

Historias de Matemáticas

Contraejemplos en Matemáticas

Counterexamples in Mathematics

Antonio Rosales Góngora

Revista de Investigación



Volumen V, Número 2, pp. 061-078, ISSN 2174-0410

Recepción: 13 Feb'15; Aceptación: 10 Jul'15

1 de Octubre de 2015

Resumen

Un contraejemplo es un ejemplo que prueba la falsedad de un enunciado. Cuando se quiere demostrar la falsedad de una afirmación es suficiente con encontrar un ejemplo que incumpla la afirmación.

Un teorema necesita a menudo de varias hipótesis, para comprender como funcionan es importante estar convencido de la necesidad de cada una de ellas. En este artículo hacemos un recorrido por algunas propiedades y teoremas matemáticos viendo la necesidad de todas las hipótesis.

Palabras Clave: contraejemplos, lógica, conjuntos, grupos, anillos, números reales, sucesiones, funciones, probabilidad.

Abstract

A counterexample is an example that proves the falsehood of a terms of reference. When the falsehood of an affirmation wants to be demonstrated it is sufficient in spite of finding an example that breaks the affirmation.

A theorem needs often from several hypotheses, to understand since they work it is important to be sure of the need of each one of them. In this article we do a tour for some properties and mathematical theorems seeing the need of all the hypotheses.

Keywords: counterexamples, logic, sets, groups, rings, royal numbers, successions, functions, probability.

1. Introducción

Un contraejemplo es la excepción que confirma la regla. Los contraejemplos juegan un papel clave para la comprensión de las matemáticas.

Un teorema es válido en todos los casos en que las hipótesis impuestas en el enunciado se verifican. La negación de un enunciado, es decir la afirmación que es falso, se demuestra por

la existencia de un caso en el que se verifican las hipótesis, son verdaderas, sin que lo sea la conclusión. La justificación matemática de la falsedad de un enunciado se consigue mediante un contraejemplo.

En la mayoría de los casos un teorema necesita varias hipótesis, para comprenderlo es importante estar convencido de la necesidad de cada una de ellas. Se debe tratar de demostrar que el enunciado obtenido suprimiendo una de las hipótesis es falso.

Algunos conceptos, cuando se estudian por primera vez, dan la impresión que un resultado debe ser verdadero, por ejemplo, toda función continua en cero es continua en un entorno de cero; si esto ocurre es porque la idea intuitiva es errónea, el dar algunos contraejemplos a esas ideas preconcebidas ayuda a rectificar esa mala intuición.

2. Lógica, conjuntos, aritmética

La lógica matemática comienza con la obra de George Boole "*The Mathematical Analysis of Logic*" (El Análisis Matemático de la Lógica) publicado en 1847. El libro es lo bastante diferente de lo hasta entonces realizado como para considerarlo el nacimiento de una nueva teoría (Teoría de la Lógica Matemática) y no como un paso más en el desarrollo de una antigua.

George Boole (1815 – 1864), de formación autodidacta, redactó su libro en solo unas semanas en la primavera de 1847, con 31 años. Es una obra que ofrece una lógica basada en la matemática, sobre todo en el álgebra, cuyas ideas capitales son la de clase, elemento de clase, y operaciones de selección de elementos de clase y su teoría está constituida por el uso de ecuaciones.

La validez de los procesos del Análisis Matemático no depende de la lectura o interpretación de los signos en él utilizados, sino solamente de las leyes que gobiernan los posibles modos de unión de esos signos.

El álgebra de Boole es una sola teoría pero tiene, al menos, dos sistemas de lectura, uno que lo pone en relación con clases y otro que lo pone en relación con enunciados.

La teoría de conjuntos formaliza, en torno a 1900 bajo el impulso de Cantor y Dedekind, la lógica matemática y la convierte en una rama más de la matemática.

Los matemáticos se apoyan en particular en la lógica de predicados de primer orden, construida con la ayuda de conectores proposicionales "no", "y", "o", "implica", "equivalente", de variables x, y, z, \dots , de proposiciones, de predicados $P(x), Q(x, y) \dots$ y de los cuantificadores universal \forall y existencial \exists .

Si se usan cuantificadores en un texto, su manejo debe hacerse con cuidado. Así, la proposición $\forall x \exists y (P(x, y))$ puede ser verdadera mientras que $\exists y \forall x (P(x, y))$, obtenida intercambiando los cuantificadores, es falsa.

Así si $P(x, y)$ representa la afirmación: x es menor o igual que y , entonces, para todo número natural x , si se elige un natural y mayor o igual que x , la afirmación $P(x, y)$ es verdadera; la proposición $\forall x \exists y (P(x, y))$ es verdadera. En cambio, como en \mathbb{N} no existe un elemento máximo, la afirmación $\exists y \forall x (P(x, y))$, es falsa.

Observemos que si $\exists y \forall x (P(x, y))$, es válida, entonces $\forall x \exists y (P(x, y))$ también lo es, es decir, la implicación $\exists y \forall x (P(x, y)) \Rightarrow \forall x \exists y (P(x, y))$, es universalmente válida. Intuitivamente podemos ver que en el término de la izquierda, y no depende de x mientras que en el de la derecha puede hacerlo; la condición es menos fuerte. Este es el motivo por el que hay que utilizar cuantificadores de distinta naturaleza para pasar de la continuidad a la continuidad uniforme.

Los conectores "y" y "o" se simbolizan respectivamente por \wedge y \vee . De este modo las proposiciones $\forall x (P(x) \wedge Q(x))$ y $\forall x (P(x)) \wedge \forall x (Q(x))$ son equivalentes, de la misma forma que

$\exists x(P(x) \vee Q(x))$ y $\exists x(P(x)) \vee \exists x(Q(x))$. No ocurre lo mismo con $\forall x(P(x) \vee Q(x))$ y $\forall x(P(x)) \vee \forall x(Q(x))$ pues la primera puede ser verdadera y la segunda falsa.

Tomando \mathbb{N} como conjunto de referencia y representamos por $P(x)$, para todo natural x , la afirmación: x es par, y por $Q(x)$ la proposición: x es impar. Entonces $\forall x(P(x) \vee Q(x))$ expresa que todo número natural es par o impar, lo que es verdadero; por el contrario $\forall x(P(x)) \vee \forall x(Q(x))$ expresa que son todos los naturales pares o todos los naturales impares, lo cual es falso.

De la misma forma $\exists x(P(x) \wedge \exists x(Q(x)))$ puede ser verdadera sin serlo $\exists x(P(x) \wedge Q(x))$. En efecto, usando las mismas hipótesis anteriores, $\exists x(P(x) \wedge \exists x(Q(x)))$ expresa que al menos existe un número natural par y al menos un natural impar, mientras que $\exists x(P(x) \wedge Q(x))$ expresa que existe simultáneamente un natural par e impar.

3. Teoría de Conjuntos

La Teoría de conjuntos, obra de los matemáticos alemanes George Cantor y Richard Dedekind, aparece a finales del XIX. En una primera aproximación, ahora calificada de ingenua, Cantor define conjunto como “una colección en un todo de determinados y distintos objetos de nuestra percepción o nuestro pensamiento, llamados los elementos del conjunto”.

En 1903 Bertrand Russel demuestra la inconsistencia de la teoría y cuestiona la definición de conjunto en la teoría de Cantor. Pero pronto la teoría axiomática de Zermello (1908) y los refinamientos de esta debidos a Fraenkel (1908), Skolem (1923), Von Newman (1925) y otros sentaron las bases de la teoría actual.

Tras veinte años de trabajo, en 1902 Gottlob Frege había terminado el segundo volumen de sus “*Las leyes fundamentales de la aritmética*” con el que trataba de dar fundamentación lógica a las matemáticas a partir de la teoría de conjuntos, terminando de imprimir el libro, Frege recibió una nota de Russel en la que le explicaba que había encontrado una paradoja a la teoría de conjuntos. Frege, sin tiempo para más, insertó la siguiente nota al final del libro:

“Difícilmente puede haber algo más indeseable para un científico que ver el derrumbe de sus cimientos justamente cuando la obra está acabada. La carta del Sr. Bertrand Russell me ha puesto en esta situación ...”

El matemático y filósofo inglés Bertrand Russel (1872 - 1970) contaba así su paradoja:

“Me parece que una clase a veces es, y a veces no es, un miembro de sí misma. La clase de las cucharitas de té, por ejemplo, no es otra cucharita de té, pero la clase de cosas que no son cucharitas de té es una de las cosas que no son cucharitas... [esto] me condujo a considerar las clases que no son miembros de sí mismas; y éstas, parecía, debían formar una clase. Me pregunté si esta clase es o no un miembro de sí misma. Si es un miembro de sí misma, debería poseer las propiedades que definen a dicha clase, que consisten en no ser miembros de sí mismas. Si no es un miembro de sí misma, no debe poseer la propiedad definitoria de la clase, y por tanto debe ser un miembro de sí misma. Así cada alternativa lleva a su opuesta y existe una contradicción.”

Si consideramos el conjunto $A = \{x \mid x \notin X\}$ entonces el objeto A pertenece o no pertenece al conjunto A .

Si A pertenece a A entonces, por definición de A , se tiene que A no pertenece a A ; si A no pertenece a A , la misma definición permite afirmar que A pertenece a A . Así las dos afirmaciones son simultáneamente verdaderas.

En definitiva había descubierto que el conjunto de los conjuntos que no son miembros de sí mismos lleva a una contradicción.

Lo que Whitehead le dijo a Russell cuando este le contó su descubrimiento es bastante gráfico: “nunca habrá otra vez una alegre y confiada mañana”. Si E denota el conjunto de todos los conjuntos y $P(E)$ el conjunto de las partes de E , entonces $P(E)$ está incluido en E pero esto contradice el teorema de Cantor que afirma que “si E es un conjunto, no existe inyección alguna de $P(E)$ en E ”, es la conocida como *paradoja de Cantor*.

Para solucionar tales problemas se elaboró una nueva teoría de conjuntos. Es la teoría de Zermello – Fraenkel, ZF, llamada así porque fue concebida por el matemático y lógico alemán Ernest Zermello en 1908 – expuesta en sus Investigaciones sobre los fundamentos de la Teoría de conjuntos – y modificada por el matemático y lógico israelí Abrahan Fraenkel en 1921 y 1922. La teoría ZF sirve de fundamento a las matemáticas, pues permite una construcción rigurosa aunque su consistencia, es decir la ausencia de paradojas, no podrá probarse como demostró Kurt Gödel en 1931.

Especialmente con vistas a la noción de cardinal, Georg Cantor utiliza básicamente correspondencias biunívocas entre conjuntos, mientras que Richard Dedekind introduce la noción de aplicación entre conjuntos.

Sabemos que dada una aplicación f de un conjunto E en un conjunto F , la imagen directa de un subconjunto A de E por f es el subconjunto $f(A) = \{f(x) \mid x \in A\}$ de F y, si M es un subconjunto de F , la imagen recíproca de M por F es el subconjunto $f^{-1}(M) = \{x \in E \mid f(x) \in M\}$ de E .

Si A y B son dos subconjuntos de E , se tiene $f(A \cup B) = f(A) \cup f(B)$, pero en general se tiene sólo la inclusión $f(A \cap B) \subset f(A) \cap f(B)$. Si f es inyectiva la inclusión se transforma en igualdad.

Considerando la aplicación $f : \mathbb{N} \rightarrow \mathbb{Z}$ de \mathbb{N} en \mathbb{Z} y los subconjuntos $A = \mathbb{N}$ y $B = -\mathbb{N}$ de \mathbb{Z} . Tendremos $A \cap B = \{\emptyset\}$ y $f(A \cap B) = \{\emptyset\}$. Sin embargo, $f(A) = \mathbb{Z}$ y $f(B) = \mathbb{Z}$ luego $f(A) \cap f(B) = \mathbb{Z}$, así $f(A \cap B) \neq f(A) \cap f(B)$.

Sabemos que dados dos subconjuntos A de E y B de F y una aplicación f de E en F se tienen las inclusiones $A \subset f^{-1}(f(A))$, $f(f^{-1}(B)) \subset B$. La primera será una igualdad si f es inyectiva y la segunda si f es sobreyectiva de E en F .

Considerando de nuevo la aplicación $f : \mathbb{N} \rightarrow \mathbb{Z}$ de \mathbb{N} en \mathbb{Z} y el subconjunto $A = \mathbb{N}$, tendremos $f(A) = \mathbb{Z} = f(\mathbb{Z})$, por tanto $f^{-1}(f(A)) = \mathbb{Z} \neq \mathbb{N}$, por tanto $A \neq f^{-1}(f(A))$.

Intuitivamente dos conjuntos son equipolentes si tienen “el mismo número de elementos”, lo que conduce a la noción de cardinal de un conjunto, desarrollada por Cantor paralelamente a la noción de ordinal (los cardinales sirven para contar el número de elementos y los ordinales para numerarlos).

Todos los conjuntos tienen un cardinal que permite establecer que dos conjuntos son equipolentes si y sólo si tienen el mismo cardinal.

Euclides estableció que “la parte es siempre más pequeña que el todo” que, como sabemos, sólo es cierto en conjuntos finitos. En un conjunto infinito existe una parte estricta con un mismo cardinal, es decir, equipolente a él.

Así, el conjunto P de los números naturales pares es una parte estricta de \mathbb{N} y la aplicación $f : \mathbb{N} \rightarrow P$ de \mathbb{N} sobre P , así pues \mathbb{N} y P son equipolentes. De la misma forma la aplicación $f : \mathbb{R} \rightarrow \mathbb{R} \in]-1, 1[\dots f(x) = \frac{x}{1 + |x|}$ define una aplicación biyectiva entre los números reales y una parte estricta suya: $] -1, 1[$.

El conjunto \mathbb{N} parece tener “menos” elementos que \mathbb{N}^2 . Sin embargo esta conjetura no es cierta, ya que ambos conjuntos son equipolentes pues la aplicación $f : \mathbb{N}^2 \rightarrow \mathbb{N}$ dada por

$f(p, q) = a_{p+q} + q$ siendo $a_n = \sum_{i=0}^n \frac{n(n+1)}{2}$ (cociente exacto) es biyectiva.

En una carta a Dedekind fechada el 29 de Junio de 1877 Cantor escribe: “*je le vois, mais je ne le crois pas*” (“lo veo pero no me lo creo”), refiriéndose a la equipolencia que había establecido entre \mathbb{R}^2 y \mathbb{R} . Hoy sabemos que todo conjunto infinito es equipolente a su cuadrado.

Se demuestra que $\mathbb{Z}, \mathbb{N}^n, \mathbb{Q}$, el conjunto de los números algebraicos son numerables, es decir, equipolentes a \mathbb{N} ; podría pensarse que ocurre lo mismo para todos los conjuntos infinitos, pero no es así pues, por ejemplo, \mathbb{R} no es numerable.

4. Aritmética

Algunas conjeturas han necesitado varios siglos para demostrarlas. La más celebre es debida a Fermat demostrada en 1993 – 1994 por el matemático inglés Andrew Wiles y conocida hoy como Teorema de Fermat – Wiles: “*Para todo entero natural $p \geq 3$, la ecuación $x^p + y^p = z^p$ no admite soluciones (x, y, z) con x, y, z naturales no nulos*”.

Inspirándose en este futuro resultado, Leonhard Euler conjeturó más generalmente que una potencia n -ésima no puede escribirse como la suma de k potencias n -ésimas con $k < n$. Así, no se puede escribir $q^4 = a^4 + b^4 + c^4$ ni $q^5 = a^5 + b^5 + c^5 + d^5$ con a, b, c, d, q números naturales no nulos, lo cual es falso pues $2682440^4 + 15365639^4 + 18796760^4 = 20615673^4$; resultado establecido por el norteamericano Noam Elkies en 1966, lo que puso fin a la conjetura de Euler.

Tenemos también la igualdad $27^5 + 84^5 + 110^5 + 133^5 = 144^5$. Se puede demostrar que todo natural se puede escribir como la suma de al menos nueve cubos de elementos de \mathbb{N}^* , pero no se puede reducir a ocho cubos.

Los números de Fermat, llamados así porque fueron introducidos por Fermat, son los naturales $F_n = 2^{2^n} + 1$ para n natural, se tiene $F_0 = 3, F_1 = 5, F_2 = 17, F_3 = 257$. Como vemos F_0, F_1, F_2, F_3 son primos. Fermat creyó demostrar en 1658 que F_n es un número primo para todo natural n , lo cual es falso pues el número de Fermat $F_5 = 2^{32} + 1$ es divisible por 641: $640 = 5 \times 2^7 \therefore 5 \times 2^7 \equiv -1 \pmod{641}$, elevando a la cuarta se tiene: $5^4 \times 2^{28} \equiv 1 \pmod{641}$, como $5^4 \equiv 625 \equiv -16 \pmod{641}$; así multiplicando las congruencias se tiene: $-2^{32} = (-2^4) \times 2^{28} \equiv 1 \pmod{641}$, así pues 641 divide a F_5 y como $641 \neq F_5$, entonces F_5 no es primo. De hecho F_5 es el producto de los dos número primos 641 y 6700417. Landry probó en 1880 que $F_6 = 274177 \times 67280421310721$. El entero F_4 es, al igual que F_0, F_1, F_2, F_3 primo pero no se ha podido encontrar algún n mayor o igual a 5 tal que F_n sea primo. La conjetura “*para todo entero natural n , F_n es un número primo*” de Fermat parece transformarse en “*para todo entero n mayor o igual que 5, F_n no es primo*”.

5. Grupos

Évariste Galois introdujo hacia 1830 la noción de grupo. Murió poco tiempo después en un duelo a los 21 años. Sus notas son confusas y se necesitaron una quincena de años para comprender su importancia.

Dos tipos de problemas llevan a la formalización de esta teoría. El estudio de las permutaciones, es decir, el estudio de las biyecciones de un conjunto finito en el mismo; y el estudio de los grupos de transformaciones en geometría, es decir, grupo de las biyecciones del plano o del espacio que conservan algunas propiedades (isometría, semejanza ...)

Matemáticos como Camile Jordan y Leopold Kronecker vieron la importancia de funda-

mentar esta teoría sobre axiomas para darle generalidad. La formalización definitiva no se hará hasta 1893.

Una ley de composición interna sobre un conjunto E es una aplicación de $E \times E$ en E . Entre las leyes de composición internas más frecuentemente utilizadas en los conjuntos numéricos, son raras las que no son asociativas, pero existen. Por ejemplo, la ley de composición interna $*$ sobre \mathbb{N} dada por $a * b = a^b$ no es asociativa pues se tiene $(2 * 2) * 3 = 2^{2^3} = 4^3 = 64$, mientras que $2 * (2 * 3) = 2^{2^3} = 2^8 = 256$.

Tampoco es conmutativa pues $2 * 3 = 2^3 = 8 \neq 9 = 3^2 = 3 * 2$.

Sabemos que si una ley de composición interna posee un elemento neutro a derecha y un elemento neutro a izquierda entonces son iguales, lo que asegura la unicidad y existencia de elemento neutro. No obstante, una ley puede tener una infinidad de elementos neutros a un lado por lo que, según lo anterior, no posee ninguno al otro lado.

Así la ley de composición interna $*$ sobre \mathbb{N} dada por $a * b = b$ es asociativa y todo natural es un elemento neutro a la izquierda.

La noción de grupo fue axiomatizada en 1893 por el matemático Walter Von Dyck (1856 – 1934). En 1882 ya había dado una axiomatización de grupos finitos en el número vigésimo de la revista *Mathematische Annalen*. En este mismo número, su compatriota Heinrich Weber (1842 – 1913) había dado una axiomatización diferente a la de Dyck. Aquí imponía la asociatividad de la ley y la regularidad a derecha e izquierda, es decir, que cualesquiera que sean los elementos a, b y c de G , $a * c = b * c$ implica $a = b$ y $c * a = c * b \Rightarrow a = b$.

En el caso de conjuntos infinitos estas propiedades son insuficientes para obtener un grupo.

Sabemos que un conjunto G con una ley de composición interna $*$ es un grupo si verifica los axiomas I) la ley $*$ es asociativa II) La ley $*$ posee elemento neutro III) Todo elemento de G posee simétrico. Estos tres axiomas son necesarios pues existen leyes de composición internas verificando dos de ellos que no tienen estructura de grupo. Por ejemplo, la ley de composición interna sobre \mathbb{N} dada por $a * b = a + b$ es asociativa y regular a izquierda y derecha pero todo elemento no tiene simétrico.

La ley definida sobre el conjunto $E = \{e, a, b\}$ por $e * x = x * e = x$, para todo elemento x de E y $x * y = e \forall x, y \in E - \{e\}$, cumple los axiomas II y III pero no es asociativa pues $a * (a * b) = a * e = a$ y $(a * a) * b = e * b = b$.

6. Anillos y Cuerpos

La noción de anillo fue introducida, en la segunda mitad del siglo XIX, por los matemáticos alemanes Richard Dedekind (1836 – 1916) y David Hilbert (1862 – 1943) para generalizar los conjuntos de números $\mathbb{Z}, \mathbb{Q}, \mathbb{R}$ dotados de la adición y el producto. La introducción de anillos no conmutativos viene justificada por su utilidad para estudiar las matrices y algunos espacios de funciones. En esta misma época Richard Dedekind introduce la noción de ideal según una idea de Ernst Kummer (1810 – 1893), para afinar la noción de divisibilidad. En los años 20 del siglo pasado, la teoría de anillos se enriquece con la introducción de diferentes tipos de anillos como los anillos factoriales, principales o noetherianos.

Como sabemos un anillo es una terna $(A, +, *)$ donde $(A, +)$ es un grupo abeliano aditivo y $*$ es una ley de composición interna sobre A , asociativa y distributiva respecto a $+$ en A y que posee elemento neutro. Evidentemente existen anillos no conmutativos, por ejemplo el anillo infinito $M_2(\mathbb{R})$ de las matrices cuadradas de orden 2 y coeficientes reales:

$$\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix} \neq \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$$

Pese a que el teorema de Wedderburn, establecido independientemente en 1905 por el matemático escocés MacLogan Wedderburn (1882 – 1948) y el matemático norteamericano Leonard Dickson (1874 – 1954), establece que todo cuerpo finito es conmutativo, esto no es válido para los anillos.

En efecto el anillo $M_2(F_2)$ de matrices cuadradas de segundo orden con coeficientes en el cuerpo $F_2 = \mathbb{Z}/2\mathbb{Z}$, de los enteros módulo 2, no es conmutativo y es finito pues su cardinal es $2^4 = 16$.

Por otra parte la condición de existencia de elemento neutro es fundamental. Si consideramos P el conjunto de los enteros relativos pares, sabemos que es un subgrupo de $(\mathbb{Z}, +)$ estable para la multiplicación \times de \mathbb{Z} . Así $(P, +)$ es un grupo abeliano aditivo, como \times es asociativa y distributiva respecto de $+$ sobre P , entonces P es un pseudoanillo. Si P fuese un anillo, entonces (P, \times) poseería un elemento neutro q tal que $2 \times q = 2$ de donde $q = 1$ en contradicción con la pertenencia de q a P . ¿Cuántas raíces tienen un polinomio?, estamos familiarizados con la respuesta en el caso general de polinomios con coeficientes en un cuerpo (o anillo íntegro), pero si consideramos por ejemplo el anillo $A[x]$ de polinomios con coeficientes en el anillo $A = \mathbb{Z}/4\mathbb{Z}$, enteros módulo 4, resulta que el polinomio $2x + 1$ no admite ninguna raíz en A mientras que el polinomio $2x$ admite dos raíces (0 y 2) en A .

En el anillo $A[x]$ de los polinomios con coeficientes en el anillo $A = \mathbb{Z}/12\mathbb{Z}$, enteros módulo 12, la ecuación $x^2 = 4$ posee cuatro soluciones: $x_1 = 2, x_2 = 4, x_3 = 8, x_4 = 10$, y la ecuación $x^2 + 3x + 2 = 0$ también admite como raíces 2, 7, 10, 11. Así los polinomios $x^2 - 4$ y $x^2 + 3x + 2$ son de segundo grado y ambos admiten cuatro raíces en A .

Como vemos en el anillo $A[x]$ un polinomio puede tener distintas factorizaciones: $x^2 - 4 = (x - 2) \cdot (x - 10) = (x - 4) \cdot (x - 8)$ o $x^2 + 3x + 2 = (x - 2) \cdot (x - 7) = (x - 10) \cdot (x - 11)$.

Como es sabido un polinomio P con coeficientes en un cuerpo conmutativo K es irreducible sobre K (o en el anillo $K[x]$) si $gr(P) \geq 1$ y si todo divisor de P en $K[x]$ es un polinomio constante o del mismo grado que P . Además un polinomio de grado 2 ó 3 sobre un cuerpo conmutativo K es irreducible si no admite raíces en K . Esta propiedad no es válida para grado 4 pues existen polinomios de grado 4 sin raíces. Así $x^4 + 1 = (x^2 + \sqrt{2}x + 1) \cdot (x^2 - \sqrt{2}x + 1)$ nos muestra que el polinomio $x^4 + 1$ es reducible sobre \mathbb{R} , sin embargo, $x^4 + 1$ no admite raíces reales pues $x^4 + 1 > 0$ para todo número real x .

En general, el producto de dos polinomios de grados dos con coeficientes reales de discriminante estrictamente negativo es un polinomio de grado cuatro sobre \mathbb{R} sin raíces reales.

7. Espacios Vectoriales

La noción de espacio vectorial fue introducida en 1840 por Arthur Cayley (1821 – 1895) y Hermann Grassmann (1809 – 1877). Cayley consideró las n -uplas de reales como un elemento y definió operaciones sobre estos objetos.

Los espacios vectoriales fueron formalizados en 1888 por Giuseppe Peano, convirtiéndose en un elemento natural de la geometría, pero también de otras muchas ramas como el análisis funcional. El álgebra lineal permite utilizar la intuición geométrica en teorías matemáticas desprovistas, aparentemente, de soporte intuitivo.

Como sabemos un espacio vectorial sobre un cuerpo conmutativo K , con elemento unidad 1, es un grupo abeliano $(E, +)$ con una ley de composición externa de dominio $K: K \times E \rightarrow E$ tal que $(\lambda, x) \rightarrow \lambda x$ verificando los cuatro axiomas siguientes:

$$\text{I. } \lambda(x + y) = \lambda x + \lambda y$$

$$\text{II. } (\lambda + \mu)x = \lambda x + \mu x$$

$$\text{III. } \lambda(\mu x) = (\lambda\mu)x$$

$$\text{IV. } 1 \cdot x = x$$

Estos cuatro axiomas son necesarios, es decir, ninguno de estos axiomas es consecuencia de los otros.

En efecto, consideremos un espacio vectorial E de dimensión finita $n \geq 2$ sobre el cuerpo de los complejos, un subespacio vectorial D de dimensión 1 de E y una ley de composición externa

$$* \text{ de } \mathbb{C} \times E \rightarrow E \text{ dada por } (\lambda, x) \rightarrow \lambda * x = \begin{cases} \lambda x & \text{si } x \in D \\ \bar{\lambda}x & \text{si } x \notin D \end{cases}$$

Podemos comprobar que se verifican los axiomas II, III y IV, sin embargo no se verifica el axioma I pues si consideramos un vector y_0 no nulo de D y un vector $z_0 \notin D$ (existe pues $h \geq 2$), entonces el vector $y_0 + z_0 \notin D$ pues eligiendo un complejo no real λ (por ejemplo i) se tiene: $\lambda * (y_0 + z_0) = \bar{\lambda}(y_0 + z_0) = \bar{\lambda}y_0 + \bar{\lambda}z_0$, en cambio: $\lambda * y_0 + \lambda * z_0 = \lambda y_0 + \bar{\lambda}z_0$, lo que muestra que $\lambda * (y_0 + z_0) \neq \lambda y_0 + \lambda z_0$. El axioma I no se verifica. No obstante, como $\bar{1} = 1$, y para cualesquiera escalares λ y μ , $\overline{\lambda + \mu} = \bar{\lambda} + \bar{\mu}$ y $\overline{\lambda\mu} = \bar{\lambda}\bar{\mu}$ es claro que se verifican los axiomas II, III y IV.

Si consideramos un espacio vectorial real E no nulo y definimos la ley de composición externa $*$ de \mathbb{R} sobre E de la siguiente forma $\lambda * x = x, \forall x \in E, \forall \lambda \in \mathbb{R}$. Eligiendo un vector no nulo x_0 de E y haciendo $\lambda = \mu = 1$ tendremos: $(\lambda + \mu) * x_0 = x_0$. En cambio, $\lambda * x_0 + \mu * x_0 = x_0 + x_0 = 2x_0$ por tanto no se cumple el axioma II.

Sin embargo, $\forall x, y \in E, \forall \lambda, \mu \in \mathbb{R}$ se tiene $\lambda * (x + y) = x + y = \lambda * x + \lambda * y$; $\lambda * (\mu * x) = \mu * x = x = (\lambda\mu) * x$ y $1 * x = x$ por tanto se verifican los axiomas I, III y IV.

Considerando ahora un espacio vectorial complejo $E \neq \{\emptyset\}$ y definiendo la ley de composición externa de dominio \mathbb{C} sobre E : $\forall x \in E, \forall \lambda \in \mathbb{C}, \lambda * x = \text{Re}(\lambda)x$ (siendo $\text{Re}(\lambda)$ la parte real de λ). Para un vector no nulo x_0 de E se tiene: $i^2 * x_0 = (-1) * x_0 = -x_0$ y $i * (i * x_0) = i * 0 = 0$. Así pues no se verifica el tercer axioma. Claramente se verifica I, y como $\text{Re}(1) = 1$ entonces $\text{Re}(\lambda + \mu) = \text{Re}(\lambda) + \text{Re}(\mu)$, se verifican también II y IV.

Si ahora consideramos $E = \mathbb{R}^2$ y definimos la ley de composición externa de dominio \mathbb{R} sobre E : $\forall z = (x, y) \in E, \forall \lambda \in \mathbb{R}, \lambda * z = (\lambda x, 0)$, no se verifica el axioma IV pues $1 * (0, 1) = (0, 0) \neq (0, 1)$; aunque es fácil verificar I, II y III.

8. Números Reales

Pese a que los números reales se han utilizado desde muy antiguo, no es hasta el siglo XIX cuando se definen de manera rigurosa. Durante mucho tiempo se contentaron con justificaciones intuitivas basadas en la "evidencia geométrica". La necesidad de definir la noción de continuidad y límite aparece hacia 1820 con Bernhart Bolzano y Augustin Louis Cauchy, y se desarrolla con Karl Weierstrass hacia 1850. Precisamente Weierstrass propuso en 1863 la primera construcción de los números reales. No la publicó hasta 1872 cuando Charles Meray y George Cantor (sucesiones de Cauchy de números racionales) y Richard Dedekind (cortaduras de números racionales) publicaron otras.

En la antigüedad, la densidad de los números racionales en los reales llevó a pensar que cualquier punto (la abscisa de cualquier punto) era racional. La regla, el compás y el teorema de Pitágoras permitieron dibujar un punto de abscisa $\sqrt{2}$ que no es racional. Surgen así los irracionales.

Debido a la creencia de que los puntos de una recta eran racionales o irracionales expresables por radicales, hubo un intento en 1683 por parte de Tschirnhaus de expresar por radicales las raíces de un polinomio con coeficientes enteros. Seguramente debido a que eran muy pocos los números utilizados que careciesen de representación racional o radical. Uno era π y el otro e , introducido en 1614 por Navier, de los que no se sabían si eran racionales o irracionales.

En 1737 Euler prueba que e y e^2 son irracionales y establece la fórmula $e^{i\pi} + 1 = 0$ que liga el estudio de π y e .

Lambert, utilizando aproximaciones por fracciones continuas como Euler, prueba en 1761 la irracionalidad de π , de e y de $\tan x$ para todo racional no nulo.

Bezout continúa intentando expresar las raíces de un polinomio por radicales. Así en 1762 escribe una raíz de $P(x)$ en la forma $A_1\rho + A_2\rho^2 + \dots + A_{n-1}\rho^{n-1}$ siendo ρ una raíz n -ésima de la unidad distinta de uno, verificando por tanto la relación $1 + \rho + \rho^2 + \dots + \rho^{n-1} = 0$.

El estudio de la resolución de ecuaciones de grado menor o igual que 4, aparecido en los trabajos de Waring, Lagrange y Vandermonde entre 1770 y 1772 lleva a las fórmulas de resolución de ecuaciones de segundo, tercer y cuarto grado con la introducción de la famosa resolvente de Lagrange. Para las de quinto grado, la resolvente le lleva a una ecuación de sexto grado lo que le hace dudar de la posibilidad de resolver por radicales una ecuación de grado mayor o igual que cinco.

La imposibilidad de resolver por radicales una ecuación de quinto grado la propuso P. Ruffini en cuatro artículos publicados de 1799 a 1813. Estos trabajos llevaron a Cauchy, en 1814, a introducir las sustituciones que preparan a Galois la noción de grupo y le permitirá obtener el criterio necesario de resolubilidad por radicales vislumbrado por Abel.

Otra consecuencia de los artículos de Ruffini es el descubrimiento de que existen números algebraicos, es decir, raíces de polinomios con coeficientes enteros que no son expresables por radicales.

El conjunto de los reales \mathbb{R} , era utilizado aunque no estuviese construido formalmente y así surge, tras los trabajos de Ruffini, la conjetura de Liouville de la existencia de irracionales no algebraicos.

Liouville, generalizando en 1844 un resultado de Lagrange, encuentra que hay número irracionales que no pueden ser algebraicos, se les llama trascendentes o irracionales trascendentes. El ejemplo clásico de irracional trascendente es $10^{-1!} + 10^{-2!} + 10^{-3!} + \dots$

Durante todo el siglo XVIII los grandes progresos del análisis, en particular el estudio de las funciones desarrolladas en serie, se obtienen apoyándose en una noción intuitiva de límite y de convergencia de una sucesión (serie) a un límite. Muchos de los cálculos que se hacen se apoyan en cálculos sobre series formales. Además, los matemáticos de este siglo no llegan a desarrollar satisfactoriamente las bases del cálculo infinitesimal. Una bases rigurosa es la de Lagrange con su libro *Teoría de funciones analíticas*, en el que, como subtítulo, dice que contiene los principios del cálculo diferencial, quitando toda consideración de infinitésimos, evanescencias, límites y fluxiones y reduciendo el análisis algebraico de las cantidades finitas.

Se trata de un intento de algebraizar el análisis, apoyándose en los desarrollos de funciones en series de Taylor. El siglo XIX se caracterizó por una vuelta a las preocupaciones de rigor y de los fundamentos.

Una nueva época se abre a raíz del artículo de Gauss sobre el estudio de la convergencia de la serie hipergeométrica en 1813. Bolzano, muerto en 1848, se interesa en los fundamentos del análisis. Los progresos posteriores de rigor en análisis están motivados por la preocupación de los matemáticos por la enseñanza.

El primero de ellos es Cauchy cuyo *Cours d'analyse de l'École Polytechnique*, en 1821, es un modelo para sus contemporáneos; no obstante, no llega a hacer una clara distinción entre

nociones como convergencia y convergencia uniforme. El problema de la representación en series trigonométricas de funciones conocidas lleva a Fourier a dar una definición más general del concepto de función. Abel muestra la inexactitud de un teorema de Cauchy que afirmaba que si una serie de funciones continuas es convergente en el entorno de un punto, su suma es una función continua, publicando un contraejemplo. Weierstrass establece claramente el concepto de convergencia uniforme: es el autor de la definición de límite con los ϵ y δ utilizados hoy en día.

Para dar un curso de análisis, Richard Dedekind elabora en 1863 una teoría publicada en 1872 con el título Continuidad y números irracionales, donde define los números reales como cortaduras en el conjunto de los racionales.

Todo lo anterior nos permite ver como los matemáticos comprenden que la solución de los fundamentos del análisis pasaba por una definición de los números reales desprendida de cualquier recurso a la intuición geométrica.

Los primeros intentos de hacer una obra en análisis que represente lo mismo que los Elementos de Euclides en Geometría, fueron los de Martin Ohm en 1822 y Bolzano en 1830. Weierstrass, en 1863, expone, en su curso, una teoría de los números reales pero la construcción más acabada de los números reales aún se hace esperar 10 años.

En el desarrollo de las ideas que conducen a la aritmetización del análisis, Bolzano juega un papel importante. En un artículo publicado en 1817 titulado *“Demostración puramente analítica del teorema: entre dos valores cualesquiera que dan dos resultados de signo opuesto se encuentra al menos una raíz real de la ecuación”*, establece explícitamente la necesidad, para justificar el teorema, de no recurrir a evidencias geométricas, ni a las nociones de tiempo y movimiento. En el artículo dice que

“No hay nada que objetar ni contra la justificación ni contra la evidencia de este teorema geométrico. Pero es igualmente claro que hay una falta intolerable contra el “buen método” que consiste en tratar de deducir las verdades de las matemáticas puras o generales (es decir de la aritmética, del álgebra o del análisis) de resultados que pertenecen a una parte aplicada o particular (la geometría)”.

Se ve la diferencia con las matemáticas del siglo pasado. La geometría se vuelve una parte aplicada de las matemáticas. Bolzano critica las demostraciones propuestas con anterioridad, enuncia y demuestra el criterio de Cauchy para convergencia de series, explicita la noción de límite superior (cota superior) de un conjunto construyendo por dicotomía dos sucesiones adyacentes que convergen hacia este límite superior. La demostración del teorema de los valores intermedios resultara de eso, o más bien, la existencia de un límite común a las dos sucesiones. Bolzano demuestra la unicidad del límite, trata de establecer la existencia pero sólo demuestra que la hipótesis de una magnitud invariable teniendo la propiedad de aproximar los términos de nuestra serie no contiene nada imposible: esto se debe a que esta hipótesis permite determinar esta magnitud con la precisión deseada.

Hay que hacer notar que Bolzano es el único, en la primera mitad del siglo XIX, en considerar este problema. Cauchy se contentó con decir, en su curso de análisis de 1821, tras enunciar el criterio precedente, *“Recíprocamente, cuando estas condiciones diversas se cumplen, la convergencia de la serie está asegurada”*.

Bolzano es un precursor para los matemáticos hasta tal punto que Weierstrass, Dedekind y Cantor completaron el trabajo de aritmetización del análisis entre 1860 y 1870.

El conjunto de los números racionales está estrictamente incluido en el de los números reales pues existen números reales que no son racionales, como probó Euler en 1737 al demostrar que e , base de los logaritmos neperianos, es irracional.

Entre los dos se sitúan en particular dos conjuntos de números, el conjunto de los números

algebraicos A (formado por los números que son raíces de un polinomio no nulo con coeficientes enteros) y el de los números construibles Γ (formado por los números reales que pueden construirse con regla y compás a partir de $(0, 0)$ y $(1, 0)$). Los conjuntos de los números racionales, números trascendentes y construibles son subcuerpos del cuerpo de los números reales y verifican las inclusiones estrictas: $\mathbb{Q} \subset \Gamma \subset A \subset \mathbb{R}$. Las desigualdades son estrictas pues existen números reales construibles que no son racionales como, por ejemplo, $\sqrt{2}$.

Existen números reales algebraicos no construibles como $\sqrt[3]{2}$ pues es raíz del polinomio $x^3 - 2$, irreducible sobre \mathbb{Q} , y no construible pues su grado no es potencia de dos.

Existen números reales que no son algebraicos como e que es trascendente como estableció en 1873 el matemático francés Charles Hermite (1822 – 1901). El matemático alemán Carl Lindemann (1852 – 1939) demostró en 1882 que π es trascendente, lo que demostró la imposibilidad de la cuadratura del círculo.

9. Sucesiones Numéricas

La noción de sucesión aparece muy pronto en matemáticas. El primer ejemplo celebre es la sucesión de Fibonacci introducida en el siglo XIII. La formalización precisa de la noción de convergencia data de principios del siglo XIX.

La importancia de las sucesiones en análisis procede del hecho que los problemas topológicos de \mathbb{R} , y más generalmente de un espacio métrico, pueden ser tratados con la ayuda de sucesiones.

En 1874 Karl Weierstrass estableció el teorema: “De toda sucesión acotada se puede extraer una subsucesión convergente”, aunque no publicó la demostración. Bernhard Bolzano lo había enunciado hacia 1830 pero su trabajo no fue descubierto hasta 1930.

El teorema no significa, evidentemente, que toda sucesión acotada sea convergente pues, por ejemplo, la sucesión $\{x_n\}_{n \geq 0}$ de término general $x_n = (-1)^n$ está acotada en \mathbb{R} pues $\forall n \in \mathbb{N}, |x_n| = 1 \leq 1$; sin embargo es divergente pues para un número real λ dado y $\forall \epsilon \in \mathbb{N}$ se tiene $|x_{n+1} - x_n| = 2, |x_n - \lambda| \geq 1$ o $|x_{n+1} - \lambda| \geq 1$.

Se sabe que una sucesión de números reales converge en \mathbb{R} si, y sólo si, es una sucesión de Cauchy. Bolzano utilizó esta propiedad de las sucesiones de números reales en 1817 en su demostración del teorema de los valores intermedios, pero es Cauchy quien introduce esta noción en su Curso de Análisis de 1821, para el estudio de series. Bolzano y Cauchy admiten esta propiedad sin demostración.

El interés de esta propiedad radica en que no es necesario el conocimiento previo del límite para mostrar que una sucesión converge.

Una sucesión $\{x_n\}$ es de Cauchy si $\forall \epsilon > 0, \exists N \in \mathbb{N} | \forall p, q \geq N, |x_q - x_p| < \epsilon$; haciendo $p = n$ y $q = n + 1$ tenemos que, para una sucesión de Cauchy $\{x_n\}$, la sucesión $(x_{n+1} - x_n) \rightarrow 0$.

Sin embargo el recíproco no es cierto. En efecto, si consideramos la sucesión $\{x_n\}_{n \geq 1}$ de término general $x_n = 1 + \frac{1}{2} + \frac{1}{3} + \dots + \frac{1}{n} = \sum_{p=1}^n \frac{1}{p}$, se tiene para todo natural mayor o igual

que uno: $x_{2n} - x_n = \sum_{p=n+1}^{2n} \frac{1}{p} \geq n \frac{1}{2n} = \frac{1}{2}$, luego no es de Cauchy, pero: $0 < x_{n+k} - x_n =$

$$\sum_{p=n+1}^{n+k} \frac{1}{p} \leq \frac{k}{n}, \text{ así la sucesión } (x_{n+k} - x_n) \rightarrow 0.$$

Como sabemos si λ es un número real y f una función definida en un entorno de λ y conti-

na en λ , entonces, para toda sucesión $\{x_n\}$ que converge hacia λ , la sucesión de las imágenes $f(x_n)$ converge hacia $f(\lambda)$; sin embargo si la función no fuese continua en λ , podríamos encontrar una función, por ejemplo $f(x) = \frac{1}{x}$ definida en $\mathbb{R} - \{0\}$ y continua en su dominio, y una sucesión, por ejemplo $x_n = \frac{1}{n}$ convergente hacia 0 tal que la sucesión de sus imágenes diverge pues $f(x_n) = n$.

A menudo se piensa que una sucesión de límite nulo debe ser monótona o monótona a partir de un término, pero no hay motivo para ello. Por ejemplo, la sucesión $\{u_n\}_{n \geq 1}$ de término general $u_n = \frac{(-1)^n}{n}$ tiende hacia 0, $\forall n \geq 1, u_{2n} \geq u_{2n+1} \geq u_{2n-1}$ así la sucesión $\{u_n\}$ no es ni creciente ni decreciente a partir de un rango. En este caso, la sucesión se alterna alrededor de su límite 0. Esto no es necesario pues el término general de la sucesión puede ser de signo constante. En efecto, la sucesión $u_n = \begin{cases} \frac{1}{n} & \text{si } n \text{ par} \\ \frac{1}{n^2} & \text{si } n \text{ impar} \end{cases}$ da, para $n \geq 1, u_n \leq \frac{1}{n}$ por tanto $u_n \rightarrow 0$. Además, $\forall n \geq 1$ es $u_{2n} = \frac{1}{4n^2} < \frac{1}{2n+1} = u_{2n+1}$ y $u_{2n+1} = \frac{1}{2n+1} > \frac{1}{4n^2+8n+4} = u_{2n+2}$, como vemos la sucesión no es creciente ni decreciente a partir de un término dado.

10. Funciones de una variable real, continuidad y límites

La noción de continuidad fue planteada en el siglo XVII, hay que esperar hasta Bolzano en 1817 y Cauchy en 1821 para obtener una definición satisfactoria. La definición actual se debe a Karl Weierstrass hacia 1860.

Las múltiples relaciones y estructuras de las que se pueden dotar a \mathbb{R} (orden, topología, estructura de cuerpo) interfieren las unas con las otras y dan a las funciones de una variable real, es decir, a las aplicaciones de un subconjunto de \mathbb{R} en \mathbb{R} , un cierto número de propiedades que, en general, no pueden, más que parcialmente, generalizarse a las funciones sobre un subconjunto de \mathbb{R}^n , $n \geq 2$, como por ejemplo el teorema de los valores intermedios.

La continuidad es una propiedad de las funciones ligada a la Topología. En \mathbb{R} la topología proviene de una métrica obtenida por la distancia canónica, lo que permite decir que una aplicación de A en \mathbb{R} es continua en un punto a de A si, y sólo si, para toda sucesión $\{x_n\}$ de puntos de A que converge hacia a , la sucesión de sus imágenes $\{f(x_n)\}$ converge hacia $f(a)$.

En 1829 el matemático alemán Gustav Lejeune Dirichlet (1805 – 1859) introdujo la función que lleva su nombre $f(x) = \begin{cases} 1 & \text{si } x \text{ es racional} \\ 0 & \text{si } x \text{ es irracional} \end{cases}$, de \mathbb{R} en \mathbb{R} , para demostrar que no todas las funciones son continuas en algún punto. En efecto, la función anterior no es continua en ningún punto pues, dado un número real a , si es racional (respectivamente, irracional) tomando $\varepsilon = 1$, como entre dos números reales distintos existe una infinidad de racionales e irracionales, existirá un irracional (respectivamente un racional) x tal que $|x - a| < \delta$ y $|f(x) - f(a)| = |0 - 1| = 1$ (respectivamente $|f(x) - f(a)| = |1 - 0| = 1$), luego $|f(x) - f(a)| \geq \varepsilon$ y por tanto f es discontinua en a , para cualquier a .

Durante mucho tiempo se utilizó el teorema de los valores intermedios “una función definida sobre un intervalo I verifica el teorema de los valores intermedios si $\forall a, b \in I$, el segmento $[f(a), f(b)]$ está contenido en $f([a, b])$ ”, sin más justificación que invocar a la evidencia geométrica. El primer intento de prueba se debe a Bolzano en 1817.

En 1868 el matemático francés Ossian Bonnet demostró que la función derivada de una función definida y derivable sobre un intervalo verifica sobre este intervalo el teorema de los valores intermedios, incluso si no es continua. Este resultado se conoce como teorema de Darboux pues fue Gaston Darboux quien dio el primer ejemplo de función discontinua verificando el teorema de los valores intermedios.

La función $f : \mathbb{R} \rightarrow \mathbb{R}$ dada por $f(x) = \begin{cases} \cos \frac{1}{x} & \text{si } x \neq 0 \\ 1 & \text{si } x = 0 \end{cases}$, es continua en $a \neq 0$, y discontinua en $a = 0$ pues, $\forall n \in \mathbb{N}^*$, la sucesión de término general $u_n = \frac{1}{n\pi}$ converge hacia 0 pero $f(u_n) = \cos(n\pi) = (-1)^n$ no converge en \mathbb{R} aunque verifica el teorema de los valores intermedios. En efecto, si a y b son dos números reales tales que $a < b$, es claro que $f([a, b]) \subset [-1, 1]$.

Si $0 \notin [a, b]$ entonces es $a > 0$ o $b < 0$, así $[a, b]$ está contenido en $]0, \infty[$ o $]-\infty, 0[$ de donde se deduce que f es continua sobre $[a, b]$; aplicando el teorema de los valores intermedios a la restricción de f al intervalo $[a, b]$, tendremos $[f(a), f(b)] \subset f([a, b])$.

Para el caso donde $0 \in [a, b]$ tenemos, suponiendo $b \neq 0$, la sucesión $\{u_n\}_{n \geq 1}$, de término general $u_n = \frac{1}{n\pi}$. Es estrictamente decreciente y converge hacia 0. Como $b > 0$ existen índices naturales $n \geq 1$ tal que $u_n < b$, así $a \leq 0 < u_{n+1} < u_n < b$. Elegimos uno de esos n , como $\cos(n\pi, (n+1)\pi) = [-1, 1]$ entonces $f([u_{n+1}, u_n]) = [-1, 1]$: $[-1, 1] = f([u_{n+1}, u_n]) \subset f([a, b]) \subset [-1, 1] \Rightarrow f([a, b]) = [-1, 1]$.

Como $f(a)$ y $f(b)$ pertenecen a $[-1, 1]$: $[f(a), f(b)] \subset [-1, 1] \subset f([a, b])$.

Si $b = 0$ entonces $a < 0$ y cambiando b por a y $\{u_n\}_{n \geq 1}$ por su opuesta, se demuestra, como en el caso anterior, que $[f(a), f(b)] \subset [-1, 1] = f([a, b])$, por tanto la función verifica el teorema de los valores intermedios.

Sabemos que una función es continua en un punto $a \in \mathbb{R}$ cuando la función admite a $f(a)$ como límite por la derecha y por la izquierda de a . Si no coinciden los límites laterales la función es discontinua, así en la función $f : \mathbb{R} \rightarrow \mathbb{R}$, dada por $f(x) = \begin{cases} 1 & \text{si } x = 0 \\ 0 & \text{si } x \neq 0 \end{cases}$, los límites laterales son nulos mientras que el valor numérico de la función en 0 es 1, por tanto la función es discontinua en $x = 0$.

No todas las funciones tienen límites, incluso existen funciones que no tienen límite a izquierda ni límite a la derecha de un punto. En efecto, la función $f : \mathbb{R}^* \rightarrow \mathbb{R}$ dada por $f(x) = \text{sen}\left(\frac{1}{x}\right)$ no tiene límite a la derecha de 0.

En efecto, para un número real l consideramos $\varepsilon_0 = 1$ y sea $\eta > 0$ un número real. Elegimos un natural N tal que $\forall n \geq N, a_n < \eta$; en particular $a_N < \eta$, así $0 < a_N < \eta$ y $0 < b_N < \eta$. Tomando $c = \frac{f(b_N) + f(a_N)}{2}$ se tiene $f(b_N) < c < f(a_N)$, $c - f(b_N) = 1$ y $f(a_N) - c = 1$. Así tendremos, en los casos $l \leq c$ y $c < l$, $|f(a_N) - l| \geq 1$ o $|f(b_N) - l| \geq 1$. Hemos encontrado un número real x tal que $0 < x < \eta$ y $|f(x) - l| > \varepsilon_0$ luego la función no admite límite a la derecha de cero. Como la función es impar, ocurre lo mismo a la izquierda.

Incluso hay funciones que no tienen límite a la izquierda ni a la derecha de ningún punto. En efecto, si consideramos la función de Dirichlet, $f : \mathbb{R} \rightarrow \mathbb{R}$ cuya expresión ya hemos visto anteriormente, $f(x) = \begin{cases} 1 & \text{si } x \text{ es racional} \\ 0 & \text{si } x \text{ es irracional} \end{cases}$, y sean a y l números reales. Para $\varepsilon_0 = \frac{1}{2}$ sea $\eta > 0$. Elegimos un racional r y un irracional α en $]a, a + \eta[$ entonces $f(r) - f(\alpha) = 1$. Tomando $c = \frac{f(b) + f(r)}{2}$ se tiene $f(\alpha) < c < f(r)$; $c - f(\alpha) = \frac{1}{2}$ y $f(r) - c = \frac{1}{2}$; así, razonando en los casos $l \leq c$ y $c < l$ se ve que $|f(r) - l| \geq \frac{1}{2}$ o $|f(\alpha) - l| \geq \frac{1}{2}$. Hemos encontrado un número real x tal que $a < x < a + \eta$ y $|f(x) - l| > \varepsilon_0$, por tanto la función no admite a l por límite a la derecha de a . Reemplazando $]a, a + \eta[$ por $]a - \eta, a[$ y operando de la misma forma, se obtiene que f no admite límite a la izquierda.

11. Derivabilidad

La noción de derivabilidad aparece a finales del siglo XVII, al mismo tiempo que el cálculo integral, bajo el impulso de Newton y Leibniz. Antes de ellos, Descartes se había interesado en el problema de las tangentes a una curva y Pierre de Fermat había introducido un concepto muy próximo al de derivada investigando los extremos.

En la segunda mitad del XIX nace el interés por las propiedades de las funciones derivadas, es entonces cuando se introducen los contraejemplos que desafían la intuición generalmente admitida.

La noción de derivabilidad en un punto a es un problema local, está ligada al comportamiento de la función en las proximidades de a . Nos podemos preguntar si ello conduce a propiedades para la función sobre todo un entorno de a (sobre todos los puntos del entorno). La respuesta es no, existen funciones derivables en cero y discontinuas en todo real distinto de cero.

Consideremos la aplicación $f : \mathbb{R} \rightarrow \mathbb{R}$ dada por $f(x) = \begin{cases} x^2 & \text{si } x \text{ es racional} \\ 0 & \text{si } x \text{ es irracional} \end{cases}$, entonces $\forall x \in \mathbb{R}, |f(x)| \leq x^2$, así, $\forall x \in \mathbb{R} - \{0\}: \left| \frac{f(x)-f(0)}{x} \right| = \left| \frac{f(x)}{x} \right| \leq |x|$ luego $\frac{f(x)-f(0)}{x-0}$ tiene límite cero en cero, así f es derivable en cero y $f'(0) = 0$.

Sea a un número real no nulo. Consideremos $\{r_n\}_{n \geq 0}$ la sucesión de valores decimales aproximados por defecto de a e introducimos la sucesión $\{\alpha_n\}_{n \geq 0}$ de término general $\alpha_n = r_n - (1/10^n)$. Ambas sucesiones convergen hacia a . $\forall n \in \mathbb{N}$, r_n es racional y α_n es irracional, así $f(r_n) = (r_n)^2$ y $f(\alpha_n) = 0$, lo que demuestra que la sucesión $\{f(r_n)\}$ converge hacia $a^2 > 0$ mientras que la sucesión $\{f(\alpha_n)\}$ converge hacia cero. Así la función es discontinua en a .

Nos preguntamos sobre un problema parecido ¿una función puede ser derivable en todos los números reales salvo en uno?

Si la función es continua la respuesta es fácil, por ejemplo $f(x) = |x|$ es continua en \mathbb{R} y no derivable en $a = 0$ pues $f'(a^+) = 1 \neq f'(a^-) = -1$.

El problema se complica si buscamos la función sin derivadas laterales en el punto.

Si consideramos $f : \mathbb{R} \rightarrow \mathbb{R}$ dada por $f(x) = \begin{cases} x \operatorname{sen} \left(\frac{1}{x} \right) & \text{si } x \neq 0 \\ 0 & \text{si } x = 0 \end{cases}$ tenemos que $\forall x \in \mathbb{R} - \{0\}, |f(x)| = |x| \left| \operatorname{sen} \left(\frac{1}{x} \right) \right| \leq |x|$, desigualdad también válida para $x = 0$, de manera que $\lim_{x \rightarrow 0} f(x) = f(0)$ lo que prueba que f es continua en cero.

Además si $a \in \mathbb{R} - \{0\}$ entonces es derivable en a y por tanto continua en a , así la función es continua en \mathbb{R} y derivable en todo real no nulo, con $f'(x) = \operatorname{sen} \left(\frac{1}{x} \right) - \frac{1}{x} \cos x$.

Tenemos así $\forall x \in \mathbb{R} - \{0\}, \frac{f(x)-f(0)}{x-0} = \frac{f(x)}{x} = \operatorname{sen} \left(\frac{1}{x} \right)$.

La sucesión $\{a_n\}_{n \geq 0}$ de término general $a_n = \frac{1}{\left(\frac{\pi}{2} + n\pi\right)}$ toma valores en $]0, \infty[$ y converge hacia 0; la sucesión $b_n = -a_n$, que toma valores en $] - \infty, 0[$, converge hacia 0. Entonces: $\frac{f(a_n)-f(0)}{a_n-0} = (-1)^n$ y $\frac{f(b_n)-f(0)}{b_n-0} = (-1)^{n+1}$ son términos generales de sucesiones divergentes, por tanto no existe la derivada de la función a derecha ni izquierda.

Como sabemos la derivabilidad en un punto implica la continuidad en ese punto, el recíproco es falso como hemos visto con $f(x) = |x|$. Hasta la mitad del siglo XIX se pensaba, generalmente, que una función continua era derivable salvo quizás en algún punto. Ampère creyó haberlo demostrado en 1806. Bolzano dio, hacia 1830, un ejemplo de función continua pero derivable en ningún punto pero sus escritos no se conocieron.

En 1854 Riemann propuso, sin demostración, la función $R(x) = \sum_{n=1}^{+\infty} \frac{\text{sen}(n^2x)}{n^2}$. Karl Weierstrass se declara incapaz de demostrarlo. Es necesario esperar hasta 1871 para saber que $R(x)$ no es derivable salvo en algunos puntos. En 1872 Weierstrass demostró que si a y b son números reales tales que $a > 0$ y $b > 0$ y $ab > 1 + \frac{3\pi}{2}$, la función $f : x \rightarrow f(x) = \sum_{n=1}^{+\infty} b^n \cos(a^n x)$ es continua en \mathbb{R} y no es derivable en punto alguno de \mathbb{R} . En 1903 el matemático japonés Teiji Takagi (1875 – 1960) propuso la función $f : x \rightarrow f(x) = \sum_{n=0}^{+\infty} b^n g(a^n x)$ donde g es la función de \mathbb{R} en \mathbb{R} dada por $x \rightarrow g(x) = d(x, \mathbb{Z})$, siendo a y b números reales tales que $0 < b < 1$ y $a \geq 4$.

Sabemos que toda función f definida y continua sobre un intervalo I es la función derivada sobre I de una función definida y derivable sobre I , lo que significa que tiene primitivas sobre I . En efecto, eligiendo un punto a del intervalo I , la aplicación $F : x \rightarrow F(x) = \int_a^x f(t)dt$ de I en \mathbb{R} es derivable sobre I y $F'(x) = f(x), \forall x \in I$. Esto no se generaliza para funciones discontinuas, por ejemplo $f : \mathbb{R} \rightarrow \mathbb{R}$ dada por $f(x) = \begin{cases} 1 & \text{si } x = 0 \\ 0 & \text{si } x \neq 0 \end{cases}$ es continua en todo número real no nulo y discontinua en cero pues es constante en el entorno de todo real no nulo y, como la sucesión $\{1/n\}_{n \geq 1}$ converge hacia cero y $f(1/n) = 0$ para todo natural $n \geq 1$, la sucesión $\{f(1/n)\}$ converge hacia $0 \neq f(0)$.

Si suponemos que f admite una primitiva F sobre \mathbb{R} , entonces F será derivable en \mathbb{R} tal que $F'(x) = f(x), \forall x \in \mathbb{R}$. Para todo número positivo el teorema de los incrementos finitos aplicado a F sobre $[0, x]$ no da un $c \in]0, x[$ tal que $F(x) - F(0) = (x - 0)F'(c) = x f(c) = 0$, entonces $\frac{F(x) - F(0)}{x - 0} = 0$, es decir $F'(0^+) = 0$, lo cual contradice el $f(0) = 1$ de la definición. En conclusión, la función f no admite primitiva sobre \mathbb{R} .

Una de las utilizaciones más frecuentes de la función derivada de una función definida y derivable sobre un intervalo, es el estudio de sus variaciones sobre ese intervalo.

Como sabemos una función f , definida y continua sobre un intervalo I y derivable en el interior I_0 de I , es creciente sobre I si, y sólo si, su función derivada f' es positiva o nula sobre I_0 .

La importancia de establecer las hipótesis sobre un intervalo es crucial. En efecto, si consideramos $f : \mathbb{R} - \{0\} \rightarrow \mathbb{R}$ dada por $f(x) = -\frac{1}{x}$ está definida y es derivable en $\mathbb{R} - \{0\}$ y, $\forall x \in \mathbb{R} - \{0\}$, se tiene $f'(x) = \frac{1}{x^2} \geq 0$; sin embargo f no es creciente sobre $\mathbb{R} - \{0\}$ pues, por ejemplo, $1 > -1$ y $f(1) = -1 < f(-1) = 1$. Ello es debido a que $\mathbb{R} - \{0\}$ no es un intervalo con extremos finitos.

Se demuestra que si una función está definida y es continua sobre un intervalo I y derivable en el interior de I , y si su función derivada es estrictamente positiva en el interior de I , entonces la función es estrictamente creciente sobre I . Esto es una condición suficiente pero no necesaria como podemos ver si consideramos la aplicación $f : x \rightarrow f(x) = x^3$ de \mathbb{R} en \mathbb{R} . La función es derivable sobre \mathbb{R} y, $\forall x \in \mathbb{R}, f'(x) = 3x^2$, en particular, $f'(0) = 0$. Cero es el único punto donde se anula la derivada.

Por otra parte, para todo número real $t: t^2 + t + 1 = \left(t + \frac{1}{2}\right)^2 + \frac{3}{4} > 0$.

Así, si $x, y \in \mathbb{R}, x \neq y$, la factorización por x^2 si x es no nulo, y por y^2 si y es no nulo, da la desigualdad estricta $y^2 + xy + x^2 > 0$. Entonces, $\forall x, y \in \mathbb{R} | x < y$, tenemos $y^3 - x^3 = (y - x)(y^2 + xy + x^2) > 0$, por tanto $f(x) < f(y)$, es decir, la función $f(x) = x^3$ es estrictamente creciente sobre \mathbb{R} .

Dada una función f definida sobre un intervalo I y dado un punto a del interior de I , si f es derivable en a y admite un extremo relativo en a , entonces $f'(a) = 0$. Sin embargo, el recíproco

no es cierto.

La función $f : x \rightarrow f(x) = x^3$ de \mathbb{R} en \mathbb{R} es derivable sobre \mathbb{R} y, $\forall x \in \mathbb{R}, f'(x) = 3x^2$, en particular, $f'(0) = 0$. Se tiene así que $\forall x, y \in \mathbb{R} \mid x < 0 < y, f(x) < f(0) = 0 < f(y)$ por tanto no admite extremos relativos en cero. Tenemos así un ejemplo de una función derivable sobre \mathbb{R} cuya derivada en cero es nula pero que no admite extremo relativo en cero.

12. Integración

El cálculo integral apareció en la antigüedad con Arquímedes, con ocasión del cálculo de áreas y volúmenes. La noción de integral de una función surge con Isaac Newton y Gottfried Leibniz a finales del siglo XVII. Agustín Louis Cauchy es el primero en construir la integral, limitándose a funciones continuas sobre un intervalo. Bernhard Riemann propuso una construcción más general basada en las particiones. En 1902 Henri Lebesgue, partiendo del concepto de medida, definió una clase más amplia de funciones integrables.

Una función f definida sobre el segmento $[a, b]$ es integrable en el sentido de Riemann sobre $[a, b]$ si, para todo número real $\varepsilon > 0$, existen aplicaciones φ y ψ de $[a, b]$ en \mathbb{R} en escalera sobre $[a, b]$ tales que $\varphi \leq f \leq \psi$ sobre $[a, b]$ e $I_{[a,b]}(\varphi - \psi) \leq \varepsilon$.

Si f es una función definida e integrable en el sentido de Riemann sobre el segmento $[a, b]$, f está acotada sobre $[a, b]$ y la cota superior de los $I_{[a,b]}(\varphi)$ para $\varphi \leq f$ es escalera sobre $[a, b]$ es igual a la cota inferior de los $I_{[a,b]}(\psi)$ para $\psi \geq f$ en escalera sobre $[a, b]$, el valor común es la integral de f sobre $[a, b]$, $\int_a^b f(t) dt$.

En la práctica la casi totalidad de funciones usuales acotadas sobre un segmento son integrables en el sentido de Riemann. Sin embargo, se pueden encontrar ejemplos de funciones acotadas que no son integrables en el sentido de Riemann. Por ejemplo, la función $f : \mathbb{R} \rightarrow \mathbb{R}$

dada por $f(x) = \begin{cases} 1 & \text{si } x \text{ es racional} \\ 0 & \text{si } x \text{ es irracional} \end{cases}$ está acotada sobre \mathbb{R} sin embargo no es integrable en

el sentido de Riemann. En efecto. Para $a < b$ sea φ una aplicación de $[a, b]$ en \mathbb{R} en escalera sobre el segmento $[a, b]$ tal que $\varphi \leq f$ (resp. $\varphi \geq f$) sobre $[a, b]$. Elegimos una subdivisión $\sigma(x_0, x_1, \dots, x_n)$ de $[a, b]$ adaptada a φ y notamos, para todo $i \in [0, n-1]$, α_i el valor constante de φ sobre $]x_i, x_{i+1}[$. Si i perteneciese a $[0, n-1]$, el intervalo abierto $]x_i, x_{i+1}[$ contiene al menos un irracional u , así se tiene $\alpha_i = \varphi(u) \leq f(u) = 0$. Por tanto $I_{[a,b]}(\varphi) \leq 0$. Así cualesquiera que sean las aplicaciones φ y ψ de $[a, b]$ en \mathbb{R} en escalera sobre $[a, b]$ tales que $\varphi \leq f \leq \psi$ sobre $[a, b]$, se tiene $I_{[a,b]}(\psi - \varphi) = I_{[a,b]}(\psi) - I_{[a,b]}(\varphi) \geq b - a$. Como $b - a > 0$, f no es integrable en el sentido de Riemann sobre el segmento $[a, b]$.

Como f es nula salvo sobre un conjunto de medida nula, f es integrable en el sentido Lebesgue sobre $[a, b]$ y su integral es nula.

Si una función es integrable en el sentido de Riemann sobre un segmento, su cuadrado lo es también. Sin embargo el recíproco no es cierto. En efecto, la función $f : \mathbb{R} \rightarrow \mathbb{R}$ definida

por $f(x) = \begin{cases} 1 & \text{si } x \text{ es racional} \\ -1 & \text{si } x \text{ es irracional} \end{cases}$, al igual que en el ejemplo anterior, no es integrable en el

sentido de Riemann sobre $[0, 1]$. Sin embargo, f^2 es constantemente igual a 1 sobre $[0, 1]$, por tanto integrable en el sentido de Riemann en $[0, 1]$.

13. Probabilidades

Se considera a menudo que el cálculo de probabilidades comienza hacia la mitad del siglo XVI con los intercambios epistolares entre Pierre de Fermat y Blaise Pascal. Poco después, las obras de Christian Huygens, y después las de Jacques Bernoulli, establecen los primeros elementos de la teoría. Pierre Simeon Laplace publica en 1812 la teoría analítica de las probabilidades donde presenta la síntesis de esta teoría y en la que utiliza las nuevas herramientas matemáticas desarrolladas en el siglo XVIII. Las primeras teorías sobre los procesos estocásticos se desarrollan hacia 1900, en particular con Andrei Markov. Será necesario esperar hasta 1929 con los trabajos de Andrei Kolmogorov para asistir a la formalización de la teoría que usamos actualmente.

Sabemos que dos sucesos, A y B , son independientes si $p(A \cap B) = p(A)p(B)$ y, si n es un entero mayor o igual que 2, los sucesos A_1, A_2, \dots, A_n son independientes en conjunto si $p(A_1 \cap A_2 \cap \dots \cap A_n) = p(A_1)p(A_2) \dots p(A_n)$. Sin embargo existen sucesos independientes dos a dos que no lo son en su conjunto. Por ejemplo, si lanzamos un dado tetraédrico con cuatro caras numeradas de 1 a 4. Al considerar los sucesos A obtener 1 ó 2, B obtener 1 ó 3, C obtener 2 ó 3. Cada uno de estos sucesos tiene probabilidad de salir igual a $\frac{1}{2}$. Cada una de las intersecciones $A \cap B, A \cap C, B \cap C$ corresponde respectivamente a obtener 1, 2, 3. Su probabilidad es $\frac{1}{4} = \left(\frac{1}{2}\right)^2$, por tanto los sucesos son independientes dos a dos.

Sin embargo $A \cap B \cap C = \emptyset$ luego su probabilidad es cero: $0 \neq \left(\frac{1}{2}\right)^3 = \frac{1}{8}$.

Si nos dejamos guiar por la intuición de la noción de independencia, se podría esperar que un suceso independiente de otros dos, independientes ellos mismos entre sí, sea independiente su intersección. Sin embargo, con los mismos sucesos del experimento anterior, tenemos que el suceso A es independiente de B y de C , pero $B \cap C = \{3\}$ así $p(B \cap C) = \frac{1}{4}$ y $A \cap B \cap C = \emptyset$. Tendremos así que $p(A \cap (B \cap C)) = 0 \neq \frac{1}{8} = p(A)p(B \cap C)$.

A es independiente de los sucesos B y C , independientes entre sí pero no es independiente de su intersección.

Una experiencia aleatoria es en general interesante porque de ella se sigue un beneficio o una pérdida, es decir, se asocia a cada resultado un número real. Esto lleva a las definiciones de variable aleatoria discreta y continua, función de densidad, variable aleatoria de densidad, ley de probabilidad. Asociada a las variables aleatorias puede venir la esperanza matemática pero ¿existe siempre?

No, por ejemplo si consideramos la variable aleatoria discreta x tomando sus valores en el conjunto $\{2^n, n \in \mathbb{N}^*\}$ y definida por $p(x = 2^n) = \frac{1}{2^n}$, es una ley de probabilidad pues $\sum_{n=1}^{\infty} \frac{1}{2^n} = 1$, sin embargo $2^n p(x = 2^n) = 1, \forall n \geq 1$, así la serie $\sum_n 2^n p(x = 2^n)$ diverge, por tanto x no tiene esperanza matemática.

Este resultado es conocido como paradoja de San Petesburgo y fue enunciado por Nicolás Bernoulli (1685 – 1726) de la siguiente forma: Un jugador participa en el juego siguiente. Se lanza una moneda repetidamente hasta que aparezca la primera cara. Una vez aparece gana un centavo si la cara aparece en el primer lanzamiento, 2 centavos si aparece en el segundo, 4 si aparece en el tercero, ... doblando en premio en cada lanzamiento. Así ganaría 2^{k-1} centavos si la moneda debe lanzarse k veces, ¿cuánto estaría dispuesto a poner para participar en el juego?

Como la esperanza matemática es infinita, no importa cuánto pague para entrar en el juego, saldría ganando a largo plazo.

La variable aleatoria también puede ser que tenga función de densidad e igualmente no tener esperanza matemática. Así, la variable aleatoria x con función de densidad $f: \mathbb{R} \rightarrow \mathbb{R}$

dada por $f(t) = \frac{1}{\pi(1+t^2)}$; como $f(t) \sim_{t \rightarrow \infty} \frac{1}{\pi} x \frac{1}{t^2}$ y es par, f es integrable sobre \mathbb{R} por el criterio de Riemann.

La función $F(t) = \frac{1}{\pi} \arctan$ es una primitiva de f , así

$$\int_{-\infty}^{+\infty} f(t) dt = \frac{1}{\pi} \left(\lim_{+\infty} \arctan - \lim_{-\infty} \arctan \right) = \frac{1}{\pi} \left(\frac{\pi}{2} - \frac{-\pi}{2} \right) = 1$$

luego f es una densidad de probabilidad. Sin embargo la esperanza de x no está definida pues $\tan f(t) \sim_{t \rightarrow +\infty} \frac{1}{\pi} x \frac{1}{t}$ no es integrable sobre \mathbb{R} .

La ley de probabilidad definida por la función f es conocida como ley de Cauchy.

Referencias

- [1] APOSTOL, T. M. *Análisis Matemático*, Editorial Reverté, Barcelona, 1974.
- [2] BOYER, C. B. *Historia de la matemática*, Alianza Editorial, Madrid, 1986.
- [3] BOURBAKI, N. *Elementos de historia de las matemáticas*, Alianza Editorial, Madrid, 1976.
- [4] COLLETE, J. P. *Historia de las matemáticas*, Siglo XXI Editorial, Madrid, 1985.
- [5] KLINE, M. *El pensamiento matemático de la antigüedad a nuestros días*, Alianza Editorial, Madrid, 1992.
- [6] LORENZO (DE), J. *La matemática y el problema de su historia*, Editorial Technos, Madrid, 1977.
- [7] NEWMAN, J. R. *El mundo de las matemáticas*, Editorial Grijalbo, Barcelona, 1985.
- [8] QUEYSANNE, M. *Álgebra Básica*, Editorial Vicens Vives, Barcelona, 1973.
- [9] SPIVAK, M. *Calculus*, Editorial Reverté, Barcelona, 1978.
- [10] TATON, R. *Historia general de la ciencia*, VIII, Editorial Orbis, Barcelona, 1988.

Sobre el autor:

Nombre: Antonio Rosales Góngora

Correo electrónico: anrogo58@yahoo.es

Institución: Instituto de Educación Secundaria Bahía de Almería, Almería, España.