e-lecto libris

textos universitarios

Juan Luis Varona Malumbres

Recorridos por la Teoría de Números

2.ª edición



matemáticas

4

RSME



Juan Luis Varona es Catedrático de Universidad en el Departamento de Matemáticas y Computación de la Universidad de La Rioja. Se licenció en Matemáticas en la Universidad de Zaragoza en 1985 y defendió su tesis doctoral en la Universidad de Cantabria en 1988. Tiene una amplia experiencia docente en Análisis Matemático, Ecuaciones Diferenciales, Métodos Numéricos y Teoría de Números, así como en los lenguajes de programación habituales en Matemáticas. Ha publicado más de 50 artículos de investigación sobre Análisis Armónico, Análisis Numérico y Teoría de Números, y ha participado en diversos proyectos de investigación financiados por el Gobierno Español. Es director de La Gaceta de la Real Sociedad Matemática Española.

Su correo electrónico es jvarona@unirioja.es y la dirección de su página web http://www.unirioja.es/cu/jvarona/.

Últimos títulos publicados en esta serie

1. B. Cascales, J.M. Mira, J. Orihuela, M. Raja
   Análisis funcional
   Papel y PDF

2. V. Calvo, A. Peris, F. Rodenas
   Diagonalización y cálculo multivariable con Mathematica
   Papel y PDF

3. G. Vera
   Variable compleja. Problemas y complementos
   Papel y PDF

4. J.L. Varona
   Recorridos por la teoría de números
   Papel y PDF

5. G. Colomé, R.M. Miró-Roig
   Álgebra lineal
   Papel y PDF

6. J. Alaminos, C. Aparicio, J. Extremera, P. Muñoz, A.R. Villena
   Cálculo
   Papel y PDF

7. F. Bombal, L.R. Marín, G. Vera
   Problemas de Análisis Matemático 1. Cálculo diferencial
   Papel

8. F. Bombal, L.R. Marín, G. Vera
   Problemas de Análisis Matemático 2. Cálculo integral
   Papel

Juan Luis Varona Malumbres

Recorridos por la Teoría de Números

Segunda edición

e-lecto libris

Comité editorial de la Real Sociedad Matemática Española
Guillermo Curbera (director),
Manuel Abellanas, Sergio Amat, David Arcoya, José Bonet, Rafael Crespo,
Guadalupe Gómez, Clara Grima, Elvira Mayordomo, Vicente Muñoz,
Rafael Sendra y Elena Vázquez

Editor por Ediciones e-LectoLibris
Pablo Mira Carrillo

A.M.S. Mathematics Subject Classification (2010):
Primaria: 11-01
Secundaria: 11Axx, 11Bxx, 11Dxx, 11Jxx, 11Mxx, 11Nxx, 94Axx
Código IBIC: PBH
Código CDU: 511

Diseño de cubierta: Juan Pedro Cascales Sandoval, Ana Martínez Martínez
y Ediciones Electolibris S.L.
Diseño y composición: Ediciones Electolibris S.L.
Compuesto con LaTeX con tipos Latin Modern

© Juan Luis Varona Malumbres

© Ediciones Electolibris, S. L. (2019)
C.I.F. B-73749186
Pablo Neruda, 7
30820 Murcia (España)
www.electolibris.es

© Real Sociedad Matemática Española (R.S.M.E.)
C.I.F. G-28833523

Todos los derechos reservados.
Queda prohibida, salvo excepción prevista en la Ley, sin la autorización por escrito de los titulares de los correspondientes derechos de la propiedad intelectual o de sus cesionarios, cualquier forma de reproducción, distribución, comunicación pública y modificación, parcial o total, en cualquier tipo de soporte o medio de esta obra. No obstante, los propietarios del *Copyright* de este libro digital autorizan a los compradores individuales del mismo a realizar copias impresas de él para su uso personal.

Segunda edición, mayo 2019

ISBN: 978-84-946150-4-7
Depósito legal: MU 523-2019

Nota editorial

Los dos productos identificados por sus respectivos ISBN más abajo, publicados por Ediciones Electolibris S. L., tienen por título

Recorridos por la Teoría de Números, segunda edición

y autor Juan Luis Varona Malumbres.

ISBN: 978-84-946150-3-0 Versión en papel.

ISBN: 978-84-946150-4-7 El libro digital, en formato PDF, contiene:

- sistema de navegación interno e hipervínculos internos para referencias cruzadas, índices y lista bibliográfica;
- hiperenlaces de acceso para las soluciones de todos los problemas propuestos a través de aplicación web (requiere suscripción aparte y acceso a internet).

Consulte las dos páginas siguientes para obtener información sobre los iconos que le permiten acceder a estos servicios.

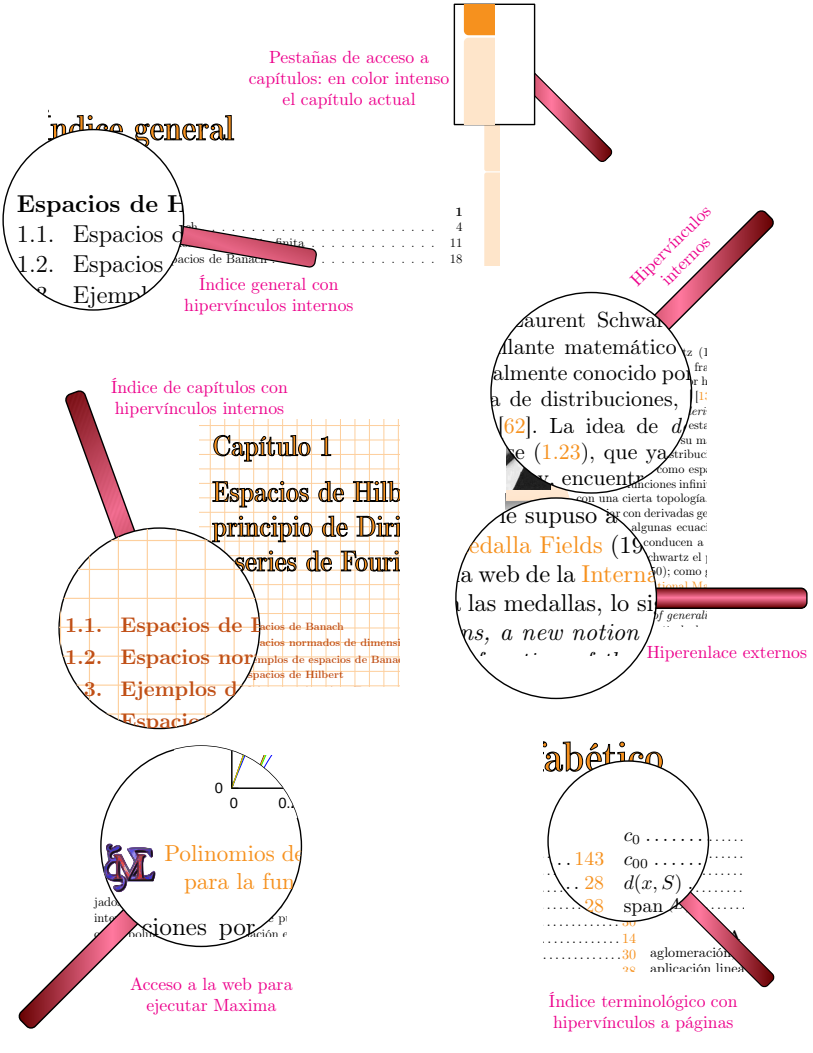
Ediciones Electolibris S.L. no posee el Copyright de todas las imágenes incluidas en este texto. Hemos utilizado todos nuestros medios para asegurarnos, más allá de la duda razonable, que estas fotografías son de dominio público. Si cree que usted tiene derechos sobre alguna de ellas, por favor contacte con ediciones@electolibris.es y trataremos de eliminarla o de citar la fuente adecuadamente.

Fotografía de cubierta original de Konrad Jacobs. Fuente: Archives of the Mathematisches Forschungsinstitut Oberwolfach.

Puede obtener información adicional sobre esta obra y otras publicadas por Ediciones Electolibris S.L. en






Pestañas de acceso a capítulos: en color intenso el capítulo actual
ndice general
Espacios de H
1.1. Espacios d
1.2. Espacios
Índice general con hipervínculos internos
Hipervínculos internos
aurent Schwar
llante matemático
almente conocido po
a de distribuciones,
62]. La idea de d
e (1.23), que ya
encuentr
le supuso a
edalla Fields (19
a web de la Interna
las medallas, lo si
ns, a new notion
Hiperenlace externos
Índice de capítulos con hipervínculos internos
Capítulo 1
Espacios de Hilb
principio de Diri
series de Fouri
1.1. Espacios de B
1.2. Espacios nor
3. Ejemplos d
Espacio
Polinomios de
para la fun
siones por
Acceso a la web para ejecutar Maxima
abético
c_0
c_{00}
$d(x, S)$
span
aglomeración
Índice terminológico con hipervínculos a páginas

A María José, Luis y David

Índice general

Índice general **7**

Prefacio a la primera edición **1**

Prefacio a la segunda edición **13**

1. Aritmética básica **15**
- 1.1. Divisibilidad . . . 18
  - 1.1.1. La identidad de Bézout . . . 21
- 1.2. Números primos . . . 24
  - 1.2.1. Factorización de un entero en producto de primos . . . 28
  - 1.2.2. Primeras consecuencias de la descomposición en factores primos . . . 29
- 1.3. El sistema de numeración posicional. Bases . . . 32
  - 1.3.1. Criterios de divisibilidad . . . 39
- 1.4. Combinatoria . . . 44
  - 1.4.1. Variaciones . . . 45
  - 1.4.2. Permutaciones . . . 46
  - 1.4.3. Combinaciones . . . 48
- 1.5. Coeficientes binomiales y binomio de Newton . . . 51
  - 1.5.1. El triángulo de Pascal . . . 55
- 1.6. Algunas ecuaciones diofánticas . . . 57
  - 1.6.1. Ecuaciones diofánticas lineales . . . 61
  - 1.6.2. Ternas pitagóricas . . . 67
  - 1.6.3. Una pequeña miscelánea de ecuaciones diofánticas . . . 75
- 1.7. Ejercicios . . . 81

2. Congruencias **87**
2.1. Definiciones y primeras propiedades 90
2.2. La ecuación con congruencias lineal 93
2.2.1. ¿Es muy costoso este algoritmo? 96
2.3. La indicatriz de Euler y el teorema de Euler-Fermat 98
2.3.1. Una expresión de la solución de una ecuación con congruencias lineal 100
2.3.2. Una fórmula explícita para evaluar la indicatriz de Euler . . 103
2.3.3. Primos de Wieferich 105
2.4. Números pseudoaleatorios 106
2.5. Primos en progresiones aritméticas 109
2.5.1. Progresiones aritméticas de primos 111
2.6. Sistemas lineales de congruencias: el teorema chino del resto 114
2.6.1. Módulos no primos entre sí 120
2.7. Teorema de Wilson 123
2.7.1. Primos de Wilson 125
2.8. Números pseudoprimos 125
2.8.1. Números de Carmichael 127
2.8.2. Test de Miller-Rabin 129
2.9. Criptografía 134
2.9.1. Un criptosistema de clave privada basado en la exponenciación discreta 137
2.9.2. Avances hacia la criptografía de clave pública 139
2.9.3. El criptosistema RSA 141
2.9.4. Ataques al criptosistema RSA 145
2.9.5. Firma digital 148
2.10. Teoría de códigos 149
2.10.1. Códigos detectores de errores 150
2.10.2. Funciones *hash* 154
2.10.3. La información es digital: códigos binarios 158
2.10.4. Códigos correctores 159
2.11. Ejercicios 162

3. Congruencias no lineales **167**
3.1. Congruencias polinómicas 169
3.1.1. Congruencias polinómicas con un módulo compuesto 172
3.2. Restos cuadráticos y ley de reciprocidad cuadrática 176
3.2.1. El símbolo de Legendre 177
3.2.2. La ley de reciprocidad cuadrática 182

3.2.3. El símbolo de Jacobi . . . 187
3.2.4. Cálculo rápido del símbolo de Jacobi . . . 191
3.2.5. Leyes de reciprocidad para potencias mayores que dos . . . 193
3.3. Pseudoprimos de Euler . . . 194
3.3.1. El test de Solovay-Strassen frente al de Miller-Rabin . . . 197
3.4. Tests de primalidad para números de formas especiales . . . 200
3.4.1. Test de Pépin . . . 201
3.4.2. Test de Proth . . . 208
3.4.3. Test de Lucas-Lehmer . . . 213
3.5. Raíces primitivas . . . 218
3.5.1. Índices modulares (logaritmos discretos) . . . 222
3.5.2. Aplicación a la resolución de ecuaciones . . . 225
3.6. Criptosistema Elgamal . . . 227
3.6.1. Firma digital con Elgamal . . . 231
3.7. Ejercicios . . . 233

4. Aproximación de un número real por racionales 237
4.1. Teorema de Dirichlet y consecuencias . . . 240
4.1.1. Un criterio de irracionalidad . . . 244
4.2. Orden de aproximación. Teorema de Liouville y consecuencias . . . 245
4.2.1. El teorema y sus consecuencias . . . 246
4.2.2. Un criterio de trascendencia . . . 248
4.3. Refinamientos del teorema de Liouville . . . 250
4.3.1. El correspondiente criterio de trascendencia . . . 252
4.3.2. Reinterpretación en términos de la derivabilidad de una función patológica . . . 255
4.3.3. Soluciones de ecuaciones diofánticas . . . 256
4.4. Fracciones de Brocot . . . 258
4.4.1. Aproximaciones óptimas . . . 262
4.5. Teorema de Hurwitz . . . 263
4.5.1. La mejor constante . . . 268
4.6. La constante de Lagrange de un irracional . . . 270
4.7. Cubrimientos de la recta real . . . 273
4.8. Ejercicios . . . 274

5. Fracciones continuas 279
5.1. La sucesión de Fibonacci . . . 281
5.2. Pares de sucesiones recurrentes . . . 286
5.3. Fracciones continuas simples . . . 288

5.4. Fracciones continuas simples finitas . . . 291
5.4.1. Unicidad del desarrollo en fracción continua finita . . . 293
5.5. Convergencia de fracciones continuas simples . . . 294
5.5.1. Unicidad del desarrollo en fracción continua infinita . . . 297
5.5.2. Un ejemplo numérico concreto . . . 299
5.5.3. ¿Las fracciones continuas proporcionan buenas aproximaciones? . . . 301
5.6. Aproximación por racionales mediante desarrollos en fracción continua . . . 303
5.6.1. La constante de Lagrange de un irracional . . . 310
5.7. Fracciones continuas simples periódicas: irracionales cuadráticos . . . 313
5.7.1. La estructura del periodo . . . 317
5.8. El desarrollo de e en fracción continua simple . . . 319
5.9. Ejercicios . . . 326

6. Irracionalidad y trascendencia 333
6.1. Algunos resultados sobre irracionalidad . . . 337
6.1.1. Expresiones trigonométricas . . . 339
6.1.2. Logaritmos . . . 341
6.2. Irracionalidad de e . . . 342
6.2.1. Introducción histórica . . . 342
6.2.2. Una demostración elemental . . . 344
6.3. Irracionalidad de π . . . 345
6.3.1. Algunas adaptaciones y generalizaciones . . . 348
6.4. Los números trascendentes existen . . . 349
6.4.1. No sólo existen, sino que hay muchísimos . . . 351
6.4.2. Pero son difíciles de encontrar . . . 359
6.5. Trascendencia de e . . . 360
6.6. Polinomios simétricos . . . 363
6.6.1. El teorema fundamental de los polinomios simétricos . . . 366
6.6.2. Algunas consecuencias sencillas, pero importantes . . . 368
6.7. Trascendencia de π . . . 371
6.7.1. Esquema de la demostración . . . 374
6.7.2. La demostración . . . 375
6.8. Extensiones y problemas abiertos . . . 378
6.8.1. El séptimo problema de Hilbert . . . 380
6.8.2. Otros resultados sobre trascendencia . . . 382
6.8.3. Tampoco es fácil probar la irracionalidad . . . 383
6.8.4. Problemas abiertos . . . 384

6.9. Polinomios de Bernoulli y $\zeta(2k)$ 385
6.10. Ejercicios 393

7. Funciones aritméticas 399
7.1. Funciones aritméticas notables 401
7.2. Números perfectos y amigos. Sucesiones alicuatorias 406
7.3. Producto de Dirichlet y transformada de Möbius 410
7.3.1. La fórmula de inversión de Möbius 413
7.3.2. Algunas relaciones entre funciones aritméticas 414
7.4. Funciones multiplicativas 416
7.4.1. Inversas de funciones completamente multiplicativas 420
7.4.2. Fórmula del producto para funciones multiplicativas 422
7.5. Derivadas de funciones aritméticas 423
7.6. Generalización del producto de Dirichlet 426
7.7. Más tipos de convoluciones y sus inversiones de Möbius 433
7.7.1. Fórmulas de inversión de Möbius con flujos 439
7.8. Ejercicios 444

8. Medias de funciones aritméticas 451
8.1. El orden de magnitud de algunas funciones aritméticas 456
8.1.1. El orden de la función $\tau(n)$ 456
8.1.2. La funciones $\sigma(n)$ y $\varphi(n)$ 459
8.2. Fórmulas de sumación 463
8.3. Estimaciones asintóticas de funciones 466
8.3.1. Algunas fórmulas elementales 468
8.4. Orden medio de las funciones divisor 472
8.4.1. Parámetro nulo (número de divisores) 473
8.4.2. Parámetro positivo 476
8.4.3. Parámetro negativo 477
8.5. Orden medio de la indicatriz de Euler 478
8.5.1. ¿Cuál es la probabilidad de que dos números tomados al azar sean primos entre sí? 481
8.5.2. Series de Dirichlet 482
8.6. Estimaciones elementales relacionadas con la función de Möbius . . 482
8.6.1. ¿Cuál es la probabilidad de que un número tomado al azar sea libre de cuadrados? 485
8.6.2. Una cota no muy fina 485
8.7. Resultados equivalentes al teorema de los números primos 487
8.7.1. Las funciones de Chebyshev 489

8.7.2. Equivalencia con las medias de la función de von Mangoldt 491
8.8. Equivalencia con las medias de la función de Möbius 494
8.9. Ejercicios 500

9. Distribución de los números primos: resultados elementales 505
9.1. Productos de Euler 507
9.1.1. La serie de los inversos de los números primos 510
9.1.2. La identidad de Euler aplicada a las series de Dirichlet . . . 511
9.2. Teorema de Chebyshev 512
9.2.1. Cinco lemas y el teorema 515
9.2.2. Cotas para el tamaño del k-ésimo primo 520
9.3. Postulado de Bertrand 521
9.3.1. Distancias entre primos 525
9.3.2. Funciones generadoras de primos 527
9.4. Fórmulas de Mertens 531
9.5. El número de divisores primos de un entero 535
9.5.1. ¿Cuántos divisores primos tiene un entero típico? 539
9.5.2. ¿Cuántos divisores tiene un entero típico? 546
9.5.3. Propiedades estadísticas 547
9.6. Ejercicios 548

10. El teorema de los números primos 553
10.1. Recordatorio de análisis complejo 557
10.2. Definición de la zeta de Riemann y resultados preparatorios 564
10.2.1. Extensión de la zeta de Riemann a la banda crítica 566
10.2.2. Extensión de algunas propiedades a la frontera derecha de la banda crítica 569
10.2.3. Sumas de logaritmos de primos 576
10.3. Un teorema tauberiano 578
10.4. El final de la demostración del teorema de los números primos . . . 587
10.5. Algunas consecuencias del teorema de los números primos 592
10.5.1. ¿Cuál es la probabilidad de que un número tomado al azar sea primo? 593
10.5.2. Fórmula asintótica para el n-ésimo primo 594
10.5.3. El tamaño asintótico de la función $\tau(n)$ 595
10.5.4. El tamaño asintótico de la función $\varphi(n)$ 599
10.6. El tamaño del error al aproximar $\pi(x)$ 602
10.6.1. Qué se sabe 604
10.6.2. Qué se sabría si la hipótesis de Riemann estuviera probada 607

10.7. Las aproximaciones que propuso Riemann 611
10.7.1. Trabajo preparatorio 614
10.7.2. Aparecen los ceros de la función zeta 618
10.7.3. Aproximación numérica y gráfica incorporando sucesivos ceros 624
10.8. Ejercicios 630

Apéndice. La construcción de los números 635
A.1. Conceptos básicos de teoría de conjuntos 637
A.1.1. Relaciones 638
A.1.2. Aplicaciones entre conjuntos 639
A.2. Definición de cuerpo 640
A.3. Los números naturales 642
A.3.1. Los axiomas de Peano 643
A.3.2. Un modelo de los números naturales 647
A.4. Los números enteros 650
A.5. Los números racionales 655
A.6. Los números reales 658
A.6.1. Cortaduras de Dedekind 661
A.6.2. Unicidad de los números reales 671
A.6.3. Otras construcciones de los números reales 676
A.7. Los números complejos 681
A.8. Cuaterniones y octoniones 686
A.9. Y aún hay más 689

Bibliografía 693

Índice alfabético 701

Prefacio a la primera edición

Éste es un libro de matemáticas, y ése es su contenido principal. Pero no hemos querido que lo sea en exclusivo, sino que hemos pretendido mostrar las matemáticas como obra de la actividad humana, en diversos momentos de la historia, y con descubrimientos que poco a poco han dado lugar a lo que ahora sabemos. Es cierto que la matemática puede estar incluida en sí misma, desvinculada del mundo exterior —quizás sea la única ciencia en la que eso es posible—, pero, en opinión del que esto escribe, no hay ninguna buena razón para estudiar matemáticas de ese modo; por el contrario, las matemáticas desprovistas de su entorno histórico y cultural, de sus aplicaciones, de los fracasos y éxitos en pos de un bonito resultado previamente intuido..., pierden gran parte de su atractivo y, desde luego, resultan más áridas para el estudiante. Así, aunque el libro no desdeña el rigor matemático, todo él está impregnado de pequeñas referencias históricas en relación a los resultados que se estén estudiando en cada momento y a los matemáticos que los descubrieron.

Otro objetivo que se ha perseguido al escribir este libro ha sido darle una gran flexibilidad de uso. Consta de diez temas y un apéndice, que son —en la medida de lo posible— lecciones sueltas sin excesiva dependencia unas de otras, a modo de recorridos por diversas partes del extenso territorio de las matemáticas en general y la teoría de números en particular. Con pocas excepciones, cada tema (y también el apéndice) puede ser leído con independencia de todos las demás; cuando esto no es así, los requisitos previos se detallan al comienzo. Eso hace que algunos conceptos aparezcan repetidos: se ha hecho conscientemente; es el precio que hay que pagar para conseguir temas autocontenidos.

El libro está dirigido, de manera primordial, a estudiantes de matemáticas a nivel de pregraduados, y es fruto de la experiencia de varios años enseñando estas materias —en diversas asignaturas y con varios planes de estudios distintos—, en los que versiones previas de lo que aquí aparece se han utilizado como apuntes de

clase. Las explicaciones son bastante detalladas, de modo que sean útiles para el lector autodidacta. Se incluyen numerosos ejercicios, todos ellos resueltos. Por otra parte, un profesor que quiera usar este libro puede elegir los temas que desee (o porciones suyas) para seguirlos en un curso, dependiendo de su interés personal y las circunstancias del curso.

No todos los temas tienen el mismo nivel. Según sean sus conocimientos previos, el lector puede optar por saltarse algunos temas por considerarlos demasiado sencillos; o renunciar a otros que requieren desenvolverse con soltura en ciertas técnicas matemáticas. En particular, el más avanzado es el dedicado al teorema de los números primos, que presupone que el lector está familiarizado con el análisis complejo; pero no queremos asustar al lector con ese tema: con unos pocos conocimientos de análisis complejo se puede seguir sin dificultades, ¡y además es realmente precioso! Ningún matemático debería perderse el placer de entender qué tiene que ver la distribución de los números primos con los ceros de la función zeta de Riemann.

La independencia entre los temas no es óbice para que —cuando aparezcan— se señalen las relaciones entre los distintos conceptos y resultados tratados en diferentes partes del libro.[1] Una de las más bellas sorpresas que nos proporcionan las matemáticas son los nexos entre campos aparentemente muy distantes, y las diversas maneras de resolver un mismo problema. Merece la pena detenerse en ello siempre que sea posible.

La estructura del libro es como sigue. Primero hay un tema de carácter introductorio, del que muchos lectores podrán prescindir si lo desean. En particular, allí se abordan los conceptos más elementales (primos, divisores...) y se prueba el teorema fundamental de la aritmética: cada número natural se descompone de manera única —salvo el orden— como producto de primos. Luego, el libro tiene tres bloques diferenciados:

- Los temas 2 y 3 están dedicados a las congruencias, y el segundo requiere algunos conceptos y propiedades que se estudian en el primero.
- Los temas 4 a 6 tratan sobre aproximación diofántica, trascendencia e irracionalidad, pero cada uno de los tres temas puede estudiarse por separado sin necesidad de los demás.
- La tercera parte —temas 7 a 10— está dedicada a las funciones aritméticas y al teorema de los números primos.

Los contenidos de estos tres bloques son totalmente independientes unos de otros. Finalmente, el libro incluye un apéndice en el que se expone cómo —partiendo de la teoría de conjuntos— se van construyendo rigurosamente los distintos conjuntos

[1] A veces, como notas a pie de página.

de números, desde los naturales hasta los complejos, pasando por los enteros, los racionales y los reales.

¿Qué estudia la teoría de números?

A grandes rasgos, podríamos decir que la teoría de números es la rama de las matemáticas que se ocupa de las propiedades de los números; sin duda, su concepto fundamental es el de número primo. Ya de forma rigurosa, comenzó a ser cultivada desde muy antiguo. Baste destacar que cuatro de los trece libros de los *Elementos* de Euclides [42] (que fueron escritos a principios del siglo III a.C.) están dedicados a ella. Aun así, la teoría de números no ha perdido actualidad, y continúan existiendo multitud de problemas cuya solución no se conoce. En palabras de Gauss (considerado unánimemente como uno de los mayores genios matemáticos de todos los tiempos, y a quien se conoce con el sobrenombre de «príncipe de las matemáticas»), las matemáticas son «la reina de las ciencias», y la teoría de números es «la reina de las matemáticas».

Y es que la teoría de números es una rama de las matemáticas muy peculiar. Las demostraciones de algunos resultados de teoría de números son muy ingeniosas, y fáciles —si es que esta palabra puede aplicarse a las matemáticas en general— de entender para quien las lee ya desarrolladas. Muchas veces, la «idea feliz» que ha generado la demostración ha sido fruto de alguna de las mejores mentes matemáticas de la historia. Estas pruebas son las que —según Erdős, uno de los matemáticos más prolíficos de la historia, y cuyo tema principal de trabajo era la teoría de números— figurarían en «El Libro», en el que un supuesto creador había desarrollado la totalidad de las matemáticas, cada teorema con su prueba óptima (véase [2]). Por el contrario, también es común que algunos enunciados aparentemente sencillos tengan demostraciones complicadísimas, que requieren el uso de herramientas matemáticas muy avanzadas y —aparentemente— muy alejadas de lo que se persigue. El paradigma es la demostración de Wiles del último teorema de Fermat (para $n \geq 3$, no existen enteros positivos x, y y z tales que $x^n + y^n = z^n$), que —finalmente— se logró en 1995, tras tres siglos y medio de esfuerzos de multitud de matemáticos. ¿Quizás alguien encontrará en el futuro una demostración fácil de entender, con matemáticas de menor complejidad? (¿digna de aparecer en «El Libro»?). Quién sabe; es lo que Fermat afirmaba —en el famoso margen de su ejemplar de la *Aritmética* de Diofanto— que había logrado, aunque no proporcionó tal demostración —porque no le cabía en el margen—, si es que realmente la tenía y no era errónea.

En teoría de números existen problemas abiertos cuyo enunciado es fácil de entender para alguien que tenga conocimientos matemáticos de nivel escolar. Esto es prácticamente imposible en otras ramas de la matemática, en las que, para conse-

guir comprender de qué se está hablando, se pueden necesitar años de estudio. Así, los problemas de la teoría de números son el principal ejemplo que tenemos los matemáticos para explicar a la sociedad que las matemáticas siguen siendo una ciencia viva: existe investigación matemática, y se ocupa de temas muy interesantes (otras ciencias, como la medicina, la biología, la astronomía... no tienen este problema; es más, sus logros aparecen muy a menudo en los medios de comunicación). A este respecto, citemos unos pocos problemas cuya solución no se conoce:

- La conjetura de Goldbach: cada número par mayor que dos puede escribirse como suma de dos primos. En 1742, Goldbach comunicó a Euler por carta que había observado esa propiedad. Pero nadie hasta ahora ha logrado probarla o refutarla.
- Denominamos perfectos a los números que son igual a la suma de sus divisores propios (por ejemplo, $6 = 1 + 2 + 3$); ¿existe algún número perfecto impar?
- Si llamamos primos gemelos a los que tienen la forma p y $p + 2$, ¿existen infinitos de ellos?
- Para $s = 5, 7, 9, 11, \ldots$, ¿son irracionales los números $\zeta(s) = \sum_{k=1}^{\infty} k^{-s}$? (Se sabe que $\zeta(3)$ es irracional, pero no si es algebraico o trascendente.)

Y hay muchos más: el lector interesado puede consultar [94] o [49]. ¡No se sentirá defraudado!

Históricamente, el desarrollo de las matemáticas se he movido por una doble vertiente: la vertiente utilitaria, respondiendo a la necesidad de resolver problemas reales y prácticos, y la vertiente más teórica que prioriza el conocimiento profundo de las situaciones, el placer de la comprensión de los conceptos y estructuras, aunque se alejen de las necesidades prácticas. Es verdad que gran parte de los temas de los que se ocupa la teoría de números no parecen muy aplicables de manera inmediata. Por ejemplo, Hardy —uno de los más famosos matemáticos de la primera mitad del siglo XX— estaba orgulloso de que la teoría de números sólo fuera útil como entretenimiento mental; su estudio proporcionaba placer en sí mismo, pero —según él— no iba a resultar nunca útil para ninguna otra cosa. ¡Pues estaba completamente equivocado! En 1978, años después de la muerte de Hardy, apareció una importante conexión entre la teoría de números y la criptografía. Utilizando bien conocidas propiedades de los números primos, Rivest, Shamir y Adleman inventaron (¿o descubrieron?) lo que ahora se conoce como criptosistema RSA, que suponía un gran avance sobre los sistemas criptográficos anteriormente conocidos. Su método no requería que el emisor y el receptor de un mensaje cifrado compartieran una clave secreta, lo cual era muy útil, pues proteger la clave es uno de los principales puntos débiles de los sistemas criptográficos; dio lugar a una revolución total en la criptografía. Por poner un ejemplo, este tipo de técnicas ha demostrado su

efectividad en temas como el comercio electrónico, que —obviamente— en la época de Hardy no existía. Como consecuencia, permítasenos decir que la investigación básica es —además de bella— útil, pero que no podemos saber cuándo surgirán sus aplicaciones en otros campos. En todo caso, lo que sí podemos asegurar es que la investigación que no se hace nunca será útil.

Diversos campos dentro de la teoría de números

Las interconexiones de la teoría de números con otras ramas de la matemática son profundas y —muchas veces— insospechadas; la teoría de números se sirve de ellas en el estudio de sus propios temas y, a su vez, contribuye con todo tipo de aportaciones y nuevos desafíos. Así, a grandes rasgos, se puede hablar de la teoría algebraica de números y de la teoría analítica de números. Aunque —como ocurre siempre en matemáticas— no hay una frontera entre ambas, la primera usa, fundamentalmente, álgebra; y, la segunda, análisis matemático. Podemos decir que el principal logro de la teoría algebraica fue la demostración del último teorema de Fermat. El mayor éxito de la teoría analítica fue la prueba del teorema de los números primos.

En este libro, la mayor parte de los temas podríamos enmarcarlos en la denominada teoría de números elemental, pues no se usan demasiadas técnicas procedentes de otros campos de las matemáticas (o las que se usan no son muy elevadas). Con todo, unos temas tienen un carácter más algebraico (los tres primeros, aunque realmente no se puede decir que sean propios de la teoría de números algebraica), y otros más analítico; es en los contenidos analíticos donde hemos llegado a demostrar resultados más profundos y avanzados —con el culmen en el último tema—. De hecho, a veces incluso ocultaremos el álgebra subyacente en algunas argumentaciones y resultados; seguro que hay a quien no le parece buena idea, pero ha sido una decisión premeditada, para no asustar a los lectores recelosos de los planteamientos excesivamente algebraicos. En estos casos, habitualmente lo hemos reseñado por medio de notas complementarias. Sí que queremos reconocer que este sesgo del libro evitando los razonamientos algebraicos ha hecho que no tratemos algunos temas de indudable actualidad, como es el caso de las curvas elípticas y su uso en criptografía.

A menudo se distingue entre teoría de números aditiva y multiplicativa, según el tipo de propiedades que se intente probar. La teoría multiplicativa estudia propiedades de números primos, divisores, funciones aritméticas...; prácticamente todo el libro trata de teoría multiplicativa. Por el contrario, el ejemplo típico de problema aditivo es la conjetura de Goldbach antes citada.

También existe lo que se ha dado en llamar teoría de números computacional; una de sus labores es desarrollar eficaces algoritmos de cálculo, por ejemplo, méto-

dos de factorización de enteros. Cada cierto tiempo, en los foros especializados se divulga un nuevo éxito de la teoría de números computacional: para alguna propiedad aritmética (ser primo, por poner otro ejemplo sencillo), se ha encontrado un entero enorme —mayor que todos los conocidos hasta entonces— que la cumple. Aunque es incluso emocionante (y podemos dar fe de ello, como atestigua [13]), apenas nos preocuparemos aquí de este tipo de resultados. Sólo en algunos lugares nos hemos permitido unas breves pinceladas sin demasiado detalle, y siempre sin preocuparnos de los récords actuales, que cambian constantemente y que, sin duda, quien esté interesado en ello podrá encontrar por sí mismo gracias a internet.

Puestos a añadir adjetivos a la teoría de números por el tipo problemas que se abordan y las técnicas que se usan, en ocasiones se habla así mismo de teoría de números geométrica y de teoría de números probabilística o combinatoria. Pero no insistiremos más en ello. En realidad, aunque todas estas ramas de la teoría de números presentan problemas muy interesantes, están poco o nada representadas en este libro, salvo en ejemplos muy elementales que no muestran ideas avanzadas ni la potencia de los métodos que allí se usan. Y lo mismo ocurre con otros temas que ni siquiera hemos mencionado. En nuestro periplo por la teoría de números, forzosamente había que dejar prometedores caminos sin recorrer y bellos parajes sin explorar.

El espíritu del libro

En ningún momento hemos querido dar una mera acumulación de resultados o técnicas de cálculo. Por el contrario, el libro está lleno de comentarios que intentan resaltar lo que se quiere hacer y por qué, así como su relevancia dentro del maravilloso mundo de las matemáticas. Las demostraciones son una parte importante del texto, pues no sólo son imprescindibles en algo que se precie de ser verdadera matemática, sino que —en gran medida— es en su elegancia donde se observa la belleza de lo que se está estudiando. No se trata de confiar ciegamente en los resultados que se enuncian, sino de entender por qué son verdad, y qué argumentos los justifican; muchas veces, incluso resulta interesante analizar cómo se han ido logrando demostraciones diversas de un mismo teorema.

Aun así, no hemos demostrado todo lo que aparece en el libro. No hace falta recordar a nadie que las matemáticas, además de bellas, acostumbran a ser difíciles y laboriosas. De algunos resultados que presentamos sólo se conocen demostraciones que salen fuera del alcance de los conocimientos aquí adquiridos, o que —aunque asequibles al nivel del libro— lo hubieran alargado en demasía. Eso sí, nunca hemos dejado algo sin demostrar dentro de la línea argumental principal de los diversos

temas tratados; sólo nos lo hemos permitido con resultados complementarios que posteriormente ya no hemos usado como fuente de otros nuevos.

Es, generalmente, al final de muchos temas (o secciones) donde hemos incluido referencias a resultados más avanzados, cuya demostración no incluimos. A menudo, también hemos ido mencionando diversos problemas abiertos relacionados con lo que se esté estudiando. En nuestra opinión, esto siempre es una buena motivación que puede animar al lector ya avezado a profundizar en alguna de las materias aquí expuestas. Sobre todo pensando en estos lectores, al principio de cada tema citamos siempre varios textos con contenidos relacionados. Algunos llevan a cabo un tratamiento similar; otros tienen enfoques diferentes, o son más especializados. En todo caso, consideramos que pueden ser un buen punto de partida para quien quiera seguir estudiando. Cuando ha sido posible, hemos preferido incluir referencias en castellano; en particular, el lector podrá ver que hay un gran número de referencias a los excelentes artículos de teoría de números que desde 1998 publica *La Gaceta de la Real Sociedad Matemática Española*.

Tal como ya hemos comentado, hemos querido enmarcar los resultados que presentamos en el contexto histórico en el que fueron demostrados y, en cada caso, procurando mencionar cuál es el matemático que ostenta el perpetuo honor de haber probado —por primera vez— cada teorema importante. Pero éste no es un libro de historia, y no hemos pretendido ser fuente de referencias históricas; aunque en numerosas ocasiones hemos hecho alusión a los autores originales de muchos de los resultados que reseñamos, así como a las fechas en que éstos fueron probados, lo habitual es que no proporcionemos la demostración original y que no citemos el artículo en el que ésa fue publicada. Si el lector está interesado en este tipo cuestiones históricas, le recomendamos [34], que proporciona un conocimiento enciclopédico de toda la teoría de números hasta 1918.

Desperdigados por las páginas de texto, se han incluido retratos de numerosos matemáticos. De los más recientes son incluso fotografías; y, por el contrario, de los anteriores al siglo XV lo habitual es que se trate sólo de imágenes sugestivas, representaciones idealizadas realizadas mucho después de la vida del personaje en cuestión, pero que han tenido éxito histórico (de hecho, el primer matemático retratado en vida fue el italiano Luca Pacioli, a finales del siglo XV, en un célebre cuadro en el que se lo representa explicando uno de los teoremas de Euclides). Pero no siempre ha sido posible incluir imágenes de los matemáticos que nos aparecerán en el libro; aunque lo importante de los matemáticos, y lo que les hace pasar a la historia, son sus resultados —y eso es producto de la mente—, es una pena que, de muchos, nos nos haya quedado una imagen que los represente físicamente. En cada caso, para averiguar dónde se encuentran los retratos de estos matemáticos basta buscar su nombre en el índice alfabético situado al final del libro: si, entre

los números de página que indican en qué lugares es citado, hay algún número en itálica, allí estará la correspondiente imagen. Algunos matemáticos que han hecho contribuciones importantes en distintos temas están representados en el libro no una, sino dos veces, con imágenes diferentes; más aun: dos figuras eminentes de la teoría de números —Euler y Gauss— se merecen sin duda aparecer en tres ocasiones (una de las veces representados en un billete).

Hemos utilizado todos nuestros medios para asegurarnos, más allá de la duda razonable, que todas las imágenes publicadas son de dominio público o, en algún sentido, se podían utilizar sin cortapisas legales. Todas las que no han sido elaboradas expresamente para este libro han sido obtenidas, de buena fe, en sitios de internet muy conocidos (el «MacTutor History of Mathematics Archive» ha sido la fuente más habitual) o, en su caso, páginas personales de los matemáticos retratados. Pero no hemos conservado una relación detallada de dónde ha sido localizada cada imagen. Si alguien cree que tiene derechos sobre alguna de las imágenes incluidas, le rogamos se ponga en contacto con este autor para poder eliminarla en sucesivas ediciones del libro.

Al final de cada tema hay bastantes ejercicios propuestos al lector, todos ellos resueltos en otro apartado del libro. La idea principal de los ejercicios es que el estudiante que sigue el libro pueda practicar sobre los conceptos y resultados que se explican, para ayudarle a entenderlos con todo rigor. Aunque algunos temas se prestan a ello, no hemos querido proponer muchos que sólo sean de cálculo (es decir, aplicaciones mecánicas de fórmulas o algoritmos diversos explicados en el texto); por el contrario, sólo hemos incluido tales ejercicios en contadas ocasiones.

En suma, el que esto escribe no ha ahorrado esfuerzos para intentar conseguir que este libro sea suficientemente atractivo y asequible. Creemos pues que un lector interesado en las matemáticas podrá seguir el texto sin excesivas dificultades. Pero, eso sí, el tiempo y la dedicación deberá aportarlos él mismo.

Prerrequisitos

Dado que el primer tema proporciona una visita general por los conceptos y propiedades de la aritmética elemental —y que para muchos lectores será un mero repaso posiblemente innecesario—, los conocimientos matemáticos requeridos para seguir el libro son escasos. Sí que se supone que el lector tiene conocimientos básicos de cálculo (análisis real en una variable). También se da por conocido el teorema fundamental del álgebra: cada polinomio de grado n tiene n raíces complejas (contando su multiplicidad).

Con eso se puede seguir casi todo el libro sin excesiva dificultad. Sin embargo, en algunos temas el lector necesitará estar familiarizado con otras técnicas más

avanzadas; sobre todo, en el tema 10, que presupone conocimientos de análisis complejo. Cuando éste sea el caso, estará indicado al comienzo del tema.

Notación

No hemos considerado necesario incluir en el libro un apartado explicando la notación y los símbolos usados en el texto (en realidad, dado que ordenar alfabéticamente la notación matemática no es sencillo, esos apartados no siempre son sencillos de manejar). Pero sí que queremos hacer ahora unos pocos comentarios al respecto.

A lo largo del texto, representaremos mediante $\mathbb{N}$ al conjunto de los números naturales; y no consideraremos al 0 como un natural. Asimismo, $\mathbb{Z}$ será el conjunto de los números enteros, $\mathbb{Q}$ el de los racionales, $\mathbb{R}$ el de los reales y $\mathbb{C}$ el de los complejos. Usaremos $\mathbb{Z}[x]$ para denotar el conjunto de los polinomios en x con coeficientes en $\mathbb{Z}$; lo mismo haremos con los otros conjuntos de números. En cuanto a los números racionales, y aunque no lo indiquemos expresamente, siempre que aparezca un racional h/k daremos por hecho que el denominador es positivo, es decir, que $h \in \mathbb{Z}$ y $k \in \mathbb{N}$. Y una advertencia importante más: cuando escribamos números con decimales, el separador decimal será un punto, no una coma.

Para cualquier x real, $\lfloor x \rfloor$ denotará la parte entera de x, esto es, el mayor entero menor o igual que x. Esta notación es muy usual en computación y temas cercanos a ella. En los libros de teoría de números es más común $[x]$, pero los corchetes se utilizan en muchos otros contextos, luego pueden dar lugar a confusión. Eso ha hecho que hayamos decidido usar $\lfloor x \rfloor$. Aún debemos comentar otra discrepancia con la notación habitual en otros textos. Está bastante extendido que el máximo común divisor de dos números a y b se denote mediante (a, b); nosotros hemos preferido $\mathrm{mcd}(a, b)$, que no crea ningún conflicto.

En cada tema se explicará la notación específica que allí se vaya a usar (y que a veces diferirá de la de otros temas). Por ejemplo, a partir del tema 7 se adopta el convenio de que p siempre denota un número primo; esto facilita las cosas, pues no hay necesidad de recordar continuamente algunas hipótesis. Sin embargo, en los temas 4 y 5, p puede ser cualquier entero; se ha hecho así ya que lo más común es emplear p/q para representar un número racional. A falta de cualquier aclaración, la notación que se siga será siempre estándar.

Finalmente, si algún lector se siente atemorizado por la notación «O grande» y «o pequeña», le recomendamos que lea las detalladas explicaciones que podrá encontrar en la introducción del tema 8.

Agradecimientos

En primer lugar, quiero mostrar mi profundo agradecimiento a Mario Pérez y a Francisco José Ruiz («Pacho»), profesores del Departamento de Matemáticas de la Universidad de Zaragoza. Con toda generosidad, me facilitaron sus magníficos apuntes de clase —muchos de ellos escritos a mano con la primorosa letra de Mario—, que yo pude utilizar el primer año que impartí clase de teoría de números, antes de embarcarme en el proyecto de este libro. Partes de lo que ellos me proporcionaron han sido el germen de algunos de los temas aquí incluidos.

Quiero continuar con Manolo Benito, cuyo entusiasmo por la teoría de números —siempre pensando en problemas realmente difíciles— es incomparable. Formalmente, fui el Director de su Tesis Doctoral, pero estoy convencido que yo he aprendido más matemáticas de él que él de mí. Me ha prestado su desinteresada ayuda en varias partes del presente libro y, además, efectuó una completa y detallada lectura que ayudó a eliminar numerosos gazapos e imprecisiones.

Hay muchos otros compañeros y amigos a quienes debo estar agradecido. Quiero mencionar expresamente, y por orden alfabético, a Julián Aguirre, Juan Arias de Reyna, Manuel Bello, Julio Bernués, Óscar Ciaurri, Javier Cilleruelo, Javier Duoandikoetxea, Antonio Durán, Luis Español, Emilio Fernández, Jesús Laliena, Francisco Luquin, Luis Navas, José María Pérez («Chema»), Adolfo Quirós, Luz Roncal, Javier Soria y Anitha Srinivasan. Todos ellos me han proporcionado su ayuda o me han dado útiles consejos.

Así mismo, debo dar las gracias a mis alumnos de teoría de números. Durante varios años, impartí mis clases con versiones previas de este libro, y lo que yo observaba sobre sus temas de interés y su proceso de aprendizaje —espero no haberme equivocado demasiado— me ha sido muy útil a la hora de enfocar y redactar el texto. Además, con su ayuda pude corregir muchas erratas. De forma particular quiero mencionar a Edgar Labarga, cuya atenta labor descubriendo erratas fue encomiable. Por supuesto, seguro que permanecen algunas, pero todas las que queden son ya culpa de este autor.[2]

En un aspecto más editorial, también quiero agradecer a Guillermo Curbera sus constantes ánimos —desde que él era Editor General de la RSME— para proseguir con la escritura de este texto, y por su confianza en mí para que el proyecto llegara

[2] Los «sólo» con tilde que aparezcan en el libro no son gazapos, sino muestra de la discrepancia del autor con las normas de la Real Academia Española (y otras Academias de la Lengua) publicadas en la *Ortografía de la lengua española* de 2010, y que, modificando la normativa hasta entonces vigente, indican que «solo» no debe llevar tilde tanto si es adjetivo como si es adverbio. Sin la posibilidad de discriminar según la tilde, una frase tan sencilla como «fue solo a comprar pan» (o «demostró solo el teorema», para ponernos en el contexto de este libro) es ambigua. ¿Qué necesidad había de modificar algo que funcionaba bien y dejarlo peor?

a buen fin; la casualidad ha querido que ahora él sea responsable de la colección «Textos Universitarios» de RSME-Electolibris en la que se publica este libro. Así mismo, estoy agradecido a Electolibris (en particular a Bernardo Cascales y José Manuel Mira, a quienes proporcioné algunas versiones preliminares) por su interés por publicar el libro dentro de su sello editorial.

Quiero concluir mostrando mi reconocimiento al revisor anónimo designado por la Real Sociedad Matemática Española. Su extremadamente cuidadosa lectura del texto y sus útiles y numerosas indicaciones han contribuido a introducir en el libro una gran cantidad de pequeñas mejoras. Sin duda, este texto es ahora mejor que antes de que pasara por sus manos.

Logroño (España), septiembre de 2013

Prefacio a la segunda edición

La primera edición de este libro se publicó en 2014 y, tras agotarse, Ediciones Electolibris llevó a cabo algunas reimpresiones. En el verano de 2018, Salvador Sánchez-Pedreño, de Electolibris, me sugirió que, en lugar de hacer una simple reimpresión, actualizáramos el texto para publicar una segunda edición que recogiera las novedades producidas en algunos de los puntos que se tratan en el libro y, de paso, ampliara un poco algunos temas. Agradezco a Salvador tanto sus esfuerzos como su confianza en el libro.

La principal diferencia de esta edición respecto a la primera ha sido la inclusión de una sección completa dedicada al cálculo de $\zeta(2k)$, lo cual ha originado unos cuantos retoques por todo el libro. El resto de los cambios introducidos, aunque numerosos, han sido más pequeños: actualizaciones de resultados, comentarios cortos, aclaraciones en explicaciones, inclusión de más retratos de matemáticos, algunos ejercicios nuevos... Por supuesto, también se han efectuado unas cuantas correcciones de erratas.

Creemos que los esfuerzos dedicados han merecido la pena, pero es el lector el único que puede juzgarlo.

Logroño (España), septiembre de 2018

Tema 1
Aritmética básica

1.1. Divisibilidad
1.2. Números primos
1.3. El sistema de numeración posicional. Bases
1.4. Combinatoria
1.5. Coeficientes binomiales y binomio de Newton
1.6. Algunas ecuaciones diofánticas
1.7. Ejercicios



Tales de Mileto
(≈ 624 a.C. – ≈ 547 a.C.)



Demócrito de Abdera
(≈ 460 a.C. – ≈ 370 a.C.)

El primer objeto de interés de la teoría de números es el estudio de los números naturales (que consideramos sin el 0)

$$\mathbb{N} = \{1, 2, 3, 4, 5, 6, 7, 8, \dots\},$$

y de los enteros

$$\mathbb{Z} = \{\dots, -4, -3, -2, -1, 0, 1, 2, 3, 4, \dots\}.$$

Un concepto básico es la divisibilidad y sus propiedades; en particular, el teorema fundamental de la aritmética. Casi toda esta parte es clásica y, esencialmente, se encuentra ya en los *Elementos* de Euclides [42], que fue escrito allá por el siglo III a.C. En sus *Elementos*, Euclides recogió, de manera sistemática y enciclopédica, toda la geometría y la aritmética elemental griega, ordenando los resultados previos de Tales, Pitágoras, Hipócrates, Demócrito, Eudoxo y Teeteto. Su estilo axiomático-deductivo de exposición y demostración rigurosa, que ordena los resultados en una secuencia jerárquica lógica, sigue siendo el mismo que los matemáticos usamos actualmente. Los *Elementos* han sido el núcleo central de la matemática elemental ¡durante dos milenios!, haciendo de Euclides un gran maestro de autoridad indiscutida; en resumen, un magistral libro de texto, el principal vehículo de transmisión del saber matemático a lo largo de la historia de la ciencia y de la educación.

Estudiaremos todo lo que concierne a la divisibilidad con cierta profundidad, aunque sin detenernos demasiado en algunos detalles. Damos por hecho que el lector ya tiene conocimiento intuitivo de los números, y que sabe desenvolverse con las herramientas del álgebra básica.

En matemáticas, siempre hay que partir de algo que se considera ya conocido, o establecido previamente. Aquí, damos por supuesto que el lector está familiarizado con las propiedades aritméticas de $\mathbb{N}$ y, por ende, de su extensión a $\mathbb{Z}$; nos referimos, centrándonos en el caso de $\mathbb{Z}$, a su comportamiento con la suma $+$ y el producto $\cdot$

(símbolo éste que se suele omitir en favor de la simple yuxtaposición de términos siempre que no dé lugar a confusión). Estas propiedades son la conmutatividad $a+b=b+a$ y $ab=ba$, asociatividad $(a+b)+c=a+(b+c)$ y $(ab)c=a(bc)$, propiedad distributiva $a(b+c)=ab+ac$, existencia de elemento neutro (0 para la suma y 1 para el producto, es decir, $0+a=a$ y $1\cdot a=a$), existencia de inverso respecto a la suma (el inverso aditivo de a se denota $-a$, y se le suele denominar opuesto) pero no respecto al producto (sólo lo tienen 1 y -1),... En el lenguaje del álgebra, todo esto hace que $\mathbb{Z}$ tenga estructura de anillo, pero aquí no insistiremos más en ello.

De igual modo suponemos conocida la relación de orden $\leq$ en $\mathbb{Z}$, así como la aplicación *valor absoluto* (también llamada *módulo*), que denotamos mediante $|\cdot|$. Destaquemos sus principales propiedades: las desigualdades se conservan al hacer sumas y al multiplicar por enteros positivos; se tiene, además, la *desigualdad triangular* $|a+b|\leq |a|+|b|$. También básico es el hecho de que $\mathbb{N}$ (es decir, los enteros positivos) es un conjunto *bien ordenado*: cualquier subconjunto no vacío de $\mathbb{N}$ tiene un primer elemento; a dicha propiedad se alude diciendo que $\mathbb{N}$ satisface el *principio de buena ordenación*, o el *principio del mínimo*. En $\mathbb{Z}$, lo que ocurre es algo muy similar: cualquier subconjunto de $\mathbb{Z}$ acotado inferiormente tiene un primer elemento.

Una característica fundamental de los enteros es que satisfacen el denominado *principio de inducción matemática* o *axioma de inducción*. Supongamos que estamos interesados en cierta propiedad que denominamos S, y que cada entero k puede satisfacer o no S. Pues bien, el principio de inducción dice que

- si existe un $k_0\in\mathbb{Z}$ que satisface S,
- si asumiendo que k satisface S se puede probar que $k+1$ también lo satisface,

entonces la propiedad S es cierta para todos los enteros $k\geq k_0$. (Lo más común es que sea $k_0=0$ o 1.) Formalmente, resulta fácil comprobar que el principio de inducción matemática es equivalente al principio de buena ordenación.

En ocasiones, el principio de inducción se enuncia de manera ligeramente distinta, que a veces resulta más útil al efectuar demostraciones. Por simplicidad para enunciarlo supongamos que empezamos con $k_0=1$. Esta variante —a la que a menudo se alude con el nombre de «inducción completa»— afirma que

- si 1 satisface S,
- si asumiendo que $1,2,\ldots,k$ satisfacen S se puede probar que $k+1$ también lo satisface,

entonces la propiedad S es cierta para todos los enteros. Ambos principios de inducción son equivalentes, y nosotros no distinguiremos entre ellos cuando queramos usarlos para demostrar algo.

En realidad, y en cuanto al desarrollo lógico interno de las matemáticas se refiere, gran parte de lo descrito en los párrafos anteriores entronca de manera muy profunda con la definición axiomática de $\mathbb{N}$ (mediante los axiomas de Peano) o su construcción a partir de la teoría de conjuntos. Una vez dado este primer paso, la construcción de $\mathbb{Z}$, y luego de los racionales $\mathbb{Q}$, ya es sencilla. No nos preocuparemos de nada de esto en el presente tema; a quien esté interesado en profundizar en ello, le remitimos al apéndice del presente libro.

Este tema comienza con el estudio de la divisibilidad y sus propiedades. Por supuesto, también abordamos con detalle el concepto de número primo, así como algunos de los resultados que de él se derivan; entre ellos, el denominado teorema fundamental de la aritmética, que convierte a los números primos en los ladrillos con los que se construye el edificio de la teoría de números. Aparece entonces una sección en la que comentamos con detalle nuestro sistema de numeración posicional. Y, tras ella, un estudio básico de combinatoria. Concluimos el tema con un pequeño paseo por las ecuaciones diofánticas.

1.1. Divisibilidad

Dados dos números enteros a y d, se dice que d *divide* a a, que d es un *divisor* de a, o que a es *divisible* por d, y lo escribimos $d\,|\,a$, si existe otro número entero q tal que $a = qd$. Si $d\,|\,a$ también decimos que a es un *múltiplo* de d. Usamos $d\mathbb{Z}$ para denotar el conjunto de los múltiplos de d, y $d\mathbb{N}$ para el de los múltiplos positivos, es decir,

$$d\mathbb{Z} = \{dm : m \in \mathbb{Z}\}, \qquad d\mathbb{N} = \{dm : m \in \mathbb{N}\};$$

así, $d\,|\,a$ es lo mismo que decir $a \in d\mathbb{Z}$ (o que $a \in d\mathbb{N}$, si estamos tratando sólo con números positivos). Si d no divide a a escribimos $d\nmid a$.

Hechos elementales como los siguientes se deducen inmediatamente de la definición:

Proposición 1.1.

Los números enteros satisfacen las siguientes propiedades:

- *Si* $d\,|\,a$ *entonces* $d\,|\,ac$ *para cualquier* c.
- *Si* $d\,|\,a$ *y* $a\,|\,b$, *entonces* $d\,|\,b$.
- *Si* $d\,|\,a$ *y* $d\,|\,b$, *entonces* $d\,|\,(a+b)$.
- *Todos los enteros dividen a* 0.
- 0 *no divide a ningún entero distinto de* 0.
- 1 *divide a todos los enteros.*
- *Cualquier entero es divisor de sí mismo.*

DEMOSTRACIÓN. Son todo consecuencias directas de la definición. Como muestra, comprobemos simplemente la primera propiedad: si $d\,|\,a$, existirá un $q \in \mathbb{Z}$ tal que $a = qd$; en consecuencia, $ac = (qc)d$ y $d\,|\,ac$. El resto lo dejamos para el lector, caso de que desee entretenerse comprobando las propiedades hasta su último detalle. □

El 0 sólo es divisor de sí mismo, y cualquier entero es divisor de 0; con eso, no hay nada más que decir sobre la posible divisibilidad entre dos enteros a y b cuando uno de ellos es el 0. En otro caso, dado que

$$b\,|\,a \;\Leftrightarrow\; -b\,|\,a \;\Leftrightarrow\; -b\,|\,-a \;\Leftrightarrow\; b\,|\,-a,$$

para muchas cuestiones podemos asumir que alguno de ellos es positivo; normalmente se supone que $b > 0$ (el posible divisor se considera positivo).

Un resultado básico pero fundamental —y conviene resaltarlo— es el siguiente:

Teorema 1.2. Algoritmo de la división.
Dados $a, b \in \mathbb{Z}$ con $b \geq 1$, existen unos enteros $q, r \in \mathbb{Z}$ tales que

$$a = bq + r, \qquad 0 \leq r < b. \tag{1.1}$$

Además, $b\,|\,a$ si y sólo si $r = 0$.

DEMOSTRACIÓN. Consideremos la «sucesión doble» $\{a - bq\}_{q=-\infty}^{\infty}$, que se extiende indefinidamente en sentido positivo y negativo. De ella, tomemos el menor número no negativo, y llamémoslo r. Así, $r = a - qb$ para un cierto q, y se cumple $r \geq 0$ y $r - b = a - (q+1)b < 0$. Con eso, ya hemos probado (1.1). La parte $b\,|\,a \;\Leftrightarrow\; r = 0$ es obvia. □

Al utilizar el algoritmo de la división y escribir $a = bq + r$ con $0 \leq r < b$, al número q se le llama *cociente* de la división de a por b, y el número r es el *resto* (acabamos de enunciar el archiconocido «dividendo es igual a divisor por cociente más resto» de nuestros años escolares; por descontado, el resto debía quedar menor que el divisor pues, de lo contrario, se podía tomar un cociente mayor). En el teorema hemos dado una demostración rigurosa de este hecho; en ella se ve que el algoritmo de la división es una fácil consecuencia del principio de buena ordenación. Por supuesto, la última parte del teorema se lee diciendo que a divide a b si y sólo si al efectuar la división de a por b el resto es 0. Finalmente, comentemos que cuando tenemos $a = bq + r$ como en (1.1), se denota $a \equiv r \pmod{b}$; esto da lugar a las *congruencias*, que en este libro serán tratadas aparte, en el tema siguiente.

Dado $a \in \mathbb{Z}$, es muy fácil señalar el conjunto de todos sus múltiplos: $\{ma : m \in \mathbb{Z}\}$. Pero señalar sus divisores (basta pensar en los positivos) es un problema mucho más complicado, del que surge la difícil teoría de números multiplicativa.

En este párrafo vamos a hacer las primeras consideraciones sobre este problema. Una herramienta imprescindible es el *máximo común divisor* de dos enteros a y b no ambos nulos que, como su nombre indica, es:

$$\mathrm{mcd}(a,b) = \mathrm{máx}\{d \in \mathbb{Z} : d\,|\,a,\ d\,|\,b\}.$$

La existencia del máximo común divisor es inmediata a partir del principio de buena ordenación; para ello, observe el lector que lo mismo da decir «cualquier conjunto acotado inferiormente tiene mínimo» que «cualquier conjunto acotado superiormente tiene máximo» (se pasa de una afirmación a otra multiplicando todos los elementos del conjunto por -1). Además, es claro que $\mathrm{mcd}(a,b)$ es un número entero entre 1 y $\mathrm{mín}\{|a|,|b|\}$. Nótese que, si $a = 0$, $\mathrm{mcd}(a,b) = |b|$ (y, análogamente, $\mathrm{mcd}(a,b) = |a|$ si $b = 0$) y esto no tiene demasiado interés, así generalmente supondremos que $a, b \in \mathbb{Z} \setminus \{0\}$; a menudo, incluso que $a, b \in \mathbb{N}$ (si cambiamos a o b por $-a$ o $-b$, el máximo común divisor no cambia). Por otra parte, a veces se dice, por completitud, que $\mathrm{mcd}(0,0) = +\infty$.

Pero no sólo nos contentamos con mostrar la existencia de máximo común divisor, sino que existe un método sencillo y rápido para hallarlo. Se conoce con el nombre de *algoritmo de Euclides* y se basa en el hecho evidente de que, si $a = bq + r$, cualquier número que divida a a y b también divide a b y r (y viceversa); en consecuencia,

$$\mathrm{mcd}(a,b) = \mathrm{mcd}(b,r). \tag{1.2}$$

Veamos cómo podemos usar eso para encontrar el máximo común divisor de dos números enteros r_1 y r_2, que podemos considerar positivos, y donde suponemos $r_1 \geq r_2$. Queremos encontrar $d = \mathrm{mcd}(r_1, r_2)$. El algoritmo de la división y la propiedad (1.2) nos aseguran que

$$r_1 = r_2 q_1 + r_3, \quad \text{con } 0 \leq r_3 < r_2, \quad d = \mathrm{mcd}(r_2, r_3). \tag{1.3}$$

Si $r_3 \neq 0$, repetimos el proceso, ahora con r_2 y r_3; esto es,

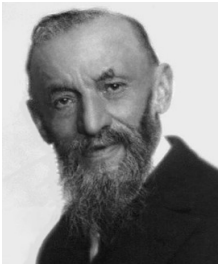
$$r_2 = r_3 q_2 + r_4, \quad \text{con } 0 \leq r_4 < r_3, \quad d = \mathrm{mcd}(r_3, r_4).$$

Y así sucesivamente. Los restos r_j son no negativos y cada vez menores. Luego llegará un momento en que el resto será nulo. Es decir, se tendrá

$$r_{n-2} = r_{n-1} q_{n-2} + r_n, \qquad \text{con } 0 \leq r_n < r_{n-1}, \quad d = \mathrm{mcd}(r_{n-1}, r_n), \tag{1.4}$$



Étienne Bézout
(1730–1783)



Giuseppe Peano
(1858–1932)

$$r_{n-1} = r_n q_{n-1} + r_{n+1}, \qquad \text{con } 0 \leq r_{n+1} < r_n, \quad d = \text{mcd}(r_n, r_{n+1}), \tag{1.5}$$

$$r_n = r_{n+1} q_n + 0, \qquad \text{con } r_{n+2} = 0 < r_{n+1}. \tag{1.6}$$

Como r_{n+1} divide a r_n de manera exacta, es obvio que $\text{mcd}(r_n, r_{n+1}) = r_{n+1}$. Por tanto, $d = r_{n+1}$, con lo que concluye el algoritmo de Euclides.

Nota 1.3. Seguro que el lector es plenamente consciente de que, en matemáticas, no podemos asegurar que algo sea cierto porque lo sea en muchos casos. El máximo común divisor nos proporciona un bonito ejemplo de esto: para n desde 1 hasta 533 359 (¡muchos casos!), el máximo común divisor entre $n^5 + 5$ y $(n+1)^5 + 5$ es $\text{mcd}(n^5 + 5, (n+1)^5 + 5) = 1$; pero, para $n = 533\,360$, el máximo común divisor es 1 968 751. ◁

1.1.1. La identidad de Bézout

Una propiedad muy útil de los enteros es la denominada *identidad de Bézout*. De ella vamos a ver dos demostraciones. La primera está basada en el principio de buena ordenación, y posiblemente pueda parecer más sencilla. La segunda está basada en el algoritmo de Euclides que acabamos de describir, y tiene una ventaja práctica: la propiedad que perseguimos afirma que existen dos números que satisfacen cierta identidad, y esta segunda demostración no sólo prueba que existen, sino que proporciona un procedimiento eficaz para encontrarlos.

Teorema 1.4. Identidad de Bézout.
Dados $a, b \in \mathbb{Z}$ no ambos nulos, existen $\lambda_1, \lambda_2 \in \mathbb{Z}$ tales que

$$\text{mcd}(a, b) = \lambda_1 a + \lambda_2 b.$$

PRIMERA DEMOSTRACIÓN. Tomemos $A = \{ax + by : x, y \in \mathbb{Z}\}$ y sea m el menor número positivo del conjunto A, es decir,

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

Así pues, $m = ax_0 + by_0$ para ciertos enteros x_0, y_0. Comprobemos que $m \,|\, a$ y $m \,|\, b$. Al dividir a entre m sabemos que existen q y $0 \leq r < m$ tales que $a = mq + r$, y en consecuencia $r = a - mq = a - q(ax_0 + by_0) = a(1 - qx_0) + b(-qy_0) \in A$. Pero m era el menor elemento positivo de A, luego forzosamente $r = 0$, y por tanto $m \,|\, a$. Un razonamiento análogo muestra que también $m \,|\, b$.

Hemos visto que m es divisor común de a y b. Comprobemos que es el máximo. Pero esto es fácil: si $d > 0$ es un divisor común de a y b, al ser $m = ax_0 + by_0$, también se cumplirá $d \,|\, m$, luego $d \leq m$. Ahora, basta tomar $\lambda_1 = x_0$ y $\lambda_2 = y_0$. □

SEGUNDA DEMOSTRACIÓN. Sin pérdida de generalidad (y simplemente para hacer la demostración más clara), podemos considerar que a y b son positivos. Sean $r_1 = \text{máx}\{a, b\}$, $r_2 = \text{mín}\{a, b\}$ y $d = \text{mcd}(a, b) = \text{mcd}(r_1, r_2)$. Una vez calculado d según el algoritmo mostrado en (1.3)–(1.6), vamos a recorrer ese proceso en sentido inverso, del final hacia el principio.

Al ser $d = r_{n+1}$, la expresión (1.5) la podemos escribir como

$$a_1 r_n + b_1 r_{n-1} = d \tag{1.7}$$

(con $a_1 = -q_{n-1}$ y $b_1 = 1$, pero no estamos interesados en sus valores concretos). Ahora, escribamos (1.4) en la forma $r_n + q_{n-2} r_{n-1} - r_{n-2} = 0$; multiplicándolo por $-a_1$ y sumándolo a (1.7) obtenemos

$$a_2 r_{n-1} + b_2 r_{n-2} = d \tag{1.8}$$

para ciertos a_2 y b_2 (concretamente, $a_2 = b_1 - a_1 q_{n-1}$ y $b_2 = a_1$, pero, de nuevo, eso no tiene importancia). Del mismo modo, la expresión anterior a (1.4) sería $r_{n-1} + q_{n-3} r_{n-2} - r_{n-3} = 0$; multiplicando por $-a_2$ y sumándolo a (1.8) quedaría

$$a_3 r_{n-2} + b_3 r_{n-3} = d.$$

Siguiendo con el retroceso, vamos obteniendo una serie de valores a_j y b_j tales que

$$a_j r_{n-j+1} + b_j r_{n-j} = d.$$

Al final, con $j = n - 1$ queda

$$a_{n-1} r_2 + b_{n-1} r_1 = d,$$

de donde se sigue el teorema. □

Un hecho básico —pero importante— que se deduce de este teorema es el siguiente:

Corolario 1.5.
Si a, b y d son números enteros tales que $d|ab$ y $\mathrm{mcd}(a,d)=1$, entonces $d|b$.

Demostración. Si $a=0$ o $b=0$, el resultado es trivial, o sea que podemos suponer $a\neq 0\neq b$. Por la identidad de Bézout, existen enteros λ_1 y λ_2 tales que $\lambda_1 a+\lambda_2 d=1$. Multiplicando ambos lados de esa ecuación por b tenemos $\lambda_1 ab+\lambda_2 bd=b$. Y, como d divide a los dos sumandos de esta igualdad, se sigue que $d|b$. □

La definición de máximo común divisor se extiende a una cantidad finita cualquiera de números enteros: dados $a_1,a_2,\ldots,a_n\in\mathbb{Z}$, no todos nulos, se define

$$\mathrm{mcd}(a_1,a_2,\ldots,a_n)=\mathrm{máx}\{d\in\mathbb{Z}: d|a_j \text{ para } 1\leq j\leq n\}.$$

Para calcular el máximo común divisor de más de dos enteros se puede usar el algoritmo de Euclides acompañado del siguiente resultado:

Lema 1.6.
Sean $a_1,a_2,\ldots,a_{n-2},a_{n-1},a_n$ números enteros, no todos nulos. Entonces

$$\mathrm{mcd}(a_1,a_2,\ldots,a_{n-2},a_{n-1},a_n)=\mathrm{mcd}(a_1,a_2,\ldots,a_{n-2},\mathrm{mcd}(a_{n-1},a_n)).$$

Demostración. Cualquier divisor común de los n enteros $a_1,a_2,\ldots,a_{n-2},a_{n-1},a_n$ divide, en particular, a a_{n-1} y a_n, luego también a $\mathrm{mcd}(a_{n-1},a_n)$. Por otra parte, cualquier divisor común de los $n-1$ enteros $a_1,a_2,\ldots,a_{n-2}$ y $\mathrm{mcd}(a_{n-1},a_n)$ dividirá a los n enteros, pues al dividir $\mathrm{mcd}(a_{n-1},a_n)$ debe dividir a a_{n-1} y a_n (obsérvese aquí que puede ocurrir $a_{n-1}=a_n=0$ y entonces $\mathrm{mcd}(0,0)=+\infty$, pero esto sólo da lugar a casos triviales que no causan problemas). Así pues, como los dos conjuntos $\{a_1,a_2,\ldots,a_{n-2},a_{n-1},a_n\}$ y $\{a_1,a_2,\ldots,a_{n-2},\mathrm{mcd}(a_{n-1},a_n)\}$ comparten todos sus divisores, el mayor de ellos será el mismo. □

La identidad de Bézout también se extiende a un número finito de enteros: existen $\lambda_1,\lambda_2,\ldots,\lambda_n\in\mathbb{Z}$ tales que

$$\mathrm{mcd}(a_1,a_2,\ldots,a_n)=\lambda_1 a_1+\lambda_2 a_2+\cdots+\lambda_n a_n. \tag{1.9}$$

La demostración es inmediata por inducción sin más que aplicar el lema 1.6 de manera sucesiva.

1.2. Números primos

Decimos que un entero entero $p > 1$ es *primo* si sus únicos divisores positivos son 1 y p. Una definición equivalente es que un número entero $p \geq 1$ es primo si tiene exactamente dos divisores positivos. Los demás enteros mayores que 1 se denominan *compuestos*. Debe quedar claro para las diferentes fórmulas que aparecerán en este libro que 1 no es un número primo.

Decimos que dos números $a, b \in \mathbb{Z} \setminus \{0\}$ son *primos entre sí*, *relativamente primos*, o *coprimos*, si $\mathrm{mcd}(a, b) = 1$. Un ejemplo importante de esta situación, y que resulta inmediato sin más que usar la definición de número primo, es la siguiente: si $a \in \mathbb{Z} \setminus \{0\}$ y p es primo y $p \nmid a$, entonces $\mathrm{mcd}(a, p) = 1$.

Nos topamos ahora con otro de los numerosos resultados prácticamente inmediatos que nos están apareciendo. No obstante su sencillez, lo exponemos con cuidado para poder llevar a cabo un desarrollo riguroso del tema.

Lema 1.7.
Cada entero mayor que 1 *tiene un divisor primo.*

Demostración. Supongamos que no es cierto. En tal caso, sea n el menor entero mayor 1 que no tiene ningún divisor primo (tal n debe existir pues $\mathbb{Z}$ es un conjunto bien ordenado). Como $n \mid n$ y n no tiene divisores primos, n no puede ser primo. Por tanto, n podrá escribirse $n = ab$ con $1 < a < n$ y $1 < b < n$. Como $a < n$ y n es el menor entero mayor que 1 sin divisores primos, a deberá tener un divisor primo p. Pero, entonces, $p \mid a$ y $a \mid n$, luego $p \mid n$, y eso contradice que n no tenía divisores primos. □

Como primera consecuencia, este lema proporciona un método sencillo (pero laborioso) de calcular todos los primos menores que un n dado. Se trata de la *criba de Eratóstenes*, que discurre como sigue: tras escribir, en una lista ordenada, todos los enteros entre 2 y n, se van tachando sucesivamente los múltiplos de 2, 3, 5, 7, etcétera, salvo ellos mismos (para tal fin, basta usar como base para las eliminaciones los no tachados previamente), y los que quedan sin eliminar son los primos. Este proceso se acelera de manera considerable si nos damos cuenta de que, cuando estamos eliminando los múltiplos de un cierto p, podemos comenzar no desde el principio, sino desde p^2; la razón es que si un número $n < p^2$ es múltiplo de p, también será múltiplo de un $q < p$, luego n ya habrá sido eliminado antes.

2	3	5	7	11	13	17	19	23	29
31	37	41	43	47	53	59	61	67	71
73	79	83	89	97	101	103	107	109	113
127	131	137	139	149	151	157	163	167	173
179	181	191	193	197	199	211	223	227	229
233	239	241	251	257	263	269	271	277	281
283	293	307	311	313	317	331	337	347	349
353	359	367	373	379	383	389	397	401	409
419	421	431	433	439	443	449	457	461	463
467	479	487	491	499	503	509	521	523	541
547	557	563	569	571	577	587	593	599	601
607	613	617	619	631	641	643	647	653	659
661	673	677	683	691	701	709	719	727	733
739	743	751	757	761	769	773	787	797	809
811	821	823	827	829	839	853	857	859	863
877	881	883	887	907	911	919	929	937	941
947	953	967	971	977	983	991	997	1009	1013
1019	1021	1031	1033	1039	1049	1051	1061	1063	1069
1087	1091	1093	1097	1103	1109	1117	1123	1129	1151
1153	1163	1171	1181	1187	1193	1201	1213	1217	1223

Cuadro 1.1: Tabla con los primeros 200 números primos.

Con este procedimiento, es fácil construir tablas de números primos, como la que presentamos en el cuadro 1.1.

Pero, además, del lema se obtiene el siguiente resultado, que ya aparecía enunciado —aunque de una forma alternativa «hay más números primos que cualquier cantidad propuesta de números primos» que evitaba cualquier alusión al infinito— en los *Elementos* de Euclides:

Teorema 1.8. Euclides.
Existen infinitos números primos.

Demostración. Supongamos que sólo existe una cantidad finita de primos, y sean $p_1, p_2, \ldots, p_r$ todos ellos. Tomemos ahora

$$N = p_1 p_2 \cdots p_r + 1.$$



Eratóstenes de Cirene
(276 a.C. – 194 a.C.)



Euclides de Alejandría
(≈ 325 a.C. – ≈ 265 a.C.)

Este N no es ninguno de los p_j (es mayor que todos ellos), luego no puede ser primo. Entonces, el lema 1.7 asegura N debe tener un divisor primo p. Pero este p que divide a N no puede ser ninguno de los p_j pues, si $p_j \,|\, N$, de $N - p_1 p_2 \cdots p_r = 1$ se deduciría que $p_j \,|\, 1$, lo cual nunca es cierto. De este modo, hemos encontrado un primo que no está en nuestra lista, que suponíamos completa, así que hemos llegado a una contradicción. □

Nota 1.9. Además del razonamiento que hemos usado —y que es, esencialmente, el de Euclides— existen otros métodos de demostrar el teorema anterior, algunos de ellos basados en argumentos aparentemente muy lejanos del objetivo. Así, por ejemplo, en [2, capítulo 1] se muestran seis formas distintas de probar la existencia de infinitos números primos, en [82, capítulo 1] se recogen diez, y en [69, § 1.1] una docena. Comentamos a continuación un par de ellas.

Un sencillo argumento permite reducir la existencia de infinitos números primos a encontrar una sucesión de enteros positivos y primos entre sí dos a dos; dejamos al lector que complete los detalles (ejercicio 1.8). En particular, esto dio origen a la que posiblemente sea la demostración más antigua distinta a la de Euclides: en 1730, Christian Goldbach escribió a Euler comunicándole que los números de Fermat $2^{2^n}+1$ eran coprimos dos a dos (nosotros estudiaremos los números de Fermat en el tema 3, y esa propiedad la veremos en el ejercicio 3.20). Este mismo razonamiento fue redescubierto —ya en el siglo XX— por György Pólya, a quien se le atribuye a menudo.

Un argumento más sofisticado —pero totalmente distinto— es el de Euler, quien probó la existencia de infinitos primos como consecuencia inmediata de otro resultado suyo: que la serie de los inversos de los números primos es divergente. En este libro lo veremos en el tema 9 (teorema 9.2). ◁

Los números primos tienen una propiedad realmente importante: si un primo divide a un producto de dos enteros, forzosamente debe dividir a uno de ellos. Veámoslo:

Proposición 1.10.
Sean a y b dos números enteros y p un primo. Si $p|ab$, entonces $p|a$ o $p|b$.

Demostración. Si $a = 0$ o $b = 0$ el resultado es trivial. Si, por el contrario, $a, b \in \mathbb{Z} \setminus \{0\}$, es un caso particular del corolario 1.5. En efecto, si $p \nmid a$, entonces $\mathrm{mcd}(a, p) = 1$, y por tanto $p|b$. □

Es más, esta propiedad que acabamos de mostrar caracteriza a los números primos, como se puede ver en el ejercicio 1.7.

Nota 1.11. En teoría de anillos, el término *primo* tiene un significado específico. Allí, un elemento no nulo y que tampoco es una unidad se dice *elemento primo* si, siempre que divide a un producto ab, debe dividir forzosamente a a o a b. Es decir, la definición de primo en teoría de anillos es justo la propiedad que nosotros hemos probado en la proposición 1.10. Además de que, con este significado, el opuesto aditivo de un elemento primo es también primo (en otras palabras, cuando consideramos el anillo $\mathbb{Z}$, los números $-2, -3, -5, -7, -11, \ldots$ también serían primos).

Por otra parte, lo que nosotros hemos llamado número primo es lo que en teoría de anillos se denomina *elemento irreducible*. En concreto, se dice que un elemento (no nulo y que tampoco es una unidad) de un anillo es irreducible cuando sus únicos divisores son triviales, es decir, cuando sólo es divisible por unidades del anillo y por él mismo multiplicado por unidades. (Quizás convenga aquí precisar que, en un anillo, se dice que un elemento u es una unidad si tiene inverso respecto al producto, esto es, si existe u^{-1} tal que $u^{-1}u = 1$; en $\mathbb{Z}$ las unidades sólo son 1 y -1.)

En un anillo genérico, puede haber elementos primos que no son irreducibles y elementos irreducibles que no son primos. Si el anillo es un *dominio de integridad* (es decir, si no existen divisores de cero, entendiendo como tal elementos a y b no nulos tales que $ab = 0$), los elementos primos siempre son irreducibles; en el ejercicio 1.7 se propone al lector que lo demuestre. Incluso en dominios de integridad, lo contrario puede no ser cierto. En $\mathbb{Z}$, sin embargo, ambos conceptos son equivalentes —proposición 1.10 y ejercicio 1.7—. Así pues, nosotros no nos preocuparemos para nada de la distinción entre primos e irreducibles, sino que aludiremos siempre a primos. Y, por supuesto, cuando hablemos de primos nos seguiremos refiriendo exclusivamente a números positivos. ◁

1.2.1. Factorización de un entero en producto de primos

Prácticamente, la propiedad mostrada en la proposición 1.10 es lo único que se necesita para probar por inducción el teorema siguiente:

Teorema 1.12. Teorema fundamental de la aritmética.
Todo entero mayor o igual que 2 *se puede escribir de forma única, salvo el orden, como producto de números primos.*

Antes de dar la demostración, detengámonos un poco. Es claro que los primos que aparezcan en el «producto de números primos» al que nos referimos en el teorema pueden repetirse (por ejemplo, $49 = 7 \cdot 7$). Dado un número $n \geq 2$, un primo p aparecerá en su factorización si $p \mid n$; y en este caso, estará exactamente $\nu(p)$ veces si $p^{\nu(p)} \mid n$ y $p^{\nu(p)+1} \nmid n$. Entonces, si $p_1, p_2, \ldots, p_k$ son los primos que dividen a n, lo que afirma el teorema fundamental de la aritmética es que

$$n = p_1^{\nu(p_1)} \cdots p_k^{\nu(p_k)}$$

(*factorización canónica*) y que si tenemos otro $m \geq 2$ y $m = q_1^{\nu(q_1)} \cdots q_r^{\nu(q_r)}$, entonces $n = m$ si y sólo si $\{p_1, \ldots, p_k\} = \{q_1, \ldots, q_r\}$ y los correspondientes exponentes son iguales.

Asimismo, deseamos adelantar algunas consecuencias que luego desarrollaremos de manera más completa. Conocida la factorización de un número $n \geq 2$, es muy fácil saber cuáles son sus divisores positivos. Son todos los productos que se pueden formar con los primos p que aparecen en su factorización y sus potencias hasta orden $\nu(p)$. Con un ejemplo lo expresamos mejor: como $60 = 2^2 \cdot 3 \cdot 5$, los divisores de 60 son, exactamente,

$$\{1, 2, 3, 5, 2 \cdot 3, 2 \cdot 5, 3 \cdot 5, 2^2, 2^2 \cdot 3, 2^2 \cdot 5, 2 \cdot 3 \cdot 5, 2^2 \cdot 3 \cdot 5\},$$

es decir,

$$\{1, 2, 3, 5, 6, 10, 15, 4, 12, 20, 30, 60\}.$$

Por otro lado, conocidas las factorizaciones de n y m es fácil ver que $\mathrm{mcd}(n, m)$ es el número cuya factorización consta de los primos comunes a n y m, elevados a la mayor potencia común. Por ejemplo, $252 = 2^2 \cdot 3^2 \cdot 7$, así que $\mathrm{mcd}(60, 252) = 2^2 \cdot 3 = 12$. Sin embargo, este método de calcular el máximo común divisor no quita validez al algoritmo de Euclides, sino muy al contrario. Con el método ahora

descrito se necesita conocer la descomposición en factores de n y m; esto puede ser muy laborioso si los números son grandes. El algoritmo de Euclides no tiene tal problema.

Demostración del teorema fundamental de la aritmética. Veamos primero que cada entero $n \geq 2$ se puede escribir como producto de primos. Hagámoslo por inducción. Para $n = 2$ es cierto, pues 2 es un número primo. Ahora, supongámoslo cierto para todos los enteros menores (estrictamente) que n. Si n es primo, no hay nada que probar. Si n no es primo, tendremos $n = dd'$ con $1 < d \leq d' < n$. Por hipótesis de inducción, d y d' son ambos productos de primos, luego $n = dd'$ también lo es.

Usemos de nuevo inducción para probar que esta representación es única. Para $n = 2$ es evidente, así que supongámoslo cierto para enteros menores que n. Debemos probar que si tenemos dos descomposiciones en factores primos $n = p_1 \cdots p_k = q_1 \cdots q_r$ (no hemos escrito exponentes pues, en cada uno de los productos, los primos pueden estar repetidos), entonces $k = r$ y existe una permutación σ en $\{1, \ldots, k\}$ tal que $p_j = q_{\sigma(j)}$. Como p_k divide a $q_1 \cdots q_r$, la proposición 1.10 asegura que p_k divide a alguno de los factores q_t, esto es, existe un $t_0 \in \{1, \ldots, r\}$ tal que $p_k \mid q_{t_0}$. Pero q_{t_0} es primo, luego $p_k = q_{t_0}$. Sea entonces

$$\frac{n}{p_k} = p_1 \cdots p_{k-1} = \prod_{\substack{t=1 \\ t \neq t_0}}^{r} q_t < n.$$

Por hipótesis de inducción, $k - 1 = r - 1$ (luego $k = r$), y existe una biyección $\sigma\colon \{1, \ldots, k-1\} \to \{1, \ldots, r\} \setminus \{t_0\}$ tal que $p_j = q_{\sigma(j)}$ para $j = 1, \ldots, k-1$. Extendamos σ definiendo $\sigma(k) = t_0$. Así, ya tenemos la permutación σ en $\{1, \ldots, k\}$ que buscábamos. □

1.2.2. Primeras consecuencias de la descomposición en factores primos

La descomposición en factores primos proporciona una herramienta teórica muy cómoda para analizar la divisibilidad. En todo el libro se usará implícitamente muchas veces. Vamos ahora a dar unos pocos resultados enunciados explícitamente.

Comencemos diciendo que, si conocemos la descomposición en factores primos de dos enteros a y b, es inmediato conocer el de su producto ab. Así mismo, podemos saber si $a \mid b$ o no; y, en caso afirmativo, dar la expresión de b/a. Lo exponemos en el siguiente teorema, cuya demostración es inmediata con lo que ya sabemos:

Teorema 1.13.
Sean a y b enteros positivos con

$$a = p_1^{r_1} p_2^{r_2} \cdots p_k^{r_k} \qquad y \qquad b = p_1^{s_1} p_2^{s_2} \cdots p_k^{s_k},$$

donde los p_j son números primos todos ellos distintos y los exponentes cumplen $r_j \geq 0$ y $s_j \geq 0$ (estamos permitiendo que los exponentes sean nulos para lograr escribir n y m con los mismos primos; obviamente, los posibles factores de la forma $p_j^0 = 1$ no contribuyen al producto, y es lo mismo que si no estuvieran; y, si todos los exponentes son nulos, el número en cuestión es 1). Con esta notación, se cumple:

- $ab = p_1^{s_1+r_1} p_2^{s_2+r_2} \cdots p_k^{s_k+r_k}$.
- $b \,|\, a$ *si y sólo si* $r_j \geq s_j$, $j = 1, 2, \ldots k$; *y, en esas circunstancias,* $a/b = p_1^{r_1-s_1} p_2^{r_2-s_2} \cdots p_k^{r_k-s_k}$.

También puede usarse la descomposición en factores primos para calcular el máximo común divisor de dos números. Pero no nos vamos a contentar sólo con eso. Aunque indudablemente de menor importancia en el desarrollo de la teoría de números, parejo al máximo común divisor se define el *mínimo común múltiplo* de dos enteros a y b no nulos como

$$\operatorname{mcm}(a, b) = \min\{m \in \mathbb{N} : a \,|\, m,\ b \,|\, m\};$$

si a o b son nulos se puede decir que, entonces, $\operatorname{mcm}(a, b) = 0$. Como ocurría con el *máximo común divisor*, el principio de buena ordenación asegura la existencia de mínimo común múltiplo. Además, es claro que $\operatorname{mcm}(a, b)$ está entre 0 y $|a| \cdot |b|$. Para calcular el mínimo común múltiplo no hay un método similar al algoritmo de Euclides, pero no hay que preocuparse por ello; en particular porque —recordémoslo— describir los múltiplos es más sencillo que encontrar los divisores.

De manera natural, la definición de mínimo común múltiplo se extiende a un número finito cualquiera de números enteros: dados $a_1, a_2, \ldots, a_n \in \mathbb{Z}$, no todos nulos, se toma

$$\operatorname{mcm}(a_1, a_2, \ldots, a_n) = \min\{m \in \mathbb{N} : a_j \,|\, m \ \text{ para } 1 \leq j \leq n\}$$

(y $\operatorname{mcm}(a_1, a_2, \ldots, a_n) = 0$ si algún $a_j = 0$). El mínimo común múltiplo de más de dos números se puede ir calculado de manera sucesiva sin más que usar que, si $a_1, a_2, \ldots, a_{n-2}, a_{n-1}, a_n$ son números enteros no todos nulos, se cumple

$$\operatorname{mcm}(a_1, a_2, \ldots, a_{n-2}, a_{n-1}, a_n) = \operatorname{mcm}(a_1, a_2, \ldots, a_{n-2}, \operatorname{mcm}(a_{n-1}, a_n));$$

la demostración de este hecho es similar a la del lema 1.6.

Ya hemos dicho que la analogía entre los conceptos de máximo común divisor y mínimo común múltiplo se rompía por la existencia del algoritmo de Euclides, aplicable únicamente al cálculo del primero. Sin embargo, el teorema fundamental de la aritmética devuelve este paralelismo. En efecto, dados a y b dos enteros positivos, escribámoslos como

$$a = p_1^{r_1} p_2^{r_2} \cdots p_k^{r_k} \qquad \text{y} \qquad b = p_1^{s_1} p_2^{s_2} \cdots p_k^{s_k},$$

donde los p_j son números primos todos ellos distintos y los exponentes cumplen $r_j \geq 0$ y $s_j \geq 0$ (se aplican aquí las mismas consideraciones que en el teorema). Entonces

$$\begin{aligned} \operatorname{mcd}(a,b) &= p_1^{\min\{r_1,s_1\}} p_2^{\min\{r_2,s_2\}} \cdots p_k^{\min\{r_j,s_j\}}, \\ \operatorname{mcm}(a,b) &= p_1^{\max\{r_1,s_1\}} p_2^{\max\{r_2,s_2\}} \cdots p_k^{\max\{r_j,s_j\}}, \end{aligned}$$

cuya demostración es inmediata a partir del teorema fundamental de la aritmética.

En otras palabras:

- El máximo común divisor de dos números se calcula tomando sus factores primos comunes con el menor exponente (nótese aquí que los factores primos p_j no comunes son los que tienen $r_j = 0$ o $s_j = 0$, que proporcionan un $\min\{r_j, s_j\} = 0$).
- El mínimo común múltiplo de dos números se calcula tomando factores primos comunes y no comunes con el mayor exponente.

De manera general, para más de dos números se tiene:

Teorema 1.14.
Para cada entero $a \geq 1$ y cada primo p, sea $\nu(a,p)$ la multiplicidad de p como factor de a. Así, la descomposición canónica en factores primos de a se puede expresar como $a = \prod_{p \text{ primo}} p^{\nu(a,p)}$ (si un primo p no aparece en la descomposición en factores de a, $\nu(a,p) = 0$ y $p^{\nu(a,p)} = 1$; ahora, podemos ignorar todos los factores «1» y tenemos un producto finito). Entonces,

$$\operatorname{mcd}(a_1, a_2, \ldots, a_k) = \prod_{p \text{ primo}} p^{\min\{\nu(a_1,p),\nu(a_2,p),\ldots,\nu(a_k,p)\}}$$

y

$$\operatorname{mcm}(a_1, a_2, \ldots, a_k) = \prod_{p \text{ primo}} p^{\max\{\nu(a_1,p),\nu(a_2,p),\ldots,\nu(a_k,p)\}}.$$

Tampoco es necesario que insistamos en la demostración de este teorema. Si fuera imprescindible hacerlo, bastaría aplicar un proceso de inducción sobre el número de factores primos.

Finalicemos la sección con unas advertencias. Y es que nunca se peca de exceso de insistencia por recalcar que lo que hemos visto en este apartado es una herramienta teórica, no práctica. Esto es así, por ejemplo, en cuanto a usar el teorema 1.13 para determinar si un entero b divide a otro entero a o no. Salvo cuando las descomposiciones en factores primos ya las conocemos o son muy fáciles de calcular, es mucho más rápido ver si la división a/b es exacta que hallar las descomposiciones en factores primos.

Lo mismo hay que advertir en cuanto al cálculo del mcd y el mcm se refiere. De nuevo con la honrosa excepción de números muy pequeños o cuya descomposición en factores primos es muy fácil de encontrar, el algoritmo de Euclides proporciona un método de calcular el mcd mucho más eficiente que el basado en la descomposición en factores primos. Si lo deseáramos, para calcular el mcd podríamos usar la fórmula mostrada en el ejercicio 1.21.

1.3. El sistema de numeración posicional. Bases

A lo largo de la historia —e incluso la prehistoria—, la humanidad ha utilizado diversas formas de representar y escribir los números. No es el objetivo de este libro incluir una historia de los números, ni contar cómo han sido representados en distintas civilizaciones, pero sí que vamos a incluir una pequeña descripción de dónde proviene y cómo nos ha llegado el sistema de numeración posicional en base 10 que en la actualidad empleamos casi sin darnos cuenta.[1] Si el lector quiere profundizar en el tema le sugerimos que acuda a textos especializados como [56] o [39] (en particular, este último ha sido la fuente principal para escribir las notas que siguen).

Aunque no sabemos datar ni atribuir con detalle los avances, sí que podemos afirmar que nuestro actual sistema de numeración es de origen hindú y cuajó durante el periodo de esplendor cultural e intelectual que tuvo lugar en todo el valle del Ganges desde mediados del siglo III hasta mediados del siglo VI.

El sistema de numeración hindú se basa en el principio posicional y el uso del cero. El sistema tiene base 10, por lo que además del 0 se usan otras 9 cifras más.

[1] Además de este sistema, también permanecen entre nosotros los denominados *números romanos* —se utilizan, por ejemplo, para numerar siglos— y la *numeración sexagesimal* de procedencia babilónica —que empleamos al descomponer un minuto en 60 segundos—. Pero ambos son de uso muy restringido, y nunca se utilizan para representar números genéricos ni para operar.



Muhammad ibn Musa
al-Khwarizmi
(≈ 780 – ≈ 850)



Vigila (siglo X)



Gerberto de Aurillac
(≈ 945–1003)

El valor de cada una de estas cifras depende del lugar que ocupa en el número. Así, cuando escribimos 444, la cifra 4 toma tres valores distintos: el primer 4 vale 400, el segundo 4 vale 40 y el tercero vale 4. Sin duda alguna, la cifra más escurridiza —y conceptualmente la más importante— es el 0. Aparentemente parece un sinsentido tener un símbolo para representar la nada, pero esa idea es precisamente la que hace funcionar al sistema de numeración posicional. Si no dispusiéramos del cero: ¿cómo podríamos representar un número compuesto por exactamente 2 centenas y 5 unidades? El cero nos permite indicar que no hay decenas en ese número: 205.

Incluso, poco a poco, los matemáticos hindúes fueron dominando los números en general y, en particular, el cero, que explicaron como el resultado de restar un número de sí mismo. El cero dejó de ser sólo el símbolo para representar el «espacio vacío» en el sistema de numeración posicional, y se convirtió en un número por derecho propio; existe constancia de ello en un texto de Brahmagupta (598–670). Los matemáticos hindúes establecieron que un número multiplicado por cero da como resultado cero, mientras que un número menos cero es igual al mismo número y, por último, que un número dividido por cero es igual a infinito. También usaron los números negativos, que los griegos habían evitado y que luego llevaría muchos siglos introducir y hacerlos de uso común en Europa.

El sistema hindú de numeración permite escribir cualquier número, por grande que sea, usando solamente diez símbolos. Puesto que hay infinitos números, poder escribir cualquiera de ellos de manera sencilla y usando sólo diez símbolos es un logro intelectual realmente prodigioso. Para escribir estas diez cifras a partir de las cuales se puede representar cualquier número, los hindúes tuvieron la buena idea de no usar letras de su alfabeto. Usaron símbolos que derivaban de la escritura con otros alfabetos. Con el tiempo, tales símbolos fueron evolucionando, y ya en el siglo VII se parecían bastante a los nuestros actuales.

Tras su conversión al Islam en el siglo VI, los árabes iniciaron una expansión hacia el Oriente que los llevó a territorio hindú. A través de esa frontera se filtró el sistema hindú de numeración, de forma que ya a finales del siglo VIII los musulmanes lo habían asimilado. Del primer cuarto del siglo IX data una de las mejores aritméticas que compusieron los árabes. Su autor fue Al-Jwarismi o Al-Khwarizmi (ca. 780–850), el gran matemático y astrónomo árabe, uno de los padres del álgebra.[2]

La forma de representar los números que los árabes aprendieron en el Oriente circuló hasta la otra punta de sus dominios; y así, a través del norte de África, el sistema de numeración hindú llegó a España en el siglo IX. Además de los musulmanes, los mercaderes y comerciantes judíos pudieron también tener un papel importante en esta trashumancia que llevó los números del Oriente al Occidente.

De finales del siglo X es el denominado *Códice Vigilanus* (o *Códice Albeldense*) en el que el monje Vigila dejó escritos los números indo-arábigos —véase la figura 1.1—; también existe constancia de que Gerberto de Aurillac (el futuro papa Silvestre II) los llevó al otro lado de los Pirineos. Desde entonces hasta que éstos fueron de uso común en Europa bien entrado el Renacimiento, pasaron más de cinco siglos. Durante ese tiempo se produjo una batalla entre «algoristas», esto es, aquéllos que usaban para las cuentas el sistema hindú de numeración y los numerales árabes, y «abacistas», esto es, aquéllos que hacían las cuentas con el viejo sistema romano mejorado con artilugios como el ábaco, o las mesas de cálculo.

Lo que al inicio del Renacimiento vino a decantar la victoria del lado de los «algoristas» frente a los «abacistas» fue el desarrollo del comercio durante la baja Edad Media. El incremento de la actividad comercial durante los siglos XII y XIII hizo necesaria una contabilidad cada vez más complicada, y la simplificación de las operaciones aritméticas que permite el sistema hindú se tornó imprescindible. Surgió así, a partir del siglo XIII, una intensa relación entre el comercio y la aritmética hindú escrita con cifras árabes. Esta relación quedó reflejada en el Renacimiento con la impresión, nada más desarrollada la imprenta, de numerosas aritméticas

[2] Abu Ja'far Muhammad ibn Musa al-Khwarizmi trabajó en Bagdad en la Casa de la Sabiduría. El libro de aritmética de Al-Khwarizmi lo conocemos por una versión latina del siglo XIII, pues no se ha conservado en su versión árabe. En él se explica la forma de representar números usando el sistema y las cifras hindúes, y la forma de hacer las *cuatro operaciones*: suma, resta, multiplicación y división. Del nombre de Al-Khwarizmi derivan las palabras «algoritmo» y «guarismo». La primera se usó para describir el procedimiento de cálculo usando el sistema hindú de numeración, aunque hoy viene a significar un sistema organizado de instrucciones y operaciones que permiten resolver un determinado problema. La segunda, hoy en cierto desuso, es sinónimo de cifra. Al-Khwarizmi empleó la palabra árabe «sifr», que significa vacío, para referirse al cero. La palabra castellana «cifra», con que desde el Renacimiento nombramos a los dígitos, deriva directamente de «sifr» (y lo mismo ha ocurrido en otras lenguas). Esto sugiere la importancia conceptual del cero: su nombre árabe ha acabado siendo usado para referirse al resto de los números.

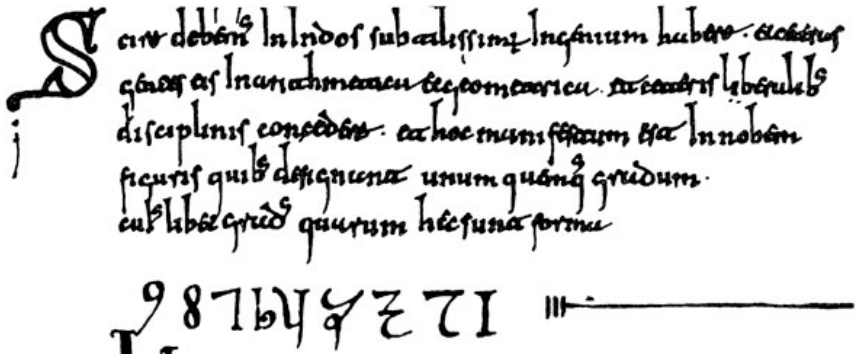


Figura 1.1: El *Códice Vigilanus* o *Códice Albeldense* fue compuesto por el monje Vigila entre los años 974 y 976 en el monasterio de San Martín de Albelda en La Rioja, y se guarda hoy en la Biblioteca del monasterio de El Escorial. El códice de Vigila es un manuscrito magnífico por su hermosura (contiene numerosas ilustraciones de gran colorido), por su antigüedad, y por su significado. Entre textos de contenido variado, el códice incluye algunos rudimentos de aritmética que, en buena parte, son extractos de diversos libros de las *Etimologías* de San Isidoro de Sevilla. Allí, Vigila inserta una glosa sobre las cifras indo-arábigas que es, precisamente, el registro más antiguo que se conserva donde las nueve cifras —no está el 0— aparecen escritas con los caracteres con que hoy las seguimos representando (el 1, 2, 6, 7, 8 y 9 son idénticos a los actuales, mientras que el 3, 4 y 5 han sufrido cierta variación). Es posible que el monje Vigila aprendiera los números que acabó incluyendo en su monumental manuscrito en el monasterio de Santa María de Ripoll, cerca de Vic. En particular, eso ilustra el esfuerzo de los monjes cristianos del noreste de la península ibérica por asimilar la cultura científica de Al-Andalus y su transmisión posterior al resto de Europa. Unos años antes de que Vigila visitara Ripoll, allí aprendió los números Gerberto de Aurillac, que después sería Papa con el nombre de Silvestre II. Gerberto fue quien primero impulsó la introducción en Europa de los numerales árabes, aunque como parte de un nuevo tipo de ábaco que usaba el principio posicional pero no el cero, innecesario en un ábaco.

mercantiles, entre las que vale la pena mencionar el *Liber Abaci* de Leonardo de Pisa (ca. 1170–1240), también llamado Fibonacci. Él fue la figura más destacada en la difusión de la aritmética hindú en Europa.[3]

[3] Fibonacci aprendió la aritmética hindú, y la ayuda inestimable que ésta prestaba al comercio, en sus diversos viajes por Egipto, Siria, Bizancio y Sicilia por encargo de su padre, un comerciante



Fibonacci (Leonardo de Pisa) (1170–1250)



Gottfried Wilhelm von Leibniz (1646–1716)

Abandonando ya la historia, es claro que el uso del 10 como base de nuestro sistema de numeración posicional es una mera «anécdota» que seguramente proviene de que en las manos tengamos diez dedos. En realidad, podríamos haber usado cualquier entero $m \geq 2$. Y, de hecho, los ordenadores actuales no usan internamente la base 10, sino que funcionan en base 2. La razón de ello es que en base 2 sólo existen dos símbolos, que en un dispositivo electrónico o informático se pueden representar con la ayuda de diversas propiedades físicas que admiten dos estados claramente diferenciados (a menudo, sólo esos dos): carga eléctrica positiva/negativa, campo magnético positivo/negativo, abierto/cerrado, encendido/apagado, blanco/negro... Aunque, en realidad, la aritmética binaria con sus dígitos 0 y 1 es muy anterior a la existencia de los ordenadores modernos, pues ya fue descrita por el matemático alemán Gottfried Leibniz en 1703.

El resultado teórico que sustenta el sistema de numeración posicional es el siguiente, cuya demostración, como veremos, sólo requiere usar inducción y el algoritmo de la división:

---

de Pisa. Sobre todo ello escribió un texto de título *Liber Abaci*, aparecido en 1202. El *Liber Abaci* estuvo muy influido por fuentes árabes y, a pesar de su título, enseña cómo hacer las cuentas con el sistema de numeración hindú y los numerales árabes. El libro contiene también ejemplos prácticos para ayuda de comerciantes y mercaderes sobre cálculo del precio de los bienes, estimación de beneficios, cambio de monedas, etc. En la sección tercera del *Liber Abaci*, Fibonacci propuso, para calcular el número de conejos engendrados por una pareja a lo largo de los meses, la sucesión de números 1, 1, 2, 3, 5, 8, 13, 21, 34, 55,... en la que cada término es la suma de los dos precedentes. La sucesión de Fibonacci es extremadamente útil, no sólo para contar conejos, sino como modelo para estudiar el crecimiento de multitud de seres vivos, ya sea una simple célula, un trigal o un enjambre de abejas. La sucesión de Fibonacci tiene por costumbre aparecer de manera sorprendente en diferentes áreas de las matemáticas y la ciencia. Por ejemplo, el cociente de un número de Fibonacci y su predecesor tiende a la razón áurea de los griegos. Estudiaremos con más detalle esta sucesión en el tema 5; en concreto, en la sección 5.1 (página 281 y siguientes).

Teorema 1.15.
Sea un entero $b \geq 2$. Cada entero positivo n se puede representar de manera única en la forma

$$n = d_k b^k + d_{k-1} b^{k-1} + \cdots + d_2 b^2 + d_1 b + d_0 \tag{1.10}$$

donde k es un entero no negativo tal que

$$b^k \leq n < b^{k+1}$$

y $d_0, d_1, \ldots, d_k$ son enteros que verifican

$$1 \leq d_k \leq b-1$$

y

$$0 \leq d_j \leq b-1 \qquad \text{para } j = 0, 1, 2, \ldots, k-1.$$

Antes de abordar la demostración, y aun a riesgo de ser demasiado insistentes sobre cosas que sin duda el lector ya conoce, quizás conviene precisar un poco cómo se relaciona esto con nuestra notación posicional habitual. En realidad, lo único que sucede es que el desarrollo (1.10) —que llamamos *representación de n en base b*— es común denotarlo mediante yuxtaposición de los coeficientes:

$$n = d_k d_{k-1} \ldots d_2 d_1 d_{0(b}.$$

Por supuesto, los coeficientes d_j son los denominados *dígitos* de n en base b. Usualmente, manejamos la base 10, y cuando ése es el caso ni siquiera escribimos el subíndice «(10» de la notación. Tras estos comentarios, vamos ya con la prueba del teorema.

Demostración. Utilizaremos inducción sobre k. Comencemos con $k = 0$. En ese caso, $1 = b^0 \leq n < b$ y la única representación de n en base b es $n = d_0$.

Efectuemos ahora el paso de inducción: suponiendo el resultado cierto hasta $k-1$ (nos referimos tanto a la existencia como a la unicidad), vamos a probarlo para k. Sea pues n que verifica $b^k \leq n < b^{k+1}$. Al aplicar el algoritmo de la división a n entre b^k, tendremos

$$n = d_k b^k + r, \qquad \text{con } \ 0 \leq r < b^k.$$

Comprobemos que d_k verifica $1 \leq d_k \leq b-1$. En efecto; por ser

$$0 < b^k - r \leq n - r = d_k b^k \leq n < b^{k+1},$$

dividiendo por b^k obtenemos $0 < d_k < b$; y, como d_k y b son enteros, forzosamente deberá cumplirse $1 \leq d_k \leq b-1$. Analicemos ahora r. Si $r = 0$, la representación en base b de n ya está encontrada: $n = d_k b^k$. Si, por el contrario, $r > 0$, al ser $r < b^k$ se cumplirá $b^{k'} \leq r < b^{k'+1}$ para un cierto $k' \leq k-1$. Por hipótesis de inducción, r tiene la representación

$$r = d_{k'} b^{k'} + d_{k'-1} b^{k'-1} + \cdots + d_2 b^2 + d_1 b + d_0.$$

Si $k' < k-1$, añadamos los valores $d_{k'+1} = d_{k'+2} = \cdots = d_{k-1} = 0$. Con esto, de $n = d_k b^k + r$ se sigue (1.10). Nos falta comprobar que esta representación en base b es única.

Supongamos que, además de (1.10), existe otra representación

$$n = e_l b^l + e_{l-1} b^{l-1} + \cdots + e_2 b^2 + e_1 b + e_0$$

con l un entero no negativo, $b^l \leq n < b^{l+1}$ y $e_0, e_1, \ldots, e_l$ enteros que verifican $1 \leq e_l \leq b-1$ y $0 \leq e_j \leq b-1$ para $j = 0, 1, 2, \ldots, l-1$. Si $l \geq k+1$, entonces

$$n < b^{k+1} \leq b^l \leq e_l b^l \leq n,$$

que es imposible. Si $l \leq k-1$, argumentamos de manera simétrica, intercambiando los papeles de l y k:

$$n < b^{l+1} \leq b^k \leq d_k b^k \leq n,$$

que de nuevo es imposible. En consecuencia, forzosamente $l = k$. Supongamos ahora que $d_k < e_k$. En ese caso,

$$\begin{aligned} n &= d_k b^k + d_{k-1} b^{k-1} + \cdots + d_2 b^2 + d_1 b + d_0 \\ &\leq d_k b^k + (b-1) b^{k-1} + \cdots + (b-1) b^2 + (b-1) b + (b-1) \\ &= d_k b^k + b^k - 1 < (d_k + 1) b^k \leq e_k b^k \leq n, \end{aligned}$$

que no es posible. Por simetría, tampoco puede ocurrir $e_k < d_k$, así que debe ser $d_k = e_k$. Tomemos ahora $n' = n - d_k b^k = n - e_k b^k$. Por una parte, este n' satisface $n' < b^k$ (luego verificará $b^{k'} \leq n' < b^{k'+1}$ para algún $k' < k$); por la otra, n' tiene dos representaciones

$$\begin{aligned} n' &= d_{k-1} b^{k-1} + \cdots + d_2 b^2 + d_1 b + d_0 \\ &= e_{k-1} b^{k-1} + \cdots + e_2 b^2 + e_1 b + e_0. \end{aligned}$$

Así, por hipótesis de inducción deberá ser $d_j = e_j$ para $j = 0, 1, \ldots, k-1$, y el teorema queda demostrado. □

Nota 1.16. Saliéndonos ya de los enteros, si extendemos (1.10) a

$$d_k b^k + d_{k-1} b^{k-1} + \cdots + d_2 b^2 + d_1 b + d_0 + \sum_{j=-1}^{-\infty} d_j b^j$$

(que siempre da lugar a una serie convergente) y lo denotamos

$$d_k d_{k-1} \ldots d_2 d_1 d_0 . d_{-1} d_{-2} d_{-3} \ldots{}_{(b},$$

obtenemos una manera de dar el desarrollo en base b de cualquier número real positivo. No es difícil comprobar que esos desarrollos son finitos o periódicos si y sólo si el número en cuestión es racional. Cuando el número es racional, será una cierta fracción irreducible h/q; en ese caso, el desarrollo será finito si y sólo si todos los divisores primos del denominador q son también divisores primos de la base b.

En particular, en base 10, una fracción irreducible h/q tiene un desarrollo decimal finito si y sólo si $q = 2^r 5^s$ para ciertos exponentes r y s no negativos. Si q tiene algún divisor primo distinto de 2 o 5, el desarrollo decimal es periódico. En base 2, sólo los números racionales cuyo denominador es de la forma 2^r tienen una representación finita. Es interesante caer en la cuenta de que ésta es la razón por la que los ordenadores —que trabajan en aritmética binaria— no pueden representar de manera exacta con un número finito de decimales muchos números que nosotros —en notación decimal— sí podemos (por ejemplo, nuestro $1/5 = 0.2$, que en base 2 es $0.001100110011\ldots{}_{(2}$); por supuesto, esto da lugar a los consabidos errores de truncamiento e imprecisiones de cómputo típicos del cálculo numérico. ◁

1.3.1. Criterios de divisibilidad

Cuando disponemos de un número natural N escrito en una determinada base b, a veces es fácil determinar si ese N es o no divisible por otro número natural $q > 1$ sin necesidad de hacer la división, sino simplemente observando el «aspecto» del desarrollo en base b de N, o una propiedad fácilmente comprobable. Esto es lo que se denomina un *criterio de divisibilidad* por q.[4]

Por ejemplo, un número expresado en base 10 cumple lo siguiente:

- es divisible por 10 si y sólo si su desarrollo decimal acaba en el dígito 0;
- es divisible por 5 si y sólo si su desarrollo decimal acaba en 0 o 5;
- es divisible por 2 si y sólo si su desarrollo decimal acaba en 0, 2, 4, 6 u 8.

[4] Aunque, para ser honestos, desde que se popularizaron las calculadoras —e incluso los ordenadores— tales criterios ya no son tan usados como en épocas anteriores.

La razón es que un número $N \in \mathbb{N}$ siempre puede escribirse como $N = D10 + d$, siendo $0 \leq d < 10$ el dígito final (el menos significativo) del desarrollo decimal de N. Así, como los tres valores $q = 10$, 5 y 2 dividen a la base 10, es claro que q divide a N si y sólo si q divide al dígito final d.

Es fácil adaptar eso a una base genérica b. En realidad, en base b, el número b se expresa precisamente como $10_{(b}$, así que un $N \in \mathbb{N}$ será $N = D10_{(b} + d = Db + d$ con $0 \leq d < b$; ese d es el «dígito» final en base b. Entonces, dado un q que divide a la base b, de nuevo se cumple $q \,|\, N$ si y sólo si $q \,|\, d$.

Cuando estamos analizando la divisibilidad por un número q que divide a 100, a 1000, o, en general, a 10^m (con $m \geq 1$), puede razonarse de manera similar. Escribamos $N \in \mathbb{N}$ como $N = D10^m + d$ con $0 \leq d < 10^m$; por ser q un divisor de 10^m, es obvio que $q \,|\, N$ si y sólo si $q \,|\, d$. Por citar un par de ejemplos, se cumple:

- N es divisible por 25 si y sólo si el número formado por las dos últimas cifras de su desarrollo decimal es divisible por 25 (pues $25 \,|\, 100$);
- N es divisible por 8 si y sólo si el número formado por las tres últimas cifras de su desarrollo decimal es divisible por 8 (pues $8 \,|\, 1000$).

Por supuesto, si tenemos un número N expresado en base b, y q es un divisor de $b^m = (10_{(b})^m$ (con $m \geq 1$), el criterio que determina la divisibilidad por q es el mismo: basta determinar si el número formado por las m últimas cifras (en base b) de N es o no divisible por q.

Volviendo a la base 10, también son muy conocidos y sencillos los criterios de divisibilidad por 3, 9 y 11 (si el lector no los recuerda, que no se preocupe: los vamos a ver en seguida). ¿Qué tienen de particular estos tres números?: pues que son justo los divisores del número anterior y el posterior a nuestra base 10. Es más, prácticamente con el mismo esfuerzo vamos a establecer criterios de divisibilidad entre divisores de $10^m - 1$ y de $10^m + 1$. Para aplicar estos criterios bastará tener el número $N > 0$ (cuya divisibilidad por cierto q queremos estudiar) escrito en la forma

$$N = \sum_{j=0}^{n} d_j 10^{m \cdot j}; \tag{1.11}$$

tal descomposición existe y es única si $0 \leq d_j < 10^m$ y $0 < d_n < 10^m$. De este modo, reduciremos la divisibilidad de N por q a la divisibilidad por q de ciertas combinaciones sencillas de los d_j. (Obsérvese además que (1.11) es como si escribiésemos N en base 10^m, donde los «dígitos» de la base 10^m son los 10^m símbolos $00\ldots00$, $00\ldots01$, $00\ldots02$, $99\ldots98$, $99\ldots99$, todos ellos de «longitud» m; con la salvedad de que en d_n, el «dígito» más significativo, no hace falta incluir los «0» que se usan para completar las m posiciones.)

Comencemos con las reglas de divisibilidad por divisores de $10^m - 1$:

Proposición 1.17.
Sea $m \geq 1$ y q un divisor de $10^m - 1$. Dado $N = \sum_{j=0}^{n} d_j 10^{mj}$, tomemos $N' = \sum_{j=0}^{n} d_j$. En estas circunstancias, $q \,|\, N$ si y sólo si $q \,|\, N'$.

Demostración. Al ser $N - N' = \sum_{j=0}^{n} (10^{mj} - 1) d_j$, el resultado estará probado si logramos establecer que

$$q \,|\, (10^{mj} - 1) \quad \text{para } j \in \mathbb{N}.$$

Pero esto es fácil por inducción: el caso $j = 1$ es verdad por hipótesis; y, si lo suponemos cierto para j, de

$$(10^{(m+1)j} - 1) - (10^{mj} - 1) = 10^{mj+m} - 10^{mj} = 10^{mj}(10^m - 1)$$

se desprende que también se cumple para $j + 1$. □

Ejemplo. Tomando $m = 1$ obtenemos las muy conocidas reglas de divisibilidad por 3 y 9:

- Un número es divisible por 3 si y sólo si la suma de sus dígitos es divisible por 3.
- Un número es divisible por 9 si y sólo si la suma de sus dígitos es divisible por 9. ◊

Ejemplo. Tomando $m = 3$, y como $1000 - 1 = 3^3 \cdot 37$, obtenemos un criterio de divisibilidad por 37. Apliquémoslo a un número concreto:

$$37 \,|\, 64\,975\,256 \;\Leftrightarrow\; 37 \,|\, (64 + 975 + 256) \;\Leftrightarrow\; 37 \,|\, 1295 \;\Leftrightarrow\; 37 \,|\, 296,$$

y esto último ya no queda más remedio que comprobar su veracidad efectuando la división. ◊

El caso correspondiente a los divisores de $10^m + 1$ requiere un poco más de trabajo. Comencemos con un lema, cuya demostración se deja para el lector (ejercicio 1.29):

Lema 1.18.
Sea $m \geq 1$ y q un divisor de $10^m + 1$. Además, sea $N = D10^m + d$, con $D, d \in \mathbb{Z}$. Con esta notación, $q \,|\, N$ si y sólo si $q \,|\, (D - d)$.

Aplicando el lema de manera reiterada es claro que llegamos a esto:

Proposición 1.19.

Sea $m \geq 1$ y q un divisor de $10^m + 1$. Dado $N = \sum_{j=0}^{n} d_j 10^{mj}$, tomemos $N' = \sum_{j=0}^{n} (-1)^j d_j$. En estas circunstancias, $q \,|\, N$ si y sólo si $q \,|\, N'$.

Ejemplo. Tomando $m = 1$ obtenemos el criterio de divisibilidad por 11: un número es divisible por 11 si y sólo si la diferencia entre la suma de las cifras que ocupan los lugares pares y la suma de las cifras de lugares impares es múltiplo de 11. Así, por ejemplo:

$$11 \,|\, 135\,802\,458 \;\Leftrightarrow\; 11 \,|\, \big((8+4+0+5+1) - (5+2+8+3)\big),$$

y eso es cierto pues $(8+4+0+5+1) - (5+2+8+3) = 0$. ◊

Ejemplo. Tomando $m = 3$ tenemos $1001 = 7 \cdot 11 \cdot 13$. De este modo surge la regla de divisibilidad por 7, que se aplica de manera similar a la que acabamos de mostrar para 11 pero tomando grupos de tres cifras:

$$7 \,|\, 86\,419\,746 \;\Leftrightarrow\; 7 \,|\, \big((86+746) - 419\big) \;\Leftrightarrow\; 7 \,|\, 413.$$

El mismo criterio sirve para analizar la divisibilidad por 13 (por supuesto, también se puede aplicar a la divisibilidad por 11, pero obviamente es mucho más sencillo el que hemos obtenido con $m = 1$). ◊

Si q_1 y q_2 son primos entre sí, la divisibilidad de un número por $q_1 q_2$ equivale a su divisibilidad por q_1 y q_2. Así, por ejemplo, $6 \,|\, N$ si y sólo si $2 \,|\, N$ y $3 \,|\, N$; y también $18 \,|\, N$ si y sólo si $2 \,|\, N$ y $9 \,|\, N$. De este modo, bastaría tener criterios de divisibilidad para primos o potencias de primos.

Entonces, de entre los números ≤ 25, para los únicos que aún no disponemos de un criterio de divisibilidad son 17, 19 y 23. Incluso para estos tres números es factible encontrar un criterio con lo que hemos visto hasta ahora. Bastaría tener en cuenta las descomposiciones

$$\begin{aligned} 10^8 + 1 &= 17 \cdot 5882353, \\ 10^9 + 1 &= 7 \cdot 11 \cdot 13 \cdot 19 \cdot 52579, \\ 10^{11} + 1 &= 11^2 \cdot 23 \cdot 4093 \cdot 8779 \end{aligned}$$

y aplicar la proposición 1.19 con $m = 8$ (para 17), $m = 9$ (para 19) y $m = 11$ (para 23). Pero, obviamente —y siguiendo con la notación de la proposición—, esto sólo puede tener alguna utilidad para números N mayores que 10^8, 10^9 y 10^{11} (respectivamente), y además el correspondiente N' acaba teniendo ese tamaño. Así pues, no nos libramos de una división relativamente «grande» en la comprobación final de si N' es o no divisible por 17, 19 o 23.

La idea para intentar manejar números menores es buscar criterios que sirvan para divisores de números cercanos a 10^m, no necesariamente $10^m \pm 1$, sino $10^m + k$ con $k \neq 0$ (y pequeño). De este modo se pueden obtener adaptaciones del lema 1.18 que llevan a los correspondientes criterios de divisibilidad (véase el ejercicio 1.30). De manera «simétrica», también se podría analizar el caso de los divisores de $K \cdot 10^m + 1$ (con $K \geq 1$ pero pequeño). E incluso combinar ambas ideas (ejercicio 1.31).

De este modo, un criterio de divisibilidad muy general es el siguiente:

Teorema 1.20.
Sea q un divisor de un número de la forma $K \cdot 10^m + k$, con $K \geq 0$, $m \geq 0$, $k \neq 0$ y $\mathrm{mcd}(q, k) = 1$ (y todos ellos números enteros). Entonces, si para cada $N = \sum_{j=0}^{n} d_j 10^{m \cdot j}$ denotamos $N' = \sum_{j=0}^{n} d_j (-k)^j K^{n-j}$, se cumple que $q \mid N$ si y sólo si $q \mid N'$.

Demostración. Comencemos probando que $\mathrm{mcd}(q, 10^m) = 1$. En efecto, supongamos que, para algún $t > 1$, $t \mid q$ y $t \mid 10^m$. Como $\mathrm{mcd}(q, k) = 1$ será $t \nmid k$, luego $t \nmid (K \cdot 10^m + k)$; y esto es absurdo ya que $t \mid q$ y $q \mid (K \cdot 10^m + k)$.

Dado N como en el enunciado, escribámoslo como $N = E_0 \cdot 10^m + e_0$ (con $E_0 = \sum_{j=1}^{n} d_j 10^{m(j-1)}$ y $e_0 = d_0$) y tomemos $N_1 = kE_0 - Ke_0$. Entonces $kN - 10^m N_1 = e_0(K \cdot 10^m + k)$ y así $q \mid (kN - 10^m N_1)$ (puesto que, por hipótesis, $q \mid (K \cdot 10^m + k)$). Por tanto $q \mid kN \Leftrightarrow q \mid 10^m N_1$. Además, $\mathrm{mcd}(q, k) = 1$, luego $q \mid N \Leftrightarrow q \mid kN$; y $\mathrm{mcd}(q, 10^m) = 1$, luego $q \mid N_1 \Leftrightarrow q \mid 10^m N_1$. En consecuencia, $q \mid N \Leftrightarrow q \mid N_1$.

Reiterando el proceso, $q \mid N_1 \Leftrightarrow q \mid N_2$ con $N_2 = kE_1 - Ke_1$ siendo $E_1 = \sum_{j=2}^{n} d_j k 10^{m(j-2)}$ y $e_1 = kd_1 - Kd_0$ (ya que $N_1 = \sum_{j=1}^{n} d_j k 10^{m(j-1)} - Kd_0$).

De nuevo $q \mid N_2 \Leftrightarrow q \mid N_3$ con $N_3 = kE_2 - Ke_2$ siendo $E_2 = \sum_{j=3}^{n} d_j k^2 10^{m(j-3)}$ y $e_2 = k^2 d_2 + (-k)Kd_1 + K^2 d_0$ (ya que, $N_2 = \sum_{j=2}^{n} d_j k^2 10^{m(j-2)} + (-k)Kd_1 + K^2 d_0$, por el caso anterior).

Continuando de esta forma tendremos que

$$q \mid N \;\Leftrightarrow\; q \mid N_1 \;\Leftrightarrow\; q \mid N_2 \;\Leftrightarrow\; \cdots \;\Leftrightarrow\; q \mid N_n,$$

donde $N_n = N'$, si n es par, o $N_n = -N'$, en otro caso. □

Veamos unos cuantos casos particulares de este resultado:

Ejemplo. Tomando $K = m = 1$ y $k = -1$ se obtienen los criterios de divisibilidad por 3 y 9; con $K = k = 1$ y $m = 3$ aparecen los correspondientes a 7, 11 y 13, ya que en este caso $K \cdot 10^m + k = 7 \cdot 11 \cdot 13$. Uno que no hemos mencionado hasta ahora es el que comparten 27 y 37, que surge de $999 = 3^3 \cdot 37$ (es decir, basta tomar $K = 1$, $k = -1$ y $m = 3$). Ya más en el espíritu generalizador del teorema podemos

citar los casos $23\,|\,2003$ (y también $29\,|\,2003$), $67\,|\,1005$ y $19\,|\,3002$, que dan lugar a los correspondientes criterios de divisibilidad que no detallamos. ◊

Pero la época en que los matemáticos efectuaban los cálculos con lápiz y papel —en la que los criterios de divisibilidad podían ahorrar mucho tiempo— ya ha pasado, y no merece la pena seguir insistiendo ni analizar con cuidado cuándo es mejor buscar un criterio o —simplemente— dividir. Ahora que las operaciones aritméticas nos las hacen los ordenadores, todo esto ya no tiene mucha utilidad práctica. En todo caso, nos ha servido de excusa para jugar un rato con los números. ¿A qué matemático no le gusta?

1.4. Combinatoria

La combinatoria es una parte esencial de la matemática discreta que tiene aplicación práctica en diversas ramas de las ciencias, e incluso en las letras y en las artes. Existen multitud de problemas combinatorios lo suficientemente atractivos y elementales como para que la combinatoria haya sido ampliamente usada en los aspectos más lúdicos de la matemática, así como en su divulgación entre no especialistas. Además, hoy en día es uno de los temas de investigación matemática más activos, pues presta importantes servicios a la ciencia de los ordenadores. Pero los aspectos más avanzados de la combinatoria quedan totalmente fuera de las pretensiones de este libro, así que aquí nos vamos a contentar con arañar en la superficie, presentando sólo algunos de sus problemas más clásicos —y más básicos—. Dados n elementos, se trata de estudiar cuántas configuraciones se pueden conseguir con ellos, con la condición de que tales configuraciones cumplan cierta propiedad previamente exigida.

Aunque tampoco nos vamos a ocupar de ello, no podemos dejar de citar —pero nos conformaremos con eso— la importancia de la combinatoria en teoría de probabilidad (para ser más precisos, en teoría de probabilidad discreta). Seguro que el lector recuerda que la probabilidad de que ocurra un suceso es «el número de casos favorables dividido entre el número de casos posibles» (naturalmente, estamos dando por hecho que todos los casos son equiprobables). Y es la combinatoria la herramienta esencial para calcular tanto el número de casos favorables como el de los posibles.

La pregunta clásica en combinatoria elemental se formula como sigue: «de entre n elementos, ¿de cuántas maneras podemos coger k de ellos, de manera que...?». Ahora, habrá que tener en cuenta si en los grupos[5] de k elementos que formemos

[5] En esta sección emplearemos la palabra «grupo» a menudo, siempre con un sentido genérico de «colección de objetos», nunca con el sentido específico al que alude la teoría de grupos.

distinguimos o no el orden, si cada elemento lo podemos coger una o varias veces, si los n elementos de partida son distintos entre sí o no, etcétera.

Así, se acostumbra a distinguir tres diferentes clases de configuraciones de objetos, que dan lugar a las tres ramas básicas tradicionales: variaciones, permutaciones y combinaciones. Además, podremos agregarles la coletilla de «con repetición» o «sin repetición».

1.4.1. Variaciones

Llamamos *variaciones sin repetición* (o *variaciones simples*) de n elementos, tomados de k en k, a los diferentes grupos que con ellos se pueden formar, de tal modo que en cada grupo haya k elementos distintos, y que un grupo se diferencie de los demás en alguno de los elementos o en la colocación de los mismos. Denotaremos su número con V_n^k.

Obviamente, en la definición anterior estamos suponiendo implícitamente que n y k son enteros positivos. Además, podemos dar por sentado que $k \leq n$ pues, si fuera $k > n$, es obvio que no se puede formar ningún grupo de k elementos todos ellos distintos ($V_n^k = 0$ si $k > n$). Sí que tiene sentido el caso trivial $k = 0$ (esta vez, incluso con $n = 0$), y se cumple $V_n^0 = 1$, ya que podemos formar un único grupo con 0 elementos: el vacío.

En general, ¿cuánto vale V_n^k? Para calcularlo, formemos los grupos de k elementos tomándolos uno a uno. El primer elemento puede ser cualquiera de los n que tenemos a nuestra disposición; en particular —si sólo buscábamos grupos de un elemento—, $V_n^1 = n$. El segundo elemento lo podemos elegir de $n - 1$ formas distintas (pues no podemos coger el ya elegido antes); así, $V_n^2 = n(n-1)$. Del mismo modo tendremos $V_n^3 = n(n-1)(n-2)$. Y, en general,

$$V_n^k = n(n-1)(n-2)\cdots(n-k+1), \qquad k \leq n$$

(obsérvese que esta fórmula es válida incluso con $k = 0$ si se considera, como es habitual, que un producto de 0 factores vale 1).

Nota 1.21. En el lenguaje de la teoría de conjuntos, las variaciones sin repetición se definen como sigue: Dado E un conjunto de n elementos, las variaciones sin repetición de n elementos tomados de k en k son las aplicaciones inyectivas del conjunto $\{1, 2, \ldots, k\}$ en E. De este modo, V_n^k indica cuántas aplicaciones inyectivas de $\{1, 2, \ldots, k\}$ en E existen. ◁

Llamamos *variaciones con repetición* de n elementos, tomados de k en k, a los diferentes grupos que con ellos se pueden formar, de tal modo que en cada grupo haya k elementos, pudiendo alguno repetirse una o varias veces, y teniendo en cuenta

que dos grupos se consideran distintos si difieren en algún elemento o en el orden en que están colocados. Denotaremos su número con VR_n^k.

Al contrario de lo que ocurría con las variaciones sin repetición, los grupos de k elementos se pueden formar aunque ya no se cumpla $k \leq n$. Así pues, k y n son ahora enteros positivos arbitrarios. El caso $k = 0$ también está permitido, y es claro que $VR_n^0 = 1$.

Computar el número de variaciones con repetición es muy sencillo. Al formar los grupos de k elementos a los que aludimos en la definición, el primer elemento puede ser cualquiera de los n; el segundo, también (luego existen n^2 grupos de 2 elementos); y el tercero (así que hay n^3 grupos de 3 elementos)... Prosiguiendo de igual modo,

$$VR_n^k = n^k,$$

fórmula que también es válida para $k = 0$, ya que $n^0 = 1$, el grupo vacío.

Nota 1.22. Como antes, también podemos definir esto en términos de la teoría de conjuntos. Dado E un conjunto de n elementos, las variaciones con repetición de n elementos tomados de k en k son las aplicaciones de $\{1, 2, \ldots, k\}$ en E. Así pues, hay exactamente $VR_n^k = n^k$ aplicaciones de $\{1, 2, \ldots, k\}$ en E. ◁

1.4.2. Permutaciones

Llamamos *permutaciones* de n elementos a los diversos grupos que con ellos se pueden formar, de modo que todos los elementos entren en cada grupo, y un grupo se diferencie de otro solamente en el orden de colocación de los elementos. Usaremos P_n para denotar su número.

En otras palabras, las permutaciones de n elementos son las variaciones sin repetición de n elementos tomados de n en n. Por tanto,

$$P_n = V_n^n = n!,$$

válido para $n \geq 0$.

Nota 1.23. Por supuesto, de nuevo podemos describir las permutaciones usando la teoría de conjuntos. Sin más que aplicar lo que hemos dicho para las variaciones sin repetición, y si tomamos E un conjunto de n elementos, las permutaciones de n elementos son las aplicaciones inyectivas de $\{1, 2, \ldots, n\}$ en E; o, lo que es equivalente, las aplicaciones inyectivas de E en sí mismo. Pero, si nos fijamos un poco, nos damos cuenta de que cualquier aplicación inyectiva de un conjunto finito en sí mismo es forzosamente biyectiva. En consecuencia, las permutaciones de n elementos son las aplicaciones biyectivas de un conjunto de n elementos en sí mismo. Hay, pues, $n!$ de tales aplicaciones. ◁

Definir *permutaciones con repetición* es un poco más delicado, y ya no tiene ningún paralelismo con las variaciones con repetición.

Esta vez, tenemos n elementos entre los cuales hay un cierto número a_1 de una misma clase, otro número a_2 de otra clase, un tercer número a_3 de otra, y así sucesivamente, hasta a_k elementos de cierta k-ésima clase. En esta situación, llamamos permutaciones con repetición a las diferentes formas en que se pueden ordenar esos n elementos, de modo que una ordenación se considera distinta de otra si tienen dos elementos distintos ocupando la misma posición. Usaremos $P_n^{a_1,a_2,\ldots,a_k}$ para denotar su número.[6]

Obviamente, aquí n y los a_j son enteros positivos, y se tiene que cumplir $a_1 + a_2 + \cdots + a_k = n$. Si $a_j = 1$ para todo j nos encontramos ante el caso de las permutaciones ordinarias. Si eso no ocurre, ¿cuánto vale $P_n^{a_1,a_2,\ldots,a_k}$?

Para calcularlo, consideremos momentáneamente que los a_j elementos de la clase j son distintos entre sí, y eso para cada j. De este modo, tendríamos que los n elementos son diferentes, y el número de ordenaciones posibles con ellos sería $P_n = n!$. Por otra parte, con los a_j elementos de la clase j se forman $P_{a_j} = a_j!$ permutaciones.

Volvamos a la situación de partida comenzando con los a_1 elementos de la primera clase, que de nuevo hacemos indistinguibles entre sí. De este modo, muchas de las $P_n = n!$ ordenaciones son realmente la misma; solamente quedan P_n/P_{a_1} distintas. Lo mismo ocurre al devolver su identidad común a los a_2 elementos de la segunda clase: el nuevo número de ordenaciones distintas se obtiene dividiendo por P_{a_2}. Y así sucesivamente. De este modo llegamos a la expresión

$$P_n^{a_1,a_2,\ldots,a_k} = \frac{P_n}{P_{a_1} P_{a_2} \cdots P_{a_k}} = \frac{n!}{a_1!\, a_2! \cdots a_k!},$$

que proporciona el número de permutaciones con repetición.

Las permutaciones con repetición aparecen en la denominada fórmula de Leibniz:

$$(x_1 + x_2 + \cdots + x_k)^n = \sum_{a_1+a_2+\cdots+a_k=n} \frac{n!}{a_1!\, a_2! \cdots a_k!}\, x_1^{a_1} x_2^{a_2} \cdots x_k^{a_k}. \qquad (1.12)$$

Obsérvese que el coeficiente de $x_1^{a_1} x_2^{a_2} \cdots x_k^{a_k}$ es $P_n^{a_1,a_2,\ldots,a_k}$. Comprobemos la fórmula (1.12) razonando de modo similar al procedimiento seguido para calcular el valor de $P_n^{a_1,a_2,\ldots,a_k}$.

[6] Muchas veces, para simplificar la notación, se acostumbra a no incluir en el exponente los elementos no repetidos; así, en lugar de $P_{15}^{4,3,3,2,1,1,1}$ se escribiría $P_{15}^{4,3,3,2}$. Pero, en esta pequeña introducción, no hace falta que nos preocupemos por ello.

Para cada j, distingamos momentáneamente los elementos de tipo x_j añadiéndoles un subíndice adicional, y transformándolos en $x_{j,1}, x_{j,2}, \ldots, x_{j,n}$. Entonces, sin más que aplicar la propiedad distributiva del producto, tenemos

$$(x_{1,1} + x_{2,1} + \cdots + x_{k,1})(x_{1,2} + x_{2,2} + \cdots + x_{k,2}) \cdots (x_{1,n} + x_{2,n} + \cdots + x_{k,n}) = \sum_{a_1+a_2+\cdots a_k=n} x_{1,1} \cdots x_{1,a_1} x_{2,1} \cdots x_{2,a_2} \cdots x_{k,1} \cdots x_{k,a_k}.$$

Si de nuevo identificamos $x_{j,1} = x_{j,2} = \cdots = x_{j,n} = x_j$ (y esto, para cada j), obtenemos precisamente $P_n^{a_1,a_2,\ldots,a_k}$ sumandos $x_1^{a_1} x_2^{a_2} \cdots x_1^{a_k}$, de donde se sigue (1.12).

El caso más importante es el que corresponde a $k = 2$, que da lugar al denominado binomio de Newton:

$$(x+y)^n = \sum_{k=0}^{n} \frac{n!}{k!\,(n-k)!}\, x^k y^{n-k},$$

que analizaremos con más detalle un poco más adelante.

1.4.3. Combinaciones

Se llaman *combinaciones sin repetición* (o *combinaciones simples*) de n elementos, tomados de k en k, a los diferentes grupos que con ellos se pueden formar, de modo que cada grupo tenga k elementos, y teniendo en cuenta que dos grupos se consideran distintos si difieren en algún elemento. Es habitual no usar la coletilla «sin repetición», y aludir sólo a *combinaciones* de n elementos tomados de k en k. Su número se representa mediante C_n^k.

Para evaluar C_n^k notemos primero que, forzosamente, k y n deben ser enteros no negativos. Si $k > n$, es obvio que $C_n^k = 0$, luego podemos asumir que $0 \le k \le n$. El caso $k = 0$ también es evidente: $C_n^0 = 1$, lo que resulta válido incluso cuando $k = n = 0$.

En otro caso, ¿cuánto vale C_n^k? Dados n y k enteros positivos, supongamos que hemos logrado obtener todas las combinaciones de n elementos tomados de k en k. Cada una de estas combinaciones —hay C_n^k de ellas— es un grupito de k elementos. Si permutamos los k elementos de cada uno de estos grupitos, lo que obtenemos son todas las variaciones sin repetición de n elementos tomados de k en k. Es decir, $C_n^k \cdot P_k = V_n^k$. Por tanto,

$$C_n^k = \frac{V_n^k}{P_k} = \frac{n(n-1)\cdots(n-k+1)}{k!}.$$

Si consideramos que $0! = 1$, esa expresión también es válida con $k = 0$.

Una manera alternativa de computar C_n^k es darse cuenta de que las combinaciones de n elementos, tomados de k en k, son lo mismo que las permutaciones con repetición de n elementos, k de un tipo y los restantes $n-k$ de otro tipo. Así pues,

$$C_n^k = P_n^{k,n-k} = \frac{n!}{k!(n-k)!} = \frac{n(n-1)\cdots(n-k+1)}{k!}$$

que, por supuesto, coincide con la expresión anterior.

Por razones obvias, a los números C_n^k los denominamos *números combinatorios* y, por su presencia en el binomio de Newton, también los conocemos como coeficientes binomiales. En realidad, estos números son tan habituales en matemáticas que tienen una notación específica para ellos; más que C_n^k (o expresiones similares), los denotamos mediante $\binom{n}{k}$. Así pues, el número de combinaciones de n elementos tomados de k en k es

$$\binom{n}{k} = \frac{n(n-1)\cdots(n-k+1)}{k!} = \frac{n!}{k!\,(n-k)!}, \quad 0 \le k \le n,$$

y $\binom{n}{k} = 0$ si $k > n$. Pero dejémoslos de momento —ya habrá tiempo en la sección 1.5 de estudiarlos con algo más de profundidad— y volvamos a la combinatoria.

Nota 1.24. Usando teoría de conjuntos, las combinaciones se definen de manera elemental. Partamos de un conjunto de n elementos. Entonces, las combinaciones de n elementos tomados de k en k son todos los posibles subconjuntos de k elementos que se pueden obtener. (Recuerde el lector que, en un conjunto —tal como se entiende en matemáticas—, los elementos no se encuentran ordenados de ninguna manera.) Así pues, un conjunto de n elementos tiene $C_n^k = \binom{n}{k}$ subconjuntos de k elementos. ◁

Llamamos *combinaciones con repetición* de n elementos, tomados de k en k, a los diferentes grupos que con ellos se pueden formar, de tal modo que en cada grupo haya k elementos, pudiendo alguno repetirse una o varias veces, y teniendo en cuenta que dos grupos se consideran distintos si difieren en algún elemento. Usaremos CR_n^k para denotar su número.

Como sucede con las variaciones con repetición, los grupos de k elementos se pueden formar aunque no se cumpla $k \le n$. Entonces, k y n son enteros, con $n \ge 1$ y $k \ge 1$. Si $k = 0$, es obvio que $CR_n^0 = 1$ para cualquier $n \ge 1$ (incluso podemos llegar a admitir que $n = k = 0$, también con $CR_0^0 = 1$).

En otro caso, ¿cuánto vale CR_n^k? Vamos a demostrar que

$$CR_n^k = C_{n+k-1}^k = \binom{n+k-1}{k}, \quad n, k \ge 1.$$

Para ello, construiremos una biyección entre las combinaciones con repetición de un conjunto $\{x_1, x_2, \ldots, x_n\}$, tomadas de k en k, y las combinaciones ordinarias (sin repetición) de otro conjunto $\{y_1, y_2, \ldots, y_{n+k-1}\}$, tomadas también de k en k. (En la explicación que sigue, como el «tomadas de k en k» es el mismo en unas y en otras, ya no lo explicitaremos más.)

Una combinación con repetición de $\{x_1, x_2, \ldots, x_n\}$ será un «grupo» $x_{r_1} x_{r_2} \cdots x_{r_k}$ con $1 \leq r_1 \leq r_2 \leq \cdots \leq r_k \leq n$ (vamos a usar notación multiplicativa, como si fuesen variables de un polinomio en n variables); además, esto representa una de tales combinaciones de manera única. De modo similar, una combinación sin repetición de $\{y_1, y_2, \ldots, y_{n+k-1}\}$ será un «grupo» $y_{s_1} y_{s_2} \cdots y_{s_k}$ con $1 \leq s_1 < s_2 < \cdots < s_k \leq n+k-1$. A cada combinación con repetición de las x_r asignémosle una combinación sin repetición de las y_s mediante

$$x_{r_1} x_{r_2} x_{r_3} \cdots x_{r_k} \mapsto y_{r_1} y_{r_2+1} y_{r_3+2} \cdots y_{r_k+n-1}.$$

Para ver que esto hace corresponder exactamente las combinaciones con repetición de las x_r con las combinaciones sin repetición de las y_s fijémonos en los respectivos conjuntos de subíndices

$$R = \{(r_1, r_2, r_3, \cdots, r_k) : 1 \leq r_1 \leq r_2 \leq \cdots \leq r_k \leq n\}$$

y

$$S = \{(s_1, s_2, s_3, \cdots, s_k) : 1 \leq s_1 < s_2 < \cdots < s_k \leq n+k-1\}.$$

Entre R y S definimos la aplicación

$$(r_1, r_2, r_3, \ldots, r_k) \mapsto (r_1, r_2+1, r_3+2, \ldots, r_k+n-1)$$

que no sólo está bien definida (es decir, lleva una k-tupla de R a una de S), sino que es inyectiva y suprayectiva (la comprobación es inmediata). En consecuencia, los conjuntos R y S tienen el mismo cardinal y, por tanto, $CR_n^k = C_{n+k-1}^k$, tal como buscábamos.

Concluyamos la sección mencionando un ejemplo típico en el que aparecen las combinaciones con repetición. Recordemos que, si una función de n variables $f : \mathbb{R}^n \to \mathbb{R}$ tiene derivadas parciales hasta orden k y son continuas, entonces el valor de esas derivadas parciales es independiente del orden de derivación.[7] En

[7] Bajo esas hipótesis, se dice que la función es de clase C^k en $\mathbb{R}^n$. Puede demostrarse que el valor de las derivadas parciales es independiente del orden de derivación con hipótesis algo más débiles que la pertenencia de f a la clase C^k; pero aquí sólo estamos dando un ejemplo, así que no prestaremos atención a estos detalles.

tales circunstancias, ¿cuántas derivadas parciales de orden k tiene una función $f(x_1, x_2, \ldots, x_n)$? Estas derivadas parciales se expresan mediante

$$\frac{\partial^k}{\partial x_1^{k_1} \partial x_2^{k_2} \cdots \partial x_n^{k_n}} f(x_1, x_2, \ldots, x_n)$$

con $k_j \geq 0$, para $1 \leq j \leq n$ y $k_1 + k_2 + \cdots + k_n = k$, y donde (como es habitual) si alguno de los k_j es 0 significa que no hemos derivado respecto a la variable x_j. Así, las derivadas parciales de orden k se corresponden con las combinaciones con repetición de las variables $x_1, x_2, \ldots, x_n$ tomadas de k en k. En consecuencia, existen $CR_n^k = \binom{n+k-1}{k}$ derivadas parciales de orden k.

1.5. Coeficientes binomiales y binomio de Newton

Aunque tanto los coeficientes binomiales como el binomio de Newton nos han aparecido con anterioridad, vamos ahora a prescindir de ello, de modo que en esta sección estudiaremos sus propiedades partiendo de cero. Lo que vamos a mostrar aquí será riguroso, aunque básico; con casi total seguridad, el lector ya estará familiarizado con todo ello. Si lo que desea es encontrar un estudio más profundo, que muestre muchas más propiedades e identidades relacionadas con los coeficientes binomiales, le animamos a que consulte [47, capítulo 5].

Dados n y k dos enteros no negativos, con $0 \leq k \leq n$, definimos el *coeficiente binomial* $\binom{n}{k}$ como[8]

$$\binom{n}{k} = \frac{n(n-1)\cdots(n-k+1)}{k!} = \frac{n!}{k! \cdot (n-k)!} \tag{1.13}$$

(con el convenio habitual $0! = 1$). Cuando $k > n$, se define $\binom{n}{k} = 0$.

De (1.13) se sigue que[9]

$$\binom{n}{k} = \frac{n!}{k! \cdot (n-k)!} = \frac{n!}{(n-k)! \cdot (n-(n-k))!} = \binom{n}{n-k}, \quad 0 \leq k \leq n.$$

[8] Aunque la expresión $\frac{n!}{k! \cdot (n-k)!}$ tiene una escritura más compacta que $\frac{n(n-1)\cdots(n-k+1)}{k!}$, es evidente que, como procedimiento operacional para evaluar $\binom{n}{k}$, es mucho más laboriosa.

[9] Como consecuencia, si k es próximo a n, lo más cómodo es calcular $\binom{n}{k}$ mediante

$$\binom{n}{k} = \binom{n}{n-k} = \frac{n(n-1)\cdots(k+1)}{(n-k)!}.$$

Si se quiere que esta expresión (la parte derecha) también tenga sentido para enteros $k > n$, tenemos que definir $\binom{n}{k} = 0$ para $k < 0$. Este hecho se asume habitualmente.

Pero, sin duda, la propiedad más característica de los coeficientes binomiales es

$$\binom{n}{k} = \binom{n-1}{k-1} + \binom{n-1}{k}, \qquad 1 \leq k \leq n. \tag{1.14}$$

En realidad, su demostración es muy sencilla, como vemos a continuación:

$$\begin{aligned}
\binom{n-1}{k-1} + \binom{n-1}{k} &= \frac{(n-1)!}{(k-1)!\cdot[(n-1)-(k-1)]!} + \frac{(n-1)!}{k!\cdot[(n-1)-k)]!} \\
&= (n-1)!\cdot\left(\frac{1}{(k-1)!\cdot(n-k)!} + \frac{1}{k!\cdot[(n-k)-1)]!}\right) \\
&= (n-1)!\cdot\frac{k+(n-k)}{k!\cdot(n-k)!} \\
&= \frac{n!}{k!\cdot(n-k)!} = \binom{n}{k}.
\end{aligned}$$

La fórmula (1.14) es una relación recurrente que permite dar una demostración por inducción para el desarrollo en potencias de $(x+y)^n$. Veámoslo:

Teorema 1.25. Binomio de Newton.
Dado un entero n no negativo, se verifica

$$(x+y)^n = \sum_{k=0}^{n} \binom{n}{k} x^{n-k} y^k. \tag{1.15}$$

Demostración. En primer lugar, démonos cuenta de que basta probar la identidad

$$(1+z)^n = \sum_{k=0}^{n} \binom{n}{k} z^k. \tag{1.16}$$

En efecto; tomando $z = y/x$ y multiplicando por x^n se obtiene (1.15) (la fórmula del binomio es trivial si $x = 0$, luego no hace falta preocuparse por este caso).

Así pues, probemos (1.16) por inducción. Si $n = 0$ o $n = 1$, es obvio. Supongámoslo cierto para $n-1$, es decir,

$$(1+z)^{n-1} = \sum_{k=0}^{n-1} \binom{n-1}{k} z^k,$$

y comprobémoslo para n:

$$\begin{aligned}(1+z)^n &= (1+z)\sum_{k=0}^{n-1}\binom{n-1}{k}z^k = \sum_{k=0}^{n-1}\binom{n-1}{k}z^k + \sum_{k=0}^{n-1}\binom{n-1}{k}z^{k+1}\\ &= \sum_{k=0}^{n-1}\binom{n-1}{k}z^k + \sum_{k=1}^{n}\binom{n-1}{k-1}z^k\\ &= 1 + \sum_{k=1}^{n-1}\left[\binom{n-1}{k} + \binom{n-1}{k-1}\right]z^k + z^n\\ &= \sum_{k=0}^{n}\binom{n}{k}z^k,\end{aligned}$$

donde en el último paso hemos usado (1.14) y $\binom{n}{0} = \binom{n}{n} = 1$. □

Nota 1.26. El desarrollo del ahora denominado «binomio de Newton» era bien conocido antes de que Sir Isaac Newton viviera. En realidad, la principal contribución de Newton en esto fue extender el desarrollo

$$(1+x)^n = \sum_{k=0}^{n}\binom{n}{k}x^k$$

a potencias fraccionarias o negativas. Para ello, se generaliza la definición (1.13) al caso $\alpha \in \mathbb{R}$, $k \in \mathbb{Z}$, $k \geq 0$, mediante[10]

$$\binom{\alpha}{k} = \frac{\alpha(\alpha-1)\cdots(\alpha-k+1)}{k!}. \tag{1.17}$$

Entonces, $\binom{\alpha}{k}$ tiene sentido para cualquier entero $k \geq 0$, y el desarrollo de $(1+x)^\alpha$ se convertiría de este modo en una serie infinita.

Así pues, lo que Newton hizo fue asumir que, para α arbitrario, se cumplía

$$(1+x)^\alpha = \sum_{k=0}^{\infty}\binom{\alpha}{k}x^k, \tag{1.18}$$

[10] La correspondiente expresión mediante factoriales que aparece en (1.13) también se puede generalizar pero, para nuestros propósitos, no lo necesitamos. No obstante, comentemos que esto se hace utilizando la denominada función Γ de Euler. Para números enteros no negativos esta función cumple $\Gamma(k+1) = k!$, y con este sentido podemos decir que la función Γ extiende el factorial.

pero nunca lo probó rigurosamente. Actualmente sabemos que (1.18) es el desarrollo de Taylor de $(1+x)^\alpha$, que es convergente siempre que $|x|<1$. Cuando $\alpha = n$, un entero positivo, esta fórmula se convierte en la habitual con $\sum_{k=0}^{n}$, ya que (1.17) da $\binom{n}{k} = 0$ para $k > n$. ◁

Ya hemos comentado que las identidades en las que aparecen coeficientes binomiales son muy numerosas, y que aquí no nos vamos a dedicar a estudiarlas. Pero no nos resistimos a mostrar alguna. Es más, de este modo vamos a ver que el binomio de Newton —que acabamos de demostrar a partir de una propiedad de los coeficientes binomiales— va a resultar a su vez muy útil para deducir nuevas identidades relacionadas con estos coeficientes. En concreto, vamos a probar lo siguiente:

Proposición 1.27. Chu Shih-Chieh, 1303, y Vandermonde, 1772.
Sean $r, s \in \mathbb{Z}$, $r, s \geq 0$.[11] *Entonces, para cada entero no negativo n se cumple*

$$\sum_{k=0}^{n} \binom{r}{k}\binom{s}{n-k} = \binom{r+s}{n}. \tag{1.19}$$

Demostración. En el desarrollo en potencias de $(1+x)^{r+s}$, el coeficiente de x^n es $\binom{r+s}{n}$. Por otra parte, si hacemos[12]

$$(1+x)^{r+s} = (1+x)^r(1+x)^s = \left(\sum_{k=0}^{\infty}\binom{r}{k}x^k\right)\left(\sum_{j=0}^{\infty}\binom{s}{j}x^j\right),$$

el coeficiente de x^n al multiplicar las series es $\sum_{k=0}^{n}\binom{r}{k}\binom{s}{n-k}$. □

En honor de Alexandre Vandermonde, que escribió un artículo sobre ella en el siglo XVIII, a la fórmula (1.19) se la denomina *identidad de Vandermonde* o *convolución de Vandermonde*. Sin embargo, ya era conocida en el siglo XIV por Chu Shih-Chieh (este matemático chino nos aparecerá de nuevo un poco más adelante en relación con el triángulo de Pascal), por lo que a (1.19) también se le llama *identidad de Chu-Vandermonde*.

Nota 1.28. La identidad de Chu-Vandermonde puede probarse de forma alternativa por medio de una bonita interpretación combinatoria. Dados dos enteros positivos r y s, el lado derecho de (1.19) representa el número de maneras de elegir

[11] En realidad, estas hipótesis sobre r y s son superfluas si, para el caso más general $r, s \in \mathbb{R}$, asumimos la definición (1.17) y la identidad (1.18).

[12] Incluso aunque r y s sean enteros positivos y las series que siguen sean finitas pues $\binom{r}{k} = 0$ si $k > r$ y $\binom{s}{j} = 0$ si $j > s$, para identificar coeficientes resulta más cómodo trabajar con las series infinitas.

n personas entre r mujeres y s hombres. En el lado izquierdo, cada término de la suma es el número de maneras de elegir k mujeres y $n-k$ hombres; al sumar sobre k contamos cada posible elección exactamente una vez. ◁

1.5.1. El triángulo de Pascal

La fórmula (1.14) permite calcular los $\binom{n}{k}$ de manera recurrente, sin más que tener en cuenta que $\binom{n}{0} = \binom{n}{n} = 1$ para cada n. Lo habitual es colocar los $\binom{n}{k}$ en forma de triángulo, como sigue:

$$
\begin{array}{ccccccccccccccccc}
 & & & & & & & \binom{0}{0} & & & & & & & & & \\
 & & & & & & \binom{1}{0} & & \binom{1}{0} & & & & & & & & \\
 & & & & & \binom{2}{0} & & \binom{2}{1} & & \binom{2}{2} & & & & & & & \\
 & & & & \binom{3}{0} & & \binom{3}{1} & & \binom{3}{2} & & \binom{3}{3} & & & & & & \\
 & & & \binom{4}{0} & & \binom{4}{1} & & \binom{4}{2} & & \binom{4}{3} & & \binom{4}{4} & & & & & \\
 & & \binom{5}{0} & & \binom{5}{1} & & \binom{5}{2} & & \binom{5}{3} & & \binom{5}{4} & & \binom{5}{5} & & & & \\
 & \binom{6}{0} & & \binom{6}{1} & & \binom{6}{2} & & \binom{6}{3} & & \binom{6}{4} & & \binom{6}{5} & & \binom{6}{6} & & & \\
\binom{7}{0} & & \binom{7}{1} & & \binom{7}{2} & & \binom{7}{3} & & \binom{7}{4} & & \binom{7}{5} & & \binom{7}{6} & & \binom{7}{7} & &
\end{array}
$$

$$
\begin{array}{cccccccccccccccccc}
\binom{8}{0} & & \binom{8}{1} & & \binom{8}{2} & & \binom{8}{3} & & \binom{8}{4} & & \binom{8}{5} & & \binom{8}{6} & & \binom{8}{7} & & \binom{8}{8}
\end{array}
$$

En la izquierda y en la derecha de este triángulo siempre tenemos $\binom{n}{0} = \binom{n}{n} = 1$. Y, en cuanto a su interior, la fórmula (1.14) indica que para calcular $\binom{n}{k}$ (con $1 < k < n$) basta sumar los dos valores $\binom{n-1}{k-1}$ y $\binom{n-1}{k}$ que aparecen, sobre él, en la fila precedente (a izquierda y derecha, respectivamente). De este modo, lo que se van obteniendo es

$$
\begin{array}{ccccccccccccccccc}
 & & & & & & & 1 & & & & & & & & & \\
 & & & & & & 1 & & 1 & & & & & & & & \\
 & & & & & 1 & & 2 & & 1 & & & & & & & \\
 & & & & 1 & & 3 & & 3 & & 1 & & & & & & \\
 & & & 1 & & 4 & & 6 & & 4 & & 1 & & & & & \\
 & & 1 & & 5 & & 10 & & 10 & & 5 & & 1 & & & & \\
 & 1 & & 6 & & 15 & & 20 & & 15 & & 6 & & 1 & & & \\
1 & & 7 & & 21 & & 35 & & 35 & & 21 & & 7 & & 1 & &
\end{array}
$$

$$
\begin{array}{cccccccccccccccccc}
1 & & 8 & & 28 & & 56 & & 70 & & 56 & & 28 & & 8 & & 1
\end{array}
$$

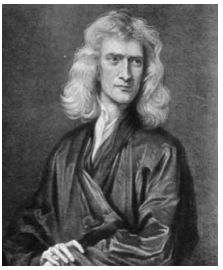
A esta disposición de números en forma triangular se le suele denominar *triángulo de Pascal*. Tal nombre se debe a que, a mediados del siglo XVII, Blaise Pascal escribió un influyente tratado en el que describía el triángulo y algunas de sus propiedades; por ejemplo, él se percató de que los términos del desarrollo del binomio



Blaise Pascal
(1623–1662)



Omar Khayyam
(1048–1131)



Sir Isaac Newton
(1643–1727)

(las filas del triángulo) coincidían con los números combinatorios (que le habían aparecido al resolver problemas probabilísticos). Pero, en realidad, el triángulo era conocido mucho antes. En Europa, ya lo había usado —un siglo antes que Pascal— el matemático italiano Niccolò Fontana «Tartaglia»;[13] por esta razón, en Italia se suele llamar triángulo de Tartaglia.

En otras culturas, el triángulo había surgido con varios siglos de antelación. En China, aparece representado en un libro del matemático Chu Shih-Chieh (cuyo nombre también se transcribe como Zhu Shijie) publicado en 1303; véase la figura 1.2. En ese libro, Chu lo atribuye a Yang Hui (o Yanghui), quien lo habría usado cuarenta años antes. En la literatura china actual, dicho triángulo se sigue denominando habitualmente triángulo de Yang Hui. Pero, en una obra suya, el mismo Yang Hui afirma que él lo aprendió de Jia Xian, un matemático chino del siglo XI (conocido asimismo como Chia Hsien).

El triángulo también había sido descrito por el astrónomo persa Omar Khayyam (o al-Khayyam, y, transcrito de otra forma, Omar Jayyam o incluso Omar ibn al-Jayyam), que vivió a caballo entre los siglos XI y XII. Antes que él, ya al-Karaji (también conocido como al-Karkhi), un matemático que vivió entre los siglos X y XI en Bagdad, había descubierto la regla para construir el triángulo de Pascal. Asimismo, existen referencias a que el triángulo era conocido por matemáticos hindúes hacía varios siglos.

Pero no es éste un libro de historia, así que no proseguiremos con este breve relato cuyo principal objetivo era —simplemente— contribuir a que el lector no olvide que también han existido matemáticas fuera de la cultura occidental de la que somos herederos directos.

[13] El apelativo «Tartaglia» por el que se le conoce habitualmente alude —en italiano— a su tartamudez.

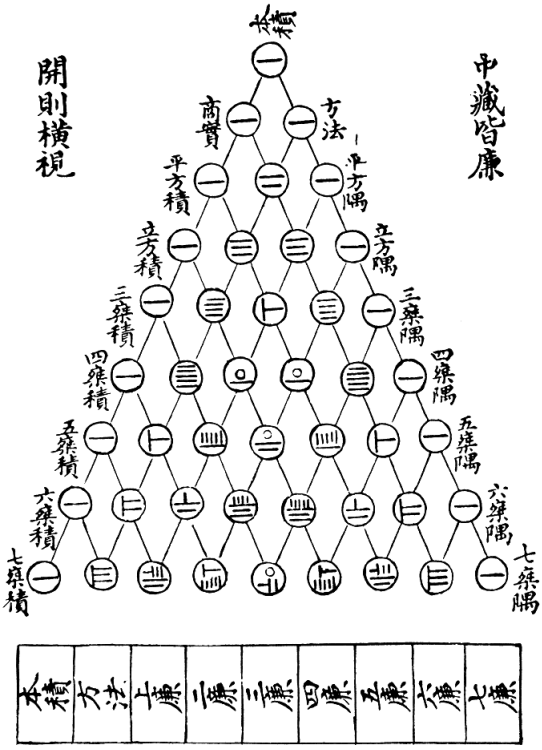


Figura 1.2: El ahora denominado triángulo de Pascal, tal como aparece, con numeración china, en una obra de Chu Shih-Chieh publicada en 1303. Si el lector conoce la numeración china o hace un esfuerzo por entenderla, observará la errata $\binom{7}{3} = 34$ en lugar de 35 (pero su simétrico $\binom{7}{4}$ es correcto).

1.6. Algunas ecuaciones diofánticas

Una *ecuación diofántica* es una ecuación en dos o más variables de la forma

$$f(x_1, \ldots, x_k) = c$$

para la cual queremos encontrar sus soluciones racionales, enteras, o enteras positivas (es decir, que las variables $x_1, \ldots, x_k$ tomen ese tipo de valores). Usualmente,

la función $f(x_1, \ldots, x_k)$ es un polinomio con coeficientes racionales o enteros, y c es una constante racional o entera (c puede estar dada, pero también puede jugar el papel de parámetro móvil).

El nombre de diofánticas que se aplica a estas ecuaciones hace honor a Diofanto. La vida de este matemático griego se desconoce por completo; a través de referencias indirectas se ha podido establecer que vivió en Alejandría, posiblemente en el siglo III d.C.[14] Sin embargo, ha llegado hasta nosotros un texto escrito por él llamado *La Aritmética*, en el que se plantean y resuelven numerosos problemas redactados en términos de ecuaciones o sistemas de ecuaciones; en concreto, *La Aritmética* comprendía trece libros, de los que nos han llegado seis de fuente griega —abarcan 189 problemas— y cuatro más que se descubrieron, en 1972, en un manuscrito árabe —otros 101 problemas—. Diofanto demostró gran creatividad a la hora de encontrar soluciones, aunque sólo admitía las enteras positivas o racionales (los números negativos tardarían muchos siglos en ser admitidos por los matemáticos europeos). Este hecho es lo que ha ocasionado que a las ecuaciones en las que ése es el tipo de soluciones que se busca se denominen ecuaciones diofánticas.

La Aritmética de Diofanto es una de las obras más importantes que se conservan de la antigüedad, y ha servido como referencia y base para estudios e investigaciones posteriores. Grandes matemáticos a lo largo de toda la historia han recurrido a ella: Hipatia (que vivió en el siglo IV en Alejandría y es la primera mujer matemática de la que tenemos conocimiento), Viète, Bachet, Fermat, Frénicle, Descartes, Euler, Jacobi, Lagrange, Legendre, Dirichlet... Fue precisamente en una versión de *La Aritmética* publicada por Bachet en 1621, en la que Fermat realizó su famosa anotación —seguramente la anotación más famosa de la historia— en un margen al lado del problema 8 del libro II, que dio lugar al último teorema de Fermat y que no fue resuelto hasta 1995 por Andrew Wiles. Concluimos esta introducción histórica recomendando [35], una reedición moderna y ampliamente comentada de *La Aritmética* (junto con *Sobre los números poligonales*, otro tratado de Diofanto).

[14] Sí que se conoce la edad de la muerte de Diofanto, porque él mismo se encargó de que se supiera a través de su epitafio:

> Transeúnte, ésta es la tumba de Diofanto: es él quien con esta sorprendente distribución te dice el número de años que vivió. Su niñez ocupó la sexta parte de su vida; después, durante la doceava parte su mejilla se cubrió con el primer vello. Pasó aún una séptima parte de su vida antes de tomar esposa y, cinco años después, tuvo un precioso niño que, una vez alcanzada la mitad de la edad de su padre, pereció de una muerte desgraciada. El padre tuvo que sobrevivirle, llorándole, durante cuatro años. De todo esto se deduce su edad.

Invitamos al lector a que calcule cuántos años vivió Diofanto.



Hipatia de Alejandría
(≈ 370 – 415)



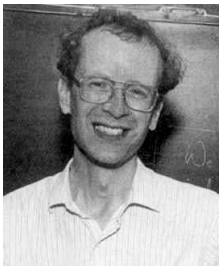
Claude Gaspar Bachet
de Méziriac
(1581–1638)



René Descartes
(1596–1650)



Pierre de Fermat
(1601–1665)



Andrew Wiles (1953–)

La resolubilidad de las ecuaciones diofánticas es un problema de investigación importante, y se está muy lejos de encontrar soluciones satisfactorias. Cuando nos enfrentamos a una ecuación diofántica, existen dos cuestiones básicas:

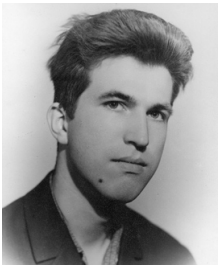
- ¿Existen soluciones?
- ¿Cómo encontrarlas todas?[15]

Obviamente, la segunda es mucho más complicada que la primera, pero ya la primera es difícil de responder en general. Para muchos tipos de ecuaciones, no se conoce la respuesta. Posiblemente una de las mejores maneras de hacernos una idea de la

[15] Diofanto, en su *Aritmética*, no se llega a plantear esto. Ante cada ecuación o sistema de ecuaciones, Diofanto se dedica a mostrar un procedimiento para buscar soluciones particulares (unas veces enteras y otras racionales, y siempre positivas), pero sin preocuparse en absoluto de cómo hallar todas las soluciones.



David Hilbert
(1862–1943)



Yuri Vladimirovich
Matiyasevich (1947–)

complejidad del problema que tenemos entre manos es remontarnos al Segundo Congreso Internacional de Matemáticos,[16] celebrado en París el año 1900.

En ese histórico congreso parisino, David Hilbert pronunció su famosa conferencia *Los problemas futuros de la matemática* en la que formulaba y razonaba 23 problemas matemáticos cuya solución no era conocida. Esta conferencia programática ha tenido una gran influencia en el desarrollo de las matemáticas del siglo XX. Fueron muchos los investigadores que trataron de resolver alguno de los problemas abiertos que había nombrado Hilbert; en el camino, generaron una gran cantidad de nuevas matemáticas. Algunos de estos problemas han sido resueltos —ya sea en sentido positivo o negativo—, pero otros aún siguen constituyendo un reto.[17]

Con un alarde de optimismo, pero también reconociendo la importancia de las ecuaciones diofánticas, el décimo problema de Hilbert preguntaba si existía un algoritmo universal para resolver ecuaciones diofánticas. Fue en 1970 cuando el joven matemático ruso Yuri Matiyasevich probó que tal método aplicable a resolver cualquier ecuación diofántica no podía existir.

Existen muchísimos tipos de ecuaciones diofánticas bien estudiadas, y también otras muchas en las que, sin duda, en el futuro se producirán nuevas e importan-

[16] El primero de estos congresos tuvo lugar en Zurich en 1897, y desde entonces se celebran —con pocas excepciones— cada cuatro años. Su interesante historia, que encantará a cualquier matemático, se relata en [31].

[17] A lo largo de libro comentaremos otros cuatro problemas de los 23 de la lista de Hilbert: el primero, el séptimo, el octavo y el noveno. En concreto, y respectivamente, nos aparecerán la hipótesis del continuo de Cantor (nota 6.15, en la página 356); la trascendencia de ciertos números de la forma α^{β} (apartado 6.8.1, en la página 380); la distribución de los números primos, que incluye la conjetura de Goldbach y la hipótesis de Riemann (nota 8.9, en la página 468, y apartado 10.6.2, en la página 607 y siguientes); y cómo dar una formulación general sobre leyes de reciprocidad (apartado 3.2.5, en la página 193). De los que ahora citamos, permanecen abiertas la conjetura de Goldbach y la hipótesis de Riemann, y parcialmente lo referido a las leyes de reciprocidad.

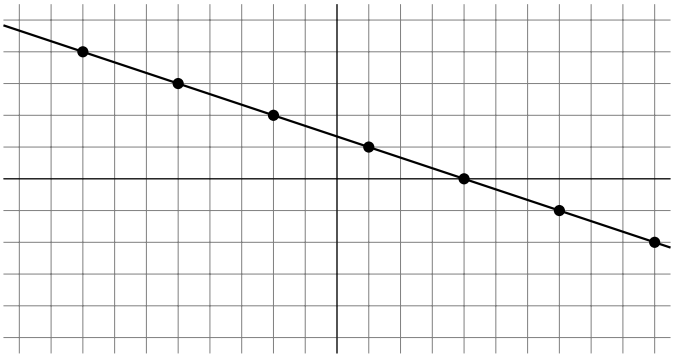


Figura 1.3: Las soluciones enteras de la ecuación diofántica $x + 3y = 4$ son los puntos de coordenadas enteras de la recta $y = -x/3 + 4/3$.

tes aportaciones. En general, las ecuaciones diofánticas se resuelven con técnicas específicas, usando herramientas especialmente adaptadas a cada tipo. En esta sección prácticamente nos contentaremos con analizar dos de los casos más sencillos: la ecuación diofántica lineal y las ternas pitagóricas. Pero, al final, añadiremos un breve apartado en el que —ya sin incluir demostraciones— daremos muestras de algunos otros tipos de ecuaciones diofánticas de interés.

1.6.1. Ecuaciones diofánticas lineales

Se denominan ecuaciones diofánticas lineales a las que tienen la forma

$$a_1x_1 + \cdots + a_kx_k = c,$$

donde los coeficientes a_j son constantes enteras dadas, y c es otro entero. Se trata de saber si esa ecuación tiene o no —dependiendo del valor de c— soluciones enteras, es decir, si existen $x_j \in \mathbb{Z}$, $j = 1, \ldots, k$ que satisfacen la ecuación; nos interesará también conocer si existen soluciones positivas.

Comencemos con el ejemplo más sencillo: la ecuación de dos variables. En ese caso, y tal como se puede ver en la figura 1.3, una ecuación diofántica lineal se corresponde con una recta en el plano, y las soluciones de la ecuación con los puntos de coordenadas enteras de la recta. Pero, en lo que sigue, no utilizaremos ya esta

interpretación geométrica, sino que nos serviremos de procedimientos puramente aritméticos. El primer matemático que describió el método general para resolver la ecuación diofántica lineal $ax + by = c$ fue el indio Brahmagupta en el siglo VII —ya nos hemos topado con él cuando hablábamos de nuestro sistema de numeración en la sección 1.3.

Teorema 1.29.
Sean a, b dos enteros no ambos nulos, $d = \mathrm{mcd}(a, b)$ y c otro entero. La ecuación diofántica

$$ax + by = c \tag{1.20}$$

no tiene soluciones enteras si $d \nmid c$, y tiene infinitas soluciones enteras si $d \mid c$. Además, si $x = x_0$, $y = y_0$ es una solución particular, todas las soluciones enteras tienen la forma

$$x = x_0 + (b/d)n, \quad y = y_0 - (a/d)n, \tag{1.21}$$

donde n recorre los enteros.

Demostración. Comencemos suponiendo que $d \nmid c$. Como $d \mid a$ y $d \mid b$, si existiesen $x, y \in \mathbb{Z}$ tal que $ax + by = c$, debería cumplirse $d \mid ax + by = c$, que es justo lo contrario de lo que hemos supuesto.

Demos pues por hecho que $d \mid c$. En ese caso, la identidad de Bézout (teorema 1.4) asegura que (1.20) tiene solución. Si x_0, y_0 es una solución particular, tomemos x e y como en (1.21). Entonces,

$$ax + by = a\big(x_0 + (b/d)n\big) + b\big(y_0 - (a/d)n\big) = ax_0 + by_0 + 0 = c,$$

luego efectivamente es solución. Nos falta ver que todas las soluciones enteras de (1.20) deben tener la forma (1.21). Dado que $ax_0 + by_0 = c$, si también $ax + by = c$, deberá ser $ax_0 + by_0 = ax + by$. Escribamos esta igualdad en la forma $a(x - x_0) = b(y_0 - y)$ que, dividiendo por $d = \mathrm{mcd}(a, b)$, queda

$$(a/d)(x - x_0) = (b/d)(y_0 - y). \tag{1.22}$$

De $d = \mathrm{mcd}(a, b)$ se sigue que $\mathrm{mcd}(a/d, b/d) = 1$; entonces, por el corolario 1.5, forzosamente $(a/d) \mid (y_0 - y)$. Así pues, $y_0 - y = (a/d)n$ para algún $n \in \mathbb{Z}$, y por tanto $y = y_0 - (a/d)n$. Finalmente, sustituyendo este valor en (1.22) llegamos a $x = x_0 + (b/d)n$. □

Este teorema se puede generalizar a más de dos variables como sigue:

Teorema 1.30.
Para $k \geq 2$, sean $a_1, a_2, \ldots, a_k$ números enteros no nulos, $c \in \mathbb{Z}$ y $d = \mathrm{mcd}(a_1, a_2, \ldots, a_k)$. La ecuación diofántica lineal

$$a_1x_1 + a_2x_2 + \cdots + a_kx_k = c \tag{1.23}$$

no tiene soluciones enteras si $d \nmid c$, y tiene infinitas soluciones enteras si $d \mid c$.

Demostración. Como en el teorema 1.29, es claro que, si $d \nmid c$, la ecuación no puede tener soluciones. Para el recíproco, podríamos haber partido directamente de (1.9), con lo cual el resultado lo íbamos a obtener de manera inmediata. Pero vamos a ser un poco más constructivos, y a proporcionar un método que, si lo deseáramos, nos permitiría encontrar explícitamente todas las soluciones de la ecuación (por medio de pasos sucesivos).

Así pues, dando por hecho que $d \mid c$, vamos a probar la existencia de infinitas soluciones por inducción sobre k. El caso $k = 2$ es el teorema 1.29. Ahora, supuesto cierto el resultado para $k - 1$, escribamos (1.23) como

$$a_1x_1 + a_2x_2 + \cdots + a_{k-1}x_{k-1} = c - a_kx_k.$$

y tomemos $d' = \mathrm{mcd}(a_1, a_2, \ldots, a_{k-1})$. Por hipótesis de inducción, la ecuación

$$a_1x_1 + a_2x_2 + \cdots + a_{k-1}x_{k-1} = d'u$$

tiene una cantidad infinita de soluciones $x_1, x_2, \ldots, x_{k-1} \in \mathbb{Z}$ para cada $u \in \mathbb{Z}$. Para que esto resulte útil, deberemos ser capaces de probar la existencia de u y x_k tales que $c - a_kx_k = d'u$, es decir,

$$d'u + a_kx_k = c.$$

Pero, por el lema 1.6, $\mathrm{mcd}(d', a_k) = d$, y sabemos que $d \mid c$. Así pues, el teorema 1.29 nos asegura que existen infinitos u y x_k enteros que son solución de esa ecuación (e incluso nos dice cómo encontrarlos). □

Como consecuencia del teorema, es claro que se cumple lo siguiente:

Corolario 1.31.
Dados $k \geq 2$ y $a_1, a_2, \ldots, a_k \in \mathbb{Z} \setminus \{0\}$, la ecuación diofántica lineal

$$a_1x_1 + a_2x_2 + \cdots + a_kx_k = c$$

tiene solución para cada entero c si y sólo si $\mathrm{mcd}(a_1, a_2, \ldots, a_k) = 1$.

En ocasiones, al resolver una ecuación diofántica interesan únicamente las soluciones no negativas. Supongamos que enviar una carta por correo para un determinado destino cuesta c céntimos, y que disponemos de sellos de valor a y b. Para saber qué sellos de cada tipo usar, deberemos resolver la ecuación $ax + by = c$; pero no nos vale para nada que x e y sean enteros negativos, pues no tiene sentido usar una cantidad negativa de sellos de algún tipo. Así, por ejemplo, si queremos franquear una carta con 3 céntimos y tenemos sellos de valor 5 y 2, pese a que la ecuación $5x + 2y = 3$ tiene la solución $x = 1$, $y = -1$, en el ejemplo de la carta no sirve para nada. En cambio, es claro que cualquier $5x + 2y = c$ con $c \geq 4$ tiene soluciones x, y no negativas.

Pero este requisito adicional no es banal, y complica extraordinariamente lo que se puede decir sobre la existencia o no de soluciones. En el siguiente teorema damos una condición suficiente que garantiza la existencia de soluciones no negativas de una ecuación diofántica lineal.

Teorema 1.32.
Para $k \geq 2$, sean $a_1, a_2, \ldots, a_k$ enteros positivos con $\operatorname{mcd}(a_1, a_2, \ldots, a_k) = 1$, *y $c \in \mathbb{Z}$. Si*

$$c \geq (a_k - 1) \sum_{j=1}^{k-1} a_j \tag{1.24}$$

entonces existen enteros no negativos $x_1, x_2, \ldots, x_k$ tales que $a_1x_1 + a_2x_2 + \cdots + a_kx_k = c$. (Por supuesto, el papel de a_k en (1.24) lo puede asumir cualquier otro de los a_j.)

Demostración. Por el teorema 1.30, existen $z_1, z_2, \ldots, z_k \in \mathbb{Z}$ tales que $a_1z_1 + a_2z_2 + \cdots + a_kz_k = c$. Apliquemos el algoritmo de la división a cada uno de los enteros $z_1, z_2, \ldots, z_{k-1}$ dividido entre a_k. Así, tendremos

$$z_j = a_kq_j + x_j, \quad \text{con } 0 \leq x_j < a_k, \qquad j = 1, 2, \ldots, k-1.$$

Tomemos ahora

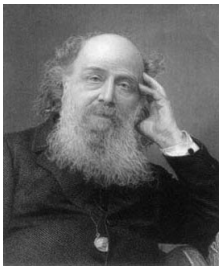
$$x_k = z_k + \sum_{j=1}^{k-1} a_jq_j.$$

Entonces,

$$\begin{aligned} c &= a_1z_1 + \cdots + a_{k-1}z_{k-1} + a_kz_k \\ &= a_1(a_kq_1 + x_1) + \cdots + a_{k-1}(a_kq_{k-1} + x_{k-1}) + a_kz_k \end{aligned}$$



Ferdinand Georg Frobenius (1849–1917)



James Joseph Sylvester (1814–1897)

$$= a_1x_1 + \cdots + a_{k-1}x_{k-1} + a_k\left(z_k + \sum_{j=1}^{k-1} a_jq_j\right)$$
$$= a_1x_1 + \cdots + a_{k-1}x_{k-1} + a_kx_k,$$

luego efectivamente $x_1, \ldots, x_{k-1}, x_k$ satisfacen la ecuación. Comprobemos que son números no negativos. Ya sabemos que $x_j \geq 0$ para $j = 1, \ldots, k-1$. En cuanto a x_k, apliquemos (1.24) y que $x_j \leq a_k - 1$ para $j = 1, \ldots, k-1$. De ese modo,

$$(a_k - 1)\sum_{j=1}^{k-1} a_j \leq c = a_1x_1 + \cdots + a_{k-1}x_{k-1} + a_kx_k \leq (a_k - 1)\sum_{j=1}^{k-1} a_j + a_kx_k,$$

así que $a_kx_k \geq 0$; y, como $a_k \geq 1$, forzosamente $x_k \geq 0$. □

Sigamos suponiendo que tenemos $a_1, a_2, \ldots, a_k$ siendo $\text{mcd}(a_1, a_2, \ldots, a_k) = 1$. Del teorema anterior se deduce que, para c suficientemente grande, siempre será posible encontrar enteros no negativos x_j, $j = 1, 2, \ldots, k$, que hagan $a_1x_1 + a_2x_2 + \cdots + a_kx_k = c$. ¿Cuál es el mayor entero c que no puede representarse de esa manera? Denotemos a este número mediante $g(a_1, a_2, \ldots, a_k)$. El teorema 1.32 nos está dando una cota superior,

$$g(a_1, a_2, \ldots, a_k) < (a_k - 1)\sum_{j=1}^{k-1} a_j \tag{1.25}$$

(recordemos que el papel de a_k en la expresión de la derecha se puede intercambiar por el de cualquiera de los a_j). Determinar $g(a_1, a_2, \ldots, a_k)$ se denomina *problema diofántico lineal de Frobenius* o, simplemente, problema de Frobenius. Y a $g(a_1, a_2, \ldots, a_k)$ se le llama *número de Frobenius*.

Precisemos aquí que, cuando la ecuación $a_1x_1 + a_2x_2 + \cdots + a_kx_k = c$ admite soluciones no negativas para cualquier $c \geq 0$, decimos que $g(a_1, a_2, \ldots, a_k) = -1$. En particular, esto ocurre si alguno de los a_j vale 1 (la ecuación admite la solución trivial dada por $x_j = c$ y $x_i = 0$ para $i \neq j$); en este caso, el valor $g(a_1, a_2, \ldots, a_k) = -1$ se corresponde muy bien con la cota (1.25) antes mostrada (con $a_j = 1$ en el lugar de a_k).

Pero, para valores de c pequeños, a veces no se puede resolver la ecuación; con $k = 2$, recordemos el ejemplo de los sellos de correos antes citado, que podemos expresar diciendo que $g(5, 2) = 3$. Es más común plantear el problema en términos de monedas en lugar de sellos. Así, por ejemplo, si tenemos monedas de valor 2, 5 y 10, con ellas se puede reunir cualquier cantidad de dinero excepto 1 y 3; es decir, $g(2, 5, 10) = 3$. Por este motivo, el problema de Frobenius también se conoce con el nombre de *problema de las monedas*, o del *cambio de moneda*, si se interpreta en términos de cambios entre las monedas de dos países. (Sin embargo, al efectuar pagos con monedas no es necesario exigir $x_j \geq 0$ si permitimos que, para pagar cierta cantidad, haya «vueltas», tal como hacemos habitualmente; esas vueltas equivalen a $x_j < 0$.)

El problema de Frobenius es un problema abierto de considerable dificultad. Existen cotas mejores que la mostrada en (1.25) para estimar el valor del número de Frobenius, pero no se conocen fórmulas precisas para el cálculo de $g(a_1, a_2, \ldots, a_k)$ cuando $k \geq 3$. El problema de Frobenius sólo está resuelto de manera satisfactoria cuando $k = 2$, caso para el que James Joseph Sylvester probó el siguiente resultado:

Teorema 1.33. Sylvester, 1884.
Sean a_1 y a_2 enteros positivos y primos entre sí. En este caso, el número de Frobenius es

$$g(a_1, a_2) = a_1a_2 - a_1 - a_2 = (a_1 - 1)(a_2 - 1) - 1.$$

Demostración. Tal como hemos visto en la demostración del teorema 1.32 (antes de usar la hipótesis (1.24)), para cada $c \in \mathbb{Z}$ existen enteros x_1 y x_2 tales que

$$c = a_1x_1 + a_2x_2, \quad \text{con } 0 \leq x_1 < a_2. \tag{1.26}$$

Veamos que, si exigimos $0 \leq x_1 < a_2$, ésta es la única solución de la ecuación diofántica $c = a_1x_1 + a_2x_2$. Supongamos que existen otros dos enteros x_1' y x_2' que cumplen

$$c = a_1x_1' + a_2x_2', \quad \text{con } 0 \leq x_1' < a_2.$$

En tal caso, tendríamos $a_1(x_1 - x_1') = a_2(x_2' - x_2)$; y, como $\mathrm{mcd}(a_1, a_2) = 1$, deberá cumplirse $a_2 \,|\, (x_1 - x_1')$. Pero $|x_1 - x_1'| < a_2$, luego forzosamente $x_1 = x_1'$. De aquí es inmediato que también $x_2 = x_2'$, y por tanto la representación de c dada en (1.26) es única.

Una consecuencia de esta unicidad es que, si existe otra solución

$$c = a_1 x_1' + a_2 x_2', \quad \text{con } 0 \leq x_1'$$

(sin exigir $x_1' < a_2$), deberá ser $x_1' \geq x_1$ y, por tanto, $x_2' < x_2$ (ya que a_1 y a_2 son positivos). Así pues, dado un $c \in \mathbb{Z}$ y su correspondiente ecuación $c = a_1 x_1 + a_2 x_2$, las consideraciones anteriores muestran que la ecuación tiene soluciones no negativas si y sólo si el x_2 de la representación (1.26) es no negativo.

Una vez preparado el terreno, abordemos el cálculo de $g(a_1, a_2)$. Si $c \in \mathbb{Z}$ es tal que $c = a_1 x_1 + a_2 x_2$ no tiene soluciones no negativas, en la representación (1.26) deberá ser $x_2 \leq -1$, y c deberá cumplir

$$c = a_1 x_1 + a_2 x_2 \leq a_1(a_2 - 1) + a_2(-1) = a_1 a_2 - a_1 - a_2.$$

En consecuencia, $g(a_1, a_2) \leq a_1 a_2 - a_1 - a_2$.

Para concluir que $g(a_1, a_2) = a_1 a_2 - a_1 - a_2$ nos basta tomar $c = a_1 a_2 - a_1 - a_2$ y comprobar que la ecuación $c = a_1 x_1 + a_2 x_2$ no tiene soluciones con x_1, x_2 enteros no negativos. Y, para ello, que la representación según (1.26) no proporciona un $x_2 \geq 0$. Pero esto es inmediato, ya que la solución de

$$a_1 a_2 - a_1 - a_2 = a_1 x_1 + a_2 x_2 \quad \text{con } 0 \leq x_1 < a_2$$

(que es única) es $x_1 = a_2 - 1$, $x_2 = -1$. □

Se puede encontrar mucha más información sobre el problema de Frobenius, y los intentos por resolverlo, en el artículo [80].

1.6.2. Ternas pitagóricas

Como es bien conocido, el teorema de Pitágoras establece que, si x e y son los dos catetos de un triángulo rectángulo, y z es su hipotenusa, los tres lados están relacionados mediante la fórmula $x^2 + y^2 = z^2$. En la figura 1.4 damos una justificación gráfica de este teorema; como demostración sin palabras que es, no merece la pena dedicar unas líneas a explicarlo.

Ya sin que sea necesario aludir a la geometría, se dice que una terna de enteros (x, y, z) es *pitagórica* si se cumple

$$x^2 + y^2 = z^2. \tag{1.27}$$

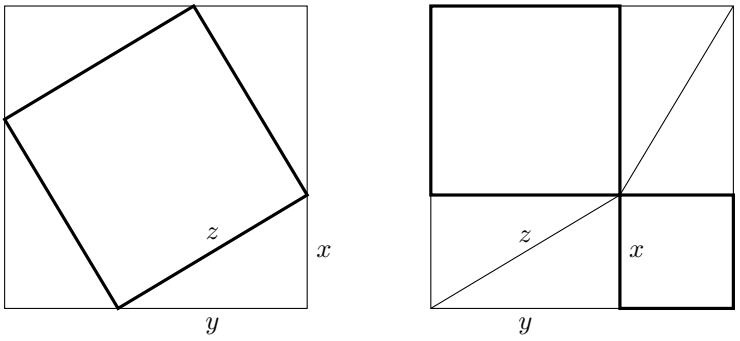


Figura 1.4: Demostración sin palabras del teorema de Pitágoras.

Debido al teorema de Pitágoras, el estudio de esta ecuación diofántica se remonta a la antigüedad, y ya los babilonios se habían ocupado de ella (véase la figura 1.5).

Mediante tanteo es sencillo encontrar soluciones particulares, como $(3, 4, 5)$. Pero nos preguntamos si podemos idear un algoritmo para obtener todas. La respuesta es que sí, y en este apartado nos vamos a ocupar de hacerlo. Pero, antes de abordarlo de manera sistemática, hagamos un breve recorrido histórico.

Parece ser que el propio Pitágoras (siglo VI a.C.) desarrolló un método que permitía encontrar infinitos triángulos rectángulos cuyos lados eran números enteros; en concreto, podía encontrarlos de tal modo que uno de sus catetos tuviera longitud impar prefijada. Dado un entero impar $x > 1$, se puede construir un triángulo rectángulo con catetos x e y e hipotenusa z sin más que tomar

$$y = \frac{x^2 - 1}{2}, \quad z = \frac{x^2 + 1}{2}.$$

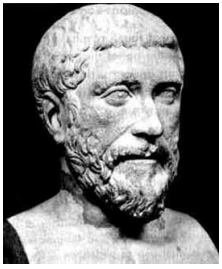
Gracias a la ayuda que ahora nos proporciona la notación algebraica, para nosotros es inmediato comprobar que, efectivamente, y y z son también enteros positivos y se cumple $x^2 + y^2 = z^2$.

Platón (siglo IV a.C.) desarrolló un método que permitía prefijar un cateto de longitud par. Así, si $x > 2$ es un entero par, construimos un triángulo rectángulo con catetos x e y e hipotenusa z mediante

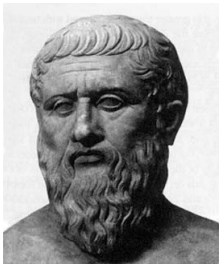
$$y = \left(\frac{x}{2}\right)^2 - 1, \quad z = \left(\frac{x}{2}\right)^2 + 1.$$



Figura 1.5: La tablilla denominada Plimpton 322 es una tablilla de arcilla de origen babilónico que data, aproximadamente, del año 1800 a.C. En escritura cuneiforme y con los números en notación hexadecimal babilónica, la tablilla parece contener una lista de ternas pitagóricas.



Pitágoras de Samos
(≈ 569 a.C. – ≈ 475 a.C.)



Platón
(427 a.C. – 347 a.C.)

Euclides (siglo III a.C.) sabía que, si n y m son enteros positivos de la misma paridad y $m > n$, con ellos se genera la terna pitagórica dada por

$$x = mn, \quad y = \frac{m^2 - n^2}{2}, \quad z = \frac{m^2 + n^2}{2}.$$

La parametrización de Pitágoras se corresponde con el caso $n = 1$ de la de Euclides; la de Platón, con el caso $n = 2$ (pero dividiendo previamente entre 2 la parametrización de Euclides).

Diofanto (siglo III d.C.), utilizando métodos más aritméticos que los de Euclides (cuyas descripciones eran «muy geométricas»), llegó prácticamente a lo mismo que su antecesor. Pero, en realidad, Diofanto no sólo busca soluciones enteras de las ahora denominadas «ecuaciones diofánticas», sino también soluciones racionales; así, por ejemplo, ante la pregunta de «descomponer un cuadrado dado en dos cuadrados», él mismo lo aplica al número $16 = 4^2$ y encuentra

$$16 = \left(\frac{16}{5}\right)^2 + \left(\frac{12}{5}\right)^2.$$

Si nos restringimos a enteros positivos, el método que implícitamente usaba Diofanto para generar ternas pitagóricas $x^2 + y^2 = z^2$ consistía en tomar $m > n$ y, con ellos,

$$x = 2mn, \quad y = m^2 - n^2, \quad z = m^2 + n^2.$$

Posiblemente fue Leonardo de Pisa «Fibonacci» (siglo XIII) el primero que se dio cuenta que de esa forma, y multiplicando por contantes, se podían obtener todas las ternas pitagóricas. Con las precisiones que luego veremos (teorema 1.34), ésta es la parametrización de las ternas pitagóricas que seguimos usando actualmente. Se conoce tanto con el nombre de *parametrización de Diofanto* como con *parametrización de Euclides*.

Por descontado, en los más de tres milenios de historia que tiene el problema, muchos más matemáticos han hecho aportaciones interesantes, y han planteado nuevas cuestiones relacionadas con las ternas pitagóricas. Podríamos eternizamos con ello, así que dejamos ya la historia y vamos con la resolución de la ecuación diofántica.

Empecemos con propiedades y observaciones sencillas sobre la ecuación (1.27). Si (x, y, z) es una terna pitagórica, cualquier combinación de signos $(\pm x, \pm y, \pm z)$ es una terna pitagórica. Y si algún elemento es 0 la ecuación es trivial. Por lo tanto, podemos concentrarnos en buscar soluciones $x, y, z > 0$.

Además, si (x, y, z) es una terna pitagórica, también lo es (kx, ky, kz) para cada $k \in \mathbb{N}$. Y si $\mathrm{mcd}(x, y, z) = d$, entonces $(x/d, y/d, z/d)$ es de nuevo una terna pitagórica. Estos dos hechos demuestran que podemos limitarnos a encontrar ternas pitagóricas (x, y, z) con $\mathrm{mcd}(x, y, z) = 1$, pues conocidas éstas obtenemos todas. Estas ternas se llaman ternas pitagóricas *primitivas*.

En cuanto a las ternas primitivas, es sencillo comprobar que, si se cumple $x^2 + y^2 = z^2$, entonces $\mathrm{mcd}(x, y, z) = 1$ si y sólo si $\mathrm{mcd}(x, y) = 1$ (ejercicio 1.41).

Así pues, podemos decir que las ternas pitagóricas primitivas son las que cumplen $\mathrm{mcd}(x, y) = 1$.

Por último, observemos que si x e y son impares, entonces $x^2 + y^2 = 4k + 2$ para algún k (es una simple comprobación a partir de $x = 2r + 1$, $y = 2s - 1$). Pero ningún cuadrado perfecto puede tener la forma $4k + 2$ (ejercicio 1.9). Por lo tanto, x e y tienen que ser de distinta paridad y, como el orden no importa en nuestro problema, podemos suponer que y es par. Así lo haremos en todo lo que sigue.

Teorema 1.34.

Tres enteros positivos x, y, z forman una terna pitagórica primitiva (es decir, x, y, z son primos entre sí y cumplen $x^2 + y^2 = z^2$) con y par si y sólo si existen enteros positivos m y n primos entre sí, de distinta paridad y con $m > n$, tales que

$$\begin{cases} x = m^2 - n^2, \\ y = 2mn, \\ z = m^2 + n^2. \end{cases} \tag{1.28}$$

Además, esta representación es única.

Demostración. Comencemos viendo que cada terna pitagórica primitiva (x, y, z) con y par se puede expresar como en (1.28). Al ser y par, no lo puede ser x, pues entonces z también sería par y la terna no sería primitiva. Así pues, x y z serán impares. Tomemos $r = (z + x)/2$ y $s = (z - x)/2$, que serán enteros positivos. Escribamos ahora $x^2 + y^2 = z^2$ como $y^2 = z^2 - x^2 = (z + x)(z - x)$; entonces,

$$\left(\frac{y}{2}\right)^2 = \left(\frac{z + x}{2}\right)\left(\frac{z - x}{2}\right) = rs.$$

Así pues, tenemos

$$\begin{cases} x = r - s, \\ z = r + s, \\ y^2 = 4rs. \end{cases} \tag{1.29}$$

Es claro que $r > s > 0$ y son números de distinta paridad (pues x y z son impares). Además, $\mathrm{mcd}(r, s)$ divide a los tres números x, y, z (que cumplen $\mathrm{mcd}(x, y, z) = 1$), luego $\mathrm{mcd}(r, s) = 1$. Precisamente por ser $\mathrm{mcd}(r, s) = 1$, de $y^2 = 4rs$ se deduce que ambos números r y s deben ser cuadrados perfectos (es inmediato sin más que pensar en el teorema fundamental de la aritmética). Tomamos ahora $r = m^2$ y $s = n^2$ y (1.29) se convierte en (1.28).

Analicemos ahora el recíproco. Es una mera formalidad ver que cada trío de números x, y, z definidos como en (1.28) forma una terna pitagórica:

$$\begin{aligned} x^2+y^2 &= (m^2-n^2)^2+(2mn)^2 = m^4-2m^2n^2+n^4+4m^2n^2 \\ &= m^4+2m^2n^2+n^4 = (m^2+n^2)^2 = z^2. \end{aligned}$$

Veamos que tal terna es primitiva. Supongamos que $\mathrm{mcd}(x,y,z)>1$; si ése es el caso, existirá un primo p que divide a $\mathrm{mcd}(x,y,z)$. No puede ser $p=2$, ya que $x=m^2-n^2$ es impar (como n y m tienen distinta paridad, también es distinta la paridad de n^2 y m^2). Además, de $p\,|\,x$ y $p\,|\,z$ se sigue que $p\,|\,(z+x)=2m^2$ y $p\,|\,(z-x)=2n^2$. De aquí deducimos que $p\,|\,m$ y $p\,|\,n$, pero eso es imposible pues $\mathrm{mcd}(m,n)=1$.

Finalmente, comprobemos la unicidad de la representación (1.28). Supongamos que una terna (x,y,z) en las condiciones del teorema se puede expresar como en (1.28) no sólo con m, n sino también con ciertos M, N (primos entre sí, de distinta paridad y con $M>N>0$). Sumando las dos igualdades

$$m^2-n^2=M^2-N^2, \qquad m^2+n^2=M^2+N^2$$

obtenemos $2m^2=2M^2$, luego $m=M$. Y de la igualdad $2mn=2MN$ se deduce que también $n=N$. □

Como consecuencia inmediata del teorema, es claro que para obtener todas las ternas pitagóricas (primitivas y no primitivas) basta hacer

$$\begin{cases} x=k(m^2-n^2), \\ y=2kmn, \\ z=k(m^2+n^2), \end{cases}$$

con $0<n<m$, primos entre sí y de distinta paridad, y $k\in\mathbb{N}$.

Daremos ahora una demostración alternativa del teorema 1.34, mucho más geométrica:

Demostración geométrica. Si (x,y,z) es una terna pitagórica, el punto $(\frac{x}{z},\frac{y}{z})$ del plano $\mathbb{R}^2$ es un punto de la circunferencia unidad con coordenadas racionales. De hecho, encontrar todas las ternas pitagóricas primitivas equivale a encontrar todos esos puntos. Así pues, ¿cómo hallar todos los puntos (a,b) de la circunferencia unidad que tienen coordenadas racionales?

Si $(a,b)\neq(-1,0)$ es uno de tales puntos, la recta que pasa por (a,b) y $(-1,0)$ tiene como ecuación

$$\frac{y}{x+1}=\frac{b}{a+1},$$

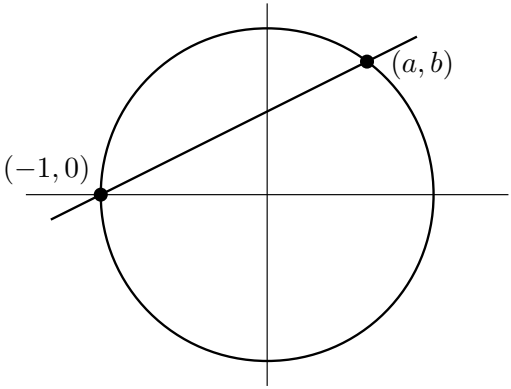


Figura 1.6: Intersección de la recta que pasa por los puntos $(-1,0)$ y (a,b) con la circunferencia unidad.

y su pendiente es el número racional $\lambda = b/(a+1)$; ilustramos esto en la figura 1.6. Recíprocamente, si λ es un número racional, entonces la recta de pendiente λ que pasa por $(-1,0)$ interseca a la circunferencia unidad exactamente en un punto (además del propio $(-1,0)$, claro), y resolviendo el sistema

$$\begin{cases} y = \lambda(x+1), \\ x^2 + y^2 = 1 \end{cases}$$

se comprueba que tal punto tiene coordenadas racionales (a,b) dadas por

$$\begin{cases} a = (1-\lambda^2)/(1+\lambda^2), \\ b = 2\lambda/(1+\lambda^2). \end{cases} \tag{1.30}$$

Así pues, prescindiendo de $(-1,0)$, estamos asociando exactamente un punto de coordenadas racionales (a,b) en la circunferencia unidad para cada pendiente racional λ. Los puntos de coordenadas positivas son los que se corresponden con una recta cuya pendiente cumple $0 < \lambda < 1$. Como $\lambda \in \mathbb{Q}$ y $0 < \lambda < 1$, podremos poner $\lambda = n/m$ con $m, n \in \mathbb{Z}$, primos entre sí y $m > n > 0$. De este modo, (1.30) puede expresarse como

$$\begin{cases} a = (m^2 - n^2)/(m^2 + n^2), \\ b = 2mn/(m^2 + n^2). \end{cases}$$

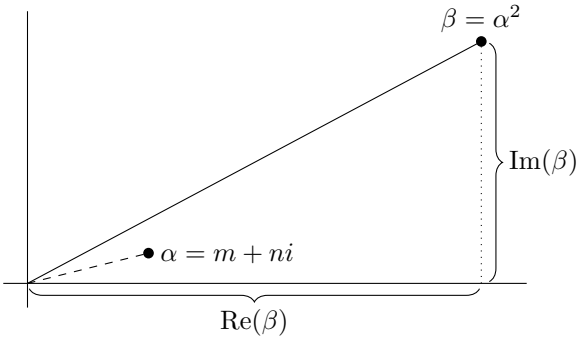


Figura 1.7: El complejo $\beta = \alpha^2$ con su parte real y su parte imaginaria.

Tenemos pues $(a, b) = (x/z, y/z)$, un punto de coordenadas racionales de la circunferencia unidad. Para poder asignar

$$\begin{cases} x = m^2 - n^2, \\ y = 2mn, \\ z = m^2 + n^2 \end{cases}$$

de modo que (x, y, z) sea una terna primitiva hay que imponer que $\mathrm{mcd}(m^2 - n^2, m^2 + n^2) = \mathrm{mcd}(2mn, m^2 + n^2) = 1$. Pero esto ocurre —no daremos ya los detalles— siempre que m y n tienen distinta paridad y $\mathrm{mcd}(m, n) = 1$. □

Nota 1.35. Para concluir, mostremos otra forma de llegar a la parametrización de las ternas pitagóricas primitivas. Esta vez, lo haremos utilizando números complejos.

Sea $\alpha = m + ni$ un número complejo cuyas partes real e imaginaria (m y n, respectivamente) son números enteros positivos; recordemos que el módulo de α es $|\alpha| = \sqrt{m^2 + n^2}$. Tomemos ahora un nuevo número complejo $\beta = \alpha^2$, que resulta ser

$$\beta = (m + ni)^2 = (m^2 - n^2) + 2mni,$$

y cuyo módulo es $|\beta| = |\alpha|^2 = m^2 + n^2$. Impongamos ahora la condición $0 < n < m$ para que $\mathrm{Re}(\beta) = m^2 - n^2$ sea positivo. Sin más que tener en cuenta la geometría del plano complejo, y tal como se muestra en la figura 1.7, el teorema de Pitágoras nos dice que $|\beta|^2 = (\mathrm{Re}(\beta))^2 + (\mathrm{Im}(\beta))^2$. Ahora, tomando

$$x = m^2 - n^2, \quad y = 2mn, \quad z = m^2 + n^2,$$

esa relación se convierte en $x^2 + y^2 = z^2$. ◁

1.6.3. Una pequeña miscelánea de ecuaciones diofánticas

No queremos acabar sin citar unas cuantas ecuaciones diofánticas que han tenido algún tipo de significación o que —simplemente— tienen un aspecto lo suficientemente sencillo como para poderlas comentar en unas líneas. En todo caso, considerémoslas únicamente como meros ejemplos.

La mayoría de las ecuaciones que presentamos consisten en diversos tipos de igualdades entre sumas de potencias, y es fundamentalmente la teoría aditiva de números la que se encarga de estudiar estos problemas. Pero también veremos otras en las que aparece alguna variable como exponente. Tales ecuaciones se denominan ecuaciones diofánticas exponenciales, y sólo se pueden abordar con métodos ad hoc o mediante prueba y error.

En primer lugar, no podemos dejar de citar diversos casos particulares de la ecuación de Fermat

$$x^n + y^n = z^n$$

con $n = 3, 4, 5, 6, \ldots$. Fermat dejó escrito que había encontrado una demostración maravillosa de que esa ecuación no tenía soluciones enteras positivas para ningún $n \geq 3$ —y tal afirmación se viene conociendo históricamente con el nombre de *último teorema de Fermat*—, pero realmente él sólo nos proporcionó la demostración correspondiente al caso $n = 4$.

Así las cosas, probar el *último teorema de Fermat* se reduce a demostrar la no existencia de soluciones de la ecuación $x^p + y^p = z^p$ con p primo. En efecto, supongamos que, para cierto $n > 2$ no primo, existe (x, y, z) solución positiva de $x^n + y^n = z^n$. Si $n = 2^k$ con $k \geq 2$, tendremos que $n = 4 \cdot 2^{k-2}$ y $(x^{2^{k-2}}, y^{2^{k-2}}, z^{2^{k-2}})$ será solución de la ecuación $x^4 + y^4 = z^4$, lo cual es imposible por el resultado de Fermat. Y si n tiene un factor primo impar p, entonces $(x^{n/p}, y^{n/p}, z^{n/p})$ será solución de $x^p + y^p = z^p$.

El primer caso en ser resuelto fue el correspondiente a $p = 3$. Fue Euler quien, alrededor de 1740, logró demostrar que la ecuación $x^3 + y^3 = z^3$ no tenía soluciones enteras no nulas.

Más de medio siglo después, a principios del siglo XIX, la matemática francesa Sophie Germain planteó un método general de abordar el problema para los primos p tales que $2p+1$ también era primo (tales primos se conocen ahora con el nombre de *primos de Germain*). Hasta donde ella llegó, su método no lograba demostrar la no existencia de soluciones para ningún p concreto, pero imponía fuertes restricciones a la existencia de soluciones.

El método de Germain consiguió su primer éxito en 1825, cuando Lejeune Dirichlet y Adrien Marie Legendre probaron —por separado, y ambos basándose en lo desarrollado por Germain— que la ecuación $x^5 + y^5 = z^5$ no tenía soluciones



Marie Sophie Germain
(1776–1831)



Gabriel Lamé
(1795–1870)



Ernst Eduard Kummer
(1810–1893)

enteras no nulas. Pocos años más tarde, en 1839, Gabriel Lamé lograba demostrar que la ecuación $x^7+y^7=z^7$ tampoco tenía soluciones enteras no triviales (esta vez, $p=7$ no es un primo de Germain).

Tras los éxitos que se estaban cosechando, siguieron varios intentos de demostración del caso general, pero ninguno llegaba a buen puerto. La siguiente contribución importante fue la de Ernst Kummer, quien en 1847 se dio cuenta de que el impedimento que frenaba el progreso de sus colegas estaba basado en la no unicidad de la descomposición en factores primos en ciertos anillos. Por ejemplo, así como en $\mathbb{Z}$ el número 6 sólo se descompone como $6=2\cdot 3$, en el anillo $\mathbb{Z}[\sqrt{-5}\,]=\{a+b\sqrt{-5}\,:a,b\in\mathbb{Z}\}$ se tiene

$$6=2\cdot 3=(1+\sqrt{-5})(1-\sqrt{-5}),$$

que son dos descomposiciones esencialmente distintas. Los esfuerzos por esquivar el problema dieron lugar a la teoría de ideales (que nosotros no abordaremos). A partir de entonces se fueron sucediendo pequeños avances y demostraciones erróneas —algunas de matemáticos consagrados—, pero la historia que sigue hasta llegar a la demostración de Wiles en 1995 es ya muy larga y mucho más técnica. Así que, en lo que a nosotros respecta, dejaremos aquí la historia del último teorema de Fermat; si el lector sigue interesado en ella, le recomendamos el excelente libro [96].

Nota 1.36. Resulta interesante citar una conjetura que, si fuera demostrada, de ella se deduciría, como fácil corolario, el último teorema de Fermat. Se trata de la denominada conjetura de Beal, que fue propuesta por el banquero y matemático aficionado Andrew Beal en 1993. Dicha conjetura afirma que si $a^n+b^m=c^k$, siendo a,b,c enteros positivos, y n,m,k enteros mayores que 2, entonces $\mathrm{mcd}(a,b,c)>1$.

La conjetura de Beal es una generalización del último teorema de Fermat, que corresponde al caso $n=m=k>2$. Si $a^n+b^n=c^n$ con $n>2$, entonces, o las

bases son coprimas o comparten un factor común; si comparten un factor común, se puede sacar de cada una de ellas para obtener una ecuación más pequeña, con bases coprimas. Como la conjetura de Beal afirma que el caso $\mathrm{mcd}(a, b, c) = 1$ nunca ocurre, de aquí se deduciría que la ecuación de Fermat $a^n + b^n = c^n$ no puede tener soluciones.

Desde 2003, Beal ofrece un premio de un millón de dólares a quien demuestre la conjetura que lleva su nombre, o a quien encuentre un contraejemplo. ◁

En otro orden de cosas, y también relacionándolo con el último teorema de Fermat, Euler conjeturó en 1769 que, si para algún exponente entero $n \geq 2$ y un número de sumandos $k \geq 2$, se cumplía

$$a_1^n + a_2^n + \cdots + a_k^n = c^n \tag{1.31}$$

con $a_1, a_2, \ldots, a_k$ y c enteros positivos, forzosamente debía ser $k \geq n$. Es decir que, cuando $n \geq 2$, al menos se necesitan n potencias n-ésimas para obtener una suma que es también una potencia n-ésima (recordemos que el propio Euler había probado en 1740 que la ecuación diofántica $x^3 + y^3 = z^3$ no tenía soluciones no nulas).

La conjetura se demostró falsa en 1966, año en el que Leon J. Lander y Thomas R. Parkin realizaron una búsqueda por ordenador que les permitió encontrar el contraejemplo

$$27^5 + 84^5 + 110^5 + 133^5 = 144^5.$$

Una vez comprobado que la conjetura era falsa al menos para exponente $n = 5$, faltaba conocer si era cierta o no para otros valores de n. Desde luego, el caso que más intrigaba era $n = 4$. ¿Podría ser cierto que no existían enteros positivos x, y, z y t tales que $x^4 + y^4 + z^4 = t^4$? La respuesta a esta pregunta se resistió todavía bastante. Fue en 1987 cuando Noam Elkies probó —utilizando ciertas construcciones geométricas denominadas curvas elípticas y ayudado por potentes ordenadores— que la conjetura era falsa también en este caso: encontró la solución

$$2\,682\,440^4 + 15\,365\,639^4 + 18\,796\,760^4 = 20\,615\,673^4$$

e incluso demostró que existía una clase infinita de soluciones. El año siguiente, Roger Frye encontró

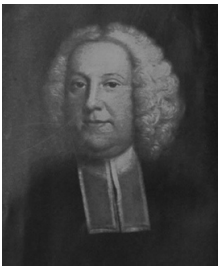
$$95\,800^4 + 217\,519^4 + 414\,560^4 = 422\,481^4$$

y demostró que era la menor solución posible.

En contraposición con el éxito en la búsqueda de soluciones para las potencias cuartas y quintas en la ecuación (1.31), aún no se sabe qué ocurre para potencias sextas o superiores. ¿Estaba Euler totalmente equivocado o había algo cierto debajo de su conjetura?



Noam D. Elkies (1966–)



John Pell (1611–1685)

Pero, como el lector puede imaginar, no todas las ecuaciones diofánticas están relacionadas con el último teorema de Fermat o sus generalizaciones. A lo largo de la historia, han sido muchas otras las que han despertado el interés de los matemáticos, y el mayor o menor éxito al resolverlas varía desde el fracaso absoluto hasta conocer exactamente todas las soluciones.

A modo de despedida, y ya prácticamente sin dar ningún detalle, mencionaremos unos pocos ejemplos más de ecuaciones diofánticas que tienen o han tenido cierto interés:

- Ecuación de Pell, que es como se denomina a la ecuación diofántica

$$x^2 - ny^2 = \pm 1.$$

  Debe su nombre al matemático inglés John Pell (1611–1685), pero ya había sido estudiada por los matemáticos hindúes Brahmagupta (siglo VII) y Bhaskara II (siglo XII).[18]

- Ecuación de Markov

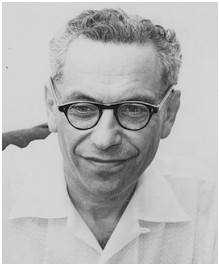
$$u^2 + v^2 + w^2 = 3uvw,$$

  cuya importancia radica en que —tal como descubrió Andrei Markov en 1880— está relacionada con la denominada aproximación por racionales. Se conocen bastantes cosas sobre esta ecuación y sus soluciones, que tienen interesantes propiedades —en este libro, nos volverá a aparecer en el tema 4 (sección 4.6)—, pero aún ofrece problemas abiertos. En concreto, la conjetura de Markov —que permanece sin demostrar— afirma que si (u_1, v_1, w) y (u_2, v_2, w) son dos soluciones de la ecuación de Markov con $u_1 \leq v_1$ y $u_2 \leq v_2$,

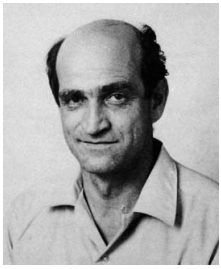
[18] El «II» de Bhaskara II se usa exclusivamente para diferenciarlo de otro matemático hindú con el que comparte nombre, pero que vivió en el siglo VII.



Adolf Hurwitz
(1859–1919)



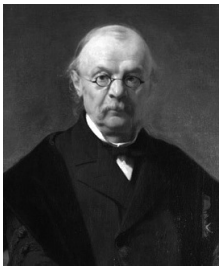
Pál Erdős (1913–1996)



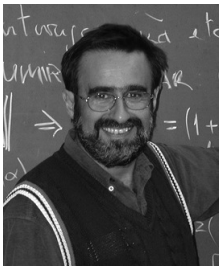
Ernst Gabor Straus
(1922–1983)

forzosamente debe cumplirse $u_1 = v_1$ y $u_2 = v_2$ (véase [98]). Por otra parte, también vale la pena destacar que, en 1907, Adolf Hurwitz demostró que ninguna ecuación de la forma $u^2 + v^2 + w^2 = kuvw$ tiene soluciones enteras positivas salvo que $k = 3$ (la ecuación de Markov) o $k = 1$.

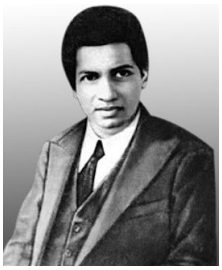
- Ecuación
$$\frac{4}{n} = \frac{1}{x} + \frac{1}{y} + \frac{1}{z},$$
que da pie a la conjetura de Erdős-Straus: para cada $n \geq 2$, existe una solución con x, y y z enteros positivos. Por ejemplo, con $n = 1801$ se tiene la solución
$$\frac{4}{1801} = \frac{1}{451} + \frac{1}{295\,364} + \frac{1}{3\,249\,004}.$$
No podemos dejar aquí de citar que las sumas de fracciones de la forma $1/m$ con $m \in \mathbb{N}$ se denominan *fracciones egipcias*, pues son el tipo de fracciones que se admitían en el Egipto de la época de los faraones. Acabamos de ver que, pese a su aparente simplicidad, las fracciones egipcias aún nos deparan problemas abiertos.
- Conjetura de Catalan: las únicas potencias consecutivas no triviales son 8 y 9, es decir, la única solución con enteros de la ecuación diofántica exponencial
$$x^n - y^m = 1, \qquad x > 0, \quad y > 0, \qquad n > 1, \quad m > 1,$$
es $x = 3$, $n = 2$, $y = 2$, $m = 3$. Esta conjetura —que databa de 1844— fue resuelta por el matemático rumano Preda Mihăilescu en 2002, por lo que ahora también se conoce con el nombre de teorema de Mihăilescu.
- Otro ejemplo típico de ecuación diofántica exponencial es
$$2^n - 7 = x^2,$$



Eugène Charles Catalan (1814–1894)



Preda V. Mihăilescu (1955–)



Srinivasa Ramanujan (1887–1920)



Trygve Nagell (1895–1988)



Leo Moser (1921–1970)

que sólo tiene soluciones naturales n y x cuando $n = 3$, 4, 5, 7 o 15. Esto fue conjeturado en 1913 por el matemático indio Srinivasa Ramanujan, y probado en 1949 por el matemático noruego Trygve Nagell. (En consecuencia, a menudo se alude a esa ecuación con el nombre de ecuación de Ramanujan-Nagell.)

- Problema de Brocard o ecuación de Brocard-Ramanujan,

$$n! + 1 = m^2,$$

cuyo estudio fue propuesto por Henri Brocard en 1876 e, independientemente, por Ramanujan en 1913. Sus únicas soluciones conocidas son

$$4! + 1 = 5^2, \qquad 5! + 1 = 11^2, \qquad 7! + 1 = 71^2.$$

- Ecuación de Erdős-Moser

$$1^k + 2^k + \cdots + (m-1)^k = m^k,$$

cuya única solución conocida es la trivial $1^1+2^1 = 3^1$. Erdős conjeturó que ésa era su única solución, y diversos resultados teóricos y búsquedas con computador (entre ellas la realizada por Leo Moser en 1953) así parecen avalarlo.

Naturalmente, nos dejamos muchas otras ecuaciones diofánticas en el tintero. Simplemente esperamos que el lector se haya hecho una idea de la inmensidad del campo que ahora abandonamos.

1.7. Ejercicios

1.1. Prueba por inducción que $n \leq 2^{n-1}$ para todo $n \in \mathbb{N}$.

1.2. Prueba por inducción la igualdad

$$1 + 2 + \cdots + n = \frac{n(n+1)}{2}, \qquad n \in \mathbb{N}.$$

1.3. Demuestra que la suma de los cubos de los primeros n enteros positivos es igual al cuadrado de la suma de esos primeros n números.

1.4. Demuestra la siguiente identidad entre polinomios:

$$\sum_{k=0}^{2n} x^{2k} = \left(\sum_{k=0}^{2n} x^k\right)\left(\sum_{k=0}^{2n} (-1)^k x^k\right).$$

1.5. Sean a y b dos enteros no nulos. Demuestra que existen $\mu_1, \mu_2 \in \mathbb{Z}$ tales que

$$d = \mu_1 a + \mu_2 b$$

si y sólo si $\mathrm{mcd}(a, b) \,|\, d$.

1.6. Sea n un número compuesto. Prueba que n tiene un divisor primo menor o igual que $\sqrt{n}$.

1.7. Sea un entero $p > 1$ que satisface la siguiente propiedad:

$$p \,|\, ab \;\Rightarrow\; p \,|\, a \;\text{ o }\; p \,|\, b$$

(por supuesto, a y b denotan números enteros). Sin usar el teorema fundamental de la aritmética, prueba que los únicos divisores positivos de p son 1 y p (es decir, que p es número primo).

1.8. Intenta reconstruir el argumento que reduce la prueba de la existencia de infinitos números primos a encontrar alguna sucesión de números enteros positivos primos entre sí dos a dos. Halla alguna sucesión con esa propiedad. *Aclaración.* En la nota 1.9 hemos mencionado que una sucesión que lo cumple son los números de Fermat; demostrar que son coprimos dos a dos requiere cierto trabajo (nosotros lo aplazamos al tema 3, donde estudiaremos los números de Fermat con más detenimiento). En cambio, existe otra sucesión —la denominada sucesión de Sylvester— para la que esa propiedad es inmediata. ¿Se te ocurre cuál es?

1.9. Sea k un entero positivo. Prueba que el número $4k+2$ no puede ser un cuadrado.

1.10. Prueba que, si $n>1$, el entero n^4+4 nunca es primo.

1.11. Demuestra que la suma de los cubos de tres enteros consecutivos cualesquiera es divisible por 9.

1.12. Sea $n\in\mathbb{Z}$. Prueba que n^5-5n^3+4n es divisible por 120.

1.13. Demuestra que, para todo n entero no negativo, la suma $A_n=11^{n+2}+12^{2n+1}$ es divisible por 133.

1.14. Supongamos que tenemos una serie de objetos idénticos dispuestos en forma de cuadrado (igual número de filas que de columnas), y que los movemos para colocarlos formando un rectángulo en el que el número inicial de filas ha aumentado en 5. ¿Cuántos objetos tenemos?

1.15. Halla todos los enteros positivos n tales que 3^n+5^n es múltiplo de $3^{n-1}+5^{n-1}$.

1.16. Halla todas las soluciones enteras de la ecuación $(n+m)p=nm$ con la condición adicional de que p sea primo.

1.17. Encuentra todos los racionales $r>0$ tales que $r+\frac{1}{r}$ es un número natural.

1.18. Prueba que las sumas parciales de la serie armónica, esto es, $\sum_{k=1}^{n}1/k$, no proporcionan un número entero para ningún $n\geq 2$.

1.19. Sean $a,b\in\mathbb{N}$ con $\mathrm{mcd}(a,b)=1$, y tomemos $d=\mathrm{mcd}(a+b,a^2+b^2)$. Demuestra que $d=1$ o 2.

1.20. Sean a, b, n y m enteros positivos. Da una demostración o un contraejemplo de cada una de las afirmaciones siguientes:

(a) Si $a^n\,|\,b^n$, entonces $a\,|\,b$.

(b) Si $n^n\,|\,m^m$, entonces $n\,|\,m$.

(c) Si $n^m \,|\, m^n$, entonces $n \,|\, m$.

(d) Si $a^n \,|\, 2b^n$ y $n > 1$, entonces $a \,|\, b$.

1.21. Sean a y b dos enteros no nulos. Prueba la siguiente relación entre el mínimo común múltiplo y el máximo común divisor:

$$\operatorname{mcm}(a,b) = \frac{|ab|}{\operatorname{mcd}(a,b)}.$$

1.22. Sean a, b y c tres enteros no nulos. Prueba que

$$\operatorname{mcm}(a,b,c) = \frac{|abc| \operatorname{mcd}(a,b,c)}{\operatorname{mcd}(a,b)\operatorname{mcd}(b,c)\operatorname{mcd}(c,a)}.$$

1.23. Sean a y b dos enteros no nulos y $d = \operatorname{mcd}(a,b)$. Prueba que

$$\operatorname{mcm}\left(\frac{a}{d}, \frac{b}{d}\right) = \frac{\operatorname{mcm}(a,b)}{d}.$$

1.24. Sean a y b dos enteros, p un primo y $r \geq 1$. Prueba que, si p^r divide a $\operatorname{mcm}(a,b)$ y a $a+b$, entonces p^r divide a $\operatorname{mcd}(a,b)$.

1.25. Sean a y b dos enteros. Demuestra que

$$\operatorname{mcd}(a,b) = \operatorname{mcd}(a+b, \operatorname{mcm}(a,b)).$$

1.26. Sea p un primo y a, b enteros positivos. Decide razonadamente si las siguientes afirmaciones son verdaderas o falsas:

(a) Si $\operatorname{mcd}(a,p^2) = p$, entonces $\operatorname{mcd}(a^2,p^2) = p^2$.

(b) Si $\operatorname{mcd}(a,p^2) = p$ y $\operatorname{mcd}(b,p^2) = p^2$, entonces $\operatorname{mcd}(ab,p^4) = p^3$.

(c) Si $\operatorname{mcd}(a,p^2) = p$ y $\operatorname{mcd}(b,p^2) = p$, entonces $\operatorname{mcd}(ab,p^4) = p^2$.

(d) Si $\operatorname{mcd}(a,p^2) = p$, entonces $\operatorname{mcd}(a+p,p^2) = p$.

1.27. Prueba que cada entero positivo n se puede escribir de manera única en la forma

$$n = 3^k + b_{k-1}3^{k-1} + \cdots + b_2 3^2 + b_1 3 + b_0$$

con $b_j \in \{0, 1, -1\}$ para $j = 0, 1, 2, \ldots, k-1$.

1.28. Imagina que tienes una balanza y pesas de 1, 3, 9, 27, 81, 243 y 729 gramos (una de cada tipo), y que regentas un puesto de venta de azúcar a granel (en gramos) para cantidades menores que un kilo. Demuestra que puedes pesar cualquier cantidad de azúcar que solicite un cliente, y explica cómo hacerlo.

1.29. Prueba el lema 1.18.

1.30. Sea $m \geq 1$, $k \neq 0$, y q un divisor de $10^m + k$ que, además, satisface $\mathrm{mcd}(q,k) = 1$. Sea asimismo $N = D10^m + d$ con $D, d \in \mathbb{Z}$. Demuestra que $q \,|\, N$ si y sólo si $q \,|\, (kD - d)$. Como aplicación directa, prueba que

$$23 \,|\, (1000D + d) \;\Leftrightarrow\; 23 \,|\, (D - 2d)$$

y úsalo para comprobar (sin hacer la división) que el número 13 069 292 668 es divisible por 23.

1.31. Sea q un divisor de $K10^m + k$ (con $m \geq 1$ y $k \neq 0$) que satisface $\mathrm{mcd}(q,k) = 1$. Dado $N = D10^m + d$ con $D, d \in \mathbb{Z}$, prueba que

$$q \,|\, N \;\Leftrightarrow\; q \,|\, (kD - Kd).$$

1.32. Sea $b \geq 2$ la base de un sistema de numeración. Demuestra que los números $10101_{(b}$, $101010101_{(b}$, $1010101010101_{(b}$, ... (siempre con $4n+1$ «dígitos» en base b) nunca son primos.

1.33. En la plantilla de un equipo de fútbol hay 3 porteros, 8 defensas, 5 centrocampistas y 6 delanteros. Además de con el consabido portero, al entrenador le gusta jugar con 4 defensas, 3 centrocampistas y 3 delanteros. ¿Cuántas alineaciones distintas puede hacer?

1.34. De todos es sabido que algunas personas sólo compran libros para adornar la estantería del salón de casa, y los eligen por el color de sus lomos. Si una de tales personas dispone de 8 libros rojos, 4 amarillos, 6 verdes y 3 azules, ¿de cuántas maneras puede decorar su estantería?

1.35. Comprueba rigurosamente que los elementos de las filas del triángulo de Pascal crecen desde los extremos hacia el centro.

1.36. Comprueba que

$$(k+1) \cdot \binom{n+1}{k+1} = (n+1) \cdot \binom{n}{k}, \qquad 0 \leq k \leq n.$$

1.37. Prueba la identidad

$$\sum_{k=0}^{n} \binom{n}{k}^2 = \binom{2n}{n}.$$

1.38. Demuestra que cada entero $n \geq 1$ satisface la identidad

$$\sum_{k=0}^{n}(k+1)\binom{n}{k} = 2^{n-1}(n+2).$$

1.39. Demuestra que, para cada entero $n \geq 2$, se verifica

$$\sum_{k=0}^{n}k(k-1)\binom{n}{k} = 2^{n-2}n(n-1).$$

1.40. Halla todas las soluciones enteras de la ecuación diofántica

$$15x + 21y + 35z = 13.$$

1.41. Sean x, y, z tres enteros positivos que verifican $x^2 + y^2 = z^2$. Prueba que

$$\text{mcd}(x, y, z) = 1 \;\Leftrightarrow\; \text{mcd}(x, y) = 1.$$

1.42. Sea un triángulo rectángulo cuyos lados son números enteros. Demuestra las siguientes propiedades:

(a) El producto de las longitudes de los dos catetos es divisible por 12.

(b) El producto de las longitudes de los tres lados es divisible por 60.

1.43. Prueba que, en una terna pitagórica, un elemento es divisible por 3, otro por 4, y otro por 5.
Aclaración. Puede ocurrir que sea el mismo elemento el que tenga dos de esos divisores, como en las ternas $(8, 15, 17)$, $(7, 24, 25)$ y $(20, 21, 29)$; o incluso los tres, como en $(11, 60, 61)$.

1.44. Demuestra que la única solución de la ecuación diofántica

$$2^n + 1 = m^2,$$

con n y m enteros positivos, es $n = m = 3$.

1.45. Encuentra todas las soluciones de la ecuación diofántica

$$n^x + n^y = n^z,$$

donde x, y, z denotan números enteros y n un entero positivo.

Tema 2
Congruencias

2.1. Definiciones y primeras propiedades
2.2. La ecuación con congruencias lineal
2.3. La indicatriz de Euler y el teorema de Euler-Fermat
2.4. Números pseudoaleatorios
2.5. Primos en progresiones aritméticas
2.6. Sistemas lineales de congruencias: el teorema chino del resto
2.7. Teorema de Wilson
2.8. Números pseudoprimos
2.9. Criptografía
2.10. Teoría de códigos
2.11. Ejercicios



Imagen de Carl Friedrich Gauss (1777–1855) en un billete alemán

El lenguaje de las congruencias que introducimos en este tema es extremadamente útil en teoría de números. Fue desarrollado por el genial Carl Friedrich Gauss, sin duda uno de los mejores matemáticos de la historia. Muchos de los resultados que aquí exponemos aparecen en su célebre libro *Disquisitiones Arithmeticae* (a menudo se dice de él que es «el primer libro de teoría de números»), publicado en 1801. Antes que Gauss, otros dos insignes matemáticos, Leonhard Euler y Joseph Louis Lagrange, habían hecho considerables aportaciones, aunque ellos no las expresaron en términos de congruencias. Pero aún podemos remontarnos mucho más atrás pues —posiblemente a partir del siglo II a.C.— los matemáticos chinos e indios se servían de algunas de las técnicas que relatamos con el objetivo de elaborar calendarios y predecir acontecimientos astronómicos.

Cuando hablamos en términos de congruencias, siempre consideramos fijo un entero m mayor que 1, que es el *módulo* de las congruencias. En ese contexto, lo único importante de un número es su *resto* al dividir por m (en ocasiones se usa *residuo* como sinónimo de resto). Así, dos números enteros que dan el mismo resto se consideran indistinguibles. La aritmética modular se conoce también con el llamativo nombre de *aritmética del reloj*. La motivación de este nombre es clara, pues el comportamiento de las horas responde al de las congruencias módulo 12.

Pese a su relativa sencillez, la aritmética modular es una herramienta poderosa, y su uso proporciona una gran ayuda en muy diferentes tipos de problemas. Algunas veces, permite resolverlos completamente. Otras, se trata de problemas muy complicados, cuya solución general no se conoce o requiere técnicas sofisticadas; sin embargo, aplicando aritmética modular se consigue resolver de manera sencilla algunos casos particulares. Por ejemplo, algunos métodos de factorización y tests para comprobar la posible primalidad de un número están basados en técnicas de aritmética modular, y también interviene con éxito al analizar muy diversos tipos de problemas diofánticos (ecuaciones en las que se quiere hallar sus soluciones enteras).

Además, las congruencias tienen aplicaciones en campos que, en principio, parecen más lejanos. No podemos dejar de citar su gran importancia en el desarrollo de las telecomunicaciones, de la mano de la criptografía y la teoría de códigos. Muchas técnicas criptográficas se basan en el uso adecuado de las congruencias; sin duda, el paradigma de esto es el criptosistema conocido —por las iniciales de los nombres de sus descubridores— como RSA (nosotros lo analizaremos en la sección 2.9). Asimismo, su concurso es fundamental para construir códigos detectores y correctores de errores (tal como veremos en la sección 2.10).

Coincidiendo con la llegada de los ordenadores, y gracias a las aplicaciones que acabamos de describir, muchos temas relacionados con la teoría de números y la matemática discreta han experimentado un gran auge. Materias que se consideraban de escaso interés práctico, de repente se volvieron muy útiles. De nuevo tenemos que poner como ejemplo el desarrollo de la criptografía y la teoría de códigos, cuyo crecimiento en el último tercio del siglo XX ha sido espectacular. Como gran parte de la teoría subyacente estaba ya estudiada, los avances fueron —tras los primeros balbuceos— rápidos y con paso firme. Es más, profundizando en los estudios teóricos, surgieron nuevas ideas de aplicabilidad inmediata. Nadie podía esperar esto sólo unos años antes, en los que la disciplina se consideraba puramente teórica.

Pero los matemáticos que se enfrentaron a lo que aquí exponemos —como ya hemos mencionado, a Euler, Lagrange y Gauss debemos los principales logros— lo hicieron por puro placer intelectual. Y ése mismo va a ser nuestro objetivo durante la mayor parte de este tema. Sólo en algunas secciones nos ocuparemos de exponer varias de las importantes aplicaciones ya citadas, y que surgieron dos siglos más tarde.

No es éste el único tema del presente libro dedicado a las congruencias, sino que también el siguiente está dedicado a ellas; sin esta división en dos, hubiera resultado un tema demasiado largo. Aun así, hemos tenido que hacer una selección de lo que incluíamos, necesariamente basada en nuestros gustos personales.

Prácticamente todos los textos de teoría de números tienen uno o varios capítulos dedicados al estudio de las congruencias; por ejemplo, [3, 5, 26, 41, 50, 55, 71, 78, 87, 88, 95]. En algunos libros se hace hincapié en los aspectos computacionales; entre los que hemos citado, es obligatorio destacar [88]. Aún más específicos son [30] y —en español y de carácter eminentemente divulgativo— [44]. Por supuesto, el empleo de la informática tiene una gran importancia práctica, pero no ha sido éste nuestro objetivo, y todos los ejercicios que proponemos pueden realizarse a mano o con la ayuda de una pequeña calculadora. Para entender lo que ocurre —y eso es lo que nosotros perseguimos— no hace falta un ordenador. En realidad, es sólo una cuestión de tamaño: con la misma facilidad con que manejamos números de 3 o 4 cifras operando con lápiz y papel, con la ayuda de un ordenador podemos utilizarlos

con varias decenas de dígitos. En este sentido, las aplicaciones relacionadas con la seguridad criptográfica y las comunicaciones pueden requerir varios cientos de cifras.

En cuanto a las numerosas aplicaciones de las congruencias, sólo en algunas secciones nos hemos dedicado a ellas; es más, incluso disponiendo de todos los ingredientes matemáticos necesarios para haber tratado estas aplicaciones con más profundidad, no hemos querido hacerlo. Nuestra pretensión ha sido mostrar la utilidad de las congruencias mediante una visión general de algunas de sus aplicaciones más destacadas, y no dar un tratamiento exhaustivo de ellas, que corresponde a otros libros más especializados. Pero hemos dejado sin mencionar muchas más aplicaciones de las que citamos, tanto relacionadas con la «matemática seria» como con cuestiones más mundanas como la elaboración de calendarios, matemáticas recreativas de todo tipo, trucos de magia...

Con respecto a la criptografía, recomendamos [65, 60, 99] (también [88] le dedica importantes recursos); asimismo, un texto divulgativo muy ameno es [97]. La teoría de códigos es —quizás— aún más amplia, y nuestro tratamiento ha sido aún más superficial; si el lector quiere ampliar sus conocimientos en esto, puede emplear los libros [53, 68, 85, 101]. Por último, tampoco queremos prescindir de recomendar [27]: de carácter eminentemente divulgativo, tiene dos excelentes capítulos dedicados a la codificación de la información y el resto —aunque dedicado a otros temas— no desmerece en absoluto.

2.1. Definiciones y primeras propiedades

Fijado un número natural $m \geq 1$, se dice que dos números $a, b \in \mathbb{Z}$ son *congruentes* módulo m si $a - b \in m\mathbb{Z}$, donde $m\mathbb{Z}$ es el conjunto de todos los enteros múltiplos de m, es decir,

$$m\mathbb{Z} = \{\, mk : k \in \mathbb{Z} \,\}.$$

Es fundamental darse cuenta que $a - b \in m\mathbb{Z}$ si y sólo si a y b dan el mismo resto al dividir por m, caracterización que a menudo resulta más efectiva. En todo caso, esto se describe con la notación $a \equiv b \pmod{m}$. Cuando $m = 1$ todo es trivial, pues todos los números son congruentes módulo 1. Así que, en general, supondremos que $m \geq 2$ (quizás a veces es incluso necesario para evitar enunciados inconsistentes, aunque no lo digamos expresamente; en tal caso, encárguese el lector de asumirlo).

Dado m, las congruencias módulo m constituyen una relación de equivalencia, pues se cumplen las propiedades

- Reflexiva: $a \equiv a \pmod{m}$.
- Simétrica: $a \equiv b \pmod{m} \;\Leftrightarrow\; b \equiv a \pmod{m}$.
- Transitiva: $a \equiv b \pmod{m}$, $b \equiv c \pmod{m} \Rightarrow a \equiv c \pmod{m}$.

Su demostración es inmediata, así que no merece la pena detenerse en ellas. Por si el lector se siente inseguro, comprobemos la transitividad: de $a-b=km$ y $b-c=k'm$ se sigue que $a-c=(a-b)+(b-c)=(k+k')m$.

Además, esta relación es compatible con la suma y el producto:

- $a\equiv b \pmod m,\ c\equiv d \pmod m \ \Rightarrow\ a+c\equiv b+d \pmod m$;
- $a\equiv b \pmod m,\ c\equiv d \pmod m \ \Rightarrow\ ac\equiv bd \pmod m$.

Podríamos seguir dando muchas otras propiedades cuya demostración tiene carácter elemental; a este respecto, véase, por ejemplo, el ejercicio 2.1.

Dado $a\in\mathbb{Z}$, a su clase de equivalencia (es decir, al conjunto de todos los enteros congruentes con a módulo m)

$$\widehat{a}=\{\,x\in\mathbb{Z}: x\equiv a \pmod m\,\}=a+m\mathbb{Z}$$

la llamamos *clase de restos* de a módulo m. Y usamos $\mathbb{Z}_m$ para denotar el conjunto de todas las clases de equivalencia módulo m. Está claro que las clases de restos $\widehat{1},\widehat{2},\dots,\widehat{m}$ son disjuntas y su unión es $\mathbb{Z}$ (lo mismo ocurre con $\widehat{0},\widehat{1},\dots,\widehat{m-1}$, que a veces resulta más cómodo de manejar). Por lo tanto, el *cardinal* de $\mathbb{Z}_m$ es m. Se denomina *representante* a un elemento de esa clase. Está claro que cada clase de $\mathbb{Z}_m$ tiene un representante en $[1,m]$ (y también en $[0,m-1]$).

Si elegimos un representante de cada clase de equivalencia, decimos que tenemos un *sistema completo de restos* módulo m. Hay infinitos sistemas de restos completos; lo más habitual es tomar $\{1,2,\dots,m\}$, o bien $\{0,1,\dots,m-1\}$.

En $\mathbb{Z}_m$ hay estructura aditiva y multiplicativa. Tanto la suma como el producto son meras herencias de las de $\mathbb{Z}$. Interesa mucho más la estructura multiplicativa, así que detengámonos en ella (por supuesto, lo mismo habría que hacer en el caso de la suma). De manera completamente natural, el producto en $\mathbb{Z}_m$ se define como

$$\widehat{a}\cdot\widehat{b}=\widehat{a\cdot b},\qquad \widehat{a},\widehat{b}\in\mathbb{Z}_m.$$

Esta definición tiene sentido (no depende de los representantes elegidos), pues

$$a\equiv a_1 \pmod m,\quad b\equiv b_1 \pmod m \ \Rightarrow\ ab\equiv a_1b_1 \pmod m;$$

es decir: $\widehat{a}=\widehat{a_1}$, $\widehat{b}=\widehat{b_1}\ \Rightarrow\ \widehat{a\cdot b}=\widehat{a_1\cdot b_1}$. Respecto a la suma, el elemento neutro es $\widehat{0}$ (o $\widehat{m}$, que es la misma clase); y, respecto al producto, $\widehat{1}$.

En $\mathbb{Z}_m$ siempre hay inverso respecto a la suma, pues $\widehat{a}+\widehat{(-a)}=\widehat{0}$; eso no es sorprendente, sino que es lo mismo que ocurre en $\mathbb{Z}$. Con respecto al producto, la situación es nueva: en $\mathbb{Z}$, sólo 1 y -1 tienen inverso multiplicativo (ambos son su propio inverso); en cambio, en $\mathbb{Z}_m$ existen otros elementos con inverso respecto al

producto. Como muestra, un botón: en $\mathbb{Z}_5$, la clase $\widehat{1}$ es su propia inversa, y lo mismo ocurre con la clase $\widehat{-1} = \widehat{4}$; pero, y ésta es la novedad, $2 \cdot 3 \equiv 1 \pmod 5$, así que $\widehat{2}$ y $\widehat{3}$ son inversos en $\mathbb{Z}_5$. Por otra parte, no siempre todos los elementos de $\mathbb{Z}_m \setminus \{\widehat{0}\}$ tienen inverso multiplicativo. Por ejemplo, la clase $\widehat{2}$ de $\mathbb{Z}_4$ no tiene inversa (ni tampoco la clase $\widehat{2}$ de $\mathbb{Z}_6$ o $\mathbb{Z}_8$, aunque sí la de $\mathbb{Z}_5$, como hemos visto antes, o también la de $\mathbb{Z}_3$, $\mathbb{Z}_7$ o $\mathbb{Z}_9$).

El siguiente resultado es clave para analizar lo que ocurre:

Proposición 2.1.

Sea un entero $a \neq 0$, y otro entero $m \geq 2$ tal que $\operatorname{mcd}(a, m) = 1$. *Entonces, la ecuación $ax \equiv b \pmod m$ tiene una única solución*[1] *módulo m.*

Demostración. Los m números $a, 2a, 3a, \ldots, ma$ son incongruentes módulo m. En efecto, si $ja \equiv ka \pmod m$, entonces $m \mid (j-k)a$. Pero como $\operatorname{mcd}(a, m) = 1$, se deduce que $m \mid (j-k)$, lo que no puede ser salvo que $j = k$. De este modo, al variar x entre 1 y m, mediante ax se van obteniendo los m restos módulo m que puede haber; en particular, para algún x aparecerá $ax \equiv b \pmod m$. Finalmente, y en cuanto a la unicidad de la solución se refiere, el motivo es similar: si dos x dieran el mismo resto, ya no obtendríamos los m distintos. □

Quedémonos con el caso $b = 1$. Como consecuencia del resultado anterior, los elementos $\widehat{a}$ con $\operatorname{mcd}(a, m) = 1$ tienen inverso (por supuesto, estamos ya prescindiendo de los a múltiplos de m o, lo que es lo mismo, de $\widehat{a} = \widehat{0}$). Esto tiene indudable interés, así que asignemos una notación para el conjunto de tales elementos:

$$\mathbb{Z}_m^* = \{\widehat{a} = a + m\mathbb{Z} : \operatorname{mcd}(a, m) = 1\}. \tag{2.1}$$

Dentro de $\mathbb{Z}_m^*$, el producto $\widehat{a} \cdot \widehat{b} = \widehat{ab}$ continúa estando bien definido: si $\widehat{a}, \widehat{b} \in \mathbb{Z}_m^*$, entonces $\widehat{ab} \in \mathbb{Z}_m^*$, ya que también $\operatorname{mcd}(ab, m) = 1$. Y el elemento neutro (la identidad) es la clase $\widehat{1}$. Resumiendo, hemos encontrado un subconjunto $\mathbb{Z}_m^*$ en el que siempre hay inverso multiplicativo.

Cuando $m = p$, un número primo, cualquier $a \in \{1, 2, \ldots, p-1\}$ verifica $\operatorname{mcd}(a, p) = 1$. Entonces, siempre existe solución única de $ax \equiv b \pmod p$. En particular (tomando $b = 1$), todos los elementos de $\mathbb{Z}_p \setminus \{\widehat{0}\}$ tienen inverso multiplicativo; y, por consiguiente, $\mathbb{Z}_p^* = \mathbb{Z}_p \setminus \{\widehat{0}\}$. Al único elemento de $\{1, 2, \ldots, p-1\}$ que resuelve la ecuación $ax \equiv 1 \pmod m$ se le denota mediante a^{-1}.

[1] De momento, sólo aseguramos su existencia, y no damos ningún procedimiento para encontrar tal solución (salvo, en todo caso, probar todos los $x \in \{1, \ldots, m\}$ para ver cuál sirve); más adelante veremos que se puede hacer mediante el algoritmo de Euclides. Otro método de cálculo nos aparecerá en el teorema 2.8.

Nota 2.2. Demos ahora una descripción alternativa de lo que hemos visto, pero empleando, esta vez, un lenguaje más algebraico. Comencemos aclarando que, además de con $\mathbb{Z}_m$, al conjunto de las clases de restos módulo $m \geq 2$ también se le suele denotar con $\mathbb{Z}/m\mathbb{Z}$; esto responde al cociente de $\mathbb{Z}$ entre $m\mathbb{Z}$, un concepto general que se puede definir siempre que dispongamos de una relación de equivalencia, y que resulta habitual entre estructuras algebraicas. Ahondando en ello, podemos decir que la compatibilidad de $\equiv$ con $+$ y $\cdot$ hace que $(\mathbb{Z}_m, +, \cdot)$ tenga estructura de anillo (conmutativo), heredada de la de $\mathbb{Z}$. Este anillo puede no ser un cuerpo, pues no siempre todos los elementos tienen inverso respecto al producto; en otras palabras, el conjunto $\mathbb{Z}_m \setminus \{\widehat{0}\}$ no siempre forma un grupo con la multiplicación. Esto no es nuevo: en $\mathbb{Z}$, sólo 1 y -1 tienen inverso multiplicativo. Lo que quizás para algunos resultará más sorprendente es que en $\mathbb{Z}_m$ puede ocurrir que $\widehat{a} \cdot \widehat{b} = \widehat{0}$ aunque $\widehat{a}$ y $\widehat{b}$ sean no nulos (lo que se expresa diciendo que existen *divisores de cero*); por ejemplo, $\widehat{2} \cdot \widehat{3} = \widehat{0}$ en $\mathbb{Z}_6$. Para evitar esto, lo que hacemos es quedarnos con $\mathbb{Z}_m^*$, que sí que es grupo con la multiplicación.

Estas situaciones patológicas no ocurren cuando el módulo es un número primo p. Continuando con el lenguaje habitual del álgebra, estamos ahora diciendo que $(\mathbb{Z}_p, +, \cdot)$ no sólo es un anillo, sino que tiene estructura de cuerpo. En particular, obsérvese que en $\mathbb{Z}_p$ (como en cualquier otro cuerpo) no puede haber divisores de cero; esto es, se cumple

$$\widehat{a}, \widehat{b} \in \mathbb{Z}_p \setminus \{\widehat{0}\} \;\Rightarrow\; \widehat{a}\widehat{b} \neq \widehat{0}.$$

La comprobación de este hecho es inmediata por reducción al absurdo: si $\widehat{a}\widehat{b} = \widehat{0}$, multiplicando por el inverso de uno de los factores obtenemos que el otro debe ser nulo.

Los $\mathbb{Z}_p$ son los ejemplos más sencillos de cuerpos finitos, y los únicos que nos van a aparecer aquí. Pero no está de más mencionar que el estudio de los cuerpos finitos en general no sólo da lugar a una elegante teoría matemática, sino que de él surgen importantes aplicaciones de gran interés, como la construcción de sistemas criptográficos, códigos correctores de errores, algoritmos de factorización o test de primalidad. En particular, sobre los cuerpos finitos se pueden definir las denominadas curvas elípticas, que resultan muy útiles en todo esto. Sin embargo, un análisis más profundo de estos temas queda fuera de los objetivos del presente libro. ◁

2.2. La ecuación con congruencias lineal

De una ecuación con congruencias de la forma $ax \equiv b \pmod{m}$ se dice que es lineal. La proposición 2.1 nos mostraba que, cuando $\mathrm{mcd}(a, m) = 1$, esa ecuación

tiene una única solución módulo m. Pero, ¿cómo encontrarla? Resulta muy interesante dar un algoritmo detallado para hacerlo, puesto que tales ecuaciones aparecen a menudo (más aún: querremos saber si, para llevarlo a la práctica, tal algoritmo es o no muy costoso).

Vamos a resolver estas ecuaciones de manera un poco más general a como teníamos en la proposición 2.1. En concreto, dados $a \neq 0$ y $m \geq 2$, estudiaremos la ecuación

$$ax \equiv b \pmod{m}, \qquad \text{con} \quad \mcd(a,m) = d \quad \text{y} \quad d \mid b. \tag{2.2}$$

Es importante caer en la cuenta de que basta buscar el algoritmo para resolver

$$ax \equiv d \pmod{m}, \qquad \text{con} \quad \mcd(a,m) = d, \tag{2.3}$$

ya que, si s es solución de (2.3) (o sea, $m \mid (as - d)$) y $b/d = b_1$ entonces sb_1 es solución de (2.2) (pues $m \mid (as - d)b_1 = asb_1 - b$). Además, podemos suponer que a verifica $0 < a < m$. Si no fuera así, sustituiríamos a por su único número equivalente módulo m que lo cumple. A todos los efectos, esa ecuación con congruencias sería la misma.

Una vez aclarado el contexto, veamos cómo resolver (2.3), con $0 < a < m$. Para ello vamos a usar el algoritmo de Euclides (el mismo que vimos en la sección 1.1 de este libro), que no por elemental resulta menos útil. En el algoritmo de Euclides, lo que se persigue es calcular $d = \mcd(m,a)$. Recordemos que, si $m = aq + r$, cualquier número que divida a m y a también divide a a y r (y viceversa); por lo tanto, $\mcd(m,a) = \mcd(a,r)$.

Aplicando esto reiteradamente (y mientras los restos que se van obteniendo son no nulos), tenemos

$$\begin{aligned}
m &= aq_1 + r_1, & 0 \leq r_1 < a, \qquad & d = \mcd(m,a) = \mcd(a,r_1), && (2.4)\\
a &= r_1q_2 + r_2, & 0 \leq r_2 < r_1, \qquad & d = \mcd(r_1,r_2), && (2.5)\\
r_1 &= r_2q_3 + r_3, & 0 \leq r_3 < r_2, \qquad & d = \mcd(r_2,r_3),\\
r_2 &= r_3q_4 + r_4, & 0 \leq r_4 < r_3, \qquad & d = \mcd(r_3,r_4),
\end{aligned}$$

y así sucesivamente. Como los restos r_j son no negativos y cada vez menores, llegará un momento en que el resto será nulo. Es decir, se tendrá

$$\begin{aligned}
r_{n-2} &= r_{n-1}q_n + r_n, & 0 \leq r_n < r_{n-1}, \qquad & d = \mcd(r_{n-1},r_n), && (2.6)\\
r_{n-1} &= r_nq_{n+1} + r_{n+1}, & 0 \leq r_{n+1} < r_n, \qquad & d = \mcd(r_n,r_{n+1}), && (2.7)\\
r_n &= r_{n+1}q_{n+2} + 0, & r_{n+2} = 0 < r_{n+1}, &
\end{aligned}$$

Como r_{n+1} divide a r_n de manera exacta, es obvio que $\mcd(r_n,r_{n+1}) = r_{n+1}$. Por tanto, $d = r_{n+1}$, con lo que concluye el algoritmo de Euclides.

Ya sólo queda recorrer el proceso anterior en sentido inverso —del final hacia el principio, despejando siempre el correspondiente resto— e ir haciendo sustituciones sucesivas. Al ser $d = r_{n+1}$, la expresión (2.7) la podemos poner como

$$x_1 r_n + y_1 r_{n-1} = d, \quad \text{con } x_1 = -q_{n+1},\ y_1 = 1. \tag{2.8}$$

Ahora, escribamos (2.6) en la forma $r_n = r_{n-2} - q_n r_{n-1}$; sustituyendo en (2.8) obtenemos

$$x_2 r_{n-1} + y_2 r_{n-2} = d, \quad \text{con } x_2 = y_1 - x_1 q_n,\ y_2 = x_1. \tag{2.9}$$

Análogamente, la expresión anterior a (2.6) sería $r_{n-1} = r_{n-3} - q_{n-1} r_{n-2}$; sustituyendo en (2.9) quedaría

$$x_3 r_{n-2} + y_3 r_{n-3} = d, \quad \text{con } x_3 = y_2 - x_2 q_{n-1},\ y_3 = x_2.$$

Siguiendo con el retroceso, vamos obteniendo una serie de valores x_j e y_j tales que

$$x_j r_{n-j+1} + y_j r_{n-j} = d, \quad \text{con } x_j = y_{j-1} - x_{j-1} q_{n-j+2},\ y_j = x_{j-1};$$

y así hasta $j = n - 1$, que quedaría

$$x_{n-1} r_2 + y_{n-1} r_1 = d, \quad \text{con } x_{n-1} = y_{n-2} - x_{n-2} q_3,\ y_{n-1} = x_{n-2}.$$

Un paso más: usando (2.5) en la forma $r_2 = a - r_1 q_2$ y sustituyendo, obtenemos

$$x_n r_1 + y_n a = d, \quad \text{con } x_n = y_{n-1} - x_{n-1} q_2,\ y_n = x_{n-1}.$$

Por fin, poniendo (2.4) como $r_1 = m - a q_1$ y sustituyendo, llegamos a

$$x_{n+1} a + y_{n+1} m = d, \quad \text{con } x_{n+1} = y_n - x_n q_1,\ y_{n+1} = x_n. \tag{2.10}$$

Así pues, $x_{n+1} a - d$ es múltiplo de m, luego x_{n+1} es solución de $ax \equiv d \pmod{m}$.

Nota 2.3. En realidad, el proceso que acabamos de efectuar es prácticamente idéntico al que usábamos en el tema anterior para llegar a la demostración de la identidad de Bézout: si a y m son dos números enteros no nulos, y $d = \text{mcd}(m, a)$, seguro que existen otros dos enteros x e y tales que

$$xa + ym = d.$$

Es más, el procedimiento desarrollado entre (2.4) y (2.10) nos muestra cómo encontrar x e y, pues son, respectivamente, el x_{n+1} e y_{n+1} de (2.10). (La única diferencia entre la identidad de Bézout y la ecuación con congruencias lineal es que, si lo que pretendemos es resolver la ecuación, sólo buscamos x, y el valor de y no tiene importancia.) ◁

Ejemplo. Resolver $811x \equiv 1 \pmod{28\,585}$.

Solución. Apliquemos el algoritmo de Euclides a 28 585 y 811. Así, tenemos,

$$\begin{aligned} 28\,585 &= 811 \cdot 35 + 200, \\ 811 &= 200 \cdot 4 + 11, \\ 200 &= 11 \cdot 18 + 2, \\ 11 &= 2 \cdot 5 + 1, \\ 2 &= 1 \cdot 2. \end{aligned}$$

Entonces, podemos poner

$$\begin{aligned} 1 &= 11 - 2 \cdot 5 = 11 - (200 - 11 \cdot 18) \cdot 5 = 11(1 + 18 \cdot 5) - 5 \cdot 200 \\ &= 11 \cdot 91 - 5 \cdot 200 = (811 - 200 \cdot 4) \cdot 91 - 5 \cdot 200 \\ &= -200(4 \cdot 91 + 5) + 91 \cdot 811 = -200 \cdot 369 + 91 \cdot 811 \\ &= -(28\,585 - 811 \cdot 35) \cdot 369 + 91 \cdot 811 = 811(35 \cdot 369 + 91) - 369 \cdot 28\,585 \\ &= 811 \cdot 13\,006 - 369 \cdot 28\,585. \end{aligned}$$

De este modo hemos llegado a

$$811 \cdot 13\,006 = 1 + 369 \cdot 28\,585,$$

luego $x = 13\,006$ es solución de la ecuación $811x \equiv 1 \pmod{28\,585}$. Y, evidentemente, también será solución cualquier $x = 13\,006 + 28\,585k$ con $k \in \mathbb{Z}$. ◊

2.2.1. ¿Es muy costoso este algoritmo?

En algunas aplicaciones (por ejemplo, en cuestiones criptográficas, tal como veremos más adelante) resulta importante poder resolver rápidamente ecuaciones en congruencias lineales con m y a números muy grandes; digamos que de unos pocos cientos de cifras. ¿Es esto factible en la práctica? ¿Cuántos pasos del algoritmo de Euclides se necesitan para ello?

Aunque no lo vamos a probar de manera rigurosa, es fácil convencerse de que el método explicado funciona muy bien, en el sentido de que es muy rápido: el número de pasos que se necesitan hasta que concluye el algoritmo es siempre menor que una constante multiplicativa (fija) por el número de dígitos de m. Casi sin esfuerzo adicional, podríamos incluso dar un valor concreto —y bien ajustado— para esta cota; pero para nuestros objetivos no necesitamos tanto. Nos vamos a contentar con hacernos una idea de cuál va a ser, como máximo, tal número de pasos.

Cuanto más pequeños sean los cocientes que van apareciendo, más pasos habrá que dar en el algoritmo de Euclides. El caso extremo es cuando todos los cocientes valen 1. Para analizar esto, vamos a emplear la sucesión de Fibonacci, que se define de manera recursiva mediante[2]

$$f_1 = f_2 = 1, \qquad f_j = f_{j-1} + f_{j-2}, \quad j \geq 3.$$

Si f_{n+1} y f_{n+2} son dos términos sucesivos de la sucesión de Fibonacci (con $n > 1$), la relación de recurrencia nos muestra que lo que va ocurriendo al aplicar el algoritmo de Euclides es

$$\begin{aligned} f_{n+2} &= f_{n+1} \cdot 1 + f_n, \\ f_{n+1} &= f_n \cdot 1 + f_{n-1}, \\ &\dots \\ f_4 &= f_3 \cdot 1 + f_2, \\ f_3 &= f_2 \cdot 2. \end{aligned}$$

Salvo en el último paso, el cociente siempre es 1; y hemos dado exactamente n pasos hasta que se obtiene un cociente con resto 0.

Si el m que aparece en (2.3) es menor que f_{n+2}, el número de pasos al efectuar el algoritmo de Euclides con m y a (con $a < m$) es menor que el requerido con f_{n+2} y f_n, pues éste era —recordémoslo— el caso más lento (y en él se daban n pasos).

Los números de Fibonacci crecen muy rápido: de la relación de recurrencia se sigue que $f_{n+2} \geq 2f_n$, y con esto es fácil deducir que $f_{n+2} \geq 2^{n/2}$ (con algo más de esfuerzo, se pueden conseguir cotas más ajustadas, pero para nuestra argumentación no las necesitamos). En consecuencia, si $m < 2^{n/2}$, bastan n pasos para aplicar el algoritmo de Euclides con m y a. Esto implica que el número de pasos es siempre menor o igual que $2\log_2(m)$. El logaritmo de un número es, salvo constante multiplicativa, como el número de dígitos de ese número (esto es cierto independientemente de en qué base tomemos el logaritmo: sólo cambia la constante; y seguro que el lector recuerda que $\lfloor \log_{10}(m) \rfloor$ da el número de dígitos de m menos 1). Así pues, el número de pasos requeridos es, salvo constante multiplicativa, menor que el número de dígitos de m.

En teoría de números computacional —y sin entrar en mayores precisiones— un algoritmo se suele considerar suficientemente rápido cuando el número de pasos que emplea es, como mucho, una función polinómica del número de dígitos de la entrada. Obviamente, más rápido cuanto más pequeño es el grado de ese polinomio.

[2] En la sección 5.1 estudiamos con más detalle esta conocida sucesión, cuyo origen está en un problema de conejos que aparece en el *Liber Abaci* de Fibonacci (siglo XIII).



Imagen de Leonhard Euler (1707–1783) en un billete suizo

En nuestro caso, el algoritmo de Euclides nos ha proporcionado un polinomio de grado uno, lo cual es una situación muy satisfactoria. Como contrapartida, conviene reseñar que no se conoce ningún algoritmo para factorizar enteros que pueda ejecutarse en un tiempo polinomial (aunque tampoco está descartado que tal algoritmo pueda existir y alguien lo halle en el futuro).

2.3. La indicatriz de Euler y el teorema de Euler-Fermat

Según hemos demostrado (véase la proposición 2.1 y la definición (2.1)), $\mathbb{Z}_m^*$ está formado por los elementos de $\mathbb{Z}_m$ con inverso multiplicativo (en $\mathbb{Z}_m$, claro). Tiene un gran interés saber cuántos elementos con inverso multiplicativo hay; es decir, nos estamos preguntando por el *cardinal* de $\mathbb{Z}_m^*$, que es la cantidad de enteros positivos menores o iguales que m y primos con m. Vamos a destacar adecuadamente este número mediante la siguiente notación:

$$\varphi(m) = \operatorname{card}\{k \in \mathbb{Z} : 1 \leq k \leq m,\ \operatorname{mcd}(k,m) = 1\}.$$

En realidad, ésta es una función muy habitual en teoría de números, y se conoce con el nombre de *función indicatriz de Euler*; aparece, a menudo, en otras partes del libro (en concreto, en el tema 7 se estudia con mucho más detalle).

Estábamos diciendo que $\mathbb{Z}_m^*$ tiene $\varphi(m)$ elementos. Elijamos un representante en cada una de estas clases de equivalencia; a esto lo llamamos *sistema reducido de restos* módulo m. El sistema reducido más habitual consiste en tomar, de entre $\{1, 2, \ldots, m\}$, los relativamente primos con m.

Observemos también que si $\{b_1, \ldots, b_{\varphi(m)}\}$ es un sistema reducido de restos módulo m y $\operatorname{mcd}(a, m) = 1$, entonces $\{ab_1, \ldots, ab_{\varphi(m)}\}$ también es un sistema

reducido de restos módulo m. En efecto, basta tener en cuenta que $ab_1, \ldots, ab_{\varphi(m)}$ son incongruentes entre sí (ya que $\mcd(a, m) = 1$ y $b_j \not\equiv b_k \pmod{m}$ si $j \neq k$), que $\mcd(ab_j, m) = 1$ para cada j, y que son $\varphi(m)$ elementos.

Esta observación va a ser todo lo que usemos para probar el teorema que presentamos a continuación (la afirmación general se atribuye a Euler y el caso particular correspondiente a $m = p$, un número primo, a Fermat). No sólo es muy conocido, sino que, si hubiese un resultado con el nombre de «teorema fundamental de la aritmética modular» debería ser, sin duda, éste.

Teorema 2.4. Teorema de Euler-Fermat o congruencia de Euler-Fermat.
Si $\mcd(a, m) = 1$, *entonces* $a^{\varphi(m)} \equiv 1 \pmod{m}$.

Demostración. Tal como acabamos de comentar, si $\{b_1, b_2, \ldots, b_{\varphi(m)}\}$ es un sistema reducido de restos módulo m, entonces $\{ab_1, ab_2, \ldots, ab_{\varphi(m)}\}$ también lo es. Por tanto,

$$b_1 b_2 \cdots b_{\varphi(m)} \equiv a^{\varphi(m)} b_1 b_2 \cdots b_{\varphi(m)} \pmod{m}. \tag{2.11}$$

Como $\mcd(b_j, m) = 1$ para cada j, también $\mcd(b_1 b_2 \cdots b_{\varphi(m)}, m) = 1$. Esto implica que $1 \equiv a^{\varphi(m)} \pmod{m}$: en efecto, al ser $b_1 b_2 \cdots b_{\varphi(m)}(1 - a^{\varphi(m)})$ múltiplo de m y $\mcd(b_1 b_2 \cdots b_{\varphi(m)}, m) = 1$, el factor $1 - a^{\varphi(m)}$ debe ser forzosamente múltiplo de m. (Un argumento alternativo consiste en multiplicar ambos lados de (2.11) por $b_1^{-1} b_2^{-1} \cdots b_{\varphi(m)}^{-1}$, siendo b_j^{-1} el inverso de b_j módulo m; es la proposición 2.1 la que nos asegura la existencia de tales inversos.) □

Nota 2.5. En términos de la teoría de grupos, el teorema de Euler-Fermat asegura que todo elemento del grupo $\mathbb{Z}_m^*$ tiene *orden* finito menor o igual que el cardinal del grupo (el orden de un elemento es cuántas veces hay que multiplicarlo por sí mismo hasta que se obtiene el elemento neutro del grupo). En realidad, aún se verifica algo más: el orden de un elemento siempre divide al cardinal del grupo; véase el ejercicio 2.8.

Esta misma propiedad —incluso enunciada de manera algo más general— se puede demostrar para cualquier grupo finito. En teoría de grupos, al cardinal de un grupo se le suele denominar *orden* del grupo. Con esta nomenclatura, lo que sucede es que el orden de un subgrupo siempre divide al orden del grupo; tal resultado se conoce con el nombre de teorema de Lagrange. Cuando se alude al orden de un elemento, este orden coincide con el del subgrupo que se obtiene con las sucesivas potencias de ese elemento (hasta que se llega al elemento neutro del grupo). ◁

Particularicemos el teorema de Euler-Fermat para algunos casos especialmente útiles. Esto requiere, simplemente, que nos fijemos en cómo se calcula la función $\varphi(m)$ para algunos valores de m (más adelante —teorema 2.10— veremos cómo evaluar $\varphi(m)$ en general).

En primer lugar es evidente que, ante un número primo p, todos los números menores que p son primos con p; así pues, $\varphi(p) = p - 1$. Entonces, tenemos esto:

Corolario 2.6. Teorema pequeño[3] de Fermat.
Sea p un número primo y a un entero. Se cumple:

(a) Si $p \nmid a$, entonces $a^{p-1} \equiv 1 \pmod{p}$.

(b) En general, sea cierta o no la hipótesis anterior $p \nmid a$, siempre se cumple $a^p \equiv a \pmod{p}$.

Demostración. Si $p \nmid a$, el teorema de Euler-Fermat asegura que $a^{p-1} \equiv 1 \pmod{p}$, luego ya hemos probado el primer apartado. Además, en ese caso, $a^p = a^{p-1}a \equiv a \pmod{p}$. Por el contrario, si $p \mid a$, entonces tanto a^p como a son múltiplos de p, luego $a^p - a$ también lo es. □

En realidad, es fácil dar una demostración alternativa del teorema pequeño de Fermat que no se apoye en el de Euler-Fermat: véase el ejercicio 2.7.

Otro interesante caso particular del teorema de Euler-Fermat es cuando se toma $m = pq$ con p y q primos distintos. Para poder continuar, debemos evaluar $\varphi(pq)$. De los $pq - 1$ enteros positivos menores que pq, únicamente no son primos con pq los múltiplos de p (esto es, p, $2p$, $3p$, …, $(q-1)p$), y los múltiplos de q (esto es, q, $2q$, $3q$, …, $(p-1)q$). En estas circunstancias, $\varphi(pq) = pq - 1 - (q-1) - (p-1)$, es decir, $\varphi(pq) = (p-1)(q-1)$. Así que ya tenemos probado el siguiente resultado:

Corolario 2.7. Teorema de la media de Fermat.
Si p y q son primos distintos que no dividen a a, entonces $a^{(p-1)(q-1)} \equiv 1 \pmod{pq}$.

2.3.1. Una expresión de la solución de una ecuación con congruencias lineal

En la sección 2.2 hemos visto que el algoritmo de Euclides permite encontrar la solución de una ecuación con congruencias lineal $ax \equiv b \pmod{m}$, con $\mathrm{mcd}(a, m) = 1$, de una manera muy eficiente. Ese algoritmo requiere efectuar una serie de pasos

[3] El calificativo de «pequeño» aplicado a este resultado se usa para distinguirlo del —mucho más famoso— último teorema de Fermat, que afirma que no existen soluciones enteras positivas de $x^n + y^n = z^n$ si $n \geq 3$, y que fue demostrado por Wiles en 1995.

sencillos, pero no proporciona una fórmula directa; no basta con echar un solo vistazo para poder decir, convencidos, «ésa es la solución». Además, disponer de una fórmula resulta útil en muchas ocasiones: imaginemos, por ejemplo, que esa fórmula se sustituye en otra, y que, tras manipularlas convenientemente, las expresiones que aparecen se simplifican muchísimo; es difícil hacer lo mismo con un algoritmo.

Ahora, gracias al teorema de Euler-Fermat, también podemos dar una expresión explícita de la solución de la ecuación con congruencias lineal; esta expresión sólo requiere saber evaluar la indicatriz de Euler. Conviene recalcar que aquí vamos a proporcionar una bonita fórmula, pero que, como procedimiento operacional, el método ahora mostrado es más lento que el que surgía a partir del algoritmo de Euclides.

Teorema 2.8.
Si $\mathrm{mcd}(a,m)=1$*, la solución (única módulo* m*) de la ecuación* $ax\equiv b \pmod m$ *es*

$$x = ba^{\varphi(m)-1}.$$

Demostración. Como $a^{\varphi(m)}\equiv 1 \pmod m$, sin más que tomar $x=ba^{\varphi(m)-1}$ se tiene $ax=ba^{\varphi(m)}\equiv b \pmod m$, tal como buscábamos (es obvio, además, que cualquier otro $x\equiv ba^{\varphi(m)-1} \pmod m$ también es solución). La unicidad de la solución ya la conocemos (proposición 2.1) pero, por completitud, demostrémosla ahora de nuevo: si x e y son soluciones de la ecuación, se cumplirá $m|(x-y)a$ y por tanto, al ser $\mathrm{mcd}(a,m)=1$, forzosamente $m|(x-y)$, es decir, $x\equiv y \pmod m$. □

Ejemplo. Aplicando el teorema, hallar la solución general de $2x\equiv 9 \pmod{15}$.

Solución. Como $\mathrm{mcd}(2,15)=1$ y $\varphi(15)=\varphi(3\cdot 5)=2\cdot 4=8$, el teorema nos dice que una solución (que además es única módulo 15) es

$$x = 9\cdot 2^{\varphi(15)-1} = 9\cdot 2^7 = 1152.$$

Para cualquier uso posterior, resultaría más cómodo emplear la solución que sea no negativa y menor que 15. Ésta se obtiene sin más que calcular el resto de 1152 módulo 15, que es 12. Además, así podemos expresar la solución general como $x=12+15k$, $k\in\mathbb{Z}$. ◊

Ejemplo. Resolver $5x\equiv 24 \pmod{77}$ mediante la fórmula dada en el teorema.

Solución. Es claro que $\mathrm{mcd}(5,77)=1$, luego la solución es única módulo 77. Además, por ser $\varphi(77)=\varphi(7\cdot 11)=6\cdot 10=60$, el teorema nos asegura que esta solución es $x=24\cdot 5^{59}$. Naturalmente, no hace falta que calculemos ese valor, que

es enorme, sino que nos basta conocerlo módulo 77; para ello, lo más cómodo es ir operando directamente módulo 77, evitando así que aparezcan números tan grandes. Además, el mejor método de evaluar 5^{59} no es multiplicar 5 por sí mismo 59 veces, sino que conviene usar un truco denominado genéricamente *cálculo de potencias mediante cuadrados sucesivos* o *algoritmo de cuadrados iterados*. Apliquemos ambas ideas; así, para obtener $5^{59} \pmod{77}$, descompongamos $59 = 2^5 + 2^4 + 2^3 + 2^1 + 1$ (obsérvese que esto se corresponde con su representación en base 2: $59 = 111011_{(2)}$). Entonces, $5^{59} = 5^{2^5} \cdot 5^{2^4} \cdot 5^{2^3} \cdot 5^2 \cdot 5$. Como $5^{2^k} = \left(5^{2^{k-1}}\right)^2$, tenemos

$$\begin{aligned} 5^2 &\equiv 25 \pmod{77}, \\ 5^{2^2} &\equiv 25 \cdot 25 \equiv 9 \pmod{77}, \\ 5^{2^3} &\equiv 9 \cdot 9 \equiv 4 \pmod{77}, \\ 5^{2^4} &\equiv 4 \cdot 4 \equiv 16 \pmod{77}, \\ 5^{2^5} &\equiv 16 \cdot 16 \equiv 25 \pmod{77}. \end{aligned}$$

De aquí que

$$5^{59} \equiv 25 \cdot 16 \cdot 4 \cdot 25 \cdot 5 \equiv 31 \pmod{77},$$

y por tanto $24 \cdot 5^{59} \equiv 24 \cdot 31 \equiv 51 \pmod{77}$. Así, finalmente, la solución de la ecuación es $x = 51 + 77k$, $k \in \mathbb{Z}$. ◊

Queremos insistir en que utilizar $x = ba^{\varphi(m)-1}$ para resolver $ax \equiv b \pmod{m}$ (con $\mathrm{mcd}(a, m) = 1$, por supuesto) no es una buena idea en la práctica. En el ejemplo de la página 96 resolvíamos $811x \equiv 1 \pmod{28\,585}$ a partir del algoritmo de Euclides para calcular el máximo común divisor de 28 585 y 811. Aunque delicado de aplicar, el método era computacionalmente muy sencillo, y no requería evaluar la función φ. Sin embargo —y dando por hecho que ya conocemos $\varphi(28\,585) = 22\,864$—, imagine el lector que queremos usar directamente la fórmula

$$x = 811^{\varphi(28\,585)-1} = 811^{22\,863},$$

bien sea con el método «completamente directo» del primer ejemplo de este apartado, como con el método «simplificado» del segundo ejemplo. Volviendo al cálculo de $\varphi(28\,585)$, en este caso no era demasiado complicado hacerlo, pues la descomposición de 28 585 en factores primos es $28\,585 = 5 \cdot 5717$ (sólo dos factores) y vimos que, entonces, $\varphi(28\,585) = (5-1)(5717-1) = 22\,864$. En general, todavía no sabemos cómo evaluar $\varphi(m)$ (salvo aplicando directamente la definición), pero adelantamos que la manera más efectiva requiere descomponer m en factores primos, y eso podría ser impracticable si m es grande.

2.3.2. Una fórmula explícita para evaluar la indicatriz de Euler

La indicatriz de Euler aparece a menudo —por ejemplo, la acabamos de ver en el teorema 2.8— así que merece la pena buscar un método que permita evaluarla de una manera suficientemente cómoda. Hasta ahora hemos visto cómo calcular $\varphi(m)$ cuando m es un primo o un producto de dos primos distintos. Nos interesaría encontrar una fórmula que proporcionara su valor para enteros positivos arbitrarios. En concreto, una expresión que permitiera calcular $\varphi(m)$ sin más que conocer la descomposición en factores de m. Antes de seguir, conviene hacer notar que —por si la definición de $\varphi(m)$ no está clara cuando $m = 1$— se toma $\varphi(1) = 1$ (por definición, si así preferimos pensarlo).

Existen diversos procedimientos para llegar a tal expresión genérica.[4] Aquí vamos a seguir un método bastante directo. Con la ayuda de cierta propiedad conjuntista conocida con el nombre de principio de inclusión-exclusión (o también, a veces, principio de clasificación cruzada), calcularemos el cardinal del conjunto mediante el que se define $\varphi(m)$.

Si tenemos dos conjuntos finitos A y B, es evidente que

$$\operatorname{card}(A \cup B) = \operatorname{card}(A) + \operatorname{card}(B) - \operatorname{card}(A \cap B),$$

ya que los elementos de $A \cap B$ están contados dos veces en $\operatorname{card}(A) + \operatorname{card}(B)$. Con más de dos conjuntos, la correspondiente expresión se complica, y su demostración ya no es tan obvia (nosotros vamos a dar una en la que los números combinatorios y el binomio de Newton juegan un papel esencial), pero la idea subyacente es la misma:

Lema 2.9. Principio de inclusión-exclusión.
Si $A_1, A_2, \ldots, A_k$ son conjuntos finitos, entonces

$$\operatorname{card}\left(\bigcup_{r=1}^{k} A_r\right) = \sum_{1\le r\le k} \operatorname{card}(A_r) - \sum_{1\le r<s\le k} \operatorname{card}(A_r \cap A_s) + \sum_{1\le r<s<t\le k} \operatorname{card}(A_r \cap A_s \cap A_t) - \cdots + (-1)^{k+1} \operatorname{card}(A_1 \cap A_2 \cap \cdots \cap A_k).$$

[4] En el tema 7 estudiaremos con mucho más detalle la indicatriz de Euler, y veremos métodos alternativos de llegar a la misma fórmula. En concreto, véase la deducción de (7.14), en la página 419; o la que se consigue a partir del teorema 7.15. Ambas demostraciones son menos laboriosas que la que ahora vamos a presentar, pero requieren una maquinaria teórica que aún no hemos visto.

DEMOSTRACIÓN. Basta ver que, en la parte derecha de la igualdad que queremos probar, cada $x \in \bigcup_{r=1}^{k} A_r$ aparece contado una sola vez. Supongamos que ese x pertenece exactamente a n de tales conjuntos y no a más (obviamente, n cumple $1 \leq n \leq k$). Entonces:

- En $\sum_{1\leq r\leq k} \operatorname{card}(A_r)$, x estará contado n veces.
- En el sumatorio correspondiente a $\sum_{1\leq r<s\leq k}$ aparecerá $\binom{n}{2}$ veces.
- En el sumatorio con 3 subíndices (r, s y t), $\binom{n}{3}$ veces.
- Y así sucesivamente hasta el sumatorio con n subíndices, en el que estará $\binom{n}{n} = 1$ vez.
- A partir de entonces, ningún otro sumatorio con más subíndices lo cuenta.

Entre todos los sumatorios, nuestro x que pertenece a n conjuntos ha sido contado estas veces (obsérvese el uso del binomio de Newton):

$$\begin{aligned} n - \binom{n}{2} + \binom{n}{3} - \cdots + (-1)^{n+1}\binom{n}{n} + 0 \\ = 1 - \left(1 - n + \binom{n}{2} - \binom{n}{3} + \cdots + (-1)^n \binom{n}{n}\right) = 1 - (1-1)^n = 1, \end{aligned}$$

tal como queríamos comprobar. □

Aplicando el lema, conseguimos esto:

Teorema 2.10.
La indicatriz de Euler cumple $\varphi(1) = 1$ y, para un entero $m > 1$ cuya descomposición canónica en factores primos distintos es $m = p_1^{a_1} p_2^{a_2} \cdots p_k^{a_k}$ (con $a_r \geq 1$ para cada r), se tiene

$$\varphi(m) = m \prod_{r=1}^{k} \left(1 - \frac{1}{p_r}\right).$$

DEMOSTRACIÓN. Cuando $m = 1$ no hay nada que probar. Así pues, sea $m = p_1^{a_1} p_2^{a_2} \cdots p_k^{a_k} > 1$. Denotemos A y A_r a los conjuntos

$$\begin{aligned} A &= \{j \in \mathbb{Z} : 1 \leq j \leq m\}, \\ A_r &= \{j \in A : p_r \,|\, j\}, \quad r = 1, 2, \ldots, k. \end{aligned}$$

Por definición de $\varphi(m)$, es claro que

$$\varphi(m) = \operatorname{card}\left(A \setminus \bigcup_{r=1}^{k} A_r\right) = \operatorname{card}(A) - \operatorname{card}\left(\bigcup_{r=1}^{k} A_r\right).$$

Notemos ahora que $\operatorname{card}(A) = m$, $\operatorname{card}(A_r) = m/p_r$ (comenzando por 1 y hasta llegar a m, exactamente un número de cada p_r es divisible por p_r), $\operatorname{card}(A_r \cap A_s) = m/(p_r p_s)$, etcétera. Entonces, sin más que aplicar el principio de inclusión-exclusión llegamos a

$$\begin{aligned}\varphi(m) &= m - \sum_{1\le r\le k} \frac{m}{p_r} + \sum_{1\le r<s\le k} \frac{m}{p_r p_s} \\ &\quad - \sum_{1\le r<s<t\le k} \frac{m}{p_r p_s p_t} + \cdots + (-1)^k \frac{m}{p_1 p_2 \cdots p_k} \\ &= m\left(1 - \frac{1}{p_1}\right)\left(1 - \frac{1}{p_2}\right)\cdots\left(1 - \frac{1}{p_k}\right).\end{aligned}$$

□

2.3.3. Primos de Wieferich

El teorema pequeño de Fermat prueba que, cuando p es un primo y a es un entero no divisible por p, el *cociente de Fermat*

$$q_p(a) = \frac{a^{p-1} - 1}{p}$$

es un entero. Particularizando en el caso $a = 2$, $q_p(2)$ es un entero para cualquier p primo impar. En 1909, Wieferich encontró una relación entre los números $q_p(2)$ divisibles por p y el último teorema de Fermat. En concreto, demostró que si p es un primo impar que no divide a xyz (el producto de tres enteros), se cumplía lo siguiente:

$$x^p + y^p = z^p \;\Rightarrow\; 2^{p-1} \equiv 1 \pmod{p^2}.$$

Con otras palabras, los primos impares p tales que $2^{p-1} - 1$ no sólo es divisible por p, sino también por p^2, eran los únicos exponentes para los que la ecuación diofántica $x^p + y^p = z^p$ podía tener soluciones enteras tales que $p \nmid xyz$. Así, tales primos, que pasaron a denominarse primos de Wieferich, adquirieron cierta relevancia.

Si se lograra identificar todos los primos de Wieferich, o caracterizarlos de alguna manera, eso podía constituir un importante paso en el camino hacia la demostración

del último teorema de Fermat. Prácticamente nada interesante se pudo probar sobre ellos. Pero algunos razonamientos heurísticos[5] hacen pensar que, dado un primo p, éste va a ser un primo de Wieferich con probabilidad $1/p$.

Eso implicaba que los primos de Wieferich iban a ser bastante numerosos. Pero los únicos conocidos eran $p = 1093$ y 3511, y no se lograban encontrar tantos como cabía esperar, lo cual añadía interés al asunto. La ayuda de potentes ordenadores, y mucho tiempo de cálculo, ha permitido intensificar esta búsqueda. Así, se ha logrado comprobar que no existe ningún otro menor que $4 \cdot 10^{14}$. Pero se desconoce si existen o no infinitos primos de Wieferich. Si el lector quiere conocer más detalles de esta historia, en [29] puede encontrar bastante información y referencias adicionales.

2.4. Números pseudoaleatorios

En ocasiones, hay que escoger entre varias posibilidades, y hacerlo de tal manera que todas tengan la misma probabilidad de resultar elegidas. Si hay 2 posibles elecciones, podemos tirar una moneda al aire. Si hay más posibilidades —digamos que m— podemos escribir los números enteros de 1 a m en papelitos, doblarlos, introducirlos en una bolsa, revolverlos, y escoger uno al azar; o usar un bombo con bolas... Pero hay que ser muy cuidadoso para garantizar que, realmente, todos los números tienen las mismas posibilidades de salir elegidos. Además, estos procedimientos son largos si queremos efectuar muchos sorteos, o si m es grande. Para ayudar a resolver estos problemas, durante las primeras décadas del siglo XX se publicaron diversas tablas de *números aleatorios* (previamente obtenidos con cuidado); y más adelante incluso se construyeron «máquinas de aleatorización» basadas en diversos procedimientos mecánicos. Estos generadores de números aleatorios resultaban muy útiles para llevar a cabo simulaciones de fenómenos naturales, o para obtener muestras de datos.

Actualmente, y ya que podemos contar con la ayuda de ordenadores, que efectúan cálculos con gran rapidez, es mucho más cómodo utilizar algún procedimiento numérico que genere números aleatorios. De hecho, los números aleatorios en sí mismos son imprescindibles en diversos algoritmos de cálculo numérico —a los que se suele aludir como métodos Monte Carlo—. Incluso en teoría de números puede resultar importante disponer de buenos procedimientos para escoger números de manera

[5] Con esto se quiere indicar que la argumentación usada para llegar a la conclusión parece bastante sensata, aunque no tiene rigor matemático. Por ejemplo, se han podido usar hipótesis o conjeturas no probadas, o se han descartado casos sin suficiente justificación. Así pues, un razonamiento heurístico no es una demostración, sino más bien un método de conjeturar.

aleatoria; por ejemplo, se utilizan en el criptosistema Elgamal y en el esquema de firma digital del mismo nombre que veremos en el tema siguiente (sección 3.6).

A decir verdad, a menudo los números aleatorios que se necesitan tienen que ser números reales entre 0 y 1. Esto es sólo una pequeña variación sobre el ejemplo de los enteros entre 1 y m: bastaría tomar un m suficientemente grande y dividir por él. Así que no nos preocuparemos de ello, y supondremos que queremos obtener, de manera aleatoria, enteros entre 1 y m.

Pero los ordenadores son deterministas, así que es imposible que con ellos elijamos números de forma verdaderamente aleatoria. Lo que sí podemos hacer es buscar métodos de forma que —en la práctica— la elección se comporte como si realmente fuese aleatoria; si uno quiere recalcar esto, a tales números se les denomina *números pseudoaleatorios* (así lo haremos aquí aunque, habitualmente, el prefijo «pseudo» se suele suprimir). De todas maneras, si nos paramos a pensarlo, lo mismo ocurre al lanzar una moneda: las ecuaciones de la física nos podrían decir qué lado va a caer hacia arriba (al menos, si no nos adentramos en los terrenos de la incertidumbre fruto de la mecánica cuántica y sólo usamos física newtoniana), pero las circunstancias —las condiciones iniciales de las ecuaciones— son tantas, y sus variaciones pequeñas influyen tanto, que un sorteo efectuado con una moneda lo podemos considerar completamente legal.

Existen numerosos métodos de obtener números de manera pseudoaleatoria. Muchos de ellos emplean congruencias. Pero nuestra única intención en esta sección es dar una muestra de una de las aplicaciones de las congruencias, y no hacer un estudio exhaustivo. Así que sólo vamos a ver unos pocos ejemplos y resultados, y no incluiremos demostraciones. El lector que quiera informarse más profundamente de esto puede consultar, por ejemplo, [59, capítulo 3].

Uno de los procedimientos más usados para generar números pseudoaleatorios es el *método de congruencias lineales*, que fue propuesto por Derrick Henry Lehmer en 1949, y se comporta como sigue. Supongamos que los números que queremos obtener tienen que estar entre 0 y $m-1$, ambos incluidos (resulta más cómodo describirlo así que con números entre 1 y m). Fijamos asimismo dos enteros a y c, con $2 \leq a < m$ y $0 \leq c < m$. Entonces, partiendo de un x_0 que cumpla $0 \leq x_0 < m$ vamos generando números pseudoaleatorios —tantos como vayamos necesitando— de manera recursiva mediante

$$x_{k+1} \equiv ax_k + c \pmod{m}, \quad 0 \leq x_{k+1} < m. \tag{2.12}$$

Al x_0 inicial se le denomina *semilla*, y se puede elegir a partir de algún parámetro físico que tenga cierta aleatoriedad; por ejemplo, el tiempo en microsegundos que lleva encendido el ordenador que efectúa los cálculos (módulo m). En la práctica, además, si se desea se pueden descartar los primeros números que se hayan obtenido.

Hemos dado un método de generar números. Esto no quiere decir que sea bueno; es importante analizar sus virtudes y defectos. Para estudiar este tipo de cosas existen herramientas estadísticas que «miden» —no vamos a explicar a qué nos referimos con ello— la aleatoriedad de lo que se ha obtenido. Dependiendo de qué se quiera hacer con esos números, esto puede no tener mucha importancia. Pero otras veces una mala elección de un método puede ser desastrosa: el resultado de una simulación puede estar viciado por culpa de la sucesión pseudoaleatoria usada. Aquí no nos vamos a ocupar de ello en demasía, pero sí que vamos a hacer un pequeño análisis.

A simple vista, el método de congruencias lineales tiene un problema. Cada número sólo depende del anterior luego, en cuanto un número aparezca por segunda vez, a partir de entonces toda la sucesión se repetirá de manera cíclica. Forzosamente, esto ocurrirá enseguida si m es pequeño (módulo m, sólo hay m números distintos), o sea que, como primera medida, hay que tomar m grande. Sería deseable que m fuese considerablemente mayor que la cantidad de números pseudoaleatorios que se quieran obtener; en este contexto, estamos usando «grande» y «pequeño» de manera relativa a lo que se pretende, y así seguiremos haciéndolo, sin precisar más.

Surge una dificultad si deseamos hacer muchos sorteos aleatorios (luego necesitamos que m sea grande), pero de modo que en cada sorteo sólo haya una cantidad pequeña de resultados posibles (que denominamos m'). En realidad, esto no es problema, pues no es difícil idear procedimientos para, tras hallar cada x_k con $0 \leq x_k < m$, proporcionar un x'_k que cumpla $0 \leq x'_k < m'$. Por ejemplo, mediante $x'_k = \lfloor x_k m'/m \rfloor$. Queda claro, por tanto, que no hace falta preocuparse de esto, y que podemos prescindir de la distinción entre m y m'.

La longitud del *periodo* de un generador de números pseudoaleatorios como (2.12) es el número de iteraciones hasta que se produce la primera repetición. Según estábamos comentando, esta longitud puede ser, como mucho, m. Obviamente, nos interesaría que así ocurriera. ¿Sucede esto siempre? Es fácil convencerse de que no es así en general:

Ejemplo. Analizar el comportamiento de los números generados por (2.12) si se toma $m = 12$, $a = 2$, $c = 5$ y $x_0 = 9$.

Solución. No hay más que fijarse en que

$$\begin{aligned} x_1 &\equiv 2 \cdot 9 + 5 \equiv 11 \pmod{12}, \\ x_2 &\equiv 2 \cdot 11 + 5 \equiv 3 \pmod{12}, \\ x_3 &\equiv 2 \cdot 3 + 5 \equiv 11 \pmod{12}, \end{aligned}$$

y a partir de aquí todo se repite, con un periodo de longitud 2. ◊

Visto que, en el ejemplo, la primera repetición se ha producido enseguida, es lógico preguntarse qué módulos m se pueden usar. Asimismo, los otros parámetros a y c, ¿cómo hay que escogerlos? El siguiente resultado nos responde a estas cuestiones (su demostración puede consultarse en [59, capítulo 3]):

Teorema 2.11.
Sean tres enteros m, a y c, con $2 \leq a < m$ y $0 \leq c < m$, y partamos de un x_0 que cumpla $0 \leq x_0 < m$. Entonces, el generador de números pseudoaleatorios

$$x_{k+1} \equiv ax_k + c \pmod{m}, \quad 0 \leq x_{k+1} < m$$

tiene periodo de longitud m si y sólo si se cumplen las siguientes condiciones:

- $\mathrm{mcd}(c, m) = 1$ *(lo cual excluye la posibilidad $c = 0$),*
- $a \equiv 1 \pmod{p}$ *para todo p divisor primo de m,*
- $a \equiv 1 \pmod{4}$ *si 4 divide a m.*

Si m es un producto de primos distintos (es decir, todos ellos con multiplicidad uno), es imposible que exista un a con $2 \leq a < m$ y que satisfaga la segunda condición. En efecto, si $m = p_1 p_2 \cdots p_l$ (con $l \geq 1$ y donde los p_j son primos distintos), un a que verificase $a \equiv 1 \pmod{p_j}$ para todo j tendría que tener la forma $a = p_1 p_2 \cdots p_l N + 1$ con $N \in \mathbb{Z}$, y ese número es menor que 2 (si $N \leq 0$) o mayor que m (si $N \geq 1$). Añadiendo la condición tercera, la imposibilidad análoga surge si m tiene la forma $m = 2^2 p_1 p_2 \cdots p_l$ con los p_j primos impares distintos.

En otro caso, encontrar a y c que satisfagan las condiciones del teorema es sencillo. Por ejemplo, si

$$m = 2^b p_1^{b_1} p_2^{b_2} \cdots p_l^{b_l}, \quad b \geq 0,\ b_j \geq 1,$$

siendo los p_j primos impares distintos (y $b > 2$ o algún $b_j > 1$), basta elegir cualquier c positivo que no tenga factores comunes con m, y como a se puede tomar $a = 2^{\min\{2,b\}} p_1 p_2 \cdots p_l N + 1$, con N un número natural arbitrario no demasiado grande para no sobrepasar m ($N = 1$, por ejemplo).

2.5. Primos en progresiones aritméticas

Uno de los más bonitos problemas de la teoría analítica de números es el de la existencia de infinitos primos en progresiones aritméticas. Aunque su solución completa no es fácil, veremos ahora que las congruencias son de gran ayuda para resolver de manera sencilla algunos casos. Comencemos planteando el problema.

En una progresión aritmética $a, q+a, 2q+a, \ldots, kq+a, \ldots$, de primer término un entero positivo a y diferencia otro entero $q > 1$, si $\mcd(a,q) > 1$ está claro que todos los números son compuestos (salvo quizás el primero). Luego una condición necesaria para que existan infinitos primos en la progresión es que $\mcd(a,q) = 1$. En 1837, Lejeune Dirichlet probó el recíproco en el siguiente sentido: si $\mcd(a,q) = 1$, existen infinitos primos en la sucesión $a, q+a, 2q+a, \ldots, kq+a, \ldots$ Lo hizo con técnicas de variable compleja y utilizando ingeniosos métodos de corte: los caracteres de Dirichlet.

Hoy en día, una de las demostraciones más «sencillas» (sólo con funciones reales y métodos de cálculo elemental) se debe a Harold N. Shapiro y se puede encontrar muy bien explicada en [5, capítulo 7]. Nosotros no la reproduciremos en este libro; simplemente analizamos aquí un par de casos particulares. Utilizando congruencias, el problema es: si $\mcd(a,q) = 1$, ¿existen infinitos primos p tales que $p \equiv a \pmod{q}$?

Comencemos examinando el caso $q = 4$ y $a = 3$. Se trata de saber si existen infinitos primos p tales que $p \equiv 3 \pmod 4$; dicho de otra forma, $p \equiv -1 \pmod 4$. Sirve un argumento similar al usado por Euclides para probar la existencia de infinitos primos:

Teorema 2.12.
Existen infinitos primos de la forma $4k-1$.

Demostración. Supongamos que existe sólo un número finito de estos primos: $p_1, \ldots, p_N$. Sea $A = 4p_1p_2\cdots p_N - 1$. Este A no puede ser primo, pues tiene forma $4k-1$ y es mayor que todos los primos de esa forma. Además, como $A \not\equiv 1 \pmod 4$, no todos los factores primos de A pueden ser congruentes con 1 módulo 4, ya que el producto de dos números de la forma $4k+1$ es de la misma forma. En consecuencia, algún factor primo de A debe tener la forma $4k-1$, y por tanto alguno de los primos $p_1, \ldots, p_N$ debería dividir a A, lo cual no es cierto. □

Tomemos ahora $q = 4$ y $a = 1$. Ya no se puede usar un argumento tan elemental como el anterior. Tenemos que recurrir al teorema de Euler-Fermat:

Teorema 2.13.
Existen infinitos primos de la forma $4k+1$.

Demostración. Vamos a ver que, cualquiera que sea el $n \in \mathbb{N}$ que escojamos, hay un primo $p > n$ que cumple $p \equiv 1 \pmod 4$; obviamente, esto equivale a decir que existen infinitos primos de la forma $4k+1$. Así pues, dado n, tomemos $A = (n!)^2 + 1$.

Los números $2, 3, 4, \ldots, n$ no dividen a A, luego el factor primo más pequeño de A es un $p > n$. Comprobemos que $p \equiv 1 \pmod 4$. Comencemos observando que

$$p \,|\, A \;\Rightarrow\; (n!)^2 \equiv -1 \pmod p \;\Rightarrow\; (n!)^{p-1} \equiv (-1)^{(p-1)/2} \pmod p.$$

Pero, por el teorema de Euler-Fermat (la hipótesis $\mcd(n!, p) = 1$ se cumple por ser $p > n$), $(n!)^{p-1} \equiv 1 \pmod p$. Así que $(-1)^{(p-1)/2} \equiv 1 \pmod p$. Y esto sólo sucede si $(p-1)/2$ es par, es decir, si $p - 1 \in 4\mathbb{Z}$, o $p \equiv 1 \pmod 4$. □

Los razonamientos elementales empleados en los dos teoremas anteriores se pueden extender a más casos. Por ejemplo, en [94, § 15] podemos ver cómo el argumento usado para $4k - 1$ se adapta fácilmente para mostrar la infinidad de primos de la forma $6k-1$. También existen demostraciones sencillas basadas en congruencias para las progresiones aritméticas de tipo $qk + 1$ con q arbitrario (aunque con distinto grado de dificultad dependiendo de q), así como para $5k-1$, $8k-1$, $8k-3$ y $8k+3$, entre otras (véase [95, capítulos 6 y 9]). Pero parecen estar lejos de proporcionar un argumento general válido para cualquier progresión.

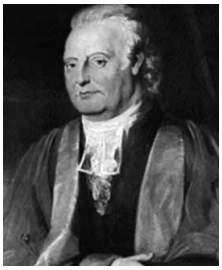
2.5.1. Progresiones aritméticas de primos

Un problema también relacionado con primos y progresiones aritméticas —pero que realmente no está conectado con el que hemos expuesto antes— es el de demostrar si existen o no progresiones aritméticas de primos de cualquier longitud. Explicado con más detalle, se trata de ver si, para cualquier entero positivo k (la longitud), existe un primo p y otro entero positivo d tal que los números

$$p, p + d, p + 2d, p + 3d, \ldots, p + (k - 1)d$$

son todos ellos primos. Por ejemplo, 31, 37, 43 es una progresión de longitud 3; y 13, 43, 73, 103 una de longitud 4. Con ayuda de ordenadores se han encontrado progresiones aritméticas de primos de longitud mayor que 20.

La existencia de progresiones aritméticas de primos arbitrariamente largas era una conjetura antigua. Fue alrededor de 1770 cuando Joseph Louis Lagrange y Edward Waring se preguntaron cuán grande podía ser k de modo que hubiera progresiones de longitud k. En 1924, Godfrey Harold Hardy y John Edensor Littlewood precisaron aún más la conjetura, dando, para cada k, estimaciones de cuántas sucesiones de longitud k (y primos hasta cierto tamaño) debía haber. Los primeros progresos teóricos no se produjeron hasta 1939, cuando van der Corput probó que existían infinitas progresiones de longitud 3; ése era un resultado interesante que, en principio, no aclaraba si la conjetura era o no cierta, pero el método usado en



Edward Waring
(1736–1798)



Godfrey Harold Hardy
(1877–1947)



John Edensor
Littlewood (1885–1977)

la prueba quizás podía dar nuevas ideas para demostrar si existían progresiones arbitrariamente largas.

Por otra parte, Pál Erdős y Pál Turán habían conjeturado, en 1936, que cualquier conjunto de enteros de densidad asintótica positiva debería contener progresiones aritméticas de longitud k, fuese quien fuese k; a estos efectos, y dado A un conjunto infinito de enteros, la densidad asintótica de A es, por definición,

$$\limsup_{x\to\infty} \frac{\operatorname{card}\{a \in A : a \leq x\}}{x}.$$

Los casos $k = 1$ y $k = 2$ son triviales, y el primer éxito en pos de la conjetura de Erdős-Turán fue de Klaus Roth en 1953, quien demostró el caso $k = 3$; años más tarde, en 1969, Endre Szemerédi lo probó para $k = 4$. El caso general tuvo que esperar hasta 1977 cuando, de nuevo Szemerédi, dio una ingeniosa demostración puramente combinatoria, aunque extremadamente complicada, de lo que desde entonces se conoce como teorema de Szemerédi. En los siguientes años, otros matemáticos proporcionaron demostraciones alternativas con diferentes enfoques.

Pero los primos no son un conjunto de densidad asintótica positiva: si denotamos $\pi(x) = \operatorname{card}\{p \text{ primo} : p \leq x\}$, se cumple $\pi(x)/x \to 0$ cuando $x \to \infty$ (eso es inmediato si utilizamos el teorema de los números primos,[6] según el cual

$$\limsup_{x\to\infty} \frac{\pi(x)}{x/\log(x)} = 1,$$

aunque no hace falta apoyarse en un resultado tan potente para probar que $\frac{\pi(x)}{x} \to 0$). Así pues, no se puede aplicar el teorema de Szemerédi para demostrar la existencia de progresiones aritméticas de primos de cualquier longitud.

[6] En este libro, el teorema de los números primos lo probaremos en el tema 10.



Johannes Gualtherus van der Corput (1890–1975)



Endre Szemerédi (1940–)

Finalmente, fue en 2004, utilizando métodos de análisis armónico, cuando Ben Green y Terence Tao probaron que sí existen progresiones aritméticas de primos de cualquier longitud (lo que ahora se conoce con el nombre de teorema de Green-Tao). Por sus numerosas contribuciones en análisis armónico —incluida la prueba de este resultado— Tao recibió la Medalla Fields[7] en el Congreso Internacional de Matemáticos celebrado en Madrid el año 2006.

De hecho, Green y Tao demostraron que, para números $\leq x$, hay aproximadamente $x^2/(\log(x))^k$ progresiones de primos de longitud k, salvo constante multiplicativa.

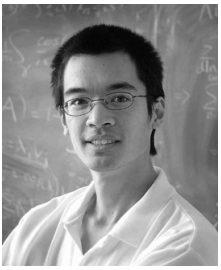
Nota 2.14. El teorema de Green-Tao sería un caso especial de la conjetura de Erdős sobre progresiones aritméticas, otra de las numerosas conjeturas debidas a Erdős.[8] En concreto, dicha conjetura —que permanece abierta— afirma que si A es un conjunto de enteros positivos tal que la suma de los inversos de los números de A es divergente, entonces A contiene progresiones aritméticas arbitrariamente

[7] Las Medallas Fields se entregan desde 1936, y están consideradas como el mayor premio que puede recibir un matemático (no existe Premio Nobel de matemáticas). Las otorga la Unión Matemática Internacional en los Congresos Internacionales de Matemáticos, que se celebran cada cuatro años (un interesante libro sobre la vertiente humana de estos congresos es [31]). Los galardonados siempre son matemáticos cuya edad no supera los 40 años; la motivación de este límite de edad es que se intenta «reconocer el trabajo matemático ya hecho, pero también animar y estimular posteriores logros de los premiados». Mucho más recientemente, otro premio vino a suplir la ausencia de Premio Nobel de matemáticas: el Premio Abel, creado por la Academia Noruega de las Ciencias y las Letras en honor del matemático noruego Niels Henrik Abel, y que se otorgó por primera vez el año 2003. Endre Szemerédi —que nos ha aparecido unas líneas más arriba— recibió el Premio Abel en 2012.

[8] Pál Erdős fue un matemático húngaro inmensamente prolífico —sólo comparable a Euler— y famoso por su excentricidad. Hizo contribuciones en combinatoria, teoría de grafos, teoría de números, análisis clásico, teoría de aproximación, teoría de conjuntos y probabilidad, y propuso



Ben Joseph Green (1977–)



Terence Tao (1975–)

largas. Dado que la serie de los inversos de los números primos es divergente —un resultado debido a Euler y cuya demostración veremos en el tema 9—, alguien que lograra demostrar la conjetura de Erdős sobre progresiones aritméticas obtendría el teorema de Green-Tao como un simple corolario. ◁

2.6. Sistemas lineales de congruencias: el teorema chino del resto

En esta sección vamos a discutir qué ocurre con los sistemas de congruencias simultáneas. Sabemos resolver perfectamente una sola ecuación con congruencias lineal. ¿Qué ocurre si queremos que se satisfagan varias a la vez? En concreto, nos ocuparemos de sistemas (lineales) de congruencias en los que aparece una sola incógnita, pero en los cuales cada ecuación se debe satisfacer con un módulo diferente. Este tipo de problemas apareció en las antiguas civilizaciones china e india en conexión con la elaboración de calendarios y el cálculo de acontecimientos astronómicos.

múltiples problemas por cuya solución ofrecía recompensas económicas; posiblemente la conjetura de Erdős sobre progresiones aritméticas es uno de los más notables.

Tras una vida incansable dedicada a las matemáticas —publicó alrededor de 1500 artículos, casi todos ellos de manera conjunta con alguno de sus más de 500 coautores—, murió en 1996, a la edad de 83 años, mientras asistía a una conferencia en Varsovia. Colaboradores y amigos inventaron el número de Erdős como un homenaje con tintes de humor matemático: Erdős tiene asignado el número 0, todos aquéllos que colaboraron con él en algún artículo tienen el 1, alguien que haya colaborado con alguno de sus colaboradores tiene el 2, y así sucesivamente. Con seriedad matemática y de manera recursiva, se puede expresar así: para cada matemático Y, su número de Erdős E_Y se define como

$$E_Y = 1 + \min\{E_{C_1}, \ldots, E_{C_n}\},$$

siendo $\{C_1, \ldots, C_n\}$ el conjunto de todos los coautores de Y, y donde $E_{\text{Erdős}} = 0$.



Yang Hui
(≈ 1238 – ≈ 1298)



Aryabhata (476–550)



Nicómaco de Gerasa
(≈ 60 – ≈ 120)

Así, gran parte de las técnicas que expondremos eran ya conocidas por matemáticos de los que la historia ha conservado muy pocos datos; a veces, ni siquiera se sabe a ciencia cierta la época en la que vivieron. Entre los matemáticos chinos que se ocuparon de esto podemos destacar a Sun Zi (vivió alrededor del siglo V, y también es conocido como «Maestro Sun»), Qin Jiushao (siglo XIII) y Yang Hui (siglo XIII); entre los indios, a Aryabhata (siglo VI), Brahmagupta (siglo VII), Bhaskara I (siglo VII), Mahavira (siglo IX) y Bhaskara II (siglo XII).[9] También aparecen sistemas de congruencias en la obra del matemático árabe Ibn al-Haytham[10] (siglo XI). En cuanto a la civilización occidental, algunos ejemplos sueltos aparecen en los trabajos del griego Nicómaco (siglo I) y de Fibonacci (siglo XIII); pero, desde luego, no ofrecen una teoría sistemática, de la que sí disponían los matemáticos chinos. En la matemática europea, fueron Euler, Lagrange y Gauss —en los siglos XVIII y XIX— quienes desarrollaron completamente el tema.

Comencemos mostrando un problema propuesto por el matemático chino Sun Zi: encontrar un número que tenga resto 1 al dividirlo por 3, resto 2 al dividirlo por 5, y resto 3 al dividirlo por 7. Esto se puede expresar mediante el siguiente sistema de congruencias:

$$\begin{cases} x \equiv 1 \pmod 3, \\ x \equiv 2 \pmod 5, \\ x \equiv 3 \pmod 7. \end{cases} \tag{2.13}$$

Otro ejemplo clásico, esta vez de origen indio, se puede ver en el ejercicio 2.21 (el mismo ejemplo, aunque sin aludir a cestas y a huevos como en su redacción india, aparece en el *Liber Abaci* de Fibonacci).

[9] Los añadidos «I» y «II» en los dos matemáticos de nombre Bhaskara se usan únicamente para poder distinguirlos.

[10] Véase la nota al pie numerada como 12, en la página 123.

En algunas ocasiones, este tipo de sistemas no tiene solución. Por ejemplo, no existe ningún x que satisfaga

$$\begin{cases} x \equiv 1 \pmod 3, \\ x \equiv 0 \pmod 6, \end{cases} \tag{2.14}$$

ya que cualquier x múltiplo de 6 debe ser también múltiplo de 3. Sin embargo, cada una de las ecuaciones, separadamente, tiene solución.

Otras veces, hallar la solución es muy sencillo: Para el sistema[11]

$$\begin{cases} x \equiv 2 \pmod 3, \\ x \equiv 2 \pmod 5, \\ x \equiv 2 \pmod 7, \end{cases}$$

cualquier x de la forma $x = 3 \cdot 5 \cdot 7k + 2$, con $k \in \mathbb{Z}$, es solución (en particular, la solución $x = 2$, que se corresponde con $k = 0$, es trivial). No parece haber ningún truco similar aplicable a (2.13).

Así pues, ¿cuándo tiene este tipo de ecuaciones solución? ¿Si las hay, cómo hallarlas? ¿Y cuántas hay? En el ejemplo mostrado en (2.14), los módulos 3 y 6 no son primos entre sí. Siempre que en un sistema haya dos módulos no primos entre sí, es fácil poner ejemplos en los que el correspondiente par de ecuaciones no tiene solución simultánea. Cuando los módulos son primos entre sí dos a dos, el siguiente resultado responde a todas nuestras preguntas:

Teorema 2.15. Teorema chino del resto.
Sean $m_1, m_2, \ldots, m_r$ enteros positivos, primos entre sí dos a dos, es decir, tales que $\operatorname{mcd}(m_j, m_k) = 1$ siempre que $j \neq k$. Entonces, para cualesquiera enteros arbitrarios $b_1, b_2, \ldots, b_r$, el sistema de congruencias

$$\begin{cases} x \equiv b_1 \pmod{m_1}, \\ x \equiv b_2 \pmod{m_2}, \\ \quad \ldots \\ x \equiv b_r \pmod{m_r} \end{cases} \tag{2.15}$$

posee exactamente una solución módulo el producto $m_1 m_2 \cdots m_r$.

[11] Obsérvese que en la argumentación que sigue al teorema 2.11 lo que hemos hecho —sin decirlo— ha sido resolver un sistema de este tipo.

DEMOSTRACIÓN. Sea $M = m_1 m_2 \cdots m_r$ y $M_j = M/m_j$, $j = 1, 2, \ldots, r$. Por ser $\mcd(m_j, m_k) = 1$ para $j \neq k$, se deduce fácilmente que $\mcd(M_j, m_j) = 1$ para todo j. Entonces, la proposición 2.1 prueba que cada M_j tiene un único inverso M'_j módulo m_j. (A menudo se usa a^{-1} para denotar el inverso de a módulo m; pero este inverso depende del módulo, y aquí estamos empleando módulos distintos, o sea que la notación $(\cdot)^{-1}$ hubiera resultado confusa.) Sea ahora el entero

$$x = b_1 M_1 M'_1 + b_2 M_2 M'_2 + \cdots + b_r M_r M'_r,$$

y analicemos su comportamiento módulo cada uno de los m_j. Es claro que $M_k \equiv 0 \pmod{m_j}$ para $k \neq j$, luego todos los sumandos de x son nulos módulo m_j, salvo quizás el j-ésimo, así que

$$x \equiv b_j M_j M'_j \equiv b_j \pmod{m_j}.$$

En consecuencia, hemos probado que x satisface todas las congruencias del sistema.

Nos falta demostrar que la solución es única módulo M, pero esto es rutinario. En primer lugar, veamos que si x es solución e y cumple $y \equiv x \pmod{M}$, también y es solución. En efecto, en esas condiciones tendremos $M \,|\, (x - y)$ y $m_j \,|\, M$, luego $m_j \,|\, (x - y)$ y por tanto $y \equiv x \equiv b_j \pmod{m_j}$ para todo j. Comprobemos ahora el recíproco, es decir, que todas las posibles soluciones son de esa manera. Si dos enteros x e y son soluciones del sistema, deberán cumplir $x \equiv y \pmod{m_j}$ para todo j, o sea, $m_j \,|\, (x - y)$. Como los m_j son primos entre sí dos a dos, tendremos también $M \,|\, (x - y)$, o sea, $x \equiv y \pmod{M}$. □

Es importante fijarse en la demostración, pues proporciona el algoritmo para calcular explícitamente las soluciones este tipo de sistemas. Destaquémoslo:

Corolario 2.16.

En las condiciones del teorema anterior, denotemos $M = m_1 m_2 \cdots m_r$, $M_j = M/m_j$, $j = 1, 2, \ldots, r$, y M'_j el inverso de M_j módulo m_j. Entonces, la solución general de (2.15) es

$$x = b_1 M_1 M'_1 + b_2 M_2 M'_2 + \cdots + b_r M_r M'_r + kM, \quad k \in \mathbb{Z} \tag{2.16}$$

(o, dicho de otra forma, las soluciones de (2.15) son los elementos del conjunto $b_1 M_1 M'_1 + b_2 M_2 M'_2 + \cdots + b_r M_r M'_r + M\mathbb{Z}$).

De este modo, ya podemos resolver fácilmente sistemas lineales de congruencias:

Ejemplo. Hallar todas las soluciones de (2.13), e identificar la menor solución no negativa.

SOLUCIÓN. Con la notación de la demostración del teorema y su corolario, $M = 3 \cdot 5 \cdot 7 = 105$, $M_1 = 105/3 = 35$, $M_2 = 105/5 = 21$ y $M_3 = 105/7 = 15$. El inverso de M_1 módulo 3 es $M_1' = 2$, ya que $35 \equiv 2 \pmod 3$ y $2 \cdot 2 = 4 \equiv 1 \pmod 3$ (o, directamente, $35 \cdot 2 = 70 \equiv 1 \pmod 3$, aunque este modo de razonar emplea números mayores, y en consecuencia es más difícil encontrar el inverso «a ojo»). De manera similar, el de M_2 módulo 5 es $M_2' = 1$ (pues $21 \equiv 1 \pmod 5$, y 1 es su propio inverso); y el de M_3 módulo 7 es $M_3' = 1$. En estas condiciones, la solución de (2.13) es

$$x = 1 \cdot 35 \cdot 2 + 2 \cdot 21 \cdot 1 + 3 \cdot 15 \cdot 1 = 157,$$

que es única módulo 105. Finalmente, como $157 \equiv 52 \pmod{105}$, y parece más cómodo usar 52 que 157, llegamos a la conclusión de que las soluciones de (2.13) son precisamente las que tienen la forma $x = 52 + 105k$ con $k \in \mathbb{Z}$. Por supuesto, la menor solución no negativa es $x = 52$. ◊

Nota 2.17. En lugar de aplicar directamente (2.16), un método alternativo de hallar la solución del sistema consiste en ir resolviendo, de manera sucesiva, cada ecuación individual del sistema. Lo más sencillo es explicarlo por medio de un ejemplo: en seguida lo hacemos, de nuevo con el sistema (2.13). En este proceso, lo único que tenemos que recordar es que, cuando hemos encontrado una solución particular de un subsistema, la solución general del subsistema se obtiene sumando múltiplos del producto de los módulos involucrados en el subsistema. Al llevar a la práctica el proceso, conviene empezar por las ecuaciones de módulo más alto.

Veamos cómo hacerlo con (2.13). Las soluciones de $x \equiv 3 \pmod 7$ son los enteros de la forma $x = 3 + 7k$, con $k \in \mathbb{Z}$. De estas, buscamos una que cumpla $x \equiv 2 \pmod 5$, lo cual se consigue con $k = 2$, que da $x = 17$. Entonces, las soluciones del sistema formado por las ecuaciones segunda y tercera son $x = 17 + 35k$. De entre estos números, para que se cumpla $x \equiv 1 \pmod 3$ elegimos $k = 1$, que da $x = 52$. Entonces, la solución general del sistema es $x = 52 + 105k$ con $k \in \mathbb{Z}$. ◁

El siguiente resultado aparenta ser bastante más general que el teorema chino del resto, pero en realidad sólo requiere un pequeño paso adicional.

Teorema 2.18.
Sean $m_1, m_2, \ldots, m_r$ enteros positivos, primos entre sí dos a dos. Sean, asimismo, $b_1, b_2, \ldots, b_r$ enteros arbitrarios, y $a_1, a_2, \ldots, a_r$ enteros tales que

$$\operatorname{mcd}(a_j, m_j) = 1 \quad \textit{para } j = 1, 2, \ldots, r.$$

Entonces, el sistema de congruencias

$$\begin{cases} a_1x \equiv b_1 \pmod{m_1}, \\ a_2x \equiv b_2 \pmod{m_2}, \\ \quad \dots \\ a_rx \equiv b_r \pmod{m_r} \end{cases}$$

posee exactamente una solución módulo $m_1m_2\cdots m_r$.

Demostración. Sea a'_j el inverso de a_j módulo m_j, que existe, pues $\mcd(a_j, m_j) = 1$. Entonces, la congruencia $a_jx \equiv b_j \pmod{m_j}$ es equivalente a la congruencia $x \equiv a'_jb_j \pmod{m_j}$. Ahora, basta aplicar el teorema chino del resto. □

El teorema chino del resto se puede usar para convertir una congruencia cuyo módulo es un número compuesto grande (cuya descomposición en factores conocemos) en un sistema de varias congruencias con módulos más pequeños, en los que resulta más sencillo operar.

Ejemplo. Convertir en sistema la ecuación $11x \equiv 72 \pmod{380}$ y resolverla.

Solución. En primer lugar, démonos cuenta que la solución de esa ecuación es, simplemente, $x = 11^{-1}72 + 380k$, donde 11^{-1} es el inverso de 11 módulo 380; pero los números que estamos manejando no son muy pequeños, e identificar ese 11^{-1} no es inmediato. Así pues, vamos a simplificar el problema de manera que aparezcan módulos menores.

De la descomposición en factores $380 = 2^2\cdot5\cdot19$ se sigue que 380 divide a $11x-72$ si y sólo si lo dividen 4, 5 y 19; por tanto, la ecuación propuesta es equivalente al sistema

$$\begin{cases} 11x \equiv 72 \pmod 4, \\ 11x \equiv 72 \pmod 5, \\ 11x \equiv 72 \pmod{19}. \end{cases}$$

Además, $11 \equiv 3 \pmod 4$, $11 \equiv 1 \pmod 5$, y análogamente 72 se puede reducir respecto a los tres módulos involucrados. Entonces, el sistema queda

$$\begin{cases} 3x \equiv 0 \pmod 4, \\ x \equiv 2 \pmod 5, \\ 11x \equiv 15 \pmod{19}. \end{cases}$$

Para aplicar directamente (2.16) tenemos que eliminar los coeficientes 3 y 11, multiplicando las ecuaciones en las que aparecen por sus correspondientes inversos. En

realidad, $3x \equiv 0 \pmod 4$ es equivalente a $x \equiv 0 \pmod 4$ sin necesidad de buscar $3^{-1} \pmod 4$. Finalmente, como $11^{-1} \pmod{19}$ es 7, y $7 \cdot 15 = 105 \equiv 10 \pmod{19}$, el sistema queda reducido a

$$\begin{cases} x \equiv 0 \pmod 4, \\ x \equiv 2 \pmod 5, \\ x \equiv 10 \pmod{19}. \end{cases}$$

Entonces, con la notación de (2.16), $M = 380$, $M_1 = 95$, $M_2 = 76$, $M_3 = 20$ y

$$\begin{aligned} M_1' &\equiv (95)^{-1} \equiv 3^{-1} \equiv 3 \pmod 4, \\ M_2' &\equiv (76)^{-1} \equiv 1^{-1} \equiv 1 \pmod 5, \\ M_3' &\equiv (20)^{-1} \equiv 1^{-1} \equiv 1 \pmod{19}. \end{aligned}$$

Así que la solución general tanto del sistema como de la ecuación de partida es

$$x = 0 \cdot 95 \cdot 3 + 2 \cdot 76 \cdot 1 + 10 \cdot 20 \cdot 1 + 380k, \quad k \in \mathbb{Z},$$

o sea, $x = 352 + 380k$, $k \in \mathbb{Z}$. ◊

2.6.1. Módulos no primos entre sí

¿Qué ocurre si dos m_j no son primos entre sí? Ya hemos comentado que, según cuáles sean los b_j, puede no haber solución del sistema. Pero es claro que también puede haberla. ¿Cuándo existe solución, y cómo hallarla?

El teorema que da condiciones necesarias y suficientes con un número arbitrario de ecuaciones tiene una demostración un tanto enrevesada, así que vamos a analizar primero el caso correspondiente a sólo dos ecuaciones. Su enunciado es muy similar, pero en su demostración se observa mucho mejor por qué unas veces hay solución y otras no. Aquí juega un papel importante el máximo común divisor de los módulos y, cuando se cumplen las condiciones requeridas, es el algoritmo de Euclides —puesto en forma de identidad de Bézout, tal como comentamos en la nota 2.3 de la página 95— el que permite probar la existencia de solución. Cuando existe, la solución es única módulo el mínimo común múltiplo de los módulos.

Proposición 2.19.

Sean dos enteros $m, n > 1$ y otros dos enteros arbitrarios b, c. Entonces, el sistema

$$\begin{cases} x \equiv b \pmod m, \\ x \equiv c \pmod n \end{cases}$$

tiene solución si y sólo si

$$b \equiv c \pmod{\operatorname{mcd}(m,n)}.$$

*Además, si x es solución del sistema, entonces y es solución si y sólo si $x \equiv y \pmod{\operatorname{mcm}(m,n)}$ (por supuesto, «*mcm*» denota el mínimo común múltiplo).*

Demostración. Supongamos que x es una solución del sistema. Entonces, $x = b + mk$ (para algún $k \in \mathbb{Z}$) y $x = c + nj$ (para algún $j \in \mathbb{Z}$), luego $b + mk = c + nj$ y $c - b = mk - nj$. Así, es evidente que $c - b$ es múltiplo de $\operatorname{mcd}(m,n)$, es decir, $c \equiv b \pmod{\operatorname{mcd}(m,n)}$.

Recíprocamente, si $c \equiv b \pmod{\operatorname{mcd}(m,n)}$, será $c - b = s\operatorname{mcd}(m,n)$ con $s \in \mathbb{Z}$. La identidad de Bézout nos asegura que existen enteros k_1 y j_1 tales que $\operatorname{mcd}(m,n) = mk_1 - nj_1$. Entonces, $c - b = mk - nj$ y, por consiguiente, $x = b + mk = c + nj$ es solución del sistema de congruencias.

Finalmente, cuando hay soluciones, hallémoslas todas. Si x es una solución del sistema, y también lo será si y sólo si $y \equiv b \equiv x \pmod{m}$ y $y \equiv c \equiv x \pmod{n}$. Esto es, si y sólo si $x - y$ es un múltiplo tanto de m como de n, o, equivalentemente, si $x - y$ es múltiplo de $\operatorname{mcm}(m,n)$, lo cual completa la demostración. □

Con r ecuaciones, lo que sucede es lo siguiente:

Teorema 2.20.
Sean $m_1, m_2, \ldots, m_r$ enteros positivos, y $b_1, b_2, \ldots, b_r$ enteros arbitrarios. Entonces, el sistema de congruencias

$$\begin{cases} x \equiv b_1 \pmod{m_1}, \\ x \equiv b_2 \pmod{m_2}, \\ \quad \ldots \\ x \equiv b_r \pmod{m_r} \end{cases} \tag{2.17}$$

tiene solución si y sólo si

$$\operatorname{mcd}(m_j, m_k) \mid (b_j - b_k) \quad \textit{siempre que } j \neq k. \tag{2.18}$$

Cuando esta condición se cumple, la solución del sistema es única módulo $\operatorname{mcm}(m_1, m_2, \ldots, m_r)$.

Demostración. Comencemos comprobando que (2.18) es una condición necesaria para la existencia de solución del sistema. En efecto, si x es tal solución, las filas j y

k del sistema se pueden escribir como $m_j \,|\, (x - b_j)$ y $m_k \,|\, (x - b_k)$. En consecuencia, $\mcd(m_j, m_k)$ dividirá $x - b_j$ y a $x - b_k$, y por tanto a su resta, es decir, $\mcd(m_j, m_k) \,|\, (b_j - b_k)$.

Para analizar el recíproco veamos que, sin pérdida de generalidad, podemos suponer que todos los m_j son potencias de primos. Para ello, si la descomposición en factores primos de m es $m = p_1^{a_1} \cdots p_s^{a_s}$, la ecuación $x \equiv b \pmod{m}$ es equivalente a que, simultáneamente, se cumplan las ecuaciones $x \equiv b \pmod{p_1^{a_1}}$, $\ldots$, $x \equiv b \pmod{p_s^{a_s}}$. De este modo, el sistema (2.17) es equivalente a otro —posiblemente con más ecuaciones— en el que todos los módulos son potencias de primos. Este proceso no altera el cumplimiento o no de la condición (2.18) ya que, fijados un par j, k, es claro $\mcd(m_j, m_k) \,|\, (b_j - b_k)$ si y sólo si $\mcd(p_j^{a_j}, p_k^{a_k}) \,|\, (b_j - b_k)$ para todos los primos p_j tal que $p_j^{a_j} \parallel m_j$ y $p_k^{a_k} \parallel m_k$ (con $p^a \parallel b$ queremos decir $p^a \,|\, b$ y $p^{a+1} \nmid b$). Finalmente, con este procedimiento tampoco se modifica el mínimo común múltiplo de los módulos.

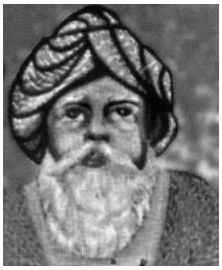
Así pues, sean $m_j = p_j^{a_j}$ con p_j primo y $a_j \geq 1$, para $j = 1, 2, \ldots, r$, y supongamos que se cumple (2.18). Ahora, $\mcd(p_j^{a_j}, p_k^{a_k}) = 1$ (y divide a $b_j - b_k$ trivialmente) salvo si $p_j = p_k$. Si esto ocurre para alguna pareja j, k, tomemos el primo $p = p_j = p_k$ con exponente $\max\{a_j, a_k\}$. Para fijar ideas, supongamos que $a_j \geq a_k$ (lo mismo daría si fuera al revés); por consiguiente, $\mcd(p_j^{a_j}, p_k^{a_k}) = p^{a_k}$. Vamos a ver que, entonces, la ecuación $x \equiv b_k \pmod{p_k^{a_k}}$ es redundante, y por tanto podemos eliminarla del sistema. En efecto, cualquier x que satisfaga la ecuación $x \equiv b_j \pmod{p_j^{a_j}}$ también cumplirá $p^{a_k} \mid p^{a_j} \mid (x - b_j)$. Como, por (2.18), $p^{a_k} \mid (b_j - b_k)$, tendremos $p^{a_k} \,|\, ((x - b_j) + (b_j - b_k)) = (x - b_k)$, luego ese mismo x es solución de $x \equiv b_k \pmod{p_k^{a_k}}$.

Este proceso de eliminar ecuaciones —que, además, no altera el mínimo común múltiplo de los módulos— puede hacerse tantas veces como primos repetidos haya. Cuando se han suprimido todas las ecuaciones en las que los primos coincidían, el mínimo común múltiplo de los módulos que quedan es, ahora, el producto de todos ellos. Para no complicar la notación, seguimos llamando r al número de ecuaciones que nos han quedado.

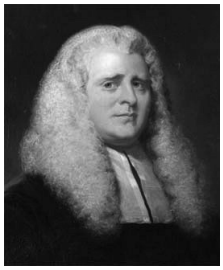
En resumen, hemos reducido (2.17) a un sistema de la forma

$$\begin{cases} x \equiv b_1 \pmod{p_1^{a_1}}, \\ x \equiv b_2 \pmod{p_2^{a_2}}, \\ \quad \ldots \\ x \equiv b_r \pmod{p_r^{a_r}}, \end{cases}$$

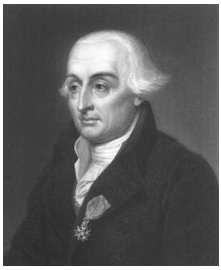
donde los p_j son primos distintos. Como $\mcd(p_j^{a_j}, p_k^{a_k}) = 1$, este sistema cumple las hipótesis del teorema chino del resto (teorema 2.15), luego tiene solución, que es



Ibn al-Haytham
(965–1040)



John Wilson
(1741–1793)



Joseph Louis Lagrange
(1736–1813)

única módulo $\prod_{j=1}^{r} p_j^{a_j}$. Ya sólo queda recordar que este producto coincide con el $\operatorname{mcm}(m_1, m_2, \ldots, m_r)$ que aparece en el enunciado. □

Como procedimiento práctico de resolución de este tipo de sistemas podemos hacer como en la nota 2.17. La única diferencia es que para pasar de la solución particular de un subsistema a su solución general, esto se hace módulo el mínimo común múltiplo de los módulos que aparecen en el subsistema.

2.7. Teorema de Wilson

Basándose en lo que ocurría en un número grande de casos, Wilson conjeturó que, si p era primo, $(p-1)!+1$ siempre iba a ser divisible por p; sin embargo, no pudo probarlo. Mucho antes —alrededor del año 1000— el matemático árabe Ibn al-Haytham[12] ya era consciente de esa propiedad. Fue Lagrange en 1773 —pocos años después de que Wilson lo intuyera— el primero que logró encontrar una demostración. No obstante no haber sido él quien lo probó, siempre se alude a este bonito resultado con el nombre de teorema de Wilson.

En muchos libros de teoría de números, la demostración de este teorema se hace a partir del de Lagrange;[13] véase, por ejemplo, [5, § 5.5 y 5.6]. El teorema de Lagrange al que ahora aludimos trata sobre congruencias polinómicas, y nosotros lo hemos postergado hasta el tema 3 (concretamente, es el teorema 3.1). Así, podríamos haber

[12] Haciendo alusión a su ciudad de nacimiento en Persia, a Ibn al-Haytham a veces se le conoce como al-Basri; y, refiriéndose al lugar de Egipto en el que posteriormente vivió, al-Misri. Así mismo, también es conocido por su nombre latinizado Alhazen o Alhacén (que procede del «al-Hasan» de su nombre completo, Abu Ali al-Hasan ibn al-Haytham).

[13] Este teorema de Lagrange no tiene nada que ver con el que aparece en la nota 2.5 (página 99), pese a que ambos se conocen con el mismo nombre.

seguido aquí ese mismo camino, aunque cambiando el orden en la exposición de los resultados (de hecho, una vez que en el tema 3 veamos el teorema de Lagrange, en el ejercicio 3.7 proponemos al lector que lo use para deducir el de Wilson).

Pero, no sólo para poder seguir el orden lógico escogido, sino también en aras de lograr la independencia de los resultados mostrados (y que, en consecuencia, el lector pueda prescindir de los que le interesen menos), vamos a hacerlo de otra manera. Más concretamente, sólo vamos a usar el siguiente lema —bastante sencillo, por cierto—, en el que únicamente intervienen congruencias cuadráticas:

Lema 2.21.
Si p es primo, entonces $x^2 \equiv 1 \pmod{p}$ si y sólo si $x \equiv \pm 1 \pmod{p}$.

Demostración. La implicación «si» es evidente, o sea que analicemos la parte «sólo si». Supongamos que $x^2 \equiv 1 \pmod{p}$, es decir, p divide a $x^2-1 = (x-1)(x+1)$. Como p es primo, deberá cumplirse $p \mid (x-1)$ o $p \mid (x+1)$, y por tanto $x \equiv 1 \pmod{p}$ o $x \equiv -1 \pmod{p}$. □

Una vez visto el lema, vamos con el teorema:

Teorema 2.22. Teorema de Wilson.
Si p es primo, entonces $(p-1)! \equiv -1 \pmod{p}$.

Demostración. El resultado es claramente cierto cuando $p = 2$ o $p = 3$, así que supongamos $p \geq 5$. Como p es primo, cualquier $a \in \{1, 2, \ldots, p-1\}$ verifica $\mathrm{mcd}(a, p) = 1$. Entonces, la proposición 2.1 nos asegura que existe un único $a^{-1} \in \{1, 2, \ldots, p-1\}$ tal que $aa^{-1} \equiv 1 \pmod{p}$. Por el lema, $a = a^{-1}$ si y sólo si $a = 1$ o $a = p-1$. De este modo, podemos partir el conjunto $\{2, 3, \ldots, p-2\}$ en $(p-3)/2$ pares de enteros $\{a_j, a_j^{-1}\}$ tales que $a_j a_j^{-1} \equiv 1 \pmod{p}$ para $j = 1, 2, \ldots, (p-3)/2$. En consecuencia,

$$(p-1)! = 1 \cdot 2 \cdot 3 \cdots (p-2)(p-1) = (p-1) \prod_{j=1}^{(p-3)/2} a_j a_j^{-1} \equiv p-1 \equiv -1 \pmod{p},$$

tal como queríamos probar. □

Es interesante comprobar que el inverso del teorema de Wilson también es cierto:

Teorema 2.23.
Si un entero $n > 1$ verifica $(n-1)! \equiv -1 \pmod{n}$, entonces n es primo.

DEMOSTRACIÓN. Si n no fuera primo, existiría un primo $p < n$ que divide a n; en particular, $p \mid (n-1)!$. Pero, por hipótesis, $p \mid ((n-1)!+1)$, luego $p \mid 1$, lo cual es imposible. □

Por ejemplo, si tomamos $n = 6$ tenemos $(6-1)! = 120 \equiv 0 \pmod 6$, y de aquí se deduce que 6 es un número compuesto (aunque esto era ya evidente antes de empezar).

El teorema de Wilson y su inverso proporcionan un verdadero test de primalidad: para decidir si un entero $n > 1$ es primo o compuesto, basta comprobar si $(n-1)! \equiv -1 \pmod n$ o no. Desafortunadamente, este criterio no tiene ninguna utilidad práctica, ya que calcular $(n-1)!$ módulo n es demasiado trabajoso.

2.7.1. Primos de Wilson

Concluyamos la sección mencionando un par de problemas abiertos. El teorema de Wilson establece que, cuando p es primo, entonces $p \mid ((p-1)!+1)$. En esas circunstancias, el *cociente de Wilson*

$$w_p = \frac{(p-1)!+1}{p}$$

es un entero. A los p para los cuales w_p es divisible por p los denominamos primos de Wilson; es decir, nos estamos preguntando por los p tales que $(p-1)! \equiv -1 \pmod{p^2}$. Del mismo modo que ocurre con los primos de Wieferich antes mencionados, razonamientos heurísticos hacen pensar que un p es de Wilson con probabilidad $1/p$. Sin embargo, pese a las intensas búsquedas por ordenador realizadas, los únicos primos de Wilson conocidos son $p = 5$, 13 y 563, muchos menos que los esperados (y no hay ningún otro menor que $5 \cdot 10^8$). No se sabe a ciencia cierta si hay o no infinitos de ellos. Para más detalles, véase [29].

Otra interesante cuestión —planteada por Javier Soria en 2003— es cuándo los cocientes de Wilson w_p constituyen un número primo. Se tienen unos cuantos ejemplos, como los correspondientes a $p = 5$, 7, 11, 29, 773, 1321 y 2621, y se piensa que habrá infinitos w_p primos (de nuevo apoyándose en razonamientos heurísticos). Pero no se sabe probar rigurosamente si eso es así.

2.8. Números pseudoprimos

Es interesante observar que el teorema pequeño de Fermat nos proporciona un test de primalidad para cada base a: dado un número m, si $a^m \not\equiv a \pmod m$ entonces m no es primo. En otras palabras, si encontramos un a tal que $a^m \not\equiv a$

(mód m), entonces m tiene que ser forzosamente un número compuesto. Esto evita tener que hallar la descomposición en factores de m, algo que, cuando m es grande, requeriría un trabajo inmenso.

Para evitar trivialidades, añadamos la hipótesis $\mcd(a, m) = 1$ —esto siempre es sencillo y rápido de comprobar, sin más que aplicar el algoritmo de Euclides—. En estas circunstancias, podemos escribir $a^m \equiv a \pmod{m}$ en su forma equivalente $a^{m-1} \equiv 1 \pmod{m}$.

Demos a esto un nombre: cuando un número m cumple $a^{m-1} \equiv 1 \pmod{m}$ para cierto a tal que $\mcd(a, m) = 1$, decimos que m pasa el test de pseudoprimalidad en base a. Un número primo pasa el test de pseudoprimalidad en todas las bases. Pero, ¿será verdad la implicación contraria?, es decir, ¿qué ocurre si m no es primo?

De momento, veamos un ejemplo de cómo se puede usar esto para detectar números compuestos:

Ejemplo. Demostrar que 437 es un número compuesto, comprobando para ello que no pasa el test de pseudoprimalidad en base 2.

Solución. Es inmediato que $\mcd(437, 2) = 1$, así que dediquémonos a evaluar $2^{436} \pmod{437}$. Por supuesto, utilizaremos el cálculo de potencias mediante cuadrados sucesivos. Descompongamos $436 = 2^8 + 2^7 + 2^5 + 2^4 + 2^2$ (recordemos que esto se corresponde con la representación binaria $436 = 110\,110\,100_{(2)}$), luego $2^{436} = 2^{2^8} \cdot 2^{2^7} \cdot 2^{2^5} \cdot 2^{2^4} \cdot 2^{2^2}$. Como $2^{2^k} = \left(2^{2^{k-1}}\right)^2$, tenemos

$$\begin{aligned}
2^{2^2} &\equiv 4 \cdot 4 \equiv 16 \pmod{437},\\
2^{2^3} &\equiv 16 \cdot 16 \equiv 256 \pmod{437},\\
2^{2^4} &\equiv 256 \cdot 256 \equiv 423 \pmod{437},\\
2^{2^5} &\equiv 423 \cdot 423 \equiv 196 \pmod{437},\\
2^{2^6} &\equiv 196 \cdot 196 \equiv 397 \pmod{437},\\
2^{2^7} &\equiv 397 \cdot 397 \equiv 289 \pmod{437},\\
2^{2^8} &\equiv 289 \cdot 289 \equiv 54 \pmod{437}.
\end{aligned}$$

Entonces

$$\begin{aligned}
2^{436} &\equiv 54 \cdot 289 \cdot 196 \cdot 423 \cdot 16 \pmod{437}\\
&\equiv 358 \not\equiv 1 \pmod{437},
\end{aligned}$$

luego efectivamente 437 no es pseudoprimo en base 2, y en consecuencia no puede ser un número primo. ◊

Hay que reconocer que, en este ejemplo, posiblemente sea más sencillo encontrar directamente la descomposición $437 = 19 \cdot 23$. Pero —y estamos apelando a la credulidad del lector, pues no vamos a desarrollar lo que sigue con más profundidad— cuando el número m es grande, calcular $a^{m-1} \pmod m$ resulta mucho menos costoso que factorizar m; desde luego, podemos suponer que disponemos de algún ordenador, con su correspondiente programa informático, que nos ayuda a efectuar todos estos cálculos.

Si nos fijamos en una base concreta —por ejemplo, 2, como antes—, es fácil convencerse de que este criterio sólo sirve para probar que un número es compuesto, pero no que es primo. En efecto; tal como descubrió Pierre Frédéric Sarrus en 1819, el número compuesto $341 = 11 \cdot 31$ verifica $2^{340} \equiv 1 \pmod{341}$ (véase el ejercicio 2.30), lo cual indica que la pseudoprimalidad en base 2 no es una condición suficiente de primalidad.

En general, cuando un número compuesto m satisface $a^{m-1} \equiv 1$ con $\mathrm{mcd}(a, m) = 1$, se dice que m es *pseudoprimo* en base a.

Volviendo al caso de $m = 341$, si en vez de $a = 2$ tomamos $a = 7$ obtenemos $7^{340} \not\equiv 1 \pmod{341}$ (es decir, no es pseudoprimo en base 7), luego eso sí hubiera detectado que 341 es un número compuesto. Surge ahora una pregunta evidente: ¿cualquier número compuesto se puede detectar de esta manera? Más aún, si m es compuesto, ¿cómo saber cuál va a ser el correspondiente a que sirve para probarlo?

En la práctica, si uno hace experimentos (con la ayuda de un ordenador) se da cuenta de que la mayoría de los números compuestos son fáciles de detectar viendo que no son pseudoprimos para algún a; por ejemplo, probando sólo con $a = 2$ y 3. Esto hace concebir esperanzas de que los métodos basados en pseudoprimalidad sirvan realmente para encontrar números primos.

2.8.1. Números de Carmichael

La respuesta —quizás sorprendente— a las preguntas que planteábamos es que existen números compuestos que son pseudoprimos en todas las bases, y que por tanto los tests basados en pseudoprimalidad no van a permitir asegurar la primalidad. A los números m que son pseudoprimos en base a para cualquier base a (con $\mathrm{mcd}(a, m) = 1$) los llamamos *números de Carmichael*.

El primero de ellos es 561. Éste es, en efecto, un número compuesto: $561 = 3 \cdot 11 \cdot 17$. Veamos que, sin embargo, verifica $a^{561-1} \equiv 1 \pmod{561}$ para cualquier a con $\mathrm{mcd}(a, 561) = 1$. En primer lugar, observemos que, si $\mathrm{mcd}(a, 561) = 1$, también se cumplirá $\mathrm{mcd}(a, 3) = \mathrm{mcd}(a, 11) = \mathrm{mcd}(a, 17) = 1$ (es más: son hechos equivalentes, aunque eso no lo necesitamos en nuestra argumentación). Entonces,

el teorema pequeño de Fermat nos asegura que

$$a^2 \equiv 1 \pmod 3, \quad a^{10} \equiv 1 \pmod{11}, \quad a^{16} \equiv 1 \pmod{17}.$$

Como consecuencia,

$$\begin{aligned} a^{560} = (a^2)^{280} &\equiv 1 \pmod 3, \\ a^{560} = (a^{10})^{56} &\equiv 1 \pmod{11}, \\ a^{560} = (a^{16})^{35} &\equiv 1 \pmod{17}. \end{aligned}$$

Así pues, $a^{560} - 1$ es divisible por 3, 11 y 17, luego también por su producto, 561. Es decir,

$$a^{560} \equiv 1 \pmod{561},$$

tal como buscábamos.

Nota 2.24. Antes de continuar, comentemos por qué ha funcionado el proceso anterior que nos ha permitido probar que $561 = 3{\cdot}11{\cdot}17$ es un número de Carmichael. Si nos fijamos, lo único que hemos aplicado es que $561 - 1$ es múltiplo de $3 - 1$, $11 - 1$ y $17 - 1$. Así, también sería de Carmichael cualquier entero m de la forma $m = p_1 p_2 \cdots p_r$ con $m - 1$ múltiplo de $p_1 - 1$, $p_2 - 1$,..., $p_r - 1$. De hecho, esto caracteriza a los números de Carmichael, pues se cumple el siguiente resultado que no probaremos, y que se conoce con el nombre de criterio de Korselt: Un entero positivo n es un número de Carmichael si y sólo si n es libre de cuadrados y satisface $(p-1)\,|\,(n-1)$ para cada p divisor primo de n. (La demostración puede consultarse en [60] o [88].)

Cuando Alwin Korselt probó en 1899 el criterio que ahora lleva su nombre, lo hizo como respuesta a la pregunta de caracterizar los números n que satisfacían la propiedad $a^{n-1} \equiv 1 \pmod n$ siempre que $\mcd(a, n) = 1$, y que había sido propuesta como problema en una revista matemática. Pero Korselt no exhibía ningún ejemplo, y probablemente pensaba que no iba a existir ningún número con esas propiedades. Fue Robert D. Carmichael en 1910 quien encontró el primer ejemplo, el número 561. Además de 561, los números de Carmichael menores que diez mil son éstos:

$$\begin{aligned} 1105 &= 5 \cdot 13 \cdot 17, & 1729 &= 7 \cdot 13 \cdot 19, & 2465 &= 5 \cdot 17 \cdot 29, \\ 2821 &= 7 \cdot 13 \cdot 31, & 6601 &= 7 \cdot 23 \cdot 41, & 8911 &= 7 \cdot 19 \cdot 67, \end{aligned}$$

todos ellos descubiertos por él en 1912. El menor número de Carmichael con cuatro factores primos es $41\,041 = 7 \cdot 11 \cdot 13 \cdot 41$. Parece claro que Korselt no dedicó demasiados esfuerzos a intentar encontrar ejemplo alguno. ◁

Si sólo hubiera un número finito de números de Carmichael, y los conociéramos todos, de nuevo los test de primalidad basados en el teorema pequeño de Fermat podrían ser realmente útiles. Pero, aunque no es fácil probarlo, existen infinitos números de Carmichael, y además tampoco es sencillo identificarlos. Así que nos quedamos con lo que teníamos al principio, que no es poco: los tests basados en el cálculo de a^{m-1} (mód m) son útiles muchas veces para comprobar que m es compuesto, pero nunca pueden servir para asegurar que m es primo (por eso se denominan tests de primalidad no seguros o tests de seguridad probabilísticos). Si tenemos un número m del que queremos saber si es primo o compuesto, y no logramos encontrar una base en la que no sea pseudoprimo, forzosamente hay que recurrir a otros métodos que permitan dilucidar si m es primo o compuesto. Tales métodos existen (en particular, hallar la descomposición en factores es uno) pero, cuando m es grande, son mucho más costosos en cuanto al tiempo de cálculo que requieren; en consecuencia, merece la pena intentar primero descartar la primalidad de m intentando ver que no es pseudoprimo para algún a.

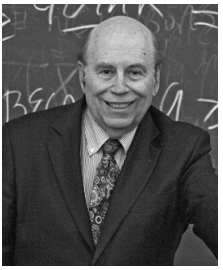
Sin embargo, hay que decir que, en ocasiones, los tests no seguros se emplean como si realmente sirvieran para probar si un número m es primo; los hay mejores que el que acabamos de exponer (en el apartado siguiente veremos uno), pero para explicar lo que ahora nos interesa podemos pensar que estamos usando el que ya tenemos. Primero se mira si m es un número de Carmichael de los ya conocidos (y tabulados). En segundo lugar, se comprueba si m es pseudoprimo en base a para bastantes —no precisaremos a cuántos nos referimos— valores de a. Si así ocurre, es bastante probable que m sea primo (tanto más cuanto más valores de a hayamos usado). Por supuesto, esto no es una demostración rigurosa de primalidad. Pero, en algunas circunstancias en las que se quiere averiguar la primalidad de m, es lo único que se tiene, pues todos los métodos rigurosos conocidos de demostrar primalidad son inaplicables con m grande. Para ciertos usos prácticos, esta no seguridad matemática puede no tener excesiva importancia. Sin demasiados esfuerzos de cálculo, la probabilidad de que se esté tomando como primo un número que realmente no lo sea es mucho menor que cualquier otra posible fuente de error involucrada en el proceso en el que se va a usar tal número. Un matemático nunca estará satisfecho con esto, pero cuando no hay otra cosa mejor, tampoco está mal contentarse con lo que uno tiene...

2.8.2. Test de Miller-Rabin

Para compensar lo pesimistas que nos hemos podido poner, comencemos anunciando que el test de primalidad basado en pseudoprimos se puede mejorar de manera considerable respecto a como lo hemos expuesto hasta ahora, y además no es



Robert Daniel Carmichael (1879–1967)



Michael Oser Rabin (1931–)

difícil ver el porqué. Recordemos una propiedad importante de los números primos (y que ya hemos probado en el lema 2.21): si p es primo, la ecuación en congruencias $x^2 \equiv 1 \pmod{p}$ sólo tiene (módulo p) las soluciones $x = 1$ o $x = -1$.

Estamos utilizando que dado cualquier a tal que $\mcd(a, m) = 1$, cuando m es primo, se cumple

$$a^{m-1} \equiv 1 \pmod{m}.$$

El caso $m = 2$ no tiene interés (y el resto de los números pares no son primos), o sea que asumamos a partir de ahora que m es impar. Entonces, podemos escribir la pseudoprimalidad en base a como

$$(a^{(m-1)/2})^2 \equiv 1 \pmod{m}.$$

Si m es primo, esto implicará (de hecho, son condiciones equivalentes) que

$$a^{(m-1)/2} \equiv \pm 1 \pmod{m}. \tag{2.19}$$

Para m genérico (no necesariamente primo), ésta es una condición más fuerte que $a^{m-1} \equiv 1 \pmod{m}$, así que parece más difícil que un número compuesto «engañe» a este nuevo test. (Además, supone un ahorro de cálculo, pues evaluar $a^{(m-1)/2}$ requiere menos trabajo que llegar a a^{m-1}.)

Lo cierto es que esta nueva condición mejora sustancialmente lo que teníamos hasta ahora: el número de Carmichael $m = 561$ no pasa el test (2.19) con $a = 5$, lo cual prueba que 561 es compuesto. (Véase el ejercicio 2.31.)

Tal como hizo Gary L. Miller en 1975 —y Michel O. Rabin precisó poco más tarde, dándole la forma probabilística actual que veremos más adelante—, la idea se puede llevar aún más lejos, y conseguir así una prueba aún más exigente. Como

$m > 2$ es impar, podremos poner $m-1 = 2^s t$ con $s \geq 1$ entero y $t \geq 1$ entero impar. Si $s = 1$, ya no tiene sentido hablar de $a^{(m-1)/4}$, y no podemos escribir $a^{(m-1)/2}$ como $(a^{(m-1)/4})^2$; esto nos impide seguir el proceso.

Pero supongamos que m y a satisfacen (2.19), y que $s \geq 2$. Si lo que hemos obtenido es $a^{(m-1)/2} \equiv -1 \pmod{m}$, tampoco seguimos profundizando. En cambio, si teníamos $a^{(m-1)/2} \equiv 1 \pmod{m}$, podemos volver a acordarnos de la ecuación $x^2 \equiv 1 \pmod{m}$ y escribir eso como

$$(a^{(m-1)/4})^2 \equiv 1 \pmod{m};$$

si m es primo, forzosamente deberá ser

$$a^{(m-1)/4} \equiv \pm 1 \pmod{m}. \tag{2.20}$$

Supongamos ahora que m y a satisfacen (2.20), y que $s \geq 3$. En el «caso -1», dejamos de insistir. En el caso «caso $+1$», de nuevo podemos escribir

$$(a^{(m-1)/8})^2 \equiv 1 \pmod{m},$$

y si m es primo se deberá cumplir

$$a^{(m-1)/8} \equiv \pm 1 \pmod{m}.$$

Podemos seguir con esto mientras vayamos obteniendo $a^{(m-1)/2^j} \equiv 1 \pmod{m}$, $j = 1, 2, \ldots, s-1$. Si con $j = s-1$ continúa siendo

$$a^{2t} = a^{(m-1)/2^{s-1}} \equiv 1 \pmod{m},$$

de ahí obtenemos que, si m es primo,

$$a^t \equiv \pm 1 \pmod{m}.$$

Como t es impar, ya no podemos proseguir.

Aún conviene hacer una observación adicional sobre los posibles casos

$$a^{(m-1)/2^j} \equiv 1 \pmod{m} \quad \text{para } j = 1, 2, \ldots, s-1$$

(es decir, antes de llegar al último paso $a^t \equiv \pm 1 \pmod{m}$, que se correspondería con $j = s$). Si eso ocurriera para alguno de tales j, nos quedamos con el mayor de ellos, y lo expresamos como $m \,|\, (a^{(m-1)/2^j} - 1)$. Entonces,

$$m \,|\, (a^{(m-1)/2^j} - 1) = (a^{(m-1)/2^{j+1}} + 1)(a^{(m-1)/2^{j+1}} - 1).$$

Si m fuese primo, debería dividir a alguno de los factores de la derecha, así que $a^{(m-1)/2^{j+1}} \equiv \pm 1 \pmod{m}$. Si $j = s-1$, hemos llegado a $a^t \equiv \pm 1 \pmod{m}$; si $j < s-1$, la posibilidad «+1» está descartada por la maximalidad de j, luego forzosamente $a^{(m-1)/2^{j+1}} \equiv -1 \pmod{m}$, que es otro de los casos ya contemplados. Así pues, salvo para $a^t \equiv \pm 1 \pmod{m}$, no hace falta preocuparse de los $a^{(m-1)/2^j} \equiv 1 \pmod{m}$.

Recopilemos todo este proceso con un pequeño cambio de notación: vamos a usar $k = s - j$. Decimos que un número impar m, con $m - 1 = 2^s t$ para $s \geq 1$ entero y $t \geq 1$ entero impar, pasa el *test de Miller-Rabin*[14] en base a (con $\mcd(a, m) = 1$) si

- $a^t \equiv 1 \pmod{m}$, o
- $a^{2^k t} \equiv -1 \pmod{m}$ para algún k con $0 \leq k \leq s-1$.

Si m es un primo impar, seguro que pasa el test de Miller-Rabin en base a para cualquier a con $\mcd(a, m) = 1$.

Si m es un número impar y queremos saber si es primo, lo sometemos al test. Para ahorrar trabajo, lo hacemos comenzando por los exponentes más bajos (en cuanto un exponente detecte que m pasa el test, no hace falta proseguir). Primero comprobamos si $a^t \equiv \pm 1 \pmod{m}$; si se cumple, el número ha pasado el test, y por tanto podría ser primo. Si no se cumple, vamos probando los sucesivos $a^{2^k t}$ desde $k = 1$ en adelante a ver si alguno cumple $a^{2^k t} \equiv -1 \pmod{m}$ (eso sólo requiere elevar al cuadrado el anterior); si ocurre, en ese momento el número acaba de pasar el test y nos detenemos. Si llegamos al final y volvemos a obtener $a^{2^{s-1} t} \not\equiv -1 \pmod{m}$, el número m no ha pasado el test, y podemos estar seguros de que no es primo. (Nótese además que, si en algún paso, menos en el $k = 0$, se obtiene $a^{2^k t} \equiv 1 \pmod{m}$, eso ya asegura que no se va pasar el test en los sucesivos k, y no hace falta continuar: ya podemos asegurar que m es compuesto.)

Como vemos en el ejemplo que sigue, de hecho existen números compuestos que «engañan» al test. En esas circunstancias, cuando un número compuesto m pasa el test de Miller-Rabin en base a (con $\mcd(a, m) = 1$), decimos que m es *pseudoprimo fuerte* en base a.

Ejemplo. Comprobar que $2047 = 23 \cdot 89$ es pseudoprimo fuerte en base 2.

Solución. Como $2047 - 1 = 2 \cdot 1023$, el test de Miller-Rabin se reduce a analizar (2.19). Para evaluar 2^{1023}, en lugar de utilizar el cálculo de potencias mediante cuadrados sucesivos, vamos a usar un método específico más rápido. Por ser

[14] En otros libros lo denominan test de Rabin-Miller, test de Miller o test de Rabin.

$1023 = 11 \cdot 93$, de ahí se sigue que

$$2^{1023} = (2^{11})^{93} = 2048^{93} \equiv 1^{93} \equiv 1 \pmod{2047},$$

tal como buscábamos. ◊

Aunque los pseudoprimos fuertes son mucho más raros que los pseudoprimos, sigue habiendo infinitos de ellos. Eso sí, esta vez no hay pseudoprimos fuertes en todas las bases (es decir, no hay nada análogo a los números de Carmichael). Esto es consecuencia del siguiente resultado que no probaremos (la demostración se puede encontrar, por ejemplo, en [88]):

Teorema 2.25. Rabin, 1976.
Si un entero impar m es un número compuesto, entonces m pasa el test de Miller-Rabin para, a lo más, $(m-1)/4$ bases b con $2 \leq b \leq m-1$.

Este teorema garantiza que, aplicando el test de Miller-Rabin el suficiente número de veces, conseguimos un test de primalidad seguro. Pero aplicarlo $(m-1)/4$ veces para analizar la primalidad de m es demasiado trabajo, mucho más incluso que intentar factorizar m (que sólo requiere buscar factores menores que $\sqrt{m}$).

En la práctica, se usa de manera probabilística. Dado un entero positivo (e impar) m, elegimos al azar k enteros distintos, todos positivos menores que m. Si m es compuesto, la probabilidad de que pase el test de Miller-Rabin en cada una de esas k bases es menor que $(1/4)^k$. Por ejemplo, con $k = 50$ esa probabilidad es menor que 10^{-30} (y, con $k = 100$, menor que 10^{-60}). Tales números son extremadamente pequeños; de hecho, es bastante más probable que el ordenador que estemos usando tenga algún tipo de error. Cuando un número m pasa el suficiente número de tests, no es seguro que m sea primo, pero «casi». Todo esto hace que éste sea uno de los tests de primalidad probabilísticos más utilizados (en el tema siguiente, concretamente en la sección 3.3, veremos otro de tales tests, el denominado test de Solovay-Strassen).

Es más, se piensa que, si m es compuesto, seguro que existe algún $b < 2(\log_2(n))^2$ (con $\log_2$ estamos denotando el logaritmo en base 2) tal que m no pasaría el test de Miller-Rabin en base b. Tal conjetura no está demostrada, pero Miller probó que se podía deducir de la denominada hipótesis de Riemann generalizada (una famosa conjetura en teoría analítica de números que nosotros no vamos a ver, pero que, como su nombre indica, está relacionada con la hipótesis de Riemann, que sí comentaremos en los últimos temas de este libro). En tales circunstancias, el test de Miller-Rabin se convertiría en un test de primalidad seguro muy eficiente, pues

podría detectar la primalidad de un número en un tiempo que, a lo más, crece como un polinomio (de grado 5) con el número de dígitos.

2.9. Criptografía

Comencemos aclarando que se denomina criptografía a la ciencia que trata de escribir mensajes de manera que sólo puedan ser descifrados por sus destinatarios legítimos. El criptoanálisis es el conjunto de técnicas destinadas a esclarecer mensajes cifrados sin poseer la clave. En ocasiones, se utiliza criptología para englobar la criptografía y el criptoanálisis.

A lo largo de la historia, cifrar mensajes para mantener su contenido oculto siempre ha tenido una gran importancia. Por supuesto, esos mensajes los tiene que poder interpretar el destinatario, pero no otras personas no autorizadas. Donde más importancia tiene esto es en el terreno militar (al menos, hasta que se han popularizado las redes informáticas y el comercio electrónico). Un general quiere comunicarse con los oficiales a sus órdenes, que dirigen otros flancos de su ejército, pero no desea que el enemigo conozca sus planes. Así, envía mensajeros con el texto cifrado mediante algún tipo de clave; aunque caigan prisioneros, el enemigo —que no conoce la clave— no deberá ser capaz de interpretar el mensaje. Naturalmente, los ejércitos contrarios tienen el objetivo opuesto: descifrar el mensaje. Son los criptoanalistas los que se deben encargar de esta tarea, empleando para ello su habilidad y toda clase de trucos. La historia nos enseña que, cuando el número de textos ocultos a los que los criptoanalistas han tenido acceso es suficientemente elevado, ingeniosas técnicas —y muchos esfuerzos— han permitido averiguar la clave. A partir de entonces, tal clave pierde toda su efectividad. En consecuencia, es conveniente cambiar la clave de vez en cuando.

Uno de los puntos débiles de la mayoría de los sistemas criptográficos es mantener la clave oculta. Si el enemigo la conoce, descifrará todos los mensajes sin ninguna dificultad. Lo habitual es que, para descifrar, haya que usar la misma clave que se ha usado para cifrar. En consecuencia, el general tiene que transmitir esa clave a los mandos de todos los flancos con los que desea comunicarse; y renovarla cada cierto tiempo. Esto es muy poco seguro. En el último apartado de esta sección veremos cómo los sistemas de clave pública solucionaron —en la década de 1970— el casi eterno problema de la distribución de claves.

Comencemos dando un ejemplo de un criptosistema que ya era usado por Julio César para cifrar mensajes, y que por esa razón se suele denominar *cifra del César*. Se basa en un simple desplazamiento entre las letras. En concreto, para cifrar un

mensaje sustituyamos cada letra por la correspondiente que aparece en la segunda fila de esta tabla:

```
A B C D E F G H I J K L M N Ñ O P Q R S T U V W X Y Z
D E F G H I J K L M N Ñ O P Q R S T U V W X Y Z A B C
```

De este modo, la frase

```
ESTE MENSAJE ES SECRETO
```

se codificaría como

```
HVWH OHPVDMH HV VHFUHWR
```

No hace falta que aclaremos lo que hay que hacer para descifrar esto.

Aquí parece que no hay matemáticas; pero sí las hay. Asociamos cada letra de la `A` a la `Z` con un equivalente numérico de 1 a 27, manteniendo el orden (es decir, `A` es 1, `B` es 2, etcétera). En el ejemplo mostrado, cifrar el mensaje es sumar 3 módulo 27. Y descifrarlo es restar 3 módulo 27. El 3 es la clave.

Este sistema de cifrado tiene tres defectos muy grandes:

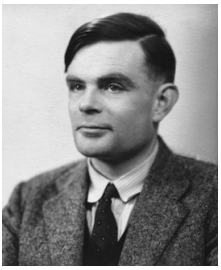
- Con saber cómo descifrar una sola letra, se descifra todo el código.
- Sólo hay 27 maneras de codificar; es decir, 27 posibles claves distintas. Quien quiera efectuar un criptoanálisis de un mensaje cifrado, le basta probar con todas. Casi seguro que sólo una producirá un texto con sentido.
- Que los espacios conserven su sitio da un pista muy importante (existen muy pocas palabras de una sola letra).

El tercer problema desaparece si suprimimos los espacios (o si hacemos que el espacio se comporte como una letra más). Los otros dos defectos se podrían solucionar si la asociación entre las letras originales y la cifradas no fuese un mero desplazamiento, sino una permutación. Esto consigue que, ahora, el número de claves diferentes sea 27!, un número demasiado grande como para intentar probarlas todas en un tiempo razonable (si tuviéramos un dispositivo automático que analizara un millón de posibilidades por segundo, tardaría más de $3 \cdot 10^{14}$ años en chequear todas las claves). Sin embargo, este método no resiste un criptoanálisis basado en el porcentaje de veces que cada letra aparece en un idioma. ¿Cómo se soluciona eso? Ideando un método en el que cada letra no siempre se codifique igual. Por ejemplo, podríamos usar unas cuantas tablas e ir alternando entre ellas (de manera cíclica): la primera letra se codifica con la tabla uno, la segunda con la dos, etc. Esto dificultaría considerablemente la tarea de los criptoanalistas, que deberían pensar otro tipo de ataques.

He aquí la eterna lucha entre criptógrafos y criptoanalistas. Cuando los criptoanalistas averiguan cómo romper un sistema, los criptógrafos toman buena nota de



François Viète
(1540–1603)



Alan Mathison Turing
(1912–1954)



Marian Adam Rejewski
(1905–1980)

ello y diseñan uno nuevo que ya no tenga la debilidad detectada. A lo largo de la historia se han usado muchos sistemas criptográficos distintos, y su seguridad (o la falta de ella) ha resultado vital en numerosas ocasiones. De carácter divulgativo —contiene muy pocas matemáticas— y lectura muy amena es el libro [97], que proporciona una completa visión sobre la evolución de la criptografía.

Por cierto, es importante dejar sentado que la seguridad de un sistema no puede depender de que el método usado no sea conocido; sólo de su oposición a ser descifrado cuando quien quiere hacerlo conoce todos los detalles del método salvo la clave. Cuando una organización decide usar un sistema criptográfico, tiene que explicarlo a las personas que lo van a usar, o programarlo en ordenadores; quizás editar manuales o construir máquinas codificadoras... Demasiadas cosas como para pensar que todo ello se va a poder mantener oculto.

Nota 2.26. Aunque en este libro sólo vamos a hablar de criptografía, no de criptoanálisis, no queremos dejar de citar a dos importantes matemáticos que tuvieron gran éxito como criptoanalistas. El primero es el francés François Viète, que fue capaz de descifrar un sistema de codificación utilizado por el rey español Felipe II. El segundo es el inglés Alan Turing —el padre de la informática y de la computación moderna—, cuya contribución en la decodificación y ruptura de los códigos secretos alemanes durante la Segunda Guerra Mundial permitió, sin duda, acortar la guerra (para ser justos, no hay que olvidarse de mencionar que los éxitos criptográficos de Turing estuvieron basados, de manera muy importante, en los logros previos del matemático polaco Marian Rejewski, mucho menos conocido que el inglés). ◁

Matemáticamente, un texto destinado a ser cifrado se puede asimilar a un entero no negativo acotado por una constante prefijada. En el ejemplo de la cifra del César (o en su generalización mediante permutaciones), cada vez ciframos una sola letra, luego podemos considerar que el texto a cifrar es un número entero del intervalo

[1, 27]; y efectuar el cifrado tantas veces como sea necesario, hasta que acabemos el mensaje. Se consiguen números mayores —la utilidad de esto quedará de manifiesto más adelante— agrupando letras del mensaje. Por ejemplo, si agrupamos 4 letras cada vez, yuxtaponiendo sus correspondientes valores (desde 01 —obsérvese que lo forzamos a que tenga dos dígitos— hasta 27) obtendríamos números de 8 dígitos. También se podría hacer $\sum_{j=0}^{3} a_{j+1}27^j$, donde a_j es el valor de la j-ésima letra de cada grupo de cuatro. Además, no hay ninguna razón para que, en estos procesos de cifrado, sólo usemos letras: sin ninguna dificultad podíamos haber incluido espacios, signos de puntuación, caracteres acentuados, haber distinguido entre mayúsculas y minúsculas, etc. En todo caso, esto son pequeños detalles que carecen de importancia, y no merece la pena que nos detengamos en ellos.

Aquí, aparte del código conocido como cifra del César ya mostrado, sólo vamos a presentar un par de sistemas criptográficos. Por supuesto, existen muchos otros. Para estudiarlos desde una perspectiva matemática recomendamos el excelente texto [60]; otro interesante libro es [99] que, sobre todo, hace mucho hincapié en la descripción de los algoritmos que hay que usar. Además, ésta es un área muy activa, y los nuevos descubrimientos, mejoras, y todo tipo de novedades, son bastante habituales.

Nota 2.27. Además del problema de la inseguridad de las claves, los criptosistemas clásicos tienen otro inconveniente, que es la gran cantidad de claves necesarias cuando hay muchas partes involucradas. Esto puede no tener importancia en el terreno militar (sólo hay dos bandos), pero sí en aplicaciones comerciales y electrónicas. Imaginemos que en un país hay 100 bancos que se tienen que comunicar entre sí, y que cada par de bancos lo debe hacer de manera secreta, de modo que todos los demás bancos no se enteren de las transacciones particulares entre ellos dos (algo muy razonable). En estas circunstancias, cada banco necesitaría tener 99 claves, una por cada uno de los otros 99 bancos con los que se tiene que comunicar. En total, $\binom{100}{2} = 4950$ claves. Y, por supuesto, cada banco necesitaría compartir una clave particular con cada cliente, etcétera. No olvidemos, por último, que cada vez nos enfrentamos a la vulnerabilidad del intercambio de claves. ◁

2.9.1. Un criptosistema de clave privada basado en la exponenciación discreta

Una vez hechas las precisiones anteriores acerca de que sólo nos tenemos que preocupar de cómo codificar números enteros, vamos a dar un ejemplo de otro sistema criptográfico basado en congruencias. En concreto, mostraremos el inventado por Pohlig y Hellman a mediados de la década de 1970, y que requiere efectuar

exponenciaciones discretas[15] módulo un número primo. (En realidad, éste es un ejemplo puramente académico, y nunca ha tenido importancia criptográfica real.) Nótese que, al contrario que en otras partes del libro, e no va a ser aquí el habitual número $e = \sum_{n=0}^{\infty} 1/n! = \lim_n (1+1/n)^n$ (que ahora no va a aparecer), sino un número entero.

Supongamos que tenemos un primo impar p y un entero positivo e, de modo que $\mcd(e, p-1) = 1$. Estos dos números, e y p, son la clave del sistema. Aunque la parte importante es e; p se puede dar por conocido.

Es habitual llamar P al texto sin cifrar —pues en inglés se denomina *plaintext* ('texto llano' o 'texto claro')— y C al texto ya cifrado; nosotros seguiremos esta misma nomenclatura. Así pues, si tenemos un texto —ya trasladado a número— P que cumple $0 < P < p$, su correspondiente versión cifrada C será el menor entero positivo C tal que $C \equiv P^e \pmod{p}$. Esto es, cifrar consiste en aplicar la siguiente operación:

$$C \equiv P^e \pmod{p}, \qquad \text{con } 0 < C < p$$

(obsérvese que, como $P \not\equiv 0 \pmod{p}$ y p es primo, también $C \not\equiv 0 \pmod{p}$). Un receptor legítimo de C tiene que poder descifrarlo. Así pues, veamos cómo recuperar P a partir de C, supuesto que conocemos el primo p y la clave e.

Como $\mcd(e, p-1) = 1$, existe d tal que $ed \equiv 1 \pmod{p-1}$. Este d, que es único módulo $p-1$ (por ejemplo, único si exigimos que $0 < d < p-1$), puede calcularse usando el algoritmo de Euclides, tal como vimos en la sección 2.2 (la fórmula $d \equiv e^{\varphi(p-1)-1} \pmod{p-1}$ del teorema 2.8 proporciona un método mucho más lento). Veamos que, para obtener P, basta calcular C^d módulo p, y que

$$C^d \equiv P \pmod{p}.$$

En efecto. Como $ed \equiv 1 \pmod{p-1}$, se tendrá $ed = 1 + (p-1)k$ para algún $k \in \mathbb{Z}$. Además, el teorema pequeño de Fermat nos asegura que $P^{p-1} \equiv 1 \pmod{p}$. Entonces, $C^d \equiv P^{ed} \equiv P^{(p-1)k+1} \equiv (P^{p-1})^k P \equiv 1^k P \equiv P \pmod{p}$, tal como queríamos comprobar.

Este método de cifrado es bastante seguro, siempre que el criptoanalista que lo esté atacando no conozca e (no importa que el primo p sea público). Si los números p y e son grandes, efectuar P^e módulo p «revuelve» muy bien el texto inicial, pues —recordando la metáfora del reloj— eso produce muchas «vueltas» módulo p. Y, de todas esas vueltas, lo único que se usa es el resto, algo casi despreciable con el tamaño (como número en $\mathbb{Z}$) de P^e. Desde luego, esto no es una justificación matemática

[15] La palabra «discreta» se emplea para hacer hincapié en que ahora no estamos tratando con la clásica función exponencial en $\mathbb{R}$ o $\mathbb{C}$, que sería la exponenciación continua.

de la fiabilidad del método, pero sí que da una idea de por qué este sistema es muy resistente al criptoanálisis. Una explicación más técnica es que, aunque conozcamos P, C y p, no hay ningún procedimiento rápido de encontrar e tal que $C \equiv P^e$ (mód p); en otras palabras, que no hay métodos eficientes de calcular logaritmos discretos. Para que todo esto sea seguro frente a los mejores ataques conocidos actualmente —los cuales se hacen con ayuda de potentes ordenadores, claro— el primo p tiene que tener unos pocos cientos de dígitos; y ser elegido con cuidado para que $p-1$ no tenga únicamente factores primos pequeños (lo cual permitiría aplicar ciertas técnicas especializadas que no vamos a comentar). Por ejemplo, es muy útil usar un p que sea de la forma $p = 2q + 1$ con q otro primo; en estas circunstancias se dice que p es un primo 1-seguro.[16] Se pueden ver más detalles en [88, § 8.3].

2.9.2. Avances hacia la criptografía de clave pública

En todos los métodos en los que la clave para cifrar y descifrar coinciden, hay algo que escapa al control de los criptógrafos: mantener la clave oculta. El proceso de enviar la clave es muy inseguro, pues tal clave no puede estar cifrada (salvo por otra clave anterior, que tendría el mismo problema...).

Parecía que la distribución de claves era inevitable en cualquier sistema criptográfico; hasta principios de la década de 1970, era como si tal afirmación fuera un axioma de la criptografía. Sin embargo, Diffie y Hellman eran conscientes de una anécdota que parecía desafiar esto. Supongamos que Alicia y Benito viven en un país en el que los empleados postales tienen la fea costumbre de leer toda la correspondencia desprotegida que reparten. Si Alicia y Benito quiere comunicarse por correo sin que ningún cartero pueda leer sus mensajes, la solución clásica es encerrar el mensaje en una caja de acero y ponerle un candado del que ambos tienen llave; ahora, ya pueden enviar la caja por correo.

Pero existe una solución más imaginativa que no requiere que ambos compartan la llave, sino que cada uno va a tener una llave de la que no va a distribuir ninguna copia. Así, para lograr que Alicia envíe un mensaje confidencial a Benito, pueden proceder del siguiente modo:

- Alicia mete el mensaje en una caja, la cierra con un candado y la envía por correo.
- Benito, cuando recibe la caja, le añade otro candado, y la devuelve por correo a Alicia.

[16] Aunque esto ya no tiene nada que ver con la criptografía, los primos q tal que $2q+1$ es primo también tienen nombre: son los primos de Germain que ya nos habían aparecido en el tema 1 (apartado 1.6.3) en relación con la ecuación diofántica $x^n + y^n = z^n$ y la búsqueda de una prueba del último teorema de Fermat.



De izquierda a derecha, Ralph C. Merkle (1952–), Martin E. Hellman (1945–) y Whitfield Diffie (1944–)

- Alicia recibe la caja con los dos candados; retira el suyo y la vuelve a enviar a Benito.
- Finalmente, Benito abre la caja y lee el mensaje.

No ha habido ningún intercambio de llaves; y, sin embargo, nadie más ha podido ver el contenido de la caja.

Adaptar esto a la vida real —y que pudiera ser aprovechado para hacer seguras las comunicaciones en las entonces recientes redes de ordenadores— no era tarea fácil. Diffie y Hellman —a los que se unió Merkle— llegaron a la conclusión de que requerían utilizar funciones matemáticas que, tras aplicarlas dos veces, el resultado obtenido fuese independiente del orden. Y que, además, fuesen difíciles de invertir; las llamaron *funciones de una sola vía* (o *funciones unidireccionales*). En 1976, tras muchos esfuerzos, lograron dar con un procedimiento mediante el que Alicia y Benito podían generar una clave privada común mediante el intercambio —por un canal inseguro— de información que a nadie más le iba a ser útil para obtener tal clave. Tal procedimiento se conoce habitualmente como protocolo de Diffie-Hellman o protocolo de Diffie-Hellman-Merkle.

Se eligen un primo p suficientemente grande y un entero positivo r (menor que p) que sea lo que denominamos *raíz primitiva* de p, esto es, que la congruencia $r^m \equiv 1$ (mód p) no se cumpla para ningún entero positivo m menor que $\varphi(p) = p - 1$ (el teorema pequeño de Fermat asegura que $r^{p-1} \equiv 1$ (mód p), pero no dice nada de si eso se puede lograr o no con otros exponentes).[17] Los números p y r no necesitan ser secretos. Ahora, Alicia escoge —al azar— un entero grande a; y Benito otro, b (ambos menores que $p - 1$). No comunican a nadie tales números. Alicia calcula $A = r^a$ (mód p) (con este «=» en lugar de «≡» queremos indicar que A es precisamente el resto de r^a al dividir por p), y Benito calcula $B = r^b$ (mód p); obsérvese que $A \neq 1 \neq B$, ya que r es una raíz primitiva de p. Ambos se intercambian A y B de cualquier forma, por muy insegura que ésta sea. Entonces, tanto Alicia como Benito pueden calcular una clave privada común $K = r^{ab}$ (mód p); Alicia lo hace mediante $K = B^a \equiv r^{ab}$ (mód p), y Benito con $K = A^b \equiv r^{ab}$ (mód p). En resumen, ambos han logrado encontrar una clave común K que nadie más conoce.

Si un espía ha interceptado todas la comunicaciones, habrá descubierto p, r, A y B, pero no a ni b. Entonces, para hallar K tiene que intentar calcular a a partir de $A = r^a$ (mód p) (o análogamente con b y B). La seguridad de este método criptográfico se basa en que no hay métodos eficientes de hacerlo: la exponenciación discreta es sencilla de calcular, pero la función inversa —el logaritmo discreto— no lo es.

Nota 2.28. Aquí no lo hemos usado aún, pero no está de más decir que el estudio de raíces primitivas ha dado lugar a numerosos resultados interesantes e importantes aplicaciones; véase, por ejemplo, [5, capítulo 10] o [88, capítulo 9]. En este libro, dedicaremos a las raíces primitivas la sección 3.5, en el tema 3. Además, dominar las técnicas involucradas en ello es fundamental para intentar atacar sistemas criptográficos basados en la exponenciación discreta (como el de Pohlig y Hellman que hemos visto en el apartado 2.9.1), o para prevenir tales ataques mediante una adecuada elección de parámetros y claves. ◁

2.9.3. El criptosistema RSA

El protocolo de Diffie-Hellman era un avance importantísimo que solucionaba el problema de la vulnerabilidad de la distribución de claves, y que inauguraba la era de la *criptografía de clave pública*. Pero aún tenía la dificultad de que Alicia y Benito tenían que ponerse previamente de acuerdo en que se iban a comunicar,

[17] Veremos mucho más sobre raíces primitivas en el tema siguiente (en concreto, a ellas dedicamos la sección 3.5). De momento, nos podemos contentar con asegurar al lector que existen, aunque en general no sean fáciles de encontrar.

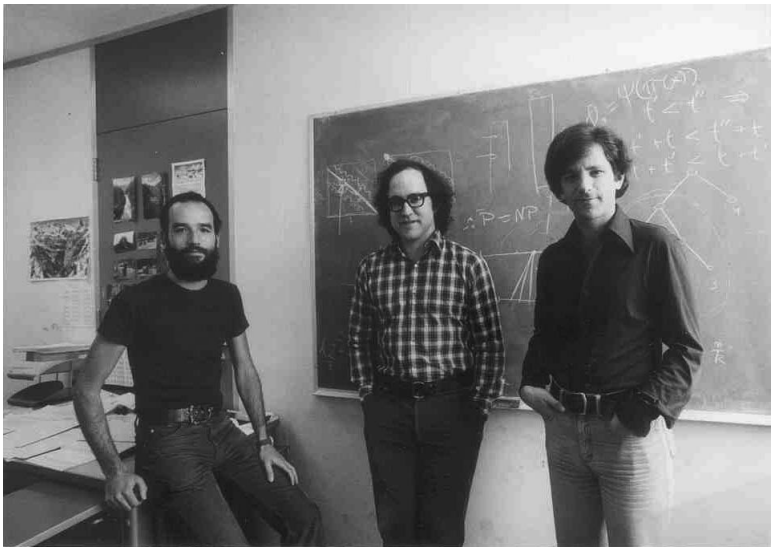
y enviarse —ambos— A y B. Uno de ellos no podía mandar un mensaje seguro al otro hasta haberse intercambiado esos dos valores (aunque sea por un canal inseguro). Además, el protocolo de Diffie-Hellman no soluciona el problema de la gran cantidad de claves necesarias cuando son bastantes los interlocutores que se tienen que intercambiar mensajes cifrados (relatado en la nota 2.27). Es cierto que eso es un problema menor comparado con el de la inseguridad de las claves, pero no deja de ser un inconveniente.

En su afán de encontrar un proceso más cómodo, Diffie y Hellman adoptaron un enfoque distinto. Se trataba de intentar conseguir claves asimétricas, esto es, que las claves para cifrar y descifrar fuesen distintas.[18] Las claves para cifrar serían públicas, y las podría conocer cualquiera (sólo sirven para cifrar mensajes, no para descifrarlos); las privadas, sólo el receptor del mensaje. Supongamos que Alicia quiere enviar un mensaje cifrado a Benito. Para ello, Alicia lo codifica con la clave de cifrar de Benito, y se lo envía —y queremos recalcar que cualquier otra persona haría igual: Benito no necesita una clave para cada interlocutor—. Por supuesto, Benito no ha comunicado a nadie su clave para descifrar; así, cuando reciba el mensaje, sólo él podrá descifrarlo, y no hay peligro de que sea interceptado por el camino. Naturalmente, la seguridad de todo esto radica en que de la clave para cifrar no se pueda deducir fácilmente la clave para descifrar. La idea parecía prometedora, si se lograba poner en práctica. Diffie y Hellman no encontraron ningún procedimiento que se ajustase a ello.

En 1978, Rivest, Shamir y Adleman concibieron un criptosistema que funcionaba de la manera requerida. Su método —conocido como RSA— es similar al de Pohlig y Hellman explicado en el apartado 2.9.1, pero con una diferencia fundamental. Supongamos que tenemos dos primos distintos p y q suficientemente grandes, y tomamos $m = pq$. Para cifrar, se usan congruencias módulo m, y no hace falta conocer los factores. En cambio, para descifrar hay que usar p y q. Así, m se puede hacer público, siempre que p y q se mantengan en secreto.

A simple vista, parece que esto tiene un fallo: si conocemos m, lo descomponemos en factores y averiguamos p y q. ¿Dónde está, pues, la seguridad de esto? En que hallar la descomposición en factores de un número es un problema que —matemáticamente hablando— seguro que tiene solución, pero que en la práctica es intratable. Si p y q tienen, digamos, algo más de cien dígitos, los mejores algoritmos de fac-

[18] Queremos insistir en esto. Ambas claves no sólo deben ser distintas, sino que de una no se pueda deducir la otra. Obsérvese que en el método de Pohlig y Hellman que hemos explicado en el apartado 2.9.1 la clave para cifrar es e, y la de descifrar d; pero e y d están ligadas por la relación $ed \equiv 1 \pmod{p-1}$, así que se pasa fácilmente de una a otra. Podemos decir que, esencialmente, son la misma clave.



De izquierda a derecha, Adi Shamir (1952–), Ronald Linn Rivest (1947–) y Leonard Adleman (1945–)

torización conocidos, ejecutados en todos los ordenadores del mundo trabajando conjuntamente, tardarían siglos en encontrar p y q a partir de su producto.

Pero dejémonos ya de preliminares y veamos cómo funciona el criptosistema RSA. Benito, que quiere recibir mensajes cifrados, elige p y q dos primos grandes distintos, calcula su producto $m = pq$ y la indicatriz de Euler $\varphi(m) = (p-1)(q-1)$; escoge, además, otro entero positivo e tal que $\mathrm{mcd}(\varphi(m), e) = 1$. Entonces, hace que m y e —su clave pública— sean accesibles a cualquiera que desee conocerlos. Los dos factores p y q y el valor de $\varphi(m)$, que no comunica a nadie, constituyen la parte privada de su clave.

Supongamos que Alicia quiere enviar un mensaje secreto a Benito; el texto del mensaje —o una parte de él, si es demasiado grande— estará representado mediante un entero P que verifica $0 < P < m$. Los dos únicos números menores que m que no son primos con m son p y q; como p, q y m son muy grandes, podemos dar por prácticamente seguro que P será primo con m (lo contrario sería equivalente a encontrar un factor no trivial de m probando a dividirlo por un número al azar). El texto cifrado es el único entero C que cumple

$$C \equiv P^e \pmod{m}, \qquad \text{con } 0 < C < m, \tag{2.21}$$

que se encuentra sin dificultad utilizando el método de cuadrados sucesivos (tal como hemos visto en los ejemplos de las páginas 101 y 126).[19] Entonces, Alicia envía C a Benito. ¿Cómo hace este último para recuperar P? Como $\mcd(\varphi(m), e) = 1$, seguro que existe un entero d que satisface

$$ed \equiv 1 \pmod{\varphi(m)}; \tag{2.22}$$

además, d es fácil de encontrar mediante el algoritmo de Euclides. Con esto, Benito sólo tiene que calcular C^d módulo m ya que —según veremos a continuación— se cumple

$$C^d \equiv P \pmod{m}. \tag{2.23}$$

En efecto, ed se podrá escribir como $ed = 1 + \varphi(m)k$ para algún $k \in \mathbb{Z}$. Además, el teorema de Euler-Fermat asegura que $P^{\varphi(m)} \equiv 1 \pmod{m}$ (en realidad, basta usar el teorema de la media de Fermat, ya que $m = pq$ es un producto de dos primos). Entonces,

$$C^d \equiv P^{ed} \equiv P^{\varphi(m)k+1} \equiv (P^{\varphi(m)})^k P \equiv P \pmod{m},$$

tal como queríamos comprobar.

Obsérvese que para descifrar C y encontrar P hemos usado d y m. Un espía que capture C y quiera descifrarlo se encuentra con la dificultad de que no conoce d. En consecuencia, tiene que intentar averiguarlo. Hallar el d que satisface (2.22) sería factible mediante el algoritmo de Euclides; ¡pero no conoce $\varphi(m)$, así que no puede aplicarlo!

Sí que tiene a su disposición m; y por tanto puede calcular $\varphi(m)$ mediante la fórmula del teorema 2.10. Para eso tiene que encontrar la descomposición en factores de m. Pero m está elegido de manera que esto sea extremadamente laborioso; en concreto, $m = pq$ con p y q dos primos muy grandes. Benito, que conoce p y q, puede calcular $\varphi(m) = (p-1)(q-1)$ sin dificultad (y también e). Pero —insistimos— si p y q son muy grandes, ningún espía será capaz de encontrar la descomposición $m = pq$ en un tiempo razonable.

Nota 2.29. Hemos narrado la historia de la criptografía de clave pública tal como se desarrolló por medio de la «literatura abierta». Pero la criptografía es algo en lo que los gobiernos y sus servicios secretos están muy interesados y que, por su naturaleza, desean mantener oculto a la luz de otros gobiernos. En consecuencia, los resultados que se obtienen no se divulgan hasta muchos años más tarde.

[19] En la práctica, es común tomar $e = 2^{16} + 1 = 65\,537$, que en base 2 es $10\,000\,000\,000\,000\,001_{(2}$ y resulta especialmente cómodo para calcular P^e por medio de potencias sucesivas. Nótese que e es parte de la clave pública y desvelarlo no ocasiona ningún problema.

Documentos hechos públicos por el gobierno británico en 1997 prueban que la criptografía de clave pública había sido previamente desarrollada por matemáticos que trabajaban para sus servicios secretos. En concreto, el correspondiente proceso comienza en 1969, cuando James Ellis concibe la idea de las claves asimétricas. En 1973, Clifford Cocks se plantea la posibilidad de utilizar números primos y la factorización como base del sistema, y descubre un sistema criptográfico prácticamente idéntico al RSA de Rivest, Shamir y Adleman. Finalmente, Malcolm Williamson describe en 1974 un sistema de intercambio de claves como el que más tarde publicarían Diffie y Hellman. Curiosamente, los criptógrafos del servicio secreto británico descubrieron el funcionamiento de un algoritmo como el RSA antes que el intercambio de claves mediante el protocolo de Diffie-Hellman, mientras que «en el mundo exterior» fue al revés. ◁

2.9.4. Ataques al criptosistema RSA

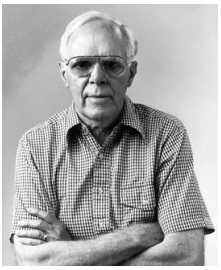
Es importante darse cuenta de que la seguridad del criptosistema RSA depende de la potencia de los algoritmos de factorización conocidos. Si alguien, en el futuro, descubriera métodos de factorización mucho más rápidos que los mejores actualmente disponibles[20] comprometería la seguridad del método.

Para intentar evitar sorpresas, los que se ocupan de técnicas criptográficas están muy interesados en conocer los avances en los conocimientos sobre factorización de enteros. Ésta es la razón por la cual algunos organismos desafían a la comunidad matemática mediante recompensas monetarias a los primeros que logren hallar la descomposición en factores de ciertos números enormes. Varios de estos premios han sido ya cobrados por grupos que, contando con la ayuda de potentes ordenadores y complejos algoritmos, han unido sus esfuerzos computaciones a través de internet.[21] Desde luego, esto ha dado lugar a nuevos desafíos con números aún mayores.

Nota 2.30. En realidad, ni siquiera se sabe si existe o no un algoritmo de factorización cuyo tiempo de respuesta esté acotado por una función polinomial del

[20] En este libro no nos hemos ocupado de ello, pero hallar buenos algoritmos de factorización es uno de los objetivos más importantes de la teoría de números computacional. En realidad, son muchos los algoritmos conocidos; por citar algunos, mencionemos los métodos ρ y $p-1$ de Pollard, los basados en curvas elípticas, la criba cuadrática multipolinomial, la criba de cuerpos... Una pequeña descripción de esto, enfocada hacia la criptografía, puede verse en [99, § 5.6]. Varios de estos métodos de factorización —los más sencillos— están muy bien explicados en [88]; los más complejos se pueden consultar en [83] y [30].

[21] La factorización más famosa ha sido, sin duda, la del número conocido como RSA129 (su nombre hace alusión a que tiene 129 dígitos). El reto para encontrar su descomposición en factores fue propuesto por Martin Gardner en su columna «Mathematical Games» de la revista *Scientific American* en 1977, coincidiendo con la invención del criptosistema RSA, e incluso antes de que Rivest, Shamir y Adleman lo publicaran en una revista científica. Gracias a los grandes avances



Martin Gardner
(1914–2010)



Peter Shor (1959–)

número de dígitos que tiene el número que se quiere factorizar. Actualmente no se conoce ningún algoritmo así, pero no podemos saber si alguien descubrirá uno en el futuro. Y es que, a veces, la investigación matemática depara sorpresas.

Un problema similar —pero no equivalente— al de factorizar en tiempo polinomial es asegurar si un número es primo o compuesto en un tiempo polinomial. No se sabía si podría existir un método que permitiera hacerlo hasta que, en 2002, los indios Agrawal, Kayal y Saxena sorprendieron a la comunidad matemática mundial exhibiendo un algoritmo sencillo que lo lograba. Ese test de primalidad —al que se alude con el nombre de AKS— utiliza matemáticas mucho más elementales que los asintóticamente más rápidos conocidos hasta entonces. Se puede ver en [1] o [30, § 4.5]. ◁

Estrictamente, romper el criptosistema RSA no requiere averiguar $\varphi(m)$ (y para ello p y q), sino d. Pero, con los conocimientos matemáticos actuales, el mejor camino de llegar a d es pasando por $\varphi(m)$. ¿Existirá un método más rápido? Quién sabe... Tampoco es necesariamente cierto que para hallar $\varphi(m)$ haga falta conocer los dos factores de m. Pero sí que son problemas equivalentes, como muestra el ejercicio 2.33.

Aun así —siempre ocurre lo mismo en criptografía—, existen diversos tipos de ataques para intentar reventar el criptosistema RSA y, por el contrario, también

en los métodos de factorización de enteros durante esos años, la factorización de RSA129 se logró en 1994, mucho antes de lo previsto por Gardner.

Y ya que citamos a Martin Gardner, no podemos dejar de decir que él ha sido —posiblemente— el mayor divulgador de matemáticas de la historia. Además de su contribución mensual en *Scientific American* durante casi 25 años —que era traducida al español en la revista *Investigación y Ciencia*—, publicó cerca de un centenar de libros de divulgación, matemática recreativa, magia..., muchos de ellos también traducidos al español. Al lector le será fácil encontrarlos, y seguro que disfruta con ellos.

recomendaciones a la hora de elegir sus parámetros (los números p, q, e y d que dan lugar a las claves privadas y públicas) para prevenir tales ataques. Nosotros ya no vamos a profundizar en ello; si el lector está interesado en conocer más detalles puede acudir a un texto especializado, como [65, § 8.2]. Pero sí queremos decir que al elegir p y q es conveniente tomar dos precauciones que evitan los ataques más serios. La primera es tomar p y q de manera que $\mathrm{mcd}(p-1, q-1)$ sea pequeño; y, la segunda, que $p-1$ y $q-1$ contengan algún factor primo grande. Estas condiciones se satisfacen de manera óptima cuando p y q tienen la forma $p = 2p_1+1$ y $q = 2q_1+1$ con p_1 y q_1 también primos. Cuando esto ocurre, los primos p y q se denominan 1-seguros (usar primos 1-seguros es una conveniencia común de distintas técnicas criptográficas, y a nosotros ya nos habían aparecido en el apartado 2.9.1).

Nota 2.31. Actualmente, la criptografía tiene casi más importancia en el terreno comercial que en el militar. Es evidente que la seguridad de muchos sistemas informáticos —los de los bancos, por citar el ejemplo más evidente— requieren que sus comunicaciones estén encriptadas de manera segura. El criptosistema RSA es la base de muchos sistemas comerciales, pero no se suele usar él solo. Su gran ventaja es que sólo una persona tiene que conocer la parte privada de la clave, y por tanto no puede ser interceptada en ningún intercambio de claves, que no hace falta realizar. Pero tiene el inconveniente de que efectuar las operaciones (2.21) y (2.23) es relativamente costoso —en cuanto al tiempo que se requiere para hacerlo— si lo comparamos con otros sistemas que sí necesitan que el emisor y el receptor de mensajes cifrados compartan la clave. Así, es muy común utilizar el criptosistema RSA únicamente para distribuir las claves de un criptosistema de clave privada. Una vez que la clave ha viajado de forma segura, este otro sistema puede no tener ninguna vulnerabilidad. ◁

No podemos terminar sin mencionar la llamada computación cuántica que, si alguna vez dejara de ser pura especulación y pudiera ser llevada a la práctica, revolucionaría toda la criptografía actualmente existente. Existen algoritmos que, basados en las propiedades de la mecánica cuántica, funcionarían de manera muy eficiente en un hipotético ordenador cuántico. En lo que a nosotros respecta, en 1994, Peter Shor concibió un algoritmo que permitiría factorizar números enormes sin dificultad. Pero, como decimos, nadie ha sido capaz de construir tales ordenadores, ni se sabe si existirán en el futuro. Todo esto, si funcionara, daría al traste con la seguridad del criptosistema RSA (y también con la de otros sistemas que aquí ni siquiera hemos mencionado). Afortunadamente, la misma mecánica cuántica proporcionaría nuevos métodos criptográficos que no se verían afectados. Pero nos estamos apartando de nuestro tema; nos queda únicamente comentar que en [97] se puede encontrar bastante información no técnica sobre esto.

2.9.5. Firma digital

En cuanto al envío y recepción de mensajes se refiere, hay una cosa que hemos pasado por alto. Si Benito recibe un supuesto mensaje de Alicia, ¿cómo puede estar seguro él de que ha sido ella quién se lo ha enviado, y no alguien que está suplantándola? Tradicionalmente, este tipo de seguridad (?) se lograba con la firma manuscrita de Alicia al pie del documento en papel, o por medio de distintos tipos de sellos. Obviamente, esto no tiene sentido en un mundo que se comunica electrónicamente, así que hay que buscar procedimientos alternativos.

Se necesita, pues, algún tipo de método (diseñado para ser usado en comunicaciones entre ordenadores) mediante el cual, cualquiera que reciba una comunicación de Alicia, puede asegurarse de si realmente ha sido ella quien la ha mandado. Esto es lo que se denomina *firma digital*.

Si las comunicaciones entre Alicia y Benito están cifradas mediante una clave privada simétrica que sólo conocen ellos dos, el problema está solucionado (aunque entonces nos reencontraríamos con el problema de la distribución de claves): Benito recibe el mensaje cifrado y lo descifra con la clave; si sale algo con sentido, es que el mensaje lo ha cifrado Alicia, la única persona con la que Benito comparte la clave. Pero tal proceso no sirve con claves públicas, pues cualquiera conoce las claves de cifrar, luego no hay impedimento alguno para que un suplantador envíe mensajes correctamente cifrados.

Afortunadamente, muchos métodos criptográficos de clave pública se pueden adaptar a este nuevo cometido. Y no resulta difícil: esencialmente, se trata de invertir los papeles de las claves de cifrar y descifrar. Cuando Alicia quiere enviar un mensaje, usa su propia clave privada para operar como si lo estuviera descifrando —éste es el proceso de firmar—; tras ello, lo envía a Benito. Cuando éste lo recibe, hace como si lo estuviera cifrando con la clave pública de Alicia. Si esto llega a buen fin, es que la clave pública se ha acoplado bien con la clave privada; como la clave privada sólo la conoce Alicia, es que ha sido ella la que ha enviado el mensaje. (Tal como hemos descrito, el mensaje ha viajado sin ningún tipo de protección; lo normal hubiera sido que, tras firmar el mensaje, Alicia lo cifrara con la clave pública de Benito antes de enviárselo. Luego, Benito deberá descifrarlo con su propia clave privada antes de comprobar que está firmado por Alicia.)

Veamos cómo funciona todo esto con el criptosistema RSA (fíjese el lector que sólo llevamos a cabo el proceso de firma, no el cifrado de seguridad explicado antes entre paréntesis). Sean p y q dos primos grandes distintos, $m = pq$ y $\varphi(m) = (p-1)(q-1)$; asimismo, sea otro entero positivo e tal que $\mathrm{mcd}(\varphi(m), e) = 1$, y d el inverso de e módulo $\varphi(m)$. En estas circunstancias, m y e son la clave pública de Alicia; los números p, q, $\varphi(m)$ y d son su clave privada. Si Alicia quiere firmar un

mensaje P (un número natural menor que m), calcula

$$S \equiv P^d \pmod{m}, \qquad \text{con } 0 < S < m,$$

y lo envía a Benito. Como $ed \equiv 1 \pmod{\varphi(m)}$ (y eso sólo ocurre con el d que conoce Alicia), se tendrá $ed = \varphi(m)k + 1$ para algún entero k. Entonces, el teorema de Euler-Fermat nos asegura que

$$S^e \equiv P^{ed} \equiv P^{\varphi(m)k+1} \equiv (P^{\varphi(m)})^k P \equiv P \pmod{m}.$$

Así, lo único que tiene que hacer Benito es evaluar S^e módulo m. Sólo recuperará el mensaje P que alguien le quería enviar si realmente proviene de Alicia, porque ningún suplantador habrá podido calcular el S correcto que hace que funcione el método.

Esto tiene, además, otra consecuencia: Alicia no puede negar que es ella la que ha mandado el mensaje, pues —de nuevo el mismo motivo— nadie más ha podido producir S a partir de P.

Para finalizar, huelga decir que se puede combinar la criptografía con la firma digital. De ese modo Alicia y Benito lograrían enviarse mensajes cifrados y firmados.

2.10. Teoría de códigos

La teoría de códigos es una importante rama de las matemáticas cuyas aplicaciones son innumerables. Con la llegada de los ordenadores y la comunicación electrónica —que ha originado una creciente demanda de todo tipo de códigos en el mundo industrial y financiero—, su desarrollo ha sido espectacular. No es el objetivo de este libro profundizar en su estudio, ni siquiera dar una una visión suficientemente general. Pero sí que queremos exponer algunos ejemplos sencillos, prácticamente sólo para convencer al lector de que —al igual que en criptografía— la teoría de números en general y las congruencias en particular resultan ser dos herramientas básicas en teoría de códigos.

A decir verdad, hay otras herramientas matemáticas —por ejemplo, la teoría de probabilidad— que también tienen mucho que decir en el campo de la codificación y las transmisión de la información. Aunque nosotros ya no nos ocuparemos de ello, sí que queremos citar que aquí intervienen conceptos como medida de la información, incertidumbre, entropía, capacidad de un canal de comunicación, teoremas de Shannon, ruido... y, finalmente, una inmensa pléyade de siglas para designar todo tipo de códigos y formatos. Por otra parte, el uso de códigos en la vida real no sólo requiere el estudio de sus aspectos matemáticos, sino también, y de manera

esencial, la fabricación de instrumentos que puedan aplicarlos de manera rápida, barata y fiable. Estos aspectos más tecnológicos también quedan fuera de nuestras pretensiones. Pero dejémonos de hablar sobre lo que no vamos a abordar y volvamos a nuestra senda.

La teoría de códigos se encarga de codificar la información, sea la que sea. En cuanto se quiere hacer de una manera sistemática y ordenada, el método usado se va a poder modelar mediante algún tipo de matemáticas. Por ejemplo, para transmitir un mensaje a través de un canal de comunicación, forzosamente se requiere un proceso previo de codificación de la información mediante un código comprensible tanto por el emisor como por el receptor (y que debería estar adaptado a las características del canal). Además, cuando la información se almacena, se manipula, o se transmite, a veces surgen errores. ¿Cómo intentar evitarlos? El primer objetivo es poderlos detectar. Si lo logramos, quizás podamos hacer algo para corregirlos de alguna manera (por ejemplo, repitiendo algún proceso que ha fallado o acudiendo a una fuente anterior); en todo caso, siempre podríamos descartar la información errónea encontrada. Para esto se emplean los denominados *códigos detectores de errores*.

Un objetivo más ambicioso es disponer de códigos que no sólo detecten errores, sino que también permitan corregirlos de manera automática. La ventaja de poder emplear tales códigos es innegable: evitarían tener que repetir procesos costosos e incluso permitirían recuperar información que de otra forma estaría perdida de manera irremediable. Para quien lo oye por primera vez, puede parecer cosa de brujas pero, de hecho, los *códigos correctores* existen, y aquí mostraremos algunos ejemplos.

Aunque ya no nos ocuparemos de ellos, también queremos citar la existencia de los denominados *códigos compresores*, que permiten comprimir los datos de manera que —en algún sentido— la información pueda almacenarse en menos espacio, o que se pueda transmitir más rápidamente.

2.10.1. Códigos detectores de errores

Aunque quizás al lector le pueda parecer lo contrario, el uso de códigos para detectar errores no es reciente. La realidad es que se han usado desde antiguo; por ejemplo, para detectar errores en las operaciones aritméticas.

De hecho, la denominada *prueba del nueve* que se usa para detectar errores en las cuatro operaciones elementales —suma, resta, multiplicación y división— ya está explicada en el *Liber Abaci* de Fibonacci, publicado en 1202.[22] La prueba del

[22] Tal hemos comentado en el tema anterior —véase la sección 1.3—, este libro tuvo una gran influencia en la adopción del sistema de numeración posicional en base 10 en Europa.

nueve era muy común cuando las operaciones se realizaban «a mano»,[23] pero cayó en desuso con la llegada de las calculadoras electrónicas.

La fundamentación matemática de la prueba del nueve es muy sencilla. En base 10, los números enteros los expresamos como

$$n = \sum_{j=0}^{k} d_j 10^j = d_k \dots d_2 d_1 d_0.$$

Al ser $10 \equiv 1 \pmod 9$ —y por tanto también $10^j \equiv 1^j \equiv 1 \pmod 9$—, es claro que

$$n \equiv \sum_{j=0}^{k} d_j \pmod 9.$$

En otras palabras, un número es congruente módulo 9 con la suma de sus dígitos (en base 10). (De hecho, hay una analogía a esto para cualquier base de numeración, como se muestra en el ejercicio 2.5.) En lo que sigue, y para entendernos de manera informal, diremos que, al efectuar la suma de los dígitos de un número, lo hemos «reducido» módulo 9; es más, si al reducirlo sale un número de más de un dígito, podemos volverlo a reducir y lo que obtengamos seguirá siendo congruente módulo 9 con el original; y así sucesivamente. De este modo, dado $n = \sum_{j=0}^{k} d_j 10^j$, sumando dígitos (varias veces si es necesario), llegaremos fácilmente a $n \equiv n' \pmod 9$ con $0 \leq n' < 9$.

La suma, la resta y el producto se comportan bien al tomar congruencias; por tanto, si tenemos

$$a \equiv a' \pmod 9, \quad 0 \leq a' < 9, \qquad b \equiv b' \pmod 9, \quad 0 \leq b' < 9,$$

deberá cumplirse

$$a \pm b \equiv a' \pm b' \pmod 9, \quad ab \equiv a'b' \pmod 9.$$

Así, en cualquiera de los tres casos, supongamos que al efectuar la correspondiente operación hemos llegado a $c = a \pm b$ o $c = ab$, y queremos comprobar si nos hemos equivocado. Para ello, calculamos c' la reducida módulo 9 de c, es decir, $c \equiv c' \pmod 9$ con $0 \leq c' < 9$. Como las operaciones entre reducidas conservan las congruencias módulo 9, si el resultado c es correcto forzosamente deberá cumplirse, respectivamente, $c' \equiv a' \pm b' \pmod 9$ o $c' \equiv a'b' \pmod 9$. Comprobar esto es muy sencillo, pues a', b', c' son números pequeños. En caso de que no se cumpla, es señal

[23] El que esto escribe lo recuerda perfectamente de sus primeros años escolares.

inequívoca de que hemos cometido algún error. El proceso de calcular a', b' y c' (sumando dígitos), y ver si c' responde a la operación entre a' y b' es lo que se denomina prueba del nueve; si los números involucrados en una operación no pasan la prueba, seguro que ha habido un error. La prueba del nueve es, pues, un código detector de errores. Existe la posibilidad de que una operación errónea pase la prueba pero, de hecho, la prueba del nueve no se deja engañar por los errores más comunes cometidos al operar «a mano».

Para la división (con resto) la prueba del nueve es similar, sin más que pensar $a/b = c+r/b$ como $a = cb+r$ y tomar congruencias en esta última expresión; es decir, que si a', b', c' y r' son las correspondientes reducidas de a, b, c y r, se deberá cumplir

$$a' \equiv c'b' + r' \pmod 9.$$

Si eso no es cierto, nos hemos equivocado al dividir.

En realidad, no importa si los números que estamos manejando no son enteros; pueden ser también números reales escritos en notación decimal con una cantidad finita de decimales. Si estamos manejando t decimales, podemos multiplicar por 10^t ambos lados de la igualdad que se corresponde con la operación efectuada. De este modo —que no afecta a las congruencias— habremos transformado todos los números involucrados en enteros.

Ejemplo. Supongamos que al sumar 8 373 049 y 6 054 792 hemos obtenido

$$8\,373\,049 + 6\,054\,792 = 14\,327\,841.$$

Para intentar detectar un posible error procedemos así:

$$\begin{aligned} 8\,373\,049 &\equiv 8+3+7+3+0+4+9 \equiv 7 \pmod 9,\\ 6\,054\,792 &\equiv 6+0+5+4+7+9+2 \equiv 6 \pmod 9,\\ 14\,327\,841 &\equiv 1+4+3+2+7+8+4+1 \equiv 3 \pmod 9. \end{aligned}$$

Pero $7+6 \not\equiv 3 \pmod 9$, luego seguro que hemos cometido algún error; en consecuencia, debemos repetir la suma para corregirlo. Como procedimiento operacional para calcular los restos módulo 9, hay un truco muy útil: al sumar de memoria, por ejemplo, $1+4+3+2+7+8+4+1$, se pueden ir buscando subgrupos que sumen 9, y eso ya se descarta mentalmente. Además, cada vez que, al ir sumando dígitos, alcanzamos un número de dos cifras, se suman ambas antes de continuar. ◊

Tal como ya hemos comentado, la prueba del nueve no detecta todos los errores —ningún código detector de errores lo hace—. Por ejemplo, no detecta bailes de

números (errores de transposición). Pero al efectuar operaciones manualmente no es probable que se dé un baile de números, y por tanto no es un defecto importante. En cambio, no tendría ninguna utilidad si se utilizara para chequear errores producidos al copiar manualmente listas de números, pues allí un error de transposición es bastante habitual, y pasaría desapercibido.

Pero dejemos ya la prueba del nueve y vamos con otros ejemplos de códigos. Se pueden encontrar muchos más en [27, capítulos 9 y 10].

Ejemplo. Desde los años 70 del siglo XX, las editoriales han venido identificando sus libros con el denominado International Standard Book Number (ISBN). Éste era un código de 10 dígitos

$$x_1x_2x_3x_4x_5x_6x_7x_8x_9x_{10}$$

en el que x_{10} es un dígito de control definido como

$$x_{10} \equiv \sum_{j=1}^{9} jx_j \pmod{11}, \quad \text{con } 0 \leq x_{10} \leq 10, \tag{2.24}$$

y con la salvedad de que, si la fórmula anterior da $x_{10} = 10$, en lugar de un «10» se usa una «X» cuando el ISBN aparece escrito (algunas editoriales, para evitar este engorro, descartan los ISBN que dan $x_{10} = 10$ y no los utilizan para codificar sus libros). Los dígitos x_1 a x_9 sirven para identificar el país o el idioma en el que está publicado el libro, la editorial, y el libro en cuestión; el x_{10} se utiliza exclusivamente para detectar errores al transmitir el ISBN o al teclearlo en un ordenador que automáticamente chequea si se cumple (2.24).[24] El ISBN es capaz de detectar errores de un solo dígito, y también los errores de transposición (ejercicio 2.34).

Ante la escasez de números ISBN disponibles, en 2007 se comenzó a usar un código de 13 dígitos

$$x_1x_2x_3x_4x_5x_6x_7x_8x_9x_{10}x_{11}x_{12}x_{13}$$

denominado ISBN-13 (a partir de entonces, el tradicional se denominó ISBN-10). En un número ISBN de 13 dígitos correctamente codificado, el dígito de control x_{13}

[24] En la práctica, los 10 dígitos se separan con guiones en cuatro grupos —no siempre de la misma longitud— que se corresponden con los anteriores conceptos; por ejemplo, en 84-239-6813-8, el 84 indica que el libro ha sido publicado en España. Pero estas separaciones no tienen ninguna influencia en el comportamiento matemático del código, así que aquí no las estamos teniendo en cuenta.

debe cumplir[25]

$$x_{13} \equiv 10 - \sum_{j\in\{1,3,\ldots,11\}} x_j - 3 \sum_{j\in\{2,4,\ldots,12\}} x_j \pmod{10}, \quad \text{con } 0 \leq x_{13} \leq 9.$$

Este nuevo método aumenta la cantidad de libros que se pueden codificar, pero como código detector de errores es peor que el ISBN-10, pues no siempre detecta transposiciones de un par dígitos consecutivos (ejercicio 2.36). ◊

Ejemplo. Desde 1990, el Documento Nacional de Identidad (DNI) español —un número entero positivo— incorpora un carácter de verificación que es simplemente el resto del número del DNI módulo 23. Este carácter añadido son las letras del alfabeto, con la precaución de que no se usan todas (no se emplea la letra «O» porque se puede confundir con un cero, ni tampoco «I», «Ñ» o «U»), y de que no se asignan en orden, sino que a los restos $0, 1, 2, \ldots, 22$ les corresponden las letras

T, R, W, A, G, M, Y, F, P, D, X, B, N, J, Z, S, Q, V, H, L, C, K, E

respectivamente. Habitualmente, al DNI seguido de su letra de verificación se le denomina Número de Identificación Fiscal (NIF) pues, desde que se ideó el método, se usa para identificar a todos los ciudadanos españoles a efectos fiscales. ◊

2.10.2. Funciones *hash*

Un mecanismo relacionado con la detección de errores —pero añadiendo una inequívoca componente criptográfica— es el uso de las denominadas *funciones hash*. Dado un mensaje o un archivo informático, una función hash se encarga de asignarle una especie de marca que lo identifique de manera casi unívoca, pero ocupando mucho menos que el original. La idea no es recuperar el mensaje o archivo original a partir de su marca, sino asegurarse de que no ha sido manipulado maliciosamente o se ha corrompido de manera accidental.

Como es habitual, supongamos de antemano que tenemos ya representados los mensajes, archivos informáticos, documentos... mediante números enteros no negativos. Entonces, una función hash es, simplemente, una función

$$H : \mathbb{N} \cup \{0\} \to M$$

donde M es un subconjunto de los enteros no negativos. La denominación de «hash» —que en inglés significa «picar y mezclar»— alude a que algo así es lo que hacen

[25] Esta fórmula se venía usando ya para definir los códigos de barras EAN-13 —EAN son las siglas de European Article Number— que incorporan la mayoría de los productos que compramos en los supermercados.

tales funciones cuando se aplican sobre un archivo informático x para extraer de él su marca $H(x)$, que se suele denominar suma de hash. En español, en lugar de «función hash» se usa a veces el término «función resumen» —la marca es, de alguna manera, una especie de resumen—, pero realmente lo más común es utilizar su designación en inglés. Y, en lugar de «suma de hash», también se emplea «suma de verificación» (o su versión inglesa *checksum*).

Dado que el conjunto de partida de una función hash H es mayor que el de llegada, es lógico que existan valores $x, y \in \mathbb{N}$ tales que $H(x) = H(y)$; es lo que se denominan colisiones de hash. Pero esas colisiones deben ser, en cierto modo, azarosas, y no debe ser fácil descubrir cómo lograrlas. Así, lo esencial para que una función hash H sea útil es que, dado $h = H(x)$, resulte virtualmente imposible encontrar y tal que $H(y) = h$. Con más detalles, las funciones hash buenas deben cumplir los siguientes requisitos criptográficos:

- Resistencia al cálculo de preimágenes: dado $h \in M$, no debe ser fácil encontrar y tal que $H(y) = h$.
- Resistencia (débil) a las colisiones de hash: dado x y su suma de hash $h = H(x)$, no debe ser fácil encontrar $y \neq x$ tal que $H(y) = H(x)$.
- Resistencia fuerte a las colisiones de hash: esta precaución extra pide que no sea fácil encontrar ningún par $x \neq y$ tal que $H(x) = H(y)$.

Una función hash que no cumpla este tipo de propiedades resulta inútil: por ejemplo, si la resistencia a las colisiones de hash falla, sería sencillo reemplazar un archivo original x por otro manipulado y con la misma suma de hash, y eso impediría detectar la alteración del archivo. En particular, las funciones hash son códigos detectores de errores, pero los ejemplos de códigos detectores de errores que hemos mostrado en el apartado anterior no sirven como funciones hash, pues no cumplen el imprescindible requisito de ser resistentes a las colisiones de hash, ni tampoco lo son al cálculo de preimágenes.

Hay bastantes funciones hash distintas que se utilizan en la práctica. Una de las más comunes es el algoritmo de reducción criptográfico denominado MD5 (abreviatura de Message-Digest Algorithm 5), que fue diseñado por Ronald Rivest en 1991. En él, las posibles sumas de hash $h \in M$ son los enteros entre 0 y $2^{128} - 1$; se suelen representar en notación hexadecimal, es decir, en base 16 con «a», «b», «c», «d», «e» y «f» para nuestros números decimales «10», «11», «12», «13», «14» y «15», respectivamente (de esa manera, un entero h que cumple $0 \leq h < 2^{128}$ se representa con 32 dígitos hexadecimales). Es relativamente común que los archivos informáticos se almacenen acompañados de sus correspondientes sumas MD5, que se pueden usar para comprobar que el archivo no se ha alterado accidentalmente

por un error informático o en una transmisión de datos (por ejemplo, al descargar un archivo de internet); también proporcionan al usuario confianza suficiente de que el archivo es igual que el publicado por su creador legítimo, y que no ha sido alterado de manera maliciosa por otro usuario o por algún tipo de virus informático o similar. Pero, a decir verdad, los avances matemáticos y de la potencia de cómputo han hecho que MD5 ya no sea seguro contra ataques premeditados, y cuando esa prevención es crucial hay que utilizar otros métodos más seguros (de hecho, MD5 surgió para reemplazar un método anterior denominado MD4 cuando se vio que éste era débil).

Similares a MD4 y MD5 son los algoritmos de la familia SHA (siglas de Secure Hash Algorithm), que es un sistema de funciones hash criptográficas diseñado por la Agencia de Seguridad Nacional de los Estados Unidos. En SHA-0 y SHA-1, sus resúmenes son, por construcción, números menores de 2^{160} (esto se corresponde con 40 caracteres hexadecimales), y ya no se consideran suficientemente seguros; mucho más seguro es SHA-2, cuyos resúmenes $h = H(x)$ están en el intervalo $0 \leq h < 2^{256}$ (es decir, 64 caracteres hexadecimales).

No explicaremos cómo está definida la función hash en MD5 ni en los SHA; los algoritmos que se emplean están descritos con detalle en muchos textos dedicados a la criptografía (por ejemplo, en [65, capítulo 9]). En su lugar, vamos a mostrar un algoritmo hash basado en aritmética modular que fue ideado por Adi Shamir, y que tiene claras reminiscencias de lo que se hace en el criptosistema RSA.

Elijamos dos primos grandes distintos p y q, y un número r tal que el menor k positivo que satisface $r^k \equiv 1 \pmod{pq}$ no sea muy pequeño (el teorema de Euler-Fermat o el teorema de la media de Fermat aseguran que $r^{(p-1)(q-1)} \equiv 1 \pmod{pq}$, pero quizás eso se puede lograr con exponentes menores); lo más sencillo es pensar que r es una raíz primitiva módulo pq, es decir, que no existe ningún $k < \varphi(pq) = (p-1)(q-1)$ que cumpla $r^k \equiv 1 \pmod{pq}$.[26] Con esto, podemos definir una función hash tomando, para cada $x \in \mathbb{N}$, el resto de r^x módulo pq; con un pequeño abuso de notación, lo podemos expresar así:

$$H(x) = r^x \pmod{pq}. \tag{2.25}$$

Puesto que no existen métodos eficientes de calcular logaritmos discretos, dado h no hay forma de encontrar y tal que

$$r^y \equiv h \pmod{pq},$$

luego H es resistente al cálculo de preimágenes.

[26] Las raíces primitivas las estudiaremos con mucho más detalle en el tema siguiente (sección 3.5).

Incluso aunque se conozca x y su suma de hash $h = H(x)$, tampoco va a ser posible hallar $y \neq x$ tal que

$$r^y \equiv r^x \pmod{pq},$$

y por tanto H también va a ser resistente a las colisiones de hash. Para comprobar esta resistencia, veamos en primer lugar que encontrar tal y que produjera la colisión de hash con x supondría dar con un y que verificase

$$y \equiv x \pmod{\varphi(pq)}; \tag{2.26}$$

en efecto, si existiera $y \not\equiv x \pmod{\varphi(pq)}$ tal que $r^y \equiv r^x \pmod{pq}$ podríamos escribir $y - x = c\varphi(pq) + k$ con $0 < k < \varphi(pq)$ y tendríamos

$$1 \equiv r^{y-x} \equiv r^{c\varphi(pq)+k} \equiv \left(r^{\varphi(pq)}\right)^c r^k \equiv 1^c r^k \equiv r^k \pmod{pq},$$

que no es posible por ser r raíz primitiva módulo pq, luego estamos forzados a que se cumpla (2.26). Pero no hay forma de resolver (2.26) si no se conoce $\varphi(pq)$, y para calcular $\varphi(pq)$ hay que encontrar los dos factores p y q a partir de su producto, que como bien sabemos es un problema intratable (la función hash H usa el producto pq y r, que se pueden adjuntar a $h = H(x)$ si por comodidad de manejo resulta conveniente, pero no se deben desvelar los dos factores p y q).

Hemos mostrado este ejemplo de función hash por su simplicidad conceptual, pero hemos de reconocer que, en la práctica, no se usa. La razón es que, pese a que es muy segura, la función hash definida en (2.25) es lenta comparada con otras funciones hash más difíciles de describir, pero que requieren mucho menos tiempo de cómputo. Calcular exponenciales modulares es mucho más rápido que calcular logaritmos discretos (y de aquí la seguridad del método), pero a su vez mucho más lento que someter números enteros a otro tipo de algoritmos también resistentes al cálculo de preimágenes y a las colisiones de hash.

De hecho, dado que existen funciones hash muy rápidas y seguras, esto les proporciona una utilidad adicional. Efectuar los cálculos necesarios para firmar digitalmente un mensaje grande es mucho más trabajoso que evaluar la función hash; además, firmar un mensaje largo requiere un proceso previo de descomponer el mensaje en trozos del tamaño adecuado, y luego aplicar el algoritmo de firma digital a cada trozo (por ejemplo, en la firma digital RSA del apartado 2.9.5, lo que se cifran son enteros positivos menores que pq, pero las funciones hash operan sobre todo $\mathbb{N}$). En consecuencia, es muy habitual utilizar la firma digital en conjunción con las funciones hash: en lugar de firmar el mensaje en sí mismo, primero se calcula su suma de hash y luego se firma únicamente esta suma, con lo cual se consigue un importante ahorro de tiempo de cálculo.

2.10.3. La información es digital: códigos binarios

Desde que los ordenadores entraron en escena, podemos suponer que tanto la codificación como la transmisión de la información se hacen de manera digital, pues los ordenadores trabajan internamente en base 2. Ya sean números, letras, imágenes..., todo acaba codificado como cadenas de ceros y unos (*bits*) dentro del ordenador. En consecuencia, muchos de los códigos que se emplean actualmente están diseñados para almacenar y transmitir información digital, es decir, cadenas de bits de determinada longitud. Estos códigos basados en ceros y unos se denominan a menudo *códigos binarios*.

Veamos un ejemplo elemental pero que ha tenido una gran trascendencia, y que incluía un bit adicional para detectar errores:

Ejemplo. Tal como se creó en la década de los sesenta del siglo XX, el código ASCII —siglas en inglés de American Standard Code for Information Interchange— era un código para representar los caracteres del alfabeto latino tal como se usa en inglés moderno y en otras lenguas occidentales. Era una evolución de los códigos utilizados entonces en telegrafía, y lo que se pretendía con él era disponer de un código estándar que permitiera la comunicación entre dispositivos digitales, así como el almacenamiento de la información y su manipulación.

Mediante 7 bits, el código ASCII permitía codificar $2^7 = 128$ caracteres, lo que era suficiente para poder representar las letras mayúsculas y minúsculas del inglés, los números del 0 al 9, los signos de puntuación, varios símbolos (como el de dólar), y ciertos caracteres de control (incluso algunos ya obsoletos por provenir de la telegrafía). Aparte de estos 7 bits —denotémoslos $x_1, x_2, \ldots, x_7$—, los códigos ASCII originales —el estándar evolucionó con el tiempo— empleaban un bit adicional x_8 (bit de paridad) que se usaba para detectar errores en la transmisión. El bit de paridad se escogía 0 o 1 según se necesitase para que junto a los 7 bits anteriores se formara siempre una cadena con un número par de unos (y de ceros); en otras palabras, se tomaba $x_8 \equiv x_1 + x_2 + \cdots + x_7 \pmod 2$ para que un código ASCII bien formado siempre cumpliera

$$x_1 + x_2 + \cdots + x_7 + x_8 \equiv 0 \pmod 2. \tag{2.27}$$

Si en la transmisión de datos se producía un error, un $x_j = 0$ iba a pasar a ser $x_j = 1$ (o viceversa), y la redundancia procedente del bit de paridad iba a permitir detectar la existencia del error, pues ya no se cumpliría (2.27) (por supuesto, no se detectarían dos errores, pero la existencia de dos errores en una misma cadena de 8 bits era bastante improbable). ◊

Pero, a decir verdad, lo que acabamos de relatar no duró mucho. El código ASCII tradicional de 7 bits más uno de paridad se quedó obsoleto, y surgieron

como setas muchos códigos que además permitían representar los caracteres que existían en otras lenguas pero no en el inglés; en particular, el octavo bit se fue empleando para poder codificar los caracteres de las lenguas de Europa Occidental, en un proceso evolutivo bastante caótico. Este uso de los 8 bits se fue estandarizando poco a poco, aunque no se llegó a un estándar sino a varios. Además, con 8 bits era imposible llegar a una buena solución; incluso sólo con lenguas de caracteres latinos, siempre quedaban símbolos que ya no cabían. Para usarlos, había que emplear otra codificación de 8 bits que eliminara unos caracteres y pusiera otros en su lugar. En la práctica, utilizar a la vez caracteres de dos codificaciones distintas era enormemente dificultoso.

¡Y no digamos nada de lo que ocurría con las lenguas que se escriben con caracteres griegos, cirílicos, árabes, chinos...! Algunas, como el chino, tienen ellas solas muchos más caracteres distintos de las que caben en una tabla de $2^8 = 256$ caracteres. Resultaba imposible manejar tales lenguas con codificaciones de 8 bits.

Este lamentable estado de cosas tomó otro rumbo a finales de la década de los ochenta. Con el objetivo de facilitar el tratamiento informático, la transmisión y la visualización de textos de múltiples lenguajes y disciplinas técnicas, además de textos clásicos de lenguas muertas, se inició un ambicioso proyecto para diseñar un estándar universal de codificación de caracteres, el denominado Unicode. En Unicode, cada carácter se representa con 32 bits, lo cual garantiza que habrá espacio suficiente para albergar todas las necesidades de caracteres distintos prevista (al principio del proyecto se pensó en 16 bits, pero pronto se vio que no era suficiente). Ya se han publicado diversas versiones del estándar, que continúa evolucionando incorporando nuevos caracteres de escrituras muy diversas (cada vez más extrañas desde el punto de vista occidental). Afortunadamente, parece que el proceso va por buen camino, y Unicode se va abriendo paso.

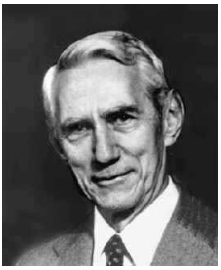
2.10.4. Códigos correctores

En los apartados anteriores hemos visto códigos que permitían detectar la presencia de algún error. Por ejemplo, añadiendo un bit de paridad. Este bit de paridad era un chivato de la presencia de un error, pero no permitía corregirlo. ¿Habrá algún método que no sólo informe de la existencia del error, sino que permita corregirlo? Vamos a ver que sí.

No está dentro de los objetivos de este libro realizar un estudio mínimamente profundo de los códigos correctores de errores, así que nos contentaremos con mostrar un ejemplo sencillo. Pero, aun así, no queremos dejar de insistir en su importancia práctica. Y es que actualmente los utilizamos por doquier en nuestra vida diaria. Por ejemplo, se utilizan códigos correctores en la lectura de los datos



Richard Wesley Hamming (1915–1998)



Claude Elwood Shannon (1916–2001)

grabados en un disco duro o en un disco compacto; esto es lo que hace que un CD de música suene bien aunque esté sucio o incluso arañado. Otra aplicación de tales códigos es la recepción correcta de mensajes digitales enviados a través de medios muy ruidosos; en particular, resultan imprescindibles para que podamos recibir las fotos enviadas por las sondas espaciales.

En lo que sigue, vamos a seguir tratando con información digital; es decir, nuestros datos siempre serán cadenas de ceros y unos (bits) de algún código binario. Hay una manera burda de detectar y corregir un error de un bit: ¡repetir ese bit tres veces! Si se produce un error en la transmisión de ese trío de bits que debían ser idénticos, la mayoría decide: si han llegado dos unos y un cero, el bit era 1 (y si han llegado dos ceros y un uno, el bit era 0).

Pero ya hemos dicho que ese método es muy burdo. En el código ASCII del ejemplo del apartado anterior, el bit de paridad adicional que permitía detectar un error aumentaba la longitud de los códigos en un factor 8/7. Repitiendo cada bit tres veces, estamos aumentando la longitud del código en un factor 3. Tanto para detectar como para corregir errores hay que introducir redundancia en el código. Pero nos interesa que la longitud del código no aumente demasiado; cuanto más aumente, más cuesta almacenar la información o transmitirla (tanto en tiempo como en dinero). Dando por hecho que esto se va a poder hacer, hay que llegar a un compromiso entre la cantidad de errores que puede detectar o corregir un código y la cantidad de redundancia que estamos dispuestos a introducir; además, hay que tener en cuenta si los errores son muy habituales o poco probables (si casi nunca hay errores, no necesitamos un código que permita corregir bastantes). Las matemáticas que se ocupan de esto han experimentado un importante avance en los últimos años, y los códigos más usados están fuera del alcance de lo que aquí estamos explicando. Aun así, existen códigos correctores no triviales que se pueden explicar fácilmente mediante el uso de congruencias. En concreto, vamos a mostrar un código corrector

cuya redundancia va a consistir en añadir tres bits de paridad para cada grupo de cuatro bits; es decir, la longitud de los códigos se está aumentando por un factor 7/4.

Para codificar o transmitir un mensaje de cuatro bits (x_1, x_2, x_3, x_4) vamos a emplear tres bits adicionales x_5, x_6, x_7 definidos como

$$\begin{aligned} x_5 &\equiv x_1 + x_2 + x_3 \pmod 2, \\ x_6 &\equiv x_1 + x_2 + x_4 \pmod 2, \\ x_7 &\equiv x_1 + x_3 + x_4 \pmod 2. \end{aligned}$$

Por simplicidad, a los elementos $(x_1, x_2, \ldots, x_7) \in \mathbb{Z}_2^7$ (donde estamos empleando $\mathbb{Z}_2^7$ para denotar $\mathbb{Z}_2 \times \mathbb{Z}_2 \times \cdots \times \mathbb{Z}_2$ siete veces) los llamaremos palabras. Así, las palabras válidas de nuestro código son los elementos $(x_1, x_2, \ldots, x_7) \in \mathbb{Z}_2^7$ tales que

$$\begin{aligned} x_1 + x_2 + x_3 + x_5 &\equiv 0 \pmod 2, \\ x_1 + x_2 + x_4 + x_6 &\equiv 0 \pmod 2, \\ x_1 + x_3 + x_4 + x_7 &\equiv 0 \pmod 2. \end{aligned} \tag{2.28}$$

Vamos a ver que este código es capaz no sólo de detectar, sino también de corregir, un error en cada palabra, es decir, en cada grupo de 7 bits. Lo ideó Richard W. Hamming hacia 1950, y es sólo uno de entre la familia de *códigos de Hamming*. En esta familia de códigos hay muchos otros, que son más o menos adecuados dependiendo de cuántos errores nos interese poder detectar y/o corregir. Con Hamming se inicia la teoría moderna de códigos detectores y correctores de errores, una pieza fundamental en el desarrollo de la teoría de la información iniciada por Claude E. Shannon poco antes.

Para entender el método de corrección de Hamming se puede plantear una analogía geométrica. Pensemos que las palabras válidas son los centros de una serie de círculos disjuntos en un plano (en nuestro caso hay $2^4 = 16$ palabras válidas y, por tanto, el mismo número de círculos). Un error en una palabra supone que tal palabra se ha desplazado del centro del círculo; pero, como no hay errores muy grandes —de un bit como máximo—, la palabra no habrá salido de su círculo. Para recuperarnos del error, basta llevar la palabra a su sitio correcto, el centro del círculo que ocupa.

Esta analogía geométrica se corresponde muy bien con el concepto de distancia. En $\mathbb{Z}_2^7$ —y si denotamos $x = (x_1, x_2, \ldots, x_7)$, y análogamente con y o z— usaremos la distancia dada por

$$d(x, y) = \sum_{j=1}^{7} |x_j - y_j|. \tag{2.29}$$

Por supuesto, esto satisface las propiedades habituales de las distancias, entre ellas la desigualdad triangular

$$d(x,z) \leq d(x,y) + d(y,z).$$

La clave es que entre dos palabras válidas distintas x e y (es decir, tales que sus correspondientes componentes satisfacen (2.28)), siempre se cumple $d(x,y) \geq 3$ —dejamos al lector que lo compruebe por sí mismo (ejercicio 2.37)—. Así, decimos que la *distancia de Hamming* del código es 3.

Si al transmitir una palabra x se produce un error de un bit, en lugar de x aparece x' que tiene una componente $x'_j \neq x_j$, y el resto iguales; en consecuencia $d(x,x') = 1$. Quien recibe x' —que ya no va a ser una palabra válida— quiere poder recuperar x. Para ello, mide la distancia de x' con todas las palabras válidas. La palabra original x es la única que satisface $d(x,x') < 2$, pues todas las demás palabras válidas $y \neq x$ están a distancia $d(y,x') \geq 2$. En efecto, si para una de éstas se cumpliera $d(y,x') < 2$, por la desigualdad triangular tendríamos

$$3 \leq d(x,y) \leq d(x,x') + d(x',y) < 1 + 2 = 3,$$

que es absurdo. Esto —con la hipótesis que hemos impuesto de que sólo se puede producir un error por palabra— permite adivinar x. Recuperar x por este procedimiento de comparar x' con todas las palabras válidas funciona, pero no es muy eficiente. De hecho, existen procedimientos mucho más efectivos —se lo aseguramos al lector—, pero ya no vamos a mostrarlos.

Basándose en las ideas de Hamming, y poco después que él, Marcel Golay encontró un código corrector de 24 bits especialmente eficiente. El código de Golay tenía muchas palabras admisibles, y era capaz de corregir hasta tres errores e incluso detectar (aunque no corregir) cuatro errores. Se usó para transmitir las imágenes de Júpiter y Saturno captadas por las naves espaciales Voyager en 1979 y 1980.

Detenemos aquí nuestro avance por el inmenso campo de la teoría de códigos en general y los códigos correctores de errores en particular, y sólo nos queda decir al lector que, si le ha interesado el tema, puede proseguir su estudio en alguno de los muchos textos especializado en teoría de códigos, como [53, 101] (el primero más elemental que el segundo). En español queremos recomendar el artículo [67] y los textos [68, 85].

2.11. Ejercicios

2.1. Demuestra las siguientes propiedades (naturalmente, damos por hecho que $a, b, x, y, m \in \mathbb{Z}$ y $m \geq 2$):

(a) $a \equiv b \pmod{m} \Rightarrow \mathrm{mcd}(a,m) = \mathrm{mcd}(b,m)$;
(b) $ax \equiv ay \pmod{m}$, $\mathrm{mcd}(a,m) = 1 \Rightarrow x \equiv y \pmod{m}$;
(c) $a \equiv b \pmod{m}$, $|b-a| < m \Rightarrow a = b$.

2.2. Sea n un entero impar. Prueba que $n^2 \equiv 1 \pmod{8}$.

2.3. Comprueba que ningún número de la forma $10d + 2$ (es decir, que, en base 10, acaba en 2) puede ser un cuadrado perfecto.

2.4. Prueba que, si n y m son enteros y $m > 2$, entonces $3^m \pm 1 \neq 2^n$.

2.5. Prueba que todo número natural escrito en base $m+1$ es congruente con la suma de sus cifras módulo m; incluso que lo mismo es cierto módulo cualquier divisor de m.

2.6. Muestra que la hipótesis $\mathrm{mcd}(a,m) = 1$ del teorema de Euler-Fermat es esencial, en el sentido de que, si $\mathrm{mcd}(a,m) = d > 1$, *nunca* se cumple $a^{\varphi(m)} \equiv 1 \pmod{m}$.

2.7. Demuestra el teorema pequeño de Fermat (corolario 2.6) sin usar el teorema de Euler-Fermat, sino el binomio de Newton.

2.8. Sea $m \geq 2$ un entero fijo. Dado $a \in \mathbb{Z}$ que verifica $\mathrm{mcd}(a,m) = 1$, el orden de a es, por definición,

$$\mathrm{ord}(a) = \mathrm{mín}\{\, k \geq 1 : a^k \equiv 1 \pmod{m}\}.$$

Prueba que $\mathrm{ord}(a)$ divide a $\varphi(m)$.

2.9. ¿Cuántos enteros entre 1 y 1000 (ambos incluidos) quedan si descartamos los múltiplos de 3, los múltiplos de 8, y los múltiplos de 11?

2.10. Calcula de cuántas maneras se pueden ordenar los elementos del conjunto $\{1, 2, \ldots, n\}$ de forma que ningún número k esté en la posición k. (A esto se le denomina el *número de desordenaciones*, o de *desarreglos*, de n elementos.)
Sugerencia. Usa el principio de inclusión-exclusión.

2.11. En el servicio de correos han contratado un nuevo cartero, totalmente inexperto. ¿De cuántas formas distintas puede entregar 4 cartas a 4 destinatarios sin que ninguno reciba la carta a él dirigida?

2.12. Sea p un primo y $a, k \in \mathbb{N}$ con $k \geq 1$. Demuestra que $\left(a^{p^k} - a^{p^{k-1}}\right)/p^k$ es un entero. (Por si acaso: a^{p^k} significa $a^{(p^k)}$.)

2.13. Sea $222\ldots223$ el desarrollo decimal de un número de 1986 dígitos, todos ellos «2» salvo el último. Prueba que no es primo.
Sugerencia. Demuestra que es divisible por 61.

2.14. Sean a, b y c números primos distintos. Demuestra que

$$(ab)^{c-1} + (bc)^{a-1} + (ca)^{b-1} - 1$$

es divisible por abc.

2.15. Dado un primo p, la suma de las potencias p-ésimas de tres enteros se resta de la potencia p-ésima de la suma de estos números. ¿Esta diferencia es siempre divisible por p?

2.16. Para k entero positivo, encuentra las dos últimas cifras del desarrollo decimal de 2^{20k+1}.

2.17. Sea n un número congruente con 7 módulo 8. Prueba que n no se puede escribir como suma de tres cuadrados.[27]

2.18. Analiza la longitud del periodo del generador de números pseudoaleatorios (2.12) cuando $\mcd(a, m) = 1$ y $c = 0$. Busca una cota superior a tal longitud.

2.19. Idea un procedimiento para encontrar enteros aleatorios grandes con un dado.

2.20. Sea p un primo impar y n un entero no divisible por p; prueba que $n^{(p-1)/2} \equiv \pm 1 \pmod{p}$. ¿Qué ocurre si n es divisible por p?

2.21. Una cesta contiene n huevos. Si se van retirando en grupos de 2, 3, 4, 5 o 6, entonces el número de huevos que quedan en la cesta es, respectivamente, 1, 2, 3, 4 o 5. En cambio, si se retiran de 7 en 7, no queda ningún huevo. ¿Cuál es el menor número de huevos con el que puede ocurrir lo anteriormente descrito?

2.22. Encuentra todos los $x, y \in \mathbb{Z}$ tales que

$$\begin{cases} 128x + 41y = 1414, \\ x \equiv 1 \pmod{3}, \\ y \equiv 17 \pmod{5}. \end{cases}$$

[27] En 1770, Lagrange demostró que todo entero positivo puede ser escrito como suma de cuatro cuadrados; el ejercicio muestra que tres cuadrados no son suficientes. El *teorema de los cuatro cuadrados* de Lagrange se generaliza mediante el denominado *problema de Waring*, que fue demostrado por Hilbert en 1909: para cada $k \in \mathbb{N}$, existe un $g(k)$ tal que todo entero positivo puede ser escrito como la suma de $g(k)$ potencias k-ésimas. Pese a que, como decimos, la existencia de los $g(k)$ ya está establecida, sólo se conoce su valor exacto para unos pocos valores de k.

2.23. Halla todas las soluciones del sistema

$$\begin{cases} x^2 \equiv 1 \pmod 3, \\ x \equiv 2 \pmod 4. \end{cases}$$

2.24. Prueba este resultado, que proporciona aún más información que el recíproco del teorema de Wilson (teorema 2.23): Si $n \geq 5$ es un número compuesto, siempre se cumple $(n-1)! \equiv 0 \pmod n$.

2.25. Se llaman primos gemelos a dos números p y $p+2$ ambos primos. Prueba que dos enteros n y $n+2$, $n > 1$, son primos gemelos si y sólo si

$$4\big((n-1)!+1\big) + n \equiv 0 \pmod{n(n+2)}.$$

2.26. Sea p un primo impar. Demuestra que la ecuación $x^2 + 1 \equiv 0 \pmod p$ tiene solución si y sólo si $p \equiv 1 \pmod 4$.
Sugerencia. Comprueba previamente que, si tomamos $k = (p-1)/2$, se cumple $(p-1)! \equiv (-1)^k (k!)^2 \pmod p$.

2.27. Demuestra o refuta la siguiente propiedad:

$$a \equiv b \pmod m \;\Rightarrow\; x^a \equiv x^b \pmod m \text{ para todo } x \in \mathbb{Z}$$

(con a, b y m estamos denotando números enteros positivos, $m > 1$). Si el resultado anterior no es cierto, sustituye las hipótesis por otras que sí impliquen la tesis $x^a \equiv x^b \pmod m$.

2.28. Evalúa $4170^{3630} \pmod{23}$.

2.29. Para cada entero no negativo n, calcula el resto de la división de $3^{2n+5}+2^{4n+1}$ entre 7.

2.30. Prueba que $341 = 11 \cdot 31$ es pseudoprimo en base 2, pero que no lo es en base 7.
Sugerencia. Se pretende que el lector no use «fuerza bruta» para calcular 2^{340} y 7^{340}, sino que emplee argumentos más sutiles.

2.31. Prueba que $m = 561$ no cumple $a^{(m-1)/2} \equiv \pm 1 \pmod m$ con $a = 5$.

2.32. Analiza si 561 pasa el test de Miller-Rabin en base 2, 3 y 5.

2.33. Sea m un entero positivo que es producto de dos números primos distintos. Prueba que los divisores de m son las raíces de la ecuación

$$x^2 - (m + 1 - \varphi(m))x + m = 0,$$

y que, por tanto, $\varphi(m)$ determina los factores primos de m.

2.34. Demuestra que el ISBN de 10 dígitos detecta tanto errores en un dígito como transposiciones de un par de dígitos (no necesariamente consecutivos). Ten en cuenta que el dígito de control juega, en cuanto a la posibilidad de error, el mismo papel que los demás, y por tanto también puede estar afectado.

2.35. En este ejercicio vamos a comprobar que el ISBN de 10 dígitos también sirve para poder recuperar un número borrado. Supón que en el ISBN 84-8871?-32-0 se ha perdido un dígito, que indicamos con un símbolo de interrogación. ¿Cuál es el valor del dígito perdido?

2.36. Convéncete de que el código ISBN-13 no siempre detecta transposiciones de dígitos comprobando que, si en $x_1x_2\dots x_{12}x_{13}$ los dígitos x_j y x_{j+1} (con $1 \leq j \leq 11$) se diferencian en 5 y se intercambian, producirán el mismo dígito de control x_{13}.

2.37. Comprueba que entre dos palabras válidas distintas $x, y \in \mathbb{Z}_2^7$ del código de Hamming caracterizado por (2.28), siempre se cumple $d(x, y) \geq 3$, donde d es la distancia definida en (2.29).

Tema 3
Congruencias no lineales

3.1. Congruencias polinómicas
3.2. Restos cuadráticos y ley de reciprocidad cuadrática
3.3. Pseudoprimos de Euler
3.4. Tests de primalidad para números de formas especiales
3.5. Raíces primitivas
3.6. Criptosistema Elgamal
3.7. Ejercicios

Todo el tema anterior estaba dedicado a la teoría de congruencias, y a mostrar unas cuantas de sus numerosas aplicaciones. Pero, para no hacerlo demasiado largo, no hemos podido incluir en él todo lo que nos hubiera gustado. Sin salirnos del nivel de este libro, nos hemos dejado tanto algunos aspectos teóricos habituales en otros textos como algunas aplicaciones interesantes. Quizás el lector ya lo haya echado en falta.

Es por eso que incluimos aquí un nuevo tema sobre congruencias. Separar en dos los conceptos, técnicas y aplicaciones relacionados con las congruencias no es tarea fácil, pero lo hemos intentado. Por supuesto, en este segundo tema hemos intentado reunir la mayor parte del material que puede resultar más avanzado o dificultoso. La frontera de separación —no muy precisa— ha intentado ser la linealidad de lo que se estudia; expliquemos qué queremos decir con esto.

En el tema anterior, reiteradamente aparecían ecuaciones modulares de tipo $ax \equiv b \pmod{m}$, donde a y b son números enteros constantes, y x una incógnita. Este tipo de ecuaciones —o sistemas de ellas—, en las que la o las variables sólo se multiplican por constantes o se suman, se denominan lineales. Obtener toda clase de información sobre esas o similares ecuaciones y sus soluciones, y sacar partido de lo que de allí surgía, eran labores cotidianas en el tema anterior. De hecho, exprimiendo eso de manera adecuada era como llegábamos a obtener interesantes aplicaciones a la generación de números pseudoaleatorios, la criptografía, o la teoría de códigos, por ejemplo.

Pero —salvo algún caso puntual— allí no abordábamos situaciones en las que la variable estuviese elevada a una potencia, o en el exponente. Las ecuaciones en las que la incógnita aparece de esa manera se denominan no lineales.

Pues bien, en este tema vamos a estudiar los aspectos no lineales de las congruencias, bien sea porque directamente queremos analizar teóricamente o resolver ecuaciones no lineales, o porque de alguna manera intervienen en las aplicaciones que veamos. De este modo, los tres pilares teóricos básicos de este tema serán los siguientes:

- Estudio de ecuaciones en congruencias de tipo polinómico, es decir, $f(x) \equiv 0 \pmod{m}$ con f un polinomio de grado n.
- Un análisis mucho más pormenorizado de un polinomio particular: la ecuación $x^2 \equiv b \pmod{m}$. No siempre esta ecuación tiene solución, y los b para los que sí existe solución se denominarán restos cuadráticos (o residuos cuadráticos). En este contexto, el principal resultado es el teorema de reciprocidad cuadrática.
- Estudios relacionados con ecuaciones de tipo $a^x \equiv b \pmod{m}$. Esto dará lugar a los conceptos de raíces primitivas e índices modulares (o logaritmos discretos).

De modo similar a como hacíamos en el tema anterior, mostraremos algunas aplicaciones prácticas de lo estudiado. En particular, veremos algunos tests de primalidad y un método criptográfico de clave pública distinto al RSA que explicamos entonces.

Por supuesto, para seguir el tema anterior no hace falta éste, pero lo contrario no es del todo posible. El lector que quiera seguir este tema debe estar, al menos, familiarizado con el lenguaje y la notación de las congruencias —que ya no repetiremos—, y conocer algunas de sus propiedades más básicas y elementales. En particular, resulta fundamental conocer la función φ de Euler y el teorema de Euler-Fermat (teorema 2.4). Si hay que concretar más los requisitos, desenvolverse bien con las tres primeras secciones del tema anterior sería —en principio— suficiente.

La bibliografía básica del presente tema coincide en general con la del anterior, pues tanto los conceptos que aquí tratamos como su desarrollo son habituales en la mayoría de los textos de teoría de números. Así, podemos recomendar con carácter general los libros [3, 5, 41, 55, 71, 78, 87, 88]. En ocasiones —como al estudiar cuestiones criptográficas— habremos de ser más precisos, pero ya nos encargaremos de ello cuando surja el caso.

Como no podría ser de otra forma, continuaremos dejándonos cosas en el tintero. Así, por ejemplo, en este tema veremos cómo la teoría de congruencias resulta útil para elaborar algunos test de primalidad muy eficientes, pero no hemos querido dedicar unas páginas extra a explicar con detalle algunos algoritmos de factorización basados en el empleo de congruencias. Además, algunos conceptos más avanzados que no hemos incluido, pues hubieran requerido una introducción algebraica más amplia, son los siguientes:

- Congruencias módulo no sólo un entero, sino módulo un polinomio de coeficientes enteros.
- Cuerpos finitos y teoría de curvas elípticas, muy usados en criptografía y teoría de códigos.

Si el lector lo desea, puede acudir a alguno de los numerosos libros que los tratan; entre ellos, algunos de los textos antes citados. A este respecto, queremos también recomendar [45], un excelente artículo que describe de manera particularmente amena las aplicaciones de la teoría computacional de números y de las curvas elípticas a la criptografía.

3.1. Congruencias polinómicas

Empalmando con lo que comentábamos en la introducción, en la proposición 2.1 y el teorema 2.8 del tema anterior estudiábamos las ecuaciones con congruencias

lineales, es decir, las que tienen la forma $ax \equiv b \pmod{m}$. Del mismo tipo son las congruencias que aparecen en el teorema chino del resto. Para todas ellas hemos sabido establecer cuáles son sus soluciones —y cómo encontrarlas— de una manera suficientemente satisfactoria.

Pero también nos podemos plantear ecuaciones con la forma $f(x) \equiv 0 \pmod{m}$, donde f es un polinomio de grado n, o incluso sus correspondientes sistemas. En cuanto sea $n > 1$, estas congruencias polinómicas ya no son lineales, y su estudio es mucho más complicado que el de las lineales —no podíamos esperar lo contrario, pues lo mismo sucede sin usar congruencias—; aquí sólo vamos a ver unos primeros resultados y algunos ejemplos.

En los casos más fáciles, tales ecuaciones se pueden intentar resolver mediante un simple tanteo, como vemos a continuación:

Ejemplo. Resolver las ecuaciones $x^2 \equiv 2 \pmod{5}$ y $x^2 \equiv 1 \pmod{5}$.

Solución. Sea $x = 5k+j$, con $k \in \mathbb{Z}$ y $j = 0, 1, 2, 3$ o 4. Así, $x^2 = 25k^2+10kj+j^2$, con lo cual, al tomar j los valores 0, 1, 2, 3 o 4, el correspondiente x^2 es congruente módulo 5 con 0, 1, 4, 4 o 1, respectivamente. En consecuencia, la ecuación $x^2 \equiv 2 \pmod{5}$ no tiene ninguna solución. Por el contrario, la ecuación $x^2 \equiv 1 \pmod{5}$ tiene como soluciones todos los x de la forma $5k + 1$ o $5k + 4$, con $k \in \mathbb{Z}$. (Nótese, además, que el lema 2.21 —en el tema 2— ya nos aseguraba que las soluciones de la ecuación $x^2 \equiv 1 \pmod{5}$ son, precisamente, los $x \equiv \pm 1 \pmod{5}$, tal como acabamos de comprobar.) ◊

Este ejemplo es ilustrativo de lo que puede ocurrir pero, obviamente, el proceso seguido para resolver las ecuaciones dista de ser un buen método en general. Antes de seguir, démonos cuenta de que una ecuación $f(x) \equiv 0 \pmod{m}$ está bien planteada en $\mathbb{Z}_m$, en el sentido de que, si un $x \in \mathbb{Z}$ es solución, también es solución toda su clase de equivalencia en $\mathbb{Z}_m$. Esto es así ya que todas las operaciones (sumas y productos) que aparecen en un polinomio están bien definidas en $\mathbb{Z}_m$, luego $f(x + km) \equiv f(x) \pmod{m}$ para todo $k \in \mathbb{Z}$. Para abreviar, usaremos la expresión «número de soluciones en $\mathbb{Z}_m$» con el mismo significado que «número de soluciones no congruentes módulo m».

El principal resultado que vamos a proporcionar en esta sección es el teorema de Lagrange. Este teorema nos dice que, cuando el módulo es un primo p, el número de soluciones en $\mathbb{Z}_p$ de una congruencia polinómica es menor o igual que el grado del polinomio. Esto recuerda, claramente, al teorema fundamental del álgebra (un polinomio en $\mathbb{C}[x]$ tiene tantas raíces en $\mathbb{C}$ —contando su multiplicidad— como su grado). Caiga el lector en la cuenta de que el teorema sólo nos informa del número de soluciones, no de cómo encontrarlas. Y, realmente, no hay ningún procedimiento general para ello.

Teorema 3.1. Lagrange, 1773.
Dado un primo p, sea $f(x) = c_0 + c_1x + \cdots + c_nx^n$ un polinomio de grado n y coeficientes enteros, y con $c_n \not\equiv 0 \pmod p$. Entonces, la congruencia polinómica $f(x) \equiv 0 \pmod p$ tiene a lo sumo n soluciones en $\mathbb{Z}_p$.

DEMOSTRACIÓN. Usaremos inducción sobre n, el grado del polinomio. Si $n = 1$, ya sabemos que $c_0 + c_1x \equiv 0 \pmod p$, con $c_1 \not\equiv 0 \pmod p$, sólo tiene una solución (proposición 2.1 con $m = p$ y $\mcd(c_1, p) = 1$). Para realizar el paso de inducción, demos el resultado por cierto para polinomios de grado $n-1$. ¿Será también verdad para polinomios de grado n? Supongamos que no, es decir, que un polinomio $f(x) = c_0 + c_1x + \cdots + c_nx^n$, con $c_n \not\equiv 0 \pmod p$, tiene $n+1$ raíces $x_0, x_1, \ldots, x_n$ no congruentes módulo p.

Como $x^j - x_0^j = (x - x_0)(x^{j-1} + x_0x^{j-2} + x_0^2x^{j-3} + \cdots + x_0^{j-1})$, podemos poner

$$f(x) - f(x_0) = \sum_{j=0}^{n} c_jx^j - \sum_{j=0}^{n} c_jx_0^j = \sum_{j=1}^{n} c_j(x^j - x_0^j) = (x - x_0)g(x),$$

donde $g(x)$ es un polinomio de coeficientes enteros, grado $n-1$ y coeficiente director c_n. Entonces,

$$f(x_k) - f(x_0) = (x_k - x_0)g(x_k) \equiv 0 \pmod p, \qquad k = 1, 2, \ldots, n,$$

ya que $f(x_k) \equiv f(x_0) \equiv 0 \pmod p$. Pero $x_k \not\equiv x_0 \pmod p$ si $k \neq 0$, y el módulo p es primo, así que, forzosamente,

$$g(x_k) \equiv 0 \pmod p, \qquad k = 1, 2, \ldots, n,$$

lo cual está en contradicción con la hipótesis de inducción. □

Obsérvese el «a lo sumo» del teorema anterior. No es cierto, en general, que haya n soluciones. En realidad, esto ya lo sabíamos pues en el ejemplo previo al teorema hemos visto que la ecuación $x^2 \equiv 2 \pmod 5$ no tiene ninguna solución.

Además, el teorema se convierte en una trivialidad cuando $n \geq p$. Como $\mathbb{Z}_p$ tiene p elementos, es imposible que en él haya más de p raíces. En este caso, además, conviene acudir al teorema pequeño de Fermat: basta aplicar $x^p \equiv x \pmod p$ y se reduce el grado del polinomio. A este respecto, véase el ejercicio 3.5.

Por otra parte, que el número de raíces es menor o igual que el grado no es cierto, en general, para módulos compuestos. Para convencerse de ello, basta observar lo que ocurre con la ecuación $x^3 - x \equiv 0 \pmod 6$, que tiene más de 3 soluciones en $\mathbb{Z}_6$:

Ejemplo. Hallar todas las soluciones de la ecuación $x^3 - x \equiv 0 \pmod 6$.

SOLUCIÓN. Descompongamos $x^3 - x = x(x-1)(x+1)$. Esto son tres números consecutivos; uno de ellos, será múltiplo de 3, y al menos uno deberá ser par. El producto de los tres seguro que es múltiplo de 6. Por consiguiente, cualquier número entero es solución de $x^3 - x \equiv 0 \pmod 6$. Así, en $\mathbb{Z}_6$ hay 6 soluciones: las clases de equivalencia de 0, 1, 2, 3, 4 y 5. ◊

Del teorema de Lagrange se puede extraer fácilmente la siguiente consecuencia (a primera vista, parece que del teorema sólo se desprende que p debe dividir al coeficiente director del polinomio; pero, entonces, ese coeficiente es nulo en $\mathbb{Z}_p$, y el segundo coeficiente se convierte en el director):

Corolario 3.2.

Sea p un número primo, y $f(x) = c_0 + c_1 x + \cdots + c_n x^n$ un polinomio de grado n y coeficientes enteros. Si la ecuación $f(x) \equiv 0 \pmod p$ tiene más de n soluciones en $\mathbb{Z}_p$, entonces cada coeficiente c_j es divisible por p.

En el ejercicio 3.6 proponemos al lector que complete la demostración de este resultado.

3.1.1. Congruencias polinómicas con un módulo compuesto

Nuestro objetivo ahora va a ser estudiar la ecuación polinómica con congruencias $f(x) \equiv 0 \pmod m$ cuando el módulo m no es un primo. Como veremos en la demostración del teorema que sigue, el teorema chino del resto va a permitir reducir su estudio al de $f(x) \equiv 0$ módulo potencias de primos.

Teorema 3.3.

Sea un entero $m > 1$ y $m = p_1^{a_1} p_2^{a_2} \cdots p_k^{a_k}$ su factorización canónica como producto de primos distintos. Sea asimismo f un polinomio con coeficientes enteros. Entonces, la congruencia

$$f(x) \equiv 0 \pmod m \tag{3.1}$$

tiene solución si y sólo si cada una de las congruencias

$$f(x) \equiv 0 \pmod{p_j^{a_j}}, \quad j = 1, 2, \ldots, k, \tag{3.2}$$

la tiene. Además, si ν denota el número de soluciones de (3.1) en $\mathbb{Z}_m$, y ν_j el número de soluciones de (3.1) en $\mathbb{Z}_{p_j^{a_j}}$, $j = 1, 2, \ldots, k$, entonces $\nu = \nu_1 \nu_2 \cdots \nu_k$.

DEMOSTRACIÓN. Supongamos que (3.1) tiene una solución s. Entonces, $m \mid f(s)$ y, como $p_j^{a_j} \mid m$, también $p_j^{a_j} \mid f(s)$. Por consiguiente, s es solución de todas las congruencias de (3.2).

Recíprocamente, supongamos que, para cada $j = 1, 2, \ldots, k$, existe s_j solución de $f(x) \equiv 0 \pmod{p_j^{a_j}}$. Como los $p_j^{a_j}$ son primos entre sí dos a dos, el teorema chino del resto nos asegura que existe un entero s solución del sistema

$$s \equiv s_j \pmod{p_j^{a_j}}, \quad j = 1, 2, \ldots, k.$$

Entonces, $f(s) \equiv f(s_j) \equiv 0 \pmod{p_j^{a_j}}$ para todo j. Así que $f(s)$ es divisible por cada $p_j^{a_j}$, y en consecuencia también por su producto, m; es decir, $f(s) \equiv 0 \pmod{m}$.

En cuanto al número de soluciones de las ecuaciones, el teorema chino del resto nos muestra que cada k-tupla $(s_1, s_2, \ldots, s_k)$ de soluciones de las ecuaciones individuales de (3.2) proporciona un único entero s (módulo m) que satisface el sistema formado por todas ellas. Es fácil convencerse de que esto constituye una biyección. Para ello, recordemos que toda solución s del sistema se consigue mediante

$$s = s_1 M_1 M_1' + s_2 M_2 M_2' + \cdots + s_k M_k M_k'$$

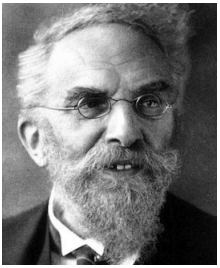
con $M_j = m/p_j^{a_j}$ y M_j' el inverso de M_j módulo $p_j^{a_j}$; de aquí que $s \equiv s_j \pmod{p_j^{a_j}}$ para cada $j = 1, 2, \ldots, k$. Entonces, el proceso es inyectivo, ya que si tenemos dos k-tuplas $(s_1, s_2, \ldots, s_k)$ y $(\overline{s_1}, \overline{s_2}, \ldots, \overline{s_k})$ distintas —digamos que $s_j \not\equiv \overline{s_j} \pmod{p_j^{a_j}}$— entonces sus correspondientes s y $\overline{s}$ verificarán $s \equiv s_j \not\equiv \overline{s_j} \equiv \overline{s} \pmod{p_j^{a_j}}$, así que $s \not\equiv \overline{s} \pmod{m}$ (pues $p_j^{a_j} \nmid m$). La suprayectividad es inmediata, ya que cualquier s solución del sistema se obtiene trivialmente a partir de la k-tupla $(s, s, \ldots, s)$.

Con esto, ya estamos en condiciones de efectuar el recuento. Cada s_j recorre las ν_j soluciones (módulo $p_j^{a_j}$) de $f(x) \equiv 0 \pmod{p_j^{a_j}}$. Por tanto, el número de enteros s que satisfacen el sistema (3.2) —y en consecuencia, según hemos probado, su equivalente (3.1)— es $\nu_1 \nu_2 \cdots \nu_k$. □

Tal como comentábamos en el preámbulo, este teorema reduce el estudio de congruencias polinómicas módulo un número compuesto al de $f(x) \equiv 0 \pmod{p^a}$ con p primo y $a \geq 1$. El caso $a = 1$ ya lo hemos analizado en el teorema de Lagrange. Fue Kurt Hensel, a principios del siglo XX, quien logró aclarar qué es lo que ocurre cuando $a \geq 2$. Con este mismo objetivo, pongamos primero de manifiesto ciertas relaciones entre las soluciones de $f(x) \equiv 0$ módulo p^a y p^{a-1}. (A la postre, esto nos permitirá dejar todo reducido al estudio de $f(x) \equiv 0 \pmod{p}$ y al de ciertas congruencias lineales.)

Dados un primo p y un entero $a \geq 2$, supongamos que s es una solución de

$$f(x) \equiv 0 \pmod{p^a}, \tag{3.3}$$



Kurt Hensel
(1861–1941)

donde s se ha elegido de forma que $0 \leq s < p^a$. Como $p^{a-1} \,|\, p^a$, es claro que s también será solución de

$$f(x) \equiv 0 \pmod{p^{a-1}}; \tag{3.4}$$

es posible que s no verifique $0 \leq s < p^{a-1}$. Para conseguir el equivalente a s módulo p^{a-1} basta hallar el resto al dividir:

$$s = qp^{a-1} + r, \quad \text{con } 0 \leq r < p^{a-1}.$$

Este r, que también es solución de (3.4), se dice que ha sido generado por s. De este modo hemos visto que cada solución s de (3.3), siendo $0 \leq s < p^a$, genera una solución r de (3.4) con $0 \leq r < p^{a-1}$.

De manera inversa, supongamos que (3.4) tiene una solución r, escogida de modo que $0 \leq r < p^{a-1}$. Nos preguntamos cuándo existirá una solución s de (3.3) que cumpla $0 \leq s < p^a$ y que genere r según el proceso descrito en el párrafo anterior. Cuando esto ocurra, se dice que r se puede subir de p^{a-1} a p^a.

El siguiente teorema aclara cuándo esto es posible y cuándo no.

Teorema 3.4. Lema de Hensel.
Dados p primo, $a \geq 2$, y $f(x)$ un polinomio con coeficientes enteros, supongamos que r es una solución de

$$f(x) \equiv 0 \pmod{p^{a-1}}; \tag{3.5}$$

que cumple $0 \leq r < p^{a-1}$. En estas condiciones, se tiene

(a) Si $f'(r) \not\equiv 0 \pmod{p}$, entonces r puede subirse de p^{a-1} a p^a de forma única. Es decir, existe un único s que cumple $0 \leq s < p^a$, genera r y es solución de

$$f(x) \equiv 0 \pmod{p^a}. \tag{3.6}$$

(b) Si $f'(r) \equiv 0 \pmod{p}$ y $f(r) \equiv 0 \pmod{p^a}$, entonces r puede subirse de p^{a-1} a p^a de p formas distintas.

(c) Si $f'(r) \equiv 0 \pmod{p}$ y $f(r) \not\equiv 0 \pmod{p^a}$, entonces r no puede subirse de p^{a-1} a p^a.

Demostración. Sea n el grado de f. La fórmula de Taylor nos muestra que

$$f(x+h) = f(x) + f'(x)h + \frac{f''(x)}{2!}\,h^2 + \cdots + \frac{f^{(n)}(x)}{n!}\,h^n, \tag{3.7}$$

lo que resulta válido para cada x y cada h. Además, para cada j entre 1 y n, el polinomio $f^{(j)}(x)/j!$ también tiene coeficientes enteros (ejercicio 3.8). En (3.7), sustituyamos $x = r$ y $h = qp^{a-1}$, donde q es un entero que especificaremos más adelante. Al ser $a \geq 2$, los sumandos de (3.7) que contienen h^2 y potencias superiores de h son múltiplos de p^a. En consecuencia, de (3.7) se deduce que

$$f(r + qp^{a-1}) = f(r) + f'(r)qp^{a-1} \pmod{p^a}. \tag{3.8}$$

Como r es solución de (3.5), tendrá que ser $f(r) = kp^{a-1}$ para algún $k \in \mathbb{Z}$; entonces, (3.8) queda

$$f(r + qp^{a-1}) = (f'(r)q + k)p^{a-1} \pmod{p^a}.$$

En estas circunstancias, tomemos

$$s = r + qp^{a-1}. \tag{3.9}$$

Este s es solución de (3.6) si y sólo si q satisface la congruencia lineal

$$f'(r)q + k \equiv 0 \pmod{p}. \tag{3.10}$$

Una vez llegado hasta aquí, veamos cómo surgen los tres casos del enunciado.

Si $f'(r) \not\equiv 0 \pmod{p}$, la congruencia (3.10) —en la que q es la incógnita— tiene solución única módulo p (proposición 2.1); elijamos esa solución q de modo que $0 \leq q < p$. Entonces, el s de (3.9) es solución de (3.6) y cumple $0 \leq s < p^a$, con lo que queda probado el apartado (a).

Si, por el contrario, $f'(r) \equiv 0 \pmod{p}$, la ecuación (3.10) queda reducida a esta otra trivial: $k \equiv 0 \pmod{p}$, en la que la incógnita (q) ha desaparecido. En estas circunstancias, existen dos posibilidades. Si $p \,|\, k$, la congruencia es cierta, y cualquier q es solución; así pues, sirve $q = 0, 1, \ldots, p-1$ (que, como antes, se llevan a (3.9) y generan las correspondientes soluciones de (3.6)). Si $p \nmid k$, no hay manera de elegir tal q. Finalmente, recordemos que $f(r) = kp^{a-1}$ así que decir $p \,|\, k$ es equivalente a afirmar $f'(r) \equiv 0 \pmod{p}$ (como se hace en (b)), y lo mismo ocurre con sus correspondientes negaciones (como en (c)). □

Obsérvese que la demostración del lema de Hensel da un método práctico para obtener soluciones de (3.6) si se tiene una solución r de (3.5) (esto es, para subir r de p^{a-1} a p^a). En este método, lo único que hay que hacer es tomar $k = f(r)/p^{a-1}$ y resolver la congruencia lineal $f'(r)q + k \equiv 0 \pmod{p}$ (con q como incógnita).

Además, este proceso se puede realizar tantas veces como sea necesario. De este modo, la resolución de una congruencia polinómica $f(x) \equiv 0 \pmod{p^a}$ queda reducida al estudio de $f(x) \equiv 0 \pmod{p}$ (aquí entra en juego el teorema de Lagrange) y a congruencias lineales (que se resuelven con el algoritmo de Euclides o aplicando el teorema 2.8).

3.2. Restos cuadráticos y ley de reciprocidad cuadrática

El caso más sencillo de polinomios no lineales es $f(x) = x^2 - b$. El teorema de Lagrange nos dice que, cuando p es un número primo, la ecuación con congruencias

$$x^2 \equiv b \pmod{p} \tag{3.11}$$

tiene, como máximo, dos soluciones (que serían dos «raíces cuadradas» de b módulo p, aunque esta nomenclatura no se suele usar). Pero parece razonable esperar que, para esta ecuación tan sencilla, podamos obtener más información. Comencemos analizando los casos triviales.

Si $p = 2$, todo es evidente, así que en esta sección p será siempre un primo impar. Por otra parte, si $b \equiv 0 \pmod{p}$ la solución de (3.11) —única módulo p, pues en $\mathbb{Z}_p$ no hay divisores de 0— es $x = 0$. En consecuencia, podremos desde ahora asumir que $b \not\equiv 0 \pmod{p}$; o, incluso, que $0 < b \leq p-1$. Una vez hechas esas consideraciones, nótese además que, si x satisface (3.11), también lo hace $-x$; así que el número de soluciones de la ecuación es forzosamente 0 o 2.

Cuando (3.11) tiene solución, se dice que b es *resto cuadrático* módulo p. Si, por el contrario, (3.11) no tiene solución, diremos que b no es un resto cuadrático módulo p. Hay dos problemas básicos que interesa resolver:

- Dado un primo p, saber identificar siempre si un b es o no resto cuadrático módulo p.
- Dado un entero b, identificar todos los primos p tales que b es un resto cuadrático módulo p.

Es fácil dar un primer resultado, que nos dice cuántos restos cuadráticos módulo p hay. Para contarlos, lo estamos haciendo como elementos de $\mathbb{Z}_p$; esto es, si $b \equiv c \pmod{p}$, no los distinguimos.

Teorema 3.5.
Sea p un primo impar. Entre los enteros $1, 2, \ldots, p-1$, el número de restos cuadráticos módulo p es $(p-1)/2$; los otros $(p-1)/2$ números no lo son. En concreto, los restos cuadráticos son los que corresponden a las clases (módulo p) de los números

$$1^2,\ 2^2,\ 3^2,\ \ldots,\ \left(\frac{p-1}{2}\right)^2. \tag{3.12}$$

Demostración. En primer lugar, comprobemos que los $(p-1)/2$ números de (3.12) son distintos módulo p. En efecto, si dos enteros x e y con $1 \leq x \leq (p-1)/2$ y

$1 \leq y \leq (p-1)/2$ verifican $x^2 \equiv y^2 \pmod{p}$, entonces $(x-y)(x+y) \equiv 0 \pmod{p}$. Pero $1 < x+y < p$, y p es primo (o sea, en $\mathbb{Z}_p$ no hay divisores de 0), así que forzosamente tiene que ser $x - y \equiv 0 \pmod{p}$, luego $x = y$.

Para concluir la demostración basta darse cuenta de que, cualquiera que sea el número $x \not\equiv 0 \pmod{p}$, su cuadrado debe ser congruente módulo p con alguno de los números de (3.12), ya que $(p-k)^2 \equiv k^2 \pmod{p}$. □

A partir de aquí, las cosas —aunque no son complicadas— requieren un poco más de trabajo; de hecho, han sido los esfuerzos de grandes matemáticos como Euler, Legendre, Gauss y Jacobi los que han logrado llevar la teoría a su estado actual. En lo que sigue haremos un desarrollo razonablemente completo y autocontenido, pero dejaremos sin demostrar algunos resultados que no son fundamentales en nuestras argumentaciones. El lector puede ampliar lo que aquí se explica en muchísimos libros de teoría de números. Por ejemplo, [5, capítulo 9] o [88, capítulo 11].

3.2.1. El símbolo de Legendre

Esencial en todo esto es el denominado *símbolo de Legendre* que, para p un primo impar y $b \not\equiv 0 \pmod{p}$, se define como

$$(b\,|\,p) = \begin{cases} 1, & \text{si } b \text{ es resto cuadrático módulo } p, \\ -1, & \text{si } b \text{ no es resto cuadrático módulo } p. \end{cases} \tag{3.13}$$

Cuando $b \equiv 0 \pmod{p}$, se toma $(b \mid p) = 0$. Con este nuevo lenguaje, nuestro objetivo de conocer si b es o no resto cuadrático módulo p se ha convertido en saber identificar $(b\,|\,p)$. Antes de seguir, conviene advertir al lector que quiera consultar otros textos que, en muchos de ellos, el *símbolo de Legendre* aparece denotado[1] como $\left(\frac{b}{p}\right)$. Fijémonos además en que, para p fijo, la función $(\cdot\,|\,p)$ está bien definida en $\mathbb{Z}_p$, ya que $(b\,|\,p) = (c\,|\,p)$ siempre que $b \equiv c \pmod{p}$.

El siguiente resultado nos proporciona una manera directa de calcular $(b\,|\,p)$; de hecho, es el que hace que la definición (3.13) sea útil. Para que su enunciado no resulte confuso, aclaremos previamente que $b^{(p-1)/2}$ siempre vale 1, -1 o 0 módulo p: en efecto, el teorema de Euler-Fermat (incluso basta aplicar el caso particular denominado teorema pequeño de Fermat) nos asegura que, cuando $b \not\equiv 0 \pmod{p}$,

$$(b^{(p-1)/2}+1)(b^{(p-1)/2}-1) = b^{p-1} - 1 \equiv 0 \pmod{p}, \tag{3.14}$$

y por tanto el primo p debe dividir a uno de los dos factores $b^{(p-1)/2}+1$ o $b^{(p-1)/2}-1$, es decir, $b^{(p-1)/2} \equiv \pm 1 \pmod{p}$.

[1] Aquí hemos preferido no seguir esa notación que, en ocasiones, puede ser confundida con una fracción.

Teorema 3.6. Criterio de Euler.
Sea p un primo impar y b un entero. Entonces, $(b\,|\,p) \equiv b^{(p-1)/2} \pmod{p}$.

Demostración. Si $b \equiv 0 \pmod{p}$ el resultado es trivial, pues $(b\,|\,p) = 0$ y $b^{(p-1)/2}$ es múltiplo de p. El caso $(b\,|\,p) = 1$ también es muy sencillo: por definición, existe un entero x tal que $x^2 \equiv b \pmod{p}$, luego, por el teorema de Euler-Fermat (de nuevo basta usar el teorema pequeño de Fermat),

$$b^{(p-1)/2} \equiv (x^2)^{(p-1)/2} = x^{p-1} \equiv 1 \pmod{p},$$

así que efectivamente $b^{(p-1)/2} \equiv 1 = (b\,|\,p) \pmod{p}$.

Finalmente, analicemos el caso $(b\,|\,p) = -1$. Como el polinomio $f(x) = x^{(p-1)/2} - 1$ tiene grado $(p-1)/2$, la congruencia

$$f(x) \equiv 0 \pmod{p} \tag{3.15}$$

tiene, a lo sumo, $(p-1)/2$ soluciones (teorema 3.1). El teorema 3.5 nos asegura que hay exactamente $(p-1)/2$ restos cuadráticos módulo p, y del razonamiento usado en el caso $(b\,|\,p) = 1$ se desprende que todos ellos son soluciones de (3.15), luego los otros $(p-1)/2$ no restos cuadráticos no pueden serlo. Así pues, un b que cumpla $(b \mid p) = -1$ no puede ser raíz de f, es decir, no puede cumplir $b^{(p-1)/2} \equiv 1 \pmod{p}$. Como, según hemos argumentado en (3.14), sólo existen las posibilidades $b^{(p-1)/2} \equiv 1 \pmod{p}$ o $b^{(p-1)/2} \equiv -1 \pmod{p}$, forzosamente deberá ser $b^{(p-1)/2} \equiv -1 \pmod{p}$. □

Tal como estábamos comentando, ahí tenemos una fórmula explícita para conocer el valor de $(b\,|\,p)$. Recordemos además al lector que calcular $b^{(p-1)/2}$ módulo p con p grande no es tan duro como parece a simple vista, pues se puede aplicar el cálculo de potencias mediante cuadrados sucesivos (también llamado algoritmo de cuadrados iterados), que ya utilizamos a menudo en el tema anterior (como en los ejemplos de las páginas 101 y 126), y que volveremos a usar más adelante en este mismo tema (en el ejemplo de la página 195). Pero, además, el criterio de Euler permite manejar $(b\,|\,p)$ de una manera más cómoda, lo que ayuda a probar diversas propiedades del símbolo de Legendre. Por ejemplo, sirve para deducir que[2]

$$(bc\,|\,p) = (b\,|\,p)(c\,|\,p) \tag{3.16}$$

para todo $b, c \in \mathbb{Z}$ (ejercicio 3.10).

[2] En el lenguaje del tema 7, esta fórmula nos dice que, para p un primo impar fijo, la función $\chi(n) = (n \mid p)$ es completamente multiplicativa; dicha función $(\cdot \mid p)$ se denomina *carácter cuadrático* módulo p.

Otra consecuencia inmediata es el siguiente resultado, que se obtiene sin más que tomar $b = -1$:

$$(-1\,|\,p) = (-1)^{(p-1)/2} = \begin{cases} 1, & \text{si } p \equiv 1 \pmod 4, \\ -1, & \text{si } p \equiv 3 \pmod 4. \end{cases} \tag{3.17}$$

Como procedimiento puramente operacional destinado a evaluar $(b \mid p)$, aún se puede mejorar el criterio de Euler. Como paso previo en este camino, resulta fundamental el siguiente resultado:

Teorema 3.7. Lema de Gauss.
Sea p un primo impar y b un entero. Supongamos que $b \not\equiv 0 \pmod p$ y, para cada uno de los números

$$b,\ 2b,\ 3b,\ \ldots,\ \frac{p-1}{2}\,b,$$

consideremos su mínimo resto positivo módulo p. Si s indica cuántos de estos restos son mayores que $p/2$, entonces $(b|p) = (-1)^s$.

Demostración. Descompongamos el conjunto $T = \{1, 2, 3, \ldots, (p-1)/2\}$ en dos subconjuntos disjuntos $C = \{C_1, C_2, \ldots, C_r\}$ y $D = \{D_1, D_2, \ldots, D_s\}$ (con $s = (p-1)/2 - r$) según el siguiente criterio:

- En C nos quedamos con los $t \in T$ tales que el resto de tb módulo p es menor que $p/2$.
- Y en D tomamos los $t \in T$ tales que el resto de tb módulo p es mayor que $p/2$.

Así pues, $C_j b \equiv c_j \pmod p$ con $0 < c_j < p/2$ $(j = 1, 2, \ldots, r)$ y $D_j b \equiv d_j \pmod p$ con $p/2 < d_j < p$ $(j = 1, 2, \ldots, s)$, y todos los números

$$c_1, c_2, \ldots, c_r, p - d_1, p - d_2, \ldots, p - d_s \tag{3.18}$$

son positivos y menores que $p/2$. Además, son números distintos:

- Si $c_j = c_k$ tendríamos $C_j b \equiv C_k b \pmod p$, pero eso (dado que $b \not\equiv 0 \pmod p$ con p primo) implicaría que $C_j = C_k$.
- Si $p - d_j = p - d_k$ también sería $d_j = d_k$ y de nuevo obtendríamos $D_j = D_k$.
- Si $c_j = p - d_k$, deberá ser $c_j + d_k = p \equiv 0 \pmod p$, de donde $0 \equiv c_j + d_k \equiv (C_j + D_k)b \pmod p$, y por tanto también $C_j + D_k \equiv 0 \pmod p$; pero eso es imposible ya que que C_j y D_k son números positivos menores que $p/2$.

Así pues, los números listados en (3.18) son todos los enteros entre 1 y $(p-1)/2$ (ambos incluidos), y su producto es $\left(\frac{p-1}{2}\right)!$. Pero también

$$C_1C_2\cdots C_rD_1D_2\cdots D_s = 1\cdot 2\cdots\frac{p-1}{2} = \left(\frac{p-1}{2}\right)!.$$

En consecuencia, módulo p se cumple

$$\begin{aligned} b^{(p-1)/2}\left(\frac{p-1}{2}\right)! &= (C_1b)(C_2b)\cdots(C_rb)(D_1b)(D_2b)\cdots(D_sb)\\ &\equiv c_1c_2\cdots c_rd_1d_2\cdots d_s\\ &\equiv (-1)^s c_1c_2\cdots c_r(p-d_1)(p-d_2)\cdots(p-d_s)\\ &\equiv (-1)^s\left(\frac{p-1}{2}\right)! \pmod p. \end{aligned}$$

Al ser p primo, podemos simplificar $\left(\frac{p-1}{2}\right)!$ módulo p, y por tanto $b^{(p-1)/2}\equiv(-1)^s \pmod p$. Aplicando el criterio de Euler, $(b\,|\,p)\equiv b^{(p-1)/2} \pmod p$, obtenemos $(b\,|\,p)\equiv(-1)^s \pmod p$ y, de aquí, $(b\,|\,p)=(-1)^s$. □

De momento, el último avance en el cálculo efectivo de $(b\,|\,p)$ requiere simplemente observar que, en la práctica, no es necesario conocer el valor exacto de s para usarlo en $(-1)^s$, sino únicamente su paridad. Y esto se puede conseguir —de manera relativamente sencilla— con la siguiente fórmula:

Teorema 3.8.
El número s que aparece en el lema de Gauss cumple[3]

$$s \equiv \sum_{t=1}^{(p-1)/2}\left\lfloor\frac{tb}{p}\right\rfloor + (b-1)\,\frac{p^2-1}{8} \pmod 2, \tag{3.19}$$

donde $\lfloor\cdot\rfloor$ denota la parte entera de un número real.

Demostración. Siguiendo con la notación que teníamos en la demostración del teorema anterior, escribamos las relaciones $C_jb\equiv c_j \pmod p$ con $0<c_j<p/2$ $(j=1,2,\ldots,r)$ y $D_jb\equiv d_j \pmod p$ con $p/2<d_j<p$ $(j=1,2,\ldots,s)$ como

$$C_jb = \left\lfloor\frac{C_jb}{p}\right\rfloor p + c_j \quad\text{y}\quad D_jb = \left\lfloor\frac{D_jb}{p}\right\rfloor p + d_j.$$

[3] Es fácil comprobar que el $(p^2-1)/8$ que aparece en la fórmula es siempre un entero (véase, por ejemplo, el ejercicio 2.2, en el tema anterior).

Sean asimismo

$$m = \left\lfloor \frac{b}{p} \right\rfloor + \left\lfloor \frac{2b}{p} \right\rfloor + \cdots + \left\lfloor \frac{p-1}{2} \frac{b}{p} \right\rfloor,$$
$$c = c_1 + c_2 + \cdots + c_r \quad \text{y} \quad d = d_1 + d_2 + \cdots + d_s.$$

Sumando estas expresiones, es claro que

$$(C_1 + C_2 + \cdots + C_r + D_1 + D_2 + \cdots + D_s)b = pm + c + d. \tag{3.20}$$

Recordemos que la suma de una progresión aritmética $x_1 + x_2 + \cdots + x_k$ es $\frac{x_k + x_1}{2} k$. Entonces,

$$\begin{aligned} &C_1 + C_2 + \cdots + C_r + D_1 + D_2 + \cdots + D_s \\ &\qquad = 1 + 2 + \cdots + \frac{p-1}{2} = \frac{\frac{p-1}{2} + 1}{2} \frac{p-1}{2} = \frac{p+1}{4} \frac{p-1}{2} = \frac{p^2-1}{8} \end{aligned}$$

y (3.20) queda

$$\frac{p^2-1}{8} b = pm + c + d. \tag{3.21}$$

Por otra parte, tengamos en cuenta que los números listados en (3.18) son todos los enteros entre 1 y $(p-1)/2$ (ambos incluidos), con lo cual su suma es

$$\begin{aligned} \frac{p^2-1}{8} &= 1 + 2 + \cdots + \frac{p-1}{2} \\ &= c_1 + c_2 + \cdots + c_r + (p - d_1) + (p - d_2) + \cdots + (p - d_s) \\ &= c + ps - d. \end{aligned} \tag{3.22}$$

Restando (3.21) y (3.22) llegamos a

$$(b-1)\frac{p^2-1}{8} = mp - sp + 2d,$$

luego

$$(b-1)\frac{p^2-1}{8} - mp \equiv -sp \pmod 2.$$

Ahora, sin más que fijarse en que $p \equiv -1 \pmod 2$, obtenemos (3.19). □

En particular, si en (3.19) tomamos $b = 2$, todos los sumandos $\lfloor tb/p \rfloor$ se anulan. De ahí se sigue que

$$(2\,|\,p) = (-1)^{(p^2-1)/8} = \begin{cases} 1, & \text{si } p \equiv \pm 1 \pmod 8, \\ -1, & \text{si } p \equiv \pm 3 \pmod 8. \end{cases} \tag{3.23}$$

Existen demostraciones alternativas de este resultado que no son una mera aplicación del lema de Gauss; pero, claro, resultan considerablemente más laboriosas (véase, por ejemplo, [5, teorema 9.5]).

Merece la pena finalizar este apartado recalcando que, como consecuencia de (3.16), el símbolo de Legendre $(\cdot\,|\,p)$ para p un primo fijo está completamente determinado por sus valores en -1, 2 y los primos impares. En concreto, si b es un entero no divisible por p y escribimos $b = \pm 2^{c_0} q_1^{c_1} q_2^{c_2} \cdots q_k^{c_k}$ (con los q_j primos distintos entre sí y distintos de p), entonces

$$(b\,|\,p) = (\pm 1\,|\,p)(2\,|\,p)^{c_0}(q_1\,|\,p)^{c_1}(q_2\,|\,p)^{c_2} \cdots (q_k\,|\,p)^{c_k}.$$

Aquí, además, los valores de $(\pm 1\,|\,p)$ y $(2\,|\,p)$ están ya dados en (3.17) y (3.23). Así pues, si asumimos que somos capaces de encontrar la descomposición en factores primos,[4] calcular $(b\,|\,p)$ se reduce a conocer $(q\,|\,p)$ con q primo impar.

Aún podemos simplificar esto con una observación adicional: a^2 siempre es —trivialmente— resto cuadrático módulo p, y por tanto $(a^2\,|\,p) = 1$ (salvo que a sea múltiplo de p, que no tiene interés ahora). Así pues, si b es un entero no divisible por p, podemos agrupar todos sus factores con exponente par y escribirlo como $b = \pm 2^{\varepsilon} a^2 q_1 q_2 \cdots q_k$ con $\varepsilon = 0$ o 1 y $k \geq 0$ (donde los q_j denotan primos distintos entre sí y distintos de p, y $k = 0$ significaría que no hay ningún q_j). De este modo, tendremos

$$(b\,|\,p) = (\pm 1\,|\,p)(2\,|\,p)^{\varepsilon}(q_1\,|\,p)(q_2\,|\,p) \cdots (q_k\,|\,p),$$

que da el símbolo de Legendre $(b\,|\,p)$ a partir de los símbolos de los primos que aparecen un número impar de veces en la descomposición en factores de b.

3.2.2. La ley de reciprocidad cuadrática

Hasta ahora hemos resuelto —de forma bastante satisfactoria— el primer problema de los dos que planteábamos en la página 176. El segundo es mucho más difícil.

[4] No nos referimos a que tal descomposición exista —eso es el teorema fundamental de la aritmética—, sino a que podamos hallarla de manera efectiva, lo cual no siempre es fácil si los números involucrados son «grandes».

Su solución depende de un resultado notable, conocido como ley de reciprocidad cuadrática. Dados p y q primos impares distintos, que la ecuación en congruencias

$$x^2 \equiv q \pmod{p}$$

tenga solución (es decir, que q sea un resto cuadrático módulo p) está de hecho relacionado —de una manera muy precisa que ya aclararemos— con que también la tenga la ecuación

$$x^2 \equiv p \pmod{q}$$

(es decir, que p sea un resto cuadrático módulo q).

Fue Euler —alrededor de 1745— el primero que lo intuyó (aunque descrito de una manera más complicada a como aquí lo expondremos); tenía evidencias numéricas de que la ley de reciprocidad cuadrática debía cumplirse, pero no logró probarlo. En 1785 lo redescubrió Legendre, quien tampoco encontró una demostración correcta. Poco después, Gauss entra en escena. En 1795 —cuando tenía 18 años de edad, y aparentemente sin conocer los esfuerzos de sus predecesores— vuelve a adivinar el resultado. En sus propias palabras, el teorema le estuvo atormentando todo un año, en el que le absorbió sus mayores esfuerzos, hasta que consiguió probarlo. Gauss estaba fascinado con la ley de reciprocidad cuadrática; a lo largo de su vida dio otras siete demostraciones, basadas en distintos procedimientos. Desde entonces, muchos matemáticos han disfrutado buscando pruebas alternativas; se conocen más de 200 (en [63, apéndice B] aparece una lista de 192 demostraciones, cada una con autor, año de descubrimiento y método utilizado; y desde la publicación de ese libro se han encontrado algunas otras).

Nosotros vamos a enunciar la ley de reciprocidad cuadrática mediante la siguiente formulación simétrica dada por Legendre, que es como se establece usualmente, y nos contentaremos con proporcionar una sola demostración. En concreto, una dada por el joven matemático Ferdinand Gotthold Max Eisenstein en 1844, que reduce la prueba a contar los puntos de coordenadas enteras de dos triángulos, y que es una simplificación de una de las originales de Gauss:

Teorema 3.9. Ley de reciprocidad cuadrática.
Si p y q son primos impares, se cumple

$$(p\,|\,q)(q\,|\,p) = (-1)^{(p-1)(q-1)/4}. \tag{3.24}$$

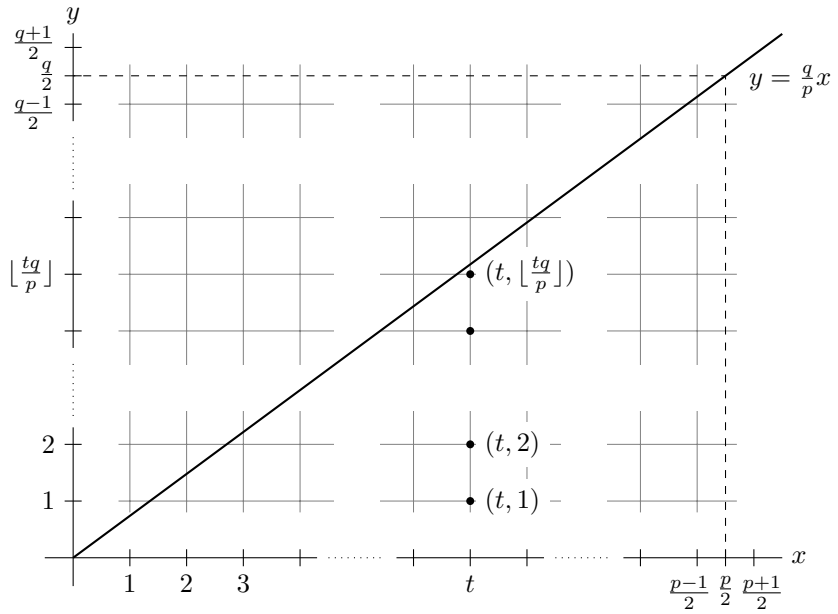


Figura 3.1: Puntos de coordenadas enteras del rectángulo $1 \le x \le \frac{p}{2}$, $1 \le y \le \frac{q}{2}$. Aparecen marcados los puntos (t, n) con $n = 1, 2, \ldots, \lfloor \frac{tq}{p} \rfloor$.

Demostración. Tomemos $b = q$ en el teorema 3.8. Como $(q-1)(p^2-1)/8$ es par, el lema de Gauss queda

$$(q\,|\,p) = (-1)^s \quad \text{con} \quad s = \left\lfloor \frac{q}{p} \right\rfloor + \left\lfloor \frac{2q}{p} \right\rfloor + \cdots + \left\lfloor \frac{p-1}{2}\frac{q}{p} \right\rfloor.$$

Además, intercambiando los papeles de p y q tendremos

$$(p\,|\,q) = (-1)^{s'} \quad \text{con} \quad s' = \left\lfloor \frac{p}{q} \right\rfloor + \left\lfloor \frac{2p}{q} \right\rfloor + \cdots + \left\lfloor \frac{q-1}{2}\frac{p}{q} \right\rfloor.$$

Por tanto, $(p\,|\,q)(q\,|\,p) = (-1)^{s+s'}$ y la demostración estará completa si vemos que

$$s + s' = \frac{p-1}{2}\frac{q-1}{2}.$$

Para comprobarlo, vamos a ayudarnos de la gráfica de figura 3.1.

Por una parte, es claro que en el rectángulo $1 \leq x \leq \frac{p}{2}$, $1 \leq y \leq \frac{q}{2}$ hay $\frac{p-1}{2}\frac{q-1}{2}$ puntos de coordenadas enteras. Por la otra, vamos a ver que por debajo de la recta $y = \frac{q}{p}x$ hay s de tales puntos; y, por encima de ella, s' (no puede haber ninguno sobre la recta ya que $\mcd(p,q) = 1$).

En la región del rectángulo situada por debajo de esa recta, detengámonos en la recta vertical $x = t$ (con t un entero entre 1 y $p/2$), cuya intersección con $y = \frac{q}{p}x$ es $(t, \frac{tq}{p})$. Sobre esa recta vertical están los puntos $(t,1)$, $(t,2)$,... hasta $(t, \lfloor \frac{tq}{p} \rfloor)$; es decir, hay $\lfloor \frac{tq}{p} \rfloor$ puntos de coordenadas enteras. Contando los puntos de todas las rectas verticales, en esta región hay

$$\left\lfloor \frac{q}{p} \right\rfloor + \left\lfloor \frac{2q}{p} \right\rfloor + \cdots + \left\lfloor \frac{tq}{p} \right\rfloor + \cdots + \left\lfloor \frac{p-1}{2}\frac{q}{p} \right\rfloor = s$$

puntos. Finalmente, de manera similar (o intercambiado los papeles de p y q) se obtiene que en la región del rectángulo por encima de la recta $y = \frac{q}{p}x$ hay s' puntos. □

Si nos fijamos, la fórmula (3.24) es una forma resumida de decir lo siguiente (válido para p y q primos impares):

- Si p y q son congruentes con 3 módulo 4, entonces p es resto cuadrático módulo q si y sólo si q no es resto cuadrático módulo p.
- En otro caso, p es resto cuadrático módulo q si y sólo si q es resto cuadrático módulo p.

Como vemos a continuación, aplicar la ley de reciprocidad cuadrática resulta a menudo muy útil para calcular el símbolo de Legendre con poco trabajo:

Ejemplo. Calcular $(7|73)$.

Solución. Los números 73 y 7 son ambos primos, y verifican $7 \equiv 3 \pmod 4$ y $73 \equiv 1 \pmod 4$, así que

$$(7|73) = (73|7) = (3|7) = -(7|3) = -(1|3) = -1$$

(nótese que hemos aplicado la ley de reciprocidad cuadrática en la primera y la tercera igualdades). ◊

Ejemplo. Calcular $(33|97)$.

Solución. Como 97 es primo pero 33 no, descomponemos $33 = 11 \cdot 3$. Entonces, aplicando las propiedades del símbolo de Legendre (entre ellas, la ley de reciprocidad cuadrática, teniendo en cuenta que $97 \equiv 1 \pmod 4$ y $11 \equiv 3 \pmod 4$), obtenemos

$$(33|97) = (3|97)(11|97) = (97|3)(97|11) = (1|3)(9|11) = 1 \cdot 1 = 1. \qquad \text{◊}$$

La ley de reciprocidad cuadrática da, además, una respuesta parcial al segundo problema de los planteados en la página 176. En efecto, si sólo consideramos los b que sean primos (es decir, $b = q$, el de la ley de reciprocidad cuadrática), la fórmula (3.24) representa una especie de dualidad entre el primer problema y el segundo. Conociendo $(p\,|\,q)$, es inmediato calcular $(q\,|\,p)$.

Ejemplo. Identificar todos los primos p (impares) tales que 5 es resto cuadrático módulo p (para evitar trivialidades, podemos suponer $p \neq 5$).

Solución. Como $5 \equiv 1 \pmod 4$, la ley de reciprocidad cuadrática siempre da $(p \mid 5)(5 \mid p) = 1$, y por tanto $(5 \mid p) = (p \mid 5)$ para cualquier primo impar p. De aquí que

$$(5\,|\,p) = (p\,|\,5) = \begin{cases} 1, & \text{si } p \equiv 1 \text{ o } 4 \pmod 5, \\ -1, & \text{si } p \equiv 2 \text{ o } 3 \pmod 5. \end{cases}$$

Así pues, 5 es resto cuadrático módulo p si y sólo si $p \equiv 1$ o $4 \pmod 5$. ◊

Ejemplo. Identificar todos los primos p (impares) tales que 7 es resto cuadrático módulo p (para evitar trivialidades, podemos suponer $p \neq 7$).

Solución. Como $7 \equiv 3 \pmod 4$, la ley de reciprocidad cuadrática da $(7\,|\,p) = (p\,|\,7)$ cuando $p \equiv 1 \pmod 4$ y $(7\,|\,p) = -(p\,|\,7)$ cuando $p \equiv 3 \pmod 4$. Analicemos ambos casos por separado. Si $p \equiv 1 \pmod 4$, tenemos

$$(7\,|\,p) = (p\,|\,7) = \begin{cases} 1, & \text{si } p \equiv 1,\ 2 \text{ o } 4 \pmod 7, \\ -1, & \text{si } p \equiv 3,\ 5 \text{ o } 6 \pmod 7; \end{cases}$$

y, si $p \equiv 3 \pmod 4$,

$$(7\,|\,p) = -(p\,|\,7) = \begin{cases} 1, & \text{si } p \equiv 3,\ 5 \text{ o } 6 \pmod 7, \\ -1, & \text{si } p \equiv 1,\ 2 \text{ o } 4 \pmod 7. \end{cases}$$

Finalmente, con el teorema chino del resto (teorema 2.15 y corolario 2.16, ambos en el tema anterior) podemos combinar cada pareja de condiciones $x \equiv a \pmod 4$ y $x \equiv b \pmod 7$ en una sola $x \equiv c \pmod{28}$. Aplicando este procedimiento, llegaríamos a

$$(7\,|\,p) = \begin{cases} 1, & \text{si } p \equiv 1,\ 3,\ 9,\ 19,\ 25 \text{ o } 27 \pmod{28}, \\ -1, & \text{si } p \equiv 5,\ 11,\ 13,\ 15,\ 17 \text{ o } 23 \pmod{28}, \end{cases}$$

y el caso $(7\,|\,p) = 1$ es el que proporciona los primos p buscados. ◊

Incluso se puede aplicar la ley de reciprocidad cuadrática al análisis del segundo problema cuando b no es primo pero conocemos su descomposición en factores primos. Aunque —hay que reconocerlo— el resultado se va haciendo más y más engorroso según aumenta el número de factores de b.

Ejemplo. Identificar todos los primos p (impares y con $5 \neq p \neq 7$) tales que 35 es resto cuadrático módulo p.

Solución. Queremos ver cuándo $(35\,|\,p) = 1$. Al ser $35 = 5 \cdot 7$ y por tanto

$$(35\,|\,p) = (5\,|\,p)(7\,|\,p),$$

es claro que

$$(35\,|\,p) = 1 \;\Leftrightarrow\; (5\,|\,p) = (7\,|\,p) = 1 \text{ o } (5\,|\,p) = (7\,|\,p) = -1.$$

Entonces, sin más que aplicar lo que calculamos en los dos ejemplos anteriores, llegamos a que 35 es resto cuadrático módulo p si y sólo si

$$p \equiv 1 \text{ o } 4 \pmod 5 \quad \text{y} \quad p \equiv 1,\ 3,\ 9,\ 19,\ 25 \text{ o } 27 \pmod{28},$$

o

$$p \equiv 2 \text{ o } 3 \pmod 5 \quad \text{y} \quad p \equiv 5,\ 11,\ 13,\ 15,\ 17 \text{ o } 23 \pmod{28}.$$

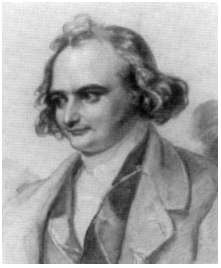
Y, si lo deseáramos, podríamos aplicar el teorema chino del resto para combinar los restos módulo 5 y 28 y expresar así el resultado como un conjunto de congruencias módulo $5 \cdot 28 = 140$. ◊

3.2.3. El símbolo de Jacobi

Según hemos visto hasta ahora, para determinar si un número compuesto b es o no un resto cuadrático módulo p hay que considerar varios casos que dependen del carácter cuadrático de los factores de b. Se puede evitar gran parte de estos cálculos usando una generalización del símbolo de Legendre introducida por Jacobi. De momento, comencemos explicando en qué consiste esta generalización y analizando algunas de sus propiedades básicas, y ya veremos un poco más adelante por qué sirve de ayuda para agilizar los cálculos.

Si $P > 1$ es un entero impar cuya descomposición en factores primos distintos es $P = \prod_{j=1}^{s} p_j^{a_j}$, y b es un número entero, el *símbolo de Jacobi* es, por definición,

$$(b\,|\,P) = \prod_{j=1}^{s} (b\,|\,p_j)^{a_j}, \tag{3.25}$$



Carl Gustav Jacob Jacobi (1804–1851)



Leopold Kronecker (1823–1891)

donde $(b\,|\,p_j)$ es el símbolo de Legendre. Si $P = 1$, se toma $(b\,|\,P) = 1$. Resulta obvio que $(b\,|\,P)$ sólo puede valer 1, -1 o 0 (esto último si y sólo si $\mathrm{mcd}(b, P) > 1$).

A partir de su definición, es claro que, si escribimos $P = p_1 p_2 \cdots p_s$, donde los factores primos impares p_j no son necesariamente distintos, la expresión (3.25) queda

$$(b\,|\,P) = \prod_{j=1}^{s} (b\,|\,p_j).$$

Además, son inmediatas de comprobar las siguientes propiedades, en las que P y Q denotan enteros positivos impares, y b y c enteros arbitrarios:

- $(b\,|\,P) = (c\,|\,P)$ si $b \equiv c \pmod{P}$,
- $(bc\,|\,P) = (b\,|\,P)(c\,|\,P)$,
- $(b\,|\,PQ) = (b\,|\,P)(b\,|\,Q)$,
- $(b^2\,|\,P) = 1$ si $\mathrm{mcd}(b, P) = 1$.

También son ciertas las siguientes fórmulas, análogas a las del símbolo de Legendre:

Proposición 3.10.

Si P es un entero positivo impar, entonces

$$(-1\,|\,P) = (-1)^{(P-1)/2}, \tag{3.26}$$

$$(2\,|\,P) = (-1)^{(P^2-1)/8}. \tag{3.27}$$

Demostración. Escribamos $P = p_1 p_2 \cdots p_s$, donde los factores primos p_j no son necesariamente distintos, como

$$P = \prod_{j=1}^{s} \big(1 + (p_j - 1)\big) = 1 + \sum_{j=1}^{s} (p_j - 1) + \sum_{j \neq k} (p_j - 1)(p_k - 1) + \cdots$$

Cada factor $p_j - 1$ es par, así que

$$P \equiv 1 + \sum_{j=1}^{s} (p_j - 1) \pmod 4,$$

y de aquí

$$\frac{1}{2}(P-1) \equiv \sum_{j=1}^{s} \frac{1}{2}(p_j - 1) \pmod 2. \tag{3.28}$$

Con esto, aplicando (3.17) a cada factor se sigue (3.26):

$$(-1\,|\,P) = \prod_{j=1}^{s} (-1\,|\,p_j) = \prod_{j=1}^{s} (-1)^{(p_j-1)/2} = (-1)^{(P-1)/2}.$$

Para probar (3.27) escribamos

$$P^2 = \prod_{j=1}^{s} p_j^2 = \prod_{j=1}^{s} \big(1 + (p_j^2 - 1)\big) = 1 + \sum_{j=1}^{s} (p_j^2 - 1) + \sum_{j \neq k} (p_j^2 - 1)(p_k^2 - 1) + \cdots.$$

Como cada p_j es impar, $p_j^2 - 1 = (p_j + 1)(p_j - 1)$ es múltiplo de 8, así que

$$P^2 \equiv 1 + \sum_{j=1}^{s} (p_j^2 - 1) \pmod{64},$$

y por tanto

$$\frac{1}{8}(P^2 - 1) \equiv \sum_{j=1}^{s} \frac{1}{8}(p_j^2 - 1) \pmod 8;$$

en particular, también

$$\frac{1}{8}(P^2 - 1) \equiv \sum_{j=1}^{s} \frac{1}{8}(p_j^2 - 1) \pmod 2.$$

Con esto, de (3.23) obtenemos

$$(2\,|\,P) = \prod_{j=1}^{s} (2\,|\,p_j) = \prod_{j=1}^{s} (-1)^{(p_j^2-1)/8} = (-1)^{(P^2-1)/8}. \qquad \square$$

Con el símbolo de Jacobi ya no es verdad que $(b\,|\,P)$ valga 1 o -1 según que la ecuación $x^2 \equiv b \pmod P$ tenga o no soluciones (por supuesto, aquí nos estamos refiriendo a cuando $b \not\equiv 0 \pmod P$). Puede que $(b\,|\,P)$ sea 1 únicamente porque en (3.25) hay un número par de factores -1. Se propone al lector que busque un ejemplo que ilustre este comportamiento (ejercicio 3.14). Tampoco es cierto en general que el número de soluciones sea 0 o 2 (acuérdese el lector del teorema 3.3).

Esto nos puede hacer pensar que $(b\,|\,P)$ ya no va a tener ninguna utilidad para determinar restos cuadráticos, pero no es así. La realidad es que, aunque el valor del símbolo de Jacobi ya no está relacionado con el resto cuadrático, va a servir como procedimiento operacional para calcular otros símbolos (esta vez, de Legendre) que sí se corresponden con restos cuadráticos. Y es que la ley de reciprocidad cuadrática también se extiende a los símbolos de Jacobi:

Teorema 3.11. Ley de reciprocidad cuadrática para símbolos de Jacobi.
Si P y Q son enteros positivos impares con $\mcd(P,Q)=1$, *se cumple*

$$(P\,|\,Q)(Q\,|\,P) = (-1)^{(P-1)(Q-1)/4}.$$

Demostración. Representemos las descomposiciones de P y Q en factores primos como $P = p_1p_2\cdots p_s$ y $Q = q_1q_2\cdots q_t$ (nótese que, por ser $\mcd(P,Q)=1$, se cumplirá $p_j \neq q_k$ para todo j,k). Aplicando la definición de símbolo de Jacobi, la propiedad $(bc\,|\,P) = (b\,|\,P)(c\,|\,P)$, y la ley de reciprocidad cuadrática para símbolos de Legendre, tenemos

$$(P\,|\,Q)(Q\,|\,P) = \prod_{j=1}^{s}\prod_{k=1}^{t}(p_j\,|\,q_k)(q_k\,|\,p_j) = (-1)^r$$

con

$$r = \sum_{j=1}^{s}\sum_{k=1}^{t}(p_j-1)(q_k-1)/4.$$

Escribamos este r como

$$r = \sum_{j=1}^{s}\sum_{k=1}^{t}\frac{1}{2}(p_j-1)\frac{1}{2}(q_k-1) = \left(\sum_{j=1}^{s}\frac{1}{2}(p_j-1)\right)\left(\sum_{k=1}^{t}\frac{1}{2}(q_k-1)\right)$$

y tengamos en cuenta que, tal como hemos visto en (3.28) (con su análogo para Q),

$$\frac{1}{2}(P-1) \equiv \sum_{j=1}^{s}\frac{1}{2}(p_j-1) \pmod 2, \qquad \frac{1}{2}(Q-1) \equiv \sum_{k=1}^{t}\frac{1}{2}(q_k-1) \pmod 2.$$

En consecuencia,

$$r \equiv \frac{1}{2}(P-1)\frac{1}{2}(Q-1) \pmod 2,$$

y esto prueba el teorema. □

Nota 3.12. Concluyamos mencionando que aún existen extensiones del símbolo de Jacobi. Por ejemplo, el denominado *símbolo de Kronecker*, en el que, con algunas condiciones, se llega a definir $(b\,|\,P)$ cuando P es un entero que puede ser par. Pero, a decir verdad, sus propiedades —que no vamos a exponer— ya no son tan útiles. Si lo desea, el lector puede ver algo de esto en [88, § 11.3]. ◁

3.2.4. Cálculo rápido del símbolo de Jacobi

Fijémonos en las siguientes propiedades del símbolo de Jacobi, todas ellas ya probadas en el apartado anterior (aquí, n denota un entero positivo impar):

(a) $(a\,|\,n) = (b\,|\,n)$ si $a \equiv b \pmod n$.

(b) $(1\,|\,n) = 1$ y $(0\,|\,n) = 0$.

(c) Fórmula $(2\,|\,n) = (-1)^{(n^2-1)/8}$, de la cual se deduce lo siguiente: $(2a\,|\,n) = (a\,|\,n)$ si $n \equiv \pm 1 \pmod 8$; en otro caso, $(2a\,|\,n) = -(a\,|\,n)$.

(d) Ley de reciprocidad cuadrática: si m y n son ambos impares, entonces $(m\,|\,n) = (n\,|\,m)$ salvo que m y n sean ambos congruentes con 3 módulo 4, en cuyo caso $(m\,|\,n) = -(n\,|\,m)$.

Con ellas, podemos calcular el símbolo de Jacobi de manera muy eficiente, siguiendo un método similar al algoritmo de Euclides para hallar el máximo común divisor de dos números. En particular, con esto queremos decir que no vamos a tener que factorizar ninguno de los dos elementos que aparecen en $(\,\cdot\,|\,\cdot\,)$ (sólo sacar los factores «2», que eliminaremos aplicando la propiedad (c)). Las propiedades (a) y (c) permiten reducir sucesivamente un símbolo de Jacobi a otros en los que sus elementos son mucho menores, hasta que alguna vez llegamos a poder aplicar (b).

Continuando el paralelismo con el cálculo del máximo común divisor, podemos utilizar la factorización de un elemento si la sabemos, pero no hay ninguna necesidad de ello. Si conocemos la descomposición en factores primos de dos números a y b, calcular $\mathrm{mcd}(a,b)$ es inmediato, pero el algoritmo de Euclides proporciona un método alternativo —y eficiente— de calcular $\mathrm{mcd}(a,b)$ sin necesidad de efectuar factorizaciones. Con el símbolo de Jacobi vamos a poder hacer lo mismo. Esto es

importante si realmente estamos interesados en evaluar el símbolo de Jacobi, pues no se conoce ningún algoritmo rápido para factorizar enteros.[5]

Por el contrario, no hay ningún método similar que pueda usarse para calcular símbolos de Legendre $(a \mid p)$ con p primo impar. Si no queremos usar que, con p primo, el símbolo de Legendre es el de Jacobi, y luego usar las propiedades de este último, posiblemente vamos a tener que hallar la descomposición de a en factores primos. Esto hace que los cálculos usando exclusivamente el símbolo de Legendre sean en general mucho más lentos que si nos ayudamos del de Jacobi. Y, de hecho, ésta es la razón por la que Jacobi extendió el símbolo de Legendre introduciendo el que lleva su nombre.

No vamos a describir explícitamente con fórmulas el algoritmo para evaluar el símbolo de Jacobi del que estamos hablando. Para nuestros objetivos, no necesitamos tanto detalle (si el lector quiere, lo puede encontrar, por ejemplo, en [88, § 11.3]). Sólo queremos convencernos de que es posible hacerlo de manera razonablemente rápida y sin factorizar, y para eso nada mejor que servirnos de algún ejemplo.

Ejemplo. Sin efectuar factorizaciones (salvo dividir números pares por 2), calcular el símbolo de Jacobi $(1001\,|\,9907)$.

Solución. Utilizando las propiedades del símbolo de Jacobi (obsérvese que sólo sacamos factores 2) tenemos:

$$\begin{aligned}(1001\,|\,9907) &= (9907\,|\,1001) = (898\,|\,1001) = (2\,|\,1001)(449\,|\,1001)\\ &= (449\,|\,1001) = (1001\,|\,449) = (103\,|\,449) = (449\,|\,103)\\ &= (37\,|\,103) = (103\,|\,37) = (29\,|\,37) = (37\,|\,29)\\ &= (8\,|\,29) = (4\,|\,29)(2\,|\,29) = -1.\end{aligned}$$ ◊

Ejemplo. Dado que 9907 es primo, tiene sentido hablar del símbolo de Legendre $(1001\,|\,9907)$. Calcularlo sin ayuda del símbolo de Jacobi.

Solución. Como 1001 no es primo, la primera igualdad

$$(1001\,|\,9907) = (9907\,|\,1001)$$

—que antes obteníamos mediante la ley de reciprocidad cuadrática— ya no tiene sentido como símbolos de Legendre, pues $(9907\,|\,1001)$ no es un símbolo de Legendre.

[5] Dicho de manera más precisa, aquí nos referimos a que ni siquiera se dispone de un algoritmo que funcione en un tiempo que crece, a lo más, polinomialmente con el número de dígitos del entero que hay que factorizar (aunque tampoco está descartado que tal algoritmo pueda existir).

Podemos salir del atolladero descomponiendo en factores $1001 = 7 \cdot 11 \cdot 13$, usando

$$(1001\,|\,9907) = (7\,|\,9907)(11\,|\,9907)(13\,|\,9907),$$

y aplicando la ley de reciprocidad cuadrática a cada uno de estos tres factores. Aun a riesgo de resultar demasiado insistentes, recalquemos una vez más que encontrar la descomposición en factores de 1001 no era necesario si usábamos símbolos de Jacobi; es más, ni siquiera hacía falta preocuparse de si 9907 o 1001 eran o no números primos. Pero, eso sí, una vez conocido que $1001 = 7 \cdot 11 \cdot 13$, el proceso es ya sencillo, pues

$$\begin{aligned}(7\,|\,9907) &= -(9907\,|\,7) = -(2\,|\,7) = -1,\\ (11\,|\,9907) &= -(9907\,|\,11) = -(7\,|\,11) = (11\,|\,7) = (4\,|\,7) = 1,\\ (13\,|\,9907) &= (9907\,|\,13) = (1\,|\,13) = 1,\end{aligned}$$

con lo cual $(1001\,|\,9907) = -1$. ◊

3.2.5. Leyes de reciprocidad para potencias mayores que dos

En principio, de manera similar a como uno se plantea estudiar lo que ocurre con restos de potencias cuadráticas, también podríamos interesarnos por los restos de potencias mayores que 2. Posiblemente una de las razones por las que Gauss estaba tan interesado en buscar cada vez nuevas demostraciones de la ley de reciprocidad cuadrática era porque intentaba encontrar una que se pudiera extender a restos de potencias superiores. Se dio cuenta de que para intentar encontrar leyes de reciprocidad cúbicas o bicuadráticas tenía que salir de los enteros, y tener en cuenta las raíces terceras o cuartas de la unidad en $\mathbb{C}$, el cuerpo de los números complejos. De hecho, Gauss introdujo los ahora denominados enteros gaussianos

$$\mathbb{Z}[i] = \{x + yi : x, y \in \mathbb{Z}\}$$

(por supuesto, aquí i denota la unidad imaginaria) porque los necesitaba en su estudio sobre restos bicuadráticos.

Tras los intentos y resultados parciales de Gauss y Legendre para potencias 3 y 4, fue Eisenstein en 1844 el primero que proporcionó una prueba completa de las correspondientes leyes de reciprocidad cúbica y bicuadrática. Tienen cierta similitud con la cuadrática, pero no podemos dar expresiones concretas sin utilizar extensiones de $\mathbb{Z}$ adjuntando las raíces de la unidad —y eso requiere un mayor tratamiento algebraico del que aquí estamos usando—, así que no las enunciaremos.



Ferdinand Gotthold Max Eisenstein (1823–1852)



Emil Artin (1898–1962)

Para potencias superiores, Eisenstein se topó con la dificultad que suponía la ausencia de factorización única en ciertos anillos.[6] Fue por entonces cuando Kummer introdujo lo que ahora ha llegado a ser la teoría de ideales en un anillo. Esto permitió a Eisenstein retomar su avance (1850), y poco después el mismo Kummer (1857) consiguió nuevas mejoras.

No proseguiremos con este relato meramente histórico, pero, si lo hubiéramos hecho, habrían salido los nombres de otros matemáticos eminentes, pues la búsqueda de leyes de reciprocidad ha tenido una gran repercusión en el desarrollo de la teoría algebraica de números. De hecho, el noveno de los problemas que propuso Hilbert en el Segundo Congreso Internacional de Matemáticos, celebrado en París el año 1900,[7] trataba sobre cómo formular leyes de reciprocidad con un planteamiento bastante general; tal como lo propuso Hilbert, el problema fue resuelto de manera parcial por Emil Artin tres décadas más tarde. Queremos finalizar citando la excelente monografía [63], en la que se pueden encontrar todas esas leyes de reciprocidad a las que hemos aludido y muchas más.

3.3. Pseudoprimos de Euler

Ya hemos comentado que, al contrario de lo que ocurría con el símbolo de Legendre, con el símbolo de Jacobi puede ser $(b \mid n) = 1$ aunque b no sea un resto

[6] Problemas similares a los de la demostración de sucesivos casos particulares del último teorema de Fermat, según ya hemos comentado en el apartado 1.6.3, en el primer tema de este libro.

[7] Ya hemos aludido a la conferencia de Hilbert en este congreso, a su lista de 23 problemas abiertos, y a su enorme influencia en la investigación matemática posterior, en el primer tema de este libro (y nos volverá a salir varias veces más). Véase la página 60.

cuadrático módulo n. Hay otra propiedad que los símbolos de Jacobi y Legendre no comparten. Si tenemos un número n que no es primo y calculamos $b^{(n-1)/2}$ módulo n, el valor que obtengamos puede coincidir o no con el símbolo de Jacobi $(b\,|\,n)$. Para convencernos de que ambas situaciones son posibles —y que, por tanto, para el símbolo de Jacobi no existe un resultado similar al criterio de Euler del teorema 3.6—, nada mejor que verlo:

Ejemplo. Sea $n = 1105$. Comprobar si se verifica o no la igualdad $(b\,|\,n) \equiv b^{(n-1)/2} \pmod{n}$ con $b = 2$ y 3.

Solución. Comencemos con $b = 2$. Al ser $1105 \equiv 1 \pmod 8$, de (3.27) obtenemos $(2\,|\,1105) = 1$. Por otra parte, para hallar $2^{(1105-1)/2}$ módulo 1105 vamos a usar el cálculo de potencias mediante cuadrados sucesivos (también llamado algoritmo de cuadrados iterados), que ya utilizamos varias veces en el tema anterior (como en los ejemplos de las páginas 101 y 126). Descompongamos

$$(1105-1)/2 = 552 = 2^9 + 2^5 + 2^3 \tag{3.29}$$

(recordemos que esto se corresponde con su representación en base 2: $552 = 1000101000_{(2)}$). Entonces, $2^{552} = 2^{2^9} \cdot 2^{2^5} \cdot 2^{2^3}$. Como $2^{2^k} = \left(2^{2^{k-1}}\right)^2$, sucesivamente vamos obteniendo

$$\begin{aligned}
&2^{2^2} \equiv 4 \pmod{1105}, && 2^{2^{2^2}} \equiv 16 \pmod{1105},\\
&2^{2^{2^3}} \equiv 256 \pmod{1105}, && 2^{2^{2^4}} \equiv 256\cdot 256 \equiv 341 \pmod{1105},\\
&2^{2^{2^5}} \equiv 341\cdot 341 \equiv 256 \pmod{1105}, && 2^{2^{2^6}} \equiv 256\cdot 256 \equiv 341 \pmod{1105},\\
&2^{2^{2^7}} \equiv 341\cdot 341 \equiv 256 \pmod{1105}, && 2^{2^{2^8}} \equiv 256\cdot 256 \equiv 341 \pmod{1105},\\
&2^{2^{2^9}} \equiv 341\cdot 341 \equiv 256 \pmod{1105}.
\end{aligned}$$

Con eso,

$$2^{(1105-1)/2} = 2^{552} \equiv 256\cdot 256\cdot 256 \equiv 1 \pmod{1105}$$

luego $(2\,|\,1105) = 1 \equiv 2^{(1105-1)/2} \pmod{1105}$.

En cuanto a $b = 3$, de (3.29) se obtiene que $3^{552} = 3^{2^9} \cdot 3^{2^5} \cdot 3^{2^3}$. Y usando $3^{2^k} = \left(3^{2^{k-1}}\right)^2$ tenemos

$$\begin{aligned}
&3^{2^2} \equiv 9 \pmod{1105}, && 3^{2^{2^2}} \equiv 81 \pmod{1105},\\
&3^{2^{2^3}} \equiv 81\cdot 81 \equiv 1036 \pmod{1105}, && 3^{2^{2^4}} \equiv 1036\cdot 1036 \equiv 341 \pmod{1105},\\
&3^{2^{2^5}} \equiv 341\cdot 341 \equiv 256 \pmod{1105}, && 3^{2^{2^6}} \equiv 256\cdot 256 \equiv 341 \pmod{1105},
\end{aligned}$$

$$3^{2^7} \equiv 341 \cdot 341 \equiv 256 \pmod{1105}, \qquad 3^{2^8} \equiv 256 \cdot 256 \equiv 341 \pmod{1105},$$
$$3^{2^9} \equiv 341 \cdot 341 \equiv 256 \pmod{1105}.$$

Entonces,

$$3^{(1105-1)/2} = 3^{552} \equiv 256 \cdot 256 \cdot 1036 \equiv 781 \pmod{1105}.$$

Como $(3\,|\,1105)$ sólo puede valer 0, 1 o -1, no es necesario molestarse en calcularlo: esta vez es imposible que $(3\,|\,1105) \equiv 3^{(1105-1)/2} \pmod{1105}$. ◊

Cuando n es primo, el símbolo de Jacobi coincide con el de Legendre, así que siempre se va a cumplir $(b\,|\,n) \equiv b^{(n-1)/2} \pmod{n}$. Si n es compuesto, esa igualdad puede no ser cierta (y, de hecho, no lo es muy a menudo). Como vimos en el apartado 3.2.4, evaluar el símbolo de Jacobi $(b\,|\,n)$ es «rápido»;[8] evaluar $b^{(n-1)/2}$ módulo n también, sin más que usar el cálculo de potencias mediante cuadrados sucesivos. Así pues, comprobar si se cumple o no $(b\,|\,n) \equiv b^{(n-1)/2} \pmod{n}$ no requiere demasiado trabajo. Desde luego, mucho menos que —por ejemplo— intentar factorizar n para ver si es o no primo.

De este modo, cada entero positivo b nos proporciona un test de primalidad: si no sabemos si un número n es primo o compuesto, podemos elegir un número b al azar, calcular el símbolo de Jacobi $(b\,|\,n)$ y compararlo con el $b^{(n-1)/2}$ módulo n de la fórmula de Euler; si difieren, n es compuesto. Si coinciden para bastantes valores de b, es bastante probable (pero no seguro) que n sea un primo.

Un número compuesto n (impar) que satisface

$$(b\,|\,n) \equiv b^{(n-1)/2} \pmod{n}, \tag{3.30}$$

donde b es un entero positivo, se denomina *pseudoprimo de Euler* en base b. Así, del ejemplo anterior se desprende que 1105 no puede ser primo, pues $(3\,|\,n) \not\equiv 3^{(n-1)/2} \pmod{n}$. Pero $(2\,|\,n) \equiv 2^{(n-1)/2} \pmod{n}$, luego 1105 es un pseudoprimo de Euler en base 2.

Así mismo, y por los estudios teóricos de Robert M. Solovay y Volker Strassen en relación con los pseudoprimos de Euler (en particular, el teorema 3.17 que veremos un poco más adelante, y que ellos probaron en 1977), cuando n satisface (3.30) se dice que n pasa el test de Solovay-Strassen en base b.

[8] Como es habitual en este libro, no precisaremos a qué nos referimos con «rápido». En todo caso, siempre alude a que el algoritmo correspondiente es mejor —en cuanto al tiempo de cálculo que requiere— que el que se emplearía usando algún otro con el que lo estemos comparando (o lo vayamos a comparar).



Robert Martin Solovay (1938–)



Volker Strassen (1936–)

Aunque no lo vamos a demostrar, no hay ningún número n que sea pseudoprimo de Euler en base b para todos los b (lo probó D. H. Lehmer en 1976). De hecho, siempre existe algún b con $1 < b < n$ y $\mathrm{mcd}(b, n) = 1$ tal que $(b \mid n) \not\equiv b^{(n-1)/2} \pmod{n}$. Así que, teóricamente, este criterio siempre podría servirnos para demostrar la primalidad; pero no hay forma de saber qué b es el que nos va a resultar útil para un n concreto, e irlos probando todos hasta n es una tarea hercúlea si n es grande: requiere más trabajo que factorizar n.

En la práctica, probar con unos cuantos valores de b proporciona bastante seguridad de que un número n que ha pasado todos esos tests es primo (y, si no ha pasado alguno, la seguridad total de que es compuesto). Cualquier algoritmo diseñado para comprobar la primalidad —incluso aunque queramos la certeza matemática— lo primero que hace es chequear la pseudoprimalidad con varios b. Esto detectará casi todos los números compuestos; de hecho, y según se desprende de los resultados de Solovay y Strassen, la probabilidad de que un número compuesto n pase k tests (para k valores de b elegidos al azar entre 2 y $n-1$) es menor que $1/2^k$. Si, tras eso, queremos la certeza absoluta de que ese n es primo, no queda más remedio que acudir a algún método seguro; tales métodos existen, pero son más complicados y requieren mucho más tiempo que los tests no seguros —también denominados tests de primalidad probabilísticos— por medio de pseudoprimos.

3.3.1. El test de Solovay-Strassen frente al de Miller-Rabin

En el tema anterior —en concreto, en la sección 2.8— ya habíamos visto otro test de primalidad no seguro, esta vez basado en el teorema pequeño de Fermat: cuando n es primo, $b^{n-1} \equiv 1 \pmod{n}$ siempre que $\mathrm{mcd}(b, n) = 1$. De nuevo no podemos evitar que, a veces, un número n que satisface esa expresión para un cierto b sea

compuesto, y en ese caso decíamos que n es *pseudoprimo* en base b. Y los números n que eran pseudoprimos para todo b eran los denominados números de Carmichael.

Es sencillo comprobar que el test de primalidad basado en pseudoprimos de Euler es mejor (en el sentido de que menos números compuestos pasan el test) que el basado en pseudoprimos:

Proposición 3.13.

Sea n un número compuesto impar y b un entero que cumple $\mathrm{mcd}(n,b)=1$. *Si n es pseudoprimo de Euler en base b, también es pseudoprimo en base b.*

Demostración. Sea n pseudoprimo de Euler en base b, es decir, n y b satisfacen (3.30). Elevando al cuadrado esa congruencia, obtenemos

$$(b^{(n-1)/2})^2 \equiv (b\,|\,n)^2 \pmod{n}.$$

Pero el símbolo de Jacobi $(b\,|\,n)$ siempre vale 1 o -1 (como $\mathrm{mcd}(n,b)=1$, no puede ser $(b\,|\,n)=0$), y por tanto la expresión anterior queda $b^{n-1}\equiv 1 \pmod{n}$; en otras palabras, n es pseudoprimo en base b. □

Lo contrario no es cierto. Por ejemplo, y tal como se propone en el ejercicio 3.15, 341 es pseudoprimo en base 2, pero no es pseudoprimo de Euler en base 2.

En la sección 2.8 —en concreto, en el apartado 2.8.2—, a la condición $b^{n-1}\equiv 1 \pmod{n}$ mediante la que definíamos los números pseudoprimos le añadíamos condiciones adicionales por medio de lo que se venía a denominar *test de Miller-Rabin* en base b. Estas condiciones hacían más difícil que un número compuesto no fuese detectado como tal, pero no eran infalibles; por ejemplo, el número $2047 = 23\cdot 89$ lograba engañar al test en base 2. Para cada base b, los números compuestos que pasan el test de Miller-Rabin se denominaban *pseudoprimos fuertes*.

En la proposición 3.13 acabamos de ver que los números pseudoprimos de Euler son también pseudoprimos, luego el test basado en pseudoprimos de Euler detecta números compuestos mejor que el basado en pseudoprimos. Aunque ya no es tan sencillo de comprobar —y, de hecho, no lo vamos a hacer (si el lector lo desea puede ver la demostración en [88, § 11.4], por ejemplo)—, con los pseudoprimos fuertes ocurre al revés:

Proposición 3.14.

Sea n un número compuesto impar y b un entero que cumple $\mathrm{mcd}(n,b)=1$. *Si n es pseudoprimo fuerte en base b, también es pseudoprimo de Euler en base b.*

En general, lo contrario no es cierto, como muestra el siguiente ejemplo:

Ejemplo. Comprobar que 1105 es pseudoprimo de Euler en base 2, pero no es pseudoprimo fuerte en base 2.

SOLUCIÓN. Que 1105 es pseudoprimo de Euler en base 2 ya lo hemos visto en un ejemplo en la página 195.

Para comprobar que 1105 no es pseudoprimo fuerte, vamos a ver que no pasa el test de Miller-Rabin en base 2. Descompongamos $1105 - 1 = 2^4 \cdot 69$, y escribamos $69 = 2^6 + 2^2 + 1$. Entre los cálculos que efectuábamos en el ejemplo antes citado, obteníamos

$$2^{2^2} \equiv 16 \pmod{1105} \qquad \text{y} \qquad 2^{2^6} \equiv 341 \pmod{1105}.$$

Entonces,

$$2^{69} = 2^{2^6} \cdot 2^{2^2} \cdot 2^1 \equiv 341 \cdot 16 \cdot 2 \equiv 967 \pmod{1105},$$

y, a partir de aquí,

$$\begin{aligned} 2^{138} &= \left(2^{69}\right)^2 \equiv 967^2 \equiv 259 \pmod{1105}, \\ 2^{276} &= \left(2^{138}\right)^2 \equiv 259^2 \equiv 781 \pmod{1105}, \\ 2^{552} &= \left(2^{276}\right)^2 \equiv 781^2 \equiv 1 \pmod{1105}. \end{aligned}$$

Como

$$2^{69} \not\equiv \pm 1,\ 2^{138} \not\equiv -1,\ 2^{276} \not\equiv -1,\ 2^{552} \not\equiv -1 \pmod{1105},$$

el número 1105 no ha pasado el test de Miller-Rabin en base 2. ◊

Pero los conceptos de número pseudoprimo de Euler y número pseudoprimo fuerte son equivalentes para algunos tipos de números, como muestran las dos proposiciones siguientes (cuya demostración proponemos al lector en los ejercicios 3.16 y 3.17):

Proposición 3.15.

Sea $n \equiv 3 \pmod 4$ un número compuesto, y b un entero con $\mcd(n, b) = 1$. Si n es pseudoprimo de Euler en base b, también es pseudoprimo fuerte en base b.

Proposición 3.16.

Sea n un número compuesto impar, y b un entero con $\mcd(n, b) = 1$. Si n es pseudoprimo de Euler en base b y $(b \,|\, n) = -1$, también es pseudoprimo fuerte en base b.

Concluyamos esta sección mostrando el resultado de Solovay y Strassen al que antes aludíamos, y que es similar al que —en el teorema 2.25— exponíamos para el test de Miller-Rabin:

Teorema 3.17. Solovay y Strassen, 1977.
Si un entero impar n es un número compuesto, entonces n pasa el test de Solovay-Strassen para, a lo más, $\varphi(n)/2$ bases b menores que n y primas con él.

Así pues, aplicando el test de Solovay-Strassen el suficiente número de veces, conseguiríamos un test de primalidad seguro. Pero, de nuevo, tal número de veces requeriría un trabajo de cálculo superior al de factorizar, luego el test de Solovay-Strassen carece de utilidad como método seguro (de hecho, ni siquiera sabemos cuánto es $\varphi(n)/2$ pues, sin conocer la descomposición en factores de n, no se puede calcular $\varphi(n)$; podríamos estimarlo y dar una cota superior, pero las únicas cotas para $\varphi(n)$ son del tamaño de n, como vemos en el apartado 8.1.2 de este libro).

En su lugar, sí que se puede usar como test probabilístico. Del teorema 3.17 se puede inferir que, si n es un entero compuesto impar, y elegimos una base $b < n$ al azar (como debe ser $\text{mcd}(b, n) = 1$, hay $\varphi(n)$ bases posibles), la probabilidad de que n sea pseudoprimo de Euler en base b es menor que $1/2$. Y, en consecuencia, si elegimos al azar k enteros distintos, todos positivos menores que n, la probabilidad de que n pase el test de Solovay-Strassen en cada una de esas k bases es menor que $(1/2)^k$.

Dado un número n y una base b, el tiempo que se tarda en comprobar si n pasa el test de Solovay-Strassen es similar al que se tarda en comprobar si pasa el de Miller-Rabin. Entonces, como los números compuestos «engañan» mejor al de Solovay-Strassen, todo esto hace que el test probabilístico de Solovay-Strassen sea algo peor que el de Miller-Rabin y, por tanto, menos usado en la práctica.

3.4. Tests de primalidad para números de formas especiales

La definición de primalidad es muy clara, y teóricamente es posible comprobar si un número m es o no primo simplemente mirando si m es o no divisible por algún entero k en el intervalo $2 \leq k \leq \sqrt{m}$.

Pero, si m es un entero «muy grande»,[9] probar todos esos k uno a uno puede resultar impracticable, pues requeriría un tiempo enorme; para ser más precisos, digamos que el tiempo que emplearíamos crece exponencialmente con el número

[9] En este libro, a menudo hablamos de enteros «grandes», pero nunca aclararemos qué queremos decir con ello, pues dependerá de las herramientas que tengamos a nuestra disposición: no es lo mismo que sólo podamos hacer divisiones con lápiz y papel o que dispongamos de un ordenador potentísimo; con mejores medios podemos alcanzar mayores objetivos, pero siempre habrá números de tamaño suficiente como para que el trabajo sea inviable en la práctica.

de dígitos de m. En las secciones 2.8 del tema anterior y 3.3 de éste veíamos métodos que podían garantizar que un número era primo con una probabilidad bastante alta, pero sin asegurarlo —eran los llamados tests de primalidad no seguros o probabilísticos, y a los números que pasaban tales tests los llamábamos pseudoprimos.

Pero no estamos ahora interesados en pseudoprimos ni en tests no seguros, sino en tests seguros que realmente garanticen la primalidad (una visión más amplia de lo que aquí apuntamos se encuentra en [20]). Factorizar es uno de tales métodos, pero —incluso con los mejores algoritmos de factorización conocidos— es muy lento. Existen algoritmos más rápidos para comprobar con rigor si un entero m genérico es o no primo. Por ejemplo, en el tema anterior (nota 2.30, en la página 145) hablábamos del test de primalidad AKS, que permitía comprobar si un número es primo en un tiempo que crece polinomialmente con el número de dígitos. En la práctica, se emplean también otros métodos (que ni siquiera enunciaremos), pero no son sencillos y, además, conseguir demostrar rigurosamente lo que afirman suele requerir conocimientos matemáticos elevados (el lector puede encontrar tales tipos de métodos en las monografías [30, 83]).

Para números de formas especiales —por ejemplo, $m = 2^{2^n} + 1$ o $m = 2^n - 1$— existen tests de primalidad específicamente diseñados que son mucho más rápidos que los genéricos. Y no sólo eso, sino que son más sencillos, tanto en su enunciado como a la hora de dar la demostración que garantiza el funcionamiento del método. En esta sección vamos a presentar algunos de estos tests de primalidad válidos para ciertos tipos de números. Como veremos, van a ser el teorema de Euler-Fermat, el símbolo de Jacobi y la ley de reciprocidad cuadrática las herramientas esenciales que nos permitirán demostrar tales tests.

Antes de comenzar, recalquemos que los algoritmos que mostraremos permiten dilucidar con total seguridad si un número es o no primo; pero que, si la conclusión es que no es primo, no proporcionan sus factores. Recordemos asimismo que, aunque demostrar la primalidad y factorizar son problemas relacionados, no son el mismo, sino que factorizar es en general mucho más complicado que comprobar la primalidad.

3.4.1. Test de Pépin

Se denominan números de Fermat a los de la forma $F_n = 2^{2^n} + 1$ con $n \geq 0$ y, cuando son primos, se conocen como primos de Fermat (en el ejercicio 3.18 se propone al lector que pruebe que, si el exponente m en $2^m + 1$ no tiene la forma $m = 2^n$, el número $2^m + 1$ no puede ser primo). Los primeros números de Fermat son

$$F_0 = 3, \quad F_1 = 5, \quad F_2 = 17, \quad F_3 = 257, \quad F_4 = 65\,537,$$

y los cinco son primos. En 1640, Fermat había conjeturado (aunque admitió no haberlo conseguido probar) que todos los $F_n = 2^{2^n} + 1$ eran primos. Casi un siglo después, en 1732, Euler encontró la descomposición

$$F_5 = 4\,294\,967\,297 = 641 \cdot 6\,700\,417,$$

lo cual mostraba que Fermat estaba equivocado. De hecho, no se conoce ningún F_n con $n > 4$ que sea primo.

La literatura dedicada a los números de Fermat es abundante (nos contentamos aquí con citar [61], un libro dedicado al 400 aniversario del nacimiento de Fermat), y también son bastantes las propiedades que podríamos enunciar o demostrar con escaso trabajo; en los ejercicios 3.19 y 3.20 aparecen unas pocas. Pero nuestro objetivo aquí es —casi en exclusiva— mostrar un test de primalidad para los números de Fermat; un test que es una consecuencia sencilla de la ley de reciprocidad cuadrática que veíamos en el apartado 3.2.2.

No obstante, y pese a que no queremos alejarnos de este camino, no nos resistimos a mencionar una de las principales propiedades de los primos de Fermat. Y es que, de manera quizás muy sorprendente, los primos de Fermat tienen un papel importante en geometría. Esta conexión con la geometría radica en que un polígono regular de m lados es construible con regla y compás si y sólo si

$$m = 2^k p_1 p_2 \cdots p_n \tag{3.31}$$

con $k \geq 0$ y siendo los p_j primos de Fermat distintos (incluido el caso $n = 0$, que significa que no hay ningún factor p_j y por tanto m es una potencia de 2). Que la condición (3.31) es suficiente fue probado por Gauss en 1796. En particular, Gauss demostró que el polígono regular de 17 lados se podía construir con regla y compás; desde la época griega, ése era el primer polígono regular cuya construcción con regla y compás se volvía factible. La necesidad de (3.31) no fue probada explícitamente por Gauss, y generalmente se atribuye a Pierre Wantzel, quien lo publicó en 1836.

Expresado con palabras menos geométricas y más algebraicas, el anterior resultado se expresa diciendo que las funciones trigonométricas $\operatorname{sen}(2\pi/m)$ y $\cos(2\pi/m)$ se pueden calcular con un número finito de sumas, multiplicaciones y raíces cuadradas si y sólo si m tiene la forma dada en (3.31).[10] El lector ha podido ya suponer que no abordaremos la demostración de estas propiedades, que se pueden encontrar

[10] Por ejemplo, el propio Gauss, en sus *Disquisitiones Arithmeticae*, proporcionó la fórmula

$$\cos\left(\frac{2\pi}{17}\right) = \frac{-1}{16} + \frac{1}{16}\sqrt{17} + \frac{1}{16}\sqrt{34 - 2\sqrt{17}} + \frac{1}{8}\sqrt{17 + 3\sqrt{17} - \sqrt{34 - 2\sqrt{17}} - 2\sqrt{34 + 2\sqrt{17}}}$$

correspondiente al heptadecágono regular.

en muchos textos sobre teoría de cuerpos, sino que dejamos definitivamente de lado la caracterización de los polígonos construibles con regla y compás y volvemos a nuestra senda.

Como decíamos, nuestro interés ahora es dar un test «rápido» para analizar la primalidad de los números de Fermat. Seguro que no se le escapa al lector que los F_n son enormes (expresado en base 2, F_n tiene exactamente $2^n + 1$ dígitos), así que los algoritmos genéricos para probar la primalidad de números arbitrarios en seguida dejan de ser aplicables. En cambio, el test que aquí mostramos es capaz de chequear números mucho mayores —insistimos en que no precisaremos cuánto, pues depende de las herramientas que tengamos a nuestra disposición, y de todos es sabido que su potencia aumenta día a día—. Pero quizás resulta desilusionante oír que, pese a la velocidad y capacidad de cálculo de los ordenadores actuales, de momento no se ha encontrado ningún F_n primo desde que el propio Fermat probara que sí lo eran los cinco primeros. ¿Serán los únicos? Nadie lo sabe.

Sin más preámbulos, enunciemos ya el test de primalidad al que nos referimos, obra del matemático (y sacerdote jesuita) francés Théophile Pépin:

Teorema 3.18. Test de Pépin, 1877.
Sea $F_n = 2^{2^n} + 1$ con $n \geq 1$. El número F_n es primo si y sólo si

$$3^{(F_n-1)/2} \equiv -1 \pmod{F_n}. \tag{3.32}$$

Antes de dar la demostración, vamos a ver el concepto de orden módulo m (aunque la definición depende de m, no lo diremos explícitamente cada vez que hablemos del orden; eso sí, la notación ord_m que utilizaremos evitará toda posible ambigüedad). Sea $m \geq 2$ un entero fijo. Para $a \in \mathbb{Z}$ que verifica $\mathrm{mcd}(a, m) = 1$, el orden de a es, por definición,

$$\mathrm{ord}_m(a) = \min\{\, k \geq 1 : a^k \equiv 1 \pmod{m}\}.$$

Obsérvese que, por el teorema de Euler-Fermat, $a^{\varphi(m)} \equiv 1 \pmod{m}$, así que el orden de a está bien definido y verifica $1 \leq \mathrm{ord}_m(a) \leq \varphi(m)$.

Una propiedad importante del orden es la siguiente:

Proposición 3.19.
Sea $m \geq 2$ un entero fijo, $a \in \mathbb{Z}$ tal que $\mathrm{mcd}(a, m) = 1$, y t otro entero no negativo. Entonces, $a^t \equiv 1 \pmod{m}$ si y sólo si $\mathrm{ord}_m(a)$ divide a t.

Demostración. En primer lugar, es claro que, si $\mathrm{ord}_m(a) \mid t$, podremos poner $t = \mathrm{ord}_m(a)q$ con q un entero positivo, y entonces

$$a^t = \left(a^{\mathrm{ord}_m(a)}\right)^q \equiv 1^q \equiv 1 \pmod{m}.$$

Para probar el recíproco, supongamos que $\mathrm{ord}_m(a) \nmid t$. Entonces, $t = \mathrm{ord}_m(a)q + r$ para algún r que cumple $1 \leq r < \mathrm{ord}_m(a)$. Pero, de aquí,

$$1 \equiv a^t = \left(a^{\mathrm{ord}_m(a)}\right)^q a^r \equiv 1^q \cdot a^r = a^r \pmod{m},$$

y esto es absurdo, pues $\mathrm{ord}_m(a)$ es el menor entero positivo k que satisface $a^k \equiv 1 \pmod{m}$. □

Dado que $a^{\varphi(m)} \equiv 1 \pmod{m}$ cuando $\mathrm{mcd}(a, m) = 1$ (teorema de Euler-Fermat), consecuencia de la proposición anterior es que $\mathrm{ord}_m(a)$ siempre divide a $\varphi(m)$.[11]

Para nuestros objetivos, vamos a utilizar también otra propiedad más específica:

Lema 3.20.

Sea $m \geq 2$ un entero fijo y $a \in \mathbb{Z}$ tal que $\mathrm{mcd}(a, m) = 1$. Supongamos además que, para ciertos enteros positivos d y k, y cierto entero no negativo r, se cumple $\mathrm{ord}_m(a^{d^r}) = d^k$. Entonces, $\mathrm{ord}_m(a) = d^{k+r}$.

En particular, si $a^{2^r} \equiv -1 \pmod{m}$ y $m \geq 3$, entonces $\mathrm{ord}_m(a) = 2^{r+1}$.

Demostración. Antes de empezar, observemos que, por ser $\mathrm{mcd}(a, m) = 1$, también $\mathrm{mcd}(a^{d^r}, m) = 1$, y tiene sentido hablar del orden de a^{d^r}. Una vez entrados en materia, fijémonos en que

$$a^{d^{k+r}} = (a^{d^r})^{d^k} \equiv 1 \pmod{m}, \tag{3·33}$$

luego $\mathrm{ord}_m(a) \leq d^{k+r}$.

Supongamos que no es cierta la igualdad, es decir, que $\mathrm{ord}_m(a) = s < d^{k+r}$. Por cumplirse (3·33), de la proposición 3.19 se sigue que el exponente d^{k+r} debe ser múltiplo de s. Así, las descomposiciones en factores primos de d y s deberán ser de la forma

$$\begin{aligned} d &= p_1^{b_1} p_2^{b_2} \cdots p_n^{b_n}, \quad b_j \geq 1, \\ s &= p_1^{c_1} p_2^{c_2} \cdots p_n^{c_n}, \quad 0 \leq c_j \leq (k+r)b_j, \end{aligned}$$

con algún $c_j < (k+r)b_j$ para que $s < d^{k+r}$. Si tomamos

$$s' = p_1^{\max\{c_1, rb_1\}} p_2^{\max\{c_2, rb_2\}} \cdots p_n^{\max\{c_n, rb_n\}} = \mathrm{mcm}(s, d^r),$$

[11] En realidad, el orden de un entero módulo m ya lo habíamos definido en el tema anterior, y esta propiedad era el ejercicio 2.8 (quizás el lector quiera recordar también la nota 2.5, en la página 99, en la que se hablaba del concepto de orden en el contexto de la teoría de grupos en general).

este s' cumplirá $s' < d^{k+r}$, $s \mid s'$, $d^r \mid s'$ y $s' \mid d^{r+k}$. Así pues, podemos escribir $s' = d^r u$ con $1 \leq u < d^k$. De $a^s \equiv 1 \pmod m$ se deduce que también $a^{d^r u} = a^{s'} = (a^s)^{s'/s} \equiv 1 \pmod m$. En estas circunstancias, tenemos

$$\begin{aligned} 1 &\equiv (a^{d^r})^{d^k} = a^{d^{r+k} - d^r u + d^r u} = a^{(d^k - u)d^r} a^{d^r u} \equiv (a^{d^r})^{d^k - u} \cdot 1 \\ &\equiv (a^{d^r})^{d^k - u} \pmod m, \end{aligned}$$

y esto es absurdo pues $1 \leq d^k - u < d^k = \operatorname{ord}_m(a^{d^r})$.

El caso particular es sólo tomar $d = 2$ y observar que $(a^{2^r})^2 \equiv (-1)^2 \equiv 1 \pmod m$, es decir, que el orden de a^{2^r} es $2 = 2^1$ (nótese que la condición $m \geq 3$ es importante, pues con $m = 2$ tendríamos $a^{2^r} \equiv -1 \equiv 1 \pmod 2$). □

Con esto, ya podemos probar el test de Pépin:

Demostración del teorema 3.18. Comencemos suponiendo que F_n es primo. La ley de reciprocidad cuadrática nos muestra que

$$(3 \mid F_n)(F_n \mid 3) = (-1)^{(3-1)(F_n - 1)/4} = 1;$$

por tanto, $(3 \mid F_n) = (F_n \mid 3)$. Pero, al ser $2 \equiv -1 \pmod 3$, se cumplirá

$$F_n \equiv 2^{2^n} + 1 \equiv (-1)^{2^n} + 1 \equiv 1 + 1 \equiv -1 \pmod 3,$$

y por consiguiente $(F_n \mid 3) = (-1 \mid 3) = -1$. A su vez, el criterio de Euler (teorema 3.6) establece que $(3 \mid F_n) \equiv 3^{(F_n - 1)/2} \pmod{F_n}$. De aquí se sigue (3.32).

Recíprocamente, supongamos que se cumple (3.32). Si F_n no fuese primo, tendría un divisor primo $p \leq \sqrt{F_n}$. Pero (3.32) implica que también

$$3^{(F_n - 1)/2} \equiv -1 \pmod p.$$

Entonces, por el lema 3.20 (basta usar el caso particular con $m = p$, $a = 3$ y $r = 2^n - 1$), $\operatorname{ord}_p(3) = F_n - 1$. Y, como además se debe cumplir $\operatorname{ord}_p(3) \mid \varphi(p) = p - 1$, de aquí se sigue que $(F_n - 1) \mid (p - 1)$, lo cual es absurdo pues $p \leq \sqrt{F_n} < F_n$. □

Veamos con un ejemplo que, utilizando el cálculo de potencias mediante cuadrados sucesivos, aplicar el test de Pépin no resulta demasiado trabajoso.

Ejemplo. Comprobar si F_6 es o no un número primo.

Solución. Tenemos que calcular $3^{(F_6 - 1)/2} = 3^{2^{2^6}/2} = 3^{2^{2^6 - 1}} = 3^{2^{63}}$ módulo $F_6 = 18\,446\,744\,073\,709\,551\,617$. Usando $3^{2^k} = \left(3^{2^{k-1}}\right)^2$ vamos obteniendo (para que la

lista que sigue no ocupe demasiado, omitimos parte de los cálculos intermedios)

$$\begin{aligned}
3^2 &\equiv 9 \pmod{F_6},\\
3^{2^2} &\equiv 9^2 \equiv 81 \pmod{F_6},\\
3^{2^3} &\equiv 81^2 \equiv 6561 \pmod{F_6},\\
3^{2^4} &\equiv 6561^2 \equiv 43\,046\,721 \pmod{F_6},\\
3^{2^5} &\equiv 43\,046\,721^2 \equiv 1\,853\,020\,188\,851\,841 \pmod{F_6},\\
3^{2^6} &\equiv 1\,853\,020\,188\,851\,841^2 \equiv 8\,733\,085\,925\,571\,693\,938 \pmod{F_6},\\
3^{2^7} &\equiv 8\,733\,085\,925\,571\,693\,938^2 \equiv 16\,487\,702\,282\,945\,150\,519 \pmod{F_6},\\
\cdots &\equiv \cdots\\
3^{2^{62}} &\equiv 12\,642\,481\,979\,558\,500\,686^2 \equiv 4\,082\,312\,613\,475\,963\,515 \pmod{F_6},\\
3^{2^{63}} &\equiv 4\,082\,312\,613\,475\,963\,515^2 \equiv 11\,860\,219\,800\,640\,380\,469 \pmod{F_6}.
\end{aligned}$$

Así pues,

$$3^{(F_6-1)/2} = 3^{2^{63}} \equiv 11\,860\,219\,800\,640\,380\,469 \not\equiv -1 \pmod{F_6}$$

y el test de Pépin asegura que F_6 es un número compuesto. ◊

El test de Pépin sólo sirve para comprobar si los números de Fermat son o no primos pero, cuando logramos establecer con él que un determinado F_n no es primo, no nos dice nada acerca de sus factores. Encontrar factores de los números de Fermat también ha sido —y es— un tema de interés para numerosos matemáticos. Es más, no podemos olvidar que Euler probó la no primalidad de F_5 encontrando un divisor, siglo y medio antes de que Pépin estableciera el test que lleva su nombre.

De hecho, Euler demostró en 1732 que todos los divisores de los números de Fermat deben tener una forma muy simple; en concreto, que los divisores (mayores que 1) de F_n deben ser de la forma $2^{n+1}k+1$ con $k \in \mathbb{N}$. Esto le permitió hallar, por simple tanteo, el factor $641 = 10 \cdot 2^6 + 1$ de F_5. Además, Édouard Lucas mejoró el resultado de Euler mostrando en 1878 que cada divisor de F_n debe tener la forma

$$2^{n+2}k + 1. \tag{3·34}$$

con k un entero positivo. Comprobémoslo (obsérvese el enunciado equivalente en términos de congruencias):



Thomas Clausen
(1801–1885)



Ivan Mikheevich
Pervushin (1827–1900)

Proposición 3.21.
Para $n \geq 2$, todo divisor q de F_n satisface $q \equiv 1 \pmod{2^{n+2}}$.

Demostración. Sin pérdida de generalidad podemos suponer que q es primo, pues cada número es producto de primos y, al multiplicar números congruentes con 1 módulo 2^{n+2}, se vuelven a obtener números con esa propiedad (es decir, si $q = q_1 \cdots q_k$ con los q_j números primos, y hemos demostrado la proposición para los primos, se tendrá $q_j \equiv 1 \pmod{2^{n+2}}$ para todo j, y su producto también cumplirá $q = q_1 \cdots q_k \equiv 1 \pmod{2^{n+2}}$, con lo cual estaría probado el caso general).

Sea pues q un divisor primo de F_n, lo que se puede expresar como $2^{2^n} \equiv -1 \pmod{q}$. Por el caso particular del lema 3.20, de aquí se deduce que 2 tiene orden 2^{n+1} módulo q y, por tanto, que 2^{n+1} divide a $\varphi(q) = q-1$. Al ser $n \geq 2$, se sigue que $q \equiv 1 \pmod 8$. Ahora, como el símbolo de Legendre cumple $(2\,|\,q) = (-1)^{(q^2-1)/8}$ (recordemos que esa propiedad era consecuencia de la ley de reciprocidad cuadrática), de aquí se concluye que $(2\,|\,q) = 1$, y por lo tanto (por definición del símbolo de Legendre) que existe un entero a tal que $a^2 \equiv 2 \pmod{q}$. Como 2 tiene orden 2^{n+1} módulo q, entonces a tiene orden 2^{n+2} módulo q, de donde 2^{n+2} divide a $\varphi(q) = q - 1$, y esto equivale a decir que $q \equiv 1 \pmod{2^{n+2}}$. □

El factor $641 = 5 \cdot 2^7 + 1$ de F_5 es pequeño y, por tanto, fácil de descubrir usando la caracterización (3·34). Pero merece la pena reflexionar un poco sobre lo que cuesta encontrar factores de F_n con este método. Por ejemplo, $F_6 = 2^{2^6} + 1 = 18\,446\,744\,073\,709\,551\,617$ es un número de 20 cifras y, si no es primo, sólo podemos asegurar que tendrá un factor primo menor que $\sqrt{F_6} \approx 4.294\,967 \cdot 10^9$. Hay 16 millones de números k que satisfacen $2^8k + 1 \leq \sqrt{F_6}$, y todos esos números $2^8k + 1$ son potenciales divisores de F_6. Chequearlos uno tras otro con lápiz y papel es un

trabajo titánico (en particular, mucho mayor que efectuar los 63 cuadrados sucesivos que requiere el test de Pépin, según hemos visto en el ejemplo anterior).

Aun así, la condición (3·34) —en realidad, su versión débil debida a Euler— permitió hallar factores de varios números de Fermat mayores que F_5. En 1854, Thomas Clausen encontró el factor $274\,177 = 1071 \cdot 2^8 + 1$ de F_6, probando de esa manera que F_6 tampoco era un número primo —el test de Pépin todavía no se conocía—. En su época, el hallazgo de Clausen pasó desapercibido: él lo comunicó a Gauss por carta, pero dicha carta no se hizo pública hasta 1964. Dicho factor de F_6 no sería redescubierto hasta 1880; lo encontró, tras varios meses de trabajo, el ya octogenario Fortuné Landry.

Durante el siglo XIX aún se encontraron algunos otros divisores de números de Fermat:

- En 1877 y principios de 1878 Ivan Mikheevich Pervushin presentó dos artículos en los que mostraba, respectivamente, que $F_{12} = 2^{2^{12}} + 1$ era divisible por $7 \cdot 2^{14} + 1 = 114\,689$ y $F_{23} = 2^{2^{23}} + 1$ lo era por $5 \cdot 2^{25} + 1 = 167\,772\,161$.
- Paul Seelhoff probó en 1886 que $2\,748\,779\,069\,441 = 5 \cdot 2^{39} + 1$ divide a F_{36}.
- Allan J. C. Cunningham[12] encontró en 1899 dos factores de F_{11}, a saber: $319\,489 = 39 \cdot 2^{13} + 1$ y $974\,849 = 119 \cdot 2^{13} + 1$.

Y muy pocos más han sido los factores de números de Fermat descubiertos —en la primera mitad del siglo XX— sin ayuda de ordenadores.

A partir de la segunda mitad de siglo XX —y ya siempre con ayuda informática— la cantidad de factores encontrados ha crecido considerablemente. Tras tres siglos de búsqueda, ya se conocen más de doscientos divisores primos de distintos F_n. Actualmente continúa habiendo matemáticos —tanto profesionales como aficionados— interesados en la búsqueda de divisores de los números de Fermat, e incluso en encontrar la descomposición completa en factores primos de cada F_n. Internet les brinda una manera excelente de colaborar, mantenerse al día y comunicar los nuevos hallazgos, que sin duda seguirán en el futuro.

3.4.2. Test de Proth

Se denominan números de Proth a los de la forma

$$k \cdot 2^n + 1, \quad k, n \in \mathbb{N},$$

[12] A él alude el denominado «proyecto Cunningham», que mencionaremos más adelante.

con k impar[13] y $k < 2^n$ (es decir, que la parte más significativa en cuanto al tamaño del número la proporciona el 2^n). Su nombre alude al matemático autodidacta francés François Proth, que descubrió un test para comprobar su primalidad (lo veremos en el teorema 3.22). A los números de Proth primos los llamamos primos de Proth.

Es sencillo comprobar que los primeros números de Proth son

$$P_0 = 2^1 + 1 = 3, \qquad P_1 = 2^2 + 1 = 5, \qquad P_2 = 2^3 + 1 = 9,$$
$$P_3 = 3 \cdot 2^2 + 1 = 13, \qquad P_4 = 2^4 + 1 = 17, \qquad P_5 = 3 \cdot 2^3 + 1 = 25,$$
$$P_6 = 2^5 + 1 = 33, \qquad P_7 = 5 \cdot 2^3 + 1 = 41, \qquad P_8 = 3 \cdot 2^4 + 1 = 49.$$

Y, ya sin detallar su forma, los primos de Proth menores que 10 000 son 3, 5, 13, 17, 41, 97, 113, 193, 241, 257, 353, 449, 577, 641, 673, 769, 929, 1153, 1217, 1409, 1601, 2113, 2689, 2753, 3137, 3329, 3457, 4481, 4993, 6529, 7297, 7681, 7937, 9473, 9601 y 9857.

Hay un par de casos particulares de números de Proth que tienen nombre propio. El primer caso especial es el de los números de Fermat, que son números de Proth con $k = 1$ (y n una potencia de 2); a ellos dedicamos el apartado anterior. El otro caso con nombre propio son los números de Cullen, que son números de Proth con $k = n$; su denominación alude al matemático y jesuita irlandés James Cullen, quien en 1905 notó que los números primos de la forma $n \cdot 2^n + 1$ —que se denominan primos de Cullen— eran bastante escasos. Los primeros primos de Cullen son los correspondientes a $n = 1$, 141, 4713, 5795 y 6611. Esa especial escasez de primos de Cullen entre los primeros valores n puede estar justificada por ciertas propiedades de divisibilidad como las que aparecen en los ejercicios 3.21 y 3.22 (y otras similares que podrían probarse): dado que para valores de n pequeños la densidad de primos es mayor, en esos casos es más fácil que tales propiedades proporcionen un divisor de $n \cdot 2^n + 1$. Pero no se sabe si hay infinitos números de Cullen primos. Otra pregunta sin respuesta conocida es si existe algún primo p tal que $p \cdot 2^p + 1$ sea también primo.

Merece asimismo la pena mencionar que, dado que los factores de los números de Fermat tienen la forma (3.34), estos factores son números de Proth (salvo por el detalle $k < 2^n$, como en el factor de F_6 descubierto por Clausen que vimos en el apartado anterior). De este modo, los tests destinados a comprobar la primalidad de los números de Proth tienen también interés si queremos analizar la primalidad de los factores de los números de Fermat; en particular, para intentar hallar la factorización completa como producto de primos de los números de Fermat.

[13] En realidad, la condición «k impar» es superflua, pues los factores 2 de k siempre se pueden incorporar al 2^n; así, esa exigencia sólo sirve para identificar a los números de Proth en función de n y k de manera única.

Para comprobar la primalidad de un número de Proth disponemos del denominado test de Proth, que afirma lo siguiente:

Teorema 3.22. Test de Proth, 1878.
Sea $m = k \cdot 2^n + 1$ con $0 < k < 2^n$. Si, para algún entero a, se cumple

$$a^{(m-1)/2} \equiv -1 \pmod{m} \tag{3.35}$$

entonces m es primo. Además, si a verifica $(a \mid m) = -1$ (estamos usando el símbolo de Jacobi[14]), entonces m es primo si y sólo si se cumple (3.35).

Demostración. Supongamos que se cumple (3.35) (en particular, eso implica que $\mcd(a, m) = 1$). Obsérvese que, al ser $k < 2^n$, forzosamente $\sqrt{m} < 2^n$. Así, si m no fuese primo, tendría un divisor primo $p \leq \sqrt{m}$, luego también $p < 2^n$. Como $p \mid m$, (3.35) implica que $a^{(m-1)/2} \equiv -1 \pmod{p}$, y esto lo podemos expresar como

$$(a^k)^{2^{n-1}} \equiv -1 \pmod{p}.$$

Entonces, por el lema 3.20, $\ord_p(a^k) = 2^n$. Además sabemos que $x^{\varphi(p)} \equiv 1 \pmod{p}$ siempre que $\mcd(x, p) = 1$ (teorema de Euler-Fermat), luego $\ord_p(a^k) \mid \varphi(p) = p - 1$ (proposición 3.19); de aquí se sigue que $2^n \mid (p - 1)$, lo cual es absurdo ya que teníamos $p < 2^n$.

Para finalizar, supongamos que m es primo y que a satisface $(a \mid m) = -1$. Como, por el criterio de Euler, $(a \mid m) \equiv a^{(m-1)/2} \pmod{m}$, es claro que se cumplirá (3.35). □

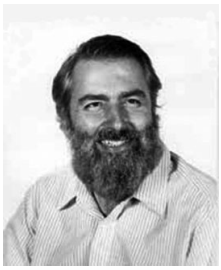
Es importante darse cuenta de que el test de Proth es útil en la práctica, ya que, si m es primo, existe un 50 % de posibilidades de que $(a \mid m) = -1$ y, por tanto, la primalidad de m sea detectada. La manera habitual de encontrar un a útil es ir tomando como a sucesivos números primos pequeños hasta que encontremos alguno tal que el símbolo de Jacobi cumpla $(a \mid m) = -1$.

Queremos también recalcar que —de manera similar a como los números de Proth son una extensión de los números de Fermat—, el test de Proth es una extensión del test de Pépin sin más que tomar $a = 3$. Esto es así porque, tal como se deduce fácilmente de lo que hacíamos en la demostración del test de Pépin (aunque esta vez aplicando la ley de reciprocidad cuadrática para el símbolo de Jacobi), $a = 3$ siempre cumple $(3 \mid F_n) = -1$.

[14] Cuando m es primo, el símbolo de Jacobi $(a \mid m)$ coincide con el de Legendre. Pero si nos hubiéramos quedado en el símbolo de Legendre y no hubiéramos definido el de Jacobi, el enunciado de este teorema no tendría sentido.



Wacław Sierpiński
(1882–1969)



John Lewis Selfridge
(1927–2010)

Nota 3.23. Unos números en cierta manera relacionados con los de Proth son los de Sierpiński. Un número de Sierpiński es un entero positivo impar k tal que todos los números del conjunto

$$\{k \cdot 2^n + 1 : n \in \mathbb{N}\}$$

son compuestos (obsérvese que ahora no estamos incluyendo la condición $k < 2^n$). Su nombre alude al matemático polaco Wacław Sierpiński, quien en 1960 probó que existen infinitos k con tal propiedad, aunque sin nombrar ninguno concreto. No se conocen muchos números de Sierpiński; al escribir esto, los más pequeños que se han encontrado son

$$78\,557, \quad 271\,129, \quad 271\,577, \quad 322\,523, \quad 327\,739.$$

Fue John Selfridge en 1962 quien probó que 78 557 es un número de Sierpiński; logró demostrar que, cuando $k = 78\,557$, todos los números de la forma $k \cdot 2^n + 1$ tienen un divisor en el conjunto $\{3, 5, 7, 13, 19, 37, 73\}$.

Determinar el menor número de Sierpiński es un problema abierto, el denominado problema de Sierpiński. En 1967, Sierpiński y Selfridge conjeturaron que $k = 78\,557$ era el menor número de Sierpiński, lo cual resolvería el problema. Para probarlo, hay que ver que todos los k impares menores que 78 557 no son números de Sierpiński; es decir, que para cada uno de ellos existe un n tal que $k \cdot 2^n + 1$ es primo. En 2002 se inició una búsqueda intensiva (colaborando a través de internet) para descartar todos esos $k < 78\,557$; en ese momento quedaban diecisiete candidatos por eliminar, y en 2007 ya se habían reducido a seis. Parecía que el problema de Sierpiński podría ser resuelto en no mucho tiempo, pero desde entonces, y a fecha actual (2018), el único avance ha sido descartar uno de esos seis números —en 2016

se demostró que $10\,223 \cdot 2^{31\,172\,165} + 1$ era primo, lo cual eliminaba a $k = 10\,223$ como posible número de Sierpiński—, así que las esperanzas de una solución cercana se han ralentizado considerablemente. ◁

Hay bastantes tipos de números similares a los de Proth cuya primalidad (existen test específicos que ya no mostraremos), u otras propiedades relacionadas con la divisibilidad, ha suscitado el interés tanto de matemáticos profesionales como aficionados. En particular, a menudo han dado lugar a búsquedas colaborativas con ayuda de internet, muy del agrado de los «cazadores de primos». No queremos despedir este apartado sin nombrar algunos (en lo que sigue, b siempre denota un entero ≥ 2):

- Números de Woodall, que son los de la forma $n \cdot 2^n - 1$.
- Números de Cullen o de Woodall generalizados, con la forma $n \cdot b^n + 1$ o $n \cdot b^n - 1$, respectivamente (y $n + 2 > b$ en ambos casos).
- Números de Sierpiński de primera clase, que son los de la forma $n^n + 1$.
- Números de Mersenne $2^n - 1$, para los que se aplica el test de Lucas-Lehmer que veremos en el siguiente apartado.
- Números de la forma $k \cdot 2^n - 1$, cuya primalidad se puede analizar con métodos basados en las ideas de Lucas y Lehmer adaptadas por Hans Riesel.
- Números de Cunningham, que son los de la forma $b^n \pm 1$. El denominado proyecto Cunningham (iniciado en 1925 y aún en marcha) pretende encontrar factores de números $b^n \pm 1$ para $b = 2, 3, 5, 6, 7, 10, 11, 12$ y potencias n grandes.
- Números de la forma $k \cdot 3^n \pm 1$ con $k < 3^n$ (o $k \cdot b^n \pm 1$ con b primo y $k < b^n$). En particular, para analizar la primalidad de $k \cdot 3^n \pm 1$ se pueden usar tests basados en la ley de reciprocidad cúbica; aquí merece la pena mencionar la metodología desarrollada, a finales del siglo XX, por Pedro Berrizbeitia y sus colaboradores (se puede ver una descripción detallada en [20]).
- Números de Riesel, que son los impares k tales que todos los números del conjunto $\{k \cdot 2^n - 1 : n \in \mathbb{N}\}$ son compuestos. En 1956, Riesel probó que existen infinitos k con esa propiedad, y determinar el menor de todos ellos es el denominado problema de Riesel. (Obsérvese el paralelismo con los números de Sierpiński y el problema de Sierpiński de la nota 3.23.)

Seguro que, si lo desea, el lector puede encontrar en internet todo tipo de información actualizada sobre los avances computacionales en el estudio de cada uno de estos tipos de números.

3.4.3. Test de Lucas-Lehmer

Para concluir nuestro periplo por los tests de primalidad para números de formas especiales, nos falta aún estudiar los más conocidos, los denominados números de Mersenne, que son los de la forma $M_p = 2^p - 1$ con p primo. Y que, cuando M_p es primo, se denominan primos de Mersenne. En este libro nos aparecerán también en la sección 7.2 en conexión con los números perfectos.

Los números de Mersenne han desempeñado un papel crucial en la historia de los tests de primalidad. Surgen de manera natural al buscar primos de la forma $a^n - 1$ con a un entero mayor que 1. De la igualdad

$$a^n - 1 = (a-1)(a^{n-1} + a^{n-2} + \cdots + a + 1), \quad n \in \mathbb{N},$$

se deduce que $a^d - 1$ divide a $a^n - 1$ para cada divisor d de n (basta escribir $a^n - 1 = (a^d)^k - 1$ y aplicar la fórmula anterior con a^d en lugar de a, y k en lugar de n). En consecuencia, $a^n - 1$ sólo puede ser primo cuando $a = 2$ y n es un primo. Así pues, y dado que el objetivo es estudiar la primalidad, los números de Mersenne no se definen como $2^n - 1$ con n un entero positivo, sino directamente añadiendo el requisito de que n sea primo (y entonces lo denotamos p).

La primalidad de $2^p - 1$ para $p = 2, 3, 5, 7$ ya era conocida desde la época griega. Quizás se podía pensar que $2^p - 1$, con p primo, era siempre primo; sin embargo, $2^{11} - 1 = 2047 = 23 \cdot 89$ es un número compuesto. La primalidad de $2^{13} - 1$ se conoce, al menos, desde 1456. Ya en 1588, Pietro Cataldi había verificado que $2^{17} - 1$ y $2^{19} - 1$ eran primos, y aventuró que lo mismo ocurría para $p = 23$, 29, 31 y 37. Se equivocó con 23, 29 y 37, como podemos ver sin más que dar sus descomposiciones en factores primos:

$$\begin{aligned} 2^{23} - 1 &= 8\,388\,607 = 47 \cdot 178\,481, \\ 2^{29} - 1 &= 536\,870\,911 = 233 \cdot 1103 \cdot 2089, \\ 2^{37} - 1 &= 137\,438\,953\,471 = 223 \cdot 616\,318\,177 \end{aligned}$$

(la primera y la tercera las encontró Fermat en 1640, y la segunda Euler en 1738, casi un siglo más tarde que las otras dos).

Nota 3.24. Incluso sin hallar su descomposición en factores, para descartar la primalidad de $2^p - 1$ se puede utilizar el teorema pequeño de Fermat, que afirma que, si q es un número primo y a un entero, forzosamente debe cumplirse $a^q \equiv a$ (mód q) (de hecho, recordemos que en la sección 2.8 llamábamos pseudoprimo en base a a un número q que pasara ese test). Calcular esas potencias puede hacerse de forma relativamente rápida (factible «a mano» para M_p con $p = 23$, 29 y 37)



Marin Mersenne
(1588–1648)



Édouard Lucas
(1842–1891)



Derrick Henry Lehmer
(1905–1991)

mediante cuadrados sucesivos. Así,

$$2^{2^{23}-1} \equiv 2 \pmod{2^{23}-1},$$
$$2^{2^{29}-1} \equiv 2 \pmod{2^{29}-1},$$
$$2^{2^{37}-1} \equiv 2 \pmod{2^{37}-1},$$

lo cual no aporta nada para decidir la primalidad de $2^{23}-1$, $2^{29}-1$ y $2^{37}-1$ (en otras palabras, los tres números han resultado ser pseudoprimos en base 2), pero

$$3^{2^{23}-1} \equiv 877\,681 \pmod{2^{23}-1},$$
$$3^{2^{29}-1} \equiv 195\,496\,587 \pmod{2^{29}-1},$$
$$3^{2^{37}-1} \equiv 36\,787\,319\,437 \pmod{2^{37}-1},$$

y eso implica que los tres números $2^{23}-1$, $2^{29}-1$ y $2^{37}-1$ son compuestos (no son pseudoprimos en base 3). En cambio, para M_{31}, por muchos valores de a que probemos siempre se va obteniendo $a^{2^{31}-1} \equiv a \pmod{2^{31}-1}$, así que ese método no permite confirmar ni descartar la primalidad de M_{31}. ◁

Volviendo a nuestro recorrido histórico, es en 1644 cuando el monje francés Marin Mersenne entra en escena en relación a los números que ahora llevan su nombre. Mersenne afirmó que, para $p \leq 257$, el número $2^p - 1$ es primo si y sólo si $p = 2$, 3, 5, 7, 13, 17, 19, 31, 67, 127 o 257. Esa lista resultó incorrecta en ambos sentidos: algunos primos no están en ella, y algunos de los números que sí están no son primos. La lista correcta para $p \leq 257$ no se completó hasta mediados del siglo XX, ya con ayuda de ordenadores: los valores de p para los que M_p es primo son $p = 2$, 3, 5, 7, 13, 17, 19, 31, 61, 89, 107 y 127 (el siguiente primo surge con $p = 521$).

Un nuevo paso importante en este relato es cuando Euler, para estudiar la primalidad de M_p, logra establecer dos condiciones que deben cumplir sus divisores. Estas condiciones son las que aparecen en las dos proposiciones que ahora exponemos:

Proposición 3.25.
Todo divisor primo q de M_p satisface $q \equiv 1 \pmod{p}$.

Demostración. Si q es un divisor primo de M_p, entonces $2^p \equiv 1 \pmod{q}$. Por la proposición 3.19, eso implica que $\mathrm{ord}_q(2) \mid p$; pero p es primo y $\mathrm{ord}_q(2) \neq 1$ (pues $q \neq 1$), luego forzosamente debe ser $\mathrm{ord}_q(2) = p$. Así mismo, el orden módulo q siempre divide a $\varphi(q)$, por lo que p divide a $\varphi(q) = q - 1$. □

Proposición 3.26.
Si p es un primo impar, todo divisor primo q de M_p satisface $q \equiv \pm 1 \pmod{8}$.

Demostración. Sea q un divisor primo de M_p. El símbolo de Legendre cumple $(2 \mid q) = (-1)^{(q^2-1)/8}$, así que demostrar $q \equiv \pm 1 \pmod{8}$ es equivalente a probar $(2 \mid q) = 1$. Usemos ahora $2^p \equiv 1 \pmod{q}$, de donde $\mathrm{ord}_q(2) \mid p$; y por ser p primo y $\mathrm{ord}_q(2) \neq 1$ (pues $q \neq 1$), eso implica que $\mathrm{ord}_q(2) = p$ y por tanto p divide a $\varphi(q) = q - 1$. Además, p es impar, luego p divide a $(q-1)/2$ y en consecuencia $2^{(q-1)/2} \equiv 1 \pmod{q}$. Con esto, del criterio de Euler $(b \mid q) \equiv b^{(q-1)/2} \pmod{q}$ (teorema 3.6) se sigue que $(2 \mid q) = 1$, tal como queríamos demostrar. □

Nótese que, si admitimos que p es un primo impar, los resultados de las dos proposiciones anteriores se combinan fácilmente. En primer lugar, las condiciones $q \equiv 1 \pmod{p}$ y $q \equiv 1 \pmod{8}$ se expresan, de manera conjunta, como $q \equiv 1 \pmod{8p}$ (pues $q - 1$ es divisible por p y por 8 si y sólo si lo es por $8p$). En segundo lugar, las condiciones $q \equiv 1 \pmod{p}$ y $q \equiv -1 \pmod{8}$ se unifican en una del tipo $q \equiv c \pmod{8p}$ mediante el teorema chino del resto.

Con estas consideraciones (y usando que $8 \cdot 31 = 248$), Euler concluyó que los posibles divisores primos q de M_{31} debían satisfacer $q \equiv 1 \pmod{248}$ o $q \equiv 63 \pmod{248}$. Así, y teniendo en cuenta que si $M_{31} = 2^{31} - 1 = 2\,147\,483\,647$ no fuese primo debería tener un divisor primo $q \leq \sqrt{M_{31}} < 46\,341$, ya no le resultó demasiado trabajoso comprobar que $M_{31} = 2^{31} - 1 = 2\,147\,483\,647$ es, en efecto, primo.

En 1876, Édouard Lucas presentó un algoritmo que permite determinar, de una manera asombrosamente eficiente, la primalidad de un número de Mersenne; tal algoritmo se conoce ahora con el nombre de test de Lucas-Lehmer, y lo veremos un poco más adelante. Con él, Lucas pudo demostrar en 1876 que el número de Mersenne $M_{127} = 2^{127} - 1$ (en base 10 tiene 39 dígitos) es primo; tal número era demasiado grande para estar al alcance de los métodos de sus predecesores. Fue un



Frank Nelson Cole (1861–1926)



Raphael Mitchel Robinson (1911–1995)

logro impresionante, y de hecho M_{127} ha sido el mayor primo de Mersenne encontrado sin ayuda de ordenadores (fue el mayor pero no fue el último temporalmente hablando, pues la comprobación de la primalidad de M_{61}, M_{89} y M_{107}, también efectuada manualmente, fue posterior).

Resulta anecdótico el caso de M_{67}. Con el test de Lucas-Lehmer es relativamente sencillo probar que M_{67} es un número compuesto. Pero eso no nos proporciona sus divisores. En 1903, Frank Cole, en una reunión de la American Mathematical Society, salió a la pizarra y en silencio se dedicó a efectuar multiplicaciones durante más de una hora hasta comprobar la siguiente igualdad:

$$193\,707\,721 \cdot 761\,838\,257\,287 = 2^{67} - 1;$$

la audiencia le premió con una gran ovación. Según él mismo explicó, había dedicado a factorizar M_{67} «tres años de domingos».

Pasaron 75 años desde que Lucas estableciera la primalidad de M_{127} hasta que se logró descubrir un primo mayor. En 1951, A. Ferrier, acompañado sólo de una calculadora mecánica, y utilizando tests de primalidad adecuados, probó que el número $(2^{148}+1)/17$ (con 44 dígitos) es primo, el mayor que se ha encontrado sin ayuda electrónica. Pero ya la era de la computación electrónica estaba comenzando, y la capacidad de cálculo se potenció hasta límites insospechados. Así, pocas semanas después, J. C. P. Miller y D. J. Wheeler inauguraban la extensa lista de primos hallados con ordenador al comunicar que habían encontrado los primos $k \cdot M_{127} + 1$ con $k = 114$, 124, 388, 408, 498, 696, 738, 774, 780, 934 y 978 (el mayor de 42 dígitos), y un nuevo récord (de 79 dígitos),

$$180M_{127}^2 + 1.$$

Un año más tarde, en 1952, Raphael Robinson comprobó la primalidad de M_{521}, M_{607}, M_{1279}, M_{2203} y M_{2281} (que tienen, respectivamente, 157, 183, 386, 664 y 687 cifras).

Desde entonces, el tamaño del mayor primo conocido ha ido aumentando sin cesar, y casi siempre ha sido un número de Mersenne (entre 1989 y 1992 fue un número de la forma $k \cdot 2^n - 1$ con 65 087 dígitos). La razón de que generalmente sea un número de Mersenne el que ostenta el honor de ser «el mayor primo conocido» es que el test de Lucas-Lehmer que se usa para probar la primalidad de los números de Mersenne no sólo es muy eficiente, sino que se adapta muy bien a la aritmética binaria en la que trabajan los potentes ordenadores actuales. Se conocen ya medio centenar de primos de Mersenne, algunos de más de diez millones de dígitos, pero se ignora si hay una cantidad infinita de ellos o no. ¿Qué nuevas ideas se utilizarán para poder averiguarlo?

Volviendo al test para analizar la primalidad de los M_p, fue Lucas (tal como decíamos) el que ideó la versión original del test en 1876, y con él probó la primalidad de M_{127}. En la formulación dada por Lucas, el test estaba presentado de manera distinta a como se muestra ahora, que es un refinamiento posterior (de 1930) debido a Derrick Henry Lehmer:

Teorema 3.27. Test de Lucas-Lehmer.
Sea $M_p = 2^p - 1$ con p un primo impar, y definamos los números S_k de manera recurrente mediante

$$S_0 = 4, \qquad S_{k+1} = S_k^2 - 4, \quad k \geq 0.$$

Entonces, el número M_p es primo si y sólo si $S_{p-2} \equiv 0 \pmod{M_p}$.

Sin embargo, no lo vamos a demostrar.

La ley de reciprocidad cuadrática para el símbolo de Jacobi proporciona ingredientes útiles para demostrar el test de Lucas-Lehmer de manera similar a como se prueba el test de Pépin (por ejemplo, las propiedades del ejercicio 3.25), pero dar un demostración completa y rigurosa requiere más herramientas algebraicas de las que hemos estudiado en este libro, así que no la abordaremos. De hecho, muchos textos sólo dan demostraciones de casos particulares, o se contentan con una implicación (que $S_{p-2} \equiv 0 \pmod{M_p}$ es condición suficiente para que M_p sea primo). Se puede encontrar una demostración moderna y completa —que utiliza algo de teoría de cuerpos finitos y de congruencias en un anillo más general que el de los enteros— en [19, 20], o en [86].

Con esas herramientas, dar un prueba del test de Lucas-Lehmer no es muy distinto que darla del test de Pépin. Pero —quizás el lector ya lo ha notado si se ha fijado en las fechas que hemos ido citando— el desarrollo histórico fue al revés del orden de dificultad argumental creciente en el que aquí están presentados los tests de Pépin y de Lucas-Lehmer. Es a Lucas al que hay que atribuirle el mérito de descubrir el primer test de este tipo (para números de Mersenne) en 1876, y Pépin lo adaptó en 1877 para números de Fermat. Es más, la redacción primitiva del resultado de Pépin era muy parecida a la de Lucas: el test de Pépin para números de Fermat consistía en tomar $S_0 = 3$ y $S_{k+1} = S_k^2$, y la condición de primalidad para F_n era $S_{2^n-1} \equiv -1 \pmod{F_n}$ (véase el ejercicio 3.23).[15] Luego, en 1878, Proth hizo una nueva adaptación para los números que ahora llevan su nombre. Resulta curioso que el test original sea el más complicado de demostrar.

3.5. Raíces primitivas

Dado un entero $m \geq 1$ (o $m > 1$, pues el caso $m = 1$ sólo da lugar a trivialidades) y otro entero a que verifica $\mcd(a, m) = 1$, el teorema de Euler-Fermat nos asegura que $a^{\varphi(m)} \equiv 1 \pmod{m}$. Entonces, en la sucesión de potencias positivas de a,

$$a, a^2, a^3, \ldots,$$

existe una potencia suficiente avanzada que cumple $a^k \equiv 1 \pmod{m}$. Como sabemos —pero lo recordamos por si el lector ha omitido las partes del libro en las que esto se explicaba—, al menor de los exponentes con esa propiedad lo denominamos *orden* de a módulo m, y se denota $\ord_m(a)$. Así pues, tenemos

$$\ord_m(a) = \min\{\, k \geq 1 : a^k \equiv 1 \pmod{m} \,\} \leq \varphi(m).$$

De hecho, ya habíamos estudiado algunas de sus propiedades. Por ejemplo, en el ejercicio 2.8 (en el tema 2) veíamos que $\ord_m(a)$ siempre divide a $\varphi(m)$. Y, en el presente tema, probábamos algo más: que $a^t \equiv 1 \pmod{m}$ si y sólo si $\ord_m(a)$ divide a t (proposición 3.19). También —aunque nosotros no lo vamos a necesitar—, para el lector interesado recordamos que el orden de un elemento se puede estudiar dentro del contexto más general de la teoría de grupos, tal como veíamos en la nota 2.5 (página 99).

[15] Resulta curioso constatar que, pese a los paralelismos de ambos test, y a los éxitos cosechados por el de Lucas-Lehmer encontrando una cantidad considerable de primos de Mersenne, el de Pépin no ha servido para descubrir ningún primo de Fermat tras los clásicos F_0, F_1, F_2, F_3 y F_4. De hecho, se tienen pocas esperanzas de encontrar más.

El objetivo de esta sección es profundizar en el estudio de las propiedades del orden módulo m (sin salirnos del contexto de los enteros), y de ello extraeremos algunas aplicaciones interesantes. En particular estamos interesados en cómo calcular $\mathrm{ord}_m(a)$, y en cuándo se cumple

$$\mathrm{ord}_m(a) = \varphi(m). \tag{3.36}$$

Para ciertas aplicaciones, los números a que satisfacen (3.36) son muy útiles, así que les vamos a asignar un nombre: un número a que cumple $\mathrm{ord}_m(a) = \varphi(m)$ se denomina *raíz primitiva* módulo m (a menudo omitiremos la referencia a m si no da lugar a confusión).

En realidad, también las raíces primitivas nos han aparecido ya en este libro. Cuando en el tema 2 hablábamos de criptografía de clave pública (apartado 2.9.2), necesitábamos una raíz primitiva módulo un primo para generar una clave pública segura mediante el protocolo de Diffie-Hellman. Más adelante veremos algunas otras aplicaciones.

Comencemos con un teorema que da una importante caracterización de las raíces primitivas. Para que su enunciado resulte claro, recordemos del tema 2, secciones 2.1 y 2.3, que $\mathbb{Z}_m^*$ denota las clases de equivalencia de los elementos de $\mathbb{Z}_m$ que son primos con m, y que $\mathbb{Z}_m^*$ siempre tiene $\varphi(m)$ elementos. Cuando tomamos un representante de cada una de estas clases de equivalencia, decimos que tenemos un sistema reducido de restos módulo m.

Teorema 3.28.
Sea $m \geq 2$ un entero fijo, y $a \in \mathbb{Z}$ tal que $\mathrm{mcd}(a, m) = 1$. Entonces, a es raíz primitiva módulo m si y sólo si los números

$$a, a^2, \ldots, a^{\varphi(m)}$$

son un sistema reducido de restos módulo m.

Demostración. Comencemos suponiendo que a es raíz primitiva módulo m. Para que los $\varphi(m)$ números $a, a^2, \ldots, a^{\varphi(m)}$ formen un sistema reducido de restos módulo m nos basta ver que no hay repetidos módulo m, es decir, que $a^j \not\equiv a^k \pmod{m}$ cuando $1 \leq j < k \leq \varphi(m)$. Como $\mathrm{mcd}(a, m) = 1$, también $\mathrm{mcd}(a^j, m) = 1$, y por tanto a^j tiene inverso $b = (a^j)^{-1}$ módulo m. Entonces, si se cumpliera $a^j \equiv a^k \equiv a^{k-j}a^j \pmod{m}$, multiplicando esa equivalencia por b obtendríamos $1 \equiv a^{k-j} \pmod{m}$, y esto no es posible por ser a raíz primitiva.

Recíprocamente, supongamos que $a, a^2, \ldots, a^{\varphi(m)}$ forman un sistema reducido de restos módulo m. Eso quiere decir que son todos distintos módulo m y, como siempre $a^{\varphi(m)} \equiv 1 \pmod{m}$, no puede haber ningún exponente $k < \varphi(m)$ tal que $a^k \equiv 1 \pmod{m}$. □

Este teorema es la clave de la importancia de las raíces primitivas. Nos acaba de mostrar que, cuando m tiene una raíz primitiva, para encontrar un sistema reducido de restos bastaría calcular los $\varphi(m)$ términos de una simple progresión geométrica. Cuando las raíces primitivas existen, esto proporciona una herramienta sencilla y potente. Pero, ¿existen siempre las raíces primitivas?

Analicemos los primeros valores de m:

- El caso $m = 1$ es trivial.
- Para $m = 2$, el número 1 es una raíz primitiva.
- Con $m = 3$, $\varphi(3) = 2$ y $a = 2$ es una raíz primitiva.
- Con $m = 4$, $\varphi(4) = 2$ y $a = 3$ es una raíz primitiva.
- Con $m = 5$, $\varphi(5) = 4$ y $a = 2$ es una raíz primitiva (módulo 5, las potencias sucesivas de 2 son $2, 4, 3, 1$).
- Con $m = 6$, $\varphi(6) = 2$ y $a = 5$ es una raíz primitiva.
- Con $m = 7$, $\varphi(7) = 6$ y $a = 3$ es una raíz primitiva (módulo 7, las potencias sucesivas de 3 son $3, 2, 6, 4, 5, 1$).

Sin embargo, no siempre existen raíces primitivas:

Ejemplo. Comprobar que no hay raíces primitivas módulo 8.

Solución. Los valores $a = 2$, 4 o 6 no nos sirven como raíces primitivas, pues no cumplen $\mcd(a, 8) = 1$. Ahora, como $\varphi(8) = 4$, para ver que no hay ninguna raíz primitiva basta comprobar que $\ord_8(1) = 1$ y $\ord_8(3) = \ord_8(5) = \ord_8(7) = 2$, lo cual es inmediato. ◊

Así pues, ¿para qué valores de m existen raíces primitivas módulo m? Hay una respuesta precisa a esta pregunta, y es la siguiente:

Teorema 3.29.
Dado m un entero positivo, existe alguna raíz primitiva módulo m si y sólo si

$$m = 1,\ 2,\ 4,\ p^k \ o\ 2p^k,$$

donde p indica un número primo impar y $k \geq 1$.

También es posible saber el número de raíces primitivas, como se muestra a continuación:

Teorema 3.30.
Si m (un entero positivo) tiene una raíz primitiva a, entonces m tiene exactamente $\varphi(\varphi(m))$ raíces primitivas incongruentes, y están dadas por los números del conjunto

$$\{\, a^k : 1 \leq k \leq \varphi(m) \ y \ \operatorname{mcd}(k, \varphi(m)) = 1 \,\}.$$

Las demostraciones de estos dos teoremas no son complicadas, pero no las incluimos para no alargarnos demasiado. Entre otros muchos textos dedicados a la teoría de números, la del primero se puede consultar en [5, § 10.2–10.8] o [88, teorema 9.15], y la del segundo en [5, § 10.9] o [88, corolario 9.4.1 y teorema 9.5]. Pero es que, además, nuestro principal interés no es efectuar un estudio teórico completo sobre raíces primitivas, sino sacar partido de ellas en algunas aplicaciones. A este respecto, y pese a su indudable interés, podemos pensar en los dos resultados anteriores como meros «teoremas de existencia» (no nos dicen cómo encontrar las raíces primitivas cuya existencia garantizan), y haber omitido su demostración no va a mermar el rigor en el desarrollo lógico-deductivo del resto del tema.

Así pues, dado un módulo m, se sabe si tiene o no raíces primitivas; y, si las tiene, cuántas. Pero encontrarlas es mucho más complicado (no hay que dejarse «engañar» por el teorema 3.30, que en realidad no proporciona la primera raíz a partir de la cual se podrían obtener las demás). De hecho, no se conoce ningún método directo para encontrar tales raíces y, por regla general, hallar una puede requerir una gran cantidad de cálculos, especialmente cuando m es grande. En cambio, sí que existen algunos resultados que, para m de algunas formas concretas, nos proporcionan una raíz primitiva; un ejemplo de esto es el ejercicio 3.29.

Nota 3.31. Aquí nos hemos planteado si, dado un entero m fijo, existen raíces primitivas módulo m; y cuántas y cuáles son éstas. Tal como hemos visto, sabemos responder al «cuántas», pero no al «cuáles». De manera simétrica, podríamos fijar un entero a y preguntarnos por los enteros m en los que a es una raíz primitiva módulo m. Para simplificar las cosas, preguntémonos sólo por los primos p en los que a es una raíz primitiva módulo p; ¿cuántos y cuáles son estos p? Esta vez, no hay respuestas concluyentes ni para el «cuántos» ni para el «cuáles». De hecho, ni siquiera se sabe con seguridad si, dado a, hay o no infinitos primos p en los que a es raíz primitiva módulo p. Pero sí hay una conjetura.

En 1927, Emil Artin conjeturó que, si a es un entero positivo que no es un cuadrado perfecto, entonces a es raíz primitiva módulo p para infinitos primos p. La conjetura de Artin va aún más lejos y pronostica que

$$\lim_n \frac{\text{card}\{p \le n : a \text{ es raíz primitiva módulo } p\}}{\text{card}\{p \le n\}} = C(a) \tag{3·37}$$

para cierta constante $C(a)$ concreta (y cuyo valor está predicho por la conjetura). Por ejemplo, cuando a es cualquier número libre de cuadrados, $C(a)$ es

$$C(a) = \prod_{q \text{ primo}} \left(1 - \frac{1}{q(q-1)}\right) = 0.373\,955\,813\,6\ldots,$$

que se denomina constante de Artin. En otros casos, el valor de $C(a)$ es un múltiplo racional de la constante de Artin.

Pensando la probabilidad de un suceso en términos de «casos favorables dividido entre casos posibles», la expresión (3·37) tiene una clara interpretación probabilística: la probabilidad de que a sea raíz primitiva módulo p para un primo p elegido al azar es $C(a)$. Por ejemplo, y en cierto sentido, $a = 2$ (y lo mismo con cualquier otro a libre de cuadrados) es raíz primitiva módulo p para, aproximadamente, el 37 % de los primos p.

Desde que Artin formuló su conjetura se han conseguido demostrar algunos resultados parciales, pero aún queda mucho camino por recorrer. ◁

3.5.1. Índices modulares (logaritmos discretos)

Volviendo a lo que veíamos en el teorema 3.28, cuando un entero $m \ge 2$ tiene una raíz primitiva g, los números $g, g^2, \ldots, g^{\varphi(m)}$ forman un sistema reducido de restos módulo m. Según esto, para cada entero a que cumple $\text{mcd}(a, m) = 1$, existe un único entero k con $1 \le k \le \varphi(m)$ tal que

$$g^k \equiv a \pmod{m}.$$

En estas circunstancias, decimos que k es el *índice* de a en base g módulo m, y se denota[16]

$$k = \text{índ}_g(a).$$

[16] Alternativamente, podíamos haber dicho que los números $1, g, g^2, \ldots, g^{\varphi(m)-1}$ forman un sistema reducido de restos módulo m (es lo mismo salvo el orden, puesto que $g^{\varphi(m)} \equiv 1 \pmod{m}$), y haber definido el índice como el único entero k con $0 \le k \le \varphi(m) - 1$ tal que $g^k \equiv a \pmod{m}$ (con el habitual $g^0 = 1$). Tal como lo hemos hecho en el texto se cumple $\text{índ}_g(1) = \varphi(m)$; y como en esta nota tendríamos $\text{índ}_g(1) = 0$.

Antes de continuar, fijémonos en que, si en lugar de estar tratando con enteros módulo m, estuviéramos hablando de números reales positivos, el k que verifica $g^k = a$ sería el logaritmo en base g de a. Ésta es la razón por la cual el índice también se conoce con el nombre de *logaritmo discreto*.

De hecho, tal como muestra el siguiente teorema cuya demostración dejamos para el lector (ejercicio 3.28), las propiedades de los índices modulares son similares a las de los logaritmos:

Teorema 3.32.
Sea un entero $m \geq 2$ que tiene una raíz primitiva g, y usemos a y b para denotar enteros que verifican $\mcd(a,m) = \mcd(b,m) = 1$. *Se cumple:*

(a) $\text{índ}_g(ab) \equiv \text{índ}_g(a) + \text{índ}_g(b) \pmod{\varphi(m)}$.
(b) $\text{índ}_g(a^n) \equiv n\,\text{índ}_g(a) \pmod{\varphi(m)}$ *para cualquier* $n \geq 1$.
(c) $\text{índ}_g(g) = 1$ *e* $\text{índ}_g(1) = \varphi(m)$ *(luego* $\text{índ}_g(1) \equiv 0 \pmod{\varphi(m)}$*).*[17]
(d) $\text{índ}_g(-1) = \varphi(m)/2$ *si* $m > 2$.
(e) Si g' es otra raíz primitiva módulo m, entonces

$$\text{índ}_g(a) \equiv \text{índ}_{g'}(a)\,\text{índ}_g(g') \pmod{\varphi(m)}.$$

De acuerdo a las propiedades (a) y (b) del teorema, si g es una raíz primitiva módulo m y la descomposición en factores primos de un entero $a > 1$ es $a = p_1^{n_1} p_2^{n_2} \cdots p_k^{n_k}$, entonces (supuesto que $\mcd(a,m) = 1$)

$$\text{índ}_g(a) \equiv n_1\,\text{índ}_g(p_1) + n_2\,\text{índ}_g(p_2) + \cdots + n_k\,\text{índ}_g(p_k) \pmod{\varphi(m)}.$$

En consecuencia, para calcular los índices módulo m nos basta saber calcular $\text{índ}_g(p)$ cuando p es primo. Por otra parte, usar una raíz primitiva g u otra g' no tiene mucha importancia, ya que la propiedad (e) nos permite fácilmente cambiar de una a otra. Así, se suele utilizar la más pequeña, que denotaremos $g(m)$.

En el cuadro 3.1 mostramos, para los números m entre 3 y 50 que tienen raíz primitiva, cuál es su menor raíz primitiva $g(m)$, así como el valor del índice $\text{índ}_{g(m)}(p)$ para primos p entre 2 y $m-1$. Dado que $\text{índ}_{g(m)}(p)$ no existe cuando $\mcd(m,p) > 1$, esos casos los hemos indicado con un «$-$» en el lugar correspondiente.

En general, incluso teniendo en cuenta las simplificaciones que acabamos de comentar, el cálculo de índices es muy laborioso. Aun suponiendo que ya conocemos una raíz primitiva g (algo que no era sencillo de encontrar en general, según

[17] Si hubiéramos definido el índice como comentábamos en la anterior nota al pie, directamente tendríamos $\text{índ}_g(1) = 0$.

m	$g(m)$	2	3	5	7	11	13	17	19	23	29	31	37	41	43	47
3	2	1														
4	3	–	1													
5	2	1	3													
6	5	–	–	1												
7	3	2	1	5												
9	2	1	–	5	4											
10	7	–	3	–	1											
11	2	1	8	4	7											
13	2	1	4	9	11	7										
14	3	–	1	5	–	4	3									
17	3	14	1	5	11	7	4									
18	11	–	–	5	4	1	2	3								
19	2	1	13	16	6	12	5	10								
22	13	–	8	4	7	–	1	9	3							
23	5	2	16	1	19	9	14	7	15							
25	2	1	7	–	5	16	19	13	18	11						
26	15	–	4	9	11	7	–	2	5	10						
27	2	1	–	5	16	13	8	15	12	11						
29	2	1	5	22	12	25	18	21	9	20						
31	3	24	1	20	28	23	11	7	4	27	9					
34	3	–	1	5	11	7	4	–	14	15	13	9				
37	2	1	26	23	32	30	11	7	35	15	21	9				
38	21	–	13	16	6	12	5	10	–	2	17	15	9			
41	6	26	15	22	39	3	31	33	9	36	7	28	32			
43	3	27	1	25	35	30	32	38	19	16	41	34	7	6		
46	5	–	16	1	19	9	14	7	15	–	18	6	21	12	5	
47	5	18	20	1	32	7	11	16	45	5	35	3	42	15	13	
49	3	26	1	29	–	40	33	25	35	38	18	7	32	15	6	5
50	27	–	7	–	5	16	19	13	18	11	2	8	9	4	15	17

Cuadro 3.1: Tabla que muestra, para los números m entre 3 y 50 que tienen raíz primitiva, cuál es su menor raíz primitiva $g(m)$, así como el valor del índice $\text{índ}_{g(m)}(p)$ para primos $p \leq m-1$. (Los símbolos «$-$» indican que el correspondiente $\text{índ}_{g(m)}(p)$ módulo m no existe por ser $\text{mcd}(m,p) > 1$.)

hemos comentado antes), no hay ningún procedimiento relativamente rápido para calcular $\text{índ}_g(a)$ módulo m. Por supuesto, podemos ir probando valores de k hasta que encontramos $g^k \equiv a \pmod{m}$, pero este proceso es prácticamente interminable cuando m es grande. Y, en realidad, ésta es una de las razones por las que las congruencias han adquirido una gran importancia a partir de las décadas de los setenta y ochenta del siglo XX. Recordemos que disponemos de un método eficaz de calcular potencias módulo un entero: el algoritmo de cuadrados iterados que hemos aplicado a menudo tanto en el tema anterior como en éste. Así pues, calcular potencias módulo un entero es rápido, pero el proceso inverso de evaluar el logaritmos discreto —es decir, el cálculo de índices— requiere mucho trabajo. Esto da lugar a una función de una sola vía de las que hablábamos en el apartado 2.9.2 del tema anterior, y que resultaban muy útiles en criptografía de clave pública. De hecho —tal como veremos un poco más adelante—, la seguridad del criptosistema denominado Elgamal está basada en la dificultad de evaluar el logaritmo discreto.

3.5.2. Aplicación a la resolución de ecuaciones

Análogamente a como ocurre con los logaritmos, el uso de índices modulares permite transformar productos en sumas, y potencias en productos. Si nos resulta sencillo calcular los índices, esto puede ser útil para resolver ecuaciones modulares:

Ejemplo. Encontrar los valores de x que satisfacen cada una de las siguientes ecuaciones:

(a) $7x \equiv 5 \pmod{13}$;
(b) $x^4 \equiv 6 \pmod{13}$;
(c) $7x^5 \equiv 9 \pmod{13}$;
(d) $5^x \equiv 8 \pmod{13}$.

SOLUCIÓN. Vamos a utilizar el cuadro 3.1, que muestra que $g = 2$ es una raíz primitiva módulo 13, y además da los valores de $\text{índ}_2(p)$ para los primos $p \leq 12$. Por supuesto, también usaremos las propiedades del teorema 3.32. Con esto:

(a) Tomando índices en base 2 (nótese que $\varphi(13) = 12$), $7x \equiv 5 \pmod{13}$ resulta ser equivalente a $\text{índ}_2(7) + \text{índ}_2(x) \equiv \text{índ}_2(5) \pmod{12}$. Pero, según el cuadro 3.1, esto es $11 + \text{índ}_2(x) \equiv 9 \pmod{12}$, o sea $\text{índ}_2(x) \equiv -2 \equiv 10 \pmod{12}$, y de aquí que $x = 2^{10} = 1024 \equiv 10 \pmod{13}$. En consecuencia, los x que resuelven la ecuación propuesta son los que verifican $x \equiv 10 \pmod{13}$.
(b) La ecuación se puede escribir como $4\,\text{índ}_2(x) \equiv \text{índ}_2(6) \pmod{12}$. A su vez, como $\text{índ}_2(6) \equiv \text{índ}_2(2)+\text{índ}_2(3) \equiv 1+4 \equiv 5 \pmod{12}$, esto queda $4\,\text{índ}_2(x) \equiv 5 \pmod{12}$. Pero la ecuación $4y \equiv 5 \pmod{12}$ no tiene solución ya que $\text{mcd}(4, 12) \nmid 5$.

(c) Transformemos primero la ecuación en $\text{índ}_2(7)+5\,\text{índ}_2(x)\equiv\text{índ}_2(3^2)\pmod{12}$. Usando $\text{índ}_2(7) = 11$ e $\text{índ}_2(3^2) = 2\,\text{índ}_2(3) = 8$, la nueva ecuación queda $5\,\text{índ}_2(x) \equiv 9 \pmod{12}$. Como $5^{-1} \equiv 5 \pmod{12}$, multiplicando por 5 llegamos a $\text{índ}_2(x) \equiv 45 \equiv 9 \pmod{12}$, de donde $x = 2^9 = 512 \equiv 5 \pmod{13}$. Así pues, la solución buscada es $x \equiv 5 \pmod{13}$.

(d) Al tomar índices en base 2 y usar $\text{índ}_2(5^x) \equiv x\,\text{índ}_2(5) = 9x \pmod{12}$ e $\text{índ}_2(8) = \text{índ}_2(2^3) \equiv 3\,\text{índ}_2(2) = 3 \pmod{12}$, la ecuación queda $9x \equiv 3 \pmod{12}$. Como $\text{mcd}(9, 12) = 3 \mid 3$, esta ecuación lineal tiene solución pese a que no existe 9^{-1} módulo 12. De hecho, tiene 3 soluciones módulo 12, que son $x = 3$, $x = 7$ y $x = 11$. Y estos x que satisfacen $x \equiv 3$, $x \equiv 7$ o $x \equiv 11 \pmod{12}$ son también las soluciones de $5^x \equiv 8 \pmod{13}$. ◊

En el ejemplo anterior, las ecuaciones de tipo $ax \equiv b \pmod{m}$ como la que aparece en (a) son las denominadas ecuaciones en congruencias lineales, y ya sabíamos cuándo tienen solución y cómo resolverlas. De hecho, dedicamos a ello un parte significativa del tema anterior (en particular, la proposición 2.1, toda la sección 2.2, y el teorema 2.8).

Las ecuaciones que aparecen en (b) y (c) son casos particulares de las ecuaciones polinómicas que hemos abordado en la sección 3.1. Allí estudiábamos cuántas soluciones podían tener como máximo, siendo el teorema de Lagrange (teorema 3.1) el resultado más importante; pero no disponíamos de un método efectivo para resolverlas. Las ecuaciones de la forma

$$ax^n \equiv b \pmod{m} \tag{3.38}$$

se denominan binomiales, y el uso de raíces primitivas e índices las transforma en ecuaciones lineales. En efecto, si m tiene una raíz primitiva g, y si $\text{mcd}(a, m) = \text{mcd}(b, m) = 1$, la ecuación anterior es equivalente a $\text{índ}_g(a) + n\,\text{índ}_g(x) \equiv \text{índ}_g(b) \pmod{\varphi(m)}$; o lo que, denotando $y = \text{índ}_g(x)$, resulta ser lo mismo que la ecuación

$$ny \equiv \text{índ}_g(b) - \text{índ}_g(a) \pmod{\varphi(m)},$$

que es lineal en la nueva incógnita y. Como tal ecuación lineal que es, tiene solución si y sólo si $\text{índ}_g(b) - \text{índ}_g(a)$ es divisible por $d = \text{mcd}(n, \varphi(m))$, en cuyo caso tiene exactamente d soluciones $y_1, \ldots, y_d$ (módulo $\varphi(m)$). Entonces, las soluciones de la ecuación (3.38) serán $x_j = g^{y_j}$, $j = 1, \ldots, d$ (módulo m).

Aún más novedoso es el ejemplo (d), en el que la incógnita está en el exponente. En general, supongamos que tenemos una ecuación en congruencias exponencial, esto es, una ecuación de la forma

$$a^x \equiv b \pmod{m}. \tag{3.39}$$

Si m tiene una raíz primitiva g, y si $\mathrm{mcd}(a,m) = \mathrm{mcd}(b,m) = 1$, la ecuación es equivalente a la congruencia lineal

$$x\,\mathrm{índ}_g(a) \equiv \mathrm{índ}_g(b) \pmod{\varphi(m)}, \tag{3.40}$$

Esta ecuación tiene solución si y sólo si $\mathrm{índ}_g(b)$ es divisible por

$$d = \mathrm{mcd}(\mathrm{índ}_g(a), \varphi(m)),$$

en cuyo caso tiene exactamente d soluciones (módulo $\varphi(m)$), que serán asimismo las soluciones de (3.39).

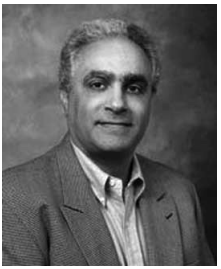
Resolver (3.39) es lo que se denomina calcular el logaritmo discreto. Pero llevar a la práctica el procedimiento mostrado en el ejemplo (d) no siempre es fácil. En particular, cuando m es grande, no es posible tener una tabla de índices módulo m como la que debería aparecer en la correspondiente línea del cuadro 3.1. Esto hace que el problema del logaritmo discreto sea irresoluble en un tiempo razonable cuando los números involucrados son grandes; para ser más precisos, no se conoce ningún algoritmo que permita resolver el problema del logaritmo discreto módulo m en un tiempo que crezca (a lo más) polinomialmente con el número de dígitos de m. Ni siquiera cuando m es primo y por tanto $\varphi(m) = m - 1$ de manera inmediata (lo cual hace que la relación módulo $\varphi(m)$ de (3.40) sea más asequible ¡si previamente hubiésemos podido calcular esos índices!), que de hecho es el caso más común.

3.6. Criptosistema Elgamal

Una de las secciones del tema anterior estaba dedicada a la criptografía. Allí veíamos un ejemplo de un antiguo sistema de cifrar mensajes —el denominado cifra del César—, mostrábamos que en la práctica los mensajes se podían reducir a números, y que tanto cifrar como descifrar eran simples operaciones modulares: sumar y restar la clave. Sus flaquezas debidas a su simplicidad hacían de la cifra del César un criptosistema muy poco seguro, y buscábamos estrategias para complicarlo y hacerlo más resistente al criptoanálisis. En definitiva, lo que se espera de un buen sistema criptográfico es que las personas que interceptan un mensaje cifrado que no iba destinado a ellas —y que por tanto no conocen la clave— no puedan descifrar tal mensaje.

En todos los métodos criptográficos clásicos, la persona que cifra y la que descifra deben ponerse de acuerdo en una clave; lo cifrado con una clave, también se descifra con ella. Si el emisor y el receptor de un mensaje cifrado comparten una clave que no conoce nadie más, existían métodos criptográficos muy seguros. Pero el

requisito previo de compartir una clave que debe ser secreta es una gran debilidad de estos métodos. Intercambiar una clave equivale a enviar un mensaje previo (aún no cifrado); ¿cómo evitar que la clave sea interceptada por un espía?



Taher Elgamal (1955–)

Es ahí donde surge la criptografía de clave pública, también llamada de claves asimétricas. La idea es que haya una clave para cifrar —la clave pública que puede conocer cualquiera— y otra clave privada que sólo se usa para descifrar. Si Alicia quiere enviar un mensaje secreto a Benito, Alicia cifra el mensaje con la clave pública de Benito, y lo envía a éste por cualquier canal de comunicación, por inseguro que sea. Sólo Benito conoce su clave privada, y por tanto sólo él logra descifrar el mensaje.

No es inmediato cómo llevar esa idea a la práctica, pero podemos poner un ejemplo. Imaginemos que tenemos dos primos «muy grandes»[18] p y q, y que hacemos su producto $m = pq$. A partir de p y q es inmediato calcular m; por el contrario, si sólo conocemos m es computacionalmente muy difícil encontrar sus dos factores primos p y q. Si dispusiéramos de un método criptográfico en el que bastara usar m para cifrar pero se necesitaran p y q para descifrar, Alicia y Benito ya tendrían lo que buscaban. Benito elegiría dos primos grandes p y q (que guardaría para sí mismo) y haría público su producto m. Alicia usaría m para cifrar mensajes y enviárselos a Benito, y sólo éste podría descifrarlos.

A grandes rasgos, eso era esencialmente lo que hacía el criptosistema RSA que veíamos con detalle en el apartado 2.9.3. La facilidad de multiplicar frente a la dificultad de descomponer en factores es un ejemplo de lo que se ha venido a denominar funciones de una sola vía o unidireccionales: se trata de procedimientos inversos pero que sólo en un sentido (el de multiplicar) se pueden llevar a la práctica cuando los números involucrados son suficientemente grandes.

La exponenciación modular y el cálculo de logaritmos discretos dan lugar a otra función de una sola vía, ya que encontrar el logaritmo discreto módulo un primo grande es incomparablemente más trabajoso que calcular potencias modulares, para lo cual —recordémoslo— disponíamos del cálculo de potencias mediante cuadrados sucesivos (o algoritmo de cuadrados iterados) que permitía hacerlo de manera relativamente rápida. En lo que sigue vamos a ver cómo esto se puede aprovechar para construir un nuevo método criptográfico de clave pública, el denominado criptosistema Elgamal.

[18] Por supuesto, «ser grande» es relativo, y depende del contexto y de la capacidad de cálculo que disponemos, pero en cada caso concreto hay unos números que son bastante grandes como para que sean útiles en esa situación.

Nota 3.33. La criptografía basada en la dificultad del cálculo de logaritmos discretos tiene una ventaja de la que carece la criptografía basada en la dificultad de factorizar enteros. Y es que la primera puede plantearse en cualquier grupo. De esta manera surgen multitud de nuevas posibilidades, porque muchos algoritmos se adaptan fácilmente. Para empezar a trabajar basta con que los elementos del grupo tengan una representación apta para el tratamiento informático y que la operación de grupo pueda implementarse de manera eficiente. Luego habrá que preguntarse por la dificultad del logaritmo discreto en ese grupo concreto. En particular, esto se utiliza en cierta aritmética que se puede definir sobre las denominadas curvas elípticas, que aquí no trataremos. Pero sí que queremos comentar que, sobre curvas elípticas, los algoritmos de ataque al logaritmo discreto están mucho menos desarrollados que sobre los enteros y, por tanto, se pueden conseguir los mismos niveles de seguridad utilizando claves de tamaño más pequeño. Esto hace la criptografía con curvas elípticas especialmente atractiva para los dispositivos de capacidad reducida. Y se supone que esta ventaja irá aumentando en el futuro. ◁

Pero volvamos ya a los números enteros. El criptosistema Elgamal fue inventado en 1985 por Taher Elgamal, un matemático estadounidense de origen egipcio. Y, como decíamos, así como la seguridad del RSA se basa en la dificultad para factorizar enteros, la seguridad del criptosistema Elgamal se basa en la dificultad de encontrar el logaritmo discreto módulo un primo grande. Sin más preámbulos, vamos a ver cómo funciona.

Supongamos que Benito quiere poder recibir mensajes secretos. Para ello, se provee de un primo p (suficientemente grande), una raíz primitiva r (módulo p), y un entero a con $0 \leq a \leq p-1$. Asimismo, calcula

$$b \equiv r^a \pmod{p}, \qquad \text{con } 0 \leq b \leq p-1.$$

Benito no comunica a nadie el exponente a, que es su clave privada; pero deja que cualquiera pueda conocer la terna (p, r, b), que es su clave pública.

Cuando Alicia quiere enviar un mensaje cifrado a Benito, coge la clave pública (p, r, b) de éste, y además elige, al azar, un entero k con $1 \leq k \leq p-2$. Como es habitual, podemos suponer que el mensaje que quiere enviar es un entero P que cumple $0 \leq P \leq p-1$ (si es necesario, descompone el mensaje en bloques del tamaño requerido). Para cifrar P, Alicia calcula los enteros γ y δ como sigue:

$$\gamma \equiv r^k \pmod{p}, \qquad \text{con } 1 \leq \gamma \leq p-1, \tag{3.41}$$

y

$$\delta \equiv P \cdot b^k \pmod{p}, \qquad \text{con } 0 \leq \delta \leq p-1. \tag{3.42}$$

El mensaje cifrado correspondiente a P es el par ordenado (γ, δ), que transmite a Benito. No se preocupa de que el canal de transmisión sea seguro (es decir, que no toma ninguna precaución para evitar que algún espía pueda «escuchar» el mensaje). Obsérvese que, en esencia, el mensaje cifrado es el $\delta \equiv P \cdot b^k \pmod{p}$, pero hay que transmitirlo junto con el número auxiliar γ (aunque no tiene demasiada importancia, sí que queremos hacer notar que esto hace que la longitud del mensaje cifrado sea el doble que la del mensaje sin cifrar, con lo cual la transmisión es más lenta).

Benito recibe el mensaje cifrado (γ, δ). ¿Qué hacer para descifrarlo y recuperar P? Lo importante ahora es que, como Benito conoce a, puede calcular

$$\beta \equiv \gamma^a \pmod{p}, \qquad \text{con } 1 \leq \beta \leq p-1$$

y su inverso (módulo p)

$$\beta^{-1} \equiv \gamma^{p-1-a} \pmod{p}, \qquad \text{con } 1 \leq \beta^{-1} \leq p-1$$

(obsérvese que $\gamma^a \cdot \gamma^{p-1-a} = \gamma^{p-1} \equiv 1 \pmod{p}$ por el teorema pequeño de Fermat). Una vez conocido β^{-1}, Benito es capaz de obtener P sin más que evaluar $\beta^{-1}\delta$ módulo p, ya que

$$\beta^{-1}\delta \equiv P \pmod{p}, \tag{3.43}$$

y esto descifra el mensaje. Veamos que, efectivamente, tal como queríamos comprobar, la relación (3.43) está garantizada:

$$\beta^{-1}\delta \equiv (\gamma^a)^{-1} P b^k \equiv (r^{ka})^{-1} P b^k \equiv \left((r^a)^k\right)^{-1} P b^k \equiv (b^k)^{-1} P b^k \equiv P \pmod{p}.$$

Si un espía que conoce p, r, b, γ y δ quiere encontrar P, lo que en principio necesitaría es hallar el a que hace $r^a \equiv b \pmod{p}$. Es decir, se trata de saber calcular el logaritmo discreto, pero no se dispone de ningún método efectivo de hacerlo cuando los números involucrados son suficientemente grandes (a este respecto, véanse los comentarios del final del apartado 3.5.2). Por supuesto, esto no asegura que no haya otro camino alternativo de llegar a P —y de hecho no se ha probado la equivalencia entre la seguridad de Elgamal y el problema del cálculo de logaritmos discretos en $\mathbb{Z}_p$—, pero por lo que hasta ahora se sabe no parece haber otra forma más rápida de lograrlo, siempre que se adopten las debidas precauciones.

Lo que no podemos dejar de lado es que el criptosistema Elgamal requiere tomar una precaución importante —crucial—, que hasta ahora no hemos mencionado. Esta precaución es que Alicia tiene que elegir al azar un k cada vez que va a cifrar un mensaje.[19] Y es que si utiliza el mismo k para cifrar diversos P, el sistema es muy

[19] En el tema anterior —sección 2.4— vimos métodos para generar números pseudoaleatorios que ahora pueden resultar muy útiles.

vulnerable. La razón es sencilla de entender. Supongamos que hemos usado el mismo k para cifrar dos mensajes P_1 y P_2, y que sus correspondientes mensajes cifrados son los pares (γ_1, δ_1) y (γ_2, δ_2); en particular, se cumplirá

$$\delta_1 \equiv P_1 b^k, \quad \delta_2 \equiv P_2 b^k \qquad (\text{mód } p).$$

De aquí,

$$\delta_1/\delta_2 \equiv P_1/P_2 \pmod{p} \tag{3.44}$$

(donde estamos usando «/» de manera informal con el sentido de «multiplicar por el inverso módulo p»), y cualquiera que conociese P_1 podría calcular fácilmente P_2. No hay que discurrir demasiado para imaginar situaciones en las que un mensaje previamente secreto P_1 deja de serlo por algún motivo; en ese caso, la relación (3.44) permitiría descifrar todos los mensajes futuros sin necesidad de usar la clave privada.

3.6.1. Firma digital con Elgamal

Como ocurría con el criptosistema RSA (véase el apartado 2.9.5, en el tema anterior), a partir del criptosistema Elgamal también vamos a poder diseñar un esquema de firma digital. Aunque, esta vez, el método va a ser más enrevesado (con RSA, el procedimiento de cifrar y el de firmar son iguales salvo por el uso inverso de claves privadas y públicas, pero con Elgamal no va a ser así). Recordemos qué es lo que se pretende con una firma digital.

Supongamos que Alicia desea enviar mensajes de manera que cualquiera que los reciba pueda estar seguro de que realmente proceden de ella, y no de alguien que intenta suplantarla. Además (aunque realmente es lo mismo interpretado de manera distinta), los receptores de mensajes de Alicia quieren tener un método que haga que Alicia no pueda negar falsamente que un mensaje ha sido enviado por ella. Esto es lo que se va a conseguir si Alicia añade a sus mensajes una firma digital que sólo ha podido calcular ella.

La idea del esquema de firma digital Elgamal es que, antes de enviar su mensaje, Alicia lo va a firmar usando para ello su clave Elgamal privada. Y que, por otra parte, existirá un procedimiento basado en la clave pública de Alicia tal que, cuando alguien reciba el mensaje, podrá comprobar que efectivamente viene de Alicia, pues está firmado con la clave privada que sólo conoce ella (de lo contrario, las dos claves no hubieran «encajado»). Vamos con ello.

Supongamos que Alicia, que desea firmar los mensajes que envía, tiene clave privada a, y que su clave pública es la terna (p, r, b), de modo tal que p es un primo grande, r una raíz primitiva módulo p, $0 \leq a \leq p-1$, y se cumple

$$b \equiv r^a \pmod{p}, \qquad \text{con } 0 \leq b \leq p-1.$$

Además de congruencias módulo p, en la firma digital se van a utilizar también congruencias módulo $\varphi(p) = p-1$; en el mecanismo interno del método, serán importantes en las operaciones entre exponentes que hacen funcionar la comprobación de la firma.

Para firmar un mensaje P (con $0 \leq P \leq p-1$), Alicia comienza eligiendo al azar un entero k que satisfaga $\mathrm{mcd}(k, p-1) = 1$.[20] En el tema anterior veíamos que, por ser $\mathrm{mcd}(k, p-1) = 1$, existe k' tal que $k'k \equiv 1 \pmod{p-1}$, y además Alicia dispone de un procedimiento eficiente basado en el algoritmo de Euclides para encontrarlo (en concreto, véanse la proposición 2.1 y la sección 2.2).

Entonces, la firma del mensaje P es el par (γ, s) definido como sigue:

$$\gamma \equiv r^k \pmod{p}, \qquad \text{con } 0 \leq \gamma \leq p-1, \tag{3.45}$$

y

$$s \equiv (P - a\gamma)k' \pmod{p-1}, \qquad \text{con } 0 \leq s \leq p-2. \tag{3.46}$$

Nótese que los dos números γ y s dependen del entero aleatorio k, y además para calcular s se ha usado a, que sólo Alicia conoce. Cuando Alicia envía el mensaje P, le adjunta su firma (γ, s).

Cualquiera que reciba el mensaje P junto con su firma (γ, s) puede hacer lo siguiente. Calcula V_1 y V_2 mediante

$$\begin{aligned} V_1 &\equiv \gamma^s b^\gamma \pmod{p}, \qquad \text{con } 0 \leq V_1 \leq p-1, \\ V_2 &\equiv r^P \pmod{p}, \qquad \text{con } 0 \leq V_2 \leq p-1. \end{aligned}$$

Ahora, ya sólo le resta chequear si $V_1 = V_2$ o no; si el mensaje está correctamente firmado con la clave privada de Alicia (y por tanto proviene de ella), esa igualdad será cierta; en otro caso, no. Veamos el porqué de esto.

Cuando (γ, s) provienen efectivamente de efectuar los cálculos (3.45) y (3.46), utilizando el teorema pequeño de Fermat se tendrá

$$\begin{aligned} V_1 &\equiv \gamma^s b^\gamma \equiv \gamma^{(P-a\gamma)k'} b^\gamma \equiv \left(\gamma^{k'}\right)^{P-a\gamma} b^\gamma \equiv \left(r^{kk'}\right)^{P-a\gamma} b^\gamma \\ &\equiv r^{P-a\gamma} b^\gamma \equiv r^P (r^a)^{-\gamma} b^\gamma \equiv r^P b^{-\gamma} b^\gamma \equiv r^P \equiv V_2 \pmod{p}, \end{aligned}$$

y como $0 \leq V_1 \leq p-1$ y $0 \leq V_2 \leq p-1$, deberá ser $V_1 = V_2$.

Si un suplantador de Alicia (que por tanto no conoce a) intenta firmar P, podrá elegir el entero aleatorio k y calcular $\gamma \equiv r^k \pmod{p}$ mediante la clave pública (p, r, b). Para completar la firma necesitaría encontrar $s \equiv (P - a\gamma)k' \pmod{p-1}$;

[20] Obsérvese que al usar Elgamal para cifrar —no como ahora en la firma digital— no hace falta que k cumpla $\mathrm{mcd}(k, p-1) = 1$.

hallar k' no es difícil, pero averiguar a partiendo de lo que él conoce —que es r y b en $b \equiv r^a \pmod{p}$— sería calcular el logaritmo discreto módulo p, que es un problema inabordable. Sin a, la posibilidad de dar con el s adecuado se desvanece; el suplantador podría elegir s al azar, pero la probabilidad de que tal s funcione y haga $V_1 = V_2$ es $1/p$, que es prácticamente nula cuando p es grande.

Debemos recalcar una cuestión importante. Y es que —del mismo modo que ocurría con la criptografía basada en el Elgamal—, Alicia debe elegir un entero k diferente cada vez que firma un mensaje. De lo contrario es muy fácil descubrir su clave privada a. En el ejercicio 3.30 se propone al lector que vea cómo hacerlo. En general, cuando alguien utiliza el esquema de firma digital Elgamal de manera sistemática, debe tener mucho cuidado y seguir un buen procedimiento aleatorio (pseudoaleatorio) para elegir los sucesivos k de cada firma. Así se asegura que no se va a poder deducir k o información parcial sobre k. Ya no lo vamos a analizar con más detalle, ni precisaremos a qué nos referimos, pero malas selecciones de k pueden dar lugar a fugas de información que facilitan que un atacante deduzca la clave secreta a. Muchos más detalles sobre seguridad de la firma digital Elgamal se pueden encontrar en [65, § 11.5.2].

Nota 3.34. En todo este apartado, los mensajes que ha enviado Alicia al firmar con el esquema Elgamal están sin cifrar. Por supuesto, si Alicia quiere enviar un mensaje secreto a Benito, puede cifrarlo con la clave pública de Benito antes de firmarlo. Cuando Benito lo recibe, comprueba que contiene una firma de Alicia correcta y luego lo descifra con su clave privada. ◁

3.7. Ejercicios

3.1. Sea $f(x)$ un polinomio con coeficientes enteros. Prueba que, si a, b y m son tres números enteros (con $m > 1$) y $a \equiv b \pmod{m}$, también deberá cumplirse $f(a) \equiv f(b) \pmod{m}$. (En otras palabras, f está bien definido sobre cada $\mathbb{Z}_m$.)

3.2. Como aplicación del ejercicio anterior (¡se pide hacerlo de esa manera!), demuestra que la ecuación $x^5 - 8x^4 - x^2 + 5x - 3 = 0$ no tiene soluciones enteras.

3.3. Demuestra[21] que no existe ningún polinomio $f \in \mathbb{Z}[x]$ de grado mayor o igual que 1 tal que $f(n)$ sea un número primo para todo $n \in \mathbb{Z}$.

[21] Este bonito resultado —y su extensión del ejercicio siguiente— es debido a Christian Goldbach, quien lo probó en 1752. Por otra parte, en el ejercicio 9.21 del tema 9 guiamos al lector para que construya otra demostración de este hecho, en la que ya no aparecen explícitamente las congruencias.

Sugerencia. Toma $a \in \mathbb{Z}$, usa el ejercicio 3.1 módulo $p = f(a)$, y construye una demostración por reducción al absurdo.

3.4. Extiende el resultado del ejercicio anterior mostrando que tampoco existe ningún polinomio $f \in \mathbb{Z}[x]$ de grado mayor o igual que 1 tal que $f(n)$ sea primo para todo $n \geq n_0$ a partir de un cierto $n_0 \in \mathbb{Z}$ (en particular, para $n \in \mathbb{N}$).

3.5. Acota el número de soluciones que puede tener la ecuación

$$x^{45} + 12x^{36} + 6x^{26} + 11x^{14} - bx^4 + 8x - 2 \equiv 0 \pmod{23},$$

en la que b representa un número entero arbitrario. Resuélvela cuando $b = 6$.
Sugerencia. No te quedes en una mera aplicación del teorema de Lagrange (teorema 3.1).

3.6. Sea $f(x)$ un polinomio de grado n con coeficientes enteros y p un número primo. Prueba que, si la congruencia $f(x) \equiv 0 \pmod{p}$ tiene más de n soluciones (en $\mathbb{Z}_p$), entonces cada coeficiente de $f(x)$ es divisible por p.

3.7. Prueba que, para cada primo p, todos los coeficientes del polinomio

$$f(x) = (x-1)(x-2)\cdots(x-p+1) - x^{p-1} + 1$$

son divisibles por p. A partir de aquí, da una demostración del teorema de Wilson —si p es primo, entonces $(p-1)! \equiv -1 \pmod{p}$— que sea distinta a la que hacíamos en el teorema 2.22, en el tema 2.
Sugerencia. Aplica el teorema pequeño de Fermat a $x^{p-1} - 1$ y luego el ejercicio anterior.

3.8. Sea $f(x)$ un polinomio con coeficientes enteros y grado n y, como es habitual, denotemos $f^{(j)}$ a su derivada j-ésima (con j un entero positivo). Prueba que el polinomio $f^{(j)}(x)/j!$ también tiene coeficientes enteros.

3.9. Sea p un primo impar y la ecuación $ax^2 + bx + c \equiv 0 \pmod{p}$ con a, b, c enteros y $p \nmid a$; además, denotemos $d = b^2 - 4ac$. Prueba lo siguiente: si $d \equiv 0 \pmod{p}$, la ecuación tiene una única solución módulo p; si d es resto cuadrático módulo p, tiene dos soluciones; si no lo es, no tiene ninguna.
Sugerencia. Efectúa el cambio de variable $y = 2ax + b$.

3.10. Demuestra la fórmula (3.16).

3.11. Sea b un entero no divisible por el primo impar p. Prueba que

$$\sum_{k=1}^{p-1} (kb \,|\, p) = 0.$$

3.12. Sea a un resto cuadrático módulo un primo p de la forma $p = 4n + 3$ (para algún entero $n \geq 0$). Demuestra que las soluciones de la ecuación $x^2 \equiv a \pmod{p}$ son los x que cumplen $x \equiv \pm a^{n+1} \pmod{p}$.

3.13. Sea a un resto cuadrático módulo un primo p de la forma $p = 8n + 5$ (para algún entero $n \geq 0$). Demuestra que las soluciones de la ecuación $x^2 \equiv a \pmod{p}$ son los x que cumplen $x \equiv \pm a^{n+1} \pmod{p}$, o los que cumplen $x \equiv \pm 2^{2n+1} a^{n+1} \pmod{p}$.

3.14. Busca una ecuación cuadrática $x^2 \equiv b \pmod{P}$ que no tenga soluciones, pero que el símbolo de Jacobi cumpla $(b \,|\, P) = 1$.

3.15. Comprueba que 341 es pseudoprimo en base 2, pero no es pseudoprimo de Euler en base 2.

3.16. Demuestra la proposición 3.15.

3.17. Demuestra la proposición 3.16.

3.18. Prueba que una condición necesaria para que $2^m + 1$ con $m \in \mathbb{N}$ sea primo es que el exponente m tenga la forma 2^n con $n \in \mathbb{N} \cup \{0\}$ (y que, por tanto, nos encontremos ante un número de Fermat).

3.19. Prueba que los números de Fermat satisfacen las siguientes relaciones de recurrencia:

(a) $F_n = (F_{n-1} - 1)^2 + 1$,
(b) $F_n = F_{n-1}^2 - 2(F_{n-2} - 1)^2$,
(c) $F_n = F_0 \cdots F_{n-1} + 2$,
(d) $F_n = F_{n-1} + 2^{2^{n-1}} F_0 \cdots F_{n-2}$.

3.20. Prueba que dos números de Fermat distintos son siempre primos entre sí (este resultado se conoce a menudo con el nombre de teorema de Goldbach, quien se dio cuenta de ello en 1730). Basándote en ello, da una demostración de la existencia de infinitos números primos.

3.21. Sea $n = p - 1$ o $n = p - 2$ con p primo impar. Prueba que el número de Cullen $n \cdot 2^n + 1$ es divisible por p. Aplicando esa propiedad, ¿qué números de Cullen podrían ser primos para $n \leq 30$?
Sugerencia. Utiliza el teorema pequeño de Fermat.

3.22. Prueba que número de Cullen $n \cdot 2^n + 1$ es divisible por $p = 2n - 1$ si p es un primo de la forma $8j \pm 3$. Para $n \leq 30$, de entre los candidatos a primos de Cullen que quedaban al aplicar el ejercicio anterior, ¿cuáles quedarían ahora?
Sugerencia. Utiliza el símbolo de Legendre $(2 \,|\, p)$ y el criterio de Euler.

3.23. Comprueba que una manera alternativa de enunciar el test de Pépin es la siguiente: para $n \geq 1$, y dado el número de Fermat $F_n = 2^{2^n} + 1$, definamos los números S_k de manera recurrente mediante

$$S_0 = 3, \qquad S_{k+1} = S_k^2, \quad k \geq 0;$$

entonces, F_n es primo si y sólo si $S_{2^n-1} \equiv -1 \pmod{F_n}$.

3.24. Generaliza las proposiciones 3.25 y 3.26 al caso en que q no es primo.

3.25. Sea el número de Mersenne $M_p = 2^p - 1$ con p primo impar. Prueba que $(3 \,|\, M_p) = (-2 \,|\, M_p) = -1$, donde $(\cdot \,|\, M_p)$ es el símbolo de Jacobi.

3.26. Sea $m \geq 2$ un entero fijo, $a \in \mathbb{Z}$ tal que $\operatorname{mcd}(a, m) = 1$, y r, s dos enteros no negativos. Prueba que $a^r \equiv a^s \pmod{m}$ si y sólo si $r \equiv s \pmod{\operatorname{ord}_m(a)}$.

3.27. Sea p un primo de la forma $p = 8n + 1$ para algún n entero positivo, y r una raíz primitiva módulo p. Demuestra lo siguiente:

(a) Las soluciones de la ecuación $x^2 \equiv 2 \pmod{p}$ son los x que cumplen $x \equiv \pm(r^{7n} + r^n) \pmod{p}$.

(b) Las soluciones de la ecuación $x^2 \equiv -2 \pmod{p}$ son los x que cumplen $x \equiv \pm(r^{7n} - r^n) \pmod{p}$.

Sugerencia. Comprueba previamente que $r^{4n} \not\equiv 1 \pmod{p}$ y $r^{6n} + r^{2n} \equiv 0 \pmod{p}$.

3.28. Demuestra el teorema 3.32.

3.29. Sea q un primo de la forma $q = 4p + 1$ con p un primo impar. Demuestra que 2 es raíz primitiva módulo q.
Sugerencia. Comprueba primero que 2 no es resto cuadrático módulo q.

3.30. Supón que, utilizando la firma digital Elgamal con clave pública (p, r, b) y clave privada a, se usa dos veces el mismo entero k para firmar dos mensajes P_1 y P_2, siendo estas firmas los pares (γ_1, s_1) y (γ_2, s_2) respectivamente. Demuestra que, si $s_1 \not\equiv s_2 \pmod{p-1}$ (lo cual, al azar, es bastante probable), se puede encontrar k. Y que, una vez que k ha sido descubierto, es fácil también averiguar la clave privada a.

Tema 4
Aproximación de un número real por racionales

4.1. Teorema de Dirichlet y consecuencias
4.2. Orden de aproximación. Teorema de Liouville y consecuencias
4.3. Refinamientos del teorema de Liouville
4.4. Fracciones de Brocot
4.5. Teorema de Hurwitz
4.6. La constante de Lagrange de un irracional
4.7. Cubrimientos de la recta real
4.8. Ejercicios

Estamos acostumbrados a identificar el conjunto de los números reales, $\mathbb{R}$, con una recta, a la que denominamos *recta real*. Si en la recta fijamos dos puntos 0 y 1, cada número real se corresponde con un único punto de la recta. Algunos de estos puntos representan un número racional; otros no. Los conjuntos $\mathbb{Q}$ y $\mathbb{R} \setminus \mathbb{Q}$ se entremezclan de manera muy complicada, y entre cada dos puntos cualesquiera de uno de estos conjuntos, siempre existen infinitos puntos del otro.

Para muchas cuestiones prácticas, los números que no son racionales resultan problemáticos, pues no podemos tratarlos convenientemente. Por ejemplo, si tenemos que efectuar operaciones (aunque sólo sean las cuatro elementales) en las que aparecen bastantes números irracionales, deberemos ir arrastrándolos todos, pues —normalmente— no podremos escribir el resultado como una sola expresión más compacta. Además, la representación decimal de los números irracionales tiene infinitos dígitos (y no es periódica), luego no existe un procedimiento genérico y finito de determinarlos todos (ni de escribirlos o almacenarlos, si pensamos en un papel o un ordenador). En todo esto, los racionales se comportan mucho mejor. Sabemos efectuar operaciones elementales con racionales, y vuelve a salir un nuevo racional cuya expresión podemos encontrar sin problemas. Además, la representación decimal de un racional es —siempre— finita o periódica, así que podemos describir fácilmente todos los decimales. Es más, para determinar cualquier número racional, bastan dos datos: numerador y denominador. No hay ninguna estructura similar aplicable a los irracionales.

Así pues, resulta útil poder aproximar números reales genéricos por racionales. El conjunto $\mathbb{Q}$ es denso en $\mathbb{R}$, luego cualquier número real se puede aproximar por racionales tanto como deseemos. Es más, existe un procedimiento trivial de llevarlo a la práctica. Dado $r \in \mathbb{R}$, sea su expresión decimal

$$r = d_0 + \sum_{j=1}^{\infty} \frac{d_j}{10^j} = d_0.d_1d_2d_3d_4d_5\ldots,$$

donde $d_0 \in \mathbb{Z}$ y $d_j \in \{0,1,2,3,4,5,6,7,8,9\}$, para todo $j \geq 1$. Entonces, los números racionales

$$d_0 + \sum_{j=1}^{n} \frac{d_j}{10^j} = d_0 + \frac{d_1d_2\ldots d_n}{10^n}, \quad n \geq 1,$$

aproximan r todo lo que deseemos, tanto más cuanto mayor sea n. Pero ésta no es una buena forma de aproximar, pues los denominadores que van apareciendo son enormes. Por poner un ejemplo, $314/100 = 3.14$ es una de tales aproximaciones de $\pi = 3.141\,592\ldots$; pero $22/7 = 3.142\,857\ldots$ es una aproximación algo mejor (el



Christiaan Huygens (1629–1695)



Peter Gustav Lejeune Dirichlet (1805–1859)



Joseph Liouville (1809–1882)

valor absoluto del error es menor), y su denominador es mucho más pequeño (y sigue siendo más pequeño si simplificamos $314/100 = 157/50$, no vayamos a pensar que es éste el problema). Por tanto, parece más adecuado usar $22/7$ como aproximación racional de π.

Estos tipos de aproximación también pueden resultar útiles para racionales. Por ejemplo, si tenemos un racional h/k, y k es enorme, en algunas circunstancias puede resultar conveniente encontrar otro racional p/q más fácilmente tratable, y de tal modo que el error $|h/k - p/q|$ sea pequeño (dentro de un cierto margen que nos podamos permitir).

Históricamente, ya en el siglo XVII, el matemático y astrónomo holandés Christiaan Huygens se enfrentó a estos problemas para fabricar planetarios; también Achille Brocot (un relojero francés) en el XIX, para construir relojes. En ambos casos, las velocidades relativas entre diversas partes del mecanismo se deben ajustar —ya sea de manera exacta o aproximada— a un número racional. Ahora, hay que construir engranajes que se encarguen de transmitir el movimiento. Esto es más sencillo cuanto más pequeños sean el numerador y denominador de la fracción que se debe reproducir mediante engranajes, pues estos números se corresponden con la cantidad de dientes que deben tener.

Resumiendo, estamos interesados en encontrar aproximaciones por racionales de manera que el denominador sea lo más pequeño posible. La pregunta clave en todo este asunto es ¿cómo se relaciona el tamaño permitido a los denominadores con el error de la aproximación? Más aún, ¿todos los números reales se pueden aproximar igual, o unos mejor que otros?; ¿sirve esto para identificar distintos tipos de números? Veremos que sí: permite distinguir entre números racionales, irracionales algebraicos, y trascendentes. Aquí tiene una gran importancia la «velocidad de aproximación por racionales» que puede conseguirse, en un sentido que ya precisa-

remos. Comenzando por el alemán Lejeune Dirichlet y el francés Joseph Liouville a mediados del siglo XIX, también grandes matemáticos más recientes han trabajado en estos temas; los frutos logrados fueron meritorios, lo que les hizo recibir importantes honores.

El tipo de técnicas que aquí estudiamos entra dentro de lo que se suele denominar *aproximación diofántica* o *análisis diofántico*. Diofanto de Alejandría fue un matemático griego del que sólo nos han llegado referencias a través de fuentes indirectas, y que vivió —aproximadamente— en el siglo III d.C. Publicó trece libros, de los que se conservan seis; es en sus obras donde aparece, por primera vez en la historia, una cierta notación algebraica. En su honor, el adjetivo «diofántico» se aplica, con bastante generalidad, cuando se estudian propiedades que tienen que ver con enteros o racionales; el caso más típico es el de encontrar las soluciones enteras (o racionales) de una ecuación, las únicas que Diofanto se preocupaba en buscar. Así, estas ecuaciones se conocen con el nombre de ecuaciones diofánticas. Quizás sorprendentemente, las ecuaciones diofánticas y la aproximación diofántica se relacionan en algo más que en el nombre; veremos algo sobre ello en la sección 4.3.

Aparte de lo que exponemos en este tema, otra útil herramienta en el estudio de la aproximación racional son las fracciones continuas; pero esto lo dejamos para el tema siguiente. Además —y ya lo hemos comentado antes—, todo esto emparenta de manera muy cercana con la naturaleza aritmética de los números. En particular, los resultados de este tema daban lugar a criterios de irracionalidad y trascendencia. Veremos mucho más sobre ello en el tema 6, aunque —a menudo— utilizando técnicas distintas.

Lo que aquí se presenta aparece también en bastantes textos dedicados a la teoría de números en general, como [41, 54, 78], cada uno con su enfoque específico, y haciendo más hincapié en unos u otros resultados. En [8], un libro de muy agradable lectura, se pueden encontrar numerosas referencias, notas históricas, y extensiones; del mismo autor, pero más avanzado, es [9]. Otro delicioso libro es [22]; prácticamente, no contiene demostraciones, lo que compensa con numerosas propuestas para el lector y una buena bibliografía. Finalmente, si el lector busca una manera rápida y rigurosa de avanzar por la teoría, llegando hasta los resultados más potentes, puede consultar [75, capítulo 1].

4.1. Teorema de Dirichlet y consecuencias

Comencemos con un interesante resultado que, aunque es evidente, da mucho juego. Como se ve, tiene un enunciado muy informal, que alude a la ilustración que mostramos en la figura 4.1.



Figura 4.1: El principio del palomar.

Teorema 4.1. Principio de Dirichlet o principio del palomar.
Si hay un palomar con más palomas que agujeros, en algún agujero tendrán que alojarse, al menos, dos palomas.

Se utiliza mucho en matemática recreativa —en los ejercicios aparecen varios problemas que se resuelven usándolo— pero, también, en temas más académicos. Sin más preámbulos, apliquémoslo a la teoría de números.

Teorema 4.2. Dirichlet, 1842.
Dado un número real ξ y un número natural N, existe alguna fracción p/q (con $p \in \mathbb{Z}$ y $q \in \mathbb{N}$) tal que $1 \leq q \leq N$ y

$$\left|\xi - \frac{p}{q}\right| < \frac{1}{qN}.$$

Demostración. Consideremos los $N+1$ números $m\xi - \lfloor m\xi \rfloor$, $m = 0, 1, 2, \ldots, N$; todos ellos verifican $0 \leq m\xi - \lfloor m\xi \rfloor < 1$. Si tomamos los N intervalos $[0, 1/N)$, $[1/N, 2/N)$, $\ldots$, $[(N-1)/N, 1)$, como hay más números que intervalos, tendrá que haber un intervalo con, al menos, dos números. Digamos que esto ocurre con los correspondientes a m y m', siendo $m < m'$. Entonces,

$$|(m'\xi - \lfloor m'\xi \rfloor) - (m\xi - \lfloor m\xi \rfloor)| < 1/N,$$

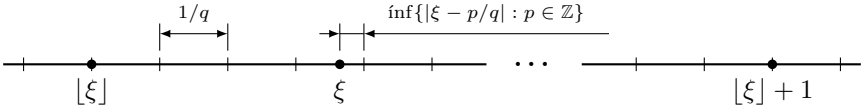


Figura 4.2: Para $q \in \mathbb{N}$ fijo, puntos racionales p/q que aproximan a un irracional ξ.

es decir,

$$\left| \xi - \frac{\lfloor m'\xi \rfloor - \lfloor m\xi \rfloor}{m' - m} \right| < \frac{1}{(m'-m)N}.$$

Por tanto, basta tomar $p = \lfloor m'\xi \rfloor - \lfloor m\xi \rfloor$ y $q = m' - m$. □

Nota 4.3. En el enunciado del teorema de Dirichlet podemos exigir que p/q sea irreducible, ya que si $p/q = p_0/q_0$ con $q_0 < q$ entonces

$$\left| \xi - \frac{p_0}{q_0} \right| = \left| \xi - \frac{p}{q} \right| < \frac{1}{qN} < \frac{1}{q_0 N}.$$

De aquí que, si p/q satisface la desigualdad, también lo hace la fracción p_0/q_0. ◁

El teorema de Dirichlet nos permitirá extraer interesantes consecuencias. Pero, antes, enunciemos un lema, que nos resultará útil en varias ocasiones.

Lema 4.4.

Dado $\xi \in \mathbb{R}$, sea $q \in \mathbb{N}$ fijo. Entonces:

(a) Para cada constante $C_q > 0$, sólo puede existir una cantidad finita de números $p \in \mathbb{Z}$ (quizás ninguno) tales que $|\xi - p/q| < C_q$.

(b) Es más, si $\xi \in \mathbb{R} \setminus \mathbb{Q}$, existe un C_q suficientemente pequeño tal que $|\xi - p/q| > C_q$ para todo $p \in \mathbb{Z}$.

Demostración. En primer lugar, nótese que

$$\left| \xi - \frac{p}{q} \right| < C_q \;\Leftrightarrow\; -C_q < \frac{p}{q} - \xi < C_q \;\Leftrightarrow\; (\xi - C_q)q < p < (\xi + C_q)q,$$

y sólo puede haber una cantidad finita de números $p \in \mathbb{Z}$ que cumplan esto, de donde se sigue (a). Para el apartado (b), es fácil darse cuenta de que, fijado q,

$$\text{ínf} \left\{ \left| \xi - \frac{p}{q} \right| : p \in \mathbb{Z} \right\} > 0$$

(vale la pena pensar en la representación del conjunto anterior en la recta real, tal como se observa en la figura 4.2); esto implica el resultado. □

Con esto, vamos ya con las consecuencias del teorema de Dirichlet. Comencemos con un interesante resultado que muestra lo bien que se puede aproximar un irracional mediante racionales (más adelante, en la sección 4.5, veremos que se puede hacer aún mejor).

Corolario 4.5. Dirichlet, 1842.

Si $\xi \in \mathbb{R} \setminus \mathbb{Q}$, entonces existen infinitas fracciones p/q tales que

$$0 < \left| \xi - \frac{p}{q} \right| < \frac{1}{q^2}$$

(como en el teorema, si se quiere, se puede añadir que las fracciones p/q sean irreducibles).

Demostración. Como $\xi \in \mathbb{R} \setminus \mathbb{Q}$, cualquier p/q cumple $0 < |\xi - p/q|$; así pues, sólo hay que preocuparse de la otra desigualdad. Elegimos un $N_0 \in \mathbb{N}$ cualquiera. Por el teorema de Dirichlet, existe p_0/q_0 tal que $1 \leq q_0 \leq N_0$ y

$$\left| \xi - \frac{p_0}{q_0} \right| < \frac{1}{q_0 N_0} \leq \frac{1}{q_0^2}.$$

Ahora, el apartado (b) del lema 4.4 nos garantiza que existirá un $N_1 \in \mathbb{N}$ grande, de forma que

$$\left| \xi - \frac{p}{q} \right| > \frac{1}{N_1} \qquad \text{para todo } p \in \mathbb{Z}, \text{ y todo } q \in \{1, 2, \ldots, N_0\}.$$

De nuevo por el teorema, existirá p_1/q_1 tal que $1 \leq q_1 \leq N_1$ y

$$\left| \xi - \frac{p_1}{q_1} \right| < \frac{1}{q_1 N_1},$$

de donde necesariamente $q_1 > N_0$ (luego $q_1 \neq q_0$) y además

$$\left| \xi - \frac{p_1}{q_1} \right| < \frac{1}{q_1^2}.$$

Reiterando el proceso, obtenemos infinitas fracciones. □

En la construcción de la demostración anterior se ve que $q \to \infty$. Necesariamente tiene que ser siempre así, ya que, para cada q, sólo puede haber una cantidad finita de $p \in \mathbb{Z}$ que cumplan

$$\left| \xi - \frac{p}{q} \right| < \frac{1}{q^2}$$

(basta aplicar el apartado (a) del lema 4.4 con $C_q = 1/q^2$).

Nota 4.6. Si $\xi \in \mathbb{Q}$, ya no se cumple la tesis del corolario. Es más, si $\xi = h/k$ y $\xi \neq p/q$, entonces

$$\left|\xi - \frac{p}{q}\right| = \left|\frac{h}{k} - \frac{p}{q}\right| = \frac{|qh - pk|}{kq} \geq \frac{1}{kq}, \tag{4.1}$$

donde la desigualdad se sigue del hecho $0 \neq |qh - pk| \in \mathbb{Z}$. Por lo tanto, si $0 < |\xi - p/q| < 1/q^2$, entonces $1/(kq) < 1/q^2$ y $q < k$. Luego sólo hay una cantidad finita de fracciones p/q que pueden verificar $0 < |\xi - p/q| < 1/q^2$ (el denominador q debe cumplir $q < k$ y, para cada q, el apartado (a) del lema sólo permite una cantidad finita de posibles p). ◁

4.1.1. Un criterio de irracionalidad

Aún podemos sacar más partido del teorema de Dirichlet si combinamos el corolario 4.5 con (4.1). En efecto; de este modo obtenemos lo siguiente:

Teorema 4.7. Criterio de irracionalidad.
Sea $\xi \in \mathbb{R}$. Entonces, $\xi \in \mathbb{R}\setminus\mathbb{Q}$ si y sólo si existe una sucesión de racionales $\{p_n/q_n\}_{n=1}^{\infty}$ tal que[1]

$$0 < \left|\xi - \frac{p_n}{q_n}\right| = o\left(\frac{1}{q_n}\right).$$

De una manera informal podemos decir que, si un número se puede aproximar *rápidamente* por racionales, ese número es irracional. Veamos una aplicación de este resultado:

Teorema 4.8. Euler, 1737.
El número e es irracional.

Demostración. A partir de la expresión $e = \sum_{k=0}^{\infty} 1/k!$ consideremos, para cada $n \in \mathbb{N}$, la n-ésima suma parcial; esto es, la fracción $\sum_{k=0}^{n} 1/k! = p_n/n!$. Entonces

$$\begin{aligned} 0 < \left|e - \sum_{k=0}^{n} \frac{1}{k!}\right| &= \frac{1}{(n+1)!} + \frac{1}{(n+2)!} + \frac{1}{(n+3)!} + \cdots \\ &= \frac{1}{(n+1)!}\left(1 + \frac{1}{n+2} + \frac{1}{(n+2)(n+3)} + \cdots\right) \end{aligned}$$

[1] Dadas dos sucesiones $\{a_n\}_{n=1}^{\infty}$ y $\{b_n\}_{n=1}^{\infty}$, la notación $a_n = o(b_n)$ significa que $a_n/b_n \to 0$ cuando $n \to \infty$. Por ejemplo, $1/q_n^2 = o(1/q_n)$ cuando $n \to \infty$ (ya que entonces $q_n \to \infty$).

$$\leq \frac{1}{(n+1)!}\left(1+\frac{1}{n+2}+\frac{1}{(n+2)^2}+\cdots\right)$$
$$= \frac{1}{(n+1)!}\cdot\frac{1}{1-1/(n+2)} = \frac{1}{n!}\cdot\frac{n+2}{(n+1)^2}$$
$$= o\left(\frac{1}{n!}\right). \qquad \square$$

Nota 4.9. En realidad, no hace falta utilizar el criterio de irracionalidad del teorema 4.7 para demostrar, a partir de la serie $e=\sum_{k=0}^{\infty}1/k!$, que e es irracional. En el tema 6 (concretamente, en el teorema 6.7) vemos cómo hacerlo sin más que usar la serie y una observación elemental: no existe ningún entero entre 0 y 1. Por otra parte, ninguna de estas demostraciones se parece en nada a la original del genial matemático suizo Leonhard Euler, que utilizaba desarrollos en fracción continua; una demostración basada en fracciones continuas —similar a la de Euler— se puede encontrar en la sección 5.8 del tema 5. A este respecto, el lector puede estar también interesado en los comentarios históricos de la sección 6.2 del tema 6. $\triangleleft$

4.2. Orden de aproximación. Teorema de Liouville y consecuencias

Sea una función $\varphi\colon \mathbb{N}\to(0,\infty)$. Decimos que un número real ξ tiene una *aproximación por racionales de orden* $\varphi(q)$ si existe alguna constante C, que sólo puede depender de ξ y de la función φ, tal que la desigualdad

$$0<\left|\xi-\frac{p}{q}\right|<C\varphi(q)$$

tiene infinitas soluciones p/q (con $p\in\mathbb{Z}$ y $q\in\mathbb{N}$). El caso más frecuente es $\varphi(q)=1/q^r$ con algún $r>0$. Obsérvese que estamos midiendo con qué velocidad se puede aproximar por racionales el número ξ; una velocidad relativa al tamaño del denominador.[2]

El teorema de Dirichlet (y sus corolarios) demuestra que cualquier irracional tiene orden de aproximación por racionales $1/q^2$; sin embargo, cada número racional tiene aproximación por racionales de orden $1/q$, pero no $1/q^2$. En concreto, en el

[2] Es posible probar que, dada cualquier función $\varphi\colon\mathbb{N}\to(0,\infty)$, existe algún número ξ cuyo orden de aproximación por racionales es $\varphi(q)$. La manera más fácil de hacerlo es por medio de fracciones continuas: véase el ejercicio 5.28, en el tema 5.

corolario 4.5 hemos probado que, si $\xi \in \mathbb{R} \setminus \mathbb{Q}$, la desigualdad

$$0 < \left| \xi - \frac{p}{q} \right| < \frac{1}{q^2}$$

tiene infinitas soluciones, es decir, sirve $C = 1$. Probaremos más adelante (teorema de Hurwitz) que siempre puede ponerse $C = 1/\sqrt{5}$ y que, en general, ésta es la mejor constante (la más pequeña).

Ahora vamos a ver el teorema de Liouville, que proporciona información más fina sobre el orden de aproximación por racionales de un número algebraico. Primero, precisemos los conceptos:

Decimos que un número $\alpha \in \mathbb{R}$ es *algebraico* si es raíz de un polinomio con coeficientes en $\mathbb{Z}$. Cuando éste es el caso, existe (salvo constante multiplicativa) un único polinomio irreducible con coeficientes enteros $f(x) \in \mathbb{Z}[x]$ tal que $f(\alpha) = 0$. Al grado de f se le llama orden de α. A los números no algebraicos los denominamos *trascendentes*.

Obviamente, los números racionales son los algebraicos de orden 1. Los algebraicos de orden 2 —los también denominados *irracionales cuadráticos*—, son los reales de la forma $(a + b\sqrt{d})/c$ con $a, b \in \mathbb{Z}$, $c, d \in \mathbb{N}$, y d libre de cuadrados.

4.2.1. El teorema y sus consecuencias

El teorema de Liouville afirma lo siguiente:

Teorema 4.10. Liouville, 1844.
Si $\alpha \in \mathbb{R}$ es algebraico de orden $n \geq 1$, entonces existe alguna constante $C(\alpha) > 0$ tal que, para cualquier fracción p/q (siendo $p \in \mathbb{Z}$ y $q \in \mathbb{N}$), con $\alpha \neq p/q$, se verifica

$$\left| \alpha - \frac{p}{q} \right| > \frac{C(\alpha)}{q^n}.$$

Aunque, a primera vista, el enunciado de este teorema no parece decir nada especialmente llamativo, veremos que tiene importantes aplicaciones. El caso $n = 1$ (incluido en el enunciado por completitud) alude a valores de α racionales, y no aporta novedades a lo que ya conocíamos (véase la demostración del teorema); cuando $n \geq 2$, la hipótesis $\alpha \neq p/q$ es superflua. Antes de abordar la demostración, analicemos una primera consecuencia.

Bajo las condiciones del teorema de Liouville, la desigualdad

$$0 < \left| \alpha - \frac{p}{q} \right| < \frac{C(\alpha)}{q^n}$$

no tiene soluciones p/q (con $p \in \mathbb{Z}$, $q \in \mathbb{N}$). Asimismo, la desigualdad

$$0 < \left|\alpha - \frac{p}{q}\right| < \frac{1}{q^{n+1}}$$

sólo puede tener un número finito de soluciones, ya que cualquier solución debe cumplir

$$\frac{C(\alpha)}{q^n} < \left|\alpha - \frac{p}{q}\right| < \frac{1}{q^{n+1}},$$

luego $q < 1/C(\alpha)$, y sólo hay una cantidad finita de q (y por tanto, por el apartado (a) del lema 4.4, de p/q) que verifican eso. El mismo argumento permite probar que, para cualquier $C > 0$ y $\varepsilon > 0$, la desigualdad

$$0 < \left|\alpha - \frac{p}{q}\right| < \frac{C}{q^{n+\varepsilon}}$$

sólo puede tener un número finito de soluciones. Es decir:

Corolario 4.11.

Si $\alpha \in \mathbb{R}$ es algebraico de orden $n \geq 2$, entonces α no tiene orden de aproximación por racionales $1/q^{n+\varepsilon}$ para ningún $\varepsilon > 0$.

Demostración del teorema de Liouville. Si $n = 1$, es decir, $\alpha = h/k$ racional, en la nota 4.6 de la página 244 hemos visto que

$$\left|\alpha - \frac{p}{q}\right| \geq \frac{C}{q}$$

con $C = 1/k$. Por tanto, basta tomar una constante $C(\alpha)$ menor que esta C.

Supongamos pues que $n \geq 2$, y sea $f \in \mathbb{Z}[x]$ polinomio irreducible con $f(\alpha) = 0$. Así, tendremos $f(x) = \lambda x^n + a_{n-1}x^{n-1} + \cdots + a_1 x + a_0$, con $\lambda \in \mathbb{N}$ y $a_k \in \mathbb{Z}$; o, también, $f(x) = \lambda(x - \alpha)\prod_{k=2}^{n}(x - \beta_k)$, con $\beta_k \in \mathbb{C} \setminus \mathbb{Q}$.

Dados cualesquiera $p \in \mathbb{Z}$ y $q \in \mathbb{N}$, es trivial que $\alpha \neq p/q$, ya que $\alpha \notin \mathbb{Q}$. Si $|\alpha - p/q| \geq 1$, es claro que

$$\left|\alpha - \frac{p}{q}\right| \geq \frac{1}{q^n},$$

y, para estos p/q, cualquier $C(\alpha) < 1$ sirve. Así que centrémonos en los casos en que $|\alpha - p/q| < 1$.

Cuando esto ocurre, de la desigualdad triangular se deduce que $|p/q| < 1 + |\alpha|$. Ahora, por una parte, $f(p/q) \neq 0$ (las raíces de f son irracionales), luego

$$\left|f\left(\frac{p}{q}\right)\right| = \frac{|\lambda p^n + a_{n-1}p^{n-1}q + \cdots + a_1 pq^{n-1} + a_0 q^n|}{q^n} \geq \frac{1}{q^n}.$$

Por otra,

$$\prod_{k=2}^{n}\left|\frac{p}{q}-\beta_k\right|<\prod_{k=2}^{n}(1+|\alpha|+|\beta_k|),$$

así que

$$\frac{1}{q^n}\le\left|f\left(\frac{p}{q}\right)\right|=\lambda\left|\alpha-\frac{p}{q}\right|\cdot\prod_{k=2}^{n}\left|\frac{p}{q}-\beta_k\right|<\lambda\prod_{k=2}^{n}(1+|\alpha|+|\beta_k|)\cdot\left|\alpha-\frac{p}{q}\right|.$$

Basta entonces tomar

$$C(\alpha)=\frac{1}{\lambda}\prod_{k=2}^{n}(1+|\alpha|+|\beta_k|)^{-1}$$

(nótese que esta constante es, obviamente, menor que 1, luego es válida para el caso $|\alpha-p/q|\ge 1$ analizado al principio de la demostración). □

Una vez probado el teorema, vamos a continuar sacando partido de él.

Nota 4.12. Ya sabíamos que los racionales tienen orden de aproximación por racionales exactamente $1/q$; los irracionales, al menos $1/q^2$. Del teorema de Liouville se deduce que los irracionales cuadráticos tienen orden, exactamente, $1/q^2$; es decir, no tienen orden $\varphi(q)$ si $\varphi(q)=o(1/q^2)$. La demostración de este hecho —similar a la del corolario 4.11— se deja para el lector (ejercicio 4.11). ◁

4.2.2. Un criterio de trascendencia

Sin duda la principal aplicación del teorema de Liouville es la siguiente:

Corolario 4.13. Criterio de trascendencia.
Supongamos que $\alpha\in\mathbb{R}$ y que existe $K>0$ tal que, para todo $m\in\mathbb{N}$, la desigualdad

$$0<\left|\alpha-\frac{p}{q}\right|<\frac{K}{q^m} \tag{4.2}$$

tiene alguna solución p/q con $p\in\mathbb{Z}$ y $q\in\mathbb{N}$, $q\ge 2$. Entonces, α es trascendente. (Por claridad, recalquemos que la constante K del enunciado sólo puede depender de α, no de m, ni de p o q.)

Demostración. Si α no fuera trascendente, sería algebraico de un cierto orden $n\in\mathbb{N}$. Por el teorema de Liouville, existe $C(\alpha)>0$ tal que, si $\alpha\ne p/q$, entonces

$|\alpha - p/q| > C(\alpha)q^{-n}$. Luego, si además se cumple (4.2), deberá ser $C(\alpha)q^{-n} < Kq^{-m}$, es decir,

$$q^{m-n} < \frac{K}{C(\alpha)}.$$

Fijémonos únicamente en los $m \geq n$. Como $q \geq 2$, resulta $2^{m-n} < KC(\alpha)^{-1}$, de donde

$$m < n + \log_2\left(\frac{K}{C(\alpha)}\right).$$

Y esto contradice la hipótesis. □

Para dejar patente la utilidad de este criterio, nos falta un paso fácil: plasmar todo esto en un ejemplo. Por supuesto, lo hizo Liouville, en lo que constituyó la primera prueba rigurosa de que los números trascendentes existían.

Corolario 4.14. Liouville, 1844.

Si $a \in \mathbb{N}$, $a \geq 2$, entonces el número $\alpha = \sum_{k=1}^{\infty} a^{-k!}$ es trascendente.

Demostración. Vamos a encontrar un K que, para cada m, la desigualdad

$$0 < \left|\alpha - \frac{p}{q}\right| < \frac{K}{q^m}$$

tendrá una solución $\frac{p}{q}$ que también identificaremos. En particular, dado m, tomemos

$$\sum_{k=1}^{m-1} \frac{1}{a^{k!}} = \frac{1}{a^{1!}} + \frac{1}{a^{2!}} + \cdots + \frac{1}{a^{(m-1)!}} = \frac{p}{q}$$

con $q = a^{(m-1)!}$. Entonces

$$\begin{aligned}\left|\alpha - \frac{p}{q}\right| &= \sum_{k=m}^{\infty} \frac{1}{a^{k!}} < \sum_{k=m!}^{\infty} \frac{1}{a^k} = \frac{1}{a^{m!}} \sum_{k=0}^{\infty} \left(\frac{1}{a}\right)^k \\ &= \frac{1}{a^{(m-1)!\cdot m}} \frac{1}{1-1/a} = \frac{K}{q^m}\end{aligned}$$

con $K = a/(a-1)$. Así, el corolario 4.13 implica la trascendencia de α. □

En el corolario anterior, el caso más usual es el correspondiente a $a = 10$, que se suele denominar *número de Liouville*. Es el número cuya expresión decimal en base 10 es como sigue:

$$\sum_{k=1}^{\infty} \frac{1}{10^{k!}} = 0.1100010\ldots010\ldots010\ldots,$$

donde los dígitos «1» de la parte fraccionaria están colocados en las posiciones 1!, 2!, 3!, 4!, 5!, ..., y el resto de los decimales son todos «0».

Nota 4.15. En general, se denominan *números de Liouville* a aquéllos cuya trascendencia puede probarse a partir teorema de Liouville. En el ejercicio 4.12 se prueba que, si a es un entero mayor o igual que 2, cada número de la forma

$$\alpha = \sum_{k=1}^{\infty} \frac{d_k}{a^{k!}}$$

con $d_k \in \{0, 1, 2, \ldots, a-1\}$ y con infinitos d_k no nulos, es de Liouville. ◁

4.3. Refinamientos del teorema de Liouville

Hasta ahora, hemos visto lo siguiente (por supuesto, la expresión informal «orden de aproximación por racionales exactamente $1/q^\nu$» se refiere a que ya no es cierto con orden $1/q^{\nu+\varepsilon}$ para ningún $\varepsilon > 0$):

- Los racionales tienen orden de aproximación por racionales exactamente $1/q$.
- Los irracionales, orden de aproximación por racionales al menos $1/q^2$.
- Los irracionales cuadráticos, orden de aproximación exactamente $1/q^2$.
- Los algebraicos de orden n no tienen orden de aproximación $1/q^{n+\varepsilon}$ para ningún $\varepsilon > 0$.

Extrapolando, podríamos llegar a pensar que el orden de aproximación por racionales de los números algebraicos de orden n debería ser $1/q^n$. ¿Será esto verdad?

Más modestos en nuestras suposiciones (y olvidándonos ya del caso de los racionales), planteémoslo del siguiente modo. Sabemos que

- Si $\xi \in \mathbb{R} \setminus \mathbb{Q}$, la desigualdad $0 < |\xi - p/q| < 1/q^2$ tiene infinitas soluciones.
- Si ξ es irracional cuadrático, la desigualdad $0 < |\xi - p/q| < 1/q^{2+\varepsilon}$ tiene infinitas soluciones cuando $\varepsilon \leq 0$, y sólo una cantidad finita cuando $\varepsilon > 0$.
- Si ξ es algebraico de orden n, entonces, para cada $\varepsilon > 0$, la desigualdad $0 < |\xi - p/q| < 1/q^{n+\varepsilon}$ sólo tiene una cantidad finita de soluciones.

Para números algebraicos de orden $n \geq 3$, ¿existe algún $\nu(n)$ tal que la desigualdad

$$0 < \left|\xi - \frac{p}{q}\right| < \frac{1}{q^{\nu(n)+\varepsilon}}$$

tiene infinitas soluciones cuando $\varepsilon \leq 0$ y sólo una cantidad finita cuando $\varepsilon > 0$? Incluso sin aspirar a que sea $\nu(n) = n$, podríamos albergar la ilusión de que $\nu(n)$

sirviera para clasificar los números algebraicos según su orden. Es interesante hacer notar al lector que las propiedades que hemos demostrado hasta ahora no nos han resultado útiles para probar que algún número algebraico concreto tiene orden 3 o mayor (sí que hemos logrado detectar números trascendentes); ¿va a servir el orden de aproximación por racionales para identificar el orden de *algún* número algebraico no cuadrático?

En principio, para cada ξ algebraico de orden $n \geq 3$, seguro que tal ν existe, aunque podría depender de ξ. Basta observar que ν sería

$$\begin{aligned}\nu &= \sup\{\, r > 0 \,:\, 0 < |\xi - p/q| < 1/q^r \text{ tiene infinitas soluciones}\,\} \\ &= \inf\{\, r > 0 \,:\, 0 < |\xi - p/q| < 1/q^r \text{ tiene una cantidad finita de soluciones}\,\}.\end{aligned}$$

Además, sabemos que $2 \leq \nu \leq n$ (pues $r = 2$ pertenece al conjunto sobre el que estamos tomando el supremo, y cada $r = n + \varepsilon$, con $\varepsilon > 0$, al conjunto sobre el que estamos tomando el ínfimo).

Pues bien. Sin demostraciones, expongamos a continuación cuál fue el desarrollo histórico de los resultados que se iban obteniendo sobre el valor de este ν, en los que se suceden resultados debidos a Axel Thue, Carl Siegel, Aleksandr Gelfond, Freeman Dyson y Klaus Roth. En los cuatro teoremas que siguen, $\xi \in \mathbb{R}$ es un número algebraico de orden $n \geq 2$, y ν es, en cada teorema, el exponente para el que se garantizaba que $0 < |\xi - p/q| < 1/q^{\nu+\varepsilon}$ (con $\varepsilon > 0$ suficientemente pequeño) sólo podía tener una cantidad finita de soluciones. El lector puede consultar [9], donde encontrará referencias, demostraciones, e incluso generalizaciones, de lo que aquí aparece.

Teorema 4.16. Thue, 1909.
Para todo $\varepsilon > 0$, la desigualdad

$$0 < \left|\xi - \frac{p}{q}\right| < \frac{1}{q^{\varepsilon+1+n/2}}$$

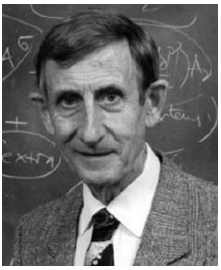
tiene sólo una cantidad finita de soluciones. Es decir, $\nu \leq 1 + n/2$.

Teorema 4.17. Siegel, 1921.
En lugar de $1 + n/2$, sirve $2\sqrt{n}$; más aún, $\nu < 2\sqrt{n}$.

Teorema 4.18. Gelfond y Dyson, 1947.
$\nu < \sqrt{2n}$.



Aleksandr Osipovich Gelfond (1906–1968)



Freeman J. Dyson (1923–)

Finalmente, los matemáticos se vieron sorprendidos cuando Roth [89] probó el siguiente resultado, en el que encuentra el valor exacto de ν (lo que hizo que Roth recibiera la Medalla Fields[3]):

Teorema 4.19. Roth, 1955.
$\nu = 2$.

Su demostración es bastante intrincada y no aparece en muchos libros. Aparte de en el artículo original antes citado, puede consultarse en [66, capítulo 6] o [75, capítulo 1, §5]. Comentemos asimismo que resulta común reunir la autoría de los resultados anteriores en lo que se denomina teorema de Thue-Siegel-Roth (que es, simplemente, el teorema 4.19).

Recapitulando, nos damos cuenta de que, pese a que hemos logrado identificarlo, ν tiene el mismo valor para todos los irracionales algebraicos, luego no sirve para caracterizar su orden; es más, *nunca* nos va a permitir detectar que algún número algebraico tiene orden ≥ 3. El orden de aproximación por racionales sólo permite distinguir los racionales de los irracionales, y —a veces— los irracionales algebraicos de los trascendentes, pero no nos da más información.

4.3.1. El correspondiente criterio de trascendencia

Quizás lo anterior nos haya hecho caer en un cierto desánimo, pues daba la impresión de que —después de tanto trabajo realizado— no era posible sacarle rentabilidad. Pero no es para tanto. El teorema de Thue-Siegel-Roth sí que proporciona, obviamente, el siguiente resultado:

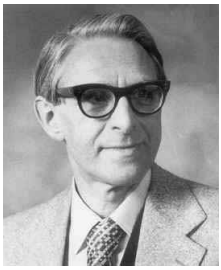
[3] Para más información sobre estos galardones, véase, en la página 113, la nota al pie numerada como 7.



Axel Thue (1863–1922)



Carl Ludwig Siegel (1896–1981)



Klaus Friedrich Roth (1925–2015)

Corolario 4.20. Criterio de trascendencia.
Supongamos que $\xi \in \mathbb{R} \setminus \mathbb{Q}$ y que existe $\varepsilon > 0$ tal que la ecuación

$$\left|\xi - \frac{p}{q}\right| < \frac{1}{q^{2+\varepsilon}}$$

tiene infinitas soluciones racionales p/q. Entonces, ξ es trascendente.

Aplicando este criterio podemos mostrar la trascendencia de algunos números expresados —como hemos hecho hasta ahora— en forma de serie.

Corolario 4.21.
Para cualesquiera números naturales $a \geq 2$ y $m \geq 3$, el número $\sum_{k=1}^{\infty} a^{-m^k}$ es trascendente.

Demostración. Denotemos $\xi = \sum_{k=1}^{\infty} a^{-m^k}$. Es claro que la expresión decimal de ξ en base a no es periódica, luego ξ no es racional. Así, para demostrar que ξ es trascendente, basta encontrar un $\varepsilon > 0$ tal que la ecuación $|\xi - p/q| < 1/q^{2+\varepsilon}$ tenga infinitas soluciones p/q.

Para cada $j = 1, 2, \ldots$, tomamos $p_j = a^{m^j}(a^{-m} + a^{-m^2} + \cdots + a^{-m^j})$, $q_j = a^{m^j}$ (ambos $p_j, q_j \in \mathbb{N}$), y el racional p_j/q_j. Así,

$$\left|\xi - \frac{p_j}{q_j}\right| = \sum_{k=1}^{\infty} a^{-m^{j+k}} = \sum_{k=1}^{\infty} \left(a^{m^j}\right)^{-m^k} = \sum_{k=1}^{\infty} (q_j)^{-m^k}$$
$$< \sum_{k=1}^{\infty} {q_j}^{-mk} = \frac{1}{q_j^m - 1} < \frac{1}{q_j^{m-1/2}},$$

donde en la última desigualdad se ha usado $q_j \geq 8$ y $m \geq 3$. Entonces, hemos encontrado un $\varepsilon = m - 5/2 > 0$ tal que todas las fracciones p_j/q_j son soluciones de $|\xi - p/q| < 1/q^{2+\varepsilon}$. □

Observemos que el criterio de trascendencia del corolario 4.13 es menos fino que el del corolario 4.20; en particular, no sirve para demostrar la trascendencia de estos números $\sum_{k=1}^{\infty} a^{-m^k}$. En cambio, sí que permitía probar la trascendencia de $\sum_{k=1}^{\infty} a^{-k!}$. En el fondo, esto es debido a que el exponente m^k crece menos rápido que el $k!$. Sin precisar demasiado, podemos decir que, cuanto más lentamente converge la serie, más difícil es probar que da lugar a un número trascendente: con $\sum_{k=1}^{\infty} a^{-k!}$, bastaba el teorema de Liouville; con $\sum_{k=1}^{\infty} a^{-m^k}$, hemos tenido que basarnos en el teorema de Thue-Siegel-Roth, un resultado mucho más complicado de establecer.

Nota 4.22. Del teorema de Thue-Siegel-Roth se puede deducir que, si $\xi \in \mathbb{R}$ es algebraico de orden ≥ 2, entonces, para cada $\varepsilon > 0$ existe un $C(\xi, \varepsilon) > 0$ tal que cualquier racional p/q cumple

$$\left|\xi - \frac{p}{q}\right| > \frac{C(\xi, \varepsilon)}{q^{2+\varepsilon}};$$

véase el ejercicio 4.15. Pero esto (al contrario de lo que ocurría con el teorema de Liouville) no tiene por qué ser verdad con $\varepsilon = 0$. En particular, para $n = 2$ el resultado de Liouville es más fuerte. ◁



Aleksandr Yakovlevich Khinchin (1894–1959)

Concluyamos preguntándonos hasta qué punto se cumple el recíproco del corolario 4.20 (¿podrá llegar a establecer una condición necesaria sobre los números trascendentes?). Ya hemos encontrado ejemplos de números trascendentes detectables con ese criterio. Pero, ¿son muy abundantes o muy escasos? Como estamos tratando con conjuntos infinitos, tenemos que aclarar a qué nos referimos con esta abundancia o escasez. En el lenguaje de la teoría de la medida, un conjunto tiene muy pocos elementos cuando su medida es 0; y es casi el total cuando es su complementario el que tiene medida nula. En estos términos, lo que realmente sucede es que en muy escasas ocasiones se pueden detectar números trascendentes por tener un orden de aproximación por racionales $1/q^{2+\varepsilon}$ con $\varepsilon > 0$, pues muy pocos números tienen ese orden de aproximación. Es más, esto ya se conocía desde tres décadas antes de que Roth estableciera su teorema, pues Aleksandr Khinchin había probado el siguiente resultado,[4] cuya demostración se puede consultar en [58, teorema 32]:

[4] Quizás aquí resulte adecuado recordar al lector que el conjunto de los números algebraicos tiene medida nula y, por tanto, casi todos los números son trascendentes. Como cualquier conjunto numerable es de medida 0, esto es consecuencia inmediata de que los números algebraicos son numerables (en este libro, véase el teorema 6.17).

Teorema 4.23. Khinchin, 1926.
Sea $\varphi\colon \mathbb{N} \to (0,\infty)$ una función monótona decreciente. Si $\sum_{q=1}^{\infty} q\varphi(q) = \infty$ entonces, para casi todos los $\xi \in \mathbb{R}$ («casi todos» significa «todos salvo un conjunto de medida nula»), la desigualdad

$$\left|\xi - \frac{p}{q}\right| < \varphi(q) \tag{4.3}$$

tiene infinitas soluciones racionales p/q. Por el contrario, si $\sum_{q=1}^{\infty} q\varphi(q) < \infty$, entonces (4.3) tiene, para casi todos los ξ, sólo un número finito de soluciones.

Sin pérdida de generalidad, podemos prescindir de lo que ocurre con $q = 1$, y analizar sólo $\sum_{q=2}^{\infty} q\varphi(q)$; esto evita molestias en la notación si en el denominador de $\varphi(q)$ queremos poner un $\log(q)$ (que, en todo caso, se podría sustituir por $\log(1+q)$). Así, los ejemplos más significativos son

$$\left|\xi - \frac{p}{q}\right| < \frac{1}{q^2 \log(q)},$$

que proporciona una serie divergente (y, por tanto, un número infinito de soluciones para casi todos los ξ), y

$$\left|\xi - \frac{p}{q}\right| < \frac{1}{q^2 \log^{1+\varepsilon}(q)}, \quad \varepsilon > 0,$$

que da lugar a una serie convergente (luego una cantidad finita de soluciones para casi todos los ξ). En el caso del criterio de trascendencia del corolario 4.20, $\varphi(q) = q^{-2-\varepsilon}$ (con $\varepsilon > 0$), y $\sum_{q=1}^{\infty} q^{-1-\varepsilon} < \infty$, así que sólo un conjunto de medida nula de números reales pueden tener orden de aproximación por racionales $q^{-2-\varepsilon}$; en particular, la mayoría de los números trascendentes no tienen aproximación por racionales de orden $1/q^{2+\varepsilon}$ para ningún $\varepsilon > 0$, luego tal criterio nunca podrá servir para detectarlos.

4.3.2. Reinterpretación en términos de la derivabilidad de una función patológica

En opinión del que escribe resulta interesante observar que el teorema de Thue-Siegel-Roth se puede reformular de manera bastante sencilla ¡en términos de la derivabilidad de una función de variable real!, un hecho que en principio puede parecer bastante sorprendente.

En concreto, dado un número real ν, usemos f_ν para denotar la función

$$f_\nu(x) = \begin{cases} 0, & \text{si } x \in \mathbb{R} \setminus \mathbb{Q}, \\ 1/q^\nu, & \text{si } x = p/q \in \mathbb{Q}, \text{ irreducible.} \end{cases}$$

Mediante aproximaciones por sucesiones es inmediato comprobar que, para $\nu \leq 0$, la función f_ν no es continua en ningún número real. En cambio, si $0 < \nu \leq 2$, f_ν es discontinua en los racionales y continua en los irracionales; este extraño comportamiento hace que la función f_1 sea —cuando se estudia el concepto de continuidad en análisis matemático— un útil y bien conocido ejemplo de función patológica cuyas propiedades no resultan nada evidentes para la intuición. Además, tampoco es difícil demostrar que, para $0 < \nu \leq 2$, f_ν no es derivable en ningún número real.

En lo que a nosotros concierne, lo más interesante surge cuando $\nu > 2$. En este caso, f_ν sigue siendo discontinua (y en consecuencia no derivable) en los racionales, y continua en los irracionales. En cuanto a la derivabilidad en los irracionales, f_ν es derivable en unos pero no en otros, sin que exista un criterio claro que permita discernir entre unos y otros; además, ambas posibilidades ocurren en una cantidad infinita de números. De hecho, la derivabilidad en cada irracional ξ está relacionado por la aproximación por racionales de ξ.

Y esto es así hasta tal punto que el teorema de Thue-Siegel-Roth es equivalente —en el sentido de que, si asumimos que uno de los dos teoremas es cierto, el otro se deduce de él con argumentos sencillos— al siguiente:

Teorema 4.24.
Sea $\nu > 2$. Si ξ es un número algebraico irracional, entonces f_ν es derivable en ξ.

Los detalles —que no son difíciles— se pueden ver en [102].

4.3.3. Soluciones de ecuaciones diofánticas

Mostremos ahora cómo el tipo de resultados que nos proporcionan el teorema de Liouville y sus refinamientos se relaciona —de una manera bastante natural— con el número de soluciones de ciertas ecuaciones diofánticas. Lo vamos a ver con un ejemplo: la ecuación diofántica

$$x^3 - 2y^3 = 1, \quad x, y \in \mathbb{Z}.$$

La única solución con $y = 0$ es $(x, y) = (1, 0)$. Por tanto, basta que analicemos las soluciones (x, y) tales que $y \neq 0$; así la ecuación es

$$\left(\frac{x}{y}\right)^3 - 2 = \frac{1}{y^3}.$$

Si tomamos $\zeta = e^{2\pi i/3} = \cos(2\pi/3) + i\,\mathrm{sen}(2\pi/3)$ (que es una raíz cúbica primitiva de la unidad), la ecuación se puede factorizar como

$$\left(\frac{x}{y} - \sqrt[3]{2}\right)\left(\frac{x}{y} - \zeta\sqrt[3]{2}\right)\left(\frac{x}{y} - \zeta^2\sqrt[3]{2}\right) = \frac{1}{y^3}.$$

En módulo, los factores segundo y tercero están acotados inferiormente por una constante positiva, que es fácil de determinar (nótese que $\zeta, \zeta^2 \in \mathbb{C} \setminus \mathbb{R}$). Por tanto, todas las soluciones (x, y) de la ecuación diofántica deberán cumplir

$$\left|\frac{x}{y} - \sqrt[3]{2}\right| \leq \frac{C}{|y|^3} \tag{4.4}$$

para alguna constante C independiente de x e y. Pero el teorema de Thue-Siegel-Roth nos asegura que sólo hay una cantidad finita de soluciones x/y de la desigualdad (4.4) (en realidad, como $\sqrt[3]{2}$ es un número algebraico de orden 3, para llegar a la misma conclusión basta usar el teorema de Thue, cuya demostración es más sencilla). Por tanto, la ecuación $x^3 - 2y^3 = 1$ sólo puede tener un número finito de soluciones enteras.



Alan Baker
(1939–2018)

Pero, con esto, no tenemos ninguna estimación de cuántas soluciones puede haber. Para acotar su número necesitaríamos conocer constantes concretas en el teorema de Thue-Siegel-Roth (en particular, del $C(\xi, \varepsilon)$ de la nota 4.22). Algunos resultados hay sobre esto, en gran parte debidos a Alan Baker, quien también —como Roth— recibió la Medalla Fields.[5] Por ejemplo, Baker logró dar una constante concreta para $\sqrt[3]{2}$.

Concluyamos la sección mencionando que identificar las constantes no es el único tema de investigación actual relacionado con la aproximación diofántica. Otro foco de interés —que mejoraría los resultados conocidos— es ver si (para $n \geq 3$) el $q^{2+\varepsilon}$ que nos ha aparecido reiteradamente se puede cambiar por $q^2\psi(q)$, con $\psi(q)$ una función de crecimiento más lento que

[5] Para conocer algo más del trabajo que le hizo merecedor del premio, véase la sección 6.8.

cualquier función potencial. En particular, se piensa que, para cualquier $\xi \in \mathbb{R}$ algebraico de orden ≥ 3, existirá un ε_0 de modo que, para cada $\varepsilon > \varepsilon_0$,

$$\left|\xi - \frac{p}{q}\right| > \frac{1}{q^2 \log^{\varepsilon}(q)}$$

para todo número racional p/q (con $q > 1$). Como es habitual, esto implicaría que ξ no tiene aproximación por racionales de orden $\varphi(q) = 1/(q^2 \log^{\varepsilon}(q))$ para ningún $\varepsilon > \varepsilon_0$. Pero todo esto es, sólo, una conjetura, y permanece sin demostrar ni refutar.

4.4. Fracciones de Brocot

Como consecuencia del teorema de Dirichlet, en el corolario 4.5 hemos probado que, cuando $C = 1$, para cada $\xi \in \mathbb{R} \setminus \mathbb{Q}$ existen infinitas fracciones $\frac{p}{q}$ tales que

$$0 < \left|\xi - \frac{p}{q}\right| < \frac{C}{q^2} \tag{4.5}$$

(recordemos además, que, por el teorema 4.19, esto no es cierto para ningún C si, en la desigualdad (4.5), sustituimos q^2 por q^ν con $\nu > 2$). Si, en lugar de tomar $C = 1$, elegimos un valor de C menor, ¿seguirá teniendo (4.5) infinitas soluciones p/q para cada $\xi \in \mathbb{R} \setminus \mathbb{Q}$?, ¿hasta qué límite podemos disminuir la constante C? Más aún, ¿cómo buscar las fracciones p/q?

Existen diversos métodos de abordar estas cuestiones. Aquí nos decantaremos por uno de los que consideramos más sencillos, que se basa en el uso de fracciones de Brocot. En realidad, tales fracciones fueron introducidas, independientemente, por el relojero francés Achille Brocot (1860) y el matemático alemán Moritz Stern (1858), por lo que también son conocidas por el nombre de Stern-Brocot. Las fracciones de Brocot son parecidas a las de Farey, que fueron definidas en 1816 por el geólogo inglés John Farey (obsérvese la curiosa coincidencia: ni Farey ni Brocot eran matemáticos). Las fracciones de Farey son mucho más habituales en los libros (véase, por ejemplo [78, 54]) pero, en algunas cuestiones, las de Brocot son más fáciles de usar. En particular, nosotros vamos a emplear las de Brocot. No podemos dejar de citar que tienen múltiples propiedades, y que resultan de utilidad en muy distintos campos; el lector que quiera profundizar en su estudio puede consultar [47] (donde se las denomina árboles de Stern-Brocot) o [14] (donde se alude a ellas con el nombre de sucesiones de Brocot). Pero aquí sólo veremos su definición y unas

pocas propiedades; prácticamente, las que vayamos a usar para aplicarlas a nuestro problema, que, finalmente, resolveremos en la sección 4.5.

Dadas dos fracciones $\frac{a}{b}$ y $\frac{c}{d}$, decimos que su *mediación* es $\frac{a+c}{b+d}$. En este proceso, permitimos que una de las fracciones sea $\frac{1}{0}$ pero, como hasta ahora, el denominador no puede ser negativo. Es fácil comprobar que, al efectuar la mediación de dos fracciones, se consigue una fracción intermedia (por ejemplo, obsérvese que $a/b < (a+c)/(b+d) \Leftrightarrow ab+ad < ab+cb \Leftrightarrow ad < cb \Leftrightarrow a/b < c/d$). Es decir,

$$\frac{a}{b} < \frac{c}{d} \;\Rightarrow\; \frac{a}{b} < \frac{a+c}{b+d} < \frac{c}{d}.$$

Para cada $n = 0, 1, 2, \ldots$, denotaremos por $\mathscr{B}_n$ al conjunto (ordenado) formado por las fracciones de Brocot de orden n, que se construyen según el proceso que indicamos a continuación; en cada $\mathscr{B}_n$, las fracciones están ordenadas en orden creciente. En primer lugar, $\mathscr{B}_0$ consta de las dos fracciones $\frac{0}{1}$ y $\frac{1}{0}$. Para construir $\mathscr{B}_1$ tomamos las dos fracciones de $\mathscr{B}_0$ y su mediación, que colocamos entre ellas; así, $\mathscr{B}_1$ tiene las fracciones $\frac{0}{1}$, $\frac{1}{1}$ y $\frac{1}{0}$. En general, dado $\mathscr{B}_n$, $\mathscr{B}_{n+1}$ consta de todas las fracciones de $\mathscr{B}_n$ y, además, de sus mediaciones (obsérvese que, al hacer mediación y colocar la nueva fracción entra las dos que la han generado, $\mathscr{B}_{n+1}$ se crea ya ordenado). En el cuadro 4.1 se muestran las primeras fracciones de Brocot; resulta útil representarlas en forma de árbol, pues esto ilustra cómo se han ido generando mediante el proceso de mediación.[6]

En lo que sigue, muchas veces aludiremos a fracciones de Brocot consecutivas. Con esto nos referiremos siempre a dos (o más) fracciones $\frac{a}{b}$ y $\frac{c}{d}$ consecutivas en un mismo $\mathscr{B}_n$ y con $\frac{a}{b} < \frac{c}{d}$ (es decir, que, aunque no lo digamos expresamente, las supondremos en orden creciente). Normalmente, no especificaremos de qué $\mathscr{B}_n$ se trata (a veces, n deberá ser lo bastante grande como para que $\mathscr{B}_n$ tenga el número de fracciones que se le exige). Por otra parte, si alguna vez tenemos problemas con la fracción $\frac{1}{0}$, que es la última de cada $\mathscr{B}_n$, podemos prescindir de ella; para conservar la simetría, eliminaríamos también las $\frac{0}{1}$. De este modo, la segunda y penúltima fracción de cada $\mathscr{B}_n$ (que son $\frac{1}{n}$ y $\frac{n}{1}$) asumen ahora el papel de primera y última.

[6] Por completitud, digamos que, dado $n \in \mathbb{N}$, las fracciones de Farey de orden n son las formadas por la fracciones irreducibles entre 0 y 1 de denominador menor o igual que n, ordenadas de menor a mayor. Por ejemplo, las de orden 4 son $\mathscr{F}_4 = \left\{\frac{0}{1}, \frac{1}{4}, \frac{1}{3}, \frac{1}{2}, \frac{2}{3}, \frac{3}{4}, \frac{1}{1}\right\}$. Cuando Farey las introdujo en 1816, lo hizo comentando que había observado una «curiosa propiedad» que no había sabido demostrar: si $\frac{a}{b} < \frac{a'}{b'} < \frac{a''}{b''}$ son tres fracciones consecutivas de $\mathscr{F}_n$, entonces $\frac{a'}{b'} = \frac{a+a''}{b+b''}$. Fíjese el lector que esto, para fracciones de Brocot, es prácticamente la definición; para fracciones de Farey hay que probarlo.

$\mathscr{B}_0$:	$\frac{0}{1}$															$\frac{1}{0}$
$\mathscr{B}_1$:	$\frac{0}{1}$							$\frac{1}{1}$								$\frac{1}{0}$
$\mathscr{B}_2$:	$\frac{0}{1}$				$\frac{1}{2}$			$\frac{1}{1}$				$\frac{2}{1}$				$\frac{1}{0}$
$\mathscr{B}_3$:	$\frac{0}{1}$		$\frac{1}{3}$		$\frac{1}{2}$		$\frac{2}{3}$	$\frac{1}{1}$		$\frac{3}{2}$		$\frac{2}{1}$		$\frac{3}{1}$		$\frac{1}{0}$
$\mathscr{B}_4$:	$\frac{0}{1}$	$\frac{1}{4}$	$\frac{1}{3}$	$\frac{2}{5}$	$\frac{1}{2}$	$\frac{3}{5}$	$\frac{2}{3}$	$\frac{3}{4}$	$\frac{1}{1}$	$\frac{4}{3}$	$\frac{3}{2}$	$\frac{5}{3}$	$\frac{2}{1}$	$\frac{5}{2}$	$\frac{3}{1}$	$\frac{4}{1}$	$\frac{1}{0}$

Cuadro 4.1: Fracciones de Brocot de orden ≤ 4.

Veamos ahora una sencilla propiedad, pero que resulta básica en todo el tratamiento de fracciones de Brocot:[7]

Proposición 4.25.

Dadas dos fracciones de Brocot $\frac{a}{b}$ y $\frac{c}{d}$ consecutivas, siempre se cumple

$$cb - ad = 1. \tag{4.6}$$

Como consecuencia,

$$\frac{c}{d} - \frac{a}{b} = \frac{1}{bd}.$$

Demostración. Probémoslo por inducción. Obviamente, el resultado es verdad para fracciones de $\mathscr{B}_1$ (y también de $\mathscr{B}_2$, si así lo deseamos, para no tener que considerar la fracción $\frac{1}{0}$). Suponiéndolo cierto en $\mathscr{B}_{n-1}$, veamos que también lo es en $\mathscr{B}_n$. Por construcción, dos fracciones consecutivas de $\mathscr{B}_n$ serán de la forma

$$\frac{a}{b} \quad \text{y} \quad \frac{a+c}{b+d}$$

o

$$\frac{a+c}{b+d} \quad \text{y} \quad \frac{c}{d},$$

[7] Empalmando con los comentarios de la anterior nota al pie, es interesante añadir que lo que ahora vamos a exponer también es cierto para fracciones de Farey, pero más complicado de demostrar. De hecho, en lo que a nosotros respecta, ésta es la principal razón por la que el uso de fracciones de Brocot lleva más rápidamente a los resultados buscados. Por el contrario, que cada racional entre 0 y 1 es una fracción de Farey es directo; comprobar que también es una fracción de Brocot requiere algo de trabajo, pero esta propiedad no la vamos a usar en las argumentaciones de este tema. No obstante, proponemos al lector que intente probarlo (véase el ejercicio 4.21).

con $\frac{a}{b}$ y $\frac{c}{d}$ dos consecutivas de $\mathscr{B}_{n-1}$. Por hipótesis de inducción, sabemos que $cb - ad = 1$. En consecuencia,

$$(a+c)b - a(b+d) = cb - ad = 1$$

y también

$$c(b+d) - (a+c)d = cb - ad = 1,$$

tal como queríamos comprobar. □

En particular, esta propiedad permite probar que, en el proceso de construir las fracciones de Brocot, todas las fracciones que van saliendo son irreducibles. En efecto:

Proposición 4.26.

Tal como las hemos construido, las fracciones de Brocot aparecen siempre en forma irreducible.

Demostración. Por inducción sobre n, probemos que esto es cierto para todas la fracciones de $\mathscr{B}_n$ Para $n = 1$, es claro que $\frac{0}{1}$ y $\frac{1}{1}$ son irreducibles. Supongámoslo cierto para $n-1$, y comprobémoslo para n. Por construcción, las fracciones de $\mathscr{B}_n$ son las de $\mathscr{B}_{n-1}$ (y por tanto son irreducibles), o son mediación $\frac{a+c}{b+d}$ de dos fracciones consecutivas $\frac{a}{b}$ y $\frac{c}{d}$ de $\mathscr{B}_{n-1}$; éste es el caso que debemos analizar. Sea $m = \mathrm{mcd}(a+c, b+d)$; comprobemos que $m = 1$. Por una parte, por la proposición 4.25, tenemos

$$\frac{a+c}{b+d} - \frac{a}{b} = \frac{cb-ad}{(b+d)b} = \frac{1}{(b+d)b}.$$

Por otra parte, descomponiendo $a + c = m\alpha$ y $b + d = m\beta$, queda

$$\frac{a+c}{b+d} - \frac{a}{b} = \frac{m\alpha}{m\beta} - \frac{a}{b} = \frac{m\alpha b - m\beta a}{m\beta b} = \frac{(\alpha b - \beta a)m}{(b+d)b}.$$

Así, $1 = (\alpha b - \beta a)m$ (y todos los números son enteros), luego $m = 1$. □

Nota 4.27. A veces resulta cómodo quedarse únicamente con las fracciones de Brocot que pertenecen al intervalo $[0, 1]$. Así aparece lo que se suele denominar fracciones de Brocot reducidas, que representaremos mediante $\mathscr{B}_n^*$. Obviamente, prescindimos de $\mathscr{B}_0^*$. Fijándonos en el cuadro 4.1, cada $\mathscr{B}_n^*$ es la mitad izquierda de $\mathscr{B}_n$, incluida la fracción central $\frac{1}{1}$. Es claro que las fracciones de cada $\mathscr{B}_n^*$ se obtienen a partir de las de $\mathscr{B}_{n-1}^*$ mediante mediación (al igual que ocurría, por definición, con $\mathscr{B}_n$ y $\mathscr{B}_{n-1}$). En los ejercicios se propone demostrar algunas propiedades de las fracciones de Brocot reducidas. En particular, el ejercicio 4.18 muestra que, según

va creciendo n, las fracciones de cada $\mathscr{B}_n^*$ están «cada vez más juntas»; o, dicho de otra manera, entre ellas no quedan huecos «grandes», en un sentido que se precisa en el ejercicio. Obviamente, esto no ocurre en la parte derecha de los $\mathscr{B}_n$. ◁

4.4.1. Aproximaciones óptimas

Vamos ahora a ver que las fracciones de Brocot sirven para conseguir buenas aproximaciones con denominadores pequeños. En primer lugar, precisemos los conceptos (además, recordemos que los denominadores los suponemos siempre positivos sin necesidad de indicarlo expresamente). Dado $\xi \in \mathbb{R}$, decimos que una fracción irreducible $p/q \in \mathbb{Q}$ es una *aproximación óptima de primera especie* de ξ si, para cualquier otra fracción $h/k \in \mathbb{Q}$ con $k \leq q$, se verifica

$$\left|\xi - \frac{h}{k}\right| > \left|\xi - \frac{p}{q}\right|.$$

Lema 4.28.

Sean dos fracciones a/b y c/d con $a/b < c/d$ y $cb - ad = 1$. Si una fracción irreducible h/k verifica $a/b < h/k < c/d$, entonces $k \geq b + d$. Además, la igualdad $k = b + d$ sólo se da cuando $h = a + c$.

Demostración. Existen tres posibilidades:

$$\frac{h}{k} = \frac{a+c}{b+d}, \quad \frac{a}{b} < \frac{h}{k} < \frac{a+c}{b+d} \quad \text{o} \quad \frac{a+c}{b+d} < \frac{h}{k} < \frac{c}{d}.$$

En el primer caso, el lema está probado. Analicemos únicamente el segundo, pues el tercero es similar. Como $h/k > a/b$, deberá cumplirse

$$\frac{h}{k} - \frac{a}{b} = \frac{hb - ak}{bk} \geq \frac{1}{bk}$$

(pues $hb - ak$ tiene que ser un entero positivo). Además, de $h/k < (a + c)/(b + d)$ y $cb - ad = 1$, se deduce que

$$\frac{h}{k} - \frac{a}{b} < \frac{a+c}{b+d} - \frac{a}{b} = \frac{(a+c)b - a(b+d)}{(b+d)b} = \frac{1}{(b+d)b}.$$

Por consiguiente,

$$\frac{1}{bk} < \frac{1}{(b+d)b},$$

luego $k > b + d$. □

Con esto, ya podemos probar el siguiente resultado:

Teorema 4.29.
Sea un número real $\xi > 0$ y un entero $n > \xi$. Entonces, la fracción de $\mathscr{B}_n$ más cercana a ξ es una aproximación óptima de primera especie de ξ. Con una salvedad: si ξ equidista de dos fracciones de $\mathscr{B}_n$, la aproximación óptima de primera especie es la de menor denominador; y si las dos tienen el mismo denominador, forzosamente nos encontramos ante el caso trivial $\xi = n - 1/2$ con $\frac{n-1}{1} < \xi < \frac{n}{1}$.

Demostración. Si ξ es una fracción de $\mathscr{B}_n$, el resultado es obvio. Si no, ξ estará comprendido entre dos fracciones consecutivas a/b y c/d de $\mathscr{B}_n$. Cualquier fracción h/k más cercana a ξ que a/b y c/d deberá cumplir $a/b < h/k < c/d$. En estas condiciones, el lema anterior nos asegura que $k \geq b + d > \max\{b, d\} \geq \min\{b, d\}$, lo que prueba el teorema. □

Nota 4.30. El nombre «aproximación óptima de primera especie» sugiere que tiene que haber otro tipo de «aproximaciones óptimas». En efecto; decimos que una fracción irreducible $p/q \in \mathbb{Q}$ es una *aproximación óptima de segunda especie* de $\xi \in \mathbb{R}$ si, para cualquier otra fracción $h/k \in \mathbb{Q}$ con $k \leq q$, se verifica $|k\xi - h| > |q\xi - p|$. Toda aproximación óptima de segunda especie lo es también de primera. En efecto, si p/q es aproximación óptima de segunda especie y $0 < k \leq q$, se tiene

$$\left|\xi - \frac{h}{k}\right| = \frac{1}{k}\,|k\xi - h| > \frac{1}{k}\,|q\xi - p| \geq \frac{1}{q}\,|q\xi - p| = \left|\xi - \frac{p}{q}\right|,$$

luego p/q es aproximación óptima de primera especie. Con respecto a la implicación contraria, véase el ejercicio 4.22. En este tema no nos aparecerán este tipo de aproximaciones, pero sí que nos surgirán en el tema siguiente: lo mismo que las fracciones de Brocot permiten obtener aproximaciones de primera especie, las fracciones continuas nos proporcionarán aproximaciones de segunda especie (esto lo veremos en el apartado (d) del teorema 5.12). ◁

4.5. Teorema de Hurwitz

El objetivo de esta sección es mejorar, todo lo posible, el corolario 4.5. Con este afán, veamos un primer resultado sobre aproximación de un número irracional mediante racionales; con él, tenemos recorrido casi todo el camino para llegar al teorema 4.33, en el que mostraremos la primera mejora. En concreto, probemos ahora que, mediante fracciones de Brocot, podemos conseguir lo siguiente:

Proposición 4.31.

Sea $\xi \in \mathbb{R} \setminus \mathbb{Q}$, $\xi > 0$. Para cada entero positivo $n > \xi$, existe una fracción $\frac{p}{q} \in \mathscr{B}_n$ tal que

$$0 < \left| \xi - \frac{p}{q} \right| < \frac{1}{2q^2}. \tag{4.7}$$

Además, este $\frac{p}{q}$ se obtiene como sigue: ξ estará entre dos términos consecutivos $\frac{a}{b}$ y $\frac{c}{d}$ de $\mathscr{B}_n$; pues bien, una de estas dos fracciones satisface la desigualdad requerida.

Demostración. En primer lugar, obsérvese que, como $n > \xi > 0$, el par de fracciones de Brocot consecutivas a ξ no será $\frac{n}{1} < \xi < \frac{1}{0}$ en ningún $\mathscr{B}_n$. Y, en segundo lugar, que, si ξ cumple $\frac{n-1}{1} < \xi < \frac{n}{1}$ (del final de algún $\mathscr{B}_n$), forzosamente deberá ser

$$0 < \left| \xi - \frac{n-1}{1} \right| < \frac{1}{2 \cdot 1^2} \qquad \text{o} \qquad 0 < \left| \xi - \frac{n}{1} \right| < \frac{1}{2 \cdot 1^2},$$

y, por tanto, una de las dos fracciones satisface (4.7).

Una vez vistos los casos triviales, sea, pues, $\frac{a}{b} < \xi < \frac{c}{d}$, con $\frac{a}{b}, \frac{c}{d} \in \mathscr{B}_n$ consecutivas, asumiendo desde ahora que no nos encontramos ante el caso $\frac{n-1}{1} < \xi < \frac{n}{1}$ ya analizado. Supongamos que ninguna de las dos fracciones $\frac{a}{b}$ o $\frac{c}{d}$ cumple (4.7). Entonces, tendríamos

$$\frac{1}{2b^2} \leq \xi - \frac{a}{b} \quad \text{y} \quad \frac{1}{2d^2} \leq \frac{c}{d} - \xi.$$

Sumándolas,

$$\frac{1}{2b^2} + \frac{1}{2d^2} \leq \frac{c}{d} - \frac{a}{b} = \frac{bc - ad}{bd} = \frac{1}{bd},$$

donde en la última igualdad hemos usado que las fracciones son consecutivas, y por tanto cumplen (4.6). Pero $\frac{1}{2b^2} + \frac{1}{2d^2} \leq \frac{1}{bd}$ se puede poner como

$$\frac{d^2 + b^2 - 2bd}{2b^2d^2} = \frac{(b-d)^2}{2b^2d^2} \leq 0,$$

lo cual es absurdo pues $b \neq d$ (por construcción de las fracciones de Brocot, es evidente que los denominadores de dos fracciones consecutivas deben ser diferentes, salvo que se trate de $\frac{n-1}{1}$ y $\frac{n}{1}$ en $\mathscr{B}_n$, caso ya descartado). Así pues, o bien $\frac{a}{b}$, o bien $\frac{c}{d}$ (quizás los dos), satisfacen la desigualdad (4.7). □

Demostremos, además, un lema. Con el único fin de simplificar su enunciado, emplearemos el siguiente lenguaje: dado ξ un irracional positivo, a las dos fracciones consecutivas $\frac{a}{b}$, $\frac{c}{d}$ de $\mathscr{B}_n$ tales que $\frac{a}{b} < \xi < \frac{c}{d}$ las denominaremos «fracciones contiguas a ξ en $\mathscr{B}_n$».

Lema 4.32.
Dado $\xi \in \mathbb{R} \setminus \mathbb{Q}$, $\xi > 0$, y $n > 0$, sean $\frac{a}{b}$, $\frac{c}{d}$ las fracciones contiguas a ξ en $\mathscr{B}_n$. Entonces, existe $m > n$ para el cual las fracciones $\frac{a'}{b'}$, $\frac{c'}{d'}$ contiguas a ξ en $\mathscr{B}_m$ satisfacen $\frac{a'}{b'} \neq \frac{a}{b}$ y $\frac{c'}{d'} \neq \frac{c}{d}$.

Demostración. Analicémoslo con $\frac{a}{b}$ y $\frac{a'}{b'}$; para $\frac{c}{d}$ y $\frac{c'}{d'}$ sería análogo. Denotemos mediante δ a la distancia entre $\frac{a}{b}$ y ξ, es decir $\delta = \xi - \frac{a}{b}$. Por construcción de las sucesiones de Brocot, en $\mathscr{B}_{n+1}$, la fracción contigua a $\frac{a}{b}$ por la derecha es $\frac{a+c}{b+d}$. En $\mathscr{B}_{n+2}$ es $\frac{2a+c}{2b+d}$. En general, en $\mathscr{B}_{n+k}$ es $\frac{ka+c}{kb+d}$. Según la proposición 4.25, la distancia entre $\frac{ka+c}{kb+d}$ y $\frac{a}{b}$ es $\frac{ka+c}{kb+d} - \frac{a}{b} = \frac{1}{(kb+d)b}$. Es claro que existe un k tal que $\frac{1}{(kb+d)b} < \delta$. Por tanto,

$$\frac{a}{b} < \frac{ka+c}{kb+d} < \xi,$$

luego podemos tomar $\frac{a'}{b'} = \frac{ka+c}{kb+d}$. □

Con esto, ya estamos en condiciones de abordar el primer teorema que responde a una de las preguntas planteadas sobre la constante C de (4.5): se puede tomar $C = 1/2$, mejorándose de este modo el resultado del corolario 4.5.

Teorema 4.33.
Si $\xi \in \mathbb{R} \setminus \mathbb{Q}$, existen infinitas fracciones $\frac{p}{q}$ tales que

$$0 < \left| \xi - \frac{p}{q} \right| < \frac{1}{2q^2}. \tag{4.8}$$

Demostración. En primer lugar, es claro que nos podemos restringir a $\xi \in (0,1)$ (luego, en particular, a $\xi > 0$, tal como exige la proposición anterior). En efecto, si quisiéramos aproximar $\theta \notin (0,1)$, el asunto se traslada a $(0,1)$ tomando $\theta = \xi + k$, con $k = \lfloor \theta \rfloor \in \mathbb{Z}$ y $\xi \in (0,1)$, ya que

$$\left| \theta - \frac{p+kq}{q} \right| = \left| \xi - \frac{p}{q} \right|.$$

De acuerdo a la proposición 4.31, para cada $n = 1, 2, \ldots$ existe una fracción de $\mathscr{B}_n$ que satisface (4.8). Debemos probar que infinitas de ellas son distintas. Esto fallaría únicamente en el caso de que hubiera una fracción $\frac{a}{b}$ (o $\frac{c}{d}$) tal que, para todo m a partir de un cierto n, $\frac{a}{b}$ forma parte del par de fracciones contiguas a ξ en todos los $\mathscr{B}_m$. Pero, por el lema anterior, esto no es posible. □

Como podemos observar, la demostración anterior proporciona —teniendo en cuenta la proposición 4.31 y el lema 4.32— un método constructivo para obtener infinitas fracciones que satisfacen (4.8).

Veamos ahora que la constante $C = 1/2$ para (4.5) aún se puede mejorar. En lo que sigue, la vamos a rebajar hasta $C = 1/\sqrt{5}$. Como antes, la demostración que vamos a efectuar de este hecho es constructiva, pues nos proporcionará un método para encontrar infinitas fracciones p/q tales que $|\xi - p/q| < 1/(\sqrt{5}\,q^2)$.

Proposición 4.34.

Sea $\xi \in \mathbb{R} \setminus \mathbb{Q}$, $\xi > 0$. Para cada entero positivo $n > \xi$, sean dos términos consecutivos $\frac{a}{b}$ y $\frac{c}{d}$ de $\mathscr{B}_n$ tal que $\frac{a}{b} < \xi < \frac{c}{d}$. Entonces, la ecuación

$$0 < \left|\xi - \frac{p}{q}\right| < \frac{1}{\sqrt{5}\,q^2}. \tag{4.9}$$

se satisface si tomamos como $\frac{p}{q}$ alguna de estas tres fracciones: $\frac{a}{b}$, $\frac{c}{d}$ o $\frac{a+c}{b+d}$. (Nótese, además, que la fracción $\frac{a+c}{b+d}$ está localizada: $\frac{a+c}{b+d} \in \mathscr{B}_{n+1}$; en realidad, las tres fracciones están en $\mathscr{B}_{n+1}$.)

Demostración. Partimos de $\frac{a}{b}, \frac{c}{d} \in \mathscr{B}_n$ consecutivas tal que $\frac{a}{b} < \xi < \frac{c}{d}$. Distinguiremos dos casos, según sea $\xi < \frac{a+c}{b+d}$ o $\frac{a+c}{b+d} < \xi$; analizaremos únicamente el primero, pues el otro es similar.

Supongamos que ninguna de las tres fracciones $\frac{a}{b}$, $\frac{c}{d}$ y $\frac{a+c}{b+d}$ satisface la desigualdad (4.9). Entonces,

$$\xi - \frac{a}{b} \geq \frac{1}{\sqrt{5}\,b^2}, \quad \frac{a+c}{b+d} - \xi \geq \frac{1}{\sqrt{5}(b+d)^2} \quad \text{y} \quad \frac{c}{d} - \xi \geq \frac{1}{\sqrt{5}\,d^2}. \tag{4.10}$$

Recordemos además que $bc - ad = 1$, pues $\frac{a}{b}$ y $\frac{c}{d}$ son fracciones consecutivas. Sumando las dos primeras desigualdades de (4.10), queda

$$\frac{1}{\sqrt{5}\,b^2} + \frac{1}{\sqrt{5}\,(b+d)^2} \leq \frac{a+c}{b+d} - \frac{a}{b} = \frac{(a+c)b - (b+d)a}{(b+d)b} = \frac{1}{(b+d)b};$$

esto se puede expresar como

$$(b+d)^2 + b^2 \leq \sqrt{5}\,b(b+d). \tag{4.11}$$

Sumando la primera y la tercera desigualdades de (4.10),

$$\frac{1}{\sqrt{5}\,b^2} + \frac{1}{\sqrt{5}\,d^2} \leq \frac{c}{d} - \frac{a}{b} = \frac{1}{bd},$$

de donde

$$b^2 + d^2 \leq \sqrt{5}\, bd. \tag{4.12}$$

Ahora, sumando (4.11) y (4.12) obtenemos

$$3b^2 + 2bd + 2d^2 \leq \sqrt{5}\, b^2 + 2\sqrt{5}\, bd.$$

Esta expresión, multiplicándola por 2 y pasando todo al lado izquierdo, podemos escribirla como

$$\Big((1-\sqrt{5})b + 2d\Big)^2 = (6 - 2\sqrt{5})b^2 + 4(1-\sqrt{5})bd + 4d^2 \leq 0.$$

Pero esto es imposible, puesto que $\sqrt{5}$ es irracional luego no puede ser $(1-\sqrt{5})b + 2d = 0$. □

Con esto llegamos al resultado principal de esta sección, que fue probado por el matemático alemán Adolf Hurwitz en 1891:

Teorema 4.35. Hurwitz, 1891.
Si $\xi \in \mathbb{R} \setminus \mathbb{Q}$, existen infinitas fracciones $\frac{p}{q}$ tales que

$$0 < \left|\xi - \frac{p}{q}\right| < \frac{1}{\sqrt{5}\, q^2}. \tag{4.13}$$

Demostración. Ahora, nos apoyamos en la proposición 4.34 igual que, al demostrar el teorema 4.33, nos apoyábamos en la proposición 4.31.

Para garantizar que, de ese modo, aparecen infinitas fracciones distintas, podemos usar el lema 4.32 dos veces. Así, a partir de $\frac{a}{b} < \xi < \frac{c}{d}$, con $\frac{a}{b}, \frac{c}{d} \in \mathscr{B}_n$ consecutivas, garantizamos la existencia de $\frac{a'}{b'}, \frac{c'}{d'} \in \mathscr{B}_m$ consecutivas (con $m > n$) y $\frac{a''}{b''}, \frac{c''}{d''} \in \mathscr{B}_l$ también consecutivas (ahora, $l > m$) que satisfacen

$$\frac{a}{b} < \frac{a'}{b'} < \frac{a''}{b''} < \xi < \frac{c''}{d''} < \frac{c'}{d'} < \frac{c}{d}.$$

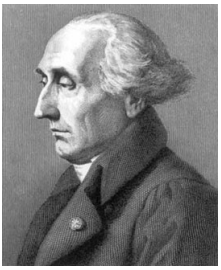
El $\frac{a+c}{b+d}$ que la proposición 4.34 propone como posible valor para $\frac{p}{q}$ cumplirá

$$\frac{a}{b} < \frac{a+c}{b+d} \leq \frac{a'}{b'} < \frac{a''}{b''} \quad \text{o} \quad \frac{c''}{d''} < \frac{c'}{d'} \leq \frac{a+c}{b+d} < \frac{c}{d}.$$

Por tanto, $\frac{a+c}{b+d}$ puede coincidir con $\frac{a'}{b'}$ o con $\frac{c'}{d'}$, pero no con $\frac{a''}{b''}$ o con $\frac{c''}{d''}$; ni tampoco con $\frac{a''+c''}{b''+d''}$, que es una fracción comprendida entre $\frac{a''}{b''}$ y $\frac{c''}{d''}$. □



Adolf Hurwitz
(1859–1919)



Joseph Louis Lagrange
(1736–1813)

Nota 4.36. La demostraciones que hemos efectuado en esta sección son similares a las basadas en fracciones de Farey que podemos encontrar en [78, § 6.2] o [54, § 6.10]. Nosotros seguimos el espíritu de [15] que, sirviéndose de las fracciones de Brocot, simplifica la correspondiente prueba del teorema de Hurwitz. Por otra parte, también son comunes las demostraciones basadas en fracciones continuas, como la de [54, § 10.4]. ◁

4.5.1. La mejor constante

Obviamente, ahora nos surge la cuestión de si la constante $C = 1/\sqrt{5}$ se puede mejorar. Es decir, si existe otra C menor de modo que, para cada ξ irracional, la inecuación $|\xi - p/q| < C/q^2$ tiene infinitas soluciones racionales p/q. El siguiente teorema (obsérvese que, por comodidad, usaremos $M = 1/C$) demuestra, entre otras cosas, que esto no es posible:

Teorema 4.37.
Sea $\phi = \frac{1+\sqrt{5}}{2}$, número que se conoce con el nombre de razón áurea,[8] *y sea M un número real positivo. Entonces:*

(a) Si $M > \sqrt{5}$, sólo puede existir una cantidad finita de racionales p/q tales que

$$\left|\phi - \frac{p}{q}\right| < \frac{1}{Mq^2}. \tag{4.14}$$

(b) Si $M \geq \frac{3+\sqrt{5}}{2}$, no hay ninguna solución racional de (4.14).

[8] Esta constante nos aparecerá de nuevo en el tema 5; véase, en la página 285, la nota al pie numerada como 5.

DEMOSTRACIÓN. Es fácil comprobar que $\left(x-\frac{1+\sqrt{5}}{2}\right)\left(x-\frac{1-\sqrt{5}}{2}\right)=x^2-x-1$. En particular, este polinomio no se anula para ningún racional p/q. Entonces,

$$0\neq\left|\frac{p}{q}-\frac{1+\sqrt{5}}{2}\right|\cdot\left|\frac{p}{q}-\frac{1+\sqrt{5}}{2}+\sqrt{5}\right|=\left|\frac{p^2}{q^2}-\frac{p}{q}-1\right|=\frac{|p^2-pq-q^2|}{q^2}.$$

Como $|p^2-pq-q^2|$ es un entero no negativo, deberá ser $|p^2-pq-q^2|\geq 1$. Es decir,

$$\left|\frac{p}{q}-\frac{1+\sqrt{5}}{2}\right|\cdot\left|\frac{p}{q}-\frac{1+\sqrt{5}}{2}+\sqrt{5}\right|\geq\frac{1}{q^2}.$$

Aplicando la desigualdad triangular, tendremos

$$\frac{1}{q^2}\leq\left|\frac{p}{q}-\frac{1+\sqrt{5}}{2}\right|\cdot\left(\left|\frac{p}{q}-\frac{1+\sqrt{5}}{2}\right|+\sqrt{5}\right).$$

Si p/q es solución de (4.14), de aquí se deduce que

$$\frac{1}{q^2}<\frac{1}{Mq^2}\left(\frac{1}{Mq^2}+\sqrt{5}\right)$$

y, por tanto,

$$M<\frac{1}{Mq^2}+\sqrt{5}. \tag{4.15}$$

Supongamos ahora que M es tal que (4.14) tiene infinitas soluciones p/q. Desde luego, para cada q sólo puede haber una cantidad finita de valores de p capaces de satisfacer (4.14). Así que, forzosamente, deberá haber infinitos denominadores distintos. De este modo, podemos garantizar que existe una sucesión de soluciones $\{p_j/q_j\}_{j=1}^{\infty}$ tal que $q_j\to\infty$ cuando $j\to\infty$. Sustituyendo en (4.15), tendremos que

$$M<\frac{1}{Mq_j^2}+\sqrt{5}$$

para cada j. Tomando límites cuando $j\to\infty$ llegamos a $M\leq\sqrt{5}$, con lo que queda demostrada la parte (a).

Analicemos ahora la parte (b) en la que afirmamos que, si M es demasiado grande, no sólo es que no hay infinitas soluciones p/q de (4.14), sino que no hay ninguna. En primer lugar, (4.15) implica que, si existe alguna solución, M deberá cumplir $M<1+\sqrt{5}$. Pero, con un poco más de cuidado, vamos a mejorar este

resultado. Si existe alguna solución, (4.15) impone la condición $M < 1/M+\sqrt{5}$. Pero

$$\begin{aligned} M < 1/M + \sqrt{5} &\Leftrightarrow M^2 < 1 + \sqrt{5}\, M \Leftrightarrow (M - \sqrt{5}/2)^2 < 9/4 \\ &\Leftrightarrow |M - \sqrt{5}/2|^2 < 9/4 \Leftrightarrow -3/2 < M - \sqrt{5}/2 < 3/2 \\ &\Leftrightarrow (-3+\sqrt{5})/2 < M < (3+\sqrt{5})/2 \Leftrightarrow M < (3+\sqrt{5})/2 \end{aligned}$$

puesto que, al ser $M > 0$, la desigualdad de la izquierda en la penúltima equivalencia es trivial. En resumen, para que exista alguna solución de (4.14), debe ser $M < (\sqrt{5}+3)/2$, tal como queríamos probar. □

4.6. La constante de Lagrange de un irracional

Acabamos de ver que, cuando $\xi = (1+\sqrt{5})/2$, la desigualdad

$$\left|\xi - \frac{p}{q}\right| < \frac{1}{Mq^2}$$

tiene una cantidad infinita de soluciones racionales p/q si y sólo si $M \leq \sqrt{5}$. Para cualquier ξ irracional, existirá un valor $M(\xi)$ tal que la desigualdad anterior tiene una cantidad infinita de soluciones si $M < M(\xi)$, y sólo una cantidad finita si $M > M(\xi)$. El teorema de Hurwitz garantiza que $M(\xi) \geq \sqrt{5}$ para todo $\xi \in \mathbb{R} \setminus \mathbb{Q}$. Este $M(\xi)$ se denomina *constante de Lagrange* de ξ (aunque hay quien lo llama *constante de Markov*). El teorema 4.37 prueba que la constante de Lagrange de la razón áurea es $M(\phi) = \sqrt{5}$. Pero, ¿cuánto vale $M(\xi)$ para otros ξ? Éste es un interesante problema que tiene conexiones insospechadas con otras áreas de la teoría de números. A continuación, y sin hacer demostraciones, comentaremos algunos de los resultados conocidos; para más información véase, por ejemplo, [54, § 10.5, pp. 257–260].

Decimos que dos números reales ξ y θ son *equivalentes* si existen enteros a, b, c y d, con $|ad - bc| = 1$, tales que[9]

$$\theta = \frac{a\xi + b}{c\xi + d},$$

lo que se denomina una *transformación modular* entre ξ y θ. Es sencillo compro-

[9] En el teorema 5.19 del tema 5 (teorema de Serret) se puede ver una importante caracterización de números equivalentes basada en fracciones continuas, lo que resulta de gran ayuda en la teoría. En realidad, el desarrollo que se hace en el libro [54] antes citado también utiliza fracciones continuas.

bar que esto define una relación de equivalencia en $\mathbb{R}$ (véase el ejercicio 4.23). Más laborioso es ver que cualquier transformación modular puede obtenerse, por composiciones sucesivas, a partir de tres básicas:

$$T(\xi) = \xi + 1, \qquad T^{-1}(\xi) = \xi - 1 \qquad \text{y} \qquad S(\xi) = 1/\xi; \tag{4.16}$$

una demostración de este hecho puede encontrarse en [75, capítulo 1, § 2].

Si ξ y θ son dos irracionales equivalentes, sus constantes de Lagrange $M(\xi)$ y $M(\theta)$ coinciden (ejercicio 4.27). Por ejemplo, para la razón áurea $\phi = (1+\sqrt{5})/2$, se tiene

$$(\sqrt{5}-1)/2 = ((1+\sqrt{5})/2)^{-1} = \frac{0\cdot\phi+1}{1\cdot\phi+0},$$

luego la constante de Lagrange de $(\sqrt{5}-1)/2$ también es $\sqrt{5}$; y lo mismo ocurre con cualquier otro número equivalente con ϕ.

Sorprendentemente —y según descubrió Hurwitz—, se cumple lo siguiente: si ξ no es equivalente con la razón áurea, entonces $M(\xi) \geq \sqrt{8}$. El valor $\sqrt{8}$ se alcanza cuando $\xi = 1+\sqrt{2}$. Nótese que, en particular, ¡la constante de Lagrange de un número nunca toma valores entre $\sqrt{5}$ y $\sqrt{8}$! Existen números cuya constante de Lagrange es infinita; sin embargo, tampoco las constantes superiores a $\sqrt{8}$ son siempre la constante de Lagrange de algún número. Así pues, ¿cuánto puede valer $M(\xi)$? Se denomina *espectro de Lagrange* al conjunto de los posibles valores de $M(\xi)$, es decir, a[10]

$$\mathscr{L} = \{M(\xi) : \xi \in \mathbb{R}\}.$$

El objetivo, ahora, sería describir el espectro de Lagrange, identificando los ξ que generan cada elemento de $\mathscr{L}$. De nuevo aparecen interesantes comportamientos que uno, a simple vista, no podía esperar. He aquí una pincelada:

Se llama *número de Markov* a un número $u \in \mathbb{N}$ para el cual existen $v, w \in \mathbb{N}$ con $u \geq v \geq w$ y tales que satisfacen la ecuación de Markov

$$u^2 + v^2 + w^2 = 3uvw.$$

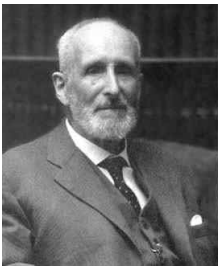
En 1880, Andrei Markov se topó con esta ecuación en cierto estudio sobre mínimos de formas cuadráticas binarias (es decir, funciones de tipo $f(x,y) = ax^2+2bxy+cy^2$) que estaba realizando. Analizó entonces la ecuación que ahora lleva su nombre y descubrió que, cuando la terna (u_1, v_1, w_1) es solución de la ecuación, también lo son (u_2, v_1, w_1), (u_1, v_2, w_1) y (u_1, v_1, w_2) con

$$u_2 = 3v_1w_1 - u_1, \quad v_2 = 3u_1w_1 - v_1, \quad w_2 = 3u_1v_1 - w_1;$$

[10] En muchos textos, se usa la notación $|\xi - p/q| < M/q^2$ y eso lleva a que el espectro de Lagrange acabe siendo el conjunto de nuestros $1/M(\xi)$. Por supuesto, esto no tiene ninguna importancia.



Andrei Andreevich Markov (1856–1922)



Oskar Perron (1880–1975)

demostró además que este proceso permitía encontrar todas las soluciones partiendo de $u = v = w = 1$, que es la solución más sencilla.[11] Los primeros números de Markov —hay infinitos de ellos— son 1, 2, 5, 13, 29, 34, 89, 169, 194, 233, 433. En 1921, el matemático alemán Oskar Perron se dio cuenta de que si u es un número de Markov y tomamos

$$\alpha_u = \frac{1}{2u}\left(\sqrt{9u^2-4} + u + \frac{2v}{w}\right),$$

entonces la constante de Lagrange de α_u es

$$M(\alpha_u) = \frac{\sqrt{9u^2-4}}{u}.$$

Así pues, ¡hemos relacionado el valor de la constante de Lagrange con las soluciones de una ecuación diofántica! Es más, si $M \in \mathscr{L}$ y $M < 3$, es seguro que $M = M(\alpha_u)$ para algún número de Markov u. Para valores por encima de 3, la estructura del espectro de Lagrange es mucho más complicada. Conviene aquí aclarar que existe otro conjunto —definido a partir de los mínimos de formas cuadráticas binarias que Markov estudiaba— que se denomina *espectro de Markov* (y los números de ese espectro se llaman *constantes de Markov*); el espectro de Markov contiene al de Lagrange, y ambos coinciden para valores menores que 3. Si el lector se ha visto interesado por esto, le recomendamos que lea las excelentes descripciones que aparecen en [22, módulo 6], [81, capítulo 10, § 5] o [98]; las demostraciones se pueden encontrar en [32].

[11] Aunque para lo que ahora nos interesa no tiene influencia, ya hemos mencionado la ecuación de Markov en el tema 1 de este libro (en concreto, en el apartado 1.6.3); allí comentábamos la denominada conjetura de Markov, un difícil problema abierto relacionado con las soluciones de la ecuación.

Por último, también merece la pena destacar que, en su estudio, Perron encontró una fórmula que proporciona la constante de Lagrange de un número a partir de su desarrollo en fracción continua. Nosotros la veremos en el tema 5 (en concreto en la proposición 5.17).

4.7. Cubrimientos de la recta real

Finalizaremos el tema mostrando que los resultados de las secciones anteriores pueden aplicarse a la teoría de cubrimientos. Para lo que vamos a emplear aquí, llamamos *cubrimiento* de la recta real $\mathbb{R}$ a un conjunto de intervalos tales que su unión es todo $\mathbb{R}$.

Los números racionales son densos en los reales, de manera que podría pensarse que cualquier familia de intervalos $\{(r-\varepsilon_r, r+\varepsilon_r)\}_{r\in\mathbb{Q}}$, con $\varepsilon_r > 0$ para todo $r \in \mathbb{Q}$, es un cubrimiento de $\mathbb{R}$. Pero esto es falso y se comprueba de manera elemental: como $\mathbb{Q}$ es numerable,[12] podemos poner $\mathbb{Q} = \{r_n : n \in \mathbb{N}\}$. Si

$$\{(r_n - \varepsilon_n, r_n + \varepsilon_n) \,:\, n \in \mathbb{N}\}$$

cubre $\mathbb{R}$, entonces necesariamente tiene que ser $\sum_{n=1}^{\infty} \varepsilon_n = \infty$. En efecto, para poder cubrir todo $\mathbb{R}$, la suma de las longitudes de los intervalos debe ser ∞ (y aun así no está garantizado, claro). Así pues, es seguro que los intervalos $(r_n - 2^{-n}, r_n + 2^{-n})$ no cubren $\mathbb{R}$.

Los teoremas de aproximación por racionales que hemos visto se pueden reinterpretar en términos de propiedades de cubrimiento. Por ejemplo, los intervalos de la forma

$$\left(\frac{p}{q} - \frac{1}{\sqrt{5}\,q^2}, \frac{p}{q} + \frac{1}{\sqrt{5}\,q^2}\right)$$

cubren $\mathbb{R}$; en realidad, cada número real pertenece a infinitos de esos intervalos (es evidente a partir del teorema de Hurwitz). Sin embargo, el apartado (b) del teorema 4.37 asegura que, si tomamos $M = (3+\sqrt{5})/2$, los intervalos de la forma

$$\left(\frac{p}{q} - \frac{1}{Mq^2}, \frac{p}{q} + \frac{1}{Mq^2}\right)$$

no cubren $\mathbb{R}$: la razón áurea $(1+\sqrt{5})/2$ no pertenece a ninguno de ellos.

[12] Éste es un hecho bien conocido; pero, si lo desea, el lector puede encontrar una demostración en la sección 6.4 (concretamente, véase el teorema 6.12).

Sobre cubrimientos de esta clase, el resultado más fino se debe al matemático hindú A. V. Prasad que, como caso particular de un teorema más general, en 1948 demuestra lo siguiente (véase [79]):

Teorema 4.38. Prasad, 1948.
Los intervalos de la forma

$$\left[\frac{p}{q}-\frac{1}{Mq^2},\frac{p}{q}+\frac{1}{Mq^2}\right]$$

cubren $\mathbb{R}$ *si* $M=(3+\sqrt{5})/2$ *(o menor, claro), y no lo cubren si* M *es mayor.*

4.8. Ejercicios

4.1. Proporciona un enunciado «más matemático» para el principio de Dirichlet (también denominado del palomar).

4.2. Sea n un entero positivo. Prueba que n tiene un múltiplo que, en base 10, se expresa como una ristra de varios dígitos «7» seguida de una ristra de varios «0».

4.3. Demuestra que en una fiesta siempre hay, al menos, dos personas que conocen al mismo número de invitados.

4.4. Demuestra que hay, al menos, dos habitantes de Nueva York que tienen el mismo número de pelos en la cabeza.

4.5. En el juego de la quiniela, una apuesta es una secuencia de 14 signos que pueden ser «1», «X» o «2». La quiniela premiada es una secuencia concreta de tales signos. El número de aciertos de una apuesta es el número de signos que coinciden, en su posición, con la quiniela premiada. ¿Cuántas apuestas tenemos que hacer para garantizar que en alguna de ellas hay, al menos, 5 aciertos? Con una apuesta menos, ¿cuántos aciertos podemos garantizar?

4.6. Sean 17 enteros positivos tales que ninguno de ellos tiene un factor primo mayor que 7. Demuestra que hay al menos una pareja de estos números cuyo producto es un cuadrado perfecto.

4.7. Sean p/q y h/k dos números racionales que satisfacen la desigualdad

$$\left|\frac{p}{q}-\frac{h}{k}\right|<\frac{1}{qk}.$$

Prueba que $p/q=h/k$.

4.8. Aunque no es necesario para su demostración, vamos a profundizar un poco en lo que hemos hecho al probar el teorema de Dirichlet. En concreto, comprueba lo siguiente:

(a) El caso $\xi = h/k \in \mathbb{Q}$ con $k \leq N$ es trivial.

(b) En otro caso, los $N+1$ números $m\xi - \lfloor m\xi \rfloor$, $m = 0, 1, 2, \ldots, N$, que aparecen en la demostración son distintos.

4.9. Sean dos enteros $a, m \geq 2$. Demuestra, a partir del teorema 4.7, que

$$\sum_{k=1}^{\infty} a^{-k^m}$$

es irracional. ¿Se te ocurre otra manera de probarlo?

4.10. Sea un polinomio $P(x) \in \mathbb{Z}[x]$ de grado mayor o igual que 2 y coeficiente director positivo. Demuestra que, para cada entero $a \geq 2$, el número

$$\sum_{k=1}^{\infty} a^{-P(k)}$$

es irracional. (Observa que este ejercicio generaliza el anterior.)

4.11. A partir del teorema de Liouville, completa la demostración de que «los irracionales cuadráticos no tienen orden de aproximación por racionales $\varphi(q)$ si $\varphi(q) = o(1/q^2)$», tal como se afirma en la nota 4.12 (página 248).

4.12. Sea a un entero mayor o igual que 2. Demuestra que cada número

$$\alpha = \sum_{k=1}^{\infty} d_k/a^{k!}$$

con $d_k \in \{0, 1, 2, \ldots, a-1\}$ para todo k, y con infinitos d_k no nulos, es de Liouville, y por tanto trascendente.

4.13. Demuestra que, para toda $\{d_n\}_{n=1}^{\infty}$ sucesión creciente de números naturales, y todo natural $a > 1$, el número

$$\alpha = a^{-d_1} + a^{-d_2 2!} + \cdots + a^{-d_n n!} + \cdots$$

es trascendente.

4.14. En el criterio de trascendencia del corolario 4.13 (que se deducía del teorema de Liouville) aparece una constante K. En cambio, en el criterio de trascendencia del corolario 4.20 (que se deducía del teorema de Thue-Siegel-Roth), no aparece tal constante; en su correspondiente lugar hay un «1». Aparentemente, este criterio hubiera sido más fuerte con un enunciado del tipo «si, para algún $C > 0$, la desigualdad $|\xi - p/q| < C/q^{2+\varepsilon}$ tiene infinitas soluciones racionales, entonces ξ es trascendente». ¿Es éste, en realidad, un resultado más fuerte?

4.15. Comprueba lo que se afirma en la nota 4.22 (página 254). Por otra parte, ¿es factible invertir la cadena deductiva? Es decir, a partir de la existencia de los $C(\xi, \varepsilon)$, ¿se puede inferir que $\nu = 2$? (tal como se afirma en el teorema de Thue-Siegel-Roth).

4.16. ¿Es cierto, en general, que la mediación de dos fracciones irreducibles proporciona una nueva fracción irreducible?

4.17. Sea $n \geq 2$. Comprueba que las tres fracciones menores (es decir, las de la izquierda, según la representación del cuadro 4.1) de $\mathscr{B}_n^*$ son

$$\frac{0}{1}, \quad \frac{1}{n} \quad \text{y} \quad \frac{1}{n-1};$$

y las tres mayores son

$$\frac{n-2}{n-1}, \quad \frac{n-1}{n} \quad \text{y} \quad \frac{1}{1}.$$

4.18. Comprueba que la mayor distancia entre términos consecutivos de $\mathscr{B}_n^*$ es $1/n$, que se alcanza únicamente en los extremos.

4.19. Sean $\frac{a}{b}$ y $\frac{c}{d}$ dos fracciones de Brocot consecutivas. Demuestra que, si el racional $\frac{h}{k}$ verifica $\frac{a}{b} < \frac{h}{k} < \frac{c}{d}$, se cumplirá $h + k \geq a + b + c + d$.

4.20. Sea h/k una fracción positiva tal que $h/k \notin \mathscr{B}_{n-1}$. Prueba que $h + k > n$.
Sugerencia. Ayúdate del ejercicio anterior.

4.21. Sea h/k un racional positivo. Demuestra que $h/k \in \mathscr{B}_{h+k-1}$.
Sugerencia. Observa el ejercicio anterior.

4.22. En la nota 4.30 hemos demostrado que toda aproximación óptima de segunda especie lo es también de primera. ¿Es cierto el recíproco?

4.23. Según hemos definido en la sección 4.6, dos números reales ξ y θ se dicen equivalentes si existen enteros a, b, c y d, con $ad - bc = \pm 1$, tales que

$\theta = (a\xi + b)/(c\xi + d)$. Comprueba que esto es, efectivamente, una relación de equivalencia.

4.24. Muestra cómo se pueden describir las transformaciones modulares —y las propiedades que caracterizan la relación de equivalencia del ejercicio anterior— por medio de matrices.

4.25. Demuestra que todos los números racionales son equivalentes.
Sugerencia. Recuerda la identidad de Bézout.

4.26. Sean ξ y θ dos números irracionales relacionados mediante una transformación modular $\theta = (a\xi + b)/(c\xi + d)$, $ad - bc = \pm 1$. Asimismo, sea $M > 0$ una constante real, y $\varepsilon > 0$ tal que $M - \varepsilon/2 > 0$. Supón que, para q suficientemente grande,

$$\xi - \frac{p}{q} = \frac{\delta}{q^2} \quad \text{con} \quad 0 < |\delta| < \frac{1}{M - \varepsilon/2}.$$

Comprueba que, en estas circunstancias, y si $cp + dq$ es lo suficientemente grande, se tiene

$$\theta - \frac{ap + bq}{cp + dq} = \frac{(ad - bc)\delta}{(cp + dq + c\delta/q)(cp + dq)}$$

y

$$\left|\theta - \frac{ap + bq}{cp + dq}\right| < \frac{1}{(M - \varepsilon)(cp + dq)^2}.$$

4.27. Prueba que, si dos números irracionales ξ y θ son equivalentes, sus constantes de Lagrange $M(\xi)$ y $M(\theta)$ coinciden.
Sugerencia. Apóyate en el ejercicio anterior. Un camino alternativo es el que sigue: Comenzaríamos probando que las tres transformaciones $T(\xi)$, $T^{-1}(\xi)$ y $S(\xi)$ definidas en (4.16) conservan la constante de Lagrange; entonces, bastaría aplicar que cualquier transformación modular se obtiene a partir de T, T^{-1} y S mediante composiciones. Pero éste es un resultado que, aunque lo hemos comentado en el texto, no hemos demostrado.

4.28. Con la notación de (4.16), sea $J = T^{-1}STST^{-1}S$ (composición de transformaciones modulares). ¿Qué hace $J(\xi)$?

4.29. En este ejercicio planteamos un resultado debido a Kronecker, un primer paso —que nosotros ya no continuaremos— en lo que se denomina *aproximación simultánea*. Dados $\xi_1, \ldots, \xi_n \in \mathbb{R}$ y $N \in \mathbb{N}$, demuestra que existen

$p_1, \ldots, p_n \in \mathbb{Z}$ y $q \in \mathbb{N}$ tales que $q \leq N^n$ y

$$\left| \xi_j - \frac{p_j}{q} \right| < \frac{1}{qN} \quad \text{para } j = 1, \ldots, n.$$

Sugerencia. En el cubo n-dimensional $[0,1)^n$, considera los puntos

$$(m\xi_1 - \lfloor m\xi_1 \rfloor, \ldots, m\xi_n - \lfloor m\xi_n \rfloor), \quad m = 0, 1, \ldots, N^n.$$



Tema 5
Fracciones continuas

5.1. La sucesión de Fibonacci

5.2. Pares de sucesiones recurrentes

5.3. Fracciones continuas simples

5.4. Fracciones continuas simples finitas

5.5. Convergencia de fracciones continuas simples

5.6. Aproximación por racionales mediante desarrollos en fracción continua

5.7. Fracciones continuas simples periódicas: irracionales cuadráticos

5.8. El desarrollo de e en fracción continua simple

5.9. Ejercicios



John Wallis
(1616–1703)



Lord William
Brouncker (1620–1684)



Johann Bernoulli
(1667–1748)

El objetivo fundamental de estas páginas es estudiar lo que se denomina fracciones continuas, así como algunas de sus propiedades y aplicaciones. Éste es un tema habitual en teoría de números, así que lo que aquí se aborda se puede encontrar en muchos textos; por ejemplo, [3, 26, 54, 55, 58, 66, 75, 78, 88, 95]. Por supuesto, cada libro tiene su enfoque particular, e incide más en unos aspectos que en otros.

A grandes rasgos, una fracción continua es un inmenso castillo de fracciones anidadas de modo que, en el denominador de cada fracción, vuelve a aparecer una nueva fracción (sumada a otros números). Esto conduce a procesos infinitos que pueden usarse para aproximar números reales por medio de su desarrollo en fracción continua.

De hecho, muchos de los resultados de aproximación de números reales por racionales pueden obtenerse tanto mediante fracciones de Brocot (o de Farey), como en el tema anterior, como usando fracciones continuas; por ejemplo, en [54, § 10.4] o [58, § 7] se prueba el teorema de Hurwitz por medio de fracciones continuas. Así, en este libro podríamos haber optado por prescindir de las fracciones de Brocot. Pero, a juicio del que escribe, tanto las fracciones de Brocot como las fracciones continuas son bellos conceptos matemáticos interesantes por sí mismos, luego merece la pena que nos hayamos detenido en ambos.

Históricamente, las fracciones continuas —o algo equivalente a ellas— ya habían sido utilizadas por los matemáticos hindúes; por ejemplo, Aryabhata (nacido en el año 476) las usó para encontrar soluciones enteras de ecuaciones lineales con varias indeterminadas (esencialmente, este mismo procedimiento se sigue usando en la actualidad). Entre los matemáticos occidentales, podemos citar a Rafael Bombelli, quien —en 1572— utilizó fracciones continuas para aproximar raíces cuadradas. Aunque, a decir verdad, Bombelli realmente no pensaba en términos de fracciones

continuas, y posiblemente fue Pietro Cataldi, en 1613, el primero que concibió la idea precisa de fracción continua. En el siglo XVII, también se sirvieron de ellas John Wallis, Lord William Brouncker, Gottfried Leibniz, Johann Bernoulli, Christiaan Huygens... Finalmente, uno de los mayores impulsores del uso de desarrollos en fracción continua y que más partido sacaron de ellos fue Leonhard Euler quien, entre otras cosas, probó que una fracción continua simple infinita siempre proporciona un número irracional. No podemos aquí evitar citar [43], una edición facsímil acompañada de una traducción anotada y comentada del *Introductio in Analysin Infinitorum* de Euler, uno de cuyos capítulos está dedicado a las fracciones continuas.

Como paso preparatorio a nuestro estudio sobre fracciones continuas, comenzaremos introduciendo la conocida sucesión de Fibonacci. El verdadero nombre de Fibonacci era Leonardo de Pisa, y vivió entre los siglos XII y XIII. Él fue uno de los principales divulgadores de los números hindo-arábigos y el sistema de notación posicional que usamos actualmente, pero que, en su época, todavía era prácticamente desconocido en Europa.[1] La sucesión que ahora lleva su nombre aparecía en su libro *Liber Abaci*, editado en 1202, en relación con un problema de reproducción de conejos (véase el ejercicio 5.1). Sin embargo, su expresión mediante una fórmula recurrente no fue introducida hasta mucho más tarde: lo hizo Albert Girard, alrededor de 1634. Aquí veremos unas pocas propiedades, pero no podemos dejar de citar que ha sido objeto de multitud de estudios, y que aparece en problemas de todo tipo, muchos de ellos relacionados con la naturaleza. Actualmente, sigue siendo un tema matemático de interés; basta decir que hay una revista (el *Fibonacci Quarterly*, que se publica desde 1963) especializada en el estudio de la sucesión de Fibonacci y temas afines. Una monografía —de fácil lectura— dedicada por completo a esta sucesión es [103].

5.1. La sucesión de Fibonacci

Se denomina *sucesión de Fibonacci* a la que, partiendo de los dos primeros términos iguales a 1, cada término se construye sumando los dos anteriores, esto es

$$1,\ 1,\ 2,\ 3,\ 5,\ 8,\ 13,\ 21,\ 34,\ 55,\ 89,\ 144,\ \ldots.$$

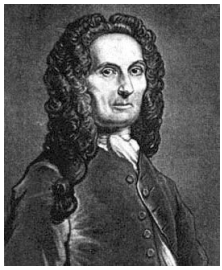
Si la denotamos $\{f_n\}_{n=1}^{\infty}$, su definición se reduce a la fórmula recursiva

$$\begin{aligned} f_1 &= f_2 = 1, \\ f_n &= f_{n-1} + f_{n-2}, \quad n \geq 3; \end{aligned}$$

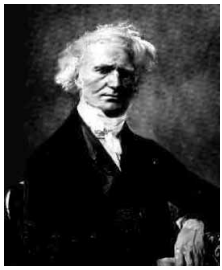
[1] Pueden verse más detalles en la sección 1.3 de este libro.



Fibonacci (Leonardo de Pisa) (1170–1250)



Abraham de Moivre (1667–1754)



Jacques Binet (1786–1856)

individualmente, a los f_n los llamamos números de Fibonacci. Nos dedicaremos ahora a ver, de dos formas distintas —y se propone otra más en el ejercicio 5.3—, cómo puede obtenerse una expresión explícita para f_n; en concreto, ésta:

$$f_n = \frac{1}{\sqrt{5}}\left(\left(\frac{1+\sqrt{5}}{2}\right)^n - \left(\frac{1-\sqrt{5}}{2}\right)^n\right) \tag{5.1}$$

que se suele denominar fórmula de Binet. Aunque, como decimos, esta fórmula se atribuye habitualmente a Jacques Binet —y, de hecho, la probó en 1843—, él no fue el primero en hallarla: Abraham de Moivre la había encontrado alrededor de 1730.

Primer método. Supongamos que existen cuatro números reales α, β, c, d tales que[2]

$$f_n = c\alpha^n + d\beta^n;$$

en principio, no tendrían por qué existir, pero, si los encontramos, es que existen. Así,

$$\begin{aligned} f_n = f_{n-1} + f_{n-2} &\Leftrightarrow c\alpha^n + d\beta^n = c\alpha^{n-1} + d\beta^{n-1} + c\alpha^{n-2} + d\beta^{n-2} \\ &\Leftrightarrow c\alpha^{n-2}(\alpha^2 - \alpha - 1) = -d\beta^{n-2}(\beta^2 - \beta - 1). \end{aligned}$$

Una forma de asegurar esto es tomar α y β como dos raíces (iguales o distintas) del polinomio[3] $x^2 - x - 1 = 0$, es decir,

$$\alpha, \beta = \frac{1 \pm \sqrt{5}}{2}.$$

[2] Esta técnica, conocida como *método de los coeficientes indeterminados*, no sólo es útil para resolver ecuaciones definidas en términos de recurrencias. También es común emplearla para hallar soluciones de ecuaciones diferenciales.

[3] En general, la clave para resolver una ecuación en recurrencias lineal, es decir, de tipo $a_k r_n + a_{k-1} r_{n-1} + \cdots + a_1 r_{n-k+1} + a_0 r_{n-k} = 0$ (donde los coeficientes a_j son conocidos, $a_k \neq 0$, y lo

Sólo con esto ya se tiene la relación de recurrencia (c y d pueden ser cualesquiera). Ahora (para fijar c y d) hay que imponer $f_1 = f_2 = 1$, es decir

$$\begin{cases} 1 = c\alpha + d\beta, \\ 1 = c\alpha^2 + d\beta^2. \end{cases} \tag{5.2}$$

Salvo en el caso trivial $\alpha = \beta = 1$ (ante el que no nos encontramos), es fácil comprobar que, para que el sistema (5.2) tenga solución, debe ser $\alpha \neq \beta$ (ejercicio 5.2). Podemos entonces suponer $\alpha = \frac{1+\sqrt{5}}{2}$ y $\beta = \frac{1-\sqrt{5}}{2}$. Resolviendo el sistema,

$$c = \frac{\beta - 1}{(\beta - \alpha)\alpha} = \frac{1}{\sqrt{5}}, \qquad d = \frac{\alpha - 1}{(\alpha - \beta)\beta} = -\frac{1}{\sqrt{5}}.$$

Y de aquí se sigue (5.1).

Segundo método. Definamos la *función generatriz* de los números de Fibonacci mediante[4]

$$f(z) = \sum_{n=1}^{\infty} f_n z^n.$$

Como $f_n < 2^n$ (por inducción: $f_{n-1} < 2^{n-1}$, $f_{n-2} < 2^{n-2} \Rightarrow f_n < 2^{n-1} + 2^{n-2} < 2^n$), la serie que define f es convergente para $|z| < 1/2$; es más, como función compleja, f es holomorfa en el disco de centro 0 y radio $1/2$ (recuérdese que el radio de convergencia R de una serie de potencias viene dado por fórmula de Cauchy-Hadamard $R^{-1} = \limsup_n \sqrt[n]{|f_n|}$). Además,

$$f(z) = z + z^2 + \sum_{n=3}^{\infty} f_n z^n = z + z^2 + \sum_{n=3}^{\infty} f_{n-1} z^n + \sum_{n=3}^{\infty} f_{n-2} z^n$$

que se busca es la sucesión $\{r_n\}_{n=1}^{\infty}$), siempre la da un polinomio: la solución de la ecuación en recurrencias se obtiene de manera directa a partir de las raíces del polinomio $a_k x^k + a_{k-1} x^{k-1} + \cdots + a_1 x + a_0 = 0$. Lo mismo sucede con la ecuación diferencial $a_k y^{(k)} + a_{k-1} y^{(k-1)} + \cdots + a_1 y' + a_0 y = 0$.

[4] La técnica que vamos a emplear se puede usar para hallar expresiones explícitas de muchas otras sucesiones definidas por medio de recurrencias. Casi sin ninguna variación, sirve siempre que la sucesión satisface una recurrencia lineal homogénea con coeficientes constantes. La función generatriz que se obtiene es una función racional, y la expresión explícita para los términos de la sucesión recurrente descansa sobre las raíces del denominador de la función generatriz. En realidad, el uso de funciones generatrices es útil en situaciones mucho más generales. Es el método más potente que se conoce para estudiar sucesiones de números de todo tipo; manipulando las series infinitas que «generan» las sucesiones es fácil llegar a descubrir muchas propiedades. Una buena referencia para todo esto es [47, capítulo 7].

$$= z + z^2 + z\sum_{n=3}^{\infty} f_{n-1}z^{n-1} + z^2\sum_{n=3}^{\infty} f_{n-2}z^{n-2}$$
$$= z + z^2 + z(f(z) - z) + z^2 f(z);$$

por tanto, $f(z)(1 - z - z^2) = z$, luego la función generatriz es

$$f(z) = \frac{-z}{z^2 + z - 1}.$$

Obviamente, si desarrollamos esta f en serie de potencias de z, los coeficientes que obtengamos deberán ser los f_n (el desarrollo en serie de potencias es único). Para ello, sean a y b las raíces de $z^2 + z - 1 = 0$, es decir

$$a = \frac{-1 - \sqrt{5}}{2}, \quad b = \frac{-1 + \sqrt{5}}{2}.$$

Entonces, descomponiendo f en fracciones simples, y usando la fórmula de la suma de una progresión geométrica (recordémosla: $\sum_{n=0}^{\infty} t^n = 1/(1-t)$, válida para $|t| < 1$), cuando $|z| < |a|$ y $|z| < |b|$ se tiene

$$f(z) = \frac{-z}{z^2 + z - 1} = \frac{-z}{(z-a)(z-b)} = \frac{z}{b-a}\left(\frac{1}{z-a} - \frac{1}{z-b}\right)$$
$$= \frac{z}{b-a}\left(\frac{1}{b}\frac{1}{1 - z/b} - \frac{1}{a}\frac{1}{1 - z/a}\right)$$
$$= \frac{z}{b-a}\left(\frac{1}{b}\sum_{n=0}^{\infty}\frac{z^n}{b^n} - \frac{1}{a}\sum_{n=0}^{\infty}\frac{z^n}{a^n}\right) = \frac{1}{b-a}\left(\sum_{n=0}^{\infty}\frac{z^{n+1}}{b^{n+1}} - \sum_{n=0}^{\infty}\frac{z^{n+1}}{a^{n+1}}\right)$$
$$= \frac{1}{b-a}\left(\sum_{n=1}^{\infty}\frac{1}{b^n}z^n - \sum_{n=1}^{\infty}\frac{1}{a^n}z^n\right) = \sum_{n=1}^{\infty}\left(\frac{b^{-n} - a^{-n}}{b-a}\right)z^n,$$

luego

$$f_n = \frac{1}{b-a}\left(b^{-n} - a^{-n}\right), \quad n \geq 1.$$

Así, tras comprobar que $b - a = \sqrt{5}$, $\frac{1}{b} = \frac{1+\sqrt{5}}{2}$, y $\frac{1}{a} = \frac{1-\sqrt{5}}{2}$, de nuevo obtenemos la fórmula de Binet (5.1).

El escocés Robert Simson, en 1753, se dio cuenta de que, aunque los números de Fibonacci crecían indefinidamente, el cociente entre dos consecutivos parecía



Robert Simson
(1687–1768)



Luca Pacioli
(1445–1517)

aproximarse más y más a la *razón áurea* $\phi = \frac{1+\sqrt{5}}{2}$, una constante ubicua.[5] Dicho de otra manera, el crecimiento de los términos de la sucesión de Fibonacci se parece bastante —en cierto sentido— al de una progresión geométrica de razón $(1+\sqrt{5})/2$. A la vista de la fórmula de Binet para f_n, la observación de Simson es fácil de comprobar:

$$\lim_n \frac{f_n}{f_{n-1}} = \lim_n \frac{\left(\frac{1+\sqrt{5}}{2}\right)^n - \left(\frac{1-\sqrt{5}}{2}\right)^n}{\left(\frac{1+\sqrt{5}}{2}\right)^{n-1} - \left(\frac{1-\sqrt{5}}{2}\right)^{n-1}}$$
$$= \frac{1+\sqrt{5}}{2} \lim_n \frac{1 - \left(\frac{1-\sqrt{5}}{1+\sqrt{5}}\right)^n}{1 - \left(\frac{1-\sqrt{5}}{1+\sqrt{5}}\right)^{n-1}} = \frac{1+\sqrt{5}}{2}, \qquad (5.3)$$

la razón áurea.[6]

[5] También se alude a la razón áurea llamándola *número áureo*, *proporción áurea*, *canon áureo*, *regla de oro* o *divina proporción*, nombre este último con el que la denominó el matemático renacentista italiano Luca Pacioli en su libro *De divina proportione*, bellamente ilustrado con figuras de Leonardo da Vinci y publicado en 1509. Ya nos hemos topado con la razón áurea en el tema anterior en relación con el teorema de Hurwitz que, en principio, no parecía tener conexión con esto. La volveremos a ver luego (página 290), cuando calculemos su desarrollo en fracción continua. Aunque aquí no estamos interesados en ello, no podemos de dejar de mencionar su importancia en arquitectura, pintura y geometría; por ejemplo —y como curiosidad—, aparece en el lado de un pentágono estrellado, lo cual era conocido desde la época griega. Tampoco nos resistimos a comentar su propiedad más conocida: si los dos lados x e y (con $x > y$) de un rectángulo están en proporción áurea, y de ese rectángulo se quita un cuadrado de lado y, los lados del rectángulo sobrante, y y $x - y$, están de nuevo en proporción áurea.

[6] Aunque realmente no tiene importancia para lo que nosotros perseguimos, no nos podemos resistir a comentar la conexión entre el límite (5.3) y la serie $f(z) = \sum_{n=0}^{\infty} f_n z^n$ que definíamos en

Lo que sigue permite saber que el límite del cociente de términos consecutivos existe para una clase bastante general de sucesiones de este tipo.

5.2. Pares de sucesiones recurrentes

Un *par regular de sucesiones recurrentes* (normalmente, no especificaremos el «regular») es un par de sucesiones $\{P_k\}_{k=1}^{\infty}$ y $\{Q_k\}_{k=1}^{\infty}$ definidas de la forma[7]

$$\begin{aligned} P_k &= q_k P_{k-1} + P_{k-2}, \quad P_1, P_2 \in \mathbb{R}, \\ Q_k &= q_k Q_{k-1} + Q_{k-2}, \quad Q_1, Q_2 \in [0, \infty), \end{aligned}$$

donde $q_k \in [1, \infty)$ para todo k, y con la condición de que

$$\Delta = P_2 Q_1 - P_1 Q_2 > 0.$$

La exigencia $\Delta > 0$ no es muy restrictiva (véase el ejercicio 5.6) ya que, si $\Delta = 0$, existe una constante λ tal que $P_k = \lambda Q_k$ para todo k o $Q_k = \lambda P_k$ para todo k; y, si $\Delta < 0$, basta considerar la sucesión $\{-P_k\}_{k=1}^{\infty}$ en lugar de $\{P_k\}_{k=1}^{\infty}$.

Observemos que $\Delta > 0$ fuerza a que Q_1 y Q_2 no sean simultáneamente nulos, de donde $Q_k > 0$ para $k \geq 3$; de aquí se sigue

$$\begin{aligned} Q_k &> Q_{k-1}, && \text{para } k \geq 5, \\ Q_k &> 2Q_{k-2}, && \text{para } k \geq 6, \\ Q_k &\to \infty && \text{cuando } k \to \infty. \end{aligned}$$

Veamos ahora una serie de interesantes propiedades:

---

el segundo método para comprobar la fórmula de Binet (5.1). Allí veíamos —a partir de la fórmula de recurrencia de los números de Fibonacci— que esa serie convergía al menos cuando $|z| < 1/2$. De hecho, si llamamos R al radio de convergencia de esa serie, una vez probado (5.1) es fácil ver, aplicando la conocida fórmula $R^{-1} = \limsup_n \sqrt[n]{|f_n|}$ (fórmula de Cauchy-Hadamard), que R es algo mayor que $1/2$, concretamente $R = 2/(1+\sqrt{5}) = (\sqrt{5}-1)/2 = 0.618\ldots$ Aunque es la expresión con raíces n-ésimas antes dada la que con seguridad proporciona el valor del radio de convergencia de una serie, es bien sabido que, si existe $\lim_n |f_{n+1}/f_n|$, también se tiene $R^{-1} = \lim_n |f_{n+1}/f_n|$; así pues, la fórmula límite (5.3) de nuevo permite obtener $R = (\sqrt{5}-1)/2$.

[7] Un caso más general —que ya no se denomina regular— es si definimos

$$\begin{aligned} P_k &= q_k P_{k-1} + p_k P_{k-2}, \quad P_1, P_2 \in \mathbb{R}, \\ Q_k &= q_k Q_{k-1} + p_k Q_{k-2}, \quad Q_1, Q_2 \in [0, \infty), \end{aligned}$$

pero sus propiedades son algo peores. Del mismo modo que —más adelante en este mismo tema— los pares regulares de sucesiones recurrentes se relacionan con las fracciones continuas simples, las no regulares lo hacen con las no simples (también denominadas generalizadas). En teoría de números, son las simples las que tienen más interés.

Proposición 5.1.
Dado $\{P_k\}_{k=1}^{\infty}$ y $\{Q_k\}_{k=1}^{\infty}$ un par de sucesiones recurrentes, se verifica

(a) $P_kQ_{k-1} - P_{k-1}Q_k = -(P_{k-1}Q_{k-2} - P_{k-2}Q_{k-1})$, $k \geq 3$.

(b) $P_kQ_{k-1} - P_{k-1}Q_k = (-1)^k\Delta$, $k \geq 2$.

(c) $\dfrac{P_k}{Q_k} - \dfrac{P_{k-1}}{Q_{k-1}} = \dfrac{(-1)^k\Delta}{Q_kQ_{k-1}}$ *si* $Q_k \neq 0 \neq Q_{k-1}$ *(luego para todo* $k \geq 4$*).*

(d) $P_kQ_{k-2} - P_{k-2}Q_k = (-1)^{k-1}q_k\Delta$, $k \geq 3$.

(e) $\dfrac{P_k}{Q_k} - \dfrac{P_{k-2}}{Q_{k-2}} = q_k\dfrac{(-1)^{k-1}\Delta}{Q_kQ_{k-2}}$ *si* $Q_k \neq 0 \neq Q_{k-2}$ *(luego para todo* $k \geq 5$*).*

Demostración. Comencemos con (a). Sin más que usar la definición y aplicar las propiedades de los determinantes —con ellos, la demostración adquiere una sencillez y una elegancia innegables—, se tiene

$$\begin{aligned} P_kQ_{k-1} - P_{k-1}Q_k &= \begin{vmatrix} P_k & P_{k-1} \\ Q_k & Q_{k-1} \end{vmatrix} = \begin{vmatrix} q_kP_{k-1} + P_{k-2} & P_{k-1} \\ q_kQ_{k-1} + Q_{k-2} & Q_{k-1} \end{vmatrix} \\ &= \begin{vmatrix} P_{k-2} & P_{k-1} \\ Q_{k-2} & Q_{k-1} \end{vmatrix} = -(P_{k-1}Q_{k-2} - P_{k-2}Q_{k-1}). \end{aligned}$$

La parte (b) es inmediata usando (a) e inducción; y (c) es consecuencia inmediata de (b). Veamos ahora (d):

$$\begin{aligned} P_kQ_{k-2} - P_{k-2}Q_k &= \begin{vmatrix} P_k & P_{k-2} \\ Q_k & Q_{k-2} \end{vmatrix} = \begin{vmatrix} q_kP_{k-1} + P_{k-2} & P_{k-2} \\ q_kQ_{k-1} + Q_{k-2} & Q_{k-2} \end{vmatrix} \\ &= q_k \begin{vmatrix} P_{k-1} & P_{k-2} \\ Q_{k-1} & Q_{k-2} \end{vmatrix} = q_k(-1)^{k-1}\Delta. \end{aligned}$$

Finalmente, (e) se sigue de (d). □

Saquemos más partido de esta proposición. La propiedad (e) implica

$$\frac{P_{2k}}{Q_{2k}} - \frac{P_{2k-2}}{Q_{2k-2}} < 0 \quad \text{y} \quad \frac{P_{2k+1}}{Q_{2k+1}} - \frac{P_{2k-1}}{Q_{2k-1}} > 0;$$

en otras palabras, los cocientes P_k/Q_k de orden par (es decir, de subíndice par) van decreciendo, y los de orden impar van creciendo. Además, de la propiedad (c) se deduce que

$$0 < \frac{P_{2k}}{Q_{2k}} - \frac{P_{2k-1}}{Q_{2k-1}} \longrightarrow 0 \quad \text{cuando } k \to \infty. \tag{5.4}$$

Entonces, $\{P_{2k}/Q_{2k}\}_{k=1}^{\infty}$ es una sucesión decreciente y acotada inferiormente (por cualquier cociente de orden impar, como P_1/Q_1), luego tiene límite; también lo tiene $\{P_{2k-1}/Q_{2k-1}\}_{k=1}^{\infty}$, por ser creciente y acotada superiormente (por P_2/Q_2, por ejemplo). Finalmente, de la convergencia a 0 en (5.4) se sigue que esos dos límites deben coincidir. Con todo esto, es claro que acabamos de probar el siguiente resultado:

Teorema 5.2.
Si $\{P_k\}_{k=1}^{\infty}$, $\{Q_k\}_{k=1}^{\infty}$ es un par de sucesiones recurrentes, entonces existe $\lim_k P_k/Q_k = \ell \in \mathbb{R}$; es más, nos encontramos ante esta situación:

$$\frac{P_1}{Q_1} < \frac{P_3}{Q_3} < \frac{P_5}{Q_5} < \cdots \longrightarrow \ell \longleftarrow \cdots < \frac{P_6}{Q_6} < \frac{P_4}{Q_4} < \frac{P_2}{Q_2}$$

(si $Q_1 = 0$ o $Q_2 = 0$, suprimimos el primer término de la izquierda o de la derecha).

Para interpretar la sucesión de Fibonacci $\{f_n\}_{n=1}^{\infty}$ por medio de pares de sucesiones recurrentes, basta tomar $q_k = 1$ para todo k. Entonces, las sucesiones $\{P_k\}_{k=1}^{\infty}$ y $\{Q_k\}_{k=1}^{\infty}$ vienen dadas por

$$\begin{aligned} P_k &: \quad 1,\ 1,\ 2,\ 3,\ 5,\ 8,\ 13,\ 21,\ 34,\ \ldots, \\ Q_k &: \quad 0,\ 1,\ 1,\ 2,\ 3,\ 5,\ 8,\ 13,\ 21,\ \ldots \end{aligned}$$

Así, $P_k = f_k$ y $Q_k = f_{k-1}$ (para $k = 1$, consideramos $f_0 = 0$). El límite de los cocientes P_k/Q_k del teorema anterior es, según (5.3),

$$\lim_k \frac{P_k}{Q_k} = \lim_k \frac{f_k}{f_{k-1}} = \frac{1+\sqrt{5}}{2}.$$

5.3. Fracciones continuas simples

Una *fracción continua* es una expresión de la forma

$$q_0 + \cfrac{p_1}{q_1 + \cfrac{p_2}{q_2 + \cfrac{p_3}{q_3 + \cdots}}} \tag{5.5}$$

(finita o infinita), donde $q_k, p_k \in \mathbb{R}$ con $q_k \neq 0$ para todo $k \geq 1$ (aunque enseguida nos interesará imponer condiciones mucho más restrictivas). Con significado evidente, hablaremos de fracciones continuas finitas e infinitas. Antes de proseguir,

obsérvese que, si la fracción continua es finita, nos proporciona un número real bien definido (salvo que se anule algún denominador, lo cual nunca ocurre si exigimos, por ejemplo, $p_k \geq 0$ y $q_k > 0$ para $k \geq 1$). Pero, si es infinita, tendremos que aclarar qué queremos decir con este proceso infinito.

Cuando $p_k = 1$ para todo k y $q_k > 0$ para todo $k \geq 1$, decimos que estamos ante una *fracción continua simple*. Salvo en la nota 5.10 (página 300) y en los ejercicios 5.38 a 5.44 (ahí abordaremos el caso $q_0 \in \mathbb{Z}$, $q_k, p_k \in \mathbb{N}$ para todo $k \geq 1$), sólo vamos a usar fracciones continuas simples, así que frecuentemente las llamaremos fracciones continuas, sin más; en estas circunstancias, a las de tipo (5.5) se les suele denominar *fracciones continuas generalizadas*. En muchas ocasiones, escribir (5.5) no es cómodo, y requiere mucho espacio, de aquí que sea habitual utilizar notaciones alternativas. En nuestro caso, usaremos

$$\langle q_0, q_1, q_2, q_3, \ldots \rangle = q_0 + \cfrac{1}{q_1 + \cfrac{1}{q_2 + \cfrac{1}{q_3 + \cdots}}}$$

(la misma de [71, 78]); aunque nosotros no las seguiremos, otras notaciones comunes son

$$\begin{aligned} \langle q_0, q_1, q_2, q_3, \ldots \rangle &= q_0 + \frac{1}{q_1+}\frac{1}{q_2+}\frac{1}{q_3+}\cdots \\ &= q_0 + \frac{1}{\lfloor q_1 \rfloor} + \frac{1}{\lfloor q_2 \rfloor} + \frac{1}{\lfloor q_3 \rfloor} + \cdots . \end{aligned}$$

Llamaremos *cociente incompleto* k-ésimo al número q_k; y *convergente* k-ésimo al número

$$\langle q_0, q_1, q_2, \ldots, q_k \rangle = q_0 + \cfrac{1}{q_1 + \cfrac{1}{q_2 + \cfrac{1}{\ddots + \cfrac{1}{q_k}}}}$$

que, como toda fracción continua finita, es un número perfectamente definido (un racional, si $q_0 \in \mathbb{Z}$ y $q_k \in \mathbb{N}$ para todo $k \geq 1$). Se dice que una fracción continua infinita $\langle q_0, q_1, q_2, q_3, \ldots \rangle$ es *convergente* si existe

$$\lim_k \langle q_0, q_1, q_2, \ldots, q_k \rangle \tag{5.6}$$

(para evitar malentendidos, recalquemos que no permitimos que el límite anterior

valga infinito). Frecuentemente, el término «fracción continua» designa también este límite. Para tranquilidad del lector, adelantemos acontecimientos: las fracciones continuas que nos interesan son las que cumplen

$$q_0 \in \mathbb{Z} \quad \text{y} \quad q_k \in \mathbb{N} \quad \text{para todo } k \geq 1; \tag{5.7}$$

probaremos que, bajo estas hipótesis, cualquier fracción continua es convergente.

Nota 5.3. Sin la condición (5.7), las fracciones continuas que surgen son más usuales en análisis matemático —como forma de aproximar funciones y, sobre todo, en conexión con los aproximantes de Padé[8]— que en teoría de números; consúltense, por ejemplo, [57] o [75]. Nosotros sólo veremos tales fracciones continuas en los ejercicios 5.33 a 5.35, en los que, precisamente, mostramos que no siempre convergen. ◁

Comencemos analizando un ejemplo sencillo: $q_k = 1$ para todo k. Aquí aparecerán, de nuevo, la sucesión de Fibonacci y, ¡cómo no!, la razón áurea.

Ejemplo. Comprobar que $\langle 1,1,1,1,\ldots\rangle = \frac{1+\sqrt{5}}{2}$.

Primera solución. Sea $\{f_k\}_{k=1}^{\infty}$ la sucesión de Fibonacci. Entonces,

$$\langle 1,1\rangle = 1 + \frac{1}{1} = \frac{2}{1} = \frac{f_3}{f_2},$$

$$\langle 1,1,1\rangle = 1 + \frac{1}{1+\dfrac{1}{1}} = 1 + \frac{1}{2/1} = \frac{3}{2} = \frac{f_4}{f_3},$$

$$\langle 1,1,1,1\rangle = 1 + \frac{1}{\langle 1,1,1\rangle} = 1 + \frac{2}{3} = \frac{5}{3} = \frac{f_5}{f_4}.$$

Comprobemos por inducción que

$$\langle \underbrace{1,1,1,\ldots,1}_{n}\rangle = \frac{f_{n+1}}{f_n}.$$

Basta ver que, supuesto eso cierto, también se cumple cambiando n por $n+1$:

$$\langle \underbrace{1,1,1,\ldots,1}_{n+1}\rangle = 1 + \frac{1}{\langle \underbrace{1,1,1,\ldots,1}_{n}\rangle} = 1 + \frac{f_n}{f_{n+1}} = \frac{f_{n+1}+f_n}{f_{n+1}} = \frac{f_{n+2}}{f_{n+1}}.$$

Ahora, usando (5.3), es claro que $\langle 1,1,1,1,\ldots\rangle = \lim_n \frac{f_{n+1}}{f_n} = \frac{1+\sqrt{5}}{2}$. ◊

[8] Los aproximantes de Padé aproximan funciones por medio de funciones racionales, de manera análoga a como el desarrollo de Taylor lo hace mediante polinomios.

Segunda solución. También podemos llegar al mismo resultado sin aplicar (5.3), que requiere conocer la expresión explícita de los f_n. Para ello, usemos que existe $\lim_n \frac{f_{n+1}}{f_n} = \xi \in \mathbb{R}$ (lo cual está garantizado por el teorema 5.2); además, $\xi \geq 1$, pues la sucesión de Fibonacci es creciente. Entonces,

$$\frac{f_{n+1}}{f_n} = \frac{f_n + f_{n-1}}{f_n} = 1 + \frac{f_{n-1}}{f_n} = 1 + \frac{1}{f_n/f_{n-1}}$$

luego $\xi = 1 + \frac{1}{\xi}$; es decir, $\xi^2 - \xi - 1 = 0$ y, de aquí, $\xi = \frac{1+\sqrt{5}}{2}$. ◊

Tercera solución. Finalmente, este método ni siquiera usa los f_n. La razón áurea $\phi = (1+\sqrt{5})/2$ es raíz de la ecuación $x^2 - x - 1 = 0$ (la otra raíz es $(1-\sqrt{5})/2$, que es negativa). Podemos escribir $\phi^2 - \phi - 1 = 0$ como $\phi = 1 + 1/\phi$ (es decir, $\phi = \langle 1, \phi \rangle$). Ahora, sustituimos el «ϕ» que aparece en el denominador de $\phi = 1 + 1/\phi$ de nuevo por $1 + 1/\phi$, con lo cual queda

$$\phi = 1 + \frac{1}{\phi} = 1 + \frac{1}{1 + \dfrac{1}{\phi}},$$

y así sucesivamente. Por tanto,

$$\phi = \langle 1, \phi \rangle = \langle 1, 1, \phi \rangle = \cdots = \langle 1, 1, 1, 1, \ldots \rangle.$$

Realmente, para que esta demostración sea rigurosa tenemos que probar que el proceso $\langle 1, 1, 1, 1, \ldots \rangle$ tiene sentido y que converge a ϕ. Pero esto se desprende de los resultados sobre fracciones continuas que veremos en la sección 5.5 (concretamente, del corolario 5.7); véase, además, el ejercicio 5.24. ◊

5.4. Fracciones continuas simples finitas: el algoritmo de la división

Recordemos que hemos impuesto $q_0 \in \mathbb{Z}$ y $q_k \in \mathbb{N}$ para todo $k \geq 1$. Trivialmente, toda fracción continua (simple) finita

$$\langle q_0, q_1, q_2, \ldots, q_{n-1}, q_n \rangle = q_0 + \cfrac{1}{q_1 + \cfrac{1}{q_2 + \cfrac{1}{\ddots \; + \cfrac{1}{q_{n-1} + \cfrac{1}{q_n}}}}} \tag{5.8}$$

representa un número racional.

Recíprocamente, todo número racional se puede representar mediante una fracción continua finita; para ello, basta usar el algoritmo de Euclides para la división. Veámoslo:

Sea un número racional $\frac{r_0}{r_1}$ (aunque no hace falta, podemos pensar que es una fracción irreducible) con $r_1 > 0$. Entonces,

$$\begin{aligned}
r_0 &= r_1 q_0 + r_2, \quad q_0 \in \mathbb{Z}, \quad 0 \leq r_2 < r_1; \quad \text{si } r_2 \neq 0,\\
r_1 &= r_2 q_1 + r_3, \quad q_1 \in \mathbb{N}, \quad 0 \leq r_3 < r_2; \quad \text{si } r_3 \neq 0,\\
r_2 &= r_3 q_2 + r_4, \quad q_2 \in \mathbb{N}, \quad 0 \leq r_4 < r_3; \quad \text{si } r_4 \neq 0,\\
&\dots\\
r_{n-1} &= r_n q_{n-1} + r_{n+1}, \quad q_{n-1} \in \mathbb{N}, \quad 0 \leq r_{n+1} < r_n; \quad \text{si } r_{n+1} \neq 0,\\
r_n &= r_{n+1} q_n,
\end{aligned}$$

donde en el último paso hemos supuesto que la división r_n/r_{n+1} es, esta vez, exacta. Como los r_k son enteros no negativos cada vez menores, esto sucederá, necesariamente, alguna vez. La primera de estas ecuaciones se puede expresar como

$$\frac{r_0}{r_1} = q_0 + \frac{r_2}{r_1} = q_0 + \frac{1}{r_1/r_2};$$

la segunda,

$$\frac{r_1}{r_2} = q_1 + \frac{r_3}{r_2} = q_1 + \frac{1}{r_2/r_3};$$

y así sucesivamente. Sustituyendo, llegamos a

$$\begin{aligned}
\frac{r_0}{r_1} &= q_0 + \frac{1}{r_1/r_2} = q_0 + \cfrac{1}{q_1 + \cfrac{1}{r_2/r_3}} = q_0 + \cfrac{1}{q_1 + \cfrac{1}{q_2 + \cfrac{1}{r_3/r_4}}} = \cdots\\
&= q_0 + \cfrac{1}{q_1 + \cfrac{1}{q_2 + \cfrac{1}{\ddots\, q_{n-1} + \cfrac{1}{r_n/r_{n+1}}}}} = q_0 + \cfrac{1}{q_1 + \cfrac{1}{q_2 + \cfrac{1}{\ddots\, q_{n-1} + \cfrac{1}{q_n}}}}\\
&= \langle q_0, q_1, q_2, \dots, q_{n-1}, q_n \rangle.
\end{aligned}$$

Nótese que, en el proceso anterior, $r_0 \in \mathbb{Z}$ y $r_1 \in \mathbb{N}$, luego q_0 puede ser nulo o incluso un entero negativo (si $r_0 < 0$). Pero el algoritmo de la división nos proporciona $r_2 < r_1$, luego $q_1 \geq 1$; y así sucesivamente para el resto de los q_k. Por tanto, $q_0 \in \mathbb{Z}$ y $q_k \in \mathbb{N}$ para todo $k \geq 1$, tal como exigíamos.

5.4.1. Unicidad del desarrollo en fracción continua finita

Trivialmente,

$$\begin{aligned}\langle q_0, q_1, q_2, \ldots, q_{n-1}, q_n, 1\rangle &= q_0 + \cfrac{1}{q_1 + \cfrac{1}{q_2 + \cfrac{1}{\ddots \; q_n + \cfrac{1}{1}}}} \\ &= \langle q_0, q_1, q_2, \ldots, q_{n-1}, q_n + 1\rangle.\end{aligned}$$

Por tanto, la representación de un número racional como fracción continua (simple) finita no es única. El siguiente resultado dice que, salvo este detalle, sí lo es (la distinción del caso $n = 0$ o $m = 0$ es sólo para que el número 1 tenga la fracción continua $1 = \langle 1\rangle$).

Proposición 5.4.

Sean $\langle q_0, q_1, \ldots, q_n\rangle$ y $\langle q'_0, q'_1, \ldots, q'_m\rangle$ dos fracciones continuas finitas con $q_0, q'_0 \in \mathbb{Z}$ y $q_k, q'_k \in \mathbb{N}$ para todo $k \geq 1$. Asumamos además que, si $n \neq 0$, el último término es $q_n > 1$; y, si $m \neq 0$, es $q'_m > 1$. En estas condiciones, si

$$\langle q_0, q_1, \ldots, q_n\rangle = \langle q'_0, q'_1, \ldots, q'_m\rangle \tag{5·9}$$

entonces $n = m$ y $q_k = q'_k$ para todo k.

Demostración. Podemos suponer $n \leq m$. Llamemos ξ al valor común (5·9).

Caso $n = 0$. Entonces, $\xi = q_0 \in \mathbb{Z}$. Si fuera $m > 0$ tendríamos

$$q'_0 < \langle q'_0, q'_1, \ldots, q'_m\rangle = q'_0 + \frac{1}{\langle q'_1, \ldots, q'_m\rangle} < q'_0 + 1,$$

luego $\langle q'_0, q'_1, \ldots, q'_m\rangle \in \mathbb{Q} \setminus \mathbb{Z}$. Por tanto, es imposible que (5·9) sea cierto. Así que debe ser $m = 0$ y, trivialmente, $\xi = q_0 = q'_0$.

Caso $n \geq 1$. Se tiene

$$\xi = q_0 + \frac{1}{\langle q_1, \ldots, q_n\rangle} = q'_0 + \frac{1}{\langle q'_1, \ldots, q'_m\rangle}.$$

De aquí, $q_0 < \xi < q_0+1$ y $q_0' < \xi < q_0'+1$, luego $q_0 = q_0' = \lfloor \xi \rfloor$. Además, despejando,

$$\langle q_1, \ldots, q_n \rangle = (\xi - q_0)^{-1} = (\xi - \lfloor \xi \rfloor)^{-1}$$

y, análogamente, $\langle q_1', \ldots, q_m' \rangle = (\xi - \lfloor \xi \rfloor)^{-1}$. Por tanto, a partir de (5.9) hemos deducido que también deberá cumplirse

$$\langle q_1, \ldots, q_n \rangle = \langle q_1', \ldots, q_m' \rangle,$$

luego hemos reducido el problema a otro en el que las fracciones tienen un término menos cada una. En estas circunstancias, es claro que la prueba se sigue sin más que hacer inducción sobre n. □

5.5. Convergencia de fracciones continuas simples

Pensemos ahora cómo evaluar una fracción continua. Si es finita, como en (5.8), basta efectuar las operaciones indicadas, comenzando por la parte de abajo. Pero supongamos que estamos interesados en aproximar el valor de una fracción continua infinita $\langle q_0, q_1, q_2, \ldots \rangle$, esto es, el límite de sus correspondientes fracciones continuas finitas (sus «convergentes»). Así, una vez calculada $\langle q_0, q_1, q_2, \ldots, q_n \rangle$, podríamos estar interesados en evaluar $\langle q_0, q_1, q_2, \ldots, q_n, q_{n+1} \rangle$. Pronto descubrimos con amargura que todo el trabajo hecho para calcular la primera no se puede aprovechar para la segunda: hay que evaluar de nuevo todo el castillo de fracciones comenzando por abajo. Esto no es muy deseable (y no es lo que ocurre, por ejemplo, cuando se define el valor de una serie $\sum_{k=1}^{\infty} c_k$ como el límite de sus sumas parciales $S_n = \sum_{k=1}^{n} c_k$, ya que ahora $S_{n+1} = S_n + c_{n+1}$), así que una buena idea es buscar un método que permita evaluar las sucesivas $\langle q_0, q_1, q_2, \ldots, q_n \rangle$ de una manera más eficiente.

Con este objetivo, introduzcamos un poco de notación. Dada $\langle q_0, q_1, q_2, \ldots \rangle$ una fracción continua simple (finita o infinita), denotaremos[9]

$$\begin{aligned} &P_{-1} = 1, \quad P_0 = q_0, \quad P_1 = q_0 q_1 + 1, \quad P_k = q_k P_{k-1} + P_{k-2}, \ k \geq 1, \\ &Q_{-1} = 0, \quad Q_0 = 1, \quad Q_1 = q_1, \qquad\quad Q_k = q_k Q_{k-1} + Q_{k-2}, \ k \geq 1. \end{aligned} \tag{5.10}$$

Así, $\{P_k\}_{k=0}^{\infty}$ y $\{Q_k\}_{k=0}^{\infty}$ es un par regular de sucesiones recurrentes (quizá un par de sucesiones «finitas»), tal como las definimos en la sección 5.2; su correspondiente

[9] Aunque es redundante dar tres pares de valores iniciales (el tercer par se deduce de los otros dos), lo estamos haciendo así pues, dependiendo de las circunstancias, es más cómodo usar las condiciones de subíndices -1 y 0 o las de subíndices 0 y 1.

Δ es $\Delta = P_1Q_0 - P_0Q_1 = 1$. Nótese que los subíndices de (5.10) están desplazados en uno respecto a los de la sección 5.2, lo cual puede tener influencia al interpretar algunos resultados, como la proposición 5.1 y el teorema 5.2. Por ejemplo, utilizaremos muy a menudo la propiedad (b) de la proposición 5.1, que en las presentes circunstancias queda

$$P_kQ_{k-1} - P_{k-1}Q_k = (-1)^{k-1}, \qquad k \geq 0. \tag{5.11}$$

Tal como habíamos convenido, vamos a seguir suponiendo que los q_k cumplen $q_0 \in \mathbb{Z}$ y $q_k \in \mathbb{N}$ para $k \geq 1$. Pero, por conveniencia, vamos ahora a permitir que, en las fracciones continuas finitas, aparezca un último término que puede ser un número real positivo.

Lema 5.5.

Sea $\langle q_0, q_1, q_2, \ldots\rangle$ una fracción continua simple (finita o infinita). Entonces, para cada $\xi \in \mathbb{R}$, $\xi > 0$, se cumple

$$\langle q_0, q_1, q_2, \ldots, q_n, \xi\rangle = \frac{\xi P_n + P_{n-1}}{\xi Q_n + Q_{n-1}}.$$

Demostración. Lo probaremos por inducción sobre n. Para $n = 0$,

$$\langle q_0, \xi\rangle = q_0 + \frac{1}{\xi} = \frac{\xi q_0 + 1}{\xi} = \frac{\xi P_0 + P_{-1}}{\xi Q_0 + Q_{-1}},$$

tal como buscábamos. Si lo suponemos cierto para n, para $n+1$ tenemos (obsérvese que el paso fundamental consiste en aplicar la fórmula de recurrencia definida en (5.10))

$$\begin{aligned}\langle q_0, q_1, q_2, \ldots, q_n, q_{n+1}, \xi\rangle &= \langle q_0, q_1, q_2, \ldots, q_n, q_{n+1} + 1/\xi\rangle \\ &= \frac{(q_{n+1} + 1/\xi)P_n + P_{n-1}}{(q_{n+1} + 1/\xi)Q_n + Q_{n-1}} = \frac{(q_{n+1}P_n + P_{n-1}) + (1/\xi)P_n}{(q_{n+1}Q_n + Q_{n-1}) + (1/\xi)Q_n} \\ &= \frac{P_{n+1} + (1/\xi)P_n}{Q_{n+1} + (1/\xi)Q_n} = \frac{\xi P_{n+1} + P_n}{\xi Q_{n+1} + Q_n}.\end{aligned}$$

□

Este lema es muy importante. Como consecuencia suya, nos va a aparecer el tan deseado procedimiento para calcular los sucesivos P_n/Q_n de una manera cómoda; en concreto, ¡mediante la relación de recurrencia (5.10)! De momento, no vamos a usar el lema con $\xi \in \mathbb{R}$ y positivo (aunque, a decir verdad, eso ya ha resultado útil en la demostración del propio lema), sino sólo con $\xi = q_n \in \mathbb{N}$. Se utilizará con toda la generalidad posible un poco más adelante, en la demostración del teorema 5.9.

Teorema 5.6.
Sea $\langle q_0, q_1, q_2, \ldots\rangle$ una fracción continua simple (finita o infinita). Entonces, su convergente n-ésimo es

$$\langle q_0, q_1, q_2, \ldots, q_n\rangle = \frac{P_n}{Q_n},$$

donde los P_n y Q_n vienen dados por la relación de recurrencia (5.10). Además, $\operatorname{mcd}(P_n, Q_n) = 1$, luego el cociente P_n/Q_n siempre proporciona —directamente— una fracción irreducible.

Demostración. Por el lema (con $n-1$ en lugar de n y $\xi = q_n$), resulta

$$\langle q_0, q_1, q_2, \ldots, q_n\rangle = \frac{q_n P_{n-1} + P_{n-2}}{q_n Q_{n-1} + Q_{n-2}} = \frac{P_n}{Q_n}.$$

Finalmente, que P_n/Q_n es irreducible se debe a la relación $P_nQ_{n-1} - P_{n-1}Q_n = (-1)^{n-1}$ (véase (5.11)). Con esto, si $d \mid P_n$ y $d \mid Q_n$, deberá cumplirse $d \mid (-1)^{n-1}$, luego $\operatorname{mcd}(P_n, Q_n) = 1$. □

Así pues, hemos probado que los convergentes de una fracción continua siempre forman un par regular de sucesiones recurrentes con $\Delta = 1$.

En primer lugar, esto, en combinación con (5.10), nos proporciona una manera alternativa de calcular los sucesivos convergentes de una fracción continua. Así, para evaluar $\langle q_0, q_1, q_2, \ldots, q_n, q_{n+1}\rangle$ ya se puede aprovechar lo previamente calculado para $\langle q_0, q_1, q_2, \ldots, q_n\rangle$. En concreto, de acuerdo a la relación de recurrencia (5.10), cada convergente se puede hallar conociendo los dos anteriores.

Además, podremos aplicar todos los resultados que ya conocemos sobre pares regulares de sucesiones recurrentes. En particular, la convergencia de sus cocientes. Así, sin más que aplicar el teorema 5.2 tenemos lo siguiente:

Corolario 5.7.
Toda fracción continua simple $\langle q_0, q_1, q_2, \ldots\rangle$ con $q_0 \in \mathbb{Z}$ y $q_k \in \mathbb{N}$ para todo $k \geq 1$ converge a un número real. Además, los convergentes de orden par crecen, y los de orden impar decrecen, según este esquema:

$$\frac{P_0}{Q_0} < \frac{P_2}{Q_2} < \frac{P_4}{Q_4} < \cdots \longrightarrow \langle q_0, q_1, q_2, \ldots\rangle \longleftarrow \cdots < \frac{P_5}{Q_5} < \frac{P_3}{Q_3} < \frac{P_1}{Q_1}.$$

Es de destacar que el corolario anterior no sólo muestra la convergencia, sino que sirve para proporcionar una cota del error que se comete al aproximar la fracción

continua por uno sus convergentes. En efecto, si tomamos P_n/Q_n como aproximación de $\langle q_0, q_1, q_2, \ldots\rangle$, el corolario muestra que

$$\left|\langle q_0, q_1, q_2, \ldots\rangle - \frac{P_n}{Q_n}\right| < \left|\frac{P_{n-1}}{Q_{n-1}} - \frac{P_n}{Q_n}\right|,$$

ya que $\langle q_0, q_1, q_2, \ldots\rangle$ es un valor intermedio entre P_{n-1}/Q_{n-1} y P_n/Q_n.

5.5.1. Unicidad del desarrollo en fracción continua infinita

Ya hemos visto que, con una pequeña excepción, el desarrollo en fracción continua finita es único. El siguiente resultado nos proporciona la unicidad del desarrollo en fracción continua infinita:

Proposición 5.8.

Si dos fracciones continuas simples infinitas convergen al mismo número real, entonces son la misma. Es decir, tienen los mismos cocientes incompletos y en el mismo orden.

Demostración. Sea $\xi = \langle q_0, q_1, q_2, \ldots\rangle$. Veamos que los q_k están determinados, unívocamente, por ξ. Los dos primeros cocientes incompletos son q_0 y $q_0 + \frac{1}{q_1}$, así que

$$q_0 < \xi < q_0 + \frac{1}{q_1}.$$

Por tanto $q_0 = \lfloor \xi \rfloor$ y q_0 está determinado por ξ. Además,

$$\xi = q_0 + \frac{1}{\langle q_1, q_2, \ldots\rangle}$$

luego

$$\langle q_1, q_2, \ldots\rangle = (\xi - q_0)^{-1}.$$

Así que basta repetir el argumento (inducción) y $q_1, q_2, \ldots$ están todos determinados por ξ. □

Así como los racionales se pueden expresar por medio de una fracción continua finita, los irracionales también se pueden desarrollar en fracción continua, pero infinita. Obsérvese que, al hacer esa afirmación, estamos diciendo implícitamente que el desarrollo converge al número que lo origina, y no a otro.[10] Y es que lo que sucede es lo siguiente:

[10] Este detalle es importante, y no es una obviedad: piénsese en el desarrollo de Taylor de $f(x) = e^{-1/x^2}$ (con $f(0) = 0$) en torno al origen, que es $\sum_{n=0}^{\infty} 0 \cdot x^n$, ya que $f^{(n)}(0) = 0$ para todo n.

Teorema 5.9.
Todo número irracional tiene un desarrollo en fracción continua simple infinita (que converge al número en cuestión), y este desarrollo es único.

Demostración. Por la proposición anterior, la unicidad ya está probada, luego sólo tenemos que mostrar la existencia. Para ello, dado $\xi \in \mathbb{R} \setminus \mathbb{Q}$, vamos a construir una fracción continua infinita $\langle q_0, q_1, q_2, \ldots \rangle$ que —según demostraremos— converge a ξ.

Tomamos $q_0 = \lfloor \xi \rfloor$; entonces, $\xi = q_0 + \frac{1}{\xi_1}$, con $\xi_1 > 1$, $\xi_1 \notin \mathbb{Q}$. Sea ahora $q_1 = \lfloor \xi_1 \rfloor$, de donde $\xi_1 = q_1 + \frac{1}{\xi_2}$, con $\xi_2 > 1$, $\xi_2 \notin \mathbb{Q}$. Así, hemos conseguido

$$\xi = q_0 + \cfrac{1}{q_1 + \cfrac{1}{\xi_2}}.$$

Siguiendo del mismo modo, obtenemos $q_0 \in \mathbb{Z}$, $q_1, q_2, \ldots \in \mathbb{N}$ tales que, para todo $n \in \mathbb{N}$,

$$\xi = q_0 + \cfrac{1}{q_1 + \cfrac{1}{q_2 + \cfrac{1}{\ddots \cfrac{}{q_n + \cfrac{1}{\xi_{n+1}}}}}} = \langle q_0, q_1, q_2, \ldots, q_n, \xi_{n+1} \rangle,$$

donde $\xi_{n+1} > 1$, $\xi_{n+1} \notin \mathbb{Q}$. Veamos que $\xi = \langle q_0, q_1, q_2, \ldots \rangle$.

Para esto, bastaría probar que, para cada $n \in \mathbb{N}$,

$$\frac{P_{2n}}{Q_{2n}} < \xi < \frac{P_{2n-1}}{Q_{2n-1}}, \tag{5.12}$$

ya que, según sabemos por el corolario 5.7,

$$\frac{P_{2n}}{Q_{2n}} \longrightarrow \langle q_0, q_1, q_2, \ldots \rangle \longleftarrow \frac{P_{2n-1}}{Q_{2n-1}}$$

Este desarrollo es válido para todo $x \in \mathbb{R}$, pero —salvo en $x = 0$— no converge a e^{-1/x^2}, sino a la función nula. (Seguramente el lector recordará que, al contrario de lo que nos acaba de suceder con una función de variable real, esta patología ya no ocurre con funciones de variable compleja. Y es que, como función compleja, $f(z) = e^{-1/z^2}$ con $f(0) = 0$ ni siquiera es continua en $z = 0$.)

cuando $n \to \infty$, y de aquí se deduciría $\xi = \langle q_0, q_1, q_2, \ldots \rangle$. Así pues, comprobemos (5.12). Por el lema 5.5,

$$\xi = \langle q_0, q_1, q_2, \ldots, q_{2n}, \xi_{2n+1} \rangle = \frac{\xi_{2n+1} P_{2n} + P_{2n-1}}{\xi_{2n+1} Q_{2n} + Q_{2n-1}}.$$

Además,

$$\begin{aligned} \frac{P_{2n}}{Q_{2n}} < \xi \;&\Leftrightarrow\; \frac{P_{2n}}{Q_{2n}} < \frac{\xi_{2n+1} P_{2n} + P_{2n-1}}{\xi_{2n+1} Q_{2n} + Q_{2n-1}} \\ &\Leftrightarrow\; P_{2n}(\xi_{2n+1} Q_{2n} + Q_{2n-1}) < Q_{2n}(\xi_{2n+1} P_{2n} + P_{2n-1}) \\ &\Leftrightarrow\; P_{2n} Q_{2n-1} - P_{2n-1} Q_{2n} < 0 \;\Leftrightarrow\; \frac{P_{2n}}{Q_{2n}} < \frac{P_{2n-1}}{Q_{2n-1}}; \end{aligned}$$

y esta desigualdad es cierta (por ejemplo, obsérvese de nuevo el corolario 5.7). Análogamente se demuestra $\xi < \frac{P_{2n-1}}{Q_{2n-1}}$, de donde se sigue (5.12), tal como buscábamos. □

5.5.2. Un ejemplo numérico concreto

Es interesante fijarse en que la demostración del teorema 5.9 proporciona un procedimiento constructivo para hallar el desarrollo en fracción continua simple. A continuación, lo aplicaremos a un ejemplo:

Ejemplo. Hallar el desarrollo en fracción continua simple de[11]

$$\pi = 3.141\,592\,653\,589\,793\,238\,462\ldots$$

Solución. Los sucesivos q_k (cocientes incompletos), ξ_k (con la notación de la demostración del teorema anterior), y P_k/Q_k (convergentes), quedan como sigue:

$$\begin{aligned} q_0 = \lfloor \pi \rfloor, \quad \pi = 3 + \frac{1}{\xi_1}, \quad \xi_1 = 7.062\,513\,305\,931\ldots, \\ \langle 3 \rangle = 3; \\ q_1 = \lfloor \xi_1 \rfloor, \quad \xi_1 = 7 + \frac{1}{\xi_2}, \quad \xi_2 = 15.996\,594\,406\,685\ldots, \\ \langle 3, 7 \rangle = 3 + \frac{1}{7} = \frac{22}{7} = 3.142\,857\,142\,857\ldots; \end{aligned}$$

[11] No queremos confundir al lector. Lo que sigue *no* es un método para calcular π, pues vamos a partir de una aproximación numérica previamente conocida (y establecida por otros métodos). Pero sí que sirve para encontrar buenas aproximaciones racionales de π con denominadores relativamente pequeños.

$$q_2 = \lfloor \xi_2 \rfloor, \quad \xi_2 = 15 + \frac{1}{\xi_3}, \qquad \xi_3 = 1.003\,417\,231\,013\ldots,$$
$$\langle 3, 7, 15\rangle = 3 + \frac{1}{7 + 1/15} = \frac{333}{106} = 3.141\,509\,433\,962\ldots;$$
$$q_3 = \lfloor \xi_3 \rfloor, \quad \xi_3 = 1 + \frac{1}{\xi_4}, \qquad \xi_4 = 292.634\,591\,014\,395\ldots,$$
$$\langle 3, 7, 15, 1\rangle = \frac{355}{113} = 3.141\,592\,920\,353\ldots;$$
$$q_4 = \lfloor \xi_4 \rfloor, \quad \xi_4 = 292 + \frac{1}{\xi_5}, \qquad \xi_5 = 1.575\,818\,089\,628\ldots,$$
$$\langle 3, 7, 15, 1, 292\rangle = \frac{103\,993}{33\,102} = 3.141\,592\,653\,011\ldots$$

En particular, obsérvese que las fracciones $\frac{22}{7}$ y $\frac{355}{113}$ son —teniendo en cuenta el tamaño de los denominadores— muy buenas aproximaciones de π (esto se debe a que los cocientes incompletos que siguen a esos convergentes, 15 y 292, son números bastante grandes: los motivos, y en concreto la fórmula (5.13), se explican en la nota 5.16, en la página 309). La fracción $\frac{22}{7}$ ya había sido usada por Arquímedes (siglo III a.C.) quien, aproximando una circunferencia mediante polígonos inscritos y circunscritos, demostró que $\frac{223}{71} < \pi < \frac{22}{7}$.[12] El cociente $\frac{355}{113}$ como aproximación de π fue descubierto por el matemático chino Zu Chongzhi alrededor del año 500 (también conocía $\frac{22}{7}$, y sabía que era un valor menos preciso); en el ejercicio 5.13 se analiza el error $|\pi - \frac{355}{113}|$ con más detalle. Acabemos mencionando que, con un poco más de esfuerzo, hubiéramos podido llegar a

$$\pi = \langle 3, 7, 15, 1, 292, 1, 1, 1, 2, 1, 3, 1, 14, 2, 1, 1, 2, 2, 2, 2, 1, 84, 2, 1, 1, 15, \ldots\rangle. \qquad \Diamond$$

Nota 5.10. El desarrollo de π en fracción continua simple no sigue ninguna pauta regular. Cuando esto ocurre, las fracciones continuas generalizadas pueden resultar útiles. Por ejemplo, Lord Brouncker encontró en 1655 la expresión

$$\frac{4}{\pi} = 1 + \cfrac{1^2}{2 + \cfrac{3^2}{2 + \cfrac{5^2}{2 + \cdots}}},$$

[12] Pueden verse los argumentos de Arquímedes en su obra *Medida del círculo*, reeditada —en español— y comentada en [7].



Arquímedes de Siracusa (287 a.C. – 212 a.C.)



Zu Chongzhi (429–501)

que es equivalente a la que proponemos en el ejercicio 5.43 (otra demostración sencilla se puede encontrar en [64]). Existen muchas otras fracciones continuas de números relacionados con π. Como muestra, citemos únicamente un sorprendente desarrollo descubierto por el matemático hindú Srinivasa Ramanujan, y del que se ha llegado a afirmar que es «la fórmula más bella de Ramanujan»:

$$\sqrt{\frac{\pi e}{2}} = \left(1 + \frac{1}{1\cdot 3} + \frac{1}{1\cdot 3\cdot 5} + \frac{1}{1\cdot 3\cdot 5\cdot 7} + \cdots\right) + \cfrac{1}{1+\cfrac{1}{1+\cfrac{2}{1+\cfrac{3}{1+\cdots}}}};$$

ni la fracción continua ni la serie se pueden expresar por separado en función de π y e, ¡pero su suma proporciona una combinación asombrosa! Si el lector se siente fascinado con las historias y las fórmulas relacionadas con π, no puede perderse [105], un excelente librito que une la divulgación con el rigor matemático (en particular, la citada fórmula de Ramanujan aparece en [105, p. 63]); también recomendamos el artículo [48]. Pero, en cuanto a lo que ahora nos concierne, olvidémonos de estas expresiones, y volvamos a las fracciones continuas simples. ◁

5.5.3. ¿Las fracciones continuas proporcionan buenas aproximaciones?

Un número irracional siempre se puede aproximar por racionales de manera cada vez más precisa; por ejemplo, bastaría tomar las fracciones que se obtienen al truncar el desarrollo decimal del irracional: cuantos más dígitos significativos tomemos,

mayor será la precisión alcanzada. Pero, claro, de este modo obtenemos fracciones con denominadores cada vez mayores (y también numeradores mayores, si los consideramos en valor absoluto). De lo que se trata es de conseguir un compromiso entre el tamaño del denominador y el error que se comete al aproximar; sin precisar lo que queremos decir con ello, esto da lugar a aproximaciones «mejores» y «peores». Las fracciones que provienen de truncar el desarrollo decimal son fáciles de obtener, pero no hay ninguna razón para pensar que puedan ser particularmente «buenas» (y, de hecho, no lo son en general).

En el tema anterior analizábamos detalladamente la aproximación de números irracionales mediante racionales, y comprobábamos que las fracciones de Brocot eran muy útiles para conseguir «buenas» aproximaciones. Concretamente, en la proposición 4.34 veíamos cómo aproximar un irracional ξ por fracciones de Brocot p/q de modo que $|\xi - p/q| < 1/(\sqrt{5}\,q^2)$; y esto se podía hacer infinitas veces, de donde se obtenía el teorema de Hurwitz, que aquí recordamos:

Teorema 5.11. Hurwitz, 1891.
Si $\xi \in \mathbb{R} \setminus \mathbb{Q}$, existen infinitas fracciones $\frac{p}{q}$ tales que

$$0 < \left|\xi - \frac{p}{q}\right| < \frac{1}{\sqrt{5}\,q^2}.$$

Esto era lo mejor a lo que se podía aspirar en general, pues era fácil ver que, si tomábamos $M > \sqrt{5}$, para la razón áurea $\phi = (1+\sqrt{5})/2$ sólo puede existir una cantidad finita de racionales p/q tales que $|\phi - p/q| < 1/(Mq^2)$. De hecho, si $M \geq (3+\sqrt{5})/2$, no existe ningún racional p/q que cumpla $|\phi - p/q| < 1/(Mq^2)$. (Véase el teorema 4.37.)

¿Se puede conseguir una demostración alternativa del teorema de Hurwitz usando fracciones continuas? Es decir, si $\{P_n/Q_n\}_{n=0}^{\infty}$ son los convergentes del desarrollo en fracción continua simple de ξ, ¿se satisfacen aproximaciones del tipo $|\xi - P_n/Q_n| < 1/(\sqrt{5}\,Q_n^2)$?, ¿son ciertas para todos los n, o para muchos de ellos? Lo abordaremos en la siguiente sección (véase el apartado (g) del teorema 5.12); es más, los resultados que probaremos y el proceso seguido recordarán —sin duda— a los correspondientes del tema anterior obtenidos mediante fracciones de Brocot. De momento, veamos qué ocurre con el ejemplo que estábamos tratando: el desarrollo en fracción continua simple de π.

Fijándonos en los cinco convergentes que hemos calculado en el ejemplo, resulta

$$1^2 \cdot |\pi - 3| = 0.141\,592\ldots,$$

$$7^2 \cdot \left|\pi - \frac{22}{7}\right| = 0.061\,959\ldots,$$
$$106^2 \cdot \left|\pi - \frac{333}{106}\right| = 0.935\,055\ldots,$$
$$113^2 \cdot \left|\pi - \frac{355}{113}\right| = 0.003\,406\ldots,$$
$$33\,102^2 \cdot \left|\pi - \frac{103\,993}{33\,102}\right| = 0.633\,219\ldots.$$

Como $1/\sqrt{5} = 0.447\,213\ldots$, el primero, segundo y cuarto convergentes cumplen la desigualdad $|\pi - P_n/Q_n| < 1/(\sqrt{5}\,Q_n^2)$. Entre otras muchas propiedades de las fracciones continuas, en la siguiente sección veremos que eso no es casual.

5.6. Aproximación por racionales mediante desarrollos en fracción continua

Sea $\xi = \langle q_0, q_1, q_2, \ldots\rangle$ una fracción continua simple infinita. Como hasta ahora, para cada $n \geq 0$, denotemos

$$\langle q_0, q_1, q_2, \ldots, q_n\rangle = \frac{P_n}{Q_n},$$

que siempre es una fracción irreducible. Sabemos que $\{P_n\}_{n=0}^{\infty}$ y $\{Q_n\}_{n=0}^{\infty}$ forman un par regular de sucesiones recurrentes, según la fórmula (5.10). Observemos que $1 \leq Q_0 \leq Q_1$ (puede darse la igualdad) y que $Q_n < Q_{n+1}$ para todo $n \geq 1$.

El siguiente resultado prueba las excelentes propiedades de los convergentes $\{P_n/Q_n\}_{n=0}^{\infty}$ como aproximaciones racionales de ξ. En particular, obsérvense (c), (e) y (g); esta última es, además, una nueva demostración del teorema de Hurwitz (la existencia de infinitas fracciones es clara al ser $Q_n < Q_{n+1}$ para todo $n \geq 1$).

Teorema 5.12.
Sea $\xi = \langle q_0, q_1, q_2, \ldots\rangle$ una fracción continua simple infinita, y P_n/Q_n sus convergentes. Se cumplen las siguientes propiedades (h/k denota siempre un racional con $k \in \mathbb{N}$):

(a) $$\frac{1}{Q_n Q_{n+2}} \leq \frac{1}{Q_n(Q_n + Q_{n+1})} < \left|\xi - \frac{P_n}{Q_n}\right| < \frac{1}{Q_n Q_{n+1}},$$ *para todo $n \geq 0$.*

(b) $\left|\xi - \dfrac{P_{n+1}}{Q_{n+1}}\right| < \left|\xi - \dfrac{P_n}{Q_n}\right|$; *es más,* $|\xi Q_{n+1} - P_{n+1}| < |\xi Q_n - P_n|$, *para todo* $n \geq 0$.

(c) *Si* $n \geq 1$, $\frac{P_n}{Q_n}$ *es la mejor aproximación de* ξ *entre todos los* $\frac{h}{k}$ *con* $k \leq Q_n$. *Es decir,*

$$n \geq 1, \quad k \leq Q_n, \quad \frac{h}{k} \neq \frac{P_n}{Q_n} \;\Rightarrow\; \left|\xi - \frac{h}{k}\right| > \left|\xi - \frac{P_n}{Q_n}\right|.$$

(d) *De hecho, se cumple esta propiedad*[13] *más fuerte que (c):*

$$n \geq 0, \quad k < Q_{n+1}, \quad \frac{h}{k} \neq \frac{P_n}{Q_n} \;\Rightarrow\; |\xi k - h| > |\xi Q_n - P_n|.$$

(e) *Si* $\frac{h}{k}$ *satisface* $|\xi - \frac{h}{k}| < \frac{1}{2k^2}$ *(esto es, la cota que aparecía en la proposición 4.31 del tema anterior), entonces* $\frac{h}{k} = \frac{P_n}{Q_n}$ *para algún* $n \geq 0$.

(f) *De cada dos convergentes consecutivos de* ξ, *al menos uno cumple*

$$\left|\xi - \frac{P_n}{Q_n}\right| < \frac{1}{2Q_n^2}$$

(salvo, quizás, los dos primeros).

(g) *De cada tres convergentes consecutivos de* ξ, *al menos uno cumple*

$$\left|\xi - \frac{P_n}{Q_n}\right| < \frac{1}{\sqrt{5}\,Q_n^2}$$

(en particular, esto demuestra el teorema de Hurwitz).

Antes de abordar la demostración del teorema, veamos algunas consecuencias inmediatas. De la cota superior que aparece en (a) se deduce que una fracción

[13] Las aproximaciones por racionales a un real ξ tal como en (d) y (c) tienen nombre; ya han aparecido en la sección 4.4 del tema anterior (en relación con las fracciones de Brocot), pero recordémoslo ahora. Se dice que una fracción irreducible a/b es una *aproximación óptima de primera especie* si se verifica $|\xi - h/k| > |\xi - a/b|$ para toda fracción $h/k \neq a/b$ con denominador $k \leq b$. Y es *aproximación óptima de segunda especie* si $|\xi k - h| > |\xi b - a|$ para toda fracción $h/k \neq a/b$ con $k \leq b$. Así pues, los convergentes de las fracciones continuas son aproximaciones óptimas de primera y segunda especie (nótese que, a partir de $n = 1$, la hipótesis $k < Q_{n+1}$ es menos exigente que $k \leq Q_n$). En la nota 4.30 (página 263) ya habíamos probado que toda aproximación óptima de segunda especie lo es también de primera; esto es lo mismo que vamos a hacer implícitamente ahora, cuando deduzcamos (c) a partir de (d).

continua simple infinita siempre proporciona un número irracional. En efecto, dado un racional $\xi = r/s$, es claro que, si $h/k \neq r/s$, se cumplirá

$$\left|\frac{r}{s} - \frac{h}{k}\right| = \frac{|rk - hs|}{sk} \geq \frac{1}{sk}$$

(el numerador $|rk - hs|$ no es nulo, pues $r/s \neq h/k$, luego al menos tiene que valer 1). Entonces, si r/s tuviera un desarrollo en fracción continua infinita $r/s = \langle q_0, q_1, q_2, \ldots\rangle$, tendríamos

$$\frac{1}{sQ_n} \leq \left|\frac{r}{s} - \frac{P_n}{Q_n}\right| < \frac{1}{Q_n Q_{n+1}},$$

lo cual no es posible cuando $Q_{n+1} \geq s$ (y seguro que esto ocurre a partir de un n, pues $Q_n \to \infty$).

En consecuencia, esto proporciona el siguiente resultado:

Corolario 5.13. Criterio de irracionalidad.

Si un número tiene un desarrollo en fracción continua simple infinita, ese número es irracional.

Merece la pena mencionar que en este criterio se basó la demostración original de Euler de la irracionalidad de e. Lo analizaremos en la última sección de este tema, y nos aparecerá de nuevo en la sección 6.2 del tema 6.

Más aún. Habíamos visto que un racional tiene siempre un único desarrollo en fracción continua simple finita (que no acabe en 1); que, recíprocamente, una fracción continua simple finita es siempre un número racional; y que cualquier irracional tiene un único desarrollo en fracción continua simple infinita. Ahora hemos completado la única duda sobre unicidad que nos quedaba: no es posible que un racional tenga un desarrollo finito y otro infinito,[14] pues hemos probado que cualquier fracción continua simple infinita es un número irracional (ya sabíamos que convergía).

Por su interés, destaquemos esto de manera adecuada:

Corolario 5.14. Unicidad del desarrollo en fracción continua simple.

Si dos fracciones continuas simples convergen al mismo número real ξ, ambas son infinitas (y ξ irracional) o ambas finitas (y ξ racional). Si son infinitas, todos sus cocientes incompletos deben coincidir. Si son finitas, las dos tienen la misma longitud y los mismos cocientes incompletos, con la salvedad de $\langle q_0, q_1, q_2, \ldots, q_{n-1}, q_n, 1\rangle = \langle q_0, q_1, q_2, \ldots, q_{n-1}, q_n + 1\rangle$.

[14] Nótese que, en principio, podía haber ocurrido que un mismo número hubiera tenido un desarrollo finito y otro infinito, de manera similar a como muchos números racionales tienen dos representaciones decimales distintas; por ejemplo, $1.73 = 1.729\,999\,999\ldots$

Por fin, vamos ya con la demostración del teorema:

Demostración del teorema 5.12. En todo lo que sigue, emplearemos reiteradamente las propiedades de los pares regulares de sucesiones recurrentes (proposición 5.1 y teorema 5.2, con el subíndice desplazado en 1).

Comencemos con la cota superior para $|\xi - P_n/Q_n|$ de (a). Tenemos

$$\frac{P_0}{Q_0} < \frac{P_2}{Q_2} < \frac{P_4}{Q_4} < \cdots < \xi < \cdots < \frac{P_5}{Q_5} < \frac{P_3}{Q_3} < \frac{P_1}{Q_1}$$

y

$$\frac{P_n}{Q_n} - \frac{P_{n+1}}{Q_{n+1}} = \frac{(-1)^{n+1}}{Q_n Q_{n+1}},$$

ya que $P_{n+1}Q_n - P_nQ_{n+1} = (-1)^n$ (esto ya lo habíamos adaptado a la notación de los convergentes en (5.11)). Por tanto,

$$\left|\xi - \frac{P_n}{Q_n}\right| < \left|\frac{P_n}{Q_n} - \frac{P_{n+1}}{Q_{n+1}}\right| = \frac{1}{Q_n Q_{n+1}}.$$

Analicemos a continuación las cotas inferiores de (a) para $|\xi - P_n/Q_n|$. Denotemos $\xi_n = \langle q_{n+1}, q_{n+2}, \ldots\rangle$, con lo que $\xi = \langle q_0, q_1, q_2, \ldots, q_n, \xi_n\rangle$. Por el lema 5.5 tenemos

$$\xi = \frac{\xi_n P_n + P_{n-1}}{\xi_n Q_n + Q_{n-1}},$$

luego

$$\left|\xi - \frac{P_n}{Q_n}\right| = \left|\frac{\xi_n P_n + P_{n-1}}{\xi_n Q_n + Q_{n-1}} - \frac{P_n}{Q_n}\right| = \frac{|P_{n-1}Q_n - P_nQ_{n-1}|}{(\xi_n Q_n + Q_{n-1})Q_n}.$$

Ahora bien, por una parte, $|P_{n-1}Q_n - P_nQ_{n-1}| = 1$; y, por otra, de la definición se deduce que $\xi_n < q_{n+1} + 1$, luego

$$\begin{aligned}\xi_n Q_n + Q_{n-1} &< Q_n + q_{n+1}Q_n + Q_{n-1} = Q_n + Q_{n+1}\\ &\leq q_{n+2}Q_{n+1} + Q_n = Q_{n+2}.\end{aligned}$$

Por lo tanto, $|\xi - P_n/Q_n| > 1/(Q_n(Q_n + Q_{n+1})) \geq 1/(Q_nQ_{n+2})$.

Para abordar (b), escribamos las desigualdades de (a) en la forma

$$\frac{1}{Q_nQ_{n+2}} < \left|\xi - \frac{P_n}{Q_n}\right| \quad \text{y} \quad \left|\xi - \frac{P_{n+1}}{Q_{n+1}}\right| < \frac{1}{Q_{n+1}Q_{n+2}}.$$

Multiplicando la segunda por Q_{n+1} y la primera por Q_n resulta

$$|\xi Q_{n+1} - P_{n+1}| < \frac{1}{Q_{n+2}} < |\xi Q_n - P_n|,$$

que es la segunda desigualdad de (b). Veamos ahora que, en efecto, esto implica la primera desigualdad de (b):

$$\begin{aligned} |\xi Q_{n+1} - P_{n+1}| < |\xi Q_n - P_n| \;\Rightarrow\; \left|\xi - \frac{P_{n+1}}{Q_{n+1}}\right| &= \frac{1}{Q_{n+1}}\,|\xi Q_{n+1} - P_{n+1}| \\ < \frac{1}{Q_{n+1}}\,|\xi Q_n - P_n| \le \frac{1}{Q_n}\,|\xi Q_n - P_n| &= \left|\xi - \frac{P_n}{Q_n}\right|. \end{aligned}$$

Probemos que (c) se deduce de (d): Si $n \ge 1$, entonces $Q_n < Q_{n+1}$. Luego

$$\begin{aligned} k \le Q_n, \quad \frac{h}{k} \ne \frac{P_n}{Q_n} \;&\Rightarrow\; |\xi k - h| > |\xi Q_n - P_n| \\ &\Rightarrow\; \left|\xi - \frac{h}{k}\right| = \frac{1}{k}\,|\xi k - h| > \frac{1}{Q_n}\,|\xi Q_n - P_n| = \left|\xi - \frac{P_n}{Q_n}\right|. \end{aligned}$$

El caso $n = 0$ de (c) no tiene excesivo interés; además, puede fallar de manera trivial, tal como se muestra en el ejercicio 5.9.

Así pues, analicemos la propiedad (d). Sean $n \ge 0$ y $h/k \ne P_n/Q_n$ con $k < Q_{n+1}$. Planteemos el sistema de ecuaciones

$$\begin{cases} P_n\alpha + P_{n+1}\beta = h, \\ Q_n\alpha + Q_{n+1}\beta = k, \end{cases}$$

cuyas soluciones, teniendo en cuenta que $P_{n+1}Q_n - P_nQ_{n+1} = (-1)^n$, son

$$\begin{aligned} \alpha &= (-1)^n(kP_{n+1} - hQ_{n+1}), \\ \beta &= (-1)^{n+1}(kP_n - hQ_n) \end{aligned}$$

(aplicando la fórmula de Cramer, que da la solución de un sistema lineal de ecuaciones mediante cocientes de determinantes,[15] resulta inmediato).

Puesto que $h/k \ne P_n/Q_n$, debe ser $\beta \ne 0$. Y, como $k < Q_{n+1}$ y P_{n+1}/Q_{n+1} es irreducible, también tenemos $h/k \ne P_{n+1}/Q_{n+1}$, luego $\alpha \ne 0$. Así pues, $\alpha, \beta \in$

[15] Cramer introdujo los determinantes en 1750, precisamente para solucionar los sistemas de ecuaciones lineales.

$\mathbb{Z}\setminus\{0\}$; además, α y β tienen distinto signo, porque lo contrario lleva a contradicción con $k < Q_{n+1}$ y $Q_n\alpha + Q_{n+1}\beta = k$.

También sabemos que los números

$$\xi Q_{n+1} - P_{n+1} = Q_{n+1}\Big(\xi - \frac{P_{n+1}}{Q_{n+1}}\Big)$$

y

$$\xi Q_n - P_n = Q_n\Big(\xi - \frac{P_n}{Q_n}\Big)$$

tienen distinto signo. Por tanto, $\alpha(\xi Q_n - P_n)$ y $\beta(\xi Q_{n+1} - P_{n+1})$ tienen el mismo. Como consecuencia, de la igualdad

$$\begin{aligned}\xi k - h &= \xi(Q_n\alpha + Q_{n+1}\beta) - (P_n\alpha + P_{n+1}\beta)\\ &= \alpha(\xi Q_n - P_n) + \beta(\xi Q_{n+1} - P_{n+1})\end{aligned}$$

se deduce que

$$|\xi k - h| = |\alpha(\xi Q_n - P_n)| + |\beta(\xi Q_{n+1} - P_{n+1})| > |\alpha|\,|\xi Q_n - P_n| \geq |\xi Q_n - P_n|,$$

tal como buscábamos.

Abordemos ahora (e). Para algún $n \geq 0$ tendremos $Q_n \leq k < Q_{n+1}$, ya que $Q_n \to \infty$ cuando $n \to \infty$. Si no fuera $h/k = P_n/Q_n$ entonces, por una parte tendríamos $1 \leq |P_n k - Q_n h|$, y, por otra, $|\xi k - h| > |\xi Q_n - P_n|$ (propiedad (d)). Así, por hipótesis,

$$\left|\xi - \frac{P_n}{Q_n}\right| = \frac{1}{Q_n}\,|\xi Q_n - P_n| < \frac{1}{Q_n}\,|\xi k - h| = \frac{k}{Q_n}\,\left|\xi - \frac{h}{k}\right| < \frac{1}{2kQ_n}.$$

Entonces,

$$\frac{1}{kQ_n} \leq \frac{|P_n k - Q_n h|}{kQ_n} = \left|\frac{P_n}{Q_n} - \frac{h}{k}\right| \leq \left|\frac{P_n}{Q_n} - \xi\right| + \left|\xi - \frac{h}{k}\right| < \frac{1}{2kQ_n} + \frac{1}{2k^2}.$$

Multiplicando por $Q_n k$ se llega a una contradicción: $1 < 1/2 + (1/2)(Q_n/k) \leq 1$. Luego debe ser $h/k = P_n/Q_n$.

Finalmente, las demostraciones de (f) y (g) son análogas a las que hacíamos en el tema anterior mediante fracciones de Brocot. Echemos un vistazo a (f), dejando (g) para el lector (ejercicio 5.10). Si fuera

$$\left|\frac{P_n}{Q_n} - \xi\right| \geq \frac{1}{2Q_n^2} \quad \text{y} \quad \left|\frac{P_{n+1}}{Q_{n+1}} - \xi\right| \geq \frac{1}{2Q_{n+1}^2},$$

entonces, teniendo en cuenta que $P_n/Q_n - \xi$ y $\xi - P_{n+1}/Q_{n+1}$ tienen el mismo signo,

$$\frac{1}{Q_nQ_{n+1}} = \left|\frac{P_n}{Q_n} - \frac{P_{n+1}}{Q_{n+1}}\right| = \left|\frac{P_n}{Q_n} - \xi\right| + \left|\frac{P_{n+1}}{Q_{n+1}} - \xi\right| \geq \frac{1}{2Q_n^2} + \frac{1}{2Q_{n+1}^2}.$$

Multiplicando por $2Q_n^2Q_{n+1}^2$, se llega a una contradicción:

$$2Q_nQ_{n+1} \geq Q_{n+1}^2 + Q_n^2 \;\Rightarrow\; (Q_{n+1} - Q_n)^2 \leq 0 \;\Rightarrow\; Q_n = Q_{n+1},$$

falso si $n \geq 1$. □

Nota 5.15. El teorema también sirve, con ligeras modificaciones, para una fracción continua simple finita $\xi = \langle q_0, q_1, q_2, \ldots, q_N\rangle$, con $N \geq 1$ y $q_N > 1$. En este caso, tenemos:

- El apartado (a) se cumple para todo $n \geq N-2$; para $n = N-1$, la desigualdad de la derecha es una igualdad.
- El apartado (b) se cumple para todo $n \geq N-1$.
- El apartado (c) se cumple para todo $n \geq N$.
- El apartado (d) se cumple para todo $n \geq N-1$; $|\xi Q_N - P_N| = 0$.
- Los apartados (e), (f) y (g) no precisan modificación.

Todo se comprueba siguiendo la demostración del caso infinito. ◁

Nota 5.16. Volviendo al caso de la fracción continua de π (ejemplo de la página 299), estamos ahora en condiciones de analizar por qué las aproximaciones 22/7 y 355/113 son particularmente buenas. El apartado (a) del teorema 5.12 nos dice

$$\frac{1}{Q_n(Q_n + Q_{n+1})} < \left|\xi - \frac{P_n}{Q_n}\right| < \frac{1}{Q_nQ_{n+1}},$$

o equivalentemente, por ser $Q_{n+1} = q_{n+1}Q_n + Q_{n-1}$,

$$\frac{1}{Q_n^2(1 + q_{n+1} + Q_{n-1}/Q_n)} < \left|\xi - \frac{P_n}{Q_n}\right| < \frac{1}{Q_n^2(q_{n+1} + Q_{n-1}/Q_n)}.$$

Esto implica que

$$\frac{1}{(q_{n+1}+2)Q_n^2} < \left|\xi - \frac{P_n}{Q_n}\right| < \frac{1}{q_{n+1}Q_n^2}. \tag{5.13}$$

De aquí es claro que, para $q_0, q_1, \ldots, q_n$ dados (luego Q_n también dado), cuanto mayor sea q_{n+1}, más precisa será la aproximación que P_n/Q_n hace de ξ. En el caso de $\pi = \langle 3, 7, 15, 1, 292, 1, \ldots\rangle$, los convergentes $P_1/Q_1 = 22/7$ y $P_3/Q_3 = 355/113$

llevan aparejados $q_2 = 15$ y $q_4 = 292$, dos números relativamente grandes, lo que hace que esas fracciones sean tan cercanas a π.

Podemos también reflexionar sobre lo que ocurrirá en el caso contrario, es decir, cuando los cocientes incompletos son siempre números pequeños. Es entonces cuando las aproximaciones que se consiguen mediante fracciones continuas son peores. En particular, he aquí el motivo por el que es precisamente la razón áurea $\phi = \frac{1+\sqrt{5}}{2} = \langle 1, 1, 1, 1, \ldots \rangle$ la que sirve para probar que la constante $M = \sqrt{5}$ del teorema de Hurwitz no se puede mejorar, resultado que ya veíamos en el teorema 4.37 (tema 4), y que ahora surge de nuevo sin más que fijarse en que los convergentes de ϕ (que, como sabemos, vienen dados por cocientes de números de Fibonacci) cumplen

$$\lim_n f_n^2 \left| \phi - \frac{f_{n+1}}{f_n} \right| = \frac{1}{\sqrt{5}},$$

que resulta inmediato sin más que aplicar la fórmula de Binet. En seguida lo comprobaremos de nuevo, esta vez mediante un procedimiento distinto aplicable a cualquier fracción continua. ◁

5.6.1. La constante de Lagrange de un irracional

Dado un irracional ξ, en la sección 4.6 nos preguntábamos cuál era la mayor constante $M(\xi)$ tal que la desigualdad

$$\left| \xi - \frac{p}{q} \right| < \frac{1}{(M(\xi) - \varepsilon) q^2}$$

tiene infinitas soluciones racionales p/q para cada $\varepsilon > 0$. A este valor $M(\xi)$ lo llamamos *constante de Lagrange* de ξ. La propiedad (c) del teorema 5.12 nos asegura que las mejores aproximaciones de ξ por fracciones de denominador acotado (es decir, no mayor que cierta constante preestablecida) se consiguen mediante los convergentes del desarrollo en fracción continua de ξ. Es más, la propiedad (g) del mismo teorema prueba que $M(\xi) \geq \sqrt{5}$ para cualquier ξ (esto, que lo hemos comentado reiteradamente, es lo que habitualmente se conoce como teorema de Hurwitz).

Pero —tal como hizo el matemático alemán Oskar Perron en 1921— aún podemos extraer más información de la fracción continua de ξ. Lo vemos a continuación:

Proposición 5.17. Perron, 1921.
Sea un irracional $\xi = \langle q_0, q_1, q_2, \ldots \rangle$. Entonces,

$$M(\xi) = \limsup_{n \to \infty} \big(\langle q_{n+1}, q_{n+2}, \ldots \rangle + \langle 0, q_n, q_{n-1}, \ldots, q_2, q_1 \rangle \big).$$

Demostración. Ya sabemos que las mejores aproximaciones se consiguen con los convergentes de la fracción continua. Analicemos qué pasa con ellos. Si tomamos $\xi_{n+1} = \langle q_{n+1}, q_{n+2}, \ldots\rangle$, tenemos

$$\xi = \langle q_0, q_1, q_2, \ldots, q_n, \xi_{n+1}\rangle = \frac{\xi_{n+1}P_n + P_{n-1}}{\xi_{n+1}Q_n + Q_{n-1}},$$

luego, aplicando (5.11),

$$\xi - \frac{P_n}{Q_n} = \frac{\xi_{n+1}P_n + P_{n-1}}{\xi_{n+1}Q_n + Q_{n-1}} - \frac{P_n}{Q_n} = \frac{P_{n-1}Q_n - P_nQ_{n-1}}{(\xi_{n+1}Q_n + Q_{n-1})Q_n} = \frac{(-1)^n}{\lambda_n Q_n^2}$$

con

$$\lambda_n = \xi_{n+1} + \frac{Q_{n-1}}{Q_n}.$$

Empleando reiteradas veces la fórmula de recurrencia de $\{Q_k\}_{k=1}^{\infty}$, el sumando Q_{n-1}/Q_n va quedando como

$$\frac{Q_{n-1}}{Q_n} = \frac{1}{Q_n/Q_{n-1}} = \frac{1}{q_n + \dfrac{Q_{n-2}}{Q_{n-1}}} = \frac{1}{q_n + \dfrac{1}{q_{n-1} + \dfrac{Q_{n-3}}{Q_{n-2}}}} = \cdots$$

$$= \langle 0, q_n, q_{n-1}, \ldots, q_2, q_1\rangle.$$

De aquí que

$$M(\xi) = \limsup_{n\to\infty} \lambda_n = \limsup_{n\to\infty} \big(\langle q_{n+1}, q_{n+2}, \ldots\rangle + \langle 0, q_n, q_{n-1}, \ldots, q_2, q_1\rangle\big). \qquad \square$$

A modo de ejemplo, obsérvese que, para la razón áurea

$$\phi = \frac{1+\sqrt{5}}{2} = \langle 1, 1, 1\ldots\rangle$$

se obtiene

$$M(\phi) = \limsup_{n\to\infty} \Big(\langle 1, 1, \ldots\rangle + \langle 0, \underbrace{1, 1, \ldots, 1, 1}_{n}\rangle\Big) = \frac{1+\sqrt{5}}{2} + \left(\frac{1+\sqrt{5}}{2} - 1\right) = \sqrt{5},$$

resultado que ya conocíamos (en el apartado (a) del teorema 4.37 veíamos una demostración que no requería tantas herramientas).

Consecuencia de la proposición anterior es este teorema:

Teorema 5.18.
Sean ξ y θ dos irracionales cuyos desarrollos en fracción continua simple son idénticos a partir de un término, esto es,

$$\xi = \langle p_0, p_1, \ldots, p_n, q_1, q_2, \ldots\rangle, \qquad \theta = \langle p'_0, p'_1, \ldots, p'_m, q_1, q_2, \ldots\rangle.$$

Entonces, $M(\xi) = M(\theta)$.

Resulta muy interesante analizar qué números M resultan ser $M = M(\xi)$ para algún irracional ξ; al conjunto de tales M lo llamamos *espectro de Lagrange*. Ya hemos comentado algo sobre esto —que queda fuera del alcance de este libro— en la sección 4.6; a un nivel divulgativo, volvemos a recomendar la lectura de [54, § 10.5] o [22, módulo 6]. Sí que queremos dejar constancia de que, en esto, las fracciones continuas resultan esenciales. Unas líneas más arriba hemos visto cómo la expresión de la proposición 5.17 permitía hallar la constante de Lagrange de la razón áurea. Este mismo cálculo se puede llevar a la práctica para muchos otros irracionales (por ejemplo, véase el ejercicio 5.23), y además se pueden obtener diversas consecuencias teóricas sobre la estructura del espectro de Lagrange.

Siguiendo con lo que veíamos en la sección 4.6, allí llamábamos números *equivalentes* a los relacionados mediante una *transformación modular*; esto es, ξ y θ son equivalentes si existen enteros a, b, c, d, con $|ad - bc| = 1$, tales que

$$\theta = \frac{a\xi + b}{c\xi + d}.$$

Este tipo de fórmulas aparece de manera natural cuando tratamos con desarrollos en fracción continua simple. En efecto, si tomamos $\theta = \langle q_0, q_1, q_2, \ldots, q_n, \xi\rangle$, el lema 5.5 prueba que

$$\theta = \langle q_0, q_1, q_2, \ldots, q_n, \xi\rangle = \frac{\xi P_n + P_{n-1}}{\xi Q_n + Q_{n-1}};$$

y sabemos que se cumple $P_nQ_{n-1} - P_{n-1}Q_n = (-1)^{n-1}$ (véase (5.11)). En particular, si, como es habitual, escribimos $\theta = \langle q_0, q_1, q_2, \ldots, q_n, \theta_{n+1}\rangle$, el número θ es equivalente a su «parte final» θ_{n+1}. Así que resulta razonable pensar que las fracciones continuas también tienen ahora algo importante que aportar.

Y no nos equivocamos, tal como muestra el siguiente teorema debido a Joseph Serret:



Joseph Alfred Serret
(1819–1885)



Rafael Bombelli
(1526–1572)

Teorema 5.19. Serret, 1847.
Una condición necesaria y suficiente para que dos números irracionales sean equivalentes es que sus desarrollos en fracción continua simple sean idénticos a partir de un término.

Aunque no es excesivamente complicada, no incluimos la demostración de este resultado. Si el lector está interesado, puede encontrarla en [54, § 10.5, teorema 5.3]; también resulta instructivo hojear [55, capítulo 6, § 63], donde se proporcionan detalles adicionales. Recomendamos, así mismo, la demostración que aparece en [75, capítulo 1, § 2].

Una de las consecuencias del teorema de Serret (en conjunción con el teorema 5.18) es que dos irracionales equivalentes tienen la misma constante de Lagrange. En realidad, nosotros ya habíamos comentado este hecho en la sección 4.6, y su demostración aparecía propuesta —sin necesidad de usar fracciones continuas— en el ejercicio 4.27.

5.7. Fracciones continuas simples periódicas: irracionales cuadráticos

Comencemos la sección con unos ejemplos. Ya hemos probado que el desarrollo en fracción continua de la razón áurea es $\frac{1+\sqrt{5}}{2} = \langle 1, 1, 1, 1, \ldots \rangle$. Veamos algunos otros. Los orígenes de esto se remontan al matemático italiano Rafael Bombelli y su obra *L'Algebra Opera*, cuyos tres primeros volúmenes fueron publicados en 1572 (había dos volúmenes más previstos, que quedaron sin acabar por la muerte del

autor). Allí, aunque sin identificarlas como tal, se utilizaban fracciones continuas para calcular raíces cuadradas.

Ejemplo (Bombelli, 1572)**.** Desarrollar $\sqrt{2}$ en fracción continua.

Solución. Como $\lfloor\sqrt{2}\rfloor = 1$, tenemos

$$\sqrt{2} = 1 + \frac{1}{\xi} = \langle 1, \xi\rangle,$$

y queremos encontrar ξ. Despejando,

$$\xi = \frac{1}{\sqrt{2}-1} = \sqrt{2}+1 = 2 + \frac{1}{\xi} = \langle 2, \xi\rangle,$$

y, quizás sorprendentemente, ¡nos ha vuelto a aparecer ξ! Entonces, reiterando,

$$\sqrt{2} = \langle 1, \xi\rangle = \langle 1, 2, \xi\rangle = \langle 1, 2, 2, \xi\rangle = \cdots.$$

Es decir, $\sqrt{2} = \langle 1, 2, 2, 2, \ldots\rangle = \langle 1, \overline{2}\rangle$, donde $\overline{\cdot}$ indica una parte periódica que se repite indefinidamente. ◊

Ejemplo. Desarrollar $\sqrt{3}$ en fracción continua.

Solución. Al ser $\lfloor\sqrt{3}\rfloor = 1$, tomamos

$$\sqrt{3} = 1 + \frac{1}{\xi} = \langle 1, \xi\rangle;$$

busquemos una expresión adecuada para ξ, con la esperanza de que, como en el ejemplo anterior, ella misma vuelva a aparecer de nuevo. Despejando,

$$\xi = \frac{1}{\sqrt{3}-1} = \frac{\sqrt{3}+1}{2}.$$

Como $\lfloor(\sqrt{3}+1)/2\rfloor = 1$, escribimos

$$\xi = 1 + \frac{\sqrt{3}-1}{2} = 1 + \frac{1}{2/(\sqrt{3}-1)} = 1 + \frac{1}{1+\sqrt{3}} = 1 + \frac{1}{2+\dfrac{1}{\xi}} = \langle 1, 2, \xi\rangle,$$

de donde

$$\sqrt{3} = \langle 1, \xi\rangle = \langle 1, 1, 2, \xi\rangle = \langle 1, 1, 2, 1, 2, \xi\rangle = \cdots.$$

Por tanto, $\sqrt{3} = \langle 1, 1, 2, 1, 2, 1, 2, \ldots\rangle = \langle 1, \overline{1, 2}\rangle$. ◊

En estos ejemplos, el truco seguido ha sido, siempre, ir desarrollando la parte que hemos denominado «ξ», hasta que ha surgido algo que ya había salido antes (ξ de nuevo). Es decir, hemos obtenido una *fracción continua simple periódica*, esto es, una fracción continua $\langle q_0, q_1, q_2, \ldots\rangle$ para la que existen p y m tales que $q_k = q_{k+p}$ para $k \geq m$. Lo denotamos así:

$$\langle q_0, q_1, q_2, \ldots, q_{m-1}, \overline{q_m, \ldots, q_{m+p-1}}\rangle.$$

Cuando $m = 0$, se dice que la fracción continua es *puramente periódica*.

Podríamos seguir mostrando otros muchos ejemplos (y así se hace en los ejercicios), pero contentémonos con esos; y vamos a especular un poco. Parece que lo que sucede es que, cuando tenemos un número en el que aparece una $\sqrt{\cdot}$, su fracción continua correspondiente es periódica. ¿Ocurrirá esto siempre? ¿Qué pasa si en vez de una $\sqrt{\cdot}$ hay dos? ¿Y si en lugar de $\sqrt{\cdot}$ tenemos $\sqrt[3]{\cdot}$, o una raíz de cualquier otro orden? Si el número es muy «raro», como π (véase la página 299), no parece que la periodicidad vaya a surgir de ninguna parte; ¿es así?

Realmente, la respuesta a estos interrogantes, debida a Joseph Louis Lagrange, es sencilla, como probamos a continuación: vamos a ver que, análogamente a como las fracciones continuas finitas caracterizaban a los números racionales, las fracciones continuas periódicas caracterizan los números irracionales cuadráticos. Recordemos que un irracional cuadrático —en otras palabras, un número algebraico de orden 2— es un número que es raíz de un polinomio irreducible $Ax^2 + Bx + C = 0$ con coeficientes A, B y C enteros ($A \neq 0$). De manera equivalente, son los reales de la forma $(a + b\sqrt{d})/c$ con $a, b \in \mathbb{Z}$, $c, d \in \mathbb{N}$, $b \neq 0$, $d \neq 1$ y d libre de cuadrados.

Teorema 5.20. Lagrange, 1766.
Una fracción continua simple es periódica si y sólo si representa un irracional cuadrático.

Demostración. En primer lugar, probemos que las fracciones continuas puramente periódicas representan números irracionales cuadráticos. Si

$$\theta = \langle \overline{q_0, q_1, q_2, \ldots, q_{m-1}}\rangle$$

es una de tales fracciones continuas, podemos escribir $\theta = \langle q_0, q_1, q_2, \ldots, q_{m-1}, \theta\rangle$. Entonces, según el lema 5.5 tenemos

$$\theta = \frac{\theta P_{m-1} + P_{m-2}}{\theta Q_{m-1} + Q_{m-2}},$$

y es claro que esto da lugar a una ecuación cuadrática en θ. Ahora, para una fracción continua periódica cualquiera

$$\xi = \langle q_0, q_1, q_2, \ldots, q_{m-1}, \overline{q_m, \ldots, q_{m+p-1}} \rangle,$$

denotemos θ a la parte puramente periódica, es decir, $\theta = \langle \overline{q_m, \ldots, q_{m+p-1}} \rangle$, de donde

$$\xi = \langle q_0, q_1, q_2, \ldots, q_{m-1}, \theta \rangle. \tag{5.14}$$

Ya hemos probado que θ es un irracional cuadrático; de (5.14) se sigue que también lo es ξ.

Recíprocamente, sea ξ una raíz de un polinomio irreducible $Ax^2 + Bx + C$ con coeficientes enteros, $A \neq 0$. Como ξ no puede ser un racional, su desarrollo en fracción continua será infinito. Denotemos

$$\xi = \langle q_0, q_1, q_2, \ldots, q_n, \ldots \rangle, \quad \xi_n = \langle q_n, q_{n+1}, q_{n+2}, \ldots \rangle.$$

Por el lema 5.5,

$$\xi = \frac{\xi_n P_{n-1} + P_{n-2}}{\xi_n Q_{n-1} + Q_{n-2}}.$$

Sustituyendo esta expresión en $A\xi^2 + B\xi + C = 0$, se sigue que ξ_n satisface una nueva ecuación con coeficientes enteros, esto es,

$$A_n \xi_n^2 + B_n \xi_n + C_n = 0 \tag{5.15}$$

con

$$\begin{aligned} A_n &= AP_{n-1}^2 + BP_{n-1}Q_{n-1} + CQ_{n-1}^2, \\ B_n &= 2AP_{n-1}P_{n-2} + B(P_{n-1}Q_{n-2} + P_{n-2}Q_{n-1}) + 2CQ_{n-1}Q_{n-2}, \\ C_n &= AP_{n-2}^2 + BP_{n-2}Q_{n-2} + CQ_{n-2}^2. \end{aligned}$$

Estos coeficientes cumplen algunas propiedades que nos resultarán útiles (no merece la pena que nos detengamos para demostrarlas con detalle; seguro que el lector llega a ellas con facilidad si desea intentarlo). Es inmediato que $C_n = A_{n-1}$; más trabajoso —pero elemental— es comprobar[16] que $B_n^2 - 4A_nC_n = B^2 - 4AC$ para todo n (en otras palabras, que todas las ecuaciones (5.15) tienen el mismo discriminante).

[16] Utilizando lenguaje matricial ya no es tan trabajoso: las expresiones que relacionan A_n, B_n, C_n con A, B, C se pueden escribir de la forma

$$\begin{pmatrix} 2A_n & B_n \\ B_n & 2C_n \end{pmatrix} = \begin{pmatrix} P_{n-1} & Q_{n-1} \\ P_{n-2} & Q_{n-2} \end{pmatrix} \begin{pmatrix} 2A & B \\ B & 2C \end{pmatrix} \begin{pmatrix} P_{n-1} & P_{n-2} \\ Q_{n-1} & Q_{n-2} \end{pmatrix}.$$

De $|\xi - P_{n-1}/Q_{n-1}| < 1/Q_{n-1}^2$ (véase (a) del teorema 5.12) se deduce que

$$P_{n-1} = \xi Q_{n-1} + \frac{\varepsilon_n}{Q_{n-1}}$$

para cierto ε_n que cumple $|\varepsilon_n| < 1$. Sustituyendo en la expresión de A_n, y teniendo en cuenta que $A\xi^2 + B\xi + C = 0$, obtenemos

$$\begin{aligned} A_n &= A\Big(\xi Q_{n-1} + \frac{\varepsilon_n}{Q_{n-1}}\Big)^2 + B\Big(\xi Q_{n-1} + \frac{\varepsilon_n}{Q_{n-1}}\Big)Q_{n-1} + CQ_{n-1}^2 \\ &= (A\xi^2 + B\xi + C)Q_{n-1}^2 + 2A\varepsilon_n\xi + A\,\frac{\varepsilon_n^2}{Q_{n-1}^2} + B\varepsilon_n \\ &= 2A\varepsilon_n\xi + A\,\frac{\varepsilon_n^2}{Q_{n-1}^2} + B\varepsilon_n. \end{aligned}$$

En particular, $|A_n| \le 2|A\xi| + |A| + |B|$, luego los $|A_n|$ están acotados. También lo están los $|C_n|$ (ya que $C_n = A_{n-1}$) y los $|B_n|$ (por ser $B_n^2 - 4A_nC_n = B^2 - 4AC$). Por tanto, los A_n, B_n y C_n (que son enteros) sólo pueden tomar un número finito de valores, así que tenemos una cantidad finita de posibles polinomios $A_nx^2 + B_nx + C_n$, y un número finito de posibles valores para los números ξ_n. De aquí que, forzosamente, deben existir dos iguales, $\xi_n = \xi_{n+p}$. Por unicidad del desarrollo en fracción continua infinita, los dos desarrollos $\xi_n = \langle q_n, q_{n+1}, \ldots\rangle$ y $\xi_{n+p} = \langle q_{n+p}, q_{n+p+1}, \ldots\rangle$ deben coincidir, esto es, $q_n = q_{n+p}$, $q_{n+1} = q_{n+p+1}$, $\ldots$; en otras palabras, la fracción continua de ξ es periódica. □

5.7.1. La estructura del periodo

Tal como hemos comentado, entre las fracciones continuas periódicas, podemos distinguir las *puras*, que tienen la forma

$$\xi = \langle \overline{q_0, q_1, \ldots, q_{m-1}, q_m}\rangle,$$

y las que no lo son, que se suelen denominar mixtas. En el primer artículo que publicaba, Évariste Galois —con sólo 17 años de edad— caracterizó cuándo una fracción continua periódica es pura:

Como

$$\begin{vmatrix} P_{n-1} & Q_{n-1} \\ P_{n-2} & Q_{n-2} \end{vmatrix} = \begin{vmatrix} P_{n-1} & P_{n-2} \\ Q_{n-1} & Q_{n-2} \end{vmatrix} = P_{n-1}Q_{n-2} - Q_{n-1}P_{n-2} = (-1)^n$$

y el determinante de un producto de matrices es el producto de sus determinantes, se sigue que $4A_nC_n - B_n^2 = 4AC - B^2$.



Leonhard Euler
(1707–1783)



Évariste Galois
(1811–1832)

Teorema 5.21. Galois, 1828.
La fracción continua de un irracional cuadrático ξ es puramente periódica si y sólo si $\xi > 1$ y $-1 < \xi' < 0$, donde ξ' denota el conjugado de ξ. (Recordemos que el conjugado de $r + s\sqrt{t}$, con $r, s, t \in \mathbb{Q}$ y $t > 0$, es $r - s\sqrt{t}$.)

La demostración, que no incluimos por no alargamos demasiado, puede consultarse en [26, § 7.6], [78, § 7.7] o [88, § 12.4]. Por otra parte, si tenemos $\xi = \langle \overline{q_0, q_1, \ldots, q_{m-1}, q_m} \rangle$, prescindimos del q_0 del periodo, y lo invertimos, no es difícil comprobar que

$$\langle \overline{q_m, q_{m-1}, \ldots, q_1} \rangle = -1/\xi'.$$

Aplicando estos resultados al irracional cuadrático $\lfloor \sqrt{d} \rfloor + \sqrt{d}$, se llega a

Teorema 5.22.
Sea d un entero positivo que no es un cuadrado perfecto. Entonces, la fracción continua simple de $\sqrt{d}$ es

$$\sqrt{d} = \langle q_0, \overline{q_1, q_2, \ldots, q_{m-1}, 2q_0} \rangle;$$

además, la parte $\{q_1, q_2, \ldots, q_{m-1}\}$ es capicúa.

Como ilustración de este resultado, véanse los desarrollos que aparecen en el cuadro 5.1.

En contraste con todo lo que acabamos de ver sobre el desarrollo en fracción continua simple de los números algebraicos de orden 2, resulta sorprendente lo poco que se conoce sobre las fracciones continuas de números algebraicos de orden 3 o mayor. Por ejemplo, no conocemos la fracción continua de uno tan sencillo como $\sqrt[3]{2}$;

$\sqrt{2} = \langle 1, \overline{2} \rangle$	$\sqrt{14} = \langle 3, \overline{1, 2, 1, 6} \rangle$
$\sqrt{3} = \langle 1, \overline{1, 2} \rangle$	$\sqrt{15} = \langle 3, \overline{1, 6} \rangle$
$\sqrt{5} = \langle 2, \overline{4} \rangle$	$\sqrt{17} = \langle 4, \overline{8} \rangle$
$\sqrt{6} = \langle 2, \overline{2, 4} \rangle$	$\sqrt{18} = \langle 4, \overline{4, 8} \rangle$
$\sqrt{7} = \langle 2, \overline{1, 1, 1, 4} \rangle$	$\sqrt{19} = \langle 4, \overline{2, 1, 3, 1, 2, 8} \rangle$
$\sqrt{8} = \langle 2, \overline{1, 4} \rangle$	$\sqrt{20} = \langle 4, \overline{2, 8} \rangle$
$\sqrt{10} = \langle 3, \overline{6} \rangle$	$\sqrt{21} = \langle 4, \overline{1, 1, 2, 1, 1, 8} \rangle$
$\sqrt{11} = \langle 3, \overline{3, 6} \rangle$	$\sqrt{22} = \langle 4, \overline{1, 2, 4, 2, 1, 8} \rangle$
$\sqrt{12} = \langle 3, \overline{2, 6} \rangle$	$\sqrt{23} = \langle 4, \overline{1, 3, 1, 8} \rangle$
$\sqrt{13} = \langle 3, \overline{1, 1, 1, 1, 6} \rangle$	$\sqrt{24} = \langle 4, \overline{1, 8} \rangle$

Cuadro 5.1: Desarrollo en fracción continua de los primeros números de la forma $\sqrt{d}$.

ni tampoco sabemos si sus cocientes incompletos están acotados o no. En realidad, esto no se sabe de ningún número algebraico de orden ≥ 3. Es más, ni siquiera si hay alguno de tales números que tenga cocientes incompletos acotados; y lo mismo con no acotados.

5.8. El desarrollo de e en fracción continua simple

Sobre todo por su interés histórico, sería un lástima despedir el tema sin mostrar el desarrollo en fracción continua de e, un notable resultado debido al inigualable matemático suizo Leonhard Euler. El tipo de expresiones que veremos aquí son las que él mismo utilizó para probar la irracionalidad de e y de e^2 (actualmente, es más habitual probar la irracionalidad de e a partir de su desarrollo en serie, pues resulta un método más corto: véanse, en los temas 4 y 6, los teoremas 4.8 y 6.7).

En realidad, las demostraciones que vamos a dar no son las originales de Euler; con el paso de los años, sus geniales ideas procedentes de su maravillosa intuición han sido tamizadas para hacerlas más rigurosas y asequibles. Afortunadamente, el lector también puede, si así lo desea, acudir a las fuentes originales: le basta con consultar [43]. Además, nos aparecerá esto de nuevo cuando, en la sección 6.2 del tema 6, abordemos de manera explícita la irracionalidad de e. Allí damos más comentarios sobre los métodos que siguió Euler, así como sobre ciertas generalizaciones efectuadas por otros autores.

Comencemos con un lema previo.

Lema 5.23.
Para $c \in \mathbb{R} \setminus \{0, -1, -2, -3, \dots\}$, sea la función

$$f_c(x) = \sum_{k=0}^{\infty} \frac{x^k}{c(c+1)(c+2)\cdots(c+k-1)k!}, \qquad x \in \mathbb{R} \tag{5.16}$$

(cuando $k = 0$, en el denominador hay un producto de 0 factores, así que, como es habitual, asumimos que el término independiente de (5.16) es 1). Se cumple

$$\begin{aligned} \frac{c}{y}\frac{f_c(y^2)}{f_{c+1}(y^2)} &= \left\langle \frac{c}{y}, \frac{c+1}{y}, \dots, \frac{c+n}{y}, \xi_{n+1} \right\rangle \\ &= \frac{c}{y} + \cfrac{1}{\cfrac{c+1}{y} + \cfrac{1}{\ddots \cfrac{c+n}{y} + \cfrac{1}{\xi_{n+1}}}} \end{aligned} \tag{5.17}$$

con $\xi_{n+1} = \dfrac{c+n+1}{y}\dfrac{f_{c+n+1}(y^2)}{f_{c+n+2}(y^2)}$.

Demostración. Antes de comenzar, comentemos que el radio de convergencia de la serie que define $f_c(x)$ es ∞ luego, efectivamente, (5.16) tiene sentido para todo $x \in \mathbb{R}$. Las funciones $f_c(x)$ satisfacen la relación

$$\begin{aligned} f_c(x) - f_{c+1}(x) &= \sum_{k=0}^{\infty} \frac{x^k}{(c+1)(c+2)\cdots(c+k-1)k!}\left(\frac{1}{c} - \frac{1}{c+k}\right) \\ &= \sum_{k=1}^{\infty} \frac{x^k}{c(c+1)(c+2)\cdots(c+k-1)(c+k)(k-1)!} \\ &= \frac{x}{c(c+1)} f_{c+2}(x). \end{aligned}$$

En consecuencia,

$$\frac{f_c(x)}{f_{c+1}(x)} = 1 + \frac{x}{c(c+1)}\frac{f_{c+2}(x)}{f_{c+1}(x)},$$

que, tomando $x = y^2$, podemos escribir como

$$\frac{c}{y}\frac{f_c(y^2)}{f_{c+1}(y^2)} = \frac{c}{y} + \frac{y}{c+1}\frac{f_{c+2}(y^2)}{f_{c+1}(y^2)} = \frac{c}{y} + \cfrac{1}{\cfrac{c+1}{y}\cfrac{f_{c+1}(y^2)}{f_{c+2}(y^2)}} = \frac{c}{y} + \frac{1}{\xi_1}.$$

Reiterando se llega a (5.17). □

Tras el lema, el teorema:

Teorema 5.24.
Para cada $q \in \mathbb{N}$ se cumple

$$\frac{e^{2/q}+1}{e^{2/q}-1} = \langle q, 3q, 5q, \ldots\rangle = q + \cfrac{1}{3q + \cfrac{1}{5q + \cdots}}.$$

Demostración. En el lema, para $c = 1/2$ y $x = y^2$ tenemos

$$\begin{aligned} f_{1/2}(y^2) &= \sum_{k=0}^{\infty} \frac{1}{\frac{1}{2}(\frac{1}{2}+1)(\frac{1}{2}+2)\cdots(\frac{1}{2}+k-1)} \frac{y^{2k}}{k!} \\ &= \sum_{k=0}^{\infty} \frac{2^k}{1\cdot 3\cdot 5\cdots(2k-1)} \frac{2^k y^{2k}}{2\cdot 4\cdot 6\cdots(2k)} \\ &= \sum_{k=0}^{\infty} \frac{(2y)^{2k}}{(2k)!} = \cosh(2y). \end{aligned}$$

De manera análoga,

$$\begin{aligned} 2y f_{3/2}(y^2) &= \sum_{k=0}^{\infty} \frac{1}{\frac{3}{2}(\frac{3}{2}+1)(\frac{3}{2}+2)\cdots(\frac{3}{2}+k-1)} \frac{2y^{2k+1}}{k!} \\ &= \sum_{k=0}^{\infty} \frac{(2y)^{2k+1}}{(2k+1)!} = \operatorname{senh}(2y). \end{aligned}$$

En consecuencia, si recordamos que $\cosh(t) = (e^t + e^{-t})/2$ y $\operatorname{senh}(t) = (e^t - e^{-t})/2$, tenemos

$$\frac{f_{1/2}(y^2)}{2y f_{3/2}(y^2)} = \frac{e^{2y}+e^{-2y}}{e^{2y}-e^{-2y}} = \frac{e^{4y}+1}{e^{4y}-1}.$$

Por otra parte, (5.17) queda

$$\frac{1}{2y}\frac{f_{1/2}(y^2)}{f_{3/2}(y^2)} = \left\langle \frac{1}{2y}, \frac{3}{2y}, \frac{5}{2y}, \ldots, \frac{2n+1}{2y}, \xi_{n+1}\right\rangle.$$

Sea ahora $2y = 1/q$ con $q \in \mathbb{N}$; así, los cocientes incompletos de la fracción continua finita anterior son de la forma $\frac{2k+1}{2y} = (2k+1)q \in \mathbb{N}$; es decir,

$$\frac{e^{2/q}+1}{e^{2/q}-1} = \langle q, 3q, 5q, \ldots, (2n+1)q, \xi_{n+1}\rangle.$$

Comprobemos que, además, $\xi_{n+1} > 1$ para todo n. En efecto,

$$\xi_{n+1} = \frac{c+n+1}{y} \frac{\displaystyle\sum_{k=0}^{\infty} \frac{y^{2k}}{(c+n+1)\cdots(c+n+k)k!}}{\displaystyle\sum_{k=0}^{\infty} \frac{y^{2k}}{(c+n+2)\cdots(c+n+k+1)k!}}$$

$$= \frac{\displaystyle\sum_{k=0}^{\infty} \frac{y^{2k}}{(c+n+2)\cdots(c+n+k)k!}}{\displaystyle\sum_{k=0}^{\infty} \frac{y^{2k}}{(c+n+2)\cdots(c+n+k)k!}\frac{y}{c+n+k+1}} > 1$$

ya que

$$\frac{y}{c+n+k+1} = \frac{1}{(2q)(1/2+n+k+1)} < 1,$$

luego los sumandos del numerador son mayores que los del denominador. Así pues, $\langle q, 3q, 5q, \ldots\rangle$ es el desarrollo fracción continua simple infinita de $(e^{2/q}+1)/(e^{2/q}-1)$, y sabemos que a él debe converger (teorema 5.9), tal como queríamos probar. □

Si en el teorema anterior hacemos $q = 2$, se tiene que

$$\xi = \frac{e+1}{e-1} = \langle 2, 6, 10, \ldots, 4n+2, \ldots\rangle. \tag{5.18}$$

Esta fracción continua simple no es finita, así que —recordemos el criterio de irracionalidad del corolario 5.13— el número ξ no puede ser racional. Por lo tanto, tampoco e. Con este tipo de métodos se pueden obtener aún más consecuencias sobre la naturaleza aritmética de e y otros números relacionados. También es directo deducir de ahí algunos otros desarrollos en fracción continua, como el de $(e-1)/2$ (véase el ejercicio 5.31). Pero no estamos ahora interesados en ello. Ni tampoco nos vamos a contentar con las expresiones que acabamos de encontrar. Queremos hallar, realmente, el desarrollo en fracción continua de e:

Teorema 5.25. Euler, 1737.
El desarrollo de e en fracción continua simple es

$$e = \langle 2, 1, 2, 1, 1, 4, 1, 1, 6, 1, \ldots, 1, 2n, 1, \ldots\rangle.$$

Demostración. Tomemos $e' = \langle 2,1,2,1,1,4,1,1,6,1,\ldots,1,2n,1,\ldots\rangle$, y sean $\{P_n/Q_n\}_{n=0}^{\infty}$ los convergentes de esta fracción continua; queremos probar que $e' = e$. Además, como en (5.18), sea $\xi = \frac{e+1}{e-1} = \langle 2,6,10,\ldots,4n+2,\ldots\rangle$ y $\{R_n/S_n\}_{n=0}^{\infty}$ sus correspondientes convergentes. Probaremos que

$$P_{3n+1} = R_n + S_n, \qquad Q_{3n+1} = R_n - S_n. \tag{5.19}$$

Si damos ese hecho por cierto, tendremos

$$e' = \lim_n \frac{P_{3n+1}}{Q_{3n+1}} = \lim_n \frac{R_n+S_n}{R_n-S_n} = \lim_n \frac{R_n/S_n+1}{R_n/S_n-1} = e,$$

tal como buscábamos. Así, para concluir la demostración, basta probar (5.19); vamos con ello.

Los convergentes $\{R_n/S_n\}_{n=0}^{\infty}$ cumplen

$$\begin{cases} R_0/S_0 = 2/1, \quad R_1/S_1 = 13/6, \quad R_2/S_2 = 132/61, \quad R_3/S_3 = 1861/860, \\ R_n = (2+4n)R_{n-1} + R_{n-2}, \quad S_n = (2+4n)S_{n-1} + S_{n-2}, \quad n \geq 2. \end{cases}$$

Las relaciones de recurrencia en los $\{P_n/Q_n\}_{n=0}^{\infty}$ son más complicadas. Sus primeros términos son

$$\begin{aligned} &P_0/Q_0 = 2/1, \quad P_1/Q_1 = 3/1, \quad P_2/Q_2 = 8/3, \quad P_3/Q_3 = 11/4, \\ &P_4/Q_4 = 19/7, \quad P_5/Q_5 = 87/32, \quad P_6/Q_6 = 106/39, \quad P_7/Q_7 = 193/71. \end{aligned}$$

Además,

$$P_{3n-3} = P_{3n-4} + P_{3n-5}, \tag{5.20a}$$
$$P_{3n-2} = P_{3n-3} + P_{3n-4}, \tag{5.20b}$$
$$P_{3n-1} = 2nP_{3n-2} + P_{3n-3}, \tag{5.20c}$$
$$P_{3n} = P_{3n-1} + P_{3n-2}, \tag{5.20d}$$
$$P_{3n+1} = P_{3n} + P_{3n-1} \tag{5.20e}$$

(y la relación entre los Q_n es la misma). Manipulando las expresiones anteriores según el esquema «(5.20a)−(5.20b)+2·(5.20c)+(5.20d)+(5.20e)» (y análogamente para los Q_n) obtenemos

$$\begin{aligned} P_{3n+1} &= (4n+2)P_{3n-2} + P_{3n-5}, \\ Q_{3n+1} &= (4n+2)Q_{3n-2} + Q_{3n-5}. \end{aligned}$$

Una vez que hemos encontrado fórmulas de recurrencia para $\{R_n\}_{n=0}^{\infty}$, $\{S_n\}_{n=0}^{\infty}$, $\{P_{3n+1}\}_{n=0}^{\infty}$ y $\{Q_{3n+1}\}_{n=0}^{\infty}$, es una mera comprobación por inducción ver que

$$P_{3n+1} + Q_{3n+1} = 2R_n, \qquad P_{3n+1} - Q_{3n+1} = 2S_n.$$

Y esto es equivalente a (5.19). □

El siguiente resultado nos lleva ya rondando un rato. Dejemos de él —por fin— constancia escrita en forma de teorema:

Teorema 5.26. Euler, 1737.
Los números e y e^2 son irracionales.

Demostración. La fracción continua simple e es infinita; por tanto, e es irracional. Además, su fracción continua simple tampoco es periódica, luego —por el teorema de Lagrange— e no es un irracional cuadrático; en particular e^2 es irracional. □

Nota 5.27. A lo largo de toda la sección, hemos sido cuidadosos y no hemos escrito ninguna fracción continua cuyos cocientes incompletos no fueran enteros positivos. Esto hubiera ocurrido, por ejemplo, si en (5.17) hacemos el desarrollo infinito, o si en el teorema 5.24 tomamos $q \notin \mathbb{N}$. El ejercicio 5.36 nos pide que analicemos esa posibilidad. ◁

Una vez llegados hasta aquí, aprovechemos para probar el siguiente resultado sobre aproximación por racionales:

Teorema 5.28.
Para todo racional p/q con $q > 1$, se cumple

$$\left| e - \frac{p}{q} \right| > \frac{c}{q^2 \log(q)}, \tag{5.21}$$

donde c es una constante positiva independiente de p y q.

Demostración. Sin pérdida de generalidad, basta probarlo a partir de un cierto q_0. En efecto, para cada $q > 1$ fijo, seguro que existe $c_q > 0$ tal que $|e - p/q| > c_q/(q^2 \log(q))$ para todo $p \in \mathbb{Z}$. Así, si logramos demostrar (5.21) con una cierta c' válida para todo $q > q_0$, podríamos tomar $c = \min\{c_2, c_3, \ldots, c_{q_0}, c'\}$.

Además, vamos a probar previamente que los convergentes P_n/Q_n de la fracción continua de e verifican

$$Q_n + Q_{n-1} < \frac{3Q_n \log(Q_n)}{\log(2)}.$$

En la fracción continua $e = \langle q_0, q_1, q_2, q_3, \ldots\rangle$, para $k \geq 1$ se cumple $q_{3k-1} = 2k$ y $q_{3k} = q_{3k+1} = 1$. Partiendo de $Q_0 = 1$ y $Q_1 = q_1$, la fórmula de recurrencia de los Q_n dice $Q_{n+1} = q_{n+1}Q_n + Q_{n-1}$; por tanto,

$$Q_n + Q_{n+1} = (1 + q_{n+1})Q_n + Q_{n-1} < (2 + q_{n+1})Q_n.$$

Queremos acotar eso por $\frac{3Q_n \log(Q_n)}{\log(2)}$. Si $q_{n+1} = 1$ (es decir, cuando $n = 3k - 1$ o $n = 3k$), tenemos

$$Q_n + Q_{n+1} < (2 + q_{n+1})Q_n = 3Q_n \leq \frac{3Q_n \log(Q_n)}{\log(2)}$$

(en cuanto $Q_n \geq 2$).

Analicemos ahora el caso $n = 3k-2$. Como $q_{3k-1} = 2k$, la fórmula de recurrencia es $Q_{3k-1} = 2kQ_{3k-2} + Q_{3k-3}$. Por inducción, vamos a probar que $Q_{3k-1} \geq 2^k k!$. Para $k = 1$ queda $P_2/Q_2 = 8/3$, luego $Q_2 = 3 > 2$. Supuesto cierto para $k-1$, para k resulta

$$\begin{aligned} Q_{3k-1} &= 2kQ_{3k-2} + Q_{3k-3} = Q_{3k-1} = 2k(q_{2k-2}Q_{3k-2} + Q_{3k-4}) + Q_{3k-3} \\ &> 2k(q_{2k-2}Q_{3k-2} + 2^{k-1}(k-1)!) + Q_{3k-3} > 2^k k!. \end{aligned}$$

Como consecuencia, de aquí se deduce que $Q_{3k-1} > 2^{k+1}$ cuando $k \geq 2$. Más aún, al ser $Q_n = Q_{3k-2} > Q_{3k-4} > 2^k$, se sigue que $\log(Q_n) > k\log(2)$, luego $k < \frac{\log(Q_n)}{\log(2)}$. Así, $q_{3k-1} = 2k < 2\,\frac{\log(Q_n)}{\log(2)}$, con lo cual

$$\begin{aligned} Q_n + Q_{n+1} &< (2 + q_{n+1})Q_n + Q_{n-1} \\ &< \left(2 + 2\,\frac{\log(Q_n)}{\log(2)}\right) Q_n < \frac{3Q_n \log(Q_n)}{\log(2)}, \end{aligned} \tag{5.22}$$

tal como buscábamos.

Ya estamos en condiciones de abordar (5.21). Primero, supongamos que $p/q = P_n/Q_n$ para algún n. Aplicando el apartado (a) del teorema 5.12 y la desigualdad (5.22), tenemos

$$\left|e - \frac{P_n}{Q_n}\right| > \frac{1}{Q_n(Q_n + Q_{n+1})} > \frac{\log(2)}{3}\,\frac{1}{Q_n^2 \log(Q_n)}; \tag{5.23}$$

es decir, hemos probado (5.21) para los convergentes de la fracción continua. Finalmente, aplicando la idea de que «los P_n/Q_n son la mejor aproximación posible»

(con más rigor, que son aproximaciones óptimas de segunda especie), de aquí podemos deducir el caso general. En efecto, dado una fracción p/q tomemos n tal que $Q_n \leq q < Q_{n+1}$. Usando el apartado (d) del teorema 5.12 y la desigualdad (5.23), obtenemos

$$|eq - p| > |eQ_n - P_n| = Q_n \left| e - \frac{P_n}{Q_n} \right| \geq \frac{\log(2)}{3Q_n \log(Q_n)} \geq \frac{\log(2)}{3q \log(q)}.$$

Por tanto, $|e - p/q| > c/(q^2 \log(q))$ con $c = \log(2)/3$. □

Merece la pena que hagamos algunos comentarios sobre el significado de lo que acabamos de probar. En primer lugar, obsérvese la similitud de (5.21) con varios de los resultados del tema 4; en particular, con lo que aparece en la nota 4.6 o en el teorema de Liouville (teorema 4.10). Además, en el lenguaje de la definición del comienzo de la sección 4.2, del teorema 5.28 se deduce que e no tiene aproximación por racionales de orden $\varphi(q) = 1/(q^2(\log(q))^{1+\varepsilon})$ para ningún $\varepsilon > 0$. En efecto, si la tuviera, para infinitos p/q debería cumplirse

$$\frac{c}{q^2 \log(q)} < \left| e - \frac{p}{q} \right| < \frac{C}{q^2(\log(q))^{1+\varepsilon}};$$

de aquí, $(\log(q))^{\varepsilon} < C/c$ para infinitos q. Y esto es, obviamente, imposible. Este hecho es más fino que gran parte de los que utilizábamos en el tema 4, pero no llega a ser suficiente para probar la trascendencia de e a partir de los criterios que surgen del teorema de Liouville (corolario 4.13) o —incluso— del de Roth (corolario 4.20).

5.9. Ejercicios

5.1. Un hombre acomoda una pareja de conejos recién nacidos en un recinto vallado. A partir del segundo mes de vida, la pareja de conejos engendra una pareja de retoños, macho y hembra, y sólo una, cada mes. La nueva pareja, a su vez, engendra una nueva pareja a partir del segundo mes, y así sucesivamente. Se supone, además, que no muere ningún conejo. ¿Cuántas parejas de conejos habrá al cabo de un año?[17] ¿Y cuántas al cabo de n meses?

5.2. Comprueba que, tal como se afirma en el primer método para encontrar la expresión del término general de la sucesión de Fibonacci, para que el

[17] Realmente, lo único que hizo Fibonacci en relación con la sucesión conocida por su nombre fue —a grandes rasgos— plantear esa pregunta y explicar la solución.

sistema (5.2) tenga solución, el α y β que allí aparecen deben ser distintos (salvo en el caso trivial $\alpha = \beta = 1$).

5.3. Encuentra la expresión que da el término general de la sucesión de Fibonacci mediante un tercer método: por inducción.

5.4. Demuestra que no existe ningún triángulo (no degenerado) cuyos tres lados sean números de Fibonacci distintos.

5.5. Los números de Lucas $\{L_n\}_{n=1}^{\infty}$ se definen de manera recurrente mediante

$$L_1 = 1, \quad L_2 = 3, \qquad \text{y} \qquad L_n = L_{n-1} + L_{n-2}, \quad n \geq 3.$$

Encuentra una fórmula que proporcione el término general L_n de manera explícita.

5.6. Sean $\{P_k\}_{k=1}^{\infty}$ y $\{Q_k\}_{k=1}^{\infty}$ un par regular de sucesiones recurrentes, tal como se definen en la sección 5.2, y sea $\Delta = P_2Q_1 - P_1Q_2$. Comprueba que la hipótesis $\Delta > 0$ no es restrictiva en el siguiente sentido:

(a) Si $\Delta = 0$, existe una constante $\lambda \in \mathbb{R}$ tal que $P_k = \lambda Q_k$ para todo k o $Q_k = \lambda P_k$ para todo k.

(b) Si $\Delta < 0$, basta considerar $\{-P_k\}_{k=1}^{\infty}$ en lugar de $\{P_k\}_{k=1}^{\infty}$.

5.7. Comprueba que, si aplicamos el algoritmo de la sección 5.4 a dos fracciones $r_0/r_1 = r_0'/r_1'$ (una irreducible y otra no), la fracción continua que se obtiene es la misma.

5.8. Demuestra que los convergentes P_n/Q_n de una fracción continua simple $\langle q_0, q_1, q_2, \ldots \rangle$ se pueden obtener mediante el siguiente producto de matrices[18]:

$$\begin{pmatrix} P_n & P_{n-1} \\ Q_n & Q_{n-1} \end{pmatrix} = \begin{pmatrix} q_0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} q_1 & 1 \\ 1 & 0 \end{pmatrix} \cdots \begin{pmatrix} q_n & 1 \\ 1 & 0 \end{pmatrix}.$$

5.9. Comprueba que, cuando $\xi - \lfloor \xi \rfloor \geq 1/2$, el apartado (c) del teorema 5.12 falla para $n = 0$. Y que, sin embargo, esto no proporciona un contraejemplo para el apartado (d) con $n = 0$.

5.10. Completa la demostración del apartado (g) del teorema 5.12.

[18] En realidad, muchos de los resultados y demostraciones de este tema se podían haber formulado en términos de matrices y determinantes; sirviéndonos de sus propiedades, podríamos haber acortado algunas pruebas (en la línea de lo que hemos hecho en la nota al pie número 16, dentro de la demostración del teorema 5.20).

5.11. Da sentido a la siguiente expresión para la razón áurea, y pruébala:

$$\frac{1+\sqrt{5}}{2} = \sqrt{1+\sqrt{1+\sqrt{1+\sqrt{1+\cdots}}}}.$$

5.12. Da sentido y evalúa la siguiente expresión:

$$\sqrt{2+\sqrt{2+\sqrt{2+\sqrt{2+\cdots}}}}.$$

5.13. A partir del desarrollo en fracción continua $\pi = \langle 3, 7, 15, 1, 292, 1, \ldots\rangle$, deduce (sin recurrir al valor decimal $\pi = 3.141\,592\,653\ldots$) que

$$\left|\pi - \frac{355}{113}\right| < \frac{1}{10^6}.$$

5.14. Sean dos fracciones continuas simples

$$r = \langle q_0, q_1, \ldots, q_{j-1}, q_j, q_{j+1}, \ldots\rangle, \qquad s = \langle q_0, q_1, \ldots, q_{j-1}, q'_j, q'_{j+1}, \ldots\rangle,$$

con $q'_j > q_j$ (por supuesto, estamos asumiendo $q_0 \in \mathbb{Z}$ y $q_k \in \mathbb{N}$ para todo $k \geq 1$, y lo mismo para las q'_k). Analiza si $r > s$, $r < s$, o qué ocurre.

5.15. Caracteriza las desigualdades entre fracciones continuas simples, esto es, da una regla que permita identificar si, dadas dos fracciones continuas $r = \langle q_0, q_1, q_2, \ldots\rangle$ y $s = \langle q'_0, q'_1, q'_2, \ldots\rangle$ (como es habitual, $q_0, q'_0 \in \mathbb{Z}$ y $q_k, q'_k \in \mathbb{N}$ para todo $k \geq 1$), se verifica $r = s$, $r < s$ o $r > s$. Ten en cuenta que tanto r como s pueden ser finitas o infinitas (pero no necesariamente ambas del mismo tipo).

5.16. Supuesto conocido el desarrollo en fracción continua simple de un número real ξ, encuentra el de $-\xi$.

5.17. Encuentra el número cuyo desarrollo en fracción continua es $\langle 1, 2, 3, \overline{4, 5}\rangle$.

5.18. Continúa con los ejemplos de desarrollos en fracción continua de $\sqrt{2}$ y $\sqrt{3}$ mostrados en la sección 5.7, hasta llegar al de $\sqrt{10}$ en el cuadro 5.1 (página 319).

5.19. Desarrolla $\sqrt{N^2+1}$, $N \in \mathbb{N}$, en fracción continua.

5.20. Desarrolla $\sqrt{N^2-1}$, $N \in \mathbb{N} \setminus \{1\}$, en fracción continua.

5.21. Desarrolla $\sqrt{N^2+2}$, $N \in \mathbb{N}$, en fracción continua.

5.22. Desarrolla $\sqrt{N^2-2}$, $N \in \mathbb{N} \setminus \{1,2\}$, en fracción continua.

5.23. Calcula la constante de Lagrange de $\sqrt{2}$.

5.24. Sea $\xi = a+b\sqrt{c}$, con a, b, c números racionales positivos, $2a \in \mathbb{N}$ y $cb^2-a^2 = 1$. Encuentra el desarrollo en fracción continua simple de ξ.

5.25. Halla el polinomio de segundo grado del que $\langle 0, 1, 2, \overline{3}\rangle$ es raíz.

5.26. Demuestra la siguiente relación entre fracciones continuas y series alternadas:

$$\langle q_0, q_1, q_2, \ldots, q_n\rangle = q_0 + \frac{1}{Q_1} - \frac{1}{Q_2Q_1} + \cdots + (-1)^{n-1}\frac{1}{Q_nQ_{n-1}}$$

(dicho con más precisión, entre los convergentes P_n/Q_n de la fracción continua y las sumas parciales de la serie; y, por tanto, entre sus correspondientes límites).

5.27. Dado un número natural $a > 1$, sea la fracción continua

$$\xi = \langle 0, a, a^{2!}, \ldots, a^{n!}, \ldots\rangle$$

y, como es habitual, usemos P_n/Q_n para denotar sus convergentes. Prueba que, para cada $n \geq 1$, se tiene

$$\left|\xi - \frac{P_n}{Q_n}\right| < \frac{1}{Q_n^{n/2}}.$$

Como consecuencia, utiliza el teorema de Liouville o alguno de sus corolarios (sección 4.2 en el tema 4) para demostrar que ξ es trascendente.

5.28. Sea $\varphi\colon \mathbb{N} \to (0,\infty)$ una función arbitraria. Prueba que existe un número ξ tal que la desigualdad

$$\left|\xi - \frac{p}{q}\right| < \varphi(q)$$

tiene infinitas soluciones racionales p/q (así, siguiendo el lenguaje de la sección 4.2 del tema anterior, podemos decir que el orden de aproximación por racionales de ξ es $\varphi(q)$).

Sugerencia. Define ξ mediante su fracción continua.

5.29. Dado $d \in \mathbb{N}$, sean α, β las dos raíces de $x^2 = dx+1$, con α positiva. Encuentra la fracción continua simple de α, así como la relación de recurrencia que

satisface la sucesión de los denominadores de sus convergentes, y demuestra que estos denominadores están dados por

$$Q_n = \frac{\alpha^{n+1} - \beta^{n+1}}{\alpha - \beta} \quad \text{para todo } n \in \mathbb{N} \cup \{0\}.$$

Comprueba que la sucesión de Fibonacci $Q_n = Q_{n-1} + Q_{n-2}$, $Q_{-1} = 0$, $Q_0 = 1$, es un caso particular, para el cual se cumple

$$Q_n = \frac{1}{\sqrt{5}} \left(\left(\frac{1+\sqrt{5}}{2} \right)^{n+1} - \left(\frac{1-\sqrt{5}}{2} \right)^{n+1} \right).$$

5.30. Sea $\langle q_0, q_1, q_2, \ldots \rangle$, con $q_0 \in \mathbb{Z}$ y $q_k \in \mathbb{N}$ para todo $k \geq 1$, una fracción continua simple, y $\{Q_n\}_{n=0}^{\infty}$ la sucesión de los denominadores de sus convergentes. En estas circunstancias:

(a) Deduce que

$$Q_n \geq \left(\frac{1+\sqrt{5}}{2} \right)^{n-1}.$$

(b) Si, además, $\sup_{n\in\mathbb{N}} |q_n| = A < \infty$, prueba que

$$Q_n \leq \left(\frac{A + \sqrt{A^2+4}}{2} \right)^{n}.$$

Sugerencia. Resuelve primero el ejercicio anterior, y fíjate en las relaciones de recurrencia de los denominadores de los convergentes, así como en sus expresiones explícitas.

5.31. Halla el desarrollo en fracción continua simple de $(e-1)/2$.

5.32. Usando el desarrollo de $(e-1)/2$ del ejercicio anterior (en lugar del desarrollo de e del teorema 5.25), simplifica la demostración del teorema 5.28.

5.33. En este ejercicio analizamos fracciones continuas simples cuyos cocientes incompletos no satisfacen la condición $q_0 \in \mathbb{Z}$, $q_k \in \mathbb{N}$ $(k = 1, 2, \ldots)$ que hemos asumido habitualmente. En concreto, dados $q_0 \in \mathbb{R}$, y $q_1, q_2, \ldots$ reales positivos, ¿cuál deberá ser la definición de $\langle q_0, q_1, q_2, \ldots \rangle$ en este caso? Comprueba que se sigue verificando la fórmula de recurrencia de los P_n y Q_n, y que la fracción continua converge si y sólo si $Q_n Q_{n-1} \to \infty$ cuando $n \to \infty$.

5.34. En la situación de ejercicio anterior, demuestra que los denominadores de los convergentes P_n/Q_n satisfacen las siguientes desigualdades:

(a) $Q_n \le (1+q_1)(1+q_2)\cdots(1+q_n)$,
(b) $Q_{2n-1} \ge q_1 + q_3 + \cdots + q_{2n-1}$,
(c) $Q_{2n} \ge 1 + q_1(q_2 + q_4 + \cdots + q_{2n})$,

todas ellas para $n = 1, 2, 3, \ldots$.

5.35. Continuando con lo planteado en el ejercicio 5.33, prueba que la fracción continua $\langle q_0, q_1, q_2, \ldots\rangle$ converge si y sólo si $\sum_{k=1}^{\infty} q_k = \infty$.
Sugerencia. Apóyate en los resultados de los dos ejercicios anteriores.

5.36. Generaliza el teorema 5.24 al caso $q \in \mathbb{R}$, $q > 0$.

5.37. Demuestra que, si no se exige que los q_k sean enteros (pero sí positivos para $k \ge 1$), las fracciones continuas de la forma

$$\langle q_0, q_1, q_2, q_3, \ldots\rangle = q_0 + \frac{1}{q_1+}\,\frac{1}{q_2+}\,\frac{1}{q_3+}\cdots$$

describen un caso muy general, ya que (aquí, los p_k denotan números reales positivos)

$$q_0 + \frac{p_1}{q_1+}\,\frac{p_2}{q_2+}\,\frac{p_3}{q_3+}\cdots = q_0 + \frac{\lambda_1 p_1}{\lambda_1 q_1+}\,\frac{\lambda_1\lambda_2 p_2}{\lambda_2 q_2+}\,\frac{\lambda_2\lambda_3 p_3}{\lambda_3 q_3+}\cdots,$$

donde los λ_k son números arbitrarios no nulos. En particular, muestra cómo hay que elegir los λ_k de modo que $\lambda_{k-1}\lambda_k p_k = 1$ para todo k.

5.38. Sean $q_0 \in \mathbb{Z}$ y $q_k, p_k \in \mathbb{N}$, $k = 1, 2, \ldots$. En este ejercicio generalizamos la definición de fracción continua simple a las no simples, esto es, al caso

$$q_0 + \frac{p_1}{q_1+}\,\frac{p_2}{q_2+}\,\frac{p_3}{q_3+}\cdots = q_0 + \cfrac{p_1}{q_1 + \cfrac{p_2}{q_2 + \cfrac{p_3}{q_3 + \cdots}}},$$

que se suelen denominar fracciones continuas generalizadas. Da la definición de fracción continua en esa situación, y comprueba que la sucesión P_n/Q_n dada por las recurrencias

$$\begin{aligned} P_n &= q_n P_{n-1} + p_n P_{n-2}, & P_{-1} &= 1, & P_0 &= q_0,\\ Q_n &= q_n Q_{n-1} + p_n Q_{n-2}, & Q_{-1} &= 0, & Q_0 &= 1, \end{aligned}$$

es la sucesión de sus convergentes n-ésimos.

5.39. Siguiendo con las hipótesis y la notación del ejercicio 5.38, demuestra la fórmula

$$\frac{P_{n+1}}{Q_{n+1}} - \frac{P_n}{Q_n} = (-1)^n \frac{p_1 p_2 \cdots p_{n+1}}{Q_n Q_{n+1}}, \quad n \in \mathbb{N}.$$

5.40. Sean dos enteros positivos m y r, con $0 < r < (m+1)^2 - m^2 = 2m+1$. Prueba que $\sqrt{m^2+r}$ tiene el siguiente desarrollo en fracción continua generalizada periódica[19]:

$$\sqrt{m^2+r} = m + \cfrac{r}{2m + \cfrac{r}{2m + \cfrac{r}{2m + \cdots}}}.$$

5.41. Este ejercicio y el siguiente abordan la relación entre fracciones continuas generalizadas y series. En la situación del ejercicio 5.38, demuestra que los convergentes P_n/Q_n son las sumas parciales n-ésimas de la serie

$$q_0 + \frac{p_1}{Q_1} - \frac{p_1 p_2}{Q_1 Q_2} + \frac{p_1 p_2 p_3}{Q_2 Q_3} - \frac{p_1 p_2 p_3 p_4}{Q_3 Q_4} + \cdots.$$

5.42. Dada $\{c_n\}_{n=1}^{\infty} \subset \mathbb{N}$ una sucesión estrictamente creciente, demuestra que

$$\sum_{k=1}^{\infty} \frac{(-1)^{k+1}}{c_k} = \cfrac{1}{c_1 + \cfrac{c_1^2}{(c_2 - c_1) + \cfrac{c_2^2}{(c_3 - c_2) + \cfrac{c_3^2}{(c_4 - c_3) + \cdots}}}}.$$

5.43. Apoyándote en el ejercicio anterior, deduce las expresiones

$$\log(2) = \cfrac{1}{1 + \cfrac{1}{1 + \cfrac{4}{1 + \cfrac{9}{1 + \cdots}}}}, \qquad \frac{\pi}{4} = \cfrac{1}{1 + \cfrac{1}{2 + \cfrac{9}{2 + \cfrac{25}{2 + \cdots}}}}.$$

5.44. Comprueba, mediante un ejemplo, que un número racional puede tener un desarrollo infinito en fracción continua generalizada.

[19] Dada su gran velocidad de convergencia, los desarrollos de este tipo proporcionan una manera fácil de aproximar rápidamente el valor numérico de una raíz cuadrada.

Tema 6
Irracionalidad y trascendencia

6.1. Algunos resultados sobre irracionalidad
6.2. Irracionalidad de e
6.3. Irracionalidad de π
6.4. Los números trascendentes existen
6.5. Trascendencia de e
6.6. Polinomios simétricos
6.7. Trascendencia de π
6.8. Extensiones y problemas abiertos
6.9. Polinomios de Bernoulli y $\zeta(2k)$
6.10. Ejercicios

En este tema probaremos la irracionalidad y trascendencia de algunos números. En particular, de e y de π. Algunos de los resultados que aquí mostramos son redundantes. Por ejemplo, la trascendencia de e o de π implica que son números irracionales; pero las demostraciones de irracionalidad son —además de hermosas e interesantes por sí mismas— mucho más sencillas que las de trascendencia, por eso hemos incluido ambas.

En lo que sigue veremos una mezcla de técnicas y demostraciones diversas. Algunas casi inmediatas para nuestros ojos (el paradigma es el de la irracionalidad de $\sqrt{2}$). Otras, bastante más complicadas (con la trascendencia de π como ejemplo). Históricamente, los primeros ejemplos de números irracionales datan de la época griega, pero la existencia de los números trascendentes no fue confirmada hasta 1844, cuando se descubrió el primero. Por otra parte, merece la pena destacar la multitud de ideas felices aquí involucradas; es admirable ver cómo puntos de partida aparentemente distantes conducen al resultado buscado. El genio de grandes matemáticos está implícito. No podemos dejar de citar que las figuras de Pitágoras, Leonhard Euler y Johann Heinrich Lambert —en cuanto a irracionalidad se refiere—, y Joseph Liouville, Charles Hermite y Georg Cantor —en los estudios sobre trascendencia— han sido claves en todo este asunto. No siempre las demostraciones que incluyamos serán similares a las originales; a veces, nuevas y maravillosas ideas de matemáticos posteriores han dado lugar a pruebas más cortas o más elegantes, y éstas son las que habitualmente hemos usado. Por otra parte, y pese a todos los logros ya conseguidos, éste no es un viaje finalizado: aún queda mucho trabajo que hacer en la investigación de la naturaleza aritmética de los números.

Aunque quizás con diferentes enfoques, gran parte de lo que aquí tratamos puede también consultarse en [8, 77, 78, 91]. En particular, [77, 91] contienen abundantes notas históricas y referencias a los trabajos originales; [8] no tiene muchas demostraciones, pero su lectura es muy agradable. Otro precioso libro, [23], muestra numerosos ejemplos de números trascendentes, proporcionando una visión general sobre el tema. Un poco de todo (incluyendo muchas citas históricas) se puede encontrar en [81], también muy ameno. Más avanzados son las monografías [9, 92, 24], y pueden ser útiles a quien quiera profundizar en el tema.

Algunos de los resultados que exponemos ya nos han aparecido en los dos temas anteriores; y es que la aproximación por racionales en general, y las fracciones continuas en particular, son muy útiles a la hora de abordar la irracionalidad o la trascendencia. En aras de lograr temas autocontenidos, varios de estos resultados los hemos incluido de nuevo. Por completitud, recordemos los conceptos de número irracional y número trascendente. Un número es *irracional* cuando no es racional, esto es, cuando no puede escribirse de la forma h/k, con $h \in \mathbb{Z}$ y $k \in \mathbb{N}$. Y un número



Tartaglia (Niccolò Fontana) (1499–1557)



Girolamo Cardano (1501–1576)



Carl Friedrich Gauss (1777–1855)

es *trascendente* cuando no es *algebraico*; ¿qué son, pues, los números algebraicos? Se denominan algebraicos a los números que son raíz de un polinomio no nulo con coeficientes en $\mathbb{Z}$ (o en $\mathbb{Q}$, da igual). Con más detalle, si α es raíz de un polinomio $f(x) \in \mathbb{Q}[x]$ no nulo de grado n, y no es raíz de ningún otro polinomio en $\mathbb{Q}[x]$ de grado menor, decimos que α es algebraico de orden n. Queremos hacer notar al lector que, obviamente, todos los números racionales son algebraicos ($r \in \mathbb{Q}$ es raíz del polinomio $x - r \in \mathbb{Q}[x]$), pero lo contrario no es cierto: los números racionales son precisamente los algebraicos de orden 1.

Concluyamos esta introducción mencionado que, en todo lo que sigue, usaremos frecuentemente el denominado *teorema fundamental del álgebra*; se conoce por este nombre al hecho de que cada polinomio de grado $n \geq 1$ tiene exactamente n raíces complejas, si contamos cada una tantas veces como su multiplicidad. Éste es un resultado importante, y su demostración no es elemental (sí que lo es una vez que se ha conseguido probar que tiene al menos una raíz). Varios matemáticos creyeron haberlo demostrado antes de que el genial Carl Friedrich Gauss lo abordara. Proporcionó una prueba en 1799 y otras dos en 1816; pero la de 1799 tenía cierta falta de rigor que él mismo corrigió en 1849. En 1806, el matemático aficionado Jean Robert Argand dio una prueba (esencialmente rigurosa) que incluía la posibilidad de que los coeficientes del polinomio fuesen números complejos.

Actualmente, existen demostraciones muy sencillas una vez que se ha llegado al nivel matemático suficiente (en particular, las que utilizan análisis complejo), y el teorema fundamental del álgebra es siempre parte integrante de alguno de los cursos de los estudios de matemáticas. Así que podemos dar por hecho que el lector ya conoce su demostración, que nosotros no incluiremos.

Nota 6.1. Un problema diferente —y que aquí sólo vamos a comentar ahora, pero que ya no desarrollaremos con más profundidad— es qué números se pueden expre-



Niels Henrik Abel
(1802–1829)



Évariste Galois
(1811–1832)

sar por medio de radicales. Esto está relacionado con el hecho de que el número en cuestión sea o no algebraico, pero no son conceptos coincidentes. Los números que se expresan con combinaciones de radicales (suma, resta, multiplicación y extracción de raíces, partiendo de números enteros o racionales), como $\sqrt[3]{2+\sqrt{13}}-\sqrt[7]{5}$, son algebraicos. Podría ser razonable pensar si también ocurre lo contrario. Por ejemplo, todos sabemos que las soluciones de la ecuación $ax^2+bx+c=0$ son $(-b\pm\sqrt{b^2-4ac})/(2a)$. Y también las soluciones de la ecuaciones de grado 3 y 4 se pueden expresar por medio de radicales, tal como probaron, en el siglo XVI, los matemáticos italianos Scipione del Ferro, Niccolò Fontana (más conocido por su apodo «Tartaglia»), Girolamo Cardano y Ludovico Ferrari. En todo caso, la primera solución publicada para las ecuaciones de tercer grado (y también el primer cálculo explícito con números complejos) se encuentra en el influyente libro *Ars Magna* de Cardano, que vio la luz en 1545. La rocambolesca historia de tales descubrimientos —comparable a una novela de intriga— se puede leer en [38, capítulo 6].

Por supuesto, los esfuerzos para encontrar soluciones por radicales de la ecuación de grado quinto fueron enormes. Hasta que, en 1824, el noruego Niels Abel conmovió al mundo matemático demostrando que las ecuaciones de grado 5 o superior no se podían resolver, en general, por medio de radicales. La búsqueda había estado condenada al fracaso desde el primer momento. En 1832, el francés Évariste Galois perfeccionó las ideas de Abel y logró aclarar cuándo una ecuación podía ser resuelta por radicales y cuándo no. Tanto Abel como Galois fueron dos genios que murieron muy jóvenes; según la leyenda (que en parte es ficticia), Galois ni siquiera tuvo tiempo de escribir con detalle sus descubrimientos, sino que los redactó apresuradamente la noche antes de morir en un duelo. ◁

6.1. Algunos resultados sobre irracionalidad

Antes de abordar resultados más complicados, recordemos un hecho elemental, que sirve para establecer la racionalidad o irracionalidad de un número si conocemos su expresión decimal completa (o, análogamente, su expresión en cualquier base). El número

$$d_0 + \sum_{j=1}^{\infty} \frac{d_j}{10^j} = d_0.d_1d_2d_3d_4d_5\ldots$$

(donde $d_0 \in \mathbb{Z}$ y $d_j \in \{0,1,2,3,4,5,6,7,8,9\}$ para cada $j \geq 1$) es racional si y sólo si su desarrollo decimal es finito o periódico. Por supuesto, lo mismo ocurre si, en lugar de expresar el número en base 10, usamos cualquier otra base.

En realidad, y aunque este criterio de irracionalidad basado en la representación decimal sea una equivalencia, no es tan útil como parece a primera vista, pues sólo se puede aplicar para números que nos han sido dados con tal expresión. Por ejemplo, sirve para determinar que el número de Champernowne (también conocido como número de Mahler)

$$M = 0.12345678910111213141516171819202 1\ldots$$

es irracional; en efecto, M tiene cadenas de ceros consecutivos arbitrariamente largas, luego su desarrollo decimal no es periódico.[1] Pero no aporta nada si queremos analizar, por ejemplo, la irracionalidad de $\sqrt{2}$, pues no tenemos su desarrollo decimal completo.

Que $\sqrt{2}$ es irracional es conocido desde la antigüedad. En concreto, aparece demostrado —a principios del siglo III a.C.— en los Elementos de Euclides. También lo sabían Pitágoras y su escuela en el siglo VI a.C.; pero la historia cuenta —y muchos lo dan por cierto— que los pitagóricos intentaron mantenerlo oculto, pues la existencia de números irracionales les creaba un gran conflicto lógico en su concepción de las matemáticas (por ejemplo, la diagonal del cuadrado unidad resultaba ser un segmento inconmensurable con los racionales). La demostración sólo requiere

[1] En 1933, David Champernowne —un estudiante inglés, que más adelante se dedicaría a la economía— analizó este número y probó que es *normal* en base 10; a grandes rasgos, esto significa que, si troceamos su desarrollo decimal en cadenas de k dígitos, cualquier cadena concreta de k dígitos aparece con proporción 10^{-k}. En 1946, Kurt Mahler —alemán de origen, abandonó su país al llegar los nazis al poder, y más tarde se nacionalizó británico— llevó a cabo una tarea mucho más difícil: demostrar la trascendencia de M (algo que aquí no abordaremos). Una prueba alternativa a partir del teorema de Thue-Siegel-Roth (que nosotros veremos —también sin demostración— al final de la sección 6.4, y que ya hemos comentado —con bastante más detalle— en el tema 4) puede consultarse en [23, § 1.6 y § 1.7].

del teorema de factorización única en $\mathbb{N}$, y —vista desde la perspectiva actual— es muy sencilla.

Supongamos que $\sqrt{2} = h/k$, fracción irreducible. Elevando al cuadrado obtenemos $h^2 = 2k^2$, luego h par. Tomemos $h = 2m$, con lo cual $4m^2 = 2k^2$. Simplificando, $k^2 = 2m^2$, luego k par. Así pues, hemos probado que tanto h como k son pares. Pero esto no es posible, pues partíamos de una fracción h/k irreducible. Hemos llegado a un absurdo, que proviene de haber supuesto que $\sqrt{2}$ es racional; luego $\sqrt{2}$ debe ser irracional.

Argumentos similares permiten probar la irracionalidad de $\sqrt{d}$, siempre que d no sea un cuadrado perfecto. También se podrían usar con raíces cúbicas, etc. Pero, ¿cómo demostrar, por ejemplo, que $\sqrt{2} + \sqrt{3}$ es irracional? Para ello, el siguiente resultado viene en nuestra ayuda.

Teorema 6.2. Teorema de las raíces racionales.
Supongamos que un número racional h/k (fracción irreducible) no nulo es raíz de un polinomio con coeficientes enteros

$$a_n x^n + a_{n-1} x^{n-1} + \cdots + a_2 x^2 + a_1 x + a_0, \quad a_n \neq 0.$$

Entonces, h divide a a_0 y k divide a a_n.

Demostración. El racional h/k es raíz del polinomio, es decir,

$$a_n \frac{h^n}{k^n} + a_{n-1} \frac{h^{n-1}}{k^{n-1}} + \cdots + a_2 \frac{h^2}{k^2} + a_1 \frac{h}{k} + a_0 = 0.$$

Multiplicando por k^n y despejando el primer sumando obtenemos

$$a_n h^n = -(a_{n-1} h^{n-1} + \cdots + a_2 h^2 k^{n-3} + a_1 h k^{n-2} + a_0 k^{n-1}) k.$$

Como h (luego también h^n) y k no tienen factores comunes, k divide a a_n. Si, en vez de despejar el primer sumando, despejamos el último, podemos poner

$$a_0 k^n = -(a_n h^{n-1} + a_{n-1} h^{n-2} k + \cdots + a_2 h k^{n-2} + a_1 k^{n-1}) h;$$

de aquí, h divide a a_0. □

Con esto, es ya muy fácil analizar si algunos números son o no irracionales.

Ejemplo. Comprobar que $\sqrt{2} + \sqrt{3}$ es irracional.



David Gawen Champernowne (1912–2000)



Kurt Mahler (1903–1988)



Pafnuty Lvovich Chebyshev (1821–1894)

SOLUCIÓN. Tomemos $x = \sqrt{2}+\sqrt{3}$; elevando al cuadrado, $x^2 = 2+3+2\sqrt{6}$, luego $x^2 - 5 = 2\sqrt{6}$. Elevando de nuevo al cuadrado, $x^4 - 10x^2 + 25 = 24$. Así pues, $\sqrt{2}+\sqrt{3}$ es una raíz del polinomio $x^4 - 10x^2 + 1 = 0$. Por el teorema anterior, las únicas posibles raíces racionales de ese polinomio son 1 o -1. Así que, si $\sqrt{2}+\sqrt{3}$ fuera racional, debería ser una de ellas. Como ninguna de las dos posibilidades es cierta, $\sqrt{2}+\sqrt{3}$ es irracional. ◊

Dos consecuencias directas del teorema de las raíces racionales son los siguientes corolarios, cuya demostración se propone como ejercicio (véanse los ejercicios 6.1 y 6.2).

Corolario 6.3.

Supóngase que un racional r es raíz de un polinomio mónico con coeficientes enteros. Entonces, r debe ser un número entero. Además, si $r \neq 0$, r debe dividir al término independiente del polinomio.

Corolario 6.4.

Sea a un entero positivo. Entonces, $\sqrt[n]{a}$ es racional si y sólo si a es la potencia n-ésima de un entero positivo.

6.1.1. Expresiones trigonométricas

Probaremos ahora algunas otras consecuencias del teorema 6.2; esta vez, no tan inmediatas como las que acabamos de enunciar. Con este objetivo, vamos previamente a introducir los polinomios de Chebyshev. Aunque aquí sólo los usaremos en relación con lo que nos interesa, no podemos dejar de citar que tales polinomios aparecen en multitud de campos de la matemática, y son útiles para resolver una

gran cantidad de problemas. Es muy probable que el lector ya se haya topado con ellos alguna vez.

Los *polinomios de Chebyshev* $\{T_n(x)\}_{n=0}^{\infty}$ se definen mediante la relación de recurrencia

$$\begin{cases} T_0(x) = 1, \quad T_1(x) = x, \\ T_{n+1}(x) = 2xT_n(x) - T_{n-1}(x), \quad n \geq 1. \end{cases} \tag{6.1}$$

Es claro que $T_n(x)$ es un polinomio de grado n, con coeficientes enteros, y su coeficiente director es 2^{n-1} (salvo para $n = 0$). Pero su principal propiedad —que sirve, además, como definición alternativa—, es la siguiente.[2] Para $x \in [-1, 1]$,

$$T_n(x) = \cos(n\theta), \quad \text{con } x = \cos(\theta) \tag{6.2}$$

(en particular, de aquí se deduce que $\cos(n \arc\cos(x))$ es, siempre, un polinomio en x, lo que puede resultar —al menos— llamativo). En efecto, veamos que los $T_n(x)$ definidos como en (6.2) cumplen (6.1). Para $n = 0$, es evidente que $T_0(x) = \cos(0) = 1$; y, para $n = 1$, $T_1(x) = \cos(\arc\cos(x)) = x$. Comprobemos que satisfacen la relación de recurrencia. Para ello, basta usar que

$$\cos((n+1)\theta) + \cos((n-1)\theta) = 2\cos(\theta)\cos(n\theta),$$

lo cual puede probarse mediante trigonometría elemental.

Teorema 6.5.
Sea $\theta = r\pi$ con $r \in \mathbb{Q}$. Entonces, $\cos(\theta)$, $\operatorname{sen}(\theta)$ y $\operatorname{tg}(\theta)$ son números irracionales, con las siguientes excepciones:

(a) $\cos(\theta) = 0, \pm 1$ *o* $\pm 1/2$.
(b) $\operatorname{sen}(\theta) = 0, \pm 1$ *o* $\pm 1/2$.
(c) $\operatorname{tg}(\theta) = 0, \pm 1$ *o no está definida.*

Demostración. En primer lugar, a partir de los polinomios de Chebyshev, definimos $f_n(x) = 2T_n(x/2)$. Así, (6.1) se convierte en

$$\begin{cases} f_0(x) = 2, \quad f_1(x) = x, \\ f_{n+1}(x) = xf_n(x) - f_{n-1}(x), \quad n \geq 1. \end{cases}$$

[2] Esta reformulación es fuente de muchas otras propiedades; en particular, permite probar fácilmente la ortogonalidad de los polinomios de Chebyshev (véase el ejercicio 6.8). De aquí surge su importancia en teoría de aproximación: series de Fourier, interpolación, fórmulas de cuadratura, etcétera. En este libro, los polinomios de Chebyshev nos volverán a salir en otro tema, concretamente en el ejercicio 7.42, donde veremos una curiosa fórmula que los involucra.

De aquí se sigue que, para $n \geq 1$, cada f_n es un polinomio de grado n, mónico y con coeficientes enteros. Además, (6.2) se expresa ahora como

$$f_n(2\cos(\theta)) = 2\cos(n\theta). \tag{6.3}$$

Una vez definidos los polinomios f_n y establecidas las propiedades que usaremos, podemos continuar. Nótese que, si $r = 0$, estamos entre las excepciones permitidas, luego podemos suponer que $r \neq 0$.

(a) Tenemos $\theta = r\pi$ con $r = h/k \in \mathbb{Q}$ ($h \in \mathbb{Z} \setminus \{0\}$ y $k \in \mathbb{N}$). Entonces, $k\theta = h\pi$, y de (6.3) se deduce que

$$f_k(2\cos(\theta)) = 2\cos(k\theta) = 2\cos(h\pi) = \pm 2.$$

Por tanto, $2\cos(\theta)$ es raíz del polinomio $f_k(x) \pm 2 = 0$, que es mónico y con coeficientes enteros. Si $2\cos(\theta)$ fuera racional, el corolario 6.3 asegura que debe ser un entero. Como, por otra parte, $\cos(\theta) \in [-1, 1]$, las únicas posibilidades son $\cos(\theta) = 0$, ± 1 o $\pm 1/2$.

(b) Basta aplicar $\operatorname{sen}(\theta) = \cos(\pi/2 - \theta)$, el hecho de que θ es un múltiplo racional de π si y sólo si lo es $\pi/2 - \theta$, y el apartado (a).

(c) Si $\operatorname{tg}(\theta)$ fuera racional, la identidad $\cos(2\theta) = (1 - \operatorname{tg}^2(\theta))/(1 + \operatorname{tg}^2(\theta))$ prueba que también lo sería $\cos(2\theta)$; además, 2θ es un múltiplo racional de π si y sólo si también lo es θ. Veamos de qué posibles valores de «tg» pueden provenir las excepciones para «cos» del apartado (a). Cuando $\cos(2\theta) = 0$, es claro que $\operatorname{tg}(\theta) = \pm 1$; cuando $\cos(2\theta) = 1$, $\operatorname{tg}(\theta) = 0$; cuando $\cos(2\theta) = -1$, $\operatorname{tg}(\theta)$ no está definida; finalmente, cuando $\cos(2\theta) = \pm 1/2$, debe ser $(1 - \operatorname{tg}^2(\theta))/(1 + \operatorname{tg}^2(\theta)) = \pm 1/2$, y de aquí es fácil deducir que $\operatorname{tg}(\theta)$ toma alguno de los valores $\pm\sqrt{3}$ o $\pm 1/\sqrt{3}$, con lo cual es irracional. □

Nota 6.6. La demostración anterior es, esencialmente, la de [78, § 6.3, teorema 6.16], aunque el resultado ya era conocido anteriormente. Existen otras pruebas que usan técnicas muy distintas; por ejemplo, la de [26, § 6.5, teorema 6.11].

Por otra parte, es fácil ver que $\cos(\pi r)$ es algebraico para cada cada $r \in \mathbb{Q}$ (y análogamente con $\operatorname{sen}(\pi r)$ y $\operatorname{tg}(\pi r)$); véase el ejercicio 6.15. ◁

6.1.2. Logaritmos

Acabemos la sección mostrando otro tipo de expresiones para las que es fácil discernir si proporcionan o no un racional; esta vez, referido a logaritmos. Dado un número positivo b, usemos $\log_b$ para indicar el logaritmo en base b. Con esta notación, y si tenemos $b, c \in \mathbb{Q}$, $b, c > 0$, ¿cómo saber si $\log_b(c)$ es racional o

irracional? Para averiguarlo, basta usar la descomposición en factores primos de los numeradores y denominadores de b y c; y aplicar que —según asegura el teorema fundamental de la aritmética— la descomposición canónica de cada entero es única. Veamos un ejemplo:

Ejemplo. Comprobar si $\log_{10}(25)$ es racional o irracional.

Solución. Supongamos que $\log_{10}(25) = h/k$, un racional. Esto quiere decir que $10^{h/k} = 25$, o $10^h = 25^k$. Descomponiendo 10 y 25 en factores primos, tendríamos $2^h \cdot 5^h = 5^{2k}$. Por la unicidad de la descomposición en factores primos, esto sólo es posible con $h = 0$ y $h = 2k$, luego $h = k = 0$, y por tanto h/k no estaría definido. Como consecuencia, $\log_{10}(25)$ es irracional. ◊

Los ejercicios que hay al final del presente tema muestran más resultados de este tipo.

6.2. Irracionalidad de e

A decir verdad, en este libro ya hemos demostrado varias veces que e es irracional. En primer lugar, en el tema 4 (véase el teorema 4.8). Allí lo hacíamos aplicando un criterio de irracionalidad (teorema 4.7) que era consecuencia del teorema de Dirichlet. De hecho, nuestro objetivo entonces era —más que probar la irracionalidad de e— mostrar un ejemplo de cómo aplicar el teorema. En segundo lugar, en el tema 5 (sección 5.8) a partir de ideas completamente diferentes: el desarrollo de e en fracción continua simple (como es infinito, e es irracional).

Aquí damos otra prueba basada en hechos más simples (aunque, para ser honestos, hay que reconocer que la demostración en sí no es muy distinta a la del tema 4, pues usa la misma manipulación de la serie que define e). Podemos decir que la demostración de la irracionalidad de e es relativamente sencilla (mucho más que la de π, por ejemplo); esto se debe a la rápida convergencia de la serie $e = \sum_{k=0}^{\infty} 1/k!$. Posiblemente fue Joseph Fourier quien, alrededor de 1815, se dio cuenta de que, directamente a partir de esa serie, se podía demostrar que e es un número irracional.

6.2.1. Introducción histórica

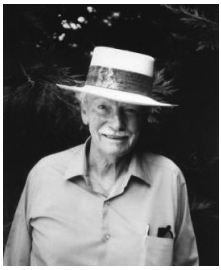
En aras de la justicia histórica, estamos obligados a hacer algunos comentarios sobre los métodos del genial Leonhard Euler, quien, en 1737, fue el primero en probar la irracionalidad de e (también probó la de e^2). Para ello, utilizó ideas similares a las que podemos encontrar en la sección 5.8 del tema 5.



Jean Baptiste Joseph Fourier (1768–1830)



Johann Heinrich Lambert (1728–1777)



Ivan Niven (1915–1999)

Curiosamente, fue Euler quien descubrió la serie de potencias de e^x (la dedujo a partir de la serie para el binomio); incluso él mismo introdujo el símbolo e para la base de los logaritmos naturales. Sin embargo, su demostración de la irracionalidad de e estaba basada en el desarrollo en fracción continua[3]

$$\frac{e-1}{2} = \cfrac{1}{1+\cfrac{1}{6+\cfrac{1}{10+\cfrac{1}{14+\cfrac{1}{18+\cdots}}}}}, \tag{6.4}$$

bastante más complicado de establecer que la serie (al menos, con ojos actuales). Euler sabía que una fracción continua simple infinita siempre define un número irracional, luego (6.4) probaba que e es irracional. En realidad, esta misma fracción continua sirve para probar que e^2 es irracional: según demostró Johann Heinrich Lambert en 1766 (véase el teorema 5.20), una fracción continua simple es periódica si y sólo si converge a un irracional cuadrático; como la fracción continua (6.4) no es periódica, $(e-1)/2$ no es un irracional cuadrático, de donde se deduce que e^2 es irracional (¿por qué?; véase el ejercicio 6.23).

Lambert dio justificaciones rigurosas de los desarrollos de Euler y, además, los utilizó para probar que e^x es irracional para cada x racional no nulo (y también la irracionalidad de $\operatorname{tg}(x)$, que nos aparecerá en la siguiente sección, cuando tratemos la irracionalidad de π). Una demostración de la irracionalidad de e^x (con $x \in \mathbb{Q}\backslash\{0\}$) basada en fracciones continuas se puede encontrar en [70]. En ella se usa el siguiente

[3] Este desarrollo ya nos ha aparecido en el tema anterior: véase el ejercicio 5.31.

desarrollo, debido también —esencialmente— a Euler:

$$e^x = \cfrac{1}{1 - \cfrac{2x}{2 + x + \cfrac{x^2}{6 + \cfrac{x^2}{10 + \cfrac{x^2}{14 + \cdots}}}}} \tag{6.5}$$

(los números 6, 10, 14,... continúan indefinidamente con la forma $4k+2$, sin que aparezcan más «x» en su correspondiente posición), que converge —aunque no sea fácil probarlo rigurosamente— para cualquier $x \in \mathbb{C}$. Actualmente, podemos considerar que estas demostraciones son —casi— una interesante curiosidad académica (e histórica, por supuesto), pues son bastante más complicadas que las basadas en la serie numérica cuya suma es e, como la que aquí presentamos. Pero, desde luego, las ideas de Euler son siempre magníficas, y es una delicia leerlo; gracias a [43], no hay ninguna razón para privarse de ese placer.

6.2.2. Una demostración elemental

La demostración de la irracionalidad de e que vamos a dar parte de la conocida serie $e = \sum_{k=0}^{\infty} 1/k!$. Aparte de eso, sólo usa un hecho completamente elemental, al que a veces se alude con el nombre de *principio fundamental de la teoría de números*: no existe ningún entero entre 0 y 1. ¡Es sorprendente que esta observación tan simple resulte útil a menudo! (por ejemplo, también la usaremos en el teorema 6.9). Vamos con ello:

Teorema 6.7. Euler, 1737.
El número e es irracional.

Demostración. Recordemos que $e = \sum_{k=0}^{\infty} 1/k!$. Supongamos que $e = m/n \in \mathbb{Q}$. Resulta obvio que $n!(m/n - \sum_{k=0}^{n} 1/k!)$ es un entero. Además, $e - \sum_{k=0}^{n} 1/k!$ es positivo. Entonces,

$$\begin{aligned} 0 < n!\left(\frac{m}{n} - \sum_{k=0}^{n}\frac{1}{k!}\right) &= n!\left(e - \sum_{k=0}^{n}\frac{1}{k!}\right) \\ &= n!\left(\frac{1}{(n+1)!} + \frac{1}{(n+2)!} + \frac{1}{(n+3)!} + \frac{1}{(n+4)!} + \cdots\right) \end{aligned}$$

$$= \frac{n!}{(n+1)!}\left(1+\frac{1}{n+2}+\frac{1}{(n+2)(n+3)}+\frac{1}{(n+2)(n+3)(n+4)}+\cdots\right)$$
$$\leq \frac{1}{n+1}\left(1+\frac{1}{n+2}+\frac{1}{(n+2)^2}+\frac{1}{(n+2)^3}+\cdots\right)$$
$$= \frac{1}{n+1}\cdot\frac{1}{1-1/(n+2)} = \frac{n+2}{(n+1)^2} < 1,$$

y hemos encontrado que el entero

$$n!\left(\frac{m}{n}-\sum_{k=0}^{n}\frac{1}{k!}\right)$$

está entre 0 y 1, lo cual es absurdo. □

6.3. Irracionalidad de π

En esta sección vamos a probar que π es un número irracional. La irracionalidad de π fue probada por Lambert en 1761. Su demostración —bastante laboriosa— comenzaba estableciendo el siguiente desarrollo en fracción continua:

$$\operatorname{tg}(x) = \cfrac{x}{1-\cfrac{x^2}{3-\cfrac{x^2}{5-\cdots}}}. \tag{6.6}$$

Luego mostraba, por un argumento de descenso infinito, que si $x \neq 0$ es racional, entonces el lado derecho de (6.6) es irracional. Como $\operatorname{tg}(\pi/4) = 1$, esto implica que π es irracional. La principal dificultad de esta prueba radica en establecer rigurosamente la convergencia de diversos procesos infinitos que aparecen; por ejemplo, de la fracción continua (6.6) (que no es una fracción continua simple como las que usábamos en el tema 5). Diversos autores (entre ellos, Gauss) simplificaron la demostración de Lambert, aunque conservando las ideas originales. Remitimos al lector a [62], donde —junto con referencias y notas históricas— puede encontrar una demostración completa basada en este tipo de argumentos; además, una deducción sencilla de (6.6) se puede ver en [70]. Por otra parte, no podemos dejar de comentar que la fracción continua (6.6) está relacionada con (6.5), la fracción continua que se podía usar como punto de partida de una demostración de la irracionalidad de e^x (con $x \in \mathbb{Q}$), tal como vimos en la sección anterior. Concretamente, (6.5) se obtiene

de (6.6) aplicando la identidad

$$e^x = \cfrac{1}{1-\cfrac{2}{1-\cfrac{1}{i\,\mathrm{tg}(ix/2)}}},$$

válida para todo $x \in \mathbb{C}$, y que es fácil de comprobar teniendo en cuenta que $i\,\mathrm{tg}(ix/2) = -\,\mathrm{tgh}(x/2)$ (nótese que, en (6.6), cada «trozo» de la forma $\frac{p_j}{q_j+p_{j+1}}$ se puede cambiar por $\frac{2p_j}{2q_j+2p_{j+1}}$ sin alterar el valor de la fracción continua).

Actualmente, los libros suelen mostrar pruebas que son variaciones de argumentos de Hermite, completamente diferentes de la original de Lambert. Así ocurre con la que aquí presentamos, que es una elegante demostración debida a Ivan Niven [76], y que sólo usa cálculo elemental (véase también [78, §6.3] o [50, §4.7]). Comencemos con un lema; lo vamos a enunciar de manera bastante general, para poderlo aprovechar también en las pruebas de la trascendencia de e y de π (ahora sólo usaremos $c = 0$ y, además, nos bastará con la consecuencia final).

Lema 6.8.

Sean $f(x) \in \mathbb{Z}[x]$, $n \in \mathbb{N}$, $c \in \mathbb{Z}$ y $\widetilde{f}(x) = (x-c)^n f(x)$. En estas circunstancias, las derivadas de $\widetilde{f}$ cumplen

$$\widetilde{f}^{(j)}(c) = 0 \text{ si } j < n,$$
$$\widetilde{f}^{(j)}(c) \text{ es múltiplo de } j! \text{ si } j \geq n.$$

En particular, $n!$ divide a $\widetilde{f}^{(j)}(c)$ para todo $j \geq 0$.

Demostración. El polinomio $f(x)$ lo podemos escribir como una suma finita $f(x) = \sum_{k\geq 0} c_k(x-c)^k$ con $c_k \in \mathbb{Z}$ (si $f(x)$ viene dado como $\sum_{k\geq 0} b_k x^k$, basta darse cuenta de que $\{(x-c)^k\}_{k\geq 0}$ es una base del espacio de los polinomios; de hecho, los c_k se pueden hallar haciendo el desarrollo de Taylor en torno a c). Entonces,

$$\widetilde{f}(x) = \sum_{k\geq 0} c_k(x-c)^{k+n} = \sum_{j\geq n} c_{j-n}(x-c)^j$$

(de nuevo, las sumas son finitas). Comparando con la bien conocida fórmula del desarrollo de Taylor de $\widetilde{f}$ en torno a c,

$$\widetilde{f}(x) = \sum_{j=0}^{\infty} \frac{\widetilde{f}^{(j)}(c)}{j!}(x-c)^j,$$

es claro que

$$\widetilde{f}^{(j)}(c) = 0 \qquad \text{si } j < n,$$
$$\widetilde{f}^{(j)}(c) = j!\, c_{j-n} \quad \text{si } j \geq n,$$

y esto demuestra el lema. (Como curiosidad, nótese que también será $\widetilde{f}^{(j)}(c) = 0$ para todo j suficientemente grande, pero esto no tiene influencia en lo que perseguimos.) □

Así, abordemos el resultado que buscábamos:

Teorema 6.9. Lambert, 1761.
El número π es irracional.

Demostración. Supongamos que no lo fuera, es decir, $\pi = a/b$ con $a, b \in \mathbb{N}$. Sea, entonces,

$$F(x) = \frac{x^n(a - bx)^n}{n!},$$

donde n lo elegiremos más adelante. Por el lema (con $f(x) = (a - bx)^n$ y $c = 0$), $F^{(j)}(0) \in \mathbb{Z}$ para todo $j \geq 0$. Además,

$$F(x) = \frac{b^n x^n(\pi - x)^n}{n!} = F(\pi - x),$$

luego $F^{(j)}(x) = (-1)^j F^{(j)}(\pi - x)$. Por lo tanto, también $F^{(j)}(\pi) \in \mathbb{Z}$ para todo $j \geq 0$. Definamos ahora

$$I = \int_0^\pi F(x)\,\mathrm{sen}(x)\,dx. \tag{6.7}$$

Probaremos que, en cualquier caso, $I \in \mathbb{Z}$; y que, si n es grande, $0 < I < 1$. Esto es, obviamente, una contradicción, que proviene de haber supuesto que π es racional. Vamos con ello.

Comprobemos, en primer lugar, que $I \in \mathbb{Z}$. Sea

$$G(x) = F(x) - F''(x) + F^{(4)}(x) - \cdots + (-1)^n F^{(2n)}(x).$$

Entonces,

$$G''(x) = F''(x) - F^{(4)}(x) - \cdots + (-1)^{n-1} F^{(2n)}(x) + 0,$$

ya que F es un polinomio de grado $2n$. Sumando ambas expresiones, $F = G + G''$.

Sustituyendo en (6.7) y usando que $\frac{d}{dx}\big(G'(x)\operatorname{sen}(x)-G(x)\cos(x)\big)=G(x)\operatorname{sen}(x)+G''(x)\operatorname{sen}(x)$, tenemos[4]

$$\begin{aligned} I &= \int_0^\pi \big(G(x)\operatorname{sen}(x)+G''(x)\operatorname{sen}(x)\big)\,dx \\ &= \big(G'(x)\operatorname{sen}(x)-G(x)\cos(x)\big)\Big|_{x=0}^{\pi} = G(\pi)+G(0)\in\mathbb{Z}. \end{aligned}$$

Finalmente, veamos que $0<I<1$ cuando n es suficientemente grande. La función $y=x(\pi-x)$ definida en el intervalo $(0,\pi)$ es positiva y tiene su máximo en $x=\pi/2$, así que

$$0<F(x)=\frac{b^n}{n!}\big(x(\pi-x)\big)^n\le\frac{b^n\pi^{2n}}{2^{2n}n!},\quad x\in(0,\pi).$$

También, $\operatorname{sen}(x)>0$ para $x\in(0,\pi)$. Con esto, de (6.7) deducimos que

$$0<I\le\pi\,\frac{b^n\pi^{2n}}{2^{2n}n!}<1,$$

si n grande. □

6.3.1. Algunas adaptaciones y generalizaciones

Diversos matemáticos han adaptado la idea de la demostración del teorema 6.9 para conseguir pruebas sencillas de los siguientes hechos:

- π^2 es irracional;
- e^r es irracional para cada $r\in\mathbb{Q}\setminus\{0\}$.

No reproducimos aquí las correspondientes demostraciones; si el lector quiere verlas, puede hacerlo en [2] o [50].

Además, el mismo Niven [77, § 2.2] adaptó su demostración para probar que $\cos(r)$ es un número irracional[5] para cada $r\in\mathbb{Q}\setminus\{0\}$. En particular, de ahí se obtenía como corolario que π es irracional, ya que, en caso contrario, $\cos(\pi)=-1$ debería ser irracional. Usando las relaciones trigonométricas

$$1-2\operatorname{sen}^2(r)=\cos(2r),\qquad \cos(2r)=\frac{1-\operatorname{tg}^2(r)}{1+\operatorname{tg}^2(r)}$$

es fácil probar que la irracionalidad de $\cos(r)$ para todo $r\in\mathbb{Q}\setminus\{0\}$ implica la

[4] Obviamente, lo mismo se obtiene integrando $\int_0^\pi G''(x)\operatorname{sen}(x)\,dx$ por partes dos veces.

[5] Obsérvese que este estudio no es comparable con el del teorema 6.5; allí analizábamos la irracionalidad de $\cos(r\pi)$ cuando $r\in\mathbb{Q}$.

de sen(r) y tg(r). Como consecuencia inmediata, también arc cos(r), arc sen(r) y arc tg(r) deberán ser irracionales cuando la variable r es un racional no nulo.

En realidad, también es cierto que e^r y cos(r) son trascendentes cuando $r \in \mathbb{Q} \setminus \{0\}$, pero éste es un resultado que requiere matemáticas considerablemente más avanzadas. A este respecto, véase el corolario 6.33, cuyo enunciado es todavía más general; eso sí, aclaremos de antemano que está basado en un teorema cuya demostración no incluimos en este texto.

Concluyamos nuestro recorrido por los números irracionales mencionando que la irracionalidad de cos(r) también se puede deducir con otros métodos. Por ejemplo, en el ejercicio 6.14 proponemos al lector otra estrategia para probar, a partir de su desarrollo en serie, que cos(1) es irracional (lo cual es fácilmente adaptable a cualquier $r \in \mathbb{Q} \setminus \{0\}$).

6.4. Los números trascendentes existen

Corría el año 1844 cuando Joseph Liouville logró probar la existencia de números trascendentes. Antes, tal existencia se conjeturaba —había sido Euler el primero en especular que los números trascendentes existían—, pero nadie podía estar seguro de ello. La demostración de Liouville fue constructiva, pues su método lograba encontrar números concretos que eran trascendentes (por ejemplo, $\alpha = \sum_{k=1}^{\infty} 1/2^{k!}$).

Ya hemos probado este resultado en la sección 4.2 del tema 4 pero, para contribuir a que éste sea un tema autocontenido, lo probaremos aquí de nuevo. Eso sí, daremos una demostración ligeramente distinta, basada en el teorema del valor medio; y avanzaremos más rápido, sin detenernos en tanto detalle.

Teorema 6.10. Liouville, 1844.
Sea $\alpha \in \mathbb{R}$ un número algebraico de orden $n \geq 2$. Entonces, existe un $K > 0$ tal que cualquier racional p/q cumple

$$\left|\alpha - \frac{p}{q}\right| > \frac{K}{q^n}. \tag{6.8}$$

Como consecuencia, para todo $\varepsilon > 0$, la desigualdad

$$\left|\alpha - \frac{p}{q}\right| < \frac{1}{q^{n+\varepsilon}} \tag{6.9}$$

sólo tiene una cantidad finita de soluciones racionales p/q.

DEMOSTRACIÓN. Basta estudiar los $p/q \in [\alpha-1, \alpha+1]$ pues, en otro caso, $|\alpha - p/q| > 1 \geq 1/q^n$ y (6.8) se cumpliría con cualquier $K \leq 1$.

Como α es algebraico de orden n, será raíz de un polinomio

$$f(x) = a_n x^n + a_{n-1}x^{n-1} + \cdots + a_2 x^2 + a_1 x + a_0 \in \mathbb{Z}[x], \quad a_n \neq 0.$$

Este f no tiene raíces racionales ya que, si $r \in \mathbb{Q}$ es raíz de f, entonces $f(x)/(x-r)$ es un polinomio en $\mathbb{Q}[x]$, sigue teniendo α como raíz, y es de grado $n-1$; si así ocurriera, α no sería algebraico de orden n. En consecuencia,

$$\left| f\left(\frac{p}{q}\right)\right| = \frac{1}{q^n}\left|a_n p^n + a_{n-1}p^{n-1}q + \cdots + a_2 p^2 q^{n-2} + a_1 p q^{n-1} + a_0 q^n\right| \geq \frac{1}{q^n},$$

pues el numerador es un entero no nulo. Por el teorema del valor medio,

$$f(\alpha) - f\left(\frac{p}{q}\right) = f'(\eta)\left(\alpha - \frac{p}{q}\right)$$

para algún valor de η intermedio entre α y p/q. Entonces,

$$\frac{1}{q^n} \leq \left| f\left(\frac{p}{q}\right)\right| = \left| f(\alpha) - f\left(\frac{p}{q}\right)\right| \leq \max_{x\in[\alpha-1,\alpha+1]} |f'(x)| \left|\alpha - \frac{p}{q}\right|,$$

de donde se sigue (6.8).

Llegar a (6.9) es ya fácil: combinando las desigualdades (6.8) y (6.9) tendríamos

$$\frac{K}{q^n} < \left|\alpha - \frac{p}{q}\right| < \frac{1}{q^{n+\varepsilon}},$$

pero sólo puede haber una cantidad finita de q tales que $K/q^n < 1/q^{n+\varepsilon}$ (es decir, $Kq^\varepsilon < 1$); y, de aquí, sólo un cantidad finita de p/q. □

Corolario 6.11. Liouville, 1844.
El número $\alpha = \sum_{k=1}^{\infty} \frac{1}{2^{k!}}$ es trascendente.

DEMOSTRACIÓN. En primer lugar, observemos que la representación en base 2 de α es

$$\sum_{k=1}^{\infty} \frac{1}{2^{k!}} = 0.1100010\ldots010\ldots010\ldots{}_{(2}$$

donde los dígitos «1» están colocados en las posiciones $k!$ y el resto de los dígitos son todos «0». Como esa representación no es finita ni periódica, α no es un número

racional. Supongamos ahora que α es algebraico de orden $n \geq 2$. Para cada entero $m \geq 2$, sea

$$\frac{p}{q} = \sum_{k=1}^{n+m-1} \frac{1}{2^{k!}},$$

luego $q = 2^{(n+m-1)!}$. Entonces

$$\left|\alpha - \frac{p}{q}\right| = \sum_{k=n+m}^{\infty} \frac{1}{2^{k!}} < \sum_{k=(n+m)!}^{\infty} \frac{1}{2^k} = \frac{1}{2^{(n+m)!}} \left(1 + \frac{1}{2} + \frac{1}{2^2} + \cdots\right)$$
$$= \frac{2}{2^{(n+m)!}} = \frac{2}{q^{n+m}} \leq \frac{2}{q^{n+2}} < \frac{1}{q^{n+1}},$$

y esto se cumple para infinitos p/q (uno por cada m), lo cual contradice al teorema anterior (con $\varepsilon = 1$). □

6.4.1. No sólo existen, sino que hay muchísimos

La sorpresa saltó en 1874, cuando Georg Cantor dio un argumento según el cual «casi todos» los números son trascendentes, en un sentido que ahora precisaremos.

Comencemos aclarando un poco las cosas, y poniéndonos en las circunstancias de aquella época. Nuestra intuición nos hace pensar que «el todo es mayor que cada una de sus partes», y tal propiedad estaba considerada como verdad inamovible. En un conjunto infinito, eso no tenía por qué ocurrir. Por ejemplo, un subconjunto de los enteros positivos es el de los números pares y, sin embargo, los elementos de ambos conjuntos se pueden emparejar mediante la regla $n \mapsto 2n$; hemos encontrado así un conjunto que tiene el mismo número de elementos que una de sus partes. A primera vista, esa contradicción con lo que nos muestra la intuición puede ser considerada paradójica.

Ya desde la época griega, los matemáticos no estaban acostumbrados a tratar directamente con conjuntos infinitos; cuando alguien se atrevía a cruzar esa frontera, los ataques por falta de rigor podían ser furiosos. Con la nomenclatura que se usaba, el infinito existía «en potencia», pero no «en acto». Fue Richard Dedekind en 1872 quien propuso utilizar ese extraño comportamiento de los conjuntos infinitos de una manera provechosa. Así, definió *conjunto infinito* como aquél que se puede emparejar con alguna de sus partes; por supuesto, al decir emparejar nos referimos a que exista una aplicación biyectiva entre el conjunto y su subconjunto.

La idea ahora es que no podemos confiar en nuestra intuición —que obviamente nunca trata con conjuntos infinitos— y que se deber recurrir a la razón para analizar



Richard Dedekind
(1831–1916)



Georg Cantor
(1845–1918)



David Hilbert
(1862–1943)

lo que ocurre. En todo caso, una vez que ya disponemos de una definición, se puede profundizar en el estudio de los conjuntos infinitos y sus propiedades. Y, desde luego, fue Cantor quien llevó la voz cantante en este estudio.[6]

Numerar un conjunto significa probar que existe una biyección entre $\mathbb{N}$ y ese conjunto. Es decir, que podemos disponer los elementos del conjunto en forma de sucesión (con todos sus términos distintos) o —con el mismo significado— en forma de «lista»; la posición en la lista (o en la sucesión) da la biyección con $\mathbb{N}$. Cuando esto se puede conseguir, se dice que el conjunto es *numerable*. En caso contrario, que es *no numerable*.

Pues bien, en 1873, Cantor demostró que $\mathbb{Q}$ es numerable, y que también lo es el conjunto de los números algebraicos; el año siguiente logró probar que $\mathbb{R}$ es no numerable. Con lo cual, los números trascendentes forman un conjunto no numerable. Y todo esto, ¡sin necesidad de encontrar ningún número trascendente! Más que matemáticas, parecía magia.

Aunque los razonamientos de Cantor son impecables y fáciles de seguir, en su época —todo hay que decirlo— no fueron bien aceptados. Como ya hemos comentado, la teoría de conjuntos no estaba aún desarrollada a finales del siglo XIX (preci-

[6] En realidad, fue algo aparentemente alejado —el estudio de las series de Fourier— lo que motivó el interés de Cantor por la teoría de conjuntos infinitos. Una serie de Fourier es algo del estilo $f(x) = \sum_{n=1}^{\infty} a_n \operatorname{sen}(nx)$, y habían sido introducidas por Fourier —en el primer cuarto del siglo XIX— en relación con sus estudios sobre la propagación del calor, pero su tratamiento distaba mucho de ser riguroso. Además, diversos argumentos físicos y matemáticos sobre la convergencia de estas series habían originado respuestas contradictorias, lo cual —por supuesto— no era del agrado de los matemáticos. Era necesario poner orden. En el fondo, y sin precisar lo que queremos decir, la respuesta es que la serie se maneja bien siempre y cuando el conjunto de los x en los que la función f no es «buena» no sea demasiado «raro». Estudiar esto requería construir una teoría que permitiera examinar cuidadosamente la estructura de los subconjuntos infinitos de la recta real. Es por ello por lo que Cantor comenzó —en 1873— a elaborar una teoría sobre conjuntos infinitos, tema en el que continuó trabajando los años posteriores.

samente fue Cantor su principal impulsor), y el manejo de conjuntos infinitos daba lugar a numerosos problemas conceptuales; muchos matemáticos no se encontraban a gusto con las soluciones propuestas, y seguían presos de sus ideas preconcebidas. Tuvo que discurrir cierto tiempo hasta que las nuevas generaciones de matemáticos —con el gran David Hilbert a la cabeza— aceptaran los planteamientos de Cantor y reconocieran su genio creador.

Nuestro objetivo en esta sección es probar los resultados de Cantor, con unos argumentos que son, esencialmente, los suyos. Antes de tomar contacto con los números algebraicos y los trascendentes, vamos a ver que los números racionales son numerables, y que los números reales no lo son.

Para demostrar que los racionales son numerables, basta buscar una forma de numerar todos los racionales positivos. En efecto, supongamos que hemos dispuesto todos los racionales positivos en forma de sucesión $\{r_n\}_{n=1}^{\infty}$. Entonces, $\mathbb{Q}$ quedaría numerado si tomamos primero 0 y luego, sucesivamente, los r_n y $-r_n$, alternándolos.

$$\begin{array}{ccccccccc}
\frac{1}{1} & \rightarrow & \frac{2}{1} & & \frac{3}{1} & & \frac{4}{1} & & \frac{5}{1} & \cdots \\
 & \swarrow & & \swarrow & & \swarrow & & \swarrow & & \\
\frac{1}{2} & & \frac{2}{2} & & \frac{3}{2} & & \frac{4}{2} & & \frac{5}{2} & \cdots \\
 & \swarrow & & \swarrow & & \swarrow & & & & \\
\frac{1}{3} & & \frac{2}{3} & & \frac{3}{3} & & \frac{4}{3} & & \frac{5}{3} & \cdots \\
 & \swarrow & & \swarrow & & & & & & \\
\frac{1}{4} & & \frac{2}{4} & & \frac{3}{4} & & \frac{4}{4} & & \frac{5}{4} & \cdots \\
 & \swarrow & & & & & & & & \\
\frac{1}{5} & & \frac{2}{5} & & \frac{3}{5} & & \frac{4}{5} & & \frac{5}{5} & \cdots \\
\vdots & & \vdots & & \vdots & & \vdots & & \vdots & \ddots
\end{array}$$

Así pues, numeremos los racionales positivos; para ello, dispongámoslos en forma de cuadro, de modo que en las filas avancen los numeradores y en las columnas los denominadores. Ahora, lo recorremos en diagonal según indican la flechas: (comenzamos por $\frac{1}{1}$ y, cada vez, cuando llegamos a una fracción del tipo $\frac{1}{n}$, cambiamos de diagonal y seguimos con $\frac{n+1}{1}$). Algunos racionales están repetidos, como $\frac{1}{1}$ y $\frac{2}{2}$, o $\frac{1}{2}$ y $\frac{2}{4}$; en estos casos, cuando un racional ya ha aparecido antes, nos lo saltamos en el recorrido que estamos describiendo (en otras palabras, podríamos decir que nos saltamos todas las fracciones no irreducibles). De este modo, hemos logrado colocar todos los racionales positivos en forma de lista, tal como buscábamos. En resumen, hemos probado esto:

Teorema 6.12. Cantor, 1873.
El conjunto $\mathbb{Q}$ es numerable.

Para demostrar que los reales no son numerables se emplea otro «argumento diagonal».[7] En realidad, probaremos que el intervalo $(0,1)$ no es numerable. Esto es suficiente, pues $(0,1)$ es un subconjunto de $\mathbb{R}$ (además, es fácil dar una biyección entre el intervalo $(0,1)$ y $\mathbb{R}$: basta usar la función $y=(2x-1)/(x-x^2)$ o, también, $y=\operatorname{tg}((x-1/2)\pi)$).

Cualquier número $r\in(0,1)$ tiene una expresión decimal de la forma

$$r=\sum_{j=1}^{\infty}\frac{d_j}{10^j}=0.d_1d_2d_3d_4d_5\ldots,$$

donde los dígitos d_j cumplen $d_j\in\{0,1,2,3,4,5,6,7,8,9\}$. No permitimos que haya infinitos $d_j=9$ seguidos, con lo cual la expresión decimal de cada número es única (sin esta salvedad, la unicidad no es cierta: por ejemplo, $0.4999999\ldots$ y $0.5000000\ldots$ representan ambos al racional $1/2$).

Para demostrar la no numerabilidad del intervalo $(0,1)$ emplearemos reducción al absurdo. Supongamos que hemos logrado numerar $(0,1)$, de tal manera que tenemos una sucesión $\{r_n\}_{n=1}^{\infty}$ en la que están todos los reales entre 0 y 1. De este modo, podemos escribir

$$\begin{aligned}r_1&=0.d_1^1d_2^1d_3^1d_4^1d_5^1\ldots,\\ r_2&=0.d_1^2d_2^2d_3^2d_4^2d_5^2\ldots,\\ r_3&=0.d_1^3d_2^3d_3^3d_4^3d_5^3\ldots,\\ r_4&=0.d_1^4d_2^4d_3^4d_4^4d_5^4\ldots,\end{aligned}$$

y así sucesivamente. Recorriendo el cuadro anterior en diagonal, construyamos un número

$$r=0.d_1d_2d_3d_4d_5\ldots,$$

de forma que

$$d_1\neq d_1^1,\quad d_2\neq d_2^2,\quad d_3\neq d_3^3,\quad d_4\neq d_4^4,\quad d_5\neq d_5^5,\quad\ldots$$

(de paso, no elegimos ningún $d_j=0$ o $d_j=9$, para que no pueda haber problemas con la no unicidad en la representación decimal, y para evitar que r sea 0 o 1). Ahora, por una parte, $r\in(0,1)$; y, por otra, este r no está en la lista ya que $d_j\neq d_j^j$ por construcción, luego $r\neq r_j$ para todos los j. Esto es un absurdo, que proviene de haber supuesto que el intervalo $(0,1)$ es un conjunto numerable.

Así, hemos probado lo siguiente:

[7] Para ser fieles a la historia, debemos reseñar que la argumentación que sigue no es la original de Cantor de 1874, que requería matemáticas relativamente avanzadas, sino una demostración, igualmente suya, de 1891.

Teorema 6.13. Cantor, 1874.
El conjunto $\mathbb{R}$ es no numerable.

A quien no tenga un conocimiento previo del tema, seguro que todo esto le suena muy extraño (¡lo mismo que les ocurrió a la mayoría de los matemáticos contemporáneos de Cantor!). Es más, la cardinalidad de los conjuntos infinitos proporciona muchos otros resultados antiintuitivos. Queremos insistir en que no debemos preocuparnos por ello: en la realidad cotidiana, nunca nos topamos con conjuntos infinitos, sino que éstos son una creación matemática efectuada por nuestra mente (lo cual —¡no nos equivoquemos!— no quiere decir que en la práctica no resulten útiles, pues son la mejor manera que tenemos de modelizar el mundo físico en el que vivimos). Nuestra intuición sólo puede opinar sobre conjuntos finitos, pues no tiene ninguna experiencia previa con conjuntos infinitos. La misma definición de *conjunto infinito* parece absurda a primera visita —recordémosla: se dice que un conjunto es infinito si se puede poner en biyección con algún subconjunto propio—. Y precisamente la definición es así pues eso nunca se puede conseguir con conjuntos finitos (lo que, a su vez, hace que a nuestra intuición le parezca tan rara).

Aunque no es el objetivo que aquí perseguimos, en las dos notas que siguen comentaremos un par de estos extraños resultados a los que estamos aludiendo.

Nota 6.14. La cardinalidad de los conjuntos infinitos tampoco está relacionada con la dimensión: por ejemplo, el plano $\mathbb{R}^2$ tiene exactamente el mismo número de puntos que la recta $\mathbb{R}$. El mismo Cantor se vio sorprendido por este hecho; cuando, tras años de intentar probar el resultado contrario, en 1877 demostró que existía una biyección entre $\mathbb{R}^d$ y $\mathbb{R}$, en una carta a Richard Dedekind escribía «¡lo veo, pero no lo creo!».

En realidad, la demostración es bastante sencilla, como veremos a continuación; lo hacemos sólo para $d = 2$. Basta comprobar que el conjunto de los pares (x, y) con $0 < x, y \leq 1$ se puede poner en biyección con el intervalo $(0, 1]$. Antes de seguir, aclaremos que a cada número le hacemos corresponder su desarrollo decimal de manera que no acabe en una sucesión infinita de ceros; por ejemplo, 0.7 sería $0.69999\ldots$ De este modo, la representación decimal de cada número es única.

Describamos la biyección entre $(0, 1] \times (0, 1]$ y $(0, 1]$ con un ejemplo: dado un par (x, y), supongamos que los desarrollos decimales de x e y son como sigue, y agrupemos los dígitos de manera que cada grupo contenga varios o ningún dígito «0» y un único dígito no nulo al final:

$$\begin{aligned} x &= 0.30143100708\ldots &&= 0.3 \quad 01 \quad 4 \; 3 \; 1 \; 007 \quad 08 \quad \ldots \\ y &= 0.0092051230008\ldots &&= 0.009 \quad 2 \; 05 \; 1 \; 2 \quad 3 \; 0008 \quad \ldots \end{aligned}$$



Ernst Zermelo (1871–1953)



Abraham Adolf Fraenkel (1891–1965)

Ahora, asociemos a (x, y) el número $z \in (0, 1]$ cuyo desarrollo decimal se escribe comenzando por el primer grupo de x, luego el primero de y, el segundo de x, y así sucesivamente. En nuestro ejemplo, tendríamos

$$z = 0.3\ 009\ 01\ 2\ 4\ 05\ 3\ 1\ 1\ 2\ 007\ 3\ 08\ 0008\ \ldots$$

El método de construcción de los grupos evita la posibilidad de que z acabe en una sucesión infinita de ceros. Por último, a partir de la expresión de z, es inmediato ver cuál debe ser la preimagen (x, y), luego el proceso es biyectivo, tal como queríamos comprobar. ◁

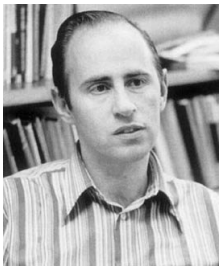
Nota 6.15. La siguiente sorpresa relacionada con los conjuntos infinitos entra de lleno en la estructura lógica de las matemáticas. Como ya sabemos, un conjunto es numerable cuando es biyectivo con $\mathbb{N}$; en ese caso, también se dice que su *cardinal* es $\aleph_0$. Del mismo modo, decimos que $\mathbb{R}$ (o cualquier conjunto biyectivo con él) tiene el *cardinal del continuo* (muchas veces se usa el sinónimo *potencia del continuo*), o que su cardinal es c. En cierto sentido, el teorema 6.13 prueba que $\aleph_0$ es estrictamente menor que c. Pero, ¿hay algún otro cardinal entre $\aleph_0$ y c?; o, en otras palabras, ¿existe algún subconjunto infinito de $\mathbb{R}$ cuyo cardinal no es $\aleph_0$ ni c? Cantor pensaba que no lo había, pero no consiguió probarlo; esta suposición se denominó *hipótesis del continuo*.

Esta pregunta fue una de las que planteó David Hilbert en la célebre conferencia que impartió en el Segundo Congreso Internacional de Matemáticos, celebrado en París en el año 1900, en la que formulaba 23 problemas matemáticos cuya solución no era conocida. De hecho, el primer problema de la lista era demostrar si la hipótesis del continuo era o no cierta.[8]

[8] Podemos ver otros cuatro problemas de la lista en las páginas 60, 193, 380 y 468 (o también el apartado 10.6.2 en la página 607 y siguientes) de este libro.



Kurt Gödel
(1906–1978)



Paul Joseph Cohen
(1934–2007)

Esta pregunta resultó ser mucho más intrincada de lo que parecía. La realidad —descubierta por Paul J. Cohen en 1963, tras completar la respuesta parcial dada por Kurt Gödel en 1938— es que esto es lo que se conoce como un *indecidible*: con los axiomas habituales de las matemáticas, tanto es posible afirmar una cosa como la contraria, pues ninguna de las dos se puede probar a partir de los axiomas de la teoría de conjuntos (los habitualmente denominados axiomas de Zermelo-Fraenkel, véase [46]). Así pues, existen «teorías de conjuntos cantorianas» en las que la hipótesis del continuo es cierta (pues la hemos podido incluir como un axioma adicional independiente de los otros), y «teorías de conjuntos no cantorianas», en las que es falsa. Por su descubrimiento, Cohen recibió la Medalla Fields.[9] ◁

Tras varios incisos, abordemos por fin la demostración de la numerabilidad del conjunto de todos los números algebraicos. Por simplicidad, veamos primero la numerabilidad de $\mathbb{Z}[x]$:

Lema 6.16.
El conjunto $\mathbb{Z}[x]$ es numerable.

Demostración. Para numerar $\mathbb{Z}[x]$ vamos a establecer una serie de «rondas» mediante las cuales nos aparecerán, en un cierto orden biyectivo con $\mathbb{N}$, todos los polinomios $f(x)$ con coeficientes enteros. En cada ronda hay que descartar los polinomios que ya han aparecido en alguna ronda anterior. En lo que sigue, todos los a_j denotan números enteros.

- Primera ronda: $f(x) = a_0$ con $|a_0| < 1$. Esta vez, el único polinomio es $f(x) = 0$ (el polinomio nulo).

[9] La nota al pie número 7 del tema 2 (página 113) proporciona más detalles sobre este importante galardón matemático.

- Segunda ronda: $f(x) = a_0 + a_1 x$ con $|a_j| < 2$, $j = 0, 1$. Salen los siguientes $3^2 = 9$ polinomios, que ordenamos así: $-1-x$, $-1+0x$, $-1+x$, $0-x$, $0+0x$, $0+x$, $1-x$, $1+0x$, $1+x$. Hay que descartar $0+0x$, que ya ha aparecido antes.
- Tercera ronda: $f(x) = a_0 + a_1 x + a_2 x^2$ con $|a_j| < 3$, $j = 0, 1, 2$. En principio, son 5^3 polinomios, y se ordenan con el mismo proceso que antes, descartando los repetidos.
- Ronda n-ésima: $f(x) = a_0 + a_1 x + \cdots + a_{n-1} x^{n-1}$ con $|a_j| < n$, $j = 0, 1, \ldots, n-1$. Por curiosidad —pues esto no tiene influencia en nuestra argumentación— mencionemos que, esta vez, aparecen $(2n-1)^n$ polinomios, no todos nuevos.

Y esto, para cada $n \in \mathbb{N}$. Así, todos los polinomios de $\mathbb{Z}[x]$ aparecen en algún lugar de la lista que hemos construido. □

Teorema 6.17. Cantor, 1874.
Los números algebraicos[10] *forman un conjunto numerable.*

Demostración. Ya hemos probado que el conjunto de todos los polinomios con coeficientes en $\mathbb{Z}$ es numerable. Quitando el polinomio nulo, también $\mathbb{Z}[x] \setminus \{0\}$ es numerable. Ahora, a partir de la lista de polinomios

$$\mathbb{Z}[x] \setminus \{0\} = \{f_1(x), f_2(x), f_3(x), \ldots\},$$

podemos construir

$$\big\{ \{\text{raíces de } f_1\},\ \{\text{raíces de } f_2\},\ \{\text{raíces de } f_3\},\ \ldots \big\}$$

(no incluyendo, cada vez, las raíces que ya hayan aparecido antes) y ya tenemos una lista con todos los números algebraicos. □

Por fin, y dado que la unión de dos conjuntos numerables es numerable (se puede formar una lista intercalando los de las dos listas previas), de este teorema se deduce lo siguiente:

Corolario 6.18. Cantor, 1874.
El conjunto de los números trascendentes es no numerable.

[10] Lo mismo da que nos refiramos a algebraicos reales o complejos. Depende de si —en los polinomios que aparecen en la demostración— tomamos raíces reales o complejas.

Nota 6.19. Una vez que los conceptos de cardinalidad están claros, es incluso sencillo percatarse de que el teorema de Liouville permite detectar una cantidad no numerable de números trascendentes —los denominados *números de Liouville*—. Por ejemplo, en la nota 4.15 (página 250) y el ejercicio 4.12 —ambos del tema 4— se prueba que todos los números de la forma

$$\alpha = \sum_{j=1}^{\infty} \frac{d_j}{10^{j!}} \tag{6.10}$$

con $d_j \in \{0, 1, 2, \ldots, 9\}$ para cada j, y con infinitos d_j no nulos, son de Liouville, y por tanto trascendentes. Ahora, basta observar que el mismo argumento diagonal que usamos para probar que los números irracionales del intervalo $(0, 1)$ son no numerables sirve para probar que los números de la forma (6.10) constituyen un conjunto no numerable. ◁

6.4.2. Pero son difíciles de encontrar

A pesar del resultado de Cantor que muestra que, en cierto sentido, los trascendentes son mucho más numerosos que los algebraicos, sigue sin ser fácil encontrar números trascendentes, o probar la trascendencia de muchos números que aparecen habitualmente en matemáticas (o que podemos construir nosotros con un poco de habilidad). En este tema daremos algunas demostraciones específicas para varios números concretos como e o π (y existen diversas generalizaciones que también comentaremos); pero siempre son ingeniosos métodos dedicados en exclusiva a constantes de algún tipo. Aparte de eso, los únicos instrumentos genéricos son el teorema de Liouville antes citado, y el teorema de Thue-Siegel-Roth ([89]), que ya hemos mencionado en la sección 4.3 del tema 4. Para $n > 2$, este teorema es una clara mejora del de Liouville, pero su demostración es complicada y muy técnica; ni la vimos antes ni la veremos ahora (en [22, módulo 10] se muestra un esquema, y en [89], [66, capítulo 6] o [75, capítulo 1, § 5] está la demostración completa). Simplemente, por completitud, concluyamos esta sección enunciándolo:

Teorema 6.20. Thue-Siegel-Roth, 1955.
Si $\alpha \in \mathbb{R}$ es algebraico, entonces, para cada $\varepsilon > 0$, la ecuación

$$\left|\alpha - \frac{p}{q}\right| < \frac{1}{q^{2+\varepsilon}}$$

sólo tiene una cantidad finita de soluciones racionales p/q.



Charles Hermite (1822–1905)



Henri Padé (1863–1953)

Esto nos proporciona un criterio de trascendencia: si para algún $\varepsilon > 0$ hay infinitas soluciones racionales, el número α debe ser trascendente. En el tema 4 hemos analizado cómo este resultado permitía probar la trascendencia de algunos números expresados en forma de serie, y para los que el teorema de Liouville no resultaba de ninguna utilidad. Éste es el caso de $\sum_{k=1}^{\infty} a^{-m^k}$, donde a y m son enteros tales que $a \geq 2$ y $m \geq 3$. Pero no hay manera de aplicar el teorema de Thue-Siegel-Roth para probar la trascendencia de dos números tan importantes para la matemática como e o π.

En realidad, son relativamente muy pocos los números trascendentes que se podrían llegar a identificar mediante este criterio. En términos de teoría de la medida, sólo un conjunto de medida nula.

6.5. Trascendencia de e

Ya hemos probado que e es un número irracional. Llegaremos ahora más lejos, mostrando que también es trascendente. Esto —que fue demostrado por Charles Hermite en 1873, justo un año antes de que Cantor diera su espectacular resultado sobre la gran abundancia de números trascendentes— supuso un hecho notable, pues era la primera demostración de la trascendencia de un número importante en las matemáticas. Hasta entonces, las demostraciones de la trascendencia de algunos números se basaban en el teorema de Liouville (lo hemos visto en la sección anterior), y eran números construidos expresamente para ello. El método de Hermite, que más tarde serviría de germen para probar la trascendencia de otros números, estaba basado en propiedades analíticas de la función exponencial, y tendía un nexo entre el análisis y las propiedades aritméticas de los números. El grado de dificultad del trabajo de Hermite se puede observar en el siguiente texto, extraído de una

carta suya: «No me atrevo a intentar demostrar la trascendencia del número π. Si otros lo intentan, nadie se alegrará tanto como yo de que lo consigan, pero créeme, querido amigo, no dejará de costarles un cierto esfuerzo». En 1882, Ferdinand von Lindemann hizo este esfuerzo, y logró adaptar el planteamiento de Hermite al estudio de la trascendencia de π (lo veremos en la sección 6.7).

La demostración que mostramos aquí está basada en la original, aunque contiene simplificaciones posteriores debidas a otros autores (entre ellos, Karl Weierstrass y David Hilbert).[11]

Lema 6.21. Identidad de Hermite.

Dado R un polinomio con coeficientes reales, denotemos $I(x) = \int_0^x e^{x-t}R(t)\,dt$ y $H(x) = \sum_{j\geq 0} R^{(j)}(x)$ (que es una suma finita). Entonces

$$I(x) = \int_0^x e^{x-t}R(t)\,dt = e^x H(0) - H(x).$$

Demostración. Como

$$H = R + R' + R'' + \cdots,$$

Entonces,

$$H' = R' + R'' + \cdots.$$

Por tanto, $R = H - H'$ y, en consecuencia,

$$\begin{aligned} I(x) &= e^x \int_0^x (e^{-t}H(t) - e^{-t}H'(t))\,dt = e^x\left(-e^{-t}H(t)\right)\Big|_{t=0}^{x} \\ &= -H(x) + e^x H(0). \end{aligned}$$

□

Con esto, ya podemos probar la trascendencia de e:

Teorema 6.22. Hermite, 1873.
El número e es trascendente.

[11] En realidad, la prueba original de Hermite estaba mucho más integrada en lo que ahora denominamos aproximantes de Padé, hasta el punto de que se la considera como uno de los precedentes de esa teoría. Es más, en el marco de los aproximantes de Padé no resulta difícil entender cuál es la idea que subyace bajo las integrales involucradas en la demostración de la trascendencia de e y que, tal como las hemos presentado nosotros, parecen apariciones milagrosas. Por si el lector lo desconoce, digamos que los aproximantes de Padé de una función son sus aproximaciones por medio de funciones racionales, generalizando de este modo la aproximación por polinomios proporcionada por el desarrollo de Taylor.

DEMOSTRACIÓN. Supongamos que e fuera algebraico. En ese caso, existirían $m \in \mathbb{N}$ y $q_0, q_1, \ldots, q_m \in \mathbb{Z}$, con $q_0 \neq 0$, tales que

$$q_0 + q_1 e + q_2 e^2 + \cdots + q_m e^m = 0.$$

Para p un primo que verifique $p > |q_0|$ y $p > m$, tomemos $R(x)$ el polinomio

$$R(x) = x^{p-1}(x-1)^p(x-2)^p \cdots (x-m)^p,$$

y definamos $I(x)$ como en el lema 6.21. Asimismo, sea

$$A = q_0 I(0) + q_1 I(1) + \cdots + q_m I(m).$$

En estas circunstancias, probaremos que existen constantes λ y μ, independientes de p, tales que

$$(p-1)! \leq |A| \leq \lambda \mu^p. \tag{6.11}$$

En particular, $(p-1)! \leq \lambda\mu^p$, y esto no es posible para p suficientemente grande, ya que $\mu^p/(p-1)! \to 0$ cuando $p \to \infty$. Así, hemos llegado a un absurdo, y el teorema estaría probado.

En primer lugar, abordemos la desigualdad izquierda de (6.11). Según el lema 6.21 (y denotando H como allí), tenemos

$$\begin{aligned} A &= \sum_{k=0}^{m} q_k I(k) = \sum_{k=0}^{m} (q_k e^k H(0) - q_k H(k)) \\ &= H(0) \sum_{k=0}^{m} q_k e^k - \sum_{k=0}^{m} q_k H(k) = -\sum_{k=0}^{m} q_k H(k). \end{aligned} \tag{6.12}$$

Usaremos ahora, para cada k, la descomposición $H(k) = \sum_{j \geq 0} R^{(j)}(k)$ (como R es un polinomio, estas sumas son finitas). Para $k \neq 0$, apliquemos el lema 6.8 (página 346) con $n = p$, $c = k$ y $\widetilde{f}(x) = R(x) = (x-k)^p f(x)$, donde f es el polinomio $f(x) = R(x)/(x-k)^p$; entonces,

$$\begin{aligned} &R^{(j)}(k) = 0 \text{ si } j < p, \\ &R^{(j)}(k) \text{ es múltiplo de } p! \text{ si } j \geq p. \end{aligned} \tag{6.13}$$

Para $k = 0$, apliquemos el lema 6.8 con $n = p-1$, $c = 0$ y $\widetilde{f}(x) = R(x) = x^{p-1} f(x)$, siendo $f(x) = R(x)/x^{p-1}$; así,

$$\begin{aligned} &R^{(j)}(0) = 0 \text{ si } j < p-1, \\ &R^{(p-1)}(0) \text{ es múltiplo de } (p-1)!, \\ &R^{(j)}(0) \text{ es múltiplo de } p! \text{ si } j \geq p. \end{aligned} \tag{6.14}$$

Además, si multiplicamos $(x-1)^p(x-2)^p\cdots(x-m)^p$ para obtener su desarrollo en potencias de x, es claro que su término independiente es $(-1)^{mp}(m!)^p$ (y éste será el coeficiente de x^{p-1} en $R(x)$). Por tanto,

$$R^{(p-1)}(0) = (p-1)!\,(-1)^{mp}(m!)^p,$$

luego $R^{(p-1)}(0)$ no puede ser múltiplo de $p!$ (recordemos que hemos tomado $p > m$ y primo). Escribamos la expresión (6.12) destacando el sumando correspondiente a $k = 0$, esto es,

$$A = -q_0\big(R(0) + R'(0) + \cdots + R^{(p-1)}(0) + \cdots\big) - \sum_{k=1}^{m} q_k H(k).$$

Por ser $q_0 \neq 0$ y, además, $p > |q_0|$, todos los sumandos excepto $-q_0R^{(p-1)}(0)$ son múltiplos de $p!$, y todos son múltiplos de $(p-1)!$. Así pues, A es múltiplo de $(p-1)!$ pero no de $p!$; en particular, $A \neq 0$. Como consecuencia, $(p-1)! \leq |A|$.

Concluyamos probando la desigualdad de la derecha en (6.11). Para $0 \leq k \leq m$, tenemos

$$\begin{aligned} |I(k)| &= \left|\int_0^k e^{k-t}R(t)\,dt\right| \leq e^k k \cdot \sup_{0\leq t\leq k}|R(t)| \\ &\leq e^m m \cdot \sup_{0\leq t\leq m}|R(t)| \leq e^m m \cdot m^{p-1+mp} = e^m(m^{1+m})^p. \end{aligned}$$

De aquí,

$$\begin{aligned} |A| &\leq |q_0|\cdot|I(0)| + |q_1|\cdot|I(1)| + \cdots + |q_m|\cdot|I(m)| \\ &\leq \left(\sum_{k=0}^{m}|q_k|\right)e^m(m^{1+m})^p = \lambda\mu^p, \end{aligned}$$

tal como queríamos demostrar. □

6.6. Polinomios simétricos

Nuestro siguiente objetivo es probar la trascendencia de π. La demostración de este hecho está inspirada en la de la trascendencia de e, pero es más complicada y requiere utilizar más herramientas. En particular, usaremos diversos resultados sobre polinomios simétricos. En esta sección, introduciremos todo lo necesario para —en la siguiente— abordar la demostración de la trascendencia de π.

Comencemos recordando las *reglas de Cardano*. Dado un polinomio $f(x) = a_\nu x^\nu + a_{\nu-1}x^{\nu-1} + \cdots + a_1 x + a_0$ (con $a_\nu \neq 0$) cuyas raíces en $\mathbb{C}$ —repetidas según su multiplicidad— son $\gamma_1, \gamma_2, \ldots, \gamma_\nu$, podemos poner

$$f(x) = a_\nu(x-\gamma_1)(x-\gamma_2)\cdots(x-\gamma_{\nu-1})(x-\gamma_\nu).$$

Multiplicando e igualando coeficientes obtenemos las relaciones

$$\begin{aligned} a_{\nu-1} &= -a_\nu(\gamma_1+\gamma_2+\cdots+\gamma_\nu),\\ a_{\nu-2} &= (-1)^2 a_\nu(\gamma_1\gamma_2+\gamma_1\gamma_3+\cdots+\gamma_{\nu-1}\gamma_\nu),\\ &\cdots\\ a_0 &= (-1)^\nu a_\nu \gamma_1\gamma_2\cdots\gamma_\nu, \end{aligned}$$

habitualmente denominadas reglas de Cardano. Obsérvese el papel que juega el coeficiente director a_ν en estas fórmulas; aunque, en muchas ocasiones, no tiene demasiada importancia, pues se puede prescindir de él sin más que utilizar polinomios mónicos (esto es, con $a_\nu = 1$).

Supongamos ahora que tenemos un polinomio $f(x) = bx^\nu + \cdots \in \mathbb{Z}[x]$ (sólo hemos detallado el término de mayor grado, un $b \neq 0$), con su raíces $\gamma_1, \gamma_2, \ldots, \gamma_\nu \in \mathbb{C}$. De las reglas de Cardano se deduce que

$$\begin{aligned} &b(\gamma_1+\gamma_2+\cdots+\gamma_\nu) \in \mathbb{Z},\\ &b(\gamma_1\gamma_2+\gamma_1\gamma_3+\cdots+\gamma_{\nu-1}\gamma_\nu) \in \mathbb{Z},\\ &\cdots\\ &b\gamma_1\gamma_2\cdots\gamma_\nu \in \mathbb{Z}. \end{aligned}$$

Mostremos también la siguiente variante: si $f(x) \in \mathbb{Q}[x]$ y $\gamma_1, \gamma_2, \ldots, \gamma_n \in \mathbb{C}$ son todas sus raíces (repetidas según su multiplicidad, por supuesto), entonces

$$\begin{aligned} &\gamma_1+\gamma_2+\cdots+\gamma_\nu \in \mathbb{Q},\\ &\gamma_1\gamma_2+\gamma_1\gamma_3+\cdots+\gamma_{\nu-1}\gamma_\nu \in \mathbb{Q},\\ &\cdots\\ &\gamma_1\gamma_2\cdots\gamma_\nu \in \mathbb{Q}. \end{aligned} \tag{6.15}$$

Ejemplo. Dado que el polinomio $f(x) = x^4 - 2 \in \mathbb{Q}[x]$ se puede expresar como

$$f(x) = (x^2-\sqrt{2})(x^2+\sqrt{2}) = (x-\sqrt[4]{2})(x+\sqrt[4]{2})(x-i\sqrt[4]{2})(x+i\sqrt[4]{2}),$$

resulta claro que $f(x)$ es irreducible en $\mathbb{Q}[x]$, y que sus raíces en $\mathbb{C}$ son $\pm\sqrt[4]{2}$ y $\pm i\sqrt[4]{2}$. Si denotamos dichas raíces mediante $\gamma_1, \gamma_2, \gamma_3, \gamma_4$, las reglas de Cardano aseguran que

$$\begin{aligned}
&\gamma_1 + \gamma_2 + \gamma_3 + \gamma_4 = 0,\\
&\gamma_1\gamma_2 + \gamma_1\gamma_3 + \gamma_1\gamma_4 + \gamma_2\gamma_3 + \gamma_2\gamma_4 + \gamma_3\gamma_4 = 0,\\
&\gamma_1\gamma_2\gamma_3 + \gamma_1\gamma_2\gamma_4 + \gamma_1\gamma_3\gamma_4 + \gamma_2\gamma_3\gamma_4 = 0,\\
&\gamma_1\gamma_2\gamma_3\gamma_4 = -2,
\end{aligned}$$

lo cual puede comprobarse de manera directa sin ninguna dificultad. ◊

Cuando $\gamma \in \mathbb{C}$ es un número algebraico y f el polinomio no nulo de menor grado con coeficientes enteros (o racionales, da igual) tal que $f(\gamma) = 0$, además de $\gamma_1 = \gamma$, las otras raíces de f (repetidas según su multiplicidad) serán $\gamma_2, \ldots, \gamma_\nu \in \mathbb{C}$, donde ν es el grado de f. En estas circunstancias, a menudo se dice que $\gamma_2, \ldots, \gamma_\nu \in \mathbb{C}$ son las *raíces conjugadas* de γ_1; o, también, que $\gamma_1, \gamma_2, \ldots, \gamma_\nu \in \mathbb{C}$ son *números algebraicos conjugados*.[12]

En realidad, cuando tenemos un número algebraico $\gamma_1 \in \mathbb{C}$, junto con $\gamma_2, \ldots, \gamma_\nu \in \mathbb{C}$ tales que $\gamma_1, \gamma_2, \ldots, \gamma_\nu \in \mathbb{C}$ son raíces de $f \in \mathbb{Q}[x]$ de grado ν, para muchos de los resultados que veremos no hace falta asegurar que f sea irreducible. Así pues, y salvo que digamos lo contrario, habitualmente nos nos preocuparemos de que el polinomio f del que un número algebraico es raíz sea irreducible (es decir, puede que exista otro polinomio con coeficientes en $\mathbb{Q}$ y grado menor del que γ_1 es raíz, junto con algunas otras de entre $\gamma_2, \ldots, \gamma_\nu$).

Por fin, establezcamos qué entendemos por *polinomio simétrico*: Decimos que un polinomio en n variables $P(x_1, x_2, \ldots, x_n)$ es simétrico si, para cualquier permutación σ de los índices $\{1, 2, \ldots, n\}$, se tiene

$$P(x_{\sigma(1)}, x_{\sigma(2)}, \ldots, x_{\sigma(n)}) = P(x_1, x_2, \ldots, x_n).$$

Ejemplos de polinomios simétricos son los siguientes:

$$\begin{aligned}
&x_1 + x_2 + \cdots + x_n,\\
&x_1x_2 + x_1x_3 + \cdots + x_{n-1}x_n,\\
&\cdots\\
&x_1x_2\cdots x_n,
\end{aligned}$$

[12] Esto es algo distinto al uso más específico que se da habitualmente en $\mathbb{C}$ a la palabra «conjugados», que se refiere sólo a la pareja de números complejos $a \pm bi$ (con $a, b \in \mathbb{R}$, $b \neq 0$), y que son las raíces del polinomio $(x-a)^2 + b^2$, irreducible en $\mathbb{R}[x]$.

los denominados *polinomios simétricos elementales*. No hay que esforzarse mucho para darse cuenta de la similitud entre estas expresiones y las que nos aparecían en las reglas de Cardano.

6.6.1. El teorema fundamental de los polinomios simétricos

El resultado básico en este campo es el que habitualmente se conoce con el nombre de *teorema fundamental de los polinomios simétricos*: cualquier polinomio simétrico puede escribirse como combinación de los polinomios simétricos elementales. Su demostración, así como otras propiedades relacionadas, se puede encontrar en bastantes textos de álgebra o de teoría de números. Por citar alguno, véase [78, A.2, p. 484]. En realidad, la prueba no es complicada, así que aquí presentaremos una, sin detenernos más que en lo imprescindible para llevarla a cabo.

Comencemos introduciendo un poco de notación. Dado un monomio en n variables $ax_1^{\lambda_1}x_2^{\lambda_2}\cdots x_n^{\lambda_n}$ (siendo a una constante no nula y $\lambda_j \in \mathbb{N}\cup\{0\}$ para $1 \leq j \leq n$), su grado es $\lambda_1 + \lambda_2 + \cdots + \lambda_n$. El grado de un polinomio $P(x_1, x_2, \ldots, x_n)$ es el mayor de los grados de los monomios que lo forman. Se dice que un polinomio es *homogéneo* cuando todos sus monomios tienen el mismo grado. Obviamente, cualquier polinomio se puede escribir como suma de polinomios homogéneos de distinto grado. Además, si el polinomio es simétrico, sus componentes homogéneas también lo son, pues las permutaciones de las variables no cambian el grado.

Con esto, vayamos ya con el teorema:

Teorema 6.23. Teorema fundamental de los polinomios simétricos. *Si $P(x_1, x_2, \ldots, x_n)$ es un polinomio simétrico en n variables con coeficientes en $\mathbb{Z}$ (o en $\mathbb{Q}$), entonces puede escribirse*

$$P(x_1, x_2, \ldots, x_n) = Q(y_1, y_2, \ldots, y_n)$$

para cierto polinomio Q con coeficientes en $\mathbb{Z}$ (o en $\mathbb{Q}$) del mismo grado que P o menor, y donde

$$\begin{aligned} y_1 &= x_1 + x_2 + \cdots + x_n, \\ y_2 &= x_1x_2 + x_1x_3 + \cdots + x_{n-1}x_n, \\ &\cdots \\ y_n &= x_1x_2\cdots x_n. \end{aligned}$$

Demostración. Como cada polinomio es suma de polinomios homogéneos, podemos suponer que $P(x_1, x_2, \ldots, x_n)$ es un polinomio homogéneo. Entre monomios no

nulos —en particular, entre los monomios que forman el polinomio homogéneo P— podemos definir el siguiente orden:

$$ax_1^{\lambda_1}x_2^{\lambda_2}\cdots x_n^{\lambda_n} \succ bx_1^{\mu_1}x_2^{\mu_2}\cdots x_n^{\mu_n}$$

si

$$\lambda_1=\mu_1,\quad \lambda_2=\mu_2,\quad \ldots,\quad \lambda_k=\mu_k,\quad \text{y}\quad \lambda_{k+1}>\mu_{k+1}$$

para algún $k\geq 0$. Sea $ax_1^{\lambda_1}x_2^{\lambda_2}\cdots x_n^{\lambda_n}$ el monomio de mayor orden de los que forman P. Deberá ser $\lambda_1\geq\lambda_2\geq\cdots\geq\lambda_n$ (como P es homogéneo, cualquier permutación de las variables debe dar otro monomio de los de P, y estamos eligiendo el de mayor orden). Consideremos ahora

$$Q_1(y_1,y_2,\ldots,y_n)=ay_1^{\lambda_1-\lambda_2}y_2^{\lambda_2-\lambda_3}\cdots y_{n-1}^{\lambda_{n-1}-\lambda_n}y_n^{\lambda_n}$$

que, al sustituir en las variables x_j, es un polinomio homogéneo del mismo grado que P, y su monomio de mayor orden es

$$ax_1^{\lambda_1-\lambda_2}(x_1x_2)^{\lambda_2-\lambda_3}\cdots(x_1\cdots x_{n-1})^{\lambda_{n-1}-\lambda_n}(x_1\cdots x_n)^{\lambda_n}=ax_1^{\lambda_1}x_2^{\lambda_2}\cdots x_n^{\lambda_n}.$$

Por consiguiente, el nuevo polinomio simétrico

$$P_1(x_1,x_2,\ldots,x_n)=P(x_1,x_2,\ldots,x_n)-ay_1^{\lambda_1-\lambda_2}y_2^{\lambda_2-\lambda_3}\cdots y_{n-1}^{\lambda_{n-1}-\lambda_n}y_n^{\lambda_n}$$

sigue siendo homogéneo, y al restar se cancela el monomio de mayor orden. Es decir, el monomio de mayor orden de P_1 tiene orden estrictamente menor que el de P. Este proceso se lo podemos volver a aplicar a P_1. Como en cada paso disminuimos el orden, en un número finito de pasos llegaremos a un polinomio nulo. Reuniendo todos los sumandos del tipo $Q_1(y_1,y_2,\ldots,y_n)$ que aparecen, se obtiene el polinomio $Q(y_1,y_2,\ldots,y_n)$ buscado.

Para finalizar, veamos que el grado de Q es menor o igual que el de P. Para ello, basta darse cuenta de que, tal como lo hemos construido, Q_1 tiene el mismo grado que P como polinomio en las variables x_j, y que su grado como polinomio en las variables y_j —que es λ_1, y esto es lo que nos interesa— puede ser menor, pero no mayor. □

Ejemplo. Sea el polinomio $P(x,y,z)=a+bx+by+bz+cx^2+cy^2+cz^2+dxy+dxz+dyz+xyz$ que es, obviamente, simétrico. Podemos escribir

$$\begin{aligned}P(x,y,z)&=a+b(x+y+z)+c(x+y+z)^2+(d-2c)(xy+xz+yz)+xyz\\&=Q(x+y+z,xy+xz+yz,xyz),\end{aligned}$$

con $Q(u,v,w) = a + bu + cu^2 + (d-2c)v + w$. En particular, nótese que en este ejemplo el grado de Q es estrictamente menor que el de P: el polinomio P tiene grado 3, y Q grado 2. ◊

Corolario 6.24.

Si $P(x_1, x_2, \ldots, x_n)$ es un polinomio simétrico en n variables con coeficientes racionales, y $\gamma_1, \gamma_2, \ldots, \gamma_n \in \mathbb{C}$ son las raíces de un polinomio en $\mathbb{Q}[x]$, entonces

$$P(\gamma_1, \gamma_2, \ldots, \gamma_n) \in \mathbb{Q}.$$

Demostración. Sea $f \in \mathbb{Q}[x]$ el polinomio mónico cuyas raíces son $\gamma_1, \gamma_2, \ldots, \gamma_n$. Definamos y_j a partir de las γ_j como en el teorema 6.23, es decir,

$$\begin{aligned} y_1 &= \gamma_1 + \gamma_2 + \cdots + \gamma_n, \\ y_2 &= \gamma_1\gamma_2 + \gamma_1\gamma_3 + \cdots + \gamma_{n-1}\gamma_n, \\ &\cdots \\ y_n &= \gamma_1\gamma_2\cdots\gamma_n. \end{aligned}$$

Por las reglas de Cardano —en concreto, bastaría aplicar (6.15)—, los y_j son números racionales (salvo el signo, coinciden con los coeficientes de f). Por tanto, aplicando el teorema fundamental de los polinomios simétricos (teorema 6.23), tenemos

$$P(\gamma_1, \gamma_2, \ldots, \gamma_n) = Q(y_1, y_2, \ldots, y_n) \in \mathbb{Q}.$$ □

6.6.2. Algunas consecuencias sencillas, pero importantes

En nuestro camino para demostrar la trascendencia de π, también emplearemos algunos resultados sobre números algebraicos. En realidad, estos resultados van a ser consecuencia de los que ya hemos visto sobre polinomios simétricos.

Lema 6.25.

Sea α un número algebraico e $i = \sqrt{-1}$ la unidad imaginaria. Entonces, $i\alpha$ es algebraico.

Demostración. Sea $f(x) \in \mathbb{Q}[x]$ tal que $f(\alpha) = 0$; sin pérdida de generalidad, lo podemos suponer mónico. Descomponiendo f en factores, tendremos

$$f(x) = \prod_{j=1}^{n} (x - \alpha_j)$$

para ciertos números complejos α_j, $j = 1, 2, \ldots, n$, y donde $\alpha_1 = \alpha$. Tomemos el polinomio

$$g(x) = \prod_{j=1}^{n} (x - i\alpha_j)(x + i\alpha_j) = \prod_{j=1}^{n} (x^2 + \alpha_j^2). \tag{6.16}$$

Es claro que $g(i\alpha) = 0$; así, si probamos que $g(x) \in \mathbb{Q}[x]$, ya tendremos que $i\alpha$ es algebraico.

Desarrollando la expresión de la derecha en (6.16), aparece un sumatorio de la forma

$$g(x) = \sum_{k=0}^{n} P_k(\alpha_1, \alpha_2, \ldots, \alpha_n) x^{2k},$$

donde los P_k resultan ser polinomios simétricos en $\alpha_1, \alpha_2, \ldots, \alpha_n$ (en realidad, todos esos α_j salen con la forma α_j^2, pero eso no tiene ahora importancia) y cuyos coeficientes son números enteros. A su vez, los números $\alpha_1, \alpha_2, \ldots, \alpha_n$ son las raíces de $f(x) \in \mathbb{Q}[x]$ (en principio, no podemos asegurar que f sea irreducible, pero no hace falta). Así pues, los P_k están en las circunstancias del corolario 6.24, y por tanto $P_k(\alpha_1, \alpha_2, \ldots, \alpha_n) \in \mathbb{Q}$, tal como queríamos probar. □

A decir verdad, podíamos haber probado el lema sin necesidad de habernos apoyado en resultados sobre polinomios simétricos. Y es que, una vez conocido $f(x) \in \mathbb{Q}[x]$ tal que $f(\alpha) = 0$, resulta fácil construir un polinomio que se anula en $i\alpha$; pero no de cualquier manera, sino siguiendo una serie de pasos que nos garantizan que los coeficientes del nuevo polinomio siguen estando en $\mathbb{Q}$. Veámoslo:

Demostración alternativa del lema 6.25. Como antes, sea $f(x) \in \mathbb{Q}[x]$ el polinomio mónico que satisface $f(\alpha) = 0$, y sean α_j, $j = 1, 2, \ldots, n$, todas sus raíces complejas, con $\alpha_1 = \alpha$. Escribamos

$$f(x) = \prod_{j=1}^{n} (x - \alpha_j) = \sum_{k=0}^{n} a_k x^k$$

y tomemos $\widetilde{f}(x) = (-1)^n \sum_{k=0}^{n} (-1)^k a_k x^k$ (el coeficiente $(-1)^n$ lo añadimos para que $\widetilde{f}$ sea mónico). Como $f(x) \in \mathbb{Q}[x]$, es obvio que también $\widetilde{f}(x) \in \mathbb{Q}[x]$; además, podemos averiguar fácilmente sus raíces:

$$\begin{aligned} \widetilde{f}(-\alpha_j) = (-1)^n \sum_{k=0}^{n} (-1)^k a_k (-\alpha_j)^k &= (-1)^n \sum_{k=0}^{n} a_k \alpha_j^k \\ &= (-1)^n f(\alpha_j) = 0, \quad j = 1, 2, \ldots, n. \end{aligned}$$

En consecuencia, $\widetilde{f}(x) = \prod_{j=1}^{n}(x+\alpha_j)$. Multiplicando los polinomios f y $\widetilde{f}$ tendremos un nuevo polinomio

$$f(x)\widetilde{f}(x) = \prod_{j=1}^{n}(x^2-\alpha_j^2) = \sum_{k=0}^{n} b_k x^{2k} \in \mathbb{Q}[x].$$

Entonces, el polinomio

$$g(x) = \sum_{k=0}^{n}(-1)^k b_k x^{2k}$$

está en $\mathbb{Q}[x]$ y verifica

$$g(i\alpha) = \sum_{k=0}^{n}(-1)^k b_k (i\alpha)^{2k} = \sum_{k=0}^{n} b_k \alpha^{2k} = f(\alpha)\widetilde{f}(\alpha) = 0$$

luego, efectivamente, $i\alpha$ es algebraico. □

Al lector que tenga suficientes conocimientos de teoría de cuerpos, el lema anterior le ha podido parecer una trivialidad. Es más, el lema también es consecuencia inmediata del siguiente hecho:

$$\alpha, \beta \text{ algebraicos } \Rightarrow \ \alpha+\beta \text{ y } \alpha\beta \text{ algebraicos}, \tag{6.17}$$

que es una propiedad muy bien conocida en el marco de la teoría de cuerpos. En nuestra descarga, podríamos haber señalado que hemos incluido el lema por completitud, pensando en los lectores con menos conocimientos algebraicos.

Pero no sólo es eso. Lo que en realidad ocurre es que se suele probar (6.17) usando métodos que tienen sus cimientos en el uso de polinomios simétricos. Al estudiar teoría de cuerpos en general, esto puede quedar más o menos escondido, pero la idea subyacente es la misma.

Así pues, y aunque no lo vayamos a usar en el resto del libro, vamos ahora a probar (6.17), un resultado importante por sí mismo. Es más, lo vamos a demostrar de manera bastante más general:

Teorema 6.26.
Sean α y β dos números algebraicos, y sea $\zeta(x,y)$ un polinomio con coeficientes en $\mathbb{Q}$. En estas condiciones, $\zeta(\alpha,\beta)$ es un número algebraico.

DEMOSTRACIÓN. Sean $f, g \in \mathbb{Q}[x]$ mónicos tales que $f(\alpha) = g(\beta) = 0$, y descompongamos (en $\mathbb{C}[x]$)

$$f(x) = \prod_{j=1}^{n}(x - \alpha_j), \qquad \alpha_j \in \mathbb{C},\ j = 1, \dots, n, \qquad \alpha_1 = \alpha,$$

$$g(x) = \prod_{k=1}^{m}(x - \beta_k), \qquad \beta_k \in \mathbb{C},\ k = 1, \dots, m, \qquad \beta_1 = \beta.$$

Tomemos ahora

$$h(x) = \prod_{j=1}^{n}\prod_{k=1}^{m}\big(x - \zeta(\alpha_j, \beta_k)\big).$$

Obviamente, $h(\zeta(\alpha, \beta)) = 0$; en consecuencia, nos basta con ver que $h(x) \in \mathbb{Q}[x]$.

Si desarrollamos $h(x)$, quedará

$$h(x) = \sum_{l=0}^{nm} P_l(\alpha_1, \dots, \alpha_n, \beta_1, \dots, \beta_m)x^l$$

para ciertos polinomios P_l simétricos en $\alpha_1, \dots, \alpha_n, \beta_1, \dots, \beta_m$ y de coeficientes enteros. Pero $\alpha_1, \dots, \alpha_n, \beta_1, \dots, \beta_m$ son las raíces de $f(x)g(x) \in \mathbb{Q}[x]$. Aplicando el corolario 6.24 obtenemos $P_l(\alpha_1, \dots, \alpha_n, \beta_1, \dots, \beta_m) \in \mathbb{Q}$, y por tanto $h(x) \in \mathbb{Q}[x]$, como queríamos probar. □

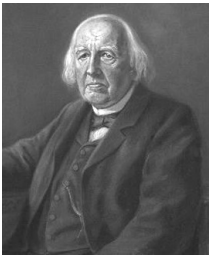
6.7. Trascendencia de π

La trascendencia de π fue establecida por Ferdinand von Lindemann en 1882. El problema de determinar la naturaleza aritmética de π es muy antiguo, pues está relacionado con la cuadratura del círculo, que provenía de la época griega. Se trataba de determinar si, mediante regla y compás, se podía construir un cuadrado de área igual a la de un círculo dado. Suponiendo que el círculo tiene radio 1, esto es equivalente a construir un cuadrado de lado $\sqrt{\pi}$. Con regla y compás, sólo se pueden efectuar operaciones algebraicas; más concretamente, sumas, restas, multiplicaciones, divisiones, y raíces cuadradas. Entonces, la trascendencia de π da la respuesta: ni π ni —por tanto— $\sqrt{\pi}$ son algebraicos, luego no se pueden construir con regla y compás, así que la cuadratura del círculo es imposible.[13] ¡Un argumento analítico

[13] Tres eran los problemas clásicos de la matemática griega, cuya solución había permanecido abierta desde que fueron planteados, y cuyo estudio tuvo una gran influencia en el desarrollo de la



Ferdinand von Lindemann (1852–1939)



Karl Weierstrass (1815–1897)

de teoría de números ha logrado resolver un problema geométrico planteado dos milenios antes!

Como ya hemos comentado, la demostración de la trascendencia de π está inspirada en la de Hermite de la trascendencia de e, aunque es más complicada. Para que se pueda observar mejor el paralelismo entre ambas demostraciones, comentaremos —mediante notas a pie de página— algunas de sus similitudes y diferencias.

Previamente, un lema, en el que se muestra una propiedad de los coeficientes de las derivadas de un polinomio. Incluso prescindiendo del resultado en sí, merece la pena fijarse en su bonita demostración.

Lema 6.27.
Sean $f(x) \in \mathbb{Z}[x]$ y $j \geq 0$. Entonces, todos los coeficientes de $f^{(j)}(x)$ son múltiplos de $j!$.

Demostración. Por linealidad, basta hacerla para $f(x) = x^s$ con $s \in \mathbb{N} \cup \{0\}$ cualquiera. Ahora, distingamos dos casos: Si $j > s$, $f^{(j)}(x)$ es el polinomio nulo y

geometría. Los otros dos eran la duplicación del cubo (construir un cubo de volumen doble al de uno dado) y la trisección del ángulo (dividir un ángulo cualquiera en tres partes iguales); en ambos se requería hacerlo con regla y compás. En 1837, Pierre Wantzel se había servido de las recientes ideas de la teoría de Galois para demostrar la irresolubilidad de estos dos problemas. Wantzel probó que un número construible con regla y compás debía ser raíz de un polinomio de coeficientes enteros y grado 2^n (para algún entero positivo n). Así, la duplicación del cubo es imposible, pues eso equivale a construir $\sqrt[3]{2}$, que es raíz del polinomio $x^3 - 2$, de grado 3 e irreducible. En cuanto a la trisección del ángulo, si $\pi/3$ se pudiera trisecar, $2\cos(\pi/9)$ sería construible; pero argumentos trigonométricos sencillos (o el uso del polinomio de Chebyshev $T_3(x)$) muestran que $2\cos(\pi/9)$ es raíz de $x^3 - 3x - 1$, de nuevo un polinomio irreducible de grado 3.

el resultado es —trivialmente— cierto. Si $j \leq s$, entonces

$$f^{(j)}(x) = s(s-1)\cdots(s-j+1)x^{s-j} = j!\binom{s}{j}x^{s-j};$$

como $\binom{s}{j} \in \mathbb{N}$, obtenemos lo que buscábamos. □

Y otro lema, cuyo objetivo es, únicamente, descargar la demostración del teorema:

Lema 6.28.

Sean $\alpha_1, \alpha_2, \ldots, \alpha_m \in \mathbb{C}$ tales que

$$\begin{aligned}
\sigma_1 &= \alpha_1 + \alpha_2 + \cdots + \alpha_m \in \mathbb{Q},\\
\sigma_2 &= \alpha_1\alpha_2 + \alpha_1\alpha_3 + \cdots + \alpha_{m-1}\alpha_m \in \mathbb{Q},\\
&\cdots\\
\sigma_m &= \alpha_1\alpha_2\cdots\alpha_m \in \mathbb{Q},
\end{aligned}$$

y elijamos $b \in \mathbb{Z}$ de modo que $b\sigma_1, b\sigma_2, \ldots, b\sigma_m \in \mathbb{Z}$. Entonces, para cualquier polinomio $\Phi \in \mathbb{Z}[x]$ de grado $\leq d$, se cumple

$$b^d \sum_{k=1}^{m} \Phi(\alpha_k) \in \mathbb{Z}.$$

Demostración. Consideremos $P(x_1, \ldots, x_m) = \sum_{k=1}^{m} \Phi(x_k)$, un polinomio en m variables $x_1, \ldots, x_m$ y coeficientes enteros; trivialmente, $P(x_1, \ldots, x_m)$ es un polinomio simétrico de grado $\leq d$. Si denotamos

$$\begin{aligned}
y_1 &= x_1 + x_2 + \cdots + x_m,\\
y_2 &= x_1x_2 + x_1x_3 + \cdots + x_{m-1}x_m,\\
&\cdots\\
y_m &= x_1x_2\cdots x_m,
\end{aligned}$$

según el teorema 6.23 podremos escribir

$$\sum_{k=1}^{m} \Phi(x_k) = Q(y_1, y_2, \ldots, y_m)$$

para cierto polinomio Q con coeficientes en $\mathbb{Z}$ y grado $\leq d$. Luego

$$b^d \sum_{k=1}^{m} \Phi(\alpha_k) = b^d Q(\sigma_1, \sigma_2, \ldots, \sigma_m) = b^d Q\Big(\frac{b\sigma_1}{b}, \frac{b\sigma_2}{b}, \ldots, \frac{b\sigma_m}{b}\Big) \in \mathbb{Z},$$

ya que $b\sigma_1, b\sigma_2, \ldots, b\sigma_m \in \mathbb{Z}$. □

6.7.1. Esquema de la demostración

La demostración de la trascendencia de π es un poco larga. Para facilitar su comprensión, veamos primero un esquema.

Esquema de la demostración. Supongamos que π es algebraico. Tras esto, los pasos que seguiremos serán los siguientes:

(a) Encontraremos un $a \in \mathbb{N}$ y unos números $\alpha_1, \dots, \alpha_m \in \mathbb{C}$ tales que[14]

$$0 = a + e^{\alpha_1} + \cdots + e^{\alpha_m},$$

(b) y $\alpha_1, \dots, \alpha_m$ son las raíces de un polinomio $\Psi(x) \in \mathbb{Z}[x]$, esto es,

$$\Psi(x) = b(x-\alpha_1)\cdots(x-\alpha_m), \quad b \in \mathbb{Z} \setminus \{0\}.$$

(c) Tomaremos $R(x) = b^{mp-1}x^{p-1}\Psi(x)^p$ con p un primo grande, y $H(x) = R(x) + R'(x) + R''(x) + \cdots$. Usando la identidad de Hermite, probaremos

$$-aH(0) - \sum_{k=1}^{m} H(\alpha_k) = \sum_{k=1}^{m} e^{\alpha_k} \int_0^{\alpha_k} e^{-t} R(t)\, dt.$$

(d) Veremos que $(p-1)! \leq \left|-aH(0) - \sum_{k=1}^{m} H(\alpha_k)\right|$

(e) y que

$$\left| \sum_{k=1}^{m} e^{\alpha_k} \int_0^{\alpha_k} e^{-t} R(t)\, dt \right| \leq \lambda \mu^p$$

para ciertas constantes λ y μ que no dependen de p.

En consecuencia, $(p-1)! \leq \lambda\mu^p$; pero esto no puede ser cierto cuando p es suficientemente grande, así que hemos llegado a una contradicción. □

[14] En la demostración de la trascendencia de e, partíamos de algo parecido:

$$0 = q_0 + q_1 e + q_2 e^2 + \cdots + q_m e^m; \quad q_0, q_1, \dots, q_m \in \mathbb{Z},$$

que se obtenía directamente de haber supuesto que e es algebraico. En principio, con π hubiéramos encontrado un polinomio en π que se anula, pero la demostración no requiere eso, sino de nuevo —quizás sorprendentemente— una expresión (parecida a un polinomio) en e (y es que no aparece π en la identidad de Hermite, sino e); para obtenerla, resultará crucial utilizar la relación $e^{i\pi} + 1 = 0$ que liga π con e. Así, en la demostración correspondiente a la trascendencia de π, los exponentes $1, 2, \dots, m$ que teníamos con e están cambiados por $\alpha_1, \alpha_2, \dots, \alpha_m$. Ahora, hay que conseguir que las $\alpha_1, \alpha_2, \dots, \alpha_m$ sean las raíces de un polinomio con coeficientes en $\mathbb{Z}$ (el $\Psi(x)$ de la parte (b), a partir del cual se define $R(x)$ en la (c)); con $1, 2, \dots, m$ esto era automático. Una vez hallados los correspondientes polinomios $R(x)$, ambas demostraciones tienen la misma estructura —aunque, con π, los detalles son más complicados.

6.7.2. La demostración

Vamos por fin con la demostración en sí; pero, previamente, recordemos el enunciado:

Teorema 6.29. Lindemann, 1882.
El número π es trascendente.

Demostración. Seguiremos los pasos (a), ..., (e) del esquema anterior. Comencemos suponiendo que π es algebraico, con lo cual, por el lema 6.25, también $\gamma = i\pi$ será algebraico.

(a) Sean $\gamma_2, \dots, \gamma_\nu$ los conjugados de γ y $\gamma_1 = \gamma$. Como $e^\gamma + 1 = e^{i\pi} + 1 = 0$, resulta trivial que $\prod_{j=1}^{\nu}(1 + e^{\gamma_j}) = 0$. De aquí,

$$0 = \prod_{j=1}^{\nu}(1 + e^{\gamma_j}) = \prod_{j=1}^{\nu}(e^{0\cdot\gamma_j} + e^{1\cdot\gamma_j}) = \sum_{\varepsilon_1=0,1}\sum_{\varepsilon_2=0,1}\cdots\sum_{\varepsilon_\nu=0,1} e^{\varepsilon_1\gamma_1+\varepsilon_2\gamma_2+\cdots+\varepsilon_\nu\gamma_\nu}.$$

Algunos de los exponentes $\varepsilon_1\gamma_1 + \varepsilon_2\gamma_2 + \cdots + \varepsilon_\nu\gamma_\nu$ pueden ser 0 (desde luego, al menos hay uno: el que resulta de $\varepsilon_1 = \varepsilon_2 = \cdots = \varepsilon_\nu = 0$). En cualquier caso, habrá una cierta cantidad de exponentes que se anulan: sea a este número (obviamente, $a \in \mathbb{N}$). La cantidad de exponentes que no se anulan será $m = 2^\nu - a$. De este modo, la expresión anterior quedará así:

$$0 = a + e^{\alpha_1} + \cdots + e^{\alpha_m}.$$

(b) Veamos que estos exponentes $\alpha_1, \alpha_2, \dots, \alpha_m$ son las raíces de cierto polinomio $\Psi(x) \in \mathbb{Z}[x]$ de grado m. De momento, sea el polinomio

$$f(x) = \prod_{\varepsilon_1=0,1}\prod_{\varepsilon_2=0,1}\cdots\prod_{\varepsilon_\nu=0,1}\big(x - (\varepsilon_1\gamma_1 + \varepsilon_2\gamma_2 + \cdots + \varepsilon_\nu\gamma_\nu)\big);$$

para cada j, el coeficiente de x^j es, a su vez, un polinomio en $\gamma_1, \gamma_2, \dots, \gamma_\nu$. Pero este polinomio es simétrico en $\gamma_1, \gamma_2, \dots, \gamma_\nu$; luego, por el corolario 6.24, es un número racional. Es decir, $f(x) \in \mathbb{Q}[x]$.

Las raíces de $f(x)$ son $\alpha_1, \alpha_2, \dots, \alpha_m$ y, además, 0, con multiplicidad a. Luego $f(x)/x^a \in \mathbb{Q}[x]$ y sus raíces son $\alpha_1, \alpha_2, \dots, \alpha_m$. Multiplicamos por un entero adecuado y obtenemos un polinomio

$$\Psi(x) = b(x - \alpha_1)(x - \alpha_2)\cdots(x - \alpha_m) = bx^m + \cdots \in \mathbb{Z}[x]$$

de grado m cuyas raíces son $\alpha_1, \alpha_2, \ldots, \alpha_m$, como queríamos. En particular, las reglas de Cardano muestran que $\alpha_1, \alpha_2, \ldots, \alpha_m$ y b cumplen las hipótesis del lema 6.28 (para no confundir al lector, aclaremos que no aplicaremos el lema a $\Psi(x)$, sino a otro polinomio $\Phi(x)$ que nos aparecerá más adelante).

(c) Tomemos un número primo p (grande: ya veremos cuánto) y definamos

$$\begin{aligned} R(x) &= b^{mp-1}x^{p-1}\Psi(x)^p \\ &= b^{mp+p-1}x^{p-1}(x-\alpha_1)^p(x-\alpha_2)^p\cdots(x-\alpha_m)^p; \end{aligned}$$

se trata de un polinomio con coeficientes en $\mathbb{Z}$ y grado $mp+p-1$. Definamos también

$$H(x) = \sum_{j\geq 0} R^{(j)}(x)$$

(pero que es una suma finita, claro). Según la identidad de Hermite[15], para $k = 1, 2, \ldots, m$ se tiene

$$e^{\alpha_k}H(0) - H(\alpha_k) = \int_0^{\alpha_k} e^{\alpha_k - t}R(t)\,dt.$$

Sumando en $k = 1, 2, \ldots, m$, y teniendo en cuenta que $0 = a + e^{\alpha_1} + e^{\alpha_2} + \cdots + e^{\alpha_m}$, resulta

$$-aH(0) - \sum_{k=1}^{m} H(\alpha_k) = \sum_{k=1}^{m} e^{\alpha_k}\int_0^{\alpha_k} e^{-t}R(t)\,dt.$$

Con esto, vamos a llegar a una contradicción, que consistirá en probar que

$$(p-1)! \leq \left|-aH(0) - \sum_{k=1}^{m} H(\alpha_k)\right| = \left|\sum_{k=1}^{m} e^{\alpha_k}\int_0^{\alpha_k} e^{-t}R(t)\,dt\right| \leq \lambda\mu^p$$

para ciertas constantes λ y μ que no dependen de p.

[15] Obsérvese que, ahora, $\alpha_k \in \mathbb{C}$, luego la integral entre 0 y α_k hay que entenderla como una integral sobre un camino: el segmento que une 0 y α_k, por ejemplo (aunque daría lo mismo cualquier otro camino, ya que el integrando es una función holomorfa, luego el valor de la integral es independiente del camino, únicamente depende de los puntos inicial y final). La demostración de la identidad de Hermite (lema 6.21) es —en este caso— exactamente la misma. Es más, tal como está enunciado —y aunque podría sobreentenderse— en el lema 6.21 no habíamos indicado $x \in \mathbb{R}$; así que, realmente, podemos pensar que $x \in \mathbb{C}$ y que todas las integrales que aparecen (incluidas las de la demostración) están hechas sobre caminos.

(d) Por una parte, examinemos $H(0) = R(0) + R'(0) + \cdots$. Usando el lema 6.8, es sencillo comprobar que[16]

$$\begin{aligned} R^{(j)}(0) &= 0, \quad \text{si } j = 0, 1, \ldots, p-2, \\ R^{(p-1)}(0) &= (p-1)!\, b^{mp-1}\Psi(0)^p, \\ R^{(j)}(0) &= p! A_j, \quad \text{con } A_j \in \mathbb{Z}, \quad \text{si } j \geq p. \end{aligned}$$

Luego $H(0) = (p-1)!\, b^{mp-1}\Psi(0)^p + p!\, A$ para cierta $A \in \mathbb{Z}$.

Por otra parte, para cada $k = 1, 2, \ldots, m$, de nuevo el lema 6.8 nos lleva a[17]

$$R^{(j)}(\alpha_k) = 0, \quad \text{si } j = 0, 1, \ldots, p-1.$$

Así que

$$H(\alpha_k) = \sum_{j \geq p} R^{(j)}(\alpha_k).$$

Ahora bien, $R(x) = b^{mp-1}x^{p-1}\Psi(x)^p$, y $x^{p-1}\Psi(x)^p$ es un polinomio con coeficientes en $\mathbb{Z}$, luego, por el lema 6.27, $R^{(p)}(x), R^{(p+1)}(x), \ldots$ son polinomios cuyos coeficientes no sólo están en $\mathbb{Z}$, sino que son todos múltiplos de $p!\, b^{mp-1}$. Entonces, podemos escribir

$$\sum_{j \geq p} R^{(j)}(x) = p!\, b^{mp-1}\Phi(x) \tag{6.18}$$

para cierto polinomio $\Phi(x) \in \mathbb{Z}[x]$. De esta manera, $H(\alpha_k) = p!\, b^{mp-1}\Phi(\alpha_k)$.

¿Cuál es el grado de Φ? Como R tiene grado $mp + p - 1$, de (6.18) se deduce que Φ tiene grado $mp - 1$. Así, según el lema 6.28 (con $d = mp - 1$),

$$\sum_{k=1}^{m} H(\alpha_k) = p!\, b^{mp-1} \sum_{k=1}^{m} \Phi(\alpha_k) = p!\, B, \quad \text{con } B \in \mathbb{Z}.$$

Recapitulando,

$$-aH(0) - \sum_{k=1}^{m} H(\alpha_k) = -a(p-1)!\, b^{mp-1}\Psi(0)^p - p!\,(aA + B),$$

donde $a, b, A, B, \Psi(0) \in \mathbb{Z}$. En realidad, A y B dependen de p, pero es importante que a, b y $\Psi(0)$ no. Así que, tomando p grande —es suficiente con elegirlo mayor que

[16] Esto es análogo a como hacíamos para llegar a (6.14); no repetimos los detalles.

[17] Basta proceder como en la deducción de (6.13).

$|a|$, $|b|$ y $|\Psi(0)|$, para que no pueda dividir a ninguno de los tres—, se deduce que $-aH(0) - \sum_{k=1}^{m} H(\alpha_k)$ es múltiplo de $(p-1)!$, pero no de $p!$ (por tanto, no es 0). Luego

$$(p-1)! \leq \left| -aH(0) - \sum_{k=1}^{m} H(\alpha_k) \right| = \left| \sum_{k=1}^{m} e^{\alpha_k} \int_0^{\alpha_k} e^{-t} R(t)\, dt \right|.$$

(e) Sean

$$M = \max\{|\alpha_1|, \ldots, |\alpha_m|\} \quad \text{y} \quad C = \max_{|z| \leq M} |\Psi(z)|.$$

Entonces, para cada $k = 1, 2, \ldots, m$ (y si denotamos $[0, \alpha_k]$ el segmento del plano complejo que une 0 con α_k), tenemos

$$\begin{aligned} \left| e^{\alpha_k} \int_0^{\alpha_k} e^{-t} R(t)\, dt \right| &\leq e^M M \sup_{t \in [0,\alpha_k]} |e^{-t} R(t)| \\ &\leq e^{2M} M \sup_{t \in [0,\alpha_k]} |b^{mp-1} e^{p-1} \Psi(t)^p| \leq e^{2M} M^p |b|^{mp-1} C^p. \end{aligned}$$

Sumando en $k = 1, 2, \ldots, m$ y aplicando la desigualdad triangular, resulta

$$\left| \sum_{k=1}^{m} e^{\alpha_k} \int_0^{\alpha_k} e^{-t} R(t)\, dt \right| \leq m \frac{e^{2M}}{|b|} \left(|b|^m MC \right)^p.$$

De aquí se deduce que

$$(p-1)! \leq \lambda \mu^p \quad \text{con} \quad \lambda = m \frac{e^{2M}}{|b|}, \quad \mu = |b|^m MC.$$

Pero esto no puede ser cierto si p es grande, ya que $\mu^p/(p-1)! \to 0$ cuando $p \to \infty$. De este modo, hemos llegado a una contradicción, con lo que queda demostrado el teorema. □

6.8. Extensiones y problemas abiertos

Como colofón a los resultados sobre irracionalidad y trascendencia estudiados en este tema, en esta sección vamos a ver —sin demostraciones— algunas extensiones de los resultados anteriores; analizaremos, así mismo, algunos problemas abiertos. Comencemos comentando que, en realidad, Lindemann no sólo probó la trascendencia de π (teorema 6.29), sino un resultado más general:

Teorema 6.30. Lindemann, 1882.
Si $\alpha \in \mathbb{C}$ es un número algebraico no nulo, entonces e^α es trascendente.

De aquí, no sólo se puede deducir la trascendencia de e (sin más que tomar $\alpha = 1$), sino también la de π. Para ello, basta utilizar la bien conocida relación $e^{i\pi} = -1$. Como -1 es, obviamente, un número algebraico, el teorema asegura que $i\pi$ no puede ser algebraico (luego es trascendente). Por el lema 6.25, π tiene que ser trascendente.

Además, si usamos log para denotar la determinación principal del logaritmo (que está definida en todo $\mathbb{C}$ menos el 0), del teorema también obtenemos lo siguiente:

Corolario 6.31.
Si $\alpha \in \mathbb{C} \setminus \{0, 1\}$ es un número algebraico, entonces $\log(\alpha)$ es trascendente.

Demostración. Supongamos que $\beta = \log(\alpha)$ fuese algebraico. Por el teorema anterior aplicado a β, el número $e^{\log(\alpha)} = \alpha$ sería trascendente, en contra de la hipótesis. □

En su trabajo, Lindemann enunció un teorema aún más general que el teorema 6.30. No incluía su demostración, pero indicaba que podía hacerse siguiendo el mismo método. Lo probó Weierstrass tres años más tarde, realizando —además— ciertas simplificaciones en las demostraciones de Hermite y Lindemann. Es el siguiente, que habitualmente se conoce con el nombre de teorema de Lindemann-Weierstrass:[18]

Teorema 6.32. Lindemann-Weierstrass, 1885.
Sean $\alpha_1, \alpha_2, \ldots, \alpha_m \in \mathbb{C}$ números algebraicos distintos y $\beta_1, \beta_2, \ldots, \beta_m \in \mathbb{C}$ números algebraicos no todos nulos. Entonces,

$$\beta_1 e^{\alpha_1} + \beta_2 e^{\alpha_2} + \cdots + \beta_m e^{\alpha_m} \neq 0.$$

No incluimos la demostración, que puede consultarse en [91, § 5 del tema 4] o [92, § 7 del tema 2]. Pero —eso sí— veamos cómo tanto el teorema 6.30 como su corolario pueden deducirse directamente de éste. También implica la trascendencia de algunas razones trigonométricas (y, de nuevo, como en el corolario 6.31, de logaritmos):

[18] Obsérvese la similitud entre la expresión $\beta_1 e^{\alpha_1} + \beta_2 e^{\alpha_2} + \cdots + \beta_m e^{\alpha_m}$ que se analiza en este teorema con $a + e^{\alpha_1} + \cdots + e^{\alpha_m}$, que se usaba —en la sección 6.7— en la demostración de la trascendencia de π. Es más, la relación $a + e^{\alpha_1} + \cdots + e^{\alpha_m} = 0$ (con $a \in \mathbb{N}$ y los α_j números algebraicos), que se obtenía tras suponer que π era algebraico, era la que finalmente conducía al absurdo que implicaba la trascendencia de π. Esto puede hacer intuir al lector —acertadamente— que la prueba del teorema que ahora presentamos sigue las mismas líneas generales que la de la trascendencia de π.

Corolario 6.33.
Si $\alpha \in \mathbb{C}$ es un número algebraico no nulo, entonces

$$e^{\alpha}, \quad \cos(\alpha), \quad \operatorname{sen}(\alpha), \quad \operatorname{tg}(\alpha)$$

son trascendentes.[19] *Si, además, $\alpha \neq 1$, entonces $\log(\alpha)$ es trascendente.*

Demostración. Por comodidad, numeremos (a), (b), ..., (e) las cinco afirmaciones que hacemos. Así,

(a) Para cada $n \in \mathbb{N}$, los números $0, \alpha, 2\alpha, \ldots, n\alpha$ son algebraicos y distintos. Luego, según el teorema, para cualesquiera $\beta_0, \beta_1, \beta_2, \ldots, \beta_n \in \mathbb{Q}$, no todos nulos, se tiene

$$\beta_0 e^0 + \beta_1 e^{\alpha} + \beta_2 e^{2\alpha} + \cdots + \beta_n e^{n\alpha} \neq 0.$$

Así pues, e^{α} no es raíz de ningún polinomio en $\mathbb{Q}[x]$; en otras palabras, es trascendente.

(b) La relación $\cos(\alpha) = (e^{i\alpha} + e^{-i\alpha})/2$ podemos escribirla como

$$(2\cos(\alpha))e^0 - e^{i\alpha} - e^{-i\alpha} = 0.$$

En estas condiciones, el teorema asegura que $\cos(\alpha)$ no puede ser algebraico.

(c) De modo similar, escribimos $\operatorname{sen}(\alpha) = (e^{i\alpha} - e^{-i\alpha})/(2i)$ como

$$(2i\operatorname{sen}(\alpha))e^0 - e^{i\alpha} + e^{-i\alpha} = 0,$$

y el teorema prueba que $2i\operatorname{sen}(\alpha)$ no puede ser algebraico; por tanto, $\operatorname{sen}(\alpha)$ no es algebraico.

(d) A partir de $\operatorname{tg}(\alpha) = (e^{i\alpha} - e^{-i\alpha})/(ie^{i\alpha} + ie^{-i\alpha})$ obtenemos

$$(i\operatorname{tg}(\alpha) - 1)e^{i\alpha} + (i\operatorname{tg}(\alpha) + 1)e^{-i\alpha} = 0.$$

Según el teorema, alguno de los dos coeficientes no es algebraico. De aquí se deduce que $\operatorname{tg}(\alpha)$ no es algebraico.

(e) Usemos ahora que $1 \cdot e^{\log(\alpha)} - \alpha e^0 = 0$. Si $\alpha \neq 1$, entonces $\log(\alpha) \neq 0$. Aplicando el teorema, $\log(\alpha)$ no puede ser algebraico. □

6.8.1. El séptimo problema de Hilbert

En su influyente conferencia del Congreso Internacional de Matemáticos de 1900 en París, Hilbert planteó una lista de 23 problemas abiertos, lo que supuso un gran

[19] Merece la pena recordar que, aunque restringiéndonos al caso $\alpha \in \mathbb{Q} \setminus \{0\}$, en el apartado 6.3.1 analizábamos la irracionalidad de diversos números de aspecto similar.



Rodion Osievich Kuzmin (1891–1949)



Theodor Schneider (1911–1988)

estímulo a la investigación matemática.[20] En el séptimo problema, Hilbert conjeturó que, si α es un número algebraico distinto de 0 y 1, y β es un número algebraico no racional, entonces el número α^β es trascendente. En particular, mencionó los números $2^{\sqrt{2}}$ y $e^\pi = i^{-2i}$ (pues $i^{-2i} = (e^{i\pi/2})^{-2i} = e^\pi$; alternativamente, también se puede hacer $(-1)^{-i} = (e^{i\pi})^{-i} = e^\pi$).

El primer avance se produjo en 1929, cuando el matemático ruso Aleksandr Gelfond desarrolló un nuevo método analítico —basado en cierto tipo de interpolación— que le permitió dar una prueba parcial del problema de Hilbert. Concretamente, demostró que la conjetura era cierta cuando el exponente β pertenece a lo que se denomina un cuerpo cuadrático imaginario. En particular, esto le permitió probar que $e^\pi = i^{-2i}$ es trascendente.

En 1930, el también ruso Rodion Kuzmin extendió el método de Gelfond, adaptándolo al caso en que β perteneciera a un cuerpo cuadrático real. Como consecuencia, se obtenía la trascendencia de $2^{\sqrt{2}}$.

En 1934, Gelfond desarrolló otro método analítico para obtener resultados sobre trascendencia; con él, fue capaz de demostrar la conjetura completa. También lo logró el alemán Theodor Schneider ese mismo año, de manera independiente y utilizando un método distinto. Actualmente, es común aludir a este resultado por el nombre de teorema de Gelfond-Schneider:

Teorema 6.34. Gelfond-Schneider, 1934.
Sean α y β dos números algebraicos, con $0 \neq \alpha \neq 1$ y $\beta \notin \mathbb{Q}$. Entonces, α^β es un número trascendente.

[20] En este tema, ya nos hemos topado con uno de estos problemas en la nota 6.15, en la página 356. A lo largo del libro aparecen otros tres, en las páginas 60, 193 y 468.

6.8.2. Otros resultados sobre trascendencia

La resolución completa del séptimo problema de Hilbert no detuvo la investigación sobre trascendencia de números; los métodos usados pudieron ser generalizados, lo que llevó a obtener nuevos resultados.

Además, otra línea de investigación, heredera de los métodos de Hermite y Lindemann había continuado desarrollándose, y recibió un renovado impulso procedente de las ideas de los métodos de Gelfond y Schneider. Así, el estudio de la trascendencia emparentó con temas como las ecuaciones diferenciales y las funciones enteras. Sin entrar en detalles, los principales resultados en este sentido fueron debidos a Carl Siegel, quien, entre 1929 y 1949, introdujo el estudio de las propiedades aritméticas de lo que denominó E-funciones (véase [93]). Nuevos ímpetus fueron aportados a partir 1954, cuando Andrei Shidlovskii logró extender la teoría de Siegel (véase [91, 92]). Tampoco podemos dejar de citar a Alan Baker (véase [9]) quien, en 1966, obtuvo una generalización del teorema de Gelfond-Schneider que proporcionó, por ejemplo, la trascendencia de los números de la forma

$$e^{\beta_0}\alpha_1^{\beta_1}\cdots\alpha_n^{\beta_n},$$

donde los α_k y β_k son algebraicos no nulos. Por sus trabajos sobre ecuaciones diofánticas[21] y números trascendentes, Baker recibió la Medalla Fields.[22]

También merece la pena destacar las aportaciones de Gregory Chudnovsky, realizadas a finales de la década de 1970 mediante el uso de curvas elípticas. En general, y sin dar detalles, una curva elíptica viene dada por las soluciones de cierto tipo de ecuaciones algebraicas (como $y^2 = 4x^3 - 4$).[23] Interpretadas de manera adecuada sobre los complejos, de estas «curvas» surge una función compleja, denominada función $\wp$ de Weierstrass (en el ejemplo anterior, $\wp$ verifica la ecuación diferencial $\wp'(z)^2 = 4\wp(z)^3 - 4$). Una de las principales propiedades de $\wp$ es que tiene un doble periodo, que genera un retículo W en $\mathbb{C}$, así como un polo de orden dos en cada punto de W; además, $\wp$ puede expresarse como

$$\wp(z) = \frac{1}{z^2} + \sum_{w\in W\setminus\{0\}} \left(\frac{1}{(z-w)^2} - \frac{1}{w^2}\right).$$

[21] Ya hemos comentado algo sobre esto en la sección 4.3.

[22] Véase la nota al pie número 7 del tema 2 (página 113).

[23] Las curvas elípticas se pueden definir sobre cualquier cuerpo y deben su nombre a que, sobre $\mathbb{R}$, son una curva en el plano (como $y^2 = 4x^3 - 4$, con $x, y \in \mathbb{R}$). Las curvas elípticas han resultado ser útiles en muchos otros campos, en particular cuando se definen sobre cuerpos discretos (como los $\mathbb{Z}_p$ con p primo del tema 2 de este libro). Se usan, por ejemplo, como herramientas destinadas a analizar la primalidad y factorización de enteros (sirven para construir algoritmos computacionalmente muy eficientes), así como en criptografía. También desempeñan un papel esencial en la demostración de Andrew Wiles del último teorema de Fermat.



Roger Apéry
(1916–1994)



Marin Mersenne
(1588–1648)



Lorenzo Mascheroni
(1750–1800)

Chudnovsky demostró que, para una clase bastante amplia de curvas elípticas, cada elemento no nulo del retículo es trascendente (nótese que esto es análogo a lo que ocurre con la función e^z: cada periodo no nulo —es decir, cada $2k\pi i$ con $k \in \mathbb{Z}\backslash\{0\}$— es trascendente). A partir de este hecho se puede demostrar la trascendencia de la función Γ de Euler

$$\Gamma(z) = \int_0^\infty t^{z-1} e^{-t}\, dt, \quad z \in \mathbb{C},\ \mathrm{Re}(z) > 0$$

(muy conocida por verificar $\Gamma(n+1) = n!$, $n \in \mathbb{N}$) para ciertos valores de la variable, como $\Gamma(1/4)$, $\Gamma(1/3)$ y $\Gamma(1/6)$. Para profundizar en esto, consúltese [23] o [24]. Contrastando con el intrincado camino recorrido hasta llegar a los resultados anteriores, determinar la naturaleza aritmética de $\Gamma(1/2)$ es mucho más directo: de la conocida fórmula $\Gamma(1/2) = \sqrt{\pi}$ se sigue que $\Gamma(1/2)$ es trascendente, por serlo π.

6.8.3. Tampoco es fácil probar la irracionalidad

Desde que comenzamos la sección 6.8, sólo hemos hablado de trascendencia. Pero, aunque parezca más sencilla, también la irracionalidad sigue siendo un tema de investigación actual. Se han producido avances provenientes de generalizaciones de las ideas de Roger Apéry, quien en 1978 anunció que había probado la irracionalidad de $\zeta(3) = \sum_{n=1}^{\infty} 1/n^3$ (su prueba puede verse en [4]); poco más tarde, Frits Beukers [21] presentó una demostración simplificada (véase también [28]). Este tipo de técnicas, que relacionaban las irracionalidad de algunas series con polinomios ortogonales y aproximantes de Padé, sirvieron de inspiración a otros matemáticos. Citemos, por ejemplo, a Peter Borwein, que en 1991 demostró el carácter irracional



Takakazu Seki
(1642–1708)



Jacob Bernoulli
(1654–1705)



Johann Faulhaber
(1580–1635)

del número definido mediante la serie

$$\sum_{k=1}^{\infty} \frac{1}{q^k - 1} \quad \text{con } q \in \mathbb{N},\ q > 1$$

(y algunas otras similares). En particular, eso extendía un resultado de Pál Erdős, quien en 1948 lo había probado para el caso $q = 2$ (este número, que se suele denominar constante de Erdős-Borwein, es la suma de los inversos de los números de Mersenne). Por supuesto, otros investigadores han continuado por esta senda.

6.8.4. Problemas abiertos

Concluyamos el tema diciendo que —claramente— no todo está hecho en este campo. Por ejemplo, continúa sin conocerse la naturaleza aritmética de los números 2e, 2^π, π^e o $e+\pi$ (¡ni siquiera si son o no racionales!). Ni de la constante de Euler $\gamma = \lim_{n\to\infty}(\sum_{k=1}^{n} 1/k - \log(n))$, un número de gran importancia en teoría de números,[24] y que también se conoce por el nombre de constante de Euler-Mascheroni (un interesante libro esencialmente dedicado a γ es [51]).

Tampoco se sabe si los números $\zeta(2k+1)$, para $k \geq 2$ y entero, son racionales (recordemos que, para $s > 1$, $\zeta(s) = \sum_{n=1}^{\infty} 1/n^s$); como ya hemos comentado, sí es conocido que $\zeta(3)$ es irracional, pero no si es algebraico. Sin embargo, la trascendencia de $\zeta(2k)$ no ofrece dudas, pues $\zeta(2k)$ es siempre un múltiplo racional

[24] A nosotros nos aparecerá a menudo en el tema 8. En particular, en la proposición 8.10 vemos que ese límite realmente existe.

de π^{2k}. En efecto: en 1740, Euler encontró la expresión[25]

$$\zeta(2k) = \sum_{n=1}^{\infty} \frac{1}{n^{2k}} = \frac{(-1)^{k+1}(2\pi)^{2k}}{2(2k)!} B_{2k},$$

donde los B_j son los denominados números de Bernoulli, unos números racionales fácilmente calculables, y que ya habían sido estudiados tanto por Jacob Bernoulli como por el matemático japonés Takakazu Seki (en ambos casos, las publicaciones de sus estudios fueron póstumas: 1713 en el caso de Bernoulli y un año antes en el de Seki). La forma más habitual de definirlos es mediante la función generatriz

$$\frac{x}{(e^x - 1)},$$

es decir, usando el desarrollo de Taylor

$$\frac{x}{e^x - 1} = \sum_{j=0}^{\infty} B_j \frac{x^j}{j!}. \tag{6.19}$$

A partir de esta serie, multiplicando por $e^x - 1 = \sum_{j=1}^{\infty} x^j/j!$ e igualando coeficientes, es sencillo obtener relaciones de recurrencia para los B_j. Conviene aclarar que ni Bernoulli ni Seki usaron los B_j para calcular $\zeta(2k)$, sino que los empleaban para evaluar las sumas finitas $1^p + 2^p + 3^p + \cdots + n^p$ (con el exponente p un entero positivo), para las que ya Johann Faulhaber había encontrado fórmulas compactas en numerosos casos particulares de p.[26]

A la vista del valor de $\zeta(2k)$, de manera inocente se puede especular con la posibilidad de que los números $\zeta(2k+1)$ sean múltiplos racionales de π^{2k+1}, pero no hay ninguna base teórica ni numérica de que así ocurra; en realidad, y pese a todos los esfuerzos realizados, muy poco es lo que se conoce sobre la naturaleza de estos números.

6.9. Polinomios de Bernoulli y $\zeta(2k)$

A principios del siglo XVIII se sabía que $\sum_{n=1}^{\infty} 1/n = \infty$ (comentamos algo más sobre esta serie en el tema 9, al principio de la sección 9.1) y que $\sum_{n=1}^{\infty} 1/n^2 < \infty$, pero no se conocía una expresión compacta para el valor de esta última suma.

[25] Actualmente hay muchos procedimientos distintos de deducir esta fórmula. Uno de los más sencillos es el que figura en la siguiente sección de este tema.

[26] En los ejercicios 6.37 y 6.38 podemos ver fórmulas de sumación para este tipo de expresiones.

Encontrarla era un problema por el que habían estado interesados bastantes de los más conocidos matemáticos de la época, entre ellos Pietro Mengoli (fue él quien, en 1644, lo planteó por primera vez), los hermanos Jacob y Johann Bernoulli, Gottfried Leibniz y John Wallis; a menudo se alude a él como el *problema de Basilea* (la ciudad donde vivía la familia Bernoulli y también Euler). Sí se sabía que

$$\sum_{n=1}^{\infty} \frac{1}{n(n+1)} = 1,$$

y de hecho se puede calcular fácilmente con un procedimiento ideado por Mengoli que, esencialmente, es así:

$$\begin{aligned}
\sum_{n=1}^{\infty} \frac{1}{n(n+1)} &= \sum_{n=1}^{\infty} \left(\frac{1}{n} - \frac{1}{n+1}\right) = \lim_{N\to\infty} \sum_{n=1}^{N} \left(\frac{1}{n} - \frac{1}{n+1}\right) \\
&= \lim_{N\to\infty} \left(\left(1 - \frac{1}{2}\right) + \left(\frac{1}{2} - \frac{1}{3}\right) + \cdots + \left(\frac{1}{N} - \frac{1}{N+1}\right)\right) \\
&= \lim_{N\to\infty} \left(1 + \left(-\frac{1}{2} + \frac{1}{2}\right) + \left(-\frac{1}{3} + \frac{1}{3}\right) + \cdots + \left(-\frac{1}{N} + \frac{1}{N}\right) - \frac{1}{N+1}\right) \\
&= \lim_{N\to\infty} \left(1 - \frac{1}{N+1}\right) = 1.
\end{aligned}$$

Aquí tenemos lo que se denomina una *serie telescópica*, en la que, tras descomponer cada sumando de manera adecuada, parte de cada sumando se cancela con parte del siguiente sumando. Pero esta técnica no funcionaba con $\sum_{n=1}^{\infty} 1/n^2$. Mediante procedimientos totalmente distintos —que aquí no veremos, pero que se pueden encontrar en [38, capítulo 9]— fue Euler en 1734 el primero que llegó a la conclusión de que $\sum_{n=1}^{\infty} 1/n^2 = \pi^2/6$; de hecho, lo consiguió probar de varias formas distintas. Es más, en 1740 Euler logró sumar $\sum_{n=1}^{\infty} 1/n^{2k}$ con k cualquier entero positivo.

Las sumas que encontró Euler son los valores para $s = 2k$ de la función que actualmente se denota

$$\zeta(s) = \sum_{n=1}^{\infty} \frac{1}{n^s}, \qquad s > 1,$$

y se denomina función «zeta» de Riemann. Aunque para lo que nosotros pretendemos ahora no tiene relevancia, esta función —extendida para valores complejos de s— juega un papel crucial al estudiar la distribución de los números primos (se ve en el tema 10 de este libro). Hay bastantes maneras distintas de sumar las series $\zeta(2k)$, con distintas técnicas y distinto grado de dificultad (a veces, no por la



Pietro Mengoli
(1626–1686)



Brook Taylor
(1685–1731)

dificultad en sí, sino por los requisitos matemáticos previos para poder hacer la correspondiente demostración con el rigor requerido). El objetivo de esta sección va a ser mostrar uno de estos procedimientos para calcular $\zeta(2k)$, y hemos elegido uno de los que consideramos más sencillos: utiliza una función generatriz (es decir, un desarrollo de Taylor), integración por partes, un fórmula trigonométrica elemental, y la cancelación de términos en una serie telescópica (esto último, que no es habitual en la mayoría de las demostraciones conocidas, la hace particularmente sencilla en cuanto a los prerrequisitos para poder abordarla). Esencialmente, es la misma que hay en [25] (en ese artículo se pueden ver referencias a otras demostraciones).

Como ya hemos comentado, al calcular $\zeta(2k)$ aparecen los números de Bernoulli (los hemos definido en (6.19)). En el proceso que vamos a seguir ahora usaremos no sólo los números de Bernoulli, sino los denominados *polinomios de Bernoulli* $\{B_k(t)\}_{k=0}^{\infty}$. Hay varias formas de definirlos; la más usual es a partir de la función generatriz

$$\frac{x}{e^x-1}\,e^{xt}=\sum_{k=0}^{\infty}B_k(t)\frac{x^k}{k!}; \tag{6.20}$$

es decir, se está haciendo el desarrollo de Taylor de $xe^{xt}/(e^x-1)$ (como función de x) en torno a $x=0$.[27] Observe el lector que, aunque hemos dicho que los $B_k(t)$ son polinomios, realmente no lo hemos probado aún, pero será evidente en seguida.

A partir de la definición (6.20) y la unicidad de los desarrollos en serie de potencias, es fácil demostrar muchas propiedades de los polinomios de Bernoulli; enunciamos las que vamos a usar y las proponemos como ejercicio al lector. En primer

[27] Aunque conocer el radio de convergencia de esta serie no tiene demasiada importancia, el primer valor (complejo) de x que anula el denominador de la función generatriz es $x=2\pi i$, así que la serie está bien definida para $|x|<2\pi$, independientemente del valor de t.

lugar, multiplicando (6.20) por $(e^x-1)/x$ se obtiene que

$$t^n = \frac{1}{n+1}\sum_{k=0}^{n}\binom{n+1}{k}B_k(t), \qquad n=0,1,2,\dots \tag{6.21}$$

(ejercicio 6.33). Esto prueba que $B_0(t)=1$, un polinomio de grado 0, y, además, permite evaluar sucesivamente todos los $B_k(t)$; en particular, prueba que $B_k(t)$ es un polinomio de grado k. Los primeros polinomios de Bernoulli son

$$B_0(t)=1, \qquad B_1(t)=t-\frac{1}{2}, \qquad B_2(t)=t^2-t+\frac{1}{6}.$$

Así mismo, derivando (6.20) respecto a t e igualando coeficientes se obtiene que

$$B_k'(t) = kB_{k-1}(t), \qquad k\geq 1 \tag{6.22}$$

(ejercicio 6.34). Otra propiedad que nos interesa, y que también se deduce de (6.20), es la siguiente relación de simetría cuya comprobación se propone en el ejercicio 6.35:

$$B_k(1-t) = (-1)^k B_k(t), \qquad k\geq 0. \tag{6.23}$$

Una vez definidos los polinomios de Bernoulli, los números de Bernoulli se definen como $B_k = B_k(0)$ (o, lo que es equivalente, tomando $t=0$ en la función generatriz (6.20)). Como $x/(e^x-1)+x/2$ es una función par (ver que al cambiar x por $-x$ vuelve a salir lo mismo es una simple comprobación), su desarrollo en serie de potencias tendrá sólo potencias pares. En consecuencia, $B_1=-1/2$ y $B_{2k+1}=0$ para cualquier $k\in\mathbb{N}$. Usando (6.23) se sigue que

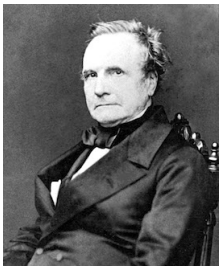
$$B_{2k+1}(1) = B_{2k+1}(0) = 0, \qquad k\geq 1. \tag{6.24}$$

También es cierto que $B_{2k}(1)=B_{2k}(0)$, pero estos números ya no son nulos. En cualquier caso, sin más que aplicar (6.22) se obtiene

$$\int_0^1 B_k(t)\,dt = \frac{1}{k+1}(B_{k+1}(1)-B_{k+1}(0)) = 0, \quad k\geq 1. \tag{6.25}$$

Los polinomios de Bernoulli cumplen muchísimas otras propiedades de gran interés,[28] pero ya tenemos todas las que vamos a utilizar en nuestro objetivo, así

[28] Hasta tal punto es así que el considerado como primer algoritmo publicado, y específicamente diseñado para ser implementado en un ordenador, estaba dedicado a calcular números de Bernoulli. Data de 1842, su autora fue Ada Lovelace, y su destino era poder ser usado en la «máquina analítica» de Charles Babbage, que nunca llegó a construirse, pero que representó un paso importante en la historia de la informática.



Charles Babbage
(1791–1871)



Ada Lovelace
(1815–1852)

que no profundizaremos por este camino.[29] A este respecto, sí que resulta relevante comentar que también se pueden definir los polinomios de Bernoulli como la familia de polinomios $\{B_k(t)\}_{k=0}^{\infty}$ que cumple tres de las propiedades que hemos visto:

$$B_0(0) = 1, \quad B_k'(t) = kB_{k-1}(t), \quad \int_0^1 B_k(t)\,dt = 0, \quad k \geq 1.$$

Posiblemente, esta definición alternativa de los polinomios de Bernoulli es «más elemental», pero menos usual.

Vamos ahora a calcular unas integrales auxiliares:

Lema 6.35.
Para $k \geq 0$ y $m \geq 1$, sea

$$I_{k,m} = \int_0^1 B_{2k}(t)\cos(m\pi t)\,dt.$$

Se verifica

$$I_{k,m} = \begin{cases} 0, & m = 1,3,5,\ldots, \\ \dfrac{(-1)^{k-1}(2k)!}{m^{2k}\pi^{2k}}, & m = 2,4,6,\ldots. \end{cases} \tag{6.26}$$

Demostración. En primer lugar, es inmediato que

$$I_{0,m} = \int_0^1 \cos(m\pi t)\,dt = 0, \quad m = 1,2,3,\ldots.$$

[29] No obstante, los ejercicios finales del tema proponen al lector que demuestre algunas otras propiedades, en particular las que conducen a fórmulas de sumación para expresiones del tipo $1^p + 2^p + 3^p + \cdots + n^p$ (con p un entero positivo).

Para $k \geq 1$, integremos por partes dos veces, aplicando (6.22) en las dos ocasiones. Como $\operatorname{sen}(m\pi t)$ se anula en $t = 0$ y 1, obtenemos

$$\begin{aligned} I_{k,m} &= \int_0^1 B_{2k}(t)\cos(m\pi t)\,dt \\ &= \frac{1}{m\pi}\Big(B_{2k}(t)\operatorname{sen}(m\pi t)\Big)\Big|_{t=0}^1 - \frac{2k}{m\pi}\int_0^1 B_{2k-1}(t)\operatorname{sen}(m\pi t)\,dt \\ &= 0 - \frac{2k}{m\pi}\int_0^1 B_{2k-1}(t)\operatorname{sen}(m\pi t)\,dt \\ &= \frac{2k}{m^2\pi^2}\Big(B_{2k-1}(t)\cos(m\pi t)\Big)\Big|_{t=0}^1 - \frac{2k(2k-1)}{m^2\pi^2}\int_0^1 B_{2k-2}(t)\cos(m\pi t)\,dt. \end{aligned}$$

Cuando $k = 1$, los dos polinomios que aparecen en la última línea son $B_1(t) = t - 1/2$ y $B_0(t) = 1$, y esto proporciona el caso especial

$$\begin{aligned} I_{1,m} &= \int_0^1 B_2(t)\cos(m\pi t)\,dt \\ &= \frac{2}{m^2\pi^2}\Big((t-1/2)\cos(m\pi t)\Big)\Big|_{t=0}^1 - \frac{2}{m^2\pi^2}\int_0^1 \cos(m\pi t)\,dt \\ &= \begin{cases} 0, & m = 1,3,5,\ldots, \\ \dfrac{2}{m^2\pi^2}, & m = 2,4,6,\ldots. \end{cases} \end{aligned}$$

En cambio, $B_{2k-1}(t)$ se anula en $t = 0$ y 1 cuando $k \geq 2$ (por (6.24)), así que obtenemos la relación de recurrencia

$$I_{k,m} = -\frac{2k(2k-1)}{m^2\pi^2}\,I_{k-1,m}, \qquad k \geq 2,$$

que, por inducción sobre k, da lugar a (6.26). □

Realmente, en lugar del $I_{k,m}$ que hemos calculado en el lema, usaremos

$$I'_{k,m} = \int_0^1 \big(B_{2k}(t) - B_{2k}(0)\big)\cos(m\pi t)\,dt, \qquad k \geq 0, \quad m \geq 1$$

(la razón es que nos interesa tener un polinomio $B_{2k}(t) - B_{2k}(0)$ que se anula en $t = 0$). Al ser $\int_0^1 \cos(m\pi t)\,dt = 0$ para cada $m \in \mathbb{N}$, es claro que $I'_{k,m} = I_{k,m}$. Así que de nuevo

$$I'_{k,m} = \begin{cases} 0, & m = 1,3,5,\ldots, \\ \dfrac{(-1)^{k-1}(2k)!}{m^{2k}\pi^{2k}}, & m = 2,4,6,\ldots. \end{cases}$$

En nuestro camino para sumar $\sum_{n=1}^{\infty} 1/n^{2k}$, además de los polinomios de Bernoulli y el lema usaremos la identidad trigonométrica

$$\cos(mx) = \frac{\operatorname{sen}(\frac{2m+1}{2}x) - \operatorname{sen}(\frac{2m-1}{2}x)}{2\operatorname{sen}(\frac{x}{2})}, \tag{6.27}$$

que es elemental. Servirá para conseguir transformar la suma de $\sum_{n=1}^{\infty} 1/n^{2k}$ en la de una serie telescópica.

Con todo esto, ya podemos abordar el resultado principal de esta sección:

Teorema 6.36. Euler, 1740.
Para cada entero positivo k se tiene

$$\zeta(2k) = \sum_{n=1}^{\infty} \frac{1}{n^{2k}} = \frac{(-1)^{k+1}(2\pi)^{2k}}{2(2k)!} B_{2k}, \tag{6.28}$$

donde los B_{2k} son los números de Bernoulli de índice par.

Demostración. Denotemos $b_{2k}(t) = B_{2k}(t) - B_{2k}(0)$.

Usando que $I'_{k,2m+1} = 0$ para cada m, aplicando la identidad trigonométrica (6.27), y cancelando términos de la serie telescópica que aquí nos aparece, tenemos

$$\begin{aligned}
\frac{(-1)^{k-1}(2k)!}{\pi^{2k}} \sum_{m=1}^{\infty} \frac{1}{(2m)^{2k}} &= \sum_{m=1}^{\infty} I'_{k,2m} = \sum_{m=1}^{\infty} I'_{k,2m} + \sum_{m=0}^{\infty} I'_{k,2m+1} \\
&= \sum_{m=1}^{\infty} I'_{k,m} = \sum_{m=1}^{\infty} \int_0^1 b_{2k}(t)\cos(m\pi t)\,dt \\
&= \lim_{N\to\infty} \sum_{m=1}^{N} \left(\int_0^1 b_{2k}(t) \frac{\operatorname{sen}(\frac{2m+1}{2}\pi t)}{2\operatorname{sen}(\frac{\pi t}{2})}\,dt - \int_0^1 b_{2k}(t) \frac{\operatorname{sen}(\frac{2m-1}{2}\pi t)}{2\operatorname{sen}(\frac{\pi t}{2})}\,dt \right) \\
&= \lim_{N\to\infty} \int_0^1 b_{2k}(t) \frac{\operatorname{sen}(\frac{2N+1}{2}\pi t)}{2\operatorname{sen}(\frac{\pi t}{2})}\,dt - \int_0^1 b_{2k}(t) \frac{\operatorname{sen}(\frac{\pi t}{2})}{2\operatorname{sen}(\frac{\pi t}{2})}\,dt.
\end{aligned}$$

Analicemos la última línea.

Comencemos comprobando que

$$\lim_{N\to\infty} \int_0^1 b_{2k}(t) \frac{\operatorname{sen}(\frac{2N+1}{2}\pi t)}{2\operatorname{sen}(\frac{\pi t}{2})}\,dt = 0.$$

La función

$$f(t) = \frac{b_{2k}(t)}{2\,\mathrm{sen}(\frac{\pi t}{2})}, \qquad t \in (0,1],$$

extendida con continuidad a $t = 0$ (nótese que $b_{2k}(0) = 0$), es derivable en $[0,1]$ con derivada continua. Además, y por simplificar la notación, tomemos $R = (2N+1)\pi/2$. Integrando por partes,

$$\int_0^1 f(t)\,\mathrm{sen}(Rt)\,dt = -\frac{\cos(R)}{R}f(1) + \frac{1}{R}f(0) + \int_0^1 f'(t)\frac{\cos(Rt)}{R}\,dt.$$

Como $f'(t)$ es acotada, cada término tiende a 0 cuando $R \to \infty$, así que el límite cuando $N \to \infty$ tiende 0.

Para concluir, usando (6.25) y recordando que $B_{2k}(0) = B_{2k}$, el valor de la integral final es

$$\int_0^1 b_{2k}(t)\frac{\mathrm{sen}(\frac{\pi t}{2})}{2\,\mathrm{sen}(\frac{\pi t}{2})}\,dt = \frac{1}{2}\int_0^1 (B_{2k}(t) - B_{2k}(0))\,dt = -\frac{1}{2}\int_0^1 B_{2k}(0)\,dt = -\frac{B_{2k}}{2}.$$

En consecuencia,

$$\frac{(-1)^{k-1}(2k)!}{2^{2k}\pi^{2k}}\sum_{m=1}^{\infty}\frac{1}{m^{2k}} = 0 + \frac{B_{2k}}{2}$$

y ya hemos encontrado el valor de $\zeta(2k)$. □

Los lectores familiarizados con la teoría de series de Fourier habrán podido reconocer que las integrales $I_{k,2m}$ que evaluamos en el lema 6.35 son los coeficientes de Fourier de los polinomios de Bernoulli de índice par $B_{2k}(t)$. El desarrollo de Fourier de $B_{2k}(t)$ —lo encontró Adolf Hurwitz en 1890— es

$$B_{2k}(t) = \frac{2(-1)^{k-1}(2k)!}{(2\pi)^{2k}}\sum_{n=1}^{\infty}\frac{\cos(2\pi nt)}{n^{2k}}, \quad t \in [0,1), \qquad k \geq 1,$$

y sustituyendo $t = 0$ obtenemos (6.28). Esto proporciona una demostración de (6.28) muy rápida pero, claro, requiere tener un conocimiento previo sobre series de Fourier y —en particular— de su convergencia puntual. En cambio, la demostración presentada en el teorema 6.36 utiliza una serie telescópica cuya convergencia se estudia sin dificultad y sin requisitos técnicos avanzados.

Nota 6.37. Resulta tentador pensar si el procedimiento seguido en esta sección, o uno similar, se puede utilizar para evaluar $\zeta(2k+1)$. La respuesta es que sí, pero a medias. Se toma

$$J_{k,m} = \int_0^1 B_{2k+1}(t)\,\mathrm{sen}(m\pi t)\,dt$$

y de forma similar a como se hace en el lema 6.35 se obtiene que

$$J_{k,m} = \begin{cases} 0, & m = 1, 3, 5, \ldots, \\ \dfrac{(-1)^{k-1}(2k+1)!}{m^{2k+1}\pi^{2k+1}}, & m = 2, 4, 6, \ldots. \end{cases}$$

Además, se usa la identidad trigonométrica

$$\operatorname{sen}(mx) = -\frac{\cos(\frac{2m+1}{2}x) - \cos(\frac{2m-1}{2}x)}{2\operatorname{sen}(\frac{x}{2})},$$

que, procediendo como en la demostración del teorema 6.36, da lugar a una serie telescópica. De nuevo el límite en N es nulo, pero la integral final no se sabe evaluar en términos elementales. Lo que queda es

$$\zeta(2k+1) = \frac{(-1)^{k-1}2^{2k}\pi^{2k+1}}{(2k+1)!}\int_0^1 B_{2k+1}(t)\frac{\cos(\frac{\pi t}{2})}{\operatorname{sen}(\frac{\pi t}{2})}\,dt, \quad k = 1, 2, 3, \ldots.$$

Así pues, la serie $\sum_{n=1}^{\infty} 1/n^{2k+1}$ se transforma en una integral, pero no se llega a una expresión compacta. ◁

6.10. Ejercicios

6.1. Demuestra el corolario 6.3.

6.2. Demuestra el corolario 6.4.

6.3. Sea $\alpha = (2 + 5^{1/3})^{1/2}$. ¿Es α racional?

6.4. Veamos que no todos los números con «aspecto irracional» son realmente irracionales. Para ello, prueba la igualdad[30]

$$\sqrt{A \pm \sqrt{B}} = \sqrt{\frac{A + \sqrt{A^2 - B}}{2}} \pm \sqrt{\frac{A - \sqrt{A^2 - B}}{2}}$$

analizando, previamente, qué condiciones sobre A y B garantizan que ninguna de las subexpresiones que aparecen da lugar a números complejos. Aplícalo al caso $A = 3$, $B = 1$ y signo «+».

[30] Esta fórmula ya era conocida —en el siglo XII— por el matemático hindú Bhaskara II (el apelativo «II» se usa exclusivamente para distinguirlo de otro matemático de igual nombre, pero que vivió cinco siglos antes).

6.5. Analiza la racionalidad o irracionalidad de $\sqrt{7+4\sqrt{3}}+\sqrt{7-4\sqrt{3}}$.

6.6. Estudia si el número real $\sqrt[3]{7+5\sqrt{2}}+\sqrt[3]{-7+5\sqrt{2}}$ es racional o irracional.

6.7. Comprueba que[31]

$$\sqrt[3]{10+\sqrt{108}}-\sqrt[3]{-10+\sqrt{108}}=2.$$

¿Puedes encontrar una expresión similar para el número que aparece en el ejercicio anterior?
Sugerencia. Lee la nota a pie de página.

6.8. Prueba que los polinomios de Chebyshev $\{T_n(x)\}_{n=0}^{\infty}$ satisfacen la relación de ortogonalidad

$$\int_{-1}^{1} T_n(x)T_m(x)(1-x^2)^{-1/2}\,dx=\begin{cases}0, & \text{si } n\neq m,\\ \pi, & \text{si } n=m=0,\\ \pi/2, & \text{si } n=m>0.\end{cases}$$

6.9. Definamos los polinomios $\{f_n(x)\}_{n=1}^{\infty}$ que aparecen en la demostración del teorema 6.5 directamente mediante

$$\begin{cases}f_1(x)=x, \quad f_2(x)=x^2-2,\\ f_{n+1}(x)=xf_n(x)-f_{n-1}(x), \quad n\geq 2.\end{cases}$$

Comprueba que $f_n\in\mathbb{Z}[x]$ y es mónico. De nuevo sin usar los polinomios de Chebyshev como intermediarios, prueba que, para $x\in[-1,1]$, estos polinomios verifican $f_n(2x)=2\cos(n\theta)$, con $x=\cos(\theta)$; en particular, deduce la relación de recurrencia a partir de alguna expresión trigonométrica.[32]

[31] Este tipo de expresiones surgen de manera natural al resolver la ecuación cúbica disminuida $x^3+mx=n$, cuya solución es

$$x=\sqrt[3]{\frac{n}{2}+\sqrt{\frac{n^2}{4}+\frac{m^3}{27}}}-\sqrt[3]{-\frac{n}{2}+\sqrt{\frac{n^2}{4}+\frac{m^3}{27}}}\,.$$

En concreto, el ejemplo del ejercicio procede del *Ars Magna* de Cardano (siglo XVI); él mismo apuntaba que esa expresión aparentemente complicada no es más que el número «2» camuflado.

[32] Así, podríamos haber completado la demostración del teorema 6.5 sin usar polinomios de Chebyshev, como se hace en [78, § 6.3]. Hay que reconocer que es algo más rápido, pero se pierde una bonita interconexión con otros campos de la matemática. De todas formas, recordemos que los polinomios f_n son —simplemente— una pequeña reformulación de los de Chebyshev: $f_n(x)=2T_n(x/2)$.

6.10. Los polinomios de Chebyshev de segunda clase, denotados $\{U_n(x)\}_{n=0}^{\infty}$, se definen como

$$\begin{cases} U_0(x) = 1, \quad U_1(x) = 2x, \\ U_{n+1}(x) = 2xU_n(x) - U_{n-1}(x), \quad n \geq 1. \end{cases}$$

Prueba que $U_n(x)$ es un polinomio de grado n, con coeficientes enteros y coeficiente director 2^n, así como su «definición alternativa»: para $x \in [-1, 1]$, verifican

$$U_n(x) = \operatorname{sen}((n+1)\theta)/\operatorname{sen}(\theta), \quad \text{con } x = \cos(\theta).$$

6.11. Prueba que los polinomios de Chebyshev de segunda clase $\{U_n(x)\}_{n=0}^{\infty}$ satisfacen la relación de ortogonalidad

$$\int_{-1}^{1} U_n(x)U_m(x)(1-x^2)^{1/2}\,dx = \begin{cases} 0, & \text{si } n \neq m, \\ \pi/2, & \text{si } n = m. \end{cases}$$

6.12. Intenta probar el teorema 6.5 mediante polinomios de Chebyshev de segunda clase.

6.13. Analiza cómo se puede usar la fórmula de De Moivre

$$\cos(k\theta) + i\operatorname{sen}(k\theta) = e^{ik\theta} = \left(e^{i\theta}\right)^k = (\cos(\theta) + i\operatorname{sen}(\theta))^k$$

para hallar los polinomios de Chebyshev de primera y segunda clase. Aplícalo al caso $k = 5$.

6.14. Partiendo de la serie de Taylor de $\cos(x)$ y $\operatorname{sen}(x)$, comprueba que $\cos(1)$ y $\operatorname{sen}(1)$ son números irracionales (para evitar malentendidos: «1» está expresado en radianes).

6.15. Prueba que $\cos(\pi r)$ es un número algebraico para cada $r \in \mathbb{Q}$.

6.16. Sean $a, b, c \in \mathbb{Z}$. Averigua si el número $\cos(a^\circ b' c'')$ es raíz o no de un polinomio con coeficientes enteros y, en caso afirmativo, busca tal polinomio. (Como es habitual, estamos utilizando los símbolos «°», «′» y «″» para indicar grados, minutos y segundos sexagesimales.)

6.17. Halla sendos polinomios de grado estrictamente menor que 90 y coeficientes enteros que tengan a $\cos(1^\circ)$ y $\operatorname{sen}(1^\circ)$ como raíz.

6.18. Prueba que, cuando la variable r toma valores racionales, $\cos^2(r\pi)$ es racional si y sólo si

$$\cos(r\pi) \in \left\{0,\ \pm 1,\ \pm\frac{1}{2},\ \pm\frac{1}{\sqrt{2}},\ \pm\frac{\sqrt{3}}{2}\right\}.$$

6.19. Comprueba si son irracionales o no los siguientes números:

$$\log_{7/4}(6), \quad \log_{4/25}(125/8) \quad \text{y} \quad \log_{9/2}(3/4).$$

6.20. Sean a y b dos números reales positivos. Prueba que, si b no es una potencia racional de a, entonces $\log_a(b)$ es irracional.

6.21. En 1744, Euler enunció —sin demostración— que, si a y b son racionales, y b no es una potencia racional de a, entonces $\log_a(b)$ es trascendente. Utilizando el teorema de Gelfond-Schneider (esto es, de la solución al séptimo problema de Hilbert), comprueba que este hecho es cierto, incluso con hipótesis más generales, admitiendo la posibilidad de que a y b sean algebraicos.

6.22. Analiza la naturaleza aritmética —irracionalidad y trascendencia— de

$$\log(3)/\log(2).$$

6.23. Supón que sabes que $(e-1)/2$ no es un irracional cuadrático. ¿Cómo puedes usar esto para demostrar que e^2 es irracional?

6.24. Las habitaciones del denominado «hotel de Hilbert» están numeradas con $1, 2, 3, \ldots$, y hay tantas como números naturales, todas ellas individuales. Cierta noche, cuando todas las habitaciones están ocupadas, llega al hotel una persona que se quiere alojar. El hotelero, muy acostumbrado a sortear dificultades, logra hacerle hueco. ¿Cómo lo consigue? Un poco más tarde, llega un autobús con 40 viajeros, a los que también encuentra habitación; ¿cómo? Y ya al filo de la medianoche, aparece otro autobús ¡con tantos asientos como números naturales!, y lleno de pasajeros. ¿Cómo acomodar ahora a todos ellos?

6.25. Sigamos con el hotel de Hilbert del ejercicio anterior. Por la mañana, todos los huéspedes se van, y el hotel se queda vacío. Pero, entonces, llega un tren de Hilbert. Dicho tren tiene tantos vagones de pasajeros como números naturales; y cada vagón tiene una cantidad numerable de asientos. Por supuesto, el tren está lleno. ¿Cómo alojar a todos los pasajeros en el hotel?

6.26. Es bien sabido que, tras su paso por la vida terrenal, todos los matemáticos —o al menos la gran mayoría— van al cielo, donde sus almas disfrutan de vida eterna. Supongamos que, cada mes, las almas de estos matemáticos plantean un nuevo problema de teoría de números. Y que, cada año, se encuentra la solución de uno de los aún abiertos, siempre el más antiguo. ¿Cuántos problemas quedarán sin resolver «al final de los tiempos»? ¿Y si, en vez de uno, cada año se resuelven dos, siempre los planteados los dos últimos meses del año anterior?

6.27. Sea β un número algebraico no nulo. Demuestra que $1/\beta$ también es algebraico.

6.28. Sean α y β números algebraicos, con $\beta \neq 0$. Demuestra que α/β también es algebraico.

6.29. Como consecuencia de los resultados de la sección 6.8, prueba que si $\alpha \in \mathbb{C}$ es un número algebraico no nulo, entonces $\cosh(\alpha)$ y $\operatorname{senh}(\alpha)$ son trascendentes.

6.30. Demuestra o refuta la siguiente afirmación:

$$\alpha, \beta \text{ números trascendentes positivos } \Rightarrow \ \alpha^{\beta} \text{ trascendente.}$$

6.31. Sean a, b y c las raíces de $x^3 + 5x + 1 = 0$. Encuentra el polinomio cuyas raíces son $a/(b+c)$, $b/(a+c)$ y $c/(a+b)$.

6.32. Encuentra todos los números enteros n para los que el polinomio $f(x) = x^5 - nx - n - 2$ se puede expresar como producto de dos polinomios no constantes con coeficientes enteros.

6.33. Demuestra la identidad (6.21).

6.34. Prueba que los polinomios de Bernoulli satisfacen (6.22).

6.35. Comprueba la relación de simetría (6.23).

6.36. Prueba que los polinomios de Bernoulli satisfacen la fórmula aditiva

$$B_k(s+t) = \sum_{j=0}^{k} \binom{k}{j} B_j(s) t^{k-j}$$

y, como consecuencia, la siguiente expresión que proporciona los polinomios a partir de los números de Bernoulli:

$$B_k(t) = \sum_{j=0}^{k} \binom{k}{j} B_j t^{k-j}.$$

6.37. Prueba que los polinomios de Bernoulli cumplen $B_k(t+1) - B_k(t) = kt^{k-1}$ y, a partir de esta identidad, deduce la fórmula de sumación de Bernoulli

$$\sum_{j=1}^{n} j^p = \frac{B_{p+1}(n+1) - B_{p+1}(0)}{p+1}, \quad p = 1, 2, 3, \ldots.$$

6.38. Demuestra la denominada fórmula de Faulhaber:

$$\sum_{j=1}^{n} j^p = \frac{n^{p+1}}{p+1} + \frac{n^p}{2} + \sum_{k=2}^{p} \binom{p+1}{k} n^{p+1-k} B_k, \qquad p = 2, 3, 4, \ldots$$

Sugerencia. Fíjate en los dos ejercicios anteriores.



Tema 7
Funciones aritméticas

7.1. Funciones aritméticas notables
7.2. Números perfectos y amigos. Sucesiones alicuatorias
7.3. Producto de Dirichlet y transformada de Möbius
7.4. Funciones multiplicativas
7.5. Derivadas de funciones aritméticas
7.6. Generalización del producto de Dirichlet
7.7. Más tipos de convoluciones y sus inversiones de Möbius
7.8. Ejercicios

Abordaremos aquí lo que se conoce con el nombre de funciones aritméticas, así como algunas de sus propiedades; en todo esto se profundizará en el tema siguiente. Es de destacar que éste es uno de los temas centrales de la teoría de números; más concretamente, de la teoría de números analítica. Como consecuencia, la literatura matemática es pródiga en textos que tratan lo que aquí se estudia. Por ejemplo, [3, 5, 12, 41, 71, 87]; sin olvidarnos de [50] que tiene, además, multitud de detalles históricos y referencias a trabajos originales. Damos por supuesto que el lector tiene conocimientos básicos de aritmética; en particular, que está familiarizado con los conceptos de número primo, divisores, el teorema fundamental de la aritmética, etcétera. Comencemos estableciendo la notación que vamos a seguir.

En todo el tema, y aunque no lo mencionemos expresamente, las variables n, m, k y d siempre indicarán números naturales (generalmente, d llevará implícito el referirse a un «divisor» de algún número); lo mismo ocurrirá si tienen algún subíndice. La letra p (sola o con subíndice) siempre aludirá a números primos. Así, por ejemplo —y esto es importante recordarlo—, una expresión del tipo $\sum_p f(p)$ significa que el sumatorio se efectúa a través de todos los primos. Para evitar dudas, recalquemos que el 1 no se considera un número primo.[1] En los sitios en los que aparezca un primo elevado a un exponente, daremos por supuesto que el exponente es un entero mayor o igual que 1. Finalmente, cuando en una fórmula aparezcan varios «p con subíndices», asumiremos que se refieren a primos distintos; de este modo, $n = p_1^{a_1} \cdots p_k^{a_k}$ denota la *factorización canónica* de un entero $n > 1$.

Usaremos $d \,|\, n$ para indicar que d divide a n; y $d \nmid n$ significa que no lo divide. En un sumatorio, $\sum_{d|n} f(d)$ quiere decir que el sumatorio se extiende a todos los divisores de n, desde 1 al mismo n. También nos aparecerán sumatorios del tipo $\sum_{ab=n}$ o similar; eso significa que el sumatorio se efectúa sobre las dos índices a y b, pero ligados por la condición $ab = n$.

En cuanto a notación, concluyamos mencionado que usaremos $\operatorname{card}(A)$ (o, simplemente, $\operatorname{card} A$) para denotar el cardinal de un conjunto A (es decir, su número de elementos); en otros libros, también son usuales las notaciones $\#A$ y $|A|$.

Por definición, una *función aritmética*[2] es, simplemente, una función $f\colon \mathbb{N} \to \mathbb{R}$; o, sin ningún esfuerzo adicional, podemos permitir $f\colon \mathbb{N} \to \mathbb{C}$. Utilizaremos $\mathscr{F}$ para denotar el conjunto de todas las funciones aritméticas.

[1] Cuando en álgebra se define el concepto de número primo en un contexto más general que el de los números naturales, en la definición de primo es importante exigir que no sea una unidad.

[2] Obviamente, una función aritmética es exactamente lo mismo que una sucesión. Pero, claro, aquí nos interesarán las que sean relevantes para la teoría de números, y el tipo de estudio que se realiza está orientado al objetivo que se persigue. Así que está justificado utilizar un nombre distinto.

Aunque algo hay de teoría general de funciones aritméticas, dedicaremos especial atención a unas cuantas funciones que tienen algún interés especial; por ejemplo, por su contribución al estudio de las propiedades de la divisibilidad de enteros y la distribución de primos. Concretamente, analizaremos fundamentalmente las funciones a las que a veces se alude —de un manera informal— con el nombre de *funciones aritméticas notables*. Vamos con ellas.

7.1. Funciones aritméticas notables

En primer lugar, definamos la *función de Möbius* que —como veremos— desempeña un papel muy destacado en teoría de números, ya que surge de manera natural en algunas expresiones (como en el teorema 7.6 y los ejemplos que siguen). La función de Möbius, denotada siempre con la letra μ (incluso es común denominarla función «mu» de Möbius), es

$$\mu(n) = \begin{cases} 1, & \text{si } n = 1, \\ (-1)^k, & \text{si } n \text{ es producto de } k \text{ primos distintos}, \\ 0, & \text{si } n \text{ tiene un divisor primo elevado al cuadrado.} \end{cases}$$

Esta función ya aparece implícitamente en trabajos de Leonhard Euler anteriores a 1750, pero fue August Möbius, en 1832, el primero que se dedicó a investigar sistemáticamente sus propiedades.

Aunque ahora no veamos ninguna razón para ello, adelantemos que la importancia de la función μ deriva de la propiedad que mostramos a continuación. Su demostración —fíjese el lector en ella— es fácil de entender, pero sumamente ingeniosa. En un par de líneas, aparecen involucrados el teorema fundamental de la aritmética, números combinatorios, y el binomio de Newton.

Teorema 7.1.
La función de Möbius verifica

$$\sum_{d|n} \mu(d) = \begin{cases} 1, & \text{si } n = 1, \\ 0, & \text{si } n > 1. \end{cases} \tag{7.1}$$

DEMOSTRACIÓN. Si $n = 1$, es trivial. En otro caso, sea $n = \prod_{j=1}^{k} p_j^{a_j}$. Entonces

$$\sum_{d|n} \mu(d) = \sum_{\substack{d|n \\ d \text{ libre de} \\ \text{cuadrados}}} \mu(d) = \mu(1) + \sum_{\substack{A\subseteq\{1,\dots,k\} \\ A\neq\varnothing}} \mu\Big(\prod_{j\in A} p_j\Big)$$

$$= 1 + \sum_{\substack{A\subseteq\{1,\dots,k\} \\ A\neq\varnothing}} (-1)^{\operatorname{card}(A)} = \sum_{j=0}^{k} \binom{k}{j} (-1)^j = (1-1)^k = 0. \qquad \square$$

Otra interesante función aritmética es la que se conoce con el nombre de *indicatriz de Euler*,[3] quien la introdujo alrededor de 1760. Esta función —también llamada función «phi» de Euler o (en inglés) función *totient*— cuenta el número de enteros positivos menores o iguales que n que son primos con n. En otras palabras,

$$\varphi(n) = \operatorname{card}\{m \le n : \operatorname{mcd}(m,n) = 1\} = \sum_{\substack{m\le n \\ \operatorname{mcd}(m,n)=1}} 1.$$

Dos funciones relacionadas con los divisores de un número son las siguientes:

$$\sigma(n) = \sum_{d|n} d = \text{suma de los divisores de } n,$$

$$\tau(n) = \operatorname{card}\{d \in \mathbb{N} : d\,|\,n\} = \sum_{d|n} 1 = \text{número de divisores de } n$$

(en muchos libros, la función $\tau(n)$ se denota $d(n)$). A veces, $\sigma(n)$ se generaliza así:

$$\sigma_a(n) = \sum_{d|n} d^a \quad (\text{con } a \in \mathbb{R}).$$

Casi sin darnos cuenta, σ_a también abarca a τ como caso particular, pues $\sigma_0 = \tau$. En general, a las funciones σ_a se les suele denominar *funciones divisor*.

No resulta difícil establecer una expresión que proporcione directamente $\sigma(n)$ y $\tau(n)$ a partir de la descomposición en factores primos de n, como vemos a continuación (en el ejercicio 7.3 se propone hallar la correspondiente fórmula para $\sigma_a(n)$):

[3] Esta función ya nos ha aparecido en el tema 2 de este libro, pues $\varphi(n)$ tiene gran importancia al estudiar congruencias módulo n.

Teorema 7.2.
Las funciones σ y τ verifican $\sigma(1) = \tau(1) = 1$; y, para $n = p_1^{a_1} \cdots p_k^{a_k} > 1$,

$$\sigma(n) = (1 + p_1 + p_1^2 + \cdots + p_1^{a_1}) \cdots (1 + p_k + p_k^2 + \cdots + p_k^{a_k}) \tag{7.2}$$

$$= \frac{p_1^{a_1+1} - 1}{p_1 - 1} \cdots \frac{p_k^{a_k+1} - 1}{p_k - 1} \tag{7.3}$$

y

$$\tau(n) = (a_1 + 1) \cdots (a_k + 1). \tag{7.4}$$

(Si, como es habitual, interpretamos que un producto sin factores vale 1, estas fórmulas también son válidas para el caso $n = 1$.)

Demostración. Es inmediato que $\sigma(1) = \tau(1) = 1$, luego analicemos el caso $n > 1$. Al desarrollar la parte derecha de (7.2) aparecen, como sumandos, todos los posibles divisores de $p_1^{a_1} \cdots p_k^{a_k}$; por tanto, obtenemos $\sigma(n)$. Además, $\tau(n)$ es el número de sumandos, de donde se obtiene (7.4). Finalmente, para llegar a (7.3) basta usar la fórmula para la suma de una progresión geométrica (véase el ejercicio 7.1). □

La siguiente función que abordamos es la *función de Liouville*, que se define como

$$\lambda(n) = \begin{cases} 1, & \text{si } n = 1, \\ (-1)^{a_1 + \cdots + a_k}, & \text{si } n = p_1^{a_1} \cdots p_k^{a_k} \end{cases}$$

(obsérvese que el exponente $a_1 + \cdots + a_k$ es el número de divisores primos de n, contando su multiplicidad; esto también sirve para el caso $n = 1$, que no tiene ningún divisor primo, luego $\lambda(1) = (-1)^0 = 1$).

Nota 7.3. La función de Liouville sirve como ejemplo de que una conjetura puede ser cierta para muchos números y aun así ser falsa. En 1919, el matemático húngaro György Pólya observó que, aparentemente, la mayoría de los números naturales (al menos la mitad) menores que cualquier número dado tienen una cantidad impar de factores primos, y conjeturó que eso iba a ocurrir siempre. El número de factores primos de k es par o impar según sea $\lambda(k) = 1$ o $\lambda(k) = -1$; por tanto, la conjetura de Pólya se expresa diciendo que

$$L(n) = \sum_{k=1}^{n} \lambda(k) \leq 0$$

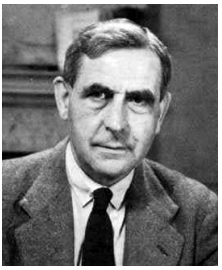
para cualquier entero $n \geq 2$.



Joseph Liouville
(1809–1882)



György Pólya
(1887–1985)



Albert Edward Ingham
(1900–1967)

En 1942, Albert E. Ingham ideó un ingenioso método que, según él suponía, podía servir para encontrar un contraejemplo. Usando métodos derivados de los de Ingham, pero ya con ayuda informática, C. Brian Haselgrove refutó la conjetura en 1958 y estimó que debía existir un contraejemplo alrededor de $1.845 \cdot 10^{361}$, pero sin darlo explícitamente. Poco más tarde, en 1960, y también basándose en el trabajo de Ingham, R. Sherman Lehman descubrió el primer contraejemplo explícito, con $n = 906\,180\,359$. Ya en 1980, Minoru Tanaka comprobó que el contraejemplo más pequeño es $n = 906\,150\,257$. Se ignora si $L(n)$ cambia de signo una cantidad infinita de veces. ◁

Ahora, otra función con nombre propio: se llama *función de von Mangoldt* a

$$\Lambda(n) = \begin{cases} \log(p), & \text{si } n = p^m \text{ para algún } m \geq 1, \\ 0, & \text{en otro caso.} \end{cases}$$

Posiblemente el lector haya pensado que esta función no parece tener ningún interés especial (¿qué puede tener que ver un logaritmo con la teoría de números, en general, y con los números primos, en particular?), y que se podrían haber definido muchas otras parecidas. Sin embargo, la función de von Mangoldt desempeña una importante labor en el estudio de la distribución de primos, como veremos un poco más adelante. De momento, merece la pena que comentemos una peculiaridad suya; y es que $\Lambda(1) = 0$, pero $f(1) = 1$ para el resto de las funciones aritméticas que hemos definido hasta ahora.

Definamos asimismo la función $\pi(n)$, que cuenta el número de primos menores o iguales que n, esto es,

$$\pi(n) = \operatorname{card}\{p \leq n : p \text{ primo}\} = \sum_{p \leq n} 1.$$

Uno de los objetivos de la teoría de números es conocer la distribución de sus componentes básicos, los números primos, algo muy lejos de resultar sencillo. Esta vez, es evidente que $\pi(n)$ juega un papel fundamental en esto, así que estudiar sus propiedades tiene gran interés. En particular, un resultado clave es lo que se conoce como teorema de los números primos, que establece que

$$\lim_n \frac{\pi(n)}{n/\log(n)} = 1. \tag{7·5}$$

Originariamente, esto fue conjeturado por Adrien Marie Legendre (1798) y Carl Friedrich Gauss (1849); tras años de intentos de diversos matemáticos, fueron Jacques Hadamard y Charles de la Vallée-Poussin en 1896 quienes, de manera independiente, lograron demostrarlo. Sin demasiada dificultad, se puede probar que (7·5) es equivalente[4] a

$$\lim_n \frac{1}{n}\sum_{k=1}^{n} \mu(k) = 0 \tag{7·6}$$

y a

$$\lim_n \frac{1}{n}\sum_{k=1}^{n} \Lambda(k) = 1. \tag{7·7}$$

consiguiéndose de este modo procedimientos alternativos de abordar la demostración del teorema de los números primos (obviamente, las demostraciones de (7·6) y (7·7) tampoco son simples). Pero, de momento, aquí hemos definido la función $\pi(n)$ únicamente por completitud, pues no nos volverá a aparecer en todo el tema. La equivalencia entre (7·5), (7·6) y (7·7) la veremos en el tema 8; finalmente, el teorema de los números primos lo probaremos en el tema 10.

Prácticamente como notación, usaremos también las siguientes funciones. En primer lugar, la función cero,

$$0(n) = 0 \quad \text{para todo } n \in \mathbb{N}.$$

La función δ, definida como

$$\delta(n) = \begin{cases} 1, & \text{si } n = 1, \\ 0, & \text{si } n > 1 \end{cases}$$

(a veces se le llama función identidad y se la denota como I). La función unidad,

$$u(n) = 1 \quad \text{para todo } n \in \mathbb{N}$$

[4] Para profundizar en lo que queremos decir con esta equivalencia, léanse los comentarios al principio de la sección 8.7 del tema 8.

(también se la suele denotar, simplemente, como 1). Y, finalmente, para $a \in \mathbb{R}$, la función N^a, que es

$$N^a(n) = n^a \quad \text{para todo } n \in \mathbb{N}.$$

Si $a = 1$ usaremos $N = N^1$, esto es, $N(n) = n$ para todo $n \in \mathbb{N}$; y, si $a = 0$, se tiene $N^0 = u$.

7.2. Números perfectos y amigos. Sucesiones alicuatorias

Antes de continuar con el objetivo principal de este tema, hagamos, en esta sección, una pequeña digresión. Muchas propiedades aritméticas se definen en términos de las funciones aritméticas que hemos definido. La función más usada para esto es σ, aunque suele ser más cómodo emplear la variante $s(n) = \sigma(n) - n$, que representa la suma de los *divisores propios* de n, esto es, de todos sus divisores (incluyendo el 1) excepto él mismo; con el mismo significado, también se dice que $s(n)$ es la suma de las *partes alícuotas* de n. Como veremos, los conceptos que vamos a definir dan lugar a numerosos problemas abiertos de enunciado bastante sencillo; si el lector quiere profundizar en esto, le aconsejamos que hojee [49, sección B].

Denominamos *perfecto* a cada número natural que sea igual a la suma de sus partes alícuotas. Es decir, n es perfecto si y sólo si $s(n) = n$ (o sea, $\sigma(n) = 2n$). Si $s(n) > n$, decimos de n que es *abundante*; si $s(n) < n$, *defectuoso*. Los primeros números perfectos son 6, 28 y 496; los primeros abundantes, 12, 18, 20, 24 y 30. Se desconoce si existen números perfectos impares,[5] aunque, en caso de que los hubiera, diversos resultados teóricos y búsquedas con ordenador aseguran que deberían ser enormes. En cambio, los pares están mucho mejor caracterizados. Ya en los *Elementos* de Euclides aparece un modo de encontrar números perfectos pares; y fue Euler quien demostró que todos surgen de esa manera:

Teorema 7.4. Euclides y Euler.
Sea $n \in \mathbb{N}$ un número par. Entonces, n es perfecto si y sólo si

$$n = 2^{p-1}(2^p - 1)$$

y $2^p - 1$ es primo (además, esto requiere que p sea primo).

[5] Posiblemente éste sea el problema matemático más antiguo cuya respuesta aún no se conoce.

DEMOSTRACIÓN. En primer lugar, comprobemos que esos números son perfectos. Al ser 2^p-1 primo, los divisores de $n=2^{p-1}(2^p-1)$ son $1,\ 2,\ 2^2,\ \ldots,\ 2^{p-1}$, y $1\cdot(2^p-1),\ 2\cdot(2^p-1),\ 2^2\cdot(2^p-1),\ \ldots,\ 2^{p-1}\cdot(2^p-1)$. Entonces,

$$\begin{aligned}\sigma(n) &= 1+2+2^2+\cdots+2^{p-1}+(1+2+2^2+\cdots+2^{p-1})(2^p-1)\\ &= \frac{2^p-1}{2-1}+\frac{2^p-1}{2-1}(2^p-1)=2^p(2^p-1)=2n.\end{aligned}$$

Recíprocamente, veamos que cualquier número perfecto par tiene esa forma. Obviamente, podemos escribir $n=2^{k-1}n'$ con $k\geq 2$ y n' impar. Del teorema 7.2 es inmediato deducir[6] que $\sigma(n)=(2^k-1)\sigma(n')$. Por tanto, al ser $\sigma(n)=2n$, deberá cumplirse

$$2^k n'=2n=\sigma(n)=(2^k-1)\sigma(n'),$$

luego $(2^k-1)\,|\,n'$. Así pues, $n'=(2^k-1)n''$ y $\sigma(n')=2^k n''$. Los dos números n' y n'' dividen a n', y además son sus únicos divisores, ya que

$$n'+n''=(2^k-1)n''+n''=2^k n''=\sigma(n').$$

En consecuencia, $n''=1$ y $n'=2^k-1$ es primo.

Que, en estas circunstancias, k debe ser primo (y por eso en el enunciado hemos usado p en vez de k) lo dejamos para el lector (ejercicio 7.4). □

Los números de la forma 2^p-1 con p primo son conocidos como *números de Mersenne*; y, si son primos, *primos de Mersenne*. En este libro ya los estudiamos en el tema 3 (concretamente, en el apartado 3.4.3). Hay tests especializados —en particular, el test de Lucas-Lehmer, que aparece enunciado en el teorema 3.27— muy eficientes dedicados a comprobar si un número de Mersenne es o no primo. Esto hace que, habitualmente, el récord de «el primo más grande conocido» siempre lo ostente uno de esos primos. Con el auge de internet, estas búsquedas se han sistematizado, y en ellas no sólo colaboran los especialistas en teoría de números computacional, sino también bastantes aficionados entusiastas que dedican a ello el tiempo libre de sus ordenadores. Fruto de tales búsquedas, en 2017 se encontró el primo de Mersenne número cincuenta; unos cuantos de estos primos tienen millones de dígitos.

Pero no se sabe si hay o no infinitos primos de Mersenne (y, como consecuencia, si existen infinitos números perfectos pares). Ni tampoco si hay infinitos números de Mersenne compuestos. Pueden verse más detalles en [49, sección A3].

[6] De hecho, la función σ satisface $\sigma(rs)=\sigma(r)\sigma(s)$ siempre que $\mathrm{mcd}(r,s)=1$ —lo cual se expresa diciendo que σ es una función multiplicativa—, pero esto no lo veremos hasta algo más adelante en este mismo tema.



Euclides de Alejandría
(≈ 325 a.C. – ≈ 265 a.C.)



Thabit ibn Qurra
(836–901)



Pierre de Fermat
(1601–1665)

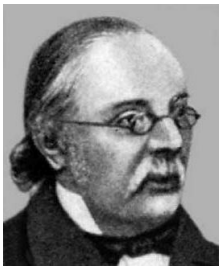


René Descartes
(1596–1650)

Se denominan *números amigos* a una pareja (n, m) tales que $s(n) = m$ y $s(m) = n$ (o, lo que es equivalente, $\sigma(n) = \sigma(m) = n + m$). La menor pareja de números amigos es $(220, 284)$, que ya era conocida por Pitágoras. Otros dos ejemplos son $(2620, 2924)$ y $(5020, 5564)$. En el siglo IX, el matemático Thabit ibn Qurra, quien trabajaba como astrónomo para el califa de Bagdad, desarrolló un método que permitió encontrar algunos otros (véase el ejercicio 7.7). Los matemáticos europeos de los siglos XVII y XVIII (entre ellos Fermat, Descartes y Euler) hallaron bastantes más; en particular la pareja $(12285, 14595)$, en la que ambos números son impares. Pero, curiosamente, el segundo menor par de números amigos, $(1184, 1210)$, no fue descubierto hasta 1866, cuando lo encontró Niccolò Paganini,[7] un estudiante italiano de 16 años de edad. Actualmente se conocen varios miles.

En la segunda mitad del siglo XX se propusieron diversos procesos que, a partir de una pareja de amigos, generan otras parejas de números que son amigos con bas-

[7] No tiene nada que ver con el violinista y compositor del mismo nombre, que había muerto un cuarto de siglo antes.



Eugène Charles Catalan (1814–1894)



Leonard Eugene Dickson (1874–1954)

tante probabilidad. Cuando se han aplicado, mediante el uso de ordenador, a tablas de amigos ya conocidos, se han obtenido más parejas nuevas que las previamente conocidas. Esto es un argumento heurístico en favor de la existencia de infinitos pares de amigos, pero no se tiene una demostración rigurosa. Tampoco se sabe si existe alguna pareja de números amigos tal que uno de los números sea par y el otro impar.

Una generalización es lo que se denominan ciclos de *números sociables* (o *pandilla* de números). Son tuplas $(a_1, \ldots, a_l)$ tales que $s(a_1) = a_2$, $s(a_2) = a_3$, $\ldots$, $s(a_l) = a_1$. Antes de 1970 sólo se conocían dos ciclos de longitud l mayor que dos (obviamente, los de longitud uno son los números perfectos y los de longitud dos son los amigos): uno tenía longitud 5 y, el otro, 28. Ambos habían sido descubiertos en 1918 por Paul Poulet —un matemático autodidacta apasionado por la aritmética—. Con la llegada de los ordenadores, se han encontrado bastantes más (de longitudes 4, 5, 6, 8, 9; la gran mayoría, de longitud 4). A estas alturas, a nadie sorprenderá que digamos que no se sabe si existen o no infinitos de estos ciclos. Ni tampoco hay nada establecido sobre qué longitudes son posibles y cuáles no. Se desconoce, por ejemplo, si existe algún ciclo de longitud 3.

Partiendo de un número n e iterando la función s se llegan a definir las *sucesiones de sumas de partes alícuotas* o, para abreviar, *sucesiones alicuatorias*: dado un entero positivo n construimos la sucesión $\{s^k(n)\}_{k\in\mathbb{N}}$ definida recursivamente mediante $s^1(n) = s(n)$, $s^2(n) = s(s^1(n))$, $\ldots$, $s^k(n) = s(s^{k-1}(n))$.

Para una de tales sucesiones, existen cuatro posibilidades: (a) Que la sucesión termine en 1 (siendo el número anterior un primo). (b) Que la sucesión llegue a un número perfecto (y a partir de entonces permanezca constante). (c) Que llegue a un par de amigos o un ciclo. (d) Que no esté acotada. Eugène Catalan en 1887 y Leonard Dickson en 1913 conjeturaron que la posibilidad (d) nunca ocurría. Pese a

todo el esfuerzo que se ha hecho (y el que escribe ha colaborado en esto: véase [18, 13]), la conjetura de Catalan-Dickson continúa sin ser demostrada o refutada. En el momento de redactar estas líneas, la primera sucesión cuyo comportamiento no se conoce es la que empieza en $n = 276$.

7.3. Producto de Dirichlet y transformada de Möbius

Dadas dos funciones aritméticas f y g, su suma y producto punto a punto,

$$(f+g)(n) = f(n) + g(n), \quad (f \cdot g)(n) = f(n)g(n), \qquad n \in \mathbb{N},$$

proporcionan dos nuevas funciones aritméticas. Obviamente, ambas operaciones son conmutativas y asociativas. El elemento neutro de $+$ es la función 0, y el de $\cdot$ es la función u. El inverso respecto a $+$ de una función f siempre existe, y es la función $-f$, definida de la manera obvia. Sin embargo, respecto a $\cdot$, la función $1/f$ —que debería ser la inversa de f— sólo puede definirse si $f(n)$ no se anula para ningún $n \in \mathbb{N}$.

Otra manera de «multiplicar» es la siguiente. Si f y g son dos funciones aritméticas, se llama *producto* (o convolución) *de Dirichlet* a la función aritmética $f * g$ dada por[8]

$$(f*g)(n) = \sum_{d|n} f(d)g\left(\frac{n}{d}\right) = \sum_{ab=n} f(a)g(b). \tag{7.8}$$

Dar esta definición —que puede parecer extraña— no es un capricho, sino que expresiones del tipo $\sum_{d|n} f(d)g(n/d)$ surgen de manera natural en diversos problemas de teoría de números, y varias funciones aritméticas tienen interrelaciones de ese tipo. Fue el matemático Eric Temple Bell (también conocido como escritor de novelas de ciencia ficción) quien, en 1915, tuvo la idea de tratarlas como un nuevo tipo de multiplicación.

Ejemplo. Comprobar que $\mu * u = \delta$.

Solución. Es, simplemente, la fórmula (7.1) con otro lenguaje:

$$\mu * u(n) = \sum_{d|n} \mu(d)u(n/d) = \sum_{d|n} \mu(d) = \delta(n). \qquad \Diamond$$

[8] Aunque no esté en las convenciones que hemos adoptado al principio del tema, en la expresión que sigue, deben ser, obviamente, $a, b \in \mathbb{N}$. Podíamos haber escrito $\sum_{dd'=n} f(d)g(d')$, que parece más directo (pues $d' = n/d$), pero haber usado a y b la hace —a nuestro entender— una expresión más simétrica.

A continuación, vamos a analizar las propiedades del producto de Dirichlet. Aunque no utilizaremos más esta nomenclatura, muchos lectores reconocerán que el teorema siguiente está probando que $(\mathscr{F}, +, *)$ tiene estructura de anillo conmutativo y con unidad (el resto de las propiedades que se deben cumplir para que $(\mathscr{F}, +, *)$ sea un anillo, como las que corresponden a la suma y la propiedad distributiva, o ya las hemos comentado, o son inmediatas). Es más, si a las dos operaciones $+$ y $*$ añadimos la multiplicaciones por escalares (reales o complejos), entonces $\mathscr{F}$ tiene estructura de álgebra.

Teorema 7.5.
El producto de Dirichlet cumple las siguientes propiedades:

(a) Es conmutativo y asociativo: $f * g = g * f$ y $(f * g) * h = f * (g * h)$ para todo $f, g, h \in \mathscr{F}$. (Como consecuencia de la asociatividad, se puede usar $f * g * h$, sin paréntesis.)

(b) El elemento neutro (o, en otras palabras, la identidad de $*$) es la función δ, esto es, $\delta * f = f$ para todo $f \in \mathscr{F}$.

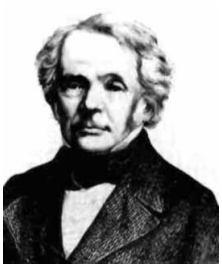
(c) Decimos que función $f \in \mathscr{F}$ es inversible (respecto a $*$) si existe otra función $f^{-1} \in \mathscr{F}$ (la denotamos así y la denominamos función inversa de f) tal que $f * f^{-1} = \delta$. Pues bien: una función $f \in \mathscr{F}$ es inversible si y sólo si $f(1) \neq 0$. Si éste es el caso, su inversa es[9]

$$f^{-1}(1) = \frac{1}{f(1)},$$
$$f^{-1}(n) = -\frac{1}{f(1)} \sum_{\substack{d|n \\ d<n}} f\left(\frac{n}{d}\right) f^{-1}(d), \quad \text{para } n > 1. \tag{7.9}$$

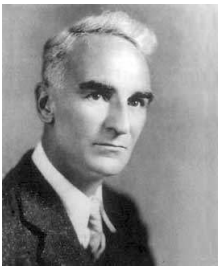
(d) Si $g = f^{-1}$ entonces $f = g^{-1}$.

(e) Si f y g son inversibles, entonces $f * g$ es inversible, y $(f * g)^{-1} = g^{-1} * f^{-1}$.

[9] Obsérvese que esta fórmula es recursiva: partiendo de $f^{-1}(1) = 1/f(1)$, para calcular el resto de los $f^{-1}(n)$ hay que conocer previamente los $f^{-1}(m)$ con $m < n$ (en realidad, no todos, sólo los divisores de n). Por otra parte, hay a quien le puede parecer que (7.9) es una expresión sofisticada, pero es justo lo que se requiere para que, cuando $n > 1$, la comprobación $f * f^{-1}(n) = 0$ funcione (véase la última línea de la demostración de (c)). Cualquiera que intentara averiguar cómo tiene que ser $f^{-1}(n)$ para que $f * f^{-1}(n) = 0$ $(n > 1)$ iba a llegar sin dificultades a la misma expresión.



August Ferdinand Möbius (1790–1868)



Eric Temple Bell (1883–1960)

DEMOSTRACIÓN. (a) La conmutatividad es inmediata (fijándose en la expresión de la derecha en (7.8)), y la asociatividad se deduce de

$$\begin{aligned}(f*g)*h(n) &= \sum_{ab=n}(f*g)(a)h(b)\\ &= \sum_{ab=n}\left(\sum_{cd=a} f(c)g(d)\right)h(b) = \sum_{bcd=n} f(c)g(d)h(b).\end{aligned}$$

(b) Es obvio, pues, en $\delta * f(n) = \sum_{d|n}\delta(d)f(n/d)$, el único sumando no nulo es el correspondiente a $d=1$, para el cual $\delta(1)=1$.

(c) Si existe inversa f^{-1}, entonces $1=\delta(1)=(f*f^{-1})(1)=f(1)f^{-1}(1)$, luego debe ser $f(1)\neq 0$. Recíprocamente, si $f(1)\neq 0$, definamos f^{-1} según (7.9) —no hay ningún problema en hacerlo así—. Comprobemos que, efectivamente, $f*f^{-1}=\delta$:

$$\delta(1)=1=f(1)f^{-1}(1)=f*f^{-1}(1)$$

y, para $n>1$,

$$\begin{aligned}f*f^{-1}(n) &= \sum_{d|n} f\left(\frac{n}{d}\right)f^{-1}(d)\\ &= f(1)f^{-1}(n)+\sum_{\substack{d|n\\ d<n}} f\left(\frac{n}{d}\right)f^{-1}(d) = 0=\delta(n).\end{aligned}$$

(d) Es inmediato por conmutatividad: si $f*g=\delta$ entonces también $g*f=\delta$.

(e) Basta comprobar que $(f*g)*(g^{-1}*f^{-1})=\delta$, y eso es inmediato usando la propiedad asociativa:

$$(f*g)*(g^{-1}*f^{-1}) = f*g*g^{-1}*f^{-1} = f*\delta*f^{-1} = f*f^{-1}=\delta. \qquad \square$$

Ejemplo. Con la definición de funciones inversas que acabamos de dar, la fórmula (7.1) y el ejemplo de la página 410 se expresan diciendo que $\mu^{-1} = u$ o, de manera equivalente, $u^{-1} = \mu$. ◊

7.3.1. La fórmula de inversión de Möbius

Algunas de las funciones aritméticas que hemos presentado las hemos definido mediante una expresión de tipo $\sum_{d|n} f(d)$, como $\sigma_a(n) = \sum_{d|n} d^a$. En otras ocasiones, tenemos una función aritmética f previamente definida y, por diversas razones, resulta interesante calcular $\sum_{d|n} f(d)$, que también se puede expresar como $(f * u)(n)$. Esta operación, efectuada para cada $f \in \mathscr{F}$, tiene un nombre: al operador

$$f \in \mathscr{F} \longmapsto f * u \in \mathscr{F}$$

se le llama *transformada de Möbius*. El siguiente teorema prueba que la transformada de Möbius tiene un operador inverso, que consiste en efectuar la convolución con μ.

Teorema 7.6. Fórmula de inversión de Möbius.
Sean f y g dos funciones aritméticas. Entonces,

$$g(n) = \sum_{d|n} f(d) \;\Leftrightarrow\; f(n) = \sum_{d|n} g(d)\mu\left(\frac{n}{d}\right), \tag{7.10}$$

esto es, $g = f * u$ si y sólo si $f = g * \mu$.

Demostración. Ya sabemos que $\mu * u = \delta$ (teorema 7.1 o ejemplo de la página 410) y, por conmutatividad, $u * \mu = \delta$. Con esto, a partir de $g = f * u$, efectuando la convolución con μ tendríamos $g * \mu = f * u * \mu = f * \delta = f$. Recíprocamente, de $f = g * \mu$ se seguiría $f * u = g * \mu * u = g * \delta = g$. □

Nota 7.7. La demostración anterior es impecable, pero a algunos les puede parecer un tanto algebraica. Mostraremos ahora otra forma de probar (7.10) —aunque, en el fondo, es la misma— en la que no hacemos intervenir explícitamente el producto de Dirichlet. Es verdad que esta nueva demostración utiliza menos herramientas. ¿Quiere eso decir que es más sencilla? Posiblemente sea cuestión de gusto.

En realidad, analizaremos

$$g(n) = \sum_{d|n} f\left(\frac{n}{d}\right) \;\Leftrightarrow\; f(n) = \sum_{d|n} \mu(d) g\left(\frac{n}{d}\right), \tag{7.11}$$

que es equivalente a (7.10) (basta cambiar d por n/d en los sumatorios). Probaremos únicamente «⇒»; la implicación «⇐» es similar. En $\sum_{d|n}\mu(d)g(n/d)$, sustituyamos la expresión izquierda de (7.11), agrupemos los términos tales que $kd=m$ y usemos (7.1):

$$\begin{aligned}\sum_{d|n}\mu(d)g\left(\frac{n}{d}\right) &= \sum_{d|n}\mu(d)\left(\sum_{k|(n/d)} f\left(\frac{n}{kd}\right)\right)\\ &= \sum_{\substack{k,d\\ kd|n}}\mu(d)f\left(\frac{n}{kd}\right) = \sum_{m|n}\sum_{kd=m}\mu(d)f\left(\frac{n}{m}\right)\\ &= \sum_{m|n}\left(\sum_{d|m}\mu(d)\right)f\left(\frac{n}{m}\right) = 1\cdot f\left(\frac{n}{1}\right)+0 = f(n),\end{aligned}$$

tal como buscábamos. ◁

7.3.2. Algunas relaciones entre funciones aritméticas

Vamos ahora a mostrar bastantes relaciones entre las funciones aritméticas notables que hemos definido; varias más aparecerán en los ejercicios. Como veremos, el uso de la fórmula de inversión de Möbius nos facilitará bastante las cosas.

Ejemplo. La función τ cumple

$$\tau(n)=\sum_{d|n}1=\sum_{d|n}1\cdot 1=\sum_{d|n}u(d)u(n/d),$$

es decir, $\tau = u*u$. Ahora, la fórmula de inversión muestra que $u=\tau*\mu$, o sea $\sum_{d|n}\tau(d)\mu(n/d)=1$ para todo $n\in\mathbb{N}$. ◊

Ejemplo. La función σ_a (con $a\in\mathbb{R}$) cumple

$$\sigma_a(n)=\sum_{d|n}d^a=\sum_{d|n}d^a\cdot 1=\sum_{d|n}N^a(d)u(n/d),$$

esto es, $\sigma_a = N^a * u$. Luego, invirtiendo, $N^a=\sigma_a*\mu$; en otras palabras,

$$\sum_{d|n}\sigma_a(d)\mu(n/d)=n^a$$

para todo $n\in\mathbb{N}$. Por otra parte, nótese que el ejemplo anterior es un caso particular de éste, el correspondiente a $a=0$. ◊

Algo más laboriosas de demostrar son las relaciones que exponemos en las dos proposiciones siguientes. En ambas, probaremos de manera directa una igualdad, y la fórmula de inversión de Möbius nos proporcionará —automáticamente— otra.

Proposición 7.8.

Recordando que N denota la función $N(n) = n$ para todo $n \in \mathbb{N}$, se tiene

(a) $N = \varphi * u$, *es decir,*

$$n = \sum_{d|n} \varphi(d).$$

(b) $\varphi = N * \mu$, *es decir,*

$$\frac{\varphi(n)}{n} = \sum_{d|n} \frac{\mu(d)}{d}. \tag{7.12}$$

Demostración. Sólo hace falta demostrar (a), porque entonces (b) se deduce de la fórmula de inversión (nótese que (7.12) responde más directamente a $\varphi = \mu * N$, pero, por la conmutatividad de $*$, da lo mismo). Para cada $d \,|\, n$, sea $A_d = \{\, k \in \mathbb{N} : 1 \le k \le n,\ \operatorname{mcd}(k,n) = d \,\}$. Es fácil comprobar que los conjuntos A_d son disjuntos y que su unión es $\bigcup_{d|n} A_d = \{1, 2, \ldots, n\}$ (demuéstralo: ejercicio 7.10). Además,

$$k \in A_d \ \Leftrightarrow\ \frac{k}{d} \in \mathbb{N},\ 1 \le \frac{k}{d} \le \frac{n}{d},\ \operatorname{mcd}\left(\frac{k}{d}, \frac{n}{d}\right) = 1.$$

Así que la aplicación

$$\begin{aligned} A_d &\longrightarrow \left\{\, q \in \mathbb{N} : 1 \le q \le n/d,\ \operatorname{mcd}(q, n/d) = 1 \,\right\} \\ k &\longmapsto k/d \end{aligned}$$

es biyectiva. Entonces, $\operatorname{card}(A_d) = \varphi(n/d)$ y

$$n = \sum_{d|n} \operatorname{card}(A_d) = \sum_{d|n} \varphi(n/d) = \sum_{d|n} \varphi(d). \qquad \square$$

La demostración del siguiente resultado muestra cómo la función de von Mangoldt surge de manera natural del teorema fundamental de la aritmética.

Proposición 7.9.

Sea log *la función logaritmo definida sobre* $\mathbb{N}$ *(para que sea función aritmética). Entonces*

(a) $\log = \Lambda * u$, *es decir,*

$$\log(n) = \sum_{d|n} \Lambda(d).$$

(b) $\Lambda = \log * \mu$, *es decir,*

$$\Lambda(n) = \sum_{d|n} \mu(d) \log\left(\frac{n}{d}\right) = -\sum_{d|n} \mu(d)\log(d). \tag{7.13}$$

Demostración. Lo mismo que en la proposición anterior, sólo hace falta demostrar el apartado (a). Con esto, la primera igualdad de (7.13) se deduce de la fórmula de inversión; y, para la segunda, basta usar $\log(n/d) = \log(n) - \log(d)$ y

$$\log(n) \sum_{d|n} \mu(d) = \log(n)\delta(n) = 0$$

(nótese que, para $n = 1$, se anula $\log(n)$). Así pues, vamos con (a). Para $n = 1$, es trivial. En otro caso, sea $n = p_1^{a_1} \cdots p_k^{a_k}$, de donde, tomando logaritmos,

$$\log(n) = \sum_{j=1}^{k} a_j \log(p_j).$$

Para cada d que divide a n, se tiene

$$\Lambda(d) \neq 0 \;\Leftrightarrow\; d = p_j^m,\; 1 \leq m \leq a_j,\; 1 \leq j \leq k.$$

Por lo tanto,

$$\sum_{d|n} \Lambda(d) = \sum_{j=1}^{k} \sum_{m=1}^{a_j} \Lambda(p_j^m) = \sum_{j=1}^{k} a_j \log(p_j) = \log(n). \qquad \square$$

7.4. Funciones multiplicativas

Comencemos con dos definiciones. A una función aritmética f la denominamos *multiplicativa* si no es la función nula y satisface

$$f(nm) = f(n)f(m) \quad \text{siempre que } \operatorname{mcd}(n, m) = 1.$$

Y, de manera más exigente, decimos que función aritmética no nula f es *completamente multiplicativa* cuando

$$f(nm) = f(n)f(m) \quad \text{para todo } n, m \in \mathbb{N}.$$

Una función multiplicativa f queda determinada por sus valores sobre las potencias de primos; en efecto, si conocemos cada $f(p^a)$, para $n = p_1^{a_1} \cdots p_k^{a_k}$ tendremos

$$f(n) = f(p_1^{a_1} \cdots p_k^{a_k}) = f(p_1^{a_1}) \cdots f(p_k^{a_k}).$$

Y una función completamente multiplicativa queda determinada por sus valores sobre los primos; en este caso, conociendo los $f(p)$ se sigue que

$$f(n) = f(p_1^{a_1} \cdots p_k^{a_k}) = f(p_1)^{a_1} \cdots f(p_k)^{a_k}.$$

Ejemplo. Es evidente, a partir de su definición, que las funciones N^a, δ y u son completamente multiplicativas. Lo mismo ocurre con la función de Liouville, λ. ◊

Ejemplo. Comprobar que μ es multiplicativa, pero no completamente multiplicativa.

Solución. Que μ no es completamente multiplicativa es inmediato: por ejemplo, se cumple $\mu(9) = 0 \neq 1 = \mu(3)^2$. Para comprobar que es multiplicativa, obsérvese que, dadas dos descomposiciones en factores primos $n = p_1^{a_1} \cdots p_k^{a_k}$ y $m = q_1^{b_1} \cdots q_l^{b_l}$, se tiene $\mathrm{mcd}(n, m) = 1$ si y sólo si $p_j \neq q_i$ para todo i, j. Con esto, basta aplicar la definición de μ, analizando los diversos casos que se presentan (nueve casos, al combinar los tres tipos de valores de $\mu(n)$ con los tres de $\mu(m)$). El caso más interesante en cuando $n, m > 1$, $n = p_1 \cdots p_k$, $m = q_1 \cdots q_l$ y $p_j \neq q_i$ para todo i, j; entonces, $\mu(nm) = (-1)^{k+l} = \mu(n)\mu(m)$. ◊

Ejemplo. Comprobar que la función de von Mangoldt Λ no es multiplicativa.

Solución. Es muy fácil encontrar contraejemplos: $\Lambda(2) = \log(2)$, $\Lambda(3) = \log(3)$, pero $\Lambda(6) = 0 \neq \Lambda(2)\Lambda(3)$. Mostremos otro, más sencillo si cabe: $\log(2) = \Lambda(2) = \Lambda(2 \cdot 1) \neq \Lambda(2)\Lambda(1)$ pues $\Lambda(1) = 0$. (Nótese, además, que Λ no satisface la propiedad (a) del teorema siguiente, luego no puede ser multiplicativa.) ◊

Teorema 7.10.
Sean f y g dos funciones aritméticas. Entonces,

(a) f multiplicativa $\Rightarrow f(1) = 1$.
(b) f y g multiplicativas $\Rightarrow f * g$ multiplicativa.
(c) f y $f * g$ multiplicativas $\Rightarrow g$ multiplicativa.
(d) f multiplicativa $\Rightarrow$ existe f^{-1} y es multiplicativa.

Demostración. (a) Como $f(n) = f(n \cdot 1) = f(n)f(1)$, basta elegir algún $n \in \mathbb{N}$ tal que $f(n) \neq 0$ (seguro que existe alguno pues —recordémoslo— en la definición de función multiplicativa se incluye la condición de que no sea la función nula).

(b) Sean n, m sin divisores comunes (por supuesto, nos referimos a «mayores que 1»). Entonces, $c \mid nm$ si y sólo si $c = ab$ con $a \mid n$ y $b \mid m$; además, esta factorización de c es única, y a, b tampoco tienen divisores comunes (ni $n/a, m/b$). Por tanto,

$$\begin{aligned}(f*g)(nm) &= \sum_{c\mid nm} f(c)g\left(\frac{nm}{c}\right) = \sum_{\substack{a\mid n\\ b\mid m}} f(ab)g\left(\frac{n}{a}\cdot\frac{m}{b}\right)\\ &= \sum_{\substack{a\mid n\\ b\mid m}} f(a)f(b)g\left(\frac{n}{a}\right)g\left(\frac{m}{b}\right)\\ &= \left(\sum_{a\mid n} f(a)g\left(\frac{n}{a}\right)\right)\left(\sum_{b\mid m} f(b)g\left(\frac{m}{b}\right)\right) = (f*g)(n)(f*g)(m).\end{aligned}$$

(c) Por inducción sobre el producto nm, probaremos que, cuando $\mathrm{mcd}(n, m) = 1$, se cumple $g(nm) = g(n)g(m)$. Si $nm = 1$, entonces $n = m = 1$. Como f y $f * g$ son multiplicativas, se tiene que $f(1) = 1$ y $(f * g)(1) = 1$. Luego $1 = (f * g)(1) = f(1)g(1) = g(1)$. De aquí,

$$g(nm) = g(1) = 1 = g(n)g(m).$$

Sean ahora n, m con $\mathrm{mcd}(n, m) = 1$ y $nm > 1$. Por hipótesis de inducción, asumamos que, para $ab < nm$ con $\mathrm{mcd}(a, b) = 1$, el resultado es cierto (es decir, que se cumple $g(ab) = g(a)g(b)$). Entonces,

$$\begin{aligned}(f*g)(n)(f*g)(m) = (f*g)(nm) &= \sum_{\substack{a\mid n\\ b\mid m}} f(ab)g\left(\frac{n}{a}\cdot\frac{m}{b}\right)\\ &= \sum_{\substack{a\mid n\\ b\mid m\\ ab>1}} f(a)f(b)g\left(\frac{n}{a}\right)g\left(\frac{m}{b}\right) + g(nm)\\ &= \sum_{\substack{a\mid n\\ b\mid m}} f(a)f(b)g\left(\frac{n}{a}\right)g\left(\frac{m}{b}\right) - g(n)g(m) + g(nm)\\ &= (f*g)(n)(f*g)(m) - g(n)g(m) + g(nm),\end{aligned}$$

luego $g(nm) = g(n)g(m)$, tal como se quería probar.

(d) Por (a), existe f^{-1} (recordar (c) del teorema 7.5). Ahora, basta aplicar (c) a f y $f * f^{-1}$, que son multiplicativas. □

Nota 7.11. Los lectores que tengan conocimientos de estructuras algebraicas, habrán notado —sin duda— el siguiente hecho: En el teorema anterior hemos probado que el producto de Dirichlet conserva la propiedad «ser multiplicativa», y que cualquier función multiplicativa tiene inversa. Por tanto, con el producto de Dirichlet, las funciones multiplicativas tienen estructura de grupo. ◁

Como consecuencia del teorema, se tiene esto:

Corolario 7.12.

Las funciones μ, φ, τ y σ_a son multiplicativas.[10]

Demostración. Basta tener en cuenta, respectivamente, las siguientes relaciones: $\delta = \mu * u$ (y aplicar que δ y u son multiplicativas, y (c) del teorema); $\varphi = N * \mu$ ($N = N^1$ y μ son multiplicativas, y se aplica (b)); $\tau = u * u$ (u multiplicativa y (b)) o $u = \tau * \mu$ (u y μ multiplicativas, y (c)); $\sigma_a = N^a * u$ (N^a y u multiplicativas, y (b)). □

Ejemplo. Utilizando que la función φ es multiplicativa, y hallando previamente su valor sobre primos y sobre potencias de primos, comprueba que, para $n \geq 1$,

$$\varphi(n) = n \prod_{p \mid n} \left(1 - \frac{1}{p}\right). \tag{7.14}$$

[10] Que la función μ es multiplicativa ya lo hemos comprobado antes, aunque de otra manera: directamente a partir de su definición. Lo mismo se podía haber hecho con otras, pero se requiere una expresión explícita de la función (con ello nos referimos a una expresión que nos proporcione $f(n)$ a partir de la descomposición en factores de n); en particular, hemos obtenido expresiones explícitas para τ y σ (y σ_a en los ejercicios), y de ahí se deduce fácilmente que esas funciones son multiplicativas, aunque no lo hemos comentado hasta ahora.

Este procedimiento, que no requiere la maquinaria que hemos desarrollado, tiene el inconveniente de que utiliza una notación más farragosa —multitud de subíndices y puntos suspensivos, por ejemplo—, y que hay que tener mucho cuidado de no equivocarse. Por otra parte, también es más fácil encontrar una expresión explícita de una función si, por algún otro método, hemos probado previamente que es multiplicativa (ya que, entonces, basta determinar $f(p^a)$).

Por ejemplo, en este tema estableceremos el valor de $\varphi(n)$ a partir de $\varphi(p^a)$ (véase (7.16)), y, para ello, estamos ahora probando que la función φ es multiplicativa. En [5, teoremas 2.4 y 2.5] se sigue el camino contrario: primero se prueba la expresión explícita para φ (apoyándose, para ello, en la fórmula (7.12)) y, a partir de ahí, sus propiedades (entre ellas, que es una función multiplicativa). Es más, nosotros aún veremos otra forma de probar (7.14) que utiliza más herramientas (la fórmula del producto del teorema 7.15), lo cual hace que la demostración final sea mucho más corta.

Como contrapartida, [5, teorema 2.13] no obtiene directamente la expresión explícita para σ_a, sino que lo hace comprobando que es multiplicativa (al ser $\sigma_a = N^a * u$) y hallando su valor sobre potencias de primos.

Por último, no podemos dejar de decir que, ya en el tema 2 de este libro, también hemos mostrado una manera directa de calcular $\varphi(n)$ a partir de su definición $\varphi(n) = \operatorname{card}\{m \leq n : \operatorname{mcd}(m, n) = 1\}$. Una conocida herramienta conjuntista —el principio de inclusión-exclusión— nos permitía, sin dificultad, evaluar el cardinal de ese conjunto; en concreto, véase el teorema 2.10.

SOLUCIÓN. Recordemos que, por definición, $\varphi(n) = \text{card}\{m \leq n : \text{mcd}(m,n) = 1\}$. Si $n = p$ (un primo), es claro que $\varphi(p) = p-1$, pues todos los números menores que p son primos con p (es más, es evidente que se verifica la siguiente equivalencia: n es primo si y sólo si $\varphi(n) = n-1$).

Si $n = p^a$ (potencia de un primo), los únicos números menores o iguales que p^a que no son primos con él son los que tienen algún factor p, es decir, p, $2p$, $3p$, $\ldots$, $(p^{a-1}-1)p$ y $p^{a-1}p$, que son p^{a-1} números. Los $p^a - p^{a-1}$ números restantes son primos con p^a; por tanto

$$\varphi(p^a) = p^a - p^{a-1}. \tag{7.15}$$

Finalmente, como φ es multiplicativa, para $n = p_1^{a_1}p_2^{a_2}\cdots p_k^{a_k}$ se tiene

$$\begin{aligned}\varphi(n) &= \varphi(p_1^{a_1}p_2^{a_2}\cdots p_k^{a_k}) = \varphi(p_1^{a_1})\varphi(p_2^{a_2})\cdots\varphi(p_k^{a_k}) \\ &= (p_1^{a_1} - p_1^{a_1-1})(p_2^{a_2} - p_2^{a_2-1})\cdots(p_k^{a_k} - p_k^{a_k-1}) \\ &= p_1^{a_1}p_2^{a_2}\cdots p_k^{a_k}\left(1-\frac{1}{p_1}\right)\left(1-\frac{1}{p_2}\right)\cdots\left(1-\frac{1}{p_k}\right) \\ &= n\prod_{p|n}\left(1-\frac{1}{p}\right),\end{aligned} \tag{7.16}$$

tal como buscábamos. ◊

7.4.1. Inversas de funciones completamente multiplicativas

Ya sabemos que cualquier función aritmética f tal que $f(1) \neq 0$ tiene inversa, e incluso tenemos una fórmula que nos proporciona f^{-1}: véase (7.9). Esto es aplicable a cualquier f multiplicativa, que forzosamente debe cumplir $f(1) = 1$. Si la función es completamente multiplicativa, su inversa es particularmente sencilla de obtener, como se ve a continuación:

Teorema 7.13.
Sea f una función multiplicativa. Entonces, f es completamente multiplicativa si y sólo si $f^{-1} = \mu \cdot f$.

DEMOSTRACIÓN. En primer lugar, recordemos que no tenemos que preocuparnos por la existencia de inversa, pues el teorema 7.10 nos lo garantiza. Así, comencemos suponiendo que f es completamente multiplicativa y definamos $g = \mu \cdot f$, es decir,

$g(n) = \mu(n)f(n)$ para todo $n \in \mathbb{N}$; comprobemos que g es la inversa de f. Para cada n,

$$(g * f)(n) = \sum_{d|n} \mu(d)f(d)f\left(\frac{n}{d}\right) = f(n)\sum_{d|n}\mu(d)$$
$$= f(n)\delta(n) = \delta(n),$$

luego, efectivamente, $g = f^{-1}$.

De manera inversa, supuesto que $f^{-1} = \mu \cdot f$, veamos que f es completamente multiplicativa. Basta comprobar que $f(p^a) = f(p)^a$ para todo p primo y para todo $a \in \mathbb{N}$. Si $a = 1$ es trivial. Si $a \geq 2$,

$$0 = \delta(p^a) = ((\mu \cdot f) * f)(p^a) = \sum_{d|p^a} \mu(d)f(d)f\left(\frac{p^a}{d}\right)$$
$$= \mu(1)f(1)f(p^a) + \mu(p)f(p)f(p^{a-1}) + 0 = f(p^a) - f(p)f(p^{a-1}),$$

luego $f(p^a) = f(p)f(p^{a-1})$. Reiterando se sigue que $f(p^a) = f(p)^a$. □

Ejemplo.

(a) N^a es completamente multiplicativa, luego $(N^a)^{-1} = \mu \cdot N^a$.

(b) $\varphi = N * \mu$, luego $\varphi^{-1} = N^{-1} * \mu^{-1}$ (obsérvese que, ahora, con N^{-1} estamos denotando la función inversa de N, no N^a con $a = -1$). De aquí, $\varphi^{-1} = (\mu \cdot N) * u$, es decir,

$$\varphi^{-1}(n) = \sum_{d|n} d\mu(d). \tag{7.17}$$

(c) $\sigma_a = N^a * u$, así que $\sigma_a^{-1} = (N^a)^{-1} * u^{-1} = (\mu \cdot N^a) * \mu$; en detalle,

$$\sigma_a^{-1}(n) = \sum_{d|n} d^a\mu(d)\mu\left(\frac{n}{d}\right).$$

En el caso particular correspondiente a $a = 0$ se tiene $\tau = u * u$ y $\tau^{-1} = \mu * \mu$.

(d) λ es completamente multiplicativa, luego $\lambda^{-1} = \mu \cdot \lambda = |\mu|$. Es decir:

$$\lambda^{-1}(n) = \begin{cases} 1, & \text{si } n \text{ está libre de cuadrados,} \\ 0, & \text{si } n \text{ es divisible por algún cuadrado mayor que 1.} \end{cases}$$ ◊

Acabamos de ver que la función $\lambda^{-1} = \mu \cdot \lambda = |\mu|$ identifica los números libres de cuadrados. Otro bonito resultado es el siguiente, que muestra cómo la función $\lambda * u$ sirve para identificar cuadrados:

Proposición 7.14.
La función $\lambda * u$ es

$$(\lambda * u)(n) = \begin{cases} 1, & \text{si } n \text{ es un cuadrado,} \\ 0, & \text{si } n \text{ no es un cuadrado.} \end{cases}$$

Demostración. Ambas funciones, $\lambda * u$ y la de la derecha, son multiplicativas, luego, para ver que son iguales, sólo hace falta comprobar que coinciden sobre cada p^a (p primo, $a \in \mathbb{N}$). Y esto es sencillo:

$$(\lambda * u)(p^a) = \sum_{d|p^a} \lambda(d) = \sum_{k=0}^{a} \lambda(p^k) = \sum_{k=0}^{a} (-1)^k = \begin{cases} 1, & \text{si } a \text{ es par,} \\ 0, & \text{si } a \text{ es impar} \end{cases}$$
$$= \begin{cases} 1, & \text{si } p^a \text{ es un cuadrado,} \\ 0, & \text{si } p^a \text{ no es un cuadrado.} \end{cases}$$

□

7.4.2. Fórmula del producto para funciones multiplicativas

Sin detenernos en preámbulos, abordemos el resultado principal de este apartado. Después veremos un par de ejemplos de uso.

Teorema 7.15. Fórmula del producto.
Si f es una función multiplicativa, entonces

$$\sum_{d|n} \mu(d) f(d) = \prod_{p|n} \big(1 - f(p)\big). \tag{7.18}$$

Demostración. Sea $g(n) = \sum_{d|n} \mu(d) f(d)$, es decir, $g = (\mu \cdot f) * u$. Como μ, f y u son multiplicativas, también lo son $\mu \cdot f$ y el producto de Dirichlet $g = (\mu \cdot f) * u$. Por otra parte, la expresión $\prod_{p|n}(1 - f(p))$ también define una función multiplicativa (ejercicio 7.30). Sabemos que, para probar que dos funciones multiplicativas son iguales, basta ver que coinciden sobre las potencias de primos. Entonces, sólo tenemos que comprobar

$$g(p^a) = 1 - f(p), \quad p \text{ primo}, \ a \in \mathbb{N}.$$

Pero esto es inmediato: $g(p^a) = \sum_{d|p^a} \mu(d) f(d) = \sum_{k=0}^{a} \mu(p^k) f(p^k) = 1 \cdot f(1) + (-1) \cdot f(p) + 0 = 1 - f(p)$. □

Nota 7.16. De manera similar se prueba que, si f es multiplicativa, entonces

$$\sum_{d|n} f(d) = \prod_{p|n} \sum_{p^k|n} f(p^k)$$

(para evitar dudas, aclaremos que, en $\sum_{p^k|n}$, el primo p es fijo; sólo varía el exponente k, desde 0 hasta el máximo posible de modo que $p^k \,|\, n$). La demostración de esta identidad de deja para el lector (véase el ejercicio 7.31). ◁

Ejemplo. Usando la fórmula del producto, deducir

$$\varphi(n) = n \prod_{p|n} \left(1 - \frac{1}{p}\right)$$

(expresión que —recordémoslo— ya habíamos obtenido mediante un procedimiento diferente; véase (7.16)).

Solución. Como $\varphi = N * \mu = \mu * N$, se tiene

$$\varphi(n) = (\mu * N)(n) = n \sum_{d|n} \mu(d) \frac{1}{d} = n \prod_{p|n} \left(1 - \frac{1}{p}\right),$$

donde en el último paso hemos usado la fórmula del producto. ◊

Ejemplo. Comprobar que

$$\varphi^{-1}(n) = \prod_{p|n} (1 - p).$$

Solución. Usando la relación $\varphi^{-1} = (\mu \cdot N) * u$ de (7.17) y la fórmula del producto, llegamos a

$$\varphi^{-1}(n) = ((\mu \cdot N) * u)(n) = \sum_{d|n} \mu(d) d = \prod_{p|n} (1 - p). \qquad ◊$$

7.5. Derivadas de funciones aritméticas

Con funciones aritméticas, se puede definir algo similar a las derivadas de funciones de variable real. Es claro que la definición no puede ser la misma, pues a un $n \in \mathbb{N}$ no nos podemos aproximar —tanto como queramos— por otros enteros, luego ni siquiera tiene sentido el concepto de «límite cuando tendemos a n». Pero, al fin y al cabo, lo que importa de una definición no es la definición en sí, sino que tenga

propiedades y consecuencias interesantes. Así, vamos a dar una definición que, en principio, puede parecer arbitraria, pero que, según veremos (teorema 7.17), tiene propiedades similares[11] a las de las derivadas de funciones de variable real. De paso, también usaremos una notación parecida.

Dada una función aritmética f, se llama *derivada* de f a la función aritmética f' definida mediante

$$f'(n) = f(n)\log(n), \quad n \in \mathbb{N}.$$

Obviamente, para derivada segunda se usa f''; así, $f''(n) = f(n)\log^2(n)$, $n \in \mathbb{N}$.

Ejemplo. Comprobar que las derivadas de las funciones δ y u son $\delta' = 0$ y $u' = \log = \Lambda * u$.

Solución. Para la primera, por definición de derivada, $\delta'(n) = \delta(n)\log(n)$; si $n > 1$, $\delta(n) = 0$; y, si $n = 1$, $\log(n) = 0$. Por tanto, $\delta'(n) = 0$ para todo $n \in \mathbb{N}$. Para la segunda, es claro que $u'(n) = u(n)\log(n) = \log(n)$; y $\log = \Lambda * u$ ya lo hemos demostrado antes (véase el apartado (a) de la proposición 7.9). ◊

Teorema 7.17.
Sean f y g dos funciones aritméticas. Entonces:

(a) $(f+g)' = f' + g'$.

(b) $(f*g)' = f'*g + f*g'$.

(c) $(f^{-1})' = -f'*(f*f)^{-1}$, *si f es inversible (es decir, si $f(1) \neq 0$).*

Demostración.

(a) Es inmediato.

(b) Para cada $n \in \mathbb{N}$, se tiene

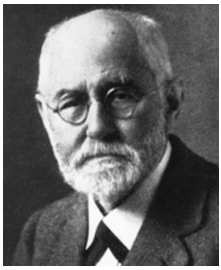
$$\begin{aligned}(f*g)'(n) &= \sum_{d|n} f(d)g\left(\frac{n}{d}\right)\log(n)\\ &= \sum_{d|n} f(d)g\left(\frac{n}{d}\right)\log(d) + \sum_{d|n} f(d)g\left(\frac{n}{d}\right)\log\left(\frac{n}{d}\right)\\ &= (f'*g)(n) + (f*g')(n),\end{aligned}$$

donde —y queremos recalcarlo expresamente— el hecho crucial que conduce a la demostración es la propiedad $\log(n) = \log(d) + \log(n/d)$.

[11] En particular, en la demostración del apartado (b) del teorema 7.17 se pone de manifiesto cómo el logaritmo que aparece en la definición de derivada es esencial a la hora de probar la fórmula de la derivada del producto.



Atle Selberg
(1917–2007)



Hans Carl Friedrich von
Mangoldt (1854–1925)

(c) Basta derivar $\delta = f * f^{-1}$. En efecto, usando $\delta' = 0$ y aplicando (b) obtenemos $0 = f' * f^{-1} + f * (f^{-1})'$, luego

$$(f^{-1})' = -f' * f^{-1} * f^{-1} = -f' * (f * f)^{-1}. \qquad \square$$

Nota 7.18. Tal como ya hemos comentado en su demostración, el teorema anterior es cierto gracias a que los logaritmos transforman productos en sumas, esto es, $\log(nm) = \log(n) + \log(m)$. A las funciones aritméticas $L\colon \mathbb{N} \to \mathbb{R}$ que verifican $L(nm) = L(n) + L(m)$ para todo $n, m \in \mathbb{N}$ se las denomina *completamente aditivas*. Con tales funciones, se podría definir una nueva «derivada» $f' = f \cdot L$, y tendría propiedades similares a las expuestas en el teorema. En realidad, no se gana mucho con eso, ya que, si L es una función completamente aditiva y creciente, deberá existir una constante multiplicativa C tal que $L(n) = C\log(n)$ para todo $n \in \mathbb{N}$ (en el ejercicio 7.34 se propone al lector que lo compruebe).

Más aún, lo mismo es cierto con funciones *aditivas*, esto es, las que satisfacen

$$L(nm) = L(n) + L(m) \quad \text{siempre que } \operatorname{mcd}(n, m) = 1;$$

cualquier función aditiva creciente L también debe verificar $L = C\log$ para alguna constante C. Pero probar este resultado es considerablemente más difícil que el correspondiente a funciones completamente aditivas; una demostración —muy simplificada sobre la original de Pál Erdős en 1946— se puede encontrar en [41, §8.33, p. 265]. Para profundizar algo más sobre esto, véanse los comentarios y referencias del citado libro y de [90, p. 133]. $\triangleleft$

Utilizando derivadas podemos obtener rápidamente una fórmula debida a Atle Selberg en la que aparecen las funciones de Möbius y de von Mangoldt, y que tiene gran interés pues se utiliza como punto de partida de una demostración elemental

(en el sentido de que no usa variable compleja) del teorema de los números primos; así se hace, por ejemplo, en [71]. Nosotros abordaremos este teorema en el tema 10, aunque con otra demostración completamente distinta.

Teorema 7.19. Identidad de Selberg.
Se verifica $\Lambda' + \Lambda * \Lambda = \mu * u''$ *o, en otras palabras,*

$$\Lambda(n)\log n + \sum_{d|n} \Lambda(d)\Lambda\left(\frac{n}{d}\right) = \sum_{d|n} \mu(d)\log^2\left(\frac{n}{d}\right).$$

Demostración. Derivando la igualdad $u' = \Lambda * u$, resulta

$$u'' = \Lambda' * u + \Lambda * u' = \Lambda' * u + \Lambda * \Lambda * u;$$

y multiplicando —referido al producto de Dirichlet $*$, claro— por $\mu = u^{-1}$,

$$\mu * u'' = \Lambda' + \Lambda * \Lambda,$$

que es lo que había que demostrar. □

7.6. Generalización del producto de Dirichlet

Hasta ahora, en este tema sólo nos han aparecido funciones de variable discreta; a partir de este momento, trataremos también con funciones de variable continua. Tener funciones definidas en un subintervalo de $\mathbb{R}$ permite, entre otras cosas, poder usar sobre ellas diversas técnicas típicas del análisis matemático.

En esta sección veremos una serie de resultados que —en cierto sentido— se parecen a muchos de los que ya hemos obtenido para funciones de variable discreta. Aunque pensamos que lo que aquí vamos a exponer tiene interés en sí mismo, somos conscientes de su carácter bastante abstracto. No obstante, en el siguiente tema podremos sacar partido de la serie de herramientas y resultados que hemos agrupado en esta sección. Allí los usaremos para extraer bonitas consecuencias sobre cosas más concretas. Comencemos poniendo de manifiesto cuál es nuestro principal objetivo con todo esto. En otras palabras, intentemos, al menos, motivar lo que vamos a hacer.

Una de las cuestiones que interesan a la teoría de números es el comportamiento (en un sentido que, de momento, no precisaremos) de las funciones aritméticas. Si f es una de tales funciones, muy a menudo ocurre que, según avanza la variable n, los valores $f(n)$ que va tomando sean demasiado erráticos como para poder extraer

consecuencias interesantes. En estos casos, es común estudiar las sumas parciales

$$F(x) = \sum_{n \leq x} f(n) \tag{7.19}$$

(o sus medias, si dividimos (7.19) por x), que suelen tener mejores propiedades. Como prevención, aclaremos que el sumatorio $\sum_{n \leq x}$ de (7.19) significa $\sum_{n=1}^{\lfloor x \rfloor}$. Normalmente lo usaremos con $x \geq 1$, pero, si alguna vez aparece $\sum_{n \leq x}$ con $x < 1$, se considerará que la suma es 0.

Por ver un ejemplo más concreto, pensemos en la función $\mu(n)$, que toma valores 0, 1 y -1 sin ningún orden aparente. Si no tenemos un n concreto y lo hemos descompuesto en factores, no podemos evaluar $\mu(n)$; en particular, conocer $\mu(n-1)$ o $\mu(n+1)$ no nos proporciona ninguna información sobre cuánto puede valer $\mu(n)$. Sin embargo, cabe esperar que sus sumas $\sum_{n \leq x} \mu(n)$ tengan un comportamiento más suave. Idealmente, para x grande, uno podría llegar a estimar $\sum_{n \leq x} \mu(n)$ con un cierto error sin necesidad de calcularlo explícitamente. Pero no adelantemos acontecimientos; de este tipo de problemas nos ocuparemos en el siguiente tema (con respecto a μ véase, en concreto, la sección 8.6).

Con funciones aritméticas nos aparecían sumatorios con la forma $\sum_{d|n}$, y los expresábamos, de una manera bastante general, por medio del producto de Dirichlet. Este tratamiento ha resultado ser muy útil en el desarrollo teórico que hemos seguido a lo largo de este tema. Ahora estamos interesados en sumatorios del tipo $\sum_{n \leq x}$, así que una buena idea es dar una definición similar a la del producto de Dirichlet de dos funciones aritméticas, y que se adapte bien a las nuevas circunstancias. Esta vez, en lugar de dos funciones aritméticas, intervendrán una función aritmética y una función de variable real.

Si α es una función aritmética y $F\colon [1,\infty) \to \mathbb{C}$, la *convolución* de α y F es la función $\alpha \circ F\colon [1,\infty) \to \mathbb{C}$ definida mediante

$$(\alpha \circ F)(x) = \sum_{n \leq x} \alpha(n) F\left(\frac{x}{n}\right); \tag{7.20}$$

en ocasiones, a este proceso se le denomina convolución de Dirichlet generalizada.

Las sumas parciales de una función aritmética se pueden expresar como una convolución. Concretamente, si llamamos U a la función que hace $U(x) = 1$ para todo $x \geq 1$, es claro que $\sum_{n \leq x} f(n) = (f \circ U)(x)$. (De manera análoga a como teníamos $\sum_{d|n} f(d) = (f * u)(n)$.)

Nota 7.20. La definición de $\circ$ que acabamos de dar generaliza el producto de Dirichlet $*$ en el siguiente sentido: Si $F = 0$ en $[1,\infty) \setminus \mathbb{N}$, entonces $\alpha \circ F = \alpha * F$ en $\mathbb{N}$ y $\alpha \circ F = 0$ en $[1,\infty) \setminus \mathbb{N}$ (obsérvese que, cuando escribimos $\alpha * F$, estamos

cometiendo un pequeño abuso de notación, pues estamos usando la misma F para referirnos a la función aritmética que resulta de restringir la F original a $\mathbb{N}$).

En efecto. Si $x \in [1,\infty)\setminus\mathbb{N}$, también $x/n \in [1,\infty)\setminus\mathbb{N}$, luego todos los $F(x/n)$ de (7.20) son nulos, y $(\alpha\circ F)(x)=0$. Y, si $x\in\mathbb{N}$, sólo hace falta darse cuenta de que los únicos sumandos no nulos del $\sum_{n\le x}$ de (7.20) son los correspondientes a los d que dividen a x (por supuesto, nos estamos refiriendo a división entera). Por tanto $\sum_{n\le x}$ se transforma en $\sum_{d|x}$, que es el sumatorio que aparece en la definición de producto de Dirichlet. ◁

Ejemplo. Para cualquier función $F\colon [1,\infty)\to\mathbb{C}$, se tiene $\delta\circ F = F$. ◊

Vamos ahora a analizar las propiedades de $\circ$. Puesto que $\circ$ es una cierta generalización de $*$, nos interesa conocer qué propiedades de $*$ se conservan; en particular, ¿qué es lo que se puede adaptar del teorema 7.5? En primer lugar, es claro que la conmutatividad no tiene sentido con $\circ$. Por otra parte, ya hemos visto —en el ejemplo anterior— que δ toma el papel de un cierto elemento neutro, lo que podríamos llamar *elemento neutro por la izquierda*. Veamos ahora cómo adaptar la asociatividad, que se transforma en algo así como una *propiedad asociativa mixta* que relaciona $\circ$ y $*$.

Proposición 7.21.
Sean α y β funciones aritméticas y $F\colon[1,\infty)\to\mathbb{C}$. Entonces, $\alpha\circ(\beta\circ F) = (\alpha * \beta)\circ F$.

Demostración. Para cada $x\in[1,\infty)$, se tiene

$$\begin{aligned}(\alpha\circ(\beta\circ F))(x) &= \sum_{n\le x}\alpha(n)(\beta\circ F)\left(\frac{x}{n}\right) = \sum_{n\le x}\alpha(n)\sum_{m\le x/n}\beta(m)F\left(\frac{x}{nm}\right)\\ &= \sum_{\substack{n,m\\ nm\le x}}\alpha(n)\beta(m)F\left(\frac{x}{nm}\right) = \sum_{k\le x}\left(\sum_{nm=k}\alpha(n)\beta(m)\right)F\left(\frac{x}{k}\right)\\ &= \sum_{k\le x}(\alpha*\beta)(k)F\left(\frac{x}{k}\right) = ((\alpha*\beta)\circ F)(x).\end{aligned}$$

□

Como consecuencia, también se puede generalizar la existencia de inversa, en el sentido de que las funciones inversas respecto de $*$ adoptan un cierto papel de funciones inversas con $\circ$; en concreto, las inversas respecto de $*$ son *inversas por la izquierda* respecto de $\circ$. Veámoslo a continuación:

Proposición 7.22.
Si α es una función aritmética inversible y $F,G\colon[1,\infty)\to\mathbb{C}$, entonces

$$G=\alpha\circ F \;\Leftrightarrow\; F=\alpha^{-1}\circ G.$$

Es decir,

$$G(x) = \sum_{n\le x} \alpha(n) F\left(\frac{x}{n}\right) \text{ para todo } x \;\Leftrightarrow\; F(x) = \sum_{n\le x} \alpha^{-1}(n) G\left(\frac{x}{n}\right) \text{ para todo } x.$$

Demostración. Supongamos que $G = \alpha \circ F$. Entonces, por la proposición anterior,

$$\alpha^{-1} \circ G = \alpha^{-1} \circ (\alpha \circ F) = (\alpha^{-1} * \alpha) \circ F = \delta \circ F = F,$$

como queríamos comprobar. El recíproco se demuestra de manera similar. □

El siguiente caso particular es especialmente importante:

Teorema 7.23. Fórmula de inversión de Möbius generalizada.
Sean $F, G\colon [1,\infty) \to \mathbb{C}$ y α una función aritmética completamente multiplicativa. Entonces

$$G = \alpha \circ F \;\Leftrightarrow\; F = (\mu \cdot \alpha) \circ G.$$

Es decir,

$$G(x) = \sum_{n\le x} \alpha(n) F\left(\frac{x}{n}\right) \;\forall x \;\Leftrightarrow\; F(x) = \sum_{n\le x} \mu(n)\alpha(n) G\left(\frac{x}{n}\right) \;\forall x.$$

Demostración. El teorema 7.13 asegura que, por ser α completamente multiplicativa, existe α^{-1} y es $\alpha^{-1} = \mu \cdot \alpha$. Basta ahora aplicar la proposición 7.22. □

A continuación, vamos a ver una fórmula general que relaciona las sumas parciales de funciones aritméticas arbitrarias f y g con las de su producto de Dirichlet $f * g$.

Teorema 7.24.
Sean f, g funciones aritméticas y $h = f * g$ su producto de Dirichlet; además, para $x \ge 1$, tomemos

$$F(x) = \sum_{n\le x} f(n), \quad G(x) = \sum_{n\le x} g(n) \quad y \quad H(x) = \sum_{n\le x} h(n).$$

Entonces,

$$H(x) = \sum_{n\le x} f(n) G\left(\frac{x}{n}\right) = \sum_{n\le x} g(n) F\left(\frac{x}{n}\right).$$

Demostración. Si definimos $U(x) = 1$ para $x \geq 1$, podemos escribir $F = f \circ U$ y $G = g \circ U$. Entonces, por la propiedad asociativa mixta entre $*$ y $\circ$ (proposición 7.21), tenemos

$$f \circ G = f \circ (g \circ U) = (f * g) \circ U = h \circ U = H,$$
$$g \circ F = g \circ (f \circ U) = (g * f) \circ U = h \circ U = H,$$

tal como buscábamos. □

Nota 7.25. En realidad, la demostración anterior se puede hacer también directamente, sin utilizar $\circ$ ni sus propiedades (aunque, como es habitual, cuanta menos maquinaria usemos, más intrincada es la demostración). Comprobemos la primera igualdad, pues la otra es similar. Por una parte,

$$H(x) = \sum_{n \leq x} \sum_{d|n} f(d) g\left(\frac{n}{d}\right) = \sum_{dc \leq x} f(d) g(c);$$

y, por otra,

$$\sum_{n \leq x} f(n) G\left(\frac{x}{n}\right) = \sum_{n \leq x} f(n) \sum_{m \leq x/n} g(m) = \sum_{nm \leq x} f(n) g(m).$$

Luego, efectivamente,

$$H(x) = \sum_{n \leq x} f(n) G(x/n).$$ ◁

Un interesante caso particular es el siguiente:

Corolario 7.26.
Sea f una función aritmética. Entonces, para $x \geq 1$, se cumple

$$\sum_{n \leq x} \sum_{d|n} f(d) = \sum_{n \leq x} f(n) \left\lfloor \frac{x}{n} \right\rfloor.$$

Demostración. Basta tomar, en el teorema, $g = u$ (es decir, $g(n) = 1$ para todo $n \in \mathbb{N}$). Con esto,

$$h(n) = (f * u)(n) = \sum_{d|n} f(d) \qquad \text{y} \qquad G(x) = \lfloor x \rfloor.$$ □

Particularizando aún más, veamos cómo queda la identidad anterior si las funciones que tomamos son la de Möbius y la de von Mangoldt.

Teorema 7.27.
Para $x \geq 1$, se cumple

$$\sum_{n\leq x} \mu(n) \left\lfloor \frac{x}{n} \right\rfloor = 1$$

y

$$\sum_{n\leq x} \Lambda(n) \left\lfloor \frac{x}{n} \right\rfloor = \log(\lfloor x \rfloor !).$$

Demostración. Aplicando el corolario 7.26 con $f = \mu$, y el teorema 7.1, obtenemos

$$\sum_{n\leq x} \mu(n) \left\lfloor \frac{x}{n} \right\rfloor = \sum_{n\leq x} \sum_{d|n} \mu(d) = \sum_{n\leq x} \delta(n) = 1.$$

De modo similar, a partir del corolario 7.26 con $f = \Lambda$, y el apartado (a) de la proposición 7.9, llegamos a

$$\sum_{n\leq x} \Lambda(n) \left\lfloor \frac{x}{n} \right\rfloor = \sum_{n\leq x} \sum_{d|n} \Lambda(d) = \sum_{n\leq x} \log(n) = \log(\lfloor x \rfloor !). \qquad \square$$

Concluyamos la sección con una versión generalizada del teorema 7.24. Utilizaremos este resultado en el tema siguiente, para estudiar las sumas parciales de ciertos productos de Dirichlet (concretamente, se aplica en la demostración del teorema 8.32).

Teorema 7.28.
Sean f, g funciones aritméticas, $h = f * g$ su producto de Dirichlet y, para $x \geq 1$, tomemos

$$F(x) = \sum_{n\leq x} f(n), \quad G(x) = \sum_{n\leq x} g(n) \quad y \quad H(x) = \sum_{n\leq x} h(n).$$

Entonces, para cualesquiera dos números reales $r, s \in [1, \infty)$ tales que $rs = x$, se cumple

$$\sum_{cd\leq x} f(d)g(c) = \sum_{n\leq s} f(n)G\left(\frac{x}{n}\right) + \sum_{n\leq r} g(n)F\left(\frac{x}{n}\right) - F(s)G(r).$$

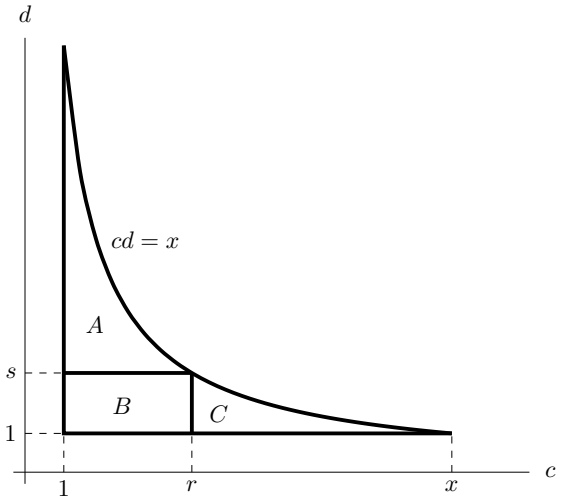


Figura 7.1: Región del plano cd comprendida entre la hipérbola $cd = x$ y las rectas $c = 1$ y $d = 1$; en el dibujo aparece descompuesta en tres partes, denotadas como A, B y C.

Demostración. Partamos de tres números reales $x, r, s \geq 1$ fijos, y con $rs = x$. Tal como indica la figura 7.1, tomemos —en el plano cd— la región limitada por la hipérbola $cd = x$ y las rectas $c = 1$ y $d = 1$. Asimismo, descompongamos esa región hiperbólica en las tres partes que en la figura denotamos como A, B y C.

La suma $\sum_{cd\leq x} f(d)g(c)$ está extendida a los puntos de coordenadas enteras de la región hiperbólica. Por otra parte, cambiando n por d es claro que

$$\sum_{n\leq s} f(n)G\left(\frac{x}{n}\right) = \sum_{d\leq s} f(d)G\left(\frac{x}{d}\right) = \sum_{d\leq s} f(d) \sum_{c\leq x/d} g(c)$$

es $\sum f(d)g(c)$ extendida a los puntos de coordenadas enteras de $B \cup C$. Análogamente ocurre con $\sum_{n\leq r} g(n)F(x/n) = \sum_{c\leq r} g(c) \sum_{d\leq x/c} f(d)$ y los puntos de $A \cup B$ (esta vez, n hace el papel de c). Los puntos de B intervienen dos veces; así pues, su contribución debemos descontarla restando la expresión

$$F(s)G(r) = \sum_{d\leq s,\, c\leq r} f(d)g(c),$$

lo que prueba el teorema. □

7.7. Más tipos de convoluciones y sus inversiones de Möbius

A lo largo de las secciones anteriores de este tema ya hemos visto dos fórmulas de inversión de Möbius, la segunda más general que la primera.

En la básica, y tras haber definido el producto de Dirichlet de funciones aritméticas mediante

$$(\alpha * \beta)(n) = \sum_{d|n} \alpha(d)\beta(n/d) = \sum_{d|n} \alpha(n/d)\beta(d),$$

en el teorema 7.6 establecíamos que, si f y g son dos funciones aritméticas, se tiene

$$g(n) = \sum_{d|n} f(d) = \sum_{d|n} f\left(\frac{n}{d}\right) \;\Leftrightarrow\; f(n) = \sum_{d|n} \mu\left(\frac{n}{d}\right) g(d) = \sum_{d|n} \mu(d) g\left(\frac{n}{d}\right);$$

esto es, $g = f * u$ si y sólo si $f = g * \mu$ (por supuesto, u es la función unidad dada por $u(n) = 1$ para todo n y μ es la función de Möbius).

Para la más sofisticada, definíamos previamente la convolución de la función aritmética α y $F\colon [1,\infty) \to \mathbb{C}$ como una nueva función $\alpha \circ F\colon [1,\infty) \to \mathbb{C}$ dada por

$$(\alpha \circ F)(x) = \sum_{n \leq x} \alpha(n) F\left(\frac{x}{n}\right)$$

(recuerde el lector que si ponemos $F(n) = f(n)$ para $n \in \mathbb{N}$ y $F(x) = 0$ para $x \notin \mathbb{N}$, entonces $(\alpha \circ F)(n) = (\alpha * f)(n)$ para $n \in \mathbb{N}$ y $(\alpha \circ F)(x) = 0$ para $x \notin \mathbb{N}$, y en este sentido podemos decir que estamos generalizando). Luego, en la proposición 7.21 probábamos la propiedad asociativa mixta, que afirma que, para α y β funciones aritméticas y $F\colon [1,\infty) \to \mathbb{C}$, se cumple $\alpha \circ (\beta \circ F) = (\alpha * \beta) \circ F$. Y concluíamos con el teorema 7.23 en el que probábamos que, si α es una función aritmética completamente multiplicativa y $F, G\colon [1,\infty) \to \mathbb{C}$, entonces

$$G(x) = \sum_{n \leq x} \alpha(n) F\left(\frac{x}{n}\right) \;\Leftrightarrow\; F(x) = \sum_{n \leq x} \mu(n)\alpha(n) G\left(\frac{x}{n}\right),$$

es decir, $G = \alpha \circ F$ si y sólo si $F = (\mu \cdot \alpha) \circ G$, lo cual se denominaba fórmula de inversión de Möbius generalizada. En realidad, en la proposición 7.22 veíamos que la inversión también se podía realizar si α no era completamente multiplicativa pero tenía inversa, con la única diferencia de que en lugar de $\mu \cdot \alpha$ había que usar α^{-1}, pero de momento no nos vamos a preocupar por ello.

Antes de seguir, vamos a introducir dos cambios. En primer lugar, dejaremos de denotar las funciones de variable real con letras mayúsculas F y G, y volveremos a usar f y g. En segundo lugar, obsérvese que, si f es una función $f\colon (0,\infty)\to\mathbb{C}$ que cumple $f(x)=0$ para $0<x<1$, la convolución de α (función aritmética) y f es la misma si definimos

$$(\alpha\circ f)(x)=\sum_{n=1}^{\infty}\alpha(n)f\left(\frac{x}{n}\right) \tag{7.21}$$

(como $f(x/n)=0$ cuando $n>x$, sólo se añaden sumandos nulos al tomar $\sum_{n=1}^{\infty}$ en vez de $\sum_{n\leq x}$). Por descontado, este pequeño cambio se propaga a la fórmula de inversión de Möbius generalizada. Así, si α es una función aritmética completamente multiplicativa y $f\colon (0,\infty)\to\mathbb{C}$ cumple $f(x)=0$ para $0<x<1$, entonces

$$g(x)=\sum_{n=1}^{\infty}\alpha(n)f\left(\frac{x}{n}\right)$$

también cumple $g(x)=0$ para $0<x<1$ y, además,

$$f(x)=\sum_{n=1}^{\infty}\mu(n)\alpha(n)g\left(\frac{x}{n}\right);$$

y recíprocamente.

Más aún, incluso sin necesidad de que haya un número finito de sumandos no nulos (que es para lo que estábamos usando que f y g se anulaban en el intervalo $(0,1)$), si se logra justificar adecuadamente que las series convergen, y que los cambios de orden en los sumatorios de la propiedad asociativa mixta se pueden realizar, se tendrá, en general, que

$$g(x)=\sum_{n=1}^{\infty}\alpha(n)f\left(\frac{x}{n}\right)\ \Leftrightarrow\ f(x)=\sum_{n=1}^{\infty}\mu(n)\alpha(n)g\left(\frac{x}{n}\right). \tag{7.22}$$

Ni siquiera hace falta que el dominio de f y g sean los reales positivos: puede ser cualquier conjunto que garantice que $f(x/n)$ y $g(x/n)$ siempre tengan sentido. Convencerse de todo esto es sencillo: basta repasar mentalmente las demostraciones de las proposiciones 7.21 (la propiedad asociativa mixta) y 7.22, y el teorema 7.23.

Pero en la literatura matemática aparecen otras fórmulas de inversión de Möbius en las cuales las funciones f y g no se evalúan en x/n, sino en otras combinaciones de x y n. De hecho, la historia de los distintos tipos de fórmulas de inversión de

Möbius se remonta a un artículo de Möbius, publicado en 1832, en el que analiza el problema de invertir una serie de Taylor

$$g(x) = a_1 x + a_2 x^2 + a_3 x^3 + a_4 x^4 + \cdots$$

con el sentido de poder expresar

$$x = b_1 g(x) + b_2 g(x^2) + b_3 g(x^3) + b_4 g(x^4) + \cdots .$$

Así, Möbius encuentra una fórmula de inversión que, con nuestra notación actual y para α una función aritmética completamente multiplicativa, sería

$$g(x) = \sum_{n=1}^{\infty} \alpha(n) f(x^n) \;\Leftrightarrow\; f(x) = \sum_{n=1}^{\infty} \mu(n)\alpha(n) g(x^n). \tag{7.23}$$

Por supuesto, para hablar con rigor de esta inversión, deberíamos preocuparnos de la convergencia de las series, pero de momento no lo vamos a hacer. (En particular, aunque quizás resulte sorprendente, la «primera» fórmula de inversión no usaba $f(x/n)$ sino $f(x^n)$.)

Un ejemplo debido al propio Möbius es el siguiente (en el ejercicio 7.41 se le propone al lector que analice otro):

Ejemplo. El desarrollo en serie de Taylor de la función arco tangente es

$$\operatorname{arc\,tg}(x) = \sum_{m=0}^{\infty} \frac{(-1)^m}{2m+1} x^{2m+1}, \quad -1 < x \leq 1.$$

Es una simple rutina comprobar que la función aritmética α dada por $\alpha(n) = 0$ si n es par y $\alpha(n) = (-1)^m/(2m+1)$ si $n = 2m+1$ es completamente multiplicativa. Entonces, $\alpha^{-1} = \mu\alpha$ y la fórmula de inversión nos da

$$x = \sum_{m=0}^{\infty} \frac{(-1)^m \mu(2m+1)}{2m+1} \operatorname{arc\,tg}(x^{2m+1}).$$

Evaluando en $x = 1$ (para ser rigurosos, se necesitaría cierto análisis adicional sobre la convergencia, pero lo omitimos), y usando $\operatorname{arc\,tg}(1) = \pi/4$, se sigue que

$$\frac{4}{\pi} = \sum_{m=0}^{\infty} \frac{(-1)^m \mu(2m+1)}{2m+1},$$

una bonita serie relacionada con π. ◊

Sin conocer el trabajo de Möbius, Pafnuty Chebyshev encuentra, en 1851, una fórmula de inversión que, de nuevo con nuestra notación y para α función aritmética completamente multiplicativa, equivaldría a

$$g(x) = \sum_{n=1}^{\infty} \alpha(n) f(nx) \;\Leftrightarrow\; f(x) = \sum_{n=1}^{\infty} \mu(n)\alpha(n) g(nx). \tag{7.24}$$

En particular, Chebyshev lo aplica al cálculo de inversas de series de Fourier de una onda triangular y otra cuadrada. En general, esto sería así (lo planteamos con senos pero se podría hacer igual con series de cosenos):

Ejemplo. Supongamos que, para una función $g : [0, 1) \to \mathbb{R}$, extendida a $\mathbb{R}$ por periodicidad (es decir, $g(x + k) = g(x)$ para todo $k \in \mathbb{N}$), se tiene el desarrollo de Fourier

$$g(x) = \sum_{n=1}^{\infty} \alpha(n) \operatorname{sen}(2\pi n x). \tag{7.25}$$

Entonces, si usamos $\operatorname{sen}_{2\pi}$ para denotar la función $\operatorname{sen}_{2\pi}(x) = \operatorname{sen}(2\pi x)$ podemos escribir (7.25) como

$$g(x) = \sum_{n=1}^{\infty} \alpha(n) \operatorname{sen}_{2\pi}(nx).$$

Si α es completamente multiplicativa y no tenemos en cuenta los problemas de convergencia, la fórmula de inversión (con $f = \operatorname{sen}_{2\pi}$) nos daría

$$\operatorname{sen}(2\pi x) = \sum_{n=1}^{\infty} \mu(n)\alpha(n) g(nx). \tag{7.26}$$

Tomando valores particulares de x para los cuales $\operatorname{sen}(2\pi x)$ tenga una expresión en forma de radicales, esta fórmula da lugar a preciosas series. ◊

Habitualmente, los coeficientes de una serie de Fourier no dan lugar a una función α completamente multiplicativa. Sin exigir eso, sino sólo que $\alpha(1) \neq 0$ para que α tenga inversa, la expresión (7.26) sigue siendo válida sin más que usar $\alpha^{-1}(n)$ en lugar de $\mu(n)\alpha(n)$; pero, claro, no suele resultar fácil encontrar una expresión explícita que proporcione $\alpha^{-1}(n)$ para todo n (para calcular $\alpha^{-1}(n)$ siempre se puede usar la fórmula recursiva (7.9) que vimos en el teorema 7.5, pero no es un método muy eficiente).

Un bonito ejemplo en el que los coeficientes de las series de Fourier son funciones completamente multiplicativas es el siguiente:

Ejemplo. Aunque hay maneras alternativas de hacerlo, los polinomios de Bernoulli $B_n(x)$ se definen habitualmente por medio de la función generatriz

$$\frac{te^{tx}}{e^t-1} = \sum_{n=0}^{\infty} B_n(x)\,\frac{t^n}{k!}.$$

Los dos primeros son $B_0(x)=1$ y $B_1(x)=x-1/2$; en general, el polinomio $B_n(x)$ es mónico y tiene grado n.[12] En 1890, Adolf Hurwitz encontró el desarrollo de Fourier

$$B_{2k}(x) = \frac{2(-1)^{k-1}(2k)!}{(2\pi)^{2k}} \sum_{n=1}^{\infty} \frac{\cos(2\pi nx)}{n^{2k}}, \quad x\in[0,1), \quad k\geq 1;$$

$$B_{2k+1}(x) = \frac{2(-1)^{k-1}(2k+1)!}{(2\pi)^{2k+1}} \sum_{n=1}^{\infty} \frac{\operatorname{sen}(2\pi nx)}{n^{2k+1}}, \quad x\in[0,1), \quad k\geq 0.$$

Analicemos sólo el caso de los pares (el otro sería similar). El desarrollo se puede extender periódicamente a todo $\mathbb{R}$ sin más que tomar $B_{2k}(x-\lfloor x\rfloor)$, y así tenemos

$$B_{2k}(x-\lfloor x\rfloor) = \frac{2(-1)^{k-1}(2k)!}{(2\pi)^{2k}} \sum_{n=1}^{\infty} \frac{\cos(2\pi nx)}{n^{2k}}, \quad x\in\mathbb{R}, \qquad k\geq 1.$$

Entonces, dado que $\alpha(n)=n^{-2k}$ es una función completamente multiplicativa, la fórmula de inversión (es decir, (7.26) pero con cosenos) nos da

$$\cos(2\pi x) = \frac{(2\pi)^{2k}}{2(-1)^{k-1}(2k)!} \sum_{n=1}^{\infty} \frac{\mu(n)}{n^{2k}}\, B_{2k}(x-\lfloor x\rfloor), \quad x\in\mathbb{R}.$$

Se pueden encontrar muchos más detalles en [72]. ◊

[12] La utilidad de los polinomios de Bernoulli y sus propiedades va mucho más allá de lo que aquí citamos. Se puede encontrar mucha más información en numerosos textos de teoría de números; por ejemplo, [5, § 12.11 y siguientes] o [47, capítulos 7 y 9]. Ahora sólo comentaremos que los denominados números de Bernoulli son $B_n = B_n(0)$ y, tal como demostró Euler en 1740, proporcionan el valor exacto de la función $\zeta(s)$ para los enteros positivos pares:

$$\zeta(2k) = \frac{(-1)^{k-1}(2\pi)^{2k}}{2(2k)!}\, B_{2k}, \qquad k\in\mathbb{N};$$

en la sección 6.9 de este libro presentamos una sencilla demostración de esta fórmula (que, además, usa los polinomios de Bernoulli). Aparecerán de nuevo en la sección 8.2, en relación con la fórmula de sumación de Euler-Maclaurin.

Otro tipo de fórmula de inversión en la que aparece la función de Möbius es

$$g(x)=\sum_{n=1}^{\infty}\alpha(n)f(x^{1/n}) \;\Leftrightarrow\; f(x)=\sum_{n=1}^{\infty}\mu(n)\alpha(n)g(x^{1/n}). \tag{7.27}$$

Sin duda su uso más famoso es cuando Riemann, en 1859, empleó

$$\mathrm{li}(x)=\sum_{n=1}^{\infty}\frac{1}{n}\,\mathrm{Ri}(x^{1/n}) \;\Leftrightarrow\; \mathrm{Ri}(x)=\sum_{n=1}^{\infty}\frac{\mu(n)}{n}\,\mathrm{li}(x^{1/n})$$

en su estudio de la distribución de primos (no hace falta que nos preocupemos ahora de lo que son las funciones li y Ri; lo veremos en el tema 10, cuando abordemos el teorema de los números primos, pero sí que merece la pena señalar que la importancia de esa inversión radica en que tanto $\mathrm{li}(x)$ como $\mathrm{Ri}(x)$ aproximan muy bien la función $\pi(x)$ que cuenta el número de primos menores o iguales que x).

¿Cómo se puede probar toda esta variedad de fórmulas de inversión de Möbius? En realidad, si suponemos conocida una de estas fórmulas de inversión (7.22), (7.23), (7.24) o (7.27), no es difícil deducir las demás (siempre que nos encontremos en los dominios adecuados). Basta que lo veamos en un caso:

Ejemplo. Supuesta conocida la fórmula de inversión de Möbius (7.22) para funciones definidas en $(0,\infty)$, deducir (7.27) para funciones definidas en $(1,\infty)$. (O, de manera alternativa, en los intervalos $(-\infty,\infty)$ y $(0,\infty)$, respectivamente.)

Solución. Vamos a ver únicamente la implicación «$\Rightarrow$», dado que el razonamiento con «$\Leftarrow$» sería exactamente igual. Partimos de

$$g(x)=\sum_{n=1}^{\infty}\alpha(n)f(x^{1/n}).$$

Si tomamos $y=\log(x)$ (que es una biyección entre los intervalos involucrados) y definimos $\widetilde{f}(y)=f(e^y)=f(x)$, $\widetilde{g}(y)=g(e^y)=g(x)$, la expresión anterior queda

$$\widetilde{g}(y)=\sum_{n=1}^{\infty}\alpha(n)f((e^y)^{1/n})=\sum_{n=1}^{\infty}\alpha(n)f(e^{y/n})=\sum_{n=1}^{\infty}\alpha(n)\widetilde{f}(y/n),$$

y su inversión de Möbius según (7.22) es

$$\widetilde{f}(y)=\sum_{n=1}^{\infty}\mu(n)\alpha(n)\widetilde{g}(y/n)=\sum_{n=1}^{\infty}\mu(n)\alpha(n)g(e^{y/n})=\sum_{n=1}^{\infty}\mu(n)\alpha(n)g((e^y)^{1/n}).$$

Deshaciendo los cambios, hemos llegado a

$$f(x) = \sum_{n=1}^{\infty} \mu(n)\alpha(n)g(x^{1/n}),$$

tal como buscábamos. ◊

Por otra parte, se pueden dar demostraciones directas de las fórmulas de inversión (7.23), (7.24) o (7.27) usando técnicas muy similares a las que vimos para (7.22) en la sección anterior de este tema. Incluso, de manera análoga a (7.21), se pueden definir sendas convoluciones $(\alpha \circ f)(x)$ mediante

$$\sum_{n=1}^{\infty} \alpha(n)f(x^n), \qquad \sum_{n=1}^{\infty} \alpha(n)f(nx) \qquad \text{o} \qquad \sum_{n=1}^{\infty} \alpha(n)f(x^{1/n}),$$

que luego dan lugar a las correspondientes propiedades asociativas mixtas. Si lo pensamos un poco, esto es más general (y quizás menos problemático) que lo que se ha hecho en el ejemplo anterior al relacionar las inversiones (7.22) y (7.27): allí usábamos $y = \log(x)$, y al final deshacíamos el cambio con $x = e^y$, y lo mismo se podría haber hecho con cualquier biyección $y = h(x)$, pero para eso necesitamos que las funciones sean «realmente inversas» en los dominios en los que se mueven las variables modificadas por n (por ejemplo, no es claro que se pueda aplicar al caso $y = \cos(x)$, $x = \arc\cos(y)$, dado que $\cos(\arc\cos(y)) = y$ pero no es cierto en general que $\arc\cos(\cos(x)) = x$: incluso aunque nos restrinjamos a $x \in [0, \pi)$, nos saldremos del intervalo al tomar nx).

7.7.1. Fórmulas de inversión de Möbius con flujos

En 1885, el matemático italiano Ernesto Cesàro se dio cuenta de que había una mecánica subyacente que se podía sistematizar, y que servía no sólo para $f(x/n)$, $f(x^n)$, $f(nx)$ o $f(x^{1/n})$, sino siempre que usásemos $f(\varphi(n, x))$ con una función φ que cumpliese dos propiedades que en seguida explicamos. Así pues, vamos a ver en qué consiste este esquema general, que en particular proporciona las cuatro fórmulas de inversión (7.22)–(7.27) como simples casos particulares. Hay muchos otros ejemplos que encajan dentro de esta estrategia, pero aquí no analizaremos casi ninguno más; si el lector se siente interesado, puede consultar [16, 17].

Para no estar forzados a emplear funciones definidas sobre $(0, \infty)$ (o algún otro intervalo predeterminado), vamos a usar Δ para denotar un subconjunto $\Delta \subset \mathbb{R}$. Llamamos *flujo* de $\mathbb{N}$ sobre Δ a una función $\varphi\colon \mathbb{N} \times \Delta \to \Delta$ que verifique

(a) $\varphi(1,x)=x$ para cada $x\in\Delta$,
(b) $\varphi(n,\varphi(m,x))=\varphi(nm,x)$ para cada $x\in\Delta$, $n,m\in\mathbb{N}$.

Entonces, para α función aritmética, φ un flujo, y $f:\Delta\to\mathbb{C}$, definimos la φ-convolución de α y f como la nueva función

$$(\alpha\odot_\varphi f)(x)=\sum_{n=1}^{\infty}\alpha(n)f(\varphi(n,x)),$$

supuesto, claro está, que la serie converja.

Si, como es habitual, δ denota la función aritmética dada por $\delta(1)=1$ y $\delta(n)=0$ para $n>1$, la propiedad (a) implica que $\delta\odot_\varphi f=f$ para cualquier función f; en efecto: para cada $x\in\Delta$, se cumple

$$(\delta\odot_\varphi f)(x)=\sum_{n=1}^{\infty}\delta(n)f(\varphi(n,x))=1\cdot f(x)+0=f(x).$$

Con la φ-convolución, la asociatividad mixta, que se deduce de la propiedad (b), es como sigue (nótese que $\alpha*\beta$ es el producto de Dirichlet ordinario, no una variante relacionada con φ):

Proposición 7.29.
Sean α,β dos funciones aritméticas, φ un flujo sobre Δ y $f:\Delta\to\mathbb{C}$. Entonces,

$$\alpha\odot_\varphi(\beta\odot_\varphi f)=(\alpha*\beta)\odot_\varphi f,$$

que es válido en cada $x\in\Delta$ tal que

$$\sum_{n,m\in\mathbb{N}}\left|\alpha(n)\beta(m)f(\varphi(nm,x))\right|<\infty. \tag{7.28}$$

Demostración. Aplicando la propiedad (b) del flujo φ, se tiene

$$\begin{aligned}(\alpha\odot_\varphi(\beta\odot_\varphi f))(x)&=\sum_{n=1}^{\infty}\alpha(n)(\beta\odot_\varphi f)(\varphi(n,x))\\&=\sum_{n=1}^{\infty}\alpha(n)\sum_{n=1}^{\infty}\beta(m)f(\varphi(m,\varphi(n,x)))\\&=\sum_{n,m\in\mathbb{N}}\alpha(n)\beta(m)f(\varphi(nm,x))\end{aligned}$$

$$= \sum_{k\in\mathbb{N}} \Big(\sum_{nm=k} \alpha(n)\beta(m) \Big) f(\varphi(k,x))$$
$$= \sum_{k\in\mathbb{N}} (\alpha * \beta)(k) f(\varphi(k,x)) = ((\alpha * \beta) \odot_\varphi f)(x),$$

donde el intercambio en los sumatorios está justificado por la convergencia absoluta (7.28). □

Nota 7.30. Obsérvese que la hipótesis (7.28) también se puede escribir como

$$\sum_{k=1}^{\infty} \Big(\sum_{nm=k} |\alpha(n)\beta(m)| \Big) |f(\varphi(k,x))| < \infty;$$

o, de manera más corta, $\sum_{k=1}^{\infty} (|\alpha| * |\beta|)(k) |f(\varphi(k,x))| < \infty$. ◁

En este contexto, el teorema de inversión para φ-convoluciones es el siguiente (en realidad, los resultados del teorema se refieren a cada $x \in \Delta$ en el que se cumplan las correspondientes hipótesis, aunque no nos molestaremos en detallarlo):

Teorema 7.31.
Sea α una función aritmética con inversa α^{-1}, un flujo φ, y una función $f : \Delta \to \mathbb{C}$. Si las series

$$\sum_{n=1}^{\infty} |\alpha(n) f(\varphi(n,x))| \quad y \quad \sum_{n,m\in\mathbb{N}} |\alpha(n)\alpha^{-1}(m) f(\varphi(nm,x))| \tag{7.29}$$

son convergentes y definimos

$$g(x) = (\alpha \odot_\varphi f)(x), \tag{7.30}$$

entonces se cumple

$$f(x) = (\alpha^{-1} \odot_\varphi g)(x). \tag{7.31}$$

Recíprocamente, si tenemos $g : \Delta \to \mathbb{C}$ tal que las series

$$\sum_{n=1}^{\infty} |\alpha(n) g(\varphi(n,x))| \quad y \quad \sum_{n,m\in\mathbb{N}} |\alpha(n)\alpha^{-1}(m) g(\varphi(nm,x))| \tag{7.32}$$

son convergentes y definimos f como en (7.31), entonces se cumple (7.30).



Peter Gustav Lejeune Dirichlet (1805–1859)



Ernesto Cesàro (1859–1908)

Demostración. Dadas α y f, la convergencia de la serie de la izquierda en (7.29) implica que la función $g = \alpha \odot_\varphi f$ existe. Ahora, por ser convergente la serie de la derecha en (7.29), la proposición 7.29 nos asegura que

$$\alpha^{-1} \odot_\varphi g = \alpha^{-1} \odot_\varphi (\alpha \odot_\varphi f) = (\alpha^{-1} * \alpha) \odot_\varphi f = \delta \odot_\varphi f = f.$$

La situación recíproca es similar. □

Antes de seguir, merece la pena que nos detengamos un poco en analizar la hipótesis

$$\sum_{n,m\in\mathbb{N}} \left|\alpha(n)\alpha^{-1}(m)f(\varphi(nm,x))\right| < \infty$$

—es la que permite el uso de la asociatividad mixta— en el caso más interesante, que es cuando α es una función aritmética completamente multiplicativa, con lo cual su inversa es $\alpha^{-1} = \mu\alpha$. En primer lugar, se podrían agrupar los sumandos que cumplen $nm = k$ y dejarla de manera análoga a como se ha hecho en la nota 7.30. Y, en segundo lugar, y más importante, podemos acotar $|\mu(m)| \leq 1$ y escribir

$$|\alpha(n)\alpha^{-1}(m)| = |\alpha(n)\mu(m)\alpha(m)| \leq |\alpha(n)\alpha(m)| = |\alpha(nm)|,$$

con lo cual

$$\sum_{nm=k} |\alpha(n)\alpha^{-1}(m)| \leq \sum_{nm=k} |\alpha(nm)| = \tau(k)|\alpha(k)|,$$

donde $\tau(k)$ es el número de divisores de k. De este modo, la hipótesis que pide la acotación de la serie de la derecha en (7.29) se puede sustituir por

$$\sum_{k=1}^{\infty} \tau(k)\left|\alpha(k)f(\varphi(k,x))\right| < \infty,$$

Pero, además, dado que $\tau(k) \geq 1$, esta serie mayora a la de la izquierda en (7.29). Como consecuencia, podemos sustituir la convergencia de las dos series que aparecen en (7.29) por la única condición

$$\sum_{k=1}^{\infty} \tau(k) \big|\alpha(k) f(\varphi(k,x))\big| < \infty, \tag{7.33}$$

que es más cómoda de comprobar. Por supuesto, lo mismo se podría hacer con la convergencia de las series (7.32), que se sustituiría por

$$\sum_{k=1}^{\infty} \tau(k) \big|\alpha(k) g(\varphi(k,x))\big| < \infty. \tag{7.34}$$

En (7.33) y (7.34) resulta muy útil poder utilizar una cota para el tamaño de la función τ. Aunque no lo vamos a demostrar ahora, lo que se cumple es que, sea cual sea el $\delta > 0$ que tomemos,

$$\tau(k) \leq C_\delta k^\delta \quad \text{para todo } k \in \mathbb{N} \tag{7.35}$$

para alguna constante C_δ (sin embargo, $\tau(k)$ no está acotada por ninguna potencia de $\log(k+1)$); en este libro, nosotros lo veremos en el teorema 8.2 del siguiente tema.

En todo caso, cuando α es una función aritmética completamente multiplicativa, el teorema anterior —prescindiendo de la convergencia de las series— se podría resumir mediante la siguiente fórmula de inversión de Möbius:

$$g(x) = \sum_{n=1}^{\infty} \alpha(n) f(\varphi(n,x)) \;\Leftrightarrow\; f(x) = \sum_{n=1}^{\infty} \mu(n)\alpha(n) g(\varphi(n,x)). \tag{7.36}$$

Sobraría ya decir que de aquí se deducen las cuatro fórmulas de inversión (7.22)–(7.27) sin más que tomar, respectivamente, $\varphi(n,x) = x/n$, $\varphi(n,x) = x^n$, $\varphi(n,x) = nx$ o $\varphi(n,x) = x^{1/n}$; es inmediato que estas cuatro funciones responden a la definición de flujo.

El mismo Cesàro mencionó otros ejemplos de flujo ya no tan obvios. En particular, que, cuando $\beta(n)$ es cualquier función completamente multiplicativa, son flujos las funciones $\varphi(n,x) = \beta(n)x$ o $\varphi(n,x) = x^{\beta(n)}$; y también

$$\varphi(n,x) = \frac{x}{1 - x\log(\beta(n))}, \qquad \varphi(n,x) = \frac{x\beta(n)}{1 - kx(\beta(n)-1)}$$

(con k una constante). Comprobarlo es un mero cálculo carente de dificultad.

Pero, ¿qué decir de la la convergencia de las dos series que hay en (7·36), y de las que permiten el uso de la asociatividad mixta? (recordemos que es de la asociatividad mixta de donde se deduce que esas series no sólo convergen, sino que son inversas). Ya hemos visto las condiciones (7·33) y (7·34) que lo garantizan, pero en ocasiones puede ser mucho más sencillo. Y es que, afortunadamente, a veces no hay que preocuparse de la convergencia: eso ocurre siempre que las series sean sumas con un número finito de sumandos no nulos. En los cuatro casos típicos, es claro que esto sucede en las siguientes circunstancias:

- Para $\varphi(n,x) = x/n$ con $\Delta = (0,\infty)$, cuando tomamos funciones que se anulan en un intervalo $(0,a)$ (pues $x/n < a$ cuando n es suficientemente grande).
- Para $\varphi(n,x) = x^n$ y $\Delta = (1,\infty)$, con funciones que se anulan a partir de un $K \in \mathbb{R}$ (pues $x^n > K$ cuando n suficientemente grande).
- Para $\varphi(n,x) = nx$ y $\Delta = (0,\infty)$, con funciones que se anulan a partir de un $K \in \mathbb{R}$.
- Para $\varphi(n,x) = x^{1/n}$ y $\Delta = (1,\infty)$, con funciones que se anulan en $(1,a)$ para algún $a > 1$.

No obstante, no son éstos los únicos casos que se pueden estudiar. En el ejercicio 7.42 planteamos una situación que no responde a ninguna de estas premisas: el flujo no es ninguno de los anteriores, las series involucradas tienen infinitos sumandos no nulos, pero se puede dar una demostración rigurosa de la correspondiente fórmula de inversión de Möbius de tipo (7·36).

7.8. Ejercicios

7.1. Demuestra la fórmula para la suma de una progresión geométrica:

$$t + tr + tr^2 + tr^3 + \cdots + tr^k = \frac{tr^{k+1} - t}{r - 1},$$

que es válida para todo $t \in \mathbb{C}$ y cualquier $r \in \mathbb{C} \setminus \{1\}$. ¿Cuánto vale la suma si $r = 1$? ¿Qué ocurre cuando $k \to \infty$?

7.2. Prueba que

$$\sum_{d|n} |\mu(d)| = 2^{\omega(n)},$$

donde $\omega(n)$ es el número de divisores primos distintos que tiene n.

7.3. De manera análoga a como se hace en el teorema 7.2, encuentra una fórmula explícita para $\sigma_a(n)$.

7.4. Supón que $2^k - 1$ es un número primo. Prueba que k debe ser primo. ¿Es cierto el recíproco?
Sugerencia. Si has estudiado el apartado 3.4.3 de este libro, puedes ignorar este ejercicio.

7.5. Demuestra que, si m es un número perfecto o abundante, y m divide propiamente a n, entonces n es abundante.

7.6. Sea un primo p y un entero positivo k. Demuestra que p^k es defectuoso.

7.7. Dado un entero $k > 1$, supón que los números $u = 3 \cdot 2^k - 1$, $v = 3 \cdot 2^{k-1} - 1$ y $w = 9 \cdot 2^{2k-1} - 1$ son primos. Tal como descubrió Thabit ibn Qurra en el siglo IX, prueba que, entonces, $n = 2^k uv$ y $m = 2^k w$ forman un par de números amigos. Aplícalo para encontrar las siguientes parejas: $(220, 284)$, $(17296, 18416)$ y $(9363584, 9437056)$.

7.8. Sea k un entero positivo tal que $p = 3^{k+1} - 4$ es primo, y toma $n = 3^k p - 1$. Prueba que, si $n/2$ también es primo, entonces $\sigma(n) = \sigma(n+1)$. Encuentra tres ejemplos concretos.

7.9. Para cada $r \in \mathbb{N}$, denotemos μ^r y μ^{-r} a las funciones aritméticas

$$\mu^r(n) = (\overbrace{\mu * \mu * \cdots * \mu}^{r \text{ veces}})(n), \qquad n \in \mathbb{N},$$

$$\mu^{-r}(n) = (\overbrace{\mu^{-1} * \mu^{-1} * \cdots * \mu^{-1}}^{r \text{ veces}})(n) = (\overbrace{u * u * \cdots * u}^{r \text{ veces}})(n), \qquad n \in \mathbb{N};$$

tomemos, además, $\mu^0 = \delta$. Demuestra que $\mu^r \neq \mu^s$ si r y s son dos enteros distintos. (Con lenguaje algebraico, estaríamos probando que las potencias —respecto al producto de Dirichlet— de la función de Möbius forman un grupo cíclico infinito; por supuesto, de igual forma lo podríamos haber expresado en términos de las potencias de $u = \mu^{-1}$.)

7.10. Para cada $d \mid n$, definamos $A_d = \{\, k \in \mathbb{N} : 1 \leq k \leq n,\ \operatorname{mcd}(k, n) = d \,\}$. Comprueba que los conjuntos A_d son disjuntos y que su unión es $\cup_{d|n} A_d = \{1, 2, \ldots, n\}$.

7.11. Demuestra que $\varphi * \tau = \sigma$.

7.12. Sean f y g funciones multiplicativas. ¿Podemos garantizar que $f + g$ y $f \cdot g$ (suma y producto punto a punto) son multiplicativas? ¿Y si f y g son completamente multiplicativas? Pruébalo o busca contraejemplos.

7.13. Comprueba que las funciones μ, φ, τ y σ_a no son completamente multiplicativas.

7.14. Prueba las siguientes propiedades de la indicatriz de Euler:
(a) $\varphi(nm) = \varphi(n)\varphi(m)\frac{d}{\varphi(d)}$, donde $d = \mathrm{mcd}(n, m)$.
(b) Si $a \mid b$, entonces $\varphi(a) \mid \varphi(b)$.
(c) $\varphi(n)$ es par para $n \geq 3$. Además, si n posee r factores primos impares distintos, entonces $2^r \mid \varphi(n)$.

7.15. Demuestra que

$$\mu(n)\log(n) = -\sum_{d|n} \mu(d)\Lambda\left(\frac{n}{d}\right).$$

7.16. Supón que pretendemos probar un teorema análogo al teorema 7.10, pero con funciones completamente multiplicativas (tanto en las hipótesis como en las tesis). ¿Serían ciertos los correspondientes resultados?

7.17. Caracteriza las funciones aritméticas tales que ella y su inversa son completamente multiplicativas.

7.18. Sean f y g funciones aritméticas no nulas. Prueba que $f * g$ tampoco puede ser la función nula. (En otras palabras, que el producto de Dirichlet no tiene divisores de cero.)

7.19. Dada una función aritmética, su soporte (al que aludimos con «sop») es, por definición, el conjunto de los enteros positivos sobre los que la función no se anula. Además de la función δ y sus múltiplos, ¿existe alguna función aritmética f que sea inversible y tal que $\mathrm{sop}(f)$ y $\mathrm{sop}(f^{-1})$ sean conjuntos finitos?

7.20. Sea f una función aritmética no nula. ¿Es posible que tanto el soporte de f como de $u * f$ sean finitos?

7.21. Sea f una función multiplicativa tal que $f(2) \neq 0$. Prueba que la función $g(n) = \frac{1}{f(2)} f(2n)$, $n \in \mathbb{N}$, también es multiplicativa.

7.22. Generaliza el ejercicio anterior probando que si f es una función multiplicativa y c es un entero positivo fijo con $f(c) \neq 0$, entonces la función

$$g(n) = \frac{1}{f(c)} f(cn), \qquad n \in \mathbb{N},$$

también es multiplicativa.

7.23. Sea f una función multiplicativa. Prueba que

$$f(n)f(m) = f(\mathrm{mcd}(n, m))f(\mathrm{mcm}(n, m)), \qquad n, m \in \mathbb{N}.$$

Recíprocamente, si una función aritmética satisface la igualdad anterior, ¿podemos garantizar que es multiplicativa? Si la respuesta es negativa, ¿podemos lograrlo imponiendo alguna hipótesis adicional?

7.24. Profundicemos un poco más en lo que estudiábamos en el ejercicio anterior; y es que la igualdad que allí aparecía sirve para definir una nueva clase de funciones: una función aritmética f que verifica

$$f(n)f(m) = f(\mathrm{mcd}(n,m))f(\mathrm{mcm}(n,m)), \qquad n,m \in \mathbb{N},$$

se denomina *semimultiplicativa*. Demuestra que una función aritmética f no nula es semimultiplicativa si y sólo si existe una constante no nula c_f, un entero positivo a_f, y una función multiplicativa $\widetilde{f}$ tales que

$$f(n) = c_f \widetilde{f}(n/a_f)$$

para todo $n \in \mathbb{N}$ (aquí estamos interpretando que $\widetilde{f}(x) = 0$ cuando $x \notin \mathbb{N}$).

7.25. Sea f una función multiplicativa que se anula sobre los números no libres de cuadrados. Prueba que su inversa es

$$f^{-1}(n) = \begin{cases} 1, & \text{si } n = 1, \\ \lambda(n) \prod_{j=1}^{k} f(p_j)^{a_j}, & \text{si } n = p_1^{a_1} \cdots p_k^{a_k}. \end{cases}$$

7.26. Sea un primo p y un número real r. Asimismo, sea f la función aritmética definida como $f(1) = 1$, $f(p) = r$ y $f(n) = 0$ en otro caso. Halla f^{-1}.

7.27. Dado $c \in \mathbb{R} \setminus \{0\}$, sea f_c la siguiente función aritmética:

$$f_c(n) = \begin{cases} 0, & \text{si } n \text{ par}, \\ (-1)^m c/(2m+1)^2, & \text{si } n = 2m+1. \end{cases}$$

Encuentra su inversa.
Sugerencia. Piensa primero el caso $c = 1$.

7.28. Sea f la función aritmética definida mediante $f(n) = (-1)^{n+1}$. Prueba que su función inversa es

$$g(n) = \begin{cases} \mu(n), & \text{si } n \text{ impar}, \\ 2^{a-1}\mu(m), & \text{si } n \text{ es par con } n = 2^a m,\ a \geq 1, \text{ y } m \text{ impar}. \end{cases}$$

7.29. Sea f una función multiplicativa. Prueba que f es completamente multiplicativa si y sólo si $f(g * h) = (fg) * (fh)$ para cualesquiera funciones $g, h \in \mathscr{F}$.

7.30. Sea f una función aritmética que no siempre vale 1 sobre los primos. Demuestra que la fórmula

$$\prod_{p|n} (1 - f(p))$$

define una función multiplicativa. ¿Podemos asegurar que esa función sea completamente multiplicativa? ¿Puede serlo alguna vez?

7.31. Prueba lo que se afirma en la nota 7.16 de la página 423.

7.32. Demuestra que

$$\frac{n}{\varphi(n)} = \sum_{d|n} \frac{\mu(d)^2}{\varphi(d)}.$$

7.33. Sean f y g dos funciones crecientes y completamente aditivas. Prueba que, en esas condiciones,

$$f(n)g(2) = f(2)g(n) \quad \text{para todo } n \in \mathbb{N}.$$

7.34. Sea f una función completamente aditiva que verifica $f(n+1) \geq f(n)$ para todo $n \in \mathbb{N}$. Demuestra que, entonces, existe una constante real $C \geq 0$ tal que $f(n) = C \log(n)$ para todo $n \in \mathbb{N}$.

7.35. Sea $f\colon \mathbb{N} \to \mathbb{N} \cup \{0\}$ una función creciente y completamente aditiva. Prueba que f debe ser la función nula.

7.36. Busca un ejemplo de una función completamente aditiva que tome valores en $\mathbb{N} \cup \{0\}$ y que no sea creciente.

7.37. Siguiendo con el ejercicio anterior, busca otro ejemplo; naturalmente, la función que encuentres ahora no deberá ser un múltiplo constante de la que hallaste allí.

7.38. Sea $\psi(x)$ la función de Chebyshev

$$\psi(x) = \sum_{n \leq x} \Lambda(n), \text{ y } H(x) = \sum_{n \leq x} \mu(n) \log(n).$$

Demuestra que, para $x \geq 1$,

$$H(x) = -\sum_{n \leq x} \mu(n) \psi\left(\frac{x}{n}\right).$$

Sugerencia. Observa la fórmula del ejercicio 7.15.

7.39. Prueba que, cuando $x \geq 1$, se cumple

$$\sum_{n \leq x} \lambda(n) \left\lfloor \frac{x}{n} \right\rfloor = \lfloor \sqrt{x} \rfloor .$$

7.40. En los preámbulos del teorema 7.28 hemos afirmado que tal resultado es una generalización del teorema 7.24. ¿Por qué es esto así?

7.41. Sea la serie de Taylor

$$-\log(1-x) = \sum_{n=1}^{\infty} \frac{x^n}{n}, \qquad |x| < 1.$$

Aplica una fórmula de inversión de Möbius para deducir, a partir de ella (y sin pararte a justificar las hipótesis que garantizan la validez de la inversión), que

$$e^x = \prod_{n=1}^{\infty} (1-x^n)^{-\mu(n)/n}, \qquad |x| < 1.$$

7.42. Los polinomios de Chebyshev $\{T_n(x)\}_{n=0}^{\infty}$ cumplen $T_n(x) = \cos(n \arccos(x))$ cuando $x \in [-1,1]$,[13] y de ahí se deduce fácilmente que $T_{nm}(x) = T_n(T_m(x))$ para todo $n, m \in \mathbb{N}$ y todo $x \in \mathbb{R}$ (¿por qué?). Asumido lo anterior y la cota (7.35), analiza la fórmula de inversión

$$g(x) = \sum_{n=1}^{\infty} \frac{1}{n^s} f(T_n(x)) \;\Leftrightarrow\; f(x) = \sum_{n=1}^{\infty} \frac{\mu(n)}{n^s} g(T_n(x))$$

para $s > 1$ y funciones acotadas definidas sobre $\mathbb{R}$ (o sobre $[-1,1]$).

[13] Los detalles los vimos en el tema 6 de este libro; en concreto, véase el apartado 6.1.1.

Tema 8
Medias de funciones aritméticas

8.1. El orden de magnitud de algunas funciones aritméticas
8.2. Fórmulas de sumación
8.3. Estimaciones asintóticas de funciones
8.4. Orden medio de las funciones divisor
8.5. Orden medio de la indicatriz de Euler
8.6. Estimaciones elementales relacionadas con la función de Möbius
8.7. Resultados equivalentes al teorema de los números primos
8.8. Equivalencia con las medias de la función de Möbius
8.9. Ejercicios

Antes de comenzar, aclaremos que este tema tiene como requisito previo el tema anterior; al menos, que el lector conozca la notación (asumimos las mismas convenciones que allí explicamos), y que esté familiarizado con los conceptos y las funciones aritméticas en él definidas. El tipo de resultados que aquí mostraremos puede encontrarse en numerosos libros, pues éste es un objetivo clásico de la teoría analítica de números; como referencias, citemos los mismos textos que en el tema anterior, [3, 5, 12, 50, 71, 87].

Cuando nos encontramos ante una sucesión $\{a_n\}_{n=1}^{\infty}$, lo primero que nos puede interesar de ella es si tiene límite, ya sea finito o infinito. Y, si es infinito, qué orden de infinitud tiene; por ejemplo, si existe algún $\alpha > 0$ tal que el límite de a_n/n^α es una constante no nula. Lo mismo podemos preguntarnos cuando el límite de $\{a_n\}_{n=1}^{\infty}$ es cero; aunque, esta vez, con un $\alpha < 0$. En cualquiera de estos casos, y suponiendo que $a_n/n^\alpha \to \ell \neq 0$, más información se obtiene si se proporciona una buena cota del error que se comete al aproximar a_n/n^α por ℓ, es decir, de $|a_n/n^\alpha - \ell|$.

Obviamente, la respuesta a estas preguntas puede ser negativa: no existe ningún $\alpha \in \mathbb{R}$ tal que $a_n/n^\alpha \to \ell \neq 0$ (por supuesto, en vez de n^α, podríamos haber usado alguna otra función de las habituales para comparar, como $n^\alpha \log^\beta(n)$). En este caso, quizás podemos obtener alguna información más débil: ¿cómo se comportan las medias $(a_1 + a_2 + \cdots + a_n)/n$? (que a menudo se denominan medias de Cesàro). A este respecto, es bien conocido el siguiente criterio de convergencia de sucesiones:

$$a_n \to \ell \;\Rightarrow\; \frac{a_1 + a_2 + \cdots + a_n}{n} \to \ell \tag{8.1}$$

(pero el recíproco no es cierto); véanse los ejercicios 8.1 y 8.2. Así, es posible que no exista el límite de la izquierda pero sí el de la derecha; en este sentido decimos que obtener este último sería encontrar «información más débil» sobre la sucesión $\{a_n\}_{n=1}^{\infty}$. En vez de $(a_1 + a_2 + \cdots + a_n)/n$, también podemos buscar límites (finitos y no nulos) de expresiones del tipo $(a_1 + a_2 + \cdots + a_n)/n^\alpha$. De nuevo nos preguntamos por cotas del error para $|(a_1 + a_2 + \cdots + a_n)/n^\alpha - \ell|$. En realidad, no hay ninguna necesidad de usar el n^α dividiendo, pues lo mismo nos da acotar $|(a_1 + a_2 + \cdots + a_n) - \ell n^\alpha|$.

Merece la pena comentar que, aunque sin darnos cuenta de ello, en nuestra vida diaria manejamos habitualmente dos conceptos uno de los cuales se define a partir de las medias del otro: el clima de una región es un promedio del tiempo atmosférico a lo largo de los años. Salvo en los pronósticos a corto plazo que dan los servicios de meteorología, es imposible predecir el tiempo atmosférico que habrá un día concreto; en cambio, eso no tiene importancia en el clima, que se nutre de los datos históricos.

En este tema, pretendemos obtener este tipo de información para las funciones aritméticas. Es decir, cuando la sucesión es $\{f(n)\}_{n=1}^{\infty}$ con f una función $f\colon \mathbb{N} \to \mathbb{R}$. En particular, para las *funciones aritméticas notables* que definíamos en el tema anterior (sección 7.1). No hay más que echar un vistazo a las definiciones de las funciones μ (función de Möbius), φ (indicatriz de Euler), τ (número de divisores), σ (suma de los divisores), λ (función de Liouville) y Λ (función de von Mangoldt) para darse cuenta de que presentan grandes oscilaciones, luego no pueden tener límite. Así, perdida la posibilidad de encontrar el tamaño de las funciones, cobra interés el estudio de sus medias. Además, la información que obtengamos sobre estas medias de funciones aritméticas será ya suficiente para responder a algunas preguntas de indudable atractivo, como ¿cuál es la probabilidad de que dos números tomados al azar sean primos entre sí?

Entre las funciones aritméticas antes mencionadas no hemos incluido la función $\pi(n)$, que cuenta el número de primos menores que n y que resulta fundamental en teoría de números. De hecho, el estudio del orden de magnitud de $\pi(n)$ es el que más atención despierta, y parte del interés de los resultados que se obtienen sobre las otras funciones radica en que se usan para conseguir información sobre $\pi(n)$. La razón de este aparente olvido es que abordar este estudio requiere herramientas más avanzadas que las que aquí se van a usar, y nos ocuparemos de ello en otro tema.

Nos aparecerán muchas veces expresiones del tipo $\sum_{n\leq x} a(n)$, con $x \geq 1$ un número real; recordemos que su significado es, simplemente, $\sum_{n=1}^{\lfloor x \rfloor} a(n)$ (para $x < 1$, si se usa alguna vez, convenimos que $\sum_{n\leq x} a(n) = 0$). Con esta idea, definiremos las medias de funciones aritméticas como funciones de variable real; en lugar de definirlas únicamente sobre los naturales. Que sean funciones de variable continua en vez de discreta tiene ventajas, pues esto permite aplicar las herramientas del análisis matemático.

Concluyamos esta introducción describiendo —y comentando— los denominados símbolos de Landau: «O grande» y «o pequeña» (también llamados «O mayúscula» y «o minúscula»), que nos aparecerán muy a menudo. Realmente, fue el alemán Paul Bachmann quien comenzó a usarlos en 1894, y se popularizaron en los años siguientes gracias a su compatriota Edmund Laundau y a algunos otros matemáticos que les siguieron. Actualmente es una notación muy extendida, pero siempre resulta extraña a quien se topa con ella por primera vez. Es muy útil para comparar, asintóticamente, el tamaño de dos funciones, mostrando sólo la información conocida o la relevante en ese momento, prescindiendo de lo que no interesa. Aquí sólo la usaremos cuando $x \to \infty$, pero se puede definir con límites hacia cualquier otro valor (esto es muy común, por ejemplo, en desarrollos de Taylor). Se pueden encontrar explicaciones más detalladas de todo esto en [47, capítulo 9].

Sean $f, g\colon [a,\infty) \to \mathbb{R}$. Entonces:

(a) Se escribe $f \sim g$ (cuando $x \to \infty$) si $\lim_{x\to\infty} \frac{f(x)}{g(x)} = 1$.

(b) Se escribe $f = o(g)$ (cuando $x \to \infty$) si $\lim_{x\to\infty} \frac{f(x)}{g(x)} = 0$.

(c) Se escribe $f = O(g)$ (cuando $x \to \infty$) si existen $M > 0$ y $b \geq a$ tales que $|f(x)| \leq Mg(x)$ para todo $x \geq b$.

A partir de ahora, la coletilla «cuando $x \to \infty$» se dará por sobreentendida, y no siempre lo indicaremos explícitamente. Muchas veces aparecerán expresiones del tipo $f(x) = h(x) + o(g(x))$ (análogamente con O); esto significa que $f(x) - h(x) = o(g(x))$.

Comentemos —con la ayuda de un ejemplo— la utilidad de este tipo de notaciones (y, posiblemente, su porqué). Cuando se escribe $\sqrt{x^6+x^4+2} = x^3 + o(x^2)$ estamos diciendo, de manera abreviada, que $\sqrt{x^6+x^4+2} = x^3 + f(x)$ para una cierta función f (ya sea desconocida, o que no es necesario explicitar pues no tiene importancia) que verifica $f(x)/x^2 \to 0$ cuando $x \to \infty$. La regla nemotécnica consiste en cambiar el «0» por una letra «o» (que se lee «o pequeña»). Por otra parte, obsérvese que la notación $\sqrt{x^6+x^4+2} = x^3 + o(x^2)$ está dando más información que si sólo escribiéramos $\sqrt{x^6+x^4+2} \sim x^3$.

Es importante darse cuenta de que, si $f = o(g)$ (es decir, si $\lim_{x\to\infty} f(x)/g(x) = 0$), deberá ser $|f(x)| < g(x)$ a partir de un b. En particular, se cumplirá $f = O(g)$; pero una $f = O(g)$ no tiene por qué ser $f = o(g)$. Olvidándonos del módulo, del b y de la constante multiplicativa, en ambos casos $f = o(g)$ y $f = O(g)$, la función f es más pequeña que la g; pero este «más pequeña» es mucho más acentuado en el caso «o» que en el «O». Así pues, la notación «O grande» se está usando para referirse a algo que, habitualmente (y cuando la variable x esté cerca de su límite), será mayor que una «o pequeña».

Conviene observar que, si $f \sim g$, y ambas son funciones no negativas (al menos, a partir de cierto valor de la variable), se cumplirá $f = O(g)$ y $g = O(f)$. Es más, si existe

$$\lim_{x\to\infty} f(x)/g(x) = \ell \in \mathbb{R} \setminus \{0\}, \tag{8.2}$$

entonces se tiene $f = O(g)$ y $g = O(f)$ (¡naturalmente!: (8.2) es lo mismo que afirmar $f \sim \ell g$). Pero de $f = O(g)$ y $g = O(f)$ no se puede asegurar (8.2), sino sólo la existencia de dos constantes positivas K_1 y K_2 tales que $K_1 g(x) \leq f(x) \leq K_2 g(x)$ para todo x suficientemente grande (aunque nosotros no lo usaremos, esto es habitual denotarlo como $f(x) \asymp g(x)$). Por otra parte, (8.2) es equivalente a decir



Paul Gustav Heinrich Bachmann (1837–1920)



Edmund Landau (1877–1938)



Paul David Gustav du Bois-Reymond (1831–1889)

$f(x) - \ell g(x) = o(g(x))$. Pero, claro, más información proporcionaría una estimación del tipo $f(x) - \ell g(x) = o(h(x))$ (o incluso $O(h(x))$), siempre que $h(x)$ sea lo suficientemente pequeña. Este tipo de aproximaciones será el que más interesados estamos en encontrar.

En el contexto de las O, una forma habitual de expresar que una función f es acotada es escribir $f = O(1)$. De manera análoga, para decir que $\lim_{x\to\infty} f(x) = 0$ podemos poner $f = o(1)$.

Finalmente, comentemos que una notación alternativa para $f = O(g)$ es $f(x) \ll g(x)$ (aunque no es tan versátil, pues no sirve para escribir algo similar a $f(x) = h(x) + O(g(x))$). Y $f = o(g)$ se expresa $f(x) \prec g(x)$; por ejemplo, podemos decir que, cuando $x \to \infty$, $x^\alpha \prec x^\beta$ si y sólo si $\alpha < \beta$. En este libro no emplearemos la notación «$\prec$» (que fue introducida por el alemán Paul du Bois-Reymond en 1871), salvo ahora mismo, que nos resulta muy cómoda para describir la jerarquía de los órdenes de infinitud. Seguro que el lector ya la conoce, pero no está de más recordar que

$$1 \prec \log(\log(x)) \prec \log(x) \prec x^\varepsilon \prec x \prec x^a \prec x^{\log(x)} \prec b^x \prec x^x \prec c^{d^x},$$

donde ε, a, b, c, d son constantes arbitrarias que verifican $0 < \varepsilon < 1$ y $a, b, c, d > 1$. De manera similar, las siguientes funciones (excepto la constante igual a 1) tienden hacia 0:

$$\frac{1}{c^{d^x}} \prec \frac{1}{x^x} \prec \frac{1}{b^x} \prec \frac{1}{x^{\log(x)}} \prec \frac{1}{x^a} \prec \frac{1}{x} \prec \frac{1}{x^\varepsilon} \prec \frac{1}{\log(x)} \prec \frac{1}{\log(\log(x))} \prec 1;$$

en otras palabras, son infinitésimos de distinto orden.

8.1. El orden de magnitud de algunas funciones aritméticas

Antes de abordar la parte esencial de este tema, vamos a estudiar el tamaño, cuando n es grande, de algunas de las funciones aritméticas más importantes; en concreto, las funciones $\tau(n)$, $\sigma(n)$ y $\varphi(n)$. El procedimiento requerido va a ser relativamente sencillo, pero los resultados obtenidos serán necesariamente limitados. Y es que, en general, estas funciones tienen oscilaciones importantes (con la salvedad de que $\varphi(n)$ es creciente).

No obstante, intentaremos comparar su orden de magnitud con expresiones de tipo $n^\alpha \log^\beta(n)$, y probaremos diversas cotas. También mostraremos algunos resultados en los que aparecen límites superiores o inferiores, pero que —por supuesto— no son ciertos para el límite en sí. No podía ser de otra manera, pero queremos recalcar que los distintos tipos de estimaciones que vamos a obtener para las funciones aritméticas son mucho más pobres —en el sentido de que se ajustan peor— que los que luego probaremos para las medias de esas mismas funciones. Con las funciones, sólo conseguiremos cotas; para las medias, verdaderas aproximaciones asintóticas.

Históricamente, los resultados más finos conocidos sobre el tamaño de las funciones aritméticas son posteriores a los de los correspondientes sobre sus medias; y, los métodos empleados, más complicados. Pero, conceptualmente, parece más natural analizar las funciones en sí antes que sus medias, y así vamos a hacerlo. En realidad, algunos resultados sólo los enunciaremos, pues probarlos requiere más herramientas de las que ahora tenemos.

Esperamos que, al menos, esto sirva al lector para convencerle de que abordar el estudio de las medias de funciones aritméticas no es algo esotérico sino que, ante la imposibilidad de conseguir estimaciones precisas sobre las funciones, resulta ser una buena idea.

8.1.1. El orden de la función $\tau(n)$

Comencemos observando que $\tau(n) \geq 2$ para todo $n > 1$ y que $\tau(n) = 2$ siempre que n es un primo. En consecuencia, el límite inferior de $\tau(n)$ es

$$\liminf_n \tau(n) = 2. \tag{8.3}$$

Encontrar buenas cotas superiores para el orden de magnitud de $\tau(n)$ no es tan directo. A este respecto, vamos a ver dos resultados, uno negativo y otro positivo:

(a) Cuando $n \to \infty$, el orden de magnitud de $\tau(n)$ no es menor que ninguna potencia de $\log(n)$.

(b) Sin embargo, el orden de $\tau(n)$ sí que es menor que cualquier potencia positiva de n.

En particular, de (a) se deduce que $\limsup_n \tau(n) = \infty$. Junto con (8.3), esto muestra la imposibilidad de conseguir una estimación asintótica «sencilla»[1] para $\tau(n)$.

Antes de proseguir, veamos un lema que nos va a probar que, si una función multiplicativa tiende a 0 a través de la subsucesión constituida por las potencias de números primos, entonces esa función tiende a 0 (a este respecto, compárese el lema con el ejercicio 8.5, que tiene un enunciado parecido).

Lema 8.1.

Sea f una función multiplicativa. Si $f(p^m) \to 0$ cuando $p^m \to \infty$, entonces $f(n) \to 0$ cuando $n \to \infty$.

Demostración. Por hipótesis, es claro que, dado cualquier $\varepsilon > 0$, se cumple

$$|f(p^m)| < A \quad \text{para todo } p \text{ y } m, \tag{8.4}$$
$$|f(p^m)| < 1 \quad \text{si } p^m > B, \tag{8.5}$$
$$|f(p^m)| < \varepsilon \quad \text{si } p^m > N(\varepsilon), \tag{8.6}$$

donde A y B son constantes independientes de p, m y ε, y $N(\varepsilon)$ depende sólo de ε. Cada entero $n > 1$ tendrá una descomposición en factores primos $n = p_1^{a_1} \cdots p_k^{a_k}$ y, como f es multiplicativa, podemos poner $f(n) = f(p_1^{a_1} \cdots p_k^{a_k}) = f(p_1^{a_1}) \cdots f(p_k^{a_k})$. Para analizar estos factores, introduzcamos un poco de notación. Sean

$$S = \{\, p^m \in \mathbb{N} : p \text{ primo}, m \geq 1, p^m \leq B \,\},$$
$$T(\varepsilon) = \{\, n = p_1^{a_1} \cdots p_k^{a_k} \in \mathbb{N} : p_j^{a_j} \leq N(\varepsilon) \text{ para todo } j \,\}.$$

Ambos conjuntos tienen cardinal finito (y el de $T(\varepsilon)$ depende sólo de ε). Por ser $T(\varepsilon)$ un conjunto finito, existe su máximo

$$P(\varepsilon) = \text{máx}\{n : n \in T(\varepsilon)\}.$$

Así, para demostrar que $f(n) \to 0$, nos basta con preocuparnos por los $n > P(\varepsilon)$. Además, sin pérdida de generalidad, podemos suponer que $\varepsilon < 1$.

[1] Hay que ser cuidadoso cuando se emplean este tipo de argumentos de «no existencia»: puede parecer una obviedad —y desde luego lo es—, pero una función siempre se aproxima de manera muy precisa por otra: ella misma. Es verdad que esto no sirve para nada, pero en matemáticas hay que ser preciso con lo que se afirma.

Con esto, dado $n = p_1^{a_1} \cdots p_k^{a_k} > P(\varepsilon)$, distingamos entre sus potencias de primos en S y fuera de S, y, usando (8.4), escribamos

$$|f(n)| = \prod_{p_j^{a_j} \in S} |f(p_j^{a_j})| \cdot \prod_{p_j^{a_j} \notin S} |f(p_j^{a_j})| \leq A^{\mathrm{card}(S)} \cdot \prod_{p_j^{a_j} \notin S} |f(p_j^{a_j})|. \tag{8.7}$$

Podríamos seguir acotando ya que, por (8.4), cada uno de los factores que queda cumple $|f(p_j^{a_j})| < 1$. Pero eso no es bastante para lo que pretendemos. En su lugar, fijémonos en que, al ser $n > P(\varepsilon)$, forzosamente tiene al menos un factor de la forma p^a con $p^a > N(\varepsilon)$ y que, por (8.6), $f(p^a) < \varepsilon$. Ahora, empalmando con (8.7) obtenemos $f(n) < A^{\mathrm{card}(S)}\varepsilon$. Esto se puede hacer para cada $\varepsilon > 0$, lo cual prueba que $f(n) \to 0$. □

Una vez probado el lema, vayamos con el teorema. En realidad, el lema sólo se usa para demostrar (8.9), y se podía haber dado una prueba directa —a partir de la fórmula $\tau(p_1^{a_1} \cdots p_k^{a_k}) = (1+a_1) \cdots (1+a_k)$—, pero usar el lema permite simplificar la demostración. Es más, no sólo ésta, sino también las del siguiente apartado.

Teorema 8.2.
La afirmación

$$\tau(n) = O(\log^{\nu}(n)) \quad \textit{cuando } n \to \infty \tag{8.8}$$

es falsa para cada número real ν, pero sí que

$$\tau(n) = O(n^{\delta}) \quad \textit{cuando } n \to \infty \tag{8.9}$$

para cada $\delta > 0$.

Demostración. Comencemos con (8.8). Por supuesto, basta verlo para los ν grandes, ya que, si (8.8) no es cierto con los grandes, con más razón fallará con los pequeños. Observemos que, si n tiene la forma $n = 2^m$ (siendo m cualquier entero positivo), se cumplirá $\tau(n) = m+1 \sim \log_2(n) = \log(n)/\log(2)$. Si $n = (2 \cdot 3)^m$, será $\tau(n) = (m+1)^2 \sim (\log(n)/\log(6))^2$; y así sucesivamente. Entonces, dado $\nu \geq 1$, y si $\{p_j\}_{j=1}^{\infty}$ denota la sucesión de primos ordenados de manera creciente, sea k el entero positivo que hace $k \leq \nu < k+1$, y tomemos $n = (2 \cdot 3 \cdots p_{k+1})^m$. De este modo,

$$\tau(n) = (m+1)^{k+1} \sim \Big(\frac{\log(n)}{\log(2 \cdot 3 \cdots p_{k+1})}\Big)^{k+1} > K(\log(n))^{k+1},$$

donde K es una constante independiente de m. En consecuencia, la estimación (8.8) no puede ser cierta.

A continuación, veamos (8.9), y para ello vamos a apoyarnos en el lema con $f(n) = \tau(n)/n^\delta$. Esta f es multiplicativa (por serlo τ) y, además, en los enteros de la forma p^m verifica

$$f(p^m) = \frac{m+1}{p^{m\delta}} \leq \frac{2m}{p^{m\delta}} = \frac{2}{p^{m\delta}} \frac{\log(p^m)}{\log(p)} \leq \frac{2}{\log(2)} \frac{\log(p^m)}{(p^m)^\delta}.$$

Como eso tiende a 0 cuando $p^m \to \infty$, el lema nos asegura que $f(n) \to 0$ cuando $n \to \infty$. Y, de aquí, $\tau(n) = o(n^\delta)$, de donde se sigue (8.9).[2] □

Existen estudios más precisos sobre el «tamaño superior» de $\tau(n)$. A este respecto, mencionemos que, en 1907, el matemático sueco Severin Wigert demostró que

$$\limsup_n \frac{\log(\tau(n)) \log(\log(n))}{\log(n)} = \log(2). \tag{8.10}$$

Este resultado es difícil, y no lo probaremos en este tema, sino en el tema 10 (sección 10.5). La demostración original de Wigert utilizaba el teorema de los números primos y, años más tarde, el genial matemático hindú Srinivasa Ramanujan logró dar una prueba más elemental, como la que se puede encontrar en [50, § 18.1, teorema 317]. Aunque aún no estamos en condiciones de abordar la demostración allí presentada, sí que podríamos hacerlo más adelante en este mismo libro, cuando hayamos ampliado nuestros conocimientos sobre la distribución de números primos (nos referimos, claro, a antes de llegar al tema 10 en el que demostraremos el teorema de los números primos y, como consecuencia, también (8.10)). La parte más delicada de la demostración que aparece en [50] consiste en usar una estimación del tipo $\vartheta(x) > Ax$, donde $\vartheta(x)$ es la función de Chebyshev, la misma función que nosotros definiremos en el apartado 8.7.1 (en cambio, una vez que se tiene el teorema de los números primos, esa estimación es inmediata).

8.1.2. La funciones $\sigma(n)$ y $\varphi(n)$

Las irregularidades en el comportamiento de la funciones $\sigma(n)$ y $\varphi(n)$ son menos pronunciadas que las de $\tau(n)$. Tal como veremos, el tamaño de $\sigma(n)$ y $\varphi(n)$ es,

[2] Hemos probado (8.9) con $o(n^\delta)$ en lugar de con $O(n^\delta)$ así que, en principio, puede parecer que hemos obtenido un resultado más fuerte del que figura en el enunciado. Pero, como tenemos un «para cada $\delta > 0$», démonos cuenta de que $n^{\delta'} = o(n^\delta)$ cuando $0 < \delta' < \delta$. Así, es equivalente afirmar $\tau(n) = O(n^\delta)$ para todo $\delta > 0$ o $\tau(n) = o(n^\delta)$ para $\delta > 0$.

en cierto sentido, casi n. En realidad, el orden de ambas funciones está bastante relacionado ya que, para cierta constante positiva A, se verifica

$$A < \frac{\sigma(n)\varphi(n)}{n^2} \leq 1, \quad n \in \mathbb{N}. \tag{8.11}$$

De aquí se sigue que $\sigma(n)/n$ y $n/\varphi(n)$ tienen el mismo orden de magnitud. Esto se podía haber usado para deducir estimaciones sobre una de las funciones a partir de las de la otra. Pero no vamos a seguir este camino pues, aunque no es difícil probar (8.11), para hacerlo con rigor y comodidad, es mejor que lo dejemos hasta el tema 9; en concreto, véase el ejercicio 9.2.

Comencemos estudiando la función $\sigma(n)$. Para cada $n > 1$, se cumple

$$\sigma(n) > n,$$

ya que $1 \mid n$ y $n \mid n$. Cuando n es primo, $\sigma(n) = n + 1$, así que

$$\liminf_n \frac{\sigma(n)}{n} = 1. \tag{8.12}$$

Respecto a su «tamaño superior», tenemos:

Teorema 8.3.
Cuando $n \to \infty$, la función $\sigma(n)$ verifica

$$\sigma(n) = O(n^{1+\delta})$$

para cada $\delta > 0$.

Demostración. Apliquemos el lema 8.1 con $f(n) = \sigma(n)/n^{1+\delta}$, que es una función multiplicativa (por serlo σ). Como $\sigma(p^m) = 1+p+p^2+\cdots+p^m = (p^{m+1}-1)/(p-1)$, sin más que utilizar $2p^m \leq p^{m+1}+1$ (luego $p^{m+1}-1 = 2p^{m+1}-p^{m+1}-1 \leq 2p^{m+1}-2p^m$), tenemos

$$f(p^m) = \frac{p^{m+1}-1}{p^{(1+\delta)m}(p-1)} \leq \frac{2(p^{m+1}-p^m)}{p^{(1+\delta)m}(p-1)} = \frac{2p^m}{p^{(1+\delta)m}} = 2\big(p^m\big)^{-\delta},$$

que tiende a 0 cuando $p^m \to \infty$, tal como buscábamos. □

Con $\varphi(n)$ vamos a obtener, con dificultad similar, resultados del mismo calado (aunque, en cierto sentido, simétricos a los de $\sigma(n)$). Es evidente que, para cada $n > 1$, la función $\varphi(n)$ cumple

$$\varphi(n) < n.$$

Por otra parte, dado cualquier $\varepsilon > 0$, tomemos primos $p > 1/\varepsilon$, y potencias del tipo $n = p^m$. Entonces,

$$\varphi(n) = p^m - p^{m-1} = n\left(1 - \frac{1}{p}\right) > n(1-\varepsilon),$$

luego

$$\limsup_n \frac{\varphi(n)}{n} = 1. \tag{8.13}$$

Además, se tiene:

Teorema 8.4.
La función $\varphi(n)$ verifica

$$\lim_n \frac{\varphi(n)}{n^{1-\delta}} = \infty$$

para cada $\delta > 0$.

Demostración. Tomemos la función multiplicativa $f(n) = n^{1-\delta}/\varphi(n)$. Por el lema 8.1, basta probar que $f(p^m) \to 0$ cuando $p^m \to \infty$. Pero

$$\frac{1}{f(p^m)} = \frac{\varphi(p^m)}{p^{(1-\delta)m}} = p^{m\delta}\left(1 - \frac{1}{p}\right) \geq \frac{p^{m\delta}}{2},$$

que tiende a ∞ cuando $p^m \to \infty$, y en consecuencia $f(p^m) \to 0$. □

Hasta ahora, hemos visto resultados de demostración bastante sencilla. Como ocurría con $\tau(n)$, los más finos requieren herramientas que aún no hemos visto. Para $\sigma(n)$, Thomas Hakon Grönwall —otro matemático sueco— demostró, en 1913, que su «tamaño superior» es

$$\limsup_n \frac{\sigma(n)}{n\log(\log(n))} = e^\gamma, \tag{8.14}$$

donde γ es la constante de Euler (la definiremos en el apartado (a) de la proposición 8.10). El correspondiente teorema para el «tamaño inferior» de $\varphi(n)$ ya había sido probado por Landau en 1903; en concreto, él estableció que

$$\liminf_n \frac{\varphi(n)\log(\log(n))}{n} = e^{-\gamma}. \tag{8.15}$$



Thomas Hakon Grönwall (1877–1932)



Niels Henrik Abel (1802–1829)



Thomas Jan Stieltjes (1856–1894)

Al igual que con (8.10), veremos la demostración de este resultado no ahora, sino en el tema 10 (sección 10.5), cuando hayamos probado el teorema de los números primos. Existen, así mismo, demostraciones que no usan dicho teorema, como las de [50]. Allí, (8.14) es [50, § 18.3, teorema 323]; y (8.15) es [50, § 18.4, teorema 328]. La demostración de ambos está en [50, § 22.9]. En ella se usa (8.11), una cota del tipo $\vartheta(x) \leq Ax$, y algunos resultados sobre distribución de primos como los que nosotros veremos —en el tema 9— al estudiar las fórmulas de Mertens; en concreto, la constante e^{γ} de (8.14) y (8.15) surge de

$$\prod_{p \leq x} \left(1 - \frac{1}{p}\right) \sim \frac{e^{-\gamma}}{\log(x)}$$

(véase el teorema 9.20 y los comentarios que le siguen).

Nota 8.5. De (8.14) y (8.15) se sigue, respectivamente,

$$\limsup_n \frac{\sigma(n)}{n} = \infty \qquad \text{y} \qquad \liminf_n \frac{\varphi(n)}{n} = 0. \tag{8.16}$$

Compárense estas expresiones con (8.12) y (8.13). Además, de aquí se deduce que la cota del teorema 8.3 no es cierta con $\delta = 0$. Pero probar (8.16) de manera directa es más sencillo que demostrar (8.14) y (8.15). En realidad, basta calcular sólo uno de los dos límites que aparecen en (8.16) ya que, entonces, de (8.11) se deduciría el otro. Y la comprobación de $\limsup_n \sigma(n)/n = \infty$ está propuesta como ejercicio en el tema 9; en concreto, véase el ejercicio 9.5. ◁

8.2. Fórmulas de sumación

En esta sección, mostremos un interesante resultado que nos será de gran utilidad para estimar el tamaño de algunas funciones.

Teorema 8.6. Identidad de Abel.
Sea a una función aritmética, $A(t) = \sum_{n\le t} a(n)$, y $f\colon [y,x] \to \mathbb{C}$ una función de variable real de clase $C^{(1)}$ (es decir, continua y con derivada continua en todo el intervalo), con $0 < y < x < \infty$. Entonces,

$$\sum_{y<n\le x} a(n)f(n) = A(x)f(x) - A(y)f(y) - \int_y^x A(t)f'(t)\,dt.$$

Nota 8.7. Antes de proceder a demostrarla, comentemos que la identidad de Abel es algo así como una versión discreta de la fórmula de integración por partes

$$\int_y^x u\,dv = uv\Big|_y^x - \int_y^x v\,du$$

en la que u es f y, en la parte discreta, $\int$ se sustituye por $\sum$, y dv por a, de donde v sería $\sum a = A$. En realidad, si el lector está familiarizado con la integral de Riemann-Stieltjes, seguro que cae en la cuenta de que, partiendo de

$$\sum_{y<n\le x} a(n)f(n) = \int_y^x f(t)\,dA(t),$$

la integración por partes está plenamente justificada dentro de dicha teoría, lo que da lugar a una demostración directa de la identidad de Abel. Así se hace en algunos textos como [12], pero aquí vamos a dar una prueba alternativa que no usa integrales de Riemann-Stieltjes. ◁

Demostración del teorema 8.6. El caso $\lfloor y\rfloor = \lfloor x\rfloor$ es inmediato, así que suponemos que $\lfloor y\rfloor < \lfloor x\rfloor$ (aunque, en realidad, no hace falta; bastaría con que interpretemos que la suma de un sumatorio sin sumandos es 0).

Descompongamos en trozos teniendo en cuenta que, en los intervalos $[\lfloor y\rfloor, y]$ y $[\lfloor x\rfloor, x]$, y en los de la forma $[j, j+1)$, la función $A(t)$ es constante. Usaremos

también que $A(y) = A(\lfloor y\rfloor)$ y $A(x) = A(\lfloor x\rfloor)$. De este modo,

$$\begin{aligned}
\int_y^x A(t)f'(t)\,dt &= -\int_{\lfloor y\rfloor}^{y} A(t)f'(t)\,dt + \int_{\lfloor x\rfloor}^{x} A(t)f'(t)\,dt + \int_{\lfloor y\rfloor}^{\lfloor x\rfloor} A(t)f'(t)\,dt \\
&= -A(y)\big(f(y) - f(\lfloor y\rfloor)\big) + A(x)\big(f(x) - f(\lfloor x\rfloor)\big) + \sum_{j=\lfloor y\rfloor}^{\lfloor x\rfloor-1} \int_j^{j+1} A(t)f'(t)\,dt \\
&= A(x)f(x) - A(y)f(y) + A(\lfloor y\rfloor)f(\lfloor y\rfloor) - A(\lfloor x\rfloor)f(\lfloor x\rfloor) \\
&\qquad + \sum_{j=\lfloor y\rfloor}^{\lfloor x\rfloor-1} A(j)\big(f(j+1) - f(j)\big) \\
&= A(x)f(x) - A(y)f(y) + \sum_{j=\lfloor y\rfloor}^{\lfloor x\rfloor-1} A(j)f(j+1) - \sum_{j=\lfloor y\rfloor+1}^{\lfloor x\rfloor} A(j)f(j) \\
&= A(x)f(x) - A(y)f(y) + \sum_{n=\lfloor y\rfloor+1}^{\lfloor x\rfloor} A(n-1)f(n) - \sum_{n=\lfloor y\rfloor+1}^{\lfloor x\rfloor} A(n)f(n) \\
&= A(x)f(x) - A(y)f(y) - \sum_{n=\lfloor y\rfloor+1}^{\lfloor x\rfloor} \big(A(n) - A(n-1)\big)f(n) \\
&= A(x)f(x) - A(y)f(y) - \sum_{y<n\le x} a(n)f(n),
\end{aligned}$$

tal como buscábamos. □

Una consecuencia inmediata es la siguiente:

Corolario 8.8. Fórmula de sumación de Euler.

Sea una función $f\colon [y,x] \to \mathbb{C}$ de clase $C^{(1)}$, con x e y cumpliendo $-\infty < y < x < \infty$. Entonces

$$\sum_{y<n\le x} f(n) = (\lfloor x\rfloor - x)f(x) - (\lfloor y\rfloor - y)f(y) + \int_y^x f(t)\,dt + \int_y^x (t - \lfloor t\rfloor)f'(t)\,dt.$$

Demostración. Supongamos en primer lugar que $0 < y < x < \infty$. Tomemos $a(n) = 1$ para todo $n \in \mathbb{N}$, con lo cual $A(t) = \sum_{1\le n\le t} a(n) = \lfloor t\rfloor$. Entonces, la identidad de Abel queda

$$\sum_{y<n\le x} f(n) = \lfloor x\rfloor f(x) - \lfloor y\rfloor f(y) - \int_y^x \lfloor t\rfloor f'(t)\,dt; \tag{8.17}$$

ahora, basta observar que $0 = -xf(x)+yf(y)+\int_y^x f(t)\,dt+\int_y^x tf'(t)\,dt$ y sumárselo a (8.17).

Para el caso general $-\infty < y < x < \infty$, sea N un entero positivo tal que $-N < y$, y definamos $g\colon [y+N, x+N] \to \mathbb{C}$ mediante $g(t) = f(t-N)$. Así, aplicando la fórmula de sumación de Euler para g en el intervalo $[y+N, x+N]$ (que ya hemos probado, pues $0 < y+N < x+N < \infty$), tenemos

$$\begin{aligned}\sum_{y<n\le x} f(n) &= \sum_{y+N<n\le x+N} g(n)\\ &= (\lfloor x+N\rfloor - (x+N))g(x+N) - (\lfloor y-N\rfloor - (y+N))g(y+N)\\ &\qquad + \int_{y+N}^{x+N} g(t)\,dt + \int_{y+N}^{x+N} (t-\lfloor t\rfloor)g'(t)\,dt\\ &= (\lfloor x\rfloor - x)f(x) - (\lfloor y\rfloor - y)f(y)\\ &\qquad + \int_y^x f(t)\,dt + \int_y^x (t-\lfloor t\rfloor)f'(t)\,dt. \qquad \square\end{aligned}$$



Colin Maclaurin
(1698–1746)

Aunque no la vamos a usar en lo que sigue, merece la pena señalar que, a menudo, se denomina fórmula de sumación de Euler no al resultado que acabamos de presentar, sino a otro que también relaciona una suma con una integral, pero de manera más precisa (es decir, con un menor término de error para la expresión que aproxima la diferencia entre la suma y la integral). De hecho, también se alude a este resultado con el nombre de fórmula de sumación de Euler-Maclaurin (la descubrieron Leonhard Euler y Colin Maclaurin, de forma independiente, alrededor de 1735).

Ya que en la sección 6.9 de este libro hemos introducido los números y los polinomios de Bernoulli, donde los usábamos para sumar $\zeta(2k)$, estamos en condiciones de presentar esta fórmula de sumación más general sin mucho esfuerzo. Recordemos que los polinomios de Bernoulli $\{B_k(t)\}_{k=0}^{\infty}$ se definen a partir de la función generatriz

$$\frac{x}{e^x-1}\,e^{xt} = \sum_{k=0}^{\infty} B_k(t)\frac{x^k}{k!}$$

(es decir, del desarrollo de Taylor de $xe^{xt}/(e^x-1)$, como función de x, en torno a $x=0$); y los números de Bernoulli son $B_k = B_k(0)$.

Con esta notación, y dados dos enteros a y b (con $a < b$), una función f suficientemente diferenciable en el intervalo $[a, b]$, y otro entero positivo m, la fórmula de sumación de Euler-Maclaurin establece que

$$\sum_{a\le n<b} f(n) = \int_a^b f(t)\,dt + \sum_{k=1}^{m} \frac{B_k}{k!}\left(f^{(k-1)}(b) - f^{(k-1)}(a)\right) + R_m$$

donde

$$R_m = (-1)^{m+1} \int_a^b \frac{B_m(t - \lfloor t \rfloor)}{m!}\, f^{(m)}(t)\,dt.$$

Dependiendo de cuántos números de Bernoulli usemos (es decir, del m), estamos aproximando la suma y la integral con mayor o menor precisión (recordemos además que $B_k = 0$ para $k \ge 3$ e impar, luego los correspondientes sumandos en $\sum_{k=1}^{m}$ son nulos). Se puede encontrar una demostración de esta fórmula en [47, § 9.5 y 9.6] donde, además, se ve cómo aplicarla para conseguir diversas estimaciones asintóticas.

Merece así mismo la pena comentar que si tomamos $f(t) = t^{m-1}$ con m entero positivo, el resto R_m se anula, pues $f^{(m)}(t)$ es la función nula. Entonces, la fórmula de sumación de Euler-Maclaurin permite encontrar fórmulas exactas para las sumas de tipo $\sum_{a\le n<b} n^{m-1}$. (Realmente, no hace falta recurrir a la sumación de Euler-Maclaurin para sumar ese tipo de expresiones usando polinomios de Bernoulli: en el tema 6, los ejercicios 6.37 y 6.38 proporcionan fórmulas alternativas.)

8.3. Estimaciones asintóticas de funciones

Nuestro objetivo en este tema es analizar el comportamiento asintótico del tamaño de funciones aritméticas. Como paso previo, en esta sección veremos algunas estimaciones asintóticas elementales que nos aparecerán insistentemente. Su demostración se basa, de manera esencial, en la fórmula de sumación de Euler.

Antes de comenzar, conviene aclarar que lo que aquí se expone está planteado con bastante generalidad, pues permitimos que distintos parámetros que aparecen sean números complejos; en concreto, los que denotamos z, s o a. Si al lector le intimidan los complejos, puede optar por suponer que son números reales. Si éste es el caso, donde aparezca $\mathrm{Re}(z)$, $\mathrm{Re}(s)$ o $\mathrm{Re}(a)$, olvídese de ellas y piense que son, simplemente, z, s o a. Para los objetivos del presente tema, no tiene ninguna influencia.

Primero vamos a introducir la función «zeta» de Riemann. La expresión más conocida para esta función es[3]

$$\zeta(z) = \sum_{n=1}^{\infty} \frac{1}{n^z}, \quad \text{válida para } \operatorname{Re}(z) > 1. \tag{8.18}$$

En efecto, cuando $\operatorname{Re}(z) > 1$, esta serie es absolutamente convergente, ya que, si denotamos $r = \operatorname{Re}(z) > 1$, se cumple $|n^z| = n^r$, luego

$$\sum_{n=1}^{\infty} \left| \frac{1}{n^z} \right| = \sum_{n=1}^{\infty} \frac{1}{n^r} < \infty.$$

Pero nos interesa disponer de una expresión más general, válida en un conjunto más amplio del plano complejo. Así, asumiremos que, por definición, la función zeta de Riemann es

$$\zeta(z) = \frac{z}{z-1} - z \int_1^{\infty} \frac{t - \lfloor t \rfloor}{t^{z+1}}\, dt, \quad \text{válida para } \operatorname{Re}(z) > 0,\ z \neq 1. \tag{8.19}$$

Démonos cuenta que esto tiene sentido para cualquier $z \neq 1$ con $\operatorname{Re}(z) > 0$, ya que la integral converge absolutamente: tomando $r = \operatorname{Re}(z) > 0$, se tiene

$$\int_1^{\infty} \left| \frac{t - \lfloor t \rfloor}{t^{z+1}} \right| dt = \int_1^{\infty} \frac{t - \lfloor t \rfloor}{t^{r+1}}\, dt \leq \int_1^{\infty} \frac{1}{t^{r+1}}\, dt < \infty.$$

Obviamente, uno de nuestros objetivos va a ser comprobar que las expresiones (8.18) y (8.19) son compatibles, esto es, que coinciden cuando $\operatorname{Re}(z) > 1$. Para ello nos serviremos de una expresión intermedia,

$$\zeta(z) = \lim_{x \to \infty} \left(\sum_{n \leq x} \frac{1}{n^z} - \frac{x^{1-z}}{1-z} \right), \quad \operatorname{Re}(z) > 0,\ z \neq 1; \tag{8.20}$$

En el apartado (b) de la proposición 8.10 veremos que, efectivamente, las dos expresiones (8.19) y (8.20) dan lugar a la misma función. Con esto, ya está claro cómo sale (8.18): en (8.20), descomponemos el límite en dos sumandos, y, al ser $\operatorname{Re}(z) > 1$, queda

$$\zeta(z) = \lim_{x \to \infty} \sum_{n \leq x} \frac{1}{n^z} - \lim_{x \to \infty} \frac{x^{1-z}}{1-z} = \sum_{n=1}^{\infty} \frac{1}{n^z} - 0.$$

[3] Así la hemos usado ya en el tema 6, por ejemplo; allí, en la sección 6.9, calculábamos $\zeta(2k)$ para k entero positivo.

Nota 8.9. Aunque en este tema no lo usaremos, aún se puede extender más la definición de la función zeta de Riemann; de este modo se llega a una función analítica en $\mathbb{C} \setminus \{1\}$, con un polo simple en 1. Para $\mathrm{Re}(z) < 0$, la extensión se hace por medio de la ecuación funcional

$$\zeta(z) = 2^z \pi^{1-z} \Gamma(1-z) \zeta(1-z) \operatorname{sen}(\pi z/2),$$

donde Γ es la función Gamma de Euler, que para $\mathrm{Re}(z) > 0$ se define como

$$\Gamma(z) = \int_0^\infty e^{-t} t^{z-1}\, dt$$

(y se va extendiendo a $\mathrm{Re}(z) < 0$ mediante sucesivas iteraciones con $\Gamma(z+1) = z\Gamma(z)$).

En particular, la ecuación funcional hace que $z = -2, -4, -6, \ldots$ sean ceros de la función $\zeta(z)$, los denominados «ceros triviales». No es muy difícil probar que todos los ceros no triviales de $\zeta(z)$ están en la franja $0 \leq \mathrm{Re}(z) \leq 1$, que se denomina banda crítica (nosotros lo veremos con detalle en el tema 10). La hipótesis de Riemann conjetura que todos ellos son de la forma $z = 1/2 + it$, con $t \in \mathbb{R}$ (es decir, están en la recta crítica). Desde que fue formulada en 1859, numerosos matemáticos han intentado, sin éxito, probarla (David Hilbert, en su famosa conferencia de 1900, la colocó en la lista de los problemas[4] más importantes a los que los matemáticos del siglo XX deberían enfrentarse). En el camino se han producido algunos avances; por ejemplo, en 1915, Godfrey Harold Hardy probó que existen infinitos ceros en la recta crítica. Actualmente, el problema sigue abierto, y posiblemente sea el más perseguido de todas las matemáticas. Sus repercusiones en teoría de números son enormes. Destaquemos simplemente que su demostración permitiría dar resultados más precisos sobre la distribución de los números primos. Algo de esto nos aparecerá en el tema 10, el dedicado al teorema de los números primos (en concreto, en el apartado 10.6.2); véase también el apartado 9.3.1 del tema 9. ◁

8.3.1. Algunas fórmulas elementales

Vamos ya con los resultados de esta sección; como va a ser la primera vez que manejamos las «O grandes», haremos las demostraciones con todo detalle.

[4] En este libro, ya nos hemos referido a la conferencia de Hilbert y su lista de 23 problemas en la sección 1.6 (concretamente, en la página 60), el apartado 3.2.5 (página 193), la nota 6.15 (página 356) y en el apartado 6.8.1 (página 380), donde nos han aparecido otros cuatro de esos problemas relacionados, respectivamente, con la solución de ecuaciones diofánticas, las leyes de reciprocidad, la hipótesis del continuo, y la trascendencia de ciertos números.

Además, observe el lector que es ahora cuando vamos a definir la *constante de Euler* (denotada γ) con rigor. Dicha constante ya nos ha aparecido en este libro, y lo volverá a hacer más adelante (en particular, en los ejercicios de este tema se muestran otras expresiones para γ), pues tiene gran importancia en teoría de números. Un interesante libro por completo dedicado a ella es [51].

Proposición 8.10.

Cuando $x \to \infty$, se verifican las siguientes estimaciones asintóticas:

(a) Las sumas parciales de la serie armónica cumplen

$$\sum_{n\leq x} \frac{1}{n} = \log(x) + \gamma + O\left(\frac{1}{x}\right),$$

donde γ es la constante de Euler (también llamada de Euler-Mascheroni)

$$\gamma = 1 - \int_1^\infty \frac{t - \lfloor t \rfloor}{t^2}\, dt = \lim_{x\to\infty}\left(\sum_{n\leq x} \frac{1}{n} - \log(x)\right).$$

(b) Si $s \in \mathbb{C} \setminus \{1\}$ y $\mathrm{Re}(s) > 0$, entonces

$$\sum_{n\leq x} \frac{1}{n^s} = \frac{x^{1-s}}{1-s} + \zeta(s) + O\left(\frac{1}{x^{\mathrm{Re}(s)}}\right),$$

donde la función $\zeta(s)$ cumple

$$\zeta(s) = \frac{s}{s-1} - s\int_1^\infty \frac{t - \lfloor t \rfloor}{t^{s+1}}\, dt = \lim_{x\to\infty}\left(\sum_{n\leq x} \frac{1}{n^s} - \frac{x^{1-s}}{1-s}\right).$$

(c) Si $s \in \mathbb{C}$, $\mathrm{Re}(s) > 1$, se tiene

$$\sum_{n>x} \frac{1}{n^s} = O(x^{1-\mathrm{Re}(s)}),$$

(d) Si $\mathrm{Re}(a) \geq 0$,

$$\sum_{n\leq x} n^a = \frac{x^{a+1}}{a+1} + O(x^{\mathrm{Re}(a)}),$$

(e) Finalmente,

$$\sum_{n\leq x} \log(n) = \log(\lfloor x \rfloor !) = x\log(x) - x + O(\log(x)).$$

Demostración. Como estamos estudiando lo que ocurre cuando $x \to \infty$ podemos suponer que x es lo suficientemente grande; en particular, en lo que sigue asumiremos que $x \geq 1$.

(a) Usamos la fórmula de sumación de Euler con $f(t) = 1/t$ e $y = 1$. Así,

$$\begin{aligned}\sum_{n\leq x} \frac{1}{n} &= 1 + \sum_{1<n\leq x} \frac{1}{n} = 1 + \frac{\lfloor x\rfloor - x}{x} + \int_1^x \frac{dt}{t} - \int_1^x \frac{t-\lfloor t\rfloor}{t^2}\,dt \\ &= \log(x) + 1 - \int_1^\infty \frac{t-\lfloor t\rfloor}{t^2}\,dt + \int_x^\infty \frac{t-\lfloor t\rfloor}{t^2}\,dt - \frac{x-\lfloor x\rfloor}{x} \\ &= \log(x) + 1 - \int_1^\infty \frac{t-\lfloor t\rfloor}{t^2}\,dt + O\left(\frac{1}{x}\right),\end{aligned}$$

donde en $O(1/x)$ hemos metido los dos últimos sumandos de la línea anterior, que están adecuadamente acotados (nótese que $|t - \lfloor t\rfloor| \leq 1$ y $\int_x^\infty 1/t^2\,dt = 1/x$). Despejando obtenemos

$$1 - \int_1^\infty \frac{t-\lfloor t\rfloor}{t^2}\,dt = \sum_{n\leq x} \frac{1}{n} - \log(x) + O\left(\frac{1}{x}\right)$$

y, tomando límites cuando $x \to \infty$, queda

$$1 - \int_1^\infty \frac{t-\lfloor t\rfloor}{t^2}\,dt = \lim_{x\to\infty}\left(\sum_{n\leq x} \frac{1}{n} - \log(x)\right),$$

que es la constante de Euler.

(b) En la fórmula de sumación de Euler, tomamos ahora $f(t) = t^{-s}$ e $y = 1$, con lo cual

$$\begin{aligned}\sum_{n\leq x} \frac{1}{n^s} &= 1 + \sum_{1<n\leq x} \frac{1}{n^s} = 1 + \frac{\lfloor x\rfloor - x}{x^s} + \int_1^x \frac{dt}{t^s} - s\int_1^x \frac{t-\lfloor t\rfloor}{t^{s+1}}\,dt \\ &= 1 + \frac{x^{1-s}}{1-s} - \frac{1}{1-s} - s\int_1^x \frac{t-\lfloor t\rfloor}{t^{s+1}}\,dt + O(x^{-\operatorname{Re}(s)}) \\ &= \frac{x^{1-s}}{1-s} + \frac{s}{s-1} - s\int_1^x \frac{t-\lfloor t\rfloor}{t^{s+1}}\,dt + O(x^{-\operatorname{Re}(s)}) \\ &= \frac{x^{1-s}}{1-s} + \zeta(s) + O(x^{-\operatorname{Re}(s)}),\end{aligned}$$

donde en el último paso hemos usado la definición de $\zeta(s)$, es decir, (8.19). Ahora, despejando y tomando límites,

$$\zeta(s) = \lim_{x\to\infty}\left(\sum_{n\le x}\frac{1}{n^s} - \frac{x^{1-s}}{1-s}\right).$$

Y de aquí, si $\operatorname{Re}(s) > 1$, resulta ser $\zeta(s) = \sum_{n=1}^{\infty} 1/n^s$.

(c) Según el apartado anterior,

$$\begin{aligned}\sum_{n>x}\frac{1}{n^s} &= \sum_{n=1}^{\infty}\frac{1}{n^s} - \sum_{n\le x}\frac{1}{n^s} = \zeta(s) - \left(\frac{x^{1-s}}{1-s} + \zeta(s) + O(x^{-\operatorname{Re}(s)})\right)\\ &= -\frac{x^{1-s}}{1-s} + O(x^{-\operatorname{Re}(s)}) = O(x^{1-\operatorname{Re}(s)})\end{aligned}$$

(en el último paso, obsérvese que $x^{1-s} = O(x^{1-\operatorname{Re}(s)})$, y que el sumando $O(x^{-\operatorname{Re}(s)})$ ya no es necesario, pues queda englobado por $O(x^{1-\operatorname{Re}(s)})$).

(d) En la fórmula de sumación de Euler, tomemos $f(t) = t^a$ e $y = 1$. De este modo,

$$\begin{aligned}\sum_{n\le x} n^a &= 1 + \sum_{1<n\le x} n^a = 1 - (x - \lfloor x\rfloor)x^a + \int_1^x t^a\,dt + a\int_1^x (t-\lfloor t\rfloor)t^{a-1}\,dt\\ &= \frac{x^{a+1}}{a+1} + O(x^{\operatorname{Re}(a)}).\end{aligned}$$

(e) Tomando esta vez $f(t) = \log(t)$ e $y = 1$, queda

$$\begin{aligned}\sum_{n\le x}\log(n) &= 0 + \sum_{1<n\le x}\log(n)\\ &= -(x-\lfloor x\rfloor)\log(x) + \int_1^x \log(t)\,dt + \int_1^x \frac{t-\lfloor t\rfloor}{t}\,dt\\ &= O(\log(x)) + x\log(x) - x + 1 + \int_1^x \frac{t-\lfloor t\rfloor}{t}\,dt\\ &= x\log(x) - x + O(\log(x)).\end{aligned}$$

□

8.4. Orden medio de las funciones divisor

Comenzaremos mostrando una estimación asintótica —acompañada de una cota del término del error— para las medias de la función σ, esto es, $\frac{1}{x}\sum_{n\leq x}\sigma(n)$. Naturalmente, el $1/x$ no hace falta arrastrarlo, pues si lo que encontramos es una expresión asintótica para $\sum_{n\leq x}\sigma(n)$, bastaría dividirla por x. A partir de ahora, lo haremos así habitualmente, sin necesidad de repetir esta aclaración.

Además, como estamos estudiando lo que ocurre cuando $x\to\infty$, podemos suponer que, en todas las expresiones en las que aparezca, será $x\geq 1$, tanto en esta sección como en el resto del tema.

Teorema 8.11.
Cuando $x\to\infty$, se verifica[5]

$$\sum_{n\leq x}\sigma(n)=\frac{1}{2}\zeta(2)x^2+O(x\log(x)).$$

Demostración. Al ser $\sigma(n)=\sum_{d\mid n}d$, se tiene

$$\sum_{n\leq x}\sigma(n)=\sum_{n\leq x}\sum_{d\mid n}d; \tag{8.21}$$

ésta es una suma doble extendida a n y a d. Como $d\,|\,n$, podemos escribir $n=cd$, y efectuar la suma sobre todos los pares de enteros positivos c y d tales que $cd\leq x$, esto es,

$$\sum_{n\leq x}\sigma(n)=\sum_{\substack{c,d\\ cd\leq x}}d \tag{8.22}$$

(otras veces que usemos un argumento similar, ya no detallaremos c,d en el sumatorio, y escribiremos únicamente $\sum_{cd\leq x}$). Podemos ahora entender esta suma como $c\leq x$ y $d\leq x/c$, con lo cual

$$\sum_{n\leq x}\sigma(n)=\sum_{c\leq x}\sum_{d\leq x/c}d.$$

Gracias al apartado (d) de la proposición 8.10, sabemos estimar el sumatorio interno

[5] Además —y aunque aquí no es relevante— el valor de $\zeta(2)$ es conocido: $\zeta(2)=\pi^2/6$ (de hecho, puede verse una demostración en la sección 6.9 del presente libro). Este mismo comentario se aplica a otros resultados de ésta y las siguientes secciones.

$\sum_{d\le x/c} d$; así,

$$\sum_{n\le x}\sigma(n)=\sum_{c\le x}\sum_{d\le x/c} d=\sum_{c\le x}\left(\frac{1}{2}\left(\frac{x}{c}\right)^2+O\left(\frac{x}{c}\right)\right).$$

Por la desigualdad triangular, podemos meter el $\sum$ dentro de la O. Entonces, aplicando (b) y (a) de la proposición 8.10, tenemos

$$\begin{aligned}\sum_{n\le x}\sigma(n)&=\frac{x^2}{2}\sum_{c\le x}\frac{1}{c^2}+O\left(x\sum_{c\le x}\frac{1}{c}\right)\\&=\frac{x^2}{2}\left(\frac{-1}{x}+\zeta(2)+O\left(\frac{1}{x^2}\right)\right)+O(x\log(x))\\&=\frac{1}{2}\zeta(2)x^2+O(x\log(x))\end{aligned}$$

(en el último paso, hemos despreciado todos los sumandos que —asintóticamente— son más pequeños que el error $x\log(x)$). □

8.4.1. Parámetro nulo (número de divisores)

Mostremos ahora una fórmula asintótica para las sumas de la función τ. Es interesante comentar que, en la demostración, podíamos haber seguido el mismo razonamiento que en el teorema anterior. Pero, de ese modo, hubiéramos obtenido un resultado más débil:

$$\sum_{n\le x}\tau(n)=x\log(x)+O(x) \tag{8.23}$$

(se invita al lector —ejercicio 8.10— a que se convenza de ello por sí solo). Así que usaremos un proceso mejor adaptado —el *método de la hipérbola* ideado por Lejeune Dirichlet— que llevará a una estimación más fina:

Teorema 8.12. Dirichlet, 1838.
Cuando $x\to\infty$, se verifica

$$\sum_{n\le x}\tau(n)=x\log(x)+(2\gamma-1)x+O(\sqrt{x}).$$

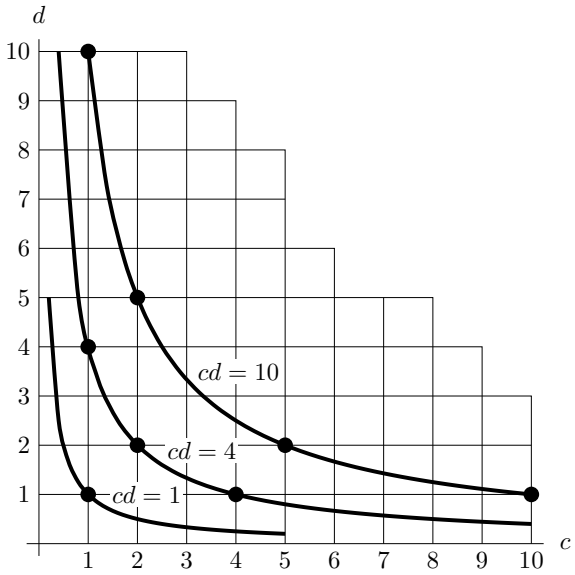


Figura 8.1: Puntos (c, d) de coordenadas enteras en las hipérbolas $cd = n$, con $n = 1, \ldots, \lfloor x \rfloor$.

Demostración. Puesto que $\tau(n) = \sum_{d|n} 1$, comenzando —de momento— como en (8.21) y (8.22) de la demostración anterior, tenemos

$$\sum_{n\le x} \tau(n) = \sum_{n\le x} \sum_{d|n} 1 = \sum_{cd\le x} 1 = \sum_{n\le x} \sum_{cd=n} 1.$$

donde en la última expresión hemos agrupado todos los sumandos tales que $cd = n$. Esto se puede interpretar como una suma extendida a ciertos puntos (de coordenadas enteras positivas) del plano cd. Para cada $n = 1, \ldots, \lfloor x \rfloor$, los puntos $cd = n$ pertenecen a una hipérbola, tal como sugiere la figura 8.1, en la que estos puntos aparecen destacados.

Vamos ahora a sacar partido de la geometría de la figura, y de su simetría respecto a la bisectriz $d = c$; representamos esta situación en la figura 8.2. Tenemos tres tipos de puntos: los de la bisectriz, los de un lado de la bisectriz y los del otro; por simetría, hay tantos puntos a un lado de la bisectriz como al otro, luego contaremos los de un lado (en concreto, los que hay en el recinto delimitado por líneas gruesas) y multiplicaremos por dos. El número de puntos de la diagonal es

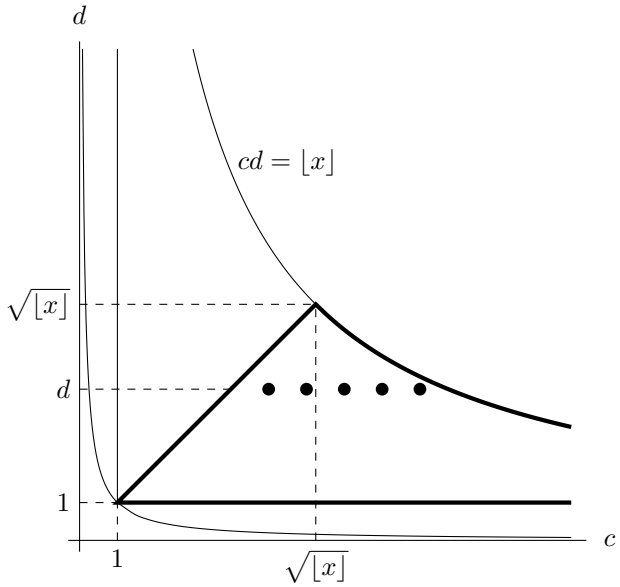


Figura 8.2: Puntos de coordenadas enteras a una cierta altura d.

$\sum_{d^2\leq x} 1$. Respecto a la zona delimitada por líneas gruesas, en la figura aparecen remarcados los puntos a un cierta altura d; su número es $\lfloor x/d\rfloor - d$. Por tanto, en esa zona hay $\sum_{d^2\leq x}(\lfloor x/d\rfloor - d)$ puntos. Así,

$$\sum_{n\leq x}\tau(n) = \sum_{d^2\leq x} 1 + 2\sum_{d^2\leq x}\left(\left\lfloor\frac{x}{d}\right\rfloor - d\right).$$

Entonces, aplicando (a) y (d) de la proposición 8.10, tenemos

$$\begin{aligned}\sum_{n\leq x}\tau(n) &= \sum_{d\leq\sqrt{x}} 1 + 2\sum_{d\leq\sqrt{x}}\left(\left\lfloor\frac{x}{d}\right\rfloor - d\right)\\ &= \lfloor\sqrt{x}\rfloor + 2\sum_{d\leq\sqrt{x}}\left(\frac{x}{d} - d + O(1)\right) = 2x\sum_{d\leq\sqrt{x}}\frac{1}{d} - 2\sum_{d\leq\sqrt{x}} d + O(\sqrt{x})\\ &= 2x\left(\log(\sqrt{x}) + \gamma + O\left(\frac{1}{\sqrt{x}}\right)\right) - 2\left(\frac{1}{2}x + O(\sqrt{x})\right) + O(\sqrt{x})\\ &= x\log(x) + (2\gamma - 1)x + O(\sqrt{x}).\end{aligned}$$

□

Un problema muy interesante es conocer hasta cuánto se puede rebajar θ de modo que se cumpla

$$\sum_{n\leq x}\tau(n) = x\log(x) + (2\gamma - 1)x + O(x^{\theta}).$$

El teorema anterior muestra que $\theta = 1/2$ sirve. En 1916, Landau probó que, para que eso sea cierto, debemos tomar $\theta \geq 1/4$. Diversos matemáticos han ido perfeccionando estos resultados, acercando las cotas superiores e inferiores de θ, pero el valor límite continúa sin conocerse.

8.4.2. Parámetro positivo

Ya sabemos lo que ocurre con las medias de las funciones $\sigma_1 = \sigma$ y $\sigma_0 = \tau$. Analicemos ahora el caso correspondiente a σ_a con a positivo distinto de 1:

Teorema 8.13.
Si $a > 0$ y $a \neq 1$, entonces, cuando $x \to \infty$,

$$\sum_{n\leq x}\sigma_a(n) = \frac{1}{a+1}\zeta(a+1)x^{a+1} + O(x^b),$$

donde $b = \max\{1, a\}$.

Demostración. Procederemos como en el teorema 8.11. Usando, esta vez, (d) y (b) de la proposición 8.10, tenemos

$$\begin{aligned}
\sum_{n\leq x}\sigma_a(n) &= \sum_{n\leq x}\sum_{d|n} d^a = \sum_{cd\leq x} d^a = \sum_{c\leq x}\sum_{d\leq x/c} d^a \\
&= \sum_{c\leq x}\left(\frac{1}{a+1}\left(\frac{x}{c}\right)^{a+1} + O\left(\frac{x^a}{c^a}\right)\right) = \frac{x^{a+1}}{a+1}\sum_{c\leq x}\frac{1}{c^{a+1}} + O\left(\sum_{c\leq x}\frac{x^a}{c^a}\right) \\
&= \frac{x^{a+1}}{a+1}\left(\frac{x^{-a}}{-a} + \zeta(a+1) + O\left(\frac{1}{x^{a+1}}\right)\right) \\
&\qquad + O\left(x^a\left(\frac{x^{1-a}}{1-a} + \zeta(a) + O\left(\frac{1}{x^a}\right)\right)\right) \\
&= \frac{1}{a+1}\zeta(a+1) + O(x) + O(x^a) = \frac{1}{a+1}\zeta(a+1) + O(x^b). \qquad \square
\end{aligned}$$

8.4.3. Parámetro negativo

Finalmente, estudiemos el orden medio de σ_a con a negativo; o, en otras palabras, el de σ_{-a} con a positivo.

Teorema 8.14.
Cuando $x \to \infty$, se tiene

$$\sum_{n\le x} \sigma_{-1}(n) = \zeta(2)x + O(\log(x)) \tag{8.24}$$

y, si $a > 0$ y $a \neq 1$,

$$\sum_{n\le x} \sigma_{-a}(n) = \zeta(a+1)x + O(x^b), \tag{8.25}$$

donde $b = \max\{0, 1-a\}$.

Demostración. Comencemos con $a > 0$ cualquiera. Así,

$$\begin{aligned}\sum_{n\le x} \sigma_{-a}(n) &= \sum_{n\le x}\sum_{d|n} d^{-a} = \sum_{cd\le x} d^{-a} = \sum_{d\le x} d^{-a} \sum_{c\le x/d} 1\\ &= \sum_{d\le x} \frac{1}{d^a}\left(\frac{x}{d} + O(1)\right) = x\sum_{d\le x}\frac{1}{d^{a+1}} + O\left(\sum_{d\le x}\frac{1}{d^a}\right)\\ &= x\left(\frac{x^{-a}}{-a} + \zeta(a+1) + O\left(\frac{1}{x^{a+1}}\right)\right) + O\left(\sum_{d\le x}\frac{1}{d^a}\right)\\ &= \zeta(a+1)x + O(x^{1-a}) + O\left(\sum_{d\le x}\frac{1}{d^a}\right).\end{aligned}$$

Ahora, analicemos por separado los casos $a = 1$ y $a \neq 1$. Si $a = 1$, entonces

$$\sum_{d\le x}\frac{1}{d^a} = \sum_{d\le x}\frac{1}{d} = O(\log(x))$$

y se obtiene (8.24). Si $a \neq 1$, entonces

$$\sum_{d\le x}\frac{1}{d^a} = \frac{x^{1-a}}{1-a} + \zeta(a) + O\left(\frac{1}{x^a}\right) = O(x^b),$$

de donde se sigue (8.25). □

8.5. Orden medio de la indicatriz de Euler

El objetivo de esta sección es encontrar estimaciones asintóticas para el orden medio de la indicatriz de Euler. Para ello, necesitaremos algunos resultados previos.

Lema 8.15.

Sean $f \geq 0$ y $g \geq 0$ funciones aritméticas tales que las series $\sum_{n=1}^{\infty} f(n)$ y $\sum_{m=1}^{\infty} g(m)$ convergen. Entonces,

$$\sum_{n=1}^{\infty} f(n) \cdot \sum_{m=1}^{\infty} g(m) = \sum_{k=1}^{\infty} \sum_{nm=k} f(n)g(m) = \sum_{k=1}^{\infty} (f * g)(k).$$

(En realidad, el resultado es válido aunque las series diverjan a $+\infty$, y con el convenio usual de que $0 \cdot (+\infty) = 0$ si surge el caso).

Demostración. Al ser $f \geq 0$ y $g \geq 0$, es claro que, para cada $N \in \mathbb{N}$, se cumple

$$\sum_{k=1}^{N} \left(\sum_{nm=k} f(n)g(m) \right) \leq \sum_{n=1}^{N} f(n) \cdot \sum_{m=1}^{N} g(m) \leq \sum_{k=1}^{N^2} \left(\sum_{nm=k} f(n)g(m) \right)$$

Ahora, basta tomar límites con $N \to \infty$. La observación entre paréntesis es inmediata. □

Para funciones aritméticas arbitrarias podemos probar un resultado similar. Pero, esta vez, las hipótesis exigen convergencia absoluta de las series.

Lema 8.16.

Sean f y g funciones aritméticas y supongamos que las series $\sum_{n=1}^{\infty} |f(n)|$ y $\sum_{m=1}^{\infty} |g(m)|$ convergen. Entonces, la serie

$$\sum_{k=1}^{\infty} \left(\sum_{nm=k} f(n)g(m) \right)$$

también converge, y

$$\sum_{n=1}^{\infty} f(n) \cdot \sum_{m=1}^{\infty} g(m) = \sum_{k=1}^{\infty} \sum_{nm=k} f(n)g(m) = \sum_{k=1}^{\infty} (f * g)(k).$$

Demostración. Vamos a hacerla de dos maneras distintas. La primera, usando un conocido resultado de análisis: una serie que converge absolutamente es incondicionalmente convergente (esto es, se pueden reordenar sus términos y su suma no

cambia). La segunda (que sólo esbozamos, pues es un proceso estándar), apoyándonos en el lema anterior.

Primera prueba. Si las dos series convergen absolutamente, podemos multiplicarlas y reordenar convenientemente sus términos sin que ello altere su suma. Así, agrupando los términos para los que $nm = k$, tenemos

$$\sum_{n=1}^{\infty} f(n) \cdot \sum_{m=1}^{\infty} g(m) = \sum_{k=1}^{\infty} \sum_{nm=k} f(n)g(m).$$

Segunda prueba. Si f y g son reales, tomamos sus partes positiva y negativa:

$$f^+(n) = \max\{f(n), 0\}, \quad f^-(n) = \max\{-f(n), 0\}$$

(y lo mismo con g). Entonces, $|f| = f^+ + f^-$, $f = f^+ - f^-$ y el resultado se deduce del lema anterior (todas las series que aparecen son convergentes). Si f y g son complejas, se toman partes reales e imaginarias y se aplica lo ya probado para funciones reales. □

A continuación, usaremos el lema para encontrar el valor de la suma $\sum_{n=1}^{\infty} \frac{\mu(n)}{n^2}$. Esto nos será útil en nuestro objetivo de encontrar estimaciones asintóticas para el orden medio de la indicatriz de Euler; pero, además, es un resultado que tiene interés en sí mismo.

Teorema 8.17.
La función de Möbius verifica

$$\sum_{n=1}^{\infty} \frac{\mu(n)}{n^2} = \frac{1}{\zeta(2)}. \tag{8.26}$$

Demostración. Apliquemos el lema anterior con las funciones $f(n) = \mu(n)/n^2$ y $g(m) = 1/m^2$. Entonces,

$$\begin{aligned}\sum_{n=1}^{\infty} \frac{\mu(n)}{n^2} \cdot \sum_{m=1}^{\infty} \frac{1}{m^2} &= \sum_{k=1}^{\infty} \sum_{nm=k} \frac{\mu(n)}{(nm)^2} = \sum_{k=1}^{\infty} \frac{1}{k^2} \sum_{nm=k} \mu(n)u(m) \\ &= \sum_{k=1}^{\infty} \frac{1}{k^2} (\mu * u)(k) = \sum_{k=1}^{\infty} \frac{1}{k^2} \delta(k) = 1,\end{aligned}$$

de donde, despejando $\sum_{n=1}^{\infty} \mu(n)/n^2$, se sigue el resultado. □

Corolario 8.18.
Cuando $x \to \infty$, la función de Möbius verifica

$$\sum_{n\le x} \frac{\mu(n)}{n^2} = \frac{1}{\zeta(2)} + O\left(\frac{1}{x}\right). \tag{8.27}$$

Demostración. Según (8.26),

$$\sum_{n\le x} \frac{\mu(n)}{n^2} = \frac{1}{\zeta(2)} + \sum_{n>x} \frac{-\mu(n)}{n^2} = \frac{1}{\zeta(2)} + O\left(\sum_{n>x} \frac{1}{n^2}\right) = \frac{1}{\zeta(2)} + O\left(\frac{1}{x}\right),$$

donde en el último paso hemos usado (c) de la proposición 8.10. □

Por fin, y tal como demostró Franz Mertens, llegamos a lo siguiente:

Teorema 8.19. Mertens, 1874.
Cuando $n \to \infty$, las sumas de la indicatriz de Euler satisfacen

$$\sum_{n\le x} \varphi(n) = \frac{1}{2\zeta(2)}\, x^2 + O(x\log(x)).$$

Demostración. Por la proposición 7.8 del tema 7, sabemos que

$$\varphi(n) = (N * \mu)(n) = \sum_{d|n} \mu(d)\frac{n}{d}.$$

De aquí, aplicando la proposición 8.10 y el corolario 8.18, obtenemos

$$\begin{aligned}
\sum_{n\le x} \varphi(n) &= \sum_{n\le x}\sum_{cd=n} \mu(d)c = \sum_{cd\le x} \mu(d)c \\
&= \sum_{d\le x} \mu(d) \sum_{c\le x/d} c = \sum_{d\le x} \mu(d)\left(\frac{1}{2}\left(\frac{x}{d}\right)^2 + O\left(\frac{x}{d}\right)\right) \\
&= \frac{1}{2}\, x^2 \sum_{d\le x} \frac{\mu(d)}{d^2} + x\, O\left(\sum_{d\le x} \frac{1}{d}\right) \\
&= \frac{1}{2}\, x^2 \left(\frac{1}{\zeta(2)} + O\left(\frac{1}{x}\right)\right) + O(x\log(x)) \\
&= \frac{1}{2\zeta(2)}\, x^2 + O(x\log(x)).
\end{aligned}$$

□

8.5.1. ¿Cuál es la probabilidad de que dos números tomados al azar sean primos entre sí?

Como consecuencia del teorema 8.19, estamos en condiciones de proporcionar la siguiente interpretación probabilística:

Corolario 8.20.
La probabilidad de que dos enteros positivos menores o iguales que N sean primos entre sí es

$$\frac{1}{\zeta(2)} + O(N^{-1}\log(N)).$$

Demostración. El número de pares de enteros positivos $\{r, s\}$ tales que $1 \leq r \leq s \leq N$ es el siguiente: en N casos r y s son iguales; y hay otros $\binom{N}{2}$ casos en los que r y s son distintos uno del otro. En total, $N + \binom{N}{2} = N(N+1)/2$. Por otra parte, por definición de φ, el número de tales pares en los que r y s son primos entre sí es $\sum_{s\leq N} \varphi(s) = \sum_{s\leq N} \operatorname{card}\{r \leq s : \operatorname{mcd}(r,s) = 1\}$.

Entonces, la probabilidad —casos favorables dividido entre casos posibles— de que dos enteros positivos menores o iguales que N sean primos entre sí es

$$\begin{aligned}\frac{\sum_{s\leq N}\varphi(s)}{N(N+1)/2} &= \frac{1}{2\zeta(2)}\,\frac{N^2}{N(N+1)/2} + \frac{1}{N(N+1)/2}\,O(N\log(N)) \\ &= \frac{1}{\zeta(2)}\left(1 - \frac{1}{N+1}\right) + O(N^{-1}\log(N)) \\ &= \frac{1}{\zeta(2)} + O(N^{-1}\log(N)).\end{aligned}$$

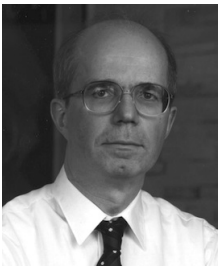
□

Habitualmente, al resultado anterior se alude una vez que se ha hecho tender $N \to \infty$. Así se dice que «la probabilidad de que dos enteros positivos sean primos entre sí es $6/\pi^2$» (lógicamente, aquí hemos usado que $\zeta(2) = \pi^2/6$). O, alternativamente, «la probabilidad de que una fracción tomada al azar sea irreducible es $6/\pi^2$». De hecho, es este tipo de interpretaciones probabilísticas de claro significado aritmético —que también aparecen en otras partes de este capítulo y los dos siguientes— el que a menudo incita a estudiar el comportamiento asintótico de las funciones aritméticas o de sus medias.

No está de más citar que fue el eminente matemático ruso Pafnuty Chebyshev —destacado tanto por sus contribuciones a la teoría de números como por ser uno de los padres de la teoría de probabilidades, por citar sólo lo que ahora nos concierne— el primero que se planteó y resolvió la cuestión planteada en este apartado.



Franz Mertens
(1840–1927)



Andrew M. Odlyzko
(1949–)



Herman J. J. te Riele
(1947–)

8.5.2. Series de Dirichlet

Casi sin ningún esfuerzo, el resultado del lema 8.16 se puede generalizar a lo que se denominan series de Dirichlet, que son series de la forma

$$F(s) = \sum_{n=1}^{\infty} \frac{f(n)}{n^s}$$

con f una función aritmética y s perteneciente a una región del plano complejo (un semiplano, concretamente). Por ejemplo, la función zeta de Riemann $\zeta(s) = \sum_{n=1}^{\infty} \frac{1}{n^s}$ es la serie de Dirichlet de la función $f(n) = 1$ para todo $n \in \mathbb{N}$; es absolutamente convergente en el semiplano $\mathrm{Re}(s) > 1$. Mediante series de Dirichlet se pueden encontrar resultados similares a (8.26) en los que el «2» se sustituye por una «s»; y también usando —en vez de la función de Möbius— otras funciones aritméticas. A este respecto, véanse los ejercicios 8.11 a 8.19.

8.6. Estimaciones elementales relacionadas con la función de Möbius

El orden medio de la función de Möbius es mucho más difícil de establecer que los que hemos visto hasta ahora para las funciones divisor y la indicatriz de Euler. Ya hemos comentado —en la sección 7.1 del tema 7— que se cumple

$$\lim_{x\to\infty} \frac{1}{x} \sum_{n\leq x} \mu(n) = 0, \tag{8.28}$$

y que este resultado es equivalente al teorema de los números primos. Es más, Franz Mertens, en 1897, conjeturó que, para cada $x > 1$, se cumplía

$$\sum_{n\leq x} \mu(n) < \sqrt{x}.$$

En 1985, Andrew Odlyzko y Herman te Riele demostraron que la conjetura de Mertens no era cierta, pero sin encontrar ningún contraejemplo concreto (que se sigue sin conocer, y que es uno de los retos más apasionantes de la teoría de números computacional). Actualmente se piensa que lo que será cierto es que

$$\sum_{n\leq x} \mu(n) = o(x^{1/2+\varepsilon}) \quad \text{para todo } \varepsilon > 0,$$

pero permanece sin confirmar ni refutar. Y éste sería, además, un resultado importante, pues se sabe que su demostración equivaldría a la de la hipótesis de Riemann (véase la nota 8.9 en la página 468).

En esta sección no abordaremos nada de esto, sino que nuestro objetivo ahora va a ser más modesto. En concreto, probaremos varias estimaciones asintóticas elementales en las que aparece la función de Möbius.

Comencemos recordando que ya hemos probado una de tales estimaciones asintóticas. En concreto, en el corolario 8.18 vimos que

$$\sum_{n\leq x} \frac{\mu(n)}{n^2} = \frac{1}{\zeta(2)} + O\left(\frac{1}{x}\right).$$

Éste era un útil resultado, al que ya sacamos partido en la demostración del teorema 8.19, en el que estimábamos el tamaño de las medias de la indicatriz de Euler. Ahora lo usaremos de nuevo; esta vez, para encontrar una estimación asintótica de las medias de los valores absolutos de la función de Möbius:

Teorema 8.21.
Cuando $x \to \infty$, se verifica

$$\sum_{n\leq x} |\mu(n)| = \frac{1}{\zeta(2)}\, x + O(\sqrt{x}).$$

Demostración. En realidad, vamos a probar un enunciado equivalente:

$$\sum_{n\leq x^2} |\mu(n)| = \frac{1}{\zeta(2)}\, x^2 + O(x).$$

Sean $F(x) = \sum_{n \le x^2} |\mu(n)|$ y $G(x) = \lfloor x^2 \rfloor$. Comprobaremos que F y G están ligadas por la relación $G = u \circ F$; entonces, la fórmula de inversión implica que $F = \mu \circ G$, y de aquí se deducirá el resultado (estamos ahora usando notación y resultados del tema anterior, en concreto, $\alpha \circ F$ se define en (7.20), y la fórmula de inversión es la que figura en el teorema 7.23). Vamos con ello.

Como

$$|\mu(n)| = \begin{cases} 1, & \text{si } n \text{ está libre de cuadrados,} \\ 0, & \text{si } n \text{ es divisible por algún cuadrado mayor que 1,} \end{cases}$$

resulta que

$$F(x) = \operatorname{card}\{\, n \le x^2 : n \text{ libre de cuadrados} \,\}.$$

Para cada $d \le x$, sea

$$A_d = \{\, n \le x^2 : d^2 \text{ es el mayor divisor cuadrado de } n \,\}.$$

Los conjuntos A_d son disjuntos y verifican $\{n : n \le x^2\} = \bigcup_{d \le x} A_d$. Además, la aplicación

$$\begin{aligned} A_d &\longrightarrow \{\, m \le x^2/d^2 : m \text{ libre de cuadrados} \,\} \\ n &\longmapsto n/d^2 \end{aligned} \tag{8.29}$$

es una biyección (compruébese: ejercicio 8.20), luego $\operatorname{card}(A_d) = F(x/d)$. Por consiguiente,

$$\begin{aligned} G(x) = \lfloor x^2 \rfloor &= \operatorname{card}\{n : n \le x^2\} \\ &= \operatorname{card}\left(\bigcup_{d \le x} A_d \right) = \sum_{d \le x} \operatorname{card}(A_d) = \sum_{d \le x} F\left(\frac{x}{d}\right) = (u \circ F)(x). \end{aligned}$$

La fórmula de inversión generalizada (teorema 7.23) nos asegura que, entonces, $F = \mu \circ G$. Y así, por el corolario 8.18 deducimos

$$\begin{aligned} F(x) = (\mu \circ G)(x) &= \sum_{d \le x} \mu(d) \left\lfloor \frac{x^2}{d^2} \right\rfloor = \sum_{d \le x} \mu(d) \left(\frac{x^2}{d^2} + O(1) \right) \\ &= x^2 \sum_{d \le x} \frac{\mu(d)}{d^2} + O(x) = x^2 \left(\frac{1}{\zeta(2)} + O\left(\frac{1}{x}\right) \right) + O(x) = \frac{x^2}{\zeta(2)} + O(x), \end{aligned}$$

tal como buscábamos. □

8.6.1. ¿Cuál es la probabilidad de que un número tomado al azar sea libre de cuadrados?

En términos probabilísticos, el resultado anterior se puede interpretar como sigue:

Corolario 8.22.

La probabilidad de que un entero positivo menor o igual que N no tenga factores cuadráticos es

$$\frac{1}{\zeta(2)} + O(N^{-1/2}).$$

Demostración. Un entero n no tiene factores cuadráticos si y sólo si $|\mu(n)| = 1$. Por tanto, la probabilidad (casos favorables dividido entre casos posibles) a la que alude el enunciado es

$$\frac{\sum_{n\leq N} |\mu(n)|}{N} = \frac{1}{\zeta(2)} + O(N^{-1/2}). \qquad \square$$

Esto, una vez que hemos hecho tender $N \to \infty$, y tenido en cuenta que $\zeta(2) = \pi^2/6$, se suele expresar diciendo que «la probabilidad de que un entero positivo no tenga factores cuadráticos es $6/\pi^2$». O también que «la densidad asintótica de los números libres de cuadrados es $6/\pi^2$».

8.6.2. Una cota no muy fina

A continuación, mostremos una cota para $|\sum_{n\leq x} \mu(n)/n|$. Como podemos ver fijándonos en su demostración, la cota no parece estar muy bien ajustada, pues, para obtenerla, se utiliza la desigualdad triangular de una manera muy burda. Aun así, resulta útil en algunas ocasiones (por ejemplo, en la demostración del teorema 8.30).

Teorema 8.23.

Para todo $x \geq 1$, se cumple que

$$\left| \sum_{n\leq x} \frac{\mu(n)}{n} \right| \leq 1. \tag{8.30}$$

Demostración. Si $1 \leq x < 2$, (8.30) es cierto de manera trivial, ya que entonces $\sum_{n\leq x} \mu(n)/n = 1$. Para $x \geq 2$, usemos la relación

$$\sum_{n\leq x} \mu(n) \left\lfloor \frac{x}{n} \right\rfloor = 1$$

que probamos en el tema anterior (concretamente, en el teorema 7.27). Con ella,

$$x\left|\sum_{n\le x}\frac{\mu(n)}{n}\right| = \left|1+\sum_{n\le x}\mu(n)\left(\frac{x}{n}-\left\lfloor\frac{x}{n}\right\rfloor\right)\right|$$
$$= \left|1+(x-\lfloor x\rfloor)+\sum_{2\le n\le x}\mu(n)\left(\frac{x}{n}-\left\lfloor\frac{x}{n}\right\rfloor\right)\right|.$$

Ahora, apliquemos la desigualdad triangular a través de todos los sumandos de $\sum_{2\le n\le x}$, usando $|\mu(n)|\le 1$ y $|x/n-\lfloor x/n\rfloor|\le 1$ (de paso, obsérvese que, al acotar dentro del sumatorio, se está perdiendo toda posible cancelación de términos proveniente de la oscilación de $\mu(n)$; esto sugiere que, con otro método, se podría lograr un resultado más fino). Así,

$$x\left|\sum_{n\le x}\frac{\mu(n)}{n}\right| \le 1+(x-\lfloor x\rfloor)+\sum_{2\le n\le x}1 = 1+(x-\lfloor x\rfloor)+(\lfloor x\rfloor-1) = x,$$

y de aquí se sigue (8.30). □

Merece la pena caer en la cuenta de lo interesante que hubiera sido poder demostrar fácilmente una acotación más fina que (8.30). En efecto, el siguiente resultado pone de manifiesto que, si hubiéramos llegado a probar $\lim_{x\to\infty}\sum_{n\le x}\mu(n)/n = 0$, eso hubiera implicado (8.28) y, por tanto, el teorema de los números primos.

Proposición 8.24.

Si fuera $\lim_{x\to\infty}\sum_{n\le x}\frac{\mu(n)}{n}=0$, *entonces también* $\lim_{x\to\infty}\frac{1}{x}\sum_{n\le x}\mu(n)=0$.

Demostración. Usemos la identidad de Abel con

$$a(n)=\mu(n)/n,\ A(x)=\sum_{n\le x}\mu(n)/n,\ f(t)=t \text{ e } y=1.$$

Así, resulta

$$\sum_{n\le x}\mu(n) = 1+\sum_{1<n\le x}\frac{\mu(n)}{n}\cdot n = 1+A(x)f(x)-A(1)f(1)-\int_1^x A(t)\,dt$$
$$= xA(x)-\int_1^x A(t)\,dt.$$

Por tanto,

$$\frac{1}{x}\sum_{n\le x}\mu(n) = \sum_{n\le x}\frac{\mu(n)}{n}-\frac{1}{x}\int_1^x A(t)\,dt$$

y basta demostrar que

$$\lim_{x\to\infty} \frac{1}{x}\int_1^x A(t)\,dt = 0. \tag{8.31}$$

Por hipótesis, $A(x) \to 0$ cuando $x \to \infty$, luego, dado $\varepsilon > 0$, existe $c > 0$ tal que $|A(x)| < \varepsilon/2$ para todo $x > c$. Además, por el teorema anterior, $|A(x)| \le 1$ para $x \ge 1$. Por tanto, si $x > c$, se tiene

$$\left|\frac{1}{x}\int_1^x A(t)\,dt\right| \le \frac{1}{x}\int_1^c |A(t)|\,dt + \frac{1}{x}\int_c^x |A(t)|\,dt \le \frac{c-1}{x} + \frac{\varepsilon}{2}\,\frac{x-c}{x}.$$

Como consecuencia, existe $b > 0$ tal que, para $x > b$,

$$\left|\frac{1}{x}\int_1^x A(t)\,dt\right| < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon,$$

y esto prueba (8.31). □

Nota 8.25. Aunque nosotros no lo probaremos, también es cierto el recíproco. En consecuencia, $\lim_{x\to\infty} \sum_{n\le x} \frac{\mu(n)}{n} = 0$ es, como $\lim_{x\to\infty} \frac{1}{x}\sum_{n\le x} \mu(n) = 0$, equivalente al teorema de los números primos.

La convergencia de la serie $\sum_{n=1}^{\infty} \frac{\mu(n)}{n}$ tiene su propia historia. Euler pensó que había probado que

$$\sum_{n=1}^{\infty} \frac{\mu(n)}{n} = 0 \tag{8.32}$$

pero su argumento asumía que la serie converge, y eso no es —ni mucho menos— obvio. La primera demostración correcta de (8.32) data de 1897 y se debe a Hans von Mangoldt. Que es un resultado equivalente al teorema de los números primos lo probó Landau en 1911. ◁

8.7. Resultados equivalentes al teorema de los números primos

Antes que nada, pongamos de manifiesto que los resultados de ésta y la siguiente sección no se van a usar ni en el resto del tema, ni en el tema siguiente. Si el lector lo desea —y es fácil que sea así en una primera lectura— puede saltarse ambas secciones sin que ello cree lagunas más adelante. No obstante, sí que le recomendamos leer los comentarios del principio, hasta que comience el apartado 8.7.1.

Nuestro objetivo ahora va a ser demostrar que los tres hechos

$$\lim_{x\to\infty} \frac{1}{x} \sum_{n\leq x} \mu(n) = 0, \tag{8.33}$$

$$\lim_{x\to\infty} \frac{1}{x} \sum_{n\leq x} \Lambda(n) = 1, \tag{8.34}$$

$$\lim_{x\to\infty} \frac{\pi(x)}{x/\log(x)} = 1 \tag{8.35}$$

son equivalentes. Realmente, aunque aquí no los hemos probado, esos tres resultados son ciertos (es decir, existen los límites y valen lo que se afirma), luego parece una perogrullada decir que son equivalentes. No causa problemas lógicos afirmar que la hipótesis de Riemann es equivalente a que $\sum_{n\leq x} \mu(n) = o(x^{1/2+\varepsilon})$ para todo $\varepsilon > 0$ ya que, en este caso, ninguno de los dos resultados se sabe si es verdadero o falso. La equivalencia quiere decir que si se demostrara o refutara uno, lo mismo le ocurriría al otro, pues se sabe cómo probar cada uno de ellos usando el otro (junto con todo tipo de matemáticas lícitas, claro). Pero no es obvio qué queremos decir con que tres verdades conocidas son equivalentes. Así pues, antes de seguir, conviene que aclaremos a qué nos referimos cuando afirmamos que (8.33), (8.34) y (8.35) son equivalentes.

Lo que se persigue con todo esto es llegar a (8.35), el teorema de los números primos, que es el resultado realmente importante. Examinando una tabla de primos, este hecho fue conjeturado por Adrien Marie Legendre (1798) y Carl Friedrich Gauss (1849), pero ninguno de los dos tuvo éxito en probarlo. En 1851, Chebyshev dio un paso importante, pues logró demostrar que, si el límite que aparece en (8.35) existía, su valor era 1; además, encontró cotas superiores e inferiores de $\pi(x)\log(x)/x$, pero no pudo establecer que el límite realmente existiera.

Por otra parte, Leonhard Euler, en 1737, había relacionado los números primos con la función $\zeta(s) = \sum_{n=1}^{\infty} n^{-s}$ con s un número real mayor que 1. Bernhard Riemann, en 1859, tuvo la idea de usar valores complejos de s y dio un ingenioso método analítico que conectaba la distribución de los primos con las propiedades de $\zeta(s)$ (en particular, de dónde se encontraban sus ceros). Esto iba a ser un impulso fundamental, pero las matemáticas necesarias para llevar a buen término los proyectos de Riemann no estaban aún lo suficientemente desarrolladas cuando murió, prematuramente, en 1866.

En los años siguientes, las herramientas del análisis matemático fueron ampliándose. Así, en 1896 —de manera independiente y casi simultáneamente—Jacques

Hadamard y Charles de la Vallée-Poussin consiguieron demostrar (8.35). Esto fue, sin duda, un gran éxito de la teoría analítica de números.

Como ya hemos comentado, esta demostración del teorema de los números primos utilizaba técnicas analíticas avanzadas, y muy alejadas de lo que, en principio, se quería abordar. En particular, usa variable compleja, prolongación analítica, integrales sobre contornos... ¡para llegar a un resultado sobre números primos! Muchos matemáticos estaban interesados en encontrar otra demostración esencialmente distinta, basada en herramientas más cercanas a la teoría de números, lo que se llamaba una demostración *elemental*. Se pensaba que tal demostración iba a necesitar tal cúmulo de nuevas ideas que sería revolucionaria para toda la teoría de números. Nótese que, en este contexto, «demostración elemental» se usa como contrapartida a «demostración analítica»; no quiere decir que sea más fácil, sino que usa herramientas menos avanzadas.

Aquí entran en juego (8.33) y (8.34), pues desde cualquiera de estos dos hechos se logra pasar al teorema de los números primos sin usar teoría de variable compleja. Luego eran, en algún sentido, «elementalmente equivalentes» al teorema de los números primos.

Acabemos la historia diciendo que, finalmente, en 1949, Atle Selberg y Pál Erdős sorprendieron a la comunidad matemática con una demostración elemental, que no usaba la función $\zeta(s)$ ni variable compleja.[6] En teoría, esta demostración es accesible a cualquiera que esté familiarizado con el cálculo elemental. En la práctica, la demostración es bastante intrincada y, desde luego, no produjo la revolución que se pronosticaba.

Ahora que ya existen tanto demostraciones analíticas como elementales, se puede decir que, gran parte de las veces que se alude a «resultados equivalentes al teorema de los números primos», se hace como consecuencia del bagaje histórico-cultural acumulado durante tantos años de persecución del teorema.

8.7.1. Las funciones de Chebyshev

Vamos a comenzar probando la equivalencia de (8.34) con el teorema de los números primos; para ello introduciremos las denominadas *funciones de Chebyshev*.

[6] Selberg y Erdős presentaron sus demostraciones por separado, aunque previamente habían colaborado en el tema (y la demostración de Erdős se basaba en un resultado de Selberg); finalmente, acabaron discutiendo sobre a quién de los dos se debía atribuir la primacía de una demostración elemental. En todo caso, los resultados de Selberg eran cruciales en ambas pruebas, gracias a lo cual se hizo acreedor del mayor galardón que puede recibir un matemático: recibió la Medalla Fields en 1950 (si el lector quiere conocer más detalles sobre las Medallas Fields, puede encontrarlos en la página 113, en la nota al pie numerada como 7).

Tal como hizo Chebyshev en 1948, usaremos $\psi(x)$ para designar a las sumas parciales de la función de von Mangoldt, es decir,

$$\psi(x) = \sum_{n\leq x} \Lambda(n), \qquad x \geq 1.$$

Así, (8.34) puede expresarse mediante

$$\lim_{x\to\infty} \frac{\psi(x)}{x} = 1. \tag{8.36}$$

Como $\Lambda(n) = 0$ salvo cuando n es la potencia de un primo (y entonces vale $\Lambda(p^m) = \log(p)$), tenemos

$$\psi(x) = \sum_{n\leq x} \Lambda(n) = \sum_{\substack{p,m \\ p^m\leq x}} \Lambda(p^m) = \sum_{m=1}^{\infty} \sum_{p\leq x^{1/m}} \log(p).$$

Nótese que, aunque hayamos escrito $\sum_{m=1}^{\infty}$, sólo hay un número finito de sumandos. De hecho, el sumatorio interno $\sum_{p\leq x^{1/m}}$ es vacío si $x^{1/m} < 2$ (o $x < 2^m$); es decir, cuando $m > \log_2(x)$. Por lo tanto,

$$\psi(x) = \sum_{m\leq \log_2(x)} \sum_{p\leq x^{1/m}} \log(p).$$

Chebyshev también introdujo la función $\vartheta(x)$, que definió como

$$\vartheta(x) = \sum_{p\leq x} \log(p), \qquad x \geq 1.$$

Con esta notación, podemos escribir

$$\psi(x) = \sum_{m\leq \log_2(x)} \vartheta(x^{1/m}). \tag{8.37}$$

El siguiente resultado relaciona los dos cocientes $\psi(x)/x$ y $\vartheta(x)/x$:

Proposición 8.26.
Para $x \geq 1$ se cumple

$$0 \leq \frac{\psi(x)}{x} - \frac{\vartheta(x)}{x} \leq \frac{\log^2(x)}{2\log(2)\sqrt{x}}.$$

Demostración. Por (8.37), se tiene

$$0 \leq \psi(x) - \vartheta(x) = \sum_{2 \leq m \leq \log_2(x)} \vartheta(x^{1/m}).$$

Además, de la definición de $\vartheta(x)$ se deduce que

$$\vartheta(x) \leq \sum_{n \leq x} \log(x) \leq x \log(x).$$

Con esto,

$$0 \leq \psi(x) - \vartheta(x) = \sum_{2 \leq m \leq \log_2(x)} \vartheta(x^{1/m}) \leq \sum_{2 \leq m \leq \log_2(x)} x^{1/m} \log(x^{1/m})$$
$$\leq \log_2(x) x^{1/2} \log(x^{1/2}) = \frac{\sqrt{x} \log^2(x)}{2 \log(2)},$$

de donde —dividiendo por x— se sigue el resultado. □

Es importante darse cuenta de que la proposición anterior implica que

$$\lim_{x \to \infty} \left(\frac{\psi(x)}{x} - \frac{\vartheta(x)}{x} \right) = 0.$$

En particular, si uno de los cocientes $\psi(x)/x$ o $\vartheta(x)/x$ tiene límite cuando $x \to \infty$, el otro también lo tendrá, y ambos límites deben coincidir.

Como consecuencia, establecer (8.34) (en su forma (8.36)) es equivalente a probar

$$\lim_{x \to \infty} \frac{\vartheta(x)}{x} = 1. \tag{8.38}$$

Como curiosidad, comentemos ahora que, cuando en el tema 10 de este libro probemos finalmente el teorema de los números primos, ésta será la caracterización que usemos (aunque, para mayor comodidad, daremos una demostración bastante autocontenida, sin apoyarnos en lo que ahora estamos viendo).

8.7.2. Equivalencia con las medias de la función de von Mangoldt

A continuación, usaremos la identidad de Abel para dar dos expresiones que relacionan las funciones $\vartheta(x)$ y $\pi(x)$:

Proposición 8.27.
Para $x \geq 2$, se cumple

$$\vartheta(x) = \pi(x)\log(x) - \int_2^x \frac{\pi(t)}{t}\,dt \tag{8.39}$$

y

$$\pi(x) = \frac{\vartheta(x)}{\log(x)} + \int_2^x \frac{\vartheta(t)}{t\log^2(t)}\,dt. \tag{8.40}$$

Demostración. Sea $a(n)$ la función

$$a(n) = \begin{cases} 1, & \text{si } n \text{ es primo,} \\ 0, & \text{si } n \text{ no es primo.} \end{cases}$$

Así, con la notación de la identidad de Abel se tiene

$$A(x) = \sum_{n\leq x} a(n) = \pi(x).$$

Además,

$$\vartheta(x) = \sum_{p\leq x} \log(p) = \sum_{1<n\leq x} a(n)\log(n).$$

Tomemos ahora $f(x) = \log(x)$ e $y = 1$. Entonces, la identidad de Abel nos muestra que

$$\vartheta(x) = \sum_{1<n\leq x} a(n)\log(n) = \pi(x)\log(x) - \pi(1)\log(1) - \int_1^x \frac{\pi(t)}{t}\,dt.$$

De aquí se sigue (8.39), ya que $\pi(t) = 0$ para $t < 2$.

Para probar (8.40), de nuevo vamos a aplicar la identidad de Abel, ahora con $\overline{a}(n) = a(n)\log(n)$, $f(x) = 1/\log(x)$ e $y = 3/2$. Como

$$\overline{A}(x) = \sum_{n\leq x} \overline{a}(n) = \vartheta(x)$$

y

$$\pi(x) = \sum_{3/2<n\leq x} \overline{a}(n)/\log(n),$$

esta vez obtenemos

$$\pi(x) = \frac{\vartheta(x)}{\log(x)} - \frac{\vartheta(3/2)}{\log(3/2)} + \int_{3/2}^x \frac{\vartheta(t)}{t\log^2(t)}\,dt.$$

Al ser $\vartheta(t) = 0$ para $t < 2$, esto lo podemos escribir en la forma (8.40). □

Ya estamos en condiciones de probar la equivalencia entre el teorema de los números primos y el hecho de que las medias de la función de von Mangoldt tienden a 1 cuando $x \to \infty$. En realidad, en lugar de usar (8.34), emplearemos las dos condiciones (8.36) y (8.38) que —tal como acabamos de demostrar— son a su vez equivalentes a (8.34).

Teorema 8.28.
Las siguientes relaciones son equivalentes:

$$\lim_{x\to\infty} \frac{\pi(x)}{x/\log(x)} = 1, \tag{8.41}$$

$$\lim_{x\to\infty} \frac{\vartheta(x)}{x} = 1, \tag{8.42}$$

$$\lim_{x\to\infty} \frac{\psi(x)}{x} = 1. \tag{8.43}$$

Demostración. Recordemos que la proposición 8.26 nos ha mostrado que «(8.42)⇔(8.43)». Entonces, sólo necesitamos comprobar «(8.41)⇔(8.43)». Además, escribamos (8.39) y (8.40) como

$$\frac{\vartheta(x)}{x} = \frac{\pi(x)}{x/\log(x)} - \frac{1}{x}\int_2^x \frac{\pi(t)}{t}\,dt \tag{8.44}$$

y

$$\frac{\pi(x)}{x/\log(x)} = \frac{\vartheta(x)}{x} + \frac{\log(x)}{x}\int_2^x \frac{\vartheta(t)}{t\log^2(t)}\,dt. \tag{8.45}$$

Con esto,

(8.41) ⇒ (8.42): Por (8.44), basta que veamos que (8.41) implica

$$\lim_{x\to\infty} \frac{1}{x}\int_2^x \frac{\pi(t)}{t}\,dt = 0. \tag{8.46}$$

De (8.41), $\pi(t)/t = O(1/\log(t))$, luego

$$\frac{1}{x}\int_2^x \frac{\pi(t)}{t}\,dt = O\left(\frac{1}{x}\int_2^x \frac{1}{\log(t)}\,dt\right).$$

Ahora bien,

$$\int_2^x \frac{1}{\log(t)}\,dt = \int_2^{\sqrt{x}} \frac{1}{\log(t)}\,dt + \int_{\sqrt{x}}^x \frac{1}{\log(t)}\,dt \le \frac{\sqrt{x}}{\log(2)} + \frac{x-\sqrt{x}}{\log(\sqrt{x})},$$

de donde, dividiendo por x y tomando límites cuando $x \to \infty$, se obtiene (8.46).

(8.42) ⇒ (8.41): Por (8.45), nos basta comprobar que (8.42) implica

$$\lim_{x\to\infty} \frac{\log(x)}{x} \int_2^x \frac{\vartheta(t)}{t\log^2(t)}\, dt = 0. \tag{8.47}$$

De (8.42), $\vartheta(t) = O(t)$, así que

$$\frac{\log(x)}{x} \int_2^x \frac{\vartheta(t)}{t\log^2(t)}\, dt = O\left(\frac{\log(x)}{x} \int_2^x \frac{dt}{\log^2(t)}\right).$$

Entonces, tras acotar

$$\int_2^x \frac{dt}{\log^2(t)} = \int_2^{\sqrt{x}} \frac{dt}{\log^2(t)} + \int_{\sqrt{x}}^x \frac{dt}{\log^2(t)} \leq \frac{\sqrt{x}}{\log^2(2)} + \frac{x-\sqrt{x}}{\log^2(\sqrt{x})},$$

es inmediato deducir (8.47). □

8.8. Equivalencia con las medias de la función de Möbius

Vamos por fin a probar la equivalencia entre (8.33) y el teorema de los números primos. Denotemos $M(x)$ a las sumas parciales de la función de Möbius, es decir,

$$M(x) = \sum_{n\leq x} \mu(n), \qquad x \geq 1$$

(en muchos sitios, a $M(x)$ se le denomina *función de Mertens*). Con esta notación, (8.33) se expresa como $\lim_{x\to\infty} M(x)/x = 0$.

De manera auxiliar, usaremos la función

$$H(x) = \sum_{n\leq x} \mu(n)\log(n), \qquad x \geq 1.$$

Resultará útil la siguiente proposición —de nuevo una sencilla consecuencia de la identidad de Abel—, en la que se muestra que, cuando $x \to \infty$, el comportamiento de $M(x)/x$ se puede asimilar al de $H(x)/(x\log(x))$.

Proposición 8.29.

Si $M(x)$ y $H(x)$ son las funciones que acabamos de definir, se cumple

$$\lim_{x\to\infty} \left(\frac{M(x)}{x} - \frac{H(x)}{x\log(x)}\right) = 0.$$

DEMOSTRACIÓN. En la identidad de Abel, tomemos $f(t) = \log(t)$, $a(n) = \mu(n)$ (luego $A(x) = M(x)$) e $y = 1$. Así, para $x > 1$ tenemos

$$H(x) = \sum_{n \le x} \mu(n)\log(n) = M(x)\log(x) - \int_1^x \frac{M(t)}{t}\,dt.$$

Escribamos esto como

$$\frac{M(x)}{x} - \frac{H(x)}{x\log(x)} = \frac{1}{x\log(x)} \int_1^x \frac{M(t)}{t}\,dt;$$

entonces, la proposición se reduce a comprobar que

$$\lim_{x\to\infty} \frac{1}{x\log(x)} \int_1^x \frac{M(t)}{t}\,dt = 0.$$

Pero esto es fácil ya que, a partir de la estimación obvia $|M(t)| \le t$, se obtiene

$$\left| \frac{1}{x\log(x)} \int_1^x \frac{M(t)}{t}\,dt \right| \le \frac{1}{x\log(x)} \int_1^x dt = \frac{1}{\log(x)},$$

que tiende a 0 cuando $x \to \infty$. □

Abordemos ya la demostración de la equivalencia entre (8.33) y (8.35). Por el teorema 8.28, ya sabemos que el teorema de los números primos equivale a que $\lim_{x\to\infty} \psi(x)/x = 1$. Entonces, lo que en realidad vamos a probar —y nos basta con ello— es

$$\lim_{x\to\infty} \frac{\psi(x)}{x} = 1 \;\Leftrightarrow\; \lim_{x\to\infty} \frac{M(x)}{x} = 0.$$

Como el proceso no es muy corto, lo vamos a hacer en dos partes (cada implicación en un teorema).

Teorema 8.30.
Supongamos probado que

$$\lim_{x\to\infty} \frac{\psi(x)}{x} = 1.$$

Entonces,

$$\lim_{x\to\infty} \frac{M(x)}{x} = 0.$$

DEMOSTRACIÓN. Por la proposición anterior, basta ver que $\lim_{x\to\infty} \psi(x)/x = 1$ permite deducir

$$\lim_{x\to\infty} \frac{H(x)}{x\log(x)} = 0;$$

es decir, que, para cada $\varepsilon > 0$, existe una constante b tal que

$$\frac{|H(x)|}{x\log(x)} < \varepsilon \quad \text{cuando } x \geq b. \tag{8.48}$$

Resultará importante el hecho de que la función $H(x)$ cumple

$$H(x) = -\sum_{n\leq x} \mu(n)\psi\left(\frac{x}{n}\right) \tag{8.49}$$

(véase el ejercicio 7.38, en el tema anterior).

Por hipótesis, dado $\varepsilon > 0$, existe una constante $a \geq 1$ tal que $|\psi(x)/x - 1| < \varepsilon/2$ para todo $x \geq a$. Por tanto,

$$|\psi(x) - x| < \frac{\varepsilon}{2}\, x \quad \text{siempre que } x \geq a. \tag{8.50}$$

Sin pérdida de generalidad, podemos suponer $\varepsilon < 1$.

Ante cualquier $x \geq a$, tomemos $y = \lfloor x/a \rfloor$, y separemos la parte derecha de (8.49) en dos sumatorios, $\sum_{n\leq y}$ y $\sum_{y<n\leq x}$. En el primero, al ser $n \leq y \leq x/a$, se debe cumplir $x/n \geq a$; entonces, por (8.50),

$$\left|\psi\left(\frac{x}{n}\right) - \frac{x}{n}\right| < \frac{\varepsilon}{2}\,\frac{x}{n}, \qquad n \leq y.$$

Sumando y restando x/n en cada término, descompongamos

$$\sum_{n\leq y} \mu(n)\psi\left(\frac{x}{n}\right) = x\sum_{n\leq y} \frac{\mu(n)}{n} + \sum_{n\leq y} \mu(n)\left(\psi\left(\frac{x}{n}\right) - \frac{x}{n}\right).$$

Ahora, apliquemos la desigualdad triangular, la cota (8.30), y estimemos $\sum_{n\leq y} 1/n$ por medio de la integral $\int_1^y (1/t)\,dt$ (también podríamos haber usado el apartado (a) de la proposición 8.10); así,

$$\left|\sum_{n\leq y} \mu(n)\psi\left(\frac{x}{n}\right)\right| \leq x\left|\sum_{n\leq y} \frac{\mu(n)}{n}\right| + \sum_{n\leq y}\left|\psi\left(\frac{x}{n}\right) - \frac{x}{n}\right|$$

$$\leq x + \frac{\varepsilon}{2}\sum_{n\leq y}\frac{x}{n} \leq x + \frac{\varepsilon}{2}\,x\left(1 + \int_1^y \frac{1}{t}\,dt\right)$$
$$\leq x + \frac{\varepsilon}{2}\,x(1+\log(y)) \leq x + \frac{\varepsilon}{2}\,x(1+\log(x)).$$

En el segundo sumatorio, fijémonos que $y \leq x/a < y+1$ y que $y \leq n+1$ (pues $y < n$ y ambos son números enteros). Entonces, $x/n \leq x/(y+1) < a$. Como ψ es una función creciente, la desigualdad $x/n < a$ implica que $\psi(x/n) \leq \psi(a)$. Por consiguiente,

$$\left|\sum_{y<n\leq x}\mu(n)\psi\left(\frac{x}{n}\right)\right| \leq \sum_{y<n\leq x}\psi(a) \leq x\psi(a).$$

Juntando las cotas de los dos sumatorios $\sum_{n\leq y}$ y $\sum_{y<n\leq x}$, obtenemos

$$|H(x)| \leq x + \frac{\varepsilon}{2}\,x(1+\log(x)) + x\psi(a) < (2+\psi(a))x + \frac{\varepsilon}{2}\,x\log(x)$$

y, por tanto,

$$\frac{|H(x)|}{x\log(x)} < \frac{2+\psi(a)}{\log(x)} + \frac{\varepsilon}{2}.$$

Para concluir, elijamos $b \geq a$ tal que $(2+\psi(a))/\log(x) < \varepsilon/2$ para todo $x \geq b$. De aquí se sigue (8.48). □

Antes de abordar el recíproco del teorema anterior, establezcamos previamente un lema.

Lema 8.31.

Sea $g(n) = \tau(n) - \log(n) - 2\gamma$, con γ la constante de Euler. Entonces, cuando $x \to \infty$, se cumple

$$\psi(x) = x - \sum_{cd\leq x}\mu(d)g(c) + O(1)$$

(como es habitual, $\sum_{cd\leq x}$ indica que la suma se extiende a todos los pares de enteros positivos c y d tales que $cd \leq x$; no lo recordaremos más en lo que sigue).

Demostración. Partamos de las identidades

$$\lfloor x\rfloor = \sum_{n\leq x} 1, \qquad \psi(x) = \sum_{n\leq x}\Lambda(n), \qquad 1 = \sum_{n\leq x}\left\lfloor\frac{1}{n}\right\rfloor.$$

En ellas sustituyamos, respectivamente,

$$1 = \sum_{d|n}\mu(d)\tau\left(\frac{n}{d}\right), \qquad 1 = \sum_{d|n}\mu(d)\log\left(\frac{n}{d}\right), \qquad \left\lfloor\frac{1}{n}\right\rfloor = \sum_{d|n}\mu(d).$$

Así, obtenemos (con $c = n/d$)

$$\begin{aligned}\lfloor x\rfloor - \psi(x) - 2\gamma &= \sum_{n\le x}\left(1 - \Lambda(n) - 2\gamma\left\lfloor\frac{1}{n}\right\rfloor\right)\\ &= \sum_{n\le x}\sum_{d\mid n}\left(\tau\left(\frac{n}{d}\right) - \log\left(\frac{n}{d}\right) - 2\gamma\right)\\ &= \sum_{cd\le x}\mu(d)\big(\tau(c) - \log(c) - 2\gamma\big) = \sum_{cd\le x}\mu(d)g(c).\end{aligned}$$

Como $\lfloor x\rfloor = x + O(1)$, esto implica el resultado. □

Finalmente, he aquí el recíproco del teorema 8.30:

Teorema 8.32.
Supongamos probado que

$$\lim_{x\to\infty}\frac{M(x)}{x} = 0.$$

Entonces,

$$\lim_{x\to\infty}\frac{\psi(x)}{x} = 1.$$

Demostración. Por el lema, y siendo g la misma función que aparece allí, basta demostrar que la hipótesis $\lim_{x\to\infty} M(x)/x = 0$ permite deducir

$$\sum_{cd\le x}\mu(d)g(c) = o(x) \quad \text{cuando } x\to\infty$$

ya que, entonces, $\psi(x) = x + o(x)$, y esto prueba $\lim_{x\to\infty}\psi(x)/x = 1$. Más concretamente, lo que vamos a ver es que, para cualquier $\varepsilon > 0$, cuando x es suficientemente grande se verifica

$$\frac{1}{x}\left|\sum_{cd\le x}\mu(d)g(c)\right| < \varepsilon. \tag{8.51}$$

Comencemos aplicando el teorema 7.28 con $f = \mu$ y g la función que hemos denotado así en el lema. De este modo si, para $x \ge 1$, tomamos $G(x) = \sum_{n\le x} g(n)$, y dos reales $r, s \in [1,\infty)$ con $rs = x$, se tiene

$$\sum_{cd\le x}\mu(d)g(c) = \sum_{n\le s}\mu(n)G\left(\frac{x}{n}\right) + \sum_{n\le r} g(n)M\left(\frac{x}{n}\right) - M(s)G(r). \tag{8.52}$$

Estimemos la función G. El teorema 8.12 nos asegura que, cuando $x \to \infty$,

$$\sum_{n\le x} \tau(n) = x\log(x) + (2\gamma - 1)x + O(\sqrt{x}).$$

Además, el apartado (e) de la proposición 8.10 nos mostraba

$$\sum_{n\le x} \log(n) = \log(\lfloor x\rfloor !) = x\log(x) - x + O(\log(x)).$$

Con estas dos estimaciones,

$$\begin{aligned} G(x) &= \sum_{n\le x} g(n) = \sum_{n\le x} \tau(n) - \sum_{n\le x} \log(n) - 2\gamma \sum_{n\le x} 1 \\ &= x\log(x) + (2\gamma-1)x + O(\sqrt{x}) - \big(x\log(x) - x + O(\log(x))\big) - 2\gamma x + O(1) \\ &= O(\sqrt{x}) + O(\log(x) + O(1) = O(\sqrt{x}). \end{aligned}$$

En consecuencia, existe una constante A tal que $G(x) \le A\sqrt{x}$ para todo $x \ge 1$ (cuando x es mayor que un cierto a, esto es cierto por definición de «O»; y, para $x \in [1, a]$, la desigualdad es obvia, pues G es acotada sobre cualquier intervalo acotado). Si aplicamos la desigualdad $G(x) \le A\sqrt{x}$ para estimar el primer sumatorio de la derecha en (8.52), obtenemos

$$\left|\sum_{n\le s} \mu(n) G\left(\frac{x}{n}\right)\right| \le A\sum_{n\le s} \sqrt{\frac{x}{n}} = Ax^{1/2} \sum_{n\le s} n^{-1/2} \le Bx^{1/2}s^{1/2} = \frac{Bx}{\sqrt{r}}$$

para alguna constante B (la suma $\sum_{n\le s} n^{-1/2}$ se puede acotar mediante una integral o usando la proposición 8.10); por conveniencia, podemos imponer $B \ge A$. Con esto, y dado $\varepsilon > 0$, elijamos $r \ge 1$ tal que $B/\sqrt{r} < \varepsilon/3$ (obsérvese que r depende de ε, pero no de x). En estas condiciones,

$$\left|\sum_{n\le s} \mu(n) G\left(\frac{x}{n}\right)\right| < \frac{\varepsilon}{3}\, x \quad \text{para todo } x \ge 1. \tag{8.53}$$

Vamos ahora a intentar acotar el segundo sumatorio de la derecha en (8.52). Apliquemos la hipótesis de que $M(x) = o(x)$. Así, dado el mismo ε que antes y un cierto número positivo K (que en seguida precisaremos), existe C (dependiente sólo de ε) tal que $|M(x)|/x < \varepsilon/K$ cuando $x \ge C$. Por consiguiente, si suponemos que $x/n \ge C$ para todo $n \le r$ (o sea, exijamos $x \ge rC$), podremos poner

$$\left|\sum_{n\le r} g(n) M\left(\frac{x}{n}\right)\right| < \sum_{n\le r} |g(n)|\, \frac{\varepsilon}{K}\, \frac{x}{n} = \frac{\varepsilon x}{K} \sum_{n\le r} \frac{|g(n)|}{n}.$$

Tomemos ahora $K = 3\sum_{n\le r} |g(n)|/n$. Entonces,

$$\left| \sum_{n\le r} g(n) M\left(\frac{x}{n}\right) \right| < \frac{\varepsilon}{3}\, x \quad \text{siempre que } x \ge rC. \tag{8.54}$$

Por fin, acotemos el último término de la derecha de (8.52). Como $A \le B$ y hemos elegido r de modo que $B/\sqrt{r} < \varepsilon/3$, el sumando $M(s)G(r)$ está dominado por

$$|M(s)G(r)| \le A\sqrt{rs} \le B\sqrt{rs} < \frac{\varepsilon}{3}\, rs = \frac{\varepsilon}{3}\, x. \tag{8.55}$$

Combinando (8.53), (8.54) y (8.55), obtenemos

$$\left| \sum_{cd\le x} \mu(d) g(c) \right| < \frac{\varepsilon}{3}\, x + \frac{\varepsilon}{3}\, x + \frac{\varepsilon}{3}\, x = \varepsilon x,$$

desigualdad válida siempre que $x \ge rC$. Recordemos una vez más que r era cualquier número tal que $B/\sqrt{r} < \varepsilon/3$, es decir, $r > 9B^2/\varepsilon^2$. Así pues, hemos probado que todos los $x > 9B^2C/\varepsilon^2$ satisfacen (8.51), tal como buscábamos. □

8.9. Ejercicios

8.1. Prueba (8.1) con $\ell \in \mathbb{R}$, y busca un contraejemplo para el recíproco.

8.2. Prueba (8.1) con $\ell = \pm\infty$, y busca un contraejemplo para el recíproco.

8.3. Sea una sucesión $\{a_n\}_{n=1}^{\infty}$ tal que $(a_1 + a_2 + \cdots + a_n)/n^\alpha$ tiene límite (finito). Comprueba que $a_n = o(n^\alpha)$.

8.4. Sea una sucesión $\{a_n\}_{n=1}^{\infty}$ tal que $|(a_1 + a_2 + \cdots + a_n)/n^\alpha|$ está acotada. Comprueba que $a_n = O(n^\alpha)$.

8.5. Sea f una función multiplicativa tal que, para cada primo p, verifica $f(p^m) \to 0$ cuando $m \to \infty$.[7] ¿Puedo garantizar que $f(n) \to 0$ cuando $n \to \infty$?

8.6. Para $x \ge 1$, sea $A(x) = \sum_{n\le x} a(n)$, donde $a(n)$ es una función aritmética. Demuestra que

$$\sum_{n\le x} (x-n)a(n) = \int_1^x A(t)\, dt.$$

[7] Compárese este enunciado con el del lema 8.1.

8.7. En este ejercicio y los que siguen vamos a ver algunas formulaciones alternativas para la constante de Euler. Con ese fin, y como paso previo, comprueba que, para cualquier $n \in \mathbb{N}$,

$$\int_0^1 \frac{1-(1-x)^n}{x}\,dx = H_n$$

con $H_n = \sum_{k=1}^n 1/k$. Utiliza esa expresión para demostrar que

$$\gamma = \int_0^1 \frac{1-e^{-x}}{x}\,dx - \int_1^\infty \frac{e^{-x}}{x}\,dx.$$

Sugerencia. En la integral que da H_n, haz el cambio de variable $x \mapsto x/n$ y manipula convenientemente lo que se obtiene.

8.8. Prueba que

$$\int_0^x \frac{1-e^{-t}}{t}\,dt - \int_x^\infty \frac{e^{-t}}{t}\,dt - \log(x) = \gamma$$

para cualquier $x > 0$.

8.9. Prueba que

$$\int_0^\infty e^{-x}\log(x)\,dx = -\gamma.$$

¿Cómo se relaciona esa integral con la función Gamma de Euler que hemos visto en la nota 8.9 (página 468)?
Sugerencia. Divide el intervalo $(0,\infty)$ en dos trozos e intenta aprovechar la fórmula del ejercicio 8.7.

8.10. Comprueba que, si intentamos estimar $\sum_{n\le x}\tau(n)$ procediendo como en la demostración del teorema 8.11, lo que se consigue —en lugar de una prueba del teorema 8.12— es (8.23).

8.11. Dada una función aritmética f, se llama serie de Dirichlet de f a la función

$$F(s) = \sum_{n=1}^\infty \frac{f(n)}{n^s}, \quad s \in \mathbb{C}.$$

Comprueba que, si $F(s)$ converge absolutamente para un $c \in \mathbb{R}$, también converge absolutamente para todo $s \in \mathbb{C}$ tal que $\mathrm{Re}(s) \ge c$.

8.12. Dadas f y g funciones aritméticas, sean

$$F(s) = \sum_{n=1}^{\infty} \frac{f(n)}{n^s}, \quad \mathrm{Re}(s) > a,$$
$$G(s) = \sum_{n=1}^{\infty} \frac{g(n)}{n^s}, \quad \mathrm{Re}(s) > b,$$

sus series de Dirichlet, convergentes absolutamente en la región indicada. Asimismo, sea $h = f * g$ el producto de Dirichlet de f y g. Demuestra que, en estas circunstancias, la serie de Dirichlet de h es

$$H(s) = \sum_{n=1}^{\infty} \frac{h(n)}{n^s} = \sum_{n=1}^{\infty} \frac{(f*g)(n)}{n^s} = F(s)G(s),$$

y además converge absolutamente en el semiplano $\mathrm{Re}(s) > \max\{a, b\}$.

8.13. Demuestra que, cuando $\mathrm{Re}(s) > 1$, se cumple

$$\sum_{n=1}^{\infty} \frac{\mu(n)}{n^s} = \frac{1}{\zeta(s)}.$$

8.14. Demuestra que, cuando $\mathrm{Re}(s) > 1$, se cumple

$$\sum_{n=1}^{\infty} \frac{\tau(n)}{n^s} = \zeta(s)^2.$$

8.15. Demuestra que, cuando $\mathrm{Re}(s) > 2$, se cumple

$$\sum_{n=1}^{\infty} \frac{\varphi(n)}{n^s} = \frac{\zeta(s-1)}{\zeta(s)}.$$

8.16. Demuestra que, cuando $\mathrm{Re}(s) > \max\{1, 1 + \mathrm{Re}(a)\}$, se cumple

$$\sum_{n=1}^{\infty} \frac{\sigma_a(n)}{n^s} = \zeta(s)\zeta(s-a).$$

8.17. Demuestra que, cuando $\mathrm{Re}(s) > 1$, se cumple

$$\sum_{n=1}^{\infty} \frac{\lambda(n)}{n^s} = \frac{\zeta(2s)}{\zeta(s)}.$$

8.18. Sea f completamente multiplicativa, y supón que $F(s) = \sum_{n=1}^{\infty} f(n)n^{-s}$ converge absolutamente para $\mathrm{Re}(s) > c$. Demuestra que $\sum_{n=1}^{\infty} \mu(n)f(n)n^{-s}$ también converge absolutamente para $\mathrm{Re}(s) > c$ y, además,

$$\sum_{n=1}^{\infty} \frac{\mu(n)f(n)}{n^s} = \frac{1}{F(s)}.$$

8.19. Demuestra que, cuando $\mathrm{Re}(s) > 1$, se cumple

$$\sum_{n=1}^{\infty} \frac{|\mu(n)|}{n^s} = \frac{\zeta(s)}{\zeta(2s)}.$$

8.20. Comprueba que la aplicación (8.29) que se usa en la demostración del teorema 8.21 está bien definida y es, realmente, una biyección.

8.21. Demuestra que, cuando $x \to \infty$, se tiene la fórmula asintótica

$$\sum_{n\leq x} \frac{\log(n)}{n} = \frac{1}{2}\log^2(x) + A + O\left(\frac{\log(x)}{x}\right)$$

para cierta constante A que debes identificar.

8.22. Demuestra las siguiente fórmula asintótica (válida cuando $x \to \infty$) y halla la constante B:

$$\sum_{2\leq n\leq x} \frac{1}{n\log(n)} = \log(\log(x)) + B + O\left(\frac{1}{x\log(x)}\right).$$

8.23. Demuestra que, cuando $x \to \infty$, se verifican las siguientes fórmulas asintóticas:

$$\sum_{n\leq x} \frac{\tau(n)}{n} = \frac{1}{2}\log^2(x) + 2\gamma\log(x) + O(1);$$

y, si $s > 0$, $s \neq 1$,

$$\sum_{n\leq x} \frac{\tau(n)}{n^s} = \frac{1}{1-s}x^{1-s}\log(x) + \zeta(s)^2 + O(x^{1-s}).$$

Observación. Ten en cuenta que éste es un ejercicio largo y que para resolverlo deberás aplicar cuidadosamente algunas de las técnicas estudiadas en este tema, así como apoyarte en bastantes resultados ya vistos. En particular, te resultará útil el teorema 8.12 y, bajo ciertas circunstancias del parámetro s, también el ejercicio 8.14.

8.24. Sea $s \in \mathbb{C}$ con $\operatorname{Re}(s) > 1$. Prueba que, cuando $x \to \infty$, se verifica

$$\sum_{n \le x} \frac{\mu(n)}{n^s} = \frac{1}{\zeta(s)} + O(x^{-\operatorname{Re}(s)+1}).$$

8.25. Sea un entero $k \ge 2$. A un natural n se le llama k-libre si no existe ningún primo p tal que $p^k \mid n$ (cuando $k = 2$, esto se suele expresar diciendo que «n es libre de cuadrados»). Se define $\mu_k(n) = 1$ si n es k-libre y $\mu_k(n) = 0$ si no lo es. Prueba que, cuando $x \to \infty$, se cumple

$$\sum_{n \le x} \mu_k(n) = \frac{x}{\zeta(k)} + O(x^{1/k})$$

(es decir, la densidad asintótica de los números k-libres es $1/\zeta(k)$).

Tema 9
Distribución de los números primos: resultados elementales

9.1. Productos de Euler
9.2. Teorema de Chebyshev
9.3. Postulado de Bertrand
9.4. Fórmulas de Mertens
9.5. El número de divisores primos de un entero
9.6. Ejercicios

A estas alturas del libro no hay ninguna necesidad de explicar que, por definición, un entero mayor que 1 es primo cuando sólo es divisible por 1 y por sí mismo. Los primos son los ladrillos que sostienen y construyen la estructura multiplicativa del edificio de los números enteros, y ya desde el primer tema del libro nos hemos topado con ellos por doquier. No en vano el teorema fundamental de la aritmética afirma que cualquier entero mayor que 1 se puede descomponer de manera única como producto de primos.

Disponemos de un método eficaz para ir calculando, en orden, todos los números primos: tras tomar todos los enteros positivos menores que uno dado, la criba de Eratóstenes es un sencillo proceso de eliminación —seguro que el lector lo recuerda perfectamente— que deja al final sólo los números que son primos. Esto permite encontrar sin dificultad tablas de primos; en el primer tema de este libro ya hemos mostrado una: véase el cuadro 1.1, en la página 25.

Con la necesaria paciencia, estas tablas pueden ser todo lo grandes que queramos, pues es bien conocido que existen infinitos números primos. La primera demostración de este hecho ya aparece en los *Elementos* de Euclides, y, de nuevo, es muy sencilla. Supongamos que sólo hubiera una cantidad finita de primos $p_1, p_2, \ldots, p_n$, y sea $N = p_1 p_2 \cdots p_n + 1$. Por una parte, ningún p_k divide a N, luego N debe ser primo. Por otra, N es mayor que todos los p_k, luego no puede ser ninguno de ellos. Y esto es absurdo, pues partíamos de que habíamos tomado todos los primos.

Pero, ¿son los primos muy numerosos entre los enteros? En otras palabras, si $\pi(x)$ es el número de primos menores o iguales que x, ¿cuánto vale $\pi(x)$? Obviamente, interesaría poder dar la respuesta con un procedimiento más eficiente que tener que computar explícitamente todos los primos hasta x (además, conviene ahora precisar que la criba de Eratóstenes es eficaz, pero forzosamente lenta si x es grande). Si no somos capaces de proporcionar respuestas exactas, ¿podemos conocer $\pi(x)$ de manera aproximada? Al estudiar la distribución de los números primos, el resultado central es —sin duda— el teorema de los números primos. Aunque ya lo hemos comentado en varios temas, recordemos que este teorema afirma que

$$\lim_{x\to\infty} \frac{\pi(x)}{x/\log(x)} = 1. \tag{9.1}$$

Su demostración no es fácil, y la postergamos al tema 10.

Ahora nos dedicaremos a mostrar diversos resultados elementales relacionados con la distribución de primos. En este sentido, con «elemental» queremos decir que, en las demostraciones que aquí aparezcan, no vamos a utilizar teoría de funciones analíticas, ni tampoco —por supuesto— el teorema de los números primos. Así, varios de los resultados de este tema son consecuencias inmediatas de (9.1). Pero

las demostraciones que aquí daremos no lo usan y, además, requieren mucha menos herramienta matemática.

Hasta que lleguemos a la sección 9.4, este tema no utiliza lo estudiado en otras partes del libro (únicamente requeriremos los conceptos de función aritmética multiplicativa y completamente multiplicativa, pero lo recordaremos en su momento). A partir de entonces, se necesitarán algunos conocimientos —no demasiados— de lo que aparece en los dos temas anteriores. Si el lector ha optado por saltárselos le recomendamos que, al menos, hojee las definiciones del principio, y se familiarice con la notación. Un resultado que se usará a menudo es la proposición 8.10, que muestra algunas fórmulas sobre estimaciones asintóticas de funciones. En cuanto a la notación, recordemos que p siempre denota un número primo, y que $\sum_p$ (o $\prod_p$) indica que la suma (o el producto) se efectúa sobre números primos.

El tipo de estudio que hacemos aquí se puede encontrar —de manera más o menos extensa— en multitud de libros de teoría de números. Por ejemplo, [3, capítulo 10], [41, capítulo 5], [50, capítulo 22], [54, capítulo 5], [71, capítulo 8], [87, capítulo 12], [95, capítulo 3]; también merece la pena destacar [5], aunque la organización del material difiere bastante de la que hemos seguido. Y, hablando de primos, es obligatorio citar [82], que trata todo tipo de temas relacionados con números primos de una forma muy amena.

No queremos despedir esta introducción sin comentar lo mucho que queda aún por saber sobre los números primos. Existen numerosos problemas de enunciado sencillo, y planteados hace ya mucho tiempo, cuya respuesta no se conoce. Por ejemplo, en una conferencia dada en el Congreso Internacional de Matemáticos de 1912, Edmund Landau mencionó cuatro conjeturas antiguas:

- Cada número par mayor que dos es suma de dos primos (conjetura de Goldbach).
- Entre cada dos cuadrados consecutivos siempre existe un número primo (conjetura de Legendre).
- Existen infinitos primos gemelos (esto es, números de la forma p y $p+2$, ambos primos).
- Existen infinitos primos de la forma n^2+1.

En nuestros días, los cuatro problemas continúan abiertos.

9.1. Productos de Euler

Por aproximarnos al desarrollo histórico —aunque sin seguirlo fielmente—, comencemos el tema mostrando un interesante resultado que muestra la abundancia de los números primos, y que fue descubierto en 1737 por el genial matemático suizo



Nicolás de Oresme
(1323–1382)



Sir Isaac Newton
(1643–1727)



Gottfried Wilhelm von
Leibniz (1646–1716)

Leonhard Euler. Los razonamientos de Euler sobre lo que ahora vamos a tratar se pueden ver en [43], edición traducida y comentada de su *Introductio in Analysin Infinitorum* de 1748.

A estas alturas, todos sabemos que $\sum_{n=1}^{\infty} 1/n = \infty$. En realidad, ahora solemos probarlo comparando la serie $\sum_{n=1}^{\infty} 1/n$ con la integral $\int_1^{\infty} (1/x)\,dx$, que es divergente. Pero la demostración original de Nicolás de Oresme —realizada hacia 1360, mucho antes de que Isaac Newton y Gottfried von Leibniz inventaran el cálculo infinitesimal— no requiere tales herramientas, sino que sólo necesita agrupar los términos de la serie de una manera adecuada:

$$\begin{aligned}\sum_{n=1}^{\infty} \frac{1}{n} &= 1 + \frac{1}{2} + \left(\frac{1}{3} + \frac{1}{4}\right) + \left(\frac{1}{5} + \frac{1}{6} + \frac{1}{7} + \frac{1}{8}\right) + \cdots \\ &\geq 1 + \frac{1}{2} + \left(\frac{1}{4} + \frac{1}{4}\right) + \left(\frac{1}{8} + \frac{1}{8} + \frac{1}{8} + \frac{1}{8}\right) + \cdots \\ &= 1 + \frac{1}{2} + \frac{1}{2} + \frac{1}{2} + \cdots,\end{aligned}$$

que es claramente divergente.[1]

Euler se preguntó cuánto sumaba la serie de los inversos de los números primos. Obviamente, hay menos primos que números naturales. Pero, ¿son los primos suficientemente numerosos para que $\sum_p 1/p$ sea divergente? O, por el contrario, ¿es

[1] Curiosamente, la demostración de Nicolás de Oresme fue olvidada y, en el siglo XVII, los hermanos Jacob y Johann Bernoulli se enzarzaron en la búsqueda del carácter de la serie $\sum_{n=1}^{\infty} 1/n$. Fue Johann —el hermano menor— quien logró probar de nuevo la divergencia de esa serie. Su demostración, esencialmente distinta de la de Nicolás de Oresme (pero bastante más complicada y difícil de justificar rigurosamente), apareció en un libro de Jacob publicado en 1689, quien reconocía la prioridad de su hermano. La historia completa puede leerse en [38, capítulo 8].

esa serie convergente? Si éste fuera el caso, podríamos decir que los primos son asintóticamente más numerosos que los cuadrados, puesto que $\sum_n 1/n^2 < \infty$.

En el camino para dilucidar esta cuestión, Euler probó el siguiente teorema que, aunque lo vamos a enunciar de manera bastante general, en seguida nos restringiremos al caso particular que nos va a interesar. Pero queremos recalcar que éste es un resultado muy útil, hasta el punto que de él se suele decir que es la versión analítica del teorema fundamental de la aritmética. La argumentación que vamos a usar es esencialmente la de Euler, aunque adaptada al rigor y la notación actuales. Recordemos que, por definición, una función aritmética (esto es, $f\colon \mathbb{N} \to \mathbb{R}$ o $\mathbb{C}$) es multiplicativa si $f(nm) = f(n)f(m)$ siempre que $\mathrm{mcd}(n,m) = 1$; y es completamente multiplicativa si $f(nm) = f(n)f(m)$ para todo $n, m \in \mathbb{N}$.

Teorema 9.1. Identidad de Euler.
Sea f una función multiplicativa tal que la serie $\sum_n f(n)$ es absolutamente convergente. Entonces,

$$\sum_{n=1}^{\infty} f(n) = \prod_p \left(1 + f(p) + f(p^2) + \cdots\right). \tag{9.2}$$

Si f es completamente multiplicativa, se tiene, además,

$$\sum_{n=1}^{\infty} f(n) = \prod_p \frac{1}{1 - f(p)}. \tag{9.3}$$

En ambos casos, al producto se le denomina producto de Euler de la serie.

Demostración. Comencemos mostrando cómo se deduce (9.3) de (9.2). No hay más que darse cuenta de que, si f es completamente multiplicativa, entonces $f(p^k) = f(p)^k$ y, por tanto, para llegar a (9.3) basta sumar una serie geométrica:

$$1 + f(p) + f(p^2) + \cdots = 1 + f(p) + f(p)^2 + \cdots = \frac{1}{1 - f(p)}.$$

Así, ya sólo tenemos que ocuparnos de (9.2).

Para cada número real $x \geq 2$, consideremos el producto finito

$$P(x) = \prod_{p \leq x} \left(1 + f(p) + f(p^2) + \cdots\right), \tag{9.4}$$

donde cada factor es una serie. Como $\sum_n f(n)$ es absolutamente convergente, en (9.4) podemos multiplicar y reordenar términos con toda libertad, sin que cambie el valor de $P(x)$. Cada uno de estos términos tiene la forma

$$f(p_1^{a_1})f(p_2^{a_2})\cdots f(p_k^{a_k}) = f(p_1^{a_1}p_2^{a_2}\cdots p_k^{a_k}),$$

ya que f es multiplicativa. Por tanto, si tomamos

$$A_x = \{\, n \in \mathbb{N} : \text{todos los factores primos de } n \text{ son} \le x \,\}$$

podemos escribir $P(x) = \sum_{n\in A_x} f(n)$ y

$$\sum_{n=1}^{\infty} f(n) - P(x) = \sum_{n\in B_x} f(n),$$

donde $B_x = \mathbb{N} \setminus A_x$ es el conjunto de los enteros positivos que tienen algún factor primo mayor que x. En consecuencia,

$$\left|\sum_{n=1}^{\infty} f(n) - P(x)\right| = \left|\sum_{n\in B_x} f(n)\right| \le \sum_{n\in B_x} |f(n)| \le \sum_{n>x} |f(n)|.$$

Como, por hipótesis, $\sum_{n=1}^{\infty} |f(n)|$ es convergente, estamos en condiciones de asegurar que $\sum_{n>x} |f(n)| \to 0$ cuando $x \to \infty$. Luego $P(x) \to \sum_{n=1}^{\infty} f(n)$ cuando $x \to \infty$, de donde se sigue (9.2). □

9.1.1. La serie de los inversos de los números primos

Euler aplicó la fórmula (9.3) a la función $f(n) = n^{-s}$, con $s > 1$, que es completamente multiplicativa, obteniendo así una expresión en forma de producto infinito para la serie

$$\zeta(s) = \sum_{n=1}^{\infty} \frac{1}{n^s}, \quad s > 1,$$

función que ahora se conoce con el nombre de zeta de Riemann.[2] En concreto, si usamos la identidad (9.3), lo que obtenemos es

$$\zeta(s) = \sum_{n=1}^{\infty} \frac{1}{n^s} = \prod_p \frac{1}{1-1/p^s} = \prod_p \left(1 - \frac{1}{p^s}\right)^{-1}, \quad s > 1.$$

[2] En las secciones 6.9 y 8.3 se dan bastantes más detalles sobre esta función. Pero, por otra parte, conviene aquí señalar que los argumentos que ahora vamos a mostrar fueron el origen de la fructífera relación entre $\zeta(s)$ y los números primos.

Pero Euler no se contentó con eso, sino que lo utilizó para llegar a una atractiva consecuencia. Tomando logaritmos, la igualdad anterior queda

$$\log\left(\sum_{n=1}^{\infty}\frac{1}{n^s}\right) = -\sum_{p}\log\left(1-\frac{1}{p^s}\right), \quad s>1.$$

Por otra parte, argumentos sencillos de análisis matemático prueban que $0 \leq -\log(1-x) \leq x+x^2$ cuando $0 \leq x \leq 1/2$; así, podemos decir que $-\log(1-x) = x + R(x)$ para cierta función $R(x)$ tal que $|R(x)| \leq x^2$ cuando $0 \leq x \leq 1/2$. Escribamos ahora

$$\log\left(\sum_{n=1}^{\infty}\frac{1}{n^s}\right) = \sum_{p}\frac{1}{p^s} + \sum_{p} R\left(\frac{1}{p^s}\right) \tag{9.5}$$

y hagamos tender $s \to 1$ en esta expresión. El último sumando converge sin ningún problema, ya que

$$\sum_{p}|R(p^{-s})| \leq \sum_{p} p^{-2s} \leq \sum_{m=1}^{\infty} m^{-2s} < \infty$$

para cualquier $s > 1/2$. En cambio, $\sum_{n=1}^{\infty} n^{-s} \to \infty$ cuando $s \to 1$, pues la serie armónica $\sum_{n=1}^{\infty} n^{-1}$ diverge. Así, la única manera de que la igualdad (9.5) siga siendo cierta cuando $s \to 1$ es que $\sum_p p^{-1} = \infty$.

Por lo tanto, hemos probado lo siguiente:

Teorema 9.2. Euler, 1737.
La serie de los inversos de los primos es divergente, es decir,

$$\sum_{p}\frac{1}{p} = \infty.$$

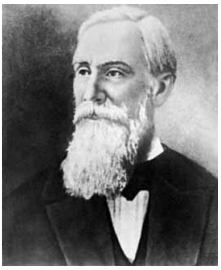
En particular, este teorema proporciona ¡una nueva demostración de la existencia de infinitos primos!

9.1.2. La identidad de Euler aplicada a las series de Dirichlet

Concluyamos mencionando que el teorema 9.1 tiene una atractiva aplicación a las series de Dirichlet de funciones aritméticas, a las que ya hemos aludido en el



Leonhard Euler
(1707–1783)



Pafnuty Lvovich
Chebyshev (1821–1894)

apartado 8.5.2 de la página 482 (véanse también los ejercicios 8.11 a 8.19). Por definición, una serie de Dirichlet es

$$\sum_{n=1}^{\infty} \frac{f(n)}{n^s}$$

con f una función aritmética y s perteneciente a una región del plano complejo. Sin más que aplicar la identidad de Euler, las expresiones sobre series de Dirichlet de funciones aritméticas se traducen en sus correspondientes productos de Euler. Se anima al lector a que lo compruebe por sí mismo (ejercicios 9.3 y 9.4).

9.2. Teorema de Chebyshev

El objetivo de esta sección es probar un resultado más débil que el teorema de los números primos, pero de demostración mucho más sencilla. Comencemos dando una expresión para la descomposición de $n!$ en factores primos mediante una fórmula que habitualmente se conoce como identidad de Legendre, pero que en ocasiones también se alude a ella con el nombre de fórmula de De Polignac.[3]

[3] Alphonse de Polignac fue un matemático francés del siglo XIX. A él se debe la conjetura de De Polignac, que propuso en 1849 y afirma que todo número par es igual a la diferencia de dos números primos consecutivos de infinitas formas (o, de manera equivalente, que para todo entero par n existe una infinidad de primos consecutivos cuya diferencia vale n). Del mismo modo que ocurre con la existencia de infinitos primos gemelos (que es un caso particular), la conjetura de De Polignac no ha sido probada ni refutada.

Proposición 9.3. Identidad de Legendre.
Para cada entero positivo n, se tiene

$$n! = \prod_{p \le n} p^{\alpha(n,p)},$$

con

$$\alpha(n,p) = \sum_{m=1}^{\infty} \left\lfloor \frac{n}{p^m} \right\rfloor = \sum_{m \le \frac{\log(n)}{\log(p)}} \left\lfloor \frac{n}{p^m} \right\rfloor. \tag{9.6}$$

Demostración. Es obvio que $n!$ tiene la forma $n! = \prod_{p \le n} p^{\alpha(n,p)}$; lo único que tenemos que hacer es identificar los exponentes $\alpha(n,p)$. Lo haremos de dos maneras distintas. Pero, antes, notemos que los dos sumatorios que aparecen en (9.6) son, en realidad, el mismo, pues $\lfloor n/p^m \rfloor = 0$ cuando $p^m > n$, esto es, cuando $m > \log(n)/\log(p)$.

Primera prueba. Teniendo en cuenta la descomposición $n! = 1 \cdot 2 \cdots n$, cada divisor primo p de $n!$ (trivialmente, debe ser $p \le n$) tiene un exponente que se calcula como sigue:

- Cada factor de la forma kp, con $1 \le k \le \lfloor n/p \rfloor$ para que $kp \le n$, contribuye con un factor p.
- Cada kp^2, con $1 \le k \le \lfloor n/p^2 \rfloor$, contribuye con otro factor p más, aparte del que ya está contado antes.
- Cada kp^3, con $1 \le k \le \lfloor n/p^3 \rfloor$, contribuye con otro más, aparte de los anteriores.

Y así sucesivamente. En total, son $\sum_{m=1}^{\infty} \lfloor n/p^m \rfloor$.

Segunda prueba. Vamos ahora a usar la identidad

$$\log(n!) = \sum_{k \le n} \Lambda(k) \left\lfloor \frac{n}{k} \right\rfloor,$$

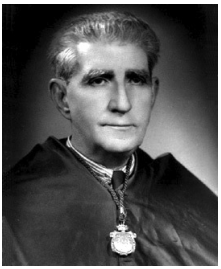
que ya probamos en el teorema 7.27. Con esto,

$$\begin{aligned} \log(n!) &= \sum_{k \le n} \Lambda(k) \left\lfloor \frac{n}{k} \right\rfloor = \sum_{\substack{p,m \\ p^m \le n}} \log(p) \left\lfloor \frac{n}{p^m} \right\rfloor \\ &= \sum_{p \le n} \left(\sum_{p^m \le n} \left\lfloor \frac{n}{p^m} \right\rfloor \right) \log(p) = \sum_{p \le n} \alpha(n,p) \log(p) = \log\left(\prod_{p \le n} p^{\alpha(n,p)} \right), \end{aligned}$$

de donde se sigue la fórmula buscada. □



Julio Rey Pastor
(1888–1962)



José Barinaga
(1890–1965)

Antes de proseguir, vamos a desviarnos un poco de nuestro camino mostrando un ejemplo en el que se aplica la identidad de Legendre. En realidad —y aunque somos plenamente conscientes de la universalidad de las matemáticas— estamos a punto de «cometer un localismo», pues vamos a hacer intervenir a dos de los no muy numerosos matemáticos españoles de la primera mitad del siglo XX. Lo que sigue es una cuestión sobre divisibilidad propuesta por Julio Rey Pastor en la *Revista de la Sociedad Matemática Española*, y cuya solución, debida a José Barinaga —que la obtuvo siendo aún alumno de la Facultad de Ciencias— apareció publicada en el número de noviembre de 1911.

Ejemplo. Demostrar que $n^k!$ es divisible por $(n!)^{(n^k-1)/(n-1)}$, donde n y k son dos enteros positivos, con $n > 1$.

SOLUCIÓN. Por la identidad de Legendre, tenemos

$$n! = \prod_{p\le n} p^{\alpha(n,p)}, \qquad n^k! = \prod_{p\le n} p^{\alpha(n^k,p)}$$

con

$$\alpha(n,p) = \sum_{m=1}^{\infty} \left\lfloor \frac{n}{p^m} \right\rfloor, \qquad \alpha(n^k,p) = \sum_{m=1}^{\infty} \left\lfloor \frac{n^k}{p^m} \right\rfloor.$$

La afirmación $(n!)^{(n^k-1)/(n-1)} \mid n^k!$ quedará demostrada si comprobamos que, para cada primo p divisor de n, se cumple

$$\alpha(n,p)\,\frac{n^k-1}{n-1} \le \alpha(n^k,p). \tag{9.7}$$

Si a y b son los enteros que hacen $p^a \le n < p^{a+1}$ y $p^b \le n^k < p^{b+1}$, podemos escribir

$$\alpha(n,p) = \sum_{m=1}^{a} \left\lfloor \frac{n}{p^m} \right\rfloor, \qquad \alpha(n^k,p) = \sum_{m=1}^{b} \left\lfloor \frac{n^k}{p^m} \right\rfloor.$$

Tomemos $c = \lfloor b/a \rfloor$ y descompongamos

$$\begin{aligned}\alpha(n^k,p) = \sum_{m=1}^{b} \left\lfloor \frac{n^k}{p^m} \right\rfloor &\ge \sum_{m=1}^{a} \left\lfloor \frac{n^k}{p^m} \right\rfloor + \sum_{m=1}^{a} \left\lfloor \frac{n^k}{p^a p^m} \right\rfloor \\ &\quad + \sum_{m=1}^{a} \left\lfloor \frac{n^k}{p^{2a} p^m} \right\rfloor + \cdots + \sum_{m=1}^{a} \left\lfloor \frac{n^k}{p^{(c-1)a} p^m} \right\rfloor \\ &\ge \sum_{m=1}^{a} \left\lfloor \frac{n^k}{p^m} \right\rfloor + \sum_{m=1}^{a} \left\lfloor \frac{n^{k-1}}{p^m} \right\rfloor + \sum_{m=1}^{a} \left\lfloor \frac{n^{k-2}}{p^m} \right\rfloor + \cdots + \sum_{m=1}^{a} \left\lfloor \frac{n}{p^m} \right\rfloor.\end{aligned}$$

Por otra parte, teniendo en cuenta que $(n^k-1)/(n-1) = n^{k-1}+n^{k-2}+n^{k-3}+\cdots+1$ y la desigualdad $\lfloor r/s \rfloor t \le \lfloor rt/s \rfloor$ (válida para r, s, t enteros positivos arbitrarios), tenemos

$$\begin{aligned}\alpha(n,p)\,\frac{n^k-1}{n-1} &= \sum_{m=1}^{a} \left\lfloor \frac{n}{p^m} \right\rfloor n^{k-1} + \sum_{m=1}^{a} \left\lfloor \frac{n}{p^m} \right\rfloor n^{k-2} \\ &\quad + \sum_{m=1}^{a} \left\lfloor \frac{n}{p^m} \right\rfloor n^{k-3} + \cdots + \sum_{m=1}^{a} \left\lfloor \frac{n}{p^m} \right\rfloor \cdot 1 \\ &\le \sum_{m=1}^{a} \left\lfloor \frac{n^k}{p^m} \right\rfloor + \sum_{m=1}^{a} \left\lfloor \frac{n^{k-1}}{p^m} \right\rfloor + \sum_{m=1}^{a} \left\lfloor \frac{n^{k-2}}{p^m} \right\rfloor + \cdots + \sum_{m=1}^{a} \left\lfloor \frac{n}{p^m} \right\rfloor,\end{aligned}$$

de donde se sigue (9.7). ◊

9.2.1. Cinco lemas y el teorema

Tras la leve digresión producida por el ejemplo, volvamos a nuestra ruta. En lo que aquí nos atañe, la identidad de Legendre nos va a servir para probar el siguiente resultado, que nos resultará útil un poco más adelante.

Lema 9.4.

Para cada $n \in \mathbb{N}$,

$$\binom{2n}{n} \quad \textit{divide a} \quad \prod_{p \le 2n} p^{\left\lfloor \frac{\log(2n)}{\log(p)} \right\rfloor}.$$

Demostración. Según la identidad de Legendre,

$$(2n)! = \prod_{p \le 2n} p^{\alpha(2n,p)}$$

y

$$(n!)^2 = \prod_{p \le n} p^{2\alpha(n,p)} = \prod_{p \le 2n} p^{2\alpha(n,p)},$$

donde en la última igualdad hemos usado que $\alpha(n,p) = 0$ para $n < p \le 2n$. Por tanto,

$$\binom{2n}{n} = \frac{(2n)!}{n!\,n!} = \prod_{p \le 2n} p^{\alpha(2n,p) - 2\alpha(n,p)},$$

y vamos a acotar los exponentes $\alpha(2n,p) - 2\alpha(n,p)$. Notemos que, para cada $x \in \mathbb{R}$,

$$\lfloor 2x \rfloor - 2\lfloor x \rfloor = \begin{cases} 0, & \text{si } \lfloor x \rfloor \le x < \lfloor x \rfloor + 1/2, \\ 1, & \text{si } \lfloor x \rfloor + 1/2 \le x < \lfloor x \rfloor + 1; \end{cases}$$

en particular, $\lfloor 2x \rfloor - 2\lfloor x \rfloor \le 1$. Con esto,

$$\begin{aligned} \alpha(2n,p) - 2\alpha(n,p) &= \sum_{1 \le m \le \frac{\log(2n)}{\log(p)}} \left(\left\lfloor \frac{2n}{p^m} \right\rfloor - 2 \left\lfloor \frac{n}{p^m} \right\rfloor \right) \\ &\le \sum_{1 \le m \le \frac{\log(2n)}{\log(p)}} 1 = \left\lfloor \frac{\log(2n)}{\log(p)} \right\rfloor, \end{aligned}$$

lo que demuestra el lema. □

De momento, y en lo que a esta sección se refiere, lo único que nos interesará del lema anterior es que

$$\binom{2n}{n} \le \prod_{p \le 2n} p^{\left\lfloor \frac{\log(2n)}{\log(p)} \right\rfloor},$$

lo cual es una consecuencia inmediata; sí que usaremos la parte «divide a» cuando probemos el postulado de Bertrand. Veamos ahora otros cuatro lemas, todos ellos de comprobación sencilla.

Lema 9.5.

Para cada $n \in \mathbb{N}$, se cumple

$$2^n \le \binom{2n}{n}.$$

DEMOSTRACIÓN. Por inducción sobre n. Para $n = 1$ es cierto. Suponiéndolo cierto para n, comprobémoslo para $n+1$:

$$2^{n+1} \leq 2\binom{2n}{n} = 2\,\frac{(2n)!}{(n!)^2} = \frac{(2n+2)!}{(n+1)(2n+1)(n!)^2}$$
$$\leq \frac{(2n+2)!}{((n+1)!)^2} = \binom{2n+2}{n+1}. \qquad \square$$

Lema 9.6.
Para cada $n \in \mathbb{N}$,

$$\prod_{n+1<p\leq 2n+1} p \qquad \textit{divide a} \qquad \binom{2n+1}{n+1}.$$

DEMOSTRACIÓN. Cada primo p que cumple $n+1 < p \leq 2n+1$ divide a $(2n+1)!$, pero no divide a $(n+1)!$ ni a $n!$, luego divide a $\binom{2n+1}{n+1} = \frac{(2n+1)!}{(n+1)!\,n!}$. $\square$

Lema 9.7.
Para cada $n \in \mathbb{N}$, se cumple

$$\binom{2n+1}{n+1} \leq 4^n.$$

DEMOSTRACIÓN. Se puede probar por inducción sobre n. O bien del siguiente modo:

$$2 \cdot 4^n = 2^{2n+1} = (1+1)^{2n+1} \geq \binom{2n+1}{n} + \binom{2n+1}{n+1} = 2\binom{2n+1}{n+1},$$

donde la desigualdad que hemos usado resulta inmediata sin más que desarrollar $(1+1)^{2n+1}$ mediante el binomio de Newton. $\square$

Lema 9.8.
Para cada $n \in \mathbb{N}$, se cumple

$$\prod_{p\leq n} p \leq 4^n.$$

DEMOSTRACIÓN. Lo probaremos mediante inducción completa sobre n. Para $n = 1$ y $n = 2$ es cierto. Supongámoslo cierto para $1, 2, \ldots, n-1$ y probémoslo para n; para ello, distingamos dos casos, según sea n par o impar.

Si n es par (luego $n \geq 4$), tendremos $\prod_{p\leq n} p = \prod_{p\leq n-1} p \leq 4^{n-1} < 4^n$. Si n es impar, escribamos $n = 2m+1$. Así,

$$\prod_{p\leq n} p = \left(\prod_{p\leq m+1} p\right)\left(\prod_{m+1<p\leq n} p\right) \leq 4^{m+1}\binom{2m+1}{m+1} \leq 4^{2m+1} = 4^n,$$

donde en la primera desigualdad se ha empleado la hipótesis de inducción y el lema 9.6 y, en la segunda, el lema 9.7. □

Finalmente, aquí tenemos el teorema principal de esta sección: una demostración elemental de que el tamaño de $\pi(n)$ es, salvo constantes multiplicativas, como $n/\log(n)$, un resultado debido al matemático ruso Pafnuty Chebyshev. Con un esfuerzo mayor —y el mismo Chebyshev lo hizo—, se pueden mejorar las constantes concretas que nos van a aparecer en el teorema, pero no nos vamos a preocupar de ello.

Dar buenas constantes y, sobre todo, aproximarlas a 1 todo lo que se pudiera para valores de n suficientemente grandes, era un reto importante hasta finales del siglo XIX, pero no ahora, puesto que ese tipo de estimaciones se pueden obtener como consecuencia inmediata del teorema de los números primos (que fue probado en 1896). Dado que aquí no nos vamos a apoyar en el teorema de los números primos, el interés actual de esto radica en la sencillez de los argumentos empleados —que no son los originales de Chebyshev—. Conviene mencionar que Chebyshev fue uno de los matemáticos que con más ahínco intentaron probar —sin éxito— el teorema de los números primos. En su camino, logró demostrar varios resultados parciales, siendo éste uno de ellos; aunque nosotros no lo veremos, también probó que si el límite que aparece en (9.1) existía, su valor tenía que ser 1.

Teorema 9.9. Chebyshev, 1852.
Existen constantes $C_1, C_2 > 0$ tales que, para todo $n \geq 2$,

$$C_1 \frac{n}{\log(n)} \leq \pi(n) \leq C_2 \frac{n}{\log(n)}.$$

En particular, sirven las constantes

$$C_1 = \frac{\log(2)}{3} \quad y \quad C_2 = \frac{2}{e} + 4\log(2).$$

DEMOSTRACIÓN. Veamos primero la estimación inferior. Usando, sucesivamente, los lemas 9.5 y 9.4, tenemos (recuerde el lector que, por definición, $a^x = e^{x\log(a)}$)

$$2^n \le \binom{2n}{n} \le \prod_{p\le 2n} p^{\left\lfloor \frac{\log(2n)}{\log(p)} \right\rfloor} \le \prod_{p\le 2n} p^{\frac{\log(2n)}{\log(p)}} = \prod_{p\le 2n} (2n) = (2n)^{\pi(2n)},$$

luego

$$\pi(2n) \ge \frac{n\log(2)}{\log(2n)} = \frac{\log(2)}{2}\cdot\frac{2n}{\log(2n)}.$$

Para argumentos impares,

$$\begin{aligned}\pi(2n+1) \ge \pi(2n) &\ge \frac{\log(2)}{2}\cdot\frac{2n}{\log(2n)} \ge \frac{\log(2)}{2}\cdot\frac{2n}{2n+1}\cdot\frac{2n+1}{\log(2n+1)}\\ &\ge \frac{\log(2)}{3}\cdot\frac{2n+1}{\log(2n+1)}.\end{aligned}$$

Por tanto, basta tomar, por ejemplo, $C_1 = \log(2)/3$.

A continuación, demostremos la estimación superior. Para cada $n \ge 2$, podemos poner

$$(\sqrt{n})^{\pi(n)-\pi(\sqrt{n})} \le \prod_{\sqrt{n}<p\le n} p \le \prod_{p\le n} p \le 4^n; \tag{9.8}$$

aquí, en la primera desigualdad hemos usado $\sqrt{n} < p$ y que el número de factores es $\pi(n) - \pi(\sqrt{n})$; y, en la última, el lema 9.8. Ahora, descompongamos $\pi(n) = \pi(\sqrt{n}) + (\pi(n) - \pi(\sqrt{n}))$. El primer sumando lo acotamos con $\pi(\sqrt{n}) \le \sqrt{n}$; y, el segundo, con la desigualdad que queda al tomar logaritmos en (9.8), es decir, con

$$\pi(n) - \pi(\sqrt{n}) \le \frac{n\log(4)}{\frac{1}{2}\log(n)}.$$

Además, para $x \ge 1$, se verifica $\log(x)/\sqrt{x} \le 2/e$ (véase el ejercicio 9.9). Con todo esto,

$$\pi(n) \le \sqrt{n} + \frac{n\log(4)}{\frac{1}{2}\log(n)} \le \sqrt{n}\cdot\frac{2}{e}\frac{\sqrt{n}}{\log(n)} + \frac{4n\log(2)}{\log(n)} = C_2\frac{n}{\log(n)}$$

con $C_2 = 2/e + 4\log(2)$. □

9.2.2. Cotas para el tamaño del k-ésimo primo

Seguidamente, aplicaremos el teorema 9.9 para obtener cotas superiores e inferiores del tamaño del k-ésimo número primo. Obviamente, sustituyendo las constantes C_1 y C_2 que allí aparecen por valores concretos, es fácil encontrar expresiones «más bonitas» que la que mostramos en (9.9). Pero hemos preferido dejarlo así, ya que esto permite observar muy bien cómo mejoras en las constantes C_1 y C_2 —y ya hemos dicho que esto es posible— conducen a mejoras en la estimación del tamaño del primo k-ésimo.

Teorema 9.10.
Para $k \geq 1$, el k-ésimo primo p_k verifica

$$\frac{1}{C_2}\, k \log(k) < p_k < \frac{2}{C_1}\, k \left(\log(k) + \frac{1}{C_1} \right). \tag{9.9}$$

donde C_1 y C_2 son las dos constantes del teorema 9.9.

Demostración. Sea $n = p_k$, luego $k = \pi(n)$ y, además, $n \geq 2$. De la estimación superior del teorema 9.9 obtenemos

$$k = \pi(n) \leq C_2 \frac{n}{\log(n)} = C_2 \frac{p_k}{\log(p_k)},$$

de donde se sigue la cota inferior de (9.9):

$$p_k \geq \frac{1}{C_2}\, k \log(p_k) > \frac{1}{C_2}\, k \log(k).$$

Obtengamos ahora la cota superior. La estimación inferior del teorema 9.9 nos permite asegurar

$$k = \pi(n) \geq C_1 \frac{n}{\log(n)} = C_1 \frac{p_k}{\log(p_k)},$$

y por tanto

$$p_k \leq \frac{1}{C_1}\, k \log(p_k). \tag{9.10}$$

Para $x > 1$, se cumple $\log(x) \leq (2/e)\sqrt{x}$ (y la igualdad sólo se alcanza con $x = e^2$, tal como vemos en el ejercicio 9.9), así que

$$p_k \leq \frac{1}{C_1}\, k \log(p_k) < \frac{2}{C_1 e}\, k \sqrt{p_k}.$$

Cambiando $\sqrt{p_k}$ de miembro, queda $\sqrt{p_k} < 2k/(C_1 e)$ y, tomando logaritmos,

$$\frac{1}{2}\log(p_k) < \log(k) + \log\left(\frac{2}{C_1 e}\right).$$

Ahora, sustituyendo en (9.10) obtenemos

$$p_k < \frac{2}{C_1}\, k \left(\log(k) + \log\left(\frac{2}{C_1 e}\right)\right).$$

Finalmente (usando de nuevo $\log(x) \le (2/e)\sqrt{x}$ y $\sqrt{x} \le x$),

$$\log\left(\frac{2}{C_1 e}\right) \le \frac{2}{e}\sqrt{\frac{2}{C_1 e}} \le \frac{4}{C_1 e^2} \le \frac{1}{C_1}$$

y aparece la cota superior buscada. □

Como consecuencia inmediata, la serie $\sum_{k=1}^{\infty} 1/p_k$ tiene el mismo carácter que $\sum_{k=2}^{\infty} 1/(k\log(k))$, que es divergente. Así, hemos obtenido otra demostración de que la serie de los inversos de los primos es divergente, el conocido resultado de Euler que ya habíamos probado en el teorema 9.2.

Nota 9.11. Las dos series $\sum_{n=1}^{\infty} 1/n$ y $\sum_{k=1}^{\infty} 1/p_k$ son divergentes. En cierto sentido, podemos decir que la primera es «más divergente» que la segunda, pues tiene «muchos más» sumandos. Reflexionando un poco más nos damos cuenta de que $\sum_{k=1}^{\infty} 1/p_k$ está ya muy cerca —mucho más que $\sum_{n=1}^{\infty} 1/n$— de la sutil frontera que separa las series convergentes de las divergentes. Y es que $\sum_{k=2}^{\infty} 1/(k\log(k)) = \infty$, pero

$$\sum_{k=2}^{\infty} \frac{1}{k\log^{1+\varepsilon}(k)} < \infty$$

para cualquier $\varepsilon > 0$ (en los ejercicios 9.10 y 9.11 se propone al lector un análisis más profundo sobre todo esto). Existe otra manera de cuantificar si una serie es «muy divergente» o «divergente pero no mucho», y es acudir al tamaño de sus sumas parciales. Esto requiere algún esfuerzo adicional y —en particular en cuanto a la serie de los inversos de los números primos se refiere— no lo veremos hasta la sección 9.4 (véase la nota 9.19). ◁

9.3. Postulado de Bertrand

El objetivo principal de esta sección es probar el postulado de Bertrand y algunas de sus consecuencias. Comentaremos también —ya sin demostración— algunos otros resultados más finos, aunque no tan sencillos. Pero de momento —y para no perder

la costumbre— comencemos con un lema (obsérvese que éste que ahora mostramos es más fuerte que el lema 9.5 que veíamos antes).

Lema 9.12.
Para cada $n \in \mathbb{N}$,

$$\frac{4^n}{2n} \leq \binom{2n}{n}.$$

Demostración. De todos los coeficientes que aparecen en el desarrollo del binomio de Newton, el mayor de todos es el del centro (o los dos centrales, si el exponente del binomio es impar, lo cual da un número par de sumandos); eso es claro a la vista del triángulo de Pascal, no obstante, una demostración rigurosa se puede ver en el ejercicio 1.35 (tema 1). En particular, con exponente $2n$ se tiene que $\binom{2n}{k} \leq \binom{2n}{n}$ para $0 \leq k \leq n$. Con esto,

$$\begin{aligned} 4^n = (1+1)^{2n} &= \sum_{k=0}^{2n} \binom{2n}{k} = 2 + \sum_{k=1}^{2n-1} \binom{2n}{k} \\ &\leq \binom{2n}{n} + (2n-1)\binom{2n}{n} = 2n\binom{2n}{n}. \end{aligned}$$

□

Y otro más:

Lema 9.13.
Si $n \geq 3$ y $2n/3 < p \leq n$, entonces

$$p \quad \textit{no divide a} \quad \binom{2n}{n}.$$

Demostración. Utilizaremos la expresión

$$\binom{2n}{n} = \frac{(n+1)(n+2)\cdots(n+n)}{1 \cdot 2 \cdots n}.$$

Ahora, distingamos dos posibilidades:

- Si p no divide a ningún $n+k$ $(1 \leq k \leq n)$, entonces el lema se sigue de manera inmediata. (Para analizar este caso con más detalle, véase el ejercicio 9.12.)
- En otro caso, supongamos que $n+k = mp$ para ciertos números k y m. Como $mp = n+k \leq 2n < 3p$, debe ser $m < 3$. Como $p \leq n < n+k = mp$, debe ser $1 < m$. Es decir, la única posibilidad es $m = 2$ y $n+k = 2p$. Entonces, p divide a $(n+1)(n+2)\cdots(n+n)$ una sola vez. Pero también divide a $1 \cdot 2 \cdots n$, luego no divide a $\binom{2n}{n}$. □

En 1845, observando una gran cantidad de datos reunidos en tablas, el matemático francés Joseph Bertrand se dio cuenta de que, aparentemente, entre cada entero positivo y su doble siempre había un número primo. Fue Chebyshev —alrededor de 1850— quien logró demostrar este hecho, que actualmente se conoce como postulado de Bertrand. La prueba que presentamos aquí es de Erdős (1932).

Teorema 9.14. Postulado de Bertrand.
Para cada $n \in \mathbb{N}$, existe algún número primo p tal que $n < p \leq 2n$.

Demostración. Para $n = 1$ y $n = 2$ el resultado es cierto, así que podemos suponer $n \geq 3$. Vamos a efectuar la demostración por reducción al absurdo. Así pues, supongamos que no existe ningún primo p con $n < p \leq 2n$.

Por el lema 9.13, cualquier divisor primo p de $\binom{2n}{n}$ debe cumplir $p \leq 2n/3$ o $n < p \leq 2n$. En consecuencia, de nuestro supuesto se deduce que todos los divisores primos de $\binom{2n}{n}$ deben ser $p \leq 2n/3$.

Además, en el lema 9.4 vimos que $\binom{2n}{n}$ divide a $\prod_{p\leq 2n} p^{\left\lfloor \frac{\log(2n)}{\log(p)} \right\rfloor}$. Con lo cual, si $a \geq 2$ y p^a divide a $\binom{2n}{n}$, se tiene

$$p^2 \leq p^a \leq p^{\left\lfloor \frac{\log(2n)}{\log(p)} \right\rfloor} \leq p^{\frac{\log(2n)}{\log(p)}} = 2n.$$

En particular, $p \leq \sqrt{2n}$ y $p^a \leq 2n$.

Conviene recalcar que, bajo la suposición de que no exista un primo entre n y $2n$, hemos probado dos cosas:

- Todos los factores primos p de $\binom{2n}{n}$ satisfacen $p \leq 2n/3$.
- Todos los factores primos p de $\binom{2n}{n}$ tales que, además, p^a (con $a \geq 2$) divide a $\binom{2n}{n}$, cumplen
$$p \leq \sqrt{2n} \quad \text{y} \quad p^a \leq 2n.$$

De aquí,

$$\binom{2n}{n} \leq \prod_{\substack{p\leq 2n/3 \\ p\mid \binom{2n}{n}}} p \cdot \prod_{\substack{p\leq \sqrt{2n} \\ p\mid \binom{2n}{n}}} p^a \leq \left(\prod_{p\leq 2n/3} p \right) \cdot (2n)^{\sqrt{2n}}, \tag{9.11}$$

donde en el segundo producto de la expresión central estarán incluidos todos los factores primos no simples, contados con su multiplicidad a. (Obsérvese que algunos factores primos de $\binom{2n}{n}$ aparecen contados en los dos $\prod$. Además, al efectuar el segundo «$\leq$» estamos incluyendo muchos primos p que no son factores de $\binom{2n}{n}$. Pero

no importa, pues la cota superior que hemos encontrado nos va a ser útil, pese a que seguro que hay cotas más ajustadas.)

Vamos a acotar (9.11) aún más. En concreto, usaremos la cota inferior $4^n/(2n) \le \binom{2n}{n}$ y la cota superior $\prod_{p\le 2n/3} p = \prod_{p\le \lfloor 2n/3\rfloor} p \le 4^{2n/3}$ de los lemas 9.12 y 9.8, respectivamente. Entonces, empalmando con (9.11), deberá ser

$$\frac{4^n}{2n} \le 4^{2n/3}(2n)^{\sqrt{2n}};$$

de aquí, $4^{n/3} \le (2n)^{1+\sqrt{2n}}$ y, tomando logaritmos, $\frac{2\log 2}{3}\, n \le (1+\sqrt{2n})\log(2n)$.

En resumen, hemos probado que, si no existe un primo p tal que $n < p \le 2n$, entonces

$$\log(2) \le \frac{3}{(2n)}(1+\sqrt{2n})\log(2n) = 3\left(\frac{1}{\sqrt{2n}}+1\right)\frac{\log(2n)}{\sqrt{2n}}. \tag{9.12}$$

Si esto no es cierto, ya hemos llegado a un absurdo. Y, en efecto, así es; al menos a partir de un n, como vemos a continuación. La desigualdad (9.12) no se cumple para $n = 2^9 = 512$:

$$3\left(\frac{1}{\sqrt{2n}}+1\right)\frac{\log(2n)}{\sqrt{2n}} = 3\left(\frac{1}{32}+1\right)\frac{10\log(2)}{32} = \frac{990}{1024}\log(2) < \log(2).$$

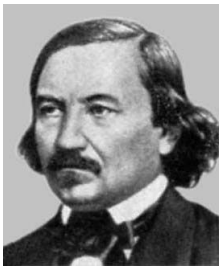
Y, como la función $\frac{\log(x)}{\sqrt{x}}$ es decreciente en el intervalo $[e^2,\infty)$ (ejercicio 9.9), deducimos que (9.12) tampoco se cumple para ningún $n \ge 512$.

Así pues, hemos probado el postulado de Bertrand para todo $n \ge 512$. Nos falta ver que también es cierto para $n < 512$. Y esto lo prueban los primos

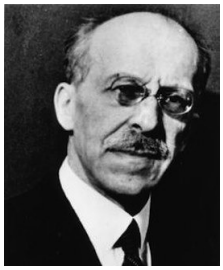
$$2,\ 3,\ 5,\ 7,\ 13,\ 23,\ 43,\ 83,\ 163,\ 317,\ 631,$$

elegidos de manera que el doble de cada uno de ellos es mayor que el siguiente. Si tomamos un $n \le 512$, sea q el mayor primo de la lista menor o igual que n, y p el siguiente primo en la lista (luego $p > n$). Como $p < 2q$, tendremos $q \le n < p < 2q \le 2n$. □

El intervalo $n < p \le 2n$ en el que el postulado de Bertrand asegura la existencia de un primo se puede acortar si no tenemos en cuenta los primeros números, aunque no es sencillo hacerlo si queremos concretar cuáles son los primeros números que aceptamos descartar. A este respecto, y prácticamente como curiosidad, citaremos los tres resultados siguientes, que supusieron sucesivas mejoras:



Joseph Bertrand
(1822–1900)



Issai Schur (1875–1941)



Emiliano Aparicio
(1926–1998)

- Para $n \geq 24$, siempre existe un primo p que cumple $n < p \leq \frac{5}{4}n$ (Issai Schur, 1929).
- Para $n \geq 48$, siempre existe un primo p que cumple $n < p \leq \frac{9}{8}n$ (Robert Breusch, 1930).
- Para $n \geq 115$, siempre existe un primo p que cumple $n < p \leq \frac{10}{9}n$ (Emiliano Aparicio, 1951).

9.3.1. Distancias entre primos

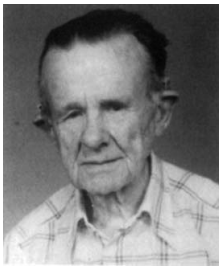
Del mismo modo que ya hemos hecho otras veces, utilicemos $\{p_k\}_{k=1}^{\infty}$ para denotar la sucesión de todos los números primos. Una manera equivalente de enunciar el postulado de Bertrand es decir que el primo p_{k+1} siempre está en el intervalo abierto $(p_k, 2p_k)$; o, en otras palabras, que $p_{k+1} - p_k < p_k$. Este resultado —y lo mismo podemos decir de las mejoras de Schur, Breusch y Aparicio antes citadas— es, asintóticamente, muy pobre, y se conocen algunos otros sustancialmente mejores, aunque de demostración indudablemente más complicada. Guido Hoheisel (1930) y Albert E. Ingham (1937) probaron que existe una constante K tal que la diferencia entre dos primos consecutivos es

$$p_{k+1} - p_k < K p_k^{5/8}$$

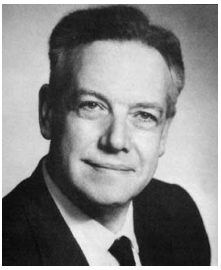
para todo k. Posteriormente, diversos matemáticos han logrado rebajar el exponente, mejorando así el resultado. Por ejemplo, en 1986, Charles J. Mozzochi demostró que

$$p_{k+1} - p_k < K p_k^{1051/1920},$$

y no ha sido ésta la última mejora, aunque, para lo que aquí perseguimos, no merece la pena dar más detalles. El margen de mejora es aún mucho, pues, esencialmente,



Harald Cramér (1893–1985)



Edward Maintland Wright (1906–2005)



Ludvig Oppermann (1817–1883)

el mayor obstáculo conocido para seguir bajando es

$$\limsup_k \frac{p_{k+1}-p_k}{\log(p_k)} = \infty \tag{9.13}$$

(lo probó Erik Westzynthius en 1931).

Por otra parte, Harald Cramér probó en 1920 que si la hipótesis de Riemann[4] fuese cierta, entonces se tendría

$$p_{k+1} - p_k < K p_k^{1/2} \log(p_k).$$

Por supuesto, desconocemos si se cumple o no esta cota, pero pone de manifiesto la importancia de la función zeta de Riemann en las propiedades de la distribución de los números primos. Se piensa que la distancia entre primos consecutivos debe ser incluso menor; algo así como $p_{k+1} - p_k \leq (1 + o(1)) \log^2(k)$ (que es compatible con (9.13)) o, quizás, un resultado algo más débil: $p_{k+1} - p_k = O(\log^2(k))$, que se denomina conjetura de Cramér (la propuso en 1936). Todas estas conjeturas —así como muchas otras fruto de la fascinación que los matemáticos sienten por los números primos— permanecen abiertas.

En particular, una llamativa conjetura relacionada con las diferencias entre primos consecutivos es la que propuso Dorin Andrica en 1985. La conjetura de Andrica afirma que

$$\sqrt{p_{k+1}} - \sqrt{p_k} < 1 \tag{9.14}$$

para todo $k \geq 1$ (véase además el ejercicio 9.15). Pese a los muchos cálculos realiza-

[4] Ya hemos aludido a ella en la nota 8.9 de la página 468, y veremos más detalles en el apartado 10.6.2 (página 607 y siguientes).

dos con ayuda de ordenadores, la mayor diferencia $\sqrt{p_{k+1}} - \sqrt{p_k}$ encontrada hasta ahora es $\sqrt{11}-\sqrt{7} = 0.670\,873\ldots$, lo cual hace confiar en que la conjetura sea cierta.

Hay, de hecho, conjeturas sobre distribución de números primos que llevan bastante más tiempo abiertas. Merece la pena mencionar las conjeturas de Oppermann (propuesta en 1882) y de Brocard (de principios de siglo XX). La conjetura de Oppermann recuerda a la de Legendre que ya hemos comentado en la introducción de este tema («siempre hay un primo entre n^2 y $(n+1)^2$»), y afirma que, para cada entero $n > 1$, siempre hay un primo entre $n^2 - n$ y n^2, y otro entre n^2 y $n^2 + n$. A su vez, la conjetura de Brocard afirma que entre p_k^2 y p_{k+1}^2 siempre hay, al menos, cuatro números primos (salvo para $k = 1$). En el ejercicio 9.16 proponemos al lector que compruebe la relación entre estas tres conjeturas.

9.3.2. Funciones generadoras de primos

Hasta ahora, hemos encontrado algunas expresiones que nos dan información sobre el k-ésimo primo. Pero, si lo que queremos es conocerlo exactamente, lo que se obtiene de ellas dista de ser satisfactorio —y esto es así incluso aunque las conjeturas mencionadas estuvieran realmente probadas—, pues sólo informan del tamaño aproximado, y con un término de error considerable. Pero los matemáticos han sido mucho más ambiciosos. Así, en su intento por doblegar la irregular distribución de los números primos, han tratado de encontrar —sin éxito— fórmulas simples que generen todos los números primos. O, en su defecto, que siempre den lugar a números primos, aunque no sean todos.

Algunas expresiones proporcionan muchos números primos. Por ejemplo, Euler observó en 1772 que el denominado trinomio de Euler $P_E(x) = x^2 + x + 41$ da primos distintos para $x = 0, 1, 2, \ldots, 39$ (en 1798, Lagrange propuso el polinomio $P_L(x) = x^2 - x + 41$, que da esos mismos primos para $x = 1, 2, \ldots, 40$, y además también es primo para $x = 0$, aunque este último coincide con el de $x = 1$). Asimismo, el polinomio

$$P_D(x) = x^2 - 79x + 1601 = (x - 40)^2 + (x - 40) + 41,$$

proporciona números primos cuando $x = 0, 1, 2, \ldots, 79$, los mismos primos que con el trinomio de Euler pero cada uno de ellos dos veces. (Si nos fijamos, lo que sucede con $P_L(x)$ y $P_D(x)$ se reduce a sencillas transformaciones sobre el resultado de Euler: $P_E(x) = x(x + 1) + 41$ y $P_L(x) = x(x - 1) + 41 = P_E(x - 1)$, así que $P_L(x)$ es primo para $x = 0, 1, \ldots, 40$; además, $P_E(x) = P_L(-x)$, y se sigue que $P_E(x)$ es primo para $x = -40, -39, \ldots, 39$; finalmente, $P_D(x) = P_E(x - 40)$, luego $P_D(x)$ deberá ser primo para $x = 0, 1, 2, \ldots, 79$.)

Pero no existe ningún polinomio $P \in \mathbb{Z}[x]$ no constante tal que $P(n)$ sea primo para todo $n \in \mathbb{N}$. Este hecho, del que se dio cuenta Christian Goldbach[5] en 1752, es relativamente sencillo: véanse los ejercicios 9.20 y 9.21, en los que se propone al lector que lo demuestre.

Puesto que no es posible hacerlo mediante polinomios, hay que resignarse a ser mucho más modestos, y nos conformaremos con hallar otro tipo de expresiones que siempre produzcan números primos; es lo que se denomina *funciones generadoras de primos*. El postulado de Bertrand nos va a servir para dar la que mostramos a continuación, ideada por el matemático inglés Edward Maintland Wright.

Teorema 9.15. Wright, 1951.
Existe un número real $\alpha \in (1,2)$ tal que los términos de la sucesión $\{\alpha_n\}_{n=0}^{\infty}$ definida —de manera recursiva— mediante

$$\alpha_0 = \alpha, \quad \alpha_{n+1} = 2^{\alpha_n}, \quad n = 0, 1, 2, 3, \ldots \tag{9.15}$$

cumplen que $\lfloor \alpha_n \rfloor$ es primo para todo $n = 1, 2, 3, \ldots$ En otras palabras, la parte entera de $2^{2^{\cdot^{\cdot^{\cdot^{2^{\alpha}}}}}}$ es un primo para cualquier número de iteraciones de la exponencial.

Demostración. Tomamos $p_1 = 3$ y elegimos una sucesión de números primos $\{p_n\}_{n=1}^{\infty}$ tales que

$$2^{p_n} < p_{n+1} \leq 2^{1+p_n};$$

el postulado de Bertrand garantiza que esta sucesión existe. Lo que perseguimos es encontrar un α tal que, al aplicar (9.15), dé $\lfloor \alpha_n \rfloor = p_n$ para todo $n \in \mathbb{N}$. Como $p_n \geq 3$, no puede ser $p_{n+1} = 2^{1+p_n}$. Tampoco es posible $p_{n+1} = 2^{1+p_n} - 1$, porque $1 + p_n = 2k_n$ y $2^{1+p_n} - 1 = (2^{k_n} + 1)(2^{k_n} - 1)$. Por consiguiente,

$$2^{p_n} < p_{n+1} < 1 + p_{n+1} < 2^{1+p_n}. \tag{9.16}$$

Usemos $\log_2(x)$ para denotar el logaritmo en base 2 de x; asimismo, $\log_2^{\circ n}(x)$ será

$$\log_2^{\circ n}(x) = \underbrace{\log_2 \log_2 \ldots \log_2}_{n \text{ veces}}(x).$$

[5] Goldbach es mucho más conocido por su famosa conjetura «cada número par mayor que dos puede escribirse como suma de dos primos», formulada en 1742, y que aún permanece abierta. Sí que se han conseguido algunos resultados parciales; por ejemplo, Ivan Vinogradov probó en 1937 que cualquier número impar suficientemente grande es suma de tres primos, y el matemático peruano Harald Helfgott ha logrado demostrar, en 2013, que realmente todos los impares mayores que 5 son suma de tres primos, lo cual era conocido como conjetura débil de Goldbach (véase [52]).

Entonces, definamos

$$u_n = \log_2^{\circ n}(p_n), \quad v_n = \log_2^{\circ n}(1+p_n).$$

Por (9.16), se tiene $p_n < \log_2(p_{n+1}) < \log_2(1+p_{n+1}) < 1+p_n$, y, aplicando $\log_2^{\circ n}$, queda

$$u_n < u_{n+1} < v_{n+1} < v_n.$$

Así pues, la sucesión $\{u_n\}_{n=1}^{\infty}$ es creciente y acotada superiormente, luego tiene límite. Sea $\alpha = \lim_n u_n$. Como $u_1 = \log_2(3)$ y $v_1 = \log_2(4) = 2$ resulta que $1 < \alpha < 2$. Además,

$$\underbrace{\log_2 \log_2 \dots \log_2}_{n \text{ veces}}(p_n) < \alpha < \underbrace{\log_2 \log_2 \dots \log_2}_{n \text{ veces}}(1+p_n),$$

luego $p_n < \alpha_n < 1+p_n$. Por lo tanto, $\lfloor \alpha_n \rfloor = p_n$, un primo. □

Acabamos de probar la existencia de un α tal que la sucesión $\{\alpha_n\}_{n=0}^{\infty}$ generada a partir de él cumple que $\lfloor \alpha_n \rfloor$ es primo para todo $n \in \mathbb{N}$. Pero no se conoce, de manera exacta, ningún número α concreto que tenga tal propiedad. Además, su valor depende fuertemente de los primos $\{p_n\}_{n=1}^{\infty}$ que hemos tomado en la demostración. Podemos asignarlos de manera única si convenimos en que p_{n+1} sea, siempre, el menor primo entre 2^{p_n} y 2^{1+p_n}. Por el postulado de Bertrand, tales primos existen, pero no tenemos ninguna fórmula explícita para conocerlos directamente.

Aunque no tiene especial interés, sí que se puede dar un valor aproximado de α. Por ejemplo, el método de construcción seguido muestra que, además de cumplir $\alpha \in (1,2)$, se verifica $\alpha \in (\log_2(3), 2)$. Más aún, $u_n < \alpha < v_n$ para todo $n \in \mathbb{N}$, luego, avanzando en el cálculo de sucesivos u_n y v_n, nos aproximaríamos al valor de α. Pero esto no es tan sencillo de hacer como parece, pues los p_n que van apareciendo (a partir de los cuales se define u_n y v_n) son realmente enormes.

Por otra parte —e incluso sin tener en consideración las dificultades técnicas—, es importante darse cuenta de que este método no tiene ninguna utilidad práctica como función generadora de primos: los primos $\lfloor \alpha_n \rfloor$ que obtendríamos serían, precisamente, los mismos p_n que hemos usado para definir α, luego ¡ya los conocíamos previamente!

Nota 9.16. Existen otras fórmulas similares a las del teorema 9.15 (véase [37]). La más conocida —y base de todas ellas— es la del matemático estadounidense William H. Mills, quien, en 1947, probó que existe un número real $\theta \in (1,\infty)$ tal que $\lfloor \theta^{3^n} \rfloor$ es primo para todo $n \in \mathbb{N}$. Tal θ —construido exactamente como hizo

él— es la denominada constante de Mills. En realidad, esto está muy relacionado con las mejoras del postulado de Bertrand que comentábamos en el apartado 9.3.1. Es más, una vez que podemos apoyarnos en este tipo de resultados más profundos, la demostración de la afirmación de Mills es algo más sencilla que la que hacíamos en el teorema 9.15. Históricamente, el resultado de Mills es anterior; el mérito de Wright fue darse cuenta de que, modificando un poco lo que había hecho Mills, podía encontrar —usando resultados elementales— otra función generadora de primos.

Veamos cómo se probaría la existencia del θ antes citado. Imaginemos que —previamente— hemos demostrado que, para algún $w \leq 2/3$, y para enteros N mayores que un cierto N_0, siempre existe un primo p tal que $N < p < N + N^w$. Asumiendo esto como cierto, construyamos —recursivamente— la siguiente sucesión de primos $\{p_n\}_{n=1}^{\infty}$. Tomamos como p_1 el primer primo tal que $p_1^3 \geq N_0$. A partir de éste, y para cada n, elegimos p_{n+1} el menor primo del intervalo $(p_n^3, p_n^3 + p_n^{3w})$. En estas circunstancias,

$$p_n^3 < p_{n+1} < 1 + p_{n+1} < 1 + p_n^3 + p_n^{3w} \leq 1 + p_n^3 + p_n^2 < (1 + p_n)^3.$$

Ahora, elevando a la potencia 3^{-n-1} obtenemos

$$p_n^{3^{-n}} < p_{n+1}^{3^{-n-1}} < (1 + p_{n+1})^{3^{-n-1}} < (1 + p_n)^{3^{-n}}.$$

Así, la sucesión $\left\{p_n^{3^{-n}}\right\}_{n=1}^{\infty}$ es creciente y acotada superiormente, luego converge. Sea θ su límite, que cumplirá $p_n^{3^{-n}} < \theta < (1 + p_n)^{3^{-n}}$ para todo n. Entonces,

$$p_n < \theta^{3^n} < 1 + p_n$$

luego $\lfloor \theta^{3^n} \rfloor = p_n$, un primo, con lo que queda demostrado el resultado de Mills.

Sin precisar el proceso de construcción, el número θ con la propiedad de que $\lfloor \theta^{3^n} \rfloor$ es primo para todo $n \in \mathbb{N}$ no es único. Por ejemplo, si θ lo cumple, también lo hace θ^3, etc. Es más, Wright probó en 1954 que los θ que tienen esa propiedad forman un conjunto no numerable[6] pero de medida cero (lo mismo ocurre con los α que se comportan como en el teorema 9.15).

Finalmente, no podemos dejar de mencionar que, como función generadora de primos, la expresión $\lfloor \theta^{3^n} \rfloor$ adolece de los mismos problemas que la del teorema 9.15: los primos que se van obteniendo son enormes (ahora no tanto como antes, pero también inmanejables) y hay que conocerlos previamente para construir θ. ◁

[6] Si el lector quiere recordar qué significa esto, puede consultar la sección 6.4.

9.4. Fórmulas de Mertens

En esta sección mostraremos tres resultados que dan estimaciones asintóticas de sumas del tipo $\sum_{p\leq x} f(p)$, donde —recordémoslo una vez más— con $\sum_{p\leq x}$ indicamos que las suma se extiende a todos los primos p menores o iguales que x. Nos interesa lo que ocurre cuando $x\to\infty$ así que, cada vez que aparezca, podemos suponer $x\geq 2$, aunque no lo indiquemos expresamente. Los tres resultados que mostraremos son debidos al alemán Franz Mertens, quien los probó en 1874.

Teorema 9.17. Mertens.
Cuando $x\to\infty$, se tiene la fórmula asintótica

$$\sum_{p\leq x}\frac{\log(p)}{p}=\log(x)+O(1).$$

Demostración. Sabemos —véase el teorema 7.27 y la proposición 8.10— que

$$\sum_{n\leq x}\Lambda(x)\left\lfloor\frac{x}{n}\right\rfloor=\log(\lfloor x\rfloor!)=x\log(x)-x+O(\log(x)).$$

Entonces, efectuando manipulaciones sencillas (entre las que se encuentra el uso de las estimaciones de la proposición 8.10), y aplicando el teorema 9.9, tenemos

$$\begin{aligned}
x\log(x)-x+O(\log(x))&=\sum_{n\leq x}\Lambda(x)\left\lfloor\frac{x}{n}\right\rfloor=\sum_{p\leq x}\sum_{m}\log(p)\left\lfloor\frac{x}{p^m}\right\rfloor\\
&=\sum_{p\leq x}\left\lfloor\frac{x}{p}\right\rfloor\log(p)+\sum_{p\leq x}\sum_{m\geq 2}\left\lfloor\frac{x}{p^m}\right\rfloor\log(p)\\
&=\sum_{p\leq x}\frac{x}{p}\log(p)+O\left(\sum_{p\leq x}\log(p)\right)+O\left(\sum_{p\leq x}\sum_{m\geq 2}\frac{x}{p^m}\log(p)\right)\\
&=x\sum_{p\leq x}\frac{\log(p)}{p}+O(\pi(x)\log(x))+O\left(x\sum_{p\leq x}\left(\frac{\log(p)}{p^2}\sum_{n\geq 0}\frac{1}{p^n}\right)\right)\\
&=x\sum_{p\leq x}\frac{\log(p)}{p}+O(x)+O\left(x\sum_{p\leq x}\frac{\log(p)}{p^2}\right)=x\sum_{p\leq x}\frac{\log(p)}{p}+O(x).
\end{aligned}$$

Despejando resulta

$$x\sum_{p\le x}\frac{\log(p)}{p}=x\log(x)+O(x),$$

de donde se sigue el enunciado. □

A continuación, un resultado que no sólo prueba que la serie de los inversos de los primos diverge (lo cual ya nos ha aparecido dos veces a lo largo del tema), sino que muestra —de una manera bastante precisa— el tamaño de las sumas parciales de la serie:

Teorema 9.18. Mertens.
Existe una constante A tal que, cuando $x\to\infty$, se verifica la fórmula asintótica

$$\sum_{p\le x}\frac{1}{p}=\log(\log(x))+A+O\left(\frac{1}{\log(x)}\right).$$

Demostración. Sean las funciones $f(t)=1/\log(t)$ (con $t\in(1,\infty)$),

$$a(n)=\begin{cases}\frac{\log(n)}{n}, & \text{si } n \text{ es primo},\\ 0, & \text{si } n \text{ no es primo},\end{cases}$$

y

$$A(x)=\sum_{n\le x}a(n)=\sum_{p\le x}\frac{\log(p)}{p}.$$

Entonces,

$$a(n)f(n)=\begin{cases}\frac{1}{n}, & \text{si } n \text{ es primo},\\ 0, & \text{si } n \text{ no es primo},\end{cases}$$

Según la identidad de Abel (teorema 8.6) en el intervalo $[2,x]$, para $x>2$ se tiene

$$\begin{aligned}\sum_{p\le x}\frac{1}{p}&=\sum_{n\le x}a(n)f(n)=\frac{1}{2}+\sum_{2<n\le x}a(n)f(n)\\&=\frac{1}{2}+A(x)f(x)-A(2)f(2)-\int_2^x A(t)f'(t)\,dt\\&=\frac{A(x)}{\log(x)}+\int_2^x\frac{A(t)}{t\log^2(t)}\,dt.\end{aligned}$$

Como, por el teorema 9.17, $A(x) = \log(x) + O(1)$, resulta

$$\begin{aligned}\sum_{p\le x}\frac{1}{p} &= 1 + O\left(\frac{1}{\log(x)}\right) + \int_2^x \frac{dt}{t\log(t)} + \int_2^x \frac{A(t)-\log(t)}{t\log^2(t)}\,dt\\ &= 1 + O\left(\frac{1}{\log(x)}\right) + \log(\log(x)) - \log(\log(2)) + \int_2^x \frac{A(t)-\log(t)}{t\log^2(t)}\,dt\\ &= \log(\log(x)) + O\left(\frac{1}{\log(x)}\right) + 1 - \log(\log(2))\\ &\qquad + \int_2^\infty \frac{A(t)-\log(t)}{t\log^2(t)}\,dt - \int_x^\infty \frac{A(t)-\log(t)}{t\log^2(t)}\,dt\\ &= \log(\log(x)) + A + O\left(\frac{1}{\log(x)}\right),\end{aligned}$$

ya que, en las dos últimas integrales, $A(t) - \log(t) = O(1)$. □

Nota 9.19. En la nota 9.11 (página 521) aludíamos a que las dos series $\sum_{n=1}^\infty 1/n$ y $\sum_{k=1}^\infty 1/p_k$ son divergentes, y tratábamos de precisar «cuánto más divergente» era la primera que la segunda. Ya estamos en condiciones de responder a esto por medio de un análisis de los tamaños de sus sumas parciales. En el tema anterior (apartado (a) de la proposición 8.10) veíamos que

$$\sum_{n\le x}\frac{1}{n} = \log(x) + O(1).$$

En el teorema 9.18 acabamos de probar que

$$\sum_{p\le x}\frac{1}{p} = \log(\log(x)) + O(1).$$

Además, si con primos se quieren conseguir sumas parciales de tamaño $\log(x)$, basta tomar

$$\sum_{p\le x}\frac{\log(p)}{p} = \log(x) + O(1),$$

tal como hemos visto en el teorema 9.17. ◁

Finalmente, veamos un resultado del mismo tipo que los anteriores, pero en el que no aludimos a sumas parciales de una serie, sino a productos parciales de un producto infinito. Antes de comenzar, obsérvese que, forzosamente, $\prod_p (1 - 1/p) = 0$.

En efecto, tomando logaritmos, aparece la serie $\sum_p \log(1-1/p)$; cuando p es grande, $\log(1-1/p)$ es del tamaño de $-1/p$, y ya sabemos que $\sum_p (-1)/p = -\infty$; de aquí se deduce que $\prod_p (1-1/p) = e^{-\infty} = 0$. El teorema que sigue proporciona información adicional: nos muestra el tamaño de sus productos parciales.

Teorema 9.20. Mertens.
Existe una constante positiva B tal que, cuando $x \to \infty$, se verifica la fórmula asintótica

$$\prod_{p\le x}\left(1-\frac{1}{p}\right) = \frac{B}{\log(x)} + O\left(\frac{1}{\log^2(x)}\right).$$

Demostración. Argumentos sencillos de análisis matemático muestran que

$$\lim_{y\to 0^+} \frac{\log(1-y)+y}{y^2} = -\frac{1}{2},$$

luego existe alguna constante $K>0$ tal que

$$\left|\log\left(1-\frac{1}{p}\right)+\frac{1}{p}\right| \le \frac{K}{p^2}$$

para todo p primo. Como consecuencia, la serie

$$\sum_p \left(\log\left(1-\frac{1}{p}\right)+\frac{1}{p}\right) \tag{9.17}$$

converge; llamemos S a su suma. Además,

$$\left|\sum_{p>x} \log\left(1-\frac{1}{p}\right)+\frac{1}{p}\right| \le K\sum_{p>x}\frac{1}{p^2} \le K\sum_{n>x}\frac{1}{n^2} = O\left(\frac{1}{x}\right).$$

Por lo tanto, usando el teorema 9.18,

$$\begin{aligned}\log\left(\prod_{p\le x}\left(1-\frac{1}{p}\right)\right) &= \sum_{p\le x}\log\left(1-\frac{1}{p}\right)\\ &= -\sum_{p\le x}\frac{1}{p} + \sum_p\left(\log\left(1-\frac{1}{p}\right)+\frac{1}{p}\right) - \sum_{p>x}\left(\log\left(1-\frac{1}{p}\right)+\frac{1}{p}\right)\end{aligned}$$

$$= -\log(\log(x)) - A + O\left(\frac{1}{\log(x)}\right) + S + O\left(\frac{1}{x}\right)$$
$$= -\log(\log(x)) - A + S + O\left(\frac{1}{\log(x)}\right)$$
$$= -\log(\log(x)) - A + S + \log\left(1 + O\left(\frac{1}{\log(x)}\right)\right).$$

De aquí,

$$\prod_{p\le x}\left(1-\frac{1}{p}\right) = \frac{e^{-A+S}}{\log(x)}\left(1+O\left(\frac{1}{\log(x)}\right)\right) = \frac{B}{\log(x)} + O\left(\frac{1}{\log^2(x)}\right). \qquad \square$$

El valor de la constante B que aparece en el teorema 9.20 se conoce; concretamente, $B = e^{-\gamma}$ con γ la constante de Euler. La demostración se puede consultar en [50, § 22.8]. Como consecuencia de este hecho, también podemos dar una expresión para el valor de la constante A del teorema 9.18. Para ello, fijémonos en que —tal como aparece en la demostración del teorema 9.20— la constante B vale $B = e^{-A+S}$, donde S es el valor de la suma (9.17). Entonces, $e^{-\gamma} = B = e^{-A+S}$, luego

$$A = \lim_{x\to\infty}\left(\sum_{p\le x}\frac{1}{p} - \log(\log(x))\right) = \gamma + \sum_{p}\left(\log\left(1-\frac{1}{p}\right) + \frac{1}{p}\right); \qquad (9.18)$$

este número, cuyo valor aproximado es $A = 0.261\ldots$, se conoce con el nombre de constante de Mertens.

9.5. El número de divisores primos de un entero

En esta sección intentaremos *medir* el número de divisores primos de un entero, y dar una estimación de cuántos hay. Aunque por el camino veremos algunos otros hechos interesantes, el principal resultado es el teorema 9.27, en el cual veremos cuántos factores primos *suele* tener un entero.

Comencemos definiendo dos funciones aritméticas que no nos han aparecido hasta ahora. Dado n un entero positivo, $\omega(n)$ denotará el número de factores primos distintos de n; y $\Omega(n)$ el número total de factores primos, es decir, contando cada uno según su multiplicidad. Así, si la factorización canónica de $n > 1$ es $n = p_1^{a_1}\cdots p_k^{a_k}$, entonces

$$\omega(n) = k \qquad \text{y} \qquad \Omega(n) = a_1 + \cdots + a_k.$$

Por definición, asumimos que $\omega(1) = \Omega(1) = 0$.

Las dos funciones $\omega(n)$ y $\Omega(n)$ son bastante irregulares. Por ejemplo, si $n = p$ primo, es claro que $\omega(n) = \Omega(n) = 1$. Sin embargo, si $n = p_1 p_2 \cdots p_k$ es el producto de los k primeros primos, entonces

$$\omega(n) = \Omega(n) = k \to \infty \quad \text{cuando } n \to \infty.$$

Pese a los ejemplos anteriores, $\omega(n)$ y $\Omega(n)$ no siempre tienen un tamaño similar. Así, si $n = 2^k$, $\omega(n) = 1$ para todo k, pero

$$\Omega(n) = k = \log_2(n) = \frac{\log(n)}{\log(2)} \to \infty \quad \text{cuando } n \to \infty.$$

No obstante, las medias de las funciones $\omega(n)$ y $\Omega(n)$ sí se comportan de manera regular, y tienen un tamaño bastante similar, como veremos en el teorema que aparecerá a continuación.

Antes, para descargar un poco su demostración, extraigamos de ella el siguiente lema técnico:

Lema 9.21.

Asintóticamente cuando $x \to \infty$, se cumple

$$\sum_{\substack{p^m > x \\ m \geq 2}} \frac{1}{p^m} = O(x^{-1/2}).$$

Demostración. Es claro que

$$p^m > x \;\Leftrightarrow\; m\log(p) > \log(x) \;\Leftrightarrow\; m > \frac{\log(x)}{\log(p)} \;\Leftrightarrow\; m \geq \left\lfloor \frac{\log(x)}{\log(p)} \right\rfloor + 1,$$

ya que m es un entero. Entonces, si denotamos

$$a(x,p) = \text{máx}\left\{ \left\lfloor \frac{\log(x)}{\log(p)} \right\rfloor + 1, 2 \right\}$$

tenemos

$$\sum_{\substack{p^m > x \\ m \geq 2}} \frac{1}{p^m} = \sum_p \sum_{m \geq a(x,p)} \frac{1}{p^m} = \sum_p \frac{1/p^{a(x,p)}}{1 - 1/p} = \sum_p \frac{1}{p^{a(x,p)-1}(p-1)} \leq 2 \sum_p \frac{1}{p^{a(x,p)}}$$

(pues $p - 1 \geq p/2$ para todo p primo). Ahora, démonos cuenta de la relación $-\text{máx}\{r,s\} = \text{mín}\{-r,-s\}$, con lo cual

$$p^{-a(x,p)} = \text{mín}\left\{ p^{-\left\lfloor \frac{\log(x)}{\log(p)} \right\rfloor - 1}, p^{-2} \right\} \leq \text{mín}\{x^{-1}, p^{-2}\},$$

donde en el último paso hemos usado que

$$p^{\left\lfloor \frac{\log(x)}{\log(p)} \right\rfloor + 1} \geq p^{\frac{\log(x)}{\log(p)}} = x$$

(¡convénzase el lector por sí solo de la última igualdad!). Si a lo anterior unimos el uso del teorema 9.9 y la proposición 8.10 (apartado (c)), conseguimos

$$\begin{aligned}\sum_{\substack{p^m > x \\ m \geq 2}} \frac{1}{p^m} &\leq 2 \sum_p \text{mín}\left\{\frac{1}{x}, \frac{1}{p^2}\right\} = 2 \sum_{p \leq \lfloor \sqrt{x} \rfloor} \frac{1}{x} + 2 \sum_{p \geq \lfloor \sqrt{x} \rfloor + 1} \frac{1}{p^2} \\ &\leq 2\pi(\sqrt{x}) \, \frac{1}{x} + 2 \sum_{n = \lfloor \sqrt{x} \rfloor + 1}^{\infty} \frac{1}{n^2} \\ &= \frac{1}{x} \, O\left(\frac{x^{1/2}}{\log(x^{1/2})}\right) + O(x^{-1/2}) = O(x^{-1/2}).\end{aligned}$$

□

Teorema 9.22.
Asintóticamente cuando $x \to \infty$, se cumple

$$\sum_{n \leq x} \omega(n) = x \log(\log(x)) + Ax + O\Big(\frac{x}{\log(x)}\Big), \tag{9.19}$$

$$\sum_{n \leq x} \Omega(n) = x \log(\log(x)) + Cx + O\Big(\frac{x}{\log(x)}\Big), \tag{9.20}$$

con A la misma constante que aparece en el teorema 9.18 y

$$C = A + \sum_p \frac{1}{p(p-1)}.$$

Demostración. Comencemos con (9.19). Aplicando los teoremas 9.18 y 9.9, tenemos

$$\begin{aligned}\sum_{n \leq x} \omega(n) &= \sum_{n \leq x} \sum_{p \mid n} 1 = \sum_{p \leq x} \left\lfloor \frac{x}{p} \right\rfloor = \sum_{p \leq x} \Big(\frac{x}{p} + O(1)\Big) = x \sum_{p \leq x} \frac{1}{p} + O(\pi(x)) \\ &= x \log(\log(x)) + Ax + x \, O\Big(\frac{1}{\log(x)}\Big) + O\Big(\frac{x}{\log(x)}\Big) \\ &= x \log(\log(x)) + Ax + O\Big(\frac{x}{\log(x)}\Big).\end{aligned}$$

Para probar (9.20), aclaremos en primer lugar que $\sum_{p^m|n}$ (o similar) alude a un sumatorio en el que varían los índices p y m. Así,

$$\sum_{n\le x}\Omega(n)=\sum_{n\le x}\sum_{p^m|n}1=\sum_{p^m\le x}\left\lfloor\frac{x}{p^m}\right\rfloor=\sum_{p\le x}\left\lfloor\frac{x}{p}\right\rfloor+\sum_{\substack{p^m\le x\\ m\ge 2}}\left\lfloor\frac{x}{p^m}\right\rfloor$$
$$=\omega(n)+\sum_{\substack{p^m\le x\\ m\ge 2}}\Big(\frac{x}{p^m}+O(1)\Big)=\omega(n)+\sum_{\substack{p^m\le x\\ m\ge 2}}\frac{x}{p^m}+O\Big(\sum_{\substack{p^m\le x\\ m\ge 2}}1\Big).$$

Aplicando el teorema 9.9,

$$\sum_{\substack{p^m\le x\\ m\ge 2}}1=\sum_{p^2\le x}1+\sum_{p^3\le x}1+\cdots+\sum_{p\le\left\lfloor\frac{\log(x)}{\log(2)}\right\rfloor\le x}1\le\frac{\log(x)}{\log(2)}\sum_{p^2\le x}1$$
$$=\frac{\log(x)}{\log(2)}\,\pi(\sqrt{x})=\log(x)\,O\Big(\frac{\sqrt{x}}{\log(\sqrt{x})}\Big)=O(\sqrt{x}),$$

con lo cual

$$\sum_{n\le x}\Omega(n)=\omega(n)+\sum_{\substack{p^m\le x\\ m\ge 2}}\frac{x}{p^m}+O(\sqrt{x}).$$

En esta expresión, es fácil estimar el sumando central por medio de una serie convergente:

$$\sum_{\substack{p^m\le x\\ m\ge 2}}\frac{1}{p^m}\le\sum_{m=2}^{\infty}\sum_{p}\frac{1}{p^m}=\sum_{p}\Big(\frac{1}{p^2}+\frac{1}{p^3}+\cdots\Big)=\sum_{p}\frac{1}{p(p-1)}=T<\infty;$$

además, esta acotación está bien ajustada, en el sentido de que el error cometido es muy pequeño. En efecto, el lema 9.21 nos asegura que

$$\sum_{\substack{p^m\le x\\ m\ge 2}}\frac{1}{p^m}=T-\sum_{\substack{p^m>x\\ m\ge 2}}\frac{1}{p^m}=T+O(x^{-1/2}).$$

Con todo esto,

$$\sum_{n\le x}\Omega(n)=\sum_{n\le x}\omega(n)+x(T+O(x^{-1/2}))+O(\sqrt{x})$$
$$=x\log(\log(x))+(A+T)x+O\Big(\frac{x}{\log(x)}\Big).$$

donde en la última igualdad hemos usado (9.19) y suprimido los sumandos no significativos. □

Nótese que, si usamos la expresión (9.18) y descomponemos

$$\frac{1}{p(p-1)} = \frac{1}{p-1} - \frac{1}{p},$$

el valor de la constante C que aparece en (9.20) es

$$C = \gamma + \sum_p \left(\log\left(1 - \frac{1}{p}\right) + \frac{1}{p-1} \right).$$

9.5.1. ¿Cuántos divisores primos tiene un entero típico?

En 1913, el hindú Srinivasa Ramanujan, que había estudiado matemáticas por su cuenta y que no tenía contacto con ninguna institución universitaria de la metrópoli, escribió a Godfrey Harold Hardy —uno de los más prestigiosos matemáticos ingleses de su época— comunicándole el fruto de sus investigaciones; ya había escrito antes a algunos otros, pero Hardy era el primero en prestar atención a la misiva de Ramanujan.

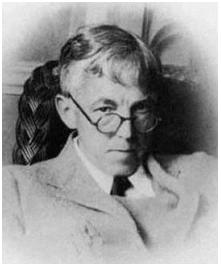
Los escritos que Ramanujan envió a Hardy contenían una larga lista de resultados matemáticos de alto nivel, pero no incluían demostración alguna ni indicaciones de cómo obtenerlos. Entre ellos había resultados que Hardy ya conocía, aunque formulados de manera nada común, así como otros que no le resultaban familiares, pero que pudo demostrar (o refutar) con mayor o menor dificultad. Finalmente, también había extraños resultados aparentemente ciertos, y cuya demostración no imaginaba cómo abordar. En opinión de Hardy, esos escritos sólo podían ser obra de un genio.

Hardy consiguió llevar a Ramanujan a Inglaterra en 1914, y allí establecieron una fructífera colaboración que, debido a la mala salud de Ramanujan y a su temprana muerte en 1920, no pudo durar mucho tiempo. En este apartado expondremos algunos de los interesantes resultados surgidos de esa colaboración entre Hardy y Ramanujan.

En ellos se prueba que, en cierto sentido que ya precisaremos en el enunciado, casi todos los enteros n tienen alrededor de $\log(\log(n))$ divisores primos distintos. Y —quizás sorprendentemente— si contamos cada primo tantas veces como sea su multiplicidad, ¡ocurre lo mismo! La demostración que damos no es la original, sino una posterior de Pál Turán (1934). Pero, antes, necesitaremos algunas cosas.



Srinivasa Ramanujan (1887–1920)



Godfrey Harold Hardy (1877–1947)



Pál Turán (1910–1976)

En primer lugar, con intención de hacer más cortas las demostraciones que siguen, probaremos un resultado previo que, además, tiene cierto interés en sí mismo. En concreto, comencemos dando una estimación para el valor de las medias de $\omega(n)^2$.

Teorema 9.23.
Asintóticamente cuando $x \to \infty$, se cumple

$$\sum_{n\le x} \omega(n)^2 = x\log^2(\log(x)) + O\big(x\log(\log(x))\big).$$

Demostración. Podemos escribir

$$\omega(n)^2 = \Big(\sum_{p|n} 1\Big)^2 = \Big(\sum_{p_1|n} 1\Big)\Big(\sum_{p_2|n} 1\Big) = \sum_{\substack{p_1p_2|n\\ p_1\neq p_2}} 1 + \sum_{p|n} 1$$
$$= \sum_{\substack{p_1p_2|n\\ p_1\neq p_2}} 1 + \omega(n) = \sum_{p_1p_2|n} 1 - \sum_{p^2|n} 1 + \omega(n),$$

de donde

$$\sum_{n\le x} \omega(n)^2 = \sum_{n\le x}\sum_{p_1p_2|n} 1 - \sum_{n\le x}\sum_{p^2|n} 1 + \sum_{n\le x} \omega(n)$$
$$= \sum_{p_1p_2\le x} \left\lfloor \frac{x}{p_1p_2} \right\rfloor - \sum_{p^2\le x} \left\lfloor \frac{x}{p^2} \right\rfloor + \sum_{n\le x} \omega(n).$$

Como

$$\sum_{p^2\le x}\left\lfloor\frac{x}{p^2}\right\rfloor\le\sum_{p^2\le x}\frac{x}{p^2}\le x\sum_{p}\frac{1}{p^2}=O(x)$$

y

$$\sum_{p_1p_2\le x}\left\lfloor\frac{x}{p_1p_2}\right\rfloor=x\sum_{p_1p_2\le x}\frac{1}{p_1p_2}+O(x),$$

usando (9.19) deducimos

$$\sum_{n\le x}\omega(n)^2=x\sum_{p_1p_2\le x}\frac{1}{p_1p_2}+O\big(x\log(\log(x))\big). \tag{9.21}$$

Ahora, démonos cuenta de que

$$\Big(\sum_{p\le\sqrt{x}}\frac{1}{p}\Big)^2\le\sum_{p_1p_2\le x}\frac{1}{p_1p_2}\le\Big(\sum_{p\le x}\frac{1}{p}\Big)^2 \tag{9.22}$$

y, por el teorema 9.18, $\sum_{p\le y}1/p=\log(\log(y))+O(1)$; entonces, tanto el lado derecho como el izquierdo de (9.22) son

$$\big(\log(\log(x))+O(1)\big)^2=\log^2(\log(x))+O\big(\log(\log(x))\big).$$

Empalmando con (9.21) se llega a la estimación del enunciado. □

También usaremos el siguiente resultado, que es, en esencia, lo que en teoría de probabilidad se conoce como desigualdad de Chebyshev:

Proposición 9.24. Desigualdad de Chebyshev.
Sea un conjunto finito $M\subseteq\mathbb{N}$ y una función $f\colon M\to\mathbb{R}$. Asimismo, sean μ y t dos números reales, con $t>0$. Entonces,

$$\operatorname{card}\big\{\,n\in M:|f(n)-\mu|\ge t\,\big\}\le\frac{1}{t^2}\sum_{n\in M}(f(n)-\mu)^2.$$

Demostración. Cuando $|f(n)-\mu|\ge t$, es claro que $1\le(f(n)-\mu)^2/t^2$. Por tanto,

$$\begin{aligned}\operatorname{card}\big\{\,n\in M:|f(n)-\mu|\ge t\,\big\}&=\sum_{\substack{n\in M\\|f(n)-\mu|\ge t}}1\\&\le\sum_{\substack{n\in M\\|f(n)-\mu|\ge t}}\frac{(f(n)-\mu)^2}{t^2}\le\frac{1}{t^2}\sum_{n\in M}(f(n)-\mu)^2.\end{aligned}$$

□

Con esto, abordemos un resultado que es —ya— muy parecido al teorema de Hardy-Ramanujan.

Proposición 9.25.
Para cada $\varepsilon > 0$,

$$\operatorname{card}\big\{\, n \le x : |\omega(n) - \log(\log(x))| \ge \log^{1/2+\varepsilon}(\log(x)) \,\big\} = o(x)$$

cuando $x \to \infty$.

Demostración. Apliquemos la desigualdad de Chebyshev a $M = \{n \in \mathbb{N} : n \le x\}$ y la función $f(n) = \omega(n)$, con $\mu = \log(\log(x))$. Así, para cada $t > 0$ se cumple

$$\operatorname{card}\big\{\, n \le x : |\omega(n) - \log(\log(x))| \ge t \,\big\} \le \frac{1}{t^2} \sum_{n \le x} \big(\omega(n) - \log(\log(x))\big)^2.$$

Estimemos el valor de esta suma. Utilizando los teoremas 9.22 y 9.23, tenemos

$$\begin{aligned}
&\sum_{n\le x} \big(\omega(n) - \log(\log(x))\big)^2 \\
&\qquad = \sum_{n\le x} \omega(n)^2 - 2\log(\log(x)) \sum_{n\le x} \omega(n) + \sum_{n\le x} \log^2(\log(x)) \\
&\qquad = x\log^2(\log(x)) + O\big(x\log(\log(x))\big) - 2\log(\log(x))\big(x\log(\log(x)) + O(x)\big) \\
&\qquad\qquad + (x + O(1))\log^2(\log(x)) \\
&\qquad = O\big(x\log(\log(x))\big).
\end{aligned}$$

Ahora, para $\varepsilon > 0$ cualquiera, sea $t = \log^{1/2+\varepsilon}(\log(x))$. Entonces

$$\begin{aligned}
&\operatorname{card}\big\{\, n \le x : \big|\omega(n) - \log(\log(x))\big| \ge t \,\big\} \\
&\qquad \le \frac{1}{\log^{1/2+\varepsilon}(\log(x))}\, O\big(x\log(\log(x))\big) = O\Big(\frac{x}{\log^{2\varepsilon}(\log(x))}\Big) = o(x),
\end{aligned}$$

tal como queríamos probar. □

El correspondiente resultado para $\Omega(n)$ también es cierto; esto será una consecuencia directa de que, en virtud de la fórmulas (9.19) y (9.20), el tamaño de las medias de $\omega(n)$ y de $\Omega(n)$ es parecido.

Proposición 9.26.
Para cada $\varepsilon > 0$,

$$\operatorname{card}\{\, n \le x : \left|\Omega(n) - \log(\log(x))\right| \ge \log^{1/2+\varepsilon}(\log(x)) \,\} = o(x)$$

cuando $x \to \infty$.

Demostración. En primer lugar, obsérvese que la proposición 9.25 garantiza que, para todo $\varepsilon > 0$,

$$\operatorname{card}\{\, n \le x : \left|\omega(n) - \log(\log(n))\right| \ge \tfrac{1}{2}\log^{1/2+\varepsilon}(\log(x)) \,\} = o(x).$$

En efecto, dado $\varepsilon > 0$, basta elegir un $\varepsilon' > 0$ tal que

$$\tfrac{1}{2}\log^{1/2+\varepsilon}(\log(x)) \ge \log^{1/2+\varepsilon'}(\log(x))$$

y aplicar la proposición a ε'; como nos interesa lo que ocurre cuando $x \to \infty$, podemos suponer que x es suficientemente grande, así que sirve cualquier $\varepsilon' < \varepsilon$, como $\varepsilon' = \varepsilon/2$.

En lo que sigue, por simplicidad, escribiremos $b(x) = \log^{1/2+\varepsilon}(\log(x))$. Si

$$\left|\Omega(n) - \log(\log(x))\right| \ge b(x),$$

deberá ser

$$\left|\omega(n) - \log(\log(x))\right| \ge b(x)/2 \quad \text{o} \quad \left|\Omega(n) - \omega(n)\right| \ge b(x)/2$$

(quizás los dos, pero eso no causa problemas). Ya hemos visto

$$\operatorname{card}\{\, n \le x : \left|\omega(n) - \log(\log(x))\right| \ge b(x)/2 \,\} = o(x),$$

así que basta comprobar que también

$$\operatorname{card}\{\, n \le x : |\Omega(n) - \omega(n)| \ge b(x)/2 \,\} = o(x).$$

Por el teorema 9.22,

$$\sum_{n \le x} (\Omega(n) - \omega(n)) = O(x).$$

Entonces,

$$\operatorname{card}\{\, n \le x : \Omega(n) - \omega(n) \ge b(x)/2 \,\} = O\left(\frac{x}{b(x)}\right) = o(x)$$

(véase el ejercicio 9.22), tal como buscábamos. □

En los enunciados de las proposiciones 9.25 y 9.26, $\omega(n)$ y $\Omega(n)$ se comparan con $\log(\log(x))$; y se usa $\log^{1/2+\varepsilon}(\log(x))$ como tamaño permitido para la diferencia. Pero parece más lógico intentar comparar con algo que dependa de n, no de x; y lo mismo se puede decir de la diferencia que consideramos admisible. Como vemos a continuación, esto también va a ser cierto: bastará con cambiar x por n en cada $\log(\log(x))$ (también cuando aparece elevado a $1/2+\varepsilon$), manteniendo el resto. En realidad, la idea que permite hacer esto es sencilla. Si $n \le x$, hay dos posibilidades, según sea $n < x^{1/e}$ o $n \ge x^{1/e}$. Los $n < x^{1/e}$ forman un conjunto cuyo cardinal es $o(x)$, luego no son significativos. Y, cuando $x^{1/e} \le n \le x$, se tiene

$$\log(\log(x)) - 1 \le \log(\log(n)) \le \log(\log(x)),$$

luego da la impresión que no habrá problemas para sustituir cada $\log(\log(x))$ por $\log(\log(n))$. Así, tras comprobar estos detalles, llegaremos al teorema de Hardy-Ramanujan que estábamos persiguiendo.

Antes de seguir, démonos cuenta de que este tipo de expresiones no tienen sentido para $n = 1$ o 2, ya que entonces $\log(\log(n)) < 0$ y no se puede elevar a $1/2+\varepsilon$. Pero son sólo dos valores, que no tienen ninguna importancia para ver si el cardinal del conjunto involucrado es $o(x)$ o no. Así pues, en todos los conjuntos que aparecen hasta el final del tema, excluiremos, si es que estorban, los números $n = 1$ y 2, aunque no lo mencionaremos expresamente.

Teorema 9.27. Hardy-Ramanujan, 1917.
Para cada $\varepsilon > 0$, asintóticamente cuando $x \to \infty$ se cumple

$$\operatorname{card}\big\{\, n \le x : \big|\omega(n) - \log(\log(n))\big| \ge \log^{1/2+\varepsilon}(\log(n)) \,\big\} = o(x)$$

y

$$\operatorname{card}\big\{\, n \le x : \big|\Omega(n) - \log(\log(n))\big| \ge \log^{1/2+\varepsilon}(\log(n)) \,\big\} = o(x).$$

Demostración. Analizaremos sólo el caso correspondiente a ω, pues la argumentación para Ω es la misma. Sabemos

$$\operatorname{card}\big\{\, n \le x : \big|\omega(n) - \log(\log(x))\big| \ge \log^{1/2+\varepsilon}(\log(x)) \,\big\} = o(x) \quad \text{para todo } \varepsilon > 0,$$

y queremos probar

$$\operatorname{card}\big\{\, n \le x : \big|\omega(n) - \log(\log(n))\big| \ge \log^{1/2+\varepsilon}(\log(n)) \,\big\} = o(x) \quad \text{para todo } \varepsilon > 0.$$

Como ya hemos comentado, los $n \le x^{1/e}$ no son significativos, pues $\text{card}\{n \in \mathbb{N} : n \le x^{1/e}\} = \lfloor x^{1/e} \rfloor = o(x)$. Por consiguiente, la primera afirmación es equivalente a

$$\begin{aligned}\text{card}\big\{\, n \in [x^{1/e}, x] : \big|\omega(n) - \log(\log(x))\big| + 1 \ge \log^{1/2+\varepsilon}(\log(x)) + 1 \,\big\} \\ = o(x) \quad \text{para todo } \varepsilon > 0\end{aligned} \tag{9.23}$$

(en seguida veremos para qué hemos añadido los dos «+1»); y la segunda a

$$\begin{aligned}\text{card}\big\{\, n \in [x^{1/e}, x] : \big|\omega(n) - \log(\log(n))\big| \ge \log^{1/2+\varepsilon}(\log(n)) \,\big\} \\ = o(x) \quad \text{para todo } \varepsilon > 0.\end{aligned} \tag{9.24}$$

Así —y puesto que ambas expresiones se refieren a «para todo $\varepsilon > 0$»— el teorema estará probado si vemos que (9.23) para $\varepsilon/2$ implica (9.24) para ε. Vamos con ello.

Al ser $n \in [x, x^{1/e}]$, se cumple $\log(\log(x)) - 1 \le \log(\log(n)) \le \log(\log(x))$. Entonces,

$$\begin{aligned}\big|\omega(n) - \log(\log(n))\big| &\le \big|\omega(n) - \log(\log(x))\big| + \big|\log(\log(x) - \log(\log(n))\big| \\ &\le \big|\omega(n) - \log(\log(x))\big| + 1.\end{aligned}$$

Además, nos interesa lo que ocurre cuando $x \to \infty$, luego podemos suponer que x es tan grande como queramos; y lo mismo para n, pues $n \ge x^{1/e}$. Así,

$$\big(\log(\log(x))\big)^{1/2+\varepsilon/2} + 1 \le \big(\log(\log(n)) + 1\big)^{1/2+\varepsilon/2} \le \big(\log(\log(n))\big)^{1/2+\varepsilon}.$$

Con todo esto, si se cumple (9.23) para $\varepsilon/2$, también deberá ser cierto

$$\text{card}\big\{\, n \in [x^{1/e}, x] : \big|\omega(n) - \log(\log(n))\big| \ge \log^{1/2+\varepsilon/2}(\log(x)) + 1 \,\big\} = o(x)$$

(pues hemos sustituido $|\omega(n) - \log(\log(x))| + 1$ por un valor menor) y, finalmente,

$$\text{card}\big\{\, n \in [x^{1/e}, x] : \big|\omega(n) - \log(\log(n))\big| \ge \log^{1/2+\varepsilon}(\log(n)) \,\big\} = o(x),$$

tal como queríamos comprobar. □

Ejemplo. Estima cuántos divisores primos debe tener —aproximadamente— un número típico de tamaño 10^k, con $k = 10, 20, \ldots, 120$.

Solución. Sin más que utilizar el teorema, el número de factores primos de un número de tamaño 10^k es, generalmente, $\log(\log(10^k)) = \log(k) + \log(\log(10))$; además, esto es así tanto si contamos cada divisor primo según su multiplicidad como si no. Toda esta información se resume en la siguiente tabla:

Tamaño	Div. primos	Tamaño	Div. primos	Tamaño	Div. primos
10^{10}	3.136 62	10^{50}	4.746 06	10^{90}	5.333 84
10^{20}	3.829 76	10^{60}	4.928 38	10^{100}	5.439 20
10^{30}	4.235 23	10^{70}	5.082 53	10^{110}	5.534 51
10^{40}	4.522 91	10^{80}	5.216 06	10^{120}	5.621 52

◊

9.5.2. ¿Cuántos divisores tiene un entero típico?

Aplicando los resultados sobre $\omega(n)$ y $\Omega(n)$, podemos obtener información sobre la función $\tau(n)$ que cuenta el número de divisores de n. En el teorema 7.2 veíamos que, si $n = p_1^{a_1} p_2^{a_2} \cdots p_k^{a_k}$, entonces

$$\tau(n) = (a_1 + 1)(a_2 + 1) \cdots (a_k + 1).$$

Para el mismo n, se tiene

$$\omega(n) = k \qquad \text{y} \qquad \Omega(n) = a_1 + a_2 + \cdots + a_k.$$

Entonces, sin más que caer en la cuenta de que $2 \leq a + 1 \leq 2^a$ para todo $a \in \mathbb{N}$, deducimos la desigualdad

$$2^{\omega(n)} \leq \tau(n) \leq 2^{\Omega(n)}.$$

Como consecuencia

$$\omega(n) \leq \log_2(\tau(n)) \leq \Omega(n)$$

y los resultados del teorema 9.27 se traducen de manera inmediata a resultados sobre $\tau(n)$. En particular obtenemos que, para cada $\varepsilon > 0$,

$$\operatorname{card}\big\{ n \leq x : \big|\log_2(\tau(n)) - \log(\log(n))\big| < \log^{1/2+\varepsilon}(\log(n)) \big\} = x - o(x).$$

Manipulemos la desigualdad de la expresión anterior. Es clara su equivalencia a

$$\log(\log(n)) - \log^{1/2+\varepsilon}(\log(n)) < \log_2(\tau(n)) < \log(\log(n)) + \log^{1/2+\varepsilon}(\log(n));$$

y, por tanto, a

$$2^{\log(\log(n)) - \log^{1/2+\varepsilon}(\log(n))} < \tau(n) < 2^{\log(\log(n)) + \log^{1/2+\varepsilon}(\log(n))}.$$

Sin pérdida de generalidad, podemos considerar que ε es pequeño, pues es el caso interesante; en concreto, $0 < \varepsilon < 1/2$. Asintóticamente, la parte $\log^{1/2+\varepsilon}(\log(n))$ es mucho más pequeña que $\eta \log(\log(n))$ (con $\eta > 0$), luego una condición del tipo

$$2^{(1-\eta)\log(\log(n))} < \tau(n) < 2^{(1+\eta)\log(\log(n))}$$

es menos exigente (hay más valores de n que la satisfacen). Así, hemos probado lo siguiente:

Teorema 9.28. Hardy-Ramanujan, 1917.
Para cada $\eta > 0$,

$$\operatorname{card}\left\{ n \le x : 2^{(1-\eta)\log(\log(n))} < \tau(n) < 2^{(1+\eta)\log(\log(n))} \right\} = x - o(x)$$

cuando $x \to \infty$.

De manera más coloquial podemos decir que, si η es positivo, entonces

$$2^{(1-\eta)\log(\log(n))} < \tau(n) < 2^{(1+\eta)\log(\log(n))}$$

para gran parte de los $n \in \mathbb{N}$, y el teorema anterior es el encargado de precisar el significado del «gran parte». Si, para hacernos una idea, tomamos $\eta = 0$, queda que el tamaño de $\tau(n)$ suele ser $2^{\log(\log(n))} = (\log(n))^{\log(2)}$.

Ejemplo. Estimar el número aproximado de divisores de un número típico de tamaño 10^k, con $k = 10, 20, \ldots, 120$.

Solución. Según acabamos de comentar, el número de divisores de un número de tamaño 10^k es, habitualmente, alrededor de $2^{\log(\log(10^k))} = k^{\log(2)}(\log(10))^{\log(2)}$. Entonces, tenemos

Tamaño	Divisores	Tamaño	Divisores	Tamaño	Divisores
10^{10}	8.794 60	10^{50}	26.835 21	10^{90}	40.331 69
10^{20}	14.219 16	10^{60}	30.450 14	10^{100}	43.387 35
10^{30}	18.833 51	10^{70}	33.883 89	10^{110}	46.350 49
10^{40}	22.989 64	10^{80}	37.169 80	10^{120}	49.231 99

◊

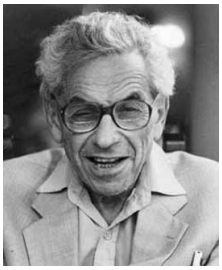
9.5.3. Propiedades estadísticas

El tipo de resultados que hemos estudiado en esta sección tiene continuidad en otros que revelan —aún más— las propiedades estadísticas de la función $\omega(n)$. Por ejemplo, utilizando argumentos sencillos, Pál Erdős y Mark Kac (1939) probaron que

$$\frac{\omega(n) - \log(\log(n))}{\sqrt{\log(\log(n))}} \tag{9.25}$$

es, asintóticamente, una distribución normal de media 0 y varianza 1. En otras palabras, el teorema de Erdős-Kac establece que el conjunto de números naturales n tales que la expresión (9.25) es $\le u$ tiene una densidad asintótica igual a

$$\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{u} e^{-t^2/2}\, dt.$$



Pál Erdős (1913–1996)



Mark Kac (1914–1984)

Aunque el origen del término es posterior, se puede considerar que este resultado es el origen de la teoría de números probabilística. Si el lector está interesado en el tema, puede consultar [90], donde encontrará referencias, generalizaciones y comentarios históricos.

9.6. Ejercicios

9.1. Sea $n \in \mathbb{N}$ cualquiera. Prueba que existen, al menos, n números consecutivos no primos.

9.2. Demuestra que, para todo $n \in \mathbb{N}$, se cumple

$$\frac{1}{\zeta(2)} < \frac{\sigma(n)\varphi(n)}{n^2} \leq 1.$$

9.3. Sea f una función multiplicativa y $s \in \mathbb{C}$ tal que la serie $\sum_n f(n)n^{-s}$ es absolutamente convergente. Entonces,

$$\sum_{n=1}^{\infty} \frac{f(n)}{n^s} = \prod_p \left(1 + \frac{f(p)}{p^s} + \frac{f(p^2)}{p^{2s}} + \frac{f(p^3)}{p^{3s}} + \cdots\right).$$

Si f es completamente multiplicativa, se tiene, además,

$$\sum_{n=1}^{\infty} \frac{f(n)}{n^s} = \prod_p \frac{1}{1 - f(p)p^{-s}}.$$

Demuestra ambas afirmaciones.

9.4. Como consecuencia del ejercicio anterior, y de las expresiones que se dan en los ejercicios 8.13 a 8.19, comprueba que, en las regiones que se indica (a estos efectos, usamos la abreviatura $\mathrm{Re}^+(a) = \max\{0, \mathrm{Re}(a)\}$), se cumple

$$\frac{1}{\zeta(s)} = \sum_{n=1}^{\infty} \frac{\mu(n)}{n^s} = \prod_p (1-p^{-s}), \qquad \mathrm{Re}(s) > 1,$$

$$\frac{\zeta(s-1)}{\zeta(s)} = \sum_{n=1}^{\infty} \frac{\varphi(n)}{n^s} = \prod_p \frac{1-p^{-s}}{1-p^{1-s}}, \qquad \mathrm{Re}(s) > 2,$$

$$\zeta(s)\zeta(s-a) = \sum_{n=1}^{\infty} \frac{\sigma_a(n)}{n^s} = \prod_p \frac{1}{(1-p^{-s})(1-p^{a-s})}, \qquad \mathrm{Re}(s) > 1 + \mathrm{Re}^+(a),$$

$$\frac{\zeta(2s)}{\zeta(s)} = \sum_{n=1}^{\infty} \frac{\lambda(n)}{n^s} = \prod_p \frac{1}{1+p^{-s}}, \qquad \mathrm{Re}(s) > 1,$$

$$\frac{\zeta(s)}{\zeta(2s)} = \sum_{n=1}^{\infty} \frac{|\mu(n)|}{n^s} = \prod_p (1+p^{-s}), \qquad \mathrm{Re}(s) > 1.$$



9.5. Para cada $n \in \mathbb{N}$ sea $\sigma(n) = \sum_{d|n} d$ la suma de los divisores de n. Dado $x \geq 2$ y $m = \prod_{p \leq x} p$, prueba que

$$\frac{\sigma(m)}{m} > \sum_{p \leq x} \frac{1}{p}.$$

Finalmente, concluye que

$$\limsup_{n\to\infty} \frac{\sigma(n)}{n} = \infty \quad \text{y} \quad \liminf_{n\to\infty} \frac{\sigma(n)}{n} = 1.$$



9.6. Prueba que, para cada $x \geq 2$,

$$\mathrm{mcm}(1, 2, \ldots, \lfloor x \rfloor) = \prod_{p \leq x} p^{\left\lfloor \frac{\log(x)}{\log(p)} \right\rfloor} = \exp\left(\sum_{n \leq x} \Lambda(n) \right).$$



9.7. Para cada entero $n > 1$, escribamos su descomposición canónica mediante $n = \prod_p n^{\nu(n,p)}$ (obsérvese que casi todos los exponentes serán nulos, luego el producto sólo tiene un número finito de factores que no son 1). Con esta notación, sea ξ la función $\xi(n) = \prod_p \nu(n,p)!$ para $n > 1$ y $\xi(1) = 1$. Analiza si son ciertas o no las siguientes afirmaciones:

(a) ξ es multiplicativa.
(b) ξ es completamente multiplicativa.
(c) $\xi(n)\xi(m)\,|\,\xi(nm)$ para todo $n, m \in \mathbb{N}$.
(d) $\xi(n)\xi(m) \leq \xi(nm)$ para todo $n, m \in \mathbb{N}$.

9.8. Caracteriza los $n \in \mathbb{N}$ para los cuales la función ξ del ejercicio anterior cumple $\xi(n) = 1$.

9.9. Prueba[7] que la función $y = \log(x)/\sqrt{x}$, definida para $x \geq 1$, está acotada superiormente por $2/e$, y que la cota sólo se alcanza en $x = e^2$. Antes de ese valor, la función es creciente; y después, decreciente.

9.10. Prueba que, si $\{p_k\}_{k=1}^{\infty}$ es la sucesión de todos los primos, se cumple

$$\sum_{k=2}^{\infty} \frac{1}{p_k \log(k)} < \infty.$$

Busca series también convergentes, pero cuyo término general sea mayor que el de la serie anterior. Y demuestra que, sin embargo,

$$\sum_{k=3}^{\infty} \frac{1}{p_k \log(\log(k))} = \infty.$$

9.11. Continuando con $\{p_k\}_{k=1}^{\infty}$ la sucesión de todos los primos, prueba que

$$\sum_{k=2}^{\infty} \frac{1}{p_k \log(p_k)} < \infty.$$

9.12. Situándonos en la demostración del lema 9.13, comprueba que, en realidad, nunca se da la primera posibilidad de las dos que contempla la demostración.

9.13. Sea p_k el k-ésimo primo. Empleando —exclusivamente— el argumento de Euclides de su demostración de la existencia de infinitos primos, prueba que $p_k < 2^{2^k}$ para todo $k \in \mathbb{N}$. Aplicando esta desigualdad, deduce que

$$\pi(n) \geq \big(\log(\log(n)) - \log(2)\big)/\log(2)$$

cuando $n \geq 3$ (para que $\log(\log(n))$ tenga sentido).

[7] Resulta curioso observar cómo este resultado elemental sobre el comportamiento de la función de variable real $y = \log(x)/\sqrt{x}$ resulta útil en las demostraciones que hemos dado de tres de los resultados importantes de este tema: los teoremas 9.9, 9.10 y 9.14.

9.14. Sea p_k el k-ésimo primo. Sin apoyarte en el teorema de Chebyshev, prueba que $p_k \leq 2^k$ para todo $k \in \mathbb{N}$. Usa luego esa desigualdad para obtener la siguiente estimación —bastante burda— para $\pi(n)$:

$$\pi(n) \geq K \log(n),$$

donde K es una constante que debes identificar.

9.15. Demuestra que la conjetura de Andrica formulada en (9.14) equivale a afirmar que

$$p_{k+1} - p_k < 2\sqrt{p_k} + 1,$$

para todo $k \geq 1$, donde p_k denota el k-ésimo primo.

9.16. Comprueba que, si la conjetura de Oppermann sobre distribución de números primos fuera verdadera (véase el apartado 9.3.1), también lo serían las conjeturas de Legendre y de Brocard.

9.17. Sea $\{p_k\}_{k=1}^{\infty}$ la sucesión de todos los primos (ordenados de forma creciente). Prueba que el número

$$\alpha = \sum_{k=1}^{\infty} p_k 2^{-k^2}$$

está bien definido y que, además, proporciona la siguiente función generadora de primos:

$$p_k = \lfloor 2^{k^2} \alpha \rfloor - 2^{2k-1} \lfloor 2^{(k-1)^2} \alpha \rfloor, \quad \text{para todo } k \in \mathbb{N}.$$

9.18. Sea p_k el k-ésimo primo. Prueba que

$$k < p_k < k^2, \qquad k \geq 2.$$

9.19. Continuemos usando $\{p_k\}_{k=1}^{\infty}$ para denotar la sucesión de números primos (ordenados de manera creciente). Encuentra constantes positivas α y β tales que

$$n^{\alpha n} < \prod_{k=1}^{n} p_k < n^{\beta n}, \qquad n \geq 2.$$

9.20. Sea un polinomio $P \in \mathbb{Z}[x]$ y, dado $k \in \mathbb{Z}$, tomemos $m = P(k)$. Prueba que m divide a $P(k + rm)$ para todo $r \in \mathbb{Z}$.

9.21. Demuestra[8] que no existe ningún polinomio $P \in \mathbb{Z}[x]$ (no constante) tal que $P(n)$ sea primo para todo $n \in \mathbb{N}$; es más, tampoco existe si sólo exigimos que $P(n)$ sea primo a partir de un cierto n_0.
Sugerencia. Apóyate en el ejercicio anterior.

9.22. Sea $f(n)$ una función aritmética no negativa, $a(x)$ y $b(x)$ dos funciones reales positivas, y supongamos que $\sum_{n\le x} f(n) = O(a(x))$ cuando $x \to \infty$. Prueba que, en estas circunstancias,

$$\operatorname{card}\{\, n \le x : f(n) \ge b(x) \,\} = O\left(\frac{a(x)}{b(x)}\right).$$

[8] Tal como ya hemos comentado en el texto, esto fue probado por Goldbach en 1752. En lo que a nosotros respecta, el enunciado de este ejercicio coincide con el de los ejercicios 3.3 y 3.4 del tema 3, donde proponíamos efectuar las correspondientes demostraciones por medio de congruencias.

Tema 10
El teorema de los números primos

10.1. Recordatorio de análisis complejo
10.2. Definición de la zeta de Riemann y resultados preparatorios
10.3. Un teorema tauberiano
10.4. El final de la demostración del teorema de los números primos
10.5. Algunas consecuencias del teorema de los números primos
10.6. El tamaño del error al aproximar $\pi(x)$
10.7. Las aproximaciones que propuso Riemann
10.8. Ejercicios



Adrien Marie Legendre
(1752–1833)



Carl Friedrich Gauss
(1777–1855)



Georg Friedrich
Bernhard Riemann
(1826–1866)



Jacques Hadamard
(1865–1963)



Charles Jean de la
Vallée-Poussin
(1866–1962)

Usemos $\pi(x)$ para denotar la cantidad de números primos menores o iguales que x, es decir,

$$\pi(x) = \text{card}\{p \text{ primo} : p \leq x\}, \qquad x \geq 1.$$

El objetivo casi único de este tema es demostrar el teorema de los números primos —o teorema del número primo, como lo denominan en muchos textos— que, como ya sabemos, afirma que

$$\pi(x) \sim \frac{x}{\log(x)}, \qquad x \to \infty,$$

donde $f(x) \sim g(x)$ significa que $\lim_{x\to\infty} f(x)/g(x) = 1$. Este hecho había sido conjeturado independientemente por Adrien Marie Legendre (1798) y Carl Friedrich Gauss (1849) a partir de sus observaciones en una tabla de primos.

Mucho antes, en 1737 —y tal como hemos visto con detalle en la sección 9.1 de este libro—, Leonhard Euler, tras desarrollar $(1-1/p^s)^{-1}$ en serie geométrica, había encontrado la identidad

$$\zeta(s)=\sum_{n=1}^{\infty}\frac{1}{n^s}=\prod_{p}\left(1-\frac{1}{p^s}\right)^{-1},\quad s>1, \tag{10.1}$$

y de ahí, tomando límites cuando $s\to 1$, había mostrado que $\sum_p 1/p=\infty$ (por supuesto, $\prod_p$ y $\sum_p$ indican que el producto o la suma involucradas habrá que hacerlos sobre todos los números primos, y esta notación la usaremos en todo el tema). En particular, como la serie de los inversos de los números primos divergía, esto implicaba la existencia de un número infinito de primos. Pero nada parecía indicar que (10.1) pudiera tener ninguna utilidad para estimar la cantidad de primos $\leq x$, es decir, para estimar $\pi(x)$.

Después de los primeros escarceos de Gauss y Legendre, la historia moderna del teorema de los números primos comienza con el matemático alemán Bernhard Riemann, quien en 1859 escribió un artículo de ocho páginas en el que, sin dar demostraciones, relacionaba la función $\pi(x)$ con los ceros complejos de la función ζ; para ello, y como paso previo, había que considerar la serie (10.1) para $\mathrm{Re}(s)>1$, y luego extender la función ζ —por prolongación analítica— al plano complejo. De hecho, y aunque lo hace sin rigor, Riemann llega a dar una fórmula explícita para $\pi(x)$ en la que intervienen dichos ceros; ¡una idea sorprendente! Actualmente, la función ζ se denomina, con toda justicia, función zeta de Riemann.

Pasó aproximadamente medio siglo hasta que el francés Jacques Hadamard y el belga Charles de la Vallée-Poussin lograran utilizar la ideas de Riemann para completar, en 1896 y de manera independiente, la demostración del teorema de los números primos. Fue un hito de la denominada teoría de números analítica.

Merece la pena insistir en la importancia de la teoría de variable compleja en la prueba del teorema de los números primos. No es raro en matemáticas que algunas demostraciones se puedan hacer, de manera particularmente sencilla, aplicando análisis complejo[1] —como dijo Hadamard, «el camino más corto entre dos verdades del análisis real pasa a menudo por el análisis complejo»—, pero en este caso resulta especialmente sorprendente, pues se trataba de una propiedad de los números primos cuya demostración dependía fuertemente del uso de la teoría de funciones de variable compleja.

[1] El paradigma es la demostración del teorema fundamental del álgebra, que resulta muy sencilla utilizando variable compleja. Se usa el teorema de Liouville (de análisis complejo), según el cual las únicas funciones derivables en todo $\mathbb{C}$ y acotadas son las constantes. Entonces, si un polinomio $P(z)$ no tiene ceros en $\mathbb{C}$, la función $1/P(z)$ será derivable en todo $\mathbb{C}$ y acotada; por tanto, será constante. Como consecuencia, cualquier polinomio no constante debe tener raíces en $\mathbb{C}$.

En los años siguientes a la demostración de Hadamard y De la Vallée-Poussin no había ningún indicio de que se pudiera encontrar una demostración elemental del teorema de los números primos, donde la descripción «elemental» se emplea como contrapartida a «analítica», que es como se alude a las demostraciones que dependen de la teoría de funciones de variable compleja. Se pensaba que si tal demostración elemental era descubierta, iba a modificar profundamente las perspectivas que se tenían sobre la matemática. Así, por ejemplo, el matemático inglés Godfrey Harold Hardy afirmaba, en 1921, que si alguien realizara una demostración elemental del teorema de los números primos, «es momento de que los libros sean reorganizados y la teoría reescrita».

La demostración elemental llegó en 1949, de la mano del húngaro Pál Erdős y el noruego Atle Selberg. Pese a que, en cierta medida, Erdős y Selberg habían compartido sus avances en pos de dicha demostración elemental, presentaron sus demostraciones por separado, en el contexto de una amarga disputa entre los dos matemáticos. En todo caso, las aportaciones de Selberg eran cruciales en ambas pruebas, y Selberg acabó recibiendo la Medalla Fields en 1950.[2] Aunque verdaderamente estas demostraciones no utilizaban análisis complejo, en la práctica no suponían ventajas sobre las «no elementales», y no originaron ninguna revolución en las matemáticas.

Con el tiempo, las demostraciones analíticas del teorema de los números primos se han ido simplificando considerablemente (el primero que lo hizo fue Edmund Landau en 1903) y, aunque siguen utilizando de manera crucial la variable compleja, son mucho más accesibles que las originales. El lector puede hacerse una idea de esta evolución en el excelente artículo [11]. Nosotros nos basaremos, fundamentalmente, en la demostración dada por Donald J. Newman en 1980 en su artículo [73] (y que luego reprodujo, incluyendo algún detalle adicional, en [74]), junto con modificaciones provenientes de [104].

Desde luego, otros textos incluyen otras demostraciones, cada una con sus peculiaridades, y con diferente grado de detalle. Así, por ejemplo, podemos encontrar demostraciones analíticas en [3, 5, 12, 26, 33, 40, 69, 100] (sin olvidar al ya mencionado [74] que contiene dos), y demostraciones elementales en [5, 33, 50, 71, 87, 100].

En el presente libro, y tras la demostración del teorema de los números primos —que concluiremos en la sección 10.4—, dedicaremos el resto del tema a ver algunas consecuencias del teorema y a intentar analizar —esta vez ya sin incluir demostraciones— el porqué de la ineludible presencia de los ceros de la función zeta de Riemann en las aproximaciones a $\pi(x)$.

[2] Se pueden ver más detalles sobre las Medallas Fields —el mayor galardón que puede recibir un matemático— en la página 113, en la nota al pie numerada como 7.

Finalmente, queremos concluir esta introducción recomendando [36], un emocionante y maravilloso relato sobre los números primos, la función zeta de Riemann, y los numerosos matemáticos que han intervenido en su fascinante historia.

10.1. Recordatorio de análisis complejo

Como el uso de la teoría de variable compleja va a ser esencial en la demostración que vamos a dar del teorema de los números primos, comenzamos con una pequeña introducción al análisis complejo. Si un lector nunca ha estudiado variable compleja, posiblemente esta introducción le resulte muy escasa, y debería acudir previamente a alguno de los muchos textos existentes. Pero sí que puede resultar útil como recordatorio para quien haya estudiado previamente análisis complejo y lo tenga algo olvidado, o como mera recopilación de los resultados más importantes que utilizaremos más adelante.

El análisis complejo se ocupa de estudiar funciones en las que tanto la variable dependiente como la independiente son números complejos, y es una rama de las matemáticas que tiene sus raíces en el siglo XIX e incluso algo antes. Como antecedente hay que mencionar de nuevo a Euler, que en su libro *Introductio in Analysin Infinitorum*, publicado en 1748 (esta magnífica obra puede verse con todo su esplendor en [43], una edición facsímil acompañada de una traducción comentada), prueba que la función exponencial satisface

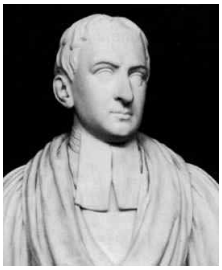
$$e^{it} = \cos(t) + i\,\mathrm{sen}(t) \tag{10.2}$$

para cualquier número real t, una expresión que inequívocamente identificamos como propia de la variable compleja.[3] Un caso especial de la fórmula (10.2) es lo que se conoce como identidad de Euler,

$$e^{i\pi} + 1 = 0.$$

Esta expresión relaciona los cinco números más importantes de las matemáticas (0, 1, e, π e i), y muchos la consideran la más bella fórmula matemática de la historia. Tras Euler, los nombres más destacados en el desarrollo del análisis complejo son, posiblemente, Gauss, Riemann, Cauchy y Weierstrass.

[3] En realidad, ya antes, en 1714, Roger Cotes había obtenido que $\log(\cos(t) + i\,\mathrm{sen}(t)) = it$; por supuesto, esta expresión no es del todo correcta pero esencialmente es equivalente a la de Euler. Y, en 1722, Abraham de Moivre había publicado su famosa fórmula $(\cos(t) + i\,\mathrm{sen}(t))^n = \cos(nt) + i\,\mathrm{sen}(nt)$ que, cuando ya disponemos de (10.2), es una consecuencia inmediata sin más que usar $(e^{it})^n = e^{int}$.



Roger Cotes
(1682–1716)



Augustin Louis Cauchy
(1789–1857)



Karl Weierstrass
(1815–1897)

En general, el análisis complejo estudia funciones

$$f\colon U \subseteq \mathbb{C} \to \mathbb{C}$$

donde U es algún subconjunto abierto del plano complejo $\mathbb{C}$. El concepto más importante del análisis complejo es el de función holomorfa. Una función f se dice *holomorfa* en un abierto U de $\mathbb{C}$ si está definida en todos los puntos de U y es derivable en ellos, es decir, si existe

$$f'(z) = \lim_{w \to z} \frac{f(w) - f(z)}{w - z} \in \mathbb{C}$$

para todo $z \in U$. Luego veremos el motivo, pero adelantemos que a las funciones holomorfas también se las denomina *analíticas*. Además, las funciones holomorfas en todo $\mathbb{C}$ se conocen con el nombre de *enteras*.

La derivabilidad compleja tiene consecuencias mucho más fuertes que la derivabilidad ordinaria con funciones de variable real o, incluso, que la diferenciabilidad de funciones de $\mathbb{R}^2$ en $\mathbb{R}^2$ (hay una diferencia esencial entre usar $\mathbb{C}$ o $\mathbb{R}^2$, y es que en $\mathbb{R}^2$ no se puede dividir, y por tanto no se puede definir la derivada).[4] Por ejemplo,

[4] Lo que se cumple es que, cuando una función $f\colon \mathbb{R}^2 \to \mathbb{R}^2$ dada por sus dos componentes $f(x,y) = (u(x,y), v(x,y))$ la interpretamos como función $f\colon \mathbb{C} \to \mathbb{C}$ mediante $f(x+iy) = u(x,y) + iv(x,y)$, la función compleja es derivable en un punto $z_0 = x_0 + iy_0$ si y sólo si la función de $\mathbb{R}^2$ es diferenciable en (x_0, y_0) y se verifica

$$\frac{\partial u}{\partial x}(x_0,y_0) = \frac{\partial v}{\partial y}(x_0,y_0), \qquad \frac{\partial u}{\partial y}(x_0,y_0) = -\frac{\partial v}{\partial x}(x_0,y_0)$$

(condiciones de Cauchy-Riemann). En ese caso, la derivada de f en z_0 es $f'(z_0) = \frac{\partial u}{\partial x}(x_0,y_0) + i\frac{\partial v}{\partial x}(x_0,y_0)$.

las funciones holomorfas siempre son infinitamente diferenciables, y eso no ocurre con las funciones derivables reales.

Además, las funciones holomorfas siempre tienen desarrollo en serie de potencias. En concreto, si f es holomorfa en U y $z_0 \in U$, en un entorno de z_0 se cumplirá

$$f(z) = \sum_{n=0}^{\infty} a_n (z - z_0)^n; \tag{10.3}$$

los coeficientes a_n se pueden expresar como

$$a_n = \frac{f^{(n)}(z_0)}{n!}.$$

Obsérvese que esto es como en el desarrollo de Taylor de funciones reales, pero, a su vez, hay diferencias importantes: en $\mathbb{R}$ necesitaríamos estar tratando con funciones infinitamente diferenciables y, además, no se puede garantizar que el desarrollo de Taylor de una función real converja a esa función.[5] Recíprocamente, cualquier desarrollo de la forma (10.3) convergente en un entorno de z_0 da lugar a una función holomorfa en ese entorno. Esto explica la doble denominación de funciones holomorfas y funciones analíticas que se emplea: en principio, holomorfa hace alusión a la derivabilidad, y analítica a la existencia de desarrollo en serie; pero, como acabamos de ver, en variable compleja ambos hechos (¡considerados siempre en un entorno de un punto!) son equivalentes.

La mayoría de las funciones elementales, incluyendo los polinomios, la función exponencial y las trigonométricas son funciones holomorfas (y, por supuesto, restringidas a $\mathbb{R} \subseteq \mathbb{C}$ dan lugar a la correspondiente función real «original»). Damos por hecho que el lector ya conoce cómo definir con rigor tales funciones (hay diversas maneras de hacerlo, según de dónde se parta en la argumentación lógica), y sólo queremos recordar la gran relación que hay entre la función exponencial y el seno y coseno; por ejemplo, si tomamos $z = x + iy$ y hacemos $e^z = e^x e^{iy}$, la fórmula (10.2) nos da

$$e^z = e^x(\cos(y) + i\,\mathrm{sen}(y)). \tag{10.4}$$

Las funciones e^z, $\cos(z)$ y $\mathrm{sen}(z)$ son enteras (holomorfas en todo $\mathbb{C}$). Señalemos así mismo que el cociente de funciones holomorfas da lugar a una nueva función holo-

[5] El ejemplo arquetípico es la función $f(x) = e^{-1/x^2}$ (y $f(0) = 0$), que es infinitamente derivable como función real y cumple $f^{(n)}(0) = 0$ para todo n; su desarrollo de Taylor en torno a $x_0 = 0$ es $\sum_{n=0}^{\infty} 0 \cdot x^n$, que no converge a $f(x)$ (salvo en $x = 0$). En cambio, si consideramos $f(z) = e^{-1/z^2}$ como función compleja, esa función ni siquiera es continua en $z = 0$ (y por tanto tampoco derivable), luego ya no podemos hablar de que sea holomorfa ni de que tenga un desarrollo en serie de potencias de tipo (10.3).

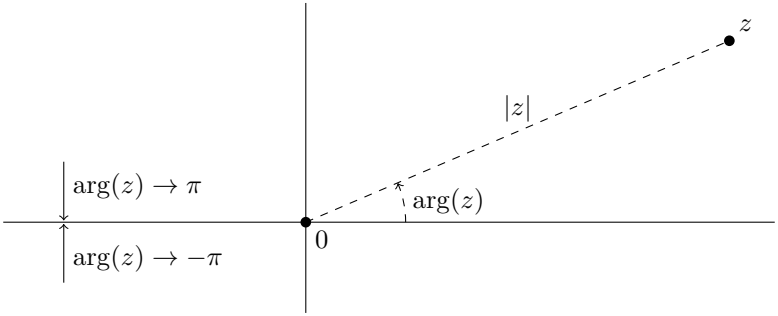


Figura 10.1: En el plano complejo, el argumento presenta una discontinuidad en el semieje de los reales negativos.

morfa en el mismo dominio salvo en los puntos en los que se anula el denominador; esto se aplica, por ejemplo, a cocientes de polinomios y de funciones trigonométricas.

Más sutil —pero muy importante— es la función logaritmo. Comencemos fijándonos en que, para cada $z = x + iy \in \mathbb{C}$, y si tomamos $|z| = \sqrt{x^2 + y^2}$, podremos escribir $z = |z|(\cos(\theta) + i\,\mathrm{sen}(\theta))$ para algún θ; aquí, por supuesto, $|z|$ es el *módulo* y θ el *argumento*. Como el seno y el coseno son periódicos, θ no es único, pero podemos fijarlo eligiéndolo siempre en algún intervalo real de longitud 2π, el que queramos; lo habitual es tomar argumentos en el intervalo $(-\pi, \pi]$, y entonces lo denominamos *argumento principal*, $\arg(z)$. Así las cosas, si nos fijamos en (10.4) nos damos cuenta de que, si queremos conseguir un $w \in \mathbb{C}$ tal que $e^w = z$ (y de esta manera w será el logaritmo de z), bastará tomar $w = \log(|z|) + i\theta$ con θ un argumento de z. Con un argumento elegido en $(a, 2\pi + a]$, en la semirrecta con vértice en el origen y argumento a se perdería incluso la continuidad. En consecuencia, para que el logaritmo pueda ser una función holomorfa, hay que suprimir del dominio una semirrecta con vértice en el origen; lo normal es quitar el semieje de los reales negativos, tal como se ve en la figura 10.1. De esta manera, la función

$$\log(z) = \log(|z|) + i\arg(z)$$

(donde $\log(|z|)$ alude al logaritmo real) es holomorfa en $\mathbb{C} \setminus \{z \in \mathbb{R} : z \leq 0\}$; se denomina *logaritmo principal*.

Una vez que disponemos de logaritmos, se puede definir la raíz cuadrada como $\sqrt{z} = e^{(1/2)\log(z)}$, y de nuevo es holomorfa en $\mathbb{C} \setminus \{z \in \mathbb{R} : z \leq 0\}$. Lo mismo podríamos hacer para definir la función z^α con exponente $\alpha \in \mathbb{C}$ fijo: $z^\alpha = e^{\alpha \log(z)}$. Y, por supuesto, α^z con $\alpha \in \mathbb{C} \setminus \{0\}$ fijo será $\alpha^z = e^{z \log(\alpha)}$.

Una herramienta esencial en análisis complejo es la integración a lo largo de curvas; sin dar demasiados detalles, recordemos los conceptos y resultados principales. Una *curva* (o un *camino*) C en $\mathbb{C}$ es la imagen de una función continua $\varphi\colon [t_1, t_2] \to \mathbb{C}$ (donde $[t_1, t_2]$ es un intervalo real), y exigiremos además que φ sea derivable a trozos con derivada continua (es decir, que exigimos derivada continua en $[t_1, t_2]$ salvo en un número finito de valores del parámetro, que corresponden a esquinas de la curva). Sea ahora $U \subseteq \mathbb{C}$ un abierto, y $f\colon U \subseteq \mathbb{C} \to \mathbb{C}$ una función continua. Si suponemos que la curva C está contenida en U, la integral de f a lo largo de C es, por definición

$$\int_C f(z)\,dz = \int_{t_1}^{t_2} f(\varphi(t))\varphi'(t)\,dt$$

(que, descomponiendo en parte real e imaginaria, tiene sentido como integral real). Cuando la curva C es cerrada (es decir, $\varphi(t_1) = \varphi(t_2)$) a menudo se indica con $\oint_C f(z)\,dz$. En ocasiones se hace integración sobre curvas que no están acotadas, sino que alguna de sus ramas se va hacia el infinito; cuando esto ocurre, se puede interpretar como límite de integrales sobre curvas acotadas. Nótese así mismo que, a menudo, el camino sobre el que se integra se describe en términos del contorno de un recinto en el plano complejo, recorrido en el sentido adecuado (de hecho, incluso puede darse el caso que el contorno se componga de varios caminos «separados», lo cual ocurre cuando el recinto tiene agujeros); en este caso, se suele hablar de integral de contorno.

Una propiedad importante es que —siguiendo con la notación anterior—, si C es un camino de longitud L, y f está acotada mediante $|f(z)| \leq M$ para todo $z \in C$, entonces

$$\left|\int_C f(z)\,dz\right| \leq ML.$$

Con funciones holomorfas, la integral tiene muy buenas propiedades, la primera de la cuales es el teorema integral de Cauchy, que asegura que, si f es holomorfa en un abierto U simplemente conexo (es decir, que no tiene agujeros), entonces

$$\oint_C f(z)\,dz = 0$$

para cualquier curva cerrada C en U. Si además tenemos en cuenta que cuando una curva se recorre en sentido contrario la integral cambia de signo, de aquí se sigue que la integral a lo largo de cualquier camino en U que vaya de un punto z_1 a otro z_2 es independiente del camino, y sólo depende de los puntos z_1 y z_2.

Aún más importante es la fórmula integral de Cauchy. Tal como probó Cauchy alrededor de 1830, si f es una función holomorfa en un abierto U, y C es una curva cerrada simple (es decir, que no se corta a sí misma) que rodea a un punto $z_0 \in U$, se cumplirá

$$f(z_0) = \frac{1}{2\pi i} \int_C \frac{f(z)}{z - z_0}\, dz$$

cuando la curva C se recorre en *sentido positivo* (el contrario a las manecillas del reloj). Aunque se necesita una justificación teórica algo más rigurosa, de aquí se sigue que, si derivamos n veces respecto de z_0 y cada vez metemos la derivada dentro del signo integral, se cumplirá

$$f^{(n)}(z_0) = \frac{n!}{2\pi i} \int_C \frac{f(z)}{(z - z_0)^{n+1}}\, dz;$$

en particular, esto proporciona una forma de calcular los desarrollos (10.3). También se usa para probar que cualquier función holomorfa (derivable en un abierto) es analítica (desarrollable en serie), propiedad que ya hemos mencionado antes. Pero esto es sólo un par de las numerosísimas consecuencias de la fórmula integral de Cauchy, y no es este libro el lugar adecuado para continuar exponiéndolas.

Otro resultado fundamental en teoría de variable compleja es el principio de prolongación analítica. Dos funciones reales —incluso infinitamente derivables— pueden coincidir en un intervalo pero ser distintas fuera de él. Eso no es posible con funciones holomorfas: si dos funciones holomorfas coinciden en un conjunto con puntos de acumulación, forzosamente deben ser la misma en todo el dominio de definición (para evitar trivialidades, en este párrafo damos por hecho que las definiciones están siempre dadas sobre abiertos conexos). No hay ninguna necesidad de que las dos funciones tengan la misma expresión algebraica. De hecho, a menudo ocurre que una expresión algebraica de una función f_1 es válida en un abierto U_1, y otra expresión de otra función f_2 en un abierto U_2; si $f_1 = f_2$ en algún conjunto con puntos de acumulación (incluido, obviamente, en $U_1 \cap U_2$), realmente tenemos una función holomorfa f sobre $U_1 \cup U_2$ (definida como $f = f_1$ en U_1 y $f = f_2$ en U_2), y esta f se denomina prolongación analítica. Esto ocurre, por ejemplo, con la función ζ de Riemann, que se define como $\zeta(z) = \sum_{n=1}^{\infty} 1/n^z$ cuando $\mathrm{Re}(z) > 1$, pero que por prolongación analítica (es decir, mediante otras expresiones válidas en otros dominios que tienen intersección con el de la definición original) se extiende a una función holomorfa en $\mathbb{C} \setminus \{1\}$.

Más generales que las funciones holomorfas son las que su derivabilidad falla en puntos sueltos (sin puntos de acumulaciones) —singularidades aisladas, dicho de manera más técnica—. En las denominadas funciones meromorfas, se permite que

la función tenga polos. Se dice que una función f tiene un *polo* de orden k en z_0 (con k un entero positivo) si es holomorfa en un abierto alrededor de z_0 (sin el punto z_0) y se cumple

$$\lim_{z\to z_0}(z-z_0)^k f(z) = a \neq 0.$$

Cuando esto ocurre, en lugar de un desarrollo como en (10.3), se tiene

$$f(z) = \sum_{n=-k}^{\infty} a_n (z-z_0)^n, \tag{10.5}$$

válido en un entorno alrededor de z_0. Las expresiones de tipo (10.5) se denominan desarrollos de Laurent (aunque aquí no lo vamos a necesitar, los desarrollos de Laurent también pueden tener la forma $\sum_{n=-\infty}^{\infty} a_n(z-z_0)^n$, que se corresponde con *singularidades esenciales*). Obsérvese que los polos son justo lo contrario que los ceros: una función f tiene un cero de orden k en z_0 si y sólo si $1/f$ tiene un polo de orden k en z_0.

Con funciones cuya analiticidad falla en singularidades aisladas, un resultado importante es el teorema de los residuos. En el desarrollo de Laurent (10.5), se denomina *residuo* de f en z_0 al coeficiente a_{-1}, y se denota $\mathrm{Res}(f, z_0)$; se puede calcular como

$$\mathrm{Res}(f, z_0) = \frac{1}{(k-1)!}\lim_{z\to z_0}\frac{d^{k-1}}{dz^{k-1}}\big((z-z_0)^k f(z)\big).$$

Supongamos que f es una función analítica en un abierto U simplemente conexo (sin agujeros), excepto en un número finito de puntos que constituyen singularidades aisladas de la función, y sea C una curva simple cerrada recorrida en sentido positivo, contenida en U y que no pasa por ninguna de las singularidades. El teorema de los residuos asegura que, en estas condiciones,

$$\oint_C f(z)\,dz = 2\pi i \sum_j \mathrm{Res}(f, z_j),$$

donde los z_j son las singularidades que caen dentro de la curva C. (Obsérvese que la fórmula integral de Cauchy es un caso particular de este resultado, pues si f es holomorfa en U, la nueva función $f(z)/(z-z_0)$ tendrá una singularidad en z_0.)

Hay todavía un par de resultados que queremos exponer, pues los utilizaremos en lo que sigue. El primero es un test de convergencia de series denominado criterio M de Weierstrass, y es relativamente sencillo: si tenemos una sucesión de funciones $\{f_n\}_{n=1}^{\infty}$ definidas en un conjunto A, y para cada f_n existe una constante positiva

M_n tal que $|f_n(z)| \leq M_n$ para todo $z \in A$, y se cumple $\sum_{n=1}^{\infty} M_n < \infty$, entonces la serie

$$\sum_{n=1}^{\infty} f_n(z)$$

converge uniformemente en A. Más técnico y específico es lo siguiente: si las funciones $\{f_n\}_{n=1}^{\infty}$ son analíticas en un abierto $A \subseteq \mathbb{C}$, y convergen a f uniformemente sobre compactos contenidos en A, entonces f es analítica en A (por supuesto, aunque este resultado lo hemos enunciado como límite de sucesiones, el comportamiento es el mismo si usamos series, sin más que tomar la sucesión de sumas parciales, o si se trata de productos infinitos).

Y concluimos esta introducción/recordatorio con una aclaración sobre notación. La definición de funciones analíticas (u holomorfas, que es lo mismo) se hace siempre sobre conjuntos abiertos. Cuando decimos que alguna función es analítica en un conjunto que no es abierto, nos referimos a que ese conjunto está incluido en un abierto en el que la función es analítica. En particular, así debe interpretarse cuando decimos que una función es analítica en $\{z \in \mathbb{C} : \mathrm{Re}(z) \geq 0\}$, lo cual nos aparecerá en varias ocasiones a lo largo del tema.

10.2. Definición de la zeta de Riemann y resultados preparatorios

Comencemos nuestro camino hacia el teorema de los números primos definiendo la función zeta de Riemann sobre $\mathbb{C}$; como es tradicional con dicha función, a menudo utilizaremos s para denotar la variable, en lugar del z que empleábamos en la sección anterior.

En principio, la manera más sencilla y habitual de introducir la función zeta de Riemann es por medio de la serie

$$\zeta(s) = \sum_{n=1}^{\infty} \frac{1}{n^s}, \qquad \mathrm{Re}(s) > 1. \tag{10.6}$$

Obsérvese aquí que los sumandos son funciones $n^{-s} = \exp(-s\log(n))$ analíticas (en todo $\mathbb{C}$), y que la serie converge uniformemente en compactos incluidos en $\{s \in \mathbb{C} : \mathrm{Re}(s) > 1\}$, así que $\zeta(s)$ es analítica en esa región del plano complejo. Pero la expresión (10.6) por sí sola no sirve para nuestros objetivos, sino que hay que manipularla de manera adecuada. Esencial en la demostración del teorema de

los números primos es la representaciónde la función zeta de Riemann como un producto infinito en el que aparecen los números primos,

$$\zeta(s) = \prod_p \frac{1}{1-p^{-s}}, \qquad \mathrm{Re}(s) > 1. \tag{10.7}$$

Esta fórmula ya la dedujimos con todo detalle en el tema anterior, en concreto en la sección 9.1; allí lo hacíamos en el contexto de los productos de Euler en general, aunque tomando números reales $s > 1$ en vez de números complejos que satisfacen $\mathrm{Re}(s) > 1$ (dado que $|n^s| = n^{|\mathrm{Re}(s)|}$, eso no supone ninguna diferencia en las argumentaciones). Reproducimos la prueba particularizada a la función $\zeta(s)$ que ahora nos interesa.

Por el teorema fundamental de la aritmética, cada entero positivo se puede expresar de manera única en producto de primos; así, si usamos $\{p_k\}_{k=1}^{\infty}$ para denotar la sucesión de los números primos, para cada entero positivo n podremos poner $n = \prod_k p_k^{a_k}$ (esto es siempre un producto finito pues casi todos los exponentes a_k son nulos). Entonces, dado que la serie $\zeta(s) = \sum_{n=1}^{\infty} n^{-s}$ converge absolutamente cuando $\mathrm{Re}(s) > 1$, podemos reordenar términos y escribir

$$\zeta(s) = \sum_{n=1}^{\infty} n^{-s} = \sum_{a_1,a_2,\ldots\geq 0} \left(p_1^{a_1} p_2^{a_2} \cdots\right)^{-s} = \prod_p \left(\sum_{a\geq 0} p^{-as}\right) = \prod_p \frac{1}{1-p^{-s}},$$

tal como buscábamos.

Con esto, ya estamos en condiciones de dar un primer resultado interesante:

Proposición 10.1.
La función $\zeta(s)$ no se anula cuando $\mathrm{Re}(s) > 1$.

Demostración. Partamos de $\zeta(s)$ como en (10.7), y veamos que su módulo $|\zeta(s)|$ no se anula. Utilizando la desigualdad triangular en cada factor, y dado que $|p^{-s}| = p^{-\mathrm{Re}(s)}$, podemos poner

$$|\zeta(s)| = \prod_p \left|\frac{1}{1-p^{-s}}\right| \geq \prod_p \frac{1}{1+|p^{-s}|} = \prod_p \frac{1}{1+p^{-\mathrm{Re}(s)}},$$

luego basta ver que esta última expresión es positiva siempre que $\mathrm{Re}(s) > 1$.

Para $x \geq 0$ se cumple que $\log(1+x) \leq x$ (esta desigualdad es clara si observamos que $\log(1+x) = \int_1^{1+x} t^{-1}\,dt$ y el valor de la integral es menor que el máximo del integrando multiplicado por la longitud del intervalo), así que $\log(1+p^{-\mathrm{Re}(s)}) \leq$

$p^{-\operatorname{Re}(s)}$ para cualquier primo p. En consecuencia,

$$\sum_p \log\bigl(1+p^{-\operatorname{Re}(s)}\bigr) \le \sum_p p^{-\operatorname{Re}(s)} \le \sum_{n=1}^{\infty} n^{-\operatorname{Re}(s)} < \infty$$

y por tanto

$$\begin{aligned} 0 &< \exp\biggl(-\sum_p \log\Bigl(1+p^{-\operatorname{Re}(s)}\Bigr)\biggr) \\ &= \exp\biggl(\log\Bigl(\prod_p \bigl(1+p^{-\operatorname{Re}(s)}\bigr)^{-1}\Bigr)\biggr) = \prod_p \frac{1}{1+p^{-\operatorname{Re}(s)}}, \end{aligned}$$

tal como queríamos probar. □

No está de más decir aquí que algunos de los temas que interesan sobre la función zeta de Riemann son el cálculo de los valores $\zeta(2k)$ con k entero positivo —en este libro, lo hemos hecho en la sección 6.9— y el estudio de la naturaleza aritmética de $\zeta(2k+1)$ —un problema abierto, según se explica en el apartado 6.8.4—. Pero ninguna de estas cuestiones tiene influencia en la distribución de los números primos, así que no nos volveremos a preocupar por ellas.

10.2.1. Extensión de la zeta de Riemann a la banda crítica

En relación con el teorema de los números primos, no nos podemos contentar con tener la función $\zeta(s)$ definida sobre los s que cumplen $\operatorname{Re}(s) > 1$, sino que se requiere emplear una extensión suya definida sobre los complejos s cuya parte real está entre 0 y 1, la denominada banda crítica. En particular, un hecho relevante será que la función $\zeta(s)$ extendida no se anula cuando $\operatorname{Re}(s) = 1$; obsérvese que tanto la serie (10.6) como el producto infinito (10.7) divergen cuando $\operatorname{Re}(s) \le 1$, luego no podemos utilizar esas expresiones.

Así pues, vamos a ver cómo extender $\zeta(s)$ para valores de s que cumplan $\operatorname{Re}(s) > 0$. De hecho, esto que vamos a hacer ahora es esencialmente lo mismo que hacíamos cuando, por primera vez, definíamos la zeta de Riemann en la sección 8.3 (véase la fórmula (8.19)), y donde aplicábamos directamente la identidad de Abel (más en concreto, la fórmula de sumación de Euler). La diferencia es que allí no nos preocupábamos de la analiticidad, una propiedad que resulta crucial para sacar partido de $\zeta(s)$ en el contexto de las funciones de variable compleja. En todo caso, no nos preocupa repetirnos un poco, dado que, en aras de la legibilidad, queremos

avanzar de la forma más directa posible, intentando evitar —en lo posible— las dependencias entre los distintos temas del libro.

Como $s = 1$ realmente no tiene remedio, la función que vamos a extender no es $\zeta(s)$, sino $\zeta(s) - 1/(s-1)$:

Proposición 10.2.

La función $\zeta(s) - \frac{1}{s-1}$ *se extiende de manera analítica a la región* $\{s \in \mathbb{C} : \mathrm{Re}(s) > 0\}$.

Demostración. Partamos de $\zeta(s) = \sum_{n=1}^{\infty} 1/n^s$, válida siempre que sea $\mathrm{Re}(s) > 1$. Dado que

$$\int_1^{\infty} \frac{dx}{x^s} = \lim_{R\to\infty} \frac{x^{-s+1}}{-s+1}\Big|_{x=1}^{R} = \frac{1}{s-1}, \quad \mathrm{Re}(s) > 1,$$

y como todas las series e integrales que aparecen son absolutamente convergentes, podemos poner

$$\begin{aligned}\zeta(s) - \frac{1}{s-1} &= \sum_{n=1}^{\infty} \frac{1}{n^s} - \int_1^{\infty} \frac{dx}{x^s} = \sum_{n=1}^{\infty} \frac{1}{n^s} - \sum_{n=1}^{\infty} \int_n^{n+1} \frac{dx}{x^s} \\ &= \sum_{n=1}^{\infty} \left(\frac{1}{n^s} - \int_n^{n+1} \frac{dx}{x^s} \right) = \sum_{n=1}^{\infty} \int_n^{n+1} \left(\frac{1}{n^s} - \frac{1}{x^s} \right) dx,\end{aligned}$$

que de nuevo es válido, en principio, cuando $\mathrm{Re}(s) > 1$.

Ahora viene el paso importante, que consiste en probar que

$$\sum_{n=1}^{\infty} \int_n^{n+1} \left(\frac{1}{n^s} - \frac{1}{x^s} \right) dx \tag{10.8}$$

es una función analítica en la región $\{s \in \mathbb{C} : \mathrm{Re}(s) > 0\}$, y que por tanto será la extensión de $\zeta(s) - \frac{1}{s-1}$ buscada. Sabemos que, si una sucesión de funciones analíticas converge uniformemente sobre compactos a otra función, esta última función también es analítica, así que basta probar que cada

$$f_n(s) = \int_n^{n+1} \left(\frac{1}{n^s} - \frac{1}{x^s} \right) dx$$

es analítica en $\{s \in \mathbb{C} : \mathrm{Re}(s) > 0\}$, y que la serie (10.8) converge uniformemente sobre compactos contenidos en esa región. Es claro que cada $f_n(s)$ es analítica en todo $\mathbb{C}$ por serlo las funciones $n^{-s} = \exp(-s\log(n))$ y $x^{-s} = \exp(-s\log(x))$ que se

usan para definirla, así que sólo necesitamos preocuparnos por la convergencia de la serie.

Como

$$\int_n^x u^{-s-1}\,du = -\frac{1}{s}u^{-s}\Big|_{u=n}^{x} = \frac{1}{s}\left(\frac{1}{n^s}-\frac{1}{x^s}\right),$$

podemos poner

$$\begin{aligned}\left|\int_n^{n+1}\left(\frac{1}{n^s}-\frac{1}{x^s}\right)dx\right| &= \left|s\int_n^{n+1}\int_n^x \frac{du}{u^{s+1}}\,dx\right| \le |s|\int_n^{n+1}\int_n^x \frac{du}{|u^{s+1}|}\,dx\\ &= |s|\int_n^{n+1}\int_n^x \frac{du}{u^{1+\mathrm{Re}(s)}}\,dx \le |s|\int_n^{n+1}\int_n^{n+1} \frac{du}{u^{1+\mathrm{Re}(s)}}\,dx\\ &\le |s|\frac{1}{n^{1+\mathrm{Re}(s)}}\int_n^{n+1}\int_n^{n+1} du\,dx = \frac{|s|}{n^{1+\mathrm{Re}(s)}}.\end{aligned}$$

En consecuencia,

$$\sum_{n=1}^{\infty}\left|\int_n^{n+1}\left(\frac{1}{n^s}-\frac{1}{x^s}\right)dx\right| \le \sum_{n=1}^{\infty}\frac{|s|}{n^{1+\mathrm{Re}(s)}} < \infty \quad \text{cuando } \mathrm{Re}(s)>0,$$

y el criterio M de Weierstrass asegura que (10.8) converge uniformemente sobre los compactos contenidos en $\{s\in\mathbb{C}:\mathrm{Re}(s)>0\}$. □

Nota 10.3. Otro procedimiento para extender la definición de $\zeta(s)$ dada en (10.6) —y que es válida para $\mathrm{Re}(s)>1$— hasta $\mathrm{Re}(s)>0$ es usar la identidad

$$\begin{aligned}\left(1-\frac{2}{2^s}\right)\zeta(s) &= 1+\frac{1}{2^s}+\frac{1}{3^s}+\cdots-2\left(\frac{1}{2^s}+\frac{1}{4^s}+\frac{1}{6^s}+\cdots\right)\\ &= \sum_{m=1}^{\infty}\left(\frac{1}{(2m-1)^s}-\frac{1}{(2m)^s}\right),\end{aligned}$$

que en principio es válida cuando $\mathrm{Re}(s)>1$, y comprobar que la serie que estamos obteniendo converge para todo s con parte real positiva (obsérvese que si s es un número real positivo, la expresión anterior da lugar a la serie alternada $\sum_{n=1}^{\infty}(-1)^{n-1}/n^s$, que es convergente, pero no tiene sentido hablar de series alternadas si $s\notin\mathbb{R}$). De hecho, vamos a ver que la serie

$$\sum_{m=1}^{\infty}\left(\frac{1}{(2m-1)^s}-\frac{1}{(2m)^s}\right)$$

es absolutamente convergente en cada compacto de la forma $\{s \in \mathbb{C} : \mathrm{Re}(s) \geq \delta, |s| \leq R\}$ para cualesquiera $\delta > 0$ y $R > 0$. En efecto,

$$\left|\frac{1}{(2m-1)^s} - \frac{1}{(2m)^s}\right| = \left|s\int_{2m-1}^{2m} \frac{dx}{x^{s+1}}\right| \leq \frac{|s|}{(2m-1)^{\mathrm{Re}(s)+1}} \leq \frac{R}{(2m-1)^{\delta+1}},$$

y como

$$\sum_{m=1}^{\infty} \frac{R}{(2m-1)^{\delta+1}} < \infty,$$

el criterio M de Weierstrass asegura la convergencia uniforme en el compacto. Usando que la convergencia uniforme sobre compactos de funciones analíticas da como límite una función analítica, hemos extendido $(1-2^{1-s})\zeta(s)$ a una función analítica en $\mathrm{Re}(s) > 0$; por ende, lo mismo ocurrirá con $\zeta(s)$ salvo en el punto $s = 1$ y en el resto de los $s \in \mathbb{C}$ tales que $2^{1-s} = 1$ (dado que, por definición, $2^{1-s} = e^{(1-s)\log(2)}$, y $e^z = 1$ si y sólo si $z = 2k\pi i$ con $k \in \mathbb{Z}$, estos valores de s serán $s = 1 + (2k\pi i)/\log(2)$ con $k \in \mathbb{Z}$). Así pues, este procedimiento de extensión de $\zeta(s)$ no es tan efectivo como el de la proposición 10.2. ◁

10.2.2. Extensión de algunas propiedades a la frontera derecha de la banda crítica

El objetivo de este apartado es ver que algunas propiedades que cumple $\zeta(s)$ —o ciertas funciones relacionadas que nos aparecerán enseguida— sobre $\{s \in \mathbb{C} : \mathrm{Re}(s) > 1\}$ se pueden extender a $\{s \in \mathbb{C} : \mathrm{Re}(s) \geq 1\}$. En particular, en la proposición 10.1 ya vimos que la zeta de Riemann no tenía ceros cuando $\mathrm{Re}(s) > 1$; demostraremos ahora que tampoco los tiene en la frontera $\mathrm{Re}(s) = 1$. Esta propiedad es esencial, y de hecho la ausencia de ceros de $\zeta(s)$ en la recta vertical $\mathrm{Re}(s) = 1$ es la clave —tras un cuidadoso tratamiento con técnicas de variable compleja— de la demostración analítica del teorema de los números primos. La probaron tanto Hadamard como De la Vallée-Poussin en sus correspondientes demostraciones del teorema de los números primos en 1896.

Hasta poder abordar todo eso necesitamos cierto trabajo preparatorio. Nos vamos a ayudar de la función $\Phi(s)$ que se define como

$$\Phi(s) = \sum_{p} \frac{\log(p)}{p^s}, \qquad \mathrm{Re}(s) > 1. \tag{10.9}$$

En compactos incluidos en $\{s \in \mathbb{C} : \mathrm{Re}(s) > 1\}$, ésta es una serie que converge uniformemente, luego $\Phi(s)$ así definida es analítica en $\{s \in \mathbb{C} : \mathrm{Re}(s) > 1\}$.

Una de las virtudes de usar $\zeta(s)$ como en (10.7) es que, al tomar logaritmos (logaritmos complejos, no reales), se transforma en una suma que, derivando, podremos relacionar fácilmente con $\Phi(s)$. En efecto, tomando logaritmos en (10.7) queda

$$-\log(\zeta(s)) = \sum_p \log(1-p^{-s})$$

y, derivando respecto a s, obtenemos

$$\begin{aligned} -\frac{\zeta'(s)}{\zeta(s)} &= \sum_p \frac{1}{1-p^{-s}}\, p^{-s}\log(p) = \sum_p \frac{\log(p)}{p^s-1} \\ &= \sum_p \left(\frac{\log(p)}{p^s} + \frac{\log(p)}{p^s(p^s-1)}\right) = \Phi(s) + \sum_p \frac{\log(p)}{p^s(p^s-1)}. \end{aligned}$$

Así pues,

$$\Phi(s) = -\frac{\zeta'(s)}{\zeta(s)} - \sum_p \frac{\log(p)}{p^s(p^s-1)}, \qquad \mathrm{Re}(s) > 1.$$

Fijémonos ahora en los dos sumandos de la derecha. Respecto al primero, ya sabíamos que $\zeta(s)$ estaba definida y era analítica en $\{s \in \mathbb{C} : \mathrm{Re}(s) > 0\} \setminus \{1\}$ (proposición 10.2), así que lo mismo podemos decir de $\zeta'(s)/\zeta(s)$, salvo en los ceros de ζ en los que se perderá la analiticidad. Respecto al segundo sumando, es claro que

$$\sum_p \frac{\log(p)}{p^s(p^s-1)}$$

está definida y es analítica en $\{s \in \mathbb{C} : \mathrm{Re}(s) > 1/2\}$ (de nuevo por la convergencia uniforme sobre compactos). Estos argumentos hacen que

$$\Phi(s) = -\frac{\zeta'(s)}{\zeta(s)} - \sum_p \frac{\log(p)}{p^s(p^s-1)}, \qquad \mathrm{Re}(s) > 1/2, \tag{10.10}$$

sea una extensión de $\Phi(s)$, analítica en $\{s \in \mathbb{C} : \mathrm{Re}(s) > 1/2\}$ salvo en $s = 1$ (polo de ζ) y en los ceros de ζ en $\{s \in \mathbb{C} : \mathrm{Re}(s) > 1/2\}$ (si los hubiera).

Antes de proseguir —y dejando momentáneamente de lado la extensión de Φ que acabamos de construir—, veamos que la función Φ tal como se define en (10.9) cumple una propiedad que, pese a su extraña apariencia inicial, nos resultará útil más adelante:

Lema 10.4.

Sea un número real $\sigma > 1$. Entonces, cualquiera que sea $\alpha \in \mathbb{R}$, la función Φ cumple

$$0 \leq \sum_{r=-2}^{2} \binom{4}{2+r} \Phi(\sigma + ir\alpha). \tag{10.11}$$

Demostración. Observemos que

$$\begin{aligned} 0 \leq \left(2\,\mathrm{Re}\left(p^{i\alpha/2}\right)\right)^4 &= \left(p^{i\alpha/2} - p^{-i\alpha/2}\right)^4 = \sum_{r=0}^{4} \binom{4}{r} p^{i(4-r)\alpha/2} p^{-ir\alpha/2} \\ &= \binom{4}{0} p^{2i\alpha} + \binom{4}{1} p^{i\alpha} + \binom{4}{2} 1 + \binom{4}{3} p^{-i\alpha} + \binom{4}{4} p^{-2i\alpha} \\ &= \sum_{r=-2}^{2} \binom{4}{2+r} \frac{1}{p^{ir\alpha}}. \end{aligned}$$

Ahora, multiplicando esta desigualdad por $\log(p)/p^{\sigma}$, queda

$$0 \leq \sum_{r=-2}^{2} \binom{4}{2+r} \frac{\log(p)}{p^{\sigma+ir\alpha}}.$$

Finalmente, sumando sobre los números primos obtenemos (10.11). □

Para sacar más partido a la función Φ, nos serviremos así mismo de un par de propiedades sencillas sobre ceros de funciones analíticas. Veámoslas.

En primer lugar, supongamos que f es una función analítica que tiene un cero de orden $k \in \mathbb{N}$ en z_0, es decir, que en algún entorno de z_0, f se puede expresar mediante el desarrollo

$$f(z) = a_k(z - z_0)^k + a_{k+1}(z - z_0)^{k+1} + a_{k+2}(z - z_0)^{k+2} + \cdots \tag{10.12}$$

con $a_k \neq 0$. Entonces, de aquí se deduce que

$$\lim_{z \to z_0} \frac{f(z)}{(z - z_0)^{k-1}} = 0 \quad \text{pero} \quad \lim_{z \to z_0} \frac{f(z)}{(z - z_0)^k} = a_k \neq 0. \tag{10.13}$$

En segundo lugar, si sabemos que f analítica tiene un cero en z_0, una manera de determinar su orden k es mediante la expresión

$$\lim_{z \to z_0} (z - z_0) \frac{f'(z)}{f(z)} = k. \tag{10.14}$$

En efecto, de (10.12) se sigue que

$$f'(z) = ka_k(z-z_0)^{k-1} + (k+1)a_{k+1}(z-z_0)^k + (k+2)a_{k+2}(z-z_0)^{k+1} + \cdots,$$

y por tanto, simplificando $(z-z_0)^k$ de numerador y denominador,

$$\begin{aligned}\lim_{z\to z_0}(z-z_0)\frac{f'(z)}{f(z)} &= \lim_{z\to z_0}\frac{ka_k(z-z_0)^k + (k+1)a_{k+1}(z-z_0)^{k+1}+\cdots}{a_k(z-z_0)^k + a_{k+1}(z-z_0)^{k+1}+\cdots}\\ &= \lim_{z\to z_0}\frac{ka_k + (k+1)a_{k+1}(z-z_0) + (k+2)a_{k+2}(z-z_0)^2+\cdots}{a_k + a_{k+1}(z-z_0) + a_{k+2}(z-z_0)^2+\cdots} = k,\end{aligned}$$

tal como buscábamos.

Una vez visto cómo caracterizar un cero de una función analítica y su orden mediante (10.13) y (10.14), ya estamos en condiciones de dar un nuevo paso, que muestra que los ceros de ζ en la recta $\mathrm{Re}(z) = 1$, si los hubiera, siempre están emparejados (cada uno con su conjugado).

Lema 10.5.

Si $1+i\alpha$ (con $\alpha \in \mathbb{R}\setminus\{0\}$) es un cero de orden k de la zeta de Riemann, entonces $1-i\alpha$ es otro cero, y del mismo orden.

Demostración. Si, como es habitual, denotamos al conjugado de $a+bi$, con $a, b \in \mathbb{R}$, mediante $\overline{a+bi} = a-bi$, el complejo $e^z = e^{x+iy} = e^x(\cos(y) + i\,\mathrm{sen}(y))$, con $z = x+iy$, $x, y \in \mathbb{R}$, cumple $\overline{e^z} = e^x(\cos(y) - i\,\mathrm{sen}(y)) = e^{\overline{z}}$. Y de aquí, por ser $n^s = e^{s\log(n)}$, también se cumplirá $\overline{n^s} = n^{\overline{s}}$. En consecuencia,

$$\overline{\zeta(s)} = \overline{\sum_{n=1}^{\infty}\frac{1}{n^s}} = \sum_{n=1}^{\infty}\frac{1}{\overline{n^s}} = \sum_{n=1}^{\infty}\frac{1}{n^{\overline{s}}} = \zeta(\overline{s}),$$

válido en $\{s \in \mathbb{C} : \mathrm{Re}(s) > 1\}$ (que es cuando tiene sentido la serie que acabamos de usar); y la igualdad se extiende a $\{s \in \mathbb{C}\setminus\{1\} : \mathrm{Re}(s) \geq 1\}$ por continuidad.

Con esto, supongamos que $1+i\alpha$, $\alpha \neq 0$, es un cero de orden k de ζ. Como $\zeta(\overline{s}) = \overline{\zeta(s)}$, $1-i\alpha$ deberá ser otro cero. Veamos que del mismo orden. Si nos acercamos a $1+i\alpha$ mediante números de la forma $1+\varepsilon+i\alpha$ con $\varepsilon > 0$, la caracterización del orden dada en (10.13) muestra que

$$\lim_{\varepsilon\to 0^+}\left(\frac{1}{\varepsilon}\right)^{k-1}\zeta(1+\varepsilon+i\alpha) = 0, \qquad \lim_{\varepsilon\to 0^+}\left(\frac{1}{\varepsilon}\right)^{k}\zeta(1+\varepsilon+i\alpha) \neq 0.$$

Entonces,

$$\lim_{\varepsilon\to 0^+}\left(\frac{1}{\varepsilon}\right)^{k-1}\zeta(1+\varepsilon-i\alpha)=\overline{\lim_{\varepsilon\to 0^+}\left(\frac{1}{\varepsilon}\right)^{k-1}\zeta(1+\varepsilon+i\alpha)}=\overline{0}=0$$

y

$$\lim_{\varepsilon\to 0^+}\left(\frac{1}{\varepsilon}\right)^{k}\zeta(1+\varepsilon-i\alpha)=\overline{\lim_{\varepsilon\to 0^+}\left(\frac{1}{\varepsilon}\right)^{k}\zeta(1+\varepsilon+i\alpha)}\neq 0,$$

luego el cero $1-i\alpha$ también tiene orden k. □

Nota 10.6. Como parte de la demostración del lema anterior hemos probado que

$$\overline{\zeta(s)}=\zeta(\overline{s})$$

en $\{s\in\mathbb{C}:\operatorname{Re}(s)>1\}$. Si, para ello, en lugar de usar la expresión $\zeta(s)=\sum_{n=1}^{\infty}1/n^s$, que es válida cuando $\operatorname{Re}(s)>1$, utilizamos

$$\zeta(s)=\left(1-\frac{2}{2^s}\right)^{-1}\sum_{m=1}^{\infty}\left(\frac{1}{(2m-1)^s}-\frac{1}{(2m)^s}\right),$$

que, tal como hemos visto en la nota 10.3, es válida cuando $\operatorname{Re}(s)>0$ (excepto cuando $2^{1-s}=1$), obtenemos que $\overline{\zeta(s)}=\zeta(\overline{s})$ en $\{s\in\mathbb{C}:\operatorname{Re}(s)>0\}\setminus\{1+2k\pi i/\log(2):k\in\mathbb{Z}\}$. Esto mismo —e incluso más— se puede conseguir, de una forma alternativa, mediante un trabajo algo más teórico, como vemos a continuación.

Comencemos observando que, dada un función meromorfa $f(z)$ en un conjunto $A\subseteq\mathbb{C}$ simétrico respecto al eje real (es decir, que cumple $z\in A\iff\overline{z}\in A$), la función $\overline{f(\overline{z})}$ también es meromorfa en A (sin la simetría respecto al eje real, lo que obtendríamos es que esta nueva función sería meromorfa en $\overline{A}$). Para comprobarlo, veamos que, para cada $z_0\in A$, existe un entorno de z_0 en el que $\overline{f(\overline{z})}$ tiene un desarrollo en serie válido. Como $\overline{z_0}\in A$ y f es meromorfa en A, existirá un entorno de $\overline{z_0}$ en el que

$$f(z)=\sum_{n=-k}^{\infty}a_n(z-\overline{z_0})^n;$$

y de aquí se desprende inmediatamente que

$$\overline{f(\overline{z})}=\sum_{n=-k}^{\infty}\overline{a_n}(z-z_0)^n$$

en el correspondiente «entorno conjugado».

Entonces, basta utilizar que las funciones meromorfas $\zeta(s)$ y $\overline{\zeta(\overline{s})}$ son iguales si $\mathrm{Re}(s) > 1$ (lo cual es inmediato por la expresión de la serie $\zeta(s) = \sum_{n=1}^{\infty} 1/n^s$), luego son iguales en $\{s \in \mathbb{C} : \mathrm{Re}(s) > 0\} \setminus \{1\}$ por prolongación analítica (o incluso en $\mathbb{C} \setminus \{1\}$, pues a todo ese conjunto se puede extender ζ de manera holomorfa, aunque aquí no lo hemos probado).

De todas formas, no está de más advertir que hay que ser cuidadoso con los argumentos de extensión analítica. Por ejemplo, fíjese el lector en que los logaritmos cumplen algo como lo anteriormente expuesto en $\mathrm{Re}(s) > 0$ pero no en $\mathrm{Re}(s) < 0$, que depende de la determinación del logaritmo que usemos. Lo que sucede es que con logaritmos no se puede usar la argumentación anterior pues las dos extensiones no son analíticas en la misma zona de $\mathbb{C}$; a cada una de ellas le falla una semirrecta distinta. ◁

Una vez allanado el terreno, éste es el resultado que conseguimos:

Proposición 10.7.
En $\{s \in \mathbb{C} : \mathrm{Re}(s) \geq 1\}$, $\zeta(s)$ no tiene ceros y la función

$$\Phi(s) - \frac{1}{s-1}$$

es analítica.

Demostración. Comencemos analizando el punto $s = 1$, y para ello fijémonos en la expresión (10.10). Dado que $\zeta(s)$ tiene un polo en $s = 1$, tiene que haber un entorno de 1 en el que no se anula; así obtenemos que

$$\Phi(s) = -\frac{\zeta'(s)}{\zeta(s)} - \sum_p \frac{\log(p)}{p^s(p^s-1)} \tag{10.15}$$

es analítica en un entorno de 1 salvo en $s = 1$. Vamos a ver que este polo se elimina restando $1/(1-s)$, con lo cual ya tendremos que $\Phi(s) - \frac{1}{s-1}$ es analítica en un entorno de 1.

Por la proposición 10.2 podemos escribir

$$\zeta(s) = \frac{1}{s-1} + g(s)$$

donde $g(s)$ es una función analítica en un entorno de 1 (de hecho, en $\{s \in \mathbb{C} : \mathrm{Re}(s) > 0\}$), con lo cual

$$\zeta'(s) = \frac{-1}{(s-1)^2} + g'(s).$$

Y sin más que sustituir esta $\zeta'(s)$ en (10.15) obtenemos

$$\begin{aligned}\Phi(s)-\frac{1}{1-s} &= -\frac{-1/(s-1)^2+g'(s)}{\zeta(s)}-\frac{1}{s-1}-\sum_p \frac{\log(p)}{p^s(p^s-1)}\\ &= -\frac{-1/(s-1)+(s-1)g'(s)+\zeta(s)}{(s-1)\zeta(s)}-\sum_p \frac{\log(p)}{p^s(p^s-1)}\\ &= \frac{(s-1)g'(s)+g(s)}{1+(s-1)g(s)}-\sum_p \frac{\log(p)}{p^s(p^s-1)},\end{aligned}$$

que efectivamente es analítica en un entorno de 1.

Veamos a continuación qué pasa en $\{s\in\mathbb{C}:\operatorname{Re}(s)\geq 1\}\setminus\{1\}$. Supongamos que $1+i\alpha$, con $\alpha\in\mathbb{R}\setminus\{0\}$, es un cero de ζ, y sea k su orden. Para ver que eso no es posible usaremos un ingenioso procedimiento en el que intervendrán los lemas 10.4 y 10.5. Sea j el orden de $1+2i\alpha$ como cero de ζ, o tomemos $j=0$ si ζ no se anula en $1+2i\alpha$. Por el lema 10.5, $1-i\alpha$ y $1-2i\alpha$ también serán ceros de orden k y j respectivamente (y de nuevo estamos usando $j=0$ con el significado de no ser cero).

Vamos ahora a aproximarnos a los cuatro puntos $1\pm i\alpha$ y $1\pm 2i\alpha$ por la derecha, tomando $1+\varepsilon\pm i\alpha$ y $1+\varepsilon\pm 2i\alpha$, con $\varepsilon>0$, y haciendo $\varepsilon\to 0$; además, también lo vamos a hacer para $s=1$. Comencemos con este último (que en realidad no requiere $s>0$). Dado que $\Phi(s)-\frac{1}{s-1}$ es analítica en un entorno de 1, es claro que $\lim_{\varepsilon\to 0}\varepsilon(\Phi(1+\varepsilon)-1/\varepsilon)=0$, luego

$$\lim_{\varepsilon\to 0}\varepsilon\Phi(1+\varepsilon)=1.$$

Para $1\pm i\alpha$ y $1\pm 2i\alpha$, utilicemos que el orden de un cero se puede detectar mediante (10.14) (en un «cero de orden 0», la expresión (10.14) resulta trivial). Entonces, usando (10.15) y dado que

$$\sum_p \frac{\log(p)}{p^s(p^s-1)}$$

converge cuando $\operatorname{Re}(s)>1/2$, tendremos

$$\lim_{\varepsilon\to 0^+}\varepsilon\Phi(1+\varepsilon\pm i\alpha)=-\lim_{\varepsilon\to 0^+}\varepsilon\frac{\zeta'(1+\varepsilon\pm i\alpha)}{\zeta(1+\varepsilon\pm i\alpha)}+0=-k$$

y análogamente

$$\lim_{\varepsilon\to 0^+}\varepsilon\Phi(1+\varepsilon\pm 2i\alpha)=-j.$$

Utilicemos ahora el lema 10.4 con $\sigma = 1+\varepsilon$, que nos proporciona la relación

$$0 \leq \sum_{r=-2}^{2} \binom{4}{2+r} \Phi(1+\varepsilon+ir\alpha),$$

válida para cualquier $\varepsilon > 0$. Multiplicando ambos lados de esta desigualdad por ε y haciendo tender $\varepsilon \to 0^+$ llegamos a

$$\begin{aligned} 0 &\leq \lim_{\varepsilon\to 0^+} \sum_{r=-2}^{2} \binom{4}{2+r} \varepsilon\Phi(1+\varepsilon+ir\alpha) \\ &= -\binom{4}{0} j - \binom{4}{1} k + \binom{4}{2} - \binom{4}{3} k - \binom{4}{4} j = -2j - 8k + 6, \end{aligned}$$

así que forzosamente debe ser $k = 0$, con lo cual $1 + i\alpha$ no es un cero de ζ.

Una vez probado que $\zeta(s)$ no tiene ningún cero con $\operatorname{Re}(s) = 1$, la analiticidad de $\Phi(s) - \frac{1}{s-1}$ es ya inmediata por (10.10). □

10.2.3. Sumas de logaritmos de primos

Vamos ahora a usar la función $\vartheta(x)$ dada por

$$\vartheta(x) = \sum_{p\leq x} \log(p), \qquad x > 0, \tag{10.16}$$

que fue definida por Chebyshev en sus intentos de probar el teorema de los números primos y que nosotros ya introdujimos en el apartado 8.7.1. De hecho, allí veíamos algunas propiedades de esta función, así como su relación con el teorema de los números primos. Pero no vamos a utilizar ahora nada de lo que hacíamos en la sección 8.7 —toda ella dedicada a mostrar resultados equivalentes al teorema de los números primos—, sino que, con el objetivo de conseguir una demostración lo más directa posible, probaremos aquí lo que vayamos a necesitar.

En nuestro camino hacia el teorema de los números primos, la función $\vartheta(x)$ va a ser muy importante; lo que interesa de ella es su comportamiento cuando $x \to \infty$. El primer resultado —que, además, nos va a servir para explicar hacia dónde nos dirigimos— es relativamente sencillo:

Proposición 10.8.
Cuando $x \to \infty$, la función $\vartheta(x)$ cumple $\vartheta(x) = O(x)$.

Demostración. Vamos a comenzar probando que

$$\vartheta(2x) - \vartheta(x) \leq 4x$$

para cada $x > 0$. Como esa desigualdad es trivial cuando $0 < x < 1$ (pues entonces $\vartheta(2x) = \vartheta(x) = 0$), podemos no preocuparnos por ese caso.

Desarrollando mediante el binomio de Newton tenemos

$$2^{2x} = (1+1)^{2x} \geq (1+1)^{2\lfloor x \rfloor} = \sum_{j=0}^{2\lfloor x \rfloor} \binom{2\lfloor x \rfloor}{j} \geq \binom{2\lfloor x \rfloor}{\lfloor x \rfloor} = \frac{(2\lfloor x \rfloor)!}{\lfloor x \rfloor!\, \lfloor x \rfloor!}.$$

Entonces, como

$$\prod_{\lfloor x \rfloor < p \leq 2\lfloor x \rfloor} p \quad \text{divide a} \quad \frac{(2\lfloor x \rfloor)!}{\lfloor x \rfloor!\, \lfloor x \rfloor!}$$

(todos esos primos p están como factores en el numerador pero no en el denominador), deberá cumplirse

$$\prod_{\lfloor x \rfloor < p \leq 2\lfloor x \rfloor} p \leq \frac{(2\lfloor x \rfloor)!}{\lfloor x \rfloor!\, \lfloor x \rfloor!} \leq 2^{2x}$$

y, tomando logaritmos,

$$\sum_{\lfloor x \rfloor < p \leq 2\lfloor x \rfloor} \log(p) \leq 2x \log(2).$$

De aquí que (obsérvese que $x < p$ es equivalente a $\lfloor x \rfloor < p$, y que el cambio de $p \leq 2x$ por $p \leq 2\lfloor x \rfloor$ se reduce a, como máximo, añadir un sumando)

$$\begin{aligned} \vartheta(2x) - \vartheta(x) = \sum_{x < p \leq 2x} \log(p) &\leq \sum_{\lfloor x \rfloor < p \leq 2\lfloor x \rfloor} \log(p) + \log(2x) \\ &\leq 2x \log(2) + \log(2x) \leq 2x + 2x = 4x, \end{aligned}$$

tal como buscábamos.

El paso final de la demostración utiliza reiteradamente la desigualdad anterior, $\vartheta(2x) - \vartheta(x) \leq 4x$, de manera ingeniosa. Si usamos x en lugar de $2x$, hemos probado que

$$\vartheta(x) - \vartheta(x/2) \leq 2x$$

y, por idéntico motivo,

$$\vartheta\left(\frac{x}{2^k}\right) - \vartheta\left(\frac{x}{2^{k+1}}\right) \leq \frac{x}{2^{k-1}} \quad \text{para cada } k \in \mathbb{N}.$$

Como consecuencia,

$$\vartheta(x) = \sum_{k=0}^{\infty} \left(\vartheta\left(\frac{x}{2^k}\right) - \vartheta\left(\frac{x}{2^{k+1}}\right)\right) \leq \sum_{k=0}^{\infty} \frac{x}{2^{k-1}} = 4x,$$

así que efectivamente $\vartheta(x) = O(x)$. □

Algo más potente y preciso es el hecho de que

$$\int_1^{\infty} \frac{\vartheta(x) - x}{x^2}\, dx < \infty. \tag{10.17}$$

Decimos que es más preciso que la proposición 10.8 puesto que, si fuese $\vartheta(x) \sim Cx$, que es compatible con el enunciado de la proposición, la integral (10.17) sería divergente salvo para $C = 1$. Pero demostrar la convergencia de esa integral no es tan sencillo, y vamos a tener que recurrir de nuevo a la variable compleja; lo haremos en la sección 10.3, utilizando lo que se denomina un *teorema tauberiano*.

Una vez establecido (10.17), ya no resultará complicado ver que, de hecho, $\vartheta(x) \sim x$; y de ahí en seguida se deducirá el teorema de los números primos. En realidad, ya vimos en la sección 8.7 cómo se podía dar ese paso final de llegar al teorema de los números primos partiendo de $\vartheta(x) \sim x$; pero no lo vamos a aplicar ahora sin más, sino que reharemos lo que vayamos a usar. (En concreto, en el teorema 8.28 probábamos que $\vartheta(x) \sim x$ y $\pi(x) \sim x/\log(x)$ eran resultados equivalentes.)

10.3. Un teorema tauberiano

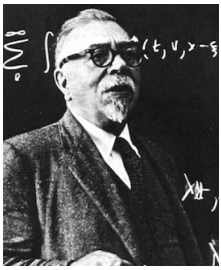
Es bien conocido —y muy sencillo de comprobar— que si una sucesión de números $\{s_k\}_{k=1}^{\infty}$ tiene límite S entonces sus medias de Cesàro convergen al mismo límite,

$$\lim_n \frac{s_1 + s_2 + \cdots + s_n}{n} = S;$$

por expresarlo de alguna forma, un resultado «fuerte» implica uno «débil» (y a menudo se recurre a probar resultados débiles cuando los fuertes no son sencillos de demostrar o fallan: de hecho, ésta es la motivación que se encuentra detrás del tema 8, como allí vimos). Lo mismo ocurre con muchos otros métodos de calcular límites, sumar series o evaluar integrales; las convergencias fuertes siempre implican convergencias más débiles. Otro ejemplo de este tipo es el teorema de Abel,



Alfred Tauber
(1866–1942)



Norbert Wiener
(1894–1964)

que establece que si una serie $\sum_{n=0}^{\infty} a_n x^n$ converge para $x \in (-1,1)$, y $\sum_{n=0}^{\infty} a_n$ converge, entonces[6]

$$\lim_{x\to 1^-} \sum_{n=0}^{\infty} a_n x^n = \sum_{n=0}^{\infty} a_n.$$

Lo extraordinario es cuando se produce lo contrario, cuando una convergencia fuerte se puede deducir de una débil. Por supuesto, si tal fuese el comportamiento sistemático de dos tipos de convergencia, ambos serían equivalentes, pero no es eso lo que aquí queremos decir, sino que lo que interesa es que una convergencia débil implique la convergencia fuerte en determinados casos especiales, expresables por medio de hipótesis adicionales.

El teorema original que Alfred Tauber probó en 1897 estaba descrito en forma de sumabilidad Abel, y decía que si la serie

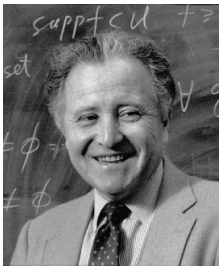
$$f(x) = \sum_{n=0}^{\infty} a_n x^n$$

converge en el intervalo $(-1,1)$, $f(x) \to A$ cuando $x \to 1^-$, y los coeficientes de la serie cumplen $a_n = o(1/n)$, entonces $\sum_{n=0}^{\infty} a_n$ converge a A.[7] Con medias de Cesàro, Hardy probó en 1910 que, si tenemos una sucesión $\{s_n\}_{n=1}^{\infty}$ tal que

$$\lim_{n} \frac{s_1 + s_2 + \cdots + s_n}{n} = S,$$

[6] El teorema de Abel, aplicado a las series de Taylor de $\log(1+x)/x$ y $\operatorname{arc\,tg}(x)$, se usa para probar que $\sum_{n=0}^{\infty}(-1)^k/(k+1) = \log(2)$ y $\sum_{n=0}^{\infty}(-1)^k/(2k+1) = \operatorname{arc\,tg}(1) = \pi/4$.

[7] Aunque no tiene influencia en lo que estamos tratando, resulta interesante comentar que, en 1911, Littlewood mejoró el resultado original de Tauber sustituyendo la hipótesis $a_n = o(1/n)$ por $a_n = O(1/n)$, que es más débil (pero a costa de una demostración considerablemente más complicada).



Donald J. Newman
(1930–2007)



Pierre Simon Laplace
(1749–1827)

y además $s_n - s_{n-1} = O(1/n)$, entonces también $\lim_n s_n = S$.[8] En honor de Tauber, los resultados en los que, a partir de algún tipo de convergencia débil de sucesiones, series o integrales se deduce, con alguna hipótesis añadida, otra más fuerte, se denominan teoremas tauberianos. (Así mismo, se suele decir que los teoremas tauberianos son inversos de los teoremas abelianos.)

Los teoremas tauberianos son muy importantes para lo que ahora nos concierne, pues un ingrediente que forma parte de todas las demostraciones analíticas del teorema de los números primos es el uso de algún tipo de teorema tauberiano.

Muchas demostraciones analíticas usan el denominado teorema tauberiano de Wiener-Ikehara, que fue probado por Norbert Wiener y su estudiante Shikao Ikehara en 1932, y sirvió para simplificar considerablemente las demostraciones del teorema de los números primos que entonces se conocían. En concreto, el teorema de Wiener-Ikehara proporcionó la primera prueba que —aparte de argumentos de analiticidad sobre la función zeta de Riemann— sólo necesitaba usar que $\zeta(s)$ no tenía ceros en la línea $\mathrm{Re}(s) = 1$ del plano complejo; las anteriores pruebas requerían, además, estimaciones sobre el crecimiento de $\zeta(s)$ en esa línea o cerca de ella. Aun así, demostrar el teorema tauberiano de Wiener-Ikehara y aplicarlo seguía teniendo considerables dificultades técnicas.

En su lugar, Donald J. Newman utilizó el teorema tauberiano de Albert E. Ingham sobre series de Dirichlet, que data de 1935 y dice lo siguiente: Supongamos que $|a_n| \leq 1$ para todo n y formemos la serie $\sum_n a_n n^{-z}$, que claramente converge a una función analítica $F(z)$ cuando $\mathrm{Re}(z) > 1$; si, de hecho, $F(z)$ es analítica cuando

[8] Este resultado tiene una descripción equivalente en términos de sumabilidad de series: supongamos que tenemos una serie $\sum_{k=1}^{\infty} a_k$ cuya convergencia queremos estudiar, y sea $\{s_n\}_{n=1}^{\infty}$ la sucesión de sus sumas parciales, es decir, $s_n = \sum_{k=1}^{n} a_k$ para cada $n \in \mathbb{N}$. En esas condiciones, si $\lim_n (s_1 + s_2 + \cdots + s_n)/n = S$ y $a_n = O(1/n)$, también se cumplirá $\sum_{k=1}^{\infty} a_k = S$.

$\mathrm{Re}(z) \geq 1$, entonces $\sum_n a_n n^{-z}$ converge para $\mathrm{Re}(z) \geq 1$. El mérito de Newman fue proporcionar —en [73], publicado en 1980— una demostración de este resultado utilizando sólo la fórmula integral de Cauchy y estimaciones sencillas en una integral de contorno

$$\int_\Gamma f(z) N^z \left(\frac{1}{z} + \frac{z}{R^2}\right) dz, \qquad N \in \mathbb{N},$$

donde Γ es cierto contorno finito en el plano complejo. Dicha demostración era mucho más sencilla que la original de Ingham.

Nosotros, siguiendo a [104], utilizaremos el método de Newman pero adaptado a integrales de Laplace (también denominadas transformadas de Laplace)[9]

$$g(z) = \int_0^\infty f(t) e^{-zt}\, dt$$

en lugar de a series de Dirichlet.

En concreto, el teorema tauberiano que aquí utilizaremos es el siguiente:

Teorema 10.9.
Sea $f : [0, \infty) \to \mathbb{R}$ una función acotada y localmente integrable (es decir, que cumple $\int_0^x |f(t)|\, dt < \infty$ para cada $x > 0$). Si la función

$$g(z) = \int_0^\infty f(t) e^{-zt}\, dt$$

definida en $\{z \in \mathbb{C} : \mathrm{Re}(z) > 0\}$ tiene extensión analítica a

$$\{z \in \mathbb{C} : \mathrm{Re}(z) \geq 0\},$$

entonces

$$\int_0^\infty f(t)\, dt$$

existe (y es igual a $g(0)$).

Demostración. Para $T > 0$, sea

$$g_T(z) = \int_0^T f(t) e^{-zt}\, dt, \qquad z \in \mathbb{C};$$

[9] El estudio de las transformadas de Laplace constituye, por sí mismo, un amplio campo de las matemáticas, con aplicaciones muy distintas de lo que ahora nos interesa. En particular, las transformadas de Laplace se utilizan como método auxiliar para resolver ecuaciones diferenciales.

para cada z, existe $g_T'(z) = -\int_0^T f(t)te^{-zt}\,dt$ (se puede derivar respecto de z bajo el signo integral), y por tanto g_T es una función holomorfa en todo $\mathbb{C}$. Con esta notación, y por definición de integral impropia,

$$\int_0^\infty f(t)\,dt = \lim_{T\to\infty}\int_0^T f(t)\,dt = \lim_{T\to\infty} g_T(0),$$

así que el teorema estará demostrado si comprobamos que existe $\lim_{T\to\infty} g_T(0)$ y su valor es $g(0)$.

La base de la demostración es que, si A es un abierto (que podemos suponer simplemente conexo) en el que g es analítica, entonces, por la fórmula integral de Cauchy,

$$g(0) - g_T(0) = \frac{1}{2\pi i}\int_\Gamma \frac{(g(z) - g_T(z))h(z)}{z-0}\,dz$$

para cualquier camino simple cerrado Γ contenido en A que tenga a 0 en su interior y cualquier función h holomorfa en A que cumpla $h(0) = 1$. La parte técnica delicada es tomar un camino Γ y una función h adecuadas que permitan deducir de ahí que $\lim_{T\to\infty} g_T(0) = g(0)$.

Observemos las dos ilustraciones de la figura 10.2, comenzando por la del lado izquierdo. Por hipótesis, la función $g(z)$ es analítica en $\{z \in \mathbb{C} : \mathrm{Re}(z) \geq 0\}$; es decir, ese conjunto está incluido en un abierto A en el que g es analítica (en particular, alrededor de cada z cuya parte real es nula existe un pequeño entorno en el que g es analítica). Para cada $R > 0$ (arbitrariamente grande), sea el semicírculo

$$D_R = \{z \in \mathbb{C} : |z| \leq R\} \cap \{z \in \mathbb{C} : \mathrm{Re}(z) \geq 0\},$$

sombreado en la ilustración de la izquierda. Como D_R es un subconjunto compacto de A, existirá un $\delta > 0$ (dependiente de R) tal que

$$K_R = \{z \in \mathbb{C} : |z| \leq R\} \cap \{z \in \mathbb{C} : \mathrm{Re}(z) \geq -\delta\} \subset A;$$

el recinto K_R aparece sombreado en la ilustración derecha de la figura 10.2.

Asimismo, sea C_R la frontera de K_R recorrida en «sentido positivo» (el contrario de las agujas del reloj); véase de nuevo la parte derecha de la figura 10.2. La curva C_R la descomponemos en dos partes,

$$C_R^+ = C_R \cap \{z \in \mathbb{C} : \mathrm{Re}(z) > 0\},$$
$$C_R' = C_R \cap \{z \in \mathbb{C} : \mathrm{Re}(z) \leq 0\},$$

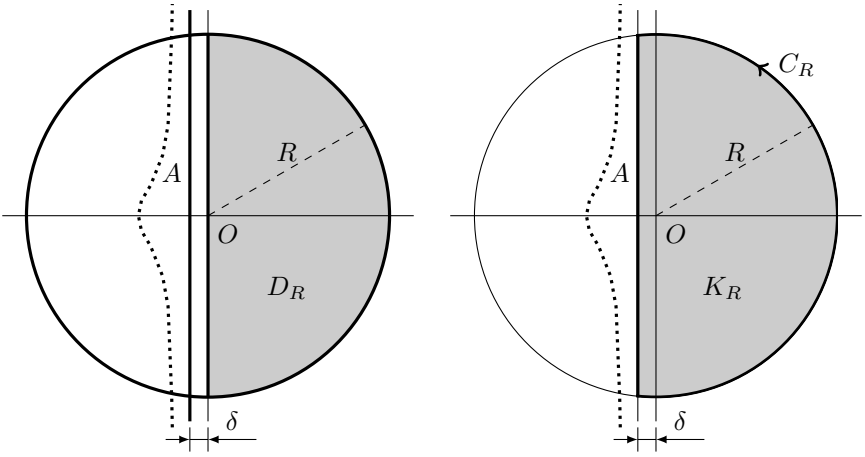


Figura 10.2: Semicírculo $D_R = \{z \in \mathbb{C} : |z| \leq R\} \cap \{z \in \mathbb{C} : \mathrm{Re}(z) \geq 0\}$ (sombreado en la ilustración de la izquierda) con una banda de anchura δ a su izquierda. Juntos forman el recinto K_R (sombreado en la ilustración de la derecha), cuya frontera es la curva C_R.

que son, respectivamente, la semicircunferencia de la derecha, y el resto (que a su vez comprende el segmento en el que $\mathrm{Re}(z) = -\delta$ y los dos trocitos de circunferencia arriba y abajo). Por supuesto, todas estas curvas también las consideramos recorridas en sentido positivo.

Para cada R y T, la función

$$h_{R,T}(z) = \big(g(z) - g_T(z)\big)e^{zT}\left(1 + \frac{z^2}{R^2}\right)$$

es holomorfa dentro de K_R y en su frontera (es decir, en un abierto que los contiene). En consecuencia, por la fórmula integral de Cauchy,

$$g(0) - g_T(0) = h_{R,T}(0) = \frac{1}{2\pi i}\int_{C_R} \frac{h_{R,T}(z)}{z - 0}\, dz$$

para cada $R > 0$ y cada $T > 0$.

En nuestro objetivo de ver que $\lim_{T\to\infty} g_T(0) = g(0)$, descompongamos

$$g(0) - g_T(0) = \frac{1}{2\pi i}\int_{C_R^+} \frac{h_{R,T}(z)}{z}\, dz + \frac{1}{2\pi i}\int_{C_R'} \frac{h_{R,T}(z)}{z}\, dz$$

con lo que, por la desigualdad triangular,

$$|g(0) - g_T(0)| \le \frac{1}{2\pi}\left|\int_{C_R^+} \frac{h_{R,T}(z)}{z}\,dz\right| + \frac{1}{2\pi}\left|\int_{C'_R} \frac{h_{R,T}(z)}{z}\,dz\right|. \tag{10.18}$$

Vamos a acotar esos dos sumandos. Recordemos para ello que f es una función acotada, así que podemos tomar $B = \sup_{t\ge 0}|f(t)| < \infty$.

En C_R^+, acotemos el módulo del integrando $h_{R,T}(z)/z$ analizando los factores separados por «$\times$» en la descomposición

$$\frac{h_{R,T}(z)}{z} = \big(g(z) - g_T(z)\big) \times e^{zT}\left(1 + \frac{z^2}{R^2}\right)\frac{1}{z}.$$

En el primero usemos que

$$\begin{aligned} |g(z) - g_T(z)| &= \left|\int_T^\infty f(t)e^{-zt}\,dt\right| \le B\int_T^\infty |e^{-zt}|\,dt \\ &= B\int_T^\infty e^{-t\,\mathrm{Re}(z)}\,dt = \frac{B}{\mathrm{Re}(z)}e^{-T\,\mathrm{Re}(z)}; \end{aligned}$$

y, en el segundo (obsérvese que $z\overline{z} = |z|^2 = R^2$ en C_R^+),

$$\begin{aligned} \left|e^{zT}\left(1 + \frac{z^2}{R^2}\right)\frac{1}{z}\right| &= e^{T\,\mathrm{Re}(z)}\left|1 + \frac{z}{\overline{z}}\right|\frac{1}{R} \\ &= e^{T\,\mathrm{Re}(z)}\left|\frac{\overline{z} + z}{\overline{z}}\right|\frac{1}{R} = e^{T\,\mathrm{Re}(z)}\frac{2\,\mathrm{Re}(z)}{R^2}. \end{aligned}$$

Así, multiplicando ambas cotas,

$$\left|\frac{h_{R,T}(z)}{z}\right| \le \frac{2B}{R^2}, \qquad z \in C_R^+.$$

Ahora, parametricemos C_R^+ como $z(t) = Re^{it}$, con $-\frac{\pi}{2} < t < \frac{\pi}{2}$, con lo cual $|z'(t)| = R$. Entonces,

$$\begin{aligned} \left|\int_{C_R^+} \frac{h_{R,T}(z)}{z}\,dz\right| &= \left|\int_{-\pi/2}^{\pi/2} \frac{h_{R,T}(z(t))}{z(t)}z'(t)\,dt\right| \\ &\le \int_{-\pi/2}^{\pi/2}\left|\frac{h_{R,T}(z(t))}{z(t)}z'(t)\right|dt \le \frac{2B}{R^2}\int_{-\pi/2}^{\pi/2} R\,dt = \frac{2\pi B}{R} \end{aligned} \tag{10.19}$$

y ya hemos acotado el primer sumando de (10.18).

Para evaluar la integral de la derecha en (10.18), descompongámosla como

$$\begin{aligned}\int_{C'_R} \frac{h_{R,T}(z)}{z}\,dz = &\int_{C'_R} g(z)e^{zT}\left(1+\frac{z^2}{R^2}\right)\frac{dz}{z}\\ &-\int_{C'_R} g_T(z)e^{zT}\left(1+\frac{z^2}{R^2}\right)\frac{dz}{z}\end{aligned} \tag{10.20}$$

y vamos a analizar ambos sumandos por separado.

Como $g_T(z)$ es holomorfa en todo $\mathbb{C}$, podemos desplazar el camino de integración de la integral que contiene a g_T en (10.20) (sin pasar por $z=0$) y la integral seguirá teniendo el mismo valor. Así, si cambiamos C'_R por la semicircunferencia izquierda $C_R^- = \{z\in\mathbb{C} : |z| = R,\ \mathrm{Re}(z) < 0\}$, se cumplirá

$$\int_{C'_R} g_T(z)e^{zT}\left(1+\frac{z^2}{R^2}\right)\frac{dz}{z} = \int_{C_R^-} g_T(z)e^{zT}\left(1+\frac{z^2}{R^2}\right)\frac{dz}{z}.$$

Esta integral sobre C_R^- se acota de modo similar a como hemos hecho con la de $h_{R,T}(z)/z$ sobre C_R^+. En efecto, para $z\in C_R^-$ se tiene (obsérvese que, esta vez, $\mathrm{Re}(z)$ es negativo)

$$\begin{aligned}|g_T(z)| &= \left|\int_0^T f(t)e^{-zt}\,dt\right| \le B\int_0^T |e^{-zt}|\,dt = B\int_0^T e^{-t\,\mathrm{Re}(z)}\,dt\\ &= \frac{B}{-\mathrm{Re}(z)}e^{-T\,\mathrm{Re}(z)} - \frac{B}{-\mathrm{Re}(z)} \le \frac{B}{-\mathrm{Re}(z)}e^{-T\,\mathrm{Re}(z)}\end{aligned}$$

y

$$\left|e^{zT}\left(1+\frac{z^2}{R^2}\right)\frac{1}{z}\right| \le e^{T\,\mathrm{Re}(z)}\frac{-2\,\mathrm{Re}(z)}{R^2},$$

con lo cual (tras proceder como con $h_{R,T}(z)/z$) se llega a

$$\left|\int_{C'_R} g_T(z)e^{zT}\left(1+\frac{z^2}{R^2}\right)\frac{dz}{z}\right| = \left|\int_{C_R^-} g_T(z)e^{zT}\left(1+\frac{z^2}{R^2}\right)\frac{dz}{z}\right| \le \frac{2\pi B}{R}. \tag{10.21}$$

Vamos ahora con la integral que contiene a g en (10.20). Para ello, descompongamos $C'_R = C_R\cap\{z\in\mathbb{C} : \mathrm{Re}(z)\le 0\}$ en tres trozos: el segmento sobre la recta $\mathrm{Re}(z) = -\delta$, que denotamos L, y los dos pequeños arcos de circunferencia arriba y abajo, que denotamos S_1 y S_2. Estos tres trozos no presentarán problemas dado que, por una parte, la longitud de S_1 y S_2 es muy pequeña; y, en cuanto a L,

allí se tiene $|e^{zT}| = e^{-\delta T}$, que es muy pequeño cuando T es grande. Pero hay que comprobar los detalles.

En $S_1 \cup S_2$ se cumple

$$\left| e^{zT} \left(1 + \frac{z^2}{R^2}\right) \frac{1}{z} \right| \leq e^{T \operatorname{Re}(z)} \frac{-2 \operatorname{Re}(z)}{R^2} \leq \frac{2\delta}{R^2}.$$

Además, $g(z)$ es holomorfa (luego continua) en la banda de anchura δ y $|z| \leq R$, así que $|g(z)|$ estará acotada en el compacto $S_1 \cup S_2$ por una constante que sólo puede depender de R. Y tengamos también en cuenta que la longitud de $S_1 \cup S_2$ es muy pequeña (2δ, aproximadamente). Entonces, dado R, podemos escoger $\delta > 0$ suficientemente pequeño para que

$$\left| \int_{S_1 \cup S_2} g(z) e^{zT} \left(1 + \frac{z^2}{R^2}\right) \frac{dz}{z} \right|$$

sea tan pequeño como deseemos (nótese que δ no depende de T). Hagámoslo de modo que

$$\left| \int_{S_1 \cup S_2} g(z) e^{zT} \left(1 + \frac{z^2}{R^2}\right) \frac{dz}{z} \right| \leq \frac{2\pi B}{R}. \tag{10.22}$$

Nos queda por estimar la integral de $g(z)e^{zT}(1 + z^2/R^2)/z$ sobre L, y la idea aquí es usar que $g(z)(1 + z^2/R^2)$ y $1/z$ son independientes de T y que, cuando $T \to \infty$, e^{zT} tiende a 0 rápidamente. En concreto, $|g(z)(1 + z^2/R^2)|$ estará acotada (de hecho en todo K_R) por una constante que sólo puede depender de R, $1/z$ verifica $|1/z| \leq 1/\delta$ en L, e^{zT} cumple $|e^{zT}| = e^{-\delta T}$, y la longitud de L es $\leq 2R$. De este modo, gracias al factor $e^{-\delta T}$ podemos escoger T suficientemente grande para que

$$\left| \int_{L} g(z) e^{zT} \left(1 + \frac{z^2}{R^2}\right) \frac{dz}{z} \right| \leq \frac{2\pi B}{R}. \tag{10.23}$$

Finalmente, llevando sobre (10.18) las desigualdades (10.19), (10.21), (10.22) y (10.23), obtenemos que, para T suficientemente grande, se cumplirá

$$|g(0) - g_T(0)| \leq \frac{1}{2\pi} \left(\frac{2\pi B}{R} + \frac{2\pi B}{R} + \frac{2\pi B}{R} + \frac{2\pi B}{R} \right) = \frac{4B}{R}.$$

En consecuencia,

$$\limsup_{T \to \infty} |g(0) - g_T(0)| \leq \frac{4B}{R}$$

y, como R es arbitrario, esto implica que $\lim_{T \to \infty} g_T(0) = g(0)$, tal como queríamos demostrar. □

10.4. El final de la demostración del teorema de los números primos

Vamos ahora a utilizar de nuevo la función $\Phi(s)$ que introdujimos en el apartado 10.2.2. Lo primero que vamos a hacer con ella es darle una formulación alternativa:

$$\Phi(s) = \sum_p \frac{\log(p)}{p^s} = s\int_0^\infty \vartheta(e^t)e^{-st}\,dt, \qquad \mathrm{Re}(s) > 1. \tag{10.24}$$

Antes de probar esta expresión, obsérvese la similitud de la integral ahí mostrada con la que aparece en el teorema 10.9 (es decir, estamos representando $\Phi(s)$ mediante una transformada de Laplace). Así pues, es gracias a (10.24) que el teorema 10.9 tiene relevancia en nuestro camino para probar el teorema de los números primos.

Básicamente, la fórmula (10.24) surge de una mera aplicación de la identidad de Abel que veíamos en el teorema 8.6. En concreto, la identidad de Abel afirma que, si $a\colon \mathbb{N} \to \mathbb{R}$ es una función aritmética y $f\colon [y,x] \to \mathbb{C}$ (con $0 < y < x < \infty$) una función de variable real continua y con derivada continua en todo el intervalo, se cumple

$$\sum_{y<n\le x} a(n)f(n) = A(x)f(x) - A(y)f(y) - \int_y^x A(t)f'(t)\,dt$$

con $A(t) = \sum_{n\le t} a(n)$.

En nuestro caso, tomemos $a(n) = \log(n)$ si n es primo y $a(n) = 0$ si no lo es, con lo cual $A(t) = \sum_{p\le t}\log(p) = \vartheta(t)$, y $f(t) = t^{-s}$. Entonces, para $y = 1$ (y dado que $a(1) = A(1) = 0$), obtenemos

$$\sum_{p\le x} \log(p)p^{-s} = \vartheta(x)x^{-s} - \int_1^x \vartheta(t)(-s)t^{-s-1}\,dt.$$

En esta fórmula, hagamos tender $x \to \infty$. Por ser $\vartheta(x) = O(x)$ (proposición 10.8) y $\mathrm{Re}(s) > 1$, se cumplirá $\vartheta(x)x^{-s} \to 0$; de este modo llegamos a

$$\Phi(s) = s\int_1^\infty \vartheta(t)t^{-s-1}\,dt, \qquad \mathrm{Re}(s) > 1.$$

Finalmente, el cambio de variable $t \mapsto e^t$ transforma esta expresión en (10.24).

Esa fórmula y el teorema tauberiano de la sección anterior son los ingredientes que nos permiten obtener —ya fácilmente— el siguiente resultado:

Proposición 10.10.
La integral

$$\int_1^\infty \frac{\vartheta(x)-x}{x^2}\,dx \tag{10.25}$$

es convergente.

Demostración. En el teorema 10.9, tomemos la función $f(t)=\vartheta(e^t)e^{-t}-1$. Como $\vartheta(t)$ es monótona creciente y $\vartheta(t)=O(t)$, es claro que $f(t)$ es acotada y localmente integrable. La función $g(z)$ que allí aparece será

$$g(z)=\int_0^\infty (\vartheta(e^t)e^{-t}-1)e^{-zt}\,dt, \qquad \operatorname{Re}(z)>0.$$

Para extender g analíticamente a $\{z\in\mathbb{C}:\operatorname{Re}(z)\geq 0\}$, manipulemos esta expresión.

Si descomponemos

$$g(z)=\int_0^\infty \vartheta(e^t)e^{-t}e^{-zt}\,dt-\int_0^\infty e^{-zt}\,dt=\int_0^\infty \vartheta(e^t)e^{-(z+1)t}\,dt-\frac{1}{z},$$

la fórmula (10.24) (con $s=z+1$) nos permite reconocer que

$$g(z)=\frac{\Phi(z+1)}{z+1}-\frac{1}{z}.$$

En la proposición 10.7 veíamos que $\Phi(s)-\frac{1}{s-1}$ es analítica en $\{s\in\mathbb{C}:\operatorname{Re}(s)\geq 1\}$ o, lo que es lo mismo, que $\Phi(z+1)-\frac{1}{z}$ es analítica en $\{z\in\mathbb{C}:\operatorname{Re}(z)\geq 0\}$. Entonces, como el denominador $z+1$ no se anula, también es analítica en $\{z\in\mathbb{C}:\operatorname{Re}(z)\geq 0\}$ la función

$$\frac{\Phi(z+1)-1/z-1}{z+1}=\frac{z\Phi(z+1)-1-z}{z(z+1)}=\frac{\Phi(z+1)}{z+1}-\frac{1}{z}=g(z).$$

Así pues, esta expresión proporciona una extensión analítica de g a $\{z\in\mathbb{C}:\operatorname{Re}(z)\geq 0\}$. Ahora, aplicando el teorema 10.9 y haciendo el cambio de variable $x=e^t$, obtenemos que

$$\int_0^\infty f(t)\,dt=\int_0^\infty \frac{\vartheta(e^t)-e^t}{e^t}\,dt=\int_1^\infty \frac{\vartheta(x)-x}{x^2}\,dx$$

es una integral convergente. □

Sólo de de la proposición 10.10 no es posible deducir que $\vartheta(x) \sim x$,[10] que es nuestro siguiente objetivo, pero casi. De hecho, lo único que vamos a tener que usar para probarlo, además de la convergencia de (10.25), es que ϑ es una función creciente (con el significado habitual de «no decreciente»). Es lo que hacemos en la demostración del siguiente teorema:

Teorema 10.11.
Cuando $x \to \infty$, la función $\vartheta(x)$ cumple

$$\lim_{x\to\infty} \frac{\vartheta(x)}{x} = 1. \tag{10.26}$$

DEMOSTRACIÓN. Por ser la integral (10.25) convergente, para cada $\varepsilon > 0$ existe un x_0 (suficientemente grande) tal que

$$\left| \int_N^M \frac{\vartheta(t) - t}{t^2}\, dt \right| < \varepsilon$$

cuando $M \geq N \geq x_0$. De aquí, y utilizando además que ϑ es creciente, probaremos que

$$\limsup_{x\to\infty} \frac{\vartheta(x)}{x} \leq 1 \tag{10.27}$$

y

$$\liminf_{x\to\infty} \frac{\vartheta(x)}{x} \geq 1. \tag{10.28}$$

[10] En general, la implicación

$$\int_1^\infty \frac{f(x) - x}{x^2}\, dx < \infty \;\Rightarrow\; \lim_{x\to\infty} \frac{f(x)}{x} = 1$$

no es cierta, tal como muestra el siguiente contraejemplo:

$$f(x) = x + \sum_{n=1}^{\infty} x^2 \chi_{[n,n+2^{-n}]}(x),$$

donde χ_I denota la función característica del intervalo I (es decir, $\chi_I(x) = 1$ si $x \in I$ y $\chi_I(x) = 0$ si $x \notin I$). En efecto; pese a que $\lim_{x\to\infty} f(x)/x$ no existe, la función f sí satisface

$$\int_1^\infty \frac{f(x) - x}{x^2}\, dx = \sum_{n=1}^{\infty} \int_n^{n+2^{-n}} dx = \sum_{n=1}^{\infty} 2^{-n} = 1 < \infty.$$

Dado que el límite inferior es siempre menor o igual que el superior, (10.27) y (10.28) prueban el teorema.

Comencemos con (10.27), demostrándolo por reducción al absurdo. Supongamos que no es cierto, es decir, que

$$\limsup_{x\to\infty} \frac{\vartheta(x)}{x} = B > 1; \tag{10.29}$$

entonces, eligiendo cualquier β que cumpla $B > \beta > 1$, tomemos

$$\varepsilon = \int_1^\beta \frac{\beta - t}{t^2}\, dt > 0.$$

Para este ε existirá x_0 (suficientemente grande) tal que, cuando $x \geq x_0$,

$$\left| \int_x^{\beta x} \frac{\vartheta(t) - t}{t^2}\, dt \right| < \varepsilon.$$

Y, por (10.29), existe $x \geq x_0$ tal que $\vartheta(x) \geq \beta x$. Entonces, por ser ϑ creciente, para este x tendremos

$$\varepsilon > \int_x^{\beta x} \frac{\vartheta(t) - t}{t^2}\, dt \geq \int_x^{\beta x} \frac{\vartheta(x) - t}{t^2}\, dt \geq \int_x^{\beta x} \frac{\beta x - t}{t^2}\, dt = \int_1^\beta \frac{\beta - t}{t^2}\, dt = \varepsilon$$

(la igualdad entre las dos últimas integrales se sigue del cambio de variable $t \mapsto tx$), que es una contradicción.

Probar (10.28) es similar, suponiendo que

$$\liminf_{x\to\infty} \frac{\vartheta(x)}{x} = B < \beta < 1$$

y tomando

$$\varepsilon = -\int_\beta^1 \frac{\beta - t}{t^2}\, dt > 0.$$

Como antes, existirá x tal que

$$\left| \int_{\beta x}^x \frac{\vartheta(t) - t}{t^2}\, dt \right| < \varepsilon$$

y $\vartheta(x) \leq \beta x$. Finalmente, para este x,

$$\varepsilon > -\int_{\beta x}^x \frac{\vartheta(t) - t}{t^2}\, dt \geq -\int_{\beta x}^x \frac{\vartheta(x) - t}{t^2}\, dt \geq -\int_{\beta x}^x \frac{\beta x - t}{t^2}\, dt = -\int_\beta^1 \frac{\beta - t}{t^2}\, dt = \varepsilon,$$

que de nuevo es absurdo. □

Nota 10.12. Es interesante fijarse en que, si en vez de $\vartheta(t)$ usamos una $f(t)$ arbitraria que cumpla las «mismas» propiedades, repitiendo la demostración del teorema estaríamos probando un «pequeño teorema tauberiano» (utilizamos «tauberiano» con el sentido de que, sin la hipótesis sobre el crecimiento de f, el resultado no es cierto en general), esta vez sobre límites; a este respecto, véase el ejercicio 10.4. Si, además, esto lo combinamos con el teorema tauberiano de la sección 10.3, podemos deducir un nuevo teorema tauberiano, tal como se propone al lector en el ejercicio 10.5. ◁

Y ya podemos dar el paso definitivo:

Teorema 10.13. Hadamard y De la Vallée-Poussin, 1896.
Para $x \geq 1$, sea $\pi(x)$ el número de primos menores o iguales que x. Entonces,

$$\lim_{x\to\infty} \frac{\pi(x)\log(x)}{x} = 1.$$

Demostración. Basta ver que

$$\liminf_{x\to\infty} \frac{\pi(x)\log(x)}{x} \geq 1 \tag{10.30}$$

y

$$\limsup_{x\to\infty} \frac{\pi(x)\log(x)}{x} \leq 1. \tag{10.31}$$

Por una parte, es claro que

$$\vartheta(x) = \sum_{p\leq x} \log(p) \leq \sum_{p\leq x} \log(x) = \pi(x)\log(x),$$

luego

$$\frac{\pi(x)\log(x)}{x} \geq \frac{\vartheta(x)}{x}, \qquad x \geq 1.$$

Con esto, de (10.26) se sigue (10.30).

Por la otra, usemos que, para cualquier $\varepsilon \in (0,1)$,

$$\begin{aligned}\vartheta(x) &\geq \sum_{x^{1-\varepsilon}<p\leq x} \log(p) \geq \sum_{x^{1-\varepsilon}<p\leq x} (1-\varepsilon)\log(x)\\ &= (1-\varepsilon)\big(\pi(x)-\pi(x^{1-\varepsilon})\big)\log(x) \geq (1-\varepsilon)\big(\pi(x)-x^{1-\varepsilon}\big)\log(x),\end{aligned}$$

luego

$$\frac{\vartheta(x)}{x} \geq (1-\varepsilon)\left(\frac{\pi(x)\log(x)}{x} - \frac{\log(x)}{x^{\varepsilon}}\right), \qquad x \geq 1.$$

Tomando límites superiores, usando (10.26), y teniendo en cuenta que $\log(x)/x^{\varepsilon} \to 0$ cuando $x \to \infty$, sacamos que para cada $\varepsilon \in (0,1)$ deberá cumplirse

$$1 \geq (1-\varepsilon)\limsup_{x\to\infty} \frac{\pi(x)\log(x)}{x} - 0.$$

Y eso implica (10.31). □

10.5. Algunas consecuencias del teorema de los números primos

En las siguientes secciones continuaremos profundizando en el teorema de los números primos, pero ahora, una vez probado el teorema, vamos a hacer un descanso. En lugar de seguir avanzando en el conocimiento de $\pi(n)$, nos detenemos momentáneamente para mostrar algunas consecuencias. En concreto, vamos a ver cuatro de estas consecuencias.

En el primer apartado de esta sección, veremos un resultado que tiene una bonita interpretación probabilística: ¿cuál es la probabilidad de que un entero positivo tomado al azar sea primo?

En el segundo apartado mostraremos que, si utilizamos p_n para denotar el n-ésimo primo, con el teorema de los números primos resulta fácil probar que

$$p_n \sim n\log(n) \quad \text{cuando } n \to \infty.$$

Es obligado destacar que esto es mucho más preciso que la existencia de dos constantes positivas C_1 y C_2 tales que

$$C_1 n\log(n) \leq p_n \leq C_2 n\log(n) \quad \text{para todo } n > 2,$$

un resultado debido a Chebyshev (lo encontró en sus intentos por demostrar el teorema de los números primos) y que ya demostrábamos en el tema 9 (véase el teorema 9.10).

El objetivo de los dos últimos apartados de esta sección va a ser mostrar que el teorema de los números primos es útil para estimar el orden de magnitud de algunas

funciones aritméticas. En concreto, allí demostraremos —como consecuencia relativamente fácil del teorema de los números primos— dos resultados que también se pueden probar sin usar dicho teorema, pero de forma considerablemente más complicada.

El primero es que la función $\tau(n)$, que cuenta el número de divisores de n, cumple

$$\limsup_{n\to\infty} \frac{\log(\tau(n))\log(\log(n))}{\log(n)} = \log(2).$$

Y, el segundo, que la indicatriz de Euler, esto es, la función $\varphi(n)$ que cuenta cuántos enteros positivos menores o iguales que n son primos con n, cumple

$$\liminf_{n\to\infty} \frac{\varphi(n)\log(\log(n))}{n} = e^{-\gamma}.$$

El primero de estos dos resultados fue demostrado por Severin Wigert en 1917, y el segundo por Edmund Landau en 1903. Ambos los hemos enunciado ya en este libro, aunque sin demostrarlos: son, respectivamente, las ecuaciones (8.10) y (8.15), que mostrábamos en la sección 8.1.

Una vez que disponemos del teorema de los números primos, estamos en condiciones de probarlos sin demasiado esfuerzo. Queremos insistir en que, tal como ya comentábamos en la sección 8.1, existen demostraciones de estos resultados que no utilizan el teorema de los números primos (por ejemplo, las que se dan en [50]). Pero —siguiendo con el símil que estamos empleando a menudo— ese camino tenía unas pendientes muy empinadas, así que no nos merecía la pena transitar por él. En su momento nos limitábamos a observar los riscos que se perfilaban en el horizonte sin trepar por la parte más empinada, las demostraciones. En cambio, ahora que ya hemos probado el teorema de los números primos y nos hemos situado en un lugar más alto, será mucho menos trabajoso llegar a las cimas que entonces divisábamos.

10.5.1. ¿Cuál es la probabilidad de que un número tomado al azar sea primo?

El teorema de los números primos responde a la pregunta «¿son muy abundantes los números primos?» mostrando —de manera asintótica— cuántos números primos hay hasta n. Ahora nos planteamos algo muy similar, «¿cuál es la probabilidad de que un número sea primo?».

El teorema de los números primos nos va a permitir contestar a eso. Previamente, hay que dar un enunciado matemático riguroso:

Corolario 10.14.
La probabilidad de que un entero positivo menor o igual que n sea primo es

$$\frac{1}{\log(n)} + o\left(\frac{1}{\log(n)}\right).$$

Demostración. La probabilidad de un suceso es el número de casos favorables dividido entre el número de casos posibles. Obviamente, al elegir un entero positivo menor o igual que n hay n casos posibles. Y, según el teorema de los números primos, los casos favorables son

$$\pi(n) = \frac{n}{\log(n)} + o\left(\frac{n}{\log(n)}\right).$$

Obviamente, esto prueba el resultado. □

Cuando n es grande, el sumando $o(1/\log(n))$ es despreciable. Con este sentido podemos decir que, para n suficientemente grande, la probabilidad de que un entero positivo menor o igual que n sea primo es

$$\frac{1}{\log(n)}.$$

Es ésta, en suma, una bonita interpretación probabilística del teorema de los números primos. Pero, al mismo tiempo, es prácticamente el comienzo de la historia del teorema de los números primos: cuando Gauss, utilizando una inmensa tabla de primos que él mismo había construido, observó que la densidad de los primos en un entorno de n era aproximadamente $1/\log(n)$, llegó a la conclusión de que el número de primos menores que x debería ser

$$\pi(x) \sim \int_2^x \frac{dt}{\log(t)} \sim \frac{x}{\log(x)}.$$

Y, más aún, esta formulación a partir de una integral será clave, de nuevo, en las secciones 10.6 y 10.7 de este tema.

10.5.2. Fórmula asintótica para el n-ésimo primo

Otra consecuencia sencilla del teorema del números primo es una fórmula asintótica para el primo n-ésimo:

Corolario 10.15.
Sea p_n el n-ésimo primo. Entonces

$$\lim_n \frac{p_n}{n\log(n)} = 1. \tag{10.32}$$

DEMOSTRACIÓN. Tomando logaritmos en la fórmula $\lim_{x\to\infty} \pi(x)\log(x)/x = 1$ del teorema de los números primos obtenemos que

$$\log(\pi(x)) + \log(\log(x)) - \log(x) \to 0$$

cuando $x \to \infty$, y por tanto

$$\frac{\log(\pi(x))}{\log(x)} + \frac{\log(\log(x))}{\log(x)} \to 1.$$

Aquí, el segundo sumando tiende a 0, con lo cual

$$\lim_{x\to\infty} \frac{\log(\pi(x))}{\log(x)} = 1.$$

Y si multiplicamos este límite por $\lim_{x\to\infty} \pi(x)\log(x)/x = 1$ queda

$$\lim_{x\to\infty} \frac{\pi(x)\log(\pi(x))}{x} = 1.$$

Finalmente, tomando valores $x = p_n$, para los cuales $\pi(p_n) = n$, esta expresión se convierte en (10.32). □

En realidad, aún podemos afinar lo que hemos obtenido en este corolario pues, de hecho, (10.32) es un resultado equivalente al teorema de los números primos. En el ejercicio 10.6 se pide al lector que lo pruebe.

10.5.3. El tamaño asintótico de la función $\tau(n)$

Estudiemos ahora el tamaño de la función $\tau(n)$ que cuenta el número de divisores de cada entero positivo n. Es claro que, si n es primo, $\tau(n) = 2$, en cambio, si $n = 2^k$, $\pi(n) = k + 1$; así pues, $\tau(n)$ presenta muchas oscilaciones y no es posible fijar el tamaño de $\tau(n)$. Nos tenemos que contentar con dar cotas (sí que es posible precisar el orden medio de $\tau(n)$, tal como hacíamos en el tema 8).

El teorema en el que analizamos el tamaño asintótico de $\tau(n)$ tiene dos partes. La primera proporciona una cota superior, y es relativamente sencilla de probar. La

segunda muestra que esa cota es buena —en el sentido de que, si se rebaja, existen infinitos n que no la cumplen—, y es ahí donde el teorema de los números primos desempeña un papel importante.

Adelantamos que, en la demostración de esa segunda parte, no sólo usaremos el teorema de los números primos en la forma $\lim_{x\to\infty} \pi(x)/(x/\log(x)) = 1$ sino que, además, utilizaremos que la función $\vartheta(x) = \sum_{p\le x} \log(p)$ que definíamos en (10.16) cumple $\lim_{x\to\infty} \vartheta(x)/x = 1$, tal como probábamos en el teorema 10.11 (de hecho, en el tema 8, en concreto en el teorema 8.28, veíamos que eso es un resultado equivalente al teorema de los números primos).

Por otra parte, sugerimos al lector que compare este teorema sobre el tamaño asintótico de $\tau(n)$ con el teorema 9.28 que veíamos en el apartado 9.5.2, y que proporciona interesante información adicional sobre $\tau(n)$.

Teorema 10.16. Wigert, 1917.
Sea $\tau(n)$ la función que cuenta el número de divisores de cada número natural n, y tomemos $\varepsilon > 0$ fijo. Entonces:

(a) Existe un entero $N(\varepsilon)$ tal que, para todo $n \ge N(\varepsilon)$, se cumple

$$\tau(n) < 2^{(1+\varepsilon)\log(n)/\log(\log(n))} = n^{(1+\varepsilon)\log(2)/\log(\log(n))}.$$

(b) Sin embargo, infinitos valores de n verifican

$$\tau(n) > 2^{(1-\varepsilon)\log(n)/\log(\log(n))} = n^{(1-\varepsilon)\log(2)/\log(\log(n))}.$$

Demostración. Comencemos con la parte (a). Escribamos $n = p_1^{a_1}\cdots p_k^{a_k}$ (su descomposición en factores primos), con lo que $\tau(n) = \prod_{j=1}^{k}(a_j+1)$. Vamos a dividir este producto en dos partes; por un lado tomaremos los $p_j < f(n)$ y por otro los $p_j \ge f(n)$, donde $f(n)$ es una función que especificaremos más adelante. Así, $\tau(n) = P_1(n)P_2(n)$ con

$$P_1(n) = \prod_{p_j<f(n)} (a_j+1), \qquad P_2(n) = \prod_{p_j\ge f(n)} (a_j+1).$$

Para estimar $P_2(n)$ usemos la desigualdad $(a+1) \le 2^a$, con lo cual $P_2(n) \le 2^{S(n)}$ con

$$S(n) = \sum_{p_j\ge f(n)} a_j.$$

Entonces,

$$n = \prod_{j=1}^{k} p_j^{a_j} \geq \prod_{p_j \geq f(n)} p_j^{a_j} \geq \prod_{p_j \geq f(n)} f(n)^{a_j} = f(n)^{S(n)},$$

de donde $\log(n) \geq S(n)\log(f(n))$, y por tanto $S(n) \leq \log(n)/\log(f(n))$ y

$$P_2(n) \leq 2^{\log(n)/\log(f(n))}. \tag{10.33}$$

Para estimar $P_1(n)$ comencemos escribiéndolo como

$$P_1(n) = \exp\left(\sum_{p_j < f(n)} \log(a_j + 1)\right).$$

Probemos ahora que $\log(a_j + 1) < 2\log(\log(n))$ si n es suficientemente grande; en efecto, al ser $n \geq p_j^{a_j} \geq 2^{a_j}$, es claro que $\log(n) \geq a_j \log(2)$, luego

$$1 + a_j \leq 1 + \frac{\log(n)}{\log(2)} < \log^2(n)$$

para n suficientemente grande (de hecho, esa desigualdad se cumple siempre que $n \geq 8$), y por tanto $\log(1 + a_j) < \log(\log^2(n)) = 2\log(\log(n))$. De aquí se sigue que, para esos n,

$$P_1(n) < \exp\left(2\log(\log(n)) \sum_{p_j < f(n)} 1\right) \leq \exp\big(2\log(\log(n))\pi(f(n))\big).$$

Vamos así mismo a aplicar que, para cierta constante C, $\pi(x) < Cx/\log(x)$ para todo $x \geq 2$; esto se deduce trivialmente del teorema de los números primos, pero realmente no hace falta usar un resultado tan fuerte, sino que ya lo probábamos de manera independiente en el tema 9 (en concreto, véase el teorema 9.9, donde obteníamos esa desigualdad con $C = 2/e + 4\log(2)$). De este modo tenemos

$$P_1(n) < \exp\left(\frac{2Cf(n)\log(\log(n))}{\log(f(n))}\right) = 2^{cf(n)\log(\log(n))/\log(f(n))} \tag{10.34}$$

con $c = 2C/\log(2)$.

Combinando las desigualdades (10.33) y (10.34) llegamos a que

$$\tau(n) = P_1(n)P_2(n) < 2^{g(n)}$$

con

$$g(n) = \frac{\log(n) + cf(n)\log(\log(n))}{\log(f(n))} = \frac{\log(n)}{\log(\log(n))} \frac{1 + \frac{cf(n)\log(\log(n))}{\log(n)}}{\frac{\log(f(n))}{\log(\log(n))}}.$$

Elijamos por fin $f(n)$ que cumpla

$$\frac{f(n)\log(\log(n))}{\log(n)} \to 0 \qquad \text{y} \qquad \frac{\log(f(n))}{\log(\log(n))} \to 1$$

cuando $n \to \infty$; por ejemplo,

$$f(n) = \frac{\log(n)}{\log^2(\log(n))}.$$

Con esta f,

$$g(n) = \frac{\log(n)}{\log(\log(n))} \frac{1 + o(1)}{1 + o(1)} = \frac{\log(n)}{\log(\log(n))}(1 + o(1)) < (1 + \varepsilon)\frac{\log(n)}{\log(\log(n))}$$

si $n \geq N(\varepsilon)$ para cierto $N(\varepsilon)$, lo cual prueba la parte (a).

Para demostrar la parte (b), hay que buscar entre los enteros que tienen gran cantidad de factores primos. Dado x, supongamos que n es el producto de todos los números primos $\leq x$, es decir,

$$n = 2 \cdot 3 \cdot 5 \cdots p_{\pi(x)}. \tag{10.35}$$

Como la función τ es multiplicativa y, sobre los primos, $\tau(p) = 2$,

$$\tau(n) = \tau(2)\tau(3)\tau(5)\cdots\tau(p_{\pi(x)}) = 2^{\pi(x)}.$$

Entonces, y si escribimos el teorema de los números primos como

$$\pi(x) = (1 + o(x))x/\log(x),$$

tendremos

$$\tau(n) = 2^{\pi(x)} = 2^{(1+o(x))x/\log(x)}.$$

Escribamos así mismo $\vartheta(x)/x \to 1$ cuando $x \to \infty$ (visto en el teorema 10.11) como $\vartheta(x) = (1 + o(x))x$, con lo cual el n de (10.35) cumplirá

$$\log(n) = \sum_{p \leq x} \log(p) = \vartheta(x) = (1 + o(x))x.$$

De aquí,

$$x = \frac{\log(n)}{1+o(1)} = (1+o(1))\log(n)$$

y

$$\begin{aligned}\log(x) &= \log(\log(n)) + \log(1+o(1)) \\ &= \left(1 + \frac{\log(1+o(1))}{\log(\log(n))}\right)\log(\log(n)) = (1+o(1))\log(\log(n)).\end{aligned}$$

Entonces, $x/\log(x) = (1+o(1))\log(n)/\log(\log(n))$ y por tanto

$$\tau(n) = 2^{(1+o(1))\log(n)/\log(\log(n))} = n^{(1+o(1))\log(2)/\log(\log(n))}.$$

Como $1+o(1) > 1-\varepsilon$ siempre que $n \geq N(\varepsilon)$ para cierto $N(\varepsilon)$, esto implica (b). □

De manera equivalente, el teorema que acabamos de ver se puede reescribir como sigue (en el ejercicio 10.8 se pide al lector que compruebe los detalles):

Corolario 10.17. Wigert, 1917.
La función $\tau(n)$ que cuenta el número de divisores de cada número natural n cumple

$$\limsup_{n\to\infty} \frac{\log(\tau(n))\log(\log(n))}{\log(n)} = \log(2). \tag{10.36}$$

Nota 10.18. Una consecuencia de (10.36) es que $\tau(n) = o(n^\delta)$ para cada $\delta > 0$, algo que ya habíamos probado (véase (8.9)) sin utilizar el teorema de los números primos. ◁

10.5.4. El tamaño asintótico de la función $\varphi(n)$

Mediante razonamientos parecidos a los del apartado anterior, vamos a obtener también desigualdades asintóticas sobre el tamaño de la indicatriz de Euler $\varphi(n)$. Recordemos que $\varphi(n)$ es, por definición,

$$\varphi(n) = \operatorname{card}\{m \in \mathbb{N} : 1 \leq m \leq n,\ \operatorname{mcd}(m,n) = 1\}$$

y que, según habíamos visto tanto en el tema 2 (teorema 2.10) como en el tema 7 (fórmula (7.14)), se puede evaluar mediante

$$\varphi(n) = n\prod_{p\mid n}\left(1-\frac{1}{p}\right), \qquad n > 1.$$

Por otra parte, en el teorema 9.20 hemos demostrado que, cuando $x \to \infty$, se verifica la expresión asintótica

$$\prod_{p\leq x}\left(1-\frac{1}{p}\right)=\frac{B}{\log(x)}+O\left(\frac{1}{\log^2(x)}\right) \tag{10.37}$$

para una cierta constante positiva B. Comentábamos además (sin probarlo) que $B=e^{-\gamma}$ con γ la constante de Euler.

Utilizando esta fórmula y el teorema de los números primos se llega al siguiente resultado (del mismo modo que en el teorema 10.16, el teorema de los números primos sólo se usa para probar la segunda parte):

Teorema 10.19. Landau, 1903.
Sea $\varphi(n)$ la indicatriz de Euler, usemos B para denotar la constante que aparece en (10.37) y tomemos $\varepsilon>0$ fijo. Entonces:

(a) Existe un entero $N(\varepsilon)$ tal que, para todo $n\geq N(\varepsilon)$, se cumple

$$\varphi(n)>(1-\varepsilon)\frac{Bn}{\log(\log(n))}.$$

(b) Sin embargo, infinitos valores de n verifican

$$\varphi(n)<(1+\varepsilon)\frac{Bn}{\log(\log(n))}.$$

Demostración. Comencemos por la parte (a). Descompongamos

$$\frac{\varphi(n)}{n}=\prod_{p\leq x}\left(1-\frac{1}{p}\right)=P_1(n)P_2(n)$$

con

$$P_1(n)=\prod_{\substack{p\mid n\\ p\leq\log(n)}}\left(1-\frac{1}{p}\right),\qquad P_2(n)=\prod_{\substack{p\mid n\\ p>\log(n)}}\left(1-\frac{1}{p}\right),$$

y analicemos por separado $P_1(n)$ y $P_2(n)$. Es claro que

$$P_2(n)>\prod_{\substack{p\mid n\\ p>\log(n)}}\left(1-\frac{1}{\log(n)}\right)=\left(1-\frac{1}{\log(n)}\right)^{f(n)} \tag{10.38}$$

donde $f(n)$ es el número de primos que dividen a n y son mayores que $\log(n)$. Como

$$n \geq \prod_{p|n} p > \prod_{\substack{p|n \\ p>\log(n)}} p \geq (\log(n))^{f(n)},$$

de aquí se deduce que $\log(n) > f(n)\log(\log(n))$, luego $f(n) < \log(n)/\log(\log(n))$. Entonces, y dado que $(1 - 1/\log(n)) < 1$, de (10.38) se sigue que

$$P_2(n) > \left(1 - \frac{1}{\log(n)}\right)^{\log(n)/\log(\log(n))} = \left(\left(1 - \frac{1}{\log(n)}\right)^{\log(n)}\right)^{1/\log(\log(n))}. \tag{10.39}$$

Ahora, teniendo en cuenta que $(1 - 1/t)^t \to e^{-1}$ cuando $t \to \infty$, la expresión de la derecha en (10.39) tiende a 1 cuando $n \to \infty$, y por tanto

$$P_2(n) \geq 1 + o(1) \quad \text{cuando } n \to \infty.$$

En cuanto a $P_1(n)$, de (10.37) se deduce que

$$P_1(n) \geq \prod_{p \leq \log(n)} \left(1 - \frac{1}{p}\right) = \frac{B}{\log(\log(n))}(1 + o(1)).$$

Juntando las cotas asintóticas de $P_2(n)$ y $P_1(n)$ llegamos a

$$\frac{\varphi(n)}{n} = P_1(n)P_2(n) \geq \frac{B}{\log(\log(n))}(1 + o(1)) > (1 - \varepsilon)\frac{B}{\log(\log(n))}$$

si $n \geq N(\varepsilon)$, lo cual demuestra la parte (a).

Para probar (b), para $x > 2$ cualquiera tomemos $n = \prod_{p \leq x} p$. Aplicando (10.37) tenemos

$$\frac{\varphi(n)}{n} = \prod_{p \leq x}\left(1 - \frac{1}{p}\right) = \frac{B}{\log(x)} + O\left(\frac{1}{\log^2(x)}\right).$$

Además, por el teorema de los números primos en su forma $\vartheta(x) \sim x$ (teorema 10.11), para estos $n = \prod_{p \leq x} p$ se cumplirá

$$\log(n) = \sum_{p \leq x} \log(p) = \vartheta(x) = (1 + o(1))x,$$

luego $\log(\log(n)) = (1+o(1))\log(x)$. Entonces,

$$\frac{\varphi(n)}{n} = \frac{B(1+o(1))}{\log(\log(n))} + O\left(\frac{1}{\log^2(\log(n))}\right) = \frac{B(1+o(1))}{\log(\log(n))} \leq (1+\varepsilon)\frac{B}{\log(\log(n))}$$

si $n \geq N(\varepsilon)$ para cierto $N(\varepsilon)$, y ya hemos probado (b). □

Como ocurría en el apartado anterior, de manera equivalente al teorema podemos escribir lo siguiente (el ejercicio 10.9 propone al lector que lo compruebe):

Corolario 10.20. Landau, 1903.
La indicatriz de Euler $\varphi(n)$ cumple

$$\liminf_{n\to\infty} \frac{\varphi(n)\log(\log(n))}{n} = B$$

donde B es la constante que aparece en (10.37).

10.6. El tamaño del error al aproximar $\pi(x)$

Tras nuestra pequeña excursión para mostrar algunas consecuencias del teorema de los números primos, volvamos a la senda principal. El resultado central de este tema es que

$$\lim_{x\to\infty} \frac{\pi(x)}{x/\log(x)} = 1,$$

y ya lo hemos probado. De manera natural, ahora nos podemos preguntar algo más: asintóticamente cuando $x \to \infty$, ¿cuál es el tamaño del error que se comete al aproximar $\pi(x)$ por $x/\log(x)$? Obviamente, el teorema de los números primos equivale a decir que

$$\left|\pi(x) - \frac{x}{\log(x)}\right| = o\left(\frac{x}{\log(x)}\right)$$

cuando $x \to \infty$; ¿hasta cuánto se puede rebajar el tamaño de esa o pequeña?

Antes de intentar responder a esa pregunta, es mejor caer en la cuenta de que, si queremos estimar el error de la aproximación, conviene no aproximar $\pi(x)$ por $x/\log(x)$ sino por la función *logaritmo integral* que se define como

$$\mathrm{li}(x) = \text{v. p.} \int_0^x \frac{dt}{\log(t)} = \lim_{\varepsilon\to 0^+}\left(\int_0^{1-\varepsilon} + \int_{1+\varepsilon}^x\right)\frac{dt}{\log(t)}, \qquad x > 1.$$

Obsérvese que estamos obligados a tomar el valor principal de la integral dado que, si no, la integral sería divergente en $t = 1$.

Nota 10.21. Aclaremos que hay otra función —muy similar— que también se conoce con el nombre de logaritmo integral, y es

$$\mathrm{Li}(x) = \int_2^x \frac{dt}{\log(t)}, \qquad x \geq 2,$$

esto es, $\mathrm{Li}(x) = \mathrm{li}(x) - \mathrm{li}(2)$ donde $\mathrm{li}(2) = 1.045\,163\ldots$ Para distinguirlas, a veces $\mathrm{Li}(x)$ se denomina logaritmo integral europeo o euleriano (también logaritmo integral *offset*), y $\mathrm{li}(x)$ logaritmo integral americano. Sobre $\mathrm{li}(x)$, el europeo tiene la ventaja de que en su dominio de integración se elimina la singularidad que hay en $t = 1$ (y que obliga a que la integral que define $\mathrm{li}(x)$ deba ser un valor principal), pero su extensión a los complejos es menos natural. Dado que $\mathrm{Li}(x)$ y $\mathrm{li}(x)$ se diferencian en una constante, para lo que ahora interesa, que es su tamaño para valores grandes de la variable x, es irrelevante usar uno u otro. En lo que sigue emplearemos $\mathrm{li}(x)$, que posiblemente es el más común de los dos. Pero, por otra parte —y por si el lector acude a otras fuentes—, vale la pena advertir que, en algunos sitios —sobre todo los menos recientes—, nuestro li se denota Li. ◁

Veamos por qué conviene aproximar $\pi(x)$ con $\mathrm{li}(x)$ en lugar de con $x/\log(x)$. En primer lugar, el teorema fundamental del cálculo muestra que la derivada de $\mathrm{li}(x)$ es $1/\log(x)$, así que, aplicando la regla de l'Hôpital,

$$\lim_{x\to\infty} \frac{\mathrm{li}(x)}{x/\log(x)} = \lim_{x\to\infty} \frac{1/\log(x)}{(\log(x)-1)/\log^2(x)} = 1,$$

luego realmente da lo mismo enunciar el teorema de los números primos como $\pi(x) \sim \mathrm{li}(x)$ o como $\pi(x) \sim x/\log(x)$. En segundo lugar, para convencernos de que $\mathrm{li}(x)$ proporciona mejores aproximaciones a $\pi(x)$ que $x/\log(x)$ podemos servirnos de un experimento numérico: obsérvese el cuadro 10.1, en el que se muestra que los errores con $|\pi(x) - \mathrm{li}(x)|$ son realmente mucho menores que con $|\pi(x) - x/\log(x)|$.

De hecho, si nos fijamos en la columna correspondiente a los valores de $\sqrt{x}$, es razonable pensar que, cuando $x \to \infty$, se cumplirá

$$|\pi(x) - \mathrm{li}(x)| = O\left(x^r\right)$$

con $r = 1/2$ o incluso menor (la razón de incluir en la tabla los valores de $xe^{-\sqrt{\log(x)}}$ se verá más adelante). ¿Será eso cierto? Y, además, ¿cómo intentar probarlo? Con lo visto hasta ahora, estamos aún muy lejos, pues, igual que cuando aproximábamos $\pi(x)$ por $x/\log(x)$, el teorema de los números primos sólo garantiza un error de tamaño $o(x/\log(x))$, es decir,

$$\pi(x) = \mathrm{li}(x) + o\left(\frac{x}{\log(x)}\right), \quad x \to \infty.$$

x	$\pi(x)$	$\pi(x) - \frac{x}{\log(x)}$	$\mathrm{li}(x) - \pi(x)$	$\sqrt{x}$	$xe^{-\sqrt{\log(x)}}$
10^2	25	3	5	10	12
10^3	168	23	10	32	72
10^4	1 229	143	17	100	481
10^5	9 592	906	38	316	3 361
10^6	78 498	6 116	130	1 000	24 309
10^7	664 579	44 158	339	3 162	180 477
10^8	5 761 455	332 774	754	10 000	1 367 847
10^9	50 847 534	2 592 592	1 701	31 623	10 543 124
10^{10}	455 052 511	20 758 029	3 104	100 000	82 418 874
10^{11}	4 118 054 813	169 923 159	11 588	316 228	652 094 200
10^{12}	37 607 912 018	1 416 705 193	38 263	1 000 000	5 213 406 648

Cuadro 10.1: Valores de $\pi(x)$ y errores de aproximación (los datos mostrados se han redondeado al entero más cercano).

En lo que sigue en esta sección, no vamos a incluir demostraciones —para darlas, necesitaríamos hacer una incursión más profunda en el terreno de la variable compleja—, sino que lo que queremos es, exclusivamente, informar al lector de qué es lo que los matemáticos conocen hasta ahora, y que se dé cuenta de que realmente no se sabe mucho más. Anticipamos que, si se pretende intentar probar que $\pi(x)$ se aproxima por $\mathrm{li}(x)$ con un error de tamaño $O(x^r)$ con algún $r < 1$, habría que ampliar, previamente, el conocimiento sobre la posición de los ceros complejos de $\zeta(z)$; cuanto mayor es la región libre de ceros, menor es la cota del término del error que se consigue demostrar (es decir, menor el r en $O(x^r)$). En particular, veremos hasta dónde se sabría llegar si la hipótesis de Riemann fuese cierta. Pero esto es una frontera que sólo se vislumbra en la lejanía, pues nada parece indicar que la hipótesis de Riemann esté cerca de ser probada.

10.6.1. Qué se sabe

El camino que nosotros hemos seguido en este tema es muy directo para llegar al teorema de los números primos, pero —tal como estábamos diciendo— no permite mostrar cómo aparecen acotaciones del error al aproximar $\pi(x)$ mediante $\mathrm{li}(x)$. Así pues, ya no podemos continuar por ahí. No importa, pues ya no tenemos intención de probar nada más, sino, simplemente, de informarnos —a grandes rasgos— del estado actual de los conocimientos sobre el tema, hasta dónde se podría seguir, y qué problemas permanecen abiertos. Las correspondientes demostraciones de lo que

incluimos en este apartado y en el siguiente se pueden consultar, por ejemplo, en [3, capítulo 11], [12, capítulo 8] o [69, capítulo 5].

Las mejoras en las estimaciones del término del error siempre están en función del tamaño de la región libre de ceros de zeta. Y, claro, ahí está, en el horizonte, la hipótesis de Riemann que afirma que todos los ceros no triviales de $\zeta(s)$ están en la recta $\mathrm{Re}(s) = 1/2$ del plano complejo (la recta crítica). Si la hipótesis de Riemann estuviera confirmada, lo que se sabe demostrar a ciencia cierta del tamaño de $\pi(x)$ aumentaría considerablemente.

En la práctica, los intentos de aproximar $\pi(x)$ lo hacen pasando por algunos resultados equivalentes al teorema de los números primos. En este tema, nosotros lo hemos hecho por medio de

$$\lim_{x\to\infty} \frac{\vartheta(x)}{x} = 1,$$

pero a menudo se utilizan otras caracterizaciones. Así, por ejemplo, en el apartado 8.7.1 vimos que un resultado equivalente al teorema de los números primos era

$$\lim_{x\to\infty} \frac{\psi(x)}{x} = 1,$$

donde $\psi(x)$ denota las sumas parciales de la función de von Mangoldt $\Lambda(n)$, esto es,

$$\psi(x) = \sum_{n\leq x} \Lambda(n), \qquad x \geq 1.$$

En esta sección usaremos la función $\psi(x)$. (Tanto conceptual como históricamente, las funciones $\psi(x)$ y $\vartheta(x)$ están muy relacionadas: son las denominadas funciones de Chebyshev, pues fue él quien las introdujo en sus intentos para demostrar el teorema de los números primos.)

Además, y tal como comentábamos, las estimaciones del término del error en $|\pi(x) - \mathrm{li}(x)|$ están en función del tamaño de la región libre de ceros de zeta. Así que lo primero que tenemos que hacer es ver qué es lo que se conoce a este respecto.

La función de von Mangoldt se relaciona con la zeta de Riemann por medio de la serie de Dirichlet

$$-\frac{\zeta'(s)}{\zeta(s)} = \sum_{n=1}^{\infty} \Lambda(n) n^{-s},$$

válida cuando $\mathrm{Re}(s) > 1$ (no es ésta ninguna de las series de Dirichlet que vimos en el tema 8, en los ejercicios 8.13 a 8.19; pero a estas alturas su demostración no es difícil, y en el ejercicio 10.14 se sugiere al lector que lo intente). Partiendo de esa expresión y de ciertas propiedades que ligan la función $\psi(x)$ con los ceros de la

función $\zeta(s)$ (que no probaremos, pero que no son muy diferentes a lo que hemos hecho con $\vartheta(x)$), se logra demostrar lo siguiente:

Teorema 10.22.
Existe un $K > 0$ tal que todos los ceros s de la función ζ cumplen

$$\mathrm{Re}(s) < 1 - \eta(|\,\mathrm{Im}(s)|)$$

con $\eta(t) = K/\log(t+2)$.

El teorema anterior proporciona una región libre de ceros de $\zeta(s)$ dentro de la banda crítica. Es un resultado típico de acuerdo a lo que actualmente se sabe hacer, y no se conoce nada mucho mejor. Pero realmente no es ningún avance considerable sobre lo que hemos visto en las secciones anteriores de este tema, cuando probábamos el teorema de los números primos. Obsérvese que, cuando $\mathrm{Im}(s) \to \pm\infty$, la región $\mathrm{Re}(s) > 1 - K/\log(|\,\mathrm{Im}(s)| + 2)$ del plano complejo se aproxima asintóticamente (por la izquierda) a la recta vertical $\mathrm{Re}(s) = 1$. (De hecho, el gráfico de la curva $\mathrm{Re}(s) = 1 - K/\log(|\,\mathrm{Im}(s)| + 2)$ es similar al de la línea punteada en la figura 10.2.)

Una vez establecido el teorema 10.22, de él se deduce el clásico término del error de De la Vallée-Poussin:

Teorema 10.23.
Asintóticamente cuando $x \to \infty$, la función $\psi(x)$ cumple

$$\psi(x) = x + O\big(x \exp\big(-c\sqrt{\log(x)}\,\big)\big)$$

para cierta constante $c > 0$.

Para hacernos una idea del tamaño de $x\exp(-c\sqrt{\log(x)})$ para x grande, nótese que $x\exp(-c\sqrt{\log(x)})$ tiende a ∞ más rápido que $x^{1-\varepsilon}$ para cualquier $\varepsilon > 0$, pero crece más lento que $x/\log^r(x)$ para cualquier número real r. Una tabla de valores de $x\exp(-c\sqrt{\log(x)})$ con $c = 1$ se puede observar en la última columna del cuadro 10.1.

Las estimaciones para $\psi(x)$ se pasan a $\pi(x)$ utilizando la relación

$$\pi(x) = \mathrm{li}(x) + \frac{\psi(x) - x}{\log(x)} + \int_2^x \frac{\psi(t) - t}{t\log^2(t)}\,dt + O(\sqrt{x}/\log(x))$$

(que no probaremos). Con esto, del teorema anterior se sigue este otro:

Teorema 10.24. De la Vallée-Poussin, 1899.
Asintóticamente cuando $x \to \infty$, la función $\pi(x)$ cumple

$$\pi(x) = \mathrm{li}(x) + O\big(x \exp\big(-c\sqrt{\log(x)}\,\big)\big)$$

para cierta constante $c > 0$.

Tal como se ve en el cuadro 10.1 (con $c = 1$), es un error mucho mayor de lo que los experimentos numéricos parecen indicar.

El resultado de De la Vallée-Poussin de 1899 que acabamos de exponer prueba que

$$\pi(x) = \mathrm{li}(x) + O\big(x \exp(-c \log^a(x))\big), \qquad x \to \infty, \tag{10.40}$$

con $a = 1/2$ y c alguna constante positiva. Desde entonces, las mejoras que se han conseguido sólo han logrado encontrar que (10.40) sigue siendo cierto para algunos valores mayores de a (pero sin llegar a $a = 1$). El primero que mejoró el término del error de De la Vallée-Poussin fue Littlewood, y luego su método fue perfeccionado por Ivan Vinogradov y su escuela, que lograron bajar la cota hasta $a = 3/5 - \varepsilon$ para cualquier $\varepsilon > 0$. (Los métodos elementales de Selberg para probar el teorema de los números primos logran $a = 1/6 - \varepsilon$.)

Estos son los resultados que se conocen. Pero no son nada espectaculares sino que, por el contrario, no llegan a ser $\pi(x) = \mathrm{li}(x) + O(x^{1-\varepsilon})$ para ningún $\varepsilon > 0$, y por tanto sólo suponen una leve mejora de lo que hemos probado en este tema.

10.6.2. Qué se sabría si la hipótesis de Riemann estuviera probada

Si tuviéramos una región libre de ceros mayor —insistimos en que no está probado que tal región exista—, el tamaño asintótico del error se podría mejorar:

Teorema 10.25.
Supongamos que existe un número $r \in [1/2, 1)$ tal que $\zeta(s) \neq 0$ siempre que $\mathrm{Re}(s) > r$. Entonces

$$\psi(x) = x + O(x^{r+\varepsilon}), \qquad x \to \infty,$$

para cada $\varepsilon > 0$.

Al pasar de $\psi(x)$ a $\pi(x)$, lo que se logra es lo siguiente (obsérvese que esto es (10.40) con $a = 1$ y con $c = 1 - r - \varepsilon$):

Teorema 10.26.
Supongamos que existe un número $r \in [1/2, 1)$ tal que $\zeta(s) \neq 0$ siempre que $\mathrm{Re}(s) > r$. Entonces

$$\pi(x) = \mathrm{li}(x) + O(x^{r+\varepsilon}), \qquad x \to \infty,$$

para cada $\varepsilon > 0$.

De hecho, la estimación asintótica anterior es muy precisa pues —tal como probó Ingham en 1932— se sabe que el supremo de las partes reales de los ceros de $\zeta(s)$ coincide con el ínfimo de los números r tales que el error de $|\pi(x) - \mathrm{li}(x)|$ es $O(x^r)$.

En 1859, Riemann conjeturó que todos los ceros no triviales de la función $\zeta(s)$ debían estar en la recta crítica $\mathrm{Re}(s) = 1/2$, aunque aparentemente él no hizo ningún esfuerzo por probarlo. Muchos otros insignes matemáticos sí que lo han intentado, pero siempre sin éxito. Esto, que se ha venido a denominar hipótesis de Riemann, es uno de los problemas abiertos más importantes de la matemática. En 1900, en el célebre Congreso Internacional de Matemáticos que se celebró en París, David Hilbert lo colocó el octavo en su lista de 23 problemas que deberían guiar la investigación de los matemáticos durante el siglo XX,[11] y a fecha actual sigue tan desafiante como entonces.

Si la hipótesis de Riemann fuese cierta, en el teorema 10.25 valdría $r = 1/2$, y entonces tendríamos que

$$\psi(x) = x + O(x^{1/2+\varepsilon}), \qquad x \to \infty,$$

para cualquier $\varepsilon > 0$. Con mayor esfuerzo, se consigue probar algo más. Así, lo que se tiene es lo siguiente:

Teorema 10.27.
Si la hipótesis de Riemann fuese cierta, entonces

$$\psi(x) = x + O(x^{1/2} \log^2(x)), \qquad x \to \infty.$$

Al trasladarlo a $\pi(x)$, lo que se consigue es esto otro:

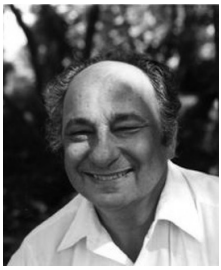
[11] En realidad, el octavo problema preguntaba —de manera más general— por la distribución de los números primos, e incluía tanto la hipótesis de Riemann como la conjetura de Goldbach («cada número par mayor que dos puede escribirse como suma de dos primos», formulada en 1742 y que también permanece abierta). En este libro ya nos hemos topado con otros problemas de esta lista: véanse las páginas 60, 356 y 380.



Ivan Matveevich Vinogradov (1891–1983)



Niels Fabian Helge von Koch (1870–1924)



Lowell Schoenfeld (1920–2002)

Teorema 10.28.
Si la hipótesis de Riemann fuese cierta, entonces

$$\frac{\pi(x)-\operatorname{li}(x)}{\sqrt{x}/\log(x)} = \frac{\psi(x)-x}{\sqrt{x}} + O(1), \qquad x \to \infty,$$

y

$$\pi(x) = \operatorname{li}(x) + O(x^{1/2}\log(x)), \qquad x \to \infty.$$

Obsérvese que los datos experimentales del cuadro 10.1 sugieren que un error de tamaño $O(x^{1/2})$ es muy plausible; pero —de acuerdo a lo que se sabe actualmente— no podemos garantizar que sea cierto. De manera indirecta, esto parece abogar en favor de la veracidad de la hipótesis de Riemann.

Pero todavía podemos incidir más en la relación entre la hipótesis de Riemann y el tamaño del error al aproximar $\pi(x)$ por $\operatorname{li}(x)$. Helge von Koch[12] probó en 1901 que la hipótesis de Riemann es equivalente a que

$$|\pi(x)-\operatorname{li}(x)| \leq C\sqrt{x}\log(x)$$

para alguna constante C. En 1976, Lowell Schoenfeld precisó el resultado anterior, y demostró que la hipótesis de Riemann equivale a

$$|\pi(x)-\operatorname{li}(x)| \leq \frac{1}{8\pi}\sqrt{x}\log(x)$$

para toda $x \geq 2657$.

[12] Este matemático sueco es mucho más conocido por haber sido el descubridor —en 1904— de uno de los primeros objetos fractales en ser descritos, el copo de nieve de von Koch.

Las evidencias numéricas en favor de la hipótesis de Riemann son muy grandes. En 1914, Hardy probó que existen infinitos ceros de la función zeta de Riemann en la recta crítica $\mathrm{Re}(s) = 1/2$. Posteriormente otros matemáticos han precisado este resultado, estableciendo sucesivas estimaciones para la proporción de ceros en esa línea, es decir, cotas inferiores —cada vez algo mayores— para el valor asintótico de

$$\frac{\mathrm{card}\{\text{ceros de } \zeta(s) \text{ en } \mathrm{Re}(s) = 1/2,\ |\mathrm{Im}(s)| < T\}}{\mathrm{card}\{\text{ceros de } \zeta(s) \text{ en } 0 < \mathrm{Re}(s) < 1,\ |\mathrm{Im}(s)| < T\}}$$

cuando $T \to \infty$. Así, por ejemplo, se conoce que más de $2/5$ de los ceros están sobre la recta crítica.

Otros esfuerzos se han destinado a demostrar que no existe ningún cero s en rectángulos $\{1/2 < \mathrm{Re}(s) < 1,\ |\mathrm{Im}(s)| < T\}$ dentro de la banda crítica para valores de T enormes, cada vez mayores según se intensifican las energías dedicadas a ello. Citemos así mismo que, mediante métodos numéricos, se han llegado a calcular miles de millones de ceros de $\zeta(s)$ ordenados en sentido creciente según el valor absoluto de su parte imaginaria, y todos están en la recta crítica. Muchos han sido, en suma, los intentos de obtener información sobre los ceros de $\zeta(s)$. Nosotros no insistiremos más en ello, pero, si el lector así lo desea, puede ver más detalles sobre los esfuerzos numéricos para comprobar la hipótesis de Riemann en [6] (y su extensa lista de referencias).

Pese a todo el trabajo invertido, y aunque es difícil predecir el futuro, la hipótesis de Riemann está, aparentemente, lejos de poder ser probada. De hecho, no se conoce ninguna región de la banda crítica libre de ceros que, asintóticamente, esté separada de la línea $\mathrm{Re}(s) = 1$; en particular, no se sabe si la hipótesis «supongamos que existe un número $r \in [1/2, 1)$ tal que $\zeta(s) \neq 0$ siempre que $\mathrm{Re}(s) > r$» que hemos usado en los teoremas 10.25 y 10.26 es o no cierta.

Si se lograra probar la hipótesis de Riemann, sus repercusiones serían enormes, y no sólo en el término del error en $\pi(x)$. La función zeta de Riemann tiene una profunda conexión con los números primos, y son numerosos los problemas que se verían afectados favorablemente. Así, por ejemplo, ya en el apartado 9.3.1 hemos comentado que —tal como probó Harald Cramér en 1920—, si la hipótesis de Riemann es cierta, entonces existe una constante K tal que la diferencia entre dos primos consecutivos es

$$p_{k+1} - p_k < K p_k^{1/2} \log(p_k).$$

para todo k (por supuesto, estamos usando $\{p_k\}_{k=1}^{\infty}$ para denotar la sucesión de todos los números primos).

Pero no sólo se conocen resultados condicionados a la hipótesis de Riemann en el ámbito estricto de los números primos, sino en otros problemas de la matemática. En

particular, y de manera muy notable, la hipótesis de Riemann implica la existencia de determinadas cotas muy precisas para muchas funciones aritméticas notables, además de la propia función $\pi(x)$ que cuenta los números primos.

Bastantes de los resultados que se sabe cómo probar si se asume que la hipótesis de Riemann es verdadera son, de hecho, equivalentes a ella. Uno de los más interesantes se debe a Littlewood, quien en 1912 demostró que la hipótesis de Riemann es equivalente a que

$$M(x) = \sum_{n \le x} \mu(n) = O(x^r), \qquad x \to \infty,$$

para cada $r > 1/2$. (Mertens había conjeturado en 1897 que $M(x) \le \sqrt{x}$ para todo x, pero Odlyzko y te Riele probaron en 1985 que dicha conjetura no es cierta.)

Para concluir, mencionemos que otro resultado equivalente —y con un enunciado llamativo— es que la función $\sigma(n)$ (suma de los divisores de n) cumpla

$$\sigma(n) < e^{\gamma} n \log(\log(n))$$

para todo $n > 5040$, donde γ es la constante de Euler; fue probado por Guy Robin en 1984. Si el lector quiere conocer más resultados condicionados a la hipótesis de Riemann puede acudir al ya citado [6] y —sobre todo— a [10], que expone 19 problemas equivalentes a la hipótesis de Riemann.

10.7. Las aproximaciones que propuso Riemann

El objetivo de esta sección es intentar autoconvencernos de que hay alguna razón —expresable mediante fórmulas— que relaciona la distribución de los números primos y los ceros complejos de la función zeta de Riemann.

El lector recordará que, en la demostración que hemos dado del teorema de los números primos, hemos usado que cierta región del plano complejo no contiene ceros de $\zeta(z)$. Pero no hemos usado los ceros en sí, sólo que cierta región no los contiene. En lo que sigue, haremos intervenir a esos ceros.

Para ello vamos a remontarnos —una vez más— al famoso artículo de Riemann de 1859, un artículo de sólo ocho páginas llenas de sorprendentes ideas pero que no contenía ninguna demostración. Gauss había tomado como punto de partida el hecho empírico de que la frecuencia de los números primos en la proximidad de un número grande x es casi exactamente $1/\log(x)$, así que el número de primos hasta x debería ser, aproximadamente,

$$\pi(x) \approx \frac{1}{\log(2)} + \frac{1}{\log(3)} + \cdots + \frac{1}{\log(\lfloor x \rfloor)} \approx \int_2^x \frac{dt}{\log(t)} = \mathrm{Li}(x) \approx \mathrm{li}(x).$$

Aquí y en lo que sigue —y como contrapartida a los significados precisos tanto de $f(x) \sim g(x)$ como de las o pequeñas y O grandes—, utilizaremos $f(x) \approx g(x)$ para indicar que $g(x)$ es una aproximación de $f(x)$ en algún sentido, pero sin precisar cuál.

Las investigaciones de Riemann sobre los números primos le sugieren que se puede mejorar la estimación de que la probabilidad de que un número grande x sea primo es $1/\log(x)$; él se percata de que la aproximación $1/\log(x)$ se consigue mucho mejor si se contabiliza no sólo que x sea primo, sino que sea potencia de primo, contando el cuadrado de un primo como la mitad de un primo, el cubo de un primo como un tercio de un primo, etc. Sin más que fijarse en que

$$\operatorname{card}\{p \text{ primo} : p^k \le x\} = \operatorname{card}\{p \text{ primo} : p \le x^{1/k}\} = \pi(x^{1/k}), \tag{10.41}$$

esto conduce a la aproximación

$$\pi(x) + \frac{1}{2}\,\pi(x^{1/2}) + \frac{1}{3}\,\pi(x^{1/3}) + \frac{1}{4}\,\pi(x^{1/4}) + \cdots \approx \operatorname{li}(x). \tag{10.42}$$

Recordemos ahora la fórmula de inversión de Möbius

$$g(x) = \sum_{n=1}^{\infty} \alpha(n) f(x^{1/n}) \;\Leftrightarrow\; f(x) = \sum_{n=1}^{\infty} \mu(n)\alpha(n) g(x^{1/n}),$$

que es válida cuando α es una función completamente multiplicativa (en nuestro caso, $\alpha(n) = 1/n$), y que vimos en el tema 7; en concreto, ésta es la fórmula (7.27). Aplicándola, (10.42) equivale a

$$\pi(x) \approx \operatorname{li}(x) + \frac{\mu(2)}{2}\operatorname{li}(x^{1/2}) + \frac{\mu(3)}{3}\operatorname{li}(x^{1/3}) + \frac{\mu(4)}{4}\operatorname{li}(x^{1/4}) + \cdots. \tag{10.43}$$

Según las ideas de Riemann, tomar varios sumandos en la expresión de la derecha debería proporcionar una mejor aproximación a $\pi(x)$ que la más simple $\pi(x) \approx \operatorname{li}(x)$ derivada de las ideas de Gauss, y que nosotros hemos venido utilizando en la sección 10.6. No se quedó aquí el trabajo de Riemann —de hecho, estamos al principio— pero, antes de proseguir, vamos a insistir un poco en esta aproximación.

En honor de Riemann, la función de la derecha en (10.43) se denota mediante[13]

$$\operatorname{Ri}(x) = \sum_{n=1}^{\infty} \frac{\mu(n)}{n} \operatorname{li}(x^{1/n})$$

[13] La notación $\operatorname{Ri}(x)$ que nosotros vamos a usar no es universal; por ejemplo, en otros sitios usan simplemente $R(x)$.

x	$\pi(x)$	$\pi(x) - \frac{x}{\log(x)}$	$\mathrm{li}(x) - \pi(x)$	$\mathrm{Ri}(x) - \pi(x)$
10^2	25	3	5	1
10^3	168	23	10	0
10^4	1 229	143	17	−2
10^5	9 592	906	38	−5
10^6	78 498	6 116	130	29
10^7	664 579	44 158	339	88
10^8	5 761 455	332 774	754	97
10^9	50 847 534	2 592 592	1 701	−79
10^{10}	455 052 511	20 758 029	3 104	−1 828
10^{11}	4 118 054 813	169 923 159	11 588	−2 318
10^{12}	37 607 912 018	1 416 705 193	38 263	−1 476

Cuadro 10.2: Aproximación a $\pi(x)$ mediante $x/\log(x)$, $\mathrm{li}(x)$ y $\mathrm{Ri}(x)$ (valores redondeados al entero más cercano).

y se conoce con el nombre de función de Riemann (no hay que confundirla con la zeta de Riemann $\zeta(s)$). En principio, no había forma de evaluar los infinitos sumandos $\mathrm{li}(x^{1/n})$ de la definición de $\mathrm{Ri}(x)$ pero, en 1884, el matemático danés J. P. Gram simplificó las cosas al probar que

$$\mathrm{Ri}(x) = 1 + \sum_{n=1}^{\infty} \frac{1}{n\zeta(n+1)} \frac{\log^n(x)}{n!}, \tag{10.44}$$

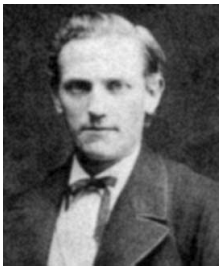
una serie que converge muy rápido incluso para valores grandes de la variable x; podemos ver una demostración de (10.44) en [83, capítulo 2, p. 50]. También Ramanujan estudió la función de Riemann, y dio para ella la siguiente aproximación, esta vez en forma de integral:

$$\mathrm{Ri}(x) = \int_0^{\infty} \frac{(\log(x))^t}{t\Gamma(t+1)\zeta(t+1)}\, dt + o(1) \tag{10.45}$$

(en el integrando, Γ es la función Gamma de Euler, que en este tema nos aparecerá algo más adelante, véase (10.48)); resulta curioso observar que en el denominador interviene la zeta de Riemann.

Ya mostrábamos cómo $\mathrm{li}(x)$ aproxima $\pi(x)$ en el cuadro 10.1. Para ver cómo la función de Riemann $\mathrm{Ri}(x)$ aproxima $\pi(x)$, observemos el cuadro 10.2.

Para valores «pequeños» de x, siempre se cumple $\pi(x) \leq \mathrm{li}(x)$; incluso, si uno se fija en la columna $\mathrm{li}(x) - \pi(x)$ de los cuadros 10.1 o 10.2, da la impresión de que, al



Jørgen Pedersen Gram (1850–1916)



John Edensor Littlewood (1885–1977)

crecer x, las diferencias $\mathrm{li}(x) - \pi(x)$ crecen continuamente hasta infinito. Además, dado que la aproximación a $\pi(x)$ dada en (10.43) comienza por $\mathrm{li}(x) - \mathrm{li}(\sqrt{x})/2$, de nuevo es razonable pensar en la posibilidad de que $\pi(x) - \mathrm{li}(x) \leq 0$ siempre. Pero esto es completamente falso: Littlewood probó en 1914 que $\pi(x) - \mathrm{li}(x)$ cambia de signo infinitas veces. Encontrar el primer cambio de signo es un tema de interés que ha dado lugar a numerosas investigaciones; pero, pese a los esfuerzos realizados, no se conoce ningún x que cumpla $\mathrm{li}(x) < \pi(x)$ (sí se sabe que existe alguno menor que ciertas cotas superiores concretas, pero son verdaderamente enormes).

A la vista del cuadro 10.2, es claro que $\mathrm{Ri}(x) - \pi(x)$ cambia de signo a menudo. Esto hace que, para valores no demasiado grandes de x, siempre se cumpla

$$|\,\mathrm{Ri}(x) - \pi(x)| \leq |\,\mathrm{li}(x) - \pi(x)|. \tag{10.46}$$

Pero tampoco es cierta en general esta desigualdad, sino que incluso puede ser $\mathrm{li}(x) = \pi(x)$ mientras $|\,\mathrm{Ri}(x) - \pi(x)|$ es bastante grande; el lector ya se imaginará que —aunque su existencia está garantizada— no se conocen valores de x concretos que violen la desigualdad (10.46).

10.7.1. Trabajo preparatorio

Aunque Riemann nunca demostró las aproximaciones antes citadas (ni tampoco el teorema de los números primos, claro), hizo algo posiblemente más sorprendente: descubrió una fórmula explícita exacta para $\pi(x)$. Para explicar dicha fórmula —que fue probada por Hans von Mangoldt en 1895— debemos profundizar algo más en el trabajo de Riemann. Así pues, vamos a ver someramente lo que hizo Riemann en su artículo de 1859; pero no seguiremos exactamente ni sus pasos, ni su razonamiento, ni su contenido (unas veces pecaremos por exceso y otras por defecto). Si el lector

está interesado en ver el desarrollo de Riemann con mayor precisión histórica, puede consultar alguno de los numerosos libros o artículos que ofrecen una perspectiva mucho más fiel que la nuestra; en español, citemos el ya mencionado [6].

En nuestro caso, el objetivo que perseguimos no es reproducir fielmente el trabajo de Riemann, sino intentar comprender *por qué* las raíces complejas de la función zeta de Riemann están ligadas, de manera inexorable, con la distribución de los números primos (como es habitual, estamos empleando los términos «ceros» y «raíces» de manera indistinta). Tal como el lector ya habrá supuesto, tampoco daremos demostraciones de lo que ahora expongamos, pues tales demostraciones quedan lejos del alcance de este libro; el rigor matemático se puede encontrar, por ejemplo, en el excelente tratado [40], que además comienza explicando con detalle el artículo de Riemann.

Esencial en la argumentación de Riemann con la función $\zeta(s)$ es su expresión a partir de los números primos que encontrara Euler, es decir,

$$\zeta(s) = \sum_{n=1}^{\infty} \frac{1}{n^s} = \prod_{p} \frac{1}{1 - 1/p^s}; \tag{10.47}$$

pero, por su parte, Riemann considera valores complejos de s (con $\mathrm{Re}(s) > 1$ para que la serie converja), lo cual da lugar una función analítica en el semiplano $\mathrm{Re}(s) > 1$. Pero todo esto ya lo hemos visto al demostrar el teorema de los números primos.

También intervendrá la función Gamma de Euler, que para $\mathrm{Re}(s) > 0$ se define como

$$\Gamma(s) = \int_0^{\infty} e^{-t} t^{s-1}\, dt. \tag{10.48}$$

Esta función no sólo es importante en teoría de números, sino que surge en muchos otros campos de la matemática, y tiene un montón de propiedades interesantes; comentaremos brevemente algunas. Integrando por partes, de la definición se obtiene fácilmente que $\Gamma(s+1) = s\Gamma(s)$, una relación muy útil. En primer lugar, y como $\Gamma(1) = \int_0^{\infty} e^{-t}\, dt = 1$, esa relación implica que

$$\Gamma(n+1) = n\Gamma(n) = n(n-1)\Gamma(n-1) = \cdots = n!\,\Gamma(1) = n!, \qquad n \in \mathbb{N}$$

(y por eso se dice que la función Gamma interpola el factorial). En segundo lugar, permite ir extendiendo Γ a $\mathrm{Re}(s) < 0$ mediante sucesivas iteraciones; en concreto, se toma $\Gamma(s) = \Gamma(s+1)/s$ como definición de Γ en $-n-1 < \mathrm{Re}(s) \leq -n$ una vez que ya está previamente definida en $-n < \mathrm{Re}(s) \leq -n+1$. Con esta técnica, Γ se extiende a todo el plano complejo y resulta una función holomorfa en $\mathbb{C} \setminus \{0, -1, -2, -3\}$ (en

esos puntos hay polos simples).[14] Aparte del valor de Γ en los enteros, también es bien conocido que $\Gamma(1/2) = \sqrt{\pi}$ (y usando $\Gamma(s+1) = s\Gamma(s)$ podemos encontrar el valor de Γ en los «semienteros»); mencionemos así mismo que $\Gamma'(1) = -\gamma$, donde γ es la constante de Euler (esa igualdad está reflejada en el ejercicio 8.9 de este libro). Un par de importantes propiedades de la función Γ extendida a $\mathbb{C}$ son la fórmula de reflexión de Euler

$$\Gamma(s)\Gamma(1-s) = \frac{\pi}{\operatorname{sen}(\pi s)}$$

y la fórmula de duplicación de Legendre

$$\Gamma(s)\Gamma\left(s+\frac{1}{2}\right) = 2^{1-2s}\sqrt{\pi}\,\Gamma(2s)$$

(que se puede generalizar con más factores).

Nótese ahora que, efectuando el cambio de variable $t = nx$ en (10.48), se tiene

$$\Gamma(s) = n^s \int_0^\infty e^{-nx} x^{s-1}\,dx$$

para cada $n \in \mathbb{N}$. Con esto, y si damos por hecho que el intercambio entre el sumatorio y la integral se puede justificar, podemos poner

$$\begin{aligned}\zeta(s)\Gamma(s) &= \sum_{n=1}^\infty \frac{1}{n^s}\,\Gamma(s) = \sum_{n=1}^\infty \int_0^\infty e^{-nx}x^{s-1}\,dx = \int_0^\infty \sum_{n=1}^\infty e^{-nx}x^{s-1}\,dx \\ &= \int_0^\infty \frac{e^{-x}}{1-e^{-x}}\,x^{s-1}\,dx = \int_0^\infty \frac{x^{s-1}}{e^x-1}\,dx,\end{aligned}$$

válido cuando $\operatorname{Re}(s) > 1$.

Además, mediante ciertas técnicas de variable compleja que no detallaremos (integrales sobre caminos), Riemann se da cuenta de que se puede extender la función ζ a todo el plano complejo de manera que se cumple la ecuación funcional

$$\zeta(s) = 2^s\pi^{s-1}\operatorname{sen}\left(\frac{\pi s}{2}\right)\Gamma(1-s)\zeta(1-s),$$

[14] Existen expresiones alternativas para definir la función Gamma de Euler directamente en un rango más grande que el de (10.48), como

$$\Gamma(s) = \lim_n \frac{n!\,n^s}{s(s+1)\cdots(s+n)} = \frac{1}{s}\prod_{n=1}^\infty \frac{(1+1/n)^s}{1+s/n},$$

que es válida para $s \in \mathbb{C}\setminus\{0,-1,-2,-3,\dots\}$.

lo cual relaciona los valores de ζ en los puntos s y $1-s$. De la ecuación funcional se deduce que, en el semiplano $\mathrm{Re}(s) < 0$, la función ζ se anula exclusivamente en los enteros negativos $s = -2, -4, -6, \ldots$ (todos ellos ceros simples), que se denominan ceros triviales. Sabemos que el resto de sus ceros forzosamente deben estar en la banda crítica caracterizada por $0 \leq \mathrm{Re}(s) \leq 1$. De estos ceros, podemos dar un par de propiedades más: en primer lugar, cada uno de ellos está acompañado de su complejo conjugado (como $\overline{\zeta(s)} = \zeta(\overline{s})$, si $\zeta(\rho) = 0$, también deberá ser $\zeta(\overline{\rho}) = 0$); en segundo lugar, la ecuación funcional implica que cada cero ρ se corresponde con un cero $1-\rho$.[15]

Riemann define asimismo la función

$$\xi(s) = \frac{1}{2}\, s(s-1)\Gamma\left(\frac{s}{2}\right)\pi^{-s/2}\zeta(s) \tag{10.49}$$

que es entera (holomorfa en todo el plano complejo), dado que el factor s anula el polo de $\Gamma(s/2)$ en 0, el factor $(s-1)$ anula el polo de $\zeta(s)$ en 1, y los ceros triviales de $\zeta(s)$ anulan los demás polos de $\Gamma(s/2)$. En particular, la función ξ tiene como ceros, exactamente, los ceros no triviales de la función ζ. Además, la ecuación funcional de ζ equivale a decir $\xi(s) = \xi(1-s)$ (ejercicio 10.16).

Es a partir de las propiedades de la función $\xi(s)$ como Riemann llega a conjeturar que todos los ceros no triviales de $\zeta(s)$ deben estar en la recta crítica $\mathrm{Re}(s) = 1/2$ —lo que actualmente se conoce como hipótesis de Riemann—, pero realmente esto no va a intervenir en lo que sigue, por tanto prescindiremos de ello.

El siguiente paso va a ser expresar $\xi(s)$ como producto infinito a partir de sus raíces. Esto es análogo a como un polinomio $P(z)$ que cumple $P(0) \neq 0$ (si $P(0) = 0$ se toma $P(z)/z$, que es otro polinomio), cuyas raíces son $z_1, \ldots, z_n$ (cada una incluida tantas veces como su multiplicidad), y cuyo coeficiente director es c, se puede expresar mediante

$$P(z) = c(z-z_1)\cdots(z-z_n) = P(0)\left(1-\frac{z}{z_1}\right)\cdots\left(1-\frac{z}{z_n}\right);$$

o —con una analogía aún mayor— a como Euler había expresado el seno, cuyas raíces son 0 y $\pm k\pi$, $k \in \mathbb{N}$, mediante

$$\mathrm{sen}(z) = z\prod_{k=1}^{\infty}\left(1-\frac{z}{k\pi}\right)\left(1+\frac{z}{k\pi}\right) = z\prod_{k=1}^{\infty}\left(1-\frac{z^2}{k^2\pi^2}\right)$$

[15] Nótese que para los ceros en la recta crítica, es decir, cuando $\mathrm{Re}(\rho) = 1/2$ (y esto ocurre con todos los ceros no triviales conocidos), se cumple $\overline{\rho} = 1-\rho$, y en este caso ambas propiedades son la misma.

(aunque sin demostrarla de manera rigurosa, Euler había partido de esa expresión para concluir que $\sum_{k=1}^{\infty} 1/k^2 = \pi^2/6$). Riemann sabía que eso no era tan fácil en general, incluso para funciones enteras; por ejemplo, si una función entera $f(z)$ con $f(0) \neq 0$ tiene raíces $\{z_k\}_{k=1}^{\infty}$ puede ocurrir que

$$f(z) = h(z) \prod_{k=1}^{\infty} \left(1 - \frac{z}{z_k}\right)$$

con cualquier función $h(z)$ que no se anule en $\mathbb{C}$ (y esto es posible sin más que pensar en $h(z) = e^{g(z)}$ con $g(z)$ una función entera arbitraria). Pero, mediante consideraciones heurísticas sobre $\xi(s)$ y la distribución de sus ceros, Riemann llega a una expresión que equivale al desarrollo como producto infinito

$$\xi(s) = \xi(0) \prod_{\substack{\rho \in \mathbb{C} \\ \xi(\rho)=0}} \left(1 - \frac{s}{\rho}\right), \tag{10.50}$$

donde se sobreentiende que cada factor aparece en el producto tantas veces como la multiplicidad de ρ como cero de ξ, y que el producto se hace en orden creciente del módulo de los ceros. Con rigor, esta fórmula la probó Hadamard en 1893, en un trabajo en el que investiga el problema general de encontrar expresiones en producto infinito para funciones enteras.

10.7.2. Aparecen los ceros de la función zeta

Una vez hecha la labor preparatoria, volvamos a $\pi(x)$. Retomemos la función que cuenta no sólo los primos $\leq x$, sino que también tiene en cuenta que las potencias de primos puedan ser $\leq x$, y que, según veíamos en (10.41) y (10.42), era

$$\sum_{p^n \leq x} \frac{1}{n} = \sum_{n=1}^{\infty} \frac{1}{n} \pi(x^{1/n}).$$

Pero no vamos a utilizar exactamente esta expresión, sino una sutil variante. Tal como resulta habitual en análisis de Fourier, en lugar de $\pi(x)$ emplearemos una función que coincide con ella en los puntos de continuidad pero que, en los saltos, tome el valor medio. La función $\pi(x)$ tiene un salto de altura 1 en cada primo, así que sustituiremos $\pi(x)$ por

$$\pi_0(x) = \frac{1}{2} \lim_{\varepsilon \to 0^+} \big(\pi(x+\varepsilon) + \pi(x-\varepsilon)\big) = \begin{cases} \pi(p) - 1/2, & \text{si } x = p \text{ primo}, \\ \pi(x), & \text{en otro caso.} \end{cases}$$



Jean Baptiste Joseph Fourier (1768–1830)



Robert Hjalmar Mellin (1854–1933)

Entonces, definamos ahora

$$\Pi_0(x) = \sum_{n=1}^{\infty} \frac{1}{n}\, \pi_0(x^{1/n}), \qquad x \geq 1. \tag{10.51}$$

Obsérvese que, como $\pi_0(x) = 0$ para $1 < x < 2$, esta serie es siempre finita, luego no tiene nunca problemas de convergencia; de manera precisa, lo que ocurre es que, si nos restringimos a $x < 2^{N+1}$, se cumple $x^{1/n} < 2$ para todo $n \geq N+1$ y, por tanto, $\Pi_0(x) = \sum_{n=1}^{N} \pi_0(x^{1/n})/n$. Además, también es claro que $\Pi_0(x) = 0$ para $1 < x < 2$.

Ahora Riemann recurre al producto de Euler (10.47); tomando logaritmos y haciendo el desarrollo en serie de Taylor de cada $\log(1 - p^{-s})$, obtiene

$$\log(\zeta(s)) = -\sum_{p} \log(1 - p^{-s}) = \sum_{p} \sum_{n=1}^{\infty} \frac{1}{n}\, p^{-ns}.$$

A partir de aquí, sustituyendo p^{-ns} por la expresión $s \int_{p^n}^{\infty} x^{-s-1}\, dx$ y usando (10.51), llega a

$$\log(\zeta(s)) = s \int_0^{\infty} \Pi_0(x) x^{-s}\, \frac{dx}{x}, \tag{10.52}$$

que expresa $\log(\zeta(s))$ a partir de $\Pi_0(x)$. Pero, en principio, el objetivo es el contrario, es decir, usar $\log(\zeta(s))$ para llegar a $\Pi_0(x)$. Veamos cómo hacerlo.

Cambiando s por $-s$, la expresión (10.52) responde a la fórmula

$$F(s) = \int_0^{\infty} f(x) x^{s}\, \frac{dx}{x},$$

que es una de las numerosas transformadas integrales que se estudian en matemáticas por muy diversos motivos; en concreto, ésta, que se puede aplicar para $k_1 < \mathrm{Re}(s) < k_2$ si la integral $\int_0^\infty |f(x)|x^{k-1}\,dx$ es finita para los números reales $k = k_1$ y k_2, se conoce actualmente con el nombre de transformada de Mellin en honor al matemático finlandés Hjalmar Mellin,[16] y su fórmula de inversión —que recupera la función de partida bajo determinadas circunstancias de las que aquí no nos preocuparemos— viene dada por la integral

$$f(x) = \frac{1}{2\pi i}\int_{a-i\infty}^{a+i\infty} F(s)x^{-s}\,ds,$$

que existe para $k_1 < a < k_2$. En nuestro caso, al interpretar (10.52) como una transformada de Mellin (y teniendo en cuenta que la función $\Pi_0(x)$ se anula en el intervalo $(0,1)$, luego una de las restricciones sobre $\mathrm{Re}(s)$ se trivializa), la fórmula de inversión queda

$$\Pi_0(x) = \frac{1}{2\pi i}\int_{a-i\infty}^{a+i\infty} \log(\zeta(s))x^s\,\frac{ds}{s}, \qquad a > 1,$$

que proporciona $\Pi_0(x)$ en términos de la función $\log(\zeta(s))$. Integrando por partes, esta expresión se puede escribir como

$$\Pi_0(x) = \frac{-1}{2\pi i}\frac{1}{\log(x)}\int_{a-i\infty}^{a+i\infty} \frac{d}{ds}\left(\frac{\log(\zeta(s))}{s}\right)x^s\,ds \tag{10.53}$$

(pues no resulta difícil comprobar que $\frac{\log(\zeta(s))}{s}\,x^s$ tiende a 0 cuando $s \to a \pm i\infty$).

[16] En general, una transformada integral es una manera de asociar a «cada» función f otra función F definida mediante

$$F(y) = \int_a^b f(x)K(x,y)\,dx,$$

donde $K(x,y)$ es el denominado núcleo de la transformada (en este tema, ya nos hemos topado con una: la transformada de Laplace, en el teorema 10.3). El mayor interés en el campo de las transformadas integrales es proporcionar otra transformada integral —la transformada inversa— que permita recuperar f a partir de F. Posiblemente, la transformada integral más usada es la de Fourier, dada, junto con su inversa, por

$$\widehat{g}(y) = \int_{-\infty}^{\infty} g(x)e^{-ixy}\,dx \;\Leftrightarrow\; g(x) = \frac{1}{2\pi}\int_{-\infty}^{\infty}\widehat{g}(y)e^{iyx}\,dy.$$

La transformada de Mellin de $f(x)$ en s es la transformada de Fourier de $g(x) = f(e^x)$ en is. De hecho, Riemann no emplea transformadas de Mellin, sino que halla la fórmula de inversión que va a usar descomponiendo en parte real e imaginaria y usando transformadas de Fourier seno y coseno (que son casos particulares de la transformada de Fourier dado que $e^{ixy} = \cos(xy) + i\,\mathrm{sen}(xy)$).

Ahora, aplicando (10.49) y (10.50) (y usando $s\Gamma(s) = \Gamma(s+1)$), descomponemos

$$\begin{aligned}\log(\zeta(s)) &= -\log(s-1) - \log(s/2) - \log(\Gamma(s/2)) + \log(\pi^{s/2}) + \log(\xi(s)) \\ &= -\log(s-1) - \log\left(\Gamma\left(\frac{s}{2}+1\right)\right) + \frac{s}{2}\log(\pi) + \log(\xi(0)) \\ &\quad + \sum_{\substack{\rho\in\mathbb{C}\\ \xi(\rho)=0}} \log\left(1-\frac{s}{\rho}\right).\end{aligned}$$

Sustituyendo en (10.53), podemos poner

$$\begin{aligned}\Pi_0(x) = &\frac{-1}{2\pi i}\frac{1}{\log(x)}\int_{a-i\infty}^{a+i\infty}\frac{d}{ds}\left(\frac{-\log(s-1)}{s}\right)x^s\,ds \\ &+\frac{-1}{2\pi i}\frac{1}{\log(x)}\int_{a-i\infty}^{a+i\infty}\frac{d}{ds}\left(\frac{-\log(\Gamma(s/2+1))}{s}\right)x^s\,ds \\ &+\frac{-1}{2\pi i}\frac{1}{\log(x)}\int_{a-i\infty}^{a+i\infty}\frac{d}{ds}\left(\frac{\log(\pi)}{2}\right)x^s\,ds \\ &+\frac{-1}{2\pi i}\frac{1}{\log(x)}\int_{a-i\infty}^{a+i\infty}\frac{d}{ds}\left(\frac{\log(\xi(0))}{s}\right)x^s\,ds \\ &+\frac{-1}{2\pi i}\frac{1}{\log(x)}\int_{a-i\infty}^{a+i\infty}\frac{d}{ds}\left(\frac{\sum_\rho \log(1-s/\rho)}{s}\right)x^s\,ds,\end{aligned}$$

y hay que calcular cada una de estas cinco integrales (no daremos los detalles de cómo hacerlo, que son largos y se pueden ver en [40]). El valor de la primera es

$$\frac{-1}{2\pi i}\frac{1}{\log(x)}\int_{a-i\infty}^{a+i\infty}\frac{d}{ds}\left(\frac{-\log(s-1)}{s}\right)x^s\,ds = \text{v. p.}\int_0^x\frac{dt}{\log(t)} = \text{li}(x);$$

el de la segunda,

$$\frac{-1}{2\pi i}\frac{1}{\log(x)}\int_{a-i\infty}^{a+i\infty}\frac{d}{ds}\left(\frac{-\log(\Gamma(s/2+1))}{s}\right)x^s\,ds = \int_x^\infty\frac{dt}{t(t^2-1)\log(t)};$$

la tercera es obvio que es nula pues en el integrando se deriva una constante; la cuarta vale

$$\frac{-1}{2\pi i}\frac{1}{\log(x)}\int_{a-i\infty}^{a+i\infty}\frac{d}{ds}\left(\frac{\log(\xi(0))}{s}\right)x^s\,ds = -\log(2);$$

y ya sólo queda la última, en cuyo integrando están los ceros no triviales de la función ζ.

Para expresar de manera precisa lo que ocurre con ella debemos introducir un poco de notación. Para números reales u, v con $v \neq 0$ o $v = 0$ y $u < 0$, definamos la función $\operatorname{lie}(u, v)$ mediante la integral compleja[17]

$$\operatorname{lie}(u, v) = \int_{-\infty+iv}^{u+iv} \frac{e^z}{z}\, dz; \tag{10.54}$$

además, para $v = 0$ y $u > 0$ tomemos

$$\operatorname{lie}(u, 0) = \operatorname{li}(e^u),$$

donde $\operatorname{li}(x)$ es el logaritmo integral.

Con esta notación, y tras un poco de esfuerzo, se puede probar que el último sumando que nos quedaba es[18]

$$\frac{-1}{2\pi i}\frac{1}{\log(x)} \int_{a-i\infty}^{a+i\infty} \frac{d}{ds}\left(\frac{\sum_\rho \log(1 - s/\rho)}{s}\right) x^s\, ds \\ = -\sum_\rho \operatorname{lie}\big(\operatorname{Re}(\rho)\log(x), \operatorname{Im}(\rho)\log(x)\big),$$

donde los ρ del sumatorio son los ceros no triviales de la función zeta de Riemann (además, la serie no es absolutamente convergente, y por tanto debe ser sumada en un orden conveniente: de acuerdo con los valores absolutos de $\operatorname{Im}(\rho)$ crecientes).

[17] Tanto Riemann como numerosos textos (incluida la monografía [40]) usan un simple

$$\operatorname{li}(e^{u+iv}) = \int_{-\infty+iv}^{u+iv} \frac{e^z}{z}\, dz$$

para lo que nosotros estamos denotando como $\operatorname{lie}(u, v)$. Pero eso es un claro abuso de notación que, cuando se propaga en otras expresiones (suele aparecer como $\operatorname{li}(x^\rho) = \operatorname{li}(e^{\rho\log(x)})$ donde ρ son los ceros no triviales de la función zeta), puede dar lugar a problemas de interpretación. Por ejemplo, obsérvese que, aunque $e^{u+iv+2\pi i} = e^{u+iv}$, utilizando teoría de variable compleja no es difícil probar

$$\int_{-\infty+iv+2\pi i}^{u+iv+2\pi i} \frac{e^z}{z}\, dz - \int_{-\infty+iv}^{u+iv} \frac{e^z}{z}\, dz = \int_{u+iv}^{u+iv+2\pi i} \frac{e^z}{z}\, dz,$$

que no es nulo en general. Más adelante, en la nota al pie numerada como 20 veremos cómo se puede evitar el abuso de notación de li pero usando en su lugar una verdadera función de variable compleja, no la función de dos variables reales lie que hemos usado ahora para salir del paso.

[18] Con el habitual abuso de notación que hemos explicado en la anterior nota al pie, se suele escribir $\operatorname{lie}(\operatorname{Re}(\rho)\log(x), \operatorname{Im}(\rho)\log(x)) = \operatorname{li}(x^\rho)$.

De este modo, hemos llegado a la importante fórmula

$$\Pi_0(x) = \mathrm{li}(x) - \sum_{\substack{\rho\in\mathbb{C}\\ \xi(\rho)=0}} \mathrm{lie}\big(\mathrm{Re}(\rho)\log(x), \mathrm{Im}(\rho)\log(x)\big) + \int_x^\infty \frac{dt}{t(t^2-1)\log(t)} - \log(2). \tag{10.55}$$

Recordemos que las raíces de ξ son las mismas que las de ζ pero eliminado los llamados ceros triviales $\rho = -2,\ -4,\ -6,\ldots$, que no se usan en la fórmula. Las raíces no triviales son números complejos cuyas partes reales están entre 0 y 1.

Nota 10.29. Aunque no tiene importancia para lo que a nosotros nos interesa, se puede comprobar que

$$\int_x^\infty \frac{dt}{t(t^2-1)\log(t)} = -\sum_{n=1}^\infty \mathrm{li}(x^{-2n}).$$

Como precisamente $\{-2n\}_{n=1}^\infty$ son los ceros triviales de ζ, podemos pensar que también están interviniendo en el valor de $\Pi_0(x)$ en (10.55) a través de esta expresión.[19] ◁

Aún queda un último paso, que consiste en aplicar la fórmula de inversión de Möbius

$$F(x) = \sum_{n=1}^\infty \alpha(n) f(x^{1/n}) \ \Leftrightarrow\ f(x) = \sum_{n=1}^\infty \mu(n)\alpha(n) F(x^{1/n})$$

(donde $\alpha(n)$ es una función completamente multiplicativa y $\mu(n)$ es la función de Möbius). Al aplicarla sobre (10.51) llegamos a

$$\pi_0(x) = \sum_{n=1}^\infty \frac{\mu(n)}{n}\,\Pi_0(x^{1/n}), \tag{10.56}$$

que proporciona $\pi_0(x)$ como una serie en la que los sumandos están dados a partir de una función Π_0 en la que —según (10.55)— ¡intervienen los ceros de la zeta de

[19] De hecho, con el abuso de notación de las dos anteriores notas al pie, la fórmula (10.55) tiene así una expresión muy simple:

$$\Pi_0(x) = \mathrm{li}(x) - \log(2) - \sum_{\substack{\rho\in\mathbb{C}\\ \zeta(\rho)=0}} \mathrm{li}(x^\rho)$$

(obsérvese que ahora estamos usando todos los ceros de ζ, no sólo los de ξ).

Riemann! Como ocurría en (10.51), y dado que $\Pi_0(x) = 0$ para $1 < x < 2$, la serie (10.56) es finita para cada x, cumpliéndose

$$\pi_0(x) = \sum_{n=1}^{N} \frac{\mu(n)}{n} \Pi_0(x^{1/n})$$

siempre que nos restrinjamos a $x < 2^{N+1}$.

10.7.3. Aproximación numérica y gráfica incorporando sucesivos ceros

La fórmula (10.56) (o su versión con un número finito de términos para cada x) es exacta, pero directamente en esa forma no sirve para evaluar $\pi_0(x)$, ya que no es posible evaluar de manera exacta los sumandos $\Pi_0(x^{1/n})$: por una parte, intentar hacerlo mediante (10.51) supondría usar la misma función $\pi_0(x)$ que queremos calcular; y, por otra, la expresión alternativa (10.55) tiene infinitos sumandos en los que ni siquiera conocemos los ρ que allí aparecen.

En cambio, si el objetivo es aproximar $\pi_0(x)$, esa expresión (10.55) sí que va a resultar muy útil. Comencemos fijándonos en que la integral

$$\int_x^\infty \frac{dt}{t(t^2-1)\log(t)}$$

que allí aparece es una función decreciente en x para $x \geq 2$, y su cota superior es

$$\int_2^\infty \frac{dt}{t(t^2-1)\log(t)} = 0.140\,010\ldots;$$

así pues, tanto la contribución de esta integral como de la constante $-\log(2)$ que la acompaña son insignificantes para el valor de $\Pi_0(x)$ cuando x es grande (en particular, y si tenemos en cuenta la nota 10.29, estamos diciendo que los llamados ceros triviales $\rho = -2, -4, -6, \ldots$ no aportan prácticamente nada a la fórmula (10.55)). Como consecuencia, podemos aproximar

$$\Pi_0(x) \approx \mathrm{li}(x) - \sum_\rho \mathrm{lie}(\mathrm{Re}(\rho)\log(x), \mathrm{Im}(\rho)\log(x))$$

donde los ρ del sumatorio son los ceros no triviales de la función zeta de Riemann (ordenados según los valores absolutos de $\mathrm{Im}(\rho)$ crecientes). En una primera tentativa, podríamos quedarnos sólo con

$$\Pi_0(x) \approx \mathrm{li}(x).$$

Siguiendo esta idea, si en (10.56) usamos $\mathrm{li}(x)$ en lugar de $\Pi_0(x)$ tendremos

$$\pi_0(x) \approx \sum_{n=1}^{\infty} \frac{\mu(n)}{n} \mathrm{li}(x^{1/n});$$

pero la expresión de la derecha ya la conocemos, pues es la función $\mathrm{Ri}(x)$ que denominábamos función de Riemann. De hecho, dado que $\pi_0(x)$ y $\pi(x)$ son prácticamente la misma función (son iguales salvo que $\pi_0(x) = \pi(x) - 1/2$ cuando $x = p$ primo), esto nos lleva a

$$\pi(x) \approx \sum_{n=1}^{\infty} \frac{\mu(n)}{n} \mathrm{li}(x^{1/n}) = \mathrm{Ri}(x),$$

la aproximación de Riemann que dábamos en (10.43). Así pues, sólo con esto no hemos encontrado nada nuevo.

Para conseguir algo mejor, no vamos a prescindir totalmente del sumatorio $\sum_\rho$ en la aproximación a $\Pi_0(x)$, sino que iremos añadiendo —en sucesivos pasos— la contribución de los ceros de la función zeta de Riemann. Además, recuperaremos en (10.55) la integral y la constante $-\log(2)$ que antes descartábamos; por simplicidad, lo denotaremos

$$w(x) = \int_x^{\infty} \frac{dt}{t(t^2-1)\log(t)} - \log(2).$$

Teniendo en cuenta que cada cero no trivial de la función zeta de Riemann va acompañado de su conjugado, usemos $\{\rho_j\}_{j\geq 1}$ para denotar la sucesión de ceros con parte imaginaria positiva (ordenados de manera creciente según el valor de la parte imaginaria), y recordemos que $0 < \mathrm{Re}(\rho_j) < 1$ para todo j. Tomando juntos los sumandos correspondientes a ρ_j y $\overline{\rho_j}$, la expresión (10.55) quedará

$$\Pi_0(x) = \mathrm{li}(x) - \sum_{j\geq 1} S(\rho_j, x) + w(x) \tag{10.57}$$

con

$$\begin{aligned} S(\rho,x) &= \mathrm{lie}\big(\mathrm{Re}(\rho)\log(x), \mathrm{Im}(\rho)\log(x)\big) + \mathrm{lie}\big(\mathrm{Re}(\overline{\rho})\log(x), \mathrm{Im}(\overline{\rho})\log(x)\big) \\ &= \mathrm{lie}\big(\mathrm{Re}(\rho)\log(x), \mathrm{Im}(\rho)\log(x)\big) + \mathrm{lie}\big(\mathrm{Re}(\rho)\log(x), -\mathrm{Im}(\rho)\log(x)\big). \end{aligned}$$

Manipulando la integral que define la función lie, no es difícil comprobar que la contribución conjunta de ρ y $\overline{\rho}$ es

$$S(\rho,x) = \mathrm{Ei}(\rho\log(x)) + \mathrm{Ei}(\overline{\rho}\log(x)) = 2\,\mathrm{Re}\big(\mathrm{Ei}(\rho\log(x))\big)$$

donde Ei es la función denominada *integral exponencial*, que para x real positivo se define como

$$\mathrm{Ei}(x) = -\,\mathrm{v.\,p.}\int_{-x}^{\infty} \frac{e^{-t}}{t}\,dt,$$

y se extiende a los complejos como función holomorfa en $\mathbb{C}$ menos el semieje real negativo $(-\infty, 0]$.[20]

Si queremos hacer aproximaciones numéricas, no podemos tomar todos los sumandos en (10.57), sino que debemos truncar la serie quedándonos con un número finito de ceros de $\zeta(s)$. Cada pareja de ceros de más que tomemos mejorará la aproximación de $\Pi_0(x)$. Si usamos los k primeros ceros (y sus conjugados) estaremos aproximando $\Pi_0(x)$ mediante

$$\Pi_{0,k}(x) = \mathrm{li}(x) - \sum_{j=1}^{k} S(\rho_j, x) + w(x);$$

y, a su vez, $\pi_0(x)$ de (10.56) (recordemos que, aunque allí escribíamos $\sum_{n=1}^{\infty}$, esa serie es $\sum_{n=1}^{N}$ para $x < 2^{N+1}$, y lo mismo ocurrirá ahora) se aproximará por

$$\begin{aligned}\pi_{0,k}(x) &= \sum_{n=1}^{\infty} \frac{\mu(n)}{n}\,\Pi_{0,k}(x^{1/n}) = \sum_{n=1}^{N} \frac{\mu(n)}{n}\,\Pi_{0,k}(x^{1/n}) \\ &= \sum_{n=1}^{N} \frac{\mu(n)}{n}\left(\mathrm{li}(x^{1/n}) - \sum_{j=1}^{k} S(\rho_j, x^{1/n}) + w(x^{1/n})\right)\end{aligned}$$

[20] Obsérvese que lo que está ocurriendo es que, con la notación de (10.54) y para $x > 1$,

$$\mathrm{lie}(u\log(x), v\log(x)) = \mathrm{Ei}((u+iv)\log(x))$$

(además, la función li también se relaciona con Ei mediante $\mathrm{li}(x) = \mathrm{Ei}(\log(x))$, $x > 1$). Pero, por otra parte, no es la función Ei la única que se puede emplear para identificar lie y, en consecuencia, $S(\rho, x)$; sin dar detalles, veamos otras que también son comunes en la literatura matemática especializada. En lugar de Ei, podíamos haber usado la función Gamma incompleta, que para $x \geq 0$ (número real) se define como

$$\Gamma(a, x) = \int_{x}^{\infty} t^{a-1}e^{-t}\,dt$$

y luego se extiende a valores complejos mediante prolongación analítica. En nuestro caso interesa $a = 0$, que se relaciona con Ei mediante $\mathrm{Ei}(z) = -\Gamma(0, -z)$. Otra descripción alternativa es usando una de las denominadas integrales exponenciales generalizadas, que se denotan con $E_b(z)$ y se relacionan con la Gamma incompleta mediante $E_b(z) = z^{b-1}\Gamma(1-b, z)$. De aquí se sigue que

$$S(\rho, x) = 2\,\mathrm{Re}(\mathrm{Ei}(\rho\log(x))) = -2\,\mathrm{Re}(\Gamma(0, -\rho\log(x))) = -2\,\mathrm{Re}(E_1(-\rho\log(x)));$$

si hojeamos distintos textos, nos podemos topar con cualquiera de estas notaciones.

$$= \sum_{n=1}^{N} \frac{\mu(n)}{n} \operatorname{li}(x^{1/n}) - \sum_{n=1}^{N} \frac{\mu(n)}{n} \sum_{j=1}^{k} S(\rho_j, x^{1/n}) + \sum_{n=1}^{N} \frac{\mu(n)}{n} w(x^{1/n})$$
$$= \sum_{n=1}^{N} \frac{\mu(n)}{n} \operatorname{li}(x^{1/n}) - \sum_{j=1}^{k} \sum_{n=1}^{N} \frac{\mu(n)}{n} S(\rho_j, x^{1/n}) + \sum_{n=1}^{N} \frac{\mu(n)}{n} w(x^{1/n}).$$

En la última línea de esta cadena de igualdades, tomemos

$$\sum_{n=1}^{N} \frac{\mu(n)}{n} \operatorname{li}(x^{1/n}) = \sum_{n=1}^{\infty} \frac{\mu(n)}{n} \operatorname{li}(x^{1/n}) - \sum_{n=N+1}^{\infty} \frac{\mu(n)}{n} \operatorname{li}(x^{1/n})$$
$$= \operatorname{Ri}(x) - \sum_{n=N+1}^{\infty} \frac{\mu(n)}{n} \operatorname{li}(x^{1/n})$$

con $\operatorname{Ri}(x)$ como en (10.44) o (10.45), que resultan relativamente sencillas de evaluar. Si N es suficientemente grande para que

$$\sum_{n=N+1}^{\infty} \frac{\mu(n)}{n} \operatorname{li}(x^{1/n}) \tag{10.58}$$

sea pequeño, podemos aproximar

$$\sum_{n=1}^{N} \frac{\mu(n)}{n} \operatorname{li}(x^{1/n}) \approx \operatorname{Ri}(x), \tag{10.59}$$

y entonces

$$\pi_{0,k}(x) \approx \pi^*_{0,k}(x) = \operatorname{Ri}(x) - \sum_{j=1}^{k} \sum_{n=1}^{N} \frac{\mu(n)}{n} S(\rho_j, x^{1/n}) + \sum_{n=1}^{N} \frac{\mu(n)}{n} w(x^{1/n}).$$

En resumen, para $x < 2^{N+1}$ (y N grande) podemos escribir

$$\pi_0(x) \approx \pi_{0,k}(x) \approx \pi^*_{0,k}(x) = \operatorname{Ri}(x) + \sum_{n=1}^{N} \frac{\mu(n)}{n} w(x^{1/n}) + \sum_{j=1}^{k} T_j(x) \tag{10.60}$$

donde

$$T_j(x) = -\sum_{n=1}^{N} \frac{\mu(n)}{n} S(\rho_j, x^{1/n})$$

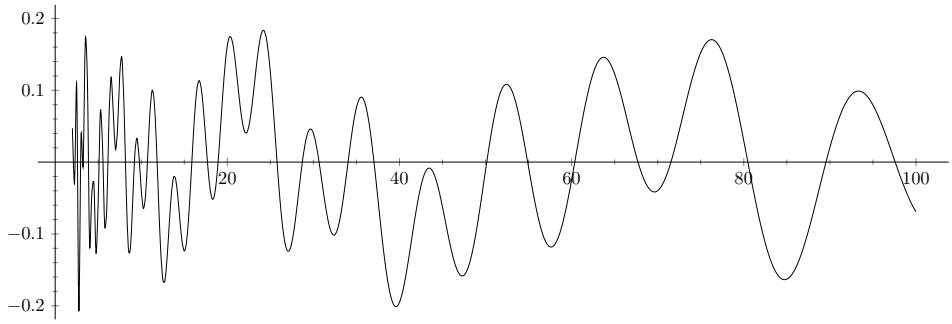


Figura 10.3: Gráfica de la función $T_5(x)$ (realizada con $N = 154$).

es la contribución del j-ésimo par de raíces de la función zeta. Para cada j, $T_j(x)$ es una función oscilante que varía suavemente con x (con oscilaciones cada vez más espaciadas según va creciendo x); según aumenta el valor de j, las funciones $T_j(x)$ presentan, a grandes rasgos, oscilaciones más rápidas pero de menor amplitud. Por poner un ejemplo que no tiene nada de particular, en la figura 10.3 podemos ver una representación gráfica de la función $T_5(x)$.

En (10.60), y con N suficientemente grande para que (10.58) sea pequeño, la parte del león de la aproximación de $\pi_0(x)$ la proporciona $\mathrm{Ri}(x)$; luego está el sumando $\sum_{n=1}^{N} \mu(n) w(x^{1/n})/n$ que no resulta complicado estimar con precisión;[21] y finalmente cada pareja de raíces contribuye a afinar poco a poco el valor de $\pi_0(x)$ (y por tanto de $\pi(x)$).

[21] En la práctica —y tal como se demuestra en [84]—, para evaluar $\sum_{n=1}^{N} \mu(n) w(x^{1/n})/n$ se puede usar que

$$\sum_{n=1}^{N} \frac{\mu(n)}{n}\, w(x^{1/n}) \approx \frac{1}{2\log(x)} \sum_{n=1}^{N} \mu(n) + \frac{1}{\pi} \operatorname{arc\,tg}\left(\frac{\pi}{\log(x)}\right),$$

con el sentido de que la diferencia entre ambas expresiones tiende a 0 cuando $N \to \infty$; de hecho estas aproximaciones son particularmente buenas cuando se usa N tal que $\sum_{n=1}^{N} \mu(n) = -2$ (como ocurre con $N = 154$ o 865). Dado que $w(x)$ está definida como la integral de una función que no tiene una primitiva elemental, es mucho más rápido evaluar la expresión de la derecha que la de la izquierda.

Nótese además que el sumando $-\log(2)$ de la definición de $w(x)$ no tiene en realidad ninguna importancia. Esto es así puesto que, tras usar la fórmula de inversión de Möbius, aparecería en la forma $-\log(2)\sum_{n=1}^{\infty} \mu(n)/n$, y $\sum_{n=1}^{\infty} \mu(n)/n = 0$ es una consecuencia del teorema de los números primos, que ya hemos probado (véase la nota 8.25, en la página 487 de este libro).

x	$\pi(x)-\frac{x}{\log(x)}$	$\mathrm{li}(x)-\pi(x)$	$\mathrm{Ri}(x)-\pi(x)$	$\varepsilon_1(x)$	$\varepsilon_{10}(x)$	$\varepsilon_{40}(x)$
10^2	3	5	1	1	0	0
10^3	23	10	0	1	0	0
10^4	143	17	−2	0	1	1
10^5	906	38	−5	−2	0	−3
10^6	6 116	130	29	24	19	10
10^7	44 158	339	88	61	42	40
10^8	332 774	754	97	70	55	−25
10^9	2 592 592	1 701	−79	75	−100	−101
10^{10}	20 758 029	3 104	−1 828	−1 250	−346	112
10^{11}	169 923 159	11 588	−2 318	−2 144	2 550	1 313
10^{12}	1 416 705 193	38 263	−1 476	−5 867	−3 526	−1 258

Cuadro 10.3: Errores de aproximación a $\pi(x)$ cuando intervienen los ceros de la función zeta de Riemann, con la notación $\varepsilon_k(x)=\pi_{0,k}(x)-\pi(x)$ (valores redondeados al entero más cercano).

Obsérvese el cuadro 10.3, en el que se amplía lo ya visto en los cuadros 10.1 y 10.2. Junto a las aproximaciones a $\pi(x)$ dadas por $x/\log(x)$ y $\mathrm{li}(x)$ se muestra, de manera numérica, cómo la función de Riemann $\mathrm{Ri}(x)=\pi_{0,0}(x)$ y las sucesivas $\pi_{0,k}(x)$ se van aproximando —muy lentamente— a $\pi(x)$ según tomamos mayores valores de k.

Pero, para realmente darse cuenta de cómo $\pi_{0,k}(x)$ aproxima cada vez mejor $\pi(x)$ según vamos aumentando el valor de k, incluso para valores pequeños de la variable x, nada mejor, de nuevo, que recurrir a unos gráficos. Precisamente porque x no es grande, no usaremos la aproximación (10.59), sino que vamos a representar

$$\pi_{0,k}(x)=\sum_{n=1}^{N}\frac{\mu(n)}{n}\,\mathrm{li}(x^{1/n})+\sum_{n=1}^{N}\frac{\mu(n)}{n}\,w(x^{1/n})+\sum_{j=1}^{k}T_j(x).$$

En concreto, en la figura 10.4 se muestran las funciones $\pi_{0,1}(x)$, $\pi_{0,10}(x)$ y $\pi_{0,40}(x)$ en el intervalo $2\le x\le 100$; merece la pena fijarse en cómo se van aproximando hasta los saltos de la función $\pi(x)$. Pese al tiempo transcurrido desde que este tipo de gráficos pudieron verse por primera vez (fue en 1970 en el artículo [84]), para el que esto escribe no dejan de ser sorprendentes.

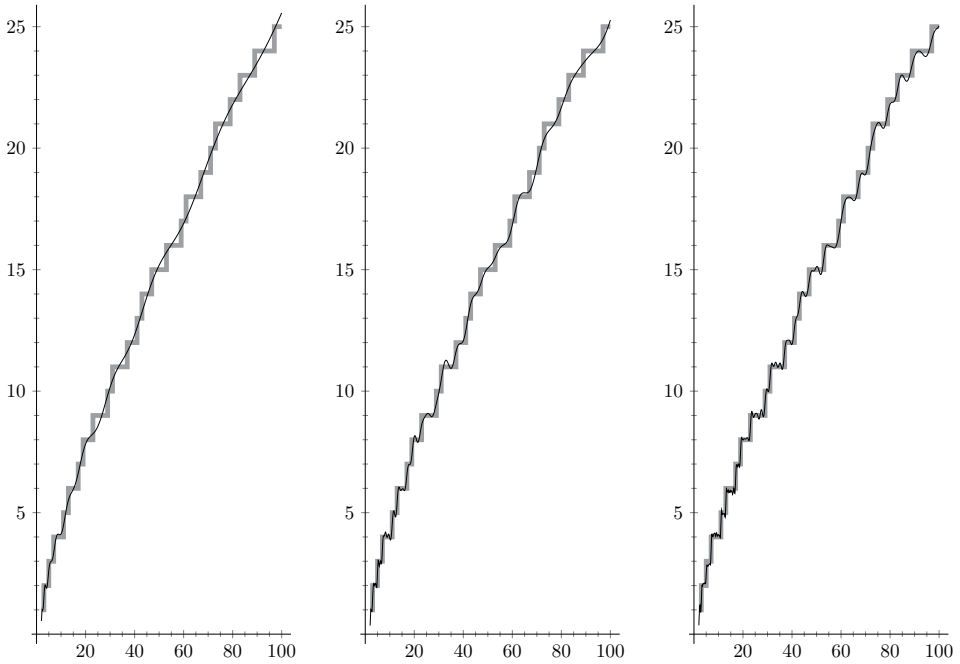


Figura 10.4: De izquierda a derecha, gráficas de las funciones $\pi_{0,k}(x)$ con $k = 1$, 10 y 40 (realizadas con $N = 154$), superpuestas sobre la función $\pi(x)$ (en gris).

10.8. Ejercicios

10.1. Tal como demostrábamos en la sección 8.6 de este libro (en particular, obsérvese la proposición 8.24), el teorema de los números primos se puede demostrar a partir de $\lim_{x\to\infty} \sum_{n\leq x} \mu(n)/n = 0$ (y, de hecho, son resultados equivalentes, según comentábamos en la nota 8.25). Vamos ahora a ver que esa serie no es absolutamente convergente. Más aún, prueba que, si d es un entero positivo y libre de cuadrados (incluyendo el caso $d = 1$), entonces

$$\sum_{n=1}^{\infty} \frac{\mu(dn)}{dn}$$

no converge absolutamente.

10.2. Comprueba que las raíces interesantes de la función $\zeta(s)$, es decir, las raíces $s = \alpha + i\beta$ con $0 < \alpha < 1$, pueden ser caracterizadas en una forma elemental por las dos ecuaciones

$$\sum_{n=1}^{\infty} \frac{(-1)^{n-1}}{n^{\alpha}} \cos(\beta \log(n)) = 0, \qquad \sum_{n=1}^{\infty} \frac{(-1)^{n-1}}{n^{\alpha}} \operatorname{sen}(\beta \log(n)) = 0.$$

Sugerencia. Fíjate en la nota 10.3.

10.3. Prueba que $\zeta(s) < 0$ cuando $0 < s < 1$ (número real).

10.4. Sea $f\colon [1,\infty) \to [0,\infty)$ una función monótona no decreciente y $A \in \mathbb{R}$ tal que

$$\int_1^{\infty} \frac{f(x) - Ax}{x^2}\, dx < \infty.$$

Prueba que, entonces,

$$\lim_{x\to\infty} \frac{f(x)}{x} = A.$$

10.5. Sea $f\colon [1,\infty) \to [0,\infty)$ una función monótona no decreciente que cumple $f(x) = O(x)$ cuando $x \to \infty$, y definamos

$$g(s) = s \int_1^{\infty} \frac{f(x)}{x^{s+1}}\, dx, \qquad \operatorname{Re}(s) > 1.$$

Prueba que, si existe una extensión analítica de g a $\{s \in \mathbb{C} : \operatorname{Re}(s) \geq 1\} \setminus \{1\}$ tal que en $s = 1$ tiene un polo simple con residuo 1, entonces[22]

$$\lim_{x\to\infty} \frac{f(x)}{x} = 1.$$

Sugerencia. Define $F(t) = e^{-t} f(e^t) - 1$ e intenta aprovechar el teorema tauberiano que hemos visto en el texto para llegar a la conclusión de que

[22] El resultado presentado en el ejercicio se conoce con el nombre de teorema tauberiano de Wiener-Ikehara (en realidad, una de sus variantes). Con este teorema tauberiano se puede dar una demostración del teorema de los números primos por medio de su resultado equivalente $\lim_{x\to\infty} \psi(x)/x = 1$, donde $\psi(x) = \sum_{n\leq x} \Lambda(n)$ (ésta es una de las denominadas funciones de Chebyshev que estudiábamos en el apartado 8.7.1). Para ello se usa la relación

$$-\frac{\zeta'(s)}{\zeta(s)} = s \int_1^{\infty} \frac{\psi(x)}{x^{s+1}}\, dx, \qquad \operatorname{Re}(s) > 1,$$

que nosotros no hemos probado pues hemos llegado al teorema de los números primos por un camino distinto.

$\int_1^\infty (f(x)-x)/x^2\,dx$ es una integral convergente. Finalmente, hay que demostrar algo similar a (10.26).

10.6. Prueba que estos tres resultados son equivalentes:

(a) $\pi(x)\log(x) \sim x$ cuando $x \to \infty$ (es decir, el teorema de los números primos).

(b) $\pi(x)\log(\pi(x)) \sim x$ cuando $x \to \infty$.

(c) $p_n \sim n\log(n)$ cuando $n \to \infty$ (donde p_n denota el n-ésimo primo).

Sugerencia. Analiza lo que se hace en la demostración del corolario 10.15 y completa lo que falte.

10.7. Sea $c > 1$. Demuestra que existe un número $x^*(c)$ tal que el intervalo (x, cx) contiene un número primo para todo $x > x^*(c)$.

10.8. Usa el teorema 10.16 para probar el corolario 10.17.

10.9. Usa el teorema 10.19 para probar el corolario 10.20.

10.10. Prueba que, para cada entero positivo n, la función $\mathrm{li}(x)$ cumple[23]

$$\mathrm{li}(x) = \sum_{k=1}^{n} \frac{(k-1)!\,x}{\log^k(x)} + O\left(\frac{x}{\log^{n+1}(x)}\right), \qquad x \to \infty.$$

Sugerencia. En la integral que define $\mathrm{li}(x)$, quédate con el intervalo (e, ∞) e integra por partes sucesivas veces.

10.11. Demuestra que, para $x > 1$, se cumple

$$\mathrm{li}(x) = \int_1^x \frac{1 - t^{-1}}{\log(t)}\,dt + \log(\log(x)) + \gamma,$$

donde γ es la constante de Euler.

Sugerencia. En la definición de $\mathrm{li}(x)$ como un valor principal, haz el cambio $t \mapsto 1/t$ en la integral entre 0 y $1-\varepsilon$. Además, ayúdate de la expresión para la constante de Euler del ejercicio 8.8 (con un cambio de variable para que aparezcan logaritmos).

[23] Esta expresión es un ejemplo de una serie $\sum_{k=1}^{\infty} (k-1)!\,x/\log^k(x)$ no convergente pero que, al tomar un número finito de sumandos, la serie truncada tiene un término del error que es del mismo tamaño que el primer sumando omitido. Tales series se denominan *desarrollos asintóticos*, y son muy útiles en teoría de aproximación.

10.12. Prueba que, para $x > 1$,

$$\mathrm{li}(x) = \sum_{n=1}^{\infty} \frac{\log^n(x)}{n \cdot n!} + \log(\log(x)) + \gamma$$

con γ la constante de Euler y, en particular,

$$\mathrm{li}(e) = \sum_{n=1}^{\infty} \frac{1}{n \cdot n!} + \gamma.$$



10.13. Prueba que el desarrollo en serie de Dirichlet de $\log(\zeta(s))$ es

$$\log(\zeta(s)) = \sum_{n=1}^{\infty} \Lambda_1(n) n^{-s}, \qquad \mathrm{Re}(s) > 1,$$

donde

$$\Lambda_1(n) = \begin{cases} 1/m, & \text{si } n = p^m \text{ para algún primo } p, \\ 0, & \text{en otro caso.} \end{cases}$$



10.14. Comprueba el siguiente desarrollo en serie de Dirichlet:

$$-\frac{\zeta'(s)}{\zeta(s)} = \sum_{n=1}^{\infty} \Lambda(n) n^{-s}, \qquad \mathrm{Re}(s) > 1,$$

donde $\Lambda(n)$ es la función de von Mangoldt.
Sugerencia. Apóyate en el ejercicio anterior.



10.15. Sea un número real $\sigma > 1$. Demuestra que, para cualquier $t \in \mathbb{R}$, se cumple

$$\left|(\zeta(\sigma))^3 (\zeta(\sigma + it))^4 \zeta(\sigma + 2it)\right| \geq 1$$

(esto se denomina desigualdad de Hadamard y De la Vallée-Poussin).
Sugerencia. Fíjate en la parte real del desarrollo en serie de Dirichlet de $\log(\zeta(s))$ y luego usa la desigualdad $2(1 + \cos(\theta))^2 \geq 0$.



10.16. Dando por hecho que disponemos de la ecuación funcional de la función zeta de Riemann, y que, para la función Gamma, conocemos la fórmula de reflexión de Euler y la fórmula de duplicación de Legendre, comprueba que la función ξ definida en (10.49) satisface la ecuación funcional $\xi(s) = \xi(1 - s)$.



Apéndice

La construcción de los números

A.1. Conceptos básicos de teoría de conjuntos
A.2. Definición de cuerpo
A.3. Los números naturales
A.4. Los números enteros
A.5. Los números racionales
A.6. Los números reales
A.7. Los números complejos
A.8. Cuaterniones y octoniones
A.9. Y aún hay más

El rigor exigido por las matemáticas tiene que pagar muchos peajes, y no poder aceptar directamente lo que nos dice la intuición es, sin duda, uno de ellos. Por ejemplo, todos tenemos una idea intuitiva de la recta real, pero ¿qué son los números reales?, ¿de dónde salen?, ¿cómo se construyen?

Por otra parte, en la construcción del edificio de las matemáticas siempre hay que obrar sobre cimientos previamente establecidos. Por mucho que queramos descender, hay que partir de conceptos y propiedades primitivas —o como queramos llamarlos— que damos por preexistentes. Leopold Kronecker, en su celebérrima frase «Dios hizo los números naturales; todo lo demás es obra del hombre», proponía basar las matemáticas en los números naturales, que asumía como entes primigenios.

Nosotros descenderemos un poco más, y construiremos los números sobre la teoría de conjuntos, que asumiremos como preestablecida. En realidad, en el edificio actual de las matemáticas aún se puede construir la teoría de conjuntos sobre bases anteriores (usando los denominados axiomas de Zermelo-Fraenkel), pero eso queda ya lejos de nuestras pretensiones, y no nos vamos a preocupar de ello. Y, desde luego, tampoco vamos a llegar a los terrenos de la lógica matemática y su axiomática.

Así pues, partimos de que en la matemática «básica» existe un concepto denominado *conjunto*, y cada conjunto está constituido por una colección de *elementos* (distintos entre sí) que pertenecen a él. Puede ocurrir que un conjunto no tenga ningún elemento, y entonces se denomina *conjunto vacío* y se denota $\varnothing$. Queremos insistir en que no estamos dando una definición formal de conjunto —eso requeriría hacerlo en función de algo previo—, sino que únicamente estamos usando la idea intuitiva de que un conjunto es una simple reunión de cosas denominadas elementos.[1]

Al contrario que en todos los temas de este libro, en este apéndice ya no vamos a dar referencias bibliográficas. Son muchos los libros de texto que incluyen —como aquí— una introducción a la teoría de conjuntos, o muestran —como nosotros— cómo ir construyendo poco a poco los distintos tipos de números. En particular, la construcción más complicada es, sin duda, la de los números reales, y se expone en numerosos textos de análisis matemático. Seguro que, si el lector busca más información sobre lo que aquí se explica, la encontrará por sí solo sin dificultades, y sin que unas pocas referencias —entre las numerosas que se podían haber dado— le pudieran haber servido de especial ayuda. Lo que sí haremos es incluir comentarios de cuándo y cómo surgieron las distintas ideas y conceptos que tratemos, y quiénes fueron los matemáticos involucrados en ello.

Sin más preámbulos, vamos ya con nuestro objetivo en este apéndice: dar una pequeña introducción a la teoría de conjuntos y, con ella, abordar las definiciones

[1] Prácticamente la única restricción que debemos poner a esta idea intuitiva es que un conjunto no puede ser miembro de sí mismo, pues eso daría lugar a contradicciones lógicas de las que, de nuevo, y pese a ser bien conocidas, aquí no nos ocuparemos.

axiomáticas y las construcciones de los distintos conjuntos de números: naturales, enteros, racionales, reales, y complejos. Esos números que, por otra parte, tan habituados estamos a manejar, y que posiblemente el lector nunca se había planteado antes que había que definir, construir... —lo que sea— de alguna manera. En su día, a todos los matemáticos nos ha pasado lo mismo; y, antes de que el rigor actual se instalara en las matemáticas, muchos no se lo plantearon nunca. En todo caso, si el lector aún necesita convencerse de la necesidad lógica de introducir los números, le sugerimos que se haga la pregunta ¿qué es un número? y trate de responderla.

A.1. Conceptos básicos de teoría de conjuntos

Si usamos X para denotar un conjunto, la expresión $x \in X$ significa que el elemento x pertenece al conjunto X; y $x \notin X$, que no pertenece. Una notación común es señalar los elementos de un conjunto entre llaves, como $X = \{a, b, c\}$ o

$$X = \{\, x : x \text{ cumple una determinada propiedad} \,\}$$

(donde el símbolo «:» se lee «tal que»). Dados dos conjuntos X e Y, su *intersección* $X \cap Y$ y su *unión* $X \cup Y$ son dos nuevos conjuntos definidos mediante

$$X \cap Y = \{\, a : a \in X \text{ y } a \in Y \,\},$$
$$X \cup Y = \{\, a : a \in X \text{ o } a \in Y \,\}.$$

Si todos los elementos de un conjunto X están en otro conjunto Y, se dice que X es un *subconjunto* de Y y se usa la notación $X \subset Y$ o $X \subseteq Y$; su negación es $X \not\subset Y$ o $X \not\subseteq Y$. Dos conjuntos X e Y son iguales cuando se cumple $X \subseteq Y$ e $Y \subseteq X$, y se indica mediante $X = Y$; si no son iguales (que se expresa diciendo que son distintos) se usa $X \neq Y$. Para indicar que $X \subset Y$ pero $X \neq Y$ se emplea $X \subsetneq Y$. Por último, cuando $X \subset Y$, el conjunto de los elementos de Y que no están en X es

$$Y \setminus X = \{\, a : a \in Y \text{ y } a \notin X \,\}.$$

A menudo, si, por algún motivo y en el marco en el que nos encontremos, el conjunto Y es una especie de conjunto «total» presente implícitamente, se suele hablar del complementario de X con el mismo significado que $Y \setminus X$.

Al hablar de conjuntos, hay muchas cosas que podríamos explicar pero que ni siquiera mencionaremos, y algunas de las que sí vamos a ver se podrían abordar con más rigor y profundidad, pero no queremos hacer un tratamiento más profundo. En realidad, sólo pretendemos pasar de puntillas por la teoría de conjuntos para, con

ellos, poder construir los distintos conjuntos de números en las siguientes secciones de este apéndice. Pero sí que hay dos conceptos importantes que debemos exponer: relaciones y aplicaciones.

A.1.1. Relaciones

Dados dos conjuntos X e Y, su producto cartesiano es

$$X \times Y = \{(x, y) : x \in X,\, y \in Y\}$$

donde (x, y) está denotando el «par ordenado» formado por x e y (por supuesto, esta definición de producto cartesiano se puede generalizar con más de dos conjuntos). Un subconjunto de $X \times Y$ se denomina una *relación*, y las que más nos interesan son las relaciones en las que el conjunto X y el Y son el mismo.

Cuando tenemos $R \subset X \times X$ una relación en X, el hecho $(a, b) \in R$ se suele denotar mediante $a\, R\, b$. En tales relaciones, hay tres posibles propiedades muy interesantes:

- Propiedad reflexiva: cuando todos los $a \in X$ cumplen $a\, R\, a$.
- Propiedad simétrica: si, siempre que $a\, R\, b$, también se cumple $b\, R\, a$ (expresado de manera abreviada: $a\, R\, b \Rightarrow b\, R\, a$).
- Propiedad transitiva: si, siempre que $a\, R\, b$ y $b\, R\, c$, también se cumple $a\, R\, c$ (es decir: $a\, R\, b$, $b\, R\, c \Rightarrow a\, R\, c$).

Una relación que cumpla estas tres propiedades se denomina *relación de equivalencia*. En lugar de usar R, las relaciones de equivalencia es común denotarlas con $\sim$.

El ejemplo más sencillo de relación de equivalencia es la relación de igualdad (cada elemento está relacionado sólo consigo mismo). Y, por otra parte, si en una relación de equivalencia identificamos los elementos relacionados, la relación de equivalencia se va a transformar en una especie de igualdad. Veamos cómo hacer esto.

Supuesto que en X tenemos una relación de equivalencia $\sim$, agrupemos cada elemento $a \in X$ con todos los que están relacionados con él. De esta manera, para cada a se forma un conjunto

$$\widehat{a} = \{x \in X : a \sim x\}$$

que es la denominada *clase de equivalencia* de a. Una consecuencia de ser $\sim$ relación de equivalencia es que, si tenemos dos elementos $a, b \in X$, sus respectivas clases de equivalencia $\widehat{a}$ y $\widehat{b}$ son iguales ($\widehat{a} = \widehat{b}$) o disjuntas ($\widehat{a} \cap \widehat{b} = \varnothing$); así, las distintas clases de equivalencia forman lo que se denomina una *partición* de X (por definición, una

partición de un conjunto es una serie de subconjuntos que son disjuntos dos a dos y cuya unión da lugar al conjunto). El conjunto de clases de equivalencia es un nuevo conjunto, que se denomina *conjunto cociente* y se denota

$$X/\sim = \{\,\widehat{a} : a \in X\}.$$

Así, si $a, b \in X$ cumplen $a \sim b$, sus clases de equivalencia serán $\widehat{a} = \widehat{b}$ y, por tanto, serán el mismo elemento en $X/\sim$. La relación de equivalencia $\sim$ en X se ha transformado en la igualdad en el conjunto $X/\sim$.

Volviendo a una relación genérica R en un conjunto X, en lugar de la propiedad simétrica también se habla de esta otra:

- Propiedad antisimétrica: cuando, si se cumple tanto $a\ R\ b$ como $b\ R\ a$, forzosamente debe ser $a = b$ (con símbolos: $a\ R\ b$, $b\ R\ a \Rightarrow a = b$).

Una relación R reflexiva, antisimétrica y transitiva se denomina *relación de orden*, y es habitual denotarla mediante $\leq$; se dice asimismo que $(X, \leq)$ es un *conjunto ordenado*. Con el mismo significado que $a \leq b$ también se emplea $b \geq a$; si además de $a \leq b$ queremos asegurar que $a \neq b$, se puede usar $a < b$ o $b > a$.

En un conjunto ordenado, si siempre se cumple $a \leq b$ o $b \leq a$, el orden se denomina *total*; en otro caso nos encontramos ante un orden *parcial*.

Cuando S es un subconjunto de un conjunto ordenado X, decimos que $x \in X$ es una *cota superior* de S si $x \geq t$ para cualquier $t \in S$. Si existe, a la menor de las cotas superiores se le llama *supremo* de S; y, si el supremo está en S, se dice además que es el *máximo* de S. Por supuesto, de manera análoga se habla de *cota inferior*, *ínfimo* y *mínimo*. Acabemos esta retahíla de definiciones diciendo que un *conjunto bien ordenado* (o que cumple el *principio de buena ordenación*) es un conjunto ordenado tal que todo subconjunto no vacío tiene un mínimo.

Nótese que, para que un conjunto pueda estar bien ordenado, necesariamente debe tener un orden total. Esto es así porque, en particular, cualquier subconjunto $\{a, b\}$ debe tener mínimo, luego deberá ser $a \leq b$ o $b \leq a$.

A.1.2. Aplicaciones entre conjuntos

Dados dos conjuntos X e Y, una aplicación es una «regla» que a cada elemento de X asocia un elemento de Y (y sólo uno). Si a la regla la denominamos f, esto se expresa con

$$f\colon X \to Y; \tag{A.1}$$

más en concreto, si f asocia $x \in X$ con $y \in Y$ se escribe $y = f(x)$ (y se dice que y es la *imagen* de x), y también se puede usar la notación $x \mapsto y$. En (A.1), X se

denomina *dominio*, y el conjunto $f(X) = \{f(x) : x \in X\}$ es la *imagen* o el *recorrido* de f.[2]

Hay algunos tipos de aplicaciones que tienen un nombre específico:

- Una aplicación $f\colon X \to Y$ se dice *inyectiva* si $f(a) \neq f(b)$ siempre que $a \neq b$ (o, expresado de manera equivalente: $f(a) = f(b) \Rightarrow a = b$).
- Una aplicación $f\colon X \to Y$ se dice *suprayectiva* (o *sobreyectiva*) si para cada $y \in Y$ existe $x \in X$ tal que $y = f(x)$.
- Una aplicación $f\colon X \to Y$ se dice *biyectiva* si es inyectiva y biyectiva a la vez.

En lugar de aplicaciones, y con el mismo significado, también se habla de *funciones*; comúnmente, el término función se utiliza cuando se trata de una aplicación entre conjuntos de números (que veremos más adelante). Como deformación del inglés *map*, en algunos países hispanos también se usa el término *mapeo*; igualmente por su notación inglesa, a veces se dice que una aplicación es 1-1 con el significado de que es una aplicación inyectiva.

A.2. Definición de cuerpo

Recordemos que el objetivo de este apéndice es construir, poco a poco, los distintos tipos de números. En los conjuntos de números que vamos a manejar siempre va a haber dos operaciones, la suma y el producto. Por descontado, estas operaciones cumplirán una serie de propiedades que será conveniente reseñar. Existen, además, dos operaciones inversas de la suma y el producto, que son la resta y la división; pero no siempre es posible restar o dividir.

De hecho, una de las razones por las que un conjunto de números se nos puede quedar corto, y sentimos la necesidad de extenderlo, es porque dentro de ese conjunto de números no se puede realizar una de esas operaciones inversas. Por ejemplo, con los naturales no siempre se puede restar, y para lograrlo construimos los enteros. Con los enteros no siempre se puede dividir, y para lograrlo construimos los racionales.

Cuando llegamos a la que —en este contexto— es la situación ideal, es porque el conjunto de números que tenemos es lo que se denomina un *cuerpo* (a veces se emplea también la denominación *campo*, como traducción literal del inglés *field*). Concretando un poco más, un cuerpo será, por definición, un conjunto, con dos operaciones en él, que tiene una serie de propiedades que en seguida detallaremos.

[2] Con un rigor que no necesitamos, una aplicación f se define, a partir de conceptos anteriores, como una relación $f \subset X \times Y$ tal que, para todo $x \in X$, existe un único $y \in Y$ que cumple $(x, y) \in f$.

Por supuesto, la de cuerpo no es la única estructura algebraica importante; cuando tenemos un conjunto con una o varias operaciones que cumplen una serie de propiedades, es habitual que las matemáticas le hayan asignado un nombre específico. Si sólo tenemos en cuenta una operación, existen los conceptos de semigrupo y grupo, con dos operaciones, además de cuerpo hay otra estructura denominada anillo, con tres podemos hablar de álgebras..., y siempre con muchas variantes (por ejemplo, según exijamos o no asociatividad o conmutatividad). Una de las labores más interesantes —y difíciles— de los matemáticos es intentar clasificar esas estructuras. Pero no es eso lo que a nosotros nos ocupa.

En realidad, ni siquiera nos vamos a preocupar de definir estas estructuras algebraicas (hay muchísimos libros de texto que lo hacen), ni aun las que responden a propiedades que cumplen los distintos tipos de números que vamos a ir construyendo, aunque a veces las nombraremos. Sólo nos vamos a ocupar de la que —en lo que a nosotros respecta— es el culmen: la estructura de cuerpo.

Y, además, nos ocuparemos de ella únicamente por el mero hecho de dar un listado conjunto de todas las propiedades mediante las que se define un cuerpo, como referencia para otras partes de este apéndice. Que el lector lo disculpe si lo considera totalmente innecesario.

Así pues, por definición, un *cuerpo* es un conjunto F (de objetos de cualquier especie), junto con dos operaciones $+$ y $\cdot$ (llamadas suma y producto) definidas sobre F (es decir, dos reglas que asocian a cada par de elementos $(a,b) \in F \times F$ otros elementos $a+b$ y $a \cdot b$ de F) de manera que se cumplen las siguientes propiedades:

- Propiedad asociativa para la suma: $(a+b)+c = a+(b+c)$ para todo $a,b,c \in F$.
- Propiedad conmutativa para la suma: $a+b = b+a$ para todo $a,b \in F$.
- Existe algún elemento $0 \in F$ (denominado elemento neutro de la suma) tal que
  - $a+0 = a$ para todo $a \in F$;
  - para todo $a \in F$, existe algún elemento $b \in F$ tal que $a+b = 0$ (ese b es el inverso respecto a la suma de a, que se denota mediante $-a$ y se suele aludir a él diciendo que es el elemento opuesto de a).
- Propiedad asociativa para el producto: $(a \cdot b) \cdot c = a \cdot (b \cdot c)$ para todo $a,b,c \in F$.
- Propiedad conmutativa para el producto: $a \cdot b = b \cdot a$ para todo $a,b \in F$.
- Existe algún elemento $1 \in F$ (denominado elemento neutro del producto) tal que $1 \neq 0$ y
  - $a \cdot 1 = a$ para todo $a \in F$;
  - para todo $a \in F$ con $a \neq 0$, existe algún elemento $b \in F$ tal que $a \cdot b = 1$ (ese b se llama inverso de a y se denota mediante a^{-1}).

- Propiedad distributiva del producto respecto a la suma: $a\cdot(b+c) = (a\cdot b)+(a\cdot c)$ para todo $a, b, c \in F$.

De todas estas propiedades, la más difícil de conseguir es la existencia de inverso respecto al producto; cuando esa propiedad falla, lo que se tiene, en lugar de cuerpo, se denomina *anillo* (a su vez, y según se cumplan o no determinadas propiedades adicionales que ni siquiera mencionaremos, hay distintos tipos de anillos que difieren enormemente en su comportamiento multiplicativo).

Adelantando acontecimientos, y por poner algún ejemplo, no son cuerpos los números naturales (sección A.3) ni los enteros (sección A.4), pero sí lo son los racionales (sección A.5), los reales (sección A.6) y los complejos (sección A.7).

De otra naturaleza, tampoco forman cuerpo los polinomios (bien sea con coeficientes racionales, reales o complejos), pero sí lo forman las denominadas funciones racionales, es decir, los cocientes de polinomios.

A.3. Los números naturales

No estamos diciendo nada nuevo al escribir que los números naturales son los números que usamos para contar, esto es, $1, 2, 3, 4, \ldots$, y que también se emplean para ordenar, esta vez con el nombre de «primero», «segundo», «tercero», «cuarto»... La humanidad los ha venido usado desde tiempo inmemorial, y desde niños aprendemos a manejarlos y a operar con ellos.

No hay un acuerdo universal sobre si debemos considerar el 0 también como número natural. Históricamente, el 0 como número tuvo un origen muy posterior al de los otros números. Los babilonios, en el siglo VII a.C., tenían un símbolo para el cero, pero sólo para no dejar confusos espacios vacíos al representar cantidades en su sistema de numeración posicional en base 60 con escritura cuneiforme. Sabemos que el 0 sí que era considerado como un número más por el matemático hindú Brahmagupta en el siglo VII; y, posiblemente, las civilizaciones olmeca y maya lo usaban, también como número en sí mismo, ya varios siglos antes. En cambio, los griegos de los que deriva la cultura occidental no tenían el concepto de 0; y, desde luego, tampoco había manera de representarlo con números romanos, de uso común en Europa hasta que el sistema posicional en base diez —de origen indio y adoptado por los árabes— comenzó a tomar auge lentamente a partir del siglo XIII.

Ahora, por descontado, ya no tenemos ninguna dificultad en considerar al 0 como un número más. Pero, ¿lo incluimos en los naturales o no? Realmente, es una cuestión sin trascendencia, y lo normal es tomarlo o no como número natural según lo que parezca más cómodo para lo que estamos haciendo. En teoría de números, posiblemente lo más habitual es considerar los números naturales sin el 0 (que será,



Giuseppe Peano
(1858–1932)



Georg Cantor
(1845–1918)

entonces, un número entero, junto con los negativos); y así lo hemos hecho en este libro, con lo cual tendríamos

$$\mathbb{N} = \{1, 2, 3, 4, 5, 6, \dots\}.$$

Sin embargo, en este apéndice nos va a resultar más cómodo usarlos con el 0; así que, para evitar toda confusión, los vamos a denotar de manera distinta:

$$\mathbb{N}_0 = \{0, 1, 2, 3, 4, 5, 6, \dots\}.$$

La razón de este cambio es que, en lo que sigue, vamos a definir y construir los números naturales —y las operaciones con ellos— a partir de conceptos matemáticos previos, y con el 0 (que es el elemento neutro de la suma) resulta algo más sencillo. Bien es verdad que las diferencias no son muchas —y que realmente se podría hacer tanto con $\mathbb{N}$ como con $\mathbb{N}_0$—, pero, en ocasiones, poder usar el 0 como natural simplifica la cadena de definiciones que vamos a ir dando (por ejemplo, para definir la relación de orden $\leq$ a partir de $+$ como «$a \leq b$ si y sólo si $a+c=b$ para algún c» hay que permitir que pueda ser $c=0$; o, alternativamente, definir previamente $<$, que no es relación de orden pues no cumple la propiedad reflexiva). No obstante, comentaremos cómo se haría con $\mathbb{N}$ en lugar de con $\mathbb{N}_0$.

A.3.1. Los axiomas de Peano

Entre los matemáticos, la necesidad de introducir un formalismo para la aritmética —más allá del uso intuitivo de los números— comenzó a gestarse a mediados del siglo XIX. Tras varios intentos de proporcionar una definición axiomática de los números naturales, fue la propuesta en 1889 por el matemático italiano Giuseppe

Peano —basada en las ideas de la teoría de conjuntos que por aquel entonces desarrollaba Georg Cantor— la que finalmente tuvo éxito. Esencialmente, es la misma que se sigue usando en la actualidad.

Los axiomas de Peano (también conocidos por el nombre de *postulados de Peano*) definen los números naturales como un conjunto —que vamos a denotar $\mathbb{N}_0$—, un elemento —que va a asumir el papel de cero, y que lógicamente denotaremos como 0— y un concepto «siguiente» (o «sucesor») —que a la postre va a ser una aplicación denotada mediante s— de manera que se cumple:

1. El 0 es número natural.
2. El siguiente de un número natural es también un número natural.
3. No existe un número natural cuyo siguiente sea 0.
4. Si los siguientes de dos números naturales son iguales, entonces los dos números son iguales.
5. Si S es un conjunto de números naturales tal que 0 es de S y, siempre que un número natural es de S también su siguiente está en S, entonces S es el conjunto de números naturales.

Con un lenguaje más algebraico, estos cinco axiomas también podemos escribirlos así:

1. $0 \in \mathbb{N}_0$.
2. Existe una aplicación $s : \mathbb{N}_0 \to \mathbb{N}_0$ que además va a cumplir los tres axiomas que siguen.
3. No existe ningún $n \in \mathbb{N}_0$ tal que $s(n) = 0$.
4. $s(n) = s(m) \Rightarrow n = m$ (es decir, s es inyectiva).
5. Si S es un conjunto de números naturales tal que $0 \in S$ y, siempre que $n \in S$ también $s(n) \in S$, entonces $S = \mathbb{N}_0$.

De hecho, aún podemos resumir todo esto diciendo simplemente que los números naturales son una terna $(\mathbb{N}_0, 0, s)$ con $\mathbb{N}_0$ un conjunto, $0 \in \mathbb{N}_0$, $s : \mathbb{N}_0 \to \mathbb{N}_0 \setminus \{0\}$ una aplicación inyectiva, y de manera que se cumpla el quinto de los axiomas anteriores, que se denomina *axioma de inducción* o principio de inducción.

Aunque pueda resultar obvio, merece la pena recalcar que el axioma de inducción captura la idea de inducción matemática, y es el que permite que podamos hacer demostraciones por inducción. Para ello, se toma S como el conjunto de los números naturales que cumplen una determinada propiedad que queremos demostrar, se ve que 0 cumple la propiedad (es decir, $0 \in S$), y que, si un número la cumple, también lo hace el siguiente (o sea, que $n \in S$ implica $s(n) \in S$); entonces, y como consecuencia del axioma de inducción, son todos los números naturales los que cumplen la propiedad ($S = \mathbb{N}_0$).

Nota A.1. Si hubiésemos optado por definir los naturales sin el cero, en los axiomas de Peano bastaría cambiar $\mathbb{N}_0$ por $\mathbb{N}$ y 0 por 1. Y es que, en realidad, el primer axioma por sí mismo sólo sirve para asegurar que $\mathbb{N}_0$ no es el conjunto vacío, y el nombre que demos al primer elemento no tiene, de momento, ninguna importancia. Donde realmente se distingue si estamos definiendo los naturales empezando en el 0 o en el 1 es cuando definamos la suma y el producto, tal como veremos un poco más adelante, en la nota A.2. ◁

Con los axiomas de Peano, y una vez que tenemos que 0 es el primer elemento de $\mathbb{N}_0$, introducimos la notación de los demás: a $s(0)$ lo llamamos 1, a $s(1)$ lo llamamos 2, a $s(2)$ lo llamamos 3, y así sucesivamente. Ahora, una vez que tenemos los números naturales, tenemos que ver cómo operar con ellos.

La suma (o adición) en $\mathbb{N}_0$ es una operación $+\colon \mathbb{N}_0 \times \mathbb{N}_0 \to \mathbb{N}_0$ que se define recursivamente mediante

$$\begin{aligned} a + 0 &= a, \\ a + s(b) &= s(a + b). \end{aligned}$$

Por ejemplo, $a+1 = a+s(0) = s(a+0) = s(a)$, y $a+2 = a+s(1) = s(a+1) = s(s(a))$. El hecho $a+0 = a$ (que forma parte de la definición) tiene nombre: 0 es el elemento neutro para la suma.

Una vez definida la suma, el producto (o multiplicación) $\cdot\colon \mathbb{N}_0 \times \mathbb{N}_0 \to \mathbb{N}_0$ se define, de nuevo recursivamente, mediante

$$\begin{aligned} a \cdot 0 &= 0, \\ a \cdot s(b) &= a + a \cdot b \end{aligned}$$

(con más detalle sobre el orden de las operaciones, $a+a\cdot b$ significa $a+(a\cdot b)$; aunque seguro que el lector ya lo sabía, quizás no esté de más recordarlo cuando estamos precisando las definiciones). Esta vez, que 1 es el elemento neutro del producto requiere una pequeña comprobación: $a\cdot 1 = a\cdot s(0) = a+(a\cdot 0) = a+0 = a$. Aunque en esta sección no lo vamos a hacer, lo normal —siempre que no pueda dar lugar a confusión— es escribir el producto de a y b como simple yuxtaposición de términos, es decir, es común usar ab con el mismo significado que $a \cdot b$.

Nota A.2. Si se construyen los naturales empezando en 1, la definición de la suma se comienza por $a + 1 = s(a)$, y la del producto por $a \cdot 1 = a$. ◁

De los axiomas de Peano y las definiciones recursivas que acabamos de dar se demuestran con facilidad las propiedades básicas de $+$ y $\cdot$. En particular, si usamos a, b, c para denotar números naturales cualesquiera, se cumple lo siguiente:

- Propiedad asociativa de la suma: $a+(b+c) = (a+b)+c$ (como consecuencia, no hace falta indicar el orden de esta operaciones mediante paréntesis, y podemos escribir $a + b + c$ sin ningún temor).
- Propiedad conmutativa de la suma: $a + b = b + a$.
- Propiedad asociativa del producto: $a\cdot(b\cdot c) = (a\cdot b)\cdot c$ (y de nuevo los paréntesis son innecesarios).
- Propiedad conmutativa del producto: $a \cdot b = b \cdot a$.
- Cancelación (simplificación) en la suma: si $a + c = b + c$ entonces $a = b$
- Cancelación (simplificación) en el producto: si $a \cdot c = b \cdot c$ con $c \neq 0$ entonces $a = b$.
- Propiedad distributiva del producto respecto a la suma: $a\cdot(b+c) = (a\cdot b)+(a\cdot c)$.

No vamos a demostrar todas estas propiedades, sino que, para hacernos una idea de lo que habría que hacer en cada caso, vamos simplemente a comprobar la propiedad asociativa de la suma. Para ello, dados dos números $a, b \in \mathbb{N}_0$, definamos el conjunto

$$S = \{n \in \mathbb{N}_0 : a + (b + n) = (a + b) + n\}.$$

Es claro que $0 \in S$, pues $(a + b) + 0 = a + b = a + (b + 0)$. Además, si suponemos que $n \in S$, se tendrá

$$(a + b) + s(n) = s((a + b) + n) = s(a + (b + n)) = a + s(b + n) = a + (b + s(n)),$$

luego también $s(n) \in S$. Entonces, el axioma de inducción asegura que $S = \mathbb{N}_0$, y por tanto $a+(b+c) = (a+b)+c$ es cierto para cualquier $c \in \mathbb{N}_0$, tal como queríamos comprobar.

Otro ingrediente importante de los números naturales es el orden. La relación de orden $\leq$ en $\mathbb{N}_0 \times \mathbb{N}_0$ se define como sigue: para cada $a, b \in \mathbb{N}_0$, decimos que $a \leq b$ si y sólo si existe algún $c \in \mathbb{N}_0$ tal que $a + c = b$.

Por descontado, hay que comprobar que, efectivamente, esa definición da lugar a una relación de orden, y para ello hay que ver que se cumplen las propiedades reflexiva, antisimétrica y transitiva. Como cuando estábamos demostrando las propiedades de la suma y el producto, tales comprobaciones hay que hacerlas a partir de los axiomas de Peano; aunque laborioso, y posiblemente tedioso, es un trabajo sencillo, que no haremos para no alargarnos demasiado.

La relación de orden en $\mathbb{N}_0$ es estable con la suma y el producto, es decir, que para $a, b, c \in \mathbb{N}_0$ se cumple:

- $a \leq b \Rightarrow a + c \leq b + c$,
- $a \leq b \Rightarrow a \cdot c \leq b \cdot c$.

Demostrar estas propiedades —o muchas otras igualmente elementales que podríamos haber dado— se logra de nuevo sin dificultad utilizando los axiomas de Peano y, sobre todo, el principio de inducción. Una vez más, no vamos a dar las demostraciones.

Sí que queremos ver con detalle la siguiente propiedad (cuya demostración ilustra de nuevo el tipo de técnica que se usa constantemente):

Proposición A.3.

$(\mathbb{N}_0, \leq)$ es un conjunto bien ordenado (es decir, cualquier subconjunto de $\mathbb{N}_0$ no vacío tiene mínimo).

Demostración. Vamos a probarlo por reducción al absurdo. Supongamos que existe un subconjunto $A \subset \mathbb{N}_0$, $A \neq \varnothing$, que no tiene mínimo. Entonces, definimos el conjunto

$$S = \{n \in \mathbb{N}_0 : n \leq a \text{ para todo } a \in A\}.$$

Observemos en primer lugar que, si existiese $n \in A \cap S$, ese n sería $n = \min S$, y hemos supuesto que S no tiene mínimo; por tanto, $A \cap S = \varnothing$.

Vamos ahora a sacar partido del axioma de inducción. Es claro que $0 \in S$, pues $0 + a = a$ así que $0 \leq a$ para todo $a \in A$. Por otra parte, si $n \in S$, al ser $A \cap S = \varnothing$ tendremos que $n < a$ para todo $a \in A$, y (por la definición del orden en $\mathbb{N}_0$) en cada caso existirá un número natural $n_a \geq 1$ tal que $n + n_a = a$, con lo cual siempre $s(n) = n + 1 \leq n + n_a = a$, y por tanto $s(n) \in S$. En estas condiciones, el axioma de inducción nos asegura que $S = \mathbb{N}_0$. Pero esto es absurdo, pues de $A \cap S = \varnothing$ se deduce que entonces $A = \varnothing$, que es una clara contradicción con la suposición de partida $A \neq \varnothing$. □

A.3.2. Un modelo de los números naturales

Cuando damos una teoría axiomática basada en algo preexistente, podemos ponernos a deducir propiedades a partir de los axiomas que hemos introducido. Pero, ¿cómo podemos estar seguros de que al final no va a aparecer algún tipo de contradicción que haga que esos axiomas no tengan sentido? Es lo que se denomina que la teoría axiomática sea consistente. Lo que hay que hacer para solucionar esto es dar un modelo basado en lo que hayamos asumido como preexistente. Sin eso, podríamos haber estado divagando en el vacío.

Los axiomas de Peano introducen los números naturales a partir de la teoría de conjuntos. Tras exponer los axiomas, hemos dado algunas definiciones adicionales (suma, producto y orden) y probado bastantes propiedades. Pero, ¿son consistentes los axiomas de Peano? Es decir, ¿existe un modelo basado en la teoría de conjuntos que cumpla los axiomas de Peano? ¿O hemos estado construyendo castillos en el

aire? Vamos pues a ver cómo construir un modelo de los números naturales a partir de la teoría de conjuntos.

Nota A.4. Al disponer de un modelo, estaríamos reduciendo la consistencia de los números naturales a la de la teoría de conjuntos. ¿Y la consistencia de la teoría de conjuntos? En algún momento hay que dejar de descender, y hay que asumir algo como cierto sin posibilidad de ser demostrado. De hecho, y de manera desconcertante para los matemáticos de la época, Kurt Gödel probó en 1930 sus famosos teoremas de incompletitud que, simplificando un poco, afirman lo siguiente:

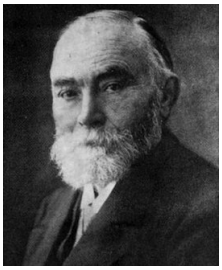
- Primer teorema de incompletitud: En cualquier formalización consistente de las matemáticas que sea lo bastante fuerte para definir el concepto de números naturales, se puede construir una afirmación que ni se puede demostrar ni se puede refutar dentro de ese sistema.
- Segundo teorema de incompletitud: Ningún sistema consistente que sea lo bastante fuerte para definir el concepto de números naturales se puede usar para demostrar su propia consistencia.

Estos resultados fueron devastadores para la aproximación filosófica a las matemáticas conocida como programa de formalización de Hilbert. David Hilbert —el matemático más influyente de comienzos del siglo XX— había propuesto que la consistencia de los sistemas más complejos, tales como el análisis real, se tenía que probar en términos de sistemas más sencillos. Finalmente, la consistencia de todas las matemáticas se podría reducir a la aritmética básica. Pero, en su primer teorema, Gödel demostró que la aritmética es incompleta; y, en el segundo, que dicha aritmética no puede probar su consistencia, y de ese modo será imposible que pueda ser usada para demostrar la consistencia de cualquier otro sistema de axiomas. El paraíso ideado por Hilbert no existía. Tras Gödel, las matemáticas se han tenido que acostumbrar a ello. ◁

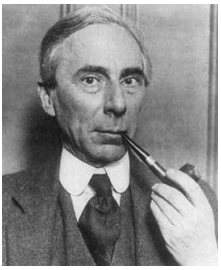
Pero dejemos la lógica y volvamos a los números naturales. Un modelo de los axiomas de Peano sería una tripleta $(N, 0, s)$ —con N un conjunto, $0 \in N$ y $s\colon N \to N$ una aplicación— que satisficiera dichos axiomas. No es difícil ver que, si tenemos dos modelos, deben ser esencialmente el mismo, en el sentido de que dados dos modelos $(N_A, 0_A, s_A)$ y $(N_B, 0_B, s_B)$ de los axiomas de Peano, la aplicación $f\colon N_A \to N_B$ definida mediante

$$f(0_A) = 0_B,$$
$$f(s_A(n)) = s_B(n).$$

es una biyección. Ya tenemos que, esencialmente, sólo puede haber un modelo de los axiomas de Peano. ¡Pero aún no tenemos ninguno!



Friedrich Ludwig
Gottlob Frege
(1848–1925)



Bertrand Arthur
William Russell
(1872–1970)



János von Neumann
(1903–1957)

Los lógicos Gottlob Frege y Bertrand Russell construyeron modelos de los axiomas de Peano basándose en la idea intuitiva de que un número natural es el cardinal de un conjunto. O, dicho de otra forma, cada número natural es la «esencia» que comparten todos los conjuntos biyectivos (y finitos) cuyo cardinal es ese número natural. Entonces, el número natural en sí se definiría como la clase de equivalencia de tales conjuntos; además, y en cierto modo, 0 se correspondería con el conjunto vacío (realmente, con su clase de equivalencia), y s aplicado a un conjunto consistiría en añadirle un elemento que no estuviese ya en el conjunto (realmente, esto se hace para cada conjunto de la clase de equivalencia). Pero esta idea tiene dificultades lógicas serias, pues, de manera similar a como las leyes de la axiomática de Zermelo-Fraenkel en las que se basa la matemática actual hacen que —dicho de manera informal— «el conjunto de todos los conjuntos no es un conjunto», las clases de equivalencia de conjuntos biyectivos tampoco forman un conjunto.

En la actualidad, el modelo que se usa habitualmente es otro, debido esta vez al genial matemático húngaro János von Neumann (luego nacionalizado estadounidense con el nombre de John), quien no sólo hizo importantes contribuciones en multitud de campos de la matemática, sino que posiblemente sea aún más conocido por su destacado papel en los inicios de la computación —los ordenadores actuales usan la denominada arquitectura von Neumann— y en la fabricación de la primera bomba atómica.

El modelo ideado por von Neumann comienza con la definición de 0 como el conjunto vacío, $\varnothing$, y además se define un operador s actuando sobre los conjuntos mediante

$$s(A) = A \cup \{A\}.$$

Ahora, se define el conjunto de los números naturales, $\mathbb{N}_0$, como la intersección de todos los conjuntos cerrados bajo la acción de s (es decir, de todos los conjuntos C tales que $s(C) \subset C$) que contienen al conjunto vacío. Con esto, cada número natural (entendido como un conjunto) es el conjunto de los números naturales menores que él:

$$\begin{aligned} 0 &= \varnothing, \\ 1 &= s(0) = \varnothing \cup \{\varnothing\} = \{\varnothing\} = \{0\}, \\ 2 &= s(1) = \{\varnothing\} \cup \{\{\varnothing\}\} = \{\varnothing, \{\varnothing\}\} = \{0, 1\}, \\ 3 &= s(2) = \{\varnothing, \{\varnothing\}\} \cup \{\{\varnothing, \{\varnothing\}\}\} = \{\varnothing, \{\varnothing\}, \{\varnothing, \{\varnothing\}\}\} = \{0, 1, 2\}, \end{aligned}$$

y así sucesivamente. El conjunto $\mathbb{N}_0$, junto con 0 y la función «siguiente» $s\colon \mathbb{N}_0 \to \mathbb{N}_0$, satisfacen los axiomas de Peano.

A.4. Los números enteros

Así como los naturales son, digamos, evidentes, a los matemáticos les costó mucho tratar a los números negativos en pie de igualdad con los positivos. Históricamente, el uso de los números negativos es muy posterior al uso de fracciones e incluso de los irracionales positivos.

Como ejemplo de las complicaciones provenientes del no uso de números negativos, tanto para el griego Diofanto como para los algebristas europeos del Renacimiento, una ecuación de tipo $x^2+bx+c=0$ (en la que para nosotros los parámetros b y c pueden ser tanto positivos como negativos, y el método de resolución es el mismo) no era única, sino que debía ser analizada en cuatro casos distintos

$$x^2 + bx + c = 0, \quad x^2 + bx = c, \quad x^2 + c = bx, \quad x^2 = bx + c,$$

con b y c siempre positivos (incluso más casos si admitimos la posibilidad de que b o c valgan 0), y cada uno de los casos tenía su correspondiente método de resolución. Fue el matemático holandés Simon Stevin, a finales del siglo XVI, el primero que reconoció la validez de los números negativos al aceptarlos como resultado de los problemas con que trabajaba. Además, reconoció la igualdad entre la sustracción de un número positivo y la adición de un número negativo (es decir, $a - b = a + (-b)$, con $a, b > 0$). Por todo ello —y de manera similar a como podemos pensar en Brahmagupta como padre del cero— Stevin es considerado el padre de los números negativos (de hecho, Stevin hizo muchas más contribuciones al mundo de los números, como comentaremos cuando abordemos los números reales).



Brahmagupta
(598–670)



Simon Stevin
(1548–1620)

Pero nosotros sí que tenemos ya la idea intuitiva de lo que son los números negativos, y no tenemos ningún problema ideológico o práctico para usarlos, así que no vamos a preocuparnos de ello, sino que abandonamos la senda histórica y volvemos a lo que aquí queremos, que es dar una construcción rigurosa.

El caso es que en $\mathbb{N}_0$ podemos sumar números, pero no siempre podemos restar. Por eso surge la necesidad de crear unos nuevos números con los que poder restar cualquier pareja de números naturales; estos nuevos números van a ser los *números enteros*, y el conjunto de tales números lo denotaremos con $\mathbb{Z}$.

Antes de proseguir, vamos a analizar la situación con un ejemplo. Si tenemos 7 unidades de un producto y entregamos 4 nos quedan 3; y eso se expresa $7-4=3$, que significa $7=3+4$. Por el contrario, si tenemos 7 unidades y tenemos que entregar 12, no podemos hacerlo; aunque entreguemos todo lo que tenemos seguimos debiendo 5. Todos sabemos que esto lo vamos a acabar expresando con $7-12=-5$; pero, ¿qué es eso de -5?, ¿qué número es ése?

El resultado final, 3, consecuencia de tener 7 y entregar 4, sería el mismo con muchas otras parejas de naturales a y b siempre que $a=b+3$; así, el número 3 se asocia a todas las parejas a y b tales que $a=b+3$. Del mismo modo, el aún inexistente -5 correspondiente a tener 7 unidades y querer entregar 12 sería el mismo que con muchas otras parejas de naturales a y b siempre que $a+5=b$. Así, la idea que surge es intentar definir los números enteros por medio de parejas de naturales. Y, como muchas parejas representarán el mismo número entero, hay que identificarlas por medio de una relación de equivalencia. Que dos parejas (a,b) y (c,d) representasen el mismo entero z, querría decir, si tuviese ya sentido, que $a-b=z$ y $c-d=z$. Es decir, (a,b) y (c,d) representarán el mismo entero si $a-b=c-d$. ¿Cómo expresar esa aún inexistente resta con operaciones que sí tengan ya sentido? Es fácil: basta escribirla en su forma equivalente $a+d=b+c$. Eso es lo que tendremos que usar al dar la definición.

Una vez vistas las ideas, vamos ya con el desarrollo formal. Tomamos el conjunto

$$\mathbb{N}_0 \times \mathbb{N}_0 = \{(a, b) : a, b \in \mathbb{N}_0\}$$

y en él definimos la relación $\sim$ dada por

$$(a, b) \sim (c, d) \;\Leftrightarrow\; a + d = b + c.$$

Es totalmente elemental comprobar que esta relación es de equivalencia, y eso nos permite definir los números enteros como el conjunto cociente

$$\mathbb{Z} = \mathbb{N}_0 \times \mathbb{N}_0/\sim.$$

Tenemos ahora que definir la suma y el producto en $\mathbb{Z}$. Y, claro, hay que hacerlo de manera inteligente para conseguir lo que se pretende al sumar y multiplicar enteros. Dado que (a, b) es la manera que tenemos en $\mathbb{Z}$ de indicar la resta $a - b$ y, del mismo modo, (c, d) representa $c-d$, no hay más que pensar en $(a-b)+(c-d) = (a + c) - (b + d)$ para dar una definición formal adecuada de la suma:

$$(a, b) + (c, d) = (a + c, b + d).$$

De manera similar, para definir el producto pensemos en que lo que queremos conseguir es $(a - b) \cdot (c - d) = (a \cdot c + b \cdot d) - (a \cdot d + b \cdot c)$; con eso en perspectiva, la definición formal del producto tiene que ser así:

$$(a, b) \cdot (c, d) = (a \cdot c + b \cdot d, a \cdot d + b \cdot c).$$

Por supuesto, para que estas definiciones dadas en $\mathbb{N}_0 \times \mathbb{N}_0$ tengan sentido en $\mathbb{Z} = \mathbb{N}_0 \times \mathbb{N}_0/\sim$, ahora hay que probar que tanto la suma como el producto que acabamos de definir son compatibles con la relación de equivalencia; es decir, que, si tenemos $(a_1, b_1) \sim (a_2, b_2)$ y $(c_1, d_1) \sim (c_2, d_2)$, se cumple

$$\begin{aligned}(a_1, b_1) + (c_1, d_1) &\sim (a_2, b_2) + (c_2, d_2),\\ (a_1, b_1) \cdot (c_1, d_1) &\sim (a_2, b_2) \cdot (c_2, d_2)\end{aligned}$$

(sin estos requisitos, la suma y el producto no estarían bien definidas en $\mathbb{N}_0 \times \mathbb{N}_0/\sim$). Son sólo comprobaciones rutinarias y elementales, así que nos las ahorraremos.

También hay que definir la relación de orden en $\mathbb{Z}$. Para definir cuándo $(a, b) \leq (c, d)$, pensemos una vez más en (a, b) como $a - b$ y en (c, d) como $c - d$. Entonces, basta fijarse en que $a - b \leq c - d$ equivale a $a + d \leq b + c$ para darse cuenta de que la definición que buscamos tiene que ser

$$(a, b) \leq (c, d) \;\Leftrightarrow\; a + d \leq b + c.$$

Como en la suma y el producto, hay que comprobar la compatibilidad de esta definición con la relación de equivalencia; es decir, que, si tenemos $(a_1, b_1) \sim (a_2, b_2)$ y $(c_1, d_1) \sim (c_2, d_2)$, se cumple

$$(a_1, b_1) \leq (c_1, d_1) \Leftrightarrow (a_2, b_2) \leq (c_2, d_2).$$

De nuevo, ver que eso es cierto no tiene ninguna dificultad.

En cierto modo, $\mathbb{N}_0$ «se incluye» en $\mathbb{Z}$ asociando a cada $a \in \mathbb{N}_0$ la clase de equivalencia de $(a, 0)$. El número natural 0 en $\mathbb{Z}$ es $(0, 0)$ (en realidad, toda su clase de equivalencia dada por $\{(a, a) : a \in \mathbb{N}_0\}$), que es el elemento neutro de la suma en $\mathbb{Z}$. Y el natural 1 es $(1, 0)$ (y su clase de equivalencia), que es el elemento neutro del producto.

Por supuesto, demostrar las propiedades asociativa y conmutativa tanto de la suma como del producto carece también de dificultad. Y lo mismo ocurre con la propiedad distributiva del producto respecto a la suma.

A diferencia de lo que sucede en $\mathbb{N}_0$, en $\mathbb{Z}$, dado un número, siempre se puede encontrar otro que sumado al primero dé cero (el elemento neutro de la suma). Si tenemos (a, b), basta tomar (b, a) y se cumple

$$(a, b) + (b, a) = (a + b, a + b) \sim (0, 0).$$

Este número (b, a) decimos que es el opuesto de (a, b), y lo denotamos $-(a, b)$. Esto permite definir la resta

$$(a, b) - (c, d) = (a, b) + (-(c, d)) = (a, b) + (d, c) = (a + d, b + c).$$

O, de manera equivalente,

$$(a, b) - (c, d) = (r, s) \Leftrightarrow (a, b) = (r, s) + (c, d),$$

que siempre se puede conseguir. Y con esto hemos logrado nuestro objetivo de extender $\mathbb{N}_0$ a un conjunto de números $\mathbb{Z}$ en el que la resta sí fuese siempre posible.

Ya es hora de desprendernos de esta notación para $\mathbb{Z}$ en términos de pares de números naturales y dar la habitual. Para cada número entero, siempre podemos tomar un representante de su clase de equivalencia de la forma $(a, 0)$ o $(0, a)$ donde $a \in \mathbb{N}_0$. En lugar del $(a, 0)$ de $\mathbb{Z}$ (en realidad, en lugar de toda su clase de equivalencia) usamos de nuevo la notación a; si $a \neq 0$, estos números se denominan positivos. Y, en lugar del $(0, a)$ de $\mathbb{Z}$ (de toda su clase de equivalencia), usamos ahora la notación $-a$; si $a \neq 0$, estos son los denominados números negativos. Por supuesto, el número $(0, 0)$ de $\mathbb{Z}$ se vuelve a denotar como 0 (y es el mismo número que -0). Además, podemos considerar que $\mathbb{N}_0$ es, en cierto modo, una especie de

subconjunto de $\mathbb{Z}$ sin más que identificar $a \in \mathbb{N}_0$ con $(a, 0)$; obsérvese que, al hacer esto, las operaciones suma y producto de $\mathbb{Z}$ extienden las de $\mathbb{N}_0$, y lo mismo ocurre con el orden (con rigor, lo que tenemos es que la aplicación que lleva $\mathbb{N}_0$ a $\mathbb{Z}$ asignando $a \mapsto (a, 0)$ es inyectiva y compatible con las operaciones y el orden, lo cual se suele expresar diciendo que $\mathbb{N}_0$ se incrusta en $\mathbb{Z}$).

Una vez construidos los enteros, podríamos ahora dedicarnos a probar todo tipo de propiedades, pero no lo vamos a hacer. Tan solo queremos mencionar las que son diferentes de las de $\mathbb{N}_0$. En primer lugar, y al contrario que los naturales, los enteros ya no son un conjunto bien ordenado (por ejemplo, el conjunto de los enteros negativos no tiene mínimo); en realidad lo que ocurre es algo muy parecido: cualquier subconjunto no vacío de $\mathbb{Z}$ acotado inferiormente tiene mínimo (y, multiplicando todo por -1, si es acotado superiormente, tiene máximo). Sí que sigue siendo cierto que el orden en $\mathbb{Z}$ es un orden total. Otras novedades surgen de la existencia ahora de dos tipos de números, positivos y negativos. Al multiplicar enteros, se cumple la denominada «regla de los signos»:

$$\begin{aligned} \text{positivo} \cdot \text{positivo} &= \text{positivo}, & \text{positivo} \cdot \text{negativo} &= \text{negativo}, \\ \text{negativo} \cdot \text{positivo} &= \text{negativo}, & \text{negativo} \cdot \text{negativo} &= \text{positivo}. \end{aligned}$$

En lo que se refiere al orden, hay que destacar que cualquier número negativo es menor que cualquier positivo. Además, que el orden sea estable con las operaciones hay que entenderlo de manera algo diferente:

- $a \leq b \Rightarrow a + c \leq b + c$,
- $a \leq b$, $c \geq 0 \Rightarrow a \cdot c \leq b \cdot c$.

En cambio, al multiplicar por un número negativo, las desigualdades se invierten:

- $a \leq b$, $c < 0 \Rightarrow a \cdot c \geq b \cdot c$.

Hay también en los enteros una función importante que no tiene ningún interés en los números naturales. Se trata de la función denominada *valor absoluto* (o *módulo*), que se define mediante

$$|a| = \begin{cases} a, & \text{si } a \geq 0, \\ -a, & \text{si } a < 0. \end{cases}$$

Sin más que comprobar los distintos casos que pueden surgir, es fácil ver que se cumple $|a + b| \leq |a| + |b|$, propiedad que —por sus implicaciones geométricas en otros contextos— se denomina *desigualdad triangular*.

Finalicemos ya nuestro periplo por los enteros con una novedad en cuanto a la estructura algebraica: la existencia de elemento opuesto le confiere a $\mathbb{Z}$, con la suma y el producto, estructura de anillo.

A.5. Los números racionales

La extensión de los números enteros, $\mathbb{Z}$, a los *números racionales*, $\mathbb{Q}$, tiene un claro paralelismo con la extensión de $\mathbb{N}_0$ a $\mathbb{Z}$ que hemos visto en la sección anterior. Antes, no siempre podíamos restar en $\mathbb{N}_0$, y nos inventamos un nuevo tipo de números para conseguirlo. Ahora, no siempre podemos dividir en $\mathbb{Z}$, y nos vamos a inventar un nuevo tipo de números para conseguirlo.

Para definir $\mathbb{Z}$ tomábamos pares de números naturales, y aplicábamos una relación de equivalencia. La definición de $\mathbb{Q}$ sigue los mismos pasos, pero de manera aún más clara: los números racionales van a ser pares de números enteros, que se corresponden con el numerador y el denominador de cada fracción; además, hay que tomar clases de equivalencia para identificar las fracciones que representan el mismo número. Vamos con ello.

Como el denominador de una fracción no puede ser nulo, en lugar de usar $\mathbb{Z}\times\mathbb{Z}$ tomamos el conjunto

$$\mathbb{Z}\times(\mathbb{Z}\setminus\{0\}) = \{(a,b) : a,b\in\mathbb{Z},\ b\neq 0\},$$

y en él definimos la relación $\sim$ dada por

$$(a,b)\sim(c,d) \ \Leftrightarrow\ ad = bc,$$

que —según se demuestra sin dificultad— es de equivalencia. Obsérvese la idea subyacente: (a,b) y (c,d) acabarán siendo, respectivamente, los racionales a/b y c/d; aún no podemos hablar de la igualdad $a/b = c/d$ (¡pues esas fracciones aún no existen!), pero si eso tuviese sentido equivaldría a decir que $ad = bc$, así que eso es lo que estamos utilizando para definir la relación de equivalencia. Ahora, definamos $\mathbb{Q}$ como el conjunto cociente

$$\mathbb{Q} = \mathbb{Z}\times(\mathbb{Z}\setminus\{0\})/\!\sim.$$

La suma y el producto en $\mathbb{Q}$ se definen de la manera obvia a partir de las reglas de operar con fracciones que conocemos desde el colegio. Queremos que —cuando las fracciones tengan sentido— sea

$$\frac{a}{b}+\frac{c}{d} = \frac{ad+bc}{bd}, \qquad \frac{a}{b}\cdot\frac{c}{d} = \frac{ac}{bd},$$

y para lograr eso nuestra definición formal de suma y producto tiene que ser

$$(a,b)+(c,d) = (ad+bc,bd), \qquad (a,b)\cdot(c,d) = (ac,bd).$$

Por supuesto, para que estas definiciones estén bien dadas en $\mathbb{Q} = \mathbb{Z} \times (\mathbb{Z} \setminus \{0\})/\sim$ hay que probar que son compatibles con la relación de equivalencia, lo cual se reduce a simples comprobaciones rutinarias.

Los enteros $\mathbb{Z}$ se incrustan en $\mathbb{Q}$ identificando $a \in \mathbb{Z}$ con $(a, 1)$ en $\mathbb{Q}$ y, como no podía ser de otra manera, la suma y el producto en $\mathbb{Q}$ que acabamos de definir son extensiones de las correspondientes operaciones en $\mathbb{Z}$. Además, dado que $(-a, -b) \sim (a, b)$, y como por definición debe ser $b \neq 0$, siempre que tomemos (a, b) en $\mathbb{Q}$ podemos asumir que $b > 0$ (es decir, que los denominadores de las fracciones los podemos suponer positivos), lo cual simplifica algunas cosas.

Definamos ahora la relación de orden en $\mathbb{Q}$. Para ello tenemos que expresar la aún inexistente desigualdad $a/b \leq c/d$ en términos de desigualdades con enteros; pero, como podemos suponer que $b > 0$ y $d > 0$, eso es equivalente a $ad \leq cb$ (obsérvese que esta desigualdad sale de la que queríamos obtener multiplicando por bd, y que haber supuesto que los denominadores son positivos hace que la desigualdad no cambie de sentido, lo cual nos permite no tener que discutir diversos casos). Así pues, definimos

$$(a, b) \leq (c, d) \;\Leftrightarrow\; ad \leq cb \qquad (\text{con } b, d > 0).$$

De nuevo se comprueba sin dificultad que esta relación es de orden, que es compatible con la relación de equivalencia (de hecho, ya hemos usado la relación de equivalencia al definir $\leq$ sólo para denominadores positivos), y que extiende la de $\mathbb{Z}$. Como en $\mathbb{Z}$, la relación de orden en $\mathbb{Q}$ es total, y también tenemos números positivos y negativos.

Una vez construido todo el entramado, podemos desprendernos de la notación auxiliar (a, b) que hemos utilizado para definir los números racionales y emplear ya, a partir de ahora, las habituales $\frac{a}{b}$ o a/b, donde a se denomina numerador y b denominador (como hemos dicho, siempre podemos suponer que el denominador es positivo). Además, en lugar de $a/1$ usaremos simplemente a. Y una última advertencia en cuanto a la notación: en vez de servirnos de la relación de equivalencia $(a, b) \sim (c, d)$ usaremos ya $a/b = c/d$ sin tapujos.

Nota A.5. Todos sabemos que, en cada clase de fracciones equivalentes de $\mathbb{Q}$, existe una fracción especialmente señalada que se suele tomar, a menudo, como representante. Tal fracción es la denominada irreducible, en la que el máximo común divisor del numerador y el denominador vale 1. Pero, con las construcciones y conceptos que hemos ido introduciendo en este apéndice, no podemos hablar de ella, pues —aunque todo lo que se requiere es elemental— tendríamos que haber desarrollado algo más la aritmética en $\mathbb{N}_0$ de lo que hemos hecho en la sección A.3.

Por supuesto, con lo que aparece en las primeras páginas del tema 1 de este libro, la existencia (e incluso la unicidad) de fracciones irreducibles es ya inmediata. ◁

Huelga decir que las propiedades de la suma y el producto en $\mathbb{Z}$ (asociatividad, conmutatividad, elementos neutros, propiedad distributiva...) se extienden a $\mathbb{Q}$; y que, además, el orden es estable con las operaciones en el mismo sentido que en $\mathbb{Z}$. Pero —y esto es lo que pretendíamos desde el principio— en $\mathbb{Q}$ hay una novedad sobre lo que ocurría en $\mathbb{Z}$: se puede dividir.

Conceptualmente hablando, comencemos notando que cualquier $a/b \in \mathbb{Q}$ no nulo tiene un inverso respecto al producto. En efecto, dado que

$$\frac{a}{b} \cdot \frac{b}{a} = \frac{ab}{ba} = \frac{1}{1} = 1,$$

el inverso de a/b es $(a/b)^{-1} = b/a$. Con esto, ya vamos a poder dividir por un racional no nulo. Dividir a/b entre c/d consiste en encontrar un número racional tal que a/b multiplicado por ese racional dé c/d; y esto es inmediato, pues basta multiplicar a/b por el inverso de c/d, así:

$$\frac{a}{b} \Big/ \frac{c}{d} = \frac{a}{b} \cdot \left(\frac{c}{d}\right)^{-1} = \frac{a}{b} \cdot \frac{d}{c} = \frac{ad}{bc}.$$

La existencia de inverso en el producto (y por tanto de división) hace que la estructura algebraica de $\mathbb{Z}$ (recordemos que era anillo) se quede ahora corta, y $\mathbb{Q}$ con la suma y el producto pasa a ser un cuerpo (véase la sección A.2). Además, la estructura de orden es estable con las operaciones, es decir, para $x, y, t \in \mathbb{Q}$ se cumple

- $x \leq y \Rightarrow x + t \leq y + t$;
- $x \leq y$, $t \geq 0 \Rightarrow tx \leq ty$.

Así pues, hemos probado que $\mathbb{Q}$ es lo que se denomina un cuerpo ordenado.[3]

Acabemos esta sección dedicada a la construcción de los números racionales demostrando una importante propiedad de $\mathbb{Q}$ (y que no es cierta en general para cualquier cuerpo ordenado). Se trata de la denominada *propiedad arquimediana*:

Proposición A.6.

Sea $x \in \mathbb{Q}$, $x > 0$. Entonces, para cualquier $y \in \mathbb{Q}$ existe $n \in \mathbb{N}_0$ tal que $nx > y$.

[3] Puesto que la relación de orden es total, se podría decir que es un «cuerpo totalmente ordenado» pero, realmente, y tal como veremos al comienzo del apartado A.6.2, cuando se dan las definiciones con más detalle, el concepto «cuerpo ordenado» incluye el requisito de que, salvo el 0, los números son positivos o negativos; en el caso de $\mathbb{Q}$, eso se corresponde con que el orden sea total.

Demostración. Si $y \leq 0$ el resultado es trivial, pues basta tomar $n = 1$. Asumamos entonces que $y > 0$. Queremos probar que existe $n \in \mathbb{N}_0$ que cumple $n > y/x$. El cociente de números racionales y/x será un número racional a/b con a y b enteros positivos. Así pues, $n > y/x = a/b$ equivale a decir $nb > a$. Y eso se consigue sin más que tomar $n = a + 1$, ya que

$$(a+1)b = ab + b \geq a + b \geq a + 1 > a,$$

tal como pretendíamos. □

A.6. Los números reales

Los números racionales nos ofrecen ya muchas posibilidades. Pero en seguida puede uno darse cuenta de que no son suficientes para alguna de las tareas que uno desearía encomendar a los números: medir longitudes. De hecho, fueron los pitagóricos en el siglo IV a.C. los que observaron esa importante carencia de los números racionales, lo cual, dicho sea de paso, supuso una enorme crisis en la concepción de las matemáticas que ellos tenían. La idea de los pitagóricos era que todo se tenía que reducir a proporciones numéricas; y, por supuesto, tales proporciones equivalían a nuestras fracciones. Posiblemente, su profunda decepción fue fruto de su descubrimiento de que la diagonal del cuadrado era inconmensurable con el lado, es decir, que la diagonal no se puede expresar como un número racional de veces la longitud del lado. Su razonamiento pudo ser —aproximadamente— como sigue.

En el cuadrado unidad, el teorema de Pitágoras nos dice que su diagonal d cumple $d^2 = 1^2 + 1^2 = 2$. Tal como se observa en la figura A.1, esa longitud d (que con la notación actual sería $d = \sqrt{2}$) se corresponde con un punto de la recta horizontal, que da lugar en la recta a un segmento de esa misma longitud. Pero —y esto ya lo sabían los pitagóricos— no existe ningún número racional que su cuadrado valga 2. En efecto, si hubiese un número racional a/b (con a y b enteros positivos) tal que $(a/b)^2 = 2$, tendríamos $a^2 = 2b^2$, luego a debería ser un entero par. Ahora, tomando $a = 2a_1$, obtenemos $4a_1^2 = 2b^2$, de donde $b^2 = 2a_1^2$, y b también debería ser par; digamos que $b = 2b_1$. En estas condiciones, $a/b = a_1/b_1$ y $(a_1/b_1)^2 = 2$, luego podemos volver a repetir el proceso, con lo cual tendremos $a_1 = 2a_2$ y $b_1 = 2b_2$, y así indefinidamente. Pero es imposible que los enteros a y b se puedan dividir por 2 tantas veces como queramos de forma que sigan saliendo números enteros, así que nuestra suposición inicial $a^2/b^2 = 2$ no puede ser cierta.

Así pues, está claro que, para poder medir cualquier longitud, tiene que haber otros números además de los racionales (y tales números se denominarían irracionales). Si pensamos en una recta en la que está marcado el segmento unidad (con

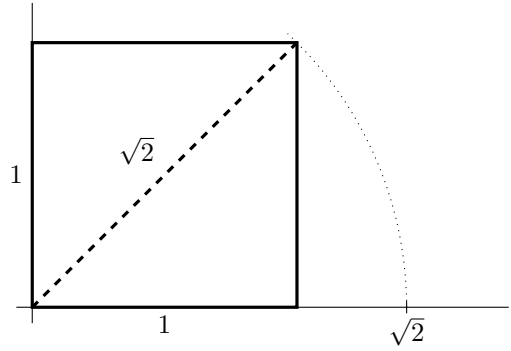
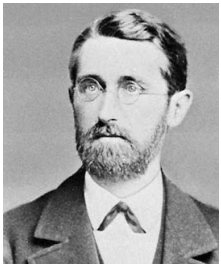


Figura A.1: La diagonal de un cuadrado es un número de la recta real.



Pitágoras de Samos
(≈ 569 a.C. – ≈ 475 a.C.)



Richard Dedekind
(1831–1916)

0 y 1 en sus extremos), y cada punto de la recta queremos que sea un «número», aparentemente los racionales rellenan esa recta bastante bien, pero aún hay muchos puntos que no se corresponden con un número racional; ¡hay que extender los números a un conjunto más amplio que el de los racionales! Estos números que deberían corresponderse con los puntos de la recta es lo que queremos que dé lugar a los *números reales* (y la recta es la denominada *recta real*).

En realidad, ya los mismos matemáticos griegos supieron zafarse de la crisis pitagórica. A este respecto, merece la pena destacar a Eudoxo (siglo IV a.C.), cuyo trabajo sobre la teoría de la proporción muestra una amplia comprensión de los números y permite el tratamiento de las cantidades continuas, no únicamente de los números enteros o números racionales. También resulta pertinente mencionar a Stevin quien, además de aceptar los números negativos, llevó a cabo, a finales

del siglo XVI, una enérgica defensa de que otros números como raíces cuadradas, irracionales, etc., debían ser tratados como números ordinarios, sin hacer distinción de su naturaleza (sin embargo, no aceptó los números imaginarios que por entonces estaban apareciendo, lo cual frenó su desarrollo). Además, y aunque no fue Stevin el inventor de los decimales (que habían manejado mucho antes que él los chinos y los árabes), fue él quien introdujo su uso en matemáticas, dando para ellos una notación ya muy parecida a la nuestra.

Pero, volviendo al presente, ¿cómo podemos extender los racionales? Tenemos la idea intuitiva de la recta real, pero se trata de dar una definición rigurosa basada en conceptos matemáticos previos.

Si pensamos en que cada número racional tiene una desarrollo decimal, y que estos desarrollos de los racionales no pueden ser cualesquiera (siempre son finitos o periódicos), ingenuamente se nos podría ocurrir pensar que un número real sería lo que se obtiene a partir de una expresión decimal arbitraria, como

$$d_0 + 0.d_1d_2d_3d_4\ldots = d_0 + \sum_{j=1}^{\infty} \frac{d_j}{10^j}, \quad d_0 \in \mathbb{Z}, \;\; d_j \in \{0, 1, 2, \ldots, 9\} \text{ para } j \geq 1.$$

Pero, ese número, ¿qué es? Porque para hablar de él necesitaríamos haber demostrado que la serie converge a algo. Quizás podíamos haber optado por definir que los números reales son esas expresiones. Pero tal camino tiene muchas dificultades. Por ejemplo, ¿cómo definir la suma y el producto de expresiones decimales infinitas?

Así que esa idea, aunque a la postre sí se puede usar para representar los números reales, no es, como método riguroso de definirlos, tan directa como parecía, y en la práctica son pocos los matemáticos que tienen este método como su preferido. Hay tres procedimientos equivalentes que se usan habitualmente para definir los números reales a partir de los racionales, y los tres tienen su origen en el siglo XIX, que es cuando el análisis matemático alcanzó el rigor que exigimos en la actualidad. Estos métodos son los siguientes:

- Cortaduras de Dedekind.
- Clases de equivalencia de sucesiones de Cauchy de racionales.
- Clases de equivalencia de pares de sucesiones monótonas convergentes.

En los apartados que siguen desarrollaremos con detalle el primer método, y comentaremos los otros dos.

Adelantando acontecimientos, lo que queremos es construir un cuerpo ordenado (como $\mathbb{Q}$), pero con una propiedad adicional: que cualquier conjunto acotado superiormente tenga supremo. De hecho, en algunos textos no se construyen los números reales, sino que se da de ellos una definición axiomática basada en propiedades que

se deben cumplir, y la propiedad que acabamos de citar se denomina *axioma del supremo*. Ese método incluso permite demostrar que, esencialmente (salvo isomorfismos, si empleamos una palabra más técnica), sólo puede existir un conjunto que cumple todos los axiomas que se exigen, y tal conjunto sería el de los números reales. Pero definir los reales simplemente mediante axiomas es engañoso, ya que, sin una construcción formal a partir de conceptos previos, nada garantiza que exista un modelo de esa definición axiomática, que por tanto podría haber sido un mero castillo en el aire.

A.6.1. Cortaduras de Dedekind

Uno de los «defectos» de $\mathbb{Q}$ es que un subconjunto de $\mathbb{Q}$ acotado superiormente puede no tener supremo. Por ejemplo, en el caso del aún inexistente $\sqrt{2}$ que citábamos antes, hay infinidad de racionales a/b tales que $(a/b)^2 < 2$, pero ninguno tal que $(a/b)^2 = 2$. La idea de las cortaduras de Dedekind es que los números reales sean esos supremos inexistentes. Pero, como tales supremos no existen, en realidad lo que vamos a hacer es que los reales sean, por definición, los subconjuntos acotados superiormente, que nos gustaría que tuviesen supremo. Como podemos pensar en muchos conjuntos distintos que deberían compartir supremo, vamos a elegir uno especial, que es lo que vamos a llamar una *cortadura*. Intuitivamente, y tal como propuso Richard Dedekind en 1858, si pensamos en la recta real, una cortadura se corresponde con cortar la recta real en dos quitando un punto y quedarse con los racionales de la parte izquierda. Entonces, $\mathbb{R}$ será el conjunto de todas estas cortaduras.

Con más concreción, llamamos *cortadura* a un conjunto α de racionales que satisface estas cuatro propiedades:

1. Si $x \in \alpha$ e $y \in \mathbb{Q}$ cumple $y < x$, entonces $y \in \alpha$.
2. $\alpha \neq \varnothing$.
3. $\alpha \neq \mathbb{Q}$.
4. Para cada $x \in \alpha$ existe $y \in \alpha$ que cumple $y > x$ (es decir, en α no hay máximo).

Por definición, un número real es, simplemente, una cortadura; y el conjunto de todos los números reales se denota con $\mathbb{R}$.

Por ejemplo, el nuevo número $\sqrt{2}$ (hasta ahora inexistente) estaría dado por la cortadura

$$\alpha = \{x \in \mathbb{Q} : x < 0 \text{ o } x^2 < 2\}$$

(por supuesto, es inmediato comprobar que se cumplen las cuatro propiedades exigidas).

Los números reales son, pues, conjuntos de racionales, y un número racional no es un número real. Pero podemos incrustar $\mathbb{Q}$ en $\mathbb{R}$ asociando a cada $q \in \mathbb{Q}$ la cortadura $\alpha_q = \{x \in \mathbb{Q} : x < q\}$. Es simplemente anecdótico que $q \notin \alpha_q$, pero eso es lo que siempre ocurre. Es una consecuencia del cuarto requisito de la definición de número real que, aunque puede resultar fastidioso, es indispensable para evitar ambigüedades. Así, por ejemplo, sin ese requisito el número 1 de $\mathbb{R}$ estaría definido tanto por $\{x \in \mathbb{Q} : x < 1\}$ como por $\{x \in \mathbb{Q} : x \leq 1\}$.

Una vez definido $\mathbb{R}$, hay que definir la suma, el producto y la relación de orden. Comencemos con la relación de orden, que es particularmente sencilla. Dadas dos números reales α y β (o sea, dos cortaduras), decimos que $\alpha \leq \beta$ si se cumple $\alpha \subseteq \beta$ (inclusión de conjuntos). Por descontado, $\alpha < \beta$ indica que $\alpha \leq \beta$ pero $\alpha \neq \beta$; y, con su significado obvio que no hace falta detallar, también se usan los símbolos $\geq$ y $>$. Nótese que la inclusión de conjuntos es siempre una relación de orden, luego, sin necesidad de comprobar las propiedades, podemos asegurar que efectivamente estamos definiendo una relación de orden en $\mathbb{R}$. En cambio, la inclusión de conjuntos no proporciona, en general, un orden total, así que ya no es automático asegurar que la relación de orden en $\mathbb{R}$ es total. Pero, dado que la relación de orden en $\mathbb{Q}$ es total, la primera propiedad de la definición de cortadura implica que la relación de orden en $\mathbb{R}$ también es total.

El 0 de $\mathbb{R}$ se define como $\{x \in \mathbb{Q} : x < 0\}$. Además del cero, hay otros dos tipos de números reales: los positivos, que cumplen $\alpha > 0$, y los negativos, que cumplen $\alpha < 0$ (la relación de orden es total, y eso implica que cualquier $\alpha \neq 0$ debe ser positivo o negativo).

El siguiente resultado es muy importante, pues —como comentábamos al principio de este apartado— es lo que diferencia a $\mathbb{R}$ de $\mathbb{Q}$: un conjunto acotado de números racionales no siempre tiene supremo en $\mathbb{Q}$ (por ejemplo, no lo tiene el conjunto $\{x \in \mathbb{Q} : x^2 < 2\}$), pero un conjunto acotado de números reales siempre tiene supremo en $\mathbb{R}$. Se expresa diciendo que $\mathbb{R}$ es *completo*.

Proposición A.7. Completitud de $\mathbb{R}$.

Si A es un conjunto de números reales, $A \neq \varnothing$, y A está acotado superiormente, entonces existe supremo de A.

Demostración. Sea

$$\beta = \{x \in \mathbb{Q} : x \in \alpha \text{ para algún } \alpha \text{ de } A\},$$

y vamos a probar que β es un número real (una cortadura) y que es el supremo de A. Demostrar que β es un número real requiere ver que cumple las cuatro propiedades que caracterizan la definición de cortadura. Vayamos con ello:

1. Sea $x \in \beta$ y $y \in \mathbb{Q}$, $y < x$. Como $x \in \alpha$ para alguna cortadura α, también $y \in \alpha$, luego $y \in \beta$.
2. Que $\beta \neq \varnothing$ es obvio, pues $A \neq \varnothing$ y por tanto existe $\alpha \in A$ que a su vez será $\alpha \neq \varnothing$.
3. Como A está acotado superiormente, existe un número real γ que cumple $\alpha \leq \gamma$ para todo $\alpha \in A$. Por ser $\gamma \neq \mathbb{Q}$, podemos tomar $x \in \mathbb{Q}$ tal que $x \notin \gamma$. Pero $\alpha \leq \gamma$ significa que $\alpha \subseteq \gamma$, luego $x \notin \alpha$, y esto es así para todos los α de A. En consecuencia, $x \notin \beta$ y $\beta \neq \mathbb{Q}$.
4. Supongamos que $x \in \beta$. Como $x \in \alpha$ para alguna cortadura α de A, existirá $y \in \alpha$ tal que $y > x$. Claramente, este y también cumple $y \in \beta$.

Una vez probado que β es un número real, veamos que es el supremo de A, es decir, que es la cota superior mínima:

- Sea $\alpha \in A$. Por construcción de β es obvio que $\alpha \subseteq \beta$, y precisamente ésa es la definición que hemos dado de $\alpha \leq \beta$. Así pues, β es cota superior de A.
- Supongamos que γ es una cota superior de A. ¿Se cumplirá forzosamente que $\beta \leq \gamma$? Para cada α en A, decir $\alpha \leq \gamma$ equivale a $\alpha \subseteq \gamma$; por construcción de β esto implica que $\beta \subseteq \gamma$, que a su vez equivale a afirmar $\beta \leq \gamma$, tal como queríamos comprobar. □

Tras el análisis de la relación de orden en $\mathbb{R}$, vamos ahora con las operaciones. Comencemos con la suma. Dadas α y β dos cortaduras, se define

$$\alpha + \beta = \{\, x \in \mathbb{Q} : x = r + s \text{ para algún } r \in \alpha \text{ y algún } s \in \beta \,\}.$$

Antes de continuar, hay que comprobar que $\alpha + \beta$ es una nueva cortadura, pues de lo contrario no tiene sentido hablar de él como número real. Y, para ello, hay que ver que el conjunto $\alpha + \beta$ cumple las cuatro propiedades que caracterizan las cortaduras. Eso es rutinario, pero, esta vez, lo vamos a comprobar con detalle:

1. Sea $y \in \mathbb{Q}$ que cumple $y < x$ para algún $x \in \alpha + \beta$. Entonces, $x = r + s$ para algún $r \in \alpha$ y algún $s \in \beta$, y por tanto $y < r + s$, de donde $y - r < s$. Esto implica que $y - r \in \beta$ (pues $s \in \beta$, que es una cortadura), y, en consecuencia, $y = r + (y - r)$ está en $\alpha + \beta$.
2. Es evidente que $\alpha + \beta \neq \varnothing$, ya que $\alpha \neq \varnothing$ y $\beta \neq \varnothing$.
3. Al ser $\alpha \neq \mathbb{Q}$ y $\beta \neq \mathbb{Q}$, existirán $r, s \in \mathbb{Q}$ tales que $r \notin \alpha$ y $s \notin \beta$. Por la primera propiedad de las cortaduras, cualquier $x \in \alpha$ deberá ser $x < r$, y cualquier $y \in \beta$ deberá ser $y < s$. Así pues, $x + y < r + s$ para cualesquiera $x \in \alpha$ e $y \in \beta$, con lo cual $x + y \notin \alpha + \beta$ y por tanto $\alpha + \beta \neq \mathbb{Q}$.

4. Sea $x \in \alpha + \beta$, con su descomposición $x = r + s$, $r \in \alpha$, $s \in \beta$. Por ser α y β cortaduras, existirán $r' \in \alpha$ y $s' \in \beta$ tales que $r' > r$ y $s' > s$. Entonces, $x < r' + s'$ y $r' + s' \in \alpha + \beta$, con lo cual $\alpha + \beta$ no tiene elemento máximo.

Con lo que hemos visto en la demostración de la proposición A.7 y la comprobación de que la suma de números reales es una nuevo número real, el lector ya se habrá hecho una idea del nivel de detalle al que hay que llegar para poder construir los números reales, definir las operaciones, demostrar sus propiedades, etc., con todo rigor. Pero, al mismo tiempo, también se habrá dado cuenta de que, una vez dadas las definiciones, la dificultad de comprobar los detalles es, en general, muy pequeña. Así pues, a partir de ahora avanzaremos con mayor rapidez, sin pararnos a efectuar todas las comprobaciones.

Comenzamos nuestro avance rápido con tres conocidas propiedades:

- Propiedad asociativa: si α, β y γ son números reales, entonces $\alpha + (\beta + \gamma) = (\alpha + \beta) + \gamma$.
- Propiedad conmutativa: si α y β son números reales, entonces $\alpha + \beta = \beta + \alpha$.
- El número real $0 = \{x \in \mathbb{Q} : x < 0\}$ es el elemento neutro de la suma, es decir, $\alpha + 0 = \alpha$ para cualquier número real α.

El opuesto de un número real α sería un $-\alpha$ tal que $\alpha + (-\alpha) = 0$. Pero, ¿cómo definir $-\alpha$? Tenemos que hacerlo de manera que $-\alpha$ sea una cortadura, y esto requiere un pequeño truco. En principio, uno podría intentar definir

$$-\alpha = \{x \in \mathbb{Q} : -x \notin \alpha\}$$

(obsérvese que $-x \notin \alpha$ significa, intuitivamente, que $-x > \alpha$, o sea, $x < -\alpha$). Pero esta definición no es adecuada, pues ese conjunto a menudo tiene máximo y, por tanto, no es una cortadura. Por ejemplo, si $\alpha = \{x \in \mathbb{Q} : x < 3\}$, esa definición daría lugar a

$$\{x \in \mathbb{Q} : -x \notin \alpha\} = \{x \in \mathbb{Q} : -x \geq 3\} = \{x \in \mathbb{Q} : x \leq -3\},$$

que obviamente no es una cortadura pues falla el cuarto requisito de la definición de número real. Para sortear el problema conviene fijarse en que, aunque una cortadura α nunca tiene (por definición) un elemento máximo, su complementario

$$\mathbb{Q} \setminus \alpha = \{x \in \mathbb{Q} : x \notin \alpha\}$$

puede tener un elemento mínimo $x_0 \in \mathbb{Q}$, y en ese caso $\{x \in \mathbb{Q} : -x \notin \alpha\}$ tendrá a $-x_0$ como elemento máximo. Una vez identificado el problema, basta definir

$$-\alpha = \{\, x \in \mathbb{Q} : -x \notin \alpha \text{ y } -x \text{ no es el elemento mínimo de } \mathbb{Q} \setminus \alpha \,\}$$

y ya resulta sencillo comprobar que este conjunto sí es una cortadura.

Ahora hay que demostrar que, efectivamente,

$$\alpha + (-\alpha) = 0.$$

Esto se hace por doble inclusión, comprobando que $\alpha + (-\alpha) \subseteq 0$ y $0 \subseteq \alpha + (-\alpha)$. La primera es muy sencilla, y no merece la pena detenerse en ella. En cambio, probar la segunda es bastante más laborioso. Como paso previo, fijémonos en que, si tenemos una cortadura α y un racional positivo q, existirán $x \in \alpha$ e $y \notin \alpha$ tales que $y - x = q$ (intuitivamente basta tomar x lo suficientemente cerca del «final» de la cortadura α e $y = x + q$, y para comprobarlo con detalle se usa la propiedad arquimediana de los racionales, que hemos mostrado en la proposición A.6); además, podemos suponer que y no es el elemento mínimo de $\mathbb{Q} \setminus \alpha$ (pues si lo fuera se toma $x' \in \alpha$ que cumpla $x' > x$ y se usa x' en vez de x e $y + (x' - x)$ en vez de y). Con eso, veamos que, efectivamente, $0 \subseteq \alpha + (-\alpha)$. Sea $z \in 0$. Según el paso previo con $q = -z > 0$, existe $x \in \alpha$ e $y \notin \alpha$ tales que $y - x = -z$, y además y no es el elemento mínimo de $\mathbb{Q} \setminus \alpha$. Entonces, podemos escribir $z = x + (-y)$ con $x \in \alpha$ e $y \in -\alpha$, lo cual prueba que $z \in \alpha + (-\alpha)$, tal como queríamos comprobar.

Antes de abordar el producto de números reales, resulta conveniente comentar algunas propiedades de los elementos positivos de $\mathbb{R}$ (recordemos que $\alpha \in \mathbb{R}$ es positivo si cumple $\alpha > 0$ y que, interpretado como cortaduras, eso significa que $\{x \in \mathbb{Q} : x < 0\} \subsetneq \alpha$). Es fácil comprobar que, si α y β son positivos, también lo es $\alpha + \beta$. También lo es que, para cualquier número real α, siempre se cumple una (y sólo una) de las tres condiciones siguientes:

- $\alpha = 0$;
- α es positivo;
- $-\alpha$ es positivo.

En particular, y procediendo por descarte ante las tres alternativas $\alpha = 0$, $\alpha > 0$ o $\alpha < 0$ que tiene cualquier número real, estamos asegurando que decir que $-\alpha$ es positivo (es decir, $-\alpha > 0$) equivale a que α sea negativo ($\alpha < 0$).

Nótese además que, unos párrafos más arriba, definíamos $\beta > \alpha$ (en realidad, usamos $\alpha < \beta$) con el significado de que $\alpha \subseteq \beta$ pero $\alpha \neq \beta$. ¿Equivale esto a decir $\beta - \alpha > 0$? Que eso es cierto es una consecuencia inmediata de que la relación de orden es estable con la suma; es decir, que se verifica

$$\alpha, \beta, \gamma \in \mathbb{R},\ \alpha < \beta \ \Rightarrow\ \alpha + \gamma < \beta + \gamma.$$

Comprobar esta propiedad es sencillo, y no nos vamos a detener en ello para no alargarnos.

Definir el producto de números reales es más complicado que definir la suma. Comencemos definiéndolo primero para números reales positivos. Si α y β son dos números reales positivos, su producto es, por definición,

$$\alpha \cdot \beta = \left\{ z \in \mathbb{Q} : z \leq 0 \text{ o } z = x \cdot y \text{ para algún } x \in \alpha \text{ e } y \in \beta \text{ con } x, y > 0 \right\}.$$

Por supuesto, hay que comprobar que este conjunto $\alpha \cdot \beta$ es una cortadura, pero es de nuevo rutinario y ya no lo vamos a hacer. Una vez definido $\alpha \cdot \beta$ para $\alpha, \beta > 0$, es claro que también se cumple $\alpha \cdot \beta > 0$.

Ahora nos ayudaremos del valor absoluto de un número real, que se define como

$$|\alpha| = \begin{cases} \alpha, & \text{si } \alpha \geq 0, \\ -\alpha, & \text{si } \alpha < 0, \end{cases}$$

con lo cual $|\alpha|$ es positivo siempre que $\alpha \neq 0$.

Con eso, definimos el producto de dos números reales arbitrarios α y β como

$$\alpha \cdot \beta = \begin{cases} 0, & \text{si } \alpha = 0 \text{ o } \beta = 0, \\ |\alpha| \cdot |\beta|, & \text{si } \alpha \text{ y } \beta \text{ son ambos positivos o ambos negativos}, \\ -(|\alpha| \cdot |\beta|), & \text{si } \alpha \text{ y } \beta \text{ son uno positivo y otro negativo}. \end{cases}$$

Puesto que todos los casos se corresponden con operaciones previamente definidas, esta vez no hay que probar que $\alpha \cdot \beta$ es una cortadura. Pero sí que habría que demostrar que la relación de orden es estable con el producto, es decir, que

$$\alpha, \beta, \gamma \in \mathbb{R},\ \alpha < \beta,\ \gamma > 0 \ \Rightarrow\ \gamma \cdot \alpha < \gamma \cdot \beta$$

(inmediato usando que el producto de reales positivos es positivo) y, con el consabido detalle, las siguientes propiedades:

- Propiedad asociativa: si α, β y γ son números reales, entonces $\alpha \cdot (\beta \cdot \gamma) = (\alpha \cdot \beta) \cdot \gamma$.
- Propiedad conmutativa: si α y β son números reales, entonces $\alpha \cdot \beta = \beta \cdot \alpha$.
- El número real $1 = \{x \in \mathbb{Q} : x < 1\}$ es el elemento neutro del producto, es decir, $\alpha \cdot 1 = \alpha$ para cualquier número real α.

Ahora hay que definir el inverso de un número real $\alpha \neq 0$, que deberá ser otro número real, denotado α^{-1}, que cumpla

$$\alpha \cdot \alpha^{-1} = 1.$$

Se comienza haciéndolo para los α positivos, en cuyo caso se toma

$$\alpha^{-1} = \{x \in \mathbb{Q} : x \leq 0 \text{ o } x > 0 \text{ y } 1/x \notin \alpha \text{ pero } 1/x \text{ no es el mínimo de } \mathbb{Q} \setminus \alpha\}.$$

Como hasta ahora, primero hay que comprobar que α^{-1} es un número real (una cortadura), y luego que efectivamente se cumple $\alpha \cdot \alpha^{-1} = 1$. Se consigue sin más que seguir un proceso similar al de la definición de $-\alpha$ y la comprobación de que $\alpha + (-\alpha) = 0$. Además, es inmediato que α^{-1} es positivo, como α.

Para los α negativos, se define

$$\alpha^{-1} = -(|\alpha|^{-1}),$$

que ya tiene sentido por ser $|\alpha|$ positivo; además, α^{-1} es automáticamente un número real, pues ya sabemos que tanto el proceso de tomar inversos de reales positivos como el de tomar opuestos da lugar a números reales. Fijémonos ahora en que, para $\alpha < 0$,

$$|\alpha^{-1}| = |-(|\alpha|^{-1})| = -(-(|\alpha|^{-1})) = |\alpha|^{-1}.$$

Entonces, como tanto α como α^{-1} son negativos, la definición de producto de números reales nos da

$$\alpha \cdot \alpha^{-1} = |\alpha| \cdot |\alpha^{-1}| = |\alpha| \cdot |\alpha|^{-1} = 1,$$

luego efectivamente α^{-1} es el inverso de α también cuando $\alpha < 0$.

Para concluir con la construcción de los reales, la última propiedad que falta verificar es la propiedad distributiva del producto respecto a la suma, esto es, que para cualesquiera $\alpha, \beta, \gamma \in \mathbb{R}$, se cumple

$$\alpha \cdot (\beta + \gamma) = (\alpha \cdot \beta) + (\alpha \cdot \gamma).$$

La demostración es laboriosa pues hay que considerar muchos casos, según los números involucrados sean positivos, negativos o nulos. Pero, si uno se embarca en comprobar los detalles, se da cuenta de que unos casos se pueden reducir a otros, y el trabajo a realizar se alivia considerablemente (además de que, si algunos de los reales α, β o γ es 0, el resultado es trivial). En todo caso, el proceso a seguir carece de dificultad, así que no nos detendremos en él.

Con todo eso, $\mathbb{R}$ con la suma y el producto tiene estructura de cuerpo; además, la relación de orden es total y estable con las operaciones. Recordemos asimismo que, en $\mathbb{R}$, cualquier subconjunto acotado superiormente (y no vacío) tiene supremo, y eso lo indicábamos diciendo que $\mathbb{R}$ es completo. De manera resumida, esto se

expresa diciendo que $\mathbb{R}$ es un cuerpo ordenado y completo (de hecho, es el único, como veremos en el apartado siguiente).

Una vez construidos los números reales, ya no hace falta tener en mente que están definidos con rigor a partir de cortaduras, sino que podemos asociarlos a nuestra idea intuitiva de número, y pensar en ellos como los puntos de la recta real. Por supuesto, a los naturales, los enteros y los racionales los consideramos como tipos especiales de números reales.

Antes de acabar este apartado queremos ver algunas propiedades que muestran cómo se distribuyen en la recta real $\mathbb{R}$ sus subconjuntos $\mathbb{N}$, $\mathbb{Z}$ y $\mathbb{Q}$. Intuitivamente son elementales, pero —ya lo hemos dicho muchas veces— para dar demostraciones correctas no se puede usar la intuición, sino el rigor matemático a partir de lo previamente establecido.

Comencemos con la propiedad arquimediana para $\mathbb{R}$:

Proposición A.8.
Dados dos números reales α y β, con $\alpha > 0$, existe algún número natural n tal que $n\alpha > \beta$.

Demostración. Dados $\alpha, \beta \in \mathbb{R}$ con $\alpha > 0$, supongamos que $n\alpha \leq \beta$ para todo $n \in \mathbb{N}$. Vamos a ver que esto conduce a una contradicción. Según hemos supuesto, el conjunto $S = \{n\alpha : n \in \mathbb{N}\}$ estará acotado (por β). Al ser $S \neq \varnothing$, la proposición A.7 nos dice que existirá un $s \in \mathbb{R}$ que es el supremo de S (es decir, la cota superior mínima). Entonces, $n\alpha \leq s$ para todo $n \in \mathbb{N}$ y, de aquí, también $(n+1)\alpha \leq s$ para todo $n \in \mathbb{N}$ (ya que $n+1 \in \mathbb{N}$ cuando $n \in \mathbb{N}$). Pero esto equivale a decir que $n\alpha \leq s - \alpha$ para todo $n \in \mathbb{N}$, lo cual es absurdo pues s era la mínima cota superior (nótese que $s - \alpha < s$ por ser $\alpha > 0$). □

Consecuencia de la propiedad arquimediana son otras dos importantes propiedades en las que interviene $\mathbb{N}$:

Proposición A.9.
$\mathbb{N}$ no está acotado superiormente.

Demostración. Basta tomar $\alpha = 1$ en la proposición anterior. □

Proposición A.10.
Para todo número real $\varepsilon > 0$ existe $n \in \mathbb{N}$ que cumple $1/n < \varepsilon$.

Demostración. Si no existiera ningún $n \in \mathbb{N}$ con $1/n < \varepsilon$, entonces tendríamos $1/n \geq \varepsilon$ para todo $n \in \mathbb{N}$, que equivale a $n \leq 1/\varepsilon$ para todo $n \in \mathbb{N}$, lo cual es absurdo por la proposición anterior. □

A continuación, una importante propiedad que permite definir la parte entera de los números reales.

Proposición A.11.

Para cada número real α existe un número entero que es el mayor entero menor o igual que α. Este número (que es único) se denomina parte entera de α, se denota $\lfloor\alpha\rfloor$, y cumple $\lfloor\alpha\rfloor \leq \alpha < \lfloor\alpha\rfloor + 1$.

Demostración. Dado α, sea el conjunto $S = \{n \in \mathbb{Z} : n \leq \alpha\}$. Por la propiedad arquimediana de $\mathbb{R}$, siempre existe $m \in \mathbb{N}$ tal que $m > -\alpha$; ese m cumplirá $-m < \alpha \leq \alpha$, luego $-m \in S$ y $S \neq \varnothing$. Además, S está acotado superiormente (por α). La completitud de $\mathbb{R}$ (proposición A.7) nos dice que un conjunto no vacío y acotado superiormente tiene supremo, así que tomemos $s = \sup S$ (en principio, s es un número real). Como s es la mínima cota superior, $s - 1$ ya no es cota superior de S, luego existe $a \in S$ que cumple $s - 1 < a \leq s$. Vamos a ver que este a (que ya es un entero) es el número que buscamos. Es claro que a es cota superior de S pues, si hubiese algún $n \in S$ tal que $n > a$, tendríamos $a < n \leq s < a + 1$, de donde se deduciría que $0 < n - a < 1$, lo cual es imposible pues no existe ningún entero positivo menor que 1. Así pues, a es cota superior de S y está en S, es decir, $a = \max S$. Con esto, ya lo tenemos: como $a \in S$, $a \leq \alpha$; y, como a es máximo y $a < a + 1$, forzosamente $a + 1 \notin S$, es decir, $a + 1 > \alpha$.[4]

Finalmente, analicemos la unicidad de la parte entera. Si $a, a' \in \mathbb{Z}$ cumplen $a \leq \alpha < a + 1$ y $a' \leq \alpha < a' + 1$, tendríamos $a \leq \alpha < a' + 1$ y $a' \leq \alpha < a + 1$; de aquí, $a < a' + 1$ y $a' < a + 1$, lo cual implica $a \leq a'$ y $a' \leq a$; es decir, $a = a'$. □

Una fácil consecuencia es ésta:

Proposición A.12.

Si α y β son dos reales que cumplen $\beta - \alpha > 1$, entonces existe $k \in \mathbb{Z}$ que cumple $\alpha < k < \beta$.

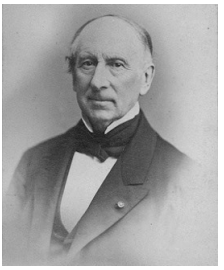
Demostración. Sea a el mayor entero que satisface $a \leq \alpha$ (es decir, $a = \lfloor\alpha\rfloor$), para el cual se tiene $a \leq \alpha < a+1$. Entonces, $\alpha < a+1 \leq \alpha+1 < \beta$, luego $k = a+1$ cumplirá $\alpha < k < \beta$. □

Por fin, un resultado que muestra que, en la recta real, los números racionales y los que no lo son están, por decirlo de alguna manera, totalmente entremezclados. Ya no vamos a dar todos los detalles de la demostración, sino simples indicaciones.

[4] Posiblemente el lector habrá intuido que el número real $s = \sup S$ es, en realidad, un entero de S, y por tanto $s = \max S$, lo cual simplificaría la demostración pues no haría falta distinguir entre s y a; esto es inmediato si aplicamos que cualquier conjunto de enteros acotado superiormente tiene máximo, pero como, aunque sí que lo hemos mencionado en la sección A.4, no lo hemos probado con detalle, hemos preferido no emplear este argumento.



Bernard Bolzano
(1781–1848)



Augustin Louis Cauchy
(1789–1857)

Proposición A.13.

Se cumple lo siguiente:

- *Sean α y β dos números reales, con $\alpha < \beta$. Entonces, existe $r \in \mathbb{Q}$ tal que $\alpha < r < \beta$.*
- *Sean r y s dos números racionales, con $r < s$. Entonces, existe $\alpha \in \mathbb{R} \setminus \mathbb{Q}$ tal que $r < \alpha < s$.*
- *Sean α y β dos números reales, con $\alpha < \beta$. Entonces, existe $\gamma \in \mathbb{R} \setminus \mathbb{Q}$ tal que $\alpha < \gamma < \beta$.*

La primera parte se deduce de que, por la proposición A.10, existe $n \in \mathbb{N}$ con $1/n < \beta - \alpha$, de donde $n\beta - n\alpha > 1$ y, por la proposición A.12, existe $k \in \mathbb{Z}$ con $n\alpha < k < n\beta$; ahora se toma $r = k/n$. Para la segunda parte basta usar que existe un irracional entre 0 y 1 (o entre cualquier otra pareja fija de racionales), desplazar y escalar. La tercera parte es consecuencia de las otras dos (entre α y β podemos intercalar racionales dos veces).

Con los reales a nuestra disposición, tenemos los cimientos para construir todo el edificio del análisis matemático partiendo de las definiciones de límite y continuidad de funciones mediante los consabidos ε y δ. Como referencia histórica, podemos señalar que fue Bernard Bolzano en 1817 el primero que dio una definición de límite razonablemente rigurosa, y Augustin Louis Cauchy en 1820 el primero que usó la notación de ε y δ; pocos años más tarde, fue Karl Weierstrass quien formuló las definiciones tal como las usamos ahora (aunque, a decir verdad, todo eso fue anterior a la existencia de una construcción rigurosa de los números reales como la que aquí hemos visto).

Un teorema paradigmático en el desarrollo riguroso del análisis matemático es el de Bolzano, que asegura que, si una función continua en un intervalo es positiva en un punto y negativa en otro, debe anularse en un punto intermedio. Para probar el

teorema de Bolzano es imprescindible usar, de alguna manera, que $\mathbb{R}$ es completo, y un resultado similar no es cierto en $\mathbb{Q}$ (por ejemplo, la función $f\colon \mathbb{Q} \to \mathbb{Q}$ dada por $f(x) = x^2 - 2$ es negativa en $x = 1$ y positiva en $x = 2$, pero no se anula en ningún $x \in \mathbb{Q}$).

Finalmente, tras los conceptos de límite y continuidad, siguen la derivación, integración, teorema fundamental del cálculo..., que sin duda el lector ya conocerá, y que, además, explicándolos nos desviaríamos demasiado de los objetivos de este libro.

A.6.2. Unicidad de los números reales

Como acabamos de comentar, una vez que tenemos los números reales, con su propiedad característica de ser un cuerpo ordenado completo, podemos desarrollar el análisis matemático. Hasta tal punto que, si tuviésemos otro cuerpo ordenado completo, sobre él podríamos construir «otro» análisis matemático. En este apartado vamos a ver que esto no es posible, pues, en cierto modo que ahora precisaremos, $\mathbb{R}$ es el único cuerpo ordenado completo.

Está claro que la unicidad no podemos tomarla en sentido literal, pues si definimos $\widetilde{\mathbb{R}} = \{(a, a) : a \in \mathbb{R}\}$ y, en él,

$$\begin{aligned}(a, a) \oplus (b, b) &= (a + b, a + b),\\ (a, a) \odot (b, b) &= (a \cdot b, a \cdot b),\\ (a, a) \preceq (b, b) &\iff a \leq b,\end{aligned}$$

es trivial que $\widetilde{\mathbb{R}}$ con las operaciones $\oplus$ y $\odot$ y el orden $\preceq$ también es cuerpo ordenado completo. Pero este ejemplo no introduce ninguna novedad sobre el $\mathbb{R}$ que ya teníamos. Es, simplemente, un cambio de nombre y de notación. Así que, al hablar de unicidad, tenemos que descartar estos procesos de cambio de nombre.

En matemáticas, eso se hace hablando de unicidad salvo isomorfismos; de hecho, es tan común que, a menudo, la aclaración «salvo isomorfismos» se omite. En lo que ahora nos concierne, debemos primero establecer con precisión qué entendemos por un cuerpo ordenado completo, y a qué nos referimos con un isomorfismo entre dos de ellos.

Llamamos *cuerpo ordenado* a un cuerpo $(F, +, \cdot)$ con un subconjunto P de F (los elementos «positivos») de manera que se cumplen las siguientes propiedades:

- Cada $a \in F$ satisface una y sólo una de estas alternativas: $a = 0$, $a \in P$ o $-a \in P$ (donde, por supuesto, $-a$ denota el opuesto de a).
- Si a y b están en P, entonces $a + b$ está en P.
- Si a y b están en P, entonces $a \cdot b$ está en P.

A partir de la definición genérica de cuerpo ordenado, resulta natural introducir una notación de orden como las que teníamos en $\mathbb{Q}$ o $\mathbb{R}$, y para eso decimos que: $a > b$ si $a - b \in P$; $a < b$ si $b - a \in P$; $a \geq b$ si $a = b$ o $a - b \in P$; $a \leq b$ si $a = b$ o $b - a \in P$. (Además, es obvio que da lo mismo usar la notación $a > 0$ que $a \in P$, y $a < 0$ que $-a \in P$.)

Con eso, ya tiene sentido decir que un subconjunto $A \subseteq F$ es acotado superiormente si existe $x \in F$ tal que $x \geq a$ para todo $a \in A$; tal x se denomina cota superior. Además, una cota superior x se denomina supremo de A si $x \leq y$ para toda $y \in F$ que sea cota superior de A. Finalmente, un cuerpo ordenado se denomina *completo* cuando cualquier subconjunto no vacío que sea acotado superiormente tiene supremo.

Supongamos ahora que F_1 con sus operaciones $+$, $\cdot$, su conjunto de elementos positivos P_1, y su orden $<$ es un cuerpo ordenado; y que F_2 con sus respectivos $\oplus$, $\odot$, P_2, y $\prec$ es otro cuerpo ordenado. Un *isomorfismo* (de cuerpos ordenados) entre F_1 y F_2 es una aplicación $f\colon F_1 \to F_2$ con las siguientes propiedades:

(a) Si $x \neq y$ entonces $f(x) \neq f(y)$ (es decir, f es inyectiva).

(b) Para cada $z \in F_2$ existe $x \in F_1$ tal que $f(x) = z$ (es decir, f es suprayectiva).

(c) Para cualesquiera $x, y \in F_1$, se cumple

$$f(x + y) = f(x) \oplus f(y),$$
$$f(x \cdot y) = f(x) \odot f(y).$$

(d) Si $x < y$, entonces $f(x) \prec f(y)$.

Una vez aclarado a qué nos referíamos al decir que $\mathbb{R}$ es el único cuerpo ordenado completo, y antes de abordar la demostración de esta afirmación, veamos algunas propiedades que deben cumplir todos los cuerpos ordenados.

Supongamos que F con sus operaciones $(F, +, \cdot)$, su conjunto de elementos positivos P_F y su relación de orden $<$ es un cuerpo ordenado, y usemos 0_F y 1_F para designar, respectivamente, los elementos neutros de la suma y del producto en F (por definición de cuerpo podemos asegurar que $0_F \neq 1_F$). Entonces, si para cada $n \in \mathbb{N}$ denotamos

$$\underbrace{1_F + \cdots + 1_F}_{n \text{ veces}} = n_F,$$

está claro que F contiene a un $\mathbb{N}_F = \{n_F : n \in \mathbb{N}\}$ que es análogo a los naturales ordinarios (si lo queremos con cero, como $\mathbb{N}_0$, bastaría tomar $\mathbb{N}_F \cup \{0_F\}$). Obsérvese que no puede ocurrir que $n_F = m_F$ si $n \neq m$ ya que F es un cuerpo ordenado: en

efecto, si suponemos que $n < m$ (lo cual puede hacerse sin pérdida de generalidad) entonces tendremos

$$m_F - n_F = (m-n)_F = \underbrace{1_F + \cdots + 1_F}_{m-n \text{ veces}} \geq 1_F > 0_F$$

(que $1_F > 0_F$ requiere su pequeña demostración: si fuese $1_F < 0_F$, entonces $-1_F > 0_F$ y $1_F = (-1_F)^2 > 0_F$, que es contradictorio). Además, $\mathbb{N}_F$ no está acotado superiormente (sirve la misma demostración que en la proposición A.9).

También tenemos algo parecido a los enteros, sin más que tomar

$$\mathbb{Z}_F = \{n_F : n \in \mathbb{N}\} \cup \{0_F\} \cup \{-n_F : n \in \mathbb{N}\}.$$

Finalmente, también hay en F algo parecido a los racionales ordinarios: para cada $h \in \mathbb{Z}_F$ y $k \in \mathbb{N}_F$ tiene sentido hablar de $h/k = h \cdot k^{-1}$ (donde k^{-1} es el inverso multiplicativo de k).

Con todo esto, vamos ya con el importante resultado que perseguíamos en este apartado:

Proposición A.14.

Si F es un cuerpo ordenado completo, entonces F es isomorfo a $\mathbb{R}$.

Demostración. Vamos a construir una aplicación $f\colon \mathbb{R} \to F$ que sea un isomorfismo. En F, denotemos su suma con $\oplus$, su producto con $\odot$ y su relación de orden con $\prec$ (para no recargar la notación, usamos $-$ y $/$ para la resta y el cociente tanto en F, como en $\mathbb{R}$); además, emplearemos 0_F y 1_F para designar, respectivamente, los elementos neutros de la suma y del producto en F. Así mismo, denotemos

$$\underbrace{1_F \oplus \cdots \oplus 1_F}_{n \text{ veces}} = n_F, \qquad n > 0.$$

Comencemos definiendo f para los enteros (por supuesto, estamos usando $\mathbb{Z} \subset \mathbb{R}$) como sigue:

$$f(n) = \begin{cases} 0_F, & \text{si } n = 0, \\ n_F, & \text{si } n > 0, \\ -|n|_F, & \text{si } n < 0. \end{cases}$$

A partir de aquí, resulta fácil comprobar que

$$\begin{aligned} f(m+n) &= f(m) \oplus f(n), \\ f(m \cdot n) &= f(m) \odot f(n), \\ m < n \;&\Rightarrow\; f(m) \prec f(n) \end{aligned}$$

para todos los enteros m y n. A continuación, definamos f para los números racionales mediante

$$f(m/n) = f(m)/f(n).$$

Esto tiene sentido, pues, en primer lugar, $f(n) = n_F \neq 0_F$ cuando $n > 0$ por ser F un cuerpo ordenado; y, en segundo lugar, si $m/n = j/k$, entonces $mk = nj$, luego $f(m) \odot f(k) = f(n) \odot f(j)$ y $f(m)/f(n) = f(j)/f(k)$. Con esto, de nuevo resulta sencillo comprobar que

$$\begin{aligned} f(r+s) &= f(r) \oplus f(s), \\ f(r \cdot s) &= f(r) \odot f(s), \\ r < s \ &\Rightarrow \ f(r) \prec f(s) \end{aligned}$$

para todos los racionales r y s.

Falta aún definir $f(x)$ para $x \in \mathbb{R}$ arbitrario. Tal como se define el número real x mediante cortaduras, este x está determinado por todos los racionales menores que él; la imagen mediante f de esta cortadura será el conjunto

$$A_x = \{f(r) : r \in \mathbb{Q},\ r < x\}.$$

Este A_x es no vacío y acotado superiormente (basta tomar $r_0 \in \mathbb{Q}$ que cumpla $r_0 > x$, con lo cual $f(r_0) \succ f(r)$ para todos los $f(r)$ de A_x). Al ser F un cuerpo ordenado completo, el conjunto A_x tiene supremo; definimos $f(x)$ como el supremo de A_x.

Ahora surge un pequeño problema pues, si x es racional, $f(x)$ está definido de dos maneras distintas: como racional y como real. Antes de proseguir debemos demostrar que estas dos definiciones coinciden. Es decir, tenemos que ver que, si $x \in \mathbb{Q}$, entonces

$$\sup A_x = f(x) \tag{A.2}$$

(este $f(x)$ de la derecha se refiere, por supuesto, a $f(x) = f(m/n) = f(m)/f(n)$ con $x = m/n$). Que se cumple (A.2) no es inmediato, y requiere detenerse un poco.

Observemos que, como $\mathbb{N}_F$ no está acotado superiormente, los mismos argumentos que se usan para demostrar la proposición A.13 sirven —trasladados a F— para probar que, si $a, b \in F$ cumplen $a \prec b$, entonces existe $r \in \mathbb{Q}$ tal que $a \prec f(r) \prec b$. Con esto, abordemos (A.2).

Para $r \in \mathbb{Q}$ con $r < x$, sabemos que $f(r) \prec f(x)$. Así pues, todos los elementos de A_x son $\prec f(x)$ y, por tanto, el supremo cumple

$$\sup A_x \prec f(x) \qquad \text{o} \qquad \sup A_x = f(x).$$

Supongamos que fuese $\sup A_x \prec f(x)$ lo que realmente sucede. Entonces, existirá $r \in \mathbb{Q}$ tal que $\sup A_x \prec f(r) \prec f(x)$. Pero $f(r) \prec f(x)$ sólo ocurre cuando $r < x$, lo cual significa que $f(r) \in A_x$, y esto es una clara contradicción con $\sup A_x \prec f(r)$. Por consiguiente, la única posibilidad aceptable es $\sup A_x = f(x)$, hemos probado (A.2), y ya tenemos bien definida la aplicación $f\colon \mathbb{R} \to F$.

Nos falta ahora ver que la aplicación f que acabamos de construir es un isomorfismo. Y, para ello, hay que comprobar que se satisfacen las condiciones (a)–(d) de la definición. Vamos con ello (lo haremos en desorden).

(d) Si x, y son dos números reales con $x < y$, es claro que $A_x \subseteq A_y$ y, por tanto,

$$f(x) = \sup A_x \preceq \sup A_x = f(y).$$

Pero queremos ver que $f(x) \prec f(y)$, así que debemos descartar la posibilidad $f(x) = f(y)$. Para ello, usemos que existen números racionales r, s tales que $x < r < s < y$. Como $f(r) \prec f(s)$, de aquí se sigue que $f(x) \preceq f(r) \prec f(s) \preceq f(y)$, y ya tenemos (d).

(a) Es obvio aplicando (d) pues, si $x \neq y$, deberá ser $x < y$ (luego $f(x) \prec f(y)$) o $y < x$ (luego $f(y) \prec f(x)$).

(b) Sea $a \in F$; ¿cómo encontrar $x \in \mathbb{R}$ tal que $f(x) = a$? Tomemos

$$A = \{r \in \mathbb{Q} : f(r) \prec a\}.$$

El conjunto A es no vacío, y también es acotado superiormente (existe $s \in \mathbb{Q}$ con $f(s) \succ a$, luego $f(s) \succ f(r)$ cuando $r \in A$, y por tanto $s > r$ para cada $r \in A$), así que tiene supremo en $\mathbb{R}$. Tomemos $x = \sup A$ y vamos a demostrar que $f(x) = a$. Lo vamos a hacer eliminando las dos alternativas

$$f(x) \prec a \qquad \text{o} \qquad a \prec f(x).$$

En el primer caso debería existir $r \in \mathbb{Q}$ con $f(x) \prec f(r) \prec a$; pero esto implica que $x < r$ y $r \in A$, lo cual es absurdo ya que $x = \sup A$. En el segundo caso debería existir $r \in \mathbb{Q}$ con $a \prec f(r) \prec f(x)$; esto implica que $r < s$ y, por ser $x = \sup A$, de aquí se sigue que $r < s$ para algún $s \in A$; entonces, $f(r) \prec f(s) \prec a$, que de nuevo es absurdo. En consecuencia, sólo queda la posibilidad $f(x) = a$, y hemos probado (b).

(c) Comencemos demostrando que, para todo $x, y \in \mathbb{R}$, se cumple $f(x + y) = f(x) \oplus f(y)$. Supongamos que no es cierto, es decir, que existen $x, y \in \mathbb{R}$ para los cuales $f(x + y) \neq f(x) \oplus f(y)$. Entonces, se cumplirá

$$f(x + y) \prec f(x) \oplus f(y) \qquad \text{o} \qquad f(x) \oplus f(y) \prec f(x + y).$$

En el primer caso, existirá $r \in \mathbb{Q}$ tal que

$$f(x+y) \prec f(r) \prec f(x) \oplus f(y),$$

y eso implica que $x+y<r$. De aquí se sigue que r podrá escribirse como

$$r = r_1 + r_2, \qquad r_1, r_2 \in \mathbb{Q}, \qquad x < r_1,\ y < r_2.$$

Ahora, usando que para los racionales sí podemos asegurar que $f(r_1+r_2) = f(r_1) \oplus f(r_2)$, llegamos a

$$f(r) = f(r_1+r_2) = f(r_1) \oplus f(r_2) \succ f(x) \oplus f(y),$$

que es absurdo pues teníamos $f(r) \prec f(x) \oplus f(y)$. El otro caso se analiza de manera similar.

Finalmente el mismo tipo de argumentación permite probar que $f(x \cdot y) = f(x) \odot f(y)$ para todo $x, y \in \mathbb{R}$, lo cual concluye la demostración de (c) y de la proposición. □

A.6.3. Otras construcciones de los números reales

La construcción que hemos hecho de los números reales por medio de las cortaduras de Dedekind es una de las que más se usan en la actualidad, pero no la única posible. En todo caso, como lo que pretendemos es llegar a un cuerpo ordenado completo, la proposición A.14 nos asegura que lo que podamos obtener por otros métodos deberá ser forzosamente una construcción de los reales isomorfa a la que aquí hemos visto. En la práctica, esto quiere decir que, cuando ya estamos usando los números reales sin tener en mente cómo están construidos (sino asimilándolos a la idea intuitiva de recta real, por ejemplo), estaremos manejando el «mismo» conjunto de números reales, con idénticas propiedades.

Por supuesto, para cualquier construcción alternativa el proceso debe involucrar los mismos pasos que los aquí seguidos:

1.º Dar una definición de número real a partir de los números racionales.

2.º En el conjunto de los números reales, definir la relación de orden y las operaciones.

3.º Comprobar la completitud y todas las propiedades que caracterizan un cuerpo ordenado.

Según el procedimiento elegido, son unos u otros los pasos más costosos pero, al final, siempre hay una importante labor rutinaria de comprobaciones.

Por ejemplo, ya hemos comentado que una manera de definir los números reales sería intentar hacerlo a partir de la idea preconcebida de tomar números con infinitos decimales,[5]

$$d_0 + 0.d_1d_2d_3d_4\ldots = d_0 + \sum_{j=1}^{\infty} \frac{d_j}{10^j}, \quad d_0 \in \mathbb{Z},\ d_j \in \{0,1,2,\ldots,9\} \text{ para } j \geq 1.$$

Intuitivamente, esto es una especie de proceso de paso al límite a partir de sus «aproximaciones» con un número finito de decimales, que son números racionales preexistentes, pero para dar una definición rigurosa hay que tener más cuidado (en particular, obsérvese que, como los reales no existen aún, no podemos hablar de la convergencia de esa serie). En su lugar, lo que hacemos es definir número real como un par $(d_0, \{d_j\}_{j=1}^{\infty})$ con $d_0 \in \mathbb{Z}$ y $d_j \in \{0,1,2,\ldots,9\}$ para $j \geq 1$. A partir de aquí, definir la relación de orden es sencillo:

$$(a_0, \{a_j\}_{j=1}^{\infty}) < (b_0, \{b_j\}_{j=1}^{\infty})$$

si $a_0 < b_0$ o $a_j = b_j$ para $0 \leq j < n$ y $a_n < b_n$. Con eso, no es difícil demostrar que cualquier conjunto acotado de números reales tiene supremo. Lo problemático es definir la suma y el producto (si intentamos usar el proceso normal de sumar o multiplicar números con decimales, en seguida nos damos cuenta de que no hay manera de comenzar operando con decimales infinitamente lejos). Para dar las definiciones, introduzcamos antes un poco de notación. Dado un número real $\alpha = (a_0, \{a_j\}_{j=1}^{\infty})$ y un entero positivo n, utilicemos α_n para denotar el número racional

$$\alpha_n = a_0 + \sum_{j=1}^{n} \frac{a_j}{10^j};$$

y, recíprocamente, si tenemos un número racional $r = a_0 + \sum_{j=1}^{n} a_j/10^j$ (con $a_0 \in \mathbb{Z}$ y $a_j \in \{0,1,2,\ldots,9\}$ para $j \geq 1$), denotemos mediante r' al número real

$$r' = (a_0, \{a'_j\}_{j=1}^{\infty}), \quad \text{con } a'_j = a_j \text{ para } 1 \leq j \leq n \text{ y } a'_j = 0 \text{ para } j > n.$$

Entonces, para dos números reales α y β, su suma se define como

$$\alpha + \beta = \sup\{(\alpha_n + \beta_n)' : n \in \mathbb{N}\}$$

[5] Obsérvese que ésta no es, realmente, la notación habitual de un número decimal, pues, tal como lo estamos tomando ahora, la parte decimal no es negativa aunque lo sea el número. Es una modificación sin importancia, pero simplifica las definiciones.

(el supremo existe pues previamente hemos visto que cualquier conjunto acotado lo tiene); y el producto se define de manera similar. Quedaría ahora comprobar todas las propiedades de cuerpo, tarea que ya no vamos a realizar (como con las cortaduras, lo más complicado es demostrar la existencia de inverso respecto al producto).

Hay matemáticos que defienden esta construcción como la más natural, por ser la más intuitiva y la que más se aproxima a nuestro manejo cotidiano de los reales. Pero no es ése el consenso general, sino que habitualmente se considera que es un mero ejercicio intelectual destinado a dar una definición cercana a la representación de los números reales como números con infinitos decimales —tal como aquí mismo hemos hecho—, así que la construcción anterior no se suele usar. Sí que hay otras dos construcciones de los números reales que son relativamente comunes, y son las basadas en estas dos ideas:

- Sucesiones de Cauchy de números racionales.
- Pares de sucesiones monótonas convergentes (que también se pueden interpretar como intervalos encajados).

Vamos a comentarlas ambas.

En $\mathbb{Q}$ (igual que cualquier otro espacio en el que exista la noción de módulo, norma o similar), una sucesión $\{r_n\}_{n=1}^{\infty}$ se dice de Cauchy si, para cualquier $\varepsilon > 0$, existe n_0 tal que $|r_n - r_m| < \varepsilon$ cuando $n, m \geq n_0$.

Seguro que al lector le suena que, en $\mathbb{R}$, cualquier sucesión de Cauchy es convergente (pero, claro, eso aún no lo podemos usar, pues estamos intentando construir $\mathbb{R}$, así que de momento no existe). En cambio, una sucesión de Cauchy de racionales puede no converger en $\mathbb{Q}$. Por ejemplo, la sucesión definida mediante

$$r_1 = 1, \qquad r_{n+1} = r_n - \frac{r_n^2 - 2}{2r_n}, \quad n \geq 1. \tag{A.3}$$

es de Cauchy (no es difícil comprobarlo) pero no puede converger en $\mathbb{Q}$: si se cumpliera $r_n \to \ell$, también tendríamos $r_{n+1} \to \ell$ y, tomando límites en (A.3), llegaríamos a $\ell^2 = 2$, lo cual sabemos que no es posible en $\mathbb{Q}$.[6]

[6] Expliquemos el origen de la recurrencia (A.3). Si queremos resolver una ecuación $f(x) = 0$, podemos intentar hacerlo mediante el método de Newton (también llamado, por su interpretación geométrica, método de la tangente) en el que, partiendo de un valor inicial x_1, se itera mediante

$$x_{n+1} = x_n - f(x_n)/f'(x_n), \quad n \geq 1.$$

Ante determinadas hipótesis no muy restrictivas, esta sucesión $\{x_n\}_{n=1}^{\infty}$ converge a un número real ξ que es raíz de f, esto es, que cumple $f(\xi) = 0$. Cuando se aplica a $f(x) = x^2 - 2$ y se parte de un número positivo, el proceso converge hacia $\xi = \sqrt{2}$, la raíz positiva de f; la definición recurrente (A.3) es, precisamente, el método de Newton aplicado a $f(x) = x^2 - 2$.

Así las cosas, la idea es decir que, por definición, un número real es una sucesión de Cauchy de racionales, lo cual, de alguna manera, representará a ese quizás inexistente límite en $\mathbb{Q}$. Pero esto tiene un problema: muchas sucesiones dan lugar al mismo límite; y una solución: tomar clases de equivalencias de sucesiones de Cauchy que convergerían al mismo número real. Esto se hace como sigue.

Dos sucesiones de Cauchy de racionales, $\{a_n\}_{n=1}^{\infty}$ y $\{b_n\}_{n=1}^{\infty}$ se dicen equivalentes si

$$\lim_n (a_n - b_n) = 0$$

(es muy sencillo comprobar que, efectivamente, esta relación es de equivalencia). Entonces, por definición, cada clase de equivalencia es un número real; y $\mathbb{R}$ es el conjunto de todos ellos.

Cuando tenemos dos números reales α y β, representados por dos sucesiones de Cauchy de racionales $\alpha = \{a_n\}_{n=1}^{\infty}$ y $\beta = \{b_n\}_{n=1}^{\infty}$, su suma y su producto se definen, respectivamente, como

$$\alpha + \beta = \{a_n + b_n\}_{n=1}^{\infty}, \qquad \alpha \cdot \beta = \{a_n b_n\}_{n=1}^{\infty}.$$

Para que esto tenga sentido hay que probar que $\{a_n + b_n\}_{n=1}^{\infty}$ y $\{a_n b_n\}_{n=1}^{\infty}$ son también sucesiones de Cauchy de racionales; y, además, que la definición es compatible con la relación de equivalencia (es decir, que si se eligen para α y β otros dos representantes de su clase de equivalencia, las operaciones con los nuevos representantes y con los antiguos dan lugar a sucesiones equivalentes y, por tanto, al mismo número real). Hay también que definir la relación de orden: dados $\alpha = \{a_n\}_{n=1}^{\infty}$ y $\beta = \{b_n\}_{n=1}^{\infty}$, decimos que $\alpha \leq \beta$ si existe cierto n_0 tal que $a_n \leq b_n$ para todo $n \geq n_0$ (y decimos que $\alpha < \beta$ si $\alpha \leq \beta$ pero $\alpha \neq \beta$). Por supuesto, de nuevo hay que demostrar que esto es compatible con la relación de equivalencia.

Una vez dadas las definiciones, se concluye la construcción comprobando todos los detalles; es decir, viendo que la suma, el producto y la relación de orden así definidas dan lugar a un cuerpo ordenado completo. La parte conceptualmente más importante es demostrar que toda sucesión de Cauchy de números reales —¡de los números reales que acabamos de definir!— es convergente, lo cual se traduce en que todo conjunto acotado de números reales tiene supremo (en cambio, en la construcción de los números basada en cortaduras, primero se prueba que todo conjunto acotado superiormente tiene supremo y, a partir de allí, que toda sucesión de Cauchy en $\mathbb{R}$ es convergente).

La última construcción de los reales que vamos a abordar es la basada en pares de sucesiones monótonas convergentes. Por supuesto, una sucesión $\{a_n\}_{n=1}^{\infty}$ se llama monótona creciente si verifica $a_{n+1} \geq a_n$ para todo n; y monótona decreciente si verifica $a_{n+1} \leq a_n$ para todo n. Así mismo, se denomina par de sucesiones monótonas

convergentes a dos sucesiones $\{a_n\}_{n=1}^{\infty}$ y $\{b_n\}_{n=1}^{\infty}$ tales que la primera es monótona creciente, la segunda monótona decreciente, $a_n \leq b_n$ para todo n, y, además,

$$\lim_n (b_n - a_n) = 0$$

(para evitar posibles equívocos, obsérvese que al llamar «convergentes» al par de sucesiones de racionales no se quiere decir que esas sucesiones converjan en $\mathbb{Q}$, y de hecho lo que tiene interés es que estas sucesiones no tienen por qué converger en $\mathbb{Q}$, aunque a la postre sí convergerán en $\mathbb{R}$, cuando lo hayamos definido).

Como los reales aún no existen, tomemos $(\{a_n\}_{n=1}^{\infty}, \{b_n\}_{n=1}^{\infty})$ un par de sucesiones monótonas convergentes de números racionales. La intuición aplicada a la recta real nos dice que las dos sucesiones $\{a_n\}_{n=1}^{\infty}$ y $\{b_n\}_{n=1}^{\infty}$, que se acercan cada vez más, van a converger, una por la derecha y otra por la izquierda, a un punto común de la recta real. Así las cosas, ese par de sucesiones monótonas convergentes es, por definición, un número real. Pero, claro, de la misma manera que en la construcción basada en sucesiones de Cauchy, hay muchos pares de sucesiones monótonas convergentes que darían lugar al mismo número real, y esto se soluciona, una vez más, tomando previamente clases de equivalencia de pares de sucesiones monótonas convergentes.

De manera alternativa, un par de sucesiones monótonas convergentes de números racionales se puede entender como una sucesión de intervalos (cerrados) encajados. Si, dados $a, b \in \mathbb{Q}$ con $a \leq b$, llamamos intervalo $[a, b]$ a

$$[a, b] = \{r \in \mathbb{Q} : a \leq r \leq b\},$$

el par $(\{a_n\}_{n=1}^{\infty}, \{b_n\}_{n=1}^{\infty})$ se puede interpretar como

$$[a_1, b_1] \supseteq [a_2, b_2] \supseteq [a_3, b_3] \supseteq \cdots$$

con $\lim_n (b_n - a_n) = 0$. Intuitivamente, la intersección de todos esos intervalos es un único punto de la recta real, así que, por definición, un número real sería ahora una sucesión de intervalos encajados. Y, como muchas sucesiones de intervalos encajados darían lugar al mismo número real, de nuevo hay que tomar clases de equivalencia.

En cualquier caso, ahora habría que definir las operaciones y el orden, y finalmente comprobar todos los detalles que hacen que estas construcciones den lugar a un cuerpo ordenado completo.

Nota A.15. En el apartado A.6.2 hemos visto que, salvo isomorfismos, $\mathbb{R}$ es el único cuerpo ordenado completo. Si sustituimos el requisito de completitud (es decir, que cualquier subconjunto acotado superiormente tenga supremo) por el hecho de que las sucesiones de Cauchy sean convergentes, la unicidad yo no es cierta; para

conseguir unicidad, hay que añadir una hipótesis adicional: exigir que el cuerpo cumpla la propiedad arquimediana. De este modo, lo que se cumple es que $\mathbb{R}$ es, salvo isomorfismos, el único cuerpo ordenado y arquimediano en el que las sucesiones de Cauchy son convergentes. Lo mismo ocurre si la completitud se sustituye, en lugar de por la convergencia de las sucesiones de Cauchy, por el requisito de que cualquier sucesión de intervalos cerrados encajados cuya longitud tiende a 0 tiene como intersección un punto; la unicidad salvo isomorfismos se consigue, de nuevo, exigiendo la propiedad arquimediana. ◁

A.7. Los números complejos

Con los números reales, hay algunas cosas que no se pueden hacer. Por ejemplo, no hay ningún número real cuyo cuadrado valga -1. Pero, ¿por qué debería haberlo? En siglo XVI, cuando los matemáticos italianos estaban intentando resolver ecuaciones cúbicas—en cuanto a su relación con los números complejos merece la pena citar a Girolamo Cardano y a Rafael Bombelli—, se dieron cuenta de que era muy útil que los números negativos tuviesen raíces cuadradas. Ellos sabían que una solución de la ecuación

$$x^3 = 3px + 2q$$

venía dada por la fórmula[7]

$$x = \sqrt[3]{q + \sqrt{q^2 - p^3}} + \sqrt[3]{q - \sqrt{q^2 - p^3}}.$$

[7] Recordemos cómo se llega a la fórmula que proporciona la solución de $x^3 = 3px + 2q$. Tomemos $x = u + v$, con lo cual la ecuación queda

$$u^3 + 3u^2v + 3uv^2 + v^3 = 3p(u + v) + 2q,$$

es decir,

$$3uv(u + v) + u^3 + v^3 = 3p(u + v) + 2q.$$

Igualando coeficientes, esto nos sugiere tomar $p = uv$ y $q = (u^3 + v^3)/2$. Despejando v en la primera expresión y sustituyendo en la segunda obtenemos $q = u^3/2 + p^3/(2u^3)$, lo que, a su vez, da lugar a

$$u^6 - 2qu^3 + p^3 = 0,$$

una ecuación cuadrática en u^3 cuyas soluciones son $u^3 = q \pm \sqrt{q^2 - p^3}$. Siguiendo un procedimiento similar, para v^3 se obtiene la misma expresión (lógico, dado que los papeles de u y v son simétricos), así que, si nos quedamos con la de signo «+» para u^3 y la de signo «−» para v^3, tenemos

$$u = \sqrt[3]{q + \sqrt{q^2 - p^3}}, \qquad v = \sqrt[3]{q - \sqrt{q^2 - p^3}},$$

y sustituyendo en $x = u + v$ se llega a la solución de $x^3 = 3px + 2q$.

En el caso de la ecuación $x^3 = 15x + 4$ (es decir, con $p = 5$ y $q = 2$), la fórmula anterior da lugar a

$$x = \sqrt[3]{2 + \sqrt{-121}} + \sqrt[3]{2 - \sqrt{-121}}.$$

En principio, y antes de haber definido los números complejos, eso no parece tener ningún sentido; pero es muy sencillo comprobar que la ecuación $x^3 = 15x + 4$ tiene tres soluciones reales, que son 4, $-2 + \sqrt{3}$ y $-2 - \sqrt{3}$, ninguna de las cuales parece desprenderse de manera directa de la fórmula. Aun sin estar convencido de que las raíces de los números negativos tuviesen sentido, Bombelli se dio cuenta de que, si aplicaba las reglas habituales de la aritmética a tales expresiones, lograba probar que[8]

$$\sqrt[3]{2 + \sqrt{-121}} + \sqrt[3]{2 - \sqrt{-121}} = 4.$$

¡De manera sorprendente, y usando como paso intermedio unos números que sólo estaban en la imaginación, la fórmula para resolver la ecuación cúbica conducía a una solución real![9]

Desde luego, todo eso no estaba bien fundado, y de ahí que costara mucho que esos números tan raros —Descartes los llamó «imaginarios»— fuesen plenamente aceptados por los matemáticos. En el siglo XVIII, muchos los usaban ya de manera regular —entre ellos, Gottfried von Leibniz, Johann Bernoulli, Johann Lambert, Joseph Louis Lagrange y, ¡cómo no!, el gran Leonhard Euler—, pero aun así seguían generando controversia. La satisfacción lógica sobre los números complejos entraba,

[8] Veamos cómo lo hizo. Elevando al cubo $2 + \sqrt{-1}$ y $2 - \sqrt{-1}$ mediante las mismas reglas que se aplican a los números ordinarios se obtiene

$$\begin{aligned}\left(2 + \sqrt{-1}\right)^3 &= 8 + 12\sqrt{-1} + 6\left(\sqrt{-1}\right)^2 + \left(\sqrt{-1}\right)^3 = 8 + 12\sqrt{-1} - 6 - \sqrt{-1} \\ &= 2 + 11\sqrt{-1} = 2 + \sqrt{-121}, \\ \left(2 - \sqrt{-1}\right)^3 &= 8 - 12\sqrt{-1} + 6\left(\sqrt{-1}\right)^2 - \left(\sqrt{-1}\right)^3 = 8 - 12\sqrt{-1} - 6 + \sqrt{-1} \\ &= 2 - 11\sqrt{-1} = 2 - \sqrt{-121},\end{aligned}$$

de donde $\sqrt[3]{2 + \sqrt{-121}} + \sqrt[3]{2 - \sqrt{-121}} = \left(2 + \sqrt{-1}\right) + \left(2 - \sqrt{-1}\right) = 4$.

[9] Obsérvese en particular que, históricamente, fueron las ecuaciones cúbicas las que contribuyeron al desarrollo de los números complejos, y no —como nos podría parecer a primera vista— el hecho de que la ecuación cuadrática $ax^2 + bx + c = 0$ tiene soluciones complejas $x = \left(-b \pm \sqrt{b^2 - 4ac}\right)/(2a)$ cuando $4ac > b^2$; en la época, que estas ecuaciones cuadráticas no tuviesen solución no constituía ningún problema. Insistiendo en ese argumento, nótese que resolver la ecuación cuadrática $x^2 = bx + c$ equivale a buscar la intersección de la parábola $y = x^2$ con la recta $y = bx + c$, y el caso de discriminante negativo que daría lugar a números complejos se corresponde con una parábola y una recta que no se cortan, lo cual es una situación geométrica muy natural.

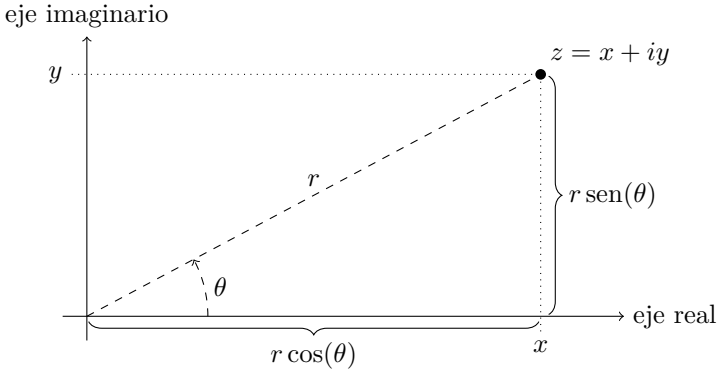


Figura A.2: El complejo $z = x + iy$ en el plano de Argand.

a finales del siglo XVIII, más en el terreno de la filosofía que en el de las matemáticas. El mismo Gauss, ya en 1821, señalaba que «la verdad metafísica de $\sqrt{-1}$ es elusiva».

Para entender de alguna manera los números complejos, sin duda resulta útil su representación geométrica como puntos del plano, que tiene sus orígenes en los trabajos de 1797 del noruego-danés Caspar Wessel y en los del suizo Jean Robert Argand en 1806. No obstante, esta representación de los complejos, en lo que ahora denominamos plano de Argand (o diagrama de Argand), no tuvo su impulso definitivo hasta que la usó Gauss en 1831. En el plano complejo —y aunque, ante la confianza de que el lector ya sabe de qué hablamos, estamos usando la notación antes de darla formalmente—, la para nosotros bien conocida fórmula

$$z = x + iy = r(\cos(\theta) + i\,\mathrm{sen}(\theta)), \qquad r = |z| = \sqrt{x^2 + y^2}, \quad \theta = \arg(z),$$

se representa con toda naturalidad (véase la figura A.2).

Finalmente, fue el irlandés William Rowan Hamilton en 1833 quien da la primera definición algebraica rigurosa de los complejos, asimilándolos a pares de números reales con una suma y un producto diseñados a propósito. Esencialmente, es ésta la definición que vamos a ver ahora.

El conjunto de los números complejos, que se denota con $\mathbb{C}$, se define como

$$\mathbb{C} = \mathbb{R} \times \mathbb{R}$$

con la suma

$$(a, b) + (c, d) = (a + c, b + d)$$

y el producto

$$(a,b)\cdot(c,d)=(ac-bd,ad+bc).$$

Es un mero ejercicio comprobar que tales operaciones son conmutativas y asociativas, y que además se cumple la propiedad distributiva del producto respecto a la suma. Así mismo, $\mathbb{R}$ se «incluye» en $\mathbb{C}$ sin más que asignar, a cada $a\in\mathbb{R}$, el complejo $(a,0)$, y se denota $(a,0)=a$. Obsérvese que, al operar un número real t con un complejo (a,b), lo que se obtiene es $t+(a,b)=(t+a,b)$ y $t\cdot(a,b)=(ta,tb)$.

Antes de proseguir, detengámonos en un complejo muy importante que tiene una notación propia: $(0,1)=i$ se denomina unidad imaginaria, y cumple $i\cdot i=(0,1)\cdot(0,1)=(-1,0)=-1$. Ésta es la razón por la que podemos decir que $i=\sqrt{-1}$. Además, podemos ahora dar una notación adicional de los números complejos, sin más que tomar

$$(a,b)=a+bi,\qquad a,b\in\mathbb{R}.$$

Esta nueva notación es útil pues se comporta bien con la suma y el producto. Para la suma es obvio,

$$(a+bi)+(c+di)=(a,b)+(c,d)=(a+c,b+d)=(a+c)+(b+d)i,$$

y la parte interesante es con el producto, para el cual el desarrollo

$$(a+bi)\cdot(c+di)=ac+adi+bci+bdi^2=(ac-bd)+(ad+bc)i$$

se corresponde con lo que hemos usado para definir el producto.

Queremos comprobar que $\mathbb{C}$ es un cuerpo. El elemento neutro de la suma es $0=0+0i$, y el del producto $1=1+0i$ (es decir, los heredados de $\mathbb{R}$), y todas las propiedades que caracterizan la definición de cuerpo se van comprobando sin dificultad. La única que merece la pena comentar es cómo se llega al cociente entre complejos (en particular, a ver que siempre existe $1/z$ cuando $z\neq 0$, que es lo que exige la definición de cuerpo). Dividir $z=a+bi\in\mathbb{C}$ por un número real $t=t+0i$ no nulo es sencillo: $z/t=a/t+(b/t)i$. Pero, ¿cómo dividir z/w cuando $z,w\in\mathbb{C}$, $w\neq 0$?

Antes de verlo, demos un par de definiciones. Para un complejo $z=a+bi$, su conjugado es, por definición, $\overline{z}=a-bi$, y su módulo $|z|=\sqrt{a^2+b^2}$. Una propiedad importante es que

$$z\cdot\overline{z}=(a+bi)(a-bi)=a^2+b^2=|z|^2.$$

El cociente de dos números complejos, z dividido por w, deberá ser un nuevo complejo u tal que $uw=z$, es decir

$$\frac{z}{w}=u\ \Leftrightarrow\ uw=z.$$

¿Existe el cociente?; ¿cómo hallarlo? Por supuesto, vamos a imponer $w \neq 0$. Multiplicando la igualdad $uw = z$ por $\overline{w}$ queda $uw\overline{w} = z\overline{w}$. Pero $w\overline{w} = |w|^2$ es un número real positivo, y ya sabemos cómo dividir complejos por números reales. Entonces, dividiendo $u \cdot |w|^2 = z\overline{w}$ por $|w|^2$ tenemos $u = z\overline{w}/|w|^2$. Así pues, el cociente de dos complejos es

$$\frac{z}{w} = \frac{z\overline{w}}{|w|^2}.$$

Una vez que sabemos que el cociente existe, no hace falta ser tan cuidadosos, sino que para calcular z/w basta multiplicar el numerador y el denominador por $\overline{w}$.

Así pues, $\mathbb{C}$ es un cuerpo. Pero no es un cuerpo ordenado, pues en $\mathbb{C}$ no hay ninguna relación de orden que sea estable con las operaciones. Una manera de demostrar que tal relación de orden no puede existir es ver que no puede existir el conjunto P de elementos positivos: si existiera, i o $-i$ deberían estar en P, con lo cual su cuadrado debería ser positivo; pero ambos cumplen $i^2 = (-i)^2 = -1$, que no es positivo.

Con $\mathbb{C}$, no sólo se pueden resolver ecuaciones cuadráticas o cúbicas, sino que esto se puede generalizar a polinomios de cualquier grado. Ésta es posiblemente la propiedad más importante de $\mathbb{C}$, que se denomina teorema fundamental del álgebra: para $n \geq 1$, la ecuación

$$a_n z^n + a_{n-1} z^{n-1} + \cdots + a_1 z + a_0 = 0, \qquad a_j \in \mathbb{R},\ a_n \neq 0, \tag{A.4}$$

tiene exactamente n raíces en $\mathbb{C}$ (si contamos cada una tantas veces como su multiplicidad); en realidad, basta afirmar que cualquier polinomio de grado n tiene al menos una raíz, y de allí en seguida se obtiene que deben ser n las raíces (pues si un polinomio $P(z)$ tiene una raíz ζ, entonces $P(z)/(z-\zeta)$ es un nuevo polinomio de grado uno menos[10]).

Fue Albert Girard el primero que sugirió —alrededor de 1620— que las ecuaciones de grado n tenían n raíces, pero su premonición del teorema fundamental del álgebra estaba planteada de forma vaga y sin rigor. Hubo que esperar hasta 1797 para que Carl Friedrich Gauss diera —en su tesis doctoral— la primera prueba esencialmente correcta del teorema fundamental del álgebra. Dar una demostración queda fuera de los objetivos de este libro (de hecho, el libro lo asume a veces como resultado conocido), así que nos contentaremos con comentar lo que seguramente ya conocerá el lector: actualmente existen diversos tipos de demostraciones, que se

[10] Aunque realmente $P(z)/(z-\zeta)$ puede no tener coeficientes reales, sino complejos. Si $P(z)$ tiene coeficientes reales y al dividir se quiere obtener un nuevo polinomio con coeficientes reales, cuando ζ es complejo hay que dividir $P(z)$ por $(z-\zeta)(z-\overline{\zeta})$, donde $\overline{\zeta}$ es el conjugado de ζ, que forzosamente también es raíz de $P(z)$.

pueden encontrar en numerosos textos, y posiblemente la más sencilla es la que utiliza el teorema de Liouville de variable compleja.[11]

Hay una última cosa que queremos añadir: $\mathbb{C}$ se construía porque algunos polinomios con coeficientes reales no tenían raíces en $\mathbb{R}$. ¿Podrá ocurrir que algunos polinomios con coeficientes en $\mathbb{C}$ no tengan raíces en $\mathbb{C}$ y de nuevo se pueda construir un conjunto de números mayor? Esto ya no ocurre, y se expresa diciendo que $\mathbb{C}$ es algebraicamente cerrado. De hecho, el teorema fundamental del álgebra sigue siendo cierto cuando en (A.4) tomamos coeficientes $a_j \in \mathbb{C}$ (y, por ejemplo, la demostración basada en el teorema de Liouville sigue siendo válida).

A.8. Cuaterniones y octoniones

William Rowan Hamilton, que ya nos ha aparecido en relación con la definición de los números complejos, buscaba formas de extender los números complejos (que pueden interpretarse como puntos en un plano) a un número mayor de dimensiones. Pese a sus esfuerzos, no pudo hacerlo para 3 dimensiones, pero en 1843 sí que logró dar una extensión a 4 dimensiones, creando de este modo unos nuevos números denominados cuaterniones (o, también, cuaternas de Hamilton). Tienen aplicaciones en diversos campos de la matemática y de la física, pero nosotros no los usamos en este libro; no obstante, vamos a comentar en qué consisten.

Por definición, el conjunto de los cuaterniones, que se denota con $\mathbb{H}$, se define como

$$\mathbb{H} = \mathbb{R} \times \mathbb{R} \times \mathbb{R} \times \mathbb{R}$$

con la suma

$$(a_1, b_1, c_1, d_1) + (a_2, b_2, c_2, d_2) = (a_1 + a_2, b_1 + b_2, c_1 + c_2, d_1 + d_2)$$

y el producto

$$\begin{aligned}(a_1, b_1, c_1, d_1) \cdot (a_2, b_2, c_2, d_2) \\ = (a_1a_2 - b_1b_2 - c_1c_2 - d_1d_2, a_1b_2 + a_2b_1 + c_1d_2 - c_2d_1, \\ a_1c_2 + a_2c_1 - b_1d_2 + b_2d_1, a_1d_2 + a_2d_1 + b_1c_2 - b_2c_1).\end{aligned}$$

La suma de cuaterniones es conmutativa y asociativa. El producto también es asociativo, y distributivo respecto a la suma, pero no es conmutativo. Así mismo, y tal

[11] El teorema de Liouville afirma que una función $f\colon \mathbb{C} \to \mathbb{C}$ derivable en todo $\mathbb{C}$ y acotada (en módulo) es, forzosamente, constante. Si $P(z)$ es un polinomio que no se anula nunca, entonces $f(z) = 1/P(z)$ es una función acotada y derivable en todo $\mathbb{C}$, con lo cual, por el teorema de Liouville, $P(z)$ sólo puede ser un polinomio constante.

como cabía esperar, el elemento neutro de la suma es el cuaternión $(0,0,0,0)$, y el del producto es $(1,0,0,0)$.

Los cuaterniones se parecen mucho a los números complejos pero, en lugar de tener una unidad imaginaria, tienen tres,

$$i=(0,1,0,0),\quad j=(0,0,1,0),\quad k=(0,0,0,1),$$

que se comportan como sigue:

$$\begin{gathered} i^2=j^2=k^2=-1,\\ ij=k,\quad ji=-k,\qquad jk=i,\quad kj=-i,\qquad ki=j,\quad ik=-j. \end{gathered} \tag{A.5}$$

Si a esto unimos que un número real a se puede identificar con el cuaternión $(a,0,0,0)$, y que un cuaternión se multiplica por un número real t mediante

$$t(a,b,c,d)=(ta,tb,tc,td),$$

de aquí se sigue que los cuaterniones se pueden también representar como

$$(a,b,c,d)=a+bi+cj+dk.$$

Entonces, el producto de cuaterniones se obtiene fácilmente sin más que conocer las reglas (A.5), mucho más fáciles de recordar que la definición de producto. Prácticamente como curiosidad, obsérvese además la posibilidad de escribir $a+bi+cj+dk=(a+bi)+(c+di)j$, lo cual nos permite pensar en los cuaterniones como una «complejificación» adicional hecha sobre los números complejos; de hecho, podíamos haber tomado desde el principio $\mathbb{H}=\mathbb{C}\times\mathbb{C}$ con ciertas operaciones suma y producto definidas en $\mathbb{C}\times\mathbb{C}$.

Por definición, el conjugado de un cuaternión $q=a+bi+cj+dk$ es $\overline{q}=a-bi-cj-dk$, y su norma es

$$\|q\|=\sqrt{q\overline{q}}=\sqrt{\overline{q}q}=\sqrt{a^2+b^2+c^2+d^2}.$$

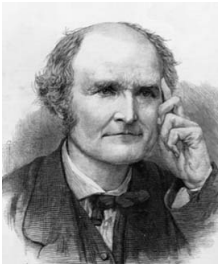
Entonces, de manera similar a como hacíamos en $\mathbb{C}$ se puede probar que el inverso multiplicativo de q es $q^{-1}=\overline{q}/\|q\|^2$ (válido siempre que $q\neq 0$).

En consecuencia, y una vez visto cómo calcular inversos para el producto, merece la pena resaltar que a $\mathbb{H}$ sólo le falla una propiedad para tener estructura de cuerpo: la conmutatividad del producto.

Existe una manera alternativa de definir los cuaterniones, que consiste en utilizar matrices reales 4×4. Un cuaternión $a+bi+cj+dk$ se puede representar mediante



Sir William Rowan Hamilton (1805–1865)



Arthur Cayley (1821–1895)

la matriz

$$\begin{pmatrix} a & b & c & d \\ -b & a & -d & c \\ -c & d & a & -b \\ -d & -c & b & a \end{pmatrix} \tag{A.6}$$

con $a, b, c, d \in \mathbb{R}$.[12] Por supuesto, la suma y el producto de cuaterniones se convierten en suma y producto de esas matrices. Con esta notación, el conjugado de un cuaternión se corresponde con la matriz traspuesta. Además, el determinante de la matriz es $a^2 + b^2 + c^2 + d^2$ (el cuadrado de la norma); este determinante es distinto de 0 siempre que la matriz sea no nula, lo cual garantiza que la matriz tiene inversa, que será a su vez el inverso multiplicativo del cuaternión. Por supuesto, si se quiere seguir este procedimiento de definición de los cuaterniones hay que comprobar que la suma y producto de matrices de la forma (A.6) vuelve a dar una matriz de esa forma; y que, si una matriz tiene la forma (A.6), también la tiene su inversa.

[12] La definición de los cuaterniones por medio de matrices reales 4×4 no es única. Por ejemplo, también se puede usar

$$\begin{pmatrix} a & -b & d & -c \\ b & a & -c & -d \\ -d & c & a & -b \\ c & d & b & a \end{pmatrix}$$

u otras variantes (por supuesto, una vez escogida una definición matricial, hay que seguirla para todos los cuaterniones que se vayan a usar). Además, esta idea de definir los cuaterniones a partir de matrices también sirve para definir los números complejos. De este modo, el complejo $a + bi$ se podría definir como la matriz real $\left(\begin{smallmatrix} a & b \\ -b & a \end{smallmatrix}\right)$; esta asociación entre complejos y matrices 2×2 tampoco es única, y se podría haber definido $a + bi$ mediante $\left(\begin{smallmatrix} a & -b \\ b & a \end{smallmatrix}\right)$.

Lo mismo se puede hacer con matrices 2×2 pero de coeficientes complejos. De esta manera, el cuaternión $a+bi+cj+dk$ se representa mediante la matriz compleja

$$\begin{pmatrix} a+bi & c+di \\ -c+di & a-bi \end{pmatrix},$$

y las operaciones con cuaterniones se reducen de nuevo a operaciones entre matrices.

Impulsados por el éxito de Hamilton, los matemáticos intentaron encontrar extensiones a otras dimensiones. Pronto aparecieron los octoniones, que son una extensión no asociativa de los cuaterniones, y que se pueden describir como 8-tuplas de números reales en los que se define una operación suma y una operación producto. Fueron descubiertos por John T. Graves en 1843, e independientemente por Arthur Cayley, quien lo publicó por primera vez en 1845. Son llamados, a veces, octavas de Cayley o números de Cayley.

El producto de octoniones no es conmutativo ni asociativo. A causa de esta no asociatividad, los octoniones, a diferencia de los cuaterniones, no admiten una representación matricial. Pero sí que cada octonión no nulo tiene un inverso para el producto y, por tanto, se puede dividir.

En cierto sentido, ya no se puede seguir extendiendo los conjuntos de números. Está demostrado que las únicas extensiones de $\mathbb{R}$ con un producto en el que siempre se puede dividir (por elementos no nulos, claro) son los complejos, los cuaterniones o los octoniones. No hay extensiones de $\mathbb{R}$ con división en dimensiones distintas de 2, 4 o 8.

Esto tiene, además, implicaciones topológicas, pues la existencia de tales extensiones confiere a algunos espacios topológicos ciertas propiedades especiales. El ejemplo más típico es que, si para cada $n \geq 2$ tomamos

$$S^{n-1} = \{(x_1, \ldots, x_n) \in \mathbb{R}^n : x_1^2 + \cdots + x_n^2 = 1\}$$

la superficie esférica en $\mathbb{R}^n$, las únicas que son *paralelizables* —un concepto que no explicaremos— son S^1, S^3 y S^7. La demostración de tal resultado se basa en la existencia de complejos, cuaterniones y octoniones en $\mathbb{R}^2$, $\mathbb{R}^4$ y $\mathbb{R}^8$ como únicas extensiones de los reales.

A.9. Y aún hay más

Los conjuntos de números que hemos ido construyendo están encajados unos en otros en una cadena de inclusiones:

$$\mathbb{N}_0 \subset \mathbb{Z} \subset \mathbb{Q} \subset \mathbb{R} \subset \mathbb{C} \subset \mathbb{H} \subset \{\text{octoniones}\}.$$

Pero se pueden definir otros conjuntos de números entre $\mathbb{Z}$ y $\mathbb{C}$ que ya no responden a esta cadena de inclusiones.

Un ejemplo son los números algebraicos, que son, por definición, las raíces —reales o complejas, según prefiramos exigir— de polinomios con coeficientes en $\mathbb{Z}$. Los algebraicos reales están entre $\mathbb{Q}$ y $\mathbb{R}$, y los algebraicos complejos entre $\mathbb{Q}$ y $\mathbb{C}$; de hecho, tanto unos como otros tienen estructura de cuerpo. Pero no es a eso a lo que nos referimos ahora.

En realidad, estamos pensando ahora en extensiones de tipo

$$\mathbb{Z}[\xi] = \{a + b\xi : a, b \in \mathbb{Z}\}$$

con ξ un número real o complejo tal que $\xi^2 \in \mathbb{Z}$. Por supuesto, la suma y el producto en $\mathbb{Z}[\xi]$ van a ser los ordinarios de $\mathbb{R}$ o $\mathbb{C}$, lo cual garantiza las propiedades asociativa, conmutativa y distributiva, y la existencia de elementos neutros, sin necesidad de comprobar nada. Alternativamente, también se puede hacer una definición formal

$$\mathbb{Z}[\xi] = \mathbb{Z} \times \mathbb{Z}$$

con la suma

$$(a, b) + (c, d) = (a + c, b + d)$$

y el producto

$$(a, b) \cdot (c, d) = (ac + bd\xi^2, ad + bc)$$

(que tiene sentido en $\mathbb{Z}[\xi]$ pues hemos exigido que $\xi^2 \in \mathbb{Z}$), aunque ahora habría que comprobar las propiedades.

Uno de los ejemplos más importantes es el de los enteros gaussianos

$$\mathbb{Z}[i] = \{a + bi : a, b \in \mathbb{Z}\}$$

que, como su nombre indica, fueron introducidos por el propio Gauss (lo hizo en 1832, y los utilizó en sus estudios sobre la ley de reciprocidad bicuadrática). En $\mathbb{C}$, el inverso de un entero gaussiano no nulo $a + bi$ es

$$\frac{1}{a + bi} = \frac{a - bi}{(a + bi)(a - bi)} = \frac{a - bi}{a^2 + b^2},$$

que sólo es un entero gaussiano cuando $a^2 + b^2 = 1$. Así pues, los únicos enteros gaussianos que tienen inverso (en $\mathbb{Z}[i]$) son 1, -1, i y $-i$.

De manera similar se puede definir

$$\mathbb{Z}[\sqrt{d}] = \{a + b\sqrt{d} : a, b \in \mathbb{Z}\}$$

con $d \in \mathbb{N}$ libre de cuadrados, o

$$\mathbb{Z}[\sqrt{-d}\,] = \{a + b\sqrt{-d} : a, b \in \mathbb{Z}\},$$

también con $d \in \mathbb{N}$ libre de cuadrados. En el primer caso, el inverso en $\mathbb{R}$ de $a + b\sqrt{d} \neq 0$ es

$$\frac{1}{a + b\sqrt{d}} = \frac{a - b\sqrt{d}}{(a + b\sqrt{d})(a - b\sqrt{d})} = \frac{a - b\sqrt{d}}{a^2 - db^2},$$

que está en $\mathbb{Z}[\sqrt{d}\,]$ cuando $a^2 - db^2 = 1$. Y, en el segundo, la condición que se obtiene para que el inverso de $a + b\sqrt{-d}$ en $\mathbb{C}$ sea un inverso en $\mathbb{Z}[\sqrt{-d}\,]$ es que $a^2 + db^2 = 1$.

Con la suma y el producto, este tipo de extensiones $\mathbb{Z}[\sqrt{d}\,]$ o $\mathbb{Z}[\sqrt{-d}\,]$ da lugar a anillos. Quizás resulta sorprendente que, aunque la definición de $\mathbb{Z}[\sqrt{d}\,]$ es la misma independientemente del d libre de cuadrados que se tome, las propiedades del correspondiente $\mathbb{Z}[\sqrt{d}\,]$ pueden variar bastante; y lo mismo ocurre con $\mathbb{Z}[\sqrt{-d}\,]$. De hecho, esto dio lugar, antes de que los matemáticos se percataran de ello, a algunas demostraciones erróneas de resultados importantes (el principal problema es que no siempre la factorización de un número como producto de primos es única). Esta riqueza de propiedades da como resultado los diversos tipos de anillos que mencionábamos en la sección A.2, pero que nosotros ya no abordaremos porque requeriría adentrarnos demasiado en el terreno del álgebra.

Casi lo mismo se puede hacer usando, en lugar de un complejo ξ que cumpla $\xi^2 \in \mathbb{Z}$, uno que cumpla $\xi^n \in \mathbb{Z}$ para algún exponente n fijo. Esta vez, para tener una multiplicación, habría que tomar el conjunto

$$\{a_1 + a_2\xi + a_2\xi^2 + \cdots + a_n\xi^{n-1} : a_1, a_2, \ldots, a_n \in \mathbb{Z}\}.$$

Por ejemplo, con $n = 3$, usando $\xi = \sqrt[3]{2}$ tendríamos

$$\{a + b\sqrt[3]{2} + c\sqrt[3]{4} : a, b, c \in \mathbb{Z}\}.$$

Y tomando $\xi = e^{2\pi i/3} = \cos(2\pi/3) + i\,\mathrm{sen}(2\pi/3)$, con lo cual $\xi^2 = e^{4\pi i/3} = \cos(4\pi/3) + i\,\mathrm{sen}(4\pi/3)$ y $\xi^3 = e^{2\pi i} = 1$, estaríamos extendiendo $\mathbb{Z}$ adjuntándole las raíces cúbicas de 1.

La utilidad de todas estas extensiones es inmensa, tanto en las matemáticas en sí mismas como en sus muy diversas aplicaciones prácticas. Pero en algún momento hay que detenerse y es lo que hacemos en este momento, así que este libro no se ha aventurado en nada de lo relacionado con estos campos, sino que nos hemos contentado con mirar desde fuera.

Bibliografía

[1] M. Agrawal, N. Kayal y N. Saxena, PRIMES is in P, *Annals of Math. (2)* **160** (2004), 781–793. Citado en pág. 146.

[2] M. Aigner y G. M. Ziegler, *EL LIBRO de las demostraciones*, Nivola, Tres Cantos (Madrid), 2005. Citado en págs. 3, 26 y 348.

[3] E. Aparicio, *Teoría de los números*, Servicio Editorial de la Universidad del País Vasco, Bilbao, 1993. Citado en págs. 89, 169, 280, 400, 452, 507, 556 y 605.

[4] R. Apéry, Irrationalité de $\zeta(2)$ et $\zeta(3)$, *Astérisque* **61** (1979), 11–13. Citado en pág. 383.

[5] T. M. Apostol, *Introducción a la teoría analítica de números*, Reverté, 1980. Citado en págs. 89, 110, 123, 141, 169, 177, 182, 221, 400, 419, 437, 452, 507 y 556.

[6] J. Arias de Reyna y J. van de Lune, Un primer encuentro con la hipótesis de Riemann y su comprobación numérica, *La Gaceta de la RSME* **13** (2010), 109–133. Citado en págs. 610, 611 y 615.

[7] Arquímedes, *Obras escogidas*, 2 vols., RSME, ICM2006 y Patrimonio Nacional, Sevilla, 2006. Editor: A. J. Durán. (Edición facsímil del manuscrito conservado en la Biblioteca del Monasterio de El Escorial, traducida, anotada y con estudios complementarios.) Citado en pág. 300.

[8] A. Baker, *Breve introducción a la teoría de números*, Alianza Universidad, 1986. Citado en págs. 240 y 334.

[9] A. Baker, *Transcendental number theory*, 2.ª ed., Cambridge University Press, 1990. Citado en págs. 240, 251, 334 y 382.

[10] M. Balazard, Un siècle et demi de recherches sur l'hypothèse de Riemann, *Gaz. Math.* **126** (2010), 7–24. Citado en pág. 611.

[11] P. H. Bateman y H. G. Diamond, A hundred years of prime numbers, *Amer. Math. Monthly* **103** (1996), 729–741. Citado en pág. 556.

[12] P. H. Bateman y H. G. Diamond, *Analytic number theory: An introductory course*, World Scientific, 2004. Citado en págs. 400, 452, 463, 556 y 605.

[13] M. Benito, W. Creyaufmüller, J. L. Varona y P. Zimmermann, Aliquot sequence 3630 ends after reaching 100 digits, *Experiment. Math.* **11** (2002), 201–206. Citado en págs. 6 y 410.

[14] M. Benito y J. J. Escribano, *Sucesiones de Brocot*, Ochoa, Logroño, 1998. Citado en pág. 258.

[15] M. Benito y J. J. Escribano, An easy proof of Hurwitz theorem, *Amer. Math. Monthly* **109** (2002), 916–918. Citado en pág. 268.

[16] M. Benito, L. M. Navas y J. L. Varona, Möbius inversion formulas for flows of arithmetic semigroups, *J. Number Theory* **128** (2008), 390–412. Citado en pág. 439.

[17] M. Benito, L. M. Navas y J. L. Varona, Möbius inversion from the point of view of arithmetical semigroup flows, en *Proceedings of the «Segundas Jornadas de Teoría de Números»* (Madrid, 2007), Bibl. Rev. Mat. Iberoamericana, Madrid, 2008, p. 63–81. Citado en pág. 439.

[18] M. Benito y J. L. Varona, Sucesiones alicuatorias, *La Gaceta de la RSME* **2** (1999), 357–365. Citado en pág. 410.

[19] P. Berrizbeitia, Pruebas determinísticas de primalidad, *La Gaceta de la RSME* **4** (2001), 447–456. Citado en pág. 217.

[20] P. Berrizbeitia, *Algoritmos deterministas de primalidad*, XVII Escuela Venezolana de Matemáticas, Asociación Matemática Venezolana, Centro de Estudios Avanzados, IVIC, Caracas, 2004. Citado en págs. 201, 212 y 217.

[21] F. Beukers, A note on the irrationality of $\zeta(2)$ and $\zeta(3)$, *Bull. London Math. Soc.* **11** (1979), 268–272. Citado en pág. 383.

[22] E. B. Burger, *Exploring the number jungle: A journey into diophantine analysis*, American Mathematical Society, Providence, RI, 2000. Citado en págs. 240, 272, 312 y 359.

[23] E. B. Burger y R. Tubbs, *Making transcendence transparent: An intuitive approach to classical transcendental number theory*, Springer, 2004. Citado en págs. 334, 337 y 383.

[24] G. V. Chudnovsky, *Contributions to the theory of transcendental numbers*, American Mathematical Society, Providence, RI, 1984. Citado en págs. 334 y 383.

[25] Ó. Ciaurri, L. M. Navas, F. J. Ruiz y J. L. Varona, A simple computation of $\zeta(2k)$, *Amer. Math. Monthly*, **122** (2015), 444–451. Citado en pág. 387.

[26] J. Cilleruelo y A. Córdoba, *La teoría de los números*, Biblioteca Mondadori, Madrid, 1992. Citado en págs. 89, 280, 318, 341 y 556.

[27] COMAP, *Las matemáticas en la vida cotidiana*, 3.ª ed., Addison-Wesley y Universidad Autónoma de Madrid, Madrid, 1999. Citado en págs. 90 y 153.

[28] A. Córdoba, Disquisitio numerorum, *La Gaceta de la RSME* **4** (2001), 249–260. Citado en pág. 383.

[29] R. Crandall, K. Dilcher y C. Pomerance, A search for Wieferich and Wilson primes, *Math. Comp.* **66** (1997), 433–449. Citado en págs. 106 y 125.

[30] R. Crandall y C. Pomerance, *Prime numbers: A computational perspective*, 2.ª ed., Springer, 2005. Citado en págs. 89, 145, 146 y 201.

[31] G. P. Curbera, *Mathematicians of the world, unite!: The International Congress of Mathematicians — a human endeavor*, A K Peters, Wellesley, MA, 2009. Citado en págs. 60 y 113.

[32] T. W. Cusick y M. E. Flahive, *The Markoff and Lagrange spectra*, American Mathematical Society, Providence, RI, 1989. Citado en pág. 272.

[33] J. M. De Koninck y F. Luca, *Analytic number theory: Exploring the anatomy of integers*, American Mathematical Society, Providence, RI, 2012. Citado en pág. 556.

[34] L. E. Dickson, *History of the theory of numbers*, 3 vols., Carnegie Institute of Washington, 1919, 1920 y 1923. Reimpresión: Chelsea, 1971. Citado en pág. 7.

[35] Diofanto de Alejandría, *La Aritmética y el libro Sobre los números poligonales*, 2 vols., Nivola (Colección Epistéme n.os 6 y 7), Tres Cantos (Madrid), 2008. Versión en castellano, introducción, notas y apéndices de M. Benito, E. Fernández y M. Sánchez. Citado en pág. 58.

[36] M. du Sautoy, *La música de los números primos: El enigma de un problema matemático abierto*, Acantilado, Barcelona, 2007. Citado en pág. 557.

[37] U. Dudley, History of a formula for primes, *Amer. Math. Monthly* **76** (1969), 23–28. Citado en pág. 529.

[38] W. Dunham, *Viaje a través de los genios: Biografías y teoremas de los grandes matemáticos*, Pirámide, 2002. Citado en págs. 336, 386 y 508.

[39] A. J. Durán, G. Ifrah y A. Manguel, *Vida de los números*, T Ediciones, 2006. Citado en pág. 32.

[40] H. M. Edwards, *Riemann's Zeta Function*, Academic Press, 1974. Reimpresión: Dover, 2001. Citado en págs. 556, 615, 621 y 622.

[41] P. Erdős y J. Surányi, *Topics in the theory of numbers*, Springer, 2003. Citado en págs. 89, 169, 240, 400, 425 y 507.

[42] Euclides, *Elementos*, edición de los 13 libros en 3 vols., traducida y anotada por M. L. Puertas Castaños, y con introducción de L. Vega, Gredos, Madrid, 1991 a 1996. Citado en págs. 3 y 16.

[43] L. Euler, *Introducción al análisis de los infinitos*, 2 vols., Sociedad Thales y RSME, Sevilla, 2000. Editores: A. J. Durán y F. J. Pérez Fernández. (Edición facsímil del original de 1748, traducida, anotada y con estudios complementarios.) Citado en págs. 281, 319, 344, 508 y 557.

[44] J. L. Gómez Pardo, Aspectos computacionales de los números primos (I) y (II), *La Gaceta de la RSME* **4** (2001), 649–673, y **5** (2002), 197–227. Citado en pág. 89.

[45] J. L. Gómez Pardo, Criptografía y curvas elípticas, *La Gaceta de la RSME* **5** (2002), 737–777. Citado en pág. 169.

[46] J. L. Gómez Pardo, Un paseo alrededor de la teoría de conjuntos, *La Gaceta de la RSME* **11** (2008), 45–96. Citado en pág. 357.

[47] R. L. Graham, D. E. Knuth y O. Patashnik, *Concrete mathematics: A foundation for computer science*, 2.ª ed., Addison-Wesley, 1994. Citado en págs. 51, 258, 283, 437, 453 y 466.

[48] J. Guillera, Historia de las fórmulas y algoritmos para π, *La Gaceta de la RSME* **10** (2007), 159–178. Citado en pág. 301.

[49] R. K. Guy, *Unsolved problems in number theory*, 3.ª ed., Springer, 2004. Citado en págs. 4, 406 y 407.

[50] G. H. Hardy y E. M. Wright, *An introduction to the theory of numbers*, 5.ª ed., Oxford University Press, 1979. Citado en págs. 89, 346, 348, 400, 452, 459, 462, 507, 535, 556 y 593.

[51] J. Havil, *Gamma: Exploring Euler's constant*, Princeton University Press, 2003. Citado en págs. 384 y 469.

[52] H. Helfgott, La conjetura débil de Goldbach, *La Gaceta de la RSME* **16** (2013), 709–726. Citado en pág. 528.

[53] R. Hill, *A first course in coding theory*, Oxford University Press, 1986. Citado en págs. 90 y 162.

[54] Hua[1] Loo Keng, *Introduction to number theory*, Springer-Verlag, 1982. Citado en págs. 240, 258, 268, 270, 280, 312, 313 y 507.

[55] A. Hurwitz y N. Kritikos, *Lectures on number theory*, Springer-Verlag, 1986. Citado en págs. 89, 169, 280 y 313.

[56] G. Ifrah, *Historia universal de las cifras*, 6.ª ed., Espasa, Madrid, 2008. Citado en pág. 32.

[57] W. B. Jones y W. J. Thron, *Continued fractions: Analytic theory and applications*, Addison-Wesley, 1980. Citado en pág. 290.

[58] A. Ya. Khinchin, *Continued fractions*, University of Chicago Press, 1964. Reimpresión: Dover, 1997. Citado en págs. 254 y 280.

[59] D. E. Knuth, *The art of computer programming, Vol. 2: Seminumerical algorithms*, 3.ª ed., Addison-Wesley, 1998. Citado en págs. 107 y 109.

[60] N. Koblitz, *A course in number theory and cryptography*, 2.ª ed., Springer-Verlag, 1994. Citado en págs. 90, 128 y 137.

[61] M. Křížek, F. Luca y L. Somer, *17 lectures on Fermat numbers: From number theory to geometry*, Springer-Verlag, 2001. Citado en pág. 202.

[62] M. Laczkovich, On Lambert's proof of the irrationality of π, *Amer. Math. Monthly* **104** (1997), 439–443. Citado en pág. 345.

[63] F. Lemmermeyer, *Reciprocity laws: From Euler to Eisenstein*, Springer-Verlag, 2000. Citado en págs. 183 y 194.

[64] P. Levrie, A short derivation of Lord Brouncker's continued fraction for π, *Math. Intelligencer* **29** (2007), n.º 2, 8–9. Citado en pág. 301.

[65] A. J. Menezes, P. C. van Oorschot y S. A. Vanstone, *Handbook of applied cryptography*, CRC Press, 1997. Citado en págs. 90, 147, 156 y 233.

[66] S. J. Miller y R. Takloo-Bighash, *An invitation to modern number theory*, Princeton University Press, 2006. Citado en págs. 252, 280 y 359.

[67] C. Munuera, Códigos correctores de errores (o cuántas preguntas son necesarias para conocer un número), *La Gaceta de la RSME* **6** (2003), 714–731. Citado en pág. 162.

[68] C. Munuera y J. G. Tena, *Codificación de la información*, Serie Manuales y Textos Universitarios **25**, Universidad de Valladolid, 1997. Citado en págs. 90 y 162.

[69] W. Narkiewicz, *The development of prime number theory: From Euclid to Hardy and Littlewood*, Springer, 2000. Citado en págs. 26, 556 y 605.

[1] Quizás esta explicación satisfaga la curiosidad del lector sobre la aparentemente extraña ordenación alfabética usada con este autor: en chino, el apellido precede al nombre.

[70] J. A. Nathan, The irrationality of e^x for nonzero rational x, *Amer. Math. Monthly* **105** (1998), 762–763. Citado en págs. 343 y 345.

[71] M. B. Nathanson, *Elementary methods in number theory*, Springer, 2000. Citado en págs. 89, 169, 289, 400, 426, 452, 507 y 556.

[72] L. M. Navas, F. J. Ruiz y J. L. Varona, The Möbius inversion formula for Fourier series applied to Bernoulli and Euler polynomials, *J. Approx. Theory* **163** (2011), 22–40. Citado en pág. 437.

[73] D. J. Newman, Simple analytic proof of the prime number theorem, *Amer. Math. Monthly* **87** (1980), 693–696. Citado en págs. 556 y 581.

[74] D. J. Newman, *Analytic number theory*, Springer, 2000. Citado en pág. 556.

[75] E. M. Nikishin y V. N. Sorokin, *Rational approximations and orthogonality*, American Mathematical Society, Providence, RI, 1991. Citado en págs. 240, 252, 271, 280, 290, 313 y 359.

[76] I. Niven, A simple proof that π is irrational, *Bull. Amer. Math. Soc.* **53** (1947), 509. Citado en pág. 346.

[77] I. Niven, *Irrational numbers*, Carus Monographs n.º 11, The Mathematical Association of America, 1956. Reimpreso en 2005. Citado en págs. 334 y 348.

[78] I. Niven, H. S. Zuckerman y H. L. Montgomery, *An introduction to the theory of numbers*, 5.ª ed., Wiley, 1991. Citado en págs. 89, 169, 240, 258, 268, 280, 289, 318, 334, 341, 346, 366 y 394.

[79] A. V. Prasad, Note on a theorem of Hurwitz, *J. London Math. Soc.* **23** (1948), 169–171. Citado en pág. 274.

[80] J. L. Ramírez Alfonsín, El problema diofántico de Frobenius, *La Gaceta de la RSME* **16** (2013), 117–129. Citado en pág. 67.

[81] P. Ribenboim, *My numbers, my friends: Popular lectures on number theory*, Springer, 2000. Citado en págs. 272 y 334.

[82] P. Ribenboim, *The little book of bigger primes*, 2.ª ed., Springer, 2004. Citado en págs. 26 y 507.

[83] H. Riesel, *Prime numbers and computer methods for factorization*, 2.ª ed., Birkhäuser, 1994. Citado en págs. 145, 201 y 613.

[84] H. Riesel y G. Göhl, Some calculations related to Riemann's prime number formula, *Math. Comp.* **24** (1970), 969–983. Citado en págs. 628 y 629.

[85] J. Rifà y L. Huguet, *Comunicación digital: Teoría matemática de la información; Codificación algebraica; Criptología*, Masson, 1991. Citado en págs. 90 y 162.

[86] Ø. J. Rødseth, A note on primality tests for $N = h \cdot 2^n - 1$, *BIT* **34** (1994), 451–454. Citado en pág. 217.

[87] H. E. Rose, *A course in number theory*, 2.ª ed., Oxford University Press, 1996. Citado en págs. 89, 169, 400, 452, 507 y 556.

[88] K. H. Rosen, *Elementary number theory and its applications*, 5.ª ed., Pearson/Addison-Wesley, 2005. Citado en págs. 89, 90, 128, 133, 139, 141, 145, 169, 177, 191, 192, 198, 221, 280 y 318.

[89] K. F. Roth, Rational approximation of algebraic numbers, *Mathematika* **2** (1955), 1–20; Corrigendum, *Mathematika* **2** (1955), 168. Citado en págs. 252 y 359.

[90] I. Ruzsa, Erdős and the integers, *J. Number Theory* **79** (1999), 115–163. Citado en págs. 425 y 548.

[91] A. B. Shidlovskii, *Aproximaciones diofánticas y números trascendentes*, Servicio Editorial de la Universidad del País Vasco, Bilbao, 1989. Citado en págs. 334, 379 y 382.

[92] A. B. Shidlovskii, *Transcendental numbers*, Walter de Gruyter, 1989. Citado en págs. 334, 379 y 382.

[93] C. L. Siegel, *Transcendental numbers*, Princeton University Press, 1949. Citado en pág. 382.

[94] W. Sierpiński, *A selection of problems in the theory of numbers*, Pergamon Press, 1964. Citado en págs. 4 y 111.

[95] W. Sierpiński, *Elementary theory of numbers*, North-Holland y Polish Scientific Publishers, 1988. Citado en págs. 89, 111, 280 y 507.

[96] S. Singh, *El enigma de Fermat*, Planeta, Barcelona, 1998. Citado en pág. 76.

[97] S. Singh, *Los códigos secretos*, Debate, Madrid, 2000. Citado en págs. 90, 136 y 147.

[98] A. Srinivasan, ¿Debemos intentar resolver la conjetura de Markoff?, *La Gaceta de la RSME* **16** (2013), 313–330. Citado en págs. 79 y 272.

[99] D. R. Stinson, *Cryptography: Theory and practice*, 3.ª ed., Chapman & Hall/CRC, 2006. Citado en págs. 90, 137 y 145.

[100] G. Tenenbaum y M. Mendès France, *The prime numbers and their distribution*, American Mathematical Society, Providence, RI, 2000. Citado en pág. 556.

[101] J. H. van Lint, *Introduction to coding theory*, 3.ª ed., Springer-Verlag, 1998. Citado en págs. 90 y 162.

[102] J. L. Varona, Differentiability of a pathological function, diophantine approximation, and a reformulation of the Thue-Siegel-Roth theorem, *Austral. Math. Soc. Gaz.* **36** (2009), 353–361. Citado en pág. 256.

[103] N. N. Vorobiev, *Fibonacci numbers*, Birkhäuser, 2002. Citado en pág. 281.

[104] D. Zagier, Newman's short proof of the prime number theorem, *Amer. Math. Monthly* **104** (1997), 705–708. Citado en págs. 556 y 581.

[105] A. V. Zhúkov, *El omnipresente número* π, Editorial URSS, Moscú, 2005. Citado en pág. 301.

Índice alfabético

Los números de página en *itálica* se refieren a los retratos de matemáticos incluidos en el libro.

ábaco, 34
Abel, Niels Henrik, 113, *336*, 336, *462*, 463, 491, 494, 532, 566, 578, 579, 587
abundante, *véase* número abundante
adición, *véase* suma
aditiva, *véase* función aditiva
Adleman, Leonard, 4, 142, *143*, 145
Agrawal, Manindra, 146, 693
Aigner, Martin, 693
AKS, *véase* test de primalidad AKS
al-Basri, *véase* al-Haytham
al-Haytham, Ibn, 115, *123*, 123
al-Jayyam, *véase* Khayyam
al-Jwarismi, *véase* al-Khwarizmi
al-Karaji, 56
al-Karkhi, *véase* al-Karaji
al-Khayyam, *véase* Khayyam
al-Khwarizmi, Muhammad ibn Musa, *33*, 34
al-Misri, *véase* al-Haytham
álgebra, 411, 641
algebraicamente cerrado, 686
algebraico, *véase* número algebraico
algoritmo de cuadrados iterados, *véase* cálculo de potencias mediante cuadrados sucesivos
algoritmo de Euclides, 20, 21, 23, 28–32, 92, 94, 96–98, 100–102, 120, 126, 138, 144, 175, 191, 232, 292
algoritmo de la división, 19, 20, 292
algoritmo de Shor, 147
Alhacén, *véase* al-Haytham
Alhazen, *véase* al-Haytham
alícuota, *véase* parte alícuota
American Standard Code for Information Interchange, *véase* ASCII
amigos, *véase* números amigos
Andrica, Dorin, 526, 551
anillo, 17, 27, 76, 93, 194, 411, 641, 642, 654, 657, 691
antisimétrica, *véase* propiedad antisimétrica
Aparicio, Emiliano, *525*, 525, 693
Apéry, Roger, *383*, 383, 693
aplicación biyectiva, 640
aplicación inyectiva, 640, 672
aplicación sobreyectiva, *véase* aplicación suprayectiva
aplicación suprayectiva, 640, 672
Apostol, Tom M., 693
aproximación diofántica, 2, 240, 257
aproximación óptima de primera especie, 262, 263, 276, 304

aproximación óptima de segunda especie, 263, 276, 304
aproximación por racionales, 78, 238–240, 245–248, 250–252, 254, 256, 258, 263–268, 273, 280, 302–304, 324, 326, 329, 334
aproximación racional, *véase* aproximación por racionales
aproximación simultánea, 277
aproximantes de Padé, 290, 361, 383
árabes, 34, 36, 123, 660
árboles de Stern-Brocot, 258
Argand, Jean Robert, 335, 683
argumento, 560
argumento principal, 560
Arias de Reyna, Juan, 693
aritmética binaria, 36, 39
aritmética del reloj, 88
Arquímedes de Siracusa, 300, *301*, 693
arquitectura von Neumann, 649
Artin, Emil, *194*, 194, 222
Aryabhata, *115*, 115, 280
ASCII, 158, 160
asociativa, *véase* propiedad asociativa
axioma de inducción, 17, 644, 646, 647
axioma del supremo, 661
axiomas de Peano, 18, 643–648, 650
axiomas de Zermelo-Fraenkel, 357, 636, 649

Babbage, Charles, 388, *389*
babilonios, 32, 68, 69
Bachet de Méziriac, Claude Gaspar, 58, *59*
Bachmann, Paul Gustav Heinrich, 453, *455*
Baker, Alan, *257*, 257, 382, 693
Balazard, Michel, 693
banda crítica, 468, 566, 569, 606, 610, 617
Barinaga, José, *514*, 514
base de un sistema de numeración, 32, 37, 102, 151, 155, 158, 163, 195, 249, 337, 350
Bateman, Paul T., 694
Beal, Daniel Andrew, 76, 77
Bell, Eric Temple, 410, *412*
Benito, Manuel, 694
Bernoulli, Jacob, *384*, 385–389, 391, 392, 397, 437, 465, 466, 508, 698
Bernoulli, Johann, *280*, 281, 386, 508, 682
Berrizbeitia, Pedro, 212, 694
Bertrand, Joseph, 516, 521, 523, 524, *525*, 525, 528–530
Beukers, Frits, 383, 694
Bézout, Étienne, *21*, 21, 23, 95, 120, 277
Bhaskara I, 115
Bhaskara II, 78, 115, 393
Binet, Jacques, *282*, 282, 285, 286, 310
binomio de Newton, 48, 49, 51–53, 103, 163, 401
bit, 158–162
bit de paridad, 158, 159
biyectiva, *véase* aplicación biyectiva
Bois-Reymond, *véase* du Bois-Reymond
Bolzano, Bernard, *670*, 670, 671
Bombelli, Rafael, 280, *313*, 313, 314, 681, 682
Borwein, Peter, 383, 384
Brahmagupta, 33, 62, 78, 115, 642, 650, *651*
Breusch, Robert Hermann, 525
Brocard, Henri, 80, 527, 551
Brocot, Achille, 239, 258–266, 268, 276, 280, 302, 304, 308, 694
Brouncker, Lord William, *280*, 281, 300, 697
Burger, Edward B., 694

cálculo de potencias mediante cuadrados sucesivos, 102, 126, 144, 178, 195, 196, 205, 208, 214, 225, 228
camino, 561
campo, 640
canon áureo, *véase* razón áurea
Cantor, Georg, 60, 334, 351, *352*, 352, 353, 355, 356, 358–360, *643*, 644
capicúa, 318
carácter cuadrático, 178
caracteres de Dirichlet, 110

Cardano, Girolamo, *335*, 336, 364–366, 394, 681
cardinal, 91, 98, 356
cardinal del continuo, 356
Carmichael, Robert Daniel, 127–129, *130*, 133, 198
Catalan, Eugène Charles, 79, *80*, *409*, 409, 410
Cataldi, Pietro Antonio, 213, 281
Cauchy, Augustin Louis, 283, 286, 557, *558*, 558, 561–563, 581, 660, *670*, 670, 678–681
Cayley, Arthur, *688*, 689
ceros de la zeta de Riemann, 2, 468, 555, 569, 570, 572, 574, 575, 580, 604–606, 608, 610, 611, 615, 617, 618, 622–626, 628, 629
César, *véase* cifra del César
César, Cayo Julio, 134, 136, 137, 227
Cesàro, Ernesto, 439, *442*, 443, 452, 578, 579
Champernowne, David Gawen, 337, *339*
Chebyshev, Pafnuty Lvovich, *339*, 339, 340, 372, 394, 395, 436, 448, 449, 459, 481, 488–490, *512*, 512, 518, 523, 541, 551, 576, 592, 605, 631
Chia Hsien, *véase* Jia Xian
Chin Chiu-Shao, *véase* Qin Jiushao
chinos, 54, 56, 88, 115, 300, 660
Chu Shih-Chieh, 54, 56, 57
Chudnovsky, Gregory Volfovich, 382, 383, 694
Ciaurri, Óscar, 695
cifra del César, 134, 136, 137, 227
Cilleruelo, Javier, 695
clase de equivalencia, 91, 638
clase de restos, 91
Clausen, Thomas, *207*, 208, 209
clave privada, 140, 141, 143, 147, 148, 228, 229, 231–233, 236
clave pública, 134, 139, 141, 143, 148, 169, 219, 225, 228, 229, 231, 233, 236
claves asimétricas, 142, 145, 228
cociente, 19
cociente de Fermat, 105
cociente de Wilson, 125
cociente incompleto, 289, 297, 319
Cocks, Clifford, 145
Códice Vigilanus, 34
código de Golay, 162
código de Hamming, 161, 166
códigos binarios, 158, 160
códigos compresores, 150
códigos correctores, 150, 159, 161, 162
códigos detectores de errores, 150, 152, 155, 161
coeficiente binomial, 51
Cohen, Paul Joseph, *357*, 357
Cole, Frank Nelson, *216*, 216
colisión de hash, 155, 157
COMAP, 695
combinaciones, 45, 48–50
combinaciones con repetición, 49–51
combinaciones simples, *véase* combinaciones sin repetición
combinaciones sin repetición, 48, 50
combinatoria, 44, 49
completamente aditiva, *véase* función completamente aditiva
completamente multiplicativa, *véase* función completamente multiplicativa
completo, 662, 668, 671–673, 676, 679, 680
comportamiento asintótico, *véase* estimación asintótica
compuesto, *véase* número compuesto
computación cuántica, 147
condiciones de Cauchy-Riemann, 558
Congreso Internacional de Matemáticos, 60, 113, 194, 380, 608
congruencia de Euler-Fermat, *véase* teorema de Euler-Fermat
congruentes, *véase* números congruentes
conjetura de Andrica, 526, 551
conjetura de Artin, 222

conjetura de Beal, 76, 77
conjetura de Brocard, 527, 551
conjetura de Catalan, 79
conjetura de Catalan-Dickson, 410
conjetura de Cramér, 526
conjetura de De Polignac, 512
conjetura de Erdős sobre progresiones aritméticas, 113, 114
conjetura de Erdős-Straus, 79
conjetura de Erdős-Turán, 112
conjetura de Goldbach, 4, 5, 60, 507, 528, 608
conjetura de Legendre, 507, 527, 551
conjetura de Markov, 78, 272
conjetura de Mertens, 483, 611
conjetura de Oppermann, 527, 551
conjetura de Pólya, 403
conjetura débil de Goldbach, 528
conjugado de un número algebraico, *véase* números algebraicos conjugados
conjugado de un número complejo, *véase* números complejos conjugados
conjunto, 636
conjunto bien ordenado, 17, 19–21, 30, 639, 647, 654
conjunto cociente, 639
conjunto infinito, 351–353, 355
conjunto no numerable, 352, 354, 355, 358, 359, 530
conjunto numerable, 254, 273, 352–354, 356–358
conjunto ordenado, 639
conjunto vacío, 636
conmutativa, *véase* propiedad conmutativa
constante de Artin, 222
constante de Erdős-Borwein, 384
constante de Euler, 384, 461, 469, 497, 501, 535, 600, 611, 616, 632, 633
constante de Euler-Mascheroni, *véase* constante de Euler
constante de Lagrange, 270–273, 277, 310, 312, 313, 329
constante de Markov, 270, 272
constante de Mertens, 535
constante de Mills, 530
convergencia de fracción continua, 294, 296, 330, 331
convergente, 289, 290, 294, 296, 302–304, 330–332
convolución de Dirichlet, 410, 419
convolución de Dirichlet generalizada, 427
convolución de Vandermonde, 54
copo de nieve de von Koch, 609
coprimos, *véase* números primos entre sí
Córdoba, Antonio, 695
Corput, *véase* van der Corput
cortadura, 661
cortaduras de Dedekind, 660, 661
cota inferior, 639
cota superior, 639, 661, 662, 667, 672
Cotes, Roger, 557, *558*
Cramer, Gabriel, 307
Cramér, Harald, *526*, 526, 610
Crandall, Richard E., 695
Creyaufmüller, Wolfgang, 694
criba cuadrática multipolinomial, 145
criba de cuerpos, 145
criba de Eratóstenes, 24, 506
criptoanálisis, 134–136, 227
criptografía, 4, 5, 134, 136, 225, 227, 382
criptografía de clave pública, 141, 228
criptología, 134
criptosistema Elgamal, 107, 225, 227–231
criptosistema RSA, 4, 89, 141–143, 145–148, 156, 228, 231
criterio de divisibilidad, 39–41, 43, 44
criterio de Euler, 178, 179, 195, 235
criterio de irracionalidad, 244, 245, 305, 322, 342
criterio de Korselt, 128
criterio de trascendencia, 248, 252, 253, 255, 276, 360
criterio M de Weierstrass, 563, 569
cuadratura, 340

cuadratura del círculo, 371
cuántica, *véase* mecánica cuántica
cuántico, *véase* ordenador cuántico
cuaterna de Hamilton, *véase* cuaternión
cuaternión, 686, 689
cubrimiento, 273, 274
cuerpo, 93, 203, 370, 382, 640–642, 657, 687, 690
cuerpo arquimediano, *véase* propiedad arquimediana
cuerpo cuadrático imaginario, 381
cuerpo cuadrático real, 381
cuerpo ordenado, 657, 668, 671–673, 676, 679–681, 685
Cullen, James, 209, 212, 235
Cunningham, Allan Joseph Champneys, 208, 212
Curbera, Guillermo P., 695
curva elíptica, 5, 77, 93, 145, 229, 382, 383
Cusick, Thomas W., 695

da Vinci, Leonardo, 285
De Koninck, Jean Marie, 695
de la Vallée-Poussin, Charles Jean, 405, 489, *554*, 555, 556, 569, 591, 606, 607, 633
de Moivre, Abraham, *282*, 282, 395, 557
de Oresme, *véase* Oresme
de Polignac, Alphonse, 512
décimo problema de Hilbert, 60
Dedekind, Richard, *352*, 355, *659*, 660, 661, 676
defectuoso, *véase* número defectuoso
del Ferro, Scipione, 336
Demócrito de Abdera, *16*, 16
demostración elemental, 425, 489, 556, 607
densidad asintótica, 112, 485, 504
derivada, 423, 424
desarreglos, *véase* desordenaciones
desarrollo asintótico, 632
desarrollo de Laurent, 563
desarrollo de Taylor, 54, 290, 297, 346, 361, 453, 465, 559
Descartes, René, 58, *59*, *408*, 408, 682
descomposición canónica, 104, 342, 549
desigualdad de Chebyshev, 541
desigualdad de Hadamard y De la Vallée-Poussin, 633
desigualdad triangular, 17
desordenaciones, 163
diagrama de Argand, *véase* plano de Argand
Diamond, Harold G., 694
Dickson, Leonard Eugene, *409*, 409, 410, 695
Diffie, Whitfield, 139, *140*, 140–142, 145, 219
Dilcher, Karl, 695
Diofanto de Alejandría, 3, 58, 59, 70, 240, 650, 695
Dirichlet, Peter Gustav Lejeune, 58, 75, 110, *239*, 240–245, 258, 274, 275, 342, 410, 411, 419, 426–429, 431, 433, 440, *442*, 445, 446, 473, 482, 501, 502, 511, 512, 580, 581, 605, 633
distancia de Hamming, 162
distribución normal, 547
dividir, 657
divina proporción, *véase* razón áurea
divisibilidad, 16, 18, 39
divisible, 18
división, 19, 655, 657
divisor, 18, 20, 24, 28
divisor de cero, 27, 93, 446
divisor propio, 4, 406
DNI, 154
Documento Nacional de Identidad, *véase* DNI
dominio, 640
dominio de integridad, 27
du Bois-Reymond, Paul David Gustav, *455*, 455
du Sautoy, Marcus, 695
Dudley, Underwood, 695
Dunham, William, 695
duplicación del cubo, 372

Durán, Antonio J., 695
Dyson, Freeman J., 251, *252*

EAN, 154
ecuación cúbica disminuida, 394
ecuación de Brocard-Ramanujan, 80
ecuación de Erdős-Moser, 80
ecuación de Fermat, 75, 77
ecuación de Markov, 78, 79, 271, 272
ecuación de Pell, 78
ecuación de Ramanujan-Nagell, 80
ecuación diofántica, 57–59, 64, 68, 78, 81, 85, 139, 240, 256, 272, 382
ecuación diofántica exponencial, 75, 79
ecuación diofántica lineal, 61, 63
ecuación funcional, 468, 616, 617, 633
Edwards, Harold M., 696
egipcios, 79
Eisenstein, Ferdinand Gotthold Max, 183, 193, *194*, 194, 697
elemental, *véase* demostración elemental
elemento, 636
elemento irreducible, 27
elemento neutro, 17, 91, 411, 641, 645, 653, 657, 664, 666, 684, 687
elemento neutro por la izquierda, 428
elemento primo, 27
Elementos de Euclides, 3, 16, 337, 406, 506
Elgamal, *véase* criptosistema Elgamal
Elgamal, Taher, 107, 225, 227, *228*, 228–231, 233, 236
Elkies, Noam D., 77, *78*
Ellis, James H., 145
enteros gaussianos, 193, 690
entropía, 149
equivalentes, 270, 271, 276, 277, 312, 313
Eratóstenes de Cirene, 24, *26*, 506
Erdős, Pál, 3, *79*, 79–81, 112–114, 384, 425, 489, 523, 547, *548*, 556, 696, 699
Escribano, José Javier, 694
espectro de Lagrange, 271, 272, 312
espectro de Markov, 272
estadística, 547
estimación asintótica, 453, 456, 457, 466, 469, 472, 473, 478, 479, 482, 483, 503, 594, 602, 605–608, 610
Euclides de Alejandría, 3, 7, 16, 20, 21, 23, 25, *26*, 26, 30, 31, 69, 70, 92, 94, 96–98, 100–102, 110, 120, 126, 138, 144, 175, 191, 232, 292, 337, 406, *408*, 506, 550, 696, 697
Eudoxo de Cnido, 16, 659
Euler, Leonhard, 4, 8, 26, 53, 58, 75, 77, 88, 89, *98*, 98–101, 103, 104, 110, 111, 113–115, 143, 144, 149, 156, 163, 169, 177–179, 183, 194–201, 203, 204, 206, 215, 218, 235, 244, 245, 281, 305, *318*, 319, 322, 324, 334, 342–344, 349, 383–386, 391, 396, 401, 402, 406, 408, 437, 446, 453, 461, 464–466, 468, 469, 478–480, 482, 483, 487, 488, 497, 501, 507–511, *512*, 512, 521, 527, 535, 555, 557, 565, 566, 593, 599, 600, 602, 611, 613, 615–619, 632, 633, 682, 696–698
European Article Number, *véase* EAN
exponenciación discreta, 137, 138, 141

factorización, 28
factorización canónica, 28, 172, 400, 535
Farey, John, 258–260, 268, 280
Faulhaber, Johann, *384*, 385, 398
Felipe II (rey de España), 136
Fermat, Pierre de, 3, 5, 26, 58, *59*, 75–78, 82, 98–101, 105, 106, 110, 111, 125, 128, 138, 139, 141, 144, 149, 156, 163, 169, 171, 177, 194, 197, 201–204, 207–209, 213, 218, 230, 232, 234–236, 382, *408*, 408, 697, 699
Ferrari, Ludovico, 336
Ferrier, Aimé, 216
Ferro, *véase* del Ferro

Fibonacci (Leonardo de Pisa), 35, *36*, 36, 70, 97, 115, 150, 281, *282*, 282–284, 286, 288, 290, 291, 310, 326, 327, 330, 700
firma digital, 148, 149, 157, 231, 233, 236
firma digital Elgamal, 231, 233
firma digital RSA, 148, 157
Flahive, Mary E., 695
flujo, 439, 440, 443, 444
forma cuadrática binaria, 271, 272
fórmula de Binet, 282, 285, 286, 310
fórmula de Cauchy-Hadamard, 283, 286
fórmula de Cramer, 307
fórmula de cuadratura, 340
fórmula de De Moivre, 395, 557
fórmula de De Polignac, 512
fórmula de duplicación de Legendre, 616, 633
fórmula de Faulhaber, 398
fórmula de inversión de la transformada de Fourier, 620
fórmula de inversión de la transformada de Mellin, 620
fórmula de inversión de Möbius, 413–415, 433, 434, 436, 438, 439, 443, 444, 449, 612, 623, 628
fórmula de inversión de Möbius generalizada, 429, 433, 434
fórmula de Leibniz, 47
fórmula de reflexión de Euler, 616, 633
fórmula de sumación de Abel, *véase* identidad de Abel
fórmula de sumación de Bernoulli, 397
fórmula de sumación de Euler, 464–466, 566
fórmula de sumación de Euler-Maclaurin, 437, 465, 466
fórmula de Taylor, 174
fórmula del producto, 422, 423
fórmula integral de Cauchy, 562, 563, 581
fórmulas de Mertens, 462, 531
Fourier, Jean Baptiste Joseph, 340, 342, *343*, 352, 392, 436, 437, 618, *619*, 620, 698
fracción continua, 240, 245, 263, 268, 270, 273, 280, 281, 285, 288–290, 293, 294, 296, 297, 300, 301, 305, 310, 312, 313, 318, 319, 322, 328–332, 334, 342, 343, 345
fracción continua finita, 288, 289, 291, 293
fracción continua generalizada, 286, 289, 300, 331, 332
fracción continua infinita, 288, 289, 297, 298
fracción continua periódica, 317
fracción continua simple, 280, 286, 289, 290, 294–296, 299–302, 305, 310, 312, 313, 318, 319, 322, 327–331, 342, 343, 345
fracción continua simple finita, 291, 293, 295, 305, 309, 315
fracción continua simple infinita, 281, 297, 298, 303, 305, 343
fracción continua simple periódica, 313, 315, 343
fracción continua simple periódica mixta, 317
fracción continua simple puramente periódica, 315, 317, 318
fracción egipcia, 79
fracción irreducible, 656, 657
fracciones de Brocot, 258–266, 268, 276, 280, 302, 304, 308
fracciones de Brocot reducidas, 261
fracciones de Farey, 258–260, 268, 280
Fraenkel, Abraham Adolf, *356*, 357, 636, 649
Frege, Friedrich Ludwig Gottlob, *649*, 649
Frénicle de Bessy, Bernard, 58
Frobenius, Ferdinand Georg, *65*, 65, 66, 698
Frye, Roger E., 77
función unidireccional, *véase* función de una sola vía
función, 640
función aditiva, 425
función aditiva creciente, 425
función analítica, 558, 559, 562–564

función aritmética, 2, 400, 401, 410, 413–417, 423–431, 433–436, 440, 445, 453, 463, 478, 500, 502, 507, 509, 511, 535, 552
función completamente aditiva, 425, 448
función completamente multiplicativa, 178, 416, 417, 420, 421, 429, 433–437, 442, 443, 445–448, 503, 509, 510, 548, 612, 623
función de Chebyshev, 448, 459
función de Euler, *véase* función indicatriz de Euler
función de Liouville, 403, 417, 453
función de Mertens, 494, 611
función de Möbius, 401, 425, 430, 431, 433, 438, 445, 453, 479, 480, 482, 483, 494, 623
función de Riemann, 613, 625, 629
función de una sola vía, 140, 225, 228
función de von Mangoldt, 404, 415, 417, 425, 430, 431, 453, 490, 491, 493, 605, 633
función de Weierstrass, 382
función entera, 558, 559, 617, 618
función Gamma de Euler, 53, 383, 468, 501, 613, 615, 616, 633
función Gamma incompleta, 626
función generadora de primos, 527–530, 551
función generatriz, 283, 385, 387, 437, 465
función hash, 154, 156, 157
función holomorfa, 558–562, 564
función identidad, 405, 411, 428, 440
función indicatriz de Euler, 98, 101, 103, 104, 143, 402, 446, 453, 478–480, 482, 483, 593, 599, 600, 602
función inversa, 411, 419, 420, 424, 428
función inversa por la izquierda, 428
función meromorfa, 562
función multiplicativa, 407, 416, 417, 419, 420, 422, 445–448, 509, 548
función patológica, 255, 256
función resumen, *véase* función hash
función semimultiplicativa, 447
función unidad, 405, 433
función zeta de Riemann, 2, 386, 467, 468, 482, 510, 526, 555–557, 562, 564–566, 568, 569, 572, 580, 605, 610, 611, 613, 615, 622, 624, 625, 633
funciones aritméticas notables, 401, 453
funciones de Chebyshev, 489, 605, 631
funciones divisor, 402, 472, 482

Galois, Évariste, 317, *318*, 318, *336*, 336, 372
Gamma de Euler, *véase* función Gamma de Euler
Gamma incompleta, *véase* función Gamma incompleta
Gardner, Martin, 145, *146*, 146
Gauss, Carl Friedrich, 3, 8, *88*, 88, 89, 115, 177, 179, 180, 182, 183, 193, 202, 208, *335*, 335, 345, 405, 488, *554*, 554, 555, 557, 594, 611, 612, 683, 685, 690
Gelfond, Aleksandr Osipovich, 251, *252*, 381, 382, 396
generatriz, *véase* función generatriz
Gerberto de Aurillac, *33*, 34, 35
Germain, Marie Sophie, 75, *76*, 76, 139
Girard, Albert, 281, 685
Gödel, Kurt, *357*, 648
Göhl, Gunnar, 698
Golay, Marcel J. E., 162
Goldbach, Christian, 4, 5, 26, 60, 233, 235, 507, 528, 552, 608, 696
Gómez Pardo, José Luis, 696
grado, 170–172, 233, 234, 246, 275, 335, 336, 366, 372, 373, 395
Graham, Ronald L., 696
Gram, Jørgen Pedersen, 613, *614*
Graves, John Thomas, 689
Green, Ben Joseph, 113, *114*, 114
Grönwall, Thomas Hakon, 461, *462*
grupo, 93, 204, 218, 229, 419, 641
grupo cíclico infinito, 445

Guillera, Jesús, 696
Guy, Richard K., 696

Hadamard, Jacques, 283, 286, 405, 489, *554*, 555, 556, 569, 591, 618, 633
Hamilton, Sir William Rowan, 683, 686, *688*, 689
Hamming, Richard Wesley, *160*, 161, 162, 166
Hardy, Godfrey Harold, 4, 5, 111, *112*, 468, 539, *540*, 542, 544, 547, 556, 579, 610, 696, 697
Haselgrove, Colin Brian, 404
hash, *véase* función hash
Havil, Julian, 696
Helfgott, Harald Andrés, 528, 696
Hellman, Martin E., 137, 139, *140*, 140–142, 145, 219
Hensel, Kurt, 173, *174*, 174, 175
heptadecágono regular, 202
Hermite, Charles, 334, 346, *360*, 360, 361, 372, 374, 376, 379, 382
Hilbert, David, *60*, 60, 164, 194, *352*, 353, 361, 380–382, 396, 468, 608, 648
Hill, Raymond, 696
hindúes, 32–35, 56, 280, 393, 642
Hipatia de Alejandría, 58, *59*
hipérbola, 473, 474
Hipócrates de Quíos, 16
hipótesis de Riemann, 60, 133, 468, 483, 488, 526, 604, 605, 607–611, 617
hipótesis de Riemann generalizada, 133
hipótesis del continuo, 60, 356, 357, 468
Hoheisel, Guido, 525
homogéneo, *véase* polinomio homogéneo
l'Hôpital (Marqués), Guillaume de, 603
hotel de Hilbert, 396
Hua Loo Keng, 696
Huguet, Llorenç, 698
Hurwitz, Adolf, *79*, 79, 246, 263, 267, *268*, 268, 270, 271, 273, 280, 285, 302–304, 310, 392, 437, 694, 697, 698
Huygens, Christiaan, *239*, 239, 281

ideal, 76, 194
identidad, *véase* función identidad
identidad de Abel, 463, 491, 494, 532, 566, 587
identidad de Bézout, 21, 23, 95, 120, 277
identidad de Chu-Vandermonde, 54
identidad de Euler, 509, 511, 512, 555, 557
identidad de Hermite, 361, 374, 376
identidad de Legendre, 512, 514, 515
identidad de Selberg, 426
identidad de Vandermonde, 54
Ifrah, Georges, 695, 697
Ikehara, Shikao, 580, 631
imagen, 639, 640
incertidumbre, 149
inconmensurable, 337
indecidible, 357
indicatriz de Euler, *véase* función indicatriz de Euler
índice, 168, 222, 225
indios, *véase* hindúes
inducción, *véase* axioma de inducción
ínfimo, 639
infinito, *véase* conjunto infinito
infinitos primos, 25, 26, 82, 235, 506, 507, 511, 550, 555
infinitos primos en progresiones aritméticas, 109
infinitos primos gemelos, 4, 507, 512
Ingham, Albert Edward, *404*, 404, 525, 580, 581, 608
integral de Riemann-Stieltjes, 463
integral exponencial, 626
integral exponencial generalizada, 626
International Standard Book Number, *véase* ISBN
interpolación, 340
interpretación probabilística, 222, 481, 485, 592, 594
intersección, 637

inversión de Möbius, *véase* fórmula de inversión de Möbius
inversión de Möbius generalizada, *véase* fórmula de inversión de Möbius generalizada
inverso, 17, 93
inyectiva, *véase* aplicación inyectiva
irracional, *véase* número irracional
irracional cuadrático, *véase* número irracional cuadrático
ISBN, 153, 154, 166
ISBN-10, *véase* ISBN
ISBN-13, 153, 166
Isidoro de Sevilla, 35
isomorfismo, 671, 672

Jacobi, Carl Gustav Jacob, 58, 177, 187, *188*, 190–192, 194–196, 201, 210, 235, 236
Jayyam, *véase* Khayyam
Jia Xian, 56
Jones, William B., 697

Kac, Mark, 547, *548*
Kayal, Neeraj, 146, 693
Khayyam, Omar, *56*, 56
Khinchin, Aleksandr Yakovlevich, *254*, 254, 255, 697
Knuth, Donald Ervin, 696, 697
Koblitz, Neal, 697
Koch, *véase* von Koch
Koninck, *véase* De Koninck
Korselt, Alwin Reinhold, 128
Kritikos, Nikolaos, 697
Křížek, Michal, 697
Kronecker, Leopold, *188*, 191, 277, 636
Kummer, Ernst Eduard, *76*, 76, 194
Kuzmin, Rodion Osievich, *381*, 381

Laczkovich, Miklós, 697
Lagrange, Joseph Louis, 58, 88, 89, 99, 111, 115, *123*, 123, 124, 164, 170–173, 175, 176, 234, *268*, 270–273, 277, 310, 312, 313, 315, 324, 329, 527, 682, 695
Lambert, Johann Heinrich, 334, *343*, 343, 345–347, 682, 697
Lamé, Gabriel, *76*, 76
Landau, Edmund, 453, *455*, 461, 476, 487, 507, 556, 593, 600, 602
Lander, Leon J., 77
Landry, Fortuné, 208
Laplace, Pierre Simon, *580*, 581, 587
Laurent, Pierre Alphonse, 563
Legendre, Adrien Marie, 58, 75, 177, 178, 182, 183, 185, 187, 188, 190, 192–194, 196, 210, 235, 405, 488, 507, 512, 514, 515, 527, 551, *554*, 554, 555, 616, 633
Lehman, R. Sherman, 404
Lehmer, Derrick Henry, 107, 197, 212, 213, *214*, 215–218, 407
Leibniz, Gottfried Wilhelm von, *36*, 36, 47, 281, 386, *508*, 508, 682
lema de Gauss, 179, 180, 182
lema de Hensel, 174, 175
Lemmermeyer, Franz, 697
Leonardo de Pisa, *véase* Fibonacci
Levrie, Paul, 697
ley de reciprocidad bicuadrática, 193, 690
ley de reciprocidad cuadrática, 168, 176, 182, 183, 185–187, 190, 193, 201, 202, 210
ley de reciprocidad cúbica, 193
libre de cuadrados, *véase* número libre de cuadrados
Lindemann, Ferdinand von, 361, 371, *372*, 375, 378, 379, 382
Lint, *véase* van Lint
Liouville, Joseph, *239*, 240, 245–250, 254, 256, 275, 276, 326, 329, 334, 349, 350, 359, 360, 403, *404*, 417, 453, 555, 686
Littlewood, John Edensor, 111, *112*, 579, 607, 611, *614*, 614, 697

logaritmo discreto, 139, 141, 156, 168, 223, 225, 227, 230, 233
logaritmo integral, 602, 603
logaritmo principal, 560
Lovelace, Ada, 388, *389*
Luca, Florian, 695, 697
Lucas, Édouard, 206, 212, 213, *214*, 215–218, 327, 407
Lune, *véase* van de Lune

M de Weierstrass, *véase* criterio M de Weierstrass
Maclaurin, Colin, 437, *465*, 465, 466
Maestro Sun, *véase* Sun Zi
Mahavira, 115
Mahler, Kurt, 337, *339*
Mangoldt, *véase* von Mangoldt
Manguel, Alberto, 695
mapeo, 640
Markoff, *véase* Markov
Markov, Andrei Andreevich, 78, 79, 270, 271, *272*, 272, 695, 699
Mascheroni, Lorenzo, *383*, 384, 469
Matiyasevich, Yuri Vladimirovich, *60*, 60
máximo, 639
máximo común divisor, 20, 23, 28, 30, 31, 83
mayas, 642
MD4, 156
MD5, 155, 156
mecánica cuántica, 147
Medalla Fields, 113, 252, 257, 357, 382, 489, 556
mediación, 259, 276
medias de Cesàro, 452, 578, 579
medida de la información, 149
Mellin, Robert Hjalmar, *619*, 620
Mendès France, Michel, 699
Menezes, Alfred J., 697
Mengoli, Pietro, 386, *387*
Merkle, Ralph C., *140*, 140
Mersenne, Marin, 212, 213, *214*, 214, 215, 217, 218, 236, *383*, 384, 407
Mertens, Franz, 462, 480, *482*, 483, 494, 531, 532, 534, 535, 611
método de congruencias lineales, 107, 108
método de factorización, 88, 93, 145, 169
método de la hipérbola, 473
método de la tangente, 678
método de los coeficientes indeterminados, 282
método de Newton, 678
método Monte Carlo, 106
Mihăilescu, Preda V., 79, *80*
Miller, Gary Lee, 129, 130, 132, 133, 165, 197–200
Miller, Jeff C. P., 216
Miller, Steven J., 697
Mills, William H., 529, 530
mínimo, 639
mínimo común múltiplo, 30, 31, 83
Möbius, August Ferdinand, 401, 410, *412*, 413–415, 425, 429–431, 433–435, 438, 439, 443–445, 449, 453, 479, 480, 482, 483, 494, 612, 623, 628, 694, 698
módulo, 88, 560, *véase* valor absoluto
Moivre, *véase* de Moivre
Monte Carlo, *véase* método Monte Carlo
Montgomery, Hugh L., 698
Moser, Leo, *80*, 80, 81
Mozzochi, Charles J., 525
mu de Möbius, *véase* función de Möbius
multiplicación, *véase* producto
multiplicativa, *véase* función multiplicativa
múltiplo, 18, 20, 24
Munuera, Carlos, 697

Nagell, Trygve, *80*, 80
Narkiewicz, Władysław, 697
Nathan, Joseph Amal, 698
Nathanson, Melvyn B., 698
Navas, Luis Manuel, 694, 695, 698

Neumann, *véase* von Neumann
Newman, Donald J., 556, *580*, 580, 581, 698, 700
Newton, Sir Isaac, 48, 49, 51–53, *56*, 103, 107, 163, 401, *508*, 508, 678
Niccolò Fontana, *véase* Tartaglia
Nicómaco de Gerasa, *115*, 115
NIF, 154
Nikishin, Evgenii Mikhailovich, 698
Niven, Ivan, *343*, 346, 348, 698
no numerable, *véase* conjunto no numerable
noveno problema de Hilbert, 60, 194
numerable, *véase* conjunto numerable
numeración hindú, 32
numeración posicional, 642, *véase* sistema de numeración posicional
numeración sexagesimal, 32
número abundante, 406, 445
número algebraico, 4, 239, 246–248, 250–252, 254, 256, 258, 315, 318, 319, 335, 341, 349, 353, 357–359, 365, 368, 370, 379–382, 384, 395–397, 690
número áureo, 285, *véase* razón áurea
número combinatorio, 49
número complejo, 637, 681, 689
número compuesto, 24
número de Carmichael, 127–130, 133, 198
número de Cayley, *véase* octonión
número de Champernowne, 337
número de Cullen, 209, 235
número de Cullen generalizado, 212
número de Cunningham, 212
número de desarreglos, *véase* número de desordenaciones
número de desordenaciones, 163
número de divisores, 402, 453
número de Erdős, 114
número de Fermat, 26, 82, 201–203, 207–209, 218, 235, 236
número de Fibonacci, *véase* sucesión de Fibonacci
número de Frobenius, 65, 66
Número de Identificación Fiscal, *véase* NIF
número de Liouville, 249, 250, 275, 359
número de Mahler, 337
número de Markov, 271, 272
número de Mersenne, 212, 213, 215, 218, 236, 384, 407
número de Proth, 208, 209, 212, 218
número de Riesel, 212
número de Sierpiński, 211, 212
número de Sierpiński de primera clase, 212
número de Woodall, 212
número de Woodall generalizado, 212
número defectuoso, 406, 445
número entero, 3, 16, 637, 650, 651
número imaginario, 660, 682
número irracional, 2, 4, 238, 239, 244, 245, 250, 256, 263–267, 270, 271, 275, 277, 281, 298, 301, 302, 305, 310, 312, 313, 319, 324, 334, 337, 338, 340, 342–345, 347–349, 359, 360, 380, 383, 384, 393–396, 658
número irracional cuadrático, 246, 248, 250, 275, 313, 315, 343, 396
número libre de cuadrados, 222, 421, 485, 504, 630, 691
número natural, 3, 16, 637, 642, 647
número negativo, 650, 651
número normal, 337
número perfecto, 4, 213, 406, 407, 409, 445
número primo, 3, 24–26, 28, 82, 200, 201, 203, 208–210, 212, 213, 215–217, 235, 506, 555, 593, 594
número pseudoprimo, 125, 127, 129, 165, 198, 201, 213, 214, 235
número pseudoprimo de Euler, 194, 196–200, 235
número pseudoprimo fuerte, 132, 198, 199
número racional, 3, 18, 291, 293, 341, 342, 395, 637, 655
número real, 3, 637, 658, 659, 661

número trascendente, 2, 4, 60, 239, 246, 248, 249, 252–255, 275, 326, 329, 334, 335, 337, 346, 349–353, 358–361, 363, 368, 371, 372, 374, 375, 379–384, 396, 397, 468
números aleatorios, 106
números algebraicos conjugados, 318, 365
números amigos, 406, 408, 409, 445
números complejos conjugados, 684
números congruentes, 19, 88, 90, 168
números coprimos, *véase* números primos entre sí
números de Bernoulli, 385, 387, 388, 391, 397, 437, 465, 466
números de Fibonacci, 282
números de Liouville, 359
números de Lucas, 327
números primos entre sí, 24, 481
números pseudoaleatorios, 106, 107, 109, 164, 230, 233
números relativamente primos, *véase* números primos entre sí
números romanos, 32
números sociables, 409

O grande, 9, 453
O mayúscula, *véase* O grande
o minúscula, *véase* o pequeña
o pequeña, 9, 453, 454
octava de Cayley, *véase* octonión
octavo problema de Hilbert, 60, 608
octonión, 686, 689
Odlyzko, Andrew M., *482*, 483, 611
olmecas, 642
Oorschot, *véase* van Oorschot
Oppermann, Ludvig, *526*, 527, 551
orden, *véase* relación de orden
orden (de un elemento), 99, 163, 203, 204, 218, 219
orden (de un grupo), 99
orden (de un número algebraico), 246, 247, 250, 251, 315, 318, 319, 349
orden de aproximación por racionales, 245–248, 250, 251, 258, 275, 326, 329
orden parcial, *véase* relación de orden parcial
orden total, *véase* relación de orden total
ordenador cuántico, 147
Oresme, Nicolás de, *508*, 508
ortogonalidad, 340, 394, 395

Pacioli, Luca, 7, *285*, 285
Padé, Henri, 290, *360*, 361, 383
Paganini, Niccolò, 408
palomar, *véase* principio del palomar
pandilla de números, *véase* números sociables
papa Silvestre II, *véase* Gerberto de Aurillac
par de sucesiones monótonas convergentes, 678, 679
par de sucesiones recurrentes, 286–288, 327
par regular de sucesiones recurrentes, 286, 294, 303, 327
paralelizable, *véase* superficie esférica paralelizable
parametrización de Diofanto, 70
parametrización de Euclides, 70
Parkin, Thomas R., 77
parte alícuota, 406, 409
parte entera, 669
partición, 638
Pascal, Blaise, 54, 55, *56*, 57
Patashnik, Oren, 696
Peano, Giuseppe, 18, *21*, *643*, 643–648, 650
Pell, John, *78*, 78
pentágono estrellado, 285
Pépin, Théophile, 201, 203, 205, 206, 208, 210, 217, 218, 236
perfecto, *véase* número perfecto
periodo, 108, 109, 164
permutaciones, 45–47
permutaciones con repetición, 47, 49
Perron, Oskar, *272*, 272, 273, 310
persas, 56, 123

Pervushin, Ivan Mikheevich, *207*, 208
phi de Euler, *véase* función indicatriz de Euler
Pitágoras de Samos, 16, 67, 68, *69*, 70, 74, 334, 337, 408, 658, *659*
planetario, 239
plano de Argand, 683
Platón, 68, *69*, 70
Plimpton 322, 69
Pohlig, Stephen C., 137, 141, 142
Polignac, *véase* de Polignac
polinomio homogéneo, 366
polinomio simétrico, 363, 365–368, 370
polinomio simétrico elemental, 366
polinomios de Bernoulli, 385, 387–389, 391, 392, 397, 437, 465, 466
polinomios de Chebyshev, 339, 340, 394, 449
polinomios de Chebyshev de primera clase, 395
polinomios de Chebyshev de segunda clase, 395
polinomios ortogonales, 340, 383, 394, 395
Pollard, John M., 145
polo, 382, 468, 563, 570, 616, 617, 631
Pólya, György, 26, 403, *404*
Pomerance, Carl, 695
postulado de Bertrand, 516, 521, 523–525, 528–530
postulados de Peano, *véase* axiomas de Peano
potencia del continuo, 356
Poulet, Paul, 409
Prasad, Akhauri Vindhyachal, 274, 698
Premio Abel, 113
primer problema de Hilbert, 60, 356
primo, *véase* número primo
primo 1-seguro, 139, 147
primo de Cullen, 209, 235
primo de Fermat, 201–203, 218, 236
primo de Germain, 75, 76, 139
primo de Mersenne, 213, 215–217, 407
primo de Proth, 209, 210
primo de Wieferich, 105, 106, 125
primo de Wilson, 125
primos entre sí, *véase* números primos entre sí
primos gemelos, 4, 165, 507, 512
príncipe de las matemáticas, 3
principio de buena ordenación, *véase* conjunto bien ordenado
principio de clasificación cruzada, 103
principio de Dirichlet, *véase* principio del palomar
principio de inclusión-exclusión, 103, 163, 419
principio de inducción, *véase* axioma de inducción
principio de prolongación analítica, 562
principio del mínimo, *véase* conjunto bien ordenado
principio del palomar, 241, 274
principio fundamental de la teoría de números, 344
probabilidad, 44, 481, 485, 541, 594
problema de Basilea, 386
problema de Brocard, 80
problema de Frobenius, 65–67
problema de las monedas, 66
problema de Riesel, 212
problema de Sierpiński, 211, 212
problema de Waring, 164
problema del cambio de moneda, *véase* problema de las monedas
problema diofántico lineal de Frobenius, *véase* problema de Frobenius
problemas de Hilbert, *véase* primer p. H., séptimo p. H., octavo p. H., noveno p. H., décimo p. H.
producto, 640, 645, 652, 655, 660, 665, 666, 684, 686, 690
producto de Dirichlet, 410, 411, 419, 426–429, 431, 433, 440, 445, 446, 502
producto de Euler, 507, 509, 512, 565, 619
programa de formalización de Hilbert, 648

progresión aritmética, 109
propiedad antisimétrica, 639, 646
propiedad arquimediana, 665, 668, 681
propiedad asociativa, 17, 411, 641, 646, 653, 657, 664, 666, 684, 686, 689, 690
propiedad asociativa mixta, 428, 433, 434, 439, 440
propiedad conmutativa, 17, 411, 641, 646, 653, 657, 664, 666, 684, 686, 687, 689
propiedad distributiva, 17, 642, 646, 653, 657, 667, 684, 686, 690
propiedad reflexiva, 638, 646
propiedad simétrica, 638
propiedad transitiva, 638, 646
proporción áurea, *véase* razón áurea
Proth, François, 208–210, 212, 218
protocolo de Diffie-Hellman, 140–142, 145, 219
protocolo de Diffie-Hellman-Merkle, *véase* protocolo de Diffie-Hellman
proyecto Cunningham, 208, 212
prueba del nueve, 150–153
prueba M de Weierstrass, *véase* criterio M de Weierstrass
pseudoaleatorios, *véase* números pseudoaleatorios
pseudoprimo, *véase* número pseudoprimo
pseudoprimo de Euler, *véase* número pseudoprimo de Euler
pseudoprimo fuerte, *véase* número pseudoprimo fuerte

Qin Jiushao, 115
quiniela, 274
Qurra, *véase* Thabit ibn Qurra

Rabin, Michael Oser, 129, *130*, 130, 132, 133, 165, 197–200
racional, *véase* número racional
raíces conjugadas, 365
raíces de la zeta de Riemann, *véase* ceros de la zeta de Riemann
raíz primitiva, 141, 168, 219–223, 236
Ramanujan, Srinivasa, *80*, 80, 301, 459, 539, *540*, 542, 544, 547, 613
Ramírez Alfonsín, Jorge Luis, 698
razón áurea, 36, 268, 271, 273, 285, 290, 291, 302, 310–313, 328
recorrido, 640
recta crítica, 468, 605, 608, 610, 617
recta real, 238, 352, 636, 661
recurrencia lineal, 283
reflexiva, *véase* propiedad reflexiva
regla de l'Hôpital, 603
regla de oro, *véase* razón áurea
regla y compás, 202, 203, 371, 372
reglas de Cardano, 364–366
Rejewski, Marian Adam, *136*, 136
relación, 638
relación de equivalencia, 90, 638, 652, 653, 656, 679
relación de orden, 639, 646, 652, 656, 662, 672, 679
relación de orden parcial, 639
relación de orden total, 639, 654, 656, 662
relativamente primos, *véase* números primos entre sí
reloj, 88, 138, 239, 562
representación de un entero en una base, *véase* base
representante de una clase de equivalencia, 91, 98, 219
residuo, *véase* resto
residuo (de una función compleja), 563
residuo cuadrático, *véase* resto cuadrático
resistencia a las colisiones de hash, 155, 157
resistencia al cálculo de preimágenes, 155–157
resta, 650, 651, 653
resto, 19, 88
resto cuadrático, 168, 176, 177, 183, 185–187, 190, 234–236
retículo, 382, 383
Rey Pastor, Julio, *514*, 514

Ribenboim, Paulo, 698
Riele, *véase* te Riele
Riemann, Georg Friedrich Bernhard, 2, 60, 133, 386, 438, 463, 467, 468, 482, 483, 488, 510, 526, *554*, 555–558, 562, 564, 566, 568, 569, 572, 580, 604, 605, 607–619, 622, 624, 625, 629, 633, 693, 696, 698
Riesel, Hans Ivar, 212, 698
Rifà, Josep, 698
Rivest, Ronald Linn, 4, 142, *143*, 145, 155
Robin, Guy, 611
Robinson, Raphael Mitchel, *216*, 217
Rødseth, Øystein J., 698
Rose, Harvey E., 699
Rosen, Kenneth H., 699
Roth, Klaus Friedrich, 112, 251, 252, *253*, 254–257, 276, 326, 337, 359, 360, 699
RSA, *véase* criptosistema RSA
RSA129, 145
Ruiz, Francisco José, 695, 698
Russell, Bertrand Arthur William, *649*, 649
Ruzsa, Imre Z., 699

Sarrus, Pierre Frédéric, 127
Sautoy, *véase* du Sautoy
Saxena, Nitin, 146, 693
Schneider, Theodor, *381*, 381, 382, 396
Schoenfeld, Lowell, *609*, 609
Schur, Issai, *525*, 525
Seelhoff, Paul Peter Heinrich, 208
Segunda Guerra Mundial, 136
Seki, Takakazu, *384*, 385
Selberg, Atle, *425*, 425, 426, 489, 556, 607
Selfridge, John Lewis, *211*, 211
semigrupo, 641
semilla, 107
semimultiplicativa, *véase* función semimultiplicativa
séptimo problema de Hilbert, 60, 380–382, 396
serie alternada, 329
serie armónica, 82, 469
serie de Dirichlet, 482, 501, 502, 511, 512, 580, 581, 605, 633
serie de Fourier, 340, 352, 392, 436
serie de Gram, 613
serie de Taylor, 435, 449, 579, 619
serie telescópica, 386, 387, 391, 392
Serret, Joseph Alfred, 270, 312, *313*, 313
SHA, 156
Shamir, Adi, 4, 142, *143*, 145, 156
Shannon, Claude Elwood, 149, *160*, 161
Shapiro, Harold N., 110
Shidlovskii, Andrei Borisovich, 382, 699
Shor, Peter, *146*, 147
Siegel, Carl Ludwig, 251, 252, *253*, 254–257, 276, 337, 359, 360, 382, 699
Sierpiński, Wacław, *211*, 211, 212, 699
Silvestre II (papa), *véase* Gerberto de Aurillac
símbolo de Jacobi, 187, 190–192, 194–196, 201, 210, 235, 236
símbolo de Kronecker, 191
símbolo de Legendre, 177, 178, 182, 185, 187, 188, 192, 194, 196, 210, 235
símbolos de Landau, 453
simétrica, *véase* propiedad simétrica
simétrico, *véase* polinomio simétrico
Simson, Robert, 284, *285*, 285
Singh, Simon, 699
singularidad aislada, 562, 563
singularidad esencial, 563
sistema completo de restos, 91
sistema de numeración posicional, 32, 33, 36, 150
sistema reducido de restos, 98, 219
sobreyectiva, *véase* aplicación sobreyectiva
sociables, *véase* números sociables
Solovay, Robert Martin, 133, 196, *197*, 197, 199, 200
solución de la ecuación cúbica, 336, 394, 681, 682

Somer, Lawrence, 697
soporte, 446
Soria, Javier, 125
Sorokin, Vladimir Nikolaevich, 698
Srinivasan, Anitha, 699
Stern, Moritz, 258
Stevin, Simon, 650, *651*, 659, 660
Stieltjes, Thomas Jan, *462*, 463
Stinson, Douglas Robert, 699
Strassen, Volker, 133, 196, *197*, 197, 199, 200
Straus, Ernst Gabor, *79*, 79
subconjunto, 637
sucesión alicuatoria, 406, 409
sucesión de Cauchy, 660, 678–681
sucesión de Fibonacci, 36, 97, 281, 283, 286, 288, 290, 291, 310, 326, 327, 330
sucesión de intervalos encajados, 680, 681
sucesión de Lucas, *véase* números de Lucas
sucesión de Sylvester, 82
sucesión monótona creciente, 679
sucesión monótona decreciente, 679
sucesión recurrente, 286, 327
sucesiones de Brocot, 258
suma, 640, 645, 652, 655, 660, 663, 683, 686, 690
suma de hash, 155, 157
suma de los divisores, 402, 453
sumabilidad Abel, 579
sumas parciales, 427, 429
Sun Zi, 115
superficie esférica paralelizable, 689
suprayectiva, *véase* aplicación suprayectiva
supremo, 639, 661, 662, 667, 672
Surányi, János, 696
Sylvester, James Joseph, *65*, 66, 82
Szemerédi, Endre, 112, *113*

tabla de índices, 224
tabla de números primos, 25, 506
tabla de raíces primitivas, 224
tablilla de arcilla, 69
Takloo-Bighash, Ramin, 697
Tales de Mileto, *16*, 16
Tanaka, Minoru, 404
Tao, Terence, 113, *114*, 114
Tartaglia (Niccolò Fontana), 56, *335*, 336
Tauber, Alfred, *579*, 579, 580
tauberiano, *véase* teorema tauberiano
Taylor, Brook, 54, 174, 290, 297, 346, 361, 385, *387*, 387, 395, 435, 449, 453, 465, 559, 579, 619
Tchebychef, *véase* Chebyshev
te Riele, Herman J. J., *482*, 483, 611
Teeteto de Atenas, 16
Tena, Juan Gabriel, 697
Tenenbaum, Gérald, 699
teorema chino del resto, 114, 116, 118, 170, 172, 186, 187, 215
teorema de Abel, 578, 579
teorema de Bolzano, 670, 671
teorema de Chebyshev, 512, 518, 551
teorema de Dirichlet, 240–245, 258, 275, 342
teorema de Erdős-Kac, 547
teorema de Euler-Fermat, 98–101, 110, 111, 144, 149, 156, 163, 169, 177, 201, 203, 204, 218
teorema de factorización única, 338
teorema de Fermat (pequeño), *véase* teorema pequeño de Fermat
teorema de Fermat (último), *véase* último teorema de Fermat
teorema de Gelfond y Dyson, 251
teorema de Gelfond-Schneider, 381, 382, 396
teorema de Goldbach, 235
teorema de Green-Tao, 113, 114
teorema de Hardy-Ramanujan, 542, 544
teorema de Hurwitz, 246, 263, 267, 268, 270, 273, 280, 285, 302–304, 310
teorema de la media de Fermat, 100, 144, 156
teorema de Lagrange, 99, 123, 124, 170–173, 175, 176, 234
teorema de las raíces racionales, 338, 339
teorema de Lindemann-Weierstrass, 379

teorema de Liouville, 245–248, 250, 254, 256, 275, 276, 326, 329, 349, 359, 360
teorema de Liouville (de análisis complejo), 555, 686
teorema de los cuatro cuadrados, 164
teorema de los números primos, 2, 5, 112, 405, 426, 438, 459, 462, 468, 483, 486–489, 491, 493–495, 506, 512, 518, 554–557, 564, 565, 569, 576, 580, 591–593, 596, 600, 604–606, 628, 630, 631
teorema de los residuos, 563
teorema de Mihăilescu, 79
teorema de Pitágoras, 67, 68, 74, 658
teorema de Prasad, 274
teorema de reciprocidad cuadrática, *véase* ley de reciprocidad cuadrática
teorema de Roth, *véase* teorema de Thue-Siegel-Roth
teorema de Serret, 270, 313
teorema de Siegel, 251
teorema de Szemerédi, 112
teorema de Thue, 251, 257
teorema de Thue-Siegel-Roth, 252, 254–257, 276, 326, 337, 359, 360
teorema de Wilson, 123–125, 165, 234
teorema del número primo, *véase* teorema de los números primos
teorema fundamental de la aritmética, 16, 28, 29, 31, 81, 182, 342, 401, 415, 506, 509, 565
teorema fundamental de la aritmética modular, 99
teorema fundamental de los polinomios simétricos, 366
teorema fundamental del álgebra, 8, 170, 335, 555, 685
teorema fundamental del cálculo, 603
teorema integral de Cauchy, 561
teorema pequeño de Fermat, 100, 105, 125, 128, 129, 138, 141, 163, 171, 177, 197, 213, 230, 232, 234, 235
teorema tauberiano, 578, 580, 581, 587, 591, 631
teorema tauberiano de Wiener-Ikehara, 580, 631
teoremas de incompletitud de Gödel, 648
teoría de aproximación, 340, 632
teoría de códigos, 149, 162
teoría de conjuntos, 2, 18, 352, 357, 636, 647
teoría de conjuntos cantoriana, 357
teoría de cubrimientos, 273
teoría de cuerpos, 203
teoría de grupos, 99, 204, 218
teoría de la información, 161
teoría de números aditiva, 5
teoría de números algebraica, 5
teoría de números analítica, 5, 555
teoría de números combinatoria, 6
teoría de números computacional, 5, 97, 145, 407
teoría de números elemental, 5
teoría de números geométrica, 6
teoría de números multiplicativa, 5, 20
teoría de números probabilística, 6, 548
teoría de probabilidad, 541
terna pitagórica, 61, 67, 72, 85
terna pitagórica primitiva, 70, 71, 74
test de Lucas-Lehmer, 212, 213, 215–218, 407
test de Miller-Rabin, 129, 132, 133, 165, 197–200
test de Pépin, 201, 203, 205, 206, 208, 210, 217, 218, 236
test de primalidad, 88, 93, 125–127, 129, 146, 169, 196, 197, 200, 201, 212
test de primalidad AKS, 146, 201
test de primalidad no seguro, 129, 133, 196, 197, 201
test de primalidad probabilístico, *véase* test de primalidad no seguro
test de primalidad seguro, 133, 201
test de Proth, 208, 210

test de Solovay-Strassen, 133, 196, 197, 200
Thabit ibn Qurra, *408*, 408, 445
Thron, Wolfgang J., 697
Thue, Axel, 251, 252, *253*, 254–257, 276, 337, 359, 360
totient, *véase* función indicatriz de Euler
transformación modular, 270, 271, 277, 312
transformada de Fourier, 620
transformada de Laplace, 581, 587, 620
transformada de Mellin, 620
transformada de Möbius, 410, 413
transitiva, *véase* propiedad transitiva
trascendente, *véase* número trascendente
tren de Hilbert, 396
triángulo de Pascal, 54–57
triángulo de Tartaglia, 56
triángulo de Yang Hui, 56
trinomio de Euler, 527
trisección del ángulo, 372
Tsu Chung-Chi, *véase* Zu Chongzhi
Tubbs, Robert, 694
Turán, Pál, 112, 539, *540*
Turing, Alan Mathison, *136*, 136

último teorema de Fermat, 3, 5, 58, 75–78, 100, 105, 106, 139, 194, 382
unicidad, 28, 293, 297, 298, 305
unicidad salvo isomorfismos, 671, 680, 681
Unicode, 159
unidad imaginaria, 684, 687
unión, 637

Vallée-Poussin, *véase* de la Vallée-Poussin
valor absoluto, 17, 654, 666
valor principal de una integral, 602
van de Lune, Jan, 693
van der Corput, Johannes Gualtherus, 111, *113*
van Lint, Jacobus Hendricus, 699
van Oorschot, Paul C., 697
Vandermonde, Alexandre-Théophile, 54
Vanstone, Scott A., 697
variable compleja, 110
variaciones, 45
variaciones con repetición, 45, 46, 49
variaciones simples, *véase* variaciones sin repetición
variaciones sin repetición, 45, 46, 48
Varona, Juan Luis, 694, 695, 698, 699
vía, *véase* función de una sola vía
Viète, François, 58, *136*, 136
Vigila, *33*, 34, 35
Vinci, *véase* da Vinci
Vinogradov, Ivan Matveevich, 528, 607, *609*
von Leibniz, *véase* Leibniz
von Lindemann, *véase* Lindemann
von Koch, Niels Fabian Helge, *609*, 609
von Mangoldt, Hans Carl Friedrich, 404, 415, 417, *425*, 425, 430, 431, 453, 487, 490, 491, 493, 605, 614, 633
von Neumann, János, *649*, 649
Vorobiev, Nicolai Nicolaevich, 700
Voyager, 162

Wallis, John, *280*, 281, 386
Wantzel, Pierre Laurent, 202, 372
Waring, Edward, 111, *112*, 164
Weierstrass, Karl, 361, *372*, 379, 382, 557, *558*, 563, 569, 670
Wessel, Caspar, 683
Westzynthius, Erik, 526
Wheeler, David J., 216
Wieferich, Arthur Josef Alwin, 105, 106, 125, 695
Wiener, Norbert, *579*, 580, 631
Wigert, Carl Severin, 459, 593, 596, 599
Wiles, Andrew, 3, 58, *59*, 76, 100, 382
Williamson, Malcolm J., 145
Wilson, John, *123*, 123–125, 165, 234, 695
Woodall, Herbert J., 212
Wright, Edward Maintland, *526*, 528, 530, 696

Yang Hui, 56, *115*, 115

Zagier, Don, 700
Zermelo, Ernst, *356*, 357, 636, 649
zeta de Riemann, *véase* función zeta de Riemann
Zhu Shijie, *véase* Chu Shih-Chieh
Zhúkov, Alexandr Vladímirovich, 700
Ziegler, Günter M., 693
Zimmermann, Paul, 694
Zu Chongzhi, 300, *301*
Zuckerman, Herbert S., 698

Este libro es un curso introductorio sobre teoría de números. Pero no sólo contiene las matemáticas como tal, y explicadas con todo rigor, sino que todo el texto está impregnado de pequeñas referencias históricas sobre los resultados expuestos y los matemáticos que los descubrieron, así como comentarios sobre sus relaciones con otros campos de las matemáticas y aplicaciones.
Uno de los objetivos que se ha perseguido al escribir este libro ha sido darle una gran flexibilidad de uso. Consta de diez temas y un apéndice, que son —en la medida de lo posible— lecciones sueltas sin excesiva dependencia unas de otras, a modo de recorridos por diversas partes del extenso territorio de las matemáticas en general y la teoría de números en particular.
Está dirigido, de manera primordial, a estudiantes de matemáticas a nivel de pregraduados, y es fruto de varios años enseñando estas materias, en las que el autor tiene amplia experiencia docente y de investigación. Las explicaciones son bastante detalladas, de modo que sean útiles para el lector autodidacta. Se incluyen numerosos ejercicios, todos ellos con soluciones, a las que se puede acceder a través de una aplicación web. Por otra parte, un profesor que quiera usar este libro puede elegir los temas que desee (o porciones suyas) para seguirlos en un curso, dependiendo de su interés personal y las circunstancias del curso.
Al adquirir el libro en papel se regala la versión electrónica (PDF) y la subscripción para el acceso a la solución de los problemas en la web. La versión electrónica (con acceso a soluciones) se puede adquirir por separado.
En esta segunda edición se han corregido numerosas erratas y se han añadido nuevos materiales y ejercicios.

ISBN 978-84-946150-4-7