

Índice

34. [La teoría de números en el siglo XIX](#)
35. [El resurgimiento de la geometría proyectiva](#)
36. [La geometría no euclídea](#)
37. [La geometría diferencial de Gauss y Riemann](#)
38. [Las geometrías proyectiva y métrica](#)
39. [La geometría algebraica](#)
40. [La introducción del rigor en el análisis](#)
41. [La fundamentación de los números reales y transfinitos](#)
42. [Los fundamentos de la geometría](#)
43. [La matemática en torno al 1900](#)
44. [La teoría de funciones de una o varias variables reales](#)
45. [Ecuaciones integrales](#)
46. [El análisis funcional](#)
47. [La teoría de series divergentes](#)
48. [El análisis tensorial y la geometría tensorial](#)
49. [La aparición del álgebra abstracta](#)
50. [Los orígenes de la topología](#)
51. [Los fundamentos de la matemática](#)

Capítulo 34

La teoría de números en el siglo XIX

Contenido:

1. *Introducción*
 2. *La teoría de congruencias*
 3. *Números algebraicos*
 4. *Los ideales de Dedekind*
 5. *La teoría de las formas*
 6. *Teoría analítica de números*
- Bibliografía.*

Es cierto que Fourier pensaba que el objeto principal de las matemáticas era el uso público y la explicación de los fenómenos naturales; pero un filósofo como él debía saber que el único objetivo de la ciencia es el honor del espíritu humano y que bajo esta perspectiva un problema de [la teoría de] números es tan valioso como un problema sobre el sistema del mundo.

C. G. J. Jacobi

1. Introducción

Hasta el siglo XIX, la teoría de números era una serie de resultados aislados, aunque muchas veces brillantes. Una nueva era se inició con el *Disquisitiones Arithmeticae (Disquisiciones Aritméticas)*¹, de Gauss, que compuso a la edad de veinte años. Esta gran obra había sido enviada a la Academia Francesa en 1800 pero la rechazaron y entonces Gauss la publicó él mismo. En este libro estableció la notación; sistematizó y extendió la teoría existente; clasificó los problemas que debían ser estudiados y los métodos conocidos para atacarlos e introdujo nuevos métodos. En el trabajo de Gauss sobre la teoría de números hay tres ideas principales: la teoría de congruencias, la introducción de los números algebraicos y la teoría de las formas como idea básica en el análisis diofántico. Este trabajo no solamente inició la teoría moderna de números, sino que determinó las direcciones de trabajo en la materia hasta el presente. Las *Disquisitiones* son difíciles de leer, pero Dirichlet las explicó.

Otro avance importante en el siglo XIX es la teoría analítica de números, que además de usar el álgebra se vale del análisis para tratar problemas concernientes a los enteros. Dirichlet y Riemann se erigieron como figuras en esta innovación.

2. La teoría de congruencias

A pesar de que la noción de congruencia no se originó con Gauss — ya que aparece en los trabajos de Euler, Lagrange y Legendre—, introdujo la notación en la primera sección del *Disquisitiones*, aplicándola a partir de ahí sistemáticamente. La idea básica es

simple: el número 27 es congruente con el 3 módulo 4,

$$27 = 3 \text{ módulo } 4,$$

ya que $27 - 3$ es exactamente divisible por 4. (La palabra módulo con frecuencia es abreviada como mód.) En general, si a , b y m son enteros:

$$a = b \text{ módulo } m$$

si $a - b$ es (exactamente) divisible por m o si a y b tienen los mismos restos (o residuos), en la división por m . Entonces se dice que b es un residuo de a módulo m y a es un residuo de b módulo m . Como Gauss muestra, todos los residuos de a módulo m , para una a y m fijas, están dados por $a + km$, donde $k = 0, \pm 1, \pm 2, \dots$

Las congruencias con respecto al mismo módulo se tratan hasta cierto punto como ecuaciones. Algunas congruencias pueden ser añadidas, sustraídas y multiplicadas. Es posible además preguntarnos acerca de la solución de congruencias donde haya incógnitas. Así, ¿qué valores de x satisfacen $2x = 25$ módulo 12?

Esta ecuación no tiene soluciones ya que $2x$ es par y $2x - 25$ es impar. De aquí que $2x - 25$ no puede ser múltiplo de 12. El teorema básico sobre congruencias polinomiales, que Gauss vuelve a probar en la segunda sección, ya lo había establecido Lagrange.²

Una congruencia del grado n ,

$$Ax^n + Bx^{n-1} + \dots + Mx + N = 0 \text{ módulo } p$$

cuyo módulo es un número primo p que no divide a A no puede tener más de n raíces no congruentes.

En la tercera sección Gauss toma los restos de potencias. Proporciona una prueba en términos de congruencias del teorema menor de Fermat, que, enunciado en términos de potencias, se lee: si p es primo y a no es múltiplo de p entonces:

$$a^{p-1} \equiv 1 \text{ módulo } p.$$

El teorema se sigue de su estudio de congruencias de grado superior, a saber,

$$x^n \equiv a \text{ módulo } m$$

donde a y m son primos entre sí. Este estudio fue continuado por muchos autores después de Gauss.

La cuarta sección de las *Disquisitiones* trata de los residuos cuadráticos. Si p es primo y a no es múltiplo de p y si existe una x tal que $x^2 \equiv a \text{ mód. } p.$, entonces a es residuo cuadrático de p ; si no es así, a no es un residuo cuadrático de p . Después de demostrar algunos teoremas subordinados sobre residuos cuadráticos, Gauss proporcionó la primera demostración rigurosa de la ley de reciprocidad cuadrática (cap. 25, sec. 4). Euler, a su vez, había dado un enunciado completo muy parecido al de Gauss en un ensayo de

sus *Opuscula Analytica (Opúsculos Analíticos)* de 1783 (cap. 25, sec. 4). Sin embargo, en el párrafo 151 de las *Disquisitiones*, Gauss afirma que nadie ha presentado el teorema en una forma tan simple como él lo ha hecho. Se refiere a otros trabajos de Euler, incluyendo otro ensayo en el *Opuscula*, y al trabajo de Legendre de 1785. De estos ensayos Gauss dice, correctamente, que las pruebas eran incompletas.

Se supone que Gauss descubrió una prueba de la ley en 1796 cuando tenía 19 años. Aportó otra demostración en las *Disquisitiones* y más tarde publicó otras cuatro. Entre sus ensayos inéditos se encontraron otras dos. Gauss afirma que buscó muchas pruebas porque deseaba hallar una que pudiera usarse para establecer el teorema de reciprocidad bicuadrática (véase más adelante). La ley de reciprocidad bicuadrática, que Gauss llamó la joya de la aritmética, es un resultado básico sobre congruencias. Después de que Gauss diera sus pruebas, otras 50 más fueron proporcionadas por matemáticos posteriores.

Gauss también trató las congruencias de polinomios. Si A y B son dos polinomios en x con, digamos, coeficientes reales, entonces se sabe que es factible encontrar polinomios únicos Q y R tales que:

$$A = BQ + R,$$

donde el grado de R es menor que el grado de B . Se dice entonces que dos polinomios A_1 y A_2 son congruentes módulo un tercer polinomio P si tienen el mismo resto R de la división por P .

Cauchy se valió de esta idea³ para definir los números complejos usando congruencias polinomiales. Así, si $f(x)$ es un polinomio con coeficientes reales, entonces con la división por $x^2 + 1$:

$$f(x) \equiv a + bx \pmod{x^2 + 1}$$

ya que el residuo es de grado menor que el divisor. En este caso, a y b son necesariamente reales en virtud del proceso de división. Si $g(x)$ es otro polinomio:

$$g(x) \equiv c + dx \pmod{x^2 + 1}.$$

Ahora, Cauchy señala que si A_1 , A_2 y B son polinomios cualesquiera y si

$$A_1 = BQ_1 + R_1 \text{ y } A_2 = BQ_2 + R_2,$$

entonces,

$$A_1 + A_2 \equiv R_1 + R_2 \pmod{B}, \text{ y } A_1A_2 \equiv R_1R_2 \pmod{B}.$$

Ahora podemos ver fácilmente que

$$f(x) + g(x) \equiv (a + c) + (b + d)x \pmod{x^2 + 1}$$

y ya que

$$x^2 \equiv -1 \pmod{x^2 + 1},$$

que

$$f(x)g(x) \equiv (ac - bd) + (ad + bc)x \pmod{x^2 + 1}.$$

Vemos así que los números $a + bx$ y $c + dx$ se combinan como números complejos; esto es, tienen las propiedades formales de los números complejos, tomando x el lugar de i . Cauchy también demostró que todo polinomio $g(x)$ no congruente con 0 módulo $x^2 + 1$ tiene inverso, ello significa que existe un polinomio $h(x)$ tal que $h(x)g(x)$ es congruente con 1 módulo $x^2 + 1$.

Cauchy introdujo i por x , siendo i para él una cantidad indeterminada real. Entonces demostró que para cualquier

$$f(i) = a_0 + a_1i + a_2i^2 + \dots$$

se tiene

$$f(i) \equiv a_0 - a_2 + a_4 - \dots + (a_1 - a_3 + a_5 - \dots) i \pmod{i^2 + 1}.$$

De aquí que cualquier expresión con números complejos se comporta como una de las formas $c + di$ y todo el aparato necesario para trabajar con expresiones complejas está disponible. Para Cauchy, entonces, los polinomios en i , con su concepción de i ,

toman el lugar de los números complejos y es posible situar dentro de una clase todos aquellos polinomios teniendo el mismo residuo módulo $i^2 + 1$. Estas clases son los números complejos.

Es interesante que en 1847 Cauchy aún tuviera sus dudas acerca de $\sqrt{-1}$. Dice: «En la teoría de equivalencias algebraicas sustituida por la teoría de números imaginarios, la letra i cesa de representar al signo simbólico $\sqrt{-1}$, que repudiamos completamente y el cual podemos abandonar sin remordimiento ya que no se sabe lo que este supuesto signo significa ni qué sentido atribuirle. Por el contrario, representamos por la letra i una cantidad real pero indeterminada, y al sustituir el signo $=$ por el signo \equiv transformamos lo que ha sido llamado una ecuación imaginaria en una equivalencia algebraica relativa a la variable i y al divisor $i^2 + 1$. Ya que el divisor permanece igual en todas las fórmulas, se evita el escribirlo.»

En la segunda década del siglo, Gauss procedió a buscar las leyes de reciprocidad aplicables a las congruencias de grado superior. Estas leyes introducen de nuevo residuos de congruencias. Así, para la congruencia

$$x^4 \equiv q \pmod{p}$$

se define q como un residuo bicuadrático de p si existe un valor entero de x satisfaciendo la ecuación. Llegó a la ley de reciprocidad cuadrática (véase más abajo) y a una ley de reciprocidad cúbica. Gran parte de este trabajo apareció en los artículos de 1808 a 1817

y el propio teorema sobre residuos bicuadráticos aparece en ensayos de 1828 y 1832⁴.

Para obtener elegancia y sencillez en su teoría de residuos cúbicos y bicuadráticos, Gauss hizo uso de enteros complejos, esto es, números de la forma $a + bi$ con a y b enteros o cero. En el trabajo de Gauss sobre residuos bicuadráticos fue necesario considerar el caso donde el módulo p es primo de la forma $4n + 1$ y Gauss necesitaba los factores complejos en los cuales los números primos de la forma $4n + 1$ podían ser descompuestos. Para obtener éstos, Gauss se percató de que se debía ir más allá del dominio de los enteros ordinarios para introducir los enteros complejos. A pesar de que Euler y Lagrange habían introducido tales enteros en la teoría de números, fue Gauss quien les dio su importancia.

Mientras que en la teoría ordinaria de los enteros las unidades son $+1$ y -1 , en la teoría de Gauss de enteros complejos las unidades son ± 1 y $\pm i$. A un entero complejo se le llama compuesto si es el producto de dos de tales enteros, donde ninguno es una unidad. Si tal descomposición no es posible, el entero es llamado primo. Así, $5 = (1 + 2i)(1 - 2i)$ es compuesto, mientras que 3 es un primo complejo.

Gauss demostró que los enteros complejos tienen esencialmente las mismas propiedades que los enteros ordinarios. Euclides había demostrado (cap. 4, sec. 7) que cada entero es descomponible unívocamente en un producto de primos. Gauss demostró que esta descomposición única, a la que se llama frecuentemente el teorema fundamental de la aritmética, se mantiene también para enteros

complejos siempre y cuando no pensemos en las cuatro unidades como factores diferentes. Esto es, si $a = be = (ib) (-ic)$, las dos descomposiciones son la misma. Gauss demostró igualmente que el proceso de Euclides para encontrar el máximo común divisor de dos enteros es aplicable a los enteros complejos.

Muchos teoremas para primos ordinarios siguen siendo válidos para los primos complejos. Así, el teorema de Fermat se extiende en la forma: si p es un primo complejo $a + bi$, y k cualquier entero complejo no divisible por p , entonces

$$k^{Np-1} \equiv 1 \text{ módulo } p$$

donde Np es la norma $(a^2 + b^2)$ de p . También existe una ley de reciprocidad cuadrática para enteros complejos, establecida por Gauss en su ensayo de 1828.

En términos de los enteros complejos, Gauss fue capaz de enunciar con simplicidad la ley de reciprocidad bicuadrática. Un entero no par se define como uno no divisible entre $1 + i$. Un entero primariamente no par es un entero no par $a + bi$ tal que b es par y $a + b - 1$ es par. Así, -7 y $-5 + 2i$ son números primariamente no pares. La ley de reciprocidad para residuos bicuadráticos establece que si α y β son dos primos primariamente no pares y A y B son sus normas, entonces

$$\left(\frac{\alpha}{\beta}\right) = (-1)^{(1/4)(A-1)(1/4)(B-1)} \left(\frac{\beta}{\alpha}\right)_4$$

El símbolo $(\alpha/\beta)_4$ tiene el siguiente significado: si p es cualquier primo complejo y k es cualquier residuo bicuadrático no divisible por p , entonces $(k/p)_4$ es la potencia de i^e de i , que satisface la congruencia

$$k^{(Np-1)/4} \equiv i \text{ módulo } p$$

donde Np denota la norma de p . Esta ley es equivalente al enunciado: los caracteres bicuadráticos de dos números primos primarios no pares uno con respecto al otro son idénticos, esto es $(\alpha/\beta)_4 = (\beta/\alpha)_4$ pero si ninguno de los primos satisface la congruencia, entonces los dos caracteres bicuadráticos son opuestos, es decir $(\alpha/\beta)_4 = -(\beta/\alpha)_4$.

Gauss enunció este teorema de reciprocidad, pero no publicó su demostración. Esta fue proporcionada por Jacobi en sus clases en Königsberg en 1836-1837. Ferdinand Gotthold Eisenstein (1823-1852), un alumno de Gauss, publicó cinco demostraciones de la ley, de las cuales las dos primeras aparecieron en 1844.⁵

Para la reciprocidad cúbica, Gauss encontró que podía obtener una ley usando los «enteros» $a + b\rho$, donde ρ es una raíz de $x^2 + x + 1 = 0$ y a y b son dos enteros (rationales) ordinarios, pero Gauss no publicó este resultado, el cual se halló entre sus papeles después de su muerte. La ley de reciprocidad cúbica la establece por primera vez Jacobi⁶ y es demostrada por él en sus clases en Königsberg. La primera demostración publicada se le debe a Eisenstein.⁷ Al conocer

esta demostración, Jacobi⁸ reclamó que era precisamente la que él había dado en sus clases, pero Eisenstein negó indignadamente cualquier plagio.⁹ También hay leyes de reciprocidad para congruencias de grado mayor que cuatro.

3. Números algebraicos

La teoría de enteros complejos es un paso en dirección de un dominio muy vasto, la teoría de los números algebraicos. Ni Euler ni Lagrange vieron las ricas posibilidades que abrió su trabajo con enteros complejos. Tampoco lo hizo Gauss.

La teoría surgió de los intentos por demostrar la aserción de Fermat acerca de $x^n + y^n = z^n$. Los casos $n = 3, 4$ y 5 ya han sido discutidos (cap. 25, sec. 4). Gauss intentó demostrar la aserción para $n = 7$, pero falló. Tal vez porque estaba a disgusto con su fracaso, dijo en una carta de 1816 a Heinrich W. M. Olbers (1758-1840): *«Confieso, por supuesto, que el teorema de Fermat como una proposición aislada tiene poca importancia para mí, ya que multitud de tales proposiciones, que uno no puede demostrar o refutar, pueden ser fácilmente formuladas.»*

Este caso particular de $n = 7$ fue resuelto por Lamé en 1839¹⁰, y Dirichlet estableció la aserción para $n = 14$.¹¹

Sin embargo, la proposición general quedaba sin probar.

Ernst Eduard Kummer (1810-1893), quien se desvió de la teología a las matemáticas, se impuso como tarea probarla. Se convirtió en alumno de Gauss y Dirichlet y más tarde trabajó como profesor en Breslau y Berlín. A pesar de que su trabajo principal versó sobre

teoría de números, hizo bellos descubrimientos en geometría que tuvieron su origen en problemas ópticos; también realizó importantes contribuciones al estudio de la refracción de la luz en la atmósfera.

Kummer tomó $x^p + y^p$, donde p es primo, y lo factorizó en

$$(x + y)(x + ay) \dots (x + a^{p-1}y),$$

donde a es una raíz imaginaria p -ésima de la unidad. Ello significa que a es una raíz de

$$a^{p-1} + a^{p-2} + \dots + a + 1 = 0. \quad (1)$$

Esto lo llevó a extender la teoría de enteros complejos de Gauss a los números algebraicos en tanto que éstos son introducidos por ecuaciones tales como (1), esto es, números de la forma

$$f(a) = a_0 + a_1a + \dots + a_{p-2}a^{p-2},$$

donde cada a_i es un entero (racional) ordinario. (Ya que a satisface (1), los términos en a^{p-1} pueden ser reemplazados por términos con menores potencias). Kummer llamó a los números $f(a)$ enteros complejos.

En 1843, Kummer dio definiciones apropiadas de enteros, enteros primos, divisibilidad, y así sucesivamente (proporcionaremos las definiciones habituales dentro de un momento) y entonces cometió

el error de suponer que la factorización única se mantiene en la clase de los números algebraicos que él había introducido. Señaló, mientras le transmitía su manuscrito a Dirichlet en 1843, que esta suposición era necesaria para demostrar el teorema de Fermat. Dirichlet le informó de que la factorización única es válida sólo para ciertos primos p . Incidentalmente, Cauchy y Lamé cometieron el mismo error de suponer la factorización única para los números algebraicos. En 1844, Kummer¹² reconoció lo correcto de la crítica de Dirichlet.

Para restaurar la factorización única, Kummer creó la teoría de números ideales en una serie de ensayos iniciados en 1844.¹³ A fin de entender su idea consideremos el dominio de $a + b\sqrt{-5}$, donde a y b son enteros. En este dominio

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

y se demuestra fácilmente que los cuatro factores son enteros primos. Por tanto la descomposición única no tiene lugar. Introduzcamos, para este dominio, los números ideales $\alpha = \sqrt{2}$, $\beta_1 = (1 + \sqrt{-5})/\sqrt{2}$, $\beta_2 = (1 - \sqrt{-5})/\sqrt{2}$. Vemos que $6 = \beta_1 \cdot \beta_2$

Entonces 6 está ahora expresado unívocamente como el producto de cuatro factores, todos números ideales en cuanto se refiere al dominio $a + b\sqrt{-5}$.¹⁴

En términos de estos ideales y otros primos la factorización en el dominio es única (aparte de los factores formados por unidades). Con los números ideales es posible demostrar algunos de los

resultados de la teoría de números ordinaria en todos los dominios que previamente carecían de factorización única.

Los números ideales de Kummer, aunque números ordinarios, no pertenecen a la clase de números algebraicos que él había introducido. Más aún, los números ideales no fueron definidos en ninguna forma general. En cuanto se refiere al teorema de Fermat, Kummer tuvo éxito al demostrar con sus números ideales que era cierto para algunos números primos. De los primeros cien enteros, sólo el 37, 59 y 67 no eran cubiertos por la demostración de Kummer. Entonces, en su ensayo de 1857,¹⁵ extendió sus resultados a estos primos excepcionales. Tales resultados los amplió aún más Dimitri Mirimanoff (1861-1945), profesor de la universidad de Ginebra, quien perfeccionó los métodos de Kummer.¹⁶ Mirimanoff demostró que el teorema de Fermat es cierto para todo n hasta 256 si x , y y z son primos con ese exponente n .

Mientras que Kummer trabajó con números algebraicos formados a partir de las raíces de la unidad, Richard Dedekind (1831-1916), un alumno de Gauss —quien dedicó cincuenta años de su vida a enseñar en un instituto técnico de Alemania— estudió el problema de la factorización única de una manera enteramente nueva y fresca. Dedekind publicó sus resultados en el suplemento 10 a la segunda edición de la *Zaklentheorie (Teoría de Números)* de Dirichlet, que Dedekind editó. Difundió estos resultados en los suplementos de las tercera y cuarta ediciones del mismo libro.¹⁷ Allí creó la teoría moderna de los números algebraicos.

La teoría de Dedekind de los números algebraicos es una

generalización de los enteros complejos de Gauss y los números algebraicos de Kummer, pero la generalización está de alguna manera divorciada de los enteros complejos de Gauss. Un número r que es una raíz de

$$a_0x^n + a_1x^{n-1} + \dots + a_{n-1}x + a_n = 0, \quad (2)$$

donde las a_i son enteros ordinarios (positivos o negativos), y que no es una raíz de una ecuación de grado menor que n es llamado un número algebraico de grado n . Si el coeficiente de la máxima potencia de x en (2) es 1, las soluciones son llamadas enteros algebraicos de grado n . La suma, diferencia y producto de enteros algebraicos son enteros algebraicos, y si un entero algebraico es un número racional, es un entero ordinario.

Debemos señalar que con las nuevas definiciones un entero algebraico puede contener fracciones ordinarias. Así $(-3 + \sqrt{-115})/2$ es un entero algebraico de segundo grado porque se trata de una raíz de $x^2 + 13x + 71 = 0$. Por otro lado, $(1 - \sqrt{-5})/2$ es un número algebraico de grado 2 pero no un entero algebraico, porque es una raíz de $2x^2 - 2x + 3 = 0$.

Dedekind introdujo después el concepto de cuerpo de números, que es una colección F de números reales o complejos tal que si α y β pertenecen a F , entonces también lo hacen $\alpha + \beta$, $\alpha - \beta$, $\alpha\beta$ y, si $\beta \neq 0$, α/β . Todo cuerpo de números contiene a los números racionales, ya que si α pertenece, entonces también α/α ó 1 y consecuentemente $1 + 1$, $1 + 2$, y de ahí en adelante. No es difícil

demostrar que el conjunto de todos los números algebraicos forma un cuerpo.

Si se empieza con el cuerpo de los números racionales y θ es un número algebraico de grado n , entonces el conjunto formado por la combinación de θ consigo mismo y los números racionales bajo las cuatro operaciones es también un cuerpo de grado n . Este cuerpo puede ser descrito alternativamente como el cuerpo mínimo conteniendo los números racionales y θ . También se le denomina un cuerpo de extensión de los números racionales. Tal cuerpo no consiste en todos los números algebraicos y es un campo de números algebraicos específico. La notación $R(\theta)$ es ahora común. Aunque fuera razonable esperar que los miembros de $R(\theta)$ son los cocientes $f(\theta)/g(\theta)$, donde $f(x)$ y $g(x)$ son polinomios cualesquiera con coeficientes racionales, se puede demostrar que si θ es de grado n , entonces cualquier miembro a de $R(\theta)$ se expresa en la forma

$$a = a_0\theta^{n-1} + a_1\theta^{n-2} + \dots + a_{n-1}$$

donde las a_i son números racionales ordinarios. Más aún, existen enteros algebraicos $\theta_1, \theta_2, \dots, \theta_n$ de este cuerpo tales que los enteros algebraicos del cuerpo son de la forma

$$A_1\theta_1 + A_2\theta_2 + \dots + A_n\theta_n$$

donde las A_i son enteros ordinarios positivos y negativos.

Un anillo, concepto introducido por Dedekind, es esencialmente

cualquier colección de números tales que si α y β pertenecen, lo hacen asimismo $\alpha + \beta$, $\alpha - \beta$, $\alpha \cdot \beta$. El conjunto de todos los enteros algebraicos forma un anillo, al igual que lo hace el conjunto de todos los enteros algebraicos de cualquier cuerpo de números algebraicos específico.

Se dice que el entero algebraico α es divisible por el entero algebraico β si existe un entero algebraico γ tal que $\alpha = \beta\gamma$. Si γ es un entero algebraico que divide cualquier otro entero de un cuerpo de números algebraicos, entonces γ es llamada una unidad de ese cuerpo. Estas unidades, que incluyen al $+1$ y -1 , son una generalización de las unidades $+1$ y -1 de la teoría de números ordinaria. El entero algebraico α es primo si no es cero o una unidad y si cualquier factorización de α en $\beta\gamma$, donde β y γ pertenecen al mismo cuerpo de números algebraicos, implica que β o γ es una unidad en ese cuerpo.

Vemos ahora hasta qué punto se mantiene el teorema fundamental de la aritmética. En el anillo de *todos* los enteros algebraicos no hay primos. Consideremos el anillo de los enteros en un cuerpo $R(\theta)$ específico de números algebraicos, digamos el cuerpo $a + b\sqrt{-5}$, donde a y b son dos números racionales ordinarios. En este cuerpo la factorización única no se mantiene. Por ejemplo,

$$21 = 3 \cdot 7 = (4 + \sqrt{-5})(4 - \sqrt{-5}) = (1 + 2\sqrt{-5})(1 - 2\sqrt{-5})$$

Cada uno de estos últimos cuatro factores es primo en el sentido de que no se expresa como un producto de la forma $(c + d\sqrt{-5})(e + f\sqrt{-5})$

con c, d, e y f enteros.

Por otra parte, consideremos el cuerpo $a + b\sqrt{6}$, donde a y b son números racionales ordinarios. Si se aplican las cuatro operaciones algebraicas a estos números se obtienen números de la misma clase. Si a y b son restringidos a los enteros, se obtienen los enteros algebraicos (de grado 2) de este dominio. En este dominio podemos tomar como definición equivalente de unidad que el entero algebraico M es una unidad si $1/M$ es también un entero algebraico. Así $1, -1, 5 - 2\sqrt{6}$ y $5 + 2\sqrt{6}$ son unidades. Todo entero es divisible por cualquiera de las unidades. Además, un entero algebraico del dominio es primo si es divisible únicamente por él mismo y las unidades. Ahora,

$$6 = 2 \cdot 3 = \sqrt{6} \cdot \sqrt{6}.$$

Parecería como si no existiera una descomposición única en primos. Pero los factores mostrados no son primos. De hecho

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

Cada uno de los últimos cuatro factores es primo en el dominio y la descomposición única se cumple en el dominio.

En el anillo de los enteros de un cuerpo de números algebraicos la factorización específica de los enteros algebraicos en primos siempre es posible, pero la factorización *única* no siempre lo es. De hecho, para los dominios de la forma $a + b\sqrt{-D}$, donde D puede tener

cualquier valor entero positivo no divisible por un cuadrado, el teorema de factorización única es válido sólo cuando $D = 1, 2, 3, 7, 11, 19, 43, 67$ y 163 , al menos para D hasta 10^9 .¹⁸ Así, los propios números algebraicos no poseen la propiedad de la factorización única.

4. Los ideales de Dedekind

Después de generalizar la noción de número algebraico, Dedekind se dedicó a restaurar la factorización única en los cuerpos de números algebraicos mediante un esquema bastante diferente al de Kummer. En lugar de números ideales, Dedekind introdujo las clases de números algebraicos que llamó ideales, en honor a los números ideales de Kummer.

Antes de definir los ideales de Dedekind, señalemos cuál es el pensamiento subyacente. Considérense los enteros ordinarios. En lugar del entero 2, Dedekind toma la clase de los enteros $2m$, donde m es cualquier entero. Esta clase consiste de todos los enteros divisibles por 2. De la misma manera, el 3 es reemplazado por la clase de todos los enteros $3n$ divisibles por 3. El producto 6 se convierte en la colección de todos los números $6p$, donde p es cualquier entero. Entonces el producto $2 \cdot 3 = 6$ es reemplazado por la afirmación de que la clase $2m$ «multiplicada» por la clase $3n$ es igual a la clase $6p$. Más aún, la clase $2m$ es un factor de la clase $6p$, a pesar del hecho de que la clase anterior contiene a la última. Estas clases son ejemplos en el anillo de los enteros ordinarios de lo que Dedekind llama ideales. Para seguir el trabajo de Dedekind, uno

debe acostumbrarse a pensar en términos de clases de números.

Con mayor generalidad, Dedekind definió los ideales como sigue: sea A un cuerpo de números algebraicos específico. Un conjunto de enteros A de K se dice que forma un ideal si cuando α y β son dos enteros cualesquiera en el conjunto, los enteros $\mu\alpha$ y $\nu\beta$, donde μ y ν son otros dos enteros cualesquiera en A , también pertenecen al conjunto. Alternativamente, un ideal A se dice que es generado por los enteros algebraicos $\alpha_1, \alpha_2, \dots, \alpha_n$ de A si A consiste en todas las sumas

$$\lambda_1\alpha_1 + \lambda_2\alpha_2 + \dots + \lambda_n\alpha_n$$

donde los λ_i son enteros cualesquiera de un cuerpo K . Este ideal es denotado mediante $(\alpha_1, \alpha_2, \dots, \alpha_n)$. El ideal cero consiste en el número 0 únicamente y por consiguiente es denotado por (0) . El ideal unidad es el generado por el número 1, esto es, (1) . Un ideal A es llamado principal si es generado por un entero único α , de tal forma que (α) consta de todos los enteros algebraicos divisibles por α . En el anillo de los enteros algebraicos todo ideal es un ideal principal.

Un ejemplo de un ideal en el campo $a + b\sqrt{5}$ de los números algebraicos, donde a y b son números racionales ordinarios, es el ideal generado por los enteros 2 y $1 + \sqrt{-5}$. Este ideal consiste en todos los enteros de la forma $2\mu + (1 + \sqrt{-5})\nu$, donde μ y ν son enteros arbitrarios del cuerpo. El ideal también es un ideal principal, ya que está generado sólo por el número 2 en virtud del hecho de que $(1 + \sqrt{-5})2$ debe pertenecer también al ideal generado

por 2.

Dos ideales $(\alpha_1, \alpha_2, \dots, \alpha_p)$ y $(\beta_1, \beta_2, \dots, \beta_q)$ son iguales si cada miembro del primer ideal es un miembro del segundo y conversamente. Para atacar el problema de la factorización debemos primero considerar el producto de dos ideales. El producto del ideal $A = (\alpha_1, \alpha_2, \dots, \alpha_s)$ y el ideal $B = (\beta_1, \beta_2, \dots, \beta_t)$ de K está definido como el ideal

$$AB = (\alpha_1\beta_1, \alpha_1\beta_2, \alpha_2\beta_1, \dots, \alpha_i\beta_j, \dots, \alpha_s\beta_t).$$

Es casi evidente que este producto es conmutativo y asociativo. Con la definición podemos decir que A divide a B si existe un ideal C tal que $B = AC$. Se escribe A/B y A es llamado un factor de B . Como ya se había sugerido mediante nuestro ejemplo de los enteros ordinarios, los elementos de B están *incluidos* en los elementos de A y la divisibilidad ordinaria es reemplazada por la inclusión de clases.

Los ideales que son los análogos de los números primos ordinarios son llamados ideales primos. Tal tipo de ideal P es definido como uno que no tiene factores que no sean él mismo y el ideal (1) , de tal forma que P no está contenido en ningún otro ideal de K . Por esta razón un ideal primo es también llamado maximal. Todas estas definiciones y teoremas condujeron a los teoremas básicos para ideales de un cuerpo de números algebraicos K . Cualquier ideal es divisible únicamente por un número finito de ideales y si un ideal primo divide el producto AB de dos ideales (de la misma clase de

números) divide a A o a B . Finalmente, el teorema fundamental en la teoría de ideales es que todo ideal puede ser factorizado unívocamente en ideales primos.

En nuestros primeros ejemplos de cuerpos de números algebraicos de la forma $a + b\sqrt{D}$, siendo entero D , encontramos que algunos permitían la factorización única de los enteros algebraicos de aquellos cuerpos y otros no. La respuesta a la pregunta de cuáles lo permiten y cuáles no está dada por el teorema de que la factorización de los enteros de un cuerpo K de números algebraicos en primos es única si y sólo si todos los ideales de K son principales.

A partir de estos ejemplos del trabajo de Dedekind podemos ver que su teoría de los ideales es por supuesto una generalización de los enteros ordinarios. En particular, suministra los conceptos y propiedades en el dominio de los números algebraicos que permiten establecer la factorización única.

Leopold Kronecker (1823-1891), fue el estudiante favorito de Kummer y le sucedió como profesor en la universidad de Berlín; también se dedicó al estudio de los números algebraicos y lo desarrolló siguiendo líneas similares a las de Dedekind. La tesis doctoral de Kronecker «Sobre unidades complejas», escrita en 1845, aunque no publicada sino mucho más tarde,¹⁹ fue su primer trabajo en el tema. La tesis trata de las unidades que pueden existir en los cuerpos de números algebraicos creados por Gauss.

Kronecker creó otra teoría de cuerpos (dominios de racionalidad).²⁰ Su concepto de cuerpo es mucho más general que el de Dedekind,

ya que consideró campos de funciones racionales en cualquier número de variables (indeterminadas). Específicamente, Kronecker introdujo (1881) la noción de una indeterminada adjunta a un cuerpo siendo únicamente la indeterminada una nueva cantidad abstracta. Esta idea de extender un cuerpo añadiendo una indeterminada la convirtió en la piedra angular de su teoría de números algebraicos. Aquí utilizó el conocimiento ya construido desde Liouville, Cantor y otros, sobre la distinción entre los números algebraicos y los trascendentes. En particular, observó que si x es trascendente sobre un cuerpo K (x es una indeterminada) entonces el campo $K(x)$ obtenido al añadir la indeterminada x a K , esto es, el cuerpo más pequeño conteniendo K y x , es isomorfo al cuerpo $A[x]$ de funciones racionales de una variable con coeficientes en K .²¹ Subrayó que la indeterminada era únicamente un elemento de un álgebra y no una variable en el sentido del análisis.²² Más adelante, en 1887,²³ demostró que a cada número primo ordinario p le corresponde dentro del anillo $A(x)$ de polinomios con coeficientes racionales un polinomio primo $p(x)$ que es irreducible en el campo racional Q . Considerando dos polinomios como iguales si son congruentes módulo un polinomio primo $p(x)$ dado, el anillo de los polinomios en $Q(x)$ se convierte en un cuerpo de clases de restos poseyendo las mismas propiedades algebraicas que el cuerpo de los números algebraicos $K(\delta)$ obtenido del cuerpo K al añadir una raíz δ de $p(x) = 0$. Aquí utilizó la idea que Cauchy ya había empleado para introducir los números imaginarios al utilizar polinomios congruentes módulo $x^2 + 1$. En este mismo trabajo demostró que la

teoría de los números algebraicos es independiente del teorema fundamental del álgebra y de la teoría del sistema completo de los números reales.

En su teoría de cuerpos (en los «Grundzüge»), cuyos elementos están formados empezando con un campo K y adjuntando indeterminadas x_1, x_2, \dots, x_n , Kronecker introdujo la noción de un sistema modular que jugó el papel de los ideales en la teoría de Dedekind. Para Kronecker, un sistema modular es el conjunto M de aquellos polinomios en n variables x_1, x_2, \dots, x_n , tales que si P_1 y P_2 pertenecen al conjunto, también lo hace $P_1 + P_2$, y si P pertenece, también lo hace QP , donde Q es cualquier polinomio en x_1, x_2, \dots, x_n .

Una base del sistema modular M es cualquier conjunto de polinomios B_1, B_2, \dots de M tal que cada polinomio de M es expresable en la forma

$$R_1B_1 + R_2B_2 + \dots,$$

donde R_1, R_2, \dots son constantes o polinomios (no perteneciendo necesariamente a M). La teoría de la divisibilidad en los cuerpos generales de Kronecker fue definida en términos de sistemas modulares, un poco como lo había hecho Dedekind con los ideales.

El trabajo sobre la teoría algebraica de números llegó a su clímax en el siglo XIX con el famoso artículo de Hilbert sobre números algebraicos.²⁴ Este trabajo es esencialmente una descripción de lo que se había hecho durante el siglo. Sin embargo, Hilbert retrabajó toda esta teoría anterior y proporcionó nuevos, elegantes y

poderosos métodos para asegurar sus resultados. Empezó a crear nuevas ideas en la teoría algebraica de números alrededor de 1892 aproximadamente, y una de las nuevas creaciones sobre los cuerpos de números galoisianos fue incorporada también en el trabajo. Posteriormente, Hilbert y muchos otros autores extendieron ampliamente la teoría de números algebraicos. Sin embargo, estos últimos desarrollos, relativos a cuerpos galoisianos, cuerpos de números abelianos relativos y cuerpos de clases, cada uno de ellos estimulando una inmensa cantidad de trabajo en el siglo XX, son de interés principalmente para los especialistas.

La teoría algebraica de números, que originalmente era un esquema para investigar las soluciones de los problemas de la vieja teoría de números, se ha convertido en un fin en sí misma. Ha llegado a ocupar una posición entre la teoría de números y el álgebra abstracta, y ahora la teoría de números y el álgebra superior moderna se mezclan en la teoría algebraica de números. Por supuesto, la teoría algebraica de números también ha producido nuevos teoremas en la teoría ordinaria de números.

5. La teoría de las formas

Otra clase de problemas en la teoría de números es la representación de los enteros por medio de formas. La expresión

$$ax^2 + 2bxy + cy^2 \quad (3)$$

donde a , b y c son enteros, es una forma binaria ya que hay dos

variables implicadas y es una forma cuadrática porque es de segundo grado. Se dice que un número M está representado por la forma si para valores enteros concretos de a, b, c, x e y , la expresión anterior es igual a M . Uno de los problemas es *encontrar* el conjunto de números M que son representables por una forma dada o clase de formas. El problema inverso, dado M y dados a, b y c , o alguna clase de a, b y c , encontrar los valores de x e y que representa M , es igualmente importante. El último problema pertenece al análisis diofántico y el anterior puede ser igualmente considerado parte de la misma materia.

Euler había obtenido algunos resultados particulares sobre estos problemas. Sin embargo, Lagrange hizo el descubrimiento clave de que si un número es representable por una forma, es también representable por muchas otras formas, que llamó equivalentes. Las últimas podían ser obtenidas a partir de la forma original mediante un cambio de variables

$$x = \alpha x' + \beta y', \quad y = \gamma x' + \delta y' \quad (4)$$

donde las α, β, γ y δ son enteros y $\alpha\gamma - \beta\delta = 1$ ²⁵. En particular, Lagrange mostró que para un discriminante dado (Gauss usó la palabra determinante) $b^2 - 4ac$ existe un número finito de formas tal que cada forma con ese discriminante es equivalente a una de ese número finito. De esta manera, todas las formas con un discriminante dado pueden ser divididas en clases, consistiendo cada clase en formas equivalentes a un miembro de esa clase. Este

resultado y otros obtenidos inductivamente por Legendre atrajeron la atención de Gauss. En un paso audaz, Gauss extrajo del trabajo de Lagrange la noción de equivalencia de formas y se concentró en ella. La quinta sección de sus *Disquisitiones*, con mucho la sección más larga, está dedicada a este tema.

Gauss sistematizó y extendió la teoría de las formas. Primero definió la equivalencia de formas. Sea

$$F = ax^2 + 2bxy + cy^2$$

transformada por medio de (4) en la forma

$$F = a'x'^2 + 2b'x'y' + c'y'^2.$$

Entonces

$$b'^2 - a'c' = (b^2 - ac)(\alpha\delta - \beta\gamma)^2.$$

Si ahora $(\alpha\delta - \beta\gamma)^2 = 1$, los discriminantes de las dos formas son iguales. Entonces el inverso de la transformación (4) también contendrá coeficientes enteros (por la regla de Cramer) y transformará F en F' . Se dice que F y F' son equivalentes. Si $(\alpha\delta - \beta\gamma)^2 = 1$, se dice que F y F' son propiamente equivalentes, y si $(\alpha\delta - \beta\gamma)^2 = -1$, entonces se dice que F y F' son impropriamente equivalentes.

Gauss demostró un buen número de teoremas sobre la equivalencia

de formas. Por ejemplo, si F es equivalente a F' y F' a F'' , entonces F es equivalente a F'' . Si F es equivalente a F' , cualquier número M representable por F es representable por F' y de tantas maneras por una como por la otra. Más adelante demuestra, si F y F' son equivalentes, cómo encontrar todas las transformaciones de F en F' . También encuentra todas las representaciones de un número dado M por la forma F , siempre que los valores de x y y sean primos entre sí.

Por definición, dos formas equivalentes tienen el mismo valor para su discriminante $D = b^2 - 4ac$. Sin embargo, dos formas con el mismo discriminante no son necesariamente equivalentes. Gauss demuestra que todas las formas con un D dado pueden ser divididas en clases; los miembros de cualquier clase son propiamente equivalentes entre sí. Aunque el número de formas con un D dado es infinito, el número de clases para una D dada es finito. En cada clase puede ser tomada una forma como representativa y Gauss proporciona criterios para la elección de la representación más simple. La forma más simple de todas aquellas con determinante D tiene $a = 1$, $b = 0$ y $c = -D$. A esto llama Gauss la forma principal y la clase a la que pertenece es llamada clase principal.

Gauss, entonces, ataca la composición (producto) de formas. Si la forma

$$F = AX^2 + 2BXY + CY^2$$

se transforma en el producto de dos formas

$$f = ax^2 + 2bxy + cy^2 \quad \text{y} \quad f' = a'x'^2 + 2b'x'y' + c'y'^2$$

por la sustitución

$$X = p_1xx' + p_2xy' + p_3x'y + p_4yy'$$

$$Y = q_1xx' + q_2xy' + q_3x'y + q_4yy'$$

entonces se dice que F es transformable en ff' . Si además los seis números

$$p_1q_2 - q_1p_2, \quad p_1q_3 - q_1p_3, \quad p_1q_4 - q_1p_4, \quad p_2q_3 - q_2p_3, \quad p_2q_4 - q_2p_4, \quad p_3q_4 - q_3p_4$$

no tienen un divisor común, entonces se dice que F es compuesta de las formas f y f' .

Más adelante, Gauss demostró un teorema esencial: si f y g pertenecen a la misma clase y f' y g' pertenecen a la misma clase, entonces la forma compuesta de f y f' pertenecerá a la misma clase que la forma compuesta de g y g' . Entonces se puede hablar de una clase de formas compuesta de dos (o más) clases dadas. En esta composición de clases, la clase principal actúa como una clase unidad; esto es, si la clase K es compuesta con la clase principal, la clase resultante será K .

Gauss pasa ahora al tratamiento de las formas cuadráticas ternarias

$$Ax^2 + 2 Bxy + Cy^2 + 2 Dxz + 2 Eyz + Fz^2,$$

donde los coeficientes son enteros y realiza un estudio muy parecido al que había llevado a cabo para formas binarias. La meta, como en el caso de las formas binarias, es la representación de los enteros. Gauss no desarrolló la teoría de las formas ternarias.

El objetivo del trabajo entero de la teoría de formas fue, como se mencionó con anterioridad, obtener teoremas en teoría de números. En el curso de su tratamiento de las formas, Gauss muestra cómo la teoría puede ser usada para demostrar un gran número de teoremas sobre los enteros incluyendo muchos que habían sido previamente demostrados por autores tales como Euler y Lagrange. Así, Gauss demuestra que cualquier número primo de la forma $3n + 1$ puede ser representado como una suma de cuadrados de una manera única. Cualquier número primo de la forma $8n + 1$ o $8n + 3$ puede ser representado en la forma $x^2 + 2y^2$ (para enteros positivos x e y) en una y sólo una manera. Muestra cómo encontrar todas las representaciones de un número dado M mediante la forma dada $ax^2 + 2bxy + cy^2$, siempre que el discriminante D sea un número positivo que no es un cuadrado. Además, si F es una forma primitiva (los valores de a , b y c son primos entre sí) con el discriminante D y si p es un número primo que divide a D , entonces los números no divisibles por p que pueden ser representados por F coinciden en

que todos ellos o son residuos cuadráticos de p o no residuos de p . Entre los resultados que Gauss dedujo de su trabajo sobre formas cuadráticas ternarias está la primera prueba del teorema de que cada número puede ser representado como suma de tres números triangulares. Estos, recordamos, son los números

$$1, 3, 6, 10, 15, \dots, (n^2+n)/2, \dots$$

También volvió a demostrar el teorema ya probado por Lagrange de que cualquier entero positivo puede ser expresado como la suma de cuatro cuadrados. A propósito del primer resultado, es importante notar que Cauchy leyó un ensayo ante la Academia de París en 1815 donde establecía el resultado general (afirmado primero por Fermat) de que cada entero es la suma de k , o menos, números k -gonales.²⁶ (El número k -gonal general es $n + (n^2 - n)(k - 2)/2$).

La teoría algebraica de las formas cuadráticas binarias y ternarias —como la presentó Gauss— tiene una analogía geométrica interesante que inició el propio Gauss. En una reseña, que apareció en el *Göttingische Gelehrte Anzeigen* de 1830,²⁷ de un libro sobre formas cuadráticas ternarias, cuyo autor fuera Ludwig August Seeber (1793-1855), Gauss bosquejó la representación geométrica de las formas y de las clases de formas.²⁸ Este trabajo es el inicio de una rama llamada la teoría geométrica de números, que ganó primero importancia cuando Hermann Minkowski (1864-1909), quien trabajó como profesor en varias universidades, publicó su *Geometrie der Zahlen (Geometría de los Números, 1896)*.

El campo de las formas adquirió importancia en la teoría de números del siglo XIX. Trabajo posterior sobre formas cuadráticas binarias y ternarias y sobre formas con más variables y de grado superior,²⁹ fue realizado por multitud de autores.

6. Teoría analítica de números

Uno de los principales avances en la teoría de números es la presentación de métodos analíticos, y de resultados analíticos, para expresar y demostrar hechos acerca de los enteros. De hecho, Euler había usado el análisis en la teoría de números (véase más adelante) y Jacobi utilizó funciones elípticas para obtener resultados en la teoría de congruencias y en la teoría de formas.³⁰ Sin embargo, los usos de Euler del análisis en la teoría de números fueron menores y los resultados de Jacobi, casi productos accidentales de su trabajo en análisis.

El primer uso deliberado y profundo del análisis para atacar lo que parecía ser un claro problema de álgebra fue realizado por Peter Gustav Lejeune-Dirichlet (1805-1859), alumno de Gauss y Jacobi, profesor en Breslau y Berlín, y más adelante sucesor de Gauss en Göttingen. El gran libro de Dirichlet, *Vorlesungen über Zahlentheorie (Lecturas sobre la teoría de números)*,³¹ explicó las *Disquisitiones* de Gauss y dio sus propias contribuciones.

El problema que llevó a Dirichlet a emplear el análisis fue demostrar que cada sucesión aritmética

$$a, a + b, a + 2b, a + 3b, \dots, a + nb, \dots,$$

donde a y b son primos entre sí, contiene un número infinito de primos. Euler³² y Legendre³³ hicieron esta conjetura y en 1808 Legendre³⁴ proporcionó una demostración que era errónea. En 1837, Dirichlet³⁵ dio una demostración correcta. Este resultado generaliza el teorema de Euclides sobre la infinitud de primos en la sucesión 1, 2, 3, ... La demostración analítica de Dirichlet era larga y complicada. Específicamente, usaba lo que ahora son llamadas las series de Dirichlet,

$$\sum_{n=1}^{\infty} a_n n^{-z}$$

donde las a_n y z son complejos. Dirichlet también demostró que la suma de los recíprocos de los primos en la sucesión $a + nb$ diverge. Esto extiende un resultado de Euler sobre los primos usuales (véase más abajo). En 1841³⁶ Dirichlet demostró un teorema sobre los primos en progresiones de números complejos $a + bi$.

El problema principal que implicaba la introducción del análisis concernía a la función que representa el número de primos que no exceden a x . Así $\pi(8)$ es 4 ya que 2, 3, 5 y 7 son primos, y $\pi(11)$ es 5. Cuando x se incrementa, los primos adicionales se hacen escasos y el problema era: ¿cuál es la expresión analítica para $\pi(x)$? Legendre, quien había demostrado que ninguna expresión racional puede servir, en algún momento perdió la esperanza de que pudiera ser encontrada cualquier expresión. Entonces Euler, Legendre, Gauss y

otros supusieron que

$$\lim_{x \rightarrow \infty} \frac{\pi(x)}{\frac{x}{\log x}} = 1 \quad (5)$$

Gauss utilizó tablas de primos (de hecho estudió todos los primos hasta 3.000.000) para hacer conjeturas acerca de $\pi(x)$ e infirió³⁷ que $\pi(x)$ difiere poco de

$$\int_2^x \frac{dt}{\log t}$$

También sabía que

$$\log_{x \rightarrow \infty} \frac{\int_2^x \frac{dt}{\log t}}{\frac{x}{\log x}}$$

En 1848, Pafnuti L. Tchebycheff (1821-1894), profesor de la universidad de Petrogrado, abordó la cuestión del número de números primos menor o igual a x y dio un paso importante hacia adelante en este viejo problema. En un ensayo clave, «Sur les nombres premiers» («Sobre los números primos»),³⁸ Tchebycheff demostró que

$$A_1 < \frac{\pi(x)}{\log x} < A_2$$

donde $0,922 < A_1 < 1$ y $1 < A_2 < 1.1050$, pero no demostró que la función tiende a un límite. Esta desigualdad fue mejorada por muchos matemáticos, incluyendo a Sylvester, quien, entre otros, dudaba en 1881 de que la función tuviera un límite. En su trabajo, Tchebycheff usó lo que ahora conocemos como la función zeta de Riemann,

$$\zeta(z) = \sum_{n=1}^{\infty} \frac{1}{n^z}$$

aunque la utilizó para valores reales de z . (Esta serie es un caso especial de la serie de Dirichlet.) Incidentalmente, también demostró en el mismo ensayo que para $n > 3$ siempre existe al menos un primo entre n y $2n - 2$.

La función zeta para z real aparece en el trabajo de Euler³⁹, donde introdujo

$$\zeta(s) = \sum_{n=1}^{\infty} \frac{1}{n^s} = \prod_{n=1}^{\infty} \left(1 - \frac{1}{p_n^s}\right)^{-1}$$

Aquí los p_n son números primos. Euler utilizó la función para demostrar que la suma de los recíprocos de los números primos

diverge. Para s entero par positivo, Euler conocía el valor de $\zeta(s)$ (cap. 20, sec. 4). Entonces, en un artículo leído en 1749⁴⁰, Euler aseguró que para s real

$$\zeta(1-s) = 2(2\pi)^{-s} \cos \frac{\pi s}{2} \Gamma(s) \zeta(s)$$

Verificó la ecuación hasta el punto de que, decía, no había duda acerca de ella. Esta relación fue establecida por Riemann en el artículo de 1859, al que nos habíamos referido con anterioridad. Riemann, usando la función zeta para z complejo, intentó demostrar el teorema de los números primos, esto es, (5) antes mencionado.⁴¹ Señalo que, para llevar más lejos la investigación, se tenían que conocer los zeros complejos de $\zeta(z)$. De hecho, $\zeta(z)$ no converge para $x \leq 1$ cuando $z = x + iy$, pero los valores de ζ en la mitad del plano $x \leq 1$ están definidos por continuación analítica. Expresó la hipótesis de que todos los zeros en la banda $0 \leq x \leq 1$ están sobre la línea $x = 1/2$. Esta hipótesis sigue sin ser demostrada.⁴²

En 1896, Hadamard,⁴³ aplicando la teoría de funciones enteras (de una variable compleja), que investigó con el propósito de demostrar el teorema de los números primos y probando el hecho crucial que $\zeta(z) \neq 0$ para $x = 1$, estuvo finalmente en condiciones de demostrar el teorema de los números primos. Charles-Jean de la Vallée Poussin (1866-1962) obtuvo el resultado para la función zeta y demostró el teorema de los números primos en el mismo momento.⁴⁴ Este teorema es central en la teoría analítica de números.

Bibliografía

- Bachmann, P.: «Über Gauss' zahlentheoretische Arbeiten». Nachbncbten Köning. Ges. der Wiss. zh Gótt., 1911, 455-508; también en Gauss: Werke, 10, 1-69.
- Bell, Eric T.: The development of mathematics, 2^a ed., McGraw-Hill, 1945, caps. 9-10.
- Carmichael, Robert D.: «Some recent researches in the theory of numbers». Amer. Math. Monthly, 39, 1932, 139-160.
- Dedekind, Richard: Über die Theorie derganzen algebraischen Zahlen (reimpreso del undécimo suplemento del Zahlentheorie de Dirichlet), F. Vieweg und Sohn, 1964: -Gesammelte mathematische Werke, 3 vols., F. Vieweg und Sohn, 1930-1932, Chelsea (reimpresión), 1968: -«Sur la théorie des nombres entiers algébriques», Bull. des Sci. Math., (1), 11, 1876, 278-288; (2), 1, 1877, 17-41, 69-92, 144-164, 207-248 = Ges. math. Werke, 3, 263-296.
- Dickson, Leonard E.: History of the theory of numbers, 3 vols., Chelsea (reimpresión), 1951. Studies in the theory of numbers (1930), Chelsea (reimpresión), 1962: - y cols.: Algebraic numbers, Report of Committee on algebraic numbers, National Research Council, 1923 y 1928; Chelsea (reimpresión), 1967.
- Dirichlet, P. G. L.: Werke (1889-1897); Chelsea (reimpresión), 1969, 2 vols.
- Dirichlet, P. G. L., y Dedekind, R.: Vorlesungen über Zahlentheorie, 4.a ed., 1894 (contiene el suplemento de Dedekind); Chelsea (reimpresión), 1968.

- Gauss, C. F.: *Disquisitiones Arithmeticae*, traducción al inglés A. A. Clarke, Yale University Press, 1965.
- Hasse, H.: «Bericht über neuere Untersuchungen und Probleme aus der Theorie der algebraischen Zahlkörper», *Jahres, der Deut. Math.-Verein*, 35, 1926, 1-55 y 36, 1927, 233-311.
- Hilbert, David: «Die Theorie der algebraischen Zahlkörper», *Jahres, der Deut. Math.-Verein*, 4, 1897, 175-546 = *Gesammelte Abhandlungen*, 1, 63-363.
- Klein, Félix: *Vorlesungen über die Entwicklung der Mathematik im 19. Jahrhundert*, Chelsea (reimpresión), 1950, vol. 1.
- Kronecker, Leopold: *Werke*, 5 vols. (1895-1931), Chelsea (reimpresión), 1968. Véase especialmente vol. 2, pp. 1-10 sobre la ley de reciprocidad cuadrática.
- *Grundzüge einer arithmetischen Theorie der algebraischen Größen*, G. Reimer, 1882 = *Jour. für Math.*, 92, 1881/1882, 1-122 = *Werke*, 2, 237-388.
- Landau, Edmund: *Handbuch der Lehre von der Verteilung der Primzahlen*, B. G. Teubner, 1909, vol. 1, págs. 1-55.
- Mordell, L. J.: «An introductory account of the arithmetical theory of algebraic numbers and its recent development». *Amer. Math. Soc. Bull.*, 29, 1923, 445-463.
- Reichardt, Hans (ed.): *C. F. Gauss, Leben und Werk*, Haude und Spenersche Verlagsbuchhandlung, 1960, pp. 38-91; también B. G. Teubner, 1957.

- Scott, J. F.: A history of mathematics, Taylor and Francis, 1958, cap. 15.
- Smith, David, E.: A source book in mathematics, Dover (reimpresión), 1959, vol. 1, 107-148.
- Smith, H. J. S.: Collected mathematical papers, 2 vols. (1890-1894), Chelsea (reimpresión), 1965. El vol. 1 contiene el Report on the theory of numbers de Smith, que también ha sido publicado por separado en Chelsea, 1965.
- Vandiver, H. S.: «Fermat's last theorem», Amer. Math. Monthly, 53, 1946, 555-578.

Capítulo 35

El resurgimiento de la geometría proyectiva

Las doctrinas de la geometría pura, frecuentemente, y en muchas cuestiones, proporcionan una manera simple y natural de penetrar en los orígenes de las verdades, para aclarar la misteriosa cadena que las une, y para hacerlas conocer individual, luminosa y completamente.

Michael Chasles

Contenido:

- 1. El renovado interés por la geometría*
 - 2. Geometría euclídea sintética*
 - 3. El resurgimiento de la geometría proyectiva sintética*
 - 4. Geometría proyectiva algebraica*
 - 5. Curvas planas de orden superior y superficies*
- Bibliografía.*

1. El renovado interés por la geometría

Por más de 100 años después de la introducción de la geometría analítica por Descartes y Fermat, los métodos algebraicos y analíticos dominaron la geometría, hasta la casi exclusión de los

métodos sintéticos. Durante este período algunos autores, por ejemplo los matemáticos ingleses que persistían en intentar fundar el cálculo rigurosamente sobre la geometría, consiguieron nuevos resultados sintéticamente. Los métodos geométricos, elegantes e intuitivamente claros, siempre cautivaron algunas mentes. Especialmente, Maclaurin prefirió la geometría sintética al análisis. La geometría pura, entonces, retuvo algo de vida aún si no se encontraba en el corazón de los desarrollos más vitales de los siglos XVII y XVIII. Al principio del siglo XIX varios grandes matemáticos decidieron que la geometría sintética había sido rechazada injusta y totalmente e hicieron un esfuerzo positivo por revivir y extender su enfoque.

Uno de los nuevos campeones de los métodos sintéticos, Jean-Victor Poncelet, concedió las limitaciones de la vieja geometría pura. Dice: «Mientras que la geometría analítica ofrece por su método general característico y uniforme medios de proceder a la solución de las cuestiones que se nos presentan... mientras que llega a resultados cuya generalidad no tiene frontera, la otra [geometría sintética] procede por casualidad; su camino depende completamente de la sagacidad de aquellos que la emplean y sus resultados casi siempre están limitados a la figura particular que considera». Pero Poncelet no creía que los métodos sintéticos estuvieran necesariamente tan limitados, y propuso crear nuevos métodos que rivalizarían con el poder de la geometría analítica.

Michel Chasles (1793-1880) fue otro gran defensor de los métodos geométricos. En su *Aperçu historique sur l'origine et le*

développement des méthodes en géométrie (Resumen histórico sobre los orígenes y desarrollo de los métodos en geometría, 1837), un estudio histórico en el cual Chasles admitió que ignoraba a los escritores alemanes porque desconocía su idioma, afirma que los matemáticos de su tiempo, y anteriores, habían declarado la geometría un lenguaje muerto que en el futuro no tendría uso ni influencia. No solamente Chasles niega esto, sino que cita a Lagrange, que era enteramente un analista, afirmando en su sesenta aniversario⁴⁵, cuando había encontrado un problema muy difícil de mecánica celeste: «A pesar de que el análisis pueda tener ventajas sobre los viejos métodos geométricos, los cuales comúnmente, pero indebidamente, llamamos sintéticos, hay sin embargo problemas en el que el último aparece más ventajosamente, en parte debido a su claridad intrínseca y en parte debido a la elegancia y facilidad de sus soluciones. Hay aún algunos problemas para los cuales el análisis algebraico en alguna medida no es suficiente y que, según parece, sólo los métodos sintéticos pueden resolver.» Lagrange cita como ejemplo el muy difícil problema de la atracción que un elipsoide de revolución ejerce sobre un punto (unidad de masa) sobre su superficie o en el interior. Este problema había sido resuelto de modo puramente sintético por Maclaurin.

Chasles también proporciona un extracto de una carta que recibió de Lambert Adolphe Quetelet (1796-1874), el astrónomo y estadístico belga. Quetelet dice: «No es propio que la mayoría de nuestros matemáticos jóvenes valoren la geometría pura tan

ligeramente.» Los jóvenes se quejaban de la carencia de generalidad del método, continúa Quetelet, pero es éste un fallo de la geometría o de aquellos que cultivan la geometría, se pregunta. Para compensar esta carencia de generalidad, Chasles proporciona dos reglas a los futuros geómetras. Deberán generalizar teoremas particulares para obtener los resultados más generales, que deberán ser al mismo tiempo simples y naturales. Segundo, no habrán de estar satisfechos con la demostración del resultado si no es parte de un método general o doctrina de la que depende naturalmente. Saber cuándo uno tiene realmente la base verdadera para un teorema, dice; siempre existe una verdad principal que uno reconocerá porque otros teoremas resultarán de la simple transformación, o como consecuencia inmediata. Las grandes verdades, que son el fundamento del conocimiento, siempre tienen las características de la simplicidad y la intuición.

Otros matemáticos atacaron los métodos analíticos con lenguaje más fuerte. Carnot deseaba «liberar a la geometría de los jeroglíficos del análisis». Más tarde, en el mismo siglo, Eduard Study (1862-1922) se refirió al proceso «mecánico» de la geometría con coordenadas como el «estruendo del molino coordinado».

Las objeciones a los métodos analíticos en geometría estuvieron basadas en algo más que una preferencia o gusto personal. Había, ante todo, la pregunta genuina de si la geometría analítica era realmente geometría, ya que el álgebra era la esencia del método y los resultados, y la significación geométrica de ambas estaba escondida. Más aún, señaló Chasles, el análisis, a través de sus

procesos formales, niega todos los pequeños pasos que continuamente da la geometría. Los rápidos y tal vez penetrantes pasos del análisis no revelan el sentido de lo que se consigue. La conexión entre el punto inicial y el resultado final no es claro. Chasles pregunta: «*¿Es entonces suficiente en un estudio filosófico y básico de una ciencia saber que algo es verdadero si uno no sabe por qué es así y qué lugar debería ocupar en la serie de verdades a la que pertenece?*» El método geométrico, por otro lado, permite pruebas y conclusiones simples y evidentemente intuitivas.

Había otro argumento, el cual, mencionado primero por Descartes, aún gustaba en el siglo XIX. La geometría era considerada como la verdad sobre el espacio y el mundo exterior. El álgebra y el análisis no eran verdades significantes en sí mismas, ni siquiera acerca de números y funciones. Estos consistían únicamente en métodos para llegar a las verdades, y en eso se presentaban como, artificiales. Esta visión del álgebra y del análisis estaba desapareciendo gradualmente. Sin embargo, la crítica era aún rigurosa al principio del siglo XIX, ya que los métodos del análisis eran incompletos y todavía sin base lógica. Los geómetras cuestionaban correctamente la validez de las pruebas analíticas y las aceptaban únicamente como sugiriendo resultados. El análisis sólo podía responder mordazmente que las demostraciones geométricas eran torpes y sin elegancia.

El resultado de la controversia es que los geómetras puros reafirmaban su papel en las matemáticas. Como si se tuvieran que vengar de Descartes porque su creación de la geometría analítica los

había obligado a abandonar la geometría pura, los geómetras de principios del siglo XIX se propusieron como objetivo vencer a Descartes en su juego de la geometría. La rivalidad entre geómetras y analistas se hizo tan amarga que Steiner —un geómetra puro— amenazó con no publicar en el *Journal für Mathematik de Crelle* si éste continuaba publicando los artículos analíticos de Plücker.

El estímulo para revivir la geometría sintética vino principalmente de un solo autor, Gaspard Monge. Ya hemos discutido sus valiosas contribuciones a la geometría analítica y diferencial y sus inspiradoras clases en la Escuela Politécnica durante los años de 1795 a 1809. El mismo Monge no intentaba más que hacer regresar la geometría al rebaño de las matemáticas como un enfoque sugestivo y una interpretación de los resultados analíticos, y buscó únicamente el subrayar ambos modos de pensamiento. Sin embargo, su propio trabajo en geometría y su entusiasmo inspiró a sus alumnos Charles Dupin, François Joseph Servois, Charles Julien Brianchon, Jean Baptiste Biot (1744-1862), Lazare Nicholas Marguerite Carnot y Jean Victor Poncelet, la urgencia por revitalizar la geometría pura.

La contribución de Monge a la geometría pura fue su *Traité de géométrie descriptive* (Tratado de geometría descriptiva, 1799). Esta materia muestra cómo proyectar ortogonalmente un objeto tridimensional en dos planos (uno horizontal y otro vertical) de tal forma que a partir de esta representación sea factible deducir propiedades matemáticas del objeto. El esquema es útil en arquitectura, diseño de fortificaciones, perspectiva, carpintería y

talla de piedras, y fue el primero en tratar la proyección de una figura tridimensional en dos bidimensionales. Las ideas y métodos de la geometría descriptiva no demostraron ser una fuente de desarrollos posteriores en geometría o, en ese aspecto, en cualquier otra parte de las matemáticas.

2. Geometría euclídea sintética

A pesar de que los geómetras que reaccionaron a la inspiración de Monge se dedicaron a la geometría proyectiva, debemos detenernos para señalar algunos nuevos resultados en la geometría euclídea sintética. Estos resultados, tal vez pequeños en significación, muestran sin embargo nuevos temas y la casi infinita riqueza de esta vieja materia. De hecho, cientos de nuevos teoremas fueron obtenidos, de los cuales sólo podemos dar unos cuantos ejemplos.

Asociados con cada triángulo ABC están nueve puntos particulares, los puntos medios de los lados, los pies de las tres alturas y los puntos medios de los segmentos que unen los vértices con los puntos de intersección de las alturas. Los nueve puntos yacen sobre un círculo, llamado el círculo de los nueve puntos. Gergonne y Poncelet publicaron este teorema por primera vez⁴⁶. Se le acredita frecuentemente a Karl Wilhelm Feuerbach (1800-1834), un maestro de escuela, quien publicó su demostración en *Eigenschaften einiger merkwürdigen Punkte des geradlinigen Dreiecks* (Propiedades de algunos puntos distinguidos del triángulo rectilíneo, 1822). En este libro, Feuerbach añadió otro hecho al círculo de los nueve puntos, un excírculo es uno que es tangente a uno de los lados y a las

extensiones de los otros dos lados. (El centro de un excírculo yace sobre los bisectores de los dos ángulos exteriores y el ángulo interior más alejado.) El teorema de Feuerbach establece que el círculo de nueve puntos es tangente al círculo inscrito y a los tres excírculos.

En un pequeño libro publicado en 1816, *Über einige Eigenschaften des ebenen geradlinigen Dreiecks* (Sobre algunas propiedades de triángulos rectilíneos planos), Crelle mostró cómo determinar un punto P dentro de un triángulo tal que las líneas uniendo P a los vértices del triángulo y los lados del triángulo formen ángulos iguales. Esto es, $\angle 1 = \angle 2 = \angle 3$ en la figura 35.1.

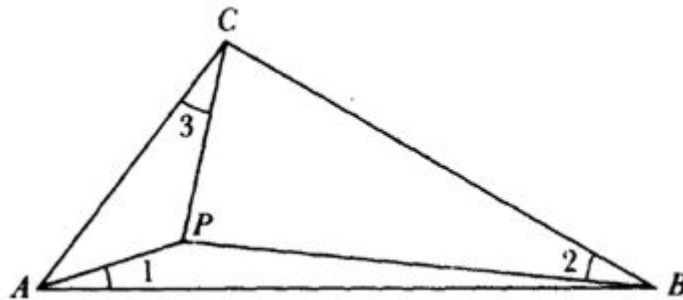


Figura 35.1

También existe un punto P diferente de P tal que $\angle PAC = \angle PCB = \angle PBA$.

Las secciones cónicas, como sabemos, fueron tratadas definitivamente por Apolonio como secciones de un cono y más tarde introducidas como lugares geométricos planos en el siglo XVII. En 1822, Germinal Dandelin (1794-1847) demostró⁴⁷ un teorema muy interesante acerca de las secciones cónicas en relación con un cono. Su teorema establece que si dos esferas están inscritas en un

cono circular de tal manera que son tangentes a un plano dado cortando al cono en una sección cónica, los puntos de contacto de las esferas con el plano son los focos de la sección cónica, y las intersecciones del plano con los planos de los círculos a lo largo de los que las esferas tocan el cono son las directrices de la cónica.

Otro tema interesante estudiado en el siglo XIX fue la solución de problemas de máximos y mínimos por métodos puramente geométricos, esto es, sin apoyarse en el cálculo de variaciones. De los varios teoremas que Jacob Steiner demostró usando únicamente métodos sintéticos, el resultado más famoso es el teorema isoperimétrico: de todas las figuras planas con un perímetro dado el círculo encierra el área máxima. Steiner proporcionó varias demostraciones⁴⁸. Desafortunadamente, Steiner supuso que existe una curva que tiene área máxima. Dirichlet intentó varias veces persuadirlo de que sus demostraciones eran incompletas en ese punto pero Steiner insistía que esto era autoevidente. Una vez, sin embargo, escribió (en el primero de los ensayos de 1842)⁴⁹: *«y la demostración se hace rápidamente si se supone que existe una figura máxima»*.

La demostración de la existencia de una curva maximizante frustró a los matemáticos por muchos años hasta que Weierstrass, en sus clases dadas durante los 1870 en Berlín, recurrió al cálculo de variaciones⁵⁰. Más tarde, Constantin Caratheodory (1873-1950) y Study⁵¹, en un artículo conjunto, rigorizaron las demostraciones de Steiner sin emplear el cálculo. Sus demostraciones (hubo dos) fueron directas en lugar de indirectas, como en el método de

Steiner. Hermann Amandus Schwarz, quien realizó una gran labor en ecuaciones en derivadas parciales y análisis y trabajó como profesor en diversas universidades, incluyendo Göttingen y Berlín, proporcionó una demostración rigurosa para el problema isoperimétrico en tres dimensiones.⁵²

Steiner también demostró (en el primero de los ensayos de 1842) que de todos los triángulos con un perímetro dado, el equilátero tiene el área máxima.

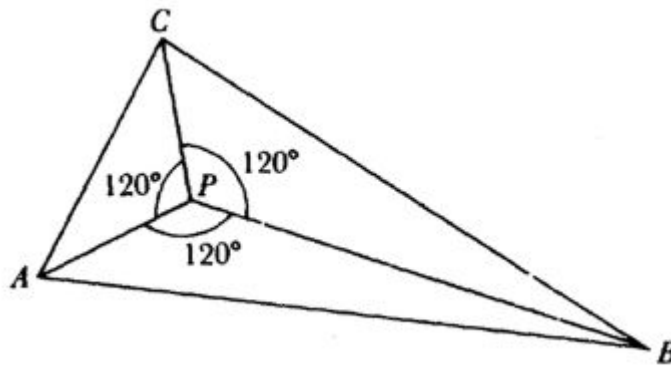


Figura 35.2

Otro de sus resultados⁵³ dice que si A , B y C son tres puntos dados (fig. 35.2) y si cada uno de los ángulos del triángulo ABC es menor de 120° , entonces el punto P para el cual $PA + PB + PC$ es un mínimo es tal que cada uno de los ángulos en P es de 120° . Si, sin embargo, un ángulo del triángulo, digamos el ángulo A , es igual o más grande que 120° , entonces P coincide con A . Este resultado había sido demostrado mucho antes por Cavalieri (*Exercitationes Geometricae Sex*, 1647), pero indiscutiblemente le era desconocido a Steiner. Steiner también extendió este resultado a n puntos.

Schwarz resolvió el siguiente problema: dado un triángulo con ángulos agudos, considérense todos los triángulos tales que cada uno tiene sus vértices sobre los tres lados del triángulo original: el problema es encontrar el triángulo con el perímetro mínimo. Schwarz demostró sintéticamente⁵⁴ lo que los vértices de este triángulo de perímetro mínimo son los pies de las alturas del triángulo dado (fig. 35.3)⁵⁵.

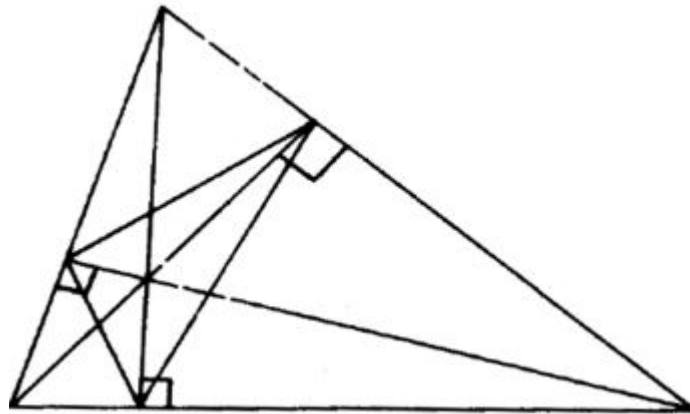


Figura 35.3

En 1899, Frank Morley, profesor de matemáticas de la Universidad Johns Hopkins, descubrió un nuevo teorema de la geometría euclídea y demostraciones de él fueron publicadas por muchos autores.⁵⁶ El teorema dice que si son trazados los trisectores del ángulo en cada vértice de un triángulo, los trisectores adyacentes se cortan en los vértices de un triángulo equilátero (fig. 35.4).

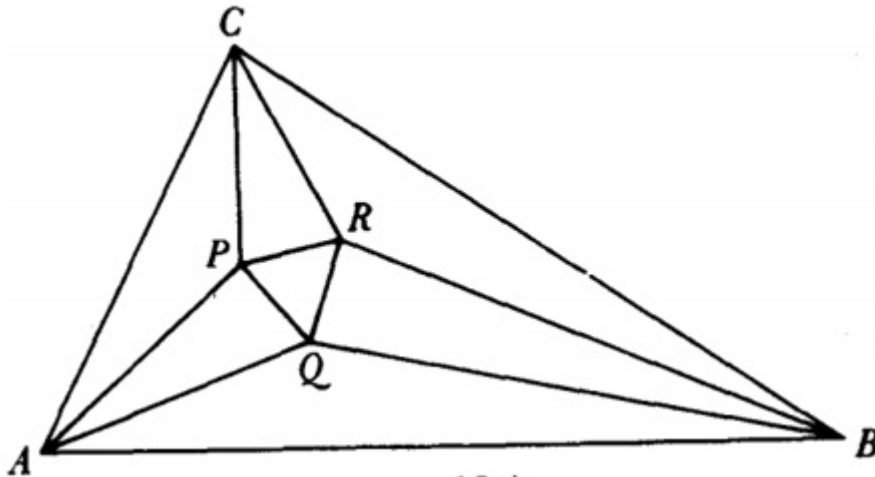


Figura 35.4

La novedad está en el uso de trisectores de ángulos. Hasta la mitad del siglo XIX ningún matemático había considerado tales líneas ya que únicamente los elementos y figuras que son constructibles eran considerados con legitimidad en la geometría euclidiana. La constructividad garantizaba la existencia. Sin embargo, la concepción de lo que significaba la existencia cambió, como veremos más claramente cuando examinemos el trabajo sobre los fundamentos de la geometría euclidiana.

Algunos esfuerzos fueron dirigidos hacia la reducción del uso de regla y compás, en la misma línea iniciada por Mohr y Mascheroni (cap. 12, sec. 2). En su *Traité* de 1822, Poncelet demostró que todas las construcciones posibles con regla y compás (excepto la construcción de arcos circulares) eran posibles simplemente con regla, siempre que nos sea dado un círculo fijo con su centro. Steiner redemonstró el mismo resultado más elegantemente en su pequeño libro *Die geometrischen Constructionen ausgeführt mittelst*

*der geraden Linie und eines festen Kreises (Las construcciones geométricas ejecutadas por medio de regla y un círculo fijo).*⁵⁷

A pesar de que Steiner escribió el libro con fines pedagógicos, asegura en el prefacio del libro que establecerá la *conjetura* que un matemático francés había expresado.

La breve muestra anterior de teoremas euclídeos establecidos mediante métodos sintéticos no debe dejar al lector con la impresión' de que no se usaron los métodos de la geometría analítica. De hecho, Gergonne proporcionó demostraciones analíticas de muchos teoremas geométricos que publicó en la revista fundada por él, los *Annales de Mathématiques*.

3. El resurgimiento de la geometría proyectiva sintética

El área principal hacia la que se volvieron Monge y sus alumnos fue la geometría proyectiva. Esta materia había tenido una explosión de actividad vigorosa, pero de breve vida, en el siglo XVII (cap. 14), pero fue menospreciada por el surgimiento de la geometría analítica, el cálculo y el análisis. Como ya hemos notado, el trabajo fundamental de Desargues de 1639 se perdió de vista hasta 1845 y el principal ensayo de Pascal sobre cónicas (1639) nunca fue recuperado. Sólo los libros de La Hire, que usaba algunos de los resultados de Desargues, estaban disponibles. Lo que los autores del siglo XIX aprendieron de los libros de La Hire, a menudo se lo atribuían a él incorrectamente. En general, estos geómetras ignoraban el trabajo de Desargues y Pascal y tuvieron que recrearlo. El resurgimiento de la geometría proyectiva fue iniciado por Lazare N. M. Carnot (1753-

1823), alumno de Monge y padre del distinguido físico Sadi Carnot. Su principal obra fue la *Géométrie de position* (Geometría de la posición, 1803) y también contribuyó al *Essai sur la théorie des transversales* (Ensayo sobre la teoría de las transversales, 1806). Monge defendió el uso conjunto del análisis y la geometría pura, pero Carnot rehusó emplear métodos analíticos y empezó la cruzada de la geometría pura. Muchas de las ideas que discutiremos en breve están sugeridas casi en su totalidad en el trabajo de Carnot. Así, el principio que Monge llamó de las relaciones contingentes, conocido también como principio de correlatividad, y más comúnmente como principio de continuidad, se encuentra ahí. Carnot, a fin de evitar que las figuras sean separadas para los distintos tamaños de los ángulos y las direcciones de las líneas, no usó los números negativos, que consideraba contradictorios, y sí introdujo en cambio un complicado esquema llamado «correspondencia de signos».

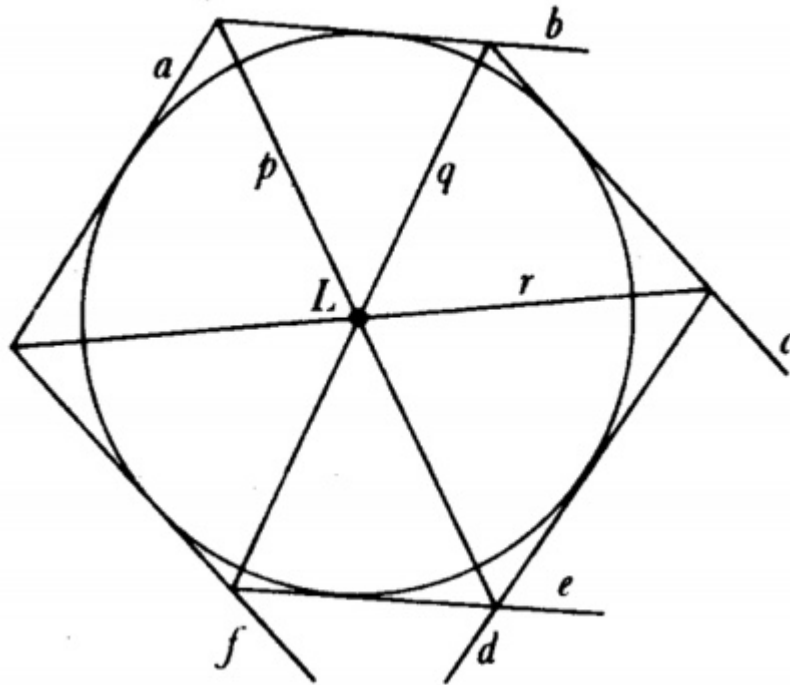


Figura 35.5

Entre los cultivadores de principios del siglo XIX de la geometría proyectiva mencionaremos únicamente a François Joseph Servois y Charles Julien Brianchon (1785-1864), quienes aplicaron aspectos de su trabajo a problemas militares. Aunque contribuyeron a la reconstrucción, sistematización y extensión de viejos resultados, el único nuevo teorema importante es el famoso resultado de Brianchon⁵⁸, que demostró mientras era estudiante en la Ecole. El teorema afirma que si tenemos seis tangentes a una cónica (fig. 35.5), formando un hexágono circunscrito, las tres rectas, cada una de las cuales une dos vértices opuestos, pasan por un punto.

Brianchon derivó ese teorema utilizando la relación polo-polar.

El resurgimiento de la geometría proyectiva recibió su ímpetu principal de Poncelet (1788-1867). Poncelet fue alumno de Monge y

también aprendió mucho de Carnot. Mientras servía como oficial en la campaña de Napoleón contra Rusia, es capturado y permanece el año 1813-1814 en una prisión rusa en Saratoff. Ahí Poncelet reconstruyó —sin la ayuda de libro alguno— lo que había aprendido de Monge y Carnot, creando nuevos resultados. Después amplió y revisó su trabajo, publicándolo bajo el título de *Traité des propriétés projectives des figures* (Tratado de las propiedades proyectivas de las figuras, 1822). Este trabajo constituye su principal contribución a la geometría proyectiva y a la erección de una nueva disciplina. En su vida posterior se vio obligado a dedicar gran parte de su tiempo al servicio gubernamental, a pesar de que tuvo algunos nombramientos de profesor por períodos limitados.

Poncelet se convirtió en el más ardiente defensor de la geometría sintética y hasta atacó el análisis. Aunque había sido amigo del gran analista Joseph Diez Gergonne (1771-1859) y publicó ensayos en los *Annales de Mathématiques*, de Gergonne, sus ataques también fueron en breve dirigidos a Gergonne. Poncelet estaba convencido de la autonomía e importancia de la geometría pura. A pesar que admitía el poder del análisis, creía que era posible otorgar el mismo poder a la geometría sintética. En un ensayo de 1818, publicado en los *Annales* de Gergonne⁵⁹, afirmó que el poder de los métodos analíticos no yacía en su uso del álgebra sino en su generalidad, y su ventaja resultaba del hecho que las propiedades métricas descubiertas para una figura típica permanecían aplicables, con la posible excepción del cambio de signo, a todas las figuras relacionadas que surgían de la típica o básica. Esta generalidad era

asegurada en la geometría sintética por el principio de continuidad (que examinaremos en breve).

Poncelet fue el primer matemático en apreciar completamente que la geometría proyectiva era una nueva rama de las matemáticas, con métodos y metas propios. Mientras que los geómetras del siglo XVII habían tratado con problemas específicos, Poncelet entreveía el problema general de buscar todas las propiedades de las figuras geométricas que eran comunes a todas las secciones de cualquier proyección de una figura, esto es, permanecían sin alteración mediante la proyección y la sección. Este es el tema que él y sus sucesores estudiaron. Ya que las distancias y los ángulos son alterados por proyección y sección, Poncelet seleccionó y desarrolló la teoría de involución y de conjuntos armónicos de puntos, pero no el concepto de razón doble. Monge había utilizado la proyección paralela en su trabajo; como Desargues, Pascal, Newton y Lambert, Poncelet se valió de la proyección central, esto es, la proyección desde un punto, y elevó este concepto a un método de aproximación a los problemas geométricos. Poncelet también consideraba la transformación proyectiva a partir de una figura del espacio en otra, por supuesto, en una forma puramente geométrica. Aquí pareció perder interés en las propiedades proyectivas y estaba más interesado por el uso del método en bajo-relieve y diseño de andamios.

Su trabajo se centra en tres ideas. La primera es la de las figuras homologas; dos figuras son homologas si se deriva una de la otra mediante una proyección y una sección, que se denomina una

perspectividad, o mediante una secuencia de proyecciones y secciones, esto es, una proyectividad. Al trabajar con figuras homologas, su plan era encontrar para una figura dada una figura homologa más simple, y, estudiándola, encontrar propiedades que son invariantes bajo proyección y sección, obteniendo propiedades de la figura más complicada. Desargues y Pascal emplearon la esencia de este método; y Poncelet, en su *Traité*, elogió la originalidad de Desargues en este y otros aspectos.

El segundo tema principal de Poncelet es el principio de continuidad. En su *Traité* lo establece de la siguiente manera: «Si una figura es derivada de otra mediante un cambio continuo y la última es tan general como la anterior, entonces cualquier propiedad de la primera figura puede ser establecida inmediatamente para la segunda.» La determinación de cuándo ambas figuras son generales no es explicada. El principio de Poncelet también afirma que si una figura degenera, como lo hace un hexágono en un pentágono cuando se hace que un lado se aproxime a cero, cualquier propiedad de la figura original será transmitida con un argumento apropiadamente redactado para la figura degenerada.

El principio no era realmente nuevo con Poncelet. En su sentido filosófico amplio se remite a Leibniz, quien estableció en 1687 que cuando las diferencias entre dos casos se pueden hacer más pequeñas que cualquier dato dado, las diferencias se pueden hacer más pequeñas que cualquier cantidad dada en el resultado. Desde los tiempos de Leibniz el principio fue reconocido y usado

constantemente. Monge empezó a valerse del principio de continuidad para establecer teoremas. Quería demostrar un teorema general, pero utilizó una posición especial de la figura para demostrarlo y entonces mantuvo que el teorema era cierto en general, aun cuando algunos elementos de las figuras se hicieran imaginarios. Así, para demostrar un teorema acerca de una recta y una superficie lo demostraba cuando la recta corta la superficie y mantiene entonces que el resultado es válido aun cuando la línea ya no corta a la superficie y los puntos de intersección son imaginarios. Ni Monge ni Carnot, quienes también aplicaron este principio, proporcionaron ninguna justificación.

Poncelet, quien acuñó el término «principio de continuidad», propuso el principio como una verdad absoluta y lo aplicó con libertad en su *Traité*. Para «demostrar» su solidez toma el teorema sobre la igualdad de los productos de los segmentos de cuerdas que se cortan en un círculo y nota que cuando el punto de intersección se mueve fuera del círculo se obtiene la igualdad de los productos de las secantes y sus segmentos externos. Más aún, cuando una secante se convierte en tangente, la tangente y su segmento externo se hacen iguales y su producto continúa siendo igual al producto de las otras secantes y su segmento externo. Todo esto era suficientemente razonable, pero Poncelet aplicó el principio para demostrar muchos teoremas y, como Monge, extendió el principio para hacer aserciones acerca de figuras imaginarias. (Mencionaremos algunos ejemplos más tarde.)

Los otros miembros de la Academia de Ciencias de París criticaron el principio de continuidad y lo consideraban como si tuviera únicamente un interés heurístico. Cauchy, en particular, criticó el principio, pero desafortunadamente la crítica fue dirigida hacia aplicaciones que Poncelet había realizado donde el principio sí funcionaba. Los críticos también mencionaron que la confianza que Poncelet y otros tenían en el principio provenía realmente del hecho de que podía ser justificado sobre una base algebraica. De hecho, las notas que Poncelet hizo en prisión muestran que utilizó el análisis para probar la solidez de su principio. Estas notas, incidentalmente, fueron escritas por Poncelet y publicadas por él en dos volúmenes titulados *Applications d'analyse et de géométrie* (Aplicaciones del análisis y la geometría, 1862-1864), que es realmente una revisión de su *Traité* de 1822, y en el último trabajo sí usa métodos analíticos. Poncelet admitió que una demostración podía estar basada sobre el álgebra, pero insistió en que el principio no dependía de tal prueba. Sin embargo, es bastante seguro que Poncelet se apoyaba en el método algebraico para ver qué sucedía y entonces asegurar los resultados geométricos usando el principio como justificación.

Chasles, en su *Aperçu*, defendió a Poncelet. La posición de Chasles era que el álgebra es una demostración a posteriori del principio. Sin embargo, él se protegía señalando que había que ser cuidadoso en no llevar de una figura a otra cualquier propiedad que dependa esencialmente de que los elementos sean reales o imaginarios. De esta forma, una sección de un cono puede ser una hipérbola, y ésta

tiene asíntotas. Cuando la sección es una elipse, las asíntotas se hacen imaginarias. De aquí que no era posible demostrar un resultado acerca de las asíntotas por sí solas, porque éstas dependían de la naturaleza particular de la sección. Como no se debían llevar los resultados para una parábola al caso de la hipérbola, porque el plano que corta no tiene una posición general en el caso de la parábola. Más adelante discute el caso de dos círculos intersecándose que tienen una cuerda en común. Cuando los círculos ya no se intersecan, la cuerda común es imaginaria. El hecho de que la cuerda común real pase por dos puntos reales es, dice, una propiedad incidental o contingente. Una cuerda se ha de definir en algún modo que no dependa del hecho de que pasa por puntos reales cuando los círculos se intersecan, sino que sea una propiedad permanente de los dos círculos en cualquier posición. Así, uno puede definirlo como el eje radical (real), que significa que es una recta tal que a partir de cualquier punto sobre ella, las tangentes a los dos círculos son iguales, o bien se puede definir por medio de la propiedad de que cualquier círculo trazado con cualquier punto de la recta como centro corta los dos círculos ortogonalmente.

Chasles también insistía en que el principio de continuidad era el adecuado para tratar los elementos imaginarios en geometría. Primero explica lo que se quiere decir por imaginario en geometría. Los elementos imaginarios pertenecen a una condición o estado de una figura en el cual ciertas partes son no existentes, siempre que estas partes sean reales en otro estado de la figura. Porque, añade,

uno no puede tener ninguna idea de las cantidades imaginarias, excepto pensando en los estados relacionados en que las cantidades son reales. Estos últimos estados son los que uno llama «accidentales» y los que proporcionan la clave para lo imaginario en geometría. Para demostrar resultados acerca de los elementos imaginarios sólo es necesario tomar la condición general de la figura en la que los elementos son reales, y entonces, de acuerdo con el principio de relaciones accidentales o el principio de continuidad, se concluye que los resultados se mantienen cuando los elementos son imaginarios. *«Así se observa que el uso y la consideración de los imaginarios está completamente justificada.»* El principio de continuidad fue aceptado durante el siglo XIX como intuitivamente claro y, por tanto, considerado como un axioma. Los geómetras lo usaron libremente y nunca juzgaron que requiriera una demostración.

A pesar de que Poncelet utilizó el principio de continuidad para aseverar resultados acerca de los puntos y líneas imaginarios, nunca proporcionó una definición general de tales elementos. Para introducir algunos puntos imaginarios proporcionó una definición geométrica que es complicada y nada clara. Entenderemos estos elementos imaginarios con más facilidad cuando sean discutidos desde un punto de vista algebraico. A pesar de la falta de claridad en el acercamiento de Poncelet, a él debe acreditársele la introducción de la noción de los puntos circulares en el infinito, esto es, dos puntos imaginarios situados sobre la recta en el infinito, y comunes a dos círculos cualesquiera⁶⁰. También presentó los

círculos esféricos imaginarios que dos esferas cualesquiera tienen en común. Más adelante demostró que dos cónicas reales que no se intersecan tienen dos cuerdas comunes imaginarias y dos cónicas se intersecan en cuatro puntos, reales o imaginarios.

La tercera idea conductora del trabajo de Poncelet es la noción de polo y polar con respecto a una cónica. El concepto data de Apolonio y fue usado por Desargues (cap. 14, sec. 3) y por otros en los trabajos del siglo XVII sobre geometría proyectiva: Euler, Legendre, Monge, Servois y Brianchon ya lo habían utilizado. Pero Poncelet proporcionó una formulación general de la transformación de polo a polar, y conversamente, y la usó en su *Traité* de 1822 y en su «*Mémoire sur la théorie générale des polaires reciproques*» («Memoria sobre la teoría general de polares recíprocas»), presentada a la Academia de París en 1824⁶¹ como un método para establecer muchos teoremas.

Uno de los objetivos de Poncelet al estudiar la polar recíproca con respecto a una cónica era establecer el principio de dualidad. Los que trabajaban en geometría proyectiva habían observado que los teoremas relativos a figuras situadas en un plano, cuando eran parafraseados reemplazando la palabra «punto» por «recta» y «recta» por «punto», no solamente tenían sentido, sino que eran ciertos. La razón para la validez de los teoremas resultantes de tal paráfrasis no era clara, y de hecho Brianchon cuestionó el principio. Poncelet pensó que la relación entre polo y polar era la razón.

Sin embargo, esta relación requería la mediación de una cónica. Gergonne⁶² insistió en que el principio era general y aplicable a

todos los enunciados y teoremas, excepto aquellos que involucraban propiedades métricas. No se necesitaba del polo y la polar como un objeto de apoyo intermedio. Introdujo el término «dualidad» para denotar la relación entre el teorema original y el nuevo. También observó que en situaciones tridimensionales el punto y el plano constituyen elementos duales y la recta es dual es sí misma.

Para ilustrar el entendimiento por Gergonne del principio de dualidad, examinemos la dualización del teorema del triángulo de Desargues. Debemos notar primero lo que es el dual de un triángulo. Un triángulo consiste de tres puntos no todos sobre la misma recta, y las rectas que los unen. La figura dual consiste de tres rectas no todas ellas pasando por el mismo punto y los puntos uniéndolas (los puntos de intersección). La figura dual es de nuevo un triángulo y, por lo mismo, el triángulo es llamado autodual. Gergonne inventó el esquema de escribir teoremas duales en columnas dobles, con el dual a un lado de la proposición original.

Ahora consideremos el teorema de Desargues, donde esta vez los dos triángulos y el punto O yacen en un plano, y veamos qué resulta cuando intercambiamos punto por recta. En el ensayo de Gergonne, de 1825-1826, ya escribió este teorema y su dual como sigue:

Teorema de Desargues

Si tenemos dos triángulos tales que las rectas uniendo vértices correspondientes pasan por el punto O , entonces los lados

Teorema Dual de Desargues

Si tenemos dos triángulos tales que los puntos que son las uniones de lados correspondientes yacen sobre

correspondientes se intersecan en tres puntos sobre una recta. una recta O , entonces los vértices correspondientes están unidos por tres rectas pasando por el mismo punto.

Aquí el teorema dual es el recíproco del teorema original.

La formulación de Gergonne del principio general de dualidad fue algo vaga y tenía deficiencias. A pesar de que estaba convencido de que era un principio universal, no podía justificarlo y Poncelet señaló correctamente las deficiencias. También disputó con Gergonne sobre la prioridad del descubrimiento del principio (que realmente pertenecía a Poncelet) y aún acusó a Gergonne de plagio. Sin embargo, Poncelet se apoyaba realmente sobre el polo y la polar y no aceptaría el que Gergonne hubiera dado un paso adelante al reconocer las amplias aplicaciones del principio. Discusiones posteriores entre Poncelet, Gergonne, Möbius, Chasles y Plücker clarificaron el principio para todos y Möbius, en su *Der barycentrische Calcul* (El Cálculo Baricéntrico), más tarde, proporcionó un buen enunciado de la relación del principio de dualidad con el polo y la polar: la noción de dualidad es independiente de las cónicas y formas cuadráticas pero coincide con el polo y la polar cuando se usan las últimas. La justificación lógica del principio general de dualidad no fue proporcionada en este momento.

Jacob Steiner (1796-1863) llevó más lejos el desarrollo sintético de la geometría proyectiva. Es el primero de la escuela de geómetras

alemanes que adoptó ideas francesas, notablemente de Poncelet y favoreció los métodos sintéticos hasta el punto de llegar a odiar el análisis. Hijo de un granjero suizo, trabajó también en la granja hasta la edad de 19 años. Aunque en gran parte era autodidacta, finalmente se convirtió en profesor en Berlín. En sus años mozos trabajó como maestro en la escuela de Pestalozzi y se impresionó ante la importancia que revestía incrementar la intuición geométrica. El principio de Pestalozzi era hacer que el estudiante creara las matemáticas con la guía del maestro y el método socrático. Steiner se fue a los extremos. Enseñó geometría pero no usaba figuras, y al preparar a los candidatos al doctorado obscurecía la sala. En su propio trabajo posterior, Steiner se dedicó a tomar teoremas y demostraciones publicadas en inglés y en otras revistas y no daba indicaciones en sus propios textos respecto a que sus resultados ya habían sido publicados. Había hecho buen trabajo original durante su juventud y buscó mantener su reputación de productividad.

Su principal trabajo, *Systematische Entwicklung der Abhängigkeit geometrischen Gestalten von einander* (Desarrollo sistemático de la dependencia de las formas geométricas una de otra, 1832) y su objetivo principal fue usar los conceptos proyectivos para construir estructuras más complicadas a partir de las más simples como puntos, rectas, haces de rectas, planos y haces de planos. Sus resultados no eran especialmente nuevos, pero sí su método.

Para ilustrar su principio vamos a examinar su propio método proyectivo, ahora común, de definir las secciones cónicas.

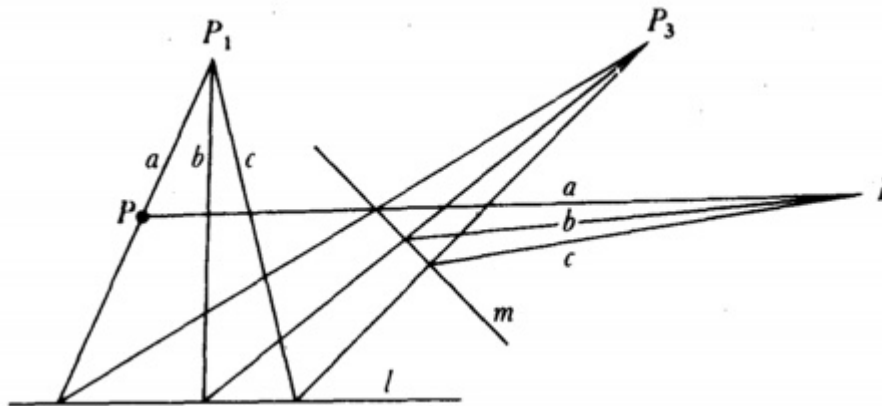


Figura 35.6

Se empieza (fig. 35.6) con dos haces de rectas (familias de rectas concurrentes), digamos p_1 y p_3 , que están relacionadas perspectivamente a través de un haz de puntos sobre una línea l , y los haces p_3 y p_2 , que están relacionados perspectivamente por medio de un haz de puntos sobre otra línea m . Entonces se dice que los haces p_1 y p_3 están relacionados proyectivamente. Las líneas marcadas a en el haz con centro en P_1 y el haz con centro en P_2 son ejemplos de líneas correspondientes en la proyectividad entre los dos haces p_1 y p_2 . Una cónica está definida ahora como el conjunto de los puntos de intersección de todos los pares de líneas correspondientes de los dos haces proyectivos. De esta manera, P es un punto sobre la cónica. Más aún, la cónica pasa por P_1 y P_2 (fig. 35.7).

De esta manera, Steiner construyó las cónicas o curvas de segundo grado por medio de formas más simples, haces de rectas. Sin embargo, no identificó sus cónicas con las secciones de un cono.

También construyó las cuádricas regladas, el hiperboloide de una hoja y el paraboloides hiperbólico, de una manera semejante, haciendo de la correspondencia proyectiva la base de sus definiciones. De hecho, su método no es suficientemente general para toda la geometría proyectiva.

En sus demostraciones utilizó la razón doble como una herramienta fundamental, ignoró los elementos imaginarios y se refirió a ellos como los «fantasmas» o las «sombras de la geometría». Tampoco utilizó cantidades con signo, a pesar de que Möbius, cuyo trabajo examinaremos en breve, ya las había introducido.

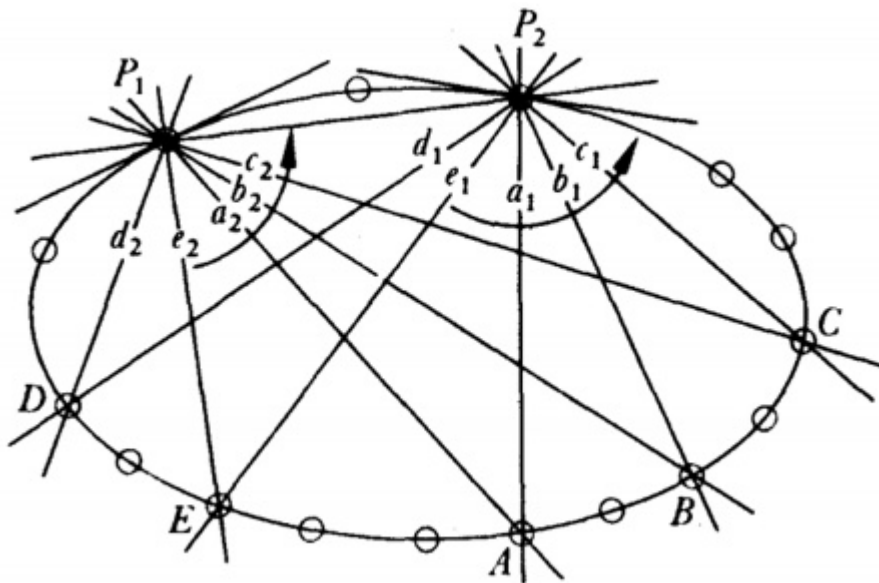


Figura 35.7

Steiner utilizó el principio de dualidad desde el inicio de su trabajo. De tal forma, dualizó la definición de una cónica para obtener una nueva estructura llamada una curva de rectas.

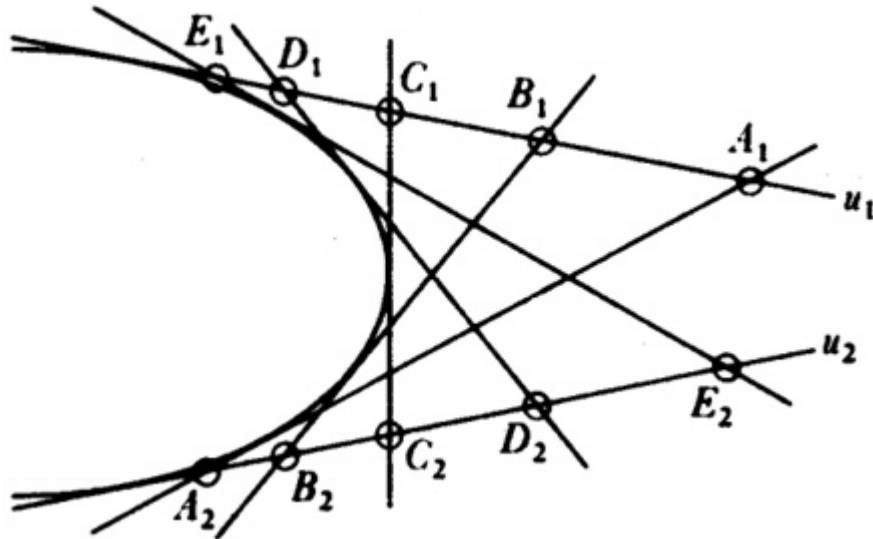


Figura 35.8

Si se empieza con dos haces de puntos relacionados proyectivamente (pero no perspectivamente), entonces la familia de líneas (fig. 35.8) uniendo los puntos correspondientes de los dos haces es llamada una cónica de rectas. Para distinguir tales familias de rectas, que también describen una curva, la curva usual como lugar geométrico de puntos es llamada curva de puntos. Las líneas tangentes a una curva de puntos son una curva de rectas y en el caso de una cónica constituyen la curva dual. Recíprocamente, cada cónica de rectas envuelve una cónica de puntos o es la colección de tangentes a una cónica de puntos.

Con la noción de Steiner de dual de una cónica de puntos, es posible dualizar muchos teoremas. Tomemos el teorema de Pascal y formemos el enunciado dual. Escribiremos el teorema a la izquierda y el nuevo enunciado a la derecha.

Teorema de Pascal

Si tomamos seis puntos A, B, C, D, E y F sobre una cónica de puntos, entonces las líneas que unen A y B y D y E se juntan en el punto P ; las líneas que unen B y C y E y F se juntan en el punto Q ; las líneas que unen C y D y F y se juntan en el punto R . Los tres puntos P, Q y R yacen sobre una línea l .

Teorema dual de Pascal

Si tomamos seis líneas a, b, c, d, e y f sobre la cónica de rectas entonces los puntos que unen a con b y d y e están unidos por la línea p ; los puntos que unen b y c y e y f están unidos por la línea q ; los puntos que unen c y d y f y a están unidos por la línea r . Las tres líneas p, q y r están sobre un punto L .

La figura 14.12 del capítulo 14 ilustra el teorema de Pascal. El teorema dual es el que Brianchon descubrió mediante la relación entre polo y polar (fig. 35.5 anterior). Steiner, como Gergonne, no hizo nada para establecer las bases lógicas del principio de dualidad. Sin embargo, desarrolló la geometría proyectiva sistemáticamente, clasificando las figuras y notando los argumentos duales a medida que avanzaba. También estudió en detalle las curvas y superficies de segundo grado.

Michel Chasles, quien dedicó su vida por completo a la geometría, siguió el trabajo de Poncelet y Steiner, aunque no conoció personalmente el trabajo de Steiner porque, como ya hemos notado, Chasles no sabía leer alemán. Chasles presentó sus propias ideas en su *Traité de géométrie supérieure* (Tratado de geometría superior,

1852) y el *Traité des sections coniques* (Tratado de las secciones cónicas, 1865). Ya que gran parte del trabajo de Chasles inintencionalmente duplicaba el de Steiner o fue superado por conceptos más generales, notaremos sólo unos cuantos resultados importantes que se le deben.

Chasles llegó a la idea de razón doble a partir de sus intentos por entender el trabajo perdido de Euclides Los Porismas (a pesar de que Steiner y Möbius ya lo habían reintroducido). Desargues también se había valido ya del concepto, pero Chasles conocía únicamente lo que La Hire había publicado sobre él. Chasles supo en algún momento que la idea también se encuentra en Pappus, ya que en la Nota IX de su *Aperçu* (p. 302) se refiere al uso de Pappus de la idea. Uno de los resultados de Chasles en esta área⁶³ es que cuatro puntos fijos de una cónica y un quinto punto de la cónica determinan cuatro rectas con la misma razón doble.

En 1828, Chasles⁶⁴ dio el teorema: dados dos conjuntos de puntos colineales en una correspondencia uno-a-uno y tales que la razón doble de cuatro puntos cualesquiera sobre una recta es igual a la de los puntos correspondientes sobre la otra, entonces las rectas uniendo los puntos correspondientes son tangentes a una cónica que toca las dos rectas dadas. Este resultado es equivalente a la definición de Steiner de una cónica de rectas, ya que la condición de razón doble asegura aquí que los dos conjuntos de puntos colineales están relacionados proyectivamente y las rectas uniendo los puntos correspondientes son las rectas de la cónica de rectas de Steiner.

Chasles señaló que en virtud del principio de dualidad, las rectas pueden ser tan fundamentales como los puntos en el desarrollo de la geometría proyectiva plana y reconoce a Poncelet y Gergonne que son claros en este punto. Chasles introdujo también nueva terminología. A la razón doble la llamó proporción anarmónica. Introdujo el término homografía para describir una transformación de un plano en sí mismo u otro plano, que lleva puntos en puntos y rectas en rectas. Este término cubre figuras relacionadas proyectivamente u homologas. Añadió la condición de que la transformación debe preservar la razón doble, pero este último hecho puede ser demostrado. A la transformación que lleva puntos en rectas y rectas en puntos la llamó correlación.

A pesar de que Chasles defendió la geometría pura, pensaba analíticamente, aunque sus demostraciones y resultados los presentaba geoméricamente. Este enfoque es llamado el «método mixto» y fue usado posteriormente por otros.

Hacia 1850 los conceptos y metas generales de la geometría proyectiva como distintos de los de la geometría euclidiana eran claros; sin embargo, la relación lógica entre las dos geometrías no estaba clara. El concepto de longitud había sido usado en geometría proyectiva de Desargues a Chasles. De hecho, el concepto de razón doble fue definido en términos de longitud. Sin embargo, la longitud no es un concepto proyectivo, porque no es invariante bajo las transformaciones proyectivas. Karl Georg Christian von Staudt (1798-1867), profesor en Erlangen, quien estaba interesado en los fundamentos, decidió liberar a la geometría proyectiva de su

dependencia de la longitud y la congruencia. La esencia de su plan, presentada en su *Geometrie der Lage* (Geometría de Posición, 1847) consistió en introducir un análogo de la longitud sobre una base proyectiva. Su esquema es llamado «el álgebra de lanzamientos». Uno elige tres puntos arbitrarios sobre una línea y les asigna los símbolos 0 , 1 y ∞ . Entonces, por medio de una construcción geométrica (que de hecho proviene de Möbius) —un «lanzamiento»—, un símbolo es asociado a cualquier punto arbitrario P .

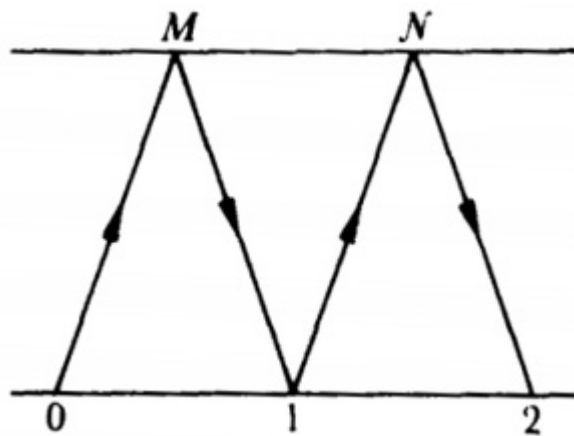


Figura 35.9

Para ver qué significa la construcción en geometría euclídea supongamos que empezamos con los puntos etiquetados 0 y 1 sobre la línea (fig. 35.9). Por el punto M sobre la línea paralela trácese $0M$ y entonces se traza $1N$ paralela a $0M$. Ahora se traza $1M$ y luego $N2$ paralela a $0M$. Es inmediato que $01 = 12$ ya que los lados opuestos de un paralelogramo son iguales. Entonces se traslada la longitud 01 a 12 mediante una construcción geométrica.

Ahora, para tratar el caso proyectivo empezamos con tres puntos: 0 , 1 , ∞ (fig. 35.10). El punto ∞ yace sobre l_∞ la recta del infinito, pero esta es una recta ordinaria en la geometría proyectiva. Ahora escójase un punto M y trácese una «paralela» a la recta 01 pasando por M . Esto significa que la recta que pasa por M debe cortar la recta 01 en ∞ y así trazamos $M\infty$. Ahora trazamos $0M$ y la prolongamos hasta l_∞ . Después trazamos 1 , una «paralela» a $0M$. Esto significa que la «paralela» que pasa por 1 debe cortar $0M$ sobre l_∞ . Así obtenemos la recta $1P$ y esto determina N . Ahora se traza $1M$ y la prolongamos hasta que toca l_∞ en Q . La recta que pasa por N y es paralela a $1M$ es QN y donde corta a 01 obtenemos un punto que es etiquetado 2 .

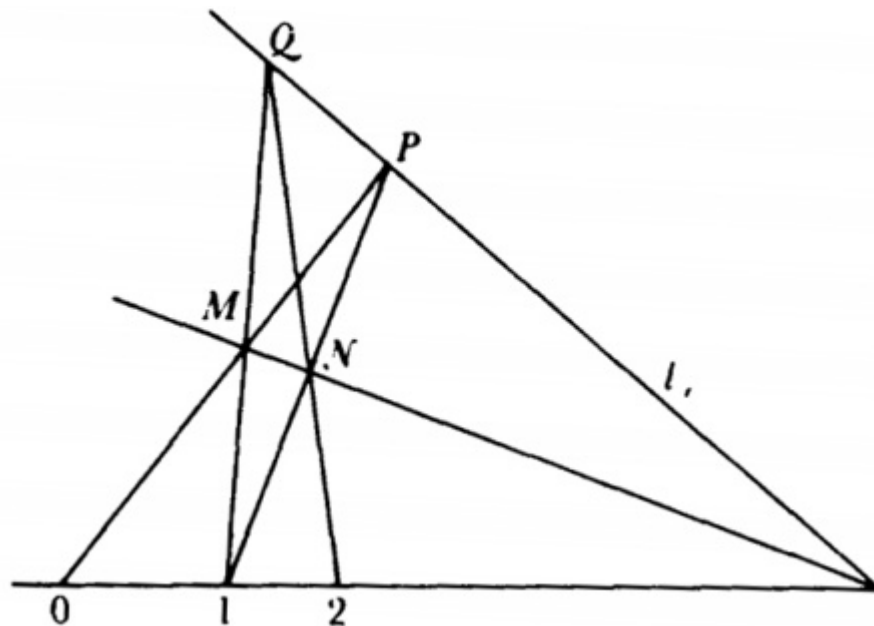


Figura 35.10

Por medio de este tipo de construcciones podemos asignar «coordenadas racionales» a los puntos sobre la recta 01∞ . Para asignar números irracionales a puntos sobre la recta se debe introducir un axioma de continuidad (cap. 41). Esta noción no era bien entendida en aquel tiempo y, como consecuencia, al trabajo de Staudt le faltó rigor.

La manera como Von Staudt asignó las coordenadas a los puntos sobre la recta no hacía uso de la longitud. Sus coordenadas, a pesar de que eran los símbolos números usuales, sirvieron como símbolos de identificación sistemática para los puntos. Para añadir o sustraer tales «números», Von Staudt no podía usar las leyes de la aritmética. En su lugar, proporcionó construcciones geométricas que definían las operaciones con estos símbolos, por ejemplo, de tal forma que la suma de los números 2 y 3 es el número 5. Estas operaciones obedecían todas las leyes usuales de los números. Así, sus símbolos o coordenadas eran susceptibles a ser tratados como números ordinarios, a pesar de que habían sido construidos geoméricamente.

Con estas etiquetas adheridas a los puntos, von Staudt podía definir la razón doble de cuatro puntos. Si las coordenadas de estos puntos son x_1 , x_2 , x_3 y x_4 . Y entonces la razón doble está definida como

$$\frac{x_1 - x_3}{x_1 - x_4} \bigg/ \frac{x_2 - x_3}{x_2 - x_4}$$

De este modo, von Staudt tenía las herramientas fundamentales para construir la geometría proyectiva sin depender de las nociones de longitud y congruencia.

Un conjunto armónico de cuatro puntos es uno para el cual la razón doble es -1 . Sobre la base de los conjuntos armónicos von Staudt aportó la definición fundamental de que dos haces de puntos están relacionados proyectivamente cuando bajo una correspondencia uno-a-uno de sus elementos, a un conjunto armónico corresponde un conjunto armónico. Cuatro rectas concurrentes forman un conjunto armónico si los puntos en los que cortan a una transversal arbitraria constituyen un conjunto armónico de puntos. Así puede ser definida también la proyectividad de dos haces de líneas. Con estas nociones von Staudt definió una colineación del plano en sí mismo como una transformación uno-a-uno de punto a punto y línea a línea y demostró que lleva un conjunto de elementos armónicos sobre un conjunto de elementos armónicos.

La principal contribución de von Staudt en su *Geometrie der Lage* fue demostrar que la geometría proyectiva es claramente más fundamental que la euclídea. Sus conceptos son lógicamente anteriores. Este libro y su *Beitrdge zur Geometrie der Lage* (Contribuciones a la geometría de la posición, 1856, 1857, 1860) revelaba la geometría proyectiva como una materia independiente de la distancia. Sin embargo, utilizó el axioma de las paralelas de la geometría euclídea, lo que desde un punto de vista lógico era una mancha, porque el paralelismo no es un invariante proyectivo. Esta mancha fue suprimida por Félix Klein⁶⁵.

4. Geometría proyectiva algebraica

Mientras los geómetras sintéticos desarrollaban la geometría proyectiva, los geómetras algebraicos seguían tratando la misma materia con sus métodos. La primera de las nuevas ideas algebraicas fue la que ahora se denomina coordenadas homogéneas. Tal esquema fue creado por August Ferdinand Möbius (1790-1868) quien, como Gauss y Hamilton, vivía de la astronomía pero dedicó considerable tiempo a las matemáticas. Aunque Möbius no tomó parte en la controversia de los métodos sintéticos contra los algebraicos, sus contribuciones estuvieron del lado algebraico.

Su esquema para representar los puntos de un plano mediante coordenadas, introducido en su trabajo principal, *Der barycentrische Calcul*, consistió en empezar con un triángulo fijo y tomar como coordenadas de cualquier punto P en el plano la cantidad de masa que debe ser colocada en cada uno de los tres vértices de un triángulo, de tal forma que P sea el centro de gravedad de las tres masas. Cuando P está fuera del triángulo, entonces una o dos de las coordenadas pueden ser negativas. Si las tres masas son multiplicadas por la misma constante, el punto P sigue siendo el centro de gravedad. De aquí que sólo las proporciones de los tres están determinadas. El mismo esquema aplicado a puntos en el espacio requiere de cuatro coordenadas. Al escribir las ecuaciones de las curvas y superficies en este sistema coordenado, las ecuaciones se hacen homogéneas; esto es, todos los

términos tienen el mismo grado. Veremos ejemplos del uso de coordenadas homogéneas en breve.

Möbius distinguió los tipos de transformaciones de un plano o espacio en otro. Si las figuras correspondientes son iguales, la transformación es una congruencia, y si son semejantes, la transformación es una semejanza. Siguiendo en generalidad viene la transformación que preserva el paralelismo, pero no la longitud o la forma y este tipo es llamado afin (noción introducida por Euler). La transformación más general que lleva rectas en rectas la llamó una colineación. Möbius demostró en su *Der barycentrische Calcul*⁶⁶ que cada colineación es una transformación proyectiva; esto es, que resulta de una sucesión de perspectivas. Su demostración suponía que la transformación es uno-a-uno y continua, pero la condición de continuidad puede ser reemplazada por una más débil. También proporcionó una representación analítica de la transformación. Como señaló Möbius, se podían considerar propiedades invariantes de las figuras bajo cada uno de los tipos anteriores de transformaciones.

Möbius introdujo elementos con signo en geometría no solamente para los segmentos lineales sino también para las áreas y volúmenes. Consecuentemente, fue capaz de dar un tratamiento completo de la noción de razón doble con signo de cuatro puntos sobre una recta.

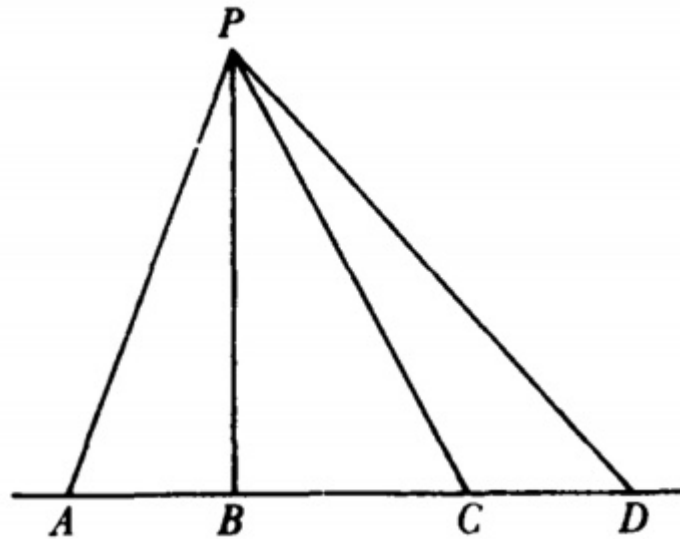


Figura 35.11

Asimismo, demostró que la razón doble de cuatro rectas de un haz es posible expresarla en términos de los senos de los ángulos en el vértice P (fig. 35.11) mediante

$$\frac{\text{sen } APB}{\text{sen } APC} / \frac{\text{sen } BPD}{\text{sen } CPD}$$

y esta proporción es la misma que la razón doble de los cuatro puntos A, B, C y D cortados sobre cualquier transversal a las rectas del haz. De aquí que la razón doble quede inalterada por proyección y sección. Möbius tenía muchas otras ideas que desarrolló lentamente y no llevó muy lejos.

El hombre que dio al enfoque algebraico de la geometría proyectiva su eficacia y vitalidad es Julius Plücker (1801-1868). Después de ser profesor de matemáticas en varias instituciones desde 1836, se convirtió en profesor de matemáticas y física en Bonn, posición que

mantuvo por el resto de su vida. Plücker era esencialmente un físico, y de hecho un físico experimental, en cuya actividad realizó muchos notables descubrimientos. A partir de 1863 se dedicó de nuevo a las matemáticas.

Plücker introdujo también las coordenadas homogéneas, pero de una manera diferente a Möbius. Su primera noción, las coordenadas trilineales,⁶⁷ fue presentada en el segundo volumen de su *Analytischgeometrische Entwicklungen* (Desarrollo de la geometría analítica, 1828 y 1831). Empieza con un triángulo fijo y toma las coordenadas de cualquier punto P como las distancias perpendiculares con signo de P a los lados del triángulo; cada distancia puede ser multiplicada por la misma constante arbitraria. Más adelante, en el segundo volumen, introdujo un caso especial que consiste en considerar una recta del triángulo como recta del infinito. Esto es equivalente a reemplazar las coordenadas cartesianas rectangulares usuales x e y por $x = x_1/x_3$ e $y = x_2/x_3$ de tal forma que las ecuaciones de las curvas se hacen homogéneas en x_1 , x_2 y x_3 . La última noción es la que fue usada más extensamente. Mediante el uso de las coordenadas homogéneas y el teorema de Euler sobre funciones homogéneas, que establece que si $f(tx_1, tx_2, tx_3) = t^n f(x_1, x_2, x_3)$ entonces

$$x_1 \frac{\partial f}{\partial x_1} + x_2 \frac{\partial f}{\partial x_2} + x_3 \frac{\partial f}{\partial x_3} = n f(x_1, x_2, x_3)$$

Plücker fue capaz de dar elegantes representaciones algebraicas de ideas geométricas. Así, si $f(x_1, x_2, x_3) = 0$ es la ecuación de una cónica con $f(x_1, x_2, x_3)$ como coordenadas de un punto sobre la cónica, entonces

$$\frac{\partial f}{\partial x_1} x'_1 + \frac{\partial f}{\partial x_2} x'_2 + \frac{\partial f}{\partial x_3} x'_3 = 0$$

puede ser interpretado, cuando x'_1, x'_2, x'_3 son las coordenadas «libres», como la ecuación de la tangente en el punto (x_1, x_2, x_3) o, cuando x_1, x_2, x_3 son las coordenadas «libres», como la ecuación de la recta polar de un punto arbitrario (x'_1, x'_2, x'_3) con respecto a la cónica.

Usando coordenadas homogéneas, Plücker proporcionó la formulación algebraica de la recta infinitamente distante, los puntos circulares en el infinito y otras nociones. En el sistema de coordenadas homogéneas (x_1, x_2, x_3) , la ecuación de la recta infinitamente distante es $x_3 = 0$. Esta recta no es excepcional en geometría proyectiva, pero en nuestra visualización de los elementos geométricos cada punto normal del plano euclídeo está en una posición finita dada por $x = x_1/x_3$ e $y = x_2/x_3$ y de esta manera estamos obligados a pensar en los puntos sobre $x_3 = 0$ como infinitamente distantes.

La ecuación de un círculo

$$(x - a)^2 + (y - b)^2 = r^2$$

se convierte con la introducción de coordenadas cartesianas homogéneas x_1, x_2, x_3 a través de $x = x_1/x_3$ e $y = x_2/x_3$ en

$$(x_1 - ax_3)^2 + (x_2 - bx_3)^2 = r^2 x_3^2$$

Ya que la ecuación de la línea infinitamente distante es $x_3 = 0$, la intersección de esta línea con el círculo está dada por

$$x_1^2 + x_2^2 = 0 \quad y \quad x_3 = 0$$

y esta es la ecuación de los puntos circulares en el ∞ . Tales puntos circulares tienen las coordenadas $(1, i, 0)$ y $(1, -i, 0)$ o tripletas correspondientes a ellos. De la misma manera la ecuación de un círculo esférico (Kugelkreis) en el infinito es

$$x_1^2 + x_2^2 + x_3^2 = 0 \quad y \quad x_1 = 0$$

Si escribimos la ecuación de una línea recta en la forma homogénea (debemos usar x, y y z en lugar de x_1, x_2 y x_3)

$$Ax + By + Cz = 0$$

y ahora se requiere que la recta pase por los puntos (x_1, y_1, z_1) y $(1, i, 0)$ entonces la ecuación no homogénea resultante de la recta es

$$x - x_0 + i(y - y_0) = 0$$

donde $x_0 = x_1/z_1$ e $y_0 = y_1/z_1$. De la misma manera, la ecuación de la recta que pasa por (x_1, y_1, z_1) y $(1, -i, 0)$ es

$$x - x_0 - i(y - y_0) = 0$$

Cada una de estas rectas es perpendicular a sí misma, ya que la pendiente es igual a su recíproco negativo. Sophus Lie las llamó rectas locas; ahora son llamadas rectas isotrópicas.

Los esfuerzos de Plücker para tratar la dualidad algebraicamente lo llevaron a una idea muy bella, las coordenadas de rectas.⁶⁸ Si la ecuación de la recta en coordenadas homogéneas es

$$ux + vy + wz = 0,$$

y si x , y y z son cantidades fijas, u , v y w o cualesquiera tres números proporcionales a ellos son las coordenadas de una recta en un plano. Entonces, de la misma manera que la ecuación $f(x_1, x_2, x_3) = 0$ representa una colección de puntos, así $f(u, v, w) = 0$ representa una colección de rectas o una curva reglada.

Con esta noción de coordenadas de rectas, Plücker fue capaz de proporcionar una formulación algebraica y una demostración del principio de dualidad. Dada cualquier ecuación $f(r, s, t) = 0$, si interpretamos r , s y t como las coordenadas homogéneas x_1 , x_2 y x_3 de un punto, entonces tenemos la ecuación de una curva de puntos,

mientras que si los interpretamos como u , v y w tenemos el dual de una curva de rectas. Cualquier propiedad demostrada mediante procesos algebraicos para la curva de puntos será, debido a que el álgebra es la misma bajo cualquier interpretación de las variables, origen de la propiedad dual de la curva de rectas.

En este segundo ensayo de 1830 y en el volumen 2 de su *Entwickelungen*, Plücker señaló también que la misma curva considerada como una colección de puntos puede asimismo ser considerada como la colección de rectas tangentes a la curva, ya que las tangentes determinan el contorno igual que lo hacen los puntos. La familia de las tangentes es una curva de rectas y tiene una ecuación en coordenadas de rectas. El grado de esta ecuación es llamado la clase de la curva, mientras que el grado de la ecuación en coordenadas de puntos es llamado el orden.

5. Curvas planas de orden superior y superficies

Los autores del siglo XVIII habían realizado algunos trabajos sobre curvas de grado superior al segundo (cap. 23, sec. 3) pero la materia estuvo latente desde 1750 a 1825. Plücker estudió curvas de grado tercero y cuarto y utilizó libremente conceptos proyectivos en su trabajo.

En su *System der analytischen Geometrie* (Sistema de geometría analítica, 1834) usó un principio que, aunque muy útil, difícilmente estaba bien fundamentado, para establecer formas canónicas de curvas. Para mostrar que la curva general de orden (grado) 4, digamos, podía ser puesta en una forma canónica, argumentó en

particular que si el número de constantes fuera el mismo en las dos formas entonces se podía convertir una en la otra. Así, argumentó que una forma ternaria (tres variables) de cuarto orden podía siempre ser puesta en la forma

$$C_4 = pqrs + \mu\Omega^2,$$

donde p , q , r y s son formas lineales y Ω es una forma cuadrática, ya que ambos miembros contienen 14 constantes. En sus ecuaciones, p era real, así como los coeficientes.

Plücker también estudió el número de puntos de intersección de curvas, un tema que también había sido considerado en el siglo XVIII. Utilizó un esquema para representar la familia de todas las curvas que pasan por las intersecciones de dos curvas C' y C'' de grado n -ésimo que había sido presentado por Lamé en un libro de 1818. Cualquier curva C_n pasando por estas intersecciones se podía expresar como

$$C_n = C'_n + \lambda C'' = 0$$

donde λ es un parámetro.

Utilizando este esquema, Plücker proporcionó una clara explicación de la paradoja de Cramer (cap. 23, sec. 3). Una curva C_n general está determinada mediante $n(n + 3)/2$ puntos, ya que este es el número de coeficientes esenciales en su ecuación. Por otro lado, como dos C_n se intersecan en n^2 puntos, por los $n(n + 3)/2$ puntos

en los cuales las dos C_n se intersecan, pasa un número infinito de otras C_n . La contradicción aparente fue explicada por Plücker⁶⁹. Dos curvas cualesquiera de grado n se cortan, por supuesto, en n^2 puntos. Sin embargo, únicamente $(n/2)(n + 3) - 1$ puntos son independientes. Esto es, si tomamos dos curvas de grado n , que pasan por los $(n/2)(n + 3) - 1$ puntos, cualquier otra curva de grado n que pase, por estos puntos pasará por los restantes $(n - 1)(n - 2)/2$ de los n^2 puntos. Así, cuando $n = 4$, 13 puntos son independientes. Cualesquiera dos curvas que pasan por los 13 puntos determinan 16 puntos, pero cualquier otra curva que pase por los 13 pasará también por los tres restantes.

Plücker atacó después⁷⁰ la teoría de las intersecciones de una curva de grado m y una curva de grado n . Consideró la última como fija y las curvas intersecándola como variables. Usando la notación abreviada C_n para la expresión para la curva de n -ésimo grado y notación similar para las otras escribió

$$C_m = C'_m + A_{m-n}C_n = 0$$

para el caso donde $m > n$, de tal forma que A_{m-n} es un polinomio de grado $m - n$. A partir de esta ecuación obtuvo el método correcto para determinar las intersecciones de una C_n con todas las curvas de grado m . Ya que, de acuerdo con esta ecuación, $m - n + 1$ curvas linealmente independientes (el número de los coeficientes en A_{m-n}) pasan por las intersecciones de C'_m y C_n , la conclusión de Plücker fue que, dados $mn - (n - 1)(n - 1)/2$ puntos arbitrarios sobre C_n , los

restantes $(n - 1)(n - 2)/2$ de los mn puntos de intersección con una C_m están determinados. El mismo resultado fue obtenido casi al mismo tiempo por Jacobi.⁷¹

En su *System* de 1834 y más explícitamente en su *Theorie der algebraischen Curven* (Teoría de curvas algebraicas, 1839), Plücker proporcionó las que ahora son llamadas fórmulas de Plücker, que relacionan el orden n y la clase k de una curva y las singularidades simples. Sea d el número de puntos dobles (puntos singulares en los cuales las dos tangentes son diferentes) y r el número de cúspides. A los puntos dobles les corresponde en la curva de rectas tangentes dobles (una tangente doble es de hecho tangente en dos puntos distintos) el número de las cuales es, digamos, t . A las cúspides les corresponde tangentes osculatrices (tangentes que cruzan la curva en puntos de inflexión) el número de las cuales es, digamos, w . Entonces Plücker pudo establecer las siguientes fórmulas duales

$$\begin{aligned} k &= n(n - 1) - 2d - 3r & n &= k(k - 1) - 2t - 3w \\ w &= 3n(n - 2) - 6d - 8r & r &= 3k(k - 2) - 6t - 8w \end{aligned}$$

El número de cualquier elemento incluye los casos reales e imaginarios.

En el caso cuando $n = 3$, $d = 0$ y $r = 0$, entonces w , el número de puntos de inflexión, es 9. Hasta los tiempos de Plücker, De Gua y Maclaurin habían demostrado que una recta que pasa por dos puntos de inflexión de la curva general de tercer grado pasa por un

tercero, y el hecho de que una C_2 general tiene tres puntos reales de inflexión había sido supuesto desde los tiempos de Clairaut. En su *System* de 1834 Plücker demostró que una C_3 tiene uno o tres puntos reales de inflexión, y en el último caso estos están sobre una recta. También llegó al resultado más general que toma en consideración elementos complejos. Una C_3 general tiene nueve puntos de inflexión de los cuales seis son imaginarios. Para derivar este resultado utilizó su principio de contar constantes para demostrar que

$$C_3 = fgh - l^3,$$

donde f , g , h y l son formas lineales, y obtuvo el resultado de De Gua y Maclaurin. Más adelante demostró (con argumentos incompletos) que los nueve puntos de inflexión de C_3 yacen sobre una recta de tal forma que hay doce de tales rectas. Ludwig Otto Hesse (1811-1874), quien trabajó como profesor en diversas universidades, completó la demostración de Plücker⁷² y mostró que las doce rectas pueden ser agrupadas en cuatro triángulos.

Como un ejemplo adicional de descubrimiento de propiedades generales de las curvas, consideraremos el problema de los puntos de inflexión de una curva $f(x,y) = 0$ de n -ésimo grado. Plücker había expresado las condiciones comunes del cálculo para un punto de inflexión cuando $y = f(x)$, a saber, $d^2y/dx^2 = 0$, en la forma apropiada para $f(x,y) = 0$ y obtuvo una ecuación de grado $3n - 4$. Ya que la curva original y la nueva curva deben tener $n(3n - 4)$

intersecciones, parecía que la curva original tenía $n(3n - 4)$ puntos de inflexión. Como este número era demasiado grande, Plücker sugirió que la curva con una ecuación de grado $3n - 4$ tiene un contacto tangencial con cada una de las n ramas infinitas de la curva original $f = 0$, de tal forma que $2n$ de los puntos comunes no son puntos de inflexión, y así obtuvo el número correcto, $3n(n - 2)$. Hesse aclaró este punto⁷³ valiéndose de coordenadas homogéneas; esto es, reemplazó x por x_1/x_3 e y por x_2/x_3 y usando el teorema de Euler sobre funciones homogéneas demostró que la ecuación de Plücker para los puntos de inflexión puede ser escrita como

$$H = \begin{vmatrix} f_{11} & f_{12} & f_{13} \\ f_{21} & f_{22} & f_{23} \\ f_{31} & f_{32} & f_{33} \end{vmatrix} = 0$$

donde los subíndices denotan derivadas parciales. Esta ecuación es de grado $3(n - 2)$ y así toca a la curva $f(x_1, x_2, x_3) = 0$ de n -ésimo grado en el número correcto de puntos de inflexión. El determinante es llamado el hessiano de f , una noción introducida por Hess.⁷⁴

Entre otros, Plücker abordó la investigación de curvas cuárticas. Fue el primero en descubrir (*Theorie der algebraischen Curven*, 1839) que tales curvas contienen 28 tangentes dobles de las cuales ocho a lo más son reales. Jacobi demostró entonces⁷⁵ que una curva de grado n -ésimo tiene en general $n(n - 2) (n^2 - 9)/2$ tangentes dobles.

En el trabajo en geometría algebraica también se consideraron figuras en el espacio. Aunque la representación de líneas rectas en

el espacio ya había sido introducida por Euler y Cauchy, Plücker en su *System der Geometrie des Raumes* (Sistema de geometría del espacio, 1846) introdujo una representación modificada

$$x = rz + \rho, y = sz + \sigma$$

en la que los cuatro parámetros r , ρ , s y σ fijan la recta. Ahora las rectas pueden ser usadas para construir todo el espacio, ya que, por ejemplo, los planos no son más que colecciones de rectas, y los puntos son intersecciones de rectas. Plücker dijo entonces que si las rectas eran consideradas como los elementos fundamentales del espacio, el espacio es tetradimensional porque son necesarios cuatro parámetros para cubrir todo el espacio con rectas. La noción de un espacio tetradimensional de puntos la rechazó como demasiado metafísica. Que la dimensión depende del elemento espacial es una idea nueva.

El estudio de figuras en el espacio incluyó superficies de tercer y cuarto grado. Una superficie reglada está generada por una recta moviéndose de acuerdo con alguna ley. El paraboloides hiperbólico (superficie con una silla) y el hiperboloides de una hoja son ejemplos de ellas, como lo es el helicoides. Si una superficie de segundo orden contiene una recta, contiene además una infinidad de rectas y es reglada. (Debe entonces ser un cono, un cilindro, el paraboloides hiperbólico, o el hiperboloides de una hoja.) Sin embargo, esto no es cierto para las superficies cúbicas.

Como ejemplo de propiedad notable de las superficies cúbicas tenemos el descubrimiento de Cayley en 1849⁷⁶ de la existencia de exactamente 27 rectas sobre cada superficie de tercer grado. No todas son necesariamente reales, pero existen superficies para las que todas son reales. Clebsch proporcionó un ejemplo de 1871⁷⁷. Estas rectas tienen propiedades especiales. Por ejemplo, cada una es cortada por otras diez. Mucho más trabajo fue dedicado al estudio de estas rectas sobre superficies cúbicas.

Entre los muchos descubrimientos relativos a superficies de cuarto orden, uno de ellos, que se debe a Kummer, merece atención especial. Kummer había trabajado con familias de rectas que representan rayos de luz y, considerando las superficies focales asociadas⁷⁸, fue llevado a introducir una superficie de cuarto grado (y de cuarta clase) con 16 puntos dobles y 16 planos dobles como superficie focal de un sistema de rayos de segundo orden. Esta superficie, conocida como superficie de Kummer, contiene como caso especial la superficie de onda de Fresnel, que representa el frente de onda de la luz propagándose en un medio anisótropo.

La labor sobre geometría proyectiva sintética y algebraica de la primera mitad del siglo XIX abrió un período brillante para las investigaciones geométricas de toda clase. Los geómetras sintéticos dominaron el período. Atacaron cada resultado nuevo para descubrir algún principio general, usualmente no demostrable geoméricamente, pero del que sin embargo derivaron un torrente de consecuencias ligadas unas a otras, y al principio general. Afortunadamente, también fueron introducidos los métodos

algebraicos y, como veremos, dominaron finalmente el campo. Sin embargo, debemos interrumpir la historia de la geometría proyectiva para considerar algunas nuevas creaciones revolucionarias que afectaron el trabajo posterior en geometría y, que de hecho, alteraron radicalmente la imagen de la matemática.

Bibliografía

- Berzolari, Luigi: «*Allgemeine Theorie der hóheren ebenen algebraischen Kurven*». Encyk. der Math. Wiss., B. G. Teubner, 1903-1915, III C 4, 313-455.
- Boyer, Cari B.: *History of Analytic Geometry, Scripta Mathematica*, 1956, caps. 8-9.
- Brill, A., y Noether, M.: «*Die Entwicklung der Theorie der algebraischen Functionen in álterer und neurerer Zeit*», Jahres. der Deut. Math.-Verein., 3, 1892/3, 109-566, 287-312 en particular.
- Cajori, Florian: *A history of mathematics*, 2.a ed., Macmillan, 1919, 289-302, 309-314.
- Coolidge, Julián L.: *A treatise on the árele and the sphere*, Oxford University Press, 1916. - *A history of geometrical methods*, Dover (reimpresión), 1963, Libro I, cap. 5 y libro II, cap. 2. - *A history of the conic sections and quadric surfaces*. Dover (reimpresión), 1968. - «*The rise and fall of proyective geometry*». American Mathematical Monthly, 41, 1934, 217-228.

- Fano, G.: «*Gegensatz von synthetischer und analytischer Geometrie in seiner historischen Entwicklung im XIX, Jahrhundert*», Encyk. der Math. Wiss., B. G. Teubner, 1907-1910, III AB4a, 221-288.
- Klein, Félix: *Elementary mathematics from an advanced standpoint*, Macmillan, 1939; Dover (reimpresión), 1945, Geometría, parte 2. Hay versión castellana, Matemática elemental desde un punto de vista superior (2 volúmenes), Madrid, Biblioteca Matemática, 1927 y 1931.
- Kotter, Ernst: «*Die Entwicklung der synthetischen Geometrie von Monge bis auf Staudt*», 1847, Jahres. der Deut. Math-Verein., vol. 5, parte II, 1896 (pub. 1901), 1-486.
- Möbius, August F.: *Der barycentrische Calcul* (1827), Georg Olms (reimpresión), 1968. También en vol. 1 de *Gesammelte Werke*, 1-388.
- Plücker, Julius: *Gesammelte wissenschaftliche Abhandlungen*, 2 vols., B. G. Teubner, 1895-1896.
- Schoenflies, A.: «*Projektive Geometrie*», Encyk. der Math. Wiss., B. G. Teubner, 1907-1910, III AB5, 389-480.
- Smith, David Eugene: *A source book in mathematics*, Dover (reimpresión), 1959, vol. 2, 315-323, 331-345, 670-676.
- Steiner, Jacob: *Geometrical constructions with a ruler* (una traducción de su libro de 1833), *Scripta Mathematica*, 1950. - *Gesammelte Werke*, 2 vols., G. Reimer, 1881-1882; Chelsea (reimpresión), 1971.

- Zacharias, M.: «*Elementargeometrie and elementare nicht-euklidische Geometrie in synthetischer Behandlung*», Encyk. der Math. Wiss., B. G. Teubner, 1907-1910, III, AB9, 859-1172.

Capítulo 36

La geometría no euclídea

... porque parece ser cierto que muchas cosas tienen, por así decirlo, una época en la que son descubiertas en muchos lugares simultáneamente, como las violetas aparecen en primavera por todas partes.

Wolfgang Bolyai

El carácter de necesidad adscrito a las verdades de la matemática, y aún la peculiar certeza a ellas atribuida, son una ilusión.

John Stuart Mill

Contenido:

- 1. Introducción*
- 2. El status de la geometría euclídea hacia 1800*
- 3. Las investigaciones sobre el axioma de las paralelas*
- 4. Presagios de las geometrías no euclídeas*
- 5. La creación de la geometría no euclídea*
- 6. El contenido técnico de la geometría no euclídea*
- 7. Las demandas de prioridad de Lobatchevsky y Bolyai*
- 8. Las implicaciones de la geometría no euclídea*

Bibliografía

1. Introducción

Entre todas las complejas creaciones técnicas del siglo XIX la más profunda, la geometría no euclídea, fue técnicamente la más simple. Esta creación dio origen a nuevas e importantes ramas de las matemáticas, pero su implicación más significativa es que obligó a los matemáticos a revisar radicalmente su comprensión de la naturaleza de las matemáticas y su relación con el mundo físico. Dio origen a problemas en los fundamentos de las matemáticas con los que el siglo XX sigue luchando. Como veremos, la geometría no euclídea fue la culminación de una larga serie de esfuerzos en el área de la geometría euclídea. El fruto de este trabajo llegó en la primera parte del siglo XIX, durante las mismas décadas en las que la geometría proyectiva estaba siendo revivida y extendida. Sin embargo, los dos dominios no estuvieron relacionados uno con otro en esta época.

1. El status de la geometría euclídea hacia 1800

A pesar de que los griegos habían reconocido que el espacio abstracto o matemático es distinto de las percepciones sensoriales del espacio, y Newton subrayó este punto⁷⁹, hasta aproximadamente 1800 todos los matemáticos estaban convencidos de que la geometría euclídea era la idealización correcta de las propiedades del espacio físico y de las figuras en ese espacio. De hecho, como ya hemos notado, hubo muchos intentos por construir la aritmética, el

álgebra y el análisis, cuyos fundamentos lógicos eran oscuros, sobre la geometría euclídea y, por ende, garantizar también la verdad de esas ramas.

De hecho, muchos proclamaron la verdad absoluta de la geometría euclídea. Por ejemplo, Isaac Barrow, quien construyó sus matemáticas (incluyendo su cálculo) sobre la geometría, da una lista de ocho razones de la certeza de la geometría: la claridad de sus conceptos, las definiciones no ambiguas, la seguridad intuitiva y la verdad universal de sus axiomas, la posibilidad clara y lo fácil de imaginar de sus postulados, el pequeño número de sus axiomas, la visión clara del modo en que las magnitudes son generadas, el fácil orden de sus demostraciones y la elusión de cosas no conocidas.

Barrow formuló la pregunta: ¿cómo estamos seguros de que los principios geométricos se aplican a la naturaleza? Su respuesta fue que éstos son derivados de la razón innata. Los objetos sensibles son únicamente los agentes que los despiertan. Más aún, los principios de la geometría han sido confirmados mediante la experiencia constante y continuará siendo así porque el mundo, diseñado por Dios, es inmutable. La geometría es entonces la ciencia perfecta y segura.

Es relevante que los filósofos de la última parte del siglo XVII y el XVIII también formularan la pregunta de cómo podemos estar seguros de que la mayor parte del cuerpo del conocimiento que la ciencia newtoniana ha producido era cierto. Casi todos, pero notablemente Hobbes, Locke y Leibniz, contestaron que las leyes matemáticas, como la geometría euclídea, eran inherentes al diseño

del universo. Leibniz dejó cierto espacio para la duda cuando distinguió entre mundos posibles y reales. Pero la única excepción significativa fue David Hume, quien en su *Treatise of human nature* (*Tratado sobre la naturaleza humana*, 1739), negó la existencia de leyes o secuencias necesarias de eventos en el universo, y afirmó que se observaba que ocurrían estas secuencias y los seres humanos concluyeron que siempre ocurrirán de la misma manera. La ciencia es puramente empírica. En particular, las leyes de la geometría euclídea no son necesariamente verdades físicas.

La influencia de Hume fue negada y claramente superada por la de Immanuel Kant. La respuesta de Kant a la pregunta de cómo podemos estar seguros de que la geometría euclídea se aplica al mundo físico', que proporcionó en su *Crítica de la razón pura* (1781), es peculiar: Kant sostenía que nuestra mente suministra ciertos modos de organización —los llamó intuiciones— del espacio y el tiempo y que la experiencia es absorbida y organizada por nuestras mentes de acuerdo con esos modos o intuiciones. Nuestras mentes están construidas de tal forma que nos obligan a ver el mundo externo sólo de una manera. Como consecuencia, ciertos principios acerca del espacio son anteriores a la experiencia; estos principios y sus consecuencias lógicas, que Kant llamó verdades sintéticas *a priori*, son las de la geometría euclídea. Conocemos la naturaleza del mundo exterior sólo en la medida en que nuestras mentes nos obligan a interpretarla. Sobre las bases antes descritas, Kant afirmó, y sus contemporáneos aceptaron, que el mundo físico debía ser euclídeo. En cualquier caso, ya sea que se recurra a la

experiencia, se apoye sobre verdades innatas, o acepte el punto de vista de Kant, la conclusión común era la unidad y necesidad de la geometría euclídea.

2. Las investigaciones sobre el axioma de la paralelas

A pesar de que la confianza en la geometría euclídea como la idealización correcta del espacio físico permaneció firme del 300 a. de C. hasta cerca de 1800, una preocupación ocupó a los matemáticos durante casi todo ese período. Los axiomas adoptados por Euclides se suponía que eran verdades autoevidentes acerca del espacio físico y de las figuras en ese espacio. Sin embargo, el axioma de las paralelas en la forma establecida por Euclides (cap. IV, sec. 3) se pensaba que era algo demasiado complicado. Nadie dudaba realmente de su verdad, y sin embargo carecía de la cualidad obligada de los otros axiomas. Aparentemente, al propio Euclides no le gustaba su propia versión del axioma de las paralelas, ya que no lo utilizó sino después de que hubo demostrado todos los teoremas que pudo sin él.

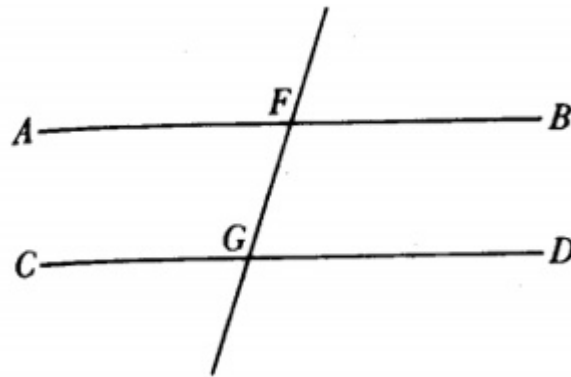
Un problema relacionado que no preocupó a tanta gente pero que, sin embargo, finalmente llegó a ser igualmente vital es si se puede suponer la existencia de líneas rectas infinitas en el espacio físico. Euclides fue cuidadoso en postular que uno únicamente puede prolongar una línea recta (finita) tan lejos como sea necesario, de tal forma que aún la recta extendida seguía siendo finita. También la redacción peculiar del axioma de las paralelas de Euclides, que dos rectas se cortarán del lado de la transversal donde la suma de los

ángulos interiores es menor que dos ángulos rectos, era una manera de evitar la aserción inequívoca de que existen pares de rectas que nunca se llegan a cortar por lejos que sean extendidas. Sin embargo, Euclides sugirió la existencia de líneas rectas infinitas ya que, de ser finitas, no podían ser extendidas tanto como fuera necesario en un contexto dado y *demonstró* la existencia de rectas paralelas.

La historia de la geometría no euclídea se inicia con los esfuerzos para eliminar las dudas acerca del axioma de las paralelas de Euclides. Desde los tiempos de los griegos hasta aproximadamente 1800, se llevaron a cabo dos intentos. Uno fue reemplazar el axioma de las paralelas por un argumento más autoevidente. El otro fue tratar de deducirlo de los otros nueve axiomas de Euclides; de ser esto posible, sería un teorema y así estaría fuera de toda duda. No daremos una reseña completamente detallada de este trabajo porque su historia se encuentra fácilmente disponible.⁸⁰

El primer intento importante fue hecho por Ptolomeo en su opúsculo sobre el axioma de las paralelas. Intentó demostrar esta aseveración deduciéndola de los otros nueve axiomas de Euclides y de los teoremas 1 a 28, que no dependen del axioma de las paralelas. Pero Ptolomeo asumió inconscientemente que dos líneas rectas no encierran un espacio y que si AB y CD son paralelas (fig. 36.1) entonces todo lo que se tenga para los ángulos interiores sobre un lado de FG debe mantenerse para el otro.

Proclo, el comentador del siglo **V**, fue muy explícito en cuanto a esta objeción al axioma de las paralelas.

*Figura 36.1*

Dice: «Este [postulado] debería ser incluso ser eliminado de los postulados; ya que es un teorema que supone demasiadas dificultades, que Ptolomeo, en cierto libro, se propuso resolver, y requiere para su demostración un gran número de definiciones así como de teoremas; y el recíproco es, de hecho, demostrado por el propio Euclides como un teorema». Proclo señala que mientras que es necesario creer que dos líneas rectas tenderán una hacia la otra del lado de la transversal donde la suma de los ángulos interiores es menor que dos ángulos rectos, no es tan claro que estas dos rectas se intersecarán de hecho en un punto finito. Esta conclusión es únicamente probable. Ya que, continúa, hay ciertas curvas que se aproximan una a la otra más y más, pero que de hecho no se juntan. Así, una hipérbola se aproxima, pero no toca, a su asíntota. ¿No sería esto verdadero para las dos líneas en el postulado de Euclides? Más adelante dice que hasta cierta suma de los ángulos interiores de un lado de la transversal, las dos líneas deben realmente cortarse; sin embargo, para un valor escasamente más

grande pero aún menor que dos ángulos rectos, las líneas podían ser asintóticas.

Proclo basó su propia demostración del postulado de las paralelas en un axioma que Aristóteles había usado para demostrar que el universo es finito. El axioma dice: «si desde un punto dos líneas rectas formando un ángulo son extendidas indefinidamente, las distancias sucesivas entre dichas líneas rectas [perpendiculares de una a la otra] excederán finalmente cualquier magnitud finita». La demostración de Proclo era correcta esencialmente, pero substituyó un axioma cuestionable por otro.

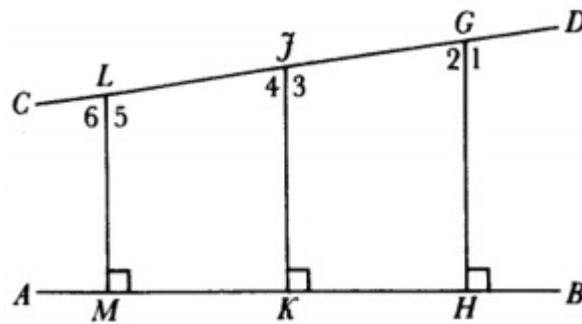


Figura 36.2

Nasir Eddin (1201-1274), el editor persa de Euclides, proporcionó análogamente una «demostración» del postulado de las paralelas de Euclides suponiendo que dos líneas no paralelas se aproximan una a la otra en una dirección y divergen en la otra. Específicamente, si AB y CD (fig. 36.2) son dos rectas cortadas por GH , JK , LM ,..., si estas últimas son perpendiculares a AB , y si los ángulos 1, 3, 5, ... son obtusos mientras que 2, 4, 6,... son agudos, entonces $GH > JK > LM$... Este hecho, afirmar Nasir Eddin, se ve claramente.

Wallis realizó cierto trabajo sobre el axioma de las paralelas en 1663, que publicó en 1693.⁸¹ Primero, reprodujo el trabajo de Nasir Eddin sobre el axioma de las paralelas que había sido traducido para él por un profesor de árabe de Oxford. Incidentalmente, así fue como el trabajo de Nasir-Eddin sobre el axioma de las paralelas fue conocido en Europa. Entonces, Wallis criticó la demostración de Nasir Eddin y ofreció su propia demostración de la aserción de Euclides. Su demostración se apoya en la suposición explícita de que para cada triángulo existe uno semejante, cuyos lados están en una razón dada con los lados del original. Wallis pensaba que este axioma era más evidente que la subdivisión arbitrariamente pequeña y la extensión arbitrariamente grande. De hecho, dijo, el axioma de Euclides de que podemos construir un círculo con un centro y radio dados presupone que existe un radio arbitrariamente grande a nuestra disposición. De aquí que se pueda suponer el resultado análogo para figuras rectilíneas tales como un triángulo.

En 1769, Joseph Fenn sugirió el más sencillo de los axiomas substitutos, a saber, que dos líneas rectas intersecándose no pueden ser ambas paralelas a una tercera línea recta. Este axioma también aparece en los comentarios de Proclo sobre la proposición 31 del libro I de los *Elementos* de Euclides. El argumento de Fenn es equivalente por completo al axioma dado en 1795 por John Playfair (1748-1819): por un punto dado P que no está sobre una recta l , existe una sola recta en el plano de P y l que no corta a l . Este axioma se usa en los libros de texto modernos (que por simplicidad dicen comúnmente: «una y sólo una recta...»).

Legendre trabajó sobre el problema del postulado de las paralelas a lo largo de un período de veinte años. Sus resultados aparecieron en libros y artículos, incluyendo las muchas ediciones de sus *Eléments de géométrie (Elementos de Geometría)*.⁸² En uno de los ataques al problema demostró el postulado de las paralelas suponiendo que existen triángulos semejantes de diferentes tamaños; de hecho esta demostración era analítica, pero supuso que la unidad de longitud no tenía importancia. Más adelante proporcionó una demostración basada en la suposición de que dados tres puntos cualesquiera no colineales existe una circunferencia pasando por los tres. En otro enfoque usó todos los postulados, excepto el de las paralelas, y demostró que la suma de los ángulos de un triángulo no puede ser mayor que dos ángulos rectos. Entonces observó que bajo estas mismas suposiciones el área es proporcional al defecto, esto es, dos ángulos rectos menos la suma de los ángulos. Por tanto, trató de construir un triángulo de dos veces el tamaño de un triángulo dado de tal manera que el defecto del más grande fuera dos veces el del dado. Procediendo de esta manera, esperó obtener triángulos con defectos mayores y menores y obtener así sumas de ángulos aproximándose a cero. Este resultado, pensó, sería absurdo, y así la suma de los ángulos tendría que ser 180° . Este hecho, a su vez, implicaría el axioma euclídeo de las paralelas. Pero Legendre encontró que la construcción se reducía a demostrar que a través de cualquier punto dado dentro de un ángulo menor de 60° , uno siempre puede trazar una línea recta que toque a ambos lados del ángulo, y esto no lo podía demostrar sin hacer uso del postulado de

las paralelas de Euclides. En cada una de las doce ediciones (12.^a ed., 1813) de la versión de Legendre de los *Elementos* de Euclides añadió apéndices que supuestamente proporcionaban demostraciones del postulado de las paralelas, pero cada uno era deficiente porque suponía algo implícito que no podía ser supuesto o bien aceptaba axiomas tan cuestionables como el de Euclides.

En el transcurso de sus investigaciones, Legendre,⁸³ usando los axiomas euclídeos con excepción del axioma de las paralelas, demostró los siguientes teoremas significativos: si la suma de los ángulos de un triángulo es dos ángulos rectos, entonces es así en cada triángulo. También, si la suma es menor que dos ángulos rectos en un triángulo, así es también en cada triángulo. Más adelante, proporciona la demostración de que si la suma de los ángulos de cualquier triángulo es dos ángulos rectos, se mantiene el postulado de las paralelas de Euclides. Este trabajo sobre la suma de los ángulos de un triángulo tampoco tuvo frutos, porque Legendre fue incapaz de demostrar (sin la ayuda del axioma de las paralelas o un axioma equivalente) que la suma de los ángulos de un triángulo no puede ser menos que dos ángulos rectos.

Los esfuerzos descritos hasta ahora fueron simplemente intentos por encontrar un sustituto más autoevidente del axioma de las paralelas de Euclides y muchos de los axiomas propuestos parecían ser intuitivamente más autoevidentes. Consecuentemente, sus creadores estaban satisfechos, ya que habían cumplido sus objetivos. Sin embargo, un examen más detenido mostró que estos axiomas sustitutos no eran necesariamente más satisfactorios.

Algunos de ellos hicieron aserciones acerca de lo que sucede indefinidamente lejos en el espacio. Así, requerir que hubiera un círculo pasando por tres puntos cualesquiera no en línea recta requiere círculos más y más grandes en cuanto los puntos se aproximan a estar alineados. Por otro lado, los axiomas sustitutos que no hacían intervenir el infinito directamente, por ejemplo, el axioma que establece que existen dos triángulos semejantes pero no iguales, se veían como suposiciones demasiado complejas y de ninguna manera preferibles al propio axioma de Euclides.

El segundo grupo de esfuerzos para resolver el problema del axioma de las paralelas buscaba deducir la aserción de Euclides de los otros nueve axiomas. La deducción podía ser directa o indirecta. Ptolomeo había intentado una prueba directa. El método indirecto supone una aserción contradictoria en lugar del enunciado de Euclides e intenta deducir una contradicción dentro del nuevo cuerpo de teoremas deducidos. Por ejemplo, ya que el axioma de las paralelas de Euclides es equivalente al axioma de que por un punto P que no está sobre una recta l hay una y solamente una paralela a l , existen dos alternativas a este axioma. Una es que no hay paralelas a l pasando por P y la otra es que existe más de una paralela a l pasando por P . Si al adoptar una de estas posibilidades en lugar del axioma de «una paralela» se demuestra que el nuevo conjunto conduce a una contradicción, entonces estas alternativas tendrían que ser rechazadas y la aserción de «una paralela» quedaría *demostrada*.

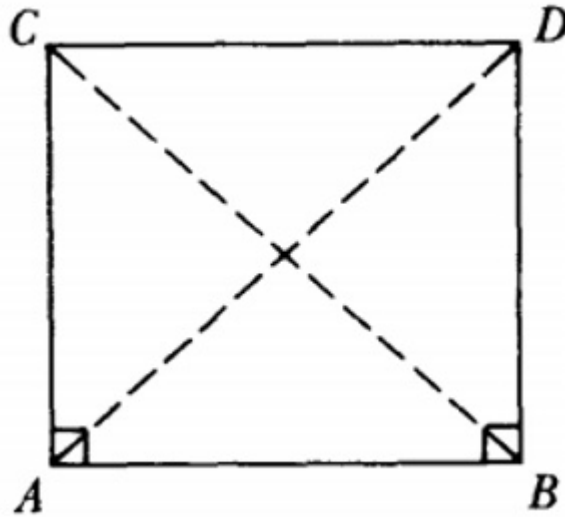


Figura 36.3

Gerolamo Saccheri (1667-1733), sacerdote jesuita y profesor de la universidad de Padua, realizó el esfuerzo más significativo en esta dirección. Estudió cuidadosamente los trabajos de Nasir Eddin y Wallis y adoptó su propio enfoque. Saccheri empezó con el cuadrilátero $ABCD$ (fig. 36.3) en el que A y B son ángulos rectos y $AC = BD$. Es fácil demostrar que $\angle C = \angle D$. Ahora, el axioma de Euclides es equivalente a la aserción de que los ángulos C y D son ángulos rectos. Entonces Saccheri consideró dos alternativas posibles:

1. la hipótesis del ángulo obtuso: $\angle C$ y $\angle D$ son obtusos;
2. la hipótesis del ángulo agudo: $\angle C$ y $\angle D$ son agudos.

Sobre la base de la primera hipótesis (y los otros nueve axiomas de Euclides), Saccheri demostró que los ángulos C y D deben ser ángulos rectos. Así, de esta hipótesis dedujo una contradicción.

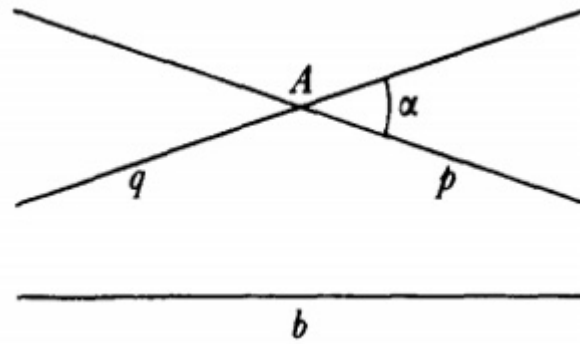


Figura 36.4

A continuación, Saccheri consideró la segunda hipótesis y demostró muchos teoremas interesantes. Continuó hasta que llegó al siguiente teorema: dado cualquier punto A y una recta b (fig. 36.4), con la hipótesis del ángulo agudo existe en el haz (familia) de rectas que pasan por A dos rectas p y q que dividen el haz en dos partes. La primera de éstas consiste de las líneas que intersecan a b , y la segunda consiste de las líneas (que forman un ángulo α) que tienen una perpendicular común con b en algún sitio a lo largo de b . Las rectas p y q ellas mismas son asintóticas a b . De este resultado y un argumento sumamente largo, Saccheri dedujo que p y b tendrían una perpendicular común en su punto común, el cual está en el infinito. A pesar de que no había obtenido contradicción alguna, Saccheri encontró esta conclusión, y otras, tan repugnantes que decidió que la hipótesis del ángulo agudo debía ser falsa.

Quedaba únicamente la hipótesis de que los ángulos C y D de la figura 36.3 eran ángulos rectos. Saccheri había demostrado previamente que cuando C y D son ángulos rectos, la suma de los ángulos de cualquier triángulo es igual a 180° y que este hecho

implica el postulado de las paralelas de Euclides. Por tanto, se sintió justificado cuando concluyó su apoyo a Euclides y así publicó su *Euclides ab Omni Naevo Vindicatus* (*Euclides vindicado de todo reproche*, 1733). Sin embargo, puesto que Saccheri no obtuvo contradicción alguna sobre la base de la hipótesis del ángulo agudo, el problema del axioma de las paralelas permanecía abierto.

Los esfuerzos por encontrar un sustituto adecuado para el axioma euclídeo de las paralelas o demostrar que la aserción euclídea debía ser un teorema fueron tan numerosas y tan fútiles que en 1759 D'Alembert llamó al problema del axioma de las paralelas «el escándalo de los elementos de la geometría».

4. Presagios de las geometrías no euclídeas

En su disertación de 1763, Georg S. Klügel (1739-1812), profesor de matemáticas de la universidad de Helmstadt, quien conocía el libro de Saccheri, observó sagazmente que la certeza con la que las gentes aceptaban la verdad del axioma euclidiano de las paralelas estaba apoyada en la experiencia. Esta observación introdujo por primera vez la idea de que la experiencia, por encima de la autoevidencia, sostenía los axiomas. Klügel expresó dudas acerca de que la aserción euclidiana pudiera ser demostrada. Se dio cuenta de que Saccheri no había llegado a contradicción alguna, sino simplemente a resultados que parecían estar en oposición con la experiencia.

El ensayo de Klügel le sugirió ideas sobre el axioma de las paralelas a Lambert. En su libro *Theorie der Parallellinien* (*Teoría de las líneas*

paralelas) escrito en 1766, pero publicado en 1786,⁸⁴ Lambert, de alguna manera como Saccheri, consideró un cuadrilátero con tres ángulos rectos y consideró las posibilidades de que un cuarto ángulo fuese recto, obtuso, o agudo. Lambert descartó la hipótesis del ángulo obtuso porque lo llevó a una contradicción. Sin embargo, contrariamente a Saccheri, Lambert no concluyó que había obtenido una contradicción de la hipótesis del ángulo agudo.

Lambert dedujo de ambas hipótesis de los ángulos obtuso y agudo, respectivamente, dos consecuencias que es importante mencionar aunque la primera no le llevó a contradicción. Su resultado más sorprendente es que bajo cualquiera de las hipótesis el área de un polígono de n lados es proporcional a la diferencia entre la suma de los ángulos y $\pi n - 4$ ángulos rectos. (Saccheri poseía este resultado para un triángulo.) También notó que la hipótesis del ángulo obtuso originaba teoremas como los que se obtenían para figuras sobre la superficie de una esfera. Conjeturó que los teoremas que se seguían de la hipótesis del ángulo agudo se aplicarían a figuras sobre una esfera de radio imaginario. Esto lo condujo a escribir un ensayo⁸⁵ sobre las funciones trigonométricas de ángulos imaginarios, esto es, iA , donde A es un ángulo real e $i = \sqrt{-1}$, lo que en efecto introducía las funciones hiperbólicas (cap. 19, sec. 2). Más tarde veremos un poco más claro el significado de las observaciones de Lambert.

Los puntos de vista de Lambert sobre la geometría eran bastante avanzados. Se dio cuenta que cualquier conjunto de hipótesis que no conducía a contradicciones ofrecía una geometría posible. Tal geometría sería una estructura lógica válida a pesar de que no

tuviera relación con figuras reales. Lo último debe ser sugestivo para una geometría en particular pero no controlaba la variedad de geometrías desarrollables lógicamente. Lambert no llegó a la conclusión más radical, introducida un poco más tarde por Gauss.

Ferdinand Karl Schweikart (1780-1859), un profesor de jurisprudencia que dedicaba su tiempo libre a las matemáticas, dio otro paso hacia adelante. Trabajó sobre las geometrías no euclídeas durante el período en el que Gauss dedicó alguna atención a la materia, pero Schweikart llegó a sus conclusiones independientemente. Sin embargo, Saccheri y Lambert influenciaron sus trabajos. En un memorándum de 1816 que envió a Gauss en 1818 para su aprobación. Schweikart distinguió de hecho dos geometrías. Existía la geometría de Euclides y una geometría basada en la suposición que la suma de los ángulos de un triángulo no es igual a dos rectos. A esta última geometría la llamó geometría astral, porque podía cumplirse en el espacio de las estrellas, y sus teoremas eran los que Saccheri y Lambert habían establecido sobre la base de la hipótesis del ángulo agudo.

Franz Adolf Taurinus (1794-1874), un sobrino de Schweikart, siguió la sugerencia de su tío de estudiar la geometría astral. Aunque estableció en su *Geometriae Prima Elementa* (Geometría de Elementos Primeros, 1826) algunos nuevos resultados, sobre todo para la geometría analítica, concluyó que únicamente la geometría de Euclides podía ser verdadera para el espacio físico, pero que la geometría astral era lógicamente consistente. Taurinus demostró también que las fórmulas que eran verdaderas en esferas de radios

imaginarios son precisamente aquellas que lo eran en su geometría astral.

Los trabajos de Lambert, Schweikart y Taurinus constituyen avances que merecen una recapitulación. Los tres, así como otros autores como Klügel y Abraham G. Kastner (1719-1800), profesor de Göttingen, estaban convencidos de que el axioma de las paralelas de Euclides no podía ser demostrado, esto es, que es independiente de los otros axiomas de Euclides. Más aún, Lambert, Schweikart y Taurinus estaban convencidos de que es posible adoptar un axioma alternativo contradiciendo el de Euclides y construir una geometría lógicamente consistente. Lambert no hizo aserciones sobre la aplicabilidad de tal tipo de geometría; Taurinus pensaba que no era aplicable al mundo físico; pero Schweikart pensaba que se podía aplicar a la región de las estrellas. Estos tres autores también notaron que la geometría sobre una esfera real tiene las propiedades de la geometría basada en la hipótesis del ángulo obtuso (si se deja de lado la contradicción que resulta de la última) y que la geometría sobre la esfera de radio imaginario tiene las propiedades de una geometría basada en la hipótesis del ángulo agudo. Así, los tres reconocieron la existencia de la geometría no euclídea pero fallaron en un punto fundamental, a saber, en que la geometría euclídea no es la única geometría que describe las propiedades del espacio físico dentro de la precisión de la que la experiencia puede responder.

5. La creación de la geometría no euclídea

Ninguna de las ramas principales de la matemática, e incluso ningún resultado específico es obra de un solo hombre. A lo más, algún paso decisivo o demostración puede ser acreditada a un individuo. Este desarrollo acumulativo de la matemática se aplica especialmente a la geometría no euclídea. Si se interpreta la creación de la geometría no euclídea como el reconocimiento de que pueden existir geometrías alternativas a la de Euclides, entonces Klügel y Lambert merecen la cita. Si geometría no euclídea significa el desarrollo técnico de las consecuencias de un sistema de axiomas que contiene una alternativa al axioma de las paralelas de Euclides, entonces la mayor parte del mérito debe ser atribuido a Saccheri, y aún él se benefició del trabajo de muchos que intentaron encontrar un axioma sustitutivo más aceptable que el de Euclides. Sin embargo, el hecho más significativo acerca de la geometría no euclídea es que puede ser usada para describir las propiedades del mundo físico tan precisamente como lo hace la geometría euclídea. La última no es necesariamente la geometría del espacio físico; su verdad física no puede ser garantizada sobre una base *a priori*. Gauss logró esta comprensión por primera vez, comprensión que no requería de ningún desarrollo matemático técnico, puesto que éste ya había sido dado.

Cari Friedrich Gauss (1777-1855) fue hijo de un albañil de la ciudad alemana de Brunswick y parecía estar destinado al trabajo manual. Pero el director de la escuela en la que recibió su educación elemental estaba impresionado por la inteligencia de Gauss y lo recomendó al duque Karl Wilhelm. El duque envió a Gauss a una

escuela superior y, más adelante, en 1795, a la universidad de Göttingen. Ahí empezó Gauss a trabajar duro en sus ideas. A los 18 años inventó el método de mínimos cuadrados y a los 19 demostró que es posible construir un polígono regular de 17 lados. Estos éxitos lo convencieron de que debería cambiar de la filología a las matemáticas. En 1798 se mudó a la universidad de Helmstadt y ahí llamó la atención de Johann Friedrich Pfaff, quien se convirtió en su maestro y amigo. Después de terminar su doctorado, Gauss regresó a Brunswick donde escribió algunos de sus más famosos trabajos. Con este trabajo ganó, en 1807, el puesto de profesor de astronomía y director del observatorio de Göttingen. Residió en Göttingen por el resto de su vida, con excepción de una visita a Berlín para asistir a una reunión científica. No le gustaba enseñar y así lo afirmaba. Sin embargo, disfrutaba de la vida social, se casó dos veces y creó una familia.

El primer trabajo importante de Gauss fue su tesis doctoral, en la que demostró el teorema fundamental del álgebra. En 1801 publicó su clásico *Disquisitiones Arithmeticae*. Su trabajo matemático en geometría diferencial, las «*Disquisitiones generales circa superficies curvas*» («Investigaciones generales sobre superficies curvas», 1827), que, incidentalmente, fue el resultado de su interés en agrimensura, geodesia y cartografía, es una joya matemática (cap. 37, sec. 2). Realizó muchas otras contribuciones al álgebra, funciones complejas y teoría del potencial. Mantuvo inédito su trabajo innovativo en dos campos principales: las funciones elípticas y la geometría no euclídea.

Sus intereses en física fueron igualmente muy amplios y les dedicó gran parte de sus energías. Cuando Giuseppe Piazzi (1746-1826) descubrió el planeta Ceres en 1801, Gauss se dedicó a determinar su trayectoria. Este fue el inicio de su trabajo en astronomía, actividad que lo absorbió mucho y a la que dedicó cerca de veinte años. Una de sus grandes contribuciones en esta área es su *Theoria Motus Corporum Coelestium (Teoría del movimiento de los cuerpos celestes, 1809)*. También se hizo merecedor de alta consideración por sus investigaciones físicas sobre magnetismo teórico y experimental. Maxwell comenta en su *Electricity and Magnetism (Electricidad y Magnetismo)* que los estudios de magnetismo de Gauss reconstruyeron la ciencia por completo, en los instrumentos usados, los métodos de observación y el cálculo de los resultados. Los ensayos de Gauss sobre magnetismo terrestre son modelos de investigación física y proporcionaron el mejor método para medir el campo magnético de la tierra. Su trabajo en astronomía y magnetismo abrió un nuevo y brillante período de alianza entre las matemáticas y la física.

A pesar de que Gauss y Wilhelm Weber (1804-1891) no inventaron la idea de la telegrafía, en 1833 mejoraron técnicas anteriores con un aparato práctico que hacía rotar una aguja a derecha o izquierda dependiendo de la dirección de la corriente enviada a través del alambre. Gauss también trabajó en óptica, que había sido olvidada desde los tiempos de Euler, y sus investigaciones de 1838-1841 proporcionaron una base completamente nueva para el estudio de los problemas ópticos.

La universalidad de las actividades de Gauss es de lo más sorprendente, ya que sus contemporáneos habían empezado a encerrarse en investigaciones especializadas. A pesar del hecho de ser reconocido como el más grande de los matemáticos, al menos desde Newton, Gauss no fue tanto un gran innovador sino más bien una figura de transición entre los siglos XVIII y XIX. A pesar de llegar a algunos nuevos puntos de vista que influenciaron a otros matemáticos, estaba orientado más hacia el pasado que hacia el futuro. Félix Klein describe la posición de Gauss con estas palabras: nosotros podemos tener un cuadro del desarrollo de las matemáticas si nos imagináramos una cadena de altas montañas representando a los autores del siglo XVIII coronados por una cima imponente —Gauss— seguida de una larga y rica región llena de nuevos elementos de vida. Los contemporáneos de Gauss reconocieron su genio y, para su muerte, en 1855, era ampliamente venerado y llamado «el príncipe de los matemáticos».

Gauss publicó relativamente poco de su trabajo, ya que pulía cualquier cosa que hacía, en parte para lograr elegancia y en parte para lograr en sus demostraciones el máximo de brevedad sin sacrificar el rigor, al menos el rigor de su tiempo. En el caso de la geometría no euclídea no publicó ningún trabajo definitivo. En una carta a Bessel de 27 de enero de 1829, dijo que probablemente nunca publicaría sus hallazgos en esta materia porque temía al ridículo o, como dijo, temía al clamor de los beodos, una referencia figurativa a una tribu griega estúpida. Gauss pudo haber sido demasiado cuidadoso, pero se recordará que, a pesar de que

algunos matemáticos habían llegado gradualmente al clímax de la geometría no euclídea, el mundo intelectual, en general, seguía dominado por las enseñanzas de Kant. Lo que conocemos del trabajo de Gauss en geometría no euclídea está tomado de las cartas a sus amigos, dos breves reseñas en el *Göttingische Gelehrte Anzeigen* de 1816 y 1822 y algunas notas de 1831 encontradas entre sus papeles después de su muerte.⁸⁶

Gauss era completamente consciente de los vanos esfuerzos por establecer el postulado de las paralelas de Euclides, ya que esto era conocimiento común en Göttingen; y Kastner, el maestro de Gauss, estaba completamente familiarizado con la historia de estos esfuerzos. Gauss le comentó a su amigo Schumacher que, ya desde 1792 (por aquel entonces Gauss contaba quince años de edad) pensaba que podía existir una geometría lógica en la que el postulado de las paralelas de Euclides no se cumpliera. En 1794 Gauss había encontrado que, en su concepto de geometría no euclídea, el área de un cuadrángulo debía ser proporcional a la diferencia entre 360° y la suma de los ángulos. Sin embargo, en esta fecha posterior y aún en 1799, Gauss insistía en deducir el postulado de las paralelas de Euclides a partir de otras suposiciones más plausibles y aún creía que la geometría euclídea era la geometría del espacio físico, y ello a pesar de que podía concebir otras geometrías lógicas no euclídeas. Sin embargo, el 17 de diciembre de 1799, Gauss escribió a su amigo el matemático húngaro Wolfgang Farkas Bolyai (1775-1856).

En cuanto a mí, he logrado cierto progreso en mi trabajo. Sin embargo, el camino que he escogido no lleva del todo a la meta que buscamos [la deducción del axioma de las paralelas], y el cual me aseguras que has logrado. Me parece mejor dudar de la verdad de la propia geometría. Es cierto que he logrado lo que la gente sostendría que constituye una demostración; pero a mis ojos no prueba nada. Por ejemplo, si pudiéramos demostrar que un triángulo rectilíneo cuya área pudiera ser más grande que cualquier área es posible, entonces estaría listo para demostrar la totalidad de la geometría [euclídea] de manera absolutamente rigurosa.

La mayoría de la gente ciertamente lo dejaría como un axioma, ¡pero yo no! Sería, por supuesto, posible que el área permanezca siempre menor que un cierto límite, por lejos que los tres puntos angulares del triángulo fueran tomados.

Este pasaje muestra que, en 1799, Gauss estaba hasta cierto punto convencido de que el axioma de las paralelas no podía ser deducido de los restantes axiomas euclídeos y comenzó a tomar más en serio el desarrollo de una nueva y posiblemente aplicable geometría.

A partir de 1813, Gauss desarrolló su nueva geometría, que primero llamó geometría antieuclídea, más adelante geometría astral y finalmente geometría no euclídea. Se convenció de que era lógicamente consistente y estaba bastante seguro de que podía ser aplicable. En reseñas de 1816 y 1822 y en una carta a Bessel de 1829, Gauss reafirmó que el postulado de las paralelas no podía ser demostrado sobre la base de los otros axiomas de Euclides. Su carta a Olbers escrita en 1817⁸⁷ es un momento culminante. En ésta,

Gauss afirma: «Me estoy convenciendo más y más de que la necesidad física de nuestra geometría euclídea no puede ser demostrada, al menos no por la razón humana ni para la razón humana. Tal vez en otra vida seamos capaces de obtener una visión de la naturaleza del espacio, que es ahora inalcanzable. Hasta entonces no debemos situar la geometría en la misma clase que la aritmética, que es puramente *a priori*, sino con la mecánica.»

Para poner a prueba la aplicabilidad de la geometría euclídea y su geometría no euclídea, Gauss midió de hecho la suma de los ángulos de un triángulo formado por los picos de tres montañas: Brocken, Hohehagen y Inselsberg. Los lados de los triángulos fueron 69, 85 y 197 km. Encontró ⁸⁸ que la suma excedía 180° en $14''85$. El experimento no demostró nada, ya que el error experimental era mucho más grande que el exceso y así la suma correcta podía haber sido 180° o aún menos. Como se dio cuenta Gauss, el triángulo era pequeño y como en la geometría no euclídea el defecto es proporcional al área, únicamente un triángulo grande podía posiblemente revelar alguna diferencia significativa de la suma de ángulos de 180° .

No discutiremos los teoremas particulares de la geometría no euclídea que se deben a Gauss. No escribió una presentación deductiva completa y los teoremas que demostró son muy parecidos a los que encontraremos en los trabajos de Lobachewsky y Bolyai. Estos dos son considerados generalmente autores de la creación de la geometría no euclídea. Lo que se les tiene que atribuir será discutido posteriormente, pero ellos publicaron presentaciones

organizadas de una geometría no euclídea sobre una base sintética deductiva con el entendimiento completo de que esta nueva geometría era lógicamente tan legítima como la de Euclides.

Nikolai Ivanovich Lobachewsky (1793-1856), ruso, estudió en la universidad de Kazan y de 1827 a 1846 fue profesor y rector de esa universidad. Presentó sus puntos de vista sobre los fundamentos de la geometría en un ensayo ante el departamento de matemáticas y física de la universidad en 1826. Sin embargo, el ensayo no fue publicado jamás y se perdió. Proporcionó su enfoque de la geometría no euclídea en una serie de ensayos, los dos primeros de los cuales fueron publicados en revistas en Kazan y el tercero en el *Journal für Mathematik*.ⁿ El primero fue titulado «Sobre los fundamentos de la geometría» y apareció en 1829-1830⁸⁹. El segundo, titulado «Nuevos fundamentos de la geometría con una teoría completa de las paralelas» (1835-1837), es una presentación mejor de las ideas de Lobachewsky. Llamó a su nueva geometría, geometría imaginaria por razones que son tal vez ya visibles y que serán claras posteriormente. En 1840, publicó un libro en alemán *Geometrische Untersuchungen zur Theorie der Paraliellinien (Investigaciones geométricas sobre la teoría de las paralelas)*⁹⁰. En este libro lamenta el poco interés mostrado por sus escritos. A pesar de que perdió la vista, dictó una exposición completamente nueva de su geometría y la publicó en 1855 bajo el título de *Pangeometrie (Pangeometría)*.

Janos Bolyai (1802-1860), hijo de Wolfgang Bolyai, fue oficial del ejército húngaro. Escribió un ensayo de 26 páginas «La Ciencia del espacio absoluto»⁹¹ sobre geometría no euclídea, que él llamó

geometría absoluta. Fue publicado como un apéndice al trabajo de su padre *Tentamen Juventutem Studiosam in Elementa Matheseos* (Ensayo sobre los elementos de las matemáticas para jóvenes estudiosos). A pesar de que el libro (de dos volúmenes) apareció en 1832-1833, y, por tanto, después de una publicación de Lobachewsky, Bolyai desarrolló aparentemente sus ideas sobre geometría no euclídea alrededor de 1825 y estaba convencido en esa época de que la nueva geometría no era contradictoria. En una carta a su padre, fechada el 23 de noviembre de 1823, Janos escribió: «He hecho tan maravillosos descubrimientos que yo mismo me he perdido en el deslumbramiento.» El trabajo de Bolyai era tan parecido al de Lobachewsky que cuando Bolyai vio por primera vez el trabajo de este último en 1835, pensó que había sido copiado de su propia publicación de 1832-1833. Por otro lado, Gauss leyó el ensayo de Janos Bolyai en 1832 y escribió a Wolfgang⁹² que era incapaz de elogiarlo porque sería como elogiar su propio trabajo.

6. El contenido técnico de la geometría no euclídea

Gauss, Lobachewsky y Bolyai se habían percatado de que el axioma euclídeo de la paralelas no podía ser demostrado tomando como base los otros nueve axiomas y que algún axioma adicional era necesario para fundamentar la geometría euclídea. Ya que el axioma de las paralelas era un hecho independiente, resultaba al menos lógicamente posible adoptar un enunciado contradictorio y desarrollar las consecuencias de ese nuevo conjunto de axiomas.

Para estudiar el contenido técnico de lo que estos hombres crearon, sólo es necesario tomar el trabajo de Lobachewsky, ya que los tres hicieron casi la misma cosa. Lobachewsky proporcionó, como sabemos, diversas versiones que difieren únicamente en pequeños detalles. Nuestro resumen estará basado en su ensayo de 1835-1837. Puesto que en los *Elementos* de Euclides se demuestran muchos teoremas que no dependen en absoluto del axioma de las paralelas, tales teoremas son válidos en la nueva geometría. Lobachewsky dedica los primeros seis capítulos de su ensayo a la demostración de estos teoremas básicos. Desde un principio supone que el espacio es infinito, y así es capaz de demostrar que dos rectas no pueden intersectarse en más de un punto y que dos perpendiculares a la misma recta no se cortan.

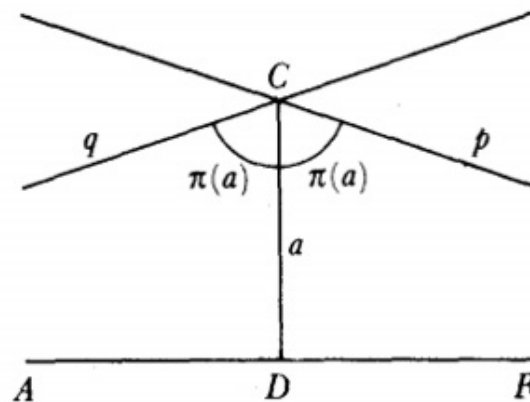


Figura 36.5

En el séptimo capítulo, Lobachewsky rechaza audazmente el postulado euclídeo de las paralelas y hace la siguiente suposición: dada una recta AB y un punto C (fig. 36.5), entonces todas las rectas que pasan por C caen dentro de dos clases con respecto a

AB , a saber, la clase de las rectas que cortan AB y la clase de las que no lo hacen. A la última pertenecen las dos rectas p y q que forman la frontera entre las dos clases. Estas dos líneas frontera son llamadas las rectas paralelas. Más precisamente, si C es un punto a una distancia perpendicular a de la recta AB , entonces existe un ángulo⁹³ $\pi(a)$ tal que todas las rectas que pasan por C y que forman con la perpendicular CD un ángulo menor que $\pi(a)$ intersecarán AB ; todas las otras rectas que pasan por C no intersecan AB .⁹⁴ Las dos rectas que forman el ángulo $\pi(a)$ con AB son las paralelas y $\pi(a)$ es llamado el ángulo de paralelismo. Las otras rectas que no son paralelas y que pasan por C y las que no cortan a AB son llamadas rectas que no intersecan, aunque en el sentido de Euclides éstas son paralelas a AB y así en este sentido la geometría de Lobachewsky contiene un número infinito de paralelas que pasan por C .

Si $\pi(a) = \pi/2$ entonces resulta el axioma euclídeo. Si no, entonces se sigue que $\pi(a)$ se incrementa y se aproxima a $\pi/2$ cuando a decrece a cero y $\pi(a)$ decrece y se aproxima a cero mientras a se hace infinito. La suma de los ángulos de un triángulo siempre es menor que π , decrece cuando se incrementa el área del triángulo y se aproxima a π cuando el área se aproxima a cero. Si los dos triángulos son semejantes, entonces son congruentes.

Ahora, Lobachewsky se vuelve hacia la parte trigonométrica de su geometría. El primer paso es la determinación de $\pi(a)$. El resultado, si un ángulo central completo es 2π , es⁹⁵

$$\tan \frac{\pi(x)}{2} = e^{-x} \quad (1)$$

de lo que se sigue que $\pi(0) = \pi/2$ y $\pi(\infty) = 0$. La relación (1) es significativa en cuanto que a cada longitud x asocia un ángulo definido. Cuando $x = 1$, $\tan [\pi(1)/2] = e^{-1}$ de tal forma que $\pi(1) = 40^\circ 24'$. Así la unidad de longitud es aquella longitud cuyo ángulo de paralelismo es $40^\circ 24'$. Esta unidad de longitud no tiene un significado físico directo. Físicamente puede ser una pulgada o una milla. Uno seleccionaría una interpretación física que haría la geometría físicamente aplicable.⁹⁶

Después, Lobachewsky deduce fórmulas conectando lados y ángulos de los triángulos planos de su geometría. En un ensayo de 1834, había definido el $\cos x$ y $\sin x$ para x real como las partes real e imaginaria de e^{ix} . La meta de Lobachewsky consistía en proporcionar un fundamento puramente analítico para la trigonometría y hacerla independiente de la geometría euclídea. Las principales fórmulas trigonométricas de su geometría son (fig. 36.6).

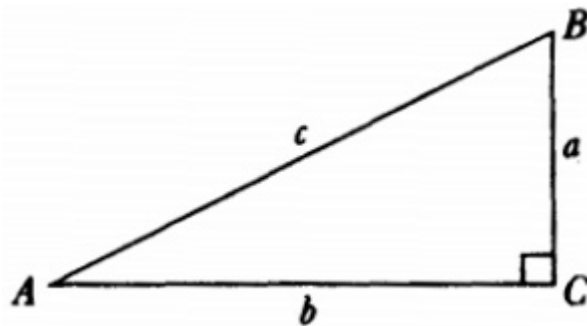


Figura 36.6

$$\cos \pi(a) = \cos \pi(c) \operatorname{sen} A$$

$$\operatorname{sen} A = \cos B \operatorname{sen} \pi(b)$$

$$\operatorname{sen} \pi(c) = \operatorname{sen} \pi(a) \operatorname{sen} \pi(b).$$

Estas fórmulas son válidas en la trigonometría esférica ordinaria siempre que los lados tengan longitudes imaginarias. Esto es, si se reemplaza a , b y c en las fórmulas usuales de la trigonometría esférica por ia , ib e ic se obtienen las fórmulas de Lobachewsky. Ya que las funciones trigonométricas de ángulos imaginarios son reemplazables por funciones hiperbólicas, se espera ver estas últimas funciones en las fórmulas de Lobachewsky. Estas pueden ser obtenidas al usar la relación

$$\tan (\pi(x)/2) = e^{-x/k}$$

Así la primera de las fórmulas anteriores puede ser convertida en

$$\operatorname{senh} \frac{a}{k} = \operatorname{senh} \frac{c}{k} \operatorname{sen} A$$

También, mientras que en la geometría esférica usual el área de un triángulo con los ángulos A , B y C es $r^2 (A + B + C - \pi)$, en la geometría no euclídea ésta es $r^2[\pi - (A + B + C)]$, lo que equivale a reemplazar r por ir en la fórmula usual.

En su primer artículo de (1829-1830), Lobachewsky derivó, trabajando con un triángulo infinitesimal, la fórmula

$$ds = \sqrt{(dy)^2 + \frac{(dx)^2}{\text{sen}^2 \pi(x)}}$$

para el elemento de arco sobre la curva $y = f(x)$ en el punto (x,y) . Así, puede ser calculada la circunferencia entera de un círculo de radio r , y es

$$C = \pi(e^r - e^{-r})$$

La expresión para el área de un círculo resulta ser

$$A = \pi(e^{r/2} - e^{-r/2})^2.$$

También proporciona teoremas sobre el área de regiones planas y curvas y volúmenes de sólidos.

Las fórmulas de la geometría euclídea resultan de las fórmulas no euclídeas cuando las magnitudes son pequeñas. Así, si usamos el hecho que

$$e^r = 1 + r + r^2/2! + \dots$$

y despreciamos para r pequeña todos los términos, excepto los dos primeros, por ejemplo,

$$C = n(e^r - e^{-r}) \sim \pi\{1 + r - (1 - r)\} = 2\pi r.$$

En el primer trabajo (1829-1830), Lobachewsky también consideró la aplicabilidad de su geometría al espacio físico. La esencia de su argumento se apoya en el paralaje de las estrellas. Supóngase que E_1 y E_2 (fig. 36.7) son las posiciones de la tierra con seis meses de diferencia y S es una estrella.

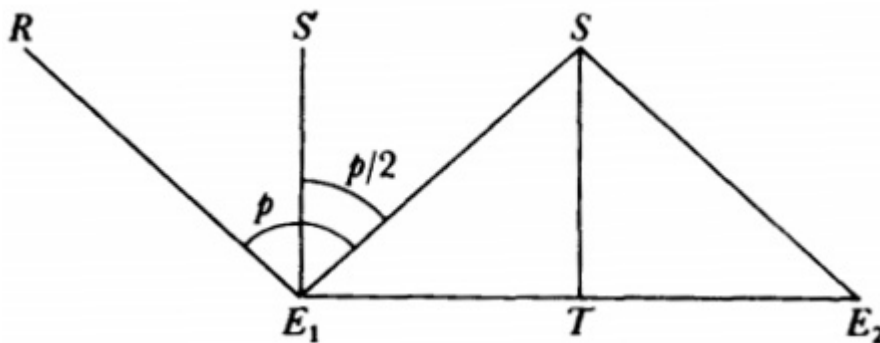


Figura 36.7

El paralaje p de S es la diferencia en las direcciones de E_1S y E_2S medida, digamos, a partir de la perpendicular E_1S' y E_2S . Si E_1R es la paralela euclídea a E_2S , entonces, ya que E_1SE_2 es un triángulo isósceles, $\pi/2 - \angle SE_1E_2$ es la mitad del cambio en la dirección de la estrella, esto es, $p/2$. Este ángulo es $1''24$ para la estrella Sirio (el valor de Lobachewsky). En tanto ese ángulo no sea cero, la recta de E_1 a la estrella no puede ser paralela a TS ya que la línea corta a TS . Sin embargo, si existiera una cota inferior para los diversos paralajes de todas las estrellas, entonces la recta desde E_1 que

forma un ángulo con E_1S' menor que esta cota inferior podía ser tomada como una paralela a TS por E_1 y esta geometría sería igualmente útil en lo que concierne a las medidas estelares. Pero, más adelante, Lobachewsky demostró que la unidad de longitud en su geometría tendría que ser, físicamente, más de un millón de veces el radio de la trayectoria de la tierra. En otras palabras, la geometría de Lobachewsky podía ser aplicable únicamente en un triángulo enormemente grande.

7. Las demandas de prioridad de Lobachewsky y Bolyai

La creación de la geometría no euclídea es comúnmente usada como ejemplo de cómo una idea se les ocurre independientemente a diferentes personas casi al mismo tiempo. Algunas veces esto es visto como coincidencia pura, y algunas como prueba del espíritu del tiempo señalando una influencia en frentes ampliamente separados. La creación de la geometría no euclídea por Gauss, Lobachewsky y Bolyai no es un ejemplo de creación simultánea, ni está bien justificado otorgarles su gran mérito a Lobachewsky y Bolyai. Es cierto, como se mencionó anteriormente, que fueron los primeros en publicar abiertamente una geometría no euclídea y su acto mostró más coraje que el de Gauss. Sin embargo, la creación de la geometría no euclídea es difícilmente una contribución suya. Ya hemos señalado que aún Gauss fue precedido por Lambert, que Schweikart y Taurinus fueron creadores independientes, y que Schweikart y Taurinus publicaron su trabajo. Más aún, a Gauss se

le debe el percatarse que la nueva geometría era aplicable al espacio físico.

Ambos, Lobachewsky y Bolyai, le deben mucho a Gauss. Johann Martin Bartels (1769-1836), un buen amigo de Gauss, fue el maestro de Lobachewsky en Kazan. De hecho, estuvieron juntos en Brunswick los años de 1805 a 1807. Posteriormente, Gauss y Bartels se mantuvieron en comunicación el uno con el otro. Es extremadamente difícil que Bartels no haya comunicado el progreso de la geometría no euclídea a Lobachewsky, quien permaneció en Kazan como colega. En particular, Bartels debió haber conocido las dudas de Gauss en cuanto a la verdad de la geometría euclídea.

En cuanto a Janos Bolyai, su padre fue un amigo cercano de Gauss y compañero mientras estudiaban en Göttingen de 1796 a 1798. Wolfgang y Gauss no solamente continuaron en comunicación el uno con el otro, sino que discutieron específicamente el tema del axioma de las paralelas, como lo indica una de las citas anteriores. Wolfgang continuó trabajando duro sobre el problema del axioma de las paralelas y envió una pretendida demostración en 1804. Gauss le mostró que la prueba era falaz. Para 1817 Gauss tenía la certeza, no sólo de que el axioma no podía ser demostrado, sino de que una geometría no euclídea lógicamente consistente y físicamente aplicable podía ser construida. Además de su comunicación de 1799 a este efecto, Gauss transmitió sus últimos pensamientos libremente a Wolfgang. Wolfgang continuó trabajando sobre el problema hasta que publicó su *Tentamen* de 1832-1833. Puesto que

le recomendó a su hijo que estudiara el problema del axioma de las paralelas, con certeza debió haberle detallado todo lo que sabía.

Hay puntos de vista contradictorios. El matemático Friedrich Engel (1861-1941) pensaba que aunque Bartels, el maestro de Lobachewsky, era amigo de Gauss, Lobachewsky difícilmente podría haber sabido a través de su relación que Gauss había dudado de la corrección física del axioma de las paralelas. Pero este hecho, en sí mismo, fue crucial. Sin embargo, Engel dudó que Lobachewsky hubiera sabido aún esto de Gauss, ya que Lobachewsky había tratado desde 1816 de demostrar el axioma euclídeo de las paralelas; así, reconociendo la inutilidad de tales esfuerzos, finalmente creó en 1826 la nueva geometría. Janos Bolyai también intentó demostrar el axioma euclídeo de las paralelas hasta alrededor de 1820 y entonces se dirigió hacia la construcción de la nueva geometría. Pero estos esfuerzos continuos por demostrar el axioma de las paralelas no implican la ignorancia del pensamiento de Gauss.⁹⁷ Ya que nadie, ni aún Gauss, había demostrado que el axioma de Euclides de las paralelas no podía ser deducido a partir de los otros nueve axiomas, tanto Lobachewsky como Bolyai pudieron haber decidido intentar resolver el problema. Después de fallar, apreciaron más rápidamente la sabiduría de los puntos de vista de Gauss en la materia.

En cuanto al contenido técnico contribuido por Lobachewsky y Janos Bolyai, a pesar de que pudieron haber creado esto independientemente de sus predecesores y entre ellos mismos, los trabajos de Saccheri y Lambert, por no decir nada de los de

Schweikart y Taurinus, eran ampliamente conocidos en Göttingen y les eran ciertamente conocidos a Bartels y Wolfgang Bolyai. Y Lobachewsky admite por inferencia su conocimiento de los trabajos anteriores cuando se refiere en su ensayo de 1835-1837 a la futilidad de los esfuerzos durante más de dos mil años por resolver la cuestión del axioma de las paralelas.

8. Las implicaciones de la geometría no euclídea

Ya hemos señalado que la creación de la geometría no euclídea fue el paso de mayores consecuencias y más revolucionario desde los tiempos de los griegos. No trataremos todas las implicaciones de la materia en este momento. En su lugar, seguiremos el curso histórico de estos hechos. El impacto de la creación y la completa realización de su significación se postergó, ya que Gauss no publicó su trabajo sobre la materia y los trabajos de Lobachewsky y Bolyai fueron ignorados durante aproximadamente treinta años. A pesar de que éstos eran conscientes de la importancia de su trabajo, los matemáticos exhibieron generalmente su reluctancia a abrigar ideas radicales. También, la materia clave en la geometría de los 1830 y 1840 fue la geometría proyectiva y por esta razón también el trabajo sobre la geometría no euclídea no atrajo a los matemáticos ingleses, franceses y alemanes. El tema atrajo la atención gracias a que el nombre de Gauss proporcionó peso a las ideas cuando las notas de Gauss y su correspondencia sobre geometría no euclídea fueron publicadas después de su muerte en 1855. Richard Baltzer (1818-1887), en un libro de 1866-1867, citó los trabajos de Lobachewsky y

Bolyai. Finalmente, desarrollos posteriores atrajeron a los matemáticos a la comprensión de la completa importancia de la geometría no euclídea.

Gauss sí vio las implicaciones más revolucionarias. El primer paso en la creación de la geometría no euclídea consistió en darse cuenta de que el axioma de las paralelas no podía ser demostrado sobre la base de los otros nueve axiomas. Era una aserción independiente, y así era posible adoptar un axioma contradictorio y desarrollar una geometría enteramente nueva. Esto lo hicieron Gauss y otros. Pero Gauss, habiéndose percatado de que la geometría euclídea no es necesariamente la geometría del espacio físico, esto es, no es necesariamente cierta, puso a la geometría en la misma clase que la mecánica y aseguraba que la cualidad de su certeza debía ser restringida a la de la aritmética (y su desarrollo en el análisis). Esta confianza en la aritmética es en sí misma curiosa. La aritmética, en esta época, no poseía un mínimo fundamento lógico. La seguridad de que la aritmética, el álgebra y el análisis ofrecían verdades acerca del mundo real provenía enteramente de su apoyo en la experiencia. La historia de la geometría no euclídea revela, de una manera sorprendente, que los matemáticos son influenciados, no tanto por el razonamiento que ellos producen, sino más bien por el espíritu de los tiempos. Saccheri había rechazado los teoremas extraños de la geometría no euclídea y concluido que Euclides había sido reivindicado. Pero cien años más tarde, Gauss, Lobachewsky y Bolyai aceptaron confiadamente la nueva geometría. Creían que su nueva geometría era lógicamente consistente y de aquí que la

geometría fuera tan válida como la de Euclides. Pero carecían de una prueba de su consistencia. A pesar de que demostraron muchos teoremas y no obtuvieron contradicciones evidentes, permanecía abierta la posibilidad de que se pudiera derivar una contradicción. De suceder esto, entonces la suposición de su axioma de las paralelas sería invalidada, como Saccheri había creído, y el axioma de las paralelas de Euclides sería una consecuencia de sus otros axiomas.

De hecho, Bolyai y Lobachewsky consideraron la cuestión de la consistencia y estaban convencidos parcialmente de ella ya que su trigonometría era la misma para una esfera de radio imaginario y la esfera es parte de la geometría euclídea. Pero Bolyai no estaba satisfecho con esta evidencia ya que la trigonometría no es en sí misma un sistema matemático completo. Así, a pesar de la ausencia de cualquier demostración de consistencia, o de la aplicabilidad de la nueva geometría, que al menos habría servido como un elemento convincente, Gauss, Bolyai y Lobachewsky aceptaron lo que sus predecesores habían considerado como absurdo. Esta aceptación fue un acto de fe. La cuestión de la consistencia de la geometría no euclídea permaneció como problema abierto otros cuarenta años.

Un punto más requiere atención y énfasis acerca de la creación de la geometría no euclídea. Existe una creencia popular de que Gauss, Bolyai y Lobachewsky se encerraron en una esquina, jugaron a cambiar los axiomas de la geometría euclídea únicamente para satisfacer su curiosidad intelectual y así crearon una nueva geometría. Y ya que esta creación ha mostrado ser enormemente

importante para la ciencia —una forma de geometría no euclídea que tenemos todavía que examinar ha sido usada en la teoría de la relatividad— muchos matemáticos han asentido a que la curiosidad puramente intelectual es una justificación suficiente para la exploración de cualquier idea matemática y que los valores de la ciencia seguirán casi seguramente como ha sucedido, a propósito, en el caso de la geometría no euclídea. Pero la historia de la geometría no euclídea no apoya esta tesis. Hemos visto que la geometría no euclídea llegó después de siglos de trabajo sobre el axioma de las paralelas. La preocupación por este axioma surgió del hecho de que debería ser, como axioma, una verdad autoevidente. Ya que los axiomas de la geometría son nuestros hechos básicos acerca del espacio físico y grandes ramas de las matemáticas y de las ciencias físicas usan las propiedades de la geometría euclídea, los matemáticos desearon estar seguros de que estaban apoyándose sobre verdades. En otras palabras, el problema del axioma de las paralelas no era únicamente un problema físico genuino, sino tan fundamental como puede serlo un problema físico.

Bibliografía

- Bonola, Roberto: *Non-Euclidean Geometría*, Dover (reimpresión), 1955.
- Dunnington, G. W.: *Cari Friedrich Gauss*, Stechert-Hafner, 1960.
- Engel, F., y Staeckel, P.: *Die Theorie der Parallellinien von Euklid bis auf Gauss*, 2 vols., B. G. Teubner, 1895.; —

Urkunden zur Geschichte der nichteuklidischen Geometrie, B. G. Teubner, 1899-1913, 2 vols. El primer volumen contiene la traducción del ruso al alemán de los ensayos de Lobachewsky de 1829-1830 y 1835-1837. El segundo es sobre el trabajo de los dos Bolyai.

- Enriques, F.: «Prinzipien der Geometrie». *Encyk. der Math. Wiss.*, B. G. Teubner, 1907-1910, III ABI, 1-129.
- Gauss, Cari F.: *Werke*. B. G. Teubner, 1900 y 1903, vol. 8, 157-268; vol. 9, 297-458.
- Heath, Thomas L.: *Euclid's Elements*, Dover (reimpresión), 1956, vol. 1, 202-220.
- Kagan, V.: *Lobachewsky and his contribution to Science*, Foreign Language Pub. House, Moscú, 1957. Hay versión castellana, *Lobachewki*, Moscú, Ed. Mir, 1986.
- Lambert, J. H.: *Opera Mathematica*, 2 vols. Orell Fussli, 1946-1948.
- Pasch, Moritz, y Max Dehn: *Vorlesungen uber neuere Geometrie*, 2.¹ ed., Julius Springer, 1926, 185-238. Hay versión castellana, M. Pasch, *Lecciones de geometría moderna*, Madrid, Junta para Ampliación de Estudios, 1912.
- Saccheri, Gerolamo: *Euclides ab Omni Naevo Vindicatus*, Trad. inglesa por G. B. Halsted en *Amer. Math. Monthly*, vols. 1-5, 1894-1898; también en Open Court Pub. Co., 1920 y Chelsea (reimpresión), 1970.
- Schmidt, Franz, y Paul Staeckel: *Briefwechsel zwischen Cari Friedrich Gauss und Wolfgang Bolyai*, B. G. Teubner, 1899;

Georg Olms (reimpresión), 1970.

- Smith, Davis F.: *A source book in mathematics*, Dover (reimpresión), 1959, vol. 2, 351-388.
- Sommerville, D. M. Y.: *The elements of non-euclidean geometry*, Dover (reimpresión), 1958.
- Staeckel, P.: «Gauss als Geometer», *Nachrichten Königl. Ges. der Wiss. zu Gött.*, 1917, Beiheft, 25-142. También en Gauss: *Werke*, X₂. Von Walterhausen, W. Sartorius: *Cari Friedrich Gauss*, S. Hirzel, 1856; Springer Verlag (reimpresión), 1965.
- Zacharias, M.: «Elementargeometrie und elementare nicht-euklidische Geometrie in synthetischer Behandlung». *Encyk. der Math. Wiss.*, B. G. Teubner, 1914-1931, III AB9, 859-1172.

Capítulo 37

La geometría diferencial de Gauss y Riemann

*Tú, naturaleza, sé mi diosa; a tus
leyes mis servicios están atados...*

Carl F. Gauss

Contenido:

1. *Introducción*
 2. *La geometría diferencial de Gauss*
 3. *El enfoque de Riemann de la geometría*
 4. *Los sucesores de Riemann*
 5. *Los invariantes de las formas diferenciales*
- Bibliografía*

1. Introducción

Debemos ahora recoger los hilos del desarrollo de la geometría diferencial, particularmente la teoría de superficies tal como fue fundada por Euler y extendida por Monge. Gauss dio el gran paso siguiente en la disciplina.

Empezando en 1816, Gauss dedicó una gran cantidad de trabajo a la geodesia y la cartografía. Su participación en mediciones físicas reales, sobre las que publicó muchísimos ensayos, estimuló su interés en la geometría diferencial y lo condujo a su ensayo definitivo de 1827 «*Disquisitiones Generales circa Superficies curvas*» («Investigaciones generales sobre superficies curvas»).⁹⁸ Sin

embargo, Gauss propuso el concepto completamente nuevo de que una superficie es un espacio en ella misma, además de contribuir al tratamiento definitivo de la geometría diferencial de superficies situadas en un espacio tridimensional. Fue este concepto el que generalizó Riemann, abriendo desde entonces nuevas perspectivas en la geometría no euclídea.

2. La geometría diferencial de Gauss

Euler ya había introducido la idea (cap. 23, sec. 7) de que las coordenadas (x, y, z) de cualquier punto sobre una superficie pueden ser representadas en términos de dos parámetros u y v ; esto es, las ecuaciones de la superficie están dadas por

$$x = x(u, v), \quad y = y(u, v), \quad z = z(u, v) \quad (1)$$

Gauss partió de usar la representación paramétrica para el estudio sistemático de las superficies. A partir de estas ecuaciones paramétricas tenemos

$$dx = a du + a' dv, \quad dy = b du + b' dv, \quad dz = c du + c' dv \quad (2)$$

donde $a = x_u$, $a' = x_v$, y así en adelante. Por conveniencia, Gauss introduce los determinantes

$$A = \begin{vmatrix} b & c \\ b' & c' \end{vmatrix} \quad B = \begin{vmatrix} c & a \\ c' & a' \end{vmatrix} \quad C = \begin{vmatrix} a & b \\ a' & b' \end{vmatrix}$$

y la cantidad

$$\Delta = \sqrt{A^2 + B^2 + C^2}$$

que él supone no es idénticamente igual a cero.

La cantidad fundamental sobre cualquier superficie es el elemento de longitud de arco, que en coordenadas (x, y, z) es

$$ds^2 = dx^2 + dy^2 + dz^2 \quad (3)$$

Gauss usa las ecuaciones (2) para escribir (3) como

$$ds^2 = E(u, v)du^2 + 2F(u, v)du dv + G(u, v)dv^2 \quad (4)$$

donde

$$E = a^2 + b^2 + c^2, \quad F = aa' + bb' + cc', \quad G = a'^2 + b'^2 + c'^2.$$

Otra cantidad fundamental es el ángulo entre dos curvas sobre una superficie. Una curva sobre una superficie está determinada por una cierta relación entre u y v , ya que entonces x , y y z se convierten en funciones de un parámetro, u o v , y las ecuaciones (1) se convierten en la representación paramétrica de una curva. Se dice en el lenguaje de las diferenciales que en un punto (u, v) una curva o la dirección de una curva emanando de un punto está dada por la proporción $du:dv$. Si entonces tenemos dos curvas o dos

direcciones partiendo de (u,v) , una dada por $du:dv$ y la otra por $du':dv'$, y si θ es el ángulo entre estas direcciones, Gauss muestra que

$$\cos \theta = \frac{E du du' + F(du dv' + du' dv) + G dv dv'}{\sqrt{E du^2 + 2F du dv + G dv^2} \sqrt{E du'^2 + 2F du' dv' + G dv'^2}} \quad (5)$$

Gauss se dedica luego al estudio de la curvatura de una superficie. Su definición de curvatura es una generalización a las superficies de la indicatriz usada para las curvas en el espacio por Euler, y usada para las superficies por Olinde Rodrigues.⁹⁹ En cada punto (x,y,z) sobre una superficie hay una normal cuya dirección está fijada. Gauss considera una esfera unitaria y selecciona un radio teniendo la dirección de la normal dirigida sobre la superficie. La elección del radio determina un punto (X,Y,Z) sobre la esfera. Si consideramos sobre la superficie cualquier región pequeña rodeando a (x,y,z) , entonces existe una región correspondiente sobre la esfera rodeando a (X,Y,Z) . La curvatura de la superficie en (x,y,z) está definida como el límite del radio del área de la región sobre la esfera al área de la región correspondiente sobre la superficie cuando las dos áreas tienden a sus puntos respectivos. Gauss evalúa esta proporción notando primero que el plano tangente en (X,Y,Z) sobre la esfera es paralelo al correspondiente a (x,y,z) sobre la superficie. De aquí que la proporción de las dos áreas sea la proporción de sus proyecciones sobre los planos tangentes respectivos. Para encontrar esta última proporción, Gauss realiza un número increíble de diferenciaciones y

obtiene un resultado que es aún básico, a saber, que la curvatura (total) K de la superficie es

$$K = \frac{LN - M^2}{EG - F^2} \quad (6)$$

donde

$$L = \begin{vmatrix} x_{uu} & y_{uu} & z_{uu} \\ x_u & y_u & z_u \\ x_v & y_v & z_v \end{vmatrix} \quad M = \begin{vmatrix} x_{uv} & y_{uv} & z_{uv} \\ x_u & y_u & z_u \\ x_v & y_v & z_v \end{vmatrix} \quad N = \begin{vmatrix} x_{vv} & y_{vv} & z_{vv} \\ x_u & y_u & z_u \\ x_v & y_v & z_v \end{vmatrix} \quad (7)$$

Entonces, Gauss demuestra que esta K es el producto de las dos curvaturas principales en (x,y,z) , ya introducidas por Euler. En 1831, Sophie Germain¹⁰⁰ introdujo la noción de curvatura media, a saber, el promedio de las dos curvaturas principales.

Ahora Gauss hace una observación extremadamente importante. Cuando la superficie está dada mediante las ecuaciones paramétricas (1), las propiedades de la superficie parecen depender de las funciones x , y y z . Fijando u , digamos $u = 3$, y dejando variar v , se obtiene una curva sobre la superficie. Se obtiene una familia de curvas para los diversos valores fijos de u . De manera análoga, fijando v se obtiene una familia de curvas. Estas dos familias son las curvas paramétricas sobre la superficie, de tal forma que cada punto está dado por un par de números (c,d) , digamos, donde $u = c$ y $v = d$ son las curvas paramétricas que pasan por un punto. Estas coordenadas no denotan necesariamente distancias más de lo que lo

hacen las latitudes y las longitudes. Pensemos en una superficie en la cual las curvas paramétricas han sido determinadas de alguna manera. Entonces, afirma Gauss, las propiedades geométricas de la superficie están determinadas sólo por las E , F y G en la expresión (4) para ds^2 . Estas funciones de u y v son todo lo que importa.

Sucede ciertamente, como es evidente a partir de (4) y (5), que las distancias y los ángulos sobre la superficie están determinados por las E , F y G . Pero la expresión fundamental de Gauss para la curvatura, la (6) arriba, depende de las cantidades adicionales L , M y N . Gauss ahora demuestra que

$$K = \frac{1}{2H} \frac{\partial}{\partial u} \left\{ \left[\frac{F}{EH} \frac{\partial E}{\partial v} - \frac{1}{H} \frac{\partial G}{\partial u} + \right. \right. \quad (8)$$

$$\left. \left. + \frac{\partial}{\partial v} \left[\frac{2}{H} \frac{\partial F}{\partial u} - \frac{1}{H} \frac{\partial E}{\partial v} - \frac{F}{EH} \frac{\partial E}{\partial u} \right] \right\}$$

donde $H = \sqrt{(EG - F^2)}$ y es igual a la Δ de Gauss definida anteriormente. La ecuación (8), llamada la ecuación característica de Gauss, muestra que la curvatura K , y, en particular, a la vista de (6), la cantidad $LN - M^2$ depende únicamente de E , F y G . Ya que E , F y G son funciones sólo de las coordenadas paramétricas sobre la superficie, la curvatura también es sólo función de los parámetros y no depende en absoluto de sí o cómo yace la superficie en el espacio tridimensional.

Gauss observó que las propiedades de una superficie dependen únicamente de E , F y G . Sin embargo, muchas otras propiedades

diferentes de la curvatura conciernen a las cantidades L , M y N , y no en la combinación $LN - M^2$ que aparece en la ecuación (6). Gaspare Mainardi (1800-1879),¹⁰¹ independientemente de Delfino Codazzi (1824-1875),¹⁰² realizó el trabajo analítico de la observación de Gauss y ambos proporcionaron dos relaciones adicionales en la forma de ecuaciones diferenciales que, junto con la ecuación característica de Gauss, teniendo k el valor en (6), determinan L , M y N en términos de E , F y G .

Más adelante, en 1867,¹⁰³ Ossian Bonnet (1819-1892) demostró el teorema de que cuando seis funciones satisfacen la ecuación característica de Gauss y las dos ecuaciones de Mainardi-Codazzi, entonces determinan una superficie única, con excepción de su posición y orientación en el espacio. Específicamente, si E , F y G y L , M y N están dadas como funciones de u y v que satisfacen la ecuación característica y las ecuaciones de Mainardi-Codazzi y si $EG - F^2 \neq 0$, entonces existe una superficie dada por las tres funciones de x , y y z como funciones de u y v que tiene la primera forma fundamental

$$E du^2 + 2F du dv + G dv^2,$$

y L , M y N están relacionadas con E , F y G a través de (7). Esta superficie está determinada unívocamente, excepto por la posición en el espacio. (Para superficies reales con coordenadas reales (u, v) , debemos tener $EG - F^2 > 0$, $E > 0$ y $G \geq 0$). El teorema de Bonnet es

el análogo del teorema correspondiente sobre curvas (cap. 23, sec. 6).

El hecho de que las propiedades de una superficie dependen sólo de E , F y G tiene muchas implicaciones, algunas de las cuales dedujo Gauss en su ensayo de 1827. Por ejemplo, si una superficie es doblada sin estirla o contraerla, las líneas coordenadas $u = \text{const.}$ y $v = \text{const.}$ permanecerán y también ds será la misma. De aquí que todas las propiedades de la superficie, en particular la curvatura, serán las mismas. Más aún, si dos superficies pueden ser puestas en una correspondencia uno-a-uno una con la otra, esto es, si $u' = \theta(u, v)$, $v' = \psi(u, v)$, donde u' y v' son las coordenadas de los puntos sobre la segunda superficie, y si el elemento de distancia en puntos correspondientes es el mismo sobre las dos superficies, esto es, si

$$E du^2 + 2F du dv + G dv^2 = E' du'^2 + 2F' du' dv' + G' dv'^2$$

donde las E , F y G son funciones de u y v y E' , F' y G' son funciones de u' y v' , entonces las dos superficies, que se dice son *isométricas*, deben tener la misma geometría. En particular, como señaló Gauss, deben tener la misma curvatura total en puntos correspondientes. A este resultado lo llamó Gauss un «teorema egregium», un teorema excelentísimo.

Como corolario, se sigue que para mover una parte de una superficie sobre otra parte (lo que significa preservar la distancia) una condición necesaria es que la superficie tenga curvatura constante. De esta manera, una parte de una esfera puede ser

movida sin distorsión sobre otra, pero esto no puede ser realizado con un elipsoide. (Sin embargo, sí puede tener lugar un plegamiento al acomodar una superficie o parte de una superficie dentro de otra bajo una aplicación isométrica). A pesar de que las curvaturas en puntos correspondientes son iguales, si dos superficies no tienen curvatura constante, no están relacionadas necesariamente isométricamente. En 1839,¹⁰⁴ Ferdinand Minding (1806-1885) demostró que si dos superficies tienen curvatura constante e igual entonces una puede ser aplicada isométricamente sobre la otra.

Otra cuestión de gran importancia que Gauss estudió en su ensayo de 1827 es la de encontrar geodésicas sobre las superficies. (Liouville introdujo, en 1850, el término geodésica, que fue tomado de la geodesia). Este problema requiere del cálculo de variaciones, que Gauss utiliza. Enfoca este tema trabajando con la representación

x, y, z y demuestra el teorema establecido por Jean Bernoulli de que la normal principal a una curva geodésica es normal a la superficie. (Así, la normal principal a un punto de un círculo de latitud sobre una esfera yace en el plano del círculo y no es normal a la esfera, mientras que la normal principal en un punto sobre el círculo de longitud es normal a la esfera). Cualquier relación entre u y v determina una curva sobre una superficie y la relación que proporciona una geodésica está determinada por una ecuación diferencial. Esta ecuación, de la que Gauss únicamente dice que es una ecuación de segundo orden en u y v , pero que no proporciona explícitamente, se escribe de muchas formas. Una es

$$\frac{d^2v}{du^2} = n \left(\frac{dv}{du}\right)^3 + (2m - v) \left(\frac{dv}{du}\right)^2 + (l - 2\mu) \frac{dv}{du} - \lambda \quad (9)$$

donde n , m , p , v , l y λ son funciones de E , F y G .

Se debe ser cuidadoso al suponer la existencia de una geodésica única entre dos puntos sobre una superficie. Dos puntos cercanos sobre una superficie esférica tienen una geodésica única que los une, pero dos puntos diametralmente opuestos están unidos por una infinidad de geodésicas. Análogamente, dos puntos sobre el mismo generador de un cilindro circular están unidos mediante una geodésica a lo largo de la generatriz, pero también por un número infinito de hélices geodésicas. Si existe un solo arco geodésico entre dos puntos de una región, el arco proporciona la trayectoria mínima entre ellos en la región. Muchos matemáticos se dedicaron al problema de determinar de hecho las geodésicas sobre superficies particulares.

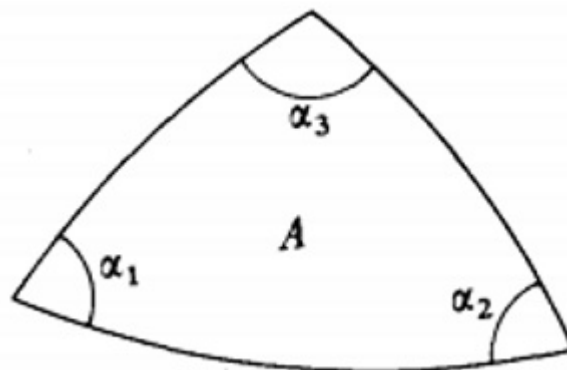


Figura 37.1

En su artículo de 1827, Gauss demostró un famoso teorema sobre curvatura para un triángulo formado por geodésicas (fig. 37.1). Sea K la curvatura variable de una superficie. $\iint_A dA$ es entonces la integral de esta curvatura sobre el área A . El teorema de Gauss, cuando se aplica al triángulo, establece que

$$\iint_A K dA = \alpha_1 + \alpha_2 + \alpha_3 - \pi$$

esto es, la integral de la curvatura sobre un triángulo geodésico es igual al exceso de la suma de los ángulos sobre 180° o, cuando el ángulo suma es menor que π , al defecto con 180° . Este teorema, dice Gauss, debe ser considerado como un teorema elegantísimo. El resultado generaliza el teorema de Lambert (cap. 36, sec. 4), que afirma que el área de un triángulo esférico iguala el producto de su exceso esférico y al cuadrado del radio, ya que en un triángulo esférico K es constante e igual a $1/R^2$.

Debemos mencionar un trabajo también importante de la geometría diferencial de Gauss. Lagrange (cap. 23, sec. 8) había tratado la aplicación conforme de una superficie de revolución en un plano. En 1822, Gauss ganó un premio, propuesto por la Real Sociedad Danesa de Ciencias a un ensayo sobre el problema de encontrar la condición analítica para transformar cualquier superficie conformemente sobre cualquier otra superficie.¹⁰⁵ Su condición, que se refiere a la vecindad de puntos correspondientes sobre las dos superficies, es igual al hecho de que una función $P + iQ$, que no

especificaremos con más detalle, de los parámetros T y U por medio de los cuales una superficie está representada es una función f de $p + iq$, que es la función correspondiente de los parámetros t y u por la cual está representada la otra superficie y $P - iQ$ es $f(p - iq)$, donde f es f o es obtenida de/reemplazando i por $-i$. La función f depende de la correspondencia entre las dos superficies, estando la correspondencia especificada por $T = T(t,u)$ y $U(t,u)$. Gauss no respondió a la cuestión de si y en qué manera una porción *finita* de la superficie puede ser aplicada conformemente sobre la otra superficie. Riemann trató este problema en su trabajo sobre funciones complejas (cap. 27, sec. 10).

El trabajo de Gauss en geometría diferencial es un punto de referencia en sí mismo. Pero sus implicaciones fueron mucho más profundas que lo que él mismo apreció. Hasta sus trabajos, las superficies habían sido estudiadas como figuras en un espacio euclídeo tridimensional. Pero Gauss demostró que la geometría de una superficie podía ser estudiada concentrándose en la misma superficie. Si se introducen las coordenadas u y v que provienen de la representación paramétrica

$$x = x(u, v), \quad y = y(u, v), \quad z = z(u, v)$$

de la superficie en el espacio tridimensional y se usan las E , F y G determinadas de este modo, entonces se obtienen las propiedades euclídeas de esa superficie. Sin embargo, dadas estas coordenadas u y v sobre la superficie y la expresión para ds^2 en términos de E , F

y G como funciones de u y v , todas las propiedades de la superficie se siguen de esta expresión. Esto sugiere dos ideas fundamentales. El primero es que la superficie puede ser considerada como un espacio en sí misma, ya que todas sus propiedades están determinadas por la ds^2 . Se puede olvidar el hecho de que la superficie yace en un espacio tridimensional. ¿Qué clase de geometría posee la superficie si es considerada como un espacio en sí misma? Si uno toma como «líneas rectas» las geodésicas sobre esa superficie, entonces la geometría es no euclídea.

De esta manera, si se estudia la superficie de la esfera como un espacio en sí misma, tiene su propia geometría, y si además se usan las latitudes y longitudes familiares como coordenadas de los puntos, la geometría sobre esa superficie es no euclídea, ya que las «líneas rectas» o geodésicas son arcos de los círculos máximos sobre la superficie. Sin embargo, la geometría de la superficie esférica es euclídea si se la considera como una superficie en el espacio tridimensional. La distancia mínima entre dos puntos sobre la superficie es entonces el segmento lineal de una geometría euclídea tridimensional (aunque no yace sobre la superficie). El trabajo de Gauss implica que existen geometrías no euclídeas al menos sobre superficies consideradas como espacios en sí mismas. No está claro si Gauss vio o no esta interpretación de su geometría de superficies. Se puede ir más allá. Se podría pensar que están determinadas las propias E , F y G para una superficie mediante las ecuaciones paramétricas (1). Pero se podría empezar con la superficie, introducir las dos familias de curvas paramétricas y entonces

escoger casi arbitrariamente funciones E , F y G de u y v . Entonces la superficie tiene una geometría determinada por estas E , F y G . Esta geometría es intrínseca a la superficie y no tiene conexión con el espacio que la rodea. Consecuentemente, la misma superficie puede tener *diferentes* geometrías, dependiendo de la elección de las funciones E , F y G .

Las implicaciones son más profundas. Si se pueden escoger diferentes conjuntos de E , F y G y, de esta manera, determinar diferentes geometrías sobre la misma superficie, ¿por qué no escoger diferentes funciones distancia en nuestro espacio tridimensional ordinario? La función de distancia común en coordenadas rectangulares es, por supuesto, $ds^2 = dx^2 + dy^2 + dz^2$ y esto es obligatorio si se empieza con la geometría euclídea, ya que es justamente el enunciado analítico del teorema pitagórico. Sin embargo, dadas las mismas coordenadas rectangulares para los puntos del espacio, se podría elegir una expresión diferente para ds^2 y obtener una geometría bastante diferente para ese espacio, una geometría no euclídea. Riemann tomó y desarrolló esta extensión a cualquier espacio de las ideas de Gauss que se obtuvo por primera vez al estudiar las superficies.

3. El enfoque de Riemann de la geometría

Las dudas acerca de lo que podemos creer de la geometría del espacio físico, originadas por el trabajo de Gauss, Lobachevski y Bolyai, estimularon una de las más grandes creaciones del siglo XIX: la geometría riemanniana. El creador fue Georg Bernhard

Riemann, el filósofo más profundo de la geometría. Aunque los detalles de los trabajos de Lobachevski y Bolyai le eran desconocidos, eran conocidos de Gauss, y Riemann ciertamente conoció las dudas de Gauss en cuanto a la verdad y necesaria aplicabilidad de la geometría euclídea. De esta manera, en el campo de la geometría, Riemann siguió a Gauss mientras que en la teoría de funciones siguió a Cauchy y Abel. Las enseñanzas del psicólogo Johann Friedrich Herbart (1776-1841) influenciaron sus investigaciones de geometría.

Gauss propuso a Riemann el tema de los fundamentos de la geometría, sobre el que debería pronunciar su conferencia calificadora, la *Habilitationsvortrag*, para el título de *Privatdozent*. Pronunció la conferencia de 1854, dirigida al cuerpo académico de Göttingen con Gauss presente, y fue publicada en 1868 con el título: «*Über die Hypothese, welche der Geometrie zu Grunde liegen*» («Sobre las hipótesis en que se funda la geometría»).¹⁰⁶

En un artículo sobre la conducción del calor, que Riemann escribió en 1861 para competir por un premio ofrecido por la Academia de Ciencias de París, y al que comúnmente se cita como el *Pariserarbeit*, Riemann encontró la necesidad de considerar más profundamente sus ideas sobre geometría, y aquí proporcionó algunas elaboraciones técnicas de su ensayo de 1854. Este ensayo de 1861, que no ganó el premio, se publicó postumamente en 1876 en sus *Obras Completas*.¹⁰⁷ En la segunda edición de sus *Werke*, Heinrich Weber explica en una nota el material, muy resumido, de Riemann.

La geometría del espacio de Riemann no fue únicamente una extensión de la geometría diferencial de Gauss. Reconsideró el enfoque total del estudio del espacio. Riemann estudió la cuestión de en qué podemos estar seguros acerca del espacio físico. ¿Qué condiciones o hechos se presuponen en la propia experiencia del espacio antes de que nosotros determinemos por experiencia los axiomas particulares que se cumplen en el espacio físico? Uno de los objetivos de Riemann fue demostrar que los axiomas particulares de Euclides eran empíricos en lugar de, como se había creído, verdades autoevidentes. Adoptó el enfoque analítico, ya que en las demostraciones geométricas podemos ser desviados por nuestras percepciones para aceptar hechos no reconocidos explícitamente. Así, la idea de Riemann era apoyarse en el análisis para empezar con lo que seguramente es *a priori* acerca del espacio y deducir las consecuencias necesarias. Cualesquiera otras propiedades del espacio serían entonces tomadas como empíricas. Gauss se había preocupado de este mismo problema, pero sólo se publicó el ensayo sobre superficies curvas de esta investigación. La búsqueda de Riemann por lo que es *a priori* lo llevó a estudiar el comportamiento local del espacio, o, en otras palabras, el enfoque de la geometría diferencial como opuesto a la consideración del espacio como un todo como se encuentra en Euclides o en la geometría no euclídea de Gauss, Bolyai y Lobachevski. Antes de examinar los detalles, debemos estar advertidos de que las ideas de Riemann expresadas en su conferencia y en su manuscrito de 1854

son vagas. Una razón es que Riemann las adaptó para su audiencia, el profesorado de Göttingen al completo.

Parte de la vaguedad surge a partir de las consideraciones filosóficas con las que Riemann inició su artículo.

Guiado en gran parte por la geometría intrínseca de Gauss de las superficies en el espacio euclídeo, Riemann desarrolló una geometría intrínseca para cualquier espacio. Prefirió tratar una geometría «-dimensional a pesar de que el caso tridimensional era claramente el importante, y habla del espacio «-dimensional como una variedad. Un punto en la variedad de « dimensiones está representado mediante la asignación de valores especiales a n parámetros variables, x_1, x_2, \dots, x_n , y el agregado de todos esos tales posibles puntos constituye la variedad n -dimensional en sí misma, de la misma manera que el agregado de los puntos sobre la superficie constituye la propia superficie. Se llama a los n parámetros variables las coordenadas de la variedad. Cuando las x varían continuamente, los puntos recorren la variedad.

Puesto que Riemann creía que sólo conocemos el espacio localmente, empezó por definir la distancia entre dos puntos genéricos cuyas coordenadas correspondientes difieren sólo en cantidades infinitesimales. Supone que el cuadrado de la distancia es

$$ds^2 = \sum_{i=1}^n \sum_{j=1}^n g_{ij} dx_i dx_j \quad (10)$$

donde las g_{ij} son funciones de las coordenadas x_1, x_2, \dots, x_n , $g_{ij} = g_{ji}$ y el lado derecho de (10) es siempre positivo para todos los valores posibles de las dx_i . Esta expresión para las ds^2 es una generalización de la fórmula de la distancia euclídea

$$ds^2 = d_1^2 + d_2^2 + \dots + d_n^2$$

Menciona la posibilidad de suponer que la ds^2 es la raíz cuarta de una función homogénea del cuarto grado en las diferenciales dx_1, dx_2, \dots, dx_n , pero no siguió con esta posibilidad. Al permitir que las g - sean funciones de las coordenadas, Riemann prevé la posibilidad de que la naturaleza del espacio pudiera variar de un punto a otro. Aunque en su ensayo de 1854 Riemann no estableció explícitamente las definiciones siguientes, indudablemente las tenía en su mente, porque se asemejan a lo que Gauss hizo para las superficies. Una curva sobre una variedad riemana está dada por el conjunto de « funciones.

$$x_1 = x_1(t) \quad x_2 = x_2(t) \quad x_n = x_n(t) \quad (11)$$

Entonces la longitud de la curva entre $t = \alpha$ y $t = \beta$ está definida por

$$l = \int_{\alpha}^{\beta} ds = \int_{\alpha}^{\beta} \frac{ds}{dt} dt = \int_{\alpha}^{\beta} \sqrt{\sum_{i,j=1}^n g_{ij} \frac{dx_i}{dt} \frac{dx_j}{dt}} dt \quad (12)$$

La curva más corta entre dos puntos dados, $t = a$ y $t = \beta$ la geodésica, se determina entonces por el método del cálculo de variaciones. En la notación de esa materia es la curva para la cual

$$\delta \int_a^\beta ds = 0$$

Se debe determinar entonces las funciones particulares de la forma (11) que proporcionan esta trayectoria mínima entre los dos puntos. En términos del parámetro de longitud de arco, las ecuaciones de las geodésicas resultan ser

$$\frac{d^2 x_i}{ds^2} + \sum_{\lambda, \mu} \left\{ \begin{matrix} \lambda, \mu \\ i \end{matrix} \right\} \frac{dx_\lambda}{ds} \frac{dx_\mu}{ds} = 0 \quad i, \lambda, \mu = 1, 2, \dots, n$$

Este es un sistema de n ecuaciones diferenciales ordinarias de segundo orden.¹⁰⁸

El ángulo θ entre dos curvas que se cortan en el punto (x_1, x_2, \dots, x_n) , estando una curva determinada por las direcciones dx_i/ds , $i = 1, 2, \dots, n$, y la otra por dx_i/ds' , $i = 1, 2, \dots, n$, donde las primas indican valores que pertenecen a la segunda dirección, está definido por la fórmula

$$\cos \theta = \sum_{i, i'}^n g_{i i'} \frac{dx_i}{ds} \frac{dx'_{i'}}{ds} \quad (13)$$

Siguiendo los procedimientos que Gauss usó para las superficies, se desarrolla una geometría métrica n -dimensional sobre la base de las definiciones anteriores. Todas las propiedades métricas están determinadas por los coeficientes en la expresión para las ds^2 .

La noción de curvatura de una variedad es el segundo concepto importante en el ensayo de Riemann de 1854. A través de ésta, Riemann buscó caracterizar el espacio euclídeo y, más generalmente, espacios en los que las figuras pueden ser movidas sin cambio en la forma o en la magnitud. La noción de Riemann de curvatura para cualquier variedad n -dimensional es una generalización de la noción de Gauss de curvatura total para las superficies. Al igual que la noción de Gauss, la curvatura de la variedad está definida en términos de cantidades determinables en la propia variedad y no hay necesidad de pensar en la variedad como si estuviera en alguna variedad de dimensión superior.

Dado un punto P en la variedad n -dimensional, Riemann considera una variedad bidimensional en el punto y en la variedad n -dimensional. Tal variedad bidimensional está formada por un único conjunto infinito de geodésicas que pasan por el punto y son tangentes a una sección plana de la variedad por el punto P . Ahora se puede describir una geodésica por medio del punto P y una dirección en ese punto. Sean $dx'_1 dx'_2, \dots, dx'_n$ la dirección de una geodésica y $dx''_1 dx''_2, \dots, dx''_n$ la dirección de otra. Entonces, la i -ésima dirección de cualquiera del conjunto infinito de geodésicas en P está dado mediante

$$dx_i = \lambda' dx'_i + \lambda'' dx''_i$$

(sujeta a la condición $\lambda'^2 + \lambda''^2 + 2\lambda'\lambda'' \cos \theta = 1$ que viene de la condición $\Sigma g_{ij}(dx/ds) (dx/ds) = 1$).

Este conjunto de geodésicas forma una variedad bidimensional que tiene una curvatura gaussiana. Ya que existe una infinidad de tales variedades bidimensionales en P , obtenemos un número infinito de curvaturas en ese punto en la variedad «-dimensional. Pero de $n(n - 1)/2$ de estas medidas de curvatura se deduce el resto. Se deriva una expresión explícita para la medida de la curvatura. Riemann hizo esto en su ensayo de 1861 y lo daremos más adelante. Para una variedad que es una superficie, la curvatura de Riemann es exactamente la curvatura total de Gauss. Hablando estrictamente, la curvatura de Riemann, como la de Gauss, es una propiedad de la métrica impuesta sobre la variedad y no de la misma variedad.

Después de que Riemann completase su investigación general de geometría n -dimensional y demostrase cómo se introduce la curvatura, consideró variedades más restringidas, en las cuales las formas espaciales finitas debían ser capaces de movimiento sin cambio de tamaño o forma y debían ser capaces de rotación en cualquier dirección. Esto lo condujo a espacios de curvatura constante.

Cuando todas las medidas de curvatura en un punto son las mismas e iguales a todas las medidas en cualquier otro punto, obtenemos lo que Riemann llama una variedad de curvatura

constante. Sobre tal variedad es posible tratar figuras congruentes. En el artículo de 1854, Riemann proporcionó los siguientes resultados, pero no dio detalles: si α es la medida de curvatura, la fórmula para el elemento infinitesimal de distancia sobre una variedad de curvatura *constante* se convierte (en un sistema de coordenadas adecuado) en

$$ds^2 = \frac{\sum_{i=1}^n dx_i^2}{\left(1 + \left(\frac{\alpha}{4}\right) \sum x_i^2\right)^2} \quad (14)$$

Riemann pensó que la curvatura α debía ser positiva o cero, de tal forma que cuando $\alpha > 0$ obtenemos un espacio esférico y cuando $\alpha = 0$ obtenemos un espacio euclídeo, e inversamente. También creía que si un espacio estaba infinitamente extendido, la curvatura debía ser cero. Sugirió, sin embargo, que podía existir una superficie real de curvatura negativa constante.¹⁰⁹

Para continuar con Riemann, para $\alpha = \alpha^2 > 0$ y $n = 3$ obtenemos una geometría esférica tridimensional, aunque no podemos visualizarla. El espacio es finito en extensión pero sin límites. Todas las geodésicas en él son de longitud constante, a saber, $2\pi/\alpha$, y regresan sobre sí mismas. El volumen del espacio es $2\pi^2/\alpha^3$. Para $\alpha^2 > 0$ y $n = 2$ obtenemos el espacio de la superficie esférica ordinaria. Las geodésicas son, por supuesto, los círculos máximos y son finitos. Aún más, dos cualesquiera se intersecan en dos puntos. De hecho, no está claro si Riemann consideró las geodésicas de una superficie de curvatura positiva constante como cortándose en uno

o dos puntos. Posiblemente pretendía lo último. Félix Klein señaló más tarde (véase el próximo capítulo) que estaban envueltas dos geometrías distintas.

Riemann también señala una distinción, de la cual se hablará más tarde, entre la falta de límite e infinidad del espacio [como en el caso de la superficie de una esfera]. Lo ilimitado, dice, tiene una credibilidad empírica mayor que cualquier otro hecho derivado empíricamente, tal como la extensión infinita.

Al final de este ensayo, Riemann nota que ya que el espacio físico es una clase especial de variedad, la geometría de ese espacio no puede ser derivada únicamente de las nociones generales de las variedades. Las propiedades que distinguen el espacio físico de cualesquiera otras variedades extendidas triplemente han de ser obtenidas únicamente por medio de la experiencia. Añade: «*Queda por resolver la cuestión de conocer en qué medida y hasta qué punto estas hipótesis acerca de variedades son confirmadas por la experiencia.*» En particular, los axiomas de la geometría euclídea pueden ser ciertos sólo aproximadamente del espacio físico. Como Lobachewsky, Riemann pensaba que la astronomía decidiría qué geometría se ajusta al espacio. Finaliza su ensayo con esta observación profética: «*O, por tanto, la realidad que subyace al espacio debe formar una variedad discreta o debemos buscar la base de sus relaciones métricas fuera de ella, en las fuerzas envolventes que actúan sobre ella... Esto nos conduce al dominio de otra ciencia, la de la física, a la que el objeto de nuestro trabajo no nos permite ir hoy.*»

William Kingdon Clifford desarrolló este punto:¹¹⁰

Mantengo de hecho:

- 1. Que las porciones pequeñas del espacio son de una naturaleza análoga a pequeñas colinas sobre una superficie que es en promedio plana.*
- 2. Que esta propiedad de ser curvada o distorsionada es transmitida continuamente de una porción del espacio a otra a la manera de una onda.*
- 3. Que esta variación de la curvatura del espacio es realmente lo que sucede en aquel fenómeno que nosotros llamamos el movimiento de la materia, ya sea ponderable o etéreo.*
- 4. Que en este mundo físico nada tiene lugar excepto esta variación, sujeta, posiblemente, a la ley de continuidad.*

Las leyes ordinarias de la geometría euclídea no son válidas para un espacio cuya curvatura cambia, no solamente con el lugar sino debido al movimiento de la materia, de un tiempo a otro. Añadió que una investigación más exacta de las leyes físicas no sería capaz de ignorar estas «colinas» en el espacio. De esta manera, Riemann y Clifford, contrariamente a la gran mayoría de los otros geómetras, sintieron la necesidad de asociar la materia con el espacio para poder determinar lo que es cierto del espacio físico. Esta línea de pensamiento conduce, por supuesto, a la teoría de la relatividad.

En su *Pariserarbeit* (1861), Riemann regresó a la cuestión de cuándo un espacio riemanniano cuya métrica es

$$ds^2 = \sum_{i,j=1}^n g_{ij} dx_i dx_j \quad (15)$$

puede ser un espacio de curvatura constante o incluso un espacio euclídeo. Sin embargo, formuló la cuestión más general de cuándo una métrica tal como (15) puede ser transformada por las ecuaciones

$$x_i = x_i(y_1, y_2, \dots, y_n), \quad i= 1,2, \dots, n \quad (16)$$

en una métrica dada

$$ds'^2 = \sum_{i,j=1}^n h_{ij} dy_i dy_j \quad (17)$$

entendiéndose por supuesto, que ds igualará a ds' , de tal forma que las geometrías de los dos espacios serían la misma excepto por la elección de coordenadas. La transformación (16) no siempre es posible porque, como señala Riemann, hay $n(n + 1)/2$ funciones independientes en (15), mientras que la transformación introduce únicamente n funciones que pueden ser usadas para convertir las g_{ij} en las h_{ij}

Para discutir la cuestión general, Riemann introdujo cantidades especiales p_{ijk} , que debemos reemplazar por los más familiares símbolos de Christoffel entendiéndose que

$$p_{ijk} = \begin{bmatrix} jk \\ i \end{bmatrix}$$

Los símbolos de Christoffel, denotados en varias formas, son

$$\Gamma_{\alpha\beta,\lambda} = \begin{bmatrix} \alpha\beta \\ \lambda \end{bmatrix} = [\alpha\beta, \lambda] = \frac{1}{2} \left(\frac{\partial g_{\alpha\lambda}}{\partial x_\beta} + \frac{\partial g_{\beta\lambda}}{\partial x_\alpha} - \frac{\partial g_{\alpha\beta}}{\partial x_\lambda} \right) \quad (18)$$

$$\Gamma_{\alpha\beta}^\lambda = \begin{bmatrix} \alpha\beta \\ \lambda \end{bmatrix} = \{\alpha\beta, \lambda\} = \sum_i g^{i\lambda} \begin{bmatrix} \alpha\beta \\ i \end{bmatrix} \quad (19)$$

donde $g^{i\lambda}$ es el cofactor dividido por g de $g_{i\lambda}$ en el determinante de g . Riemann también introdujo lo que ahora es conocido como el símbolo riemanniano de cuatro índices

$$(\mu\lambda, jk) = R_{\lambda\mu jk} = \frac{\partial \Gamma_{\lambda j, \mu}}{\partial x_k} - \frac{\partial \Gamma_{\lambda k, \mu}}{\partial x_j} + \sum_{i, \alpha} g^{i\alpha} (\Gamma_{\lambda k, \alpha} \Gamma_{\mu j, i} - \Gamma_{\lambda j, \alpha} \Gamma_{\mu k, i}) \quad (20)$$

Más adelante, Riemann muestra que una condición necesaria para que ds^2 sea transformable en ds'^2 es

$$(\alpha\delta, \beta\gamma)' = \sum_{r,k,i,h} (rk, ih) \frac{\partial x_r}{\partial y_\alpha} \frac{\partial x_i}{\partial y_\beta} \frac{\partial x_h}{\partial y_\gamma} \frac{\partial x_k}{\partial y_\delta} \quad (21)$$

donde la parte izquierda del símbolo se refiere a cantidades formadas para las métricas ds' y (21) se cumple para todos los valores de $\alpha, \beta, \gamma, \delta$, cada uno de los cuales varía de 1 a n .

Ahora Riemann se dirige a la cuestión específica de cuándo una ds^2 dada puede ser transformada en una con coeficientes constantes. Primero deriva una expresión explícita para la curvatura de una variedad. La definición general ya dada en el ensayo de 1854 hace uso de las líneas geodésicas partiendo de un punto O del espacio. Sean d y δ que determinan dos vectores o direcciones de geodésicas saliendo de O (cada dirección es especificada por las componentes de la tangente a la geodésica). Considérese entonces el haz de vectores geodésicos partiendo de O y dado por $kd + \lambda\delta$ donde k y λ son parámetros. Si se piensa en d y δ como operando sobre las $x_i = f_i(t)$ que describe cualquier otra curva, entonces existe un significado para la diferencial segunda $(kd + \lambda\delta)^2 = k^2 d^2 + 2kd\lambda\delta + \lambda^2\delta^2$. Entonces Riemann forma

$$\Omega = \delta\delta \sum g_{ij} dx_i dx_j - 2d\delta \sum g_{ij} dx_i dx_j + dd \sum g_{ij} dx_i dx_j \quad (22)$$

Aquí se entiende que las d y δ operan formalmente sobre las expresiones que las siguen (y d y δ conmutan) de tal forma que

$$\begin{aligned} \delta\delta \sum g_{ij} dx_i dx_j &= \delta \left[\sum (\delta g_{ij}) dx_i dx_j + \right. \\ &\left. + \sum g_{ij} ((\delta dx_i) dx_j + dx_i \delta dx_j) \right] \end{aligned} \quad (23)$$

y

$$\delta g_{ij} = \sum_r \left(\frac{\partial g_{ij}}{\partial x_r} \right) \delta x_r$$

Si se calcula Ω se encuentra que todos los términos involucrando diferenciales terceras de una función se anulan. Sólo los términos en que aparecen δx_i , dx_i , $\delta^2 x_i$, δdx_i y $d^2 x_i$ permanecen. Calculando estos términos y usando la notación

$$p_{ik} = dx_i \delta x_k - dx_k \delta x_i$$

Riemann obtiene

$$[\Omega] = \sum_{i,k,r,s} (ik,rs) p_{ik} p_{rs} \quad (24)$$

Ahora hagamos

$$4\Delta^2 = \sum g_{ij} dx_i dx_j \sum g_{ij} \delta x_i \delta x_j - \left(\sum g_{ij} dx_i \delta x_j \right)^2$$

Entonces la curvatura K de una variedad riemanniana es

$$K = -\frac{[\Omega]}{8\Delta^2} \quad (25)$$

La conclusión global es que la condición necesaria y suficiente para que una ds^2 dada puede ser llevada a la forma (para $n = 3$)

$$ds'^2 = c_1 dx_1^2 + c_2 dx_2^2 + c_3 dx_3^2 \quad (26)$$

donde las c_i son constantes, es que todos los símbolos $(\alpha\beta, \gamma\delta)$ sean cero. En el caso de que las c_i sean todas positivas, las ds' pueden ser reducidas a $dy_1^2 + dy_2^2 + dy_3^2$, esto es, el espacio es euclídeo. Como podemos ver por el valor de $[\Omega]$, cuando la K es cero, el espacio es esencialmente euclídeo.

Es útil notar que la curvatura de Riemann para una variedad n -dimensional se reduce a la curvatura total de Gauss de una superficie. De hecho cuando

$$ds^2 = g_{11} dx_1^2 + 2g_{12} dx_1 dx_2 + g_{22} dx_2^2,$$

de los dieciséis símbolos $(\alpha\beta, \gamma\delta)$, 12 son cero y para los cuatro restantes tenemos $(12,12) = -(12,21) = -(21,12) = (21,21)$. Así, la K de Riemann se reduce a

$$K = \frac{(12,12)}{g}$$

Usando (20) se puede demostrar que esta expresión es igual a la de Gauss para la curvatura total de una superficie.

4. Los sucesores de Riemann

El ensayo de Riemann de 1854 fue publicado en 1868, dos años después de su muerte; creó gran interés, y muchos matemáticos se apresuraron a completar y extender las ideas que él había bosquejado. Los sucesores inmediatos de Riemann fueron Beltrami, Christoffel y Lipschitz.

Eugenio Beltrami (1835-1900), profesor de matemáticas en Bolonia y otras universidades italianas, quien conocía el artículo de Riemann de 1854, pero que aparentemente no conocía el de 1861, abordó la cuestión de demostrar que la expresión general para ds^2 se reduce a la forma (14) dada por Riemann para un espacio de curvatura constante.¹¹¹ Además de este resultado y de haber demostrado otras cuantas aserciones de Riemann, Beltrami se dedicó al tema de los invariantes diferenciales, que consideraremos en la siguiente sección.

Elwin Bruno Christoffel (1829-1900), que fuera profesor de matemáticas en Zurich y más tarde en Estrasburgo, avanzó las ideas contenidas en ambos ensayos de Riemann. En dos ensayos clave,¹¹² la mayor preocupación de Christoffel consistió en

reconsiderar y amplificar el tema tratado un tanto escuetamente por Riemann en su ensayo de 1861, a saber, cuándo una forma

$$F = \sum_{i,j} g_{ij} dx_i dx_j$$

puede ser transformada en otra

$$F' = \sum_{i,j} g'_{ij} dx_i dx_j$$

Christoffel buscó condiciones necesarias y suficientes. Incidentalmente, en este ensayo introdujo los símbolos de Christoffel. Consideremos primero el caso bidimensional donde

$$F = a dx^2 + 2b dx dy + c dy^2$$

y

$$F' = A dX^2 + 2B dX dY + C dY^2$$

y supongamos que las x y las y se expresan como funciones de X y de Y , de tal forma que F se convierte en F' bajo la transformación. Por supuesto, $dx = (\partial x / \partial X) dX + (\partial x / \partial Y) dY$. Ahora, cuando x , y , dx y dy son reemplazadas en F por sus valores en X e Y y cuando se

igualan los coeficientes en esta nueva forma de F con los de F' , se obtiene

$$a \left(\frac{\partial x}{\partial X} \right)^2 + 2b \frac{\partial x}{\partial X} \frac{\partial y}{\partial X} + c \left(\frac{\partial y}{\partial X} \right)^2 = A$$

$$a \frac{\partial x}{\partial X} \frac{\partial y}{\partial Y} + b \left(\frac{\partial x}{\partial X} \frac{\partial y}{\partial Y} + \frac{\partial x}{\partial Y} \frac{\partial y}{\partial X} \right) + c \frac{\partial y}{\partial X} \frac{\partial y}{\partial Y} = B$$

$$a \left(\frac{\partial x}{\partial Y} \right)^2 + 2b \frac{\partial x}{\partial Y} \frac{\partial y}{\partial Y} + c \left(\frac{\partial y}{\partial Y} \right)^2 = C$$

Estas son tres ecuaciones diferenciales para x e y como funciones de X e Y . Si se resuelven, entonces sabemos cómo transformar a F en F' . Sin embargo, hay sólo dos funciones relacionadas. Entonces deben existir algunas relaciones entre a , b y c , por un lado, y A , B y C por el otro. Diferenciando las tres ecuaciones anteriores y con pasos algebraicos posteriores se demuestra que la relación es $K = K'$. Para el caso de n variables, Christoffel utiliza la misma técnica. Empieza con

$$F = \sum g_{rs} dx_r dx_s$$

y

$$F' = \sum g'_{rs} dy_r dy_s$$

La transformación es

$$x_i = x_i(y_1, y_2, \dots, y_n), \quad i = 1, 2, \dots, n$$

y hace $g = |g_{rs}|$. Entonces, si Δ_{rs} es el cofactor de g_{rs} en el determinante, sea $g^{pq} = \Delta_{pq}/g$. Como Riemann, introduce independientemente los cuatro símbolos índice (sin la coma)

$$(gkhi) = \frac{\partial}{\partial x_i} [gh, k] + \sum_p (\{gi, p\}[hk, h] - \{gh, p\}[ik, p])$$

Más adelante, deduce $n(n + 1)/2$ ecuaciones en derivadas parciales para las x_i como funciones de las y_i . Una típica es

$$\sum_{r,s} g_{rs} \frac{\partial x_r}{\partial y_\alpha} \frac{\partial x_s}{\partial y_\beta} = g'_{\alpha\beta}$$

Estas ecuaciones son las condiciones necesarias y suficientes para que exista una transformación mediante la cual $F = F'$.

En parte para tratar la integrabilidad de este conjunto de ecuaciones, y en parte porque Christoffel desea considerar formas

de grado superior a dos en las dx_i , lleva a cabo un número de diferenciaciones y pasos algebraicos que muestran que

$$(\alpha\delta\beta\gamma)' = \sum_{g,h,i,k} (gkhi) \frac{\partial x_g}{\partial y_\alpha} \frac{\partial x_h}{\partial y_\beta} \frac{\partial x_i}{\partial y_\gamma} \frac{\partial x_k}{\partial y_\delta} \quad (27)$$

donde las α , δ , β y γ toman todos los valores de 1 a n . Hay $n^2(n^2 - 1)/12$ ecuaciones de esta forma. Estas ecuaciones son las condiciones necesarias y suficientes para la equivalencia de dos formas diferenciales de cuarto orden. En efecto, sean

$$d^{(1)}x, d^{(2)}x, d^{(3)}x, d^{(4)}x$$

cuatro conjuntos de diferenciales de x y de la misma manera para las y . Entonces, si tenemos la forma cuadrilineal

$$G_4 = \sum_{g,k,h,i} (gkhi) d^{(1)}x_g d^{(2)}x_k d^{(3)}x_h d^{(4)}x_i$$

las relaciones (27) son necesarias y suficientes para que $G_4 = G'_4$ donde G'_4 es el análogo de G_4 en las variables y .

Esta teoría puede ser generalizada para formas diferenciales μ -múltiples. De hecho, Christoffel introduce

$$G_\mu = \sum_{i_1, \dots, i_\mu} (i_1 i_2 \dots i_\mu) \partial_1 x_{i_1} \partial_2 x_{i_2} \dots \partial_\mu x_{i_\mu} \quad (28)$$

donde el término entre paréntesis está definido en términos de las g_{rs} , así como los cuatro símbolos índice y el símbolo ∂_i es utilizado para distinguir las diferenciales del conjunto de las x_i del conjunto obtenido al aplicar ∂_j . Más adelante demuestra que

$$(\alpha_1 \alpha_2 \dots \alpha_\mu)' = \sum (i_1 \dots i_\mu) \frac{\partial x_{i_1}}{\partial y_{\alpha_1}} \dots \frac{\partial x_{i_\mu}}{\partial y_{\alpha_\mu}} \quad (29)$$

y obtiene condiciones necesarias y suficientes para que las G_μ sean transformables en las G'_μ .

A continuación proporciona un procedimiento general mediante el cual se puede derivar de una forma μ -ésima G una forma $(\mu+1)$ -ésima. El paso clave es introducir

$$(i_1 i_2 \dots i_\mu i) = \frac{\partial y}{\partial x_i} (i_1 i_2 \dots i_\mu) - \sum_\lambda (\{i i_1, \lambda\} (\lambda i_2 \dots i_\mu) + \{i i_2, \lambda\} (i_1 \lambda i_3 \dots i_\mu) + \dots) \quad (30)$$

Estos símbolos son los coeficientes de la forma $G_{\mu+1}$. El procedimiento utilizado por Christoffel es el que Ricci y Levi-Civita llamaron más tarde diferenciación covariante (cap. 48).

Mientras que Christoffel sólo escribió un artículo clave sobre geometría riemanniana, Rudolph Lipschitz, profesor de matemáticas en la universidad de Bonn, escribió un gran número, apareciendo en el *Journal für Mathematik* a partir de 1869. Aunque hay algunas generalizaciones del trabajo de Beltrami y Christoffel, la materia esencial y los resultados son los mismos que los de los dos últimos. Obtuvo algunos nuevos resultados sobre subespacios riemannianos y espacios euclídeos n -dimensionales.

Las ideas proyectadas por Riemann y las desarrolladas por estos tres sucesores inmediatos sugirieron una gran variedad de nuevos problemas, tanto en la geometría diferencial euclídea como en la riemanniana. En particular, se generalizaron los resultados ya obtenidos en el caso euclídeo para tres dimensiones a curvas, superficies y formas de dimensión superior en n dimensiones. Citaremos sólo un resultado.

En 1886, Friedrich Schur (1856-1932) demostró el teorema que lleva su nombre.¹¹³ De acuerdo con el enfoque de Riemann de la noción de curvatura, Schur habla de la curvatura de una orientación en el espacio. Tal orientación está determinada por un haz de geodésicas $\mu\alpha + \lambda\beta$ donde α y β son las direcciones de las geodésicas que pasan por un punto. Este haz forma una superficie y tiene una curvatura de Gauss que Schur llama la curvatura riemanniana de la orientación. Su teorema establece entonces que si en cada punto la curvatura riemanniana del espacio es independiente de la orientación, entonces la curvatura riemanniana

es constante en todo el espacio. La variedad es entonces un espacio de curvatura constante.

5. Los invariantes de las formas diferenciales

Está claro a partir del estudio de la cuestión de cuándo una expresión dada para ds^2 puede ser transformada mediante una transformación de la forma

$$x_i = x_i(x_1', x_2', \dots, x_n') \quad i = 1, 2, \dots, n \quad (31)$$

en otra expresión semejante, con la conservación del valor de ds^2 , que se obtienen diferentes representaciones coordenadas para la misma variedad. Sin embargo, las propiedades geométricas de la variedad deben ser independientes del sistema coordenado particular usando para representarla y estudiarla. Analíticamente, estas propiedades geométricas serían representadas por medio de los invariantes, esto es, expresiones que retienen su forma bajo el cambio de coordenadas y que consecuentemente tendrán el mismo valor en un punto dado. Las invariantes de interés para la geometría riemanniana conciernen no sólo la forma cuadrática fundamental, que contiene las diferenciales dx_i y dx_j , sino que también puede contener las derivadas de los coeficientes y de otras funciones. Por tanto, son llamados invariantes diferenciales.

Tomamos el caso bidimensional como ejemplo: si

$$ds^2 = E du^2 + 2F du dv + G dv^2 \quad (32)$$

es el elemento de la distancia de una superficie, entonces la curvatura gaussiana K está dada por la fórmula (8) anterior. Si ahora las coordenadas se cambian a

$$u' = f(u, v) \quad v' = g(u, v) \quad (33)$$

entonces se tiene el teorema de que si

$$E du^2 + 2F du dv + G dv^2$$

se transforma en

$$E' du'^2 + 2F' du' dv' + G' dv'^2,$$

entonces $K = K'$, donde K' es la misma expresión que en (8) pero en las variables acentuadas. De aquí que la curvatura gaussiana de una superficie sea un invariante escalar. El invariante K es llamado un invariante asociado a la forma (32) y concierne sólo a E , F y G y sus derivadas.

Lamé, en un contexto más limitado, inició el estudio de los invariantes diferenciales. Estaba interesado en invariantes bajo transformaciones de un sistema coordenado curvilíneo ortogonal en tres dimensiones a otro. Para las coordenadas cartesianas rectangulares demostró¹¹⁴ que

$$\Delta_1\phi = \left(\frac{\partial\phi}{\partial x}\right)^2 + \left(\frac{\partial\phi}{\partial y}\right)^2 + \left(\frac{\partial\phi}{\partial z}\right)^2 \quad (34)$$

$$\Delta_2\phi = \frac{\partial^2\phi}{\partial x^2} + \frac{\partial^2\phi}{\partial y^2} + \frac{\partial^2\phi}{\partial z^2} \quad (35)$$

son invariantes diferenciales (los llamó parámetros diferenciales). Así, si Φ es transformada en $\Phi(x',y',z)$ por una transformación ortogonal (rotación de ejes) entonces

$$\left(\frac{\partial\phi}{\partial x}\right)^2 + \left(\frac{\partial\phi}{\partial y}\right)^2 + \left(\frac{\partial\phi}{\partial z}\right)^2 = \left(\frac{\partial\phi'}{\partial x'}\right)^2 + \left(\frac{\partial\phi'}{\partial y'}\right)^2 + \left(\frac{\partial\phi'}{\partial z'}\right)^2$$

en el mismo punto cuyas coordenadas son (x,y,z) en el sistema original y (x',y',z) en el nuevo sistema coordenado. Se tiene la ecuación análoga para $\Delta_2\Phi$.

Para las coordenadas curvilíneas ortogonales en el espacio euclídeo donde ds^2 tiene la forma

$$ds^2 = g_{11} du_1^2 + g_{22} du_2^2 + g_{33} du_3^2 \quad (36)$$

Lamé demostró (*Leçons sur les coordonnées curvilignes*, 1859, véase cap. 28, sec. 5) que la divergencia del gradiente de Φ , que en coordenadas rectangulares está dado mediante el $\Delta_2\Phi$ de arriba, tiene la forma invariante

$$\Delta_2\phi = \frac{1}{\sqrt{g_{11}g_{22}g_{33}}} \left[\frac{\partial}{\partial u_1} \left(\sqrt{\frac{g_{22}g_{33}}{g_{11}}} \frac{\partial\phi}{\partial u_1} \right) + \right. \\ \left. + \frac{\partial}{\partial u_2} \left(\sqrt{\frac{g_{22}g_{33}}{g_{11}}} \frac{\partial\phi}{\partial u_2} \right) + \frac{\partial}{\partial u_3} \left(\sqrt{\frac{g_{11}g_{22}}{g_{33}}} \frac{\partial\phi}{\partial u_3} \right) \right]$$

Incidentalmente, en este mismo trabajo proporcionó Lamé condiciones sobre cuándo las ds^2 dadas por (36) determinan un sistema coordenado curvilíneo en el espacio euclídeo y, si lo hace, cómo cambiar las coordenadas rectangulares.

La investigación de los invariantes para la teoría de superficies fue hecha primero por Beltrami,¹¹⁵ que dio los dos invariantes diferenciales

$$\Delta_1\phi = \frac{1}{EG - F^2} \left\{ E \left(\frac{\partial\phi}{\partial v} \right)^2 - 2F \frac{\partial y\phi}{\partial u} \frac{\partial\phi}{\partial v} + G \left(\frac{\partial\phi}{\partial u} \right)^2 \right\}$$

y

$$\Delta_2\phi = \frac{1}{\sqrt{EG - F^2}} \left\{ \frac{\partial}{\partial u} \left(\frac{G\phi_u - F\phi_v}{\sqrt{EG - F^2}} \right) + \frac{\partial}{\partial v} \left(\frac{-F\phi_u + E\phi_v}{\sqrt{EG - F^2}} \right) \right\}$$

Estos tienen un significado geométrico. Por ejemplo, en el caso de $\Delta_1\Phi$, si $\Delta_1\Phi = 1$, las curvas $\Phi(u,v) = \text{const.}$ son las trayectorias ortogonales de una familia de geodésicas sobre la superficie.

La búsqueda de invariantes diferenciales se extendió a las formas diferenciales cuadráticas de n variables. De nuevo la razón fue que estos invariantes son independientes de las elecciones particulares de coordenadas; representan propiedades intrínsecas de la propia variedad. De esta manera resulta que la curvatura de Riemann es un invariante escalar.

Beltrami, usando un método proporcionado por Jacobi,¹¹⁶ tuvo éxito en transportar a los espacios riemannianos «-dimensionales los invariantes de Lamé.¹¹⁷ Sea g como de costumbre el determinante de las g_{ij} y sea g^{ij} el cofactor dividido por g de g_{ij} en g . Entonces Beltrami demostró que el primer invariante de Lamé se convierte en

$$\Delta_1(\phi) = \sum g^{ij} \frac{\partial \phi}{\partial x_i} \frac{\partial \phi}{\partial x_j}$$

Esta es la forma general para el cuadrado del gradiente de 0. Para el segundo de los invariantes de Lamé, Beltrami obtuvo

$$\Delta_2(\phi) = \frac{1}{g} \sum_i \frac{\partial y}{\partial x_i} \left(\sqrt{g} \sum_i g^{ij} \frac{\partial \phi}{\partial x_j} \right)$$

También introdujo el invariante diferencial mixto

$$\Delta_1(\phi\psi) = \sum_{i,j} g^{ij} \frac{\partial \phi}{\partial x_i} \frac{\partial \psi}{\partial x_j}$$

Esta es la forma general del producto escalar de los gradientes de Φ y ψ .

Por supuesto, la forma ds^2 es ella misma un invariante para el cambio de coordenadas. De esto, como vimos en la sección previa, Christoffel derivó formas diferenciales de orden superior, sus G_4 y G_{μ} , que también son invariantes. Más aún, demostró cómo de las G_{μ} se derivan las $G_{\mu+1}$ que también son un invariante. Lipschitz siguió también la construcción de dichos invariantes. El número y la variedad son grandes. Como veremos, esta teoría de invariantes diferenciales fue la inspiración del análisis tensorial.

Bibliografía

- Beltrami, Eugenio: *Opere matematiche*, 4 vols., Ulrico Hoepli, 1902-1920.
- Clifford, William K.: *Mathematical papers*, Macmillan, 1882; Chelsea (reimpresión), 1968.
- Coolidge, Julián L.: *A history of geometrical methods*, Dover (reimpresión), 1963, 355-387.
- *Encyklopadie der Mathematischen Wissenschaften*, III, Teil 3, varios artículos, B. G. Teubner, 1902-1907.
- Gauss, Karl F.: *Werke*, 4, 192-216, 217-258, Königliche Gesellschaft der Wissenschaften zu Göttingen, 1880. «Investigaciones generales sobre superficies curvas», la traducción al inglés ha sido reimpresa por Raven Press, 1965.

- Helmholtz, Hermann von: «Über die tatsächlichen Grundlagen der Geometrie». *Wissenschaftliche Abhandlungen*, 2, 610-617. — «Über die Tatsachen, die der Geometrie zum Grunde liegen», *Nachrichten König. Ges. der Wiss. zu Gött.*, 15, 1868, 193-221; *Wiss. Abh.*, 618-639. — «Über den Ursprung Sinn und Bedeutung der geometrischen Sätze». Traducción al inglés: «Sobre el origen y significación de los axiomas geométricos», contenido en: Helmholtz: *Popular Scientific Papers*, Dover (reimpresión), 1962, 223-249. También en James R. Newman: *The world of mathematics*, Simón and Schuster, 1956, vol. 1, 647-668. Hay versión castellana, *Sigma. El mundo de las matemáticas*, Barcelona, Grijalbo, 8.^a ed., 1983.
- Jammer, Max: *Concepts of space*, Harvard University Press, 1954.
- Killing, W.: *Die nicht-euklidischen Raumformen in analytischer Behandlung*, B. G. Teubner, 1885.
- Klein, F.: *Vorlesungen über die Entwicklung der Mathematik im 19. Jahrhundert*, Chelsea (reimpresión), 1950, vol. 1, 6-62; vol. 2, 147-206.
- Pierpont, James: «Some modern views of space». *Amer. Math. Soc. Bull.*, 32, 1926, 225-258.
- Riemann, Bernhard: *Gesammelte mathematische Werke*, 2.^a ed. Dover (reimpresión), 1953, 272-287 y 391-404.
- Russell, Bertrand: *An essay on the foundations of geometry* (1897), Dover (reimpresión), 1956.

- Smith, David E.: *A source book in mathematics*. Dover (reimpresión), 1959, vol. 2, 411-425, 463-475. Contiene traducciones al inglés del ensayo de 1854 de Riemann y del ensayo de Gauss de 1822.
- Staeckel, P.: «Gauss als Geometer». *Nachrichten König. Ges. der Wiss. zu Gött.*, 1917, Beiheft, 25-140; también en *Werke*, IO2.
- Weatherburn, C. E.: «The development of multidimensional differential geometry». *Australian and New Zealand Ass'n for the Advancement of Science*, 21, 1933, 12-28.

Capítulo 38

Las geometrías proyectiva y métrica

Pero siempre debiera pedirse que un asunto matemático no se considere agotado hasta que haya llegado a ser intuitivamente evidente...

Félix Klein

Contenido:

- 1. Introducción*
- 2. Las superficies como modelos de la geometría no euclídea*
- 3. Las geometrías proyectiva y métrica*
- 4. Los modelos y el problema de la consistencia*
- 5. La geometría desde el punto de vista de las transformaciones*
- 6. La realidad de la geometría no euclídea*

Bibliografía

1. Introducción

Antes y durante el trabajo sobre geometría no euclídea, el estudio de las propiedades proyectivas fue la actividad geométrica principal. Además, fue evidente a partir del trabajo de von Staudt (cap. 35, sec. 3) que la geometría proyectiva precede lógicamente a la geometría euclídea, porque trata de las propiedades cualitativas y descriptivas que intervienen en la formación misma de las figuras

geométricas y no usa las medidas de los segmentos de recta ni los ángulos. Este hecho sugirió que la geometría euclídea podría ser algún caso particular de la geometría proyectiva. Teniendo después a la mano las geometrías no euclídeas surgió la posibilidad de que éstas también, al menos las que tratan de los espacios de curvatura constante, pudieran ser casos especiales de la geometría proyectiva. Así, la relación de la geometría proyectiva con las geometrías no euclídeas, las cuales también son geometrías métricas, ya que emplean la distancia como un concepto fundamental, se convirtió en un tema de investigación. El aclarar la relación de la geometría proyectiva con las geometrías euclídea y no euclídeas es el gran logro de las obras que estamos a punto de examinar. Igualmente vital fue el establecimiento de la consistencia de las geometrías no euclídeas básicas. Las superficies como modelos de la geometría no euclídea

Las geometrías no euclídeas que parecieron ser más significativas después de la obra de Riemann fueron las de los espacios de curvatura constante. El propio Riemann había sugerido en su artículo de 1854 que un espacio de curvatura constante positiva en dos dimensiones podría ser realizado sobre la superficie de una esfera siempre y cuando se tomara como «línea recta» la geodésica sobre la esfera. Ahora se hace referencia a esta geometría no euclídea como geometría elíptica doble por razones que se aclararán más tarde. Antes del trabajo de Riemann, la geometría no euclídea de Gauss, Lobachewsky y Bolyai, a la cual Klein llamó después geometría hiperbólica, había sido introducida como la geometría en

un plano en el que las líneas rectas ordinarias (y necesariamente infinitas) son las geodésicas. La relación de esta geometría con las variedades de Riemann no estaba clara. Riemann y Minding¹¹⁸ habían pensado en las superficies de curvatura constante negativa pero ninguno de ellos relacionó éstas con la geometría hiperbólica.

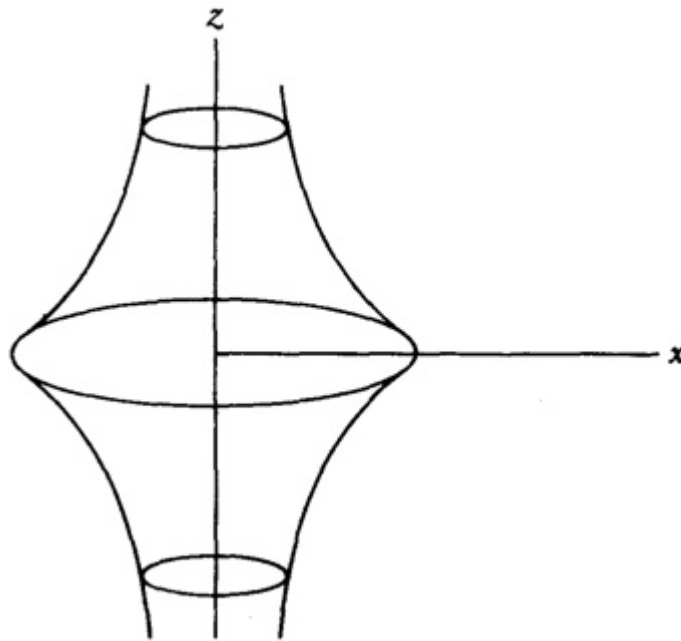


Figura 38.1

Independientemente de Riemann, Beltrami reconoció¹¹⁹ que las superficies de curvatura constante son espacios no euclídeos. También dio una representación limitada de la geometría hiperbólica sobre una superficie,¹²⁰ con la cual mostró que la geometría de una porción restringida del plano hiperbólico se cumple en una superficie de curvatura constante negativa si se toman como líneas rectas las geodésicas sobre esta superficie. Las longitudes y los ángulos sobre la superficie son las longitudes y los

ángulos de la geometría euclídea ordinaria sobre la superficie. Una de estas superficies es conocida como la pseudoesfera (fig. 38.1). Se genera al hacer rotar una curva llamada tractriz alrededor de su asíntota. La ecuación de la tractriz es

$$z = k \log \frac{k + \sqrt{k^2 - x^2}}{x} - \sqrt{k^2 - x^2}$$

y la ecuación de la superficie es

$$z = k \log \frac{\sqrt{k^2 - x^2 - y^2}}{\sqrt{x^2 + y^2}} - \sqrt{k^2 - x^2 - y^2}$$

La curvatura de la superficie es $-1/k^2$. Por tanto, la pseudoesfera es un modelo para una porción limitada del plano de Gauss, Lobachewsky y Bolyai. Sobre la pseudoesfera se puede mover una figura y con sólo curvarla se logrará que se ajuste a la superficie, al igual que una figura plana se puede ajustar a la superficie de un cilindro circular curvándola.

Beltrami había mostrado que sobre una superficie de curvatura constante negativa se puede construir un *pedazo* del plano de Lobachewsky. Sin embargo, no hay ninguna superficie analítica regular de curvatura constante negativa sobre la cual sea válida la geometría de *todo* el plano de Lobachewsky. Todas esas superficies tienen una curva singular —el plano tangente no es continuo a través de ella— de modo que una continuación de la superficie a

través de esta curva no continuará las figuras que representan las de la geometría de Lobachewsky. Este resultado se debe a Hilbert.¹²¹ En relación con lo dicho, vale la pena observar que Heinrich Liebmann (1874-1939)¹²² demostró que la esfera es la única superficie analítica cerrada (sin singularidades) de curvatura constante positiva, y por tanto es la única que puede utilizarse como un modelo para la geometría elíptica doble.

El desarrollo de estos modelos ayudó a los matemáticos a entender y ver el significado de las geometrías no euclídeas básicas. Se debe tener presente que estas geometrías, en el caso de dos dimensiones, son fundamentalmente geometrías del plano en las que las líneas y los ángulos son las líneas y los ángulos usuales de la geometría euclídea. Mientras la geometría hiperbólica había sido desarrollada en este estilo, las conclusiones aún parecían extrañas a los matemáticos y habían sido aceptadas sólo de mala gana dentro de las matemáticas. La geometría elíptica doble, sugerida por el enfoque de la geometría diferencial de Riemann, ni siquiera tuvo un desarrollo axiomático como una geometría del plano. Así, el único significado que los matemáticos podían ver para ella fue el proporcionado por la geometría sobre la esfera. Una comprensión mucho mejor de la naturaleza de estas geometrías fue asegurada por medio de otros desarrollos que buscaban relacionar la geometría proyectiva con la euclídea.

2. Las geometrías proyectiva y métrica

Aunque Poncelet había introducido la distinción entre propiedades proyectivas y métricas de las figuras y había establecido en su *Traité* de 1822 que las propiedades proyectivas eran lógicamente más fundamentales, fue von Staudt quien empezó a construir la geometría proyectiva sobre una base independiente de la longitud y la medida de ángulos (cap. 35, sec. 3). En 1853 Edmond Laguerre (1834-1886), profesor del Collège de France, aunque interesado principalmente en lo que sucede con los ángulos bajo una transformación proyectiva, de hecho amplió su meta al establecimiento de las propiedades métricas de la geometría euclídea sobre la base de conceptos proyectivos, proporcionando una base proyectiva para la medida de un ángulo.¹²³

Para obtener una medida del ángulo entre dos rectas dadas que se cortan, se pueden considerar dos líneas que pasan por el origen y paralelas respectivamente a las dos rectas dadas. Sean $y = x \tan \theta$ e $y = x \tan \theta'$ las ecuaciones de las dos rectas que pasan por el origen (en coordenadas no homogéneas). Sean $y = ix$ e $y = -ix$ las dos rectas (imaginarias) desde el origen a los puntos circulares en el infinito, es decir, los puntos $(1, i, 0)$ y $(1, -i, 0)$. Llamemos a estas cuatro rectas u, u', w y w' respectivamente. Sea Φ el ángulo entre u y u' . Entonces el resultado de Laquerre¹²⁴ es que

$$\phi = \theta' - \theta = \frac{i}{2} \log(uu', ww') \quad (1)$$

donde (uu', ww') es la razón doble de las cuatro rectas. Lo que es significativo de la expresión (1) es que puede tomarse como la

definición de la medida de un ángulo en términos del concepto proyectivo de razón doble. La función logaritmo, desde luego, es puramente cuantitativa y puede introducirse en cualquier geometría.

Independientemente de Laguerre, Cayley dio el paso siguiente. Estudió la geometría desde el punto de vista del álgebra, y de hecho estuvo interesado en la interpretación geométrica de las cuánticas (formas polinomiales homogéneas), un tema que consideraremos en el capítulo 39. Buscando mostrar que las nociones métricas pueden formularse en términos proyectivos se concentró en la relación de la geometría euclídea con la proyectiva. La obra que estamos a punto de describir se encuentra en su «*Sixth Memoir upon Quantics*» (Sexta memoria sobre cuánticas).¹²⁵

La obra de Cayley resultó ser una generalización de la idea de Laguerre. Este último había usado los puntos circulares en el infinito para definir el ángulo en el plano. Los puntos circulares son realmente una cónica degenerada. En dos dimensiones Cayley introdujo cualquier cónica en lugar de los puntos circulares y en tres dimensiones introdujo cualquier superficie cuádrlica. A estas figuras las llamó los absolutos. Cayley afirmó que todas las propiedades métricas de las figuras no son otra cosa sino las propiedades proyectivas aumentadas por el absoluto o en relación con el absoluto. Mostró entonces cómo este principio conducía a una nueva expresión para el ángulo y a una expresión para la distancia entre dos puntos.

Comienza con el hecho de que los puntos de un plano se representan mediante coordenadas homogéneas. Estas coordenadas no tienen que ser vistas como distancias o razones de distancias sino como una noción fundamental aceptada que no requiere ni admite explicación. Para definir la distancia y la medida de un ángulo introduce la forma cuadrática

$$F(x, x) = \sum_{i,j=1}^3 a_{ij} x_i x_j \quad a_{ij} = a_{ji}$$

y la forma bilineal

$$F(x, y) = \sum_{i,j=1}^3 a_{ij} x_i y_j$$

La ecuación $F(x, x) = 0$ define una cónica que es el absoluto de Cayley. La ecuación del absoluto en coordenadas de rectas es

$$G(u, u) = \sum_{i,j=1}^3 A^{ij} u_i u_j = 0$$

donde A^{ij} es el cofactor de a en el determinante $|a|$ de los coeficientes de F .

Cayley define entonces la distancia δ entre dos puntos x e y , donde $x = (x_1, x_2, x_3)$ e $y = (y_1, y_2, y_3)$, por medio de la fórmula

$$\delta = \arccos \frac{F(x,y)}{[F(x,x)F(y,y)]^{1/2}} \quad (2)$$

El ángulo Φ entre dos rectas cuyas coordenadas de rectas son $u = (u_1, u_2, u_3)$ y $v = (v_1, v_2, v_3)$ se define por

$$\cos \phi = \frac{G(u,v)}{[G(u,u)G(v,v)]} \quad (3)$$

Estas fórmulas generales se hacen más sencillas si tomamos como absoluto la cónica particular $x_1^2 + x_2^2 + x_3^2 = 0$. Entonces, si (a_1, a_2, a_3) y (b_1, b_2, b_3) son las coordenadas homogéneas de dos puntos, la distancia entre ellos está dada por

$$\arccos \frac{a_1 b_1 + a_2 b_2 + a_3 b_3}{\sqrt{a_1^2 + a_2^2 + a_3^2} + \sqrt{b_1^2 + b_2^2 + b_3^2}} \quad (4)$$

y el ángulo Φ entre dos rectas cuyas coordenadas homogéneas de rectas son (u_1, u_2, u_3) y (v_1, v_2, v_3) está dado por

$$\cos \phi = \frac{u_1 v_1 + u_2 v_2 + u_3 v_3}{\sqrt{u_1^2 + u_2^2 + u_3^2} + \sqrt{v_1^2 + v_2^2 + v_3^2}} \quad (5)$$

Con respecto a la expresión para la distancia, si usamos la simplificación de que $xy = x_1 y_1 + x_2 y_2 + x_3 y_3$ entonces, si $a = (a_1, a_2, a_3)$, b y c son tres puntos sobre una recta,

$$\arccos \frac{ab}{\sqrt{aa}\sqrt{bb}} + \arccos \frac{bc}{\sqrt{bb}\sqrt{cc}} = \arccos \frac{ac}{\sqrt{aa}\sqrt{cc}}$$

Esto es, las distancias se suman como se debe. Tomando como cónica absoluto los puntos circulares en el infinito $(1, i, 0)$ y $(1, -i, 0)$, Cayley demostró que sus fórmulas para la distancia y el ángulo se reducen a las fórmulas euclídeas usuales.

Se observará que las expresiones para la longitud y el ángulo involucran la expresión algebraica para el absoluto. Generalmente la expresión analítica para cualquier propiedad métrica euclídea involucra la relación de esa propiedad con el absoluto. Las propiedades métricas no son propiedades de la figura *per se*, sino de la figura en relación con el absoluto. Esta es la idea de Cayley de la determinación proyectiva general de la métrica. El lugar del concepto métrico en la geometría proyectiva y la generalidad mayor de la última fueron descritos por Cayley como «*La geometría métrica es parte de la geometría proyectiva*».

La idea de Cayley fue retomada por Félix Klein (1849-1925) y generalizada de modo que incluyera las geometrías no euclídeas.

Klein, profesor de Göttingen, fue una de las figuras de las matemáticas en Alemania durante la última parte del siglo XIX y la primera parte del XX. Durante los años 1869-1870 estudió el trabajo de Lobachewsky, Bolyai, von Staudt y Cayley; sin embargo, aún en 1871 no conocía el resultado de Laguerre. A él le parecía que era posible subsumir las geometrías no euclídeas, la geometría

elíptica doble y la hiperbólica, en la geometría proyectiva explotando la idea de Cayley. Proporcionó un bosquejo de sus ideas en un artículo de 1871¹²⁶ y después los desarrolló en otros dos artículos.¹²⁷ Klein fue el primero en reconocer que no necesitamos superficies para obtener modelos de las geometrías no euclídeas.

Para empezar, Klein observó que Cayley no dejó claro exactamente lo que consideraba como significado de sus coordenadas. Eran simplemente variables sin interpretación geométrica alguna o eran distancias euclídeas. Pero para obtener las geometrías métricas a partir de la geometría proyectiva era necesario construir las coordenadas sobre una base proyectiva. Von Staudt había demostrado (cap. 35, sec. 3) que era posible asignar números a los puntos por medio de su álgebra de lanzamientos. Pero usó el axioma euclídeo del paralelismo. A Klein le parecía claro que se podía prescindir de este axioma y en el artículo de 1873 muestra que esto puede hacerse. Por tanto, las coordenadas y la razón doble de cuatro puntos, cuatro rectas o cuatro planos puede definirse sobre una base puramente proyectiva.

La idea más importante de Klein fue que por medio de la particularización de la naturaleza de la superficie cuádrlica absoluto de Cayley (si se considera la geometría tridimensional), podría mostrarse que la métrica, la cual según Cayley dependía de la naturaleza del absoluto, conduciría a la geometría hiperbólica y a la elíptica doble. Cuando la superficie de segundo grado es un elipsoide real, un paraboloides elíptico real o un hiperboloides real de dos hojas, se obtiene la geometría métrica de Lobachewsky. Cuando

la superficie de segundo grado es imaginaria, se obtiene la geometría no euclídea de Riemann (de curvatura constante positiva). Si el absoluto es la esfera-círculo, cuya ecuación en coordenadas homogéneas es $x^2 + y^2 + z^2 = 0$, $t = 0$, entonces se obtiene la geometría métrica euclídea usual. Así las geometrías métricas se convierten en casos especiales de la geometría proyectiva.

Para apreciar las ideas de Klein consideremos la geometría de dos dimensiones. Se escoge una cónica en el plano proyectivo; esta cónica será el absoluto. Su ecuación es

$$F = \sum_{i,j=1}^3 a_{ij}x_iy_j = 0 \quad (6)$$

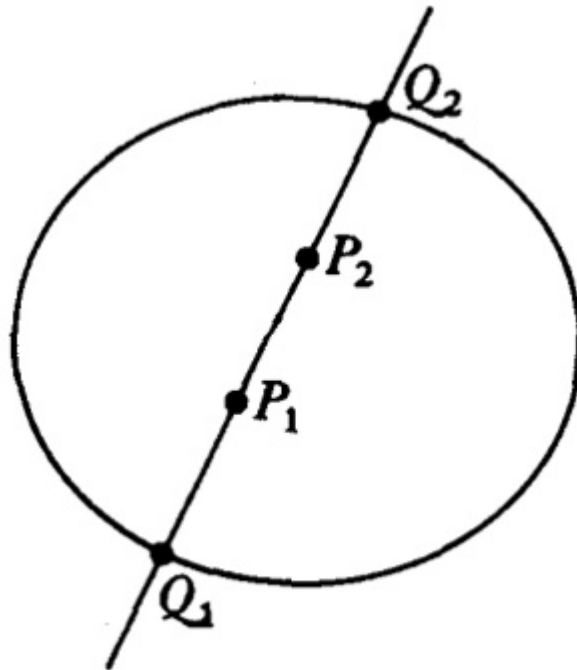


Figura 38.2

en coordenadas puntuales y

$$G = \sum_{i,j=1}^3 A^{ij} u_i u_j = 0 \quad (7)$$

en coordenadas de rectas. Para obtener la geometría de Lobachewsky, la cónica debe ser real, por ejemplo, en coordenadas homogéneas planas $x_1^2 + x_2^2 - x_3^2 = 0$; para la geometría de Riemann sobre una superficie de curvatura constante positiva es imaginaria, por ejemplo, $x_1^2 + x_2^2 - x_3^2 = 0$ y para la geometría euclídea, la cónica degenera en dos rectas que coinciden, representadas en coordenadas homogéneas por $x_1^2 + x_2^2 - x_3^2 = 0$ y sobre este lugar geométrico se escogen dos puntos imaginarios cuya ecuación es $x^2 + x^2 = 0$, esto es, los puntos circulares del infinito cuyas coordenadas homogéneas son $(1, i, 0)$ y $(1, -i, 0)$. En todos los casos la cónica tiene una ecuación real.

Para ser precisos supongamos que la cónica es la que se muestra en la figura 38.2. Si P_1 y P_2 son dos puntos de una recta, esta recta corta al absoluto en dos puntos (reales o imaginarios). Entonces la distancia se toma como

$$d = c \log (P_1 P_2, Q_1 Q_2) \quad (8)$$

donde la cantidad entre paréntesis denota la razón doble de los cuatro puntos y c es una constante. Esta razón doble puede expresarse en términos de las coordenadas de los puntos. Es más,

si hay tres puntos P_1, P_2, P_3 en la recta entonces puede mostrarse fácilmente que

$$(P_1P_2, Q_1Q_2) (P_2P_3, Q_1Q_2) = (P_1P_3, Q_1Q_2)$$

de modo que $P_1P_2 + P_2P_3 = P_1P_3$.

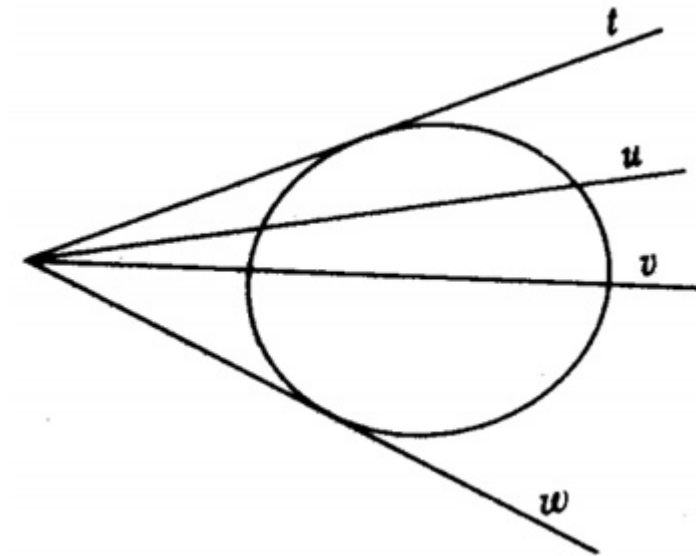


Figura 38.3

Análogamente, si u y v son dos rectas (fig. 38.3) se consideran las tangentes t y w desde su punto de intersección al absoluto (las tangentes pueden ser rectas imaginarias); entonces el ángulo entre u y v se define como

$$\Phi' = c' \log (uv, tw),$$

donde nuevamente c' es una constante y la cantidad entre paréntesis denota la razón doble de las cuatro rectas.

Para expresar los valores de d y Φ analíticamente y demostrar su dependencia de la elección del absoluto, sea dada la ecuación del absoluto por las F y G anteriores. Por definición

$$F_{xy} = \sum_{i,j=1}^3 a_{ij}x_i y_j = 0$$

Ahora puede demostrarse que si $x = (x_1, x_2, x_3)$ e $y = (y_1, y_2, y_3)$ son las coordenadas de P_1 y P_2 entonces

$$d = c \log \frac{F_{xy} + \sqrt{F_{xy}^2 - F_{xx}F_{yy}}}{F_{xy} - \sqrt{F_{xy}^2 - F_{xx}F_{yy}}}$$

Análogamente, si (u_1, u_2, u_3) y (v_1, v_2, v_3) son las coordenadas de las dos rectas, entonces puede demostrarse, usando G , que

$$\phi = c' \log \frac{G_{uv} + \sqrt{G_{uv}^2 - G_{uu}G_{vv}}}{G_{uv} - \sqrt{G_{uv}^2 - G_{uu}G_{vv}}}$$

La constante c' se toma generalmente como $i/2$, de modo que Φ sea real y un ángulo central completo mida 2π .

Klein usó las expresiones logarítmicas anteriores para el ángulo y la distancia y demostró cómo pueden obtenerse las geometrías

métricas a partir de la geometría proyectiva. Es decir, que si se parte de la geometría proyectiva, entonces por la elección del absoluto y usando las expresiones anteriores para la distancia y el ángulo pueden obtenerse las geometrías euclídea, hiperbólica y elíptica como casos especiales. La naturaleza de la geometría métrica queda determinada por la elección del absoluto. Incidentalmente, puede mostrarse que las expresiones de Klein para la distancia y el ángulo son iguales a las de Cayley.

Si se realiza una transformación proyectiva (ie., lineal) del plano proyectivo en sí mismo, la cual transforma el absoluto en sí mismo (aunque los puntos en el absoluto vayan a otros puntos), entonces, debido a que la razón doble queda inalterada por una transformación lineal, las distancias y los ángulos tampoco se alteran. Estas transformaciones lineales particulares que dejan fijo el absoluto son los movimientos rígidos o transformaciones de congruencia de la geometría métrica particular determinada por el absoluto. Una transformación proyectiva general no dejará invariante el absoluto. Así la propia geometría proyectiva es realmente más general por las transformaciones que permite.

Otra contribución de Klein a la geometría no euclídea fue la observación, que él dice hizo primero en 1871¹²⁸, pero que publicó en 1874¹²⁹, de que hay dos clases de geometría elíptica. En la geometría elíptica doble dos puntos no siempre determinan una línea recta única. Esto es evidente a partir del modelo esférico cuando los dos puntos son diametralmente opuestos. En la segunda geometría elíptica, llamada elíptica simple, dos puntos siempre

determinan una línea recta única. Cuando la investigó desde el punto de vista de la geometría diferencial, la forma diferencial ds^2 de una superficie de curvatura constante positiva es (en coordenadas homogéneas)

$$ds^2 = \frac{dx_1^2 + dx_2^2 + dx_3^2}{\{1 + (a^2/4)(x_1^2 + x_2^2 + x_3^2)\}^2}$$

En ambos casos $a^2 > 0$. Sin embargo, en el primer tipo las geodésicas son curvas de longitud finita $2\pi/a$ o, si R es el radio, $2\pi R$, y son cerradas (regresan sobre sí mismas). En el segundo las geodésicas son de longitud π/a o πR y también son cerradas.

Un modelo de una superficie que tiene las propiedades de la geometría elíptica simple, y que se le debe a Klein,¹³⁰ es proporcionado por un hemisferio que incluye la frontera. Sin embargo, deben identificarse dos puntos cualesquiera sobre la frontera que sean diametralmente opuestos. Los arcos circulares máximos en la semiesfera son las «líneas rectas» o geodésicas de esta geometría y los ángulos ordinarios en la superficie son los ángulos de la geometría. La geometría elíptica simple (a veces llamada elíptica, en cuyo caso la geometría elíptica doble es llamada esférica) también es realizada entonces en un espacio de curvatura constante positiva. Realmente no pueden unirse los pares de puntos que son identificados en este modelo, al menos en el espacio tridimensional. La superficie tendría que cruzarse a sí misma y los

puntos que coincidieran en la intersección tendrían que ser considerados distintos.

Ahora podemos ver por qué Klein introdujo la terminología de hiperbólica para la geometría de Lobachewsky, elíptica para el caso de la geometría de Riemann sobre una superficie de curvatura constante positiva, y parabólica para la geometría euclídea. Esta terminología fue sugerida por el hecho de que la hipérbola ordinaria interseca la línea del infinito en dos puntos y correspondientemente en la geometría hiperbólica cada línea interseca al absoluto en dos puntos reales. La elipse ordinaria no tiene puntos reales en común con la recta del infinito y en la geometría elíptica, de modo parecido, toda recta carece de puntos reales en común con el absoluto. La parábola ordinaria solamente tiene un punto real en común con la recta del infinito y en la geometría euclídea (como es extendida en la geometría proyectiva) cada recta tiene un punto real en común con el absoluto.

El sentido que gradualmente surgió de las contribuciones de Klein fue que la geometría proyectiva es en realidad lógicamente independiente de la geometría euclídea. Más aún, las geometrías no euclídeas y euclídea también fueron vistas como casos especiales o subgeometrías de la geometría proyectiva. De hecho, la obra estrictamente lógica o rigurosa sobre los fundamentos axiomáticos de la geometría proyectiva y sus relaciones con las subgeometrías quedaba por realizarse (cap. 42). Pero haciendo ver claro el papel básico de la geometría proyectiva, Klein abrió el camino a un

desarrollo axiomático que podría comenzar con la geometría proyectiva y obtener de ella las varias geometrías métricas.

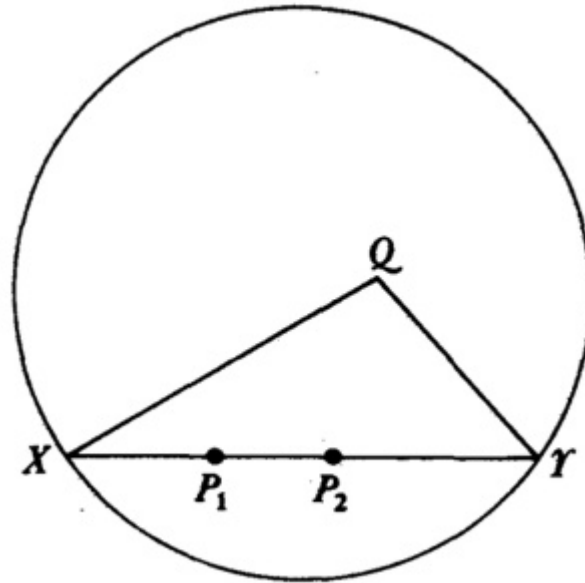
4. Los modelos y el problema de la consistencia

Al principio de los 1870 se habían introducido y estudiado intensivamente varias geometrías no euclídeas básicas, la hiperbólica y las dos geometrías elípticas. La cuestión fundamental que aún quedaba por resolver para lograr que estas geometrías fueran ramas legítimas de las matemáticas consistía en determinar si eran consistentes. Toda la obra realizada por Gauss, Lobachewsky, Bolyai, Riemann, Cayley y Klein podría haber resultado un sinsentido si hubiera contradicciones inherentes a estas geometrías.

De hecho, la prueba de la consistencia de la geometría elíptica doble en dos dimensiones estaba a mano, y posiblemente Riemann apreció este hecho, aunque no hizo afirmación explícita alguna. Beltrami¹³¹ había señalado que la geometría de Riemann en dos dimensiones con curvatura constante positiva se construye sobre una esfera. Este modelo hace posible la demostración de la consistencia de la geometría elíptica doble en dos dimensiones. Los axiomas (que no eran explícitos en esa época) y los teoremas de esta geometría son todos aplicables a la geometría de la superficie de la esfera siempre que se interprete la recta en la geometría elíptica doble como el círculo máximo sobre la superficie de la esfera. Si hubiera teoremas contradictorios en esta geometría elíptica doble entonces habría teoremas contradictorios en la geometría de la

superficie de la esfera. Ahora bien, la esfera es parte de la geometría euclídea. Por tanto, si la geometría euclídea es consistente, entonces la geometría elíptica doble también debe serlo. Para los matemáticos de los años 1870 la consistencia de la geometría euclídea difícilmente estaba sujeta a dudas porque, aparte de los puntos de vista de unos pocos autores como Gauss, Bolyai, Lobachewsky y Riemann, la geometría euclídea era todavía la geometría necesaria del mundo físico y era inconcebible que pudiera haber propiedades contradictorias en la geometría del mundo físico. Sin embargo, es importante, especialmente a la luz de desarrollos posteriores, darse cuenta de que esta demostración de la consistencia de la geometría elíptica doble depende de la consistencia de la geometría euclídea.

El método para demostrar la consistencia de la geometría elíptica doble no podía utilizarse para la geometría elíptica simple o para la geometría hiperbólica. El modelo hemisférico para la geometría elíptica simple no puede ser llevado a cabo en la geometría euclídea tridimensional aunque sí en la geometría euclídea de cuatro dimensiones. Si uno estuviera dispuesto a creer en la consistencia de esta última, entonces podría aceptar la consistencia de la geometría elíptica simple. Sin embargo, aunque la geometría «-dimensional ya había sido considerada por Grassmann, Riemann y otros, es dudoso que cualquier matemático de los años 1870 hubiese estado dispuesto a afirmar la consistencia de la geometría euclídea de cuatro dimensiones.

*Figura 38.4*

El caso de la consistencia de la geometría hiperbólica no podría llevarse a cabo sobre alguna de estas bases. Beltrami había dado la interpretación de la pseudoesfera, que es una superficie en el espacio euclídeo, pero esto sirve como modelo sólo para una región limitada de la geometría hiperbólica y, por tanto, no pudo usarse para establecer la consistencia de toda la geometría. Lobachewsky y Bolyai habían considerado este problema (cap. 36, sec. 8), pero no pudieron aclararlo. El caso es que aunque Bolyai publicó orgullosamente su geometría no euclídea, hay pruebas de que dudó de su consistencia, ya que en artículos que se encontraron después de su muerte, continuó tratando de demostrar el axioma euclídeo del paralelismo.

La consistencia de las geometrías hiperbólica y elíptica simple fue establecida por medio de nuevos modelos. El modelo para la geometría hiperbólica es debido a Beltrami.¹³² Sin embargo, la

función distancia que se usa en este modelo es debida a Klein y a él se adscribe frecuentemente el modelo. Consideremos el caso bidimensional.

Dentro del plano euclídeo (que es parte del plano proyectivo) se elige una cónica real que, indistintamente, puede ser el círculo (fig. 38.4). De acuerdo con esta representación de la geometría hiperbólica, los puntos de la geometría son los puntos interiores de este círculo. Una recta de esta geometría es una cuerda del círculo, digamos la cuerda XY (sin incluir a X e Y). Si tomamos cualquier punto Q que no esté en XY entonces podemos encontrar cualquier número de rectas que pasen por Q y que no corten a XY . Dos de estas rectas, digamos, QX y QY , separan las rectas que pasan por Q en dos clases, aquellas que cortan a XY y a las que no la cortan. En otras palabras, el axioma del paralelismo de la geometría hiperbólica se satisface por los puntos y rectas (cuerdas) interiores del círculo. Además, sea

$$\angle(a, b) = \frac{1}{2i} \log(ab, mn)$$

la medida del ángulo formado por dos rectas a y b , donde m y n son las tangentes imaginarias conjugadas desde el vértice del ángulo al círculo y (ab, mn) es la razón doble de las cuatro rectas a , b , m y n . La constante $1/2i$ asegura que un ángulo recto tenga la medida $\pi/2$. La definición de la distancia entre dos puntos está dada por la fórmula (8), esto es, $d = c \log (PP', XY)$ tomándose generalmente c como $k/2$. De acuerdo con esta fórmula, conforme P o P' se acercan

a X o Y , la distancia PP' se hace infinita. De este modo, en términos de esta distancia, una cuerda es una recta infinita de la geometría hiperbólica.

Así, con las definiciones proyectivas de distancia y medida de ángulos, los puntos, cuerdas, ángulos y otras figuras en el interior del círculo satisfacen los axiomas de la geometría hiperbólica. Entonces los teoremas de la geometría hiperbólica se aplican también a estas figuras dentro del círculo. En este modelo los axiomas y teoremas de la geometría hiperbólica son realmente afirmaciones acerca de figuras y conceptos especiales (e.g. la distancia definida a la manera de la geometría hiperbólica) de la geometría euclídea. Como los axiomas y teoremas en cuestión se aplican a estas figuras y conceptos, vistos como pertenecientes a la geometría euclídea, todas las afirmaciones de la geometría hiperbólica son teoremas de la geometría euclídea. Entonces, si hubiese alguna contradicción en la geometría hiperbólica, esta contradicción sería una contradicción dentro de la geometría euclídea. Pero si la geometría euclídea es consistente, entonces la geometría hiperbólica también debe serlo. Así la consistencia de la geometría hiperbólica se reduce a la consistencia de la geometría euclídea.

El hecho de que la geometría hiperbólica sea consistente implica que el axioma euclídeo del paralelismo es independiente de los otros axiomas euclídeos. Si este no fuera el caso, esto es, si el axioma euclídeo del paralelismo se pudiera obtener de los otros axiomas, sería también un teorema de la geometría hiperbólica, pues además

del axioma del paralelismo, los otros axiomas de la geometría euclídea son los mismos que los de la geometría hiperbólica. Pero este teorema contradiría el axioma del paralelismo de la geometría hiperbólica y entonces ésta sería inconsistente. La consistencia de la geometría elíptica simple en dos dimensiones puede demostrarse de la misma manera que para la geometría hiperbólica, porque esta geometría elíptica también se realiza dentro del plano proyectivo y con la definición proyectiva de distancia.

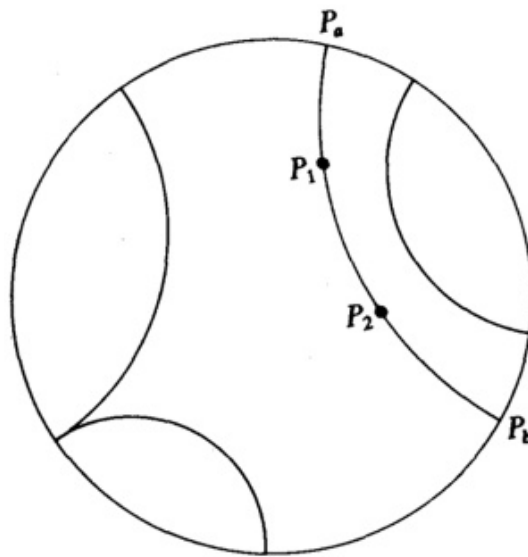


Figura 38.5

Independientemente y en relación con su trabajo sobre funciones automorfas, Poincaré ¹³³ dio otro modelo que también establece la consistencia de la geometría hiperbólica. Una forma en la cual este modelo de Poincaré para la geometría hiperbólica plana ¹³⁴ puede expresarse toma el absoluto como un círculo (fig. 38.5). Dentro del absoluto las líneas rectas de la geometría son arcos de círculo que

cortan al absoluto ortogonalmente y líneas rectas que pasan por el centro del absoluto. La longitud de cualquier segmento P_1P_2 está dada por $\log (P_1P_2, P_aP_b)$, donde $(P_1P_2, P_aP_b) = (P_1P_bP_2P_b)/(P_1P_a, P_2P_b)$, P_a y P_b son los puntos en los cuales el arco que pasa por P_1 y P_2 corta al absoluto, y las longitudes P_1P_b , P_2P_b etc. son las cuerdas. El ángulo entre dos «rectas» de este modelo que se intersecan es el ángulo euclídeo normal entre los dos arcos. Dos arcos circulares que sean tangentes en un punto en el absoluto son «rectas» paralelas. Puesto que en este modelo los axiomas y teoremas de la geometría hiperbólica también son teoremas especiales de la geometría euclídea, el argumento dado antes a propósito del modelo de Beltrami puede aplicarse aquí para establecer la consistencia de la geometría hiperbólica. Los análogos en dimensiones mayores de los modelos anteriores son también válidos.

5. La geometría desde el punto de vista de las transformaciones

El éxito de Klein en subsumir las varias geometrías métricas en la geometría proyectiva lo condujo a buscar el caracterizar las varias geometrías no solamente sobre la base de las propiedades no métricas y métricas y las distinciones entre las métricas, sino desde el punto de vista más amplio de lo que estas geometrías, y otras geometrías que ya habían aparecido en escena, intentaban lograr. Proporcionó esta caracterización en un discurso de 1872, «*Vergleichende Betrachtungen über neuere geometrische Forschungen*» (Una Revisión Comparativa de las Investigaciones Recientes en Geometría),¹³⁵ con ocasión de su admisión en la

facultad de la universidad de Erlangen, y las opiniones allí expresadas se conocen como el *Erlanger Programm*.

La idea básica de Klein es que cada geometría puede ser caracterizada por un grupo de transformaciones y que una geometría trata realmente de los invariantes por este grupo de transformaciones. Más aún, una subgeometría de una geometría es la colección de los invariantes bajo un subgrupo de transformaciones del grupo original. Con esta definición todos los teoremas de una geometría correspondientes a un grupo dado continúan siendo teoremas en la geometría del subgrupo.

Aunque Klein no da en su artículo las formulaciones analíticas de los grupos de transformaciones que discute, daremos algunas para ser explícitos. De acuerdo con su noción de una geometría, la geometría proyectiva, digamos en dos dimensiones, es el estudio de los invariantes bajo el grupo de transformaciones desde los puntos de un plano a los de otro plano, o a los del mismo plano (colineaciones). Cada transformación es de la forma

$$\begin{aligned}x'_1 &= a_{11}x_1 + a_{12}x_2 + a_{13}x_3 \\x'_2 &= a_{21}x_1 + a_{22}x_2 + a_{23}x_3 \\x'_3 &= a_{31}x_1 + a_{32}x_2 + a_{33}x_3\end{aligned}\tag{9}$$

en donde se presuponen las coordenadas homogéneas, los a_{ij} son números reales y el determinante de los coeficientes debe ser

diferente de cero. En coordenadas no homogéneas las transformaciones se representan por

$$\begin{aligned}x' &= \frac{a_{11}x + a_{12}y + a_{13}}{a_{31}x + a_{32}y + a_{33}} \\y' &= \frac{a_{21}x + a_{22}y + a_{23}}{a_{31}x + a_{32}y + a_{33}}\end{aligned}\tag{10}$$

y nuevamente el determinante de los a_{ij} debe ser diferente de cero. Los invariantes bajo el grupo proyectivo son, por ejemplo, linealidad, colinealidad, razón doble, conjuntos armónicos y la propiedad de ser una sección cónica.

Un subgrupo del grupo proyectivo es la colección de las transformaciones afines.¹³⁶ Este subgrupo se define como sigue: fijese una recta l_∞ cualquiera en el plano proyectivo. Los puntos de l_∞ son llamados puntos ideales o puntos del infinito y l_∞ es llamada la recta del infinito. Los otros puntos y rectas del plano proyectivo son llamados puntos ordinarios y estos son los puntos usuales del plano euclídeo. El grupo afin de colineaciones es el subgrupo del grupo proyectivo que deja invariante l_∞ (aunque no necesariamente punto a punto) y la geometría afin es el conjunto de propiedades y relaciones invariantes por el grupo afin. Algebraicamente, en dos dimensiones y en coordenadas homogéneas, las transformaciones afines se representan por las ecuaciones (9) anteriores, pero en las cuales $a_{31} = a_{32} = 0$, y con la misma condición para el determinante.

En coordenadas no homogéneas las transformaciones afines están dadas por

$$\begin{aligned} x' &= a_{11}x + a_{12}y + a_{13} \\ y' &= a_{21}x + a_{22}y + a_{23} \end{aligned} \quad \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} \neq 0$$

Bajo una transformación afín, las líneas rectas van a líneas rectas y las líneas rectas paralelas a rectas paralelas. Sin embargo, las longitudes y medidas de ángulos se alteran. La geometría afín fue observada primero por Euler y después por Möbius en su *Der barycentrische Calcul*. Es útil en el estudio de la mecánica de las deformaciones.

El grupo de cualquier geometría métrica es el mismo que el grupo afín, excepto que el determinante anterior debe tener el valor +1 ó -1. La primera de las geometrías métricas es la geometría euclídea. Para definir al grupo de esta geometría se empieza con l_∞ y se supone que hay una involución fija sobre l_∞ . Se requiere que esta involución no tenga puntos reales dobles sino que tenga a los puntos circulares en el ∞ como puntos dobles (imaginarios). Luego consideramos todas las transformaciones proyectivas que no solamente dejan fija la l_∞ sino que llevan cualquier punto de la involución al punto correspondiente de ella, lo cual implica que cada punto circular va sobre sí mismo. Algebraicamente estas transformaciones del grupo euclídeo se representan en coordenadas no homogéneas (bidimensionales) por

$$x' = \rho(x \cos \theta - y \sin \theta + a) \quad \rho = \pm 1$$

$$y' = \rho(x \sin \theta + y \cos \theta + \beta)$$

Los invariantes son longitud, medida de ángulos y medida y forma de cualquier figura.

La geometría euclídea, según se usa el término en esta clasificación, es el conjunto de invariantes bajo esta clase de transformaciones. Las transformaciones son rotaciones, traslaciones y reflexiones. Para obtener los invariantes asociados con figuras semejantes, se introduce el subgrupo del grupo afin conocido como el grupo métrico parabólico. Este grupo se define como la clase de las transformaciones proyectivas que dejan invariantes la involución sobre l_∞ , y esto significa que cada par de puntos correspondientes va a algún par de puntos correspondientes. En coordenadas no homogéneas las transformaciones del grupo métrico parabólico son de la forma

$$x' = ax - by + c$$

$$y' = bex + aey + d$$

donde $a^2 + b^2 = 0$, y $e^2 = 1$. Estas transformaciones conservan la medida de los ángulos.

Para caracterizar la geometría métrica hiperbólica regresamos a la geometría proyectiva y consideramos una cónica no degenerada real

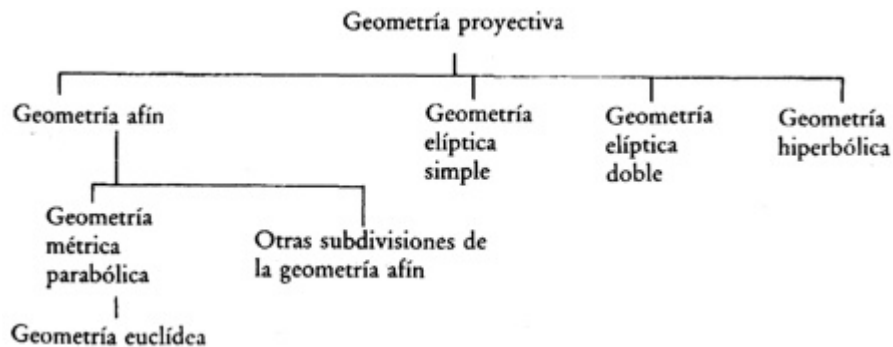
arbitraria (el absoluto) en el plano proyectivo. El subgrupo del grupo proyectivo que deja esta cónica invariante (aunque no necesariamente punto a punto) es llamado grupo métrico hiperbólico y la geometría correspondiente es la geometría métrica hiperbólica. Los invariantes son aquellos que están asociados con la congruencia.

La geometría elíptica simple es la geometría correspondiente al subgrupo de las transformaciones proyectivas que deja invariante una elipse imaginaria dada (el absoluto) en el plano proyectivo. El plano de la geometría elíptica es el plano proyectivo real y los invariantes son aquellos asociados con la congruencia.

Incluso la geometría elíptica doble puede abarcarse desde este punto de vista de las transformaciones, pero se debe comenzar con las transformaciones proyectivas en tres dimensiones para caracterizar la geometría métrica bidimensional. El subgrupo de transformaciones consiste en aquellas transformaciones proyectivas tridimensionales que transforman una esfera dada (superficie) S de la porción finita del espacio en sí misma. La superficie esférica S es el «plano» de la geometría elíptica doble. Nuevamente los invariantes son aquellos asociados con la congruencia.

En cuatro geometrías métricas, esto es, la euclídea, la hiperbólica y las dos geometrías elípticas, las transformaciones que son permitidas en el correspondiente subgrupo son las que generalmente se llaman movimientos rígidos y estas son las únicas geometrías que permiten movimientos rígidos.

Klein introdujo varias clasificaciones intermedias que no repetiremos aquí. El esquema que sigue muestra las relaciones de las principales geometrías.



Klein también consideró geometrías más generales que la proyectiva. En esta época (1872) se estaba distinguiendo gradualmente la geometría algebraica como una disciplina separada y caracterizó a esta geometría introduciendo las transformaciones que en tres dimensiones \mathbf{y} en coordenadas no homogéneas se escriben como

$$x' = \Phi(x, y, z) \quad y' = \psi(x, y, z) \quad z' = \chi(x, y, z).$$

Requirió que las funciones Φ , ψ , χ fueran racionales y univaluadas y que debería ser posible resolver las ecuaciones para x , y y z en términos de funciones racionales univaluadas de x' , y' y z' . Estas transformaciones son llamadas transformaciones de Cremona y los invariantes por ellas son el objeto de estudio de la geometría algebraica (cap. 39).

Klein también proyectó el estudio de los invariantes bajo transformaciones continuas uno-a-uno con inversas continuas. Esta es la clase ahora llamada de los homeomorfismos y el estudio de los invariantes bajo tales transformaciones es el objeto de estudio de la topología (cap. 50). Aunque Riemann también había considerado lo que ahora se reconoce que son problemas topológicos en su trabajo con las superficies de Riemann, la previsión de la topología como una geometría importante fue un paso audaz en 1872.

Desde los días de Klein, ha sido posible llevar a cabo adiciones y especializaciones más amplias en la clasificación de Klein. Pero no toda la geometría puede ser incorporada en el esquema de Klein. Hoy, la geometría algebraica y la geometría diferencial no caen dentro de este esquema.¹³⁷ A pesar de que el enfoque de Klein de la geometría no demostró abarcar todo, sí proporcionó un método sistemático de clasificación y estudio de una buena parte de la geometría y sugirió numerosos problemas de investigación. Su «definición» de geometría guió el pensamiento geométrico durante aproximadamente cincuenta años. Más aún, su énfasis en los invariantes por las transformaciones fue más allá de las matemáticas, generalmente a la mecánica y a la física matemática. El problema físico de la invariancia bajo transformaciones o el problema de expresar las leyes físicas de una manera independiente del sistema de coordenadas fue de mucha importancia en el pensamiento físico después de que se observó la invariancia de las ecuaciones de Maxwell bajo las transformaciones de Lorentz (un

subgrupo de cuatro dimensiones de la geometría afín). Esta línea de pensamiento condujo hacia la teoría especial de la relatividad.

Mencionaremos aquí meramente estudios más avanzados sobre la clasificación de las geometrías que, al menos en su época, atrajeron considerable atención. Helmholtz y Sophus Lie (1842-1899) buscaron caracterizar las geometrías en las que son posibles los movimientos rígidos. El artículo básico de Helmholtz, «*Über die Thatsachen, die der Geometrie zum Grunde liegen*» (Sobre los hechos que subyacen a la Geometría),¹³⁸ mostró que si los movimientos de los cuerpos rígidos tienen que ser posibles en un espacio, entonces la expresión de Riemann para ds en un espacio de curvatura constante es la única posible. Lie atacó el mismo problema por medio de lo que se conoce como la teoría de los grupos de transformaciones continuas, una teoría que él ya había introducido en el estudio de las ecuaciones diferenciales ordinarias, y caracterizó los espacios en los que los movimientos rígidos son posibles por medio de los tipos de grupos de transformaciones que permiten estos espacios.¹³⁹

6. La realidad de la geometría no euclídea

El interés en las geometrías no euclídeas sintéticas clásicas y en la geometría proyectiva declinó después de la obra de Klein y Lie, en parte debido a que la esencia de estas estructuras fue expuesta de modo muy claro desde el punto de vista de las transformaciones. El sentimiento de los matemáticos en cuanto concierne al descubrimiento de teoremas adicionales era que la mina había sido

agotada. La rigorización de los fundamentos quedó por hacer y esta fue un área activa durante muchos años después de 1880 (cap. 42). Otra razón de la pérdida de interés en las geometrías no euclídeas fue su aparente falta de relevancia para el mundo físico. Es curioso que los primeros que se dedicaron a este campo, Gauss, Lobachewsky y Bolyai sí pensaban que la geometría no euclídea podría ser aplicable cuando se hubiese avanzado más en el dominio de la astronomía. Pero ninguno de los matemáticos que trabajaron en el período siguiente creyó que estas geometrías no euclídeas básicas tendrían trascendencia física. Cayley, Klein y Poincaré, aunque consideraron este asunto, afirmaron que nunca necesitaríamos mejorar o abandonar la geometría euclídea. El modelo de la pseudoesfera de Beltrami había hecho real a la geometría no euclídea en un sentido matemático (aunque no físico) porque daba una interpretación fácilmente visualizable de la geometría de Lobachewsky, pero al precio de cambiar la recta del borde de la regla por las geodésicas y la pseudoesfera. Análogamente los modelos de Beltrami-Klein y Poincaré dieron sentido a la geometría no euclídea cambiando los conceptos ya sea de recta, distancia o medida de ángulos, o los tres al mismo tiempo, y representándolos en el espacio euclídeo. Pero el pensamiento de que el espacio físico podría ser no euclídeo bajo la interpretación usual de línea recta o bajo alguna otra interpretación fue rechazado. De hecho, la mayoría de los matemáticos trataban a la geometría no euclídea como una curiosidad lógica.

Cayley fue un firme partidario del espacio euclídeo y aceptó las geometrías no euclídeas solamente en cuanto que podían ser realizadas en el espacio euclídeo por medio de nuevas fórmulas para la distancia. En 1883, en su discurso presidencial a la «British Association for the Advancement of Science» (Asociación Británica para el Progreso de la Ciencia),¹⁴⁰ dijo que los espacios no euclídeos eran a priori una idea equivocada, pero que las geometrías no euclídeas eran aceptables porque resultaban meramente de un cambio en la función distancia del espacio euclídeo. No garantizó la existencia independiente de las geometrías no euclidianas sino que las trató como una clase de estructuras euclídeas especiales, o como una manera de representar las relaciones proyectivas en la geometría euclídea. Su opinión era que:

El duodécimo [décimo] axioma de Euclides en la forma que dio Playfair no necesita demostración, sino que es parte de nuestra noción de espacio, del espacio físico de nuestra propia experiencia —esto es, el espacio con el que llegamos a conocer por medio de la experiencia, pero que es la representación subyacente en el fundamento de toda experiencia exterior.

Se puede decir que la opinión de Riemann es que, teniendo in intellectu una noción más general de espacio (de hecho una noción de espacio no euclídeo), aprendemos por medio de la experiencia que el espacio (el espacio físico de la experiencia) es, si no exactamente, al menos en el más alto grado de aproximación, espacio euclídeo.

Klein consideró el espacio euclídeo como el espacio fundamental necesario. Las otras geometrías eran meramente euclídeas con las nuevas funciones para las distancias. Las geometrías no euclídeas estaban en efecto subordinadas a la geometría euclídea.

El juicio de Poincaré fue más liberal. La ciencia siempre debería tratar de usar la geometría euclídea y variar las leyes de la física donde fuese necesario. La geometría euclídea puede no ser verdadera pero es la más conveniente. Una geometría no puede ser más verdadera que otra; solamente puede ser más conveniente. El hombre crea la geometría y después adapta las leyes físicas a ella para hacer que la geometría y las leyes físicas coincidan con el mundo. Poincaré insistió ¹⁴¹ en que, aun si la suma de los ángulos de un triángulo resultara ser mayor de 180° , sería mejor suponer que la geometría euclídea describe el espacio físico y que la luz viaja siguiendo curvas porque la geometría euclídea es más sencilla. Por supuesto que los acontecimientos probaron que estaba equivocado. No es sólo la simplicidad de la geometría lo que cuenta para la ciencia sino la simplicidad de la teoría científica entera. Claramente, los científicos del siglo XIX estaban aún atados a la tradición por sus nociones acerca de lo que tiene sentido físico. El advenimiento de la teoría de la relatividad forzó un cambio drástico en la actitud hacia la geometría no euclídea.

La ilusión de los matemáticos en el sentido de que aquello en lo que están trabajando en el momento es el asunto más importante concebible se ilustra nuevamente por su actitud hacia la geometría proyectiva. La obra que hemos examinado en este capítulo muestra

ciertamente que la geometría proyectiva es fundamental para muchas geometrías. Sin embargo, no abarca a la geometría riemanniana, evidentemente vital, y al cuerpo creciente de la geometría algebraica. De cualquier modo, Cayley afirmó en su artículo de 1859 (sec. 3) que «La geometría proyectiva es toda la geometría y recíprocamente.»¹⁴² Bertrand Russell, en su *Essay on the Foundations of Geometry* (Ensayo sobre los fundamentos de la geometría) (1897), también creía que la geometría proyectiva era necesariamente la forma a priori de cualquier geometría del espacio físico. Hermann Hankel, a pesar de la atención que dio a la historia,¹⁴³ no vaciló en decir en 1869 que la geometría proyectiva es el camino real hacia todas las matemáticas. Nuestro examen de la evolución ya realizada muestra claramente que los matemáticos pueden fácilmente verse desorientados al dejarse llevar por sus entusiasmos.

Bibliografía

- Beltrami, Eugenio: *Opere matematiche*, Ulrico Hoepli, 1902, vol. 1.
- Bonola, Roberto: *Non-Euclidean Geometry*, Dover (reimp.), 1955, 129-264.
- Coolidge, Julián L.: *A History of Geometrical Methods*, Dover (reimp.), 1963, 68-87.
- Klein, Félix: *Gesammelte mathematische Abhandlungen*, Julius Springer, 1921-1923, vols. 1 y 2.

- Pasch, Moritz, y Max Dehn: Vorlesungen über neuere Geometrie, 2.^a ed., Julius Springer, 1926,' 185-239. Hay versión castellana, M. Pasch, Lecciones de geometría moderna, Madrid, Junta de Ampliación de Estudios, 1912.
- Pierpont, James: «Non-Euclidean Geometry. A Retrospect», Amer. Math. Soc. BtiU., 36, 1930, 66-76.
- Russell, Bertrand: An Essay on the Foundations of Geometry (1897), Dover (reimp.), 1956.

Capítulo 39

La geometría algebraica

En estos días el ángel de la topología y el demonio del álgebra abstracta luchan por el alma de cada dominio de las matemáticas.

Hermann Weyl

Contenido:

- 1. Antecedentes*
- 2. La teoría de invariantes algebraicos*
- 3. El concepto de transformación birracional*
- 4. El enfoque de la teoría de funciones en geometría algebraica*
- 5. El problema de la uniformización*
- 6. El enfoque de la geometría algebraica*
- 7. El enfoque aritmético*
- 8. La geometría algebraica de superficies*

Bibliografía

Antecedentes

Mientras se estaban creando las geometrías no euclídea y riemanniana, los investigadores de la geometría proyectiva estaban dedicados a su propio tema. Como hemos visto, las dos áreas fueron ligadas por la obra de Cayley y Klein. Después de que el método algebraico llegase a ser ampliamente usado en la geometría

proyectiva, el problema de reconocer qué propiedades de las figuras geométricas eran independientes de la representación en coordenadas pedía atención, y esto motivó el estudio de los invariantes algebraicos.

Las propiedades proyectivas de las figuras geométricas son aquellas que quedan invariantes bajo transformaciones lineales de las figuras. Mientras los matemáticos estudiaban estas propiedades, se permitieron ocasionalmente considerar transformaciones de mayor grado y buscar aquellas propiedades de las curvas y superficies que son invariantes bajo estas últimas transformaciones. La clase de transformaciones, que pronto reemplazaron a las transformaciones lineales como interés preferido, es llamada birracional porque éstas se expresan algebraicamente como funciones racionales de las coordenadas y las transformaciones inversas también son funciones racionales de sus coordenadas. La concentración en las transformaciones birracionales resultó indudablemente del hecho de que Riemann las había usado en su trabajo sobre integrales y funciones abelianas y de hecho, como veremos, los primeros grandes pasos en el estudio de las transformaciones birracionales de las curvas fueron guiados por lo que había hecho Riemann. Estos dos temas formaron el contenido de la geometría algebraica durante la última parte del siglo XIX.

El término geometría algebraica es inadecuado, ya que originalmente se refería a toda la obra desde la época de Fermat y Descartes en la que el álgebra había sido aplicada a la geometría; en la última parte del siglo XIX se aplicó al estudio de los invariantes

algebraicos y las transformaciones birracionales. En el siglo XX se refiere al último campo mencionado.

2. La teoría de invariantes algebraicos

Como ya hemos observado, la determinación de las propiedades geométricas de las figuras que se representan y estudian por medio de la representación coordenada exige el discernimiento de aquellas expresiones algebraicas que permanecen invariantes bajo los cambios de coordenadas. Visto en forma alternativa, la transformación proyectiva de una figura en otra por medio de una transformación lineal conserva algunas propiedades de la figura. Los invariantes algebraicos representan estas propiedades geométricas invariantes.

El tema de los invariantes algebraicos ya había surgido previamente en la teoría de números (cap. 34, sec. 5) y particularmente en el estudio de cómo se transforman las formas cuadráticas binarias.

$$f = ax^2 + 2 bxy + xy^2 \quad (1)$$

cuando x y y son transformadas por la transformación lineal T , a saber,

$$x = \alpha x' + \beta y' \quad y = \gamma x' + \delta y', \quad (2)$$

donde $\alpha\delta - \beta\gamma = r$. La aplicación de T a f da

$$f' = a'x'^2 + 2b'x'y' + c'y'^2 \quad (3)$$

En la teoría de números las cantidades a , b , c , a' , b' , c' y δ son enteros y $r = 1$. Sin embargo, es cierto generalmente que el discriminante D de f satisface la relación

$$D' = r^2 D \quad (4)$$

Las transformaciones lineales de la geometría proyectiva son más generales porque los coeficientes de las formas y las transformaciones no están restringidos a ser enteros. El término invariantes algebraicos se usa para distinguir aquellos que surgen bajo estas transformaciones lineales más generales de los invariantes modulares de la teoría de números y, a estos efectos, de los invariantes diferenciales de la geometría riemanniana.

Para discutir la historia de la teoría de los invariantes algebraicos necesitamos algunas definiciones. La forma de n -ésimo grado en una variable

$$f(x) = a_0x^n + a_1x^{n-1} + \dots + a_n$$

se convierte en coordenadas homogéneas en la forma binaria

$$f(x_1, x_2) = a_0x_1^n + a_1x_1^{n-1}x_2 + \dots + a_nx_2^n \quad (5)$$

En tres variables las formas son llamadas ternarias; en cuatro variables, cuaternarias, etc. Las definiciones que siguen se aplican a las formas en n variables.

Supóngase que sometemos la forma binaria a una transformación T de la forma (2). Mediante T , la forma $f(x_1, x_2)$ es transformada en la forma

$$F(X_1, X_2) = A_0 X_1^n + A_1 X_1^{n-1} X_2 + A_n X_2^n.$$

Los coeficientes de F diferirán de los de f y las raíces de $F = 0$ diferirán de las de $f = 0$. Cualquier función I de los coeficientes de f que satisface la relación

$$I(A_0, A_1, \dots, A_n) = r^w I(a_0, a_1, \dots, a_n)$$

es llamada un invariante de f . Si $w = 0$, el invariante se llama invariante absoluto de f . El grado del invariante es el grado en los coeficientes y el peso es w . El discriminante de una forma binaria es un invariante, como lo ejemplifica (4). En este caso el grado es 2 y el peso es 2. La significación del discriminante de cualquier ecuación polinomial $f(x) = 0$ es que su anulación es la condición de que $f(x) = 0$ tenga raíces iguales o, geoméricamente, que el lugar geométrico de $f(x) = 0$, el cual está formado por una serie de puntos, tengan dos puntos que coincidan. Esta propiedad es claramente independiente del sistema de coordenadas.

Si dos formas binarias (o más)

$$f_1 = a_0 x_1^m + \dots + a_m x_2^m$$

$$f_1 = b_0 x_1^n + \dots + b_n x_2^n$$

son transformadas por T en

$$F_1 = A_0 X_1^m + \dots + A_m X_2^m$$

$$F_1 = B_0 X_1^n + \dots + B_n X_2^n$$

entonces cualquier función I de los coeficientes que satisfaga la relación

$$I(A_0, \dots, A_m, B_1, \dots, B_n) = r^w I(a_0, \dots, a_m, b_1, \dots, b_n) \quad (6)$$

se dice que es un invariante simultáneo de las dos formas. Así, las formas lineales $a_1 x_1 + b_1 x_2$ y $a_2 x_1 + b_2 x_2$ tienen como invariante simultáneo la resultante $a_1 b_2 - a_2 b_1$ de las dos formas. Geométricamente, la anulación de la resultante significa que las dos formas representan al mismo punto (en coordenadas homogéneas). Dos formas cuadráticas

$$f_1 = a_1 x_1^2 + 2b_1 x_1 x_2 + c_1 x_2^2$$

y

$$f_2 = a_2x_1^2 + 2b_2x_1x_2 + c_2x_2^2$$

poseen un invariante simultáneo

$$D_{12} = a_1 c_2 - 2b_1b_2 + a_2c_1$$

cuya anulación expresa el hecho de que f_1 y f_2 representan pares armónicos de puntos.

Más allá de los invariantes de una forma o de un sistema de formas hay covariantes. Cualquier función C de los coeficientes y variables de f que es un invariante por T excepto para una potencia del módulo (determinante) de T es llamada un covariante de f . Así, para las formas binarias, un covariante satisface la relación

$$C(A_0, A_1, \dots, A_n, X_1, X_2) = r^w C(a_0, a_1, \dots, a_n, x_1, x_2).$$

Las definiciones de covariantes absolutos y simultáneos son análogas a las de los invariantes. El grado de un covariante en los coeficientes es llamado su grado y el grado en sus variables es llamado orden. Los invariantes son, por tanto, covariantes de orden cero. Sin embargo, algunas veces la palabra invariante se usa para referirse a un invariante en el sentido más estricto o a un covariante.

Un covariante de f representa alguna figura que no solamente está relacionada con f , sino relacionada proyectivamente. Así, el

jacobiano de dos formas binarias cuadráticas $f(x_1, x_2)$ y $\Phi(x_1, x_2)$ a saber

$$\begin{vmatrix} \frac{\partial f}{\partial x_1} & \frac{\partial f}{\partial x_2} \\ \frac{\partial \Phi}{\partial x_1} & \frac{\partial \Phi}{\partial x_2} \end{vmatrix}$$

es un covariante simultáneo de peso 1 de las dos formas. Geométricamente, el conjunto donde el jacobiano es igual a cero representa un par de puntos que es armónico con cada uno de los pares originales representados por f y Φ y la propiedad armónica es proyectiva. El hessiano de una forma binaria introducido por Hesse¹⁴⁴,

$$\begin{vmatrix} \frac{\partial^2 f}{\partial x_1^2} & \frac{\partial^2 f}{\partial x_1 \partial x_2} \\ \frac{\partial^2 f}{\partial x_1 \partial x_2} & \frac{\partial^2 f}{\partial x_2^2} \end{vmatrix}$$

es un covariante de peso 2. Su significado geométrico es demasiado complicado para concederle espacio aquí (Cf. cap. 35, sec. 5). El concepto de hessiano y su covariancia se aplica a cualquier forma en n variables.

El trabajo sobre invariantes algebraicos fue iniciado en 1841 por George Boole (1815-1864), cuyos resultados¹⁴⁵ eran limitados. Lo que tiene mayor relevancia es que Cayley fue atraído hacia este

tema por la obra de Boole y también hizo que Sylvester se interesara en el asunto. A ellos se les unió George Salmón (1819-1904), quien fue profesor de matemáticas en el Trinity College en Dublín desde 1840 hasta 1866, y después se convirtió en profesor de teología en esa institución. Estos tres autores realizaron tanto trabajo sobre los invariantes que en una de sus cartas Hermite los apodó la trinidad invariante.

En 1841 Cayley empezó a publicar artículos matemáticos sobre el aspecto algebraico de la geometría proyectiva. El artículo de Boole de 1841 sugirió a Cayley el cálculo de los invariantes de las funciones homogéneas de grado n . Llamó a los invariantes derivados y después hiperdeterminantes; el término invariante se le debe a Sylvester¹⁴⁶. Cayley, empleando las ideas de Hesse y de Eisenstein sobre determinantes, desarrolló una técnica para generar sus «derivados». Después publicó diez artículos sobre cuánticos en las *Philosophical Transactions*, desde 1854 hasta 1878¹⁴⁷. Cuánticos fue el término que adoptó para los polinomios homogéneos en 2, 3 o más variables. Cayley llegó a interesarse tanto en los invariantes que los investigó por el propio interés que presentaban. También inventó un método simbólico para tratar los invariantes.

En el caso particular de la forma cuártica binaria¹⁴⁸

$$f = ax_1^4 + 4bx_1^3x_2 + 6cx_1^2x_2^2 + 4dx_1x_2^3 + ex_2^4$$

Cayley mostró que el hessiano H y el jacobiano de f y H son covariantes y que

$$g^2 = ae - 4bd + 3c^2$$

$$g_3 = \begin{vmatrix} a & b & c \\ b & c & d \\ c & d & e \end{vmatrix}$$

son invariantes.

A estos resultados agregaron muchos

A estos resultados agregaron muchos más Sylvester y Salmón.

Otro personaje que aportó resultados, Ferdinand Eisenstein, quien estaba más interesado en la teoría de números, ya había encontrado para la forma cúbica binaria

$$f = ax_1^3 + 3bx_1^2x_2 + 3cx_1x_2^2 + dx_2^3$$

que el covariante más sencillo de segundo grado es su hessiano H y el invariante más sencillo es

$$3b^2c^2 + 6abcd - 4b^3d - 4ae^3 - a^2d^2,$$

el cual es el determinante del hessiano cuadrático, al igual que el discriminante de f También el jacobiano de f y H es otro covariante de orden tres. Después Siegfried Heinrich Aronhold (1819-1884), quien empezó a trabajar con invariantes en 1849, proporcionó invariantes para las formas cúbicas ternarias.¹⁴⁹

El primer problema importante que afrontaron los fundadores de la teoría de invariantes fue el descubrimiento de invariantes

particulares. Esta fue la dirección de la investigación desde aproximadamente 1840 hasta 1870. Como podemos ver, se pueden construir muchas de estas funciones porque algunos invariantes tales como el jacobiano y el hessiano son ellos mismos formas que tienen invariantes y porque algunos invariantes, tomados junto con la forma original, dan un nuevo sistema de formas que tienen invariantes simultáneos. Docenas de matemáticos sobresalientes, incluyendo a los pocos que ya hemos mencionado, calcularon invariantes particulares.

La continuación del cálculo de invariantes llevó al problema principal de la teoría de invariantes, el cual surgió después de que se encontraran muchos invariantes especiales o particulares; éste consistía en encontrar un sistema completo de invariantes. Lo que esto quiere decir es encontrar para una forma de un número de variables y grado dados, el número más pequeño posible de invariantes enteros racionales o covariantes tal que cualquier otro invariante o covariante entero racional pudiera expresarse como una función entera racional con coeficientes numéricos de este conjunto completo. Cayley demostró que los invariantes y covariantes encontrados por Eisenstein para la forma cúbica binaria y los que obtuvo para la forma cuártica binaria son un sistema completo para los casos respectivos.¹⁵⁰

Esto dejó abierto el problema de un sistema completo para otras formas.

La existencia de un sistema completo finito o base para las formas binarias de cualquier grado dado fue establecida primero por Paul

Gordan (1837-1912), quien dedicó la mayor parte de su vida a este tema. Su resultado¹⁵¹ es que a cada forma binaria $f(x_1, x_2)$ le corresponde un sistema completo finito de invariantes y covariantes enteros racionales. Gordan contó con el auxilio de teoremas debidos a Clebsch y el resultado es conocido como el teorema de Clebsch-Gordan. La demostración es larga y difícil. Gordan también demostró¹⁵² que cualquier *sistema* finito de formas binarias tiene un sistema completo finito de invariantes y covariantes. Las pruebas de Gordan mostraron cómo calcular los sistemas completos.

Se obtuvieron varias extensiones limitadas de los resultados de Gordan durante los veinte años siguientes. Gordan mismo proporcionó el sistema completo para la forma cuadrática ternaria,¹⁵³ para la forma cúbica ternaria,¹⁵⁴ y para un sistema de dos y tres cuadráticas ternarias.¹⁵⁵ Para la cuártica ternaria especial

$$x_1^3x_2 + x_2^3x_3 + x_3^3x_1$$

Gordan dio un sistema completo de 54 formas fundamentales.¹⁵⁶

En 1886, Franz Mertens (1840-1927)¹⁵⁷ volvió a demostrar el teorema de Gordan para los sistemas binarios por un método inductivo. Supuso que el teorema era cierto para cualquier conjunto dado de formas binarias y después demostró que debería seguir siendo cierto cuando se aumentaba en uno el grado de una de las formas. No mostró explícitamente el conjunto finito de invariantes y covariantes independientes pero sí demostró que existía. El caso más sencillo, una forma lineal, fue el punto de partida de la

inducción y tal forma tiene como covariantes solamente potencias de sí misma.

Hilbert, en 1885, después de escribir una tesis doctoral sobre invariantes,¹⁵⁸ también volvió a demostrar en 1888¹⁵⁹ el teorema de Gordan de que cualquier sistema dado de formas binarias tiene un sistema de invariantes y covariantes completo finito. Su demostración fue una modificación de la de Mertens. Ambas pruebas eran mucho más fáciles que la de Gordan. Pero la prueba de Hilbert tampoco presentó un proceso para encontrar el sistema completo.

En 1888, Hilbert sorprendió a la comunidad matemática anunciando un enfoque totalmente nuevo del problema de demostrar que cualquier forma de grado dado y número de variables dado, y que cualquier sistema de formas dado en cualquier número de variables dado, tiene un sistema completo finito de invariantes y covariantes enteros racionales independientes.¹⁶⁰ La idea básica de este nuevo enfoque consistía en olvidarse de los invariantes por el momento y considerar la cuestión de si, dado un sistema infinito de expresiones enteras racionales en un número finito de variables, ¿bajo qué condiciones existe un número finito de estas expresiones, una base, en términos de la cual todas las demás sean expresables como combinaciones lineales con funciones enteras racionales de las mismas variables como coeficientes? La respuesta es: siempre. Más específicamente, el teorema de la base de Hilbert, que precede al resultado sobre invariantes, es como sigue: por forma algebraica

entendemos una función homogénea entera racional en n variables con coeficientes en algún dominio de racionalidad definido (cuerpo). Dada una colección de infinitas formas de cualquier grado en n variables, entonces existe un número finito (una base) F_1, F_2, \dots, F_m tal que cualquier forma F de la colección puede escribirse como

$$F = A_1F_1 + A_2F_2 + \dots + A_mF_m$$

donde A_1, A_2, \dots, A_m son formas apropiadas en las n variables (no necesariamente en el sistema infinito) con coeficientes en el mismo dominio que los coeficientes del sistema infinito.

En la aplicación de este teorema a los invariantes y covariantes, el resultado de Hilbert establece que para cualquier forma o sistema de formas existe un número finito de invariantes y covariantes enteros racionales por medio de los cuales se pueden expresar todos los demás invariantes o covariantes enteros racionales como combinación lineal de los del conjunto finito. Esta colección finita de invariantes y covariantes es el sistema completo de invariantes.

La demostración de existencia de Hilbert era tan sencilla con respecto al cálculo laborioso de Gordan, que éste no pudo evitar exclamar: «Esto no son matemáticas; es teología». Sin embargo, reconsideró el asunto y posteriormente dijo: «Me he convencido a mí mismo de que la teología también tiene sus ventajas.» De hecho, él mismo simplificó la demostración de existencia dada por Hilbert.¹⁶¹

En los años 1880 y 1890 se pensó en la teoría de los invariantes como unificadora de muchas áreas de las matemáticas. Esta teoría

fue el «álgebra moderna» del período. Sylvester dijo en 1864:¹⁶² «De la misma manera que todos los caminos conducen a Roma, así encuentro, en mi propio caso al menos, que todas las investigaciones algebraicas, tarde o temprano, van a dar al Capitolio del álgebra moderna sobre cuyo pórtico brillante se inscribe la Teoría de los Invariantes.» Pronto la teoría se convirtió en un fin en sí misma, independientemente de sus orígenes en la teoría de números y en la geometría proyectiva. Los que se dedicaban al estudio de los invariantes algebraicos persistieron en demostrar toda clase de identidades algebraicas, tuvieran o no significación geométrica. Maxwell, cuando era estudiante en Cambridge, dijo que algunos hombres veían al universo entero en términos de quinticas y cuánticas.

Por otro lado, los físicos de finales del siglo XIX no tomaron cartas en el asunto. Ciertamente Tait observó una vez acerca de Cayley, «¿No es una vergüenza que ese hombre tan sobresaliente dedique sus habilidades a tales cuestiones totalmente inútiles?» Sin embargo, el asunto sí tuvo impacto en la física, indirecta y directamente, en gran medida a través del trabajo en invariantes diferenciales.

A pesar del enorme entusiasmo por la teoría de invariantes en la segunda mitad del siglo XIX, el tema perdió atractivo tal y como fue concebido e investigado durante ese período. Los matemáticos dicen que Hilbert mató la teoría de invariantes porque había resuelto todos los problemas. Hilbert le escribió a Minkowski en 1893 que ya no se dedicaría más a ese tema, y dijo en un artículo de 1893 que

las metas generales más importantes de la teoría ya se habían alcanzado. Sin embargo, estaba lejos de ser así. El teorema de Hilbert no mostraba cómo calcular los invariantes para cualquier forma o sistema de formas dados, y por tanto no podía proporcionar ni siquiera un solo invariante significativo. La búsqueda de invariantes específicos que tuvieran significado geométrico o físico era aún importante. Incluso el cálculo de una base para las formas de grado y número de variables dados podría ser valioso.

Lo que «mató» a la teoría de invariantes en el sentido que se entendía el tema en el siglo XIX es la usual colección de factores que mataron muchas otras actividades investigadas con demasiado entusiasmo. Los matemáticos siguen a los maestros. El pronunciamiento de Hilbert y el hecho de que él mismo abandonó el campo, ejercieron gran influencia en otros. También el cálculo de invariantes específicos significantes había llegado a ser más difícil después de que se lograran los resultados más fácilmente obtenibles.

El cálculo de invariantes algebraicos no acabó con la obra de Hilbert. Emmy Noether (1882-1935), una alumna de Gordan, escribió su tesis doctoral en 1907 «Sobre Sistemas Completos de Invariantes para las Formas Bicuadráticas Ternarias». ¹⁶³ También dio un sistema completo de formas covariantes para una cuártica ternaria, 331 en total. En 1910 extendió el resultado de Gordan a n variables. ¹⁶⁴

La historia posterior de la teoría algebraica de invariantes pertenece al álgebra abstracta moderna. La metodología de Hilbert llevó a

primer plano la teoría abstracta de los módulos, anillos y cuerpos. Hilbert demostró, en este lenguaje, que todo sistema modular (un ideal en la clase de los polinomios en n variables) tiene una base que consiste en un número finito de polinomios, o que todo ideal en un dominio polinomial de n variables posee una base finita siempre que en el dominio de los coeficientes de los polinomios todo ideal tenga una base finita. De 1911 a 1919, Emmy Noether escribió muchos artículos sobre bases finitas para varios casos distintos usando la técnica de Hilbert y la suya propia. En su desarrollo del siglo XX dominó el punto de vista algebraico abstracto. Tal y como se quejó Eduard Study en su texto sobre la teoría de invariantes, hubo falta de interés por problemas específicos y solamente se investigaron los métodos abstractos.

3. El concepto de transformación birracional

Vimos en el capítulo 35 que, especialmente durante las décadas tercera y cuarta del siglo XIX, el trabajo en geometría proyectiva se dirigió hacia las curvas de grado superior. Sin embargo, antes de que este trabajo hubiese avanzado demasiado, hubo un cambio en la naturaleza del estudio. El punto de vista proyectivo significa transformaciones lineales en coordenadas homogéneas. Gradualmente entraron en escena transformaciones de segundo grado (y superior) y el interés se desvió hacia las transformaciones birracionales. Tales transformaciones, para el caso de dos coordenadas no homogéneas, son de la forma

$$x' = \Phi(x,y) \quad y' = \psi(x,y)$$

donde Φ y ψ son funciones racionales en x e y , y además x e y pueden expresarse como funciones racionales de x' e y' . En coordenadas homogéneas x_1 , x_2 y x_3 las transformaciones son de la forma

$$x'_i = F_i(x_1, x_2, x_3) \quad i = 1, 2, 3$$

y la inversa es

$$x_i = G_i(x_1, x_2, x_3) \quad i = 1, 2, 3$$

donde F_i y G_i son polinomios homogéneos de grado n en sus respectivas variables.

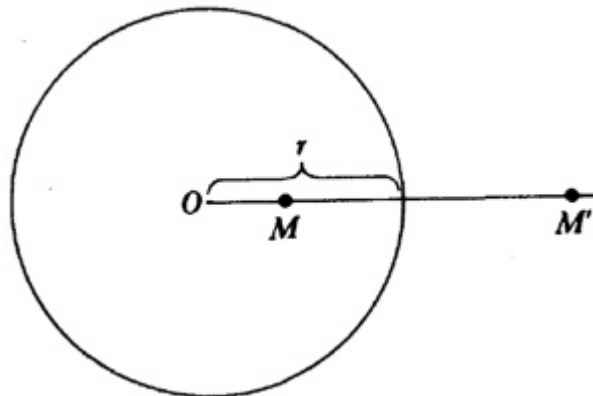


Figura 39.1

La correspondencia es uno-a-uno, excepto que cada uno de los puntos de un conjunto finito de ellos puede corresponder a una curva.

Como ejemplo de una transformación birracional tenemos la inversión con respecto a un círculo. Geométricamente, esta transformación (fig. 39.1) va de M a M' o de M' a M por medio de la ecuación que la define

$$OM \times OM' = r^2,$$

donde r es el radio del círculo. Algebraicamente, si fijamos un sistema de coordenadas en O , por el teorema de Pitágoras se llega a que

$$x' = \frac{r^2 x}{x^2 + y^2} \quad y' = \frac{r^2 y}{x^2 + y^2} \quad (7)$$

donde M es (x,y) y M' es (x',y') . Con esta transformación los círculos se transforman en círculos o líneas rectas y viceversa. La inversión es una transformación que lleva todo el plano en sí mismo y tales transformaciones birracionales son llamadas transformaciones de Cremona. Otro ejemplo de una transformación de Cremona en tres variables (homogéneas) es la transformación cuadrática

$$x_1 = x_2 x_3, \quad x_2 = x_3 x_1 \quad x_3 = x_1 x_2 \quad (8)$$

cuya inversa es

$$x_1 = x_2x_3, \quad x_2 = x_3x_1 \quad x_3 = x_1x_2$$

El término transformación birracional también se usa en un sentido más general, a saber, donde la transformación de los puntos de una curva en los de otra es birracional pero la transformación no necesita ser birracional en todo el plano. Así (en coordenadas no homogéneas) la transformación

$$X = x^2, \quad Y = y \quad (9)$$

no es inyectiva en todo el plano pero sí lleva cualquier curva C situada a la derecha del eje y en otra por una correspondencia uno-a-uno.

La transformación de inversión fue la primera de las transformaciones birracionales que aparecieron. En situaciones limitadas fue usada por Poncelet en su *Traite* de 1822 (π370) y después por Plücker, Steiner, Quetelet y Ludwig Immanuel Magnus (1790-1861). Fue estudiada extensamente por Möbius¹⁶⁵ y su uso en física fue reconocido por Lord Kelvin ¹⁶⁶ y por Liouville,¹⁶⁷ quien la llamó transformación por radios recíprocos.

En 1854, Luigi Cremona (1830-1903), que fue profesor de matemáticas en varias universidades italianas, introdujo la transformación birracional general (del plano completo en sí mismo) y escribió importantes artículos sobre ello.¹⁶⁸ Max Noether (1844-

1921), padre de Emmy Noether, demostró entonces el resultado fundamental¹⁶⁹ de que una transformación plana de Cremona puede construirse a partir de una sucesión de transformaciones cuadráticas y lineales. Jacob Rosanes (1842-1922) encontró este resultado independientemente¹⁷⁰ y demostró también que todas las transformaciones algebraicas uno-a-uno del plano deben ser transformaciones de Cremona. Las demostraciones de Noether y Rosanes fueron completadas por Guido Castelnuovo (1865-1952).¹⁷¹

4. El enfoque de la teoría de funciones en geometría algebraica

Aunque era clara la naturaleza de las transformaciones birracionales, el desarrollo del área de la geometría algebraica como el estudio de los invariantes bajo tales transformaciones fue, al menos en el siglo XIX, insatisfactorio. Se emplearon varios enfoques; los resultados eran fragmentarios y sin relación; la mayoría de las demostraciones estaban incompletas; y se obtuvieron muy pocos teoremas importantes. La variedad de los enfoques ha dado lugar a marcadas diferencias en los lenguajes usados. Las metas del área también eran vagas. Aunque el tema principal había sido la invariancia bajo transformaciones birracionales, la materia cubre la investigación de las propiedades de las curvas, superficies y estructuras de mayores dimensiones. Desde este punto de vista no hay muchos resultados importantes. Daremos unos pocos ejemplos de lo que se hizo.

Clebsch planteó uno de los primeros enfoques. (Rudolf Friedrich) Alfred Clebsch (1833-1872) estudió bajo el asesoramiento de Hesse

en Königsberg de 1850 a 1854. En la primera etapa de su trabajo estuvo interesado en la física matemática y de 1858 a 1863 fue profesor de mecánica teórica en Karlsruhe y después profesor de matemáticas en Giessen y Göttingen. Trabajó sobre problemas dejados por Jacobi en el cálculo de variaciones y la teoría de las ecuaciones diferenciales. En 1862 publicó el *Lehrbuch der Elasticitat* (*Tratado de Elasticidad*). Sin embargo, su labor principal la llevó a cabo en invariantes algebraicos y geometría algebraica.

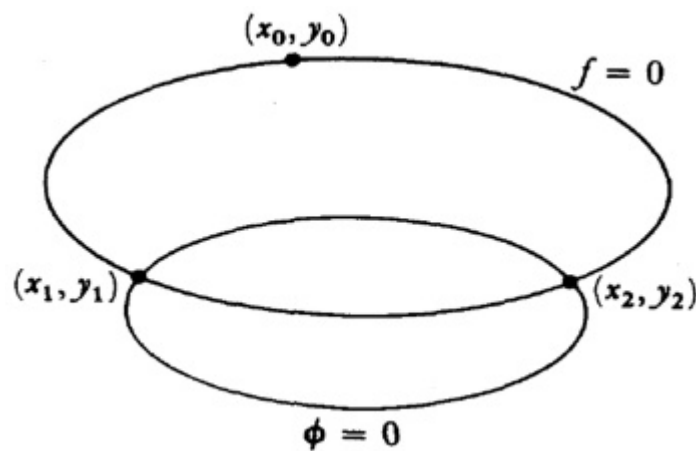


Figura 39.2

Clebsch había trabajado sobre las propiedades proyectivas de las curvas y superficies de tercero y cuarto grado hasta aproximadamente 1860. Conoció a Paul Gordan en 1863 y supo del trabajo de Riemann en la teoría de las funciones de variable compleja. Clebsch hizo entonces que esta teoría sirviera a la teoría de las curvas.¹⁷² Este enfoque es llamado trascendente. A pesar de que Clebsch llevó a cabo la conexión entre las funciones complejas y las curvas algebraicas, en una carta que envió a Gustav Roch

admitió que no podía entender la obra de Riemann sobre funciones abelianas ni las contribuciones de Roch en su disertación.

Clebsch reinterpretó la teoría de las funciones complejas de la siguiente manera: la función $f(w,z) = 0$, donde z y w son variables complejas, geométricamente requiere de una superficie de Riemann para z y un plano o una parte de un plano para w o, si se prefiere, de una superficie de Riemann a la cual se asigna un par de valores z y w a cada uno de sus puntos. Considerando sólo la parte real de z y w , la ecuación $f(w,z) = 0$ representa una curva en el plano cartesiano real. Es posible que z y w tengan aún valores complejos que satisfagan $f(w,z) = 0$, pero éstos no se dibujan. Este punto de vista de las curvas reales con puntos complejos era ya familiar a partir del trabajo en geometría proyectiva. A la teoría de las transformaciones birracionales de las superficies corresponde una teoría de las transformaciones birracionales de las curvas planas. Con la reinterpretación ahora descrita, los puntos de ramificación de la superficie de Riemann corresponden a aquellos puntos de la curva donde una línea $x = \text{const.}$, corta a la curva en dos o más puntos consecutivos, esto es, o es tangente a la curva o pasa por una cúspide. Un punto doble de la curva corresponde a un punto en la superficie donde dos hojas se tocan exactamente, sin ninguna relación adicional. Los puntos múltiples de mayor orden en las curvas también corresponden a otras peculiaridades de las superficies de Riemann.

En lo que sigue utilizaremos las siguientes definiciones (*cf.* cap. 23, sec. 3): un punto múltiple (punto singular) P de orden $k > 1$ de una

curva plana de grado n es un punto tal que una recta genérica que pasa por P corta a la curva en $n - k$ puntos. El punto múltiple es ordinario si las k tangentes en P son distintas. Al contar las intersecciones de una curva de n -ésimo grado y de otra de grado m -ésimo, se debe tener en cuenta la multiplicidad de los puntos múltiples sobre cada curva. Si es h en la curva C'' y k en C''' y si las tangentes en P de C'' son distintas de las de C''' , entonces el punto de intersección tiene multiplicidad hk . Se dice que una curva C es adjunta de una curva C' cuando los puntos múltiples de C son ordinarios o cúspides y si C' tiene un punto de orden de multiplicidad $k - 1$ en todo punto múltiple de C de orden k .

Clebsch¹⁷³ reformuló por primera vez el teorema de Abel (cap. 27, sec. 7) sobre integrales de primera clase en términos de curvas. Abel consideró una función racional fija $R(x,y)$ donde x e y están relacionadas por cualquier curva algebraica $f(x,y) = 0$, de modo que y sea una función de x . Supóngase (fig. 39.2) que $f = 0$ es cortada por otra curva algebraica

$$\Phi(x, y, a_1, a_2, \dots, a_k) = 0,$$

donde los a_i son los coeficientes de $\Phi = 0$. Sean (x_1, y_1) , (x_2, y_2) , $\dots, (x_m, y_m)$ las intersecciones de $\Phi = 0$ con $f = 0$. (El número m de ellas es el producto de los grados de f y Φ). Dado un punto (x_0, y_0) en $f = 0$, donde y_0 pertenece a una rama de $f = 0$, entonces podemos considerar la suma

$$I = \sum_{i=1}^m \int_{x_0, y_0}^{x_i, y_i} R(x, y) dx$$

Los límites superiores x_i, y_i se encuentran en $\Phi = 0$ y la integral I es una función de los límites superiores. Entonces hay un número característico p de estos límites que tienen que ser funciones algebraicas de los otros. Este número p depende solamente de f . Más aún, f puede expresarse como la suma de estas p integrales y de funciones racionales y logarítmicas de las x_p, y_p $i = 1, 2, \dots, m$. Además, si se hace variar la curva $\Phi = 0$ variando los parámetros de a_i hasta a_k , entonces x_i variará también en f se convertirá en una función de los a_i mediante los x . La función f de los a_i será racional en los a o, en el peor de los casos, involucrará funciones logarítmicas de las a_i .

Clebsch también aplicó a las curvas el concepto de Riemann de las integrales abelianas sobre superficies de Riemann, esto es, las integrales de la forma $\int g(x, y) dx$, donde g es una función racional y $f(x, y) = 0$. Para ilustrar las integrales de primera clase considérese una curva plana de cuarto grado C_4 sin puntos dobles. Aquí $p = 3$ y se tienen las tres integrales finitas en todas partes.

$$u_1 = \int \frac{x dx}{f_y} \quad u_2 = \int \frac{y dx}{f_y} \quad u_3 = \int \frac{dx}{f_y}$$

Lo que se aplica a la C_4 también se cumple para curvas algebraicas arbitrarias $f(x, y) = 0$ de orden n . En lugar de las tres integrales

finitas en todas partes hay entonces p de estas integrales (donde p es el género de $f = 0$). Cada una de ellas tiene $2p$ módulos de periodicidad (cap. 27, sec. 8). Las integrales son de la forma

$$\int \frac{\Phi(x, y)}{\frac{\partial f}{\partial y}} dx$$

donde Φ es un polinomio (un adjunto) de grado precisamente $n = 3$, el cual se anula en los puntos dobles y cúspides de $f = 0$.

La contribución siguiente de Clebsch¹⁷⁴ consistió en introducir la noción de género como un concepto para clasificar las curvas. Si la curva tiene d puntos dobles entonces el género es $p = (1/2)(n - 1)(n - 2) - d$. Previamente ya se tenía la noción de la deficiencia de una curva (cap. 23, sec. 3), esto es, el máximo número posible de puntos dobles que podía poseer una curva de grado n , a saber $(n - 1)(n - 2)/2$, menos el número que realmente posee. Clebsch demostró¹⁷⁵ que para curvas con solamente puntos múltiples ordinarios (las tangentes son todas distintas) el género es igual a la deficiencia y es invariante por una transformación birracional de todo el plano en sí mismo.¹⁷⁶ La noción de género dada por Clebsch está relacionada con la conectividad de Riemann de una superficie de Riemann. La superficie de Riemann correspondiente a una curva de género p tiene conectividad $2p + 1$.

La noción de género puede utilizarse para establecer teoremas significativos acerca de las curvas. Jacob Lüroth (1844-1910) demostró¹⁷⁷ que una curva de género 0 puede ser transformada

birrationalmente en una línea recta. Cuando el género es 1, Clebsch demostró que una curva puede transformarse birracionalmente en una curva de tercer grado.

Además de la clasificación de curvas por el género, Clebsch, siguiendo a Riemann, introdujo clases dentro de cada género. Riemann había considerado¹⁷⁸ la transformación birracional de sus superficies. Así, si $f(w,z) = 0$ es la ecuación de la superficie y si

$$w_1 = R_1(w,z), \quad z_1 = R_2(w,z)$$

son funciones racionales, y si la transformación inversa es racional, entonces $f(w,z)$ puede transformarse en $F(w_1,z_1) = 0$. Dos ecuaciones algebraicas $F(w,z) = 0$ (o sus superficies) pueden transformarse birracionalmente una en la otra solamente si ambas tienen el mismo valor p . (El número de hojas no se conserva, en general). Para Riemann no era necesaria ninguna demostración adicional. Estaba garantizada por la intuición.

Riemann (en el artículo de 1857) trató todas las ecuaciones (o las superficies) que son transformables birracionalmente una en la otra como pertenecientes a la misma clase. Tienen el mismo género p . Sin embargo, hay diferentes clases con el mismo valor p (porque los puntos de ramificación pueden diferir). La clase más general de género p es caracterizada por $3p - 3$ constantes (complejas) coeficientes en la ecuación cuando $p > 1$, por una constante cuando $p = 1$ y por cero constantes cuando $p = 0$. En el caso de funciones elípticas $p = 1$, y hay una constante. Las funciones trigonométricas,

para las cuales $p = 0$, no tienen constante arbitraria alguna. El número de constantes fue llamado por Riemann el módulo de clase. Las constantes son invariantes por transformaciones birracionales. Clebsch, de manera parecida, puso a todas aquellas curvas que se pueden obtener de una dada por medio de una transformación birracional uno-a-uno en una clase. Las de una clase tienen necesariamente el mismo género, pero puede haber diferentes clases con el mismo género.

5. El problema de la uniformización

Clebsch volvió entonces su atención a lo que se llama el problema de la uniformización para curvas. Indiquemos primero exactamente a qué se refiere este problema. Dada la ecuación

$$w^2 + z^2 = 1 \quad (10)$$

la podemos representar en forma paramétrica como

$$z = \text{sen } t, \quad w = \text{cos } t \quad (11)$$

o en la forma paramétrica

$$z = \frac{2t}{1+t^2} \quad w = \frac{1-t^2}{1+t^2} \quad (12)$$

Así, aunque (10) define w como una función multivaluada de z , podemos representar z y w como funciones univaluadas o uniformes de t . Las ecuaciones paramétricas (11) o (12) se dice que uniformizan la ecuación algebraica (10).

Para una ecuación $f(w,z) = 0$ de género 0, Clebsch¹⁷⁹ demostró que cada una de las variables puede expresarse como una función racional de un solo parámetro. Estas funciones racionales son funciones uniformizadoras. Cuando $f = 0$ se interpreta como una curva, entonces es llamada unicursal. Inversamente, si las variables w y z de $f = 0$ son expresables racionalmente en términos de un parámetro arbitrario, entonces $f = 0$ es de género 0.

En el mismo año,¹⁸⁰ Clebsch demostró que cuando $p = 1$, entonces w y z pueden expresarse como funciones racionales de los parámetros ξ y η , donde η^2 es un polinomio de grado tres o cuatro en ξ . Entonces $f(w,z) = 0$, o la curva correspondiente es llamada bicursal, un término introducido por Cayley.¹⁸¹ También es llamada elíptica, porque la ecuación $(dw/dz)^2 = \eta$ conduce a las integrales elípticas. También podemos decir que w y z se pueden expresar como funciones univaluadas doblemente periódicas de un solo parámetro a o como funciones racionales de $\rho(a)$ donde $\rho(a)$ es la función de Weierstrass. El resultado de Clebsch sobre la uniformización de curvas de género 1 por medio de funciones elípticas de un parámetro hizo posible establecer propiedades notables para tales curvas acerca de puntos de inflexión, cónicas osculatrices, tangentes desde un punto a una curva, y otros

resultados, muchos de los cuales habían sido demostrado con anterioridad aunque con mayor dificultad.

Para las ecuaciones $f(w,z) = 0$ de género 2, Alexander von Brill (1842-1935) demostró¹⁸² que las variables w y z se pueden expresar como funciones racionales de ξ y η donde η es ahora un polinomio de quinto o sexto grado en ξ .

Por tanto, las funciones de género 0, 1 y 2 pueden uniformizarse. Para la función $f(w,z) = 0$ de género mayor que 2, la idea era emplear funciones más generales, a saber, funciones automorfas. En 1882, Klein¹⁸³ dio un teorema general de uniformización, pero la demostración no era completa. En 1883, Poincaré anunció¹⁸⁴ su teorema general sobre uniformización pero él tampoco tenía una prueba completa. Tanto Klein como Poincaré continuaron trabajando duro para demostrar este teorema, pero no se obtuvo ningún resultado decisivo durante veinticinco años. En 1907 Poincaré¹⁸⁵ y Paul Koebe (1882-1945)¹⁸⁶ dieron independientemente una demostración de este teorema de uniformización. Koebe extendió después el resultado en muchas direcciones. Con el teorema sobre uniformización ya establecido rigurosamente ha llegado a ser posible un tratamiento perfeccionado de las funciones algebraicas y de sus integrales.

6. El enfoque de la geometría algebraica

Con la colaboración de Clebsch y Gordan empieza una nueva dirección de investigación en la geometría algebraica durante los años 1865-1870. Clebsch no estaba satisfecho con mostrar sólo la

trascendencia de la obra de Riemann para las curvas. Buscaba ahora establecer la teoría de las integrales abelianas sobre la base de la teoría algebraica de las curvas. En 1865, él y Gordan reunieron sus esfuerzos en este trabajo y produjeron su *Theorie der Abelschen Funktionen* (*Teoría de las Funciones Abelianas*, 1866). Se debe apreciar que en esta época la teoría más rigurosa de Weierstrass sobre integrales abelianas no era conocida y el fundamento de Riemann —su prueba de existencia basada en el principio de Dirichlet— no solamente era extraño, sino que no estaba bien establecido. También en esta época hubo considerable entusiasmo por la teoría de los invariantes de formas algebraicas (o curvas) y por los métodos proyectivos como el primer estadio, por decirlo así, del tratamiento de las transformaciones birracionales.

Aunque el trabajo de Clebsch y Gordan fue una contribución a la geometría algebraica, no estableció una teoría puramente algebraica de la teoría de Riemann sobre las integrales abelianas. Sí que usaron métodos algebraicos y geométricos como opuestos a los métodos de la teoría de funciones de Riemann, pero también usaron resultados básicos de la teoría de funciones y los métodos de la teoría de funciones de Weierstrass. Además, tomaron algunos resultados acerca de las funciones racionales y el teorema del punto de intersección como ciertos. Su contribución se sumó para que, partiendo de algunos resultados de teoría de funciones y usando métodos algebraicos, se obtuvieran nuevos resultados establecidos previamente por métodos de la teoría de funciones. Las

transformaciones racionales fueron la esencia del método algebraico.

Ellos dieron la primera demostración *algebraica* de la invariancia del género p de una curva algebraica por las transformaciones racionales, usando como definición de p el grado y número de singularidades de $f = 0$. Entonces, usando el hecho de que p es el número de integrales linealmente independientes de primera clase sobre $f(x_1, x_2, x_3) = 0$ y que estas integrales son siempre finitas, mostraron que la transformación

$$qx_i = \psi(y_1, y_2, y_3) \quad i = 1, 2, 3$$

transforma una integral de primera clase en una integral de primera clase de tal modo que p es invariante. También proporcionaron nuevas demostraciones del teorema de Abel (usando ideas y métodos de la teoría de funciones).

Su trabajo no era riguroso. En particular también ellos, en la tradición de Plücker, contaron las constantes arbitrarias para determinar el número de puntos de intersección de una C_m con una C_n . No investigaron las clases especiales de puntos dobles. La trascendencia de la obra de Clebsch-Gordan para la teoría de las funciones algebraicas consistió en expresar claramente en forma algebraica resultados tales como el teorema de Abel y usarlo en el estudio de las integrales abelianas. Colocaron la parte algebraica de la teoría de las integrales y funciones abelianas más en primer

plano, y en particular establecieron la teoría de las transformaciones sobre fundamentos propios de ella.

Clebsch y Gordan habían planteado muchos problemas y dejado muchos huecos. Los problemas van en la dirección de nuevas investigaciones algebraicas de una teoría puramente algebraica de las funciones algebraicas. El trabajo con el enfoque algebraico fue continuado por Alexander von Brill y Max Noether a partir de 1871; su artículo principal fue publicado en 1874.¹⁸⁷ Brill y Noether basaron su teoría sobre un famoso teorema residual (*Restsatz*) que en sus manos ocupó el lugar del teorema de Abel. También dieron una prueba algebraica del teorema de Riemann-Roch sobre el número de constantes que aparecen en las funciones algebraicas $F(w,z)$ que no se hacen infinitas en lugar alguno a excepción de m puntos predeterminados de una C_n . De acuerdo con este teorema, la función algebraica más general que satisface esta condición tiene la forma

$$F = C_1 F_1 + C_2 F_2 + \dots + C_\mu F_\mu + C_{\mu+1}$$

donde

$$\mu = m - p + \tau$$

τ es el número de funciones Φ linealmente independientes (de grado $n - 3$) que se anulan en los m puntos prescritos, y p es el género de

la C_n . De esta manera, si la C_n es una C_4 sin puntos dobles, entonces $p = 3$ y las Φ son líneas rectas. Para este caso cuando

$$m = 1, \text{ entonces } \tau = 2 \quad \text{y} \quad \mu = 1 - 3 + 2 = 0;$$

$$m = 2, \text{ entonces } \tau = 1 \quad \text{y} \quad \mu = 2 - 3 + 1 = 0;$$

$$m = 3, \text{ entonces } \tau = 1 \text{ o } 0 \quad \text{y} \quad \mu = 1 \text{ o } 0.$$

Cuando $p = 0$ no hay ninguna función algebraica que se haga infinita en los puntos dados. Cuando $m = 3$, existe una y sólo una función, siempre que los tres puntos dados estén sobre una línea recta. Si los tres puntos caen en una recta $v = 0$, esta recta corta a la C_4 en un cuarto punto. Escogemos una recta $u = 0$ que pase por este punto y entonces $F_1 = u/v$.

Este trabajo reemplaza la determinación de Riemann de la función algebraica más general que tenga puntos dados en los cuales se haga infinita. También el resultado de Brill-Noether trasciende el punto de vista proyectivo en la medida en que trata de la geometría de los puntos sobre la curva C_n dada por $f = 0$, cuyas relaciones mutuas no son alteradas por una transformación bircional inyectiva. Así, por primera vez, se establecieron los teoremas sobre puntos de intersección de curvas de manera algebraica. Se prescindió de contar las constantes como método.

El trabajo en geometría algebraica continuó con la investigación detallada de curvas espaciales algebraicas por Noether¹⁸⁸ y

Halphen.¹⁸⁹ Cualquier curva espacial C puede ser proyectada birracionalmente en una curva plana C_1 . Todas las C_1 que se obtienen a partir de C tienen el mismo género. Por tanto, el género de C se define como el de cualquiera de dichas C_1 , y el género de C es invariante bajo una transformación birracional del espacio.

El tema que ha recibido la mayor atención durante el transcurso del tiempo es el estudio de las singularidades de las curvas algebraicas planas. Hasta 1871, la teoría de las funciones algebraicas considerada desde el punto de vista algebraico se había limitado al estudio de curvas que tengan puntos dobles distintos o separados, y en el peor de los casos solamente cúspides (*Ruckkehrpunkte*). Se creía que las curvas con singularidades más complicadas podían ser tratadas como casos límites de curvas con puntos dobles. Pero el procedimiento efectivo de paso al límite era vago y carecía de rigor y unidad. La culminación del trabajo sobre singularidades son dos famosos teoremas sobre transformaciones. El primero afirma que toda curva algebraica plana irreducible puede ser transformada por medio de una transformación de Cremona en una que no tenga más puntos singulares que puntos múltiples con tangentes distintas. El segundo afirma que, por medio de una transformación birracional solamente en la curva, toda curva algebraica plana irreducible puede transformarse en otra que tenga solamente puntos dobles con tangentes distintas. La reducción de las curvas a estas formas más sencillas facilita la aplicación de muchos de los métodos de la geometría algebraica.

Sin embargo, las numerosas pruebas de estos teoremas, especialmente del segundo, han sido incompletas o al menos criticadas por los matemáticos (excepto por el autor). Realmente, hay dos casos del segundo teorema, curvas reales en el plano proyectivo y curvas en el sentido de la teoría de las funciones complejas, donde x e y se toman cada una de ellas en un plano complejo. Noether¹⁹⁰ usó en 1871 una sucesión de transformaciones cuadráticas que son uno-a-uno en todo el plano para demostrar el primer teorema. Generalmente se le atribuye la prueba, pero realmente él sólo indicó una demostración que fue perfeccionada y modificada por muchos autores.¹⁹¹ Kronecker, usando el análisis y el álgebra, desarrolló un método para probar el segundo teorema. Comunicó este resultado verbalmente a Riemann y Weierstrass en 1858, dio conferencias sobre él a partir de 1870 y lo publicó en 1881.¹⁹² El método usaba transformaciones birracionales, las cuales, con la ayuda de la ecuación de la curva plana dada, son uno-a-uno, y transformaban el caso singular en el «regular»; esto es, los puntos singulares se convierten precisamente en puntos dobles con tangentes distintas. El resultado, sin embargo, no fue enunciado por Kronecker y sólo se encuentra implícito en su obra. Este segundo teorema, con la intención de que todos los puntos múltiples se puedan reducir a puntos dobles por medio de transformaciones birracionales sobre las curvas, fue enunciado primero explícitamente y demostrado por Halphen en 1884.¹⁹³ Se han dado muchas otras pruebas, pero ninguna es universalmente aceptada.

7. El enfoque aritmético

Además del enfoque trascendente y del de la geometría algebraica, está el llamado enfoque aritmético de las curvas algebraicas, el cual es, sin embargo, al menos en concepto, puramente algebraico. Este enfoque es realmente un grupo de teorías que difieren grandemente en detalle, pero que tienen en común la construcción y el análisis de los integrandos de las tres clases de integrales abelianas. Este enfoque fue desarrollado por Kronecker en sus conferencias,¹⁹⁴ por Weierstrass en las suyas de 1875-1876 y por Dedekind y Heinrich Weber en un artículo escrito conjuntamente.¹⁹⁵ El enfoque se presenta en su totalidad en el texto de Kurt Hensel y Georg Landsberg: *Theorie der algebraischen Funktionen einer Variablen* (*Teoría de las Funciones Algebraicas de una Variable*, 1902).

La idea central de este enfoque proviene del trabajo sobre los números algebraicos de Kronecker y Dedekind y utiliza una analogía entre los enteros algebraicos de un cuerpo de números algebraicos y las funciones algebraicas en la superficie de Riemann de una función compleja. En la teoría de los números algebraicos se empieza con una ecuación polinomial irreducible $f(x) = 0$ con coeficientes enteros. Lo análogo para la geometría algebraica es una ecuación polinomial irreducible $f(\zeta, z) = 0$, cuyos coeficientes en las potencias de ζ son polinomios en z (digamos con coeficientes reales). En la teoría de números se considera entonces el cuerpo $R(x)$ generado por los coeficientes de $f(x) = 0$ y una de sus raíces. En geometría se considera el cuerpo de todas las $R(\zeta, z)$ que son

algebraicas y univaluadas en la superficie de Riemann. Se considera entonces en la teoría de números los números algebraicos enteros. A éstos les corresponden las funciones algebraicas $G(\zeta, z)$ que son enteras, esto es, se hacen infinitas solamente en $z = \infty$. La descomposición de los enteros algebraicos en factores primos reales y unidades corresponde respectivamente a la descomposición de la $G(\zeta, z)$ en factores tales que cada uno de ellos se anula solamente en un punto de la superficie de Riemann y en factores que no se anulan en ningún punto, respectivamente. Donde Dedekind introdujo ideales en la teoría de números para discutir la divisibilidad, en el análogo geométrico se reemplaza un factor de una $G(\zeta, z)$ que se anula en un punto de la superficie de Riemann por la colección de todas las funciones del cuerpo de $E(\zeta, z)$ que se anulan en ese punto. Dedekind y Weber usaron este método aritmético para tratar el cuerpo de las funciones algebraicas y obtuvieron los resultados clásicos.

Hilbert¹⁹⁶ continuó lo que es esencialmente el enfoque algebraico o aritmético de Dedekind y Kronecker a la geometría algebraica. Un teorema principal, el *Nullstellensatz* de Hilbert, dice que toda estructura algebraica (figura) de extensión arbitraria en un espacio de un número arbitrario de variables homogéneas x_1, \dots, x_n siempre puede representarse por un número finito de ecuaciones homogéneas

$$F_1 = 0, \quad F_2 = 0, \quad \dots, \quad F_{\mu} = 0,$$

de modo que la ecuación de cualquier otra estructura que contenga a la original puede representarse por

$$M_1F_1 + \dots + M_{\mu}E_{\mu} = 0,$$

donde las M son formas enteras homogéneas arbitrarias cuyo grado debe escogerse de modo tal que el lado izquierdo de la ecuación sea él mismo homogéneo.

Siguiendo a Dedekind, Hilbert llamó a la colección de los MF_i un módulo (el término que ahora se emplea es ideal y módulo es algo más general). Se puede enunciar el resultado de Hilbert así: toda estructura algebraica de R_n determina la anulación de un módulo finito.

8. La geometría algebraica de superficies

Casi desde el inicio de la investigación en la geometría algebraica de las curvas, se investigó también la teoría de las superficies. Aquí también el trabajo se dirigió hacia el estudio de los invariantes bajo transformaciones lineales y birracionales. Igual que la ecuación $f(x,y) = 0$, la ecuación polinomial $f(x,y,z) = 0$ tiene una doble interpretación. Si x y z toman valores reales, entonces la ecuación representa una superficie bidimensional en el espacio tridimensional. Sin embargo, si estas variables toman valores complejos, entonces la ecuación representa una variedad tetradimensional en un espacio de seis dimensiones.

El enfoque de la geometría algebraica de las superficies fue análogo al de las curvas. Clebsch empleó métodos de la teoría de funciones e introdujo¹⁹⁷ integrales dobles que hacen el papel de las integrales abelianas en la teoría de las curvas. Clebsch observó que para una superficie algebraica $f(x,y,z) = 0$ de grado m con puntos múltiples aislados y rectas múltiples ordinarias, ciertas superficies de grado $m - 4$ tendrían que jugar el papel que las curvas adjuntas de grado $m - 3$ juegan con respecto a una curva de grado m . Dada una función racional $R(x,y,z)$, donde x , y y z están relacionadas por $f(x,y,z) = 0$, si se investigan las integrales dobles

$$\iint R(x,y,z) dx dy$$

que permanecen siempre finitas cuando las integrales se extienden sobre una región bidimensional de la superficie tetradimensional, se encuentra que son de la forma

$$\iint \frac{Q(x,y,z)}{f_z} dx dy$$

donde Q es un polinomio de grado $m - 4$. $Q = 0$ es una superficie adjunta que pasa por las rectas múltiples de $f = 0$ y tiene una recta múltiple de orden $k - 1$ al menos en cada recta múltiple de / de orden k y tiene un punto múltiple de orden $q - 2$ al menos en cada punto múltiple aislado de orden q . Tal integral es llamada una integral doble de primera clase. El número de integrales linealmente

independientes de esta clase, que es el número de constantes esenciales en $Q(x,y,z)$, es llamado el género geométrico p de $f = 0$. Si la superficie no tiene rectas múltiples de puntos

$$p_g = \frac{(m-1)(m-2)(m-3)}{6}$$

Max Noether¹⁹⁸ y Hieronymus G. Zeuthen (1839-1920)¹⁹⁹ demostraron que p es invariante por las transformaciones birracionales de la superficie (no de todo el espacio).

Hasta este punto la analogía con la teoría de las curvas es buena. Las integrales dobles de primera clase son análogas a las integrales abelianas de primera clase. Pero se manifiesta ahora una primera diferencia. Es necesario calcular el número de constantes esenciales en los polinomios Q de grado $m - 4$ que en puntos múltiples de la superficie se comportan de manera tal que la integral permanece finita. Pero por medio de una fórmula se puede encontrar el número preciso de condiciones de este modo involucradas solamente para un polinomio de grado N suficientemente grande. Si en esta fórmula se pone $N = m - 4$ se podría encontrar un número diferente de p_g

Cayley²⁰⁰ llamó a este nuevo número el género numérico (aritmético) p_n de la superficie. El caso más general es cuando $p_n = p_g$. Cuando no se cumple la igualdad se tiene $p_n < p_g$ y la superficie escamada irregular, si no es llamada regular. Después Zeuthen ²⁰¹ y Noether ²⁰² establecieron la invariancia del número p_n cuando no es igual a p_g .

Picard²⁰³ desarrolló una teoría de integrales dobles de segunda clase. Estas son las integrales que se hacen infinitas a la manera de

$$\iint \left(\frac{\partial U}{\partial x} + \frac{\partial V}{\partial y} \right) dx dy \quad (13)$$

donde U y V son funciones racionales de x , y y z y $f(x,y,z) = 0$. El número de integrales diferentes de segunda clase, diferentes en el sentido de que ninguna combinación lineal de estas integrales se reduce a la forma (13), es finito; este es un invariante birracional de la superficie $f = 0$. Pero no es cierto aquí, como en el caso de las curvas, que el número de integrales abelianas distintas de segunda clase es $2p$. Este nuevo invariante de las superficies algebraicas no parece estar vinculado con el género numérico o el geométrico.

Hasta ahora se ha logrado mucho menos para la teoría de las superficies que para las curvas. Una razón es que las posibles singularidades de las superficies son mucho más complicadas. Se tiene el teorema de Picard y Georges Simart demostrado por Beppo Levi (1875-1928)²⁰⁴ de que cualquier superficie algebraica (real) puede ser transformada birracionalmente en una superficie sin singularidades, la cual, sin embargo, debe estar en un espacio de cinco dimensiones. Pero este teorema no resulta ser demasiado útil.

En el caso de las curvas, el número invariante, el género p , es susceptible de ser definido en términos de las características de la curva o la conectividad de la superficie de Riemann. En el caso de $f(x,y,z) = 0$, el número que caracteriza los invariantes birracionales

aritméticos es desconocido ²⁰⁵. No intentaremos describir más a fondo los pocos resultados limitados para la geometría algebraica de las superficies.

El tema de estudio de la geometría algebraica abarca ahora el estudio de figuras de mayores dimensiones («manifolds» o variedades) definidas por una o más ecuaciones algebraicas. Más allá de la generalización en esta dirección, otra extensión, a saber, el uso de coeficientes más generales en las ecuaciones que las definen, también ha sido llevada a cabo. Estos coeficientes pueden ser elementos de un anillo o cuerpo abstractos y se aplican los métodos del álgebra abstracta. Los diversos métodos de cultivar la geometría algebraica, al igual que la formulación algebraica abstracta introducida en el siglo XX, han llevado a agudas diferencias en el lenguaje y métodos de enfoque, de modo que una clase de investigadores encuentra muy difícil entender a otras. En este siglo, el interés ha sido puesto en el enfoque algebraico abstracto. Parece que sí ofrece formulaciones y teoremas precisos, y de ese modo las demostraciones zanján muchas controversias acerca del significado y corrección de los resultados anteriores. Sin embargo, parece que mucho del trabajo hecho tiene mucha más relación con el álgebra que con la geometría.

Bibliografía

- Barker, H. F.: «On Some Recent Advances in the Theory of Algebraic Surfaces». *Proc. Lon. Math. Soc.* (2), 12, 1912-1913, 1-40.

- Berzolari, L.: «Allgemeine Theorie der höheren ebenen algebraischen Kurven». *Encyk. der Math. Wiss.*, B. G. Teubner, 1903-1915, III C4, 313-455. «Algebraische Transformationen und Korrespondenzen». *Encyk. der Math. Wiss.*, B. G. Teubner, 1903-1915, III, 2, 2.^a parte B, 1781-2218. Util por los resultados sobre figuras de mayores dimensiones.
- Bliss, G. A.: «The Reduction of Singularities of Plane Curves by Birational Transformations». *Amer. Math. Soc. Bull.*, 29, 1923, 161-183.
- Brill, A., y Noether, M.: «Die Entwicklung der Theorie der algebraischen Funktionen». *Jahres. der Deut. Math.-Verein.*, 3, 1892-1893, 107-565.
- Castelnuovo, G., y Enriques, F.: «Die algebraischen Flächen vom Gesichtspunkte der birationalen Transformationen». *Encyk. der Math. Wiss.*, B. G. Teubner, 1903-1915, III C6b, 674-768. «Sur quelques récents résultats dans la théorie des surfaces algébriques». *Math. Ann.*, 48, 1897, 241-316.
- Cayley, A.: *Collected Mathematical Papers*. Johnson Reprint. Corp., 1963, vols. 2, 4, 6, 7, 10, 1891-1896.
- Clebsch, R. F. A.: «Versuch einer Darlegung und Würdigung seiner wissenschaftlichen Leistungen». *Math. Ann.*, 7, 1874, 1-55. Un artículo escrito por amigos de Clebsch.
- Coolidge, Julián L.: *A History of Geometrical Methods*, Dover (reimp.), 1963, 195-230, 278-292.
- Cremona, Luigi: *Opere matematiche*, 3 vols., Ulrico Hoepli,

1914-1917.

- Hensel, Kurt, y Georg Landsberg: *Theorie der algebraischen Funktionen einer Variablen* (1902), Chelsea (reimp.), 1965, 694-702 en particular.
- Hilbert, David: *Gesammelte Abhandlungen*. Julius Springer, 1933, vol. 3.
- Klein, Félix: *Vorlesungen über die Entwicklung der Mathematik im 19 Jahrhundert*, 1, 155-166, 295-319; 2, 2-26; Chelsea (reimp.), 1950.
- Meyer, Franz W.: «Bericht über den gegenwärtigen Stand der Invariantentheorie». *Jahres. der Deut. Math.-Verein.*, 1, 1890-1891, 79-292.
- National Research Council: *Selected Topics in Algebraic Geometry*, Chelsea (reimp.), 1970.
- Noether, Emmy: «Die arithmetische Theorie der algebraischen Funktionen einer Veränderlichen in ihrer Beziehung zu den übrigen Theorien and zu der Zahlentheorie». *Jahres. der Deut. Math.-Verein.*, 28, 1919, 182-203.

Capítulo 40

La introducción del rigor en el análisis

Pero sería un serio error pensar que sólo se puede encontrar certeza en las demostraciones geométricas o en el testimonio de los sentidos.

A. L. Cauchy

Contenido:

1. *Introducción*
2. *Las funciones y sus propiedades*
3. *La derivada*
4. *La integral*
5. *La teoría de series infinitas*
6. *Las series de Fourier*
7. *La situación del análisis*

Bibliografía

1. Introducción

Hacia el año 1800, los matemáticos empezaron a interesarse por la imprecisión de los conceptos y demostraciones de vastas ramas del análisis. El propio concepto de función no estaba claro; el uso de las series sin referencia a la convergencia y divergencia había producido

paradojas y desacuerdos; la controversia sobre las representaciones de las funciones por medio de series trigonométricas había introducido mayor confusión; y, desde luego, las nociones fundamentales de derivada e integral nunca habían sido definidas correctamente. Todas estas dificultades causaron finalmente insatisfacción con el status lógico del análisis.

Abel, en una carta de 1826 al profesor Christoffer Hansteen,²⁰⁶ se quejaba de «la tremenda oscuridad que incuestionablemente encuentra uno en el análisis. Carece completamente de todo plan y sistema y el hecho de que tantos lo hayan podido estudiar es notable. Lo peor del caso es que nunca ha sido tratado rigurosamente. Hay muy pocos teoremas en el análisis superior que se hayan demostrado de una manera lógicamente sostenible. En todas partes encuentra uno esta manera miserable de concluir lo general partiendo de lo especial y es extremadamente peculiar que tal procedimiento haya llevado a tan pocas de las así llamadas paradojas».

Varios matemáticos se resolvieron a poner orden en todo este caos. Las cabezas de lo que frecuentemente es llamado el movimiento crítico decidieron reconstruir el análisis solamente sobre la base de los conceptos aritméticos. Los principios del movimiento coinciden con la creación de la geometría no euclídea. Un grupo totalmente diferente, excepto Gauss, se involucró en esta última actividad y, por tanto, es difícil rastrear alguna relación directa entre ella y la decisión de fundar el análisis en la aritmética. Tal vez se llegó a esa decisión porque la esperanza de fundar el análisis sobre la

geometría, lo cual muchos matemáticos del siglo XVII afirmaron con frecuencia que sí se podía hacer, fue desechado por la complejidad creciente de los desarrollos en análisis durante el siglo XVIII. Sin embargo, Gauss ya había expresado sus dudas en cuanto a la verdad de la geometría euclídea en 1799 y en 1817 decidió que la verdad residía únicamente en la aritmética. Más aún, incluso durante la obra inicial de Gauss y otros en la geometría no euclídea, ya se habían notado imperfecciones en el desarrollo de Euclides. Por tanto, es muy probable que ambos factores causaran desconfianza en la geometría y apresuraran la decisión de fundar el análisis sobre conceptos aritméticos. Esto fue ciertamente lo que las cabezas del movimiento crítico se comprometieron a llevar a cabo.

El análisis riguroso empieza con la obra de Bolzano, Cauchy, Abel y Dirichlet, y Weierstrass lo fomentó. Cauchy y Weierstrass son los más conocidos en esta relación. Las obras básicas de Cauchy sobre los fundamentos del análisis son: *Cours d'analyse algébrique* (Curso de Análisis Algebraico)²⁰⁷, *Résumé des leçons sur le calcul infinitesimal* (Resumen de las Lecciones sobre el Cálculo Infinitesimal)²⁰⁸ y *Leçons sur le calcul différentiel* (Lecciones sobre el Cálculo Diferencial).²⁰⁹ Realmente, el rigor de Cauchy en estas obras es impreciso de acuerdo con los criterios modernos. Usó frases tales como «se acerca indefinidamente», «tan poco como se desee», «últimas razones de incrementos infinitamente pequeños» y «una variable se acerca a su límite». Sin embargo, si uno compara la *Théorie des fonctions analytiques* (Teoría de las Funciones Analíticas)²¹⁰ de Lagrange y sus *Leçons sur le calcul des fonctions*

(Lecciones sobre el Cálculo de las Funciones)²¹¹ y el libro de Lacroix, que tanta influencia tuvo, *Traité du calcul différentiel et du calcul intégral* (Tratado de Cálculo Diferencial y de Cálculo Integral)²¹² junto con el *Cours d'analyse algébrique* de Cauchy, uno empieza a ver la sorprendente diferencia entre las matemáticas del siglo XVIII y las del XIX. Lagrange, en particular, era puramente formal. Operaba con expresiones simbólicas. Los conceptos subyacentes de límite, continuidad, etc. no están ahí.

Cauchy es muy explícito en su introducción a la obra de 1821 en cuanto a que busca dar rigor al análisis. Señala que el uso libre para todas las funciones de las propiedades que se cumplen para las funciones algebraicas y el uso de series divergentes no están justificados. Aunque la obra de Cauchy no fue sino un paso en la dirección del rigor, él mismo creía y establece en su *Résumé* que había llegado a lo último en rigor dentro del análisis. Sí proporcionó los comienzos de demostraciones precisas de teoremas y limitó las afirmaciones adecuadamente, al menos para las funciones elementales. Abel, en su artículo de 1826 sobre serie binomiales (sec. 5) alabó este logro de Cauchy: «La distinguida obra [el *Cours d'analyse*] debiera ser leída por todo aquel que ame el rigor dentro de las investigaciones matemáticas.» Cauchy abandonó las representaciones explícitas de Euler y las series de potencias de Lagrange e introdujo nuevos conceptos para tratar las funciones.

2. Las funciones y sus propiedades

Los matemáticos del siglo XVIII creían en su totalidad que una función debía tener la misma expresión analítica en todas partes. Durante la parte final del siglo, en gran medida como consecuencia de la controversia sobre el problema de la cuerda vibrante, Euler y Lagrange permitieron las funciones que tienen diferentes expresiones en diferentes dominios y usaron la palabra continua donde se mantenía la misma expresión y discontinua en puntos donde la expresión cambiaba de forma (aunque en el sentido moderno toda la función podía ser continua). Mientras Euler, D'Alembert y Lagrange tenían que reconsiderar el concepto de función, no llegaron a ninguna definición ampliamente aceptada y tampoco resolvieron el problema de qué funciones podían representarse por medio de series trigonométricas. Sin embargo, la expansión gradual en la variedad y uso de las funciones forzó a los matemáticos a aceptar un concepto más amplio.

Gauss, en su trabajo inicial, interpretaba por función una expresión cerrada (analítica finita) y cuando hablaba de la serie hipergeométrica $F(\alpha, \beta, \gamma, x)$ como una función de α , β , γ y x la explicaba señalando, «en tanto se la pueda tratar como función». Lagrange ya había usado un concepto más amplio al tratar las series de potencias como funciones. En la segunda edición de su *Mécanique analytique* (1811-1815) usó la palabra función para casi cualquier clase de dependencia de una o más variables. Aún Lacroix, en su *Traité* de 1797, había introducido ya una noción más amplia. En la introducción dice: «Toda cantidad cuyo valor depende de una o varias otras es llamada una función de estas últimas, ya

sea que uno conozca o no por medio de qué operaciones es necesario pasar de las últimas a la primera cantidad.» Lacroix da como un ejemplo de raíz de una ecuación de quinto grado como una función de sus coeficientes.

La obra de Fourier abrió aún más ampliamente la discusión sobre lo que es una función. Por un lado, insistió en que las funciones no necesitan ser representables por una expresión analítica. En su *Teoría Analítica del Calor*²¹³ dice: «En general la función $f(x)$ representa una sucesión de valores u ordenadas, cada una de las cuales es arbitraria. ... No suponemos que estas ordenadas estén sujetas a una ley común; se suceden una a la otra de cualquier manera, sea la que fuere...» De hecho, sólo trató funciones con un número finito de discontinuidades en cualquier intervalo finito. Por otro lado, hasta cierto grado, Fourier estaba apoyando la aseveración que una función debe ser representable por una expresión analítica, aunque esta expresión fuese una serie de Fourier. En cualquier caso, la obra de Fourier perturbó la creencia del siglo XVIII de que todas las funciones, en el peor de los casos, eran extensiones de funciones algebraicas. Las funciones algebraicas, y aun las funciones trascendentes elementales, ya no fueron el prototipo de funciones. Como las propiedades de las funciones algebraicas ya no podían cumplirse para todas las funciones, surgió entonces la cuestión en cuanto a lo que realmente se quiere decir por función, por continuidad, diferenciabilidad, integrabilidad y otras propiedades.

En la reconstrucción positiva del análisis, que muchos matemáticos se propusieron llevar a cabo, el sistema de los números reales se dio por sentado. No se hizo ningún intento por analizar esta estructura o por construirla lógicamente. Aparentemente, los matemáticos sentían que trabajaban sobre bases seguras en cuanto a lo concerniente a esta área.

Cauchy empieza su trabajo de 1821 con la definición de una variable. «Se llama variable a una cantidad que se considera tiene que tomar sucesivamente muchos valores diferentes unos de los otros.» En cuanto al concepto de función: «Cuando se relacionan cantidades variables entre ellas de modo que estando dado el valor de una de éstas, se puedan determinar los valores de todas las otras, ordinariamente se concibe a estas cantidades diversas expresadas por medio de la que está entre ellas, la cual entonces toma el nombre de variable independiente; y las otras cantidades expresadas por medio de la variable independiente son aquellas que uno llama funciones de esta variable.» Cauchy también es explícito en que una serie infinita es una manera de especificar una función. Sin embargo, no se requiere una expresión analítica para una función.

En un artículo sobre series de Fourier, al cual retornaremos posteriormente, «*Über die Darstellung ganz willkürlicher Functionen durch Sinus-und Cosinusreihen*» («Sobre la Representación de Funciones Completamente Arbitrarias por Series de Senos y Cosenos»),²¹⁴ Dirichlet dio la definición de una función (univaluada) que es la que más frecuentemente se emplea ahora, a saber, que y

es una función de x cuando a cada valor de x en un intervalo dado le corresponde un único valor de y . Agregó que no importa si en todo este intervalo y depende de x de acuerdo a una ley o más, o si la dependencia de y de x puede expresarse por medio de operaciones matemáticas. De hecho, en 1829²¹⁵ dio el ejemplo de una función de x que tiene el valor c para todos los valores racionales de x y el valor d para todos los valores irracionales de x .

Hankel señala que los mejores libros de texto de al menos la primera mitad del siglo no sabían qué hacer con el concepto de función. Algunos definían una función esencialmente en el sentido de Euler; otros requerían que y variara con x de acuerdo con alguna ley, pero no explicaban lo que quería decir ley; algunos usaban la definición de Dirichlet; y aún otros no dieron definición alguna. Pero todos dedujeron consecuencias a partir de sus definiciones que no estaban implicadas lógicamente por las definiciones.

La distinción adecuada entre continuidad y discontinuidad surgió gradualmente. El estudio cuidadoso de las propiedades de las funciones fue iniciado por Bernhard Bolzano (1781-1848), un sacerdote, filósofo y matemático de Bohemia. Bolzano fue llevado a este trabajo cuando trataba de proporcionar una demostración puramente aritmética del teorema fundamental del álgebra en lugar de la primera demostración de Gauss (1799), la cual usaba ideas geométricas. Bolzano tenía los conceptos correctos para el establecimiento del cálculo (a excepción de una teoría de los números reales), pero su trabajo pasó desapercibido durante medio siglo. Negó la existencia de números infinitamente pequeños

(infinitésimos) y de números infinitamente grandes, ambos habían sido utilizados por los autores del siglo XVIII. En un libro de 1817, cuyo largo título empieza con *Rein analytischer Beweis* (véase la bibliografía), Bolzano dio la definición apropiada de continuidad, a saber, $f(x)$ es continua en un intervalo si en cualquier x del intervalo la diferencia $f(x + w) - f(x)$ se puede hacer tan pequeña como se desee tomando w suficientemente pequeña. Prueba que los polinomios son continuos.

Cauchy, también, abordó las nociones de límite y continuidad. Como con Bolzano, el concepto de límite estaba basado en consideraciones puramente aritméticas. En el Cours (1821), Cauchy dice: «Cuando los valores sucesivos asignados a una variable se acercan indefinidamente a un valor fijo de modo que terminan por diferir de él por tan poco como se desee, este último es llamado el límite de los otros. Así, por ejemplo, un número irracional es el límite de diversas fracciones que toman valores cada vez más aproximados a él.» Este ejemplo fue un poco desafortunado porque muchos lo tomaron como una definición de los números irracionales en términos de límite, siendo así que el límite podría no tener significado si los irracionales no estuvieran ya presentes. Cauchy lo omitió en sus obras de 1823 y 1829.

En el prefacio a su trabajo de 1821, Cauchy dice que para hablar de la continuidad de las funciones debe dar a conocer las propiedades principales de las cantidades infinitamente pequeñas. «Se dice [Cours, pág. 5] que una cantidad variable se hace infinitamente pequeña cuando su valor numérico decrece indefinidamente de

modo tal que converge al límite 0.» Llama a tales variables infinitésimos. Así clarifica Cauchy la noción de infinitésimo de Leibniz y la libera de ataduras metafísicas. Cauchy continúa: «Se dice que una cantidad variable se hace infinitamente grande cuando su valor numérico se incrementa indefinidamente de manera tal que converge al límite 0.» Sin embargo, ∞ no significa una cantidad fija, sino algo indefinidamente grande.

Ahora, Cauchy está preparado para definir la continuidad de una función. En el Cours (págs. 34-35) dice: «Sea $f(x)$ una función de la variable x , y supóngase que, para cada valor de x que se encuentre entre dos límites [cotas] dados, esta función toma constantemente un valor finito y único. Si, empezando con un valor de x contenido entre estos límites, se asigna a la variable x un incremento infinitamente pequeño, la función misma tomará como incremento la diferencia $f(x + a) - f(x)$, el cual dependerá al mismo tiempo de la nueva variable a y del valor de x . Garantizado esto, la función $f(x)$ será, entre los dos límites asignados a la variable x , una función continua de la variable si, para cada valor de x que se encuentre entre estos dos límites, el valor numérico de la diferencia $f(x + a) - f(x)$ decrece indefinidamente con el de a . En otras palabras, la función $f(x)$ permanecerá continua con respecto a x entre los dos límites dados, si, entre estos límites, un incremento infinitamente pequeño de la variable produce siempre un incremento infinitamente pequeño de la función misma.

«También decimos que la función $f(x)$ es una función continua de x en la vecindad de un valor particular asignado a la variable x ,

siempre que sea continua [la función] entre esos dos límites de x , no importe cuán cercanos estén, los cuales encierren el valor en cuestión.» Entonces dice que $f(x)$ es discontinua en x si no es continua en todo el intervalo alrededor de x .

En el *Cours* (pág. 37) Cauchy afirmó que si una función de varias variables es continua en cada una por separado, entonces es una función continua de todas las variables. Esto no es correcto.

Durante todo el siglo XIX se exploró la noción de continuidad y los matemáticos aprendieron más acerca de ella, a veces produciendo resultados que los asombraban. Darboux dio un ejemplo de una función que tomaba todos los valores intermedios entre dos valores dados al pasar de $x = a$ a $x = b$, pero que no era continua. Así, una propiedad básica de las funciones continuas no es suficiente para asegurar la continuidad.²¹⁶

El trabajo de Weierstrass sobre la rigorización del análisis mejoró el de Bolzano, Abel y Cauchy. También buscó la manera de evitar la intuición y de basarse en conceptos aritméticos. Aunque realizó este trabajo durante los años 1841-1856, cuando era maestro de escuela de educación media, mucho de ello no se conoció hasta 1859 cuando empezó a sentar cátedra en la Universidad de Berlín.

Weierstrass atacó la frase «una variable se acerca a un límite», la cual, desafortunadamente, sugiere tiempo y movimiento. Interpreta una variable simplemente como una letra que se usa para denotar cualquier valor de los de un conjunto dado que se le puede asignar a la letra. Así se elimina la idea de movimiento. Una variable continua es tal que si x es cualquier valor del conjunto de valores de

la variable y δ cualquier número positivo, hay otros valores de la variable en el intervalo $(x_0 - \delta, x_0 + \delta)$.

Para eliminar la vaguedad en la frase «se hace y permanece menor que cualquier cantidad dada», que Bolzano y Cauchy usaron en sus definiciones de continuidad y límite de una función, Weierstrass dio la definición ahora aceptada de que $f(x)$ es continua en $x = x_0$ si dado cualquier número positivo ε , existe una δ tal que para toda x en el intervalo $|x - x_0| < \delta$, $|f(x) - f(x_0)| < \varepsilon$. Una función $f(x)$ tiene un límite L en $x = x_0$ si se cumple la misma afirmación, pero con L reemplazando a $f(x_0)$. Una función $f(x)$ es continua en un intervalo de valores x si es continua en cada x del intervalo.

Durante los años en que la noción de continuidad misma se estaba refinando, los esfuerzos por establecer el análisis rigurosamente exigían la demostración de muchos teoremas sobre funciones continuas que habían sido aceptados intuitivamente. Bolzano, en su publicación de 1817, buscó demostrar que si $f(x)$ es negativa para $x = a$ y positiva para $x = b$, entonces $f(x)$ tiene un cero entre a y b . Consideró la sucesión de funciones (para x fija)

$$F_1(x), F_2(x), F_3(x), \dots, F_n(x), \dots \quad (1)$$

e introdujo el teorema que afirma que si para n suficientemente grande podemos hacer la diferencia $F_{n+r} - F_n$ menor que cualquier cantidad dada, por grande que sea r , entonces existe una magnitud fija x tal que la sucesión se acerca cada vez más a X , y ciertamente tanto como se desee. Su determinación de la cantidad x fue algo

oscura, porque no tenía una teoría particular clara del sistema de los números reales y de los números irracionales sobre la cual basarse. Sin embargo, tuvo la idea de lo que ahora llamamos la condición de Cauchy para la convergencia de una sucesión (ver más adelante).

En el desarrollo de la demostración, Bolzano estableció la existencia de una cota superior mínima para un conjunto acotado de números reales. Su enunciado preciso es: si una propiedad M no se aplica a todos los valores de una cantidad variable x , sino a todos aquellos que son más pequeños que una cierta u siempre hay una cantidad U que es la mayor de las que se puede afirmar que toda x más pequeña posee la propiedad M . La esencia de la demostración de Bolzano de este lema consistió en dividir el intervalo acotado en dos partes y seleccionar una parte particular que contenga un número infinito de miembros del conjunto. En seguida repite el proceso hasta que finalmente encierra el número que es la mínima cota superior del conjunto dado de números reales. Este método fue usado por Weierstrass en los sesenta, con el debido tributo a Bolzano, para demostrar lo que ahora se llama el teorema de Bolzano-Weierstrass. Establece que para cualquier conjunto infinito acotado de puntos existe un punto tal que en toda vecindad de él hay puntos del conjunto.

Cauchy había usado sin prueba (en una de sus demostraciones de la existencia de raíces de un polinomio) la existencia de un mínimo de una función continua definida sobre un intervalo cerrado. Weierstrass, en sus clases en Berlín, demostró que para cualquier

función continua de una o más variables definidas sobre un dominio cerrado y acotado, existe un valor mínimo y un valor máximo de la función.

En su trabajo inspirado por las ideas de Georg Cantor y Weierstrass, Heine definió la continuidad uniforme para funciones de una o varias variables²¹⁷ y después demostró²¹⁸ que una función que es continua en un intervalo cerrado y acotado de los números reales es uniformemente continua. El método de Heine introdujo y usó el siguiente teorema: sean dados un intervalo cerrado $[a,b]$ y un conjunto infinito numerable A de intervalos, todos en $[a,b]$, tales que todo punto x de $a \leq x \leq b$ sea un punto interior de al menos uno de los intervalos de A . (Los extremos a y b se consideran como puntos interiores cuando a es el extremo izquierdo de un intervalo y b el extremo derecho de otro intervalo.) Entonces un conjunto que consiste en un número finito de intervalos de A tiene la misma propiedad, a saber, todo punto del intervalo cerrado $[a,b]$ es un punto interior de al menos uno de los intervalos de este conjunto finito (a y b pueden ser extremos).

Emile Borel (1871-1956), uno de los grandes matemáticos franceses de este siglo, reconoció la importancia de poder seleccionar un número finito de intervalos recubridores, y lo estableció primero como un teorema independiente para el caso de un conjunto original de intervalos A numerable.²¹⁹ Aunque muchos matemáticos alemanes y franceses se refieren a este teorema como el teorema de Borel, ya que Heine usó en su demostración la propiedad de la continuidad uniforme, el teorema también es conocido como

teorema de Heine-Borel. El mérito del teorema, como lo señaló Lebesgue, no radica en su demostración, la cual no es difícil, sino en percibir su importancia y enunciarlo como un teorema distinto. El teorema se aplica a conjuntos cerrados en cualquier número de dimensiones y es ahora básico en la teoría de los conjuntos.

La extensión del teorema de Heine-Borel para el caso en que puede seleccionarse un conjunto finito de intervalos recubridores de un conjunto infinito no numerable, se le atribuye generalmente a Lebesgue, quien reclamó haber conocido el teorema en 1898 y haberlo publicado en sus *Leçons sur l'intégration* (Lecciones sobre la integración, 1904). Sin embargo, fue publicado primero por Pierre Cousin (1867-1933) en 1895.²²⁰

3. La derivada

D'Alembert fue el primero en ver que Newton tenía esencialmente la noción correcta de la derivada. D'Alembert dice explícitamente en la *Encyclopédie* que la derivada debe basarse en el límite de la razón de las diferencias de variables dependientes e independientes. Esta versión es una reformulación de las razones primera y última de Newton. D'Alembert no avanzó más porque sus pensamientos aún estaban atados a la intuición geométrica. Sus sucesores de los cincuenta años siguientes todavía fracasaron en dar una definición clara de la derivada. Hasta Poisson creía que hay números positivos que no son cero, pero que son más pequeños que cualquier número dado por muy pequeño que sea.

Bolzano fue el primero (1817) en definir la derivada de $f(x)$ como la cantidad $f'(x)$ a la cual se acerca indefinidamente la razón $[f(x + \Delta x) - f(x)]/\Delta x$ conforme Δx se acerca a 0 con valores positivos y negativos. Bolzano insistió en que $f'(x)$ no era un cociente de ceros o una razón de cantidades evanescentes sino un número al cual tendía la razón anterior.

En su *Résumé des leçons* (Resumen de Lecciones),²²¹ Cauchy definió la derivada de la misma manera que Bolzano. Entonces unificó esta noción con la de diferenciales de Leibniz definiendo dx como cualquier cantidad finita y dy como $f'(x)dx$.²²² En otras palabras, se introducen dos cantidades dx y dy cuya razón, por definición, es $f'(x)$. Las diferenciales tienen significado en términos de la derivada y son meramente una noción auxiliar de la cual se podría prescindir lógicamente, pero son convenientes como una manera de pensar o escribir. Cauchy también señaló lo que significaban las expresiones diferenciales utilizadas durante el siglo XVIII en términos de derivadas.

A continuación clarificó la relación entre $\Delta y/\Delta x$ y $f'(x)$ a través del teorema del valor medio, esto es, $\Delta y = f'(x + \theta\Delta x)\Delta x$, donde $0 < \theta < 1$. El teorema mismo era conocido por Lagrange (cap. 20, sec. 7). La demostración de Cauchy del teorema del valor medio usaba la continuidad de $f'(x)$ en el intervalo Δx .

Aunque Bolzano y Cauchy habían rigorizado (de alguna manera) las nociones de continuidad y de derivada, Cauchy y casi todos los matemáticos de su época creyeron, y muchos textos «demostraron» durante los siguientes cincuenta años, que una función continua

debe ser diferenciable (excepto desde luego en puntos aislados tales como $x = 0$ para $y = 1/x$). Bolzano sí entendió la distinción entre continuidad y diferenciabilidad. En su *Funktionenlebre* (Lecciones de Funciones), que escribió en 1834, pero que no terminó ni publicó,²²³ dio un ejemplo de una función continua que no tiene derivada finita en ningún punto. El ejemplo de Bolzano, al igual que sus otros trabajos, pasó desapercibido.²²⁴ Aún si hubiese sido publicado en 1834, probablemente no hubiera causado impresión, porque la curva no tenía una representación analítica, y para los matemáticos de ese período las funciones todavía eran entidades dadas por expresiones analíticas.

El ejemplo que finalmente hizo ver de manera cabal la distinción entre continuidad y diferenciabilidad fue dado por Riemann en el *Habilitationsschrift* el artículo de 1854 que escribió para calificarse como Privatdozent en Göttingen, «*Über die Darstellbarkeit einer Function durch eine trigonometrische Reihe*» («Sobre la representabilidad de una función mediante una serie trigonométrica».²²⁵ (El artículo sobre los fundamentos de la geometría (cap. 37, sec. 3) fue dado como una conferencia de concurso.) Riemann definió la siguiente función: denótese con (x) la diferencia entre x y el entero más cercano y sea $(x) = 0$ si está a la mitad entre dos enteros consecutivos. Entonces $-1/2 < (x) < 1/2$. Ahora $f(x)$ se define como

$$f(x) = \frac{(x)}{1} + \frac{(2x)}{4} + \frac{(3x)}{9} + \dots$$

Esta serie converge para todos los valores de x . Sin embargo, para $x = \rho/2n$, donde ρ es un entero primo con $2n$, $f(x)$ es discontinua y tiene un salto cuyo valor es $\pi^2/8n^2$. En todos los otros valores de x , $f(x)$ es continua. Más aún, $f(x)$ es discontinua un número infinito de veces en todo intervalo arbitrariamente pequeño. Sin embargo, $f(x)$ es integrable (sec. 4). Lo que es más, $F(x) = \int f(x)dx$ es continua para toda x , pero no tiene derivada donde $f(x)$ es discontinua. Esta función patológica no atrajo mucha atención hasta que fue publicada en 1868.

Una distinción aún más sorprendente entre continuidad y diferenciabilidad fue mostrada por el matemático suizo Charles Cellierier (1818-1889). En 1860 dio un ejemplo de una función que es continua pero en ningún sitio diferenciable, a saber,

$$f(x) = \sum_{n=1}^{\infty} a^{-n} \text{sen } a^n x$$

en la cual a es un entero positivo grande. Sin embargo, esto no se publicó hasta 1890.²²⁶ El ejemplo que atrajo la mayor atención se debe a Weierstrass. Remontándonos hasta 1861, había afirmado en sus conferencias que cualquier intento por demostrar que la diferenciabilidad se sigue de la continuidad debe ser fallido. Dio entonces, en una conferencia a la Academia de Berlín el 18 de julio de 1872, el ejemplo clásico de una función continua en ninguna parte diferenciable.²²⁷ Weierstrass comunicó su ejemplo en una

carta de 1874 a Du Bois-Reymond y el ejemplo fue publicado primero por este último.²²⁸ La función de Weierstrass es

$$f(x) = \sum_{n=0}^{\infty} b^n \cos(a^n \pi x)$$

donde a es un entero impar y b una constante positiva menor que 1 tal que $ab > 1 + (3\pi/2)$. La serie es uniformemente convergente y, por tanto, define una función continua. El ejemplo proporcionado por Weierstrass apresuró la creación de muchas más funciones que son continuas en un intervalo o en todas partes pero que no son diferenciables ya sea en un conjunto denso de puntos o en cualquier punto.²²⁹

La significación histórica del descubrimiento de que la continuidad no implica la diferenciabilidad y de que las funciones pueden tener todos los tipos de comportamiento anormal fue grandiosa. Hizo que los matemáticos fueran más aprensivos al confiar en la intuición o en el pensamiento geométrico.

4. La integral

La obra de Newton mostró cómo se podía calcular áreas invirtiendo la diferenciación. Desde luego que, aun ahora, este es el método esencial. La idea de Leibniz de área y volumen como una «suma» de elementos tales como rectángulos o cilindros [la integral definida] fue desdeñada. Cuando se empleó a fondo este último concepto en el siglo XVIII se usó de manera imprecisa.

Cauchy hizo hincapié en definir la integral como el límite de una suma, en lugar de considerarla como la inversa de la diferenciación. Hubo al menos una razón principal para el cambio. Fourier, como sabemos, trató con funciones discontinuas, y la fórmula para los coeficientes de una serie de Fourier, a saber,

$$a_n = \frac{1}{\pi} \int_0^{2\pi} f(x) \cos nx \, dx \quad b_n = \frac{1}{\pi} \int_0^{2\pi} f(x) \operatorname{sen} nx \, dx$$

necesita de las integrales de tales funciones. Fourier consideró la integral como una suma (el punto de vista leibniziano) y así no tuvo dificultad para manejar incluso $f(x)$ discontinuas. Sin embargo, se tuvo que considerar el problema del significado analítico de la integral cuando $f(x)$ es discontinua.

El ataque más sistemático de Cauchy sobre la integral definida fue hecho en su *Résumé* (1823), en donde también señala que es necesario establecer la existencia de la integral definida e indirectamente de la antiderivada o función primitiva antes de que puedan usarse. Comienza con funciones continuas.

Da la definición precisa de la integral para $f(x)$ continuas²³⁰ como el límite de una suma. Si el intervalo $[x, X]$ se subdivide por los valores de x : x_1, x_2, \dots, x_{n-1} , con $x_n = X$, entonces la integral es

$$\lim_{n \rightarrow \infty} \sum_{i=1}^n f(\xi_i)(x_i - x_{i-1})$$

donde ξ es cualquier valor de x en $[x_{i-1}, x_i]$. La definición presupone que $f(x)$ es continua sobre $[x_0, X]$ y que la longitud del subintervalo más grande tiende a cero. La definición es aritmética. Cauchy demuestra que la integral existe, no importa cómo se escojan los x_i y ξ_i , sin embargo, su prueba no era rigurosa porque no poseía la noción de continuidad uniforme. Denota el límite usando la notación propuesta por Fourier

$$\int_{\underline{x_0}}^x f(x) dx$$

en lugar de

$$\int f(x) dx \left[\begin{array}{l} x = b \\ x = b \end{array} \right]$$

empleada frecuentemente por Euler para la antidiferenciación. Cauchy define entonces

$$F(x) = \int_{x_n}^x f(x) dx$$

y muestra que $F(x)$ es continua en $[x_0, X]$. Formando

$$\frac{F(x+h) - F(x)}{h} = \frac{1}{h} \int_x^{x+h} f(x) dx$$

y usando el teorema del valor medio para integrales, Cauchy prueba que

$$F'(x) = f(x).$$

Este es el teorema fundamental del cálculo, y la presentación que hace Cauchy es la primera demostración de él. Entonces, después de demostrar que todas las primitivas de una $f(x)$ dada difieren en una constante, define la integral indefinida como

$$\int f(x) dx = \int_a^x f(x) dx + C$$

y señala que

$$\int_a^b f'(x) dx = f(b) - f(a)$$

presupone que $f'(x)$ es continua. Después, Cauchy trata las integrales singulares (impropias), donde $f(x)$ se hace infinita en algún valor de x en el intervalo de integración, o donde el intervalo de integración se extiende hasta ∞ . Para el caso en que $f(x)$ tiene

una discontinuidad en $x = c$ en cuyo valor $f(x)$ puede ser acotada o no, Cauchy define

$$\int_a^b f(x) dx = \lim_{\varepsilon_1 \rightarrow 0} \int_a^{c-\varepsilon_1} f(x) dx + \lim_{\varepsilon_2 \rightarrow 0} \int_{c+\varepsilon_2}^b f(x) dx$$

cuando estos límites existen. Cuando $\varepsilon_1 = \varepsilon_2$ obtenemos lo que Cauchy llamó el valor principal.

Las nociones de área acotada por una curva, longitud de una curva, volumen acotado por superficies y áreas de superficies se habían aceptado como entendidas intuitivamente, y se había considerado uno de los grandes logros del cálculo que estas cantidades pudiesen calcularse por medio de integrales. Pero Cauchy, teniendo presente su meta de aritmetizar el análisis, definió estas cantidades geométricas por medio de las integrales que habían sido formuladas para calcularlas. Cauchy, involuntariamente, impuso una limitación en los conceptos que definió debido a que las fórmulas del cálculo imponen restricciones sobre las cantidades involucradas. Así, la fórmula para la longitud de arco de una curva dada por $y = f(x)$ es

$$s = \int_a^b \sqrt{1 + (y')^2} dx$$

y esta fórmula presupone la diferenciabilidad de $f(x)$. Posteriormente iba a surgir la cuestión de cuáles son las definiciones más generales

de áreas, longitudes de curvas, y volúmenes (cap. 42, sec. 5).

Cauchy había demostrado la existencia de una integral para cualquier integrando continuo. También había definido la integral cuando el integrando tiene discontinuidades de salto e infinitas. Pero con el desarrollo del análisis se hizo manifiesta la necesidad de considerar integrales de funciones de comportamiento más irregular. El tema de la integrabilidad fue retomado por Riemann en su artículo en 1854 sobre series trigonométricas. Dice que, al menos para las matemáticas, es importante, aunque no para las aplicaciones físicas, considerar las condiciones más amplias bajo las cuales se cumple la fórmula integral para los coeficientes de Fourier.

Riemann generalizó la integral para abarcar funciones $f(x)$ definidas y acotadas en un intervalo $[a,b]$. Descompone este intervalo en subintervalos²³¹ $\Delta x_1, \Delta x_2, \dots, \Delta x_n$ y define la oscilación de $f(x)$ en Δx cómo la diferencia entre el valor más grande y el más pequeño de $f(x)$ en Δx . Entonces demuestra que una condición necesaria y suficiente para que las sumas

$$S = \sum_{i=1}^n f(x_i) \Delta x_i$$

donde x_i es cualquier valor de x en Δx_i , tiendan a un único límite (que la integral exista) cuando el máximo Δx_i tiende a cero es que la suma de los intervalos Δx_i en los que la oscilación de $f(x)$ es más grande que cualquier número dado λ debe tender a cero con el

tamaño de los intervalos.

Riemann señala entonces que esta condición sobre las oscilaciones le permite reemplazar las funciones continuas por funciones con discontinuidades aisladas y también por funciones que tengan un conjunto denso en todas partes de puntos de discontinuidad. De hecho, el ejemplo que dio de una función integrable con un número infinito de discontinuidades en todo intervalo arbitrariamente pequeño (sec. 3) lo ofreció para ilustrar la generalidad de su concepto de integral. Así, Riemann prescindió de la continuidad y continuidad a trozos en la definición de la integral.

En su artículo de 1854, Riemann, sin observaciones adicionales, proporciona otra condición necesaria y suficiente para que una función acotada $f(x)$ sea integrable en $[a,b]$. Requiere que primero se establezca lo que ahora se llaman las sumas superior e inferior

$$S = M_1\Delta x_1 + \dots + M_n\Delta x_n$$

$$s = m_1\Delta x_1 + \dots + m_n\Delta x_n$$

donde m_i y M_i son los valores mínimo y máximo de $f(x)$ en Δx . Después, haciendo $D_i = M_i - m_i$ Riemann afirma que la integral de $f(x)$ en $[a,b]$ existe si y sólo si

$$\lim_{\max \Delta x \rightarrow 0} \{D_1\Delta x_1 + D_2\Delta x_2 + \dots + D_n\Delta x_n\} = 0$$

para todas las elecciones de Δx_i que llenen el intervalo $[a,b]$,

Darboux completó esta formulación y demostró que la condición es necesaria y suficiente.²³² Existen muchos valores de S correspondiendo cada uno de ellos a una partición de $[a,b]$ en Δx_i . De manera análoga, existen muchos valores de s . Cada S se llama una suma superior y cada s una suma inferior. Sea J la máxima cota inferior de las S e I la mínima cota superior de las s . Se sigue que $I \leq J$. El teorema de Darboux establece entonces que las sumas S y s tienden respectivamente a J e I cuando se incrementa indefinidamente el número de Δx_i de tal manera que el máximo subintervalo tiende a cero. Se dice que una función acotada es integrable en $[a,b]$ si $J = I$.

Darboux demuestra a continuación que una función acotada será integrable sobre $[a,b]$ si y sólo si las discontinuidades de $f(x)$ constituyen un conjunto de medida cero. Por esto último quería dar a entender que los puntos de discontinuidad pueden encerrarse en un conjunto finito de intervalos cuya longitud total es arbitrariamente pequeña. Esta misma formulación de la condición de integrabilidad fue establecida por otros autores en el mismo año (1875). Los términos integral superior y la notación

$$\int_a^{\bar{b}} f(x) dx$$

para la máxima cota inferior J de las S , e integral inferior y la notación

$$\int_a^b f(x) dx$$

para la mínima cota superior I de las s fueron introducidos por Volterra.²³³

Darboux también demostró en el artículo de 1875 que el teorema fundamental del cálculo se cumple para funciones integrables en el sentido ampliado. Bonnet había dado una demostración del teorema del valor medio del cálculo diferencial que no utilizaba la continuidad de $f(x)$.²³⁴ Darboux, usando esta demostración, la cual es ahora habitual, demostró que

$$\int_a^b f'(x) dx = f(b) - f(a)$$

cuando f es integrable solamente en el sentido Riemann-Darboux. El argumento de Darboux fue que

$$f(b) - f(a) = \sum_{i=1}^n f(x_i) - f(x_{i-1})$$

donde $a = x_0 < x_1 < x_2 < \dots < x_n = b$. Por el teorema del valor medio

$$\sum f(x_i) - f(x_{i-1}) = \sum f'(t_i)(x_i - x_{i-1})$$

donde t_i es algún valor en x). Ahora, si el máximo (x_{i-1}, x_i) . Ahora si Δx_i o (x_i, x_{i-1}) tiende a cero, entonces el miembro de la derecha de esta última ecuación tiende a

$$\int_a^b f'(x) dx$$

y el miembro izquierdo es $f(b) - f(a)$

Una de las actividades favoritas de los setenta y ochenta fue construir funciones con diversos conjuntos infinitos de discontinuidades que aún así fueran integrables en el sentido de Riemann. En este sentido, H. J. S. Smith²³⁵ dio el primer ejemplo de una función no integrable en el sentido de Riemann, pero para la cual los puntos de discontinuidad eran «raros». La función de Dirichlet (sec. 2) también es no integrable en este sentido, pero es discontinua en todas partes.

La noción de integración se extendió entonces a funciones no acotadas y a varias integrales impropias. La extensión más significativa fue llevada a cabo en el siguiente siglo por Lebesgue (cap. 44). Sin embargo, para 1875, en lo que concernía al cálculo elemental, la noción de integral era suficientemente amplia y estaba rigurosamente fundamentada.

También se abordó la teoría de las integrales dobles. Los casos más sencillos habían sido tratados en el siglo XVIII (cap. 19, sec. 6). En su artículo de 1814 (cap. 27, sec. 4), Cauchy demostró que el orden

de integración en el que se calcula una integral doble $\iint f(x,y) dx dy$ sí es importante si el integrando es discontinuo en el dominio de integración. Cauchy señaló específicamente²³⁶ que las integrales repetidas

$$\int_0^1 dy \left(\int_0^1 f(x,y) dx \right) \quad \int_0^1 dx \left(\int_0^1 f(x,y) dy \right)$$

no necesitan ser iguales cuando f es no acotada.

Karl J. Thomae (1840-1921) extendió la teoría de Riemann sobre integración a funciones de dos variables.²³⁷ Entonces Thomae, en 1878,²³⁸ proporcionó un ejemplo sencillo de una función acotada para la cual la segunda integral repetida existe pero la primera carece de significado.

En los ejemplos de Cauchy y Thomae la integral doble no existe. Pero en 1883,²³⁹ Du Bois-Reymond demostró que, aun cuando la integral doble exista, las dos integrales repetidas no necesariamente existen. En el caso de las integrales dobles también la generalización más significativa fue realizada por Lebesgue.

5. La teoría de series infinitas

Los matemáticos del siglo XVIII usaron las series indiscriminadamente. A finales del siglo, algunos resultados dudosos o llanamente absurdos del trabajo con series infinitas estimularon las investigaciones sobre la validez de las operaciones con ellas. Alrededor de 1810, Fourier, Gauss y Bolzano iniciaron el

manejo exacto de las series infinitas. Bolzano enfatizó que se debe considerar la convergencia y, en particular, criticó la demostración imprecisa del teorema del binomio. Abel fue el crítico más franco de los usos antiguos de las series.

En su artículo de 1811, y su Teoría analítica del calor, Fourier dio una definición satisfactoria de la convergencia de una serie infinita, aunque en general trabajó libremente con series divergentes. En el libro (pág. 196 de la edición inglesa) describe la convergencia dando a entender que conforme n se incrementa, la suma de n términos tiende a un valor fijo de manera cada vez más próxima y debiera diferir de él solamente por una cantidad que se haga menor que cualquier magnitud dada. Aún más, reconoció que la convergencia de una serie de funciones puede obtenerse solamente en un intervalo de valores x . También enfatizó que una condición necesaria para la convergencia es que los términos tiendan a cero. Sin embargo, la serie $1 - 1 + 1 + \dots$ aún lo confundía; tomó como su suma $1/2$.

La primera investigación importante y estrictamente rigurosa de la convergencia fue hecha por Gauss en su artículo de 1812 «*Disquisitiones Generales Circa Seriem Infinitam*» (Investigaciones generales de las series infinitas)²⁴⁰ en donde estudió la serie hipergeométrica $F(\alpha, \beta, \gamma, x)$. En la mayor parte de su obra llamó convergente a una serie si los términos a partir de un α determinado decrecen hacia cero. Pero en su artículo de 1812 hizo notar que este no es el concepto correcto. Debido a que la serie hipergeométrica puede representar a muchas funciones para diferentes elecciones de

α , β y γ , le parecía deseable desarrollar un criterio exacto para la convergencia de esta serie. Se llega a ese criterio de manera laboriosa, pero clarifica la cuestión de la convergencia para los casos que fue destinada a cubrir. Demostró que la serie hipergeométrica converge para x real y compleja si $|x| < 1$ y diverge si $|x| > 1$. Para $x = 1$, la serie converge si y solo si $\alpha + \beta < \gamma$ y para $x = -1$ la serie converge si y sólo si $\alpha + \beta < \gamma + 1$. El rigor inusual desanimó el interés de los matemáticos de la época por el artículo. Lo que es más, Gauss tenía interés en las series particulares y no consideró los principios generales de la convergencia de las series. Aunque Gauss es mencionado frecuentemente como uno de los primeros en reconocer la necesidad de restringir el uso de las series a sus dominios de convergencia, evitó cualquier posición definitiva. Estaba tan interesado en resolver problemas concretos mediante cálculos numéricos que hizo uso del desarrollo de Stirling de la función gamma. Cuando investigó la convergencia de la serie hipergeométrica en 1812 indicó,²⁴¹ que lo hizo así, no para dar su propia postura sobre el tema, sino para complacer a aquellos que favorecían el rigor de los geómetras antiguos. En la presentación de su artículo²⁴² usó el desarrollo de $\log(2 - 2 \cos x)$ en cosenos de múltiplos de x a pesar de que no había demostración de la convergencia de esta serie y no podía haber demostración con las técnicas disponibles en la época. En su obra astronómica y geodésica, Gauss, al igual que los matemáticos del siglo XVIII, siguió la práctica de usar un número finito de términos de una serie infinita y desdeñar los demás. Dejaba de incluir términos cuando

veía que los términos siguientes eran numéricamente pequeños y, por supuesto, no estimaba el error.

Poisson tomó una posición particular. Rechazó las series divergentes²⁴³ y hasta dio ejemplos de cómo los cálculos con series divergentes pueden conducir a resultados falsos. Pero de todas maneras hizo uso extensivo de las series divergentes en su representación de funciones arbitrarias por medio de series de funciones trigonométricas y esféricas.

Bolzano, en su publicación de 1817, tenía la noción correcta de la condición para la convergencia de una sucesión, la condición ahora atribuida a Cauchy. Bolzano también tenía nociones claras y correctas acerca de la convergencia de las series. Pero, como ya hemos notado, su obra no llegó a ser ampliamente conocida.

El trabajo de Cauchy sobre la convergencia de series es el primer tratamiento extensivo significativo del tema. En su *Cours d'analyse* Cauchy dice: «Sea

$$S_n = u_0 + u_1 + u_2 + \dots + u_{n-1}$$

la suma de los n primeros términos [de la serie infinita que se considera], designando n un número natural. Si, para valores constantemente crecientes de n , la suma s_n se acerca indefinidamente a cierto límite s , la serie es llamada convergente, y el límite en cuestión será llamado la suma de la serie.²⁴⁴ Por el contrario, si mientras n se incrementa indefinidamente, la suma s_n no se acerca a un límite fijo, la serie será llamada divergente y no

tendrá suma.

Después de definir la convergencia y la divergencia, Cauchy establece (*Cours*, p. 125) su criterio para la convergencia, a saber, una sucesión (S_n) converge a un límite S si y sólo si $S_{n+r} - S_n$ puede hacerse en valor absoluto menor que cualquier cantidad asignable para todo r y n suficientemente grandes. Cauchy demuestra que esta condición es necesaria pero únicamente dice que si la condición se satisface, se asegura la convergencia de la sucesión. Carecía del conocimiento de las propiedades de los números reales para hacer la demostración.

Cauchy establece enseguida y demuestra casos específicos para la convergencia de series con términos positivos. Señala que u_n debe tender a cero. Otra prueba (*Cours*, 132-135) requiere que se encuentre el límite o límites hacia los cuales la expresión $(u_n)^{1/n}$ tiende conforme n se hace infinito, y designa el mayor de estos límites por k . Entonces la serie será convergente si $k < 1$ y divergente si $k > 1$. También proporciona la prueba del cociente que emplea

$$\lim_n u_{n+1}/u_n$$

Si este límite es menor que 1, la serie converge y si es mayor que 1, la serie diverge. Se proporcionan casos especiales si la razón es 1. Siguen criterios de comparación y un criterio logarítmico. Demuestra que la suma $u_n + v_n$ de dos series convergentes converge a la suma de las sumas separadas y el resultado análogo para el

producto. Las series con algunos términos negativos, demuestra Cauchy, convergen cuando la serie de los valores absolutos de los términos converge, y después deduce la prueba de Leibniz para series alternadas.

Cauchy también considera la suma de una serie

$$\sum u_n(x) = u_1(x) + u_2(x) + u_3(x) + \dots$$

en la que todos los términos son funciones continuas reales unívocas. Los teoremas sobre la convergencia de las series de términos constantes se aplican aquí para determinar un intervalo de convergencia. También considera series cuyos términos sean funciones complejas.

Lagrange fue el primero en enunciar el teorema de Taylor con resto, pero Cauchy, en sus textos de 1823 y 1829 hizo la observación importante de que la serie de Taylor infinita converge a la función de la cual se obtiene si el residuo tiende a cero. Proporciona el ejemplo

$$e^{-x^2} + e^{-1/x^2}$$

de una función cuya serie de Taylor no converge a la función. En su texto de 1823, da el ejemplo e^{-1/x^2} de una función que tiene todas las derivadas en $x = 0$ pero no tiene expansión de Taylor alrededor de $x = 0$. Aquí, por medio de un ejemplo, contradice la afirmación de Lagrange en su *Théorie des Fonctions* (cap. V, art. 30) de que si $f(x)$ tiene todas las derivadas en x_0 , entonces puede expresarse como

una serie de Taylor que converge a $f(x)$ para x cercana a x_0 . Cauchy también dio²⁴⁵ una forma alternativa para el residuo en la fórmula de Taylor.

Aquí cometió Cauchy algunos errores adicionales con respecto al rigor. En su *Cours d'analyse* (pp. 131-132) establece que $F(x)$ es continua si cuando

$$F(x) = \sum_1^{\infty} u_n(x)$$

la serie es convergente y las $u_n(x)$ son continuas. En su *Résumé des leçons*,²⁴⁶ dice que si las $u_n(x)$ son continuas y la serie converge, entonces la serie se puede integrar término a término; esto es,

$$\int_a^b F dx = \sum_1^{\infty} \int_a^b u_n dx$$

Pasó por alto la necesidad de la convergencia uniforme. También afirma que para funciones continuas²⁴⁷

$$\frac{\partial}{\partial u} \int_a^b f(x, u) dx = \int_a^b \frac{\partial f}{\partial u} dx$$

La obra de Cauchy inspiró a Abel. Escribiendo de París a su primer maestro Holmboé en 1826, Abel dijo²⁴⁸ que Cauchy «es en el

presente quien conoce cómo deben ser tratadas las matemáticas». En ese año,²⁴⁹ Abel investigó el dominio de convergencia de la serie binómica

$$1 + mx + \frac{m(m-1)}{2}x^2 + \frac{m(m-1)(m-2)}{3!}x^3 + \dots$$

con m y x complejos, y expresó admiración de que nadie hubiese investigado previamente la convergencia de esta importantísima serie. Primero demuestra que si la serie

$$f(\alpha) = v_0 + v_1\alpha + v_2\alpha^2 + \dots$$

en donde las v , son constantes y α es real, converge para un valor δ de α entonces convergerá para todo valor de α más pequeño, y $f(\alpha - \beta)$ para β acercándose a 0 tenderá a $f(\alpha)$ cuando α sea igual o menor que δ . La última parte dice que una serie de potencias convergente es una función continua de su argumento hasta δ inclusive, pues α puede ser δ .

En este mismo artículo de 1826,²⁵⁰ Abel corrigió el error de Cauchy sobre la continuidad de la suma de una serie convergente de funciones continuas. Dio el ejemplo de

$$\operatorname{sen} x - \frac{\operatorname{sen} 2x}{2} + \frac{\operatorname{sen} 3x}{3} \dots \quad (2)$$

que es discontinua cuando $x = (2n + 1)\pi$ y n es entero, aunque los

términos individuales son continuos.²⁵¹ Entonces, usando la idea de convergencia uniforme dio una demostración correcta de que la suma de una serie uniformemente convergente de funciones continuas es continua en el interior del intervalo de convergencia. Abel no aisló la propiedad de la convergencia uniforme de una serie. La noción de convergencia uniforme de una serie

$$\sum_1^{\infty} u_n(x)$$

requiere que dado cualquier ε , exista un N tal que para todo $n > N$

$$\left| S(x) - \sum_1^{\infty} u_n(x) \right| < \varepsilon$$

para todo x en algún intervalo. $S(x)$ es por supuesto la suma de la serie. Esta noción fue reconocida en y por sí misma por Stokes, un físico matemático sobresaliente,²⁵² e independientemente por Philipp L. Seidel (1821-1896).²⁵³ Ninguno de los dos proporcionó la formulación precisa. Más bien ambos demostraron que si la suma de una serie de funciones continuas es discontinua en x_0 entonces existen valores de x cercanos a x_0 para los cuales la serie converge de manera arbitrariamente lenta. Tampoco relacionaron la necesidad de la convergencia uniforme para la justificación de integrar una serie término a término. De hecho, Stokes aceptó²⁵⁴ el uso de Cauchy de la integración término a término. Finalmente,

Cauchy reconoció la necesidad de la convergencia uniforme ²⁵⁵ para asegurar la continuidad de la suma de una serie de funciones continuas pero, incluso él, en esa época, no vio el error en el uso de la integración de una serie término a término.

De hecho, Weierstrass²⁵⁶ poseía la noción de convergencia uniforme tan temprano como en 1842. En un teorema que, sin saberlo, duplica el teorema de Cauchy sobre la existencia de las soluciones en series de potencias de un sistema de ecuaciones diferenciales ordinarias de primer orden, afirma que las series convergen uniformemente y así constituyen funciones analíticas de la variable compleja. Aproximadamente por la misma época, Weierstrass usó la noción de convergencia uniforme para proporcionar condiciones para la integración de una serie término a término y condiciones para la diferenciación bajo el signo de integral.

A través del círculo de estudiantes de Weierstrass se dio a conocer la importancia de la convergencia uniforme. Heine enfatizó la noción en un artículo sobre series trigonométricas.²⁵⁷ Es posible que Heine haya aprendido la idea por medio de Georg Cantor, quien había estudiado en Berlín y después fue en 1867 a Halle, donde Heine era profesor de matemáticas.

Durante sus años como maestro de escuela de educación media, Weierstrass también descubrió que cualquier función continua sobre un intervalo cerrado del eje real se puede expresar en ese intervalo como una serie de polinomios absoluta y uniformemente convergente. Weierstrass incluyó también funciones de varias variables. Este resultado ²⁵⁸ despertó considerable interés y se

establecieron muchas extensiones de él para la representación de funciones complejas por medio de una serie de polinomios o una serie de funciones racionales en el último cuarto del siglo XIX.

Se había supuesto que los términos de una serie se pueden ordenar como se desee. En un artículo de 1837,²⁵⁹ Dirichlet demostró que en una serie absolutamente convergente es posible agrupar o reacomodar los términos y no cambiar la suma. También proporcionó ejemplos para mostrar que los términos de cualquier serie condicionalmente convergente se pueden ordenar de modo que la suma se altere. Riemann, en un artículo escrito en 1854 (véase más adelante), demostró que por medio de reordenaciones adecuadas de los términos, la suma podía ser igual a cualquier número dado. Los grandes matemáticos desarrollaron muchos más criterios para la convergencia de series infinitas durante los años treinta y todo el resto del siglo.

6. Las series de Fourier

Como sabemos, el trabajo de Fourier demostró que una gran variedad de funciones puede ser representada por medio de series trigonométricas. Quedó abierto el problema de encontrar condiciones precisas para las funciones que poseen una serie de Fourier convergente. Los esfuerzos de Cauchy y Poisson fueron infructuosos.

Dirichlet se interesó por las series de Fourier después de conocer a Fourier en París durante los años 1822-1825. En un artículo básico, «*Sur la convergence des séries trigonométriques*» (Sobre la

convergencia de las series trigonométricas),²⁶⁰ Dirichlet dio el primer conjunto de condiciones suficientes para que la serie de Fourier representando una $f(x)$ dada converja y lo haga a $f(x)$. La demostración dada por Dirichlet es un refinamiento de la que bosquejó Fourier en las secciones finales de su Teoría analítica del calor. Considérese $f(x)$ dada, ya sea periódica con período 2π o dada en el intervalo $[-\pi, \pi]$ y definida periódica en cada intervalo de longitud 2π a la izquierda y a la derecha de $[-\pi, \pi]$. Las condiciones de Dirichlet son:

$f(x)$ es unívoca y acotada.

- a. $f(x)$ es continua a trozos; esto es, sólo tiene un número finito de discontinuidades en el período (cerrado).
- b. $f(x)$ es monótona a trozos; esto es, sólo tiene un número finito de máximos y mínimos en un período.
- c. La $f(x)$ puede tener diferentes representaciones analíticas en partes diferentes del período fundamental.

El método de demostración de Dirichlet consistía en hacer una sumatoria directa de n términos e investigar lo que sucede conforme n tiende a infinito. Demostró que para cualquier valor dado de x la suma de la serie es $f(x)$ siempre que $f(x)$ sea continua en ese valor de x y es

$$(1/2)[f(x-0) + f(x+0)]$$

si $f(x)$ es discontinua en ese valor de x .

En su demostración, Dirichlet se vio forzado a discutir cuidadosamente los valores límites de las integrales

$$\int_0^a f(x) \frac{\text{sen } \mu x}{\text{sen } x} dx \quad a > 0$$

$$\int_0^b f(x) \frac{\text{sen } \mu x}{\text{sen } x} dx \quad b > a > 0$$

conforme μ crece indefinidamente. Aún se les llama a éstas integrales de Dirichlet.

Fue en relación con este trabajo cuando proporcionó la función que vale c para valores racionales de x y d para valores irracionales de x (sec. 2). Había esperado generalizar la noción de integral de modo que aún fuera posible representar una clase más amplia de funciones por medio de una serie de Fourier convergente a estas funciones, pero la función particular dada se pensó como un ejemplo de una función que no pudiera ser incluida en una noción más amplia de integral.

Riemann estudió durante un tiempo bajo la dirección de Dirichlet en Berlín y adquirió interés por las series de Fourier. En 1854 emprendió la investigación del tema en su Habilitationsschrift en Göttingen,²⁶¹ «*Über die Darstellbarkeit einer Function durch eine trigonometrische Reihe*», que tenía como finalidad encontrar condiciones necesarias y suficientes que una función debe

satisfacer para que la serie de Fourier para $f(x)$ en un punto x del intervalo $[-\pi, \pi]$ convergiera a $f(x)$.

Riemann sí demostró el teorema fundamental de que si $f(x)$ es acotada e integrable en $[-\pi, \pi]$ entonces la coeficientes de Fourier

$$a_n = \frac{1}{\pi} \int_{-\pi}^{\pi} f(x) \cos nx \, dx \quad b_n = \frac{1}{\pi} \int_{-\pi}^{\pi} f(x) \operatorname{sen} nx \, dx \quad (3)$$

se acercan a cero cuando n tiende al infinito. El teorema también demostraba que para $f(x)$ acotada e integrable, la convergencia de su serie de Fourier en un punto en $[-\pi, \pi]$ depende solamente del comportamiento de $f(x)$ en la vecindad de ese punto. Sin embargo, el problema de encontrar condiciones necesarias y suficientes para $f(x)$ de modo que su serie de Fourier converja a $f(x)$ no se había y no ha sido resuelto.

Riemann abrió otra línea de investigación. Consideró las series trigonométricas pero no exigió que los coeficientes se determinaran por la fórmula (3) para los coeficientes de Fourier. Comienza con la serie

$$\sum_1^{\infty} a_n \operatorname{sen} nx + \frac{b_0}{2} + \sum_1^{\infty} b_n \cos nx \quad (4)$$

y define

$$A_0 = (1/2)b_0 \quad A_n(x) = a_n \operatorname{sen} nx + b_n \operatorname{cos} nx$$

Entonces la serie (4) es igual a

$$f(x) = \sum_{n=0}^{\infty} A_n(x)$$

Desde luego que $f(x)$ tiene un valor solamente para aquellos valores de x para los cuales la serie converge. Refirámonos a la serie misma por Q . Ahora los términos de Q pueden aproximarse a cero para toda x o para alguna x . Estos dos casos los trata Riemann por separado.

Si a_n y b_n tienden a cero, los términos de Q tienden a cero para toda x . Sea $F(x)$ la función

$$F(x) = C + C'x + A_0 \frac{x^2}{2} - A_1 - \frac{A_2}{4} - \dots - \frac{A_n}{n^2}$$

que se obtiene por medio de dos integraciones sucesivas de Q . Riemann demuestra que $F(x)$ converge para toda x y es continua en x . Entonces la propia $f(x)$ puede integrarse. Riemann demuestra luego una serie de teoremas acerca de $F(x)$, que como consecuencia conducen a condiciones necesarias y suficientes para que una serie de la forma (4) converja a una función dada $f(x)$ de período $2n$. Proporciona, entonces, una condición necesaria y suficiente para

que la serie trigonométrica (4) converja en un valor particular de x , con a_n y b_n aproximándose todavía a cero conforme n tiende a ∞ .

Después considera el caso alterno donde

$$\lim_{n \rightarrow \infty} A_n$$

depende del valor de x y de condiciones que se cumplen cuando la serie Ω es convergente para valores particulares de x y un criterio para la convergencia en valores particulares de x .

También demuestra que una $f(x)$ dada puede ser integrable y aún así no tener una representación en serie de Fourier. Lo que es más, existen funciones no integrables a las cuales converge la serie Ω para un número infinito de valores de x tomados entre límites arbitrariamente cercanos. Finalmente, una serie trigonométrica puede converger para un número infinito de valores de x en un intervalo arbitrariamente pequeño a pesar de que a_n y b_n no tiendan a cero para todas las x .

La naturaleza de la convergencia de las series de Fourier recibió mayor atención después de la introducción del concepto de convergencia uniforme por Stokes y Seidel. Desde Dirichlet se sabía que las series, en general, eran sólo condicionalmente convergentes, si es que lo eran, y que su convergencia dependía de la presencia de términos positivos y negativos. Heine hizo notar, en un artículo de 1870,²⁶² que la demostración usual de que una $f(x)$ acotada está representada unívocamente entre $-\pi$ y π , por una serie de Fourier es incompleta porque la serie puede no ser uniformemente convergente

y así no puede integrarse término a término. Esto sugirió que, sin embargo, pueden existir series trigonométricas no uniformemente convergentes que sí representen una función. Además, una función continua podría representarse por una serie de Fourier y aún así la serie podría no ser uniformemente convergente. Estos problemas dieron origen a la aparición de una nueva serie de investigaciones que buscaban establecer la unicidad de la representación de una función mediante una serie trigonométrica y si los coeficientes son necesariamente los coeficientes de Fourier. Heine demostró, en el artículo que se mencionó antes, que una serie de Fourier que representa una función acotada que satisface las condiciones de Dirichlet es uniformemente convergente en las partes del intervalo $[-\pi, \pi]$ que quedan cuando se eliminan del intervalo entornos arbitrariamente pequeños de los puntos de discontinuidad de la función. En estos entornos la convergencia es necesariamente no uniforme. Heine demostró después que si la convergencia uniforme que se acaba de especificar se cumple para una serie trigonométrica que representa a una función, entonces la serie es única.

El segundo resultado, sobre unicidad, es equivalente a la afirmación de que si una serie trigonométrica de la forma

$$\frac{a_0}{2} + \sum_{n=1}^{\infty} (a_n \cos nx + b_n \operatorname{sen} nx) \quad (5)$$

es uniformemente convergente y representa cero donde converge, esto es, excepto en un conjunto finito P de puntos, entonces los

coeficientes son todos cero y por supuesto que entonces la serie representa a cero en todo el intervalo $[-\pi, \pi]$.

Los problemas asociados con la unicidad de las series trigonométricas y las de Fourier atrajeron a Georg Cantor, quien estudió el trabajo de Heine. Cantor empezó sus investigaciones buscando criterios de unicidad para las representaciones de las funciones en series trigonométricas²⁶³. Demostró que cuando $f(x)$ se representa por una serie trigonométrica convergente para toda x , no existe otra serie trigonométrica de la misma forma que converja análogamente para todo x y represente la misma función $f(x)$.²⁶⁴ Otro artículo proporcionó una demostración mejor de este último resultado.

El teorema de unicidad que demostró puede enunciarse de nuevo así: si, para toda x , existe una representación convergente de cero por una serie trigonométrica, entonces los coeficientes a_n y b_n son cero. Entonces Cantor demuestra, en el artículo de 1871, que la conclusión es válida aun si se prescinde de la convergencia para un número finito de valores de x . Este artículo fue el primero de una serie de ellos en los que Cantor trata los conjuntos de valores excepcionales de x . Extendió²⁶⁵ el resultado de unicidad al caso donde se permite un conjunto infinito de valores excepcionales. Para describir este conjunto definió primero que un punto p es punto límite de un conjunto de puntos S si todo intervalo que contenga a p contiene infinitos puntos de S . Después introdujo la noción de conjunto derivado de un conjunto de puntos. Este conjunto derivado consiste de los puntos límite del conjunto original. Existe

entonces un segundo conjunto derivado, esto es, el conjunto derivado del conjunto derivado, y así sucesivamente. Si el w -ésimo conjunto derivado de un conjunto dado es un conjunto finito de puntos entonces se dice que el conjunto dado es de clase n -ésima o n -ésimo orden (o de primera especie). La respuesta final de Cantor a la cuestión de si una función puede tener dos representaciones diferentes en series trigonométricas en el intervalo $[-\pi, \pi]$ o si cero puede tener una representación de Fourier que no sea cero, es que si en el intervalo una serie trigonométrica suma cero para toda x excepto las de un conjunto de puntos de clase n -ésima (en el cual no se sabe algo más acerca de la serie) entonces todos los coeficientes de la serie deben ser cero. En este artículo de 1872, Cantor sentó las bases de la teoría de los conjuntos de puntos que consideraremos en un capítulo posterior. Muchos otros investigaron el problema de la unicidad en la última parte del siglo XIX y la primera del XX.²⁶⁶

Durante aproximadamente cincuenta años después del trabajo de Dirichlet, se creyó que la serie de Fourier de cualquier función continua en a converge a la función. Pero Du Bois-Reymond²⁶⁷ dio un ejemplo de una función continua en a cuya serie de Fourier no converge en un punto particular. También construyó otra función continua cuya serie de Fourier no converge en los puntos de un conjunto denso en todas partes. Después, en 1875²⁶⁸, demostró que si una serie trigonométrica de la forma

$$a_0 + \sum_1^{\infty} (a_n \cos nx + b_n \operatorname{sen} nx)$$

converge a $f(x)$ en $[-\pi, \pi]$ y si $f(x)$ es integrable (en un sentido aún más general que el de Riemann, en el que $f(x)$ puede ser no acotada en un conjunto de primera especie) entonces la serie debe ser la serie de Fourier para $f(x)$.²⁶⁹ También demostró que cualquier serie de Fourier de una función que es integrable en el sentido de Riemann se puede integrar término a término a pesar de que la serie no sea uniformemente convergente.

Muchos autores emprendieron entonces la investigación del problema ya resuelto de una manera por Dirichlet, a saber, proporcionar condiciones suficientes para que una función $f(x)$ tenga una serie de Fourier convergente a $f(x)$. Varios resultados son clásicos. Jordán proporcionó una condición suficiente en términos del concepto de función de variación acotada, que él introdujo.²⁷⁰

Sea $f(x)$ acotada en $[a, b]$ y sea $a = x_0, x_1, \dots, x_{n-1}, x_n = b$ un modo de división (partición) de este intervalo. Sean $y_0, y_1, \dots, y_{n-1}, y_n$ los valores de $f(x)$ en estos puntos. Entonces, para cada partición

$$\sum_0^{n-1} (y_{r+1} - y_r) = f(b) - f(a)$$

denótese con t a

$$\sum_0^{n-1} |y_{r+1} - y_r|$$

Para cada modo de subdividir el intervalo $[a, b]$ existe una t . Cuando, correspondiendo a todos los modos posibles de división de $[a, b]$, las sumas t tienen una mínima cota superior, entonces f se define como de variación acotada en $[a, b]$.

La condición suficiente de Jordán establece que la serie de Fourier para la función integrable $f(x)$ converge a

$$(1/2) [f(x + 0) + f(x - 0)]$$

en cada punto para el cual exista un entorno en el que $f(x)$ sea de variación acotada.²⁷¹

Durante los sesenta y setenta se examinaron también las propiedades de los coeficientes de Fourier y entre los resultados importantes obtenidos estaba el que es llamado teorema de Parseval (quien lo enunció bajo condiciones más restringidas, cap. 29, sec. 3), según el cual si $f(x)$ y $[f(x)]^2$ son integrables en el sentido de Riemann en $[-\pi, \pi]$ entonces

$$\frac{1}{\pi} \int_{-\pi}^{\pi} [f(x)]^2 dx = 2a_0^2 + \sum_1^{\infty} (a_n^2 + b_n^2)$$

y si $f(x)$ y $g(x)$ y sus cuadrados son integrables en el sentido de Riemann entonces

$$\frac{1}{\pi} \int_{-\pi}^{\pi} f(x)g(x)dx = 2a_0\alpha_0 + \sum_1^{\infty} (a_n\alpha_n + b_n\beta_n)$$

donde a_n , b_n , α_n y β_n son los coeficientes de Fourier de $f(x)$ y $g(x)$ respectivamente.*

7. La situación del análisis

La obra de Bolzano, Cauchy, Weierstrass y otros proporcionó rigor al análisis. Esta obra liberó al cálculo y sus extensiones de toda dependencia de nociones geométricas, del movimiento y de comprensiones intuitivas. Desde el principio, estas investigaciones causaron considerable conmoción. Después de una reunión científica en la que Cauchy presentó la teoría sobre la convergencia de las series, Laplace se marchó precipitadamente a casa y permaneció allí retirado hasta que hubo examinado las series en su *Mécanique celeste*. Afortunadamente, encontró que todas eran convergentes. Cuando se llegó a conocer la obra de Weierstrass a través de sus conferencias, el efecto fue aún más notable. Las mejoras en el rigor pueden verse si se compara la primera edición

* En 1966 L. Carleson demostró que la serie de Fourier de una función de cuadrado integrable en el sentido de Lebesgue converge en casi todo punto a la función, vid. L. Carleson, «On convergence and growth of partial sums of Fourier series», *Acta Mathematica*, 116 (1966), 135-157. Un poco después, Hunt extendió el resultado usando los métodos de Carleson, a L_p con $p \neq 2$, vid. R. A. Hunt, «On the convergence of Fourier series», en *Orthogonal expansions and their continuous analogues*, Proc. Conference Southern Illinois University, 1967; Southern Illinois University Press, 1968, pp. 235-255. (Nota del traductor.)

del Cours d'analyse de Jordán (1882-1887) con la segunda (1893-1896) y la tercera (3 vols., 1909-1915). Muchos otros tratados incorporaron el nuevo rigor.

La rigorización del análisis no resultó ser el fin de la investigación en los fundamentos. Esto fue así porque prácticamente todo el trabajo presuponía el sistema de los números reales, pero este tema permanecía desorganizado. Excepto para Weierstrass, quien, como veremos, consideró el problema del número irracional durante los años cuarenta, todos los demás no creían necesario investigar los fundamentos lógicos del sistema de los números. Parecería que aún los más grandes matemáticos deben desarrollar sus capacidades para apreciar la necesidad del rigor en etapas. El trabajo sobre los fundamentos lógicos del sistema de los números reales iba a seguir en breve (cap. 41).

El descubrimiento de que las funciones continuas no tienen necesariamente derivadas, que las funciones discontinuas pueden integrarse, la nueva luz arrojada sobre las funciones discontinuas por el trabajo de Dirichlet y Riemann acerca de las series de Fourier y el estudio de la variedad y extensión de las discontinuidades de las funciones hicieron que los matemáticos se dieran cuenta de que el estudio riguroso de las funciones se extiende más allá de las que se usan en el cálculo y las ramas usuales del análisis, donde el requerimiento de la diferenciabilidad restringe generalmente la clase de funciones. El estudio de las funciones continuó en el siglo XX y llevó al desarrollo de una nueva rama de las matemáticas, conocida como la teoría de las funciones de una variable real (cap. 44).

Igual que todos los nuevos movimientos en las matemáticas, la rigorización del análisis no avanzó sin oposición. Hubo mucha controversia en cuanto a si los refinamientos en el análisis debían investigarse. Las funciones peculiares que se introdujeron fueron atacadas como curiosidades, funciones disparatadas, funciones divertidas y como juguetes matemáticos, tal vez más intrincados pero no de mayor consecuencia que los cuadrados mágicos. También se vieron como enfermedades o parte de la patología mórbida de las funciones y sin tener relación alguna con los problemas importantes de las matemáticas puras y aplicadas. Estas nuevas funciones, violando leyes que se creía perfectas, se consideraron como signos de anarquía y caos que se mofaban del orden y armonía que las generaciones previas habían buscado. Las muchas hipótesis que ahora se tenían que hacer para enunciar un teorema preciso fueron vistas como pedantes y destructoras de la elegancia del análisis clásico del siglo XVIII, «como era en el paraíso», por parafrasear a Du Bois-Reymond. Los nuevos detalles se resintieron como oscurecedores de las ideas principales.

Poincaré, en particular, desconfió de esta nueva investigación. Dijo que:²⁷²

La lógica a veces crea monstruos. Durante medio siglo, hemos visto una masa de funciones extrañas que parecen forzadas a parecerse tan poco como sea posible a funciones honestas que sirvan a algún propósito. Más continuidad, o menos continuidad, más derivadas, etc. Ciertamente, desde el punto de vista de la lógica, estas funciones extrañas son las más generales; por otro lado, aquellas

que uno se encuentra sin buscarlas, y que siguen leyes sencillas se presentan como un caso particular que no representa más que una esquina pequeña.

En épocas pasadas, cuando uno inventaba una nueva función era con un propósito práctico; hoy, uno las inventa con el propósito de mostrar defectos en los razonamientos de nuestros padres y sólo eso se deducirá de ellas.

Charles Hermite dijo en una carta a Stieltjes: «Me aparto con miedo y horror de esta lamentable plaga de funciones que no tienen derivadas.»

Du Bois-Reymond²⁷³ dio a conocer otro tipo de objeción. Su interés radicaba en que la aritmetización del análisis separaba al análisis de la geometría, y consecuentemente de la intuición y el pensamiento físico. Reducía el análisis «a un simple juego de símbolos donde los signos escritos toman la significación arbitraria de las piezas en el ajedrez o un juego de cartas».

El asunto que provocó la mayor controversia fue la prohibición de las series divergentes, principalmente por Abel y Cauchy. En una carta a Holmboé, escrita en 1826, Abel dice: ²⁷⁴

Las series divergentes son la invención del demonio, y es una vergüenza basar en ellas cualquier demostración, sea la que fuere. Usándolas, se puede obtener cualquier conclusión que le plazca a uno y por eso han producido tantas falacias y tantas paradojas... He llegado a estar prodigiosamente atento a todo esto, pues con la excepción de las series geométricas, no existe en todas las matemáticas una sola serie infinita cuya suma se

haya determinado rigurosamente. En otras palabras, las cosas más importantes en matemáticas son aquellas que tienen menor fundamentación.

Sin embargo, Abel mostró cierto interés acerca de si se había pasado por alto alguna idea buena, porque continúa en su carta del siguiente modo: «Que la mayor parte de estas cosas son correctas a pesar de eso es extraordinariamente sorprendente. Estoy tratando de encontrar una razón de ello; es una cuestión extraordinariamente interesante.» Abel murió joven y, por tanto, nunca investigó el asunto.

Cauchy, también, tuvo algunos escrúpulos en condenar a las series divergentes al ostracismo. Dice en la introducción de su Cours (1821), «He sido forzado a admitir diversas proposiciones que parecen algo deplorables, por ejemplo, que una serie divergente no puede sumarse.» A pesar de esta conclusión, Cauchy continuó usando series divergentes, según aparece en notas añadidas a la publicación en 1827 ²⁷⁵ de un ensayo para concursar, escrito en 1815, sobre ondas de agua. Decidió investigar la cuestión de por qué las series divergentes resultaron tan útiles, y de hecho finalmente llegó a estar cerca de reconocer la razón (cap. 47).

Los matemáticos franceses aceptaron la prohibición de Cauchy de las series divergentes, pero no los ingleses y alemanes. En Inglaterra, la escuela de Cambridge defendió el uso de las series divergentes apelando al principio de la permanencia de la forma (cap. 32, sec. 1). En relación con las series divergentes, el principio

fue usado primero por Robert Woodhouse (1773-1827). En *The Principies of Analytic Calculation* (Los principios del Cálculo Analítico, 1803, p. 3) señala que en la ecuación

$$\frac{1}{1-r} = 1 + r + r^2 + \dots \quad (6)$$

el signo de igualdad tiene «un significado más amplio» que sólo el de la igualdad numérica. Por tanto la ecuación se cumple tanto si la serie es divergente como si no.

Peacock también aplicó el principio de permanencia de la forma a las operaciones con series divergentes.²⁷⁶ En la página 267 dice: «Así, como para $r < 1$, la igualdad (6) anterior se cumple, entonces para $r = 1$ obtenemos $\infty = 1 + 1 + 1 + \dots$. Para $r > 1$ obtenemos un número negativo a la izquierda y, como los términos a la derecha se incrementan continuamente, una cantidad mayor que infinito en la derecha.» Peacock acepta esto. El punto que trata de aclarar es que la serie pueda representar $1/(1 - r)$ para toda r . Dice:

Si las operaciones del álgebra se consideran como generales, y los símbolos que se sujetan a ellas ilimitados en valor, será imposible evitar la formación de series divergentes al igual que convergentes; y si tales series se consideran como los resultados de operaciones que son definibles, además de las series mismas, entonces no será muy importante entrar en tal examen de la relación de los valores aritméticos de los términos sucesivos como puede ser necesario para asegurar su convergencia o divergencia; pues bajo tales circunstancias, se

deben considerar como formas equivalentes que representan a su función generadora, y como que para los propósitos de tales operaciones poseen propiedades equivalentes... El intento por excluir el uso de las series divergentes en las operaciones simbólicas necesariamente impondría un límite sobre la universalidad de las fórmulas y operaciones algebraicas, lo cual, conjuntamente, es contrario al espíritu de la ciencia... Necesariamente conduciría a una mayor y embarazosa multiplicación de casos: privaría a casi todas las operaciones algebraicas de mucha de su certeza y simplicidad.

Augustus de Morgan, aunque mucho más agudo y más consciente que Peacock de las dificultades con las series divergentes, se encontraba, sin embargo, bajo la influencia de la escuela inglesa y también se impresionó por los resultados obtenidos por el uso de las series divergentes a pesar de las dificultades en ellas. En 1844 empezó un artículo, incisivo pero aun así confuso, sobre «Series Divergentes»,²⁷⁷ con estas palabras, «Creo que será generalmente admitido que el encabezamiento de este artículo describe el único asunto que aún queda, de carácter elemental, sobre el cual existe un serio cisma entre los matemáticos en cuanto a la exactitud o inexactitud absolutas de los resultados.» La posición que De Morgan tomó ya la había declarado en su *Differential and Integral Calculus* (Cálculo Diferencial e Integral),²⁷⁸ «La historia del álgebra nos muestra que nada es más erróneo que el rechazo de cualquier método que surja naturalmente, a causa de uno o más casos

aparentemente válidos en los cuales tal método conduzca a resultados equivocados. Tales casos ciertamente debieran enseñar cautela, pero no rechazo; si se hubiera preferido lo último a lo primero, las cantidades negativas, y aún más sus raíces cuadradas, habrían sido un obstáculo eficaz para el progreso del álgebra... y esos campos inmensos del análisis en los cuales incluso los que rechazaban las series divergentes ahora se incluyen sin temor, no habrían sido descubiertos, ni mucho menos cultivados y establecidos... La consigna que yo adoptaría contra un desarrollo que me parece calculado para detener el progreso del descubrimiento estaría contenido en una palabra y un símbolo — recuérdese $\sqrt{-1}$.» Distingue entre el significado aritmético y algebraico de una serie.

El significado algebraico se cumple en todos los casos. Para dar cuenta de algunas de las conclusiones falsas obtenidas con las series divergentes dice en el artículo de 1844 (p. 187) que la integración es una operación aritmética y no algebraica y por tanto no se podría aplicar sin razonamiento adicional con las series divergentes. Pero la obtención de

$$\frac{1}{1-r} = 1 + r + r^2 + \dots$$

comenzando con $y = 1 + ry$, reemplazando y en la derecha por $1 + ry$, y continuando así, lo acepta porque es algebraico. Análogamente, de $z = 1 + 2z$ se obtiene $z = 1 + 2 + 4 + \dots$. Por tanto $-1 = 1 + 2 + 4 + \dots$ y esto es correcto. Acepta la teoría completa (como

estaba en esa época) de las series trigonométricas pero estaría dispuesto a rechazarla si se pudiera dar un ejemplo donde $1 - 1 + 1 - 1 + 1 \dots$ no sea igual a $1/2$ (véase cap. 20).

Muchos otros matemáticos ingleses prominentes proporcionaron otros tipos de justificación para la aceptación de las series divergentes, algunos retornando a un argumento de Nicholas Bernoulli (cap. 20, sec. 7), de que la serie (6) contiene un resto r^∞ o $r^\infty/(1 - r^\infty)$. Esto debe tomarse en cuenta (aunque no indicaron cómo). Otros dijeron que la suma de una serie divergente es algebraicamente verdadera pero aritméticamente falsa.

Algunos matemáticos alemanes usaron los mismos argumentos que Peacock aunque utilizaron palabras distintas, tales como operaciones sintácticas opuestas a operaciones aritméticas o literal opuesto a numérico. Martin Ohm²⁷⁹ dijo, «Una serie infinita (soslayando cualquier cuestión sobre la convergencia o divergencia) está completamente adaptada para representar una expresión dada si uno puede estar seguro de tener la ley correcta para desarrollar la serie. Del valor de una serie infinita se puede hablar sólo si converge.» Los argumentos en Alemania a favor de la legitimidad de las series divergentes fueron sostenidos durante varias décadas más.

La defensa del uso de las series divergentes no estuvo tan cercana del desastre como podría parecer, aunque muchos de los argumentos dados a favor de las series eran, tal vez, inverosímiles. Por un lado, en la totalidad del análisis del siglo XVIII, la atención al rigor o a la demostración fue mínima, y esto era aceptable porque

los resultados obtenidos casi siempre eran correctos. Así, los matemáticos llegaron a acostumbrarse a los procedimientos y argumentos imprecisos. Pero, aún más interesante para la cuestión, muchos de los conceptos y operaciones que habían causado perplejidades, tales como los números complejos, se demostró que eran correctos después de que fueran totalmente entendidos. Por eso los matemáticos pensaron que las dificultades con las series divergentes también serían aclaradas cuando se tuviera una mejor comprensión, y que las series divergentes resultarían ser legítimas. Lo que es más, las operaciones con las series divergentes estaban ligadas con frecuencia con otras operaciones poco entendidas del análisis, tales como el intercambio del orden de los límites, la integración sobre discontinuidades de un integrando y la integración sobre un intervalo infinito, de modo que los defensores de las series divergentes podían sostener que las conclusiones falsas atribuidas al uso de las series divergentes surgieron de otras fuentes de problemas.

Un argumento que pudo haberse esgrimido es que cuando una función analítica se expresa en algún dominio por una serie de potencias, lo que Weierstrass llamó un elemento, esta serie ciertamente sí lleva consigo las propiedades «algebraicas» o «sintácticas» de la función y estas propiedades son llevadas más allá del dominio de convergencia del elemento. El proceso de la continuación analítica usa este hecho. Realmente había material matemático sólido en el concepto de las series divergentes que justificaba su utilidad. Pero el reconocimiento de este material y la

aceptación final de las series divergentes tuvo que esperar una nueva teoría de las series infinitas (cap. 47).

Bibliografía

- Abel, N. H.: *Œuvres complètes*, 2 vols., 1881, Johnson Reprint Corp., 1964. - *Mémorial publié à l'occasion du centenaire de sa naissance*, Jacob Dybwad, 1902. Cartas escritas por Abel y dirigidas a él.
- Bolzano, B.: *Rein analytischer Beweis des Lehrsatzes, das zwischen je zwei Werthen, die ein entgegengesetztes Resultat gewähren, wenigstens eine reelle Wurzel der Gleichung liege*, Gottlieb Hass, Prague, 1817 = *Abh. Königl. Böhm. Ges. der Wiss.* (3), 5, 1814-1817, pub. 1818 = *Ostwald's. -Klassiker der exakten Wissenschaften*, 153, 1905, 3-45. No contenida en los *Schriften* de Bolzano. - *Paradoxes of the Infinite*, Routledge and Kegan Paul, 1950. Contiene una investigación histórica del trabajo de Bolzano. - *Schriften*, 5 vols., Königlichen Böhmischen Gesellschaft der Wissenschaften, 1930-1948.
- Boyer, Cari B.: *The Concepts of the Calculus*, Dover (reimp.), 1949, cap. 7.
- Burkhardt, H.: «Über den Gebrauch divergenter Reihen in der Zeit von 1750-1860.» *Math. Ann.*, 70, 1911, 169-206. - «Trigonometrische Reihe und Intégrale». *Encyk. der Math. Wiss.*, II A12, 819-1354, B. G. Teubner, 1904-1916.
- Cantor, Georg: *Gesammelte Abhandlungen* (1932), Georg Olms (reimp.), 1962.

- Cauchy, A. L.: *Œuvres* (2), Gauthier-Villars, 1897-1899, vols. 3 y 4.
- Dauben, J. W.: «The Trigonometric Background to Georg Cantor's Theory of Sets». *Archive for History of Exact Sciences*, 7, 1971, 181-216.
- Dirichlet, P. G. L.: *Werke*, 2 vols. Georg Reimer, 1889-1897, Chelsea (reimp.), 1969.
- Du Bois-Reymond, Paul: *Zwei Abhandlungen über unendliche und trigonometrische Reihen* (1871 y 1874), Ostwald's Klassiker, 185; Wilhelm Engelmann, 1913.
- Freudenthal, H.: «Did Cauchy Plagiarize Bolzano?». *Archive for History of Exact Sciences*, 7, 1971, 375-392.
- Gibson, G. A.: «On the History of Fourier Series». *Proc. Edinburgh Math. Soc.*, 11, 1892-1893, 137-166.
- Grattan-Guinness, I.: «Bolzano, Cauchy and the "New Analysis" of the Nineteenth Century». *Archive for History of Exact Sciences*, 6, 1970, 372-400. -*The Development of the Foundations of Mathematical Analysis from Euler to Riemann*, Massachusetts Institute of Technology Press, 1970.
- Hawkins, Thomas W., Jr.: *Lebesgue's Theory of Integration: Its Origins and Development*. University of Wisconsin Press, 1970, caps. 1-3.
- Manheim, Jerome H.: *The Génesis of Point Set Topology*. Macmillan, 1964, caps. 1-4.
- Pesin, Ivan H.: *Classical and Modern Integration Theories*. Academic Press, 1970, cap. 1.

- Pringsheim, A.: «Irrationalzahlen und Konvergenz unendlichen Prozesse». *Encyk. der Math. Wiss.*, IA3, 47-147, B. G. Teubner, 1898-1904.
- Reiff, R.: *Geschichte der unendlichen Reihen*. H. Lauppische Bbuchhandlung, 1889; Martin Sánding (reimp.), 1969.
- Riemann, Bernhard: *Gesammelte mathematische Werke*, 2.^a ed. (1902), Dover (reimp.), 1953.
- Schleinger, L.: «Über Gauss Arbeiten zur Funktionenlehre». *Nachrichten*
- *König. Ges. der Wiss. zu Gótt.*, 1912, Beiheft, 1-43. También en el *Werke* de Gauss, 10, 77 y sigs.
- Schoenflies, Arthur M.: «Die Entwicklung der Lehre von den Punktmannigfaltigkeiten». *Jahres. der Deut. Math.-Verein.*, 8, 1899, 1-250.
- Singh, A. N.: «The Theory and Construction of Non-Differentiable Functions», en E. W. Hobson: *Squaring the Circle and Other Monographs*, Chelsea (reimp.), 1953.
- Smith, David E.: *A Source Book in Mathematics*, Dover (reimp.), 1959, vol. 1, 286-291, vol. 2, 635-637.
- Stolz, O.: «B. Bolzanos Bedeutung in der Geschichte der Infinitesimalrechnung». *Math. Ann.*, 18, 1881, 255-279.
- Weierstrass, Karl: *Mathematische Werke*, 7 vols., Mayer und Muller, 1894-1927.
- Young, Grace C.: «On Infinite Derivatives». *Quart. Jour. of Math.*, 47, 1916, 127-175.

Capítulo 41

La fundamentación de los números reales y transfinitos

Dios hizo los enteros; el resto es obra del hombre.

Leopold Kronecker

Contenido:

1. *Introducción*
 2. *Números algebraicos y trascendentes*
 3. *La teoría de los números irracionales*
 4. *La teoría de los números racionales*
 5. *Otros enfoques del sistema de los números reales*
 6. *El concepto de conjunto infinito*
 7. *Los fundamentos de la teoría de conjuntos*
 8. *Car-dinales y ordinales transfinitos*
 9. *La situación de la teoría de conjuntos hacia 1900*
- Bibliografía*

1. Introducción

Uno de los hechos más sorprendentes en la historia de la matemática es que no se acometiera la fundamentación lógica del sistema de números reales hasta finales del siglo XIX. Hasta ese momento no quedaron lógicamente establecidas ni siquiera las propiedades más simples de los números racionales positivos y negativos y de los números irracionales, ni habían sido definidos

esos números. La misma fundamentación lógica de los números complejos, cuya existencia no databa de mucho antes (vid. cap. 32, sec. 1), presuponía la del sistema de números reales. Considerando el extenso desarrollo del álgebra y el análisis, y cómo en ellos se utilizan los números reales, la falta de una estructuración precisa de éstos y de sus propiedades muestra cuán ilógicamente progresa la matemática. La comprensión intuitiva de esos números parecía suficiente, y los matemáticos se contentaban con operar sobre esa base.

La rigorización del análisis impelía a remediar la falta de claridad en el sistema numérico mismo. Por ejemplo, la demostración de Bolzano (cap. 40, sec. 2) de que una función continua que es negativa para $x = a$ y positiva para $x = b$ se anula para algún valor de x entre a y b patinaba en un punto crítico porque faltaba una adecuada comprensión de la estructura del sistema de números reales. El estudio detallado de los límites también mostraba la necesidad de comprender los números reales, ya que números racionales pueden tener un límite irracional y recíprocamente. La incapacidad de Cauchy para probar la suficiencia de su criterio para la convergencia de una sucesión se derivaba igualmente de su falta de comprensión de la estructura del sistema numérico. El estudio de las discontinuidades de funciones representables mediante series de Fourier revelaba la misma deficiencia. Fue Weierstrass el primero que señaló que para establecer con precisión las propiedades de las funciones continuas necesitaba la teoría del continuo aritmético.

Otra motivación para la fundamentación del sistema numérico fue

el deseo de asegurar la verdad de la matemática. Como consecuencia de la creación de las geometrías no euclídeas, la geometría había perdido su *status* de verdad (vid. cap. 36, sec. 8), pero parecía todavía que la matemática construida sobre la aritmética ordinaria debía ser una realidad incuestionable en cierto sentido filosófico. Ya en 1817, en su carta a Oibers, Gauss²⁸⁰ había distinguido la aritmética de la geometría en que sólo la primera era puramente a priori. En su carta a Bessel del 9 de abril de 1830,²⁸¹ repetía la afirmación de que sólo las leyes de la aritmética son necesarias y verdaderas. Sin embargo, faltaba una fundamentación del sistema numérico que despejara cualquier duda sobre la verdad de la aritmética, y del álgebra y el análisis construido sobre esa base.

Merece la pena señalar que antes de que los matemáticos apreciaran la necesidad de analizar el sistema numérico mismo, el problema que les había parecido más pertinente era el de la fundamentación del *álgebra*, y en particular una explicación del hecho de que uno pueda usar letras para representar números reales y complejos, y operar con las letras por medio de las propiedades aceptadas como verdaderas para los enteros positivos. Para Peacock, de Morgan y Duncan Gregory, el álgebra de comienzos del siglo XIX era un ingenioso, pero también ingenuo, complejo de esquemas manipulato- rios con algo de sentido pero muy poca sustancia; les parecía que el núcleo de la habitual confusión residía en una fundamentación inadecuada del álgebra. Ya hemos visto cómo resolvieron el problema (cap. 32, sec. 1). A

fines de siglo, sin embargo, quedó claro que había que profundizar por el lado del análisis y clarificar la estructura de todo el sistema numérico real. De pasada se aseguraría así también la estructura lógica del álgebra, ya que resultaba intuitivamente claro que los diferentes tipos de números poseían las mismas propiedades formales. De aquí que si se podían establecer esas propiedades sobre una base firme, podrían aplicarse a las letras que representaban números cualesquiera.

2. Números algebraicos y trascendentes

Una etapa importante hacia la mejor comprensión de los números irracionales fue el trabajo de mediados del siglo XIX sobre los irracionales algebraicos y trascendentes. La distinción entre unos y otros había quedado establecida en el siglo XVIII (cap. 25, sec. 1). El interés de tal distinción aumentó con los trabajos del XIX sobre resolución de ecuaciones, que mostraron que no todos los irracionales algebraicos pueden obtenerse mediante operaciones algebraicas sobre números racionales. Además, el problema de determinar si e y π eran algebraicos o trascendentes seguía atrayendo a los matemáticos.

Hasta 1844 siguió abierta la cuestión sobre si había o no irracionales trascendentes. Ese año, Liouville²⁸² mostró que cualquier número de la forma

$$\frac{a_1}{10} + \frac{a_2}{10^{2!}} + \frac{a_3}{10^{3!}} + \dots$$

donde los a_i son enteros arbitrarios de 0 a 9, es trascendente.

Para probarlo, Liouville demostró primero algunos teoremas sobre la aproximación de irracionales algebraicos mediante números racionales. Por definición (cap. 25, sec. 1), un número algebraico es cualquier número, real o complejo, que satisface una ecuación algebraica

$$a_0x^n + a_1x^{n-1} + \dots + a_n = 0$$

donde los a_i son enteros. Una raíz es un número algebraico de grado n si satisface una ecuación de grado n y no satisface ninguna ecuación de menor grado. Algunos números algebraicos son racionales; esos son de grado uno. Liouville probó que si p/q es cualquier aproximación a un número algebraico irracional x de grado n , con p y q enteros, entonces existe un número positivo M tal que

$$\left| x - \frac{p}{q} \right| > \frac{M}{q^n}$$

Esto significa que cualquier aproximación racional a un irracional algebraico de grado n por cualquier p/q debe ser menos precisa que M/q^n . Con otras palabras, podemos decir que si x es un irracional algebraico de grado n , existe un número

$$\left| x - \frac{p}{q} \right| > \frac{M}{q^n}$$

no tiene soluciones enteras p y q para $p = n$ y tampoco por tanto para $\mu \leq n$. Así pues, x es trascendente si para un M fijo y para cada entero positivo p la desigualdad tiene alguna solución p/q . Mostrando que sus irracionales satisfacen este último criterio, Liouville demostró que son trascendentes.

La siguiente gran etapa en el reconocimiento de números trascendentes específicos fue la demostración de Hermite en 1873²⁸³ de que e es trascendente. Después de obtener este resultado, Hermite escribió a Cari Wilhelm Borchardt (1817-1880): «No me atrevo a intentar probar la trascendencia de π . Si otros lo logran nadie estará más feliz que yo con su éxito, pero créame, mi querido amigo, que no dejará de costarles un cierto esfuerzo.»

Que π es trascendente había sido ya sospechado por Legendre (cap. 25, sec. 1). Ferdinand Lindemann (1852-1939) lo demostró en 1882²⁸⁴ con un método que no difiere esencialmente del empleado por Hermite: Lindemann probó que si x_1, x_2, \dots, x_n son números algebraicos distintos, reales o complejos, y p_1, p_2, \dots, p_n son números algebraicos no todos nulos, la suma

$$p_1 e^{x_1} + p_2 e^{x_2} + \dots + p_n e^{x_n}$$

no puede ser 0. Si tomamos $n = 2$, $p_1 = 1$, y $x_2 = 0$, vemos que e^{x_1} no puede ser algebraico para un x_1 que sea algebraico y no nulo. Como x_1 puede ser 1, e es trascendente. Ahora bien, dado que $e^{i\pi} + 1 = 0$, el número $i\pi$ no puede ser algebraico, y tampoco π , puesto que i lo

es, y el producto de dos números algebraicos es algebraico. La demostración de que n es trascendente puso punto final a los famosos problemas de construcción de la geometría, ya que todos los números constructibles son algebraicos.

Permanece todavía el misterio en torno a una constante fundamental, la constante de Euler (cap. 20, sec. 4)

$$C = \lim_{n \rightarrow \infty} \left(1 + \frac{1}{2} + \dots + \frac{1}{n} - \log n \right)$$

que es aproximadamente 0,577216 y que tiene un importante papel en análisis, en particular en el estudio de las funciones gamma y zeta, y de la que no se sabe si es racional o irracional.

3. La teoría de los números irracionales

A finales del siglo XIX se afrontó directamente la cuestión de la estructura lógica del sistema de los números reales. Los números irracionales suponían la principal dificultad. Ahora bien, el desarrollo del significado y propiedades de los números irracionales presupone la construcción del sistema de los números racionales. Los distintos autores que contribuyeron a la teoría de los números irracionales, o bien supusieron que los números racionales se conocían con tanta seguridad que no precisaban una fundamentación, o bien esbozaron precipitadamente algún esquema improvisado.

Curiosamente, la construcción de una teoría de los irracionales no

requería mucho más que un cambio de punto de vista. Euclides, en el Libro V de los *Elementos*, había estudiado las razones entre magnitudes inconmensurables, y había definido la igualdad y la desigualdad entre tales razones. Su definición de igualdad (cap. 4, sec. 5) equivalía a dividir los números racionales m/n en dos clases, la de aquellos para los que m/n es menor que la razón a/b entre las magnitudes inconmensurables a y b , y la de aquellos para los que m/n es más grande. Ciertamente es que la lógica de Euclides era deficiente, ya que nunca definió la razón entre dos magnitudes inconmensurables.

Además, el desarrollo de Euclides de la teoría de las proporciones, o igualdades entre razones, sólo se podía aplicar a la geometría. Aun así, tuvo la idea esencial que podía haberse utilizado antes para definir los números irracionales. De hecho, Dedekind hizo uso del trabajo de Euclides y reconoció²⁸⁵ su deuda hacia éste; también Weierstrass pudo guiarse por la teoría de Euclides. Sin embargo, es más fácil contemplar el pasado que el futuro. El largo retraso en sacar partido de alguna reformulación de las ideas de Euclides puede explicarse ahora fácilmente: los números negativos tenían que ser completamente aceptados antes de poder disponer del sistema completo de números racionales; además, tenía que experimentarse la necesidad de una teoría de los irracionales, y esto ocurrió solamente una vez emprendida la aritmetización del análisis.

William R. Hamilton ofreció el primer tratamiento de los números irracionales en dos artículos, leídos ante la Royal Irish Academy en

1833 y 1835²⁸⁶, y publicados más tarde como «Algebra as the Science of Puré Time». Basaba, pues, en el tiempo su noción de todos los números, racionales e irracionales, una base poco satisfactoria para las matemáticas (aunque muchos, siguiendo a Kant, la hayan considerado como una intuición básica). Tras presentar una teoría de los números racionales, introdujo la idea de separar los racionales en dos clases (describiremos más detalladamente la idea en conexión con la obra de Dedekind) y definió un número irracional como una de esas particiones, aunque no completó el trabajo.

Aparte de esta obra inacabada, todas las presentaciones de los irracionales anteriores a Weierstrass utilizaron la noción de irracional como límite de una sucesión infinita de racionales. Pero el límite, si es irracional, no tiene existencia lógica hasta que se hayan definido los irracionales. Cantor²⁸⁷ señaló que este error lógico no fue advertido durante algún tiempo porque no conducía a dificultades posteriores. Weierstrass, en sus clases en Berlín a partir de 1859, ofreció una teoría de los números irracionales, señalando su necesidad. Una publicación de H. Kossak, *Die Elemente der Arithmetik* (1872), trató de presentar esta teoría, pero Weierstrass la desaprobó.

En 1869 Charles Méray (1835-1911), uno de los apóstoles de la aritmetización de las matemáticas y equivalente francés de Weierstrass, dio una definición de los irracionales basada en los racionales²⁸⁸. Georg Cantor también ofreció una teoría, que necesitaba para clarificar las ideas sobre conjuntos que había

utilizado en su trabajo de 1871 sobre series de Fourier. Fue seguida un año más tarde por la teoría de Heinrich Heine, que apareció en el *Journal für Mathematik*,²⁸⁹ y por la de Dedekind, publicada en *Stetigkeit und irrationale Zahlen*²⁹⁰

Todas estas teorías sobre los números irracionales son esencialmente parecidas; nos limitaremos a dar algunas indicaciones sobre las de Cantor y Dedekind. Cantor²⁹¹ comienza con los números racionales. En su artículo de 1883,²⁹² donde da más detalles sobre su teoría de los números irracionales, dice (p. 565) que no es necesario extenderse sobre los racionales, puesto que ya lo habían hecho Hermann Grassmann en su *Lehrbuch der Arithmetik* (1861) y J. H. T. Müller (1797-1862) en su *Lehrbuch der allgemeinen Arithmetik* (1855). De hecho, tales presentaciones no resultaron definitivas. Cantor introdujo una nueva clase de números, los números reales, que incluyen racionales e irracionales. Construyó los números reales a partir de los racionales, definiendo primeramente lo que llamaba sucesión fundamental: cualquier sucesión de racionales que cumple la condición de que, para cualquier ε prefijado, todos los términos de la sucesión, salvo a lo más un número finito, difieren entre sí en menos de ε , o bien que

$$\lim_{n \rightarrow \infty} (a_{n+m} - a_n) = 0$$

para m arbitrario. Cada sucesión fundamental es, por definición, un número real al que podemos denotar por b . Dos sucesiones

fundamentales (a_n) y (b_n) son el mismo número real si y sólo si $|a_n - b_n|$ se aproxima a cero cuando n se hace infinito.

Para tales sucesiones se presentan tres posibilidades: dado un número racional arbitrario, los términos a_n de la sucesión para n suficientemente grande son todos en valor absoluto menores que el número dado; o todos los términos a partir de un cierto n son mayores que un cierto número racional positivo ρ ; o bien todos los términos de la sucesión a partir de un cierto n son menores que un cierto número racional negativo $-\rho$. En el primer caso, $b = 0$; en el segundo $b > 0$; en el tercero $b < 0$.

Si (a_n) y (a'_n) son dos sucesiones fundamentales, denotadas por b y b' , se puede probar que $(a_n \pm a'_n)$ y (a_n/a'_n) son también sucesiones fundamentales, y definen $b + b'$ y b/b' . Además, si $b \neq 0$, entonces (a'_n/a_n) es también una sucesión fundamental, que define b'/b .

Los números reales racionales quedan incluidos en la definición de más arriba, ya que una sucesión (a_n) con todos sus términos a_n iguales al mismo número racional a define el número real racional a .

A continuación se puede definir la igualdad y desigualdad entre dos números reales cualesquiera: así, $b = b'$, $b > b'$ o $b < b'$, según que $b - b'$ sea igual, mayor o menor que 0.

El siguiente teorema es crucial. Cantor demuestra que si (b_n) es cualquier sucesión de números reales (rationales o irracionales), y si $\lim (b_{n+m} - b_n) = 0$ para m arbitrario, entonces existe un único número real b , determinado por una sucesión fundamental (a_n) de números racionales a_n tal que

$$\lim_{n \rightarrow \infty} b_n = b$$

Es decir, que la formación de sucesiones fundamentales de números reales no crea la necesidad de otros nuevos tipos de números que puedan servir como límites para esas sucesiones fundamentales, puesto que los números reales ya existentes bastan para proporcionar tales límites. En otras palabras, desde el punto de vista de la convergencia de sucesiones fundamentales (o lo que es lo mismo, sucesiones que satisfacen el criterio de convergencia de Cauchy), los números reales constituyen un sistema completo.

La teoría de Dedekind de los números irracionales, presentada en su libro de 1872 mencionado más arriba, parte de las ideas con que ya contaba en 1858. En esa época tuvo que dar clases de cálculo y constató que el sistema de números reales carecía de fundamentación lógica. Para probar que una cantidad monótonamente creciente y acotada se aproxima a un límite tenía que hacer uso, como otros autores, de la evidencia geométrica (decía que esa sigue siendo la forma en que hay que hacerlo en una primera presentación del cálculo, particularmente si uno no desea perder mucho tiempo). Además, muchos teoremas aritméticos básicos quedaban sin demostración; mostraba como ejemplo el hecho de que no se hubiera todavía probado rigurosamente que $\sqrt{3} = \sqrt{6}$.

En su exposición afirma que presupone el desarrollo de los números racionales, que comenta brevemente. Para introducir los

irracionales se pregunta primero qué significa la continuidad geométrica. Los pensadores contemporáneos y anteriores —por ejemplo Bolzano— creían que continuidad significa la existencia de al menos otro número entre dos cualesquiera, propiedad que ahora se conoce como densidad. Pero los números racionales ellos mismos también forman un conjunto denso, luego densidad no es continuidad.

Lo que le sugirió a Dedekind la definición de número irracional fue la observación de que en cada partición de la recta en dos clases de puntos tales que cada punto de la primera está a la izquierda de cada punto de la segunda, hay *uno y sólo un punto* que produce la partición. Ese hecho hace a la recta continua. Para la recta se trata de un axioma; trasladó entonces esa idea al sistema numérico. Consideremos, dice Dedekind, cualquier partición de los números racionales en dos clases tales que cualquier número de la primera es menor que cualquier número de la segunda. Tal partición de los números racionales es lo que llama una cortadura. Si las clases son denotadas por A_1 y A_2 , la cortadura se denota por (A_1, A_2) . Para algunas cortaduras, específicamente las determinadas por un número racional, hay un número máximo en A_1 o un número mínimo en A_2 . Recíprocamente, cada cortadura en los racionales en la que hay un máximo en la primera clase o un mínimo en la segunda está determinada por un número racional.

Pero hay cortaduras que no están determinadas por números racionales. Si ponemos en la primera clase todos los números racionales positivos y negativos cuyo cuadrado es menor que 2, y en

la segunda clase todos los demás racionales, esa cortadura no está determinada por ningún número racional. Para cada cortadura de ese tipo «creamos un nuevo miembro irracional a que está completamente definido por la cortadura; diremos que el número a corresponde a esa cortadura, o que produce la cortadura». Así, a cada cortadura le corresponde uno y un solo número, racional o irracional.

El lenguaje de Dedekind al introducir los números irracionales deja un poco que desear. Introduce el irracional α como correspondiente a la cortadura y definido por ella, pero no está demasiado claro de dónde viene α . Debería decir que el número irracional α no es otra cosa que la cortadura misma. De hecho, Heinrich Weber le dijo esto a Dedekind, y este le replicó en una carta de 1888 que el número irracional α no es la cortadura misma, sino algo distinto, que corresponde a la cortadura y que la produce. De modo parecido, aunque los números racionales generan cortaduras, *no* son lo mismo que ellas. Y añadía que tenemos el poder mental de crear tales conceptos.

A continuación define cuándo una cortadura es menor o mayor que otra cortadura (B_1, B_2) . Después de haber definido la desigualdad, señala que los números reales poseen tres propiedades demostrables:

1. si $\alpha > \beta$ y $\beta > \gamma$, entonces $\alpha > \gamma$.
2. Si α y γ son dos números reales diferentes, entonces hay una cantidad infinita de números diferentes entre α y γ .
3. Si α es cualquier número real, entonces los números reales

pueden dividirse en dos clases A_1 y A_2 , cada una de las cuales contiene una infinita de elementos, y cada elemento de A_1 es menor que a y cada elemento de A_2 es mayor que a . El número a mismo puede ser asignado a cualquiera de las dos clases. El conjunto de los números reales posee ahora *continuidad*, que Dedekind expresa así: si se divide el conjunto de todos los números reales en dos clases A_1 y A_2 tales que cada elemento de A_1 es menor que todos los elementos de A_2 , entonces existe un y sólo un número a que produce esa división.

Define después las operaciones con números reales. La suma de las cortaduras (A_1, A_2) y (B_1, B_2) se define así: si c es cualquier número racional, lo pondremos en la clase C_1 si hay un número a_1 en A_1 , y un número b_1 en B_1 tales que $a_1 + b_1 \geq c$. Pondremos todos los demás números racionales en la clase C_2 . Ese par de clases C_1 y C_2 constituyen una cortadura (C_1, C_2) , puesto que cada elemento de C_1 es menor que cada elemento de C_2 . La cortadura (C_1, C_2) es entonces la suma de (A_1, A_2) y (B_1, B_2) . Las otras operaciones, dice, se definen análogamente. Ahora ya puede establecer propiedades de la suma y la multiplicación, como la propiedad asociativa y conmutativa. Aunque la teoría de Dedekind de los números irracionales, con pequeñas modificaciones como la señalada más arriba, es lógicamente satisfactoria, Cantor la criticó porque las cortaduras no aparecen de manera natural en análisis.

Hay otros enfoques de la teoría de los números irracionales además de los ya mencionados o descritos. Por ejemplo, en 1696 Wallis

había identificado números racionales y números decimales periódicos. Otto Stolz (1842-1905), en su *Vorlesungen über allgemeine Arithmetik*²⁹³ mostró que cada número irracional puede representarse como un decimal no periódico, y esto puede utilizarse como propiedad definitoria.

Queda claro de estos varios enfoques que la definición lógica del número irracional es bastante refinada. Desde el punto de vista lógico un número irracional no es simplemente un solo símbolo o un par de símbolos, tal como una razón entre dos enteros, sino una colección infinita, como una sucesión fundamental de Cantor o una cortadura de Dedekind. El número irracional, definido lógicamente, es un monstruo intelectual, y podemos comprender por qué los griegos y tras ellos tantas generaciones de matemáticos los encontraron tan difíciles de manejar.

Los avances en matemáticas no son recibidos con universal aprobación. Hermann Hankel, creador él mismo de una teoría lógica de los números racionales, objetó a las teorías de los números irracionales²⁹⁴: «Cualquier intento de tratar los números irracionales formalmente y sin el concepto de magnitud (geométrica) conduce necesariamente a los artificios más abstrusos e incómodos, e incluso si son llevados a cabo con completo rigor, sobre lo que tenemos perfecto derecho a dudar, no tienen mucho valor científico.»

4. La teoría de los números racionales

La siguiente etapa en la fundamentación del sistema numérico fue

la definición y deducción de las propiedades de los números racionales. Como ya hemos señalado, algún esfuerzo en esa dirección precedió al trabajo sobre los números irracionales. La mayoría de los que escribieron sobre números racionales supusieron que la naturaleza y propiedades de los enteros ordinarios eran conocidas, y que el problema era establecer lógicamente los números negativos y las fracciones.

El primer intento fue realizado por Martin Ohm (1792-1872), profesor en Berlín y hermano del físico, en su *Versuch eines vollkommen consequenten Systems der Mathematik* (Estudio de un sistema completo y consistente de las Matemáticas, 1822). Más tarde, Weierstrass, en sus lecciones de los años sesenta, presentaba los números racionales a partir de los naturales introduciendo los racionales positivos como pares de números naturales, los enteros negativos como otro tipo de pares de números naturales, y los racionales negativos como pares de enteros negativos. Esta idea fue utilizada independientemente por Peano, y la presentaremos más tarde con más detalle en relación con la obra de este último. Weierstrass no sintió la necesidad de clarificar la lógica de los números naturales. De hecho, su teoría de los números racionales no estaba exenta de dificultades. No obstante, en sus lecciones desde 1859 en adelante, afirmó correctamente que una vez admitidos los números naturales, ya no había necesidad de más axiomas para construir los números reales.

El problema clave para la construcción del sistema de los números racionales consistía en fundamentar los enteros ordinarios por

algún procedimiento y en establecer sus propiedades. Entre los que trabajaron sobre la teoría de los enteros, algunos creían que eran tan fundamentales que no podía hacerse ningún análisis lógico de ellos. Esta posición fue adoptada por Kronecker, motivado por consideraciones filosóficas en las que después entraremos más profundamente. Kronecker también deseaba aritmetizar el análisis, esto es, fundamentar el análisis sobre los enteros; pero pensaba que no se podía ir más allá del reconocimiento de su existencia. El hombre los poseería mediante alguna especie de intuición fundamental. «Dios hizo los enteros», decía, «el resto es obra del hombre».

Dedekind ofreció una teoría de los enteros en su *Was sind und was sollen die Zahlen*.²⁹⁵ Aunque fue publicada en 1888, la obra data de los años que van de 1872 a 1878. Utilizó en ella ideas de la teoría de conjuntos que Cantor había expuesto ya para esa época, y que iban a adquirir gran importancia. Sin embargo, el enfoque de Dedekind era tan complicado que se le prestó poca atención.

La aproximación a los enteros que mejor se adaptaba a las proclividades axiomatizadoras de finales del siglo XIX consistía en introducirlos mediante un conjunto de axiomas. Desconocedor en ese momento de los resultados obtenidos por Dedekind en la obra mencionada más arriba, Giuseppe Peano (1858-1932) también lo hizo en su *Arithmetices Principia Nova Methodo Expósita* (1889).²⁹⁶ Como el enfoque de Peano es muy frecuentemente utilizado, lo expondremos aquí.

Buscando la precisión del razonamiento, utilizaba grandes dosis de

simbolismo. Así, ε significa pertenecer a; \supset significa implica; N_0 denota la clase de los números naturales, y a^+ el número natural que viene a continuación de a . Peano utilizó este simbolismo en su presentación de todas las matemáticas, en particular en su *Formulario mathematico* (5 vols., 1895-1908), y también en sus clases, lo que provocó una revuelta de los estudiantes. Trató entonces de satisfacerles aprobándoles a todos, pero esto no funcionó, y fue obligado a renunciar a su puesto de profesor en la academia militar, permaneciendo a partir de entonces en la Universidad de Turín.

Aunque la obra de Peano influyó en el desarrollo posterior de la lógica simbólica y en el intento de Frege y Russell de asentar las matemáticas sobre la lógica, su obra debe distinguirse de la de estos últimos. Peano *no* quería fundamentar las matemáticas en la lógica, que para él sólo era una disciplina auxiliar de la matemática.

Peano partía de los conceptos no definidos (vid. cap. 42, sec. 2) de «conjunto», «número natural», «sucesor» y «pertenecer a». Sus cinco axiomas para los números naturales son:

1. 1 es un número natural.
2. 1 no es el sucesor de ningún otro número natural.
3. Cada número natural a tiene un sucesor.
4. Si los sucesores de a y b son iguales lo mismo pasa con a y b .
5. Si un conjunto S de números naturales contiene a 1, y cuando contiene a algún número natural a también contiene al sucesor de a , entonces S contiene todos los números naturales.

Este último axioma es el de la inducción matemática.

Peano también adoptó los axiomas de reflexividad, simetría y transitividad para la igualdad. Esto es, $a = a$; si $a = b$, entonces $b = a$; y si $a = b$ y $b = c$, entonces $a = c$. Definió la adición estableciendo que para cada par de números naturales a y b hay una suma única tal que

$$a + 1 = a +$$

$$a + (b+) = (a + b) +$$

De modo parecido, la multiplicación quedaba definida estableciendo que para cada par de números naturales a y b hay un producto único tal que

$$a \cdot 1 = a$$

$$a \cdot (b+) = (a \cdot b) + a$$

A continuación exponía todas las propiedades acostumbradas de los números naturales.

A partir de los números naturales y sus propiedades es ya muy simple definir y establecer las propiedades de los números enteros negativos y de los números racionales. Se pueden definir primero los enteros positivos y negativos como una nueva clase de números,

cada uno de ellos como un par ordenado de números naturales. Así (a,b) , donde a y b son números naturales, es un entero. El significado intuitivo de (a,b) es $a - b$. Cuando $a > b$, el par representa un entero positivo ordinario, y cuando $a < b$, un entero negativo. Las definiciones adecuadas de las operaciones de adición y multiplicación conducen a las propiedades habituales de los enteros positivos y negativos.

Dados los enteros, se introducen los números racionales como pares ordenados de enteros. Así, si A y B son enteros, el par ordenado (A, B) es un número racional. Intuitivamente (A, B) es A/B . Una vez más, las definiciones adecuadas de adición y multiplicación de esos pares conducen a las propiedades usuales de los números racionales.

Así, una vez se hubo alcanzado el enfoque lógico de los números naturales, el problema de construir una fundamentación para el sistema numérico real quedaba resuelto. Como ya hemos señalado, las personas que trabajaron en la teoría de los irracionales suponían generalmente que los números racionales eran tan perfectamente conocidos que podían tomarlos como punto de partida, o en todo caso que bastaría con un pequeño esfuerzo para clarificarlos. Después de que Hamilton hubiera basado los números complejos en los reales, y después de que los irracionales hubieran sido definidos en función de los racionales, se creó finalmente la lógica de esta última clase de números. El orden histórico fue esencialmente el inverso del orden lógico requerido para la construcción del sistema de los números complejos.

5. Otros enfoques del sistema de los números reales

La esencia de los enfoques de la fundamentación lógica del sistema de los números reales descritos hasta el momento reside en la obtención de los enteros y sus propiedades por algún procedimiento para luego obtener, a partir de ellos, las fracciones y más tarde los números irracionales. La base lógica de este enfoque consiste en ciertas series de asertos únicamente acerca de los números naturales; por ejemplo, los axiomas de Peano. Para todos los demás números se ofrece un proceso de construcción. Hilbert llamó a ese enfoque «método genético» (puede que no conociera en ese momento los axiomas de Peano, pero sí otros similares para los números naturales). Reconociendo que el método genético puede tener valor pedagógico o heurístico, consideraba sin embargo como más seguro desde el punto de vista de la lógica la aplicación del método axiomático al sistema completo de los números reales. Antes de examinar sus razones, echemos un vistazo a sus axiomas²⁹⁷.

Introduce el término no definido «número», denotado por a, b, c, \dots , y plantea a continuación los siguientes axiomas:

I. Axiomas de conexión

I_1 . A partir del número a y el número b se obtiene por adición un determinado número c ; simbólicamente

$$a + b = c \quad \text{o} \quad c = a + b.$$

I_2 . Si a y b son números dados, existe uno y sólo un número x y existe también un y sólo un número y tales que

$$a + x = b \quad \text{y} \quad y + a = b.$$

I_3 . Hay un número determinado, denotado por 0 , tal que para cualquier a

$$a + 0 = a \quad \text{y} \quad 0 + a = a.$$

I_4 . A partir del número a y el número b se obtiene por otro método, la multiplicación, un número determinado c ; simbólicamente

$$ab = c \quad \text{o} \quad c = ab.$$

I_5 . Dados dos números arbitrarios, a y b , si a no es 0 , existe uno y sólo un número x , y también uno y sólo un número y , tales que

$$ax = b \quad \text{a} \quad ya = b.$$

I_6 . Existe un número determinado, denotado por 1 , tal que para cada a tenemos

$$a \cdot 1 = a \quad \text{y} \quad 1 \cdot a = a.$$

II. Axiomas de cálculo

$$II_1. a + (b + c) = (a + b) + c$$

$$II_2. a + b = b + a.$$

$$II_3. a(bc) = (ab)c.$$

$$II_4. a(b + c) = ab + ac.$$

$$II_5. (a + b)c = ac + bc.$$

$$II_6. ab = ba.$$

III. Axiomas de orden

III₁. Si a y b son dos números diferentes, entonces uno de ellos es siempre mayor que el otro; este último se dice que es más pequeño que el primero; simbólicamente

$$a > b \quad \text{y} \quad b < a.$$

III₂. Si $a > b$ y $b > c$, entonces $a > c$.

III₃. Si $a > b$, entonces siempre es cierto que

$$a + c > b + c \quad \text{y} \quad c + a > c + b.$$

III₄. Si $a > b$ y $c > 0$, entonces $ac > bc$ y $ca > cb$.

IV. Axiomas de continuidad

IV₁. (Axioma de Arquímedes). Si $a > 0$ y $b > 0$ son dos números arbitrarios, entonces es posible siempre sumar a consigo mismo un número suficiente de veces para tener

$$a, + a + \dots + a > b.$$

*IV*₂. (Axioma de completitud). No es posible añadir al sistema de los números ninguna colección de cosas de manera que la colección resultante siga satisfaciendo los axiomas precedentes; dicho en pocas palabras, los números forman una colección de objetos que no puede ampliarse sin que deje de cumplirse alguno de los axiomas precedentes.

Hilbert señala que esos axiomas no son independientes; se pueden deducir algunos de ellos de los otros. Afirma a continuación que la objeción contra la existencia de conjuntos infinitos (vid. sec. 6) no es válida para esta concepción de los números reales, ya que, dice, no tenemos que pensar en la colección de todas las leyes posibles según las que se pueden formar los elementos de una sucesión fundamental (sucesiones de Cantor de números racionales), sino únicamente en un sistema cerrado de axiomas y conclusiones que se pueden deducir de ellos tras un número finito de pasos lógicos. Señala desde luego que es necesario probar la consistencia de ese conjunto de axiomas, pero que desde el momento en que eso esté hecho, los objetos que los axiomas definen, es decir, los números reales, existirán desde el punto de vista matemático. Hilbert no era consciente en ese momento de la dificultad de probar la consistencia de los axiomas que había propuesto para los números reales.

A la afirmación de Hilbert de que el método axiomático es superior al genético, Bertrand Russell objetaba que ofrece la misma ventaja

que el robo sobre el trabajo honrado: se arroga de inmediato todo lo que se puede construir por argumentos deductivos a partir de un conjunto mucho menor de axiomas.

Como en el caso de casi todos los avances significativos en matemáticas, la creación de la teoría de los números reales suscitó cierta oposición. Du Bois-Reymond, a quien ya hemos citado como oponente a la aritmetización del análisis, escribió en su *Théorie générale des fonctions* de 1887:²⁹⁸

Sin duda, con ayuda de los llamados axiomas, a partir de convenios, con proposiciones filosóficas construidas *ad hoc*, extendiendo ininteligiblemente conceptos originalmente claros, se puede construir un sistema aritmético que se parece en todos los aspectos al que se obtiene a partir del concepto de magnitud, para aislar así la matemática computacional, por decirlo de algún modo, mediante un cordón sanitario de dogmas y definiciones defensivas... Pero de esa forma se podrían inventar también otros sistemas aritméticos. La aritmética ordinaria no es otra que la que corresponde al concepto de magnitud lineal.

A pesar de ataques como éste, a los matemáticos les pareció que la culminación de los trabajos sobre los números reales resolvía todos los problemas lógicos que se habían presentado en el asunto. La aritmética, el álgebra y el análisis constituían en conjunto la mayor parte de las matemáticas, y esa parte quedaba ahora firmemente asentada.

6. El concepto de conjunto infinito

La rigorización del análisis había revelado la necesidad de comprender la estructura de los conjuntos de números reales. Para abordar este problema, Cantor había introducido ya (cap. 40, sec. 6) algunas nociones acerca de conjuntos infinitos de puntos, en particular los conjuntos de primera especie. Decidió que el estudio de los conjuntos infinitos era tan importante que debía dedicarse a examinar los conjuntos infinitos como tales, lo que le capacitaría, según creía, para distinguir claramente los diferentes conjuntos infinitos de discontinuidades.

La dificultad crucial en la teoría de conjuntos es el concepto mismo de conjunto infinito. Tales conjuntos habían suscitado naturalmente la atención de los matemáticos y los filósofos desde los tiempos de la antigua Grecia en adelante, y su naturaleza misma, así como sus propiedades aparentemente contradictorias, habían impedido cualquier progreso en su comprensión. Las paradojas de Zenón habían sido quizá la primera indicación de las dificultades. Ni la divisibilidad infinita de la recta ni la concepción de esta misma como conjunto infinito de puntos discretos parecían permitir conclusiones razonables acerca del movimiento. Aristóteles consideró conjuntos infinitos, como el de los números naturales, y negó la existencia de un conjunto infinito de objetos como entidad fija. Para él, los conjuntos sólo podían ser potencialmente infinitos (vid. cap. 3, sec. 10).

Proclo, el comentador de Euclides, señaló que como el diámetro de un círculo lo divide en dos mitades y hay un número infinito de diámetros, tendría que haber dos veces ese número de mitades.

Esto les parece a muchos una contradicción, dice Proclo, y la resuelve diciendo que no se puede hablar de un infinito actual de diámetros o de mitades en un círculo, sino sólo de un número cada vez más grande de diámetros o de mitades. Con otras palabras, Proclo aceptaba el concepto aristotélico de un infinito potencial pero no actual, lo que evita el problema de una infinidad doble que iguala a una infinidad.

A lo largo de la Edad Media, los filósofos se situaron de un lado u otro con respecto a la cuestión de si puede haber una colección infinita actual de objetos. Se señaló que los puntos de dos circunferencias concéntricas podían ponerse entre sí en correspondencia uno a uno asociando los puntos del mismo radio. Y sin embargo una de las circunferencias es más larga que la otra.

Galileo se enfrentó a los conjuntos infinitos rechazándolos porque no podían ser sometidos a la razón. En su *Dos Nuevas Ciencias* (pp. 18-40 de la traducción inglesa) señala que los puntos de dos segmentos de distintas longitudes AB y CD pueden ponerse entre sí en correspondencia uno a uno (fig. 41.1) y por tanto contienen presumiblemente el mismo número de puntos. También señala que los números naturales pueden ponerse en correspondencia biunívoca con sus cuadrados asignando simplemente a cada número sus cuadrados; pero esto lleva a diferentes «cantidades» de infinito, lo que Galileo considera imposible. Todas las cantidades infinitas son la misma y no pueden compararse.

Gauss, en su carta a Schumacher del 12 de julio de 1831,²⁹⁹ dice:

«Protesto contra el uso de una cantidad infinita como entidad

actual; eso no está permitido nunca en matemáticas. El infinito es solamente una manera de hablar, en la que uno está realmente hablando de límites a los que ciertas razones pueden aproximarse tanto como se quiera, mientras que otras crecen ilimitadamente.» Cauchy, como otros antes que él, negaba la existencia de conjuntos infinitos porque el hecho de que una parte de ellos pudiera ponerse en correspondencia biunívoca con el todo le parecía contradictorio.

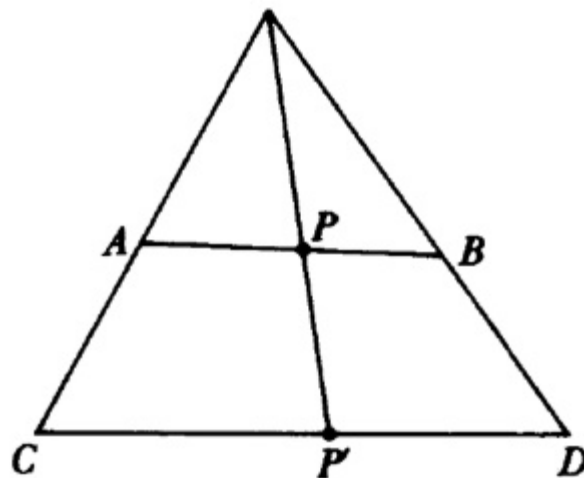


Figura 41.1

La polémica sobre los diferentes problemas concernientes a los conjuntos era interminable y recogía argumentos metafísicos e incluso teológicos. La actitud de la mayoría de los matemáticos hacia este problema consistía en ignorar lo que no podían resolver. En general, también evitaban el reconocimiento explícito de conjuntos actualmente infinitos, aunque utilizaban series infinitas, por ejemplo, y el sistema de números reales. Hablaban de los puntos de una recta evitando decir que ésta está formada por un

número infinito de puntos. Este escamoteo de los problemas incómodos era hipócrita, pero suficiente para construir el análisis clásico. Sin embargo, cuando el siglo XIX se enfrentó al problema de introducir el rigor en análisis, ya no se pudieron evitar muchas cuestiones acerca de los conjuntos infinitos.

7. Los fundamentos de la teoría de conjuntos

Bolzano, en sus *Paradojas del Infinito* (1851), publicada tres años después de su muerte, fue el primero en dar pasos decisivos hacia una definitiva teoría de conjuntos. Defendió la existencia de conjuntos actualmente infinitos, y puso el acento en la noción de equivalencia de dos conjuntos, con la que aludía a lo que más tarde se llamaría correspondencia biunívoca entre los elementos de ambos conjuntos. Esta noción de equivalencia se podía aplicar tanto a conjuntos finitos como infinitos. Señaló que en el caso de conjuntos infinitos una parte o subconjunto podía ser equivalente al total, e insistió en que esto debía ser aceptado. Así, los números reales entre 0 y 5 pueden ser puestos en correspondencia biunívoca con los números reales entre 0 y 12 mediante la fórmula $y = 12x/5$, a pesar de que el segundo conjunto de números contiene al primero. A los conjuntos infinitos se les podrían atribuir números, y habría diferentes números transfinitos para los diferentes conjuntos infinitos, si bien la asignación que Bolzano hacía de números transfinitos era incorrecta según la teoría posterior de Cantor.

La obra de Bolzano sobre el infinito era más filosófica que matemática, y no dejaba suficientemente clara la noción de lo que

más tarde se llamaría potencia o número cardinal de un conjunto. También él encontró propiedades que la parecieron paradójicas, y las cita en su libro. Decidió que los números transfinitos no eran necesarios para fundamentar el cálculo, y en consecuencia no prosiguió más lejos la tarea.

El verdadero creador de la teoría de conjuntos fue Georg Cantor (1845-1918), que había nacido en Rusia en una familia judeo-danesa, pero se trasladó a Alemania con sus padres. Su progenitor le instó a estudiar ingeniería, y Cantor ingresó en la Universidad de Berlín en 1863 con esa intención. Allí fue influido por Weierstrass y se volcó en la matemática pura. Llegó a *Privatdozent* en Halle en 1869, y a profesor en 1879. Cuando tenía sólo veintinueve años publicó su primer artículo revolucionario sobre la teoría de conjuntos infinitos en el *Journal für Mathematik*. Aunque algunas de sus proposiciones fueron consideradas erróneas por los matemáticos más viejos, su total originalidad y brillo atrajeron la atención. Siguió publicando artículos sobre la teoría de conjuntos y los números transfinitos hasta 1897.

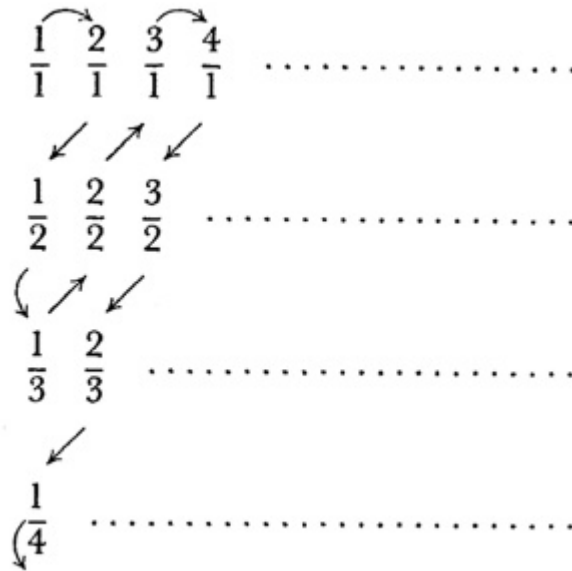
La obra de Cantor, que resolvía problemas muy antiguos y trastocaba mucho del pensamiento anterior, difícilmente podía recibir una aceptación inmediata. Sus ideas sobre los números transfinitos ordinales y cardinales suscitaron la hostilidad del poderoso Leopold Kronecker, que atacó brutalmente las ideas de Cantor durante más de una década. En cierto momento Cantor sufrió una crisis nerviosa, pero reemprendió el trabajo en 1887. Aunque Kronecker murió en 1891, sus ataques despertaron las

sospechas de los matemáticos acerca de la obra de Cantor.

Su teoría de conjuntos aparece dispersa en muchos artículos, y no intentaremos por tanto indicar el lugar específico en el que aparece cada uno de sus conceptos y teoremas. Esos artículos aparecieron en los *Mathematische Annalen* y el *Journal für Mathematik* desde 1874 en adelante.³⁰⁰ Por conjunto Cantor entiende una colección de objetos definidos y separados que puede ser concebida por la mente y sobre la que podemos decidir si un determinado objeto pertenece o no a ella. Mantiene que los que defienden conjuntos sólo potencialmente infinitos se equivocan, y refuta los argumentos previos de matemáticos y filósofos contra los conjuntos actualmente infinitos. Para Cantor un conjunto es infinito si puede ponerse en correspondencia biunívoca con una parte de sí mismo. Algunos de sus conceptos, como el de punto límite de un conjunto, o los de conjunto derivado, conjunto de primera especie, etc., fueron definidos y utilizados en un artículo sobre series trigonométricas,³⁰¹ que ya hemos descrito en el capítulo precedente (sec. 6). Un conjunto es cerrado si contiene todos sus puntos límite; es abierto si todos sus puntos son interiores, esto es, si cada uno de sus puntos puede ser incluido en un intervalo que sólo contiene puntos del conjunto; es perfecto si cada uno de sus puntos es un punto límite y además es cerrado. También definió la unión e intersección de conjuntos. Aunque estaba ante todo interesado en conjuntos de puntos de una recta o conjuntos de números reales, extendió esos conceptos a los conjuntos de puntos de un espacio euclídeo n -dimensional.

Trató después de distinguir los conjuntos infinitos según su «tamaño» y, como Bolzano, decidió que la correspondencia uno a uno debía ser el criterio básico. Dos conjuntos que pueden ponerse en correspondencia biunívoca son equivalentes, o tienen la misma potencia (más tarde, la palabra «potencia» sería reemplazada por «número cardinal»). Dos conjuntos pueden tener distinta potencia; si de dos conjuntos de objetos M y N , N puede ponerse en correspondencia biunívoca con un subconjunto de M , pero M no puede ponerse en correspondencia biunívoca con ningún subconjunto de N , la potencia de M es mayor que la de N .

Los conjuntos de números eran desde luego los más importantes, y por eso Cantor ilustra su concepto de equivalencia o potencia con tales conjuntos. Introduce el término «numerable» para cualquier conjunto que pueda ponerse en correspondencia biunívoca con los enteros positivos. Este es el conjunto infinito más pequeño. Cantor probó que el conjunto de los números racionales es numerable. Ofreció una demostración en 1874.³⁰² No obstante, su segunda demostración ³⁰³ es la más empleada, y es la que vemos a describir. Dispone los números racionales así:



Puede observarse que todos los que hay en cualquier diagonal tienen la misma suma de numerador y denominador. Parte entonces de 1/1 y sigue las flechas asignando el número 1 a 1/1, 2 a 2/1, 3 a 1/2, 4 a 1/3, y así sucesivamente. Cada número racional será alcanzado en algún momento, y a cada uno se le asignará un entero finito. Luego el conjunto de más arriba de números racionales (en el que algunos aparecen muchas veces) está en correspondencia uno a uno con los enteros. Si se suprimen las repeticiones, el conjunto de los números racionales será todavía infinito y necesariamente numerable, ya que éste es el menor conjunto infinito.

Todavía más sorprendente es la demostración de Cantor en el artículo citado de 1874 de que el conjunto de todos los números algebraicos, esto es, el conjunto de todos los números que son soluciones de ecuaciones algebraicas

$$a_0x^n + a_1x^{n-1} + \dots + a_n = 0,$$

donde los a_i son enteros, es también numerable.

Para probar esto, asigna a cada ecuación algebraica de grado n la altura N definida por

$$N = n - 1 + |a_0| + |a_1| + \dots + |a_n|,$$

donde los a_i son los coeficientes de la ecuación. La altura N es un entero. A cada N corresponden solamente un número finito de ecuaciones algebraicas, y por tanto solamente una cantidad finita de números algebraicos, digamos $\varphi(N)$. Por ejemplo, $\varphi(1) = 1$; $\varphi(2) = 2$; $\varphi(3) = 4$. A partir de $N = 1$, numera los correspondientes números algebraicos desde 1 hasta n_1 ; los números algebraicos de altura 2 son numerados desde $n_1 + 1$ hasta n_2 , y así sucesivamente. Como cada número algebraico será alcanzado en alguna etapa y se le asignará uno y sólo un entero, el conjunto de los números algebraicos es numerable.

En sus cartas a Dedekind en 1873, Cantor le planteó la cuestión de si el conjunto de los números reales podría ponerse en correspondencia uno a uno con los enteros, y algunas semanas más tarde le aseguró no era posible. Ofreció dos demostraciones. La primera (en el artículo de 1874 que ya hemos citado) era más complicada que la segunda,³⁰⁴ que es la que más a menudo se expone en la actualidad. Tiene también la ventaja, como señaló Cantor, de ser independiente de consideraciones técnicas acerca de los números irracionales.

La segunda demostración de Cantor de que los números reales son no numerables (no enumerables) comienza suponiendo que se pueden enumerar los números reales entre 0 y 1. Escribamos cada uno de ellos en forma decimal, conviniendo en que un número tal como $1/2$ se escribirá $0,4999\dots$. Si esos números reales son numerables, podremos asignar cada uno de ellos a un entero n , así:

$$1 \leftrightarrow 0, a_{11} a_{12} a_{13} \dots$$

$$2 \leftrightarrow 0, a_{21} a_{22} a_{23} \dots$$

$$3 \leftrightarrow 0, a_{31} a_{32} a_{33} \dots$$

Definamos ahora un número real entre 0 y 1 así: sea $b = 0. b_1 b_2 b_3, \dots$, donde $b_k = 9$ si $a_{kk} = 1$ y $b_k = 1$ si $a_{kk} \neq 1$. Ese número real difiere de cada uno de los escritos en la correspondencia de más arriba. Sin embargo, se suponía que esa lista contenía todos los números reales entre 0 y 1. Hemos llegado así a una contradicción. Como los números reales no son numerables y los números algebraicos son numerables, tiene que haber irracionales trascendentes. Esta es la demostración de existencia no constructiva de Cantor, que podríamos comparar con la construcción efectiva de irracionales trascendentes de Liouville (sec. 2).

En 1874 Cantor se ocupó de la equivalencia entre los puntos de una recta y los de \mathbb{R} (espacio n -dimensional) tratando de probar que una correspondencia biunívoca entre esos dos conjuntos era imposible.

Tres años más tarde comprobó que sí existe tal correspondencia, escribiendo a Dedekind: «Lo veo, pero no lo creo.»³⁰⁵²⁶

La idea utilizada para establecer esa correspondencia uno a uno ³⁰⁶²⁷ puede generalizarse fácilmente a partir de una correspondencia entre los puntos del cuadrado unidad y los del segmento (0,1). Sean (x,y) un punto del cuadrado unidad y z un punto del intervalo unidad. Representemos x e y en forma decimal, reemplazando los ceros de un decimal finito por una sucesión infinita de nueves. A continuación separemos en x e y grupos de cifras decimales que terminen con la primera cifra no nula de la sucesión. Por ejemplo,

$$x = 0, 3 \ 002 \ 03 \ 04 \ 6\dots$$

$$y = 0, 01 \ 6 \ 07 \ 8 \ 09\dots$$

Formemos entonces

$$z = 0, 3 \ 01 \ 002 \ 6 \ 03 \ 07 \ 04 \ 8 \ 6 \ 09\dots$$

eligiendo como grupos en la expresión decimal de z el primer grupo de x , a continuación el primero grupo de y , y así sucesivamente. Si las coordenadas x o y de dos puntos del cuadrado difieren en alguna cifra decimal, los z correspondientes serán diferentes. Luego para cada (x,y) hay un único z . Recíprocamente, dado un z , se pueden separar en su expresión decimal los grupos descritos anteriormente, formando x e y , de manera que para cada z hay un

único (x,y) . La correspondencia biunívoca que acabamos de establecer no es continua; *grosso modo*, eso significa que a puntos z próximos no les corresponden necesariamente puntos (x,y) próximos, ni a la inversa.

Du Bois-Reymond objetó a esta demostración:³⁰⁷ ²⁸ «Repugna al sentido común. De hecho, se trata simplemente de la conclusión de un tipo de razonamiento que permite la intervención de ficciones ideales, a las que se hace jugar el papel de cantidades genuinas aunque no sean siquiera límites de representaciones de cantidades. Ahí es donde reside la paradoja.»

8. Cardinales y ordinales transfinitos

Habiendo demostrado la existencia de conjuntos con la misma potencia y con diferentes potencias, Cantor prosiguió con este concepto de potencia de un conjunto e introdujo una teoría de números cardinales y ordinales en la que los elementos destacados son los cardinales y ordinales transfinitos. Cantor desarrolló este trabajo en una serie de artículos en los *Mathematische Annalen* desde 1879 hasta 1884, bajo el título común «Uber unendliche lineare Punkt- mannichfaltigkeiten» (sobre agregados lineales infinitos de puntos). Después escribió dos artículos definitivos en 1895 y 1897 en la misma revista.³⁰⁸

En el quinto artículo sobre agregados lineales³⁰⁹, Cantor comienza con la observación:

La descripción de mis investigaciones en la teoría de agregados ha alcanzado un estado en el que su prolongación depende de

una generalización de los enteros reales positivos más allá de sus límites actuales; una generalización en una dirección en la que, por lo que yo sé, nadie se ha aventurado todavía.

Dependo de esa generalización del concepto de número hasta tal punto que sin ella no podría dar ni siquiera pequeños pasos adelante en la teoría de conjuntos. Espero que esta situación justifique o, si es necesario, excuse la introducción de ideas aparentemente tan extrañas en mis argumentaciones. De hecho, el objetivo consiste en generalizar o extender la serie de los enteros reales más allá del infinito. Por atrevido que esto pueda parecer, tengo no sólo la esperanza, sino la firme convicción de que a su debido tiempo esta generalización será reconocida como un paso bastante simple, apropiado y natural. Aun así, soy muy consciente de que adoptando tal procedimiento me sitúo a contracorriente con respecto a las opiniones generales sobre el infinito en matemáticas y sobre la naturaleza de los números.

Señala que su teoría de los números infinitos o transfinitos es distinta del concepto de infinidad bajo el que se entiende una variable que se hace infinitamente pequeña o infinitamente grande. Dos conjuntos que están en correspondencia uno a uno tienen la misma potencia, o el mismo número cardinal. Para conjuntos finitos el número cardinal es el número usual de objetos del conjunto. Para conjuntos infinitos se introducen nuevos números cardinales. El denotó por \aleph_0 al del conjunto de los números enteros. Como los

números reales no se pueden poner en correspondencia biunívoca con los números enteros, el conjunto de los números reales debe tener otro cardinal distinto, que se denota por c , la inicial de continuo. Como en el caso del concepto de potencia, si dos conjuntos M y N son tales que N puede ponerse en correspondencia biunívoca con un subconjunto de M , pero M no puede ponerse en correspondencia biunívoca con ningún subconjunto de N , el número cardinal de M es mayor que el de N . Así, $c > \aleph_0$.

Para obtener un número cardinal mayor que otro dado,³¹⁰ se considera cualquier conjunto M que represente a ese cardinal dado. Sea entonces N el conjunto de todos los subconjuntos de M . Entre esos subconjuntos están los elementos individuales de M , los pares de elementos de M , y así sucesivamente. Ahora bien, es evidentemente posible establecer una correspondencia biunívoca entre M y un subconjunto de N , el constituido por los elementos individuales de M (considerados como subconjuntos de M , y como elementos de N). Por otra parte, no es posible establecer una correspondencia uno a uno entre M y N . Supongamos efectivamente que sí la hubiera; sea m cualquier elemento de M y consideremos aquellos m tales que el elemento de N asociado por la supuesta correspondencia *no* contiene a ese m . Sea η el conjunto de tales m , que naturalmente es un elemento de N . Cantor afirma entonces que η no está incluido en la supuesta correspondencia uno a uno, ya que si η correspondiera a algún m de M , y η contuviera a m , tendríamos una contradicción con la definición de η . Pero si m no pertenece a η , que es el elemento de N que supuestamente le

corresponde, como η era por definición el conjunto de todos los m no contenidos en los correspondientes elementos de N , tendría que contener a ese m . Así, la suposición de que hay una correspondencia biunívoca entre los elementos de M y los de N , que son los subconjuntos de M , conduce a una contradicción. Luego el número cardinal del conjunto que consiste en todos los subconjuntos de un conjunto dado es mayor que el cardinal de este último.

Cantor definió la suma de dos números cardinales como el cardinal de la unión de dos conjuntos (disjuntos) que representen a los sumandos. También definió el producto de dos cardinales cualesquiera: dados dos cardinales α y β , se toman un representante M de α y un representante N de β , y se forman los pares de elementos (m,n) donde m es un elemento de M y n de N . El producto de α y β es entonces el cardinal del conjunto de todos esos posibles pares.

Se definen también potencias de números cardinales. Si tenemos un conjunto M de m objetos y otro N de n objetos, Cantor define el conjunto m^n de las permutaciones de m objetos tomando n cada vez y permitiendo repeticiones de los m objetos iniciales. Así, por ejemplo, si $m = 3$ y $n = 2$, con los objetos m_1, m_2, m_3 , las permutaciones que resultan son

$$\begin{array}{ccc}
 m_1 m_1 & m_2 m_2 & m_3 m_1 \\
 m_1 m_2 & m_2 m_1 & m_3 m_2 \\
 m_1 m_3 & m_2 m_3 & m_3 m_3
 \end{array}$$

Cantor define α^β como el cardinal de este conjunto de permutaciones, siendo α el cardinal de M y β el de N . A continuación demuestra que $2^{\aleph_0} = c$, cardinal del continuo.

Cantor llama la atención sobre el hecho de que su teoría de números cardinales se aplica en particular a los cardinales finitos, proporcionando así «el fundamento más corto y más riguroso para la teoría de números finitos».

El siguiente concepto es el de número ordinal. Había encontrado ya la necesidad de tal concepto al introducir los sucesivos conjuntos derivados de un conjunto dado de puntos. Ahora los introduce abstractamente. Un conjunto está totalmente ordenado si para cada dos elementos uno de ellos precede al otro, de manera que dados m_1 y m_2 , o bien m_1 precede a m_2 , o m_2 precede a m_1 : la notación es $m_1 < m_2$ o $m_2 < m_1$. Además, si $m_1 < m_2$ y $m_2 < m_3$, ese orden total también implica que $m_1 < m_3$; esto es, la relación de orden es transitiva. El número ordinal de un conjunto ordenado M es el tipo de orden del orden definido en el conjunto. Dos conjuntos ordenados son semejantes si hay una correspondencia biunívoca entre ellos y si, cuando a m_1 le corresponde n_1 y a m_2 le corresponde n_2 , y $m_1 < m_2$, entonces $n_1 < n_2$. Dos conjuntos ordenados semejantes tienen el mismo tipo o número ordinal. Como ejemplos de conjuntos

ordenados podemos utilizar cualquier conjunto finito de números en cualquier orden dado. Para un conjunto finito, sea cual sea el orden, el número ordinal es el mismo y se puede tomar como símbolo para él el número cardinal del conjunto. El número ordinal del conjunto de los enteros positivos en su orden natural se denota por ω . El mismo conjunto de los enteros positivos, pero en orden decreciente, esto es,

$$\dots, 4, 3, 2, 1$$

se denota por $^*\omega$. El conjunto de todos los enteros, positivos y negativos, más el cero, en su orden natural, tiene como número ordinal $^*\omega + \omega$.

Cantor define a continuación la suma y el producto de números ordinales. La suma es el número ordinal del conjunto ordenado que se obtiene posponiendo al primero el segundo de los conjuntos dados, manteniendo en cada uno de ellos el orden original. Así, el conjunto de los enteros positivos seguido de los primeros cinco enteros, esto es,

$$1, 2, 3, \dots, 1, 2, 3, 4, 5$$

tiene como número ordinal $\omega + 5$. También se definen de manera bastante obvia la igualdad y desigualdad de números ordinales.

Y ahora es cuando introduce el conjunto de todos los ordinales transfinitos, en parte por su propia importancia, y en parte para

poder definir con precisión números cardinales transfinitos más elevados. Para introducir esos nuevos ordinales restringe los conjuntos totalmente ordenados a conjuntos bien ordenados.³¹¹³² Un conjunto está bien ordenado si tiene un primer elemento en la ordenación y también lo tiene cada uno de sus subconjuntos. Hay una jerarquía de números ordinales y números cardinales. En la primera clase, denotada por Z_1 están los ordinales finitos

$$1, 2, 3, \dots$$

En la segunda clase, denotada por Z_2 , están los ordinales

$$\omega, \omega+1, \omega+2, \dots, 2\omega, 2\omega+1, \dots, 3\omega, 3\omega+1, \dots, \omega^2, \omega^3, \dots, \omega^\omega, \dots$$

Cada uno de esos ordinales es el ordinal de un conjunto cuyo cardinal es \aleph_0

El conjunto de los ordinales que hay en Z_2 tiene a su vez un número cardinal. Ese conjunto no es numerable, y Cantor introduce para él un nuevo número cardinal \aleph_1 , que muestra que es el primer cardinal posterior a \aleph_0 .

Los ordinales de la tercera clase, denotada por Z_3 , son

$$\Omega, \Omega + 1, \Omega+2, \dots, \Omega + \Omega, \dots$$

Estos son números ordinales de conjuntos bien ordenados, cada uno de los cuales tiene \aleph_1 elementos. Sin embargo, el conjunto de

ordinales Z_3 tiene más de \aleph_1 elementos, y Cantor denota su cardinal por \aleph_2 . Esta jerarquía de ordinales y cardinales puede proseguirse indefinidamente.

Cantor había mostrado que dado cualquier conjunto siempre es posible crear otro, el conjunto de las partes del conjunto dado, cuyo cardinal es mayor que el de este último. Si el conjunto dado es \aleph_0 , el número cardinal del conjunto de sus partes es 2^{\aleph_0} . Cantor había demostrado también que $2^{\aleph_0} = c$, el cardinal del continuo. Por otra parte había introducido \aleph_1 mediante los números ordinales, y había probado que \aleph_1 es el primer cardinal que sigue a \aleph_0 . Así pues, $\aleph_1 \leq c$, pero la cuestión era si $\aleph_1 = c$, cuestión conocida como hipótesis del continuo. A pesar de intentarlo esforzadamente, Cantor no pudo dar una respuesta. Hilbert incluyó esta cuestión en una lista de problemas sobresalientes presentada en el Congreso Internacional de Matemáticos de 1900 (vid. cap. 43, sec. 5 y cap. 51, sec. 8).

Para dos conjuntos en general M y N cabe la posibilidad de que M no pueda ponerse en correspondencia biunívoca con ningún subconjunto de N , sin que tampoco N se pueda poner en correspondencia biunívoca con ningún subconjunto de M . En tal caso, para los números cardinales α y β de M y N no se puede decir que $\alpha = \beta$, $\alpha < \beta$ o $\alpha > \beta$, esto es, esos dos cardinales no son comparables. Cantor fue capaz de probar que esta situación no puede darse si los conjuntos están bien ordenados. Parecía paradójico que pudiera haber conjuntos no bien ordenados cuyos cardinales no pudieran compararse, pero tampoco pudo Cantor resolver este problema.

Ernst Zermelo (1871-1953) afrontó el problema de qué hacer para comparar los números cardinales de conjuntos que no estén bien ordenados. En 1904 ³¹² probó, ofreciendo en 1908 ³¹³ una segunda demostración, que a todo conjunto se le puede dar una buena ordenación. En la demostración tuvo que utilizar lo que se conoce ahora como axioma de elección (axioma de Zermelo), según el cual, dada cualquier colección de conjuntos no vacíos disjuntos, es posible elegir en cada uno de ellos un elemento para formar un nuevo conjunto. El axioma de elección, el teorema sobre la buena ordenación, y el hecho de que dos conjuntos cualesquiera pueden compararse en cuanto a su tamaño (esto es, que si sus cardinales son α y β , o es $\alpha = \beta$, o $\alpha < \beta$, o bien $\alpha > \beta$, son principios equivalentes.

9. La situación de la teoría de conjuntos hacia 1900

La teoría de conjuntos de Cantor constituyó un audaz paso adelante en un dominio que, como ya hemos señalado, había sido explorado intermitentemente desde el tiempo de los griegos. Exigía una aplicación estricta de argumentos puramente racionales, y afirmaba la existencia de conjuntos infinitos de potencia tan elevada como se quiera, completamente fuera del alcance de la intuición humana. Habría sido extraño que tales ideas, mucho más revolucionarias que la mayoría de las introducidas anteriormente, no hubieran encontrado oposición. Las dudas en cuanto a la coherencia de este desarrollo se vieron reforzadas por ciertos interrogantes planteados por el mismo Cantor y por algunos otros. En sus cartas a Dedekind

del 28 de julio y el 28 de agosto de 1899,³¹⁴ Cantor se preguntaba si el conjunto de todos los números cardinales podría ser realmente un conjunto, ya que entonces su cardinal sería mayor que cualquier otro. Pensó que la respuesta correcta debía ser la negativa, distinguiendo entre conjuntos consistentes e inconsistentes. Pero en 1897 Cesare Burali-Forti (1861-1931) señaló que la sucesión de *todos* los números ordinales, que está bien ordenada, debería tener como número ordinal el mayor de todos los ordinales.³¹⁵ Pero entonces ese número ordinal sería mayor que *todos* los números ordinales (Cantor había apreciado ya esta dificultad en 1895). Estos y otros problemas no resueltos, llamados paradojas, empezaron a hacerse sentir a finales del siglo XIX.

La oposición a las ideas de Cantor crecía. Kronecker, como ya hemos dicho, se opuso casi desde el principio. Félix Klein tampoco simpatizaba en absoluto con ellas. Poincaré³¹⁶ observaba críticamente: «Pero sucede que hemos encontrado ciertas paradojas, ciertas contradicciones aparentes que habrían hecho las delicias de Zenón de Elea y de la escuela de Megara... Creo, por mi parte, y no soy el único, que el punto delicado está en la introducción de objetos que no pueden definirse completamente con un número finito de palabras.» Se refería a la teoría de conjuntos como un interesante «caso patológico», y predecía (en el mismo artículo) que «las generaciones posteriores considerarán las *Mengenlehre* (de Cantor) como una enfermedad de la que uno se ha curado». Hermann Weyl decía de la jerarquía de alefs de Cantor que era como una niebla en medio de la niebla.

No obstante, muchos matemáticos prominentes quedaron impresionados por la utilización que podía ya hacerse de la nueva teoría. En el primer Congreso Internacional de Matemáticos en Zurich (1897), Adolf Hurwitz y Hadamard señalaron importantes aplicaciones al análisis de la teoría de los números transfinitos. Pronto se descubrieron nuevas aplicaciones en la teoría de la medida (cap. 44) y en topología (cap. 50). Hilbert difundió las ideas de Cantor en Alemania, en 1926 ³¹⁷ decía: «Nadie podrá expulsarnos del paraíso que Cantor creó para nosotros», y elogiaba la aritmética transfinita de Cantor como «el producto más impresionante del pensamiento matemático, una de las más bellas realizaciones de la actividad humana en el dominio de lo puramente inteligible». ³¹⁸ Bertrand Russell describió la obra de Cantor como «la que probablemente puede enorgullecer más a nuestra época».

Bibliografía

- Becker, Oskar: *Grundlagen der Mathematik in geschichtlicher Entwicklung*, Verlag Karl Alber, 1954, 217-316.
- Boyer, Cari B.: *Historia de la matemática*, Madrid, Alianza, 1986.
- Cantor, Georg: *Gesammelte Abhandlungen*, 1932, Georg Olms (reimpresión), 1962. -*Contributions to the Founding of the Theory of Transfinite Numbers*; Dover (reimpresión), sin fecha. Contiene una traducción al inglés de los dos artículos clave de 1895 y 1897 y una muy útil introducción de P. E. B. Jourdain.
- Cavadles, Jean: *Philosophie mathématique*, Hermann, 1962.

Contiene también la correspondencia entre Cantor y Dedekind, traducida al francés.

- Dedekind, R.: *Essays on the Theory of Numbers*, Dover (reimpresión), 1963. Contiene una traducción al inglés de los artículos de Dedekind «Stetigkeit und irrationale Zahlen» y «Was sind und was sollen die Zahlen». Ambos están también en sus *Werke*, 3, 314-334 y 335-391.
- Fraenkel, Abraham A.: «Georg Cantor», *Jahres, der Deut. Math.-Verein.*, 39, 1930, 189-266. Revisión histórica de la obra de Cantor.
- Helmholtz, Hermann von: *Counting and Measuring*, D. Van Nostrand, 1930. Traducción al inglés de la obra de Helmholtz, *Zahlen und Messen, Wissenschaftliche Abhandlungen*, 3, 356-391.
- Manheim, Jerome H.: *The Génesis of Point Set Topology*, MacMillan, 1964, 76-110.
- Meschkowski, Herbert: *Ways of Thought of Great Mathematicians*, Holden-Day, 1964, 91-104. -*Evolution of Mathematical Thought*, Holden-Day, 1965, caps. 4-5. -*Probleme des Unendlichen: Werk und Leben Georg Cantors*, F. Vieweg und Sohn, 1967.
- Noether, E., y J. Cavailles: *Briefwechsel Cantor-Dedekind*, Hermann, 1937.
- Peano, G.: *Opere scelte*, 3 vols., Edizioni Cremonese, 1957-1959.
- Schoenflies, Arthur M.: *Die Entwicklung der Mengenlehre und*

ihre An- wendungen, en dos partes, B. G. Teubner, 1908 y 1913.

- Smith, David Eugene: *A Source Book in Mathematics*, Dover (reimpresión), 1959, vol. 1, 35-45 y 99-106.
- Stammler, Gerhard: *Der Zahlbegriff seit Gauss*, Georg Olms, 1965.

Capítulo 42

Los fundamentos de la geometría

La geometría no es nada si no es rigurosa... Los métodos de Euclides son considerados casi universalmente como irreprochables desde el punto de vista del rigor.

H. J. S. Smith (1873)

Ha sido costumbre defender a Euclides, cuando se le ataca como libro de texto por su verbosidad, su oscuridad o su pedantería, con el argumento de su excelencia lógica, que supuestamente proporcionaría un entrenamiento incomparable a las jóvenes capacidades de razonamiento. Esta suposición pierde consistencia, sin embargo, cuando se la examina más de cerca. Sus definiciones no siempre definen, sus axiomas no siempre son indemostrables, sus

demostraciones requieren muchos axiomas de los que es enteramente inconsciente. Una prueba válida mantiene su poder demostrativo cuando no se dibuja ninguna figura, pero muchas de las demostraciones de Euclides no pasarían esa criba... El valor de su obra como obra maestra de la lógica se ha exagerado enormemente.

Bertrand Russell (1902)

Contenido:

- 1. Los defectos de Euclides*
 - 2. Contribuciones a la fundamentación de la geometría proyectiva*
 - 3. Los fundamentos de la geometría euclídea*
 - 4. Otros trabajos de fundamentación*
 - 5. Algunas cuestiones abiertas*
- Bibliografía*

1. Los defectos de Euclides

Las críticas a las definiciones y axiomas de Euclides (cap. 4, sec. 10) se remontan a sus primeros comentaristas conocidos, Pappus y

Proclo. Cuando los europeos comenzaron a interesarse por Euclides durante el Renacimiento, también apreciaron fallos. Jacques Peletier (1517-1582), en su *In Euclidis Elementa Geométrica Demonstrationum* (1557), criticó el uso que hace Euclides de la superposición para probar teoremas sobre congruencia. También el filósofo Arthur Schopenhauer manifestó en 1844 su sorpresa ante el hecho de que los matemáticos cuestionaran el postulado de las paralelas y no el axioma según el que las figuras que coinciden son iguales, argumentando que o bien las figuras coincidentes son automáticamente idénticas o iguales y entonces no se necesita ningún axioma, o bien la coincidencia es algo completamente empírico, que no pertenece a la pura intuición (*Anschauung*) sino a la experiencia sensorial externa. Además, el axioma presupone la movilidad de las figuras; pero lo que se puede mover en el espacio es materia, y queda por tanto fuera de la geometría. En el siglo XIX llegó a reconocerse generalmente que el método de superposición descansaba sobre axiomas no explicitados, o que debería reemplazarse por otro enfoque de la congruencia.

A algunos críticos no les gustaba como axioma la afirmación de que todos los ángulos rectos son iguales, y trataron de probarla, naturalmente sobre la base de los demás axiomas. Christophorus Clavius (1537-1612), uno de los editores de la obra de Euclides, señaló la ausencia de un axioma que garantizara la existencia de una cuarta proporcional para tres magnitudes dadas (cap. 4, sec. 5). Leibniz comentó acertadamente que Euclides se basaba en la intuición al afirmar (Libro I, Proposición 1) que dos circunferencias,

cada una de las cuales pasa por el centro de la otra, tienen un punto en común. Con otras palabras, Euclides suponía que una circunferencia tiene una cierta estructura continua que le hace poseer un punto en el que la otra circunferencia la corta.

También Gauss señaló las imperfecciones de la presentación que Euclides hacía de la geometría. En una carta a Wolfgang Bolyai del 6 de marzo de 1832,³¹⁹ le hacía notar que hablar de la parte del plano limitada por un triángulo exige una fundamentación adecuada. Y decía también: «En un desarrollo completo, palabras tales como «entre» deben basarse en concepto claros, cosa que puede hacerse, pero que yo no he encontrado en ningún sitio.» Gauss realizó otras críticas sobre la definición de línea recta³²⁰ y sobre la definición de plano como una superficie en la que debe estar contenida la recta que une dos puntos cualesquiera del plano.³²¹

Es bien sabido que pueden obtenerse muchas «demostraciones» de resultados falsos porque los axiomas de Euclides no indican dónde deben estar ciertos puntos con relación a otros. Está, por ejemplo, la «demostración» de que todo triángulo es isósceles: se construyen la bisectriz en A del triángulo ABC y la mediatriz del lado BC (fig. 42.1).

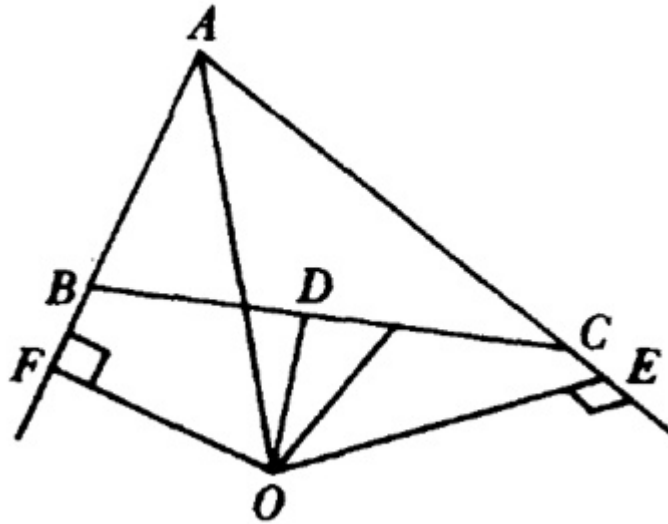


Figura 42.1

Si esas dos rectas son paralelas, la bisectriz es perpendicular a BC y el triángulo es isósceles. Supongamos, por el contrario, que esas dos rectas se cortan, digamos en O , y vamos entonces a «probar» que el triángulo sigue siendo isósceles. Dibujemos las perpendiculares OF a AB y OE a AC .

Los triángulos señalados con I son congruentes, y $OF = OE$. Los triángulos señalados con III son también congruentes, y $OB = OC$. En consecuencia, los triángulos marcados con II son congruentes, y $FB = EC$. En los triángulos señalados con I tenemos $AF = AE$. Luego $AB = AC$ y el triángulo es isósceles.

Uno podría preguntarse por la posición del punto O , y de hecho se puede probar que está en el círculo circunscrito pero fuera del triángulo. Sin embargo, si se dibuja la figura 42.2, todavía se puede «demostrar» que el triángulo ABC es isósceles.

El fallo está en que de los dos puntos E y F uno debe estar dentro y el otro fuera de los respectivos lados del triángulo. Pero esto

significa que debemos determinar la posición correcta de F con respecto a A y B y la de E con respecto a A y C antes de comenzar la demostración. Claro está que no habría que confiar en el dibujo para determinar las posiciones precisas de E y F , pero esto es precisamente lo que Euclides y los matemáticos anteriores a 1800 hacían. Se suponía que la geometría euclídea ofrecía pruebas rigurosas de teoremas sugeridos intuitivamente por las figuras, pero de hecho ofrecía demostraciones intuitivas a partir de figuras dibujadas rigurosamente.

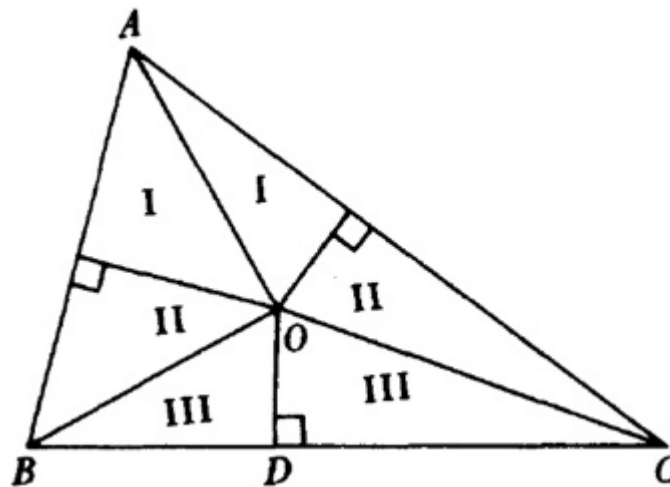


Figura 42.2

Aunque se alzaron críticas de la estructura lógica de los *Elementos* de Euclides casi desde el momento en que se escribió, no fueron ampliamente conocidas, o se consideraban sus defectos como menores. En general se tomaba a los *Elementos* como un modelo de rigor. No obstante, los trabajos sobre geometrías no euclídeas hicieron a los matemáticos conscientes del alcance de esas

deficiencias, exigiendo una actitud especialmente crítica de lo que presuponían al realizar una demostración. El reconocimiento de tales imperfecciones obligó finalmente a los matemáticos a emprender la reconstrucción de los fundamentos de la geometría euclídea y de otras geometrías afectadas de la misma debilidad. Esta actividad se difundió en el último tercio del siglo XIX.

2. Contribuciones a la fundamentación de la geometría proyectiva

En los años setenta del pasado siglo, los trabajos sobre geometría proyectiva en relación con las geometrías métricas revelaron que la fundamental es aquélla (vid. cap. 38). Quizá por esta razón el trabajo de fundamentación comenzó por la geometría proyectiva. Sin embargo, casi todos los autores estaban igualmente interesados en construir las geometrías métricas, ya fuera sobre la base de la geometría proyectiva o independientemente. Por eso los libros y artículos de finales del siglo XIX y comienzos del XX que tratan sobre los fundamentos de la geometría no pueden distribuirse según los diferentes temas.

Los trabajos sobre geometrías no euclídeas habían llevado al convencimiento de que las geometrías son construcciones humanas que se basan en el espacio físico sin ser necesariamente idealizaciones exactas de éste. Lo que implicaba que había que realizar varios cambios importantes en cualquier enfoque axiomático de la geometría. Esto fue reconocido y puesto de relieve por Moritz Pasch (1843-1930), que fue el primero en hacer

contribuciones notables a los fundamentos de la geometría. Sus *Vorlesungen über neuere Geometrie* (1.^a edición en 1882, 2.^a edición, revisada por Max Dehn, en 1926) son una obra pionera en este campo.

Pasch observó que las nociones comunes de Euclides, como las de punto y línea, en realidad no estaban definidas. Decir que un punto es lo que no tiene partes significa bien poco, porque ¿cuál es el significado de «parte»? De hecho, señalaba Pasch, como antes Aristóteles y unos pocos matemáticos más tardíos como Peacock y Boole, algunos conceptos deben quedar sin definición, o si no el proceso de definición sería interminable, o bien la matemática descansaría sobre conceptos físicos. Una vez que se seleccionan ciertos conceptos indefinidos, los demás deben definirse en términos de éstos. Así por ejemplo, en geometría pueden elegirse como términos indefinidos los de punto, recta y plano (Pasch también seleccionaba, en su primera edición, el de congruencia de segmentos de recta). La elección no es única. Como hay términos sin definición, surge la cuestión de qué propiedades de esos conceptos deben usarse para realizar demostraciones con ellos. La respuesta de Pasch es que los axiomas afirman algo acerca de esos términos indefinidos, y que son esos los únicos asertos que pueden utilizarse. Como Gergonne había dicho ya en 1818,³²² los conceptos indefinidos están implícitamente definidos por los axiomas.

En cuanto a éstos, continúa Pasch, aunque algunos pueden ser sugeridos por la experiencia, una vez que se ha seleccionado un conjunto de ellos, debe ser posible realizar todas las demostraciones

sin hacer más referencias a la experiencia o al significado físico de los conceptos. Además, los axiomas no son en absoluto verdades auto-evidentes, sino solamente supuestos destinados a proporcionar los teoremas de cualquier geometría particular. En sus *Vorlesungen* (2.^a ed., p. 90), afirma

... si la geometría ha de convertirse en una ciencia deductiva genuina, es esencial que la manera en que se realizan inferencias sea independiente tanto del significado de los conceptos geométricos como de los diagramas; todo lo que debe considerarse son las relaciones entre los conceptos geométricos aseguradas por las proposiciones y definiciones. Al llevar a cabo una deducción es tan juicioso como útil mantener presente el significado de los conceptos geométricos utilizados, pero no tiene por qué ser esencial; de hecho es precisamente cuando eso se hace necesario cuando se produce un salto en la deducción y (si no es posible colmar la deficiencia modificando el razonamiento) estamos obligados a admitir la inadecuación de las proposiciones invocadas como medio de demostración.

Pasch sí creía que los conceptos y axiomas deben basarse en la experiencia, pero esto era lógicamente irrelevante.

En sus *Vorlesungen* ofrecía axiomas para la geometría proyectiva, pero muchos de esos axiomas o sus análogos fueron igualmente importantes para la axiomatización de las geometrías euclídea y no euclídeas cuando se constituyeron como disciplinas independientes. Como ejemplo, fue el primero en ofrecer un conjunto de axiomas

para el orden en que se encuentran los puntos de una recta (que regulan el concepto «estar situado entre»). Tales axiomas deben incluirse también en un conjunto completo para cualquiera de las geometrías métricas. Trataremos más tarde sobre esos conceptos de orden.

Su método para construir la geometría proyectiva consistía en añadir punto, recta y plano del infinito a los puntos, rectas y planos propios. A continuación introducía coordenadas (sobre una base geométrica), utilizando la construcción de Von Staudt y Klein (vid. cap. 35, sec. 3), y finalmente la representación algebraica de las transformaciones proyectivas. Las geometrías euclídea y no euclídeas aparecían como casos especiales sobre una base geométrica distinguiendo los puntos y rectas propios e impropios al estilo de Félix Klein.

Un enfoque más satisfactorio de la geometría proyectiva fue el ofrecido por Peano,³²³ Este fue seguido por las obras de Mario Pieri (1860-1904), «*I Principii della geometria di posizione*»³²⁴, Federigo Enriques (1871-1946), *Lezioni di geometri proiettiva* (1898); Eliakim Hastings Moore (1862-1932);³²⁵ Friedrich H. Schur (1856-1932);³²⁶ Alfred North Whitehead (1861-1947), *The Axioms of Projective Geometry*,³²⁷ y Oswald Veblen (1880-1960) y John W. Young (1879-1932).³²⁸ Estos dos últimos ofrecieron un conjunto completamente independiente. Su clásico texto, *Projective Geometry* (2 vols., 1910 y 1918), desarrolla la organización kleiniana de la geometría, partiendo de la geometría proyectiva sobre una base estrictamente axiomática, para particularizar después esa geometría eligiendo

diferentes cuádricas absolutas (vid. cap. 38, sec. 3), obteniendo así las geometrías euclídea y no euclídeas. Sus axiomas son lo suficientemente generales como para incluir geometrías con sólo un número finito de puntos, geometrías con sólo puntos racionales, y geometrías con puntos complejos.

Convendría señalar un rasgo más de muchos de los sistemas axiomáticos para la geometría proyectiva y de los que a continuación veremos para la geometría euclídea: algunos de los axiomas de Euclides son axiomas de existencia (vid. cap. 4, sec. 3). Para garantizar la existencia lógica de ciertas figuras, los griegos utilizaban construcciones con regla y compás. Los trabajos de fundamentación del siglo XIX revisaron la noción de existencia, en parte para colmar deficiencias en el manejo que Euclides hacía de este tema, y en parte para ampliar la noción de existencia de manera que la geometría euclídea pudiera incluir puntos, rectas y ángulos no necesariamente construibles con regla y compás. Veremos qué nuevos tipos de axiomas de existencia aparecen en los sistemas que vamos ahora a examinar.

3. Los fundamentos de la geometría euclídea

En su *Sui fondamenti della geometría* (1894), Giuseppe Peano ofrecía un conjunto de axiomas para la geometría euclídea. También él recalca que los elementos básicos no deben definirse. Planteaba como principio que debe haber tan pocos conceptos indefinidos como sea posible, que para él eran punto, segmento y movimiento. La inclusión de este último puede parecer un poco sorprendente

vista la crítica del uso que hacía Euclides de la superposición; sin embargo, la objeción básica no se refiere al concepto de movimiento sino a la ausencia de una base axiomática adecuada para poder utilizarlo. Un conjunto similar de axiomas fue ofrecido por Pieri,³²⁹ discípulo de Peano, utilizando punto y movimiento como conceptos indefinidos. Otro conjunto, usando línea, segmento y congruencia de segmentos como elementos indefinidos, fue propuesto por Giuseppe Veronese (1854-1917) en sus *Fondamenti di geometria* (1891).

El sistema de axiomas para la geometría euclídea que parece más simple en sus conceptos y proposiciones, el que se mantiene más próximo a Euclides y ha conseguido mayor aceptación, es el debido a Hilbert, que no conocía el trabajo de los italianos. La primera versión apareció en sus *Grundlagen der Geometrie* (1899), pero la revisó varias veces. El sumario que viene a continuación está tomado de la séptima edición (1930) de ese libro. En su utilización de los conceptos indefinidos, cuyas propiedades quedan especificadas únicamente por los axiomas, Hilbert sigue a Pasch: no hay por qué asignar significado explícito alguno a los conceptos no definidos. Esos elementos, punto, recta, plano, etc. podrían reemplazarse, como señalaba Hilbert, por mesas, sillas, jarras de cerveza u otros objetos. Naturalmente, si la geometría trata con «cosas», los axiomas no son en absoluto verdades auto-evidentes, sino que deben ser considerados como arbitrarios aunque de hecho sean sugeridos por la experiencia.

Hilbert enumera primeramente sus conceptos indefinidos: punto,

recta, plano, pertenencia de un punto a una recta, pertenencia de un punto a un plano, situación de un punto entre otros dos, congruencia de pares de puntos, y congruencia de ángulos. El sistema de axiomas trata la geometría euclídea plana y sólida conjuntamente y los axiomas aparecen separados en grupos. El primero de ellos contiene los axiomas de existencia:

I. Axiomas de conexión o incidencia

I_1 . Para cada dos puntos A y B existe una recta a a la que pertenecen A y B .

I_2 . Para cada dos puntos A y B no hay más que una recta a la que pertenezcan A y B .

I_3 . En cada recta hay por lo menos dos puntos. Hay por lo menos tres puntos que no pertenecen a la misma recta

I_4 . Para cada tres puntos A , B y C que no están en la misma recta existe un plano a que contiene esos tres puntos. En cada plano hay (por lo menos) un punto.

I_5 . Para cada tres puntos A , B y C que no están en la misma recta no hay más que un plano que contenga los tres puntos

I_6 . Si dos puntos de una recta están en un plano a , entonces cada uno de los puntos de la recta está en a .

I_7 . Si dos planos α y β tienen un punto A en común, entonces tienen por lo menos otro punto B en común.

I_8 . Hay por lo menos cuatro puntos que no están en el mismo plano.

El segundo grupo de axiomas cubre la omisión más seria en el

conjunto de Euclides, es decir, la de axiomas acerca del orden en que se encuentran los puntos de una recta:

II. Axiomas de separación

II_1 . Si un punto B está entre los puntos A y C , entonces A , B y C son tres puntos diferentes de una recta y B también está entre C y A .

II_2 . Para dos puntos cualesquiera A y C hay por lo menos un punto B en la recta AC tal que C está entre A y B .

II_3 . Entre tres puntos cualesquiera de una recta, no hay más de uno que esté entre los otros dos.

Los axiomas II_2 y II_3 sirven para hacer la recta infinita.

Definición. Sean A y B dos puntos de una recta α . Se llama segmento AB al par de puntos A, B o B, A . Los puntos entre A y B se llaman puntos del segmento o puntos interiores al segmento. A y B se llaman extremos del segmento. Todos los demás puntos de la recta α se dice que están fuera del segmento.

II_3 . (Axioma de Pasch). Sean A, B y C tres puntos no alineados y sea α cualquier recta del plano determinado por A, B y C que no pase por ninguno de esos tres puntos. Si α pasa por algún punto del segmento AB , también debe pasar por algún punto del segmento AC o por algún punto del segmento BC .

III. Axiomas de congruencia

III_1 . Si A y B son dos puntos de una recta α y A' es un punto de α o de otra recta α' , entonces a un lado (previamente definido)

de A' sobre la recta α' se puede encontrar un punto B' tal que el segmento AB es congruente con $A'B'$, lo que se escribe $AB = A'B'$.

III₂. Si $A'B'$ y $A''B''$ son congruentes con AB , entonces $A'B' = A''B''$.

Este axioma limita la afirmación de Euclides «Cosas iguales a una misma cosa son iguales entre sí» a los segmentos de recta.

III₃. Sean AB y BC segmentos sin puntos interiores comunes de una recta α y sean $A'B'$ y $B'C'$ segmentos sin puntos interiores comunes de una recta α' . Si $AB = A'B'$ y $BC = B'C'$, entonces $AC = A'C'$.

Esto equivale a la afirmación de Euclides «Al añadir iguales a iguales se obtienen iguales», aplicada a segmentos de recta.

III₄. Sean $\angle (h,k)$ el ángulo formado por dos semirrectas h y k en un plano α y α' una recta de un plano α' en el que se supone fijado un cierto semiplano de los determinados por α' . Si h' es una semirrecta en α' que parte del punto O' , en α' hay una y sólo una semirrecta k' que parte de O' tal que $\angle (h,k)$ es congruente con $\angle (h',k')$ y todos los puntos interiores de $\angle (h',k')$ están en el semiplano fijado de α' . Cada ángulo es congruente consigo mismo.

III₅. Si para dos triángulos ABC y $A'B'C'$ tenemos $AB = A'B'$, $AC = A'C'$ y $\angle BAC = \angle B'A'C'$, entonces $\angle ABC = \angle A'B'C'$.

Este último axioma puede utilizarse para probar que $\angle ACB = \angle A'C'B'$. Se consideran los mismos dos triángulos y las mismas hipótesis. Tomando primero $AC = A'C'$ y después $AB = A'B'$,

puede concluirse que $\angle ACB = \angle A'C'B'$ sin más que aplicar el axioma a la nueva reordenación de los términos.

IV. El axioma de las paralelas

Sean α una recta y A un punto que no esté en α . En el plano determinado por α y A hay a lo más una recta que pasa por A sin cortar a α .

La existencia de al menos una recta que pasa por A sin cortar a α puede probarse por otros medios y por tanto no es preciso mencionarlo en el axioma.

V. Axiomas de continuidad

V_1 . (Axioma de Arquímedes). Dados dos segmentos cualesquiera AB y CD , en la recta determinada por A y B existen puntos A_1, A_2, \dots, A_n tales que los segmentos $AA_1, A_1A_2, A_2A_3, \dots, A_{n-1}A_n$ son congruentes con CD y B está entre A y A_n .

V_2 . (Axioma de completitud lineal). Los puntos de una recta forman un conjunto que satisface los axiomas I_1, I_2, II, III y V_1 y no puede ampliarse a un conjunto mayor de puntos que siga satisfaciendo esos axiomas.

Este axioma equivale a requerir suficientes puntos en la recta como para que puedan ponerse en correspondencia biunívoca con los números reales. Aunque este hecho había sido utilizado consciente e inconscientemente desde que se empezó a trabajar en geometría con coordenadas, hasta ese momento no se había establecido de su base lógica.

Con estos axiomas Hilbert demostró algunos de los teoremas básicos de la geometría euclídea. Otros autores completaron la tarea de mostrar que toda ella puede deducirse de los axiomas.

El carácter arbitrario de los axiomas de la geometría euclídea, esto es, su independencia con respecto a la realidad física, sacó a la palestra otro problema, el de la consistencia de esa geometría. Mientras se consideró a la geometría euclídea como la verdad acerca del espacio físico, cualquier duda sobre su consistencia parecía infundada. Pero la nueva comprensión de los conceptos no definidos y los axiomas exigía que se estableciera aquélla. El problema era aún más vital porque la consistencia de las geometrías no euclídeas se había reducido a la de la geometría euclídea (vid. cap. 38, sec. 4). Poincaré se había ocupado de este asunto en 1898,³³⁰ diciendo que podía creerse en la consistencia de una estructura axiomáticamente fundada, siempre que se pudiera dar de ella una interpretación aritmética. Hilbert mostró que la geometría euclídea es consistente proporcionando tal interpretación.

Para ello identifica (en el caso de la geometría plana) cada punto con un par ordenado (a,b) de números reales,³³¹ y cada recta con una razón $(u:v:w)$ en la que u y v no son simultáneamente 0. Un punto (a,b) está en una recta $(u:v:w)$ si

$$ua + vb + w = 0$$

La congruencia se interpreta algebraicamente mediante las expresiones de la geometría analítica para traslaciones y rotaciones;

así, dos figuras son congruentes si se puede transformar la una en la otra mediante traslación, simetría con respecto al eje x , y rotación.

Después de haber interpretado aritméticamente cada concepto haciendo ver que esa interpretación satisface los axiomas, Hilbert argumenta que los teoremas deben aplicarse también a la interpretación, puesto que son consecuencia lógica de los axiomas. Si hubiera una contradicción en la geometría euclídea, esta aparecería también en la formulación algebraica de la geometría, que es una extensión de la aritmética. Pero si la aritmética es consistente, también lo es entonces la geometría. En aquel momento todavía permanecía abierta la cuestión de la consistencia de la aritmética (vid. cap. 51).

Sería deseable probar que ninguno de los axiomas puede deducirse de algunos o todos los demás, porque en tal caso no habría por qué incluirlo como axioma. Esta noción de independencia fue planteada y discutida por Peano en su artículo de 1894 ya mencionado, e incluso antes en sus *Arithmetices Principia* (1889). Hilbert examinó la independencia de sus axiomas. Sin embargo, en su sistema no es posible mostrar que cada axioma es independiente de todos los demás, ya que el significado de algunos de ellos depende de los precedentes. Lo que sí consiguió probar es que los axiomas de cualquier grupo no pueden deducirse de los axiomas de los otros cuatro grupos. Su método consistió en proporcionar interpretaciones consistentes o modelos que satisfacen los axiomas de cada cuatro grupos sin satisfacer todos los axiomas del quinto

grupo.

Estas demostraciones de independencia tenían una relevancia especial en lo que atañe a las geometrías no euclídeas. Para establecer la independencia del axioma de las paralelas, Hilbert ofreció un modelo que satisface los otros cuatro grupos de axiomas pero no éste, utilizando los puntos interiores a una esfera euclídea y ciertas transformaciones que conservan la superficie de la esfera invariante. Así pues, el axioma de las paralelas no puede ser consecuencia de los otros cuatro grupos, porque si lo fuera, el modelo, como parte de la geometría euclídea, poseería propiedades contradictorias en cuanto al paralelismo. La misma demostración prueba que las geometrías no euclídeas son posibles, porque si el axioma euclídeo de las paralelas es independiente de los demás, la negación de ese axioma también debe ser independiente; ya que si fuera una consecuencia, el sistema entero de los axiomas euclídeos contendría una contradicción.

El sistema de axiomas de Hilbert para la geometría euclídea, publicado por primera vez en 1899, atrajo una considerable atención hacia los fundamentos de la geometría euclídea, y fueron muchos los autores que ofrecieron versiones alternativas utilizando diferentes conjuntos de elementos no definidos o ciertas variaciones en los axiomas. El mismo Hilbert, como ya hemos señalado, hizo varios cambios en su sistema antes de llegar a la versión de 1930. Entre los muchos sistemas alternativos mencionaremos sólo uno: Veblen ³³² ofreció un conjunto de axiomas basado en los conceptos no definidos de punto y orden. Mostró que cada uno de sus axiomas

es independiente de los demás, y también estableció otra propiedad, conocida con el nombre de categoricidad. Esta noción fue establecida claramente y empleada por primera vez por Edward V. Huntington (1874-1952), en un artículo dedicado al sistema de los números reales ³³³ (el nombre que él empleaba para esta noción era el de suficiencia). Un conjunto de axiomas P_1, P_2, \dots, P_n que conectan un conjunto de símbolos no definidos S_1, S_2, \dots, S_m se dice que es categórico si entre los elementos de dos colecciones cualesquiera, cada una de las cuales contiene símbolos no definidos y satisface los axiomas, se puede establecer una correspondencia biunívoca para los conceptos no definidos que preserve las relaciones establecidas por los axiomas; esto es, ambos sistemas son isomorfos. La categoricidad significa así que las diferentes interpretaciones del sistema de axiomas difieren únicamente en lenguaje. Esta propiedad no se cumpliría, por ejemplo, si se omitiera el axioma de las paralelas, ya que entonces la geometría euclídea y la hiperbólica serían interpretaciones no isomorfas del conjunto reducido de axiomas.

La categoricidad implica otra propiedad que Veblen llamaba disyuntiva y que se conoce ahora como completitud. Se dice que un sistema de axiomas es completo si es imposible añadir otro axioma que sea independiente del conjunto dado y consistente con él (sin introducir nuevos conceptos primitivos). La categoricidad implica la completitud, ya que si un conjunto de axiomas A fuese categórico y no completo se podría introducir un axioma S tal que tanto S como no- S fueran consistentes con el conjunto A . Y como el conjunto

original A es categórico, habría interpretaciones isomorfas de A con S y de A con no- S . Pero eso es imposible, porque deben cumplirse las proposiciones correspondientes de ambas interpretaciones, y 5 se aplica a una interpretación y no- S a la otra.

4. Otros trabajos de fundamentación

La nitidez alcanzada en los axiomas para la geometría euclídea sugirió investigaciones paralelas para las varias geometrías no euclídeas. Una de las características notables de los axiomas de Hilbert es que los axiomas para la geometría hiperbólica pueden obtenerse inmediatamente reemplazando el axioma euclídeo de las paralelas por el axioma de Lobachevski-Bolyai, permaneciendo inalterados los demás.

Para obtener axiomas para la geometría simple o doblemente elíptica, no basta abandonar el axioma euclídeo de las paralelas para sustituirlo por un axioma que haga que dos rectas cualesquiera tengan un punto en común (elíptica simple) o al menos un punto en común (elíptica doble), y hay que cambiar también otros axiomas. La recta de esas geometrías no es infinita, sino que tiene las propiedades de una circunferencia. Así pues, hay que reemplazar los axiomas de orden de la geometría euclídea por otros que describan las relaciones de orden entre los puntos de una circunferencia. Se han propuesto varios sistemas de axiomas de este tipo. George B. Halsted (1853-1922), en su *Rational Geometry*,³³⁴ y John R. Kline (1891-1955)³³⁵ ofrecieron bases axiomáticas para la geometría doblemente elíptica, y Gerhard

Hessenberg (1874-1925)³³⁶ propuso un sistema de axiomas para la geometría elíptica simple.

Otro tipo de investigaciones sobre la fundamentación de la geometría son las referidas a las consecuencias de la negación u omisión de uno o más axiomas de un determinado conjunto. El mismo Hilbert ya había hecho esto en sus demostraciones de independencia, ya que la esencia de tales demostraciones consiste en la construcción de un modelo o interpretación que satisfaga todos los axiomas excepto aquél cuya independencia se pretende establecer. El ejemplo más significativo de axioma negado es, naturalmente, el de las paralelas. Pero también se han obtenido resultados interesantes eliminando el axioma de Arquímedes, que en el sistema de Hilbert aparece como $V\setminus$. La geometría resultante se llama no-arquimediana; en ella hay pares de segmentos tales que al multiplicar uno de ellos por cualquier número entero, por grande que sea, no se consigue un segmento mayor que el otro. En sus *Fondamenti di geometría*, Giuseppe Veronese construyó una geometría no arquimediana, mostrando también que los teoremas de esa geometría se aproximan tanto como se quiera a los de la geometría euclídea.

Max Dehn (1878-1952) también obtuvo ³³⁷ muchos teoremas interesantes omitiendo el axioma de Arquímedes. Por ejemplo, hay una geometría en la que la suma de los ángulos de un triángulo es dos rectos, existen triángulos semejantes no congruentes y se pueden trazar infinitas paralelas a una recta dada pasando por un punto también dado.

Hilbert había señalado que para la construcción de la teoría de las áreas en el plano no se necesitaba el axioma de continuidad V_2 . Para el espacio, sin embargo, Max Oehn probó³³⁸ la existencia de poliedros que tienen el mismo volumen y no se pueden descomponer en partes mutuamente congruentes (incluso añadiéndoles poliedros congruentes). Así pues, en tres dimensiones el axioma de continuidad es necesario.

La fundamentación de la geometría euclídea fue enfocada de manera completamente diferente por algunos matemáticos. La geometría, como sabemos, había caído en desgracia cuando los matemáticos se dieron cuenta de que habían aceptado ciertos hechos inconscientemente, basándose en la intuición, y sus supuestas demostraciones eran consiguientemente incompletas. El peligro de que esto pudiera repetirse una y otra vez les hizo creer que la única base sólida para la geometría sería la aritmética. La forma en que podían erigir esa base estaba clara. De hecho, Hilbert había ofrecido una interpretación aritmética de la geometría euclídea. Lo que había que hacer ahora, para la geometría plana por ejemplo, era no interpretar un punto como un par de números (x,y) , sino *definir* como punto el par de números, como recta la razón entre tres números $(u:v:w)$, definir la pertenencia del punto (x,y) a la recta $(u:v:w)$ por la satisfacción de la ecuación $ux + vy + w = 0$, definir como circunferencia el conjunto de los (x,y) que satisfacen una ecuación del tipo $(x - a)^2 + (y - b)^2 = r^2$, y así sucesivamente. Con otras palabras, habría que utilizar los equivalentes analíticos de las nociones puramente geométricas como definiciones de los conceptos

geométricos, *y* métodos algebraicos para probar los teoremas. Como la geometría analítica contiene en forma algebraica la contrapartida exacta de cuanto existe en geometría euclídea, no había problema en cuanto a la obtención de la fundamentación aritmética. De hecho, el trabajo técnico necesario estaba ya hecho, incluso para la geometría euclídea n -dimensional, por ejemplo en la *Ausdehnungslehre* de Grassmann; y el mismo Grassmann había propuesto que su obra sirviera como fundamentación para la geometría euclídea.

5. Algunas cuestiones abiertas

Las investigaciones críticas sobre la geometría se extendieron más allá de la reconstrucción de los fundamentos. Hasta entonces se habían utilizado libremente las curvas. Las más simples, como la elipse, tenían definiciones geométricas y analíticas seguras. Pero muchas curvas se habían introducido únicamente por medio de ecuaciones y funciones. La rigorización del análisis había supuesto no sólo la ampliación del concepto de función, sino también la construcción de funciones muy peculiares, como las funciones continuas sin derivada en ningún punto. Que estas funciones atípicas fueran turbadoras desde el punto de vista geométrico es fácil de entender. La curva que representa el ejemplo de Weierstrass de una función que es continua en todo punto pero no es diferenciable en ninguno no se ajusta evidentemente al concepto usual, porque la falta de derivada significa que esa curva no puede tener tangente en ningún punto. La cuestión que surgió entonces es

si las representaciones geométricas de tales funciones son curvas y, más en general, la de qué debe entenderse por curva.

Jordán dio como definición de curva³³⁹ el conjunto de puntos representados por las funciones continuas $x = f(t)$, $y = g(t)$, para $t_0 \leq t \leq t_1$. Para algunos propósitos Jordán quería restringir sus curvas de manera que no poseyeran puntos múltiples, y requirió entonces que $f(t) \neq f(t')$ o $g(t) \neq g(t')$ para t y t' entre t_0 y t_1 es decir, que para cada (x,y) de la curva hubiese un solo t . Tales curvas reciben el nombre de curvas de Jordán.

En este mismo tratado presentó la noción de curva cerrada,³⁴⁰ requiriendo únicamente que $f(t_0) = f(t_1)$ y $g(t_0) = g(t_1)$, y estableció el teorema según el que una curva cerrada divide al plano en dos partes, una interior y otra exterior. Dos puntos de la misma región pueden unirse mediante un camino poligonal que no corta a la curva, y dos puntos situados en diferentes regiones no pueden unirse mediante ninguna línea poligonal o curva continua que no corte a la curva cerrada simple. El teorema es más potente de lo que parece a primera vista, ya que una curva cerrada simple puede tener un aspecto muy intrincado. De hecho, como las funciones $f(t)$ y $g(t)$ sólo tienen que ser continuas, resulta implicada toda la variedad de posibles complicaciones que pueden aparecer en las funciones continuas. El mismo Jordán y muchos matemáticos distinguidos propusieron demostraciones incorrectas del teorema. La primera demostración rigurosa se debe a Veblen³⁴¹.

No obstante, la definición de Jordán de curva, aun siendo satisfactoria para muchas aplicaciones, era demasiado amplia. En

1890 Peano descubrió³⁴² que una curva satisfaciendo la definición de Jordán puede pasar por todos los puntos de un cuadrado al menos una vez. Ofreció una descripción aritmética detallada de la correspondencia entre los puntos del intervalo $[0,1]$ y los puntos del cuadrado, especificando dos funciones $x = f(t)$, e $y = g(t)$, unívocas y continuas para $0 \leq t \leq 1$ y tales que x e y toman los valores correspondientes a cada punto del cuadrado unidad. Sin embargo, la correspondencia de (x,y) a t no es unívoca, ni es continua. Una correspondencia biunívoca continua de los valores de t con los valores de (x,y) es imposible; o sea, que $f(t)$ y $g(t)$ no pueden ser entonces simultáneamente continuas. Quien probó esto fue Eugen E. Netto (1846-1919)³⁴³.

La interpretación geométrica de la curva de Peano fue ofrecida por Arthur M. Schoenflies (1853-1928)³⁴⁴ y E. H. Moore³⁴⁵.

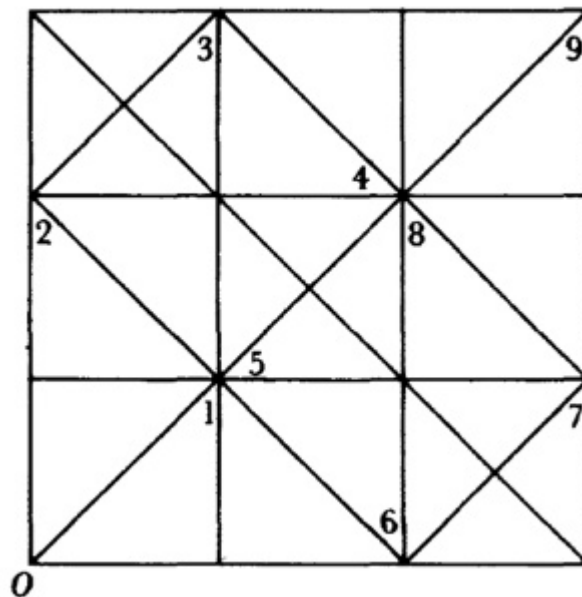


Figura 42.3

Primeramente se transforma el segmento $[0,1]$ en los nueve segmentos presentados en la figura 42.3, y a continuación en cada subcuadrado se descompone la diagonal en él contenida siguiendo el mismo método, pero haciendo que la transición desde cada subcuadrado al siguiente sea continua, y se repite el proceso *ad infinitum*, consiguiendo en el límite cubrir todo el cuadrado original. Ernesto Cesáro (1859-1906)³⁴⁶ presentó la forma analítica de las f y g de Peano.

Hilbert ofreció otro ejemplo³⁴⁷ de aplicación continua del segmento unidad sobre el cuadrado. Divídanse el segmento unidad y el cuadrado en cuatro partes iguales, tal como aparecen en la figura 42.4.

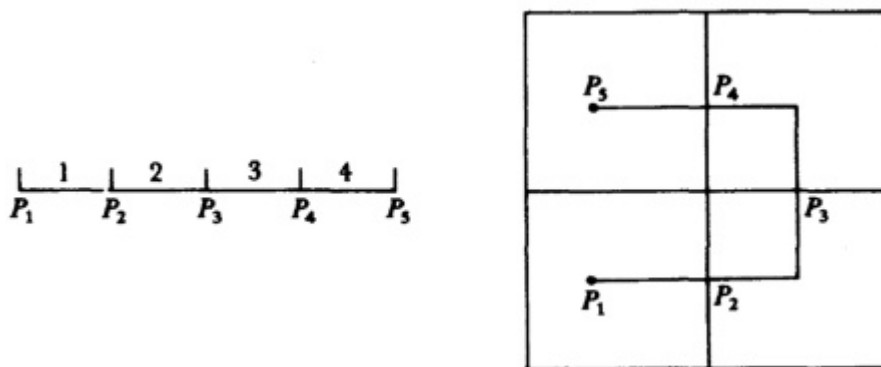


Figura 42.4

Atravesemos ahora cada subcuadrado de manera que el camino mostrado en la figura corresponda al segmento unidad. Dividamos a continuación el cuadrado unidad en 16 subcuadrados, numerándolos en la forma que muestra la figura 42.5, y unamos los centros de esos 16 subcuadrados en la forma también indicada.

El proceso continúa dividiendo cada subcuadrado en cuatro partes, y numerando éstas de manera que podamos recorrer todo el conjunto siguiendo un camino continuo. La curva deseada es el límite de las sucesivas poligonales obtenidas en cada etapa. Como los subcuadrados y las partes del segmento unidad se contraen simultáneamente hacia puntos al proseguir la subdivisión, puede verse intuitivamente que cada punto del segmento unidad se transforma así en un solo punto del cuadrado. De hecho, si fijamos un punto del segmento unidad, digamos $t = 2/3$, entonces la imagen de ese punto es el límite de las sucesivas imágenes de $t = 2/3$ que aparecen en las diferentes poligonales.

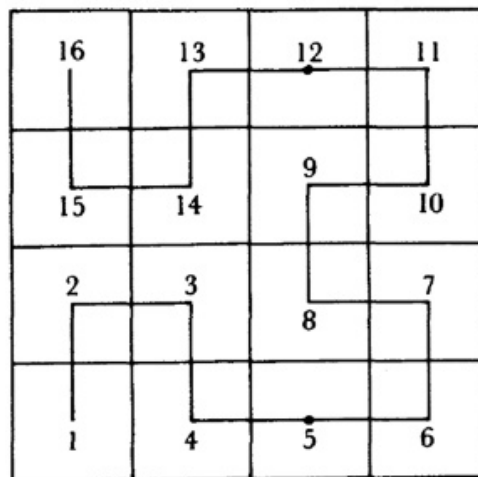


Figura 42.5

Estos ejemplos muestran que la definición de curva sugerida por Jordán no es satisfactoria, ya que una curva, según esa definición, puede llenar un cuadrado. La cuestión de qué debe entenderse por curva permanecía por tanto abierta. Félix Klein señaló en 1898 que

no había nada más oscuro que la noción de curva³⁴⁸. Y fueron los topólogos quienes afrontaron el problema (vid. cap. 50, sec. 2).

Más allá del problema de qué debe entenderse por curva, la extensión del análisis a las funciones sin derivada planteó también la cuestión de la longitud que se supone debe tener una curva. La acostumbrada fórmula del cálculo es

$$L = \int_a^b (1 + y'^2)^{1/2} dx$$

donde $y = f(x)$ exige, desde su mismo planteamiento, la existencia de derivada, y no puede aplicarse a las funciones no diferenciables. Du Bois Reymond, Peano, Ludwig Scheefer (1859-1885) y el mismo Jordán realizaron varios esfuerzos por generalizar el concepto de longitud de una curva, utilizando definiciones de integral más generales o conceptos geométricos. La definición más general se alcanzó en relación con la noción de medida, que examinaremos en el capítulo 44.

Una dificultad similar fue señalada para el concepto de área de una superficie. La idea más extendida en los textos del siglo XIX consistía en inscribir en la superficie un poliedro con caras triangulares, tomando como área de la superficie el límite de las sumas de las áreas de esos triángulos cuando sus lados tienden a 0. Analíticamente, si la superficie viene representada por

$$x = \phi(u, v) \quad y = \psi(u, v) \quad z = \chi(u, v)$$

la fórmula para el área de la superficie se convierte en

$$\int_D \sqrt{A^2 + B^2 + C^2} \, du \, dv$$

donde A , B y C son los jacobianos de y y z , x y z , y , x e y , respectivamente. De nuevo se planteaba la cuestión de qué hacer cuando x , y o z no poseen derivadas. Para complicar aún más la situación, H. A. Schwarz propuso en una carta a Hermite³⁴⁹ un ejemplo en el que una elección determinada de los triángulos del poliedro inscrito conducía a un área infinita incluso para un cilindro ordinario³⁵⁰. La teoría del área de una superficie tuvo que ser también reconsiderada en relación con la noción de medida.

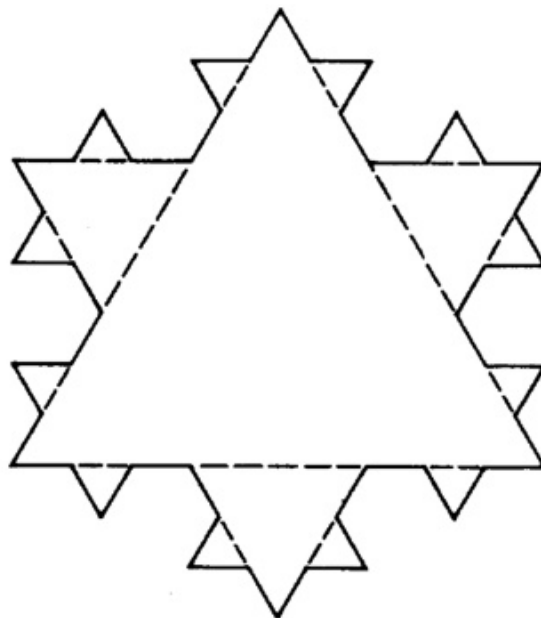


Figura 42.6

Hacia 1900 nadie había probado que una curva cerrada plana, tal como la habían definido Jordán y Peano, encerrara un área. Helge von Koch (1870-1924) complicó aún más el problema del área ofreciendo un ejemplo de una curva continua aunque no diferenciable, con perímetro infinito, que limitaba un área finita.³⁵¹ Se parte de un triángulo equilátero ABC (fig. 42.6) de lado $3s$. Sobre el tercio central de cada lado se construye un triángulo equilátero de lado $s/9$ y se elimina la base de esos nuevos triángulos, que serán tres. Sobre cada lado de longitud s de la nueva figura se construye, hacia el exterior, tomando como base el tercio central, un triángulo equilátero de lado $s/3$, y se eliminan las bases de todos esos nuevos triángulos, que serán 12. Sobre el tercio central de cada lado de la figura resultante se construye ahora un triángulo equilátero de lado $s/9$, de los que habrá 48. Los perímetros de los sucesivos polígonos son $9s$, $12s$, $16s$,... y tienden por tanto a infinito. Sin embargo, el área de la figura limitada por esos polígonos se mantiene acotada. En efecto, por la bien conocida fórmula del área de un triángulo equilátero en función de su lado, es decir $(b^2/4)\sqrt{3}$ si la longitud del lado es b , el área del triángulo original es $[(3s)^2/4]\sqrt{3}$. El área de los tres primeros triángulos añadidos es $3 \cdot (s^2/4)\sqrt{3}$. Como la longitud de los lados de los triángulos añadidos en la siguiente etapa es $s/3$ y hay 12, el área total añadida es $12 \cdot (s/3)^2 \sqrt{3}/4 = (s^2/3)\sqrt{3}$. La suma de las áreas es

$$S = \frac{9s^2}{4}\sqrt{3} + \frac{3s^2}{4}\sqrt{3} + \frac{s^2}{3}\sqrt{3} + \frac{4s^2}{27}\sqrt{3} + \dots$$

Se trata de una serie geométrica (excluyendo el primer sumando) de razón $4/9$, cuya suma es

$$S = \frac{9s^2}{4}\sqrt{3} + \frac{(3/4)s^2}{1 - 4/9}\sqrt{3} = \frac{18}{5}s^2\sqrt{3}$$

Las curvas de Peano y Hilbert también plantearon la cuestión de qué debe entenderse por dimensión. El cuadrado es por sí mismo bidimensional, pero como imagen continua de una curva, tendría que ser unidimensional. Además, Cantor había mostrado que los puntos de un segmento de recta pueden ponerse en correspondencia biunívoca con los puntos de un cuadrado (cap. 41, sec. 7). Aunque esa correspondencia no es continua en ninguno de los dos sentidos, muestra que la dimensión no es una cuestión de cantidad de puntos. Tampoco es cuestión del número de coordenadas necesarias para fijar la posición de un punto, como habían pensado Riemann y Helmholtz, ya que la curva de Peano asigna un único (x,y) a cada valor de i .

A la luz de estas dificultades puede verse que la rigorización de la geometría no bastaba para responder a todos los problemas que se habían planteado. Muchos de ellos fueron resueltos por los topólogos y analistas del presente siglo. El mismo hecho de que continuaran surgiendo problemas acerca de los conceptos fundamentales muestra una vez más que las matemáticas no crecen

como una estructura lógica. Los avances en nuevos campos, e incluso el perfeccionamiento de otros más antiguos, revelan nuevos e insospechados defectos. Más allá de la resolución de los problemas que conciernen a curvas y superficies tenemos todavía que ver si se alcanzó la última etapa en cuanto al rigor con los trabajos de fundamentación en análisis, el sistema de los números reales y la geometría básica.

Bibliografía

- Becker, Oskar: *Grundlagen der Mathematik in geschichtlicher Entwicklung*, Karl Alber, 1954, 199-212.
- Enriques, Federigo: «Prinzipien der Geometrie». *Encyk. der Math. Wiss.*, B. G. Teubner, 1907-1910, III AB1, 1-129.
- Hilbert, David: *Grundlagen der Geometrie*, 7.^a ed., B. G. Teubner, 1930. Hay versión castellana, D. Hilbert, *Fundamentos de la Geometría*, Madrid, C.S.I.C., 1991.
- Pasch, M. y M. Dehn: *Vorlesungen über neuere Geometrie*, 2.^a ed., Julius Springer, 1926, 185-271.
- Peano, Giuseppe: *Opere scelte*, 3 vols. Edizioni Cremonese, 1957-1959.
- Reichardt, Hans: *C. F. Gauss, Leben und Werke*, Haude und Spenersche, 1960, 111-150.
- Schmidt, Arnold: «Zu Hilberts Grundlegung der Geometrie», en las *Gesammelte Abhandlungen* de Hilbert, 2, 404-414.
- Singh, A. N.: «The Theory and Construction of Non-Differentiable Functions», en E. W. Hobson: *Squaring the*

Girele and Other Monographs, Chelsea (reimpresión), 1953.

Capítulo 43

La matemática en torno a 1900

No he vacilado en llamar al siglo XIX, en el Congreso de Matemáticos de París en 1900, el siglo de la teoría de funciones.

Vito Volterra

Contenido:

- 1. Las principales características de los desarrollos del siglo XIX*
- 2. El movimiento axiomático*
- 3. La matemática como creación del hombre*
- 4. La pérdida de la verdad*
- 5. La matemática como el estudio de estructuras arbitrarias*
- 6. El problema de la consistencia*
- 7. Una mirada hacia el futuro*

Bibliografía

1. Las principales características de los desarrollos del siglo XIX

Tanto en el XIX como en los dos siglos precedentes los avances en matemáticas trajeron consigo profundos cambios, apenas perceptibles en el desarrollo de un año a otro, pero trascendentes en sí mismos y en su efecto sobre futuras investigaciones. La vasta expansión de los temas y la incorporación de nuevos campos, tanto como la extensión de otros más antiguos, son evidentes. El álgebra

recibió un impulso totalmente nuevo con Galois; la geometría recobró vitalidad y se vio profundamente alterada por la introducción de las geometrías no euclídeas y el resurgimiento de la geometría proyectiva; la teoría de números se transformó en teoría analítica de números; y el análisis se amplió extraordinariamente con la introducción de la teoría de funciones complejas y la expansión de las ecuaciones diferenciales ordinarias y en derivadas parciales. Desde el punto de vista del desarrollo técnico, la teoría de funciones de variable compleja fue la más significativa de las nuevas creaciones. Pero desde el punto de vista de la importancia intelectual y de sus efectos más profundos sobre la naturaleza de la matemática, el desarrollo más renovador fue la geometría no euclídea. Como veremos, sus efectos fueron mucho más revolucionarios de lo que hasta ahora hemos indicado. El círculo en el que los estudios matemáticos parecían estar encerrados a comienzos del siglo se rompió por muchos puntos, y las matemáticas proliferaron en un centenar de ramas. El torrente de nuevos resultados contradujo agudamente la opinión dominante a finales del siglo XVIII de que el caudal de las matemáticas había quedado exhausto.

La actividad matemática también se expandió en otros aspectos durante el siglo XIX. El número de matemáticos se incrementó enormemente como consecuencia de la democratización de la enseñanza. Aunque Alemania, Francia e Inglaterra seguían siendo los centros más importantes, Italia reapareció en la arena, y los Estados Unidos, con Benjamín Peirce, G. W. Hill y Josiah Williard

Gibbs entraron en ella por primera vez. En 1863 se fundó en los USA la Academia Nacional de Ciencias. Sin embargo, a diferencia de la Royal Society de Londres, o las Academias de Ciencias de París o Berlín, la Academia Nacional americana no ha sido un lugar de encuentros científicos en los que se presentaran y revisaran los nuevos resultados, aunque sí publica una revista, los *Proceedings of the Academy*. Se organizaron otras sociedades matemáticas (vid. cap. 26, sec. 6), en las que los investigadores podían reunirse y presentar sus artículos, y que patrocinaban diferentes revistas. A fines de siglo el número de éstas dedicadas parcial o enteramente a la investigación matemática había aumentado hasta unas 950. En 1897 comenzó la costumbre de mantener un congreso internacional cada cuatro años.

Junto a la explosión de la actividad matemática se produjo una novedad menos agradable. Las diferentes disciplinas se hicieron autónomas, dando lugar cada una de ellas a su propia terminología y metodología. La búsqueda de cualquier resultado impuso el estudio de problemas más especializados y más difíciles, que requerían más ideas, y más ingeniosas, una fértil inspiración, y demostraciones más intrincadas. Para realizar cualquier avance, los matemáticos se vieron obligados a adquirir una sólida fundamentación teórica y gran habilidad técnica. La especialización se hizo notable en los trabajos de Abel, Jacobi, Galois, Poncelet y muchos otros. Aunque se puso algún acento en la interrelación entre diferentes ramas con nociones como las de grupo, transformación lineal o invariancia, el efecto general fue el de una

separación en numerosos campos distintos y alejados entre sí. A Félix Klein le pareció en 1893 que se podría superar la especialización y divergencia entre las diferentes ramas mediante los conceptos que acabamos de mencionar, pero su esperanza fue vana. Cauchy y Gauss fueron los últimos en conocer la matemática como un todo, aunque Poincaré y Hilbert fueron también matemáticos casi universales.

Desde el siglo XIX los matemáticos trabajan normalmente sólo en algún área de la matemática, subrayando naturalmente cada uno la importancia de la propia sobre las restantes. Sus publicaciones ya no son para un público amplio, sino para un círculo restringido de colegas. La mayoría de los artículos ya no contienen ninguna indicación de su conexión con otros problemas de la matemática, son difícilmente accesibles para muchos matemáticos, e incomprensibles para los no iniciados.

Más allá de sus logros en temas concretos, el siglo XIX se caracteriza por la reintroducción del rigor en las demostraciones. Dejando de lado lo que los matemáticos puedan haber pensado individualmente acerca de la solidez de sus resultados, el hecho es que desde el año 200 a. de C. hasta 1870 aproximadamente, casi toda la matemática se sostuvo sobre una base empírica y pragmática, perdiéndose de vista el concepto de demostración deductiva a partir de axiomas explícitos. Una de las revelaciones más sorprendentes de la historia de las matemáticas es precisamente esa relajación en el rigor deductivo durante los dos milenios en los que su contenido se extendió tan ampliamente.

Aunque cabe reconocer algunos esfuerzos más tempranos por rigORIZAR el análisis, en particular los de Lagrange (cap. 19, sec. 7), la actitud dominante era la representada por Lacroix (cap. 26, sec. 3). Los trabajos de Fourier pueden poner los pelos de punta a un analista de nuestros días; en cuanto a Poisson, la derivada y la integral no eran para él sino abreviaturas para un cociente de diferencias o una suma finita. El movimiento iniciado por Bolzano y Cauchy para asentar los fundamentos surgió indudablemente de la preocupación por la cantidad rápidamente creciente de matemáticas que descansaban sobre los frágiles fundamentos del cálculo, y se vio acelerado por el descubrimiento de Hamilton de los cuaterniones no conmutativos, que desafiaba los principios acríticamente aceptados acerca de los números. Pero más turbadora aún fue la creación de las geometrías no euclídeas, que no sólo destruyó la mismísima noción de la auto-evidencia de los axiomas y su aceptación demasiado superficial, sino que reveló insuficiencias en demostraciones que habían sido consideradas como las más sólidas de toda la matemática.

Los matemáticos se dieron cuenta de que habían pecado de credulidad confiando en la intuición.

Hacia 1900 parecía alcanzado el objetivo de proporcionar una fundamentación rigurosa a la matemática, y los matemáticos casi se envanecían de semejante logro. En su intervención ante el Segundo Congreso Internacional en París,³⁵² Poincaré exultaba: «¿Hemos alcanzado por fin el rigor absoluto? En cada etapa de su evolución nuestros predecesores también creyeron haberlo alcanzado. Y si

ellos se equivocaron, ¿no estaremos equivocados nosotros también?... En el análisis actual, no obstante, si tenemos el cuidado de ser rigurosos, sólo hay silogismos o invocaciones a la intuición de número puro, que no pueden engañarnos. Ahora sí se puede decir que se ha alcanzado el rigor absoluto.» Cuando se consideran los resultados clave en la fundamentación del sistema numérico y la geometría, y la erección del análisis sobre la base de ese sistema numérico, se entienden las razones para el envanecimiento. Las matemáticas contaban ahora con unos fundamentos que prácticamente cualquiera estaría dispuesto a aceptar.

La formulación precisa de conceptos básicos como los de número irracional, continuidad, integral, y derivada no fue sin embargo recibida entusiásticamente por todos los matemáticos. Muchos de ellos no entendían el nuevo lenguaje $\varepsilon - \delta$, y consideraban esas definiciones precisas como caprichosas e innecesarias para la comprensión de las matemáticas, e incluso para el rigor en las demostraciones, sintiendo que bastaba con la intuición, a pesar de las sorpresas de las funciones continuas sin derivada, las curvas que llenan un espacio, o las curvas sin longitud. Emile Picard, por ejemplo, decía a propósito del rigor en las ecuaciones en derivadas parciales: «... el verdadero rigor es productivo, y se distingue en eso de aquel otro rigor que es puramente forma y enojoso, y que arroja una sombra sobre cada uno de los problemas que toca.»³⁵³

A pesar de que la geometría también se ha visto sometida a la rigorización, una de las consecuencias de este movimiento fue que números y análisis se elevaran por encima de ella. El

reconocimiento por los matemáticos, durante y después de la creación de las geometrías no euclídeas, de que se habían dejado llevar inconscientemente por la intuición al aceptar las demostraciones de la geometría euclídea, les hizo temer que pudiera seguir siendo así en cualquier razonamiento geométrico, y preferir una matemática construida sobre los números. Muchos estuvieron a favor de ir más allá y construir toda la geometría sobre los números, lo que podía hacerse por medio de la geometría analítica en la forma descrita anteriormente. Así, la mayoría de los matemáticos hablaba de la aritmetización de las matemáticas, aunque habría sido más preciso hablar de la aritmetización del análisis. Donde Platón pudo decir «Dios geometriza eternamente», Jacobi, ya a mediados de siglo, dijo: «Dios aritmetiza permanentemente». En el Segundo Congreso Internacional, Poincaré aseguraba: «Actualmente sólo quedan en análisis los enteros y sistemas finitos e infinitos de enteros, interrelacionados por una red de relaciones de igualdad o desigualdad. La matemática, como decimos, ha sido aritmetizada.» Si Pascal había dicho: «*Tout ce qui passe la Géométrie nous passe*»,³⁵⁴ los matemáticos de 1900 preferían decir: «*Tout ce qui passe l'Arithmétique nous passe.*»

La construcción de fundamentos lógicos para las matemáticas, dejando a un lado la preferencia por una base aritmética o geométrica, completó otra etapa en la ruptura con la metafísica. La vaguedad en los fundamentos y justificaciones de las inferencias matemáticas se había ocultado en el siglo XVIII y a comienzos del XIX con alusiones a argumentos metafísicos que, aunque nunca se

explicitaban, eran mencionados como respaldo para aceptar las matemáticas. La axiomatización de los números reales y la geometría proporcionó a la matemática una base clara, independiente y autosuficiente. Ya no era preciso recurrir a la metafísica. Como dijo Lord Kelvin, «la matemática es la única metafísica aceptable».

La rigORIZACIÓN de la matemática pudo cubrir una necesidad del siglo XIX, pero también nos enseña algo acerca de su desarrollo. La estructura lógica recién implantada se suponía que garantizaba la solidez de las matemáticas, pero esa garantía tenía algo de impostura: ni un solo teorema de aritmética, álgebra o geometría euclídea fue refutado por el nuevo rigor, y los teoremas de análisis sólo tuvieron que ser formulados más cuidadosamente. De hecho, todo lo que hicieron las nuevas estructuras axiomáticas fue confirmar lo que los matemáticos sabían ya; los axiomas permitían inferir los teoremas existentes, pero no los determinaban. Lo que significa que la matemática descansa, no en la lógica, sino en intuiciones profundas. El rigor, como señalaba Jacques Hadamard, sólo sanciona las conquistas de la intuición; o como decía Hermann Weyl, la lógica es la higiene que practica el matemático para mantener sus ideas fuertes y saludables.

2. El movimiento axiomático

La rigORIZACIÓN de la matemática se consiguió axiomatizando sus diferentes ramas. La esencia de un desarrollo axiomático, de acuerdo con el modelo que hemos descrito en los capítulos 41 y 42,

consiste en partir de unos términos no definidos cuyas propiedades quedan especificadas por los axiomas; el objetivo de todo el trabajo es obtener los teoremas como consecuencia de los axiomas. Además, hay que establecer para cada sistema la independencia, consistencia y categoricidad de los axiomas, nociones que ya hemos examinado en los dos capítulos precedentes.

A comienzos del siglo XX el método axiomático no sólo permitió establecer los fundamentos lógicos de muchas ramas, viejas y nuevas, de la matemática, sino que reveló precisamente qué suposiciones subyacen bajo cada una de ellas e hizo posible la comparación y clarificación de las relaciones entre esas diferentes ramas. Hilbert estaba entusiasmado con la validez de este método. Analizando el estado presumiblemente perfecto que la matemática había alcanzado fundamentando cada una de sus ramas sobre bases axiomáticas sólidas, señalaba:³⁵⁵

De hecho, el método axiomático ha sido y sigue siendo la ayuda más conveniente e indispensable para cualquier investigación exacta en no importa qué dominio; es lógicamente inobjetable y al mismo tiempo fructífero, y garantiza una completa libertad de investigación. Proceder axiomáticamente significa pues pensar con conocimiento de lo que uno se trae entre manos. Mientras que antes, sin el método axiomático, se actuaba ingenuamente creyendo en ciertas relaciones como en un dogma, el enfoque axiomático aparta esa ingenuidad sin privarnos por eso de las ventajas de la creencia.

También en el último apartado de su «Axiomatisches Denken»,³⁵⁶ defendía este método:

Todo lo que puede ser objeto del pensamiento matemático, en cuanto esté madura la construcción de una teoría, cae bajo el método axiomático y por tanto directamente dentro de la matemática. Ahondando en capas cada vez más profundas de los axiomas... podemos obtener percepciones más profundas del pensamiento científico y constatar la unidad de nuestro conocimiento. Gracias especialmente al método axiomático, la matemática parece llamada a desempeñar un papel conductor en todo conocimiento.

La posibilidad de explorar nuevos problemas omitiendo, negando o variando de alguna otra manera los axiomas de sistemas establecidos atrajo a muchos matemáticos. Esta actividad, y la construcción de bases axiomáticas para las diferentes ramas de la matemática se conocen como movimiento axiomático. Sigue siendo todavía una actividad muy practicada. Parte de su atractivo se explica por el hecho de que, una vez establecidas bases axiomáticas sólidas para las ramas más importantes, variaciones del tipo que acabamos de describir son relativamente fáciles de introducir y explorar. Sin embargo, cualquier nuevo desarrollo en la matemática ha atraído siempre a un cierto número de personas que buscan campos abiertos a la exploración o están sinceramente convencidos de que el futuro de la matemática reside en esa área particular.

3. La matemática como creación del hombre

Desde el punto de vista del futuro desarrollo de las matemáticas, el resultado más significativo del siglo fue la obtención de un enfoque apropiado de las relaciones de las matemáticas con la naturaleza. Aunque no hemos entrado en las opiniones sobre las matemáticas de muchos de los autores cuyos trabajos hemos descrito, sí hemos dicho que los griegos, Descartes, Newton, Euler y muchos otros creían que las matemáticas proporcionan la descripción precisa de los fenómenos reales y que consideraban sus trabajos como desvelamientos de la estructura matemática del universo. Los matemáticos trataban con abstracciones, pero éstas no eran más que las formas ideales de los objetos físicos o de los sucesos. Incluso conceptos tales como funciones o derivadas estaban implicados en los fenómenos reales y servían para describirlos.

Aparte de lo que ya hemos señalado como soporte de esta opinión acerca de las matemáticas, las afirmaciones de los matemáticos sobre el número de dimensiones que cabe considerar en geometría muestran claramente lo ligada que se mantenía la matemática a la realidad. Así, en el primer libro de *El Cielo*, Aristóteles dice: «La línea tiene una magnitud, el plano dos, y el sólido tres, y más allá de esas tres no hay ninguna otra magnitud, porque ellas lo llenan todo... No se puede pasar a otra cosa, como se pasa de la longitud al área y del área al volumen.» Y en otro párrafo: «... No hay magnitud que trascienda al tres, porque no hay más que tres dimensiones», y añade: «tres es el número perfecto». En su *Algebra*, John Wallis consideraba a un espacio con más de tres dimensiones como «un

monstruo de la naturaleza, menos posible que una quimera o un centauro». Y sigue: «Longitud, Anchura y Espesor llenan todo el espacio. Tampoco Fansie puede imaginar cómo podría hacer una Cuarta Dimensión Local más allá de esas Tres.» Cardano, Descartes, Pascal y Leibniz también consideraron la posibilidad de una cuarta dimensión, rechazándola como absurda. Mientras el álgebra se mantuvo ligada a la geometría también se rechazó el producto de más de tres cantidades. Jacques Ozanam señalaba que el producto de más de tres letras representaría una magnitud «con tantas dimensiones como letras haya, pero sería sólo imaginaria, porque en la naturaleza no conocemos ninguna cantidad que tenga más de tres dimensiones».

La idea de una geometría matemática de más de tres dimensiones se rechazó incluso a comienzos del siglo XIX. Möbius, en su *Der barycentrische Calcul* (1827), señalaba que las figuras geométricas que no podían superponerse en tres dimensiones al ser cada una de ellas imagen especular de la otra, sí podrían superponerse en cuatro dimensiones. Pero añadía a continuación:³⁵⁷ «Sin embargo, como tal espacio no puede ser pensado, la superposición es imposible.» Kummer, hacia 1860, todavía se burlaba de la idea de una geometría tetradimensional. Las objeciones que todos esos autores planteaban frente a una geometría de dimensión más elevada eran razonables en tanto que se identificaba la geometría con el estudio del espacio físico.

Pero gradual e inconscientemente los matemáticos comenzaron a introducir conceptos que tenían poco o ningún significado físico

directo. Entre ellos, los más turbadores eran los números negativos y complejos. Como esos dos tipos de números no existían «realmente» en la naturaleza, eran todavía sospechosos a comienzos del siglo XIX, aunque ya se utilizaban ampliamente entonces. La representación geométrica de los números negativos como distancias en una dirección sobre una recta y la de los números complejos como puntos o vectores en el plano complejo que, como Gauss subrayó acerca de los últimos, les dio significado intuitivo haciéndolos así admisibles, pudo también retrasar la conciencia de que la matemática trata con conceptos ideados por el hombre. Pero la introducción de los cuaterniones, las geometrías no euclídeas, los elementos complejos en geometría, la geometría n -dimensional, las funciones extrañas y los números transfinitos obligaron a reconocer la artificialidad de las matemáticas.

En relación con esto ya hemos indicado el impacto de las geometrías no euclídeas (cap. 36, sec. 8), y nos detendremos ahora en el efecto producido por la geometría n -dimensional. El concepto aparece ya, de forma inocua, en los trabajos analíticos de D'Alembert, Euler y Lagrange. D'Alembert sugirió la posibilidad de considerar al tiempo como una cuarta dimensión en su artículo de la *Encyclopédie* sobre «Dimensión». Lagrange, al estudiar la reducción de formas cuadráticas a su expresión canónica, introdujo informalmente formas en n variables, y utilizó también el tiempo como una cuarta dimensión en su *Mécanique Analytique* (1788) y en su *Théorie des fonctions analytiques* (1797). En esta última obra, decía: «Podemos así considerar a la mecánica como una geometría de cuatro

dimensiones, y a la mecánica analítica como una extensión de la geometría analítica.» Los trabajos de Lagrange ponían así a un mismo nivel las tres coordenadas espaciales y una cuarta para representar el tiempo. Más tarde, George Green ya no vacilaba en su artículo de 1828 sobre teoría del potencial en considerar problemas de potencial en « dimensiones, diciendo de esa teoría: «Ya no está confinada, como antes, a las tres dimensiones del espacio.»

Esos primeros desarrollos en « dimensiones no se planteaban como una investigación propiamente geométrica; eran generalizaciones naturales de trabajos analíticos que ya no estaban ligados a la geometría. En parte, esa introducción del lenguaje «-dimensional se planteaba únicamente como algo conveniente, que ayudaba al pensamiento analítico. Era útil pensar en $(x_1, x_2 \dots x_n)$ como un punto, y en una ecuación en « variables como una hipersuperficie en el espacio «-dimensional, porque pensando en términos de lo que eso significa en la geometría tridimensional se podía introducir cierta intuición en el trabajo analítico. De hecho, Cauchy ³⁵⁸ subrayó que el concepto de espacio n -dimensional es útil en muchas investigaciones analíticas, especialmente las de teoría de números.

No obstante, el estudio serio de la geometría n -dimensional, sin que implicara un espacio físico de « dimensiones, se emprendió también en el siglo XIX; el fundador de esta geometría abstracta fue Grassmann, en su *Ausdehnungslehre* de 1844. Ahí se encuentra el concepto de geometría n -dimensional en su completa generalidad. En una nota publicada en 1845 decía Grassmann:

Mi Ausdehnungslehre construye el fundamento abstracto de la

teoría del espacio; esto es, queda libre de toda intuición espacial y es una ciencia puramente matemática; la geometría constituye únicamente su aplicación especial al espacio (físico).

Sin embargo, los teoremas que presento en él no son simples traducciones a un lenguaje abstracto de resultados geométricos; tienen una significación mucho más general, porque mientras que la geometría ordinaria se limita a las tres dimensiones del espacio (físico), la ciencia abstracta queda libre de esa limitación.

Grassmann añade que la geometría ordinaria es considerada impropriamente como una rama de la matemática pura, ya que se trata en realidad de una rama de la matemática aplicada que trata de un tema no creado por el intelecto sino exterior a él: la materia. Pero, dice, es posible crear una disciplina puramente intelectual que se ocupe de la extensión como concepto, más que del espacio percibido por las sensaciones. Así, la obra de Grassmann aparece como representativa de la corriente que afirma que el pensamiento puro puede producir construcciones arbitrarias, que pueden ser o no físicamente aplicables.

Cayley, independientemente de Grassmann, se propuso también tratar analíticamente la geometría n -dimensional «sin recurrir a nociones metafísicas». En el *Cambridge Mathematical Journal* de 1845 publicó unos «Capítulos de Geometría Analítica en N Dimensiones»,³⁵⁹ en los que ofrecía resultados analíticos en «variables, que para $n = 3$ se concretaban en teoremas ya conocidos

sobre superficies. Aunque no hizo nada especialmente nuevo en geometría n -dimensional, el concepto de ésta quedaba plenamente establecido en esos trabajos.

Cuando Riemann presentó su *Habilitationsvortrag* en 1854, «*Über die Hypothesen welche die Geometrie zu Grunde liegen*», no vaciló en tratar variedades n -dimensionales, aunque su objetivo principal era la geometría del espacio físico tridimensional. Los que siguieron el camino abierto por ese trabajo básico —Helmholtz, Lie, Christoffel, Beltrami, Lipschitz, Darboux y otros— continuaron trabajando en el espacio n -dimensional.

La noción de geometría n -dimensional encontró tenaz resistencia en algunos matemáticos incluso mucho tiempo después de haber sido introducida. Aquí, como en el caso de los números negativos y complejos, la matemática progresaba más allá de los conceptos sugeridos por la experiencia, y los matemáticos todavía tenían que asumir que su objeto podía consistir en conceptos creados por la mente, y que ya no era, si es que alguna vez lo había sido, una lectura de la naturaleza.

No obstante, en torno a 1850, fue ganando aceptación la idea de que la matemática puede introducir y tratar conceptos y teorías bastante arbitrarios que no tienen una interpretación física inmediata, pero que sin embargo pueden ser útiles, como en el caso de los cuaterniones, o satisfacer un deseo de generalidad, como en el caso de la geometría n -dimensional. Hankel, en su *Theorie der complexen Zahlensysteme* (1867, p. 10) defendió la matemática como «puramente intelectual, una teoría pura de las formas, que

tiene como objeto no la combinación de cantidades o de sus imágenes, los números, sino asuntos del pensamiento a los que pueden corresponder objetos o relaciones reales, si bien esa correspondencia no es precisa».

Cantor, en defensa de su creación de los números transfinitos como cantidades realmente existentes, aducía que la matemática se distingue de otras ciencias por su libertad para crear sus propios conceptos sin atender a la realidad transitoria. En 1883³⁶⁰ escribía: «La matemática es completamente libre en su desarrollo, y sus conceptos sólo se ven restringidos por la necesidad de ser no contradictorios y estar coordinados con los conceptos previamente introducidos mediante definiciones precisas... La esencia de la matemática reside en su libertad.» Prefería el término «matemática libre» al más usual de «matemática pura».

Esta nueva visión de las matemáticas se extendió a sus ramas más antiguas y físicamente asentadas. En su *Universal Algebra* (1898), Alfred North Whitehead decía (p. 11):

... el álgebra no depende de la aritmética para la validez de sus leyes de transformación. Si hubiera tal dependencia es obvio que tan pronto como se volvieran aritméticamente ininteligibles las expresiones algebraicas, todas las leyes referidas a ellas tendrían que perder su validez. Pero las leyes del álgebra, aunque sugeridas por la aritmética, no dependen de ella; dependen por entero de convenios mediante los que se establece que ciertos modos de agrupar los símbolos deben considerarse como idénticos, asignándose así ciertas propiedades a los

signos que forman los símbolos del álgebra.

El álgebra es un desarrollo lógico independiente del significado. «Es obvio que podemos tomar los signos que queramos y manipularlos de acuerdo con cualquier regla que decidamos asignarles» (p. 4). Whitehead indica que tales manipulaciones arbitrarias pueden ser frívolas, y que sólo son significativas las construcciones a las que se puede atribuir algún sentido, o que pueden utilizarse de alguna forma.

También la geometría cortó sus amarras con la realidad física. Como señalaba Hilbert en sus *Grundlagen* de 1899, la geometría habla de cosas cuyas propiedades quedan especificadas en los axiomas. Aunque Hilbert se refería únicamente a la estrategia con la que se deben enfocar las matemáticas con objeto de examinar su estructura lógica, apoyó y alentó sin embargo la opinión de que la matemática es algo completamente distinto de los conceptos y leyes de la naturaleza.

4. La pérdida de la verdad

La introducción y la aceptación gradual de conceptos que no tienen contrapartida inmediata en el mundo real, obligó naturalmente a reconocer que la matemática es una creación humana, en cierto modo, arbitraria, más que una idealización de las realidades de la naturaleza, obtenida únicamente a partir de ésta. Pero acompañando a este reconocimiento, y empujando de hecho a su aceptación, se produjo un descubrimiento más profundo: que la

matemática no es un conjunto de verdades acerca de la naturaleza. El desarrollo que planteó la cuestión de la verdad fue la geometría no euclídea, aunque su impacto se vio aplazado por el característico conservadurismo y estrechez de miras de la mayoría de los matemáticos. El filósofo David Hume (1711-1776) había señalado ya que la naturaleza no se acomoda a patrones fijos ni leyes necesarias; pero la opinión dominante, expresada por Kant, era que las propiedades del espacio físico eran euclídeas. Incluso Legendre, en sus *Eléments de géométrie* de 1794, creía todavía que los axiomas de Euclides eran verdades auto-evidentes.

Con respecto al menos a la geometría, la opinión que hoy parece correcta fue expresada primeramente por Gauss. A comienzos del siglo XIX ya estaba convencido de que la geometría es una ciencia empírica, que debe alinearse junto a la mecánica, mientras que la aritmética y el análisis eran para él verdades a priori. En 1830 escribía a Bessel:³⁶¹

Según mi convicción más profunda, la teoría del espacio ocupa un lugar en nuestro conocimiento a priori completamente diferente del que ocupa la aritmética pura. En todo nuestro conocimiento de la primera falta la convicción completa de necesidad (también de verdad absoluta) que caracteriza a la segunda; debemos añadir humildemente que si el número es un mero producto de nuestra mente, el espacio tiene una realidad fuera de ella cuyas leyes no podemos prescribir completamente a priori.

Sin embargo parece que Gauss se hallaba en cierta contradicción consigo mismo, porque también expresó la opinión de que toda la matemática es creación humana. En otra carta a Bessel, del 21 de noviembre de 1811, en la que hablaba de funciones de una variable compleja, decía:³⁶² «No se debe olvidar nunca que las funciones, como cualquier otra construcción matemática, son sólo nuestras propias creaciones, y que cuando la definición con la que uno comienza deja de tener sentido, no debería preguntarse qué es, sino qué tendría que suponer para que siguiera siendo significativa.»

Pese a las opiniones de Gauss acerca de la geometría, la mayoría de los matemáticos pensaban que en ella había verdades básicas. Bolyai creía que las verdades absolutas de la geometría eran los axiomas y teoremas comunes a la euclídea y la hiperbólica. No conocía la geometría elíptica, y en su época todavía no se podía concebir que muchos de esos axiomas comunes no lo eran para todas las geometrías.

En su artículo de 1854 «Sobre las hipótesis en que se basa la Geometría», Riemann creía todavía que había ciertas proposiciones acerca del espacio que serían a priori, aunque no estuviera incluida entre ellas la afirmación de que el espacio físico es verdaderamente euclídeo. Era, sin embargo, localmente euclídeo.

Cayley y Klein seguían apegados a la realidad de la geometría euclídea (vid. cap. 38, sec. 6). En su discurso presidencial a la Asociación Británica para el Avance de la Ciencia,³⁶³ decía: «... no que las proposiciones de la geometría sean sólo aproximadamente ciertas, sino que permanecen absolutamente ciertas con respecto a

ese espacio euclídeo que ha sido considerado durante tanto tiempo como el espacio físico de nuestra experiencia.» Aunque ambos habían trabajado en geometrías no euclídeas, las consideraban como novedades resultantes de introducir otras distancias en la geometría euclídea, y no eran capaces de ver que las geometrías no euclídeas son tan básicas y tan aplicables como la euclídea.

En los años noventa Bertrand Russell se planteó la cuestión de qué propiedades del espacio son necesarias para la experiencia y están implícitas en ésta. Es decir, aquellas propiedades a priori cuya negación convertiría en un sinsentido la experiencia. En su *Essay on the Foundations of Geometry* (1897) acepta que la geometría euclídea no representa un conocimiento a priori, pero defiende que la geometría proyectiva es a priori para cualquier planteamiento geométrico, conclusión que podemos entender a la luz de la importancia de ese enfoque en torno a 1900. A continuación añade como a priori para la geometría los axiomas comunes a la euclídea y todas las no euclídeas. La homogeneidad del espacio, la dimensionalidad finita y algún concepto de distancia hacen posibles las mediciones. Considera sin embargo como empíricos los hechos de que el espacio sea tridimensional y euclídeo.

Russell considera como un resultado técnico sin significación filosófica el que se puedan derivar las geometrías métricas de la proyectiva introduciendo una distancia. La geometría métrica, según él, es una rama de la matemática separada y lógicamente subsidiaria, y no es a priori. Con respecto a las geometrías euclídea y no euclídeas, se distancia de Cayley y Klein considerándolas a

todas ellas igualmente significativas. Como los únicos espacios métricos que poseen las propiedades indicadas más arriba son los euclídeos, hiperbólicos y simple o doblemente elípticos, concluye que esas son las únicas geometrías métricas posibles, y desde luego la euclídea es la única físicamente aplicable; las otras tienen importancia filosófica, mostrando que puede haber otras geometrías. Retrospectivamente podemos ahora decir que Russell sustituía el prejuicio euclídeo por un prejuicio proyectivo.

Aunque los matemáticos fueron lentos en reconocer el hecho, claramente observado por Gauss, de que no hay ninguna seguridad en la verdad física de la geometría euclídea, gradualmente fueron llegando a esa convicción y también a la sostenida por Gauss de que la verdad de la matemática reside en la aritmética y consecuentemente también en el análisis. Kronecker, por ejemplo, en su ensayo «Uber den Zahlbegriff» (Sobre el Concepto de Número),³⁶⁴ sostuvo la verdad de las disciplinas aritméticas, negándosela a la geometría. Gottlob Frege, acerca de cuya obra hablaremos más tarde, también insistió en la verdad de la aritmética.

Sin embargo, también la aritmética y el análisis construido sobre ella se hicieron pronto sospechosos. La creación de álgebras no conmutativas, en particular las de cuaterniones y matrices, plantearon naturalmente la cuestión de cómo podría estar uno seguro de que los números ordinarios poseyeran la propiedad privilegiada de la verdad acerca del mundo real. El ataque contra la veracidad de la aritmética vino en primer lugar de Helmholtz.

Después de haber insistido, en un ensayo famoso,³⁶⁵ en que nuestro conocimiento del espacio físico proviene únicamente de la experiencia y depende de la existencia de cuerpos rígidos que puedan servir, entre otras cosas, como reglas para medir, en su *Zahlen und Messen* (Contar y Medir, 1887) puso en cuestión las verdades de la aritmética. Consideraba como el principal problema de ésta el significado o la validez de la aplicación objetiva de la cantidad y la igualdad a la experiencia. La aritmética misma puede considerarse como una exposición consistente de las consecuencias de las operaciones aritméticas. Trata con símbolos y puede entenderse como un juego. Pero esos símbolos se aplican a objetos reales y a las relaciones entre ellos, y proporcionan resultados acerca del funcionamiento real de la naturaleza. ¿Cómo es esto posible? ¿Bajo qué condiciones se pueden aplicar a los objetos reales los números y las operaciones aritméticas? En particular, ¿cuál es el significado objetivo de la igualdad de dos objetos, y cómo puede tratarse la adición física como suma aritmética?

Helmholtz señalaba que la aplicabilidad de los números no es ni un accidente ni tampoco una prueba de la verdad de las leyes numéricas. *Algunos* tipos de experiencia los sugieren, y a esos es a los que se pueden aplicar. Para aplicar números a objetos reales, decía Helmholtz, éstos no deben desaparecer, o mezclarse entre sí, o dividirse. Al añadir físicamente una gota de lluvia a otra no se obtienen dos gotas de lluvia. Sólo la experiencia puede decirnos si los objetos de una colección física determinada mantienen en ella su identidad de manera que se pueda atribuir a la colección un

número definido de objetos. De igual modo, saber cuándo se puede aplicar la igualdad entre dos cantidades físicas también depende de la experiencia. Cualquier afirmación de igualdad cuantitativa debe satisfacer dos condiciones: si se intercambian los objetos, deben seguir siendo iguales. Y si el objeto a es igual a c y el objeto b es igual a c , los objetos a y b deben ser iguales. De esta manera podemos hablar de la igualdad de pesos o intervalos de tiempo, ya que para esos objetos sí se puede determinar la igualdad. Pero dos sonidos pueden ser indistinguibles para el oído de un tercero, intermedio entre ambos, aunque sí puedan distinguirse entre sí. Aquí dos cosas iguales a una tercera no son iguales entre sí. Tampoco se pueden sumar los valores de unas resistencias eléctricas conectadas en paralelo para obtener la resistencia total, ni se pueden combinar de cualquier manera los índices de refracción de diferentes medios.

Hacia finales del siglo XIX prevaleció la opinión de que todos los axiomas de las matemáticas son arbitrarios, constituyendo únicamente la base para deducir consecuencias de ellos. Como ya no se entendían como verdades acerca de los conceptos implicados en ellos, dejó de importar el significado físico de esos conceptos. Ese significado podía, a lo más, servir como guía heurística cuando los axiomas mantenían alguna relación con la realidad. Así, incluso los conceptos quedaron separados del mundo físico. Hacia 1900 las matemáticas se habían despegado de la realidad, abandonando clara e irreparablemente su pretensión a la verdad acerca de la naturaleza, y convirtiéndose en la búsqueda de las consecuencias

necesarias de axiomas arbitrarios acerca de cosas sin sentido.

La pérdida de la verdad y la aparente arbitrariedad, la naturaleza subjetiva de las ideas y resultados matemáticos, turbaron profundamente a muchos, que entendieron esto como una denigración de la matemática. Algunos de ellos adoptaron una postura mística que de alguna manera garantizara cierta realidad y objetividad a las matemáticas. Esos matemáticos suscribieron la idea de que la matemática es una realidad en sí misma, un cuerpo independiente de verdades, y que sus objetos nos vienen dados como nos vienen dados los objetos del mundo real; los matemáticos simplemente descubren los conceptos y sus propiedades. Hermite, por ejemplo, en una carta a Stieltjes,³⁶⁶ decía: «Creo que los números y las funciones del análisis no son el producto arbitrario de nuestras mentes; creo que existen fuera de nosotros con el mismo carácter de necesidad que los objetos de la realidad objetiva, y que nosotros los encontramos o descubrimos y los estudiamos como lo hacen los físicos, químicos y zoólogos.»

En el Congreso Internacional de Bolonia de 1928, Hilbert decía:³⁶⁷ «¿Qué sucedería con la verdad de nuestro conocimiento y con la existencia y el progreso de la ciencia si no hubiese ninguna verdad en las matemáticas? De hecho, actualmente aparece con demasiada frecuencia en los escritos profesionales y en lecciones orales cierto escepticismo o falta de confianza acerca del conocimiento; se trata de un tipo de ocultismo que considero perjudicial.»

Godfrey H. Hardy (1877-1947), sobresaliente analista del siglo XX, decía en 1928:³⁶⁸ «Los teoremas matemáticos son verdaderos o

falsos; su verdad o falsedad es absolutamente independiente de nuestro conocimiento de ellos. En *cierto* sentido, la verdad matemática forma parte de la realidad objetiva.» Expresó de nuevo la misma opinión en su libro *A Mathematician's Apology* (ed. de 1967, p. 123): «Creo que la realidad matemática está fuera de nosotros, que nuestra función consiste en descubrirla u observarla, y que los teoremas que describimos con grandilocuencia como nuestras «creaciones» son simplemente las anotaciones de nuestra observación.»

5. La matemática como el estudio de estructuras arbitrarias

Los matemáticos del siglo XIX se ocupaban prioritariamente del estudio de la naturaleza, y la física fue sin duda la fuente de inspiración más importante para el trabajo matemático. Los investigadores más descollantes —Gauss, Riemann, Fourier, Hamilton, Jacobi, Poincaré— y otros menos conocidos —Christoffel, Lipschitz, Du Bois-Reymond, Beltrami...— trabajaron directamente en problemas de física y en problemas matemáticos surgidos de investigaciones físicas. Incluso los autores considerados comúnmente como matemáticos puros, como por ejemplo Weierstrass, trabajaron en problemas físicos. De hecho, los problemas físicos proporcionaron más sugerencias e indicaciones para las investigaciones matemáticas que en cualquier otro siglo anterior, y para analizarlos se creó una matemática altamente compleja. Fresnel había señalado que «la naturaleza no se detiene ante dificultades de análisis», pero los matemáticos no se arredraron

y consiguieron vencer esas dificultades. La única rama importante que se había mantenido por una satisfacción intrínsecamente estética, al menos desde la obra de Diofanto, era la teoría de números.

No obstante, fue por primera vez en el siglo XIX cuando los matemáticos no sólo desarrollaron su trabajo más allá de las necesidades de la ciencia y la tecnología de su época, sino que plantearon y resolvieron cuestiones que no tenían que ver con problemas reales. La *raison d'être* de este desarrollo se podría describir así: la convicción bimilenaria de que la matemática constituía la verdad acerca de la naturaleza se había hecho pedazos. Pero las teorías matemáticas que ahora se reconocían como arbitrarias habían demostrado sin embargo su utilidad en el estudio de la naturaleza. Aunque las teorías existentes debían mucho históricamente a sugerencias de la naturaleza, quizá nuevas teorías construidas únicamente por la mente podrían ser también útiles para su representación. Los matemáticos se sintieron entonces libres para crear estructuras arbitrarias, quedando justificada una nueva libertad en la investigación matemática. Sin embargo, como algunas de las estructuras ya creadas hacia 1900, y muchas de las que vendrían después, parecían tan artificiales y alejadas de cualquier aplicación potencial, sus patrocinadores comenzaron a defenderlas como deseables en y por sí mismas.

El ascenso y aceptación gradual de la opinión de que la matemática debía ocuparse de estructuras arbitrarias que no estaban obligadas a tener nada que ver, en primera ni última instancia, con el estudio

de la naturaleza, condujo a un cisma que se conoce actualmente como el de la matemática pura frente a la aplicada. Semejante ruptura con la tradición no podía dejar de generar controversia. Nos limitaremos únicamente a citar unos pocos de los argumentos de cada bando.

Fourier había escrito, en el prefacio a su *Théorie Analytique de la Chaleur*: «El estudio profundo de la naturaleza es el campo más fértil para los descubrimientos matemáticos. Ese estudio ofrece no sólo la ventaja de un objetivo bien definido, sino también la de excluir cuestiones vagas y cálculos inútiles. Es un medio para construir el análisis en sí mismo y para descubrir qué ideas importan verdaderamente y cuáles debe preservar la ciencia. Las ideas fundamentales son aquellas que representan los acontecimientos naturales.» También subrayó la aplicación de las matemáticas a problemas socialmente útiles.

Aunque Jacobi había hecho investigaciones de primer orden en mecánica y astronomía, criticó vivamente las opiniones de Fourier. El 2 de julio de 1830 escribía a Legendre:³⁶⁹ «Es cierto que Fourier piensa que el objeto prioritario de la matemática es la utilidad pública y la explicación de los fenómenos naturales; pero un científico como él debería saber que el único objeto de la ciencia es el honor del espíritu humano y sobre esta base una cuestión de (teoría de) números es tan importante como una cuestión acerca del sistema planetario...»

A lo largo del siglo, la deriva de más y más investigadores hacia la matemática pura suscitó reacciones de protesta. Kronecker, por

ejemplo, escribía a Helmholtz: «La riqueza de su experiencia práctica con problemas sanos e interesantes dará un nuevo sentido y un nuevo ímpetu a la matemática... La especulación matemática unilateral e introspectiva conduce a campos estériles.»

Félix Klein, en su *Teoría Matemática del Giróscopo* (1897, pp. 1-2) afirmaba: «La gran necesidad del presente en la ciencia matemática es que la ciencia pura y los departamentos de ciencias físicas en los que encuentra sus más importantes aplicaciones vuelvan de nuevo a la asociación íntima que se mostró tan fructífera en las obras de Lagrange y Gauss.» Y Emile Picard, a comienzos de este siglo (*La Science moderne et son état actuel*, 1908) prevenía contra la tendencia a las abstracciones y los problemas sin interés.

Un poco después, Félix Klein insistía de nuevo sobre el asunto.³⁷⁰ Temiendo el abuso de la libertad para crear estructuras arbitrarias, subrayaba que éstas son «la muerte de la ciencia. Los axiomas de la geometría no son... proposiciones arbitrarias, sino que se deducen, en general, de nuestra percepción espacial, y están determinados en cuanto a su contenido preciso por la utilidad.» Para justificar los axiomas no euclídeos, Klein señalaba que la visualización puede verificar el axioma euclídeo de las paralelas sólo hasta ciertos límites. En otra ocasión indicaba que «quienquiera que reivindique el privilegio de la libertad debe también soportar la responsabilidad». Klein entendía por «responsabilidad» la dedicación a investigar la naturaleza.

A pesar de las advertencias, la tendencia a la abstracción, a generalizar porque sí los resultados existentes, y la búsqueda de

problemas arbitrariamente planteados, prosiguió en nuestro siglo. La razonable necesidad de estudiar toda una clase de problemas para saber más acerca de algún caso concreto, y de abstraer para quedarse con lo esencial de un problema, se convirtieron en excusas para complacerse en generalidades y abstracciones en y por sí mismas.

En parte para contrarrestar esa tendencia a la generalización, Hilbert no sólo insistió en que los problemas concretos constituyen el flujo vital de las matemáticas, sino que se tomó el trabajo de publicar en 1900 una lista de los 23 más sobresalientes (vid. bibliografía) y de citarlos en una conferencia que dio en el Segundo Congreso Internacional de Matemáticos en París. El prestigio de Hilbert llevó a muchos a ocuparse de esos problemas. Ningún honor podía sobrepasar al de resolver un problema planteado por alguien tan respetado. Pero la tendencia a la creación libre y a las abstracciones y generalizaciones no pudo contenerse. La matemática se despegó de la naturaleza y la ciencia para seguir su propio curso.

6. El problema de la consistencia

La matemática, desde el punto de vista de la lógica, era a finales del siglo XIX una colección de estructuras construida cada una de ellas sobre su propio sistema de axiomas. Como ya hemos señalado, una de las propiedades necesarias de cualquier estructura de ese tipo es la consistencia de sus axiomas. Mientras se consideró la matemática como la verdad acerca de la naturaleza, no cabía la

posibilidad de que pudieran surgir teoremas contradictorios, y pensarlo siquiera habría sido considerado como absurdo. Cuando se crearon las geometrías no euclídeas, su aparente discrepancia con la realidad planteó la cuestión de su consistencia. Como hemos visto, se respondió a esta cuestión haciendo depender la consistencia de las geometrías no euclídeas de la de la geometría euclídea.

En los años ochenta, la conciencia de que ni la aritmética ni la geometría euclídea son verdades en sí hizo imperativa la investigación de la consistencia de esas ramas de la matemática. Peano y su escuela comenzaron a considerar este problema en los años noventa. Creía que se podrían concebir pruebas claras que dilucidaran la cuestión, pero los acontecimientos demostraron que estaba equivocado. Hilbert consiguió establecer la consistencia de la geometría euclídea sobre la hipótesis de que la aritmética es consistente (cap. 42, sec. 3). Pero la consistencia de esta última no había quedado establecida, y Hilbert planteó este problema como el segundo de su lista en el Segundo Congreso Internacional de 1900; en su «Axiomatisches Denken»,³⁷¹ lo señaló como el problema básico de los fundamentos de la matemática. Muchos otros investigadores eran también conscientes de la importancia del problema. En 1904 Alfred Pringsheim (1850-1941),³⁷² afirmó que la verdad que persigue la matemática no es ni más ni menos que la consistencia. En el capítulo 51 examinaremos los trabajos sobre este problema.

7. Una mirada hacia el futuro

El ritmo de la creación matemática se ha venido incrementando sin pausa desde 1600, y esto sigue siendo cierto para el siglo XX, en el que se desarrollaron la mayoría de los campos explorados en el XIX. No obstante, los detalles de los nuevos trabajos en esos campos interesarían únicamente a los especialistas. Limitaremos por tanto nuestro examen del siglo XX a los campos que aparecieron o se hicieron por primera vez importantes en este período. Además, nos ocuparemos tan sólo de los comienzos de esas investigaciones. Los desarrollos del segundo y tercer cuartos de este siglo son todavía demasiado recientes como para poderlos evaluar correctamente. Ya hemos visto como muchas áreas exploradas vigorosa y entusiásticamente en el pasado, consideradas por sus abogados como la esencia de la matemática, resultaron ser modas pasajeras o con poca influencia en el curso general de ésta. Por muy confiados que estén los matemáticos del último medio siglo en la gran importancia de sus trabajos, el lugar que ocupen sus contribuciones en la historia de la matemática no puede decidirse todavía.

Bibliografía

- Fang, J.: *Hilbert*, Paideia Press, 1970. Esbozos de la obra matemática de Hilbert.
- Hardy, G. H.: *A Mathematician's Apology*, Cambridge University Press, 1940 y 1967. Hay versión castellana, *Autojustificación de un matemático*, Barcelona, Ariel, 1981.
- Helmholtz, H. von: *Counting and Measuring*. D. van Nostrand,

1930. Traducción al inglés de *Zählen und Messen. Wissenschaftliche Abhandlungen*, 4, 356-391. - «Über den Ursprung Sinn und Bedeutung der geometrischen Sätze»; traducción al inglés: «On the Origin and Significance of Geometrical Axioms», en Helmholtz: *Popular Scientific Lectures*, Dover (reimpresión), 1962, pp. 223-249. También en James R. Newman: *The World of Mathematics*, Simón & Schuster, 1956, vol. 1, pp. 647-668. Hay versión castellana, *Sigma, el mundo de las matemáticas*, Barcelona, Grijalbo, 8.^a ed., 1983. Véase también Helmholtz: *Wissenschaftliche Abhandlungen*, 2, 640-660.
- Hilbert, David: «Sur les problèmes futurs des mathématiques». *Comptes Rendus du Deuxième Congrès International des Mathématiciens*, Gaut-hier-Villars, 1902, 58-114. También se puede encontrar en alemán en *Nachrichten König. Ges. der Wiss. zu Gótt.*, 1900, 253-297, y en Hilbert: *Gesammelte Abhandlungen*, 3, 290-329. Hay una traducción al inglés en *Amer. Math. Soc. Bull.*, 8, 1901/1902, 437-479.
 - Klein, Félix: «Über Arithmetisierung der Mathematik». *Ges. Math. Abh.*, 2, 232-240. Traducción al inglés en *Amer. Math. Soc. Bull.*, 2, 1895/1896, 241-249.
 - Pierpoint, James: «On the Arithmetization of Mathematics». *Amer. Math. Soc. Bull.*, 5, 1898/1899, 394-406.
 - Poincare, Henri: *The Foundations of Science*, Science Press, 1913. Vid. en particular pp. 43-91.
 - Reid, Constance: *Hilbert*, Springer-Verlag, 1970. Se trata de

una biografía.

Capítulo 44

La teoría de funciones de una o varias variables reales

Si Newton y Leibniz hubieran llegado a imaginarse que las funciones continuas no tienen por qué tener necesariamente derivada (y esto es lo que ocurre, en general), nunca se habría creado el cálculo diferencial.

Emile Picard

Contenido:

1. *Los orígenes*
2. *La integral de Stieltjes*
3. *Prime-ros trabajos sobre contenido y medida*
4. *La integral de Le-besgue*
5. *Generalizaciones*

Bibliografía

1. Los orígenes

La teoría de funciones de una o varias variables reales tuvo su origen en el intento de llegar a entender y clarificar cierto número de descubrimientos extraños que se habían ido haciendo a lo largo del siglo **XIX**. La aparición de funciones continuas pero no diferenciables, de series de funciones continuas cuya suma era

discontinua, de funciones continuas que no eran monótonas a trozos, de funciones con derivadas acotadas pero que no eran integrables en el sentido de Riemann, de curvas rectificables pero que no lo eran de acuerdo con la definición de longitud de un arco de curva dada por el cálculo infinitesimal, y de funciones no integrables que eran límite de sucesiones de funciones integrables, todo ello, en fin, parecía contradecir flagrantemente el comportamiento que se esperaba de las funciones, las derivadas y las integrales. Otro tipo de motivaciones para estudiar más a fondo el comportamiento de las funciones provenía de las investigaciones sobre las series de Fourier. Esta teoría, tal como había sido construida por Dirichlet, Riemann, Cantor, Ulisse Dini (1845-1918), Jordán y otros matemáticos del siglo **XIX**, resultaba ser una herramienta muy satisfactoria para la matemática aplicada, pero lo cierto era que las propiedades de estas series, en el estado de desarrollo alcanzado, no constituían aún una teoría que pudiera satisfacer al matemático puro. Todavía se echaban de menos las relaciones de unidad, simetría y completitud que debería haber entre funciones y series.

Las investigaciones que se desarrollaban en teoría de funciones hicieron énfasis en la teoría de la integral, porque parecía que la mayoría de las incongruencias se podrían resolver extendiendo este concepto, y por tanto estos trabajos pueden considerarse en gran medida como una continuación directa de la obra de Riemann, Darboux, Du Bois-Reymond, Cantor y otros (véase cap. 40, sec. 4).

2. La integral de Stieltjes

En realidad, la primera extensión del concepto de integral tuvo su origen en un tipo de problemas completamente distintos de los que acabamos de mencionar. En 1894 publicaba Thomas Jan Stieltjes (1856-1894) su largo artículo titulado «*Recherches sur les fractions continues*»,³⁷³ trabajo de una gran originalidad en el que partía de un problema muy particular para terminar resolviéndolo con rara elegancia. Esta obra venía a sugerir problemas de un tipo completamente nuevo tanto en la teoría de funciones analíticas como en la de funciones de una variable real. En particular, y con objeto de representar el límite de una sucesión de funciones analíticas, se vio obligado Stieltjes a introducir una nueva integral que generalizase la idea de Riemann-Darboux.

Stieltjes comienza considerando una distribución positiva de masa a lo largo del eje Ox , lo que venía a generalizar el concepto de densidad puntual, que ya había sido utilizado, desde luego. Hace observar que tal distribución de masa vendrá dada por una función creciente $\phi(x)$ que exprese la masa total acumulada en el intervalo $[0,x]$ para cada $x > 0$, donde las discontinuidades de ϕ corresponderán a masas concentradas en un punto. Para tal distribución de masas sobre el intervalo $[a,b]$ define las sumas de Riemann

$$\sum_{i=0}^n f(\xi_i) (\phi(x_{i+1}) - \phi(x_i))$$

donde los puntos x_0, x_1, \dots, x_n forman una partición del intervalo $[a, b]$ y ξ_i está en $[x_i, x_{i+1}]$. Stieltjes demuestra después que cuando f es continua sobre $[a, b]$ y el subintervalo máximo de la partición tiende a cero, entonces la suma tiende a un límite que representa por

$$\int_a^b f(x) d\phi(x)$$

A pesar de que utilizó esta integral en su propia obra, Stieltjes no investigó más sobre el concepto mismo de integral, excepto para definir de manera obvia, para el intervalo $(0, \infty)$

$$\int_0^{\infty} f(x) d\phi(x) = \lim_{b \rightarrow \infty} \int_0^b f(x) d\phi(x)$$

Este concepto de integral no fue adoptado por los matemáticos hasta mucho más tarde, cuando se le encontraron gran cantidad de aplicaciones (véase cap. 47, sec. 4).

3. Primeros trabajos sobre contenido y medida

Muy distinta fue la línea de pensamiento que condujo a otra generalización diferente de la idea de integral, la integral de Lebesgue. El estudio de los conjuntos de puntos de discontinuidad de funciones planteó el problema de cómo medir la extensión o «longitud» de tales conjuntos, porque precisamente la «extensión» de

dichas discontinuidades es lo que determina la integrabilidad de la función. La teoría del contenido, y más tarde la teoría de la medida, fueron introducidas precisamente para extender la idea de longitud a conjuntos de puntos que no sean simples intervalos de la recta real usual.

El concepto de contenido se basa en la siguiente idea: considérese un conjunto E de puntos distribuidos de alguna manera sobre el intervalo $|a,b|$. En términos un poco imprecisos por el momento, supongamos que es posible encerrar o recubrir estos puntos mediante subintervalos pequeños de $[a,b]$ de manera que los puntos de E sean o bien interiores a uno de los intervalos o en el peor de los casos uno de sus extremos. Reduzcamos las longitudes de estos subintervalos cada vez más, añadiendo otros si es necesario para seguir recubriendo los puntos de E , a la vez que reducimos la suma de las longitudes de todos ellos. Al extremo inferior de las sumas de estos subintervalos que recubren los puntos de E se llama el contenido (exterior) de E . Esta formulación un poco vaga no es la del concepto adoptado finalmente de manera definitiva, pero nos puede servir para entender lo que estaban intentando hacer los matemáticos.

Una definición de contenido (exterior) fue dada por Du Bois-Reymond en su obra *Die allgemeine Funktionentheorie* (1882), por Axel Harnack (1851-1888) en su *Die Elemente der Differential-und Integralrechnung* (1881), por Otto Stolz³⁷⁴ y por Cantor.³⁷⁵ Stolz y Cantor extendieron además la noción de contenido a conjuntos de dos y de más dimensiones utilizando rectángulos, paralelepípedos,

etc. en lugar de intervalos.

La utilización de esta noción de contenido, que desgraciadamente no resultó ser satisfactoria en todos los sentidos, reveló sin embargo la existencia de conjuntos no densos en ninguna parte (es decir, contenidos en un intervalo pero no densos en ninguno de sus subintervalos) de contenido positivo, y que las funciones con tales conjuntos de puntos de discontinuidad no eran integrables en el sentido de Riemann. También había funciones con derivadas acotadas no integrables. Sin embargo, los matemáticos de la época, la década de los ochenta, seguían pensando que el concepto de integral de Riemann no se podría generalizar.

Con objeto de superar las limitaciones de la anterior teoría del contenido y de rigorizar la noción de área de una región plana, Peano introduce una definición de contenido más completa y muy mejorada en su obra *Applicazioni geometriche del cálculo infinitesimale* (1887). Peano introduce aquí las nociones de contenido interior y exterior de una región. Supongámonos situados en dos dimensiones; el contenido interior es el extremo superior de las áreas de todas las regiones poligonales contenidas en la región dada R , mientras que el contenido exterior es el extremo inferior de las áreas de todas las regiones poligonales que contengan a la región R . Si los contenidos interior y exterior coinciden, a este valor común se le llama área de la región R . Para los conjuntos lineales la idea es análoga, utilizando esta vez intervalos en vez de polígonos. Peano hace observar que, si la función $f(x)$ es no negativa sobre el intervalo $[a, b]$, entonces

$$\int_{\underline{a}}^{\underline{b}} f(dx) = C_i(R) \quad y \quad \int_{\underline{a}}^{\bar{b}} = C_e(R)$$

donde la primera integral es el extremo superior de las sumas de Riemann inferiores de f sobre $[a,b]$ y la segunda es el extremo inferior de las sumas de Riemann superiores, y $C_i(R)$ y $C_e(R)$ son los contenidos interior y exterior respectivamente de la región R limitada por el grafo de f y el eje Ox . Así pues, f es integrable si y sólo si R tiene área definida, en el sentido de que $C_i(R) = C_e(R)$.

A Jordán le corresponde el mérito de haber dado el paso más atrevido y definitivo en la teoría del contenido (*étendue*) de todo el siglo XIX. También introduce unos contenidos interior y exterior³⁷⁶ pero formula estos conceptos de una manera un poco más eficaz. Su definición para un conjunto de puntos E contenido en un intervalo $[a,b]$ comienza por el contenido exterior; hay que recubrir E por un conjunto finito de subintervalos de $[a,b]$ tales que cada punto de E sea interior o extremo de uno de esos subintervalos. Al extremo inferior de las sumas de las longitudes de todos los conjuntos de ese tipo de subintervalos que contienen al menos un punto de E se le llama contenido exterior de E . El contenido interior de E se define como el extremo superior de las sumas de los subintervalos de $[a,b]$ que contengan sólo puntos de E . Si los contenidos interior y exterior de E son iguales, se dice que E tiene contenido definido. Jordán aplicó la misma idea a conjuntos de un

espacio n -dimensional cualquiera, reemplazando los subintervalos por rectángulos y sus análogos en dimensiones mayores. Ahora ya podía demostrar Jordán lo que llamó propiedad de aditividad: el contenido de la suma de un número *finito* de conjuntos disjuntos con contenido definido es la suma de los contenidos de los mismos. Esto no se verificaba en las teorías del contenido anteriores, excepto en la de Peano.

El interés de Jordán por el contenido provenía del intento de precisar la teoría de las integrales dobles extendidas a una región plana E . La definición generalmente adoptada consistía en dividir el plano en cuadrados R_{ij} por medio de rectas paralelas a los ejes de coordenadas. Esta partición del plano induce una partición de E en regiones E_{ij} . Entonces, por definición,

$$\int_E f(x,y)dE = \lim_{\Delta x, \Delta y \rightarrow 0} \sum_{i,j} f(x_i, y_i) a(R_{ij})$$

donde $a(R_{ij})$ representa el área de E_{ij} , y la suma se extiende a todas las R_{ij} totalmente contenidas en E ya todas las R_{ij} que contengan puntos de E así como puntos exteriores a E . Para que exista la integral es necesario demostrar que la contribución de los R_{ij} que no están completamente contenidos en E puede despreciarse, es decir, que la suma de las áreas de los R_{ij} que contienen puntos frontera de E tiende a cero con las dimensiones de los R_{ij} . En general, hasta entonces se había supuesto que éste era el caso, y Jordán mismo lo hizo así en la primera edición de su *Cours d'Analyse* (vol. 2, 1883).

Sin embargo, el descubrimiento de curvas tan peculiares como la de Peano, que llena un cuadrado, forzó a los matemáticos a ser más cautelosos. Si E tiene contenido de Jordán bidimensional bien definido, entonces se pueden desprestigiar los R_{ij} que contienen la frontera de E . Jordán también consiguió demostrar resultados sobre el cálculo de integrales dobles por integración iterada.

La segunda edición del *Cours d'Analyse* (vol. 1, 1893) de Jordán incluye ya su tratamiento del contenido y su aplicación a la integración. Aunque superior a las de sus predecesores, la definición de contenido de Jordán no era del todo satisfactoria. Según ella, un conjunto abierto acotado podría no tener contenido definido, y el conjunto de los puntos racionales de un intervalo acotado no tenía contenido.

La etapa siguiente en la teoría del contenido se debe a Borel. Borel se vio conducido al estudio de la teoría de la medida, tal como él la llamó, trabajando con conjuntos de puntos en que convergen series que representan funciones de variable compleja. Su *Leçons sur la théorie des fonctions* (1898) contiene sus primeros trabajos importantes sobre el tema; Borel se dio cuenta de los defectos de las teorías anteriores del contenido y les puso remedio.

Cantor había demostrado ya que todo abierto U de la recta es unión de una familia *numerable* de intervalos abiertos disjuntos dos a dos. En vez de aproximar U encerrándolo en un conjunto finito de intervalos, Borel propone, utilizando el resultado de Cantor, definir la medida de un abierto acotado U como la suma de las longitudes de los intervalos componentes. A continuación define la medida de

la suma de una cantidad *finita o numerable* de conjuntos medibles disjuntos como la suma de sus medidas individuales, y la medida del conjunto $A-B$, siendo A y B medibles y B contenido en A , como la correspondiente diferencia de las medidas. Con estas definiciones podía atribuir una medida a los conjuntos formados sumando cualquier cantidad finita o numerable de conjuntos medibles disjuntos y a la diferencia de dos conjuntos medibles cualesquiera A y B , siempre que A contenga a B . A continuación estudia los conjuntos de medida 0 y demuestra que cualquier conjunto de medida mayor que 0 tiene que ser no numerable.

La teoría de la medida de Borel suponía una mejora indudable de las teorías del contenido de Peano y Jordán, pero no iba a ser aún la última palabra sobre el tema, ni estudió Borel sus aplicaciones a la teoría de integración.

4. La integral de Lebesgue

La generalización de la idea de medida e integral que se considera hoy como definitiva se debe a Henri Lebesgue (1875-1941), discípulo de Borel y más tarde profesor del Collège de France. Siguiendo las ideas de Borel y también las de Jordán y Peano, presentó por primera vez sus propias ideas sobre la medida y la integral en su tesis «Intégrale, longueur, aire».³⁷⁷ Su obra vino a reemplazar todas las creaciones del siglo XIX y, en particular, mejoró la teoría de la medida de Borel.

La teoría de integración de Lebesgue se basa en su definición de medida de conjuntos de puntos, y ambas ideas se aplican a

conjuntos del espacio n -dimensional. Como ejemplo, nos limitaremos al caso unidimensional. Sea E un conjunto de puntos contenido en el intervalo $[a,b]$. Los puntos de E pueden encerrarse como puntos interiores de una familia finita o *infinita numerable* de intervalos d_1, d_2, \dots contenidos en $[a,b]$. (Los extremos de $[a,b]$ pueden considerarse, en su caso, como extremos de algún d_i) Puede demostrarse que la familia de intervalos $\{d_i\}$ se puede sustituir por otra formada por intervalos $\delta_1, \delta_2, \dots$ no rampantes tales que todo punto de E o es punto interior de uno de estos intervalos o el extremo común de dos intervalos adyacentes. Sea $\Sigma\delta_n$ la suma de las longitudes de los δ_i . El extremo inferior de las $\Sigma\delta_n$ para todas las posibles familias $\{\delta_i\}$ recibe el nombre de medida exterior de E y se representa por $m_e(E)$. La medida interior de E , $m_i(E)$, viene definida como la medida exterior del conjunto $C(E)$, es decir, del complemento de E con respecto a $[a,b]$, o los puntos de $[a,b]$ que no pertenecen a E .

Con estas definiciones se puede demostrar una serie de resultados auxiliares, incluido el hecho de que $m_i(E) \leq m_e(E)$. Al conjunto E se le llama por definición medible si se verifica que $m_i(E) = m_e(E)$, y a este valor común se le llama su medida $m(E)$. Lebesgue demuestra que una unión numerable de conjuntos medibles disjuntos dos a dos es medible y que su medida es la suma de las medidas de los conjuntos componentes. También se verifica que todos los conjuntos medibles en el sentido de Jordán lo son en el sentido de Lebesgue y la medida es la misma. El concepto de medida de Lebesgue difiere del de Borel por la adjunción de un conjunto de

medida nula en el sentido de Borel. Lebesgue llamó la atención también sobre la existencia de conjuntos no medibles.

El siguiente concepto importante introducido por Lebesgue fue el de función medible. Sea E un conjunto medible acotado del eje Ox ; la función $f(x)$, definida en todos los puntos E , se llamará medible en E si el subconjunto formado por los puntos de E tales que $f(x) > A$ es medible, para toda constante A .

Por último, llegamos a la definición de Lebesgue de la integral. Sea $f(x)$ una función acotada y medible definida sobre el conjunto medible E contenido en $[a,b]$. Sean A y B los extremos inferior y superior de $f(x)$ sobre E . Dividamos el intervalo $[A,B]$ del eje Oy en n intervalos parciales

$$[A, l_1], [l_1, l_2], \dots [l_{n-1}, B]$$

donde $A = l_0$ y $B = l_n$. Sea e_r el conjunto de los puntos de E para los que $l_{r-1} \leq f(x) \leq l_r$ para $r = 1, 2, 3, \dots, n$. Entonces e_1, e_2, \dots, e_n son conjuntos medibles. Construyamos las sumas S y s definidas por

$$S = \sum_1^n l_r m(e_r) \quad s = \sum_1^n l_{r-1} m(e_r)$$

Estas sumas S y s tienen un extremo inferior J y un extremo superior I respectivamente. Lebesgue demuestra entonces que para toda función acotada medible, $I = J$, y a este valor común se le llama la integral de Lebesgue de $f(x)$ sobre E , y se representa por

$$I = \int_E f(x) dx$$

Si E es todo el intervalo $[a,b]$, entonces se usa la notación $\int_a^b f(x) dx$ usual, pero entendiendo que la integral es ahora en el sentido de Lebesgue. Si $f(x)$ es integrable en el sentido de Lebesgue y el valor de la integral es finito, entonces se dice que $f(x)$ es sumable, término introducido por Lebesgue mismo. Toda función $f(x)$ que sea integrable en el sentido de Riemann sobre $[a,b]$ lo es en el sentido de Lebesgue, pero no necesariamente al revés. Si $f(x)$ es integrable en ambos sentidos entonces los valores de las dos integrales coinciden. La generalidad de la integral de Lebesgue deriva del hecho de que una función integrable para Lebesgue no necesita ser continua casi por doquier (es decir, excepto en un conjunto de medida nula). Así por ejemplo, la función de Dirichlet, que vale 1 para los valores racionales de x , y 0 para los valores irracionales del intervalo $[a,b]$ es totalmente discontinua y, aunque no sea integrable en el sentido de Riemann sí es integrable para Lebesgue, y en este caso $\int_a^b f(x) dx = 0$.

Esta noción de integral de Lebesgue puede extenderse a funciones más generales, por ejemplo a funciones no acotadas. Si $f(x)$ es integrable en el sentido de Lebesgue pero no acotada en el intervalo de integración, la integral converge absolutamente. Las funciones no acotadas pueden ser integrables en el sentido de Lebesgue pero no en el de Riemann y recíprocamente.

Para fines prácticos la integral de Riemann suele ser suficiente. De hecho, Lebesgue mismo demostró (*Leçons sur l'intégration et la recherche des fonctions primitives*, 1904) que una función acotada es integrable para Riemann si y sólo si sus puntos de discontinuidad forman un conjunto de medida nula. Pero para la investigación teórica la integral de Lebesgue introduce importantes simplificaciones. Los nuevos teoremas se apoyan en la aditividad numerable de la medida de Lebesgue, en contraste con la aditividad finita del contenido de Jordán.

Para ilustrar la simplicidad de los teoremas utilizando la integral de Lebesgue, veamos un resultado demostrado por Lebesgue mismo en su tesis. Supongamos que $u_1(x)$, $u_2(x)$, ... son funciones sumables sobre un conjunto medible E y que $\sum u_n(x)$ converge a $f(x)$; entonces $f(x)$ es medible. Si además

$$s_n(x) = \sum_{i=1}^n u_i(x)$$

está acotada uniformemente (es decir, si $|s_n(x)| \leq B$ para todo x en E y todo n), entonces es un teorema que $f(x)$ es integrable en el sentido de Lebesgue sobre $[a, b]$ y

$$\int_a^b f(x) dx = \lim_{n \rightarrow \infty} \int_a^b s_n(x) dx$$

Si hubiéramos estado trabajando con la integral de Riemann, habríamos necesitado la hipótesis adicional de que la suma de la serie fuera integrable; este caso, para la integral de Riemann, es un teorema debido a Cesare Arzelá (1847-1912).³⁷⁸ Lebesgue hizo de este teorema la piedra angular de la exposición de su teoría y de sus *Leçons sur l'intégration*.

La integral de Lebesgue resulta especialmente útil en la teoría de series de Fourier, a la que Lebesgue mismo hizo importantes contribuciones.³⁷⁹ Según Riemann, los coeficientes de Fourier a_n y b_n de una función acotada e integrable tienden a cero cuando n tiende a infinito. La generalización debida a Lebesgue afirma que

$$\lim_{n \rightarrow \infty} \int_a^b f(x) \begin{cases} \text{sen } nx \\ \text{cos } nx \end{cases} dx = 0$$

donde ahora $f(x)$ es cualquier función, acotada o no, que sea integrable en el sentido de Lebesgue. Este resultado se suele denominar hoy lema de Riemann-Lebesgue.

En el mismo artículo de 1903 demuestra Lebesgue que si f es una función acotada representada por una serie trigonométrica, es decir,

$$f(x) = \frac{a_0}{2} + \sum_1^{\infty} (a_n \cos nx + b_n \text{sen } nx)$$

entonces los a , y los b son los coeficientes de Fourier. En 1905³⁸⁰ dio

Lebesgue una nueva condición suficiente para la convergencia de la serie de Fourier a la función $f(x)$, que incluía todas las condiciones conocidas previamente.

Lebesgue demostró también (en sus *Leçons sur les séries trigonométriques*, 1906, p. 102) que la posibilidad de integrar término a término una serie de Fourier no depende de la convergencia uniforme de la serie a la función $f(x)$ misma. Lo que ocurre es que

$$\int_{-\pi}^x f(x) dx = a_0(x + \pi) + \sum_1^{\infty} \frac{1}{n} (a_n \operatorname{sen} nx + b_n (\cos n\pi - \cos nx))$$

donde x es un punto cualquiera del intervalo $(-\pi, \pi)$ para cualquier $f(x)$ integrable en el sentido de Lebesgue, converja o no la serie original para $f(x)$. Y la nueva serie converge uniformemente al miembro de la izquierda de la ecuación sobre el intervalo $(-\pi, \pi)$

Además, el teorema de Parseval, en el sentido de que

$$\frac{1}{\pi} \int_{-\pi}^{\pi} [f(x)]^2 dx = 2a_0^2 + \sum_1^{\infty} (a_n^2 + b_n^2)$$

se verifica para cualquier $f(x)$ cuyo cuadrado sea integrable sobre (*Leçons*, 1906, p. 100). Pierre Fatou (1878-1929) demostraría³⁸¹ que

$$\frac{1}{\pi} \int_{-\pi}^{\pi} f(x)g(x)dx = 2a_0\alpha_0 + \sum_1^{\infty} (a_n\alpha_n + b_n\beta_n)$$

donde a_n , b_n y α_n , β_n son los correspondientes coeficientes de Fourier para $f(x)$ y $g(x)$, cuyos cuadrados son integrables para Lebesgue sobre $[-\pi, \pi]$. A pesar de estos importantes progresos en la teoría de series de Fourier, no se conoce ninguna propiedad de una $f(x)$ integrable en el sentido de Lebesgue en $[-\pi, \pi]$ que sea condición necesaria y suficiente para la convergencia de su serie de Fourier.

Lebesgue dedicó la mayor parte de sus esfuerzos a la conexión entre las nociones de integral y de función primitiva (o integral indefinida). Cuando Riemann introdujo su generalización de la integral, se planteó ya el problema de si la correspondencia entre integral definida y función primitiva, válida para las funciones continuas, se seguía cumpliendo o no en el caso más general. Ahora bien, es posible dar ejemplos de funciones / integrables en el sentido de Riemann y tales que $\int_a^b f(t) dt$ no tiene derivada (ni siquiera derivadas laterales, por la derecha o por la izquierda) en algunos puntos. Recíprocamente, Volterra demostró en 1881³⁸² que una función $F(x)$ puede tener una derivada acotada en un intervalo I que no sea integrable en el sentido de Riemann sobre dicho intervalo. Un sutil análisis del problema permitió a Lebesgue demostrar que si $f(x)$ es integrable en su sentido sobre $[a,b]$, entonces $F(x) = \int_a^b f(t) dt$ tiene una derivada igual a $f(x)$ casi por doquier, es decir, salvo sobre un conjunto de medida nula (*Leçons sur l'intégration*). Y recíprocamente, si una función $g(x)$ es

diferenciable sobre $[a,b]$ y si su derivada $g' = f$ está acotada, entonces f es integrable en el sentido de Lebesgue y se verifica la fórmula $g(x) - g(a) = \int_a^b f(t)$. Sin embargo, como advierte el mismo Lebesgue, la situación es mucho más complicada si g' no está acotada. En este caso g' no es necesariamente integrable, y el primer problema es el de caracterizar a las funciones g para las que g' existe casi por doquier y es integrable. Limitándose al caso en que uno de los cuatro números derivados³⁸³ de g sea siempre finito, demuestra Lebesgue que g tiene que ser necesariamente una función de variación acotada (cap. 40, sec. 6). Por último, Lebesgue demostró (en su libro de 1904) el recíproco. Una función g de variación acotada admite una derivada g' casi por doquier, la cual es integrable. Sin embargo, no se tiene necesariamente que

$$g(x) - g(a) = \int_a^x g'(t) dt \quad (1)$$

sino que la diferencia entre los dos miembros de esta ecuación es una función de variación acotada no constante pero con derivada cero casi por doquier. En cuanto a las funciones de variación acotada g para las que se verifica, tienen la siguiente propiedad: la variación total de g en un conjunto abierto U (es decir, la suma de las variaciones totales de g en cada una de las componentes conexas de U) tiende a cero con la medida de U . A estas funciones las llamó absolutamente continuas Giuseppe Vitali (1875-1932), que además las estudió con detalle.

La obra de Lebesgue también hizo progresar la teoría de las integrales múltiples. Con su definición de integral doble, se amplía el dominio de las funciones para las que la integral doble se puede calcular por integración iterada. Lebesgue ya dio un resultado de este tipo en su tesis de 1902, pero fue mejorado por Guido Fubini (1879-1943)³⁸⁴: si $f(x,y)$ es sumable sobre el conjunto medible G , entonces

- a. $f(x,y)$ como función de x y como función de y es sumable para casi todo y y casi todo x respectivamente;
- b. el conjunto de los puntos (x_0,y_0) para los que o bien $f(x,y_0)$ o bien $f(x_0,y)$ no es sumable, tiene medida cero;
- c.
$$\iint_G f(x,y)dG = \int dy(\int f(x,y) dx) = \int dx(\int f(x,y) dy)$$
- d. donde las integrales exteriores están tomadas sobre los conjuntos de puntos y (respectivamente x) para los que $f(x,y)$ como función de x (respectivamente, como función de y) son sumables.

Finalmente, en 1910³⁸⁵ llegó Lebesgue a resultados sobre integrales múltiples que generalizaban los de las derivadas para integrales simples. A cada función f integrable en toda región compacta de R^n le asoció la función de conjunto (en oposición a las funciones de variables numéricas) $F(E) = \int_E f(x) dx$ (donde x representa aquí n coordenadas) definida para cada dominio de integración E de R^n . Este concepto generaliza la integral indefinida. Lebesgue observó que la función G posee las dos propiedades siguientes:

1. Es completamente aditiva; es decir, $F(E_1 \cup E_n) = \Sigma F(E_n)$ donde los

E_n son conjuntos medibles disjuntos dos a dos.

2. Es absolutamente continua en el sentido de que $F(E)$ tiende a cero con la medida de E .

La parte esencial de este artículo de Lebesgue consistía en demostrar el recíproco de esta proposición, es decir, definir una derivada de $F(E)$ en un punto P del espacio n -dimensional. Lebesgue llegó al siguiente teorema; si $F(E)$ es absolutamente continua y aditiva, entonces tiene una derivada finita casi por doquier, y F es la integral indefinida de la función sumable que es igual a la derivada de F donde ésta exista y sea finita, y arbitraria en los puntos restantes.

La herramienta principal de la demostración es un teorema de recubrimientos debido a Vitali,³⁸⁶ que sigue siendo fundamental en esta área de la teoría de integración. Pero Lebesgue no se paró aquí, sino que indicó la posibilidad de generalizar la noción de función de variación acotada considerando funciones $F(E)$, donde E es un conjunto medible, completamente aditivas y tales que $\sum_n |F(E_n)|$ permanezca acotada para toda partición numerable de E en subconjuntos medibles E_n . Se podrían citar muchos otros teoremas del cálculo integral que han sido generalizados utilizando el concepto de integral de Lebesgue.

La obra de Lebesgue, una de las grandes contribuciones a la matemática de este siglo, fue aceptada pero, como de costumbre, no sin cierta resistencia. Ya hemos hablado (cap. 40, sec. 7) de las objeciones de Hermite a las funciones sin derivada. El mismo

Hermite trató de evitar que Lebesgue publicara una «Nota sobre las superficies no regladas aplicables sobre el plano»³⁸⁷, en la que Lebesgue estudiaba ciertas superficies no diferenciables. Muchos años más tarde contaba Lebesgue en su *Notice* (p. 14, ver bibliografía del cap.),

Darboux había dedicado su Mémoire de 1875 a la integración y a las funciones sin derivadas, por tanto no experimentaba el mismo horror que Hermite. Sin embargo, dudo de que nunca llegara a perdonarme del todo mi «Nota sobre las superficies aplicables». Seguramente pensaba que los que se enfrascan en estos estudios pierden el tiempo en vez de dedicarse a una investigación útil.

y dice también Lebesgue (*Notice*, p. 13),

Para muchos matemáticos me convertí en el hombre de las funciones sin derivadas, a pesar de que nunca en la vida me dediqué por completo al estudio o consideración de tales funciones. Y como el temor y el horror que mostraba Hermite lo compartía casi todo el mundo, siempre que intentaba tomar parte en una discusión matemática, aparecía algún analista que decía, «Esto no le interesará a usted; estamos discutiendo de funciones con derivadas». O algún geómetra diciéndolo en su lenguaje: «Estamos discutiendo de superficies que tienen planos tangentes.»

5. Generalizaciones

Ya hemos indicado las ventajas de la integración de Lebesgue tanto para generalizar viejos resultados como para formular nuevos y elegantes teoremas sobre series. En los capítulos siguientes nos encontraremos con más aplicaciones de las ideas de Lebesgue. Los desarrollos más inmediatos dentro de la teoría de funciones consistieron en diversas extensiones de la idea de integral. De ellas sólo mencionaremos una debida a Johann Radon (1887-1956), que incluye tanto la integral de Stieltjes como la de Lebesgue y que, de hecho, se conoce con el nombre de integral de Lebesgue-Stieltjes. Las generalizaciones consisten no sólo en conceptos de integral distintos y más amplios sobre conjuntos de puntos del espacio euclídeo n -dimensional, sino sobre dominios de espacios más generales, tales como espacios de funciones, por ejemplo. Las aplicaciones de estos conceptos más generales pueden encontrarse hoy en la teoría de probabilidades, en teoría espectral, en teoría ergódica y en el análisis armónico (o análisis de Fourier generalizado).

Bibliografía

- Borel, Emile: *Notice sur les travaux scientifiques de M. Emile Borel*, Gauthier-Villars, 2.^a ed., 1921.
- Bourbaki, Nicolás: *Elementos de historia de las matemáticas*, Madrid, Alianza, 1976.
- Collingwood, E. F.: «Emile Borel», *Jour. London Math. Soc.*, 34, 1959, 488-512.
- Fréchet, M.: «La vie et l'OEuvre d'Emile Borel», *L'Enseignement*

Mathématique, (2), 11, 1965, 1-94.

- Hawkins, T. W.; Jr.: *Lebesgue's theory of integration: its origins and development*, University of Wisconsin Press, 1970, caps. 4-6.
- Hildebrandt, T. H.: «On intégrais related to and extensions of the Lebesgue integral», *Amer. Math. Soc. Bull.*, 24, 1918, 113-177.
- Jordán, Camille: *OEuvres*, 4 volúmenes, París, Gauthier-Villars, 1961-1964. Lebesgue, Henri: *Measure and the integral*, Holden-Day, 1966, pp. 176-194. Traducción del francés *Le mesure des grandeurs. -Notice sur les travaux scientifiques de M. Henri Lebesgue*, Edouard Privat, 1922. *-Leçons sur l'intégration et la recherche des fonctions primitives*, Gauthier-Villars, 1904; 2.^a ed., 1928.
- McShane, E. J.: «Intégrais devised for special purposes»; *Amer. Math. Soc. Bull.*, 69, 1963, 597-627.
- Pesin, Ivan M.: *Classical and modern integration theories*, Academic Press, 1970.
- Plancherel, Michel: «Le développement de la théorie des séries trigonométriques dans le dernier quart de siècle», *L'Enseignement Mathématique*, 24, 1924-1925, 19-58.
- Riesz, F.: «L'évolution de la notion d'intégrale depuis Lebesgue», *Annales de l'Institut Fourier*, 1, 1949, 29-42.³⁸⁸

Capítulo 45

Ecuaciones integrales

*La naturaleza no se ve
desconcertada por las dificultades
del análisis.*

Augustin Fresnel

Contenido:

- 1. Introducción*
 - 2. Los comienzos de una teoría general*
 - 3. La obra de Hilbert*
 - 4. Los sucesores inmediatos de Hilbert*
 - 5. Generalizaciones de la teoría*
- Bibliografía*

1. Introducción

Una ecuación integral no es más que una ecuación en la que aparece una función incógnita bajo un signo integral, y el problema de resolver dicha ecuación consiste en determinar esa función. Como vamos a ver enseguida, algunos problemas de la física matemática conducen directamente a ecuaciones integrales, mientras que otros problemas que conducen en primer lugar a ecuaciones diferenciales ordinarias o en derivadas parciales, pueden manejarse a veces con más comodidad convirtiéndolas previamente

en ecuaciones integrales. Al principio, a la resolución de ecuaciones integrales se le llamó el problema de la inversión de integrales, y el término ecuación integral lo introdujo Du Bois-Reymond.³⁸⁹

Al igual que en otras ramas de la matemática, aparecieron problemas aislados en los que se presentaban ecuaciones integrales mucho antes de que el tema adquiriese su *status* y su metodología propios. Así, por ejemplo, Laplace estudió en 1782³⁹⁰ la ecuación integral en $g(t)$ dada por

$$f(x) = \int_{-\infty}^{\infty} e^{-xt} g(t) dt \quad (1)$$

Tal como aparece escrita, la ecuación (1) recibe hoy el nombre de transformación de Laplace de $g(t)$. Poisson³⁹¹ descubrió la expresión de $g(t)$, a saber,

$$g(t) = \frac{1}{2\pi i} \int_{\alpha-i\infty}^{\alpha+i\infty} e^{xt} f(x) dx$$

para a suficientemente grande. Otro de los resultados notables que en realidad pertenecen a la historia de las ecuaciones integrales proviene del famoso artículo de Fourier de 1811 sobre la propagación del calor (cap. 28, párr. 3). Allí nos encontramos con la ecuación

$$f(x) = \int_0^{\infty} \cos(xt)u(t) dt$$

y la fórmula de inversión

$$u(t) = \frac{2}{\pi} \int_0^{\infty} \cos(xt)f(x) dx$$

El primer caso de uso directo y consciente de una ecuación integral y su resolución se remontan a Abel.

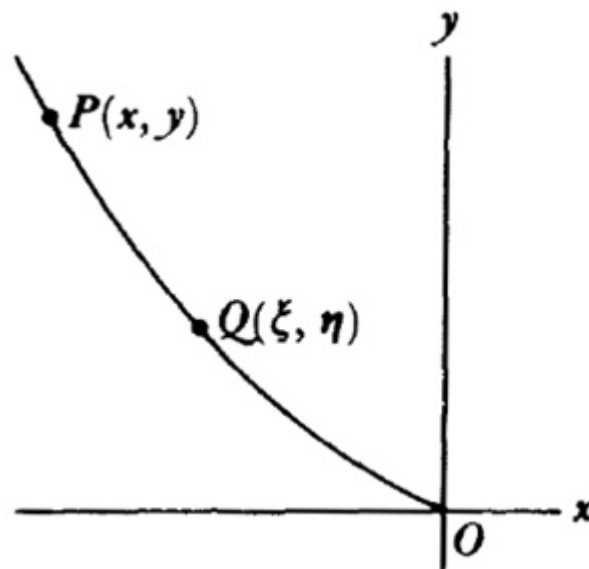


Figura 45.1

En dos de sus primeros artículos publicados, el primero en 1823³⁹² en una oscura revista y el segundo en el *Journal für Mathematik*,³⁹³ considera Abel el siguiente problema de mecánica: una partícula

material cae desde P deslizándose por una curva lisa (fig. 45.1) hasta el punto O ; la curva está situada en un plano vertical.

La velocidad adquirida en O es independiente de la forma de la curva, pero el tiempo empleado en deslizar desde P hasta O no lo es. Si llamamos a las coordenadas de un punto genérico Q entre P y O , y s es el arco PQ , entonces la velocidad de la partícula en Q vendrá dada por

$$\frac{ds}{dt} = -\sqrt{2g(x-\xi)}$$

donde g es la constante gravitatoria. Por tanto

$$t = \frac{-1}{\sqrt{2g}} \int_P^Q \frac{ds}{\sqrt{x-\xi}}$$

Ahora bien, s puede expresarse en términos de ξ . Supongamos que s es $v(\xi)$. Entonces el tiempo total de descenso T desde P hasta O vendrá dado por

$$T(x) = \frac{1}{\sqrt{2g}} \int_0^x \frac{v'(\xi)d\xi}{\sqrt{x-\xi}}$$

Claramente, el tiempo T depende de x para cualquier curva. El problema que se propone Abel es, dado T como función de x , hallar

$v(\xi)$. Si introducimos

$$f(x) = \sqrt{2g} T(x)$$

el problema se convierte en el de determinar v en la ecuación

$$f(x) = \int_0^x \frac{v'(\xi)}{\sqrt{x-\xi}} d\xi$$

Abel obtiene la solución

$$v(\xi) = \int_0^\xi \frac{f(x) dx}{\sqrt{\xi-x}}$$

Los métodos que utiliza (da dos) son particulares y no nos interesan ahora.

En realidad, de lo que trata Abel es de resolver el problema más general

$$f(x) = \int_0^x \frac{u(\xi) d\xi}{(x-\xi)^\lambda} \quad 0 < \lambda < 1 \quad (2)$$

y obtiene

$$u(z) = \frac{\text{sen } \lambda\pi}{\pi} \frac{d}{dz} \int_a^z \frac{f(x)dx}{(z-x)^{1-\lambda}}$$

Liouville, que trabajaba independientemente de Abel, resolvió algunas ecuaciones integrales concretas desde 1832³⁹⁴ en adelante. Un paso más importante dado por Liouville³⁹⁵ fue el de mostrar cómo se puede obtener la solución de ciertas ecuaciones diferenciales resolviendo ecuaciones integrales. La ecuación diferencial que había que resolver era la

$$y'' + [\rho^2 - \sigma(x)]y = 0 \quad (3)$$

sobre el intervalo $a \leq x \leq b$, con ρ como parámetro. Sea $u(x)$ la solución particular que satisface las condiciones iniciales

$$u(a) = 1 \quad u'(a) = 0. \quad (4)$$

Esta función también será solución de la ecuación no homogénea

$$y'' + \rho^2 y = \sigma(x)u(x).$$

Entonces, por un resultado básico de ecuaciones diferenciales

$$u(x) = \cos \rho(x-a) + \frac{1}{\rho} \int_a^x \sigma(\xi) \text{sen } \rho(x-\xi) u(\xi) d\xi \quad (5)$$

Así pues, si podemos resolver esta ecuación integral habremos obtenido la solución de la ecuación diferencial (3) que satisface las condiciones iniciales (4).

Liouville obtuvo la solución por un método de sustituciones sucesivas más tarde atribuido a Cari G. Neumann, cuya obra *Untersuchungen über das logarithmische und Newton'sche Potential* (1877) apareció treinta años después. No explicaremos el método de Liouville porque es prácticamente idéntico al de Volterra, del que hablaremos más adelante.

Las ecuaciones integrales estudiadas por Abel y por Liouville son de dos tipos básicos; la de Abel es de la forma

$$f(x) = \int_a^x K(x, \xi)u(\xi)d\xi \quad (6)$$

y la de Liouville de la forma

$$u(x) = f(x) + \int_a^x K(x, \xi)u(\xi)d\xi \quad (7)$$

En ambas $f(x)$ y $K(x, \xi)$ son funciones conocidas y $u(\xi)$ es la función que hay que determinar. La terminología utilizada hoy, que fue introducida por Hilbert, se refiere a estas ecuaciones llamándolas de primero y segundo tipo respectivamente, y $K(x, \xi)$ se denomina el

núcleo de la ecuación. Como hemos dicho ya, también reciben el nombre de ecuaciones de Volterra, mientras que cuando el límite superior de integración es un número fijo b , se las llama ecuaciones de Fredholm. En realidad, las ecuaciones de Volterra son a su vez casos particulares, respectivamente, de las de Fredholm, porque siempre se puede tomar $K(x, \xi) = 0$ para $\xi > x$ y considerar las ecuaciones de Volterra como ecuaciones de Fredholm. El caso especial de ecuación del segundo tipo en que $f(x) = 0$ recibe el nombre de ecuación homogénea.

A mediados del siglo XIX el principal interés en las ecuaciones integrales se centraba en torno a la resolución de los problemas de condiciones de contorno asociados a la ecuación del potencial

$$\Delta u = u_{xx} + u_{yy} = 0. \quad (8)$$

La ecuación se verifica en una región plana dada, limitada por una curva C . Si el valor en la frontera de u es alguna función $f(s)$ dada en términos de la longitud de arco s a lo largo de C , entonces una solución de este problema del potencial viene dada por

$$u(x, y) = \frac{1}{2\pi} \int_C \rho(s) \log \frac{1}{r(s; x, y)} ds$$

donde $r(s; x, y)$ es la distancia de un punto s de C a un punto cualquier (x, y) del interior o de la frontera, y $\rho(s)$ es una función

desconocida que verifica, para $s = (x, y)$ sobre C

$$f(s) = \frac{1}{2\pi} \int_C \rho(t) \log \frac{1}{r(t; x, y)} dt \quad (9)$$

Esta es una ecuación integral de primer tipo para $\rho(t)$. De manera alternativa, si uno toma como solución de (8) con la misma condición de frontera

$$v(x, y) = \frac{1}{2\pi} \int_C \phi(s) \frac{\partial}{\partial n} \left(\log \frac{1}{r(s; x, y)} \right) ds$$

donde $\partial/\partial n$ representa la derivada normal en la frontera, entonces $\phi(s)$ tiene que satisfacer la ecuación integral

$$f(s) = \frac{1}{2} \phi(s) = \frac{1}{2\pi} \int_C \phi(s) \frac{\partial}{\partial n} \left(\log \frac{1}{r(t; x, y)} \right) dt \quad (10)$$

que es una ecuación integral del segundo tipo. Neumann resolvió estas ecuaciones para regiones convexas en sus *Untersuchungen* y en publicaciones posteriores.

Otro problema de ecuaciones en derivadas parciales fue atacado por medio de ecuaciones integrales; la ecuación

$$\Delta u + \lambda u = f(x, y) \quad (11)$$

aparece en el estudio de movimientos de ondas cuando la dependencia temporal de la correspondiente ecuación hiperbólica

$$\Delta u - \frac{1}{c^2} u_{tt} = f(x, y)$$

que suele tomarse $e^{-i\omega t}$, se elimina. Se sabía (cap. 28, sec. 8) que el caso homogéneo de la ecuación (11) sometido a las condiciones de contorno tiene soluciones no triviales sólo para un conjunto discreto de valores de λ , llamados autovalores o valores característicos. En 1894³⁹⁶ estudió Poincaré el caso no homogéneo de (11) con λ complejo, y logró construir una función meromorfa en λ que representaba la solución única de (11) para cualquier λ que no fuese un autovalor, y cuyos residuos producían autofunciones para el caso homogéneo, es decir, cuando $f = 0$.

Sobre la base de estos resultados, consideró Poincaré en 1896³⁹⁷ la ecuación

$$u(x) + \lambda \int_a^b K(x, y)u(y) dy = f(x)$$

que obtuvo de (11), afirmando que la solución sería una función meromorfa en λ . Este resultado lo demostraría Fredholm en un artículo que vamos a comentar en breve.

La transformación de ecuaciones diferenciales en ecuaciones

integrales, que hemos ilustrado con los anteriores ejemplos, se convirtió en una técnica importante para resolver problemas de contorno en ecuaciones diferenciales ordinarias y en derivadas parciales, y constituyó el más fuerte impulso para el estudio de las ecuaciones integrales en sí mismas.

2. Los comienzos de una teoría general

Vito Volterra (1860-1940), sucesor de Beltrami como profesor de física matemática en Roma, fue el primero de los fundadores de una teoría general para las ecuaciones integrales. Escribió diversos artículos sobre el tema desde 1884, de los cuales los más importantes datan de 1896 y 1897.³⁹⁸ Volterra ideó un método para resolver ecuaciones integrales de segundo tipo,

$$f(s) = \phi(s) + \int_a^b K(s,t)\phi(t)dt \quad (12)$$

donde $\phi(s)$ es la incógnita y $K(s,t) = 0$ para $t > s$. Volterra escribe esta ecuación de la forma

$$f(s) = \phi(s) + \int_a^s K(s,t)\phi(t)dt$$

y su método de resolución consiste en definir

$$\begin{aligned}
 f_1(s) &= - \int_a^b K(s,t) f(t) dt \\
 &\dots \\
 f_n(s) &= - \int_a^b K(s,t) f_{n-1}(t) dt \\
 &\dots
 \end{aligned}
 \tag{13}$$

y tomar $\phi(s)$ como

$$\phi(s) = f(s) + \sum_{p=1}^{\infty} f_p(s) \tag{14}$$

Para su núcleo $K(s,t)$ pudo demostrar Volterra la convergencia de la serie (14), y si uno sustituye (14) en (12) puede ver que es una solución. Esta sustitución da

$$\phi(s) = f(s) + \int_a^b \bar{K}(s,t) f(t) dt \tag{15}$$

que puede escribirse en la forma

$$\phi(s) = f(s) - \int_a^b K(s,t) f(t) dt + \int_a^b \int_a^b K(s,r) K(r,t) f(t) dr dt + \dots$$

donde ahora el núcleo \bar{K} (llamado más tarde núcleo resolvente o

simplemente resolvente por Hilbert) es

$$\bar{K}(s,t) = -K(s,t) + \int_a^b K(s,r)K(r,t)dr - \\ - \int_a^b \int_a^b K(s,r)K(r,w)K(w,t) drdw + \dots$$

La ecuación (15) es la representación obtenida anteriormente por Liouville para una ecuación integral particular, y atribuida a Neumann. Volterra resolvió también ecuaciones integrales de primer tipo $f(s) = \int_a^b K(x,s) \phi(x) dx$ reduciéndolas a ecuaciones del segundo tipo.

En 1896 observó Volterra que una ecuación integral del primer tipo venía a ser una forma límite de un sistema de n ecuaciones lineales con n incógnitas, cuando n tiende a infinito. Erik Ivar Fredholm (1866-1927), profesor de matemáticas en Estocolmo, interesado en la resolución del problema de Dirichlet, recogió esta idea en 1900³⁹⁹ y la utilizó para resolver ecuaciones integrales del segundo tipo, es decir, ecuaciones de la forma (12), sin la restricción sobre $K(s,t)$ sin embargo.

Escribiremos la ecuación que estudió Fredholm en la forma

$$u(x) = f(x) + \lambda \int_a^b K(x,\xi)u(\xi)d\xi \quad (16)$$

aunque el parámetro λ no aparecía explícitamente en su obra; sin embargo, es más fácil de entender lo que él hizo, a la luz de trabajos posteriores, si lo hacemos explícito. Para ajustarnos fielmente a las fórmulas de Fredholm bastaría hacer $\lambda = 1$ o bien suponerlo implícitamente incluido en K .

Fredholm divide el intervalo en $x [a,b]$ en n partes iguales por medio de los puntos

$$a, x_1 = a + \delta, x_2 = a + 2\delta, \dots, x_n = a + n\delta = b$$

y a continuación reemplaza la integral que figura en (16) por la suma

$$u_n(x) = f(x) + \sum_{j=1}^n \lambda K(x, x_j) u_n(x_j) \delta \quad (17)$$

Ahora bien, la ecuación (17) se supone que se verifica para todos los valores de x del intervalo $[a,b]$. Por tanto se debe verificar para $x = x_1, x_2, \dots, x_n$. Esto conduce al sistema de n ecuaciones

$$-\sum_{j=1}^n \lambda K(x, x_j) u_n(x_j) \delta + u_n(x_i) = f(x_i) + \quad i = 1, 2, \dots, n \quad (18)$$

Este es un sistema lineal no homogéneo de n ecuaciones para

determinar las n incógnitas $u_n(x_1), u_n(x_2), \dots$

En la teoría de sistemas de ecuaciones lineales se conocía ya el siguiente resultado. Si tenemos una matriz

$$S_n = \begin{vmatrix} 1 + a_{11} & a_{12} & a_{13} & \dots & a_{1n} \\ a_{21} & 1 + a_{22} & a_{23} & \dots & a_{2n} \\ & & \dots & & \\ a_{n1} & a_{n2} & a_{n3} & \dots & 1 + a_{nn} \end{vmatrix}$$

entonces el determinante $D(n)$ de S admite el siguiente desarrollo

$$D(n) = 1 + \frac{1}{1!} \sum_{r_1} a_{r_1 r_1} + \frac{1}{2!} \sum_{r_1, r_2} \begin{vmatrix} a_{r_1 r_1} & a_{r_1 r_2} \\ a_{r_2 r_1} & a_{r_2 r_2} \end{vmatrix} +$$

$$+ \frac{1}{n!} \sum_{r_1, \dots, r_n} \begin{vmatrix} a_{r_1 r_1} & a_{r_1 r_2} & \dots & a_{r_1 r_n} \\ a_{r_2 r_1} & a_{r_2 r_2} & \dots & a_{r_2 r_n} \\ & \dots & & \\ a_{r_n r_1} & a_{r_n r_2} & \dots & a_{r_n r_n} \end{vmatrix}$$

donde r_1, r_2, \dots, r_n recorren independientemente todos los valores de 1 a n . Desarrollando el determinante de los coeficientes en (18) y haciendo tender después n a infinito, obtuvo Fredholm el determinante⁴⁰⁰

$$\begin{aligned}
 D(\lambda) = & 1 - \lambda \int_a^b K(\xi_1, \xi_1) d\xi_1 + \\
 & + \frac{\lambda^2}{2!} \int_a^b \int_a^b \begin{vmatrix} K(\xi_1, \xi_1) & K(\xi_1, \xi_2) \\ K(\xi_2, \xi_1) & K(\xi_2, \xi_2) \end{vmatrix} d\xi_1 d\xi_2 - \dots
 \end{aligned}
 \tag{19}$$

al que llamó determinante de la ecuación (16) o del núcleo K . Análogamente, considerando los adjuntos de los elementos de la fila μ y de la columna ν del determinante de los coeficientes en (18) y haciendo tender n al infinito, obtuvo Fredholm la función

$$\begin{aligned}
 D(x, y, \lambda) = & \lambda K(x, y) - \lambda^2 \int_a^b \begin{vmatrix} K(x, y) & K(x, \xi_1) \\ K(\xi_1, y) & K(\xi_1, \xi_1) \end{vmatrix} d\xi_1 + \\
 & + \frac{\lambda^3}{2} \int_a^b \int_a^b \begin{vmatrix} K(x, y) & K(x, \xi_1) & K(x, \xi_2) \\ K(\xi_1, y) & K(\xi_1, \xi_1) & K(\xi_1, \xi_2) \\ K(\xi_2, y) & K(\xi_2, \xi_1) & K(\xi_2, \xi_2) \end{vmatrix} d\xi_1 d\xi_2 - \dots
 \end{aligned}
 \tag{20}$$

Fredholm llamó a esta función $D(x, y, \lambda)$ el primer menor del núcleo K porque juega un papel análogo al de los primeros menores en el caso de n ecuaciones lineales con n incógnitas. También llamó a los ceros de la función analítica entera $D(\lambda)$ las raíces de $K(x, y)$. Aplicando la regla de Cramer al sistema de ecuaciones lineales (18) y haciendo tender n a infinito, fue como dedujo Fredholm la forma de la solución de (16). A continuación demostró que era correcta sustituyéndola directamente, y pudo al fin enunciar el siguiente

resultado: si λ no es una de las raíces de K , es decir, si $D(\lambda)$ es distinto de cero, entonces (16) tiene una y sólo una solución (continua), a saber,

$$u(x, \lambda) = f(x) + \int_a^b \frac{D(x, y, \lambda)}{D(\lambda)} f(y) dy \quad (21)$$

Por otra parte, si λ es una raíz de $K(x, y)$, entonces (16) o bien no tiene ninguna solución continua o tiene infinitas.

Fredholm obtuvo además resultados acerca de la relación entre la ecuación homogénea

$$u(x) = \lambda \int_a^b K(x, \xi) u(\xi) d\xi \quad (22)$$

y la no homogénea (16). Es casi evidente a partir de (21) que cuando λ no sea raíz de K la única solución continua de (22) es $u = 0$; por tanto, se dedicó a estudiar el caso en que λ es raíz de K . Sea $\lambda = \lambda_1$ una tal raíz. Entonces (22) tiene las infinitas soluciones

$$c_1 u_1(x) + c_2 u_2(x) + \dots + c_n u_n(x),$$

donde las c son constantes arbitrarias; las u_1, u_2, \dots, u_n , llamadas soluciones principales, son linealmente independientes, y n depende de λ_i . El número n recibe el nombre de índice de λ_i (que no es la

multiplicidad de λ_i , como cero de $D(\lambda)$). Fredholm pudo calcular el índice de cualquier raíz λ_i y demostrar que el índice nunca puede exceder de la multiplicidad (que siempre es finita). Las raíces de $D(\lambda) = 0$ reciben el nombre de valores característicos de $K(x,y)$ y al conjunto de las raíces se le llama su espectro. Las soluciones de (22) correspondientes a los valores característicos reciben el nombre de auto- funciones o funciones características.

Ahora estaba ya Fredholm en condiciones de establecer lo que ha sido llamado desde entonces el teorema de la alternativa de Fredholm. En el caso en que λ sea un valor característico de K , no solamente la ecuación integral (22) tiene n soluciones independientes, sino que la ecuación asociada o adjunta, que tiene el núcleo transpuesto, es decir,

$$u(x) = \lambda \int_a^b K(\xi, x)u(\xi)d\xi \quad (22)$$

también tiene n soluciones $\psi_1(x), \psi_2(x), \dots, \psi_n(x)$ para el mismo valor característico, y entonces la ecuación no homogénea (16) es soluble si y sólo si

$$\int_a^b f(x)\psi_i(x)dx \quad i = 1, 2, \dots, n \quad (23)$$

Estos últimos resultados siguen un paralelismo muy estrecho con la

teoría de sistemas de ecuaciones lineales, homogéneos y no homogéneos.

3. La obra de Hilbert

El interés de Hilbert por las ecuaciones integrales se despertó oyendo una conferencia dada por Erik Holmgren (n. 1872) en 1901 sobre la obra de Fredholm en este campo, que ha habido sido publicada en Suecia. David Hilbert (1862-1943), uno de los más importantes matemáticos de este siglo, que había realizado ya una obra impresionante sobre números algebraicos, teoría de invariantes y los fundamentos de la geometría, volvía ahora su atención a las ecuaciones integrales. El mismo nos dice que una investigación sobre el tema le mostró que era importante para la teoría de integrales definidas, para el desarrollo de funciones arbitrarias en series (de funciones especiales o de funciones trigonométricas), para la teoría de ecuaciones diferenciales lineales, para la teoría del potencial y para el cálculo de variaciones. Desde 1904 a 1910 escribió una serie de seis artículos en el *Nachrichten von der Königlichlichen Gesellschaft der Wissenschaften zu Göttingen*, artículos recogidos posteriormente en su libro *Grundzüge einer allgemeinen Theorie der linearen Integralgleichungen* (1912). En la última parte de esta obra aplica las ecuaciones integrales a problemas de la física matemática.

Fredholm había utilizado la analogía existente entre las ecuaciones integrales y las ecuaciones algebraicas lineales, pero en vez de llevar a cabo el proceso de paso al límite para las infinitas ecuaciones

algebraicas resultantes, se limitó a escribir audazmente los determinantes que salían de esa analogía y a comprobar que resolvían la ecuación integral. Lo primero que hizo Hilbert entonces fue llevar a cabo rigurosamente el paso al límite sobre el sistema finito de ecuaciones lineales.

Hilbert parte de la ecuación integral

$$f(s) = \xi(s) - \lambda \int_0^1 K(s,t) \phi(t) dt \quad (24)$$

$K(s,t)$ es continua. El parámetro λ aparece explícitamente y juega de hecho un papel importante en la teoría que sigue. Al igual que Fredholm, Hilbert divide el intervalo $[0,1]$ en n partes, de manera que p/n y q/n ($p, q = 1, 2, \dots, n$) representan puntos del intervalo $[0,1]$. Sean

$$K_{pq} = K\left(\frac{p}{n}, \frac{q}{n}\right) \quad f_p = f\left(\frac{p}{n}\right) \quad \phi_q = \phi\left(\frac{q}{n}\right)$$

Entonces, de (24) obtenemos el sistema de n ecuaciones con n incógnitas $\phi_1, \phi_2, \dots, \phi_n$, es decir,

$$f_p = \phi_p - \lambda \sum_{q=1}^n K_{pq} \phi_q \quad p = 1, 2, \dots, n$$

Después de hacer un repaso de la teoría general de la resolución de

los sistemas finitos de n ecuaciones lineales con n incógnitas, Hilbert vuelve a considerar directamente la ecuación (24). Para el núcleo K de (24) los autovalores están definidos como los ceros de la serie de potencias

$$\delta(\lambda) = 1 + \sum_{n=1}^{\infty} (-1)^n d_n \lambda^n$$

donde los coeficientes d_n vienen dados por

$$d_n = \frac{1}{n!} \int_0^1 \dots \int_0^1 |\{K(s_i, s_j)\}| ds_1 \dots ds_n$$

Aquí $|\{A(s_i, s_j)\}|$ es el determinante de la matriz $n \times n$, $\{A(s_i, s_j)\}$, $i, j = 1, 2, \dots, n$, y los s , son valores de t en el intervalo $[0, 1]$. Para explicar el resultado principal obtenido por Hilbert necesitamos definir las siguientes funciones

$$\Delta_p(x, y) = \frac{1}{p!} \int_0^1 \dots \int_0^1 \begin{vmatrix} 0 & x(s_1) & \dots & x(s_p) \\ y(s_1) & K(s_1, s_2) & \dots & K(s_1, s_p) \\ \dots & \dots & \dots & \dots \\ y(s_p) & K(s_p, s_1) & \dots & K(s_p, s_p) \end{vmatrix} ds_1 \dots ds_p$$

donde $x(r)$ e $y(r)$ son funciones continuas arbitrarias cuando r recorre el intervalo $[0, 1]$, y

$$\Delta(\lambda; x, y) = \sum_{p=1}^{\infty} (-1)^p \Delta_p(x, y) \lambda^{p-1}$$

Hilbert define a continuación

$$\Delta^*(A; s, t) = \lambda \Delta(\lambda; x, y) - \delta(\lambda)$$

donde ahora $x(r) = K(s, r)$ e $y(r) = K(r, t)$, y demuestra que si K se define como

$$\bar{K}(s, t) = \frac{\Delta^*(\lambda; s, t)}{-\delta(\lambda)}$$

para valores de λ tales que $\delta(\lambda) \neq 0$, entonces

$$\begin{aligned} K(s, t) &= \bar{K}(s, t) - \lambda \int_0^1 \bar{K}(s, r) K(r, t) dr = \\ &= \bar{K}(s, t) - \lambda \int_0^1 K(s, r) \bar{K}(r, t) dr \end{aligned}$$

Por último, si se toma ϕ definida por

$$\phi(r) = f(r) + \lambda \int_0^1 \bar{K}(r, t) f(t) dt \quad (25)$$

(f) es una solución de (24). Las demostraciones de las diversas etapas de esta teoría requieren un cierto número de consideraciones de paso al límite sobre expresiones que aparecían en el tratamiento de Hilbert de los sistemas de ecuaciones lineales.

Hasta aquí Hilbert había demostrado que para cualquier núcleo continuo (no necesariamente simétrico) $K(s,t)$ y para cualquier valor de λ tal que $\delta(\lambda) \neq 0$, existe la función resolvente $\hat{K}(s,t)$ con la propiedad de que la función (25) es solución de la ecuación (24).

A continuación Hilbert supone que $K(s,t)$ es simétrico, lo que le permite hacer uso de propiedades de las matrices simétricas en el caso finito, y demuestra que los ceros de $\delta(\lambda)$, es decir, los autovalores del núcleo simétrico son reales. Entonces los ceros de $\delta(\lambda)$ se ordenan en valor absoluto creciente (para valores absolutos iguales el cero positivo se toma primero, y han de tenerse en cuenta también las multiplicidades de las raíces). Las autofunciones de la ecuación (24) vienen definidas ahora por

$$\phi^k(s) = \left(\frac{\lambda_k}{\Delta^*(\lambda_k; s^*, s^*)} \right)^{1/2} \Delta^*(\lambda_k; s, s^*)$$

donde s^* se toma de manera que $\Delta^*(\lambda_k; s^*, s^*) \neq 0$ y λ_k es un autovalor cualquiera de $K(s,t)$.

Las autofunciones asociadas a los distintos autovalores pueden tomarse formando un sistema ortonormal (ortogonal y normalizado),⁴⁰¹ y para cada autovalor λ_k y cada autofunción

correspondiente a λ_k

$$\phi^k(s) = \lambda_k \int_0^1 K(s,t)\phi^k(t)dt$$

Con estos resultados Hilbert puede demostrar ya lo que se conoce como teorema del eje principal generalizado para las formas cuadráticas simétricas. Sea

$$\sum_{p=1}^n \sum_{q=1}^n k_{pq} x_p x_q \quad (26)$$

una forma cuadrática n -dimensional en las n variables x_1, x_2, \dots, x_n . Podemos escribirla en la forma (Kx, x) , donde K es la matriz de los k_{pq} , x representa el vector (x_1, x_2, \dots, x_n) y (Kx, x) es el producto escalar de los vectores Kx y x . Supongamos que K tiene los n autovalores distintos $\lambda_1, \lambda_2, \dots, \lambda_n$; entonces, para cada λ_k fijo, las ecuaciones

$$0 = \phi_p - \lambda_p \sum_{q=1}^n k_{pq} \phi_q \quad p = 1, 2, \dots, n \quad (27)$$

tienen la solución

$$\phi^k = (\phi_1^k, \phi_2^k, \dots, \phi_n^k)$$

que es única salvo un factor constante. Se tiene entonces, tal como demostró Hilbert

$$K(x, x) = \sum_{k=1}^n \frac{1}{\lambda_k} \frac{(\phi^k, x)^2}{(\phi^k, \phi^k)} \quad (28)$$

donde los paréntesis representan de nuevo el producto escalar de dos vectores.

El teorema generalizado de los ejes principales de Hilbert puede enunciarse de la manera siguiente: sea $K(s, t)$ una función continua simétrica de s y t . Sea $\phi^p(s)$ la autofunción normalizada correspondiente al autovalor λ_p de la ecuación integral (24). Entonces se verifica la siguiente relación, donde $x(s)$ e $y(s)$ son funciones continuas arbitrarias,

$$\int_a^b \int_a^b K(s, t) x(s) y(t) ds dt = \sum_{p=1}^{\alpha} \frac{1}{\lambda_p} \left(\int_a^b \phi^p(s) x(s) ds \right) \times \quad (29)$$

$$\times \left(\int_a^b \phi^p(s) y(s) ds \right)$$

donde $a = n$ ó ∞ , dependiendo del número de autovalores, y en el segundo caso la suma converge uniforme y absolutamente para todo $x(s)$ e $y(s)$ que satisfagan las condiciones

$$\int_a^b x^2(s) ds < \infty \quad \text{y} \quad \int_a^b y^2(s) ds < \infty$$

La generalización de (28) a (29) se pone de manifiesto si definimos en primer lugar $\int_a^b u(s) v(s) ds$ como el producto escalar de las dos funciones $u(s)$ y $v(s)$ y lo representamos por (u, v) . Ahora basta reemplazar $y(s)$ en (29) por $x(s)$ y en el miembro de la derecha de (28) la suma por integración.

A continuación demostró Hilbert un famoso resultado que más tarde recibió el nombre de teorema de Hilbert-Schmidt. Si $f(s)$ es tal que para alguna función continua $g(s)$ se verifica que

$$f(s) = \int_a^b K(s, t) g(t) dt \quad (30)$$

entonces

$$f(s) = \sum_{p=1}^{\infty} c_p \phi^p \quad (31)$$

donde las ϕ^p son las autofunciones ortonormales de K y

$$c_p = \int_a^b \phi^p(s) f(s) ds \quad (32)$$

Así pues, una función «arbitraria» $f(s)$ puede expresarse como un desarrollo en serie en las autofunciones de K con coeficientes c_p que son los «coeficientes de Fourier» del desarrollo.

En el trabajo anterior, Hilbert llevó a cabo un proceso de paso al límite que le permitió generalizar resultados sobre sistemas finitos de ecuaciones lineales y formas cuadráticas finitas a integrales y ecuaciones integrales. Sobre esta base decidió que un tratamiento de las formas cuadráticas infinitas, es decir, las formas cuadráticas con infinitas variables, por sí mismas, «vendría a completar de una manera esencial la teoría bien conocida de las formas cuadráticas con un número finito de variables», lo cual le llevó a ocuparse de problemas que pueden ser considerados como puramente algebraicos. Partiendo de una forma bilineal infinita

$$K(x, y) = \sum_{p, q=1}^{\infty} k_{pq} x_p y_q$$

y procediendo por paso al límite sobre resultados para formas bilineales y cuadráticas en $2n$ y n variables respectivamente, obtiene Hilbert algunos resultados básicos; los detalles del trabajo son complicados y nos limitaremos a formular solamente algunos de estos resultados. En primer lugar obtiene Hilbert una expresión para una forma resolvente $K(\lambda; x, x)$ que tiene la característica peculiar de que es la suma de varias expresiones, una para cada valor de λ de un cierto conjunto discreto, y de una integral sobre un

conjunto de λ pertenecientes a un dominio continuo. El conjunto discreto de valores de λ corresponde al espectro puntual de K , y el conjunto continuo al espectro continuo o de banda. Esta es la primera aplicación importante de los espectros continuos, que habían sido observados ya en 1896 para ecuaciones en derivadas parciales por Wilhelm Wirtinger (n. 1865).⁴⁰²

Para llegar al resultado clave sobre formas cuadráticas, introdujo Hilbert el concepto de forma acotada. La notación (x,x) representa el producto escalar del vector $(x_1, x_2, \dots, x_n, \dots)$ consigo mismo, y (x,y) el producto análogo de los vectores x e y . Entonces, la forma $K(x,y)$ se llama acotada si $|K(x,y)| \leq M$ para todo x e y tales que $(x,x) \leq 1$ e $(y,y) \leq 1$. La acotación implica la continuidad, que define Hilbert para una función de infinitas variables.

El resultado clave de Hilbert que hemos mencionado es la generalización a formas cuadráticas con infinitas variables del conocido teorema de los ejes principales de la geometría analítica. Hilbert demuestra que existe una transformación ortogonal T tal que en las nuevas variables x' , donde $x' = Tx$, la forma K se reduce a una «suma de cuadrados». Es decir, toda forma cuadrática acotada

$$K(x, x) = \sum_{p,q=1}^{\infty} k_{pq} x_p x_q$$

puede transformarse mediante una transformación ortogonal única en la forma

$$K(x, x) = \sum_{i=1}^{\infty} k_i x_i^2 + \int_{(s)} \frac{d\sigma(\mu, \xi)}{\mu} \quad (33)$$

donde los k_i son los autovalores recíprocos de K . La integral, en cuya explicación no entraremos, está extendida a un dominio continuo de autovalores o espectro continuo.

Para eliminar el espectro continuo introduce Hilbert el concepto de continuidad completa: una función $F(x_1, x_2, \dots)$ de infinitas variables se llama completamente continua en el punto $a = (a_1, a_2, \dots)$ si

$$\lim_{\substack{\varepsilon_1 \rightarrow 0 \\ \varepsilon_2 \rightarrow 0 \\ \dots}} F(a_1 + \varepsilon_1, a_2 + \varepsilon_2, \dots) = F(a_1, a_2, \dots)$$

siempre que $\varepsilon_1, \varepsilon_2, \dots$ recorran cualquier sistema de valores $\varepsilon_1^{(h)}, \varepsilon_2^{(h)}, \dots$ con límites

$$\lim_{h \rightarrow \infty} \varepsilon_1^{(h)} = 0 \quad \lim_{h \rightarrow \infty} \varepsilon_2^{(h)} = 0, \dots$$

Esta es una condición más fuerte que la continuidad introducida previamente por Hilbert.

Para que una forma cuadrática $K(x, x)$ sea completamente continua es suficiente que

$$\sum_{p, q=1}^{\infty} k_{pq}^2 < \infty$$

Con esta condición adicional pudo demostrar Hilbert que si K es una forma acotada completamente continua entonces, mediante una transformación ortogonal, puede reducirse a la forma

$$K(x, x) = \sum_j k_j x_j^2 \quad (34)$$

donde los k_j son autovalores recíprocos y (x_1, x_2, \dots) satisface la condición de que $\sum_1^\infty x_j^2$ es finita.

A continuación Hilbert se dedica a aplicar su teoría de formas cuadráticas en infinitas variables a las ecuaciones integrales. Los resultados no son nuevos, la mayoría de ellos, pero se obtienen por métodos más claros y sencillos. Hilbert comienza esta nueva etapa de su trabajo sobre ecuaciones integrales definiendo el importante concepto de sistema ortogonal completo e funciones $\{\phi_p(s)\}$. Se trata de una sucesión de funciones definidas y continuas todas ellas sobre el intervalo $|a, b|$ con las siguientes propiedades:

1. ortogonalidad:

$$\int_a^b \phi_p(s) \phi_q(s) ds = \delta_{pq} \quad p, q = 1, 2, \dots$$

2. completitud: para todo par de funciones u y v definidas sobre el intervalo

$$\int_a^b u(s)v(s)ds = \sum_{p=1}^{\infty} \int_a^b \phi_p(s)u(s)ds \int_a^b \phi_p(s)v(s)ds$$

El valor

$$u_p^* = \int_a^b \phi_p(s)u(s)ds$$

recibe el nombre de coeficiente de Fourier de $u(s)$ con respecto al sistema $\{\phi_p\}$.

Hilbert demuestra que se puede definir un sistema ortonormal completo para todo intervalo finito $[a,b]$, por ejemplo mediante el uso de polinomios. Demuestra una desigualdad generalizada de Bessel y finalmente que la condición

$$\int_a^b u^2(s) ds = \sum_1^{\infty} u_p^{*2}$$

es equivalente a la completitud.

Hilbert vuelve a considerar la ecuación integral

$$f(s) = \phi(s) + \int_a^b K(s,t)\phi(t)dt \quad (35)$$

El núcleo $K(s,t)$, no necesariamente simétrico, puede desarrollarse en una serie de «Fourier» doble por medio de los coeficientes

$$a_{pq} = \int_a^b \int_a^b K(s,t) \phi_p(s) \phi_q(t) ds dt$$

Se sigue que

$$\sum_{p,q=1}^{\infty} a_{pq}^2 \leq \int_a^b \int_a^b K^2(s,t) ds dt$$

y también que si

$$a_p = \int_a^b \phi_p(s) f(s) ds$$

es decir, si los a son los «coeficientes de Fourier» de $f(s)$, entonces $\sum_{p=1}^{\infty} a_p^2 < \infty$, Hilbert transforma a continuación la anterior ecuación integral en un sistema de infinitas ecuaciones lineales en infinitas incógnitas. La idea consiste en tratar de resolver la ecuación integral en $\emptyset(s)$ como el problema de hallar los «coeficientes de Fourier» de $\emptyset(s)$. Representando los coeficientes, aún desconocidos, por x_1, x_2, \dots obtiene las siguientes ecuaciones lineales

$$x_p + \sum_{q=1}^{\infty} a_{pq} x_q \quad p = 1, 2, \dots \quad (36)$$

y demuestra que si este sistema tiene solución única entonces la ecuación integral tiene solución continua única, y que cuando el sistema lineal homogéneo asociado a (36) tiene n soluciones linealmente independientes, la ecuación integral homogénea asociada a (35)

$$0 = \phi(s) + \int_a^b K(s,t)\phi(t)dt \quad (37)$$

tiene n soluciones linealmente independientes. En este caso la ecuación integral no homogénea original tiene una solución si y sólo si $h = 1, 2, \dots, n$, que son las n soluciones linealmente independientes de la ecuación homogénea transpuesta

$$\phi(s) + \int_a^b K(t,s)\phi(t)dt = 0$$

y que también existen cuando (37) tiene n soluciones, satisfacen las condiciones

$$\int_a^b \psi^{(h)}(s)f(s) ds = 0 \quad h = 1, 2, \dots, n \quad (38)$$

Así se obtiene el teorema de la alternativa de Fredholm: o bien la ecuación

$$f(s) = \phi(s) + \int_a^b K(s,t)\phi(t)dt \quad (39)$$

tiene una única solución para toda función f , o la ecuación homogénea asociada tiene n soluciones linealmente independientes. En el segundo caso (39) tiene una solución si y sólo si se verifican las condiciones de ortogonalidad (38).

Hilbert considera a continuación el problema de autovalores

$$f(s) = \phi(s) - \lambda \int_a^b K(s,t)\phi(t)dt \quad (40)$$

donde K es ahora simétrico. La simetría de K implica que sus «coeficientes de Fourier» determinan una forma cuadrática $K(x,x)$ que es completamente continua, y demuestra que existe una transformación ortogonal T cuya matriz es $\{l_{pq}\}$ tal que

$$K(x', x') = \sum_{p=1}^{\infty} \mu_p x_p'^2$$

donde los μ_p son los autovalores recíprocos de la forma cuadrática $K(x, x)$. Las autofunciones $\{\phi_p(s)\}$ para el núcleo $K(s, t)$ vienen definidas ahora por

$$L_p(K(s)) = \sum_{q=1}^{\infty} l_{pq} \int_a^b K(s, t) \phi_q(t) dt = \mu_p \phi_p(s)$$

donde los $\phi_q(t)$ constituyen un sistema ortonormal completo dado. Se puede demostrar que los $\phi_p(s)$ (distintos de los $\phi_q(t)$) forman también un sistema ortonormal y satisfacen

$$\phi_p(s) = \lambda_p \int_a^b K(s, t) \phi_p(t) dt$$

donde $\lambda_p = 1/\mu_p$. Así demuestra Hilbert de nuevo la existencia de autofunciones para el caso homogéneo de (40) y para todo autovalor finito de la forma cuadrática $K(x, x)$ asociada al núcleo $K(s, t)$ de (40). Hilbert demuestra de nuevo (teorema de Hilbert-Schmidt) que si $f(s)$ es una función continua cualquiera para la que existe una g tal que

$$\int_a^b K(s,t)g(t) dt = f(s)$$

entonces f es representable en serie de autofunciones de K , que es uniforme y absolutamente convergente (véase [31]). Hilbert utiliza este resultado para demostrar que la ecuación homogénea asociada a (40) no tiene ninguna solución no trivial excepto para los autovalores λ_p . Entonces el teorema de la alternativa de Fredholm afirma que: para $\lambda \neq \lambda_p$ la ecuación (40) tiene solución única; para $\lambda = \lambda_p$ tiene solución si y sólo si se satisfacen las n_p condiciones

$$\int_a^b \phi_{p+j}(s) f(s) = 0 \quad j = 1, 2, \dots, n_p$$

donde las $\phi_{p+j}(s)$ son las n autofunciones asociadas a λ_p . Por último demuestra una vez más la generalización del teorema de los ejes principales:

$$\int_a^b \int_a^b K(s,t)u(s)u(t) ds dt = \sum_{p=1}^{\infty} \frac{1}{\lambda_p} \left\{ \int_a^b u(t)\phi_p(t) dt \right\}^2$$

donde $u(s)$ es una función continua arbitraria y todas las ϕ_p asociadas a cualquier λ_p aparecen incluidas en la suma.

En este último trabajo (1906) Hilbert prescinde del uso de los

determinantes infinitos de Fredholm, demostrando directamente la relación entre las ecuaciones integrales y la teoría de sistemas ortogonales completos para el desarrollo en serie de funciones.

Hilbert aplicó también sus resultados sobre ecuaciones integrales a diversos problemas de geometría y de física. En particular, en el tercero de los seis artículos resolvió el problema de Riemann de construir una función holomorfa en un dominio limitado por una curva lisa, cuando se da la parte real o imaginaria del valor en la frontera o ambas están relacionadas por una ecuación lineal dada.

Uno de los resultados más notables de la obra de Hilbert, que apareció en los artículos de 1904 y 1905, es la formulación de los problemas de contorno de Sturm-Liouville para ecuaciones diferenciales como ecuaciones integrales. El resultado de Hilbert afirma que los autovalores y autofunciones de la ecuación diferencial

$$\frac{d}{dx} \left(p(x) \frac{du}{dx} \right) + q(x)u + \lambda u \quad (41)$$

con las condiciones de contorno $u(a) = 0$, $u(b) = 0$ (e incluso condiciones de contorno más generales) son los autovalores y autofunciones de

$$\phi(x) - \lambda \int_a^b G(x, \xi) \phi(\xi) d\xi = 0 \quad (42)$$

donde $G(x,\xi)$ es la función de Green para la ecuación (41), es decir, una solución particular de

$$\frac{d}{dx}\left(p \frac{du}{dx}\right) + q(x)u = 0$$

que satisface ciertas condiciones de diferenciabilidad y cuya derivada parcial dG/dx tiene una singularidad de salto en $x = \xi$ igual a $-1/p(\xi)$. Resultados análogos se verifican para las ecuaciones en derivadas parciales. Así pues, las ecuaciones integrales constituyen un instrumento para resolver ecuaciones diferenciales ordinarias y en derivadas parciales.

Recapitulando los principales resultados obtenidos por Hilbert, en primer lugar estableció la teoría espectral general para núcleos simétricos K . Solamente veinte años antes había requerido grandes esfuerzos matemáticos (cap. 28, sec. 8) demostrar la existencia de la frecuencia de oscilación inferior para una membrana. Con ayuda de las ecuaciones integrales, la demostración constructiva de la existencia de la serie completa de frecuencias y de las autofunciones concretas se obtenía para condiciones muy generales relativas al medio de oscilación. Emile Picard ⁴⁰³ fue el primero en obtener estos resultados utilizando la teoría de Fredholm. Otro resultado notable debido a Hilbert es el de que el desarrollo de una función en las autofunciones pertenecientes a una ecuación integral de segundo tipo depende de la resolubilidad de la correspondiente ecuación integral de primer tipo. En particular, Hilbert descubrió que el éxito

del método de Fredholm radicaba en el concepto de continuidad completa, que aplicó a las formas bilineales, estudiándolo sistemáticamente. Aquí inauguraba la teoría espectral de formas bilineales simétricas.

Una vez que Hilbert mostró cómo convertir problemas de ecuaciones diferenciales en ecuaciones integrales, se empezó a utilizar este planteamiento cada vez más para resolver problemas físicos; aquí el uso de una función de Green para la conversión ha resultado una herramienta de gran importancia. Hilbert mismo demostró,⁴⁰⁴ por su parte, en problemas de dinámica de gases, que uno puede ir directamente a las ecuaciones integrales. Este recurso directo a las ecuaciones integrales es posible porque el concepto de suma demuestra ser tan fundamental en algunos problemas físicos como el concepto de cociente incremental que conduce a las ecuaciones diferenciales en otros. Hilbert subrayó también que no eran las ecuaciones diferenciales ordinarias ni en derivadas parciales, sino las ecuaciones integrales, el punto de partida natural y necesario para la teoría de desarrollos de funciones en serie, y que los desarrollos obtenidos mediante ecuaciones diferenciales no eran más que casos particulares del teorema general de la teoría de ecuaciones integrales.

4. Los sucesores inmediatos de Hilbert

La obra de Hilbert sobre ecuaciones integrales fue simplificada por Erhard Schmidt (1876-1959), profesor en varias universidades alemanas, utilizando métodos que había introducido H. A. Schwarz

en teoría del potencial. La contribución más importante de Schmidt fue su generalización del concepto de autofunción para ecuaciones integrales con núcleo no simétrico, que data de 1907.⁴⁰⁵

Otro matemático que también se dedicó, en 1907, a continuar la obra de Hilbert fue Friedrich Riesz (1880-1956), de nacionalidad húngara.⁴⁰⁶ Hilbert había estudiado ecuaciones integrales de la forma

$$f(s) = \phi(s) + \int_a^b K(s,t)\phi(t)dt$$

donde f y K son continuas. Riesz trató de extender las ideas de Hilbert a funciones $f(s)$ más generales. Con este objeto, lo que se necesitaba era asegurarse de que los «coeficientes de Fourier» de f con respecto a un sistema ortonormal de funciones $\{\phi_p\}$ se podían calcular. También estaba interesado en investigar en qué condiciones una sucesión de números dada $\{a_p\}$ podía ser la sucesión de coeficientes de Fourier de una función f con respecto a un sistema ortonormal dado $\{\phi_p\}$.

Riesz consideró funciones cuyo cuadrado es integrable en el sentido de Lebesgue, y obtuvo el siguiente teorema: sea $\{\phi_p\}$ una sucesión ortonormal de funciones de cuadrado integrable en el sentido de Lebesgue, definidas sobre el intervalo $[a,b]$. Si $\{a_p\}$ es una sucesión de números reales, entonces la convergencia de $\sum_{p=1}^{\infty} a_p^2$ es condición necesaria y suficiente para que exista una función f tal que

$$\int_a^b f(x)\phi_p(x)dx = a_p$$

para cada ϕ_p y a_p . Se demuestra además que la función f es de cuadrado integrable. Este teorema establece una correspondencia biunívoca entre el conjunto de las funciones de cuadrado integrable y el conjunto de las sucesiones de cuadrado sumable, para cada sucesión ortonormal de funciones de cuadrado integrable.

Utilizando las funciones integrables en el sentido de Lebesgue pudo demostrar Riesz que la ecuación integral de segundo tipo

$$f(s) = \phi(s) + \int_a^b K(s,t)\phi(t)dt$$

admite solución bajo las condiciones más débiles de que $f(s)$ y $K(s,t)$ sean de cuadrado integrable. La solución es única salvo una función cuya integral de Lebesgue sobre $[a,b]$ sea nula.

El mismo año que Riesz publicaba sus primeros artículos, un profesor de la universidad de Colonia, Ernst Fischer (1875-1959) introducía el concepto de convergencia en media.⁴⁰⁷ Se dice que una sucesión de funciones $\{f_n\}$ definidas sobre el intervalo $[a,b]$ converge en media si

$$\lim_{m,n \rightarrow \infty} \int_a^b (f_n(x) - f_m(x))^2 dx = 0$$

y se dice que $\{f_n\}$ converge en media a f si

$$\lim_{n \rightarrow \infty} \int_a^b (f - f_n)^2 dx = 0$$

donde las integrales están tomadas en el sentido de Lebesgue. La función f está determinada de manera única salvo una función definida sobre un conjunto de medida cero, es decir, una función $g(x) \neq 0$, llamada una función nula, que satisface la condición $\int_a^b g^2(x) dx = 0$.

El conjunto de funciones de cuadrado integrable en el sentido de Lebesgue sobre un intervalo $[a,b]$ se representó más tarde por $L^2(a,b)$ o simplemente por L^2 . El resultado principal obtenido por Fischer fue que $L^2(a,b)$ es completo en media, es decir, si las funciones f_n pertenecen a $L^2(a,b)$ y la sucesión $\{f_n\}$ converge en media, entonces converge en media a una función f perteneciente a $L^2(a,b)$. Esta propiedad de completitud es la principal ventaja de utilizar funciones de cuadrado sumable. Fischer dedujo a continuación como corolario el teorema de Riesz mencionado anteriormente, y que hoy se conoce como teorema de Riesz-Fischer. En una nota posterior ⁴⁰⁸ subraya Fischer que era esencial el uso de las funciones de cuadrado integrable; el teorema no sería válido

para ningún subconjunto.

La determinación de una función $f(x)$ correspondiente a un conjunto de coeficientes de Fourier $\{a_n\}$ con respecto a una sucesión dada de funciones ortonormales $\{g_n\}$, o la determinación de una f tal que

$$\int_a^b g_n(x) f(x) dx = a_n \quad n = 1, 2, \dots$$

que aparecía en los artículos de Riesz de 1907, recibe hoy el nombre de problema de los momentos (entendiendo siempre la integración en el sentido de Lebesgue). En 1910⁴⁰⁹ intentó Riesz generalizar este problema, y debido a que necesitaba utilizar las desigualdades de Hölder y

$$\sum_{i=1}^n |a_i b_i| \leq \left(\sum_{i=1}^n |a_i|^p \right)^{1/p} \left(\sum_{i=1}^n |b_i|^q \right)^{1/q}$$

y

$$\left| \int_M f(x) g(x) dx \right| \leq \left(\int_M |f|^p dx \right)^{1/p} \left(\int_M |g|^q dx \right)^{1/q}$$

donde $1/p + 1/q = 1$, y otras desigualdades, se vio obligado a introducir el conjunto L^p de las funciones f medibles sobre un conjunto M tales que $|f|^p$ es integrable sobre M . Su primer teorema

importante fue el de que si una función $h(x)$ es tal que el producto $f(x)h(x)$ es integrable para toda f de L^p , entonces h pertenece a L^q y, recíprocamente, el producto de una función de L^p por otra de L^q siempre es integrable; se entiende que en todos los casos es $p > 1$ y $1/p + 1/q = 1$.

Riesz introdujo también los conceptos de convergencia fuerte y débil. La sucesión de funciones $\{f_n\}$ converge fuertemente a f (en la media de orden p) si

$$\lim_{n \rightarrow \infty} \int_a^b |(f_n(x) - f(x))|^p dx = 0$$

La sucesión $\{f_n\}$ converge débilmente a f si

$$\int_a^b |f_n(x)|^p dx < M$$

donde M es independiente de n , y si para todo x del intervalo $[a, b]$

$$\lim_{n \rightarrow \infty} \int_a^x (f_n(t) - f(t)) dt = 0$$

La convergencia fuerte implica la convergencia débil. (La definición moderna de convergencia débil: si las funciones f_n y f pertenecen a L^p y si

$$\lim_{n \rightarrow \infty} \int_a^x (f(x) - f_n(x))g(x)dx = 0$$

se verifica para toda g de L^q , entonces $\{f_n\}$ converge débilmente a f , es equivalente a la de Riesz).

En el mismo artículo de 1910 generaliza Riesz la teoría de las ecuaciones integrales

$$\phi(x) - \lambda \int_a^b K(x,t)\phi(t)dt = f(x)$$

al caso en que la f dada y la incógnita ϕ son funciones de L^p . Los resultados sobre la resolución del problema de los autovalores para esta ecuación integral son análogos a los de Hilbert. Lo más sorprendente es que, para llevar esto a cabo, Riesz introduce el concepto abstracto de operador, formula para él el concepto de continuidad completa de Hilbert, inaugurando así la teoría abstracta de operadores. En el próximo capítulo hablaremos más de este planteamiento abstracto. Entre otros resultados demostró Riesz que el espectro continuo de un operador real completamente continuo sobre L^2 es vacío.

5. Generalizaciones de la teoría

La importancia que atribuyó Hilbert a las ecuaciones integrales

puso de moda el tema a escala mundial durante un largo período de tiempo, durante el cual se produjo una masa enorme de literatura, la mayor parte de un valor efímero. Sin embargo, algunas generalizaciones se mostraron valiosas, aunque aquí solamente podremos mencionarlas.

La teoría de ecuaciones integrales que hemos presentado se refiere a ecuaciones integrales lineales, es decir, en las que la función incógnita $u(x)$ aparece linealmente. Esta teoría se ha generalizado a ecuaciones integrales no lineales donde la función incógnita aparece elevada al cuadrado o potencias más altas, o de alguna manera más complicada.

Por otra parte, en nuestro breve resumen apenas hemos dicho nada acerca de las condiciones sobre las funciones $f(x)$ y $K(x,\xi)$ que daban validez a las conclusiones. Si estas funciones no son continuas o si el intervalo $[a,b]$ se sustituye por un intervalo infinito, muchos de los resultados se ven afectados, o al menos se necesitan nuevas demostraciones. Así, incluso la transformación de Fourier

$$f(x) = \sqrt{\frac{2}{\pi}} \int_0^{\infty} \cos(x\xi)u(\xi)d\xi$$

que puede considerarse como una ecuación integral de primer tipo y que tiene como solución la transformación inversa

$$u(x) = \sqrt{\frac{2}{\pi}} \int_0^{\infty} \cos(x\xi) f(\xi) d\xi$$

tiene exactamente dos autovalores ± 1 y cada uno de ellos tiene una cantidad infinita de auto funciones. Estos casos se estudian ahora bajo el nombre de ecuaciones integrales singulares; tales ecuaciones no pueden resolverse por los métodos aplicados a las ecuaciones de Volterra y Fredholm. Además presentan una curiosa propiedad, a saber, que hay intervalos continuos de valores de A o espectros de banda para los que existen soluciones. El primer trabajo importante sobre este tema se debe a Hermann Weyl (1885-1955).⁴¹⁰

También han recibido mucha atención los teoremas de existencia para ecuaciones integrales, tanto lineales como no lineales. Por ejemplo, muchos matemáticos han formulado teoremas de existencia para ecuaciones como

$$y(x) = f(x) + \int_a^x K(x, s, y(s)) ds$$

que incluyen como caso especial la ecuación de Volterra de segundo tipo

$$y(x) = f(x) + \int_a^x K(x, s) y(s) ds$$

Desde un punto de vista histórico, el siguiente desarrollo importante consistió en la generalización de las ideas anteriores en un marco abstracto. Hilbert consideraba dada una función por sus coeficientes de Fourier, los cuales satisfacían la condición de que $\sum_{1}^{\infty} a_p^2$ es finito, por lo que introdujo también sucesiones generales de números reales $\{x_n\}$, tales que $\sum_{1}^{\infty} x_n^2$ es finito. Riesz y Fischer demostraron entonces que hay una correspondencia biunívoca entre las funciones de cuadrado integrable y las sucesiones de cuadrado sumable de sus coeficientes de Fourier. Las sucesiones de cuadrado sumable pueden ser consideradas como puntos de un espacio de infinitas dimensiones que es una generalización del espacio euclídeo n-dimensional. Así pues, las funciones pueden ser consideradas como puntos de un espacio, llamado hoy espacio de Hilbert, y la integral $\int_a^b K(x,y) u(x) dx$ puede ser interpretada como un operador que transforma $u(x)$ en sí misma o en otra función. Estas ideas sugirieron un planteamiento abstracto para el estudio de las ecuaciones integrales que encajaba bien en el también incipiente planteamiento abstracto del cálculo de variaciones. Este nuevo planteamiento se conoce hoy como análisis funcional y a él dedicaremos el próximo capítulo.

Bibliografía

- Bernkopf, M.: «*The Development of Function Spaces with Particular Reference to their Origins in Integral Equation Theory*». *Archive for History of Exact Sciences*, 3, 1966, 1-96.
- Bliss, G. A.: «*The Scientific Work of E. H. Moore*». *Amer. Math.*

Bull, 40, 1934, 501-514.

- Bocher, M.: *An Introduction to the Study of Integral Equations*, 2.^a ed., Cambridge University Press, 1913.
- Bourbaki, N.: *Elementos de historia de las matemáticas*, Madrid, Alianza, 1976.
- Davis, Harold T.: *The Present State of Integral Equations*, Indiana University Press, 1926.
- Hahn, H.: «Bericht über die Theorie der linearen Integralgleichungen». *Jahres. der Deut. Math.-Verein.*, 20, 1911, 69-117.
- Hellinger, E.: *Hilberts Arbeiten über Integralgleichungen und unendliche Gleichungs-systeme*, in Hilbert's *Gesam. Abh.*, 3, 94-145, Julius Springer, 1935. - «Begründung der Theorie quadratischer Formen von unendlichvielen Veränderliche». *Jour. für Math.*, 136, 1909, 210-271.
- Hellinger, E., y Toeplitz, O.: «*Integralgleichungen und Gleichungen mit unendlichvielen Unbekannten*». *Encyk. der Math. Wiss.*, B. G. Teubner, 1923-1927, vol. 2, part. 3, 2.^a half, 1335-1597.
- Hilbert, D.: *Grundzüge einer allgemeinen Theorie der linearen Integralgleichungen*, 1912, Chelsea (reprint), 1953.
- Reid, Constance: *Hilbert*, Springer-Verlag, 1970.
- Volterra, Vito: *Opere matematiche*, 5 vols. Accademia Nazionale dei Lincei, 1954-1962.
- Weyl, Hermann: *Gesammelte Abhandlungen*, 4 vols., Springer-Verlag, 1968.

Capítulo 46

El análisis funcional

Uno debe estar bien seguro de que ha permitido a la ciencia hacer un gran progreso, si va a sobrecargarla con una multitud de términos nuevos y a exigir que los lectores sigan una investigación que les ofrece tantas cosas extrañas.

A. L. Cauchy

Contenido:

- 1. ¿Qué es el análisis funcional?*
- 2. La teoría de funcionales*
- 3. El análisis funcional lineal*
- 4. La axiomatización de los espacios de Hilbert*

Bibliografía

1. ¿Qué es el análisis funcional?

A finales del siglo XIX se puso claramente de manifiesto que muchos campos de la matemática utilizaban transformaciones u operadores que actuaban sobre funciones; por ejemplo, incluso la simple operación de diferenciación ordinaria y su inversa la antidiferenciación o integración actúan sobre una función para

producir otra. En los problemas de cálculo de variaciones en los que se manejan integrales tales como

$$J = \int_a^b F(x, y, y') dx$$

dicha integral puede considerarse como operando sobre una clase de funciones $y(x)$, de las cuales se busca precisamente aquella que haga máxima o mínima la integral. La teoría de ecuaciones diferenciales nos ofrece otra clase de operadores; por ejemplo, el operador diferencial

$$L = \frac{d^2}{dx^2} + p(x) \frac{d}{dx} + q(x)$$

actuando sobre una clase de funciones $y(x)$ las transforma en otras funciones. Naturalmente, resolver la ecuación diferencial asociada a este operador significa buscar una función particular $y(x)$ tal que L actuando sobre ella nos dé 0, y que satisfaga posibles condiciones iniciales o de contorno dadas. Como un último ejemplo de operadores tenemos las ecuaciones integrales; el segundo miembro de la igualdad

$$f(x) = \int_a^b K(x, y)u(x) dx$$

puede ser considerado como un operador que actúa sobre las diversas $u(x)$ posibles para producir nuevas funciones, y aquí también, como en el caso de las ecuaciones diferenciales, la función $u(x)$ solución de la ecuación debe transformarse en $f(x)$.

La idea que motivó la creación del análisis funcional fue la de que todos estos operadores podrían ser estudiados dentro de una formulación abstracta de una teoría general de operadores actuando sobre clases de funciones; además, estas funciones podrían ser consideradas como elementos o puntos de un espacio, entonces el operador transformaría puntos en puntos, y en este sentido sería una generalización de las transformaciones geométricas ordinarias, tales como las rotaciones. Algunos de los operadores anteriores transforman funciones en números reales y no en otras funciones; los operadores que dan como resultado números reales o complejos reciben actualmente el nombre de funcionales, mientras que el nombre de operador se reserva normalmente para las transformaciones que aplican funciones en funciones. Así pues, el nombre de análisis funcional, introducido por Paul P. Levy (1886) cuando los funcionales constituían el centro del interés, ya no resulta hoy muy adecuado. La búsqueda de la generalidad y de la unificación de campos diversos es una de las características distintivas de la matemática del siglo XX, y el análisis funcional también tiende a alcanzar esas metas, evidentemente.

2. La teoría de funcionales

La teoría abstracta de funcionales fue iniciada por Volterra en sus trabajos sobre cálculo de variaciones. Volterra publicó una serie de artículos⁴¹¹ sobre funciones de líneas (curvas), tal como él las llamaba. Una función de línea era, para Volterra, una función real F , cuyos valores dependen de todos los valores que toman ciertas funciones $y(x)$ definidas en un intervalo $[a, b]$. Las funciones mismas se consideraban como puntos de un espacio en el cual se podían definir los entornos de un punto y el límite de una sucesión de puntos. Volterra dio, para los funcionales $F[y(x)]$, definiciones de continuidad, derivada y diferencial, pero estas definiciones no resultaban adecuadas para la teoría abstracta del cálculo de variaciones y fueron reemplazadas; de hecho, las definiciones dadas por Volterra fueron criticadas por Hadamard.⁴¹²

Antes incluso de que Volterra comenzara sus trabajos, ya había sido admitida la idea de considerar una colección de funciones $y(x)$, definidas todas ellas en algún intervalo común, como puntos de un espacio⁴¹³. Riemann hablaba ya en su tesis de una colección de funciones formando un dominio cerrado conexo (de puntos de un espacio), y Giulio Ascoli (1843-1896)⁴¹⁴ y Cesare Arzelá⁴¹⁵ intentaban, por su parte, extender a conjuntos de funciones la teoría de conjuntos de puntos de Cantor, considerando así las funciones de nuevo como puntos de un espacio; Arzelá hablaba también de funciones de línea. Hadamard sugería en el Primer Congreso Internacional de Matemáticos, en 1897,⁴¹⁶ considerar las curvas como puntos de un conjunto; él pensaba en la familia de todas las funciones continuas definidas en el intervalo $[0, 1]$, familia

que se le presentó de manera natural en sus trabajos sobre ecuaciones en derivadas parciales. Entile Borel hizo la misma sugerencia,⁴¹⁷ aunque con una finalidad diferente, la del estudio de funciones arbitrarias por medio de series.

Hadamard también se vio conducido al estudio de los funcionales⁴¹⁸ con motivo del cálculo de variaciones. El nombre de funcional es debido a él, así como el llamar a un funcional $U[y(t)]$ lineal cuando, si $y(t) = \lambda_1 y_1(t) + \lambda_2 y_2(t)$ donde λ_1 y λ_2 son constantes, entonces

$$U[y(t)] = \lambda_1 U[y_1(t)] + \lambda_2 U[y_2(t)]$$

El primer esfuerzo importante por construir una teoría abstracta de espacios de funciones y de funcionales fue el llevado a cabo por Maurice Fréchet (1878), eminente matemático francés, en su tesis doctoral de 1906.⁹ En lo que Fréchet llamó cálculo funcional, intentó unificar en términos abstractos las ideas contenidas en los trabajos de Cantor, Volterra, Arzelá, Hadamard y otros.

Con objeto de conseguir el más alto grado de generalidad para sus espacios de funciones, Fréchet echó mano de todas las ideas básicas sobre conjuntos desarrolladas por Cantor, aunque para él los puntos de los conjuntos eran ahora funciones. También formuló de una manera más general el concepto de Emite de un conjunto de puntos; este concepto no quedaba definido explícitamente, sino que venía caracterizado por unas propiedades lo bastante generales como para incluir los distintos tipos de límites que aparecían en las teorías concretas que Fréchet intentaba unificar. Introdujo así una

clase L de espacios, donde la L indica que existe un concepto de límite definido sobre cada uno de estos espacios, de manera que si B es un espacio de la clase L y tomamos en él elementos A_1, A_2, \dots arbitrarios, debe ser posible determinar si existe o no un elemento único A , llamado, cuando existe, límite de la sucesión $\{A_n\}$, tal que

- a. Si $A_i = A$ para todo i , entonces $\lim\{A_n\} = A$.
- b. Si A es el límite de $\{A_n\}$ entonces A también es el límite de toda subsucesión infinita de $\{A_n\}$.

A partir de este concepto definía Fréchet un cierto número de conceptos aplicables a cualquier espacio de la clase L ; por ejemplo, el conjunto derivado E' de un conjunto E es el conjunto de todos los puntos del espacio en cuestión que son límites de sucesiones de puntos de E , E es cerrado si E' está contenido en E ; E es perfecto si $E' = E$; un punto A perteneciente a E es un punto interior de E (en sentido restringido) si A no es límite de ninguna sucesión contenida en el complementario de E ; un conjunto E es compacto si es finito o si todo subconjunto infinito de E tiene por lo menos un elemento límite; si E es cerrado y compacto recibe el nombre de extremal («compacto» según Fréchet es lo que actualmente se llama relativamente secuencialmente compacto, y «extremal» equivale al concepto actual de secuencialmente compacto). El primer teorema⁴¹⁹ importante de Fréchet es una generalización del teorema de los intervalos cerrados encajados: si $\{E_n\}$ es una sucesión decreciente de subconjuntos cerrados de un conjunto extremal, es decir, tal que E_{n+1} está contenido en E_n para todo n , entonces la

intersección de todos los E_n es no vacía.

Fréchet pasa a considerar entonces funcionales (a los que llama operaciones funcionales), es decir, funciones con valores reales definidas sobre un conjunto E , y define la continuidad de un funcional de la manera siguiente: un funcional U se llamará continuo en un elemento A de E , si se verifica que $\lim U(A_n) = U(A)$ para toda sucesión $\{A_n\}$ contenida en E y que converja a A . Introduce también la idea de semicontinuidad de funcionales, concepto definido ya para funciones ordinarias por René Baire (1874-1932) en 1899.⁴²⁰ El funcional U se llamará semicontinuo superiormente en E si $U(A) \geq \lim \sup U(A_n)$ para las $\{A_n\}$ como en el caso anterior, y será semicontinuo inferiormente si $U(A) \leq \lim \inf U(A_n)$ ⁴²¹.

A partir de estas definiciones pudo demostrar Fréchet una serie de teoremas sobre funcionales. Por ejemplo, todo funcional continuo sobre un conjunto extremal E está acotado y alcanza un valor máximo y un valor mínimo sobre E ; todo funcional semicontinuo superiormente definido sobre un conjunto extremal E , está acotado superiormente y alcanza su máximo sobre E .

A continuación Fréchet introdujo generalizaciones de conceptos aplicables a sucesiones y conjuntos de funcionales, tales como el de convergencia uniforme, convergencia cuasi-uniforme, compacidad y equicontinuidad. Por ejemplo, la sucesión de funcionales $\{U_n\}$ converge uniformemente a U si, dado cualquier número positivo ε , $|U_n(A) - U(A)| < \varepsilon$ tomando n suficientemente grande pero independiente del punto A de E . De esta manera pudo demostrar para el caso de funcionales generalizaciones de teoremas obtenidos

previamente para funciones reales.

Una vez estudiados los espacios generales L , definió Fréchet otros espacios más especializados, tales como los espacios de entornos, redefinió los conceptos utilizados para espacios con puntos límites, y demostró teoremas análogos a los anteriores, pero a menudo con mejores resultados, debido a que los espacios eran más ricos en propiedades.

Por último introdujo Fréchet los espacios métricos. En un espacio métrico está definida una función que juega el papel de la distancia (*écart*), representada por (A, B) para cada par de puntos A y B del espacio, función que debe satisfacer las siguientes condiciones:

- a. $(A, B) = (B, A) \geq 0$.
- b. $(A, B) = 0$ si y sólo si $A = B$.
- c. $(A, B) + (B, C) \geq (A, C)$.

La condición (c) se llama desigualdad triangular. A estos espacios los llamó Fréchet de clase ε , y también pudo demostrar para ellos un cierto número de teoremas sobre sus funcionales, de una manera muy parecida a los casos de los espacios más generales.

Fréchet dio algunos ejemplos de espacios de funciones: por ejemplo, el conjunto de todas las funciones reales de una variable real continuas sobre un intervalo I , con el «*écart*» entre dos funciones cualesquiera f y g definido de la forma

$$\max_{x \in I} |f(x) - g(x)|,$$

es un espacio de clase ε ; hoy día a este «*écart*» se le denomina norma del máximo.

Otro ejemplo propuesto por Fréchet es el del conjunto de todas las sucesiones de números reales; si $x = (x_1, x_2, \dots)$ e $y = (y_1, y_2, \dots)$ son dos de estas sucesiones, el «*écart*» entre x e y se define como

$$(x, y) = \sum_{p=1}^{\infty} \frac{1}{p!} \frac{|x_p - y_p|}{1 + |x_p - y_p|}$$

obteniéndose de esta forma un espacio de dimensión infinita numerable, como Fréchet mismo hizo notar.

Usando sus espacios ε , consiguió dar Fréchet⁴²² una definición general de continuidad, diferencial y diferenciabilidad de un funcional. Aunque estas definiciones no eran del todo adecuadas para el cálculo de variaciones, su definición de diferencial vale la pena destacarla porque constituye el núcleo de lo que demostró ser más satisfactorio. Fréchet supone la existencia de un funcional lineal $L(\eta(x))$ tal que

$$F[y(x) + \eta(x)] = F(y) + L(\eta) + \varepsilon M(\eta)$$

donde $\eta(x)$ es una variación sobre la $y(x)$, $M(\eta)$ es el máximo del valor absoluto de $\eta(x)$ sobre $[a, b]$ y ε tiende a cero con M ; entonces $L(\eta)$ es, por definición, la diferencial de $F(y)$. Fréchet suponía también la continuidad de $F(y)$, lo cual es más de lo que puede asegurarse normalmente en los problemas de cálculo de variaciones.

Charles Albert Fischer (1884-1922)⁴²³ mejoró posteriormente la definición de Volterra de la derivada de un funcional, de manera que cubriese el caso de los funcionales usados en el cálculo de variaciones; la diferencial de un funcional podía definirse entonces en términos de la derivada.

Por lo que se refiere a las definiciones básicas de las propiedades de los funcionales necesarias en el cálculo de variaciones, las formulaciones finales fueron dadas por Elizabeth le Stourgeon (1881-1971).⁴²⁴ El concepto clave, el de diferencial de un funcional, es una modificación del de Fréchet: se dice que el funcional $F(y)$ tiene una diferencial en $y_0(x)$ si existe un funcional lineal $L(\eta)$ tal que para todos los arcos $y_0 + \eta$ en un entorno de y_0 , se verifica la relación

$$F(y_0 + \eta) = F(y_0) + L(\eta) + M(\eta) \cdot \varepsilon(\eta)$$

donde $M(\eta)$ es el máximo del valor absoluto de η y η' sobre el intervalo $[a, b]$ y $\varepsilon(\eta)$ tiende a cero con $M(\eta)$. También definió las diferenciales segundas.

Tanto Le Stourgeon como Fischer obtuvieron, a partir de sus definiciones de las diferenciales, condiciones necesarias para que un funcional admita un mínimo, que esta vez sí eran aplicables a los problemas del cálculo de variaciones. Por ejemplo, una condición necesaria para un funcional $F(y)$ tenga un mínimo para $y = y_0$ es que $L(\eta)$ se anule para toda $\eta(x)$ que sea continua y tenga derivada primera continua en $[a, b]$ y tal que $\eta(a) - \eta(b) = 0$. Se puede deducir

la ecuación de Euler de la condición de que se anule la diferencial primera, y usando una definición adecuada de la diferencial segunda de un funcional (que ya habían dado varios de los matemáticos que hemos mencionado), es posible deducir la necesidad de la condición de Jacobi del cálculo de variaciones.

El trabajo definitivo, por lo menos hasta 1925, en la teoría de funcionales necesaria para el cálculo de variaciones, fue llevado a cabo por Leonida Tonelli (1885-1946), profesor de las universidades de Bolonia y Pisa. Después de haber escrito una gran cantidad de artículos sobre el tema desde 1911, publicó su «*Fondamenti di calcólo delle variazioni*» (2 vols., 1922, 1924), donde enfoca el tema desde el punto de vista de los funcionales. La teoría clásica se basaba principalmente en la teoría de ecuaciones diferenciales, y la intención de Tonelli era la de reemplazar los teoremas de existencia para ecuaciones diferenciales por teoremas de existencia para minimizar integrales de curvas. A lo largo de su obra el concepto de semicontinuidad inferior de un funcional es el concepto fundamental, porque los funcionales no suelen ser continuos.

Tonelli considera en primer lugar conjuntos de curvas y da teoremas que aseguran la existencia de una curva límite de una cierta clase de curvas. Los teoremas siguientes aseguran que la integral usual, pero en la forma paramétrica

$$\int_{t_1}^{t_2} F(t, x(t), y(t), x', y') dt$$

será semicontinua inferiormente como función de $x(t)$ y de $y(t)$ (posteriormente considera las integrales no paramétricas, más fundamentales), y obtiene las cuatro condiciones necesarias clásicas del cálculo de variaciones para el tipo habitual de problemas. El segundo tomo de la obra está dedicado principalmente a teoremas de existencia para una gran variedad de problemas, obtenidos a partir del concepto de semicontinuidad. Es decir, dada una integral de la forma anterior, demuestra, imponiendo condiciones sobre ella en tanto que funcional, y sobre la clase de curvas a considerar, que existe una curva en dicha clase que minimiza la integral. Sus teoremas se refieren a máximos y mínimos absolutos y relativos.

En cierta medida, la obra de Tonelli rinde también beneficios en el terreno de las ecuaciones diferenciales, dado que sus teoremas de existencia implican la existencia de soluciones de las ecuaciones diferenciales que suministraban las curvas minimales como soluciones en el planteamiento clásico. Sin embargo, su trabajo estaba naturalmente limitado a los tipos básicos de problemas del cálculo de variaciones. Aunque este enfoque abstracto se vio continuado posteriormente por muchos otros matemáticos, lo cierto es que no se han hecho grandes progresos en la aplicación de la teoría de funcionales al cálculo de variaciones.

3. El análisis funcional lineal

La labor más importante llevada a cabo en análisis funcional fue la de intentar dar una teoría abstracta para las ecuaciones integrales, en oposición al cálculo de variaciones. Las propiedades de los

funcionales necesarias en el segundo caso son bastante particulares y no se suelen verificar para los funcionales en general. Por otro lado, el carácter no lineal de estos funcionales creaba dificultades que no se presentaban en el caso de los funcionales y operadores que incluyen como caso particular a las ecuaciones integrales. Al mismo tiempo que se hacían progresos concretos en la teoría de ecuaciones integrales, debidos a Schmidt, Fischer y Riesz, estos mismos matemáticos, y otros, comenzaron a trabajar en la elaboración de la correspondiente teoría abstracta.

La primera tentativa de elaborar una teoría abstracta de funcionales y operadores lineales fue hecha por el matemático americano E. H. Moore a partir de 1906.⁴²⁵ Moore comprobó que había ciertas características comunes entre la teoría de ecuaciones lineales con un número finito de incógnitas, la teoría de sistemas infinitos de ecuaciones con un número infinito de incógnitas y la teoría de ecuaciones integrales lineales. Basándose en estas analogías, emprendió la tarea de construir una teoría abstracta, a la que llamó Análisis General, que incluiría a las teorías concretas anteriores como casos particulares, y adoptó para ello un planteamiento axiomático. No vamos a presentar aquí los detalles porque la influencia de Moore no fue muy extensa ni consiguió una metodología realmente eficaz; además su lenguaje simbólico era complicado y difícil de seguir.

El primer paso importante hacia una teoría abstracta de funcionales y operadores lineales fue dado por Erhard Schmidt⁴²⁶ y Fréchet⁴²⁷ en 1907. Hilbert, en sus trabajos sobre ecuaciones integrales,

consideraba una función como dada por sus coeficientes de Fourier en su desarrollo con respecto a una sucesión ortonormal de funciones. Estos coeficientes, y los valores que asignaba a los x_i en su teoría de formas cuadráticas en infinitas variables, son sucesiones $\{x_n\}$ tales que $\sum_1^\infty x_n^2$ es finita. Sin embargo, Hilbert no llegó a considerar estas sucesiones como coordenadas de un punto en un espacio, ni utilizó tampoco un lenguaje geométrico. Este paso fue dado por Schmidt y Fréchet; considerando cada sucesión $\{x_n\}$ como un punto, las funciones quedaban representadas como puntos de un espacio de dimensión infinita. Schmidt consideraba también números complejos, además de reales, en las sucesiones $\{x_n\}$; un espacio de este tipo ha sido llamado desde entonces un espacio de Hilbert. Nosotros seguiremos en nuestra exposición el trabajo de Schmidt.

Los elementos de los espacios de funciones de Schmidt son sucesiones infinitas de números complejos, $z = \{z_p\}$ tales que

$$\sum_{p=1}^{\infty} |z_p|^2 < \infty$$

Schmidt introdujo la notación $\|z\|$ para $\{\sum_{p=1}^{\infty} z_p \cdot \hat{z}_p\}^{1/2}$, que recibió más tarde el nombre de norma de z . Siguiendo a Hilbert, Schmidt usó la notación (z, w) para $\sum_{p=1}^{\infty} z_p \cdot w_p$ con lo que $\|z\| = \sqrt{(z, \hat{z})}$ (modernamente se acostumbra definir directamente (z, w) como $\sum_{p=1}^{\infty} z_p \cdot w_p$). Dos elementos z y w del espacio se llaman ortogonales si y sólo si $(z, w) = 0$. Schmidt demostró entonces una forma

generalizada del teorema de Pitágoras, que asegura que si z_1, z_2, \dots, z_n son n elementos del espacio ortogonales dos a dos, entonces

$$w = \sum_{p=1}^n z_p$$

implica que

$$\|w\|^2 = \sum_{p=1}^n \|z_p\|^2$$

de lo cual se sigue que los n elementos del espacio ortogonales dos a dos son linealmente independientes. También obtuvo Schmidt la desigualdad de Bessel para este espacio general, en la forma: si $\{z_n\}$ es una sucesión ortonormal de elementos, de manera que $(z_p, \hat{z}_q) = \delta_{pq}$ y si w es un elemento cualquiera del espacio, entonces

$$\sum_{p=1}^n \|w, \hat{z}_p\|^2 \leq \|w\|^2$$

También puede demostrarse la desigualdad de Schwarz y la desigualdad triangular para la norma.

Se dice que una sucesión de elementos $\{z_n\}$ converge fuertemente a z si $\|z_n - z\|$ tiende a cero, y entonces puede demostrarse que toda sucesión «fuerte» de Cauchy, es decir, toda sucesión para la cual $\|z_p$

- $\|z_q\|$ tienda a cero cuando p y q tienden a ∞ , converge a un cierto elemento z , de forma que el espacio de las sucesiones es completo, lo cual constituye una propiedad de importancia vital.

Schmidt introdujo a continuación el concepto de subespacio (fuertemente) cerrado: un subconjunto A del espacio considerado H se llama subespacio cerrado si es un subconjunto cerrado en el sentido de la convergencia que acabamos de definir, y si además es cerrado algebraicamente, es decir, si dados dos elementos w_1 y w_2 de A entonces también $a_1w_1 + a_2w_2$ pertenece a A , siendo a_1 y a_2 dos números complejos cualesquiera. Se prueba la existencia de tales subespacios cerrados simplemente tomando cualquier sucesión $\{z_n\}$ de elementos linealmente independientes y formando todas las combinaciones lineales finitas de elementos tomados de entre los $\{z_n\}$, entonces la clausura de esta familia de elementos es un subespacio algebraicamente cerrado.

Tomemos ahora un subespacio cerrado cualquiera A ; Schmidt demuestra que si z es un elemento cualquiera del espacio, entonces existen elementos únicos w_1 y w_2 tales que $z = w_1 + w_2$, donde w_1 pertenece a A y w_2 es ortogonal a A , lo que significa que w_2 es ortogonal a todo elemento de A . Este resultado se designa actualmente con el nombre de teorema de la proyección; w_1 es la proyección de z sobre A . Además, $\|w_2\| = \min \|y - z\|$, donde y es un elemento cualquiera de A , y este mínimo se alcanza sólo para $y = w_1$; $\|w_2\|$ recibe el nombre de distancia entre z y A .

Schmidt y Fréchet observaron simultáneamente en 1907 que el espacio de las funciones de cuadrado sumable (con la integral de

Lebesgue) tiene una geometría completamente análoga a la del espacio de Hilbert de las sucesiones. Esta analogía quedó aclarada algunos meses después, cuando Riesz, haciendo uso del teorema de Riesz-Fischer (cap. 45, sec. 4) que establece una correspondencia biunívoca entre las funciones medibles de Lebesgue de cuadrado integrable y las sucesiones de números reales de cuadrado sumable, señaló que puede definirse una distancia en el conjunto L^2 de las funciones de cuadrado sumable, y puede usarse esta distancia para construir una geometría sobre este espacio de funciones. Este concepto de distancia entre dos funciones cualesquiera de cuadrado sumable del espacio L^2 sobre el intervalo $[a, b]$, también fue definido, de hecho, por Fréchet,⁴²⁸ en la forma

$$\sqrt{\int_a^b [f(x) - g(x)]^2 dx} \quad (1)$$

donde la integral está tomada en el sentido de Lebesgue, y considerando que dos funciones que solamente difieren entre sí sobre un conjunto de medida cero, son iguales. El cuadrado de esta distancia recibe también el nombre de desviación cuadrática media de las funciones. El producto escalar de f y g se define de la forma

$$(f, g) = \int_a^b f(x) \cdot g(x) dx$$

y las funciones que cumplen $(f, g) = 0$ se llaman ortogonales. La desigualdad de Schwarz

$$\int_a^b f(x) \cdot g(x) dx \leq \sqrt{\int_a^b f^2 dx} \sqrt{\int_a^b g^2 dx}$$

y otras propiedades que se verifican en el espacio de las sucesiones de cuadrado sumable, se aplican también al espacio de funciones; en particular, esta clase de funciones de cuadrado sumable constituye un espacio completo. Así pues, se puede identificar el espacio de las funciones de cuadrado sumable y el espacio de las sucesiones de cuadrado sumable que son los coeficientes de Fourier de dichas funciones con respecto a un sistema ortonormal completo de funciones fijo.

Por lo que se refiere a espacios de funciones, tenemos que recordar también (cap. 45, sec. 4) los espacios L^p , $1 < p < \infty$, introducidos por Riesz; estos espacios también son completos con la métrica

$$d(f_1, f_2) = \left(\int_a^b |f_1 - f_2|^p dx \right)^{1/p}$$

Aunque volveremos a considerar más adelante el desarrollo de la teoría de espacios abstractos en general, los resultados siguientes se refieren a funcionales y operadores. En su artículo de 1907 ya

mencionado, en el que introducía la métrica o «*écart*» para funciones del espacio L^2 , y en otro artículo del mismo año,⁴²⁹ demostró Fréchet que para todo funcional lineal continuo $U(f)$ definido sobre L^2 , existe una única función $u(x)$ en L^2 , tal que para toda f de L^2 ,

$$U(f) = \int_a^b f(x)u(x)dx$$

resultado que generaliza otro obtenido por Hadamard en 1903.⁴³⁰ En 1909,⁴³¹ Riesz generalizó a su vez este resultado, expresando $U(f)$ como una integral de Stieltjes, es decir

$$U(f) = \int_a^b f(x) du(x)$$

Riesz mismo volvió a generalizar este resultado para funcionales lineales A definidos sobre un espacio L^p , que satisfagan la condición de que para todo f en L^p ,

$$A(f) \leq M \left[\int_a^b |f(x)|^p \right]^{1/p}$$

donde M depende sólo de A ; entonces existe una función $a(x)$ en L^q , única módulo la suma de una función de integral cero, tal que para toda f en L^p

$$U(f) = \int_a^b a(x)f(x) dx \quad (2)$$

Este resultado recibe el nombre de teorema de representación de Riesz.

La parte central del análisis funcional se ocupa de la teoría abstracta de los operadores que aparecen en las ecuaciones diferenciales e integrales; esta teoría unifica la teoría de autovalores para ecuaciones diferenciales e integrales y para transformaciones lineales que actúan sobre un espacio «-dimensional. Un operador de esta clase, como por ejemplo

$$g(x) = \int_a^b k(x,y)f(y)dy$$

donde k es una función dada, hace corresponder la función g a la f y satisface algunas otras condiciones adicionales. Utilizando la notación A para el operador abstracto, y la notación $g = Af$ para la correspondencia asociada, la linealidad significa que

$$A(\lambda_1 f_1 + \lambda_2 f_2) = \lambda_1 A f_1 + \lambda_2 A f_2 \quad (3)$$

donde las λ_i son constantes reales o complejas cualesquiera. La integral indefinida $g(x) = \int_a^x f(t)dt$ y la derivada $f'(x) = Df(x)$ son operadores lineales sobre las correspondientes clases de funciones. La continuidad del operador A significa que si la sucesión de funciones f_n tiende a f en el sentido de la operación límite en el espacio de funciones, entonces Af_n debe tender a Af .

El análogo abstracto de lo que se presenta en el caso de las ecuaciones integrales cuando el núcleo $k(x, y)$ es simétrico, es la propiedad del operador A de ser autoadjunto: si, cualesquiera que sean las funciones f_1 y f_2 , se tiene que

$$(Af_1, f_2) = (f_1, Af_2)$$

donde (Af_1, f_2) representa el producto escalar de dos funciones del espacio, entonces A se llama autoadjunto. En el caso de las ecuaciones integrales, si

$$Af = \int_a^b k(x,y) f(y)dy$$

entonces

$$(Af_1, f_2) = \int_a^b \int_a^b k(x, y) f_1(y) f_2(x) dy dx$$

$$(f_1, Af_2) = \int_a^b \int_a^b k(x, y) f_2(y) f_1(x) dy dx$$

y $(Af_1, f_2) = (f_1, Af_2)$ si el núcleo es simétrico. Para los operadores autoadjuntos arbitrarios, los autovalores son todos reales, y las autofunciones correspondientes a autovalores distintos son ortogonales entre sí.

Un vigoroso impulso inicial a la teoría abstracta de operadores, que es el núcleo del análisis funcional, se debe a Riesz, en su artículo de 1910 en los *Mathematische Annalen*, donde introduce los espacios L^p (cap. 45, sec. 4). Riesz pretendía generalizar en este artículo la resolución de la ecuación integral

$$\Phi(x) - \lambda \int_a^b K(x, t) \Phi(t) dt = f(x)$$

para funciones en los espacios L^p . Riesz consideró la expresión

$$\int_a^b K(x, t) \Phi(t) dt$$

como una transformación actuando sobre la función $\Phi(t)$, la llamó

una transformación funcional y la representó por $T(\Phi(t))$. Además, dado que las $\Phi(t)$ que consideraba Riesz estaban en un cierto espacio L^p , la transformación aplicaba funciones en otras del mismo espacio, o de otro; en particular, una transformación u operador que transforme funciones de L^p en funciones de L^p se llama lineal sobre L^p si satisface (3) y si T está acotado, es decir, si existe una constante M tal que para toda f en L^p que satisfaga

$$\int_a^b |f(x)|^p dx \leq 1$$

se tiene que

$$\int_a^b |T(f(x))|^p dx \leq M^p$$

Posteriormente, al extremo inferior de tales M se le llamó la norma de T , y se representó por $\|T\|$.

Riesz introdujo también el concepto de operador adjunto o traspuesto de T : para toda g en L^q y todo operador T actuando sobre U , se tiene que

$$\int_a^b T(f(x))g(x) dx \leq M^p \quad (4)$$

define un funcional sobre L^p , donde g es una función fija y f varía sobre L^p ; por tanto, por el teorema de representación de Riesz, existe una función $\psi(x)$ en L^q , única salvo una función cuya integral sea cero, tal que

$$\int_a^b T(f(x))g(x) dx = \int_a^b f(x)\Psi(x)dx \quad (5)$$

El adjunto o traspuesto de T , representado por T^* , se define entonces como aquel operador sobre L^q que, para el T fijo depende sólo de g , y asigna a la función g la ψ de la ecuación (5), es decir, $T^*(g) = \psi$. (En notación moderna, T^* satisface $(Tf, g) = (f, T^*g)$. T^* es una transformación lineal sobre L^q y verifica que $\|T^*\| = \|T\|$. Riesz consideró entonces la solución de la ecuación

$$T(\Phi(x)) = f(x) \quad (6)$$

donde T^* es una transformación lineal sobre L^p , la función f es conocida y Φ es desconocida. Demostró que (6) tiene una solución si y sólo si

$$\left| \int_a^b f(x)g(x)dx \right| \leq M \left(\int_a^b |T^*(g(x))|^q dx \right)^{1/q}$$

para toda g en L^q . De esta manera se vio conducido al concepto de

transformación inversa u operador T^{-1} , y al mismo tiempo para T^{*-1} ; con la ayuda del operador adjunto pudo demostrar la existencia de los inversos.

En su artículo de 1910 introduce Riesz la notación

$$\Phi(x) - \lambda K(\Phi(x)) = f(x) \quad (7)$$

donde K representa $\int_a^b K(x,t)^* dt$ y donde $*$ debe reemplazarse por la función sobre la que actúa K ; sus resultados siguientes se limitan a L^2 , donde $K = K^*$. Para tratar el problema de los auto valores de ecuaciones integrales introduce el concepto de continuidad completa de Hilbert, pero formulado ahora para operadores abstractos. Un operador K sobre L^2 se llama completamente continuo si K transforma toda sucesión de funciones débilmente convergente (cap. 45, sec. 4) en otra fuertemente convergente; es decir, si $\{f_n\}$ converge débilmente, entonces $\{K(f_n)\}$ converge fuertemente. Riesz demostró que el espectro de (7) es discreto (es decir, un K simétrico no tiene espectro continuo) y que las autofunciones asociadas a los autovalores son ortogonales.

Otro enfoque de los espacios abstractos, esta vez usando el concepto básico de norma, fue iniciado unos años más tarde también por Riesz;⁴³² sin embargo, la definición general de los espacios normados fue dada de una manera casi simultánea durante los años 1920 a 1922 por Stefan Banach (1892-1945), Hans Hahn (1879-1934), Eduard Helly (1884-1943) y Norbert Wiener (1894-1964). Aunque los trabajos de estos matemáticos se

solapan en buena parte y resulta difícil decidir en cuestiones de prioridad, el hecho es que la obra de Banach es la que ha tenido mayor influencia. Su motivación fue la generalización de las ecuaciones integrales.

La característica esencial de todos los trabajos que hemos mencionado, y del de Banach en particular,⁴³³ era la de construir un espacio dotado de una norma, pero que en general no era definida a partir de un producto escalar. Mientras que en L^2 se tiene que $\|f\| = (f, f)^{1/2}$ en un espacio de Banach general no se puede definir la norma de esta manera, porque no se tiene previamente un producto escalar.

Banach parte de un espacio E cuyos elementos se representan por x, y, z, \dots , mientras que a, b, c, \dots , denotan números reales. Los axiomas que debe satisfacer el espacio E se dividen en tres grupos: el primer grupo consta de trece axiomas que expresan el hecho de que E es un grupo abeliano para la suma, que es cerrado para la multiplicación por un escalar real, y que se verifican las propiedades asociativas y distributivas usuales entre las operaciones con números reales y con elementos de E .

El segundo grupo de axiomas caracteriza lo que se llama una norma definida para los elementos (o vectores) de E . La norma es una función con valores reales definida sobre E , que representaremos por $\|x\|$: para todo número real a y todo vector x de E , la norma tiene las propiedades siguientes:

- a. $\|x\| \geq 0$.
- b. $\|x\| = 0$ si y sólo si $x = 0$.

$$c. \|ax\| = |a| \cdot \|x\|$$

$$d. \|x + y\| \leq \|x\| + \|y\|$$

El tercer grupo consta de un solo axioma, el axioma de completitud, que dice que si $\{x_n\}$ es una sucesión de Cauchy para la norma definida, es decir, si

$$\lim_{n,p \rightarrow \infty} \|x_n - x_p\| = 0$$

entonces existe un elemento x en E tal que

$$\lim_{n \rightarrow \infty} \|x_n - x\| = 0$$

Un espacio que satisface estos tres grupos de axiomas recibe el nombre de espacio de Banach o espacio vectorial normado completo. Aunque un espacio de Banach es más general que un espacio de Hilbert, porque no se supone que esté definido en él un producto escalar entre vectores que defina a su vez la norma, como consecuencia de ello se pierde el concepto fundamental de vectores ortogonales en los espacios de Banach que no sean además de Hilbert. Los grupos de condiciones primero y tercero se verifican también en los espacios de Hilbert, pero el segundo grupo es más débil que las condiciones que verifica la norma en un espacio de Hilbert. Los espacios de Banach incluyen a los espacios L^p , los espacios de funciones continuas, los espacios de funciones medibles

acotadas, entre otros, siempre que se defina sobre ellos una norma adecuada.

Utilizando el concepto de norma pudo demostrar Banach una serie de resultados conocidos para estos espacios. Uno de los teoremas clave dice lo siguiente: sea $\{x_n\}$ una sucesión de elementos de E tales que

$$\sum_{p=1}^{\infty} \|x_p\| < \infty$$

entonces $\sum_{p=1}^{\infty} x_p$ converge según la norma a un elemento x de E .

Después de demostrar una serie de teoremas, pasó a estudiar Banach los operadores definidos sobre el espacio E pero con valores en otro espacio también de Banach E_1 . Un operador F se llama continuo en x_0 con respecto a un conjunto A , si $F(x)$ está definido para todo x de A , x_0 pertenece a A y al conjunto derivado de A , y $\{x_n\}$ es una sucesión arbitraria de A con límite x_0 , entonces $F(x_n)$ converge a $F(x_0)$. También definió Banach la continuidad uniforme de un operador F con respecto a un conjunto A , para pasar a continuación al estudio de las sucesiones de operadores: la sucesión $\{F_n\}$ de operadores se llama convergente según la norma al operador F sobre un conjunto A , si para todo x de A , $\lim_{n \rightarrow \infty} F_n(x) = F(x)$.

Una clase importante de operadores introducida por Banach es la de los operadores aditivos continuos: un operador F se llama aditivo si para todo x y para todo y , $F(x + y) = F(x) + F(y)$. Un operador aditivo continuo verifica la propiedad de que para todo número real

a , $F(ax) = aF(x)$; si F es aditivo y continuo en un punto de F , entonces es continuo en todos y además es acotado, es decir, existe una constante M , que depende sólo de F , tal que para todo x en E se tiene que $\|F(x)\| \leq \|Mx\|$. Otro teorema asegura que si $\{F_n\}$ es una sucesión de operadores aditivos continuos, y si F es un operador aditivo tal que para todo x , $\lim_{n \rightarrow \infty} F_n(x) = F(x)$, entonces F es continuo y existe un M tal que para todo n se verifica que $\|F_n(x)\| \leq M \|x\|$.

En este artículo demuestra Banach algunos teoremas sobre la resolución de ecuaciones integrales, formulados de una manera abstracta. Si F es un operador continuo que aplica el espacio E en sí mismo, y si existe un número M , con $0 < M < 1$, tal que para todo x' y todo x'' en E , $\|F(x') - F(x'')\| \leq M \|x' - x''\|$, entonces existe un único elemento x en E que satisface $F(x) = x$. Aún más importante es el teorema siguiente: consideremos la ecuación

$$x + hF(x) = y \quad (8)$$

donde y es una función conocida de E , F es un operador aditivo continuo de E en E , y h un número real. Sea M el extremo inferior de los números M' que satisfacen $\|F(x)\| \leq M' \|x\|$ para todo x ; entonces para todo y , y todo valor de h que verifique $|hM| < 1$, existe una función x que satisface la ecuación (8) y además

$$x = y + \sum_{n=1}^{\infty} (-1)^n h^n F^{(n)}(y)$$

donde $F^{(n)}(y) = F(F^{(n-1)}(y))$. Este resultado es una forma del teorema del radio espectral, y una generalización del método de Volterra para resolver ecuaciones integrales.

En 1929⁴³⁴ introdujo Banach otro concepto importante del análisis funcional: el concepto de espacio dual o adjunto de un espacio de Banach, idea que también fue introducida de manera independiente por Hahn,⁴³⁵ pero el trabajo de Banach era más sistemático. Este espacio dual es el espacio de todos los funcionales lineales continuos acotados sobre el espacio dado; la norma para este espacio de funcionales se define como las cotas de los funcionales, con la cual resulta ser un espacio vectorial normado completo, es decir, un espacio de Banach. En realidad, el trabajo de Banach generaliza aquí resultados de Riesz sobre los espacios L^p y L^q , siendo $q = p/(p-1)$, porque el espacio L^q es equivalente al dual del espacio L^p en el sentido de Banach. La conexión con el trabajo de Banach se pone en evidencia por el teorema de representación de Riesz (2). En otras palabras, el espacio dual de Banach tiene la misma relación con el espacio de Banach dado E , que L^q tiene con L^p .

Banach parte de la definición de un funcional lineal continuo, es decir, una función continua con valores reales definida sobre el espacio E , y demuestra que todo funcional de tal clase está acotado. Un teorema básico, que generaliza otro debido a Hahn, es el que se conoce actualmente como teorema de Hahn-Banach: sea p un funcional con valores reales definido sobre un espacio vectorial normado completo R , que verifica, para todo x e y en R

$$a) p(x + y) \leq p(x) + p(y).$$

$$b) p(\lambda c) = \lambda p(x) \quad \text{para } \lambda \geq 0.$$

Entonces existe un funcional aditivo f sobre R que verifica

$$-p(-x) \leq f(x) \leq p(x)$$

para todo x en R . Hay toda una serie de teoremas sobre el conjunto de los funcionales continuos definidos sobre R .

Las investigaciones sobre funcionales condujeron al concepto de operador adjunto. Sean R y S dos espacios de Banach, y sea U un operador lineal continuo de R en S ; sean R^* y S^* los conjuntos de funcionales lineales acotados definidos respectivamente sobre R y S . Entonces U induce una aplicación de S^* en R^* , de la manera siguiente: si g es un elemento de S^* , entonces $g(U(x))$ está bien definido para todo x en R . Por la linealidad de U y de g este es también un funcional lineal sobre R , es decir, $g(U(x))$ es un elemento de R^* . Dicho en otras palabras, si $U(x) = y$ entonces $g(y) = f(x)$, donde f es un funcional de R^* . La aplicación inducida U^* de S^* en R^* recibe el nombre de adjunta de la U .

Utilizando este concepto demuestra Banach que si U^* tiene un inverso continuo, entonces $y = U(x)$ tiene solución para todo y en S . También, si $f = U^*(g)$ tiene solución para todo f en R^* , entonces U^{-1} existe y es continuo sobre la imagen de U , siendo la imagen de U en S el conjunto de todos los y para los cuales existe un g en S^* con la

propiedad de que $g(y) = 0$ siempre que $U^*(g) = 0$. (Esta última afirmación es una versión generalizada del teorema de la alternativa de Fredholm).

Banach aplicó su teoría de operadores adjuntos a los operadores de Riesz, que había introducido este último en su artículo de 1918. Estos son operadores U de la forma $U = I - \lambda V$ donde I es el operador identidad y V es un operador completamente continuo. Entonces puede aplicarse la teoría abstracta al espacio L^2 de funciones definidas sobre $[0,1]$ y a los operadores

$$U_\lambda(x) = x(s) - \lambda \int_0^1 K(s,t)x(t)dt$$

donde

$$\int_0^1 \int_0^1 K^2(s,t) ds dt < \infty$$

La teoría general, aplicada a

$$y(s) = x(s) - \lambda \int_0^1 K(s,t)x(t) dt \quad (9)$$

y a su ecuación traspuesta asociada

$$f(s) = g(s) - \lambda \int_0^1 K(t,s)g(t) dt \quad (10)$$

nos dice que si λ_0 es un autovalor de (9), entonces λ_0 es también un autovalor de (10) y recíprocamente. Además, (9) tiene un número finito de autofunciones linealmente independientes asociadas a λ_0 , y lo mismo ocurre con (10). También ocurre que (9) no tiene solución para todo λ cuando $\lambda = \lambda_0$; de hecho, una condición necesaria y suficiente para que (9) tenga una solución, es que se satisfagan las n condiciones

$$\int_0^1 y(t)g_0^{(p)}(t) dt = 0 \quad p = 1, 2, \dots, n$$

donde g_0^1, \dots, g_0^n es un conjunto de soluciones linealmente independientes de

$$g(s) - \lambda_0 \int_0^1 K(t,s)g(t) dt = 0$$

4. La axiomatización de los espacios de Hilbert

La teoría de espacios de funciones y operadores, durante los años veinte, parecía conducir sólo a la abstracción por la abstracción. Incluso Banach no hizo ningún uso concreto de su obra. Este estado de cosas hizo observar más tarde a Hermann Weyl: «No fue

mérito alguno, sino favor de la fortuna el que se descubriese, a partir de 1923, que la teoría espectral del espacio de Hilbert era el instrumento matemático adecuado a la mecánica cuántica». La investigación en mecánica cuántica mostró que los observables de un sistema físico se pueden representar por operadores lineales simétricos en un espacio de Hilbert, y que los autovalores y autovectores (autofunciones) del operador particular que representa la energía son los niveles de energía de un electrón en el átomo y los correspondientes estados cuánticos estacionarios del sistema. Las diferencias entre dos autovalores dan las frecuencias de los cuantos de luz emitidos, y definen así el espectro de radiación de la sustancia. En 1926 Erwin Schrödinger presentó su teoría cuántica basada en ecuaciones diferenciales, y demostró la identidad de esta última teoría con la teoría de matrices infinitas de Werner Heisenberg (1925), que éste había aplicado a la teoría cuántica. Sin embargo, faltaba una teoría general que unificase los trabajos de Hilbert y la teoría de autofunciones para ecuaciones diferenciales.

El uso de operadores en la teoría cuántica estimuló el trabajo en una teoría abstracta del espacio de Hilbert y sus operadores, tarea que fue emprendida en primer lugar por John von Neumann (1903-1957) en 1927. Su planteamiento incluye a la vez el espacio de las sucesiones de cuadrado sumable y el espacio L^2 de funciones definidas en un intervalo común.

En dos largos artículos⁴³⁶ presentó von Neumann un tratamiento axiomático del espacio de Hilbert y de los operadores sobre estos espacios. Aunque puede detectarse el origen de los axiomas de von

Neumann en los trabajos de Norbert Wiener, Weyl y Banach, la obra de von Neumann era más completa y ejerció una mayor influencia; su objetivo principal era formular una teoría general de auto valores para una amplia gama de operadores llamados hermíticos.

Von Neumann introduce un espacio L^2 de funciones complejas, medibles y de cuadrado integrable, definidas sobre un conjunto medible arbitrario E del plano complejo, así como el espacio análogo o espacio de sucesiones complejas, «es decir, el conjunto de todas las sucesiones de números complejos a_1, a_2, \dots con la propiedad de que

$$\sum_{p=1}^{\infty} |a_p|^2 < \infty$$

El teorema de Riesz-Fischer permitía demostrar entonces que existe una correspondencia biunívoca entre las funciones del primer espacio y las sucesiones del segundo espacio, definida de la forma siguiente. Elijamos en el espacio de funciones una sucesión ortonormal completa de funciones $\{\Phi_n\}$; entonces si f es una función cualquiera del espacio, los coeficientes de Fourier de f en su desarrollo con respecto a las $\{\Phi_n\}$ forman una sucesión del espacio de sucesiones, y recíprocamente, si una parte de una sucesión así, entonces hay una única función de L^2 (salvo una función cuya integral sea cero) que tiene esta sucesión como la de sus coeficientes de Fourier con respecto a las $\{\Phi_n\}$.

Además, si se define el producto escalar (f, g) en el espacio de

funciones por

$$(f, g) = \int_E f(z) \overline{g(z)} dz$$

donde $\overline{g(z)}$ es el complejo conjugado de $g(z)$, y en el espacio de sucesiones el producto escalar de dos sucesiones a y b por

$$(a, b) = \sum_{p=1}^{\infty} a_p \bar{b}_p$$

entonces, si f corresponde a a y g a b , se tiene $(f, g) = (a, b)$.

Un operador hermítico R sobre uno de estos espacios se define como un operador lineal con la propiedad de que, para todas las f y g pertenecientes a su dominio, $(Rf, g) = (f, Rg)$; análogamente, en el espacio de las sucesiones debe verificarse $(Ra, b) = (a, Rb)$.

La teoría de Von Neumann ofrece un tratamiento axiomático para ambos espacios a la vez, el de funciones y el de sucesiones; los axiomas que propone son los siguientes:

- A. H es un espacio vectorial; es decir, están definidas sobre H una suma y una multiplicación por escalares, de manera que si f_1, f_2 son elementos de H y a_1 y a_2 son números complejos cualesquiera, entonces $a_1 f_1 + a_2 f_2$ es otro elemento de H .
- B. Está definido sobre H un producto escalar, o función definida para todas las parejas de vectores f y g , y con valores complejos, representada por (f, g) , con las propiedades:

- a) $(af, g) = a(f, g)$
- b) $(f_1 + f_2, g) = (f_1, g) + (f_2, g)$
- c) $(f, g) = (g, f)$
- d) $(f, f) > 0$
- e) $(f, f) = 0$ si y sólo si $f = 0$.

Dos elementos f y g se llaman ortogonales si $(f, g) = 0$; la norma de f , representada por $\|f\|$, se define como $\sqrt{(f, f)}$ y $\|f - g\|$ define una métrica sobre el espacio H .

C. Con la métrica que acabamos de definir, H es separable. Es decir, existe en H un subconjunto denso numerable con respecto a la métrica $\|f - g\|$

D. Para todo entero positivo n , hay en H n vectores linealmente independientes.

E. H es completo. Es decir, si la sucesión $\{f_n\}$ es tal que $\|f_n - f_m\|$ tiende a cero cuando m y n tienden a ∞ , entonces hay un f en H tal que $\|f - f_n\|$ tiende a cero cuando n tiende a ∞ . (Esta convergencia es equivalente a la convergencia fuerte.)

De estos axiomas se deducen inmediatamente algunas propiedades sencillas, como son la desigualdad de Schwarz

$$\|(f, g)\| \leq \|f\| \cdot \|g\|$$

el hecho de que todo conjunto ortonormal completo de elementos de H tiene que ser numerable, la desigualdad de Parseval, es decir,

$$\sum_{p=1}^{\infty} \|(f, \Phi_p)\|^2 \geq \|f\|^2$$

A continuación von Neumann pasa a estudiar los subespacios vectoriales de H y los operadores proyección. Si M y N son subespacios cerrados de H , entonces se define $M - N$ como el conjunto de todos los elementos de M que son ortogonales a todos los elementos de N . Entonces, el teorema de la proyección dice: sea M un subespacio cerrado de H ; entonces cualquier vector f de H puede expresarse de una y sólo una manera como $f = g + h$, donde g está en M y h en $H - M$. El operador proyección asociado a M , P_M , se define de la forma $P_M(f) = g$; es decir, es el operador, definido sobre todo H , que proyecta cada elemento f en su componente según M .

En el segundo de los artículos citados introduce von Neumann dos topologías sobre H , la fuerte y la débil. La topología fuerte es simplemente la topología métrica definida a partir de la norma. La topología débil, en la que no entraremos aquí, da el sistema de entornos asociados a la convergencia débil.

Von Neumann presenta un buen número de resultados sobre operadores en el espacio de Hilbert. Se presupone que todos los operadores estudiados son lineales, y la integración se entiende generalmente en el sentido de Lebesgue-Stieltjes. Una transformación u operador lineal acotado es uno que transforma elementos de un espacio de Hilbert en otro, satisface la condición de linealidad (3), y es acotado, es decir, existe un número M tal que

para todo f del espacio sobre el que actúa el operador R ,

$$\|R(f)\| \leq M \|f\|$$

El ínfimo de los M posibles recibe el nombre de módulo (norma) de R . Esta última condición es equivalente a la continuidad del operador. También es fácil ver que la continuidad en un único punto, junto con la linealidad, es suficiente para garantizar la continuidad en todos los puntos, y por tanto la acotación.

Está también el operador adjunto de R , R^* . Para los operadores hermíticos $R = R^*$ es decir, R es autoadjunto. Si $RR^* = R^*R$ entonces se dice que R es normal. Si $RR^* = R^*R = I$, donde I es el operador identidad, entonces R es análogo a una transformación ortogonal, y se llama unitario; para todo R unitario, $\|R(f)\| = \|f\|$.

Otro de los resultados de von Neumann afirma que si R es hermítico sobre el espacio de Hilbert y satisface la condición de cierre débil, a saber, que si f_n tiende a f y $R(f_n)$ tiende a g entonces $R(f) = g$, y entonces R es acotado.

Además, si R es hermítico, entonces el operador $I - \lambda R$ tiene inverso para todo λ real o complejo exterior a un intervalo (m, M) del eje real en el que están los valores de $(R(f), f)$ cuando $\|f\| = 1$. Otro resultado más importante es el de que a todo operador lineal acotado y hermítico le corresponden otros dos operadores E_- y E_+ (*Einzeltransformationen*) con las siguientes propiedades:

$$a) \quad E_- E_- = E_-, \quad E_+ E_+ = E_+, \quad I = E_- + E_+.$$

- b) E_- y E_+ conmutan, y también conmutan con todo operador que conmute con R .
- c) RE_- y RE_+ son respectivamente negativo y positivo (RE_+ es positivo si $(RE_+(f), f) \geq 0$).
- d) Para todo f tal que $R(f) = 0$, $E_-(f) = 0$ y $E_+(f) = 1$.

Von Neumann estableció también una importante relación entre los operadores hermíticos y unitarios, la de que si U es unitario y R es hermítico, entonces $U = e^{iR}$. Más adelante también generalizó su teoría a operadores no acotados. Pese a que sus contribuciones en este campo, como las de otros matemáticos, fueron de la mayor importancia, una exposición detallada de ellas nos llevaría demasiado lejos y nos introduciría en desarrollos recientes más complejos.

Se han hecho, y se siguen haciendo, aplicaciones del análisis funcional al problema generalizado de los momentos, a la mecánica estadística, a los teoremas de existencia y unicidad en ecuaciones en derivadas parciales, a los teoremas del punto fijo, y a muchos otros. Hoy día el análisis funcional juega un papel importante en el cálculo de variaciones y en la teoría de representaciones de grupos continuos compactos, así como en álgebra, cálculo aproximado, topología y teoría de funciones de variable real. A pesar de esta diversidad de aplicaciones, hay una deplorable ausencia de aplicaciones nuevas a los grandes problemas del análisis clásico. Este fracaso desilusionó a los fundadores del análisis funcional.

Bibliografía

- Bernkopf, M.: «The Development of Function Spaces with Particular Reference to their Origins in Integral Equation Theory». *Archive for History of Exact Sciences*, 3, 1966, 1-96. - «A History of Infinite Matrices». *Archive for History of Exact Sciences*, 4, 1968, 308-358.
- Bourbaki, N.: *Elementos de historia de las matemáticas*, Madrid, Alianza, 1976.
- Dresden, Arnold: «Some Recent Work in the Calculus of Variations». *Amer. Math. Soc. Bull.*, 32, 1926, 475-521.
- Fréchet, M.: *Notice sur les travaux scientifiques de M. Maurice Fréchet*, Hermann, 1933.
- Hellinger, E., y Toeplitz: «Integralgleichungen und Gleichungen mit unendlichvielen Unbekannten». *Encyk. der Math. Wiss.*, B. G. Teubner, 1923-1927, vol. 2, part. III, 2.^a half, 1335-1597.
- Hildebrandt, T. H.: «Linear Functional Transformations in General Spaces». *Amer. Math. Soc. Bull.*, 37, 1931, 185-212.
- Levy, Paul: «Jacques Hadamard, sa vie et son Oeuvre». *L'Enseignement Mathématique* (2), 13, 1967, 1-24.
- McShane, E. J.: «Recent Developments in the Calculus of Variations». *Amer. Math. Soc. Semicentennial Publications*, 2, 1938, 69-97.
- Neumann, John von: *Collected Works*, Pergamon Press, 1961, vol. 2.
- Sanger, Ralph G.: «Functions of Lines and the Calculus of Variations». *University of Chicago Contributions to the Calculus*

of Variations for 1931-1932, University of Chicago Press, 1933, 193-293.

- Tonelli, L.: «The Calculus of Variations». *Amer. Math. Soc. Bull.*, 31, 1925, 163-172.
- Volterra, Vito: *Opere matematiche*, 5 vols., Accademia Nazionale dei Lincei, 1954-1962.

Capítulo 47

La teoría de series divergentes

Es realmente una extraña circunstancia de nuestra ciencia el que aquellas series que a principios de siglo se suponía que habían de ser desterradas de una vez y para siempre de la matemática rigurosa aparezcan, a su final, llamando a la puerta para ser readmitidas.

James Pierpont

Esta serie es divergente, por tanto algo podremos hacer con ella.

Oliver Heaviside

Contenido:

- 1. Introducción*
- 2. La utilización informal de las series divergentes*
- 3. La teoría formal de las series divergentes*
- 3. La teoría formal de las series asintóticas*
- 4. El problema de la sumabilidad de series divergentes*

Bibliografía

1. Introducción

La toma en consideración desde finales del siglo XIX en adelante de un tema como el de las series divergentes indica cuán radicalmente los matemáticos han revisado su propia consideración de la naturaleza de la matemática. Mientras que en la primera parte del siglo XIX aceptaron la exclusión de las series divergentes sobre la base de que la matemática se encontraba restringida por alguna necesidad intrínseca o por los dictados de la naturaleza a una clase fija de conceptos correctos, a finales del siglo reconocían su libertad para aceptar cualquier idea que pareciera presentar alguna utilidad. Cabe recordar que las series divergentes se utilizaron a lo largo de todo el siglo XVIII, con un reconocimiento más o menos consciente de su divergencia, porque suministraban aproximaciones útiles de funciones tomando unos pocos términos solamente. A partir de la aparición de la matemática rigurosa con Cauchy, la mayor parte de los matemáticos siguieron sus criterios y rechazaron las series divergentes por desconfiar de ellas. Sin embargo, unos pocos matemáticos (cap. 40, sec. 7) continuaron defendiendo las series divergentes porque estaban convencidos de su utilidad, bien para el cálculo de funciones o como representación analítica de las funciones de las que se derivaban. Otros las defendían como método de descubrimiento. Así, por ejemplo, De Morgan⁴³⁷ decía que *«Tenemos que admitir que muchas series son tales que no podemos utilizarlas de momento con seguridad, excepto como método de descubrimiento, cuyos resultados tendrán que ser comprobados*

posteriormente, y sin duda incluso el enemigo más acérrimo de las series divergentes hace este uso de ellas en privado».

Los astrónomos siguieron utilizando series divergentes, después incluso de que fueran excluidas, por las necesidades de su ciencia a efectos de cálculo. Dado que unos pocos de los primeros términos de tales series daban una aproximación numérica útil, los astrónomos solían ignorar el hecho de que la serie completa era divergente, mientras que los matemáticos, interesados en el comportamiento no de los primeros diez o veinte términos, sino de la serie total, no podían aceptarlas únicamente a causa de su utilidad.

Sin embargo, como ya hemos hecho observar (cap. 40, sec. 7) tanto Abel como Cauchy no dejaron de ser conscientes de que, al excluir las series divergentes, estaban eliminando algo útil. Cauchy no sólo continuó usándolas (véase más abajo) sino que escribió incluso un artículo con el título «Sur l'emploi légitime des séries divergentes»,⁴³⁸ en el que, hablando de la serie de Stirling para $\log T(x)$ o $\log m!$ (cap. 20, sec. 4), señala que la serie, aun siendo divergente para todos los valores de x , puede ser utilizada para calcular $\log T(x)$ para x positivo y suficientemente grande. De hecho, Cauchy demuestra que, fijado el número n de términos tomados, el error absoluto cometido al sumarlos es menos que el valor absoluto del siguiente término, y el error se va haciendo más pequeño según x aumenta. Cauchy trató de explicarse por qué era tan buena la aproximación dada por la serie, pero no tuvo éxito.

La utilidad de las series divergentes terminó por convencer a los

matemáticos de que debían tener alguna característica que, una vez aislada, mostraría por qué daban tan buenas aproximaciones. Como decía Oliver Heaviside en el segundo volumen de su *Electromagnetic Theory* (1899), «Debo decir unas pocas palabras sobre la diferenciación generalizada y las series divergentes... No es fácil levantar entusiasmo alguno una vez que ha sido enfriado artificialmente por los aguafiestas de los rigoristas... Tendrá que haber una teoría de series divergentes, o digamos una teoría más general de funciones que la actual, que incluya las series convergentes y divergentes en una totalidad armoniosa». Heaviside no sabía, cuando hizo esta observación, que ya se habían dado los primeros pasos para ello.

La buena disposición de los matemáticos para dedicarse al estudio de las series divergentes se vio sin duda reforzada por otra circunstancia que había ido penetrando de manera gradual en la atmósfera matemática de la época: la geometría no euclídea y las nuevas álgebras. Los matemáticos comenzaron a darse cuenta lentamente de que la matemática es obra del hombre y que la definición de convergencia de Cauchy ya no podía considerarse como una necesidad impuesta por algún poder sobrehumano. A finales del siglo XIX consiguieron aislar la propiedad esencial de aquellas series divergentes que dan aproximaciones útiles a funciones. A dichas series las llamó Poincaré asintóticas, aunque durante todo el siglo recibieron el nombre de semiconvergentes, término introducido por Legendre en su *Essai sur la théorie des nombres* (1798, p. 13), y utilizado también para las series oscilantes.

La teoría de series divergentes se ocupa de dos cuestiones importantes. La primera es la que ya hemos mencionado brevemente, a saber, que algunas de estas series pueden, con un número fijo de términos aproximar una función cada vez mejor según aumenta la variable. De hecho Legendre, en su *Traité des fonctions elliptiques* (1825-1828), ya había caracterizado tales series por la propiedad de que el error cometido al sumar hasta un término cualquiera es del orden del primer término omitido. La segunda cuestión de la teoría de las series divergentes es el concepto de sumabilidad. Es posible definir la suma de una serie de manera completamente nueva y que dé sumas finitas para series divergentes en el sentido de Cauchy.

2. La utilización informal de las series divergentes

Ya hemos tenido ocasión de hablar de la obra de matemáticos del siglo XVIII que utilizaron tanto series convergentes como divergentes. Durante el siglo XIX, antes y después de que Cauchy prohi-biera las series divergentes, hubo matemáticos y físicos que siguieron usándolas. Una aplicación nueva consistió en el cálculo de integrales por desarrollo en serie. Naturalmente los matemáticos no eran conscientes en esa época de que lo que estaban calculando eran desarrollos en series asintóticas completas o los primeros términos de desarrollos de integrales en series asintóticas.

El cálculo asintótico de integrales se remonta a Laplace por lo menos. En su *Théorie analytique des probabilités* (1812)⁴³⁹ Laplace obtiene integrando por partes el desarrollo de la función error

$$\operatorname{Erfc}(T) = \int_T^{\infty} e^{-t^2} dt = \frac{e^{-T^2}}{2T} \left\{ 1 - \frac{1}{2T^2} + \frac{1 \times 3}{(2T^2)^2} - \frac{1 \times 3 \times 5}{(2T^2)^3} + \dots \right\}$$

Hace notar que la serie es divergente, pero la utiliza para calcular $\operatorname{Erfc}(T)$ para valores grandes de T .

En el mismo libro ⁴⁴⁰ también señala Laplace que

$$\int \phi(x) \{u(x)\}^2 dx$$

para valores de s grandes depende de los valores de $u(x)$ próximos a sus puntos estacionarios, es decir, de los valores de x para los que $u'(x) = 0$. Laplace utiliza esta observación para demostrar que

$$s! \sim s^{s+1/2} e^{-s} \sqrt{2\pi} \left(1 + \frac{1}{12s} + \frac{1}{288s^2} + \dots \right)$$

resultado que también puede obtenerse a partir de la aproximación de Stirling de $\log s!$ (cap. 20, sec. 4).

En su *Théorie analytique des probabilités*,⁴⁴¹ tuvo ocasión Laplace de estudiar integrales de la forma

$$f(x) = \int_a^b g(t) e^{xh(t)} dt \quad (1)$$

donde g puede ser una función compleja, h y t son reales, y x es positivo y grande. Laplace observó que la contribución principal a la integral proviene del entorno inmediato de aquellos puntos del dominio de integración donde $h(t)$ alcanza su máximo absoluto. Esta contribución de la que habla Laplace es el primer término de lo que hoy llamaríamos un desarrollo en serie asintótica de la integral. Si $h(t)$ tiene exactamente un máximo en $t = a$, entonces el resultado de Laplace se reduce a

$$f(x) \sim g(a)e^{xh(a)} \sqrt{\frac{-\pi}{2xh''(a)}}$$

cuando x tiende a ∞ .

Si en lugar de (1) la integral que hay que calcular es

$$f(x) = \int_a^b g(t)e^{ixh(t)} dt \quad (2)$$

para t y x reales y x grande, entonces $|e^{ixh(t)}|$ es constante y no se puede aplicar el método de Laplace. En este caso es aplicable, sin embargo, un método vislumbrado por Cauchy en el más importante de sus artículos sobre la propagación de las ondas,⁴⁴² hoy llamado principio de la fase estacionaria. Este principio afirma que la contribución más importante a la integral proviene del entorno inmediato de los puntos estacionarios de $h(t)$, es decir, de aquellos

en que $h'(t) = 0$. Este principio resulta intuitivamente razonable porque podemos considerar el integrando como una corriente oscilante u onda de amplitud $|g(t)|$; si t es el tiempo, entonces la velocidad de la onda es proporcional a $xh'(t)$ y si $h'(t) \neq 0$ esta velocidad crece indefinidamente cuando x tiende a infinito. Las oscilaciones son tan rápidas entonces que durante un periodo completo $g(t)$ es aproximadamente constante y $xh(t)$ es aproximadamente lineal, de manera que la integral sobre un periodo completo se anula. Este razonamiento falla para los valores de t tales que $h'(t) = 0$. Así pues, lo más verosímil es que los puntos estacionarios de $h(t)$ den la contribución principal al valor asintótico de $f(x)$. Si r es un valor de t en el que $h'(t) = 0$ y $h''(t) < 0$, entonces

$$f(x) \sim \sqrt{\frac{2\pi}{xh''(r)}} g(r) e^{ixh(r) + i\pi/4}$$

cuando x tiende a infinito.

Este principio fue utilizado por Stokes para calcular la integral de Airy (véase más adelante) en un artículo de 1856⁴⁴³ y lo formuló explícitamente Lord Kelvin.⁴⁴⁴ Sin embargo, la primera demostración satisfactoria de dicho principio se debe a George N. Watson (1886-1965).⁴⁴⁵

Durante las primeras décadas del siglo XIX, tanto Cauchy como Poisson calcularon muchas integrales dependientes de un parámetro por desarrollo en serie de potencias del parámetro. En el caso de Poisson las integrales aparecían en problemas geofísicos de

conducción del calor y de transmisión de vibraciones elásticas, mientras que Cauchy estudiaba problemas de ondas acuáticas, óptica y astronomía. Por ejemplo, trabajando sobre la difracción de la luz⁴⁴⁶ da Cauchy expresiones en series divergentes de las integrales de Fresnel

$$\int_0^m \cos\left(\frac{\pi}{2}z^2\right) dz = \frac{1}{2} - N \cos\frac{\pi}{2}m^2 + M \operatorname{sen}\frac{\pi}{2}m^2$$

$$\int_0^m \operatorname{sen}\left(\frac{\pi}{2}z^2\right) dz = \frac{1}{2} - M \cos\frac{\pi}{2}m^2 - N \operatorname{sen}\frac{\pi}{2}m^2$$

donde

$$M = \frac{1}{m\pi} - \frac{1 \times 3}{m^5\pi^3} + \frac{1 \times 3 \times 5 \times 7}{m^9\pi^5} - \dots$$

$$N = \frac{1}{m^3\pi^2} - \frac{1 \times 3 \times 5}{m^7\pi^4} + \dots$$

A lo largo del siglo XIX se inventaron otros diversos métodos para calcular integrales, tal como el de descenso más rápido. La teoría completa de todos estos métodos y la comprensión correcta de a qué conducía la aproximación, fueran términos aislados o series completas, tuvo que esperar a la creación de la teoría de series asintóticas.

Muchas de las integrales que se desarrollaron por los métodos

anteriores aparecieron por primera vez como soluciones de ecuaciones diferenciales. Otro uso de las series divergentes fue para resolver directamente ecuaciones diferenciales, uso que puede hacerse remontar al menos a la obra de Euler⁴⁴⁷ donde, al tratar de resolver el problema de la cuerda vibrante no uniforme (cap. 22, párr. 3) da la solución en serie asintótica de una ecuación diferencial ordinaria que esencialmente se reduce a la ecuación de Bessel de orden $r/2$, para r entero.

Jacobi⁴⁴⁸ dio el desarrollo asintótico de $J_n(x)$ para x grande:

$$J_n(n) \sim \left(\frac{2}{\pi x}\right)^{\frac{1}{2}} \left[\cos\left(x - \frac{n\pi}{2} - \frac{\pi}{4}\right) \left\{ 1 - \frac{(4n^2 - 1^2)(4n^2 - 3^2)}{2!(8x)^2} + \dots \right\} - \right. \\ \left. \sin\left(x - \frac{n\pi}{2} - \frac{\pi}{4}\right) \left\{ \frac{4n^2 - 1^2}{1!8x} - \frac{(4n^2 - 1^2)(4n^2 - 3^2)(4n^2 - 5^2)}{3!(8x)^2} \right\} \right]$$

Liouville introdujo una aplicación algo distinta de las series divergentes a la resolución de ecuaciones diferenciales. Buscaba soluciones aproximadas⁴⁴⁹ de la ecuación diferencial

$$\frac{d}{dx} \left(p \frac{dy}{dx} \right) + (\lambda^2 q_0 + q_1)y = 0 \quad (3)$$

donde p , q_0 y q_1 son funciones positivas de x , λ un parámetro y se busca la solución para $a \leq x \leq b$. Aquí, en contra de lo que ocurría en los problemas de contorno donde se buscaban valores discretos de λ , Liouville estaba interesado en conseguir alguna forma

aproximada de y para valores grandes de λ . Para ello introduce las variables

$$t = \int_{x_0}^x \left(\frac{q_0}{p}\right)^{1/2} dx \quad w = (q_0 p)^{1/4} y \quad (4)$$

obteniendo

$$\frac{d^2 w}{dt^2} + \lambda^2 w = r w \quad (5)$$

donde

$$r = (q_0 p)^{-1/4} \frac{d^2}{dx^2} (q_0 p)^{1/4} - \frac{q_1}{q_0}$$

y a continuación utiliza un proceso que se reduce, en términos modernos, a la resolución por aproximaciones sucesivas de una ecuación integral del tipo de Volterra, que es la siguiente:

$$w(t) = c_1 \cos \lambda t + c_2 \operatorname{sen} \lambda t + \int_{t_0}^t \frac{\operatorname{sen} \lambda(t-s)}{\lambda} r(s) w(s) ds$$

Liouville afirma ahora que, para valores suficientemente grandes de λ , la primera aproximación a las soluciones de (5) sería

$$w \sim c_1 \cos \lambda t + c_2 \operatorname{sen} \lambda t \quad (6)$$

Utilizando ahora los valores de w y de t dados por (4) para obtener la solución aproximada de (3), resulta de (6) que

$$y \sim c_1 \frac{1}{(q_0 p)^{1/4}} \cos \left\{ \lambda \int_{x_0}^x \left(\frac{q_0}{p} \right)^{1/2} dx \right\} + \quad (7)$$

$$+ c_2 \frac{1}{(q_0 p)^{1/4}} \operatorname{sen} \left\{ \lambda \int_{x_0}^x \left(\frac{q_0}{p} \right)^{1/2} dx \right\}$$

Aunque Liouville no lo sabía, lo que había obtenido era el primer término de una solución de (3) en serie asintótica, para valores grandes de λ .

Green⁴⁵⁰ utilizó el mismo método en el estudio de la propagación de ondas en un canal. Este método se ha visto ligeramente generalizado para ecuaciones de la forma

$$y'' + \lambda^2 q(x, \lambda) y = 0 \quad (8)$$

donde λ es un parámetro positivo grande y x puede ser real o complejo. Las soluciones suelen expresarse como

$$y \sim q^{-1/4} e^{\pm i \lambda \int_0^q q^{1/2} dx} \left[1 + O\left(\frac{1}{\lambda}\right) \right] \quad (9)$$

El término de error $O(1/\lambda)$ implica que la solución exacta incluiría un término $F(x,\lambda)/\lambda$ donde $|F(x, \lambda)|$ está acotado para todo x del dominio considerado y para $\lambda > \lambda_0$. La forma del término de error es válida para un dominio restringido del plano complejo x . Ni Liouville ni Green dieron el término de error o las condiciones bajo las que sus soluciones eran válidas. La aproximación más general y precisa (9) aparece explícitamente en artículos de Gregor Wentzel (1898)⁴⁵¹, Hendrik A. Kramers (1894-1952),⁴⁵² León Brillouin (1889-1969)⁴⁵³ y Harold Jeffreys (1891),⁴⁵⁴ y se la conoce familiarmente como la solución *WKBJ*. Todos ellos fueron físicos que trabajaron en teoría cuántica con la ecuación de Schrödinger.

En un artículo leído en 1850⁴⁵⁵ estudia Stokes el valor de la integral de Airy

$$W = \int_0^{\infty} \cos \frac{\pi}{2} (w^3 - mw) dw \quad (10)$$

para $|m|$ grande. Esta integral representa la intensidad de la luz difractada en las proximidades de una cáustica. Airy había dado una serie para W en potencias de m que, a pesar de ser convergente para todo m , no resultaba útil para el cálculo de $|m|$ grande. El método de Stokes consistía en construir una ecuación diferencial de la que la integral es una solución particular, y después resolver dicha ecuación diferencial en términos de series divergentes que pudieran ser útiles para el cálculo (a estas series las llamó

semiconvergentes).

Una vez demostrado que $U = (\pi/2)^{1/3}$ satisface la ecuación diferencial de Airy

$$\frac{d^2U}{dn^2} + \frac{n}{3}U = 0, \quad n = \left(\frac{\pi}{2}\right)^{2/3} m, \quad (11)$$

demuestra Stokes que para n positivo

$$U = An^{-1/4} \left(R \cos \frac{2}{3} \sqrt{\frac{n^3}{3}} + S \operatorname{sen} \frac{2}{3} \sqrt{\frac{n^3}{3}} \right) + \quad (12)$$

$$+ Bn^{-1/4} \left(R \operatorname{sen} \frac{2}{3} \sqrt{\frac{n^3}{3}} - S \cos \frac{2}{3} \sqrt{\frac{n^3}{3}} \right)$$

donde

$$R = 1 - \frac{1 \times 5 \times 7 \times 11}{1 \times 2 \times 16^2 \times 3n^3} + \quad (13)$$

$$+ \frac{1 \times 5 \times 7 \times 11 \times 13 \times 17 \times 19 \times 23}{1 \times 2 \times 3 \times 4 \times 16^2 \times 3^2 \times n^6} \dots$$

$$S = \frac{1 \times 5}{1 \times 16 \times (3n)^{1/2}} - \quad (14)$$

$$- \frac{1 \times 5 \times 7 \times 11 \times 13 \times 17}{1 \times 2 \times 3 \times 16^3 \times (3n^3)^{3/2} \times n^6} + \dots$$

Los valores de A y B que hacen que U dé la integral vienen

determinados por un argumento especial (donde Stokes utiliza el principio de la fase estacionaria) y da también un resultado análogo para n negativo.

La serie (12) y la que corresponde a n negativo se comporta como series convergentes para cierto número de términos, pero de hecho divergen, y Stokes observó que podían ser utilizadas para el cálculo aproximado. Dado un valor de n , pueden utilizarse los términos desde el primero hasta uno que resulte más pequeño para dicho valor de n . Stokes dio un razonamiento de tipo cualitativo para mostrar por qué las series son útiles para el cálculo numérico.

Stokes mismo se encontró con una dificultad especial en la resolución de (11) para n positivo y negativo. No pudo pasar de la serie para n positivo a la serie para n negativo simplemente haciendo variar n a través de cero porque dicha serie no tenía sentido para $n = 0$. En vista de eso intentó pasar de valores de n positivos o negativos a través de valores complejos, pero esto no da la serie correcta y los multiplicadores constantes.

Lo que descubrió Stokes, no sin esfuerzo,⁴⁵⁶ fue que si para un cierto dominio de la amplitud de n , una solución general era representada por cierta combinación lineal de dos series asintóticas, cada una de las cuales es una solución, entonces no necesariamente la misma combinación lineal de los dos desarrollos asintóticos fundamentales representa la misma solución general en un entorno de ese dominio de amplitudes. Stokes vio que las constantes de la combinación lineal cambian abruptamente cuando se atraviesa las curvas $n = \text{const.}$, ahora llamadas curvas de Stokes.

Aunque Stokes se había interesado en principio en el cálculo de integrales, vio claramente que se podrían utilizar las series divergentes para resolver ecuaciones diferenciales en general. Mientras que Euler, Poisson y otros habían resuelto ecuaciones concretas de esta manera, sus resultados parecían reducirse a trucos que daban respuesta a problemas físicos concretos. Stokes dio varios ejemplos en sus artículos de 1856 y 1857.

Los trabajos anteriores sobre el cálculo de integrales y la resolución de ecuaciones diferenciales por medio de series divergentes constituyen una buena muestra de la obra de muchos matemáticos y físicos.

3. La teoría formal de las series asintóticas

Poincaré y Stieltjes lograron en 1886, de manera independiente, una definición formal y una caracterización completa de aquellas series divergentes que resultan útiles para la representación y cálculo de funciones. Poincaré llamó a estas series asintóticas, mientras que Stieltjes continuó utilizando el nombre de semiconvergentes⁴⁵⁷. Poincaré se ocupó del tema motivado por sus investigaciones sobre la resolución de ecuaciones diferenciales lineales. Impresionado por la utilidad de las series divergentes en astronomía, trató de determinar cuáles eran útiles y por qué, logrando al fin aislar y formular la propiedad esencial. Una serie de la forma

$$a_0 + \frac{a_1}{x} + \frac{a_2}{x^2} + \dots \quad (15)$$

donde los a_i son independientes de x , se dice que representa asintóticamente la función $f(x)$ para valores grandes de x , siempre que

$$\lim_{x \rightarrow \infty} x^n \left[f(x) - \left(a_0 + \frac{a_1}{x} + \dots + \frac{a_n}{x^n} \right) \right] = 0 \quad (16)$$

para $n = 0, 1, 2, 3, \dots$. La serie (15) será en general divergente, pero en casos especiales puede converger. La relación que acabamos de definir entre la función $f(x)$ y la serie se representa por

$$f(x) \sim a_0 + \frac{a_1}{x} + \frac{a_2}{x^2} + \dots$$

Tales series vienen a ser desarrollos de funciones en un entorno de $x = \infty$. En su artículo de 1886 Poincaré se limitó a valores de x reales; sin embargo, la definición sirve también para x compleja si sustituimos $x \rightarrow \infty$ por $|x| \rightarrow \infty$, aunque la validez de la representación puede verse ahora limitada a un sector del plano complejo con vértice en el origen.

Como hemos dicho, la serie (15) es asintótica a $f(x)$ en un entorno de $x = \infty$. Sin embargo, la definición ha sido generalizada, y puede decirse que la serie

$$a_0 + a_1x + a_2x^2 + \dots$$

es asintótica a $f(x)$ en $x = 0$ si

$$\lim_{x \rightarrow \infty} \frac{1}{x^n} \left[f(x) - \sum_0^{n-1} a_i x^i \right] = a_n$$

A pesar de que en el caso de algunas series asintóticas se sabe qué error se comete al truncar por un determinado término, para las series asintóticas generales no se suele conocer tal información sobre el error numérico cometido. Sin embargo, las series asintóticas pueden ser utilizadas para dar resultados numéricos bastante aproximados para valores grandes de x usando sólo aquellos términos para los cuales el valor va disminuyendo según vayamos tomando más y más términos. El orden de magnitud del error es en cualquier caso el del primer término omitido.

Poincaré demostró que la suma, diferencia, producto y cociente de dos funciones vienen representadas asintóticamente por la suma, diferencia, producto y cociente de sus correspondientes series asintóticas, siempre que el término constante de la serie divisor no sea nulo. Por otra parte, si

$$f(x) \sim a_0 + \frac{a_1}{x} + \frac{a_2}{x^2} + \dots$$

entonces

$$\int_{x_0}^x f(z) dz \rightarrow C + a_0 x + a_1 \log x - \frac{a_2}{x} - \frac{1}{2} \frac{a_3}{x^2} -$$

El uso de la integración supone una ligera generalización de la definición original, a saber,

$$\phi(x) \sim f(x) + g(x) \left(a_0 + \frac{a_1}{x} + \frac{a_2}{x^2} + \dots \right)$$

si

$$\frac{\phi(x) - f(x)}{g(x)} \sim a_0 + \frac{a_1}{x} + \frac{a_2}{x^2} + \dots$$

incluso cuando $f(x)$ y $g(x)$ no admitan ellas mismas representaciones por series asintóticas. En cuanto a la diferenciación, si se sabe que $f(x)$ tiene un desarrollo en serie asintótica, entonces se puede obtener simplemente diferenciando la serie asintótica de $f(x)$.

Si una función dada tiene un desarrollo en serie asintótica entonces es único, pero el recíproco es falso porque, por ejemplo, $(1+x)^{-1}$ y $(1+e^{-x})x(1+x)^{-1}$ tienen el mismo desarrollo asintótico.

Poincaré aplicó su teoría de las series asintóticas a las ecuaciones diferenciales, como hemos dicho, y pueden encontrarse muchos ejemplos de tal uso en su tratado de mecánica celeste *Les Méthodes nouvelles de la Mécanique Céleste*.⁴⁵⁸ El tipo de ecuaciones estudiadas en su artículo de 1886 es

$$P_n(x)y^{(n)} + P_{n-1}(x)y^{(n-1)} + \dots + P_0(x)y = 0 \quad (17)$$

donde los $P(x)$ son polinomios en x . En realidad Poincaré sólo se ocupó del caso de segundo orden, pero el método puede aplicarse perfectamente al caso general (17).

Los únicos puntos singulares de la ecuación (17) son los ceros de $P_n(x)$ y $x = \infty$. Para un punto singular regular (*Stelle der Bestimmtheit*) existen expresiones convergentes para las integrales, dadas por Fuchs (cap. 29, sec. 5). Consideremos entonces un punto singular irregular; mediante una transformación lineal este punto puede ser desplazado a ∞ , mientras que la ecuación sigue conservando su forma original. Si P_n es de grado p , la condición de que $x = \infty$ sea un punto singular regular implica que los grados de P_{n-1} , P_{n-2}, \dots, P_0 , sean como máximo $p - 1$, $p - 2, \dots, p - n$ respectivamente. Para un punto singular irregular uno (o más) de estos grados tiene que ser mayor. Poincaré demostró que para una ecuación diferencial de la forma (17), donde los grados de los P no exceden del grado de P_n existen n series de la forma

$$e^{ax} x^\alpha \left(A_0 + \frac{A_1}{x} + \dots \right)$$

que satisfacen formalmente la ecuación diferencial. También demostró que a cada una de estas series corresponde una solución exacta en forma de una integral a la que la serie es asintótica.

Los resultados de Poincaré están incluidos en el siguiente teorema debido a Jacob Horn (1867-1946).⁴⁵⁹ Horn estudia la ecuación

$$y^{(n)} + a_1(x)y^{(n-1)} + \dots + a_n(x)y^n = 0 \quad (18)$$

donde los coeficientes son funciones racionales de x que se suponen desarrollables, para x positivo y suficientemente grande, en series convergentes o asintóticas del tipo

$$a_r(x) = x^{rk} \left[a_{r,0} + \frac{a_{r,1}}{x} + \frac{a_{r,2}}{x^2} + \dots \right] \quad r = 1, 2, \dots, n$$

siendo k algún entero positivo o 0. Si para la ecuación (18) anterior las raíces m_1, m_2, \dots, m_n de su ecuación característica, es decir, de la ecuación algebraica

$$m^n + a_{1,0}m^{n-1} + \dots + a_{n,0} = 0$$

son distintas, entonces la ecuación (18) tiene n soluciones linealmente independientes y_1, y_2, \dots, y_n , desarrollables asintóticamente para valores positivos grandes de x , en la forma

$$y_r \sim e^{f_r(x)} x^{\rho_r} \sum_{j=0}^{\infty} \frac{A_{r,j}}{x^j} \quad r = 1, 2, \dots, n$$

donde $f_r(x)$ es un polinomio de grado $k + 1$ en x , cuyo coeficiente de la mayor potencia en x es $m_r/(k + 1)$, mientras que ρ_r y $A_{r,j}$ son constantes, con $A_{r,0} = 1$. Los resultados de Poincaré y Horn han sido extendidos a otros diversos tipos de ecuaciones diferenciales y

generalizados para incluir los casos en que las raíces de la ecuación característica no sean necesariamente distintas.

La existencia, forma y otras características de la solución en forma de serie asintótica cuando en la ecuación (18) se permite que la variable independiente tome valores complejos, fueron estudiadas por primera vez por Horn⁴⁶⁰. George David Birkhoff (1884-1944), uno de los primeros grandes matemáticos americanos, obtuvo un resultado general para este caso.⁴⁶¹ En este artículo no considera Birkhoff la ecuación (18) sino el sistema más general

$$\frac{dy_i}{dx} = \sum_{j=1}^n a_{ij}(x)y_j \quad i = 1, 2, \dots, n \quad (19)$$

donde, para $|x| > R$ tenemos, para cada a_{ij}

$$a_{ij}(x) \sim a_{ij}x^q + a_{ij}^{(1)}x^{q-1} + \dots + a_{ij}^{(q)} + a_{ij}^{(q+1)}\frac{1}{x} + \dots$$

y para el que la ecuación característica en a

$$|a_{ij} - \delta_{ij}\alpha| = 0$$

tiene raíces distintas. Birkhoff obtuvo soluciones en series asintóticas para las y_i que se verifican en diversos sectores del plano complejo con vértices en $x = 0$.

Mientras que los desarrollos en serie tanto de Poincaré como de los

restantes autores que acabamos de mencionar eran en potencias de la variable independiente, otras investigaciones sobre soluciones de ecuaciones diferenciales mediante series asintóticas volvieron a considerar el problema original de Liouville (sec. 2) con un parámetro. Birkhoff mismo se ocupó del problema y obtuvo un resultado general.⁴⁶² Estudió la ecuación

$$\frac{d^n z}{dx^n} + \rho a_{n-1}(x, \rho) \frac{d^{n-1} z}{dx^{n-1}} + \dots + \rho^n a_0(x, \rho) z = 0 \quad (20)$$

para $|\rho|$ grande y para x sobre el intervalo $[a, b]$. Las funciones $a_i(x, \rho)$ se suponen analíticas en el parámetro complejo ρ en $\rho = \infty$ y con derivadas de todos los órdenes en la variable real x . Estas hipótesis sobre las $a_i(x, \rho)$ implican que

$$a_i(x, \rho) = \sum_{j=0}^{\infty} a_{ij}(x) \rho^{-j}$$

y que las raíces de $w_1(x)$, $w_2(x)$, ..., $w_n(x)$ de la ecuación característica

$$w^n + a_{n-1,0}(x)w^{n-1} + \dots + a_{0,0}(x) = 0$$

son distintas para cada x . Birkhoff demuestra que hay n soluciones independientes

$$z_1(x, \rho), \dots, z_n(x, \rho)$$

de la ecuación (20) que son analíticas en ρ en una región S del plano ρ (determinada por el argumento de ρ), tales que para todo m entero y para $|\rho|$ grande

$$z_i(x, \rho) = u_i(x, \rho) + e^{[\rho \int_a^x w_i(t) dt]} E_0 \rho^{-m} \quad (21)$$

donde

$$u_i(x, \rho) = e^{[\rho \int_a^x w_i(t) dt \sum_{j=0}^{m-1} u_{ij}(x) \rho^{-j}]} \quad (22)$$

y E_0 es una función de x, ρ , y m acotada para todo x en $[a, b]$ y ρ en S ; las funciones $u_{ij}(x)$ pueden calcularse. A la vista de (22), el resultado (21) afirma que z_i viene dada por una serie en $1/\rho$ hasta $1/\rho^{m-1}$ más un término de resto, a saber, el segundo término de la derecha, que contiene $1/\rho^m$. Además, dado que m es arbitrario, pueden tomarse en la expresión de tantos términos en $1/\rho$ como se quiera. Y como E_0 está acotada, el término del resto es de orden más alto en $1/\rho$ que $u_i(x, \rho)$, y la serie completa que se puede obtener haciendo m infinito es asintótica a

$$\frac{z_i(x, \rho)}{\exp(\rho \int_a^x w(t) dt)}$$

en el sentido de Poincaré.

En el teorema de Birkhoff, la serie asintótica para ρ compleja es válida solamente en un sector S del plano complejo ρ . Aparece el fenómeno de Stokes, es decir, la prolongación analítica de $z_i(x,\rho)$ a través de una línea de Stokes no viene dada por la prolongación analítica de la serie asintótica para $z_i(x,\rho)$.

El uso de las series asintóticas o de la aproximación *WKBJ* a las soluciones de ecuaciones diferenciales planteó otro problema. Supongamos que tenemos la ecuación

$$y'' + \lambda^2 q(x)y = 0 \quad (23)$$

para x sobre el intervalo $[a,b]$ La aproximación *WKBJ* para valores grandes de λ da, en vista de (7), dos soluciones para $x > 0$ y dos para $x < 0$. Falla para los valores de x tales que $q = 0$. Uno de estos puntos recibe el nombre de punto de transición, punto crítico o punto de Stokes. Las soluciones exactas de la ecuación (23) son finitas en dichos puntos, sin embargo. El problema consiste en relacionar las soluciones *WKBJ* por cada lado del punto de transición de manera que representen la misma solución exacta sobre el intervalo $[a,b]$ sobre el que se trata de resolver la ecuación diferencial. Para formular el problema de una manera más explícita, considérese la ecuación anterior en la que $q(x)$ es real para x real y tal que $q(x) = 0$, $q'(x) \neq 0$ para $x = 0$. Supongamos también que $q(x)$ es negativa para x positiva (o al revés). Dada una combinación lineal de dos soluciones *WKBJ*, una para $x < 0$ y la otra para $x > 0$, se plantea la cuestión de cuál de las soluciones válidas para $x > 0$

habría que unir a una válida para $x < 0$. Las llamadas fórmulas de conexión dan la respuesta.

El primero que dio la idea de cómo cruzar a través de un cero de $q(x)$ fue Lord Rayleigh,⁴⁶³ idea generalizada poco más tarde por Richard Gans (1880-1954),⁴⁶⁴ que conocía el trabajo de Rayleigh. Ambos eran físicos que estudiaban la propagación de la luz a través de medios no homogéneos.

El primer tratamiento sistemático de las fórmulas de conexión se debe a Harold Jeffreys,⁴⁶⁵ independientemente del trabajo de Gans. Jeffreys estudió la ecuación

$$\frac{d^2y}{dx^2} + \lambda^2 X(x)y = 0 \quad (24)$$

donde x es real, λ es real y suficientemente grande, y $X(x)$ tiene un cero simple, digamos en $x = 0$, obteniendo fórmulas para conectar las soluciones por medio de series asintóticas de (24) para $x > 0$ y $x < 0$ por medio de una ecuación de aproximación a (24), a saber

$$\frac{d^2y}{dx^2} + \lambda^2 xy = 0 \quad (25)$$

en la que la $X(x)$ de (24) aparece sustituida por una función lineal en x . Las soluciones de (25) son

$$y_{\pm}(x) = x^{1/2} J_{\pm 1/3}(\xi) \quad (26)$$

donde $\xi = (2/3) \lambda x^{3/2}$. Los desarrollos asintóticos de estas soluciones para valores grandes de x se pueden utilizar para unir las soluciones asintóticas de (24) por ambos lados de $x = 0$. Es necesario tener en cuenta diversos detalles sobre los dominios de x y λ en el proceso de unión, pero aquí no entraremos en ellos. Más tarde se publicaron numerosos artículos generalizando las fórmulas de conexión para los casos en que en (24) $X(x)$ tenga ceros múltiples o varios ceros distintos, así como para ecuaciones más complicadas de segundo orden y de orden superior, y para valores complejos de x y de λ .

La teoría de series asintóticas, aplicada al cálculo de integrales o a la resolución aproximada de ecuaciones diferenciales, ha experimentado un enorme crecimiento en años recientes. Vale la pena subrayar que los desarrollos matemáticos vienen a demostrar que los matemáticos del siglo XVIII y XIX, especialmente Euler, que se dieron cuenta de la gran utilidad de las series divergentes y sostuvieron que estas series se podrían usar como los equivalentes analíticos de las funciones que representaban, es decir, que operaciones sobre las series corresponderían a operaciones sobre las funciones, estaban bien encaminados. Aunque no consiguieran aislar rigurosamente la idea esencial, intuitivamente y a partir de los resultados obtenidos se dieron cuenta de que las series divergentes estaban estrechamente relacionadas con las funciones que representaban.

5. El problema de la sumabilidad de series divergentes

Los trabajos sobre series divergentes que hemos analizado hasta ahora se referían únicamente al problema de encontrar series asintóticas para representar funciones, o bien conocidas explícitamente, o cuya existencia estaba asegurada implícitamente como soluciones de ecuaciones diferenciales ordinarias. Otro problema que ocupó a los matemáticos desde 1880 aproximadamente en adelante, es esencialmente el recíproco del de encontrar series asintóticas. Dada una serie divergente en el sentido de Cauchy, ¿se le podrá asignar una «suma»? Si la serie es de términos variables, esta «suma» sería una función para la cual la serie divergente podría ser o no su desarrollo asintótico. No obstante, la función podría ser considerada como la «suma» de la serie, «suma» que podría servir para algún fin útil, incluso aunque la serie no convergiera a ella o no pudiera ser utilizable para calcular valores aproximados de la función.

En cierto sentido, el problema de sumar series divergentes se planteó ya antes de que Cauchy introdujese sus definiciones de convergencia y divergencia. Los matemáticos se encontraron con series divergentes y trataron de encontrar sumas para ellas lo mismo que para las series convergentes, simplemente porque la distinción entre los dos tipos no estaba establecida con precisión, y la única pregunta era ¿cuál será la suma correcta? Así, por ejemplo, el principio de Euler (cap. 20, sec. 7) de que un desarrollo en serie de potencias de una función tiene como suma el valor de la función de la que se deriva la serie, daba una suma para la serie incluso

para valores de x para la que ésta diverge en el sentido de Cauchy. Análogamente, en su transformación de series (cap. 20, sec. 4) convertía las series divergentes en convergentes sin dudar de que prácticamente todas las series tendrían su suma. Sin embargo, a partir de que Cauchy hiciese la distinción entre convergencia y divergencia, el problema de sumar series divergentes se planteaba en un nivel distinto. La asignación de sumas a todas las series, hecha de manera relativamente ingenua durante el siglo XVIII, ya no era aceptable. Las nuevas definiciones deberían establecer lo que hoy se llama *sumabilidad* para distinguirlo del concepto de convergencia en el sentido de Cauchy.

Mirando retrospectivamente, puede verse que la noción de sumabilidad era realmente la que estaban anticipando los matemáticos del siglo XVIII y comienzos del XIX. A esto es a lo que conducen los métodos de sumación de Euler que acabamos de mencionar; de hecho, en una carta a Goldbach de 7 de agosto de 1745, en la que afirma que la suma de una serie de potencias es el valor de la función de la que se deriva la serie, Euler asegura también que toda serie debe tener una suma, pero debido a que la palabra suma implica el proceso usual de añadir, y este proceso no conduce a la suma en el caso de una serie divergente tal como la $1 - 1! + 2! - 3! + \dots$, tendríamos que utilizar la palabra «valor» para la «suma» de una serie divergente.

Poisson también introdujo lo que hoy reconoceríamos como un concepto de sumabilidad. En la definición de Euler de suma como el valor de la función de la que se deriva la serie, está implícita la idea

de que

$$\sum_{n=0}^{\infty} a_n = \lim_{x \rightarrow 1^-} \sum_{n=0}^{\infty} a_n x^n \quad (27)$$

donde la notación $x \rightarrow 1^-$ significa que x tiende a 1 por valores inferiores. Según (27) la suma de $1 - 1 + 1 - 1 + \dots$ ha de ser

$$\lim_{x \rightarrow 1^-} (1 - x + x^2 - x^3 + \dots) = \lim_{x \rightarrow 1^-} (1 + x)^{-1} = \frac{1}{2}$$

Poisson⁴⁶⁶ estudió la serie

$$\text{sen } \theta + \text{sen } 2\theta + \text{sen } 3\theta + \dots$$

que diverge excepto cuando θ es un múltiplo de n . Su idea, expresada para la serie de Fourier completa

$$\frac{a_0}{2} + \sum_n (a_n \cos n\theta + b_n \text{sen } n\theta) \quad (28)$$

es la de que habría que considerar la serie de potencias asociada

$$\frac{a_0}{2} + \sum_n (a_n \cos n\theta + b_n \text{sen } n\theta) r^n \quad (29)$$

y definir la suma de (28) como el límite de la serie (29) cuando r tiende a 1 por la izquierda. Por supuesto, Poisson no era consciente de que lo que estaba sugiriendo era una definición de suma para una serie divergente porque, como ya hemos dicho, la distinción entre convergencia y divergencia no era clara en esta época.

La definición utilizada por Poisson recibe hoy el nombre de sumabilidad en el sentido de Abel, porque también aparece sugerida en un teorema de Abel⁴⁶⁷ que afirma que si la serie de potencias

$$f(x) = \sum_{n=0}^{\infty} a_n x^n$$

tiene radio de convergencia r y converge para $x = r$ entonces

$$\lim_{x \rightarrow r^-} f(x) = \sum_{n=0}^{\infty} a_n r^n \quad (30)$$

Entonces la función $f(x)$ definida por la serie en el intervalo $-r < x \leq r$ es continua por la izquierda en $x = r$. Sin embargo, si no converge y el límite (30) existe para $r = 1$, entonces tenemos una definición de suma para la serie divergente. Esta definición formal de sumabilidad para series divergentes en el sentido de Cauchy no se introdujo hasta finales del siglo XIX en un contexto que veremos más adelante.

Uno de los motivos para reconsiderar la sumación de series divergentes, aparte de su reconocida utilidad en astronomía, fue lo

que se ha llamado el problema del valor en la frontera (*Grenzwert*) en la teoría de funciones analíticas. Una serie de potencias $\sum a_n x^n$ puede representar una función analítica dentro de un círculo de radio r pero no para valores de x en la circunferencia. El problema era saber si uno podría encontrar un concepto de suma tal que la serie de potencias pudiera tener una suma para $|x| = r$ y tal que dicha suma pudiera incluso ser el valor de $f(x)$ cuando $|x|$ tiende a r . Este intento de extender el dominio de representación de una función analítica mediante una serie de potencias fue lo que motivó a Frobenius, Hölder y Ernesto Cesáro. Frobenius⁴⁶⁸ demostró que si la serie de potencias $\sum a_n x^n$ tiene el intervalo de convergencia $-1 < x < 1$, y si

$$s_n = a_1 + a_2 + a_3 + \dots + a_n. \quad (31)$$

entonces

$$\lim_{x \rightarrow 1^-} \sum_{n=0}^{\infty} a_n x^n = \lim_{n \rightarrow \infty} \frac{s_1 + s_2 + \dots + s_n}{n+1}$$

cuando existe el límite de la derecha. Así, la serie de potencias, normalmente divergente para $x = 1$, puede tener una suma; además, si $f(x)$ es la función representada por la serie, la definición de Frobenius del valor de la serie para $x = 1$ coincide con $\lim_{x \rightarrow 1^-} f(x)$

Abstrayéndolo de su relación con las series de potencias, el artículo de Frobenius venía a sugerir una definición de sumabilidad para

series divergentes. Si $\sum a_n$ es divergente y s_n tiene el significado de (31), entonces se puede tomar como suma

$$\sum_{n=0}^{\infty} a_n = \lim_{n \rightarrow \infty} S_n = \lim_{n \rightarrow \infty} \frac{s_1 + s_2 + \dots + s_n}{n+1}$$

si este límite existe. Por ejemplo, para la serie $1 - 1 + 1 - 1 + \dots$, las S_n toman los valores $1, 1/2, 2/3, 2/4, 3/5, 1/2, 4/7, 1/2, \dots$ de manera que $\lim_{n \rightarrow \infty} S_n = 1/2$. Si $\sum a_n$ converge, entonces la «suma» de Frobenius da la suma usual. Esta idea de promediar las sumas parciales de una serie se puede encontrar ya en la vieja literatura; fue utilizada para tipos especiales de series por Daniel Bernoulli⁴⁶⁹ y Joseph L. Raabe (1801-1859).⁴⁷⁰

Poco después de que Frobenius publicara su artículo, Hölder dio una generalización. Dada la serie $\sum a_n$ sean

$$\begin{aligned} s_n^{(0)} &= s_n \\ s_n^{(1)} &= \frac{1}{n+1} s_0^{(0)} + s_1^{(0)} + \dots + s_n^{(0)} \\ s_n^{(02)} &= \frac{1}{n+1} s_0^{(1)} + s_1^{(1)} + \dots + s_n^{(1)} \\ &\dots \dots \\ s_n^{(r)} &= \frac{1}{n+1} s_0^{(r-1)} + s_1^{(r-1)} + \dots + s_n^{(r-1)} \end{aligned}$$

Entonces la suma s viene dada por

$$s = \lim_{n \rightarrow \infty} s_n^{(r)} \quad (32)$$

si este límite existe para algún r . La definición de Hölder se conoce hoy como (H, r) -sumabilidad.

He aquí un ejemplo de Hölder: considérese la serie

$$-\frac{1}{(1+x)^2} = -1 + 2x - 3x^2 + 4x^3 - \dots$$

que diverge para $x = 1$. Sin embargo, para $x = 1$ se tiene $s_0 = -1$, $s_1 = 1$, $s_2 = -2$, $s_3 = 2$, $s_4 = -3, \dots$

Entonces

$$s_0^{(1)} = -1, s_1^{(1)} = 0, s_2^{(1)} = -2/3, s_3^{(1)} = 0, s_4^{(1)} = -3/5, \dots;$$

$$s_0^{(2)} = -1, s_1^{(2)} = -1/2, s_2^{(2)} = -5/9, s_3^{(2)} = -5/12, s_4^{(2)} = -34/75, \dots$$

Es casi inmediato que $\lim_{n \rightarrow \infty} s_n^{(2)} = -1/4$, y esta es la $(H,2)$ -suma de Hölder. Se trata exactamente del mismo valor que asignó⁴⁷¹ Euler a la serie basándose en su principio de que la suma es el valor de la función de la que se deriva la serie.

Otra de las definiciones de sumabilidad hoy habituales se debe a Cesáro, profesor de la universidad de Nápoles.⁴⁷² Sea la serie $\sum_{i=0}^{\infty} a_i$ y sea $s_n = \sum_{i=0}^n a_i$. Entonces la suma de Cesáro es

$$s = \lim_{n \rightarrow \infty} \frac{S_n^{(r)}}{D_n^{(r)}} \quad r \text{ entero y } \geq 0 \quad (33)$$

donde

$$S_n^{(r)} = s_n + r s_{n-1} + \frac{r(r+1)}{2!} s_{n-2} + \dots + \frac{r(r+1) \dots (R+N-1)}{n!} s_0$$

y

$$D_n^{(r)} = \frac{(r+1)(r+2) \dots (r+n)}{n!}$$

El caso $r = 1$ incluye la definición de Frobenius. La definición de Cesáro se conoce hoy como (C, r) -sumabilidad. Los métodos de Hölder y Cesáro dan los mismos resultados: que la sumabilidad de Hölder implica la de Cesáro fue demostrado por Konrad Knopp (1882-1957) en una disertación no publicada en 1907; el recíproco fue demostrado por Walter Schnee (n. 1885).⁴⁷³

Una característica interesante de algunas de las definiciones de sumabilidad, cuando se aplican a series de potencias de radio de convergencia 1, es la de que no sólo dan una suma que coincide con $\lim_{x \rightarrow 1^-} f(x)$, donde $f(x)$ es la función de la que se deriva la serie, sino que tienen la propiedad adicional de que siguen teniendo significado en regiones donde $|x| > 1$ y en estas regiones dan la prolongación analítica de la serie de potencias original.

Otros desarrollos en la sumación de series divergentes estuvieron motivados por un problema completamente distinto, las investigaciones de Stieltjes sobre fracciones continuas. Euler⁴⁷⁴

había utilizado ya el hecho de que las fracciones continuas pueden convertirse en series convergentes o divergentes y recíprocamente. Euler buscaba (cap. 20, sec. 4 y 6) una suma para la serie divergente

$$1 - 2! + 3! - 4! + 5! - \dots \quad (34)$$

En primer lugar demostró, en su artículo sobre series divergentes⁴⁷⁵ y en cartas a Nicholas Bernoulli (1687-1759),⁴⁷⁶ que

$$x - (1!)x^2 + (2!)x^3 - (3!)x^4 + \dots \quad (35)$$

satisface formalmente la ecuación diferencial

$$x^2 \frac{dy}{dx} + y = x$$

para la que obtuvo la solución integral

$$y = \int_0^{\infty} \frac{xe^{-t}}{1+xt} dt \quad (36)$$

Entonces, utilizando reglas obtenidas transformando series convergentes en fracciones continuas, Euler transformó (35) en

$$\frac{x}{1} + \frac{x}{1+1} + \frac{x}{1+1+1} + \frac{2x}{1+1+1+1} + \frac{2x}{1+1+1+1+1} + \frac{3x}{1+1+1+1+1+1} + \frac{3x}{1+1+1+1+1+1+1} + \dots \quad (37)$$

Este trabajo tiene dos características importantes. Por un lado Euler obtiene una integral que puede ser considerada como la «suma» de la serie divergente (35); de hecho, esta es asintótica a la integral. Por otra parte mostró cómo convertir series divergentes en fracciones continuas. Utilizó la fracción continua para $x = 1$ para calcular el valor de la serie (34).

Otros trabajos de este tipo aparecieron a finales del siglo XVIII y buena parte del XIX, de los que el más notable se debe a Laguerre,⁴⁷⁷

(falta 1 página en mi original)

en 1894 y 1895.⁴⁷⁸ Esta obra, que constituye el origen de la teoría analítica de fracciones continuas, estudia cuestiones de convergencia y las relaciones con las integrales definidas y series divergentes. En estos artículos fue donde introdujo Stieltjes la integral que lleva su nombre.

Stieltjes parte de la fracción continua

$$\frac{1}{a_1 z + a_2} + \frac{1}{a_3 z + a_4} + \frac{1}{a_5 z + a_6} + \dots + \frac{1}{a_{2n} z + a_{2n+1}} + \frac{1}{a_{2n+2} z + a_{2n+3}} + \dots, \quad (40)$$

donde los a_n son números reales positivos y z es una variable compleja. Demuestra que cuando la serie

$$\sum_{n=1}^{\infty} a_n$$

diverge, la fracción continua (40) converge a una función $F(z)$ que es analítica en todo el plano complejo excepto a lo largo del eje real negativo junto con el origen, y

$$F(z) = \int_0^{\infty} \frac{d\phi(u)}{u+z} \quad (41)$$

Cuando $\sum a_n$ converge, las sumas parciales pares e impares de (40) convergen a límites distintos $F_1(z)$ y $F_2(z)$, donde

$$F_1(z) = \int_0^{\infty} \frac{dg_1(u)}{z+u} \quad F_2(z) = \int_0^{\infty} \frac{dg_2(u)}{z+u}$$

Por otra parte, se sabía que la fracción continua (40) podía desarrollarse formalmente en serie

$$\frac{C_0}{z} - \frac{C_1}{z^2} + \frac{C_2}{z^3} - \frac{C_3}{z^4} + \dots \quad (42)$$

con los C positivos. La correspondencia es también recíproca (con algunas restricciones); a toda serie (42) le corresponde una fracción

continua (40) con los a_n positivos. Stieltjes mostró cómo calcular los C_n a partir de los a_n , y en el caso de que Σa_n sea divergente, demostró que el cociente C_n/C_{n-1} aumenta; si tiene un límite finito λ , la serie converge para $|z| > \lambda$, pero si el cociente crece indefinidamente la serie diverge para todo z .

La relación entre la serie (42) y la fracción continua (40) es más complicada. Aunque la fracción continua converge si lo hace la serie, el recíproco no es cierto. Cuando la serie (42) diverge hay que distinguir dos casos, según que Σa_n sea divergente o convergente. En el primer caso, como hemos dicho, la fracción continua da uno y sólo un equivalente funcional, que puede tomarse como la suma de la serie divergente (42). Cuando Σa_n es convergente se obtienen dos funciones distintas de la fracción continua, una de los convergentes pares y la otra de los convergentes impares. Pero a la serie (42) (ahora divergente) corresponden infinitas funciones, cada una de las cuales tiene dicha serie como desarrollo asintótico.

Los resultados de Stieltjes también tuvieron la importante consecuencia siguiente: revelaban una clasificación de las series divergentes en dos clases al menos; aquellas series para las que propiamente había un único equivalente funcional cuyo desarrollo era dicha serie, y aquellas para las que había al menos dos equivalentes funcionales cuyo desarrollo era la serie. La fracción continua es solamente el intermediario entre la serie y la integral; es decir, dada la serie, se obtiene la integral a través de la fracción continua. Así, una serie divergente corresponde a una o más funciones, las cuales pueden considerarse como la suma de la serie

en un sentido nuevo de la misma.

Stieltjes también planteó y resolvió un tipo de problema inverso. Para simplificar un poco la exposición, supongamos que $\Phi(u)$ es diferenciable, de manera que la integral (41) se puede escribir como

$$\int_0^{\infty} \frac{f(u)}{z+u} du$$

A la serie divergente (42), y en el caso en que Σa_n sea divergente, corresponde una integral de este tipo. El problema recíproco es, dada la serie, hallar $f(u)$. Un desarrollo formal de la integral muestra que

$$C_n = \int_0^{\infty} f(u)u^n du \quad n = 0,1,2... \quad (43)$$

Por tanto, conociendo los C_n hay que determinar $f(u)$ de manera que satisfaga las infinitas ecuaciones (43). A este problema es al que llamó Stieltjes «problema de los momentos». No admite solución única, y Stieltjes mismo dio una función

$$f(u) = e^{-\sqrt[4]{u}} \operatorname{sen} \sqrt[4]{u}$$

que hace $C_n = 0$ para todo n . Si se impone la condición suplementaria de que $f(u)$ sea positiva entre los límites de

integración, entonces sólo es posible una única solución $f(u)$.

El desarrollo sistemático de la teoría de series sumables comienza con la obra de Borel a partir de 1895. Al principio Borel dio definiciones que generalizan las de Cesáro, para dar después, partiendo de la obra de Stieltjes, una definición integral.⁴⁷⁹ Si se aplica el proceso utilizado por Laguerre a una serie cualquiera de la forma

$$a_0 + a_1x + a_2x^2 + \dots \quad (44)$$

que tenga un radio de convergencia finito (incluyendo 0), nos vemos conducidos a la integral

$$\int_0^{\infty} e^{-zx} F(zx) dz \quad (45)$$

donde

$$F(u) = 1 + a_1u + \frac{a_2}{2!}u^2 + \dots + \frac{a_n}{n!}u^n + \dots$$

Esta integral es la expresión sobre la que construyó Borel su teoría de series divergentes; Borel la consideró como la suma de la serie (44). La serie $F(u)$ se llama la serie asociada de la serie original.

Si la serie original (44) tiene un radio de convergencia R mayor que cero, entonces la serie asociada representa una función entera, la

integral

$$\int_0^{\infty} e^{-z} F(zx) dz$$

tiene sentido si x es interior al círculo de convergencia y los valores da la integral y la serie son idénticos. Pero la integral también puede tener sentido para valores de x exteriores al círculo de convergencia y en este caso la integral da una prolongación analítica de la serie original. Borel llama a la serie *sumable* (en el sentido mencionado) en un punto x donde la integral tenga sentido.

Si la serie original (44) es divergente ($R = 0$), la serie asociada puede ser convergente o divergente. Si es convergente sobre una región del plano $u = zx$ únicamente, entendemos por $F(u)$ el valor no simplemente de la serie asociada, sino el de su prolongación analítica. Entonces la integral

$$\int_0^{\infty} e^{-z} F(zx) dz$$

puede tener un significado y se obtiene, como se ve, de la serie divergente original. Borel se ocupó también de la determinación de la región de los valores de x en los que la serie original es sumable, tanto cuando la serie original es convergente ($R > 0$) como si es divergente ($R = 0$).

Borel introdujo también el concepto de sumabilidad absoluta. La serie original se llama absolutamente sumable en un valor de x cuando

$$\int_0^{\infty} e^{-z} F(zx) dz$$

es absolutamente convergente y las sucesivas integrales

$$\int_0^{\infty} e^{-z} \left| \frac{d^{\lambda} F(zx)}{dz^{\lambda}} \right| dz \quad \lambda = 1, 2, \dots$$

tengan sentido. Borel demuestra a continuación que si una serie divergente es absolutamente sumable, puede ser manipulada precisamente como una serie convergente. En otras palabras, la serie representa una función y puede ser utilizada en lugar de la función.

Así, la suma, diferencia y producto de dos series absolutamente sumables es absolutamente sumable y representa la suma, diferencia y producto, respectivamente, de las dos funciones representadas por las series individuales. Un hecho análogo se verifica para la derivada de una serie absolutamente sumable. Además, la suma en el sentido anterior coincide con la suma usual en el caso de las series convergentes, y la resta de los k primeros términos reduce la «suma» de la serie total en la suma de esos k

términos. Borel subraya que cualquier definición satisfactoria de sumabilidad debería tener estas propiedades, aunque no todas las tienen. No exige, en cambio, que dos definiciones cualesquiera den necesariamente la misma suma.

Estas propiedades posibilitaron la aplicación inmediata de la teoría de Borel a las ecuaciones diferenciales. De hecho, si

$$P(x, y, y', \dots, y^{(n)}) = 0$$

es una ecuación diferencial que es holomorfa en x en el origen y algebraica en y y todas sus derivadas, entonces cualquier serie absolutamente sumable

$$a_0 + a_1x + a_2x^2 + \dots$$

que satisfaga formalmente la ecuación diferencial, define una función analítica que es una solución de la ecuación. Por ejemplo, la serie de Laguerre

$$1 + x + 2!x^2 + 3!x^3 + \dots$$

satisface formalmente la ecuación

$$x^2 \frac{d^2y}{dx^2} + (x-1)y = -1$$

y por tanto la función (39)

$$f(x) = \int_0^{\infty} e^{-z} \frac{1}{1-zx} dz$$

será solución de la ecuación.

Una vez que el concepto de sumabilidad fue aceptado, docenas de matemáticos introdujeron diversas definiciones nuevas que cumplieran algunas o todas las exigencias indicadas por Borel y otros. Muchas de las definiciones de sumabilidad han sido generalizadas a series dobles. Naturalmente también se han planteado y resuelto muchos problemas que involucran la noción de sumabilidad. Por ejemplo, supongamos una serie sumable por un cierto método: ¿qué condiciones adicionales habría que imponer a dicha serie para que, supuesta su sumabilidad, también fuera convergente en el sentido de Cauchy? Tales teoremas reciben el nombre de teoremas tauberianos, en honor a Alfred Tauber (n. 1866). Así, por ejemplo, Tauber demostró⁴⁸⁰ que si Σa_n es sumable en el sentido de Abel a s y na_n tiende a cero cuando n tiende a infinito, entonces converge a s . El concepto de sumabilidad nos permite, pues, dar un valor o suma a una gran variedad de series divergentes. Naturalmente, se plantea de manera necesaria la cuestión de qué es lo que se ha conseguido así. Si de una situación física concreta surge una serie dada, la adecuación de cualquier definición de suma dependería completamente de si tal suma es o no físicamente significativa, de la misma manera que la utilidad física de cualquier geometría depende

de si describe o no adecuadamente el espacio físico. La definición de suma de Cauchy es la que usualmente encaja bien, porque básicamente viene a decir que la suma es lo que se obtiene añadiendo cada vez más términos en el sentido ordinario. Pero no hay ninguna razón lógica para preferir este concepto de suma a los otros que han sido introducidos. Ciertamente, la representación de funciones mediante series se ha ampliado enormemente utilizando los nuevos conceptos. Por ejemplo, Leopold Fejér (1880-1959), alumno de H. A. Schwarz, mostró el valor de la teoría de sumabilidad dentro de la teoría de las series de Fourier. En 1904⁴⁸¹ demostró Fejér que si la función $f(x)$ está acotada en el intervalo $[-\pi, \pi]$ y es integrable en el sentido de Riemann, o si no está acotada pero la integral $\int_{-\pi}^{\pi} f(x) dx$ es absolutamente convergente, entonces en todo punto del intervalo en el que existan $f(x + 0)$ y $f(x - 0)$, la suma de Frobenius de la serie de Fourier

$$\frac{a_0}{2} + \sum_{n=1}^{\infty} (a_n \cos nx + b_n \operatorname{sen} nx)$$

es $[f(x + 0) + f(x - 0)]/2$. Las condiciones exigidas a $f(x)$ en este teorema son más débiles que en los teoremas anteriores sobre la convergencia de la serie de Fourier a $f(x)$ (véase cap. 40, sec. 6).

Este resultado fundamental de Fejér fue el comienzo de una larga serie de investigaciones fructíferas sobre la sumabilidad de series. Ya hemos tenido numerosas ocasiones de ver la necesidad de representar las funciones por medio de series. Así, al tratar la

condición inicial en la resolución de problemas de valor inicial o de contorno para ecuaciones en derivadas parciales, suele ser necesario representar la $f(x)$ inicial dada en términos de autofunciones obtenidas aplicando las condiciones de contorno a las ecuaciones diferenciales ordinarias que resultan del método de separación de variables. Estas autofunciones pueden ser funciones de Bessel, funciones de Legendre o cualesquiera otras de un cierto número de funciones especiales. Mientras que puede no tenerse la convergencia en el sentido de Cauchy de tal serie de autofunciones a la $f(x)$ dada, dicha serie puede ciertamente ser sumable a $f(x)$ en uno u otro de los sentidos de sumabilidad que hemos visto y así se satisfaga la condición inicial. Estas aplicaciones de la sumabilidad representan un gran éxito del concepto mismo.

La construcción y posterior aceptación de la teoría de series divergentes constituye otro ejemplo sorprendente de cómo crece la matemática. Viene a mostrar, ante todo, que cuando un concepto o técnica demuestra ser útil, incluso aunque su lógica sea confusa o inexistente, una investigación persistente descubrirá alguna justificación lógica a propósito. También muestra lo lejos que han llegado los matemáticos en reconocer que la matemática es obra del hombre. Las nuevas definiciones de sumabilidad no coinciden con la noción natural de ir añadiendo más y más términos continuamente, noción que Cauchy simplemente rigorizó, sino que son artificiales. Pero no obstante sirven para fines matemáticos, incluyendo la solución matemática de problemas físicos; y estos son ahora motivos suficientes para admitirlas dentro del dominio de la

matemática con todos los derechos.

Bibliografía

- Borel, Emile: *Notice sur les travaux scientifiques de M. Emile Borel*, 2.^a ed., Gauthier-Villars, 1921; *Leçons sur les séries divergentes*, Gauthier-Villars, 1901.
- Burkhardt, H.: «Trigonometrische Reihe und Intégrale». *Encyk. der Math.*
- *Wiss.*, B. G. Teubner, 1904-1916, II A12, 819-1354; «Über den Gebrauch divergenter Reihen in der Zeit von 1750-1860». *Math. Ann.*, 70, 1911, 169-206.
- Carmichael, Roben D.: «General Aspects of the Theory of Summable Series». *Amer. Math. Soc. Bull.*, 25, 1918/1919, 97-131.
- Collingwood, E. F.: «Emile Borel». *Jour. Lon. Math. Soc.*, 34, 1959, 488-512.
- Ford, W. B.: «A Conspectus of the Modera Theories of Divergent Series». *Amer. Math. Soc. Bull.*, 25, 1918/1919, 1-15.
- Hardy, G. H.: *Divergent Series*, Oxford University Press, 1949. Véanse las notas históricas.
- Hurwitz, W. A.: «A Report on Topics in the Theory of Divergent Series». *Amer. Math. Soc. Bull.*, 28, 1922, 17-36.
- Knopp, K.: «Neuere Untersuchungen in der Theorie der divergenten Reihen». *Jahres. der Deut. Math.-Verein.*, 32, 1923, 43-67.
- Langer, Rudolf E.: «The Asymptotic Solution of Ordinary Linear

- Differential Equations of the Second Order». *Amer. Math. Soc. Bull.*, 40, 1934, 545-582.
- McHugh, J. A. M.: «An Historical Survey of Ordinary Linear Differential Equations with a Large Parameter and Turning Points». *Archive for History of Exact Sciences*, 7, 1971, 277-324.
 - Moore, C. N.: «Applications of the Theory of Summability to Developments in Orthogonal Functions». *Amer. Math. Soc. Bull.*, 25, 1918/1919, 258-276.
 - Plancherel, Michel: «Le Développement de la théorie des séries trigonométriques dans le dernier quart de siècle». *L'Enseignement Mathématique*, 24, 1924/1925, 19-58.
 - Pringsheim, A.: «Irrationalzahlen und Konvergenz unendlicher Prozesse». *Encyk. der Math. Wiss.*, B. G. Teubner, 1898-1904, IA3, 47-146.
 - Reiff, R.: *Geschichte der unendlichen Reihen*, H. Lauppsche Buchhandlung, 1889, Martin Sandig (reprint), 1969.
 - Smail, L. L.: *History and Synopsis of the Theory of Summable Infinite Processes*, University of Oregon Press, 1925.
 - Van Vleck, E. B.: «Selected Topics in the Theory of Divergent Series and Continued Fractions». *The Boston Colloquium of the Amer. Math. Soc.*, 1903, Macmillan, 1905, 75-187.

Capítulo 48

El análisis tensorial y la geometría diferencial

Por tanto, o bien la realidad que sirve de base al espacio debe consistir en una variedad discreta, o bien tendremos que buscar el fundamento de su métrica en relaciones exteriores a ella, en las fuerzas de unión que actúan sobre ella. Esto nos conduce a los dominios de otra ciencia, los de la física, a los que el objetivo de nuestro trabajo no nos permite ir ahora.

Bernhard Riemann

Contenido:

- 1. Los orígenes del análisis tensorial*
- 2. El concepto de tensor*
- 3. Derivación covariante*
- 4. El desplazamiento paralelo*
- 5. Generalizaciones de la geometría riemanniana*

Bibliografía

1. Los orígenes del análisis tensorial

A menudo se habla del análisis tensorial como si se tratase de una rama totalmente nueva de la matemática, creada *ab initio*, bien para satisfacer algún objetivo concreto, o para deleite de los matemáticos. En realidad se trata solamente de una nueva variación sobre un viejo tema, a saber, el del estudio de los invariantes diferenciales asociados en principio a una geometría riemanniana. Recordemos (vid. cap. 37, sec. 5) que esos invariantes eran expresiones que conservaban su forma y su valor bajo cualquier cambio en el sistema de coordenadas, porque representaban propiedades puramente geométricas o físicas.

El estudio de los invariantes diferenciales había sido iniciado por Riemann, Beltrami, Christoffel y Lipschitz. El nuevo enfoque se debe a Gregorio Ricci-Curbastro (1853-1925), que fue profesor de matemáticas en la Universidad de Padua. Influyó mucho sobre él Luigi Bianchi, cuya obra había continuado la de Christoffel. Ricci trató de facilitar la búsqueda de propiedades geométricas y la expresión de leyes físicas en forma invariante bajo cambios del sistema de coordenadas. Realizó sus trabajos más importantes sobre el tema durante los años 1887-1896, aunque tanto él mismo como el resto de la escuela italiana continuaron trabajando en ello durante veinte años o más a partir de 1896. En esos primeros nueve años Ricci desarrolló sus planteamientos y elaboró un sistema de notación completo para la teoría, que llamó «cálculo diferencial absoluto». Presentó la primera exposición sistemática de su método en un artículo publicado en 1892,⁴⁸² aplicándolo a algunos problemas de geometría diferencial y de física.

Nueve años más tarde, y en colaboración con su famoso discípulo Tullio Levi-Civita (1873-1941), publicó un artículo general, «Métodos de cálculo diferencial absoluto», ofreciendo una formulación más definitiva de ese cálculo. El tema terminaría por conocerse como «análisis tensorial» desde que Einstein le diera ese nombre en 1916. Dados los muchos cambios de notación realizados por Ricci, y más tarde por Levi-Civita y Ricci, utilizaremos la notación aceptada actualmente de una manera bastante general.

2. El concepto de tensor

Para llegar a la idea de tensor tal como la introdujo Ricci, consideremos una función $A(x^1, x^2, \dots, x^n)$, cuyas derivadas parciales $\partial A / \partial x^i$ representaremos por A_i . Entonces la expresión

$$\sum A_i dx^i \quad (1)$$

es un invariante diferencial bajo transformaciones del tipo

$$x^i = f_i(y^1, y^2, \dots, y^n) \quad (2)$$

Supondremos que las funciones f_i tienen todas las derivadas necesarias y que la transformación es reversible, de manera que

$$y^i = g_i(x^1, x^2, \dots, x^n) \quad (3)$$

Bajo la transformación (2) la expresión (1) se convierte en

$$\sum \hat{A}_j(y^1, y^2, \dots, y^n) dy^j \quad (4)$$

Sin embargo, \hat{A}_j no será en general igual a A_j , sino que

$$\hat{A}_j = \frac{\partial \hat{A}}{\partial y^j} = A_1 \frac{\partial x^1}{\partial y^j} + A_2 \frac{\partial x^2}{\partial y^j} + \dots + A_n \frac{\partial x^n}{\partial y^j} \quad (5)$$

donde se entiende que las x^i han sido sustituidas en las A_j por sus valores en términos de las y^i . Así pues, las \hat{A} se relacionan con las A por medio de la ley de transformación (5), en la que intervienen las derivadas primeras de las ecuaciones de la transformación.

La idea básica de Ricci era que, en lugar de referirse a la forma diferencial invariante (1), sería suficiente, y más sencillo, considerar simplemente el conjunto de funciones

$$A_1, A_2, \dots, A_n$$

y llamarlas componentes de un tensor siempre que, bajo un cambio de coordenadas, el nuevo conjunto de componentes

$$\hat{A}_1, \hat{A}_2, \dots, \hat{A}_n$$

estuviera relacionado con el conjunto original por la ley de transformación (5). Este énfasis explícito en el conjunto o sistema

de funciones y en la ley de transformación es lo que caracteriza el enfoque de Ricci de la teoría de invariantes diferenciales. El conjunto de funciones A , que son, por cierto, las componentes del gradiente de la función escalar A , constituye un ejemplo de tensor covariante de rango 1. La idea de un conjunto o sistema de funciones que caracterizan una cantidad invariante no era en sí nada nuevo, porque los vectores eran ya bien conocidos en tiempos de Ricci. Efectivamente, los vectores vienen representados por sus componentes en un sistema de coordenadas, y también están sujetos a una ley de transformación del mismo tipo, si es que el vector ha de permanecer, como es debido, invariante bajo un cambio de coordenadas. Sin embargo, los nuevos sistemas que introducía Ricci eran mucho más generales, y el énfasis puesto en la ley de transformación también resultaba nuevo.

Como otro ejemplo del punto de vista introducido por Ricci, consideremos la expresión del elemento de longitud, que viene dada por

$$ds^2 = \sum_{i,j=1}^n g_{ij} dx^i dx^j \quad (6)$$

Bajo un cambio de coordenadas, el valor de la distancia ds debe permanecer invariante, por razones puramente geométricas. Sin embargo, si realizamos la transformación (2) y escribimos la nueva expresión en la forma

$$\bar{ds}^2 = \sum_{i,j=1}^n G_{ij} dy^i dy^j \quad (7)$$

entonces $g_{ij}(x^1, x^2, \dots, x^n)$ no será igual, en general, a $G_{ij}(y^1, y^2, \dots, y^n)$ (cuando los valores de las y^i representen el mismo punto que las correspondientes x^i). Lo que sí se verifica es que

$$G_{kl} = \sum_{i,j=1}^n g_{ij} \frac{\partial x^i}{\partial y^k} \frac{\partial x^j}{\partial y^l} \quad (8)$$

donde en las funciones g_{ij} se han sustituido las x^i por sus valores en términos de las y^i . Para ver que efectivamente se verifica (8) no tenemos más que sustituir dx^i en (6) por

$$dx^i = \sum_{k=1}^n \frac{\partial x^i}{\partial y^k} dy^k$$

y dx^j por

$$dx^j = \sum_{l=1}^n \frac{\partial x^j}{\partial y^l} dy^l$$

y obtener el coeficiente de $dy^k dy^l$. Así pues, aunque G_{kl} no tiene por qué coincidir con g_{kl} , sabemos cómo obtener las G_{kl} a partir de las g_{kl} . El conjunto de n^2 coeficientes g_{ik} de la forma cuadrática

fundamental constituye otro tensor, un tensor covariante de rango 2, cuya ley de transformación viene dada por (8).

Ricci introdujo también los tensores contravariantes. Consideremos la transformación inversa de la (2) que hemos manejado hasta ahora; si esa inversa es

$$y^j = g_j(x^1, x^2, \dots, x^n), \quad (9)$$

entonces

$$dy^j = \sum_{k=1}^n \frac{\partial y^j}{\partial x^k} dx^k \quad (10)$$

Si ahora consideramos a las dx^k como un conjunto de cantidades que constituyen un tensor, entonces vemos que, desde luego, $dy^j \neq dx^j$, pero podemos obtener las dy^j a partir de las dx^j mediante la ley de transformación (10). El conjunto de elementos dx^k recibe por tanto el nombre de tensor contravariante de rango 1, donde el adjetivo «contravariante» indica la presencia en la transformación de las derivadas parciales $\partial y^j / \partial x^k$ en lugar de las derivadas $\partial x^i / \partial y^j$ que aparecían en (5) y (8). Así pues, las diferenciales mismas de las variables que se transforman constituyen un tensor contravariante de rango 1.

También podemos tener un tensor de rango 2 que se transforme de manera contravariante en ambos índices. Si el conjunto de funciones $A^{kl}(x^1, x^2, \dots, x^n)$, $k, l = 1, 2, \dots, n$ se transforma al realizar el

cambio de variables (9) de manera que

$$\hat{A}^{ij} = \sum_{k,l=1}^n \frac{\partial y^i}{\partial x^k} \frac{\partial y^j}{\partial x^l} A^{kl} \quad (11)$$

entonces ese conjunto constituye un tensor contravariante de rango 2. Además, podemos tener lo que se llaman tensores mixtos, que se transforman de manera covariante en algunos índices y de manera contravariante en los demás. Por ejemplo, el conjunto de funciones A_{ij}^k , $i, j, k = 1, 2, \dots, n$, denota un tensor mixto en el que, de acuerdo con Ricci, los índices inferiores son aquellos en los que se transforma covariantemente, y el índice superior en el que se transforma contravariantemente. El tensor cuyos elementos son las A_{ij}^k se dice que es de rango 3. Podemos tener tensores covariantes, contravariantes y mixtos de rango r . Un tensor n -dimensional de rango r tendrá n^r componentes. La ecuación (21) del capítulo 37 muestra que el símbolo de cuatro índices de Riemann (rk, ih) es un tensor covariante de rango 4. Un tensor covariante de rango 1 es un vector. Levi-Civita le asociaba el vector contravariante que se define como sigue: si el conjunto de las λ_i son las componentes de un vector covariante, entonces el conjunto de las

$$\lambda^i = \sum_{k=1}^n g^{ik} \lambda_k$$

constituye el vector contravariante asociado a él; es un cociente

cuyo numerador es el menor complementario de g_{ik} en el determinante de las g_{ik} y cuyo denominador es g , el valor del determinante.

Se pueden realizar operaciones con los tensores. Por ejemplo, si tenemos dos tensores del mismo tipo, es decir, que tengan el mismo número de índices covariantes y contravariantes, podemos sumarlos, sumando las componentes con idénticos índices. Así,

$$A_i^j + B_i^j = C_i^j$$

Se debe y puede mostrar que las C_i constituyen un tensor covariante en el índice i y contravariante en el índice j .

También se pueden multiplicar dos tensores cualesquiera cuyos índices vayan de 1 al mismo n . Un ejemplo bastará para mostrar cómo funciona esa multiplicación. Si hacemos

$$A_i^h B_j^k = C_{ij}^{hk}$$

se puede comprobar que el tensor con n^4 componentes C_{ij} es covariante en los índices inferiores y contravariante en los superiores. No hay una operación de división para los tensores.

Lo que sí puede realizarse con ellos es una operación llamada «contracción», que ilustraremos con el siguiente ejemplo: dado el tensor de componentes A_{ir} , definimos las cantidades

$$B_i^h = \sum_{r=1}^n A_{ir}^{hr}$$

y puede probarse entonces que el conjunto de cantidades constituye un tensor de rango 2, covariante en el índice i , y contravariante en el índice h .

Resumiendo, un tensor es un conjunto de funciones (componentes) fijas con respecto a un sistema de referencia, o de coordenadas, que se transforman al realizar un cambio de éstas de acuerdo con ciertas leyes. Cada componente en un sistema de coordenadas es una función lineal homogénea de las componentes en otro sistema de coordenadas. Si las componentes de un tensor son iguales a las de otro en un determinado sistema de referencia, también son iguales en cualquier otro sistema de referencia. En particular, si las componentes se anulan en algún sistema de referencia, se anulan en cualquier otro. La igualdad de tensores es por tanto un invariante con respecto al cambio de coordenadas. El significado físico, geométrico, o incluso puramente matemático, que tiene un tensor en un sistema de referencia determinado, es preservado por la transformación, manteniéndolo en cualquier otro sistema de referencia. Esta propiedad es vital en la teoría de la relatividad, en la que cada observador posee su propio sistema de referencia; como las verdaderas leyes físicas son aquellas que se cumplen para todos los observadores, para reflejar esa independencia del sistema de coordenadas, esas leyes se expresan como tensores.

Disponiendo del concepto de tensor se puede reexpresar en forma

tensorial muchos de los conceptos de la geometría riemanniana. El más importante es quizá la curvatura del espacio. La curvatura riemanniana (vid. cap. 37, sec. 3) se puede expresar tensorialmente de muchas maneras. Las expresiones modernas utilizan el convenio de sumación introducido por Einstein, según el cual se sobreentiende la suma siempre que aparece un índice repetido en el producto de dos símbolos. Así, por ejemplo,

$$g^{ij}\lambda_j = \sum_{j=1}^n g^{ij}\lambda_j$$

Con esta notación, el tensor de curvatura (vid. [20] del cap. 37) es:

$$R_{\lambda\mu\rho\sigma} = \frac{\partial}{\partial x^\rho} [\mu\sigma, \lambda] - \frac{\partial}{\partial x^\sigma} [\mu\rho, \lambda] + \{\mu\rho, \varepsilon\}[\lambda\sigma, \varepsilon] - \{\mu\sigma, \varepsilon\}[\lambda\rho, \varepsilon]$$

o lo que es equivalente,

$$R_{jkl}^i = \frac{\partial}{\partial x^l} \{jk, i\} - \frac{\partial}{\partial x^k} \{jl, i\} - [\{sk, i\}\{jl, s\} - \{sl, i\}\{jk, s\}]$$

donde los corchetes denotan símbolos de Christoffel de primera especie, y las llaves símbolos de Christoffel de segunda especie. Cualquiera de esas formas es lo que se llama ahora tensor de curvatura de Riemann-Christoffel. A causa de ciertas relaciones (que no nos entretendremos en describir ahora) entre las

componentes, el número de componentes distintas de ese tensor es $n^2(n^2 - 1)/12$. Para $n = 4$, que es el caso en la teoría general de la relatividad, ese número de componentes distintas es 20. En una variedad riemanniana bidimensional todas las componentes son iguales, por ejemplo, a R_{1212} . En este caso la curvatura total de Gauss K no es sino

$$K = \frac{R_{1212}}{g}$$

donde g es el determinante de las g_{ij} , es decir, $g_{11}g_{22} - g_{12}^2$. Si todas las componentes son nulas, el espacio es euclídeo.

A partir del tensor de Riemann-Christoffel, Ricci obtuvo por contracción el que se llama ahora tensor de Ricci o tensor de Einstein. Sus componentes R_{jl} son $\sum_{k=1}^n R_{jlk}$. Este tensor fue utilizado por Einstein,⁴⁸⁴ con $n = 4$, para expresar la curvatura de su geometría riemanniana espacio-temporal.

3. Derivación covariante

Ricci también introdujo en el análisis tensorial una operación a la que él y (más tarde) Levi-Civita llamaron derivación covariante.⁴⁸⁵ Esta operación había aparecido ya en los trabajos de Christoffel y Lipschitz;⁴⁸⁶ Christoffel había presentado un método (vid. cap. 37, sec. 4) con el que se podían obtener, a partir de invariantes diferenciales formados con las derivadas de la forma fundamental para ds^2 y de ciertas funciones $\gamma(x_1, x_2, \dots, x_n)$ invariantes asociados a

derivadas de orden superior. Ricci constató la importancia de ese método para su análisis tensorial, y lo adoptó.

Mientras que Christoffel y Lipschitz realizaban la derivación covariante de la forma completa, Ricci, que ponía más énfasis en las componentes de los tensores, trabajó sobre ellas. En particular, si $A_1(x_1, x_2, \dots, x_n)$ es una componente covariante de un vector o tensor de rango 1, la derivada covariante de A_i no es simplemente su derivada con respecto a x^l , sino el tensor de rango 2

$$A_{i,l} = \frac{\partial A_i}{\partial x_j} - \sum_{j=1}^n \{il, j\} A_j \quad (12)$$

donde las llaves indican el símbolo de Christoffel de segunda especie. De igual modo, si A_{ik} es una componente de un tensor covariante de rango 2, su derivada covariante con respecto a x^l viene dada por

$$A_{ik,l} = \frac{\partial A_{ik}}{\partial x^l} \sum_{j=1}^n \{il, j\} A_{jk} - \sum_{j=1}^n \{kl, j\} A_{ij} \quad (13)$$

Para un tensor contravariante de rango 1 con componentes A^i la derivada covariante A^i_l viene dada por

$$A^i_l = \frac{\partial A^i}{\partial x^l} + \sum_{j=1}^n A^j \{jl, i\}$$

y es un tensor mixto de segundo orden. Para el tensor mixto con componentes A_i^h la derivada covariante es

$$A_{i,l}^h = \frac{\partial A_i^h}{\partial x^l} - \sum_{j=1}^n A_j^h \{il, j\} + \sum_{j=1}^n A_i^j \{jl, h\}$$

La derivada covariante de un invariante escalar y es el vector covariante cuyas componentes vienen dadas por $\gamma_i = \partial_V / \partial x_i$. A ese vector se le llama gradiente del invariante escalar.

Desde el punto de vista puramente matemático, la derivada covariante de un tensor es otro tensor cuyo rango es una unidad mayor en los índices covariantes. Se trata de un hecho importante, ya que posibilita el tratamiento de tales derivadas en el marco general del análisis tensorial. También tiene significado geométrico: supongamos que tenemos un campo vectorial constante en el plano, esto es, un conjunto de vectores, anclado cada uno de ellos en un punto distinto, pero con la misma magnitud y dirección. En tal caso, las componentes con respecto a un sistema rectangular de coordenadas son también constantes. Sin embargo, las componentes de esos vectores con respecto al sistema polar de coordenadas, es decir, una componente a lo largo del radio vector y la otra perpendicular a ese radio vector cambian de un punto a otro, porque las direcciones en que se toman esas componentes también cambian de un punto a otro. Si se obtienen las derivadas con respecto a las coordenadas, digamos r y v , de esas componentes, el

cambio expresado por esas derivadas, que ya no son nulas, refleja el cambio en las componentes debido al sistema de coordenadas, y no un cambio en los vectores mismos. Las coordenadas utilizadas en geometría riemanniana son curvilíneas; el efecto de la curvatura de esas coordenadas viene dado por los símbolos de Christoffel de segunda especie (denotados aquí mediante llaves). La derivada covariante completa de un tensor representa la tasa de cambio efectiva de la cantidad física o geométrica representada por ese tensor, así como el cambio debido al sistema de coordenadas.

En los espacios euclídeos, en los que la ds^2 se puede siempre reducir a una suma de cuadrados con coeficientes constantes, la derivada covariante se reduce a la ordinaria, ya que los símbolos de Christoffel son nulos. También es nula la derivada covariante de cada g_{ij} que aparece en la métrica riemanniana. Esto fue demostrado por el mismo Ricci,⁴⁸⁷ y se conoce como lema de Ricci.

El concepto de derivación covariante nos permite expresar con facilidad para los tensores generalizaciones de nociones ya conocidas en análisis vectorial, que pueden tratarse ahora en geometría riemanniana. Así, si $A_i(x_1, x_2, \dots, x_n)$ son las componentes de un vector n -dimensional A , entonces

$$\theta = \sum_{i,l=1}^n g^{il} A_{i,l} \quad (14)$$

donde g^{il} ha sido ya definido más arriba, es un invariante diferencial. Cuando la métrica fundamental es la de un sistema

rectangular de coordenadas en un espacio euclídeo (coordenadas cartesianas), las constantes g^{il} son nulas excepto para $i = l$, y en este caso las derivadas ordinaria y covariante son idénticas, (14) se convierte entonces en

$$\theta = \sum_{i=1}^n \frac{\partial A_i}{\partial x^i}$$

y constituye el análogo n -dimensional de lo que en tres dimensiones se llama la divergencia. Por eso (14) se llama también la divergencia del tensor cuyas componentes son las A_i . También se puede mostrar, utilizando (14), que si A es una función escalar, la divergencia del gradiente de A viene dada por

$$\Delta_2 A = \frac{1}{\sqrt{g}} \sum_{l=1}^n \frac{\partial}{\partial x^l} (\sqrt{g} A^l) \quad (15)$$

donde

$$A^l = \sum_{i=1}^n g^{il} \frac{\partial A}{\partial x^i} = \sum_{i=1}^n g^{il} A_i$$

Esta es también la expresión de Beltrami para $\Delta_2 A$ en una geometría riemanniana (vid. cap. 37, sec. 5).

Aunque Ricci y Levi-Civita dedicaron una gran parte de su artículo

de 1901 a la técnica del análisis tensorial, su principal objetivo era la búsqueda de invariantes diferenciales. Plantearon el siguiente problema general: dada una forma diferencial cuadrática positiva y un número arbitrario de funciones asociadas S , determinar todos los invariantes diferenciales absolutos que se pueden formar a partir de los coeficientes de y , las funciones S , y las derivadas de esos coeficientes y funciones hasta un orden dado m ; ofrecieron una solución completa: basta encontrar los invariantes *algebraicos* del sistema formado por la forma diferencial cuadrática fundamental y , las derivadas covariantes de las funciones asociadas S , hasta el orden m y, si $m > 1$, una cierta forma tetralineal G_4 cuyos coeficientes son los símbolos de Riemann (ih,jk) y sus derivadas covariantes hasta el orden $m - 2$.

Concluían su artículo mostrando cómo podían expresarse ciertas ecuaciones en derivadas parciales y leyes físicas en forma tensorial, haciéndolas así independientes del sistema de coordenadas. Este era el objetivo declarado por Ricci. El análisis tensorial se utilizó así para expresar la invarianza matemática de leyes físicas, muchos años antes de que Einstein lo empleara con el mismo propósito.

4. El desplazamiento paralelo

Desde 1901 hasta 1905 la investigación en análisis tensorial se vio limitada a un muy pequeño grupo de matemáticos. No obstante, la obra de Einstein cambió el panorama. Albert Einstein (1879-1955), cuando aún trabajaba como ingeniero en la oficina suiza de patentes, provocó una gran excitación en el mundo científico con el

anuncio de su teoría de la relatividad especial, o restringida.⁴⁸⁸ En 1914 aceptó un puesto en la Academia Prusiana de Ciencias en Berlín, como sucesor del célebre químico-físico Jacobus Van't Hoff (1852-1911). Dos años más tarde anunció su teoría general de la relatividad.⁴⁸⁹

Las propuestas revolucionarias de Einstein acerca de la relatividad de los fenómenos físicos despertaron un intenso interés entre los *físicos*, filósofos y matemáticos de todo el mundo. Los matemáticos se sintieron especialmente atraídos por la naturaleza de la geometría que Einstein había encontrado conveniente utilizar para la formulación de sus teorías.

La exposición de la teoría restringida, que implica las propiedades de variedades pseudoeuclídeas tetradimensionales (espacio-tiempo), se hace más cómoda con la ayuda de vectores y tensores; pero la teoría general, que depende de las propiedades de las variedades riemannianas tetradimensionales, exige el uso del cálculo tensorial especial asociado a tales variedades. Afortunadamente ese cálculo ya se había desarrollado, aunque no había atraído hasta entonces la atención de los físicos.

El trabajo de Einstein sobre la teoría restringida no utiliza de hecho geometría riemanniana ni análisis tensorial.⁴⁹⁰ Pero esa teoría restringida no tenía en cuenta la acción de la gravitación. Einstein comenzó entonces a trabajar sobre el problema de la fuerza gravitatoria y de cómo dar cuenta de su efecto imponiendo una estructura en la geometría espacio-temporal que hiciera que los objetos se movieran automáticamente a lo largo de los mismos

camino que los que se derivarían de tener en cuenta la acción de la fuerza gravitacional. En 1911 hizo pública una teoría que interpretaba de esa manera el efecto de una fuerza gravitatoria que tuviera dirección constante en todo el espacio, sabiendo desde luego que esa teoría no era realista. Hasta ese momento Einstein había utilizado únicamente los instrumentos matemáticos más simples, objetando incluso la necesidad de la «matemática elevada», de la que sospechaba que a menudo se introducía sólo para dificultar la lectura. Sin embargo, tratando de avanzar en su problema discutió sobre él en Praga con un colega, el matemático Georg Pick, que atrajo su atención hacia la teoría matemática de Ricci y Levi-Civita. En Zúrich encontró a un amigo, Marcel Grossmann (1878-1936), que le ayudó a comprender esa teoría y, con ella como base, consiguió formular la teoría general de la relatividad.

Para representar su mundo tetradimensional de tres coordenadas espaciales y una cuarta representando el tiempo, Einstein utilizó la métrica riemanniana

$$ds^2 = \sum_{i,j=1}^n g_{ij} dx_i dx_j \quad (16)$$

donde x_4 denota la coordenada temporal. Las g_i debían elegirse de manera que reflejaran la presencia de materia en las diferentes regiones del universo. Además, como la teoría se ocupa de la determinación de longitudes, tiempos, masas y otras cantidades físicas, por diferentes observadores que se mueven de manera

arbitraria unos con respecto a otros, los «puntos» del espacio-tiempo vienen representados en diferentes sistemas de coordenadas, ligado cada uno a un observador. La relación entre un sistema de coordenadas y otro viene dada por una transformación

$$x_i = \Phi_i(y_1, y_2, y_3, y_4) \quad i = 1, \dots, 4$$

Las leyes de la naturaleza son aquellas relaciones o expresiones que son iguales para todos los observadores. Así pues, son invariantes en el sentido matemático.

Desde el punto de vista de las matemáticas, la importancia de la obra de Einstein consistió, como ya hemos indicado, en la extensión del interés por el análisis tensorial y la geometría riemanniana. La primera innovación en análisis tensorial posterior a la teoría de la relatividad se debió a Levi-Civita. En 1917, mejorando una idea de Ricci, introdujo ⁴⁹¹ el concepto de desplazamiento paralelo de un vector. Gerhard Hessenberg introdujo independientemente esa noción en el mismo año.⁴⁹² Brouwer la había empleado ya en 1906 para superficies de curvatura constante. El objetivo de este concepto consiste en definir lo que se entiende por vectores paralelos en una variedad riemanniana. Se puede constatar la dificultad de hacerlo considerando la superficie de una esfera, que con la distancia determinada por arcos de círculo máximo es una variedad riemanniana. Si un vector, partiendo digamos de un paralelo y apuntando hacia el norte (el vector debe ser tangente a la superficie esférica) se mueve a lo largo de esa circunferencia no

máxima, manteniéndose paralelo a sí mismo en el espacio euclídeo tridimensional, cuando haya recorrido media circunferencia ya no será tangente a la esfera, y no pertenecerá al espacio en cuestión. Para obtener una noción de paralelismo de vectores válida en variedades riemannianas, hay que generalizar el concepto de euclídeo, aunque en el proceso se pierdan algunas de las propiedades usuales.

La idea geométrica utilizada por Levi-Civita para definir el transporte o desplazamiento paralelo se comprenderá más fácilmente para una superficie. Consideremos una curva C sobre la superficie, y desplacemos un vector anclado en un punto de C paralelamente a sí mismo en el siguiente sentido: en cada punto de C hay un plano tangente a la superficie. La envolvente de esa familia de planos es una superficie desarrollable, y cuando ésta se desarrolla sobre un plano, los vectores paralelos a lo largo de C deben ser paralelos en el plano euclídeo.

Levi-Civita generalizó esta idea a las variedades riemannianas n -dimensionales. Cuando un vector se desplaza paralelamente a sí mismo en el plano euclídeo a lo largo de una línea recta —una geodésica del plano— el vector forma siempre el mismo ángulo con la recta. De acuerdo con esto, el paralelismo en una variedad riemanniana se define así: cuando un vector se mueve a lo largo de una geodésica, debe seguir formando el mismo ángulo con la geodésica (con la tangente a la geodésica). En particular, una tangente a la geodésica se mantiene paralela a sí misma cuando se desplaza a lo largo de ésta. Por definición, el vector debe mantener

la misma magnitud. Se sobreentiende que el vector se mantiene tangente a la variedad riemanniana, incluso si ésta está inmersa en un espacio euclídeo. La definición de transporte paralelo requiere también que se mantenga el ángulo formado por dos vectores cuando ambos se desplazan paralelamente a lo largo de la misma curva C . En el caso general de transporte paralelo a lo largo de una curva cerrada arbitraria C , el vector inicial y el final no tendrán usualmente la misma dirección (euclídea). La desviación en la dirección dependerá del camino C . Consideremos por ejemplo un vector que parte de un punto P perteneciente a un paralelo C de una superficie esférica, tangente al correspondiente meridiano que pasa por P ; cuando se realiza el *transporte paralelo* de ese vector a lo largo de C , al volver a P seguirá siendo tangente a la superficie esférica, pero formará un ángulo de valor $2\pi(1 - \cos \gamma)$ con el vector de partida, donde γ es la colatitud de P .

Si se utiliza la definición general de desplazamiento paralelo a lo largo de una curva de una variedad riemanniana, se obtiene una condición analítica. La ecuación diferencial que satisfacen las componentes X^α de un vector contravariante que se desplaza paralelamente a lo largo de una curva es (sobreentendiendo la suma en los índices repetidos):

$$\frac{dX^\alpha}{dt} + \{\beta\gamma, \alpha\}X^\beta \frac{du^\gamma}{dt} = 0 \quad \alpha = 1, 2, \dots, n \quad (17)$$

donde las $u^i(t)$, $i = 1, 2, \dots, n$ definen la curva. Para un vector

covariante X_α la condición es

$$\frac{dX_\alpha}{dt} - \{a^l, j\}X_j \frac{du^l}{dt} = 0 \quad \alpha = 1, 2, \dots, n \quad (18)$$

Esas ecuaciones se pueden utilizar para definir el transporte paralelo a lo largo de cualquier curva C . La solución determinada unívocamente por los valores de las componentes en un punto dado P es un vector con valores en cada punto de C y paralelo por definición al vector inicial en P . La ecuación (18) establece que la derivada covariante de X_α es 0.

Una vez introducida la noción de desplazamiento paralelo, se puede describir en términos de él la curvatura del espacio; concretamente, en términos del cambio experimentado por un vector infinitesimal al ser transportado paralelamente a distancias infinitesimales. El paralelismo está en la base del concepto de curvatura incluso en un espacio euclídeo, ya que la curvatura de un arco infinitesimal depende del cambio de dirección de un vector tangente a lo largo del arco.

5. Generalizaciones de la geometría riemanniana

El éxito que tuvo la utilización de la geometría riemanniana en la teoría de la relatividad resucitó el interés por ella. Sin embargo, la obra de Einstein planteó una cuestión incluso más general: había incorporado el efecto gravitacional de la masa en el espacio utilizando las funciones adecuadas para las g_{ij} ; como consecuencia,

las geodésicas de su espacio-tiempo eran precisamente los caminos recorridos por objetos que se movieran libremente, como por ejemplo lo hace la Tierra en torno al Sol. A diferencia de lo que ocurría en la mecánica newtoniana, no se requería una fuerza gravitatoria para explicar el camino recorrido. La eliminación de la gravedad sugirió la posibilidad de explicar también la atracción y repulsión entre las cargas eléctricas en términos de la métrica del espacio. Tal resultado proporcionaría una teoría unificada de la gravitación y el electromagnetismo. Estos trabajos condujeron a generalizaciones de la geometría riemanniana conocidas colectivamente con el nombre de geometrías no-riemannianas.

En geometría riemanniana la ds^2 liga entre sí los diferentes puntos del espacio, especificando la relación entre ellos por la distancia que los separa. En las geometrías no-riemannianas la conexión entre diferentes puntos se especifica de otras maneras, que no tienen que basarse necesariamente en una métrica. La variedad de tales geometrías es muy amplia, y cada una de ellas tiene un desarrollo tan extenso como la misma geometría riemanniana. Aquí daremos sólo algunos ejemplos de las ideas básicas de esas geometrías.

Quien inició primeramente esta área de trabajo fue Hermann Weyl,⁴⁹³ y las geometrías que introdujo se conocen como espacios con una conexión afin. En geometría riemanniana, la demostración de que la derivada covariante de un tensor es también un tensor depende sólo de relaciones de la forma

$$\{\overline{ik, h}\} = \{ab, j\} \frac{\partial x^a}{\partial y^i} \frac{\partial x^b}{\partial y^k} \frac{\partial x^h}{\partial y^j} + \frac{\partial^2 x^j}{\partial y^i \partial y^k} \frac{\partial y^h}{\partial x^j} \quad (19)$$

en donde el miembro de la izquierda es la transformada de $\{ab, j\}$ por el cambio de coordenadas que pasa de las x^i a las y^l . Los símbolos de Christoffel, que se definen en términos de los coeficientes de la forma fundamental, satisfacen esas relaciones. Consideremos en su lugar funciones L^i_{jk} y \underline{L}^i_{jk} de las x^i e y^i respectivamente, que satisfagan las mismas relaciones (19) pero sin estar necesariamente relacionadas con la forma cuadrática fundamental. Un conjunto de funciones L^i_{jk} asociadas a un espacio V_n que tengan la propiedad de transformación (19) constituyen lo que se llama una conexión afín; las funciones mismas se llaman coeficientes de la conexión afín, y se dice que el espacio V_n posee una conexión afín, o que es un espacio afín. La geometría riemanniana resulta como caso especial cuando los coeficientes de la conexión afín son los símbolos de Christoffel de segunda especie, que se obtienen a partir del tensor fundamental del espacio. Dadas las funciones L^i_{jk} , es posible introducir conceptos tales como derivación covariante, curvatura, y otras nociones análogas a las de la geometría riemanniana. Sin embargo, en esa nueva geometría ya no se puede hablar de la magnitud de un vector.

Es un espacio con conexión afín, una curva tal que sus tangentes sean paralelas con respecto a la curva (en el sentido del desplazamiento paralelo en ese espacio) es lo que se llama un camino propio del espacio. Esos caminos propios constituyen una

generalización de las geodésicas de una variedad riemanniana. Todos los espacios conectados afinmente que tienen las mismas funciones L tienen los mismos caminos propios. La geometría de los espacios conectados afinmente puede prescindir así de la métrica riemanniana. Weyl obtuvo las ecuaciones de Maxwell a partir de las propiedades del espacio, pero la teoría en su conjunto no casaba demasiado bien con otros hechos establecidos.

Otra geometría no riemanniana, debida a Luther P. Eisenhart (1876-1965) y Veblen,⁴⁹⁴ llamada geometría de los caminos, funciona de manera un poco diferente: se parte de n^3 funciones $\Gamma_{\lambda\mu}^i$ de x^1, \dots, x^n ; el sistema de n ecuaciones diferenciales

$$\frac{d^2 x^i}{ds^2} + \sum_{\lambda\mu} \Gamma_{\lambda\mu}^i \frac{dx^\lambda}{ds} \frac{dx^\mu}{ds} = 0 \quad i = 1, 2, \dots, n \quad (20)$$

con $\Gamma_{\lambda\mu}^i = r_{\lambda\mu}^i$ define entonces una familia de curvas, llamadas caminos, que son las geodésicas de esa geometría (en la geometría riemanniana las ecuaciones (20) son precisamente las de las geodésicas); a partir de ellas se puede entonces construir la geometría de los caminos de una manera análoga a la de la geometría riemanniana.

Otra generalización distinta de la geometría riemanniana se debe a Paul Finsler (1894), quien la expuso en su excelente tesis de 1918 en Göttingen.⁴⁹⁵ La ds^2 riemanniana se sustituye por una función más general $F(x, dx)$ de las coordenadas y sus diferenciales, sobre la

que se imponen restricciones que aseguren la posibilidad de minimizar la integral $\int F(x, (dx/dt)) dt$ obteniendo así las geodésicas. No obstante, todos los intentos realizados hasta el momento por generalizar el concepto de geometría riemanniana, incorporando tanto el electromagnetismo como los fenómenos gravitatorios, han resultado infructuosos. Aun así, los matemáticos siguen trabajando en esas geometrías abstractas.

Bibliografía

- Cartan, E.: «*Les récentes généralisations de la notion d'espace*», Bull. des Sci. Math., 48, 1924, 294-320.
- Pierpont, James: «*Some modern views of space*», Amer. Math. Soc. Bull., 32, 1926, 225-258.
- Ricci-Curbastro, G.: *Opere*, 2 vol., Edizioni Cremonese, 1956-1957. Ricci-Curbastro, G. y T. Levi-Civita: «*Méthodes de calcul différentia 1 absolu et leurs applications*», Math. Ann., 54, 1901, 125-201.
- Thomas, T. Y.: «*Recent trends in geometry*», Amer. Math. Soc. *Semicentennial Publications*, II, 1938, 98-135.
- Weatherburn, C. E.: *The development of multidimensional differential geometry*, Australian and New Zealand Association for the Advancement of Science, 21, 1933, 12-28.
- Weitzenböck, R.: «*Neuere arbeiten der algebraischen Invariantentheorie. Differentialinvarianten*», *Encyk. der Math. Wiss.*, B. G. Teubner, 1902-1927, II, parte III, E1, 1-71.
- Weyl, H.: *Mathematische Analyse des Raumproblems* (1923),

Chelsea (reimpresión), 1964.

Capítulo 49

La aparición del álgebra abstracta

Quizá pueda, sin inmodestia, reclamar para mí mismo la denominación de Adán Matemático, porque creo que he dado más nombres (que han pasado a ser de uso general) a criaturas de la razón matemática que todos los restantes matemáticos de la época juntos.

J. J. Sylvester

Contenido:

- 1. El panorama del siglo XIX*
 - 2. La teoría abstracta de grupos*
 - 3. La teoría abstracta de cuerpos*
 - 4. Anillos, 1518*
 - 5. La teoría de álgebras no asociativas*
 - 6. Panorama general del álgebra abstracta*
- Bibliografía*

1. El panorama del siglo XIX

En el campo del álgebra abstracta, lo mismo que en la mayoría de

los desarrollos matemáticos del siglo XX, tanto los conceptos como las metas fundamentales se fijaron a lo largo del siglo XIX. El hecho básico de que el álgebra puede tratar sistemas de objetos que no sean necesariamente números reales o complejos, quedó definitivamente demostrado en una buena docena de creaciones del siglo XIX. Ejemplos de objetos que se combinaban mediante operaciones y bajo leyes para dichas operaciones peculiares del sistema en cuestión son los vectores, cuaterniones, matrices, formas tales como la $ax^2 + bxy + cy^2$, hipernúmeros de diversos tipos, transformaciones, y sustituciones o permutaciones. Incluso en la teoría de números algebraicos, a pesar de que trata de clases de números complejos, aparecían en escena diversidad de álgebras, porque estas clases sólo tenían algunas de las propiedades del sistema completo de los números complejos.

Estas diversas clases de objetos se distinguieron y se clasificaron de acuerdo con las propiedades de las operaciones definidas sobre ellas, y ya hemos visto que se introdujeron nociones tales como las de grupo, anillo, ideal y cuerpo, y nociones subordinadas como las de subgrupo, subgrupo invariante o extensión de un cuerpo, para identificar conjuntos concretos de propiedades. Sin embargo, a lo largo de casi todo el siglo XIX los trabajos sobre esos diversos tipos de álgebras se referían a los sistemas concretos que hemos mencionado más arriba. Solamente durante las últimas décadas del siglo se dieron cuenta los matemáticos de que podían ascender a un nivel de eficacia nuevo integrando juntas muchas álgebras hasta entonces dispersas, por abstracción de su contenido común. Así,

por ejemplo, los grupos de permutaciones, los grupos de clases de formas estudiados por Gauss, los hipernúmeros con su suma y los grupos de transformaciones podían ser estudiados todos de un solo golpe considerando un conjunto de cosas o elementos sometidos a una operación cuya naturaleza quedaba especificada únicamente por ciertas propiedades abstractas, siendo la primera y más importante de las cuales que dicha operación, aplicada a dos elementos del conjunto, produce otro elemento del mismo conjunto. Y las mismas ventajas podían conseguirse para las diversas colecciones que constituían anillos o cuerpos, evidentemente. Aunque la idea de trabajar con colecciones abstractas fue anterior a las axiomáticas de Pasch, Peano y Hilbert, los desarrollos posteriores en este sentido aceleraron indudablemente la aceptación general de los planteamientos abstractos dentro del álgebra.

Así surgió el álgebra abstracta como el estudio explícitamente consciente de clases enteras de álgebras, las cuales, individualmente consideradas, no sólo eran sistemas concretos sino que servían para fines también concretos en áreas específicas de la matemática, como ocurría con los grupos de sustituciones en la teoría de ecuaciones. La ventaja de obtener resultados que podían ser útiles en muchas ramas concretas de la matemática considerando versiones abstractas, pronto se perdió de vista, y el estudio de las estructuras abstractas y de sus propiedades se convirtió en un fin en sí mismo.

El álgebra abstracta ha sido uno de los campos más favorecidos de la matemática del siglo XX, y hoy sus dominios son muy extensos.

Nosotros sólo presentaremos aquí los comienzos de esta materia e indicaremos las posibilidades casi ilimitadas que en ella ofrece la investigación. La mayor dificultad para exponer lo que se ha ido haciendo en este campo es la de la terminología. Aparte de las dificultades usuales de que diferentes autores usen términos distintos y de que los términos cambien de significado de un período a otro, el álgebra abstracta viene marcada por la introducción de cientos de términos nuevos. Toda variación en un concepto, por pequeña que sea, aparece distinguida por un término nuevo e impresionante; un diccionario completo de los términos utilizados llenaría un libro de buen tamaño.

2. La teoría abstracta de grupos

La primera estructura abstracta que se introdujo y estudió fue la de grupo. Muchas de las ideas básicas de la teoría de grupos abstractos se pueden encontrar ya, tanto implícita como explícitamente, en época tan temprana como el 1800, por lo menos. Una de las actividades favoritas de los historiadores de la matemática, ahora que existe la teoría abstracta, es la de rastrear cuántas de las ideas abstractas aparecen ya anunciadas en las obras concretas de Gauss, Abel, Galois, Cauchy y docenas de otros. Aquí no dedicaremos espacio a esta reconstrucción del pasado. El único punto importante que vale la pena mencionar es el de que, una vez adquirida la noción abstracta, fue relativamente fácil para los fundadores de la teoría de grupos abstractos obtener ideas y resultados reformulando obras del pasado.

Antes de examinar el desarrollo del concepto abstracto de grupo, puede ser interesante saber qué es lo que pretendían los matemáticos. La definición abstracta de grupo que se suele utilizar hoy se refiere a una colección de elementos, en cantidad finita o infinita, y a una operación que, aplicada a dos elementos cualesquiera de la colección, produce otro de la misma (propiedad de cierre). La operación es asociativa; hay un elemento tal que para cualquier elemento a del grupo $ae = ea = a$; y para cada elemento a existe un elemento inverso a' tal que $a'a = aa' = e$. Cuando la operación es conmutativa, el grupo se llama conmutativo o abeliano, y la operación suele llamarse suma y representarse por $+$; el elemento e se representa entonces por 0 y se le llama elemento cero. Si la operación no es conmutativa se le suele llamar multiplicación, y el elemento e se representa por 1 y se llama identidad.

El concepto de grupo abstracto y las propiedades que lo caracterizan surgieron lentamente. Podemos recordar (cap. 31, sec. 6) que Cayley había propuesto ya la idea de grupo abstracto en 1849, pero en esta época no se reconoció su importancia. En 1858 Dedekind,⁴⁹⁶ muy por delante de su época, dio una definición abstracta de los grupos finitos, derivada de los grupos de permutaciones. En 1877⁴⁹⁷ observó de nuevo que sus módulos de números algebraicos, a los que pertenecían $\alpha + \beta$ y $\alpha - \beta$ si pertenecían α y β se podían generalizar de manera que los elementos no fueran ya números algebraicos, y la operación podía ser arbitraria con tal de tener un inverso y ser conmutativa; así pues, sugería un grupo finito abstracto conmutativo. La visión de

Dedekind del valor de la abstracción es notable. Se dio cuenta claramente, trabajando en la teoría de números algebraicos, del valor de estructuras tales como los ideales y los cuerpos; él fue el verdadero creador del álgebra abstracta.

Kronecker,⁴⁹⁸ siguiendo los trabajos sobre los números ideales de Kummer, dio también lo que podemos considerar la definición abstracta de un grupo abeliano finito, análoga a la de Cayley de 1849. Kronecker habla expresamente de elementos abstractos, una operación abstracta, la propiedad de clausura, las propiedades asociativa y conmutativa y la existencia de un inverso único para cada elemento, y a continuación demuestra varios teoremas. Entre las diversas potencias de un elemento cualquiera θ , hay una que es igual al elemento unidad 1; si ν es el mínimo exponente para el que $\theta^\nu = 1$, entonces para cada divisor μ de ν hay un elemento ϕ tal que $\phi^\mu = 1$. Si ϕ^ρ y ϕ^σ son ambos iguales a 1 y ρ y σ son los mínimos exponentes para lo que esto ocurre y son primos entre sí, entonces $(\theta\phi)^{\rho\sigma} = 1$. Kronecker también dio la primera demostración de lo que hoy se llama un teorema de la base: existe un sistema finito fundamental de elementos $\theta_1, \theta_2, \theta_3, \dots$ tales que los productos

$$\theta_1^{h_1} \theta_2^{h_2} \theta_3^{h_3} \dots \quad h_i = 1, 2, 3, \dots, n$$

representan todos los elementos del grupo exactamente una vez. Los mínimos valores posibles de n_1, n_2, n_3, \dots que corresponden a $\theta_1, \theta_2, \theta_3, \dots$ (es decir, para los cuales $\theta^i = 1$) son tales que cada uno es divisible por el siguiente y el producto n_1, n_2, n_3, \dots es igual al número

de elementos n del grupo. Además todos los factores primos de n están en n_1 .

En 1878 escribió Cayley cuatro artículos más sobre grupos abstractos finitos.⁴⁹⁹ En ellos, como en los de 1849 y 1854, subraya que un grupo puede ser considerado como un concepto general y no necesita limitarse a los grupos de sustituciones aunque, señala, todo grupo (finito) pueda ser representado como un grupo de sustituciones. Estos artículos de Cayley tuvieron más influencia que los anteriores, porque la época estaba ya madura para una abstracción que abarcaba más que los grupos de sustituciones.

En un artículo conjunto de Frobenius y Ludwig Stickelberger (1850-1936)⁵⁰⁰ se da el importante paso de reconocer que el concepto abstracto de grupo incluye las congruencias y la composición de formas de Gauss, así como los grupos de sustituciones de Galois. Se menciona además la existencia de grupos de orden infinito.

Eugen Netto, en su libro *Substitutionentheorie und ihre Anwendung auf die Algebra* (1882) se limitaba a tratar grupos de sustituciones, pero los enunciados de sus conceptos y teoremas permitían reconocer el carácter abstracto de dichos conceptos. Aparte de reunir resultados de sus predecesores, Netto trata de los conceptos de isomorfismo y homomorfismo. El primero significa una correspondencia biunívoca entre dos grupos, tal que si $ab = c$, donde a , b y c son elementos del primer grupo, entonces $a'b' = c'$, siendo a' , b' y c' los elementos correspondientes del segundo grupo. Un homomorfismo es una correspondencia en general, tal que de nuevo $ab = c$ implica $a' b' = c'$.

Hacia 1880 surgieron nuevas ideas sobre grupos. Klein, influenciado por la obra de Jordán sobre grupos de permutaciones, había demostrado en su Erlanger Programm (cap. 38, sec. 5) que se podía usar grupos infinitos de transformaciones para clasificar las geometrías. Estos grupos son además continuos en el sentido de que en cualquier grupo se incluyen transformaciones arbitrariamente pequeñas o, dicho de otra manera, los parámetros de las transformaciones pueden tomar todos los valores reales. Así, en las transformaciones que expresan giros de los ejes, el ángulo d puede tomar todos los valores reales. En sus trabajos sobre funciones automorfas, Klein y Poincaré habían utilizado otro grupo de tipo infinito, un grupo no continuo o discreto (cap. 29, sec. 6).

Sophus Lie, que había trabajado con Klein en torno a 1870, adoptó el concepto de grupo continuo de transformaciones, pero para otro objetivo que el de clasificar las geometrías. Había observado que la mayor parte de las ecuaciones diferenciales ordinarias que se habían integrado por viejos métodos eran invariantes bajo clases de grupos de transformaciones continuas, y pensó que esto podría arrojar luz sobre la resolución de ecuaciones diferenciales y clasificarlas.

En 1874 introdujo Lie su teoría general de grupos de transformaciones.⁵⁰¹ Un grupo viene representado por un sistema de ecuaciones

$$x'_i = f_i(x_1, x_2, \dots, x_n, a_1, \dots, a_{1n}) \quad i = 1, 2, \dots, n \quad (1)$$

donde las f_i son funciones analíticas de las x_i y a_i . Las a_i son parámetros mientras que las x_i son las variables, y (x_1, x_2, \dots, x_n) representa un punto del espacio n -dimensional. Tanto los parámetros como las variables pueden tomar todos los valores reales o complejos. Por ejemplo, en una dimensión, la clase de transformaciones

$$x' = \frac{ax + b}{cx + d}$$

donde a , b , c y d pueden tomar todos los valores reales, es un grupo continuo. Los grupos representados por (1) se llaman finitos, donde la palabra finito se refiere al número de parámetros; el número de transformaciones es infinito, desde luego. El caso unidimensional anterior es un grupo tri-paramétrico porque sólo importan las razones de a , b y c a d . En el caso general, el producto de dos transformaciones

$$x'_i = f_i(x_1, \dots, x_n, a_1, \dots, a_n)$$

$$x''_i = f_i(x'_1, \dots, x'_n, b_1, \dots, b_n)$$

es la transformación

$$x''_i = f_i(x_1, \dots, x_n, c_1, \dots, c_n)$$

donde las c son funciones de las a_i y las b_i . En el caso de una

variable, Lie llamó al grupo una variedad simplemente extendida, y en el de n variables una variedad múltiplemente extendida.

En un artículo de 1883 sobre grupos continuos, publicado en una obscura revista noruega,⁵⁰² introduce Lie los grupos continuos de transformaciones infinitos. Estos no vienen definidos por ecuaciones tales como la (1), sino por medio de ecuaciones diferenciales. Las transformaciones resultantes no dependen ahora de un número finito de parámetros continuos, sino de funciones arbitrarias. No hay un concepto abstracto de grupo correspondiente a estos grupos continuos infinitos, y aunque se ha investigado mucho sobre ellos, no los consideraremos aquí.

Quizá sea de interés observar que, al comienzo de sus trabajos, Klein y Lie definían un grupo de transformaciones como uno que posee sólo la propiedad de clausura. Las otras propiedades, tales como la existencia de inversa de cada transformación, elemento neutro, etc. se establecían utilizando las propiedades de las transformaciones o, como en el caso de la propiedad asociativa, se utilizaban como propiedades obvias de las transformaciones. Lie reconoció durante el curso de su trabajo que habría que postular como parte de la definición de un grupo la existencia de inverso de cada elemento.

Hacia 1880 se conocían ya cuatro tipos principales de grupos. Los grupos discontinuos de orden finito, como los grupos de sustituciones; los grupos discontinuos (o discretos) infinitos, tales como los que aparecen en la teoría de funciones automorfas; los grupos continuos finitos de Lie, como los grupos de

transformaciones de Klein y las transformaciones analíticas más generales de Lie; y los grupos continuos infinitos de Lie definidos por ecuaciones diferenciales.

Con la obra de Walther von Dyck (1856-1934), las tres raíces principales de la teoría de grupos, la teoría de ecuaciones, la teoría de números y los grupos de transformaciones infinitos, quedaron todas incluidas en el concepto abstracto de grupo. Dyck estaba influenciado por Cayley y fue discípulo de Félix Klein. En 1882 y 1883⁵⁰³ publicó artículos sobre teoría de grupos abstractos, que incluía los grupos continuos y discretos. Su definición de grupo se refiere a un conjunto de elementos y una operación que satisfaga la propiedad de clausura, la asociativa, y la existencia de inverso de cada elemento, pero no la conmutativa.

Dyck estudia de manera explícita el concepto de generadores de un grupo, que estaba implícito ya en el teorema de la base de Kronecker y explícito en los trabajos de Netto sobre grupos de sustituciones. Los generadores constituyen un subconjunto fijo de elementos independientes de un grupo, tales que todo elemento del mismo se puede expresar como producto de potencias de los generadores y sus inversos. Cuando no hay restricción alguna sobre los generadores el grupo se llama un grupo libre. Si los generadores son A_1, A_2, \dots entonces una expresión de la forma donde los son enteros positivos o negativos, se llama una palabra. Puede haber relaciones entre los generadores, que serán de la forma

$$A_1^{\mu_1} A_2^{\mu_2} \dots,$$

donde los μ son enteros positivos o negativos, se llama una palabra. Puede haber relaciones entre los generadores, que serán de la forma

$$F_i(A_i) = 1$$

es decir, una palabra o combinación de palabras igual al elemento unidad del grupo. Dyck demuestra entonces que la presencia de relaciones implica un subgrupo invariante y un grupo cociente \underline{G} de grupo libre G . En su artículo de 1883 aplica la teoría de grupos abstractos a los grupos de permutaciones, grupos finitos de rotaciones (simetrías de poliedros), grupos de la teoría de números y grupos de transformaciones.

Huntington,⁵⁰⁴ E. H. Moore,⁵⁰⁵ y Leonard E. Dickson (1874-1954)⁵⁰⁶ dieron todos ellos conjuntos de postulados independientes para el concepto de grupo abstracto; éstos, así como otros sistemas de postulados, son en realidad pequeñas variaciones unos de otros.

Habiendo conquistado ya la noción abstracta de grupo, los matemáticos se dedicaron a demostrar teoremas sobre grupos abstractos sugeridos por resultados conocidos para casos concretos. Así, Frobenius ⁵⁰⁷ demostró el teorema de Sylow (cap. 31, sec. 6) para grupos abstractos finitos: todo grupo finito cuyo orden, es decir, su número de elementos, sea divisible por la potencia v de un primo p , contiene siempre un subgrupo de orden p^v .

Aparte de buscar grupos concretos para propiedades que podían verificarse en grupos abstractos, muchos matemáticos introdujeron

conceptos directamente para los grupos abstractos. Dedekind ⁵⁰⁸ y George A. Miller (1863-1951)⁵⁰⁹ estudiaron los grupos no abelianos en que todo subgrupo es normal (o invariante). Dedekind, en su artículo de 1897 y Miller⁵¹⁰ introdujeron el concepto de conmutador y subgrupo conmutador; si s y t son elementos de un grupo G , al elemento $s^{-1} t^{-1} st$ se le llama conmutador de s y t . Tanto Dedekind como Miller utilizaron esta concepto en sus teoremas; por ejemplo, el conjunto de todos los conmutadores de los pares ordenados de elementos de un grupo G generan un subgrupo invariante de G . Hölder⁵¹¹ y E. H. Moore⁵¹² estudiaron de manera abstracta los automorfismos de un grupo, es decir, las transformaciones biunívocas de un grupo en sí mismo bajo las cuales si $ab = c$ entonces $a'b' = c'$.

El desarrollo posterior de la teoría de grupos abstractos ha seguido direcciones muy distintas. Una de ellas partía de los grupos de sustituciones en el artículo de Hölder de 1893 y pretendía hallar todos los grupos de un orden dado, problema que ya había mencionado Cayley en sus artículos de 1878.⁵¹³ El problema general se ha resistido a ser resuelto y, en consecuencia se han investigado órdenes particulares, tales como p^2q^2 donde p y q son primos. Un problema relacionado con éste ha sido el de la enumeración de los grupos intransitivos, primitivos e imprimitivos de diversos grados (el número de letras en un grupo de sustituciones).

Otra dirección de investigación ha sido la determinación de grupos compuestos o solubles y de los grupos simples, es decir, aquellos que no tienen subgrupos invariantes aparte de la identidad. Este

problema tiene su origen en la teoría de Galois, naturalmente. Hölder, después de introducir el concepto de grupo cociente,⁵¹⁴ estudió los grupos simples⁵¹⁵ y compuestos.⁵¹⁶ Entre sus resultados está el de que un grupo cíclico de orden primo es simple, y así le ocurre al grupo alternado de todas las permutaciones pares de n letras para $n \geq 5$. Se han encontrado muchos otros grupos finitos simples.

En cuanto a los grupos resolubles, Frobenius dedicó varios artículos al problema. Descubrió, por ejemplo,⁵¹⁷ que todos los grupos cuyo orden no sea divisible por el cuadrado de un primo son resolubles.⁵¹⁸ El problema general de investigar qué grupos son resolubles es parte del problema más general de determinar la estructura de un grupo dado.

En sus artículos de 1882 y 1883 había introducido Dyck la idea abstracta de grupo definido por generadores y relaciones entre ellos. Dado un grupo definido en términos de un número finito de generadores y relaciones, el problema de la identidad o de las palabras, formulado por Max Dehn,⁵¹⁹ es el de determinar si una «palabra» o producto de elementos cualesquiera es igual al elemento unidad. Puede darse cualquier conjunto de relaciones, porque, en el peor de los casos, el grupo trivial consistente sólo en la identidad las satisface. Decidir si un grupo dado por generadores y relaciones es trivial, no es trivial; de hecho, no hay ningún procedimiento efectivo para hacerlo. Para una relación demostró Wilhelm Magnus (1907) que el problema es resoluble,⁵²⁰ pero el problema general no lo es.⁵²¹ Otro famoso problema sin resolver de la teoría general de grupos es

el problema de Burnside. Todo grupo finito tiene las propiedades de ser finitamente generado y de que todo elemento tiene orden finito; en 1902 ⁵²² William Burnside (1852-1927) se preguntó si el recíproco sería cierto, es decir, si un grupo G es finitamente generado y si todo elemento tiene orden finito, ¿es G necesariamente finito? Este problema ha atraído mucha atención, pero solamente se han resuelto casos particulares. Otro problema afín, el problema del isomorfismo, consiste en determinar cuándo dos grupos, definido cada uno por generadores y relaciones, son isomorfos.

Uno de los giros sorprendentes de la teoría de grupos ha sido el de que, poco después de romper amarras la teoría abstracta, los matemáticos se dedicaron a buscar representaciones por medio de álgebras más concretas para obtener resultados relativos a los grupos abstractos. Cayley había señalado ya en su artículo de 1854 que todo grupo abstracto finito puede ser representado por un grupo de permutaciones. También hemos mencionado (cap. 31, sec. 6) que Jordán había introducido en 1878 la representación de grupos de sustituciones por transformaciones lineales. Estas transformaciones o sus matrices han demostrado ser la representación más eficaz de los grupos abstractos y reciben el nombre de representaciones lineales.

Una representación matricial de un grupo G es un homomorfismo de los elementos g de G en un conjunto de matrices cuadradas no singulares $A(g)$ de orden fijo y de elementos complejos. El homomorfismo implica que

$$A(g_i g_j) = A(g_i)A(g_j)$$

para todo g_i y g_j de G . Hay muchas representaciones matriciales posibles de un mismo grupo G , alterando el orden (o dimensión) de las matrices, e incluso para un orden dado puede variar la correspondencia establecida. También se pueden sumar dos representaciones; si para cada elemento g de G es A_g la matriz correspondiente en una representación de orden m y B_g la matriz correspondiente en otra de orden n , entonces

$$\begin{pmatrix} A_g & 0 \\ 0 & B_g \end{pmatrix}$$

es otra representación llamada suma de las representaciones por separado. Análogamente, si

$$E_g = \begin{pmatrix} B_g & C_g \\ 0 & D_g \end{pmatrix}$$

es otra representación, cuando B_g y D_g son matrices no singulares de órdenes m y n respectivamente, las B_g y D_g también son representaciones y de orden menor que la E_g . A E_g se le llama una representación graduada; tanto ella como cualquier representación equivalente a ella ($F_g^{-1} E_g F_g$ es una representación equivalente si F_g es no singular y del mismo orden que E_g) se llama reducible. Una representación no equivalente a otra graduada se llama irreducible.

La idea básica de una representación irreducible consiste en un conjunto de transformaciones lineales en n variables, es que es una representación homomorfa o isomorfa en la que es imposible elegir $m < n$ funciones lineales de las variables que se transformen entre sí bajo toda operación del grupo que representan. Una representación equivalente a una suma de representaciones irreducibles se llama completamente reducible.

Todo grupo finito tiene una representación especial llamada regular. Supongamos que los elementos del grupo son g_1, g_2, \dots, g_n . Sea a uno cualquiera de los g y consideremos una matriz $n \times n$; supongamos que $a g_i = g_j$. Entonces colocamos un 1 en el lugar (i, j) de la matriz, para todo g_i y el elemento fijo a , añadiendo ceros en todos los lugares restantes de la matriz. La matriz así obtenida corresponde al elemento a . Esta matriz está definida para cada g del grupo y el conjunto constituye una representación regular por la izquierda; análogamente, formando los productos ga obtendríamos una representación regular por la derecha, y reordenando los elementos g del grupo resultarían otras representaciones regulares. La idea de una representación regular fue introducida por Charles S. Peirce en 1879.⁵²³

La representación de grupos de sustituciones por transformaciones lineales de la forma

$$x'_i = \sum_j a_{ij} x_j \quad i = 1, 2, \dots, n$$

iniciada por Jordán, fue generalizada al estudio de representaciones de todos los grupos finitos por Frobenius, Burnside, Theodor Molien (1861-1941) e Issai Schur (1875-1941) a finales del siglo XIX y comienzos del XX. Frobenius⁵²⁴ introdujo el concepto de representación reducible y completamente reducible para grupos finitos, y demostró que una representación regular contiene todas las representaciones irreducibles. En otros artículos publicados entre 1897 y 1910, algunos conjuntamente con Schur, demostró muchos otros resultados, incluyendo el hecho de que hay sólo unas pocas representaciones irreducibles, con las que se componen todas las demás.

Burnside ⁵²⁵ obtuvo otro resultado importante, una condición necesaria y suficiente sobre los coeficientes de un grupo de transformaciones lineales en n variables para que el grupo sea reducible. El hecho de que cualquier grupo finito de transformaciones lineales es completamente reducible lo demostró por primera vez Heinrich Maschke (1853-1908).⁵²⁶ La teoría de representaciones de grupos finitos ha conducido a importantes resultados sobre grupos abstractos. Durante el segundo cuarto de este siglo la teoría de representación se generalizó a grupos continuos, pero aquí no entraremos en ello.

En el estudio de las representaciones de grupos es importante el concepto de carácter de un grupo, que puede hacerse remontar a la obra de Gauss, Dirichlet y Heinrich Weber (véase nota 35). Esta noción la formuló de manera abstracta Dedekind para grupos abelianos en la tercera edición de las *Vorlesungen über Zahlentheorie*

(1879) de Dirichlet. Un carácter de un grupo es una función $\chi(s)$ definida para todos los elementos s , tal que no es cero para ningún s y $\chi(ss') = \chi(s)\chi(s')$. Dos caracteres son distintos si $\chi(s) \neq \chi'(s)$ para un s al menos del grupo.

Esta definición fue generalizada para todos los grupos finitos por Frobenius. Después de enunciar una definición bastante complicada,⁵²⁷ dio otra más sencilla,⁵²⁸ que hoy es la habitual. La función carácter es la traza (o suma de los elementos de la diagonal principal) de las matrices de una representación irreducible del grupo. Frobenius mismo y otros aplicarían más tarde el mismo concepto a grupos infinitos.

Los caracteres de los grupos suministran entre otras cosas una determinación del número mínimo de variables en términos de las cuales se puede representar un grupo finito como un grupo de transformaciones lineales, y para grupos conmutativos permiten determinar todos los subgrupos.

Muchos matemáticos de finales del siglo XIX y comienzos del XX mostraron el entusiasmo usual por las corrientes de moda, pensando que toda la matemática que valía la pena recordar terminaría incluida en la teoría de grupos. Klein en particular, a pesar de que no le gustaba el formalismo de la teoría *abstracta* de grupos, se mostró favorable al concepto de grupo porque pensaba que unificaría la matemática. Poincaré se mostró igualmente entusiasta, diciendo⁵²⁹ «... la teoría de grupos es, por así decirlo, la totalidad de la matemática desprovista de su materia y reducida a pura forma».

3. La teoría abstracta de cuerpos

El concepto del cuerpo R generado por n cantidades a_1, a_2, \dots, a_n , es decir, el conjunto de todas las cantidades formadas sumando, restando, multiplicando y dividiendo repetidamente estas cantidades (exceptuada la división por cero), así como el concepto del cuerpo extensión obtenido añadiendo un nuevo elemento A no en R , aparecen ya en la obra de Galois. Sus cuerpos eran los dominios de racionalidad de los coeficientes de una ecuación, y las extensiones se construían añadiendo una raíz. El mismo concepto tuvo también un origen muy distinto en los trabajos sobre números algebraicos de Dedekind y Kronecker (cap. 34, sec. 3) y, de hecho, el nombre «cuerpo» (*Körper*) se debe a Dedekind.

La teoría abstracta de cuerpos la inició Heinrich Weber, que ya había adoptado el punto de vista abstracto en teoría de grupos. En 1893⁵³⁰ dio una exposición abstracta de la teoría de Galois en la que introducía cuerpos (conmutativos) como extensiones de grupos. Un cuerpo, dice Weber, es una colección de elementos sometida a dos operaciones, llamadas suma y multiplicación, que satisfacen la condición de clausura, las propiedades asociativa y conmutativa, y la distributiva; además, cada elemento debe tener un inverso único para cada operación, excepto la división por cero. Weber insiste en que los conceptos de grupo y de cuerpo son los dos más importantes del álgebra. Algo más tarde Dickson⁵³¹ y Huntington⁵³² dieron sistemas de axiomas independientes para un cuerpo.

A los cuerpos que ya se conocían en el siglo XIX, los de los números

racionales, reales y complejos, los cuerpos de números algebraicos y los cuerpos de funciones racionales en una o varias variables, añadió Kurt Hensel otro tipo, los cuerpos p -ádicos, que abrían nuevos campos en teoría de números algebraicos (*Theorie der algebraischen Zahlen*, 1908). Lo primero que observó Hensel fue que cualquier entero D se puede expresar de una y sólo una manera como suma de potencias de un número primo p . Es decir

$$D = d_0 + d_1p + \dots + d_kp^k$$

donde d_i es algún entero de 0 a $p - 1$. Por ejemplo,

$$14 = 2 + 3 + 3^2$$

$$216 = 2 \times 3^3 + 2 \times 3^4.$$

Análogamente, cualquier número racional r (distinto de cero) puede escribirse en la forma

$$r = \frac{a}{b}p^n$$

donde a y b son enteros no divisibles por p y n es un entero positivo, negativo o nulo. Hensel generalizó estas observaciones e introdujo los números p -ádicos, que son expresiones de la forma

$$\sum_{i=-\rho}^{\infty} c_i p^i \quad (2)$$

donde p es un número primo y los coeficientes c son números racionales irreducibles cuyo denominador no es divisible por p . Tales expresiones no tendrán en general valor numérico ordinario; sin embargo, son entes matemáticos por definición.

Hensel define las cuatro operaciones básicas con estos números y demuestra que forman un cuerpo. Un subconjunto de los números p -ádicos puede ponerse en correspondencia biunívoca con los números racionales y, de hecho, este subconjunto es isomorfo a los números racionales en el sentido de isomorfismo entre cuerpos. Hensel define en el cuerpo de los números p -ádicos las unidades, los números enteros p -ádicos y otros conceptos análogos a los de los números racionales ordinarios.

Definiendo los polinomios cuyos coeficientes sean números p -ádicos, puede hablar Hensel de raíces p -ádicas de ecuaciones polinómicas y extender a estas raíces todos los conceptos de los cuerpos de números algebraicos. Así, habrá enteros algebraicos p -ádicos y, más en general, números algebraicos p -ádicos, y pueden formarse cuerpos de números algebraicos p -ádicos que sean extensiones de los números «racionales» p -ádicos definidos por (2). De hecho, toda la teoría ordinaria de números algebraicos puede trasladarse a los números p -ádicos. De una manera un tanto sorprendente, la teoría de números algebraicos p -ádicos conduce a resultados sobre los números algebraicos ordinarios, y también

resulta útil al estudiar las formas cuadráticas y ha conducido al concepto de cuerpo valuado.

La variedad creciente de cuerpos indujo a Ernst Steinitz (1871-1928), que estaba muy influenciado por la obra de Weber, a emprender un estudio sistemático de los cuerpos abstractos, cosa que hizo en su artículo fundamental *Algebraischen Theorie der Körper*.⁵³³ Todos los cuerpos, según Steinitz, pueden dividirse en dos tipos principales. Sea K un cuerpo y consideremos todos los subcuerpos de K (por ejemplo, los números racionales son un subcuerpo de los números reales). Los elementos comunes a todos los subcuerpos constituyen un subcuerpo llamado el subcuerpo primo P de K . Hay dos tipos posibles de cuerpos primos; el elemento unidad e está contenido en P y, por tanto, lo están

$$e, 2e, \dots, ne, \dots$$

Estos elementos, o bien son todos distintos o existe un entero p tal que $pe = 0$. En el primer caso P tiene que contener todas las fracciones ne/me y, dado que estos elementos forman un cuerpo P ha de ser isomorfo al cuerpo de los números racionales, y se dice que K tiene característica 0.

Si en cambio se tiene $pe = 0$, es fácil demostrar que el más pequeño p que cumple esta *condición tiene* que ser primo, y el cuerpo es isomorfo al de los restos de los enteros módulo p , es decir $0, 1, 2, p - 1$. Entonces se dice que K es un cuerpo de característica p y cualquier subcuerpo suyo tiene la misma característica. En este

caso $pa = pea = 0$, es decir, todas las expresiones en K pueden reducirse módulo p .

A partir del cuerpo primo P , en cualquiera de los dos casos anteriores, se puede obtener el cuerpo original K por un proceso de adjunción de elementos. El método consiste en tomar un elemento a de K que no esté en P y formar todas las funciones racionales $R(a)$ de a con coeficientes en P y después, si es necesario, tomar un b que no esté en $R(a)$ y hacer lo mismo con b , y continuar el proceso repetidamente.

Si una parte de un cuerpo arbitrario K , pueden hacerse diversos tipos de adjunciones. Una adjunción simple se obtiene añadiendo un único elemento x . El cuerpo extendido tiene que contener todas las expresiones de la forma

$$a_0 + a_1x + \dots + a_nx^n \quad (3)$$

donde los a_i son elementos de K . Si estas expresiones son todas distintas, el cuerpo extendido es el cuerpo $K(x)$ de todas las funciones racionales en x con coeficientes en K . Tal adjunción se llama una adjunción trascendente, y $K(x)$ una extensión trascendente de K . Si algunas de las expresiones (3) son iguales, se puede demostrar que existe una relación (utilizando α en lugar de x)

$$f(\alpha) = \alpha^m + b_1\alpha^{m-1} + \dots + b_m = 0$$

con las b_i en K y $f(x)$ irreducible sobre K . Entonces las expresiones

$$C_1\alpha^{m-1} + \dots + C_m$$

con los C_i en K constituyen un cuerpo $K(\alpha)$ formado por la adjunción de α a K . Este cuerpo se llama entonces una extensión algebraica simple de K . En $K(\alpha)$, $f(x)$ tiene una raíz, y recíprocamente, si tomamos un polinomio irreducible arbitrario $f(x)$ sobre K , entonces se puede construir un $K(\alpha)$ en el que $f(x)$ tenga una raíz.

Un teorema fundamental de Steinitz afirma que todo cuerpo se puede obtener a partir de su cuerpo primo haciendo primero una serie de adjunciones transcendentales (posiblemente infinitas), y a continuación una serie de adjunciones algebraicas al cuerpo transcendente. Un cuerpo K' se llama extensión algebraica de K si se obtiene por sucesivas adjunciones algebraicas simples, y si su número es finito se dice que K es de rango finito.

No todo cuerpo se puede extender por adjunciones algebraicas. Por ejemplo, en el caso de los complejos es imposible porque todo polinomio $f(x)$ es reducible sobre este cuerpo; un cuerpo con esta propiedad se llama algebraicamente cerrado. Steinitz demostró también que para todo cuerpo K existe un único cuerpo algebraicamente cerrado K' que es algebraico sobre K en el sentido de que cualquier otro cuerpo algebraicamente cerrado sobre K (que contenga a K) contiene un subcuerpo isomorfo a K' .

Steinitz estudió también el problema de determinar en qué cuerpos se verifica la teoría de Galois. Decir que la teoría de Galois se verifica en un cuerpo significa lo siguiente: un cuerpo de Galois \underline{K}

sobre un cuerpo dado K es un cuerpo algebraico en el que todo polinomio irreducible $f(x)$ en K , o bien permanece irreducible, o se descompone en un producto de factores lineales. Para todo cuerpo de Galois \underline{K} existe un conjunto de automorfismos, cada uno de los cuales transforma los elementos de \underline{K} en otros del mismo cuerpo y tales que a $\alpha \pm \beta$ y a $\alpha\beta$ corresponden $\alpha' \pm \beta'$ y $\alpha'\beta'$ mientras que todos los elementos de K quedan invariantes (se corresponden a sí mismos). Este conjunto de automorfismos forma un grupo G , llamado grupo de Galois de \underline{K} con respecto a K . El teorema principal de la teoría de Galois afirma que hay una única correspondencia entre los subgrupos de G y los subcuerpos de K tal que a cada subgrupo G' de G corresponde el subcuerpo K' de todos los elementos invariantes por G' y recíprocamente. Se dice que la teoría de Galois se verifica para aquellos cuerpos que cumplen este teorema. El resultado de Steinitz dice esencialmente que la teoría de Galois se verifica en los cuerpos de rango finito que pueden obtenerse de un cuerpo dado por una serie de adjunciones de raíces de polinomios irreducibles $f(x)$ que no tengan raíces iguales. Los cuerpos en los que todos los $f(x)$ irreducibles no tienen raíces iguales se llaman separables (Steinitz los llamó *volkommen* o completos).

La teoría de cuerpos incluye también, como indica la clasificación de Steinitz, cuerpos finitos de característica p . Un ejemplo sencillo de este tipo es el conjunto de todos los restos módulo un número primo p . El concepto de cuerpo finito se debe a Galois; en 1830 éste publicó un importante artículo, «*Sur la théorie des nombres*». ⁵³⁴ Galois pretendía resolver en él las congruencias

$$F(x) \equiv 0 \pmod{p}$$

donde p es primo y $F(x)$ un polinomio de grado n . Galois tomó un $F(x)$ irreducible (módulo p), de manera que la congruencia no tenía raíces enteras ni irracionales. Esto le obligó a considerar otras soluciones, sugeridas por los números imaginarios. Representó una de las raíces de $F(x)$ por i (que no es $\sqrt{-1}$) y consideró la expresión

$$a_0 + a_1 i + a_2 i^2 + \dots + a_{n-1} i^{n-1} \quad (4)$$

donde los a son números enteros. Cuando estos coeficientes varían sobre los restos positivos módulo p , esta expresión toma solamente p^n valores. Sea a uno de estos valores no nulos, de los que habrá $p^n - 1$. Las potencias de a también tienen la forma (4), luego no pueden ser todas distintas. Debe haber al menos una potencia $a^m = 1$, con m mínimo; entonces habrá m valores distintos

$$1, a, a^2, \dots, a^{m-1} \quad (5)$$

Si multiplicamos estas m cantidades por una expresión β de la misma forma, obtenemos un grupo de cantidades distintas de las (5) y entre sí. Multiplicando el conjunto (5) por y hallaremos otras cantidades, hasta que obtengamos todas las de la forma (4). Por tanto, m tiene que dividir a $p^n - 1$ o bien $a^{p^n-1} = 1$ y por tanto $a^p = a$. Los p^n valores de la forma (4) constituyen un cuerpo finito. Así pues,

Galois ha demostrado en esta situación concreta que el número de elementos de un cuerpo de Galois de característica p es una potencia de p .

E. H. Moore ⁵³⁵ demostró que cualquier cuerpo finito es isomorfo a un cuerpo de Galois de orden p^n , con p primo. Existe tal cuerpo para todo primo p y todo entero positivo n y su característica es p . Joseph H. M. Wedderburn (1882-1948), profesor de la universidad de Princeton,⁵³⁶ y Dickson demostraron simultáneamente que todo cuerpo finito es conmutativo (para la multiplicación). Se ha investigado mucho tratando de determinar la estructura de los grupos aditivos de los cuerpos de Galois, y de los cuerpos mismos.

4. Anillos

Aunque las estructuras de anillo e ideal eran bien conocidas y las utilizaron Dedekind y Kronecker en sus trabajos sobre números algebraicos, la teoría abstracta es en su totalidad un producto del siglo XX. El nombre ideal ya se había utilizado (cap. 34, sec. 4), pero Kronecker utilizó la palabra «orden» para anillo, término este último introducido por Hilbert.

Antes de pasar a la historia, será bueno precisar el significado moderno de los conceptos. Un anillo abstracto es una colección de elementos que forma un grupo abeliano con respecto a una operación llamada suma y sobre los que está definida una segunda operación de multiplicación. Esta segunda operación es asociativa, pero puede ser conmutativa o no; puede haber o no un elemento unidad y además se verifica la propiedad distributiva $a(b + c) = ab +$

ac y $(b + c)a = ba + ca$.

Un ideal de un anillo R es un subanillo M tal que si a pertenece a M y r es un elemento cualquiera de R entonces ar y ra pertenecen a M . Si solamente ar pertenece a M , entonces M se llama un ideal por la derecha, y si sólo pertenece a M ra , un ideal por la izquierda. Si un ideal lo es por la derecha y por la izquierda se llama bilátero. El anillo total es un ideal llamado idea unidad, y el ideal (a) generado por un elemento a está formado por todos los elementos $ra + na$ donde r pertenece a R y n es un número entero. Si R tiene unidad, entonces

$$ra + na = ra + nea = (r + ne) a = r'a,$$

donde r' es un elemento de R . A estos ideales generados por un elemento se les llama ideales principales. Cualquier ideal distinto de 0 y de R recibe el nombre de ideal propio. Análogamente al caso de un generador, si a_1, a_2, \dots, a_m son m elementos dados del anillo R con elemento unidad, entonces el conjunto de todas las sumas $r_1 a_1 + r_2 a_2 + \dots + r_m a_m$ con coeficientes r en R constituyen un ideal de R , que se representa por (a_1, a_2, \dots, a_m) . Se trata del ideal más pequeño que contiene a los elementos a_1, a_2, \dots, a_m . Un anillo conmutativo R se llama noetheriano cuando todo ideal es de esta forma.

Cualquier ideal M en un anillo R , al ser un subgrupo del grupo aditivo del anillo, divide a dicho anillo en clases de restos. Dos elementos a y b de R son congruentes con respecto a M , $a \equiv b \pmod{M}$, si $a - b$ pertenece a M . Si T es un homomorfismo del anillo R en

el anillo R' , lo que supone que $T(a + b) = Ta + Tb$, $T(ab) = Ta \times Tb$ y $T1 = 1'$, los elementos de R que se aplican en el elemento 0 de R' constituyen un ideal llamado núcleo de T , y R' es isomorfo al anillo cociente de R por el núcleo de T . Recíprocamente, dado un ideal L en R podemos formar el anillo cociente de R por L y definir un homomorfismo de R en R cociente L que tiene como núcleo L .

En la definición de anillo no se exige la existencia de inverso de cada elemento para la multiplicación. Si existe elemento unidad y también dicho inverso (excepto para 0), el anillo se llama un anillo con división o cuerpo no conmutativo (*skew field*). Ya hemos mencionado que Wedderburn demostró en 1905 que todo anillo con división finito es un cuerpo conmutativo. Hasta 1905 las únicas álgebras con división conocidas eran los cuerpos conmutativos y los cuaterniones. Entonces Dickson introdujo otras nuevas, tanto conmutativas como no conmutativas. En 1914 el mismo Dickson⁵³⁷ y Wedderburn⁵³⁸ dieron los primeros ejemplos de cuerpos no conmutativos con centros (conjunto de todos los elementos que conmutan con todos los demás) de rango n^2 .⁵³⁹

A finales del siglo XIX se introdujo una gran variedad de álgebras lineales asociativas concretas (cap. 32, sec. 6). Estas álgebras, consideradas en abstracto, son anillos, y cuando se formuló la teoría de anillos abstractos, ésta incluyó y generalizó los trabajos sobre estas álgebras concretas. Esta teoría de álgebras lineales asociativas, y en general, todo el álgebra abstracta, recibió un nuevo impulso cuando Wedderburn, en su artículo «*On Hypercomplex Numbers*»,⁵⁴⁰ generalizó resultados anteriores de Elie Cartan (1869-

1951).⁵⁴¹ Recordemos que los números hipercomplejos son de la forma

$$x = x_1 e_1 + x_2 e_2 + \dots + x_n e_n \quad (6)$$

donde las e_i son unidades formales y los x números reales o complejos. Wedderburn sustituyó los x_i por elementos de un cuerpo arbitrario F , y denominó a estas álgebras lineales asociativas generalizadas simplemente álgebras. Para estudiarlas tuvo que abandonar los métodos de sus predecesores debido a que el cuerpo arbitrario F no tenía por qué ser algebraicamente cerrado, aunque adoptó y perfeccionó la técnica de los idempotentes de Benjamin Peirce.

En el artículo de Wedderburn, pues, un álgebra consiste en todas las combinaciones lineales de la forma (6), ahora con coeficientes en un cuerpo F . El número de las e_i , llamadas unidades básicas, es finito y se denomina el orden del álgebra. La suma de dos de tales elementos viene dada por

$$\sum x_i e_i + \sum y_i e_i = \sum (x_i + y_i) e_i$$

el producto escalar de un elemento a del cuerpo F y un elemento x del álgebra está definido por

$$a \sum x_i e_i = \sum a x_i e_i$$

y el producto de dos elementos del álgebra por

$$\left(\sum x_i e_i\right) \left(\sum y_j e_j\right) = \sum_{ij} x_i y_j e_i e_j$$

complementado por una tabla que da todos los productos $e_i e_j$ como combinaciones lineales de los e_k con coeficientes en F . Se exige que el producto sea asociativo, y siempre se puede añadir un elemento unidad (o módulo) 1 tal que $x \times 1 = 1 \times x = x$ para todo x, y entonces los elementos del álgebra incluyen los del cuerpo F de los coeficientes.

Dada un álgebra A , un subconjunto B de elementos que a su vez constituyan un álgebra, se llama una subálgebra. Si x pertenece a A e y a B y yx y xy pertenecen ambos a B , entonces B se llama una subálgebra invariante. Si el álgebra A es la suma de dos subálgebras invariantes sin elementos comunes, se llama reducible o también se dice que es la suma directa de las subálgebras.

Un álgebra simple es la que no tiene ninguna subálgebra invariante. Wedderburn también utilizó y modificó el concepto de álgebra semisimple de Cartan. Para definir este concepto hizo uso del elemento nilpotente: el elemento x es nilpotente si $x^n = 0$ para algún entero n , y se le llama propiamente nilpotente si tanto xy como yx son nilpotentes para todo y del álgebra A . Se puede demostrar que el conjunto de todos los elementos propiamente nilpotentes de un álgebra A forman una subálgebra invariante. Entonces, un álgebra

semisimple A es la que no tiene ninguna subálgebra invariante nilpotente.

Wedderburn demostró que toda álgebra semisimple se puede expresar como suma directa de álgebras irreducibles, y cada una de ellas es equivalente al producto directo de un álgebra de matrices y un álgebra con división (o álgebra primitiva en la terminología de Wedderburn). Esto significa que cada elemento del álgebra irreducible puede considerarse como una matriz cuyos elementos pertenecen al álgebra con división. Dado que las álgebras semisimples se pueden reducir a una suma directa de varias álgebras simples, este teorema supone la determinación de todas las álgebras semisimples. Otro resultado utiliza el concepto de álgebra de matrices total, que es simplemente el álgebra de todas las matrices $n \times n$. Un álgebra tal que el cuerpo F de los coeficientes sea el de los números complejos y que no contenga ningún elemento propiamente nilpotente, es equivalente a una suma directa de álgebras de matrices totales. Esta muestra de resultados obtenidos por Wedderburn puede dar una idea de las investigaciones realizadas sobre álgebras lineales asociativas generalizadas.

La teoría de anillos e ideales fue edificada sobre una base axiomática y sistematizada por Emmy Noether, una de las pocas grandes mujeres matemáticas que en el mundo han sido, que comenzó a dar clases en Gotinga en 1922. Ya se conocían muchos resultados sobre anillos e ideales cuando ella inició su trabajo, pero consiguió unificarlos en una teoría abstracta a través de una formulación adecuada de los conceptos. Así, por ejemplo, Noether

reformula el teorema de la base de Hilbert (cap. 39, sec. 2) de la manera siguiente: un anillo de polinomios en un número cualquiera de indeterminadas sobre un anillo de coeficientes con unidad y una base finita, tiene él mismo una base finita. En esta reformulación, la teoría de invariantes queda integrada dentro del álgebra abstracta.

Emanuel Lasker (1868-1941)⁵⁴² había desarrollado ya una teoría de ideales para dominios de polinomios, buscando un método para decidir si un polinomio dado pertenece o no a un ideal generado por otros r polinomios. En 1921⁵⁴³ demostró Emmy Noether que esta teoría de ideales para polinomios podía deducirse del teorema de la base de Hilbert. De esta manera se establecían unos fundamentos comunes para la teoría de ideales de números enteros algebraicos y de funciones algebraicas enteras (o polinomios). Noether y otros profundizaron mucho más en la teoría abstracta de anillos e ideales, aplicándola a anillos de operadores diferenciales y otras álgebras. Sin embargo, una exposición de los resultados de tales investigaciones nos llevaría demasiado lejos en desarrollos muy especializados.

5. La teoría de álgebras no asociativas

La moderna teoría de anillos o, para ser más exactos, una generalización de la teoría de anillos, incluye también las álgebras no asociativas, en las que la operación producto no es conmutativa ni tampoco asociativa, mientras que las restantes propiedades de las álgebras lineales asociativas siguen siendo válidas. Hoy día se conocen diversos tipos de álgebras no asociativas importantes. El

más importante históricamente es el de las álgebras de Lie. En ellas se acostumbra representar el producto de dos elementos a y b por $[a, b]$. En lugar de la propiedad asociativa la operación producto satisface las dos condiciones:

$$[a, b] = -[b, a] \text{ y } [a, [b, c]] + [b, [c, a]] + [c, [a, b]] = 0$$

La segunda propiedad recibe el nombre de identidad de Jacobi. Dicho sea de paso, el producto vectorial de dos vectores cumple las dos condiciones.

Un ideal de un álgebra de Lie L es una subálgebra L_1 tal que el producto de un elemento cualquiera de L por un elemento cualquiera de L_1 pertenece a L_1 . Un álgebra de Lie simple es la que no tiene ideales no triviales, y semisimple si no tiene ideales abelianos.

Las álgebras de Lie surgieron de los esfuerzos de Lie por estudiar la estructura de sus grupos continuos de transformaciones. Para ello introdujo Lie la idea de transformación infinitesimal.⁵⁴⁴ Hablando de una manera intuitiva, una transformación infinitesimal es la que mueve los puntos una distancia infinitesimal. Lie las representó simbólicamente por

$$x'_i = x_i + \delta t X_i(x_1, x_2, \dots, x_n) \quad (7)$$

donde δt es una cantidad infinitamente pequeña, o bien

$$\delta x_i = \delta t X_i(x_1, x_2, \dots, x_n) \quad (8)$$

La δt es consecuencia de un cambio pequeño en los parámetros del grupo. Así, por ejemplo, supongamos un grupo de transformaciones dado por

$$x_1 = \phi(x, y, a) \quad \text{e} \quad y_1 = \psi(x, y, a)$$

y sea a_0 el valor del parámetro para el que ϕ y ψ definen la transformación identidad, de manera que

$$x = \phi(x, y, a_0) \quad \text{e} \quad y = \psi(x, y, a_0)$$

Si incrementamos a_0 a $a_0 + \delta a$, entonces, por el teorema de Taylor

$$x_1 = \phi(x, y, a_0) + \frac{\partial \phi}{\partial a} \delta a + \dots$$

$$y_1 = \psi(x, y, a_0) + \frac{\partial \psi}{\partial a} \delta a + \dots$$

de manera que despreciando las potencias más altas de da nos queda

$$\delta x = x_1 - x = \frac{\partial \phi}{\partial a} \delta a \quad ; \quad \delta y = y_1 - y = \frac{\partial \psi}{\partial a} \delta a$$

Para el a_0 fijo, $\partial \phi / \partial a$ y $\partial \psi / \partial a$ son funciones de x e y , de manera que

$$\frac{\partial \phi}{\partial a} = \xi(x, y) \quad , \quad \frac{\partial \psi}{\partial a} = \eta(x, y)$$

y

$$\delta x = \xi(x, y)\delta a \quad , \quad \delta y = \eta(x, y)\delta a \quad (9)$$

Si δa vale δt obtenemos las formas (7) u (8). Las ecuaciones (9) representan una transformación infinitesimal del grupo.

Si $f(x, y)$ es una función analítica de x e y , el efecto sobre ella de una transformación infinitesimal supone reemplazar $f(x, y)$ por $f(x + \xi\delta a, y + \eta\delta a)$ y aplicando el teorema de Taylor para aproximar hasta el primer orden

$$\delta f = \left(\xi \frac{\partial f}{\partial x} + \eta \frac{\partial f}{\partial y} \right) \delta a$$

El operador

$$\xi \frac{\partial}{\partial x} + \eta \frac{\partial}{\partial y}$$

es otra manera de *representar* la transformación infinitesimal (9), porque el conocimiento de uno de ellos da el otro. Estos operadores pueden sumarse y multiplicarse en el sentido usual de los operadores diferenciales.

El número de transformaciones infinitesimales independientes, o el número de operadores independientes correspondientes, es el número de parámetros del grupo original de transformaciones. Las transformaciones infinitesimales o los correspondientes operadores, que representaremos por X_1, X_2, \dots, X_n , determinan, de hecho, el grupo de Lie de transformaciones. Pero además, y esto es igualmente importante, son ellos mismos generadores de un grupo. Aunque el producto $X_i X_j$ no es un operador lineal, la expresión denominada alternante o conmutador de X_i y X_j

$$X_i X_j - X_j X_i,$$

sí es un operador lineal, y se representa por $[X_i, X_j]$. Con esta operación producto, el grupo de operadores se convierte en un álgebra de Lie.

Lie comenzó el trabajo de determinar la estructura de sus grupos (continuos) simples finitos con r parámetros, y descubrió cuatro clases principales de álgebras. Wilhelm K. J. Killing (1847-1923)⁵⁴⁵ comprobó que esas clases eran correctas para todas las álgebras simples, pero que había además cinco casos excepcionales con 14, 52, 78, 133 y 248 parámetros. El trabajo de Killing no era totalmente riguroso y Elie Cartan intentó cubrir sus huecos.

En su tesis *Sur la structure des groupes de transformations finis et continus*,⁵⁴⁶ da Cartan una clasificación completa de todas las álgebras de Lie simples sobre el cuerpo de los complejos para los parámetros y las variables. Tal como había obtenido Killing, Cartan

descubrió que se dividían en cuatro casos generales y las cinco álgebras excepcionales. Cartan construyó explícitamente estas cinco álgebras excepcionales. En 1914⁵⁴⁷ determinó Cartan todas las álgebras simples con valores reales para los parámetros y las variables. Sus resultados son aún básicos.

El uso de representaciones para estudiar álgebras de Lie ha resultado muy eficaz, tal como en el caso de los grupos abstractos. En su tesis, y en un artículo de 1913,⁵⁴⁸ encontró Cartan representaciones irreducibles de las álgebras de Lie simples. Más tarde obtendría Hermann Weyl un resultado clave⁵⁴⁹: cualquier representación de un álgebra de Lie semisimple (sobre un cuerpo algebraicamente cerrado de característica cero) es completamente reducible.

6. Panorama general del álgebra abstracta

Nuestras breves indicaciones de los resultados obtenidos en el dominio del álgebra abstracta, ciertamente no dan una imagen completa de los desarrollos ni siquiera del primer cuarto de nuestro siglo. Puede ser útil, sin embargo, mostrar el vasto campo que se ha abierto por el uso sistemático de un enfoque abstracto.

Hasta 1900 más o menos, las diversas cuestiones de tipo algebraico estudiadas, ya fueran matrices, álgebras de formas en dos, tres o n variables, hipernúmeros, congruencias o la resolución de ecuaciones polinómicas, habían estado basadas en los sistemas de números reales o complejos. El movimiento algebraico abstracto introdujo, sin embargo, los grupos, anillos, ideales, álgebras con

división y cuerpos abstractos. Aparte de investigar las propiedades de tales estructuras abstractas y relaciones como las de isomorfismo y homomorfismo, los matemáticos han descubierto la posibilidad de tomar casi cualquier cuestión algebraica y plantearse problemas sobre ella reemplazando los números reales o complejos por cualquier otra estructura abstracta. Por ejemplo, en lugar de matrices con elementos complejos, pueden estudiarse matrices con elementos de un anillo o un cuerpo arbitrario. Análogamente se pueden considerar problemas de la teoría de números y, sustituyendo los números enteros por un anillo cualquiera, reconsiderar cualquier cuestión que antes se había investigado solamente para los enteros usuales. Pueden considerarse incluso funciones y series de potencias con coeficientes en un cuerpo arbitrario.

Todas estas generalizaciones se han hecho realmente. Ya hemos visto que Wedderburn, en su trabajo de 1907, generalizó resultados anteriores sobre álgebras lineales asociativas (hipernúmeros) reemplazando los coeficientes reales o complejos por otros de un cuerpo cualquiera. Pues bien, es posible sustituir el cuerpo por un anillo e investigar los teoremas que siguen verificándose pese a este cambio. Se ha estudiado incluso la teoría de ecuaciones con coeficientes en un cuerpo arbitrario, incluso finito.

Como otro ejemplo de la moderna tendencia a generalizar, piénsese en las formas cuadráticas. Las formas cuadráticas con coeficientes enteros fueron importantes en el estudio de la representación de enteros como sumas de cuadrados, y con coeficientes reales en el

estudio de las cónicas y las cuádricas. En el siglo XX se han estudiado las formas cuadráticas con coeficientes en cualquier cuerpo. Según se van introduciendo más estructuras abstractas, pueden ser utilizadas como base o dominio de coeficientes de otras teorías algebraicas más antiguas, y el proceso de generalización avanza indefinidamente. Este uso de conceptos abstractos exige a su vez el uso de técnicas algebraicas también abstractas; así, muchos campos anteriormente no relacionados entre sí se han visto absorbidos por el álgebra abstracta. Este ha sido el caso, en particular, con extensas partes de la teoría de números, incluida también la teoría de números algebraicos.

Sin embargo, el álgebra abstracta ha terminado por subvertir su propio papel dentro de la matemática. Sus conceptos se formularon para unificar dominios matemáticos aparentemente diversos y completamente separados, tal como hizo, por ejemplo, la teoría de grupos. Una vez formuladas las teorías abstractas, los matemáticos olvidaron los campos concretos originales y concentraron su atención únicamente en las estructuras abstractas. Con la introducción de cientos de conceptos subordinados, la materia se ha desarrollado como los hongos en un desorden de desarrollos menores que tienen poca relación unos con otros y con los campos concretos originales. La unificación ha cedido su lugar a la diversificación y a la especialización. En realidad, la mayor parte de los investigadores actuales en álgebra abstracta ignoran incluso los orígenes de las estructuras abstractas, y tampoco están interesados en las aplicaciones de sus resultados a campos concretos.

Bibliografía

- Artin, Emil: «The Influence of J. H. M. Wedderburn on the Development of Modern Algebra». *Amer. Math. Soc. Bu.ll.*, 56, 1950, 65-72.
- Bell, Eric T.: «Fifty Years of Algebra in America, 1888-1938». *Amer. Math. Soc. Semicentennial Publications*, II, 1938, 1-34.
- Bourbaki, N.: *Elementos de historia de las matemáticas*, Madrid, Alianza, 1976.
- Cartan, Elie: «Notice sur les travaux scientifiques». *CEuvres completes*, Gaut- hier-Villars, 1952-1955, part I, vol. I, 1-98.
- Dicke, Auguste: *Emmy Noether, 1882-1935*, Birkhäuser Verlag, 1970. Dickson, L. E.: «An Elementary Exposition of Frobenius's Theory of Group-Characters and Group-Determinants». *Annals of Math.*, 4, 1902, 25-49; *Linear Algebras*, Cambridge University Press, 1914.; *Algebras and Their Arithmetics* (1923), G. E. Stechert (reprint), 1938. Frobenius, F. G.: *Gesammelte Abhandlungen*, 3 vols., Springer-Verlag, 1968.
- Hawkins, Thomas: «The Origins of the Theory of Group Characters». *Archive for History of Exact Sciences*, 7, 1971, 142-170.
- MacLane, Saunders: «Some Recent Advances in Algebra». *Amer. Math. Monthly*, 46, 1939, 3-19. Also in Albert, A. A., ed.: *Studies in Modern Algebra*, The Math. Assn. of Amer., 1963, 9-34.; «Some Additional Advances in Algebra» in Albert, A. A.,

- ed.: *Studies in Modern Algebra*, The Math. Assn. of Amer., 1963, 35-58.
- Ore, Oystein: «Some Recent Developments in Abstract Algebra». *Amer. Math. Soc. Bull.*, 37, 1931, 537-548.; «Abstract Ideal Theory». *Amer. Math. Soc. Bull.*, 39, 1933, 728-745. Steinitz, Ernst: *Algebraische Theorie der Körper*, W. de Gruyter, 1910; 2.^a ed., 1930; Chelsea (reprint), 1950. La primera edición es la misma que el artículo en *Jour. für Math.*, 137, 1910, 167-309.
 - Wiman, A.: «Endliche Gruppen linearer Substitutionen». *Encyk. der Math.*
 - *Wiss.*, B. G. Teubner, 1898-1904, I, part 1, 522-554.
 - Wussing, H. L.: *Die Génesis des abstrakten Gruppenbegriffes*, VEB Deutscher Verlag der Wissenschaften, 1969.

Capítulo 50

Los orígenes de la topología

Creo que nos falta otro tipo de análisis propiamente geométrico o lineal que exprese directamente la localización, así como el álgebra expresa la magnitud.

G. W. Leibniz

Contenido:

- 1. ¿Qué es la topología?*
- 2. La topología conjuntista*
- 3. Los comienzos de la topología combinatoria*
- 4. La obra combinatoria de Poincaré*
- 5. Los invariantes combinatorios*
- 6. Los teoremas de punto fijo*
- 7. Generalizaciones y extensiones*

Bibliografía

1. ¿Qué es la topología?

Cierto número de desarrollos durante el siglo XIX cristalizaron en una nueva rama de la geometría, que hoy llamamos topología pero conocida durante largo tiempo con el nombre de «análisis situs». Para decirlo de una manera vaga, por el momento, la topología se ocupa de aquellas propiedades de las figuras que permanecen

invariantes cuando dichas figuras son plegadas, dilatadas, contraídas o deformadas de cualquier manera tal que no aparezcan nuevos puntos o se hagan coincidir puntos ya existentes. La transformación permitida presupone, en otras palabras, que hay una correspondencia biunívoca entre los puntos de la figura original y los de la figura transformada, y que la transformación hace corresponder puntos próximos a puntos próximos. Esta última propiedad se llama continuidad, y lo que se requiere es que la transformación y su inversa sean ambas continuas. Una transformación así se llama un homeomorfismo. La topología se describe a veces de una manera imprecisa como la geometría de las bandas de goma, porque si las figuras fueran de goma sería posible deformar muchas de ellas en otras figuras homeomorfas. Así una banda de goma plana puede deformarse para convertirse en, y es por tanto topológicamente lo mismo que, un círculo o un cuadrado, pero no es topológicamente lo mismo que una figura en forma de ocho, porque esto requeriría la superposición de dos puntos de la banda.

Se suele imaginar las figuras como sumergidas en un espacio ambiente. A efectos topológicos dos figuras pueden ser homeomorfas incluso aunque no sea posible transformar topológicamente todo el espacio en el que está sumergida la primera figura en el espacio que contiene a la segunda. Por ejemplo, si se toma un rectángulo de papel alargado y se unen los dos lados cortos, se obtiene una superficie cilíndrica. Si en lugar de esto se gira uno de los lados cortos 360° y después se unen los dos lados

cortos, la nueva figura obtenida es topológicamente equivalente a la anterior; sin embargo, es imposible transformar topológicamente el espacio tridimensional en sí mismo de tal manera que la primera figura se transforme en la segunda.

La topología, tal como la concebimos hoy, se divide en dos ramas separadas, en cierto sentido: la topología conjuntista (o topología general), que se ocupa de las figuras geométricas consideradas como conjuntos de puntos, donde el conjunto total es frecuentemente considerado como un espacio; y la topología combinatoria o algebraica, que trata las figuras como agregados de bloques más pequeños, exactamente lo mismo que una pared es una colección de ladrillos. Desde luego, los conceptos de la topología de conjuntos de puntos se usan también en la topología combinatoria, especialmente para estructuras geométricas muy generales.

La topología ha tenido numerosos y variados orígenes. Como en la mayor parte de las ramas de la matemática, se recorrieron muchas etapas que sólo posteriormente fueron consideradas como perteneciendo a (o capaces de ser englobadas en) una nueva teoría. En el caso que nos ocupa, la posibilidad de un estudio significativo por separado, fue al menos esbozada por Klein en su *Erlanger Programm* (cap. 38, sec. 5). Klein estaba generalizando los tipos de transformaciones estudiadas en geometría proyectiva y geometría algebraica, y ya era consciente de la importancia de los homeomorfismos por la obra de Riemann.

2. La topología conjuntista

La teoría de conjuntos de puntos tal como fue iniciada por Cantor (cap. 41, sec. 7), y continuada por Jordán, Borel y Lebesgue (cap. 44, secs. 3 y 4) no se refiere *eo ipso* a transformaciones y propiedades topológicas. Pero por otra parte, cualquier conjunto de puntos considerado como un espacio es ya de interés para la topología. Lo que distingue a un espacio de un simple conjunto de puntos es algún concepto que ligue unos puntos con otros. Así, en un espacio euclídeo, la distancia entre puntos nos dice cómo están de próximos unos a otros, y en particular nos permite definir qué es un punto límite de un conjunto.

Los orígenes de la topología conjuntista ya han sido mencionados (cap. 46, sec. 2). Fréchet, el año 1906, impulsado por el deseo de unificar la teoría de conjuntos de puntos de Cantor y el tratamiento de las funciones como puntos de un espacio, que ya se había hecho habitual en el cálculo de variaciones, inició el estudio de los espacios abstractos. La aparición del análisis funcional con la introducción de los espacios de Hilbert y de Banach dio una importancia adicional al estudio de los conjuntos de puntos como espacios. Las propiedades que se mostraron relevantes para el análisis funcional son topológicas, en gran medida por la importancia de los límites de sucesiones. Por otra parte, los operadores del análisis funcional son simplemente transformaciones de un espacio en otro.⁵⁵⁰

Como hizo notar Fréchet, la propiedad que liga los puntos unos con otros no necesita ser la distancia euclídea. El introdujo (cap. 46,

sec. 2) varios conceptos diferentes que pueden utilizarse para indicar cuándo un punto es límite de una sucesión de puntos. En particular, generalizó la idea de distancia introduciendo la clase de los espacios métricos. En un espacio métrico, por ejemplo, el espacio euclídeo bidimensional, uno habla de un entorno de un punto entendiendo por ello el conjunto de todos los puntos cuya distancia al dado es menor que un cierto número ϵ . Tales entornos son los entornos circulares; de manera análoga se podrían usar entornos cuadrados. Sin embargo, es posible también suponer que los entornos, ciertos subconjuntos de un conjunto de puntos dado, vengan definidos de alguna otra manera, incluso sin la introducción previa de una métrica. Se dice entonces que tales espacios tienen una topología de entornos; este concepto es una generalización de los espacios métricos. Félix Hausdorff (1868-1942), en su «*Grundzüge der Mengenlehre*» (Fundamentos de Teoría de Conjuntos, 1914) utilizó el concepto de entorno (que ya había usado Hilbert en 1902 en una formulación axiomática especial de la geometría euclídea plana) y edificó una teoría definitiva de los espacios abstractos basada en este concepto.

Hausdorff define un espacio topológico como un conjunto de elementos x , junto con una familia de subconjuntos U_x asociada a cada punto x . Estos subconjuntos se llaman entornos y deben satisfacer las condiciones siguientes:

- a. Para cada punto x existe el menos un entorno U_x que lo contiene.
- b. La intersección de dos entornos cualesquiera de x contiene

- otro entorno de x .
- c. Si y es un punto que pertenece a U_x , entonces existe un U_y tal que $U_y \subseteq U_x$.
 - d. Si $x \neq y$, entonces existen U_y y U_x tales que $U_x \cap U_y = \emptyset$.

Hausdorff introdujo también los axiomas de numerabilidad:

- a. Para todo punto x el conjunto de los U_x es como máximo numerable.
- b. El conjunto de todos los entornos distintos es numerable.

La tarea básica en topología conjuntista consiste en definir diversas nociones fundamentales. Por ejemplo, un punto límite de un conjunto de puntos en un espacio de entornos es un punto tal que todo entorno suyo contiene otros puntos del conjunto; un conjunto es abierto si todo punto perteneciente a él puede ser encerrado en un entorno que contiene sólo puntos del conjunto; si un conjunto contiene a todos sus puntos límites entonces se llama cerrado; un espacio o un subconjunto se llama compacto si todo subconjunto infinito suyo tiene un punto límite. Así, por ejemplo, la recta euclídea usual no es un conjunto compacto porque el subconjunto infinito formado por los puntos de abscisa un número entero, no tiene ningún punto límite. Un conjunto es conexo si, de cualquier manera que se le divida en dos subconjuntos disjuntos, uno de ellos al menos contiene puntos límites del otro; la curva $y = \operatorname{tg} x$ no es conexa, pero la curva $y = \operatorname{sen} 1/x$ más el intervalo $(-1, 1)$ del eje Y es conexa. La separabilidad, introducida por Fréchet en su tesis de

1906, es otro concepto básico; un espacio se denomina separable si tiene un subconjunto numerable cuya adherencia (el conjunto más sus puntos límites) coincide con el espacio total.

Ahora ya podemos introducir los conceptos de transformación continua y homeomorfismo. Una transformación continua es una correspondencia que a cada punto de un espacio asocia un único punto de un segundo espacio o espacio imagen, y tal que dado un entorno arbitrario de un punto imagen existe un entorno del punto original (o de cada punto original si hay más de uno) cuyos puntos se transforman en puntos del entorno dado en el espacio imagen. Este concepto no es más que una generalización de la definición $\varepsilon - \delta$ de una función continua, donde el ε determina el entorno del punto en el espacio imagen y el δ un entorno del punto original. Un homeomorfismo entre dos espacios S y T es una correspondencia biunívoca que es continua en los dos sentidos, es decir, la transformación de S a T y la de T a S son continuas. La tarea básica de la topología general es entonces la de descubrir propiedades que sean invariantes bajo transformaciones continuas y homeomorfismos. Todas las propiedades mencionadas más arriba son invariantes topológicos.

Hausdorff contribuyó con muchos resultados a la teoría de espacios métricos. En particular desarrolló el concepto de completitud que había introducido Fréchet en su tesis de 1906. Un espacio se llama completo si toda sucesión $\{a_n\}$ que cumpla la condición de que dado un ε positivo exista un N tal que $|a_n - a_m| < \varepsilon$ para todos m y n mayores que N , tiene límite. Hausdorff demostró que todo espacio

métrico puede extenderse a un espacio métrico completo de una y sólo una manera.

La introducción de los espacios abstractos planteó diversos problemas que provocaron una intensa investigación. Por ejemplo, si un espacio está definido por medio de entornos ¿es necesariamente metrizable?; es decir, ¿es posible introducir una métrica que preserve la estructura del espacio de tal manera que los puntos límites permanezcan como puntos límites? Este problema fue planteado por Fréchet. Un resultado debido a Paul S. Urysohn (1898-1924), afirma que todo espacio normal es metrizable;⁵⁵¹ un espacio normal es aquel en que dos conjuntos cerrados disjuntos cualesquiera pueden ser separados por dos abiertos disjuntos. Otro resultado importante relacionado con éste y debido también a Urysohn es el siguiente:⁵⁵² todo espacio métrico separable, es decir, todo espacio métrico que contenga un subconjunto denso numerable en el espacio, es homeomorfo a un subconjunto del cubo de Hilbert; el cubo de Hilbert consiste en el espacio de todas las sucesiones (x) de números reales tales que $0 \leq x \leq 1/i$ y donde la distancia se define como

$$d = \sqrt{\sum_{i=1}^{\infty} (x_i - y_i)^2}$$

El problema de la dimensión, ya mencionado, apareció por la demostración de Cantor de la existencia de una correspondencia

biunívoca entre la recta y el plano (cap. 41, sec. 7) y por la construcción de la curva de Peano que llena un cuadrado (cap. 42, sec. 5). Fréchet (que trabajaba ya con espacios abstractos) y Poincaré vieron la necesidad de una definición de dimensión que se aplicase también a los espacios abstractos y que, sin embargo, atribuyese a la recta y al plano las dimensiones usuales. La definición que había sido aceptada tácitamente hasta entonces era la del número de coordenadas necesarias para determinar los puntos del espacio; esta definición no era aplicable a los espacios generales.

En 1912 ⁵⁵³ Poincaré dio una definición recursiva. Un continuo (o conjunto cerrado conexo) tiene dimensión n si puede ser separado en dos partes cuya frontera común consiste en continuos de dimensión $n - 1$. Luitzen E. J. Brouwer (1881-1967) hizo notar que esta definición no era aplicable al cono de dos hojas, porque las hojas se separan por un punto. La definición de Poincaré fue mejorada por Brouwer,⁵⁵⁴ Urysohn ⁵⁵⁵ y Karl Menger (n. 1902).⁵⁵⁶

Las definiciones de Menger y Urysohn son análogas y a ambos se les atribuye la definición generalmente aceptada hoy. Su definición asigna una dimensión local. La formulación de Menger es la siguiente: el conjunto vacío tendría por definición dimensión -1 . Un conjunto M se llama n -dimensional en un punto P si n es el mínimo número para el cual existen entornos de P arbitrariamente pequeños cuyas fronteras en M tengan dimensión menor que n . El conjunto M se llama n -dimensional si su dimensión es menor o igual que n en cada uno de sus puntos, pero igual a n en un punto

al menos.

Otra definición ampliamente aceptada es la de Lebesgue;⁵⁵⁷ un espacio es n -dimensional si n es el número mínimo para el cual los recubrimientos cualesquiera mediante conjuntos cerrados de diámetro arbitrariamente pequeño contienen algún punto común a $n + 1$ conjuntos del recubrimiento. Los espacios euclídeos tienen la dimensión correcta con cualquiera de estas definiciones, y la dimensión de un espacio arbitrario es un invariante topológico.

Un resultado clave en teoría de la dimensión es el teorema debido a Menger (*«Dimensionstheorie»*, 1928, p. 295) y a A. Georg Nöbeling (1907),⁵⁵⁸ que afirma que todo espacio métrico compacto n -dimensional es homeomorfo a un subconjunto del espacio euclídeo $(2n + 1)$ -dimensional.

Otro problema planteado por los trabajos de Jordán y Peano era la definición misma de curva (cap. 42, sec. 5). La respuesta fue posible basándose en resultados de teoría de la dimensión. Menger⁵⁵⁹ y Urysohn⁵⁶⁰ definieron una curva como un continuo unidimensional, entendiendo por continuo un conjunto de puntos cerrado y conexo. (Esta definición requiere que una curva abierta, como una parábola, se cierre mediante un punto del infinito.) Esta definición excluye las curvas que llenan un espacio y hace de la propiedad de ser una curva un invariante bajo homeomorfismos.

El campo de la topología general o conjuntista ha seguido siendo extraordinariamente activo. Es relativamente fácil aquí el introducir variaciones, particularizar y generalizar las caracterizaciones axiomáticas de los diversos tipos de espacios. Se han introducido

cientos de definiciones y de teoremas, aunque en la mayor parte de los casos el valor último de estos conceptos es dudoso. Como en otros campos, los matemáticos no han dudado en zambullirse libremente y sin complejos en la topología general.

3. Los comienzos de la topología combinatoria

Tan pronto como en 1679 intentó ya Leibniz, en su «*Característica Geométrica*», formular las propiedades básicas de las figuras geométricas, utilizar símbolos especiales para representarlas sintéticamente y combinar estas propiedades mediante ciertas operaciones para producir otras. Leibniz llamó a este estudio análisis situs o *geometría situs*. En una carta a Huygens de 1679⁵⁶¹ explica que no le satisface el tratamiento de las figuras geométricas mediante coordenadas porque, aparte del hecho de que este método no era directo ni bello, se refería a magnitudes, mientras que «Creo que nos falta otro tipo de análisis propiamente geométrico o lineal que exprese directamente la localización (*situs*), así como el álgebra expresa la magnitud». Los pocos ejemplos que da Leibniz acerca de lo que se proponía construir implicaban aún propiedades métricas, aunque él aspiraba a conseguir unos algoritmos geométricos que suministrasen la solución de los problemas puramente geométricos. Debido probablemente a que Leibniz era más bien impreciso acerca del tipo de geometría que buscaba, Huygens no se mostró muy entusiasmado con sus ideas y su simbolismo. En la medida en que las cosas estaban medianamente claras, lo que Leibniz preveía era lo que llamamos hoy topología combinatoria.

Una propiedad combinatoria de ciertas figuras geométricas es la siguiente, atribuida a Euler aunque era conocida ya por Descartes en 1639 y, a través de los manuscritos no publicados de este último, por Leibniz en 1675: si se cuentan los números de vértices, caras y aristas de un poliedro convexo cualquiera, por ejemplo un cubo, entonces se tiene que $V - A + C = 2$. Este hecho fue publicado por Euler en 1750⁵⁶² y en 1751 presentó una demostración.⁵⁶³ Euler estaba interesado en esta relación para utilizarla al clasificar los poliedros; aunque había descubierto una propiedad de todos los poliedros cerrados convexos, no pensó en su posible invariancia bajo transformaciones continuas, ni tampoco definió la clase de los poliedros para los que dicha relación se verifica.

En 1811 Cauchy⁵⁶⁴ dio otra demostración suprimiendo el interior de una cara y desarrollando la figura restante sobre un plano; de esta forma se obtiene un polígono para el cual $V - A + C$ debe valer 1. Cauchy demostraba esto último triangulando la figura y contando después los cambios cuando los triángulos se van suprimiendo uno por uno. Esta demostración, a pesar de ser incorrecta porque supone que cualquier poliedro convexo cerrado es homeomorfo a una esfera, fue aceptada sin reparos por los matemáticos del siglo XIX.

Otro problema muy conocido, que fue en su tiempo una curiosidad, pero cuya naturaleza topológica fue reconocida más tarde, es el problema de los puentes de Königsberg.

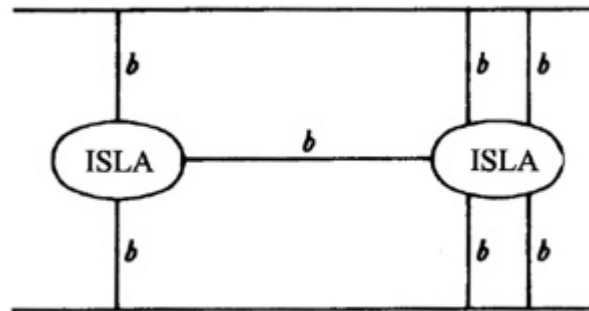


Figura 50.1

Sobre el Pregel, río que atraviesa Königsberg, hay construidos siete puentes que unen las dos islas con las riberas, puentes señalados con la letra b en la figura 50.1. Los habitantes de la ciudad se divertían intentando cruzar los siete puentes en un paseo continuo sin cruzar dos veces ninguno de ellos. Euler, que vivía entonces en San Petersburgo, oyó hablar del problema y lo resolvió en 1735⁵⁶⁵ simplificando su planteamiento al reemplazar las zonas de tierra por puntos y los puentes por segmentos o arcos, tal como aparece en la figura 50.2; la cuestión que Euler se planteó entonces fue la de si era posible describir esta figura en un movimiento continuo del lápiz sin recorrer dos veces ninguno de los trazos.

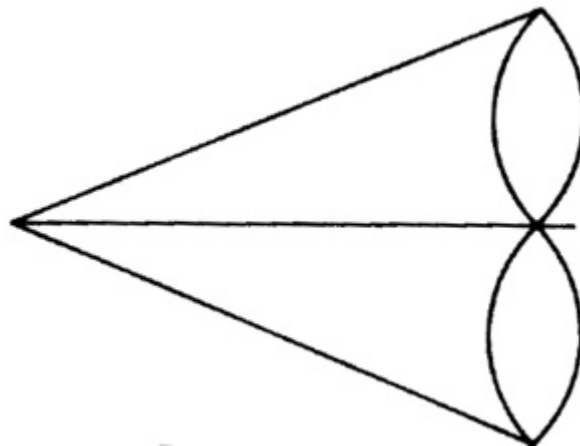


Figura 50.2

Demostó que no era posible en este caso, y dio un criterio acerca de cuándo tales caminos son o no posibles para conjuntos dados de puntos y arcos.

Gauss insistió a menudo ⁵⁶⁶ en la necesidad de estudiar las propiedades geométricas básicas de las figuras, pero no hizo ninguna contribución sobresaliente sobre el tema. En 1848, Johann B. Listing (1806-1882), alumno de Gauss en 1834 y más tarde profesor de física en Göttingen, publicó su *Vorstudien zur Topologie*, donde discutía lo que él hubiera preferido llamar geometría de posición, pero que denominó topología, ya que el primer nombre había sido utilizado el año anterior por Von Staudt para denominar la geometría proyectiva. En 1858 comenzó una nueva serie de investigaciones topológicas que aparecieron publicadas bajo el título de *Der Census raamlicher Complexe* (Panorama de los Complejos Espaciales).⁵⁶⁷ Listing buscaba leyes cualitativas para las figuras geométricas; por ejemplo, trató de generalizar la relación $V - A + C = 2$ de Euler.

El primero que formuló correctamente el carácter de las investigaciones topológicas fue Möbius, que había sido ayudante de Gauss en 1813. Möbius había clasificado ya diversos tipos de propiedades geométricas, proyectivas, afines, semejanzas y congruencias, y en 1863, en su *Theorie der elementaren Verwandtschaft* (Teoría de las relaciones elementales)⁵⁶⁸ se propuso estudiar la relación existente entre dos figuras cuyos puntos están

en una correspondencia biunívoca tal que puntos próximos corresponden a puntos próximos. Comenzó estudiando la *geometría sítas* de los poliedros, subrayando que un poliedro puede ser considerado como una colección de polígonos bidimensionales, los cuales, triangulados, convertirían al poliedro en una colección de triángulos. Esta idea resultó ser fundamental. También demostró⁵⁶⁹ que algunas superficies podían cortarse y desarrollarse como polígonos, identificando adecuadamente ciertos lados. Así, un doble anillo podía ser representado como un polígono (fig. 50.3) identificando los lados que están señalados con las mismas letras.

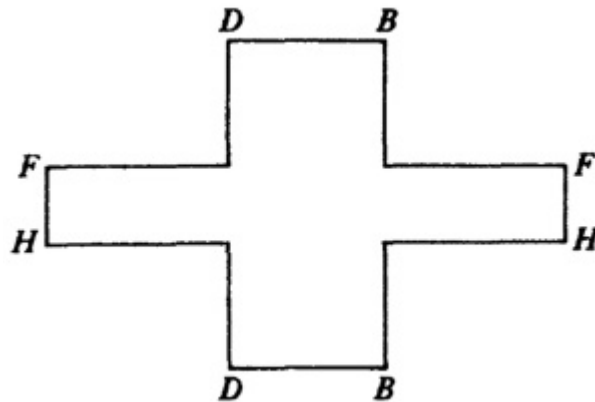


Figura 50.3

En 1858 Möbius y Listing descubrieron independientemente las superficies de una sola cara, de las cuales la banda de Möbius es la más conocida (fig. 50.4); esta figura se obtiene cogiendo una banda rectangular de papel, girando uno de sus extremos cortos 180° y uniéndolo al extremo opuesto.

Listing publicó su descubrimiento en *Der Census*, y la banda en cuestión también aparece descrita en una publicación de Möbius.⁵⁷⁰

Por lo que se refiere a esta banda, el tener una sola cara puede ser caracterizado por el hecho de que se la puede pintar totalmente con un movimiento continuo de la brocha. Si se pretende, en cambio, pintar una banda que no ha sido previamente girada, la brocha debe saltar sobre uno de sus bordes para pasar de una cara a la otra.

El hecho de tener una sola cara puede caracterizarse también utilizando un vector perpendicular a la superficie; si al desplazar este vector arbitrariamente sobre ella, manteniéndolo siempre perpendicular a la superficie, conserva el mismo sentido cuando vuelve a su posición original, entonces se dice que dicha superficie tiene dos caras, y si el sentido se invierte, que tiene una sola cara.

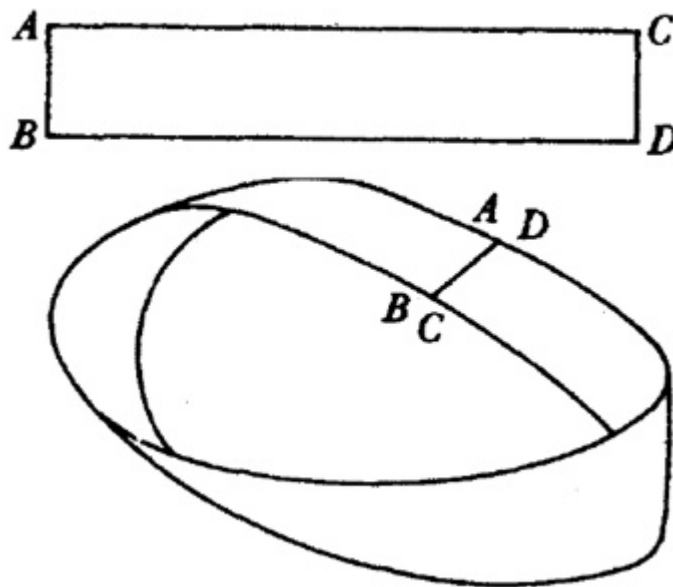


Figura 50.4

Sobre la banda de Möbius el vector perpendicular regresa al punto de partida pero sobre el «lado opuesto», es decir, con su sentido

invertido.

Otro problema que más tarde mostró su carácter topológico es el llamado problema del mapa. Consiste en demostrar que cuatro colores son suficientes para colorear cualquier mapa, de manera que países que tengan por lo menos un arco de frontera común estén coloreados con colores distintos. Francis Guthrie (m. 1899), profesor de matemáticas poco conocido, formuló por primera vez en 1852 la conjetura de que cuatro colores siempre serían suficientes, y por aquella época su hermano Frederick se lo comunicó a De Morgan. El primer artículo dedicado a este problema fue el de Cayley de 1879;⁵⁷¹ en él decía que no había podido conseguir una demostración de dicha conjetura. Muchos matemáticos han seguido buscando desde entonces una demostración, y aunque algunas de ellas fueron aceptadas durante cierto tiempo, se vio posteriormente que todas eran incorrectas y el problema aún sigue abierto[†].

El mayor impulso a las investigaciones topológicas vino de los trabajos de Riemann en teoría de funciones de variable compleja. En su tesis de 1851 sobre funciones de variable compleja y en sus estudios sobre funciones abelianas,⁵⁷² insiste en que para trabajar con funciones eran indispensables algunos teoremas del «análisis situs». En estas investigaciones se encontró con la necesidad de introducir la idea de conexión sobre las superficies de Riemann, definiéndola de la manera siguiente: «Si sobre la superficie F (con frontera) pueden dibujarse n curvas cerradas a_1, a_2, \dots, a_n que ni

[†] El problema de los cuatro colores ha sido resuelto por Appel y Haken utilizando computadores. Véase K. Appel y W. Haken, Every planar map is four colorable, I: Discharging, Illinois J. Math., 21 (1977), 429-490, y K. Appel, W. Haken y J. Koch, Every planar map is four colorable, II: Reducibility, *ibid.*, 491-567. (Nota del traductor.)

individualmente, ni combinadas, limitan completamente una parte de dicha superficie F , pero tales que con ayuda de ellas cualquier otra curva cerrada forma la frontera completa de una parte de F , entonces la superficie se llama $(n + 1)$ -conexa». Para reducir la conexión de una superficie (con frontera) establece Riemann:

Por medio de una sección [Querschnitt], es decir, de una línea sobre la superficie que vaya de un punto frontera a un punto frontera, una superficie $(n + 1)$ -conexa puede transformarse en otra n -conexa F' . Las partes de la frontera que aparecen en esta división hacen el papel de frontera incluso durante las divisiones posteriores, de manera que una sección no puede pasar más de una vez por un mismo punto pero puede terminar en uno de sus puntos previos.

... Para aplicar estas consideraciones a una superficie sin frontera, cerrada, tenemos primero que transformarla en una con frontera destacando un punto arbitrario y formando la primera división mediante este punto y una sección que comienza y termina en él, es decir, por una curva cerrada.

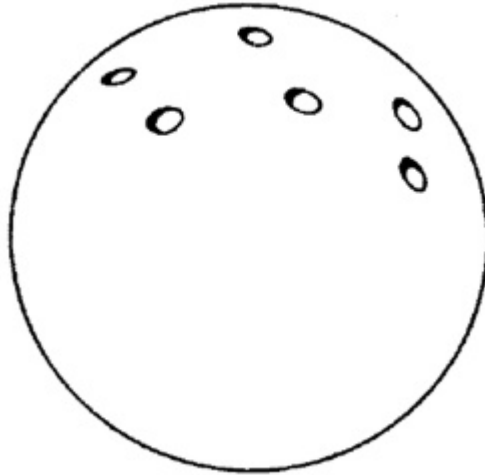


Figura 50.5

Riemann pone el ejemplo de un toro (fig. 50.5) que es una superficie 3-conexa (género 1 o número de Betti unidimensional 2) y que puede transformarse en una superficie simplemente conexa por medio de una curva cerrada abe y una sección $ab'c'$.

Riemann había clasificado de esta forma las superficies según su tipo de conexión y, como él mismo constató, había introducido así una propiedad topológica. En términos del género, nombre usado por los geómetras algebraicos de finales del siglo XIX, Riemann había clasificado las superficies cerradas por medio de su género p , siendo $2p$ el número de curvas cerradas (secciones-lazo o «*Rückkehrschnitte*») necesarias para hacer la superficie simplemente conexa, y $2p + 1$ para cortar la superficie en dos partes distintas. Riemann consideraba como intuitivamente evidente que si dos superficies de Riemann cerradas (orientables) son topológicamente equivalentes, entonces tienen el mismo género. Observó también que todas las superficies algebraicas cerradas de género cero, es

decir, simplemente conexas, son topológicamente (y conformemente y birracionalmente) equivalentes, pudiéndose aplicar topológicamente sobre una esfera.

Debido a que la estructura de las superficies de Riemann es complicada y a que figuras topológicamente equivalentes tienen el mismo género, algunos matemáticos se dedicaron a buscar estructuras más sencillas.

William K. Clifford demostró⁵⁷³ que la superficie de Riemann de una función n -valuada con w puntos de ramificación, puede transformarse en una esfera con p agujeros, siendo $p = (w/2) - n + 1$ (fig. 50.6).

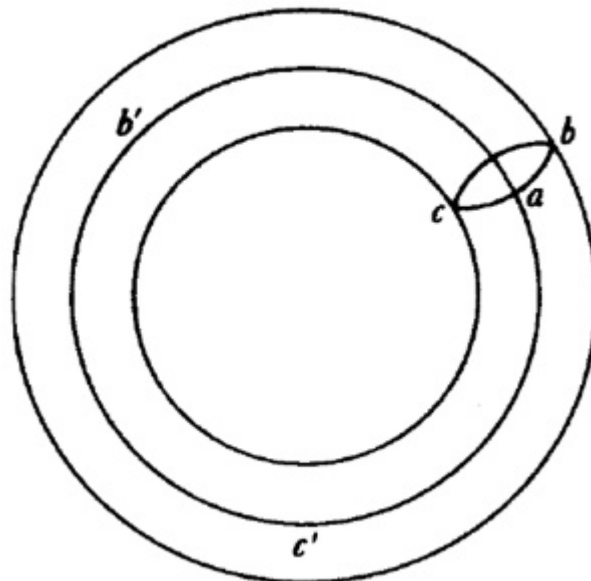


Figura 50.6

Es probable que Riemann ya conociera y usara este modelo. Klein sugirió otro modelo topológico, la esfera con p asas (fig. 50.7).⁵⁷⁴

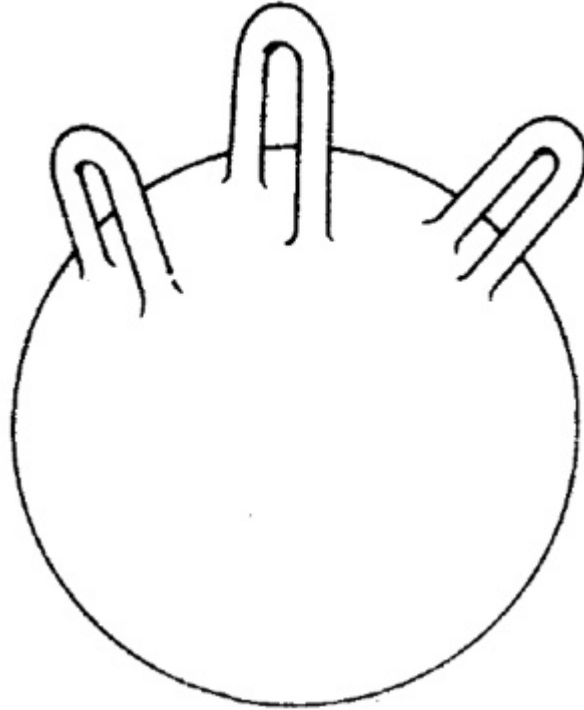


Figura 50.7

El estudio de la equivalencia topológica de las superficies cerradas se debe a un gran número de matemáticos. Para exponer el resultado principal es necesario dar la definición de superficie orientable: una superficie orientable es aquella que se puede triangular de manera que cada triángulo (curvilíneo) puede ser orientado de tal forma que cada lado común a dos triángulos tiene orientaciones opuestas inducidas sobre él por las de los dos triángulos. Así, por ejemplo, la esfera es orientable pero el plano proyectivo (ver más adelante) no lo es, lo cual fue descubierto por Klein.⁵⁷⁵ El resultado fundamental, clarificado por Klein en su trabajo, es que dos superficies cerradas orientables son homeomorfas si y sólo si tienen el mismo género. Para las superficies orientables con frontera, debe añadirse a la condición

anterior el que el número de curvas frontera sea el mismo en ambas; este teorema había sido demostrado ya por Jordán.⁵⁷⁶

La complejidad que pueden presentar las figuras cerradas, incluso las bidimensionales, fue puesta de relieve por Klein al introducir en 1882 la superficie llamada posteriormente botella de Klein (fig. 50.8) (véase la sec. 23 de la referencia en la nota 25).

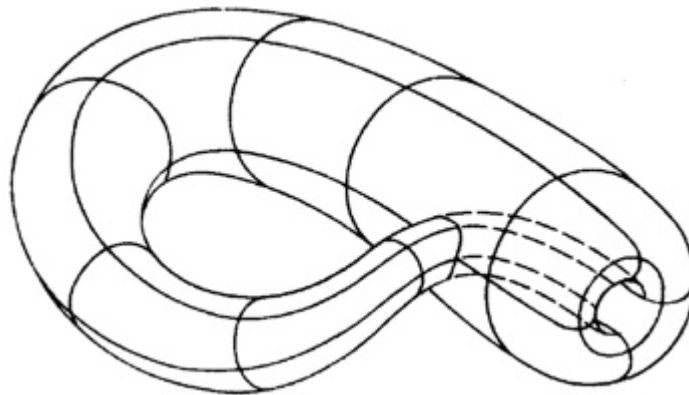


Figura 50.8

El cuello de esta botella penetra en su interior sin atravesarla y termina unido de una manera continua a la base a lo largo de C . A lo largo de D la superficie no se interrumpe y sin embargo el cuello penetra en el interior de la superficie; la superficie no tiene aristas, ni interior ni exterior; tiene una sola cara y número de conexión unidimensional 3, o bien género igual a 1; no puede ser construida en tres dimensiones.

El plano proyectivo es otro ejemplo de una superficie cerrada bastante complicada; puede representarse topológicamente por un círculo con los pares de puntos diametralmente opuestos identificados (fig. 50.9), quedando representada la recta del infinito

por la semicircunferencia CAD.

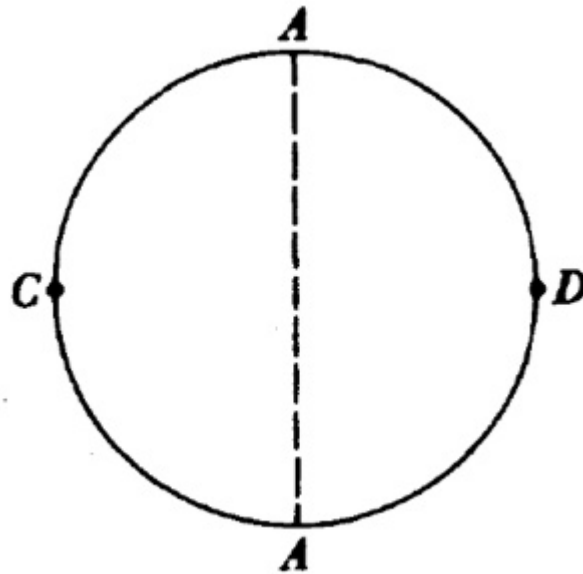


Figura 50.9

Esta superficie es cerrada y su número de conexión es 1, o bien su género 0. Puede engendrarse también pegando el borde de un círculo a lo largo de una banda de Möbius (que justamente tiene un solo borde), aunque de nuevo la figura no puede ser construida en tres dimensiones sin que resulte que coincidan puntos que deberían ser distintos.

Aún otro impulso a la investigación topológica vino de la geometría algebraica. Ya hemos mencionado (cap. 39, sec. 8) que los geómetras se habían dedicado a estudiar las «superficies» de cuatro dimensiones que representan el dominio de las funciones algebraicas de dos variables complejas, y habían introducido integrales sobre estas superficies, de manera análoga a como se hace en la teoría de funciones e integrales algebraicas sobre las

superficies de Riemann bidimensionales. Al estudiar estas figuras cuatridimensionales se investigó su tipo de conexión y se descubrió que no puede caracterizarse tales figuras por un número único, tal como el género caracteriza a las superficies de Riemann. Investigaciones llevadas a cabo por Emile Picard en torno a 1890 revelaron que para caracterizar dichas superficies serían necesarios por lo menos un número de conexión unidimensional y otro bidimensional.

La necesidad de estudiar el tipo de conexión de figuras de dimensión elevada fue reconocida por Enrico Betti (1823-1892), profesor de matemáticas en la universidad de Pisa, que decidió pasar directamente al caso de n dimensiones. Betti se había encontrado con Riemann en Italia, donde se había desplazado este último durante varios inviernos para cuidar su salud, y por él, Betti pudo tomar contacto con los trabajos de Riemann y de Clebsch. Betti⁵⁷⁷ introdujo los números de conexión para cada dimensión de 1 a $n - 1$. El número de conexión unidimensional es el número de curvas cerradas que pueden dibujarse en la estructura geométrica y que no dividen a la superficie en regiones disjuntas (el número de conexión de Riemann era una unidad mayor). El número de conexión bidimensional es el número de superficies cerradas en la figura que, de una manera colectiva, no limiten ninguna región tridimensional de la misma, y de una manera análoga se definen los números de conexión para dimensiones más altas. Las curvas, superficies y figuras de dimensión mayor, cerradas, utilizadas en estas definiciones se llaman ciclos. (Si una superficie tiene borde,

las curvas deben ser «secciones», es decir, curvas que van de un punto en el borde a otro punto del borde; por ejemplo, el número de conexión unidimensional de un cilindro hueco finito (sin extremos) es 1, porque puede trazarse una sección de un borde al otro sin desconectar la superficie). Betti demostró que el número de conexión unidimensional para las estructuras cuatridimensionales utilizadas para representar funciones algebraicas complejas $f(x, y, z) = 0$, es igual al número de conexión tridimensional.

4. La obra combinatoria de Poincaré

Hacia finales de siglo, el único campo que había sido cubierto de una manera bastante completa era la teoría de superficies cerradas; la obra de Betti era sólo el comienzo de una teoría más general. El autor que hizo el primer ataque sistemático y general a la teoría combinatoria de las figuras geométricas, y que suele ser considerado como el fundador de la topología, es Henri Poincaré (1854-1912). Poincaré fue profesor de matemáticas en la universidad de París, y se le conoce como el matemático más importante del último cuarto del siglo XIX y primeros años del siglo XX, así como el último hombre que tuvo un conocimiento universal de la matemática y de sus aplicaciones. Escribió un gran número de artículos de investigación, textos y artículos de divulgación, cubriendo casi todas las áreas básicas de la matemática y las más importantes de la física teórica, teoría del electromagnetismo, dinámica, mecánica de fluidos y astronomía; su obra máxima es *Les Méthodes nouvelles de la mécanique céleste* (3 vols., 1892-1899). Los

problemas físicos constituyeron la motivación de su investigación matemática.

Antes de ocuparse de la teoría combinatoria que vamos a describir, Poincaré había hecho contribuciones a otro campo de la topología, la teoría cualitativa de ecuaciones diferenciales (cap. 29, sec. 8); este trabajo era básicamente topológico porque se refería a la forma de las curvas integrales y a la naturaleza de los puntos singulares. Su contribución a la topología combinatoria estaba motivada por el problema de determinar la estructura de las superficies cuatridimensionales usadas para representar funciones algebraicas $f(x, y, z) = 0$, donde x, y, z son complejos. Decidió que era necesario un estudio sistemático del *análisis situs* de figuras generales n -dimensionales y, después de algunas notas en las «*Comptes Rendus*» de 1892 y 1893, publicó en 1895 ⁵⁷⁸ un trabajo básico, seguido por otros cinco largos suplementos esparcidos en varias revistas, hasta 1904. Poincaré consideraba su trabajo en topología combinatoria como una manera sistemática de estudiar la geometría «-dimensional, más bien que como un estudio de invariantes topológicos.

En su artículo de 1895, Poincaré intentaba abordar la teoría de figuras «-dimensionales usando sus representaciones analíticas, pero no hizo muchos progresos por este camino y volvió a una teoría puramente geométrica de variedades que son generalizaciones de las superficies de Riemann. Una figura es una variedad «-dimensional cerrada si cada punto posee entornos que son homeomorfos al interior «-dimensional de una $(n - 1)$ -esfera; por

ejemplo, la circunferencia (y cualquier figura homeomorfa) es una variedad unidimensional; una superficie esférica o un toro es una variedad bidimensional. Además de las variedades cerradas hay variedades con borde; un cubo y un toro sólido son variedades tridimensionales con borde; en un punto del borde un entorno es sólo parte del interior de una de dimensión 2.

El método que adoptó por último Poincaré aparece en su primer suplemento;⁵⁷⁹ aunque él usaba celdas curvas o trozos de las figuras y variedades estudiadas, nosotros vamos a exponer sus ideas en términos de simplexes y complejos, que fueron introducidos posteriormente por Brouwer. Un simplex es llanamente un triángulo n -dimensional; es decir, un simplex cero-dimensional es un punto, un simplex unidimensional es un segmento, un simplex bidimensional es un triángulo, uno tridimensional es un tetraedro, y un simplex n -dimensional es un tetraedro generalizado con $n + 1$ vértices. Las caras de dimensión menor de un simplex son a su vez simplexes. Un complejo es cualquier conjunto finito de simplex tales que la intersección de dos cualesquiera de ellos es una cara común, como máximo, y que toda cara de cada uno de los simplex del complejo es a su vez un simplex del complejo; a los simplex se les llama también celdas.

Para los fines de la teoría combinatoria se da una orientación determinada a cada simplex o celda de cualquier dimensión; por ejemplo, un 2-simplex E^2 (un triángulo) con vértices a_0 , a_1 y a_2 , puede ser orientado eligiendo un orden, digamos a_0 , a_1 , a_2 y entonces cualquier otro orden de los a_i que se obtenga a partir de

éste por un número par de permutaciones de los a_i se dice que tiene la misma orientación. Así, E^2 viene dado por (a_0, a_1, a_2) o (a_2, a_0, a_1) o (a_1, a_2, a_0) . Cualquier orden derivable del original por un número impar de permutaciones de los a_i representa el simplex orientado de manera opuesta; por ejemplo, $-E^2$ viene dado por (a_0, a_2, a_1) o (a_1, a_0, a_2) o (a_2, a_1, a_0) .

El borde de un simplex está formado por los simplex de dimensión una unidad inferior contenidos en el dado, por ejemplo, el borde de un 2-simplex consiste en tres 1-simplex. Sin embargo, el borde debe ser tomado con la orientación correcta; así, para obtener el borde orientado se adopta la regla siguiente: el simplex E^k

$$a_0 a_1 a_2 \dots a_k$$

induce la orientación

$$(-1)^i(a_0 a_1 a_2 \dots a_{i-1} a_{i+1} \dots a_k) \quad (1)$$

sobre cada uno de los simplex $(k-1)$ -dimensionales que forman el borde. El E^{k-1} puede tener la orientación dada por (1), en cuyo caso el número de incidencia que representa la orientación de E_i^{k-1} respecto de E^k , es 1, o puede tener la orientación opuesta, en cuyo caso su número de incidencia es -1. Sea 1 ó -1 el número de incidencia, el hecho básico es que el borde del borde de E^k es 0.

Dado un complejo cualquiera se puede formar combinaciones lineales de sus simplex orientados k -dimensionales; por ejemplo, si

los E_i^k son simplex k -dimensionales orientados, se tiene que

$$C^k = c_1 E_1^k + c_2 E_2^k + c_3 E_3^k \quad (2)$$

donde los c son enteros positivos o negativos, es tal combinación lineal y se llama una cadena. Los números c nos dicen simplemente cuántas veces está contado cada simplex, significando el hecho de ser negativo que la orientación del simplex en cuestión está cambiada. Por ejemplo, si nuestra figura es un tetraedro determinado por cuatro puntos $(a_0 \ a_1 \ a_2 \ a_3)$, podríamos formar la cadena $C^3 = 5 E_1^3$. El borde de una cadena es la suma de todos los simplex de dimensión inmediatamente inferior de todos los simplex de la cadena, tomado cada uno de ellos con el número de incidencia correspondiente y con la multiplicidad que aparece en (2). Debido a que el borde de una cadena es la suma de los bordes de cada uno de los simplex que aparecen en ella, el borde del borde de una cadena es cero.

Una cadena cuyo borde es *cero* se llama un ciclo; es decir, algunas cadenas son ciclos. Entre los ciclos, algunos son bordes de otras cadenas, tal como por ejemplo el borde del simplex E_1^3 es un ciclo que limita a E_1^3 . Sin embargo, si nuestra figura inicial no fuera el simplex tridimensional sino sólo su superficie, aún tendríamos el mismo ciclo, pero no sería borde en la figura considerada, es decir, en la superficie.

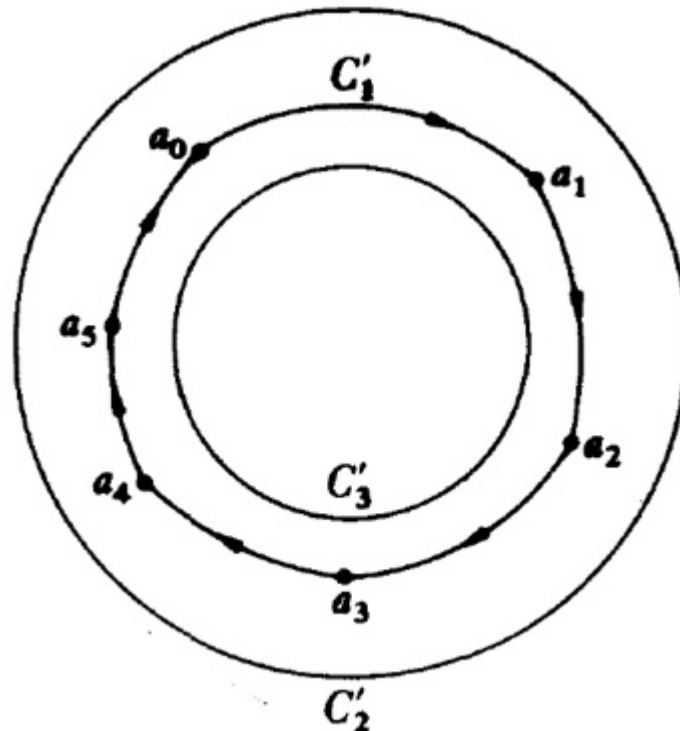


Figura 50.10

Otro ejemplo: la cadena (fig. 50.10).

$$C_1^1 = (a_0 a_1) + (a_1 a_2) + (a_2 a_3) + \dots + (a_5 a_0)$$

es un ciclo porque el borde de $a_1 a_2$, por ejemplo, es $a_2 - a_1$ y el borde de la cadena completa es cero; sin embargo C_1 no es borde de ninguna cadena bidimensional. Esto es intuitivamente evidente porque el complejo en cuestión es un anillo circular y el agujero interior no forma parte de la figura.

Es posible que dos ciclos por separado no limiten ninguna región pero sí su suma o diferencia; por ejemplo, la suma (o diferencia) de C_2^1 y de C_3^1 (fig. 50.10) limita el área correspondiente al anillo completo; dos ciclos como estos se llaman dependientes. En general,

los ciclos C_1^k, \dots, C_r^k se llaman dependientes si

$$\sum_{i=1}^r c_i C_i^k$$

es borde y no todas las c_i son cero.

Poincaré introdujo además los importantes parámetros que denominó números de Betti (en honor de Enrico Betti). Para cada dimensión de los posibles simplex de un complejo, el número de ciclos independientes de esa dimensión recibe el nombre de número de Betti correspondiente a esa dimensión (en realidad Poincaré usaba un número que es una unidad mayor que el número de conexión de Betti). Así, por ejemplo, en el caso del anillo el número de Betti cero-dimensional es 1 porque todo punto es un ciclo pero dos puntos limitan la sucesión de segmentos lineales que los unen; el número de Betti unidimensional es 1 porque hay 1-ciclos que no son bordes pero dos cualesquiera de ellos (su suma o diferencia) son borde; el número de Betti bidimensional del anillo es cero porque ninguna cadena de 2-simplex es un ciclo. Para darse cuenta de lo que significan estos números, uno puede comparar el anillo con el círculo mismo (incluido su interior); en esta última figura todo ciclo unidimensional es borde, de manera que el número de Betti unidimensional es cero.

También introdujo Poincaré en su artículo de 1899 lo que llamaba coeficientes de torsión. Es posible en estructuras más complicadas, por ejemplo en el plano proyectivo, tener un ciclo que no es borde

pero tal que dos veces dicho ciclo lo es; por ejemplo, si los simplex están orientados tal como muestra la figura 50.11, el borde de los cuatro triángulos es dos veces la recta BB (debe tenerse en cuenta que AB y BA son el mismo segmento lineal); el número 2 recibe el nombre de coeficiente de torsión, y el ciclo correspondiente

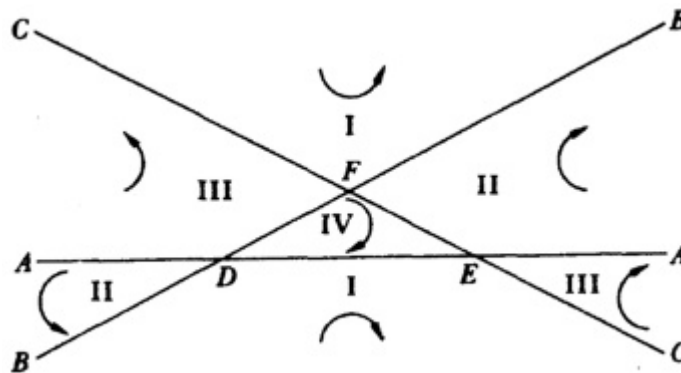


Figura 50.11

BB un ciclo de torsión; puede haber un número finito de tales ciclos de torsión independientes.

Incluso en estos ejemplos tan sencillos es claro que los números de Betti y los coeficientes de torsión de una figura geométrica distinguen de alguna manera una figura de otra, tal como se distingue el anillo circular del círculo.

En su primer suplemento (1899) y en el segundo,⁵⁸⁰ Poincaré introdujo un método para calcular los números de Betti de un complejo. Cada simplex E_q de dimensión q tiene simplex de dimensión $q - 1$ en su borde, los cuales tienen números de incidencia $+1$ ó -1 ; a un simplex $(q - 1)$ -dimensional que no forme parte del borde de E_q se le hace corresponder el número de

incidencia cero. Es posible entonces construir una matriz rectangular formada por los números de incidencia del ε_{ij}^q del j -ésimo simplex q -dimensional con respecto al i -ésimo simplex q -dimensional. Existe una matriz T_q para cada dimensión q distinta de cero; T_q tendrá tantas filas como simplex q -dimensionales haya en el complejo, y tantas columnas como simplex $(q - 1)$ -dimensionales. De esta forma, T_1 da las relaciones de incidencia de los vértices con las aristas, T_2 da las relaciones de incidencia de los simplex unidimensionales con los bidimensionales, y así sucesivamente. Mediante operaciones elementales sobre las matrices es posible hacer cero todos los elementos que no estén en la diagonal principal, y números enteros positivos o nulos los elementos de dicha diagonal. Supongamos que γ_q de estos elementos diagonales son 1; entonces demostró Poincaré que el número de Betti q -dimensional p_q (que es una unidad mayor que el número de conexión de Betti) es

$$p_q = \alpha_q - \gamma_{q+1} - \gamma_q + 1$$

donde α_q es el número de simplex q -dimensionales.

$$N(K^n) = \sum_{k=0}^n (-1)^k \alpha_k$$

Poincaré distinguía entre complejo con torsión y sin torsión; en el segundo caso todos los números que forman la diagonal principal

son 0 ó 1 para todo q , mientras que valores mayores indican la presencia de torsión.

También introdujo la característica $N(K^n)$ de un complejo n -dimensional K^n : si dicho complejo tiene a_k simplex k -dimensionales, entonces, por definición

$$N(K^n) = \sum_{k=0}^n (-1)^k p_k$$

Este número es una generalización del número de Euler $V - A + C$. Poincaré demostró que si p_k es el k -ésimo número de Betti de K^n , entonces⁵⁸¹

$$\sum_{i=1}^{n+1} x_i^2 = 1$$

resultado conocido como fórmula de Euler-Poincaré.

En su artículo de 1895 presenta Poincaré un teorema básico, con el nombre de teorema de dualidad, que se refiere a los números de Betti de una variedad cerrada. Una variedad cerrada «-dimensional, como dijimos anteriormente, es un complejo tal que cada uno de sus puntos tiene un entorno homeomorfo a una región del espacio euclídeo «-dimensional. El teorema afirma que en una variedad n -dimensional cerrada orientable, el número de Betti de dimensión p es igual al número de Betti de dimensión $n - p$; la demostración dada por Poincaré no era del todo completa, sin embargo.

En sus esfuerzos por distinguir unos complejos de otros introdujo Poincaré en 1895 otro concepto que juega actualmente un papel de una importancia considerable en topología, el grupo fundamental de un complejo, conocido también como el grupo de Poincaré o primer grupo de homotopía. La idea surge de considerar la diferencia entre regiones del plano simplemente y múltiplemente conexas; en el interior de un círculo todas las curvas cerradas pueden contraerse a un punto, pero en un anillo circular algunas de ellas, las que dejan en su interior la frontera circular más pequeña, no pueden ser contraídas a un punto, mientras que las curvas cerradas que no rodean dicha frontera interior sí pueden contraerse a un punto.

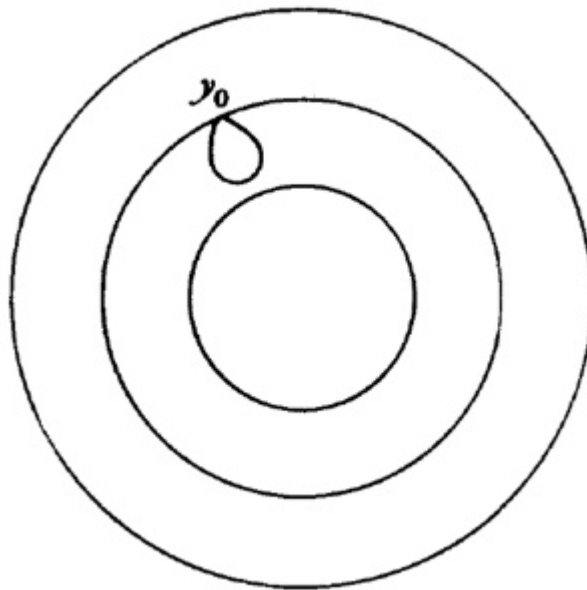


Figura 50.12

La mejor manera de introducir la idea precisa es considerar las curvas cerradas que empiezan y terminan en un punto dado y_0 del complejo; entonces, todas aquellas curvas que puedan ser

deformadas una en otra por un movimiento continuo dentro del complejo, se llaman homótopas entre sí y son consideradas como una clase. Así, las curvas cerradas que comienzan y terminan en y_0 en el anillo circular (fig. 50.12) y que no encierran la frontera interior, forman una clase; las que comienzan y terminan en y_0 y encierran la frontera interior forman otra clase; las que comienzan y terminan en y_0 y rodean n veces dicha frontera constituyen otra clase.

Entonces es posible definir una operación entre clases que geoméricamente consiste en partir de y_0 y recorrer una curva cualquiera de la primera clase, y a continuación una curva cualquiera de la segunda clase. El orden en que se toman las dos curvas así como el sentido de recorrido sobre cada una de ellas es importante. Las clases forman entonces un grupo no abeliano llamado el grupo fundamental del complejo con respecto al punto base y y representado por $\pi_1(K, y_0)$ siendo K el complejo. Para complejos razonablemente sencillos el grupo no depende en realidad del punto base, es decir, los grupos correspondientes a y_0 e y_1 , digamos, son isomorfos. Para el anillo circular el grupo fundamental es un grupo cíclico infinito, mientras que el dominio simplemente conexo que utiliza normalmente el análisis, es decir, el círculo cerrado, tiene un grupo fundamental trivial reducido al elemento identidad. De la misma forma que el círculo y el anillo circular se diferencian en sus grupos de homotopía, es posible describir complejos de mayor número de dimensiones que difieren entre sí marcadamente en este respecto.

Poincaré nos dejó algunas conjeturas importantes; por ejemplo, en su segundo suplemento afirma que dos variedades cerradas cualesquiera que tengan los mismos números de Betti y los mismos coeficientes de torsión, deben ser homeomorfas; sin embargo, en el quinto suplemento ⁵⁸² dio un ejemplo de una variedad tridimensional que tenía los mismos números de Betti y coeficientes de torsión que la esfera tridimensional (la superficie de una esfera sólida cuatridimensional) pero que no era simplemente conexa. En vista de ello, añadió la conexión simple como una condición más, demostrando que hay variedades tridimensionales con los mismos números de Betti y coeficientes de torsión, pero que tienen grupos fundamenta - y grupo fundamental sin ser por ello homeomorfas. Sin embargo, James W. Alexander (1888-1971), profesor de matemáticas en la universidad de Princeton y después en el Institute of Advanced Study, demostró⁵⁸³ que dos variedades de dimensión tres pueden tener los mismos números de Betti, coeficientes de torsión y grupo fundamental sin ser por ello homeomorfas.

Sin embargo, James W. Alexander (1888-1971), profesor de matemáticas en la universidad de Princeton y después en el Institute of Advanced Study, demostró que dos variedades de dimensión tres pueden tener los mismos números de Betti, coeficientes de torsión y grupo fundamental sin ser por ello homeomorfas.

En su quinto suplemento (1904) hizo Poincaré una conjetura un poco más restringida, consistente en que toda variedad

tridimensional cerrada, orientable y simplemente conexa es homeomorfa a la esfera de la misma dimensión. Esta famosa conjetura ha sido generalizada en la forma: toda variedad «-dimensional cerrada y simplemente conexa que tenga los mismos números de Betti y los mismos coeficientes de torsión que la esfera «-dimensional, es homeomorfa a ella; ni la conjetura de Poincaré ni la generalizada ha sido demostradas[‡].⁵⁸⁴

Otra famosa conjetura de Poincaré, llamada la «*Hauptvermutung*» (conjetura principal), afirma que si T_1 y T_2 son subdivisiones simpliciales de una misma 3-variedad, entonces T_1 y T_2 tienen subdivisiones isomorfas.⁵⁸⁵

5. Los invariantes combinatorios

El problema de establecer la invariancia de ciertas propiedades combinatorias consiste en demostrar que cualquier complejo homeomorfo a uno dado, considerados ambos como conjuntos de puntos, tiene las mismas propiedades combinatorias que el complejo dado. La demostración de que los números de Betti y los coeficientes de torsión son invariantes combinatorios fue dada por vez primera por Alexander,⁵⁸⁶ probando que si K y K_1 son subdivisiones simpliciales cualesquiera (no necesariamente rectilíneas) de dos poliedros homeomorfos (como conjuntos de puntos) P y P_1 entonces los números de Betti y los coeficientes de torsión de P y P_1 son iguales. El recíproco no es cierto, de manera que la igualdad de los números de Betti y de los coeficientes de

[‡] La conjetura de Poincaré ha sido resuelta para $n = 4$ por M. Freedman. Véase M. Freedman, *The topology of four dimensional manifolds*, J. Diff. Geom., 17 (1983), 357-454. (Nota del traductor.)

torsión de dos complejos no garantiza que dichos complejos sean homeomorfos.

Otro invariante de importancia se debe a L. E. J. Brouwer, que se vio interesado en topología a partir del estudio de problemas en teoría de funciones. Brouwer intentaba demostrar que hay $3g - 3$ clases de superficies de Riemann conformemente equivalentes, de género $g > 1$, y este problema le condujo a considerar otros problemas topológicos relacionados con él; Brouwer demostró ⁵⁸⁷ la invariancia de la dimensión de un complejo, en el sentido siguiente: si K es una subdivisión simplicial n -dimensional de un poliedro P , entonces toda subdivisión simplicial de P y toda división de ese tipo de cualquier poliedro homeomorfo a P es también un complejo n -dimensional.

Las demostraciones de este teorema, así como las del de Alexander, hacen uso de un método debido a Brouwer,⁵⁸⁸ llamado de aproximaciones simpliciales de las transformaciones continuas. Las transformaciones simpliciales (de simplex en simplex) no son otra cosa que las análogas de dimensión superior de las transformaciones continuas, mientras que las aproximaciones simpliciales de las transformaciones continuas son análogas a la aproximación lineal de las funciones continuas; si el dominio en el que se realiza la aproximación es pequeño, entonces la aproximación sirve para representar la transformación continua a efectos de demostraciones de invariancia.

6. Los teoremas de punto fijo

Aparte de servir para distinguir unos complejos de otros, los métodos combinatorios han permitido obtener los llamados teoremas del punto fijo, que son geoméricamente importantes y tienen también aplicaciones en análisis. Brouwer fue capaz, mediante la introducción de conceptos (en los que no entraremos) tales como la clase de una aplicación de un complejo en otro⁵⁸⁹ y el grado de una aplicación,⁵⁹⁰ de tratar lo que se llaman puntos singulares de los campos vectoriales sobre una variedad, en primer lugar. Consideremos la circunferencia S^1 , la superficie S^2 y la n -esfera $\Sigma_1^{n+1} x^2 = 1$ en el espacio euclídeo $(n + 1)$ -dimensional; sobre S^1 es posible tener un vector tangente en cada punto, de tal forma que las longitudes y las direcciones de estos vectores varían de manera continua a lo largo de la circunferencia y ningún vector tiene longitud cero; se dice entonces que se tiene un campo vectorial tangente continuo y sin singularidades sobre S^1 . Sin embargo no puede existir un campo así sobre S^2 , y Brouwer demostró⁵⁹¹ que lo que pasa en S^2 debe pasar en toda esfera de dimensión par, es decir, no pueden existir campos vectoriales continuos sobre una esfera de dimensión par, que no tengan al menos un punto singular.

Relacionada estrechamente con la teoría de puntos singulares está la teoría de transformaciones continuas de unos complejos en otros, y de interés especial en tales transformaciones son los puntos fijos. Si denotamos por $f(x)$ el transformado de un punto x , entonces un punto fijo es el que verifica $f(x) = x$; para todo punto x podemos considerar el vector que va de x a $f(x)$, y entonces en el caso de un

punto fijo el vector sería cero y dicho punto un punto singular del campo vectorial resultante. El teorema fundamental sobre existencia de puntos fijos es debido a Brouwer,⁵⁹² se aplica a simplex «-dimensionales (o bien homeomorfos a ellos), y afirma que toda transformación continua de un «-simplex en sí mismo tiene al menos un punto fijo. Así, por ejemplo, una transformación continua de un disco circular en sí mismo, debe tener al menos un punto fijo. En el mismo artículo demostró Brouwer que cualquier transformación continua inyectiva de una esfera de dimensión par en sí misma, que se pueda deformar en la transformación identidad, debe tener por lo menos un punto fijo.

Poco antes de su muerte en 1912, Poincaré demostró⁵⁹³ que existirían órbitas periódicas en un cierto problema restringido de los tres cuerpos, siempre que se verificase un cierto teorema topológico; dicho teorema afirma que existen al menos dos puntos fijos en una región anular contenida entre dos circunferencias, cuando se somete dicha región a una transformación topológica que transforma cada circunferencia en sí misma moviendo una en una dirección y la otra en la dirección opuesta, y que conserva el área. Este «último teorema» de Poincaré fue demostrado por George D. Birkhoff.⁵⁹⁴

Los teoremas de punto fijo fueron generalizados a espacios funcionales de dimensión infinita por Birkhoff y Olivier D. Kellogg en un trabajo conjunto,⁵⁹⁵ y fueron aplicados por Jules P. Schauder (1899-1940) ⁵⁹⁶ y conjuntamente por Schauder y Jean Leray (1906),⁵⁹⁷ para demostrar la existencia de soluciones de ecuaciones

diferenciales. Un teorema clave en tales aplicaciones dice que si T es una aplicación continua de un conjunto cerrado compacto y convexo de un espacio de Banach en sí mismo, entonces T tiene un punto fijo.

El uso de teoremas del punto fijo para demostrar la existencia de soluciones de ecuaciones diferenciales se entenderá mejor mediante un ejemplo sencillo. Consideremos la ecuación diferencial

$$\frac{dy}{dx} = F(x, y)$$

en el intervalo $0 \leq x \leq 1$, y la condición inicial $y = 0$ en $x = 0$. La solución $\Phi(x)$ satisface la ecuación

$$\phi(x) = \int_0^x F(x, \phi(x)) dx$$

Podemos entonces definir la transformación general

$$g(x) = \int_0^x F(x, f(x)) dx$$

donde $f(x)$ es una función arbitraria; esta transformación asocia la función g a la f , y se puede demostrar que es continua sobre el espacio de las funciones continuas $f(x)$ definidas en $[0,1]$. La solución Φ que buscamos es un punto fijo de este espacio de

funciones. Si se puede demostrar que este espacio funcional satisface las condiciones que garantizan la existencia de puntos fijos, entonces queda establecida la existencia de Φ ; los teoremas de punto fijo aplicables a espacios funcionales garantizan eso precisamente. El método ilustrado por este ejemplo sencillo nos permite establecer la existencia de soluciones de las ecuaciones en derivadas parciales no lineales que aparecen usualmente en el cálculo de variaciones y en hidrodinámica.

7. Generalizaciones y extensiones

Las ideas de Poincaré y Brouwer fueron aprovechadas por un cierto número de matemáticos que ampliaron el campo de la topología hasta el punto de que hoy día constituye uno de los dominios más activos de la matemática. Brouwer mismo extendió el teorema de la curva de Jordán;⁵⁹⁸ este teorema (cap. 42, sec. 5) puede formularse de la manera siguiente: sea S^2 una 2-esfera (superficie) y sea J una curva cerrada (topológicamente una S^1) en S^2 , entonces el número de Betti cero-dimensional de $S^2 - J$ es 2. Ya que este número de Betti es el número de componentes, J separa S^2 en dos regiones. La generalización de Brouwer afirma que una variedad $(n - 1)$ -dimensional separa el espacio euclídeo n -dimensional R_n en dos regiones. Alexander⁵⁹⁹ generalizó el teorema de dualidad de Poincaré e indirectamente el teorema de la curva de Jordán; el teorema de Alexander afirma que el número de Betti r -dimensional de un complejo K en la esfera n -dimensional S^n es igual al número de Betti $(n - r - 1)$ -dimensional del dominio complementario $S^n - K$, siendo $r \neq$

0 y $r \neq n - 1$. Para $r = 0$ el número de Betti de K es igual a 1 más el número de Betti $(n - 1)$ -dimensional de $S^n - K$, y para $r = n - 1$ el número de Betti de K es igual al número de Betti cero-dimensional de $S^n - K$ menos 1.⁶⁰⁰ Este teorema generaliza el teorema de la curva de Jordán, pues si tomamos como K la esfera $(n - 1)$ -dimensional S^{n-1} entonces el teorema asegura que el número de Betti $(n - 1)$ -dimensional de S^{n-1} que es 1, es igual al número de Betti cero-dimensional de $S^n - S^{n-1}$ menos 1, de manera que el número de Betti cero-dimensional de $S^n - S^{n-1}$ es 2 y S^{n-1} divide a S^n en dos regiones. Las definiciones de los números de Betti han sido modificadas y generalizadas en varios sentidos. Veblen y Alexander⁶⁰¹ introdujeron las cadenas y ciclos módulo 2, es decir, en vez de simplex orientados se emplean simplex no orientados, pero los coeficientes enteros se toman módulo 2, y los bordes de cadenas se cuentan de la misma manera; posteriormente Alexander⁶⁰² introdujo los coeficientes módulo m para cadenas y ciclos. Solomon Lefschetz (1884-1973) sugirió utilizar números racionales como coeficientes,⁶⁰³ y Lev L. Pontrjagin (1908-1960)⁶⁰⁴ hizo una generalización aún más amplia tomando como coeficientes de las cadenas elementos de un grupo abeliano; este último concepto incluye como casos particulares las clases de coeficientes mencionados más arriba, así como otra clase también utilizada, la de los números reales módulo 1. Todas estas generalizaciones, aunque conducían a teoremas más generales, no mejoran la capacidad de los números de Betti y de los coeficientes de torsión para distinguir complejos.

Otro cambio en la formulación de las propiedades combinatorias básicas, que fue llevado a cabo durante los años 1925 a 1930 por diversos matemáticos, posiblemente por sugerencia de Emmy Noether, fue el de reformular la teoría de cadenas y ciclos en el lenguaje de la teoría de grupos. Cadenas de la misma dimensión pueden sumarse unas con otras de la manera natural, es decir, sumando los coeficientes del mismo simplex, y debido a que los ciclos son cadenas pueden también sumarse ciclos y su resultado es otro ciclo. De esta forma las cadenas y los ciclos forman grupos. Dado un complejo K , para cada cadena k -dimensional existe el borde $(k - 1)$ -dimensional, y la suma de dos cadenas tiene como borde la suma de los bordes de las cadenas individuales; por tanto, la relación de cadena a borde establece un homomorfismo del grupo $C^k(K)$ de las cadenas k -dimensionales en un subgrupo $H^{k-1}(K)$ del grupo de las cadenas $(k - 1)$ -dimensionales. El conjunto de todos los k -ciclos ($k > 0$) es un subgrupo $Z^k(K)$ de $C^k(K)$ que se transforma por este homomorfismo en el elemento identidad o 0 de $C^{k-1}(A)$. Debido a que el borde de toda cadena es un ciclo, $H^{k-1}(K)$ es un subgrupo de $Z^{k-1}(K)$.

Sobre la base de estos hechos podemos formular la definición siguiente: para todo $k \geq 0$, el grupo cociente de $Z^k(A)$, es decir, del grupo de los ciclos k -dimensionales, módulo el subgrupo $H^k(K)$ de los ciclos borde, recibe el nombre de grupo de homología k -ésimo de K , representado por $B^k(K)$. El número de generadores linealmente independientes de este grupo cociente, recibe el nombre de número de Betti k -ésimo del complejo, y se representa por $p^k(K)$. El grupo k -

ésimo de homología puede contener también grupos cíclicos finitos, los cuales corresponden a los ciclos de torsión; de hecho, los órdenes de estos grupos finitos son los coeficientes de torsión. Con esta formulación de los grupos de homología de un complejo, muchos resultados antiguos pueden reformularse en el lenguaje de la teoría de grupos.

La generalización más importante de la primera parte de este siglo fue la de introducir la teoría de homología para espacios generales, tales como los espacios métricos compactos, en vez de partir de figuras que sean complejos. Las ideas básicas se deben a Paul S. Alexandroff (n. 1896),⁶⁰⁵ Leopold Vietoris (1891)⁶⁰⁶ y Eduard Cech (1893-1960).⁶⁰⁷ No daremos aquí los detalles, ya que suponen una formulación totalmente nueva de la teoría de homología. Es necesario hacer observar, sin embargo, que este trabajo marca un hito en el proceso de fusión de la topología general o conjuntista y la topología combinatoria.

Bibliografía

- Bouligand, Georges: *Les Définitions modernes de la dimension*, Hermann, 1935.
- Dehn, M., y Heegard, P.: «Analysis Situs». *Encyk. der Math. Wiss.*, B. G.
- Teubner, 1907-1910, III AB3, 153-220.
- Franklin, Philip: «The Four Color Problem». *Scripta Mathematica*, 6, 1939, 149-156, 197-210.
- Hadamard, J.: «L'OEuvre mathématique de Poincaré». *Acta*

Math., 38, 1921, 203-287.

- Manheim, J. H.: *The Génesis of Point Set Topology*, Pergamon Press, 1964.
- Osgood, William F.: «Topics in the Theory of Functions of Several Complex Variables». *Madison Colloquium*, American Mathematical Society, 1914, 111-230.
- Poincaré, Henri: *OEuvres*, Gauthier-Villars, 1916-1956, vol. 6.
- Smith, David E.: *A Source Book in Mathematics*, Dover (reimpresión), 1959, vol. 2, 404-410.
- Tietze, H., y Vietoris L.: «Beziehungen zwischen den verschiedenen Zweigen der Topologie». *Encyk. der Math. Wiss.*, B. G. Teubner, 1914-1931, III AB13, 141-237.
- Zoretti, L., y Rosenthal, A.: «Die Punktmengen». *Encyk. der Math. Wiss.*, B. G. Teubner, 1923-1927, II C9A, 855-1030.

Capítulo 51

Los fundamentos de la matemática

La lógica es invencible, porque para combatir la lógica es necesario utilizar la lógica misma.

Pierre BOUTROUX

Sabemos que los matemáticos no se preocupan más por la lógica que los lógicos por la matemática. Los dos ojos de que dispone la ciencia exacta son la matemática y la lógica; la secta matemática extrae el ojo lógico, la secta lógica extrae el ojo matemático, pensando cada uno que puede ver mejor con un ojo que con dos.

Augustos de Morgan

Contenido:

- 1. Introducción*
- 2. Las paradojas de la teoría de conjuntos*
- 3. La axiomatización de la teoría de conjuntos*
- 4. La aparición de la lógica matemática*
- 5. La escuela logicista*

6. La escuela intuicionista

7. La escuela formalista

8. Algunos desarrollos recientes

Bibliografía

1. Introducción

La actividad más profunda de la matemática del siglo XX, con gran diferencia sobre las demás, ha sido la investigación sobre sus propios fundamentos. Los problemas que le vinieron impuestos al respecto a los matemáticos, y otros que ellos asumieron voluntariamente, se refieren no sólo a la naturaleza misma de la matemática, sino también a la validez de la matemática deductiva.

Hubo varias circunstancias que convergieron para poner en primera línea los problemas de fundamentos a principios de siglo. La primera fue el descubrimiento de las contradicciones, llamadas eufemísticamente paradojas, especialmente en la teoría de conjuntos. De una de estas contradicciones, la paradoja de Burali-Forti, hemos hablado ya (cap. 41, sec. 9), y durante los primeros años de este siglo se descubrieron varias otras. Naturalmente, tales descubrimientos inquietaron profundamente a algunos matemáticos. Otro problema que había ido surgiendo gradualmente y que al fin apareció en escena en sus verdaderas proporciones, fue el de la consistencia de la matemática (cap. 43, sec. 6). En vista de las paradojas de la teoría de conjuntos, era allí donde urgía especialmente establecer la consistencia.

Durante los últimos años del siglo XIX hubo un cierto número de matemáticos que comenzaron a preocuparse de los fundamentos de su ciencia y, de una manera muy particular, de las relaciones entre la matemática y la lógica. Las investigaciones en este área (de cuyos detalles hablaremos más adelante) sugirieron a algunos matemáticos la idea de que la matemática podría fundamentarse sobre la lógica. Otros, en cambio, pusieron en cuestión la aplicación universal de los principios lógicos, el dudoso significado de algunas demostraciones existenciales, e incluso la fiabilidad de la demostración lógica como justificación de los resultados matemáticos obtenidos. Las controversias, que habían permanecido en estado latente como los rescoldos antes de 1900, se convirtieron en un fuego declarado cuando las paradojas y los problemas de consistencia añadieron combustible. Inmediatamente se volvió vital y de un interés general la cuestión de la fundamentación adecuada de toda la matemática.

2. Las paradojas de la teoría de conjuntos

Poco después del descubrimiento por Cantor y Burali-Forti de la paradoja sobre los números ordinales, aparecieron varias otras paradojas o antinomias. En realidad la palabra paradoja es ambigua e inadecuada en este contexto, porque parece referirse a una contradicción sólo aparente, mientras que lo que se habían encontrado los matemáticos eran incuestionables y palmarias contradicciones. Veamos primero en qué consistían.

Lina de las paradojas fue popularizada por Bertrand Russell (1872-1970) en 1918, con el nombre de «paradoja del barbero». El barbero de un pueblecillo, presumiendo de no tener competencia, se anuncia diciendo que él no afeita a aquellos que se afeitan a sí mismos, desde luego, pero sí afeita a todos aquellos que no se afeitan a sí mismos. Un buen día se le ocurre preguntarse si debería afeitarse a sí mismo o no. Si se afeita a sí mismo, entonces, por la primera parte de su afirmación, no debería afeitarse a sí mismo. Pero si no se afeita a sí mismo, entonces por la segunda parte debería afeitarse a sí mismo. El pobre barbero se encuentra en un verdadero atolladero lógico.

Otra paradoja formulada por Jules Richard (n. 1862),⁶⁰⁸ de la que G. G. Berry y Russell dieron una versión simplificada, fue publicada por el segundo de ellos.⁶⁰⁹ La paradoja simplificada, conocida también como «paradoja de Richard», dice así: todo número natural puede definirse utilizando palabras con un cierto número de letras, en general de diversas maneras. Por ejemplo, el número 36 puede definirse como «treinta y seis» o como «cuatro por nueve»; la primera definición contiene doce letras y la segunda catorce. No hay una manera uniforme de definir cualquier número, pero esto no es esencial. Clasifiquemos ahora todos los números naturales en dos grupos; el primero incluirá todos aquellos que puedan definirse (de una manera por lo menos) con 100 letras o menos, mientras que el segundo incluirá todos aquellos que, de cualquier manera que se definan, se necesiten para ello al menos 101 letras. Es obvio que sólo hay un número finito de números que se puedan definir con

100 letras o menos, puesto que hay exactamente 27^{100} expresiones con 100 letras (y la mayoría no tienen sentido). Existe por tanto el número mínimo del segundo grupo que, evidentemente, puede definirse por la frase «el mínimo número natural que no se puede definir con cien letras o menos». Pero esta frase tiene menos de 100 letras. Luego el mínimo número natural no definible con 100 letras o menos puede definirse con menos de 100 letras.

Otra forma de esta paradoja fue enunciada por primera vez en 1908 por Kurt Grelling (1886-1941) y Leonard Nelson (1882-1927), que la publicaron en una revista muy poco conocida.⁶¹⁰ Algunas palabras (adjetivos concretamente) se describen o aplican a sí mismas y otras no. Por ejemplo, la palabra «polisílaba» es polisílaba; «corta» es corta. En cambio la palabra «monosílaba» no es monosílaba y «larga» no es larga. Llamemos a todas aquellas que se aplican a sí mismas «autológicas», y a las que no se aplican a sí mismas «heterológicas». Así pues, « X » es heterológica si « X » no tiene la propiedad X . Si lo aplicamos ahora a la palabra heterológica, sustituyéndola por la X , nos queda que «heterológica» es heterológica si «heterológica» no es heterológica.

Cantor señalaba, en carta a Dedekind de 1899, que no se podría hablar del conjunto de todos los conjuntos sin entrar en contradicción (cap. 41, sec. 9). Algo análogo ocurre con la paradoja de Russell (The Principles of Mathematics, 1903, p. 101). La clase de todos los hombres no es un hombre, pero la clase de todas las ideas es una idea; la clase de todas las bibliotecas es una biblioteca, y la clase de todos los conjuntos de cardinal mayor que uno tiene

cardinal mayor que uno. Por tanto, algunas clases no son elementos de sí mismas y otras sí lo son. Esta clasificación las abarca a todas y los dos tipos son mutuamente excluyentes. Sea M la clase de todas las clases que son elementos de sí mismas, y A la clase de todas las clases que no son elementos de sí mismas. Podemos preguntarnos ahora si la clase N cae dentro de M o de N . Si N perteneciese a N entonces sería elemento de sí misma y por tanto debería pertenecer a M . Por otro lado, si N pertenece a M , entonces, dado que M y N son mutuamente excluyentes, N no podría pertenecer a N . Por tanto, N no es elemento de sí misma y, en virtud de su propia definición, debería pertenecer a N .

La causa de todas estas paradojas, como señalan Russell y Whitehead, radica en la definición de un objeto en términos de una clase que contiene como elemento al objeto que se está definiendo. Tales definiciones se llaman impredicativas y aparecen de manera especial en teoría de conjuntos. Como hizo observar Zermelo en 1908, este tipo de definición se utiliza también para definir el extremo superior de un conjunto acotado de números reales y otros conceptos del análisis. Así pues, el análisis contiene paradojas.

La demostración de Cantor de la no numerabilidad del conjunto de los números reales (cap. 41, sec. 7) también utiliza un conjunto definido impredicativamente. Se supone dada una correspondencia biunívoca f entre el conjunto de todos los números naturales y el conjunto M de todos los números reales (identificados a conjuntos de números naturales). A cada número natural k corresponde el conjunto $f(k)$. Ahora bien, k pertenece o no pertenece a $f(k)$. Sea N el

conjunto de todos los k tales que k no pertenece a $f(k)$; este conjunto N es un número real y, por tanto, por la correspondencia biunívoca f tiene que haber un único natural n tal que $f(n) = N$. Ahora bien, si n pertenece a N tendría que no pertenecer, por la definición misma de N ; pero si no perteneciese debería pertenecer por la misma razón. No puede existir, por tanto, la correspondencia f . La definición del conjunto N es impredicativa porque k pertenece a A si y sólo si existe un conjunto K en M tal que $K = f(k)$ y k no pertenece a K . Así pues, al definir N hacemos uso de la totalidad de los conjuntos M que contienen a N como elemento; por tanto, para definir N , N ha de estar ya en M .

Es muy fácil caer inadvertidamente en la trampa de introducir definiciones impredicativas. Así, si uno define la clase de todas las clases que contiene más de cinco elementos, ha definido una clase que se contiene a sí misma como elemento. Análogamente, la frase «el conjunto S de todos los conjuntos definibles con treinta palabras o menos», define a S impredicativamente.

Estas paradojas inquietaron a muchos matemáticos, dado que ponían en cuestión no sólo la teoría de conjuntos, sino también grandes dominios del análisis clásico. La matemática, como estructura lógica, se encontraba en un estado bien triste, y los matemáticos pensaban con añoranza en los viejos y felices días de antes de que aparecieran las paradojas.

3. La axiomatización de la teoría de conjuntos

No nos puede sorprender el hecho de que una de las primeras medidas que tomaran los matemáticos fuera la de axiomatizar la teoría de conjuntos que había formulado Cantor de una manera muy libre o informal o, como algunos prefieren decir hoy, intuitiva. La axiomatización de la geometría y de la aritmética había permitido resolver problemas lógicos en esas ramas, y parecía verosímil que la axiomatización también clarificaría las dificultades de la teoría de conjuntos. El primero en emprender esta tarea fue el matemático alemán Ernst Zermelo, que creía que las paradojas habían aparecido porque Cantor no había restringido adecuadamente el concepto de conjunto. En 1895⁶¹¹ Cantor había definido un conjunto como una colección de objetos distintos de nuestra intuición o nuestro pensamiento. Esto era bastante vago, y por tanto Zermelo esperaba que un sistema de axiomas precisos y explícitos clarificaría qué es lo que se entiende por un conjunto y cuáles son sus propiedades. Cantor mismo no había dejado de ser consciente de que su concepto de conjunto presentaba dificultades; en una carta a Dedekind de 1899⁶¹² distinguía entre conjuntos consistentes e inconsistentes. Zermelo pensó que podría restringir sus conjuntos a los consistentes de Cantor, y que éstos serían suficientes para la matemática. Su sistema de axiomas⁶¹³ contenía conceptos y relaciones fundamentales que estaban definidas implícitamente por las afirmaciones de los axiomas mismos. Entre tales conceptos los fundamentales eran el de conjunto y la relación de pertenencia de un elemento a un conjunto. No debería utilizarse ninguna propiedad de los conjuntos salvo que estuviese garantizada

por un axioma. La existencia de un conjunto infinito y las operaciones tales como la unión de conjuntos o la formación de subconjuntos también estaban especificadas en dichos axiomas. Zermelo incluyó de manera destacada el llamado «axioma de elección» (cap. 41, sec. 8).

El plan de Zermelo era el de admitir en la teoría de conjuntos sólo aquellas clases de las que verosímilmente no pudieran derivarse contradicciones. Así, por ejemplo, la clase vacía, cualquier clase finita y la clase de los números naturales parecían seguras. Dada una clase segura, ciertas otras clases formadas a partir de ella, tales como cualquiera de sus subclases, la clase de todas sus subclases y la unión de clases seguras, deberían ser a su vez clases seguras. Evitó, sin embargo, la complementación, puesto que, aunque x sea una clase segura, el complemento de x , es decir, todos los no- x en algún universo de objetos muy grande podría no ser segura.

La fundamentación de la teoría de conjuntos de Zermelo fue mejorada por Abraham A. Fraenkel (1891-1965),⁶¹⁴ y von Neumann⁶¹⁵ introdujo cambios adicionales. En el caso del sistema de Zermelo-Fraenkel, la esperanza de evitar las paradojas se basa en restringir los tipos de conjuntos que se admiten, siempre que los admitidos sean suficientes para las necesidades del análisis. La idea de von Neumann era un poco más atrevida, estableciendo una distinción entre conjuntos y clases propias. Las clases propias eran colecciones tan grandes que no estaban contenidas como elementos en ninguna otra clase o conjunto, mientras que los conjuntos eran colecciones más restringidas, que podían ser elementos de otras

clases. Así pues, los conjuntos eran exactamente las clases seguras y, como apuntaba von Neumann, no era la admisión de las clases propias lo que conducía a contradicciones, sino el tratarlas como elementos de otras clases.

La teoría de conjuntos formal de Zermelo, modificada por Fraenkel, von Neumann y otros, resulta adecuada para desarrollar la teoría de conjuntos que se necesita para todo el análisis clásico prácticamente, y evitar las paradojas en el sentido de que, hasta hoy, nadie ha descubierto ninguna dentro de la teoría. Sin embargo, la consistencia de la teoría axiomática de conjuntos no ha sido demostrada. A propósito de esta cuestión abierta de la consistencia, decía Poincaré: «Hemos puesto una valla en torno al rebaño para protegerlo de los lobos, pero lo que no sabemos es si habrá quedado algún lobo dentro ya de la valla.»

Aparte del problema de la consistencia, hay que decir que la axiomatización de la teoría de conjuntos incluía el axioma de elección, necesario para establecer importantes resultados del análisis clásico, de la topología y del álgebra abstracta. Un cierto número de matemáticos, entre los que se contaban Hadamard, Lebesgue, Borel y Baire, consideraron objetable este axioma, y en 1904, cuando Zermelo lo utilizó para demostrar el teorema de buena ordenación (cap. 41, sec. 8), una verdadera avalancha de objeciones inundó las revistas matemáticas.⁶¹⁶ Las cuestiones acerca de si este axioma era esencial para los resultados en los que se utilizaba, y de si era independiente de los restantes o no, no tardaron en

plantearse, permaneciendo durante muchos años sin ser resueltas (véase la sec. 8).

La axiomatización de la teoría de conjuntos, a pesar de que dejaba abiertas cuestiones como el problema de la consistencia y el papel del axioma de elección, podría haber tranquilizado al menos a los matemáticos con respecto a las paradojas, haciendo declinar así el interés por los fundamentos. Pero para entonces ya habían entrado en actividad y en plena polémica varias escuelas filosóficas sobre los fundamentos de la matemática, motivadas sin duda por las paradojas y el problema de la consistencia. Los partidarios de esas líneas filosóficas no encontraban satisfactorio el método axiomático tal como lo practicaban Zermelo y otros. A algunos de ellos les resultaba cuestionable porque presuponia la lógica que utilizaba, mientras que por entonces ya estaban sometidas a investigación tanto la lógica misma como sus relaciones con la matemática. Otros, más radicales, rechazaban apoyarse en cualquier tipo de lógica, especialmente al aplicarla a conjuntos infinitos. Para entender los argumentos que proponían las diversas escuelas de pensamiento, necesitamos retroceder un poco en el tiempo.

4. La aparición de la lógica matemática

Un tema que provocó nuevas controversias e insatisfacciones con la axiomatización de la teoría de conjuntos es el que se refiere al papel de la lógica en la matemática, que provenía ya de la matematización de la lógica llevada a cabo durante el siglo XIX. Esta corriente de desarrollo tiene su propia historia.

La potencia del álgebra para representar simbólicamente e incluso mecanizar argumentos geométricos, había impresionado ya a Descartes y a Leibniz, entre otros (cap. 13., sec. 8), y ambos concibieron una ciencia mucho más general que el álgebra de los números. Se trataría de una ciencia abstracta o general de razonamiento, que operaría de un modo algo parecido al álgebra ordinaria, pero que sería aplicable al razonamiento en todos los campos. Tal como lo expresaba Leibniz en uno de sus escritos, «la matemática universal es, por así decirlo, la lógica de la imaginación», y debería tratar de «todo aquello que en el dominio de la imaginación sea susceptible de determinación exacta». Con una lógica tal se podría construir cualquier edificio del pensamiento, a partir de sus elementos más simples, hasta alcanzar estructuras más y más complicadas. El álgebra universal sería parte de la lógica, pero de una lógica en forma algebraica. Descartes comenzó modestamente, intentando construir un álgebra de la lógica, de la que nos queda un bosquejo incompleto.

Leibniz, persiguiendo la misma meta general que Descartes, se propuso un programa más ambicioso. A lo largo de toda su vida había prestado atención a la lógica, y desde muy joven se había visto fascinado por el proyecto del teólogo medieval Raimundo Lulio (1235-1315), que en su libro *Ars Magna et Ultima* presentaba un ingenuo método mecánico para producir nuevas ideas combinando otras ya existentes, que implicaba la idea de una ciencia universal de la lógica que sería aplicable a todo tipo de razonamientos. Leibniz se apartó de la lógica escolástica y de Lulio, pero quedó

impresionado por la posibilidad de un cálculo muy general que le permitiría al hombre razonar en todos los campos mecánicamente y sin esfuerzo alguno. Leibniz nos dice de su plan para una lógica simbólica universal que tal ciencia, de la que el álgebra ordinaria sólo sería una pequeña parte, sólo estaría limitada por la necesidad de obedecer las leyes de la lógica formal. Podríamos llamarla, decía él, una «síntesis algebraico-lógica».

Esta ciencia general suministraría, ante todo, un lenguaje racional universal, adaptado a todo tipo de pensamiento. Los conceptos, una vez resueltos en otros primitivos, distintos y separados que no se solapen, podrían combinarse de una manera casi mecánica. Leibniz pensó también que sería necesario un simbolismo especial para evitar a la mente perderse. Aquí se puede ver claramente la influencia del simbolismo algebraico en su pensamiento. Lo que pretendía era un lenguaje simbólico capaz de expresar los pensamientos humanos sin ambigüedades, y de ayudar a realizar deducciones. Este lenguaje simbólico sería su *characteristica universalis*.

En 1666 escribió Leibniz su *De Arte Combinatoria*,⁶¹⁷ que contiene, entre otras cosas, sus primeras ideas sobre el sistema universal de razonamiento. Más tarde escribiría numerosos otros fragmentos sobre el mismo tema que no llegaron a publicarse nunca, pero que pueden verse recogidos en la edición de sus escritos filosóficos (ver nota 10). En su primer ensayo asociaba a cada concepto primitivo un número primo; así, cualquier concepto compuesto de varios otros primitivos venía representado por el producto de los primos

correspondientes. Por ejemplo, si 3 representa «animal» y 7 «racional», 21 representaría «animal racional». A continuación intentó trasladar las reglas usuales de los silogismos a este esquema pero no tuvo éxito. También intentó en otra ocasión utilizar símbolos especiales en vez de números primos, donde, una vez más, las ideas complejas se representarían por combinaciones de símbolos. En realidad, Leibniz pensaba que el número de ideas primitivas sería pequeño, cosa que se demostró errónea. Tampoco fue suficiente con una sola operación básica, la conjunción, para construir todas las composiciones de ideas primitivas.

También comenzó a trabajar Leibniz en un álgebra de la lógica propiamente dicha. Directa e indirectamente disponía en su álgebra de conceptos que hoy describiríamos como adición lógica, multiplicación lógica, identidad, negación y la clase vacía. También llamó la atención sobre el interés del estudio de relaciones abstractas, tales como la inclusión, las correspondencias biunívocas o multívocas y las relaciones de equivalencia, reconociendo que algunas de estas relaciones tenían las propiedades de simetría y transitividad. Leibniz no completó esta obra; no pasó de las reglas de los silogismos que, como él mismo tuvo que admitir, no agotan toda la lógica que utiliza la matemática. Leibniz comunicó sus ideas a L'Hôpital y a otros, pero no le prestaron atención y sus escritos lógicos permanecieron sin publicarse hasta comienzos del siglo XX, de manera que su influencia directa fue muy pequeña. Durante el siglo XVIII y comienzos del XIX algunos hicieron intentos análogos a los de Leibniz, pero no llegaron más lejos.

Una etapa más eficaz, aunque menos ambiciosa, fue la llevada a cabo por Augustus de Morgan. De Morgan publicó su *Formal Logic* en 1847 y muchos artículos, algunos de los cuales aparecieron en las *Transactions of the Cambridge Philosophical Society*, en los que trataba de corregir defectos de la lógica aritotélica y mejorarla. En su *Formal Logic* añadía un nuevo principio a la lógica aristotélica. En esta última las premisas «Algunos M son A » y «Algunos M son B » no permiten extraer ninguna conclusión; de hecho, esta lógica afirma que el término medio M ha de aparecer de manera universal, es decir, deben aparecer «Todos los M ». De Morgan sostiene, sin embargo, que de «La mayor parte de los M son A » y de «La mayor parte de los M son B » se sigue necesariamente que «Algunos A son B ». De Morgan razona este hecho de una forma cuantitativa. Si hay m de los M , y a de los M son A , y b de los M son B , entonces al menos $(a + b - m)$ de los A son B . El interés de la observación de de Morgan está en que los términos pueden ser cuantificados y, como consecuencia pudo introducir muchas más formas válidas de silogismo. La cuantificación también eliminó un defecto de la lógica aristotélica: la conclusión «Algunos A son B », que en lógica aristotélica se sigue de la premisa «Todos los A son B », implicaría la existencia de los A , que no tienen por qué existir.

De Morgan inició también el estudio de la lógica de relaciones. La lógica aristotélica está dedicada básicamente a la relación «ser», y se limita a afirmar o negar esta relación. Como hizo observar de Morgan, esta lógica no podría demostrar que si un caballo es un animal, entonces una cola de caballo es una cola de un animal, y

ciertamente no podría manejar una relación tal como la « x ama a y ». De Morgan introdujo un simbolismo adecuado al manejo de relaciones, pero no llegó demasiado lejos.

En el campo de la lógica simbólica, de Morgan es bien conocido sobre todo por las hoy llamadas leyes de Morgan. Tal como él las formuló,⁶¹⁸ el contrario de un agregado es el compuesto de los contrarios de los agregados; y el contrario de un compuesto es el agregado de los contrarios de los componentes. En simbolismo lógico estas dos leyes se escriben

$$1 - (x + y) = (1 - x)(1 - y)$$

$$1 - xy = (1 - x) + (1 - y).$$

La contribución del simbolismo a un álgebra de la lógica constituye la mayor aportación de George Boole (1815-1864), que fue en gran medida un autodidacta y llegó a ser profesor de matemáticas en el Kings College de Cork. Boole estaba profundamente convencido de que la simbolización del lenguaje rigORIZARÍA la lógica. Su obra *Mathematical Analysis of Logic*, que curiosamente apareció el mismo día que la *Formal Logic* de de Morgan, y *An Investigation of the Laws of Thought* (1854), contienen sus ideas más importantes.

El planteamiento básico de Boole consistía en hacer recaer el énfasis en la lógica extensional, es decir, en una lógica de clases, donde, por cierto, los conjuntos o clases se representaban por x , y , z ,..., mientras que los símbolos X , Y , Z ,... representaban a los miembros individuales. La clase universal venía representada por 1

y la clase vacía o nula por 0. Boole utilizó el término xy para representar la intersección de dos conjuntos x e y (operación a la que llamó elección), es decir, el conjunto de todos los elementos comunes a x y a y , y el $x + y$ para representar el conjunto de los elementos tanto de x como de y (para Boole, hablando estrictamente, la unión o adición sólo se aplicaba a conjuntos disjuntos; W. S. Jevons (1835-1882) generalizó este concepto). El complemento de x viene representado por $1 - x$ y, de una manera más general, $x - y$ es la clase de los elementos de x que no están en y . La relación de inclusión, es decir, x está contenido en y , la representó por $xy = x$, y en todos los casos el signo igual representa la identidad de dos clases.

Boole creía que nuestra mente nos otorga de modo inmediato ciertos procesos de razonamiento elementales que son los axiomas de la lógica. Por ejemplo, el principio de contradicción, que A no puede ser nunca B y no B a la vez, es un hecho axiomático, que se expresa por la fórmula

$$x(l - x) = 0$$

También resulta obvio para la mente que

$$xy = yx$$

y así, esta propiedad conmutativa de la intersección es otro axioma. Igualmente evidente es la propiedad

$$xx = x$$

axioma en el que la lógica se aparta del álgebra ordinaria. Boole aceptó también como axiomático que

$$x + y = y + x$$

y

$$x(u + v) = xu + xv.$$

Con estos axiomas, la ley del tercio excluso puede formularse ya en la forma

$$x + (1 - x) = 1$$

es decir, todo objeto es x o no x . Todo X es Y se expresa por $x(1 - y) = 0$. Ningún X es Y por $xy = 0$. Algunos X son Y por $xy \neq 0$, y algunos X no son Y por $x(1 - y) \neq 0$.

Boole planeaba deducir de los axiomas las leyes generales del razonamiento, aplicando repetidamente los procesos permitidos por dichos axiomas. Como conclusiones triviales obtuvo que $1x = x$ y que $0x = 0$. Un ejemplo de un argumento un poco más complicado es el siguiente: de

$$x + (1 - x) = 1$$

se sigue que

$$z(x + (1 - x)) = z \cdot 1$$

y por tanto que

$$zx + z(1 - x) = z.$$

Así pues, la clase de objetos z consiste en los que están en x y en los que están en $1 - x$.

Boole observó que el cálculo de clases se podía interpretar también como un cálculo de proposiciones. Así, si x e y son proposiciones en vez de clases, entonces xy sería la afirmación conjunta de x e y , y $x + y$ la afirmación de x o y o ambas. La proposición $x = 1$ significaría que x es verdadera, mientras que $x = 0$ que x es falsa; $1 - x$ significaría la negación de x . Sin embargo, Boole no llegó demasiado lejos con su cálculo de proposiciones.

Tanto a de Morgan como a Boole los podemos considerar en justicia como los reformadores de la lógica aristotélica y los iniciadores del álgebra de la lógica. El resultado de sus obras fue el de construir una ciencia de la lógica que iba a estar en adelante separada de la filosofía y unida a la matemática.

El cálculo de proposiciones progresó de la mano de Charles S. Peirce. Peirce distinguió entre una proposición y una función proposicional. Una proposición, «Juan es un hombre», sólo contiene

constantes, mientras que una función proposicional, « x es un hombre», contiene variables. Mientras que una proposición es verdadera o falsa, una función proposicional en general es verdadera para algunos valores de la variable y falsa para otros. Peirce introdujo también las funciones preposicionales de dos variables, por ejemplo, « x conoce a y ».

Todos los que habían contribuido a construir la lógica simbólica hasta este momento, estaban interesados principalmente en la lógica y en matematizarla. Con la obra de Gottlob Frege (1848-1925), profesor de matemáticas de Jena, la lógica matemática toma una dirección nueva, que es la que nos interesa en relación con los problemas de fundamentación de la matemática. Frege escribió varias obras importantes: *Begriffsschrift* (Cálculo de conceptos, 1879), *Die Grundlagen der Arithmetik* (Los Fundamentos de la Aritmética, 1884), y *Grundgesetze der Arithmetik* (Las Leyes Fundamentales de la Aritmética; vol. 1, 1893; vol. 2, 1903). Todas estas obras se caracterizan por su alto nivel de precisión y completitud en los detalles.

En el campo de la lógica propiamente dicha, Frege generalizó el uso de las variables, cuantificadores y funciones preposicionales; la mayor parte de esta obra fue llevada a cabo independientemente de sus predecesores, incluido Pierce. En el *Begriffsschrift* Frege da una fundamentación axiomática de la lógica, introduciendo muchas distinciones que más tarde adquirieron gran importancia; por ejemplo, la distinción entre el enunciado de una proposición y la afirmación de que es verdadera. La afirmación de su veracidad se

representa colocando el símbolo $|-$ delante de la proposición en cuestión. También distinguió entre un objeto x y el conjunto $\{x\}$ que contiene al elemento x solamente, y entre la pertenencia a un conjunto y la inclusión de un conjunto en otro. Utilizó variables y funciones proposicionales, lo mismo que Pierce, indicando la cuantificación de sus funciones proposicionales, es decir, el dominio de la variable o variables para la que son verdaderas. También introdujo (1879) el concepto de implicación material: A implica B significa que o bien A es verdadera y B verdadera, o A es falsa y B verdadera, o A es falsa y B falsa. Esta interpretación de la implicación es la conveniente para la lógica matemática. Frege se ocupó también de la lógica de relaciones; así, por ejemplo, la relación de orden « a es mayor que b » es muy importante en su obra. Una vez edificada la lógica sobre un sistema explícito de axiomas, procede Frege en su *Grundlagen* a atacar el verdadero problema, el de construir la matemática como extensión de la lógica, expresando los conceptos de la aritmética en términos de conceptos lógicos. Así pues las definiciones y leyes de la aritmética se derivarían de premisas puramente lógicas. Examinaremos esta construcción al hablar de la obra de Russell y Whitehead. Desgraciadamente el simbolismo de Frege era muy complicado y extraño para los matemáticos, con la consecuencia de que su obra fue poco conocida, de hecho, hasta que la descubrió Russell. Por ello resulta bastante irónico el hecho de que, justo en el momento en que el segundo volumen del *Grundgesetze* estaba en prensa, recibiera Frege una carta de Russell en la que le informaba de las paradojas

de la teoría de conjuntos. Al final del volumen 2 (p. 253) escribe Frege: «*Difícilmente puede ocurrirle a un científico algo menos deseable que ver tambalearse los fundamentos de su obra recién terminada. Me he visto en esta posición por una carta de Mr. Bertrand Russell cuando esta obra estaba casi terminada de imprimir.*»

5. La escuela logicista

Dejamos la exposición de las investigaciones sobre los fundamentos de la matemática en el punto en el que la axiomatización de la teoría de conjuntos había suministrado una fundamentación que evitaba las paradojas conocidas y sin embargo servía de base lógica para toda la matemática existente. También señalamos que este planteamiento no resultaba satisfactorio a muchos matemáticos. Todos reconocían que quedaba por demostrar la consistencia del sistema de los números reales y de la teoría de conjuntos; tal consistencia ya no podía seguir siendo una cuestión menor y, por ejemplo, el uso del axioma de elección era muy controvertido. Pero, aparte de estos problemas, estaba la cuestión más general de cuál había de ser la fundamentación correcta de la matemática. El movimiento axiomático de finales del siglo XIX y la axiomatización de la teoría de conjuntos habían procedido sobre la base de que la lógica empleada por la matemática podía considerarse como algo dado y seguro. Pero a comienzos del siglo XX aparecieron varias escuelas de pensamiento que ya no estaban de acuerdo con esta hipótesis. La escuela encabezada por Frege intentó reconstruir la

lógica y construir la matemática dentro de ella. Este plan, como hemos visto, se vio frenado por la aparición de las paradojas, pero no fue abandonado. De hecho, fue adoptado y desarrollado por Bertrand Russell y Alfred North Whitehead. Por su parte, Hilbert, consciente ya de la necesidad de establecer la consistencia, comenzó a formular sus propias ideas sistemáticas sobre la fundamentación de la matemática. Hubo aún otro grupo de matemáticos, conocidos como los intuicionistas, descontentos con los conceptos y demostraciones que se habían introducido en el análisis de finales del siglo XIX. Adoptaban una posición filosófica que no sólo no se podía reconciliar con parte de la metodología del análisis, sino que cuestionaba el papel de la lógica. El desarrollo de estas varias filosofías constituyó la tarea más importante en la fundamentación de la matemática y su resultado fue el de volver a plantear de manera completa la cuestión siempre abierta de la naturaleza de la matemática. Analizaremos por separado cada una de estas tres escuelas de pensamiento principales.

La primera de ellas se conoce como escuela logicista y su filosofía recibe el nombre de logicismo, siendo sus fundadores Russell y Whitehead. Independientemente de Frege, concibieron la idea de que la matemática es derivable exclusivamente de la lógica y, en consecuencia, es una extensión de la lógica. Las ideas básicas aparecen bosquejadas por Russell en sus *Principles of Mathematics* (1903), y se verían desarrolladas en detalle en la obra de Whitehead y Russell *Principia Mathematica* (3 vols., 1910-1913). Dado que los

Principia constituyen la versión definitiva, nos basaremos en ella para nuestra exposición.

Esta escuela parte del desarrollo de la lógica misma, de la que se seguirá la matemática, sin necesidad de ningún axioma específicamente matemático. El desarrollo de la lógica consistirá en establecer para ella un sistema de axiomas, del que se irán deduciendo teoremas que pueden utilizarse en los razonamientos sucesivos. Así pues, las leyes de la lógica se derivarán formalmente de los axiomas. Los Principia parten también de conceptos indefinidos, lo mismo que cualquier teoría axiomática, puesto que no es posible definir todos los términos sin recurrir a un regreso al infinito en las definiciones. Algunos de estos conceptos indefinidos son los de proposición elemental, de función proposicional, la afirmación de la verdad de una proposición elemental, la negación de una proposición y la disyunción de dos proposiciones.

Russell y Whitehead explican estos conceptos, aunque, como indican expresamente, tales explicaciones no forman parte del desarrollo lógico. Por una proposición entienden simplemente una sentencia cualquiera que afirma un estado de cosas o una relación: por ejemplo, Juan es un hombre, las manzanas son rojas, etc. Una función proposicional es una expresión que contiene al menos una variable, de manera que al sustituir un valor por dicha variable se obtiene una proposición. Así, por ejemplo, « x es un entero» es una función proposicional. La negación de una proposición significa «No es cierto que la proposición se verifique», de manera que si p es la proposición «Juan es un hombre», entonces la negación de p ,

representada por $\neg p$, significa «No es cierto que Juan sea un hombre», o «Juan no es un hombre». La disyunción de dos proposiciones p y q , representada por $p \vee q$, significa que «Se verifica p o se verifica q ». El significado de «o» aquí es el que se sobreentiende en la sentencia «Pueden solicitarlo hombres o mujeres»; es decir, los hombres pueden solicitarlo, las mujeres pueden solicitarlo, y unos y otros pueden solicitarlo. En cambio en la sentencia «Esta persona es un hombre o una mujer», la partícula «o» tiene el significado más común de o bien lo uno o lo otro, pero no las dos cosas a la vez. En matemáticas se utiliza el «o» en el primer sentido, aunque a veces el segundo es el único posible. Por ejemplo, «El triángulo es isósceles o el cuadrilátero es un paralelogramo», ilustra el primer sentido, pero también podemos decir «Todo número no nulo es positivo o negativo», sabiendo por la aritmética de los números reales que un número no puede ser a la vez positivo y negativo. Así pues, la afirmación $p \vee q$ significa p y q o bien p y $\neg q$ o bien $\neg p$ y q .

Una relación de la mayor importancia entre proposiciones es la de implicación, es decir, que la verdad de la primera impone la verdad de la segunda. En los Principia la implicación $p \Rightarrow q$ viene definida por $\neg p \vee q$, lo que, a su vez, significa que se verifican $\neg p$ y q o bien $\neg p$ y $\neg q$ o bien p y q . Como ejemplo considérese la implicación «Si X es un hombre, entonces X es mortal»; aquí podrían ocurrir tres cosas:

X no es hombre y X es mortal.

X es un hombre y A es mortal.

X no es hombre y A no es mortal.

Cualquiera de estas posibilidades es admisible; lo único que la implicación prohíbe es que

X es un hombre y A no es mortal.

Algunos de los postulados de los Principia son:

- a. Cualquier cosa implicada por una proposición elemental verdadera es verdadera.
- b. $(p \vee p) \Rightarrow p$
- c. $q \Rightarrow (p \vee q)$
- d. $(p \vee q) \Rightarrow (q \vee p)$
- e. $(p \vee (q \vee r)) \Rightarrow (q \vee (p \vee r))$
- f. La afirmación de p y la afirmación de $p \Rightarrow q$ permite la afirmación de q .

La independencia y la consistencia de estos postulados no puede demostrarse por los métodos usuales, inaplicables aquí. De estos postulados los autores proceden a deducir teoremas de la lógica y por último la aritmética y el análisis. Las reglas conocidas de los silogismos aristotélicos aparecen también como teoremas, por supuesto.

Para ilustrar cómo incluso la lógica misma ha sido formalizada deductivamente, fijémonos en unos pocos teoremas de la primera parte de los *Principia Mathematica*:

$$p \Rightarrow \neg p \Rightarrow \neg p \quad 2.01$$

que es el principio de *reductio ad absurdum*; en palabras, si la hipótesis de p implica que p es falsa, entonces p es falsa.

$$(q \Rightarrow r) \Rightarrow ((p \Rightarrow q) \Rightarrow (p \Rightarrow r)) \quad 2.05$$

que es una forma de silogismo; en palabras, si q implica r , entonces si p implica q , p implica r .

$$p \vee \neg p \quad (2.11)$$

que es el principio de tercio excluso: p es verdadera o falsa.

$$p \Rightarrow \neg(\neg p) \quad (2.12)$$

en palabras, p implica que $\neg p$ es falsa.

$$(p \Rightarrow q) \Rightarrow (\neg q \Rightarrow \neg p) \quad (2.16)$$

Si p implica q , no- q implica no- p .

Las proposiciones constituyen una etapa hacia las funciones proposicionales, que permiten tratar conjuntos por medio de las correspondientes propiedades, en vez de nombrar sus elementos. Así, la función proposicional « x es rojo» representa el conjunto de todos los objetos rojos.

Si los elementos de un conjunto son objetos individuales, entonces la función proposicional que se les aplica se llama de tipo 0. Si los elementos de un conjunto son ellos mismos funciones proposicionales de tipo 0, cualquier función proposicional que se les aplique se llamará de tipo 1 y, en general, cualquier función proposicional cuyas variables sean de tipos menores o iguales que n , se llamará de tipo $n + 1$.

La teoría de los tipos trata de evitar las paradojas que surgen al tratar con colecciones de objetos que contienen un elemento definido solamente en términos de la colección completa. La solución propuesta por Russell y Whitehead a esta dificultad consistía en exigir que «aquello que supone previamente todos los elementos de una colección, no puede ser un elemento de dicha colección». Para llevar a cabo esta restricción en los Principia, imponen que una función (lógica) no puede tener como uno de sus argumentos nada que esté definido en términos de la función misma. A continuación pasan a discutir las paradojas y demuestran que la teoría de tipos las evita.

Sin embargo, la teoría de los tipos nos conduce a clases de afirmaciones que hay que distinguir cuidadosamente según su tipo. Si se intenta construir la matemática dentro de la teoría de los tipos, el desarrollo se hace excesivamente complicado. Por ejemplo, en los Principia dos objetos a y b son iguales si para toda propiedad $P(x)$, $P(a)$ y $P(b)$ son proposiciones equivalentes (cada una implica la otra). Según la teoría de tipos, P puede ser de diferentes tipos, porque puede contener variables de varios niveles además de los

objetos a o b , y así la definición de igualdad debe aplicarse a todos los posibles tipos de P ; en otras palabras, hay infinitas relaciones de igualdad, una para cada tipo de la correspondiente propiedad. Análogamente, un número irracional definido por una cortadura de Dedekind es un objeto de un tipo más alto que un número racional, que a su vez es de un tipo más alto que un número natural, y así el continuo consiste de números de diferentes tipos. Para escapar a esta complejidad introdujeron Russell y Whitehead el llamado «axioma de reducibilidad», que afirma la existencia, para cada función proposicional de cualquier tipo, de una función proposicional equivalente de tipo cero.

Una vez estudiadas las funciones proposicionales, los autores se ocupan de la teoría de clases. Una clase, dicho un poco vagamente, es la colección de objetos que satisfacen alguna función proposicional. Las relaciones se consideran entonces como clases de parejas que satisfacen funciones proposicionales de dos variables; así, « x juzga a y » expresa una relación. Sobre esta base los autores están preparados ya para introducir el concepto de número cardinal.

La definición de número cardinal tiene un gran interés, dependiendo de la relación de correspondencia biunívoca entre clases introducida previamente. Si dos clases están en correspondencia biunívoca se dice que son semejantes; la relación de semejanza es reflexiva, simétrica y transitiva. Todas las clases semejantes tienen una propiedad común, que es su número de elementos. Sin embargo, las clases semejantes bien podrían tener más de una propiedad común.

Para evitar esta dificultad Russell y Whitehead definieron el número de elementos de una clase, como ya había hecho Frege, como la clase de todas las clases semejantes a la clase dada. Así, el número tres es la clase de todas las clases con tres elementos, es decir, de todas las clases x, y, z tales que $x \neq y \neq z$. Dado que la definición de número presupone el concepto de correspondencia biunívoca, podría parecer que tal definición es circular. Los autores hacen notar, sin embargo, que una relación es biunívoca cuando, si x y x' están relacionados con y , entonces x y x' han de ser idénticos, y viceversa, si x está relacionado con y e y' , entonces y e y' han de ser idénticos. Por tanto, es claro que el concepto de correspondencia biunívoca en realidad no hace uso del número 1.

Dados los números cardinales y los naturales, es posible construir ya los sistemas de los números reales y complejos, las funciones y, de hecho, todo el análisis. La geometría puede introducirse a través de los números. Aunque los detalles en los Principia difieren algo de nuestra propia exposición de los fundamentos de los distintos sistemas numéricos y de la geometría (caps. 41 y 42), es fácil ver que tales construcciones son lógicamente posibles sin axiomas adicionales.

Este era, pues, el grandioso programa de la escuela logicista. Lo que hizo por la lógica misma fue notable, como acabamos de esbozar brevemente. Lo que hizo por la matemática, y esto hemos de subrayarlo claramente, fue fundamentar la matemática en la lógica. No era necesario ningún axioma propio de la matemática; la matemática se reducía así a una extensión natural de las leyes y

contenidos de la lógica. Pero los postulados de la lógica y todas sus consecuencias tienen un carácter arbitrario y además formal. Es decir, no tienen contenido propio, sólo tienen forma. Como consecuencia, la matemática tampoco tiene contenido, sino sólo forma. El significado físico que atribuimos a los números o a los conceptos geométricos no forman parte de la matemática. Esto es lo que tenía Russell en la mente cuando dijo que la matemática es la materia en la que nunca sabemos de qué estamos hablando ni si lo que decimos es verdad. En realidad, cuando Russell comenzó a desarrollar su programa a principios de siglo, pensaba (como Frege) que los axiomas de la lógica eran verdades, pero abandonó este punto de vista desde la edición de 1937 de los *Principles of Mathematics*.

El planteamiento logicista ha sido muy criticado. Concretamente, el axioma de reducibilidad levantó considerable oposición, por su carácter bastante arbitrario. Ha sido llamado un feliz accidente y no una necesidad lógica; se ha dicho que tal axioma no tiene lugar en la matemática, y que lo que no puede demostrarse sin él, no puede considerarse como demostrado en absoluto; otros llaman a este axioma un verdadero sacrificio del intelecto. Además, el sistema de Russell y Whitehead nunca se llegó a completar, y es oscuro en numerosos detalles. Más tarde se hicieron muchos esfuerzos por simplificarlo y clarificarlo.

Otra crítica filosófica sería de la posición logicista en su totalidad es la de que, si el punto de vista logicista es correcto, entonces toda la matemática es una ciencia lógico-deductiva puramente formal,

cuyos teoremas se siguen de las leyes del pensamiento exclusivamente. Entonces parece necesario explicar cómo es posible que tal elaboración deductiva de las leyes del pensamiento sirva para representar tan gran diversidad de fenómenos naturales como la acústica, el electromagnetismo, la mecánica, etc. Por otra parte, en la creación de la matemática una intuición sensorial o imaginativa debe suministrar nuevos conceptos, sean derivados de la experiencia o no. De otra manera, ¿cómo podrían surgir conocimientos nuevos? Pero en los Principia todos los conceptos se reducen a conceptos lógicos.

La formalización del programa logicista no parece representar la matemática en ningún sentido real. Nos presenta la cáscara pero no la sustanciosa semilla. Poincaré llegó a decir, maliciosamente (*Les Fondements de la Science*, p. 483), «La teoría logicista no es estéril; engendra contradicciones». Esto no es cierto si uno acepta la teoría de tipos, pero esta teoría, como hemos dicho, resulta artificial. Weyl también atacó al logicismo, diciendo que su compleja estructura «pone a prueba la fuerza de nuestra fe apenas menos que las doctrinas de los primeros Padres de la Iglesia o de los filósofos escolásticos de la Edad Media».

A pesar de esta crítica, muchos matemáticos siguen aceptando la filosofía logicista. La construcción llevada a cabo por Russell y Whitehead tuvo importantes consecuencias en otra dirección, al conducir a una axiomatización completa de la lógica en forma totalmente simbólica, lo cual permitió a la lógica matemática hacer enormes progresos.

6. La escuela intuicionista

Otro grupo de matemáticos, más tarde llamados intuicionistas, se plantearon un enfoque radicalmente distinto de la matemática. Al igual que en el caso del logicismo, la filosofía intuicionista tuvo sus orígenes a finales del siglo XIX, cuando la rigORIZACIÓN del sistema numérico y de la geometría eran tareas de la mayor importancia. El posterior descubrimiento de las paradojas de la teoría de conjuntos vino a impulsar su desarrollo.

El primer intuicionista fue Kronecker, que expresó sus puntos de vista durante las décadas de los 1870 y 1880. Para Kronecker, el rigor impuesto por Weierstrass involucraba conceptos inaceptables, y la obra de Cantor sobre teoría de conjuntos y números transfinitos no era matemática sino misticismo. Kronecker estaba dispuesto a aceptar los números enteros porque son claros a la intuición. Estos eran «obra de Dios»; todo lo demás era obra del hombre y, por tanto, sospechoso. En su ensayo de 1887, «*Über den Zahlbegriff*» (Sobre el concepto de número),⁶¹⁹ mostró cómo algunos tipos de números, por ejemplo los números racionales, se podrían definir en términos de números enteros. Los números fraccionarios como tales eran aceptables por un convenio de notación. Kronecker quería prescindir de la teoría de números irracionales y funciones continuas; su ideal era que todo teorema del análisis se pudiera interpretar en términos de relaciones entre enteros únicamente.

Otra objeción que presentaba Kronecker a muchas partes de la matemática era la de que no daban métodos o criterios

constructivos para determinar en un número finito de pasos los objetos que manejaban. Las definiciones deberían incluir los medios necesarios para calcular el objeto definido en un número finito de pasos, y las demostraciones de existencia deberían permitir el cálculo del número cuya existencia se demostraba con un grado de aproximación arbitrario. Los algebristas solían contentarse con decir de un polinomio $f(x)$ que tenía una raíz racional, en cuyo caso $f(x)$ era reducible, y en caso contrario irreducible. En su «*Festschrift Grundzüge einer arithmetischen Theorie der algebraischen Grössen*» (Elementos de una teoría aritmética de magnitudes algebraicas)⁶²⁰ decía Kronecker, «*La definición de reducibilidad está desprovista de fundamento seguro mientras no se dé un método en virtud del cual se pueda decidir si una función dada es reducible o no*».

En otro contexto, a pesar de que las diversas teorías de números irracionales definían cuándo dos números reales a y b son iguales, o $a > b$ o $b > a$, no daban criterios para determinar qué alternativa se verificaba en un caso dado. Por tanto Kronecker planteó objeciones a tales definiciones; para él sólo eran definiciones en apariencia. La teoría de números irracionales le parecía insatisfactoria en su totalidad, y así un día le dijo a Lindemann, que acababa de demostrar que n era un irracional trascendente: «*¿De qué sirve su bella investigación sobre π ? ¿Para qué estudiar tales problemas, si los números irracionales no existen?*».

Kronecker mismo no hizo mucho por desarrollar la teoría intuicionista, excepto criticar la ausencia de métodos constructivos para determinar objetos matemáticos cuya existencia simplemente

se había demostrado. Intentó reconstruir el álgebra pero no hizo ningún esfuerzo por reconstruir el análisis. Por otra parte, Kronecker realizó una importante obra en aritmética y álgebra que no se ajusta a sus propias exigencias porque, como observó Poincaré,⁶²¹ a veces olvidaba temporalmente su propia filosofía.

En su época Kronecker no encontró partidarios de su filosofía, y durante casi veinticinco años nadie siguió sus ideas. Sin embargo, una vez descubiertas las paradojas, el intuicionismo se revitalizó y se convirtió en un movimiento serio y muy extendido. El siguiente y enérgico defensor fue Poincaré. Ya hemos mencionado su oposición a la teoría de conjuntos porque conducía a las paradojas, y tampoco aceptó el programa logicista para salvar la matemática. Ridiculizó los intentos de basar la matemática en la lógica porque reduciría la matemática a una inmensa tautología. También se burló de la muy artificial (para él) introducción de los números; así, en los Principia se define el 1 como $\hat{a}\{\exists x \times a = i' x\}$, ante lo cual Poincaré decía sarcásticamente que esta era una definición admirable para dársela a gente que nunca hubiera oído hablar del número 1.

En su obra Science et Méthode (Les Fondaments de la Science, p. 480) dice:

El logicismo tiene que ser reconstruido, y uno no puede estar demasiado seguro de lo que se pueda salvar. Es innecesario añadir que sólo se ponen en cuestión el cantorismo y el logicismo; la verdadera matemática, la que sirve a algún fin útil, puede seguir desarrollándose de acuerdo con sus propios principios, sin prestar atención alguna a las tempestades

desencadenadas en el exterior, y continuará sus conquistas usuales, etapa tras etapa, conquistas que son definitivas y que no necesitará abandonar nunca.

Poincaré rechazaba los conceptos que no se pueden definir con un número finito de palabras; por ejemplo, un conjunto construido con ayuda del axioma de elección, no está realmente definido cuando se ha elegido un elemento de cada uno de un número infinito de conjuntos. También sostenía que la aritmética no puede justificarse por ninguna fundamentación axiomática. Nuestra intuición es anterior a tal estructura y, en particular, la inducción completa o matemática procede de una intuición fundamental y no se reduce a un axioma que casualmente es útil en algunos sistemas axiomáticos. Al igual que Kronecker, Poincaré insistía en que todas las definiciones y demostraciones tenían que ser constructivas.

Poincaré estaba de acuerdo con Russell en que el origen de las paradojas estaba en la definición de colecciones o conjuntos que incluían el objeto definido. Así, el conjunto **A** de todos los conjuntos contiene a A como elemento, pero A no puede definirse mientras no esté definido cada elemento de A , y si A es uno de ellos la definición es circular. Otro ejemplo de definición impredicativa es la del valor máximo de una función continua sobre un intervalo cerrado. Tales definiciones eran frecuentes en análisis y especialmente en la teoría de conjuntos.

Otras críticas a la situación lógica de la matemática de la época se desarrollaron y discutieron en un intercambio epistolar entre Borel,

Baire, Hadamard y Lebesgue.⁶²² Borel defendía la afirmación de Poincaré de que los números naturales no pueden fundamentarse axiomáticamente, y criticaba además al axioma de elección porque exige una infinitud no numerable de elecciones simultáneas, lo cual es inconcebible para la intuición. Hadamard y Lebesgue iban más lejos, sosteniendo que incluso una infinitud numerable de elecciones arbitrarias sucesivas no es más intuitiva porque sigue exigiendo una infinitud de operaciones, que son imposibles de concebir como realizadas de manera efectiva. Para Lebesgue todas las dificultades se reducían a saber lo que uno entiende al decir que un objeto matemático existe. En el caso del axioma de elección sostenía que si uno simplemente «piensa» en una manera de elegir, ¿no puede entonces cambiar sus elecciones en el curso del razonamiento? Incluso la elección de un único elemento de un conjunto no vacío planteaba, según Lebesgue, las mismas dificultades; uno tiene que saber que el objeto «existe», lo que significa que debe indicar la elección explícitamente. Así pues, Lebesgue rechazaba la demostración de Cantor de la existencia de números trascendentes. Hadamard indicó que las objeciones de Lebesgue conducían a negar la existencia del conjunto de los números reales, y Borel llegó exactamente a la misma conclusión.

Todas las objeciones de los intuicionistas que hemos mencionado fueron esporádicas y fragmentarias. El fundador sistemático del intuicionismo moderno es Brouwer. Al igual que Kronecker, la mayor parte de su obra matemática, especialmente en topología, no estaba de acuerdo con su filosofía, pero ello no implica la menor

duda sobre la seriedad de su posición. Desde su tesis doctoral, *Sobre los Fundamentos de la Matemática* (1907), comenzó Brouwer a construir su filosofía intuicionista, y desde 1918 en adelante desarrolló y generalizó sus puntos de vista en una serie de artículos publicados en diversas revistas, incluidos los *Mathematische Annalen* de 1925 y 1926.

La posición intuicionista de Brouwer se deriva de una filosofía más general. La intuición fundamental, según él, es la presencia de percepciones en una sucesión temporal. «La matemática surge cuando la cuestión de la “paridad”, que resulta del paso del tiempo, se abstrae de todas las apariciones concretas. La forma vacía que permanece [la relación de n a $n + 1$] del contenido común de todas estas “paridades” se convierte en la intuición original de la matemática, y repetida indefinidamente crea nuevos objetos matemáticos». Así pues, por repetición ilimitada construye la mente el concepto de la sucesión de los números naturales. Esta idea de que los números naturales se derivan de la intuición del tiempo la habían defendido ya Kant, William R. Hamilton en su «*Algebra as a Science of Time*», y el filósofo Arthur Schopenhauer.

Brouwer concibe el pensamiento matemático como un proceso de construcción que edifica su propio universo independiente del de nuestra experiencia y algo así como un modelo libre, restringido únicamente en tanto que está basado en la intuición matemática fundamental. Este concepto intuitivo fundamental no debe ser concebido como una idea indefinida, tal como ocurre en las teorías axiomáticas, sino algo más bien como algo en términos de lo cual

han de ser concebidas intuitivamente todas las ideas indefinidas que aparecen en los diversos sistemas matemáticos, si es que han de servir realmente en el pensamiento matemático.

Brouwer sostiene que «en este proceso constructivo limitado por la obligación de reconocer con cuidado qué tesis son aceptables a la intuición y cuáles no, yace la única fundamentación posible de la matemática». Las ideas matemáticas están en la mente humana previamente al lenguaje, la lógica y la experiencia. Es la intuición, y no la experiencia ni la lógica la que determina la validez y aceptabilidad de las ideas. Hay que recordar, desde luego, que estas afirmaciones sobre el papel de la experiencia hay que tomarlas en el sentido filosófico y no en el histórico.

Para Brouwer los objetos matemáticos los adquirimos por construcción intelectual, donde los números básicos 1, 2, 3,... dan el prototipo de tales construcciones. La posibilidad de la repetición ilimitada de la forma vacía, es decir, la etapa de n a $n + 1$ conduce a los conjuntos infinitos. Sin embargo, el infinito de Brouwer es el infinito potencial de Aristóteles, mientras que la matemática moderna tal como la fundamenta Cantor, por ejemplo, hace un amplio uso de conjuntos infinitos en sentido actual, cuyos elementos están presentes «todos de golpe».

Con respecto a la idea intuicionista de conjunto infinito, dice Weyl, que perteneció a la escuela intuicionista, que

... la sucesión de los números que crece más allá de cualquier nivel ya alcanzado... es una variedad de posibilidades que se abre al infinito; permanece para siempre en el status de

creación, pero no es un dominio cerrado de cosas que existan por sí mismas. El haber convertido ciegamente lo uno en lo otro constituye el verdadero origen de nuestras dificultades, incluyendo las antinomias, un origen de naturaleza más fundamental que el principio del círculo vicioso de Russell, ya indicado. Brouwer nos abrió los ojos y nos mostró hasta dónde la matemática clásica, alimentada por una creencia en lo absoluto que trasciende todas las posibilidades de realización humanas, se remonta más allá de las afirmaciones que pueden pretender tener un significado real y una verdad fundada en la evidencia.

El mundo de la intuición matemática se opone al mundo de las percepciones causales. A este mundo causal, y no a la matemática, corresponde el lenguaje, que sirve allí para el entendimiento en los asuntos comunes; las palabras o expresiones verbales se utilizan para comunicar verdades. El lenguaje sirve para evocar copias de ideas en las mentes de los hombres por medio de símbolos y sonidos. Pero los pensamientos nunca pueden ser simbolizados completamente, y estas observaciones se aplican igualmente al lenguaje matemático, incluido el lenguaje simbólico. Las ideas matemáticas son independientes de la vestidura del lenguaje y, de hecho, mucho más ricas.

La lógica se refiere al lenguaje, presentando un sistema de reglas que permiten deducir unas conexiones verbales de otras y que también intentan comunicar verdades. Sin embargo, estas últimas

verdades no son tales antes de que se tenga experiencia de ellas, y tampoco está garantizado que se pueda alcanzar dicha experiencia. La lógica no es un instrumento seguro para descubrir verdades, y no puede deducir verdades que no se puedan obtener de alguna otra manera. Los principios lógicos expresan la regularidad observada a posteriori en el lenguaje, y se reducen a un instrumento para manipular el lenguaje o una teoría de la representación de dicho lenguaje. Los progresos más importantes de la matemática no se obtienen perfeccionando la forma lógica sino modificando la teoría básica misma. La lógica se apoya en la matemática y no la matemática en la lógica.

Dado que Brouwer no reconoce ningún principio lógico obligatorio a priori, tampoco reconoce la tarea matemática de deducir conclusiones de axiomas. La matemática no está obligada a respetar las reglas de la lógica, y por este motivo las paradojas carecen de importancia incluso si tuviéramos que aceptar los conceptos y construcciones matemáticas en que aparecen. Desde luego, como veremos, los intuicionistas no aceptan todos estos conceptos y demostraciones.

Sobre el papel de la lógica dice Weyl:⁶²³

Según su punto de vista [el de Brouwer] y lectura de la historia, la lógica clásica resultó por abstracción de la matemática de conjuntos finitos y sus subconjuntos... Olvidando este origen limitado se aplicó equivocadamente, más tarde, esa misma lógica para algo anterior a toda la matemática y por encima de ella, y por último se terminó aplicando, sin justificación, a la matemática de los

conjuntos infinitos. En esto consiste la caída y el pecado original de la teoría de conjuntos, por el cual resulta justamente castigada en las antinomias. Lo sorprendente no es que aparecieran tales contradicciones, sino que lo hicieran en una etapa tan avanzada del juego.

En el campo de la lógica hay algunos principios o procedimientos claros, intuitivamente aceptables, que pueden ser utilizados para obtener nuevos teoremas de otros anteriores. Estos principios forman parte de la intuición matemática fundamental. Sin embargo, no todos los principios lógicos son aceptables por la intuición básica, y debería someterse a crítica lo que ha sido aceptado desde la época de Aristóteles. Las antinomias se han producido porque los matemáticos han aplicado libremente esas leyes aristotélicas. Por tanto los intuicionistas tenían que proceder a un análisis acerca de qué principios lógicos son legítimos para que la lógica usual esté de acuerdo con las intuiciones correctas y sea capaz de expresarlas.

Como ejemplo concreto de un principio lógico que se aplica demasiado libremente, menciona Brouwer la ley de tercio excluso. Este principio, que afirma que toda proposición significativa es verdadera o falsa, es fundamental para el método de demostración indirecta. Históricamente surgió por aplicación de razonamientos a subconjuntos de conjuntos finitos, por abstracción. Fue aceptado entonces como un principio independiente a priori y se aplicó injustificadamente a conjuntos infinitos. Mientras que en el caso de conjuntos finitos es posible saber si todos los elementos tienen una propiedad comprobándolo con cada uno, eso no puede hacerse con

los infinitos. Puede ocurrir que sepamos que un elemento del conjunto infinito no posee una propiedad, o puede que por la misma construcción del conjunto sepamos o podamos demostrar que todo elemento tiene esa propiedad. En cualquier caso, no se puede usar la ley de tercio excluso para demostrar que la propiedad se verifica.

Por tanto, si uno demuestra que no todos los elementos de un conjunto infinito tienen una propiedad, entonces Brouwer rechaza que se haya demostrado la conclusión de que existe por lo menos un elemento que no tiene dicha propiedad. Así, por ejemplo, de la negación de que se verifique $a^b = b^a$ para todos los números, los intuicionistas no concluyen que existan a y b tales que $a^b \neq b^a$. Como consecuencia, muchas demostraciones clásicas de existencia no son aceptadas por los intuicionistas. La ley de tercio excluso puede ser utilizada en los casos en que la conclusión pueda alcanzarse en un número finito de etapas; por ejemplo, para decidir la cuestión de si un libro contiene erratas. En cualquier otro caso los intuicionistas rechazan la posibilidad de una decisión.

El rechazo de la ley de tercio excluso da lugar a una nueva posibilidad, la de proposiciones indecidibles. Los intuicionistas sostienen, con respecto a los conjuntos infinitos, que hay una tercera posibilidad, a saber, que una proposición no sea ni demostrable ni refutable. Como ejemplo de tal proposición, definamos k como el lugar que ocupa el primer cero que aparece seguido por las cifras 1, 2, ..., 9 en el desarrollo decimal de π . La lógica aristotélica nos dice que k existe o no existe, y los matemáticos que siguen a Aristóteles pueden razonar a partir de

estas dos posibilidades. Brouwer, en cambio, rechaza todos estos razonamientos debido a que no sabemos si podremos demostrar que k existe o no existe. Así pues, hay cuestiones matemáticas que pueden no ser decididas nunca a partir de las afirmaciones que expresan los axiomas de la matemática. Tales cuestiones pueden parecer decidibles, pero en realidad nuestra base para esperar que esto ocurra se reduce a que tratan de conceptos matemáticos.

Con respecto a los conceptos aceptables como legítimos para una discusión matemática, los intuicionistas insisten en que han de tener definiciones constructivas. Para Brouwer, como para todos los intuicionistas, el infinito existe exactamente en el sentido de que uno puede encontrar siempre un conjunto finito mayor que otro dado. Para discutir cualquier otro tipo de infinito, los intuicionistas exigirían que se dé un método de construir o definir tal infinito en un número finito de pasos. Así, Brouwer rechaza los conjuntos infinitos de la teoría de conjuntos.

La exigencia de constructibilidad es otra que excluye cualquier concepto cuya existencia se haya establecido por un razonamiento indirecto, es decir, argumentando que la no existencia conduce a una contradicción. Aparte del hecho de que la demostración de existencia pueda utilizar la rechazable ley de tercio excluso, tal demostración no es satisfactoria para el intuicionista porque exige una definición constructiva del objeto cuya existencia se está demostrando. Tal definición constructiva debe permitir determinar el objeto con cualquier grado de aproximación deseado en un número finito de pasos. La demostración de Euclides de la

existencia de infinitos números primos (cap. 4, sec. 7) no es constructiva, no permite determinar el n -ésimo primo, y por tanto no es aceptable. Si uno demostrase simplemente la existencia de enteros x, y, z, n tales que $x^n + y^n = z^n$, ningún intuicionista aceptaría la demostración. Por otra parte, la definición de número primo sí es constructiva, pues se puede aplicar para determinar si un número lo es o no en un número finito de pasos. La insistencia en las definiciones constructivas se aplica especialmente a los conjuntos infinitos; por ejemplo, un conjunto construido aplicando el axioma de elección a infinitos conjuntos, no sería aceptable.

Weyl dice de las demostraciones de existencia no constructivas (*Philosophy of Mathematics and Natural Science*, p. 51) que nos informan de que en mundo hay un tesoro sin descubrirnos su localización. La demostración a partir de postulados no puede reemplazar a la construcción sin pérdida de significado y de valor. También señala que adherirse a la filosofía intuicionista significa abandonar muchos teoremas de existencia del análisis clásico, por ejemplo, el teorema de Weierstrass-Bolzano: un conjunto infinito acotado de números reales no tiene necesariamente un punto límite. Para los intuicionistas, si una función de una variable real existe en su sentido, entonces es ipso facto continua. La inducción transfinita y sus aplicaciones al análisis y la mayor parte de la teoría de Cantor se ven condenadas categóricamente. El análisis, dice Weyl, está construido sobre arena.

Brouwer y su escuela no se han limitado a ejercer la crítica, sino que han tratado de construir una nueva matemática basada en las

construcciones que aceptan. Han tenido éxito salvando el cálculo infinitesimal con sus procesos de límite, pero sus construcciones son muy complicadas; también han reconstruido partes elementales del álgebra y la geometría. Al revés que Kronecker, Weyl y Brouwer permiten algunos tipos de números irracionales. Es evidente que la matemática de los intuicionistas tiene que diferir radicalmente de lo que los matemáticos habían aceptado de una manera casi universal antes de 1900.

7. La escuela formalista

La tercera de las principales filosofías de la matemática recibe el nombre de escuela formalista y su creador fue Hilbert, que comenzó a trabajar en esta línea filosófica en 1904. Sus motivaciones en esa época eran las de establecer una base para el sistema numérico sin utilizar la teoría de conjuntos, y a continuación demostrar la consistencia de la aritmética. La consistencia de esta última teoría era un problema abierto de importancia vital, puesto que él mismo había demostrado que la consistencia de la geometría se reducía a la de la aritmética. También trataba de combatir la opinión de Kronecker de que era necesario excluir los números irracionales de la matemática. Hilbert aceptaba sin reservar el infinito actual y elogiaba la obra de Cantor (cap. 41, sec. 9), deseando conservar el infinito, las demostraciones puramente existenciales y los conceptos tales como el de extremo superior, cuya definición parecía ser circular.

Hilbert presentó un artículo sobre sus puntos de vista en el Congreso Internacional de Matemáticos de 1904.⁶²⁴ Durante los siguientes quince años no volvió sobre el tema pero, al fin, movido por el deseo de responder a las críticas intuicionistas del análisis clásico, volvió a los problemas de fundamentos y continuó trabajando en ellos durante el resto de su carrera científica. Durante la década de los años veinte publicó varios artículos clave, y un cierto número de matemáticos fueron adoptando gradualmente sus ideas.

Su filosofía de la etapa madura incluye varias doctrinas. Siguiendo la nueva tendencia de que cualquier fundamentación de la matemática debe tener en cuenta el papel de la lógica, los formalistas sostienen que la lógica tiene que ser tratada simultáneamente con la matemática, que tiene diversas ramas, cada una de ellas con su fundamentación axiomática propia. Esta fundamentación incluye conceptos y principios tanto lógicos como matemáticos. La lógica es un lenguaje simbólico que expresa mediante fórmulas las proposiciones matemáticas y reduce el razonamiento a un proceso formal en el que unas fórmulas se siguen de otras de acuerdo con las reglas deductivas. Todos los símbolos se consideran desprovistos de significado. En su artículo de 1926⁶²⁵ dice Hilbert que los objetos del pensamiento matemático son los símbolos mismos; los símbolos son ahora la esencia y ya no representan objetos físicos idealizados. Las fórmulas pueden expresar afirmaciones intuitivamente significativas, pero estos significados no son parte de la matemática.

Hilbert conservó la ley de tercio excluso por encontrarla necesaria para la construcción del análisis. Decía,⁶²⁶ *«Prohibirle a un matemático usar el principio de tercio excluso es como prohibirle a un astrónomo usar su telescopio o a un boxeador el uso de sus puños»*. Puesto que la matemática trabaja solamente con expresiones simbólicas, a estas expresiones formales se les puede aplicar todas las reglas de la lógica aristotélica. En este nuevo sentido la matemática de conjuntos infinitos resultaría posible. Hilbert esperaba también eliminar las paradojas evitando el uso explícito de la palabra «todo».

Para formular los axiomas lógicos introduce Hilbert un simbolismo para los conceptos y relaciones tales como «y», «o», «no», «existe», etc. Afortunadamente el cálculo lógico (o lógica simbólica) se había desarrollado ya para otros fines y, como Hilbert mismo dice, tenía a mano lo que necesitaba. Todos los símbolos anteriores constituyen el material para construir las expresiones formales o fórmulas.

Para manejar el infinito, Hilbert usa, aparte de los axiomas ordinarios indiscutidos, el axioma transfinito.

$$A(\tau A) \rightarrow A(a)$$

que significa: si un predicado A se verifica para el objeto fiduciario τA , entonces se verifica para todos los objetos a . Así, si suponemos que A expresa «ser corruptible», entonces si Arístides el Justo es fiduciario y corruptible, todo el mundo es corruptible.

Una demostración matemática consistirá en el siguiente proceso: la afirmación de una fórmula; la afirmación de que esta fórmula implica otra; la afirmación de esta segunda fórmula. Una sucesión de etapas tales como esta, en la que las fórmulas afirmadas vienen precedidas por axiomas o conclusiones previas, constituye la demostración de un teorema. Otra operación permisible es la sustituir un símbolo por otro o por un grupo; así pues, las fórmulas se derivan unas de otras aplicando las reglas de manipulación de símbolos en ellas.

Una proposición es verdadera si y sólo si puede obtenerse como la última de una sucesión de proposiciones, cada una de las cuales o bien es un axioma del sistema formal o se deriva de otras anteriores por una de las reglas de deducción. Cualquiera puede comprobar si una proposición dada ha sido obtenida correctamente de una sucesión de proposiciones. Así pues, bajo el punto de vista formalista, verdad y rigor están bien definidos y son objetivos.

Para el formalista, entonces, la matemática propiamente dicha es una colección de sistemas formales, construyendo cada uno su propia lógica a la vez que su matemática, y cada uno de ellos con sus propios conceptos, sus propios axiomas, reglas deductivas tales como la igualdad y sustitución, y sus propios teoremas. El desarrollo de cada uno de estos sistemas deductivos constituye la tarea de la matemática. La matemática se convierte pues, no en una teoría sobre algo, sino en una colección de sistemas formales, en cada uno de los cuales ciertas expresiones formales se obtienen de otras por transformaciones también formales. Esto, en lo que se

refiere a la parte del programa de Hilbert que se ocupa de la matemática propiamente dicha.

Sin embargo, todavía tenemos que preguntarnos si las deducciones están libres de contradicciones. Esto no necesita observarse de modo necesario intuitivamente, pero para demostrar la no contradicción, todo lo que necesitamos demostrar es que nunca se puede llegar a una expresión formal como la $0 = 1$. (Dado que, por un teorema de lógica, cualquier otra proposición se sigue de ésta y de la $0 \neq 1$, podemos limitarnos a ésta).

Hilbert y sus discípulos Wilhem Ackermann (1896-1962), Paul Bernays (1888) y von Neumann desarrollaron, durante los años 1920 y 1930 lo que Hilbert llamó Beweistheorie (Teoría de la Demostración) o metamatemática, un método que pretendía establecer la consistencia de un sistema formal. Hilbert proponía utilizar en la metamatemática únicamente una lógica especial básica y libre de todo tipo de objeciones. Para ello utilizó razonamientos concretos y finitos de un tipo admitido universalmente y muy próximo a los principios intuicionistas. No deberían utilizarse principios controvertidos tales como la demostración de existencia por reducción al absurdo, la inducción transfinita ni el axioma de elección, sino que las demostraciones de existencia deberían ser constructivas. Dado que un sistema formal puede ser indefinido, la metamatemática debe considerar conceptos y cuestiones en los que aparecen sistemas infinitos, al menos potencialmente. No obstante, los métodos de demostración utilizados deben ser totalmente finitarios. No debe haber ninguna

referencia a una cantidad infinita de propiedades estructurales de las fórmulas ni a un número infinito de manipulaciones con ellas.

Ahora bien, la consistencia de una gran parte de la matemática clásica se puede reducir a la de la aritmética de los números naturales (o teoría de números) tal como aparece formulada en los axiomas de Peano, o a la de una teoría de conjuntos lo suficientemente rica como para que permita demostrar dichos axiomas. Por tanto, la consistencia de la aritmética de los números naturales pasó a ser el centro de atención.

Hilbert y su escuela consiguieron demostrar la consistencia de algunos sistemas formales sencillos y creyeron estar a punto de conseguir la meta de demostrar la consistencia de la aritmética y de la teoría de conjuntos. En su artículo «*Über Unendliche*»⁶²⁷ dice Hilbert,

En geometría y en la teoría física la demostración de consistencia se consigue reduciéndola a la consistencia de la aritmética. Este método falla obviamente para la demostración de la consistencia de la aritmética misma. Dado que nuestra teoría de la demostración... hace esta última etapa posible, constituye la piedra angular necesaria en la estructura de la matemática. Y, en particular, lo que ya hemos experimentado dos veces, primero en las paradojas del cálculo infinitesimal y más tarde en la teoría de conjuntos, no puede volver a ocurrir dentro del campo de la matemática.

Pero en este momento entra en escena Kurt Gödel (1906-1978). El primer trabajo importante de Gödel fue su artículo «*Über formal unentscheidbare Sätze der Principia Mathematica und verwandter Systeme I*». ⁶²⁸ En él demuestra Gödel que la consistencia de un sistema que incluya la lógica usual y la teoría de números no puede ser demostrada si se limita uno a conceptos y métodos que puedan ser representables formalmente en el sistema de la teoría de números. Lo que esto significa, en efecto, es que la consistencia de la aritmética no puede ser demostrada por la lógica más restringida admisible en la metamatemática. A propósito de este resultado decía Weyl que Dios existe porque la matemática es consistente, y el diablo existe porque no podemos demostrar dicha consistencia.

El resultado anterior de Gödel es un corolario de su más sorprendente teorema. El teorema más importante de Gödel, o teorema de incompletitud, afirma que si una teoría formal axiomatizable T que incluya la aritmética es consistente, entonces T es incompleta, es decir hay una afirmación S tal que ni S ni $\text{no-}S$ son teoremas de la teoría T . Ahora bien, o S o $\text{no-}S$ es verdadera; tenemos pues una afirmación verdadera que no es demostrable en la teoría. Este resultado se aplica al sistema de Russell-Whitehead, al de Zermelo-Fraenkel y a la axiomatización de Hilbert de la teoría de números. No deja de ser irónico que Hilbert, en su comunicación al Congreso Internacional de Bolonia de 1928 (ver nota 22), hubiera criticado las viejas demostraciones de completitud basadas en la categoricidad (cap. 42, sec. 3) mientras que tenía gran confianza en la completitud de su propio sistema. En realidad las viejas

demostraciones sobre sistemas que contenían los números naturales solamente se habían aceptado como válidas porque la teoría de conjuntos aún no había sido axiomatizada, y se utilizaba de manera intuitiva.

El fenómeno de la incompletitud constituye un importante defecto porque entonces el sistema formal no es adecuado para demostrar todas las afirmaciones que podrían serlo correctamente (sin contradicción) dentro del sistema. Para poner las cosas peor, resulta haber afirmaciones indecidibles pero intuitivamente verdaderas en algún modelo del sistema. La incompletitud no puede remediarse añadiendo S o $\neg S$ como axioma, puesto que lo que demostró Gödel es que cualquier sistema que incluya a la aritmética contiene irremediablemente una proposición indecidible. Así pues, mientras Brouwer sostenía que lo que es intuitivamente verdadero se queda corto con respecto a lo que es matemáticamente demostrable, Gödel viene a demostrar que lo intuitivamente verdadero va más allá de la capacidad de la demostración matemática.

Una de las consecuencias del teorema de Gödel es la de que no hay ningún sistema de axiomas adecuado, no sólo para toda la matemática, sino incluso para casi cualquier rama importante de la matemática, porque cualquier tal sistema axiomático sería incompleto. Existen pues afirmaciones sobre los conceptos del sistema, que no pueden ser demostradas en él, pero cuyo carácter verdadero puede mostrarse por argumentos no formales, mediante la lógica de la metamatemática. Esta consecuencia, la de que existen limitaciones sobre lo que se puede conseguir mediante la

axiomatización, contrasta nítidamente con la creencia de finales del siglo XIX de que la matemática coincide exactamente con la colección de sus ramas axiomatizadas. Este resultado de Gödel asestó un golpe mortal a cualquier pretensión de axiomatizaciones globales. Esta incapacidad del método axiomático no es en sí misma una contradicción, pero resultó sorprendente porque los matemáticos esperaban que cualquier afirmación verdadera podría ser ciertamente demostrable dentro del marco de algún sistema axiomático. Desde luego, los argumentos anteriores no excluyen la posibilidad de nuevos métodos de demostración que permitan ir más allá de la metamatemática de Hilbert.

Hilbert no estaba convencido de que este golpe destruyera su programa. Sostenía que incluso aunque se necesitara usar conceptos exteriores a un sistema formal, aún podrían ser finitarios e intuitivamente concretos, y por tanto aceptables. Hilbert era profundamente optimista; tenía una confianza sin límites en el poder del entendimiento y del razonamiento humanos. En la conferencia que dio en el Congreso Internacional de 1928,⁶²⁹ había dicho, «... *no hay límites para el entendimiento matemático... en la matemática no hay ningún Ignorabimus, sino que siempre podremos contestar a las preguntas significativas... nuestra razón no posee ningún arte secreto, sino que procede de acuerdo con reglas bien definidas y estables que son la garantía de la objetividad absoluta de sus juicios.*» Todo matemático, continúa, comparte la convicción de que cualquier problema matemático definido puede ser resuelto.

Este optimismo le dio valor y fuerza pero le impidió entender que podrían presentarse problemas matemáticos indecidibles.

El programa formalista, tuviera éxito o no, resultaba inaceptable para los intuicionistas. En 1925 Brouwer atacaba enérgicamente a los formalistas.⁶³⁰ Desde luego, decía, el tratamiento axiomático formalista evitará las contradicciones, pero de esta manera no se encontrará nada de valor matemático. Una teoría falsa no es menos falsa porque no conduzca a contradicción, lo mismo que un acto criminal es criminal esté o no condenado por un tribunal. Y añadía también sarcásticamente, «A la pregunta de dónde se encuentra el rigor matemático, las dos partes dan respuestas distintas. El intuicionista dice que en el intelecto humano; el formalista que en el papel.» Weyl atacó también el programa de Hilbert: «La matemática de Hilbert puede ser un bonito juego con fórmulas, más divertido aún que el ajedrez, pero qué relación tiene eso con el conocimiento, dado que se reconoce que sus fórmulas no tienen ningún significado material en virtud del cual pudieran expresar verdades intuitivas.» En defensa de la filosofía formalista hay que subrayar que únicamente con el fin de demostrar la consistencia, completitud y otras propiedades, se reduce la matemática a fórmulas carentes de significado. En cuanto a la matemática globalmente considerada, incluso la mayoría de los formalistas rechaza la idea de que se trate simplemente de un juego, sino que la consideran como una ciencia objetiva.

A su vez Hilbert acusó a Brouwer y a Weyl de intentar arrojar por la borda todo lo que no les convenía, promulgando de manera

dictatorial un embargo,⁶³¹ y calificó el intuicionismo de traición a la ciencia. (Sin embargo, él mismo, en su metamatemática, se limitó a principios lógicos intuitivamente claros.)

8. Algunos desarrollos recientes

Ninguna de las soluciones propuestas a los problemas básicos de los fundamentos —la axiomatización de la teoría de conjuntos, el logicismo, el intuicionismo o el formalismo— lograron el objetivo de establecer un planteamiento de la matemática universalmente aceptable. Los desarrollos realizados desde la obra de Gödel de 1931 no han alterado esencialmente el cuadro. Sin embargo, ha habido unos pocos movimientos y resultados que vale la pena destacar. Algunos matemáticos han tratado de conseguir planteamientos de la matemática de compromiso, utilizando características de dos escuelas básicas. Otros, especialmente Gerhard Gentzen (1909-1945), miembro de la escuela de Hilbert, debilitaron las restricciones sobre los métodos de demostración permitidos en la metamatemática de Hilbert y así consiguió, por ejemplo, demostrar la consistencia de la aritmética y de algunas partes del análisis, utilizando inducción transfinita (inducción sobre los ordinales transfinitos numerables).⁶³²

Entre otros resultados importantes hay dos especialmente notables. En su *The Consistency of the Axiom of Choice and the Generalized Continuum Hypothesis with the Axioms of Set Theory* (1940, ed. rev., 1951) demostró Gödel que si el sistema de axiomas de Zermelo-Fraenkel sin el axioma de elección es consistente, entonces el

sistema obtenido añadiendo esta axioma también es consistente; es decir, el axioma de elección no puede ser refutado. Análogamente, la hipótesis del continuo, de que no hay ningún número cardinal entre \aleph y 2^{\aleph} , es consistente con el sistema de Zermelo-Fraenkel (sin el axioma de elección). En 1963 Paul J. Cohen (1934), profesor de matemáticas de la universidad de Stanford, demostró⁶³³ que estos dos últimos axiomas son independientes del sistema de Zermelo-Fraenkel; es decir, que no pueden ser demostrados dentro de dicho sistema. Además, incluso si se añade el axioma de elección al sistema de Zermelo-Fraenkel, la hipótesis del continuo sigue sin poder ser demostrada. Estos resultados implican que tenemos libertad para construir nuevos sistemas de matemáticas en los que se nieguen uno o los dos de estos controvertidos axiomas.

Todos los progresos realizados desde 1930 dejan abiertos dos importantes problemas: demostrar la consistencia del análisis clásico sin restricciones y de la teoría de conjuntos, y construir la matemática sobre una base intuicionista estricta o determinar los límites de este enfoque. El origen de las dificultades en estos dos problemas es el infinito utilizado tanto en el sentido de los conjuntos infinitos como en los procesos infinitos. Este concepto, que ya creó problemas a los griegos en conexión con los inconmensurables, y que eludieron mediante el método de exhaustión, ha sido tema de discusión desde entonces, lo que hizo decir a Weyl que la matemática es en realidad la ciencia del infinito. La cuestión acerca de la base lógica adecuada para la matemática y el nacimiento en particular del intuicionismo viene a sugerir que, en

un sentido muy general, la matemática ha recorrido un círculo completo. Sus comienzos tuvieron una base intuitiva y empírica. El rigor se convirtió en una necesidad con los griegos y, aunque escasamente logrado hasta el siglo XIX, por un momento pareció alcanzado. Pero lo cierto es que todos los esfuerzos por perseguir el rigor hasta el final, han conducido a un impasse en el que ya no hay acuerdo acerca de qué es lo que realmente significa. La matemática sigue viva y con saludable vitalidad, pero sólo apoyándose en una base pragmática.

Algunos tienen la esperanza de que el aparente callejón sin salida actual tenga solución. El grupo de matemáticos franceses que escribieron con el seudónimo de Nicolás Bourbaki muestra su optimismo con las siguientes palabras:⁶³⁴ «Hace veinticinco siglos que los matemáticos vienen practicando la costumbre de corregir sus errores, viendo así su ciencia enriquecida y no empobrecida; esto les da derecho a contemplar el futuro con serenidad.»

Esté justificado o no el optimismo, el estado actual de la matemática ha sido descrito adecuadamente por Weyl:⁶³⁵ *«El problema de los fundamentos últimos y del significado último de la matemática sigue abierto; no sabemos en qué dirección hallará su solución final, ni siquiera si cabe esperar en absoluto una respuesta final objetiva. El “matematizar” muy bien pudiera ser una actividad creativa del hombre, como el lenguaje o la música, de una originalidad primaria, cuyas decisiones históricas desafíen una racionalización objetiva completa.»*

Bibliografía

- *Becker, Oskar: Grundlagen der Mathematik in geschichtlicher Entwicklung, Verlag Karl Alber, 1956, 317-401.*
- *Beth, E. W.: Mathematical Thought: An Introduction to the Philosophy of Mathematics, Gordon and Breach, 1965.*
- *Bochenski, I. M.: A History of Formal Logic, University of Notre Dame Press, 1962; Chelsea (reprint), 1970. Hay versión castellana, Historia de la lógica formal, Madrid, Gredos, 1985.*
- *Boole, George: An Investigation of the Laws of Thought (1854), Dover (reprint), 1951. Hay versión castellana, Una investigación sobre las leyes del pensamiento, Madrid, Paraninfo, 1982.; The Mathematical Analysis of Logic (1847), Basil Blackwell (reprint), 1948. Hay versión castellana, El análisis matemático de la lógica, Madrid, Cátedra, 1984.; Collected Logical Works, Open Court, 1952.*
- *Bourbaki, N.: Elementos de historia de las matemáticas, Madrid, Alianza, 1976. .*
- *Brouwer, L. E. J.: «Intuitionism and Formalism». Amer. Math. Soc. Bull, 20, 1913/1914, 81-96. Traducción al inglés de la conferencia inaugural de Brouwer como profesor de matemáticas en Amsterdam.*
- *Church, Alonzo: «The Richard Paradox». Amer. Math. Monthly, 41, 1934, 356-361.*
- *Cohen, Paul J., y Reuben Hersh: «Non-Cantorian Set Theory». Scientific American, Dec. 1967, 104-116.*
- *Couturat, L.: La Logique de Leibniz d'après des documents*

inédits, *Alean*, 1901.

- De Morgan, Augustus: *On the Syllogism and Other Logical Writings*, Yale University Press, 1966. Una colección de sus artículos, editada por Peter Heath.
- Dresden, Arnold: «*Brouwer's Contribution to the Foundations of Mathematics*». Amer. Math. Soc. Bull, 30, 1924, 31-40.
- Enriques, Federigo: *The Historie Development of Logic*, Henry Holt, 1929.
- Fraenkel, A. A.: «*The Recent Controversies About the Foundations of Mathematics*». Scripta Mathematica, 13, 1947, 17-36.
- Fraenkel, A. A., y Bar-Hillel, Y.: *Foundations of Set Theory*, North-Holland, 1958.
- Frege, Gottlob: *The Foundations of Arithmetic*, Blackwell, 1953, en inglés y en alemán; también la traducción inglesa sola, Harper and Bros., 1960. Hay versión castellana, Fundamentos de la aritmética, Barcelona, Laia, 1972.; *The Basic Laws of Arithmetic*, University of California Press, 1965.
- Gerhardt, C. I. (ed.): *Die philosophischen Schriften von G. W. Leibniz*,
- 1875-1880, vol. 7.
- Gödel, Kurt: *On Formally Undecidable Propositions of Principia Mathematica and Related Systems*, Basic Books, 1965. Hay versión castellana en K. Gödel, *Obras completas*, Madrid, Alianza, 1981, pp. 55-89.; «What Is Cantor's Continuum Problem?», Amer. Math. Monthly, 54, 1947, 515-525. Hay

versión castellana, *ibid.*, pp. 340-362.; *The Consistency of the Axiom of Choice and of the Generalized Continuum Hypothesis with the Axioms of Set Theory*, Princeton University Press, 1940; *rev. ed.*, 1951.

- Kneale, William, y Martha: *The Development of Logic*, Oxford University Press, 1962. Hay versión castellana, W. y H. Kneale, *El desarrollo de la lógica*, Madrid, Tecnos, 1972.
- Kneebone, G. T.: *Mathematical Logic and the Foundations of Mathematics*, D. Van Nostrand, 1963. Véase especialmente el *apéndice sobre los desarrollos desde 1939*.
- Leibniz, G. W.: *Logical Papers*, traducido y editado por G. A. R. Parkinson, Oxford University Press, 1966.
- Lewis, C. I.: *A Survey of Symbolic Logic*, Dover (reprint), 1960, 1-117.
- Meschkowski, Herbert: *Probleme des Unendlichen, Werk und Leben Georg Cantors*, F. Vieweg und Sohn, 1967.
- Mostowski, Andrzej: *Thirty Years of Foundational Studies*, Barnes and Noble, 1966.
- Nagel, E., y Newman, J. R.: *Gödel's Proof* New York University Press, 1958. Hay versión castellana, *El teorema de Gödel*, Madrid, Tecnos, 1979.
- Poincaré, Henri: *The Foundations of Science*, Science Press, 1946, 448-485. *Esta es una reimpresión en un volumen de Science and Hypothesis, The Value of Science, and Science and Method*. Hay versiones castellanas. *La ciencia y la*

hipótesis, El valor de la ciencia, Ciencia y método, *Madrid, Espasa Calpe*.

- Rosser, J. Barkley: «An Informal Exposition of Proofs of Gödel's Theorems and Church's Theorem». *Journal of Symbolic Logic*, 4, 1939, 53-60.
- Russell, Bertrand: *The Principles of Mathematics*, George Alien and Unwin, 1903; 2.^a ed., 1937. Hay versión castellana, *Los principios de la matemática*, Madrid, Espasa Calpe, 4.^a ed., 1983.
- Scholz, Heinrich: *Concise History of Logic*, Philosophical Library, 1961.
- Styazhkin, N. I.: *History of Mathematical Logic from Leibniz to Peano*, Massachusetts Institute of Technology Press, 1969.
- Van Heijenoort, Jean: *From Frege to Gödel*, Harvard University Press, 1967. Traducciones de artículos clave sobre lógica y fundamentos de la matemática.
- Weyl, Hermann: «Mathematics and Logic». *Amer. Math. Monthly*, 53, 1946, 2-13 = *Ges. Abh.*, 4, 268-279.; — *Philosophy of Mathematics and Natural Science*, *Princeton University Press*, 1949.
- Wilder, R. L.: «The Role of the Axiomatic Method». *Amer. Math. Monthly*, 74, 1967, 115-127.

Notas al fin del libro:

¹ Publicado 1801 = *Werke*, 1.

² *Hist. de VAcad. de Berlín*, 24, 1768, 192 sgs., pub. 1770 = *Œuvres*, 2, 655-726.

³ *Exercices d'analyse et de physique mathématique*, 4, 1847, 84 sgs. = *Œuvres* (1), 10. 312-323

y (2), 14, 93-120.

⁴ Comm. Soc. Gott., 6, 1828 y 7, 1832 = Werke, 2, 65-92 y 93-148, también pp. 165-178.

⁵ Jour. für Math., 28, 1844, 53-67 y 223-245.

⁶ Jour. für Math., 2, 1827, 66-69 = Werke, 6, 233-237.

⁷ Jour. für Math., 27, 1844, 289-310.

⁸ Jour. für Math., 30, 1846, 166-182, p. 172 = Werke, 6, 254-274.

⁹ Jour. für Math., 35, 135-274 (p. 273).

¹⁰ Jour. de Math., 5, 1840, 195-211.

¹¹ Jour. für Math., 9, 1832, 390-393 = Werke, 1, 189-194.

¹² Jour. de Math., 12, 1847, 185-212.

¹³ Jour. für Math., 35, 1847, 319-326, 327-367.

¹⁴ Con la introducción de los números ideales, 2 y 3 dejaron de ser descomponibles, ya que $2 = \alpha^2$ y $3 = \beta_1\beta_2$.

¹⁵ Abh. König. Akad. der Wiss. Berlín, 1858, 41-74.

¹⁶ Jour. für Math., 128, 1905, 45-68.

¹⁷ 4ª ed., 1894 = Werke, 3, 2-222.

¹⁸ H. M. Stark ha demostrado que los valores anteriores de D son los únicos posibles. Véase su «On the problem of unique factorization in complex quadratic fields», Proceedings of Symposia in Puré Mathematics, XII, 41-56, Amer. Math. Soc., 1969.

¹⁹ Jour. für Math., 93, 1882, 1-52 = Werke, 1, 5-71.

²⁰ «Grundzüge einer arithmetischen Theorie der algebraischen Grössen», Jour. für Math., 92, 1882, 1-122 = Werke, 2, 237-387; también publicado por G. Reimer, 1882.

²¹ Werke, 2, 253.

²² Werke, 2, 339

²³ Jour. für Math., 100, 1887, 490-510 = Werke, 3, 211-240.

²⁴ «Die Theorie der algebraischen Zahlkörper», Jabres. der Deut. Afath.-Ver., 4, 1897, 175-546 = Ges. Abh., 1, 63-363.

²⁵ Nouv. Mém. de l'Acad. de Berlín, 1773, 263-312; y 1775, 323 sgs. = (Euvres, 3, 693-795.

²⁶ Mém. de l'Acad. des Sci., París (1), 14, 1813-1815, 177-220 = Œuvres (2), 6, 320-353.

²⁷ Werke, 2, 188-196.

²⁸ Félix Klein en su *Entwicklung* (véase la bibliografía al final de este capítulo), pp. 35-39, desarrolla el bosquejo de Gauss.

²⁹ Para más detalles, véanse los trabajos de Smith y Dickson, incluidos en la bibliografía.

³⁰ Jour. für Math., 37, 1848, 61-94 y 221-254 = Werke, 2, 219-288.

³¹ Publicado en 1863; Dedekind suplemento extensamente las segunda, tercera y cuarta ediciones de 1871, 1879 y 1894.

³² Opuscula Analytica, 2, 1783.

³³ Mém. de l'Acad. des Sci., París, 1785, 465-559, pub. 1788.

³⁴ Théorie des nombres, 2.¹ ed., p. 404.

³⁵ Abh. König. Akad. der Wiss., Berlín, 1837, 45-81 y 108-110 = Werke, 1, 307-342.

³⁶ Abh. König. Akad. der Wiss., Berlín, 1841, 141-161 = Werke, 2, 509-532.

³⁷ Werke, 2, 444-447.

³⁸ Mém. Acad. Sci. St. Peters., 7, 1854, 15-33; también en Jour. de Math. (1), 17, 1852, 366-390 = (Euvres, 1, 51-70).

³⁹ Comm. Acad. Set. Petrop., 9, 1737, 160-188, pub. 1744 = Opera, (1), 14, 216-244.

⁴⁰ Hist. de l'Acad. de Berlín, 17, 1761, 83-106, pub. 1768 = Opera (1), 15, 70-90.

⁴¹ Monatsber. Berliner Akad., 1859, 671-680 = Werke, 145-155.

⁴² En 1914 Godfrey H. Hardy demostró (Comp. Rend., 158, 1914, 1012-1014 = Coll. Papen, 2, 6-9) que una infinidad de ceros de $\zeta(z)$ está sobre la línea $x = 1/2$.

⁴³ Bull. Soc. Math. de France, 14, 1896, 199-220 = (Euvres, 1, 189-210).

⁴⁴ Ann. Soc. Sci. Brnxelles (1), 20 parte II, 1896, 183-256, 281-397.

⁴⁵ Nouv. Mém. de l'Acad. de Berlín, 1773, 121-148, pub. 1775 = Œuvres, 3, 617-658

⁴⁶ Ann. de Math., 11, 1820/21, 205-220.

⁴⁷ Nouv. Mém. de l'Acad. Roy. des Sci., Bruxelles, 2, 1822, 169-202.

- 48 Jour. für Math., 18, 1838, 281-296; y 24, 1842, 83-162, 189-250; los artículos del 1842 se encuentran en su Ges. Werke, 2, 177-308
- 49 Ges. Werke, 2, 197
- 50 Werke, 7, 257-264, 301-302.
- 51 Math. Ann., 68, 1909, 133-140 = Caratheodory, Ges. Math. Schriften, 2, 3-11
- 52 Nachrichten König. Ges. der Wiss. zu Gótt., 1884, 1-13 = Ges. Math. Abh., 2, 327-340.
- 53 Monatsber. Berliner Akad., 1837, 144 = Ges. Werke, 2, 93 y 729-731
- 54 Artículo inédito, Ges. Math. Abh., 2, 344-345
- 55 La demostración de Schwarz se encuentra en Richard Courant y Herbert Robbins, What is Mathematics? Oxford University Press, 1941, 346-349. Una demostración usando el cálculo fue proporcionada por J. F. de Toschi di Fagnano (1715-1797) en el Acta Eruditorum, 1775, 297. Hubo demostraciones geométricas menos elegantes anteriores a la de Schwarz. Hay versión castellana, ¿Qué es la matemática?, Madrid, Aguilar.
- 56 Para una demostración y referencias a demostraciones publicadas véase: H. S. M. Coxeter, Introduction to Geometry, John Wiley and Sons, 1961, 23-25.
- 57 Publicada 1833 = Werke, 1, 461-522.
- 58 Jour. de l'Ecole Poly., 6, 1806, 297-311
- 59 Ann. de Math., 8, 1817/18, 141-55. Este ensayo fue reimpresso en Applications d'analyse et de géométrie de Poncelet (1862-1864), 2, 466-476
- 60 Traite, 1, 48
- 61 Jour. für Math., 4, 1829, 1-71
- 62 Ann. de Math., 16, 1825-1826, 209-231
- 63 Correspondance mathématique et physique, 5, 1829, 6-22
- 64 Correspondance mathématique et physique, 4, 1828, 363-371
- 65 Math. Ann., 6, 1873, 112-145 = Ges. Abh., 1, 311-343. También véase el cap. 38, sec. 3
- 66 Publicado en 1827 = Werke, 1, 1-388
- 67 Jour. für Math., 5, 1830, 1-36 = Wiss. Abh., 1, 124-158
- 68 Jour. für Math., 6, 1830, 107-146 = Wiss. Abh., 1, 178-219
- 69 Annales de Math., 19, 1828, 97-106 = Wiss. Abh., 1, 76-82
- 70 Jour. für Math., 16, 1837, 47-54
- 71 Jour. für Math., 15, 1836, 285-308 = Werke, 3, 329-354
- 72 Jour. für Math., 28, 1844, 97-107 = Ges. Abh., 123-135
- 73 Jour. für Math., 41, 1851, 272-284 = Ges. Abh., 263-278
- 74 Jour. für Math., 28, 1844, 68-96 = Ges. Abh., 89-122
- 75 Jour. für Math., 40, 1850, 237-260 = Werke, 3, 517-542
- 76 Cambridge and Dublin Math. Jour., 4, 1849, 118-132 = Math. Papers, 1, 445-456.
- 77 Math. Ann., 4, 1871, 284-345
- 78 Monatsber. Berlinger Akad., 1864, 246-260, 495-499
- 79 Principia, Libro I, Def. 8, Escolio.
- 80 Véase, por ejemplo, la referencia de Bonola en la bibliografía al final del capítulo.
- 81 Opera, 2, 669-678.
- 82 1ª ed., 1794.
- 83 Mém. de l'Acad. des Sc., Paris, 12, 1833, 367-410.
- 84 Magazin für reine und angewandte Mathematik, 1786, 137-164, 325-358.
- 85 Hist. de l'Acad. de Berlín, 24, 1768, 327-354, pub. 1770 = Opera Mathematica, 2, 245-269.
- 86 Werke, 8, 157-268, contiene todos los anteriores y la carta discutida más abajo.
- 87 Werke, 8, 177.
- 88 Werke, 4, 258.
- 89 Jour. für Math., 17, 1837, 295-320.
- 90 Las traducciones al inglés aparecen en Bonola. Véase la bibliografía al final de este capítulo.
- 91 La traducción al inglés aparece en Bonola. Véase la bibliografía al final de este capítulo.
- 92 Werke, 8, 220-221
- 93 El símbolo $\pi(a)$ es común y así es usado aquí. De hecho, el π de $\pi(a)$ no tiene nada que ver

con el número π .

⁹⁴ La idea de que un ángulo específico pueda estar asociado con una longitud se debe a Lambert.

⁹⁵ Esta es una formulación especial. En su trabajo de 1840, Lobatchevsky proporciona lo que equivale a la forma comúnmente dada en los libros de texto y que también da Gauss, a saber,

$$\tan \frac{1}{2}\pi(x) = e^{-x/k} \quad (a)$$

donde k es una constante, llamada constante de espacio. Con fines teóricos, el valor de la k es indiferente. Bolyai también da la forma (a).

⁹⁶ En el caso de la relación $\tan [\pi(x)/2] = e^{-x/k}$, la elección del valor de la x que debería corresponder a, digamos, $40^\circ 24'$, determinará la k

⁹⁷ Sin embargo, véase George Bruce Halsted, *Amer. Math. Monthly*, 6, 1899, 166-172; y 7, 1900, 247-252.

⁹⁸ *Comm. Soc. Gött.*, 6, 1828, 99-146 = *Werke*, 4, 217-258.

⁹⁹ *Corresp. sur l'Ecole Poly.*, 3, 1814-1816, 162-182.

¹⁰⁰ *Jour. für Math.*, 7, 1831, 1-29.

¹⁰¹ *Giomale dell'Istituto Lombardo*, 9, 1856, 385-398.

¹⁰² *Annali di Mat.* (3), 2, 1868-1869, 101-119.

¹⁰³ *Jour. de l'Ecole Poly.*, 25, 1867, 31-151.

¹⁰⁴ *Jour. für Math.*, 19, 1839, 370-387.

¹⁰⁵ *Werke*, 4, 189-216

¹⁰⁶ *Abb. der Ges. der Wiss. zu Gott.*, 13, 1868, 1-20 = *Werke*, 2.^a ed. 272-287. En los *Collected Mathematical Papers* de W. K. Clifford se encuentra en traducción al inglés. También en *Nature*, 8, 1873, 14-36 y en D. E. Smith, *A source book in mathematics*, 411-425.

¹⁰⁷ *Werke*, 2.^a ed., 1892, 391-404.

¹⁰⁸ Para el significado del símbolo de la llave véase más adelante (19). Riemann no proporcionó estas ecuaciones explícitamente.

¹⁰⁹ Tales superficies ya le eran conocidas a Ferdinand Minding (*Jour. für Math.*, 19, 1839, 370-387, 378-380, en particular), incluyendo la más reciente llamada después pseudoesfera (véase cap. 38, sec. 2). Véase también: Gauss, *Werke*, 8, 265.

¹¹⁰ *Proc. Camb. Phil. Soc.*, 2, 1870, 157-158 = *Math. Papers*, 20-22.

¹¹¹ *Annali di Mat.*, (2), 2, 1868-1869, 232-255 = *Opera Mat.*, 1, 406-429.

¹¹² *Jour. für Math.*, 70, 1869, 46-70 y 241-245 = *Ges. Math. Abh.*, 1, 352 sgs., 378

¹¹³ *Math. Ann.*, 27, 1886, 167-172 y 537-567.

¹¹⁴ *Jour. de l'Ecole Poly.*, 14, 1834, 191-288

¹¹⁵ *Gior. di Mat.*, 2, 1864, 267-282 y en artículos sucesivos en vols. 2 y 3 = *Opere Mat.*, 1, 107-198.

¹¹⁶ *Jour. für Math.*, 36, 1848, 113-134 = *Werke*, 2, 193-216.

¹¹⁷ *Mémoire dell'Accademia delle Scienze dell'Istituto di Bologna*, (2), 8, 1868, 551-590 = *Opere Mat.*, 2, 74-118.

¹¹⁸ *Jour. für Math.*, 19, 1839, 370-387.

¹¹⁹ *Annali di Mat.*, 7, 1866, 185-204 = *Opere Mat.*, 1, 262-280.

¹²⁰ *Gior. di Mat.*, 6, 1868, 248-312 = *Opere Mat.*, 1, 374-405.

¹²¹ *Amer. Math. Soc., Trans.*, 2, 1901, 86-99 = *Ces. Abh.*, 2, 437-448. La prueba y detalles históricos adicionales pueden encontrarse en el Apéndice V del *Grundlagen der Geometrie* de David Hilbert, 7.^a ed., B. G. Teubner, 1930. El teorema presupone que las rectas de la geometría hiperbólica serían las geodésicas de la superficie y las longitudes y los ángulos serían las longitudes y los ángulos euclídeos en la superficie.

¹²² *Nachrichten König. Ges. der wiss. zh Gött.*, 1899, 44-55; *Math. Ann.*, 53, 1900, 81-112; y 54, 1901, 505-517.

¹²³ *Nouvelles Anuales de Mathématiques*, 12, 1853, 56-66 = *CEuvres*, 2, 6-15.

¹²⁴ La razón doble es ella misma un número complejo. El coeficiente $i/2$ asegura que un ángulo recto tenga como medida $\pi/2$. La forma en que se calculan tales razones dobles pueden encontrarse en textos sobre geometría proyectiva. Véase, por ejemplo, William C. Graustein:

Introduction to Higher Geometry, MacMillan, 1933, cap. 8.

¹²⁵ *Phil. Trans.*, 149, 1859, 61-91 = *Coll. Math. Papers*, 2, 561-606.

¹²⁶ *Nachrichten König. Ges. der Wiss. zu Gött.*, 1871, 419-433 = *Ges. Math. Abh.*, 1, 244-253.

¹²⁷ *Math. Ann.*, 4, 1871, 573-625; y 6, 1873, 112-145 = *Ges. Math. Abh.*, 1, 254-305, 311-343.

¹²⁸ *Math. Ann.*, 4, 1871, 604. Véase también *Math. Ann.*, 6, 1873, 125; y *Math. Ann.*, 37, 1890, 554-557.

¹²⁹ *Math. Ann.*, 7, 1874, 549-557; 9, 1876, 476-482 = *Ges. Math. Abh.*, 2, 63-77.

¹³⁰ Véanse las referencias en la nota 11.

¹³¹ *Annali di Mat.* (2), 2, 1868-1869, 232-255 = *Opere Matematiche*, 1, 406-429.

¹³² *Annali di Mat.*, 7, 1886, 185-204 = *Opere Mat.*, 1, 262-280; *Gior. di Mat.*, 6, 1868, 284-312 = *Opere Mat.*, 1, 374-405.

¹³³ *Acta Math.*, 1, 1882, 1-62 = *CEuvres*, 2, 108-168; véase pág. 8 y pág. 52 del artículo.

¹³⁴ *Esta forma, atribuida a Poincaré, es parecida a una que él dio en el Bull. Soc. Math. de France*, 15, 1887, 203-216 = *CEuvre*, 11, 79-91. El modelo descrito aquí parece haber sido dado primero por Joseph Wellstein (1869-1919) en H. Weher y J. Wellstein, *Enzyklopadie der Elementar-Mathematik*, 2, 1905, 39-81.

¹³⁵ *Math. Anrt.*, 43, 1893, 63-100 = *Ges. Math. Abh.*, 1, 460-497. Una traducción al inglés puede encontrarse en el *N. Y. Math. Soc. Bull.*, 2, 1893, 215-249.

¹³⁶ Klein no señaló este subgrupo.

¹³⁷ Klein sí habla, en el caso de la geometría diferencial, del grupo de transformaciones que dejan invariante la expresión ds^2 . Esto conduce a los invariantes diferenciales (*Ges. math. Abh.*, 1, 487).

¹³⁸ *Nachrichten König. Ges. der Wiss. zu Gött.*, 15, 1868, 193-221 = *Wiss. Abh.*, 2, 618-639.

¹³⁹ *Theorie der Transformationsgruppen*, 3, 437-543, 1893.

¹⁴⁰ *Collected Math. Papers*, 11, 429-459.

¹⁴¹ *Bu. ll. Soc. Math. de Frunce*, 15, 1887, 203-216 = (*Euvres*, 11, 79-91. Expresó esta opinión nuevamente en un artículo «Les Géométries non-euclidiennes» en la *Revue Générale des Sciences*, 2, 1891, n. 23. Hay una traducción inglesa en *Nature* 45, 1892, 404-407. Véase también su *Science and Hypothesis*, cap. 3, en *The Foundations of Science*, the Science Press, 1946.

¹⁴² Cayley usó el término «geometría descriptiva» para la geometría proyectiva.

¹⁴³ *Die Entwicklung der Mathematik in den letzten Jahrhunderten* (La Evolución de las Matemáticas durante los últimos siglos), 1869; 2.^a ed., 1884.

¹⁴⁴ *Jour. für Math.*, 28, 1844, 68-96 = *Ges. Abh.*, 89-122.

¹⁴⁵ *Cambridge Mathematical Journal*, 3, 1841, 1-20; y 3, 1842, 106-119

¹⁴⁶ *Coll. Math. Papers*, I, 273.

¹⁴⁷ *Coll. Math. Papers*, 2, 4, 6, 7, 10.

¹⁴⁸ *Jour. für Math.*, 27, 1844, 89-106, 319-321

¹⁴⁹ *Jour. für Math.*, 55, 1858, 97-191; y 62, 1863, 281-345.

¹⁵⁰ *Phil. Trans.*, 146, 1856, 101-126 = *Coll. Math. Papers*, 2, 250-275.

¹⁵¹ *Jour. für Math.*, 69, 1868, 323-354.

¹⁵² *Math. Ann.*, 2, 1870, 227-280.

¹⁵³ R. Clebsch y F. Lindemann, *Vorlesungen über Geometrie*, I, 1876, p. 291.

¹⁵⁴ *Math. Ann.*, 1, 1869, 56-89, 90-128.

¹⁵⁵ Clebsch-Lindemann, p. 288.

¹⁵⁶ *Math. Ann.*, 17, 1880, 217-233.

¹⁵⁷ *Jour. für Math.*, 100, 1887, 223-230.

¹⁵⁸ *Math. Ann.*, 30, 1887, 15-29 = *Ges. Abh.*, 2, 102-116.

¹⁵⁹ *Math. Ann.*, 33, 1889, 223-226 = *Ges. Abh.*, 2, 162-164.

¹⁶⁰ *Math. Ann.*, 36, 1890, 473-534 = *Ges. Abh.*, 2, 199-257, y en artículos subsiguientes hasta 1893.

¹⁶¹ *Nachrichten König. Ges. der Wiss. zu Gött.*, 1899, 240-242.

¹⁶² *Phil. Trans.*, 154, 1864, 579-666 = *Coll. Math. Papers*, 2, 376-479, p. 380.

- ¹⁶³ *Jour. für Math.*, 134, 1908, 23-90.
- ¹⁶⁴ *Jour. für Math.*, 139, 1911, 118-154.
- ¹⁶⁵ *Theorie der Kreisverwandtschaft (Teoría de la Inversión)*, Abb. König. Sach. Ges. der Wiss., 2, 1855, 529-565 = *Werke*, 2, 243-345.
- ¹⁶⁶ *Jour. de Math.*, 10, 1845, 364-367.
- ¹⁶⁷ *Jour. de Math.*, 12, 1847, 265-290.
- ¹⁶⁸ *Gior. di Mat.*, 1, 1863, 305-311 = *Opere*, 1, 54-61; y 3, 1865, 269-280, 363-376 = *Opere*, 2, 193-218.
- ¹⁶⁹ *Math. Ann.*, 3, 1871, 165-227, p, 167 en particular.
- ¹⁷⁰ *Jour. für Math.*, 73, 1871, 97-110.
- ¹⁷¹ *Atti Accad. Torino*, 36, 1901, 861-874.
- ¹⁷² *Jour. für Math.*, 63, 1864, 189-243.
- ¹⁷³ *Jour. für Math.*, 63, 1864, 189-243.
- ¹⁷⁴ *Jour. für Math.*, 64, 1865, 43-65.
- ¹⁷⁵ *Jour. für Math.*, 64, 1865, 98-100.
- ¹⁷⁶ Si los puntos múltiples (singulares) son de orden r . entonces el género p de una curva C es $(n - 1)(n - 2)12 - (1/2) 2r(r - 1)$, donde la sumatoria se extiende sobre todos los puntos múltiples. El género es un concepto más refinado.
- ¹⁷⁷ *Math. Atin.*, 9, 1876, 163-165.
- ¹⁷⁸ *Jour. für Math.*, 54, 1857, 115-155 = *Werke*, 27 ed., 88-142.
- ¹⁷⁹ *Jour. für Math.*, 64, 1865, 43-65.
- ¹⁸⁰ *Jour. für Math.*, 64, 1865, 210-270.
- ¹⁸¹ *Proc. Lon. Math. Soc.*, 4, 1871-1873, 347-352 = *Coll. Math. Papers*, 8, 181-187.
- ¹⁸² *Jour. für March.*, 65, 1866, 269-283.
- ¹⁸³ *Math. Ann.*, 21, 1883, 141-218 = *Ges. math. Abh.*, 3, 630-710.
- ¹⁸⁴ *Bull. Soc. Math. de France*, 11, 1883, 112-125 = *CEuvres*, 4, 57-69.
- ¹⁸⁵ *Acta Math.*, 31, 1908, 1-63 = *(Euvres*, 4, 70-139.
- ¹⁸⁶ *Math. Ann.*, 67, 1909, 145-224.
- ¹⁸⁷ «Über die algebraischen Funktionen und ihre Anwendung in der Geometrie». *Math. Ann.*, 7, 1874, 269-310.
- ¹⁸⁸ *Jour. für Math.*, 93, 1882, 271-318.
- ¹⁸⁹ *Jour. de l'Ecole Poly.*, Cahier 52, 1882, 1-200 = *(Euvres*, 3, 261-455.
- ¹⁹⁰ *Nachrichten König. Ges. der Wiss. zu Gótt.*, 1871, 267-278.
- ¹⁹¹ Véase también Noether, *Math. Ann.*, 9, 1876, 166-182; y 23, 1884, 311-358.
- ¹⁹² *Jour. für Math.*, 91, 1881, 301-334 = *Werke*, 2, 193-236.
- ¹⁹³ Reproducido en un apéndice a una edición francesa (1884) del *Higher Plane Curves* de G. Salmón y en el *Traité d'analyse* de E. Picard, 2, 1893, 364 y sgs. = *(Euvres*, 4, 1-93.
- ¹⁹⁴ *Jour. für Math.*, 91, 1881, 301-334 y 92, 1882, 1-122 = *Werke*, 2, 193-387.
- ¹⁹⁵ «Theorie der algebraischen Funktionen einer Veränderlichen». *Jour. für Math.*, 92, 1882, 181-290 = *Werke* de Dedekind, I, 238-350.
- ¹⁹⁶ «Über die Theorie der algebraischen Formen». *Math. Ann.*, 36, 1890, 473-534 = *Ges. Abh.*, 2, 199-257.
- ¹⁹⁷ *Comp. Rend.*, 67, 1868, 1238-1239.
- ¹⁹⁸ *Math. Ann.*, 2, 1870, 293-316.
- ¹⁹⁹ *Math. Ann.*, 4, 1871, 21-49.
- ²⁰⁰ *Phil. Trans.*, 159, 1869, 201-229 = *Coll. Math. Papers*, 6, 329-358; y *Math. Ann.*, 3, 1871, 526-529 = *Coll. Math. Papen*, 8, 394-397.
- ²⁰¹ *Math. Ann.*, 4, 1871, 21-49.
- ²⁰² *Math. Ann.*, 8, 1875, 495-533.
- ²⁰³ *Jour. de Math.* (5), 5, 1899, 5-54, y artículos posteriores.
- ²⁰⁴ *Annali di Mat.* (2), 26, 1897, 219-253.
- ²⁰⁵ Las variedades involucradas no pueden ser caracterizadas, ni siquiera topológicamente.
- ²⁰⁶ *CEuvres*, 2, 263-265.

- 207 1821, *Œuvres* (2), III.
- 208 1823, *Œuvres* (2), IV, 1-261.
- 209 1829, *Œuvres* (2), IV, 265-572.
- 210 1797; 2.¹ ed. 1813 = *Œuvres*, 9.
- 211 1801; 2.^a ed. 1806 = *Œuvres*, 10.
- 212 3 vols., 1.^a ed., 1797-1800; 2.^a ed., 1810-1819.
- 213 Traducción en inglés, p. 430, Dover (reimp.), 1955.
- 214 *Repertorium der Physik*, 1, 1837, 152-174 = *Werke*, 1, 135-160.
- 215 *Jour. für Math.*, 4, 1829, 157-169 = *Werke*, 1, 117-132.
- 216 Considérese $y = \text{sen}(1/x)$ para $x \neq 0$ e $y = 0$ para $x = 0$. Esta función va de todos los valores desde uno que tome para un valor negativo de x hasta uno que tome para un valor positivo de x . Sin embargo, no es continua en $x = 0$.
- 217 *Jour. für Math.*, 71, 1870, 353-365.
- 218 *Jour. für Math.*, 74, 1872, 172-188.
- 219 Ann. de l'École Norm. Sup. (3), 12, 1895, 9-55.
- 220 *Acta Math.*, 19, 1895, 1-61.
- 221 1823, *Œuvres* (2), 4, 22.
- 222 Lacroix ya había definido de esta manera dy en la primera edición de su *Traite*.
- 223 *Schriften*, 1, Praga, 1930. Fue editado y publicado por K. Rychlik, Praga, 1930.
- 224 En 1922 Rychlik demostró que la función era no diferenciable en todas partes. Véase Gerhard Kowalewski, «Über Bolzanos nichtdifferenzierbaré stetige Funktion», *Acta Math.*, 44, 1923, 315-319. Este artículo contiene una descripción de la función de Bolzano.
- 225 Abh. der Ges. der Wiss. zu Gött., 13, 1868, 87-132 = *Werke*, 227-264.
- 226 *Bull. des Sci. Math.* (2), 14, 1890, 142-160.
- 227 *Werke*, 2, 71-74.
- 228 *Jour. für Math.*, 79, 1875, 21-37.
- 229 Otros ejemplos y referencias se pueden encontrar en E. J. Townsend, *Functions of Real Variables*, Henry Holt, 1928, y en E. W. Hobson, *The Theory of Functions of a Real Variable*, 2, cap. 6, Dover (reimp.), 1957.
- 230 *Résumé*, 81-84 = *Œuvres* (2), 4, 122-127.
- 231 Por brevedad se usa Δx_i para los subintervalos y sus longitudes.
- 232 Ann. de l'École Norm. Sup. (2), 4, 1875, 57-112.
- 233 *Gior. di Mat.*, 19, 1881, 333-372.
- 234 Publicado en el *Cours de calcul différentiel et intégral* de Serret, 1, 1868, 17-19.
- 235 *Proc. Lon. Math. Soc.*, 6, 1875, 140-153 = *Coll. Papen*, 2, 86-100.
- 236 *Mémoire* de 1814; véase, en particular, p. 394 de *Œuvres* (1), 1.
- 237 *Zeit. für Math. und Phys.*, 21, 1876, 224-227.
- 238 *Zeit. für Math. und Phys.*, 23, 1878, 67-68.
- 239 *Jour. für Math.*, 94, 1883, 273-290.
- 240 *Comm. Soc. Gött.*, 2, 1813 = *Werke*, 3, 125-162 y 207-229.
- 241 *Werke*, 3, 129.
- 242 *Werke*, 3, 156.
- 243 *Jour. de l'École Poly.*, 19, 1823, 404-509.
- 244 La noción correcta del límite de una sucesión fue dada por Wallis en 1695 (*Opera*, 1695, 1, 382), pero no fue adoptada.
- 245 *Exerciccs de mathématiques*, 1, 1826, 5 = *Œuvres*, (2), 6, 38-42.
- 246 1823, *Œuvres* (2), 4, p. 237.
- 247 *Exercices de mathématiques*, 2, 1827 = *Œuvres* (2), 7, 160.
- 248 *Œuvres*, 2, 259.
- 249 *Jour. für Math.*, 1, 1826, 311-339 = *Œuvres*, 1, 219-250.
- 250 *Œuvres*, 1, 224.
- 251 La serie (2) es la expansión de Fourier de $x/2$ en el intervalo $-n < x < n$. Por tanto, la serie representa la función periódica que vale $x/2$ en cada intervalo de longitud $2n$. Entonces la serie

converge a $n!2$ cuando x tiende a $(2n + 1) n$ por la izquierda y la serie converge a $-n!2$ cuando x tiende a $(2n + 1)$ por la derecha.

²⁵² Trans. Camb. Phil. Soc., 8, 1848, 533-583 = Math. and Phys. Papers, 1, 236-313.

²⁵³ Abh. der Bayer. Akad. der Wiss., 1847/1849, 379-394.

²⁵⁴ Papers, 1, 242, 255, 268 y 283.

²⁵⁵ Comp. Rend., 36, 1853, 454-459 = Œuvres (1), 12, 30-36.

²⁵⁶ Werke, 1, 67-85.

²⁵⁷ Jour. für Math., 71, 1870, 353-365.

²⁵⁸ Sitzungsbber. Akad. Wiss. zu Berlin, 1855, 633-639, 789-905 = Werke, 3, 1-37.

²⁵⁹ Abh. König. Akad. der Wiss., Berlin, 1837, 45-81 = Werke, 1, 313-342 = Jour. de Math., 4, 1839, 393-422.

²⁶⁰ Jour. für Math., 4, 1829, 157-169 = Werke, 1, 117-132.

²⁶¹ Abh. der Ges. der Wiss. zu Gött., 13, 1868, 87-132 = Werke, 227-264.

²⁶² Jour. für Math., 71, 1870, 353-365.

²⁶³ Jour. für Math., 72, 1870, 139-142 = Ges. Abh., 80-83

²⁶⁴ Jour. für Math., 73, 1871, 294-296 = Ges. Abh., 84-86.

²⁶⁵ Math. Ann., 5, 1872, 123-132 = Ges. Abh., 92-102.

²⁶⁶ Se puede encontrar detalles en E. W. Hobson, *The Theory of Functions of a Real Variable*, vol. 2, 656-698.

²⁶⁷ Nachrichten König. Ges. der Wiss. zu Gött., 1873, 571-582.

²⁶⁸ Abh. der Bayer. Akad. der Wiss., 12, 1876, 117-166.

²⁶⁹ Math. Ann., 22, 1883, 260-268.

²⁷⁰ Comp. Rend., 92, 1881, 228-230 = Œuvres, 4, 393-395 y *Cours d'analyse*, 2, 1.^a ed., 1882, cap. V.

²⁷¹ Cours d'analyse, 2.a ed., 1893, 1, 67-72

²⁷² *L'Enseignement mathématique*, 11, 1899, 157-162 = Œuvres, 2, 129-134.

²⁷³ *Théorie générale des fonctions*, 1887, 61.

²⁷⁴ Œuvres, 2, 256.

²⁷⁵ *Mém. des sav. étrangers*, 1, 1827, 3-312; véase Œuvres (1), 1, 238, 277, 286.

²⁷⁶ Report on the Recent Progress and Present State of Certain Branches of Analysis, *Bnt. Assn. for Adv. of Science*, 3, 1833, 185-352.

²⁷⁷ Trans. Camb. Philo. Soc., parte II, 1844, 182-203, pub. 1849.

²⁷⁸ Londres, 1842, p. 566.

²⁷⁹ *Aufsätze aus dem Gebiet der höherer Mathematik* (Ensayos en el Dominio de las Matemáticas Avanzadas, 1823).

²⁸⁰ Werke, 8, 177.

²⁸¹ Werke, 8, 201.

²⁸² Comp. Rend., 18, 1844, 910-911, y *Jour. de Math.* (1), 16, 1851, 133-142.

²⁸³ Comp. Rend., 77, 1873, 18-24, 74-79, 226-233, 285-293, y Œuvres, 2, 150-181.

²⁸⁴ *Matk. Ann.*, 20, 1882, 213-225.

²⁸⁵ *Essays*, 40.

²⁸⁶ *Trans. Royal Irish Academy*, 17, 1837, 293-422, y *Math. Papers*, 3, 3-96.

²⁸⁷ *Math. Ann.*, 21, 1883, 566.

²⁸⁸ *Revue des Sociétés Savants*, 4, 1869, 280-289.

²⁸⁹ *Jour. für Math.*, 74, 1872, 172-188.

²⁹⁰ Continuidad y números irracionales, 1872, en *Werke*, 3, 314-334.

²⁹¹ *Math. Ann.*, 5, 1872, 123-132 = *Ges. Abh.*, 92-102.

²⁹² *Math. Ann.*, 21, 1883, 545-591 = *Ges. Abh.*, 165-204.

²⁹³ 1886, 1, 109-119.

²⁹⁴ *Theorie der complexen Zahlensystem*, 1867, 46-47.

²⁹⁵ La naturaleza y significado de los números, en *Werke*, 3, 335-391.

²⁹⁶ *Opere scelte*, 2, 20-56, y *Rivista di Matematica*, 1, 1891, 87-102, 256-257, en *Opere scelte*, 3, 80-109.

- ²⁹⁷ *Jahres. der Deut. Math.-Verein.*, 8, 1899, 180-184; este artículo no está en las *Gesammelte Abhandlungen* de Hilbert. Aparece en sus *Grundlagen der Geometrie*, 7.^a edición, Apéndice 6.
- ²⁹⁸ Pág. 62 de la edición francesa de *Die allgemeine Funktionentheorie*, 1882.
- ²⁹⁹ *Werke*, 8, 216.
- ³⁰⁰ *Ges. Abh.*, 115-356.
- ³⁰¹ *Math. Ann.*, 5, 1872, 122-132 = *Ges. Abh.*, 92-102.
- ³⁰² *Jour. für Math.*, 77, 1874, 258-262 = *Ges. Abh.*, 115-118.
- ³⁰³ *Math. Ann.*, 46, 1895, 481-512 = *Ges. Abh.*, 283-356, en particular 294-295.
- ³⁰⁴ *Jahres. der Deut. Math.-Verein.*, 1, 1890/1891, 75-78 = *Ges. Abh.*, 278-281.
- ³⁰⁵ *Briefwechsel Cantor-Dedekind*, 34.
- ³⁰⁶ *Jour. für Math.*, 84, 1878, 242-258 = *Ges. Abh.*, 119-133.
- ³⁰⁷ Pág. 167 de la edición francesa (1887) de su *Die allgemeine Funktionentheorie*.
- ³⁰⁸ *Math. Ann.*, 46, 1895, 481-512, y 49, 1897, 207-246 = *Ges. Abh.*, 282-351; se puede encontrar una traducción al inglés de estos dos artículos en Georg Cantor, *Contributions to the Founding of a Theory of Transfinite Numbers*, Dover (reimpresión), sin fecha.
- ³⁰⁹ 1883, *Ges. Abh.*, 165.
- ³¹⁰ Cantor, *Ges. Abh.*, 278-280
- ³¹¹ *Math. Ann.*, 21, 1883, 545-586 = *Ges. Abh.*, 165-204.
- ³¹² *Math. Ann.*, 59, 1904, 514-516.
- ³¹³ *Math. Ann.*, 65, 1908, 107-128.
- ³¹⁴ *Ges. Abh.*, 445-448.
- ³¹⁵ Rendiconti del Circolo Matematico di Palermo, 11, 1897, 154-164 y 260.
- ³¹⁶ Proceedings of the Fourth Internat. Cong. of Mathematicians, Roma, 1908, 167-182. Bull. des Sci. Math. (2), 32, 1908, 168-190 = *Hay un extracto de CEuvres*, 5, 19-23.
- ³¹⁷ *Math. Ann.*, 95, 1926, 170 = *Grundlagen der Geometrie*, 7.^a edición, 1930, 274.
- ³¹⁸ *Math. Ann.*, 95, 1926, 167 = *Grundlagen der Geometrie*, 7.^a edición, 1930, 270. El artículo «Über das Unendliche», del que hemos tomado las citas de más arriba, aparece también en francés en *Acta Math.*, 48, 1926, 91-122. No está incluido en las *Gesammelte Abhandlungen* de Hilbert.
- ³¹⁹ *Werke*, 8, 222.
- ³²⁰ *Werke*, 8, 196.
- ³²¹ *Werke*, 8, 193-195 y 200.
- ³²² *Ann. de Math.*, 9, 1818-1819, 1.
- ³²³ *I principii di Geometria*, Turin, Fratelli Bocca, 1889 = *Opere Scelte*, 2, 56-91.
- ³²⁴ *Memorie della Reale Accademia delle Scienze di Torino*, (2), 48, 1899, 1-62.
- ³²⁵ *Amer. Math. Soc. Trans.*, 3, 1902, 142-158.
- ³²⁶ *Math. Ann.*, 55, 1902, 265-292.
- ³²⁷ Cambridge University Press, 1906.
- ³²⁸ *Amer. J. of Math.*, 1908, 347-378.
- ³²⁹ *Memorie della Reale Accademia delle Scienze di Torino*, (2), 1899, 173-222.
- ³³⁰ *Monist*, 9, 1898, 38.
- ³³¹ Estrictamente, usa un conjunto más limitado de números reales.
- ³³² *Amer. Math. Soc. Transi.*, 5, 1904, 343-384.
- ³³³ *Amer. Math. Soc. Transi.*, 3, 1902, 264-279.
- ³³⁴ 1904, 212-247.
- ³³⁵ *Annals of Math.*, (2), 18, 1916-1917, 31-44.
- ³³⁶ *Math. Ann.*, 61, 1905, 173-184.
- ³³⁷ *Math. Ann.*, 53, 1900, 404-439.
- ³³⁸ *Math. Ann.*, 55, 1902, 465-478.
- ³³⁹ *Cours d'Analyse*, 1.^a ed., vol. 3, 1887, 593; 2.^a ed., vol. 1, 1893, 90.
- ³⁴⁰ 1.^a ed., p. 593; 2.^a ed., p. 98.
- ³⁴¹ *Amer. Math. Soc. Trans.*, 6, 1905, 83-98, y 14, 1913, 65-72.
- ³⁴² *Math. Ann.* 36, 1890, 157-160 = *Opere Scelte*, 1, 110-115.

- 343 Jour. für Math., 86, 1879, 263-268.
 344 Jahres. der Deut. Math.-Verein, 82, 1900, 121-125.
 345 Amer. Math. Soc. Trans., 1, 1900, 72-90.
 346 *Bul. des Sci. Math.*, (2), 21, 1897, 257-266.
 347 *Math. Ann.*, 38, 1891, 459-460 = *Ges. Abh.*, 3, 1-2.
 348 *Math. Ann.*, 50, 1898, 586.
 349 *Ges. Math. Abh.*, 2, 309-311.
 350 Puede encontrarse este ejemplo en J. Pierpont, *The theory of functions of a real variable*, Dover (reimpresión), 1959, vol. 2, p. 26.
 351 *Acta Math.*, 30, 1906, 145-176.
 352 *Comp. Rendus du Deuxième Congres Intemat. des Math.*, 1900, pub. 1902, 121-122.
 353 *Amer. Math. Soc. Bull.*, 11, 1904-1905, 417; ver también cap. 40, sec. 7 y cap. 41, sec. 5 y 9.
 354 «Todo lo que trasciende la geometría trasciende nuestra comprensión».
 355 *Abh. Math. Seminar Hamburger Univ.*, 1, 1922, 157-177 = *Ges. Abh.*, 3, 157-177.
 356 *Math. Ann.*, 78, 1918, 405-415 = *Ges. Abh.*, 3, 145-156.
 357 *Ges. Werke*, 1, 172.
 358 *Comp. Rend.*, 24, 1847, 885-887 = *Œuvres*, (1), 10, 292-295.
 359 4. 119-127 = *Collected Math. Papers*, 1, 55-62.
 360 *Math. Ann.*, 21, 1883, 563-564 = *Ges. Abh.*, 182.
 361 *Werke*, 8, 201.
 362 *Werke*, 10, 363.
 363 Report of the Brit. Assn. for the Adv. of Sci., 1883, 3-37 = *Coll. Math. Papers*, 11, 429-459
 364 *Jour. für Math.*, 101, 1887, 337-353 = *Werke*, 3, 249-274.
 365 *Nachrichten Konig. Ges. der Wiss. zu Gilt.*, 15, 1868, 193-221 = *Wiss. Abh.*, 2, 618-639.
 366 C. Hermite-T. Stieltjes correspondance, *Gauthier-Villars*, 1905, 2, 398.
 367 *Atti del Congreso*, 1, 1929, 141 = *Grundlagen der Geometrie*, 7.^a ed., 323.
 368 *Mind*, 38, 1929, 1-25.
 369 *Ges. Werke*, 1, 454-455.
 370 *Elementary Mathematics from an advanced standpoint*, Mac Millan, 1939; Dover (reimp.), 1945, vol. 2, 187
 371 *Math. Ann.* 78, 1918, 405-415 = *Ges. Abh.*, 145-156
 372 *Jahres. der Deut. Math.-Verein*, 13, 1904, 381.
 373 *Ann. Fac. Sci. de Toulouse*, 8, 1894, J. 1-122 y 9, A. 1-47 = *Oeuvres completes*, 2, 402-559.
 374 *Math. Ann.*, 23, 1884, 152-156.
 375 *Math. Ann.*, 23, 1884, 453-488 = *Ges. Abh.*, 210-246.
 376 *Math.*, 8, 1892, 69-99 = (*Oeuvres*, 4, 427-457.
 377 *Annali di Mat.*, (3), 7, 1902, 231-259.
 378 *Atti della Accad. dei Lincei, Rendiconti*, (4), 1, 1885, 321-326, 532-537, 566-569.
 379 *E. g., Ann. de l'Ecole Norm. Sup.*, (3), 20, 1903, 453-485.
 380 *Math. Ann.*, 61, 1905, 251-280.
 381 *Acta Math.*, 30, 1906, 335-400.
 382 *Gior di Mat.*, 19, 1881, 333-372 = *Opere Mat.*, 1, 16-48.
 383 Los dos números derivadas por la derecha son los dos límites

$$\limsup_{h \rightarrow 0, h > 0} \frac{g(x+h) - g(x)}{h} \quad \liminf_{h \rightarrow 0, h > 0} \frac{g(x+h) - g(x)}{h}$$

Los dos números derivadas por la izquierda se definen análogamente.

- 384 *Atti della Accad. dei Lincei, Rendiconti*, (5), 16, 1907, 608-614.
 385 *Ann. de l'Ecole Norm. Sup.*, (3), 27, 1910, 361-450.
 386 *Atti Accad., Tormo*, 43, 1908, 229-246.
 387 *Comp. Rend.*, 128, 1899, 1502-05.
 388 *Sitzungsber. der Akad. der Wiss. Wien*, 122, *Abt. IIA*, 1913, 1295-1438.
 389 *Jour. für Math.*, 103, 1888, 228.
 390 *Mém. de l'Acad. des Sci. Paris*, 1782, 1-88, pub. 1785, y 1783, 423-467, pub. 1786 =

OEuvres, 10, 209-291, p. 236 en particular.

³⁹¹ *Jour. de l'Ecole Poly.*, 1823, 1-144, 249-403.

³⁹² *Magazin for Naturwidenskaberne*, 1, 1823 = *OEuvres*, 1, 11-27.

³⁹³ *Jour. für Math.*, 1, 1826, 153-157 = *OEuvres*, 1, 97-101.

³⁹⁴ *Jour. de l'Ecole Poly.*, 13, 1832, 1-69.

³⁹⁵ *Jour. de Math.*, 2, 1837, 16-35.

³⁹⁶ *Rendiconti del Circolo Matematico di Palermo*, 8, 1894, 57-155 = *OEuvres*, 9, 123-196.

³⁹⁷ *Acta Math.*, 20, 1896-1897, 59-142 = *OEuvres*, 9, 202-272. Véase también la referencia de Hellinger y Toeplitz en la bibliografía, p. 1354.

³⁹⁸ *Atti della Acad. dei Lincei, Rendiconti*, (5), 1896, 177-185, 289-300; *Atti Accad. Torino*, 31, 1896, 311-323, 400-408, 557-567; *Annali di Mat.*, (2), 25, 1897, 139-178; todos están en sus *Opere matematiche*, 2, 216-313.

³⁹⁹ *Acta Math.*, 27, 1903, 365-390.

⁴⁰⁰ Puede encontrarse una excelente exposición en Gerhard Kowalewski, *Integralgleichungen*, Walter de Gruyter, 1930, 101-134.

⁴⁰¹ Normalización significa que se modifica $\Phi^k(s)$ de tal modo que se tenga $\int_0^1 (\Phi^k)^2(s) = 1$.

⁴⁰² *Math. Ann.*, 48, 1897, 365-389.

⁴⁰³ *Rendiconti del Circolo Matematico di Palermo*, 22, 1906, 241-259.

⁴⁰⁴ *Math. Ann.*, 72, 1912, 562-577 = *Grundzuge*, cap. 22.

⁴⁰⁵ *Math. Ann.*, 63, 1907, 433-476 y 64, 1907, 161-174.

⁴⁰⁶ *Comp. Rend.*, 144, 1907, 613-619, 734-736, 1409-1411.

⁴⁰⁷ *Comp. Rend.*, 144, 1907, 1022-1024.

⁴⁰⁸ *Comp. Rend.*, 144, 1907, 1148-1150.

⁴⁰⁹ *Math. Ann.*, 69, 1910, 449-497.

⁴¹⁰ *Math. Ann.*, 66, 1908, 273-324 = *Ges. Abh.*, 1, 1-86.

⁴¹¹ *Atti della Reale Accademia dei Lincei*, (4), 3, 1887, 97-105, 141-146, 153-158 = *Opere Matematiche*, 1, 294-314, y otros del mismo año y posteriores

⁴¹² *Bull. Soc. Math. de France*, 30, 1902, 40-43 = (*Euvres*, 1, 401-404.

⁴¹³ *Werke*, p. 30

⁴¹⁴ *Memorie della Reale Accademia dei Lincei*, (3), 18, 1883, 521-586.

⁴¹⁵ *Atti della Accademia dei Lincei, Rendiconti*, (4), 5, 1889, 342-348.

⁴¹⁶ Verhandlungen des ersten internationalen Matbematiker-Kongresses, *Teubner*, 1898, 201-202.

⁴¹⁷ Verhandlungen, 204-205.

⁴¹⁸ *Comp. Rend.*, 136, 1903, 351-354 = (*Euvres*, 1, 405-408.

⁴¹⁹ *Rendiconti del Circolo Matematico di Palermo*, 22, 1906, 1-74.

⁴²⁰ *Annali di Mat.*, (3), 13, 1899, 1-122.

⁴²¹ El límite inferior es el mínimo punto límite de la sucesión $U(A_n)$.

⁴²² *Amer. Math. Soc. Trans.*, 15, 1914, 135-161.

⁴²³ *Amer. J. Math.*, 33, 1913, 369-394.

⁴²⁴ *Amer. Math. Soc. Trans.*, 21, 1920, 357-383.

⁴²⁵ Véase, por ejemplo, *Atti del IV Congresso Intemazionale dei Matematici* (1908), 2, Reale Accademia dei Lincei, 1909, 98-114, y *Amer. Math. Soc. Bull.*, 18, 1911-1912, 334-362.

⁴²⁶ *Rendiconti del Circolo Matematico di Palermo*, 25, 1908, 53-77.

⁴²⁷ *Nouvelles Annales de Mathématiques*, 8, 1908, 97-116, 289-317.

⁴²⁸ *Comp. Rend.*, 144, 1907, 1414-1416.

⁴²⁹ *Amer. Math. Soc. Trans.*, 8, 1907, 433-446.

⁴³⁰ *Comp. Rend.*, 136, 1903, 351-354 = *CEuvres*, 1, 405-408.

⁴³¹ *Comp. Rend.*, 149, 1909, 974-977, y *Ann. de l'Ecole Norm. Sup.*, 28, 1911, 33 ss.

⁴³² *Acta Math.*, 41, 1918, 71-98.

⁴³³ *Fundamenta Mathematicae*, 3, 1922, 133-181.

⁴³⁴ *Studia Mathematica*, 1, 1929, 211-216.

⁴³⁵ *Jour. für Math.*, 137, 1927, 214-229.

- 436 *Math. Ann.*, 102, 1929-1930, 49-131 y 370-427 = *Coll. Works*, 2, 3-143.
- 437 *Trans Cambridge Phil. Soc.*, 8, Part II, 1844, 182-203, pub. 1849.
- 438 *Comp. Rend.*, 17, 1843, 370-376 = *OEuvres*, (1), 8, 18-25.
- 439 Tercera edición, 1820, 88-109 = *OEuvres*, 7, 89-110.
- 440 *OEuvres*, 7, 128-131.
- 441 Tercera edición, 1820, vol. 1, parte 2, Cap. 1 = *OEuvres*, 7, 89-110.
- 442 *Mém. de l'Acad. des Sci., Inst. France*, 1, 1827, Note 16 = *OEuvres*, (1), 1, 230.
- 443 *Trans. Cambridge Phil. Soc.*, 9, 1856, 166-187 = *Math. and Phys. Papers*, 2, 329-357.
- 444 *Phil. Magaz.*, (5), 23, 1887, 252-255 = *Math. and Phys. Papers*, 4, 303-306.
- 445 *Proc. Camb. Phil. Soc.*, 19, 1918, 49-55.
- 446 *Comp. Rend.*, 15, 1842, 554-556 y 573-578 = *OEuvres*, (1), 7, 149-157.
- 447 *Novi Comm. Acad. Sci. Petrop.*, 9, 1762-1763, 246-304, pub. 1764 = *Opera*, (2), 10, 293-343.
- 448 *Astronom. Nach.*, 28, 1849, 65-94 = *Werke*, 7, 145-174.
- 449 *Jour. de Math.*, 2, 1837, 16-35.
- 450 *Trans. Camb. Phil. Soc.*, 6, 1837, 457-462 = *Math. Papers*, 225-230.
- 451 *Zeit. für Physik*, 38, 1926, 518-529.
- 452 *Zeit. für Physik*, 39, 1926, 828-840.
- 453 *Comp. Rend.*, 183, 1926, 24-26.
- 454 *Proc. London Math. Soc.*, (2), 23, 1923, 428-436.
- 455 *Trans. Camb. Phil. Soc.*, 9, 1856, 166-187 = *Math. and Phys. Papers*, 2, 329-357.
- 456 *Trans. Camb. Phil. Soc.*, 10, 1857, 106-128 = *Math. and Phys. Papers*, 4, 77-109.
- 457 *Acta Math.*, 8, 1886, 293-344 = *OEuvres*, 1, 290-332.
- 458 Vol. 2, cap. 8, 1893.
- 459 *Acta Math.*, 24, 1901, 289-308.
- 460 *Math. Ann.*, 50, 1898, 525-556.
- 461 *Amer. Math. Soc. Trans.*, 10, 1909, 436-470 = *Coll. Math. Papers*, 1, 201-235.
- 462 *Amer. Math. Soc. Trans.*, 9, 1908, 219-231 y 380-382 = *Coll. Math. Papers*, 1, 1-36.
- 463 *Proc. roy. Soc.*, A86, 1912, 207-226 = 5a. *Papers*, 6, 71-90.
- 464 *Annalen der Phys.*, (4), 47, 1915, 709-736.
- 465 *Proc. London Math. Soc.*, (2), 23, 1922-1924, 428-436.
- 466 *Jour. de l'Ecole Poly.*, 11, 1820, 417-489.
- 467 *Jour. für Math.*, 1, 1826, 311-339 = *CEuvres*, 1, 219-250.
- 468 *Jour. für Math.*, 89, 1880, 262-264 = *Ges. Abh.*, 2, 8-10.
- 469 *Comm. Acad. Sci. Petrop.* 16, 1771, 71-90.
- 470 *Jour. für Math.*, 15, 1836, 355-364.
- 471 *Math. Ann.*, 20, 1882, 535-549.
- 472 *Bull. Sci. Math.*, (2), 1890, 114-120.
- 473 *Math. Ann.*, 67, 1909, 110-125.
- 474 *Novi Comm. Acad. Sci. Petrop.*, 5, 1754-1755, 205-237, pub. 1760 = *Opera*, (1), 14, 585-617, y *Nova Acta Acad. Sci. Petrop.*, 2, 1784, 36-45, pub. 1788 = *Opera*, (1), 16, 34-43.
- 475 *Novi Comm. Acad. Sci. Petrop.*, 5, 1754-1755, 205-237, pub. 1760 = *Opera*, (1), 14, 585-617.
- 476 Euler, *Opera Posthuma*, 1, 545-549.
- 477 *Buil. Soc. Math. de France*, 7, 1879, 72-81 = *OEuvres*, 1, 428-437.
- 478 *Ann. Fac. Sci. de Toulouse*, 8, 1894, J. 1-122, y 9, 1895, A. 1-47 = *OEuvres completes*, 2, 402-559.
- 479 *Ann. de t'Ecole Norm. Sup.*, (3), 16, 1899, 9-136.
- 480 *Monatshefte für Mathematik und Physik*, 8, 1897, 273-277.
- 481 *Math. Ann.*, 58, 1904, 51-69.
- 482 *Bull. des Sci. Math.*, (2), 16, 1892, 167-189 = *Opere*, 1, 288-310.
- 483 *Math. Ann.*, 54, 1901, 125-201 = Ricci, *Opere*, 2, 185-271.
- 484 *Zeit. für Math. und Phys.*, 62, 1914, 225-261.

- 485 Atti della Accad. dei Lincei, Rendiconti, (4), 3, 1887, 15-18 = *Opere*, 1, 199-203.
- 486 Jour. für Math., 70, 1869, 46-70 y 241-245, y 71-102.
- 487 Atti della Accad. dei Lincei, Rendiconti, (4), 5, 1889, 112-118 = *Opere*, 1, 268-275.
- 488 *Annalen der Phys.*, 17, 1905, 891-921; hay traducción inglesa en la edición de Dover de A. Einstein, *The prináple of relativity* (1951).
- 489 *Annalen der Phys.*, 49, 1916, 769-822.
- 490 La métrica es $ds^2 = dx^2 + dy^2 + dz^2 - c^2 dt^2$. Se trata de un espacio de curvatura constante. Toda sección por un plano $t = \text{const.}$ es euclídea.
- 491 Rendiconti del Circolo Matematico di Palermo, 42, 1917, 173-205.
- 492 *Math. Ann.*, 78, 1918, 187-217.
- 493 *Mathematische Zeitschrift*, 2, 1918, 384-411 = *Ges. Abh.*, 2, 1-28.
- 494 Proceedings of the National Academy of Sciences, 8, 1922, 19-23.
- 495 Esta tesis se publicó hace algún tiempo: *Über Kurven und Flächen in allge- meinen Raumen*, Birkhauser Verlag, Basilea, 1951.
- 496 *Werke*, 3, 439-446.
- 497 Bull. des Sci. Math., (2), 1, 1877, 17-41, en particular p. 41 = *Werke*, 3, 262-296.
- 498 Monatsber. Berliner Akad., 1870, 881-889 = *Werke*, 1, 271-282.
- 499 *Math. Ann.*, 13, 1878, 561-565; Proc. London Math. Soc., 9, 1878, 126-133; Amer. J. of Math., 1, 1878, 50-52 y 174-176: todo el volumen 10 de sus *Collected Math. Papers*.
- 500 Jour. für Math., 86, 1879, 217-262 = Frobenius, *Ges. Abh.*, 1, 545-590.
- 501 Nachrichten König. Ges. der Wiss. zu Gött., 1874, 529-542 = *Ges. Abh.*, 5, 1-8.
- 502 *Ges. Abh.*, 5, 314-360.
- 503 *Math. Ann.* 20, 1882, 1-44 y 22, 1883, 70-118.
- 504 Amer. Math. Soc. Bull., 8, 1902, 296-300 y 388-391, y Amer. Math. Soc. Trans. 6, 1905, 181-197.
- 505 Amer. Math. Soc. Trans., 3, 1902, 485-492, y 6, 1905, 179-180.
- 506 Amer. Math. Soc. Trans., 6, 1905, 198-204.
- 507 Jour für Math., 100, 1887, 179-181 = *Ges. Abb.*, 2, 301-303.
- 508 *Math. Ann.*, 48, 1897, 548-561 - *Werke*, 2, 87-102.
- 509 Amer. Mat. Soc. Bull., 4, 1898, 510-515 = *Coll. Works*, 1, 266-269.
- 510 Amer. Math. Soc. Bull., 4, 1898, 135-139 = *Coll. Works*, 1, 254-257.
- 511 *Math. Ann.*, 43, 1893, 301-412.
- 512 Amer. Math. Soc. Bull., 1, 1895, 61-66, y 2, 1896, 33-43.
- 513 *Coll. Math. Papers*, 10, 403.
- 514 *Math. Ann.*, 34, 1889, 26-56.
- 515 *Math. Ann.*, 40, 1892, 55-88 y 43, 1893, 301-412.
- 516 *Math. Ann.*, 46, 1895, 321-422.
- 517 Sitzungsber. Akad. Wiss. zu Berlin, 1893, 337-345, y 1895, 1027-1044 = *Ges. Abh.*, 2, 565-573, 677-694.
- 518 Un resultado muy importante ha sido obtenido recientemente por Walter Feit (1930) y John G. Thompson (*Pacific Jour. of Math.*, 13, parte 2, 1963, 775-1029): todos los grupos finitos de orden impar son resolubles. Burnside ya había sugerido en 1906 que podía ser así.
- 519 *Math. Ann.*, 71, 1911, 116-144.
- 520 *Math. Ann.*, 106, 1932, 295-307.
- 521 Esto fue demostrado por P. S. Novikov en 1955. Véase American Math. Soc. Translations (2), 9, 1958, 1-122.
- 522 *Quat. Jour. of Math.*, 33, 230-238.
- 523 Amer. J. of Math., 4, 1881, 221-225.
- 524 Sitzungsber, Akad. Wiss. zu Berlin, 1897, 994-1015 = *Ges. Abh.*, 3, 82-103.
- 525 Proc. London Math. Soc., (2), 3, 1905, 430-434.
- 526 *Math. Ann.*, 52, 1899, 363-368.
- 527 Sitzungsber. Akad. Wiss. zu Berlin, 1896, 985-1021 = *Ges. Abh.*, 3, 1-37.
- 528 Sitzungsber. Akad. Wiss. zu Berlin, 1897, 994-1015 = *Ges. Abh.*, 3, 82-103.
- 529 *Acta Math.*, 38, 1921, 145.

- ⁵³⁰ Math. Ann., 43, 1893, 521-549. Para los grupos, ver Math. Ann., 20, 1882, 301-329.
- ⁵³¹ Amer. Math. Soc. Trans., 4, 1903, 13-20 y 6, 1905, 198-204.
- ⁵³² Amer. Math. Soc. Trans., 4, 1903, 31-37 y 6, 1905, 181-197.
- ⁵³³ Jour. für Math., 137, 1910, 167-309.
- ⁵³⁴ Bulletin des Sciences Mathématiques de Férussac, 13, 1830, 428-435 = (*Euvres*, 1897, 15-23).
- ⁵³⁵ N. Y. Math. Soc. Bull., 3, 1893, 73-78.
- ⁵³⁶ Amer. Math. Soc. Trans., 6, 1905, 349-352.
- ⁵³⁷ Amer. Math. Soc. Trans., 15, 1914, 31-46.
- ⁵³⁸ Amer. Math. Soc. Trans., 15, 1914, 162-166.
- ⁵³⁹ En 1958 demostraron Michel Kervaire (1927) (en los Proceedings of the National Academy of Sciences, 44, 1958, 280-283) y John Milnor (1931) (en Annals of Math., (2), 68, 1958, 444-449), utilizando ambos un resultado de Raoul Bott (1923), que las únicas álgebras con división posibles con coeficientes reales, si no se suponen las propiedades asociativa y conmutativa de la multiplicación, son los números reales, los complejos, los cuaterniones y los números de Cayley.
- ⁵⁴⁰ Proc. London Math. Soc., (2), 6, 1907, 77-118.
- ⁵⁴¹ Ann. Fac. Sci. de Toulouse, 128, 1898, 1-99 = *Vuvres*, parte II, vol. I, 7-105
- ⁵⁴² Math. Ann., 60, 1905, 20-116.
- ⁵⁴³ Math. Ann., 83, 1921, 24-66.
- ⁵⁴⁴ Archiv für Mathematik Naturvidenskab, 1, 1876, 152-193 = Ges. Abh., 5, 42-75.
- ⁵⁴⁵ Math. Ann., 31, 1888, 252-290 y vols. 33, 34 y 36.
- ⁵⁴⁶ 1894; Paris, Vuiber, 2.^a ed., 1933 = *OEuvres*, parte I, vol. 1, 137-286.
- ⁵⁴⁷ Ann. de l'École Norm. Sup., 31, 1914, 263-355 = *OEuvres*, parte I, vol. 1, 399-491.
- ⁵⁴⁸ Bull. Soc. Math. de France, 41, 1913, 53-96 = *OEuvres*, parte I, vol. 1, 355-398.
- ⁵⁴⁹ Mathematische Zeitschrift, 23, 1925, 271-309 y 24, 1926, 328-395 = *Ges. Abh.*, 2, 543-647.
- ⁵⁵⁰ Las definiciones de algunas propiedades básicas de los conjuntos de puntos, tales como la compacidad y la separabilidad, han sido distintas según los diferentes autores y todavía no han sido unificadas. Aquí utilizaremos los significados usados más comúnmente en la actualidad.
- ⁵⁵¹ Math. Ann., 94, 1925, 262-295.
- ⁵⁵² Math. Ann., 94, 1925, 309-315.
- ⁵⁵³ Revue de Métaphysique et de Morale, 20, 1912, 483-504.
- ⁵⁵⁴ Jour. für Math., 142, 1913, 146-152.
- ⁵⁵⁵ Fundamenta Mathematicae, 7, 1925, 30-137 y 8, 1926, 225-359.
- ⁵⁵⁶ Monatshefte für Mathematik und Physik, 33, 1923, 148-160 y 34, 1926, 137-161.
- ⁵⁵⁷ Fundamenta Mathematicae, 2, 1921, 256-285.
- ⁵⁵⁸ Math. Ann., 104, 1930, 71-80.
- ⁵⁵⁹ Monatshefte für Mathematik und Physik, 33, 1923, 148-160 y Math. Ann., 95, 1926, 266-306.
- ⁵⁶⁰ Fundamenta Mathematicae, 7, 1925, 30-137, parte de la p. 93.
- ⁵⁶¹ Leibniz, *Math. Schriften*, 1 Abt., vol. 2, 1850, 19-20 = Gerhardt, *Der Briefwechsel von Leibniz mit Mathematikern*, 1, 1899, 568 = Chr. Huygens, *Oeuw. Comp.*, 8, n.º 2192.
- ⁵⁶² Novi Comm. Acad. Sci. Petrop., 4, 1752-1753, 109-140, pub. 1758 = *Opera*, (1), 26, 71-93.
- ⁵⁶³ Novi Comm. Acad. Sci. Petrop., 4, 1752-1753, 140-160, pub. 1758 = *Opera*, (1), 26, 94-108.
- ⁵⁶⁴ Jour. de L'École Poly., 9, 1813, 68-86 y 87-98 = *OEuvres*, (2), 1, 7-38.
- ⁵⁶⁵ Comm. Acad. Sci. Petrop., 8, 1736, 128-140, pub. 1741. Puede encontrarse una traducción inglesa de este artículo en James R. Newman, *The world of mathematics*, Simón and Schuster, 1956, vol. 1, 573-580. Hay versión castellana, Sigma. El mundo de las matemáticas, Barcelona, Grijalbo, 8.a ed., 1983.
- ⁵⁶⁶ *Werke*, 8, 270-286.
- ⁵⁶⁷ Abh. der Ges. der Wiss. zu Gött., 10, 1861, 97-180, y como libro en 1862.
- ⁵⁶⁸ Königlich Sächsischen Ges. der Wiss. zu Leipzig, 15, 1863, 18-57 = *Werke*, 2, 433-471.
- ⁵⁶⁹ *Werke*, 2, 518-559.

- ⁵⁷⁰ Königlich Sachsische der Wiss. zu Leipzig, 17, 1865, 31-68 = *Werke*, 2, 473-512; ver también p. 519.
- ⁵⁷¹ Proceedings of the Royal Geographic Society, 1, 1879, 259-261 = *Coll. Math. Papers*, 11, 7-8.
- ⁵⁷² Jour. für Math., 54, 1857, 105-110 = *Werke*, 91-96, ver también *Werke*, 479-482.
- ⁵⁷³ Proc. London Math. Soc., 8, 1877, 292-304 = *Math. Papers*, 241-254.
- ⁵⁷⁴ *Über Riemanns Theorie der algebraischen Funktionen und ihrer Intégrale*, B. G. Teubner, 1882; reimpresión inglesa en Dover, 1963. También en Klein, *Ges. Math. Abh.*, 3, 499-573.
- ⁵⁷⁵ Math. Ann., 7, 1874, 549-557 = *Ges. Math. Abh.*, 2, 63-77.
- ⁵⁷⁶ Jour. für Math., (2), 11, 1866, 105-109 = *Oeuvres*, 4, 85-89
- ⁵⁷⁷ Annali di Mat., (2), 4, 1870-1871, 140-158.
- ⁵⁷⁸ Jour. de l'École Poly., (2), 1, 1895, 1-121 = (*Euvres*, 6, 193-288.
- ⁵⁷⁹ Rendiconti del Circolo Matematico di Palermo, 13, 1899, 285-343 = (*Euvres*, 6, 290-337. Proc. London Math. Soc., 32, 1900, 277-308 = *CEuvres*, 6, 338-370.
- ⁵⁸¹ Estos p_k son inferiores en una unidad a los de Poincaré. Aquí utilizamos el enunciado que es más familiar hoy.
- ⁵⁸² Rendiconti del Circolo Matematico di Palermo, 18, 1904, 45-110 = (*Euvres*, 6, 435-498.
- ⁵⁸³ Amer. Math. Soc. Trans., 20, 1919, 339-342.
- ⁵⁸⁴ La conjetura generalizada ha sido demostrada para $n > 5$ por Stephen Smale (Amer. Math. Soc. Bull., 66, 1960, 373-375), John R. Stallings (*ibid.*, 485-488) y E. C. Zeeman (*ibid.*, 67, 1961, 270).
- ⁵⁸⁵ La *Hauptvermutung* se ha demostrado para complejos simpliciales finitos (que son más generales que las variedades) de dimensión menor que 3, y es falsa para los de dimensión mayor que 5. Es correcta para variedades de dimensión menor o igual que 3 y está abierta para variedades de dimensión igual o mayor que 4. Ver John Milnor, *Annals of Math.*, (2), 74, 1961, 575-590.
- ⁵⁸⁶ Amer. Math. Soc. Trans., 16, 1915, 148-154.
- ⁵⁸⁷ Math. Ann., 70, 1910-1911, 161-165 y 71, 1911-1912, 305-313.
- ⁵⁸⁸ Math. Ann., 71, 1911-1912, 97-115.
- ⁵⁸⁹ Proceedings Koninklijke Akademie von Wetenschappen te Amsterdam, 12, 1910, 785-794.
- ⁵⁹⁰ Math. Ann., 71, 1911-1912, 97-115.
- ⁵⁹¹ Math. Ann., 71, 1911-1912, 97-115.
- ⁵⁹² Math. Ann., 71, 1911-1912, 97-115.
- ⁵⁹³ Rendiconti del Circolo Matematico di Palermo, 33, 1912, 375-407 = (*Euvres* 6, 499-538.
- ⁵⁹⁴ Amer. Math. Soc. Trans., 14, 1913, 14-22 = *Coll. Math. Papers*, 1, 673-681.
- ⁵⁹⁵ Amer. Math. Soc. Trans., 23, 1922, 96-115 = Birkhoff, *Coll. Math. Papers*, 3, 255-274.
- ⁵⁹⁶ *Studia Mathematica*, 2, 1930, 170-179.
- ⁵⁹⁷ Ann. de l'École Norm. Sup., 51, 1934, 45-78.
- ⁵⁹⁸ Math. Ann., 71, 1911-1912, 314-319.
- ⁵⁹⁹ Amer. Math. Soc. Trans., 23, 1922, 333-349.
- ⁶⁰⁰ Alexander formuló este teorema bajo la condición de que los coeficientes de sus cadenas fueran enteros módulo 2 (véase parágrafo siguiente). Nuestra formulación es para coeficientes enteros usuales.
- ⁶⁰¹ *Annals of Math.*, (2), 14, 1913, 163-178.
- ⁶⁰² Amer. Math. Soc. Trans., 28, 1926, 301-329.
- ⁶⁰³ *Annals of Math.*, (2), 29, 1928, 232-254.
- ⁶⁰⁴ *Annals of Math.*, (2), 35, 1934, 904-914.
- ⁶⁰⁵ *Annals of Math.*, (2), 30, 1928-1929, 101-187.
- ⁶⁰⁶ Math. Ann., 97, 1927, 454-472.
- ⁶⁰⁷ *Fundamenta Mathematicae*, 19, 1932, 149-183.
- ⁶⁰⁸ *Revue Générale des Sciences*, 16, 1905, 541.
- ⁶⁰⁹ Proc. London Math. Soc., (2), 4, 1906, 29-53.
- ⁶¹⁰ *Abhandlungen der Friesschen Schule*, 2, 1908, 301-324.
- ⁶¹¹ Math. Ann., 46, 1895, 481-512 = *Ges. Abh.*, 282-356.
- ⁶¹² *Ges. Abh.*, 443-448.

-
- ⁶¹³ Math. Ann., 65, 1908, 261-281.
- ⁶¹⁴ Math. Ann., 86, 1921-1922, 230-237, y muchos artículos posteriores.
- ⁶¹⁵ Jour. für Math., 154, 1925, 219-240 y artículos posteriores.
- ⁶¹⁶ Las opiniones de estos matemáticos aparecen en un famoso intercambio epistolar. Véase Bull. Soc. Math. de France, 33, 1905, 261-273. También en E. Borel, *Leçons sur la théorie des fonctions*, Gauthier-Villars, 4.^a ed., 1950, 150-158.
- ⁶¹⁷ Pub. 1690 - G. W. Leibniz: Die philosophischen Schriften, ed. por C. I. Gerhardt, 1875-1890, vol. 4, 27-102.
- ⁶¹⁸ Trans. Camb. Phil. Soc., 10, 1858, 173-230.
- ⁶¹⁹ Jour. für Math., 101, 1887, 337-355 = *Werke*, 3, 251-274.
- ⁶²⁰ Jour. für Math., 92, 1882, 1-122 = *Werke*, 2, 237-387.
- ⁶²¹ Acta Math., 22, 1899, 17.
- ⁶²² Ver nota 9.
- ⁶²³ Amer. Math. Monthly, 53, 1946, 2-13 = *Ges. Abh.*, 4, 268-279.
- ⁶²⁴ *Proc. Third Internat. Congress of Math.*, Heidelberg, 1904, 174-185 = *Grundlagen der Geom.*, 7.^a ed., 247-261; traducción inglesa en el *Monist*, 15, 1905, 338-352.
- ⁶²⁵ Math. Ann., 95, 1926, 161-190 = *Grundlagen der Geometrie*, 7.^a ed., 262-288. Ver nota 20.
- ⁶²⁶ Weyl, Amer. Math. Soc. Bull., 50, 1944, 637 = *Ges. Abh.*, 4, 157.
- ⁶²⁷ Math. Ann., 95, 1926, 161-190 = *Grundlagen der Geometrie*, 7.^a ed., 262-288. Hay traducción inglesa en Paul Benacerraf y Hilary Putnam, *Philosophy of Mathematics*, 134-181, Prentice-Hall, 1964.
- ⁶²⁸ Monatshefte für Mathematik und Physik, 38, 1931, 173-198, véase la bibliografía.
- ⁶²⁹ Atti del Congresso Internazionale Dei Matematici, I, 135-141 = *Grundlagen der Geometrie*, 7.^a ed., 313-323.
- ⁶³⁰ Jour. für Math., 154, 1925, 1.
- ⁶³¹ Abh. Math. Seminar der Hamburger Univ., 1, 1922, 157-177 = *Ges. Abh.*, 3, 157-177.
- ⁶³² Math. Ann., 112, 1936, 493-565.
- ⁶³³ Proceedings of the National Academy of Sciences, 50, 1963, 1143-1148; 51, 1964, 105-110.
- ⁶³⁴ Journal of Symbolic Logic, 14, 1949, 2-8.
- ⁶³⁵ Obituary Notices of Fellows of the Royal Soc., 4, 1944, 547-553 = *Ges. Abh.*, 4, 121-129, parte de la p. 126.