
EVOLUCIÓN HISTÓRICA DE LOS CONCEPTOS BÁSICOS
DEL ANÁLISIS MATEMÁTICO

Francisco Javier Pérez González

Departamento de Análisis Matemático

Universidad de Granada

Septiembre de 2021

Índice

1. Números y cantidades. El continuo y lo discreto	1
1.1. Los inicios de la aritmética	1
1.2. Números y medida de magnitudes	1
1.3. Números y cantidades en la antigua Grecia	2
1.4. Los pitagóricos y el descubrimiento de las magnitudes inconmensurables	3
1.5. Eudoxo y Euclides. Álgebra geométrica	8
1.6. De la antigua Grecia a la invención del Cálculo	10
1.7. Infinitésimos y el continuo numérico	16
1.8. El triunfo de Pitágoras	20
1.9. Cortaduras de Dedekind	22
1.10. Métodos axiomáticos y métodos constructivos	24
1.11. El regreso de los pequeños	25
1.12. Otros números: complejos, cuaterniones, octoniones.	26
2. Evolución del concepto de función y límite funcional	29
2.1. La teoría de las “razones últimas” de Newton	29
2.2. Sobre el concepto de función	30
2.3. La <i>metafísica del Cálculo</i> en D’Alembert y Lagrange	34
2.4. El premio de la Academia de Berlín de 1784	36
2.5. Cauchy y su <i>Cours d’Analyse</i> de 1821	39
2.6. El innovador trabajo de Bolzano	43
2.7. Weierstrass nos dio los $\epsilon - \delta$	45
3. Orígenes y desarrollo del concepto de derivada. La invención del Cálculo	47
3.1. Las matemáticas en Europa en el siglo XVII	47
3.2. Cálculo de tangentes y de valores extremos	48
3.2.1. El método de máximos y mínimos de Fermat	49

3.2.2.	El método de las tangentes de Fermat	51
3.2.3.	El método de Roberval y de Torricelli para las tangentes	54
3.2.4.	El triángulo diferencial de Barrow	54
3.3.	Los inventores del Cálculo	56
3.4.	Newton y el cálculo de fluxiones	57
3.5.	Leibniz y el cálculo de diferencias	63
3.6.	Desarrollo del cálculo diferencial	67
4.	Evolución de la idea de integral	68
4.1.	Problemas de cuadraturas en las matemáticas griegas	68
4.1.1.	Cuadratura de un segmento de parábola por Arquímedes	70
4.1.2.	<i>El Método</i> de Arquímedes	73
4.1.3.	Área de una espiral	74
4.2.	La integración antes del Cálculo	76
4.2.1.	Los indivisibles de Cavalieri	76
4.2.2.	Cuadratura de la cicloide por Roberval	77
4.2.3.	Parábolas e hipérbolas de Fermat	78
4.2.4.	La integración aritmética de Wallis	80
4.2.5.	El resultado fundamental de Barrow	83
4.3.	La relación fundamental entre cuadraturas y tangentes	84
4.3.1.	El Teorema Fundamental del Cálculo según Newton	84
4.3.2.	La invención del <i>calculus summatorius</i> por Leibniz	86
5.	El Cálculo y las series	89
5.1.	Los primeros desarrollos en serie	91
5.2.	Newton y las series infinitas	92
5.3.	Euler y el cálculo con series	98
5.4.	El problema de la convergencia y la divergencia	102
5.5.	El problema de la cuerda vibrante y las series trigonométricas	105

5.6. La ecuación del calor y las series de Fourier	107
6. Breve historia del infinito	111
6.1. La idea de infinito en la filosofía y la matemática Griegas	111
6.1.1. Las aporías de Zenón de Elea	111
6.1.2. Atomismo y divisibilidad infinita	113
6.2. El infinito desde la Edad Media hasta el siglo XIX	116
6.3. El infinito matemático y el nacimiento de la teoría de conjuntos	118
7. Los orígenes del Análisis Funcional	128
7.1. Sistemas de infinitas ecuaciones lineales	130
7.2. La escuela italiana: Ascolí, Arzelá, Volterra	132
7.3. Las ecuaciones integrales y su influencia en el desarrollo del Análisis Funcional	135
7.4. David Hilbert y el nacimiento de la teoría espectral	140
7.5. Riesz, Hahn, Banach. Espacios vectoriales normados	143

1. Números y cantidades. El continuo y lo discreto

1.1. Los inicios de la aritmética

Estamos tan acostumbrados a *contar* que cuesta trabajo imaginar un mundo sin números. Pero así fue, no creo que nadie lo dude, durante muchísimo tiempo. Incluso en nuestros días se tienen noticias de tribus aisladas que no saben contar más allá de cuatro o cinco objetos; cuando tienen que referirse a una cantidad mayor emplean una expresión que quiere decir “muchos”. Es frecuente también que en los lenguajes primitivos se utilicen palabras distintas para designar números iguales cuando éstos se refieren a diferentes clases de objetos. Poco a poco, conforme los primitivos grupos tribales fueron organizándose en sociedades cada vez más complejas, el proceso de contar colecciones concretas de objetos condujo al concepto de “número abstracto”. . . !Que a nosotros nos parece tan *natural!* Pero esto fue un largo proceso cuyo desarrollo es el tema de esta lección.

En el tercer milenio a.C. en el antiguo Egipto y en Mesopotamia se inventaron diversos símbolos con significado numérico y sistemas de numeración que se usaron para resolver una variedad de problemas prácticos. Las características más sorprendentes del sistema numérico babilónico son el principio de notación posicional y la base 60. También utilizaron el principio de notación posicional para representar fracciones. Esto les permitió disponer de una aritmética bastante avanzada y utilizarla en muchas situaciones prácticas, especialmente en astronomía. Sin embargo, se limitaban a dar instrucciones verbales de los cálculos a realizar, sin justificarlos de ninguna manera. Los egipcios, por su parte, inventaron un sistema de escritura numérica que usaba la base 10, pero no era posicional, sino aditivo. Usaban un símbolo especial para representar fracciones unitarias (con numerador igual a 1) y las demás fracciones se escribían como suma de estas. Contar, con la ayuda de símbolos numéricos, señala el principio de la aritmética.

Las matemáticas que se desarrollaron en Mesopotamia y en Egipto no fueron más que un conjunto de reglas para resolver problemas de la vida diaria. Casi no hay simbolismo, apenas algún pensamiento consciente sobre abstracciones, ninguna formulación metodológica general y ninguna idea de demostración o incluso de razonamiento plausible que pudiera convencer a alguien de la corrección de un procedimiento o fórmula. No hubo, de hecho, ninguna concepción de ciencia teórica de ningún tipo.

Quien quiera profundizar en estos temas puede consultar [9] y [12].

1.2. Números y medida de magnitudes

Una vez inventados los números, el paso siguiente fue usarlos para *medir magnitudes* tales como longitudes, superficies, volúmenes o tiempos. Este proceso requiere bastante ingenio. Consi-

deremos, para fijar ideas, que queremos *expresar numéricamente* la longitud de un segmento de recta \overline{AB} . Lo primero que hay que hacer es elegir una *unidad de medida* que será otro segmento \overline{OU} y comparar ambos. Puede ocurrir que \overline{AB} contenga un número exacto, m , de veces a \overline{OU} . En tal caso podemos escribir simbólicamente $\overline{AB} = m\overline{OU}$. El número m representa entonces la *medida* de \overline{AB} respecto de \overline{OU} . Lo más frecuente, no obstante, es que \overline{OU} no esté contenido un número exacto de veces en \overline{AB} . En tal caso podemos dividir \overline{OU} en un cierto número, n , de partes iguales con la esperanza de que, al tomar como nueva unidad de medida una de estas partes, $\overline{OU'}$, resulte que \overline{AB} contenga un número exacto, m , de veces a $\overline{OU'}$. Cuando esto es así se dice que los segmentos \overline{AB} y \overline{OU} son *commensurables*. Esto quiere decir que *admiten una unidad de medida común*: el segmento $\overline{OU'}$. Podemos escribir $\overline{AB} = m\overline{OU'}$. Por otra parte $\overline{OU} = n\overline{OU'}$. Podemos ahora usar los números m, n para hacernos una idea de cómo es \overline{AB} comparado con \overline{OU} ; esto es lo que se expresa diciendo que *la razón de \overline{AB} respecto de \overline{OU} es $m : n$* (léase m sobre n).

$$\frac{\overline{AB}}{\overline{OU}} = \frac{m}{n}$$

Con el paso del tiempo el símbolo $m : n$ que, en sus orígenes, como acabamos de explicar, representaba la razón de (las longitudes) de dos segmentos commensurables quedó desprovisto de su significado original y pasó a ser considerado simplemente como un *número* naciendo de esta forma los números *racionales* (cuyo nombre alude, precisamente, a que tales números representan *razones* de segmentos commensurables).

1.3. Números y cantidades en la antigua Grecia

Los griegos de la antigüedad distinguían entre “*número*” y “*cantidad*” o “*magnitud*”. Para ellos un número era un agregado de unidades. Podemos precisar más. Un número es una multiplicidad que se obtiene por repetición de un individuo – la unidad –, cuyas partes están separadas – son discontinuas – y tienen fronteras bien definidas. Todo esto se expresa diciendo que los números son de naturaleza *discreta*¹. Por otra parte, los números no tienen sentido si se separan de los objetos materiales o ideales a los que enumeran. Así, “tres árboles” tiene sentido, pero “tres” por sí mismo carece de significado. Es decir, un número es un atributo de un grupo de objetos y carece de autonomía propia.

Una “*cantidad*” puede ser, entre otras cosas, tiempo, longitud, volumen, velocidad o masa. La característica esencial de la cantidad es su *continuidad*². Una cantidad puede dividirse indefinidamente, pero no está formada por partes separadas que son réplicas de una unidad, sino que

¹La palabra “discreta” deriva de una raíz latina que significa “separar”. De esta misma raíz deriva el verbo “discernir” cuyo significado es reconocer algo como distinto o separado [3].

²La palabra “continuo” deriva de una raíz latina que significa “mantener junto” o “cohesionar”. Esta misma raíz nos da el sustantivo “continente” con el significado de una extensión de tierra no interrumpida por el mar [3].

sus componentes están unidos entre sí por fronteras comunes: donde acaba uno empieza otro. Por ejemplo, un área plana puede dividirse en trozos que, al estar unidos unos con otros, pierden su singularidad quedando como partes indiferenciadas de un todo. Por otra parte, los matemáticos griegos no estudiaron la cantidad como algo abstracto, para ellos *las cantidades tienen siempre un carácter concreto*: son una cantidad de algo.

El concepto de cantidad estaba estrechamente ligado a la Geometría. Una proporción entre dos segmentos o entre dos áreas planas es una cantidad que a veces puede expresarse con ayuda de números. Cuando dichos segmentos o áreas admiten una unidad de medida común podemos decir que la razón de uno a otro es, por ejemplo, $7 : 10$. Pero, para los griegos, $7 : 10$ no es un número sino una forma de expresar una cantidad concreta, que podría leerse algo así como “siete partes de diez”. Ellos *solamente consideraban como números los enteros positivos* y ni siquiera consideraban como número a la unidad. La unidad es, eso, “la unidad” de la que están formados los números, pero ella misma no es un número porque no está compuesta de unidades.

1.4. Los pitagóricos y el descubrimiento de las magnitudes inconmensurables

Volviendo a la situación antes descrita, parece intuitivo que, cualquiera sea el segmento \overline{AB} , dividiendo \overline{OU} en un número, n , *suficientemente grande* de partes iguales, podemos conseguir que la nueva unidad de medida, $\overline{OU'}$, esté efectivamente contenida un número exacto de veces en \overline{AB} . En otras palabras, parece que dos segmentos cualesquiera deben ser conmensurables. Pues bien, la intuición aquí nos engaña, y ese fue el extraordinario descubrimiento que realizaron los pitagóricos.

Para tratar de encontrar una unidad de medida común para dos segmentos $\alpha_0 = \overline{AB}$ y $\alpha_1 = \overline{CD}$ se procede como sigue. Supuesto que $\alpha_0 > \alpha_1$, quitamos de α_0 tantas veces como podamos α_1 hasta obtener como sobrante un segmento $\alpha_2 < \alpha_1$:

$$\alpha_0 = n_1 \alpha_1 + \alpha_2 \quad \alpha_2 < \alpha_1$$

Es decir, el segmento α_0 contiene n_1 veces al α_1 y queda un segmento sobrante $\alpha_2 < \alpha_1$. Repetimos el proceso para obtener:

$$\begin{aligned} \alpha_1 &= n_2 \alpha_2 + \alpha_3 & \alpha_3 < \alpha_2 \\ \alpha_2 &= n_3 \alpha_3 + \alpha_4 & \alpha_4 < \alpha_3 \\ &\vdots & \vdots \end{aligned}$$

Si α_0 y α_1 tienen una unidad de medida común, este proceso termina después de un número finito de etapas, es decir, hay un $k \in \mathbb{N}$ tal que $\alpha_{k-1} = n_k \alpha_k$ de manera que α_k es una unidad de medida común para los segmentos α_0 y α_1 . Si el proceso puede continuarse indefinidamente obteniendo

en cada etapa segmentos sobrantes cada vez más pequeños $\alpha_{n+1} < \alpha_n$ es porque los segmentos de partida no son commensurables.

Observa que si α_0 y α_1 son números enteros, este proceso es el algoritmo de Euclides para la división de enteros y α_k es el máximo común divisor de α_0 y α_1 .

Parece ser que fue un pitagórico, [Hipaso de Metaponto](#), quien en el siglo V a.C., al estudiar las propiedades geométricas del pentagrama, descubrió la existencia de los segmentos inconmensurables.

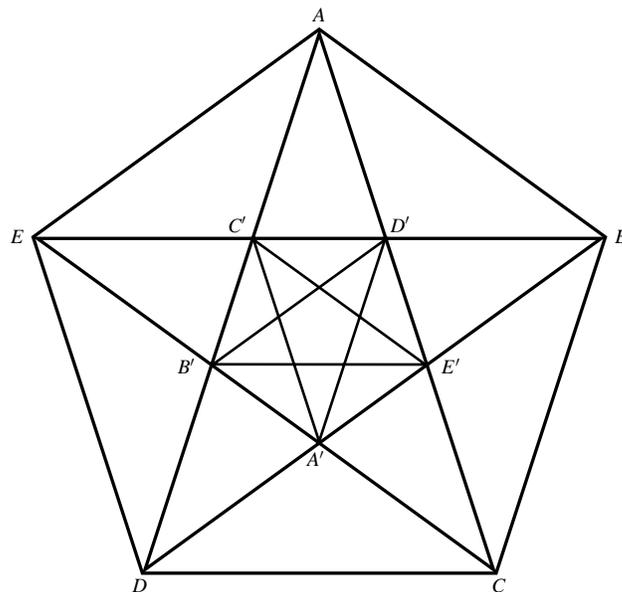


Figura 1. El pentagrama pitagórico

En el pentágono regular $ABCDE$ las diagonales se cortan formando otro pentágono regular $A'B'C'D'E'$. Debido a la simetría, cada lado de un pentágono regular es paralelo a una de sus diagonales. Por tanto los triángulos AED y $BE'C$ tienen sus lados correspondientes paralelos y, en consecuencia, sus ángulos son iguales, por lo que son semejantes. Por tanto:

$$\frac{\overline{AD}}{\overline{AE}} = \frac{\overline{BC}}{\overline{BE'}}$$

Pero $\overline{BE'} = \overline{BD} - \overline{DE'} = \overline{BD} - \overline{AE} = \overline{BD} - \overline{BC}$ pues $\overline{DE'} = \overline{AE}$ por ser lados opuestos de un paralelogramo. Y como $\overline{AE} = \overline{BC}$ y $\overline{AD} = \overline{BD}$, hemos obtenido que:

$$\frac{\overline{BD}}{\overline{BC}} = \frac{\overline{BC}}{\overline{BD} - \overline{BC}}$$

Es decir, en cualquier pentágono regular se cumple que

$$\frac{\text{diagonal}}{\text{lado}} = \frac{\text{lado}}{\text{diagonal} - \text{lado}}$$

Tenemos además:

$$\overline{BD} - \overline{EA} = \overline{BD} - \overline{DE'} = \overline{E'B} = \overline{D'B} = \overline{D'A'}$$

Es decir, la diferencia entre la diagonal y el lado del pentágono mayor es igual a la diagonal del pentágono menor. Y también:

$$\overline{BC} - \overline{D'A'} = \overline{BA'} - \overline{BE'} = \overline{E'A'}$$

Es decir, la diferencia entre el lado del pentágono mayor y la diagonal del pentágono menor es igual al lado del pentágono menor.

Pongamos $\alpha_0 = \text{diagonal} = \overline{BD}$, $\alpha_1 = \text{lado} = \overline{BC}$, $\alpha_2 = \alpha_0 - \alpha_1$. Tenemos que

$$\frac{\alpha_0}{\alpha_1} = \frac{\alpha_1}{\alpha_2}$$

Este proceso puede continuarse pues si ahora formamos la diferencia $\alpha_3 = \alpha_1 - \alpha_2$, puesto que α_2 es la diagonal del pentágono menor y α_3 es su lado, tenemos que $\alpha_3 < \alpha_2$ y poniendo $\alpha_4 = \alpha_2 - \alpha_3$:

$$\frac{\alpha_2}{\alpha_3} = \frac{\alpha_3}{\alpha_4}$$

Y dicho número es evidentemente el mismo para cualquier pentágono regular, luego:

$$\frac{\alpha_0}{\alpha_1} = \frac{\alpha_1}{\alpha_2} = \frac{\alpha_2}{\alpha_3} = \frac{\alpha_3}{\alpha_4}$$

Proceso que puede continuarse indefinidamente puesto que las diagonales de cada pentágono regular determinan otro pentágono regular. Esto significa que el proceso que hemos descrito anteriormente:

$$\begin{array}{ll} \alpha_0 = 1 \cdot \alpha_1 + \alpha_2 & \alpha_2 < \alpha_1 \\ \alpha_1 = 1 \cdot \alpha_2 + \alpha_3 & \alpha_3 < \alpha_2 \\ \alpha_2 = 1 \cdot \alpha_3 + \alpha_4 & \alpha_4 < \alpha_3 \\ \vdots & \vdots \end{array}$$

no termina nunca y por tanto los segmentos α_0 y α_1 son inconmensurables. Luego la diagonal de un pentágono regular es inconmensurable con el lado.

De la igualdad

$$\frac{\alpha_0}{\alpha_1} = \frac{\alpha_1}{\alpha_0 - \alpha_1}$$

se deduce que $\frac{\alpha_0}{\alpha_1} = \frac{1+\sqrt{5}}{2}$. Y acabamos de probar que dicho número no es cociente de números enteros, es decir, no es un número racional.

El número $\Phi = \frac{1+\sqrt{5}}{2}$ es uno de los más famosos de las Matemáticas. Se conoce como *razón áurea*. Si en *Google* buscas “razón áurea” (así, con las comillas) te saldrán casi tres mil páginas. Eso en español, porque si buscas en inglés “golden section” obtendrás casi novecientas mil páginas. El poeta Rafael Alberti dedicó un hermoso soneto a la razón áurea.

A la divina proporción

A ti, maravillosa disciplina,
media, extrema razón de la hermosura,
que claramente acata la clausura
viva en la malla de tu ley divina.
A ti, cárcel feliz de la retina,
áurea sección, celeste cuadratura,
misteriosa fontana de medida
que el Universo armónico origina.
A ti, mar de los sueños angulares,
flor de las cinco formas regulares,
dodecaedro azul, arco sonoro.
Luces por alas un compás ardiente.
Tu canto es una esfera transparente.
A ti, divina proporción de oro.

La siguiente demostración, que aparece al final del libro X de los *Elementos* de Euclides, prueba que la diagonal de un cuadrado no es conmensurable con el lado.

En efecto, si \overline{OU} es el lado y \overline{AB} la diagonal, y suponemos que ambos admiten una unidad de medida común $\overline{OU'}$, tendremos que $\overline{OU} = n\overline{OU'}$, y $\overline{AB} = m\overline{OU'}$ para convenientes números naturales m, n . Pero, en virtud del teorema de Pitágoras, $2(\overline{OU})^2 = (\overline{AB})^2$, y deducimos que $2n^2(\overline{OU'})^2 = m^2(\overline{OU'})^2$, por lo que debe ser $2n^2 = m^2$. Veamos que esto lleva a una contradicción. Podemos suponer que m y n no tienen factores comunes (si los tuvieran se los quitamos) y, en particular, m y n no son ambos pares. La igualdad $2n^2 = m^2$ nos dice que m^2 es par lo cual implica que también tiene que serlo m . Así podemos escribir $m = 2p$. Sustituyendo en la igualdad anterior y simplificando tenemos que $n^2 = 2p^2$, y de aquí se sigue, al igual que antes, que n tiene que ser par y ésta es la contradicción anunciada.

Suele afirmarse que los pitagóricos pensaban que el *número* era el *fundamento último* de toda realidad aunque ni siquiera Aristóteles tenía claro lo que eso significaba [3]. Puede que los pitagóricos atribuyeran un significado físico a esos *números*, considerados como sustrato universal subyacente a todo, y los entendieran como las *partes últimas indivisibles* de cualquier magnitud, lo que Demócrito y Leucipo llamaron *átomos*. Eso explicaría la profunda crisis que produjo entre los pitagóricos el descubrimiento de las cantidades inconmensurables, pues la existencia de tales cantidades niega directamente la interpretación atomista según la cual cualquier segmento de línea está compuesto de un número finito de *mínimas unidades indivisibles*, pues si esto fuera así está claro que tales unidades mínimas serían una unidad de medida común para cualquier par de segmentos.

El descubrimiento de la inconmensurabilidad fue un logro sin precedentes porque no fue empírico sino puramente teórico, puso de manifiesto que la geometría no es aritmética y que los objetos matemáticos no eran tan simples como se pensaba. De hecho, algunas propiedades que parecían claramente verdaderas (como que dos segmentos siempre admitían una unidad de medida común), resultaban ser falsas. Esta crisis de fundamentos hizo cuestionar la seguridad del método seguido para demostrar las propiedades de los objetos matemáticos, consistente en *hacer ver o poner en evidencia* que tales resultados eran necesariamente verdaderos. Y así, en algún momento de la segunda mitad del siglo V a.C., un grupo de matemáticos griegos establecieron un nuevo método para el descubrimiento de la verdad: el *método axiomático-deductivo*, que es esencialmente el mismo que usamos hoy. Se trata de, partiendo de unas pocas verdades evidentes (o axiomas), y a través de una serie de etapas sucesivas muy simples, obtener una cadena de afirmaciones, con la propiedad de que una cualquiera de ellas es verdadera siempre que lo sean todas las anteriores. A las leyes que rigen las formas correctas de pasar de una afirmación a otra de la cadena, se les llamó más tarde *leyes lógicas o deductivas*, y tienen un *carácter formal*, independiente del carácter de verdadero o falso de la afirmación a la que se aplica. A estas cadenas de afirmaciones lógicamente correctas, los griegos las llamaron *demostraciones* [4].

Aristóteles atribuye a los pitagóricos el mérito de haber sido los fundadores en el siglo V a.C. de la matemática como ciencia deductiva. Ellos fueron los primeros en insistir en el razonamiento deductivo como único método de demostración en matemáticas.

Hoy vivimos en un mundo digitalizado: la música que escuchas, las películas que ves, la televisión digital y tantas más cosas de uso cotidiano son, en su esencia, números. Parece que los pitagóricos no estaban del todo equivocados.

Pitágoras, junto con su maestro Tales de Mileto, y también Anaximandro y Anáximenes, sin olvidar a Demócrito y algunos más de los que queda memoria y que tú mismo puedes consultar en *Wikipedia*, todos ellos eran matemáticos y filósofos. ¿Casualidad? Ni mucho menos. Lo que hoy llamamos *Cultura Occidental* nace de una gran blasfemia, a saber, la afirmación de que la realidad puede ser comprendida y explicada racionalmente. Frente a los relatos mitológicos y a

los caprichosos dioses imprevisibles, la afirmación insolente de que la inteligencia humana puede desentrañar por sus propios medios el funcionamiento del Universo. Y ¿qué produce la inteligencia humana cuando empieza a indagar sobre la Naturaleza? Matemáticas y Filosofía. Las Matemáticas, por su propia naturaleza, tienen un campo mucho más restringido que el de la Filosofía pero, en cambio, son extraordinariamente eficaces. La palabra griega $\mu\alpha\theta\eta\mu\alpha$, que se lee *mathema*, significa *conocimiento*.

El Libro del Universo está escrito en el lenguaje de las matemáticas y sus caracteres son triángulos, círculos y otras figuras geométricas, sin las cuales es imposible entender ni una palabra; sin ellos es como girar vanamente en un oscuro laberinto.
Galileo Galilei

1.5. Eudoxo y Euclides. Álgebra geométrica

La existencia de segmentos inconmensurables era un serio problema para el desarrollo de la geometría pues, como dichos segmentos no pueden compararse, no se sabía cómo interpretar su razón. Por ejemplo, un resultado, sin duda conocido por los pitagóricos, afirma que *las áreas de dos triángulos con igual altura están en la misma proporción que sus bases*. ¿Qué sentido tiene esta afirmación si las bases no son segmentos conmensurables? El problema está en que no había una forma de comparar razones entre magnitudes inconmensurables pues el concepto de proporcionalidad pitagórico suponía que las cantidades que se hallan en proporción poseen una medida común. Surge así la necesidad de extender una teoría de la proporción, es decir de la igualdad entre dos razones, que incluyera las razones de cantidades inconmensurables. Esto fue llevado a cabo por un gran matemático **Eudoxo de Cnido** (c. 400 - 347 a.C.) que introdujo la idea de *magnitud continua* o de *cantidad*. Se trata de un concepto que, aunque no se define, permite considerar cantidades tales como longitudes, áreas, volúmenes, ángulos, pesos, tiempo... Eudoxo concebía estas magnitudes de modo geométrico y no les asignaba ningún valor numérico. De esta forma se evitaba el uso de lo que nosotros conocemos como *números irracionales*. Eudoxo definía entonces una razón de tales cantidades y a partir de ella una proporción, es decir, una igualdad de dos razones, que cubría los casos de razones conmensurables e inconmensurables. Puesto que no se utilizaba número alguno para expresar tales razones, los conceptos de razón y proporción quedaban ligados a la geometría lo que condujo al desarrollo de un *álgebra geométrica* cuya característica es tratar los problemas algebraicos por medio de construcciones geométricas. Así fueron resueltas las ecuaciones cuadráticas obteniendo sus raíces como segmentos. Por ejemplo, la fórmula del binomio

$$(a + b)^2 = a^2 + 2ab + b^2$$

se expresa como sigue: *Si se divide un segmento, como viene dado, entonces el cuadrado sobre el segmento entero es igual a los cuadrados sobre las partes más dos veces el rectángulo formado por las partes conjuntamente.* Lo que debe entenderse en el sentido de que las áreas geoméricamente combinadas de los dos cuadrados y el rectángulo sobre las partes eran igual al área del cuadrado sobre el segmento. No hay que olvidar que para los griegos los únicos números son los naturales, y por tanto una propiedad relativa a cantidades que no se consideran numéricas no puede traducirse algebraicamente usando operaciones como suma y producto que los griegos solamente usaban cuando se trataba de números naturales.

Se desarrolló así una especie de “álgebra geométrica” en la que los números se representaban por segmentos de línea y las operaciones aritméticas fueron sustituidas por construcciones geométricas. Las ecuaciones lineales y cuadráticas fueron resueltas con técnicas geométricas, evitándose así el problema de las magnitudes inconmensurables. De esta forma en las matemáticas griegas el razonamiento geométrico llegó a considerarse como el modelo de razonamiento matemático riguroso. Y así siguió siendo durante más de 2000 años.

[Euclides](#), el matemático más famoso de la Escuela de Alejandría, expone la teoría de Eudoxo en el Libro V de los *Elementos* (300 a.C.) y en ella destacan los siguientes puntos que enunciamos con nuestro simbolismo actual muy diferente de la forma puramente verbal en que fueron formulados ([12], I.4.5).

E1 (*Propiedad arquimediana*) Dadas dos cantidades siempre hay un múltiplo de una de ellas que excede a la otra. Es decir, si es $0 < a < b$ hay algún $n \in \mathbb{N}$ tal que $na > b$.

Observa que esta propiedad impide la existencia de cantidades positivas *infinitamente pequeñas*, los llamados *infinitésimos* de los que hablaremos más adelante.

E2 (*Criterio de igualdad*) Las razones $a : b$ y $c : d$ son iguales si cualesquiera sean los enteros positivos m, n se tiene que

$$ma < nb \implies mc < nd, \quad ma = nb \implies mc = nd, \quad ma > nb \implies mc > nd \quad (1)$$

Esta definición de la proporción no necesita suponer que las cantidades son conmensurables. Al mismo tiempo, permite demostrar todas las proposiciones conocidas sobre proporciones. Volveremos a considerar más adelante este elaborado criterio de igualdad que, desde luego, no aclaraba nada sobre la naturaleza de las cantidades irracionales y ponía de manifiesto la dificultad de reducir a la aritmética el estudio de las mismas.

Los libros VII, VIII y IX de los *Elementos* de Euclides tratan de la teoría de números, esto es, de las propiedades de los números naturales, de las razones entre números naturales y entre magnitudes conmensurables. Su estudio es independiente del Libro V dedicado a las magnitudes continuas propias de la geometría, es decir, Euclides trata separadamente los conceptos de número

y magnitud.

Euclides, en el Libro X de los *Elementos*, clasifica las magnitudes inconmensurables en los tipos siguientes (uso, claro está, la notación actual):

$$a \pm \sqrt{b}, \quad \sqrt{a} \pm \sqrt{b}, \quad \sqrt{a \pm \sqrt{b}}, \quad \sqrt{\sqrt{a} \pm \sqrt{b}}$$

Donde se entiende que a y b son racionales. Se demuestran también identidades -¡todo ello mediante construcciones geométricas!- que con el simbolismo actual se traducen en ([25], pg. 45):

$$\sqrt{\sqrt{a} \pm \sqrt{b}} = \sqrt{\frac{\sqrt{a} + \sqrt{a-b}}{2}} \pm \sqrt{\frac{\sqrt{a} - \sqrt{a-b}}{2}}$$

y

$$\sqrt{\sqrt{a} \pm \sqrt{a^2 - b^2}} = \sqrt{\frac{a+b}{2}} \pm \sqrt{\frac{a-b}{2}}$$

La carencia de una teoría aritmética satisfactoria de las cantidades inconmensurables hizo que los matemáticos griegos consideraran la Geometría como una ciencia más general que la Aritmética, puesto que sólo la Geometría podía manipular las razones inconmensurables, y dedicaran sus esfuerzos al estudio de la primera en detrimento de la última. La consecuencia fue que la Geometría se convirtió en la base de casi todas las matemáticas y que durante casi 2000 años, en Europa, hasta al menos el año 1600, casi todo razonamiento matemático riguroso se expresó en lenguaje geométrico.

Quizás el único matemático griego, después de los pitagóricos, que no hizo Geometría sino Aritmética fue Diofanto de Alejandría (c. 214 - 298). En su obra llamada *Aritmética*, de la que se han conservado seis libros de un total de trece, resuelve diversos tipos de ecuaciones algebraicas admitiendo como soluciones números enteros o números fraccionarios positivos, los cuales son considerados por Diofanto como auténticos números y no solamente como proporciones. Otra innovación de Diofanto fue la invención de una notación “sincopada” que constituye el primer ejemplo de simbolismo matemático.

1.6. De la antigua Grecia a la invención del Cálculo

Es sabido que la civilización Romana, tan excelente en tantos aspectos, no destacó en el estudio de las ciencias puras y, en particular, de las matemáticas. La prueba de ello es que no hay ningún matemático Romano digno de mención. No obstante, el sistema de numeración Romano se impuso extendiéndose por todo el Imperio.

Con el triunfo del Cristianismo a finales del siglo IV y la caída del Imperio Romano de Occidente en el año 476, se inicia una larga era de oscurantismo en Europa. En el año 529, el emperador

cristiano Justiniano ordenó cerrar la Academia fundada por Platón por considerarla *reducto de enseñanzas paganas de funesta influencia*. La fe y los dogmas no son demostrables lógicamente; absurdas disputas teológicas ocupan el lugar de los estudios de la Naturaleza y la Biblia es la fuente de todo conocimiento. Según San Agustín “Las palabras de las Escrituras tienen más autoridad que toda la inteligencia humana”. El racionalismo científico es sospechoso de paganismo. Entonces. . . ¿Para qué pensar?

A diferencia que en Grecia, en la India se había desarrollado principalmente la Aritmética y se conocía el sistema de numeración posicional decimal desde el siglo VI. La primera vez que el cero es tratado como un número de pleno derecho es en la obra *Brahmasphutasiddhanta* del matemático y astrónomo indio [Brahmagupta](#) (598 - 670). Esta obra también contenía el principio de la numeración decimal posicional y los métodos de cálculo del álgebra india. En ella se tratan los números negativos en términos muy parecidos a los actuales.



Figura 2. al-Jwarizmi



Figura 3. Fibonacci

La herencia matemática griega pasa a los árabes. La cultura árabe tiene una época de esplendor en los siglos VIII - XII. Al-Mamun (c. 786 - 833), sexto califa de la dinastía Abasida, fundó en Bagdad la Casa de la Sabiduría, una especie de academia con una biblioteca y un observatorio. Allí se tradujeron las obras de los matemáticos y filósofos griegos y tuvieron conocimiento de las matemáticas indias.

El más conocido matemático de la Escuela de Bagdad fue Muhammad ibn-Musa al-Jwarizmi. En su obra *Libro de la Adición y la Sustracción según el cálculo de los hindúes* se describe el sistema decimal posicional y se dan métodos para realizar cálculos aritméticos con dicho sistema.

[Leonardo de Pisa](#) (c. 1170 - 1250), más conocido como Fibonacci, aprendió en sus viajes por los países árabes del Mediterráneo a usar los métodos de al-Jwarizmi. Al regresar a Italia, publicó en 1202 el *Liber abaci*, obra que contribuyó a extender el sistema de numeración indo-árabe en Occidente. Estudiando las soluciones de una ecuación de tercer grado, Fibonacci probó que había números irracionales diferentes de los considerados por Euclides. En consecuencia, las técnicas del álgebra geométrica griega no permitían construir todas las cantidades inconmensurables.

Fibonacci dio también una interpretación de los números negativos como pérdidas o deudas, que tuvo bastante buena acogida. Pero todavía deberá pasar mucho tiempo para que los números nega-

tivos y el cero sean totalmente aceptados como números.

En este apresurado repaso que estamos dando a la historia de los números, debemos avanzar ahora casi trescientos años para llegar a la siguiente etapa protagonizada por los matemáticos italianos del Renacimiento [Niccoló Tartaglia](#) (c. 1500 - 1557), [Gerolamo Cardano](#) (1501 - 1576), [Rafael Bombelli](#) (1526 - 1572) y [Ludovico Ferrari](#) (1522 - 1565). Los dos primeros resolvieron la ecuación general de tercer grado de la cual solamente se conocían las soluciones en algunos casos particulares. En la resolución de la cúbica, Cardano tuvo en cuenta las soluciones negativas aunque las llamó “ficticias”, y comprobó que la cúbica podía tener tres soluciones.



Figura 4. Cardano

Así mismo, Cardano reconoció por primera vez la existencia de lo que ahora llamamos números complejos (a los que Napier llamó “los fantasmas de los números reales”) aunque no los aceptó como posibles soluciones. Por su parte, Bombelli fue el primero en especificar las reglas para sumar y multiplicar números complejos. Usando dichas reglas, probó que podían obtenerse soluciones reales correctas para la cúbica, incluso cuando la fórmula de Tartaglia – Cardano requería el cálculo de raíces de números negativos.

De esta época es también un opúsculo *De Thiende* (1585) (“El Décimo”) – 36 páginas – de [Simon Stevin](#) (1548 - 1620), ingeniero y matemático nacido en Brujas, en el que se introducen las fracciones decimales y se explica su uso en las operaciones aritméticas. Así mismo, en su obra *L'Arithmetique* (1585) escribió que “no hay números inexplicables, irregulares, irracionales, surds³ o absurdos”, indicando con esto que todos los números debían ser tratados por igual y no hacer distinciones entre ellos como si fueran de distinta naturaleza. Después de Stevin, la idea de que 1 era un número ganó una amplia aceptación.

A pesar de estos avances, los conceptos de “número” y “cantidad” de la antigüedad permanecen sin cambios notables hasta el siglo XVII cuando se desarrolla el simbolismo algebraico.

Lo importante del simbolismo algebraico, no es tanto el uso de los símbolos por sí mismos, sino la elaboración de reglas formales para realizar operaciones de forma simbólica. Por ejemplo, a^2 puede entenderse como una forma simplificada de escribir “el área del cuadrado de lado a ”. Eso es muy distinto de escribir $(a + b)^2 = a^2 + 2ab + b^2$. Esto último ya es una manipulación simbólica abstracta en la que las letras a , b no son más que símbolos sin una naturaleza concreta.

³La palabra griega “alogos”, $\alpha\lambda\omicron\gamma\omicron\varsigma$, usada por los griegos para designar a los números irracionales, también significa “sin discurso” y los árabes la tradujeron por *asamm*, “sordo” o “mudo”, que fue traducida al latín por *surdus*.

François Viète en su *In artem analyticem isagoge* (1591) expone una “logística speciosa” (*specis*: símbolo), o arte de calcular con símbolos, que fue un paso decisivo para el desarrollo del concepto de cantidad abstracta. No obstante, Viète consideraba que solamente las cantidades homogéneas podían compararse entre sí. Para entender esto debes tener en cuenta que, desde la antigüedad, el producto de dos cantidades, por ejemplo ab , representaba el área de un rectángulo de lados a y b . De la misma forma, abc representaba el volumen de un ortoedro. Una expresión como $ab + c$ no tenía significado porque no se podía sumar una longitud y un área: no eran cantidades homogéneas.



Figura 5. Viète



Figura 6. Fermat



Figura 7. Descartes

El siguiente paso definitivo fue el invento de la *geometría analítica* en los años 1630 por Pierre Fermat y René Descartes. La introducción de coordenadas y la representación de curvas por medio de ecuaciones supuso un cambio de perspectiva revolucionario. Piensa que, en la Antigüedad, solamente podían estudiarse aquellas curvas para las que se conocía un método de construcción con regla y compás. Ahora, por primera vez, los objetos geométricos podían estudiarse por medio del simbolismo algebraico, cuando hasta entonces lo usual había sido que el simbolismo algebraico fuera un pálido reflejo de relaciones geométricas.

Una importante creación de Descartes fue el desarrollo de un “álgebra de segmentos”. Para ello, tomando como unidad un segmento u , construyó un segmento (cantidad) c que verificaba la proporción $u : a = b : c$. Dicho segmento c representaba el producto de los dos segmentos (cantidades) a y b . También construyó segmentos que se correspondían con la suma, la diferencia y el cociente de segmentos. De esta manera, cualquier operación con cantidades se corresponde con un segmento, lo que hace que todas las cantidades sean homogéneas. Una expresión como $ab + c$ ya es correcta porque representa un segmento de línea. Esta homogeneización de todas las cantidades conduce al concepto de *cantidad abstracta* desconocido en la antigüedad.

Por esta época ya también los números eran objetos abstractos del pensamiento. Es decir, ya no eran simplemente un atributo del grupo al que contaban sino que se habían convertido en entidades

autónomas.

Descartes introdujo el término “imaginario” para referirse a aquellas soluciones de una ecuación polinómica que solamente están en “nuestra imaginación”. Como era costumbre, llamaba “soluciones falsas” a las soluciones negativas. Las “raíces verdaderas” eran las positivas.

A estos progresos en matemáticas hay que agregar los realizados en astronomía y en mecánica por Copérnico (1473-1543), Kepler (1571-1630) y Galileo (1564-1642). Todos ellos se apoyan en métodos experimentales y empíricos cuantitativos para formular sus resultados como Leyes de la Naturaleza de contenido matemático.

Al mismo tiempo, a lo largo de los dos primeros tercios del siglo XVII, se van desarrollando una gran variedad de “métodos infinitesimales”, cuyos precedentes clásicos estaban en Eudoxo y Arquímedes, para resolver multitud de problemas de tipo geométrico y analítico, como cálculo de tangentes a curvas, cálculo de áreas y de valores máximos. Los trabajos de Cavalieri (1598 - 1647), Wallis (1616 - 1703) y Barrow (1630 - 1677) entre otros muchos, establecieron las bases sobre las que dos grandes genios, Newton (1643 - 1727) y Leibniz (1646 - 1716) desarrollaron el Cálculo Infinitesimal. Más adelante veremos con algún detalle todo este proceso, pues ahora quiero considerar solamente aquellos aspectos del mismo relacionados con las ideas de número y de cantidad.

El Cálculo Infinitesimal son las matemáticas del cambio y del movimiento. Las ideas de magnitud variable y de dependencia entre magnitudes son fundamentales en estas nuevas matemáticas. Surge así el concepto de “variable” que se forma a partir de la idea de cantidad abstracta. En el libro de L'Hôpital *Analyse des infiniment petits, pour l'intelligence des lignes courbes* (1696) se lee:

Se llaman cantidades variables aquellas que aumentan o disminuyen continuamente, y por contraste cantidades constantes aquellas que permanecen igual mientras las otras cambian.

Los matemáticos de los siglos XVII y XVIII usan el término “cantidad” para referirse a cantidades generales abstractas, así como a cantidades geométricas concretas, pero siempre se consideran dichas cantidades como continuas. La noción de cantidad continua no se discute, se trataba de un concepto basado en la realidad física. Según Leibniz “Natura non facit saltus”.

La idea de cantidad es más general que la idea de número. Un segmento de línea, por ejemplo, representa una cantidad, pero él mismo no se reduce a números. La idea de número como elemento de un conjunto no existe en el siglo XVIII. Por la misma razón, un segmento no puede “separarse” de sus extremos y siempre los incluye. Los números eran interpretados como medidas. En *Arithmetica universalis* (1707) Newton escribe:

Por número entendemos no tanto una multitud de cantidades, como la razón abstracta de cualquier cantidad a otra cantidad de la misma clase que tomamos por unidad. Un entero es lo que es medido por la unidad, una fracción, aquello a lo que una parte submúltiplo de la unidad mide, y un surd, aquello que es inconmensurable con la unidad.

Esta interpretación de los números se corresponde con la consideración de las Matemáticas en los siglos XVII y XVIII como una Ciencia de la Naturaleza y, en consecuencia, los objetos matemáticos deben estar vinculados, directa o indirectamente, con la realidad física. Por ello, solamente se consideran como “verdaderos números” los que representan el resultado de una medida: los enteros y los racionales positivos. Los demás números (negativos, el 0 y los imaginarios) son necesarios y útiles para los cálculos, pero no son considerados “verdaderos números” son “ficticios”.

Los números irracionales positivos, aunque no son números en sentido estricto, tampoco son propiamente “ficticios”, porque pueden representarse por un segmento y sirven para medir cantidades geoméricamente especificadas. Los racionales e irracionales positivos son llamados “números reales” en oposición a los números imaginarios.

Los números empiezan a considerarse como entidades simbólicas sobre las que se opera con unas reglas establecidas (pero que no pueden ser libremente definidas). Por ejemplo, según Euler, $\sqrt{12}$ es un número que multiplicado por sí mismo es igual a 12, y esto es una definición simbólica de $\sqrt{12}$.

El desarrollo inicial del Cálculo, en el último tercio del siglo XVII, se basa en ideas vagas e imprecisas como “cantidad evanescente”, “razón última” o “infinitamente pequeño”. El uso de los “infinitésimos”, considerados como cantidades que, sin ser nulas, son más pequeñas que cualquier cantidad positiva imaginable, es característico de las técnicas del Cálculo.

Después de la invención del Cálculo el objetivo era usarlo para descubrir nuevos resultados. Al principio, nadie se preocupó mucho por la corrección matemática de los procedimientos empleados. La confianza en dichas técnicas descansaba en su extraordinaria eficacia para resolver multitud de problemas. Sin embargo, a finales del siglo XVIII, el uso continuado de los infinitésimos, que nadie sabía explicar, unido a la incomprensión que se tenía de los números irracionales y de los procesos de convergencia, propiciaron estudios críticos de los conceptos básicos del Cálculo, que acabaron llevando a una nueva formulación de los mismos mucho más formal y rigurosa, según los criterios actuales, pero también mucho menos intuitiva.

1.7. Infinitésimos y el continuo numérico

Estás leyendo *ahora mismo* estas palabras. Tienes un sentido preciso del “ahora”, al igual que del “pasado” y del “futuro”. Tenemos una percepción muy clara del “flujo del tiempo”. Percibimos el tiempo como un continuo: lo que separa dos instantes de tiempo es . . . tiempo. Dado un pequeño intervalo de tiempo, digamos un segundo, siempre podemos concebir otro intervalo más pequeño todavía, medio segundo o una cienbillonésima parte de un segundo. ¿Has pensado alguna vez hasta dónde es posible *dividir el tiempo*? Si aceptamos que el “instante” es lo que no tiene duración, parece difícil aceptar que el tiempo esté formado por instantes. ¿Debemos considerar entonces que hay una *unidad mínima* de tiempo, todo lo pequeña que queramos, pero que no se reduce a un instante? Estarás de acuerdo en que esa unidad mínima de tiempo sería algo así como una unidad de tiempo “infinitesimal”. ¿Cuántas unidades infinitesimales de tiempo caben en un minuto? ¿Un número finito? ¿Una cantidad infinita?

En el párrafo anterior podemos cambiar la palabra “tiempo” por “espacio” e “instante” por “punto” y llegaremos a los problemas derivados de la “infinita divisibilidad” del espacio. Tiempo y espacio son ejemplos de “continuo”. Una entidad continua, un *continuo*, es lo que no está roto ni separado ni tiene huecos, lo que puede ser indefinidamente dividido sin que pierda su naturaleza. Por ejemplo, un volumen de líquido, un segmento, un movimiento o, los ejemplos más inmediatos, el espacio y el tiempo.

Lo que relaciona espacio y tiempo es el movimiento. El Cálculo es la matemática del movimiento, del cambio continuo. El Cálculo se apoya en la geometría analítica de Descartes y Fermat y en la Aritmética. La Geometría se ocupa de cantidades continuas; la Aritmética de lo discreto. El Cálculo es la síntesis de lo “discreto” y lo “continuo”. Los “infinitésimos”, las cantidades infinitesimales, son el puente entre lo discreto y lo continuo.

Los procedimientos del Cálculo, límites, convergencia, continuidad, pueden describirse como matemáticas del *continuo numérico*. La expresión “continuo numérico” puede parecer un oxímoron, esto es, una combinación de dos palabras con significados opuestos y, en cierto sentido, es así. Los números sirven para contar grupos de cosas de igual naturaleza; por ejemplo árboles, o lo que quiera que sea, pero cada una de ellas con su propia individualidad, separadas entre sí, cosas que no tiene sentido dividir porque al hacerlo pierden su naturaleza. Todo esto se resume diciendo que los números tienen un carácter *discreto*. Los números siempre fueron considerados como lo opuesto del continuo.

La oposición continuo – discreto ha ocupado a los filósofos desde hace 2500 años y tiene como primeros representantes respectivos a [Parménides](#) (c. 510 - 450 a.C.) y a [Demócrito](#) (c. 460 - 370 a.C.).

Parménides, el filósofo más famoso de la Escuela Eleática, afirma en su hermoso poema *Sobre la Naturaleza*, que lo que Es, el Ser, es uno, ingénito, homogéneo, continuo, indivisible e inmutable. Este concepto del Ser excluye toda posibilidad de nueva generación de seres o sustancias y, por tanto, el cambio y el movimiento son mera ilusión, porque ambos presuponen que lo que no es pueda llegar a ser.

Demócrito es el representante más conocido de la Escuela Atomista cuyo materialismo se opone al idealismo de la Escuela Eleática. Demócrito mantiene que el universo está compuesto de pequeños corpúsculos invisibles, los “átomos”, que pueden poseer diferentes formas y extensiones y que por movimientos y combinaciones diversas en el vacío engendran la totalidad de lo existente.

[Zenón de Elea](#), discípulo de Parménides, es famoso por sus *aporías*, en las que trata de probar que tanto si el espacio o el tiempo son infinitamente divisibles, como si no lo son, el movimiento no existe o es imposible. Las aporías de Zenón son un extraordinario desafío, al que filósofos y matemáticos han dado diversas respuestas, sin que aún hoy se tenga conciencia clara de haberlas podido explicar de forma totalmente convincente.

Según Aristóteles, los atomistas preguntaban, en el supuesto de que una magnitud sea infinitamente divisible, qué es lo que quedaba de ella después de haberla sometido a un proceso de división exhaustivo. Y decían, si queda algo como polvo, es porque todavía no se ha completado el proceso de división, y si lo que queda son puntos o algo sin extensión, ¿cómo es posible recomponer una magnitud extensa con algo que no tiene extensión? Según ellos, la respuesta eran los átomos. La palabra griega “átomos” significa “lo que no puede dividirse”, por tanto, la Escuela Atomista negaba la infinita divisibilidad de la materia y afirmaba que cualquier magnitud contiene elementos indivisibles.

De la oposición continuo – discreto siguieron ocupándose los filósofos de la Antigüedad, [Platón](#) (c. 427 - 347 a.C.), [Aristóteles](#) (384 - 322a.C.), [Epicuro](#) (341 - 270 a.C.); y de la Edad Media [Duns Scoto](#) (c. 1266 - 1308), [Guillermo de Ockham](#) (c. 1280 - 1349), [Nicolás de Cusa](#) (1401 - 1464), entre otros. Éste último, en una supuesta demostración de la cuadratura del círculo, consideró una circunferencia como un polígono regular de infinitos lados. La idea de considerar que una curva está formado por infinitos segmentos infinitesimales de línea recta fue usada, entre otros, por Kepler, Galileo y Leibniz y está recogida en el libro de Guillaume de L'Hôpital *Analyse des Infiniment Petits pour l'Intelligence des Lignes Courbes* (1696) al cual ya nos hemos referido anteriormente.

A finales del siglo XVII, con el invento del Cálculo, resurgió la oposición entre lo continuo y lo discreto, esta vez centrada en el concepto de cantidad infinitesimal. Algunos consideraban los infinitésimos como algo real, infinitamente pequeño, parecido a los átomos de Demócrito, salvo que ahora su número era infinito. La integración se consideraba como una suma infinita de estos

infinitésimos. Una diferencial de una cantidad variable era un incremento infinitesimal de dicha variable, y un cociente o una razón de diferenciales “en el momento en que se anulan”, lo que Newton llamaba *cantidades evanescentes*, era lo que ahora llamamos una derivada, y Newton llamaba una *fluxión*.

El uso de los infinitésimos en el Cálculo demostraba ser muy eficaz y, aunque a algunos, como al mismo Newton, les hubiera gustado evitarlo, lo cierto es que no se sabía bien cómo hacerlo. Lo peor de todo, no es que el mero concepto de infinitésimo sea de por sí difícilmente sostenible, sino la forma en que los infinitésimos se manejaban en los cálculos. Podemos destacar dos características.

- Con los infinitésimos podía operarse como con cantidades finitas no nulas y, en particular, podía dividirse por ellos.
- Los infinitésimos podían ser tratados como cantidades nulas. Así, si x es una cantidad positiva y o un infinitésimo, entonces $x + o = x$.

Dependiendo del tipo de cálculo eran tratados de una forma u otra. Además, había infinitésimos de primer orden despreciables frente a cantidades finitas; de segundo orden que eran despreciables frente a los de primer orden, y así sucesivamente. Para acabar de empeorar las cosas, los infinitésimos no respetaban la [propiedad arquimediana](#), pues el producto de cualquier cantidad finita por un infinitésimo seguía siendo un infinitésimo.

Ejemplo. Un ejemplo típico es el cálculo de la diferencial de un producto de dos cantidades x e y . Se razonaba como sigue. Cuando x cambia a $x + dx$, y cambia a $y + dy$, por lo que xy se transforma en

$$(x + dx)(y + dy) = xy + xdy + ydx + dx dy$$

por lo que la diferencial de xy es $xdy + ydx + dx dy$, pero como $dx dy$ es una cantidad infinitamente pequeña con respecto a los otros términos, se sigue que la diferencial de xy es $xdy + ydx$.

Ejemplo. Veamos otro ejemplo típico. Consideremos dos cantidades x , y relacionadas por $y - x^3 = 0$. Cuando x cambia a $x + dx$, y cambia a $y + dy$, por lo que

$$0 = y + dy - (x + dx)^3 = y + dy - x^3 - 3x^2 dx - 3x(dx)^2 - (dx)^3$$

Teniendo en cuenta que $y - x^3 = 0$, deducimos:

$$dy = 3x^2 dx + 3x(dx)^2 + (dx)^3$$

Dividiendo por dx la igualdad obtenida resulta:

$$\frac{dy}{dx} = 3x^2 + 3x dx + (dx)^2$$

Y como $3x dx + (dx)^2$ es infinitamente pequeño respecto de $3x^2$, concluimos que $\frac{dy}{dx} = 3x^2$.

En lenguaje actual, lo que hemos hecho es calcular la derivada de la función $f(x) = x^3$. Y... ¡el resultado es correcto! A pesar de que hemos dividido por una cantidad que después hemos hecho igual a cero.

En 1734 el filósofo [George Berkeley](#) (1685 - 1753) publicó una obra cuyo título es *El analista, o discurso dirigido a un matemático infiel, donde se examina si el objeto, principios e inferencias del análisis moderno están formulados de manera más clara, o deducidos de manera más evidente, que los misterios de la religión y las cuestiones de la fe*. En dicha obra Berkeley, que fue obispo anglicano de Cloyne, hace una crítica de los fundamentos del Cálculo que tuvo una gran influencia. Afirmaba Berkeley que si se acepta que el Cálculo puede alcanzar soluciones exactas por medio de razonamientos erróneos, entonces debe admitirse que la fe puede alcanzar la verdad por vías místicas. Es famoso su comentario:

¿Qué son las fluxiones? Las velocidades de incrementos evanescentes. Y ¿qué son estos mismos incrementos evanescentes? Ellos no son ni cantidades finitas, ni cantidades infinitamente pequeñas, ni siquiera son nada. ¿No las podríamos llamar los fantasmas de las cantidades desaparecidas?

Los embrollos en que andaban metidos los matemáticos se reflejan en la novela de Jonathan Swift *Los viajes de Gulliver* (1726) donde aparecen los diminutos enanos de Lilliput y los enormes gigantes de Brobdingnag, y en la narración corta *Micromegas* (1752) de Voltaire.

La realidad es que los matemáticos del siglo XVIII, y hasta bien entrado el siglo XIX, estaban mucho más interesados en desarrollar y aplicar las técnicas del Cálculo, que en ocuparse de problemas de fundamentos. Entre los principales matemáticos de esta época hay que citar a [Leonard Euler](#) (1707 - 1783), [Jean d'Alembert](#) (1717 - 1783), [Joseph-Louis Lagrange](#) (1736 - 1813), [Pierre-Simon Laplace](#) (1749 - 1827), [Joseph Fourier](#) (1768 - 1830), [Carl Friedrich Gauss](#) (1777 - 1855). El espíritu de los tiempos, el Siglo de las Luces, queda bien reflejado en la siguiente frase.

Todos los efectos de la naturaleza son tan sólo las consecuencias matemáticas de un pequeño número de leyes inmutables. Laplace

En el primer tercio del siglo XIX, el ideal de Newton de “someter los fenómenos de la Naturaleza a las leyes matemáticas”, podía considerarse esencialmente realizado.

1.8. El triunfo de Pitágoras

Llegamos así al siglo XIX que, en cuanto a matemáticas se refiere, ha sido llamado el Siglo del Rigor. Veamos cómo se entendían en los primeros años de dicho siglo los conceptos básicos del Cálculo.

- Concepto de función. No existía tal como lo entendemos en la actualidad. En vez de funciones, se consideraban relaciones entre variables, es decir, ecuaciones. Las correspondencias entre variables se interpretaban en términos geométricos. No existía la idea del dominio de una variable.
- Concepto de continuidad. El concepto de continuidad puntual no había sido siquiera formulado matemáticamente. La idea de Euler de función continua, como aquella que está definida por una única expresión analítica, era todo lo que había.
- Concepto de límite. Solamente se tenían algunas ideas confusas agravadas por el uso de los infinitésimos. Los infinitésimos empezaban a considerarse como variables con límite cero.
- Concepto de número. La idea de cantidad abstracta variable, a la que podían asignarse valores concretos, no había experimentado cambios notables en casi un siglo. Los números complejos ya eran aceptados, gracias a los trabajos de Euler y, sobre todo, de Gauss, pero seguía sin tenerse una idea clara de los números irracionales, y prevalecía una interpretación geométrica de los mismos.

Esta situación iba a cambiar gracias principalmente a los trabajos de [Bernad Bolzano](#) (1781 - 1848), [Augustin Louis Cauchy](#) (1789 - 1857) y [Karl Weierstrass](#) (1815 - 1897) de los que nos ocuparemos al estudiar la formalización del concepto de límite. Ahora quiero detenerme solamente en la evolución de la idea de número real. A los tres matemáticos citados hay que agregar los nombres de [Richard Dedekind](#) (1831 - 1916) y [George Cantor](#) (1845 - 1918), fueron ellos quienes desarrollaron la teoría de los números reales. Es lógico preguntarse por qué esto no se hizo antes. Pueden darse varias razones para ello.

- En el siglo XVIII las matemáticas son consideradas una Ciencia de la Naturaleza. Las teorías matemáticas deben reflejar la realidad física. Las matemáticas son una herramienta para formular y descubrir las Leyes de la Naturaleza. Las teorías matemáticas no se inventan, se descubren.
- Los números reales estaban asociados con magnitudes y se interpretaban geoméricamente. Eran algo dado en la realidad física. A los matemáticos del siglo XVIII no les pareció necesario dar una definición matemática de los mismos.

- Observa que para precisar un número como $\sqrt{2}$ debes dar todas sus cifras decimales en su orden, es decir, un vector de infinitas componentes. Fíjate también que la condición dada por Eudoxo (1) para comparar razones inconmensurables hace intervenir a *todos* los números naturales. Esto no es casual. La idea de número irracional lleva consigo asociada la de infinito. Hasta que no se elaboraron los fundamentos de una teoría matemática del infinito, no pudo desarrollarse una teoría satisfactoria de los números reales.

En el siglo XVIII las definiciones matemáticas eran descriptivas; no creaban objetos matemáticos sino que describían algo que se suponía debía imitar una realidad externa. Por la misma razón, no podían inventarse reglas para operar con los objetos matemáticos. Las reglas había que descubrirlas, pero no podían elegirse libremente. Se consideraba que la Naturaleza imponía unas normas que las Matemáticas de alguna manera debían imitar, no se era libre para inventar una teoría matemática. La idea de una Matemática como juego lógico formal era algo impensable en el siglo XVIII.

La idea que los matemáticos tenían de su Ciencia cambió de forma radical como consecuencia de la invención en el siglo XIX de las geometrías no euclídeas por [Janos Bolyai](#) (1802 - 1860) y [Nikolai I. Lobachevsky](#) (1792 - 1856). Con la aparición de las geometrías no euclídeas, se han ido introduciendo en matemáticas nuevos conceptos y desarrollos que no tienen una contrapartida inmediata en el mundo real. Quedó claro a partir de entonces que las matemáticas no son una Ciencia de la Naturaleza, que la definición usual de las matemáticas como la ciencia que estudia la cantidad y la forma es inadecuada, y pasó a considerarse que la matemática es la ciencia que obtiene conclusiones lógicas de sistemas axiomáticos. Las matemáticas son, pues, una ciencia puramente deductiva. Una teoría matemática es un conjunto de axiomas que contienen ciertos términos indefinidos, y un sistema de reglas de inferencia lógica. El papel que juegan las definiciones en una teoría matemática consiste en crear nuevos objetos matemáticos y precisar su significado en dicha teoría. Todos los objetos que se estudian en una teoría matemática, o bien son términos indefinidos de dicha teoría o son objetos creados por medio de definiciones que remiten a los axiomas. En el XVIII los números reales son algo dado y externo que las matemáticas deben explicar, al final del XIX los números serán algo completamente diferente.

La idea de número real es el soporte de otras ideas básicas del Cálculo como las de continuidad y límite. Los procesos de convergencia dependen de la propiedad de completitud de los números reales. Por todo ello, los matemáticos eran cada vez más conscientes de que los progresos del Cálculo dependían de un mejor conocimiento de los mismos.

1.9. Cortaduras de Dedekind

A mediados del siglo XIX no era posible demostrar algunos resultados básicos del cálculo; por ejemplo, que toda función creciente y acotada tiene límite, o el teorema del valor intermedio para funciones continuas. Ello se debía a que faltaba codificar matemáticamente una propiedad fundamental de los números reales, la que ahora llamamos *completitud* y entonces se llamaba *propiedad de continuidad*. En 1872 se publicaron dos trabajos, uno de Cantor y otro de Dedekind, en los que, tomando como punto de partida el sistema de los números racionales, cada autor desarrollaba una construcción matemática de los números reales. Nos vamos a ocupar aquí del trabajo de Dedekind, titulado *Continuidad y números irracionales*. En dicho trabajo, Dedekind manifiesta su propósito de reducir los números reales a la aritmética, eliminando así todo contenido geométrico en la idea de número real. Para explicar lo que él hizo vamos a partir de la intuición de una recta.



Figura 8. Dedekind

Una recta es un ejemplo claro de continuidad. Elegido un punto como origen y un segmento como unidad, podemos hacer corresponder a cada número racional un punto de esa recta. Ya hemos visto, al hablar de las magnitudes inconmensurables, que los números racionales no agotan todos los puntos de la recta; cualquier punto que corresponda con un segmento de longitud inconmensurable con la unidad elegida no puede ser representado por un número racional, es decir, en la recta racional hay “huecos”. Por tanto, los números racionales no son suficientes para describir numéricamente “el continuo”. Se pregunta Dedekind:

¿En qué consiste esta continuidad? Todo depende de la respuesta a esta pregunta, y solamente a través de ella obtendremos una base científica para la investigación de todos los dominios continuos. Con vagas observaciones sobre la unión sin rotura de las partes más pequeñas, obviamente nada se gana; el problema es indicar una característica precisa de la continuidad que pueda servir como base para deducciones válidas. Durante largo tiempo he meditado sobre esto en vano, pero finalmente he encontrado lo que pretendía.

Dedekind se dispone a revelar el secreto, pero como su idea además de ser genial es muy sencilla, previene al lector con esta observación.

Muchos de mis lectores quedarán grandemente disgustados al saber que por esta vulgar observación se revela el secreto de la continuidad.

¿Cuál es esa *vulgar* observación? Vamos a explicarla. Todo punto en una recta R la divide en dos partes disjuntas, la parte A , formada por los puntos de la recta que están a su izquierda, y la parte B ,

formada por los puntos de la recta que están a su derecha. El propio punto podemos incluirlo bien en A o en B . Dice Dedekind:

He encontrado la esencia de la continuidad en el recíproco, es decir, en el siguiente principio: "Si todos los puntos de la recta se dividen en dos clases tales que todo punto de la primera clase queda a la izquierda de todo punto de la segunda clase, entonces existe un, y sólo un punto, que produce esta división de todos los puntos en dos clases, esta escisión de la línea recta en dos partes."

Las ideas geniales, que además son sencillas, son doblemente geniales. Igual que el tiempo es continuo porque entre dos instantes de tiempo solamente hay tiempo, la recta es continua porque entre dos puntos de ella solamente hay puntos de la misma recta. Es esta la idea que Dedekind ha sabido expresar matemáticamente de una forma insuperable. Para entenderla un poco mejor, vamos a considerar el conjunto \mathbb{Q} de los números racionales como puntos de una recta en la que hemos elegido un origen y una unidad, la recta racional.

Definición. Una *cortadura* de \mathbb{Q} es un par (A, B) , donde A y B son conjuntos no vacíos de números racionales tales que $\mathbb{Q} = A \cup B$, y todo número de A es menor que todo número de B y A no tiene máximo.

Todo número racional $r \in \mathbb{Q}$ produce una cortadura dada por

$$A = \{x \in \mathbb{Q} : x < r\}, \quad B = \{x \in \mathbb{Q} : r \geq x\}$$

Pero en la recta racional hay muchas cortaduras que no están producidas por números racionales. Es un sencillo ejercicio probar que los conjuntos

$$A = \{x \in \mathbb{Q} : x \leq 0 \text{ o } x^2 < 2\}, \quad B = \{x \in \mathbb{Q} : x > 0 \text{ y } x^2 \geq 2\}$$

definen una cortadura de \mathbb{Q} que no está producida por ningún número racional. De hecho, si te imaginas la recta racional dentro de la recta real, y tomas un número α que sea irracional, los conjuntos

$$A = \{x \in \mathbb{Q} : x < \alpha\}, \quad B = \{x \in \mathbb{Q} : r > \alpha\}$$

Definen una cortadura de \mathbb{Q} que no está producida por ningún número racional. Es decir, considerando \mathbb{Q} dentro de \mathbb{R} , vemos que cada cortadura de \mathbb{Q} está determinada por un punto que puede ser racional o irracional.

Pero claro, *está prohibido usar la recta real cuando lo que queremos es justamente construirla a partir de \mathbb{Q}* . ¿De dónde sacamos los números reales si todo lo que tenemos son los racionales? Esta es la idea genial de Dedekind.

Definición. Un número real es una cortadura de \mathbb{Q} .

El conjunto de todos los números reales se representa por \mathbb{R} . Observa el papel que desempeñan las definiciones en una teoría matemática: crean nuevos objetos de la teoría. La definición anterior dice lo que es un número real en términos exclusivamente de números racionales. Vuelve ahora a leer la definición de Eudoxo (1) para la igualdad de razones incommensurables. ¡Lo que dice (1) es que dos razones incommensurables son iguales si producen una misma cortadura en \mathbb{Q} ! Salvo esto, ningún otro parecido hay entre Dedekind y Eudoxo.

Los números racionales se construyen a partir del conjunto \mathbb{Z} de los enteros, y éstos se obtienen fácilmente a partir de los naturales. Dedekind y [Giuseppe Peano](#) establecieron una base axiomática para el conjunto \mathbb{N} de los números naturales. Ya ves, al final, Pitágoras ha regresado: todo es número.

1.10. Métodos axiomáticos y métodos constructivos

Supongo lo que estás pensando: “¡Vaya definición extraña de número real! Ahora resulta que un número es una cortadura. . . ¡nada menos que dos conjuntos infinitos de números!”. Vayamos poco a poco.

- Es una definición operativa, es decir, permite definir la suma y el producto de números reales, así como la relación de orden y demostrar que verifican las propiedades que definen un cuerpo ordenado completo. Además, todo esto se hace de forma sencilla aunque laboriosa. Si tienes curiosidad, puedes consultar el Capítulo 28 de [21].
- Lo importante de la definición es que define los números reales solamente usando los números racionales. Es decir, resuelve un problema de *existencia* en sentido matemático.

Dos cuerpos ordenados completos son matemáticamente indistinguibles pues se demuestra que entre ellos hay un isomorfismo creciente. Por tanto, *los números reales son el único cuerpo ordenado completo*. La demostración de que *existe* un cuerpo ordenado completo y es *único* es larga, laboriosa y depende de las hipótesis de partida.

Lo más usual es dar por conocidos los números racionales y a partir de ellos *construir* \mathbb{R} . Esto puede parecer extraño a primera vista, porque si sólo conocemos los números racionales, ¿de dónde van a salir los demás? De eso precisamente se ocupan los *métodos constructivos* (Cantor, Dedekind). Por ejemplo, si partimos de la intuición de que con los números reales se pueden representar *todos* los puntos de una recta, es claro que un número real queda determinado de forma única por los números racionales menores que él. Esta idea conduce a la *definición* de número real dada

por Dedekind. La definición de Cantor es mucho menos intuitiva pues, para Cantor, un número real es una clase de infinitas sucesiones de números racionales que cumplen una cierta propiedad.

Es posible probar, partiendo de estas definiciones, que el conjunto de los números reales así definidos puede dotarse de una estructura algebraica y de orden de manera que satisface los axiomas de un cuerpo ordenado completo. Este proceso es bastante laborioso; además se corre el peligro de centrar la atención en el proceso en sí mismo olvidándose de lo que se persigue. Por otra parte, las definiciones de Dedekind o de Cantor no son las únicas, hay otras definiciones de número real. Pensarás que esto no es serio. ¿Qué está ocurriendo aquí? Ocorre, sencillamente, que cualquier definición de los números reales a partir de los racionales, esto es, cualquier método constructivo de \mathbb{R} , tiene su razón última de ser en el *problema de la existencia*: ¿puede ser construido un cuerpo ordenado completo a partir de los axiomas usuales de la Teoría de Conjuntos? Pues bien, la respuesta es que sí; además, y esto es fundamental, matemáticamente, en un sentido preciso, dicho cuerpo *es único*.

Da igual, por tanto, cómo se interprete lo que es un número real, lo importante es que de cualquier forma que lo hagamos, los axiomas que definen un cuerpo ordenado completo determinan totalmente sus propiedades matemáticas. Es decir, una vez que sabemos que hay un único cuerpo ordenado completo, lo mejor es olvidar cualquier posible interpretación de cómo sean sus elementos (ningún matemático piensa que $\sqrt{2} = \{x \in \mathbb{Q} : x < 0 \text{ o } x^2 < 2\}$) y quedarnos exclusivamente con las propiedades de los mismos. Esto es precisamente lo que se hace con el método axiomático que es la forma más conveniente de iniciarse en el estudio de los números reales.

Para más detalles sobre la fundamentación del sistema de los números reales puede consultarse el capítulo 41 de [12].

1.11. El regreso de los pequeñitos

Con la reducción del continuo a lo discreto, parece que finalmente ha triunfado la Aritmética. Pero la historia continúa. Por una parte, los números naturales tuvieron un reinado efímero, pues fueron esencialmente reducidos a pura lógica como consecuencia del trabajo pionero de [Gottlob Frege](#). Por otra parte en 1960, el lógico [Abraham Robinson](#) (1918 - 1974) construyó un sistema numérico, los *hiperreales*, un cuerpo totalmente ordenado no arquimediano, que contiene una copia de los números reales y en el que hay números infinitamente pequeños y números infinitamente grandes. Las técnicas desarrolladas por Robinson se conocen con el nombre de *Análisis No Estándar*. Con dichas técnicas pueden probarse los resultados fundamentales del Cálculo de forma intuitiva y directa al estilo de Newton y Leibniz. ¡Están aquí! ¡Los infinitésimos han regresado!

1.12. Otros números: complejos, cuaterniones, octoniones.

Acabamos de recorrer el largo camino que lleva del descubrimiento de las cantidades inconmensurables hasta la formalización matemática del concepto de número real. Pero, además de los naturales, enteros, racionales e irracionales, hay otros números, algunos bien conocidos, otros no tanto.

Los números que hoy llamamos “complejos” fueron durante muchos años motivo de polémicas y controversias entre la comunidad científica. Poco a poco, por la creciente evidencia de su utilidad, acabaron por ser comúnmente aceptados, aunque no fueron bien comprendidos hasta épocas recientes. Nada hay de extraño en ello si pensamos que los números negativos no fueron plenamente aceptados hasta finales del siglo XVII.

Los números complejos hacen sus primeras tímidas apariciones en los trabajos de Cardano (1501-1576) y Bombelli (1526-1672) relacionados con el cálculo de las raíces de la cúbica o ecuación de tercer grado. Fue René Descartes (1596-1650) quien afirmó que “*ciertas ecuaciones algebraicas sólo tienen solución en nuestra imaginación*” y acuñó el calificativo “*imaginarias*” para referirse a ellas. Desde el siglo XVI hasta finales del siglo XVIII los números complejos o imaginarios son usados con recelo, con desconfianza. Con frecuencia, cuando la solución de un problema resulta ser un número complejo se interpreta esto como que el problema no tiene solución. Para Leibniz “*el número imaginario es un recurso sutil y maravilloso del espíritu divino, casi un anfibio entre el ser y el no ser.*”

Las razones de todo esto son claras. Así como los números reales responden al problema bien cotidiano de la medida de magnitudes, no ocurre nada similar con los números complejos. Mientras los matemáticos necesitaron interpretar en términos físicos sus objetos de estudio, no se avanzó mucho en la comprensión de los números complejos.

El éxito de Euler y Gauss al trabajar con números complejos se debió a que ellos no se preocuparon de la “*naturaleza*” de los mismos; no se preguntaron “¿qué es un número complejo?”, sino que se dijeron “*a ver, para qué sirven, qué puede hacerse con ellos*”. Es Gauss quien definitivamente concede a los números complejos un lugar privilegiado dentro de las matemáticas al probar en 1799 el resultado conocido como *Teorema Fundamental del álgebra* que afirma que toda ecuación polinómica de grado n con coeficientes complejos tiene, si cada raíz se cuenta tantas veces como su orden, n raíces que *también son números complejos*.

El término, hoy usado de “*números complejos*” se debe a Gauss, quien también hizo popular la letra “ i ” que Euler (1707-1783) había usado esporádicamente. En 1806 Argand interpreta los números complejos como vectores en el plano. La fecha de 1825 es considerada como el nacimiento de la teoría de funciones de variable compleja, pues se publica en dicho año la Memoria sobre la

Integración Compleja que Cauchy había escrito ya en 1814.

Los números complejos son una herramienta básica de cálculo. Son especialmente útiles para trabajar con funciones sinusoidales, y por eso se hace uso constante de ellos siempre que representamos una señal por medio de dichas funciones, y no hay que olvidar que ése es el propósito básico de los “*métodos de Fourier*”. La *Transformada de Fourier Discreta*, una herramienta fundamental en el tratamiento digital de señales, toma valores complejos. Las *transformadas de Fourier y de Laplace* son funciones complejas. La *transformada z*, al igual que otras transformadas de uso frecuente, se define como una serie de números complejos. La función exponencial compleja desempeña un papel fundamental en el estudio de los sistemas LTI (sistemas lineales invariantes en el tiempo) y también en la teoría de las ecuaciones diferenciales lineales. Los formatos digitales más frecuentes de audio e imagen son, respectivamente, MP3 y JPG. Cuesta trabajo imaginar cómo sería Internet sin estos formatos. Lo que quizás no sepas es que la codificación MP3 y la JPG se llevan a cabo con algoritmos que usan números complejos. El hecho, por extraño que pueda parecer, es que las principales herramientas para trabajar con todo tipo de señales (audio, vídeo, voz, imagen, . . .) son complejas.

El Teorema Fundamental del álgebra nos dice que el procedimiento de ir ampliando el campo numérico con el objetivo de que toda ecuación polinómica tenga soluciones alcanza su cima con los números complejos y que no aparecerán ya por este procedimiento nuevos tipos de números. Pero podemos preguntarnos si, al igual que hemos definido una multiplicación en $\mathbb{R} \times \mathbb{R}$ para obtener \mathbb{C} , es posible hacer algo *similar* en \mathbb{R}^n de manera que resulte una estructura a cuyos elementos podamos llamar *números*. Para ello tendremos que ponernos de acuerdo en qué cosa queremos que sea un *número*. Lo que está claro es que, sean lo que sean, los números se deben poder sumar y multiplicar; la adición debe tener las propiedades usuales, la multiplicación debe ser distributiva respecto de la adición y todo número distinto de cero debe tener un inverso. La codificación de estas propiedades conduce al concepto de *álgebra real no asociativa de división*: un espacio vectorial real, A , dotado de una aplicación bilineal $A \times A \rightarrow A$, $(a, b) \mapsto ab$, llamada *producto*. Debemos exigir que A tenga unidad, es decir, que haya un elemento $u \in A$, $u \neq 0$, tal que $au = ua = a$ para todo $a \in A$. Además, todo elemento $a \neq 0$ debe tener un inverso $b \in A$ tal que $ab = ba = u$.

Claramente, \mathbb{R} y \mathbb{R}^2 (con el producto complejo) son ejemplos de álgebras reales de división cuyo producto tiene las propiedades asociativas y conmutativas.

En \mathbb{R}^4 pongamos $\mathbf{1} = (1, 0, 0, 0)$, $\mathbf{i} = (0, 1, 0, 0)$, $\mathbf{j} = (0, 0, 1, 0)$, $\mathbf{k} = (0, 0, 0, 1)$ y definamos un producto en el que $\mathbf{1}$ hace de unidad y:

$$\mathbf{i}^2 = \mathbf{j}^2 = \mathbf{k}^2 = -\mathbf{1}, \quad \mathbf{ij} = -\mathbf{ji} = \mathbf{k}, \quad \mathbf{jk} = -\mathbf{kj} = \mathbf{i}, \quad \mathbf{ki} = -\mathbf{ik} = \mathbf{j}$$

extendiéndolo por bilinealidad a \mathbb{R}^4 . Es fácil probar que:

$$(a + b\mathbf{i} + c\mathbf{j} + d\mathbf{k})(a - b\mathbf{i} - c\mathbf{j} - d\mathbf{k}) = (a^2 + b^2 + c^2 + d^2)\mathbf{1}$$

De esta forma obtenemos un álgebra real de división de dimensión 4 llamada el álgebra de los **cuaterniones** de Hamilton que se representa por \mathbb{H} . El producto de dicha álgebra es asociativo pero claramente no es conmutativo.

Podemos representar los elementos de \mathbb{R}^8 en la forma

$$\left\{ \begin{pmatrix} z & \mathbf{w} \\ -\bar{\mathbf{w}} & \bar{z} \end{pmatrix} : z \in \mathbb{C}, \mathbf{w} \in \mathbb{C}^3 \right\}$$

donde, para un vector $\mathbf{w} = (w_1, w_2, w_3) \in \mathbb{C}^3$, $\bar{\mathbf{w}} = (\bar{w}_1, \bar{w}_2, \bar{w}_3)$. Se entiende que dicha matriz representa el vector de \mathbb{R}^8 cuyas coordenadas son

$$(\operatorname{Re}(z), \operatorname{Im}(z), \operatorname{Re}(w_1), \operatorname{Im}(w_1), \operatorname{Re}(w_2), \operatorname{Im}(w_2), \operatorname{Re}(w_3), \operatorname{Im}(w_3)).$$

En \mathbb{R}^8 definimos un producto por:

$$\begin{aligned} \begin{pmatrix} z_1 & \mathbf{w}_1 \\ -\bar{\mathbf{w}}_1 & \bar{z}_1 \end{pmatrix} \begin{pmatrix} z_2 & \mathbf{w}_2 \\ -\bar{\mathbf{w}}_2 & \bar{z}_2 \end{pmatrix} &= \\ &= \begin{pmatrix} z_1 z_2 - \langle \mathbf{w}_1 | \bar{\mathbf{w}}_2 \rangle & z_1 \mathbf{w}_2 + \bar{z}_2 \mathbf{w}_1 + \bar{\mathbf{w}}_1 \times \bar{\mathbf{w}}_2 \\ -\bar{z}_1 \bar{\mathbf{w}}_2 - z_2 \bar{\mathbf{w}}_1 - \mathbf{w}_1 \times \mathbf{w}_2 & \bar{z}_1 \bar{z}_2 - \langle \bar{\mathbf{w}}_1 | \mathbf{w}_2 \rangle \end{pmatrix} \end{aligned}$$

donde, para vectores $\mathbf{u} = (u_1, u_2, u_3), \mathbf{v} = (v_1, v_2, v_3) \in \mathbb{C}^3$,

$$\langle \mathbf{u} | \mathbf{v} \rangle = u_1 v_1 + u_2 v_2 + u_3 v_3$$

y

$$\mathbf{u} \times \mathbf{v} = (u_2 v_3 - u_3 v_2, u_3 v_1 - u_1 v_3, u_1 v_2 - u_2 v_1)$$

El producto así definido en \mathbb{R}^8 es bilineal, tiene unidad $\mathbf{1} = (1, 0, 0, 0, 0, 0, 0, 0)$, no es conmutativo y no es asociativo, aunque verifica las identidades $\mathbf{a}^2 \mathbf{b} = \mathbf{a}(\mathbf{a}\mathbf{b})$ y $\mathbf{b}\mathbf{a}^2 = (\mathbf{b}\mathbf{a})\mathbf{a}$ para todos $\mathbf{a}, \mathbf{b} \in \mathbb{R}^8$. Un álgebra en la que el producto verifica estas identidades se dice que es un **álgebra alternativa**. Además se verifica que

$$\begin{pmatrix} z & \mathbf{w} \\ -\bar{\mathbf{w}} & \bar{z} \end{pmatrix} \begin{pmatrix} \bar{z} & -\mathbf{w} \\ \bar{\mathbf{w}} & z \end{pmatrix} = (z\bar{z} + \langle \mathbf{w} | \bar{\mathbf{w}} \rangle) \mathbf{1}$$

de donde se sigue que es un álgebra de división. Dicha álgebra se llama álgebra de los **octoniones** de Cayley y es un álgebra real de división de dimensión 8 que se representa por \mathbb{O} . Tenemos, pues, cuatro álgebras reales de división, $\mathbb{R}, \mathbb{C} = \mathbb{R}^2$ (con el producto complejo), \mathbb{H} y \mathbb{O} , de dimensiones respectivas 1, 2, 4 y 8. La pregunta es obligada ¿hay más álgebras reales de división? La respuesta la dieron F.G. Frobenius (1878) y M. Zorn (1930) probando que, salvo isomorfismos, no hay más álgebras reales de división alternativas de dimensión finita que $\mathbb{R}, \mathbb{C}, \mathbb{H}$ y \mathbb{O} . Otro resultado definitivo fue demostrado por R. Bott y J. Milnor (1958) probando que solamente es posible definir una estructura de álgebra de división en \mathbb{R}^n para los valores $n = 1, 2, 4, 8$.

2. Evolución del concepto de función y límite funcional

Lo más específico del Análisis Matemático son los procesos de convergencia, o procesos “de paso al límite”, que en él se consideran. Aquí nos vamos a ocupar solamente del concepto de límite funcional. Dicho concepto está estrechamente relacionado con los de función y de número real; y los tres juntos constituyen el núcleo del Análisis. Por ello, la historia de su evolución es también la del desarrollo del Cálculo, de los sucesivos intentos para fundamentarlo sobre bases lógicas rigurosas. Aislar en este proceso aquellos aspectos directamente relacionados con el concepto de límite funcional, conlleva una pérdida de perspectiva que, espero, quedará compensada cuando estudiemos la evolución de los conceptos de derivada, integral y convergencia de series.

2.1. La teoría de las “razones últimas” de Newton

En las matemáticas de la Antigüedad no existía una idea de “límite” que pueda ser considerada como un precedente lejano de la actual. Lo más parecido era el método de exhaustión empleado con maestría por Arquímedes para realizar diversas cuadraturas. Pero dicho método no consistía en un límite, sino que, precisamente, lo que hacía era evitarlo y sustituirlo por un esquema de razonamiento de doble reducción al absurdo, típico de las matemáticas griegas. La matemática Griega abomina del infinito y la idea de límite connota la de infinito. Es notable, sin embargo, que cuando los matemáticos Griegos tienen que enfrentarse al infinito como, por ejemplo, Eudoxo al definir la igualdad de razones de magnitudes inconmensurables (1), lo que hace es basar su definición de *igualdad* en un álgebra de *desigualdades*.

Tenemos que llegar al siglo XVII, con la invención de las técnicas infinitesimales que preludian el descubrimiento del Cálculo, para encontrar las primeras referencias confusas de procesos de convergencia. El primer indicio del concepto de límite funcional aparece en estrecha relación con el cálculo de *fluxiones* (velocidades instantáneas) de Newton. En su teoría de las “razones últimas” expuesta en *Philosophiae Naturalis Principia Mathematica* (1687) se lee:

También puede alegarse que si las razones últimas de cantidades evanescentes son dadas, sus últimas magnitudes también serán dadas; y por tanto toda cantidad consistirá de indivisibles, en contra de lo que Euclides ha probado. . . Pero esta objeción está basada sobre una hipótesis falsa. Aquellas razones últimas con las que tales cantidades desaparecen no son en realidad razones de cantidades últimas, sino límites. . . a los que ellas pueden aproximarse tanto que su diferencia es menor que cualquier cantidad dada. . . Este asunto será entendido más claramente en el caso de cantidades indefinidamente grandes. Si dos cantidades cuya diferencia es dada son indefinidamente aumentadas, su última razón será dada, a saber, la razón de igualdad y, no obstante, las cantidades últimas o máximas de las cuales esta es la razón no serán por eso dadas.

Lo que yo entiendo que quiere decir Newton es lo que sigue. La expresión “razones últimas de

cantidades evanescentes” puede interpretarse como el límite de un cociente cuyo numerador y denominador tienen límite cero: $\lim_{x \rightarrow a} \frac{f(x)}{g(x)} = L$, donde $\lim_{x \rightarrow a} f(x) = \lim_{x \rightarrow a} g(x) = 0$. En el primer párrafo, Newton dice que el hecho de que la razón última sea dada igual a L , no quiere decir que el cociente de las últimas magnitudes, $\frac{f(a)}{g(a)}$, sea igual a L . De manera muy interesante, Newton relaciona esto con la estructura del continuo, pues la idea que expresa es que si el valor de todo límite se alcanza, entonces el continuo estaría formado por últimas partes indivisibles. En el segundo párrafo, además de insistir en la idea anterior, queda claro que por “razones últimas” Newton entendía algo muy parecido a nuestra idea actual de límite. Finalmente, Newton propone un ejemplo excelente; consideremos, dice, dos cantidades $f(x)$ y $g(x)$ cuya diferencia está dada, $f(x) - g(x) = \alpha \neq 0$, y tales que $\lim_{x \rightarrow a} f(x) = \lim_{x \rightarrow a} g(x) = +\infty$, en tal caso tendremos que su razón última será de igualdad, esto es, $\lim_{x \rightarrow a} \frac{f(x)}{g(x)} = 1$ y está claro que para ningún valor de x es $\frac{f(x)}{g(x)} = 1$ y que tampoco las magnitudes $f(x)$ y $g(x)$ tienen un último valor.

Siempre es arriesgado hacer interpretaciones de esta naturaleza, pero creo que lo dicho es esencialmente correcto y, por tanto, manifiesto mi desacuerdo con quienes afirman que Newton tenía ideas muy confusas con respecto al límite. Lo que no tenía (no podía tener) era el concepto de función (por eso habla de cantidades o magnitudes), ni el simbolismo apropiado, ni el concepto de variable real continua. . . pero la idea de límite la tenía bien clara.

Además, Newton considera que los infinitésimos no son cantidades fijas y, en los *Principia*, advierte a sus lectores que cuando hable de cantidades mínimas, o evanescentes, o de cantidades últimas, éstas no debieran entenderse como cantidades fijas que tienen un determinado valor, sino como cantidades que fueran indefinidamente disminuidas.

Estas ideas de Newton fueron desarrolladas por el matemático escocés Colin MacLaurin (1698 - 1746) que, en su gran obra *A Treatise of Fluxions* (1742), establece el cálculo sobre la base de una teoría geométrica – cinemática de límites. MacLaurin rechazaba los infinitésimos, afirmaba que los antiguos nunca reemplazaron curvas por polígonos y que la base de la geometría de Arquímedes era el concepto de límite. Lo sorprendente es que MacLaurin usa el concepto de límite como algo evidente que no precisa ser explícitamente presentado ni analizado. Esto se debe a que el cálculo de MacLaurin se sustenta sobre las ideas de espacio, tiempo y movimiento lo que le lleva a aceptar como evidentes la continuidad y la diferenciabilidad.

2.2. Sobre el concepto de función

Estamos acostumbrados a usar la idea de “función” para expresar una relación de dependencia entre varias magnitudes; por ejemplo, decimos que “los precios están en función de los costes de producción”. Toda persona con conocimientos básicos sabe que las derivadas y las integrales

son herramientas que se usan para estudiar funciones. Las funciones no solamente se estudian en Cálculo; en todas las ramas de las Matemáticas se estudian funciones de distintos tipos, y puede afirmarse que el concepto de función constituye un vínculo unificador entre todas ellas.

El desarrollo de la geometría analítica llevó a la introducción de variables continuas y a un punto de vista dinámico entre sus relaciones las cuales se expresaban por medio de ecuaciones. Una ecuación no es una función, por eso el Cálculo de Newton y Leibniz no es un cálculo de funciones. En el primer libro de Cálculo *Analyse des infiniment petits, pour l'intelligence des lignes courbes* (L'Hôpital, 1696), como ya se indica en su propio título, lo que se estudia son curvas, no funciones. Las variables asociadas a una curva eran de tipo geométrico: tangentes, subtangentes, normales, radio de curvatura... , y no se establecía una relación de dependencia de unas variables (dependientes) con respecto a otras variables (libres), porque en los problemas matemáticos y físicos que se estudiaban no era preciso hacer esa distinción⁴.

El método de fluxiones de Newton se aplica a *fluentes* (que sugiere la idea geométrica de un punto que “fluye” a lo largo de una curva), no a funciones. La principal contribución de Newton al desarrollo del concepto de función fue su uso de las series de potencias.

En 1692, Leibniz introdujo el término “función” para designar un objeto geométrico asociado con una curva. Por ejemplo, Leibniz afirmaba que “una tangente es función de una curva”. Al mismo tiempo generalizó el uso de los términos “constante”, “variable”, “coordenadas” y “parámetro”. La carencia de un término específico para representar cantidades que dependen de otras cantidades en las ecuaciones que las relacionan dio lugar al uso del término “función” (del latín *functio*; *fungor, functus sum* significa algo así como *llevar a cabo, cumplir con una obligación*) tal como aparece en la definición dada por Johann Bernoulli en 1718 ([7], pg.9), ([26], pg. 60):

Llamo función de una magnitud variable a una cantidad compuesta de cualquier manera a partir de esta magnitud variable y de constantes.

Esta fue la primera definición formal de función, aunque Johann Bernoulli no explicó lo que significaba “compuesta de cualquier manera”, posiblemente se refiere a las expresiones analíticas conocidas en su época⁵.

Descartes había clasificado las curvas en “geométricas” y “mecánicas”; las primeras son aquellas que podían ser estudiadas por sus métodos algebraicos mientras que las segundas no. Leibniz dividió las funciones y curvas en dos clases: *algebraicas*, aquellas que podían ser representadas

⁴Aunque Newton consideraba que sus “fluentes” eran dependientes del tiempo, su punto de vista era esencialmente cinemático y geométrico más que funcional.

⁵Parece que la primera aproximación a un concepto general de función como una expresión “analítica” en la que puede intervenir un proceso infinito, es debida a J. Gregory en su libro *Veritable quadrature of the circle and hyperbola* (*Vera circuli et hyperbolae quadrature*), publicado en 1667 ([26], pg. 58).

por una ecuación de un cierto orden, y *trascendentes*, las que se representan por ecuaciones de un grado indefinido o infinito.

El siguiente importante avance en el concepto de función fue debido al gran Leonhard Euler, discípulo de Johann Bernoulli. En el Capítulo I del Volumen I de su *Introductio in analysin infinitorum*, publicado en 1748, Euler hace un estudio detallado del concepto de función. Empieza definiendo lo que entiende por “constante” y “variable”. Una constante es una cantidad definida que representa siempre uno y el mismo valor. Pero

Una cantidad variable es una indeterminada o cantidad universal que puede tomar por ella misma absolutamente todos los valores determinados.

Para Euler una variable puede tomar cualquier valor numérico, positivo o negativo, racional o irracional, y explícitamente escribe: *Quin etiam cyphra et numeri imaginarii a significatu quantitatis varabilis non excluduntur*, esto es, no se excluyen ni el cero ni los números imaginarios del significado de una cantidad variable. Siguiendo a su maestro, Euler da la siguiente definición de función:

Una función de una cantidad variable es cualquier expresión analítica formada a partir de dicha cantidad variable y números o cantidades constantes.

La diferencia con la definición de J. Bernoulli está en el término “expresión analítica”. Después de dar esta definición, Euler distingue entre varios tipos de funciones según que puedan o no representarse por medio de una sola expresión analítica. También fue Euler quien usó por primera vez la notación $f(x)$ para indicar el valor de una función f en un valor x de la variable. Euler no precisaba lo que entendía por “cualquier expresión analítica” pero, sin duda, incluía las series, fracciones continuas y productos infinitos y primitivas. El concepto de función de Euler incluía tanto funciones “explícitas” como “implícitas”. Estas últimas aparecen en la resolución de ecuaciones algebraicas.



Figura 9. Euler

El libro de Euler *Introductio in analysin infinitorum* es considerado como el tercero más influyente en toda la historia de las matemáticas (el primero serían los *Elementos* de Euclides (300 adC) y el segundo los *Principia* (1687) de Newton) y tuvo una amplia difusión. En el prefacio de dicho libro, Euler, afirmaba que el Análisis Matemático es la ciencia general de las variables y sus funciones.

Esto, que hoy día nos parece una evidencia, estaba muy lejos de serlo en el siglo XVIII. De hecho, matemáticos como Newton, Leibniz, los hermanos Bernoulli y otros muchos en los siglos XVII y XVIII, se expresaban en términos de curvas, superficies, áreas, líneas tangentes.

En el Capítulo IV de su *Introductio*, afirma Euler que la forma más conveniente de expresión analítica para representar una función es por medio de una serie de potencias del tipo

$$A + Bz + Cz^2 + Dz^3 + \dots$$

Puesto que la mayoría de funciones consideradas en esta época eran analíticas salvo en algunos puntos aislados de su dominio, esta afirmación de Euler era esencialmente correcta y fue aceptada por muchos matemáticos. De esta forma la definición de función dada por Euler se concretó en la práctica en el actual concepto de función analítica. De hecho, estas son las únicas funciones estudiadas en el Volumen I de la *Introductio*. En el Volumen II, dedicado al estudio de las curvas planas, Euler hace una distinción entre *curvas continuas* y *discontinuas*, las primeras son aquellas que vienen dadas por una única expresión analítica, las segundas se corresponden con lo que hoy entendemos por función analítica definida a trozos. Este uso tan particular de los términos “continuo” y “discontinuo” dio lugar a muchas confusiones.

La necesidad de precisar el concepto de función surgió poco después, de forma muy natural, en el estudio de las vibraciones planas de una cuerda elástica tensa, sujeta por sus extremos, cuya posición inicial en el plano viene dada por una función conocida $\psi(x)$. D’Alembert (1749) y Euler (1750) obtuvieron esencialmente la misma solución, pero discreparon sobre el tipo de función inicial $\psi(x)$ permitida. Mientras que, según D’Alembert, la posición inicial debía venir dada por una función suave (derivable dos veces), Euler insistía en que la evidencia física imponía la consideración de funciones más generales (no derivables, con picos). Él mismo propuso como posición inicial de la cuerda una línea poligonal. Otro matemático, Daniel Bernoulli, propuso en 1753 una solución del problema que tenía como consecuencia que la función $\psi(x)$ podía representarse como suma de una serie trigonométrica infinita. Una situación muy similar a ésta se produjo unos años después, en 1822, como consecuencia de los trabajos de Jean B. Joseph Fourier sobre la propagación del calor.

Los detalles de toda esta historia son muy interesantes pero imposibles de resumir en unas pocas líneas. Volveremos sobre ellos más adelante. En esencia, se trata de lo siguiente. En la segunda mitad del siglo XVIII y primera del XIX, al mismo tiempo que los matemáticos seguían considerando que las funciones debían ser continuas y derivables, salvo a lo sumo en una cantidad finita de “puntos especiales” (el mismo Euler tenía esta idea), se estaban desarrollando métodos para resolver problemas cada vez más complejos que permitían representar “funciones cualesquiera” por medio de expresiones analíticas, principalmente, series de Fourier. Se suponía que una representación de este tipo debía “transmitir su regularidad” a la función representada pero, por otra parte,

ésta podía ser muy general. El corazón del problema estaba en la confusión de dos conceptos, aparentemente iguales pero muy distintos de hecho, el de función y el de su representación analítica. La separación de estos conceptos llevó a considerar una función con independencia de su representación analítica. De esta forma una función quedaba reducida a un conjunto de valores numéricos completamente independientes asociados a una o varias variables, que es la idea subyacente a la definición moderna debida a Dirichlet (1837):

“y es una función de una variable x , definida en un intervalo $a < x < b$, si para cada valor de la variable x en este intervalo le corresponde un valor concreto de la variable y . Además, es irrelevante la forma en la que esta correspondencia se establezca.”



Figura 10. Dirichlet

Esta nueva idea de función llevó a investigar nuevos tipos de funciones que, con frecuencia, tenían un comportamiento inusual. En 1854 Riemann dio un ejemplo de función integrable con infinitas discontinuidades en todo intervalo de longitud positiva. En 1872 Weierstrass sorprende a la comunidad matemática con una función continua que no es derivable en ningún punto. A estos ejemplos de funciones “patológicas” pronto les siguen otros. En el siglo XIX la necesidad de una fundamentación rigurosa del Análisis Matemático se hace evidente. El concepto de función sigue en el centro de atención y, aunque dicho concepto siguió discutiéndose casi hasta el final del siglo, hoy se reconoce a Dirichlet haber sido el primero en considerar seriamente la idea de función como una “correspondencia arbitraria”.

2.3. La metafísica del Cálculo en D’Alembert y Lagrange

D’Alembert redactó la mayoría de los artículos de matemáticas y ciencias para la obra inmortal del Siglo de las Luces la *Encyclopédie, ou Dictionnaire Raisonné des Sciences, des Arts et des Métiers* (1751 - 65). En el artículo Différentielle (1754), después de criticar la “metafísica del infinito” de Leibniz, escribe:

Newton partía de otro principio; y se puede decir que la metafísica de este gran geómetra sobre el cálculo de fluxiones es muy exacta y luminosa, aunque solamente la ha dejado entrever. Él no ha considerado nunca el cálculo diferencial como el cálculo de cantidades infinitamente pequeñas, sino como el método de las primeras y últimas razones, es decir, el método para hallar los límites de las razones.

[. . .] La suposición que se hace de las cantidades infinitamente pequeñas sólo sirve para acortar y simplificar los razonamientos; pero en el fondo el cálculo diferencial no precisa suponer la existencia de tales cantidades; y más aún, este cálculo consiste meramente en la determinación algebraica del límite de una razón.



Figura 11. D'Alembert

Durante el siglo XVIII, por una parte, el uso permanente de los infinitesimales dificultaba la comprensión de los procesos de paso al límite y, por otra parte, el recién inventado Cálculo era una herramienta maravillosa para estudiar y formular matemáticamente multitud de fenómenos naturales. Además, los resultados obtenidos eran correctos, por tanto no había por qué preocuparse mucho de la coherencia lógica de los fundamentos, ya habría tiempo para ello más adelante.

Debemos destacar, no obstante, la propuesta de Jean le Rond D'Alembert (1717 - 1783) de fundamentar el Cálculo sobre el concepto de límite: “*La théorie des limites est la base de la vraie Méthaphysique du calcul différentiel*”.

D'Alembert fue el primer matemático que afirmó haber probado que los infinitamente pequeños “*n'existent réellement ni dans la nature, ni dans les suppositions des Géomètres*”. Según D'Alembert:

Una cantidad es algo o nada; si es algo, aún no se ha desvanecido; si es nada, ya se ha desvanecido literalmente. La suposición de que hay un estado intermedio entre estos dos es una quimera.

En el artículo *Limite* (1765), también escrito para la *Encyclopédie* junto con Jean-Baptiste de La Chapelle (1710 - 1792), se da la siguiente definición de límite:

Se dice que una magnitud es el límite de otra magnitud, cuando la segunda puede aproximarse a la primera, sin llegar nunca a excederla, en menos que cualquier cantidad dada tan pequeña como se quiera suponer.

Este artículo también contiene los resultados sobre la unicidad del límite y sobre el límite del producto de dos magnitudes, por supuesto, enunciados retóricamente sin ningún tipo de símbolo para representar los límites. Dichos resultados habían aparecido en el libro de La Chapelle *Institutions de Géométrie* (1757). Tanto D'Alembert como La Chapelle tenían una idea esencialmente geométrica del concepto de límite, así el ejemplo que ponen en el citado artículo es el de la aproximación de un círculo por polígonos.

El punto de vista de D'Alembert, esencialmente correcto, no era compartido por otros matemáticos, de forma destacada, por Joseph-Louis de Lagrange (1736 - 1813) quien en su obra *Théorie des fonctions analytiques* (1797), cuyo subtítulo era nada menos que *Les Principes du Calcul Différentiel, dégagés de toute considération d'infiniment petits, d'évanouissants, de limites et de fluxions, et réduits à l'analyse algébrique des quantités finies*, pretendió establecer una fundamentación algebraica del Cálculo, eliminando toda referencia a los infinitesimales y a los límites. Lagrange criticaba la teoría de las “últimas razones” de Newton y afirmaba:

Ese método tiene el gran inconveniente de considerar cantidades en el momento en que ellas cesan, por así decir, de ser cantidades; pues aunque siempre podemos concebir adecuadamente las razones de dos cantidades en tanto en cuanto ellas permanecen finitas, esa razón no ofrece a la mente ninguna idea clara y precisa tan pronto como sus términos ambos llegan a ser nada a la vez.

Esta severa crítica va realmente dirigida contra Euler, quien concebía las cantidades infinitesimales como ceros exactos y, por tanto, un cociente de diferenciales lo interpretaba como $\frac{0}{0}$, expresión de la cual había que hallar en cada caso “su verdadero valor”. Lo llamativo es que la propuesta de Lagrange se basaba en los desarrollos en series de Taylor, considerados como una generalización del álgebra de polinomios, con lo que, de hecho, estaba usando la idea de límite que quería evitar. Por otra parte, es conocida la jactancia de Lagrange de que en su monumental *Mécanique analytique* (1772 - 88) no había usado ni necesitado ninguna figura. Lagrange seguía así la tendencia, cada vez mayor, de separar el cálculo y la geometría. De hecho, Lagrange puede considerarse un “matemático puro”; su rechazo a la teoría de fluxiones se debe a que está basada en la idea de movimiento, que no es matemática, y su rechazo de los límites es debido a la confusa formulación de dicho concepto en su tiempo.

2.4. El premio de la Academia de Berlín y otras propuestas en el último tercio del siglo XVIII

Jacques-Antoine-Joseph Cousin (1739 - 1800) escribió un libro de texto *Leçons de Calcul Différentiel et de Calcul Intégral* (1777), en el que, siguiendo la idea de D’Alembert, afirmaba fundamentar el cálculo sobre el concepto de límite, el cual, para Cousin, es el mismo que el expresado por La Chapelle en el artículo *Limite* de la *Encyclopédie* antes reseñado. En particular, no hace distinción entre cantidades variables y constantes. Desde un punto de vista operativo, Cousin introduce, sin justificación, un principio de conservación de las razones entre dos variables por paso al límite. Una novedad importante es que Cousin reconoce la necesidad de un símbolo para expresar el límite, pero no hace nada al respecto.

Roger Martin (1741 - 1811) publicó un libro de texto *Éléments de Mathématiques* (1781) con igual propósito que Cousin. La definición de límite de Martin es más precisa:

Por el límite de una cantidad variable se entiende el valor o estado hacia el cual ella siempre tiende conforme varía, sin alcanzarlo nunca; pero al cual, no obstante, puede aproximarse de manera que difiera de él por una cantidad menor que cualquier cantidad dada.

La condición de pequeñez de la diferencia está formulada en la forma que después sería la usual, además, distingue entre variable y valor constante.

En 1784 la Academia de Berlín, cuyo director era Lagrange, anunció la convocatoria de un

premio para “una teoría clara y precisa de lo que se llama el infinito en matemáticas”. El propósito de la Academia era eliminar el uso de los infinitesimales:

Es bien sabido que la geometría superior emplea regularmente lo *infinitamente grande* y lo *infinitamente pequeño*. . . La Academia, en consecuencia, desea una explicación de cómo es posible que se hayan conseguido deducir tantos teoremas correctos a partir de unos presupuestos contradictorios, así como. . . un principio verdaderamente matemático que pueda sustituir correctamente al del infinito.

El premio fue concedido en 1786 a Simon-Antoine-Jean L’Huilier (1750 - 1840) por su ensayo *Exposition élémentaire des principes des calculs supérieurs* en el cual L’Huilier desarrollaba una teoría de límites. Su definición de límite es:

Sea una cantidad variable, siempre menor o siempre mayor que una propuesta cantidad constante; pero de la cual puede diferir menos que cualquier propuesta cantidad menor que ella misma: esta cantidad constante se dice que es el límite por exceso o por defecto de la cantidad variable.

La novedad aquí está en los conceptos de “límite por exceso” y “límite por defecto”. Al introducir esta distinción, L’Huilier observaba que hasta entonces no se había tenido en cuenta el hecho de que la aproximación al límite puede realizarse tanto desde una variable con valores crecientes como desde una variable con valores decrecientes. Por ello, L’Huilier introduce los conceptos de *límite por la derecha* y de *límite por la izquierda*.

En esta obra es donde, por primera, se usa el símbolo “lím.” (con el punto, como si fuera una abreviación de “límite”) para representar el límite, aunque L’Huilier no lo hace de una forma regular.

El principal logro de L’Huilier fue extender la aplicabilidad del concepto de límite. Mientras que sus predecesores habían dado solamente un par de reglas básicas, él realizó un desarrollo más sistemático, probando las reglas del producto y del cociente para límites, y obteniendo la regla de la derivada de un producto por medio de límites.

En esta exposición estoy siguiendo muy de cerca el excelente libro de Schubring [20]. Este autor hace un notable descubrimiento. Se trata de un ensayo de 100 páginas, titulado *Compendio da Theorica dos Limites, ou Introducção ao Methodo das Fluxões*, que fue publicado por la Academia de Ciencias de Lisboa en 1794, aunque su autor Francisco de Borja Garção Stockler (1759 - 1829) lo había presentado ya en 1791. Stockler nació en Lisboa, su padre era alemán y su madre portuguesa. Estudió la carrera militar y también matemáticas en la Universidad de Coimbra. Desarrolló una gran actividad tanto política como científica. La importancia del citado libro de Stockler es que contiene el primer intento de una presentación algebraica del concepto de límite. Stockler tenía un excelente conocimiento de la literatura matemática de su época, y en su libro se apoya precisamente en los autores que hemos citado anteriormente. Pero Stockler aventaja ampliamente

a sus fuentes al separar el concepto de límite del concepto geométrico, algebraizándolo tanto para variables como para funciones. Además, es un pionero en el uso de desigualdades. Su definición de límite es la siguiente:

Una cantidad constante es llamada “Límite” de una variable si la última puede ir aumentando o disminuyendo – aunque sus valores nunca lleguen a ser igual al de la constante – de tal forma que puede aproximar la constante tanto que la diferencia llega a ser menor que cualquier cantidad dada, por pequeña que esta pueda haber sido escogida.

La definición es parecida a la de Martin, aunque hay un mayor énfasis en que el límite es un valor constante. Stockler también usa los conceptos de límites por la derecha y por la izquierda de L’Huillier.

Debemos notar que todas estas definiciones de límite que estamos dando se refieren a variables y que dichas variables suelen interpretarse como cantidades geométricas (áreas, longitudes de arco, medidas de ángulos, etc.). Además, una “cantidad constante” es interpretada generalmente como una cantidad positiva. Con frecuencia se considera que el cero tiene un carácter especial y se dan definiciones específicas para tenerlo en cuenta. Precisamente, eso es lo que hace Stockler introduciendo el concepto de “*variable sin límite de disminución*” con el significado de una variable con límite cero. De esta forma, también evita usar infinitésimos. Stockler establece como un resultado fundamental que

Toda cantidad capaz de un límite, tiene necesariamente que ser igual a su límite, más o menos una cantidad variable sin límite de disminución.

Stockler desarrolla todo un álgebra de límites y no se limita a las operaciones de suma, producto y cociente. He aquí una muestra:

Una potencia a^x , donde $a < 1$ es una constante y x una variable con valores positivos y sin límite de aumento, forma una sucesión nula.

Stockler explica el uso del símbolo “Lim.” para representar límites y lo emplea de forma operativa para permutar límites. Por ejemplo, si $b = \text{Lim. } x$ y a es constante, $\text{Lim. } (a^x) = a^b$.

Stockler no considera solamente límites de variables sino también de funciones. De forma explícita establece la permutabilidad del límite con una función:

El límite de cualquier función Fx de una variable x que es capaz de (tiene) límite, es igual al valor homólogo por la función de su límite.

Simbólicamente, Stockler expresa el teorema como sigue: Para $a = \text{Lim. } x$, se sigue que $\text{Lim. } Fx = Fa$.

2.5. Cauchy y su *Cours d'Analyse* de 1821

A principios del siglo XIX, parecía cada vez más necesario consolidar la enorme cantidad de resultados que ya se habían obtenido usando las técnicas precariamente fundamentadas del cálculo. Había llegado el momento en que se disponía de las herramientas necesarias para desvelar las sutilezas del concepto de límite, cuya lenta y trabajosa evolución a lo largo del siglo XVIII acabamos de ver. Lo que se necesitaba era dar definiciones precisas, simbólicas y operativas, que no estuvieran basadas en intuiciones geométricas ni cinemáticas. Para ello, había que precisar las expresiones vagas que solían usarse, al estilo de “aproximarse más que una cantidad dada, por pequeña que ésta sea”, y dotarlas de un significado matemático preciso que pudiera ser usado para dar demostraciones. Lo que se necesitaba era traducir las definiciones verbales de límite mediante el álgebra de desigualdades que en esa época ya se había desarrollado. Esto puede parecer fácil visto desde nuestra perspectiva actual, pero no lo era en absoluto.

Si recuerdas la actual definición de límite de una función en un punto, puedes comprobar lo abstracta que es: no queda nada en ella de la intuición inicial con la que Newton imaginaba sus “razones últimas”. Es una definición “estática” y todo en ella es aritmética: valor absoluto, desigualdades. . . ¡no contiene ninguna igualdad! Ganamos rigor a costa de la intuición. Quien realizó la hazaña de fundamentar con rigor el cálculo sobre el concepto de límite fue [Augustin - Louis Cauchy](#) (1789 - 1857). Nos vamos a centrar aquí exclusivamente en este aspecto de su obra, de la que volveremos a ocuparnos más adelante. Conviene, no obstante decir, que hay interpretaciones muy distintas de la obra de Cauchy. En particular, se ha escrito mucho sobre el uso que Cauchy hace de los infinitésimos.



Figura 12. Cauchy

Creo que la documentada exposición que hace Schubring en [20] es muy convincente. Su tesis es que Cauchy, por su propia voluntad, nunca hubiera dejado entrar a los infinitésimos en sus libros, pero que se vio en la necesidad de hacerlo por la presión del entorno de l'École Polytechnique donde desempeñaba su labor docente. De todas formas, su concepto de infinitésimo, como veremos enseguida, no es el de una cantidad no nula pero infinitamente pequeña. En su *Cours d'Analyse de l'École Polytechnique* (1821), Cauchy empieza exponiendo su concepto de número, de cantidad y seguidamente, en la página 19, aparecen las siguientes definiciones:

Se llama cantidad *variable* aquella que se considera debe recibir sucesivamente varios valores diferentes unos de otros.[...] Cuando los valores sucesivamente atribuidos a una misma variable se aproximan indefinidamente a un valor fijo, de manera que acaban por diferir de él tan poco como se quiera, éste último es llamado el *límite* de todos los otros.

[...] Cuando los valores numéricos (valores absolutos) sucesivos de una misma variable decrecen indefinidamente, de manera que quedan por debajo de todo número dado, esta variable recibe el nombre de *infinitésimo* o de cantidad *infinitamente pequeña*. Una variable de esta naturaleza tiene por límite a cero.

Cuando los valores numéricos (valores absolutos) sucesivos de una misma variable crecen más y más, de manera que permanecen por encima de todo número dado, se dice que esta variable tiene por límite el *infinito positivo*, indicado por el signo ∞ , cuando se trata de una variable positiva, y el *infinito negativo*, indicado por la notación $-\infty$, cuando se trata de una variable negativa.

Llama la atención en esta definición la idea repetida de “sucesivos valores” que algunos autores interpretan como si Cauchy considerara a las cantidades variables como sucesiones. Aunque sigue siendo una definición verbal, es mucho más precisa que las anteriores y lo importante es la forma en que Cauchy la interpreta por medio del álgebra de desigualdades. Podemos hacernos una idea de la forma de trabajar de Cauchy considerando el siguiente resultado que aparece en la página 54 del *Cours d'Analyse*. Traduzco y hago algunos comentarios que van en cursiva y entre paréntesis.

Teorema I (Cauchy - Cours D'Analyse, p.54). Si para valores crecientes de x , la diferencia

$$f(x+1) - f(x)$$

converge hacia un cierto límite k , la fracción

$$\frac{f(x)}{x}$$

convergerá al mismo tiempo hacia el mismo límite.

Demostración. Supongamos para empezar que la cantidad k tenga un valor finito, y designemos por ϵ un número tan pequeño como se quiera. Puesto que los valores crecientes de x hacen converger la diferencia

$$f(x+1) - f(x)$$

hacia el límite k , se podrá dar al número h un valor suficientemente grande para que, siendo x igual o mayor que h , la diferencia correspondiente esté constantemente comprendida entre los límites

$$k - \epsilon, \quad k + \epsilon.$$

(Este comienzo es impecable y nosotros lo haríamos exactamente igual. Con nuestras notaciones actuales, la hipótesis es que

$$\lim_{x \rightarrow +\infty} (f(x+1) - f(x)) = k.$$

Por tanto, dado $\epsilon > 0$, existe $h > 0$ tal que para todo $x \geq h$ se verifica que $|f(x+1) - f(x) - k| < \epsilon$. Eso es exactamente lo que escribe Cauchy.)

Supuesto esto, si se designa por n un número entero cualquiera, cada una de las cantidades

$$\begin{aligned} & f(h+1) - f(h) \\ & f(h+2) - f(h+1) \\ & \dots\dots\dots \\ & f(h+n) - f(h+n-1) \end{aligned}$$

y, en consecuencia, su media aritmética, a saber

$$\frac{f(h+n) - f(h)}{n}$$

se encontrará comprendida entre los límites $k - \varepsilon$, $k + \varepsilon$. Se tendrá pues

$$\frac{f(h+n) - f(h)}{n} = k + \alpha$$

siendo α una cantidad comprendida entre los límites $-\varepsilon$, $+\varepsilon$. Sea ahora

$$h + n = x$$

La ecuación precedente se convertirá en

$$\frac{f(x) - f(h)}{x - h} = k + \alpha, \quad (2)$$

y se concluirá

$$\begin{aligned} f(x) &= f(h) + (x - h)(k + \alpha) \\ \frac{f(x)}{x} &= \frac{f(h)}{x} + \left(1 - \frac{h}{x}\right)(k + \alpha) \end{aligned} \quad (3)$$

(Hasta aquí, nada que objetar. Todo es correcto.)

Además, para hacer crecer indefinidamente el valor de x , será suficiente hacer crecer indefinidamente el número entero n sin cambiar el valor de h . Supongamos, en consecuencia, que en la ecuación (3) se considera h como una cantidad constante, y x como una cantidad variable que converge hacia el límite ∞ . Las cantidades

$$\frac{f(h)}{x}, \quad \frac{h}{x},$$

encerradas en el segundo miembro, convergerán hacia el límite cero, y el propio segundo miembro hacia un límite de la forma

$$k + \alpha,$$

α estando siempre comprendida entre $-\varepsilon$ y $+\varepsilon$. Por consiguiente, la razón

$$\frac{f(x)}{x}$$

tendrá por límite una cantidad comprendida entre $k - \varepsilon$ y $k + \varepsilon$. Debiendo subsistir esta conclusión, cualquiera que sea la pequeñez del número ε , resulta que el límite en cuestión será precisamente igual a la cantidad k . En otras palabras, se tendrá

$$\lim \frac{f(x)}{x} = k = \lim [f(x+1) - f(x)]. \quad (4)$$

(Seguidamente, Cauchy pasa a considerar los casos en que $k = \infty$ y $k = -\infty$.) □

Esta demostración es notable, por su rigor y también porque *no es correcta*. ¿Serías capaz de explicar a Cauchy dónde está el error en su razonamiento? Por supuesto, lo que interesa aquí es

la forma en que Cauchy traduce los conceptos de límite por medio de desigualdades. El error es anecdótico, además, cuando Cauchy emplea este resultado lo hace siempre en casos en que la tesis es correcta; por ejemplo, para la función $\log x$, se tiene que $\lim_{x \rightarrow +\infty} (\log(x+1) - \log(x)) = 0$ y, por tanto, $\lim_{x \rightarrow +\infty} \frac{\log x}{x} = 0$ lo cual es correcto.

Contraejemplo. Considera la función $f: \mathbb{R}^+ \rightarrow \mathbb{R}$ dada por

$$f(x) = \frac{1}{1-x+E(x)}$$

Donde, como de costumbre, $E(x)$ es la parte entera de x .

Claramente $f(x+1) = f(x)$, por lo que $\lim_{x \rightarrow +\infty} (f(x+1) - f(x)) = 0$. Por otra parte para $x_n = n - \frac{1}{n^2}$ con $n \geq 2$, se tiene que $\frac{f(x_n)}{x_n} = \frac{n^4}{n^3 - 1} \rightarrow +\infty$.

Durante el siglo XVIII, el concepto de continuidad no había merecido nada más que una esporádica atención, y siempre había sido considerado desde un punto de vista filosófico, más como una ley de la naturaleza que como un concepto propiamente matemático. Generalmente la continuidad de una función se entendía en el sentido de Euler, y significaba que dicha función estaba definida por una única expresión analítica. En su *Cours D'Analyse*, Cauchy define el concepto de función continua y, lo que es notable, de función discontinua; y su definición es realmente muy minuciosa. Dice así:

Sea $f(x)$ una función de la variable x , y supongamos que, para cada valor de x comprendido entre ciertos límites dados, esta función admite constantemente un valor único y finito. Si, partiendo de un valor de x comprendido entre estos límites, se atribuye a la variable x un incremento infinitamente pequeño α , la función misma recibirá por incremento la diferencia

$$f(x+\alpha) - f(x)$$

que dependerá a la vez de la nueva variable α y del valor de x . Dicho esto, la función $f(x)$ será, entre los dos límites asignados a la variable x , función *continua* de esta variable, si, para cada valor de x intermedio entre estos límites, el valor numérico (valor absoluto) de la diferencia

$$f(x+\alpha) - f(x)$$

decrece indefinidamente con el de α . En otras palabras, *la función $f(x)$ permanecerá continua con respecto a x entre los límites dados, si, entre estos límites un incremento infinitamente pequeño de la variable produce siempre un incremento infinitamente pequeño de la función.*

Se dice también que la función $f(x)$ es, en un entorno de un valor particular atribuido a la variable x , función continua de esta variable, siempre que ella sea continua entre dos límites de x , por cercanos que estén, que encierren al valor considerado. Finalmente, cuando una función deja de ser continua en el entorno de un valor particular de la variable x , se dice entonces que ella se hace *discontinua* y que para este valor particular de x hay una *solución de continuidad*.

Cauchy da realmente dos definiciones; primero define lo que nosotros llamaríamos “continuidad en un intervalo” y, después, la continuidad puntual. La primera definición ha sido interpretada en el sentido de que lo que Cauchy entiende por continuidad es lo que ahora llamamos “continuidad uniforme”.

Seguidamente a esta definición, Cauchy pasa a estudiar la continuidad de las funciones elementales, considerando en cada caso, los límites entre los que cada función es continua. Después demuestra el teorema de los valores intermedios (teorema de Bolzano) del cual da dos demostraciones. Una que se apoya de forma decisiva en la intuición geométrica y, en una nota al final del texto, otra, que él califica de “puramente analítica”, que consiste en el método de bisección, en la que Cauchy usa, sin demostración ni comentario, que una sucesión monótona acotada es convergente, propiedad que equivale a la completitud del sistema de los números reales.

El citado texto de Cauchy, así como los libros *Résumé des leçons sur le Calcul Infinitésimal* (1823) y *Leçons sur le Calcul Différentiel* (1829), en los que se recogen los cursos impartidos por Cauchy en la École Polytechnique durante los años precedentes, tuvieron una gran influencia y establecieron nuevas exigencias de rigor. En el cálculo de Cauchy los conceptos de función y de límite son los conceptos fundamentales.

2.6. El innovador trabajo de Bolzano



Figura 13. Bolzano

Es obligado citar a [Bernhard Bolzano](#) (1781 - 1848), matemático, lógico y filósofo, profesor en la Universidad de su ciudad natal, Praga, desde 1805 a 1820. Bolzano, cuyas obras completas comprenderán, cuando terminen de editarse, alrededor de 140 volúmenes, fue un innovador en todos los campos que trabajó. En sus trabajos matemáticos anticipó muchos de los conceptos que posteriormente redescubrieron y desarrollaron matemáticos como Cauchy, Weierstrass o Cantor. Debido a su relativo aislamiento en la ciudad de Praga, en una época en la que el centro de toda la producción matemática estaba en París, la obra matemática de Bolzano fue poco conocida y no tuvo la influencia que merecía por su rigor y profundidad.

Por lo que a la continuidad de una función se refiere, Bolzano publicó en 1817 un pequeño libro de 60 páginas *Purely analytic proof of the theorem that between any two values which give results of opposite sign there lies at least one real root of the equation* [22], en el que, entre otras cosas, demuestra el teorema que ahora lleva su nombre. Bolzano empieza razonando que las demostraciones conocidas de ese teorema eran inapropiadas. La claridad de ideas con que se expresa es muy

llamativa (traduzco de [22]):

No obstante, un examen más cuidadoso muestra muy pronto que ninguna de estas pruebas puede considerarse adecuada.

I. El tipo de demostración más usual depende de una verdad pedida en préstamo a la geometría, a saber, que toda línea continua de curvatura simple cuyas ordenadas son primero positivas y después negativas (o recíprocamente) necesariamente debe intersectar en algún lugar al eje de abscisas en un punto comprendido entre aquellas ordenadas. Ciertamente, nada hay que objetar respecto a la corrección, ni tampoco a la obviedad, de esta proposición geométrica. Pero está claro que es una intolerable ofensa contra el método correcto, deducir verdades de las matemáticas puras (o generales, i.e. aritmética, álgebra, análisis) a partir de consideraciones que pertenecen simplemente a una parte aplicada (o especial), a saber, la geometría.

[...] Consideremos ahora la razón objetiva por la que una línea en las circunstancias antes mencionadas interseca el eje de abscisas. Sin duda, todo el mundo verá enseguida que esta razón descansa en nada más que en el asentimiento general, como consecuencia del cual toda función continua de x que sea positiva para algún valor de x , y negativa para otro, debe ser cero para algún valor intermedio de x . Y ésta es, precisamente, la verdad que debe ser probada.

II. No menos reprobable es la demostración que algunos han construido a partir del concepto de la continuidad de una función con la inclusión de los conceptos de tiempo y movimiento. [...] Esto es adicionalmente ilustrado por el ejemplo del movimiento de dos cuerpos, uno de los cuales está inicialmente detrás del otro y posteriormente delante del otro. Necesariamente se deduce que en un tiempo debe haber estado al lado del otro. Nadie negará que los conceptos de tiempo y movimiento son tan extraños a la matemática general como el concepto de espacio. No obstante, si estos conceptos fueran introducidos solamente por motivos de claridad, no tendríamos nada en contra de ello. [...] Por tanto, debe observarse que no consideramos que los ejemplos y aplicaciones disminuyan en lo más mínimo la perfección de una exposición científica. De otra manera, estrictamente exigimos sólo esto: que los ejemplos nunca sean empleados como argumentos en lugar de las demostraciones, y que la esencia de una deducción nunca esté basada sobre el uso meramente metafórico de frases o sobre sus ideas relacionadas, de forma que la deducción misma quedaría vacía tan pronto como éstas fueran cambiadas.

Es difícil expresarse con más claridad que como lo hace Bolzano. Su definición de continuidad, en el citado trabajo, es como sigue:

Una función $f(x)$ varía según la ley de continuidad para todos los valores de x dentro o fuera de ciertos límites, significa exactamente que: si x es algún tal valor, la diferencia $f(x + \omega) - f(x)$ puede ser hecha más pequeña que cualquier cantidad dada, supuesto que ω puede ser tomado tan pequeño como queramos.

Seguidamente, Bolzano establece un teorema previo cuyo asombroso enunciado es como sigue:

Si una propiedad M no pertenece a todos los valores de una variable x , pero sí pertenecen todos los valores que son menores que un cierto u , entonces existe siempre una cantidad U que es la mayor de aquellas de las cuales puede afirmarse que toda más pequeña x tiene la propiedad M .

Comprendes por qué califico de “asombroso” ese enunciado, ¿verdad? ¡Es la propiedad de extremo inferior! ¡En el año 1817, 55 años antes de que Dedekind y Cantor publicaran sus teorías de los números reales!

El conocido historiador de las matemáticas Ivor Grattan - Guinness, en un polémico trabajo titulado *Bolzano, Cauchy and the “New Analysis” of Early Nineteenth Century* [11], expresa su opinión de que Cauchy conocía el trabajo de Bolzano pero nunca lo reconoció. Desde luego, ni en las numerosas obras de Cauchy, ni en su correspondencia particular, se ha encontrado ninguna referencia a Bolzano, por lo que la afirmación de Grattan - Guinness, como él mismo reconoce, no está sustentada en pruebas documentales.

2.7. Weierstrass nos dio los $\varepsilon - \delta$

Una característica de los textos citados de Cauchy es que en ellos no hay ni una sola figura. Cauchy liberó al cálculo de sus ataduras geométricas, aunque todavía sus definiciones contenían términos imprecisos como “tan pequeño como queramos” y “disminuir indefinidamente hasta converger al límite cero”, o ideas de movimiento como “variable que se acerca a un límite” por no hablar de sus “infinitamente pequeños”. Para seguir avanzando era necesario acabar de una vez con las distinciones entre número y cantidad. Los números reales todavía eran considerados geométricamente y no se habían establecido sus propiedades de forma explícita. El cero y los números negativos eran vistos aún por muchos matemáticos como algo de naturaleza diferente a los números positivos.



Figura 14. Weierstrass

En definitiva, debía concretarse el significado de expresiones como “cantidad variable” y “variable continua”. También era preciso separar la idea de función de su representación analítica concreta, lo cual, como ya vimos en el Capítulo 2, fue hecho por Dirichlet en 1837 con su definición general de función como correspondencia arbitraria. Finalmente, pero no menos importante, estaban las cuestiones referentes a la convergencia de sucesiones y series numéricas y funcionales, aún mal comprendidas en la época de Cauchy, de las que nos ocuparemos en otro lugar.

En los cincuenta años que van de 1830 a 1880 se lograron desentrañar todas estas cuestiones fundamentales gracias, principalmente, a los trabajos de Dirichlet, Riemann, Weierstrass, Dedekind y Cantor. Ya conocemos una parte de este complejo proceso, la que culmina en 1872 con la fundamentación del sistema de los números reales por Dedekind y Cantor.

Fue [Karl Weierstrass](#) (1815 - 1897) quien llevó a sus últimas consecuencias el proceso de “aritmización del Análisis”. Weierstrass era un desconocido profesor de instituto, cuando en 1854 publicó un trabajo sobre las funciones abelianas que causó sensación en la comunidad matemática. Poco después, en 1856, Weierstrass ya era profesor de la Universidad de Berlín. Los cursos que Weierstrass impartió en Berlín durante más de treinta años atrajeron a numerosos matemáticos de toda Europa. Discípulos suyos fueron, entre muchos otros menos conocidos, George Cantor (1845 - 1918), Sonya Kovalevsky (1850 - 1891), Max Planck (1858 - 1947) y David Hilbert (1862 - 1943).

Weierstrass estaba convencido de que el Análisis debía ser liberado de los razonamientos geométricos y de los conceptos intuitivos de espacio, tiempo y movimiento y debía ser fundamentado sobre los enteros positivos. Acometió la tarea de revisar radicalmente los conceptos fundamentales del Análisis y a este fin dedicó algunos de sus cursos. Entre otras cosas, desarrolló en ellos una teoría aritmética de los números reales parecida a la de Cantor. Aunque Weierstrass no publicó mucho, su influencia fue enorme y sus conferencias magistrales fueron difundidas por toda Europa por sus numerosos alumnos. Weierstrass es considerado como el más grande analista del último tercio del siglo XIX y se le ha llamado “el padre del análisis moderno”. Más adelante tendremos ocasión de exponer algunas de sus contribuciones.

Por lo que al concepto de límite funcional se refiere, Weierstrass tradujo por medio de desigualdades y de valores absolutos las definiciones verbales de límite y de continuidad dadas por Cauchy y Bolzano. Para Weierstrass, una variable solamente es un símbolo que sirve para designar cualquier elemento del conjunto de valores que se le pueden atribuir. Una variable continua es aquella cuyo conjunto de valores no tiene puntos aislados. La definición de límite dada por Weierstrass, tal como la recogió en sus notas el matemático H.E. Heine (1821 - 1881) es la siguiente:

Se dice que L es el límite de una función $f(x)$ para $x = x_0$ si, dado cualquier ϵ , existe un δ_0 tal que para $0 < \delta < \delta_0$, la diferencia $f(x_0 \pm \delta) - L$ es menor en valor absoluto que ϵ .

Cuando una teoría ha sido desarrollada, llega el momento del rigor. Así el concepto de límite, fundamental en cálculo porque en él se basan los de continuidad, derivada, integral y los distintos tipos de convergencia, y es el concepto que confiere al cálculo su característica distintiva, solamente pudo ser expresado de forma rigurosa (según nuestros criterios actuales) en el último tercio del siglo XIX, después de haberse estado usando, de forma más o menos disfrazada por los infinitésimos y otros conceptos afines como el movimiento, durante doscientos años. Curiosamente, la letra griega ϵ , que usaba Cauchy con un significado de “error”, se ha convertido en el paradigma de la precisión en nuestras actuales definiciones heredadas de Weierstrass.

La flechita en la notación para límites, $\lim_{x \rightarrow x_0} f(x)$, fue introducida por G.H. Hardy (1877 - 1947) en su notable libro *A Course of Pure Mathematics* (1908).

3. Orígenes y desarrollo del concepto de derivada. La invención del Cálculo

El concepto de derivada presupone los de función y de límite funcional, los cuales, como ya hemos visto en capítulos anteriores, tuvieron una larga evolución hasta alcanzar su significado actual, por eso la definición de derivada es relativamente reciente. No obstante, técnicas en las que podemos reconocer el uso, más o menos explícito, de derivadas, se han venido usando desde el siglo XVII, incluso antes de que Newton y Leibniz, en el último tercio de dicho siglo, las formularan en términos de *fluxiones* y de *cocientes diferenciales* respectivamente. Durante los siglos XVIII y XIX las derivadas fueron ampliamente desarrolladas y aplicadas a campos muy diversos y no fueron definidas en los términos actuales hasta el último tercio del siglo XIX. Todo este proceso lo resume la historiadora de las matemáticas Judith V. Grabiner en una frase feliz: “*Primero, la derivada fue usada, después descubierta, explorada y desarrollada y, finalmente, definida*”.

En lo que sigue vamos a repasar muy someramente este proceso. Además de la referencia antes citada, he seguido de cerca los trabajos de Kirsti Andersen [2], Israel Kleiner [18] y González Urbaneja [16].

3.1. Las matemáticas en Europa en el siglo XVII

Es conocido que la carencia de una teoría aritmética satisfactoria de las cantidades inconmensurables, hizo que los matemáticos griegos consideraran la Geometría como una ciencia más general que la Aritmética, lo que condujo al desarrollo de un álgebra geométrica que fue usada por Euclides, Arquímedes y Apolonio para realizar sus cálculos. La consecuencia de esta actitud fue que durante casi 2000 años, en Europa, casi todo razonamiento matemático riguroso se expresó en lenguaje geométrico.

Ya hemos comentado en capítulos anteriores cómo la herencia matemática griega pasó a los árabes de donde regresó a Europa ya en el siglo XII. En estos siglos se desarrolló sobre todo la aritmética y los comienzos del álgebra. Pero hay que esperar hasta el siglo XVII para que en Europa empiecen a notarse cambios significativos en la forma de hacer matemáticas y a lograr avances que abren nuevas perspectivas. Las características principales de las matemáticas en el siglo XVII en Europa son las siguientes.

- Asimilación y síntesis de la tradición clásica griega y del legado árabe.
- Se sigue admirando el rigor demostrativo euclidiano pero se buscan procedimientos heurísticos. Se impone la idea de “primero descubrir y luego demostrar”.

- Progresos decisivos en el simbolismo algebraico (Viète, Stevin). Concepto de cantidad abstracta.
- Invención de la geometría analítica por Fermat y Descartes.
- Multitud de nuevas curvas, muchas de ellas curvas mecánicas, como la cicloide, que llevan consigo problemas de tangentes, cuadraturas, centros de gravedad, máximos y mínimos, rectificaciones.
- Invención de métodos infinitesimales para tratar problemas de cuadraturas, tangentes, máximos y mínimos. Libre uso del infinito.
- Inicios del estudio matemático del movimiento. Concepto de cantidad variable.
- La Revolución Científica protagonizada por Copérnico, Galileo y Kepler. Mecanicismo.
- Invención de los logaritmos por Neper. Progresos de la astronomía y de la trigonometría. Desarrollo de la óptica.
- Creación de instituciones científicas como la Royal Society (1660) en Londres y la Académie des Sciences (1666) en París y comienzo de las publicaciones científicas periódicas.

En el periodo de 1630 a 1660 empiezan a usarse técnicas en las que podemos apreciar el uso de derivadas. Suelen ser técnicas específicas para resolver problemas concretos de forma empírica, con frecuencia dichas técnicas no se justifican sino que, simplemente, se comprueba que proporcionan soluciones correctas. Los matemáticos de la época se interesaban por problemas de óptica, por ejemplo, determinar la forma de una lente que hace que todos los rayos luminosos paralelos entre sí o los que parten de un único foco, después de atravesar la lente, converjan en un único punto. Problemas físicos, como la determinación de la trayectoria de un cuerpo que se mueve alrededor de un centro y que cae al mismo tiempo hacia ese centro con aceleración constante. Otros problemas consistían en el cálculo de tangentes y de valores máximos o mínimos. Estaban, además, los problemas relacionados con la integral (cuadraturas, áreas de superficies, centros de gravedad, rectificaciones de curvas,...) que consideraremos en el capítulo correspondiente.

3.2. Cálculo de tangentes y de valores extremos

Los matemáticos de la antigüedad sabían cómo trazar tangentes a diversos tipos de curvas. El concepto de tangencia de los griegos es estático y, naturalmente, geométrico. Inicialmente, la tangente se considera como una recta que toca a la curva sin cortarla. Esta definición resultaba apropiada para la circunferencia pero no lo era para otras curvas. En el siglo III a.C., Apolonio definió la tangente a una sección cónica y procedió a determinarla en cada caso. Las técnicas para el cálculo de tangentes eran, por supuesto, geométricas. Para curvas como la espiral de Arquímedes o la conoide de Nicomedes estas técnicas no eran de gran utilidad.

Con la invención de la geometría analítica, había una enorme variedad de nuevas curvas para cuyo estudio no servían los métodos tradicionales. Los matemáticos del siglo XVII se vieron en la necesidad de inventar nuevas técnicas para calcular tangentes. Vamos a considerar algunas de las aportaciones más significativas.

3.2.1. El método de máximos y mínimos de Fermat

En 1637 Fermat escribió una memoria titulada *Methodus ad disquirendam maximam et minimam* (“Método para la investigación de máximos y mínimos”). En ella se establecía el primer procedimiento general conocido para calcular máximos y mínimos. Fermat se expresa como sigue.

Toda la teoría de la investigación de máximos y mínimos supone la consideración de dos incógnitas y la única regla siguiente:

1. Sea a una incógnita cualquiera del problema (que tenga una, dos o tres dimensiones, según convenga al enunciado).
2. Se expresará la cantidad máxima o mínima por medio de a en términos que pueden ser de cualquier grado.
3. Se sustituirá a continuación la incógnita original a por $a + e$, y se expresará la cantidad máxima o mínima por medio de a y e , en términos que pueden ser de cualquier grado.
4. Se “adigualará” para hablar como Diofanto, las dos expresiones de la cantidad máxima o mínima.
5. Se eliminarán los términos comunes de ambos lados, tras lo cual resultará que a ambos lados habrá términos afectados de e o de una de sus potencias.
6. Se dividirán todos los términos por e , o por alguna potencia superior de e , de modo que desaparecerá la e , de al menos uno de los términos de uno cualquiera de los dos miembros.
7. Se suprimirán, a continuación, todos los términos donde todavía aparece la e o una de sus potencias, y se iguala lo que queda, o bien si en uno de los miembros no queda nada, se igualará, lo que viene a ser lo mismo, los términos afectados con signo positivo a los afectados con signo negativo.
8. La resolución de esta última ecuación dará el valor de a , que conducirá al máximo o mínimo, utilizando la expresión original.

Fermat ilustraba su método hallando el punto E de un segmento AC que hace máxima el área del rectángulo $AE.EC$.

Pongamos $AC = b$.

1. Sea a uno de los segmentos, el otro será $b - a$.
2. El producto del que se debe encontrar el máximo es $ba - a^2$.
3. Sea ahora $a + e$ el primer segmento de b , el segundo segmento será $b - a - e$, y el producto de segmentos: $ba - a^2 + be - 2ae - e^2$.

4. Se debe “adigular” al precedente: $ba - a^2 + be - 2ae - e^2 \sim ba - a^2$.
5. Suprimiendo términos comunes: $be \sim 2ae + e^2$.
6. Dividiendo todos los términos por e : $b \sim 2a + e$.
7. Se suprime la e : $b = 2a$.
8. Para resolver el problema se debe tomar por tanto la mitad de b .

El recurso de hacer $e = 0$ es equivalente a lo indicado en la instrucción 7 de Fermat. Esto era precisamente lo que se hacía al aplicar el método, a pesar de que antes era necesario dividir por e , lo que resultaba algo contradictorio.

Debemos observar que el método de Fermat da una condición necesaria para los máximos y mínimos, pero esa condición no es suficiente y tampoco distingue máximos de mínimos. Es un método puramente algebraico y algorítmico, no geométrico.

Es tentador reproducir este razonamiento en términos actuales. Hagamos $a = x$, $e = \Delta x$, y pongamos $f(x) = x(b - x)$.

$$1 - 5 \quad f(x + \Delta x) - f(x) \sim 0.$$

$$6 \quad \frac{f(x + \Delta x) - f(x)}{\Delta x} \sim 0.$$

$$7, 8 \quad \left(\frac{f(x + \Delta x) - f(x)}{\Delta x} \right)_{\Delta x=0} = 0$$

Para funciones derivables podemos interpretar todo esto como que el valor de x que hace máximo o mínimo a $f(x)$ es la solución de resolver la ecuación

$$f'(x) = \lim_{\Delta x \rightarrow 0} \frac{f(x + \Delta x) - f(x)}{\Delta x} = 0$$

Sin embargo, esto significa extrapolar demasiado el contenido estricto del método. Lo que estamos haciendo es interpretar con nuestra mirada de hoy lo que hizo Fermat. En primer lugar, Fermat no pensaba en una cantidad como una función, y por eso habla de “cantidad máxima o mínima”, no de una función que alcance un máximo o un mínimo. Fermat no tiene clara la noción de variable independiente. Él está pensando en una ecuación algebraica con dos incógnitas que interpreta como segmentos, es decir, magnitudes lineales dadas. Fermat no decía nada acerca de que e fuese un infinitesimal, ni siquiera una magnitud muy pequeña, y el método no implica ningún concepto de límite, sino que es puramente algebraico. Además, la condición 6 no tiene sentido en esta interpretación. Los problemas a los que Fermat aplicó su método son problemas de construcciones geométricas más que de optimización de cantidades.

3.2.2. El método de las tangentes de Fermat

En la misma memoria antes referida, Fermat, determina la subtangente a una parábola haciendo uso de su método para máximos y mínimos. Su razonamiento es como sigue.

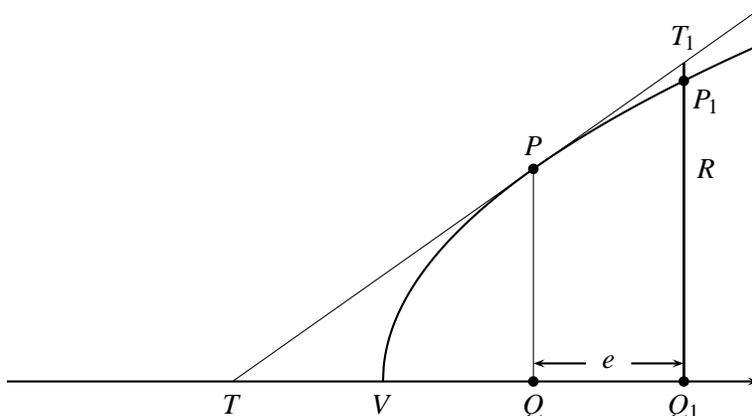


Figura 15. Cálculo de la subtangente

En la figura (16), el segmento \overline{TQ} es la subtangente a la parábola en un punto dado P . El vértice de la parábola es V . Teniendo en cuenta que los triángulos $\triangle(TQP)$ y $\triangle(TQ_1T_1)$ son semejantes, resulta

$$\frac{\overline{T_1Q_1}}{\overline{PQ}} = \frac{\overline{TQ_1}}{\overline{TQ}} \quad (5)$$

Teniendo en cuenta ahora la propiedad de la parábola

$$\frac{\overline{VQ_1}}{\overline{VQ}} = \frac{\overline{P_1Q_1}^2}{\overline{PQ}^2}$$

y que $\overline{P_1Q_1} < \overline{T_1Q_1}$, deducimos que:

$$\frac{\overline{VQ_1}}{\overline{VQ}} < \frac{\overline{TQ_1}^2}{\overline{TQ}^2} \quad (6)$$

Pongamos ahora $\overline{VQ} = a$, que es la abscisa de la parábola en P , conocida porque se conoce P . Hagamos también $\overline{TQ} = x$ que es la subtangente que queremos calcular, y $\overline{QQ_1} = e$. La desigualdad (6) se expresa por:

$$\frac{a+e}{a} < \frac{(x+e)^2}{x^2} \iff ax^2 + ex^2 < ax^2 + 2aex + ae^2$$

Fermat aplica su método de máximos y mínimos y sustituye esta desigualdad por la *adigualdad*

$$ax^2 + ex^2 \sim ax^2 + 2aex + ae^2$$

Cancelando términos y dividiendo por e obtenemos

$$x^2 \sim 2ax + ae$$

Eliminando ahora el término que queda en e , igualando y simplificando por x , se obtiene que $x = 2a$, resultado ya conocido de la Antigüedad y que expresa que la subtangente es el doble de la abscisa.

Realmente no se entiende bien la razón de por qué Fermat usa su método de máximos y mínimos para calcular tangentes y Descartes hizo una dura crítica de esta forma de proceder. Para responder a estas críticas, Fermat desarrolló, en una memoria de 1638, un procedimiento bastante general para calcular tangentes que, con notación actual, podemos resumir como sigue. Sea $P = (x, y)$ un punto de una curva $f(x, y) = 0$ y sea $P_1 = (x + e, y_1)$ otro punto de la curva próximo a P como en la figura (15). Llamemos $b = \overline{TQ}$, la subtangente en P . Teniendo en cuenta que $\overline{PQ} = y$, la igualdad (5) se escribe como

$$\frac{\overline{T_1Q_1}}{y} = \frac{y(b + e)}{b}$$

Como $\overline{T_1Q_1}$ es casi igual a $y_1 = \overline{P_1Q_1}$, Fermat escribe

$$f\left(x + e, \frac{y(b + e)}{b}\right) \sim 0$$

y a esta *adigualdad* le aplica su método para máximos y mínimos. Es fácil ver que ello conducirá a una expresión para b dada por

$$b = -\frac{y \frac{\partial f}{\partial y}(x, y)}{\frac{\partial f}{\partial x}(x, y)}$$

Que, usando que la tangente viene dada por y/b , podemos escribir, viendo y como función (implícita) de x , en la forma familiar

$$y' = -\frac{\frac{\partial f}{\partial x}(x, y)}{\frac{\partial f}{\partial y}(x, y)}$$

La idea de “*adigualdad*” en Fermat puede interpretarse algo así como “cantidades infinitamente próximas”. De alguna forma Fermat está considerando cantidades infinitesimales.

Es tentador expresar en términos actuales las ideas de Fermat para calcular tangentes. Esencialmente, dado un punto $P = (a, f(a))$ en una curva $y = f(x)$, se trata de calcular la pendiente de la curva en P . Sea $\overline{QQ_1}$ un incremento de \overline{TQ} en una cantidad E . Ya que los triángulos $\triangle(TQP)$ y $\triangle(PRT_1)$ son semejantes, se tiene

$$\frac{\overline{PQ}}{\overline{TQ}} = \frac{\overline{T_1R}}{E}$$

Pero, dice Fermat, $\overline{T_1R}$ es casi igual a $\overline{P_1R}$; por tanto tenemos la *adigualdad*

$$\frac{\overline{PQ}}{\overline{TQ}} \sim \frac{\overline{P_1Q_1} - \overline{PQ}}{E}$$

Poniendo $\overline{PQ} = f(a)$, la igualdad anterior puede escribirse como:

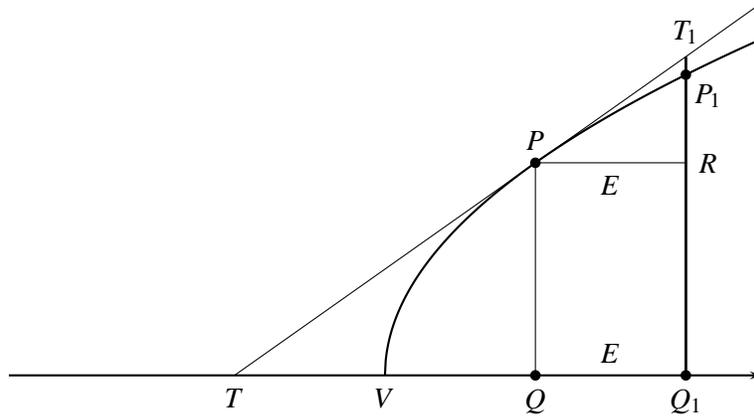


Figura 16. Cálculo de la tangente

$$\frac{f(a)}{\overline{TQ}} \sim \frac{f(a+E) - f(a)}{E}$$

Ahora, dice Fermat, se cancelan términos iguales en $f(a+E) - f(a)$, se divide por E y finalmente, se ignoran los términos que aún contengan E (lo que equivale a hacer $E = 0$), y el resultado es la pendiente de la tangente en P . Está claro que el procedimiento que indica Fermat es equivalente a calcular

$$\lim_{E \rightarrow 0} \frac{f(a+E) - f(a)}{E}$$

Naturalmente, a esta interpretación se le pueden hacer las mismas observaciones que hicimos a la interpretación análoga del método para máximos y mínimos.

Ejemplo.

Sea $f(x) = x^2 - 2x + 3$ y $a = 2$. Entonces $f(2) = 3$. Pongamos $c = TQ$ la longitud de la subtangente. Tenemos la *adigualdad*:

$$\frac{3}{c} = \frac{f(2+E) - f(2)}{E} = \frac{2E + E^2}{E} = 2 + E$$

Haciendo $E = 0$ se obtiene $3/c = 2$, por la que la subtangente es $c = 3/2$ y el valor de la pendiente de la tangente es $3/c = 2$ que, efectivamente es igual a la derivada de f en $x = 2$.

3.2.3. El método de Roberval y de Torricelli para las tangentes

En 1630 Roberval y Torricelli descubrieron independientemente un método para calcular tangentes por medio de consideraciones cinemáticas. Este método se apoya en dos ideas básicas: la primera es la de considerar una curva como la trayectoria de un punto móvil que obedece a dos movimientos simultáneamente, y la segunda es la de considerar la tangente en un punto de la curva como la dirección del movimiento en ese mismo punto. Si la razón entre las velocidades de los dos movimientos es conocida, la dirección del movimiento resultante se puede hallar mediante la ley del paralelogramo. Ya en la antigüedad, Arquímedes había usado un método análogo para trazar la tangente a su espiral.

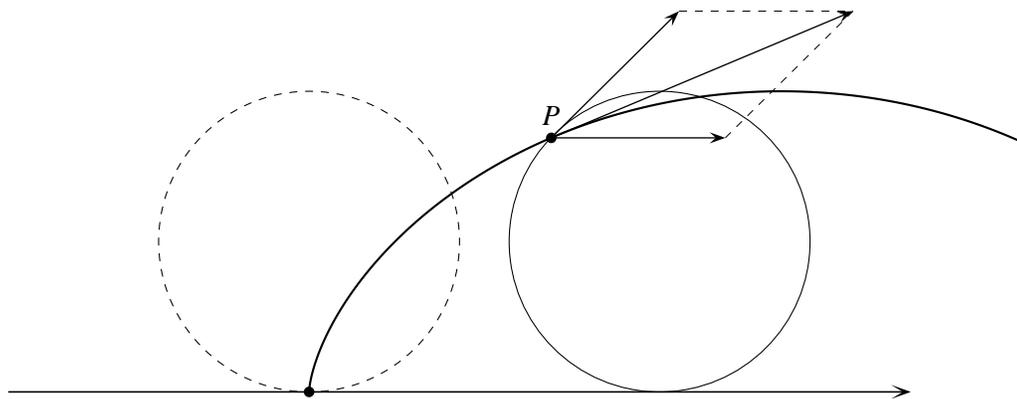


Figura 17. Tangente a la cicloide

Consideremos una cicloide, esto es la curva que describe un punto de una circunferencia que rueda sin deslizar. El punto que genera la cicloide tiene una velocidad angular igual a la velocidad de avance horizontal, por tanto, su tangente en un punto P se obtiene sumando el vector tangente a la circunferencia generadora en P y un vector horizontal en P , y ambos vectores tienen igual módulo.

Naturalmente, esta idea de la tangente solamente podía aplicarse a curvas mecánicas, si bien tenía la virtud de relacionar geometría y dinámica siguiendo las ideas de Galileo.

3.2.4. El triángulo diferencial de Barrow

Isaac Barrow (1630 - 1677) también dio un método para calcular tangentes. Barrow era un admirador de los geómetras antiguos y editó las obras de Euclides, Apolonio y de Arquímedes, a

la vez que publicaba sus propias obras *Lectiones Opticae* (1669) y *Lectiones Geometricae* (1670) en la edición de las cuales colaboró Newton. El tratado *Lectiones Geometricae* se considera una de las principales aportaciones al Cálculo. En él Barrow quiso hacer una puesta al día de todos los últimos descubrimientos, principalmente de problemas de tangentes y cuadraturas. Barrow hace un tratamiento detallado de todos estos problemas incluyendo conceptos como tiempo y movimiento y usando métodos infinitesimales y métodos de indivisibles.

Una de las herramientas a las que saca gran partido es al triángulo característico o triángulo diferencial.

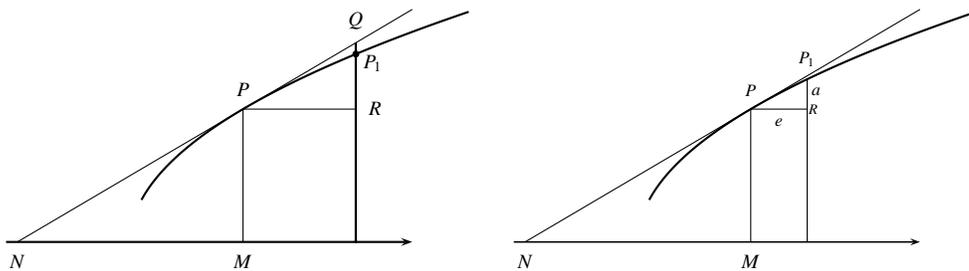


Figura 18. Triángulo diferencial

Partiendo del triángulo PRQ , que resulta de un incremento PR , como este triángulo es semejante al PNM , resulta que la pendiente de la tangente PM/MN es igual a QR/PR . Barrow afirma que cuando el arco PP_1 es muy pequeño podemos identificarlo con el segmento PQ de la tangente en P . El triángulo PP_1R de la figura de la derecha, en el cual PP_1 es considerado a la vez como un arco de la curva y como parte de la tangente, es el *triángulo característico o diferencial*. Ya había sido usado mucho antes por Pascal y otros en problemas de cuadraturas.

En la Lección X de *Lectiones*, Barrow calcula la tangente a una curva, dada por una ecuación polinómica $f(x,y) = 0$, en un punto de la misma $P = (x,y)$ de la forma siguiente. Pongamos $P_1 = (x + e, y + a)$ un punto de la curva próximo a P y sustituyamos estas coordenadas en la ecuación $f(x,y) = 0$. En palabras de Barrow:

Rechacemos todos los términos en los que no hay a o e (porque se anulan unos a otros por la naturaleza de la curva); rechacemos todos los términos en los que a o e están por encima de la primera potencia, o están multiplicados ambos (porque, siendo infinitamente pequeños, no tienen valor en comparación con el resto).

Después de estas operaciones se puede calcular el cociente a/e que es la pendiente de la curva en

el punto P .

Ejemplo.

Consideremos la curva $x^3 + y^3 = r^3$ y sigamos el método de Barrow para calcular su pendiente en un punto $P = (x, y)$ de la misma. Como el punto $P_1 = (x + e, y + a)$ está en la curva se tiene:

$$(x + e)^3 + (y + a)^3 = r^3$$

Esto es

$$x^3 + 3x^2e + 3xe^2 + e^3 + y^3 + 3y^2a + 3ya^2 + a^3 = r^3$$

Simplificamos usando que $x^3 + y^3 = r^3$ y eliminando las potencias de a y e de grado mayor que uno, y obtenemos

$$3x^2e + 3y^2a = 0$$

de donde resulta la pendiente:

$$\frac{a}{e} = -\frac{x^2}{y^2}$$

Observa que este procedimiento equivale a quedarse con la aproximación lineal de la función en el punto P y eso es como reemplazar el triángulo PRP_1 en la figura de la izquierda por el triángulo diferencial.

El método de Barrow es parecido al de Fermat, la diferencia es que Barrow considera incrementos independientes de las dos variables con el propósito de calcular el cociente a/e . Parece que Barrow no conocía directamente la obra de Fermat.

3.3. Los inventores del Cálculo

El método de Fermat para el cálculo de valores máximos o mínimos y la técnica para el cálculo de tangentes que, esencialmente, consistía en calcular el cociente:

$$\frac{f(x+h) - f(x)}{h},$$

realizando las operaciones algebraicas necesarias para desarrollar y simplificar el numerador y después dividir por h para, finalmente, hacer $h = 0$, fueron aplicados en una gran variedad de situaciones. La relación entre ambos tipos de problemas acabó siendo bien entendida: los valores extremos se obtenían en los puntos donde la pendiente de la tangente se anulaba. Así mismo, de la multitud de casos particulares estudiados, emergieron ciertas regularidades que llevaron a reformular las citadas técnicas de forma más general. De esta forma, aunque en el 1660 no se disponía de un concepto general de derivada ni se conocía la relación crucial entre problemas de tangentes y de áreas, se habían desarrollado bastantes métodos eficaces, aunque no rigurosos, para

resolver muchos tipos de problemas de cálculo. Solamente faltaba realizar la gran síntesis de todo el trabajo realizado desde 1630. Eso es lo que hicieron Newton y Leibniz.

La invención del Cálculo es uno de los grandes logros de la humanidad. El Cálculo se ha convertido en la *lingua franca* de todas las ciencias. Ha sido, y sigue siendo, una herramienta fundamental para la comprensión científica de la Naturaleza.

En el último tercio del siglo XVII, Newton (en 1664 - 1666) y Leibniz (en 1675), de forma independiente cada uno, inventaron el Cálculo. Esto quiere decir que:

- Unificaron y resumieron en dos conceptos generales, el de integral y derivada, la gran variedad de técnicas diversas y de problemas que se abordaban con métodos particulares.
- Desarrollaron un simbolismo y unas reglas formales de “cálculo” que podían aplicarse a funciones algebraicas y trascendentes, independientes de cualquier significado geométrico, que hacía fácil, casi automático, el uso de dichos conceptos generales.
- Reconocieron la relación inversa fundamental entre la derivación y la integración.

3.4. Newton y el cálculo de fluxiones



Figura 19. Newton

Los principales descubrimientos matemáticos de Newton en el campo del cálculo infinitesimal datan de los llamados *Anni Mirabiles* 1665 y 1666. La Universidad de Cambridge, en la que Newton se había graduado como *bachelor of arts* en 1664, estuvo cerrada por la peste esos dos años. Newton pasó ese tiempo en su casa de Woolsthorpe y, como él mismo reconoció cincuenta años después, ése fue el período más creativo de su vida.

Newton llamó a nuestra derivada una *fluxión* – una razón de cambio o flujo; Leibniz vio la derivada como una razón de diferencias infinitesimales y la llamó el *cociente diferencial*. Newton hizo sus primeros descubrimientos diez años antes que Leibniz quien, sin embargo, fue el primero en publicar sus resultados. A principios de 1665 descubre el teorema del binomio y el cálculo con las series infinitas. A finales de ese mismo año, el método de fluxiones, es decir, el cálculo de derivadas. En 1666 el método inverso de fluxiones y la relación entre cuadraturas y fluxiones. En esos dos años también inició las teorías de los colores y de la gravitación universal. Newton tenía 24 años, había nacido el día de Navidad de 1642.

Newton desarrolló tres versiones de su cálculo. En la obra *De Analysi per aequationes numero terminorum infinitas*, que Newton entregó a su maestro Barrow en 1669, y que puede considerarse el escrito fundacional del Cálculo, Newton usa conceptos infinitesimales de manera similar a como hacía el propio Barrow.

Una segunda presentación del Cálculo es la que realiza Newton en el libro *Methodus fluxionum et serierum infinitorum*, escrito hacia 1671 y que se publicó mucho después en 1736. Newton considera cantidades variables que van fluyendo con el tiempo, a las que llama *fluente*s. Después se introducen las razones de cambio instantáneas de las fuentes, a las que llama *fluxiones*, que son las derivadas respecto al tiempo de las fuentes. Newton representaba a las primeras por letras x, y, z, \dots y a las segundas por letras punteadas $\dot{x}, \dot{y}, \dot{z}, \dots$. Los incrementos de las fuentes x, y, z, \dots , los representa por medio de las correspondientes fluxiones en la forma $\dot{x}o, \dot{y}o, \dot{z}o, \dots$, y los llama *momentos*, donde o es entendido como un incremento infinitesimal de tiempo. Newton desarrolló una serie de algoritmos y redujo muchos problemas como determinación de tangentes, máximos y mínimos, áreas y superficies, curvaturas, longitudes de arcos, centros de gravedad etc., a dos problemas fundamentales que pueden formularse tanto en términos mecánicos como en términos matemáticos:

Problema 1 Determinación de la velocidad de movimiento en un momento de tiempo dado según un camino dado. De otro modo: dada la relación entre las cantidades fuentes, determinar la relación de las fluxiones.

Problema 2 Dada la velocidad de movimiento determinar el camino recorrido en un tiempo dado. Matemáticamente: determinar la relación entre las fuentes dada la relación entre las fluxiones.

Hay que notar que Newton no piensa en términos de funciones con el significado actual de ese término, sino que imagina curvas o superficies descritas por las variables, o sea, considera relaciones entre las fuentes del tipo $f(x, y, z, \dots) = 0$, donde f para él es una expresión analítica finita o infinita. Por tanto, el primer problema planteado puede verse como un problema de derivación implícita: supuesta conocida la expresión analítica que satisfacen las fuentes $f(x, y, z, \dots) = 0$, obtener la expresión analítica $F(x, y, z, \dot{x}, \dot{y}, \dot{z}, \dots) = 0$ que satisfacen las fluxiones. Para este problema, Newton introdujo un algoritmo que sistematizaba los cálculos necesarios. Por ejemplo, sea la curva de ecuación

$$x^3 - ax^2 + axy - y^3 = 0$$

Sustituyendo x e y por $x + \dot{x}o$ e $y + \dot{y}o$ respectivamente, tenemos:

$$(x^3 + 3\dot{x}ox^2 + 3\dot{x}^2o^2x + \dot{x}^3o^3) - a(x^2 + 2\dot{x}ox + \dot{x}^2o^2) + a(xy + \dot{x}oy + \dot{y}ox + \dot{x}\dot{y}o^2) - (y^3 + 3\dot{y}oy^2 + 3\dot{y}^2o^2y + \dot{y}^3o^3) = 0$$

Teniendo en cuenta ahora que $x^3 - ax^2 + axy - y^3 = 0$, dividiendo por o y despreciando los demás términos que contengan a o , resulta

$$3\dot{x}x^2 - 2a\dot{x}x + a\dot{x}y + ax\dot{y} - 3\dot{y}y^2 = 0$$

Esta es la relación que satisfacen las fluxiones. A partir de ella puede obtenerse la tangente a la curva $x^3 - ax^2 + axy - y^3 = 0$ en cualquier punto (x, y) de la misma, que viene dada por:

$$\frac{\dot{y}}{\dot{x}} = \frac{3x^2 - 2ax + ay}{3y^2 - ax}$$

Como ya hemos indicado, Newton aplica los resultados sobre fuentes y fluxiones a la resolución de multitud de problemas. Por ejemplo, con respecto a los problemas de máximos y mínimos, escribe:

Cuando una cantidad es la más grande o la más pequeña, en ese momento su fluir ni crece ni decrece: si creciera, eso probaría que era menor y que lo que sigue sería más grande que lo que ahora es, y recíprocamente pasaría si decreciera. Así, calcúlese su fluxión como se ha explicado en el problema 1 e iguállese a cero.

De nuevo, Newton usa el teorema fundamental del cálculo para realizar cuadraturas. Escribe:

Problema 9: Determinar el área de cualquier curva propuesta.

La resolución del problema está basada en el establecimiento de la relación entre la cantidad fuente y su fluxión (problema 2).

Newton reduce la integración al proceso inverso del cálculo de fluxiones, esto es, al cálculo de primitivas.

El problema 2, es mucho más difícil que el problema 1, pues se trata de resolver una ecuación diferencial que puede ser muy general. Newton consideró varias posibilidades resolviendo algunos casos particulares. Para ello utilizó técnicas de cálculo de primitivas y de desarrollos en serie.

En *De Quadratura Curvarum*, escrita en 1676 y publicada en 1704, Newton propone fundamentar su cálculo de fluxiones en lo que llama *razones primera y última de incrementos evanescentes*. De esa forma se refiere Newton a los cocientes de los incrementos infinitesimales de las cantidades variables, y su objetivo es determinarlos en el momento en que dichas cantidades nacen desde cero (“razón primera”) o se anulan (“razón última”). Un ejemplo ayudará a entender el significado de estas ideas. En la introducción de la citada obra, Newton calcula la fluxión de x^n . Para ello, considera un incremento o de forma que x pasa a $x + o$. Entonces x^n se convierte en

$$(x + o)^n = x^n + nox^{n-1} + \frac{n(n-1)}{2}o^2x^{n-2} + \dots$$

Los incrementos de x y x^n , a saber,

$$o \quad \text{y} \quad nox^{n-1} + \frac{n(n-1)}{2}o^2x^{n-2} + \dots$$

están entre sí en la misma razón que

$$1 \text{ a } nx^{n-1} + \frac{n(n-1)}{2}ox^{n-2} + \dots$$

Dice Newton “dejemos ahora que los incrementos se anulen y su última proporción será 1 a nx^{n-1} : por tanto, la fluxión de la cantidad x es a la fluxión de la cantidad x^n como 1 : nx^{n-1} ”.

Hay distintas interpretaciones de las razones que llevaron a Newton a exponer su cálculo de una u otra forma. La más extendida es que su intención era conseguir una fundamentación rigurosa del mismo. La primera exposición, basada en el concepto de cantidad infinitesimal, entendida como una cantidad menor que cualquier cantidad positiva pero no nula, presentaba problemas de coherencia lógica de los que Newton era muy consciente. En sus propias palabras, su cálculo estaba “*concisamente explicado más que exactamente demostrado*”.

En *Methodus Fluxionum et Serierum Infinitarum* (1671), el concepto básico es el de cantidad en movimiento o que fluye continuamente en el tiempo. Las magnitudes están generadas por el movimiento continuo y no por agregación de cantidades infinitesimales; la idea básica es la de continuidad tal como se observa en los procesos de la Naturaleza. Quizás Newton pretendía de esta forma evitar el uso de “infinitesimales estáticos o geométricos”, pero lo que realmente hizo fue sustituirlos por los infinitesimales de tiempo usados para definir los momentos de las fuentes. Conviene advertir que lo que Newton considera es la abstracción matemática análoga al tiempo, es decir, una magnitud independiente imaginaria abstracta que fluye uniformemente y con la que se relacionan todas las fuentes. Puede verse aquí un intento de Newton por evitar los problemas matemáticos del continuo (infinitesimales, indivisibles) y trasladarlos al mundo físico, a la continuidad de los procesos naturales y al movimiento. Por otra parte, Newton aceptaba como algo dado la idea intuitiva de velocidad instantánea de las fuentes, no le pareció preciso definirla.

En *Quadrature of Curves* (1676), Newton expresa su propósito de abandonar por completo el uso de cantidades infinitesimales. Manifiesta en este sentido que “*errores quam minimi in rebus mathematicis non sunt contemnendi*”, esto es, que en matemáticas ni siquiera los errores más pequeños pueden ser admitidos. Y eso es justamente lo que se hacía cuando se despreciaban en los cálculos cantidades infinitesimales. Seguidamente, enuncia su teoría de las “*razones primera y última de cantidades evanescentes*”. Estas ideas señalan claramente al concepto matemático de límite. Lo que expresa, a su manera, Newton es, en términos actuales, el límite de un cociente de funciones que se anulan. Pero estamos en el siglo XVII y se necesitarán casi 200 años para precisar matemáticamente el concepto de límite. Debemos notar que Newton usa dicho concepto a partir de la intuición mecánica del movimiento.

Por velocidad última se entiende aquella con la que el cuerpo se mueve, no antes de alcanzar el punto final y cesa, por consiguiente, el movimiento, ni tampoco después de haberlo alcanzado,

sino aquella con la que se mueve cuando lo alcanza, esto es, aquella velocidad con la que el cuerpo alcanza el punto final y aquella con la que cesa el movimiento. De igual manera, ha de entenderse por razón última de cantidades evanescentes, la razón de cantidades, no antes de que desaparezcan, ni después de desaparecidas, sino aquella con la que desaparecen.

Newton tenía su particular idea de “límite”.

Las razones últimas con las que tales cantidades desaparecen en realidad no son razones de cantidades últimas, sino límites a los que tiende a acercarse siempre las razones de cantidades continuamente decrecientes, límites a los que pueden acercarse más que una diferencia dada, pero nunca traspasarlo, ni tampoco alcanzarlo antes de que las cantidades disminuyan in infinitum.

La teoría de las razones últimas puede verse como una teoría cinemática de límites. Con esta teoría, Newton pretendía recuperar el rigor de la geometría de la Antigüedad.

[...] investigar las razones primera y última de cantidades finitas, nacientes o evanescentes, está en armonía con la geometría de los antiguos; y me he esforzado en probar que, en el método de fluxiones, no es necesario introducir en la geometría cantidades infinitamente pequeñas.

Otros autores opinan que estos tres métodos empleados por Newton responden, más que a fundamentar con rigor su cálculo, a distintos propósitos. Así, la teoría de fluxiones proporciona métodos heurísticos de descubrimiento y algoritmos útiles para el cálculo; la teoría de “razones primera y última” serviría al propósito de proporcionar demostraciones convincentes y el uso de los infinitésimos serviría para proporcionar atajos a las pruebas más rigurosas. Newton usó simultáneamente estas tres aproximaciones en la resolución de una gran variedad de problemas.

Newton realizó también contribuciones importantes en la teoría de ecuaciones, donde podemos destacar las “identidades de Newton” para la suma de las potencias de las raíces de una ecuación polinómica, y a la teoría de curvas, siendo notable su clasificación de las curvas de tercer grado.

Considerando la matemática desde el comienzo del mundo hasta la época de Newton, lo que él ha hecho es, con mucho, la mitad mejor. Leibniz

Las tres obras consideradas, escritas entre 1666 y 1676, se publicaron ya en el siglo XVIII, por eso la primera noticia impresa de la teoría de fluxiones apareció, de forma bastante circunstancial, en la obra magna de Newton *Philosophiae Naturalis Principia Mathematica*, cuya primera edición se hizo en 1687. Los *Principia* consta de tres libros escritos en el estilo tradicional a la manera de los *Elementos* de Euclides, y su lenguaje es principalmente el de la geometría sintética.

Los *Principia* están considerados como la obra científica más importante de todos los tiempos y una hazaña intelectual incomparable por sus logros y sus consecuencias. En dicha obra Newton estable los fundamentos de la mecánica y enuncia las tres célebres leyes del movimiento, así como la ley de la gravitación universal. En los dos primeros libros, se estudia el movimiento de los cuerpos en el vacío y en un medio resistente. Newton deduce matemáticamente las tres leyes que Kepler había obtenido empíricamente. En el libro III, titulado *Sobre el Sistema del Mundo*, Newton desarrolla la mecánica celeste. Hace un detallado estudio de los movimientos de la Luna, explicando las causas de las mareas. Calcula la masa del Sol con respecto a la de la Tierra, estudia la precesión de los equinoccios, predice el achatamiento de la Tierra por los polos . . .

En los *Principia* el mundo aparece como un sistema ordenado y armonioso en el que todo, los cielos, la tierra y el mar, obedecen unas pocas leyes matemáticas fundamentales. A partir de Newton quedará claro que no hay diferencias entre un mundo sublunar y otro supralunar, ni entre la Tierra y el Cielo; las leyes de la Naturaleza no hacen estas distinciones y en todas partes del Universo los procesos obedecen a las mismas leyes naturales inexorables.

El Universo newtoniano es un Cosmos diáfano y sereno ofrecido a la exploración racional del hombre. La gran obra de Newton proporcionará a la Ilustración, en el siglo XVIII, la base científica necesaria para acabar con una concepción conservadora y absolutista del poder político apoyada en dogmáticas concepciones religiosas.

El prestigio y admiración que gozó Newton en vida queda reflejado en las palabras de Alexander Pope:

*Nature, and Nature's Laws lay hid in Night:
God said, Let Newton be – and All was light.*

Y ¿qué pensaba el propio Newton de sí mismo? Escuchemos sus palabras, ya casi al final de su vida.

No sé cómo puedo ser visto por el mundo, pero a mí me parece haber sido solamente como un niño que juega al borde del mar, y que se divierte al encontrar de vez en cuando una piedra más pulida o una concha más bonita de lo normal, mientras que el gran océano de la verdad yace ante mí completamente desconocido.

Newton murió en la noche del 20 de marzo de 1727, y fue enterrado con grandes honores en la abadía de Westminster entre los grandes hombres de Inglaterra.

3.5. Leibniz y el cálculo de diferencias



Figura 20. Leibniz

Gottfried Wilhelm Leibniz (1646 - 1716) nació en Leipzig (Alemania) en el seno de una piadosa familia luterana. A los quince años entró en la Universidad de su ciudad natal donde estudió una gran variedad de materias incluyendo derecho, teología, filosofía y matemáticas. Se doctoró a la edad de 21 años en la Universidad de Altdorf, en Nuremberg, donde le fue ofrecido un puesto de profesor que él rechazó.

A lo largo de su vida, Leibniz realizó múltiples actividades. Como abogado y diplomático trabajó para el Príncipe elector arzobispo de Maguncia y, desde 1676 hasta su muerte, para los Duques de Brunswick-Luneburgo (conocidos como príncipes electores de Hanover desde 1692), lo que le llevó a viajar por gran parte de Europa. Inventó una máquina de calcular, la primera máquina de este

tipo capaz de realizar las operaciones de multiplicación, división y extracción de raíces cuadradas. Como ingeniero trabajó en prensas hidráulicas, molinos de viento y desarrolló proyectos para drenar el agua de las minas de plata de las montañas de Harz en la Baja Sajonia. Como historiador escribió la historia de la casa de Brunswick, realizando muchas investigaciones genealógicas. Trabajó también como bibliotecario en la ciudad de Hanover.

Leibniz fue un pensador profundo. Como filósofo se propuso la creación de un álgebra del pensamiento humano, algo así como un lenguaje simbólico universal para escribir los razonamientos con símbolos y fórmulas, cuyas reglas de combinación permitieran reducir todo discurso racional a cálculos rutinarios. Esto explica el gran interés de Leibniz en desarrollar una notación matemática apropiada para su cálculo; de hecho, su notación, muy superior a la de Newton, es la que usamos actualmente. Leibniz fundó la Academia de Ciencias de Berlín en 1700 y fue su primer presidente; también fue uno de los fundadores de la primera revista científica alemana, el *Acta Eruditorum*.

Aunque Leibniz publicó poco, mantuvo correspondencia con más de 600 eruditos y se han conservado sus manuscritos que están en el archivo que lleva su nombre en la ciudad de Hannover. Las contribuciones de Leibniz al álgebra (determinantes, resolución de ecuaciones), la historia natural, la geología y la lingüística son también importantes.

En 1672, estando en París en misión diplomática, Leibniz se dedicó intensamente al estudio de la matemática superior teniendo como guía al matemático y físico Christian Huygens (1629 - 1695). En los años 1673 y 1676 realizó, también en misión diplomática, dos viajes a Londres donde tuvo acceso al manuscrito de Newton *De Analysi*, circunstancia que se usó para acusar, hoy sabemos que sin motivo alguno, a Leibniz de plagio cuando se produjo la agria controversia sobre

la prioridad en el descubrimiento del Cálculo. Los progresos matemáticos realizados por Leibniz en estos cuatro años fueron extraordinarios.

En las matemáticas de Leibniz son importantes los estudios sobre sucesiones numéricas y sus sucesiones de diferencias consecutivas asociadas. Dada una sucesión de números:

$$a_1, a_2, a_3, a_4, \dots, a_{n-1}, a_n, \dots$$

Podemos formar la sucesión de sus diferencias primeras:

$$b_1 = a_1, b_2 = a_2 - a_1, b_3 = a_3 - a_2, b_4 = a_4 - a_3, \dots, b_n = a_n - a_{n-1}, \dots$$

Leibniz se había dado cuenta de la relación:

$$b_1 + b_2 + b_3 + \dots + b_n = a_n$$

lo que indica que las sucesiones de diferencias pueden sumarse fácilmente, y que el proceso de formar la sucesión de diferencias y después sumarla recupera la sucesión inicial, es decir, que se trata de operaciones inversas una de la otra. Esta sencilla idea, cuando se lleva al campo de la geometría, conduce al concepto central del cálculo de Leibniz que es el de “diferencial”, el cual tuvo para él diferentes significados en distintas épocas.

Leibniz consideraba una curva como un polígono de infinitos lados de longitud infinitesimal. Con una tal curva se asocia una sucesión de abscisas $x_1, x_2, x_3, x_4, \dots$ y una sucesión de ordenadas $y_1, y_2, y_3, y_4, \dots$ donde los puntos (x_i, y_i) están todos ellos en la curva y son algo así como los “vértices” de la poligonal de infinitos lados que forma la curva. La diferencia entre dos valores sucesivos de x es llamada la *diferencial* de x y se representa por dx , significado análogo tiene dy . El diferencial dx es una cantidad fija, no nula, infinitamente pequeña en comparación con x , de hecho es una cantidad infinitesimal. Los lados del polígono que constituye la curva son representados por ds . Resulta así el *triángulo característico* de Leibniz que es el mismo que ya había sido considerado por Barrow.

Curiosamente, los términos “abscisa”, “ordenada” y “coordenadas”, tan propios de la geometría analítica, no fueron usados nunca por Descartes sino que son debidos a Leibniz; y mientras que nosotros hablamos de “diferenciales”, Leibniz siempre hablaba de “diferencias”.

El triángulo característico tiene lados infinitesimales dx , dy , ds y se verifica la relación $(ds)^2 = (dx)^2 + (dy)^2$. El lado ds sobre la curva o polígono se hace coincidir con la tangente a la curva en el punto (x, y) . La pendiente de dicha tangente viene dada por $\frac{dy}{dx}$, que es un cociente de diferenciales al que Leibniz llamó *cociente diferencial*. Leibniz nunca consideró la derivada como un límite.

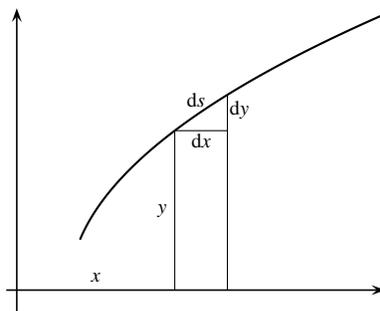


Figura 21. Triángulo característico

Leibniz investigó durante algún tiempo hasta encontrar las reglas correctas para diferenciar productos y cocientes. Dichas reglas se expresan fácilmente con su notación diferencial:

$$d(xy) = ydx + xdy, \quad d\left(\frac{x}{y}\right) = \frac{ydx - xdy}{y^2}$$

La manera en que Leibniz llegó a estas fórmulas pudo ser como sigue. Consideremos

$$z_n = \left(\sum_{j=1}^n x_j\right) \left(\sum_{j=1}^n y_j\right)$$

Entonces

$$z_{n+1} - z_n = x_{n+1} \sum_{j=1}^{n+1} y_j + y_{n+1} \sum_{j=1}^n x_j \quad (7)$$

Si interpretamos, al estilo de Leibniz, que x_j e y_j son diferencias de valores consecutivos de las cantidades x e y respectivamente, entonces los valores de dichas cantidades vendrán dados por las sumas respectivas $x = \sum_{j=1}^n x_j$ e $y = \sum_{j=1}^{n+1} y_j$, mientras que $dx = x_{n+1}$ y $dy = y_{n+1}$ por ser diferencias de valores consecutivos. De la misma forma, $z_{n+1} - z_n$ sería la diferencial de $z = xy$. Por tanto, la igualdad 7 es interpretada por Leibniz en la forma $d(xy) = xdy + ydx$, lo que lleva a la regla para la diferencial de un producto.

A partir de la regla para la diferencial de un producto, Leibniz obtuvo la regla correspondiente para la diferencial de un cociente $z = \frac{x}{y}$. Poniendo $x = zy$ se tiene que $dx = ydz + zdy$, de donde despejando dz , resulta:

$$dz = \frac{dx - zdy}{y} = \frac{dx - \frac{x}{y}dy}{y} = \frac{ydx - xdy}{y^2}$$

Además, dicha notación tiene una gran potencialidad heurística, como ya hemos visto al estudiar la derivada de una función compuesta.

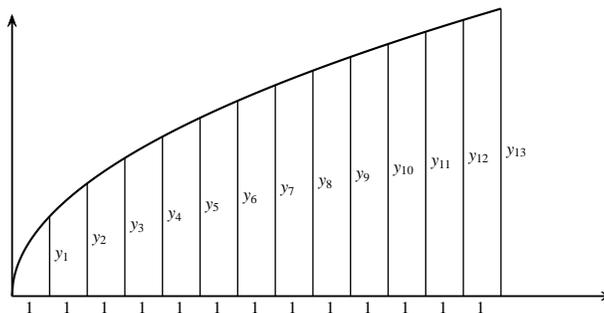


Figura 22. Aproximación de una cuadratura

Consideremos ahora una curva como la de la figura 22 con una sucesión de ordenadas trazadas a intervalos de longitud unidad. La suma de las ordenadas es una aproximación de la cuadratura de la curva (del área bajo la curva), y la diferencia entre dos ordenadas sucesivas es aproximadamente igual a la pendiente de la correspondiente tangente. Cuanto más pequeña se elija la unidad 1, tanto mejor serán estas aproximaciones. Leibniz razonaba que si la unidad pudiera ser tomada *infinitamente pequeña*, estas aproximaciones se harían exactas, esto es, la cuadratura sería igual a la suma de las ordenadas, y la pendiente de la tangente sería igual a la diferencia de dos ordenadas sucesivas. Como las operaciones de tomar diferencias y sumar son recíprocas entre sí, dedujo Leibniz que el cálculo de cuadraturas y de tangentes también eran operaciones inversas una de otra.

Las investigaciones de Leibniz sobre la integración y el origen de sus notaciones para la integral y los diferenciales, pueden seguirse con todo detalle en una serie de manuscritos del 25 de octubre al 11 de noviembre de 1675. Nos ocuparemos de ello en el capítulo dedicado a la integración. En 1676 Leibniz ya había obtenido prácticamente todos los resultados descubiertos por Newton un poco antes.

La primera publicación sobre cálculo diferencial fue el artículo de Leibniz *Nova methodus pro maximis et minimis, itemque tangentibus, quae nec fractals nec irrationales quantitates moratur, et singulare pro illis calculi genus*, que fue publicado en *Acta Eruditorum* hace ya más de tres siglos, en 1684. En este trabajo, Leibniz definía el diferencial dy de forma que evitaba el uso de las sospechosas cantidades infinitesimales. Poco después, en 1686, Leibniz publicó un trabajo con sus estudios sobre la integración.

Reconocido hoy día como un genio universal, Leibniz vivió sus últimos años en Hannover en un aislamiento cada vez mayor y murió el 14 de noviembre de 1716. A su entierro solamente asistió su secretario.

3.6. Desarrollo del cálculo diferencial

Aunque las publicaciones de Leibniz eran breves y difíciles de leer, su cálculo, más sencillo de entender que el de Newton y provisto de una excelente notación, triunfó pronto en el continente europeo logrando grandes éxitos, mientras que en Inglaterra la fidelidad a la teoría de fluxiones y a la notación newtoniana condujo a un cierto aislamiento, agravado por sentimientos nacionales y la disputa sobre la prioridad, y no consiguió éxitos comparables a los del continente.

Los hermanos Jakob y Johann Bernoulli, matemáticos y profesores de la universidad de Basilea, estudiaron los trabajos de Leibniz con quien iniciaron una productiva correspondencia. A partir de 1690 publicaron una serie de trabajos en el *Acta Eruditorum* y en otras revistas, poniendo de manifiesto que el cálculo de Leibniz era una herramienta poderosa con la que había que contar. Para divulgar dicha herramienta era preciso un buen libro de texto que explicara con detalle los pormenores del nuevo cálculo. Dicho libro apareció bien pronto, en 1696, y su autor fue el matemático y noble francés Guillaume François, marqués de L'Hôpital. El título del libro, del que ya hemos dado noticia en anteriores capítulos, era *Analyse des infiniment petits pour l'intelligence des lignes courbes*. Hoy sabemos que los resultados originales que aparecen en dicho libro son debidos no a L'Hôpital sino a su profesor Johann Bernoulli.

En su libro, L'Hôpital desarrollaba el cálculo diferencial tal como había sido concebido por Leibniz, es decir, usando cantidades infinitesimales para las que se establecían ciertas reglas de cálculo. La definición de diferencial es como sigue: “*La parte infinitamente pequeña en que una cantidad variable es aumentada o disminuida de manera continua, se llama la diferencial de esta cantidad*”. Para trabajar con infinitésimos se establece la siguiente regla: “*Dos cantidades cuya diferencia es otra cantidad infinitamente pequeña pueden intercambiarse una por la otra*”.

Los escritos de los Bernoulli, Leibniz y L'Hôpital popularizaron el cálculo leibniziano y ya en la primera década del siglo XVIII otros matemáticos se interesaron por él. La potencialidad del concepto de derivada se puso de manifiesto en las aplicaciones del cálculo a la física newtoniana.

Para no hacer excesivamente larga esta exposición, voy a resumir muy esquemáticamente los puntos clave en el desarrollo del cálculo diferencial.

- El descubrimiento en 1715 por Brook Taylor de las llamadas series de Taylor, que se convirtieron en una herramienta básica para el desarrollo del cálculo y la resolución de ecuaciones diferenciales.
- El extraordinario trabajo, tanto por su asombrosa amplitud como por sus notables descubrimientos, de Leonhard Euler (1707 - 1783) que, sin duda, es la figura principal de las matemáticas en el siglo XVIII. En sus tres grandes tratados, escritos en latín, *Introductio in analysin infinitorum* (1748), *Institutiones calculi differentiales* (1755) e *Institutiones calculi*

integralis (1768), Euler dio al cálculo la forma que conservó hasta el primer tercio del siglo XIX. El cálculo, que inicialmente era un cálculo de variables o, más exactamente, de cantidades geométricas variables, y de ecuaciones, se fue transformando, por influencia de Euler, en un cálculo de funciones.

- La propuesta de Joseph Louis Lagrange (1736 - 1813) de fundamentar el cálculo sobre un álgebra formal de series de potencias. Si bien la idea de Lagrange de evitar el uso de límites no era acertada, su propuesta, concretada en su obra *Théorie des fonctions analytiques* (1797), tuvo el efecto de liberar el concepto de derivada de sus significaciones más tradicionales. De hecho, la terminología “función derivada”, así como la notación $f'(x)$ para representar la derivada de una función f , fueron introducidas por Lagrange en dicho texto. A partir de este momento la derivada deja de ser algo de naturaleza imprecisa (fluxión o cociente diferencial) y empieza a ser considerada simplemente como una función.
- Los problemas planteados por las series de Fourier. Dichas series hacen sus primeras apariciones a mitad del siglo XVIII en relación con el problema de la cuerda vibrante, y nacen oficialmente en el trabajo de Joseph Fourier (1768 - 1830) *Théorie analytique de la chaleur* (1822). Tales series plantean problemas relacionados con las ideas centrales del análisis: el concepto de función, el significado de la integral y los procesos de convergencia.
- El proceso de “algebraización del análisis” que tiene lugar en los dos últimos tercios del siglo XIX y que culmina con la fundamentación del análisis sobre el concepto de límite (Bolzano, Cauchy, Weierstrass) y la teoría de los números reales (Dedekind, Cantor). Lo esencial de este proceso ya ha sido considerado en el tema anterior.

4. Evolución de la idea de integral

4.1. Problemas de cuadraturas en las matemáticas griegas

⁶ Los problemas de cuadraturas son problemas geométricos que consisten en lo siguiente: dada una figura, construir un cuadrado con área igual a la de la figura dada. Esta construcción debía hacerse con regla no graduada y compás, siguiendo unas normas precisas. Según lo establecido en los *Elementos* de Euclides (c. 300 a.C.) la construcción debe constar de un número finito de pasos, cada uno de ellos consistente en:

- Trazar una recta que una dos puntos.
- Trazar una circunferencia de centro y radio arbitrarios.

⁶Para escribir estas notas históricas he seguido de cerca los trabajos de Kirsti Andersen [2], Israel Kleiner [18], González Urbaneja [16] y H. J. M. Bos [6].

- Intersecar dos de las figuras anteriores.

Son famosos los problemas de la cuadratura del círculo, la trisección de un ángulo, la duplicación del cubo y la inscripción de polígonos regulares en una circunferencia. En la antigua Grecia se sabía cuadrar cualquier polígono.

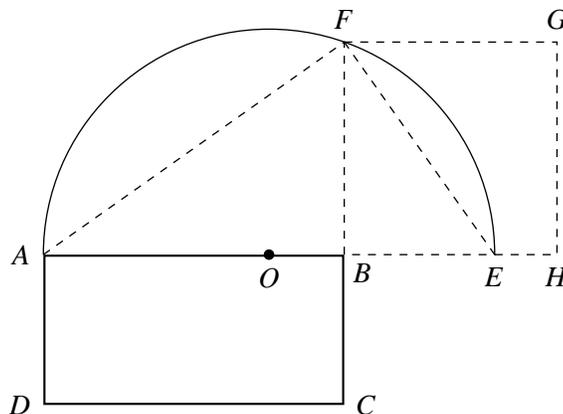


Figura 23. Cuadratura de un rectángulo

Para cuadrar el rectángulo $ABCD$ de la figura 23 se procede de la forma siguiente:

- 1) Se prolonga el lado AB y se determina sobre él un punto E tal que $BE = BC$.
- 2) Se traza con centro en el punto medio O de AE una semicircunferencia de radio OE .
- 3) Se traza por B una perpendicular a AE y se determina su punto de corte F con la semicircunferencia.
- 4) El segmento FB es el lado de un cuadrado cuya área es igual a la del rectángulo $ABCD$. Esto es consecuencia de que la altura FB de un triángulo rectángulo AFE es media proporcional entre las dos partes en que divide a la hipotenusa, es decir, $FB/AB = BE/FB$, por lo que $FB^2 = AB \cdot BE = AB \cdot BC$.

A partir de aquí es fácil obtener la cuadratura de un triángulo, lo que permite obtener la cuadratura de cualquier polígono descomponiéndolo en triángulos. Los matemáticos griegos inventaron un procedimiento, que se conoce con el nombre de “exhausción”, por el cual podían lograr la cuadratura de algunas regiones delimitadas por curvas. Se atribuye a Eudoxo de Cnido (c. 400 - 347 a.C.) la invención de este método, que fue perfeccionado posteriormente por Arquímedes (c. 287 - 212 a.C.). El siguiente es un notable ejemplo de su aplicación.

4.1.1. Cuadratura de un segmento de parábola por Arquímedes

Teorema. El área del segmento parabólico PVQ es igual a cuatro tercios el área del triángulo inscrito $\triangle PVQ$.

Demostración. Esta demostración aparece en una carta que escribe Arquímedes a su amigo Dositheus, obra que se conoce con el nombre de *Sobre la Cuadratura de la Parábola*. La demostración consiste en hacer una descomposición exhaustiva del segmento parabólico por medio de triángulos de una forma muy ingeniosa. Empezaremos explicando la construcción geométrica de la figura 24.

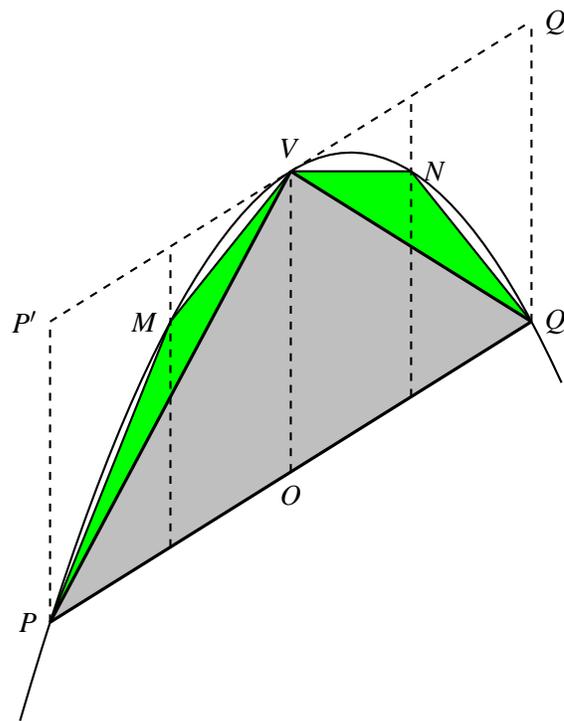


Figura 24. Cuadratura de un segmento de parábola

Una *cuerda* PQ de una parábola es un segmento que une dos de sus puntos. La región plana acotada, cuya frontera está formada por la cuerda PQ y el arco de la parábola comprendido entre los puntos P y Q se llama un *segmento parabólico*. El *vértice* de un segmento parabólico es el punto de la parábola en el cual la tangente es paralela a la cuerda que define el segmento.

Se verifica que el vértice de un segmento parabólico PVQ es el punto intersección con la parábola de la recta paralela al eje de la parábola que pasa por el punto medio $O = \frac{1}{2}(P + Q)$ del segmento PQ .

El triángulo $\triangle PVQ$ cuya base es el segmento PQ y cuyo otro vértice es el vértice V del seg-

mento parabólico le llamaremos el triángulo inscrito.

En la figura 24 se han representado también los triángulos $\triangle PMV$ y $\triangle VNQ$ inscritos, respectivamente, en los segmentos parabólicos determinados por las cuerdas PV y VQ .

La primera parte de la demostración consiste en calcular el área de los dos triángulos $\triangle PMV$ y $\triangle VNQ$. Arquímedes demuestra que

$$\lambda(\triangle VNQ) = \frac{1}{4}\lambda(\triangle VOQ), \quad \lambda(\triangle VMP) = \frac{1}{4}\lambda(\triangle VOP)$$

Por tanto

$$\lambda(\triangle VNQ) + \lambda(\triangle VMP) = \frac{1}{4}\lambda(\triangle PVQ) \quad (8)$$

Llamando S al área del triángulo $\triangle PVQ$, el área de los dos nuevos triángulos es $\frac{1}{4}S$. Naturalmente, este proceso se puede repetir ahora con cada uno de los cuatro segmentos parabólicos determinados por las cuerdas PM , MV , VN y NQ inscribiendo en ellos los respectivos triángulos, la suma de cuyas áreas será igual a $\frac{1}{16}S$. Y puede repetirse indefinidamente.

Nosotros ahora acabaríamos calculando el área del segmento parabólico por

$$\sum_{n=0}^{\infty} \frac{1}{4^n} S = \frac{4}{3} S$$

Pero Arquímedes, que no sabe de convergencia de series ni falta que le hace, razona de forma muy elegante por medio de la doble reducción al absurdo usual en la matemática griega.

Para ello hace uso de la llamada *propiedad arquimediana* o *axioma de Arquímedes*. Este axioma aparece en el libro de Arquímedes *La Esfera y el Cilindro* así como en *Sobre la Cuadratura de la Parábola* y en *Espirales*. Al parecer, dicho axioma fue ya formulado por Eudoxo. Como sabemos, la propiedad arquimediana establece que:

Dadas magnitudes cualesquiera $a > 0$ y $b > 0$, siempre es posible, por pequeña que sea a y grande que sea b , conseguir que un múltiplo conveniente de a exceda a b , es decir $na > b$ para algún número natural n .

Partiendo de la propiedad arquimediana se deduce fácilmente el siguiente resultado, llamado *principio de convergencia de Eudoxo*, en el que se basa el llamado *método de exhaustión* griego:

Si de cualquier magnitud sustraemos una parte no menor que su mitad, y si del resto sustraemos de nuevo una cantidad no menor que su mitad, y si continuamos repitiendo este procesos de sustracción, terminaremos por obtener como resto una magnitud menor que cualquier magnitud del mismo tipo dada de antemano.

Arquímedes razona como sigue. Sea K el área del segmento parabólico PVQ .

(I) Supongamos que $K > \frac{4}{3}S$; es decir, que $K - \frac{4}{3}S > 0$.

Como el área del triángulo inscrito en un segmento parabólico PVQ es la mitad del área del paralelogramo circunscrito $PP'QQ'$, la cual, a su vez, es mayor que el área del segmento, se sigue que el área del triángulo inscrito en un segmento parabólico es mayor que la mitad del área de dicho segmento, lo que permite aplicar el principio de convergencia de Eudoxo.

Por tanto, en la sucesión de áreas

$$K, K - S, K - \left(S + \frac{1}{4}S\right), K - \left(S + \frac{1}{4}S + \frac{1}{16}S\right), \dots$$

cada una es menor que la mitad de la que le precede y, por tanto, en virtud del citado principio, podemos concluir que en alguna etapa se tendrá que

$$K - \frac{4}{3}S > K - \left(S + \frac{1}{4}S + \frac{1}{16}S + \dots + \frac{1}{4^n}S\right)$$

Esto implica que

$$S + \frac{1}{4}S + \frac{1}{16}S + \dots + \frac{1}{4^n}S > \frac{4}{3}S$$

lo que es contradictorio con la igualdad, conocida por Arquímedes, que dice que:

$$S + \frac{1}{4}S + \frac{1}{16}S + \dots + \frac{1}{4^n}S = \frac{4}{3}S - \frac{1}{3} \frac{1}{4^n}S \quad (9)$$

la cual implica que $S + \frac{1}{4}S + \frac{1}{16}S + \dots + \frac{1}{4^n}S < \frac{4}{3}S$. Por tanto, no puede ser $K > \frac{4}{3}S$.

(II) Supongamos que $K < \frac{4}{3}S$; es decir, que $\frac{4}{3}S - K > 0$.

Como cada una de las áreas $S, \frac{1}{4}S, \frac{1}{16}S, \dots, \frac{1}{4^n}S$ es menor que la mitad de la que le precede y, por tanto, en virtud del principio de convergencia de Eudoxo, podemos concluir que en alguna etapa se tendrá que $\frac{1}{4^n}S < \frac{4}{3}S - K$. Entonces

$$\frac{4}{3}S - K > \frac{1}{4^n}S > \frac{1}{3} \frac{1}{4^n}S = \frac{4}{3}S - \left(S + \frac{1}{4}S + \frac{1}{16}S + \dots + \frac{1}{4^n}S\right)$$

Lo que implicaría que

$$K < S + \frac{1}{4}S + \frac{1}{16}S + \dots + \frac{1}{4^n}S$$

Que es absurdo pues la suma de la derecha es el área de un polígono inscrito en el segmento parabólico. Por tanto, no puede ser $K < \frac{4}{3}S$.

La única posibilidad es $K = \frac{4}{3}S$. □

4.1.2. El Método de Arquímedes

En su tratado *El Método*, que se creía perdido y fue descubierto en 1906, Arquímedes obtiene la cuadratura de la parábola por medios mecánicos usando el principio de la palanca. Aunque el propio Arquímedes *reconoce que esa forma de proceder no es una demostración*, merece la pena decir algo sobre ella.

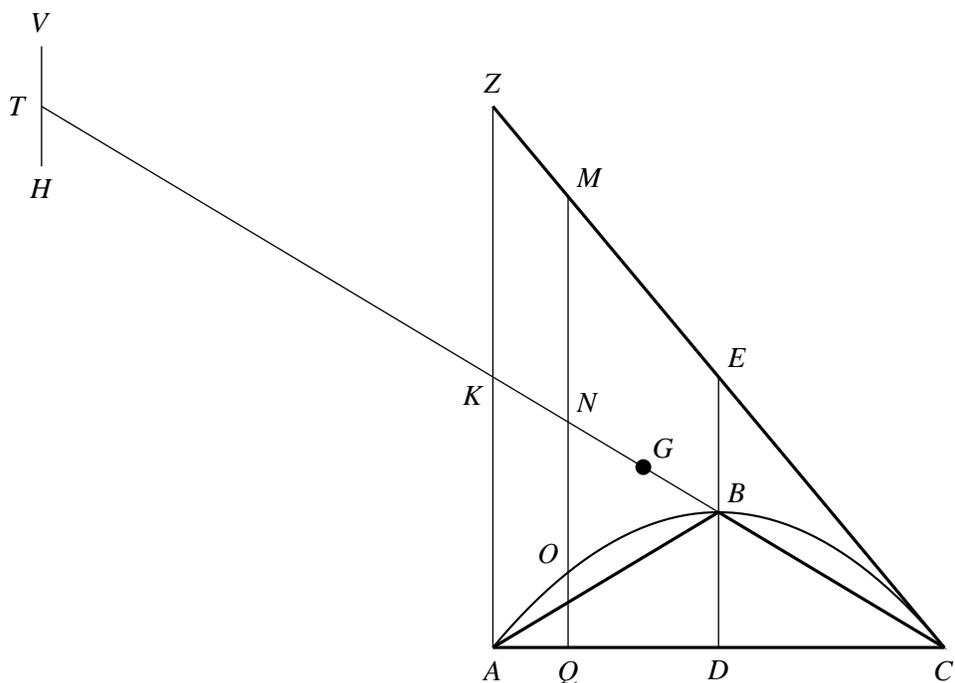


Figura 25. El Método de Arquímedes

La recta CZ es la tangente a la parábola en C , B es el vértice del segmento parabólico. El segmento AZ es perpendicular a la cuerda AC , CT es la recta que pasa por el punto C y el vértice B de forma que K es el punto medio del segmento CT . Se considera CT como un brazo de palanca con fulcro en K .

Por ser ABC una parábola, se sabe que la subtangente ED en un punto C es igual al doble de la abscisa BD (conviene imaginarse la parábola girada 90 grados), es decir, $ED = 2BD$, de donde, $EB = BD$. Deducimos, por la semejanza de triángulos en la figura, que $MN = NQ$ y $ZK = KA$.

Arquímedes demuestra en *Sobre la Cuadratura de la Parábola* que

$$\frac{CA}{AQ} = \frac{MQ}{QO}$$

Y, como también es $\frac{CA}{AQ} = \frac{CK}{KN}$, y por construcción es $TK = CK$, obtenemos que

$$\frac{TK}{KN} = \frac{MQ}{QO} \iff TK \cdot QO = KN \cdot MQ$$

Si ahora trasladamos al punto T un segmento de longitud igual a QO y lo ponemos como en la figura el segmento VH de modo que su centro de gravedad sea el punto T , la igualdad anterior nos dice que el segmento $VH = QO$ queda equilibrado por el segmento MQ , pues el producto de dichos segmentos por la longitud correspondiente del brazo de palanca con fulcro en K es la misma. Obsérvese que N es el centro de gravedad del segmento MQ . Deducimos que K es el centro de gravedad de los segmentos VH y MQ .

Análogamente puede razonarse con cualquier paralela al eje de la parábola ED , todas ellas estarán en equilibrio con los segmentos determinados sobre ellas por el segmento parabólico trasladados al punto T , de manera que el centro de gravedad de cada par de segmentos será el punto K .

Ahora bien, los segmentos paralelos a DE “*componen*” el triángulo $\triangle AZC$ y los correspondientes segmentos dentro del segmento parabólico “*componen*” dicho segmento parabólico. Por tanto el triángulo AZC “*permaneciendo en su lugar*”, estará en equilibrio respecto del punto K con el segmento parabólico trasladado hasta tener su centro de gravedad en T , de manera que el centro de gravedad del conjunto de ambos será el punto K .

Dividimos ahora CK por el punto G de forma que CK sea el triple de KG , el punto G será el centro de gravedad del triángulo AZC , y puesto que el triángulo AZC , “*permaneciendo en su lugar*” está en equilibrio, respecto del punto K , con el segmento parabólico ABC , trasladado con centro de gravedad en T , y que G es el centro de gravedad del triángulo AZC , se verifica, por consiguiente, que la razón del triángulo AZC al segmento parabólico ABC colocado alrededor del centro T es igual a la razón de TK a KG . Ahora bien, siendo TK triple de KG , el triángulo AZC será triple del segmento parabólico ABC . Además, el triángulo AZC es cuádruple del triángulo inscrito ABC , ya que ZK es igual que KA y KA es doble de BD al ser AD igual que DC . Concluimos que el segmento parabólico ABC equivale a cuatro tercios del triángulo inscrito ABC . \square

4.1.3. Área de una espiral

El siguiente ejemplo de cuadratura sigue un procedimiento que, traducido a las notaciones actuales, es prácticamente el mismo de la integral de Riemann.

La espiral de Arquímedes es la curva que describe un punto material que se mueve con velocidad uniforme a lo largo de una semirrecta que gira con velocidad angular uniforme alrededor de

su extremo. Es un ejemplo de las llamadas *curvas mecánicas*. La ecuación polar de una espiral de Arquímedes es de la forma $\rho = a\vartheta$, donde $a > 0$ es una constante.

teorema. El área del primer ciclo de una espiral es igual a una tercera parte del área del círculo circunscrito.

Demostración. Consideremos una espiral de Arquímedes de ecuación polar $\rho = a\vartheta$ y calculemos el área cuando el ángulo polar varía desde 0 a 2π , es decir, de la primera vuelta de la espiral. El radio del círculo circunscrito es $2\pi a$. Para ello dividimos este círculo en sectores de amplitud $\vartheta = 2\pi/n$, desde $\vartheta = 2\pi k/n$ a $\vartheta = 2\pi(k+1)/n$ para $k = 0, 1, \dots, n-1$. En cada sector examinamos el arco de espiral que queda dentro del mismo y acotamos el área correspondiente a dicho arco de espiral entre las áreas de dos sectores circulares. Teniendo en cuenta que el área de un sector circular de radio r y amplitud φ radianes es $\frac{1}{2}r^2\varphi$, resulta que el área de sector circular más grande inscrito en cada arco de espiral es $\frac{1}{2}(a2\pi k/n)^2(2\pi/n)$, y el área de sector circular más pequeño circunscrito a cada arco de espiral es $\frac{1}{2}(a2\pi(k+1)/n)^2(2\pi/n)$. Deducimos que el área, S , de la espiral verifica que:

$$\sum_{k=0}^{n-1} \frac{1}{2} \left(\frac{a2\pi k}{n} \right)^2 \frac{2\pi}{n} = \frac{4\pi^3 a^2}{n^3} \sum_{k=0}^{n-1} k^2 < S < \sum_{k=1}^n \frac{1}{2} \left(\frac{a2\pi k}{n} \right)^2 \frac{2\pi}{n} = \frac{4\pi^3 a^2}{n^3} \sum_{k=1}^n k^2$$

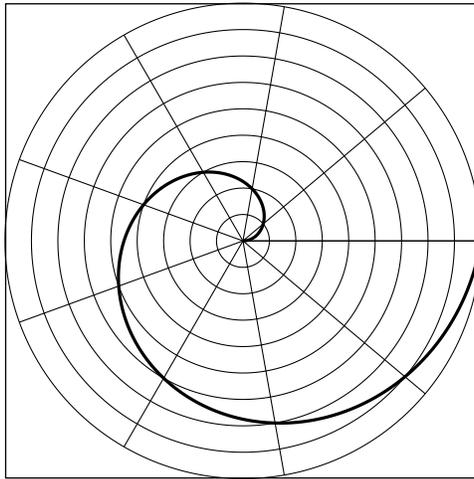


Figura 26. Cuadratura de una espiral

Arquímedes conocía que $\sum_{k=1}^n k^2 = \frac{1}{6}n(n+1)(2n+1)$. Usando este resultado podemos escribir la desigualdad anterior en la forma:

$$4\pi^3 a^2 \frac{1}{6} \left(1 - \frac{1}{n} \right) \left(2 - \frac{1}{n} \right) < S < 4\pi^3 a^2 \frac{1}{6} \left(1 + \frac{1}{n} \right) \left(2 + \frac{1}{n} \right)$$

Pongamos $K = \frac{1}{3}\pi(2\pi a)^2$ que es una tercera parte del área del círculo circunscrito. Restando K en la desigualdad anterior y haciendo operaciones sencillas, obtenemos que:

$$K \left(-\frac{3}{2n} + \frac{1}{2n^2} \right) < S - K < K \left(\frac{3}{2n} + \frac{1}{2n^2} \right);$$

y como $1/n^2 \leq 1/n$, obtenemos que $-2K/n < S - K < 2K/n$. Usando ahora el axioma de Arquímedes se concluye que $S = K$. \square

4.2. La integración antes del Cálculo

4.2.1. Los indivisibles de Cavalieri

El método de integración geométrica que se consideraba ideal durante la primera mitad del siglo XVII era el método de exhaustión que había sido inventado por Eudoxo y perfeccionado por Arquímedes. El nombre es desafortunado porque la idea central del método es la de evitar el infinito y por lo tanto este método no lleva a un “agotamiento” de la figura a determinar.

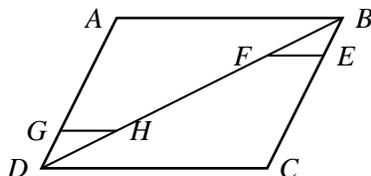
Entre los matemáticos del siglo XVII era general el deseo de encontrar un método para obtener resultados y que, a diferencia del método de exhaustión, fuera directo. Y mejor que mejor si el nuevo método, aparte de dar resultados, pudiera ser utilizado para demostrarlos.

El camino que siguieron fue el que se deriva de una concepción intuitiva inmediata de las magnitudes geométricas. Se imaginaron un área como formada, por ejemplo, por un número infinito de líneas paralelas. Kepler ya había hecho uso de métodos infinitesimales en sus obras; el interés que se tomó en el cálculo de volúmenes de toneles de vino dio como resultado un libro *Nova stereometria doliurum vinariorum* (1615). En él consideraba sólidos de revolución como si estuvieran compuestos de diversas maneras por una cantidad infinita de partes sólidas. Por ejemplo, consideraba una esfera como formada por un número infinito de conos con vértice común en el centro y base en la superficie de la esfera. Esto le conducía al resultado de que la esfera es igual en volumen al cono que tiene como altura el radio de la esfera y como base un círculo igual al área de la esfera, es decir un círculo con el diámetro de la esfera como radio.

Galileo tenía la intención de escribir un libro sobre indivisibles, pero este libro nunca se publicó.

Bonaventura Cavalieri (1598 - 1647), discípulo de Galileo y profesor en la Universidad de Bolonia, publicó en 1635 un tratado *Geometria Indivisibilibus Continuorum Nova quadam Ratione Promota* en el que, siguiendo ideas de Kepler y Galileo, desarrolló una técnica geométrica para calcular cuadraturas, llamada *método de los indivisibles*. En este método, un área de una región plana se considera formada por un número infinito de segmentos paralelos, cada uno de ellos se interpreta como un rectángulo infinitamente estrecho; un volumen se considera compuesto por un

número infinito de áreas planas paralelas. A estos elementos los llama los *indivisibles* de área y volumen respectivamente. En líneas generales los “indivisibilistas” mantenían, como expresa Cavalieri en sus *Exercitationes Geometricae Sex* (1647), que *una línea está hecha de puntos como una sarta de cuentas; el plano está hecho de líneas, como un tejido de hebras y un sólido de áreas planas como un libro de hojas.*



La forma en que se aplicaba el método o principio de Cavalieri puede ilustrarse como sigue. Para demostrar que el paralelogramo $ABCD$ tiene área doble que cualquiera de los triángulos ABD o BCD , hace notar que cuando $GD = BE$, se tiene que $GH = FE$. Por tanto los triángulos ABD y BCD están constituidos por igual número de líneas iguales, tales como GH y FE , y por tanto sus áreas deben ser iguales.

4.2.2. Cuadratura de la cicloide por Roberval

En 1630, Mersenne, propuso a sus amigos matemáticos hacer la cuadratura de la cicloide. Esta fue llevada a cabo por Gilles Personne de Roberval en 1634, utilizando esencialmente el método de los indivisibles de Cavalieri. Recuerda que la cicloide es la curva que describe un punto de una circunferencia que rueda sin deslizar.

En la figura 27, sea $QMNS$ la mitad de un arco de la cicloide generada por el círculo de radio r centrado en O . El área del rectángulo $QMNP$ es el doble del área del círculo. Construimos segmentos de línea infinitesimales horizontales, AB , con longitud determinada por la distancia horizontal entre el diámetro PQ y la circunferencia. Cada punto C de la cicloide lo sometemos a una traslación horizontal hasta el punto D , según el correspondiente segmento $AB = CD$, y así obtenemos la curva QRN , llamada compañera de la cicloide. Por la construcción realizada, las secciones horizontales del semicírculo y de la región comprendida entre la cicloide y su curva compañera son segmentos de igual longitud, por lo que dicha región tiene área igual a la mitad del círculo. Por otra parte, la curva compañera de la cicloide divide en dos partes iguales al rectángulo $QMNP$, pues, como Roberval demostró, las secciones horizontales de altura a y $2r - a$ dan en cada una de las partes en que dicha curva divide al rectángulo, segmentos iguales XY y UV . Deducimos así que el área

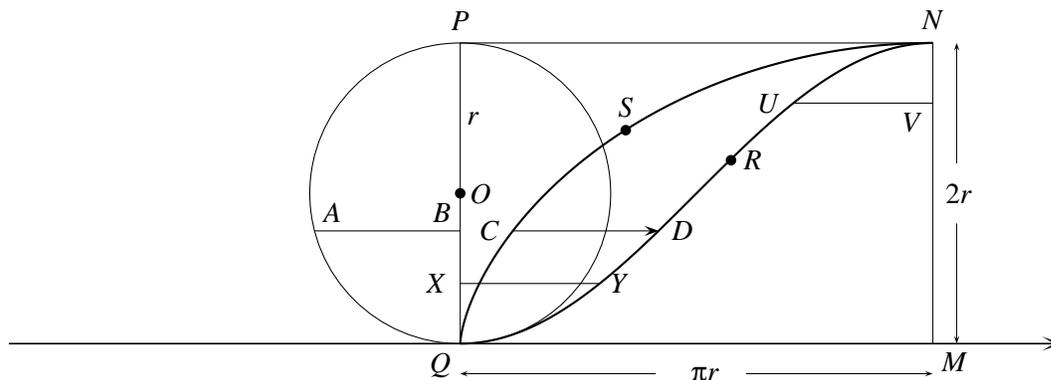


Figura 27. Cuadratura de la cicloide

encerrada por la mitad de un arco de cicloide es $\pi r^2 + \frac{1}{2}\pi r^2 = \frac{3}{2}\pi r^2$. Por tanto, concluimos que el área encerrada por un arco de la cicloide es tres veces el área del círculo que la genera.

Los matemáticos no se mostraban de acuerdo acerca del valor que había que dar a una demostración por el método de los indivisibles. La mayoría de los que se preocupaban de la cuestión consideraban el método de los indivisibles sólo como un método heurístico y creían que era aún necesaria una demostración por exhaustión.

4.2.3. Parábolas e hipérbolas de Fermat

La cuadratura de las curvas definidas por $y = x^n$ donde n es un número natural o bien un entero negativo $n \neq -1$, había sido realizada para $n = 1, 2, \dots, 9$ por Cavalieri, aunque podemos remontarnos hasta Arquímedes que había resuelto geoméricamente los casos correspondientes a $n = 1, 2, 3$. Fermat, con una ingeniosa idea, logró obtener la cuadratura de áreas limitadas por arcos de hipérbolas generalizadas $x^n y^m = 1$ ($m, n \in \mathbb{N}$).

Fermat seguía un método clásico de exhaustión, pero con una idea feliz que consistió en considerar rectángulos infinitesimales inscritos en la figura a cuadrar cuyas bases estaban en progresión geométrica. Fermat considera al principio las hipérbolas $yx^n = k$ y manifiesta:

Digo que todas estas infinitas hipérbolas, excepto la de Apolonio, que es la primera, pueden ser cuadradas por el método de la progresión geométrica, de acuerdo a un procedimiento uniforme general.

Vamos a hacernos una idea de cómo calculaba Fermat la cuadratura de la hipérbola generalizada $y = x^{-2}$ para $x \geq a$. Usaremos notación y terminología actuales.

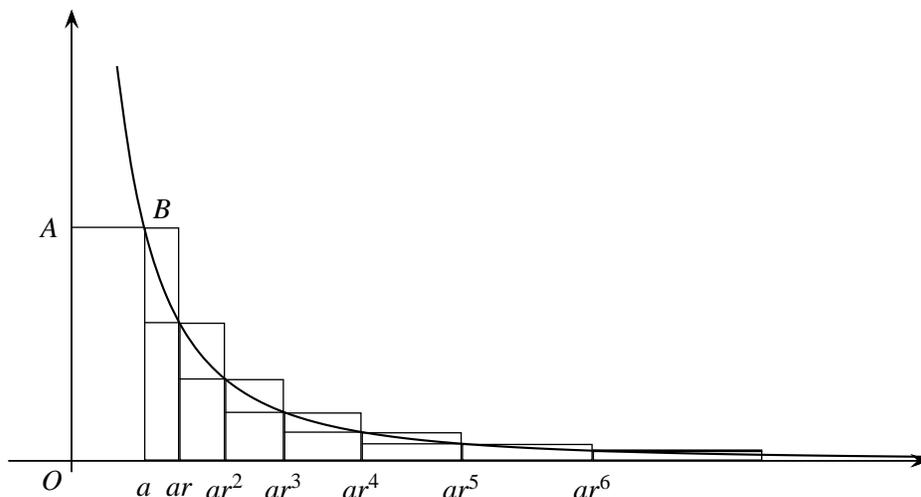


Figura 28. Cuadratura de la hipérbola de Fermat $y = x^{-2}$

Elegimos un número $r > 1$ y consideremos los puntos de abscisas a, ar, ar^2, ar^3, \dots . Los rectángulos inscritos (ver figura 28) tienen área

$$(ar - a) \frac{1}{(ar)^2} + (ar^2 - ar) \frac{1}{(ar^2)^2} + (ar^3 - ar^2) \frac{1}{(ar^3)^2} + \dots = \frac{r-1}{ar^2} \sum_{k=0}^{\infty} \frac{1}{r^k} = \frac{1}{ar}$$

El área de los rectángulos circunscritos viene dada por

$$(ar - a) \frac{1}{a^2} + (ar^2 - ar) \frac{1}{(ar)^2} + (ar^3 - ar^2) \frac{1}{(ar^2)^2} + \dots = \frac{r-1}{a} \sum_{k=0}^{\infty} \frac{1}{r^k} = \frac{r}{a}$$

Por tanto, llamando S al área bajo la curva, tenemos que

$$\frac{1}{ar} < S < \frac{r}{a}$$

Como esta desigualdad es válida para todo $r > 1$, concluimos que $S = \frac{1}{a}$. Observa que dicho valor es precisamente el área del rectángulo $OABa$.

El razonamiento de Fermat tiene detalles muy interesantes que se pierden usando la terminología y símbolos actuales. Vamos a reproducir parte de su razonamiento. Fermat se apoya en una propiedad de las progresiones geométricas de razón menor que la unidad, que enuncia como sigue:

Dada una progresión geométrica cuyos términos decrecen indefinidamente, la diferencia entre dos términos consecutivos es al más pequeño de ellos, como el mayor es a la suma de los términos restantes.

Llamemos R_1, R_2, R_3, \dots a las áreas de los sucesivos rectángulos y S a la suma de todas ellas. Como se trata de una progresión geométrica decreciente, se tiene que:

$$\frac{R_1 - R_2}{R_2} = \frac{R_1}{S - R_1}$$

Simplificando, resulta

$$S - R_1 = OA \cdot AB = \frac{1}{a}$$

Dice Fermat:

[...] *si ahora añadimos [a ambos miembros de esta igualdad] el rectángulo R_1 que a causa de las infinitas subdivisiones, se desvanece y queda reducido a nada, alcanzamos la conclusión, que podría ser fácilmente confirmada por una más prolija prueba llevada a cabo a la manera de Arquímedes... No es difícil extender esta idea a todas las hipérbolas definidas anteriormente excepto la que ha sido indicada [la hipérbola de Apolonio].*

Vemos cómo en las cuadraturas de Fermat de hipérbolas y parábolas generalizadas, subyacen los aspectos esenciales de la integral definida:

- La división del área bajo la curva en elementos de área infinitamente pequeños.
- Aproximación de la suma de esos elementos de área por medio de rectángulos infinitesimales de altura dada por la ecuación analítica de la curva.
- Un intento de expresar algo parecido a un límite de dicha suma cuando el número de elementos crece indefinidamente mientras se hacen infinitamente pequeños.

4.2.4. La integración aritmética de Wallis

Jhon Wallis (1616 - 1703) publicó en 1655 un tratado *Arithmetica infinitorum* (“La Aritmética de los infinitos”) en el que aritmetizaba el método de los indivisibles de Cavalieri. Para ilustrar el método de Wallis consideremos el problema de calcular el área bajo la curva $y = x^k$ ($k = 1, 2, \dots$) y sobre el segmento $[0, a]$ (ver figura (29)). Siguiendo a Cavalieri, Wallis considera la región PQR formada por un número infinito de líneas verticales paralelas, cada una de ellas con longitud igual a x^k . Por tanto, si dividimos el segmento $PQ = AB = a$ en n partes de longitud $h = a/n$, donde n es infinito, entonces la suma de estas infinitas líneas es del tipo

$$0^k + h^k + (2h)^k + (3h)^k + \dots + (nh)^k \tag{10}$$

Análogamente, el área del rectángulo $ABCD$ es

$$a^k + a^k + a^k + \dots + a^k = (nh)^k + (nh)^k + (nh)^k + \dots + (nh)^k \quad (11)$$

La razón entre el área de la región PQR y el rectángulo $ABCD$ es

$$\frac{\text{Área } PQR}{\text{Área } ABCD} = \frac{0^k + 1^k + 2^k + 3^k + \dots + n^k}{n^k + n^k + n^k + n^k + \dots + n^k} \quad (12)$$

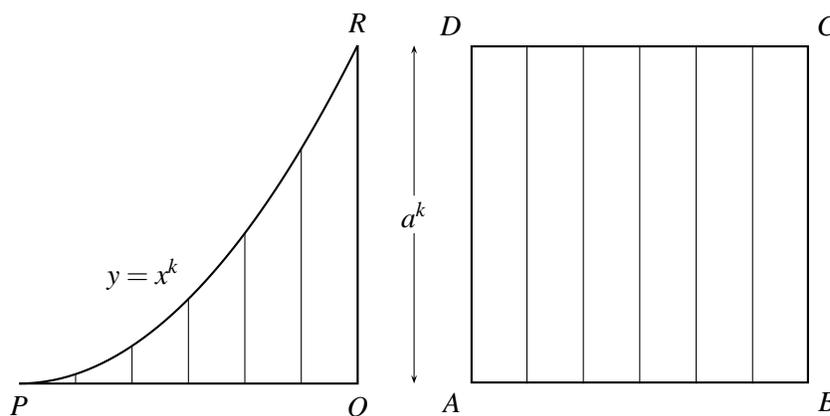


Figura 29. Comparando indivisibles

Esto lleva a Wallis a estudiar el valor de la expresión (12) para $n = \infty$ ⁷. Después de estudiar varios casos para valores de $k = 1, 2, 3$ haciendo, en cada caso, sumas para distintos valores de $n = 1, 2, 3, 4$, Wallis observa ciertas regularidades en las mismas y, con tan débil base, acaba afirmando que para $n = \infty$ y para todo $k = 1, 2, \dots$, se verifica que:

$$\frac{0^k + 1^k + 2^k + 3^k + \dots + n^k}{n^k + n^k + n^k + n^k + \dots + n^k} = \frac{1}{k+1} \quad (13)$$

Naturalmente, de aquí deduce el valor del área de la región PQR :

$$\frac{\text{Área } PQR}{\text{Área } ABCD} = \frac{\text{Área } PQR}{a^{k+1}} = \frac{1}{k+1} \Rightarrow \text{Área } PQR = \frac{a^{k+1}}{k+1} \quad k = 1, 2, 3, \dots \quad (14)$$

Este resultado ya era conocido anteriormente, pero Wallis no se paraba aquí y extendía la validez de la igualdad (13) a todos los exponentes racionales positivos. Su peculiar razonamiento tiene interés pues en él se basó Newton para obtener la serie binomial. Lo esencial del mismo puede resumirse, en términos actuales, como sigue.

⁷Fue precisamente Wallis quien introdujo en 1655 en la obra *De Sectionibus Conicis*, el símbolo del “lazo del amor”, ∞ , con el significado de “infinito”.

Definamos el índice, $\sigma(f)$, de una función f mediante la igualdad

$$\lim_{n \rightarrow \infty} \frac{f(0) + f(1) + f(2) + \dots + f(n)}{f(n) + f(n) + f(n) + \dots + f(n)} = \frac{1}{\sigma(f) + 1} \quad (15)$$

suponiendo que dicho límite tenga sentido. Por ejemplo, (13) nos dice que el índice de la función $f_k(x) = x^k$ es $\sigma(f_k) = k$ para $k = 1, 2, \dots$

Wallis observó que, dada una progresión geométrica de potencias de x como, por ejemplo $1, x^3, x^5, x^7, \dots$, la correspondiente sucesión de índices $0, 3, 5, 7, \dots$ forman una progresión aritmética. Como $\sigma(f_k) = k$, esta observación es trivial, pero le permite dar un atrevido salto adelante, de manera que mediante una audaz interpolación establece (sin demostración) que una conclusión análoga puede deducirse para la progresión geométrica

$$1, \sqrt[q]{x}, (\sqrt[q]{x})^2, \dots, (\sqrt[q]{x})^{q-1}, x$$

de manera que la sucesión de sus índices debe formar una progresión aritmética, de donde se sigue que debe ser $\sigma((\sqrt[q]{x})^p) = p/q$ para $p = 1, 2, \dots, q$. De esta forma obtiene que

$$\lim_{n \rightarrow \infty} \frac{(\sqrt[0]{n})^p + (\sqrt[1]{n})^p + (\sqrt[2]{n})^p + (\sqrt[3]{n})^p + \dots + (\sqrt[n]{n})^p}{(\sqrt[n]{n})^p + (\sqrt[n]{n})^p + (\sqrt[n]{n})^p + (\sqrt[n]{n})^p + \dots + (\sqrt[n]{n})^p} = \frac{1}{p/q + 1}$$

Wallis estaba convencido de la validez de su método, conocido posteriormente como *interpolación de Wallis*, que tuvo importancia en el siglo XVIII. Puede considerarse como un intento de resolver el siguiente problema:

Dada una sucesión P_k , definida para valores enteros de k , encontrar el significado de P_α cuando α no es un número entero.

Además, Wallis deduce que *necesariamente debe ser* $(\sqrt[q]{x})^p = x^{p/q}$. Será Newton, poco más tarde, quien siguiendo los pasos de Wallis, introducirá el uso de potencias fraccionarias y negativas.

Wallis, incluso llega a afirmar que la igualdad

$$\int_0^a x^r dx = \frac{a^{r+1}}{r+1} \quad (16)$$

no es válida solamente para exponentes r racionales, sino también para otros como $r = \sqrt{3}$ pero, naturalmente, no puede dar ninguna justificación.

Obtenida, a su manera, la cuadratura fundamental (16), Wallis intenta calcular la integral

$$\int_0^1 \sqrt{x-x^2} dx$$

Dicha integral representa el área bajo la semicircunferencia de centro $(1/2, 0)$ y radio $1/2$, su valor es, por tanto, $\pi/8$. Wallis quería obtener dicho resultado evaluando directamente la integral. No tuvo éxito en este empeño que Newton habría de resolver posteriormente, pero sus resultados le llevaron a obtener la llamada *fórmula de Wallis*

$$\frac{2}{\pi} = \frac{1 \cdot 3 \cdot 3 \cdot 5 \cdot 5 \cdot 7 \cdot 7 \cdots}{2 \cdot 2 \cdot 4 \cdot 4 \cdot 6 \cdot 6 \cdot 8 \cdots}$$

4.2.5. El resultado fundamental de Barrow

Barrow estuvo muy cerca de descubrir la relación inversa entre problemas de tangentes y de cuadraturas, pero su conservadora adhesión a los métodos geométricos le impidió hacer uso efectivo de esta relación. Veamos cómo aparece esa relación tal como se expone en la Lección X, Proposición 11 de las *Lectiones Geometricae*.

En la figura (30) se han representado dos curvas $y = f(x)$ e $y = g(x)$. El segmento AD representa el eje de abscisas donde toma valores x . La cantidad $g(x)$ representa el valor del área bajo la gráfica de f comprendida entre el punto A y x . Dado un punto de abscisa D , se trata de probar que la pendiente de la tangente a $y = g(x)$ en el punto F , es decir en el punto $(D, g(D))$, es igual a $f(D) = DE$. La demostración de Barrow es geométrica.

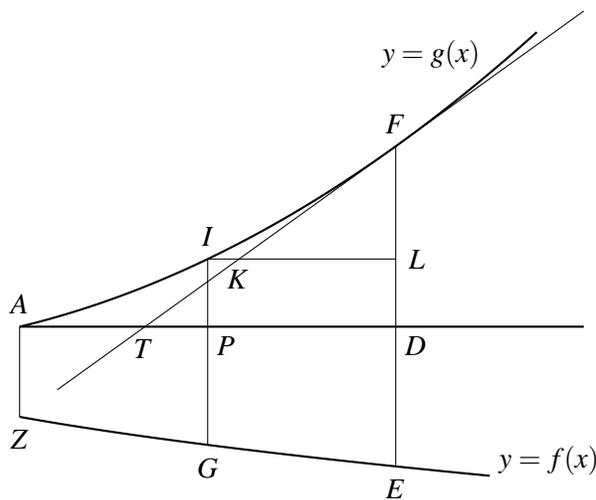


Figura 30. Teorema Fundamental

Tracemos una línea recta FT por F que corta en T a la recta AD y tal que

$$DF/TD = f(D) = DE$$

Queremos probar que FT es la tangente a $y = g(x)$ en el punto F . Para ello vamos a ver que la distancia horizontal, KL , de cualquier punto L de la recta EF a la recta FT es menor que la

distancia, IL , de dicho punto L a la curva $y = g(x)$. Esto probará que la recta FT queda siempre por debajo de $y = g(x)$.

Tenemos que:

$$FL/KL = DF/TD = DE$$

Por otra parte:

$$\text{área } ADEZ = FD$$

$$\text{área } APGZ = PI = LD$$

$$\text{área } PDEG = FD - LD = FL$$

Ya que

$$\text{área } PDEG < \text{rectángulo } PD.DE \quad (17)$$

Se sigue que

$$FL < PD.DE \longrightarrow DE > FL/PD$$

y por tanto

$$FL/KL > FL/PD \longrightarrow KL < PD = IL$$

Deducimos que el punto K queda debajo de la curva $y = g(x)$ y por tanto la recta FT queda a un lado de la curva. Para completar la demostración es necesario repetir el razonamiento tomando puntos a la derecha de EF . Esto prueba que TF es tangente a $y = g(x)$ en D y su pendiente es $DE = f(D)$. En términos actuales, lo que Barrow ha probado es que:

$$\frac{d}{dx} \int_a^x f(t)dt = f(x)$$

4.3. La relación fundamental entre cuadraturas y tangentes

4.3.1. El Teorema Fundamental del Cálculo según Newton

Newton desarrolló tres versiones de su cálculo. En la obra *De Analysi per aequationes numero terminorum infinitas*, que Newton entregó a su maestro Barrow en 1669, y que puede considerarse el escrito fundacional del Cálculo, Newton usa conceptos infinitesimales de manera similar a como hacía el propio Barrow. Este trabajo, además de contener el teorema binomial y los descubrimientos de Newton relativos a series infinitas, contiene también un claro reconocimiento de la relación inversa entre problemas de cuadraturas y de tangentes. La exposición que hace Newton de esta relación fundamental es como sigue. Supone una curva y llama z al área bajo la curva hasta el punto de abscisa x (ver figura 31). Se supone conocida la relación entre x y z . Aunque Newton

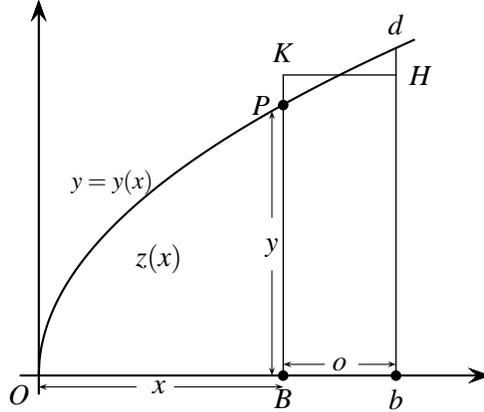


Figura 31. $z = z(x) = \text{área } OPB$

explica su método con un ejemplo, queda perfectamente claro su carácter general. El ejemplo que Newton considera es

$$z = \frac{n}{m+n} ax^{\frac{m+n}{n}} \quad (18)$$

Pongamos, por comodidad $r = \frac{m+n}{n}$. Newton se imagina que el punto $P = (x, y)$ se mueve a lo largo de la curva y razona como sigue. Incrementemos la abscisa x a $x + o$ donde o es una cantidad infinitesimal o *momento*. Tomemos $BK = v$ de forma que $ov = \text{área } BbHK = \text{área } BbPd$. El incremento del área viene dado por:

$$ov = z(x+o) - z(x) = \frac{a}{r}(x+o)^r - \frac{a}{r}x^r \quad (19)$$

Desarrollando en potencias

$$\frac{a}{r}(x+o)^r = \frac{a}{r}x^r(1+o/x)^r = \frac{a}{r}x^r \left(1 + r\frac{o}{x} + \frac{r(r-1)}{2}\frac{o^2}{x^2} + \frac{r(r-1)(r-2)}{1 \cdot 2 \cdot 3}\frac{o^3}{x^3} + \dots \right) \quad (20)$$

De (19) y (20) deducimos, después de dividir por o , que:

$$v = ax^{r-1} + \frac{a(r-1)}{2}ox^{r-2} + \frac{a(r-1)(r-2)}{1 \cdot 2 \cdot 3}o^2x^{r-3} + \dots$$

Si en esta igualdad suponemos que o va disminuyendo hasta llegar a ser nada, en cuyo caso v coincidirá con y , después de eliminar los términos que contienen o que desaparecen, resulta que:

$$y = ax^{r-1} = ax^{\frac{m}{n}} \quad (21)$$

Este es, por tanto, el valor de la ordenada de la curva en $P = (x, y)$. El proceso puede invertirse y, de hecho, ya se sabía que la cuadratura de (21) viene dada por (18).

Observemos que Newton no ha usado el significado tradicional de la integral al estilo de sus predecesores, es decir, no ha interpretado la integral como un límite de sumas de áreas infinitesimales, sino que ha probado que la expresión que proporciona la cuadratura es correcta estudiando

la variación momentánea de dicha expresión. De hecho, lo que Newton ha probado es que la razón de cambio del área bajo la curva, esto es, el cociente

$$\frac{z(x+o) - z(x)}{o}$$

se hace igual a la ordenada de la curva cuando o “se hace nada”. En términos actuales, la derivada de $z(x)$ es la función $y = y(x)$. La relación simétrica entre cuadraturas y derivadas queda así puesta claramente de manifiesto. Para calcular cuadraturas, basta con calcular una antiderivada, lo que llamamos una primitiva de la función $y = y(x)$.

4.3.2. La invención del *calculus summatorius* por Leibniz

Ya hemos comentado en el capítulo 6 (ver pg. 66) las principales ideas que guiaron a Leibniz en la invención del Cálculo:

- La creación de un simbolismo matemático que automatizara los cálculos y permitiera formular fácilmente procesos algorítmicos.
- La apreciación de que las sucesiones de diferencias pueden sumarse fácilmente, y que el proceso de formar la sucesión de diferencias y después sumarla recupera la sucesión inicial, es decir, que se trata de operaciones inversas una de la otra.
- La consideración de las curvas como polígonos de infinitos lados de longitudes infinitesimales y de las variables como sucesiones que toman valores consecutivos infinitamente próximos.

Se conservan en el archivo Leibniz en Hannover los manuscritos que contienen las investigaciones de Leibniz sobre los problemas de cuadraturas. En dichos documentos, fechados del 25 de octubre al 11 de noviembre de 1675, Leibniz investiga la posibilidad de formular simbólicamente los problemas de cuadraturas e introduce los símbolos que actualmente usamos para la integral y la diferencial. Los progresos de Leibniz se exponen de forma concisa y clara en el trabajo de H.J.M. Bos [6] que sigo muy de cerca. Algunos de los resultados de Leibniz en estos manuscritos son casos particulares de la regla de integración por partes, como, por ejemplo, la siguiente igualdad (se supone $f(0) = 0$):

$$\int_0^a x f'(x) dx = a f(a) - \int_0^a f(x) dx = a \int_0^a f'(x) dx - \int_0^a \left(\int_0^x f'(t) dt \right) dx \quad (22)$$

Por supuesto, Leibniz no la escribe así. Recuerda que la notación que usamos para la derivada se debe a J.L. Lagrange y es bastante tardía, de finales del siglo XVIII. Además, la notación que usamos para indicar los límites de integración fue introducida por J. Fourier en el primer tercio del

siglo XIX. Incluso el término “integral” no se debe a Newton ni a Leibniz. Leibniz llamó *calculus differentialis*, esto es “cálculo de diferencias”, a la parte de su cálculo que se ocupa del estudio de tangentes, y *calculus summatorius*, o sea “cálculo de sumas”, a la que se ocupa de problemas de cuadraturas. Para Leibniz una integral es una suma de infinitos rectángulos infinitesimales, el símbolo que ideó para representarlas, “ \int ” tiene forma de una “s” alargada como las que en aquel tiempo se usaban en la imprenta; además, es la primera letra de la palabra latina *summa*, o sea, “suma”. Fue Johann Bernoulli quien, en 1690, sugirió llamar *calculus integralis* al cálculo de cuadraturas, de donde deriva el término “integral” que usamos actualmente.

De hecho, Leibniz obtuvo la fórmula (22) antes de inventar su notación para las integrales y las diferenciales. Es interesante mostrar cómo lo hizo. Para ello vamos a seguir el camino opuesto al seguido por Leibniz, modificando la notación de dicha fórmula hasta llegar a escribirla como lo hizo él.

Podemos interpretar gráficamente la igualdad (22) sin más que observar la figura 32.

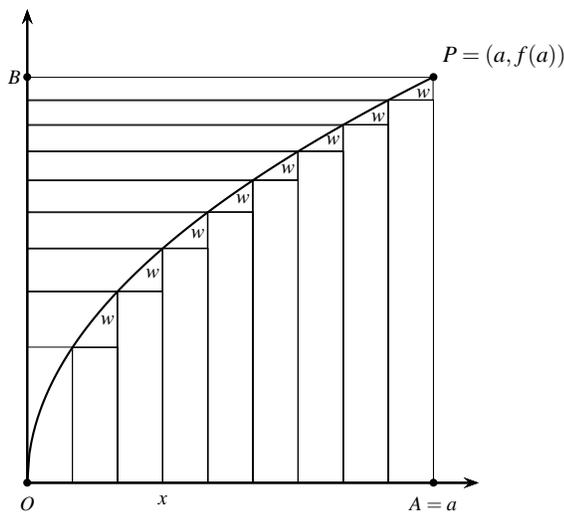


Figura 32. Áreas complementarias

El número $af(a)$ es el área del rectángulo $OAPB$, la integral $\int_0^a f(x) dx$ es el área de la parte de dicho rectángulo OAP que queda bajo la curva $y = f(x)$. Deducimos de (22) que la integral $\int_0^a xf'(x) dx$ es el área de la parte OBP de dicho rectángulo que queda por encima de la curva $y = f(x)$. Esta área es la suma de las áreas de rectángulos horizontales como los representados en la figura 32. Estos rectángulos horizontales tienen como base el valor de la abscisa correspondiente, x , y como altura la diferencia infinitamente pequeña entre dos ordenadas sucesivas, que Leibniz representa por w . Esta diferencia es lo que posteriormente se llamará diferencial de y . Podemos, pues, interpretar que $w = dy = f'(x) dx$. Por su parte, el área de la región OAP es considerada por Leibniz como la suma de las ordenadas y . Finalmente, podemos eliminar y porque para Leibniz

el valor de una variable puede obtenerse sumando sus diferencias consecutivas, por eso, y puede verse como la suma de las w . Esto equivale, en nuestra notación, a sustituir $f(x)$ por $\int_0^x f'(t) dt$ (o, al estilo de Leibniz, y por $\int dy$), lo que también hemos hecho en la igualdad (22). La forma exacta en que Leibniz escribió la igualdad 22, según se lee en [6], es:

$$\text{omn. } \overline{xw} \sqcap \text{ult. } x, \overline{\text{omn. } w}, - \overline{\text{omn. } \text{omn. } w} \quad (23)$$

Aquí \sqcap es el símbolo para la igualdad, “ult. x ” significa el *ultimus* x , el último de los x , es decir, $OA = a$. El símbolo “omn.” es la abreviatura de *omnes lineae*, “todas las líneas”, símbolo que había sido usado por Cavalieri y que Leibniz usa con el significado de “una suma”. Se usan también líneas por encima de los términos y comas donde ahora pondríamos paréntesis.

En un manuscrito posterior en algunos días, Leibniz vuelve a escribir la igualdad 23 en la forma:

$$\text{omn. } x\ell \sqcap x \text{omn. } \ell - \text{omn. } \text{omn. } \ell, \quad (24)$$

y observa que *omn.* antepuesto a una magnitud lineal como ℓ da un área; *omn.* antepuesto a un área como $x\ell$ da un volumen y así sucesivamente.

[6]. . . Estas consideraciones de homogeneidad dimensional parecen haber sido las que sugirieron a Leibniz el usar una única letra en vez del símbolo “omn.”, porque escribe a continuación: “Sería conveniente escribir “ \int ” en lugar de “omn.”, de tal manera que $\int \ell$ represente *omn.* ℓ , es decir, la suma de todas las ℓ ”. Así fue como se introdujo el signo “ \int ” [. . .] E inmediatamente a continuación escribe Leibniz la fórmula (24) utilizando el nuevo formalismo:

$$\int x\ell = x \int \ell - \int \int \ell \quad (25)$$

haciendo notar que:

$$\int x = \frac{x^2}{2} \quad \text{y} \quad \int x^2 = \frac{x^3}{3}$$

y subrayando que estas reglas se aplican a “las series en las que la razón de las diferencias de los términos a los términos mismos es menor que cualquier cantidad dada”, es decir, a las series cuyas diferencias son infinitamente pequeñas.

Una líneas más adelante nos encontramos también con la introducción del símbolo “ d ” para la diferenciación. Aparece en el contexto de un brillante razonamiento que puede resumirse de la forma siguiente: el problema de las cuadraturas es un problema de suma de sucesiones, para lo cual hemos introducido el símbolo “ \int ” y para el que queremos elaborar un *cálculo*, es decir, un conjunto de algoritmos eficaces. Ahora bien, sumar sucesiones, es decir hallar una expresión general para $\int y$ dada la y , no es posible normalmente, pero siempre lo es encontrar una expresión para las diferencias de una sucesión dada. Así pues, el cálculo de diferencias es la operación recíproca del cálculo de sumas, y por lo tanto podemos esperar dominar el cálculo de sumas desarrollando su recíproco, el cálculo de diferencias. Para citar las mismas palabras de Leibniz:

Dada ℓ y su relación con x , hallar $\int \ell$. Esto se puede obtener mediante el cálculo inverso, es decir, supongamos que $\int \ell = ya$ y sea $\ell = ya/d$; entonces de la misma manera que la \int aumenta las dimensiones, d las disminuirá. Pero la \int representa una suma y d una diferencia, y de la y dada podemos encontrar siempre y/d o ℓ , es decir, la diferencia de las y .

Así se introduce el símbolo “ d ” (o más bien el símbolo “ $1/d$ ”). [...] De hecho, pronto se da cuenta de que ésta es una desventaja notacional que no viene compensada por la ventaja de la interpretación dimensional de la \int y de d , y pasa a escribir “ $d(ya)$ ” en vez de “ ya/d ”, y de ahí en adelante son interpretadas la d y la \int como símbolos adimensionales [...].

En el resto del manuscrito Leibniz se dedica a explorar este nuevo simbolismo, al que traduce viejos resultados, y a investigar las reglas operacionales que rigen la \int y la d .

Esta larga cita, extraída del trabajo de H.J.M. Bos *Newton, Leibniz y la tradición leibniziana* ([6]), nos da una idea de cómo llegó Leibniz a la invención del cálculo. No fueron los caminos del razonamiento lógico deductivo los seguidos por Leibniz sino los de la intuición, la conjetura, el estudio de casos particulares y su generalización... Los mismos caminos que hoy siguen los matemáticos activos en sus trabajos de investigación. Pese a que los conceptos que maneja Leibniz son oscuros e imprecisos fue capaz de desarrollar algoritmos de cálculo eficaces y de gran poder heurístico. Como ya hemos indicado en el capítulo 6, el cálculo de Leibniz triunfó en el continente europeo gracias a los trabajos de los hermanos Bernoulli y al libro de texto del Marqués de L'Hôpital que divulgó las técnicas del cálculo leibniziano por toda Europa.

5. El Cálculo y las series

Las sumas de progresiones geométricas, con un número finito de sumandos, ya aparecen en la matemática griega; hemos visto que Arquímedes las usa en su cuadratura de un segmento de parábola. En la Edad Media tardía, siglo XIV, aparecen esporádicamente algunas series para calcular la distancia recorrida por cuerpos móviles cuando la velocidad cambia de un período temporal a otro. Así Richard Swineshead, apodado *el caculador de Oxford*, publicó hacia 1350 un tratado, *Liber Calculationum*, en el que, con un largo y tedioso razonamiento puramente verbal, logra calcular la suma de la primera serie no geométrica:

$$\frac{1}{2} + \frac{2}{4} + \frac{3}{8} + \dots + \frac{n}{2^n} + \dots = 2 \quad (26)$$

cuya suma representa la velocidad media de un móvil que en un intervalo de tiempo unidad y con velocidad inicial unidad va aumentando su velocidad de unidad en unidad cada vez que transcurre la mitad del tiempo restante.

Nicolás de Oresme en su *Tractatus de configurationibus qualitatum et motuum*, publicado en 1350, prueba que si se van haciendo sucesivamente las sumas $\{1 + 1/2 + \dots + 1/n\}$, entonces, afirma, “la totalidad llegará a ser infinito”. Su razonamiento es el mismo que ahora se hace para probar la divergencia de la serie armónica, a saber, Oresme razona agrupando términos de forma conveniente:

$$1 + \frac{1}{2} + \frac{1}{3} + \frac{1}{4} + \frac{1}{5} + \frac{1}{6} + \frac{1}{7} + \frac{1}{8} + \dots$$

$$> 1 + \frac{1}{2} + \left(\frac{1}{4} + \frac{1}{4}\right) + \left(\frac{1}{8} + \frac{1}{8} + \frac{1}{8} + \frac{1}{8}\right) + \dots = 1 + \frac{1}{2} + \frac{1}{2} + \frac{1}{2} + \dots$$

Así mismo, en la citada obra, Oresme probó que si k es un entero mayor que 1 y a es una cierta cantidad:

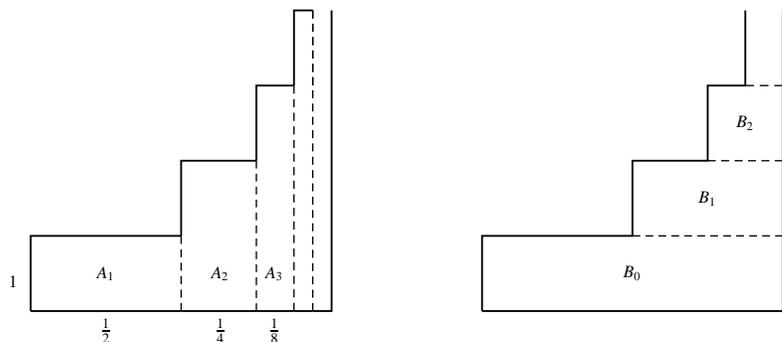
$$\frac{a}{k} + \frac{a}{k} \left(1 - \frac{1}{k}\right) + \frac{a}{k} \left(1 - \frac{1}{k}\right)^2 + \frac{a}{k} \left(1 - \frac{1}{k}\right)^3 + \dots + \frac{a}{k} \left(1 - \frac{1}{k}\right)^n + \dots = a$$

Oresme razona esta igualdad verbalmente como sigue [14]:

Si una parte proporcional (la k -ésima parte) se quita de alguna cantidad a , y si del resto que queda vuelve a quitarse la misma parte proporcional y así sucesivamente, finalmente tal cantidad será consumida enteramente de forma exacta – no más, no menos – por esta forma de sustracción.

Observa que haciendo $k = a = 4/3$ se obtiene la serie geométrica de razón $1/4$ usada por Arquímedes.

También proporciona Oresme la siguiente estrategia geométrica para calcular la suma de la serie (26).



Puesto que $\text{área}(A_n) = \frac{n}{2^n}$ y $\text{área}(B_n) = \frac{1}{2^{n-1}}$ se deduce que

$$\frac{1}{2} + \frac{2}{4} + \frac{3}{8} + \dots + \frac{n}{2^n} + \dots =$$

$$= \sum_{n=1}^{\infty} \text{área}(A_n) = \sum_{n=1}^{\infty} \text{área}(B_n) = \sum_{n=1}^{\infty} \frac{1}{2^{n-1}} = 2$$

Oresmes no pretendía usar las series para realizar cálculos sino para estudiar las paradojas relacionadas con el hecho de que un número infinito de cantidades finitas pudiera tener una suma finita, cuestiones que están relacionadas con la estructura del continuo y su infinita divisibilidad y con las aporías de Zenón de Elea.

Las progresiones geométricas establecen un puente entre lo discreto y lo continuo. Para Gregory de St. Vincent (1584 - 1667) una serie es un *continuo dividido*; en su *Opus Geometricum* (1647) escribe: “Llamo serie geométrica a una cantidad finita dividida en sucesión ininterrumpida según una razón dada cualquiera” [13]. Fue el primero en afirmar explícitamente que una serie infinita puede representar una magnitud. También le debemos el primer análisis de las paradojas de Zenón usando series. Descubrió que la cuadratura de la hipérbola $xy = k$ es la misma en $[a, b]$ que en $[c, d]$ cuando $a/b = c/d$, resultado fundamental para la comprensión de los logaritmos y que llevó al descubrimiento del logaritmo natural por Mercator.

5.1. Los primeros desarrollos en serie

En 1668, Nicholas Mercator (1620 - 1687) publicó un libro titulado *Logarithmotechnia* en el que proporcionaba un método para calcular logaritmos basado en el desarrollo en serie del logaritmo natural

$$\log(1+x) = x - \frac{x^2}{2} + \frac{x^3}{3} - \frac{x^4}{4} + \dots \quad (27)$$

el cual obtuvo usando los resultados de Gregory de St. Vincent.

A su vez, este resultado de Mercator fue mejorado por James Gregory (1638 - 1675) que obtuvo la expansión:

$$\log \frac{1+x}{1-x} = 2x + \frac{2x^3}{3} + \frac{2x^5}{5} + \dots$$

que converge más rápidamente que la anterior. A James Gregory se debe también la serie del arco-tangente:

$$\arctan x = x - \frac{x^3}{3} + \frac{x^5}{5} - \frac{x^7}{7} + \dots \quad (28)$$

Sustituyendo $x = 1$ resulta

$$\frac{\pi}{4} = 1 - \frac{1}{3} + \frac{1}{5} - \frac{1}{7} + \dots$$

Mejores representaciones de π se deducen de esta serie haciendo como A. Sahrp (1651 - 1742) $x = 1/\sqrt{3}$, con lo que

$$\frac{\pi}{6} = \frac{1}{\sqrt{3}} \left(1 - \frac{1}{3 \cdot 3} + \frac{1}{3^2 \cdot 5} - \frac{1}{3^3 \cdot 7} + \dots \right)$$

Con cuya serie calculó π con 72 cifras decimales en 1705. Una mejor aproximación de π que evita el uso de radicales y converge rápidamente, fue obtenida en 1706 por John Machin (1680 - 1752).

La idea es expresar $\pi/4 = \arctan 1$ en función de dos ángulos de tangentes racionales y cada una de ellas menor que la unidad. La serie de Machin es:

$$\frac{\pi}{4} = 4 \arctan \frac{1}{5} - \arctan \frac{1}{239} = 4 \left(\frac{1}{5} - \frac{1}{3 \cdot 5^3} + \frac{1}{5 \cdot 5^5} - \dots \right) - \left(\frac{1}{239} - \frac{1}{3 \cdot 239^3} + \frac{1}{5 \cdot 239^5} - \dots \right)$$

Con ella calculó π con 100 cifras decimales.

5.2. Newton y las series infinitas

Los principales descubrimientos matemáticos de Newton en el campo del cálculo infinitesimal datan de los llamados *Anni Mirabiles* 1665 y 1666. La Universidad de Cambridge, en la que Newton se había graduado como *bachelor of arts* en 1664, estuvo cerrada por la peste esos dos años. Newton pasó ese tiempo en su casa de Woolsthorpe y, como él mismo reconoció cincuenta años después, ése fue el período más creativo de su vida.

A principios de 1665 descubre el teorema del binomio y el cálculo con las series infinitas. A finales de ese mismo año, el método de fluxiones, es decir, el cálculo de derivadas. En 1666 el método inverso de fluxiones y la relación entre cuadraturas y fluxiones. En esos dos años también inició las teorías de los colores y de la gravitación universal. Newton tenía 24 años, había nacido el día de Navidad de 1642.

Newton había leído la obra de Wallis *Arithmetica Infinitorum*, y siguiendo las ideas de interpolación allí expuestas, descubrió la serie del binomio que hoy lleva su nombre. Dicha serie es una generalización del desarrollo del binomio, que era bien conocido para exponentes naturales, y había sido muy usado por Pascal para resolver una gran variedad de problemas.

Newton, en su intento de calcular la cuadratura del círculo, es decir, de calcular la integral $\int_0^1 (1-x^2)^{1/2} dx$, consideró dicha cuadratura como un problema de interpolación, relacionándola con las cuadraturas análogas $\int_0^1 (1-x^2)^n dx$ conocidas para exponentes naturales $n \in \mathbb{N}$. Newton tuvo la ocurrencia de sustituir el límite superior de integración por un valor genérico x . De esta forma obtuvo las siguientes cuadraturas (Newton no disponía de símbolo para la integral; usamos, claro está, la notación actual).

$$\begin{aligned} \int_0^x (1-t^2) dt &= x - \frac{1}{3}x^3 \\ \int_0^x (1-t^2)^2 dt &= x - \frac{2}{3}x^3 + \frac{1}{5}x^5 \\ \int_0^x (1-t^2)^3 dt &= x - \frac{3}{3}x^3 + \frac{3}{5}x^5 - \frac{1}{7}x^7 \\ \int_0^x (1-t^2)^4 dt &= x - \frac{4}{3}x^3 + \frac{6}{5}x^5 - \frac{4}{7}x^7 + \frac{1}{9}x^9 \end{aligned}$$

Newton observó que el primer término de cada expresión es x , que x aumenta en potencias impares, que los signos algebraicos se van alternando, y que los segundos términos $\frac{1}{3}x^3, \frac{2}{3}x^3, \frac{3}{3}x^3, \frac{4}{3}x^3$ estaban en progresión aritmética. Razonando por analogía, supuso que los dos primeros términos de $\int_0^x (1-t^2)^{1/2} dt$ deberían ser

$$x - \frac{1}{2}x^3$$

De la misma manera, procediendo por analogía, pudo encontrar algunos términos más:

$$\int_0^x (1-t^2)^{1/2} dt = x - \frac{1}{2}x^3 - \frac{1}{8}x^5 - \frac{1}{16}x^7 - \frac{1}{128}x^9 - \dots$$

Representando para $n = 0, 1, 2, \dots$ por $Q_n(x)$ el polinomio $\int_0^x (1-t^2)^n dt$, se tiene que

$$Q_n(x) = \int_0^x (1-t^2)^n dt = \sum_{k=0}^n \binom{n}{k} \frac{(-1)^k}{2k+1} x^{2k+1}$$

Donde

$$\binom{n}{k} = \frac{n(n-1)(n-2)\cdots(n-k+1)}{1 \cdot 2 \cdot 3 \cdots k}, \quad \binom{n}{0} = 1$$

Haciendo ahora en $Q_n(x)$, $n = 1/2$, se obtiene

$$Q_{1/2}(x) = x - \frac{1}{2}x^3 - \frac{1}{8}x^5 - \frac{1}{16}x^7 - \frac{1}{128}x^9 - \dots$$

Lo que llevó a Newton a concluir que

$$\int_0^x (1-t^2)^{1/2} dt = Q_{1/2}(x)$$

Donde $Q_{1/2}(x) = \sum_{n=0}^{\infty} \binom{1/2}{n} \frac{(-1)^n}{2n+1} x^{2n+1}$ es una suma con infinitos términos. A partir de aquí, Newton dedujo el desarrollo de $(1-x^2)^{1/2}$ por derivación.

$$(1-x^2)^{1/2} = 1 - \frac{1}{2}x^2 - \frac{1}{8}x^4 - \frac{1}{16}x^6 - \frac{1}{128}x^8 - \dots$$

Newton nunca publicó su teorema binomial, ni dio una demostración general del mismo. La primera vez que apareció en un texto impreso fue en 1685 en un libro de Wallis (que reconoce la autoría de Newton), titulado *Treatise of Algebra*. Newton mismo, en una carta a Henry Oldenburg, el secretario de la Royal Society, conocida como la *Epistola Prior* (junio de 1676), expone el teorema binomial, a requerimiento de Leibniz, con estas oscuras palabras:

Las extracciones de raíces resultan muy abreviadas por el teorema

$$(P + PQ)^{m/n} = P^{m/n} + \frac{m}{n}AQ + \frac{m-n}{2n}BQ + \frac{m-2n}{3n}CQ + \frac{m-3n}{4n}DQ + \text{etc}$$

donde $P + PQ$ representa una cantidad cuya raíz o potencia, o cuya raíz de una potencia se necesita calcular, siendo P el primer término de esa cantidad, Q los términos restantes divididos por el primero, y $\frac{m}{n}$ el índice numérico de las potencias de $P + PQ$. Por último $A = P^{m/n}$, $B = \frac{m}{n}AQ$, $C = \frac{m-n}{2n}BQ$ y así sucesivamente.

Newton era consciente de que su forma de razonar por analogía no era rigurosa por lo que comprobó su resultado de varias formas. Aplicó su algoritmo a diversos resultados conocidos, comprobando que las soluciones obtenidas eran siempre correctas, redescubrió la serie de Mercator para el logaritmo y obtuvo las series del arcoseno y del seno. Por cierto, que es en este contexto de la serie binomial cuando Newton usa por primera vez exponentes fraccionarios o negativos para realizar cálculos rutinarios.

Newton encontró que el método de desarrollos en serie proporcionaba un algoritmo casi universal para calcular cuadraturas y resolver multitud de problemas. En su obra *De analysi per aequationes numero terminorum infinitas*, escrita en 1669 y publicada en 1711, aunque circulaba en forma manuscrita entre los colegas y conocidos de Newton, propuso un método para cuadrar una curva consistente en tres reglas:

1. El área bajo la curva de ecuación $y = ax^{m/n}$ es $\frac{na}{m+n}ax^{\frac{m+n}{n}}$.
2. Si la ecuación $y = y(x)$ de la curva está dada por un número finito de términos $y_1 + y_2 + y_3 + \dots$, el área bajo la curva y es igual a la suma de las áreas de todos los términos y_1, y_2, y_3, \dots .
3. Si la curva tiene una forma más complicada, entonces debe desarrollarse la ecuación de la curva en una serie del tipo $\sum a_k x^{r_k}$, donde r_k es un número racional, y aplicar las reglas 1 y 2.

Debe notarse que Newton supuso que cualquier cantidad analíticamente expresada podía desarrollarse en una serie de la forma $\sum a_k x^{r_k}$, donde r_k es un número racional, serie que puede ser cuadrada término a término usando la regla 1.

Veamos un ejemplo de esta forma de proceder. Se trata de calcular $\int_0^{1/4} \sqrt{x-x^2} dx$. Newton procede como sigue

$$(x-x^2)^{1/2} = x^{1/2}(1-x)^{1/2} = x^{1/2} - \frac{1}{2}x^{3/2} - \frac{1}{8}x^{5/2} - \frac{1}{16}x^{7/2} - \frac{1}{128}x^{9/2} - \dots$$

Por tanto

$$\begin{aligned} \int_0^{1/4} (x-x^2)^{1/2} dx &= \left[\frac{2}{3}x^{3/2} - \frac{1}{5}x^{5/2} - \frac{1}{28}x^{7/2} - \frac{1}{72}x^{9/2} - \frac{5}{704}x^{11/2} - \dots \right]_0^{1/4} \\ &= \frac{2}{3 \cdot 2^3} - \frac{1}{5 \cdot 2^5} - \frac{1}{28 \cdot 2^7} - \frac{1}{72 \cdot 2^9} - \frac{5}{704 \cdot 2^{11}} - \dots \end{aligned} \quad (29)$$

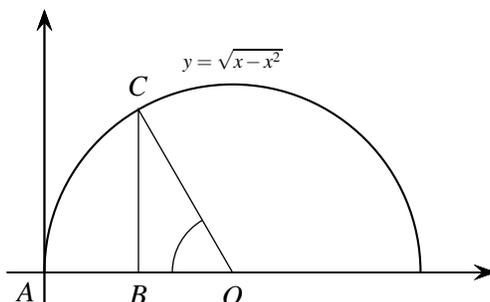


Figura 33. Cuadratura $\int_0^{1/4} \sqrt{x-x^2} dx$

En la figura 33 se ha representado el semicírculo de centro $(1/2, 0)$ y radio $1/2$. El sector circular COA tiene amplitud $\pi/3$ por lo que su área es la tercera parte de la del semicírculo, es decir, $\pi/24$. Como $BC = \sqrt{3}/4$, el área del triángulo BOC es $\sqrt{3}/32$. Por otra parte, la integral calculada en (29) es el área de la región ACB . Por tanto:

$$\int_0^{1/4} (x-x^2)^{1/2} dx + \frac{\sqrt{3}}{32} = \frac{\pi}{24}$$

Deducimos que

$$\pi = \frac{3\sqrt{3}}{4} + 24 \left(\frac{2}{3 \cdot 2^3} - \frac{1}{5 \cdot 2^5} - \frac{1}{28 \cdot 2^7} - \frac{1}{72 \cdot 2^9} - \frac{5}{704 \cdot 2^{11}} - \dots \right)$$

Y de esta forma, Newton expresa la cuadratura del círculo por medio de una serie infinita que, además, converge rápidamente.

Newton utilizó series para resolver ecuaciones diferenciales de primer orden [12]; así, para integrar

$$\dot{y} = 2 + 3x - 2y + x^2 + x^2y$$

Supone que

$$y = A_0 + A_1x + A_2x^2 + \dots$$

Entonces

$$\dot{y} = A_1 + 2A_2x + 3A_3x^2 + \dots$$

Sustituyendo estos valores en la ecuación e igualando coeficientes de las mismas potencias de x , obtiene

$$A_1 = 2 - A_0, \quad 2A_2 = 3 - 2A_1, \quad 3A_3 = 1 + A_0 - 2A_2, \dots$$

Se determinan así los A_n el hecho de que A_0 queda indeterminado y que por lo tanto existen infinitas soluciones fue advertido, pero la importancia de una constante arbitraria no fue plenamente apreciada hasta aproximadamente 1750. Leibniz resolvió algunas ecuaciones diferenciales elementales por medio de series y utilizó también el método anterior de coeficientes indeterminados.

En *De Analysi* Newton aplica su método de *aproximaciones sucesivas* para el cálculo de soluciones de una ecuación, el ahora conocido como *método de Newton*, para revertir (invertir) series. Dicho método aplicado a la serie:

$$y = x - \frac{x^2}{2} + \frac{x^3}{3} - \frac{x^4}{4} + \dots$$

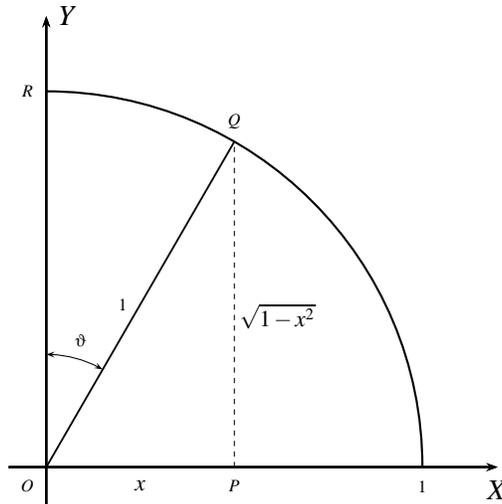
le llevó a:

$$x = y + \frac{1}{2}y^2 + \frac{1}{6}y^3 + \frac{1}{24}y^4 + \frac{1}{120}y^5 + \dots$$

Ya que $y = \log(1+x)$, o sea, $x = e^y - 1$, lo que ha obtenido Newton es el desarrollo por primera vez de la función exponencial:

$$e^y = 1 + y + \frac{1}{2}y^2 + \frac{1}{6}y^3 + \frac{1}{24}y^4 + \frac{1}{120}y^5 + \dots$$

También obtuvo por primera vez la serie del seno de la siguiente forma [14]. Consideremos la circunferencia $x^2 + y^2 = 1$ como en la figura siguiente.



El número $\vartheta = \text{arc sen } x$ es el doble del área del sector circular OQR . Newton, después de haber integrado $\sqrt{1-x^2}$ usando la serie binomial, sabía que el área del segmento $OPQR$ era

$$x - \frac{1}{6}x^3 - \frac{1}{40}x^5 - \frac{1}{112}x^7 - \dots$$

Se deduce que

$$\begin{aligned} \text{arc sen } x = \vartheta &= 2 \left(x - \frac{1}{6}x^3 - \frac{1}{40}x^5 - \frac{1}{112}x^7 + \dots \right) - x\sqrt{1-x^2} = \\ &= 2 \left(x - \frac{1}{6}x^3 + \frac{1}{40}x^5 - \frac{1}{112}x^7 + \dots \right) - x \left(1 - \frac{1}{2}x^2 - \frac{1}{8}x^4 - \frac{1}{16}x^6 - \dots \right) = \\ &= x + \frac{1}{6}x^3 + \frac{3}{40}x^5 + \frac{5}{112}x^7 + \dots \end{aligned}$$

Invirtiéndola esta serie, Newton obtiene la serie para el seno y se da cuenta de la *obvia* sucesión de los coeficientes:

$$\text{sen } x = \sum_{n=0}^{\infty} (-1)^n \frac{1}{(2n+1)!} x^{2n+1}, \quad \text{cos } x = \sum_{n=0}^{\infty} (-1)^n \frac{1}{(2n)!} x^{2n}$$

La confianza de Newton en los procesos infinitos queda reflejada en las siguientes palabras de la citada obra *De analysis*:

Todo lo que el análisis común [es decir, el álgebra] realiza por medio de ecuaciones con un número finito de términos, este nuevo método puede siempre conseguir lo mismo por medio

de ecuaciones infinitas, de tal forma que no he tenido ninguna duda en darle asimismo el nombre de análisis. Porque el razonamiento en éste no es menos cierto que en el otro; ni las ecuaciones menos exactas; aunque nosotros los mortales, cuyo poder de razonamiento está confinado dentro de estrechos límites, no podemos expresar ni tampoco concebir todos los términos de esas ecuaciones como para conocer exactamente a partir de ellas las cantidades que deseamos. . . Para terminar, podemos considerar todo esto como perteneciente al *Arte Analítica*, con cuya ayuda pueden ser determinadas de una manera exacta y geoméricamente las áreas, longitudes, etc., de curvas.

Es decir, Newton no sólo descubrió el teorema binomial sino que las series infinitas proporcionaban un método de análisis con la misma consistencia interna que el álgebra de ecuaciones finitas. De manera que, para Newton, las series infinitas no eran más que una parte del álgebra, un álgebra superior que trata de un número infinito de términos en lugar de un número finito.

5.3. Euler y el cálculo con series

Como la mayoría de los matemáticos del siglo XVIII, Euler hizo notables aportaciones al cálculo con series infinitas. Para hacernos una idea de cómo trabajaba Euler vamos a reproducir su desarrollo en serie de la función exponencial tal como aparece en el Volumen I de su famosa obra *Introductio in Analysin Infinitorum* ([8], pg. 531).

Dado $a > 1$, Euler escribe $a^\omega = 1 + k\omega$, donde ω debe considerarse un número *infinitamente pequeño* (“tan pequeño que justamente no es igual a cero”) y k es una constante que sólo depende de a . Para cualquier número real x pongamos $j = x/\omega$; entonces

$$a^x = a^{j\omega} = (1 + k\omega)^j = (1 + kx/j)^j,$$

y desarrollando por el binomio de Newton

$$a^x = 1 + \frac{j}{1!} \frac{kx}{j} + \frac{j(j-1)}{2!} \left(\frac{kx}{j}\right)^2 + \frac{j(j-1)(j-2)}{3!} \left(\frac{kx}{j}\right)^3 + \dots$$

Como ω es infinitamente pequeño, j es infinitamente grande, lo que permite a Euler suponer que

$$1 = \frac{j-1}{j} = \frac{j-2}{j} = \frac{j-3}{j} = \dots$$

Para concluir que

$$a^x = 1 + \frac{kx}{1!} + \frac{(kx)^2}{2!} + \frac{(kx)^3}{3!} + \dots$$

Y haciendo $x = 1$

$$a = 1 + \frac{k}{1!} + \frac{k^2}{2!} + \frac{k^3}{3!} + \dots$$

Puesto que para $k = 1$ se tiene que $\log a = \frac{\log(1 + \omega)}{\omega} = 1$ (pues ω es infinitamente pequeño), se sigue que $\log a = 1$. Este número Euler lo representa con el símbolo e . Por tanto

$$e = 1 + \frac{1}{1!} + \frac{1}{2!} + \frac{1}{3!} + \dots$$

Usando esta serie, Euler calcula $e=2.71828182845904523536028\dots$ con 23 cifras decimales. Naturalmente, la hipótesis de que ω es “infinitamente pequeño” puede evitarse sin más que usar límites.

Se debe también a Euler la divulgación a través de sus escritos del símbolo π para representar el cociente entre la longitud de una circunferencia y la de su diámetro; así mismo, en sus últimas obras introdujo el símbolo i para representar a la unidad imaginaria $\sqrt{-1}$. También le debemos las abreviaturas que seguimos usando para representar las funciones elementales, así como el empleo del símbolo \sum para indicar una suma, y, como indicamos en su momento, la notación $f(x)$ para indicar el valor de una función f en un número x . Ya en 1740, Euler, en una carta a Jean Bernoulli, usa exponentes imaginarios y escribe la igualdad $e^{x\sqrt{-1}} + e^{-x\sqrt{-1}} = 2 \cos x$. Las conocidas identidades de Euler aparecieron en su obra *Introductio*.

Veamos otro ejemplo del proceder de Euler; concretamente su cálculo de la suma de la serie de los inversos de los cuadrados de los números naturales. Oldenburg, en una carta a Leibniz en 1673, preguntaba por la suma de esta serie, pero Leibniz no respondió. Tampoco pudo calcularla Jacques Bernoulli. Euler parte de la serie conocida

$$\text{sen } x = x - \frac{x^3}{3!} + \frac{x^5}{5!} - \frac{x^7}{7!} + \dots$$

o bien, para $x \neq 0$:

$$\frac{\text{sen } x}{x} = 1 - \frac{x^2}{3!} + \frac{x^4}{5!} - \frac{x^6}{7!} + \dots$$

Interpretando la expresión de la derecha como un polinomio en x , sus raíces son los números $n\pi$ para n entero, poniendo $x^2 = y$, Euler considera la siguiente “ecuación”:

$$0 = 1 - \frac{y}{3!} + \frac{y^2}{5!} - \frac{y^3}{7!} + \dots$$

Cuyas soluciones son $n^2\pi^2$ con $n = 1, 2, 3, \dots$. Era sabido que la suma de los inversos de las raíces de una ecuación algebraica cuyo término independiente vale 1 es igual al opuesto del coeficiente del término de grado uno. Euler, tratando una serie como si fuera un polinomio, concluye que

$$\frac{1}{\pi^2} + \frac{1}{2^2\pi^2} + \frac{1}{3^2\pi^2} + \frac{1}{4^2\pi^2} + \dots = \frac{1}{6} \implies \sum_{n=1}^{\infty} \frac{1}{n^2} = \frac{\pi^2}{6}$$

De hecho, Euler fue un poco más allá y “factorizó” $1 - \frac{x^2}{3!} + \frac{x^4}{5!} - \frac{x^6}{7!} + \dots$ como si fuera un polinomio

por medio de un producto infinito:

$$1 - \frac{x^2}{3!} + \frac{x^4}{5!} - \frac{x^6}{7!} + \dots =$$

$$\left(1 + \frac{x}{\pi}\right) \left(1 - \frac{x}{\pi}\right) \left(1 + \frac{x}{2\pi}\right) \left(1 - \frac{x}{2\pi}\right) \left(1 + \frac{x}{3\pi}\right) \left(1 - \frac{x}{3\pi}\right) \dots =$$

$$\left(1 - \frac{x^2}{1^2\pi^2}\right) \left(1 - \frac{x^2}{2^2\pi^2}\right) \left(1 - \frac{x^2}{3^2\pi^2}\right) \dots$$

Igualando coeficientes en x^2 vuelve a obtener que $\sum_{n=1}^{\infty} \frac{1}{n^2} = \frac{\pi^2}{6}$. Euler era consciente de que sus razonamientos no estaban bien justificados, y en su obra ([10], §158) da una nueva demostración, ahora rigurosa, de este resultado.

Por otra parte, puesto que

$$\frac{\operatorname{sen} x}{x} = 1 - \frac{x^2}{3!} + \frac{x^4}{5!} - \frac{x^6}{7!} + \dots = \left(1 - \frac{x^2}{1^2\pi^2}\right) \left(1 - \frac{x^2}{2^2\pi^2}\right) \left(1 - \frac{x^2}{3^2\pi^2}\right) \dots$$

Haciendo en esta igualdad $x = \pi/2$ se obtiene

$$\frac{2}{\pi} = \left(1 - \frac{1}{4}\right) \left(1 - \frac{1}{16}\right) \left(1 - \frac{1}{36}\right) \dots = \frac{3}{4} \frac{15}{16} \frac{35}{36} \dots = \frac{1 \cdot 3 \cdot 3 \cdot 5 \cdot 5 \cdot 7 \cdot 7 \cdot 9 \dots}{2 \cdot 2 \cdot 4 \cdot 4 \cdot 6 \cdot 6 \cdot 8 \cdot 8 \dots}$$

que es la conocida fórmula de Wallis.

Euler estableció una conexión entre la serie armónica y los números primos que puede ser considerada el inicio de la teoría analítica de números. Probó la sorprendente “igualdad”

$$1 + \frac{1}{2} + \frac{1}{3} + \frac{1}{4} + \frac{1}{5} + \dots = \frac{2 \cdot 3 \cdot 5 \cdot 7 \cdot 11 \cdot 13 \dots}{1 \cdot 2 \cdot 4 \cdot 6 \cdot 10 \cdot 12 \dots} \quad (30)$$

donde, explica Euler, “el numerador en la derecha es el producto de todos los números primos y el denominador el producto de todos los números una unidad menores que los primos”.

Euler empieza su “prueba” escribiendo $S = 1 + \frac{1}{2} + \frac{1}{3} + \frac{1}{4} + \frac{1}{5} + \dots$, él sabe que S es una cantidad infinita, no es en absoluto un número, pero eso no va a detenerle y en sus cálculos trata S como si fuera un número con las reglas usuales. Así, escribe:

$$\frac{1}{2}S = S - \frac{1}{2}S = \left(1 + \frac{1}{2} + \frac{1}{3} + \frac{1}{4} + \frac{1}{5} + \dots\right) - \left(\frac{1}{2} + \frac{1}{4} + \frac{1}{6} + \frac{1}{8} + \frac{1}{10} + \dots\right) = 1 + \frac{1}{3} + \frac{1}{5} + \frac{1}{7} + \frac{1}{9} + \dots \quad (31)$$

Y observa que en la serie obtenida no hay denominadores pares. Multiplicando esta serie por $\frac{1}{3}$ obtiene:

$$\frac{1}{3} \left(\frac{1}{2}S\right) = \frac{1}{3} + \frac{1}{9} + \frac{1}{15} + \frac{1}{21} + \frac{1}{27} + \dots$$

Y sustrayendo esta serie de la anterior (31):

$$\frac{1}{2}S - \frac{1}{3} \left(\frac{1}{2}S\right) = \frac{1 \cdot 2}{2 \cdot 3}S = \left(1 + \frac{1}{3} + \frac{1}{5} + \frac{1}{7} + \frac{1}{9} + \dots\right) - \left(\frac{1}{3} + \frac{1}{9} + \frac{1}{15} + \frac{1}{21} + \frac{1}{27} + \dots\right) =$$

$$= 1 + \frac{1}{5} + \frac{1}{7} + \frac{1}{11} + \frac{1}{13} + \dots$$

Y en esta última serie los denominadores no contienen a 2 ni a 3 como factores. La siguiente etapa de este proceso sería

$$\begin{aligned} \frac{1 \cdot 2}{2 \cdot 3} S - \frac{1}{5} \frac{1 \cdot 2}{2 \cdot 3} S &= \frac{1 \cdot 2 \cdot 4}{2 \cdot 3 \cdot 5} S = \\ &= \left(1 + \frac{1}{5} + \frac{1}{7} + \frac{1}{11} + \frac{1}{13} + \dots \right) - \left(\frac{1}{5} + \frac{1}{25} + \frac{1}{35} + \frac{1}{55} + \frac{1}{65} + \dots \right) = \\ &= 1 + \frac{1}{7} + \frac{1}{11} + \frac{1}{13} + \frac{1}{17} + \dots \end{aligned}$$

Y en la serie obtenida los denominadores no contienen a 2, ni a 3, ni a 5 como factores. Para Euler el proceso ya está claro: en cada etapa removemos todos los denominadores que son divisibles por un primo y generamos una serie reducida que empieza en $1 + \frac{1}{p}$ en la que en el siguiente paso eliminaremos los denominadores divisibles por el primo p . Alguien como Euler, que no se asusta con los procesos infinitos, concluye que después de una infinitud de divisiones y sustracciones obtendríamos:

$$\frac{1 \cdot 2 \cdot 4 \cdot 6 \cdot 10 \cdot 12 \dots}{2 \cdot 3 \cdot 5 \cdot 7 \cdot 11 \cdot 13 \dots} S = 1$$

Y multiplicando en cruz resulta finalmente

$$1 + \frac{1}{2} + \frac{1}{3} + \frac{1}{4} + \frac{1}{5} + \dots = S = \frac{2 \cdot 3 \cdot 5 \cdot 7 \cdot 11 \cdot 13 \dots}{1 \cdot 2 \cdot 4 \cdot 6 \cdot 10 \cdot 12 \dots}$$

¿No es asombroso?

Para sacar algo de provecho a este ingenioso razonamiento de Euler, observemos que para todo natural $p \geq 2$:

$$\frac{p}{p-1} = \frac{1}{1-\frac{1}{p}} = \sum_{n=0}^{\infty} \frac{1}{p^n}$$

Por lo que, si representamos por \mathcal{P} el conjunto de todos los primos, podemos escribir la igualdad (30) en la forma:

$$\sum_{n=1}^{\infty} \frac{1}{n} = \prod_{p \in \mathcal{P}} \frac{1}{1-\frac{1}{p}} \quad (32)$$

Si p, q, r son números primos, los inversos de los números de la forma $p^n q^m r^l$ donde $n, m, l \in \mathbb{N} \cup \{0\}$, tienen una suma finita pues dicha suma es:

$$\sum_{n,m,r=0}^{\infty} \frac{1}{p^n q^m r^l} = \left(\sum_{n=0}^{\infty} \frac{1}{p^n} \right) \left(\sum_{m=0}^{\infty} \frac{1}{q^m} \right) \left(\sum_{l=0}^{\infty} \frac{1}{r^l} \right) = \frac{1}{1-\frac{1}{p}} \frac{1}{1-\frac{1}{q}} \frac{1}{1-\frac{1}{r}}$$

Deducimos de la igualdad 32 que hay infinitos números primos. Aunque los razonamientos de Euler pueden calificarse de *matemáticas*, esta última afirmación puede probarse de manera impecable a partir de las ideas desarrolladas usando la divergencia de la serie armónica. Es una manera bastante llamativa de probar un resultado conocido desde la antigüedad. Te lo dejo para que lo hagas tú.

Aunque la igualdad 32 carece de sentido, Euler, después de probarla, afirma ([10], §274) que también se cumple para potencias cualesquiera, aunque él está pensando en potencias enteras, es decir que para $s \in \mathbb{N}$ se verifica la igualdad:

$$\sum_{n=1}^{\infty} \frac{1}{n^s} = \prod_{p \in \mathcal{P}} \frac{1}{1 - \frac{1}{p^s}}$$

De hecho, esta igualdad es válida y hay convergencia para todo número complejo $s \in \mathbb{C}$ con $\text{Re}(s) > 1$. Dicha igualdad es conocida como *producto de Euler*. La serie de la izquierda permite definir, por extensión analítica a $\mathbb{C} \setminus \{1\}$, la famosa *función zeta* de Riemann.

5.4. El problema de la convergencia y la divergencia

En esta sección sigo muy de cerca el libro [12]. En los trabajos sobre series del siglo XVIII dominó el punto de vista formal, su justificación estaba en la utilidad de los resultados obtenidos sin que preocuparan mucho las cuestiones de convergencia o divergencia. Newton, Leibniz, Euler e incluso Lagrange consideraban las series de potencias como una extensión del álgebra de polinomios y no advertían que al extender las sumas a un número infinito de términos estaban dando lugar a nuevos problemas. Era justamente la carencia de un concepto preciso de límite lo que llevaba a los matemáticos de la época a contemplar el cálculo infinitesimal de un modo ingenuo, como una extensión del álgebra. Acabamos de ver algunos ejemplos de cómo trabajaba Euler con series sin preocuparse de su posible convergencia, no obstante obtenía resultados que posteriormente o bien han podido probarse de manera rigurosa o, cuando ello no era posible, han proporcionado las ideas clave para probar resultados de gran utilidad, especialmente en teoría analítica de números. Pero a veces se obtenían resultados muy sorprendentes, sobre todo cuando trabajaban con series divergentes. Veamos algunos ejemplos.

Sea $S = 1 - 1 + 1 - 1 + 1 - 1 + \dots$. Entonces $S - 1 = -(1 - 1 + 1 - 1 + \dots) = -S$. Luego $S = 1/2$.

A este resultado también se llega haciendo $t = 1$ en la igualdad $\frac{1}{1+t} = 1 - t + t^2 - t^3 + t^4 - t^6 + \dots$.

¿Te parece un disparate? Veamos. Si $\sum c_n$ es una serie numérica convergente, entonces la serie $\sum c_n t^n$ converge uniformemente en $[0, 1]$ y $\lim_{t \rightarrow 1} \sum_{n=1}^{\infty} c_n t^n = \sum_{n=1}^{\infty} c_n$. Este resultado se conoce como *teorema límite de Abel*.

Se dice que una serie $\sum a_n$ es sumable en el sentido de Abel, o *A-sumable*, si existe el límite $\lim_{t \rightarrow 1} \sum_{n=1}^{\infty} a_n t^n$, cuyo valor se llama la *A-suma* o *suma de Abel* de la serie. Acabamos de ver que la serie $1 - 1 + 1 - 1 + 1 - 1 + \dots$ es *A-sumable* y su *A-suma* es $1/2$. Es decir, el teorema límite de Abel permite asignar un valor numérico a algunas series que no son convergentes en sentido usual. Nada se pierde con ello porque dicho teorema nos dice que toda serie convergente en el sentido usual es Abel sumable y su *A-suma* coincide con su suma.

Sea $B = 1 - 2 + 3 - 4 + 5 - 6 + \dots$. Entonces

$$\begin{aligned} B &= 1 + (-2 + 3 - 4 + 5 - 6 + \dots) \\ &= 1 - (1 - 1 + 1 - 1 + 1 - 1 + \dots) - (1 - 2 + 3 - 4 + 5 - 6 + \dots) \\ &= 1 - A - B \end{aligned}$$

Luego $B = 1/4$. Esto es correcto si interpretamos la suma en el sentido de Abel pues se tiene que:

$$\begin{aligned} f(t) &= \sum_{n=1}^{\infty} (-1)^{n-1} n t^n = t \sum_{n=1}^{\infty} (-1)^{n-1} n t^{n-1} = t \frac{d}{dt} \sum_{n=1}^{\infty} (-1)^{n-1} t^n \\ &= t \frac{d}{dt} \sum_{n=1}^{\infty} (-1)^{n-1} t^{n-1} = t \frac{d}{dt} \sum_{n=1}^{\infty} (-t)^{n-1} = t \frac{d}{dt} \left(\frac{t}{1+t} \right) = \frac{t}{(1+t)^2} \end{aligned}$$

Claramente $\lim_{t \rightarrow 1} f(t) = 1/4$. Es decir, la serie $\sum (-1)^{n-1} n = 1 - 2 + 3 - 4 + 5 - 6 + \dots$ es A -sumable y su A -suma es $1/4$.

Hay otras formas de asociar un significado a ciertas series que no son convergentes en el sentido usual. Pero antes que nada era necesario precisar el concepto de convergencia para series.

La distinción entre convergencia y divergencia había sido ya tenida en cuenta por algunos matemáticos del siglo XVII. De hecho, los términos “convergente” y “divergente” fueron empleados en 1668 por James Gregory, pero no desarrolló sus ideas. Newton reconoció la necesidad de examinar la convergencia, pero no pasó de afirmar que, para valores pequeños de la variable, las series de potencias convergen al menos tan bien como la serie geométrica.

También Leibniz mostró cierto interés acerca de la convergencia y en una carta del 25 de octubre de 1713 indicó a Jean Bernoulli lo que hoy es un teorema, a saber, que una serie cuyos términos alternan de signo y decrecen en valor absoluto monótonamente hacia cero, es convergente.

Edward Waring (1734-98), Lucasian Professor de matemáticas en la universidad de Cambridge, probó que las series (que ahora llamamos “de Riemann”) $\sum_{n \geq 1} \frac{1}{n^\alpha}$ convergen para $\alpha > 1$ y divergen para $\alpha \leq 1$.

En 1768 D’Alembert dio un criterio para la convergencia absoluta de una serie $\sum a_n$, a saber: que exista un número $0 < \rho < 1$ tal que para todo n mayor o igual que un cierto n_0 se verifique que $|a_{n+1}|/|a_n| < \rho$.

Cauchy en su *Course d’analyse* (1821) al principio del Capítulo VI da las siguientes definiciones:

Se llama *serie* a una sucesión indefinida de cantidades $u_0, u_1, u_2, u_3, \dots$ que derivan unas de otras según una ley determinada. Estas cantidades son ellas mismas los diferentes *términos* de la serie considerada. Sea $S_n = u_0 + u_1 + u_2 + \dots + u_{n-1}$ la suma de los primeros n términos, n designando un número entero cualquiera. Si, para valores de n siempre crecientes, la suma S_n se aproxima indefinidamente a un cierto límite S , se dirá que la serie es *convergente* y el límite en cuestión se llamará *suma* de la serie. Al contrario, si mientras que n

crece indefinidamente, la suma S_n no se aproxima a ningún límite fijo, se dirá que la serie es *divergente* y no tendrá suma.

Un poco más adelante, después de estudiar la serie geométrica dice:

Para que la serie $u_0 + u_1 + u_2 + u_3 + \dots$ sea convergente es necesario y suficiente que para valores infinitamente grandes del número n las sumas $S_n, S_{n+1}, S_{n+2}, \dots$ difieran del límite S , y en consecuencia entre ellas, por cantidades infinitamente pequeñas.

La necesidad de esta condición es clara pero Cauchy no podía probar su suficiencia porque todavía no se había establecido la propiedad de completitud de los números reales. Una demostración rigurosa de lo que ahora conocemos como “condición de Cauchy” requiere la construcción previa del cuerpo de los números reales.

Cauchy establece en este libro una serie de criterios de convergencia para series de términos positivos, entre ellos, el que ahora se conoce como “criterio de Cauchy” el cual pasó inadvertido y fue redescubierto por Hadamard en 1892 y es un resultado fundamental para calcular el radio de convergencia de una serie de potencias compleja. El estudio de los criterios de convergencia para series refleja el cambio radical en la forma de entender el Análisis con respecto a la tradición del siglo XVIII. Cauchy establece la convergencia de una serie bajo convenientes hipótesis sobre su término general sin necesidad de conocer el valor de la suma de la serie.

También “demuestra” Cauchy un teorema que afirma que *la suma de una serie de funciones continuas es una función continua*. Un sencillo ejemplo prueba que Cauchy se equivoca. La serie $\sum u_n(x)$ donde

$$u_n(x) = \frac{1}{(n-1)x+1} - \frac{1}{nx+1} \quad (x > -1)$$

cuyas sumas parciales son $S_n(x) = 1 - \frac{1}{nx+1}$ tiene como suma la función $S(x) = 1$ para $x > -1$, $x \neq 0$, y $S(0) = 0$, que es discontinua en 0.

Poco tiempo después, en 1826, Abel, en un artículo sobre la convergencia de la serie binomial, observó que este resultado de Cauchy “debía tener excepciones”, que era una forma educada de decir que el teorema era falso y necesitaba hipótesis adicionales. No obstante, Abel también cometió errores parecidos en sus demostraciones. La razón de esto es el uso de los infinitesimales en las demostraciones que impedía ver las relaciones de dependencia entre las distintas variables. Naturalmente, es el concepto de *convergencia uniforme* el que se necesita para que el teorema de Cauchy sea correcto. Dicho concepto fue introducido por Weierstrass en sus clases de la Universidad de Berlin a principio de los años 1860.

5.5. El problema de la cuerda vibrante y las series trigonométricas

Supongamos que tenemos una cuerda elástica situada en el eje de abscisas y fijada por sus extremos que suponemos situados en $(0,0)$ y $(\ell,0)$, de manera que quede tensa. Ahora la sometemos a un desplazamiento formando una curva $y = f(x)$ y, seguidamente, la soltamos. Suponiendo que las vibraciones de la cuerda se producen en el plano XY . ¿Qué función $u(x,t)$ describe el movimiento de dicha cuerda en función de la abscisa $x \in [0, \ell]$ y el tiempo $t \geq 0$? Se pretende obtener la función de dos variables $u(x,t)$ que nos da el desplazamiento vertical del punto de la cuerda de abscisa x en el tiempo t a partir de la posición inicial de la cuerda dada por $y = f(x)$.

D'Alembert, en 1747, se interesó por el problema y, bajo diversas hipótesis físicas (fundamentalmente que las vibraciones sean “pequeñas”), demostró que la función u debe satisfacer las siguientes condiciones:

$$\begin{aligned}\frac{\partial^2 u}{\partial t^2}(x,t) &= a^2 \frac{\partial^2 u}{\partial x^2}(x,t) \quad (0 < x < \ell, \quad t > 0) \\ u(x,0) &= f(x) \quad 0 \leq x \leq \ell \\ \frac{\partial u}{\partial t}(x,0) &= 0 \quad 0 \leq x \leq \ell \\ u(0,t) &= u(\ell,t) = 0 \quad t \geq 0\end{aligned}$$

La primera igualdad, donde $a > 0$ es una constante que refleja algunas características físicas de la cuerda, es una ecuación diferencial en derivadas parciales de segundo orden, llamada *ecuación de ondas*. La segunda igualdad nos da la forma inicial de la cuerda, la tercera indica que la velocidad inicial de cada partícula es cero, y la cuarta que la cuerda se mantiene fija en sus extremos. De una forma muy ingeniosa D'Alembert probó que la solución viene dada por

$$u(x,t) = \frac{1}{2}\psi(at+x) - \frac{1}{2}\psi(at-x)$$

donde ψ es una función impar que verifica que $\psi(x) = f(x)$ para todo $x \in [0, \ell]$ y es periódica con período 2ℓ , esto es ψ es la extensión impar 2ℓ -periódica de f . D'Alembert creía que la función ψ , y por tanto f , debía ser una “expresión analítica”, esto es, debía venir dada por una “fórmula”, pues para él estas eran las únicas funciones permitidas. Además, ya que esta expresión analítica satisface la ecuación debe ser dos veces diferenciable, lo que implica que la función f debe ser dos veces derivable.

En 1748, Euler escribió un trabajo sobre el mismo problema en el que aunque coincidía con la solución de D'Alembert, su interpretación de la misma era muy diferente. A partir de consideraciones físicas, Euler, argumentó que la forma inicial de la cuerda podía venir dada por varias expresiones analíticas en diferentes subintervalos de $]0, \ell[$, o incluso dibujarse libremente a mano⁸. Realmente, las funciones admitidas por Euler como posición inicial de la cuerda serían lo que hoy llamamos “funciones continuas de clase C^1 a trozos”.

⁸Euler llamaba a estas funciones *discontinuas*, reservando la palabra *continua* para funciones dadas por una única expresión analítica. Esta idea de “continuidad” permaneció hasta Cauchy

Euler considera el caso particular en el perfil inicial de la cuerda viene dado por una función del tipo

$$f(x) = \sum A_n \operatorname{sen} \frac{n\pi x}{\ell}$$

aunque no aclara si el signo de sumación se refiere a un número finito o infinito de términos ([12], pg. 673).

En tal caso Euler obtiene que

$$u(x, t) = \sum A_n \operatorname{sen} \frac{n\pi x}{\ell} \cos \frac{n\pi at}{\ell}$$

En 1753 Daniel Bernoulli (1700-1782) estudió el mismo problema. Parte de la afirmación, justificada por razones físicas sobre las distintas frecuencias simultáneamente presentes en un sonido, de que la forma inicial de la cuerda siempre puede escribirse en la forma

$$f(x) = \sum_{n=0}^{\infty} a_n \operatorname{sen} \frac{n\pi x}{\ell} \quad (33)$$

porque, afirma, “existen suficientes constantes a_n para que la serie se ajuste a cualquier curva”. En consecuencia, afirma, todas las soluciones vendrán dadas por

$$u(x, t) = \sum_{n=0}^{\infty} a_n \operatorname{sen} \frac{n\pi x}{\ell} \cos \frac{n\pi at}{\ell}$$

Por supuesto, ni Euler ni D’Alembert aceptaban que cualquier función pudiera escribirse en la forma (33). La discusión entre D’Alembert, Euler y Bernoulli continuó por una década sin que se alcanzase un acuerdo. Mientras que para D’Alembert el concepto de función limitaba los posibles perfiles iniciales de la cuerda, para Euler la gran variedad de posibles perfiles iniciales forzaba a una extensión del concepto de función. La esencia del problema era la amplitud de la clase de funciones que se podían representar mediante una serie como en 33. Y también, claro está, que se necesitaba ya un nuevo concepto de función.

Una *serie trigonométrica* es cualquier serie de la forma

$$\frac{a_0}{2} + \sum_{n \geq 1} (a_n \cos(n\pi x/\ell) + b_n \operatorname{sen}(n\pi x/\ell)) \quad (34)$$

donde a_n, b_n son números complejos, $\ell > 0$ es una constante y $x \in \mathbb{R}$. Observa que se trata de una serie de funciones periódicas con período 2ℓ . Cuando hay alguna función f para la que se verifican las igualdades

$$a_n = \frac{1}{\ell} \int_{-\ell}^{\ell} \cos(n\pi x/\ell) f(x) dx \quad n = 0, 1, 2, \dots; \quad b_n = \frac{1}{\ell} \int_{-\ell}^{\ell} \operatorname{sen}(n\pi x/\ell) f(x) dx \quad n = 1, 2, \dots \quad (35)$$

se dice que la serie 34 es la *serie de Fourier* de f . Es frecuente que $\ell = \pi$ con lo cual se tiene:

$$a_n = \frac{1}{\pi} \int_{-\pi}^{\pi} \cos(nx) f(x) dx \quad n = 0, 1, 2, \dots \quad b_n = \frac{1}{\pi} \int_{-\pi}^{\pi} \operatorname{sen}(nx) f(x) dx \quad n = 1, 2, \dots \quad (36)$$

En el prólogo de su famoso libro de 1958 sobre series trigonométricas [27] Zygmund escribe:

Esta teoría ha sido una fuente de nuevas ideas para los analistas durante los dos últimos siglos y probablemente lo será en los próximos años. Muchas nociones y resultados básicos de la teoría de funciones han sido obtenidos por los matemáticos trabajando sobre series trigonométricas. Es concebible pensar que estos descubrimientos podían haber sido realizados en contextos diferentes, pero de hecho nacieron en conexión con la teoría de las series trigonométricas. No fue accidental que la noción de función aceptada ahora generalmente fuera formulada en la celebrada memoria de Dirichlet (1837) que trata de la convergencia de las series de Fourier, o que la definición de integral de Riemann en su forma general apareciera en el “Habilitationsschrift” de Riemann sobre series trigonométricas, o que la teoría de conjuntos, uno de los desarrollos más importantes de las matemáticas del siglo XIX, fuera creada por Cantor en su intento de resolver el problema de los conjuntos de unicidad para series trigonométricas. En épocas más recientes, la integral de Lebesgue se desarrolló en estrecha conexión con la teoría de series de Fourier, y la teoría de funciones generalizadas (distribuciones) con la de las integrales de Fourier.

5.6. La ecuación del calor y las series de Fourier

La máquina de vapor (James Watt, 1769), protagonista indiscutible de la Primera Revolución Industrial, así como otros problemas de interés científico o técnico, motivaron el interés por desarrollar una teoría matemática de la difusión del calor, dando así lugar a los inicios de la termodinámica. En el año 1811 la Academia de Ciencias de París propuso el problema de la propagación del calor como materia del gran premio que sería asignado en 1812. Este premio fue ganado por el matemático y físico francés Jean Baptiste-Joseph Fourier. El trabajo que presentó Fourier se titulaba *Théorie du mouvement de la chaleur dans les corps solides* (1811), y era una revisión de otro trabajo inicial, *Mémoire sur la propagation de la chaleur dans les corps solides* (1807), el cual también había sido presentado a la Academia en 1807 pero había sido rechazado. El trabajo de 1811, aunque ganador del premio, no fue publicado por la Academia porque, según el jurado formado por Lagrange, Laplace, Lacroix y Monge “la manera en la que el autor llega a sus ecuaciones no está exenta de dificultades y su análisis para integrarlas deja aún algo que desear, sea respecto a la generalidad, sea incluso del lado del rigor”. En 1822 Fourier publicó por su cuenta su libro *Théorie analytique de la chaleur* que muy pronto fue considerado como una de las grandes obras de la Física Matemática y donde incluyó parte de su trabajo de 1812. Finalmente, Fourier, siendo ya Secretario Perpetuo de la Academia, logró publicar su trabajo ganador de 1812 en dos partes aparecidas en 1824 y 1826. Fourier fue un hombre comprometido con su tiempo, consideraba el análisis infinitesimal como la principal herramienta para comprender los fenómenos naturales, firmemente convencido de que las ecuaciones diferenciales apoyadas en datos experimentales previos eran el modelo adecuado para ello, fue el prototipo de lo que hoy consideramos un “matemático aplicado”.

De los diversos problemas de conducción del calor estudiados por Fourier, vamos a considerar el caso más sencillo de una varilla de longitud ℓ , homogénea y delgada, cuyos extremos se mantienen a 0°C y cuya superficie lateral está aislada. Se supone que la temperatura de la varilla en cada sección transversal de la

misma es constante, y que la distribución de la temperatura inicial en la varilla está dada por una función conocida $f(x)$ que nos da la temperatura en el momento $t = 0$ de la sección de la varilla de abscisa x . Fourier demostró, sobre la base de principios físicos, que la función $T(x, t)$ que da el valor de la temperatura en la sección de abscisa x en el tiempo t debe satisfacer la siguiente ecuación diferencial, llamada ecuación del calor en una dimensión:

$$\frac{\partial^2 T}{\partial x^2}(x, t) = k^2 \frac{\partial T}{\partial t}(x, t) \quad 0 < x < \ell, \quad t > 0 \quad (37)$$

Sujeta a las condiciones de contorno (que expresan que la temperatura en los extremos se mantiene fija en 0°C):

$$T(0, t) = T(\ell, t) = 0 \quad t > 0;$$

y la condición inicial

$$T(x, 0) = f(x) \quad 0 \leq x \leq \ell$$

Inspirándose en las ideas de D. Bernoulli para la ecuación de ondas y usando el método de separación de variables, Fourier encontró que las soluciones más elementales venían dadas por funciones de la forma

$$T_n(x, t) = b_n e^{(-n^2\pi^2/k^2\ell^2)t} \text{sen} \frac{n\pi x}{\ell}$$

donde b_n es una constante real. Como la ecuación 37 es lineal, Fourier aplica el principio de superposición de soluciones para obtener como solución

$$T(x, t) = \sum_{n=1}^{\infty} b_n e^{(-n^2\pi^2/k^2\ell^2)t} \text{sen} \frac{n\pi x}{\ell}$$

La condición inicial $T(x, 0) = f(x)$ implica que

$$f(x) = \sum_{n=1}^{\infty} b_n \text{sen} \frac{n\pi x}{\ell} \quad 0 \leq x \leq \ell \quad (38)$$

Fourier se enfrentó entonces a la cuestión ¿puede $f(x)$ ser representada como una serie trigonométrica? En particular, ¿pueden ser determinadas las b_n ? Daniell Bernoulli se había limitado a afirmar que dicha representación era posible pero nada dijo de cómo podían calcularse los b_n . Fourier calcula dichos coeficientes de dos formas. En un primer intento desarrolla los senos y la propia función f en serie de potencias, permuta las sumas a la derecha e identifica coeficientes obteniendo un sistema de infinitas ecuaciones lineales con los b_n como incógnitas. De forma ingeniosa y muy trabajosa consigue finalmente llegar a las soluciones. También obtiene dichas soluciones de una forma mucho más sencilla y directa usando las relaciones de ortogonalidad. De una y otra forma obtuvo que

$$b_n = \frac{2}{\ell} \int_0^{\ell} f(x) \text{sen} \frac{n\pi x}{\ell} dx$$

Fourier observa que el cálculo de los coeficientes usando relaciones de ortogonalidad también puede aplicarse para una representación del tipo

$$g(x) = \sum_{n=1}^{\infty} a_n \cos \frac{n\pi x}{\ell} \quad (39)$$

La serie 38 representa una función impar, y la serie 39 representa una función par. Como toda función es suma de una función par y otra impar, Fourier concluye que cualquier función f puede ser representada en el intervalo $] -\ell, \ell[$ de la forma:

$$f(x) = \frac{a_0}{2} + \sum_{n=1}^{\infty} (a_n \cos(n\pi x/\ell) + b_n \operatorname{sen}(n\pi x/\ell))$$

y los coeficientes pueden ser determinados multiplicando por $\cos(n\pi x/\ell)$ o por $\operatorname{sen}(n\pi x/\ell)$ e integrando entre $-\ell$ y ℓ lo que lleva a las igualdades 35. Fourier observó que estas igualdades expresan los coeficientes a_n y b_n como las áreas bajo las curvas $y = \cos(n\pi x/\ell)f(x)$ e $y = \operatorname{sen}(n\pi x/\ell)f(x)$ para x entre $-\ell$ y ℓ y que dichas áreas tienen sentido para funciones muy generales, no se precisa para ello ni siquiera que f sea continua. Las funciones que Fourier considera son, como se lee en ([15], pg. 552):

En general, la función $f(x)$ representa una sucesión de valores u ordenadas, cada una de las cuales es arbitraria. Como la abscisa x puede recibir una infinidad de valores, hay un número igual de ordenadas $f(x)$. Todas tienen valores numéricos concretos, sean positivos, negativos o nulos. No se supone que estas ordenadas estén sujetas a una ley común; se suceden de cualquier manera, y cada una de ellas se da como si fuese una sola cantidad.

Este concepto de función está muy cerca del actual, aunque en la práctica Fourier siempre considera funciones de clase C^1 a trozos.

Algunas consecuencias del trabajo de Fourier fueron:

- Ampliación del concepto de función. Antes, una función debía venir dada por una única expresión analítica cerrada (D'Alembert) o, como mucho, por varias expresiones analíticas que en conjunto representaban una curva (Euler). El trabajo de Fourier probaba que una misma curva podía representarse de distintas formas por medio de series de senos o de cosenos o de senos y cosenos, lo que ayudó a separar el concepto de función de sus representaciones concretas. Después de Fourier se fue haciendo notorio que una función puede ser inventada de forma bastante arbitraria. Fue Dirichlet en 1837 quien definió el actual concepto de función.
- Cambio de perspectiva en la interpretación de la integral. Durante el siglo XVIII se consideraba la integral esencialmente como una antiderivada. La interpretación de los coeficientes de Fourier como áreas bajo curvas muy generales llevó a centrar la atención en la integral como área y a preguntarse para qué tipo de funciones tenía sentido considerar su integral. Aunque las técnicas del cálculo integral

se remontan a la antigüedad, no fue hasta 1823 que Cauchy dio la primera *definición matemática* de integral con la que probaba que los coeficientes de Fourier tenían sentido para cualquier función continua y acotada con un número finito de discontinuidades. La integral de Cauchy fue ampliamente generalizada por Riemann en su “Habilitationsschrift” titulada *On the representability of a function by trigonometric series* escrita en 1854, pero no publicada hasta después de su muerte en 1867.

El concepto de integral de Riemann se mantuvo hasta los primeros años del siglo XX hasta la llegada de la integral de Lebesgue. La primera aplicación de esta nueva integral fue a un problema en series de Fourier.

- Necesidad de precisar los conceptos fundamentales del análisis: límite, convergencia y continuidad. Lo que, como ya sabemos, fue en principio hecho por Bolzano y por Cauchy en los años 1815-1825. Naturalmente, Fourier, no pudo demostrar su afirmación de que *cualquier* función podía expresarse como suma de una serie de Fourier puesto que tal afirmación estaba muy lejos de ser cierta, aunque algunos matemáticos como el mismísimo Cauchy dieron supuestas “demostraciones”.

Fue Dirichlet el primero que estableció condiciones necesarias que debe satisfacer una función para que esto pueda hacerse. 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 series trigonométriques qui servent à représenter une fonction arbitraire entre des limites données* en 1829, probó el siguiente resultado:

Supongamos que f satisface las siguientes hipótesis:

- f es continua en $[-\pi, \pi]$, salvo a lo más en un conjunto finito de puntos en los que tiene discontinuidades de salto.
- f tiene un número finito de máximos y mínimos locales en $[-\pi, \pi]$.

Entonces se verifica que la serie de Fourier de f converge a $f(x)$ en todo punto x de continuidad de f , y converge a $\frac{1}{2}(f(x+) - f(x-))$ en todos punto de discontinuidad.

La demostración dada por Dirichlet era rigurosa y exhibía nociones claras de continuidad, convergencia e integral (en el sentido de Cauchy) así como del concepto de función. Aunque las funciones consideradas por Dirichlet eran muy generales para la época, al final de este trabajo da un ejemplo, la ahora llamada *función de Dirichlet*:

$$D(x) = \begin{cases} c, & \text{si } x \text{ es racional;} \\ d, & \text{si } x \text{ es irracional.} \end{cases}$$

con $c \neq d$, que aparece aquí por primera vez, la cual no cumple las hipótesis del teorema pues ni siquiera es integrable según Cauchy. Se trata del primer ejemplo explícito de una función que no está dada por una o

varias expresiones analíticas, ni tampoco es una curva dibujada libremente a mano y es discontinua en todo punto. Los matemáticos desde Euler, pasando por Fourier, hasta Cauchy, habían hablado de “funciones de naturaleza arbitraria”, pero en la práctica consideraban las funciones como expresiones analíticas o curvas. Dirichlet fue el primero en considerar seriamente el concepto de función como correspondencia arbitraria. Además, Dirichlet, en su definición de función en 1837, fue de los primeros en restringir explícitamente el dominio de una función a un intervalo, anteriormente se suponía que la variable independiente podía tomar cualquier valor.

6. Breve historia del infinito

Es conocida la exclamación de David Hilbert *¡El infinito! Ninguna cuestión ha conmovido tan profundamente el espíritu del hombre*. Es verdad, el infinito atrae poderosamente nuestra imaginación. Es difícil imaginar que el tiempo tuviera un comienzo y también que el espacio sea finito, porque no podemos pensar en una frontera para el espacio tras de la cual no exista más espacio, ni un origen para el tiempo antes del cual no hubiera tiempo. Cualquier respuesta a estas preguntas conduce siempre a nuevas preguntas. Un error típico consiste en creer que si algo fuera infinito debería contener todas las cosas, algo así como el Aleph borgeano. Matemáticamente, es claro que no tiene por qué ser así: los números pares son infinitos y no son todos los números. Algo infinito tampoco tiene por qué ser necesariamente muy grande. El Aleph de la narración de Borges es una pequeña esfera, un conjunto fractal contiene infinitas copias de sí mismo, el veloz Aquiles permanece corriendo sin alcanzar jamás a la tortuga que le lleva unos pocos metros de ventaja. . . .

6.1. La idea de infinito en la filosofía y la matemática Griegas

6.1.1. Las aporías de Zenón de Elea

*¡Zenón, cruel Zenón, Zenón de Elea!
Me has traspasado con la flecha alada.
Que, cuando vibra volando, no vuela.
Me crea el son y la flecha me mata.
¡Oh sol, oh sol! ¡Qué sombra de tortuga
para el alma: si en marcha Aquiles, quieto!
Paul Valery*

Un griego llamado Zenón, del que se sabe muy poco y de forma indirecta a través del *Parménides* de Platón, cuyo nacimiento se fecha hacia el año 490 a.C en la ciudad de Elea en el sur de Italia, y que fue discípulo de Parménides, sigue manteniendo desde hace 2300 años su permanente desafío a la razón.

Te recuerdo que, según Parménides, el ser es necesariamente uno, eterno, continuo, indivisible e inmu-

table. Los cambios, transformaciones y multiplicación de los seres, son meras apariencias a las cuales no responde realidad alguna. La filosofía de Parménides fue muy criticada porque choca con nuestras creencias más básicas sobre la realidad.

Zenón ideó sus paradojas o aporías (proposiciones sin salida lógica) para desacreditar a quienes negaban las ideas de Parménides, y afirmaban la realidad del cambio y la pluralidad de los seres. Aristóteles califica a Zenón de “inventor de la dialéctica”, una elaborada forma de razonamiento que consiste en probar al oponente que de sus ideas se deducen consecuencias inaceptables. Los argumentos de Zenón son realmente del tipo “reducción al absurdo”: se acepta provisionalmente una hipótesis y, razonando correctamente a partir de ella, se llega a una conclusión inaceptable, lo que obliga a rechazar la hipótesis inicial.

Vamos a exponer, en lenguaje actual, tres de las paradojas de Zenón que van dirigidas contra las dos teorías del movimiento sostenidas en la antigüedad, las cuales dependen, claro está, de la supuesta naturaleza del tiempo y del espacio. Debes tener en cuenta que Zenón no niega el movimiento sino su inteligibilidad; la afirmación de que “el movimiento se demuestra andando” no refuta a Zenón, su desafío no es a la experiencia sensible sino a la razón.

Las dos primeras paradojas parten del supuesto de que el espacio y el tiempo son *infinitamente divisibles* y el movimiento continuo y uniforme.

La dicotomía.

Para que un móvil pueda llegar a un punto dado, debe recorrer primero la mitad de la distancia; pero antes de alcanzar esa mitad debe recorrer la mitad de la mitad. Y así sucesivamente, “ad infinitum”. De este modo para alcanzar completamente cualquier distancia tendría que recorrer un número infinito de divisiones, lo cual es imposible en un tiempo finito.

Aquiles y la tortuga.

Aquiles, el de los pies ligeros, nunca alcanzará a la tortuga que avanza lentamente unos cuantos metros por delante de él. Pues cuando Aquiles alcance el punto donde estaba la tortuga, ésta ya estará un poco más adelante; y cuando de nuevo Aquiles alcance ese lugar, la tortuga habrá avanzado un poco más. Sin desanimarse, sigue corriendo, pero al llegar de nuevo donde estaba la tortuga, esta ha avanzado un poco más... De este modo, la tortuga estará siempre por delante de Aquiles.

Ambos argumentos están relacionados. Según la Dicotomía, para que haya movimiento debe haber un comienzo, pero no hay una distancia mínima con la que empezar; por tanto el movimiento no puede empezar, luego no hay movimiento. Según Aquiles, un móvil para alcanzar su destino debe cubrir primero la mitad de la distancia que lo separa, pero antes deberá recorrer la mitad de esa mitad, y así sucesivamente; luego debe recorrer infinitas divisiones lo cual es imposible en tiempo finito, por tanto nunca alcanzará su destino. Es decir, una vez empezado, el movimiento no puede parar.

La tercera paradoja, que se refiere a una flecha lanzada al aire, supone que el espacio y el tiempo están formados por *unidades mínimas indivisibles* y el movimiento es una sucesión de diminutos saltos consecutivos.

La flecha.

En un instante indivisible de tiempo la flecha debe permanecer quieta, pues si se moviera el instante contendría unidades de tiempo más pequeñas en las que dicho movimiento tendría lugar en contra de lo supuesto. Por tanto, en cada instante la flecha está quieta y, como el tiempo se compone de instantes, la flecha está siempre quieta y el movimiento no tiene lugar.

La influencia de las aporías de Zenón en filosofía, lógica y matemáticas ha sido notable y se ha escrito y se sigue escribiendo mucho sobre ellas ([17] es una de las referencias más interesantes). Despidamos a Zenón con una cita de Borges.

Zenón es incontestable, salvo que confesemos la idealidad del espacio y del tiempo. Aceptemos el idealismo, aceptemos el crecimiento concreto de lo percibido, y eludiremos la pululación de abismos de la paradoja. ¿Tocar a nuestro concepto del universo, por ese pedacito de tiniebla griega?, interrogará mi lector.

J.L. Borges, “La perpetua carrera de Aquiles y la tortuga”.

6.1.2. Atomismo y divisibilidad infinita

El filósofo Anaximandro (ca.610 - 546 a.C.) introdujo el infinito en la filosofía Griega. Afirmó que el principio de todas las cosas existentes es el *ápeiron*. Etimológicamente *ápeiron* significa *lo sin límites*. Según Anaximandro, el *ápeiron* es infinito, porque provee la energía para que en el mundo no cese la generación y corrupción, e indeterminado, porque no es concreto y no se identifica con ninguno de los elementos agua, aire, tierra, fuego. Podemos interpretarlo como la fuente de energía primordial que garantiza la transformación y la unidad del cosmos.

En el período que separa a Zenón de Elea de Aristóteles surgió la filosofía del atomismo, iniciada por Leucipo (ca. 450 - 420 a.C.) y desarrollada por Demócrito (ca. 460 - 370 a.C.). El atomismo es una filosofía materialista que se ha interpretado como una respuesta al idealismo de la Escuela Eleática (Parménides, Zenón). Los atomistas mantienen que hay dos principios fundamentales: los átomos y el vacío. Los átomos son indivisibles e invisibles, infinitos en número y de diversas formas y tamaños, perfectamente sólidos, indestructibles y permanentes. Las sustancias materiales son producidas por la unión y separación de esos átomos moviéndose en el vacío. El movimiento se produce por la reordenación de los átomos entre sí; según Aristóteles, los atomistas reducen todo cambio a un mero cambio de lugar. Los atomistas admiten la pluralidad y el movimiento y niegan la infinita divisibilidad del espacio y la materia.

El atomismo fue cuestionado por [Aristóteles](#) (384 - 322 a.C.), que realizó un análisis sistemático del continuo. Aristóteles divide las cantidades en discretas y continuas. Los números y el lenguaje hablado son

discretas y las líneas, superficies, sólidos, tiempo y espacio son continuas. La respuesta a la pregunta de si una magnitud continua (un *continuo*) es permanentemente divisible en partes cada vez más pequeñas, o hay un límite más allá del cual no puede proseguirse el proceso de división, depende de la naturaleza del infinito. Aristóteles dedica el Libro III de su *Física* a un estudio sistemático del infinito. Considera que el estudio del infinito forma parte del estudio de la naturaleza, pues lo característico de ésta es el movimiento y el cambio, y el movimiento es pensado como algo continuo, y lo que es continuo es definido con frecuencia como algo infinitamente divisible.

Primero, dice Aristóteles, “*hay que examinar en general si es o no es posible que haya un cuerpo sensible infinito*”. Después del correspondiente estudio, llega a la conclusión de que “*no existe un cuerpo que sea actualmente infinito*”. Pero “*la negación absoluta del infinito es una hipótesis que conduce a consecuencias imposibles*”.

Aristóteles expone algunas razones que apoyan la creencia en la realidad del infinito y considera los distintos sentidos de dicho término. Entre las primeras: la infinitud del tiempo, la divisibilidad de las magnitudes y la infinitud de los números; entre los segundos: lo que no puede ser recorrido o se puede recorrer pero sin llegar a un término.

Es también evidente que no es posible que lo infinito exista como un ser en acto o como una substancia y un principio. Luego lo infinito existe como un atributo.

Lo infinito es un atributo que puede predicarse de la cantidad o de determinados procesos; especialmente, los procesos de adición y de división. Aristóteles, habla en ese sentido del infinito por adición y el infinito por división o la divisibilidad infinita de un continuo.

Ahora bien, el ser se dice o de lo que es en potencia o de lo que es en acto, mientras que el infinito es o por adición o por división. Y ya se ha dicho que la magnitud no es actualmente infinita [...] Nos queda, entonces, por mostrar que el infinito existe potencialmente.

Pero la expresión “existencia potencial” no se debe tomar en el sentido en que se dice, por ejemplo, “esto es potencialmente una estatua, y después será una estatua”, pues no hay un infinito tal que después sea en acto. Y puesto que el ser se dice en muchos sentidos, decimos que el infinito “es” en el sentido en que decimos “el día es”.

Aristóteles distingue, pues, dos clases de infinito: el infinito como una totalidad completa, que llama el infinito actual y cuya existencia niega; y el infinito potencial, que concibe como un proceso secuencial de adición o de subdivisión sin final. Lo metáfora del día es muy apropiada, pues el *ser* de un día es un *estar siendo* de forma sucesiva, de manera que en ningún momento el día queda realizado plenamente como un todo. Análogamente, el infinito potencial nunca será plenamente realizado *pues no hay un infinito tal que después sea en acto*. La infinitud potencial es la forma usual en que concebimos el tiempo como una línea recta indefinidamente prolongable o la sucesión de los números que podemos ir formando por adición consecutiva de la unidad.

Esta concepción aristotélica del infinito se aceptó sin mayores cambios hasta el siglo XIX. De todas formas, Aristóteles cree que la negación del infinito actual no afecta a los matemáticos:

Esta argumentación no priva a los matemáticos de sus especulaciones por el hecho de excluir que el infinito por adición pueda recorrerse en acto. Porque no tienen necesidad de este infinito ya que no hacen uso de él, sino sólo, por ejemplo, de una línea finita que se prolongue tanto como ellos quieran.

Sobre todo esto se ha escrito y se sigue escribiendo mucho. Más interesante para nosotros es la relación del infinito con la divisibilidad infinita del continuo.

Los atomistas negaban la divisibilidad infinita. Su argumento era que si una magnitud continua fuera dividida en todo punto, entonces no quedaría nada o solamente quedarían puntos sin extensión, porque en caso contrario el proceso de división podría proseguir. Pero, decían, si quedan puntos sin extensión, entonces no es posible recomponer la magnitud original a partir de ellos, pues por la agregación de puntos sin extensión no puede lograrse nunca una magnitud finita. Concluían que en cualquier caso la magnitud inicial se ha convertido en algo incorpóreo y, por tanto, algo que tenía existencia ha dejado de ser, lo cual, evidentemente, es un imposible.

Aristóteles defendía la divisibilidad infinita pero debía refutar el argumento atomista. Su solución es muy original, pues afirma que aunque una magnitud continua puede ser dividida en *cualquier* punto, no puede ser dividida en *todo* punto. Para Aristóteles, dividir un continuo en todos sus puntos es reducirlo a lo discreto. Mientras que un continuo tiene la propiedad de densidad, es decir, entre dos cualesquiera de sus puntos siempre hay otro punto del continuo, los puntos obtenidos, después de una división infinita actual de un continuo, serían adyacentes unos con otros, y esto implica que la propiedad de densidad se habría perdido. Pero si dividimos un continuo, lo que obtenemos son dos continuos cada uno de ellos con la propiedad de densidad. Por tanto, es imposible llegar, por divisiones sucesivas, a reducir un continuo a puntos. Así, Aristóteles afirma la divisibilidad infinita pero niega la divisibilidad en todo punto, con lo que el argumento atomista deja de tener valor.

Las matemáticas griegas evitan el infinito actual. Así, Euclides, considera rectas que pueden ser prolongadas cuanto se quiera, pero no “rectas infinitas”. Igualmente, al enunciar que los números primos son infinitos, lo expresa diciendo que “*Hay más números primos que cualquier cantidad de números primos propuesta*”. De esta forma evita considerar el infinito actual de los números primos.

En *Los Elementos* Euclides expone el *método de exhaustión* (4.1) de Eudoxo de Cnido, que se utilizaba para calcular áreas (cuadraturas) de regiones planas. Es frecuente afirmar que este método consiste en una aproximación al área seguida de un proceso límite. No es así. Aunque su nombre sugiere “agotamiento” de una figura plana por polígonos inscritos, el método estaba basado en un razonamiento muy cuidadoso de doble reducción al absurdo (llamado razonamiento *apagógico*), precisamente para evitar la consideración de un infinito actual.

Mención aparte merece Arquímedes. Por una parte, probó en su obra *El arenario* que si el Universo estuviera completamente lleno de granos de arena, su número sería finito. Para lo cual desarrolla un sistema de numeración apropiado para manejar grandes números (para los griegos el número mayor era la miríada de miríadas, equivalente a 10^8) que le permite describir un número que, en base diez, tendría unos 80000 millones de millones de cifras. Pero también Arquímedes ideó métodos heurísticos⁹ que están expuestos en su obra *El Método* (ver 4.1.2), descubierta en 1906, en la que explica cómo anticipó algunos de sus descubrimientos por medio de técnicas de equilibrio usando la ley de la palanca. En estas técnicas, Arquímedes hace un uso muy libre del infinito; por ejemplo, descompone áreas planas como sumas infinitas de segmentos, es decir, reduce un continuo a elementos indivisibles, con lo cual podrían estar de acuerdo los atomistas, pero no Aristóteles.

6.2. El infinito desde la Edad Media hasta el siglo XIX

Es sabido que las religiones lo contaminan todo de irrealidad. Después del triunfo de la Iglesia Católica, las discusiones sobre el infinito adquieren una orientación marcadamente teológica.

San Agustín (354 - 430), filósofo cristiano, admite el infinito actual como atributo de Dios, pero niega que Dios creara nada infinito. En su obra *La Ciudad de Dios* escribe refiriéndose a los números:

Así que son desiguales entre sí y diferentes; cada uno es finito y todos son infinitos. ¿Y que sea posible que Dios todopoderoso no sepa los números por su infinidad, y que la ciencia de Dios llegue hasta cierta suma de números, y que ignore los demás, quién habrá que pueda decirlo, por más ignorante y necio que sea? [...] Y así que la infinidad de los números para la ciencia de Dios, que la comprende, no puede ser infinita.

En esa insólita cuadratura del círculo que fue la Escolástica, en su intento de conciliar la filosofía de Platón y Aristóteles con la revelación cristiana, destaca Santo Tomás de Aquino (ca. 1225 - 1274). La infinitud actual de Dios en todos los sentidos es un dogma Católico y Tomás de Aquino es una autoridad en tan delicada cuestión teológica. En su obra *Summa Contra Gentiles*, Capítulo 43, proporciona catorce argumentos breves para demostrar la infinitud de Dios, cada uno de ellos termina con la letanía “*Por tanto Dios es infinito*”.

Para encontrar ideas más interesantes sobre el infinito debemos referirnos a Galileo Galilei (1564 - 1642). En su obra pionera sobre la dinámica y estática de sólidos *Discorsi e dimostrazioni matematiche intorno a due nuove scienze attenenti alla meccanica e i movimenti locali* (1638), Galileo expone algunas paradojas sobre el infinito. Una de ellas es la de la “equivalencia entre una circunferencia y un punto”. Para explicarla, consideremos un rectángulo formado por dos cuadrados iguales unidos por un lado común. Recortemos en este rectángulo una semicircunferencia de centro en la mitad del lado superior del rectángulo e igual radio. La

⁹Por *método heurístico* se entiende cualquier proceso que facilite anticipar un resultado. Son métodos que se apoyan en alguna forma de intuición que conduce a la formulación de conjeturas razonables, que después deben ser probadas con métodos científicos rigurosos

figura que resulta de quitar dicha semicircunferencia al rectángulo se gira alrededor de su eje de simetría y se obtiene un sólido de revolución parecido a un cuenco. Supongamos ahora inscrito en dicho sólido un cono circular recto cuya base coincide con la del cuenco y de altura igual a la del cuenco.

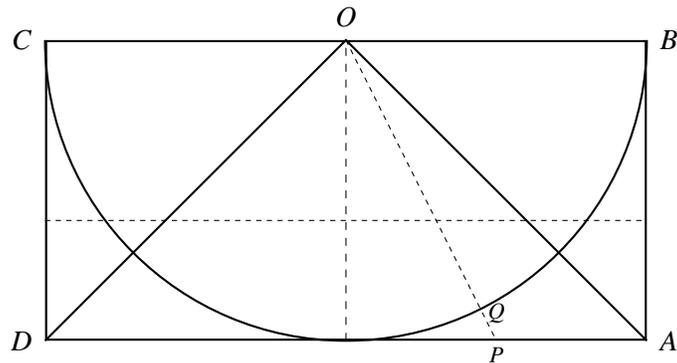


Figura 34. Paradoja circunferencia-punto

Cada plano paralelo a la base del cuenco determina en su intersección con el cono un círculo, y en su intersección con el cuenco una corona circular. Es muy fácil comprobar que dichos círculo y corona circular tienen igual área. Si ahora consideramos planos paralelos a la base del cuenco que se van acercando al borde superior del mismo, las áreas de las intersecciones de dichos planos con el cono y el cuenco son siempre iguales. El último de dichos planos da como intersección con el cuenco una circunferencia (el borde del cuenco) y con el cono un punto (el vértice del cono). Como los límites de cantidades iguales entre sí deben también ser iguales entre sí, Galileo se pregunta por qué no podemos considerar la circunferencia como igual a su centro. Si lo hacemos, llegaremos a la conclusión de que todas las circunferencias son iguales entre sí e iguales a un punto.

La misma figura anterior pone de manifiesto que la semicircunferencia está formada por tantos puntos como los que forman la poligonal $CDAB$. Pues cada semirrecta con origen en O corta a la semicircunferencia en un único punto Q y a la poligonal en otro único punto P . Así podemos emparejar los puntos de la semicircunferencia con los de la poligonal y de esta forma los agotamos todos. Por tanto ambas líneas tienen igual número infinito de puntos. Si llamamos ℓ a la longitud de AB , la semicircunferencia tiene longitud $\pi\ell$, menor que la longitud de la poligonal $CDAB$ que es igual a 4ℓ . Galileo escribe al respecto:

Estas dificultades son reales; y no son las únicas. Pero recordemos que estamos tratando con infinitos e indivisibles, los cuales trascienden nuestra comprensión finita, los primeros a causa de su magnitud, los últimos debido a su pequeñez.

[...] intentamos, con nuestras mentes finitas, discutir sobre el infinito, asignándole propiedades que damos a lo finito y limitado; pero pienso que esto es incorrecto, dado que no podemos hablar de cantidades infinitas como si fuesen mayores, menores o iguales a otras.

Otra paradoja considerada por Galileo, es la que se deduce de la observación de que para cada número natural n podemos construir un cuadrado de lado n , cuya área es igual a n^2 , de donde se deduce que hay tantos números naturales como cuadrados perfectos. Sin embargo la mayoría de los números no son cuadrados perfectos. A la vista de ello, Galileo escribe:

[...] el total de los números es infinito, y el número de cuadrados es infinito; ni es menor el número de cuadrados que el de la totalidad de números, ni el otro mayor que el anterior; y, finalmente, los atributos “igual”, “mayor” y “menor” no son aplicables al infinito, sino solo a cantidades finitas.

Una característica de las matemáticas del siglo XVII es el libre uso del infinito. En los dos primeros tercios del siglo XVII se desarrollan una variedad de métodos infinitesimales que preludian el cálculo diferencial, así como técnicas de cuadraturas basadas en la descomposición de recintos planos o de sólidos en infinitos *elementos indivisibles*. El matemático inglés John Wallis introdujo en 1655 en su obra *De Sectionibus Conicis*, el símbolo del “lazo del amor”, ∞ , con el significado de “infinito”.

La invención del Cálculo, en el último tercio del siglo XVII, ordena y sistematiza estos procedimientos, y proporciona algoritmos generales para resolver multitud de problemas que antes se abordaban con técnicas específicas para cada caso. Las cantidades infinitesimales, los casi imprescindibles infinitésimos, que ya son viejos amigos nuestros, son otra forma del infinito, en este caso, de lo infinitamente pequeño. Durante el siglo XVIII y parte del XIX, los infinitésimos se usaron de forma casi generalizada porque, a pesar de los problemas de todo tipo que planteaban, eran útiles y eficaces para resolver problemas y una herramienta heurística muy apreciada.

6.3. El infinito matemático y el nacimiento de la teoría de conjuntos

A principios del siglo XIX, la actitud de los matemáticos ante el infinito no era diferente a la mantenida por Galileo doscientos años antes. La consideración del infinito actual conducía a paradojas; en particular, la llamada *paradoja de la reflexividad*, es decir, la posibilidad de establecer una biyección entre un conjunto infinito y una parte del mismo, indicaba que la consideración del infinito actual contradecía el principio lógico de que “el todo es mayor que las partes”. Para los principales matemáticos de la época, como Gauss y Cauchy, el infinito seguía siendo un infinito potencial, un concepto sin contenido matemático, una palabra que servía para designar un proceso sin punto final. Gauss lo expresó claramente en una carta a su amigo Schumacher en 1831:

Debo protestar vehementemente contra el uso del infinito como algo completado, pues esto nunca está permitido en matemáticas. El infinito es simplemente una forma de hablar; una forma resumida para la afirmación de que existen los límites a los cuales ciertas razones pueden aproximarse tanto como se desee, mientras otras son permitidas crecer ilimitadamente.

La consideración del infinito actual como objeto matemático exige disponer de objetos matemáticos que puedan ser llamados “infinitos”. Que los números naturales son potencialmente infinitos quiere decir que son una sucesión a la que podemos agregar términos indefinidamente, muy diferente es la consideración del infinito actual de todos los números naturales (a lo que estamos ya acostumbrados y no nos causa mayor problema), que equivale a considerarlos como un todo acabado, como un conjunto formado por todos ellos. Esto indica que una teoría matemática del infinito supone la consideración de conjuntos infinitos. Es imposible separar la teoría de conjuntos y la teoría del infinito.

En esto, como en otras cosas, Bernahrd Bolzano fue un adelantado a su tiempo. En su libro *Las Paradojas del Infinito*, publicado en 1851, tres años después de su muerte, Bolzano se propone estudiar las paradojas conocidas y mostrar que, debido a la falta de precisión en el uso del término *infinito*, daban lugar a *aparentes* contradicciones. Es necesario, afirma, definir el término *infinito* y las matemáticas son el contexto apropiado para ello. Naturalmente, Bolzano, está refiriéndose al infinito actual. Con la idea de fundamentar matemáticamente la noción de infinito actual, Bolzano introduce los términos de *agregado*, *conjunto* y *multitud*, siendo en esta obra la primera vez que la palabra “conjunto” es usada con un significado matemático preciso. Un agregado es una totalidad compuesta de objetos bien definidos; un conjunto es un agregado donde el orden de sus partes es irrelevante y donde nada esencial se cambia si solo se cambia el orden (es decir, un agregado sin estructura alguna); una multitud es un conjunto cuyos miembros son individuos de una misma especie.

Bolzano considera un conjunto como un todo, sin necesidad de considerar separadamente cada uno de sus elementos. El ejemplo que propone es muy significativo a este respecto:

... puedo pensar en el conjunto, o agregado, o si se prefiere, en la totalidad de los habitantes de Praga o de Pekín sin formar una representación separada de cada habitante individual.

Bolzano abandona así el punto de vista constructivo, la idea de que un conjunto se va formando a partir de sus elementos mediante alguna clase de algoritmo.

Bolzano define una multitud infinita como aquella de la cual cualquier multitud finita solamente puede ser parte de la misma. Debemos observar que esta definición no es la tradicional en la que infinito es definido como la negación de lo finito. Con respecto a la existencia de conjuntos infinitos, Bolzano afirmó que “el conjunto de todas las verdades absolutas es un conjunto infinito”. Su idea es partir de una proposición que se sabe verdadera a la que podemos llamar A ; a partir de ella podemos formar otra “ A es verdadera” que, claramente, es diferente de la proposición A y este proceso puede proseguirse indefinidamente.

Bolzano mantiene que el criterio de validez para la existencia de conjuntos infinitos debe basarse en su naturaleza no contradictoria.

Tan pronto como disponemos de un concepto, A , el cual representa los objetos a, b, c, d, \dots y no otros, es extremadamente fácil llegar a un concepto que represente el agregado de todos estos objetos tomados juntos. Solamente se necesita combinar la idea expresada por la palabra “agregado” y el concepto A , en

la manera expresada por las palabras “el agregado de todo A ”. Esta simple observación, cuya corrección confío que será evidente para todos, elimina todas las dificultades planteadas contra la idea de un conjunto que comprende infinitos miembros.

En otras palabras, lo que Bolzano afirma es que dada cualquier propiedad existe un conjunto cuyos miembros son justamente aquellos objetos que tienen dicha propiedad. Bolzano se propone establecer un criterio de comparación para conjuntos infinitos. La paradoja de la reflexividad no le preocupa tanto como a Galileo; al contrario, el hecho de que pueda establecerse una biyección entre un conjunto y una parte de él le parece “una de las más notables características de los conjuntos infinitos”. Pero en este punto crucial Bolzano no eligió el criterio adecuado.

... el conjunto de todas las cantidades entre 0 y 5 (o menores que 5) es claramente infinito, al igual que lo es el conjunto de todas las cantidades menores que 12. Con no menos seguridad es el último conjunto mayor que el primero, pues el primero constituye solamente una parte del último [...]. Pero no menos cierto que todo esto es lo siguiente: si x representa una cantidad arbitraria entre 0 y 5, y si fijamos la razón entre x e y por la ecuación $5y = 12x$, entonces y es una cantidad entre 0 y 12; y recíprocamente, siempre que y esté entre 0 y 12, x está entre 0 y 5.

Es decir, Bolzano afirma que la aplicación dada por $y = \frac{12}{5}x$ para $x \in [0, 5]$ establece una biyección entre dicho intervalo y el intervalo $[0, 12]$. Pero, cuando se trata de conjuntos infinitos, a Bolzano no le parece que la existencia de una biyección sea criterio suficiente para afirmar que ambos conjuntos son “equinumerosos” y elige como criterio de comparación la relación de inclusión entre conjuntos. De esta forma puede comparar conjuntos infinitos pero no puede *cuantificar* el infinito y, por tanto, no logra desarrollar, pese a su intento, una aritmética del infinito.

Le estaba reservada a [Georg Cantor](#) (1845 - 1918) la gloria de ser el primer matemático que domesticara el infinito. Cantor se vio obligado a defender constantemente sus innovadoras ideas en contra de las opiniones de influyentes matemáticos de su tiempo, alguno de los cuales, como Leopold Kronecker, pasó incluso del ataque científico al ataque personal, si bien otros destacados matemáticos como Weierstrass, Dedekind o Hilbert estuvieron de su parte.

El interés de Cantor por los conjuntos infinitos de puntos y la naturaleza del continuo procede de sus tempranos trabajos en series trigonométricas. En 1869, recién llegado a la universidad de Halle, Cantor, a propuesta de su colega [H.E. Heine](#) (1821-1881), se interesó por el problema sobre la unicidad de la representación de una función como una serie trigonométrica. La interpretación dada hasta entonces a dicha representación se basaba en gran parte en que se suponía que era única, y que de hecho la serie trigonométrica era la serie de Fourier de la función. Esta creencia fue puesta en entredicho porque Weierstrass había probado que para integrar término a término una serie se requiere convergencia uniforme.

Consideremos una serie trigonométrica

$$\frac{a_0}{2} + \sum (a_n \cos(2n\pi x) + b_n \operatorname{sen}(2n\pi x))$$

donde por comodidad consideramos funciones de período 1. El problema de la unicidad equivale a probar que si dicha serie converge puntualmente a cero en $[0, 1]$ entonces sus coeficientes son todos nulos. Esto fue probado por Cantor en 1870, y en 1872 generalizó este resultado. Para ello introduce el concepto de punto de acumulación de un conjunto U de números reales, el conjunto de todos los puntos de acumulación de U lo llamó *conjunto derivado* de U y lo representó por U' . Cantor define los sucesivos conjuntos derivados $U^1 = U'$, $U^{n+1} = (U^n)'$, y demuestra que si se supone que la serie converge puntualmente y con suma cero en un conjunto de la forma $[0, 1] \setminus U$, donde U es un conjunto tal que para algún $n \in \mathbb{N}$ se tiene que $U^n = \emptyset$, entonces los coeficientes de la serie son todos nulos. Probaba así que el resultado de unicidad permanecía válido incluso si se admitía un conjunto infinito de puntos excepcionales en los que no se supone que haya convergencia siempre que estos puntos estén distribuidos de manera apropiada. En este artículo de 1872, Cantor sentó las bases de la teoría de los conjuntos de puntos y se convenció de que el estudio de los conjuntos infinitos era esencial para clasificar los posibles conjuntos de puntos excepcionales, problema que todavía sigue abierto.

Al principio de este trabajo que estamos comentando, Cantor desarrolló una teoría de los números reales basada en sucesiones de números racionales. Ese mismo año, un poco antes, Dedekind había publicado su teoría de las cortaduras. No es ésta la única ocasión en que coinciden los intereses de Cantor y Dedekind. De hecho, la contribución de Dedekind a la creación de la teoría de conjuntos es mucho más importante de lo que suele reconocerse. En su famoso trabajo *Was sind und was sollen die Zahlen* (*¿Qué son y para qué sirven los números?*) publicado en 1888, Dedekind precisa el significado de las operaciones elementales de la teoría de conjuntos *ingenua*, y da la definición general de función entre conjuntos abstractos, generalizando así la anteriormente dada por Dirichlet para funciones reales. Así mismo Dedekind da la siguiente definición:

Un sistema S se llama *infinito* cuando es semejante a una parte propia de sí mismo; en caso contrario, se dice que S es un sistema finito.

En términos actuales: un conjunto S es infinito, si hay un subconjunto propio, $\emptyset \subsetneq A \subsetneq S$, y una biyección de A sobre S . En una nota a pie de página, Dedekind, afirma haber comunicado esa definición a Cantor ya en 1882 y varios años antes a otros colegas. También fue Dedekind un precursor de las técnicas conjuntistas en Álgebra, introduciendo, entre otros, los conceptos de *cuerpo*, *ideal* y *módulo*.



Figura 35. Cantor

En una carta a Dedekind, de fecha 29 de noviembre de 1873, Cantor afirmaba, sin incluir prueba alguna, que los racionales positivos y, más generalmente, el conjunto de las sucesiones finitas de enteros positivos, podía ponerse en correspondencia biyectiva con los enteros positivos, y preguntaba si eso mismo se podía hacer con los números reales. Dedekind le respondió, a vuelta de correo, que en su opinión nada se oponía a ello, y añadió, con demostración incluida, que el conjunto de los números algebraicos sí es biyectivo con el de los enteros positivos.

Definición. Los *números algebraicos* son números, reales o complejos, que son raíces de alguna ecuación polinómica con coeficientes enteros. Por tanto, un número real o complejo x es algebraico si hay números enteros $c_k \in \mathbb{Z}$, ($k = 0, 1, 2, \dots, n$) tal que x satisface la ecuación polinómica

$$c_0 + c_1x + c_2x^2 + \dots + c_nx^n = 0$$

Los números que no son algebraicos se llaman *trascendentes*. Todo número racional es evidentemente algebraico, pero también lo son las raíces de cualquier orden de números racionales positivos y muchos más. Intuitivamente, los números algebraicos son los que pueden obtenerse a partir de los enteros por procedimientos algebraicos: suma, producto, cociente, división, raíces, iterados un número finito cualquiera de veces. En ese sentido podemos decir que los números algebraicos no están “muy alejados” de los enteros. Los números trascendentes son justamente lo contrario: son números irracionales “muy alejados” de los enteros.

Hasta 1844 siguió abierta la cuestión sobre si había o no irracionales trascendentes. Ese año, Liouville mostró que cualquier número de la forma

$$\frac{a_1}{10} + \frac{a_2}{10^{2!}} + \frac{a_3}{10^{3!}} + \dots + \frac{a_n}{10^{n!}} + \dots$$

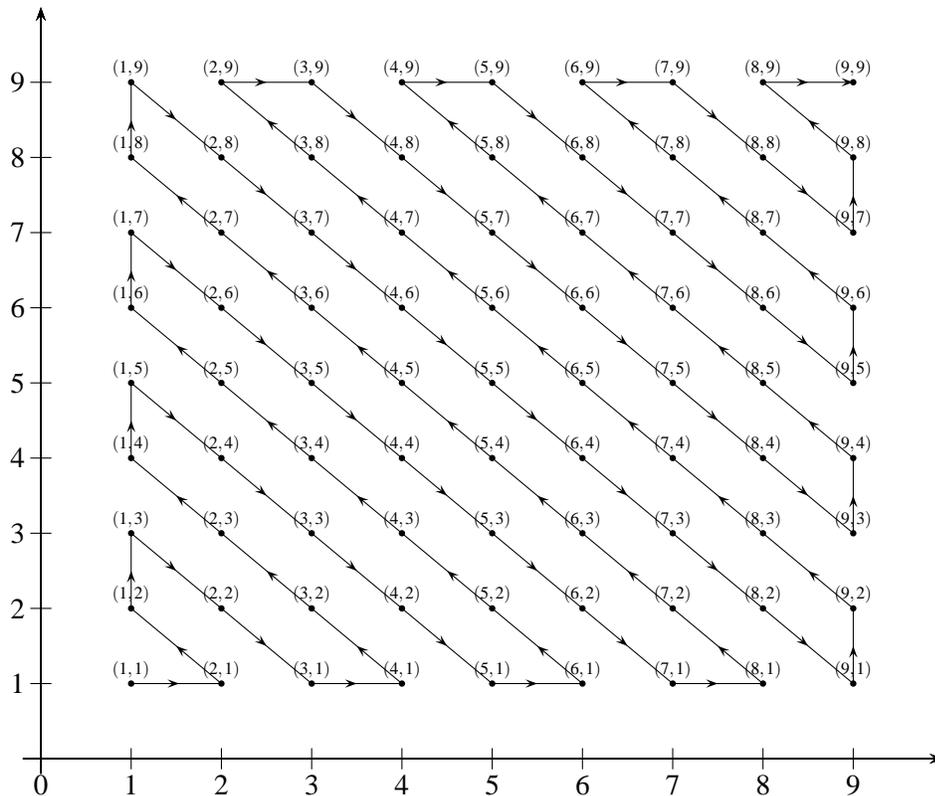
donde para todo $n \in \mathbb{N}$ es $a_n \in \{0, 1, 2, 3, 4, 5, 6, 7, 8, 9\}$, es trascendente.

demostrar que un número concreto es trascendente es difícil. La trascendencia del número e fue demostrada por Charles Hermite en 1873, y Ferdinand Lindemann logró probar la trascendencia de π en 1882 (demostrando así que el problema de la cuadratura del círculo no tenía solución).

Para facilitar la exposición que sigue voy a dar algunas definiciones de conceptos introducidos por Cantor años más tarde.

Definición. Se dice que dos conjuntos A y B son *equipotentes* si existe una aplicación biyectiva de uno de ellos sobre el otro. Un conjunto se llama *numerable* si es vacío o si existe una aplicación inyectiva de él en el conjunto de los enteros positivos.

Es fácil probar que los conjuntos numerables infinitos son precisamente los conjuntos equipotentes al conjunto \mathbb{N} de los números naturales. Por tanto, los conjuntos numerables son aquellos conjuntos cuyos elementos se pueden contar ¡aunque sean infinitos! El resultado, citado por Cantor, de que \mathbb{Q} es numerable, no deja de ser muy sorprendente y contrario a la intuición, pues si $r < s$ son números racionales cualesquiera, entre ellos dos hay siempre infinitos números racionales. Pese a ello, no hay más números racionales que números naturales. La siguiente figura indica una forma de enumerar todos los racionales positivos.



En cada diagonal se cuentan las fracciones p/q tales que $p + q$ es constante. Es claro que así se cuentan todas las fracciones positivas.

Poco después de las cartas citadas, Cantor logró demostrar que el conjunto de los números reales no es numerable. Merece la pena recordar la demostración.

Teorema. Dados dos números reales $a < b$ se verifica que el intervalo $[a, b]$ no es numerable.

Demostración. Si $[a, b]$ fuera numerable tendría que ser equipotente a \mathbb{N} . Veamos que esto no puede ocurrir. Supongamos que $\varphi : \mathbb{N} \rightarrow [a, b]$ es una biyección de \mathbb{N} sobre $[a, b]$. En particular φ es sobreyectiva por lo que deberá ser $[a, b] = \{\varphi(n) : n \in \mathbb{N}\}$. Obtendremos una contradicción probando que tiene que existir algún elemento $z \in [a, b]$ tal que $z \notin \{\varphi(n) : n \in \mathbb{N}\}$.

Para ello se procede de la siguiente forma. Dividimos el intervalo $[a, b]$ en tres intervalos cerrados de igual longitud:

$$\left[a, a + \frac{b-a}{3} \right], \left[a + \frac{b-a}{3}, b - \frac{b-a}{3} \right], \left[b - \frac{b-a}{3}, b \right]$$

y llamamos I_1 al primero de ellos (es decir el que está más a la izquierda) que no contiene a $\varphi(1)$. Dividimos ahora el intervalo I_1 en tres intervalos cerrados de igual longitud y llamemos I_2 al primero de ellos que no contiene a $\varphi(2)$.

Este proceso puede “continuarse indefinidamente” pues, supuesto que $n \in \mathbb{N}$, $n \geq 2$, y que tenemos intervalos cerrados de longitud *positiva* I_k , $1 \leq k \leq n$, tales que $I_{k+1} \subset I_k$ para $1 \leq k \leq n-1$, y $\varphi(k) \notin I_k$ para $1 \leq k \leq n$, dividimos el intervalo I_n en tres intervalos cerrados de igual longitud y llamamos I_{n+1} al primero de ellos que no contiene a $\varphi(n+1)$. De esta forma para cada $n \in \mathbb{N}$ tenemos un intervalo cerrado I_n no vacío verificándose que $I_{n+1} \subset I_n$ y $\varphi(n) \notin I_n$ para todo $n \in \mathbb{N}$. El principio de los intervalos encajados nos dice que hay algún número real z que está en *todos* los I_n . Por tanto, cualquiera sea $n \in \mathbb{N}$, por ser $z \in I_n$ y $\varphi(n) \notin I_n$, se tiene necesariamente que $z \neq \varphi(n)$, esto es, $z \notin \{\varphi(n) : n \in \mathbb{N}\}$ pero, evidentemente, $z \in [a, b]$. \square

¿Te recuerda algo la demostración anterior? ¿Quizás a la divisibilidad infinita del continuo? Pues claro, lo que estamos haciendo es dividir infinitas veces un segmento (el prototipo de *continuo*). Lo que nos dice este resultado es que, aunque lo dividamos en un infinito actual de partes, siempre nos quedarán puntos que no habremos tocado. Aristóteles afirmaba que un continuo puede dividirse en cualquier parte pero no en todas partes: hay que darle la razón en este punto.

Es fácil probar que cualquier intervalo no reducido a un punto es equipotente a \mathbb{R} . El teorema anterior demuestra que “hay muchos más números irracionales que racionales” pues mientras que podemos enumerar los racionales no podemos hacer lo mismo con los irracionales ya que no hay biyecciones de \mathbb{N} sobre $\mathbb{R} \setminus \mathbb{Q}$.

Deducimos también la siguiente estrategia *para probar que un conjunto $A \subset \mathbb{R}$ no es vacío es suficiente probar que su complemento $\mathbb{R} \setminus A$ es numerable* (¡con lo cual, de hecho, estamos probando que A es infinito no numerable!).

Puesto que los números algebraicos son un conjunto numerable, deducimos, de forma no constructiva ¡sin dar ni un solo ejemplo!, que en todo intervalo de \mathbb{R} , hay infinitos números trascendentes. Compara este razonamiento con el tradicional constructivo seguido años antes por Liouville para probar que hay números trascendentes. Cantor publicó estos resultados, el suyo y el de Dedekind, en un trabajo de tres páginas titulado *Über eine Eigenschaft des Inbegriffes aller reellen algebraischen Zahlen (Sobre una propiedad del sistema de todos los números algebraicos reales)* (1874). Es muy llamativo que el título de este trabajo, considerado como el nacimiento oficial de la teoría de conjuntos, no haga referencia alguna al resultado que hoy consideramos como el principal: la no numerabilidad de \mathbb{R} . Además la propia presentación del trabajo elude destacar

estos resultados. Posiblemente, Cantor temía la reacción que pudiera provocar un trabajo tan radicalmente innovador. Porque lo que él hacía era probar que en cualquier intervalo $[a, b] \subset \mathbb{R}$ con $a < b$ hay, en un sentido matemático preciso, más números que todos los números algebraicos juntos, de donde se deducía que en $[a, b]$ tenía que haber números trascendentes. Esta es una demostración de *existencia pura*, algo nuevo en las matemáticas.

Naturalmente, Cantor sabía muy bien que había descubierto una propiedad específica del continuo: su no numerabilidad. Disponía ya de dos tipos de conjuntos infinitos: \mathbb{N} y \mathbb{R} , claramente \mathbb{N} tenía un tamaño más pequeño que \mathbb{R} . Precisar esa idea de tamaño y elaborar una teoría de comparación de conjuntos infinitos es lo que hizo Cantor en los siguientes veinte años y, casi contra su voluntad, se vio llevado a desarrollar la teoría de números transfinitos y la teoría de conjuntos como una disciplina matemática independiente.

En 1877, Cantor probó, para su propia sorpresa, que los puntos del plano podían ponerse en correspondencia biyectiva con \mathbb{R} , y, más general, que los espacios \mathbb{R}^n son todos ellos biyectivos a la recta real. Este resultado fue de los que más desconcierto provocó entre los matemáticos contemporáneos. En 1883, en su trabajo *Fundamentos de una teoría general de conjuntos*, escribe:

La presentación de mis investigaciones hasta la fecha en teoría de conjuntos, ha alcanzado un punto donde su progreso depende de una extensión del concepto de número entero más allá de sus límites actuales. Esta extensión señala en una dirección que, por lo que yo sé, no ha sido investigada por nadie todavía.

[...] Por atrevido que esto pueda parecer, tengo que expresar, no sólo la esperanza, sino también la firme convicción de que esta extensión tendrá que ser considerada con el tiempo como absolutamente simple, adecuada y natural. Pero no se me oculta de ninguna manera el hecho de que en esta empresa me encuentro situado en una cierta oposición a concepciones muy extendidas acerca del infinito matemático, y a opiniones formuladas frecuentemente sobre la naturaleza del número.

En este trabajo Cantor introduce los *números transfinitos* o *cardinales transfinitos*. Por el mismo proceso que podemos abstraer la idea del número 5 como la clase de todos los conjuntos equipotentes a un conjunto cualquiera con cinco elementos, a, b, c, d, e , de la misma forma este proceso permite, dado un conjunto M , por doble abstracción de la naturaleza de sus elementos y del posible orden en que estén dados, asociar a M un objeto matemático, representado por $\aleph M$, que se llama *su número cardinal* o *potencia*, que es el mismo para todos los conjuntos equipotentes a M . Cuando M es finito, $\aleph M$ es el número de elementos de M ; la potencia de los conjuntos numerables (infinitos) la representó Cantor por \aleph_0 (\aleph es la primera letra del alfabeto hebreo, se pronuncia “alef”); la potencia de la recta real y de cualquier intervalo de la misma, no vacío y no reducido a un punto, se representa por c y se llama la *potencia del continuo*.

Cantor define una relación de orden entre números cardinales: si M, N son dos conjuntos, diremos que $\aleph M \preceq \aleph N$ si existe una biyección de M sobre una parte de N . Si, además, no existe ninguna biyección entre ninguna parte de M y la totalidad de N , se escribe $\aleph M \prec \aleph N$. Con esta definición se tiene que $\aleph_0 \prec c$. Para números cardinales finitos esta relación de orden es la usual. La demostración de que \preceq es una relación

de orden entre números cardinales está muy lejos de ser fácil. La dificultad estaba en probar la propiedad reflexiva, es decir, si $\aleph M \preccurlyeq \aleph N$ y también $\aleph N \preccurlyeq \aleph M$, entonces es $\aleph M = \aleph N$. Este resultado fue probado en 1898, y se conoce como teorema de Cantor - Bernstein. Se verifica, además, que \preccurlyeq es una relación de orden total, es decir, dados conjuntos M y N se verifica alguna de las relaciones $\aleph M \preccurlyeq \aleph N$ o $\aleph N \preccurlyeq \aleph M$. La demostración de este resultado exige usar el axioma de elección, también llamado axioma de Zermelo.

Todos esto está muy bien, pero ¿cuántos números cardinales infinitos hay? Hasta ahora solamente conocemos dos. Cantor ideó un procedimiento por el cual, dado un conjunto M , se puede construir un conjunto cuyo cardinal es estrictamente mayor. Para ello, definió el conjunto $\mathcal{P}(M)$ como el conjunto cuyos elementos son todos los subconjuntos de M

$$\mathcal{P}(M) = \{A : A \subset M\}$$

Teorema. No existe una biyección entre M y $\mathcal{P}(M)$. Es decir, se verifica que:

$$\aleph M \prec \aleph \mathcal{P}(M).$$

Demostración. Supongamos, para obtener una contradicción, que $\psi : M \rightarrow \mathcal{P}(M)$ es una biyección. Sea $A = \{x \in M : x \notin \psi(x)\} \in \mathcal{P}(M)$. Como ψ es una biyección existirá un $a \in M$ tal que $\psi(a) = A$. Y tenemos que: $a \in \psi(a)$ si, y sólo si, $a \notin \psi(a)$. □

Suele escribirse $\aleph \mathcal{P}(M) = 2^{\aleph M}$, igualdad que, para el caso de conjuntos finitos, es conocida.

Por tanto, los conjuntos

$$\mathcal{P}(M), \mathcal{P}(\mathcal{P}(M)), \mathcal{P}(\mathcal{P}(\mathcal{P}(M))) \dots$$

tienen todos ellos distinto número cardinal.

Cantor siguió desarrollando sus ideas en una serie de seis trabajos publicados en los años 1878 a 1884. El desarrollo de la teoría de conjuntos condujo a algunas contradicciones, las llamadas *paradojas de la teoría de conjuntos*. Ello era debido al punto de vista ingenuo adoptado respecto a los conjuntos. Se pensaba que cualquier *propiedad matemática*, $P(x)$, definía su correspondiente conjunto, a saber, el formado por los elementos para los cuales dicha propiedad es verdadera. Por “propiedad matemática”, $P(x)$, se entendía cualquier expresión simbólica bien formada que tenga libre la variable x . Esto se conoce como *axioma de abstracción*. La primera formulación explícita del mismo fue hecha por [Gottlob Frege \(1848-1925\)](#) en 1893 en su obra *Grundgesetze der Arithmetik* (“Leyes Básicas de la Aritmética”). En 1902, estando en imprenta el segundo volumen de la obra citada, recibió una carta de [Bertrand Russell \(1872-1970\)](#) en la que mostraba la siguiente contradicción derivada de dicho axioma. Consideremos la siguiente propiedad $P(x) = x \notin x$ y definamos el conjunto $W = \{x : x \notin x\}$. Entonces resulta que si $W \in W$ es porque $W \notin W$ y si $W \notin W$ debe ser $W \in W$. Una contradicción insalvable, conocida como la *paradoja de Russell*. Otra paradoja, conocida como *paradoja de Cantor*, resulta de la consideración del *conjunto de todos los conjuntos*, pues si llamamos $U = \{x : x = x\}$ a

dicho conjunto universal que contiene a cualquier conjunto, entonces $\mathcal{P}(U) \subset U$ por lo que $\# \mathcal{P}(U) \preccurlyeq \# U$ lo que contradice el resultado antes visto. Estas paradojas son de tipo lógico.

Surgieron otras paradojas de tipo semántico y se derivan de la capacidad del lenguaje para referirse a sí mismo. Una de las más conocidas es la de Berry (1906): Existen muchas expresiones en castellano que permiten describir números naturales, cada una de ellas consta de un número finito de caracteres gramaticales: letras, comas,... Considera la expresión: “*sea x el menor número natural que no puede ser descrito con menos de cien caracteres*”. Resulta que dicha expresión describe con menos de cien caracteres a un número que no puede ser así descrito.

La solución fue axiomatizar la teoría de conjuntos para evitar que pudieran formularse paradojas como las anteriores lo que, para evitar las paradojas de tipo semántico, obliga al uso de un lenguaje formalizado y, para evitar las de tipo lógico, a restringir de alguna forma la existencia de conjuntos “demasiado grandes” como los conjuntos W y U anteriores. Observa que el axioma de abstracción es en realidad un *esquema axiomático* ya que para cada elección de la propiedad $P(x)$ resulta un axioma afirmando la existencia de un cierto conjunto. [Zermelo \(1871-1953\)](#) propuso modificar dicho axioma debilitándolo de la manera siguiente:

Para toda propiedad $P(x)$ definible en la teoría y para todo conjunto A , existe un conjunto B cuyos elementos son exactamente aquellos $x \in A$ que cumplen $P(x)$.

Es decir, Zermelo postula la existencia del conjunto $B = \{x \in A : P(x)\}$.

Este esquema axiomático de Zermelo permite obtener conjuntos a partir de otros, y cuyo tamaño es menor que el de aquellos de los que han sido obtenidos. Esto implica que, necesariamente, para formar un conjunto debemos partir de otro previamente dado. Por ello hay que añadir otros axiomas que garantizan la existencia de aquellos conjuntos necesarios que no pueden obtenerse como subconjuntos de otros conjuntos dados. Esto ha dado lugar a la [teoría axiomática de Zermelo-Fraenkel](#).

En la teoría de Zermelo-Fraenkel, la no existencia de un conjunto W tal que $x \in W$ si, y sólo si $x \notin x$ es un *teorema*. Pues si un tal conjunto existiera se llega a la contradicción de que $W \in W$ si, y sólo si $W \notin W$. Observa que ni siquiera es posible considerar en esta teoría algo como $W = \{x : x \notin x\}$. Por tanto, tampoco existe en esta teoría un conjunto universal U , ya que si tal conjunto existiera también existiría el conjunto $W = \{x \in U : x \notin x\}$, lo que acabamos de ver que es contradictorio.

Observa que no podemos definir el “complementario absoluto” de un conjunto A , es decir, no existe un conjunto A^c cuyos elementos sean los conjuntos que no pertenecen a A . Si existiera, entonces también existiría el conjunto $U = A \cup A^c$ cuya existencia acabamos de ver que lleva a contradicción.

Las operaciones con números transfinitos se definen con facilidad por medio de las correspondientes operaciones conjuntistas. Por ejemplo, el producto $\#M \cdot \#N$ es, por definición, igual a $\#(M \times N)$ donde $M \times N$ es el conjunto producto cartesiano de M y N . Análogamente se define la suma $\#M + \#N$ como el número

cardinal de la unión disjunta de M y N . Estas operaciones son asociativas, conmutativas y distributivas pero, para cardinales transfinitos se cumple que

$$\aleph_M + \aleph_N = \aleph_M \cdot \aleph_N = \max\{\aleph_M, \aleph_N\}$$

Esto es, la aritmética transfinita no responde a las reglas usuales de la aritmética finita. Pero esto no quiere decir que sea contradictoria, simplemente, es diferente.

El trabajo de Cantor cambió la forma de ver las matemáticas y acabó por ser ampliamente aceptado. La visión que Cantor tenía de las matemáticas puras es muy hermosa; para él, las matemáticas puras son el reino de la libertad y las llamaba “matemáticas libres”, porque son una creación de la libertad del espíritu humano cuyas únicas limitaciones son la coherencia y la no contradicción.

7. Los orígenes del Análisis Funcional

Los orígenes del Análisis Funcional están estrechamente relacionados con la teoría de ecuaciones integrales y los sistemas de infinitas ecuaciones lineales, las ecuaciones diferenciales y el cálculo de variaciones, el desarrollo de los conceptos topológicos, la teoría conjuntista y la evolución del álgebra “moderna”. El Análisis Funcional nació en el primer tercio del siglo XX y sus orígenes están en los trabajos de David Hilbert y Ivar Fredholm en ecuaciones integrales y teoría espectral de operadores, los trabajos de Henri Lebesgue, Maurice Fréchet y Frigyes Riesz en teoría de la medida y espacios abstractos, y los trabajos de Eduard Helly, Hans Hahn y Stefan Banach sobre la teoría de dualidad. Todos estos trabajos aparecen entre los años 1900 y 1932, fecha esta última que, con la publicación del libro de Stefan Banach *Théorie des opérations linéaires*, es considerada la fecha de nacimiento del Análisis Funcional como materia de estudio por derecho propio debido a su gran aplicabilidad a multitud de problemas en diversos campos. El siguiente es un breve resumen de fechas importantes en el desarrollo de todas estas ideas:

- 1900 trabajo de Erik Ivar Fredholm sobre ecuaciones integrales.
- 1902 tesis de Henri Lebesgue sobre integración.
- 1906 trabajo de David Hilbert sobre teoría espectral.
- 1906 tesis de Maurice Fréchet sobre espacios métricos.
- 1910 y 1911 trabajos de Frigyes Riesz sobre $C[a, b]$ y L_p .
- 1912 y 1921 trabajos de Eduard Helly.
- 1922 Tesis de Stefan Banach sobre espacios normados.

- 1927 trabajo de Hans Hahn y 1929 trabajo de Banach sobre dualidad; 1927 trabajo de Banach y Hugo Steinhaus.
- 1928 libro de Fréchet *Les espaces abstraits*, y 1932 libro de Banach *Théorie des opérations linéaires*.

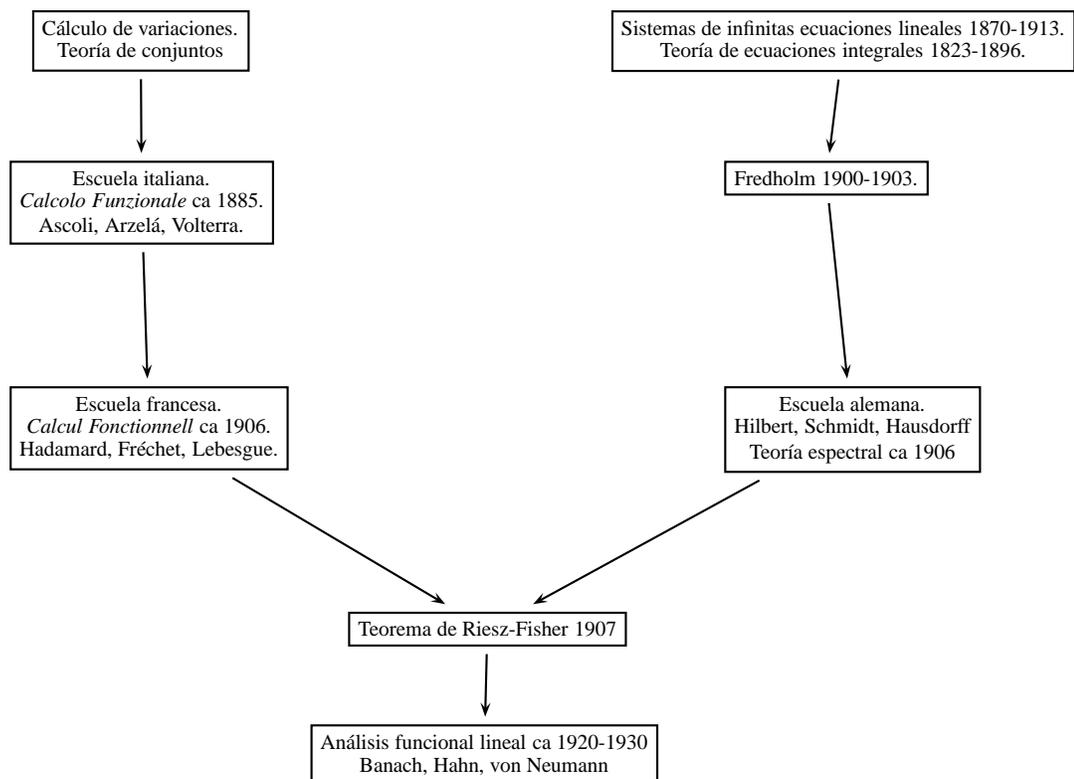


Figura 36. Los orígenes del Análisis Funcional

En los años que van de 1900 a 1932, los matemáticos descubrieron que problemas procedentes de diferentes campos compartían una serie de características comunes que permitían abordarlos en un contexto unificado, aunque para ello era necesario omitir ciertos detalles no esenciales. Naturalmente, esto condujo a un planteamiento cada vez más abstracto de los problemas, a la consideración de espacios de funciones de diversos tipos y, finalmente, al nacimiento de los *espacios abstractos* en los que hay una estructura algebraica que satisface ciertos axiomas compatible con una topología, pero en los que no se especifica para nada la naturaleza de sus elementos. Una clara ventaja de este punto de vista abstracto, es que los resultados obtenidos pueden aplicarse en cualquier contexto que satisfaga los axiomas de la teoría.

El desarrollo del Análisis Funcional marca un hito en la tendencia de las matemáticas, dominante desde principios del siglo XX, hacia la generalización y unificación, la axiomatización y la abstracción. Un aspecto

central en este camino es la emergencia del concepto fundamental de *espacio funcional*, esto es, un espacio dotado de cierta estructura algebraica y topológica cuyos *puntos* son funciones. Aunque parezca anecdótico, hubo que recorrer un largo camino para representar una función con una sola letra, digamos f , en vez de referirse a sus valores $f(x)$. Se necesitó mucho tiempo para dejar de ver una ecuación diferencial o integral solamente como eso, como una simple *ecuación* que hay que resolver, y empezar a considerarla como un *operador* entre espacios de funciones. Lo característico del Análisis Funcional es la consideración de espacios funcionales donde las propiedades individuales de las funciones no son objeto de estudio sino las propiedades estructurales de tales espacios en sus aspectos algebraicos y topológicos.

7.1. Sistemas de infinitas ecuaciones lineales

La técnica de solución de ecuaciones diferenciales por el *método de los coeficientes indeterminados*, muy usada en el siglo XVIII, conducía a sistemas de infinitas ecuaciones lineales con infinitas incógnitas. Esta técnica consiste en suponer la existencia de una solución dada como suma de una serie de potencias cuyos coeficientes hay que calcular, sustituir dicha serie en la ecuación diferencial e identificar los coeficientes de las mismas potencias lo que daba lugar a un sistema de infinitas ecuaciones lineales. Aunque los coeficientes a calcular eran infinitos, cada una de las ecuaciones lineales del sistema tenía solamente un número finito de los mismos lo que, si todo iba bien, permitía obtener una *relación recurrente* entre los coeficientes de la serie que permitía resolver el sistema de infinitas ecuaciones por técnicas convencionales para el caso de un número finito de ecuaciones. Puesto que se trataba de una técnica de cálculo, cuyos resultados debían comprobarse en cada caso, no hubo intentos de desarrollar una teoría general de tales sistemas de ecuaciones.

Parece ser que el primer sistema de infinitas ecuaciones lineales no recurrente es considerado por Fourier en su *Théorie analytique de la chaleur* (1822). Se trata de encontrar una solución de la ecuación diferencial

$$\frac{\partial^2 u}{\partial x^2}(x, y) + \frac{\partial^2 u}{\partial y^2}(x, y) = 0, \quad x > 0, \quad -\frac{\pi}{2} < y < \frac{\pi}{2}$$

con las condiciones de contorno $u(0, y) = 1$ para $-\pi/2 < y < \pi/2$, $u(x, -\pi/2) = u(x, \pi/2) = 0$ para $x > 0$, $\lim_{x \rightarrow +\infty} u(x, y) = 0$. Se trata de un modelo matemático de la temperatura estacionaria en el interior de una placa infinita de forma rectangular, cuyos bordes se mantienen a la temperatura prefijada.

Fourier, con su habitual método de separación de variables, pone $u(x, y) = f(x)g(y)$ y calcula f y g sin tener en cuenta las condiciones de contorno. Obtiene así la solución

$$u(x, y) = e^{-nx} \cos(ny)$$

donde n es cualquier valor real. Para que se cumplan las tres últimas condiciones de contorno n debe ser un entero positivo impar. Puesto que se trata de un problema lineal, Fourier usa el *principio de superposición* y

considera la solución dada por la serie:

$$u(x, y) = \sum_{n=1}^{\infty} c_n e^{-(2n-1)x} \cos(2n-1)y$$

donde los coeficientes c_n deben ser determinados de forma que

$$1 = \sum_{n=1}^{\infty} c_n \cos(2n-1)y, \quad -\frac{\pi}{2} < y < \frac{\pi}{2}$$

Ahora Fourier deriva sucesivamente la serie y sustituye $y = 0$ obteniendo el sistema de ecuaciones

$$\begin{aligned} 1 &= \sum_{n=1}^{\infty} c_n \\ 0 &= \sum_{n=1}^{\infty} (2n-1)^2 c_n \\ 0 &= \sum_{n=1}^{\infty} (2n-1)^4 c_n \\ \dots &\dots \dots \dots \end{aligned}$$

donde las incógnitas son los coeficientes c_n . Para resolver este sistema, Fourier, considera solamente las primeras k ecuaciones y k incógnitas. Un cálculo elemental proporciona las soluciones $c_1^{(k)}, c_2^{(k)}, \dots, c_k^{(k)}$ y, finalmente obtiene:

$$c_n = \lim_{k \rightarrow \infty} c_n^{(k)} = \frac{4(-1)^{n-1}}{\pi(2n-1)}$$

Con lo que queda determinada la solución $u(x, y)$ y además:

$$\frac{\pi}{4} = \sum_{n=1}^{\infty} \frac{(-1)^{n-1}}{2n-1} \cos(2n-1)y, \quad -\frac{\pi}{2} < y < \frac{\pi}{2}$$

Seguidamente, Fourier, comprueba, por cálculos directos, que sus resultados son correctos. Aunque los resultados de Fourier son correctos, sus razonamientos están lejos de serlo, de hecho cuando se sustituye el valor calculado para c_n en el sistema lineal, las series resultantes, a partir de la segunda en adelante, no son convergentes.

Esta forma de proceder de Fourier, sustituyendo un sistema de infinitas ecuaciones lineales con infinitas incógnitas, $\{x_k : k \in \mathbb{N}\}$, por sus primeras n ecuaciones lineales con las incógnitas, $\{x_k : 1 \leq k \leq n\}$, calcular las soluciones del mismo, $\{x_k^{(n)} : 1 \leq k \leq n\}$, y hacer tender n a infinito, es lo que Riesz ([19]) llama *principio de las reducidas*. Este principio puede considerarse como un paso “de lo finito a lo infinito” y tuvo un papel importante en la teoría de tales sistemas de infinitas ecuaciones. En la teoría de las ecuaciones integrales veremos una variante del mismo que puede considerarse como el “paso de lo discreto a lo continuo”.

Que el principio de las reducidas no siempre proporciona soluciones queda claro con el siguiente ejemplo

de Helly (1921). El sistema

$$\begin{array}{rcl}
 1 & = & x_1 + x_2 + x_3 + x_4 + \cdots \\
 1 & = & \quad x_2 + x_3 + x_4 + \cdots \\
 1 & = & \quad \quad x_3 + x_4 + \cdots \\
 1 & = & \quad \quad \quad x_4 + \cdots \\
 \dots & \dots & \dots \quad \dots \quad \dots
 \end{array}$$

No tiene solución porque restando a cada ecuación la que le sigue se deduce que debe ser $x_n = 0$ para todo $n \in \mathbb{N}$. Sin embargo cada sistema reducido tiene como solución $x_n = 1, x_k = 0$ para $k = 1, 2, \dots, n - 1$.

El principio de las reducidas de Fourier pasó desapercibido, y tiene que pasar casi medio siglo para que a partir de 1870 algunos matemáticos se interesen por los sistemas de infinitas ecuaciones lineales. En el último tercio del siglo XIX se desarrolla una teoría de determinantes infinitos que se aplica para resolver sistemas de infinitas ecuaciones lineales en casos particulares. Los detalles de esta historia se exponen en los dos primeros capítulos de [19]. Para aplicar el método clásico de los determinantes a los sistemas de infinitas ecuaciones lineales era necesario imponer condiciones más o menos restrictivas pero, como afirma Riesz, tales restricciones venían impuestas por el método más que por el problema en sí. En 1913 F. Riesz publicó su memoria *Les systèmes d'équations linéaires à une infinité d'inconnues* en la que desarrolla una teoría satisfactoria de tales sistemas sin usar el principio de las reducidas ni determinantes infinitos. Volveremos sobre esto más adelante.

7.2. La escuela italiana: Ascolí, Arzelá, Volterra

En 1696 Johann Bernoulli desafió a “los más brillantes matemáticos del mundo” a resolver el problema de la *braquistócrona*, o curva de más rápido descenso. Newton, Leibniz, L'Hôpital y los hermanos Johann y Jakob Bernoulli encontraron que dicha curva es la cicloide. Todos ellos dieron sus soluciones en términos geométricos o mecánicos. Nosotros sabemos que la curva buscada $y = y(x), a \leq x \leq b$, es la que hace mínima la integral:

$$\int_a^b \frac{\sqrt{1+y'(x)^2}}{\sqrt{y(x)}} dx$$

Previamente, en 1687, Newton había estudiado el problema del sólido de revolución que experimenta una mínima resistencia cuando se mueve a través de un fluido homogéneo con velocidad constante en la dirección de su eje de revolución. Si es $y = y(x), a \leq x \leq b$, la curva generatriz de dicho sólido, Newton afirma que la solución buscada hace mínima la integral:

$$\int_a^b \frac{y(x)y'(x)^3}{1+y'(x)^2} dx$$

Otro problema del mismo tipo es el siguiente. Calcular entre todas las curvas $y = y(x)$ satisfaciendo $y(a) = 0$, $y(b) = 1$ la que al girar su gráfica alrededor del eje de abscisas genera una superficie de área mínima. Es decir, queremos hacer mínima la cantidad:

$$\int_a^b y(x) \sqrt{1 + y'(x)^2} dx$$

Problemas *isoperimétricos* o de cálculo de *geodésicas* o *superficies minimales* también responden al mismo modelo: calcular la curva o superficie que maximiza o minimiza los valores que toma una expresión analítica en la que interviene una curva o superficie desconocida. Fue Euler quien en 1756 llamó a estos problemas *cálculo de variaciones*. [Volterra \(1860-1940\)](#) llamó a estas expresiones analíticas “funciones de línea” o “funciones de funciones”. [Hadamard \(1865-1963\)](#) sugirió otro nombre y llamó a estas funciones “especiales” *funcionales* y se refería al cálculo de variaciones como “el análisis de funcionales”. Parece ser que fue [Paul Levy \(1886-1971\)](#) el primero en usar la expresión “análisis funcional” en 1922.

En el Cálculo de Variaciones se trata de maximizar o minimizar una expresión del tipo

$$J(\varphi) = \int_a^b F(x, \varphi(x), \varphi'(x)) dx$$

donde F es una función suficientemente buena y la variable φ pertenece a una cierta familia \mathcal{F} de curvas definidas en el intervalo $[a, b]$. Es en este contexto donde aparece primero la idea de *campo funcional*, como conjunto de funciones admisibles, y la de distancia entre funciones. Se suponía, apoyándose por lo general en razones físicas o geométricas, que la existencia del mínimo $\min\{J(\varphi) : \varphi \in \mathcal{F}\}$ o del máximo estaba garantizada cuando la función F estaba acotada. Pero Weierstrass en 1870 probó con un ejemplo que esto no era así.

En 1887 Volterra publicó varios trabajos en los que investigaba clases especiales de funcionales, definidos como aplicaciones continuas $U : \mathcal{F} \rightarrow \mathbb{R}$, donde \mathcal{F} es una clase de funciones reales definidas en un intervalo $[a, b]$. Ya que estos parecen ser los primeros trabajos en los que se estudian funcionales en general, se considera que 1887 es el año de nacimiento del análisis funcional.

Es sabido que la compacidad garantiza la existencia de extremos de funciones continuas reales. Esto llevó a otro matemático italiano [Cesare Arzelà](#) (1847-1912) a intentar justificar el cálculo de variaciones usando conceptos de compacidad secuencial. Usando el concepto de *equicontinuidad* introducido por [Giulio Ascoli](#) (1843- 1896), probó el que seguramente es uno de los primeros resultados notables de análisis funcional conocido como *teorema de Ascoli-Arzelà* publicado en 1883: *Si una sucesión $\{f_n\}$ de funciones reales definidas en un intervalo $[a, b]$ está uniformemente acotada y es equicontinua entonces tiene alguna sucesión parcial uniformemente convergente*. Este teorema es una generalización del teorema de Bolzano-Weierstrass para sucesiones numéricas acotadas al espacio de dimensión infinita $C[a, b]$.

Puede ser conveniente hacer un poco de historia sobre la convergencia uniforme que es la convergencia

en el espacio $C[a, b]$. Ya vimos cómo Cauchy afirmaba que la suma de una serie convergente de funciones continuas es ella misma una función continua, y Abel replicaba que eso no era cierto. El asunto permaneció así por unos años hasta que en 1847 Stokes (1819-1903), en 1848 Seidel (1821-1896) y en 1853 el propio Cauchy, independientemente, demostraron que la convergencia uniforme es suficiente para asegurar la continuidad de la función límite. No obstante, Weierstrass ya había usado el concepto de convergencia uniforme en algunos manuscritos no publicados de 1841. Ninguno de ellos pudo encontrar una condición necesaria. Fue Arzelá quien pudo dar en 1884 una condición necesaria y suficiente para que la función límite puntual, f , de una sucesión $\{f_n\}$ de funciones continuas en un intervalo $[a, b]$ fuera continua: que la convergencia sea *casi-uniforme*. Es decir que para todo $\varepsilon > 0$, y para todo $p \in \mathbb{N}$ exista un conjunto finito de índices q_1, q_2, \dots, q_n mayores o iguales que p , tales que para cada $t \in [a, b]$:

$$\min \{|f_{q_i}(t) - f(t)| : 1 \leq i \leq n\} < \varepsilon.$$

El Cálculo de Variaciones está estrechamente relacionado con uno de los problemas más importantes del Análisis del siglo XIX: el *Problema de Dirichlet*. Este problema consiste en encontrar una función armónica u en un dominio $\Omega \subset \mathbb{R}^n$ y continua en $\overline{\Omega}$ cuyos valores en la frontera de Ω vienen dados por una función f :

$$\begin{aligned} \Delta u(\mathbf{x}) &= 0 & (\mathbf{x} \in \Omega) \\ u(\mathbf{x}) &= f(\mathbf{x}) & (\mathbf{x} \in \text{Fr}(\Omega)) \end{aligned}$$

donde

$$\Delta u(\mathbf{x}) = \sum_{i=1}^n \frac{\partial^2 u}{\partial x_i^2}(\mathbf{x})$$

es el operador de Laplace.

Es un problema fundamental en muchas áreas de la Matemática y la Física, y los esfuerzos para resolverlo dieron lugar a nuevas ideas y propiciaron el desarrollo de nuevas teorías. En algunos casos simples el problema de Dirichlet puede resolverse en forma explícita. Por ejemplo, la solución del problema de Dirichlet para un disco en \mathbb{R}^2 está dada por la conocida fórmula integral de Poisson. La existencia de solución no siempre está asegurada.

El *principio de Dirichlet* afirma que dicho problema es equivalente a encontrar una función u de clase C^2 en Ω y continua en $\overline{\Omega}$, verificando que $u(\mathbf{x}) = f(\mathbf{x})$ para todo $\mathbf{x} \in \text{Fr}(\Omega)$, y que es entre todas dichas funciones la que hace mínima la integral

$$E(u) = \int_{\Omega} \sum_{i=1}^n \left(\frac{\partial u}{\partial x_i} \right)^2 d\mathbf{x}$$

Este principio convierte un problema de contorno para la ecuación de Laplace en un problema de tipo variacional. Los intentos de probar la existencia y unicidad de la solución del problema de tipo variacional que se hicieron durante el siglo XIX sólo tuvieron éxito para una clase bastante restringida de dominios y las cosas

se complicaron cuando en 1860s Weierstrass demostró, por medio de un ejemplo, que el problema de contorno para la ecuación de Laplace para el caso de una condición de frontera continua podía tener solución y el correspondiente problema variacional no tenerla. Fue Hilbert quien en 1900 logró justificar el principio de Dirichlet bajo la hipótesis de que existiera al menos una función en las condiciones indicadas para el problema variacional con $E(u) < \infty$.

7.3. Las ecuaciones integrales y su influencia en el desarrollo del Análisis Funcional

¹⁰El estudio formal de las ecuaciones integrales suele remontarse a 1823 cuando Niels Abel estudiando un problema mecánico relacionado con la tautocrona resuelve la siguiente ecuación:

$$\int_0^x \frac{f(t)}{\sqrt{x-t}} dt = g(x)$$

donde g es una función conocida que da el tiempo de descenso e $y = f(x)$ es la curva buscada. En la terminología introducida por Hilbert, este es un ejemplo de *ecuación integral de primera especie*, ya que la función incógnita f aparece sólo bajo el signo integral. Una ecuación integral muy parecida fue resuelta de forma independiente por Liouville en 1832 ([24], pg 38). En 1836 Sturm y Liouville transformaron algunas ecuaciones diferenciales en ecuaciones integrales. Así toda solución de la ecuación:

$$y''(x) - q(x)y(x) + \lambda y(x) = 0$$

donde q es una función real continua en $[a, b]$ y $\lambda > 0$, también es solución de la ecuación integral

$$y(x) = a \cos \sqrt{\lambda}x + b \sen \sqrt{\lambda}x + \frac{1}{\sqrt{\lambda}} \int_a^x q(t)y(t) \sen \sqrt{\lambda}(x-t) dt$$

Esta es una *ecuación integral de segunda especie*, pues la incógnita aparece tanto dentro como fuera de la integral. En general, el problema de Sturm-Liouville:

$$\begin{aligned} p(x)u''(x) + p'(x)u'(x) + r(x)u(x) + \lambda u(x) &= 0 \\ \alpha_1 u(a) + \beta_1 u'(a) &= 0 \\ \alpha_2 u(b) + \beta_2 u'(b) &= 0 \end{aligned}$$

donde $p \in C^1([a, b])$ es una función positiva, $r \in C([a, b])$ y $\lambda \in \mathbb{C}$, es equivalente a la siguiente ecuación integral

$$u(x) + \int_a^b G(x, y)u(y) dy = 0$$

¹⁰En algunos puntos de esta sección sigo muy de cerca el trabajo de F. Bombal [5]. La referencia básica para ecuaciones integrales es el libro [23].

donde G es la *función de Green* del sistema (los detalles pueden consultarse en [1]).

En 1837 Liouville introdujo el método de iteraciones para resolver ciertos tipos de ecuaciones integrales. En 1865 A. Beer redujo el problema de Dirichlet para un dominio Ω en el plano al cálculo de una función φ verificando una ecuación integral del tipo:

$$\varphi(x) + \int_a^b k(x,y)\varphi(y) dy = f(x)$$

donde f es la condición en la frontera y k es una función continua y simétrica. Si consideramos la integral como un operador $K(\varphi)(x) = \int_a^b k(x,y)\varphi(y) dy$, podemos escribir la ecuación integral de la forma

$$(I + K)(\varphi) = f$$

Esta ecuación fue resuelta en 1877 por Carl Neumann en términos de una serie

$$\varphi = (I + K)^{-1}(f) = f - K(f) + K^2(f) - K^3(f) + \dots$$

cuya convergencia pudo probar en ciertos casos.

Fue en 1888 cuando P. du Bois-Reymond sugirió el nombre de *ecuaciones integrales* para designar este tipo de problemas, y propuso desarrollar una teoría general de tales ecuaciones como método alternativo para resolver problemas de ecuaciones diferenciales. J.M. Le Roux (1894) y V. Volterra (1896) fueron los primeros matemáticos que probaron teoremas de existencia y unicidad para una clase general de ecuaciones integrales. Aunque sus métodos eran parecidos, el trabajo de Volterra tuvo más influencia porque incluía una fórmula explícita para la solución y enfatizaba un principio básico: la analogía de las ecuaciones integrales con un sistema de ecuaciones lineales algebraico. En 1896 Volterra desarrolló una teoría general para ecuaciones integrales del tipo

$$\varphi(x) - \lambda \int_a^x k(x,t)\varphi(t) dt = f(x) \quad (40)$$

que ahora se llaman *ecuaciones de Volterra de segunda clase*. Se supone que las funciones que intervienen son continuas y que $k(x,t) = 0$ para $t > x$. Volterra usa un desarrollo en términos de iteradas para expresar la solución por medio de una ecuación integral de segunda especie.

Los trabajos decisivos en ecuaciones integrales fueron escritos en 1900, *Sur une nouvelle méthode pour la résolution du problème de Dirichlet*, y 1903, *Sur une classe d'équations fonctionnelles*, por Ivar Fredholm (1866-1927). En dichos trabajos se desarrolla una completa teoría de la ecuación integral de segunda especie

$$\varphi(x) - \lambda \int_a^b k(x,t)\varphi(t) dt = f(x) \quad (41)$$

Observa que cuando $k(x,t) = 0$ para $t > x$, esta ecuación se convierte en la ecuación integral de Volterra (40). Podemos, por tanto, razonar sobre la ecuación integral de Fredholm (41).

El método de las aproximaciones sucesivas, usado por Volterra, consiste en escribir la ecuación (41) en la forma

$$\varphi(x) = f(x) + \lambda \int_a^b k(x,t)\varphi(t) dt \quad (42)$$

y partiendo de una función inicial, por ejemplo $\varphi_0 = f$, se define por recurrencia

$$\varphi_n(x) = f(x) + \lambda \int_a^b k(x,t)\varphi_{n-1}(t) dt \quad (n = 1, 2, 3, \dots)$$

que puede escribirse en la forma

$$\varphi_n(x) = \sum_{m=0}^n \lambda^m \psi_m(x) \quad (43)$$

donde $\psi_0(x) = f(x)$ y

$$\psi_m(x) = \int_a^b k_m(x,y)f(y) dy \quad (44)$$

donde $k_1(x,y) = k(x,y)$ y

$$k_{m+1}(x,y) = \int_a^b k(x,z)k_m(z,y) dz \quad (m = 1, 2, 3, \dots)$$

es la sucesión de los *núcleos iterados*. Podemos escribir (43) en la forma

$$\varphi_n(x) = f(x) + \int_a^b \left[\sum_{m=1}^n \lambda^m k_m(x,y) \right] f(y) dy \quad (45)$$

La función dada por

$$H(x,y,\lambda) = - \sum_{n=0}^{\infty} \lambda^n k_{n+1}(x,y)$$

se llama *núcleo resolvente*, y la función

$$\psi(x) = f(x) - \int_a^b H(x,y,\lambda)f(y) dy$$

es solución de la ecuación integral (42). Por supuesto, para asegurar la convergencia de este proceso hay que hacer hipótesis adecuadas en las que no vamos a entrar.

En general, el método de las aproximaciones sucesivas permite resolver la ecuación (41) para valores pequeños de $|\lambda|$. El camino seguido por Fredholm fue diferente y consistió en definir un “determinante” y resolver la ecuación integral por analogía con un sistema algebraico de ecuaciones lineales.

Dividamos el intervalo $[a, b]$ en n partes de igual longitud $h = (b - a)/n$. Reemplacemos la integral en (41) por una suma de Riemann para obtener una ecuación aproximada

$$\varphi(x) - \lambda h \sum_{j=1}^n k(x,t_j)\varphi(t_j) = f(x) \quad x \in [a, b] \quad (46)$$

Hagamos sucesivamente en esta igualdad $x = t_1, t_2, \dots, t_n$ para obtener los valores aproximados de $\varphi(t_j)$. Obtenemos así el sistema de ecuaciones lineales en las incógnitas $\varphi(t_j)$:

$$\varphi_i - \lambda h \sum_{j=1}^n k_{ij} \varphi_j = f_i \quad i = 1, 2, \dots, n. \quad (47)$$

donde $f(s_i) = f_i$, $\varphi(s_i) = \varphi_i$, $k(s_i, s_j) = k_{ij}$. La solución de este sistema depende del valor del determinante

$$\Delta(\lambda) = \begin{vmatrix} 1 - \lambda h k_{11} & -\lambda h k_{12} & \cdots & -\lambda h k_{1n} \\ -\lambda h k_{21} & 1 - \lambda h k_{22} & \cdots & -\lambda h k_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ -\lambda h k_{n1} & -\lambda h k_{n2} & \cdots & 1 - \lambda h k_{nn} \end{vmatrix}$$

que es un polinomio en λ . Si $\Delta(\lambda) \neq 0$, el sistema (47) tiene solución única. Resolviendo este sistema y sustituyendo los valores obtenidos $\varphi_j = \varphi(t_j)$ en (46) obtenemos una solución aproximada de (41):

$$\varphi(x) \cong f(x) + \lambda \frac{Q(x, t_1, t_2, \dots, t_n, \lambda)}{\Delta(\lambda)} \quad (48)$$

donde Q y Δ son polinomios en λ .

Cuando $n \rightarrow \infty$, las sumas de Riemann en (46) tienden a la integral en (41), por lo que podemos esperar que el lado derecho de (48) converja a una solución exacta de (41). Razonando de forma parecida, Fredholm propuso como solución de (41):

$$\varphi(x) = f(x) + \lambda \int_a^b R(x, t, \lambda) dt \quad (49)$$

donde

$$R(x, t, \lambda) = \frac{D(x, t, \lambda)}{D(\lambda)}, \quad D(\lambda) = \sum_{n=0}^{\infty} \frac{(-1)^n}{n!} A_n \lambda^n, \quad D(x, t, \lambda) = \sum_{n=0}^{\infty} \frac{(-1)^n}{n!} B_n(x, t) \lambda^n$$

siendo

$$A_n = \int_a^b \cdots \int_a^b k \left(\begin{matrix} t_1, \dots, t_n \\ t_1, \dots, t_n \end{matrix} \right) dt_1 \dots dt_n$$

$$B_n(x, t) = \int_a^b \cdots \int_a^b k \left(\begin{matrix} x, t_1, \dots, t_n \\ t, t_1, \dots, t_n \end{matrix} \right) dt_1 \dots dt_n$$

$$k \left(\begin{matrix} x_1, \dots, x_n \\ t_1, \dots, t_n \end{matrix} \right) = \begin{vmatrix} k(x_1, t_1) & \cdots & k(x_1, t_n) \\ \cdots & \cdots & \cdots \\ k(x_n, t_1) & \cdots & k(x_n, t_n) \end{vmatrix}$$

El cálculo de A_n y $B_n(x, t)$ puede hacerse por medio de relaciones de recurrencia. La función $D(\lambda)$ se llama el *determinante de Fredholm* y $D(x, t, \lambda)$ el *primer menor de Fredholm* de la ecuación (41). La función $R(x, t, \lambda)$ se llama la *resolvente* o el *núcleo resolvente* de la ecuación (41).

La ecuación integral

$$\varphi(x) - \lambda \int_a^b k(t,x)\varphi(t) dt = f(x) \quad (50)$$

se llama *ecuación transpuesta* de la ecuación (41). Se verifica que ambas ecuaciones tienen el mismo determinante.

Bajo hipótesis convenientes, Fredholm justificó los procesos de convergencia, probó que $D(\lambda)$ es una función entera, y obtuvo los siguientes resultados conocidos como la *alternativa de Fredholm*:

- Si $D(\lambda) \neq 0$ entonces la ecuación (41) tiene una única solución que viene dada por (49), y la *ecuación homogénea* asociada

$$\varphi(x) - \lambda \int_a^b k(x,t)\varphi(t) dt = 0 \quad (51)$$

solo tiene la solución trivial nula $\varphi = 0$.

- Si λ_0 es un cero de $D(\lambda)$ de multiplicidad m entonces la ecuación (51) y la correspondiente ecuación homogénea de (50) tienen m soluciones linealmente independientes. En tal caso la ecuación (41) tiene solución si, y sólo si, la función f satisface que

$$\int_a^b f(x)\varphi_k(x) dx = 0 \quad (k = 1, 2, \dots, m)$$

donde las φ_k , $1 \leq k \leq m$ son soluciones linealmente independientes de la ecuación integral homogénea transpuesta.

Fredholm escribe:

Considerando la ecuación (41) como una transformación que convierte la función φ en una f , escribo esta misma ecuación $S_k\varphi(x) = f(x)$ y digo que la transformación S_k pertenece a la función $k(x,y)$.

Es claro que en este trabajo se encuentran ya los principios de una teoría de operadores entre espacios de funciones.

Todavía queda una última sorpresa y es que Fredholm aplicó los resultados obtenidos para probar muy fácilmente que el problema de Dirichlet en el plano tenía solución única para dominios cuya frontera sea una curva que tenga tangente y curvatura finita en todo punto ([23]).

Estos resultados impresionaron a la comunidad matemática y pusieron la teoría de ecuaciones integrales en el centro de interés de los matemáticos contemporáneos. Estos resultados, junto con la teoría de Sturm-Liouville, supusieron el punto de partida de la moderna teoría espectral e influyeron decisivamente en el desarrollo posterior del Análisis Funcional.

7.4. David Hilbert y el nacimiento de la teoría espectral

Los trabajos de Fredholm atrajeron la atención de matemáticos de todos los países, y entre ellos Hilbert fue uno de los más entusiastas. Entre los años 1904 y 1906, Hilbert publicó cinco trabajos sobre ecuaciones integrales, y en 1910 un sexto, todos ellos en la revista de la universidad de Göttingen, que fueron recogidos en un libro *Grundzüge einer allgemeinen Theorie der linearen Integralgleichungen* publicado en 1912. Estos trabajos están entre los que mayor influencia tuvieron en el siglo XX. En ellos formuló los teoremas y definiciones básicos de la teoría espectral (a la que él dio nombre) y de los espacio de Hilbert (que él no nombró, ni siquiera definió).

En el primer trabajo, Hilbert empieza su estudio de la ecuación integral (41). Siguiendo los mismos pasos que Fredholm, reemplaza la integral por las sumas de Riemann, y obtiene de (41) el sistema (47), que resuelve en forma de cociente de determinantes. Tomando límites vuelve a obtener, con demostraciones explícitas y rigurosas, los resultados de Fredholm. Seguidamente Hilbert considera un núcleo continuo y simétrico, es decir, $k(x,y) = k(y,x)$ y advierte la analogía entre la forma bilineal

$$\sum_{i,j=1}^n k_{ij}x_iy_j \tag{52}$$

y la integral

$$\int_a^b \int_a^b k(x,y) dx dy$$

Esto le lleva a escribir el teorema de los ejes principales para la forma bilineal (52) en una forma apropiada para pasar al límite. En este proceso probó que las raíces del determinante de Fredholm, a las que Hilbert llama *valores propios*¹¹ son reales y que si se escriben en una sucesión $\{\lambda_n\}$ repitiendo cada una tantas veces como su multiplicidad, entonces para cada n hay una *función propia*¹² φ_n :

$$\varphi_n(x) = \lambda_n \int_a^b k(x,t)\varphi_n(t) dt$$

Estas funciones propias verifican que

$$\int_a^b \varphi_n(t)\varphi_m(t) dt = 0 \quad (m \neq n)$$

Si dichas funciones propias se normalizan por la condición

$$\int_a^b \varphi_n(t)^2 dt = 1$$

¹¹Estos son los recíprocos de lo que entendemos actualmente por “valores propios”.

¹²Hilbert mantuvo la misma terminología *eigenvalues* “valores propios” o “autovalores”, y *eigenfunction* “función propia” o “autofunción”, usada en el caso finito dimensional.

y si para cada función continua x en $[a, b]$ definimos sus *coeficientes de Fourier* por

$$\langle x | \varphi_n \rangle = \int_a^b \varphi_n(t) x(t) dt$$

Hilbert probó que

$$\int_a^b \int_a^b k(s, t) x(s) y(t) ds dt = \sum_{n=1}^{\alpha} \frac{1}{\lambda_n} \langle x | \varphi_n \rangle \langle y | \varphi_n \rangle \quad (53)$$

donde α es un número finito o bien $\alpha = \infty$ dependiendo de que el número de valores propios distintos sea finito o infinito. En este último caso, la serie converge absolutamente siempre que las funciones x e y verifiquen que

$$\int_a^b x^2(t) dt < \infty, \quad \int_a^b y^2(t) dt < \infty$$

Hilbert interpreta la igualdad (53) como la generalización natural del clásico teorema de reducción de una forma cuadrática a sus ejes principales. A partir de esta igualdad prueba que cualquier función continua f que pueda expresarse en la forma

$$f(s) = \int_a^b k(s, t) g(t) dt \quad (54)$$

para alguna función continua g , verifica que su desarrollo de Fourier

$$f(s) = \sum_{n=1}^{\infty} \langle f | \varphi_n \rangle \varphi_n(s) \quad (55)$$

es absoluta y uniformemente convergente. La validez de este desarrollo para *toda* función continua f fue probada por uno de los mejores discípulos de Hilbert, E. Schmidt, en su Tesis (1905).

¹³ El artículo cuarto es uno de los más apreciados de Hilbert. En él aparece claramente el espíritu actual del Análisis Funcional y la Teoría Espectral. En efecto, en este artículo Hilbert da un salto cualitativo: abandona deliberadamente el marco de las ecuaciones integrales y trata de crear una teoría general de formas bilineales y cuadráticas de infinitas variables:

$$Q(x) = \sum_{p, q=1}^{\infty} k_{pq} x_p x_q \quad (k_{pq} = k_{qp})$$

donde $\sum_{n=1}^{\infty} x_n^2 < \infty$. Hilbert definió el espectro de la forma cuadrática Q distinguiendo el *espectro puntual* (los valores propios) del *espectro continuo*.

La forma cuadrática se llama *completamente continua* si

$$\lim_{n \rightarrow \infty} \sum_{p, q=1}^n k_{pq} x_p x_q = Q(x)$$

¹³En todo lo que sigue he copiado párrafos sueltos del antes citado trabajo de F. Bombal ([5]).

uniformemente para toda sucesión $\{x_n\}$ tal que donde $\sum_{n=1}^{\infty} x_n^2 < \infty$. Hilbert prueba que toda forma cuadrática completamente continua puede reducirse a sus ejes por una transformación ortogonal:

$$Q(x) = \sum_{n=1}^{\infty} \alpha_n x_n^2$$

Es evidente, con una perspectiva actual, que estos resultados están prefigurando la teoría de los espacios de Hilbert en su versión canónica de espacio ℓ_2 . Por otro lado, en 1906 aparece también la famosa Tesis Doctoral de M. Fréchet “*Sur quelques points du calcul fonctionnel*”, que tuvo una tremenda influencia, tanto para el desarrollo del Análisis Funcional como para el de la Topología. En su Tesis, Fréchet introduce la noción abstracta de distancia en un conjunto, lo que permite extender las nociones habituales de entornos, límites, continuidad, etc. en conjuntos abstractos. También introdujo Fréchet las nociones de compacidad, completitud y separabilidad, y las estudió en distintos espacios funcionales, $C([a, b])$, $H(D)$, $B([a, b])$, etc. mostrando la importancia de las mismas.

Estas ideas topológicas se difundieron rápidamente. No es extraño, pues, que se intentaran aplicar en el contexto de los importantes trabajos desarrollados por Hilbert. Este programa fue llevado a cabo por el mismo Fréchet y uno de los mejores discípulos de Hilbert: Erhard Schmidt quien, en un artículo publicado en 1908, definió el “espacio de dimensión infinita” ℓ_2 , con las nociones actuales de producto escalar, norma, ortogonalidad, etc. Introdujo también el lenguaje geométrico moderno, probando el teorema de la proyección ortogonal y el proceso de ortogonalización que lleva su nombre. Schmidt introdujo también por primera vez el símbolo que se usa habitualmente para la norma de un vector, él lo hizo para la norma de ℓ_2 :

$$\|x\| = \sqrt{\langle x|x \rangle} = \sqrt{\sum_{n=1}^{\infty} |x(n)|^2}$$

Otros dos jóvenes matemáticos, E. Fischer y F. Riesz, también compartieron esta visión geométrica y topológica del espacio de Hilbert, lo que les llevó a descubrir (independientemente) el llamado “teorema de Fischer-Riesz” (1907), que a su vez establece una inesperada relación de estos temas con otro gran descubrimiento de la época: la Teoría de integración de Lebesgue. El teorema en cuestión establece que, fijado un sistema ortonormal completo de funciones $\{\varphi_n : n \in \mathbb{N}\}$, la aplicación $f \rightarrow \{\langle f|\varphi_n \rangle\}$ es un isomorfismo hilbertiano entre el espacio $L_2([a, b])$ de las (clases de) funciones de cuadrado integrable Lebesgue sobre $[a, b]$ (que se define en estos trabajos), y el espacio de Hilbert ℓ_2 . Como importante subproducto, resulta que los resultados de Hilbert se pueden aplicar a cualquier ecuación integral con núcleo $k \in L_2([a, b]^2)$, objetivo perseguido infructuosamente por distintos matemáticos de la época (Hadamard y Hilbert entre ellos). Las consecuencias de este resultado estructural, hicieron ver la importancia del nuevo Análisis, y abrieron el camino hacia la introducción de los espacios L^p y ℓ^p por Riesz y, en definitiva, la aparición de la noción general de espacio normado.

7.5. Riesz, Hahn, Banach. Espacios vectoriales normados

Probablemente, uno de los mayores responsables del desarrollo del Análisis Funcional, tanto por la variedad de aplicaciones como por la profundidad y originalidad de sus contribuciones, es el matemático húngaro Frédéric Riesz (1880-1956). Ya hemos comentado el famoso teorema de Riesz-Fischer, descubierto independientemente por Fischer y Riesz en 1907. En el mismo año, Fréchet y Riesz, independientemente, obtienen la representación de cualquier forma lineal continua φ sobre el espacio $L_2([a, b])$ en la forma

$$\varphi(f) = \langle f | g \rangle = \int_a^b f(x)g(x) dx$$

para alguna (única) $g \in L_2([a, b])$.

Recuerda que para $p \geq 1$,

$$\ell_p = \left\{ x \in \mathbb{K}^{\mathbb{N}} : \sum_{n=1}^{\infty} |x(n)|^p < \infty \right\}$$

Riesz prueba que si φ es una forma lineal continua sobre ℓ_p , con $p > 1$, existe una única sucesión $z \in \ell_q$ donde $1/p + 1/q = 1$ tal que

$$\varphi(x) = \sum_{n=1}^{\infty} x(n)z(n) \quad (x \in \ell_p)$$

Estos resultados marcan el inicio de la teoría de la dualidad.

Las contribuciones fundamentales a que nos hemos venido refiriendo, preparan el camino para el desarrollo de una teoría general de espacios normados, funcionales y operadores lineales entre ellos. Esto aconteció en la Tesis de S. Banach *Sur les opérations dans les ensembles abstraits et leur application aux équations intégrales*, defendida en Junio de 1920 y publicada dos años después en *Fundamenta Mathematicae*. En la Introducción, Banach declara su intención de demostrar un serie de resultados válidos en distintos “campos funcionales”, para lo cual establece un conjunto de teoremas en un marco muy general que, por especialización, dan lugar a los distintos resultados buscados. Con sus propias palabras:

“El objetivo de este trabajo es demostrar algunos teoremas que son ciertos para diferentes espacios funcionales (champs fonctionnelles). En lugar de probar los resultados para cada espacio funcional particular, he optado por enfoque diferente: considero en general un conjunto de elementos abstractos, para los que postulo una serie de propiedades y demuestro los teoremas para esos conjuntos. Entonces pruebo que los distintos espacios funcionales particulares en los que estoy interesado, satisfacen los axiomas postulados...”

El marco general en cuestión es precisamente lo que hoy conocemos como espacio normado completo. Banach da la definición axiomática de espacio vectorial real, normado y completo y comprueba que numerosos campos funcionales verifican esos axiomas. La Tesis contiene, entre otros, el *principio de acotación uniforme* y el

teorema de las aplicaciones contractivas en espacios métricos completos, y Banach aplica con gran habilidad estos resultados a distintos espacios funcionales. Veamos un ejemplo de aplicación de éste último.

Convengamos en escribir la ecuación integral (41) en la forma:

$$\varphi = f + \lambda K\varphi \quad (56)$$

donde $K\varphi(x) = \int_a^b k(x,y)\varphi(y) dy$. Definamos $T : C([a, b]) \rightarrow C([a, b])$ por

$$T\varphi = f + \lambda K\varphi \quad (57)$$

Es evidente que resolver la ecuación integral (56) equivale a resolver la ecuación $\varphi = T\varphi$, es decir, a encontrar un punto fijo para T . Considerando en $C([a, b])$ la norma usual de la convergencia uniforme, tratemos de aplicar el teorema de las aplicaciones contractivas. Tenemos que

$$\|T\varphi - T\psi\|_\infty \leq \max_{a \leq y \leq b} |\lambda| \int_a^b |k(x,y)| |\varphi(y) - \psi(y)| dy \leq \|\varphi - \psi\|_\infty |\lambda| \max_{a \leq x \leq b} \int_a^b |k(x,y)| dy$$

Por tanto siempre que se cumpla la condición

$$|\lambda| \max_{a \leq x \leq b} \int_a^b |k(x,y)| dy < 1$$

El citado teorema garantiza la existencia y unicidad de un punto fijo para T y, por tanto, la existencia y unicidad de una solución de la ecuación integral (56).

Para terminar, veamos una aplicación del teorema de Hahn-Banach a la resolución de un sistema de infinitas ecuaciones lineales. Consideremos el sistema

$$\sum_{m=1}^{\infty} k_{nm}x_m = c_n \quad n = 1, 2, 3, \dots \quad (58)$$

Lo primero que debemos hacer es precisar el campo de trabajo. Hay que garantizar que las series convergen. Para ello, supondremos que $x = \{x_n\}$ es un vector genérico de ℓ_p , con $p > 1$, y que los vectores $A_n = \{a_{nm}\}_{m \in \mathbb{N}}$ están en ℓ_q con $1/p + 1/q = 1$. Estas condiciones garantizan que las series convergen absolutamente. Sea M el subespacio vectorial de ℓ_q generado por los A_n , esto es, $M = \text{Lin}(\{A_n : n \in \mathbb{N}\})$. Podemos definir una forma lineal φ sobre M (supuesto que las ecuaciones son linealmente independientes) por $\varphi(A_n) = c_n$ y extendiendo a M por linealidad. La condición de continuidad para dicha forma lineal viene dada porque exista una constante $\alpha > 0$ tal que $|\varphi(z)| \leq \alpha \|z\|_q$ para todo $z \in M$. Puesto que todo $z \in M$ puede escribirse de la forma $z = \sum_{j=1}^n \lambda_j A_j$, dicha condición se expresa por

$$\left| \sum_{j=1}^n \lambda_j c_j \right| \leq \alpha \left\| \sum_{j=1}^n \lambda_j A_j \right\|_q = \alpha \left(\sum_{m=1}^{\infty} \left| \sum_{j=1}^n \lambda_j k_{jm} \right|^q \right)^{1/q} \quad (59)$$

Cuando hay un $\alpha > 0$ tal que la condición anterior se cumple cualesquiera sean los números λ_j , $1 \leq j \leq n$ y cualquiera sea $n \in \mathbb{N}$, el teorema de extensión de Hanh-Banach afirma la existencia de una forma lineal continua $f : \ell_q \rightarrow \mathbb{R}$ que extiende a ϕ por lo que $f(A_n) = c_n$ para todo $n \in \mathbb{N}$. Dicha forma lineal f está definida por un único vector $u \in \ell_p$ de la forma

$$f(z) = \sum_{m=1}^{\infty} z_m u_m \quad (\forall z \in \ell_q)$$

Por tanto dicho vector u verifica que

$$f(A_n) = \sum_{m=1}^{\infty} k_{nm} u_m = c_n \quad n = 1, 2, 3, \dots$$

y, en consecuencia, hemos probado la existencia de la solución del sistema de infinitas ecuaciones lineales. Si exigimos que $\|f\| = \|\phi\|$ entonces la solución es única. Finalmente, la condición (59) no sólo es suficiente para la existencia de solución sino también necesaria ([19], pg 61).

Referencias

- [1] M.A. Al-Gwaiz. *Sturm-Liouville Theory and its Applications*. Springer-Verlag, London, 2008. 136
- [2] Kirsti Andersen. Las Técnicas del Cálculo, 1630-1660. En *Del Cálculo a la Teoría de Conjuntos, 1630-1910. Una introducción histórica*. Alianza Editorial, S.A., 1984. 47, 68
- [3] John L. Bell. *The Continuous, the Discrete and the Infinitesimal in Philosophy and Mathematics*. The Western Ontario Series in Philosophy of Science Volume 82. Springer Nature Switzerland AG 2019. 2, 7
- [4] F. Bombal. *Matemáticas y Ciencia*. Rev.R.Acad.Cienc.Exact.Fís.Nat. (Esp) Vol. 103, N°. 2, pp 279-295, 2009. 7
- [5] F. Bombal. *Análisis Funcional: Una perspectiva histórica*. <https://eprints.ucm.es/id/eprint/19972/1/bombal66.pdf> 135, 141
- [6] Bos, H.J.M. Newton, Leibniz y la Tradición Leibniziana. En *Del Cálculo a la Teoría de Conjuntos, 1630-1910. Una introducción histórica*. Alianza Editorial, S.A., 1984. 68, 86, 88, 89
- [7] U. Bottazzini. *The Higher Calculus: A History of Real and Complex Analysis from Euler to Weierstrass*. Springer-Verlag New York Inc., 1986. 31
- [8] David M. Burton. *The History of Mathematics: An Introduction, Sixth Edition*. The McGraw-Hill Companies, 2007. 98

- [9] H. D. Ebbinghaus, H. Hermes, F. Hirzebruch, M. Koecher et al. Numbers. Graduate Texts in Mathematics 123. Springer, New York, 1991. 1
- [10] Leonhard Euler Introduction to Analysis of the Infinite. Book I Springer-Verlag, New York, 1988 100, 102
- [11] I. Grattan-Guinness. Bolzano, Cauchy and the “New Analysis” of Early Nineteenth Century. *Archive for History of Exact Sciences*, 6(7):372 – 400, 1970. 45
- [12] M. Kline. *El pensamiento matemático de la antigüedad a nuestros días*. Vols. 1, 2 y 3. Alianza Editorial, S.A., Madrid, 1992. 1, 9, 25, 96, 102, 106
- [13] Jean Dhombres *Las progresiones infinitas el papel del discreto y del continuo en el siglo XVII*. LLULL, vol. 16, 1993, 43-114. 91
- [14] C.H. Edwards The Historical Development of the Calculus. Springer-Verlag New York, Inc., 1976. 90, 96
- [15] J. Fourier Théorie analytique de la chaleur, edición facsímil. Cambridge University Press, 2009. 109
- [16] González Urbaneja, P.M. *Las Técnicas del Cálculo: Fermat, Wallis y Roberval*. https://fundacionorotava.org/media/web/publication_files/publication23__a2_c016w.pdf. 47, 68
- [17] Ralph E. Kenyon, Jr. Atomism and Infinite Divisibility. <http://www.xenodochy.org/rekphd/index.html>, 1994. 113
- [18] Israel Kleiner. History of the Infinitely Small and the Infinitely Large in Calculus. *Educational Studies in Mathematics*, 48:137 – 174, 2001. 47, 68
- [19] F. Riesz. *Les systèmes d'équations linéaires a une infinité d'inconnues*. Gauthier-Villars, París, 1913. 131, 132, 145
- [20] Gert Schubring. *Conflicts between Generalization, Rigor, and Intuition*. Sources and Studies in the History of Mathematics and Physical Sciences. Springer, New York, 2005. 37, 39
- [21] Michael Spivak. *Cálculo Infinitesimal*. Reverté Ediciones S.A., México D.F., 2ªed. - 3ª Reimpresión edition, 1996. 24
- [22] S.B. Russ. A Translation of Bolzano's Paper on the Intermediate Value Theorem. *Historia Mathematica*, (7):156 – 185, 1980. 43, 44
- [23] F.G. Tricomi. *Integral equations*. Interscience Publisher Inc., New York, 1957 135, 139

- [24] V. Volterra. *Leçons sur les équations intégrales et les équations intégro-différentielles*. Gauthier-Villars, Paris, 1913. [135](#)
- [25] H. Wussing. *Lecciones de Historia de las Matemáticas*. Siglo XXI de España Editores, S.A., Madrid 1998. [10](#)
- [26] A. P. Youschkevitch. *The Concept of Function up to the Middle of the 19th Century*. *Archive for History of Exact Sciences*. Volume 16, issue 1, March 1976. Springer. [31](#)
- [27] A. Zygmund. *Trigonometric Series*. Third Edition. Volumes I & II combined. Cambridge University Press, 2002. [106](#)