodésicas tangentes a Π (determina todo el Riemann). *Ricci*: promedio direccional; defecto de volumen de conos geodésicos ($1-\frac{1}{6}R_{ij}v^iv^jr^2$): la gravedad enfoca —por eso Ricci está en Einstein, y Raychaudhuri lleva a los teoremas de singularidad—. *Escalar*: defecto de volumen de bolas.

Teorema 36.65 (Bonnet–Myers)**.** $\mathrm{Ric} \geq (n-1)k > 0 \Rightarrow$ *compacta con diámetro* $\leq \pi/\sqrt{k}$*: curvatura positiva encierra.*

Ejemplo 36.66 (Curvatura positiva y tamaño finito)**. (a) El caso óptimo.** La esfera unidad S^n tiene $\mathrm{Ric} = (n-1)g$, es decir $k = 1$, y su diámetro es exactamente π, la longitud de un meridiano de polo a polo. La cota del teorema se alcanza, y un resultado posterior de Cheng dice que ésa es la única manera de alcanzarla: si $\mathrm{Ric} \geq (n-1)k$ y el diámetro vale $\pi/\sqrt{k}$, entonces la variedad *es* la esfera de radio $1/\sqrt{k}$.

(b) Una esfera más pequeña. Para S^n de radio R se tiene $k = 1/R^2$ y el diámetro es $\pi R = \pi/\sqrt{k}$: coherente. Cuanta más curvatura, menos espacio; la curvatura positiva actúa como una fuerza que cierra el espacio sobre sí mismo.

(c) Con curvatura nula o negativa falla. El plano $\mathbb{R}^2$ tiene $\mathrm{Ric} = 0$ y diámetro infinito; el plano hiperbólico tiene curvatura -1 y también es infinito. La hipótesis $k > 0$ estricta no es un tecnicismo: es lo que fuerza a las geodésicas a reencontrarse.

(d) La consecuencia topológica. Del teorema se sigue que el recubrimiento universal también es compacto, y por tanto que π_1 es *finito*. Eso descarta de un plumazo que el toro admita ninguna métrica de curvatura de Ricci positiva, porque su grupo fundamental es $\mathbb{Z}^2$, infinito. Es un ejemplo perfecto de cómo una hipótesis sobre curvatura, que es local y analítica, impone una restricción global y topológica.

36.12.2. Segunda forma fundamental

II mide la curvatura extrínseca; autovalores = curvaturas principales κ_1, κ_2 (direcciones ortogonales: espectral). $K = \kappa_1\kappa_2 = \frac{\det \mathrm{II}}{\det \mathrm{I}}$ (Gauss: coincide con la intrínseca: Egregium preciso), $H = \frac{\kappa_1+\kappa_2}{2}$. Esfera: $\mathrm{II} = \frac{1}{R}g$; cilindro: $K = 0$ pero $H \neq 0$; $H = 0$: superficies minimales (películas de jabón: cada punto una silla perfecta, Euler–Lagrange del área). Codazzi completa las condiciones de integrabilidad: el Frenet bidimensional.

36.12.3. Ondas gravitacionales

Linealizando Einstein ($g = \eta + h$) y fijando gauge: $\Box h_{\mu\nu} = 0$: la métrica ondula a velocidad c con dos polarizaciones (+ y ×) que estiran anillos de partículas en cuadraturas. Efecto: $\frac{\Delta L}{L} \sim \frac{h}{2}$; LIGO mide $h \sim 10^{-21}$ (milésima de protón en 4 km): la medición más fina de la historia. El factor $\frac{G}{c^4}$ de la fórmula del cuadrupolo explica por qué hicieron falta agujeros negros: el espaciotiempo es rígido.

36.12.4. El mapa exponencial

$\exp_p(v)$: caminar una unidad por la geodésica de velocidad v. Difeomorfismo local (función inversa), rectas $\to$ geodésicas, lema de Gauss (esferas geodésicas ortogonales a radios): coordenadas polares geodésicas con $f \approx r - \frac{K}{6} r^3$ delatando la curvatura. En la esfera es el mapa de la ONU; en grupos de matrices es e^A: exponencial matricial, flujo geodésico y cartas normales son un solo concepto.

36.13. Última capa: Lie, Cartan y Friedmann

36.13.1. Grupos y álgebras de Lie

Definición 36.67. Un *grupo de Lie* es un grupo que es a la vez variedad, con operaciones suaves: $SO(3)$ (rotaciones), $SU(2)$, el grupo de Lorentz, los euclídeos. Su *álgebra de Lie* $\mathfrak{g}$ es el espacio tangente en la identidad, con el *corchete* $[X, Y] = XY - YX$ (para matrices), y la exponencial e^{tX} reconstruye el grupo cerca de la identidad.

Ejemplo 36.68 (Grupos de Lie y sus álgebras)**. (a) Los más pequeños.** S^1 visto como los complejos de módulo 1 es un grupo de Lie de dimensión 1, abeliano y compacto. $(\mathbb{R}, +)$ es de dimensión 1, abeliano y no compacto. Ambos tienen álgebra de Lie $\mathbb{R}$ con corchete nulo, lo que muestra que el álgebra sólo determina al grupo localmente.

(b) Las rotaciones. $SO(3)$ tiene dimensión 3: hacen falta tres parámetros (por ejemplo los ángulos de Euler) para especificar una rotación. Su álgebra de Lie está formada por las matrices antisimétricas 3×3, y el corchete es el conmutador $[A, B] = AB - BA$, que sobre los generadores reproduce exactamente el producto vectorial: $[L_x, L_y] = L_z$.

(c) La exponencial. La aplicación $\exp\colon \mathfrak{g} \to G$ conecta el álgebra con el grupo. Para $A = \left(\begin{smallmatrix} 0 & -t \\ t & 0 \end{smallmatrix}\right)$ sale $\exp(A) = \left(\begin{smallmatrix} \cos t & -\sin t \\ \sin t & \cos t \end{smallmatrix}\right)$: las matrices antisimétricas se convierten en rotaciones. Comprobación en $t = \pi/2$: da la rotación de noventa grados, como debe.

(d) Dimensiones que conviene recordar. $U(n)$ tiene dimensión real n^2, $SU(n)$ tiene $n^2 - 1$, $SO(n)$ tiene $n(n-1)/2$ y $GL(n, \mathbb{R})$ tiene n^2. Así, $SU(2)$ tiene dimensión 3 igual que $SO(3)$, y de hecho son localmente idénticos: el primero recubre al segundo dos a uno. Ese detalle es la razón de que el espín exista.

Ejemplo completo, $\mathfrak{so}(3)$: matrices antisimétricas 3×3, base L_x, L_y, L_z (generadores de giros por eje) con

$$[L_x, L_y] = L_z, \quad [L_y, L_z] = L_x, \quad [L_z, L_x] = L_y :$$

¡el producto vectorial! ($\mathfrak{so}(3) \cong (\mathbb{R}^3, \times)$: por eso el momento angular es un vector solo en 3D). La no conmutatividad de las rotaciones *es* el corchete: girar 90° en x y luego z difiere de al revés exactamente en un giro generado por $[L_x, L_z]$. Los campos de Killing de una variedad forman un álgebra de Lie con el corchete de campos: la simetría continua tiene estructura algebraica, y las representaciones de estas álgebras (capítulo 25 en versión continua) clasifican las partículas elementales: el espín es una representación de $\mathfrak{su}(2)$.

36.13.2. El método de las formas de Cartan

Alternativa eficiente a los Christoffel: elegir un *marco ortonormal* de 1-formas e^a (con $\mathrm{d}s^2 = \sum (e^a)^2$) y resolver las *ecuaciones de estructura*

$$\mathrm{d}e^a = -\omega^a{}_b \wedge e^b, \qquad \Omega^a{}_b = \mathrm{d}\omega^a{}_b + \omega^a{}_c \wedge \omega^c{}_b,$$

donde ω (conexión) es antisimétrica y Ω (curvatura) da el Riemann. Ejemplo, la esfera de radio R: $e^1 = R\,\mathrm{d}\phi$, $e^2 = R\sin\phi\,\mathrm{d}\theta$; de $\mathrm{d}e^2 = R\cos\phi\,\mathrm{d}\phi\wedge\mathrm{d}\theta$ se lee $\omega^2{}_1 = \cos\phi\,\mathrm{d}\theta$, y

$$\Omega^2{}_1 = \mathrm{d}\omega^2{}_1 = -\sin\phi\,\mathrm{d}\phi\wedge\mathrm{d}\theta = \frac{1}{R^2}\,e^1\wedge e^2 :$$

$K = \frac{1}{R^2}$ en *tres líneas* contra la página de Christoffel. El cálculo exterior del capítulo 11 puesto a producir geometría: así se computan las curvaturas de Schwarzschild y de los espacios de la física en la práctica profesional.

36.13.3. Las ecuaciones de Friedmann, deducidas

Con la métrica FRW $\mathrm{d}s^2 = -c^2\mathrm{d}t^2 + a(t)^2\big[\frac{\mathrm{d}r^2}{1-kr^2} + r^2\mathrm{d}\Omega^2\big]$ y el fluido perfecto, las ecuaciones de Einstein se reducen a

$$\Big(\frac{\dot a}{a}\Big)^2 = \frac{8\pi G}{3}\rho - \frac{kc^2}{a^2}, \qquad \frac{\ddot a}{a} = -\frac{4\pi G}{3}\Big(\rho + \frac{3p}{c^2}\Big) :$$

una EDO de primer orden (energía) y una de segundo (aceleración) para el factor de escala. Lecturas inmediatas: materia ordinaria ($\rho + 3p/c^2 > 0$) desacelera —de ahí los teoremas de Big Bang: hacia atrás, $a \to 0$ en tiempo finito—; una constante cosmológica ($p = -\rho c^2$) acelera, que es lo observado desde 1998; y el destino (expansión eterna o recolapso) depende de k y del contenido: la pregunta más grande que existe, escrita en el lenguaje de los capítulos 12 y 36 de este libro. La historia térmica del universo (nucleosíntesis, fondo de microondas) es la solución de estas dos líneas con la termodinámica acoplada.

36.13.4. Kerr y la termodinámica de agujeros negros

La solución de Kerr (1963) describe el agujero negro *rotante* (el genérico astrofísico): además del horizonte tiene una *ergosfera* donde nada puede quedarse quieto (el arrastre del espacio obliga a corrotar) y de la que se puede extraer energía de rotación (proceso de Penrose: hasta el 29 %: el motor de los cuásares). Con Hawking, el área del horizonte nunca decrece —un "segundo principio"— y la analogía se vuelve identidad: los agujeros negros tienen temperatura $T \sim \frac{1}{M}$ y entropía $S = \frac{k_B c^3 A}{4G\hbar}$ (un cuarto del área en unidades de Planck), radiando cuánticamente. Que la entropía sea *área* y no volumen es el germen del principio holográfico: la frontera actual de la física teórica, escrita con la geometría de este capítulo.

36.14. Ronda de temarios: desviación geodésica

Los programas de relatividad general ponen la *ecuación de desviación geodésica* como el corazón físico del tensor de curvatura: es la pieza que hace medible a $R^i{}_{jkl}$, y nos la debíamos en forma de ecuación.

36.14.1. La ecuación de Jacobi

Considérese una familia de geodésicas $x(\tau, s)$: para cada s fijo, $\tau \mapsto x(\tau, s)$ es geodésica; $u = \partial_\tau x$ es la velocidad y $\xi = \partial_s x$ el *vector de separación* entre geodésicas vecinas.

Teorema 36.69 (Desviación geodésica)**.**

$$\frac{D^2 \xi^{\,\mu}}{\mathrm{d}\tau^2} = -R^{\mu}{}_{\nu\alpha\beta}\, u^\nu\, \xi^\alpha u^\beta .$$

Demostración en dos líneas conceptuales. Como ∂_τ y ∂_s conmutan, $\nabla_u \xi = \nabla_\xi u$ (la torsión de Levi-Civita es cero). Entonces

$$\nabla_u \nabla_u \xi = \nabla_u \nabla_\xi u = \nabla_\xi \underbrace{\nabla_u u}_{=0 \ \text{(geodésica)}} + [\nabla_u, \nabla_\xi]\, u,$$

y el conmutador de derivadas covariantes es, por definición, la curvatura: $[\nabla_u, \nabla_\xi]u = R(u, \xi)u$, cuyas componentes dan el enunciado (con el signo de nuestro convenio). La ecuación entera es la afirmación "la curvatura mide el fallo de conmutación" promovida a ley dinámica. □ □

Ejemplo 36.70 (La esfera, cuantitativamente)**.** En la esfera de radio R ($K = \frac{1}{R^2}$), para geodésicas parametrizadas por longitud de arco la ecuación se reduce a la escalar

$$\xi'' = -K\,\xi \qquad \Longrightarrow \qquad \xi(\tau) = \xi_0 \cos \frac{\tau}{R} :$$

dos meridianos que salen paralelos del ecuador oscilan: se acercan, se cruzan en el polo (a distancia $\frac{\pi R}{2}$, donde cos se anula) y se reabren. En el plano ($K = 0$): $\xi'' = 0$, separación lineal, Euclides. En el hiperbólico ($K = -1$): $\xi = \xi_0 \cosh \tau$, separación *exponencial*: la divergencia de geodésicas que hace caótico el billar hiperbólico. Los tres comportamientos —oscilar, lineal, explotar— son el trío elíptico/parabólico/hiperbólico una vez más.

Observación 36.71 (Mareas, y cómo LIGO escucha). En el límite newtoniano $R^i{}_{0j0} \approx \partial_i \partial_j \Phi$, y la ecuación de Jacobi se convierte en $\ddot{\xi}^i = -\partial_i \partial_j \Phi\, \xi^j$: la *fuerza de marea*. Para la Tierra en el campo lunar, $\partial_r^2 \Phi = -\frac{2GM}{r^3}$ estira a lo largo de la línea Tierra–Luna y comprime en las transversales (traza nula: ¡$\Delta\Phi = 0$ en el vacío!): por eso hay *dos* mareas altas diarias, en la cara cercana y en la lejana. Y una onda gravitacional es exactamente un $R^i{}_{0j0}(t)$ oscilante viajando: LIGO mide la desviación geodésica entre sus espejos flotantes, $\frac{\delta L}{L} \sim 10^{-21}$. La ecuación de esta sección es, literalmente, el principio de funcionamiento del detector: la curvatura dejó de ser abstracta el 14 de septiembre de 2015.

Idea clave

La cadena completa del capítulo, ahora cerrada: la métrica da los símbolos de Christoffel (Koszul), estos el transporte paralelo, su holonomía la curvatura, y la curvatura vuelve al mundo físico como aceleración relativa de caídas libres vecinas. Einstein no postuló "el espacio se curva" como metáfora: postuló que la gravedad *es* desviación geodésica, y las ecuaciones de campo eligen qué curvatura produce cada distribución de materia.

36.15. Cierre definitivo: Hopf–Rinow, Frobenius y Gauss–Bonnet

Tres teoremas estructurales que completan el capítulo: cuándo las geodésicas llegan a donde deben, cuándo un sistema de planos se integra en superficies, y la joya que ata curvatura con topología, ahora con demostración.

36.15.1. Hopf–Rinow: la completitud es una sola

Teorema 36.72 (Hopf–Rinow)**.** *Para una variedad riemanniana conexa, son equivalentes: (i) alguna* $\exp_p$ *está definida en todo el tangente (las geodésicas desde p viven para siempre); (ii) lo mismo para todo p; (iii) el espacio métrico* (M, d) *es completo (las sucesiones de Cauchy convergen); (iv) los cerrados y acotados son compactos. Y cualquiera de ellas implica:* todo par de puntos se une por una geodésica minimizante*.*

Idea de la demostración. El corazón es (i)⇒minimizante: para llegar de p a q, dispárese la geodésica en la dirección "más prometedora" (la que minimiza la distancia al objetivo sobre una esferita geodésica alrededor de p: existe por compacidad de la esferita) y pruébese que la propiedad "he recorrido t y estoy a distancia $d(p,q) - t$ de q" se mantiene: el conjunto de tiempos donde vale es no vacío, cerrado (continuidad) y abierto (repetir el argumento en el punto alcanzado): es todo $[0, d(p,q)]$ por conexidad: el *método de continuidad*, la misma lógica abierto-cerrado-no vacío del capítulo 9. Las equivalencias métricas salen de comparar bolas metricas con imágenes de bolas por exp. □ □

Ejemplo 36.73. El plano sin el origen no es completo: la geodésica que apunta al hueco muere en tiempo finito, y los puntos $(-1, 0)$ y $(1, 0)$ *no* tienen geodésica minimizante (el camino recto está prohibido y los desvíos por arriba y por abajo aprietan hacia longitud 2 sin alcanzarla: el ínfimo no se realiza). La esfera y el hiperbólico son completos: allí la promesa de Hopf–Rinow es total. La hipótesis de completitud, que parecía decorado analítico, es exactamente lo que compra la existencia de los caminos óptimos: sin ella, la geometría variacional se queda sin testigos.

36.15.2. El teorema de Frobenius: cuándo los planos se pegan en hojas

Un campo de *planos* (una *distribución* k-dimensional $\mathcal{D}$: en cada punto, un subespacio k-dimensional del tangente) ¿es el campo tangente de una familia de superficies? Localmente, ¿existen "hojas" cuyas tangentes sean exactamente los planos dados?

Teorema 36.74 (Frobenius)**.** $\mathcal{D}$ *es integrable (por cada punto pasa una subvariedad* k*-dimensional tangente a* $\mathcal{D}$*) si y solo si es* involutiva*: el corchete de Lie de campos tangentes a* $\mathcal{D}$ *vuelve a ser tangente a* $\mathcal{D}$*:* $[X, Y] \in \mathcal{D}$ *para* $X, Y \in \mathcal{D}$*.*

Idea. Necesidad: sobre una hoja, X y Y son campos *de la hoja* y su corchete también (el corchete es intrínseco). Suficiencia: enderezar un campo a la vez. Un campo no nulo siempre se rectifica a ∂_1 (flujo como coordenada: capítulo 9, función inversa); la involutividad

garantiza que al rectificar el primero, los demás pueden elegirse sin componente en ∂_1 y *independientes de* x_1 (aquí entra $[\partial_1, Y] \in \mathcal{D}$), e inducción sobre k concluye con coordenadas en que $\mathcal{D} = \operatorname{gen}(\partial_1, \dots, \partial_k)$: las hojas son los planos coordenados. □ □

Ejemplo 36.75 (Un plano que no se deja)**.** En $\mathbb{R}^3$, la distribución $\mathrm{d}z - y\,\mathrm{d}x = 0$ (planos generados por $X = \partial_x + y\partial_z$ y $Y = \partial_y$) tiene $[X, Y] = -\partial_z \notin \mathcal{D}$: *no integrable*. Moverse solo en direcciones permitidas permite, con maniobras de ida y vuelta (¡el corchete es la maniobra de estacionarse en paralelo!), ganar altura z: no hay superficie que confine el movimiento. Esta "no integrabilidad" es un recurso, no un defecto: es la geometría del control (aparcar un coche: dos controles, tres grados de libertad alcanzables) y de la mecánica no holónoma. Y en el lenguaje de formas: $\mathcal{D} = \ker \omega$ es integrable $\iff \omega \wedge \mathrm{d}\omega = 0$: Frobenius es el pariente no lineal del lema de Poincaré del capítulo 11.

36.15.3. Gauss–Bonnet, demostrado

Teorema 36.76 (Gauss–Bonnet local)**.** *Para un triángulo geodésico T en una superficie, con ángulos interiores α, β, γ:*

$$\alpha + \beta + \gamma - \pi = \iint_T K\,\mathrm{d}A.$$

Demostración. Con las formas de Cartan del capítulo: en un marco ortonormal, la 1-forma de conexión ω_{12} satisface $\mathrm{d}\omega_{12} = -K\,\mathrm{d}A$ (segunda ecuación de estructura, ya demostrada). Stokes (¡capítulo 11!) sobre el triángulo:

$$\iint_T K\,\mathrm{d}A = -\oint_{\partial T} \omega_{12}.$$

La integral de ω_{12} a lo largo de una curva mide cuánto *rota el marco* respecto del transporte paralelo; sobre un lado geodésico, la velocidad no rota (definición de geodésica), así que solo contribuyen los tres vértices, donde la tangente salta el ángulo exterior: $\oint \omega_{12} = \sum(\pi - \text{ángulo interior}) - 2\pi$ (la tangente da una vuelta completa al recorrer el borde: teorema del giro de la tangente). Despejando queda el enunciado. □ □

Idea clave

Qué dice esto en cristiano. En el plano, los ángulos de un triángulo suman exactamente 180 grados. El teorema dice qué ocurre cuando la superficie está curvada: lo que sobra o falta respecto de 180 grados es *la curvatura acumulada dentro del triángulo*. En una esfera los ángulos suman de más, en una silla de montar suman de menos, y la diferencia es medible. Es un resultado asombroso porque relaciona algo que se mide en el borde (tres ángulos) con algo que vive en el interior (la curvatura integrada), y es el antepasado directo de todos los teoremas del tipo Stokes.

Ejemplo 36.77 (Medir la curvatura con un triángulo)**. (a) El octante de la esfera.** Tómese el triángulo de la esfera unidad limitado por dos meridianos separados 90 grados y el ecuador. Sus tres ángulos son rectos, así que suman $3\pi/2$ y el exceso es $3\pi/2 - \pi = \pi/2$. Como $K = 1$, el teorema predice que el área vale $\pi/2$, y en efecto el octante es la octava parte de 4π, es decir $\pi/2$. Cuadra exactamente.

(b) Triángulos pequeños. En la misma esfera, un triángulo de área 0,01 tiene un exceso de 0,01 radianes, unos 0,57 grados: casi imperceptible. Por eso la geometría euclídea funciona tan bien en una habitación y hace falta un continente para notar la curvatura de la Tierra. Gauss lo intentó midiendo el triángulo formado por los montes Hohenhagen, Brocken e Inselsberg, y el exceso quedó dentro del error experimental.

(c) Curvatura negativa. En el plano hiperbólico, con $K=-1$, el resultado se invierte: los ángulos suman *menos* de π y el defecto es igual al área. De ahí sale un hecho curioso: ningún triángulo hiperbólico puede tener área mayor que π, por grandes que sean sus lados, porque el defecto máximo es π (los tres ángulos tendiendo a cero).

(d) La versión global. Integrando sobre toda una superficie compacta se obtiene $\iint_S K\,\mathrm{d}A = 2\pi\chi(S)$. Para la esfera da $4\pi = 2\pi\cdot 2$; para el toro da $0 = 2\pi\cdot 0$, y de ahí se deduce que *ninguna* métrica en el toro puede tener curvatura siempre positiva ni siempre negativa: la curvatura total está fijada por la topología, y el geómetra sólo puede repartirla.

Corolario 36.78 (Gauss–Bonnet global)**.** *Triangulando una superficie compacta S y sumando el teorema local sobre todos los triángulos (cada arista se recorre dos veces y se cancela, cada vértice aporta 2π):*

$$\iint_S K\,\mathrm{d}A = 2\pi(V-A+C) = 2\pi\,\chi(S).$$

Observación 36.79. La cuenta final es idéntica en espíritu a la demostración del teorema de la divergencia: sumar identidades locales y dejar que las fronteras internas se maten por pares; lo que sobrevive es global y *topológico*. En la esfera ($\chi=2$): $\iint K = 4\pi$, exacto para el redondo ($K=1/R^2$ por área $4\pi R^2$) y para *cualquier* deformación abollada: la curvatura se redistribuye pero su total está congelado por la topología. En el toro ($\chi=0$): tanta curvatura positiva como negativa, siempre. Este es el prototipo de todos los "teoremas de índice" de la matemática moderna, y ya habíamos visto a su primo combinatorio: Poincaré–Hopf con los ceros del campo.

36.16. Segunda ronda: Bianchi contraída, polares y el GPS

36.16.1. La identidad que hace consistente a Einstein

Bianchi contraída dos veces:

$$\nabla^\mu\big(R_{\mu\nu} - \tfrac{1}{2}R g_{\mu\nu}\big) = 0\,:$$

el tensor de Einstein conserva *por geometría*. Como el lado material cumple $\nabla^\mu T_{\mu\nu}=0$, las ecuaciones son consistentes automáticamente: el término $-\frac{1}{2}Rg$ que a Einstein le costó años (1913–1915) es exactamente lo que Bianchi exige. Las identidades diferenciales (aquí Bianchi; en electromagnetismo $\mathrm{d}F=0$) son las leyes de conservación en estado puro.

36.16.2. La aceleración en polares, desde Christoffel

La fórmula que se memoriza con dolor,

$$\mathbf{a} = (\ddot r - r\dot\theta^2)\hat e_r + (r\ddot\theta + 2\dot r\dot\theta)\hat e_\theta,$$

es $a^k = \ddot{x}^k + \Gamma^k_{ij}\dot{x}^i\dot{x}^j$ con los Christoffel polares: $a^r = \ddot{r} - r\dot{\theta}^2$, $a^\theta = \ddot{\theta} + \frac{2}{r}\dot{r}\dot{\theta}$ (la componente física lleva el factor r). Los términos "centrífugo" y "de Coriolis" son símbolos de Christoffel: fuerzas ficticias = conexión: la antesala exacta del principio de equivalencia.

36.16.3. El GPS con números

Satélite a 20200 km, $v \approx 3.9$ km/s. Especial: $-\frac{v^2}{2c^2} \approx -8.4\times10^{-11}$: $-7\,\mu$s/día. General: $\frac{GM}{c^2}(\frac{1}{R_\oplus} - \frac{1}{r}) \approx +5.3\times10^{-10}$: $+46\,\mu$s/día. Neto $+38\,\mu$s/día = *11 km de error diario* sin corrección: los relojes se fabrican deliberadamente lentos (10.22999999543 MHz). Cada navegación del teléfono es un experimento de relatividad con resultado positivo.

36.17. Coda: espinores y la doble cubierta

Queda un secreto del grupo de rotaciones: $SO(3)$ tiene una *doble cubierta* $SU(2)$ (matrices unitarias 2×2 de determinante 1 = cuaterniones unitarios del capítulo 25): cada rotación proviene de *dos* elementos $\pm U$, y un lazo de 360° en $SO(3)$ *no* es contráctil (hay que dar 720° para volver de verdad: el truco del cinturón de Dirac lo demuestra con una correa). Los *espinores* son los objetos que se transforman bajo $SU(2)$ y no bajo $SO(3)$: cambian de signo tras una vuelta completa. No son curiosidad topológica: el electrón es un espinor (espín $\frac{1}{2}$), el principio de exclusión de Pauli —y con él toda la química— viene del signo espinorial, y los experimentos de interferometría de neutrones han medido el -1 de las 360° directamente. La cadena capítulo 25 (cuaterniones) → capítulo 36 (rotaciones y topología) → física de la materia es uno de los cierres más finos del libro.

36.18. Temas extra y ejemplos desarrollados

36.18.1. Relatividad especial como geometría

$\mathrm{d}\tau = \mathrm{d}t\sqrt{1 - v^2/c^2}$: dilatación = Pitágoras con signo menos. Gemelos: la recta *maximiza* el tiempo propio (la métrica invierte la triangular): el viajero envejece menos, sin paradoja. Muones cósmicos y GPS (-7μs/día por velocidad, $+45$ por gravedad): sin el $\mathrm{d}s^2$, kilómetros de error.

36.18.2. Cuadrivelocidad y $E = mc^2$

$u^\mu = \gamma(c, \mathbf{v})$ con $g(u,u) = -c^2$; $p^\mu = mu^\mu$: $E = \gamma mc^2 = mc^2 + \frac{mv^2}{2} + \cdots$ —la fórmula más célebre es el término constante de una serie binomial—; $E^2 = (pc)^2 + (mc^2)^2$, con $m = 0$: la luz.

36.18.3. El cono: curvatura concentrada

Papel enrollado: $K = 0$ en todas partes pero holonomía α (el déficit) alrededor del vértice: curvatura delta, análogo exacto del campo del índice. Gauss–Bonnet global: $\iint K = 2\pi\chi$: esfera 4π, toro 0 —curvatura total atada a la topología—.

36.18.4. El tensor energía–momento

$T^{\mu\nu} = (\rho + \frac{p}{c^2})u^\mu u^\nu + p g^{\mu\nu}$: densidad, flujo, tensiones; $\nabla_\mu T^{\mu\nu} = 0$ contiene continuidad y Euler: la hidrodinámica en una divergencia. En FRW da Friedmann: el temario entero en una ecuación.

36.18.5. Triángulos que delatan la curvatura

$\alpha + \beta + \gamma = \pi + KA$: esfera, más de π (el triángulo de dos polos y ecuador: 270° y $KA = \frac{\pi}{2}$ ✓); hiperbólico, menos (área $\leq \pi$). Gauss midió montañas; hoy el fondo cósmico hace el mismo experimento: geo-metría literal.

36.19. Conexión y tensor de Weyl: los objetos que faltaban por definir

El teorema de Levi-Civita afirma que existe una *única conexión* con dos propiedades, y el capítulo nunca dice qué es una conexión cualquiera; y el recuento $20 - 10 = 10$ de las componentes independientes nombra un tensor del que sólo se decía que es "lo que sobra".

Definición 36.80 (Conexión afín)**.** Una *conexión afín* sobre una variedad M es una aplicación que a cada par de campos vectoriales X, Y le asigna otro campo $\nabla_X Y$ —la *derivada covariante de Y en la dirección X*— cumpliendo, para toda función f:

1. $\nabla_{fX_1+X_2} Y = f\,\nabla_{X_1} Y + \nabla_{X_2} Y$ (es *tensorial* en la dirección);
2. $\nabla_X(Y_1 + Y_2) = \nabla_X Y_1 + \nabla_X Y_2$;
3. $\nabla_X(fY) = (Xf)\,Y + f\,\nabla_X Y$ (regla de Leibniz en el argumento derivado).

En coordenadas, una conexión queda determinada por las n^3 funciones Γ^k_{ij} definidas por $\nabla_{\partial_i}\partial_j = \Gamma^k_{ij}\,\partial_k$. Su *torsión* es el campo tensorial $T(X,Y) = \nabla_X Y - \nabla_Y X - [X,Y]$, de componentes $T^k_{ij} = \Gamma^k_{ij} - \Gamma^k_{ji}$.

Observación 36.81 (Los símbolos de Christoffel no son un tensor, y ahora se ve por qué)*.* La propiedad (3) tiene un término $(Xf)Y$ que *no* es lineal sobre funciones, y ése es exactamente el motivo de que los Γ^k_{ij} transformen con un término inhomogéneo al cambiar de coordenadas: no son las componentes de ningún tensor. En cambio la *diferencia* de dos conexiones sí lo es —el término molesto se cancela—, y la torsión también, porque es antisimétrica y el término inhomogéneo es simétrico en i, j. Es la observación que el capítulo hacía sin tener el objeto definido.

Con esto, el teorema de Levi-Civita adquiere sentido: entre *todas* las conexiones —y hay infinitas, pues difieren en un tensor arbitrario de tipo $(1,2)$— existe exactamente una que es *sin torsión* ($T = 0$) y *compatible con la métrica* ($\nabla g = 0$). Sin la definición 36.80 el enunciado cuantificaba sobre un conjunto vacío de significado.

Definición 36.82 (Descomposición de Riemann y tensor de Weyl)**.** En dimensión $n \geq 3$, el tensor de curvatura de Riemann se descompone de manera única como

$$R_{ijkl} = \underbrace{C_{ijkl}}_{\text{Weyl}} + \underbrace{\frac{2}{n-2}\left(g_{i[k}S_{l]j} - g_{j[k}S_{l]i}\right)}_{\text{parte de Ricci}} + \underbrace{\frac{2R}{(n-1)(n-2)}\, g_{i[k}g_{l]j}}_{\text{parte escalar}},$$

donde $S_{ij} = R_{ij}$ es el tensor de Ricci, R la curvatura escalar y los corchetes indican antisimetrización. El *tensor de Weyl* C_{ijkl} es, por construcción, la parte de la curvatura con *todas las trazas nulas*: $g^{ik}C_{ijkl} = 0$.

Observación 36.83 (Qué mide cada trozo, y el recuento $20 = 10 + 10$)*.* El tensor de Riemann tiene, en dimensión 4, exactamente 20 componentes independientes. El de Ricci, siendo simétrico 4×4, tiene 10; y el de Weyl se lleva los 10 restantes. Ése es el recuento que el capítulo daba sin nombrar al segundo protagonista.

La división tiene significado físico directo. Las ecuaciones de Einstein $R_{ij} - \frac{1}{2}Rg_{ij} = 8\pi T_{ij}$ determinan *sólo* la parte de Ricci a partir de la materia local: donde no hay materia, $R_{ij} = 0$ y el Riemann se reduce a su parte de Weyl. Por eso el vacío no es plano: el tensor de Weyl es el campo gravitatorio que se propaga —las ondas gravitacionales son Weyl puro— y el que produce las fuerzas de marea a distancia. Ricci mide cómo la materia *comprime* un volumen; Weyl, cómo lo *estira en una dirección y lo aplasta en otra* sin cambiar el volumen.

Dos propiedades que conviene retener: C es *invariante conforme*, es decir, no cambia al sustituir g por $\mathrm{e}^{2\varphi}g$ (salvo un factor global), y por eso una métrica es conformemente plana si y sólo si $C = 0$ para $n \geq 4$. Y en dimensión 3 el tensor de Weyl es idénticamente nulo: el Riemann queda determinado por el Ricci, y ésa es la razón matemática de que no haya gravedad en el vacío —ni ondas gravitacionales— en un mundo de dos dimensiones espaciales más el tiempo.

36.20. Ejercicios resueltos

Ejercicio 36.1. Demuestra que δ^i_j tiene las mismas componentes en toda base, que la contracción T^i_i es invariante, y que $\partial_i v^i$ *no* es invariante en general (por qué hace falta la divergencia covariante).

Solución. Un $(1,1)$ transforma $T' = A^{-1}TA$. Para la delta: $A^{-1}IA = I$. Para la traza: $\operatorname{tr}(A^{-1}TA) = \operatorname{tr}(TAA^{-1}) = \operatorname{tr} T$ (propiedad cíclica). En cambio $\partial_i v^i$ involucra *derivadas* de las componentes: al transformar, la regla del producto genera el término $v^j \partial'_i[(A^{-1})^i{}_j]$, que solo se anula si el cambio de coordenadas es lineal. En coordenadas curvilíneas ese término sobrevive, y es exactamente lo que compensa el $\Gamma^i_{ij}v^j$ de $\nabla_i v^i$. ■

Ejercicio 36.2. Deriva la métrica polar desde la cartesiana y calcula sus Christoffel.

Solución. De $x = r\cos\theta$, $y = r\sin\theta$: $\mathrm{d}x = \cos\theta\,\mathrm{d}r - r\sin\theta\,\mathrm{d}\theta$, $\mathrm{d}y = \sin\theta\,\mathrm{d}r + r\cos\theta\,\mathrm{d}\theta$; sumando cuadrados, los términos cruzados se cancelan y $\mathrm{d}s^2 = \mathrm{d}r^2 + r^2\mathrm{d}\theta^2$. Con $g = \operatorname{diag}(1, r^2)$, la única derivada no nula es $\partial_r g_{\theta\theta} = 2r$; la fórmula (36.1) da

$$\Gamma^r_{\theta\theta} = \tfrac{1}{2}g^{rr}(-\partial_r g_{\theta\theta}) = -r, \qquad \Gamma^\theta_{r\theta} = \Gamma^\theta_{\theta r} = \tfrac{1}{2}g^{\theta\theta}\partial_r g_{\theta\theta} = \frac{1}{r},$$

y el resto cero. Coordenadas curvas $\Rightarrow \Gamma \neq 0$, aunque el plano no tenga curvatura alguna. ■

Ejercicio 36.3. Escribe las geodésicas del plano en polares y comprueba que las rectas que no pasan por el origen, $r\cos(\theta-\theta_0)=d$, las satisfacen.

Solución. Las ecuaciones son $\ddot r - r\dot\theta^2 = 0$ y $\ddot\theta + \frac{2}{r}\dot r\dot\theta = 0$. Tomemos la recta $x = d$ (es decir $\theta_0 = 0$: $r\cos\theta = d$) parametrizada por $y = t$: entonces $r = \sqrt{d^2+t^2}$, $\theta = \arctan(t/d)$. Calculamos: $\dot r = \frac{t}{r}$, $\ddot r = \frac{1}{r} - \frac{t^2}{r^3} = \frac{d^2}{r^3}$; $\dot\theta = \frac{d}{r^2}$, así que $r\dot\theta^2 = \frac{d^2}{r^3} = \ddot r$ ✓. Y $\ddot\theta = -\frac{2dt}{r^4} = -\frac{2}{r}\cdot\frac{t}{r}\cdot\frac{d}{r^2} = -\frac{2}{r}\dot r\dot\theta$ ✓. Las rectas euclídeas son geodésicas se miren en las coordenadas que se miren: la geometría no depende de la descripción. ■

Ejercicio 36.4. En la esfera, muestra que el ecuador es geodésica y los demás paralelos no.

Solución. Con $\Gamma^{\phi}_{\theta\theta} = -\sin\phi\cos\phi$ (Ejemplo 36.43), la ecuación en ϕ es $\ddot\phi - \sin\phi\cos\phi\,\dot\theta^2 = 0$. En un paralelo $\phi = \phi_0$ constante, $\ddot\phi = 0$, y queda $\sin\phi_0\cos\phi_0\,\dot\theta^2 = 0$: solo se cumple con $\cos\phi_0 = 0$, es decir $\phi_0 = \frac{\pi}{2}$, el ecuador. Por eso las rutas aéreas de larga distancia siguen círculos máximos y no paralelos: son las geodésicas de la Tierra. ■

Ejercicio 36.5. Verifica la compatibilidad métrica $\nabla_r g_{\theta\theta} = 0$ en polares, y explica por qué implica que subir/bajar índices conmuta con ∇.

Solución. Para un $(0,2)$: $\nabla_r g_{\theta\theta} = \partial_r g_{\theta\theta} - 2\Gamma^{l}_{r\theta} g_{l\theta} = 2r - 2\cdot\frac{1}{r}\cdot r^2 = 0$. En general $\nabla g = 0$ significa que la métrica "pasa a través" de la derivada covariante: $\nabla_i(g_{jk}v^k) = g_{jk}\nabla_i v^k$, de modo que da igual bajar el índice antes o después de derivar. Sin esta propiedad, la identificación vector–covector se rompería al derivar y el formalismo sería inconsistente. ■

Ejercicio 36.6. Transporta paralelamente el vector $v = e_\theta|_{\text{punto inicial}}$ a lo largo del paralelo $\phi = \phi_0$ de la esfera unitaria, desde $\theta = 0$ hasta $\theta = 2\pi$, y calcula el ángulo de holonomía.

Solución. Las ecuaciones de transporte a lo largo de la curva ($\dot\phi = 0$, $\dot\theta = 1$) son

$$\dot v^{\phi} + \Gamma^{\phi}_{\theta\theta} v^{\theta} = 0 \Rightarrow \dot v^{\phi} = \sin\phi_0\cos\phi_0\, v^{\theta}, \qquad \dot v^{\theta} + \Gamma^{\theta}_{\theta\phi} v^{\phi} = 0 \Rightarrow \dot v^{\theta} = -\cot\phi_0\, v^{\phi}.$$

Derivando la primera y sustituyendo la segunda: $\ddot v^{\phi} = -\cos^2\phi_0\, v^{\phi}$: ¡oscilador armónico con frecuencia $\omega = \cos\phi_0$! El vector rota (respecto de la base local) con esa frecuencia, y tras la vuelta completa ($\theta = 2\pi$) ha girado

$$\alpha = 2\pi\cos\phi_0.$$

Comprobaciones: en el ecuador ($\phi_0 = \frac{\pi}{2}$) no gira nada ($\alpha = 0$: es geodésica); cerca del polo ($\phi_0 \to 0$) gira casi 2π. Este ángulo es medible mecánicamente: es la precesión del péndulo de Foucault, con ϕ_0 la colatitud del laboratorio —la rotación del plano del péndulo en un día es exactamente la holonomía del transporte paralelo sobre el paralelo terrestre—. ■

Ejercicio 36.7. Con el resultado del Ejemplo 36.43, verifica las ecuaciones de Einstein en el vacío ($T_{\mu\nu} = 0 \Rightarrow R_{\mu\nu} = 0$) *no* se satisfacen para la esfera 2D, y explica por qué no es contradicción.

Solución. Para la esfera, $R_{\phi\phi} = 1 \neq 0$: la esfera no es solución de vacío. No hay contradicción: las ecuaciones de Einstein rigen la métrica del *espaciotiempo* físico de dimensión 4, no la de una superficie arbitraria; la esfera es solo un espacio riemanniano de juguete. (De hecho, en dimensión 2 las ecuaciones de vacío forzarían curvatura nula, y en dimensión 4 admiten soluciones curvas no triviales como Schwarzschild: la gravedad pura necesita al menos 4 dimensiones para existir.) ■

Ejercicio 36.8. Demuestra que $F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu$ es invariante bajo el cambio de gauge $A_\mu \to A_\mu + \partial_\mu \chi$, y que la identidad $\partial_\lambda F_{\mu\nu} + \partial_\mu F_{\nu\lambda} + \partial_\nu F_{\lambda\mu} = 0$ es automática.

Solución. Bajo el cambio: $F_{\mu\nu} \to \partial_\mu(A_\nu + \partial_\nu\chi) - \partial_\nu(A_\mu + \partial_\mu\chi) = F_{\mu\nu} + \partial_\mu\partial_\nu\chi - \partial_\nu\partial_\mu\chi = F_{\mu\nu}$ por Clairaut: los campos físicos no dependen de la elección de potencial. Para la identidad, sustituimos $F = \partial A - \partial A$:

$$\partial_\lambda(\partial_\mu A_\nu - \partial_\nu A_\mu) + \partial_\mu(\partial_\nu A_\lambda - \partial_\lambda A_\nu) + \partial_\nu(\partial_\lambda A_\mu - \partial_\mu A_\lambda) = 0,$$

pues los seis términos se cancelan por pares (cruzadas iguales). Es el $\mathrm{d}\circ\mathrm{d} = 0$ del capítulo 11 con índices: las dos ecuaciones de Maxwell sin fuentes son pura geometría diferencial. ■

36.21. Ejercicios adicionales resueltos

Ejercicio 36.9. Con (36.2): triple producto y $\operatorname{rot}\operatorname{rot}$.

Solución. $[\mathbf{a}\times(\mathbf{b}\times\mathbf{c})]_i = (\delta_{il}\delta_{jm} - \delta_{im}\delta_{jl})a_j b_l c_m = b_i(\mathbf{a}\cdot\mathbf{c}) - c_i(\mathbf{a}\cdot\mathbf{b})$; con ∂: $\operatorname{rot}\operatorname{rot}\mathbf{F} = \nabla(\operatorname{div}\mathbf{F}) - \Delta\mathbf{F}$. Dos formularios en una contracción. ■

Ejercicio 36.10. Killing del plano y conservadas de las rectas.

Solución. $(1,0), (0,1), (-y,x)$: $\dot{x}$, $\dot{y}$, $x\dot{y} - y\dot{x}$: momentos lineal y angular: Noether limpio. ■

Ejercicio 36.11. Verifica $K = -1$ en $\mathbb{H}$.

Solución. $R^x{}_{yxy} = -\frac{1}{y^2}$ (los términos cuadráticos se cancelan), $R_{xyxy} = -\frac{1}{y^4}$, y $K = \frac{R_{xyxy}}{g_{xx}g_{yy}} = -1$: constante pese a la métrica variable. ■

Ejercicio 36.12. Verticales de $\mathbb{H}$ como geodésicas y distancia $(0,a)$–$(0,b)$.

Solución. $\dot{x} = 0$ satisface la ecuación en x; $d = \int_a^b \frac{\mathrm{d}y}{y} = \ln\frac{b}{a}$: el borde está a distancia infinita, duplicar la altura siempre cuesta $\ln 2$. ■

Ejercicio 36.13. Órbitas de Schwarzschild por Killing.

Solución. $E = (1 - \frac{r_s}{r})c^2\dot{t}$, $L = r^2\dot{\theta}$; la normalización da

$$\frac{\dot{r}^2}{2} + \frac{L^2}{2r^2} - \frac{GM}{r} - \frac{GML^2}{c^2 r^3} = \text{cte}:$$

Kepler más el término $\frac{1}{r^3}$: precesión de Mercurio ($43''$/siglo), órbitas inestables y captura de luz. ■

Ejercicio 36.14. Transporte paralelo por el paralelo ϕ_0 de la esfera.

Solución. $\ddot{v}^\phi = -\cos^2\phi_0 v^\phi$: rotación con frecuencia $\cos\phi_0$: holonomía $2\pi\cos\phi_0$ por vuelta —el péndulo de Foucault: su precesión diaria es esta holonomía en el paralelo terrestre—. ■

36.22. Taller de ejemplos explicados: midiendo sobre la esfera

Idea clave

Intuición primero. Todo lo abstracto de este capítulo (métrica, Christoffel, curvatura) se vuelve tangible en cuanto pones números sobre una superficie concreta. La *métrica* es la regla para medir distancias sobre la esfera; los *símbolos de Christoffel* corrigen cómo giran los ejes de referencia al moverte por ella. Al calcular la longitud del ecuador y un Christoffel a mano, verás que no hay nada místico: es aritmética con derivadas de la métrica.

Los temas de este capítulo son abstractos, pero se vuelven concretos en cuanto pones números. Trabajemos sobre la esfera de radio R, cuyo elemento de distancia (la "métrica") es

$$\mathrm{d}s^2 = R^2\,\mathrm{d}\theta^2 + R^2\sin^2\theta\,\mathrm{d}\varphi^2,$$

donde θ es la colatitud (desde el polo norte) y φ la longitud. Esta fórmula dice cómo medir longitudes sobre la superficie curva.

36.22.1. Longitud de un paralelo

Ejemplo 36.84. Calcular la longitud del ecuador.

Paso 1. Fijar el camino. El ecuador es $\theta = \dfrac{\pi}{2}$ constante, con φ recorriendo de 0 a 2π. Como θ no cambia, $\mathrm{d}\theta = 0$.

Paso 2. Simplificar la métrica sobre ese camino. Con $\mathrm{d}\theta = 0$ y $\sin\frac{\pi}{2} = 1$:

$$\mathrm{d}s^2 = R^2\sin^2\!\left(\tfrac{\pi}{2}\right)\mathrm{d}\varphi^2 = R^2\,\mathrm{d}\varphi^2 \implies \mathrm{d}s = R\,\mathrm{d}\varphi.$$

Paso 3. Integrar sobre toda la vuelta:

$$L = \int_0^{2\pi} R\,\mathrm{d}\varphi = R\cdot 2\pi = 2\pi R.$$

La circunferencia del ecuador es $2\pi R$, como debía. La métrica funcionó.

36.22.2. Un símbolo de Christoffel de la esfera

Los símbolos de Christoffel Γ corrigen cómo cambian los vectores al moverse sobre una superficie curva. Se calculan con derivadas de la métrica. Aquí va uno, despacio.

Ejemplo 36.85. Calcular $\Gamma^{\theta}_{\varphi\varphi}$ de la esfera.

Paso 1. Anotar las componentes de la métrica: $g_{\theta\theta} = R^2$ (constante) y $g_{\varphi\varphi} = R^2\sin^2\theta$.

Paso 2. Usar la fórmula (para este caso se reduce a):

$$\Gamma^{\theta}_{\varphi\varphi} = -\frac{1}{2g_{\theta\theta}}\,\frac{\partial g_{\varphi\varphi}}{\partial\theta}.$$

Paso 3. Calcular la derivada: $\dfrac{\partial}{\partial\theta}\!\left(R^2\sin^2\theta\right) = R^2\cdot 2\sin\theta\cos\theta$ (regla de la cadena).

Paso 4. Sustituir:

$$\Gamma^{\theta}_{\varphi\varphi} = -\frac{1}{2R^2}\cdot R^2 \cdot 2\sin\theta\cos\theta = -\sin\theta\cos\theta.$$

Este número mide cuánto "se curva hacia el polo" un movimiento a lo largo de un paralelo: es lo que hace que las líneas rectas sobre la esfera (los círculos máximos) parezcan curvarse.

36.23. Ejercicios propuestos

Sin solución desarrollada: solo enunciado y, al final, la respuesta. Varias respuestas son un nombre de teorema; si tu argumento llega al mismo sitio por otro camino, está bien igual.

36.1. Distingue con un ejemplo un vector contravariante de uno covariante y di cómo cambian sus componentes al cambiar de base.

36.2. ¿Cuántas componentes tiene un tensor de tipo $(1,2)$ en dimensión 4?

36.3. Escribe la métrica euclídea de $\mathbb{R}^2$ en coordenadas polares.

36.4. Calcula el determinante de la métrica esférica $\mathrm{d}s^2 = R^2\mathrm{d}\theta^2 + R^2\sin^2\theta\,\mathrm{d}\varphi^2$ y el elemento de área.

36.5. Halla $\Gamma^{\theta}_{\varphi\varphi}$ y $\Gamma^{\varphi}_{\theta\varphi}$ para la esfera de radio R.

36.6. Calcula la curvatura escalar de la esfera de radio R.

36.7. ¿Cuántas componentes independientes tiene el tensor de Riemann en dimensión 2? ¿Y en dimensión 4?

36.8. Escribe la ecuación de las geodésicas y di qué son las geodésicas de la esfera.

36.9. ¿Es el símbolo de Christoffel un tensor? Justifica con su ley de transformación.

36.10. Sube y baja índices: si $g_{ij} = \mathrm{diag}(1, r^2)$ y $V^i = (1, 1)$, calcula V_i y $\|V\|^2$.

36.11. ¿Qué distingue la derivada covariante de la derivada parcial y por qué hace falta?

36.12. Enuncia qué mide el tensor de Ricci frente al de Riemann completo.

Respuestas

1. El vector velocidad es contravariante (sus componentes se transforman con $\partial x'^i/\partial x^j$); el gradiente es covariante (se transforma con la matriz inversa, $\partial x^j/\partial x'^i$). **2.** $4^3 = 64$. **3.** $\mathrm{d}s^2 = \mathrm{d}r^2 + r^2\mathrm{d}\theta^2$, es decir $g = \mathrm{diag}(1, r^2)$. **4.** $\det g = R^4\sin^2\theta$, luego $\mathrm{d}A = R^2\sin\theta\,\mathrm{d}\theta\,\mathrm{d}\varphi$ (y el área total es $4\pi R^2$). **5.** $\Gamma^{\theta}_{\varphi\varphi} = -\sin\theta\cos\theta$ y $\Gamma^{\varphi}_{\theta\varphi} = \cot\theta$. **6.** $S = 2/R^2$: positiva y constante, y tiende a 0 cuando $R \to \infty$ (la esfera enorme se ve plana). **7.** En dimensión 2, una sola componente independiente. En dimensión 4, veinte. **8.** $\ddot{x}^i + \Gamma^i_{jk}\dot{x}^j\dot{x}^k = 0$. En la esfera son los círculos máximos. **9.** No es tensor: su ley de transformación tiene un término extra con segundas derivadas. Por eso puede anularse en un punto eligiendo coordenadas normales, cosa imposible para un tensor no nulo. **10.** $V_i = (1, r^2)$ y $\|V\|^2 = V^iV_i = 1 + r^2$. **11.** La parcial de las componentes no es tensorial porque la base también varía; la covariante añade los Γ que corrigen esa variación, y el resultado sí se transforma como tensor. **12.** Riemann guarda toda la curvatura, incluidas las deformaciones de forma (mareas); Ricci es su traza y mide solo cómo cambia el volumen de una bolita de geodésicas.

Capítulo 37

Topología algebraica

Lo que necesitas traer

Del capítulo 28: espacios topológicos, continuidad, conexión, compacidad, homotopía y el grupo fundamental, que aquí se retoma con más profundidad. Del capítulo 23: grupos, cocientes y grupos abelianos finitamente generados, que son los que clasifican la homología.

La topología general dice cuándo dos espacios son homeomorfos, pero demostrar que *no* lo son es mucho más difícil: habría que descartar todas las biyecciones continuas posibles. La topología algebraica resuelve el problema convirtiendo espacios en objetos algebraicos, de manera que espacios homeomorfos den grupos isomorfos. Entonces, para probar que dos espacios son distintos, basta con calcular sus grupos y ver que difieren, que es una comparación finita.

37.1. Invariantes: la característica de Euler

Definición 37.1. Para una triangulación de una superficie compacta con V vértices, A aristas y C caras, la *característica de Euler* es $\chi = V - A + C$. No depende de la triangulación elegida, sino solo del espacio, y por tanto es un *invariante topológico*. Para una superficie orientable de *género* g (número de asas) se cumple $\chi = 2 - 2g$.

Idea clave

Qué dice esto en cristiano. La característica de Euler es un número entero que se calcula contando trozos y que no cambia al deformar el espacio. Su valor es el mismo por muchos vértices o caras que se añadan, porque cada subdivisión añade tantas caras como aristas de más y el balance se cancela. Que sea invariante convierte una pregunta geométrica difícil (¿son el mismo espacio?) en una comparación de números. Si dos superficies tienen χ distinto, no son homeomorfas, y no hace falta examinar ninguna aplicación entre ellas.

Ejemplo 37.2 (Contar caras y distinguir superficies)**. (a) Los cinco sólidos platónicos.** Tetraedro $4-6+4=2$; cubo $8-12+6=2$; octaedro $6-12+8=2$; dodecaedro $20-30+12=2$; icosaedro $12-30+20=2$. Los cinco dan 2, porque los cinco son topológicamente una esfera. La fórmula de Euler no distingue entre ellos: los ve como el mismo objeto.

(b) La familia de las superficies orientables. Con la fórmula $\chi = 2-2g$: la esfera ($g=0$) da 2, el toro ($g=1$) da 0, el toro doble ($g=2$) da -2 y el triple ($g=3$) da -4. Como los cuatro números son distintos, las cuatro superficies son topológicamente distintas, y eso queda demostrado con cuatro restas.

(c) Las no orientables. La botella de Klein tiene $\chi=0$, igual que el toro, y sin embargo no son homeomorfas: la primera no es orientable. La característica sola no basta, hace falta añadir la orientabilidad. El plano proyectivo tiene $\chi=1$.

(d) La clasificación completa. Toda superficie compacta y conexa es homeomorfa a una esfera con g asas (si es orientable) o a una esfera con k bandas de Möbius pegadas (si no lo es), y el par formado por χ y la orientabilidad la determina por completo. Es uno de los pocos problemas de clasificación de la matemática que está totalmente resuelto; en dimensión tres costó hasta la demostración de la conjetura de Poincaré por Perelmán en 2003, y en dimensión cuatro sigue abierto.

Cuidado

[Un invariante solo sirve en una dirección] Si dos espacios tienen invariantes distintos, son distintos. Si los tienen iguales, *no se concluye nada*: el toro y la botella de Klein comparten $\chi=0$ y no son homeomorfos. Por eso se construyen familias enteras de invariantes cada vez más finos, y aun así ninguna colección conocida distingue todos los espacios. Segundo aviso: la fórmula $V-A+C=2$ exige que el poliedro sea topológicamente una esfera; para un poliedro con un agujero central (un marco de cuadro) sale $\chi=0$, y quien la aplique a ciegas obtendrá un resultado falso.

37.2. Caminos, lazos y concatenación: por qué π_1 es un grupo y cuánto vale el de la circunferencia

El invariante más antiguo de la topología algebraica se construye con un material humildísimo: funciones continuas de un intervalo en un espacio. Todo el resto del capítulo lo usa, y en el resto del capítulo aparece ya montado. Aquí se monta pieza a pieza, porque cada palabra —camino, concatenación, clase de homotopía— tiene una definición exacta, y porque la afirmación "las clases de homotopía de lazos forman un grupo" no es evidente en absoluto: hay que comprobar que la operación está bien definida sobre clases y que cumple los tres axiomas, y ninguna de las cuatro comprobaciones es automática.

Caminos y su álgebra

Definición 37.3 (Camino, camino constante, camino inverso)**.** Sea X un espacio topológico. Un *camino* en X de x a y es una aplicación continua $\gamma:[0,1]\to X$ con $\gamma(0)=x$ y $\gamma(1)=y$. Se dice que x es el *origen* y y el *extremo*. Dos casos especiales tienen nombre propio:

- el *camino constante* en x es $c_x(s) = x$ para todo $s \in [0,1]$;
- el *camino inverso* de γ es $\overline{\gamma}(s) = \gamma(1-s)$, que recorre el mismo trazado en sentido contrario y va de y a x.

Un *lazo* con base en x_0 es un camino con $\gamma(0) = \gamma(1) = x_0$.

Cuidado

Un camino *no* es un subconjunto de X: es la función, no su imagen. El camino que recorre la circunferencia una vez y el que la recorre tres veces tienen exactamente la misma imagen y son caminos distintos, y toda la teoría vive de esa diferencia. Tampoco se pide que γ sea inyectiva: un camino puede pararse, volver sobre sus pasos o pasar cien veces por el mismo punto.

Definición 37.4 (Concatenación)**.** Sean γ un camino de x a y y δ un camino de y a z —obsérvese que el extremo del primero ha de ser el origen del segundo—. Su *concatenación* $\gamma * \delta$ es el camino de x a z dado por

$$(\gamma * \delta)(s) = \begin{cases} \gamma(2s), & 0 \le s \le \frac{1}{2}, \\ \delta(2s-1), & \frac{1}{2} \le s \le 1. \end{cases}$$

Es decir: se recorre γ al doble de velocidad en la primera mitad del tiempo y δ al doble de velocidad en la segunda.

Observación 37.5 (Por qué la fórmula define de verdad una función continua). Las dos expresiones se solapan en $s = \frac{1}{2}$, y allí valen $\gamma(1) = y$ y $\delta(0) = y$: la definición es coherente. Y es continua por el *lema del pegado*: si $[0,1] = A \cup B$ con A y B cerrados y una función es continua restringida a cada uno y coincide en $A \cap B$, entonces es continua en todo el intervalo. Aquí $A = [0, \frac{1}{2}]$ y $B = [\frac{1}{2}, 1]$. Este lema se usará cuatro veces más en la sección sin volver a mencionarlo.

Cuidado

[La concatenación no es asociativa] Con caminos concretos, $(\gamma * \delta) * \varepsilon$ y $\gamma * (\delta * \varepsilon)$ son funciones *distintas*: en la primera, γ ocupa el intervalo de tiempo $[0, \frac{1}{4}]$; en la segunda, $[0, \frac{1}{2}]$. Recorren lo mismo con horarios distintos. Tampoco $\gamma * c_y$ es igual a γ, ni $\gamma * \overline{\gamma}$ es el camino constante: es el camino que va y vuelve. Por eso los caminos *no* forman un grupo, y por eso hay que pasar a clases de homotopía. Todo lo que sigue existe para arreglar exactamente estos tres defectos.

Homotopía de caminos

Definición 37.6 (Homotopía de caminos, con extremos fijos)**.** Sean γ_0, γ_1 dos caminos en X *con los mismos extremos*, de x a y. Son *homótopos como caminos*, y se escribe $\gamma_0 \simeq \gamma_1$, si existe una aplicación continua

$$H : [0,1] \times [0,1] \longrightarrow X$$

tal que, para todos $s, t \in [0,1]$,

$$H(s,0) = \gamma_0(s), \qquad H(s,1) = \gamma_1(s), \qquad H(0,t) = x, \qquad H(1,t) = y.$$

La clase de equivalencia de γ se denota $[\gamma]$.

Idea clave

Léase H como una película: para cada instante t fijo, la función $s \mapsto H(s,t)$ es un camino γ_t; en $t=0$ es el primero, en $t=1$ es el segundo, y la película es continua, sin cortes. Las dos últimas condiciones dicen que los dos extremos están clavados durante toda la película. Sin ellas cualquier camino sería homótopo a cualquier otro dentro de una misma componente conexa por caminos, sin más que deslizar un extremo hasta el otro, y el invariante no distinguiría nada. La noción general de homotopía entre aplicaciones es la de la definición 37.53; ésta es esa misma noción con la condición añadida de dejar quietos los extremos.

Proposición 37.7 (Es una relación de equivalencia)**.** *La relación $\simeq$ entre caminos de x a y es reflexiva, simétrica y transitiva.*

Demostración. Reflexiva: $H(s,t) = \gamma(s)$, que no depende de t, es continua y cumple todo. *Simétrica*: si H va de γ_0 a γ_1, entonces $H'(s,t) = H(s,1-t)$ va de γ_1 a γ_0; es continua por ser composición de H con una aplicación continua. *Transitiva*: si H lleva γ_0 a γ_1 y K lleva γ_1 a γ_2, defínase

$$L(s,t) = \begin{cases} H(s,2t), & t \le \frac{1}{2}, \\ K(s,2t-1), & t \ge \frac{1}{2}. \end{cases}$$

En $t = \frac{1}{2}$ ambas valen $\gamma_1(s)$, luego L está bien definida y es continua por el lema del pegado aplicado a $[0,1]\times[0,\frac{1}{2}]$ y $[0,1]\times[\frac{1}{2},1]$, que son cerrados. □

Lema 37.8 (Reparametrizar no cambia la clase)**.** *Sea γ un camino y $\varphi : [0,1] \to [0,1]$ continua con $\varphi(0) = 0$ y $\varphi(1) = 1$. Entonces $\gamma \circ \varphi \simeq \gamma$.*

Demostración. Basta interpolar linealmente los horarios dentro del intervalo, que es convexo:

$$H(s,t) = \gamma((1-t)\varphi(s) + t\,s).$$

El argumento de γ está en $[0,1]$ por convexidad, H es continua por composición, $H(\cdot,0) = \gamma\circ\varphi$, $H(\cdot,1) = \gamma$, y en $s=0$ y $s=1$ el argumento vale constantemente 0 y 1, luego los extremos quedan fijos. □

El grupo fundamental, construido

Teorema 37.9 (La concatenación está bien definida sobre clases)**.** *Si $\gamma_0 \simeq \gamma_1$ como caminos de x a y, y $\delta_0 \simeq \delta_1$ como caminos de y a z, entonces $\gamma_0 * \delta_0 \simeq \gamma_1 * \delta_1$. Por tanto $[\gamma]*[\delta] := [\gamma*\delta]$ es una operación bien definida sobre clases.*

Demostración. Sean F una homotopía de γ_0 a γ_1 y G una de δ_0 a δ_1. Defínase

$$H(s,t) = \begin{cases} F(2s,t), & s \le \frac{1}{2}, \\ G(2s-1,t), & s \ge \frac{1}{2}. \end{cases}$$

En $s = \frac{1}{2}$ la primera fórmula da $F(1,t) = y$ y la segunda $G(0,t) = y$ *para todo* t, porque ambas homotopías tienen los extremos clavados: aquí es donde se usa esa hipótesis, y sin ella la definición ni siquiera sería coherente. Por el lema del pegado H es continua. Y $H(\cdot,0) = \gamma_0 * \delta_0$, $H(\cdot,1) = \gamma_1 * \delta_1$, $H(0,t) = F(0,t) = x$, $H(1,t) = G(1,t) = z$. □

Teorema 37.10 ($\pi_1(X,x_0)$ es un grupo)**.** *Sea $x_0 \in X$ y sea $\pi_1(X,x_0)$ el conjunto de clases de homotopía de lazos con base en x_0. Con la operación $[\gamma][\delta] = [\gamma * \delta]$ es un grupo, con elemento neutro $[c_{x_0}]$ y con $[\gamma]^{-1} = [\overline{\gamma}]$.*

Demostración. La operación está bien definida por el teorema 37.9, y es interna porque la concatenación de dos lazos con base x_0 es un lazo con base x_0. Quedan tres cosas.

Asociatividad. Sea $\varphi : [0,1] \to [0,1]$ la función continua, lineal a trozos,

$$\varphi(s) = \begin{cases} 2s, & 0 \le s \le \frac{1}{4}, \\ s + \frac{1}{4}, & \frac{1}{4} \le s \le \frac{1}{2}, \\ \frac{s+1}{2}, & \frac{1}{2} \le s \le 1. \end{cases}$$

Es continua (en $s = \frac{1}{4}$ las dos primeras valen $\frac{1}{2}$; en $s = \frac{1}{2}$ las dos últimas valen $\frac{3}{4}$) y cumple $\varphi(0) = 0$, $\varphi(1) = 1$. Un cálculo directo, trozo a trozo, da la identidad de funciones

$$(\gamma * \delta) * \varepsilon = \big(\gamma * (\delta * \varepsilon)\big) \circ \varphi.$$

En efecto: para $s \le \frac{1}{4}$ el lado izquierdo vale $\gamma(4s)$ y el derecho $\gamma(2\varphi(s)) = \gamma(4s)$; para $\frac{1}{4} \le s \le \frac{1}{2}$ valen $\delta(4s-1)$ y $\delta(4\varphi(s)-2) = \delta(4s-1)$; y para $s \ge \frac{1}{2}$ valen $\varepsilon(2s-1)$ y $\varepsilon(4\varphi(s)-3) = \varepsilon(2s-1)$. Por el lema 37.8, las dos concatenaciones son homótopas, luego $([\gamma][\delta])[\varepsilon] = [\gamma]([\delta][\varepsilon])$.

Elemento neutro. Sea $\psi(s) = 0$ para $s \le \frac{1}{2}$ y $\psi(s) = 2s - 1$ para $s \ge \frac{1}{2}$. Es continua, $\psi(0) = 0$ y $\psi(1) = 1$, y por definición de concatenación $c_{x_0} * \gamma = \gamma \circ \psi$. El lema 37.8 da $c_{x_0} * \gamma \simeq \gamma$. Análogamente, con $\psi'(s) = 2s$ para $s \le \frac{1}{2}$ y $\psi'(s) = 1$ después, sale $\gamma * c_{x_0} \simeq \gamma$.

Elemento inverso. La idea es recorrer γ sólo hasta el instante $1-t$ y volver inmediatamente; en $t = 1$ no se recorre nada. Defínase

$$H(s,t) = \begin{cases} \gamma(2s(1-t)), & s \le \frac{1}{2}, \\ \gamma(2(1-s)(1-t)), & s \ge \frac{1}{2}. \end{cases}$$

En $s = \frac{1}{2}$ ambas valen $\gamma(1-t)$, luego H es continua por el lema del pegado. Para $t = 0$ se obtiene $\gamma(2s)$ en la primera mitad y $\gamma(2-2s) = \overline{\gamma}(2s-1)$ en la segunda, es decir, $\gamma * \overline{\gamma}$. Para $t = 1$ se obtiene $\gamma(0) = x_0$ constantemente. Y $H(0,t) = H(1,t) = \gamma(0) = x_0$. Luego $\gamma * \overline{\gamma} \simeq c_{x_0}$. Cambiando γ por $\overline{\gamma}$ sale la otra igualdad. □

Cuidado

Nótese qué es lo que *no* se ha demostrado: que la concatenación de caminos sea asociativa. No lo es, y no lo será nunca. Lo que se ha demostrado es que la obstrucción es una reparametrización, y que las reparametrizaciones son invisibles para la homotopía. Ese es el patrón de toda la topología algebraica: el objeto rígido no cumple los axiomas, y el cociente por la relación adecuada sí.

Definición 37.11 (Espacio simplemente conexo)**.** Un espacio X es *simplemente conexo* si es conexo por caminos y $\pi_1(X, x_0)$ es el grupo trivial para algún —equivalentemente, por la proposición siguiente, para todo— punto base x_0. En palabras: todo lazo se puede contraer con continuidad a un punto sin salirse del espacio.

Proposición 37.12 (Cambio de punto base)**.** *Sea h un camino de x_0 a x_1 en X. La aplicación*

$$\beta_h : \pi_1(X, x_1) \longrightarrow \pi_1(X, x_0), \qquad \beta_h([\gamma]) = [h * \gamma * \overline{h}\,],$$

es un isomorfismo de grupos. En consecuencia, si X es conexo por caminos, todos los grupos $\pi_1(X, x)$ son isomorfos y se escribe $\pi_1(X)$ sin punto base.

Demostración. Está bien definida por el teorema 37.9, y la expresión no necesita paréntesis por la asociatividad ya demostrada en clases. Es homomorfismo porque

$$[h * \gamma * \overline{h}\,][h * \delta * \overline{h}\,] = [h * \gamma * (\overline{h} * h) * \delta * \overline{h}\,] = [h * \gamma * \delta * \overline{h}\,],$$

usando que $[\overline{h} * h] = [c_{x_1}]$ por el mismo argumento del teorema 37.10. Y $\beta_{\overline{h}}$ es su inversa, porque $\beta_{\overline{h}}\beta_h([\gamma]) = [\overline{h} * h * \gamma * \overline{h} * h] = [\gamma]$. □

Cuidado

El isomorfismo depende del camino h elegido: dos caminos no homótopos dan en general isomorfismos distintos, que difieren en una conjugación. Por eso "$\pi_1(X)$" está definido salvo isomorfismo pero no canónicamente, y por eso en los enunciados que prescinden del punto base —como el que identifica $\pi_1(X)$ con las clases de aplicaciones de S^1 en X— aparece siempre la coletilla "salvo conjugación". Y si X no es conexo por caminos, la proposición no dice nada: recuérdese el contraejemplo de la circunferencia junto a un punto suelto.

El cálculo fundamental: $\pi_1(S^1) = \mathbb{Z}$

Este es el único cálculo del capítulo que no se deduce de otro, y del que cuelgan Brouwer, el teorema fundamental del álgebra y la bola peluda. Merece hacerse.

Teorema 37.13 (El grupo fundamental de la circunferencia)**.** *Sea $S^1 = \{z \in \mathbb{C} : |z| = 1\}$ con punto base 1. Entonces $\pi_1(S^1, 1) \cong \mathbb{Z}$, y el isomorfismo manda cada clase de lazo a su* número de vueltas.

Demostración. Sea $p : \mathbb{R} \to S^1$, $p(t) = \mathrm{e}^{2\pi i t}$. Es un recubrimiento: para cada $z \in S^1$, el arco abierto U de z de longitud menor que media circunferencia cumple que $p^{-1}(U)$ es la unión disjunta de los intervalos $I_n = I_0 + n$, $n \in \mathbb{Z}$, y p restringida a cada uno es un homeomorfismo sobre U. Se usará el teorema de levantamiento único de caminos y homotopías enunciado más adelante en el capítulo, que es lo único que hace falta.

Definición de la aplicación grado. Dado un lazo γ con base en 1, existe un único levantamiento $\widetilde{\gamma} : [0,1] \to \mathbb{R}$ con $p \circ \widetilde{\gamma} = \gamma$ y $\widetilde{\gamma}(0) = 0$. Como $\gamma(1) = 1$, se tiene $p(\widetilde{\gamma}(1)) = 1$, luego $\widetilde{\gamma}(1) \in \mathbb{Z}$. Defínase $\deg(\gamma) = \widetilde{\gamma}(1)$.

Está bien definida sobre clases. Si H es una homotopía de caminos de γ_0 a γ_1, se levanta a $\widetilde{H} : [0,1]^2 \to \mathbb{R}$ con $\widetilde{H}(0,0) = 0$. El borde derecho $t \mapsto \widetilde{H}(1,t)$ es continuo y toma valores en $p^{-1}(1) = \mathbb{Z}$, que es discreto; una función continua de un conexo en un discreto es constante, luego $\widetilde{\gamma}_0(1) = \widetilde{\gamma}_1(1)$ y $\deg(\gamma_0) = \deg(\gamma_1)$.

Es homomorfismo. Sean γ, δ lazos con grados m y n. El levantamiento de $\gamma * \delta$ que empieza en 0 es $\widetilde{\gamma}$ seguido de $\widetilde{\delta} + m$, porque $p(t+m) = p(t)$ para m entero y porque ese pegado es continuo y levanta la concatenación. Su valor final es $n+m$, luego $\deg(\gamma * \delta) = \deg(\gamma) + \deg(\delta)$.

Es sobreyectiva. El lazo $\omega_n(s) = \mathrm{e}^{2\pi i n s}$ se levanta a $\widetilde{\omega_n}(s) = ns$, cuyo valor en 1 es n.

Es inyectiva. Si $\deg(\gamma) = 0$, entonces $\widetilde{\gamma}$ es un lazo en $\mathbb{R}$ con base 0. Como $\mathbb{R}$ es convexo, $K(s,t) = (1-t)\widetilde{\gamma}(s)$ es una homotopía de caminos de $\widetilde{\gamma}$ al camino constante, con extremos fijos en 0. Componiendo con p se obtiene una homotopía de caminos en S^1 de γ a c_1, luego $[\gamma]$ es el neutro.

Un homomorfismo biyectivo es un isomorfismo, y $\deg : \pi_1(S^1, 1) \to \mathbb{Z}$ lo es. □

Ejemplo 37.14 (El grado, calculado en cuatro lazos)**. (a)** $\gamma(s) = \mathrm{e}^{6\pi i s}$ se levanta a $3s$: grado 3.

(b) $\gamma(s) = \mathrm{e}^{-2\pi i s}$ se levanta a $-s$: grado -1. Es el inverso del generador.

(c) Ir y volver. $\gamma(s) = \mathrm{e}^{2\pi i \sin(\pi s)}$ es un lazo, porque $\sin 0 = \sin \pi = 0$; se levanta a $\sin(\pi s)$, que empieza y acaba en 0: grado 0. Recorre media vuelta y deshace lo andado, y en efecto es contráctil.

(d) Tres adelante y una atrás. El lazo $\omega_3 * \overline{\omega_1}$ tiene grado $3 + (-1) = 2$ por el homomorfismo, luego es homótopo a ω_2 y a ningún otro ω_n. Ésta es la afirmación precisa que en la sección del grupo fundamental se enuncia sin justificar.

Observación 37.15 (La homología singular, en una definición)*.* El complejo de cadenas de la sección correspondiente se definió sobre triangulaciones. La versión que vale para cualquier espacio, sin elegir nada, es la *singular*: sea Δ^k el k-símplice estándar, un *k-símplice singular* de X es una aplicación continua $\sigma : \Delta^k \to X$ —sin ninguna condición más: puede ser constante, o no inyectiva—, y $C_k^{\mathrm{sing}}(X)$ es el grupo abeliano libre generado por todos ellos. El operador borde es la misma suma alternada de caras, y $H_k^{\mathrm{sing}}(X) = \ker \partial_k / \operatorname{im} \partial_{k+1}$. Los grupos de cadenas son enormes, pero la definición no depende de ninguna elección y es funtorial de manera evidente: eso es todo lo que se le pide.

Regla: decidir si dos lazos son homótopos, y con qué herramienta

1. **Compruébese primero que son lazos con el mismo punto base.** Si no lo son, la pregunta está mal planteada: hay que transportar con un camino h mediante β_h de la proposición 37.12, y entonces la respuesta depende de h salvo conjugación.
2. **Búsquese una homotopía explícita** antes que nada. Si el espacio es convexo, o estrellado, la interpolación lineal $H(s,t) = (1-t)\gamma_0(s) + t\gamma_1(s)$ siempre funciona y resuelve el problema en una línea.
3. **Si se sospecha que *no* son homótopos**, no se busque la homotopía: hay infinitas candidatas y no se agotan. Búsquese un invariante que los separe.
4. **Aplíquese un homomorfismo inducido.** Si $f : X \to Y$ es continua con $f(x_0) = y_0$, entonces $f_*[\gamma] = [f \circ \gamma]$ está bien definido y es homomorfismo. Si $f_*[\gamma_0] \neq f_*[\gamma_1]$ en un grupo donde sí se sabe calcular, ya está.
5. **El caso maestro es $Y = S^1$**, donde el invariante es el grado del teorema 37.13 y se calcula levantando: sígase el argumento continuo del lazo y mírese cuánto ha crecido al volver.
6. **Como último recurso, abelianícese.** Si los dos lazos tienen clases distintas en $H_1(X)$, tampoco son homótopos; el recíproco es falso, porque H_1 es el abelianizado de π_1.
7. **Compruébese el resultado en un caso conocido.** En S^1 el grado debe salir entero; en un espacio simplemente conexo, según la definición 37.11, la respuesta es siempre que sí, y si sale que no, hay un error.

Ejemplo 37.16 (Dos aplicaciones inmediatas del teorema)**. (a) S^1 no es contráctil.** Si lo fuera sería simplemente conexo, y su grupo fundamental sería trivial en lugar de $\mathbb{Z}$. En particular S^1 no es homeomorfo ni homotópicamente equivalente a un punto, ni a un intervalo, ni a $\mathbb{R}$.

(b) No hay retracción del disco sobre su borde. Si $r : D^2 \to S^1$ cumpliera $r|_{S^1} = \mathrm{id}$, aplicando π_1 a $S^1 \hookrightarrow D^2 \xrightarrow{r} S^1$ se tendría que $\mathrm{id}_{\mathbb{Z}}$ factoriza por $\pi_1(D^2) = 1$, porque el disco es convexo y por tanto simplemente conexo. Un homomorfismo que pasa por el grupo trivial es nulo, y el nulo no es la identidad de $\mathbb{Z}$. De aquí sale Brouwer, y de Brouwer todo lo demás.

37.3. Homología: contar agujeros de cada dimensión

Definición 37.17. Los *grupos de homología* $H_k(X)$ miden, hablando informalmente, los agujeros de dimensión k del espacio: H_0 cuenta las componentes conexas, H_1 los agujeros que atrapan un lazo, H_2 las cavidades que atrapan una superficie. Los *números de Betti* b_k son los rangos de esos grupos, y se cumple

$$\chi = \sum_k (-1)^k b_k.$$

Idea clave

Qué dice esto en cristiano. El grupo fundamental solo detecta agujeros de dimensión uno, los que atrapan lazos, y además puede ser no abeliano y difícil de calcular. La homología generaliza la idea a todas las dimensiones y produce grupos abelianos, que se calculan con álgebra lineal sobre la triangulación: se escribe qué borde tiene cada trozo, y se mide cuántos ciclos hay que no son bordes de nada. Un agujero es exactamente eso: algo que se cierra sobre sí mismo sin ser el borde de una región rellena.

Ejemplo 37.18 (Homología de los espacios básicos)**. (a) La esfera.** $H_0 = \mathbb{Z}$ (una componente), $H_1 = 0$ (todo lazo se contrae) y $H_2 = \mathbb{Z}$ (encierra una cavidad). Los números de Betti son $1, 0, 1$, y su suma alternada es $1 - 0 + 1 = 2 = \chi$, como debe.

(b) El toro. $H_0 = \mathbb{Z}$, $H_1 = \mathbb{Z} \times \mathbb{Z}$ (los dos lazos independientes, el que rodea el agujero y el que atraviesa el tubo) y $H_2 = \mathbb{Z}$. Los Betti son $1, 2, 1$ y $1 - 2 + 1 = 0 = \chi$. Nótese que aquí H_1 coincide con π_1, porque el grupo fundamental del toro ya era abeliano.

(c) Cuando difieren de π_1. Para la figura de ocho, π_1 es el grupo libre de dos generadores, no abeliano, mientras que $H_1 = \mathbb{Z} \times \mathbb{Z}$. La homología es la abelianización del grupo fundamental, así que pierde información a cambio de ser calculable.

(d) Torsión. En el plano proyectivo, $H_1 = \mathbb{Z}_2$: hay un lazo que no se contrae pero cuyo doble sí. Ese fenómeno, imposible de detectar contando dimensiones, es lo que hacen visible los grupos de homología frente a los meros números de Betti, y es la firma algebraica de la no orientabilidad.

Cuidado

[Homología y homotopía no son lo mismo] H_1 es la abelianización de π_1, luego dos espacios pueden tener la misma homología y grupos fundamentales distintos. Existen incluso espacios con la homología de un punto en todas las dimensiones y grupo fundamental no trivial, llamados esferas de homología; el más famoso es la esfera de Poincaré, que él construyó precisamente para refutar su propia conjetura inicial de que la homología bastaba para caracterizar la esfera. Segundo aviso: la homología es un funtor, es decir, una aplicación continua entre espacios induce homomorfismos entre sus grupos, y esa propiedad es la que hace demostrables teoremas como el del punto fijo de Brouwer.

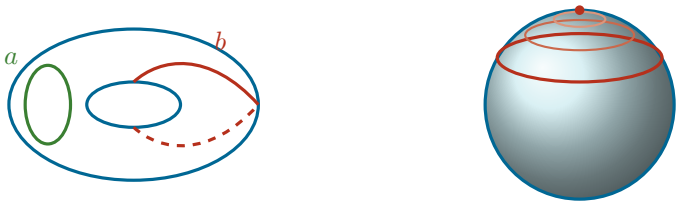
Pausa para pensar

La topología algebraica nació de un fracaso célebre. Poincaré conjeturó en 1900 que una variedad tridimensional con la homología de la esfera tenía que ser la esfera, y cuatro años después construyó él mismo un contraejemplo, la hoy llamada esfera de Poincaré, que tiene la misma homología y un grupo fundamental de orden 120. Al corregirse formuló entonces la conjetura correcta, con el grupo fundamental en vez de la homología, y esa versión resistió un siglo hasta que Perelmán la demostró en 2003 con técnicas de geometría diferencial, no de topología algebraica. Rechazó la medalla Fields y el millón de dólares del Instituto Clay.

37.4. El grupo fundamental: lazos que no se pueden deshacer

La característica de Euler es un número; la homología, una sucesión de grupos abelianos. Entre ambos vive el invariante más intuitivo de la topología algebraica: el grupo fundamental, que registra de cuántas maneras esencialmente distintas se puede dar una vuelta dentro de un espacio. Es el primer invariante no abeliano de la teoría, y por eso distingue cosas que la homología no ve.

Definición 37.19. Sea X un espacio topológico y $x_0 \in X$ un punto base. Un *lazo* es una aplicación continua $\gamma\colon [0,1] \to X$ con $\gamma(0) = \gamma(1) = x_0$. Dos lazos γ_0, γ_1 son *homótopos* si existe una aplicación continua $H\colon [0,1]\times[0,1] \to X$ con $H(s,0) = \gamma_0(s)$, $H(s,1) = \gamma_1(s)$ y $H(0,t) = H(1,t) = x_0$ para todo t. El conjunto de clases de homotopía de lazos, con el producto "recorrer primero uno y luego el otro", es un grupo llamado *grupo fundamental* y se denota $\pi_1(X, x_0)$.



toro: $\pi_1 = \mathbb{Z}\times\mathbb{Z}$ esfera: $\pi_1 = \{1\}$

Figura 37.1: Dos superficies que el grupo fundamental distingue sin ambigüedad. En el toro, el lazo a que rodea el agujero y el lazo b que atraviesa el tubo no se pueden deformar uno en otro ni contraer a un punto. En la esfera, cualquier lazo se puede ir encogiendo hasta desaparecer.

Idea clave

Qué dice esto en cristiano. Póngase una goma elástica dentro del espacio, clavada en un punto, y trátese de encogerla hasta el punto sin salirse del espacio ni romperla. Si siempre se puede, el grupo fundamental es trivial y el espacio se llama simplemente conexo. Si hay agujeros, algunas gomas se quedan enganchadas y no hay modo de encogerlas: esas son las clases no triviales. El producto es "dar primero una vuelta y luego la otra", el inverso es "dar la misma vuelta al revés", y el neutro es la goma que se encoge. Lo notable es que la deformación es continua y no rígida: da igual la forma exacta del lazo, sólo importa qué agujeros rodea y cuántas veces.

Regla: Calcular un grupo fundamental en siete pasos

1. **Comprobar que el espacio es conexo por caminos**; si lo es, el punto base no importa y se escribe $\pi_1(X)$ sin más.
2. **Buscar una retracción por deformación** a algo más simple: π_1 es invariante

por equivalencia homotópica, no sólo por homeomorfismo.

3. **Reconocer los casos base**: $\pi_1(\text{punto}) = \pi_1(\mathbb{R}^n) = \pi_1(S^n) = 1$ para $n \geq 2$, y $\pi_1(S^1) = \mathbb{Z}$.
4. **Usar productos**: $\pi_1(X \times Y) = \pi_1(X) \times \pi_1(Y)$.
5. **Usar Seifert–Van Kampen** si $X = U \cup V$ con U, V abiertos conexos por caminos y $U \cap V$ conexo por caminos: $\pi_1(X)$ es el producto libre amalgamado de $\pi_1(U)$ y $\pi_1(V)$ sobre $\pi_1(U \cap V)$.
6. **Si el espacio es un CW-complejo de una celda** 2, leer la presentación directamente: generadores son las aristas del 1-esqueleto y relaciones las palabras de borde de las caras.
7. **Contrastar con** H_1: siempre se cumple $H_1(X) \cong \pi_1(X)^{\mathrm{ab}}$, el abelianizado. Si el resultado no cuadra al abelianizar, hay un error.

Ejemplo 37.20 (Cuatro grupos fundamentales)**. (a) La circunferencia.** $\pi_1(S^1) = \mathbb{Z}$, y el isomorfismo es el *grado*: el número de vueltas netas que da el lazo, contando las horarias como negativas. Un lazo que da tres vueltas adelante y una atrás es homótopo al que da dos vueltas, y no es homótopo a ninguno de otra clase. La demostración usa el recubrimiento $\mathbb{R} \to S^1$, $t \mapsto \mathrm{e}^{2\pi i t}$, y el levantamiento único de caminos.

(b) El plano agujereado. $\mathbb{R}^2 \setminus \{0\}$ se retrae por deformación a S^1 mediante $x \mapsto x/\|x\|$, luego su grupo fundamental es también $\mathbb{Z}$. En cambio $\mathbb{R}^3 \setminus \{0\}$ se retrae a S^2 y tiene grupo trivial: en dimensión tres una goma puede rodear el origen y luego deslizarse por encima de él. La dimensión decide.

(c) El toro y la botella de Klein. $T^2 = S^1 \times S^1$ da $\pi_1(T^2) = \mathbb{Z} \times \mathbb{Z}$, generado por el meridiano a y el paralelo b con la relación $aba^{-1}b^{-1} = 1$. La botella de Klein tiene la presentación $\langle a, b \mid abab^{-1} \rangle$, que es un grupo no abeliano infinito. Sus abelianizados son $\mathbb{Z}^2$ y $\mathbb{Z} \oplus \mathbb{Z}/2\mathbb{Z}$ respectivamente, y ahí se ve la diferencia también en homología.

(d) El ocho. La figura de ocho, o cuña de dos circunferencias, tiene $\pi_1 = F_2$, el grupo libre de dos generadores: no hay ninguna relación entre "rodear el lazo izquierdo" y "rodear el derecho", y las palabras ab y ba describen lazos genuinamente distintos. Su abelianizado es $\mathbb{Z}^2$, el mismo que el del toro, de modo que la homología H_1 *no* distingue el toro del ocho mientras que π_1 sí. Éste es el ejemplo canónico de por qué vale la pena cargar con un invariante no abeliano.

Teorema 37.21 (Punto fijo de Brouwer en dimensión dos)**.** *Toda aplicación continua $f\colon D^2 \to D^2$ del disco cerrado en sí mismo tiene un punto fijo.*

Demostración. Si no lo tuviera, para cada x la semirrecta que parte de $f(x)$ y pasa por x corta a S^1 en un punto $r(x)$, y r resulta continua con $r(x) = x$ para $x \in S^1$: sería una retracción $D^2 \to S^1$. Aplicando π_1, la composición $S^1 \hookrightarrow D^2 \xrightarrow{r} S^1$ es la identidad, luego induce la identidad en $\mathbb{Z}$; pero factoriza por $\pi_1(D^2) = 1$, de modo que la identidad de $\mathbb{Z}$ sería la aplicación nula. Contradicción. □

Ejemplo 37.22 (Puntos fijos que sí y que no)**. (a) Un caso explícito.** La aplicación $f(x,y) = (\frac{y}{2}, \frac{x}{2})$ manda el disco unidad en sí mismo y su único punto fijo es el origen. La aplicación $f(x,y) = (\frac{x+1}{2}, \frac{y}{2})$ tiene punto fijo en $(1,0)$, sobre el borde.

(b) Sin compacidad, falla. En el disco *abierto*, la aplicación que empuja cada punto un poco hacia el borde no tiene punto fijo, porque el candidato estaría en la frontera, que se ha excluido. En $\mathbb{R}^2$ entero, la traslación $f(\mathbf{x}) = \mathbf{x} + (1,0)$ no tiene ninguno.

(c) Sin convexidad, falla. En la corona $\{1 \leq \|x\| \leq 2\}$, el giro de noventa grados es continuo, manda la corona en sí misma y no tiene ningún punto fijo. Lo que hace falta no es exactamente convexidad, sino que el espacio sea contráctil.

(d) En dimensiones altas. El enunciado vale para D^n con cualquier n, y la demostración es la misma sustituyendo π_1 por H_{n-1}. De él se deducen el teorema del punto fijo de Kakutani, la existencia de equilibrios de Nash y el teorema de Perron–Frobenius sobre matrices positivas.

Idea clave

Qué dice esto en cristiano. Agítese una taza de café y déjese reposar: alguna partícula del líquido ha vuelto exactamente a donde estaba. Póngase un mapa de España encima de otro mapa de España más pequeño: algún punto del mapa pequeño está justo encima del mismo punto del grande. La demostración es un ejemplo perfecto del método de la topología algebraica: se supone lo contrario, se fabrica con esa suposición una aplicación continua imposible, y la imposibilidad se detecta traduciendo el problema a grupos, donde "1 no es igual a $\mathbb{Z}$" es una obviedad. Ningún argumento analítico directo es tan corto.

Cuidado

[Contraejemplo: sin conexión por caminos el punto base importa] Si X es la unión disjunta de una circunferencia y un punto, entonces $\pi_1(X, x_0) = \mathbb{Z}$ si x_0 está en la circunferencia y $\pi_1(X, x_0) = 1$ si x_0 es el punto suelto. Escribir "$\pi_1(X)$" sin más carece de sentido. Por eso el paso 1 de la regla no es un tecnicismo: el grupo fundamental sólo ve la componente conexa por caminos del punto base, y no se entera de que existe el resto del espacio.

Cuidado

[Contraejemplo: π_1 trivial no implica contráctil] La esfera S^2 tiene grupo fundamental trivial y desde luego no es contráctil: su homología $H_2(S^2) = \mathbb{Z}$ lo impide. Más aún, $S^2 \times S^2$ y $\mathbb{C}P^2$ son ambos simplemente conexos y no son homotópicamente equivalentes. Y en sentido contrario, el *pendiente hawaiano* (unión de circunferencias de radios $1/n$ todas tangentes en un punto) tiene un grupo fundamental monstruoso, no numerable y no libre, pese a parecerse mucho a una cuña numerable de circunferencias, cuyo grupo sí es libre. La topología de punto conjunto se venga cuando se abusa de la intuición.

Pausa para pensar

Poincaré introdujo el grupo fundamental en 1895 y en 1904 formuló la pregunta que llevó su nombre durante un siglo: si una 3-variedad cerrada tiene grupo fundamental trivial, ¿es homeomorfa a S^3? La versión análoga en dimensión $n \geq 5$ la resolvió Smale en 1961, la de dimensión 4 Freedman en 1982, y la original, la de dimensión 3, la demostró Perelmán entre 2002 y 2003 con métodos que no son de topología algebraica sino de ecuaciones en derivadas parciales: el flujo de Ricci de Hamilton, una ecuación de difusión para la métrica. Perelmán rechazó la medalla Fields y el millón de dólares del Instituto Clay. La lección técnica es que un invariante puede plantear un problema sin poder resolverlo.

37.5. Recubrimientos y el grado de una aplicación

Entre el grupo fundamental de un espacio y sus recubrimientos hay una correspondencia tan ajustada como la de Galois entre cuerpos y grupos. Y una vez que se tiene, contar vueltas se convierte en una herramienta que demuestra teoremas de álgebra y de análisis.

Definición 37.23. Una aplicación continua $p\colon \tilde{X} \to X$ es un *recubrimiento* si todo punto de X tiene un entorno U tal que $p^{-1}(U)$ es unión disjunta de abiertos, cada uno aplicado homeomórficamente sobre U. Si X es conexo, localmente conexo por caminos y semilocalmente simplemente conexo, existe una biyección entre los recubrimientos conexos de X salvo isomorfismo y los subgrupos de $\pi_1(X)$ salvo conjugación; el recubrimiento *universal* corresponde al subgrupo trivial.

Idea clave

Qué dice esto en cristiano. Un recubrimiento es un espacio desplegado: la hélice infinita sobre la circunferencia, donde cada vuelta del círculo corresponde a subir un piso. El diccionario dice que la lista de todos los despliegues posibles de un espacio está codificada en su grupo fundamental, exactamente igual que la lista de cuerpos intermedios de una extensión de Galois está codificada en el grupo. Cuanto menor es el subgrupo, mayor es el recubrimiento, y el universal corresponde al subgrupo trivial y es simplemente conexo. Que la misma estructura formal aparezca en álgebra y en topología no es coincidencia.

Definición 37.24. El *grado* de una aplicación continua $f\colon S^n \to S^n$ es el entero d tal que $f_*\colon H_n(S^n) \to H_n(S^n)$ es la multiplicación por d. Para $n = 1$ coincide con el número de vueltas.

Ejemplo 37.25 (Contar vueltas para demostrar teoremas)**. (a) El teorema fundamental del álgebra.** Sea $p(z) = z^n + a_{n-1}z^{n-1} + \cdots + a_0$ sin raíces. Para R grande, p restringida al círculo $|z| = R$ tiene grado n, porque domina el término principal; para R pequeño tiene grado 0, porque p es casi constante. Como el grado es invariante por homotopía y R varía

continuamente, $n = 0$. Luego todo polinomio no constante tiene raíz. Es la demostración más limpia que existe y no usa análisis complejo.

(b) Brouwer en toda dimensión. Si $f\colon D^n \to D^n$ continua no tuviera punto fijo, se construiría una retracción $D^n \to S^{n-1}$, y aplicando H_{n-1} se obtendría que la identidad de $\mathbb{Z}$ factoriza por el grupo trivial.

(c) La bola peluda. Un campo vectorial continuo no nulo sobre S^n daría una homotopía entre la identidad y la aplicación antipodal, cuyos grados son 1 y $(-1)^{n+1}$. Para n par esos números difieren, así que no existe tal campo: en la esfera bidimensional siempre hay un punto donde el viento se para.

(d) Borsuk–Ulam. Toda aplicación continua $S^2 \to \mathbb{R}^2$ identifica algún par de puntos antipodales. En la interpretación popular: en todo instante existen dos puntos opuestos de la Tierra con la misma temperatura y la misma presión.

Cuidado

[Contraejemplo: no todo espacio tiene recubrimiento universal] El *pendiente hawaiano* —la unión de las circunferencias de radio $1/n$ tangentes en un punto común— es conexo y localmente conexo por caminos, pero no es semilocalmente simplemente conexo: en cualquier entorno del punto de tangencia hay lazos que no se contraen dentro del espacio. Carece de recubrimiento universal, y su grupo fundamental no es libre ni numerable, a diferencia del de una unión finita de círculos.

37.6. Cohomología, producto y dualidad de Poincaré

La homología cuenta agujeros y no sabe multiplicarlos. La cohomología cuenta los mismos agujeros, tiene un producto y por eso distingue espacios que la homología confunde. La diferencia entre subíndice y superíndice esconde una estructura de anillo.

Definición 37.26. La *cohomología* $H^k(X; R)$ se define dualizando el complejo de cadenas. El *producto copa* $\smile\colon H^k \times H^l \to H^{k+l}$ la convierte en un anillo graduado. El *teorema de coeficientes universales* relaciona ambas teorías mediante la sucesión exacta corta

$$0 \to \operatorname{Ext}^1(H_{k-1}X, R) \to H^k(X; R) \to \operatorname{Hom}(H_k X, R) \to 0.$$

Idea clave

Qué dice esto en cristiano. Si sólo se miran los grupos, homología y cohomología casi coinciden: sobre un cuerpo son literalmente duales. La ganancia está en el producto. Dos ciclos de dimensiones k y l que se cortan transversalmente en una variedad producen un ciclo de dimensión $k + l - n$, y el producto copa es la versión algebraica de esa intersección geométrica. Es decir: la cohomología recuerda no sólo cuántos agujeros hay, sino cómo se enlazan entre sí.

Teorema 37.27 (Dualidad de Poincaré)**.** *Si M es una variedad compacta, orientable y de dimensión n, entonces $H^k(M;\mathbb{Z}) \cong H_{n-k}(M;\mathbb{Z})$ para todo k. En particular los números de Betti son simétricos: $b_k = b_{n-k}$.*

Ejemplo 37.28 (Donde la homología no basta y el producto sí)**. (a) El toro.** $H^0 = \mathbb{Z}$, $H^1 = \mathbb{Z}^2$, $H^2 = \mathbb{Z}$, con números de Betti $1, 2, 1$, que son simétricos como exige la dualidad, y suma alternada $1 - 2 + 1 = 0 = \chi$. Los dos generadores a, b de H^1 cumplen $a \smile b = -b \smile a$ y ese producto genera H^2.

(b) Los mismos grupos, anillos distintos. El toro $S^1 \times S^1$ y el espacio $S^1 \vee S^1 \vee S^2$ —dos círculos y una esfera pegados por un punto— tienen exactamente los mismos grupos de homología: $\mathbb{Z}, \mathbb{Z}^2, \mathbb{Z}$. Pero en el segundo el producto copa de los dos generadores de H^1 es *cero*, porque los círculos no se cortan. El anillo de cohomología los distingue y los grupos por separado no. Ése es todo el argumento a favor del producto.

(c) Superficies de género g. Los números de Betti son $1, 2g, 1$ y $\chi = 2 - 2g$, que reproduce la fórmula del capítulo.

(d) Dónde falla. La botella de Klein no es orientable, y la dualidad de Poincaré con coeficientes enteros no se le aplica: tiene $H_1 = \mathbb{Z} \times \mathbb{Z}/2\mathbb{Z}$ y $H_2 = 0$, de modo que $b_0 = 1$ y $b_2 = 0$ no son simétricos. Con coeficientes en $\mathbb{Z}/2\mathbb{Z}$ la dualidad se recupera.

Cuidado

[Contraejemplo: el término de torsión no se puede ignorar] Es tentador escribir $H^k(X;\mathbb{Z}) \cong \mathrm{Hom}(H_k X, \mathbb{Z})$ y olvidarse. Con el plano proyectivo real, $H_1 = \mathbb{Z}/2\mathbb{Z}$, cuyo Hom en $\mathbb{Z}$ es cero; y sin embargo $H^2(\mathbb{RP}^2;\mathbb{Z}) = \mathbb{Z}/2\mathbb{Z}$, que aparece por el término Ext desplazado un grado. La torsión de la homología reaparece en la cohomología un escalón más arriba. Quien trabaje sólo con coeficientes racionales no verá nunca este fenómeno.

37.7. Complejos de cadenas: la definición de la homología

Hemos dicho que $H_k(X)$ "cuenta agujeros de dimensión k" y hemos anunciado sus valores para la esfera y el toro. Toca decir qué es exactamente ese grupo, porque la definición es completamente explícita y se calcula a mano.

Definición 37.29 (Complejo simplicial y grupos de cadenas)**.** Un complejo simplicial es una colección de vértices, aristas, triángulos, tetraedros… pegados por sus caras. $C_k(X)$ es el grupo abeliano libre generado por los k-símplices *orientados*, con el convenio de que intercambiar dos vértices cambia el signo.

Definición 37.30 (Operador borde)**.**

$$\partial_k[v_0, \ldots, v_k] = \sum_{i=0}^{k} (-1)^i [v_0, \ldots, \widehat{v_i}, \ldots, v_k],$$

donde el sombrero indica omisión. Por ejemplo, $\partial[v_0,v_1]=[v_1]-[v_0]$ y $\partial[v_0,v_1,v_2]=[v_1,v_2]-[v_0,v_2]+[v_0,v_1]$.

Teorema 37.31. $\partial_{k-1}\circ\partial_k=0$*: el borde de un borde es cero.*

La comprobación es un ejercicio de contabilidad: al quitar dos vértices, cada cara de codimensión dos aparece exactamente dos veces, con signos opuestos. Y es *la* identidad de la que cuelga todo el edificio; en el lenguaje de las formas diferenciales es $\mathrm{d}\circ\mathrm{d}=0$, y es la misma afirmación.

Definición 37.32 (Homología)**.**

$$Z_k=\ker\partial_k\ \ \text{(ciclos)},\qquad B_k=\operatorname{im}\partial_{k+1}\ \ \text{(bordes)},\qquad H_k=Z_k/B_k.$$

La inclusión $B_k\subset Z_k$ es exactamente $\partial\partial=0$, y por eso el cociente tiene sentido.

Idea clave

Un ciclo es algo cerrado, sin borde: un lazo, una superficie cerrada. Un borde es algo cerrado *porque* rodea a un trozo de dimensión superior. El cociente mide los ciclos que están cerrados sin razón aparente: los que rodean un agujero. Toda la intuición de "contar agujeros" cabe en esa frase, y ahora es una definición y no una metáfora.

Ejemplo 37.33 ($H_*(S^1)$ calculado a mano)**.** Triangúlese la circunferencia como el borde de un triángulo: tres vértices v_0,v_1,v_2 y tres aristas $a_0=[v_1,v_2]$, $a_1=[v_0,v_2]$, $a_2=[v_0,v_1]$. No hay 2-símplices, así que $C_2=0$.

(a) La matriz del borde. En las bases dadas,

$$\partial_1=\begin{pmatrix}0&-1&-1\\-1&0&1\\1&1&0\end{pmatrix},$$

donde la columna j recoge ∂a_j escrito en la base de vértices: $\partial a_0=v_2-v_1$, $\partial a_1=v_2-v_0$ y $\partial a_2=v_1-v_0$. Su rango es 2, porque la combinación $a_0-a_1+a_2$ está en el núcleo y dos columnas cualesquiera son independientes.

(b) H_1. $Z_1=\ker\partial_1$ tiene dimensión $3-2=1$, generado por el ciclo $a_0-a_1+a_2$, que es precisamente recorrer el triángulo entero. Como $C_2=0$, se tiene $B_1=0$ y

$$H_1(S^1)=\mathbb{Z}.$$

(c) H_0. $Z_0=C_0=\mathbb{Z}^3$, y $B_0=\operatorname{im}\partial_1$ tiene rango 2, luego $H_0=\mathbb{Z}^3/\mathbb{Z}^2\cong\mathbb{Z}$: una componente conexa.

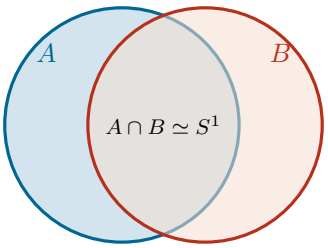
(d) Comprobación. La característica de Euler es $\chi=3-3=0$, y también $\chi=\operatorname{rango}H_0-\operatorname{rango}H_1=1-1=0$. La igualdad entre la cuenta ingenua de celdas y la cuenta de homología es un teorema, no una coincidencia, y es la versión moderna de $V-A+C=2$.

Observación 37.34. La homología *singular* sustituye los símplices de una triangulación por *todas* las aplicaciones continuas $\Delta^k \to X$. Es un objeto monstruoso —los grupos de cadenas tienen rango no numerable— pero se define para cualquier espacio, sin elegir triangulación, y es manifiestamente un invariante topológico. El teorema de comparación dice que para un complejo simplicial las dos homologías coinciden, y eso es lo que permite calcular con la finita y razonar con la general. La homología *celular*, basada en complejos CW, es el compromiso práctico: la esfera S^n se construye con una celda de dimensión 0 y otra de dimensión n, y sus grupos salen de leer el complejo $0 \to \mathbb{Z} \to 0 \to \cdots \to 0 \to \mathbb{Z} \to 0$.

37.8. Cómo se calcula de verdad: funtorialidad, Mayer–Vietoris y Künneth

Con la definición en la mano, calcular la homología de una esfera triangulando parece penoso. No hace falta: hay tres herramientas que la producen casi sin cuentas.

Definición 37.35 (Funtorialidad)**.** Una aplicación continua $f : X \to Y$ induce homomorfismos $f_* : H_k(X) \to H_k(Y)$, compatibles con la composición ($(g \circ f)_* = g_* f_*$) y con la identidad. Además, si $f \simeq g$ son homótopas, entonces $f_* = g_*$.



$$0 \longrightarrow H_2(S^2) \xrightarrow{\partial} H_1(S^1) \longrightarrow 0$$

$$0 \longrightarrow H_2(S^2) \xrightarrow{\cong} \mathbb{Z} \longrightarrow 0$$

$$H_2(S^2) \cong \mathbb{Z}, \qquad H_1(S^2) = 0$$

Figura 37.2: Dos discos que se solapan cubren la esfera, y su intersección tiene el tipo de homotopía de una circunferencia. Como los dos trozos son contráctiles, la sucesión exacta larga se reduce a un isomorfismo y da $H_2(S^2) = \mathbb{Z}$ sin triangular nada. Es el análogo homológico de Seifert–Van Kampen.

Esa última línea es la que convierte la homología en un invariante *homotópico*, y de ella salen las demostraciones de imposibilidad. Por ejemplo, la no existencia de retracción $r : D^2 \to S^1$: si existiera, con i la inclusión se tendría $r \circ i = \mathrm{id}_{S^1}$, luego $r_* i_* = \mathrm{id}$ en $H_1(S^1) = \mathbb{Z}$; pero i_* factoriza por $H_1(D^2) = 0$. Contradicción, y de ahí Brouwer en dos líneas más.

Teorema 37.36 (Sucesión exacta larga de Mayer–Vietoris)**.** *Si $X = A \cup B$ con A, B abiertos, hay una sucesión exacta*

$$\cdots \to H_k(A \cap B) \to H_k(A) \oplus H_k(B) \to H_k(X) \xrightarrow{\partial} H_{k-1}(A \cap B) \to \cdots$$

Es el análogo homológico exacto de Seifert–Van Kampen: descomponer el espacio en dos trozos cuyas homologías se conocen, y leer la del total.

Ejemplo 37.37 (La esfera, en cuatro líneas)**.** Cúbrase S^2 con dos discos abiertos ligeramente solapados, A el hemisferio norte engordado y B el sur. Entonces A y B son contráctiles y $A \cap B$ tiene el tipo de homotopía de S^1. La sucesión en grado 2 es

$$\underbrace{H_2(A) \oplus H_2(B)}_{0} \to H_2(S^2) \xrightarrow{\partial} \underbrace{H_1(S^1)}_{\mathbb{Z}} \to \underbrace{H_1(A) \oplus H_1(B)}_{0},$$

luego ∂ es un isomorfismo y $H_2(S^2) \cong \mathbb{Z}$. Bajando un grado, $H_1(S^2) = 0$. Y $H_0 = \mathbb{Z}$ por conexión. En total

$$H_*(S^2) = (\mathbb{Z}, 0, \mathbb{Z}).$$

El mismo argumento con $S^n = A \cup B$ da la recursión $H_k(S^n) \cong H_{k-1}(S^{n-1})$, y de ahí $H_k(S^n) = \mathbb{Z}$ para $k = 0, n$ y 0 en lo demás, por inducción.

Teorema 37.38 (Künneth, versión con cuerpo)**.** *Si F es un cuerpo,*

$$H_n(X \times Y; F) \cong \bigoplus_{i+j=n} H_i(X; F) \otimes H_j(Y; F).$$

Sobre $\mathbb{Z}$ hay un término correctivo de torsión, que se anula si alguna de las dos homologías es libre.

Ejemplo 37.39 (El toro sin dibujar nada)**.** $T^2 = S^1 \times S^1$, con $H_*(S^1) = (\mathbb{Z}, \mathbb{Z})$, ambos libres. Künneth da

$$H_0 = \mathbb{Z} \otimes \mathbb{Z} = \mathbb{Z}, \qquad H_1 = (\mathbb{Z} \otimes \mathbb{Z}) \oplus (\mathbb{Z} \otimes \mathbb{Z}) = \mathbb{Z}^2, \qquad H_2 = \mathbb{Z} \otimes \mathbb{Z} = \mathbb{Z}.$$

Los dos generadores de H_1 son los dos lazos del toro, el meridiano y el paralelo, y ahora aparecen como consecuencia formal de un producto tensorial, no como observación visual.

Observación 37.40 (Coeficientes: cuando cambiar de anillo cambia la respuesta)*.* La homología se puede calcular con coeficientes en cualquier grupo abeliano. Para el plano proyectivo:

$$H_*(\mathbb{RP}^2; \mathbb{Z}) = (\mathbb{Z},\ \mathbb{Z}/2,\ 0), \qquad H_*(\mathbb{RP}^2; \mathbb{Z}/2) = (\mathbb{Z}/2,\ \mathbb{Z}/2,\ \mathbb{Z}/2).$$

Con coeficientes enteros, $H_2 = 0$: el plano proyectivo no tiene clase fundamental, porque no es orientable. Con coeficientes en $\mathbb{Z}/2$ sí la tiene, y de hecho la dualidad de Poincaré se recupera. La razón es que la no orientabilidad se manifiesta como un signo, y en $\mathbb{Z}/2$ los signos desaparecen.

El teorema de coeficientes universales explica el trasvase: la torsión de H_{k-1} con coeficientes enteros reaparece en H_k al cambiar de anillo, vía el término Tor. La homología con coeficientes en un cuerpo es más simple pero pierde información; la entera la conserva toda a costa de complicarse.

Idea clave

[La torre que no se puede calcular] Junto a la homología existen los grupos de homotopía superiores $\pi_n(X)$, definidos como clases de aplicaciones $S^n \to X$ que respetan un punto

base. Son abelianos para $n \geq 2$ y el teorema de Hurewicz dice que si X es $(n-1)$-conexo entonces $\pi_n \cong H_n$: en el primer grado no trivial, homotopía y homología coinciden. A partir de ahí divergen brutalmente. La esfera S^2 tiene $H_3 = 0$ y sin embargo

$$\pi_3(S^2) = \mathbb{Z},$$

generado por la fibración de Hopf $S^3 \to S^2$. Y los grupos $\pi_n(S^k)$ con $n > k$ son en general desconocidos: no hay algoritmo, no hay patrón cerrado, y calcular unos pocos más ha ocupado carreras enteras. Esa es la razón práctica de que la topología algebraica se haga con homología: es infinitamente más débil que la homotopía, y precisamente por eso se puede calcular.

37.9. Recubrimientos, monodromía y la torre de los π_n

La correspondencia entre recubrimientos y subgrupos se enunció como un diccionario. Falta el mecanismo que la hace calculable, y es el levantamiento.

Teorema 37.41 (Levantamiento de caminos y homotopías)**.** *Sea $p : \tilde{X} \to X$ un recubrimiento. Todo camino γ en X con origen x_0 se levanta a un único camino en $\tilde{X}$ una vez elegido el punto inicial sobre x_0. Y toda homotopía de caminos se levanta a una homotopía.*

Ese teorema, usado sin enunciar al demostrar $\pi_1(S^1) = \mathbb{Z}$, es lo único que hace falta.

Definición 37.42 (Acción de monodromía)**.** Recorrer un lazo γ de X y levantarlo permuta la fibra $p^{-1}(x_0)$. Eso define una acción de $\pi_1(X, x_0)$ sobre la fibra, transitiva si $\tilde{X}$ es conexo, y el estabilizador de un punto es $p_*\pi_1(\tilde{X})$. Por órbita-estabilizador,

$$\text{número de hojas} = [\pi_1(X) : p_*\pi_1(\tilde{X})].$$

Ejemplo 37.43 (Todos los recubrimientos del círculo)**.** $\pi_1(S^1) = \mathbb{Z}$, y sus subgrupos son exactamente $n\mathbb{Z}$ para $n \geq 0$. La correspondencia da, uno por uno:

subgrupo	recubrimiento	hojas
$n\mathbb{Z}$ $(n \geq 1)$	$z \mapsto z^n,\ S^1 \to S^1$	n
0	$t \mapsto e^{2\pi i t},\ \mathbb{R} \to S^1$	∞

Y no hay más. El recubrimiento universal es $\mathbb{R}$, correspondiente al subgrupo trivial, y su grupo de transformaciones es $\mathbb{Z}$ actuando por traslaciones enteras: exactamente $\pi_1(S^1)$.

El caso proyectivo. $S^2 \to \mathbb{RP}^2$ identifica puntos antípodas: dos hojas, luego el subgrupo tiene índice 2 en $\pi_1(\mathbb{RP}^2)$; como S^2 es simplemente conexa, ese subgrupo es trivial y

$$\pi_1(\mathbb{RP}^2) = \mathbb{Z}/2.$$

Y la característica de Euler es multiplicativa en recubrimientos: $\chi(S^2) = 2 = 2 \cdot \chi(\mathbb{RP}^2)$, luego $\chi(\mathbb{RP}^2) = 1$, coherente con su descomposición celular en una celda de cada dimensión 0, 1 y 2.

La torre que no se puede calcular

Definición 37.44 (Grupos de homotopía superiores)**.** $\pi_n(X, x_0)$ es el conjunto de clases de homotopía de aplicaciones $(S^n, *) \to (X, x_0)$, con la operación de concatenar en una coordenada. Para $n \geq 2$ el grupo es *abeliano*, y la razón es visual: con dos direcciones libres las dos copias de la esfera se pueden intercambiar deslizándolas.

Teorema 37.45 (Hurewicz)**.** *Si X es $(n-1)$-conexo con $n \geq 2$, entonces $\pi_k(X) = 0$ para $k < n$ y*

$$\pi_n(X) \;\cong\; H_n(X).$$

Para $n = 1$ el enunciado se degrada a $H_1(X) \cong \pi_1(X)^{\mathrm{ab}}$.

En el primer grado no trivial, homotopía y homología coinciden. A partir de ahí divergen, y de forma espectacular.

Ejemplo 37.46 (La fibración de Hopf)**.** Para la esfera S^2 se tiene $H_3(S^2) = 0$: no hay agujeros tridimensionales. Y sin embargo

$$\pi_3(S^2) = \mathbb{Z}.$$

El generador es la **fibración de Hopf** $h : S^3 \to S^2$, cuya fibra sobre cada punto es una circunferencia y cuyas fibras están todas enlazadas dos a dos: dos circunferencias cualesquiera de la familia forman un enlace de número 1, y ese entero es exactamente el invariante que detecta la clase de homotopía.

La construcción es puro álgebra: $S^3 = \{(z_1, z_2) \in \mathbb{C}^2 : |z_1|^2 + |z_2|^2 = 1\}$ y $h(z_1, z_2) = z_1/z_2 \in \mathbb{C} \cup \{\infty\} = S^2$. Es decir, S^2 es la recta proyectiva compleja y la fibración es el paso al cociente por la fase común.

Que exista una aplicación $S^3 \to S^2$ no homótopa a una constante era, en 1931, una sorpresa mayúscula: contradecía la intuición de que "no se puede envolver algo pequeño con algo grande". Y abrió la caja: los grupos $\pi_n(S^k)$ con $n > k$ resultaron ser casi todos no nulos, sin patrón cerrado y sin algoritmo. Calcular unos pocos más ha ocupado carreras enteras; se conocen hasta $\pi_{n+k}(S^n)$ para k de dos cifras y poco más.

Idea clave

Esa es la razón práctica de que la topología algebraica se haga con *homología*. Es infinitamente más débil que la homotopía —no distingue S^2 de sus fibraciones, ni ve el enlace de Hopf— y precisamente por eso se puede calcular: hay Mayer–Vietoris, hay escisión, hay Künneth, hay una sucesión exacta larga para cada par. La homotopía no tiene nada de eso. En matemáticas, el invariante útil casi nunca es el más fino: es el más fino *entre los computables*.

37.10. Van Kampen y los complejos CW, enunciados

El capítulo pide en un ejercicio "enuncie el teorema de Seifert y Van Kampen", y el enunciado no aparece en ninguna parte; y el método de cálculo por celdas se usa sin decir qué

es una celda. Se remedian ambas cosas.

Definición 37.47 (Producto libre y producto libre amalgamado)**.** El *producto libre* $G * H$ de dos grupos es el grupo cuyos elementos son las palabras finitas alternadas $g_1 h_1 g_2 h_2 \cdots$ con letras en G y en H, con la multiplicación por yuxtaposición y sin más relaciones que las de G y las de H por separado. Si $G = \langle S_G \mid R_G \rangle$ y $H = \langle S_H \mid R_H \rangle$, entonces $G * H = \langle S_G \sqcup S_H \mid R_G \sqcup R_H \rangle$.

Dado además un grupo K con homomorfismos $\varphi : K \to G$ y $\psi : K \to H$, el *producto libre amalgamado* es el cociente

$$G *_K H = (G * H) / \langle\!\langle\, \varphi(k)\, \psi(k)^{-1} : k \in K \,\rangle\!\rangle,$$

donde se cocienta por el subgrupo normal generado por esas relaciones. En palabras: se pegan G y H sin relaciones *salvo* las que obliga a identificar la parte común K.

Teorema 37.48 (Seifert–Van Kampen)**.** *Sea $X = U \cup V$ con U, V abiertos, y supóngase que U, V y $U \cap V$ son conexos por caminos y que el punto base x_0 está en $U \cap V$. Entonces*

$$\pi_1(X, x_0) \;\cong\; \pi_1(U, x_0) \;*_{\pi_1(U\cap V,\, x_0)}\; \pi_1(V, x_0),$$

donde la amalgama se toma por los homomorfismos inducidos por las inclusiones $U \cap V \hookrightarrow U$ y $U \cap V \hookrightarrow V$.

En particular, si $U \cap V$ es simplemente conexo, $\pi_1(X) \cong \pi_1(U) * \pi_1(V)$; y si V es simplemente conexo, $\pi_1(X)$ es el cociente de $\pi_1(U)$ por el subgrupo normal generado por la imagen de $\pi_1(U \cap V)$.

Ejemplo 37.49 (Tres cálculos inmediatos)**.**

- *Figura de ocho.* Se cubre con dos entornos, cada uno retráctil a una circunferencia, y su intersección es contráctil: $\pi_1 = \mathbb{Z} * \mathbb{Z} = F_2$, el grupo libre de rango 2. Un grupo no abeliano, lo que explica que dos lazos no conmuten.
- *Esfera S^2.* Se cubre con dos discos abiertos ligeramente solapados; ambos son simplemente conexos y la intersección es un anillo, conexo. Sale $\pi_1(S^2) = 1 * 1 = 1$: la esfera es simplemente conexa.
- *Toro.* Presentado como cuadrado con lados identificados, U es el complementario de un punto —que se retrae a la figura de ocho, con $\pi_1 = \langle a, b \rangle$— y V un disco alrededor del punto; la intersección es un anillo cuyo generador va a parar al conmutador $aba^{-1}b^{-1}$. Resultado: $\pi_1(T^2) = \langle a, b \mid aba^{-1}b^{-1} \rangle = \mathbb{Z}^2$.

Definición 37.50 (Celda y complejo CW)**.** Una *n-celda* es una copia del disco abierto D^n. Un *complejo CW* se construye por etapas:

- el 0-*esqueleto* X^0 es un conjunto discreto de puntos (las 0-celdas);
- el *n-esqueleto* X^n se obtiene de X^{n-1} pegando n-celdas: para cada una se da una *aplicación de adjunción* continua $\varphi_\alpha : S^{n-1} \to X^{n-1}$, y se forma el espacio cociente $(X^{n-1} \sqcup \bigsqcup_\alpha D^n_\alpha)/\!\sim$ identificando cada punto $z \in \partial D^n_\alpha$ con $\varphi_\alpha(z)$;

- $X = \bigcup_n X^n$, con la topología en la que un conjunto es cerrado si su intersección con cada celda cerrada lo es.

El 1-*esqueleto* es, por tanto, un grafo: vértices y aristas.

Ejemplo 37.51 (Los espacios habituales, como complejos CW)**.** La esfera S^2: una 0-celda y una 2-celda, pegada por la aplicación constante (todo el borde del disco va al punto). El toro: una 0-celda, dos 1-celdas a, b y una 2-celda pegada a lo largo de la palabra $aba^{-1}b^{-1}$. El plano proyectivo $\mathbb{R}P^2$: una 0-celda, una 1-celda a y una 2-celda pegada según a^2. La botella de Klein: igual que el toro pero con $abab^{-1}$. Toda superficie compacta admite una descomposición así con una sola 2-celda, y la palabra de pegado es justamente su clasificación.

Observación 37.52 (Por qué esto hace calculable la topología)*.* Con la descomposición celular, el grupo fundamental de un complejo CW conexo se lee directamente: los generadores son las 1-celdas fuera de un árbol maximal del 1-esqueleto, y las relaciones son las palabras de pegado de las 2-celdas. Las celdas de dimensión ≥ 3 no afectan a π_1. Y la característica de Euler es simplemente $\chi = \sum_n (-1)^n \#\{n\text{-celdas}\}$. Esa es la razón por la que los complejos CW son el marco de trabajo de la topología algebraica: no son más generales que los espacios razonables, pero sí infinitamente más manejables.

37.11. Equivalencia homotópica, orientabilidad y cómo se calcula la homología contando celdas

Hasta aquí se ha dicho que la homología "no cambia al deformar el espacio" y que χ "es un invariante topológico", sin decir nunca con precisión qué deformaciones están permitidas. Toca decirlo, porque la relación que de verdad preservan todos estos invariantes no es el homeomorfismo sino algo más grosero, y esa holgura es justamente lo que los hace calculables. Después se define la orientabilidad, que es el segundo dato del teorema de clasificación de superficies y el único que no es un número, y se cierra con el procedimiento por celdas, que es como se calcula la homología en la práctica cuando no hay una descomposición cómoda en dos abiertos.

Definición 37.53 (Homotopía de aplicaciones y equivalencia homotópica de espacios)**.** Dos aplicaciones continuas $f, g : X \to Y$ son *homótopas*, y se escribe $f \simeq g$, si existe una aplicación continua $H : X \times [0,1] \to Y$ con $H(\cdot, 0) = f$ y $H(\cdot, 1) = g$. Es una relación entre *aplicaciones*, con los dos espacios fijos.

Dos *espacios* X e Y son *homotópicamente equivalentes*, y se escribe $X \simeq Y$, si existen aplicaciones continuas $f : X \to Y$ y $g : Y \to X$ tales que

$$g \circ f \simeq \mathrm{id}_X \qquad \text{y} \qquad f \circ g \simeq \mathrm{id}_Y .$$

Nótese que no se pide $g \circ f = \mathrm{id}_X$, que es lo que exigiría el homeomorfismo, sino solo que sea homótopa a la identidad. Un espacio homotópicamente equivalente a un punto se llama *contráctil*.

Un caso particular, y el único que se usa en la práctica: $A \subseteq X$ es un *retracto de deformación* de X si existe $H : X \times [0,1] \to X$ con $H(x, 0) = x$, $H(x, 1) \in A$ para todo x, y $H(a, t) = a$

para todo $a \in A$ y todo t. Es decir, si X se puede encoger continuamente sobre A sin mover nunca los puntos de A. Entonces $A \simeq X$, con f la inclusión y $g = H(\cdot, 1)$.

Cuidado

Las dos nociones se confunden todo el tiempo porque comparten el símbolo $\simeq$ y la palabra homotopía, y sin embargo tienen tipos distintos: $f \simeq g$ relaciona dos *flechas*, mientras que $X \simeq Y$ relaciona dos *objetos*. La confusión no es inocua. Que exista una aplicación $f : X \to Y$ homótopa a una constante no dice nada sobre si X e Y son equivalentes; y dos espacios homotópicamente equivalentes pueden tener dimensiones distintas, cosa que jamás ocurre entre espacios homeomorfos.

Ejemplo 37.54 (Cuatro equivalencias que conviene tener memorizadas)**. (a)** $\mathbb{R}^n \simeq \{0\}$**.** La homotopía $H(x,t) = (1-t)x$ contrae todo el espacio al origen sin salirse de él, y deja fijo el 0: el origen es un retracto de deformación de $\mathbb{R}^n$. El mismo argumento vale para cualquier convexo y, más en general, para cualquier estrellado. Por eso todos sus invariantes son los del punto: $H_0 = \mathbb{Z}$ y $H_k = 0$ para $k \geq 1$. Y $\mathbb{R}^2$ no es homeomorfo a $\mathbb{R}^3$ aunque ambos sean contráctiles; la equivalencia homotópica olvida la dimensión.

(b) La banda cilíndrica $\simeq S^1$**.** En $S^1 \times [-1,1]$, la homotopía $H\big((z,s),t\big) = (z,(1-t)s)$ desliza cada punto por su segmento vertical hasta la circunferencia central $s = 0$, dejándola fija. Luego el cilindro se retrae por deformación a S^1, y $H_*(\text{cilindro}) = H_*(S^1) = (\mathbb{Z},\mathbb{Z})$. No son homeomorfos —uno es una superficie con borde y el otro una curva— pero para la homología y para π_1 son indistinguibles.

(c) La cinta de Möbius $\simeq S^1$**.** La cinta es el cociente de $[0,1]\times[-1,1]$ que identifica $(0,s)$ con $(1,-s)$. El mismo deslizamiento vertical $s \mapsto (1-t)s$ es compatible con esa identificación, porque $-0 = 0$, y por tanto pasa al cociente: la circunferencia central es también aquí un retracto de deformación. En consecuencia la cinta de Möbius y el cilindro tienen *exactamente* los mismos grupos de homología y el mismo grupo fundamental $\mathbb{Z}$. Ningún invariante homotópico los distinguirá jamás, y sin embargo son superficies muy distintas: una es orientable y la otra no. Ese es el ejemplo que obliga a definir la orientabilidad aparte.

(d) $\mathbb{R}^2\setminus\{0\} \simeq S^1$ **y** $\mathbb{R}^3\setminus\{0\} \simeq S^2$**.** La homotopía radial $H(x,t) = (1-t)x + t\,x/\|x\|$ empuja cada punto hasta la esfera unidad sin cruzar nunca el origen, que ya no está. De aquí sale el cálculo entero de la homología del complementario de un punto, y con él la demostración de que $\mathbb{R}^2$ y $\mathbb{R}^3$ no son homeomorfos: quitar un punto deja en un caso algo con $H_1 = \mathbb{Z}$ y en el otro algo con $H_1 = 0$.

Idea clave

Por qué se trabaja con esta relación tan tosca y no con el homeomorfismo. Porque *todos* los invariantes construidos en este capítulo —π_1, H_k, H^k, χ— son invariantes de equivalencia homotópica, no solo de homeomorfismo. Eso los debilita, porque no verán la diferencia entre un disco y un punto; y a cambio los vuelve calculables, porque permiten sustituir un espacio complicado por otro mucho más pequeño con la misma respuesta. Todo el oficio del cálculo consiste en encontrar el retracto de deformación adecuado

antes de empezar a contar.

Definición 37.55 (Superficie orientable)**.** Sea S una superficie compacta y conexa, sin borde. Las dos condiciones siguientes son equivalentes, y cuando se cumplen se dice que S es *orientable*:

- *Versión diferencial.* Existe un campo de vectores normales unitarios continuo y global sobre S; es decir, se puede elegir en cada punto uno de los dos lados de la superficie de manera que la elección varíe con continuidad y no se contradiga al recorrer ningún camino cerrado.
- *Versión topológica.* S no contiene ninguna cinta de Möbius embebida; equivalentemente, todo camino cerrado devuelve un pequeño reloj dibujado sobre S girando en el mismo sentido en que partió.

Si falla, S es *no orientable*. Un camino que invierte el sentido de giro se llama *camino invertidor de orientación*, y transportar a lo largo de él el vector normal lo devuelve apuntando al lado contrario: por eso ambas versiones dicen lo mismo.

Ejemplo 37.56 (Quién es orientable y quién no)**. (a) El toro sí.** T^2 es la frontera de un sólido de $\mathbb{R}^3$, y basta tomar en cada punto la normal exterior: es continua, global y nunca se contradice. Todas las superficies compactas embebidas en $\mathbb{R}^3$ sin borde son orientables por esta misma razón, lo que explica que la intuición de "dos caras" funcione tan bien mientras uno no salga del espacio ordinario.

(b) La botella de Klein no. Su descomposición celular pega la 2-celda según la palabra $abab^{-1}$, y esa inversión de signo es exactamente un camino invertidor. En términos de la segunda caracterización: la botella de Klein contiene una cinta de Möbius, y de hecho es la suma conexa de dos planos proyectivos. Por eso no admite normal global, y por eso no se puede embeber en $\mathbb{R}^3$ sin autointersecarse.

(c) El plano proyectivo tampoco. $\mathbb{RP}^2$ se obtiene pegando un disco al borde de una cinta de Möbius, así que la contiene por construcción. Su palabra de pegado es a^2, con los dos signos iguales.

(d) Por qué hace falta este dato. El toro y la botella de Klein tienen ambos $\chi = 0$, y la cinta de Möbius y el cilindro son homotópicamente equivalentes. La orientabilidad es la información que separa esos pares y la que completa el teorema de clasificación: el par formado por χ y la orientabilidad determina la superficie salvo homeomorfismo. Algebraicamente se manifiesta como torsión —$H_1(\mathbb{RP}^2) = \mathbb{Z}/2$— y como la ausencia de clase fundamental, es decir, $H_2 = 0$ con coeficientes enteros.

Regla: calcular la homología de un complejo CW

Sea X un complejo CW finito, en el sentido de la definición 37.50. Los grupos $H_n(X)$ se obtienen así, sin triangular y sin cubrir por abiertos.

1. **Elíjase una estructura CW con el menor número de celdas posible.** Este

paso es todo el arte: cuantas menos celdas, menores son las matrices. La esfera S^n admite una descomposición con dos celdas; el toro, con cuatro. Toda superficie compacta admite una con una sola 2-celda, y su palabra de pegado es su clasificación.

2. **Escríbase el grupo de cadenas celulares** $C_n(X)$ como el grupo abeliano libre generado por las n-celdas, orientadas. Su rango es simplemente el número de n-celdas, y $C_n = 0$ en cuanto no haya celdas de esa dimensión.

3. **Escríbase la matriz del borde** $\partial_n : C_n \to C_{n-1}$. Su entrada (β, α) es el *grado* $d_{\alpha\beta}$ de la composición

$$S^{n-1} \xrightarrow{\varphi_\alpha} X^{n-1} \longrightarrow X^{n-1}/(X^{n-1} \setminus e_\beta^{n-1}) \cong S^{n-1},$$

donde φ_α es la aplicación de adjunción de la celda e_α^n y la segunda flecha colapsa todo salvo la celda e_β^{n-1}. En la práctica: recórrase la palabra de pegado y súmese $+1$ por cada aparición de la letra y -1 por cada aparición de su inversa. Así, $aba^{-1}b^{-1}$ da grado 0 en a y en b; $abab^{-1}$ da grado 0 en a y 2 en b; y a^2 da grado 2 en a.

4. **Calcúlese** $H_n(X) = \ker \partial_n / \operatorname{im} \partial_{n+1}$, reduciendo las dos matrices a forma normal de Smith si hace falta. Los factores invariantes distintos de 1 son exactamente la torsión, y el número de ceros es el número de Betti b_n.

5. **Compruébese con la característica de Euler.** Debe cumplirse

$$\chi(X) = \sum_n (-1)^n \operatorname{rango} C_n = \sum_n (-1)^n \operatorname{rango} H_n,$$

es decir, la cuenta ingenua de celdas y la cuenta de agujeros han de coincidir. La primera suma se lee del paso 2 sin hacer nada; la segunda, del paso 4. Si discrepan, hay un error de signo en alguna matriz del borde, porque la igualdad es un teorema y no una casualidad: al pasar al cociente por un subgrupo se resta su rango, y esos rangos se cancelan telescópicamente.

Ejemplo 37.57 (La esfera y el toro por celdas, y dos casos con torsión)**. (a)** $H_*(S^2)$**.** Estructura CW con una 0-celda v y una 2-celda e, pegada por la aplicación constante que manda todo ∂D^2 a v. Entonces

$$C_2 = \mathbb{Z} \text{ (generado por } e), \qquad C_1 = 0, \qquad C_0 = \mathbb{Z} \text{ (generado por } v),$$

y todas las matrices del borde son nulas por falta de sitio: ∂_2 llega a $C_1 = 0$ y ∂_1 sale de $C_1 = 0$. Por tanto

$$H_0(S^2) = \mathbb{Z}, \qquad H_1(S^2) = 0, \qquad H_2(S^2) = \mathbb{Z}.$$

Comprobación del paso 5: $\sum(-1)^n \operatorname{rango} C_n = 1-0+1 = 2$ y $\sum(-1)^n \operatorname{rango} H_n = 1-0+1 = 2$. Coincide con $\chi = 2$ y con el resultado obtenido antes por Mayer–Vietoris, por dos caminos completamente distintos.

(b) $H_*(T^2)$**.** Estructura CW con una 0-celda v, dos 1-celdas a y b, y una 2-celda e pegada según $aba^{-1}b^{-1}$. Así

$$C_2 = \mathbb{Z}, \qquad C_1 = \mathbb{Z}^2, \qquad C_0 = \mathbb{Z}.$$

La matriz ∂_1 es nula: cada arista es un lazo, luego $\partial a = v - v = 0$ y lo mismo para b. La matriz ∂_2 también es nula, y esto sí hay que calcularlo: por el paso 3, el grado en a es $(+1)+(-1)=0$ y el grado en b es $(+1)+(-1)=0$. Por tanto

$$H_0(T^2) = \mathbb{Z}, \qquad H_1(T^2) = \ker \partial_1 / \operatorname{im} \partial_2 = \mathbb{Z}^2/0 = \mathbb{Z}^2, \qquad H_2(T^2) = \ker \partial_2 = \mathbb{Z}.$$

Comprobación: $1-2+1=0$ por celdas y $1-2+1=0$ por homología, y en efecto $\chi(T^2)=0$. Los dos generadores de H_1 son el meridiano y el paralelo, y el generador de H_2 es la superficie entera, que existe precisamente porque $\partial_2 = 0$.

(c) Dónde el paso 3 deja de ser trivial: la botella de Klein. Mismas celdas que el toro, pero la palabra es $abab^{-1}$. Ahora el grado en a es $(+1)+(+1)=2$ y el grado en b es $(+1)+(-1)=0$, luego $\partial_2(e) = 2a$ y la matriz no es nula. Sale

$$H_0 = \mathbb{Z}, \qquad H_1 = \mathbb{Z}^2/\langle 2a \rangle = \mathbb{Z} \oplus \mathbb{Z}/2, \qquad H_2 = \ker \partial_2 = 0.$$

Comprobación: por celdas $1-2+1=0$; por homología rango H_0 − rango H_1 + rango H_2 $= 1-1+0=0$, porque el sumando $\mathbb{Z}/2$ es torsión y no aporta rango. Nótese que la botella tiene la misma χ que el toro y una homología distinta, y que $H_2 = 0$ es la firma algebraica de que no es orientable.

(d) El plano proyectivo. Una celda de cada dimensión 0, 1, 2, con palabra a^2: $\partial_2(e) = 2a$ y $\partial_1 = 0$. Queda $H_0 = \mathbb{Z}$, $H_1 = \mathbb{Z}/2$, $H_2 = 0$, y la comprobación da $1-1+1=1$ por celdas frente a $1-0+0=1$ por homología. Ambas valen $\chi(\mathbb{RP}^2) = 1$.

Observación 37.58. Compárese el esfuerzo. Calcular $H_*(S^2)$ triangulando exige manejar matrices de tamaño 4×6 y 6×4 para el borde del tetraedro; por celdas, las matrices son vacías y la respuesta se lee sin escribir nada. El precio es que hay que saber de antemano una descomposición celular, y que el paso 3 requiere calcular grados, que es donde se cuelan los errores de signo. Por eso el paso 5 no es un adorno: es la única comprobación barata de que las matrices están bien.

Regla: calcular χ a partir de una triangulación

1. **Triangúlese de verdad.** Todas las caras han de ser triángulos, y dos triángulos distintos han de cortarse en una arista común, en un vértice común o en nada. Si las caras son polígonos de más lados, subdivídanse antes: un polígono de s lados se parte en $s-2$ triángulos trazando $s-3$ diagonales desde uno cualquiera de sus vértices.
2. **Cuéntense V, A y C** —vértices, aristas y caras— y hágase $\chi = V - A + C$.
3. **Compruébese que la subdivisión no lo altera.** Añadir un vértice en el interior de una cara triangular suma 1 a V, 3 a A y 2 a C, y $1-3+2=0$; añadir un vértice en mitad de una arista suma 1 a V, 1 a A y 0 a C si la arista está en el borde, y 1, 3 y 2 si es interior. En todos los casos el balance es nulo, y por eso χ no depende de la triangulación.

Ejemplo 37.59 (Los cinco sólidos platónicos, con las caras ya trianguladas)**.** El conteo directo sobre las caras originales da $\chi = 2$ en los cinco casos, pero solo el tetraedro, el octaedro y el

icosaedro son triangulaciones legítimas: los otros dos tienen caras cuadradas y pentagonales y hay que subdividirlas antes de aplicar el paso 1. La tabla recoge las dos cuentas.

	V	A	C	χ	V'	A'	C'	χ
tetraedro	4	6	4	2	4	6	4	2
cubo	8	12	6	2	8	18	12	2
octaedro	6	12	8	2	6	12	8	2
dodecaedro	20	30	12	2	20	54	36	2
icosaedro	12	30	20	2	12	30	20	2

Las columnas primadas son las de la triangulación. Para el cubo, cada una de las 6 caras cuadradas se parte con una diagonal: $A' = 12 + 6 = 18$ y $C' = 6 \cdot 2 = 12$, con lo que $8 - 18 + 12 = 2$. Para el dodecaedro, cada una de las 12 caras pentagonales se parte con 2 diagonales: $A' = 30 + 24 = 54$ y $C' = 12 \cdot 3 = 36$, con lo que $20 - 54 + 36 = 2$. El tetraedro, el octaedro y el icosaedro ya eran triangulaciones y sus columnas no cambian; para el icosaedro se ha marcado con guiones justamente por eso. Los cinco dan 2 porque los cinco son topológicamente una esfera, y las dos columnas de χ coinciden porque subdividir no altera el balance.

37.12. Ejercicios resueltos

Tres problemas donde un invariante algebraico decide una cuestión geométrica.

Ejercicio 37.1. Calcule $\pi_1(S^1)$ y úselo para demostrar que no existe retracción del disco sobre su borde. Deduzca el teorema del punto fijo de Brouwer en dimensión 2.

Solución. $\pi_1(S^1) \cong \mathbb{Z}$, y el isomorfismo manda cada lazo a su número de vueltas. La demostración usa el recubrimiento $\mathbb{R} \to S^1$, $t \mapsto \mathrm{e}^{2\pi i t}$, y el levantamiento único de caminos.

Supongamos que existiese una retracción $r : D^2 \to S^1$, es decir, continua y con $r|_{S^1} = \mathrm{id}$. Componiendo con la inclusión $i : S^1 \hookrightarrow D^2$ se tendría $r \circ i = \mathrm{id}$, y aplicando π_1,

$$\mathbb{Z} = \pi_1(S^1) \xrightarrow{i_*} \pi_1(D^2) = 0 \xrightarrow{r_*} \pi_1(S^1) = \mathbb{Z}$$

con composición la identidad. Imposible: una aplicación que pasa por el grupo trivial no puede ser la identidad de $\mathbb{Z}$.

Brouwer se sigue: si $f : D^2 \to D^2$ continua no tuviese punto fijo, la semirrecta que va de $f(x)$ a x cortaría al borde en un punto $r(x)$, y r sería una retracción continua, que acabamos de descartar. ■

Ejercicio 37.2. Calcule la característica de Euler de la esfera, del toro y de la botella de Klein, y diga qué distingue cada superficie.

Solución. Con cualquier triangulación, $\chi = V - A + C$. Para la esfera, el tetraedro da $4 - 6 + 4 = 2$. Para el toro, la representación como cuadrado con lados identificados da $1 - 2 + 1 = 0$. Para la botella de Klein, también 0.

La característica separa la esfera de las otras dos, pero no distingue el toro de la botella de Klein: para eso hace falta la orientabilidad, o bien el grupo fundamental, que es $\mathbb{Z} \times \mathbb{Z}$ en el toro y $\langle a, b \mid abab^{-1} \rangle$ en la botella, no abeliano. El teorema de clasificación de superficies compactas dice que χ y la orientabilidad juntas determinan la superficie salvo homeomorfismo, y eso es todo lo que hace falta. ■

Ejercicio 37.3. Calcule los grupos de homología de S^2 y explique qué mide cada uno.

Solución. $H_0(S^2) = \mathbb{Z}$, $H_1(S^2) = 0$ y $H_2(S^2) = \mathbb{Z}$, y todos los demás se anulan.

Cada uno mide algo distinto. H_0 cuenta componentes conexas: hay una. H_1 cuenta agujeros rodeables por un lazo: no hay ninguno, todo lazo en la esfera se contrae. H_2 cuenta cavidades encerradas por una superficie: hay una.

La comparación con el toro es instructiva: allí $H_1 = \mathbb{Z} \times \mathbb{Z}$, con los dos lazos independientes, y $H_2 = \mathbb{Z}$. Y $\mathbb{R}^3$ menos un punto es homotópicamente equivalente a S^2 y por tanto tiene la misma homología: quitar un punto de $\mathbb{R}^3$ crea una cavidad pero ningún agujero de lazo, que es justo lo contrario de lo que ocurre al quitar un punto de $\mathbb{R}^2$. ■

37.13. Ejercicios propuestos

37.1. Calcule π_1 de $\mathbb{R}^2$, de $\mathbb{R}^2 \setminus \{0\}$ y de $\mathbb{R}^3 \setminus \{0\}$.

37.2. ¿Son homotópicamente equivalentes una banda cilíndrica y una circunferencia? ¿Y homeomorfos?

37.3. Calcule χ del cubo, del octaedro y del icosaedro.

37.4. ¿Cuál es χ de una superficie orientable de género g?

37.5. Calcule π_1 del toro y diga por qué es abeliano.

37.6. ¿Cuál es π_1 de la botella de Klein? ¿Es abeliano?

37.7. Enuncie el teorema de Seifert y van Kampen y aplíquelo a la unión de dos discos por su borde.

37.8. ¿Qué es un espacio recubridor? Dé el recubridor universal de S^1 y el de la banda de Möbius.

37.9. Calcule $H_n(S^3)$ para todo n.

37.10. ¿Por qué S^2 y el toro no son homeomorfos? Dé dos invariantes que lo demuestren.

37.11. Enuncie el teorema de la bola peluda y diga qué tiene que ver con el viento en la Tierra.

37.12. ¿Qué relación hay entre χ y la homología, según la fórmula de Euler–Poincaré?

Capítulo 38

Teoría de categorías

Lo que necesitas traer

Casi todo el libro. Éste es el capítulo más abstracto y va al final por una razón: sus ejemplos *son* las estructuras de los capítulos anteriores. Hacen falta al menos conjuntos y funciones (capítulo 1), grupos y anillos (capítulos 23 y 24), espacios vectoriales (capítulo 7), espacios topológicos (capítulo 28) y el producto tensorial (capítulo 25). Sin ellos, las categorías son cajas y flechas vacías.

A lo largo del libro han aparecido las mismas construcciones una y otra vez en contextos distintos: productos de conjuntos, de grupos, de espacios vectoriales y de espacios topológicos; cocientes en todos ellos; teoremas de isomorfía con enunciados calcados. La teoría de categorías toma en serio esa repetición y la convierte en objeto de estudio. Su tesis es que lo esencial de un objeto matemático no es de qué está hecho, sino cómo se relaciona con los demás.

38.1. Categorías, funtores y transformaciones naturales

Definición 38.1. Una *categoría* $\mathcal{C}$ consta de objetos, de *morfismos* entre ellos, de una composición asociativa y de una identidad para cada objeto. Un *funtor* $F : \mathcal{C} \to \mathcal{D}$ asigna objetos a objetos y morfismos a morfismos respetando composición e identidades. Una *transformación natural* entre dos funtores es una familia de morfismos que conmuta con todo lo demás.

Idea clave

Qué dice esto en cristiano. Una categoría es un mundo matemático descrito solo por sus flechas: los conjuntos con las aplicaciones, los grupos con los homomorfismos, los espacios topológicos con las funciones continuas. Un funtor es un diccionario entre dos mundos que respeta la estructura, y el ejemplo que ya conocemos es el grupo fundamental, que traduce espacios en grupos y aplicaciones continuas en homomorfismos. Las transformaciones naturales fueron el motivo original de inventar todo esto: hacían falta para dar sentido preciso a la palabra “canónico”, que los matemáticos usaban sin

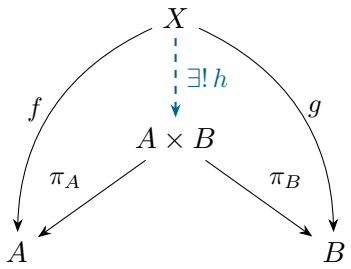


Figura 38.1: La propiedad universal del producto. Dadas f y g, existe una única h que hace conmutar los dos triángulos. Nada de esto menciona qué son los elementos de $A \times B$.

poder definirla.

Ejemplo 38.2 (Categorías, funtores y una definición por flechas)**. (a) Categorías conocidas.** Los conjuntos con las aplicaciones; los grupos con los homomorfismos; los espacios vectoriales sobre K con las aplicaciones lineales; los espacios topológicos con las continuas. También, y esto sorprende, un grupo aislado es una categoría con un solo objeto donde todos los morfismos son invertibles, y un conjunto ordenado es una categoría con un morfismo de a a b exactamente cuando $a \leq b$.

(b) Funtores familiares. El olvido, que manda un grupo a su conjunto subyacente; el grupo fundamental π_1, de espacios punteados a grupos; la homología H_n; el dual $V \mapsto V^*$, que es *contravariante* porque invierte el sentido de las flechas. Que π_1 sea funtor es exactamente lo que permite demostrar el teorema del punto fijo de Brouwer.

(c) El producto, definido sin mirar dentro. El producto de A y B es un objeto P con dos proyecciones tal que toda pareja de morfismos hacia A y B factoriza de forma *única* por P. Esa definición, que no menciona pares ordenados, da el producto cartesiano en conjuntos, el producto directo en grupos y el producto topológico en espacios. Una definición, tres teoremas.

(d) Natural frente a existente. En dimensión finita, V y V^* son isomorfos pero el isomorfismo exige elegir una base: no es natural. En cambio V y V^{**} son naturalmente isomorfos, sin elegir nada. Esa distinción, que hasta 1945 era una intuición vaga, tiene ahora un significado técnico exacto, y darlo fue el motivo por el que Eilenberg y Mac Lane inventaron la teoría.

Cuidado

[Definir por propiedad universal da unicidad, no existencia] Una propiedad universal determina el objeto salvo isomorfismo único, lo que es una garantía muy fuerte; pero no demuestra que exista. La existencia hay que construirla en cada categoría por separado, y hay categorías donde ciertos límites sencillamente no existen. Segundo aviso: la categoría de *todas* las categorías, o la de todos los conjuntos, no es un conjunto y su manejo ingenuo reproduce la paradoja de Russell; hace falta la distinción entre conjuntos y clases, o el recurso a universos de Grothendieck. La abstracción no sale gratis.

Pausa para pensar

Eilenberg y Mac Lane introdujeron las categorías en 1945 y ellos mismos las llamaron, medio en broma, *abstract nonsense*, porque parecían un lenguaje sin contenido propio. El apodo se quedó y ha ido perdiendo la ironía: Grothendieck refundó la geometría algebraica sobre esa base en los años sesenta, los topos dieron una fundamentación alternativa a la teoría de conjuntos, y hoy las categorías monoidales describen la computación cuántica y los lenguajes funcionales de programación están literalmente construidos sobre funtores y mónadas. Es probablemente el ejemplo más claro de una abstracción que nació para ordenar lo que ya se sabía y acabó permitiendo pensar cosas que antes no se podían formular.

38.2. Qué es exactamente una categoría: los axiomas, seis ejemplos concretos y las flechas invertibles

La definición de categoría con la que abre el capítulo cabe en tres líneas y es demasiado corta para trabajar con ella: no dice qué es exactamente "una composición asociativa", ni qué se exige a las identidades, ni aparece por ninguna parte la palabra *isomorfismo*, que sin embargo se usa después decenas de veces. Se despliega aquí todo, con los axiomas escritos uno a uno, con seis ejemplos completos —objetos, morfismos, composición e identidad en cada caso— y con las nociones que el resto del capítulo da por sabidas.

Los axiomas

Definición 38.3 (Categoría)**.** Una *categoría* $\mathcal{C}$ consta de los datos siguientes:

- una colección de *objetos* $A, B, C, \ldots$;
- para cada par ordenado de objetos (A, B), una colección $\mathrm{Hom}_{\mathcal{C}}(A, B)$ de *morfismos* o *flechas* de A en B, y se escribe $f : A \to B$; el objeto A es el *dominio* y B el *codominio*;
- para cada terna (A, B, C), una *composición* $\mathrm{Hom}(B, C) \times \mathrm{Hom}(A, B) \to \mathrm{Hom}(A, C)$, escrita $(g, f) \mapsto g \circ f$;
- para cada objeto A, un morfismo $\mathrm{id}_A \in \mathrm{Hom}(A, A)$.

Y se exigen tres axiomas:

1. *Asociatividad*: $h \circ (g \circ f) = (h \circ g) \circ f$ siempre que las composiciones tengan sentido.
2. *Unidad*: $f \circ \mathrm{id}_A = f$ y $\mathrm{id}_B \circ f = f$ para todo $f : A \to B$.
3. *Disyunción*: los conjuntos $\mathrm{Hom}(A, B)$ son disjuntos dos a dos, de modo que cada morfismo tiene un dominio y un codominio bien determinados.

Cuidado

El tercer axioma parece un tecnicismo y no lo es: sin él, la inclusión $\mathbb{Z} \hookrightarrow \mathbb{Q}$ y la inclusión $\mathbb{Z} \hookrightarrow \mathbb{R}$ podrían ser "el mismo morfismo", y las ecuaciones entre flechas dejarían de tener

sentido. Formalmente se consigue definiendo un morfismo como una terna (A, B, f). Y obsérvese que se ha dicho *colección* y no *conjunto*: los objetos de **Set** no forman un conjunto. Esa distinción es la que se trata en la definición 38.44, y hasta allí basta con saber que existe.

Proposición 38.4 (La identidad es única)**.** *Para cada objeto A hay a lo sumo un morfismo que cumpla el axioma de unidad.*

Demostración. Si u y u' ambos lo cumplen, entonces $u = u \circ u'$ —porque u' es identidad por la derecha— y $u \circ u' = u'$ —porque u lo es por la izquierda—. Luego $u = u'$. Es el argumento clásico del neutro de un monoide, y de hecho una categoría con un solo objeto *es* un monoide. □

Seis categorías concretas

Ejemplo 38.5 (Seis categorías, con sus objetos y sus morfismos)**.** En cada caso hay que decir cuatro cosas —objetos, morfismos, composición, identidad— y comprobar los axiomas. Se hace de una vez.

(a) Set. Objetos: los conjuntos. Morfismos $A \to B$: las aplicaciones $f : A \to B$. Composición: la composición usual de funciones. Identidad: $\mathrm{id}_A(a) = a$. La asociatividad de la composición de funciones es un hecho elemental y los axiomas se cumplen. Ésta es la categoría de referencia y el patrón del que se abstraen todas las demás.

(b) Grp. Objetos: los grupos. Morfismos: los homomorfismos de grupos. Composición y identidad: las de conjuntos, y hay que comprobar que la composición de dos homomorfismos es homomorfismo —lo es— y que id_G lo es —también—. Sin esas dos comprobaciones no habría categoría. Igual se construyen **Ab** (grupos abelianos), **Ring** (anillos y homomorfismos de anillos, que preservan el 1) y $\mathbf{Mod}_R$ (módulos sobre R y aplicaciones R-lineales).

(c) $\mathbf{Vect}_K$. Objetos: los espacios vectoriales sobre un cuerpo K. Morfismos: las aplicaciones lineales. Aquí cada $\mathrm{Hom}(V, W)$ es a su vez un espacio vectorial, y la composición es bilineal: eso es lo que convierte a $\mathbf{Vect}_K$ en una categoría *aditiva* en el sentido de la sección correspondiente.

(d) Top. Objetos: los espacios topológicos. Morfismos: las aplicaciones continuas. La composición de continuas es continua y la identidad es continua, luego hay categoría. Su variante **hTop** tiene los mismos objetos y por morfismos las *clases de homotopía* de aplicaciones continuas; que eso esté bien definido es exactamente el hecho de que la homotopía es compatible con la composición, y es el marco donde π_1 y H_n son funtores.

(e) Un conjunto preordenado como categoría. Sea $(P, \leq)$ con $\leq$ reflexiva y transitiva. Objetos: los elementos de P. Morfismos: hay exactamente un morfismo $a \to b$ si $a \leq b$, y ninguno si no. La composición existe por transitividad y la identidad por reflexividad; la asociatividad es automática porque entre dos objetos hay a lo sumo una flecha y dos flechas paralelas cualesquiera son iguales. Es el ejemplo que enseña que en una categoría los morfismos no tienen por qué ser funciones de ningún tipo.

(f) $\mathbf{Mat}_K$. Objetos: los números naturales $0, 1, 2, \ldots$. Morfismos $n \to m$: las matrices $m \times n$ con entradas en K. Composición: el producto de matrices, asociativo por un cálculo

conocido. Identidad de n: la matriz identidad I_n. Aquí un objeto no es un conjunto ni tiene elementos, y sin embargo la categoría es equivalente a la de espacios vectoriales de dimensión finita, como se ve en la sección sobre equivalencia.

Y dos más, por lo que enseñan. Un *grupo* G es una categoría con un solo objeto $*$, con $\mathrm{Hom}(*,*) = G$, composición el producto e identidad el neutro; los axiomas de categoría son entonces la asociatividad y el neutro del grupo. Y la *categoría opuesta* $\mathcal{C}^{\mathrm{op}}$ tiene los mismos objetos que $\mathcal{C}$ y $\mathrm{Hom}_{\mathcal{C}^{\mathrm{op}}}(A,B) = \mathrm{Hom}_{\mathcal{C}}(B,A)$, con la composición al revés: es la que hace que cada definición venga con su dual gratis.

Isomorfismos

Definición 38.6 (Isomorfismo)**.** Un morfismo $f : A \to B$ es un *isomorfismo* si existe $g : B \to A$ con

$$g \circ f = \mathrm{id}_A \qquad \text{y} \qquad f \circ g = \mathrm{id}_B\,.$$

Un tal g se llama *inverso* de f. Dos objetos son *isomorfos*, y se escribe $A \cong B$, si existe un isomorfismo entre ellos.

Proposición 38.7 (El inverso es único)**.** *Si $f : A \to B$ tiene inversos g y g', entonces $g = g'$. Se escribe $g = f^{-1}$.*

Demostración. Sólo se usan los axiomas de unidad y asociatividad:

$$g = g \circ \mathrm{id}_B = g \circ (f \circ g') = (g \circ f) \circ g' = \mathrm{id}_A \circ g' = g'. \qquad \square$$

Corolario 38.8. *La relación "ser isomorfo" es de equivalencia: id_A es isomorfismo con inverso id_A; si f lo es, f^{-1} lo es con inverso f; y si $f : A \to B$ y $h : B \to C$ lo son, $h \circ f$ lo es con inverso $f^{-1} \circ h^{-1}$, como se comprueba componiendo y cancelando con asociatividad.*

Ejemplo 38.9 (Quién es isomorfismo en cada categoría)**.** En **Set** son las biyecciones; en **Grp**, **Ring** y $\mathbf{Vect}_K$, los homomorfismos biyectivos, porque la inversa de un homomorfismo biyectivo es automáticamente homomorfismo; en **Top** son los *homeomorfismos*, y aquí la biyectividad ya *no* basta: la inversa de una biyección continua no tiene por qué ser continua. En un preorden visto como categoría, $f : a \to b$ es isomorfismo si y sólo si $a \leq b$ y $b \leq a$, de modo que en un orden parcial los únicos isomorfismos son las identidades. En $\mathbf{Mat}_K$ son las matrices cuadradas invertibles, y dos objetos n y m son isomorfos si y sólo si $n = m$. Y en un grupo visto como categoría de un objeto, *todos* los morfismos son isomorfismos: eso es exactamente lo que distingue a un grupo de un monoide, y una categoría con esa propiedad se llama *grupoide*.

Cuidado

Ser isomorfismo es estrictamente más fuerte que ser monomorfismo y epimorfismo a la vez, según la definición de la sección correspondiente: la aplicación $[0,1) \to S^1$ es mono y epi en **Top** y no es isomorfismo, y la inclusión $\mathbb{Z} \hookrightarrow \mathbb{Q}$ lo es en **Ring**. La costumbre conjuntista de identificar "inyectivo y sobreyectivo" con "invertible" es una propiedad

de **Set**, no un teorema general.

Funtores y transformaciones naturales, con los diagramas escritos

Definición 38.10 (Funtor covariante y contravariante)**.** Un *funtor covariante* $F : \mathcal{C} \to \mathcal{D}$ asigna a cada objeto A de $\mathcal{C}$ un objeto $F(A)$ de $\mathcal{D}$ y a cada morfismo $f : A \to B$ un morfismo $F(f) : F(A) \to F(B)$, de modo que

$$F(g \circ f) = F(g) \circ F(f) \qquad \text{y} \qquad F(\mathrm{id}_A) = \mathrm{id}_{F(A)}\,.$$

Un *funtor contravariante* hace lo mismo pero invirtiendo el sentido: $F(f) : F(B) \to F(A)$ y

$$F(g \circ f) = F(f) \circ F(g),$$

que equivale a un funtor covariante $\mathcal{C}^{\mathrm{op}} \to \mathcal{D}$.

Ejemplo 38.11 (Cuatro funtores comprobados)**. (a) El olvido** $U : \mathbf{Grp} \to \mathbf{Set}$ manda cada grupo a su conjunto subyacente y cada homomorfismo a sí mismo visto como aplicación. Las dos condiciones son inmediatas.

(b) El grupo fundamental $\pi_1 : \mathbf{Top}_* \to \mathbf{Grp}$, sobre espacios con punto base, con $\pi_1(f)[\gamma] = [f \circ \gamma]$. Que $(g \circ f)_* = g_* \circ f_*$ y que $(\mathrm{id})_* = \mathrm{id}$ es literalmente lo que se usa para demostrar Brouwer: sin funtorialidad no hay argumento.

(c) El dual, contravariante. $V \mapsto V^* = \mathrm{Hom}(V, K)$ y, para $f : V \to W$, la traspuesta $f^* : W^* \to V^*$ dada por $f^*(\varphi) = \varphi \circ f$. Entonces $(g \circ f)^* = f^* \circ g^*$: la flecha se da la vuelta, y por eso el dual es contravariante.

(d) $\mathrm{Hom}(A, -)$ **y** $\mathrm{Hom}(-, A)$**.** El primero es covariante, mandando $f : X \to Y$ a la posposición $\varphi \mapsto f \circ \varphi$; el segundo es contravariante, mandando f a la preposición $\varphi \mapsto \varphi \circ f$. Son los dos funtores sobre los que se enuncia el lema de Yoneda.

Definición 38.12 (Transformación natural)**.** Sean $F, G : \mathcal{C} \to \mathcal{D}$ funtores covariantes. Una *transformación natural* $\eta : F \Rightarrow G$ es una familia de morfismos $\eta_A : F(A) \to G(A)$, uno por cada objeto A de $\mathcal{C}$, tal que para todo morfismo $f : A \to B$ de $\mathcal{C}$ se cumple

$$G(f) \circ \eta_A = \eta_B \circ F(f),$$

es decir, el cuadrado de vértices $F(A), F(B), G(A), G(B)$ conmuta. Si además cada η_A es un isomorfismo, η es un *isomorfismo natural* y se escribe $F \cong G$.

Ejemplo 38.13 (El ejemplo que motivó la teoría)**.** Sea $\mathcal{C} = \mathbf{Vect}_K^{\mathrm{fin}}$, $F = \mathrm{id}$ y $G = (-)^{**}$ el bidual. Defínase $\eta_V : V \to V^{**}$ por $\eta_V(v)(\varphi) = \varphi(v)$: "evaluar en v". No se ha elegido ninguna base. Para $f : V \to W$ y $v \in V$, $\varphi \in W^*$:

$$(f^{**}(\eta_V(v)))(\varphi) = \eta_V(v)(f^*\varphi) = (\varphi \circ f)(v) = \varphi(f(v)) = \eta_W(f(v))(\varphi),$$

luego el cuadrado conmuta y η es natural; y en dimensión finita cada η_V es biyectiva, luego $\mathrm{id} \cong (-)^{**}$. En cambio, V y V^* tienen la misma dimensión y son isomorfos, pero no hay ninguna familia de isomorfismos $V \to V^*$ que sea natural: cada uno exige elegir una base. Ésa es la diferencia técnica entre "isomorfos" y "canónicamente isomorfos", y darle sentido preciso fue el motivo por el que se inventaron las categorías.

Objetos y construcciones definidos por flechas

Definición 38.14 (Objeto inicial y objeto final)**.** Un objeto I de $\mathcal{C}$ es *inicial* si para cada objeto X existe *exactamente un* morfismo $I \to X$. Es *final* —o *terminal*— si para cada X existe exactamente un morfismo $X \to I$. Un objeto que es a la vez inicial y final se llama *objeto cero*.

Ejemplo 38.15 (Iniciales y finales en las seis categorías)**.** En **Set**: inicial $\varnothing$ —hay una única aplicación, la vacía, desde él a cualquier conjunto— y final cualquier conjunto de un elemento. En **Grp** y $\mathbf{Vect}_K$ el grupo trivial y el espacio nulo son ambas cosas: son objetos cero. En **Ring** el inicial es $\mathbb{Z}$, porque un homomorfismo de anillos ha de mandar 1 a 1 y eso lo determina por completo sobre los enteros, y el final es el anillo nulo. En **Top**, el vacío y el punto. En un preorden, un mínimo y un máximo si existen. Y en $\mathbf{Mat}_K$ el objeto 0 es inicial y final a la vez, porque entre 0 y n hay exactamente una matriz, la vacía. La unicidad salvo isomorfismo único está demostrada en los ejercicios resueltos del capítulo, y el argumento es el mismo que el de la proposición 38.7.

Definición 38.16 (Producto y coproducto)**.** Un *producto* de dos objetos A y B es un objeto P junto con dos morfismos $\pi_A : P \to A$ y $\pi_B : P \to B$ tal que para todo objeto X y todo par de morfismos $f : X \to A$, $g : X \to B$ existe un *único* $h : X \to P$ con $\pi_A \circ h = f$ y $\pi_B \circ h = g$. Se escribe $P = A \times B$.

Un *coproducto* es la noción dual: un objeto S con $\iota_A : A \to S$ e $\iota_B : B \to S$ tal que para todo X y todo par $f : A \to X$, $g : B \to X$ existe un único $h : S \to X$ con $h \circ \iota_A = f$ y $h \circ \iota_B = g$. Se escribe $S = A \sqcup B$. Es exactamente el producto en $\mathcal{C}^{\mathrm{op}}$.

Ejemplo 38.17 (El coproducto, que es donde está la sorpresa)**.** En **Set** el coproducto es la unión disjunta, con las inclusiones: dar una función desde $A \sqcup B$ es dar una función desde A y otra desde B, sin más. En $\mathbf{Vect}_K$ el coproducto de dos espacios es la suma directa, *el mismo objeto* que el producto, porque una aplicación lineal desde $V \oplus W$ equivale al par de sus restricciones. Pero en **Grp** el coproducto es el *producto libre* $G * H$ —el grupo de las palabras alternadas en letras de G y de H, sin más relaciones que las de cada factor—, un objeto infinito y no abeliano incluso para $G = H = \mathbb{Z}/2$; y en anillos conmutativos es el producto tensorial. Invertir las flechas no es una simetría inocente: puede cambiar por completo el tamaño del objeto resultante.

Regla: comprobar que algo cumple una propiedad universal

Casi todas las verificaciones del capítulo tienen esta forma, y siempre son los mismos cuatro pasos. Sea cual sea la construcción —producto, coproducto, objeto libre, límite—:

1. **Escríbase el dato completo**: no basta con el objeto, hay que decir también los morfismos estructurales que lo acompañan (las proyecciones, las inclusiones, la unidad). Un producto sin sus proyecciones no es nada.
2. **Existencia**: dado el objeto de prueba X con sus flechas, constrúyase el morfismo h y compruébese que los triángulos conmutan. Aquí es donde se usa la descripción

concreta del objeto.

3. **Unicidad**: supóngase que h' también sirve y dedúzcase $h = h'$. Casi siempre sale de que las flechas estructurales son conjuntamente monomórficas, o directamente de cancelar.
4. **Conclúyase la unicidad salvo isomorfismo único**: si P y P' cumplen lo mismo, cada uno induce una flecha hacia el otro, la composición cumple la propiedad que caracteriza a id_P, y por la unicidad del paso 3 es id_P. Luego son isomorfos, y por un único isomorfismo compatible con la estructura.

Adviértase lo que estos cuatro pasos *no* dan: la existencia del objeto en la categoría de turno. En **Grp** el producto existe y hay que construirlo; en la categoría de cuerpos no existe, y ningún razonamiento por flechas lo fabricará.

38.3. Límites, colimites y adjunciones

Definición 38.18. Un *límite* de un diagrama es un objeto con morfismos hacia él que factoriza de manera única cualquier otro cono sobre el diagrama; los casos particulares son el producto, el ecualizador y el producto fibrado. Un *colimite* es la noción dual, con las flechas invertidas: coproducto, coecualizador y cociente. Dos funtores $F : \mathcal{C} \to \mathcal{D}$ y $G : \mathcal{D} \to \mathcal{C}$ son *adjuntos* si $\mathrm{Hom}_{\mathcal{D}}(F(X), Y) \cong \mathrm{Hom}_{\mathcal{C}}(X, G(Y))$ de manera natural.

Idea clave

Qué dice esto en cristiano. Límite es "lo más grande que encaja debajo" y colimite "lo más pequeño que encaja encima". Un producto reúne información de varios objetos; un coproducto los pega. Y la adjunción, que es la noción más profunda de toda la teoría, formaliza la idea de construcción *óptima*: el grupo libre sobre un conjunto es el grupo más económico que contiene a ese conjunto, y esa economía se expresa diciendo que los homomorfismos desde él corresponden exactamente a las aplicaciones desde el conjunto. Casi toda construcción canónica de la matemática resulta ser un adjunto.

Ejemplo 38.19 (La misma construcción en cinco categorías)**. (a) Productos.** En conjuntos, el producto cartesiano; en grupos, el producto directo; en espacios vectoriales, la suma directa finita; en espacios topológicos, el producto con la topología producto; en conjuntos ordenados, el ínfimo. Una sola definición por flechas, cinco objetos distintos.

(b) Coproductos, que sorprenden más. En conjuntos es la unión disjunta; en espacios vectoriales, otra vez la suma directa (producto y coproducto coinciden en el caso finito); pero en grupos es el *producto libre*, un objeto enorme y no conmutativo, y en anillos conmutativos es el producto tensorial. La dualidad no es simetría: invertir las flechas puede dar algo mucho más complicado.

(c) La adjunción libre–olvido. Sea F el funtor que manda un conjunto a su grupo libre y U el que olvida la estructura. Entonces $\mathrm{Hom}_{\mathrm{Grp}}(F(X), G) \cong \mathrm{Hom}_{\mathrm{Set}}(X, U(G))$: dar un

homomorfismo desde el grupo libre es exactamente dar imágenes arbitrarias a los generadores. Eso es precisamente lo que hace útiles las presentaciones por generadores y relaciones.

(d) Más adjunciones familiares. Abelianización y olvido; compleción de un espacio métrico e inclusión; producto tensorial y funtor Hom, cuya adjunción se escribe $\mathrm{Hom}(A \otimes B, C) \cong \mathrm{Hom}(A, \mathrm{Hom}(B, C))$ y es la razón formal de que una aplicación bilineal equivalga a una lineal desde el tensorial.

Cuidado

[Los adjuntos preservan límites, pero cada uno los suyos] El adjunto por la derecha preserva límites y el de la izquierda colimites, nunca al revés. De ahí sale, por ejemplo, que el producto tensorial no preserve inyecciones (no es exacto por la izquierda), lo que obliga a introducir los funtores derivados y toda el álgebra homológica. Segundo aviso: que un límite exista no está garantizado. En la categoría de cuerpos no hay productos, y en la de grupos finitos no hay coproductos infinitos. La abstracción organiza lo que existe; no lo crea.

38.4. El lema de Yoneda: un objeto es lo que hace

Si hay un resultado que justifica por sí solo la existencia de la teoría de categorías, es éste. Su enunciado parece una tautología técnica y su contenido es una tesis filosófica: un objeto queda completamente determinado por la manera en que se relaciona con todos los demás. No hay que abrirlo para saber qué es; basta con observar cómo interactúa.

Teorema 38.20 (Yoneda)**.** *Sea $\mathcal{C}$ una categoría localmente pequeña, A un objeto y $F\colon \mathcal{C} \to \mathbf{Set}$ un funtor. Entonces existe una biyección natural*

$$\mathrm{Nat}\big(\mathrm{Hom}(A, -),\, F\big) \;\cong\; F(A),$$

dada por $\eta \mapsto \eta_A(\mathrm{id}_A)$. En particular, el funtor $A \mapsto \mathrm{Hom}(A, -)$ es plenamente fiel, de modo que $\mathrm{Hom}(A, -) \cong \mathrm{Hom}(B, -)$ implica $A \cong B$.

Idea clave

Qué dice esto en cristiano. El objeto $\mathrm{Hom}(A, -)$ es la ficha de A: la lista de todas las flechas que salen de A hacia cualquier sitio. El lema dice dos cosas. La primera es que cualquier manera coherente de asignar a cada objeto un elemento de F, compatible con todas las flechas, está determinada por un único dato: adónde va la identidad de A. Toda esa maquinaria infinita colapsa en un solo elemento de $F(A)$. La segunda, que es la consecuencia famosa, es que dos objetos con la misma ficha son el mismo objeto. Un grupo no es un conjunto con una operación: es aquello que responde de cierta manera cuando se le pregunta desde fuera.

Regla: Usar Yoneda en la práctica

1. **Para demostrar que dos objetos son isomorfos**, comprobar que sus funtores Hom lo son. A menudo es mucho más fácil que construir el isomorfismo a mano.
2. **Para definir un objeto**, describir su funtor Hom: si existe algo que lo represente, es único salvo isomorfismo único. Eso es lo que significa una *propiedad universal*.
3. **Para verificar una identidad entre morfismos**, comprobarla sobre elementos generalizados, es decir, sobre flechas $X \to A$ para todo X.
4. **Para saber si un funtor es representable**, buscar un objeto A y un elemento de $F(A)$ que sea *universal*, en el sentido de que todo otro se obtenga de él de una única manera.
5. **Recordar la versión contravariante**: $\mathrm{Nat}(\mathrm{Hom}(-, A), F) \cong F(A)$ para funtores contravariantes, que es la que se usa en geometría algebraica.
6. **Comprobar la naturalidad**, que es la parte donde de verdad hay algo que demostrar: los cuadrados deben conmutar para toda flecha.
7. **No esperar contenido computacional**: Yoneda no calcula nada, reorganiza. Su valor es que convierte construcciones en verificaciones.

Ejemplo 38.21 (Objetos definidos por lo que hacen)**. (a) El producto.** $A \times B$ se define pidiendo que $\mathrm{Hom}(X, A \times B) \cong \mathrm{Hom}(X, A) \times \mathrm{Hom}(X, B)$ de forma natural en X. Por Yoneda, si tal objeto existe es único salvo isomorfismo único, y no hace falta decir qué son sus elementos. En **Set** resulta el producto cartesiano; en **Grp**, el producto directo; en **Top**, el producto con la topología producto —que se define así precisamente para que esto funcione, y no al revés—.

(b) Los conjuntos, vistos desde fuera. En **Set**, tomando A el conjunto de un elemento, $\mathrm{Hom}(\{*\}, X) \cong X$: los elementos de un conjunto son las flechas desde el punto. Por eso en cualquier categoría se llama *elemento generalizado* de X a una flecha $T \to X$, y por eso las demostraciones categóricas pueden imitar los argumentos con elementos aunque los objetos no los tengan.

(c) El anillo de polinomios. $\mathrm{Hom}_{\mathbf{Ring}}(\mathbb{Z}[x], R) \cong R$: dar un homomorfismo desde $\mathbb{Z}[x]$ es lo mismo que elegir la imagen de x, que puede ser cualquier elemento de R. El anillo de polinomios queda caracterizado por eso y nada más, sin mencionar ni una sola vez qué es un polinomio.

(d) El grupo fundamental. En topología, $[S^1, X]$, las clases de homotopía de aplicaciones desde la circunferencia, *es* $\pi_1(X)$ salvo conjugación. La circunferencia representa el funtor "lazos módulo homotopía", y ése es el sentido preciso en que S^1 es el objeto que detecta agujeros.

Cuidado

[Contraejemplo: no todo funtor es representable] El funtor que asigna a cada conjunto el conjunto de sus subconjuntos, $X \mapsto \mathcal{P}(X)$, *no* es representable de forma covariante:

no existe ningún conjunto A con $\mathrm{Hom}(A, X) \cong \mathcal{P}(X)$ naturalmente, porque para X de un elemento el lado derecho tiene dos elementos y el izquierdo tendría que tener uno o infinitos según el tamaño de A. Contravariantemente sí lo es, con $A = \{0, 1\}$, y de ahí sale la idea de *clasificador de subobjetos* que estructura los topos. La lección es que Yoneda dice qué ocurre *si* el objeto existe, no que exista.

Cuidado

[Contraejemplo: la naturalidad no es un adorno] Para grupos abelianos finitos se cumple $G \cong \mathrm{Hom}(G, \mathbb{C}^*)$, pero el isomorfismo depende de la elección de una base y *no* es natural. En cambio $G \cong \mathrm{Hom}(\mathrm{Hom}(G, \mathbb{C}^*), \mathbb{C}^*)$ sí lo es. Toda la diferencia entre "son isomorfos" y "son el mismo objeto" está en esa palabra, y Yoneda sólo funciona con isomorfismos naturales. El caso análogo en álgebra lineal es $V \cong V^*$ frente a $V \cong V^{**}$.

Pausa para pensar

Nobuo Yoneda nunca escribió el lema. Se lo explicó a Saunders Mac Lane en 1954 en el andén de la Gare du Nord de París, mientras esperaba un tren, y fue Mac Lane quien lo difundió y le puso su nombre. La conversación empezó en un café y continuó en la estación hasta que el tren se llevó a Yoneda; Mac Lane contaba que la segunda mitad, la importante, la recibió a través de la ventanilla. Es el teorema con el origen más literalmente fugaz de las matemáticas modernas, y también el que más veces se ha calificado de trivial por gente que había tardado meses en entenderlo.

38.5. Categorías abelianas y sucesiones exactas

Hay una familia de categorías donde se puede hacer casi todo lo que se hace con grupos abelianos: sumar morfismos, hablar de núcleos e imágenes, y encadenarlos en sucesiones. Se llaman abelianas, y son el marco donde vive el álgebra homológica, que es la maquinaria que convierte problemas geométricos en cálculos.

Definición 38.22. Una *categoría aditiva* es una categoría en la que cada $\mathrm{Hom}(A, B)$ es un grupo abeliano, la composición es bilineal, existe un objeto cero y existen sumas directas finitas. Es *abeliana* si además todo morfismo tiene núcleo y conúcleo, todo monomorfismo es el núcleo de alguien y todo epimorfismo es el conúcleo de alguien. Una sucesión $A \xrightarrow{f} B \xrightarrow{g} C$ es *exacta* en B si $\mathrm{im}\, f = \ker g$.

Idea clave

Qué dice esto en cristiano. La lista de axiomas parece arbitraria y no lo es: dice exactamente "pórtate como los grupos abelianos". La última condición, la que suena más rara, es la que evita el desastre. En la categoría de grupos abelianos *libres*, la multiplicación por 2 en $\mathbb{Z}$ es monomorfismo pero no es el núcleo de nada, porque el

cociente $\mathbb{Z}/2\mathbb{Z}$ no es libre y se sale de la categoría; por eso esa categoría no es abeliana y no se puede hacer homología en ella. Exigir que los monos sean núcleos es exigir que los cocientes existan y sean decentes. Y una sucesión exacta es simplemente un modo de decir "lo que sale de aquí es exactamente lo que muere allí": una contabilidad sin pérdidas.

Regla: Leer una sucesión exacta

1. $0 \to A \xrightarrow{f} B$ **exacta** significa que f es inyectiva, porque el núcleo de f es la imagen del cero.
2. $A \xrightarrow{f} B \to 0$ **exacta** significa que f es sobreyectiva.
3. $0 \to A \to B \to C \to 0$ **exacta** es una *sucesión exacta corta*: A es un subobjeto de B y C es el cociente B/A.
4. **Contar dimensiones**: en espacios vectoriales, una sucesión exacta corta obliga a $\dim B = \dim A + \dim C$, y una sucesión exacta larga a que la suma alternada sea cero.
5. **Buscar dónde se escinde**: si existe $s : C \to B$ con $g \circ s = \mathrm{id}$, entonces $B \cong A \oplus C$. En espacios vectoriales siempre ocurre; en grupos abelianos no.
6. **Aplicar un funtor** y ver qué se rompe: los exactos preservan la sucesión entera, y los que no lo son la truncan.
7. **Medir el fallo** con los funtores derivados, que son la respuesta a la pregunta "¿cuánto se pierde?".

Ejemplo 38.23 (Sucesiones exactas que se leen solas)**. (a) El teorema del rango.** La sucesión $0 \to \ker T \to V \xrightarrow{T} \operatorname{im} T \to 0$ es exacta, y contar dimensiones da $\dim V = \dim \ker T + \operatorname{rango} T$. El teorema más usado del álgebra lineal es una sucesión exacta corta disfrazada.

(b) Una que no se escinde. $0 \to \mathbb{Z} \xrightarrow{2} \mathbb{Z} \to \mathbb{Z}/2\mathbb{Z} \to 0$ es exacta y no se escinde: $\mathbb{Z}$ no es isomorfo a $\mathbb{Z} \oplus \mathbb{Z}/2\mathbb{Z}$, porque el primero no tiene torsión. En espacios vectoriales esto no puede pasar, y por eso el álgebra lineal es fácil y la de módulos no.

(c) El funtor Hom **y su fallo.** Aplicar $\mathrm{Hom}(\mathbb{Z}/2\mathbb{Z}, -)$ a la sucesión anterior deja $0 \to 0 \to 0 \to \mathbb{Z}/2\mathbb{Z}$, que ya no es exacta por la derecha. Ese defecto es $\mathrm{Ext}^1(\mathbb{Z}/2\mathbb{Z}, \mathbb{Z}) = \mathbb{Z}/2\mathbb{Z}$, y mide exactamente las extensiones no escindidas que existen. El grupo de error es un objeto con significado, no un residuo.

(d) Mayer–Vietoris. En topología, cubrir un espacio con dos abiertos produce una sucesión exacta larga que relaciona la homología del todo con la de las partes y la de su intersección. Es la herramienta de cálculo por excelencia, y funciona porque la homología toma valores en una categoría abeliana.

Cuidado

[Contraejemplo: exactitud no es conmutatividad de las operaciones] Es tentador pensar que un funtor razonable respeta las sucesiones exactas. Casi ninguno lo hace. El producto tensorial $-\otimes \mathbb{Z}/2\mathbb{Z}$ aplicado a $0 \to \mathbb{Z} \xrightarrow{2} \mathbb{Z}$ da $\mathbb{Z}/2\mathbb{Z} \xrightarrow{0} \mathbb{Z}/2\mathbb{Z}$, donde la flecha que era inyectiva se ha vuelto nula. El funtor derivado que mide ese destrozo se llama Tor, y el hecho de que $\mathrm{Tor}_1(\mathbb{Z}/2\mathbb{Z}, \mathbb{Z}/2\mathbb{Z}) = \mathbb{Z}/2\mathbb{Z}$ no sea cero es lo que hace que la fórmula de los coeficientes universales tenga un término corrector.

Pausa para pensar

Las categorías abelianas las axiomatizó Grothendieck en 1957, en el artículo conocido como Tôhoku por la revista donde apareció. Tenía veintinueve años y lo escribió, según contó después, para no tener que repetir tres veces la misma demostración —una para grupos abelianos, otra para módulos, otra para haces—. El artículo cambió el álgebra del siglo, y su motivación declarada era la pereza. Es un buen recordatorio de que la abstracción no suele nacer del amor a lo abstracto, sino de la irritación ante la repetición.

38.6. Mónadas: estructura que se compone consigo misma

Una mónada es lo que queda de una adjunción cuando se olvida la mitad. Suena a pérdida y resulta ser una ganancia, porque el objeto que queda es manejable y aparece en sitios donde no hay ninguna adjunción a la vista, empezando por la programación.

Definición 38.24. Una *mónada* sobre una categoría $\mathcal{C}$ es un funtor $T\colon \mathcal{C} \to \mathcal{C}$ junto con transformaciones naturales $\eta\colon \mathrm{Id} \Rightarrow T$ (la *unidad*) y $\mu\colon T^2 \Rightarrow T$ (la *multiplicación*) que hacen conmutar los diagramas de asociatividad $\mu \circ T\mu = \mu \circ \mu T$ y de unidad $\mu \circ T\eta = \mu \circ \eta T = \mathrm{id}$.

Idea clave

Qué dice esto en cristiano. Los axiomas son los de un monoide, escritos con funtores en lugar de con elementos: η hace de neutro y μ de producto. De ahí el nombre, que en realidad debería ser "monoide en la categoría de endofuntores". Lo que hace una mónada es envolver: $T(X)$ es "X con algo añadido". La unidad dice cómo meter un elemento desnudo dentro del envoltorio, y la multiplicación dice cómo aplastar dos capas de envoltorio en una sola. Ejemplo doméstico: si $T(X)$ son las listas de elementos de X, entonces η convierte x en la lista $[x]$ y μ convierte una lista de listas en una lista, concatenando. Los axiomas dicen que concatenar es asociativo y que la lista de un elemento no estorba.

Ejemplo 38.25 (Cuatro mónadas y de dónde salen)**. (a) De una adjunción.** Si $F \dashv U$, entonces $T = U \circ F$ es una mónada. Con F el grupo libre y U el olvido, $T(X)$ es el conjunto subyacente al grupo libre sobre X. Toda adjunción produce una mónada, y toda mónada proviene de al menos una adjunción, aunque en general de varias.

(b) La mónada de las partes. $T(X) = \mathcal{P}(X)$, con $\eta(x) = \{x\}$ y μ la unión. Sus morfismos de Kleisli $X \to \mathcal{P}(Y)$ son exactamente las relaciones: la mónada convierte funciones no deterministas en funciones ordinarias.

(c) La mónada del error. $T(X) = X \sqcup \{*\}$, donde $*$ significa "ha fallado". Componer dos funciones que pueden fallar dentro de esta mónada equivale a propagar el fallo sin escribir la comprobación cada vez. Es la razón de que las mónadas aparezcan en lenguajes de programación funcional: hacen componible lo que no lo era.

(d) La mónada de la distribución. $T(X)$ son las distribuciones de probabilidad finitas sobre X; $\eta(x)$ es la delta en x y μ es la mezcla, es decir, la ley de la probabilidad total. Los morfismos de Kleisli $X \to T(Y)$ son los núcleos de Markov, y componerlos es multiplicar matrices estocásticas. La teoría de cadenas de Markov del capítulo 22 es, vista así, cálculo en la categoría de Kleisli de esta mónada.

Cuidado

[Contraejemplo: no todo funtor con unidad es mónada] Tomar $T(X) = X \times X$ con $\eta(x) = (x, x)$ parece prometedor, pero no hay ninguna $\mu\colon T^2 \Rightarrow T$ que cumpla los axiomas de forma natural: $T^2(X) = X^4$ y cualquier elección de dos de las cuatro coordenadas rompe la asociatividad o la unidad. La existencia de μ es una condición fuerte, no un trámite, y es exactamente lo que distingue a las estructuras que se componen de las que no.

Pausa para pensar

Las mónadas se llamaron durante años "tripletas estándar", nombre que Mac Lane consideraba horrible. El actual se lo debemos a él, y lo tomó prestado de Leibniz, cuyas mónadas eran sustancias simples sin ventanas. La conexión conceptual es dudosa y el nombre se quedó. Cuando Moggi las importó a la informática en 1989 para modelar efectos computacionales, y Wadler las popularizó poco después, muchos programadores aprendieron teoría de categorías sin saberlo, por la vía de necesitar componer funciones que devuelven errores.

38.7. Monomorfismos, epimorfismos y equivalencia de categorías

Una categoría solo conoce flechas. Así que las nociones habituales —inyectivo, sobreyectivo, ser el mismo objeto— hay que redefinirlas sin mirar dentro, y el resultado no siempre coincide con la intuición conjuntista.

Definición 38.26 (Mono y epi)**.** $f : A \to B$ es **monomorfismo** si es cancelable por la izquierda: $fg = fh \Rightarrow g = h$ para todo par $g, h : C \to A$. Es **epimorfismo** si lo es por la derecha: $gf = hf \Rightarrow g = h$.

En **Set**, **Grp**, **Ab** y **Vect** los monos son exactamente los inyectivos y los epis los

sobreyectivos. Pero eso es un teorema en cada caso, no una definición, y falla en cuanto uno se aleja.

Ejemplo 38.27 (Dos sorpresas)**. (a) En Ring, la inclusión $\mathbb{Z} \hookrightarrow \mathbb{Q}$ es epimorfismo y no es sobreyectiva.** En efecto, si dos homomorfismos de anillos $g, h : \mathbb{Q} \to R$ coinciden en $\mathbb{Z}$, entonces coinciden en todo $\mathbb{Q}$: el valor en p/q está forzado por $g(p/q) \cdot g(q) = g(p)$ y por que $g(q)$ sea invertible. Un homomorfismo desde $\mathbb{Q}$ está determinado por su restricción a $\mathbb{Z}$, y eso es exactamente ser cancelable por la derecha.

(b) En Top, la aplicación $[0, 1) \to S^1$, $t \mapsto e^{2\pi i t}$, es mono y epi y no es isomorfismo. Es biyectiva y continua, luego cancelable por ambos lados; pero su inversa no es continua, y en **Top** un isomorfismo es un homeomorfismo. Así que "mono + epi" no implica "iso", aunque en **Set** y en toda categoría abeliana sí.

Definición 38.28 (Funtores fieles, plenos y esencialmente sobreyectivos)**.** $F : \mathcal{C} \to \mathcal{D}$ es **fiel** si es inyectivo sobre cada conjunto de morfismos, **pleno** si es sobreyectivo sobre ellos, y **esencialmente sobreyectivo** si todo objeto de $\mathcal{D}$ es isomorfo a algún $F(A)$.

Definición 38.29 (Equivalencia)**.** $\mathcal{C}$ y $\mathcal{D}$ son **equivalentes** si existen $F : \mathcal{C} \to \mathcal{D}$ y $G : \mathcal{D} \to \mathcal{C}$ con isomorfismos *naturales* $GF \cong \mathrm{Id}_{\mathcal{C}}$ y $FG \cong \mathrm{Id}_{\mathcal{D}}$.

Teorema 38.30 (Caracterización)**.** *F es una equivalencia si y solo si es pleno, fiel y esencialmente sobreyectivo.*

Idea clave

La diferencia entre *isomorfismo* y *equivalencia* de categorías es la lección más importante del tema. Un isomorfismo pide una biyección entre los objetos; una equivalencia solo pide que cada objeto de un lado tenga un compañero *isomorfo* al otro lado. Y eso es lo correcto, porque en matemáticas nunca se distingue entre objetos isomorfos: exigir igualdad de objetos es una condición que no respeta la práctica.

Ejemplo 38.31 (Espacios vectoriales y matrices)**.** Sea $\mathcal{D}$ la categoría cuyos objetos son los números naturales $0, 1, 2, \ldots$ y cuyos morfismos $n \to m$ son las matrices $m \times n$ sobre K, con la composición dada por el producto. Y sea $\mathcal{C} = \mathbf{Vect}_K^{\mathrm{fin}}$.

El funtor $\mathcal{D} \to \mathcal{C}$ que manda $n \mapsto K^n$ y cada matriz a su aplicación lineal es pleno (toda aplicación lineal es una matriz), fiel (matrices distintas dan aplicaciones distintas) y esencialmente sobreyectivo (todo espacio de dimensión finita es isomorfo a algún K^n). Luego es una equivalencia.

Pero no es un isomorfismo, ni puede serlo: en $\mathcal{C}$ hay una clase propia de espacios de dimensión 2 —uno por cada conjunto de dos elementos que se quiera usar de base— y en $\mathcal{D}$ hay exactamente un objeto "2". No hay biyección posible entre una clase propia y un conjunto. Y sin embargo, hacer álgebra lineal en un sitio o en el otro es literalmente lo mismo. $\mathcal{D}$ es un **esqueleto** de $\mathcal{C}$: un representante por clase de isomorfía.

Observación 38.32 (Una dualidad de verdad). Las equivalencias más valiosas invierten las flechas. La categoría de álgebras de Boole finitas es *antiequivalente* a la de conjuntos finitos, vía $B \mapsto$ átomos(B): el álgebra de $8 = 2^3$ elementos corresponde al conjunto de 3 elementos, y cada homomorfismo de álgebras corresponde a una función en sentido contrario. La misma idea reaparece por todas partes: espacios compactos de Hausdorff y C^*-álgebras conmutativas (Gelfand), variedades afines y álgebras reducidas finitamente generadas (Hilbert), esquemas afines y anillos conmutativos (Grothendieck). "Geometría y álgebra son la misma cosa con las flechas al revés" no es un eslogan: es una lista de teoremas.

38.8. Resoluciones y funtores derivados: de dónde salen Ext y Tor

Hemos usado $\mathrm{Ext}^1(\mathbb{Z}/2,\mathbb{Z}) = \mathbb{Z}/2$ y $\mathrm{Tor}_1(\mathbb{Z}/2,\mathbb{Z}/2) = \mathbb{Z}/2$ como valores conocidos. Salen de una construcción única, y conviene verla porque es el mecanismo que produce casi todas las sucesiones exactas largas de la matemática moderna.

Idea clave

El problema es siempre el mismo. Un funtor como $\mathrm{Hom}(-,M)$ o $-\otimes M$ es *exacto por un lado* pero no por el otro: aplicado a una sucesión exacta corta produce una sucesión que se rompe en un extremo. Los funtores derivados miden exactamente ese fallo, y lo hacen de manera que la sucesión rota se prolonga en una sucesión exacta *larga*. El defecto deja de ser un obstáculo y pasa a ser un invariante.

Definición 38.33 (Objetos proyectivos e inyectivos, resoluciones)**.** P es **proyectivo** si $\mathrm{Hom}(P,-)$ es exacto (equivalentemente, si toda flecha desde P se levanta a lo largo de cualquier epimorfismo). I es **inyectivo** si lo es $\mathrm{Hom}(-,I)$. Una **resolución proyectiva** de M es una sucesión exacta

$$\cdots \to P_2 \to P_1 \to P_0 \to M \to 0$$

con todos los P_i proyectivos.

En **Ab** los proyectivos son los grupos abelianos libres, y toda resolución se puede tomar de longitud 2: cada grupo es cociente de un libre, y todo subgrupo de un libre es libre.

Definición 38.34 (Funtores derivados)**.** Para F aditivo y covariante exacto por la derecha, se toma una resolución proyectiva de M, se le *quita* el propio M, se le aplica F y se toma homología:

$$L_nF(M) = H_n\big(F(P_\bullet)\big).$$

El resultado no depende de la resolución elegida. Con $F = -\otimes N$ se obtiene Tor_n, y dualmente, con $\mathrm{Hom}(-,N)$ y resoluciones proyectivas se obtiene Ext^n.

Ejemplo 38.35 (Los dos valores del capítulo, derivados de verdad)**.** Tómese la resolución proyectiva más corta posible de $\mathbb{Z}/2$ en **Ab**:

$$0 \longrightarrow \mathbb{Z} \xrightarrow{\cdot 2} \mathbb{Z} \longrightarrow \mathbb{Z}/2 \longrightarrow 0.$$

(a) $\operatorname{Ext}^*(\mathbb{Z}/2, \mathbb{Z})$**.** Se quita el $\mathbb{Z}/2$ y se aplica $\operatorname{Hom}(-, \mathbb{Z})$, lo que invierte las flechas:

$$0 \to \operatorname{Hom}(\mathbb{Z}, \mathbb{Z}) \xrightarrow{\cdot 2} \operatorname{Hom}(\mathbb{Z}, \mathbb{Z}) \to 0, \qquad \text{es decir} \qquad 0 \to \mathbb{Z} \xrightarrow{\cdot 2} \mathbb{Z} \to 0.$$

La homología en grado 0 es $\ker(\cdot 2) = 0$, luego $\operatorname{Hom}(\mathbb{Z}/2, \mathbb{Z}) = 0$ (correcto: no hay morfismos no nulos de un grupo finito a $\mathbb{Z}$). Y en grado 1 es $\mathbb{Z}/2\mathbb{Z}$:

$$\operatorname{Ext}^1(\mathbb{Z}/2, \mathbb{Z}) = \mathbb{Z}/2.$$

(b) $\operatorname{Tor}_*(\mathbb{Z}/2, \mathbb{Z}/2)$**.** Ahora se aplica $- \otimes \mathbb{Z}/2$, que respeta el sentido de las flechas:

$$0 \to \mathbb{Z} \otimes \mathbb{Z}/2 \xrightarrow{\cdot 2} \mathbb{Z} \otimes \mathbb{Z}/2 \to 0, \qquad \text{es decir} \qquad 0 \to \mathbb{Z}/2 \xrightarrow{0} \mathbb{Z}/2 \to 0,$$

porque multiplicar por 2 es la aplicación nula en $\mathbb{Z}/2$. La homología en grado 0 es $\mathbb{Z}/2$ —que es $\mathbb{Z}/2 \otimes \mathbb{Z}/2$— y en grado 1 es $\ker 0 = \mathbb{Z}/2$:

$$\operatorname{Tor}_1(\mathbb{Z}/2, \mathbb{Z}/2) = \mathbb{Z}/2.$$

Los dos números que antes aparecían por decreto son ahora dos núcleos de la aplicación "multiplicar por dos".

Observación 38.36 (Qué miden). $\operatorname{Ext}^1(A, B)$ clasifica las extensiones $0 \to B \to E \to A \to 0$ salvo equivalencia, con el elemento nulo correspondiendo a la suma directa. Que $\operatorname{Ext}^1(\mathbb{Z}/2, \mathbb{Z}) = \mathbb{Z}/2$ dice que hay exactamente dos maneras de meter $\mathbb{Z}$ dentro de algo con cociente $\mathbb{Z}/2$: la trivial $\mathbb{Z} \oplus \mathbb{Z}/2$ y la torcida $\mathbb{Z} \xrightarrow{2} \mathbb{Z}$. Y Tor mide, como su nombre indica, la torsión que el producto tensorial destruye: es idénticamente nulo si uno de los dos factores es libre de torsión.

Teorema 38.37 (Lema de la serpiente)**.** *Dado un diagrama conmutativo de filas exactas*

$$\begin{array}{ccccccccc} 0 & \to & A & \to & B & \to & C & \to & 0 \\ & & \downarrow f & & \downarrow g & & \downarrow h & & \\ 0 & \to & A' & \to & B' & \to & C' & \to & 0 \end{array}$$

existe un morfismo de conexión $\delta : \ker h \to \operatorname{coker} f$ *que hace exacta la sucesión*

$$0 \to \ker f \to \ker g \to \ker h \xrightarrow{\delta} \operatorname{coker} f \to \operatorname{coker} g \to \operatorname{coker} h \to 0.$$

Esa serpiente que baja de la fila de los núcleos a la de los conúcleos es el mecanismo que genera *todas* las sucesiones exactas largas: la de Mayer–Vietoris, la del par en homología, la de los funtores derivados. Y el lema de los cinco —si cuatro de los cinco morfismos verticales de un diagrama son isomorfismos, el quinto también— es su complemento operativo.

Cuidado

La demostración de ambos lemas se hace "persiguiendo elementos": se toma $c \in \ker h$, se levanta a B, se aplica g, se comprueba que cae en A', y se define $\delta(c)$ como su clase. Pero en una categoría abeliana abstracta *no hay elementos*. La licencia para razonar así la da el teorema de inmersión de Freyd–Mitchell: toda categoría abeliana pequeña se sumerge de manera plena, fiel y exacta en una categoría de módulos. Es decir, si un enunciado sobre diagramas exactos vale para módulos, vale en general. Es uno de los pocos casos en que un teorema profundo sirve para justificar que se puede seguir siendo ingenuo.

38.9. Categorías de funtores, la incrustación de Yoneda y las álgebras de una mónada

Definición 38.38 (Categoría de funtores y prehaces)**.** Dadas $\mathcal{C}$ y $\mathcal{D}$, la categoría $[\mathcal{C}, \mathcal{D}]$ tiene por objetos los funtores y por morfismos las transformaciones naturales. Un **prehaz** sobre $\mathcal{C}$ es un objeto de $\widehat{\mathcal{C}} = [\mathcal{C}^{\mathrm{op}}, \mathbf{Set}]$.

Esto es lo que faltaba para que el lema de Yoneda tuviera dónde vivir: $\mathrm{Nat}(\mathrm{Hom}(A, -), F)$ es un conjunto de morfismos en una categoría de funtores.

Teorema 38.39 (Incrustación de Yoneda)**.** *El funtor $y : \mathcal{C} \to \widehat{\mathcal{C}}$, $A \mapsto \mathrm{Hom}(-, A)$, es pleno y fiel. Además $\widehat{\mathcal{C}}$ tiene todos los límites y colímites pequeños, aunque $\mathcal{C}$ no tenga ninguno, y todo prehaz es colímite de representables.*

Dos consecuencias con nombre. La primera es que un objeto queda determinado, salvo isomorfismo, por el conjunto de flechas que *entran* en él: no hace falta mirar dentro, basta con saber cómo lo ven los demás. La segunda es que $\widehat{\mathcal{C}}$ es la *compleción por colímites* de $\mathcal{C}$: el modo canónico de añadir a una categoría todo lo que le falta sin romper nada.

Ejemplo 38.40 (Los grafos como prehaces)**.** Sea $\mathcal{C}$ la categoría con dos objetos V y E y dos flechas paralelas $s, t : V \to E$. Un prehaz F asigna un conjunto a cada objeto y una función a cada flecha *en sentido contrario*: dos conjuntos $F(E)$ y $F(V)$ y dos funciones $F(E) \to F(V)$.

Eso es exactamente un grafo dirigido: aristas, vértices, y las aplicaciones origen y destino. La categoría $\widehat{\mathcal{C}}$ es la de los grafos dirigidos, y los morfismos son los homomorfismos de grafos.

Los dos representables son: $y(V)$, el grafo de un solo vértice sin aristas; y $y(E)$, el grafo de una arista con sus dos extremos. Yoneda dice entonces que los homomorfismos desde $y(E)$ a un grafo G son exactamente las aristas de G, y desde $y(V)$, sus vértices. Y "todo prehaz es colímite de representables" se convierte en la afirmación evidente de que todo grafo se construye pegando vértices y aristas.

Qué es una mónada por dentro

Definición 38.41 (Kleisli y Eilenberg–Moore)**.** Sea (T, η, μ) una mónada sobre $\mathcal{C}$.

- La **categoría de Kleisli** $\mathcal{C}_T$ tiene los mismos objetos, y los morfismos $X \to Y$ son las flechas $X \to TY$ de $\mathcal{C}$. La composición es

$$g \bullet f = \mu_Z \circ Tg \circ f,$$

  y la identidad es η_X.

- Un **álgebra de Eilenberg–Moore** es un par (X, h) con $h : TX \to X$ que cumple $h \circ \eta_X = \mathrm{id}$ y $h \circ \mu_X = h \circ Th$. Forman la categoría $\mathcal{C}^T$.

Teorema 38.42. *Ambas construcciones vienen equipadas con adjunciones que reproducen T, y son respectivamente la* inicial *y la* terminal *entre todas las adjunciones que lo hacen. Eso demuestra la afirmación de que toda mónada proviene de al menos una adjunción: proviene de dos canónicas.*

Ejemplo 38.43 (Tres mónadas, leídas por sus álgebras)**. (a) La lista.** $T(X) = X^*$, las sucesiones finitas de elementos de X; η mete un elemento en una lista de longitud uno y μ aplana una lista de listas. Un álgebra es una función $h : X^* \to X$ compatible con esas dos operaciones, es decir: una manera asociativa de combinar cualquier cantidad de elementos, con un neutro. Un *monoide*. Las álgebras de la mónada de la lista son exactamente los monoides.

(b) El grupo libre. Sus álgebras son los grupos. Y en general, para la mónada asociada al funtor libre de cualquier estructura algebraica, las álgebras son esa estructura. Una mónada *es* una presentación de una teoría algebraica, y el teorema de monadicidad de Beck da el criterio exacto para reconocer cuándo un funtor olvidadizo es de ese tipo.

(c) La mónada del error. $T(X) = X \sqcup \{*\}$, donde $*$ representa "ha fallado". Un morfismo de Kleisli $X \to Y$ es una función que devuelve un Y o un fallo, y la composición $g \bullet f$ propaga el fallo automáticamente: si $f(x) = *$ entonces $(g \bullet f)(x) = *$ sin llegar a evaluar g. Eso es exactamente el operador de encadenamiento seguro de los lenguajes de programación, y explica por qué la palabra "mónada" salió de la teoría de categorías y acabó en los manuales de Haskell.

Y una más para cerrar el círculo: la mónada de las partes, $T(X) = \mathcal{P}(X)$, tiene por morfismos de Kleisli las funciones $X \to \mathcal{P}(Y)$, es decir, las *relaciones*; su composición de Kleisli es la composición de relaciones de toda la vida.

38.10. Tamaño, diagramas y conos: los cimientos que faltaban

La definición de límite del capítulo habla de "conos sobre un diagrama" sin que ninguna de las dos palabras esté definida: es, tal cual, circular. Y el lema de Yoneda exige una hipótesis de tamaño que tampoco se explicó.

Definición 38.44 (Categoría pequeña y localmente pequeña)**.** Una categoría $\mathcal{C}$ es *pequeña* si tanto su colección de objetos como la de todos sus morfismos son *conjuntos*. Es *localmente pequeña* si, para cada par de objetos A, B, la colección $\mathrm{Hom}(A, B)$ es un conjunto, aunque los objetos formen una clase propia.

Observación 38.45 (Por qué hay que distinguir, y dónde muerde). **Set**, **Grp** y **Top** *no* son pequeñas: no existe el conjunto de todos los conjuntos, por la paradoja de Russell del capítulo 1. Pero sí son localmente pequeñas: entre dos conjuntos fijos, las aplicaciones sí forman un conjunto. En cambio, un grupo visto como categoría de un objeto, o un conjunto preordenado, sí son pequeños.

La distinción no es pedantería. El lema de Yoneda afirma una *biyección*

$$\mathrm{Nat}\big(\mathrm{Hom}(A,-),\,F\big) \;\cong\; F(A),$$

y para que el miembro izquierdo sea un conjunto —y la afirmación tenga sentido— hace falta que $\mathcal{C}$ sea localmente pequeña. Del mismo modo, "$\mathcal{C}$ tiene todos los límites pequeños" significa límites sobre diagramas indexados por categorías *pequeñas*: pedirlos sobre categorías cualesquiera obliga, por un teorema de Freyd, a que la categoría sea un preorden.

Definición 38.46 (Diagrama, cono y cocono)**.** Sea J una categoría pequeña, llamada *categoría índice*. Un *diagrama* de forma J en $\mathcal{C}$ es sencillamente un funtor $D : J \to \mathcal{C}$. Se escriben D_j para los objetos y $D_u : D_j \to D_k$ para la imagen de un morfismo $u : j \to k$ de J.

Un *cono* sobre D con vértice N es un objeto N de $\mathcal{C}$ junto con una familia de morfismos $\pi_j : N \to D_j$, uno por cada objeto de J, tales que para todo $u : j \to k$ en J se cumple

$$D_u \circ \pi_j = \pi_k$$

(el triángulo conmuta). Un *cocono* es lo mismo con las flechas al revés: $\iota_j : D_j \to N$ con $\iota_k \circ D_u = \iota_j$.

Observación 38.47 (Y ahora sí: qué es un límite). Con esto, el *límite* de D es el cono *universal*: un cono (L, π_j) tal que cualquier otro cono (N, φ_j) se factoriza de manera *única* a través de él, es decir, existe un único $u : N \to L$ con $\pi_j \circ u = \varphi_j$ para todo j. El colímite es el cocono universal.

Todo lo demás son casos particulares de esta única definición, y conviene tener la tabla a mano:

J	límite	colímite
dos objetos, sin flechas	producto $A \times B$	coproducto $A \sqcup B$
vacía	objeto final	objeto inicial
dos objetos, dos flechas paralelas	igualador	coigualador
tres objetos, $\bullet \to \bullet \leftarrow \bullet$	producto fibrado	—
orden dirigido	límite inverso	límite directo

El producto de dos conjuntos, la intersección de subgrupos, el núcleo de un homomorfismo, los enteros p-ádicos y el grupo fundamental como límite directo son *la misma construcción* cambiando J. Ésa es la razón de ser del lenguaje.

Regla: Comprobar que dos funtores son adjuntos

Sean $F : \mathcal{C} \to \mathcal{D}$ y $U : \mathcal{D} \to \mathcal{C}$. Que F sea *adjunto por la izquierda* de U —se escribe $F \dashv U$— se demuestra por cualquiera de estos dos caminos, y conviene saber los dos porque en cada problema uno es mucho más corto que el otro.

Vía la biyección.

1. Constrúyase, para cada par de objetos $X \in \mathcal{C}$, $Y \in \mathcal{D}$, una aplicación
$$\Phi_{X,Y} : \mathrm{Hom}_{\mathcal{D}}(FX, Y) \longrightarrow \mathrm{Hom}_{\mathcal{C}}(X, UY),$$
y su inversa. Suele salir de una propiedad universal ya conocida.
2. Compruébese que Φ es *natural en* Y: para todo $g : Y \to Y'$ se cumple $\Phi(g \circ f) = U(g) \circ \Phi(f)$.
3. Compruébese que es *natural en* X: para todo $h : X' \to X$ se cumple $\Phi(f \circ Fh) = \Phi(f) \circ h$.

Vía unidad y counidad.

1. Dese una transformación natural $\eta : \mathrm{Id}_{\mathcal{C}} \Rightarrow UF$ (la *unidad*) y otra $\varepsilon : FU \Rightarrow \mathrm{Id}_{\mathcal{D}}$ (la *counidad*).
2. Verifíquense las dos *identidades triangulares*:
$$\varepsilon_{FX} \circ F(\eta_X) = \mathrm{id}_{FX}, \qquad U(\varepsilon_Y) \circ \eta_{UY} = \mathrm{id}_{UY}.$$

Los dos caminos son equivalentes: dada la biyección, $\eta_X = \Phi(\mathrm{id}_{FX})$ y $\varepsilon_Y = \Phi^{-1}(\mathrm{id}_{UY})$.

Ejemplo 38.48 (Libre $\dashv$ olvido, con los dos métodos)**.** Sea $F : \mathbf{Set} \to \mathbf{Grp}$ el funtor grupo libre y U el que olvida la estructura. *Con la biyección*: un homomorfismo $F(S) \to G$ queda determinado por dónde van los generadores, y esa asignación es una aplicación de conjuntos $S \to U(G)$ completamente arbitraria; la correspondencia es biyectiva y la naturalidad se comprueba en una línea en cada variable. *Con unidad y counidad*: $\eta_S : S \to UF(S)$ es la inclusión de S como generadores, y $\varepsilon_G : F(U(G)) \to G$ es el homomorfismo que multiplica las palabras. La primera identidad triangular dice que evaluar una palabra de generadores da esa palabra; la segunda, que meter los elementos de G como generadores libres y luego multiplicarlos devuelve el elemento original.

La moraleja que conviene fijar: *los funtores libres son adjuntos por la izquierda de los funtores de olvido*, y eso vale para grupos, anillos, módulos, espacios vectoriales, espacios topológicos y álgebras tensoriales. Cuando en el capítulo 25 se dice que "el objeto libre es el más general posible", se está enunciando una adjunción.

38.11. Ejercicios resueltos

Tres problemas para practicar el reflejo de pensar en flechas y no en elementos.

Ejercicio 38.1. Demuestre que el objeto inicial de una categoría, si existe, es único salvo isomorfismo único.

Solución. Sean I e I' iniciales. Por inicialidad de I existe una única $f : I \to I'$, y por inicialidad de I' una única $g : I' \to I$. Entonces $g \circ f : I \to I$ es un morfismo de I en I, y como I es inicial sólo hay uno, que es id_I; luego $g \circ f = \mathrm{id}_I$. Simétricamente $f \circ g = \mathrm{id}_{I'}$. Por tanto f es isomorfismo, y es el único posible.

El argumento es el mismo para todos los objetos definidos por propiedad universal —productos, coproductos, límites— y ésa es precisamente su virtud: se demuestra una vez y sirve siempre. En **Set** el inicial es $\varnothing$; en **Grp**, el grupo trivial; en **Ring**, el anillo $\mathbb{Z}$. ∎

Ejercicio 38.2. Demuestre que el funtor olvido $U : \mathbf{Grp} \to \mathbf{Set}$ tiene adjunto por la izquierda, y diga cuál es.

Solución. El adjunto por la izquierda es el funtor grupo libre $F : \mathbf{Set} \to \mathbf{Grp}$. La adjunción se expresa como una biyección natural

$$\mathrm{Hom}_{\mathbf{Grp}}(F(X), G) \;\cong\; \mathrm{Hom}_{\mathbf{Set}}(X, U(G)),$$

que en cristiano dice: dar un homomorfismo desde el grupo libre sobre X es lo mismo que decir adónde va cada generador, sin ninguna restricción. Eso es exactamente la propiedad universal del grupo libre, y aquí aparece como un caso de algo mucho más general.

El patrón se repite por todas partes: el anillo de polinomios es libre sobre los anillos, el espacio vectorial libre sobre un conjunto es el de combinaciones formales, y la compactificación de Stone–Čech es adjunta a un olvido. Casi toda construcción que se llame “libre” o “universal” es un adjunto por la izquierda. ∎

Ejercicio 38.3. Verifique que $\mathbf{Set}^{\mathrm{op}}$ no es equivalente a **Set**, exhibiendo una propiedad que se cumpla en una y no en la otra.

Solución. En **Set**, todo epimorfismo se escinde —eso es el axioma de elección— pero no todo monomorfismo lo hace, porque $\varnothing \hookrightarrow X$ no tiene retracción cuando $X \neq \varnothing$. En la opuesta ocurriría lo contrario: los monomorfismos se escindirían y los epimorfismos no.

Como la escisión de epimorfismos es una propiedad categórica, invariante por equivalencia, las dos categorías no pueden ser equivalentes. Nótese que el argumento no menciona elementos en ningún momento: se apoya sólo en flechas, que es lo que exige la noción de equivalencia. ∎

38.12. Ejercicios propuestos

38.1. Dé el objeto inicial y el terminal de **Set**, **Grp** y **Ring**.

38.2. ¿Qué es el producto en **Top**? ¿Y el coproducto?

38.3. ¿Es todo monomorfismo de **Grp** inyectivo? ¿Y todo epimorfismo sobreyectivo?

38.4. Dé un ejemplo de epimorfismo no sobreyectivo en **Ring**.

38.5. ¿Qué es un funtor pleno y fiel? Dé un ejemplo de cada cosa por separado.

38.6. Enuncie el lema de Yoneda y explique en una frase qué significa.

38.7. ¿Es representable el funtor $X \mapsto \mathcal{P}(X)$? Distinga el caso covariante del contravariante.

38.8. ¿Cuál es el adjunto por la izquierda del funtor olvido de espacios vectoriales a conjuntos?

38.9. ¿Preservan límites los adjuntos por la derecha? ¿Y colímites?

38.10. ¿Qué es una transformación natural? Dé un ejemplo con $V \to V^{**}$.

38.11. ¿Por qué $V \cong V^*$ no es natural en dimensión finita y $V \cong V^{**}$ sí?

38.12. Explique en qué sentido una propiedad universal define un objeto "salvo isomorfismo único" y por qué eso basta.

Epílogo: el mismo número, visto seis veces

Este libro empezó con la división con resto y termina con el lema de Yoneda. Entre una cosa y otra hay mil páginas, y es fácil que por el camino se pierda de vista qué las une. La respuesta corta es que no las une un tema: las une un movimiento. El mismo objeto, mirado desde alturas sucesivas, deja de ser lo que parecía y se convierte en otra cosa más general que lo contiene. Vale la pena hacer ese recorrido una última vez, y para que se vea bien conviene fijar un único objeto y no soltarlo. Tomemos el más humilde de todos: el conjunto de los números enteros.

Primera mirada: los enteros son cosas que se cuentan

En el capítulo 3, $\mathbb{Z}$ es una lista. Hay un cero, se puede sumar uno, se puede restar uno, y con eso se llega a todas partes. Los enteros sirven para contar ovejas y para deber dinero. Lo interesante que se hace con ellos es dividir: dados a y b, existen q y r únicos con $a = bq + r$ y $0 \leq r < |b|$. De esa única frase sale todo lo demás. Sale el algoritmo de Euclides, sale que el máximo común divisor es una combinación lineal, sale que si un primo divide a un producto divide a uno de los factores, sale la factorización única, salen las congruencias, sale el teorema chino del resto, sale el pequeño teorema de Fermat y sale RSA. Un capítulo entero, el más largo del libro, colgando de la división con resto.

En esta primera mirada, los enteros son *lo que son*. Son objetos concretos con propiedades que se descubren. El trabajo del matemático consiste en averiguar cosas sobre ellos.

Segunda mirada: los enteros son un ejemplo de algo

En el capítulo 24 ocurre el primer desplazamiento, y es el más difícil de todo el libro. Alguien se fija en que la división con resto no es una peculiaridad de $\mathbb{Z}$. Los polinomios con coeficientes en un cuerpo también la tienen: dado P y $Q \neq 0$ existen C y R únicos con $P = QC + R$ y $\deg R < \deg Q$. Y en cuanto uno tiene división con resto, tiene Euclides; y en cuanto tiene Euclides, tiene identidad de Bézout; y con ella, factorización única. Toda la maquinaria del capítulo 3 se repite palabra por palabra en un contexto donde no hay ningún número.

La conclusión es incómoda y liberadora: los teoremas no eran sobre los enteros. Eran sobre *cualquier estructura con división con resto*, y $\mathbb{Z}$ resultó ser el primer ejemplo que alguien

examinó. Nace la noción de dominio euclídeo, luego la de dominio de ideales principales, luego la de dominio de factorización única, y $\mathbb{Z}$ pasa de protagonista a caso particular. La pregunta ya no es qué propiedades tiene $\mathbb{Z}$, sino de qué familia es miembro y qué comparte con sus hermanos. El objeto no ha cambiado; ha cambiado quién pregunta.

Tercera mirada: los enteros son un grupo, y el grupo es lo que actúa

En el capítulo 23 se olvida la multiplicación y queda sólo la suma. Entonces $\mathbb{Z}$ es el grupo cíclico infinito, y resulta ser el más importante de todos por una razón que no tiene nada que ver con contar: es el grupo libre de un generador. Cualquier grupo generado por un elemento es un cociente suyo. $\mathbb{Z}/n\mathbb{Z}$ ya no es "las horas de un reloj de n casillas" sino "lo que queda de $\mathbb{Z}$ cuando se decide que n vale cero".

Y aquí aparece la idea que reorganiza el resto del libro: un grupo no es tanto un conjunto con operación como una *colección de simetrías*. Los enteros son las simetrías de la recta discreta, las traslaciones de una fila infinita de puntos. Por eso $\mathbb{Z}$ reaparece en el capítulo 37 como grupo fundamental de la circunferencia: dar un lazo en un círculo es dar un entero, el número de vueltas, y componer lazos es sumar vueltas. El mismo grupo que servía para contar ovejas cuenta ahora agujeros, y no por analogía, sino porque es literalmente el mismo grupo.

Cuarta mirada: los enteros como espacio

En el capítulo 34 los enteros dejan de ser discretos. La pregunta "¿cuántos primos hay hasta x?" no tiene respuesta aritmética, y sí tiene una respuesta analítica: la función zeta, definida por una serie que converge en el semiplano derecho, se extiende a todo el plano complejo, y sus ceros gobiernan el error del recuento de primos. Un objeto que se construyó contando resulta que se entiende mejor con integrales de contorno.

Ese salto —de lo discreto a lo continuo y de vuelta— es la firma de las matemáticas del siglo XX. En el capítulo 33 la misma idea aparece disfrazada: descomponer una función en frecuencias es exactamente descomponer una representación en irreducibles, y el análisis de Fourier clásico es la teoría de representaciones del grupo $\mathbb{Z}$ y de la circunferencia. Los coeficientes de Fourier son los caracteres. Nadie lo diría mirando una onda en un osciloscopio.

Quinta mirada: los enteros son un objeto geométrico

En el capítulo 35 se lleva la idea al extremo. Si a un anillo se le puede asociar el espacio de sus ideales primos, entonces $\mathbb{Z}$ tiene un espacio asociado: un punto por cada primo, más un punto genérico. $\operatorname{Spec}\mathbb{Z}$ es una curva, y hacer teoría de números es hacer geometría sobre ella. Los enteros, que empezaron siendo una lista, son ahora un espacio con topología, con funciones definidas encima y con una noción de proximidad en la que dos números están cerca si son congruentes módulo una potencia alta de un primo.

Esto no es una metáfora poética. Es la razón técnica por la que el último teorema de Fermat se demostró con curvas elípticas y formas modulares, y no con manipulaciones de

exponentes. La pregunta era aritmética y la respuesta fue geométrica, porque desde cierta altura las dos disciplinas son la misma.

Sexta mirada: los enteros son lo que los enteros hacen

Y llegamos al capítulo 38, donde el objeto se disuelve del todo. $\mathbb{Z}$ es el objeto inicial de la categoría de anillos: desde él sale exactamente un morfismo hacia cualquier anillo, y esa propiedad lo determina por completo, salvo isomorfismo único. No hace falta decir qué son sus elementos. No hace falta la división con resto, ni la suma, ni el orden. Basta con decir cómo se relaciona con todos los demás anillos, y el lema de Yoneda garantiza que con eso ya está dicho todo.

Ahí termina el viaje que el título anunciaba. Empezamos con objetos que se cuentan y acabamos con objetos que se definen por sus relaciones. Y conviene decir con claridad que la última mirada no es mejor que la primera. Nadie calcula el máximo común divisor de 4620 y 1386 invocando propiedades universales; se hace con Euclides, en cuatro líneas, como en el capítulo 3. Cada altura sirve para preguntas distintas, y la habilidad que de verdad distingue a quien sabe matemáticas es saber a qué altura ponerse para cada problema. Subir siempre al máximo nivel de abstracción es tan torpe como no subir nunca.

Qué se ha quedado fuera

Mucho. Este libro tiene mil páginas y sigue siendo una introducción. No hay geometría diferencial más allá de lo básico, ni sistemas dinámicos, ni teoría ergódica, ni teoría de números algebraica propiamente dicha, ni topología diferencial, ni álgebra homológica más allá de una mención, ni teoría de la computación, ni combinatoria extremal, ni análisis numérico serio de ecuaciones en derivadas parciales. De lo que sí está, casi todo se podría haber contado con el doble de profundidad.

Si hay que sugerir por dónde seguir, tres direcciones son especialmente agradecidas. Quien haya disfrutado los capítulos de álgebra encontrará que la teoría de números algebraica —cuerpos de números, anillos de enteros, grupos de clases— es la continuación natural, y que allí la factorización única vuelve a fallar y hay que reconstruirla con ideales, que es exactamente para lo que Dedekind los inventó. Quien haya disfrutado el análisis debería ir hacia las ecuaciones en derivadas parciales modernas, donde los espacios de Sobolev y el análisis funcional dejan de ser formalismo y se vuelven la única herramienta posible. Y quien haya disfrutado la topología y las categorías tiene delante la topología algebraica seria: homología, cohomología, sucesiones exactas, y la maquinaria que convierte problemas geométricos en álgebra calculable.

Una última cosa

Las matemáticas se aprenden haciéndolas. Ningún libro, ni éste ni otro mejor, sustituye el rato de estar atascado con un ejercicio, ni la sensación de que una demostración que se leyó tres veces sin entenderla de pronto se vuelve obvia. Los ejercicios propuestos de estas

páginas están ahí por eso, y el apéndice de soluciones da el resultado y una indicación, nunca el camino completo, precisamente para que el camino tenga que recorrerlo el lector.

Si algo debería quedar de estas mil páginas, no es una lista de teoremas. Es la costumbre de preguntarse, ante cualquier objeto, tres cosas: qué propiedades tiene, de qué familia es un caso particular, y qué queda de él si se olvida todo salvo cómo se relaciona con los demás. Esas tres preguntas son, en el fondo, el índice de este libro.

Apéndice A

Mapa de cierre del temario

Este apéndice recorre el temario completo y certifica qué cubre cada capítulo. Todos los resultados numéricos y fórmulas del libro (más de 150 comprobaciones: integrales, valores propios, discriminantes, polinomios mínimos, símbolos de Christoffel, curvaturas, soluciones de EDO, resolventes, ciclotómicos) fueron verificados con álgebra simbólica computacional durante la preparación del documento, y los datos de estado (problemas abiertos) contrastados con fuentes actuales.

Criterio de cierre

Cada técnica del temario aparece con (1) su definición precisa, (2) el teorema que la justifica, (3) una demostración o idea de demostración, y (4) al menos un ejercicio o ejemplo resuelto que la ejercita. La estructura de cada capítulo: teoría nuclear, *Fundamentos* (demostraciones de fondo), *Complementos*, *Cierre del temario*, *Profundización final*, *Última capa*, *Temas extra*, y dos tandas de ejercicios resueltos.

Contenido por capítulo

Cálculo (cap. 1): completitud de $\mathbb{R}$ con construcciones (Dedekind, Cantor) y cinco equivalencias; sucesiones (monótona, B–W, Cauchy, limsup/liminf); límites y continuidad (ε–δ, Bolzano, Weierstrass, Heine demostrados); derivada (reglas demostradas, TVM, L'Hôpital, Darboux, convexidad, Taylor); integral (Darboux, criterio, TFC, técnicas completas, impropias, Gamma, Dirichlet oscilante); series (todos los criterios, Raabe, condensación, Abel–Dirichlet, reordenación); series de potencias y convergencia uniforme (criterio M, Cauchy–Hadamard, Abel del borde); productos infinitos (seno de Euler, zeta); sumabilidad (Cesàro/Abel); exp construida desde cero; Weierstrass–Bernstein; hiperbólicas; Euler $e^{i\theta}$; Feynman; irracionalidad de e; fracciones continuas; Euler–Maclaurin; aplicaciones completas de derivada (trazado, tasas, Newton) e integral (sólidos, arco, superficies, trabajo); Wallis y Stirling; Riemann sums; trapecio/Simpson; Snell.

Álgebra lineal (cap. 2): sistemas y Gauss (Rouché–Frobenius, LU, elementales, unicidad de la reducida); espacios vectoriales (Steinitz demostrado, Grassmann, cambio de base); transformaciones (rango–nulidad, cuatro subespacios, Fredholm); determinantes (existencia

y unicidad vía Leibniz, Cramer, Cauchy–Binet); espectral (Cayley–Hamilton vía Schur, mínimo, primaria, Jordan demostrado, racional, funciones de matrices); euclídeo (Gram–Schmidt, QR, mínimos cuadrados, espectral demostrado, Courant–Fischer, entrelazado); complejos (hermitianas/unitarias/normales); formas cuadráticas y bilineales (Sylvester, inercia, congruencia); SVD demostrada, polar, pseudoinversa; Cholesky, normas y condición; Perron–Frobenius e iteración de potencias; circulantes y DFT; Markov/PageRank; Vandermonde; Gershgorin; proyecciones; rotaciones 3D; grafos.

Multivariable (cap. 3): topología y Weierstrass; límites (trayectorias, iterados vs. doble); diferenciabilidad (jerarquía, Clairaut demostrado, parciales continuas $\Rightarrow$ dif.); Taylor multivariable (orden 2 demostrado, multi-índices, resto integral); optimización (hessiana, Lagrange demostrado, hessiana orlada, KKT, convexidad); inversa/implícita demostradas (Banach); Fubini y cambio de variables; aplicaciones (centroides, probabilidad conjunta); Morse y rango constante; homogéneas y Euler; envolvente; Reynolds; cuádricas; impropias múltiples; Jensen; d'Alembert por características; derivada material; descenso por gradiente; curvatura implícita; cinemática (a_T, a_N, proyectil).

Vectorial (cap. 4): curvas y Frenet (fórmulas deducidas); operadores e identidades; integrales de línea y conservativos (caracterización, potenciales, índice); superficies y flujo (parametrizaciones estándar, toro, independencia); Green demostrado (ambas formas), Stokes demostrado, Gauss demostrado; deformación de contornos; formas diferenciales ($\mathrm{d}^2 = 0$, Poincaré, H^1 del plano agujereado calculado); Green identidades, valor medio, máximo, Dirichlet/Neumann; curvilíneas generales; Helmholtz; ángulo sólido; Biot–Savart; Kepler completo (leyes 1–3); Poincaré–Hopf; Maxwell (ondas, Poynting); trabajo–energía; calor deducido; Pappus; Kelvin/vorticidad; Arquímedes.

EDO/EDP (cap. 5): primer orden completo (separables, lineales, exactas, Bernoulli, Riccati, homogéneas); Picard demostrado + Gronwall; segundo orden (wronskiano/Abel, característica, resonancia, variación de parámetros, Euler–Cauchy); orden n; sistemas (matriz fundamental, exponencial, retratos, Hartman–Grobman, Lotka–Volterra); Laplace (existencia, inyectividad, convolución demostrada, escalón/delta); series (Hermite, Frobenius/Bessel); Fourier (Dirichlet vía núcleo, Gibbs, Parseval/Basilea, Fejér, complejas/FFT); EDP (calor/onda/Laplace resueltas, clasificación $B^2 - 4AC$, características y Burgers, condiciones D/N/Robin, Legendre y armónicos esféricos, tambor/Bessel, Sturm–Liouville demostrado, energía); Green de contorno; péndulo elíptico; estabilidad numérica y rigidez; hamiltonianos; línea de fases y bifurcaciones (silla–nodo, Hopf); Lyapunov; Poincaré–Bendixson y van der Pol; Floquet y resonancia paramétrica; caos (Lorenz, logístico, Feigenbaum); comparación de Sturm; modos normales; pozo cuántico; transformada de Fourier y núcleo del calor.

Tensorial (cap. 6): convenio y gramática; dual y ley de transformación; tensores y operaciones (producto universal demostrado, cociente); métrica (repertorio, subir/bajar, volumen); Levi-Civita (Koszul demostrado, tensorialidad); geodésicas (transporte, variacional demostrado, Noether/Killing demostrado); curvatura (Riemann, simetrías en normales, Ricci y su significado volumétrico, escalar, Bonnet–Myers, desviación geodésica); esfera completa; holonomía/Foucault; $\mathbb{H}$ completo; cono y Gauss–Bonnet; Egregium y segunda forma (Codazzi, minimales); variedades formales (cartas, derivaciones, pushforward/pullback); conos de luz y horizonte; Christoffel esféricos 3D; isometrías y homogéneos; Lie ($\mathfrak{so}(3)$ = producto vectorial, espín); Cartan (esfera en tres líneas); relatividad (dilatación, $E = mc^2$, electromagnetismo covariante, Schwarzschild y órbitas, ondas gravitacionales/LIGO, Friedmann deducidas, Kerr

y termodinámica); energía–momento; medios continuos.

Álgebra (cap. 7): grupos (permutaciones, cíclicos, Lagrange, normalidad, isomorfía 1–3 demostrados, correspondencia, acciones/órbita–estabilizador, Cayley, ecuación de clases, Cauchy–McKay demostrado, Sylow demostrado con orden pq, abelianos finitos, Jordan–Hölder, solubles/simples, A_5 con cuentas, semidirectos y tabla de pequeños, presentaciones e indecidibilidad, automorfismos); anillos (ideales, cocientes, maximales/primos, CRT demostrado, euclídeo–DIP–DFU con contraejemplos, fracciones, noetheriano/Hilbert); polinomios (división, irreducibilidad con cinco criterios demostrados, la máquina $K[x]/(p)$, simétricos y Newton, resultantes/Sylvester, factorización efectiva); cuerpos (característica, finitos con Frobenius, $\mathbb{F}_q^*$ cíclico demostrado); módulos sobre DIP (abelianos + Jordan unificados); Wilson demostrado; dos cuadrados con $\mathbb{Z}[i]$ demostrado; Jacobi; enteros algebraicos y Dedekind (ideales primos, grupo de clases); representaciones y caracteres (S_3 completa); Burnside; RSA completo.

Galois (cap. 8): extensiones y torre demostrada; algebraicos y mínimos; descomposición y normalidad; separabilidad; primitivo demostrado; grupos de Galois (repertorio completo con ciclotómicos y finitos); correspondencia demostrada (Dedekind, Artin, ensamblaje) con retículos trabajados (x^3-2, $\mathbb{Q}(\sqrt{2},\sqrt{3})$, x^4-2, x^4+1, x^6-2, $\mathbb{F}_{2^{12}}$); fórmulas clásicas (Cardano deducido y ejecutado, cuártica por resolvente con ejemplo); solubilidad demostrada en ambas direcciones (Kummer, ciclotómicas, resolventes de Lagrange explícitas); Abel–Ruffini completo ($x^5 - 4x + 2$) y general (ecuación general con S_n); discriminantes y cómputo de grupos (cúbicas, cuárticas, Dedekind/Chebotarév); norma y traza; regla y compás (tres problemas, Gauss–Wantzel, pentágono construido, $\cos 72^\circ$); TFA vía Galois demostrado; Hilbert 90 con ternas pitagóricas; sumas de Gauss y reciprocidad; clausura algebraica; Liouville demostrado y panorama de trascendencia; ciclotómicos completos (Φ_{105}); base normal; raíces de primos; p-ádicos (Ostrowski, Hensel con ejemplo, local–global); Lüroth; cohomología en cápsula; Krull y grupo absoluto; irreducibilidad de Hilbert; Kronecker–Weber; AES y aplicaciones. *Estado*: la irracionalidad de γ y el problema inverso de Galois siguen abiertos a la fecha.

Variable compleja: holomorfía y Cauchy–Riemann (demostrado, con conformidad y armonicidad), Cauchy–Goursat vía Green, fórmula integral (demostrada) con holomorfa=analítica, Liouville y la tercera demostración del TFA, módulo máximo, Laurent y clasificación de singularidades, teorema de los residuos con integrales reales trabajadas (racionales, Fourier, trigonométricas, y la cuarta demostración de Basilea), mapas conformes (Riemann, Yukovski, Möbius) con un problema de potencial resuelto, y el panorama (elípticas, zeta y la hipótesis de Riemann, formas modulares, Bromwich).

Cómo usar este mapa

Si un concepto del temario no aparece aquí, pertenece a un curso posterior (Lebesgue, variedades avanzadas, análisis complejo como materia propia, representaciones continuas): los señalados en la bibliografía como siguientes pasos.

Ronda de temarios (cotejo curricular). Última pasada contrastando cada capítulo con temarios oficiales (Cálculo II y Álgebra Lineal II de la Facultad de Ciencias UNAM, Ecuaciones Diferenciales I, curso básico de álgebra del Cinvestav, programas estándar de geometría diferencial y variable compleja). Incorporado en esta ronda: las tres formas del

resto de Taylor (integral, Lagrange, Cauchy) con la aplicación al logaritmo, la irracionalidad de π (Niven) y las cotas de error de trapecio y Simpson demostradas (cap. 6); espacio cociente con el primer teorema de isomorfismo, bidual canónico y anuladores con la correspondencia completa (cap. 7); envolventes de familias de curvas: círculo, parábola de seguridad y astroide (cap. 9); derivada material, teorema del transporte de Reynolds con la ecuación de continuidad deducida, y ángulo sólido con la ley de Gauss (cap. 11); ecuación de Clairaut con soluciones singulares como envolventes y d'Alembert–Lagrange (cap. 12); ecuación de desviación geodésica (Jacobi) con mareas y el principio de LIGO (cap. 36); grupos nilpotentes con la caracterización por Sylow, subgrupo de Frattini y $\mathrm{Inn}(G) \cong G/Z(G)$ con el automorfismo exterior de S_6 (cap. 25); teorema de la base normal, extensiones inseparables y cuerpos perfectos, bases de trascendencia y teorema de Steinitz (cap. 26); lema de Schwarz con los automorfismos del disco, núcleo de Poisson y problema de Dirichlet, principio de reflexión y la fórmula de reflexión de la función Γ (cap. 30).

Lebesgue y Fourier. La deuda anunciada desde el capítulo 6 queda saldada: medida exterior y criterio de Carathéodory con la σ-álgebra y la aditividad numerable; el conjunto de Vitali (no medible, vía axioma de elección) y el conjunto de Cantor (no numerable y de medida cero, con su dimensión $\log 2/\log 3$); la integral de Lebesgue con los tres teoremas de convergencia demostrados (monótona, Fatou, dominada) y el catálogo de patologías; el criterio de Lebesgue para la integrabilidad de Riemann; $\int \frac{\sin x}{x}$ como frontera absoluta/condicional; los espacios L^p con Riesz–Fischer y el estatus definitivo de las series de Fourier en L^2 (Parseval, Carleson, el contraejemplo de Kolmogórov); y la transformada de Fourier con demostraciones: gaussiana autotransformada, Riemann–Lebesgue, teorema de inversión vía amortiguamiento gaussiano, Plancherel, principio de incertidumbre de Heisenberg con su caso de igualdad, la ecuación del calor en la recta con su núcleo, el teorema del muestreo, y como síntesis final la fórmula de sumación de Poisson con la función theta.

Cierre definitivo de los capítulos de cálculo y álgebra lineal. Última pasada de completitud sobre la primera mitad del libro: función Beta con $B = \Gamma\Gamma/\Gamma$ demostrada y todas las integrales trigonométricas unificadas, y la fórmula de Stirling demostrada con cotas explícitas válidas para todo n (cap. 6); teorema de diagonalización simultánea con sus tres lecturas: circulantes/DFT, observables compatibles y explotación de simetrías (cap. 7); teorema del valor medio multivariable, su fallo vectorial con la desigualdad que lo sustituye, y el corolario $C^1 \Rightarrow$ Lipschitz local que alimenta a Picard (cap. 9); identidades de Green con unicidad de Dirichlet y Neumann, propiedad del valor medio en 3D y principio del máximo fuerte demostrados (cap. 11); teorema de Peano vía poligonales de Euler y Arzelà–Ascoli, el abanico completo de soluciones de $y' = 3y^{2/3}$, intervalos maximales con la alternativa de escape, y el criterio de existencia global por crecimiento lineal vía Gronwall (cap. 12).

Figuras y cierre definitivo de los capítulos avanzados (tensores, estructuras algebraicas, Galois, variable compleja y Lebesgue). Se añadieron doce figuras (banda de convergencia uniforme, polinomios de Taylor, elipse de la SVD, tangencia de Lagrange, astroide-envolvente, cancelación de Stokes, plano de fase del péndulo, sumas de Fourier con Gibbs, transporte paralelo en la esfera, retículo de Galois de $\mathbb{Q}(\sqrt{2},\sqrt{3})$, contorno semicircular, Riemann vs. Lebesgue) y veinticuatro recuadros de intuición. Cierre curricular de la segunda

mitad: Hopf–Rinow, teorema de Frobenius de integrabilidad y Gauss–Bonnet demostrado local y globalmente (cap. 36); anillos noetherianos con la base de Hilbert, Zorn e ideales maximales, y localización (cap. 25); teoría de Galois infinita con la topología de Krull y $\widehat{\mathbb{Z}}$, y teoría de Kummer (cap. 26); demostración del teorema del mapeo de Riemann vía Montel y el problema extremal, teoremas de Picard, y la función $\wp$ de Weierstrass con su ecuación diferencial y el puente toros = curvas elípticas (cap. 30); Fubini–Tonelli con el contraejemplo condicional, Radon–Nikodym vía von Neumann, el teorema de diferenciación de Lebesgue, y el broche final: la probabilidad como teoría de la medida, con la ley de los grandes números demostrada y el teorema del límite central completo vía funciones características (cap. 31).

Capítulo 3 (Teoría de números), escrito en clave pedagógica. A diferencia del resto del libro, este capítulo se redactó despacio y desde cero, como modelo de una exposición pensada para *aprender* y no solo para *consultar*: cada resultado nace de una pregunta o un experimento concreto, las demostraciones van completas y sin atajos, hay recuadros de "Pausa para pensar" y todos los ejemplos numéricos están desarrollados paso a paso. Contenido: divisibilidad, algoritmo de la división y de Euclides, identidad de Bézout; primos, infinitud (Euclides), lema de Euclides y factorización única; aritmética modular, inversos y ecuaciones lineales de congruencia; teorema chino del resto (con el acertijo de Sun Zi resuelto a mano); pequeño teorema de Fermat y teorema de Euler, ambos demostrados con el método de barajar residuos; el criptosistema RSA ejecutado íntegro con números pequeños (generación de llaves, cifrado y descifrado, y la demostración de por qué funciona vía Euler); y como cierre, el teorema de Fermat sobre sumas de dos cuadrados con la mitad elemental demostrada y el puente hacia los enteros de Gauss. Cinco ejercicios resueltos con todos los pasos.

Capítulo las reglas de derivación e integración (Manual de cálculo paso a paso), estilo libro de colegio. El capítulo más elemental del libro, pensado para leerse *antes* que el capítulo 6 si el repaso resultó veloz. Cada regla del cálculo va enunciada en su propia caja ("Regla: ...") y seguida de varios ejemplos resueltos con todos los pasos numerados. Derivadas: definición por límite (una vez), reglas de la constante, la potencia (con exponentes enteros, negativos y fraccionarios), múltiplo, suma, tabla de funciones elementales, producto, cociente, cadena (con abundantes ejemplos), combinaciones, derivación implícita y órdenes superiores, más una tabla resumen. Integrales: concepto de primitiva y el $+C$, regla de la potencia al revés, tabla de integrales inmediatas, y las tres técnicas escolares —sustitución, integración por partes (con la regla LIATE) y fracciones parciales— cada una con su receta en pasos y ejemplos desarrollados; cierra con la integral definida y el teorema fundamental. Siete ejercicios resueltos con cada paso escrito. Todas las derivadas e integrales verificadas simbólicamente.

Capítulos escolares de álgebra, trigonometría y límites, estilo libro de colegio. Tres capítulos elementales que forman, junto con el manual de cálculo de las reglas de derivación e integración, un curso completo desde cero. La *ruta de estudio* recomendada para un principiante es leerlos en este orden antes que los capítulos avanzados: **(13) Álgebra** —leyes de exponentes, productos notables, factorización por todos los métodos, fracciones algebraicas, ecuaciones lineales y cuadráticas (factorización y fórmula general con discriminante), y sistemas por sustitución y eliminación—; **(14) Trigonometría** —grados y radianes, SOH-CAH-TOA,

círculo unitario y tabla de valores notables, identidad pitagórica y otras, ley de senos y de cosenos, y ecuaciones trigonométricas—; **(15) Límites** —idea intuitiva, sustitución directa, la forma $0/0$ resuelta factorizando y racionalizando, límites al infinito por grados, continuidad, y la regla de L'Hôpital con el límite fundamental $\sin x/x$—. Cada regla en su caja, cada ejemplo con pasos numerados, cajas de advertencia para los errores típicos, y ejercicios resueltos. Todos los cálculos verificados simbólicamente. Con estos cuatro capítulos escolares (álgebra, trigonometría, límites y el manual de cálculo) el lector llega preparado al cálculo formal.

Talleres de ejemplos explicados (capítulos avanzados). Para acercar los capítulos avanzados al mismo estilo pedagógico de los capítulos escolares, cada uno de los capítulos avanzados termina ahora con un "Taller de ejemplos explicados" con ejercicios resueltos paso a paso y explicaciones amplias: convergencia de series con los criterios geométrico, p y del cociente (cap. 6); eliminación de Gauss y cálculo de valores y vectores propios (cap. 7); derivadas parciales, gradiente y un problema completo de multiplicadores de Lagrange (cap. 9); una integral de línea y el teorema de Green sobre un cuadrado (cap. 11); resolución de ecuaciones diferenciales separables, lineales de primer orden y de segundo orden con coeficientes constantes (cap. 12); longitud sobre la esfera y un símbolo de Christoffel calculados con la métrica (cap. 36); orden de un elemento, clases laterales y Lagrange en $\mathbb{Z}_6$ (cap. 25); grupos de Galois de x^2-2 y de x^3-2 (cap. 26); aritmética compleja y una integral real por residuos (cap. 30); un límite de integrales por convergencia dominada y un conjunto de medida cero (cap. 31). Todos los cálculos verificados simbólicamente.

Capítulo la geometría analítica (Geometría analítica, escolar) y capítulo 28 (Topología, avanzado). Dos capítulos nuevos. **Geometría analítica** sigue el formato de libro de colegio (reglas en cajas, ejemplos con pasos, ejercicios resueltos): plano cartesiano, distancia y punto medio; la recta con sus formas (punto-pendiente, pendiente-ordenada, general) y las condiciones de paralelismo y perpendicularidad; y las cuatro cónicas —circunferencia (con centro por completar cuadrados), parábola (vértice), elipse (semiejes y focos) e hipérbola (asíntotas)—. **Topología** sigue el formato de los capítulos avanzados (definiciones, teoremas y demostraciones, más cajas de intuición y un taller): espacios métricos y bolas abiertas, conjuntos abiertos y cerrados, espacios topológicos abstractos, la definición topológica de continuidad (equivalente a la de ε–δ), compacidad con cubrimientos y Heine–Borel, conexión y el teorema del valor intermedio, y homeomorfismos con la idea de invariante topológico; el taller demuestra que $(0,1)$ es abierto, que no es compacto (con un cubrimiento explícito), y que $[0,1]$ y $(0,1)$ no son homeomorfos. Todos los cálculos verificados.

Capítulos de Funciones y Logaritmos (escolares). Dos capítulos más en formato de libro de colegio que completan la base antes del cálculo. **Funciones y gráficas**: concepto de función como máquina, dominio y rango con las dos prohibiciones (no dividir entre cero, no raíz par de negativo), los tipos básicos de función, cómo graficar con tabla de valores, las transformaciones de gráficas (traslaciones, reflexiones, estiramientos), y composición e inversa de funciones. **Logaritmos y exponenciales**: funciones exponenciales y el número e, la definición de logaritmo como exponente, las leyes de los logaritmos (producto, cociente, potencia, cambio de base), resolución de ecuaciones exponenciales y logarítmicas (con la comprobación obligatoria de soluciones), y aplicaciones de crecimiento y decaimiento (interés

continuo, vida media). Cada regla en su caja, ejemplos con pasos, cajas de advertencia para los errores clásicos, y ejercicios resueltos. Todos los cálculos verificados.

Ampliación: más ejemplos por tema (capítulos escolares). A cada capítulo escolar (manual de cálculo, álgebra, trigonometría, límites, geometría analítica, funciones y logaritmos) se le añadió una sección "Más ejemplos resueltos, tema por tema" que refuerza cada sub-punto con ejercicios adicionales trabajados paso a paso: más práctica de cada regla de derivación e integración; más exponentes, productos notables, factorización, cuadráticas y sistemas; más conversiones, SOH-CAH-TOA, identidad pitagórica y ley de cosenos; más límites por sustitución, factorización, racionalización, al infinito y por L'Hôpital; más distancia, rectas, circunferencias, parábolas, elipses e hipérbolas; más evaluación, dominio, transformaciones, composición e inversa de funciones; y más logaritmos, leyes, ecuaciones exponenciales y logarítmicas y aplicaciones de crecimiento. Todos los cálculos verificados simbólicamente.

Ampliación del temario por capítulo (bloque avanzado). Se integraron los temas solicitados en los primeros cinco capítulos, cada uno con reglas y ejemplos resueltos: **(cap. 6)** regla de L'Hôpital y las siete formas indeterminadas, desglose de técnicas de integración (por partes con LIATE, fracciones parciales, sustitución trigonométrica), y curvas paramétricas y coordenadas polares (tangente, longitud de arco de la cicloide, área de la cardioide); **(cap. 7)** geometría analítica en $\mathbb{R}^2$ y $\mathbb{R}^3$ (producto cruz, rectas, planos, distancia punto-plano), espacios afines (puntos frente a vectores), programación lineal con un problema completo y el método Simplex, y el producto tensorial con el álgebra multilineal; **(cap. 9)** integrales múltiples en coordenadas cilíndricas y esféricas (volumen de la esfera), y aplicaciones físicas (masa, centro de masa, momento de inercia del disco); **(cap. 11)** introducción al cálculo tensorial (campos tensoriales, jerarquía por rango, tensor de esfuerzos, operadores); **(cap. 12)** la ecuación de Bessel y las funciones cilíndricas, la transformada Z como análogo discreto de Laplace, y los métodos numéricos para EDP (diferencias finitas con el esquema explícito del calor, e introducción al método de elementos finitos). Todos los cálculos verificados simbólica y numéricamente.

Enriquecimiento con ayudas intuitivas (sin perder formalidad). A cada tema de las ampliaciones avanzadas y de los talleres de los capítulos difíciles se le añadió una caja "Idea clave" que explica, en lenguaje llano y antes de la formalidad, de dónde sale y qué significa: por qué L'Hôpital funciona (reemplazar cada función por su tangente), la integración por partes como regla del producto al revés, el triángulo escondido de la sustitución trigonométrica, las curvas paramétricas como una película, la rebanada de pastel de las coordenadas polares, el producto cruz como dirección-y-área, los espacios afines como pizarra de posiciones y flechas, el Simplex como caminar por las aristas, el producto tensorial como reempaquetado de lo bilineal, el jacobiano como factor de estiramiento de la rejilla, el momento de inercia y su distancia al cuadrado, el tensor como máquina de ranuras, la ecuación de Bessel y el sonido del tambor, la transformada Z como Laplace discreto, y los métodos numéricos como promediar vecinos; más las cajas de intuición al frente de los cinco talleres de los capítulos avanzados finales. La formalidad (definiciones, teoremas, demostraciones y ejemplos) queda intacta; solo se añadieron las ayudas.

Teoría de grupos: enriquecimiento del capítulo 25 y capítulo 23 nuevo. Última ronda, dedicada a los grupos por petición expresa. **En el capítulo 25** se añadieron diez cajas "Idea clave" delante de los temas centrales que aún no las tenían, sin tocar una sola línea del formalismo existente: el grupo como sistema de acciones reversibles (y por qué no se pide conmutar), los ciclos como carruseles de fichas, el subgrupo cíclico como reloj y su correspondencia con los divisores, el homomorfismo como traducción que respeta la gramática (con el núcleo como "lo que la traducción olvida"), los grupos simples como átomos y la solubilidad como desmontaje en capas abelianas, Sylow como el esqueleto primo que Lagrange no garantiza, la nilpotencia como abelianidad que se propaga hacia arriba, las representaciones y el carácter como huella dactilar del grupo, los automorfismos interiores como cambios de punto de vista, y el anillo como medida de "cuánto se parece esto a los enteros".

El capítulo 23 es un curso completo de teoría de grupos escrito desde cero, en el mismo estilo pedagógico de los capítulos escolares pero llegando hasta contenido de licenciatura. Empieza con un triángulo de cartón sobre una mesa y no enuncia ningún axioma hasta haber contado a mano las seis simetrías y calculado productos concretos; a partir de ahí: la tabla de Cayley completa de D_3 (verificada entrada por entrada por ordenador) con sus tres teoremas escondidos, los cuatro axiomas en su caja con las consecuencias demostradas paso a paso, un catálogo de ejemplos y de *no*-ejemplos con el motivo exacto del fallo en cada caso, el orden de un elemento con la fórmula $n/\gcd(a,n)$ aplicada a los doce elementos de $\mathbb{Z}_{12}$, el criterio de subgrupo y la búsqueda sistemática de todos los subgrupos de D_3 y de D_4 (los diez, uno por uno), la relación diédrica $srs = r^{-1}$ como herramienta de cálculo, la notación de ciclos con multiplicaciones desarrolladas y el inventario completo de S_4 por tipo cíclico, el teorema de Lagrange con todas las clases laterales escritas explícitamente (incluido el caso no normal, donde izquierda y derecha difieren), los cocientes con una demostración numérica de por qué la operación *no* está bien definida si el subgrupo no es normal, $\mathbb{Z}/3\mathbb{Z}$ y $D_4/Z(D_4) \cong V$ calculados a mano, los criterios prácticos para probar o refutar un isomorfismo, órbita–estabilizador contando las 24 rotaciones del cubo por tres caminos distintos, el lema de Burnside con los 14 collares y los 57 cubos tricolores, la tabla de cuántos grupos hay de cada orden hasta 12 con los criterios que la explican y la clasificación completa de los de orden 6 demostrada, los productos directos, un recorrido por dónde aparecen los grupos de verdad (Rubik, los 230 grupos espaciales, química cuántica, códigos correctores, criptografía, modelo estándar) y un taller de diez ejercicios resueltos con cada paso escrito. Todas las cuentas del capítulo —la tabla de D_3 completa, los recuentos de subgrupos, los tipos cíclicos de S_4, los 20 elementos de orden 6 de S_5, los conteos de Burnside y el orden del grupo del cubo de Rubik— fueron verificadas con álgebra computacional.

Reordenamiento del libro en cinco partes. Hasta esta versión los capítulos aparecían en el orden en que se fueron escribiendo, no en el orden en que se deben leer: la teoría de Galois caía antes que el álgebra elemental, y un lector que empezara por la página uno se encontraba con demostraciones de teoría de cuerpos mucho antes que con la resolución de una ecuación de segundo grado. Esta versión corrige la estructura sin tocar ni una línea de contenido matemático. El libro se divide ahora en cinco partes ordenadas por dependencia lógica: **Parte I, Fundamentos desde cero** (álgebra, funciones, logaritmos, trigonometría, geometría analítica, límites y el manual de cálculo); **Parte II, Cálculo y álgebra lineal**; **Parte III, Análisis** (topología, variable compleja, medida de Lebesgue y Fourier); **Parte**

IV, Álgebra abstracta y teoría de números (números, grupos desde cero, estructuras algebraicas, Galois); y **Parte V, Geometría diferencial y tensores**. Como consecuencia técnica, las casi trescientas referencias cruzadas del texto ("véase el capítulo tal") dejaron de ser números escritos a mano y pasaron a ser referencias simbólicas resueltas por el propio compilador: cualquier reordenamiento futuro las actualiza solo. Los dos capítulos que antes se llamaban "Epílogo" y "Epílogo II" recuperaron su título propio, porque ya no están al final. La guía de uso del principio se reescribió para describir la ruta de estudio real, parte por parte, y para decir con claridad dónde debe empezar cada lector según lo que ya sepa.

Pase de intuición capítulo a capítulo, en el orden real de lectura. El enriquecimiento anterior se había hecho por temas sueltos, no siguiendo el recorrido del libro; esta ronda lo rehace como debe ser: capítulo por capítulo, del primero al último, añadiendo al frente de cada sección una caja de intuición que explica de dónde sale la idea *antes* de que aparezca el simbolismo, y sin borrar ni alterar una sola línea de lo ya escrito. **Parte I:** las letras como números con nombre y los productos notables como áreas (el álgebra elemental); la función como máquina de una sola salida y el dominio como lista de prohibiciones (las funciones elementales); el logaritmo como la pregunta "¿a qué exponente?" y sus tres leyes como las tres de los exponentes vistas al espejo (los logaritmos); el radián medido con la propia circunferencia y la identidad pitagórica como Pitágoras sobre el círculo unidad (la trigonometría elemental); la traducción de Descartes entre álgebra y geometría, con la pendiente como semilla de la derivada (la geometría analítica); el límite como "hacia dónde va" y el $0/0$ como una carrera, no como una respuesta (los límites elementales); y las cuatro reglas de montaje de la derivada, la regla de la cadena como engranajes y la integración por partes como la regla del producto al revés (las reglas de derivación e integración). **Parte II:** la promesa ε–N y por qué $\mathbb{Q}$ falla, y derivar como linealizar (capítulo 6); el producto de matrices como composición y la SVD como rotación, estiramiento y rotación (capítulo 7); los infinitos caminos de aproximación y el jacobiano como factor de volumen de la celda (capítulo 9); la curva como viaje frente al dibujo, y gradiente, divergencia y rotacional con lectura física (capítulo 11); el campo de direcciones y los tres regímenes de amortiguamiento (capítulo 12). **Parte III:** por qué tres distancias distintas del plano dan las mismas sucesiones convergentes (capítulo 28); la rigidez de la frontera que determina el interior y los mapas conformes (capítulo 30); Lebesgue como contar monedas por valor y no por orden de aparición, y por qué los teoremas de convergencia necesitan un techo (capítulo 31). **Parte IV:** el algoritmo de Euclides como columna vertebral y la unicidad de la factorización como teorema de verdad (capítulo 3); el orden de un elemento como "cuánto tarda en volver a casa", las clases laterales como baldosas y el cociente como mirar con lentes borrosos (capítulo 23); $K[x]$ como la copia más fiel de $\mathbb{Z}$ y la característica como "qué pasa si sumo uno muchas veces" (capítulo 25); la extensión de cuerpos como agregar la raíz que falta, la cadena de subgrupos como esqueleto de la fórmula, y regla y compás resueltos contando dimensiones (capítulo 26). **Parte V:** el convenio de Einstein como corrector ortográfico, la diferencia física entre flecha y máquina de medir flechas, el tensor como máquina de ranuras, la curvatura como el giro que sufre un vector al pasear en círculo, y la covariancia como exigencia de que las leyes no dependan de quién las escribe (capítulo 36). La formalidad queda intacta: solo se añadieron ayudas.

Capítulo nuevo: anillos, cuerpos y polinomios desde cero. La Parte IV tenía un salto: de la teoría de grupos (capítulo 23) se pasaba directamente a las estructuras algebraicas generales (capítulo 25), donde los anillos aparecían ya axiomatizados y en compañía de todo lo demás. El capítulo 24 llena ese hueco con el mismo método que el de grupos: primero el experimento — la tabla completa de multiplicación de $\mathbb{Z}_6$, el descubrimiento de que $2 \cdot 3 = 0$ sin que ni 2 ni 3 lo sean, y el hallazgo de que solo 1 y 5 tienen inverso — y solo después los axiomas, regla por regla. De ahí salen, en orden y sin saltos: el catálogo de anillos con sus contraejemplos exactos (por qué $\mathbb{N}$ no lo es, por qué los impares no lo son, por qué $2\mathbb{Z}$ no tiene unidad); unidades y divisores de cero con los grupos $U(\mathbb{Z}_8)$, $U(\mathbb{Z}_9)$, $U(\mathbb{Z}_{10})$, $U(\mathbb{Z}_{12})$ calculados a mano; dominios de integridad y la demostración de que todo dominio finito es un cuerpo; los $\mathbb{Z}_n$ con tablas de inversos completas para $\mathbb{Z}_7$ y $\mathbb{Z}_{11}$; homomorfismos, ideales y el teorema de que todo ideal de $\mathbb{Z}$ es $n\mathbb{Z}$; los cocientes, con la fabricación explícita de $\mathbb{C}$ como $\mathbb{R}[x]/(x^2+1)$; la división con resto en $K[x]$ demostrada y ejecutada; raíces, factores y el criterio de la raíz racional; irreducibilidad con su caja de herramientas completa (Eisenstein, reducción módulo p, cambio de variable) y cinco ejemplos decididos; la construcción con las manos de $\mathbb{F}_4$ y $\mathbb{F}_8$ con sus tablas verificadas; la ruptura de la factorización única en $\mathbb{Z}[\sqrt{-5}]$, que es de donde nacen los ideales; la jerarquía anillo $\supset$ dominio $\supset$ DFU $\supset$ DIP $\supset$ euclídeo $\supset$ cuerpo con un contraejemplo en cada escalón; y un taller de nueve ejercicios resueltos. Con esto, cuando el lector llegue al capítulo 25 ya habrá tocado los objetos con las manos, y el capítulo 26 podrá apoyarse en cuerpos construidos, no postulados.

Capítulo nuevo: análisis real, los cimientos y las patologías. La Parte III pasaba de la topología (capítulo 28) a la variable compleja sin detenerse en las preguntas incómodas que uno se hace *después* de creer que ya entendió el cálculo. El capítulo 29 las responde todas, con demostración: cuántos reales hay comparados con cuántos racionales (argumento diagonal de Cantor, y la existencia de trascendentes por puro conteo, sin exhibir ninguno); el conjunto de Cantor, con la cantidad de puntos de toda la recta y longitud cero, y su escalera del diablo, continua, creciente de 0 a 1 y con derivada nula en casi todo punto; la oscilación en un punto como número, y con ella el criterio de Lebesgue, que dice *exactamente* qué funciones acotadas son integrables Riemann (las que se portan mal en un conjunto de medida cero), con la función de Thomae como ejemplo de continuidad exactamente en los irracionales; el teorema de Baire y su corolario de que ninguna función puede ser continua exactamente en los racionales; el principio de contracción de Banach con su cota de error y el cálculo de la raíz de $\cos x = x$; la función de Weierstrass, continua en todo punto y derivable en ninguno, con la contabilidad exacta de la condición $ab > 1 + 3\pi/2$; los teoremas de Dini y de Arzelà–Ascoli, que explican qué hay que añadir a "cerrado y acotado" para recuperar la compacidad en dimensión infinita; la aproximación por polinomios de Bernstein y el teorema de Stone–Weierstrass, del que cuelga toda la teoría de Fourier; el teorema de reordenamiento de Riemann, con el cálculo explícito de que reordenar la serie armónica alternada tomando dos positivos por cada negativo cambia la suma de $\ln 2$ a $\frac{3}{2}\ln 2$; y la jerarquía $C^1 \subset$ Lipschitz $\subset AC \subset BV \cap C \subset C$ con un contraejemplo en cada escalón. El capítulo cierra explicando por qué la convergencia absoluta y la continuidad absoluta son la misma idea con dos disfraces, y por qué el capítulo 31 tiene que existir.

Capítulo nuevo: análisis funcional. El capítulo 32 cierra la Parte III uniendo las dos mitades del libro: el álgebra lineal del capítulo 7 y el análisis de los capítulos 29 y 31. La idea que lo organiza es que una función puede tratarse como un punto y una ecuación diferencial o integral como un operador lineal actuando sobre ese punto. De ahí salen, en orden: espacios normados y de Banach, con la observación de que $C[0,1]$ es completo con la norma del supremo y no lo es con la norma L^1 —y que su compleción es justamente L^1, lo que explica *para qué* se inventó la integral de Lebesgue—; la equivalencia de todas las normas en dimensión finita y el teorema de Riesz, que dice que la bola unidad es compacta si y solo si la dimensión es finita; los espacios de Hilbert, con el teorema de proyección sobre convexos cerrados demostrado a partir de la sola ley del paralelogramo, y con él las bases ortonormales, Bessel, Parseval y el isomorfismo con ℓ^2; las series de Fourier presentadas como el caso particular de elegir bien una base, con el problema de Basilea $\sum n^{-2} = \pi^2/6$ como subproducto; la equivalencia entre continuidad y acotación para operadores lineales, con normas calculadas de verdad (la matriz $\left(\begin{smallmatrix} 1 & 2 \\ 3 & 4 \end{smallmatrix}\right)$ en las tres normas, el operador de Volterra con $\|V\| = 2/\pi$ y sin ningún autovalor, la derivada como ejemplo de operador no acotado); los cuatro teoremas grandes —Hahn–Banach, Banach–Steinhaus, aplicación abierta y gráfica cerrada— con la observación de que los tres últimos son Baire disfrazado, y con la demostración por acotación uniforme de que existe una función continua cuya serie de Fourier diverge en un punto; el teorema de representación de Riesz, que es la notación de bras y kets de la física; los operadores compactos y el teorema espectral, que justifica de una vez todos los desarrollos en autofunciones que aparecen al separar variables; y la convergencia débil con Banach–Alaoglu, que devuelve la compacidad que Riesz había quitado y sostiene el método directo del cálculo de variaciones. Cierra con nueve ejercicios resueltos, entre ellos el cálculo de $\sum (2k+1)^{-4} = \pi^4/96$ por Parseval con comprobación cruzada, y el desplazamiento en ℓ^2 como ejemplo de operador sin autovalores cuyo adjunto tiene un disco entero de ellos.

Índice alfabético. El libro pasó de ser un texto que se lee en orden a ser también un texto que se consulta. Al final del volumen hay ahora un índice alfabético con doscientas sesenta y cinco entradas y setecientas cincuenta y nueve remisiones de página, construido marcando en el propio texto el lugar donde cada concepto aparece explicado o usado de forma significativa, no todas las veces que se menciona. Las entradas están ordenadas por el sustantivo que las identifica y no por la palabra que las acompaña —"teorema de Baire" aparece bajo la B, "función inyectiva" bajo la I, "polinomio irreducible" bajo la I—, con las tildes ordenadas como si no existieran, que es lo que hace falta para que un índice en español quede en el orden en que uno lo busca. Cada número de página es además un enlace: en el PDF se pulsa y salta. La utilidad práctica se ve mejor con un ejemplo: la entrada "límite" remite a dieciséis páginas repartidas por todo el libro, desde la definición escolar de la Parte I hasta el límite de una sucesión de tensores de la Parte V, de modo que el índice funciona también como mapa de cómo una misma idea reaparece a distintos niveles de profundidad. Nada del contenido se tocó para construirlo: las marcas de índice son invisibles en la página y el texto es exactamente el mismo de antes.

Ejercicios propuestos. A partir de esta edición los veintitrés capítulos terminan con una sección de *ejercicios propuestos*: doce enunciados sin solución desarrollada y, al final, una caja gris con las respuestas. Son distintos de los ejercicios resueltos que ya había dentro de cada

capítulo, y la diferencia es deliberada. Leer una solución bien escrita produce una sensación de comprensión que casi nunca sobrevive al papel en blanco; por eso aquí solo se da el resultado, lo justo para comprobar. El orden dentro de cada sección va de lo mecánico a lo conceptual: los primeros se resuelven aplicando una regla del capítulo, los últimos piden decidir qué teorema aplica o construir un contraejemplo. En las partes I y II casi todas las respuestas son números o expresiones; a partir de la parte III son cada vez más un nombre de teorema o una frase, de modo que conviene comparar el fondo del argumento y no su redacción. En total son doscientos setenta y seis ejercicios nuevos, y todos los que admiten verificación simbólica fueron comprobados con un sistema de álgebra computacional antes de imprimirse.

Formulario y glosario de símbolos. Los dos apéndices que siguen a este resuelven problemas distintos. El formulario reúne, en una sola tirada de páginas, todo lo que uno quiere tener delante mientras resuelve: productos notables, cónicas, identidades trigonométricas, leyes de logaritmos, tablas de derivadas e integrales, desarrollos de Maclaurin, los teoremas de Stokes y de la divergencia, las recetas de ecuaciones diferenciales y una lista de constantes. No está pensado para leerse sino para consultarse, y por eso va sin explicaciones: cada fórmula tiene su deducción en el capítulo correspondiente, y usarla sin haber pasado por ahí es justamente lo que este libro intenta evitar. El glosario de símbolos ataca un problema que casi nadie enseña de forma explícita: la notación se aprende por contagio, y eso deja huecos que después cuesta detectar. Están recogidos los símbolos de conjuntos y lógica, de números, de análisis, de álgebra y de geometría, cada uno con su lectura en voz alta y con el capítulo donde aparece por primera vez, más una nota final sobre los convenios tipográficos de este libro: coma decimal, corchetes, el cuadrado negro que cierra demostraciones y el código de colores de las cajas.

Dos capítulos nuevos al principio de la Parte I: lógica y combinatoria. La Parte I arrancaba con el álgebra, y eso dejaba dos huecos que se notaban a lo largo de todo el libro. El primero es de idioma: desde el capítulo 6 el texto usa “sea”, “basta ver que”, “recíprocamente”, cuantificadores anidados y demostraciones por contradicción, sin haber explicado nunca cómo se leen. El capítulo 1 lo remedia y es ahora el primero del libro: conjuntos (pertenencia frente a inclusión, conjunto potencia con su 2^n contado, producto cartesiano, De Morgan comprobado en un ejemplo), lógica proposicional con la tabla de verdad completa y la discusión de por qué una implicación con hipótesis falsa es verdadera, las cuatro formas de una implicación con el contrarrecíproco como única equivalente, cuantificadores con la negación de la definición ε–δ hecha paso a paso y la advertencia de que $\forall\exists$ y $\exists\forall$ dicen cosas distintas —que es toda la diferencia entre continuidad y continuidad uniforme—, y los cinco métodos de demostración (directa, contrarrecíproco, contradicción con la irracionalidad de $\sqrt{2}$, contraejemplo con n^2+n+41 fallando en el cuadragésimo intento, e inducción simple y fuerte). Cierra con una sección sobre cómo se lee una demostración, traduciendo el vocabulario del oficio.

El segundo hueco era temático: un libro que se llama “de cero a” no puede saltarse el conteo. El capítulo 5 entra entre la geometría analítica y los límites, y es autocontenido: los principios multiplicativo y aditivo con la traducción “y multiplica, o suma”; permutaciones, variaciones y permutaciones con repetición (las 151 200 palabras de MATEMATICA); combinaciones y coeficiente binomial, con la identidad $P(n,k)=\binom{n}{k}k!$ como la única cosa que hay que recordar; el binomio de Newton demostrado contando, no por inducción, y el triángulo de Pascal; inclusión–exclusión y el principio del palomar; y la probabilidad como combinatoria dividida

entre combinatoria, con la regla de Laplace, el problema del cumpleaños (0,5073 con veintitrés personas, y la explicación de por qué hay que contar parejas y no personas), la probabilidad condicionada, el teorema de Bayes con la prueba médica que deja al paciente positivo en un 16,7 %, Monty Hall, la distribución binomial y la esperanza. Termina explicando que la axiomática de Kolmogorov consiste en decir que una probabilidad es una medida de masa total uno, con lo que el capítulo 31 queda anunciado desde la Parte I. Ambos capítulos traen sus ejercicios resueltos y sus doce propuestos con respuesta, y todos los números fueron verificados simbólicamente antes de escribirse.

Segunda pasada del índice alfabético. El índice descrito más arriba se construyó cuando el libro tenía veintitrés capítulos y terminaba en la bibliografía. Todo lo añadido después —las secciones nuevas de los capítulos escolares, la ampliación de la topología, el formulario, el glosario y los dos capítulos de lógica y combinatoria— entró sin marcas, de modo que había materia explicada que el índice no encontraba. Esta pasada la incorpora: entradas para inecuación, raíz extraña, división sintética, teorema del resto y del factor, función a trozos, asíntota oblicua, función par e impar, escala de pH, escala de Richter, decibelio, interés compuesto, vida media, carbono 14, ángulo doble, caso ambiguo, fórmula de Herón, excentricidad, baricentro, ortocentro, circuncentro, recta de Euler, límite lateral, teorema del valor intermedio, bisección, aproximación lineal, razones de cambio relacionadas, optimización, sustitución trigonométrica, volumen de revolución, método de discos y de capas, longitud de arco, subbase, recta de Sorgenfrey, topología producto y cociente, axiomas de separación, Hausdorff, Urysohn, Tietze, Tychonoff, botella de Klein, curva seno del topólogo, grupo fundamental, Brouwer, factorial, coeficiente binomial, triángulo de Pascal, inclusión–exclusión, principio del palomar, espacio muestral, regla de Laplace, probabilidad condicionada, independencia, Bayes, distribución binomial, esperanza, conjunto, conjunto potencia, producto cartesiano, De Morgan, recíproco, contrarrecíproco y demostración por contradicción. El índice pasa así de setecientas cincuenta y nueve remisiones a más de ochocientas cincuenta. Como en la primera pasada, las marcas son invisibles y ni una línea del texto cambió.

Un capítulo nuevo entre la combinatoria y los límites: sucesiones y series. Faltaba el eslabón. El libro pasaba de contar objetos finitos (capítulo 5) a la definición ε–δ de límite (los límites elementales) sin haber dicho nunca qué significa sumar infinitos números, cuando esa es precisamente la pregunta que hace falta haberse hecho antes para que la idea de límite parezca necesaria y no caprichosa. El las sucesiones y series elementales llena ese hueco con material estrictamente escolar: qué es una sucesión y las dos maneras de darla (explícita y recursiva), progresiones aritméticas con el truco de Gauss demostrado en dos líneas, progresiones geométricas con el truco del desplazamiento, la interpolación de medios, el interés simple frente al compuesto, y después el salto: sumas parciales, convergencia, la fórmula $a_1/(1-r)$ con su hipótesis $|r|<1$ subrayada, y sus tres aplicaciones clásicas (la paradoja de Zenón, los decimales periódicos con la igualdad $0{,}\overline{9}=1$ explicada en serio, y la pelota que rebota infinitas veces recorriendo cuarenta metros). De ahí sale el número e como límite de $(1+1/n)^n$ —motivado por la capitalización continua, que es de donde históricamente salió—, la notación de sumatorio con las tres sumas de potencias, las sumas telescópicas, Fibonacci con la razón áurea obtenida al resolver $L=1+1/L$, y la serie armónica como

advertencia de que $a_n \to 0$ no basta. Termina con diez errores típicos, siete ejercicios resueltos y doce propuestos con respuesta.

Con esto la Parte I queda cerrada en su forma definitiva: lógica, álgebra, funciones, logaritmos, trigonometría, geometría analítica, combinatoria, sucesiones, límites y manual de cálculo. Nada de lo que la Parte II da por sabido queda ya sin explicar antes.

Los complejos, por fin, en su sitio. El libro trataba los números complejos en el capítulo 30 con todo el aparato del análisis —holomorfía, Cauchy, residuos— y los mencionaba de pasada en el capítulo 6, pero nunca había explicado desde cero qué es i. El capítulo 4 lo hace ahora dentro de la Parte I, justo después de la geometría analítica, que es donde el plano de Argand tiene sentido. Empieza por la historia verdadera y no por la escolar: los complejos no se inventaron para resolver $x^2 + 1 = 0$ —una ecuación que nadie necesita resolver— sino porque la fórmula de Cardano para la cúbica $x^3 = 15x + 4$, cuya solución es el humildísimo $x = 4$, obliga a pasar por $\sqrt{-121}$ para llegar allí. Después vienen la forma binómica, el ciclo de cuatro de las potencias de i, el conjugado como herramienta de división, el plano de Argand con el módulo convertido en Pitágoras y la conjugación en una reflexión, y el resultado central del capítulo: multiplicar por un complejo es *girar y estirar*, de donde $i^2 = -1$ deja de ser una definición arbitraria y pasa a ser el hecho geométrico de que dos cuartos de vuelta son media vuelta. De ahí salen De Moivre, las raíces n-ésimas dibujando polígonos regulares, las cuadráticas de discriminante negativo que la Parte I había dejado a medias, el teorema de las raíces conjugadas con su corolario sobre polinomios de grado impar, el enunciado del teorema fundamental del álgebra y un anticipo de la fórmula de Euler. Cierra con diez errores típicos —encabezados por el clásico $\sqrt{-4}\sqrt{-9} = 6$, que es falso—, seis ejercicios resueltos y doce propuestos.

Con este capítulo, las tres apariciones de $\mathbb{C}$ en el libro quedan encadenadas: aquí se construye y se dibuja, en el capítulo 30 se hace análisis sobre él, y en los capítulos 24 y 26 se explica por qué era el único cuerpo posible al que llegar.

Estadística descriptiva: la otra mitad de la matemática aplicada. Un libro que va "de cero a Galois" y no dice nunca qué es una desviación típica tiene un agujero raro, porque la estadística descriptiva es, con diferencia, la parte de la matemática que más gente usa en su trabajo. El la estadística descriptiva la cubre justo después de la probabilidad del capítulo 5, con el que forma pareja natural: allí se va del modelo a los datos, aquí de los datos al modelo. Está escrito con un sesgo deliberado hacia *cuándo cada medida miente* y no hacia cómo se calcula, porque calcular una media no es el problema de nadie. Así, la comparación media contra mediana se hace con seis sueldos de los cuales cinco están por debajo de la media; el coeficiente de variación aparece para explicar por qué comparar desviaciones típicas de variables con unidades distintas no significa nada; el diagrama de caja se presenta como el dibujo de la desigualdad; y la sección de regresión termina con las dos advertencias que valen por el capítulo entero: que la correlación no es causalidad —con las cuatro explicaciones alternativas que admite siempre una correlación alta— y que $r = 0$ significa "sin relación lineal", ilustrado con cinco puntos de r exactamente nulo que están *perfectamente* sobre una parábola. Incluye también datos agrupados en intervalos con interpolación de la mediana, mínimos cuadrados enunciado como teorema y conectado con la proyección ortogonal de los

capítulos 7 y 32, la regla 68–95–99,7, las puntuaciones típicas y una nota sobre el teorema central del límite. Cierra con once errores típicos, seis ejercicios resueltos y doce propuestos.

Geometría clásica: el capítulo que faltaba antes de la trigonometría. El libro hablaba de senos, de cónicas y de variedades de Riemann, pero en ningún sitio se había demostrado que la suma de los ángulos de un triángulo sea 180ř, ni que los triángulos semejantes tengan los lados proporcionales. Sin ese segundo hecho, la definición misma de seno como «cateto opuesto entre hipotenusa» está mal planteada: nada garantiza que la razón no dependa del triángulo elegido. El la geometría elemental tapa ese hueco y se coloca justo antes de la trigonometría, que es donde hace falta. De paso recoge lo que un curso escolar da por sabido y este libro nunca había escrito: áreas, volúmenes, puntos notables, el ángulo inscrito y la ley del cuadrado-cubo, que es el error de escala más caro de toda la matemática elemental.

Aritmética: por fin el cero del título. El libro se llama *De cero a Galois* y durante mucho tiempo su «cero» fue el álgebra escolar, que ya suponía saber operar con fracciones, potencias y porcentajes. Eso no es cero: es primero de secundaria. El la aritmética elemental escribe lo que faltaba —divisibilidad, mcd y mcm, fracciones, decimales periódicos, notación científica, porcentajes, proporcionalidad y error— y lo escribe con el mismo criterio que el resto del libro: justificando las reglas en lugar de dictarlas. Ahí se explica por qué $(-1)(-1) = 1$ a partir de la distributiva, por qué $a^0 = 1$ es la única opción posible y por qué dos descuentos sucesivos del 30 % y el 20 % suman 44 % y no 50 %. Con este capítulo, el título deja de ser una exageración.

Matrices: el puente que faltaba hacia la Parte II. El salto más brusco que quedaba en el libro era este: el el álgebra elemental resolvía dos ecuaciones por sustitución, y el capítulo 7 abría hablando de espacios vectoriales y aplicaciones lineales. Entre esas dos páginas faltaba todo lo que un curso escolar dedica a matrices: el producto fila por columna y por qué se define así, el determinante como cuenta antes de ser volumen, la inversa, el método de Gauss sobre la matriz ampliada y la clasificación de sistemas en compatibles determinados, indeterminados e incompatibles. El el cálculo matricial elemental cubre ese tramo y deja preparadas, sin nombrarlas todavía, las nociones de rango y de transformación lineal que la Parte II formaliza.

Vectores: la geometría que se puede calcular. El los vectores elementales traduce la geometría sintética de la geometría elemental al lenguaje de las coordenadas: ángulos con el producto escalar, perpendiculares y áreas con el producto vectorial, volúmenes con el producto mixto. Es la última pieza escolar antes de que la Parte II hable de espacios vectoriales abstractos, y quien lo lea reconocerá en el capítulo 7 que $\mathbb{R}^n$ no es más que esto mismo con más coordenadas.

Inferencia: de la muestra a la población. El capítulo 16 cierra el bloque de datos que empiezan el capítulo 5 (probabilidad) y el la estadística descriptiva (descripción). Aquí aparece la variable aleatoria, que el capítulo 31 reconocerá después como una función medible, y el teorema central del límite, cuya demostración moderna es un argumento de Fourier. Quien

solo vaya a usar matemáticas fuera de la matemática debería leer este capítulo antes que ningún otro de la Parte II.

Finanzas: la progresión geométrica con nombre y apellidos. El la matemática financiera no introduce ninguna matemática nueva —todo sale de la suma geométrica de las sucesiones y series elementales y de los porcentajes de la aritmética elemental— y sin embargo es el que más se usa fuera de las aulas. Incluye además dos guiños hacia arriba: el número e aparece como techo de la capitalización continua, y el cálculo de la TIR se topa con la imposibilidad de resolver por radicales que demuestra el capítulo 26.

Programación lineal: optimizar sin derivar. Todo el cálculo diferencial busca máximos anulando una derivada. La programación lineal (la programación lineal) hace lo contrario: cuando la función objetivo y las restricciones son lineales, el óptimo no está donde la derivada se anula —no se anula nunca, el gradiente es constante— sino en un *vértice* de la región factible. Ese teorema fundamental es geometría convexa disfrazada de ejercicio de bachillerato: la región factible es un polítopo, la función objetivo alcanza su extremo en un punto extremo, y ambas afirmaciones reaparecen con todo su aparato en el álgebra lineal de la Parte II (capítulo 7) y en la compacidad de Weierstrass del capítulo 28. El método del símplex de Dantzig no es más que la estrategia de recorrer los vértices en orden de mejora; que ese recorrido sea corto en la práctica y largo en el peor caso es, hoy, un problema abierto de complejidad computacional.

El estudio de funciones: dos signos bastan. Toda la sección de representación de curvas (el estudio de funciones) descansa en un hecho que en bachillerato se usa sin nombrarlo: el signo de f' decide la monotonía y el de f'' la convexidad. Ninguna de las dos afirmaciones es evidente, y ambas se apoyan en el teorema del valor medio, que a su vez se apoya en el de Rolle, que a su vez se apoya en Weierstrass, que a su vez se apoya en la completitud de $\mathbb{R}$ (capítulo 29). Es decir: el ejercicio más rutinario de un examen de bachillerato —"estudia el crecimiento de esta función"— es lógicamente equivalente a que los reales no tengan agujeros. Y la noción de convexidad que aquí aparece como "forma de $\cup$" es la misma que, sin derivadas y en dimensión infinita, sostiene la teoría de la optimización moderna (la programación lineal y 32).

Grafos: la topología que cabe en una servilleta. Cuando Euler resolvió los puentes de Königsberg (capítulo 8) hizo algo que en 1736 no tenía nombre: descartó por irrelevantes todas las magnitudes —distancias, ángulos, áreas— y se quedó solo con la relación de incidencia. Eso es exactamente el gesto fundacional de la topología (capítulo 28), y no es casualidad que la fórmula $V - A + C = 2$ sea el primer *invariante topológico* de la historia: el número $V - A + C$ es la característica de Euler, vale 2 para todo lo que sea deformable en una esfera y 0 para todo lo que sea deformable en un toro, de modo que contar vértices, aristas y caras de un dibujo permite averiguar si la superficie tiene agujero. La misma cantidad reaparece, con todo el aparato del álgebra homológica, como suma alternada de rangos de grupos, y ahí conecta con la teoría de grupos del capítulo 23. Un acertijo de paseantes prusianos y la clasificación de superficies resultan ser, cien años después, el mismo asunto.

Numeración y Boole: dos objetos que resultaron ser el mismo. El los sistemas de numeración contiene, sin decirlo, dos teoremas de estructura. El primero es que todo natural admite una única escritura en base b: eso es exactamente decir que el anillo $\mathbb{Z}$ del capítulo 24 tiene división euclídea, y la unicidad de la representación es la unicidad del cociente y el resto. Toda la aritmética posicional, incluidos los algoritmos que se aprenden a los ocho años, es un corolario de esa propiedad.

El segundo es que el álgebra de Boole de dos elementos es, simultáneamente, el cuerpo $\mathbb{F}_2$ del capítulo 25 visto con otras operaciones, el retículo de los subconjuntos de un conjunto de un punto, y el modelo de la lógica proposicional del capítulo 1. El teorema de representación de Stone afirma que *toda* álgebra de Boole finita es isomorfa al álgebra de los subconjuntos de un conjunto finito, de modo que la lista de axiomas de Boole no describe una estructura exótica sino, salvo isomorfismo, únicamente las uniones e intersecciones de conjuntos. La o exclusiva $\oplus$ es la suma de $\mathbb{F}_2$ y el producto lógico es su producto; el complemento a dos es aritmética módulo 2^n; y el punto flotante es la advertencia, incómoda pero necesaria, de que $\mathbb{R}$ es un objeto infinito que ninguna máquina finita contiene.

Movimientos del plano: el programa de Erlangen en miniatura. El las transformaciones del plano termina observando que las isometrías del plano forman un grupo. Eso no es una curiosidad terminológica: es el punto de partida del *programa de Erlangen* que Felix Klein enunció en 1872, según el cual una geometría *es* un grupo de transformaciones junto con el estudio de lo que ese grupo deja invariante. La geometría euclídea es el estudio de los invariantes del grupo de isometrías (la distancia, el ángulo, el área); la geometría afín, del grupo de las aplicaciones lineales invertibles más traslaciones (se pierde la distancia, sobrevive el paralelismo y la razón simple); la proyectiva, de un grupo mayor todavía (solo sobrevive la razón doble); y la topología del capítulo 28, del grupo enorme de los homeomorfismos, donde ya solo sobreviven la conexión y la compacidad.

Vista así, la cadena de capítulos escolares no es una sucesión de temas sueltos sino una torre de grupos cada vez más grandes con cada vez menos invariantes. La clasificación de los diecisiete grupos cristalográficos planos es el primer teorema de clasificación de grupos que el lector encuentra en el libro, mucho antes de los grupos abstractos del capítulo 23; y la pregunta de Galois —qué permutaciones de las raíces dejan invariantes las relaciones algebraicas— es literalmente la misma pregunta formulada sobre una figura que no se puede dibujar.

Métodos numéricos: la contracción es la idea única. Los cuatro métodos del capítulo 14 parecen distintos y son el mismo teorema visto desde cuatro ángulos. Bisección, Newton, secante e iteración de punto fijo son todos casos de buscar un punto fijo de una aplicación g, y todos convergen por la misma razón: cerca de la solución, g es una contracción, es decir, acerca los puntos entre sí. El teorema del punto fijo de Banach del capítulo 32 enuncia eso en un espacio métrico completo cualquiera, y de él salen a la vez la convergencia de estos algoritmos, el teorema de existencia y unicidad de Picard–Lindelöf para ecuaciones diferenciales del capítulo 12 y el teorema de la función implícita.

El orden de convergencia mide cuánto contrae: lineal cuando $|g'(p)| = k < 1$ (el error se multiplica por k), cuadrático cuando además $g'(p) = 0$, que es exactamente lo que consigue la construcción de Newton al cancelar el término de primer orden del desarrollo de Taylor.

Y la completitud de $\mathbb{R}$ del capítulo 29 es lo que garantiza que el límite al que se acercan las iteraciones existe de verdad: sin ella los algoritmos convergerían hacia un agujero. La cancelación catastrófica, en cambio, no es análisis sino aritmética finita, y recuerda que el los sistemas de numeración no era un paréntesis.

Aritmética modular: el primer cociente de la historia de cada estudiante. Cuando en el capítulo 3 se escribe $a \equiv b \pmod{m}$ y se comprueba que la suma y el producto "pasan al resto", lo que se está construyendo sin decirlo es el anillo cociente $\mathbb{Z}/m\mathbb{Z}$: el conjunto de las clases de equivalencia de la relación "diferir en un múltiplo de m", con las operaciones heredadas. El teorema de compatibilidad de la sección de las tres reglas de oro es exactamente la verificación de que esas operaciones *están bien definidas* sobre las clases, que es el único punto delicado de toda construcción cociente y el que reaparecerá, con las mismas palabras, al formar grupos cociente en el capítulo 23 y anillos cociente en el capítulo 24. La condición $\mathrm{mcd}(a, m) = 1$ para que a tenga inverso dice, traducida, que las unidades de $\mathbb{Z}/m\mathbb{Z}$ son las clases de los primos con m; y que $\mathbb{Z}/m\mathbb{Z}$ sea un cuerpo equivale a que m sea primo, lo que enlaza con los cuerpos finitos $\mathbb{F}_p$ sobre los que se construye toda la teoría de Galois del capítulo 26. El teorema chino del resto, que aquí se presenta como un acertijo del siglo III, es el isomorfismo $\mathbb{Z}/mn\mathbb{Z} \cong \mathbb{Z}/m\mathbb{Z} \times \mathbb{Z}/n\mathbb{Z}$, primer ejemplo en el libro de una descomposición en factores que se volverá sistemática con los teoremas de estructura. Y los criterios de divisibilidad, que parecían trucos de primaria, son evaluaciones del morfismo de anillos $\mathbb{Z} \to \mathbb{Z}/9\mathbb{Z}$ en el número escrito en base 10: el mismo morfismo, mirado desde el los sistemas de numeración, dice que cambiar de base no cambia el número, solo su nombre.

Criptografía: por qué la dificultad computacional es un objeto matemático. Los cifrados de César, afín y Vigenère son, los tres, acciones de un grupo sobre el conjunto de las letras: César es la acción de $\mathbb{Z}/26\mathbb{Z}$ por traslaciones, el afín es la del grupo de transformaciones $x \mapsto ax + b$ con a unidad —el grupo afín de la recta sobre $\mathbb{Z}/26\mathbb{Z}$, de orden $\varphi(26) \cdot 26 = 312$—, y Vigenère es una acción de un producto directo $(\mathbb{Z}/26\mathbb{Z})^L$. Que todos caigan ante el análisis de frecuencias tiene una explicación estructural: el grupo es demasiado pequeño y la acción demasiado regular, de modo que las órbitas conservan la información estadística del texto. RSA rompe el esquema porque su seguridad no descansa en el tamaño del grupo sino en una *asimetría de coste*: la función $x \mapsto x^e$ en $(\mathbb{Z}/n\mathbb{Z})^\times$ es una biyección fácil de calcular cuya inversa es fácil de calcular *solo si* se conoce $\varphi(n)$. Esa es la primera vez, en toda la historia de la matemática, en que la dificultad de un cálculo se convierte en el cimiento de una construcción útil, y es también el punto donde la teoría de números dejó de ser, en palabras de Hardy, la más inútil de las ramas.

Polares y paramétricas: la curva como aplicación, no como conjunto. El capítulo 10 contiene, escondido bajo un lenguaje escolar, el cambio de punto de vista más importante de toda la geometría moderna. Una ecuación cartesiana $F(x, y) = 0$ define una curva como un *conjunto*: el lugar de los puntos que cumplen una condición. Una parametrización $t \mapsto (x(t), y(t))$ la define como una *aplicación* de un intervalo en el plano. La advertencia sobre lo que se pierde al eliminar el parámetro —el dominio, la orientación, el número de pasadas— es la primera aparición en el libro de una distinción que será central después: en el capítulo 11, donde la integral de línea depende de la orientación y no solo del conjunto recorrido; en el

capítulo 28, donde un camino es por definición una aplicación continua $[0,1] \to X$ y dos caminos con la misma imagen pueden no ser homótopos; y en el capítulo 36, donde una curva sobre una variedad solo tiene sentido como aplicación, porque no hay ninguna ecuación cartesiana ambiente en la que escribirla. El paso de coordenadas cartesianas a polares, por su parte, es el primer cambio de coordenadas no lineal del libro; su jacobiano r es el que aparecerá en el capítulo 9 al integrar en polares, y la razón profunda de que las polares fallen en el origen —donde el ángulo no está definido— es que no existe ninguna carta global para el círculo, hecho que en el capítulo 36 se vuelve la definición misma de variedad. La ambigüedad de las polares, que aquí es un incordio de cálculo, es allí un teorema.

Apéndice B

Formulario

Todo lo que conviene tener a mano en una sola tirada de páginas. No sustituye a los capítulos: una fórmula sin su idea es un préstamo que se devuelve con intereses el día del examen. Pero cuando ya entendiste de dónde sale, tenerla escrita ahorra tiempo.

B.1. Álgebra

$$(a \pm b)^2 = a^2 \pm 2ab + b^2 \qquad (a \pm b)^3 = a^3 \pm 3a^2 b + 3ab^2 \pm b^3$$
$$a^2 - b^2 = (a+b)(a-b) \qquad a^3 \pm b^3 = (a \pm b)(a^2 \mp ab + b^2)$$
$$a^m a^n = a^{m+n} \qquad (a^m)^n = a^{mn}$$
$$a^{-n} = \frac{1}{a^n} \qquad a^{m/n} = \sqrt[n]{a^m}$$

Ecuación cuadrática $ax^2 + bx + c = 0$:

$$x = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a}, \qquad x_1 + x_2 = -\frac{b}{a}, \qquad x_1 x_2 = \frac{c}{a}.$$

El discriminante $\Delta = b^2 - 4ac$ decide: $\Delta > 0$ dos raíces reales, $\Delta = 0$ una doble, $\Delta < 0$ dos complejas conjugadas.

Valor absoluto, con $c > 0$:

$$|E| = c \iff E = \pm c, \qquad |E| < c \iff -c < E < c, \qquad |E| > c \iff E < -c \text{ o } E > c.$$

Binomio de Newton:

$$(a+b)^n = \sum_{k=0}^{n} \binom{n}{k} a^{n-k} b^k, \qquad \binom{n}{k} = \frac{n!}{k!(n-k)!}.$$

B.2. Geometría analítica

$$d = \sqrt{(x_2-x_1)^2+(y_2-y_1)^2} \qquad M = \left(\frac{x_1+x_2}{2}, \frac{y_1+y_2}{2}\right)$$

$$m = \frac{y_2-y_1}{x_2-x_1} \qquad y - y_1 = m(x-x_1)$$

$$\text{paralelas} : m_1 = m_2 \qquad \text{perpendiculares} : m_1 m_2 = -1$$

Distancia de (x_0, y_0) a $Ax + By + C = 0$: $d = \dfrac{|Ax_0 + By_0 + C|}{\sqrt{A^2+B^2}}$.

$$\text{circunferencia} : (x-h)^2 + (y-k)^2 = r^2$$

$$\text{parábola} : (x-h)^2 = 4p(y-k) \quad \text{foco a distancia } |p| \text{ del vértice}$$

$$\text{elipse} : \frac{(x-h)^2}{a^2} + \frac{(y-k)^2}{b^2} = 1, \quad c^2 = a^2 - b^2$$

$$\text{hipérbola} : \frac{(x-h)^2}{a^2} - \frac{(y-k)^2}{b^2} = 1, \quad c^2 = a^2 + b^2, \quad y - k = \pm\tfrac{b}{a}(x-h)$$

Excentricidad $e = c/a$: vale 0 en la circunferencia, está entre 0 y 1 en la elipse, vale 1 en la parábola y supera 1 en la hipérbola.

B.3. Trigonometría

$$\sin^2 x + \cos^2 x = 1, \qquad 1 + \tan^2 x = \sec^2 x, \qquad 1 + \cot^2 x = \csc^2 x.$$

$$\sin(A \pm B) = \sin A \cos B \pm \cos A \sin B \qquad \cos(A \pm B) = \cos A \cos B \mp \sin A \sin B$$

$$\sin 2A = 2 \sin A \cos A \qquad \cos 2A = \cos^2 A - \sin^2 A = 2\cos^2 A - 1 = 1 - 2\sin^2 A$$

$$\sin^2 A = \frac{1-\cos 2A}{2} \qquad \cos^2 A = \frac{1+\cos 2A}{2}$$

Valores notables:

	0	$\pi/6$	$\pi/4$	$\pi/3$	$\pi/2$	π
sin	0	$1/2$	$\sqrt{2}/2$	$\sqrt{3}/2$	1	0
cos	1	$\sqrt{3}/2$	$\sqrt{2}/2$	$1/2$	0	-1
tan	0	$\sqrt{3}/3$	1	$\sqrt{3}$	—	0

Resolución de triángulos:

$$\frac{a}{\sin A} = \frac{b}{\sin B} = \frac{c}{\sin C}, \qquad c^2 = a^2 + b^2 - 2ab\cos C, \qquad S = \tfrac{1}{2}ab\sin C = \sqrt{s(s-a)(s-b)(s-c)}.$$

B.4. Exponenciales y logaritmos

$$\log_b(xy) = \log_b x + \log_b y \qquad \log_b \frac{x}{y} = \log_b x - \log_b y$$

$$\log_b x^n = n \log_b x \qquad \log_b a = \frac{\ln a}{\ln b}$$

$$b^{\log_b x} = x \qquad \log_b b^x = x$$

Crecimiento y decaimiento: $N(t) = N_0 e^{kt}$, con semivida $T = \dfrac{\ln 2}{|k|}$. Interés: $A = P\left(1 + \frac{r}{n}\right)^{nt}$, y en el límite continuo $A = Pe^{rt}$.

B.5. Límites

$$\lim_{x\to 0} \frac{\sin x}{x} = 1, \qquad \lim_{x\to 0} \frac{1-\cos x}{x^2} = \frac{1}{2}, \qquad \lim_{x\to\infty} \left(1 + \frac{a}{x}\right)^x = e^a, \qquad \lim_{x\to 0} \frac{e^x - 1}{x} = 1.$$

Las siete indeterminaciones: $\frac{0}{0}$, $\frac{\infty}{\infty}$, $0 \cdot \infty$, $\infty - \infty$, 1^∞, 0^0, ∞^0.

B.6. Derivadas

$$(x^n)' = nx^{n-1} \qquad (e^x)' = e^x \qquad (a^x)' = a^x \ln a \qquad (\ln x)' = \frac{1}{x}$$

$$(\sin x)' = \cos x \qquad (\cos x)' = -\sin x \qquad (\tan x)' = \sec^2 x \qquad (\sec x)' = \sec x \tan x$$

$$(\arcsin x)' = \frac{1}{\sqrt{1-x^2}} \qquad (\arctan x)' = \frac{1}{1+x^2} \qquad (\sinh x)' = \cosh x \qquad (\cosh x)' = \sinh x$$

$$(fg)' = f'g + fg', \qquad \left(\frac{f}{g}\right)' = \frac{f'g - fg'}{g^2}, \qquad (f \circ g)' = f'(g(x))\, g'(x).$$

B.7. Integrales

$$\int x^n \mathrm{d}x = \frac{x^{n+1}}{n+1} + C \ (n \neq -1) \qquad \int \frac{\mathrm{d}x}{x} = \ln|x| + C$$

$$\int e^x \mathrm{d}x = e^x + C \qquad \int a^x \mathrm{d}x = \frac{a^x}{\ln a} + C$$

$$\int \sin x \,\mathrm{d}x = -\cos x + C \qquad \int \cos x \,\mathrm{d}x = \sin x + C$$

$$\int \sec^2 x \,\mathrm{d}x = \tan x + C \qquad \int \tan x \,\mathrm{d}x = -\ln|\cos x| + C$$

$$\int \frac{\mathrm{d}x}{1+x^2} = \arctan x + C \qquad \int \frac{\mathrm{d}x}{\sqrt{1-x^2}} = \arcsin x + C$$

Por partes: $\int u\,\mathrm{d}v = uv - \int v\,\mathrm{d}u$ (elige u según LIATE).
Sustituciones trigonométricas: $\sqrt{a^2-x^2} \Rightarrow x = a\sin\theta$; $\sqrt{a^2+x^2} \Rightarrow x = a\tan\theta$; $\sqrt{x^2-a^2} \Rightarrow x = a\sec\theta$.

Aplicaciones:

$$A = \int_a^b (f-g), \qquad V_{\text{discos}} = \pi\int_a^b f^2, \qquad V_{\text{capas}} = 2\pi\int_a^b xf, \qquad L = \int_a^b \sqrt{1+(f')^2}.$$

B.8. Series

$$\sum_{n=0}^{\infty} ar^n = \frac{a}{1-r}\ (|r|<1), \qquad \sum_{n=1}^{\infty}\frac{1}{n^p}\ \text{converge} \iff p>1.$$

Desarrollos de Maclaurin:

$$e^x = \sum_{n\geq 0}\frac{x^n}{n!} \qquad \sin x = x - \frac{x^3}{3!} + \frac{x^5}{5!} - \cdots$$

$$\cos x = 1 - \frac{x^2}{2!} + \frac{x^4}{4!} - \cdots \qquad \ln(1+x) = x - \frac{x^2}{2} + \frac{x^3}{3} - \cdots\ (|x|<1)$$

$$\frac{1}{1-x} = \sum_{n\geq 0} x^n\ (|x|<1) \qquad (1+x)^\alpha = \sum_{n\geq 0}\binom{\alpha}{n}x^n\ (|x|<1)$$

B.9. Álgebra lineal

Para $A = \left(\begin{smallmatrix} a & b \\ c & d\end{smallmatrix}\right)$: $\det A = ad-bc$ y $A^{-1} = \frac{1}{ad-bc}\left(\begin{smallmatrix} d & -b \\ -c & a\end{smallmatrix}\right)$.

$$\text{rango}\,A + \text{nulidad}\,A = n, \qquad \operatorname{tr} A = \sum \lambda_i, \qquad \det A = \prod \lambda_i.$$

Proyección de v sobre u: $\operatorname{proy}_u v = \dfrac{\langle v,u\rangle}{\langle u,u\rangle}u$.

B.10. Cálculo vectorial

$$\operatorname{grad} f = \left(\frac{\partial f}{\partial x}, \frac{\partial f}{\partial y}, \frac{\partial f}{\partial z}\right), \qquad \operatorname{div} F = \frac{\partial F_1}{\partial x} + \frac{\partial F_2}{\partial y} + \frac{\partial F_3}{\partial z}, \qquad \operatorname{rot} F = \nabla \times F.$$

Identidades: $\operatorname{rot}\operatorname{grad} f = 0$ y $\operatorname{div}\operatorname{rot} F = 0$.

$$\oint_{\partial D} F\cdot \mathrm{d}r = \iint_D (\operatorname{rot} F)\cdot n\,\mathrm{d}S\ \text{(Stokes)}, \qquad \iint_{\partial V} F\cdot n\,\mathrm{d}S = \iiint_V \operatorname{div} F\,\mathrm{d}V\ \text{(divergencia)}.$$

Jacobianos: polares r; cilíndricas r; esféricas $\rho^2\sin\varphi$.

B.11. Ecuaciones diferenciales

Variables separables: $\dfrac{\mathrm{d}y}{\mathrm{d}x} = g(x)h(y) \Rightarrow \displaystyle\int \frac{\mathrm{d}y}{h(y)} = \int g(x)\mathrm{d}x.$

Lineal de primer orden $y' + P(x)y = Q(x)$: factor integrante $\mu = e^{\int P}$, y $y = \frac{1}{\mu}\int \mu Q\,\mathrm{d}x$.
Lineal homogénea de coeficientes constantes $ay''+by'+cy = 0$: raíces de $ar^2+br+c=0$. Reales distintas $\Rightarrow c_1e^{r_1x}+c_2e^{r_2x}$; doble $\Rightarrow (c_1+c_2x)e^{rx}$; complejas $\alpha\pm\beta i \Rightarrow e^{\alpha x}(c_1\cos\beta x+c_2\sin\beta x)$.
Transformadas: $\mathcal{L}\{1\} = \frac{1}{s}$, $\mathcal{L}\{e^{at}\} = \frac{1}{s-a}$, $\mathcal{L}\{t^n\} = \frac{n!}{s^{n+1}}$, $\mathcal{L}\{f'\} = sF(s) - f(0)$.

B.12. Geometría del plano y del espacio

$$\text{triángulo} : S = \tfrac{1}{2}bh \qquad \text{trapecio} : S = \tfrac{(B+b)h}{2}$$
$$\text{círculo} : L = 2\pi r,\ S = \pi r^2 \qquad \text{sector de ángulo } \alpha : S = \tfrac{1}{2}r^2\alpha \ (\alpha \text{ en radianes})$$
$$\text{polígono regular} : S = \tfrac{\text{perímetro}\cdot\text{apotema}}{2} \qquad \text{corona} : S = \pi(R^2 - r^2)$$

$$\text{prisma/cilindro} : V = S_b h \qquad \text{pirámide/cono} : V = \tfrac{1}{3}S_b h$$
$$\text{esfera} : V = \tfrac{4}{3}\pi r^3,\ S = 4\pi r^2 \qquad \text{cono} : S_{\text{lat}} = \pi r g,\ g = \sqrt{r^2+h^2}$$
$$\text{cilindro} : S_{\text{lat}} = 2\pi r h \qquad \text{casquete de altura } h : S = 2\pi R h$$

Semejanza de razón k: longitudes $\times k$, áreas $\times k^2$, volúmenes $\times k^3$. Pitágoras: $a^2+b^2=c^2$; teorema de Tales; ángulo inscrito = mitad del central. Suma de ángulos interiores de un n-ágono: $(n-2)\cdot 180°$; de un polígono regular, cada uno mide $\frac{(n-2)180°}{n}$.

B.13. Números complejos

$$z = a+bi = r(\cos\theta + i\sin\theta) = r\,\mathrm{e}^{i\theta}, \qquad r = |z| = \sqrt{a^2+b^2}, \qquad \bar{z} = a-bi, \qquad z\bar{z} = |z|^2.$$

$$z_1z_2 = r_1r_2\,\mathrm{e}^{i(\theta_1+\theta_2)}, \qquad \frac{z_1}{z_2} = \frac{r_1}{r_2}\mathrm{e}^{i(\theta_1-\theta_2)}, \qquad z^n = r^n\mathrm{e}^{in\theta} \text{ (De Moivre)}.$$

Raíces n-ésimas de z: $\sqrt[n]{r}\,\mathrm{e}^{i(\theta+2k\pi)/n}$, $k = 0,\dots,n-1$; forman un polígono regular. Fórmula de Euler: $\mathrm{e}^{i\theta} = \cos\theta + i\sin\theta$, y $\mathrm{e}^{i\pi}+1 = 0$.

B.14. Coordenadas polares y curvas paramétricas

$$x = r\cos\theta, \quad y = r\sin\theta; \qquad r = \sqrt{x^2+y^2}, \quad \tan\theta = y/x \text{ (mirando el cuadrante)}.$$

$$\text{cónica desde el foco} : r = \frac{\ell}{1+e\cos\theta}; \qquad \text{cardioide} : r = a(1+\cos\theta); \qquad \text{rosa} : r = \cos(k\theta).$$

Área y longitud en polares: $S = \dfrac{1}{2}\displaystyle\int_\alpha^\beta r^2\,\mathrm{d}\theta$, $L = \displaystyle\int_\alpha^\beta \sqrt{r^2+(r')^2}\,\mathrm{d}\theta$.

En paramétricas $(x(t), y(t))$: $\dfrac{\mathrm{d}y}{\mathrm{d}x} = \dfrac{y'(t)}{x'(t)}$, $L = \displaystyle\int_a^b \sqrt{x'(t)^2+y'(t)^2}\,\mathrm{d}t$.

Tiro parabólico: $x = v_0\cos\alpha\, t$, $y = v_0\sin\alpha\, t - \frac{1}{2}gt^2$; $T = \frac{2v_0\sin\alpha}{g}$, $R = \frac{v_0^2\sin 2\alpha}{g}$, $H = \frac{v_0^2\sin^2\alpha}{2g}$.

Cicloide: $x = R(t-\sin t)$, $y = R(1-\cos t)$; un arco mide $8R$ y encierra $3\pi R^2$.

B.15. Sucesiones, progresiones y matemática financiera

$$\text{aritmética}: a_n = a_1 + (n-1)d, \qquad S_n = \frac{(a_1+a_n)n}{2}$$

$$\text{geométrica}: a_n = a_1 r^{\,n-1}, \qquad S_n = a_1\frac{r^n-1}{r-1}, \quad S_\infty = \frac{a_1}{1-r}\ (|r|<1)$$

$$\sum_{k=1}^{n} k = \frac{n(n+1)}{2}, \qquad \sum_{k=1}^{n} k^2 = \frac{n(n+1)(2n+1)}{6}, \qquad \sum_{k=1}^{n} k^3 = \left(\frac{n(n+1)}{2}\right)^2.$$

Fibonacci: $F_{n+1} = F_n + F_{n-1}$ y $F_n = \dfrac{\phi^n - (1-\phi)^n}{\sqrt{5}}$ con $\phi = \frac{1+\sqrt{5}}{2}$.

Interés simple $I = Crt$; compuesto $C_n = C_0(1+i)^n$; capitalización m veces al año, $C = C_0\left(1+\frac{i}{m}\right)^{mn}$; TAE $= \left(1+\frac{i}{m}\right)^m - 1$.

Valor actual de una renta de n términos: $A = C\dfrac{1-(1+i)^{-n}}{i}$; cuota de un préstamo: $C = A\dfrac{i}{1-(1+i)^{-n}}$.

B.16. Combinatoria

$$\text{variaciones } V_{n,k} = \frac{n!}{(n-k)!}, \qquad \text{con repetición } VR_{n,k} = n^k, \qquad \text{permutaciones } P_n = n!,$$

$$\text{combinaciones } \binom{n}{k} = \frac{n!}{k!(n-k)!}, \qquad \text{con repetición } \binom{n+k-1}{k}, \text{permutaciones con repetición } \frac{n!}{n_1!\cdots n_r!}.$$

Propiedades: $\binom{n}{k} = \binom{n}{n-k}$, $\binom{n}{k} + \binom{n}{k+1} = \binom{n+1}{k+1}$ (Pascal), $\sum_k \binom{n}{k} = 2^n$.

B.17. Probabilidad y distribuciones

$$P(A\cup B) = P(A) + P(B) - P(A\cap B), \qquad P(A\mid B) = \frac{P(A\cap B)}{P(B)}, \qquad P(\bar{A}) = 1 - P(A).$$

Teorema de Bayes: $P(A_i \mid B) = \dfrac{P(B \mid A_i)P(A_i)}{\sum_j P(B \mid A_j)P(A_j)}$.

Distribución	$P(X=k)$ o $f(x)$	$E(X)$	$\mathrm{Var}(X)$
Bernoulli	$p^k q^{1-k}$	p	pq
Binomial $B(n,p)$	$\binom{n}{k}p^k q^{n-k}$	np	npq
Poisson $P(\lambda)$	$\mathrm{e}^{-\lambda}\lambda^k/k!$	λ	λ
Uniforme $U(a,b)$	$1/(b-a)$	$(a+b)/2$	$(b-a)^2/12$
Exponencial	$\lambda \mathrm{e}^{-\lambda x}$	$1/\lambda$	$1/\lambda^2$
Normal $N(\mu,\sigma)$	$\frac{1}{\sigma\sqrt{2\pi}}\mathrm{e}^{-(x-\mu)^2/2\sigma^2}$	μ	σ^2

Tipificación: $Z = \dfrac{X-\mu}{\sigma} \sim N(0,1)$. Regla empírica: en el intervalo $\mu \pm \sigma$ cae el 68 %, en $\mu \pm 2\sigma$ el 95 % y en $\mu \pm 3\sigma$ el 99,7 %.

B.18. Estadística e inferencia

$$\bar{x} = \frac{\sum x_i}{n}, \qquad s^2 = \frac{\sum (x_i - \bar{x})^2}{n} = \overline{x^2} - \bar{x}^2, \qquad CV = \frac{s}{\bar{x}}.$$

Regresión y correlación:

$$r = \frac{s_{xy}}{s_x s_y} \in [-1,1], \qquad y - \bar{y} = \frac{s_{xy}}{s_x^2}(x - \bar{x}), \qquad R^2 = r^2.$$

Media muestral: $\bar{X} \sim N\left(\mu, \frac{\sigma}{\sqrt{n}}\right)$ (teorema central del límite). Intervalos de confianza:

$$\mu \in \bar{x} \pm z_{\alpha/2}\frac{\sigma}{\sqrt{n}}, \qquad p \in \hat{p} \pm z_{\alpha/2}\sqrt{\frac{\hat{p}(1-\hat{p})}{n}}, \qquad n \geq \left(\frac{z_{\alpha/2}\sigma}{E}\right)^2.$$

Valores críticos: $z_{0{,}05} = 1{,}645$ (90 %), $z_{0{,}025} = 1{,}96$ (95 %), $z_{0{,}005} = 2{,}576$ (99 %).

B.19. Aritmética modular y criptografía

$$a \equiv b \pmod{m} \iff m \mid a-b; \qquad \text{si } a \equiv b,\ c \equiv d:\ a \pm c \equiv b \pm d,\ ac \equiv bd.$$

$ax \equiv b \pmod{m}$ tiene solución $\iff d = \mathrm{mcd}(a,m)$ divide a b, y entonces hay d soluciones.
Existe $a^{-1} \iff \mathrm{mcd}(a,m) = 1$.
Criterios: $10 \equiv 1$ $(3,9)$ da la suma de cifras; $10 \equiv -1$ (11) da la suma alternada.
Fermat: $a^{p-1} \equiv 1 \pmod{p}$ si $p \nmid a$. Euler: $a^{\varphi(m)} \equiv 1$ si $\mathrm{mcd}(a,m) = 1$, con $\varphi(p) = p-1$ y $\varphi(pq) = (p-1)(q-1)$.
Teorema chino del resto: si $\mathrm{mcd}(m,n) = 1$, el sistema $x \equiv a$ (m), $x \equiv b$ (n) tiene solución única módulo mn.
RSA: $n = pq$, $ed \equiv 1 \pmod{(p-1)(q-1)}$, $C \equiv M^e$, $M \equiv C^d \pmod{n}$.

B.20. Sistemas de numeración

$$N = (c_k \dots c_1 c_0)_b = \sum_{j=0}^{k} c_j b^j; \qquad \text{de decimal a base } b: \text{ divisiones sucesivas entre } b.$$

Un dígito hexadecimal = 4 bits; 1 byte = 8 bits = 2 dígitos hex; con n bits se representan 2^n valores, de 0 a $2^n - 1$.
Complemento a dos con n bits: $-x$ se codifica como $2^n - x$; el rango es $[-2^{n-1},\, 2^{n-1} - 1]$.
Boole: $A + \bar{A} = 1$, $A\bar{A} = 0$, $\overline{A+B} = \bar{A}\bar{B}$, $\overline{AB} = \bar{A} + \bar{B}$ (De Morgan), $A + AB = A$ (absorción).

B.21. Métodos numéricos

$$\text{bisección : error} \le \frac{b-a}{2^{n+1}} \qquad \text{Newton} : x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)}$$

$$\text{secante} : x_{n+1} = x_n - f(x_n)\frac{x_n - x_{n-1}}{f(x_n) - f(x_{n-1})} \quad \text{punto fijo} : x_{n+1} = g(x_n), \text{ converge si } |g'| < 1$$

$$\int_a^b f \approx \frac{h}{2}(f_0 + 2f_1 + \cdots + 2f_{n-1} + f_n) \text{ (trapecios)}, \qquad h = \frac{b-a}{n},$$

$$\int_a^b f \approx \frac{h}{3}(f_0 + 4f_1 + 2f_2 + \cdots + 4f_{n-1} + f_n) \text{ (Simpson, } n \text{ par)}.$$

Órdenes del error: trapecios $O(h^2)$, Simpson $O(h^4)$; Newton converge cuadráticamente, la secante con orden $\phi \approx 1{,}618$.

B.22. Grafos

$$\sum_v \deg(v) = 2|A|; \qquad \text{árbol} : |A| = |V| - 1; \qquad \text{planar conexo} : |V| - |A| + |C| = 2 \text{ (Euler)}.$$

Camino euleriano $\iff$ hay 0 o 2 vértices de grado impar; circuito euleriano $\iff$ todos los grados son pares. En un grafo planar simple con $|V| \ge 3$: $|A| \le 3|V| - 6$. Todo grafo planar es 4-coloreable.

B.23. Álgebra abstracta

Lagrange: si $H \le G$ finito, $|G| = |H| \cdot [G : H]$, y el orden de todo elemento divide a $|G|$.
Órbita–estabilizador: $|G| = |\mathcal{O}(x)| \cdot |G_x|$.
Isomorfía: $G/\ker f \cong \operatorname{im} f$; para anillos, $A/\ker f \cong \operatorname{im} f$.
$\mathbb{Z}/m\mathbb{Z}$ es cuerpo $\iff$ m es primo; $|\mathbb{F}_{p^n}| = p^n$ y su grupo multiplicativo es cíclico.
Extensiones: $[L : K] = [L : F][F : K]$ (torre); $[K(\alpha) : K] = \deg m_\alpha$.
Galois: $|\operatorname{Gal}(L/K)| = [L : K]$ si la extensión es de Galois, y la correspondencia entre subgrupos y cuerpos intermedios invierte inclusiones. Un polinomio es resoluble por radicales $\iff$ su grupo de Galois es resoluble.

B.24. Desigualdades y estimaciones útiles

$$|x+y| \le |x| + |y|, \qquad \big|\langle x, y\rangle\big| \le \|x\|\,\|y\| \text{ (Cauchy–Schwarz)}, \qquad \|x+y\| \le \|x\| + \|y\|.$$

Medias: $\dfrac{n}{\sum 1/a_i} \le \sqrt[n]{a_1 \cdots a_n} \le \dfrac{a_1 + \cdots + a_n}{n} \le \sqrt{\dfrac{\sum a_i^2}{n}}$ (armónica $\le$ geométrica $\le$ aritmética $\le$ cuadrática).

Bernoulli: $(1+x)^n \ge 1+nx$ para $x \ge -1$. Taylor con resto de Lagrange: $f(x) = \sum_{k=0}^{n} \frac{f^{(k)}(a)}{k!}(x-a)^k + \frac{f^{(n+1)}(\xi)}{(n+1)!}(x-a)^{n+1}$.

B.25. Constantes

$$\pi = 3{,}14159265\ldots \quad e = 2{,}71828182\ldots \quad \gamma = 0{,}57721566\ldots \quad \varphi = \frac{1+\sqrt{5}}{2} = 1{,}61803398\ldots$$

$$\ln 2 = 0{,}693147\ldots \quad \ln 10 = 2{,}302585\ldots \quad \sqrt{2} = 1{,}414213\ldots \quad \sqrt{3} = 1{,}732050\ldots$$

Apéndice C

Tablas estadísticas

Ninguna de las tablas que siguen es un adorno histórico: sin ellas no se puede terminar un solo ejercicio de los capítulos de inferencia. Todos los valores están calculados con quince cifras y redondeados a la última que se muestra, de modo que pueden usarse con confianza en cualquier contraste del libro.

C.1. Normal estándar: función de distribución $\Phi(z)$

La tabla da $\Phi(z) = P(Z \le z)$ para $Z \sim N(0,1)$. Para valores negativos, úsese $\Phi(-z) = 1 - \Phi(z)$. La fila indica la parte entera y la primera decimal; la columna, la segunda.

z	.00	.01	.02	.03	.04	.05	.06	.07	.08	.09
0.0	.5000	.5040	.5080	.5120	.5160	.5199	.5239	.5279	.5319	.5359
0.1	.5398	.5438	.5478	.5517	.5557	.5596	.5636	.5675	.5714	.5753
0.2	.5793	.5832	.5871	.5910	.5948	.5987	.6026	.6064	.6103	.6141
0.3	.6179	.6217	.6255	.6293	.6331	.6368	.6406	.6443	.6480	.6517
0.4	.6554	.6591	.6628	.6664	.6700	.6736	.6772	.6808	.6844	.6879
0.5	.6915	.6950	.6985	.7019	.7054	.7088	.7123	.7157	.7190	.7224
0.6	.7257	.7291	.7324	.7357	.7389	.7422	.7454	.7486	.7517	.7549
0.7	.7580	.7611	.7642	.7673	.7704	.7734	.7764	.7794	.7823	.7852
0.8	.7881	.7910	.7939	.7967	.7995	.8023	.8051	.8078	.8106	.8133
0.9	.8159	.8186	.8212	.8238	.8264	.8289	.8315	.8340	.8365	.8389
1.0	.8413	.8438	.8461	.8485	.8508	.8531	.8554	.8577	.8599	.8621
1.1	.8643	.8665	.8686	.8708	.8729	.8749	.8770	.8790	.8810	.8830
1.2	.8849	.8869	.8888	.8907	.8925	.8944	.8962	.8980	.8997	.9015
1.3	.9032	.9049	.9066	.9082	.9099	.9115	.9131	.9147	.9162	.9177
1.4	.9192	.9207	.9222	.9236	.9251	.9265	.9279	.9292	.9306	.9319
1.5	.9332	.9345	.9357	.9370	.9382	.9394	.9406	.9418	.9429	.9441
1.6	.9452	.9463	.9474	.9484	.9495	.9505	.9515	.9525	.9535	.9545
1.7	.9554	.9564	.9573	.9582	.9591	.9599	.9608	.9616	.9625	.9633
1.8	.9641	.9649	.9656	.9664	.9671	.9678	.9686	.9693	.9699	.9706
1.9	.9713	.9719	.9726	.9732	.9738	.9744	.9750	.9756	.9761	.9767
2.0	.9772	.9778	.9783	.9788	.9793	.9798	.9803	.9808	.9812	.9817
2.1	.9821	.9826	.9830	.9834	.9838	.9842	.9846	.9850	.9854	.9857
2.2	.9861	.9864	.9868	.9871	.9875	.9878	.9881	.9884	.9887	.9890
2.3	.9893	.9896	.9898	.9901	.9904	.9906	.9909	.9911	.9913	.9916
2.4	.9918	.9920	.9922	.9925	.9927	.9929	.9931	.9932	.9934	.9936
2.5	.9938	.9940	.9941	.9943	.9945	.9946	.9948	.9949	.9951	.9952
2.6	.9953	.9955	.9956	.9957	.9959	.9960	.9961	.9962	.9963	.9964
2.7	.9965	.9966	.9967	.9968	.9969	.9970	.9971	.9972	.9973	.9974
2.8	.9974	.9975	.9976	.9977	.9977	.9978	.9979	.9979	.9980	.9981
2.9	.9981	.9982	.9982	.9983	.9984	.9984	.9985	.9985	.9986	.9986
3.0	.9987	.9987	.9987	.9988	.9988	.9989	.9989	.9989	.9990	.9990
3.1	.9990	.9991	.9991	.9991	.9992	.9992	.9992	.9992	.9993	.9993
3.2	.9993	.9993	.9994	.9994	.9994	.9994	.9994	.9995	.9995	.9995
3.3	.9995	.9995	.9995	.9996	.9996	.9996	.9996	.9996	.9996	.9997
3.4	.9997	.9997	.9997	.9997	.9997	.9997	.9997	.9997	.9997	.9998

Cuantiles usados con más frecuencia:

p	0,90	0,95	0,975	0,99	0,995	0,999	0,9995
z_p	1,2816	1,6449	1,9600	2,3263	2,5758	3,0902	3,2905

Observación C.1. Los dos que hay que saber de memoria son $z_{0,975} = 1{,}9600$ (intervalos al 95 %) y $z_{0,995} = 2{,}5758$ (al 99 %). El célebre "$\bar{x} \pm 2s/\sqrt{n}$" usa 2 en lugar de 1,96, lo que corresponde en realidad a una confianza del 95,45 %: una aproximación excelente y consciente, no un error.

C.2. t de Student: cuantiles $t_{\nu,p}$

Valor t tal que $P(T \le t) = p$ para $T \sim t_\nu$. Para un intervalo bilateral de confianza $1-\alpha$ úsese la columna $p = 1-\alpha/2$.

ν	$p=0{,}90$	0,95	0,975	0,99	0,995	0,999
1	3,0777	6,3138	12,7062	31,8205	63,6567	318,3088
2	1,8856	2,9200	4,3027	6,9646	9,9248	22,3271
3	1,6377	2,3534	3,1824	4,5407	5,8409	10,2145
4	1,5332	2,1318	2,7764	3,7469	4,6041	7,1732
5	1,4759	2,0150	2,5706	3,3649	4,0321	5,8934
6	1,4398	1,9432	2,4469	3,1427	3,7074	5,2076
7	1,4149	1,8946	2,3646	2,9980	3,4995	4,7853
8	1,3968	1,8595	2,3060	2,8965	3,3554	4,5008
9	1,3830	1,8331	2,2622	2,8214	3,2498	4,2968
10	1,3722	1,8125	2,2281	2,7638	3,1693	4,1437
11	1,3634	1,7959	2,2010	2,7181	3,1058	4,0247
12	1,3562	1,7823	2,1788	2,6810	3,0545	3,9296
13	1,3502	1,7709	2,1604	2,6503	3,0123	3,8520
14	1,3450	1,7613	2,1448	2,6245	2,9768	3,7874
15	1,3406	1,7531	2,1314	2,6025	2,9467	3,7328
16	1,3368	1,7459	2,1199	2,5835	2,9208	3,6862
17	1,3334	1,7396	2,1098	2,5669	2,8982	3,6458
18	1,3304	1,7341	2,1009	2,5524	2,8784	3,6105
19	1,3277	1,7291	2,0930	2,5395	2,8609	3,5794
20	1,3253	1,7247	2,0860	2,5280	2,8453	3,5518
21	1,3232	1,7207	2,0796	2,5176	2,8314	3,5272
22	1,3212	1,7171	2,0739	2,5083	2,8188	3,5050
23	1,3195	1,7139	2,0687	2,4999	2,8073	3,4850
24	1,3178	1,7109	2,0639	2,4922	2,7969	3,4668
25	1,3163	1,7081	2,0595	2,4851	2,7874	3,4502
26	1,3150	1,7056	2,0555	2,4786	2,7787	3,4350
27	1,3137	1,7033	2,0518	2,4727	2,7707	3,4210
28	1,3125	1,7011	2,0484	2,4671	2,7633	3,4082
29	1,3114	1,6991	2,0452	2,4620	2,7564	3,3962
30	1,3104	1,6973	2,0423	2,4573	2,7500	3,3852
35	1,3062	1,6896	2,0301	2,4377	2,7238	3,3400
40	1,3031	1,6839	2,0211	2,4233	2,7045	3,3069
45	1,3006	1,6794	2,0141	2,4121	2,6896	3,2815
50	1,2987	1,6759	2,0086	2,4033	2,6778	3,2614
60	1,2958	1,6706	2,0003	2,3901	2,6603	3,2317
80	1,2922	1,6641	1,9901	2,3739	2,6387	3,1953
100	1,2901	1,6602	1,9840	2,3642	2,6259	3,1737
120	1,2886	1,6577	1,9799	2,3578	2,6174	3,1595
∞	1,2816	1,6449	1,9600	2,3263	2,5758	3,0902

Observación C.2. La última fila es la normal: la t converge a ella, y a partir de $\nu \approx 30$ la diferencia en el cuantil del 97,5 % es menor que 0,07. Obsérvese en cambio lo brutal que es

la primera fila: con un solo grado de libertad, el cuantil al 97,5 % vale 12,7062 en lugar de 1,9600. Ese es el precio de estimar σ con dos observaciones.

C.3. Chi cuadrado: cuantiles $\chi^2_{\nu,p}$

Valor q tal que $P(Q \le q) = p$ para $Q \sim \chi^2_\nu$. Para un intervalo bilateral de σ^2 se usan las columnas $p = \alpha/2$ y $p = 1 - \alpha/2$.

ν	$p = 0{,}005$	0,025	0,05	0,50	0,95	0,975	0,995
1	0,000	0,001	0,004	0,455	3,841	5,024	7,879
2	0,010	0,051	0,103	1,386	5,991	7,378	10,597
3	0,072	0,216	0,352	2,366	7,815	9,348	12,838
4	0,207	0,484	0,711	3,357	9,488	11,143	14,860
5	0,412	0,831	1,145	4,351	11,070	12,833	16,750
6	0,676	1,237	1,635	5,348	12,592	14,449	18,548
7	0,989	1,690	2,167	6,346	14,067	16,013	20,278
8	1,344	2,180	2,733	7,344	15,507	17,535	21,955
9	1,735	2,700	3,325	8,343	16,919	19,023	23,589
10	2,156	3,247	3,940	9,342	18,307	20,483	25,188
11	2,603	3,816	4,575	10,341	19,675	21,920	26,757
12	3,074	4,404	5,226	11,340	21,026	23,337	28,300
13	3,565	5,009	5,892	12,340	22,362	24,736	29,819
14	4,075	5,629	6,571	13,339	23,685	26,119	31,319
15	4,601	6,262	7,261	14,339	24,996	27,488	32,801
16	5,142	6,908	7,962	15,338	26,296	28,845	34,267
17	5,697	7,564	8,672	16,338	27,587	30,191	35,718
18	6,265	8,231	9,390	17,338	28,869	31,526	37,156
19	6,844	8,907	10,117	18,338	30,144	32,852	38,582
20	7,434	9,591	10,851	19,337	31,410	34,170	39,997
21	8,034	10,283	11,591	20,337	32,671	35,479	41,401
22	8,643	10,982	12,338	21,337	33,924	36,781	42,796
23	9,260	11,689	13,091	22,337	35,172	38,076	44,181
24	9,886	12,401	13,848	23,337	36,415	39,364	45,559
25	10,520	13,120	14,611	24,337	37,652	40,646	46,928
26	11,160	13,844	15,379	25,336	38,885	41,923	48,290
27	11,808	14,573	16,151	26,336	40,113	43,195	49,645
28	12,461	15,308	16,928	27,336	41,337	44,461	50,993
29	13,121	16,047	17,708	28,336	42,557	45,722	52,336
30	13,787	16,791	18,493	29,336	43,773	46,979	53,672
40	20,707	24,433	26,509	39,335	55,758	59,342	66,766
50	27,991	32,357	34,764	49,335	67,505	71,420	79,490
60	35,534	40,482	43,188	59,335	79,082	83,298	91,952
80	51,172	57,153	60,391	79,334	101,879	106,629	116,321
100	67,328	74,222	77,929	99,334	124,342	129,561	140,169

Observación C.3. Para $\nu > 100$ puede usarse la aproximación de Wilson–Hilferty, $\chi^2_{\nu,p} \approx \nu\big(1 - \frac{2}{9\nu} + z_p\sqrt{\frac{2}{9\nu}}\big)^3$, que es sorprendentemente precisa: para $\nu = 100$ y $p = 0{,}95$ da 124,340 frente al valor exacto 124,342: dos milésimas de error.

C.4. F de Fisher: cuantiles al 95 % y al 99 %

Valor f con $P(F \leq f) = 0{,}95$ (tabla superior) y $= 0{,}99$ (inferior), para $F \sim F_{m,n}$; m son los grados de libertad del numerador y n los del denominador. Para la cola izquierda, $F_{m,n;\alpha} = 1/F_{n,m;1-\alpha}$.

$p = 0{,}95$

$n\backslash m$	1	2	3	4	5	6	8	10	15	20
1	161,45	199,50	215,71	224,58	230,16	233,99	238,88	241,88	245,95	248,01
2	18,51	19,00	19,16	19,25	19,30	19,33	19,37	19,40	19,43	19,45
3	10,13	9,55	9,28	9,12	9,01	8,94	8,85	8,79	8,70	8,66
4	7,71	6,94	6,59	6,39	6,26	6,16	6,04	5,96	5,86	5,80
5	6,61	5,79	5,41	5,19	5,05	4,95	4,82	4,74	4,62	4,56
6	5,99	5,14	4,76	4,53	4,39	4,28	4,15	4,06	3,94	3,87
7	5,59	4,74	4,35	4,12	3,97	3,87	3,73	3,64	3,51	3,44
8	5,32	4,46	4,07	3,84	3,69	3,58	3,44	3,35	3,22	3,15
9	5,12	4,26	3,86	3,63	3,48	3,37	3,23	3,14	3,01	2,94
10	4,96	4,10	3,71	3,48	3,33	3,22	3,07	2,98	2,85	2,77
12	4,75	3,89	3,49	3,26	3,11	3,00	2,85	2,75	2,62	2,54
15	4,54	3,68	3,29	3,06	2,90	2,79	2,64	2,54	2,40	2,33
20	4,35	3,49	3,10	2,87	2,71	2,60	2,45	2,35	2,20	2,12
25	4,24	3,39	2,99	2,76	2,60	2,49	2,34	2,24	2,09	2,01
30	4,17	3,32	2,92	2,69	2,53	2,42	2,27	2,16	2,01	1,93
40	4,08	3,23	2,84	2,61	2,45	2,34	2,18	2,08	1,92	1,84
60	4,00	3,15	2,76	2,53	2,37	2,25	2,10	1,99	1,84	1,75
120	3,92	3,07	2,68	2,45	2,29	2,18	2,02	1,91	1,75	1,66

$p = 0{,}99$

$n\backslash m$	1	2	3	4	5	6	8	10	15	20
1	4052,18	4999,50	5403,35	5624,58	5763,65	5858,99	5981,07	6055,85	6157,28	6208,73
2	98,50	99,00	99,17	99,25	99,30	99,33	99,37	99,40	99,43	99,45
3	34,12	30,82	29,46	28,71	28,24	27,91	27,49	27,23	26,87	26,69
4	21,20	18,00	16,69	15,98	15,52	15,21	14,80	14,55	14,20	14,02
5	16,26	13,27	12,06	11,39	10,97	10,67	10,29	10,05	9,72	9,55
6	13,75	10,92	9,78	9,15	8,75	8,47	8,10	7,87	7,56	7,40
7	12,25	9,55	8,45	7,85	7,46	7,19	6,84	6,62	6,31	6,16
8	11,26	8,65	7,59	7,01	6,63	6,37	6,03	5,81	5,52	5,36
9	10,56	8,02	6,99	6,42	6,06	5,80	5,47	5,26	4,96	4,81
10	10,04	7,56	6,55	5,99	5,64	5,39	5,06	4,85	4,56	4,41
12	9,33	6,93	5,95	5,41	5,06	4,82	4,50	4,30	4,01	3,86
15	8,68	6,36	5,42	4,89	4,56	4,32	4,00	3,80	3,52	3,37
20	8,10	5,85	4,94	4,43	4,10	3,87	3,56	3,37	3,09	2,94
25	7,77	5,57	4,68	4,18	3,85	3,63	3,32	3,13	2,85	2,70
30	7,56	5,39	4,51	4,02	3,70	3,47	3,17	2,98	2,70	2,55
40	7,31	5,18	4,31	3,83	3,51	3,29	2,99	2,80	2,52	2,37
60	7,08	4,98	4,13	3,65	3,34	3,12	2,82	2,63	2,35	2,20
120	6,85	4,79	3,95	3,48	3,17	2,96	2,66	2,47	2,19	2,03

C.5. Rango studentizado $q_{\alpha;k,\nu}$ para el método de Tukey

Valor crítico del rango studentizado con k grupos y ν grados de libertad del error, al nivel $\alpha = 0{,}05$. Es la tabla que hace falta para las comparaciones múltiples del capítulo de modelos lineales, donde se declara significativa una diferencia si $|\bar{y}_i - \bar{y}_j| > q_{\alpha;k,\nu}\sqrt{s^2/n}$.

$\nu \backslash k$	2	3	4	5	6	8	10
5	3,6354	4,6017	5,2183	5,6731	6,0329	6,5823	6,9947
6	3,4605	4,3392	4,8956	5,3049	5,6284	6,1222	6,4931
7	3,3441	4,1649	4,6813	5,0601	5,3591	5,8153	6,1579
8	3,2612	4,0410	4,5288	4,8858	5,1672	5,5962	5,9183
9	3,1992	3,9485	4,4149	4,7554	5,0235	5,4319	5,7384
10	3,1511	3,8768	4,3266	4,6543	4,9120	5,3042	5,5984
12	3,0813	3,7729	4,1987	4,5077	4,7502	5,1187	5,3946
14	3,0332	3,7014	4,1105	4,4066	4,6385	4,9903	5,2534
16	2,9980	3,6491	4,0461	4,3327	4,5568	4,8962	5,1498
18	2,9712	3,6093	3,9970	4,2763	4,4944	4,8243	5,0705
20	2,9500	3,5779	3,9583	4,2319	4,4452	4,7676	5,0079
24	2,9188	3,5317	3,9013	4,1663	4,3727	4,6838	4,9152
30	2,8882	3,4864	3,8454	4,1021	4,3015	4,6014	4,8241
40	2,8582	3,4421	3,7907	4,0391	4,2316	4,5205	4,7345
60	2,8288	3,3987	3,7371	3,9774	4,1632	4,4411	4,6463
120	2,8000	3,3561	3,6846	3,9169	4,0960	4,3630	4,5595

Observación C.4. Compárese la primera columna con la t: $q_{0{,}05;2,\nu} = \sqrt{2}\, t_{\nu;0{,}975}$, exactamente, porque con dos grupos el rango es la única diferencia y la fórmula se reduce a la t de dos muestras. A partir de ahí q crece con k, y esa es la corrección por comparaciones múltiples: con $k = 10$ grupos y $\nu = 20$, se tiene $q = 5{,}0079$, y la diferencia exigida es $5{,}0079/\sqrt{2} = 3{,}54$ veces la desviación típica de la diferencia, frente a las 2,09 que exigiría una t ingenua: un umbral un 69 % más alto. Usar t en lugar de q es el error que produce la mayoría de los "hallazgos" que luego no se replican.

C.6. Valores críticos de Kolmogórov–Smirnov y de Wilcoxon

Kolmogórov–Smirnov (una muestra, F completamente especificada): rechazar si $D_n > d_{n,\alpha}$.

n	5	10	15	20	25	30	> 35
$\alpha = 0{,}1$	0,5088	0,3689	0,3043	0,2651	0,2380	0,2179	$1{,}2238/\sqrt{n}$
$\alpha = 0{,}05$	0,5646	0,4094	0,3377	0,2942	0,2641	0,2418	$1{,}3581/\sqrt{n}$
$\alpha = 0{,}01$	0,6767	0,4907	0,4047	0,3525	0,3165	0,2897	$1{,}6276/\sqrt{n}$

Wilcoxon–Mann–Whitney: valor crítico u_α de la cola izquierda al nivel $\alpha = 0{,}05$ (una cola); rechazar si $U \le u_\alpha$.

$n_1 \backslash n_2$	3	4	5	6	7	8	9	10
3	0	0	1	2	2	3	4	4
4	0	1	2	3	4	5	6	7
5	1	2	4	5	6	8	9	11
6	2	3	5	7	8	10	12	14
7	2	4	6	8	11	13	15	17
8	3	5	8	10	13	15	18	20
9	4	6	9	12	15	18	21	24
10	4	7	11	14	17	20	24	27

Observación C.5. Las casillas con raya indican que con esos tamaños *no existe* ningún valor crítico al 5 %: incluso el resultado más extremo posible tiene probabilidad mayor que 0,05 bajo la hipótesis nula. Con $n_1 = n_2 = 3$, por ejemplo, $P(U = 0) = 1/20 = 0{,}05$, justo en el límite. Es la manifestación más clara de que los contrastes de permutación tienen una resolución mínima impuesta por el tamaño muestral, y de que no hay forma de esquivarla.

Apéndice D

Glosario de símbolos

La notación matemática es una lengua que casi nadie enseña explícitamente: se aprende por contagio, y por eso deja huecos. Esta lista reúne los símbolos que aparecen en el libro, con la lectura en voz alta y el capítulo donde entran en escena.

D.1. Conjuntos y lógica

$\in$, $\notin$ “pertenece a”, “no pertenece a”.

$\subset$, $\subseteq$ “está contenido en”. En este libro $\subset$ no excluye la igualdad.

$\cup$, $\cap$ unión e intersección. $\bigcup$, $\bigcap$ para familias.

$\setminus$ diferencia de conjuntos: $A \setminus B$ es “lo de A que no está en B”.

$\varnothing$ conjunto vacío.

$\times$ producto cartesiano: $A \times B$ son los pares ordenados.

$\forall$, $\exists$ “para todo”, “existe”. $\exists!$ es “existe un único”.

$\Rightarrow$, $\Leftrightarrow$ “implica”, “si y solo si”.

$\mathcal{P}(A)$ conjunto de partes de A.

$|A|$, $\#A$ cardinal. $\aleph_0$ es el cardinal de $\mathbb{N}$ y $\mathfrak{c}$ el de $\mathbb{R}$ (capítulo 29).

D.2. Números

$\mathbb{N}$, $\mathbb{Z}$, $\mathbb{Q}$, $\mathbb{R}$, $\mathbb{C}$ naturales, enteros, racionales, reales, complejos.

$\mathbb{F}_q$ cuerpo finito de q elementos (capítulo 25).

$\mathbb{Z}/n\mathbb{Z}$ enteros módulo n (capítulo 3).

$a \mid b$, $a \nmid b$ “a divide a b”, “no divide”.

gcd, mcm
: máximo común divisor y mínimo común múltiplo.

$a \equiv b \pmod{n}$
: congruencia: n divide a $a - b$.

$\varphi(n)$ función de Euler: cuántos enteros menores que n son coprimos con n.

$\left(\frac{a}{p}\right)$ símbolo de Legendre.

$|z|$, $\bar{z}$ módulo y conjugado de un complejo (capítulo 30).

$\lfloor x \rfloor$, $\lceil x \rceil$ parte entera por defecto y por exceso.

D.3. Análisis

$\to$, $\mapsto$ "tiende a" o "va de"; $\mapsto$ es "se manda a" (nivel de elementos).

sup, ínf supremo e ínfimo: la menor cota superior y la mayor cota inferior.

lím sup, lím inf
: límite superior e inferior de una sucesión.

f', $\frac{\mathrm{d}f}{\mathrm{d}x}$, $\frac{\partial f}{\partial x}$
: derivada, derivada total, derivada parcial.

∇f, grad f
: gradiente. $\nabla \cdot F = \operatorname{div} F$ divergencia, $\nabla \times F = \operatorname{rot} F$ rotacional.

Δ, ∇^2 laplaciano.

$\mathrm{d}x$ diferencial. En el capítulo 36 pasa a ser una 1-forma con derecho propio.

$\int$, $\oint$, $\iint$ integral, integral sobre una curva cerrada, integral doble.

$\sum$, $\prod$ suma y producto de una familia indexada.

$\mathcal{O}$, o notación de Landau: $f = \mathcal{O}(g)$ es "no crece más rápido que".

c.t.p. "en casi todo punto": salvo en un conjunto de medida nula (capítulo 31).

μ, λ medida en general; λ suele reservarse a la de Lebesgue.

L^p, ℓ^p espacios de funciones y de sucesiones p-integrables (capítulo 32).

$\|\cdot\|$, $\langle \cdot, \cdot \rangle$
: norma y producto interno.

X^* espacio dual: los funcionales lineales continuos sobre X.

D.4. Álgebra

$\cong$ isomorfo. $\simeq$ se reserva a homotopía o equivalencia más débil.

$\leq$ **entre grupos**
: "es subgrupo de". $N \trianglelefteq G$ o $N \lhd G$: subgrupo normal.

G/N grupo cociente. R/I anillo cociente (capítulos 23 y 24).

$[G : H]$ índice: número de clases laterales.

ker, im núcleo e imagen de un homomorfismo.

$\langle a \rangle$ subgrupo (o ideal) generado por a.

S_n, A_n, D_n, C_n
simétrico, alternante, diédrico y cíclico.

$\oplus$, $\otimes$ suma directa y producto tensorial.

$R[x]$ anillo de polinomios con coeficientes en R.

$[K : F]$ grado de la extensión de cuerpos K/F (capítulo 25).

$\mathrm{Gal}(K/F)$
grupo de Galois (capítulo 26).

$\operatorname{car} R$ característica de un anillo o cuerpo.

rango, nulidad, tr, det
rango, nulidad, traza y determinante.

D.5. Geometría y topología

$B(x, r)$ bola abierta de centro x y radio r.

$\overline{A}$, A°, ∂A
clausura, interior y frontera (capítulo 28).

τ una topología, es decir la familia de abiertos.

S^n, T^2, $\mathbb{R}P^n$
esfera n-dimensional, toro, espacio proyectivo real.

$\pi_1(X)$ grupo fundamental.

g_{ij}, g^{ij} métrica y su inversa (capítulo 36).

Γ^i_{jk} símbolos de Christoffel.

$R^i{}_{jkl}$, R_{ij}, S
tensor de Riemann, de Ricci y curvatura escalar.

δ^i_j, ε_{ijk} delta de Kronecker y símbolo de Levi-Civita.

D.6. Notación escolar de uso diario

Los símbolos de esta sección son los primeros que uno encuentra y, precisamente por eso, los que nadie explica nunca del todo.

$=$, $\neq$ igual y distinto. Conviene distinguir el "=" de identidad (cierto para todo valor: $(a+b)^2 = a^2 + 2ab + b^2$) del "=" de ecuación (cierto solo para algunos valores: $x^2 = 4$).

$\approx$, $\sim$ “aproximadamente igual” y “del mismo orden”. $\pi \approx 3{,}1416$; en cambio $a_n \sim b_n$ significa que el cociente tiende a 1.

$\equiv$ “idénticamente igual” o, en el capítulo 3, “congruente módulo”. El contexto decide, y siempre es visible cuál de los dos es.

$\propto$ “proporcional a”.

$<$, $\leq$, $\ll$ menor, menor o igual, mucho menor.

$\pm$, $\mp$ “más o menos”. En $x = \frac{-b \pm \sqrt{\Delta}}{2a}$ no es una ambigüedad: son las dos soluciones a la vez.

$\sqrt{\ }$, $\sqrt[n]{\ }$ raíz cuadrada y raíz n-ésima. $\sqrt{x^2} = |x|$, no x (el álgebra elemental).

a^n, a^{-n}, $a^{1/n}$ potencia, inverso de la potencia y raíz.

$|x|$ valor absoluto; también módulo de un vector o de un complejo, y cardinal de un conjunto. Tres significados, un símbolo: se distinguen por lo que hay dentro.

% tanto por ciento. Un aumento del 20 % multiplica por 1,20 (la matemática financiera).

∞ infinito. *No* es un número: es una forma de decir “sin cota”.

Δx incremento o variación de x. En el capítulo 6 es el germen de $\mathrm{d}x$.

$\overline{AB}$, $\widehat{ABC}$ segmento (o su longitud) y ángulo de vértice B.

$\parallel$, $\perp$ paralelo y perpendicular.

$\triangle ABC$ triángulo de vértices A, B y C.

$\square$, $\checkmark$ fin de una comprobación rutinaria; visto bueno tras verificar una cuenta.

D.7. Trigonometría, exponenciales y logaritmos

sin, cos, tan seno, coseno y tangente. En textos españoles antiguos aparece “sen”; aquí se usa siempre sin (la trigonometría elemental).

csc, sec, cot cosecante, secante y cotangente: los inversos de los tres anteriores.

arcsin, arc cos, arctan funciones inversas, con sus dominios restringidos. $\sin^{-1} x$ significa $\arcsin x$ y *no* $1/\sin x$: es la ambigüedad más peligrosa de toda la notación elemental.

$\sin^2 x$ en cambio sí significa $(\sin x)^2$. La convención es inconsistente y hay que vivir con ella.

sinh, cosh, tanh funciones hiperbólicas, definidas con e^x en el capítulo 6.

°, **rad** grados y radianes. $180° = \pi$ rad. Las fórmulas del cálculo solo valen en radianes.

e — la constante 2,71828…, base del logaritmo natural.

ln, log, $\log_a$ — logaritmo natural, decimal y en base a. Cuidado: en muchos textos de análisis log significa ln, y en informática significa $\log_2$ (los logaritmos).

D.8. Combinatoria, probabilidad y estadística

$n!$ — factorial. $0! = 1$, por convenio útil y no arbitrario (capítulo 5).

$\binom{n}{k}$ — número combinatorio, "n sobre k". También se escribe $C_{n,k}$ o $\binom{n}{k} = \frac{n!}{k!(n-k)!}$.

$V_{n,k}$, $VR_{n,k}$ — variaciones sin y con repetición.

$P(A)$ — probabilidad del suceso A. $P(A \mid B)$ es la probabilidad condicionada, "A sabiendo B".

A^c, $\bar{A}$ — suceso complementario.

X, x — la mayúscula es la variable aleatoria; la minúscula, un valor concreto que toma. La distinción parece pedante y es la clave de todo el capítulo 16.

$\mathbb{E}(X)$, $\mathrm{Var}(X)$ — esperanza y varianza. $\sigma = \sqrt{\mathrm{Var}(X)}$ es la desviación típica.

$B(n,p)$, $\mathrm{Po}(\lambda)$, $N(\mu,\sigma)$ — binomial, Poisson y normal.

$\bar{x}$, s — media y desviación típica *muestrales*, frente a μ y σ *poblacionales*. Confundirlas es el error clásico del capítulo 16.

z_α — valor crítico de la normal estándar: deja a su derecha una probabilidad α.

H_0, H_1 — hipótesis nula y alternativa.

r, R^2 — coeficiente de correlación y coeficiente de determinación.

D.9. Sucesiones, matrices y estructuras discretas

a_n, (a_n) — término general y la sucesión entera. Los paréntesis marcan la diferencia entre un número y una lista infinita.

$\sum_{i=1}^{n}$, $\sum_{i=1}^{\infty}$ — suma finita y serie.

d, r — diferencia de una progresión aritmética y razón de una geométrica (las sucesiones y series elementales).

$A = (a_{ij})$ — matriz y su elemento genérico: primero fila, después columna. Siempre en ese orden.

A^T, A^{-1} — traspuesta e inversa. A^{-1} no es "uno partido por A".

I_n, O matriz identidad de orden n y matriz nula.

$\det A$, $|A|$
determinante. La segunda notación choca con el valor absoluto, y por eso aquí se prefiere la primera (el cálculo matricial elemental).

$G = (V, A)$
grafo con conjunto de vértices V y de aristas A; $\deg(v)$ es el grado del vértice v (capítulo 8).

K_n, $K_{m,n}$
grafo completo y grafo bipartito completo.

(r, θ) coordenadas polares, frente a (x, y) cartesianas (capítulo 10).

$\vec{u}$, $\mathbf{u}$ vector. La flecha es la costumbre escolar; la negrita, la universitaria. Significan lo mismo.

$\vec{u} \cdot \vec{v}$, $\vec{u} \times \vec{v}$
producto escalar (da un número) y producto vectorial (da un vector). El punto y el aspa no son intercambiables.

D.10. El alfabeto griego

Los griegos no son decoración: cada letra arrastra un uso tradicional, y respetarlo hace las fórmulas legibles de un vistazo.

Letra	Nombre	Uso habitual en este libro
α, β, γ	alfa, beta, gamma	ángulos; exponentes; escalares
δ, Δ	delta	cantidad pequeña; incremento; discriminante $\Delta = b^2 - 4ac$
ε	épsilon	el "tan pequeño como se quiera" del análisis
θ, φ	theta, fi	ángulos; φ también es la función de Euler y el número áureo
λ	lambda	autovalor; parámetro; medida de Lebesgue; media de Poisson
μ	mu	media poblacional; medida
π	pi	la constante; también $\pi_1(X)$, grupo fundamental
ρ	ro	densidad; radio; radio espectral
σ, Σ	sigma	desviación típica; Σ es la suma; σ-álgebra
τ	tau	una topología; tiempo propio
ϕ, Φ	fi	función; Φ es la distribución normal acumulada
ω, Ω	omega	frecuencia; Ω es el espacio muestral

D.11. Abreviaturas y expresiones latinas

s.p.d.g. "sin pérdida de generalidad": el caso que se trata basta porque los otros se reducen a él por simetría. Es honesto solo si se puede explicar *cómo* se reducen.

c.q.d., q.e.d.
"como queríamos demostrar", *quod erat demonstrandum*. Aquí se sustituyen por ■.

ssi "si y solo si", el $\iff$ escrito en palabras.

i.e., e.g.
id est ("es decir") y *exempli gratia* ("por ejemplo"). Se confunden constantemente.

c.t.p. "en casi todo punto" (capítulo 31).

mcd, mcm
máximo común divisor y mínimo común múltiplo.

EDO, EDP
ecuación diferencial ordinaria y en derivadas parciales (capítulo 12).

TVM, TFC
teorema del valor medio y teorema fundamental del cálculo.

$\therefore$, $\because$ "por tanto" y "porque". Se usan poco en texto impreso y mucho en pizarra.

D.12. Convenios de este libro

Se usa la coma decimal, como es costumbre en español: 3,14 y no 3.14. Los intervalos abiertos se escriben con paréntesis, (a, b), y no con corchetes invertidos. El símbolo ■ cierra una demostración. Las cajas azules marcan ideas clave, las rojas advertencias, las verdes pausas para pensar y las moradas reglas operativas; las grises, al final de cada capítulo, contienen las respuestas de los ejercicios propuestos. Cada capítulo lleva una etiqueta simbólica, de modo que las referencias cruzadas siguen siendo correctas aunque el orden de los capítulos cambie en futuras ediciones.

Apéndice E

Soluciones a los ejercicios propuestos

Lo que sigue es el resultado de cada ejercicio y, cuando hace falta, una indicación de una línea sobre por dónde atacarlo. No es la resolución completa, y eso es deliberado: el valor de un ejercicio está en el rato que uno pasa atascado, y una solución detallada se lo roba. Si el resultado no coincide con el suyo, vuelva al enunciado antes de leer la indicación. Los ejercicios que piden demostrar algo llevan sólo la idea de la demostración.

Capítulo 1. Conjuntos, lógica y cómo se demuestra algo

1.1. $A \cap B = \{3, 5, 7\}$; $A \cup B = \{1, 2, 3, 5, 7\}$; $A \setminus B = \{1\}$; $A^c = \{2, 4, 6, 8\}$.

1.2. 16. El de dentro tiene 4 elementos, y $2^4 = 16$.

1.3. $\{n^2 : n \in \mathbb{N},\ 1 \le n \le 5\}$.

1.4. Recíproco: "si es divisible por 3, lo es por 6", falso (9). Contrarrecíproco: "si no es divisible por 3, no lo es por 6", cierto y equivalente al original.

1.5. Es la bicondicional $p \Leftrightarrow q$: verdadera sólo cuando p y q coinciden.

1.6. "Para todo estudiante existe un examen que no aprobó". La negación intercambia cuantificadores y niega el interior.

1.7. Falso: para $x = 0$ no hay ningún y con $0 \cdot y = 1$. Con $\mathbb{R} \setminus \{0\}$ sería cierto.

1.8. Inducción: el paso usa $\frac{n(n+1)(2n+1)}{6} + (n+1)^2 = \frac{(n+1)(n+2)(2n+3)}{6}$. Compruébelo en $n = 4$: ambos lados valen 30.

1.9. Por reducción al absurdo: si $q + i$ fuese racional, entonces $i = (q + i) - q$ sería diferencia de racionales y por tanto racional.

1.10. Verdadero. Transitividad de la inclusión: si $x \in A$ entonces $x \in B$ y luego $x \in C$.

1.11. Falso salvo si $A = B$ o alguno es vacío. Con $A = \{1\}$, $B = \{2\}$: $(1, 2) \neq (2, 1)$.

1.12. Ambas ciertas, y son afirmaciones distintas: la primera dice que $\varnothing$ es un *elemento* del conjunto de partes; la segunda, que es un *subconjunto* de A. Confundirlas es el error más común del capítulo.

Capítulo 2. Lógica matemática y teoría de conjuntos

2.1. 16. Una fórmula con dos variables queda determinada por su tabla de verdad, y hay $2^4 = 16$ tablas posibles.

2.2. $\exists x\, \forall y\, (y \leq x \vee \neg P(y))$.

2.3. Sí, $\{\neg, \wedge\}$ es completo, porque $p \vee q \equiv \neg(\neg p \wedge \neg q)$. $\{\Rightarrow\}$ solo no lo es: toda fórmula construida únicamente con $\Rightarrow$ es verdadera cuando todas las variables son verdaderas, así que no puede expresar $\neg p$.

2.4. $\vdash$ es sintáctico —existe una derivación formal— y $\models$ es semántico —en todo modelo de Γ es cierto φ—. El teorema de completitud de Gödel, de 1929, dice que en primer orden coinciden.

2.5. Que si todo subconjunto finito de una teoría tiene modelo, la teoría entera lo tiene. Para el grafo: cada subgrafo finito es 3-coloreable, luego cada subconjunto finito de las fórmulas que describen una 3-coloración es satisfacible, luego el conjunto entero lo es.

2.6. No. Por Löwenheim–Skolem hacia arriba tiene modelos no estándar de todo cardinal infinito, y la aritmética no estándar de Skolem es el ejemplo explícito.

2.7. Porque "no numerable" significa "no hay biyección con $\mathbb{N}$ *dentro del modelo*". Un modelo numerable de ZFC contiene un conjunto que él cree no numerable porque la biyección que lo numeraría, aunque existe fuera, no pertenece al modelo.

2.8. Ambos valen $2^{\aleph_0}$. La hipótesis del continuo afirma que no hay cardinales estrictamente entre $\aleph_0$ y $2^{\aleph_0}$.

2.9. Ni lo uno ni lo otro: es independiente. Gödel probó en 1938 que no es refutable, y Cohen en 1963, con el forcing, que no es demostrable.

2.10. $\omega + 1$, es decir, $\mathbb{N}$ con un elemento añadido al final; o el conjunto $\{1 - 1/n\} \cup \{1\}$ con el orden usual.

2.11. El problema de la parada: el conjunto de programas que terminan es recursivamente enumerable —basta ejecutarlos y esperar— pero no recursivo, porque no hay forma de decidir en tiempo finito que uno *no* va a terminar.

2.12. Falso tal como está enunciado. Gödel demostró que ningún sistema formal fijo, consistente y efectivo demuestra todas las verdades aritméticas; pero cada sentencia indecidible en un sistema es demostrable en otro más fuerte. Lo que no existe es un sistema único que lo demuestre todo, no verdades intrínsecamente incognoscibles.

Capítulo 3. Teoría de números: de la divisibilidad a la criptografía

3.1. $\gcd = 21$ y $21 = 1071 \cdot (-3) + 462 \cdot 7$. Euclides hacia atrás, o el algoritmo extendido.

3.2. $2024 = 2^3 \cdot 11 \cdot 23$.

3.3. 4. Fermat da $3^6 \equiv 1$, y $100 = 6 \cdot 16 + 4$, luego $3^{100} \equiv 3^4 = 81 \equiv 4$.

3.4. 15, porque $7 \cdot 15 = 105 = 4 \cdot 26 + 1$.

3.5. $x \equiv 23 \pmod{105}$.

3.6. $\varphi(36) = 12$ y $\varphi(3120) = 768$. Use la multiplicatividad sobre $36 = 2^2 3^2$ y $3120 = 2^4 \cdot 3 \cdot 5 \cdot 13$.

3.7. El orden es $6 = \varphi(7)$, luego sí, 3 es raíz primitiva módulo 7.

3.8. $6! \equiv 6 \equiv -1 \pmod 7$, como afirma Wilson.

3.9. $\left(\frac{5}{7}\right) = -1$: no es residuo cuadrático. Los cuadrados módulo 7 son $1, 2, 4$.

3.10. $\sigma_0(28) = 6$ y $\sigma_1(28) = 56 = 2 \cdot 28$: sí, 28 es perfecto.

3.11. $\sqrt{7} = [2; \overline{1, 1, 1, 4}\,]$.

3.12. $n = 3233$, $\varphi(n) = 3120$, $d = 2753$.

3.13. (-23) mód $7 = 5$ y (-23) mód $11 = 10$. El resto es siempre no negativo.

3.14. En 9. Las potencias de 9 alternan $9, 1, 9, 1$ y el exponente es impar.

3.15. 1. El orden de 5 módulo 13 es 4, y $4 \mid 200$.

3.16. Divisible entre 9 (las cifras suman 45); no entre 11 (el resto es 6).

3.17. La cifra 4: $4 + 7 + 4 + 3 = 18$.

3.18. No puede serlo. El producto real es 244 944, y aunque la prueba del nueve da 0 en ambos lados —no detecta este error—, la del once sí lo detecta. Es el ejemplo de por qué la prueba del nueve sólo sirve para descartar.

3.19. Tres soluciones módulo 15: $x \equiv 4, 9, 14$. Como $\gcd(9, 15) = 3$ divide a 6, hay exactamente tres.

3.20. 19, porque $11 \cdot 19 = 209 = 8 \cdot 26 + 1$.

3.21. $x \equiv 29 \pmod{36}$.

3.22. RFYJRFYNHFX.

3.23. 486. Hay $\varphi(27) = 18$ multiplicadores válidos y 27 desplazamientos.

3.24. Viernes. $10^6 \equiv 1 \pmod 7$, así que se avanza un solo día.

Capítulo 4. Números complejos desde el principio

4.1. $5 + 3i$ y $46 + i$.

4.2. $-i$. $2027 = 4 \cdot 506 + 3$ y $i^3 = -i$.

4.3. $\dfrac{5}{13} - \dfrac{14}{13}i$. Multiplique por el conjugado del denominador.

4.4. Módulo $3\sqrt{2}$, argumento $3\pi/4$. Ojo con el cuadrante: está en el segundo.

4.5. $4\big(\cos(-2\pi/3) + i\sin(-2\pi/3)\big)$, es decir, módulo 4 y argumento $-2\pi/3$.

4.6. -64. En polares es 2^6 con argumento $6 \cdot \pi/6 = \pi$.

4.7. $z = 3 \pm 5i$.

4.8. $1, i, -1, -i$; suman cero porque son los vértices de un cuadrado centrado en el origen, y en general las n raíces n-ésimas de la unidad suman 0 para $n \geq 2$.

4.9. $x = 2$, $y = 3$.

4.10. $z + \bar{z} = 2\Re z$ y $z\bar{z} = |z|^2$; ambas expresiones son invariantes por conjugación, y un complejo igual a su conjugado es real.

4.11. La recta $y = 0$, es decir, el eje real: son los puntos equidistantes de i y de $-i$.

4.12. $2 - i$ y 2. Como los coeficientes son reales, la conjugada también es raíz; el tercer factor sale dividiendo.

Capítulo 5. Combinatoria y probabilidad

5.1. $\binom{9}{4} = 126$ y $P(9,4) = 3024$. La diferencia es el $4! = 24$ del orden.

5.2. $10\,080 = \dfrac{8!}{2!\,2!}$, por las dos aes y las dos eles.

5.3. $\binom{12}{5} = 792$.

5.4. $\binom{4}{2}\binom{6}{2} = 6 \cdot 15 = 90$.

5.5. $67 = 50 + 33 - 16$, por inclusión-exclusión con los múltiplos de 12.

5.6. $\binom{8}{5}2^5 = 56 \cdot 32 = 1792$.

5.7. 3/8.

5.8. $22/52 = 11/26$. Inclusión-exclusión: $12 + 13 - 3$, restando las tres figuras de corazones.

5.9. 3/4. El complementario es que ambos sean impares, con probabilidad 9/36.

5.10. 3/8. Bayes: $\dfrac{0{,}6 \cdot 0{,}02}{0{,}6 \cdot 0{,}02 + 0{,}4 \cdot 0{,}05}$.

5.11. 127/128. Complementario de "todas caras".

5.12. $45 = \binom{10}{2}$: un apretón es exactamente un subconjunto de dos personas.

Capítulo 6. Cálculo en una variable

6.1. 3/2. Divida numerador y denominador por n^2.

6.2. Ínfimo 0, alcanzado en $n = 1$; supremo 1, no alcanzado.

6.3. La primera converge ($p = 3/2 > 1$); la segunda diverge, por el criterio integral, porque $\int \frac{\mathrm{d}x}{x \ln x} = \ln \ln x$.

6.4. 2. Derive la serie geométrica, o use $\sum nx^n = x/(1-x)^2$ en $x = 1/2$.

6.5. 1. Telescópica: $\frac{1}{n} - \frac{1}{n+1}$.

6.6. Radio 1. Converge en ambos extremos, absolutamente, porque $\sum 1/n^2$ converge.

6.7. $x - \dfrac{x^2}{2} + \dfrac{x^3}{3}$.

6.8. 1/2. Multiplique y divida por el conjugado.

6.9. No en $\mathbb{R}$ (la pendiente crece sin límite); sí en $[0,1]$, por el teorema de Heine sobre compactos.

6.10. $c = 2/\sqrt{3} = 1{,}1547$, de $3c^2 = 4$.

6.11. Converge a $\ln 2$. Al ser condicionalmente convergente, el teorema de Riemann permite reordenarla para obtener cualquier suma real, o incluso $\pm\infty$.

6.12. La sucesión de truncamientos decimales de $\sqrt{2}$, o la iteración $x_{n+1} = \frac{1}{2}(x_n + 2/x_n)$ desde $x_0 = 1$. Es de Cauchy y su límite no es racional.

Capítulo 7. Álgebra lineal

7.1. $\det A = 5$ y $A^{-1} = \frac{1}{5}\begin{pmatrix} 4 & -1 \\ -3 & 2 \end{pmatrix}$.

7.2. Rango 2: la segunda fila es el doble de la primera.

7.3. 3 y 2: es triangular, así que están en la diagonal.

7.4. $\lambda = 3$ con $(1,1)$, y $\lambda = 1$ con $(-1,1)$. Simétrica, luego autovectores ortogonales.

7.5. No: el determinante es 0 porque el tercer vector es la suma de los dos primeros.

7.6. $x = 1$, $y = 2$.

7.7. $\frac{1}{\sqrt{2}}(1,1)$ y $\frac{1}{\sqrt{2}}(1,-1)$.

7.8. Traza 5, determinante -2; los autovalores son $\frac{5\pm\sqrt{33}}{2}$, cuya suma es 5 y cuyo producto es -2.

7.9. Sobre $(1,0)$ da $(3,0)$; sobre $(1,1)$ da $(7/2, 7/2)$.

7.10. Base $\{(-2,1,0),(-3,0,1)\}$; rango 1 más nulidad 2 igual a 3 columnas.

7.11. No. El único autovalor es 1 y su autoespacio tiene dimensión 1, menor que la multiplicidad algebraica 2.

7.12. 3. El teorema del rango se aplica al número de columnas: $7 - 4$.

Capítulo 8. Grafos: la matemática de las conexiones

8.1. $\binom{6}{2} = 15$ y $\binom{10}{2} = 45$.

8.2. No existe: la suma de grados sería 15, impar, y siempre ha de ser par por el lema del apretón de manos.

8.3. 24, el doble del número de aristas.

8.4. 19. Todo árbol con n vértices tiene $n - 1$ aristas.

8.5. 13, de $V - A + C = 2$ con $V = 9$ y $C = 6$.

8.6. $6 - 12 + 8 = 2$. Se cumple.

8.7. Sí a las dos cosas. Todos los grados son pares, luego —si el grafo es conexo— existe circuito euleriano, y todo circuito es en particular un camino.

8.8. K_4 tiene $V = 4$, $A = 6$ y cumple $A \leq 3V - 6 = 6$; se dibuja como un triángulo con el cuarto vértice en el centro unido a los tres.

8.9. 2. Los árboles son bipartitos: coloree por paridad de la distancia a la raíz.

8.10. Camino $A \to C \to B \to D$ con coste 4. El directo $A \to B \to D$ cuesta 5 y $A \to C \to D$ cuesta 6.

8.11. $16 = 4^{4-2}$, por la fórmula de Cayley.

8.12. Cuatro colores siempre bastan, por el teorema de los cuatro colores. Que hagan falta cuatro depende del mapa: si ningún país toca a otro, basta uno.

Capítulo 9. Cálculo multivariable

9.1. $f_x = 2xy$ y $f_y = x^2 + 3y^2$.

9.2. $\operatorname{grad} f(1,2) = (2,4)$; crece más rápido en la dirección $\frac{1}{\sqrt{5}}(1,2)$.

9.3. $\sqrt{2}$. Es $\operatorname{grad} f(1,1) \cdot u = (1,1) \cdot \frac{1}{\sqrt{2}}(1,1)$.

9.4. $z = 2x + 2y - 2$.

9.5. Mínimo en $(2,0)$, con valor -4. El hessiano es $2I$, definido positivo.

9.6. Punto de silla: el hessiano tiene autovalores 2 y -2.

9.7. $1/4$. Las variables se separan.

9.8. $\pi/2$. En polares queda $\int_0^{2\pi}\int_0^1 r^3 \mathrm{d}r\,\mathrm{d}\theta$.

9.9. r en polares; $\rho^2 \sin\varphi$ en esféricas.

9.10. Máximo 25 en $x = y = 5$, con $\lambda = 5$.

9.11. $4\pi/3$.

9.12. $\frac{\mathrm{d}z}{\mathrm{d}t} = 2t\, f_x + 3t^2 f_y$.

Capítulo 10. Coordenadas polares y ecuaciones paramétricas

10.1. $(5, \pi/2)$, $(3,\pi)$ y $(2, -\pi/3)$.

10.2. $(2\sqrt{3},\, 2)$, $(-2,0)$ y $(0,-1)$.

10.3. La circunferencia de centro $(3,0)$ y radio 3, que pasa por el origen.

10.4. $r = 5$ y $\theta = \pi/4$.

10.5. $r = \sin 4\theta$ tiene ocho pétalos y $r = \sin 5\theta$ tiene cinco: con n par salen $2n$, con n impar salen n.

10.6. $1, 0, 1, 2$: es una cardioide con la entrante hacia arriba.

10.7. $y = x^2$, pero sólo la mitad con $x \geq 0$, porque t^2 nunca es negativo.

10.8. $\dfrac{x^2}{4} - \dfrac{y^2}{9} = 1$, una hipérbola, por la identidad $\sec^2 - \tan^2 = 1$.

10.9. $x = 1 + 4t$, $y = 2 - 4t$.

10.10. $T = 5{,}30$ s, $R = 79{,}53$ m y $H = 34{,}44$ m.

10.11. Con 65°: los ángulos complementarios dan el mismo alcance.

10.12. Dividiendo por 2 queda $r = \dfrac{3}{1 + \frac{1}{2}\cos\theta}$, luego $e = 1/2$ —una elipse— y semilado recto 3.

Capítulo 11. Cálculo vectorial

11.1. $2x + 2y + 2z$.

11.2. $(0,0,2)$: rotacional constante, el campo gira uniformemente.

11.3. Sí, con potencial $f = x^2 y$.

11.4. 2, que es $f(1,2) - f(0,0)$. Al ser conservativo, el camino da igual.

11.5. π. La fórmula de Green convierte el área en una integral de borde.

11.6. 4π. La divergencia vale 3 y el volumen es $4\pi/3$.

11.7. Igualdad de las derivadas cruzadas: cada componente del rotacional es una diferencia $f_{xy} - f_{yx}$, nula por el teorema de Schwarz.

11.8. Lo mismo, un escalón más arriba: los seis términos se cancelan por parejas.

11.9. $2\sqrt{2}\,\pi$. La rapidez es constante e igual a $\sqrt{2}$.

11.10. Vale 2π. El rotacional es nulo donde el campo está definido, pero el dominio $\mathbb{R}^2 \setminus \{0\}$ no es simplemente conexo, y el teorema no se aplica. Es el ejemplo canónico de que la topología del dominio importa.

11.11. Ambos lados valen 2π: el rotacional es $(0, 0, 2)$ y el área del disco es π.

11.12. 6.

Capítulo 12. Ecuaciones diferenciales

12.1. $y = 2\mathrm{e}^{3t}$.

12.2. $y = C\mathrm{e}^{-2t}$.

12.3. $y = Ct$.

12.4. $y = C_1\mathrm{e}^{2t} + C_2\mathrm{e}^{3t}$.

12.5. $y = C_1 \cos 2t + C_2 \sin 2t$.

12.6. $y = (C_1 + C_2 t)\mathrm{e}^{-t}$. La t aparece justamente por la raíz doble.

12.7. $y = \frac{1}{2}\mathrm{e}^{t} + C\mathrm{e}^{-t}$. Factor integrante e^{t}.

12.8. $y = \dfrac{1}{1-t}$, que explota en $t = 1$. La ecuación es perfectamente inocente y la solución sólo existe en $(-\infty, 1)$: es el ejemplo estándar de existencia local.

12.9. $y_p = \frac{1}{3}\sin t$. No hay resonancia porque la frecuencia forzante 1 no coincide con la natural 2.

12.10. $W = \mathrm{e}^{3t} \neq 0$, luego las dos soluciones son independientes y forman base.

12.11. $\dfrac{1}{(s-a)^2}$.

12.12. $y = 0$ inestable y $y = 1$ asintóticamente estable. Es el modelo logístico.

Capítulo 13. Ecuaciones en derivadas parciales

13.1. Hiperbólica, elíptica, parabólica y parabólica respectivamente. El discriminante $B^2 - 4AC$ vale 16, -4, 0 y 0.

13.2. $u(x,t) = \frac{1}{2}\big(\mathrm{e}^{-(x-2t)^2} + \mathrm{e}^{-(x+2t)^2}\big)$: la campana se parte en dos que se alejan a velocidad 2.

13.3. Las rectas $x \pm 3t =$ constante.

13.4. $u(x,t) = 3\mathrm{e}^{-4t}\sin 2x$. El dato ya es un modo propio, así que no hay serie.

13.5. Cuando $\mathrm{e}^{-4t} = 1/2$, es decir, $t = \ln 2/4 = 0{,}173$.

13.6. $\lambda_n = (n\pi/L)^2$ con $X_n(x) = \sin(n\pi x/L)$, para $n = 1, 2, 3, \ldots$

13.7. Porque el problema hacia atrás está mal planteado: los modos altos, que hacia adelante decaen como $\mathrm{e}^{-\alpha n^2 t}$, hacia atrás crecen como $\mathrm{e}^{+\alpha n^2 t}$, y un error minúsculo en el dato final se amplifica sin control.

13.8. Se comprueba derivando. Es el núcleo del calor: la temperatura que produce una unidad de calor concentrada en el origen en el instante inicial. Nótese que es positiva para todo x y todo $t > 0$, es decir, propagación instantánea a distancia infinita.

13.9. Una función armónica en un dominio acotado alcanza su máximo y su mínimo en la frontera. Consecuencia: si dos soluciones del problema de Dirichlet coinciden en el borde, su diferencia es armónica y nula en el borde, luego nula en todo el dominio.

13.10. $x^2 - y^2$ sí, porque su laplaciano es $2 - 2 = 0$; es la parte real de z^2. $x^2 + y^2$ no, su laplaciano vale 4. Toda parte real de una función holomorfa es armónica.

13.11. $u_{rr} + \dfrac{1}{r}u_r + \dfrac{1}{r^2}u_{\theta\theta} = 0$.

13.12. Porque la solución no depende continuamente de los datos. Hadamard tomó el dato $u(x,0) = 0$, $u_y(x,0) = \frac{1}{n}\sin nx$, que tiende a cero uniformemente, y sin embargo la solución $\frac{1}{n^2}\sin(nx)\sinh(ny)$ crece sin límite con n para cualquier $y > 0$.

Capítulo 14. Métodos numéricos: cuando no hay fórmula

14.1. $f(1) = -1$ y $f(2) = 7$, luego hay raíz por Bolzano. Primer paso: $f(1{,}5) = 1{,}875 > 0$, queda $[1; 1{,}5]$. Segundo: $f(1{,}25) = 0{,}203 > 0$, queda $[1; 1{,}25]$.

14.2. 9 pasos: hace falta $4/2^n < 0{,}01$, es decir, $2^n > 400$.

14.3. $x_{n+1} = \frac{x_n}{2} + \frac{7}{2x_n}$. Desde 3: 2,666667 y luego 2,645833, frente a $\sqrt{7} = 2{,}645751$.

14.4. $x_{n+1} = x_n - \dfrac{\cos x_n - x_n}{-\sin x_n - 1}$, que converge a 0,739085.

14.5. Porque $f'(0) = 0$ y la fórmula divide por cero. Con raíz triple, Newton pierde su convergencia cuadrática incluso empezando cerca.

14.6. $3 + 1/7 = 3{,}1429$, frente a 3,1623. El error es del seis por mil.

14.7. 5, frente al valor exacto 4. Los trapecios sobreestiman en funciones convexas.

14.8. Exactamente 4. Simpson integra sin error los polinomios hasta grado tres, y x^3 lo es.

14.9. Punto fijo $x = 2$; converge, porque $|g'| = 1/2 < 1$.

14.10. Punto fijo $x = 1/2$; no converge, porque $|g'| = 3 > 1$.

14.11. $\dfrac{1}{\sqrt{x+1} + \sqrt{x}}$. La resta de cantidades casi iguales pierde cifras significativas; el cociente no.

14.12. Cuatro pasos: $1 \to 3 \to 9 \to 27 \to 81$.

Capítulo 15. Optimización y análisis convexo

15.1. El disco sí; $\{xy \geq 1,\ x > 0\}$ sí, es la región sobre una hipérbola en el primer cuadrante, que es convexa; la unión de dos discos disjuntos no.

15.2. Si x, y están en todos los C_i, el segmento también, por convexidad de cada uno. Contraejemplo para la unión: $[0,1] \cup [2,3]$ en $\mathbb{R}$.

15.3. $|x|$ sí; x^3 en $\mathbb{R}$ no, porque es cóncava para $x < 0$; en $[0, \infty)$ sí.

15.4. El intervalo $[-1, 1]$: todo número entre -1 y 1 es la pendiente de una recta que queda por debajo de $|x|$ y la toca en el origen.

15.5. $x = 2$, $y = 1$, con valor 6 y multiplicador $\lambda = 4$.

15.6. Sobre la bola unidad, (0,6; 0,8), es decir, el vector normalizado. Sobre el semiplano $x \leq 1$, el punto $(1, 4)$: sólo se corrige la coordenada que se sale.

15.7. Maximizar $b^\top y$ sujeto a $A^\top y \leq c$, $y \geq 0$.

15.8. Óptimo en $(3, 1)$ con valor 11. Los precios sombra son 2 para la restricción $x + y \leq 4$ y 1 para $x \leq 3$.

15.9. Slater garantiza que no haya salto de dualidad. La convexidad sola no basta: el ejemplo estándar es minimizar e^{-x} sujeto a $x^2/y \leq 0$ con $y > 0$, convexo y con salto igual a 1.

15.10. Para cada restricción, $\lambda_i g_i(x^\star) = 0$. Que $\lambda_i = 0$ significa que esa restricción no está activa en el óptimo: sobra, y relajarla no mejoraría nada.

15.11. No basta la convexidad. Hace falta que el gradiente sea Lipschitz con constante L y tomar paso $\leq 1/L$; entonces el error decae como $O(1/k)$, y como $O(\rho^k)$ si además la función es fuertemente convexa.

15.12. Si x^* fuese local y no global, existiría y mejor, y todos los puntos del segmento $[x^*, y]$ cercanos a x^* tendrían valor menor que $f(x^*)$ por convexidad, contradiciendo la minimalidad local. Falla, por ejemplo, con $f(x) = x^4 - 4x^2 + x$.

Capítulo 16. Variables aleatorias, distribuciones e inferencia

16.1. $E(X) = 7$ y $\mathrm{Var}(X) = 35/6 = 5{,}833$.

16.2. $P(X = 2) = 0{,}3115$, $E(X) = 2$ y $\sigma = 1{,}2247$.

16.3. $k = 1/2$ y $\mu = 4/3$.

16.4. $P(X > 75) = 0{,}0668$ y $P(45 < X < 70) = 0{,}7745$.

16.5. $Q_1 = 53{,}255$, que es $60 - 0{,}6745 \cdot 10$.

16.6. $\mathrm{Var}(3X - 2) = 36$ y $\sigma = 6$. La constante no afecta a la dispersión.

16.7. $n = 166$, de $\left(\frac{2{,}576 \cdot 5}{1}\right)^2 = 165{,}89$ redondeando hacia arriba.

16.8. $z = 2{,}40$, mayor que 1,96: se rechaza. El p-valor es 0,0164.

16.9. 0,166, con media 160, desviación 9,798 y corrección de continuidad. El valor exacto es 0,1662.

16.10. No. La esperanza es 3,5 y el juego cuesta 4: se pierden 0,5 por partida.

16.11. Porque exigir más confianza obliga a abarcar más valores; la única forma de tener a la vez más confianza y menos anchura es más datos.

16.12. 9000. El margen va como $1/\sqrt{n}$, así que dividirlo por tres exige multiplicar n por nueve.

Capítulo 17. Estadística descriptiva y análisis exploratorio

17.1. Media 4,8; mediana 4; moda 4; desviación típica muestral 2,588.

17.2. La media y la mediana suben 10; la desviación típica no cambia.

17.3. Las tres se multiplican por 3.

17.4. Cuando hay valores extremos o la distribución es muy asimétrica. Con salarios, unos pocos sueldos altísimos disparan la media y dejan la mediana intacta; por eso los informes serios dan la mediana.

17.5. 0,15 y 0,30. El segundo conjunto tiene la mitad de desviación absoluta y el doble de dispersión relativa: el coeficiente de variación mide otra cosa.

17.6. Que la cola larga está a la izquierda: la mayoría de los datos se acumulan a la derecha y unos pocos valores pequeños tiran de la media hacia abajo, de modo que media $<$ mediana.

17.7. Por debajo de -3 y por encima de 45.

17.8. Sí puede no tener moda, si todos los valores aparecen una vez; y sí puede tener tres, si hay tres valores empatados en la frecuencia máxima.

17.9. Dividir por $n-1$ hace el estimador insesgado de la varianza poblacional. La razón es que las desviaciones se miden respecto de $\bar{x}$ y no de μ, lo que consume un grado de libertad y hace que la suma de cuadrados salga sistemáticamente pequeña.

17.10. Que la población es probablemente una mezcla de dos subpoblaciones distintas, y que conviene separarlas antes de resumir nada.

17.11. Porque la caja sólo usa cinco números —mínimo, cuartiles y máximo— y esos cinco son idénticos para muchas formas distintas. El histograma muestra la forma entera.

17.12. Que los estadísticos resumen y por tanto pierden información, y que ningún conjunto de números sustituye a mirar el gráfico. Anscombe construyó los cuatro conjuntos en 1973 precisamente para obligar a dibujar los datos.

Capítulo 18. Inferencia estadística

18.1. Al 95 por ciento, (117,06; 122,94); al 99, (116,14; 123,86).

18.2. $t_{15;\,0{,}975} = 2{,}131$, luego (11,281; 13,519).

18.3. $\hat{p} = 0{,}52$ y el intervalo es (0,471; 0,569). Contiene a 0,5: la encuesta no permite afirmar que haya mayoría.

18.4. $n = (1{,}96/0{,}02)^2 \cdot 0{,}25 = 2401$, tomando $p = 0{,}5$, que es el caso peor.

18.5. El de tipo I es rechazar H_0 siendo cierta y el de tipo II es no rechazarla siendo falsa. α controla el primero.

18.6. La potencia es $1-\beta$, la probabilidad de rechazar H_0 cuando es falsa. Depende del tamaño del efecto real, del tamaño muestral y del nivel α.

18.7. $t = 2{,}582$ con 58 grados de libertad y $p = 0{,}012$: se rechaza al 5 por ciento.

18.8. Significa que, *si H_0 fuese cierta*, la probabilidad de observar datos al menos tan extremos como los observados sería 0,03. Dos errores frecuentes: creer que es la probabilidad de

que H_0 sea cierta, y creer que $1-p$ es la probabilidad de que H_1 lo sea. Ninguna de las dos cosas.

18.9. Porque con una muestra suficientemente grande cualquier diferencia, por minúscula que sea, acaba siendo significativa. Significación no es magnitud: hay que mirar el tamaño del efecto.

18.10. Uno, en promedio. La corrección de Bonferroni consiste en usar α/m en cada contraste; para muchos contrastes se prefiere controlar la tasa de falsos descubrimientos con el procedimiento de Benjamini y Hochberg.

18.11. La t cuando σ es desconocida y se estima con s; la normal cuando σ se conoce. A partir de $n \approx 30$ la diferencia entre ambas es despreciable.

18.12. Porque no rechazar sólo significa que no hay evidencia suficiente, lo cual puede deberse a que H_0 sea cierta o a que la muestra sea pequeña. Para concluir equivalencia hay que hacer un contraste de equivalencia, fijando de antemano qué diferencia se considera irrelevante.

Capítulo 19. Modelos lineales, regresión múltiple y análisis de la varianza

19.1. $\hat{\beta}_1 = 100/40 = 2{,}5$ y $\hat{\beta}_0 = 20 - 2{,}5 \cdot 5 = 7{,}5$, luego $\hat{y} = 7{,}5 + 2{,}5x$.

19.2. $R^2 = 0{,}64$: la recta explica el 64 por ciento de la variabilidad de y.

19.3. Sí. Con $y = x^2$ y x simétrico alrededor de cero, la correlación lineal es nula y la dependencia es perfecta. r sólo mide relación *lineal*.

19.4. Linealidad en los parámetros, esperanza nula de los errores, varianza constante (homocedasticidad), independencia, y para la inferencia exacta, normalidad.

19.5. El gráfico de residuos frente a valores ajustados. La heterocedasticidad aparece como un abanico: la dispersión de los residuos crece o decrece con el valor ajustado.

19.6. Correlación alta entre regresores. Se detecta con el factor de inflación de la varianza; valores por encima de 5 o 10 son señal de alarma. Molesta porque infla los errores típicos y vuelve inestables los coeficientes, aunque no daña la predicción.

19.7. Es la diferencia esperada en y entre mujeres y hombres, manteniendo constantes las demás variables del modelo.

19.8. El de la media condicional estima dónde está la recta verdadera y se estrecha con n; el de predicción cubre una observación individual, incluye además la varianza del error y no se estrecha por debajo de σ por mucho que crezca n.

19.9. Porque añadir una columna sólo puede reducir la suma de cuadrados residual: el ajuste antiguo sigue siendo posible con coeficiente cero. El ajustado penaliza los grados de libertad gastados y por eso sí puede bajar.

19.10. Que ese punto influye desproporcionadamente en el ajuste. Hay que examinarlo: comprobar que no es un error de datos, y reportar el ajuste con y sin él.

19.11. Que todos los coeficientes de pendiente son simultáneamente nulos, es decir, que el modelo entero no explica nada.

19.12. Confusora: helados y ahogamientos correlacionan porque ambos suben con el calor. Causalidad inversa: se observa que quienes van al hospital enferman más, cuando es la enfermedad la que lleva al hospital.

Capítulo 20. Estadística bayesiana

20.1. Beta$(5, 15)$, con media $5/20 = 0{,}25$. La priori equivalía a ocho observaciones con dos éxitos, y arrastra el resultado por debajo del 0,3 muestral.

20.2. Para Poisson, la gamma; para la normal con varianza conocida, la normal.

20.3. Los parámetros de la beta se suman a los recuentos: partir de Beta(a, b) y observar k éxitos en n intentos da Beta$(a + k,\, b + n - k)$, exactamente como si $a - 1$ éxitos y $b - 1$ fracasos se hubiesen observado antes.

20.4. El creíble es una afirmación sobre θ dado los datos; el de confianza es una afirmación sobre el procedimiento a lo largo de repeticiones hipotéticas.

20.5. El cociente $P(\text{datos} \mid H_1)/P(\text{datos} \mid H_0)$. Un valor de 10 indica evidencia fuerte a favor de H_1, según la escala de Jeffreys.

20.6. 0,688. Con prevalencia alta el mismo test pasa de ser casi inútil a bastante concluyente: la calidad de un test no es una propiedad suya sola, sino de él y de la población.

20.7. Porque la verosimilitud entra elevada a n y domina cualquier priori propia. En el límite, dos personas con prioris distintas pero no dogmáticas acaban coincidiendo: es el argumento clásico contra la acusación de subjetividad.

20.8. Una que no integra a uno. Hay que comprobar analíticamente que la posteriori sí es propia; si no lo es, el muestreador produce números sin que nada avise.

20.9. Sirve cuando se conocen las distribuciones condicionales completas y son fáciles de muestrear; entonces nunca rechaza. Es preferible cuando la posteriori es multidimensional y las condicionales tienen forma cerrada.

20.10. Compara la varianza entre cadenas con la varianza dentro de cada cadena. Valores por encima de 1,01 indican que las cadenas todavía no han convergido a lo mismo.

20.11. Sí. Con priori uniforme y verosimilitud normal, el intervalo creíble central coincide numéricamente con el de confianza clásico. Coinciden los números y no las interpretaciones.

20.12. Es falsa por los dos lados. El bayesiano hace explícita su priori, que es criticable y revisable; el frecuentista toma decisiones igual de arbitrarias —el nivel α, el modelo, qué contrastar— pero no las escribe. La diferencia está en dónde se declara la subjetividad, no en si existe.

Capítulo 21. Métodos no paramétricos y remuestreo

21.1. Cuando la normalidad es dudosa y la muestra pequeña, y cuando los datos son ordinales o hay valores extremos que distorsionan las medias.

21.2. Contrasta si una variable tiende a tomar valores mayores que la otra, es decir, si $P(X > Y) = 1/2$. Bajo el supuesto de que ambas distribuciones tienen la misma forma y sólo difieren en un desplazamiento, se lee como igualdad de medianas.

21.3. 1, 2,5, 2,5, 4, 5. Los empates reciben el promedio de los rangos que les corresponderían.

21.4. $p = 2\,P(X \geq 7) = 0{,}344$ con $X \sim B(10; 0{,}5)$: no se rechaza.

21.5. Usa las magnitudes de las diferencias además de sus signos, mediante los rangos de los valores absolutos, y por eso tiene más potencia.

21.6. $\rho_s = 0{,}8$.

21.7. La máxima distancia vertical entre la función de distribución empírica y la teórica. Se compara con los valores críticos de la distribución de Kolmogorov, que no dependen de la distribución teórica siempre que ésta se haya fijado de antemano.

21.8. $3/\pi = 0{,}955$. Es decir, bajo normalidad Wilcoxon pierde menos de un cinco por ciento de eficiencia frente a la t, y para colas pesadas puede ser mucho más eficiente.

21.9. El bootstrap estima la variabilidad de un estadístico remuestreando de la muestra; la validación cruzada estima el error de predicción de un modelo reservando partes de los datos. Uno mide incertidumbre, el otro mide capacidad de generalizar.

21.10. Una fracción $(1 - 1/n)^n$, que tiende a $1/\mathrm{e} = 0{,}368$.

21.11. Sí, pero no el bootstrap ordinario, que destruye la dependencia. Hace falta remuestrear bloques contiguos: el bootstrap por bloques.

21.12. Se reasignan las etiquetas de grupo de todas las formas posibles —o de muchas al azar— y se mira en qué percentil cae el estadístico observado. Su validez sólo necesita que, bajo H_0, las etiquetas sean intercambiables, y esa hipótesis es sobre el diseño del experimento, no sobre la distribución de los datos.

Capítulo 22. Procesos estocásticos

22.1. $\pi = (2/7,\ 5/7) = (0{,}2857;\ 0{,}7143)$.

22.2. Irreducible: desde cualquier estado se alcanza cualquier otro. Aperiódica: los tiempos de retorno posibles no tienen un divisor común mayor que 1. Ejemplo periódico: la cadena de dos estados que alterna con probabilidad 1, de periodo 2.

22.3. Uno del que no se sale: $p_{ii} = 1$. Por ejemplo $\begin{pmatrix} 1 & 0 \\ 0{,}3 & 0{,}7 \end{pmatrix}$.

22.4. $1 - 3/10 = 0{,}7$. Con juego justo, la probabilidad de alcanzar el objetivo es la fracción del capital inicial sobre el total.

22.5. 0,214, con $\lambda = 2{,}5$ en media hora.

22.6. Exponencial de parámetro λ, con media $1/\lambda$.

22.7. Que $P(T > s + t \mid T > s) = P(T > t)$: haber esperado no cambia lo que queda por esperar. En el caso discreto la tiene la geométrica.

22.8. $\mathrm{Cov}(B(2), B(5)) = \min(2, 5) = 2$.

22.9. Es $S(t) = S_0 \exp\big((\mu - \sigma^2/2)t + \sigma B(t)\big)$. Se usa porque garantiza precios positivos y porque hace que los rendimientos logarítmicos sean normales e independientes, que es

la hipótesis de Black y Scholes.

22.10. Un proceso con $E[X_{t+s} \mid \mathcal{F}_t] = X_t$: la mejor previsión del futuro es el presente. $B(t)$ lo es, y $B(t)^2 - t$ también.

22.11. Si la cadena es irreducible, aperiódica y con espacio de estados finito —o positiva recurrente en el caso infinito— entonces existe una única distribución estacionaria y P^n converge a ella desde cualquier inicio.

22.12. Porque en un intervalo de longitud t dividido en n trozos, la suma de los valores absolutos de los incrementos crece como $n \cdot \sqrt{t/n} = \sqrt{nt} \to \infty$, mientras que la suma de sus cuadrados es $n \cdot (t/n) = t$, constante. Esa asimetría entre la primera y la segunda potencia es la razón de ser del cálculo de Itô.

Capítulo 23. Teoría de grupos desde cero: la simetría explicada con las manos

23.1. $|S_4| = 24$, $|A_4| = 12$ y $|D_4| = 8$.

23.2. Componiendo de derecha a izquierda es el 4-ciclo $(1\,2\,3\,4)$, que es impar por constar de tres transposiciones.

23.3. $10 = \operatorname{mcm}(5, 2)$. El orden de una permutación es el mínimo común múltiplo de las longitudes de sus ciclos disjuntos.

23.4. Seis, uno por cada divisor de 12: los generados por $0, 6, 4, 3, 2, 1$, de órdenes $1, 2, 3, 4, 6, 12$.

23.5. Sí a lo primero, porque $\gcd(2, 3) = 1$; no a lo segundo, porque $\mathbb{Z}/2\mathbb{Z} \times \mathbb{Z}/2\mathbb{Z}$ no tiene ningún elemento de orden 4.

23.6. Cinco: la rotación de 180° y las cuatro reflexiones.

23.7. $n_5 \equiv 1 \pmod 5$ y divide a 3, luego $n_5 = 1$; análogamente $n_3 = 1$. Ambos Sylow son normales, se cortan trivialmente y $G \cong \mathbb{Z}/3\mathbb{Z} \times \mathbb{Z}/5\mathbb{Z} \cong \mathbb{Z}/15\mathbb{Z}$.

23.8. $n_3 \in \{1, 4\}$: debe dividir a 4 y ser congruente con 1 módulo 3.

23.9. A_4 no es simple: contiene el subgrupo normal de Klein V_4. A_5 sí lo es, y de ahí sale la irresolubilidad de la quíntica.

23.10. S_4 es resoluble, con serie derivada $S_4 \rhd A_4 \rhd V_4 \rhd \{e\}$ de cocientes abelianos. S_5 no lo es, porque A_5 es simple no abeliano.

23.11. Orden $\varphi(15) = 8$, isomorfo a $\mathbb{Z}/2\mathbb{Z} \times \mathbb{Z}/4\mathbb{Z}$. No es cíclico, porque 15 no tiene raíces primitivas.

23.12. Tres: $\mathbb{Z}/8\mathbb{Z}$, $\mathbb{Z}/4\mathbb{Z} \times \mathbb{Z}/2\mathbb{Z}$ y $(\mathbb{Z}/2\mathbb{Z})^3$, una por cada partición de 3.

Capítulo 24. Anillos, cuerpos y polinomios desde cero

24.1. No lo es: $2 \cdot 3 = 0$ en $\mathbb{Z}/6\mathbb{Z}$, luego 2 y 3 son divisores de cero.

24.2. Exactamente cuando n es primo.

24.3. (x) es primo, porque $\mathbb{Z}[x]/(x) \cong \mathbb{Z}$ es dominio, pero no maximal, porque $\mathbb{Z}$ no es cuerpo. Sí es maximal (x, p) con p primo.

24.4. A $\mathbb{Z}[i]$, los enteros de Gauss: la clase de x hace de i.

24.5. Sí, por Eisenstein con $p = 2$, o simplemente porque no tiene raíces racionales y es de grado 3.

24.6. $x^4 + 1 = (x + 1)^4$ sobre $\mathbb{F}_2$. Es el ejemplo estándar de un polinomio irreducible sobre $\mathbb{Q}$ que se descompone módulo todo primo.

24.7. $x^2 + 1 = (x + 2)(x + 3)$ sobre $\mathbb{F}_5$, porque -1 es residuo cuadrático módulo 5.

24.8. No lo es: $6 = 2 \cdot 3 = (1 + \sqrt{-5})(1 - \sqrt{-5})$, y los cuatro factores son irreducibles y no asociados. Es el ejemplo que obligó a Dedekind a inventar los ideales.

24.9. Cada inclusión es estricta. $\mathbb{Z}[x]$ es DFU pero no DIP, porque el ideal $(2, x)$ no es principal.

24.10. Ocho elementos, y sí es cuerpo: $x^3 + x + 1$ es irreducible sobre $\mathbb{F}_2$ por no tener raíces. Es $\mathbb{F}_8$.

24.11. Si $f : A \to B$ es homomorfismo, entonces $A/\ker f \cong \operatorname{im} f$. Con la proyección $\mathbb{Z} \to \mathbb{Z}/n\mathbb{Z}$, el núcleo es (n) y la imagen es todo.

24.12. Todo maximal es primo, pero no al revés: en $\mathbb{Z}[x]$, el ideal (x) es primo y no maximal.

Capítulo 25. Estructuras algebraicas: grupos, anillos y cuerpos

25.1. $[\mathbb{Q}(\sqrt{2}) : \mathbb{Q}] = 2$, con base $\{1, \sqrt{2}\}$.

25.2. 4, con base $\{1, \sqrt{2}, \sqrt{3}, \sqrt{6}\}$, por la multiplicatividad de los grados.

25.3. $x^4 - 10x^2 + 1$. Eleve al cuadrado dos veces y compruebe que no factoriza sobre $\mathbb{Q}$.

25.4. Sobre $\mathbb{Q}$ es $x^3 - 2$, de grado 3; sobre $\mathbb{Q}(\sqrt[3]{2})$ es $x - \sqrt[3]{2}$, de grado 1.

25.5. No: contiene una raíz de $x^3 - 2$ pero no las otras dos, que son complejas. Su clausura normal tiene grado 6.

25.6. Nueve elementos, y sí es cuerpo, porque $x^2 + 1$ no tiene raíces en $\mathbb{F}_3$. Es $\mathbb{F}_9$.

25.7. De orden 6 no, porque todo cuerpo finito tiene orden potencia de un primo; de orden $27 = 3^3$ sí, y es único salvo isomorfismo.

25.8. $\varphi(5) = 4$. El polinomio mínimo es el ciclotómico $x^4 + x^3 + x^2 + x + 1$.

25.9. Sobre $\mathbb{Q}$ no lo es (Lindemann, 1882); sobre $\mathbb{Q}(\pi^2)$ sí, porque es raíz de $x^2 - \pi^2$.

25.10. 1, 2, 3 o 6: los divisores de 6, por la fórmula de las torres.

25.11. $\mathbb{Q}(\sqrt[4]{2}, i)$, de grado 8 sobre $\mathbb{Q}$.

25.12. Toda extensión finita es algebraica; el recíproco es falso, y el contraejemplo es la clausura algebraica de $\mathbb{Q}$ dentro de $\mathbb{C}$, que es algebraica y de grado infinito.

Capítulo 26. Teoría de Galois

26.1. Es de orden 2: la identidad y la conjugación $\sqrt{2} \mapsto -\sqrt{2}$.

26.2. El grupo de Klein $\mathbb{Z}/2\mathbb{Z} \times \mathbb{Z}/2\mathbb{Z}$, con los cuatro cambios independientes de signo.

26.3. Tres cuerpos intermedios propios: $\mathbb{Q}(\sqrt{2})$, $\mathbb{Q}(\sqrt{3})$ y $\mathbb{Q}(\sqrt{6})$, uno por cada subgrupo de orden 2. La correspondencia invierte las inclusiones.

26.4. S_3, de orden 6. El cuerpo de descomposición es $\mathbb{Q}(\sqrt[3]{2}, \omega)$, de grado 6.

26.5. El grupo diédrico D_4, de orden 8.

26.6. El de Klein, $\mathbb{Z}/2\mathbb{Z} \times \mathbb{Z}/2\mathbb{Z}$: el cuerpo de descomposición es $\mathbb{Q}(\zeta_8)$, de grado 4 y abeliano.

26.7. Cíclico de orden n, generado por el Frobenius $x \mapsto x^p$.

26.8. Porque S_5 no es resoluble, al contener el simple no abeliano A_5, y una fórmula por radicales exigiría una torre de extensiones cíclicas.

26.9. $x^5 - 6x + 3$, irreducible por Eisenstein con $p = 3$. El argumento: tiene exactamente tres raíces reales y dos complejas conjugadas, luego el grupo contiene una transposición —la conjugación compleja— y un 5-ciclo, y esos dos generan S_5.

26.10. El de 20° no, porque exigiría resolver una cúbica irreducible y el grado constructible ha de ser potencia de 2; el pentágono sí, porque 5 es primo de Fermat.

26.11. $\cos(2\pi/5) = \dfrac{\sqrt{5}-1}{4}$.

26.12. Los que cumplen $n = 2^k p_1 \cdots p_r$ con los p_i primos de Fermat distintos, por el teorema de Gauss–Wantzel. Los primeros: $3, 4, 5, 6, 8, 10, 12, 15, 16, 17$.

Capítulo 27. Teoría de la representación de grupos

27.1. Seis, todas de dimensión 1, porque el grupo es abeliano.

27.2. $\chi_k(j) = i^{\,kj}$ para $k = 0, 1, 2, 3$, es decir, las cuatro potencias de i.

27.3. Cinco irreducibles, de dimensiones $1, 1, 1, 1, 2$; en efecto $4 \cdot 1 + 4 = 8$. D_4 tiene cinco clases de conjugación.

27.4. El carácter es $(4, 2, 1, 0, 0)$ sobre las clases e, (12), $(12)(34)$, (123), (1234), contando puntos fijos. Se descompone como trivial más estándar, de dimensiones $1 + 3$.

27.5. Que toda representación de un grupo finito sobre un cuerpo de característica que no divide a $|G|$ es suma directa de irreducibles. Falla en característica p dividiendo a $|G|$: en $\mathbb{F}_2$, la representación de $\mathbb{Z}/2\mathbb{Z}$ por $\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$ tiene un subespacio invariante sin complemento invariante.

27.6. $\chi(gxg^{-1}) = \operatorname{tr}(\rho(g)\rho(x)\rho(g)^{-1}) = \operatorname{tr}(\rho(x)) = \chi(x)$, por la invariancia cíclica de la traza.

27.7. Es irreducible si y sólo si $\langle \chi, \chi \rangle = 1$.

27.8. 5, por la fórmula del gancho.

27.9. Once, una por cada partición de 6.

27.10. $V_1 \otimes V_{1/2} \cong V_{1/2} \oplus V_{3/2}$, con dimensiones $3 \times 2 = 6 = 2 + 4$.

27.11. Forman el grupo dual $\hat{G}$, isomorfo a G aunque no canónicamente. La dualidad de Pontriaguin dice que $\hat{\hat{G}} \cong G$, y ese segundo isomorfismo sí es natural.

27.12. Que las exponenciales e^{inx} son exactamente los caracteres irreducibles del grupo circunferencia, y desarrollar en serie de Fourier es descomponer en irreducibles.

Capítulo 28. Topología: la geometría de lo que no se rompe

28.1. Ninguna de las dos cosas. Interior $(0,1)$, clausura $[0,1]$, frontera $\{0,1\}$.

28.2. Dado $y \in B(x,r)$, tome $\rho = r - d(x,y) > 0$; la desigualdad triangular da $B(y,\rho) \subseteq B(x,r)$.

28.3. Todos son abiertos. Los compactos son exactamente los finitos: un conjunto infinito se cubre con sus puntos y no admite subcubrimiento finito.

28.4. Es cerrado, porque su complementario es unión de intervalos abiertos. No es compacto, porque no está acotado. Son dos propiedades independientes.

28.5. $\{(1/n, 1) : n \geq 2\}$. Ninguna cantidad finita cubre los puntos próximos a 0.

28.6. La usual es estrictamente más fina. La identidad $(\mathbb{R}, \tau_1) \to (\mathbb{R}, \tau_2)$ sí es continua, porque la preimagen de todo abierto de τ_2 está en τ_1; la inversa no.

28.7. Si $f(X) = U \sqcup V$ con U, V abiertos disjuntos no vacíos, entonces $X = f^{-1}(U) \sqcup f^{-1}(V)$ sería una desconexión de X.

28.8. No. La compacidad los distingue, y también el hecho de que $[0,1]$ tiene puntos que al quitarlos dejan el resto conexo.

28.9. No. Quitando un punto, S^1 sigue conexo y $[0,1]$ se parte en dos —salvo si se quita un extremo—, lo cual basta para descartar el homeomorfismo.

28.10. $f(X)$ es compacto en $\mathbb{R}$, luego cerrado y acotado, luego contiene su supremo.

28.11. No es conexo. Sus componentes conexas son los puntos: $\mathbb{Q}$ es totalmente disconexo, pese a no tener ningún punto aislado.

28.12. Es cerrado, pero no conexo —tiene dos ramas— ni compacto, porque no está acotado.

Capítulo 29. Análisis real: los cimientos y las patologías

29.1. Ambos son numerables: los polinomios racionales son unión numerable de conjuntos numerables, y cada uno tiene finitas raíces. De ahí que "casi todo" real sea trascendente, aunque exhibir uno cueste.

29.2. Por ejemplo, mande $0 \mapsto 1/2$, $1 \mapsto 1/3$, $1/n \mapsto 1/(n+2)$ y deje fijo el resto. El truco del hotel de Hilbert.

29.3. Medida cero, y sin embargo su cardinal es el del continuo. Es el contraejemplo que separa "pequeño en medida" de "pequeño en cardinal".

29.4. La función de Dirichlet. Falla el criterio de Lebesgue porque es discontinua en todo punto, y ese conjunto no tiene medida nula.

29.5. Sí es integrable, y su integral vale 0: sus discontinuidades son los racionales, de medida nula.

29.6. No, por el teorema de Baire: $\mathbb{R}$ es completo y no puede ser unión numerable de cerrados con interior vacío.

29.7. Sí converge, y al único punto fijo $x^* = 0{,}739085$. En $[0,1]$ se tiene $|\cos'| = |\sin x| \leq \sin 1 < 1$.

29.8. Sí, uniformemente, por el teorema de Dini: monotonía más continuidad del límite sobre un compacto.

29.9. Le falta la equicontinuidad, que falla cerca de $x = 1$: las pendientes nx^{n-1} no están acotadas uniformemente. Además el límite no es continuo, lo que ya delata el problema.

29.10. A cualquier número real, y también a $+\infty$ o $-\infty$. Teorema de reordenación de Riemann.

29.11. No lo es. La variación total diverge como la serie armónica; con $x^2 \sin(1/x)$ sí sería de variación acotada.

29.12. No. La función de Cantor es continua, monótona —luego de variación acotada— y no absolutamente continua: sube de 0 a 1 sobre un conjunto de medida nula.

Capítulo 30. Variable compleja: el puente entre los mundos

30.1. En ninguno. Con $u = x$ y $v = -y$ resulta $u_x = 1$ y $v_y = -1$, que nunca coinciden.

30.2. Con $u = x^2 - y^2$ y $v = 2xy$: $u_y = -2y$ y $v_x = 2y$, luego $u_y = -v_x$, y también $u_x = v_y = 2x$. Se cumplen.

30.3. $2\pi i$.

30.4. $-i/2$.

30.5. $1/2$, el coeficiente de z^2 en la serie de e^z.

30.6. En $\sin z/z$ es evitable (el límite vale 1); en $\mathrm{e}^{1/z}$ es esencial, y por Picard la función toma allí todo valor complejo salvo a lo sumo uno, infinitas veces.

30.7. $-\dfrac{1}{z} - 1 - z - \cdots$, desarrollando $\dfrac{-1}{1-z}$ en serie geométrica.

30.8. $\dfrac{\pi}{\sqrt{2}}$.

30.9. $\dfrac{\pi}{\mathrm{e}}$. Use e^{iz} y quédese con la parte real; el lema de Jordan controla el arco.

30.10. $\dfrac{2\pi}{\sqrt{3}}$.

30.11. Exactamente uno. Sobre $|z| = 1$ se tiene $|5z| = 5 > |z^4 + 1| \leq 2$, luego $z^4 + 5z + 1$ tiene los mismos ceros que $5z$.

30.12. Que es constante, por Liouville. Y si $|f|$ alcanza un máximo interior, también es constante, por el principio del módulo máximo.

Capítulo 31. La integral de Lebesgue y la transformada de Fourier

31.1. 0 y 1 respectivamente. Todo conjunto numerable tiene medida nula.

31.2. Enumere los racionales de $[0, 1]$ y ponga alrededor de q_n un intervalo de longitud $\varepsilon 2^{-n}$ con $\varepsilon = 1/20$: la unión es abierta, densa y de medida a lo sumo $1/10$.

31.3. Sí, y vale 0: es la indicatriz de un conjunto de medida nula. Riemann no la integra.

31.4. $\lim \int f_n = 1$ y $\int \lim f_n = 0$. No falla ningún teorema: falla la hipótesis de dominación, porque no hay ninguna integrable que acote a todas.

31.5. 0. Está dominada por la constante 1 y converge puntualmente a cero.

31.6. $\int \liminf f_n \leq \liminf \int f_n$, y el ejemplo anterior muestra que la desigualdad puede ser estricta: $0 < 1$.

31.7. $1/\sqrt{x}$ sí, con integral 2; $1/x$ no, porque diverge logarítmicamente.

31.8. Justo al revés que antes: $1/x^2$ sí, con integral 1; $1/x$ no.

31.9. $\sqrt{\pi}$. Y $\sin x/x$ *no* es Lebesgue-integrable en $(0,\infty)$, porque $\int |\sin x/x| = \infty$, aunque la integral impropia de Riemann converja a $\pi/2$.

31.10. No. Lusin afirma algo más débil y más útil: para todo $\varepsilon > 0$ existe un cerrado cuyo complementario tiene medida menor que ε y sobre el cual f es continua.

31.11. Sí, el conjunto de Vitali, y la construcción usa el axioma de elección. Sin él es consistente que todo subconjunto de $\mathbb{R}$ sea medible.

31.12. Que $f = 0$ en casi todo punto, es decir, salvo un conjunto de medida nula.

Capítulo 32. Análisis funcional: álgebra lineal en dimensión infinita

32.1. Es norma, pero no proviene de un producto interno para $n \geq 2$: falla la ley del paralelogramo con $x = (1,0)$ e $y = (0,1)$.

32.2. En $\mathbb{R}^n$ sí, todas. En dimensión infinita no: en $C[0,1]$, las normas del supremo y L^1 no lo son.

32.3. Con el supremo sí es Banach; con L^1 no, y las sucesiones que aproximan un escalón dan el contraejemplo.

32.4. Sólo ℓ^2. Es el único p para el que se cumple la ley del paralelogramo, y por tanto el único cuya norma viene de un producto escalar.

32.5. En ℓ^2: $\|x+y\|^2 + \|x-y\|^2 = 2+2 = 4 = 2(1+1)$. En ℓ^1 el lado izquierdo vale $4+4=8$ y el derecho 4.

32.6. $\|\varphi\| = 1$, alcanzada en la función constante 1.

32.7. Todo funcional acotado sobre un subespacio se extiende a todo el espacio con la misma norma. Garantiza que el dual es lo bastante rico para separar puntos.

32.8. Que una familia de operadores acotada puntualmente lo está en norma. Descansa en la completitud del dominio, vía el teorema de Baire.

32.9. No: el desplazamiento del ejercicio 11 es inyectivo y no sobreyectivo. El caso sobreyectivo lo controla el teorema de la aplicación abierta.

32.10. En norma no, por el teorema de Riesz: la compacidad de la bola caracteriza la dimensión finita. En la topología débil-$*$ del dual sí, por Banach–Alaoglu.

32.11. Es isométrico y no sobreyectivo —su imagen es el subespacio de sucesiones con primera coordenada nula—, y no tiene ningún autovalor. En dimensión finita, isométrico implicaría sobreyectivo.

32.12. Que existe una base ortonormal de autovectores, con autovalores reales que tienden a cero. Es la generalización exacta de la diagonalización de matrices simétricas.

Capítulo 33. Análisis armónico

33.1. $a_0 = \pi$ y $a_n = \dfrac{2((-1)^n - 1)}{\pi n^2}$, es decir, $-4/(\pi n^2)$ para n impar y 0 para n par. Todos los b_n se anulan por ser función par.

33.2. Porque el producto de una función par por un seno es impar y su integral sobre un intervalo simétrico se anula, y viceversa.

33.3. $\frac{1}{\pi}\int |f|^2 = \frac{a_0^2}{2} + \sum(a_n^2 + b_n^2)$. Aplicada a la onda cuadrada da $\sum_{k\geq 0} \frac{1}{(2k+1)^2} = \frac{\pi^2}{8} = 1{,}2337$.

33.4. Cuanto más derivable es f, más rápido decaen: si f es de clase C^k y periódica, los coeficientes son $o(n^{-k})$. Un salto da decaimiento $1/n$; una función analítica, un decaimiento exponencial.

33.5. No. Kolmogorov construyó en 1926 una función integrable cuya serie diverge en todo punto. Pero Carleson demostró en 1966 que para funciones de cuadrado integrable —y en particular continuas— la serie converge en casi todo punto.

33.6. $\hat{f}(\xi) = \dfrac{\sin(2\pi\xi)}{\pi\xi}$, la función seno cardinal.

33.7. Trasladar multiplica la transformada por una exponencial de módulo uno; dilatar por a la comprime por $1/a$ y la escala por $1/|a|$.

33.8. $\widehat{f * g} = \hat{f} \cdot \hat{g}$. Se usa para convertir convoluciones —que cuestan $O(n^2)$— en productos punto a punto, y de ahí el filtrado digital de señales e imágenes.

33.9. Que una señal cuyo espectro se anula por encima de la frecuencia B queda determinada por sus muestras tomadas a ritmo $2B$. El aliasing es lo que ocurre al muestrear más despacio: las frecuencias altas se disfrazan de bajas y no hay forma de distinguirlas.

33.10. Porque la transformada de tamaño n se descompone en dos de tamaño $n/2$ —las muestras pares y las impares— más una combinación lineal de coste $O(n)$, y esa recursión da $T(n) = 2T(n/2) + O(n)$.

33.11. Es $D_n(x) = \dfrac{\sin((n+\frac{1}{2})x)}{\sin(x/2)}$. Su norma L^1 crece como $\log n$ porque el numerador oscila mientras el denominador se comporta como x. Ese crecimiento es exactamente la razón de que existan funciones continuas con serie divergente en un punto.

33.12. Que el círculo es un grupo abeliano compacto, que sus caracteres son las exponenciales, y que la serie de Fourier es la descomposición de una función en los caracteres.

Capítulo 34. Teoría analítica de números

34.1. $\mu(12) = 0$ porque 4 divide a 12; $\mu(30) = -1$ porque $30 = 2 \cdot 3 \cdot 5$ tiene tres factores distintos; $\mu(1) = 1$ por convenio.

34.2. Los divisores de 30 son $1, 2, 3, 5, 6, 10, 15, 30$ y sus φ valen $1, 1, 2, 4, 2, 4, 8, 8$, que suman 30.

34.3. $\zeta(2) = \pi^2/6$ y $\zeta(4) = \pi^4/90$.

34.4. Sí es multiplicativa. $\sigma_1(12) = 1 + 2 + 3 + 4 + 6 + 12 = 28$.

34.5. $\pi(100) = 25$, frente a $100/\ln 100 = 21{,}71$ y $\mathrm{li}(100) = 30{,}13$. Con x tan pequeño ninguna

de las dos aproximaciones es buena todavía.

34.6. $(3,5)$, $(5,7)$, $(11,13)$, $(17,19)$, $(29,31)$, $(41,43)$, $(59,61)$ y $(71,73)$: ocho pares.

34.7. Que toda progresión $an+b$ con $\gcd(a,b)=1$ contiene infinitos primos. En particular hay infinitos de la forma $4k+3$, cosa que además admite una demostración elemental al estilo de Euclides.

34.8. En $s=1$, y es un polo simple de residuo 1.

34.9. Son los enteros pares negativos $-2,-4,-6,\ldots$, y salen de los polos de la función gamma en la ecuación funcional, que los convierte en ceros de ζ.

34.10. Equivale a $\pi(x)=\mathrm{li}(x)+O(\sqrt{x}\,\ln x)$, que es el término de error más pequeño compatible con la existencia de ceros.

34.11. Que entre n y $2n$ siempre hay un primo, para todo $n\geq 1$. Lo conjeturó Bertrand en 1845 y lo demostró Chebyshev en 1852; Erdős dio una demostración elemental en 1932.

34.12. Porque $\sum 1/n^2$ converge y $\sum 1/p$ diverge: los primos, siendo mucho más escasos que los enteros, son bastante más abundantes que los cuadrados. La infinitud sola no dice nada sobre la densidad; la divergencia sí.

Capítulo 35. Geometría algebraica

35.1. En $\mathbb{C}^2$ es la unión de las dos rectas $y=\pm ix$; en $\mathbb{R}^2$ es sólo el origen. Sobre $\mathbb{R}$ el ideal no es radical y el diccionario falla.

35.2. No es primo, porque $x\cdot y\in(xy)$ y ninguno de los dos factores está en el ideal. Sí es radical.

35.3. El ideal maximal (x,y).

35.4. Dimensión 1 y dos componentes irreducibles, los dos ejes.

35.5. Todo ideal de $K[x_1,\ldots,x_n]$ está generado por finitos elementos. Consecuencia: toda variedad es intersección de finitas hipersuperficies, y las cadenas descendentes de variedades se estabilizan.

35.6. Cuatro, por Bézout con $m=n=2$.

35.7. $g=\dfrac{(4-1)(4-2)}{2}=3$.

35.8. Porque una cúbica lisa tiene género 1 y sólo las curvas de género 0 admiten parametrización racional. Ésa es la razón de fondo de que el problema del logaritmo discreto sea difícil allí.

35.9. El origen. Es una cúspide: la curva llega con tangente única y retrocede, y el gradiente $(3x^2,-2y)$ se anula ahí.

35.10. Es el espacio de ideales primos de $\mathbb{Z}$: un punto cerrado por cada primo, más el ideal (0), que es el punto genérico, denso en todo el espacio.

35.11. $|E(\mathbb{F}_5)|=6$, luego $a_5=5+1-6=0$, que cumple holgadamente $|a_5|\leq 2\sqrt{5}=4{,}47$.

35.12. El proyectivo, porque si no las rectas paralelas no se cortan y falta un punto; y el cierre algebraico, porque si no faltan raíces, como en $x^2+y^2+1=0$.

Capítulo 36. Cálculo tensorial

36.1. El gradiente es covariante y la velocidad contravariante. Al cambiar de base con matriz M, las componentes contravariantes se transforman con M^{-1} y las covariantes con $M^\top$: por eso el producto $\omega_i V^i$ no depende de la base.

36.2. $4^3 = 64$.

36.3. $\mathrm{d}s^2 = \mathrm{d}r^2 + r^2\mathrm{d}\theta^2$, es decir, $g = \mathrm{diag}(1, r^2)$.

36.4. $\det g = R^4 \sin^2\theta$, luego $\mathrm{d}A = R^2 \sin\theta\,\mathrm{d}\theta\,\mathrm{d}\varphi$, que integrado da $4\pi R^2$.

36.5. $\Gamma^\theta_{\varphi\varphi} = -\sin\theta\cos\theta$ y $\Gamma^\varphi_{\theta\varphi} = \cot\theta$. No dependen de R.

36.6. $R_{\mathrm{esc}} = 2/R^2$: positiva y constante, y tanto menor cuanto mayor es la esfera.

36.7. Una en dimensión 2 y veinte en dimensión 4, de la fórmula $n^2(n^2-1)/12$.

36.8. $\ddot{x}^k + \Gamma^k_{ij}\dot{x}^i\dot{x}^j = 0$. En la esfera son los círculos máximos.

36.9. No lo es: su ley de transformación tiene un término adicional con segundas derivadas. Por eso puede anularse en un punto eligiendo bien las coordenadas, cosa imposible para un tensor no nulo.

36.10. $V_i = (1, r^2)$ y $\|V\|^2 = 1 + r^2$.

36.11. La covariante corrige el cambio de los vectores de la base al moverse por la variedad. Sin esa corrección, la derivada de un campo tensorial no sería un tensor.

36.12. El de Riemann mide toda la curvatura, incluida la deformación de forma; el de Ricci es su traza y mide sólo cómo cambia el volumen de una bola pequeña. En dimensión 2 y 3 Ricci determina a Riemann; en dimensión 4 no, y lo que sobra es el tensor de Weyl.

Capítulo 37. Topología algebraica

37.1. $\pi_1(\mathbb{R}^2) = 0$; $\pi_1(\mathbb{R}^2 \setminus \{0\}) = \mathbb{Z}$; $\pi_1(\mathbb{R}^3 \setminus \{0\}) = 0$, porque es homotópicamente equivalente a S^2, que es simplemente conexa.

37.2. Homotópicamente equivalentes sí: el cilindro se retrae sobre su círculo central. Homeomorfos no, porque tienen dimensiones distintas.

37.3. Los tres valen 2: $8 - 12 + 6$, $6 - 12 + 8$ y $12 - 30 + 20$. Todos los sólidos platónicos son homeomorfos a la esfera.

37.4. $\chi = 2 - 2g$.

37.5. $\pi_1(T^2) = \mathbb{Z} \times \mathbb{Z}$. Es abeliano porque el toro es un grupo topológico, y el grupo fundamental de un grupo topológico siempre lo es.

37.6. $\langle a, b \mid abab^{-1}\rangle$, que no es abeliano. Su abelianizado es $\mathbb{Z} \times \mathbb{Z}/2\mathbb{Z}$, y ahí se ve la no orientabilidad.

37.7. Si $X = U \cup V$ con U, V abiertos, conexos por caminos y con $U \cap V$ conexo por caminos, entonces $\pi_1(X)$ es el producto amalgamado de $\pi_1(U)$ y $\pi_1(V)$ sobre $\pi_1(U \cap V)$. Dos discos pegados por el borde dan S^2, y sale $\pi_1(S^2) = 0$.

37.8. Una aplicación que localmente es un homeomorfismo sobre copias disjuntas. El recubridor universal de S^1 es $\mathbb{R}$ con $t \mapsto \mathrm{e}^{2\pi i t}$, y el de la banda de Möbius es la banda cilíndrica infinita, con dos hojas sobre el círculo central.

37.9. $H_0 = \mathbb{Z}$, $H_3 = \mathbb{Z}$ y todos los demás nulos. En general $H_k(S^n) = \mathbb{Z}$ para $k = 0$ y $k = n$, y cero en lo demás.

37.10. La característica de Euler, que vale 2 y 0; y el grupo fundamental, trivial frente a $\mathbb{Z} \times \mathbb{Z}$. Cualquiera de los dos basta.

37.11. Que todo campo vectorial continuo tangente a S^2 se anula en algún punto. En términos del viento: en cualquier instante hay al menos un punto de la Tierra donde el viento horizontal es exactamente nulo.

37.12. $\chi = \sum(-1)^k \operatorname{rango} H_k$, la suma alternada de los números de Betti. Es lo que garantiza que χ no dependa de la triangulación elegida.

Capítulo 38. Teoría de categorías

38.1. En **Set**, $\varnothing$ y $\{*\}$. En **Grp**, el trivial en ambos casos: es objeto cero. En **Ring** con unidad, $\mathbb{Z}$ y el anillo nulo.

38.2. El producto topológico con la topología producto; el coproducto es la unión disjunta.

38.3. En **Grp**, los monomorfismos sí son exactamente los inyectivos, y —cosa poco obvia— los epimorfismos también son exactamente los sobreyectivos.

38.4. La inclusión $\mathbb{Z} \hookrightarrow \mathbb{Q}$ es epimorfismo en **Ring** y no es sobreyectiva: dos homomorfismos que coinciden en $\mathbb{Z}$ coinciden en $\mathbb{Q}$, porque la imagen de $1/n$ queda forzada.

38.5. Pleno: sobreyectivo en los conjuntos de morfismos. Fiel: inyectivo en ellos. El olvido **Grp** $\to$ **Set** es fiel y no pleno; la inclusión de los grupos abelianos en los grupos es plena y fiel.

38.6. $\mathrm{Nat}(\mathrm{Hom}(A, -), F) \cong F(A)$. Significa que un objeto queda determinado por el conjunto de morfismos que salen de él hacia todos los demás.

38.7. Covariantemente no; contravariantemente sí, representado por el conjunto de dos elementos, que actúa como clasificador de subobjetos.

38.8. El funtor espacio vectorial libre, que a cada conjunto le asigna el espacio de combinaciones lineales formales de sus elementos.

38.9. Los adjuntos por la derecha preservan límites; los adjuntos por la izquierda preservan colímites. Por eso el olvido, que suele ser adjunto por la derecha, respeta productos y no coproductos.

38.10. Una familia de morfismos $\eta_X : F(X) \to G(X)$ que conmuta con toda flecha. El ejemplo: $\eta_V : V \to V^{**}$ que manda v a la evaluación en v; conmuta con toda aplicación lineal y no requiere elegir base.

38.11. Porque construir un isomorfismo $V \to V^*$ exige elegir una base, y la elección no conmuta con los cambios de base. El de $V \to V^{**}$ se define sin elegir nada, mandando v a "evaluar en v", y por eso conmuta con todo.

38.12. Significa que dos objetos con la misma propiedad universal admiten un único isomorfismo compatible con las estructuras, no simplemente que sean isomorfos. Basta porque en la práctica matemática nunca se necesita más: todo lo que se puede decir del objeto en el lenguaje de la categoría es invariante por ese isomorfismo.

Apéndice F

Bibliografía comentada y siguientes pasos

Esta guía te da el mapa y el equipamiento básico; los siguientes libros son las montañas mismas. Están ordenados por tema y, dentro de cada tema, del más accesible al más profundo.

Cálculo y análisis

Spivak, *Calculus.* El mejor libro para *redescubrir* el cálculo de una variable con rigor. Sus problemas son legendarios.

Apostol, *Calculus* (2 vols.). Cubre además álgebra lineal y multivariable, con la integral antes que la derivada. Estilo sobrio y completo.

Marsden y Tromba, *Cálculo vectorial.* El texto estándar para los capítulos 9 y 11 de esta guía, con abundante geometría.

Rudin, *Principios de análisis matemático.* Cuando quieras las demostraciones completas de todo lo que aquí se enunció. Exigente y perfecto.

Álgebra lineal

Strang, *Introduction to Linear Algebra.* Intuición geométrica y aplicaciones; sus cursos en MIT OpenCourseWare son un complemento excelente.

Axler, *Linear Algebra Done Right.* Espacios vectoriales y operadores sin determinantes hasta el final; ideal como segunda pasada, con vistas al álgebra abstracta.

Friedberg, Insel y Spence, *Álgebra lineal.* Equilibrio entre teoría y cálculo, muy usado en cursos formales.

Ecuaciones diferenciales

Boyce y DiPrima, *Ecuaciones diferenciales y problemas con valores en la frontera.* El clásico de referencia; cubre todo el capítulo 12 con cientos de ejercicios.

Arnold, *Ordinary Differential Equations.* El punto de vista geométrico (flujos, retratos de fase); cambia la manera de pensar las EDO.

Strauss, *Partial Differential Equations.* Introducción seria a las EDP: calor, onda, Laplace, Fourier.

Cálculo tensorial y geometría

Fleisch, *A Student's Guide to Vectors and Tensors.* La entrada más suave al mundo de los índices; excelente para primeras intuiciones.

Do Carmo, *Geometría diferencial de curvas y superficies.* La geometría de superficies en $\mathbb{R}^3$, puente ideal antes de la geometría riemanniana abstracta.

Schutz, *A First Course in General Relativity.* Enseña el cálculo tensorial "con las manos" y lo aplica de inmediato a la relatividad.

Lee, *Introduction to Smooth Manifolds.* El tratamiento moderno y completo de variedades, formas diferenciales y el teorema de Stokes generalizado.

Álgebra abstracta y teoría de Galois

Gallian, *Contemporary Abstract Algebra.* Amable y lleno de ejemplos; ideal para el capítulo 25.

Dummit y Foote, *Abstract Algebra.* La enciclopedia: grupos, anillos, cuerpos y Galois con todo detalle. Libro de cabecera.

Stewart, *Galois Theory.* Monografía dedicada, histórica y muy legible; la continuación natural del capítulo 26.

Artin, *Algebra.* Álgebra con alma geométrica y lineal; su capítulo de Galois es una joya.

Variable compleja

Needham, *Visual Complex Analysis.* La geometría de las holomorfas contada con imágenes; único en su especie.

Ahlfors, *Complex Analysis.* El clásico riguroso de referencia.

Stein y Shakarchi, *Complex Analysis.* Moderno, limpio, con aplicaciones a números primos.

¿Y después?

Con lo cubierto en esta guía tienes la base para varias rutas: **análisis real y funcional** (medida e integración de Lebesgue, espacios de Hilbert — la versión rigurosa e infinito-dimensional del capítulo 7); **geometría diferencial y relatividad general** (continuación directa del capítulo 36); **teoría de números algebraica** (donde la teoría de Galois es herramienta diaria); **topología algebraica** (que convierte en teoría sistemática el fenómeno del campo con un agujero del capítulo 11); y **física matemática** (mecánica clásica geométrica, teoría cuántica), donde todos estos hilos se trenzan.

La matemática se aprende despacio y se disfruta siempre. Buen viaje.

Medida e integración

- **Bartle, *The Elements of Integration and Lebesgue Measure*.** El camino más corto y limpio a los teoremas de convergencia; ideal como primer contacto.
- **Royden–Fitzpatrick, *Real Analysis*.** El texto clásico de un curso completo de medida, con la recta real como protagonista.
- **Stein–Shakarchi, *Real Analysis* y *Fourier Analysis* (Princeton Lectures III y I).** La pareja que sigue exactamente el espíritu de nuestro capítulo de medida: medida y Fourier como una sola historia, con las aplicaciones aritméticas (sumación de Poisson, theta, zeta) desarrolladas.
- **Rudin, *Real and Complex Analysis*.** Para releer todo el libro —medida, L^p, Fourier y variable compleja— unificado en un solo edificio austero; la meta de madurez.

Teoría de números (capítulo 3)

- **Silverman, *A Friendly Introduction to Number Theory*.** Exactamente el espíritu de nuestro capítulo: lento, motivado por experimentos, escrito para entender; el mejor primer libro del tema.
- **Davenport, *The Higher Arithmetic*.** Un clásico breve y luminoso, accesible y a la vez profundo.
- **Hardy y Wright, *An Introduction to the Theory of Numbers*.** La referencia enciclopédica para cuando quieras ir mucho más lejos, incluida la demostración completa del teorema de las sumas de dos cuadrados y muchísimo más.
- **Niven, Zuckerman y Montgomery, *An Introduction to the Theory of Numbers*.** Un curso universitario estándar, riguroso y completo, ideal como segundo paso.

Combinatoria, grafos y métodos numéricos (capítulos 5, 8 y 14)

- **Graham, Knuth y Patashnik, *Concrete Mathematics.*** El libro que enseña a manipular sumas, coeficientes binomiales y funciones generatrices como quien toca un instrumento. Los márgenes con anotaciones de estudiantes son parte del encanto.
- **Wilf, *generatingfunctionology.*** Doscientas páginas dedicadas a una sola técnica, la de nuestra última sección de combinatoria, y descargable gratis desde la página del autor.
- **Bondy y Murty, *Graph Theory.*** La referencia moderna y completa; empieza donde termina nuestro capítulo y llega hasta los teoremas grandes.
- **Burden y Faires, *Análisis numérico.*** El estándar universitario, con pseudocódigo para todo y análisis de error cuidadoso.
- **Trefethen y Bau, *Numerical Linear Algebra.*** Escrito en cuarenta lecciones, cada una una joya; enseña por qué el condicionamiento y la estabilidad importan más que el número de operaciones.

Probabilidad, estadística y procesos estocásticos (capítulos 16 a 22)

- **Blitzstein y Hwang, *Introduction to Probability.*** El mejor primer libro de probabilidad que existe hoy: intuitivo, riguroso y con el curso de Harvard en abierto.
- **Wasserman, *All of Statistics.*** Un curso entero de inferencia, no paramétricos y remuestreo en cuatrocientas páginas densas y honestas; el complemento natural de nuestros capítulos de estadística.
- **Faraway, *Linear Models with R.*** Modelos lineales y generalizados desde la práctica, con diagnósticos y todos los avisos sobre lo que puede salir mal.
- **Gelman y otros, *Bayesian Data Analysis.*** La referencia del análisis bayesiano moderno, con la elección de prioris y el diagnóstico de cadenas tratados con la seriedad que merecen.
- **McElreath, *Statistical Rethinking.*** Reconstruye la estadística desde cero con mentalidad bayesiana y causal; el mejor libro para curarse de los reflejos del contraste automático.
- **Efron y Tibshirani, *An Introduction to the Bootstrap.*** De los inventores del método; explica no sólo cómo se hace sino por qué funciona.
- **Ross, *Introduction to Probability Models.*** Cadenas de Markov, Poisson y colas con centenares de ejemplos aplicados.
- **Øksendal, *Stochastic Differential Equations.*** La entrada estándar al cálculo de Itô, con Black–Scholes deducido en el camino y sin exigir teoría de la medida avanzada de entrada.

Topología y análisis funcional (capítulos 28, 32 y 33)

- **Munkres, *Topología.*** El libro de topología general por antonomasia: paciente, completo y con los contraejemplos en su sitio.
- **Brezis, *Functional Analysis, Sobolev Spaces and PDE.*** Análisis funcional escrito por alguien que lo usa para resolver ecuaciones; la segunda mitad justifica la primera.
- **Kreyszig, *Introductory Functional Analysis with Applications.*** La opción más accesible, con las demostraciones desmenuzadas.
- **Körner, *Fourier Analysis.*** Cien capítulos cortos, cada uno una historia; probablemente el libro de matemáticas más divertido de esta lista.
- **Strang y Nguyen, *Wavelets and Filter Banks.*** La entrada práctica a las ondículas, con la conexión con el procesamiento de señales hecha explícita.

Optimización convexa (capítulo 15)

- **Boyd y Vandenberghe, *Convex Optimization.*** La referencia indiscutible, gratuita en la web del autor, con la dualidad y la programación semidefinida tratadas con una claridad que no tiene rival.
- **Nocedal y Wright, *Numerical Optimization.*** Para el lado algorítmico: Newton, cuasi-Newton, puntos interiores y todo lo que hace falta para que las cosas converjan de verdad.
- **Rockafellar, *Convex Analysis.*** El clásico severo donde viven los subgradientes y la conjugación de Fenchel; difícil y definitivo.

Álgebra avanzada: representación y categorías (capítulos 27 y 38)

- **Fulton y Harris, *Representation Theory: A First Course.*** Empieza con ejemplos concretos —S_n, $SU(2)$— y llega hasta las álgebras de Lie; exactamente el orden de nuestro capítulo, ampliado.
- **Serre, *Linear Representations of Finite Groups.*** Breve, elegante e implacable; la primera parte se lee en una semana y contiene todo lo esencial, Burnside incluido.
- **Leinster, *Basic Category Theory.*** La introducción moderna más legible, con Yoneda explicado tres veces desde ángulos distintos y descargable gratis.
- **Riehl, *Category Theory in Context.*** El paso siguiente, con los ejemplos sacados de toda la matemática y un tratamiento serio de límites y adjunciones.
- **Mac Lane, *Categories for the Working Mathematician.*** La fuente original, todavía insuperada; conviene llegar a ella con ejemplos ya en la cabeza.

- **Weibel, *An Introduction to Homological Algebra.*** Cuando las sucesiones exactas de nuestro capítulo pidan continuación: funtores derivados, Tor, Ext y sucesiones espectrales.

Geometría algebraica y topología algebraica (capítulos 35 y 37)

- **Cox, Little y O'Shea, *Ideals, Varieties and Algorithms.*** Geometría algebraica computacional: bases de Gröbner, el Nullstellensatz y todo comprobable con un ordenador. El primer paso natural.
- **Reid, *Undergraduate Algebraic Geometry.*** Ciento veinte páginas que llevan de las cónicas a las curvas elípticas sin pedir esquemas.
- **Hartshorne, *Algebraic Geometry.*** El libro con el que la disciplina se aprendió durante cuarenta años; los esquemas de nuestra última sección empiezan de verdad en su capítulo II.
- **Hatcher, *Algebraic Topology.*** Grupo fundamental, recubrimientos, homología y cohomología, con dibujos por todas partes y descargable gratis. Probablemente el mejor libro de matemáticas disponible en abierto.
- **Bott y Tu, *Differential Forms in Algebraic Topology.*** Cuando quieras ver la cohomología desde el lado del cálculo; Mayer–Vietoris se vuelve una herramienta cotidiana.

Teoría analítica de números (capítulo 34)

- **Apostol, *Introduction to Analytic Number Theory.*** Funciones aritméticas, caracteres de Dirichlet y el teorema de los números primos, todo demostrado con paciencia y sin trucos ocultos.
- **Stopple, *A Primer of Analytic Number Theory.*** La entrada más amable, con muchos experimentos numéricos que invitan a programar.
- **Iwaniec y Kowalski, *Analytic Number Theory.*** La referencia de investigación, para cuando las cribas de nuestra última sección pidan el aparato completo.
- **Derbyshire, *Prime Obsession.*** Divulgación de la buena sobre la hipótesis de Riemann, con capítulos matemáticos e históricos alternados; excelente lectura de acompañamiento.

Lógica, conjuntos y computabilidad (capítulos 2 y 1)

- **Enderton, *A Mathematical Introduction to Logic.*** Completitud, compacidad e incompletitud demostradas con todo detalle y sin prisa.
- **Sipser, *Introduction to the Theory of Computation.*** Máquinas de Turing, indecidibilidad y complejidad, con una claridad expositiva que ya es legendaria.

- **Jech, *Set Theory.*** La referencia sobre ordinales, cardinales, forcing e independencia de la hipótesis del continuo.
- **Nagel y Newman, *El teorema de Gödel.*** Cien páginas de divulgación honesta que explican la idea sin mentir; el mejor punto de partida.
- **Smullyan, *Gödel's Incompleteness Theorems.*** Breve, técnico y escrito por alguien que además hacía juegos de lógica; la combinación funciona.

Índice alfabético

Abel, teorema de, 259
Abel–Ruffini, teorema de, 1362
conjunto abierto, 1445, 1502
continuidad absoluta, 1542, 1658, 1659
acción de grupo, 1169, 1181
aceleración centrípeta, 543
acimut, 551
ajuste de curvas, 810
al-Kindi, 121
Banach–Alaoglu, teorema de, 1722
alcance de un proyectil, 538
aleatorización, 1039
álgebra de sucesos, 882
elemento algebraico, 1328
número algebraico, 1505
algoritmo de Euclides, 57, 1201, 1274
aliasing, 1755, 1774
amplitud de una onda, 681, 1525
análisis de frecuencias, 121
análisis de la varianza, 1026, 1030
análisis exploratorio, 949
función analítica, 276, 1564
ángulo doble, fórmulas del, 258
ángulo ψ, 563
anillo, 100, 248, 364, 1186, 1226, 1232, 1386
anillo de coordenadas, 1817
anillo local, 1808
anillo local, 1187
ANOVA de dos factores, 1036
apalancamiento, 1043
aplicación abierta, teorema de la, 1705
apretón de manos, lema del, 396
aproximación, 790
árbol (grafo), 405
árbol de recorrido, 402
argumento de un complejo, 151
argumento diagonal, 1505
argumento principal, 143
argumento válido, 12
arista, 386
aristas paralelas, 387
aritmética modular, 91
armónica, función, 595
función armónica, 1565
armónica conjugada, 1558, 1565
armónicos esféricos, 767
Arzelà–Ascoli, teorema de, 1527, 1598, 1678
asimetría, 960
asíntota, 256
elementos asociados, 1216
asociatividad, 1142
astroide, 545, 546
autoadjunto, 1734
operador autoadjunto, 629, 1700
automorfismo de cuerpos, 1333
ecuación autónoma, 727
autovalor, 359, 710, 1678, 1879
autovector, 367, 1716
axioma de elección, 43
axiomas de separación, 1468

Baire, teorema de, 1705
balance detallado, 1123
espacio de Banach, 255, 1207, 1549, 1715
Banach–Steinhaus, teorema de, 1705
barras y estrellas, 203
base, 317, 680, 1317, 1678, 1837
base de entornos, 1487
base de numeración, 1505, 1628
base ortonormal, 1728
Bayes, teorema de, 198
Bayes empírico, 1064
intervalo BCa, 1093
criterio de Bendixson–Dulac, 720
Benjamini–Hochberg, procedimiento de, 994

Bernoulli, Jakob, 528
polinomios de Bernstein, 261, 1532
desigualdad de Bessel, 244, 695, 1694
funciones de Bessel, 767
Beta, función, 479
identidad de Bézout, 60, 352, 1196, 1276, 1324, 1332, 1392
Bézout, identidad de, 99
Bézout, teorema de, 1814
bicondicional, 4
binomial, distribución, 199
binomio de Newton, 226
binomio de Newton, 194
bisección, método de, 226, 800
biyección, 20
Black–Scholes, ecuación de, 1122
bolas y urnas, 183
Bolzano, teorema de, 196, 800
Bolzano–Weierstrass, teorema de, 226, 440, 715, 1527, 1599
bondad de ajuste, 914
corrección de Bonferroni, 1033
Bonferroni, corrección de, 994
segundo teorema del valor medio, 269
Bonnet–Myers, teorema de, 1879
bootstrap, 1082
conjunto boreliano, 1618
bosque, 387
botella de Klein, 1463
transformación de Box–Cox, 971
braquistócrona, 539
Brouwer, teorema de, 1904
Brouwer, teorema de punto fijo, 1475
Brouwer, teorema de, 1905
algoritmo de Buchberger, 1818
principio de buena ordenación, 25
ecuación de Burgers, 765
Burnside, lema de, 1169

cadena de Markov, 1108
cálculo de variaciones, 539
descomposición de Calderón–Zygmund, 1763
calendario gregoriano, 97
cambio de base, 1841
cambio de variables, demostración, 494
camino, 1894
camino en el plano complejo, 1569
camino crítico, 424
camino en un grafo, 398
camino euleriano, 398
camino mínimo, 418
camino mínimo sin pesos, 402
campana de Gauss, 902
campo conservativo, 582
campo de direcciones, 718
campo de direcciones, 678
campo de Killing, 1876
campo escalar, 577
campo vectorial, 577
cancelación catastrófica, 813
conjunto de Cantor, 1504, 1626
argumento diagonal de Cantor, 1504
Cantor, argumento diagonal de, 20
Cantor, función de, 1672
teorema de Cantor–Schröder–Bernstein, 23
cara de un grafo plano, 407
caracol de Pascal, 545, 546
carácter de Dirichlet, 1788
carácter de una representación, 1259, 1406
característica de un cuerpo, 195, 710, 1324, 1667
característica de Euler, 1482, 1893
característica de un anillo, 1210, 1232
características, 753, 784
Cardano, fórmula de, 139
cardinal, 43, 1506
cardinalidad, 1388
cardinalidad, 20
cardioide, 520, 527
casos favorables, 879
casus irreducibilis, 1392
Catalan, números de, 202
categoría, 1921, 1923
categoría abeliana, 1931
categoría aditiva, 1931
categoría localmente pequeña, 1939
categoría pequeña, 1939
Cauchy, fórmula integral de, 1575
teorema de Cauchy, 1176
sucesión de Cauchy, 225, 1495, 1617, 1688, 1883

ecuaciones de Cauchy–Riemann, 1555
Cauchy–Schwarz, desigualdad de, 339, 486, 1649, 1692
Cayley, teorema de, 1164
teorema de Cayley, 1176
Cayley, tabla de, 1141
celda, 1913
centralidad, 428
centralidad de cercanía, 428
centralidad de grado, 428
centralidad de intermediación, 428, 429
centroide, 291
ceros de la función zeta, 1786
ceros triviales, 1790
conjunto cerrado, 1502, 1548
sumación de Cesàro, 1761
CFL, condición, 786
desigualdad de Chebyshev, 205
funciones de Chebyshev, 1792
contraste χ^2, 914
chi-cuadrado de Pearson, 1084
teorema chino del resto, 100, 1168, 1209
choque, 765
símbolos de Christoffel, 1849
Church–Turing, tesis de, 42
ciclo (aritmética modular), 97
ciclo (permutación), 249, 713, 1150, 1160, 1228, 1390
ciclo hamiltoniano, 404
ciclo hamiltoniano, 404
ciclo impar, 412
cicloide, 522, 538
polinomio ciclotómico, 1392
cifrado afín, 122
cifrado de César, 121
cifrado de Vigenère, 122
cinta de una máquina, 38
circuito euleriano, 399
circulación, 579
circunferencia, 1903
circunferencia de Apolonio, 154
circunferencia en polares, 524
circunferencia osculatriz, 542
circunferencia unitaria, 228, 489, 648, 713, 1246, 1595
clase de equivalencia, 15
clase de restos, 76, 89
clase lateral, 1154, 1165, 1243
clasificación de EDP, 743, 750
clausura de un conjunto, 1368, 1434, 1533, 1716
clausura algebraica, 1398
clave privada, 123
clave pública, 123
Clebsch–Gordan, regla de, 1415
clique, 430, 431
cobertura, 1013
teorema de Cochran, 1032
grupo cociente, 1156, 1167
código de barras, 98
coeficiente binomial, 192
coeficiente de agrupamiento, 428
coeficiente de determinación, 810, 957, 1018, 1027
coeficiente de variación, 976
coeficientes de Fourier, 1684
coeficientes indeterminados, 735
cohomología, 1906
cola (estructura), 402
cola $M/M/1$, 1118
colatitud, 551
colimite, 1928
coloración de grafos, 415, 416
combinación con repetición, 183
combinación convexa, 854
combinaciones con repetición, 203
combinatoria, 181
compacidad, 1452
compacidad, 250, 439, 1389, 1433, 1502, 1514, 1626, 1689, 1883
operador compacto, 1716
comparaciones múltiples, 994
compatibilidad de congruencias, 94
complejo de cadenas, 1907
completar el cuadrado, 1349
completitud, 249
completitud, 212, 1493, 1505, 1642, 1693, 1883
complexificación, 369
componente conexa, 402
componente conexa, 387

componentes principales, 368
composición de funciones, 4
concatenación de caminos, 1894
concavidad, 239
condensación de Cauchy, 258
condición inicial, 748
condicional completa, 1070
condicionamiento, 845
condiciones de contorno, 759
condiciones KKT, 853
conectiva lógica, 4
conectivos, conjunto completo de, 52
conexión afín, 1887
conexo, 1454
conexión, 59, 326, 600, 656, 1388, 1455, 1574, 1870
aplicación conforme, 1588
congruencia, 89
congruencia, 74, 90
congruencia lineal, 99
cónica en polares, 528
conjugado complejo, 1614
conjugado de un complejo, 149
conjunto, 3
conjunto algebraico afín, 1803
conjunto cerrado, 1514
conjunto convexo, 852
conjunto de Cantor, 1512
conjunto de nivel, 501
conjunto decidible, 38
conjunto dirigido, 1488
conjunto G_δ, 1519
conjunto independiente, 430
conjunto potencia, 3
conjunto recursivamente enumerable, 38
cono, 1940
cono convexo, 866
cono de luz, 786
consistencia, 31
contenido de Jordan, 493
contenido de un polinomio, 1211
tabla de contingencia, 966
continuación analítica, 1778, 1790
continuidad, 1447
continuidad, 212, 452, 632, 698, 1440, 1512, 1574, 1647, 1699, 1887
continuidad en el plano complejo, 1560
continuidad absoluta, 1541
continuidad uniforme, 1499
contracción, 483, 804, 1064
contracción de índices, 1836
principio de contracción, 484, 1522, 1840
contracción, principio de, 1711
contradicción, demostración por, 18
demostración por contradicción, 250
contraejemplo, 19
contrarrecíproco, 11
contraste de hipótesis, 910, 991
contraste del signo, 1075
componentes contravariantes, 1834, 1839
convergencia absoluta, 244, 1533, 1655
convergencia cuadrática, 802
convergencia de un método iterativo, 820
convergencia débil, 1722
convergencia dominada, 1636
convergencia dominada, 1597, 1633
convergencia monótona, 226, 1526, 1634
convergencia uniforme, 253, 485, 702, 1524, 1633
convexidad fuerte, 872
conjunto convexo, 485, 625, 1693
convolución, 1753
convolución de Dirichlet, 1798
Cook, distancia de, 1043
teorema de Cook–Levin, 45
algoritmo de Cooley–Tukey, 1760
coordenadas cilíndricas, 448, 512, 551
coordenadas esféricas, 448, 551, 552
coordenadas normales, 1874
coordenadas polares, 525
coordenadas polares, 282, 448, 467, 514, 1597, 1846
números coprimos, 68
coproducto, 1923, 1928
corrección de un sistema, 31
corrección por continuidad, 903
corrección por población finita, 939
coeficiente de correlación, 924
correlación de Pearson, 957, 963

corte de rama, 1568
corte mínimo, 421
componentes covariantes, 1834
covarianza, 924, 957
covector, 365, 1837
V de Cramér, 965
Cramér–Rao, cota de, 997
criba, 1787
criba de Selberg, 1795
criptografía, 121
criterio de comparación, 244
criterio de divisibilidad, 96
criterio de información de Akaike, 1029
cuadrado latino, 1039
cuadratura, 812
cuantificador, 14
cuantil, 943, 976
cuartil, 943
cuasivarianza, 948
cuaterniones, 1305
cuatro colores, teorema de los, 415
cúbica, ecuación, 139
cubrimiento abierto, 1452, 1626
cuello de un grafo, 387
cuerpo, 89, 248, 296, 486, 1160, 1186, 1268, 1317, 1323
cuerpo de descomposición, 1309, 1322, 1336
cuerpo de fracciones, 1210
cuerpo de fracciones, 1301
cuerpo fijo, 1337
cuerpo finito, 1205, 1279, 1387
cuerpo ordenado, 248
curtosis, 960
curva cerrada, 567
curva de nivel, 457, 592, 712
curva elíptica, 1813, 1815
curva integral, 640
curva parametrizada, 578
curva plana, 1809
curva polar, 517
curva regular, 567
curva simple, 567
curvatura, 542, 543, 623
curvatura, 573, 1595, 1854
curvatura en polares, 563
cúspide, 542
CW, complejo, 1913

DAG (grafo dirigido acíclico), 424
datos apareados, 1002
datos de Cauchy, 784
datos faltantes, 971
De Moivre, fórmula de, 146, 156, 255
de Moivre–Laplace, teorema de, 903
De Morgan, leyes de, 4
cortaduras de Dedekind, 1206, 1269, 1317
Dedekind, teorema de, 1352
Dedekind, método de, 1352
delta de Kronecker, 1832
deltoide, 545
demostración directa, 18
demostración formal, 31
densidad conjunta, 916
conjunto denso, 1434, 1502, 1517
derivada, 228, 360, 451, 598, 654, 1387, 1500, 1513, 1588, 1647, 1707, 1839
derivada compleja, 1561
derivada covariante, 652, 1849
derivada direccional, 458, 639
derivada numérica, 804
derivada parcial, 442, 1716
derivadas cruzadas, 452
descenso por gradiente, 463, 861, 862
descomposición polar, 361
descomposición sesgo-varianza, 979
desigualdad triangular, 1522
desviación absoluta mediana, 948
determinante, 323, 468, 616, 704, 1227, 1842
transformada discreta de Fourier, 1760
dominio de factorización única, 85, 260, 1200, 1287
diagonal principal, 303
diagonalización, 333, 690, 1866
diagrama, 1940
diagrama de árbol, 890
diagrama de caja, 955, 960
diagrama de Hasse, 26
diagrama de Young, 1414
diámetro de un grafo, 428
diferenciación de Lebesgue, 1658
diferencial, 446

diferencial racional, 1824
Diffie Hellman, intercambio de claves, 117
difusión, 756
dígito de control, 98
Dijkstra, algoritmo de, 418
dimensión, 363
Dini, teorema de, 1526
dominio de ideales principales, 1207
Dirac, teorema de, 405
función de Dirichlet, 240, 1515, 1616
Dirichlet, condición de, 748
Dirichlet, teorema de, 1789
discriminante, 339, 714, 1289, 1348
discriminante de una EDP, 742
diseño de experimentos, 1039
diseño factorial, 1036
distribución a posteriori, 1047
distribución a priori, 1047
distribución beta, 1049
distribución binomial, 894
distribución en el muestreo, 982
distribución estacionaria, 1108, 1109
distribución exponencial, 900, 1118
distribución geométrica, 937
distribución hipergeométrica, 939
distribución inicial, 1110
distribución normal, 902
distribución temperada, 1758
distribución uniforme discreta, 936
divergencia (cálculo vectorial), 244, 609, 1850
divisibilidad, 63
división euclídea, 65
divisor, 1820
divisor canónico, 1824
divisor de cero, 1196, 1268
doble inclusión, 4
dominio conexo, 582
dominio de dependencia, 755, 786
dominio de influencia, 786
dominio de integridad, 1197, 1286
dominio euclídeo, 1219
principio de invariancia de Donsker, 1125
espacio dual, 1682, 1704

EAN-13, 98
teorema de Eckart–Young, 368
ecuación cuasilineal, 746
ecuación de clases, 1176
ecuación de Laplace, 749
ecuación de ondas, 749
ecuación de Poisson, 749
ecuación del calor, 749
ecuación diferencial, 297, 654, 1522, 1600, 1728
ecuación diferencial ordinaria, 662, 745
ecuación diofántica, 79
ecuación elíptica, 742
ecuación en derivadas parciales, 742
ecuación exacta, 675
ecuación funcional de ζ, 1790
ecuación hiperbólica, 742
ecuación homogénea, 746
ecuación lineal, 667, 746
ecuación parabólica, 742
ecuación semilineal, 746
ecuación trascendente, 790
ecuaciones normales, 810
ecuaciones paramétricas, 537
eficiencia relativa asintótica, 1102
álgebras de Eilenberg–Moore, 1938
convenio de Einstein, 1828, 1830
criterio de Eisenstein, 1204, 1287, 1318, 1332, 1600
elemento de longitud, 578
elemento neutro, 1142
elemento primitivo, 1209, 1322
elemento primo, 1216
teorema de eliminación, 1818
eliminación del parámetro, 537
eliminación gaussiana, 296
elipse, 279, 344, 493, 608, 684
emparejamiento, 411
ensayo de Bernoulli, 199
enteros módulo n, 1145
entorno, 1487
envolvente convexa, 854
epicicloide, 545
epígrafo, 854
epimorfismo, 1934
épsilon de máquina, 833
equicontinuidad, 1527, 1728
equivalencia de categorías, 1934

equivalencia de representaciones, 1408
equivalencia homotópica, 1914
error cuadrático medio, 979
error de redondeo, 813
error de tipo I, 991
error de tipo I, 910
error de tipo II, 991
error de truncamiento, 813
error típico, 906, 983
errores robustos de White, 1038
escalera del diablo, 1512
escalera de Tukey, 971
esfera de Riemann, 165
espacio afín, 1804
espacio cociente, 363
espacio de estados, 1110
espacio de probabilidad, 1665
espacio dual, 1838
espacio métrico, 1433, 1517, 1686
espacio muestral, 879
espacio muestral, 196
espacio probabilístico, 883
espacio recubridor, 1478
espacio topológico, 1445
espacio vectorial, 384, 1369, 1641
especificidad, 891
espectro de un operador, 360
espectro de un anillo, 1816, 1821
esperanza matemática, 200
esperanza de ganancia, 897
esperanza matemática, 893, 897
espín, 1415
espiral de Arquímedes, 520, 528
espiral equiangular, 554
espiral logarítmica, 528
espirógrafo, 545
esquema, 1821
estabilizador, 1169
estadística no paramétrica, 1075
estadístico, 982
estadístico de contraste, 986
estadísticos de orden, 1075, 1095
estado absorbente, 1132
estimador, 979
estimador consistente, 979
estimador de máxima a posteriori, 1049
estimador por núcleos, 1089
estructura de primer orden, 49
dominio euclídeo, 1286
Euclides, teorema de, 83
lema de Euclides, 84
Euclides, algoritmo extendido de, 99
Euclides, infinitud de los primos, 82
Euler, fórmula de, 161
Euler, fórmula de, 255
función indicatriz de Euler, 108, 1231
Euler, teorema de, 108
Euler, característica de, 1482
Euler, teorema de, 115
evidencia marginal, 1059
excentricidad, 528
experimento aleatorio, 879
explosión en tiempo finito, 716
exponenciación modular, 124
exponenciación rápida, 115
exponencial compleja, 144
funtor Ext, 1936
extensión de cuerpos, 1325, 1386
extensión de isomorfismos, 1390
extensión entera, 1187
extensión normal, 1337
extensión simple, 1331
extrapolación, 807

F, distribución, 1031
factor de Bayes, 1060
factor de inflación de la varianza, 1028
factor integrante, 673, 714
factorial, 191
factorial fraccionado, 1039
factorización, 84, 346, 1201, 1286, 1376
factorización única, 85
falacia, 12
falso positivo, 887, 891
falta de memoria, 901
familia normal, 1612
familia uniparamétrica, 665
lema de Farkas, 870
fasor, 171
Fatou, lema de, 1635
núcleo de Fejér, 1761

conjugada de Fenchel, 867
desigualdad de Fenchel–Young, 867
fenómeno de Gibbs, 1740
pequeño teorema de Fermat, 106, 1165, 1206, 1244, 1390
Feynman, truco de, 477
transformada rápida de Fourier, 1760
filtración, 1132
flujo de un campo, 586, 608
flujo de un campo, 640
flujo máximo, 421
ecuación de Fokker–Planck, 1128
Ford–Fulkerson, algoritmo de, 422
forma bilineal, 1723
forma binómica, 147
forma cerrada, 633
forma cuadrática, 343, 452, 1385
forma diferencial, 593, 1870
forma escalonada, 314
forma exponencial, 144
forma polar, 151
forma polar, 143, 151
forma real de Schur, 369
fórmula bien formada, 31
fórmula de Cayley, 406
fórmula de Euler (poliedros), 407
fórmula de Little, 1118
fórmula del gancho, 1414
serie de Fourier, 260, 339, 698, 1533, 1643, 1693
transformada de Fourier, 1616
fracción continua, 127
fracciones parciales, 265, 655
frecuencia absoluta, 952
frecuencia acumulada, 952
frecuencia relativa, 952
Fredholm, alternativa de, 1719
triedro de Frenet, 555
contraste de Friedman, 1092
endomorfismo de Frobenius, 1210
Frobenius, endomorfismo de, 297, 739, 1291, 1374, 1883
reciprocidad de Frobenius, 1422
Frobenius, reciprocidad de, 1416
frontera de un conjunto, 607, 716, 1434, 1547
Fubini, teorema de, 266, 467, 632, 705, 1648
función definida a trozos, 704
función aritmética, 1784
función armónica, 626, 748, 757
función completamente multiplicativa, 1780
composición de funciones, 1138
función convexa, 852
función de clase, 1417, 1428
función de densidad, 899
función de distribución, 942
función de distribución empírica, 1075, 1086
función de masa, 892
función de pérdida, 1062
función de variable compleja, 1559
función divisor, 127
función dual, 863
función entera, 1564, 1611
función euclídea, 1219
función generatriz, 201
función impar, 1741
función inversa, 4, 480, 1883
función inyectiva, 23
función L de Dirichlet, 1788
función maximal de Hardy–Littlewood, 1763
función medible, 1622, 1665
función meromorfa, 1611
función par, 698, 1741
función periódica, 1741
función φ de Euler, 115
función sobreyectiva, 23
función soporte, 867
función zeta, 1776, 1777
funcional lineal, 1679
funciones hiperbólicas, 168
funciones ortogonales, 1742
funciones trigonométricas complejas, 168
funtor, 1921, 1922
funtor contravariante, 1923
funtor derivado, 1936
funtores adjuntos, 1928

método de Galerkin, 761
Galerkin, método de, 1724
teoría de Galois, 1139, 1317
correspondencia de Galois, 1303, 1360
extensión de Galois, 1337

grupo de Galois, 1333
proceso de Galton–Watson, 1126
función Gamma, 288
Gamma, función, 478
lema de Gauss, 1211
Gauss–Bonnet, teorema de, 1884
teorema de Gauss (divergencia), 630, 1885
Gauss, suma de, 1355
Gauss, teorema de, 610
cuadratura de Gauss–Legendre, 833
teorema de Gauss–Markov, 1032
Gelman–Rubin, 1071
generador infinitesimal, 1128
género, 1820
género de una curva, 1812
género de una superficie, 1893
geodésica, 1854
geometría analítica, 181
Gibbs, fenómeno de, 1774
Gibbs, muestreador de, 1058, 1070
índice de Gini, 967
Gödel, teoremas de, 41
gradiente, 451, 588, 595, 737, 1837
grado de una extensión, 1403
grado de separabilidad, 1387
grado de un vértice, 396
grado de una aplicación, 1903, 1905
grado de una extensión, 1325
gráfica cerrada, teorema de la, 1706
gráfica de un campo escalar, 499
gráfico cuantil-cuantil, 961
grafo, 386
grafo bipartito, 411
grafo bipartito, 411
grafo complementario, 431
grafo completo, 387
grafo conexo, 398
grafo dirigido, 387, 424
grafo dirigido acíclico, 425
grafo plano, 407
grafo ponderado, 387
grafo regular, 387
grafo simple, 387
grafos coespectrales, 434
Gram–Schmidt, 353, 1727
fórmula de Grassmann, 372
función de Green, 763
teorema de Green, 639
Green, función de, 777
Green, identidades de, 636
base de Gröbner, 1818
Gronwall, desigualdad de, 703
grupo, 317, 1137, 1143, 1195, 1226, 1338, 1877
grupo abeliano, 1143
grupo alternado, 1152, 1163
grupo cíclico, 1149, 1161, 1176, 1283, 1389
grupo de Lie, 1880
grupo diédrico, 1138, 1147
grupo fundamental, 1902
grupo fundamental, 1473
grupo resoluble, 1182
grupo simétrico, 1150, 1414
grupo simétrico, como grupo de Galois, 1399
grupo simple, 1181
grupo simple, 1171, 1265

medida de Haar, 1421
Haar, ondícula de, 1757
Hahn–Banach, teorema de, 1704
Hall, teorema de, 414
base de Hamel, 372
Hasse, cota de, 1815
espacio de Hausdorff, 1468
Hausdorff, espacio de, 1469
haz, 1821
Heine–Borel, teorema de, 440, 1451, 1528, 1687
hélice, 551, 552
descomposición de Helmholtz, 641
hessiana, 854
matriz hessiana, 446, 455
espacio de Hilbert, 1694
transformada de Hilbert, 1763
Hilbert 90, 1383
hipérbola, 255
hiperplano soporte, 870
hipocicloide, 545
hipótesis alternativa, 986
hipótesis de Riemann, 1786
hipótesis nula, 910, 986
estimador de Hodges–Lehmann, 1095
hoja de un árbol, 405

hojas de un recubrimiento, 1478
holgura complementaria, 863
función holomorfa, 1533, 1556, 1564
homeomorfismo, 1458
homogénea, los dos sentidos, 670
homogeneización, 1814
homología, 1900
homología celular, 1914
homología simplicial, 1907
homomorfismo, 1153, 1168, 1199, 1228, 1232, 1376
homomorfismo de grupos, 1176
homotopía, 1488, 1902
homotopía de caminos, 1894
fibración de Hopf, 1911
L'Hôpital, regla de, 236
intervalo de máxima densidad posterior, 1062
teorema de Hurewicz, 1911
Huygens, Christiaan, 539

ideal, 82, 1199, 1232, 1268, 1392, 1599
ideal de una variedad, 1802
ideal generado, 1187
idempotente, 372, 1216
identidad de Parseval, 1739, 1745
impedancia, 171
implicación, 11
inclusión–exclusión, 195, 884
principio de inclusión–exclusión, 203
independencia de sucesos, 198
independencia de variables aleatorias, 206
independencia lineal de caracteres, 1374
independientes e idénticamente distribuidas, 981
indeterminación, 288
índice de un campo, 631
índice de un subgrupo, 1154
índice de una curva, 1603
inducción matemática, 19, 261, 366, 1202
inducción de representaciones, 1416
inducción fuerte, 25
inferencia estadística, 878
ínfimo, 1493
información de Fisher, 996
función integrable, 1623
integración numérica, 812
integración por partes, 636, 706
integral de contorno, 1569
integral de Lebesgue, 242, 1492, 1676, 1687
integral de línea, 595, 596, 1574
integral de línea de un campo escalar, 579
integral de Riemann, 1500, 1549, 1617
integral impropia, 286, 705
integral iterada, 447
interacción, 1036
intercambiabilidad, 1090
interior de un conjunto, 1434
interpolación, 807
interpolación lineal, 807
intersección de conjuntos, 4
intervalo creíble, 1047
intervalo de confianza, 984, 990, 1013
intervalo de confianza, 907
intervalo de credibilidad, 1049
intervalo de predicción, 1035
intervalo percentil, 1083
invariante de un grafo, 434
invariante topológico, 1893
fórmula de inversión de Fourier, 1758
inverso modular, 77, 99
función inyectiva, 321, 509, 691, 1198, 1504, 1705
polinomio irreducible, 212, 1201, 1269, 1318
ISBN, 98
isoclina, 718
isomorfismo, 1923
isomorfismo, 248, 357, 703, 1153, 1168, 1205, 1280, 1322, 1694, 1869
isomorfismo de grafos, 387, 434
Itô, fórmula de, 1122
Itô, integral de, 1122

jackknife, 1093
jacobiana, 835
jacobiano, 1639
jacobiano, 264, 327, 448, 468, 705, 1833, 1847
priori de Jeffreys, 1065
desigualdad de Jensen, 854
Jensen, desigualdad de, 916
forma canónica de Jordan, 367
descomposición de Jordan, 1539
teorema de Jordan–Hölder, 1305

juego justo, 897

suma compensada de Kahan, 840
Karush–Kuhn–Tucker, condiciones de, 864
Kasiski, método de, 122
tau de Kendall, 965, 1098
leyes de Kepler, 555
Kepler, primera ley de, 529
Klein, grupo de, 1146, 1159
categoría de Kleisli, 1938
Kleisli, categoría de, 1933
ecuaciones de Kolmogórov, 1128
Kolmogórov, axiomas de, 882
contraste de Kolmogórov–Smirnov, 1074
Kolmogórov–Smirnov, contraste de, 1086
König, teorema de, 413
Königsberg, puentes de, 398
construcción de Kronecker, 1309
producto de Kronecker, 370
contraste de Kruskal–Wallis, 1092
Kummer, extensión de, 1380
fórmula de Künneth, 1909
Kuratowski, teorema de, 407

ecuación de Lagrange, 720
Lagrange, teorema de, 259, 1154, 1165, 1315
lagrangiana, 863
aproximación de Laplace, 1065
Laplace, regla de, 196
Laplace, transformada de, 705
laplaciano, 747
laplaciano, 366, 451, 627, 683, 1850
serie de Laurent, 1579
Lax–Milgram, teorema de, 1723, 1724
teorema de Lax–Milgram, 761
lazo, 1894, 1902
lazo (grafo), 387
criterio de Lebesgue, 1514
Lebesgue–Stieltjes, medida de, 1662
Legendre, símbolo de, 119, 1355
teorema de Lehmann–Scheffé, 998
Leibniz, regla de, 477
Euclides, lema de, 83
lema diagonal, 47
lemniscata, 520
levantamiento de caminos, 1478
Levi-Civita, símbolo de, 1882
ley de los grandes números, 205, 897, 898, 1666
leyes de De Morgan, 885
libre, grupo, 1248
límite, 226, 439, 574, 680, 1207, 1291, 1376, 1513, 1556, 1617, 1704, 1882
límite categórico, 1928
Lindelöf, espacio de, 1464, 1465
teorema de Lindemann, 1328, 1392
Lindley, paradoja de, 1060
línea de fases, 728
linealización, 810, 811
independencia lineal, 298, 1386
fórmula de Liouville, 640
Liouville, teorema de, 256, 637, 704, 1506, 1574, 1717
condición de Lipschitz, 676
lista de adyacencia, 387
localización, 1187
logaritmo, 243, 1246, 1598
logaritmo complejo, 1568
logaritmo complejo, 163
logaritmo discreto, 117
longitud de arco, 543
longitud de arco, 285, 523, 573, 1882
curva de Lorenz, 967
teorema de Löwenheim–Skolem, 47
lowess, 1088
lugar geométrico, 154
ecuación de Lyapunov, 370
Lyapunov, estabilidad de, 727

distancia de Mahalanobis, 969
maldición de la dimensión, 831
función de von Mangoldt, 1792
contraste de Mann–Whitney, 1074
máquina de Turing, 38, 42
martingala, 1120
teorema de Maschke, 1417
matriz, 108, 232, 297, 302, 439, 651, 704, 1137, 1195, 1268, 1828
matriz ampliada, 312
matriz antisimétrica, 303
matriz cuadrada, 302
matriz de diseño, 1016
matriz de incidencia, 387

matriz de transición, 1110
matriz diagonal, 303
matriz diagonal dominante, 820
matriz identidad, 303
matriz nula, 303
matriz semidefinida positiva, 854
matriz simétrica, 303
matriz triangular, 303
máxima verosimilitud, 988, 989
ideal maximal, 1301
sucesión de Mayer–Vietoris, 1909
máximo común divisor, 60, 68, 209
mínimo común múltiplo, 71
MCMC, 1123
media aritmética, 948
media muestral, 906
media recortada, 969
mediana, 948
mediana de una variable aleatoria, 943
conjunto medible-Jordan, 493
medida, 1619
medida cero, 1549
medida nula, 269
medida cero, 1512, 1620, 1642
medida producto, 1674
método de Gauss–Seidel, 820, 821
método de Jacobi, 820
método de la potencia, 820
método de Monte Carlo, 829, 830
método de Newton, 861
método del círculo, 1793
método numérico, 790
Metropolis–Hastings, algoritmo de, 1057
algoritmo de Metropolis–Hastings, 1123
teorema minimax, 870
mínimos cuadrados, 340, 485, 810, 958, 1695
mínimos cuadrados ponderados, 1038
Möbius, función de, 1798
Möbius, transformación de, 605, 1381, 1588
Möbius, transformación de, 165
moda, 976
modelo estándar de la aritmética, 840
modelo jerárquico, 1064
modelo lineal, 1016
modelo lineal generalizado, 1031
modelo no estándar, 47
modelo normal–normal, 1056
aritmética modular, 89, 1167, 1199
módulo, 1187, 1303
módulo de un complejo, 149
módulo libre, 1311
principio del módulo máximo, 1594
modus ponens, 12
mónada, 1933
monodromía, 1911
monomorfismo, 1934
Monte Carlo por cadenas de Markov, 1057
morfismo, 1921
morfismo de variedades, 1817
movimiento browniano, 1116
muestra, 980
muestreo, 906
muestreo aleatorio simple, 981
multicolinealidad, 1023
multigrafo, 387
multiplicadores de Lagrange, 853
multiplicadores de Lagrange, 509
multiplicidad de intersección, 1823

operador nabla, 589
Nadaraya–Watson, estimador de, 1087
negación de un enunciado, 14
método de Nesterov, 872
Neumann, condición de, 748
Newton, método de, 802
Newton para sistemas, 835
Newton, método de, 463
lema de Neyman–Pearson, 1003
nilpotente, 1216
matriz nilpotente, 311
nivel de significación, 910
noetheriano, anillo, 1213
norma, 342, 439, 706, 1196, 1291, 1526, 1642, 1686, 1846
norma de operador, 1680
norma matricial subordinada, 845
extensión normal, 1310
subgrupo normal, 1155, 1172, 1264, 1384
normal, operador, 1734
normalidad, contraste de, 961
notación de subíndices, 744

NP-completo, 431
NP-completo, 45
NP-completo, 405
núcleo de una aplicación, 355, 1153, 1168, 1199, 1246
núcleo de convolución, 1754
núcleo de Dirichlet, 169
núcleo del calor, 755
Nullstellensatz, 1802
numerabilidad, axiomas de, 1464
conjunto numerable, 1318, 1326, 1502, 1503, 1627
numerable, conjunto, 20
número ciclomático, 402
número complejo, 140, 148
número cromático, 415
número de condición, 358, 376, 821, 845, 872
número de independencia, 431
número irracional, 262
número perfecto, 127
números de Betti, 1900, 1901
números de Stirling, 183
números pseudoaleatorios, 830

objeto final, 1923
objeto inicial, 1923
ondícula, 1756
operación binaria, 1140
operaciones elementales de fila, 313
operador acotado, 1680, 1700
operador borde, 1907
operador compacto, 1717
operador diferencial lineal, 668
operador lineal, 1679
operador proximal, 865
optimización, 440
órbita, 640, 685, 1138, 1252, 1361, 1877
órbita planetaria, 529
orden de convergencia, 851
orden de un elemento, 117
orden de un elemento, 1147, 1160
orden de un grupo, 1147
orden de una EDO, 662
orden de una EDP, 745
orden medio, 1792
orden monomial, 1803
orden parcial, 15, 26
orden topológico, 424
ordinal, 43
orientabilidad, 1914
superficie orientable, 645
orientación de una curva, 578
proceso de Ornstein–Uhlenbeck, 1128
ortogonalidad, 338, 508, 698, 1302, 1556, 1657, 1683, 1694, 1880
oscilación de una función, 681, 1513, 1874
oscilación amortiguada, 171
oscilador armónico, 631, 681, 1649, 1889

P frente a NP, 45
p-valor, 991
p-valor, 910
PageRank, 428, 429
palomar, principio del, 196
Pappus, teoremas de, 291
parábola, 379, 440
paradoja de Simpson, 959
parametrización por el arco, 567
parámetro, 981
parámetro de una curva, 537
identidad de Parseval, 699, 1580, 1694
parte entera, 65
parte imaginaria, 148
parte real, 148
parte real e imaginaria de una función, 1559
partición, 15, 240, 1173, 1500, 1515
partición de un entero, 1414
Pascal, triángulo de, 194
paseo aleatorio, 1116
pendiente de una recta, 232, 440, 654, 1548
péndulo isócrono, 539
percentil, 943
periodicidad, 268
período de Gauss, 1370
periodo fundamental, 1741
períodos de Gauss, 1370
contraste de permutación, 1090
permutación, 183, 351, 1163, 1227, 1338, 1876
permutación con repetición, 192
pesos λ_d, 1795
teorema de Peter–Weyl, 1421
π, definición analítica, 268

Picard, teoremas de, 1599
Picard–Lindelöf, teorema de, 484, 658, 1381, 1579, 1728
pila (estructura), 402
pivote, 314
teorema de Plancherel, 1758
plano de Argand, 142, 150
plano fase, 735
plano proyectivo, 1813
plano tangente, 447
población, 980
Poincaré–Bendixson, teorema de, 712
núcleo de Poisson, 763
Poisson, núcleo de, 1595
sumación de Poisson, 1765
Poisson, distribución de, 895
poliedro, 1893
poliedro convexo, 407
polinomio, 228, 317, 681, 1137, 1195, 1269, 1317, 1506, 1556, 1727
polinomio característico, 333, 1728
polinomio mínimo, 361, 1204, 1328, 1382
polo (singularidad), 247, 740, 1580
polo, orden de un, 1573
dualidad de Pontriaguin, 1765
potencia compleja, 163
potencia de un contraste, 924, 991, 992
potencia exterior, 1419
potencia simétrica, 1419
potencial newtoniano, 641
distribución predictiva, 1061
prehaz, 1821, 1938
presentación de un grupo, 1248
prevalencia, 891
primera forma fundamental, 473
número primo, 67, 82, 260, 1201
primos gemelos, 1788
principio de incertidumbre, 1752
principio de reflexión, 1125
principio del máximo, 629
prior conjugado, 1047, 1048
priori impropia, 1070
probabilidad, axiomas de, 882
probabilidad condicionada, 890
probabilidad condicionada, 198
probabilidad de extinción, 1126
probabilidad total, 890, 891
problema bien planteado, 759
problema de contorno, 666
problema de Dirichlet, 637
problema de la parada, 40, 42
problema de valor inicial, 666
problema mal planteado, 759
proceso de Poisson, 1117
proceso estocástico, 1110
producto cartesiano, 4
producto categórico, 1923
producto copa, 1906
producto de Euler, 1776, 1778
producto directo, 1157, 1172
producto interno, 338, 706, 1692, 1840
producto interno de funciones, 1742
producto libre, 1913
producto libre amalgamado, 1913
producto tensorial, 1870
producto tensorial de representaciones, 1419
producto vectorial, 1880
programación semidefinida, 866
progresión aritmética y geométrica, 510
promedio esférico, 777
propagación del error, 840
propiedad de Markov, 1108
proposición, 11
proyección ortogonal, 340
proyección estereográfica, 165
proyección oblicua, 372
proyección ortogonal, 1693
prueba del nueve, 97
pseudoinversa, 376
punto aislado, 1435
punto base, 1894
punto crítico, 239, 446, 455, 485
punto de acumulación, 1435
punto de equilibrio, 512
punto de ruptura, 948
punto de silla, 446
punto fijo, iteración de, 804
punto fijo, teorema de Banach, 1711
punto fijo, teorema del, 1712
punto singular, 542

punto singular de una curva, 1812

radical de un ideal, 1805
solubilidad por radicales, 1342
radio de convergencia, 247, 275, 736, 1565
radio de curvatura, 542
radio espectral, 820
radio negativo en polares, 516
transformada de Radon, 1765
Radon–Nikodym, teorema de, 1657
raíces conjugadas, 159, 160
raíces n-ésimas, 146
raíces n-ésimas de la unidad, 156
raíz de un polinomio, 1209, 1404
raíz de la unidad, 362, 1289, 1337, 1360
raíz de un polinomio, 160
raíz primitiva, 117
raíz racional, teorema de la, 1203, 1313, 1368
rama de una función multivaluada, 1568
rama principal, 163
rango, 314, 322, 488
rango de un módulo, 1311
rango de una matriz, 359
rango (estadística), 1098
rangos medios, 1098
condición de Rankine–Hugoniot, 765
teorema de Rao–Blackwell, 998
rapidez, 542, 543
razón de verosimilitudes, 886, 1000
reciprocidad cuadrática, 119
recíproco de una implicación, 11
recorrido, 955
recorrido en anchura, 402
recorrido en profundidad, 402, 403
recta de regresión, 810, 811
recta de regresión, 957
recta del infinito, 1814
recubridor, espacio, 1477
recubrimiento, 1905, 1911
recubrimiento por vértices, 431, 432
recurrencia, 1123
recurrencia lineal, 183
recursivamente axiomatizable, 51
red (net), 1488
espacio reflexivo, 1722
región crítica, 986
regla de la cadena, 234, 451, 624, 1850
regla de la suma, 879
regla de Laplace, 879, 880
regla del cociente, 283
regla del producto, 890
regla del producto, 234, 622, 1200
regresión lineal simple, 1019
regresión logística, 1031
regresión múltiple, 1017, 1028
relación binaria, 15
relación de cubrimiento, 26
relación de equivalencia, 15
Riemann, teorema de reordenamiento, 1533
representación de un grupo, 1406
representación inducida, 1422
representación irreducible, 1406
representación lineal, 1259
representación matricial, 1408
representación regular, 1413
residuo, 810, 1020
teorema de los residuos, 109, 340, 630, 1580, 1637
residuo cuadrático, 119
residuo estandarizado, 1024
residuo estudentizado, 1024
resolución proyectiva, 1936
resonancia, 333, 710
restricción de representaciones, 1416
retracto de deformación, 1914
tensor de Ricci, 1858
regresión contraída, 1038
Riemann–Lebesgue, lema de, 1647
teorema de Riemann–Roch, 1820
Riesz, lema de, 1688
Riesz, teorema de representación de, 1715
Robin, condición de, 730
Robin, condición de, 748
extrapolación de Romberg, 833
rosa polar, 520, 528
rotacional, 610
Rouché, teorema de, 1601
RSA (criptografía), 106, 1198, 1280
ruina del jugador, 1121
ruleta (curva), 545
paradoja de Russell, 23

salto de dualidad, 863
satisfacción, relación de, 50
satisfacibilidad, 31
método de Scheffé, 1033
lema de Schur, 1417
clase de Schwartz, 1758
Schwarz, reflexión de, 1596
secante, método de la, 804
semidirecto, producto, 1255
semilado recto, 528
seno, construcción, 268
sensibilidad, 891
separable, extensión, 1337, 1397
separable, polinomio, 1397
separable, espacio, 1464
separación de variables, 706, 751, 757, 1649
serie, 240, 378, 483, 658, 1291, 1384, 1492, 1497, 1576, 1626, 1695, 1886
serie armónica, 1548
serie de Dirichlet, 1778, 1784
serie de Fourier, 751, 758, 1739
serie de potencias, 247, 694, 1565
serie derivada, 1182
serie geométrica compleja, 169
serie singular, 1793
series formales, 1211
sesgo de un estimador, 979
σ-álgebra, 882, 1618
σ-finita, 1674
signo de una permutación, 1152, 1163
Silverman, regla de, 1089
símbolos de Christoffel, 1871
simetría, 123, 262, 492, 632, 739, 1162, 1186, 1227, 1362, 1519, 1593, 1705, 1850
simetría de una curva polar, 518
simplemente conexo, 1894
dominio simplemente conexo, 647
Simpson, regla de, 812
Simpson, paradoja de, 964
simulación, 830
sin pérdida de generalidad, 20
singularidad, 635, 695, 1575, 1879
singularidad aislada, 1573
sistema completo de restos, 76
sistema de congruencias, 100
sistema de ecuaciones, 1300
sistema deductivo, 31
sistema lineal, 703
sistema ortonormal, 1742
forma normal de Smith, 1303
espacio de Sobolev, 761
sobreajuste, 812
función sobreyectiva, 323, 1198, 1506
campo solenoidal, 646
solución clásica, 745
solución de una EDO, 663
solución débil, 761
solución fundamental, 641, 763
solución general, 665
solución singular, 731
recta de Sorgenfrey, 1462
coeficiente de Spearman, 1098
correlación de Spearman, 965
Spec $\mathbb{Z}$, 1816
spline cúbico, 835
spline de suavizado, 1087
integral de Riemann–Stieltjes, 269
Stieltjes, medida de, 1662
Stirling, fórmula de, 265
teorema de Stokes, 616
Stone–Weierstrass, teorema de, 1533
problema de Sturm–Liouville, 767
representaciones de $SU(2)$, 1421
$SU(2)$, 1415
subanillo, 1187, 1232
subbase, 1438, 1462
subcuerpo, 1324
subespacio topológico, 1436
subgradiente, 861, 865
subgrafo, 387
subgrafo inducido, 387
subgrupo, 1148, 1161, 1199, 1230, 1317
subrepresentación, 1408
subsucesión, 1496
sucesión, 213, 317, 706, 1440, 1494, 1511, 1642, 1707
sucesión exacta, 1909, 1932
suceso, 879
suceso, 196
suceso complementario, 196

suceso contrario, 879
sucesos incompatibles, 879
sucesos independientes, 888, 890
estadístico suficiente, 998
suma de cuadrados residual, 1027
suma de cuadrados total, 1027
suma de Gauss, 1793
suma de Riemann, 467, 1634
suma directa de representaciones, 1408
Sunzi Suanjing, 100
superficie de revolución, 523
superficie parametrizada, 472, 473
superficie parametrizada, 585
principio de superposición, 668
superposición, principio de, 747
supremo, 1493
Sylow, teoremas de, 1162, 1253, 1377, 1576
ecuación de Sylvester, 370

distribución t de Student, 914
tabla de caracteres, 1407, 1408
tabla de contingencia, 1085
tabla de frecuencias, 952
tabla de valores, 807
tabla de verdad, 11
tamaño muestral, 993
tangente en polares, 554
tangente horizontal, 521
tangente vertical, 521
Tarski, definición de verdad, 50
teorema de Tarski, 47
tasa de error por familia, 994
tautócrona, 539
tautología, 12
serie de Taylor, 239
tensor, 625, 1840
teorema central del límite, 903
teorema chino del resto, 102
teorema chino del resto, 100
teorema de Bayes, 890
teorema de Bézout, 1812
teorema de compacidad, 30, 31
teorema de completitud, 30
teorema de convolución, 1753
teorema de Dirichlet, 1784
teorema de estructura, 1303
teorema de estructura, 1293
teorema de isomorfismo, 364
teorema de los números primos, 1776, 1785
teorema del límite central, 1667
teorema del resto, 1274
teorema del valor medio, 490
teorema fundamental de la aritmética, 84
teorema fundamental del álgebra, 160
teoremas de incompletitud, 37
teoría de colas, 1119
teorema fundamental del cálculo, 573, 702, 1542, 1658
función de Thomae, 1515
tiempo de llegada, 1125
tiempo de parada, 1120
Tietze, teorema de extensión, 1470
tipificación, 902, 954
tiro parabólico, 538
teorema de Tonelli, 494
topología, 488, 600, 1343, 1433, 1518, 1883
topología cociente, 1463
topología más fina, 1439
topología producto, 1463
funtor Tor, 1936
torsión, 555
transformación lineal, 321, 1841
transformación natural, 1921, 1923
transformada de Fourier, 1752
transformada de Laplace, 485, 739
transporte paralelo, 1882
transposición, 1150, 1163
trapecios, regla de los, 812
elemento trascendente, 1328
número trascendente, 1368, 1505, 1705
matriz traspuesta, 306
trayectoria, 1110
trayectoria, 537
trayectorias ortogonales, 718
traza, 285, 333, 634, 1292, 1367, 1539, 1888
triangulación, 1482
Tricomi, ecuación de, 770
trigonometría, 255
diferencia honestamente significativa de Tukey, 1033
Tychonoff, teorema de, 1463

umbral suave, 865
unidad de un anillo, 1165
unidad modular, 77
unidad imaginaria, 147
unión de conjuntos, 4
unitario, 1734
universal, recubrimiento, 1481
urna, 880
Urysohn, lema de, 1469

validación cruzada, 1093
valor absoluto, 57, 1201, 1274, 1524
valor atípico, 949, 960
valor intermedio, teorema del, 227, 1455
teorema del valor medio, 261, 481, 1524
valor predictivo positivo, 886
valor principal de Cauchy, 287
valores ajustados, 1020
valuación p-ádica, 73
variable aleatoria, 206, 892
variable aleatoria continua, 899
variable cualitativa, 951
variable cuantitativa continua, 951
variable cuantitativa discreta, 951
variable estadística, 950
variable ordinal, 951
variables separables, 671
variación, 192
variación acotada, 269, 1539, 1659
variación con repetición, 183
variación de parámetros, 704
variación total, 1659
varianza, 205
varianza, 199, 893
variedad afín, 1802
variedad diferenciable, 1877
operador vec, 370
vector, 318, 605, 684, 1837
vector aceleración, 542
vector aleatorio, 916
vector velocidad, 542
ventaja de la casa, 897
verosimilitud, 988, 989, 996, 1049
vértice de un grafo, 386
vértices adyacentes, 387
viajante de comercio, problema del, 404
conjunto de Vitali, 1301, 1626
operador de Volterra, 736, 1716

WAIC, 1065
contraste de Wald, 1000
teorema de Wedderburn, 1305
corrección de Welch, 1002
Weyl, tensor de, 1888
contraste de Wilcoxon, 1074
Wilcoxon, rangos con signo, 1102
teorema de Wilks, 1000
intervalo de Wilson, 1003
Wilson, teorema de, 1174
Wilson, teorema de, 127
wronskiano, 681
wronskiano, 704
wronskiano, 729

incrustación de Yoneda, 1938
Yoneda, lema de, 1930

Zariski, topología de, 1803, 1816
lema de la mariposa, 1305
función zeta de Riemann, 126
ZFC, 43
lema de Zorn, 43