Universidad de Valladolid
Facultad de Ciencias

Trabajo Fin de Grado

Grado en Matemáticas

El teorema de derivación de Lebesgue, un hito del Análisis Real.

Autora: Nuria Díaz Cuadrado

Tutor: Luis A. Tristán Vega

D. Luis Alberto Tristán Vega, Profesor Titular de Análisis Matemático de la Universidad de Valladolid,

CERTIFICA:

Que el presente trabajo, **“El teorema de derivación de Lebesgue, un hito del Análisis Real”**, ha sido realizado bajo su dirección en el Departamento de Álgebra, Análisis Matemático, Geometría y Topología, por Doña Nuria Díaz Cuadrado, y constituye su Trabajo Fin de Grado para optar al título de Grado en Matemáticas.

Que le consta que el trabajo es original e inédito y, que autoriza su presentación.

Y para que conste a los efectos oportunos, firma la presente en Valladolid a veintidós de junio de dos mil dieciocho.

Fdo.: Luis A. Tristán Vega

A mi familia.
Especialmente, a mi padre.

Índice general

Introducción

Las matemáticas en general y, especialmente el Análisis Matemático, han evolucionado espoleadas por los avances en las ciencias natulares. Muchos logros realizados han surgido como respuesta a retos concretos proporcionados por la Física, Mecánica, etc. En esta memoria prestamos atención a uno de estos procesos: la idea de Fourier, que pensaba que toda serie era derivable término a término, la respuesta de Weierstrass buscando el rigor en esta teoría, proporcionando un contraejemplo y creando gran incertidumbre en la comunidad matemática y, por último, el teorema de derivación de Lebesgue, es decir, “una función monótona es diferenciable en casi todo punto.”

Para la presentación de este proceso, necesitamos poner los hechos en contexto desde el surgimiento de la idea de derivada hasta la formalización moderna de este concepto.

La formación del cálculo diferencial (e integral) surgió casi simultáneamente en dos formas diferentes:

i) La teoría de las fluxiones en los trabajos de Newton.

ii) El cálculo de las diferenciales de Leibniz.

Aunque ambos se apoyaron en la experiencia de los numerosos predecesores.

En el capítulo 1 vamos a ver a partir de Newton y Leibniz como diversos autores han ayudado a esclarecer la noción de derivada y como lo han relacionado con otros conceptos como: la continuidad, el límite, la convergencia de series, etc.

De hecho, durante mucho tiempo se creyó que las funciones continuas tenían que ser derivables excepto, quizás, en conjuntos aislados de puntos. Sin embargo, después del trabajo de Fourier y las controversias que despertó sobre cómo debían ser las funciones para que hubiera convergencia de, lo que actualmente llamamos “series de Fourier”, se empezaron a estudiar funciones “raras”, como las continuas que no son derivables en ningún punto. En la actualidad, el Análisis Funcional permite probar que estas funciones son numerosas. De hecho, el teorema de categoría de Baire asegura la abundancia de funciones continuas y no diferenciables en ningún punto, aunque construirlas es, en general, un trabajo muy costoso.

El ejemplo que atrajo la mayor atención sobre la distinción entre continuidad y diferenciabilidad se debe a Weierstrass. En el capítulo 3 estudiaremos esta función exhaustivamente, y veremos otros ejemplos anteriores y posteriores al de Weiestrass, y además como abrió la puerta a otro tipo de patologías.

Debido a estos tipos de funciones se convirtió en apremiante el problema de determinar cuándo una función continua puede ser diferenciable. El periodo de incertidumbre creado por la profusión de ejemplos patológicos fue resuelto por el teorema de Legesgue que expondremos en el capítulo 4 y sus consecuencias, que ayudaron a restablecer la confianza en la armonía del Análasis Matemático. En palabras de Riesz ([27] pag. 113):

"El teorema de Lebesgue es uno de los más llamativos y más importantes en la teoría de la variable real".

Otro proceso revolucionario en el desarrollo del Análisis, es la aparición de las funciones generalizadas como respuesta a la formulación cuántica de Dirac; los trabajos de Sóbolev y Schwartz establecen el marco adecuado donde alcanzan pleno sentido y rigor las denominadas distribuciones. En el capítulo 2 presentamos de forma concisa la teoría básica de distribuciones, así mismo, en ese capítulo daremos una visión de otras nociones de derivación.

Capítulo 1

Un poco de historia

A lo largo de la historia, encontramos varios ejemplos de conceptos matemáticos que han pasado a través de numerosos cambios.

En este capítulo nos vamos a centrar en dos: la noción de función y la noción de derivada.

El análisis de la evolución histórica del concepto de función nos muestra que éste es muy complejo, y para llegar a la definición actual se tuvo que pasar por un proceso largo y laborioso. Durante su evolución, se encuentran diferentes definiciones y cada una de ellas corresponde a diferentes niveles de abstración.

Por otro lado, el concepto de derivada surge de dos tipos distintos de problemas:

i) el problema físico, que consiste en buscar la velocidad instantánea de una partícula móvil.

ii) el problema geométrico, que consiste en buscar la recta tangente a una curva en un punto dado.

ambos conducen de forma muy natural a la noción de derivada.

Este concepto fue utilizado de manera implícita por Fermat, pero Newton y Leibniz lo formalizaron de forma más precisa. Su desarrollo completo hasta la visión actual requirió de la contribución de muchos matemáticos.

1.1. Sobre el concepto de función

La palabra *"función"* apareció por primera vez en un manuscrito de Gottfried von Leibniz (1646-1716) en el año 1673, donde hablaba todavía de lo relativo entre ordenadas y abscisas. En otros tratados publicados entre 1692 y 1694 utiliza la palabra *"funciones"* para designar cualquier clase de segmentos rectos (abscisas, ordenadas, cuerdas, secciones de tangentes, subtangentes, normales, subnormales) que dependen de un punto fijo o de los puntos de una curva dada. Por otra parte, Johann Bernoulli (1667-1748) al mantener correspondencia con Leibniz sobre este tema, y muchos otros, en 1718 definió el concepto de función de forma explícita ([32], pag. 80):

"Se llama aquí a una función de una variable una cantidad compuesta de cualquier manera de esta variable y constantes."

Nota. Johann Bernoulli al principio escribía ξ o X para denotar una función general de x, aunque en 1718 cambió a φx.

Leonhard Euler (1707-1783) define con mayor precisión el concepto de función introducido por Johann Bernoulli en su libro *Introductio in analysin infinitorum* ([10], pag. 192):

"Una función de una cantidad variable es una expresión analítica compuesta de cualquier manera a partir de la cantidad variable y de números o cantidades constantes."

Euler clasifica las funciones de la siguiente manera:

i) Funciones algebraicas, que son aquellas en las que sólo están permitidas operaciones algebraicas sobre la variable independiente y distingue, además, dos clases de funciones algebraicas: la función racional, que implica las cuatro operaciones aritméticas habituales, y la función irracional, que incluye además la radicación.

ii) Funciones trascendentes, que son aquellas que vienen dadas por una serie infinita, pues requiere repeticiones ilimitadas de las funciones algebraicas. Estas funciones son las logarítmicas, exponenciales y trigonométricas.

Por tanto, el significado aproximado de *expresión analítica* engloba estas funciones. Además, para Euler todas las funciones podían ser desarrolladas por una serie de potencias.

Uno de los escollos en la teoría de funciones de Euler fue distinguir entre función y su representación, para Euler una función continua se expresaba mediante una única expresión analítica, una función mixta se expresaba en términos de dos o más expresiones analíticas, y una función discontinua incluía funciones mixtas, pero eran más generales.

La representación de funciones se desarrolló, en parte, gracias a la controversia suscitada a propósito del problema de la cuerda vibrante entre Euler y Jean D'Alembert (1717-1783). En 1746 D'Alembert publicó la solución al problema de una cuerda tensa que vibra, diciendo que la función que determina la velocidad de cada uno de los puntos de la cuerda debían de ser expresados mediante una sola expresión analítica, pero Euler no estaba de acuerdo.

Debido a este debate de la cuerda vibrante, Euler dio otro concepto de función de naturaleza más general, que se encuentra en su *Institutiones calculi differentialis* de 1755 ([31], pag. 121):

"Si algunas cantidades dependen de otras cantidades de modo que si las últimas cambian, las primeras también lo hacen, entonces las primeras cantidades se llaman funciones de las últimas. Esta denominación es de naturaleza amplia e inlcuye cada método por el cual una cantidad pudiera ser determinada por otras. Si, por consiguiente, x denota una cantidad variable, entonces toda cantidad la cual dependa de x en cualquier manera o esté determinada por ella se llama función de ella."

Esta definición se aleja de su primera definición, centrada más en las expresiones analíticas que definen una relación entre las variables, dado que se presta mayor atención a la relación que guardan las variables. Además, conviene destacar que define explícitamente una dependencia entre variables, concepto clave para la teoría moderna de las funciones. Sim embargo, a pesar de esta definición, Euler desarrolló *Institutiones calculi differentialis* usando sólo funciones analíticas.

Nota. Euler introdujo en 1734 la notación de función, $f(x)$, que hoy en día utilizamos.

Posteriormente cuando Daniel Bernoulli (1700-1782) lee los artículos de D'Alembert y Euler sobre la cuerda vibrante, él supone que puede existir simultáneamente muchos modos de oscilación en la cuerda vibrante, e insiste en que todas las posibles curvas se pueden representar de la forma:

$$f(x) = \sum_{n=1}^{\infty} a_n \sin\left(\frac{nx\pi}{l}\right)$$

porque existen suficientes constantes a_n, como para que la serie se ajuste a cualquier curva. En consecuencia, afirma, que todos los correspodientes movimientos vendrán dados por la serie infinta:

$$y(t,x) = \sum_{n=1}^{\infty} a_n \sin\left(\frac{nx\pi}{l}\right)\cos\left(\frac{nct\pi}{l}\right)$$

de donde:

$$y(0,x) = \sum_{n=1}^{\infty} a_n \sin\left(\frac{nx\pi}{l}\right) = f(x).$$

Euler le replica que la cuestión principal radica en preguntarse si todas las curvas del conjunto de la cuerda en movimiento están contenidas o no en esta expresión analítica, pues opinaba que la determinación de los coeficientes *"es sin duda muy difícil, por no decir imposible"*.

El intento de Euler de buscar argumentos rigurosos para determinar si cualquier función, por arbitraria que fuese, admite una representación mediante una serie trigonométrica, llegó a oídos de Joseph-Louis Lagrange (1736-1813), quien en 1759 publicó el trabajo *"Investigaciones sobre la naturaleza y la propagación del sonido"*.

Tanto Euler, como D'Alembert y Daniel Bernoulli elogiaron la originalidad de los planteamientos de Lagrange y la habilidad matemática del trabajo, pero criticaron aspectos de los cálculos efectuados. Puede ser que esos comentarios adversos influyesen en la obsesión de Lagrange por probar que las funciones analíticas son precisamente aquellas que se pueden expresar por series de potencias y con esto trata de fundamentar algebraicamente el análisis infinitesimal.

En su libro *"Théorie des fonctions analytiques"* de 1797, define la función como ([31], pag. 128):

"Llamamos función a toda expresión matemática de una o varias cantidades con las cuales éstas aparecen de cualquier manera, relacionadas o no con algunas otras cantidades que son consideradas como constantes, mientras las cantidades de la función pueden tomar todos los valores posibles."

Por lo tanto, Lagrange va a operar sus razonamientos solo con expresiones analíticas al estilo de Euler.

Sin embargo, Joseph Fourier (1768-1830) empezó a esclarecer el concepto de función. Fue el primero en ver la diferencia entre *función* y *expresión analítica*, debido al estudio de la ecuación del calor. Para él, una función podía ser representada por más de una expresión analítica. Sobre la noción de función dice en *Théorie analytique de la chaleur* de 1822 ([32], pag. 80):

"En general, la función $f(x)$ representa una sucesión de valores u ordenadas, cada una de las cuales es arbitraria. Cuando se da una infinidad de valores a la abscisa x, hay un número igual de ordenadas $f(x)$. Todos tienen valores numéricos reales, ya sean positivos o negativos o nulos. No suponemos que estas ordenadas estén sujetas a la ley común; se suceden una a la otra de cualquier manera, y cada una de ellas se da como si fuera una sola cantidad."

De hecho, sólo trató funciones con un número finito de discontinuidades en cualquier intervalo finito. Por otro lado, Fourier apoyaba que la función debía ser representable por una expresión analítica, concretamente por una serie trigonométrica. Es decir, sus series pueden representar funciones que tienen diferentes expresiones analíticas en diferentes partes del intervalo $(-\pi, \pi)$.

Fourier añade también que las expresiones diferentes para cada parte del intervalo dado pueden corresponder a una gráfica compuesta de curvas juntas o disjuntas en ese intervalo.

Por su parte, Augustin-Louis Cauchy (1789-1857) en 1821 proporcionó una definición de función en su obra *Cours d'Analyse* que hace la dependencia entre variables sea central para el concepto ([20], Vol.3 pag. 1254):

"Cuando se relacionan cantidades variables entre ellas de modo que estando dado el valor de una de éstas, se puedan determinar los valores de todas las otras, ordinariamente se concibe a estas cantidades diversas expresadas por medio de la que está entre ellas, la cual entonces toma el nombre de variable independiente; y las otras cantidades expresadas por medio de la variable independiente son aquellas que uno llama funciones de esta variable."

Además, Cauchy también es explícito en que una serie infinita es una manera de especificar una función. Sin embargo, no se requiere una expresión analítica para una función.

La definición que más se acerca a la conocida hoy en día fue dada por Lejeune Dirichlet (1805-1859) en 1837 ([32], pag. 80):

"Supongo que a y b son dos valores definidos y x es una cantidad variable que supone asumir, gradualmente, todos los valores localizados entre a y b. Ahora bien, si a cada x le corresponde un único y finito y, entonces se llama a y la funcion de x para este intervalo. Además, no es necesario, que y dependa de x en todo este intervalo de acuerdo con la misma ley; ni es necesario pensar sólo en relaciones que pueden expresarse mediante operaciones matemáticas."

Esto es debido, a que Dirichlet se interesó por el desarrollo de las funciones en series trigonométricas de Fourier.

A Dirichlet le debemos la introducción de la función :

$$D(x) = \begin{cases} 1 & si \quad x \in \mathbb{Q} \\ 0 & si \quad x \in \mathbb{I}. \end{cases}$$

Sin embargo, el ejemplo de Dirichlet no fue considerado como función "válida" por la mayoría de la comunidad matemática hasta que en el año 1867, G.F. Bernhard Riemann (1826-1866) encontró uso para las funciones raras como la de Dirichlet, y facilitó una definición de función ([32], pag. 81):

"Sea D un conjunto de números reales. Una función $f\colon D \to \mathbb{R}$ es una regla que asigna un número $f(x)$ a cada elemento x de D."

Con el surgimiento de la teoría de conjuntos en la década de los 70 del siglo XIX se abrió la posibilidad de expresar una función como relación entre dos conjuntos abstractos no necesariamente numéricos.

La definición que más se repite en varios textos, fue aportada por el grupo Bourbaki en 1939 ([31], pag. 163-164):

"Sean X e Y dos conjuntos, diferentes o no. Una relación entre una variable x de E y una variable y de F se dice relación funcional de E hacia F, si, cualquiera que sea x de E, existe un elemento y de F, y uno solo, que esté en la realación considerada con x.

Se da el nombre de función a la operación que asocia así a todo elemento x de E el elemento y de F que se encuentra en la relación dada con x; se dice que y es el valor de la función para el elemento x, y que la función esta determinada por la relación funcional considerada."

1.2. Evolución de la noción de Derivada

La noción de derivada no se formuló hasta el siglo XVII (al menos en Occidente), cuando el matemático francés Fermat trato de examinar los máximos y mínimos de ciertas funciones. Aunque a finales del siglo este concepto estaba más esclarecido debido a Newton y Leibniz.

Sin embargo, seguía habiendo bastante confusión, pues muchos relacionaban la continuidad y diferenciabilidad. De hecho, hasta principios del siglo XIX se pensaba que toda función continua era derivable. Tal fue este pensamiento, que en 1806 Ampère intento probar la existencia general de las derivadas, pero su prueba era dificil de evaluar porque no está claro qué suposiciones implícitas estaba haciendo sobre lo que constituye una función.

A continuación, veremos la evolución del concepto de derivada a través de diversos autores. Desde luego no son todos, pero sí los más significativos, en lo que atañe al objetivo de esta memoria.

1.2.1. Newton

Isaac Newton (1642-1727) comenzó sus descubrimientos a mediados de la década de 1660, cuando era un estudiante en la Universidad de Cambridge. Por razones psicológicas y científicas, no compartió sus descrubrimientos en papel, por eso sus resultados no se publicaron durante décadas, como mucho se lo enviaba en cartas a sus amigos.

Su maestro fue Isaac Barrow (1630-1677), que fue quien le enseñó las ideas básicas de las matemáticas infinitesimales.

Las direcciones fundamentales de la actividad científica de Newton fueron la física, la mecánica, la astronomía y las matemáticas. Durante los años 1665 y 1666 hizo cuatro de sus principales descubrimientos:

1) El teorema binomial

2) El cálculo

3) La ley de la gravitación

4) La naturaleza de los colores

Referente al cálculo, Newton hizo varios escritos. Los tres más importantes son:

i) *"De Analysi per Aequationes Numero Terminorum Infinitas"* (sobre el análisis por medio de ecuaciones con un infinito número de términos). Escrito en 1669 y publicado en 1711. Este trabajo puede considerarse el escrito fundamental del cálculo, aquí Newton utiliza el método de los infinitesimales.

   Newton en este libro, no sólo proporcionó un método general para obtener el cambio relativo de una variable con respecto a otra, sino que mostró que el área puede obtenerse invirtiendose el proceso de obtener un cambio relativo.

ii) *"Methodus Fluxionum et Serierum Infinitarum"* (método de las fluxiones y series infinitas). Escrito en 1672, aunque publicado en 1736.

   En el método de las fluxiones se estudian las varibles, introducidas como abstracción de las diferentes formas del movimiento mecánico continuo (puntos, rectas y planos). A estas varibles se las llama *fluentes o fluyentes*, son variables dependientes y tienen un argumento en común, el tiempo. A su cambio relativo, es decir, cuando se introduce la velocidad de la corriente del fluente, se llama *fluxión* (lo que llamamos la derivada con relación al tiempo). Su notación es $\dot{x}$ para la fluxión de x. Como la fluxión constituye una variable, podemos encontrar las fluxiones de las fluxiones, es decir, la fluxión de $\dot{x}$ es $\ddot{x}$.

   Para el cálculo de las velocidades instantáneas, es decir, de las fluxiones, se exigían variaciones infinitesimales de los fluentes, denominados por Newton *momentos*. El símbolo del momento del tiempo es o, por tanto $\dot{x}o$ es el momento del fluente x.

   Los símbolos de Newton no son tan cómodos como los de Leibniz (que veremos más adelante), no obstante se siguen conservando en la mecánica.

   En la teoría de fluxiones se resuelven dos problemas principales, formulados tanto en términos mecánicos como en términos matemáticos. Veamos la formulación en términos matemáticos:

   ii.1) Determinar la relación entre las fluxiones dada la relación entre los fluentes, es decir, el problema de la diferenciación implícita de funciones.

ii.2) Determinar la relación entre los fluentes dada la relación entre las fluxiones, es decir, el problema de la integración de las ecuaciones diferenciales.

Nota. Como este trabajo trata de las derivadas, solo vamos a ver el primer problema. Para el segundo problema propuesto, si se desea tener algún conocimiento, consultar ([20], Vol.1 pag. 479).

Para el primer problema, Newton introdujo una regla uniforme: el algoritmo de diferenciación de funciones. Este algoritmo lo aclara con ejemplos. Veamos la relación entre $\dot{x}$ e $\dot{y}$ sabiendo que $y = x^n$.

Primero formemos la misma relación para los fluentes después de experimentar una variación instantánea, esto es, cuando en cada fluente se añade su momento

$$y + \dot{y}o = (x + \dot{x}o)^n,$$

segundo, desarrollando el lado derecho mediante el teorema binomial

$$y + \dot{y}o = x^n + nx^{n-1}\dot{x}o + \frac{n(n-1)}{2}x^{n-2}(\dot{x}o)^2 + \cdots$$

ahora, restando $y = x^n$, dividiendo por o y despreciando todos los términos que todavía contienen o, nos queda:

$$\dot{y} = nx^{n-1}\dot{x}.$$

En notación moderna puede escribirse como

$$\frac{dy}{dt} = nx^{n-1}\frac{dx}{dt}.$$

iii) *"Tractatus de Quadratura Curvarum"* (tratado sobre la cuadratura de las curvas). Escrito en 1676, pero no publicado hasta 1704.

En este trabajo, trata de evitar las cantidades infinitamente pequeñas como las cantidades fluentes, reemplazándolas por una teoría de las llamadas *razones primeras y últimas*. El método de la razón primera y última significa lo siguiente:

Considera la función $y = x^n$. Para obtener la fluxión de y o de x^n, se deja a x "fluir"hasta $x + o$. Entonces x^n se convierte en

$$y + o = (x + o)^n$$

$$y + o = x^n + nx^{n-1}o + \frac{n(n-1)}{2}x^{n-2}o^2 + \cdots$$

Los incrementos de x e y son, uno con respecto a otro (dividiendo ambos por o), como

$$1 \quad \text{es a} \quad nx^{n-1} + \frac{n(n-1)}{2}x^{n-2}o + \cdots$$

Haciendo tender a cero los incrementos, la última relación se convertirá en

$$1 \quad \text{es a} \quad nx^{n-1}.$$

Newton caracterizó la última relación como *"la relación de las cantidades no antes de que desaparezcan, ni después, sino con las que desaparecen"*. Al percibir esto como un cociente cuyo numerador y denominador se estaban reduciendo a cero, es decir, Newton quería capturar la relación en el instante preciso cuando el numerador y denominador desaparecen simultáneamente (Newton está aquí realmente cerca del concepto de límite).

Nota. D'Alembert fue el primero en observar que Newton tenía la noción "correcta" de la derivada, y en la *Encyclopédie* dice sobre la derivada que debe basarse en el límite de la razón de las diferencias de variables dependientes e independientes (esta versión es una reformulación de las razones primera y última de Newton).

1.2.2. Leibniz

Sin ánimo de polémica sobre la autoría del cálculo infinitesimal, diremos que Leibniz fue quien enseñó el cálculo al mundo, puesto que lo desarrolló en pocos años y, a diferencia de Newton, mantuvo una comunicación fluida con sus colegas, haciéndoles partícipes de sus descubrimientos y debatiendo sobre ellos.

En 1684 publicó por primera vez un escrito sobre cálculo diferencial en la revista *"Acta Eruditorum"* bajo el título *"Nova methodus pro maximis et minimis, itemque tangetibus, qua nec irrationales quantitates moratur"* (Nuevo método para los máximos y mínimos, asi como para las tangentes, el cual puede también aplicarse a las cantidades fraccionarias e irracionales).

Antes de hacer esta publicación, por el año 1673 era consciente del importante problemo directo e inverso de obtener tangentes a las curvas, también estaba bastante seguro de que el método inverso era equivalente al de obtener áreas y volúmenes mediante sumaciones. El desarrollo de sus ideas comienzan con notas de 1675, que no fueron publicadas.

En un manuscrito fechado el 11 de Noviembre de 1675, titulado *"Ejemplos del método inverso de las tangentes "*, utiliza $\int$ para la suma y x/d para la diferencia, diciendo que x/d es dx, la diferencia de dos valores consecutivos de x aunque, aparentemente aquí dx es una constante e igual a la unidad.

Nota. $\int$ es la estilización de la letra latina S en contraposición a la griega Σ, usada para sumas discretas.

A partir de este pensamiento, Leibniz afirma que *"la integración en cuanto a proceso de sumación es la inversa de la diferenciación"*, y como tiene que establecer una relación entre dx y dy (las diferencias de x e y respectivamente), se interesa por las expresiones del tipo $dx/dy,\ d(uv),\ d(u/v)$ etc, y pensó que $d(uv) = dudv$.

Por tanto, en una nota manuscrita del 26 de Junio de 1676, Leibniz afirma que la mejor manera de encontrar las tangentes es determinando dy/dx, donde dy y dx son diferencias, y dy/dx es un cociente. Se da cuenta, que la determinación de la tangente a la curva depende de la razón de las diferencias de ordenadas y abscisas cuando se hacen infinitamente pequeñas.

En Noviembre de ese mismo año, da las reglas generales de $dx^n = nx^{n-1}$ para n entero, mientras que el 11 de Julio de 1677 consigue, con dificultad, establecer las reglas exactas de la diferencial de la suma, diferencia, producto y cociente de dos funciones y para las potencias y raíces.

Hacia el año 1680, el dx se había convertido en la diferencia de las abscisas y dy en la diferencia de las ordenadas. Llama dy al incremento momentáneo en y cuando la ordenada se desplaza a lo largo del eje de las x.

Gracias a todas estas notas, en su manuscrito de 1864 presenta las reglas más generales de la diferenciación y se sirve de las diferenciales más que de las derivadas por medio de la notación d.

En este manuscrito (que no pasa de 10 páginas), las diferenciales $dx,\ dy$ se definen como incrementos finitos, pero en una ocasión define dy como

$$dy : dx = y \ \ : \ \text{subtangente}$$

que es el germen de la interpretación gráfica de la derivada como pendiente de una recta.

Nota. En un triángulo rectángulo, las rectas subtangente y subnormal mencionadas anteriormente en el concepto de función, son las longitudes de los dos segmentos resultantes de la división de la hipotenusa al ser cortada por la proyección ortogonal desde el vértice opuesto, dónde los dos catetos del triángulo yacen sobre la recta tangente y la recta normal, respectivamente.

También aporta las fórmulas para la diferencial de la suma, diferencia, producto y cociente de dos funciones, para las potencias y raíces y la regla de dx^n, todo ello acompañado de aplicaciones geométricas tales como la busqueda de la tangente, de los máximos o mínimos, y de los puntos de inflexión.

En este escrito no proporciona justifiaciones claras, es decir, no publica ninguna demostración, sin embargo a sus colegas les manda cartas con esos intentos. Por ejemplo, un fragmento de la carta que envia a John Wallis (1616-1703) el 30 de Marzo de 1690 dice ([23], pag. 447):

"Es útil considerar las cantidades como infinitamente pequeñas, de manera que, cuando se busca su relación, son omitidas en lugar de verse como cero, cuando aparecen junto a cantidades que son incomparablemente mayores. Si tenemos $x + dx$, entonces dx se omite. Del mismo modo, no podemos dejar que xdx y $(dx)^2$ estén al lado del otro. Así, si tenemos que diferenciar xy, escribimos:

$$(x + dx)(y + dy) - xy = xdy + ydy + dxdy,$$

pero aquí debería omitirse $dxdy$, pues es incomparablemente menor que $xdy + ydx$. Por lo tanto, en cada caso particular, el error es menor que cualquier cantidad finita."

En 1686 publicó un segundo manuscrito también en *"Acta Eruditorum"* bajo el título *"De geometria recondita et analysi indivisibilium atque infinitarum"* (sobre una geometría recóndita y el análisis de los indivisibles e infinitos), este escrito trata del cálculo integral dónde se muestra que las cuadraturas son un caso especial del método inverso de las tangentes.

En 1695 publicó la regla de diferenciación de la función exponencial general y la formula de la diferenciación multiple del producto, que hoy en día lleva su nombre:

$$(uv)^{(n)} = u^{(n)}v^{(0)} + nu^{(n-1)}v^{(1)} + \cdots nu^{(1)}v^{(n-1)} + u^{(0)}v^{(n)},$$

en la actualidad

$$(fg)^{(n)} = \sum_{k=0}^{n} \binom{n}{k} f^{(k)} g^{(n-k)}.$$

El talón de aquiles del cálculo de Leibniz fue la "igualdad" $x+dx = x$, cuya mayor crítica vino de París. Por tanto, el 2 de Febrero de 1702 escribió una carta detallada a su aliado de París, Varignum, que la publicó inmediatamente y años más tarde ya no había ninguna objeción a ello en el continente europeo.

Leibniz facilitó muchas justificaciones. Argumentó que las distancias infinitas e infinitamente pequeñas pueden usarse como conceptos ideales no como objetos reales. Es más conveniente introducir el concepto de incomparablemente pequeño, en vez de hablar siempre de magnitudes capaces de disminuir ilimitadamente. Con eso se justifican los cálculos que Leibniz mandó a Wallis, afirmando que el error es invariablemente inferior a cualquier magnitud asignable.

Para Leibniz, las reglas algebraicas que llevan los números reales a los diferenciales, se sostiene del finito al infinito y viceversa, aunque escribió muy poco sobre la validez de la reglas. Estas reglas son:

i) $d\,(ax) = adx$.

ii) $d\,(z - y + uv + x) = dz - dy + d\,(uv) + dx$.

iii) $d\,(xv) = xdv + vdx$.

1.2.3. L'Hôpital

En 1696 apareció en París el primer libro sobre el cálculo diferencial, *"Analyse des infinitement petits pour l'intelligence des lignes courbes"*, escrito por el Marqués Guillaume de L'Hôpital (1661-1704). En este libro L'Hôpital expresó las ideas de Leibniz de una forma más organizada y más clara. Gran parte del material adquirido para escribir el libro se lo entregó Johann Bernoulli. Además Bernoulli le proporcionó conferencias a cambio de que L'Hôpital le pagara una cuota. L'Hôpital nunca engañó a nadie, siempre dijo que sus fuentes de trabajo fueron Leibniz y Bernoulli, y refiriéndose a ellos escribió ([12], pag. 5):

"He hecho uso libre de sus descubrimientos para que yo francamente les devuelva lo que les plazca reclamar como suyos."

La definición de diferencial que L'Hôpital dio fue la siguiente ([12], pag. 5):

"La parte infinitamente pequeña por la cual una cantidad variable aumenta o disminuye continuamente se llama el diferencial de esa cantidad."

En la notación moderna, si $y = y(x)$ y si x está incrementado por su diferencial infinitamente pequeño dx, entonces el correspondiente cambio infinitamente pequeño en y es $dy = y(x+dx) - y(x)$, y entonces $y(x+dx) = y + dy$.

L'Hôpital utilizó la noción de infinitamente pequeño de Leibniz y la idea de las cantidades que desaparecen de Newton para derivar los resultados claves del cálculo diferencial, entre ellos la regla del producto y más tarde la regla del cociente.

Como dato curioso, tenemos que en este libro se encuentra la primera versión de *"la relga de L'Hôpital"*, donde no se utilizaba la noción de límite, y también se haya su primer ejemplo

$$\frac{\sqrt{2a^3x - x^4} - a\sqrt[3]{a^2x}}{a - \sqrt[4]{ax^3}} \quad \text{cuando} \quad x = a$$

cuya respuesta fue $16a/9$.

Dicha regla fue descubierta por Bernoulli, y fue incluida en los materiales que le entregó a L'Hôpital. Hubo una sugerencia de llamarla *"la regla de Bernoulli"*, pero el historiador matemático Dirk Struik comentó ([12], pag. 7):

"Que el buen Marqués mantenga su elegante regla. Él pagó por ella".

1.2.4. Euler

Euler fue el matemático más prolífico del siglo XVIII y probablemente de todos los tiempos. Su nombre se encuentra en todas las ramas de las matemáticas. Su obra más célebre fue *"Introductio in analysim infinitorum"* de 1748. En esta obra se encuentran las fórmulas que relacionan las funciones trigonométricas y exponenciales, expresiones del seno y del coseno en forma de productos infinitos, y el uso de los símbolos e,i,π, etc.

En lo relacionado con las diferenciales, Euler considera que las nociones de infinitamente grande e infinitamente pequeño no son tan difíciles de entender, y propone dar unas explicaciones que cree que son suficientes para entenderlo.

Una cantidad infinitamente pequeña no es otra cosa que una cantidad que va disminuyendo y, consecuentemente, es igual a cero. Por eso, el cálculo de lo infinitamente pequeño consiste para Euler en el estudio de las razones geométricas de las cantidades infinitamente pequeñas.

Por ejemplo, si dx es una cantidad infinitamente pequeña, entonces $dx = 0$ así como $adx = 0$ (donde a es una cantidad finita cualquiera), y su razón geométrica adx/dx es finita, es decir $a/1$, por este motivo según Euler, dos cantidades infinitamente pequeñas no se pueden confundir cuando se estudia su razón. Además, como lo infinitamente pequeño es de hecho nada, para Euler una cantidad finita no cambia si se le añade o quita una cantidad infinitamente pequeña.

Sea a una cantidad finita y dx una cantidad infinitamente pequeña, entonces $a \pm ndx = a$ y cuya razón geométrica es

$$\frac{a \pm ndx}{a} = 1.$$

Se deduce una regla aceptada según la cual ([10], pag. 195):

"las cantidades infinitamente pequeñas tienden a cero en comparación con las cantidades finitas, y, además, pueden ser despreciadas cuando están implicadas estas cantidades finitas."

Según Euler, el cuadrado, el cubo, etc de una cantidad infinitamente pequeña igual a cero, tambien serán igual a cero, y estas cantidades tienden a cero en comparación con cantidades finitas. Además, lo mismo ocurrirá con la cantidad infinitamente pequeña dx^n, con n mayor que 1, cuando se la compare con dx, tendremos que $dx \pm dx^n = dx$ y la razón geométrica dada por Euler es

$$\frac{dx \pm dx^{n+1}}{dx} = 1 \pm dx^n = 1$$

pues $dx^n = 0$.

Por lo tanto, Euler determina la diferencial de $y = x^2$ de la forma siguiente:

Sea w es el incremento de x, entonces el incremento de y será $2xw + w^2$ y la razón de los incrementos de x e y es $1/(2x + w)$, y como $w = 0$, se deduce que

$$\frac{dy}{dx} = \frac{d(x^2)}{dx} = \frac{2x}{1}.$$

Como Euler elimina las diferenciales, la derivada dy/dx, que significa para él el cociente $0/0$, puede ser igual a un número, de la manera siguiente:

Para todo número n, $n \times 0 = 0$, entonces $n = 0/0$, y en consecuencia la derivada resulta ser un método apropiado para determinar el cociente $0/0$.

En lo que respecta al infinito, la idea de Euler es más bien intuitiva. Lo simboliza con el símbolo ∞, el cual representa una cantidad mayor que cualquier cantidad finita. Interpreta que la relación $a/0 = \infty$ significa que una cantidad infinita multiplicada por cero o nada puede ser igual a una cantidad finita. Además, las razones $\frac{a}{dx}, \frac{a}{dx^2}, \ldots, \frac{a}{dx^n}$, con los $dx^i = 0$ con $i \in \mathbb{N}$, corresponden respectivamente al infinito de primer orden, de segundo orden,..., de orden n-ésimo, de manera que el orden de la diferencial determina el orden del infinito.

Esto se encuentra en su obra *"Institutiones calculi differentialis"* que esta divida en dos partes. La primera que consta de nueve capítulos, trata de la teoría del cálculo diferencial, y la segunda que consta de dieciocho, trata de las aplicaciones del cálculo diferencial (donde está los estudios de los problemas de máximos y mínimos de una función).

Nota. Todavía en la época de Euler la Matemática no estaba exenta del influjo de la filosofía e, incluso, de la religión. Como anécdota curiosa, mencionaremos que George Berkeley (1685-1763) filósofo y obispo de Cloyne, era un excéptico con las justificaciones del cálculo e hizo dudar a la comunidad matemática a pesar de las explicaciones fáciles de Euler.

Berkeley mostró su preocupación con la prueba de que la derivada de x^n es nx^{n-1}. En la proporción

$$\frac{(x + o)^n - x^n}{o},$$

no entendia como o al principio era distinto de cero y después de desarrollar el binomio y dividir por o quedaba

$$nx^{n-1} + \frac{n(n-1)}{2}x^{n-2}o + \cdots + o^{n-1}$$

y ahí ya o es igual a cero. Al final lo aceptó, aunque no confiaba en un procedimiento que daba respuestas correctas de una serie de errores. Ofreció su réplica ([12], pag. 10):

"Digo que, en cualquier otra ciencia, los hombres prueban sus conclusiones por sus principios, y no sus principios por sus conclusiones."

1.2.5. Lagrange

Lagrange critica el método de las fluxiones de Newton, muestra su insatisfación con las cantidades infinitesimales de Leibniz y rechaza el concepto de límite formulado por D'Alembert, y quiere sustituirlo todo por un método algebraico simple, para proporcionar al cálculo todo el rigor de las demostraciones antiguas.

En su tratado *"Théorie des fonctions analytiques"* (Teoría de las funciones analíticas) de 1797, contiene los principios del cálculo diferencial en su faceta algebraica, la de la aproximación lineal. El método que expone en este tratado, consiste en utilizar el hecho de que toda función f puede expresarse en su opinión, de la siguiente manera:

$$f(x+h) = f(x) + ph + qh^2 + rh^3 + \cdots$$

donde los coeficientes $p,\ q,\ r\ldots$ dependen de x pero son independientes de h. Intenta justificar que este desarrollo de las funciones en serie de Taylor es siempre posible, y partiendo de ese desarrollo, Lagrange obtiene que:

$$p = f'(x),\ q = \frac{1}{2!}f''(x),\ r = \frac{1}{3!}f'''(x)$$

de donde se deduce que

$$f(x+h) = f(x) + hf'(x) + \frac{h^2}{2!}f''(x) + \cdots$$

En particular, encuentra p o $f'(x)$ de $f(x)$ despreciando todos los términos del desarrollo, salvo los dos primeros, es decir, si

$$f(x+h) - f(x) = ph$$

dividiendo por h tenemos

$$p = \frac{f(x+h) - f(x)}{h} = f'(x).$$

Nota.

i) La limitación más importante de este método, es que toda función, en la acepción moderna, no tiene desarrollo de Taylor.

ii) Lagrange quiso deshacerse de infinitesimales o límites, pero lo siguió utilizando para sus funciones derivadas.

iii) A esta obra de Lagrange le debemos las notaciones que usamos hoy en día para las derivadas sucesivas de una función $f(x)$.

1.2.6. Fourier

La contribución más importante de Fourier fue la idea, intuida por Daniel Bernoulli, de que cualquier función $y = f(x)$ se puede representar por una serie trigonométrica.

Fourier llegó a esta idea mediante el estudio de lo que hoy se conoce como la "ecuación del calor", pues necesitaba escribir la función que da el dato inicial como una suma de una serie trigonométrica. Aunque hay varias maneras según el tipo del problema, en el caso general y para una función 2π-periódica, el problema consiste en, dada una función f, encontrar su serie trigonométrica

$$f(x) = \frac{a_0}{2} + \sum_{n=1}^{\infty} (a_n \cos(nx) + b_n \sin(nx))$$

cuya suma coincida con $f(x)$ para cada x.

Fourier observó que los coeficientes debían venir dados por las fórmulas

$$a_n = \frac{1}{\pi} \int_{-\pi}^{\pi} f(x) \cos(nx) \quad n \in \mathbb{N} \cup \{0\}$$

$$b_n = \frac{1}{\pi} \int_{-\pi}^{\pi} f(x) \sin(nx) \quad n \in \mathbb{N}$$

La serie trigonométrica hayando los coeficientes de esta manera, se llama "Serie de Fourier de f".

La dedución de las fórmulas se basa en el uso de sistemas lineales de infinitas ecuaciones y también un argumento de ortogonalidad. Nunca dudó de como definir los coeficientes, pues los coeficientes de las series trigonométricas son áreas definidas. Fourier daba por sentado que siempre hay un área bien definida entre la gráfica de la función y el eje de abscisas.

Tras considerar varios ejemplos, Fourier supuso que ([13], pag. 654):

"Las series ordenadas segun los cosenos y los senos de arcos múltiplos son siempre convergentes, es decir, que dando a la variable un valor cualquiera no imaginario, la suma de los términos converge cada vez más hacia un único limite fijo, que es el valor de la función desarrollada."

En cuanto a la convergencia, no le pareció necesario demostrarla puesto que le parecía fácil. Podemos leer lo siguiente ([13], pag. 654):

"Esto no resulta solamente de que los valores de los términos disminuyen continuamente; pues esta condición por sí sola no basta para establecer la convergencia de una serie. Es necesario que los valores a los que se llega aumentando continuamente el número de términos se acerquen cada vez más a un límite fijo y no se separen de éste más que en una cantidad que puede hacerse menor que toda magnitud dada: este límite es el valor de la serie. Pues bien, se demuestra rigurosamente que las sucesiones de las que se trata satisfacen esta última condición."

Fourier con todo esto dejó un problema en el que implicaba los conceptos de función (visto anteriormente), integral, suma de series y tipos de convergencia. Este problema influyó en el desarrollo de los conceptos del ánalisis matemático.

El primero en dar un resultado correcto de convergencia fue Dirichlet en 1829, para ello impuso algunas condiciones *suficientes* en $f(x)$ ([20], Vol.3 pag. 1276):

i) $f(x)$ es unívoca y acotada.

ii) $f(x)$ es continua a trozos; esto es, sólo tiene un número finito de discontinuidades en el período (cerrado).

iii) $f(x)$ es monótona a trozos; esto es, sólo tiene un número finito de máximos y mínimos en un período.

Motivado por este trabajo, Dirichlet presentó como contraejemplo la función que hoy en día lleva su nombre (vista en la página 7).

1.2.7. Cauchy

Cauchy fue uno de los trabajadores más prolíficos. En sus tres libros, el *"Cours d'analyuse de l'École Polytechnique"* de 1821, el *"Résumé des leçons sur le calcul infinitésimal "* de 1823 y las *"Leçons sur le calcul différential"* de 1829, dio al cálculo infinitesimal la forma que tiene hoy en día.

Para él, el cálculo no requería cantidades decrecientes ni infinitamente pequeñas, sino el límite, que es con lo que construyó todo su cálculo. Proporcionó la siguente definición ([12], pag. 10):

"Cuando los valores sucesivamente atribuidos a una variable se aproximan indefinidamente a un valor fijo de modo que terminen por diferir en él por tan poco como se deseé, este último se llama límite de todo lo demás."

Muchos matemáticos habían considerado un infinitésimo como un número constante muy pequeño, no obstante Cauchy lo define como una cantidad variable debido al concepto de límite ([10], pag. 312):

"Cuando los valores numéricos sucesivos de una misma variable decrecen indefinidamente de manera que disminuyen por debajo de todo número dado, esta variable resulta ser lo que se llama un infinitamente pequeño o una cantidad infinitamente pequeña. Una variable de esta especie tiene cero como límite."

Cauchy utilizó de nuevo su definición de límite para definir la continuidad de una función ([10], pag. 313):

"Sea $f(x)$ una función de la variable x, y supongamos que esta función posee un valor único y finito para cada valor de x en un entorno dado. Si para un valor de x en este intervalo, se añade un valor infinitesimal h, la función aumenta en la diferencia $f(x+h)-f(x)$, que depende, a su vez, de la nueva variable h y del valor de x. Establecido lo anterior, la función $f(x)$ será continua con respecto a x entre los límites dados si, entre esos límites, un crecimiento infinitamente pequeño de la variable produce siempre un crecimietno infinitamente pequeño de la función.
Se dice, además, que la función $f(x)$ es continua en el entorno de un valor particular atribuido a la varible, siempre que sea continua entre dos límtes de x, incluso muy proximos, que contengan al valor de que se trata."

Nota. Esta definición de la continuidad equivale a decir que $f(x)$ será continua en a si $f(x)$ se aproxima al límite $f(a)$ cuando x se aproxima al límite a.

Por lo tanto, la definición de derivada según Cauchy es ([10], pag. 313-314):

"Cuando la función $y=f(x)$ es continua entre los dos límites dados de la variable x, y se asigna a esta variable un valor comprendido entre los dos límites de que se trata, un crecimiento infinitamente pequeño atribuido a la variable produce un crecimiento infinitamente pequeño de la función. Por consiguiente, si se hace $\Delta x = i$, los dos términos de la razón de diferencias

$$\frac{\Delta y}{\Delta x} = \frac{f(x+i)-f(x)}{i}$$

serán cantidades infinitamente pequeñas. Pero, mientras que estos dos términos se aproximarán indefinida y simultáneamente al límite cero, la razón *podrá converger hacia otro límite, sea positivo o negativo. Este límite, cuando exista, tiene un valor determinado para cada valor determinado de x, pero varía con x."*

Nota. El primero en dar así la definición de derivada fue Bolzano en 1817, e insistió en que $f'(x)$ no era un cociente de ceros o una razón de cantidades que desaparecen.

Cauchy definió el concepto de diferencial en términos de la derivada ([10], pag. 314):

"Si dx es una cierta cantidad finita, entonces la correspondiente diferencial dy de $y = f(x)$ vendrá definida simplemente como $f'(x)dx$."

En otras palabras, se introduce dos cantidades dy y dx cuya razón $\frac{dy}{dx}$ por definición es $f'(x)$. Además, clarificó la relación entre $\frac{dy}{dx}$ y $f'(x)$ a traves del teorema del valor medio.

Nota. Lacroix ya había definido $dy = f'(x)dx$ en la primera edición de su *"Traitè"*.

Cauchy se sirve del concepto de diferencial expresado en términos de la derivada para definir las diferenciales de orden superior. Por ejemplo, como la diferencial $dy = f'(x)dx$ es, de hecho, una función de x y de dx, manteniendo dx fijo, la función $f'(x)dx$ tendrá por derivada $f''(x)dx$ y una diferencial de orden dos $d^2y = f''(x)dx^2$. En general, $d^ny = f^n(x)dx^n$, donde $f^n(x)$ es el coeficiente diferencial.

La noción de diferencial tiene sentido cuando está relacionada con la derivada.

Nota. Bernard Bolzano (1781-1848) si vio la distinción entre continuidad y diferenciabilidad. En su libro *"Funktionenlehre"* (Lecciones de Funciones) que escribió sobre el año 1834, pero ni terminó ni publicó, presentó una función continua que no tiene derivada finita en ningún punto (aunque lo hubiese publicado en ese año, no habría impresionado puesto que la curva no tenía una representación analítica). Esta función se trata en la sección 3.1 del capítulo 3.

Cauchy también trabajo en las series infinitas, y da un cierto rigor en el aspecto de la convergencia de las series. Sobre el concepto de serie dice ([10], pag.316):

"Una serie (sucesión) es una sucesión infinita de cantidades, u_0, u_1, $u_2, \ldots$, que se suceden en virtud de una ley determinada. Estas cantidades son los diferentes términos de la sucesión considerada. "

Y del concepto de convergencia dice ([10], pag. 316):

"Sea $s_n = u_0 + u_1 + \cdots + u_{n-1}$ *la suma de los* n *primeros términos, donde* n *es un entero (número natural). Si la suma* s_n *tiende hacia un cierto límite* s *para valores crecientes de* n*, entonces la serie se dice convergente, y el límte en cuestión se llama la suma de la serie. Por el contrario, si la suma* s_n *no se aproxima a un límite determinado cuando* n *aumenta indefinidamente, la serie es divergente y no tendrá suma."*

Cauchy muestra que el concepto de límite está relacionado con la definición de convergencia de una serie, y supone que una serie infinita puede tener una suma en el sentido de un límite.

Enuncia el criterio de convergencia que lleva su nombre ([10], pag. 316-317):

"Una sucesión S_m *converge hacia un límte* S *sólo si la diferencia de* S_{n+r} *y de* S_n*, para todo valor de* r *y de* n *suficientemente grande, puede hacerse menor en valor absoluto que cualquier cantidad dada."*

Cauchy demostró la condición necesaria a partir de la definición de convergencia, pero la condición de suficiencia no pudo demostrarla porque no disponía la definición rigurosa de números irracionales, es decir, la completitud de $\mathbb{R}$.

Cauchy asume que la suma de funciones continuas es una función continua, error que subsanaría Karl Weierstrass (1815-1897) una vez establecida la noción de convergencia uniforme.

Nota. Neils Henrik Abel (1802-1829) estuvo seis meses en París, y le describió en una carta a su mentor lo que estaba ocurriendo en ese momento ([8], pag. 160-161). Un fragmento interesante de esa carta es:

"Cauchy está loco, y no hay forma de llevarse bien con él, aunque ahora mismo él es el único que sabe cómo se deben hacer las matemáticas. Lo que está haciendo es excelente, pero muy confuso. Al principio no entendí casi nada, pero ahora lo veo un poco más claro. Está publicando una serie de memorias bajo el título "Exercise in mathematics"*. Cauchy es, por el momento, el único interesado en las matemáticas puras. Poisson, Fourier, Ampère, etc. están trabajando exclusivamente en magnetismo y otros temas físicos. El señor Laplace no está trabajando, creo."*

1.2.8. Weierstrass

Al igual que Cauchy, Weierstrass se interesó sólo en las matemáticas puras, en particular en el análisis, dando el máximo rigor posible en vez de recurrir a la intuición.

Construyó la teoría de funciones analíticas sobre la base de las series de potencias, concepto central del trabajo de Lagrange, y el proceso de prolongación analítica.

Para Weierstrass una función se define mediante su desarrollo en serie de potencias convergente, y sus valores dentro del círculo de convergencia constituyen el *elemento de la función*. Además, mediante la prolongación analítica se relacionan uno con otro los elementos de la función, surgiendo así una función analítica como conjunto de todas las prolongaciones a que da lugar un elemento de la función.

El uso de las series de potencias para representar funciones complejas dadas en forma analítica ya era conocido. Sin embargo, la contribución de Weierstrass fue asentar la noción de *"derivadas de series"*, es decir: "Probar la fórmula de derivación de series término a término bajo las condiciones de convergencia uniforme adecuadas."

Weierstrass formaliza la idea de "uniformidad" y proporciona numerosos teoremas que hoy llevan su nombre y que habían sido utilizados tácticamente antes. Por ejemplo, para probar que toda función continua en un intervalo compacto es integrable, Cauchy asume implícitamente la continuidad uniforme de la función, aunque no prueba este hecho.

Finalmente, mencionaremos que en su empeño por la formalización rigurosa del Análisis, y en respuesta a la hipóteses de Fourier, construye el contraejemplo que estudiaremos en el capítulo 3.

1.2.9. Baire

René Baire (1874-1932) fue uno de los primeros en avanzar en la teoría de conjuntos. Admiraba a Georg Cantor (1845-1918) y estudio con Vito Volterra (1860-1940). Su tésis de 1899 *"Sur les functions de varables réelles"* contiene su famoso teorema de categoría.

Baire con sus ideas de conjuntos de primera categoría contestó a la pregunta de *"¿como de discontinua puede ser una derivada?"* al probar que los puntos de discontinuidad de una derivada forman un conjunto de primera categoría y concluyó ([12], pag. 16):

"Si f es diferenciable en un intervalo I, entonces f' debe ser continua en un subconjunto denso de I."

Por lo tanto, las derivadas pueden ser discontinuas, pero no tan discontinuas que haya un intervalo abierto libre de sus puntos de continuidad. Sobre este punto aportaremos un poco más de información en la sección 2.1.

1.2.10. Lebesgue

El foco de atención principal de Henri Lebesgue (1875-1941) eran las series de Fourier y los problemas de convergencia. La necesidad de una formulación coherente de esta teoría le empujó a establecer nuevos criterios de integrabilidad por los que hoy es universalmente conocido. Su reputación analítica se estableció con su clásico de 1904 *Leçons sur l'integration*.

Lebesgue soluciona en gran medida el caos suscitado por el contrejemplo de Weierstrass. En palabras de Royden ([27] pag. 113):

"El teorema de Lebesgue ayudó a recuperar la confianza en la armonía del Análisis Matemático."

Lebesgue definió los conjuntos de medida nula, y usando el teorema de Heine-borel demostró que ([12], pag. 16-17):

"Una función acotada en $[a, b]$ *es Riemann integrable en* $[a, b]$ *si, y solo si, su conjunto de discontinuidades en* $[a, b]$ *tiene medida nula."*

Esto resolvió la cuestión de cómo de discontinua puede ser una función integrable de Riemann.

Lebesgue desarrolló una teoria de la medida para la subconjuntos de la recta real, no sólo para los subconjuntos de medida nula. Explicó la noción de una función medible y luego su integral. La diferencia estriba en que en la teoría de Riemann se divide el dominio de definición de la función, mientras que en la de Lebesgue se parte la imagen. Gracias al estudio de su integral, el teorema fundamental del cálculo pudo ser salvado y esto reparó el daño por el contraejemplo de Volterra, que en el año 1881 describe una función que era diferenciable con una derivada acotada, pero cuya derivada era tan descontrolada que su integral (Riemann) no existía.

Otra de las aportaciones de Lebesgue al esclarecimiento de los problemas abiertos en el Análisis Real es el teorema de derivabilidad que tratamos en esta memoria.

Capítulo 2

Otras formas de derivación

La mayoría de los métodos de extensión de la derivada, conservan el cociente de diferencia como la medida de cuánto una curva “se parece”a su recta tangente, y proceden manipulando la forma en que se calcula el límite de cocientes de diferencias.

En este capítulo daremos otros métodos que se van alejando de los cocientes incrementales.

2.1. Condición de derivabilidad en el sentido estricto

En primer lugar, en este apartado vamos a ver cuando unas funciones son derivadas de otras.

El teorema de Darboux establece que si una función f es derivable en todos los puntos del intervalo I, entonces la función tiene la propiedad del valor intermedio. Esto significa que si f' asume los valores de A y B, entonces también asume todos los valores que están entre A y B. Resulta curioso, que las derivadas que no son necesariamente funciones continuas, también posean está propiedad.

Teorema 2.1.1 (**Propiedad de Darboux para las derivadas**). Sea f una función definida y continua en un intervalo compacto $[a, b]$ y tal que existe la derivada, finita o infinta, en cada punto de $[a, b]$. Se supone además que $f'(a) \neq f'(b)$. Entonces, si λ es un número real comprendido entre $f'(a)$ y $f'(b)$, existe $c \in (a, b)$ tal que $f'(c) = \lambda$

Definición 2.1.2. Una función de Darboux es aquella que lleva intervalos en intervalos o que tiene la propiedad del valor medio.

Nota. Las integrales que se utilizan aquí son del tipo Lebesgue, y las demostraciones de los siguientes resultados se pueden observar en [6].

Lema 2.1.3. Si $f\colon [a,b] \to \mathbb{R}$ es una función medible no negativa y $\int_a^b f = 0$, entonces $f(x) = 0$ c.s. en $[a,b]$.

Definición 2.1.4. Sean f una función medible y acotada en $[a,b]$, I un subintervalo cerrado de $[a,b]$, y $x_0 \in \mathring{I}$. Llamamos x_0 *"un punto medio de I relativo a f"* si $\int_I f = f(x_0)\mu(I)$, donde $\mu(I)$ es la medida de Lebesgue, que en $\mathbb{R}$ es la longitud del intervalo I.

Observación 2.1.5.

i) Un punto medio depende tanto del intervalo cerrado como de la función.

ii) En una función continua cada intervalo cerrado tiene un punto medio, pero el recíproco no es cierto. Veamoslo en el siguente ejemplo:

$$f(x) = \begin{cases} 1 & si \quad x \in \mathbb{I} \\ 0 & si \quad x \in \mathbb{Q} \end{cases}$$

Claramente cada intervalo cerrado en $[0,1]$ tiene un punto medio, pero f no es continua. Además f no es una función de Darboux.

Definición 2.1.6. Sea $\{I_n\}_{n=1}^{\infty}$ una sucesión de intervalos cerrados no degenerados, y sea x un número real. Decimos que $\{I_n\}_{n=1}^{\infty}$ converge a x, y lo denotamos por $I_n \to x$, si $x \in I_n$ para cada n y $\mu(I_n) \to 0$.

Teorema 2.1.7. Sea f una función de Darboux acotada en $[a,b]$. Entonces f es una derivada en $[a,b]$ si, y solo si,

i) f es medible, en el sentido Borel

ii) para cada $x \in [a,b]$ y para cualquier sucesión I_n de subintervalos cerrados no degenerados de $[a,b]$ tal que $I_n \to x$, se tiene que $f(x_n) \to f(x)$, donde x_n es un punto medio arbitrario de I_n relativo a f.

Observación 2.1.8. La condición (ii) del teorema es mucho más débil que la continuidad; esta condición requiere que $f(x_n) \to f(x)$ se mantenga sólo para la sucesión especial x_n.

Nota. Si f es integral Riemann en $[a,b]$, entonces f es acotada y continua en casi todo punto de $[a,b]$.

2.2. Derivadas de Dini

El italiano Dini (1845-1918) estudió matemáticas en Pisa, durante el año 1865 se fue a París para ampliar sus conocimientos. Cuando regresó a Pisa, comenzó a dudar de la verdad de algunas proposiciones que había escuchado sobre la continuidad y diferenciabilidad. Estas proposiciones habían sido confirmadas en publicaciones por Hankel, Schwarz, Darboux y Weierstrass.

Dini realizó sus propias investigaciones, que publicó en *"Fundamenti per la teorica delle funzioni di variabili reali"* (Fundamentos de la teoría de funciones de variable real) en 1878.

Su estudio se basó en cuatro términos útiles y elementales, conocidos hoy en día como las derivadas de Dini. Las definió de la siguiente manera:

Definición 2.2.1. Sea f una función definida sobre un intervalo $[a, b]$ y sea x_0 un punto de $[a, b]$. Entonces

si $x_0 < b$, entonces

- Derivada superior derecha de f en x_0

$$D^+ f(x_0) = \limsup_{x \to x_0^+} \frac{f(x) - f(x_0)}{x - x_0}.$$

- Derivada inferior derecha de f en x_0

$$D_+ f(x_0) = \liminf_{x \to x_0^+} \frac{f(x) - f(x_0)}{x - x_0}.$$

si $x_0 > a$, entonces

- Derivada superior izquierda de f en x_0

$$D^- f(x_0) = \limsup_{x \to x_0^-} \frac{f(x) - f(x_0)}{x - x_0}.$$

- Derivada inferior izquierda de f en x_0

$$D_- f(x_0) = \liminf_{x \to x_0^-} \frac{f(x) - f(x_0)}{x - x_0}.$$

Nota. La definición que hemos dado es la moderna, sin embargo téngase en cuenta que la idea de límite superior y límite inferior fue posterior a las definiciones de derivadas de Dini.

Observando los límites de forma global se definen las derivadas superior e inferior:

Definición 2.2.2. Sea f una función definida en un intervalo $[a, b]$ y sea x_0 un punto de $[a, b]$. Entonces

- Derivada superior de f en x_0

$$\overline{D} f(x_0) = \limsup_{x \to x_0} \frac{f(x) - f(x_0)}{x - x_0}.$$

• Derivada inferior de f en x_0

$$\underline{D}f(x_0) = \liminf_{x \to x_0} \frac{f(x) - f(x_0)}{x - x_0}.$$

Nótese que si $D^+f(x_0) = D_+f(x_0)$, entonces existe la derivada superior de f en x_0, y si $D^-f(x_0) = D_-f(x_0)$, entonces existe la derivada inferior de f en x_0.

Dini aplicó estas ideas al teorema fundamental del cálculo, y pudo demostrar que si la función f es integrable y $F(x) = \int_a^x f$, entonces

$$F(b) - F(a) = \int_a^b DF$$

donde DF es cualquiera de sus cuatro derivadas.

También demostró que si F es una función continua y una de sus cuatro derivadas es integrable, entonces

$$\int_a^b DF = F(b) - F(a). \tag{2.1}$$

La derivada de Dini es una generalización especial del concepto de derivada para funciones continuas no necesariamente diferenciables. Estas derivadas proporcionan información a cerca de la diferenciabilidad de una función sobre un punto determinado. Concretamente, para las funciones de una sola variable, si todas las derivadas de Dini existen y además se cumple $D^+f(x_0) = D_+f(x_0) = D^-f(x_0) = D_-f(x_0)$, entonces la función f es diferenciable en el punto x_0 y el valor de su derivada coincide con el de las derivadas de Dini.

La condición de continuidad de la función es esencial en la fórmula (2.1). Notese que si

$$H(x) = \begin{cases} 1 & si \quad x \geq 0 \\ 0 & si \quad x < 0 \end{cases}$$

(la función de Heaviside), entonces H tiene derivadas de Dini en todo punto, excepto en el punto $x = 0$, y $DH(x) = 0$ si $x \neq 0$.

La función valor absoluto, $f(x) = |x|$, es una función continua para toda la recta real y diferenciable en todos los puntos salvo en el punto $x = 0$, donde la derivada no está bien definida. Sin embargo, las derivadas de Dini en dicho punto existen y son finitas.

$$D^+f(x_0) = D_+f(x_0) = 1 \quad \Rightarrow \quad \overline{D}f(x_0) = 1$$

$$D^-f(x_0) = D_-f(x_0) = -1 \quad \Rightarrow \quad \underline{D}f(x_0) = -1$$

La noción de derivada de Dini es clave en las demostraciones clásicas del teorema de Lebesgue.

2.3. Derivada por Mínimos Cuadrados

Aquí tomamos otro enfoque reinterpretando la frase *"la curva se parece a la recta"*. Un conjunto de puntos se puede decir que "parece" una recta, y una forma de verlo es en el sentido de los mínimos cuadrados. Este método se suele utilizar para encontrar la recta que mejor se ajusta a un conjunto de puntos. El método puede adaptarse para nuestro propósito de esta manera:

Sea $m(h)$ y $b(h)$ los números que minimizan la expresión

$$\int_x^{x+h} [f(t) - m(h)t - b(h)]^2 \, dt$$

para un valor dado de h (solo nos interesa $m(h)$, la pendiente de la recta de aproximación). Resolviendo el problema variacional, esto es estableciendo las derivadas parciales respecto de m y b igualadas a 0, y solucionando el sistema de ecuaciones resultante, e integrando tenemos

$$m(h) = \frac{12}{h^3} \left(\int_x^{x+h} t f(t) dt - \left[x + \frac{h}{2} \right] \int_x^{x+h} f(t) dt \right).$$

Decimos que f es diferenciable por mínimos cuadrados si $\lim_{h \to 0} m(h)$ existe, y llamamos a esa cantidad *la derivada de mínimos cuadrados* de f en x. Para más información consultar [21].

2.4. Derivadas Fraccionarias

Leibniz fue quien introdujo el símbolo $\frac{d^n}{dx^n} f(x)$, que denota la n-ésima derivada de la función f. Esto se lo mandó a L'Hôpital en una carta supuestamente diciendo que $n \in \mathbb{N}$, entonces L'Hôpital le preguntó que qué pasaba si $n = 1/2$. Leibniz no encontró respuesta alguna, excepto al caso especial $f(x) = x$.

La respuesta a la pregunta de L'Hôpital no es única, pues hay muchas generalizaciones posibles de $\frac{d^n}{dx^n} f(x)$ al caso $n \notin \mathbb{N}$. Nosotros vamos a ver dos: derivada de Riemann-Liouville y derivada de Caputo.

La idea básica detras del cálculo fraccional está relacionada con el teorema fundamental del cálculo.

Teorema 2.4.1 (Teorema fundamental del cálculo). Sea $f \colon [a, b] \to \mathbb{R}$ una función continua, y sea $F \colon [a, b] \to \mathbb{R}$ definida por

$$F(x) := \int_a^x f(t) dt.$$

Entonces, F es diferenciable en $[a, b]$ y $F' = f$.

Por lo tanto, tenemos una relación muy estrecha entre operadores diferenciales y operadores integrales. Debido a esto, el cálculo fraccional desea mantener la relación entre los operadores diferenciales y los operadores integrales en un sentido adecuadamente generalizado. Por tanto es necesario tratar con operadores integrales fraccionarios antes de llegar a los operadores diferenciales fraccionarios.

Veamos primero una serie de resultados que nos interesan.

Definición 2.4.2.

i) Denotamos por D al operador que representa una función diferenciable sobre su derivada, es decir, $Df(x) := f'(x)$.

ii) Denotamos por J_a al operador que asigna a cada función f integrable (Riemann) en el intervalo $[a, b]$ su primitiva centrada en a, es decir, $J_a f(x) := \int_a^x f(t)dt$ para $a \leq x \leq b$.

iii) Para $n \in \mathbb{N}$ usamos los símbolos D^n y J_a^n que denotan n iteraciones de D y J_a respectivamente, es decir, establecemos $D^1 := D$, $J_a^1 := J_a$, y $D^n := DD^{n-1}$ y $J_a^n := J_a J_a^{n-1}$.

Observación 2.4.3. Con la notación que acabamos de dar, el teorema fundamental del cálculo se lee de la siguiente manera

$$DJ_a f = f.$$

Esto implica que $D^n J_a^n f = f$ para $n \in \mathbb{N}$, es decir, D^n es un inverso por la izquierda de J_a^n. Lo que queremos es generalizar las condiciones del teorema fundamental del cálculo al caso fraccional $n \notin \mathbb{N}$.

Podemos sustituir la definición 2.4.2 (iii) por la siguiente fórmula explicita:

Lema 2.4.4. Sea f una función Riemann integrable en $[a, b]$. Entonces para $a \leq x \leq b$ y $n \in \mathbb{N}$ tenemos

$$J_a^n f(x) = \frac{1}{(n-1)!} \int_a^x (x-t)^{n-1} f(t)dt.$$

Una consecuencia inmediata de $D^n J_a^n f = f$ es:

Lema 2.4.5. Sea $m, n \in \mathbb{N}$ con $m > n$, y sea f una función que tenga n-ésima derivada continua en el intervalo $[a, b]$. Entonces $D^n f = D^m J_a^{m-n} f$.

Estos dos lemas son fundamentales para las generalizaciones que veremos en el sentido de Riemann-Liouville y de Caputo.

Atendiendo al lema 2.4.4 sera útil generalizar los factoriales a argumentos no enteros. Tal generalización es la siguiente:

Definición 2.4.6. La función $\Gamma : (0, \infty) \to \mathbb{R}$, definada por

$$\Gamma(x) := \int_a^\infty t^{x-1} e^{-t} dt,$$

se llama función Gamma de Euler.

La propiedad más importante de la función Gamma que nosotros vamos a utilizar es:

Teorema 2.4.7. Para $n \in \mathbb{N}$, tenemos que $(n-1)! = \Gamma(n)$.

Antes de comenzar con lo que nos interesa, definiremos algunos espacios que vamos a utilizar más adelante.

Definición 2.4.8. Sea $k \in \mathbb{N}$ y $1 \leq p$

$$\mathcal{L}^p[a,b] := \left\{ f \colon [a,b] \to \mathbb{R} : \text{si } f \text{ es medible en } [a,b] \text{ y } \int_a^b |f(x)|^p dx < \infty \right\}.$$

Este espacio se llama espacio de Lebesgue.

$$\mathcal{C}^k[a,b] := \left\{ f \colon [a,b] \to \mathbb{R},\ f \text{ tiene } k\text{-ésima derivada continua} \right\}.$$

Definición 2.4.9. Sea $n \in \mathbb{N}$, $\mathcal{A}^n[a,b]$ es el conjunto de funciones con las $(n-1)$ primeras derivadas absolutamente continuas, es decir, las funciones f para las cuales existe (en casi todo punto) una función $g \in \mathcal{L}^1[a,b]$ tal que

$$f^{(n-1)}(x) = f^{(n-1)}(a) + \int_a^x g(t) dt$$

donde $g = f^{(n)}$.

Observación 2.4.10. A partir de esta definición, queda claro que una función $f \in \mathcal{A}^1[a,b]$ posee (en casi todo punto) una derivada $f' \in \mathcal{L}^1[a,b]$.

Ahora ya vamos a responder a la pregunta de L'Hôpital. Podemos dar condiciones de operadores integrales y diferenciales fraccionarios J_a^n y D^n para $n \notin \mathbb{N}$.

2.4.1. Derivada en el sentido de Riemann-Liouville

Definición 2.4.11. Sea $n \in \mathbb{R}_+$. El operador J_a^n definido en $\mathcal{L}^1[a,b]$ por

$$J_a^n f(x) := \frac{1}{\Gamma(n)} \int_a^x (x-t)^{n-1} f(t) dt$$

para $a \leq x \leq b$, se llama el operador integral fraccionario Riemann-Liouville de orden n.

Para $n = 0$ tenemos el operador identidad, es decir, $J_a^0 = I$.

Esta definición coincide con la clásica de J_a^n para $n \in \mathbb{N}$, excepto que hemos pasado de integral de Riemann a Lebesgue. Para $n \geq 1$ la integral $J_a^n f(x)$ existe pues el integrando es el producto de una función integrable f, y una función continua $(x-t)^{n-1}$. Pero para $0 < n < 1$ se necesita el siguiente teorema para la justificación de la definición.

Teorema 2.4.12. Sea $f \in \mathcal{L}^1[a,b]$ y $n > 0$. Entonces, la integral $J_a^n f(x)$ existe para casi todo punto $x \in [a,b]$. Además, la función $J_a^n f$ es también un elemento de $\mathcal{L}^1[a,b]$.

Teorema 2.4.13. Sean $m, n \geq 0$ y $\phi \in \mathcal{L}^1[a,b]$. Entonces

$$J_a^m J_a^n \phi = J_a^{m+n} \phi$$

en casi todo punto de $[a,b]$. Si además, $\phi \in \mathcal{C}[a,b]$ o $m+n \geq 1$, entonces la identidad se mantiene en todo punto de $[a,b]$.

Corolario 2.4.14. Bajo los supuestos del teorema anterior

$$J_a^m J_a^n \phi = J_a^n J_a^m \phi.$$

Hay una manera algebraica de establecer este resultado.

Teorema 2.4.15. El operador $\{J_a^n \colon \mathcal{L}^1([a,b]) \to \mathcal{L}^1([a,b]) : n \geq 0\}$ forma un semigrupo commutativo con respecto a la composición. El operador identidad J_a^0 es el elemento neutral del semigrupo.

Después de haber dado algunas propiedades de los operadores integrales de Riemann-Liouville, ahora empezaremos con los operadores diferenciales.

Definición 2.4.16. Sean $n \in \mathbb{R}_+$ y $m = \lceil n \rceil$. El operador D_a^n, definido por

$$D_a^n f := D^m J_a^{m-n} f$$

se llama el operador diferencial fraccionario Riemann-Liouville de orden n.

Para $n = 0$ tenemos el operador identidad, es decir, $D_a^0 = I$.

Observación 2.4.17. $\lceil n \rceil$ denota al mínimo número entero superior a n, es decir, $\lceil n \rceil - 1 < n \leq \lceil n \rceil$.

La diferencia de esta definición con el resultado clásico es que ahora el operador depende de la elección del punto a. Además, m tiene que ser lo más pequeño posible, mientras que antes daba igual el valor de m siempre que se satisfaciera la desigualdad $m > n$. Algo similar tenemos aquí.

Lema 2.4.18. Sean $n \in \mathbb{R}_+$ y $m \in \mathbb{N}$ tal que $m > n$. Entonces,

$$D_a^n = D^m J_a^{m-n}.$$

El siguiente resultado proporciona una condicion suficiente para la existencia de $D_a^n f$.

Lema 2.4.19. Sean $f \in \mathcal{A}^1[a,b]$ y $0 < n < 1$. Entonces, $D_a^n f$ existe en casi todo punto de $[a,b]$. Además $D_a^n f \in \mathcal{L}^p[a,b]$ para $1 \leq p \leq 1/n$ y

$$D_a^n f(x) = \frac{1}{\Gamma(1-n)} \left(\frac{f(a)}{(x-a)^n} + \int_a^x f'(t)(x-t)^{-n} dt \right).$$

Hemos visto que los operadores integrales de Riemann-Liouville forman un semigrupo. Los operadores diferenciales clásicos $\{D^n : n \in \mathbb{N} \cup \{0\}\}$ también tiene la propiedad de semigrupo. Entonces, ¿los operadores diferenciales de Riemann-Liouville tienen esa estructura?. Veamos el siguiente resultado.

Teorema 2.4.20. Sean $n_1, n_2 \geq 0$. Además sea $\phi \in \mathcal{L}^1[a,b]$ y $f = J_a^{n_1+n_2}\phi$. Entonces

$$D_a^{n_1} D_a^{n_2} f = D_a^{n_1+n_2} f.$$

Observación 2.4.21. Para aplicar esta identidad no hace falta conocer la función ϕ explicitamente, es suficiente saber que tal función existe.

Recordemos que una de las características clave que queremos obtener era $D^n J_a^n f = f$. Las definiciones del Riemann-Liouville si tienen esa propiedad.

Teorema 2.4.22. Sea $n \geq 0$. Entonces, para cada $f \in \mathcal{L}^1[a,b]$

$$D^n J_a^n f = f$$

casi en todo punto.

Otra peculiaridad interesante de los operadores diferenciales fraccionarios, es que cumple la condición de linealidad.

Teorema 2.4.23. Sean f_1 y f_2 dos funciones definidas en $[a,b]$ tales que $D_a^n f_1$ y $D_a^n f_2$ existen en casi todo punto. Además, sean $c_1, c_2 \in \mathbb{R}$, entonces $D_a^n(c_1 f_1 + c_2 f_2)$ existe en casi todo punto, y

$$D_a^n(c_1 f_1 + c_2 f_2) = c_1 D_a^n f_1 + c_2 D_a^n f_2.$$

Cuando se trata del producto de funciones, la situación es diferente. En el caso clásico tenemos la fórmula de Leibniz que tiene dos propiedades:

i) es simétrica, es decir, podemos intercambiar f y g

ii) para evaluar la n-ésima derivada del producto fg, solo se necesita evaluar derivadas de orden menor o igual que n de ambos factores.

Teorema 2.4.24. Sea $n \in \mathbb{N}$, y sean $f, g \in \mathcal{C}^n[a,b]$. Entonces

$$D^n(fg) = \sum_{k=0}^{n} \binom{n}{k} \left(D^k f\right) \left(D^{n-k} g\right).$$

El siguiente teorema transfiere la fórmula de Leibniz a la configuración fraccionaria, y es evidente que ambas propiedades se pierden.

Teorema 2.4.25. Sea $n > 0$, y supongamos que f y g son analíticas en $(a-h, a+h)$ con $h > 0$. Entoces

$$\begin{aligned} D_a^n(fg)(x) &= \sum_{k=0}^{\lceil n\rceil -1} \binom{n}{k} \left(D_a^k f\right)(x) \left(D_a^{n-k} g\right)(x) \\ &+ \sum_{\lceil n\rceil}^{\infty} \binom{n}{k} \left(D_a^k f\right)(x) \left(J_a^{k-n} g\right)(x) \end{aligned}$$

para $a < x < a + h/2$.

La relación entre operadorer integrales fraccionarios y operadores diferenciales fraccionarios es:

Teorema 2.4.26. Sea $n > 0$. Si exite alguna $\phi \in \mathcal{L}^1[a, b]$ tal que $f = J_a^n \phi$ entonces

$$J_a^n D_a^n f = f$$

para casi todo punto.

Si f no verifica las condiciones del teorema anterior, hay diferentes representaciones para $J_a^n D_a^n f$.

Teorema 2.4.27. Sean $n > 0$ y $m = \lceil n\rceil$. Supongamos que f es tal que $J_a^{m+n} \in \mathcal{A}^m[a, b]$. Entonces,

$$J_a^n D_a^n f(x) = f(x) - \sum_{k=0}^{m-1} \frac{(x-a)^{n-k-1}}{\Gamma(n-k)} \lim_{z\to a^+} D^{m-k-1} J_a^{m-n} f(z).$$

En particular, para $0 < n < 1$, tenemos

$$J_a^n D_a^n f(x) = f(x) - \frac{(x-a)^{n-1}}{\Gamma(n)} \lim_{z\to a^+} J_a^{1-n} f(z).$$

2.4.2. Derivada en el sentido de Caputo

El operador de Caputo se construyó para mejorar algunos aspectos del operador de Riemann-Liouville. Antes de definirlo necesitamos una definición y teorema preliminar

Definición 2.4.28. Sean $n \geq 0$ y $m = \lceil n\rceil$. Entonces definimos el operador $\widehat{D}_a^n$ por

$$\widehat{D}_a^n f := J_a^{m-n} D^m f$$

cuando $D^m f \in \mathcal{L}^1[a, b]$.

Observación 2.4.29. Sea $m = n \in \mathbb{N} \cup \{0\}$, entonces $\widehat{D}_a^n f = J_a^0 D^n f = D^n f$.

Teorema 2.4.30. Sean $n \geq 0$ y $m = \lceil n \rceil$. Si $f \in \mathcal{A}^m[a,b]$. Entonces,

$$D_a^n f = D_a^n \left[f - T_{m-1}[f;a]\right]$$

en casi todo punto.

Nota. $T_{m-1}[f;a]$ denota el polinomio de Taylor de grado $m-1$ de la función f centrada en el punto a.

Definición 2.4.31. Asumimos $n \geq 0$ y f es tal que $D_a^n[f - T_{m-1}[f;a]]$ existe, donde $m = \lceil n \rceil$. Entonces definimos la función

$$D_{*a}^n f := D_a^n[f - T_{m-1}[f;a]].$$

El operador D_{*a}^n se llama el operador diferencial de Caputo de orden n.

Observación 2.4.32. Si $n \in \mathbb{N}$ y $m = n$

$$D_{*a}^n f = D_a^n[f - T_{n-1}[f;a]] = D^n f - D^n(T_{n-1}[f;a]) = D^n f.$$

Este concepto se introdujo de forma independiente por muchos autores. Por ejemplo, Caputo y Rabotnov basaron sus desarrollos en el enfoque de la definición 2.4.28, mientras que Dzherbashyon y Nersesian utilizan la definición 2.4.31 como punto de partida. Liouville no encontraba la diferencia entre el operador de Caputo y el operador de Riemann-Liouville, pues se centraba en los casos donde ambos operadores coincidían. Veamos la relación entre ambos operadores.

Lema 2.4.33. Sea $n \geq 0$ y $m = \lceil n \rceil$. Si f es tal que ambos operadores $D_{*a}^n f$ y $D_a^n f$ existen. Entonces

$$D_{*a}^n f(x) = D_a^n f(x) - \sum_{k=0}^{m-1} \frac{D^k f(a)}{\Gamma(k-n+1)}(x-a)^{k-n}.$$

Por tanto, tendremos que

$$D_{*a}^n f = D_a^n f \Leftrightarrow D^k f(a) = 0 \quad \text{para} \quad k = 0, 1, \ldots, m-1.$$

También tenemos una relación entre los operadores integrales de Riemann-Liouville y los operadores diferenciales de Caputo. Se tiene que la derivada de Caputo es una inversa por la izquierda de la integral Riemann-Liouville, mientras que por la derecha no lo es.

Teorema 2.4.34. Si f es continua y $n \geq 0$, entonces $D_{*a}^n J_a^n f = f$. Además, si $m = \lceil n \rceil$ y $f \in \mathcal{A}^m[a,b]$. Entonces

$$J_a^n D_{*a}^n f(x) = f(x) - \sum_{k=0}^{m-1} \frac{D^k f(a)}{k!}(x-a)^k.$$

Nota. Para más información sobre las derivadas fraccionarias y ver las demostraciones de estos resultados consultar [11].

2.5. Derivada débil

Como se ha mencionado en la introducción, otro hecho revolucionario en el desarrollo del Análisis Real es la inclusión de las funciones generalizadas. Hacia 1930, en su libro *"Principios de mecánica cuántica"*, Dirac (1902-1984) establece una definición de la función delta (δ) con las siguientes propiedades:

$$\delta(x) = \begin{cases} 0 & si \quad x \neq 0 \\ \infty & si \quad x = 0 \end{cases} \qquad \text{y} \qquad \int_{-\infty}^{\infty} \delta(x)dx = 1.$$

Los físicos e ingenieros se dieron cuenta de la utilidad de tales ideas y procedieron a usar la función δ en sus cálculos. Pero los matemáticos se regodeaban señalando que la función δ carecía de sentido. Entonces para aumentar la diversión de los matemáticos, Dirac procedió a derivar dicha función. Esta idea es una herramienta indispensable en la actualidad en las matemáticas aplicadas, y alrededor de 1948, fundamentalmente gracias a los trabajos de Schwartz (1915-2002) y Sóbolev (1908-1989), alcanzó sentido la δ de Dirac con la denominada teoría de distribuciones.

En este apartado, vamos a echar un breve vistazo a esta teoría.

La idea de la derivación débil es traspasar a funciones de referencia el peso de las propiedades de regularidad, descargando de esta imposición a las denominadas *funciones generalizadas* o *distribuciones*.

Definición 2.5.1. Se designa por $\mathcal{D}(\mathbb{R})$ (también por $\mathcal{C}_c^\infty(\mathbb{R})$) al conjunto de las funciones complejas definidas y de clase $\mathcal{C}^\infty$ en $\mathbb{R}$ y de soporte compacto. Sus elementos son las denominadas *funciones test*. Este conjunto dotado de las operaciones habituales de suma de funciones y producto de un escalar por una función, es un espacio vectorial, el denominado *espacio de funciones test*.

Definición 2.5.2. Sea $\{\varphi_n\}_{n=1}^\infty$ una sucesión de elementos de $\mathcal{D}(\mathbb{R})$. Se dice que $\{\varphi_n\}_{n=1}^\infty$ converge hacia 0 en $\mathcal{D}(\mathbb{R})$ si:

i) Existe un intervalo compacto $[a, b]$ tal que $sop(\varphi_n) \subset [a, b]$ para todo $n \in \mathbb{N}$.

ii) Para cada $k \geq 0$ la sucesión de derivadas $\{\varphi_n^{(k)}\}_{n=1}^\infty$ converge uniformemente hacia 0 en $\mathbb{R}$.

Se dice que $\{\varphi_n\}_{n=1}^\infty$ converge hacia $\varphi \in \mathcal{D}(\mathbb{R})$ (en el sentido de las distribuciones) si la sucesión $\{\varphi_n - \varphi\}_{n=1}^\infty$ converge hacia 0.

Definición 2.5.3. Se denomina distribución a toda aplicación lineal y continua $T\colon \mathcal{D}(\mathbb{R}) \to \mathbb{C}$. Esto es:

i) Si $\varphi, \phi \in \mathcal{D}(\mathbb{R})$ y $\alpha, \beta \in \mathbb{C}$ se verifica que

$$T(\alpha\varphi + \beta\phi) = \alpha T(\varphi) + \beta T(\phi).$$

ii) Para cada sucesión $\{\varphi_n\}_{n=1}^{\infty}$ de funciones test que converge hacia 0 en $\mathcal{D}(\mathbb{R})$ se tiene que la sucesión de números complejos $\{T(\varphi_n)\}_{n=1}^{\infty}$ converge hacia $0 \in \mathbb{C}$. Equivalentemente,

$$\text{si } \varphi_n \underset{n\to\infty}{\longrightarrow} \varphi \text{ en } \mathcal{D}(\mathbb{R}), \text{ entonces } T(\varphi_n) \underset{n\to\infty}{\longrightarrow} T(\varphi) \text{ en } \mathbb{C}.$$

Ejemplos 2.5.4.

i) Sea a un número real. La aplicación δ_a definida en $\mathcal{D}(\mathbb{R})$ por

$$\delta_a(\varphi) = \langle \delta_a, \varphi \rangle = \varphi(a)$$

es una distribución que se denomina delta de Dirac en el punto a. Se suele escribir $\delta_0 = \delta$.

ii) Sea f una función localmente integrable en $\mathbb{R}$ (i.e., integrable en cada intervalo compacto). La aplicación T_f definida en $\mathcal{D}(\mathbb{R})$ por

$$T_f(\varphi) = \langle T_f, \varphi \rangle = \int_{-\infty}^{\infty} f(x)\varphi(x)dx = \int_{sop(\varphi)} f(x)\varphi(x)dx$$

es una distribución. También se dice que f es una distribución, se representa T_f simplemente por f y se escribe

$$\langle f, \varphi \rangle = \int_{-\infty}^{\infty} f(x)\varphi(x)dx.$$

iii) Un caso particular de lo anterior es la denominada función de Heaviside, definida por $H(x) = \mathcal{X}_{[0,\infty)}$ o, en general, la función escalón en el punto $a \in \mathbb{R}$ dada por $H_a = \mathcal{X}_{[0,\infty)}$, es decir $H_a(x) = H(x-a)$, cuya distribución asociada viene dada por

$$\langle H_a, \varphi \rangle = \int_a^{\infty} \varphi(x)dx.$$

Consideremos una función f derivable con continuidad en $\mathbb{R}$ (esto implica su caracter localmente integrable). Si φ es una función test, su soporte estará contenido en un intervalo acotado, digamos $[a,b]$, y lo mismo ocurre con los soportes de sus derivadas sucesivas (en particular, debe ser $\varphi^{(k)}(a) = \varphi^{(k)}(b) = 0$ para todo $k \geq 0$). Entonces, de la fórmula de integración por partes se deduce que

$$\begin{aligned}\int_{-\infty}^{\infty} f'(x)\varphi(x)dx &= \int_a^b f'(x)\varphi(x)dx = f(x)\varphi(x)\Big|_{x=a}^{x=b} - \int_a^b f(x)\varphi'(x)dx \\ &= -\int_{-\infty}^{\infty} f(x)\varphi'(x)dx\end{aligned}$$

y, en general, si f es de clase $\mathcal{C}^k$

$$\int_{-\infty}^{\infty} f^{(k)}(x)\varphi(x)dx = (-1)^k \int_{-\infty}^{\infty} f(x)\varphi^{(k)}(x)dx.$$

Esto sugiere la siguiente definición.

Definición 2.5.5. Si T es una distribución y $k \in \mathbb{N}$, la distribución

$$\varphi \in \mathcal{D}(\mathbb{R}) \mapsto (-1)^k \left\langle T, \varphi^{(k)} \right\rangle \in \mathbb{C}$$

se denomina derivada k-ésima de T y se denota por $T^{(k)}$.

Ejemplos 2.5.6.

i) Si $a \in \mathbb{R}$ las derivadas sucesivas de δ_a vienen dadas por

$$\delta_a^{(k)}(\varphi) = (-1)^k \delta_a(\varphi^{(k)}) = (-1)^k \varphi^{(k)}(a), \quad \varphi \in \mathcal{D}(\mathbb{R}).$$

ii) La derivada de la función de Heaviside en el sentido de las distribuciones es

$$\begin{aligned}(T_H)'(\varphi) &= -\int_{-\infty}^{\infty} H(x)\varphi'(x)dx = -\int_{0}^{\infty} H(x)\varphi'(x)dx \\ &= -\varphi(x)\Big|_{x=0}^{x\to\infty} = \varphi(0), \quad \varphi \in \mathcal{D}(\mathbb{R}).\end{aligned}$$

Es decir, la derivada de H en el sentido de las distribuciones es la delta de Dirac en 0.

Capítulo 3

Ejemplos patológicos

A principios del siglo XIX, el pensamiento general de la comunidad matemática era que una función continua debía poseer derivada en un conjunto significativo de puntos, es decir la derivada podía no existir o ser infinita en algunos puntos aislados del dominio de la función, como ocurre por ejemplo con la función valor absoluto $f(x) = |x|$.

Esto es debido a que, en general, estaban más interesados en el cálculo explícito de la derivada de una función con vistas a la modelización analítica de los fenómenos naturales (*"natura non facit saltus"*, Leibniz) que en la hipotética existencia de una función continua que no poseyera derivada en ningún punto.

Es esta filosofía que asume la regularidad de la naturaleza la que llevó a equiparar continuidad y diferenciabilidad.

Usando la cita que nos proporciona Charles Hermite en una carta dirigida a Stieltjes ([20], Vol.3 pag. 1285):

"Me aparto con mucho miedo y horror de esta lamentable plaga de funciones que no tienen derivada."

No es de extrañar la reacción de los matemáticos del siglo XIX ante la existencia de estas funciones.

No obstante, la pretensión de Hermite no resuelve la asunción de "analiticidad", pues en 1821 Cauchy proporcionó un ejemplo de una función continua e indefinidamente derivable, es decir, de clase $\mathcal{C}^\infty$ en $\mathbb{R}$, pero que no coincide con su serie de Taylor en el punto $x_0 = 0$, puesto que la función a la derecha del punto es positiva. Dicha función es:

$$f(x) = \begin{cases} 0 & si \quad x \leq 0 \\ e^{-1/x} & si \quad x > 0 \end{cases}$$

además, se tiene que para un $x > 0$

$$f^{(n)}(x) = P_n(1/x)e^{-1/x}, \quad n \in \mathbb{N}$$

donde P_n es un polinomio de grado $2n$.

Sin embargo, la existencia de las funciones continuas pero no diferenciables en ningún punto ocasionó una reconsideración del concepto de la función continua y un incremento del rigor en el Análisis Matemático.

Nosotros en este capítulo vamos a presentar diversas funciones continuas no derivables en ningún punto, o derivables en conjuntos "pequeños".

3.1. Función de Bolzano

Quizás el primer caso conocido de una función continua no derivable en ningún punto sea debido al matemático cheho Bolzano, que se encuentra en el manuscrito *"Funktionenlehre"* (Lecciones de Funciones) escrito sobre el año 1834, pero la función fue descubierta alrededor del año 1830. El matemático Karel Rychlik (1885-1968) fue quien publicó la función en 1930, aunque en 1922 ya había demostrado que la función no era diferenciable en ningún punto.

La función de Bolzano se basa en una construcción geométrica en lugar de una aproximación mediante una serie infinita, y se define como el límite de una sucesión $\{B_k\}_{k\in\mathbb{N}}$ de funciones continuas y lineales a trozos. Para la construcción empezamos denotando $[a, b]$ y $[\alpha, \beta]$ como el dominio y el rango de la función respectivamente. Ahora comenzamos a definir las funciones.

La función B_1 lineal y continua en el intervalo $[a, b]$ es de la forma:

$$B_1(x) := \alpha + (x - a)\frac{\beta - \alpha}{b - a}$$

donde $B_1(a) := \alpha$ y $B_1(b) := \beta$.

Para definir la función B_2 continua y lienal a trozos, Bolzano divide el intervalo $[a, b]$ en cuatro subintervalos limitados por los puntos:
$a,\ a + \frac{3}{8}(b - a),\ \frac{1}{2}(a + b),\ a + \frac{7}{8}(b - a)$ y b, es decir

i) $I_1 := \big[a, a + \frac{3}{8}(b - a)\big]$.

ii) $I_2 := \big[a + \frac{3}{8}(b - a), \frac{1}{2}(a + b)\big]$.

iii) $I_3 := \big[\frac{1}{2}(a + b), a + \frac{7}{8}(b - a)\big]$.

iv) $I_4 := \big[a + \frac{7}{8}(b - a), b\big]$.

y a esos puntos, les asigna los siguientes valores:

i) $B_2(a) := \alpha$.

ii) $B_2(a + \frac{3}{8}(b-a)) := \alpha + \frac{5}{8}(\beta - \alpha)$.

iii) $B_2(\frac{1}{2}(a+b)) := \alpha + \frac{1}{2}(\alpha - \beta)$.

iv) $B_2(a + \frac{7}{8}(b-a)) := \beta + \frac{1}{8}(\beta - \alpha)$.

v) $B_2(b) := \beta$

La funciones B_k restantes se definen análogamente, siguiendo el mismo proceso en cada subintervalo.

Observación 3.1.1. La demostración original de Bolzano era incompleta por dos motivos:

i) La prueba de la continuidad se basaba en la afirmación errónea de que el límite de una sucesión de funciones continuas es siempre una función continua (se vuelve cierta si se requiere convergencia uniforme).

ii) Para justificar que la función no es diferenciable en ningún punto del intervalo $[a, b]$, Bolzano demostró que cuando la función no tiene una derivada en dos puntos diferentes, entonces hay un punto entre ellos donde nuevamente la derivada no existe. Esto sólo concluye que en el conjunto de puntos donde la función no es derivable es denso en $[a, b]$.

3.2. Función de Weierstrass

El matemático alemán Weierstrass en unas conferencias en el año 1861, dijo que cualquier intento de demostrar que la diferenciabilidad va unido a la continuidad debe ser fallido. Posteriormente, en la conferencia del 18 de Julio de 1872 en la academia de Berlín, presentó una función continua y no diferenciable en ningún punto. No obstante, este ejemplo fue publicado primero por Paul du Bois-Reymond (1859-1938) en 1875, pues Weierstrass se lo comunicó en una carta en 1874.

La función en cuestión venía definida por:

$$f(x) = \sum_{k=0}^{\infty} b^k \cos(a^k \pi x),$$

siendo a un entero impar mayor que 1 y b un número real con $0 < b < 1$ verifiando que $ab > 1 + 3\pi/2$.

Antes de que Weierstrass construyera su función, el matemático suizo Charles Cellérier (1818-1889) en el año 1860 construyó una función continua y no diferenciable en ningún punto, aunque no fue publicada hasta 1890. Su función es de la forma

$$C(x) = \sum_{n=1}^{\infty} \frac{\sin(a^n x)}{a^n},$$

donde a es un entero suficientemente grande. También, sobre el año 1861 Riemann (1826-1866) presentó otro ejemplo de una función continua y no diferenciable en ningún intervalo abierto, concretamente, la función tiene derivada finita en los puntos de la forma $(2q+1)\pi/(2p+1)$ con $p, q \in \mathbb{Z}$, mientras que en el resto de puntos del intervalo no es diferenciable. Dicha función es de la forma

$$R(x) = \sum_{n=1}^{\infty} \frac{\sin(n^2 x)}{n^2}.$$

Aunque existieran otras construcciones de funciones continuas y no diferenciables en ningún punto antes de la de Weierstrass, su ejemplo fue el que hizo ver claramente la distinción entre continuidad y diferenciabilidad.

Si nos fijamos bien, estas tres funciones están asociadas a un serie trigonométrica, contradiciendo la suposición de Fourier de que la derivada de una función asociada a una serie trigonométrica se obtenía derivando término a término la serie.

Teorema 3.2.1. Sea $b \in (0,1)$ y a un entero impar mayor que 1 tal que $ab > 1 + \frac{3\pi}{2}$. Entonces la función

$$f(x) = \sum_{k=0}^{\infty} b^k \cos(a^k \pi x)$$

es continua y no derivable en ningún punto.

Demostración. Primero vamos a ver que la función $f(x)$ es continua para todo $x \in \mathbb{R}$. Para ello veamos que la serie converge normalmente en $\mathbb{R}$. En efecto, sea $g_k(x) = b^k \cos(a^k \pi x)$, tomando valores absolutos tenemos que

$$|g_k(x)| = |b^k \cos(a^k \pi x)| \leq b^k \quad \forall k \in \mathbb{N} \quad \text{y} \quad \forall x \in \mathbb{R},$$

puesto que $0 \leq |\cos(a^k \pi x)| \leq 1$, y como la serie $\sum_{k=0}^{\infty} b^k$ converge, dado que $0 < b < 1$, tenemos que la serie que define a f converge normalmente. Por el criterio de Weierstrass, f converge absoluta y uniformemente en $\mathbb{R}$, y por el criterio de convergencia uniforme de Cauchy, f es continua en $\mathbb{R}$.

Ahora veamos que la función $f(x)$ no es diferenciable en ningún punto de $\mathbb{R}$. Para ello probaremos que

$$\limsup_{h \to 0} \left| \frac{f(x+h) - f(x)}{h} \right| = \infty,$$

concretamente, vamos a construir una sucesión convergente hacia 0, $\{h_n\}_{n=1}^{\infty}$, tal que

$$\lim_{n\to\infty} \left| \frac{f(x+h_n)-f(x)}{h_n} \right| = \infty.$$

Sea $x \in \mathbb{R}$ fijo y $h > 0$, entonces

$$\frac{f(x+h)-f(x)}{h} = \sum_{k=0}^{\infty} b^k \frac{\cos(a^k\pi(x+h)) - \cos(a^k\pi x)}{h} = S_n + R_n,$$

donde

$$S_n = \sum_{k=0}^{n-1} b^k \frac{\cos(a^k\pi(x+h)) - \cos(a^k\pi x)}{h}$$

y

$$R_n = \sum_{k=n}^{\infty} b^k \frac{\cos(a^k\pi(x+h)) - \cos(a^k\pi x)}{h}.$$

Comencemos estimando $|S_n|$.

Aplicando el teorema del valor medio a la función $\cos(x)$ en el itervalo I, donde $I = [a^k\pi x, a^k\pi(x+h)]$, tenemos que existe $\xi \in I$ tal que

$$\cos(a^k\pi(x+h)) - \cos(a^k\pi x) = (a^k\pi(x+h) - a^k\pi x)(-\sin(\xi\pi)) = -a^k\pi h \sin(\xi\pi),$$

por lo tanto,

$$\frac{\cos(a^k\pi(x+h)) - \cos(a^k\pi x)}{h} = -a^k\pi \sin(\xi\pi).$$

Tomando valores absolutos obtenemos que

$$\left| \frac{\cos(a^k\pi(x+h)) - \cos(a^k\pi x)}{h} \right| = \left| -a^k\pi \sin(\xi\pi) \right| \leq a^k\pi$$

puesto que $0 \leq |\sin(\xi\pi)| \leq 1$. Entonces

$$\begin{aligned} |S_n| &= \left| \sum_{k=0}^{n-1} b^k \frac{\cos(a^k\pi(x+h)) - \cos(a^k\pi x)}{h} \right| \\ &\leq \sum_{k=0}^{n-1} a^k b^k \pi = \pi \frac{1-a^n b^n}{1-ab} = \pi \frac{a^n b^n - 1}{ab-1} < \frac{\pi a^n b^n}{ab-1}. \end{aligned}$$

La segunda igualdad se debe a que es una suma finita de términos en progresión geométrica.

Ahora para poder estimar $|R_n|$ definimos los siguientes valores:

i) Sea $a^n x = \alpha_n + \beta_n$, donde $\alpha_n \in \mathbb{Z}$ y $-1/2 \leq \beta_n < 1/2$.
Nótese que α_n y β_n están unívocamente determinados pues

$$\mathbb{R} = \bigcup_{k\in\mathbb{Z}} \left[k - \frac{1}{2}, k + \frac{1}{2}\right).$$

ii) Sea $h_n = (1 - \beta_n)/a^n$.
Es obvio que $\lim_{n\to\infty} h_n = 0$, pues $(1 - \beta_n)$ está acotado y $1/a^n \to 0$ cuando $n \to \infty$.

Para $k \geq n$, consideramos

$$a^k\pi(x + h_n) = a^{k-n}a^n\pi(x + h_n) = a^{k-n}\pi(a^n x + 1 - \beta_n) = a^{k-n}\pi(1 + \alpha_n),$$

donde la segunda igualdad se debe por lo definido en ii) y la última igualdad se debe por lo definido en i).

Como a es un número impar, se verifican las siguientes igualdades

$$\cos(a^k\pi(x + h_n)) = \cos(a^{k-n}\pi(1 + \alpha_n)) = (-1)^{1+\alpha_n}, \tag{3.1}$$

donde la última igualdad se debe a las propiedades del coseno, es decir,

1) Si $1 + \alpha_n$ es impar, como a^{k-n} es impar, tenemos que

$$\cos(a^{k-n}\pi(1 + \alpha_n)) = -1.$$

2) Si $1 + \alpha_n$ es par, como a^{k-n} es impar, tenemos que

$$\cos(a^{k-n}\pi(1 + \alpha_n)) = 1.$$

También obtenemos que

$$-\cos(a^k\pi x) = -\cos(a^{k-n}a^n\pi x) = -\cos(a^{k-n}\pi(\alpha_n + \beta_n)), \tag{3.2}$$

por la suma de ángulos del coseno se tiene que

$$-\cos(a^{k-n}\pi(\alpha_n + \beta_n)) = -\cos(a^{k-n}\pi\alpha_n)\cos(a^{k-n}\pi\beta_n) \tag{3.3}$$

pues $\sin(a^{k-n}\pi\alpha_n) = 0$.

Además, por el mismo razonamiento de antes de las propiedades del coseno, llegamos a que

$$-\cos(a^{k-n}\pi\alpha_n) = -(-1)^{\alpha_n} = (-1)^{1+\alpha_n}. \tag{3.4}$$

Por lo tanto, de (3.2), (3.3) y (3.4) se sigue que

$$-\cos(a^k\pi x) = (-1)^{1+\alpha_n}\cos(a^{k-n}\pi\beta_n). \tag{3.5}$$

Estableciendo $h = h_n$, de (3.1) y (3.5), se obtiene que

$$\begin{aligned}
|R_n| &= \left| \sum_{k=n}^{\infty} b^k \frac{(-1)^{1+\alpha_n} + (-1)^{1+\alpha_n} \cos(a^{k-n}\pi\beta_n)}{h_n} \right| \\
&= \left| \frac{(-1)^{1+\alpha_n}}{h_n} \right| \sum_{k=n}^{\infty} b^k \left(1 + \cos(a^{k-n}\pi\beta_n)\right) \\
&\geq \frac{1}{h_n} \sum_{k=n}^{\infty} b^k \geq \frac{b^n}{h_n} \geq \frac{2a^n b^n}{3}.
\end{aligned}$$

Esto se debe a:

i) $\sum_{k=n}^{\infty} b^k \left(1 + \cos(a^{k-n}\pi\beta_n)\right) \geq \sum_{k=n}^{\infty} b^k$, pues $\cos(\beta_n\pi) \geq 0$ puesto que $\beta_n\pi \in [-\frac{\pi}{2}, \frac{\pi}{2}]$.

ii) $\sum_{k=n}^{\infty} b^k \geq b^n$ para cualquier $n \in \mathbb{N}$, pues todos los términos de la serie son positivos.

iii) $\frac{1}{h_n} \geq \frac{2a^n}{3}$. Como $-\frac{1}{2} \leq \beta_n < \frac{1}{2}$, tenemos que $\frac{1}{2} < 1 - \beta_n \leq \frac{3}{2}$ y $\frac{1}{2a^n} < \frac{1-\beta_n}{a^n} \leq \frac{3}{2a^n}$, por lo tanto $\frac{2a^n}{3} \leq \frac{1}{h_n} < 2a^n$.

Combinando las estimaciones de $|R_n|$ y $|S_n|$ llegamos a

$$\begin{aligned}
\left| \frac{f(x+h_n) - f(x)}{h_n} \right| &\geq |R_n| - |S_n| > \frac{2a^n b^n}{3} - \frac{\pi a^n b^n}{ab-1} \\
&= (ab)^n \left(\frac{2}{3} - \frac{\pi}{ab-1} \right) > (ab)^n.
\end{aligned}$$

La ultima desiguladad se debe a que, como $ab > 1 + \frac{3\pi}{2}$, tenemos que $ab - 1 > \frac{3\pi}{2}$ y $\frac{1}{ab-1} < \frac{2}{3\pi}$, por lo tanto $\frac{\pi}{ab-1} < \frac{2}{3}$, entonces $\left(\frac{2}{3} - \frac{\pi}{ab-1}\right) > 0$.

Entonces llegamos a la conclusión de que cuando $n \to \infty$, $h_n \to 0$ y $f'(x)$ no existe, pues

$$\lim_{n\to\infty} \left| \frac{f(x+h_n) - f(x)}{h_n} \right| = \infty.$$

□

Observación 3.2.2.

i) Esta demostración se puede consultar en [18] pag. 258-260.

ii) La demostración original de Weierstrass se basa en la construcción de dos sucesiones de incrementos tales que los cocientes incrementales asociados a una de ellas están acotados inferiormente por 2/3, mientras que los asociados a la otra están acotados superiormente por $-2/3$.

iii) Más tarde, Hardy probó que si $ab > 1$ la función de Weierstrass sigue siendo continua y no diferenciable en ningún punto. Además, cambiando la función coseno por la función seno, de forma que

$$f(x) = \sum_{k=0}^{\infty} b^k \sin(a^k \pi x)$$

se sigue verificando que f es continua y no diferenciable en ningún punto.

En la figura 3.1, se representa una función de Weierstrass, concretamente para los valores $a = 8$ y $b = 3/4$. En realidad lo que se representa es la suma parcial S_2, ya que para índices superiores las frecuencias son mayores, es decir, en cada ciclo de la función se volvería a poner la gráfica que tenemos y no apreciaríamos nada.

La escala de la gráfica está distorsionada, por razones estéticas.

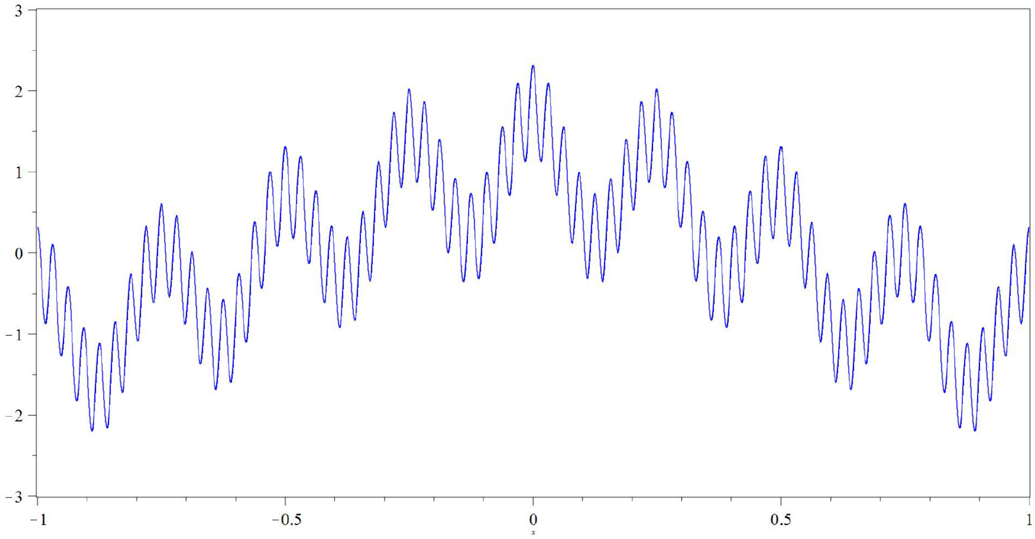

Figura 3.1: Aproximación de orden 2 a la función de Weierstrass.

3.3. Las funciones de Takagi y Van der Waerden

Alejado de las series trigonométricas, el matemático japonés Teiji Takagi (1875-1903) publicó en 1903 un ejemplo de una función continua y no diferenciable en ningún punto. Más tarde, en el año 1930 el matemático holandés

Bartel Van der Waerden (1903-1996) publicó sin percatarse una función muy parecida a la de Takagi.

Ambas funciones toman como punto de partida la función periódica ϕ definida por $\phi = \text{dist}(x, \mathbb{Z})$ con $x \in \mathbb{R}$, esto es mín $\{x - \lceil x \rceil - 1, \lceil x \rceil - x\}$ con $x \in \mathbb{R}$, donde $\lceil x \rceil$ está definido en la observación 2.4.17.

Se define la función F_a con $a > 1$ mediante la serie

$$F_a(x) = \sum_{k=1}^{\infty} \frac{\phi(a^k x)}{a^k}, \; x \in \mathbb{R}.$$

Cuando $a = 2$ se tiene la función de Takagi y en el caso $a = 10$ hablamos de la función de Van de Waerden.

Observación 3.3.1.

i) Si nos fijamos bien, la función de Takagi se basa en la representación binaria de los números enteros, mientras que la función de Var der Waerden se basa en la representación decimal.

ii) Si nos remontamos a la época de Euler, la función "dist" no se consideraría una función continua, pues no está expresada por una única expresión analítica.

3.4. Otras Patologías

La función de Weierstrass abrió la puerta a otro tipo de patologías, como la que contradice la idea intuitiva de curva como conjunto de "área cero". A continuación, vamos a ver una serie de ejemplos de curvas que se obtinen mediante una sucesión de curvas continuas que convergen a una curva límite. A estas curvas se las conoce con el nombre de *"curvas que rellenan todo el espacio"* (especificamente, la curva es un conjunto denso en el plano).

Nota. Es suficiente con rellenar el cuadrado $[0, 1] \times [0, 1]$, luego mediante traslaciones, giros y simetrías es posible teselar todo el plano.

3.4.1. La curva de Peano

El matemático italiano Guiseppe Peano (1858-1932) fue el pionero en la construcción de curvas que rellenan conjuntos planos de área mayor que cero. En el año 1890 proporcionó su construcción, que se obtiene como límite uniforme de una sucesión de curvas continuas $\gamma_n \colon [0, 1] \to \mathbb{R}^2$, construidas recurrentemente como sigue:

En primer lugar se considera el cuadrado $[0, 1] \times [0, 1]$ y una de sus diagonales, el segmento de extremos $(0, 0)$ y $(1, 1)$. A continuación se divide el cuadrado anterior en nueve cuadrados iguales y, empezando en $(0,0)$, se recorren estos cuadrados uniendo el vértice al que se ha llegado con el vértice opuesto del cuadrado adyacente encima o debajo, o a la derecha en el caso de que el vértice de partida se encuentre en el borde del cuadrado inicial, terminando finalmente en el otro vértice $(1,1)$. Para comprenderlo mejor, ver la figura 3.2, dónde las flechas indican en que sentido se recorren los segmentos, y los números el orden en que se hace. Cada uno de estos 9 segmentos se parametriza (mediante una aplicación afín) en un intervalo de la forma $[k/9,(k+1)/9]$.

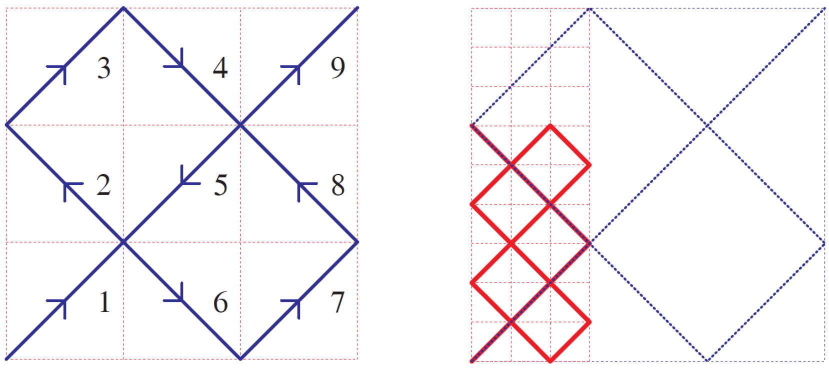

Figura 3.2: Curva γ_1 (izquierda) y construcción de γ_2 a partir de ella

En general, la curva γ_n consta de 9^n segmentos, que son diagonales de otros cuadrados semejantes al inicial y cuyos lados tienen longitud $1/3^n$.

Si nos fijamos en la figura 3.2, vemos la construcción de γ_2 a partir de γ_1, donde podemos observar que cada cuadrado de γ_1 se divide en 9 cuadrados iguales y se construye una poligonal de 9 lados cuyos puntos inicial y final coinciden con los segmentos de γ_1.

Por tanto, como cada uno de los 9^n segmentos de γ_n se parametriza por una función afín en un intervalo de la forma $[k/9^n,(k+1)/9^n]$ con $k = 0,1,\ldots,9^{n-1}$, entonces γ_n converge uniformemente, pues

$$||\gamma_n(t)-\gamma_{n+1}(t)|| \leq \sqrt{2}/3^n \text{ para cada } t \in [0,1].$$

3.4.2. La curva de Hilbert

El matemático alemán David Hilbert (1862-1943) hizo un año más tarde una variación de la curva de Peano. Hilbert describe la curva como una curva fractal continua, en la que cada una de las curvas que aproximan la curva

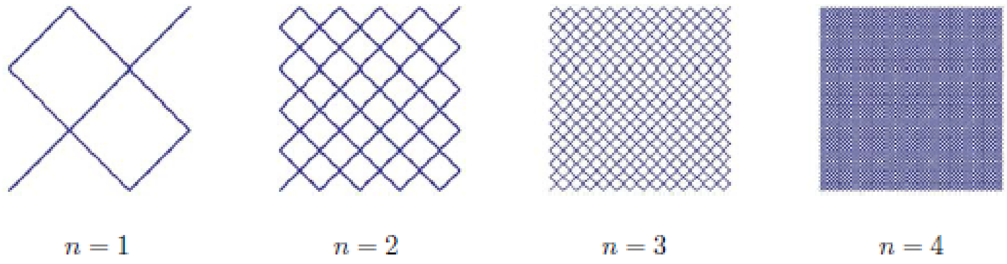

Figura 3.3: Curvas de Peano.

final es simple, es decir, no se corta a sí misma. La construcción es más simple que la de Peano:

Primero se parte de un cuadrado unidad que se divide en cuatro subcuadrados iguales y se unen sus centros con segmentos, empezando por el cuadrado inferior izquierdo y acabando por el cuadrado inferior derecho. Posteriormente, con cada subcuadrado se realiza el mismo procedimiento que en el paso anterior, uniendo sus centros con segmentos y empezando también por el cuadrado inferior izquierdo y terminando por el cuadrado inferior derecho. En el paso n-ésimo se obtienen 4^n cuadrados y de esta forma la curva esta constituida por 4^{n-1} segmentos.

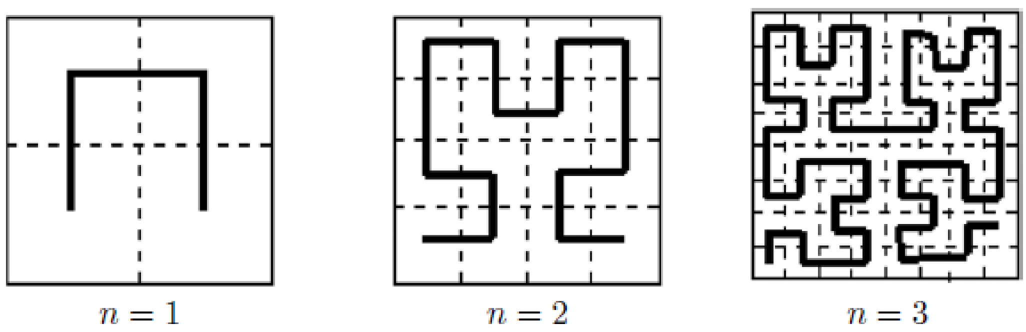

Figura 3.4: Curvas de Hilbert

3.4.3. La curva de Schoenberg

El matemático Isaac Schoenberg (1903-1990) construyó en 1938 una curva basándose en la representación binaria de los números reales. Esta curva es más difícil de visualizar geométricamente que las anteriores, pero la descripción analítica es mucho más sencilla:

Se define la función ϕ en el intervalo $[0,2]$ por

$$\phi(t) = \begin{cases} 0 & si \quad t \in [0, \frac{1}{3}], \\ 3t-1 & si \quad t \in [\frac{1}{3}, \frac{2}{3}], \\ 1 & si \quad t \in [\frac{2}{3}, \frac{4}{3}], \\ -3t+5 & si \quad t \in [\frac{4}{3}, \frac{5}{3}], \\ 0 & si \quad t \in [\frac{5}{3}, 2], \end{cases}$$

y se extiende por periodicidad a toda la recta real mediante la relación $\phi(t+2) = \phi(t)$.

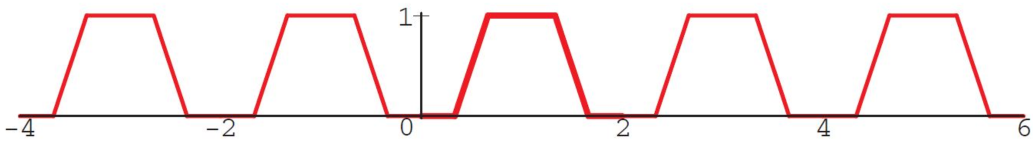

Figura 3.5: Función ϕ

Así ϕ resulta ser continua, y puesto que las dos series funcionales

$$\sum_{n=1}^{\infty} \frac{\phi(3^{2n-2}t)}{2^n} \quad y \quad \sum_{n=1}^{\infty} \frac{\phi(3^{2n-1}t)}{2^n}$$

convergen uniformemente en $\mathbb{R}$ definen sendas funciones continuas, $x(t)$ e $y(t)$, y no diferenciables en ningún punto (ver [30]). La imagen de la curva $\varphi : [0,1] \to \mathbb{R}^2$ dada por

$$\varphi(t) = (x(t), y(t))$$

es el cuadrado $[0,1] \times [0,1]$.

El estudio exhaustivo de la curva de Schoenberg puede encontrase en [1], pag. 272-273.

Capítulo 4

El teorema de Lebesgue

Después de que Weierstrass mostrara a la comunidad matemática la existencia de funciones continuas pero no diferenciables en ningún punto, quedaba la pregunta de si una función continua "buena" debía ser diferenciable. Lo más simple, sería insistir en una función monótona o monótona a trozos. Varias personas buscaron una prueba de que una función monótona continua tendría que ser diferenciable en todo punto salvo en un número finito.

Weierstrass responde con una función monótona continua que no es diferenciable en ningún número racional. Luego buscó esa función que no es diferenciable en ningún punto. Aunque nunca encontró una, creía que debía existir. Pero incluso el genio de Weierstrass se equivocó alguna vez.

Más tarde, Lebesgue publicó en 1904 (ver [25]) que una función monótona (y continua) debía ser diferenciable en casi todo punto (en sus tesis de 1903 ya aparecía este resultado).

En 1911 William Young presentó una prueba de este resultado sin asumir la continuidad. En 1932 Riez dio una prueba elemental de este resultado mediante el uso de su lema "del sol naciente". La prueba de Riesz es simple y elegante para el caso continuo. En 1965, Donald Austin proporcionó una prueba geométrica usando la teoría de la medida y la medibilidad de las derivadas de Dini [2]. Aparte del trabajo de Riesz, las pruebas del teorema de Legesgue que hemos visto implican la teoría de la medida, el teorema de recubrimiento de Vitali o el teorema de densidad de Lebesgue (ver [18], [27]).

Nuestro objetivo en este capítulo, es dar una demostración elemental del teorema de Lebesgue usando los conjuntos de medida nula y las derivadas superior e inferior de la función (estas funciones no necesitan ser medibles). Seguiremos el artículo de Botsko [5] desarrollándolo con más detalle.

4.1. Demostración del teorema de Lebesgue

Antes de llegar al teorema de Lebesgue, vamos a dar una serie de resultados que nos servirán para la demostración.

Lema 4.1.1. Sean $f\colon [a,b] \to \mathbb{R}$ una función, $P = \{x_0, x_1, \dots, x_n\}$ una partición de $[a,b]$, S un subconjunto no vacio de $\{1,2,3,\dots,n\}$ y $A > 0$. Supongamos que se verifica una de las dos condiciones siguientes:

a) $f(a) \leq f(b)$ y $\dfrac{f(x_k) - f(x_{k-1})}{x_k - x_{k-1}} < -A$

b) $f(a) \geq f(b)$ y $\dfrac{f(x_k) - f(x_{k-1})}{x_k - x_{k-1}} > A$

para cada $k \in S$. Entonces, se tiene que

$$\sum_{k=1}^{n} |f(x_k) - f(x_{k-1})| > |f(b) - f(a)| + AL,$$

donde $L = \sum_{k\in S}(x_k - x_{k-1})$.

Demostración. Probaremos el lema con el primer supuesto. Como $f(a) \leq f(b)$, tenemos que

$$\begin{aligned}
|f(b) - f(a)| = f(b) - f(a) &= \sum_{k=1}^{n} (f(x_k) - f(x_{k-1})) \\
&= \sum_{k\in S} (f(x_k) - f(x_{k-1})) + \sum_{k\notin S} (f(x_k) - f(x_{k-1})) \\
&< -A\sum_{k\in S}(x_k - x_{k-1}) + \sum_{k\notin S} (f(x_k) - f(x_{k-1})) \\
&\leq -AL + \sum_{k=1}^{n} (f(x_k) - f(x_{k-1})) \\
&\leq -AL + \sum_{k=1}^{n} |f(x_k) - f(x_{k-1})|.
\end{aligned}$$

Por lo tanto,

$$\sum_{k=1}^{n} |f(x_k) - f(x_{k-1})| > |f(b) - f(a)| + AL.$$

El segundo supuesto se reduce al caso anterior considerando $g = -f$. □

Observación 4.1.2. Si $E \subset \mathbb{R}$ no es de medida nula, por definición, existe $\varepsilon > 0$ tal que $\sum_{n=1}^{\infty} \mu(J_n) \geq \varepsilon$ para cualquier sucesión $\{J_n\}_{n=1}^{\infty}$ de intervalos abiertos que recubran E, donde μ denota la medida de Lebesgue, que en $\mathbb{R}$ es la longitud del intervalo. En otras palabras

$$\mu^*(E) = \text{ínf} \left\{ \sum_{n=1}^{\infty} \mu(J_n) : E \subset \bigcup_{n=1}^{\infty} J_n,\ J_n \text{ intervalos} \right\} \geq \varepsilon$$

donde $\mu^*(E)$ es la denominada *medida exterior de E*.

Lema 4.1.3. Sea E un subconjunto de (a, b) que no es de medida nula y sea $\varepsilon > 0$ su medida exterior asociada. Sea $\mathfrak{J}$ cualquier colección de subintervalos abiertos de $[a, b]$ que recubren E, entonces existe una subcolección finita disjunta $\{I_1, I_2, \ldots, I_N\}$ de $\mathfrak{J}$ tal que $\sum_{k=1}^{N} \mu(I_k) > \varepsilon/3$

Demostración. Puesto que $\bigcup_{I\in\mathfrak{J}} I$ es un conjunto abierto, existe una sucesión de intervalos abiertos disjuntos $\{(a_n, b_n)\}_{n=1}^{\infty}$ tal que $\bigcup_{I\in\mathfrak{J}} I = \bigcup_{n=1}^{\infty}(a_n, b_n)$. En el caso de que la unión sea finita el razonamiento es el mismo.

Como $\bigcup_{I\in\mathfrak{J}} I$ recubre E y $\bigcup_{I\in\mathfrak{J}} I = \bigcup_{n=1}^{\infty}(a_n, b_n)$, tenemos que $E \subseteq \bigcup_{n=1}^{\infty}(a_n, b_n)$, entonces

$$\sum_{n=1}^{\infty} \mu((a_n, b_n)) = \sum_{n=1}^{\infty}(b_n - a_n) \geq \varepsilon.$$

Ahora, elegimos para cada $n \in \mathbb{N}$ un subintervalo cerrado $[c_n, d_n]$ del intervalo abierto (a_n, b_n) tal que $d_n - c_n = 3(b_n - a_n)/4$. Por tanto,

$$\sum_{n=1}^{\infty}(d_n - c_n) \geq \frac{3\varepsilon}{4}.$$

Sea n un entero positivo fijo. Para cada $x \in [c_n, d_n]$, existe J_x en $\mathfrak{J}$ tal que $x \in J_x \subseteq (a_n, b_n)$. Entonces $\{J_x : x \in [c_n, d_n]\}$ es un abierto que recubre $[c_n, d_n]$.

Como $[c_n, d_n]$ es compacto, existe un número finito de intervalos abiertos $J_1, J_2, \ldots, J_p$ tal que

$$[c_n, d_n] \subseteq \bigcup_{k=1}^{p} J_k.$$

Además, podemos suponer, descartando algunos de los intervalos si es necesario, que ningún intervalo en $\{J_k\}_{k=1}^{p}$ es un subconjunto de la unión de los intervalos restantes en $\{J_k\}_{k=1}^{p}$. Es decir, cada J_i contiene un punto x_i que no pertenece a $\bigcup_{k\neq i} J_k$ y, volviendo a renumerar la familia si es necesario, podemos suponer que $x_1 < x_2 < \ldots < x_p$. Entonces, tanto $\{J_1, J_3, J_5, \ldots\}$ y $\{J_2, J_4, J_6, \ldots\}$ son subcolecciones finitas disjuntas de $\mathfrak{J}$.

Claramente se tiene que

$$\sum_k \mu(J_{2k-1}) \geq \frac{1}{2}\sum_{k=1}^{p} \mu(J_k) \qquad \text{o} \qquad \sum_k \mu(J_{2k}) \geq \frac{1}{2}\sum_{k=1}^{p} \mu(J_k)$$

por lo tanto, hemos encontrado una subcolección finita disjuna $\mathfrak{J}_n$ de $\mathfrak{J}$ tal que

$$\sum_{I\in\mathfrak{J}_n} \mu(I) \geq \frac{1}{2}\sum_{k=1}^{p} \mu(J_k) \geq \frac{1}{2}(d_n - c_n).$$

En consecuencia, para cada natural n existe $\mathfrak{J}_n$ una subcolección finita disjunta de $\mathfrak{J}$, y cada intervalo abierto de $\mathfrak{J}_n$ es un subconjunto de (a_n, b_n) tal que

$$\sum_{I\in\mathfrak{J}_n} \mu(I) \geq \frac{1}{2}(d_n - c_n).$$

Sumando ambos lados de la desigualdad anterior para $n = 1, 2 \ldots$, obtenemos que

$$\sum_{n=1}^{\infty}\sum_{I\in\mathfrak{J}_n} \mu(I) \geq \frac{1}{2}\sum_{n=1}^{\infty}(d_n - c_n) \geq \frac{3\varepsilon}{8}.$$

Por lo tanto, $\bigcup_{n=1}^{\infty} \mathfrak{J}_n = \{I_1, I_2, I_3, \ldots\}$ es una subcolección numerable disjunta de la familia $\mathfrak{J}$, para la cual $\sum_{k=1}^{\infty} \mu(I_k) \geq \frac{3\varepsilon}{8}$.

Finalmente, eligiendo N tan grande como se quiera, llegamos a

$$\sum_{k=1}^{N} \mu(I_k) > \frac{\varepsilon}{3}.$$

□

Observación 4.1.4. Sea E un subconjunto de (a, b) que no es de medida nula y sea $\varepsilon > 0$ su medida exterior asociada. Si P es un subconjunto finito de $[a, b]$, y $\mathfrak{J}$ una colección de subintervalos abiertos de $[a, b]$ que recubren $E \setminus P$, añadiendo a $\mathfrak{J}$ intervalos abiertos suficientemente pequeños centrados en cada punto de P y usando el lema 4.1.3 es fácil demostrar que existe una subcolección finita disjunta $\{I_1, I_2, \ldots, I_N\}$ de $\mathfrak{J}$ tal que

$$\sum_{k=1}^{N} \mu(I_k) > \frac{\varepsilon}{4}.$$

Teorema 4.1.5. Si I es un intervalo y $f\colon I \to \mathbb{R}$ es monótona, entonces el conjunto de las discontinuidades de f es numerable.

Demostración. Supongamos que f es creciente y que $I = [a, b]$.

Para cada $x \in (a, b)$, f tiene límite por la izquierda $f(x^-)$ y límite por la derecha $f(x^+)$, con $f(x^-) \leq f(x) \leq f(x^+)$. Que f sea discontinua en x, es equivalente a que $f(x^+) - f(x^-) > 0$.

El conjunto de las discontinuidades de f en (a, b) es la unión de los conjuntos

$$S_n = \left\{ x \in (a, b) : \ f(x^+) - f(x^-) > \frac{1}{n} \right\} \qquad n = 1, 2, 3, \ldots$$

por lo tanto, es suficiente demostrar que cada S_n es finito.

En general, si $r > 0$ el conjunto

$$S = \{x \in (a, b) : \ f(x^+) - f(x^-) > r\}$$

es finito. En efecto, supongamos que $x_1, x_2, \ldots, x_n$ son puntos de S con $x_1 < x_2 < \ldots < x_n$. Elegimos puntos $c_1, \ldots, c_n, c_{n+1}$ de (a, b) tal que

$$\begin{aligned} &c_1 < x_1, \\ &x_{i-1} < c_i < x_i \quad \text{para} \quad i = 2, \ldots, n, \\ &x_n < c_{n+1}. \end{aligned}$$

es decir, tenemos que

$$x_{i-1} < c_i < x_i \quad \text{para} \quad i = 1, \ldots, n,$$

y por consiguiente,

$$f(c_i) \leq f(x_i^-) \leq f(x_i^+) \leq f(c_{i+1})$$

asi que

$$f(x_i^+) - f(x_i^-) \leq f(c_{i+1}) - f(c_i)$$

sumando a ambos lados de esta última desigualdad desde $i = 1$ hasta $i = n$, obtenemos que

$$\sum_{i=1}^{n} (f(x_i^+) - f(x_i^-)) \leq \sum_{i=1}^{n} (f(c_{i+1}) - f(c_i)) = f(c_{n+1}) - f(c_1) \leq f(b) - f(a),$$

de donde se sigue que $f(b) - f(a) > nr$.

Esto muestra que S consta a lo sumo de $\dfrac{f(b) - f(a)}{r}$ elementos. $\square$

Teorema 4.1.6. Si f es una función creciente en $[a, b]$, entonces $f'(x)$ existe (finita o infinita) c.s. en $[a, b]$.

Demostración. Como f es continua excepto en un número numerable de puntos, entonces es suficiente con demostrar que el conjunto

$$F = \{x : x \in (a, b),\ f \text{ continua en } x \text{ y, } \overline{D}f(x) > \underline{D}f(x)\}$$

es de medida nula. Sea

$$E_{r,s} = \{x : x \in (a, b),\ f \text{ continua en } x \text{ y } \overline{D}f(x) > r > s > \underline{D}f(x)\}$$

con $r, s \in \mathbb{Q}$ y $r > s$.

El conjunto F es unión numerable de los conjuntos $E_{r,s}$. Como la unión de conjuntos de medida nula sigue siendo un conjunto de medida nula, para ver que el conjunto F es de medida nula, veremos que cada conjunto $E_{r,s}$ es de medida nula.

Para ello razonaremos por reducción al absurdo.

Supongamos que para algún r y s el conjunto $E = E_{r,s}$ no es de medida nula, y sea $\varepsilon > 0$ su medida exterior asociada.

Sean A y B dos números positivos definidos de la siguiente manera,

$$A = \frac{r - s}{2} \quad \text{y} \quad B = \frac{r + s}{2}$$

y sea $g = f - Bx$, entonces el conjunto E es de la siguiente manera

$$E = \{x : x \in (a, b),\ g \text{ continua en } x, \overline{D}f(x) > A \text{ y } \underline{D}f(x) < -A\}$$

Ahora el conjunto

$$\left\{ \sum_P |g(x_k) - g(x_{k-1})| :\ P \text{ partición de } [a, b] \right\}$$

está acotado superiormente:

$$\begin{aligned}
\sum_P |g(x_k) - g(x_{k-1})| &= \sum_P |(f(x_k) - f(x_{k-1})) - B(x_k - x_{k-1})| \\
&\leq \sum_P |f(x_k) - f(x_{k-1})| + B \sum_P |(x_k - x_{k-1})| \\
&= f(b) - f(a) + B(b - a).
\end{aligned}$$

Sea T el extremo superior de este conjunto.

Como A y ε son ambos positivos, existe una partición $P = \{x_0, x_1, \ldots, x_n\}$ de $[a, b]$ tal que

$$\sum_{k=1}^{n} |g(x_k) - g(x_{k-1})| > T - \frac{A\varepsilon}{4}. \tag{4.1}$$

Ahora sea $x \in E \setminus P$, entonces $x \in E \cap (x_{k-1}, x_k)$ para algún k.
Como $\overline{D}g(x) > A$, $\underline{D}g(x) < -A$ y g es continua en x, podemos elegir a_x y b_x con $(a_x, b_x) \subseteq (x_{k-1}, x_k)$ y satisfaciendo una de las dos alternativas siguientes:

$$g(x_{k-1}) \geq g(x_k) \quad \text{y} \quad \frac{g(x_k) - g(x_{k-1})}{x_k - x_{k-1}} > A$$

ó

$$g(x_{k-1}) \leq g(x_k) \quad \text{y} \quad \frac{g(x_k) - g(x_{k-1})}{x_k - x_{k-1}} < -A$$

Así $\mathfrak{J} = \{(a_x, b_x) : x \in E \setminus P\}$ es una colección de subintervalos abiertos de $[a, b]$ que recubren $E \setminus P$.
En virtud de lo expuesto en la observación 4.1.4, existe una subcolección finita disjunta $\{I_1, I_2, \ldots, I_N\}$ de $\mathfrak{J}$ tal que

$$\sum_{k=1}^{N} \mu(I_k) > \varepsilon/4. \tag{4.2}$$

Ahora sea $Q = \{y_0, y_1, \ldots, y_q\}$ una partición de $[a, b]$ determinada por los puntos de la partición P y los puntos finales de los intervalos $I_1, I_2, \ldots, I_N$. Para cada $[x_{k-1}, x_k]$ que contenga al menos uno de los intervalos $I_1, I_2, \ldots, I_N$ por el lema 4.1.1 tenemos que

$$\sum_{[y_{i-1}, y_i] \subseteq [x_{k-1}, x_k]} |g(y_i) - g(y_{i-1})| > |g(x_k) - g(x_k - 1)| + AL_k \tag{4.3}$$

donde la suma se toma sobre los intervalos cerrados determinados por Q que están contenidos en $[x_{k-1}, x_k]$, y L_k es la suma de las longitudes de los intervalos $I_1, I_2, \ldots, I_N$ que están contenidos en $[x_{k-1}, x_k]$.

Sumando las desigualdades (4.3) sobre $k = 1, \ldots, n$ se sigue que

$$\sum_{k=1}^{n} \left(\sum_{[y_{i-1}, y_i] \subseteq [x_{k-1}, x_k]} |g(y_i) - g(y_{i-1})| \right) = \sum_{k=1}^{q} |g(k_i) - g(y_{k-1})|$$

$$\sum_{k=1}^{n} (|g(x_k) - g(x_k - 1)| + AL_k) = \sum_{k=1}^{n} |g(x_k) - g(x_k - 1)| + A \sum_{k=1}^{N} \mu(I_k)$$

es decir,

$$\sum_{k=1}^{q} |g(k_i) - g(y_{k-1})| > \sum_{k=1}^{n} |g(x_k) - g(x_k - 1)| + A \sum_{k=1}^{N} \mu(I_k)$$

y ahora aplicando las desigualdades (4.1) y (4.2) tenemos que

$$\sum_{k=1}^{q} |g(k_i) - g(y_{k-1})| > \left(T - \frac{A\varepsilon}{4}\right) + A\left(\frac{\varepsilon}{4}\right) = T.$$

que contradice la definición de T. Por tanto, E es de medida nula. $\square$

Hemos demostrado que el conjunto F es de medida nula, por tanto las derivadas superior e inferior son iguales. Pero ahora tenemos que ver que esa derivada es finita en casi todo punto de $[a, b]$.

Teorema 4.1.7. Si f es una función creciente en $[a, b]$, entones $\underline{D}f(x)$ es finita c.s. en $[a, b]$.

Demostración. Como $\underline{D}f(x) \geq 0$ para todo $x \in (a, b)$, es suficiente demostrar que $\underline{D}f(x) < \infty$ c.s. en (a, b), es decir, que el conjunto

$$E = \{x : x \in (a, b) \text{ y } \underline{D}f(x) = \infty\}$$

es de medida nula.

Razonemos por reducción al absurdo.
Supongamos que E no es de medida nula. Sean M un número positivo arbitrariamente grande, y $\varepsilon > 0$ la medida exterior asociada de E.

Si $x \in E$, entonces $\underline{D}f(x) > M$ y existen a_x y b_x tal que $a_x < x < b_x$, con $(a_x, b_x) \subseteq (a, b)$ y

$$\frac{f(b_x) - f(a_x)}{b_x - a_x} > M.$$

Entonces $\mathfrak{J} = \{(a_x, b_x) : x \in E\}$ recubre E, y según el lema 4.1.3 contine una subcolección finita disjunta $\{I_1, I_2, \ldots, I_N\}$ tal que $\sum_{k=1}^{N} \mu(I_k) > \varepsilon/3$.
Sea $I_k = (a_k, b_k)$ para cada k. Como f es creciente, tenemos

$$f(b) - f(a) \geq \sum_{k=1}^{N} (f(b_k) - f(a_k)) > \sum_{k=1}^{N} M(b_k - a_k) > M\frac{\varepsilon}{3}.$$

Por lo tanto, $f(b) - f(a) > M\varepsilon/3$ y como $\varepsilon > 0$ y M es arbitrariamente grande, llegamos a un absurdo, pues se contradice que $f(b) - f(a)$ es finito. $\square$

En conclusión, de los teoremas 4.1.6 y 4.1.7 se obtiene:

Teorema 4.1.8 (**Teorema de Lebesgue**). Sea f una función monótona y continua en $[a, b]$. Entonces f es diferenciable c.s. en $[a, b]$.

4.2. Consideraciones finales

Desafortunadamente la derivabilidad en casi todo punto de una función continua y monótona no permite, en todos los casos, reconstruir la función a partir de su derivada, esto es, no se verifica en general la regla de Barrow

$$f(b) - f(a) = \int_a^b f'(t)dt.$$

El teorema fundamental del cálculo tiene dos facetas.

La primera parte, más familiar, trata de que la derivada de la integral indefinida de una función continua es la misma función, es decir:

Sea $f\colon [a,b] \to \mathbb{R}$ una función continua, y sea $F : [a,b] \to \mathbb{R}$ definida por

$$F(x) := \int_a^x f(t)dt.$$

Entonces, F es diferenciable en $[a,b]$ y $F' = f$.

La segunda parte va en la otra dirección, es decir, consiste en reconstruir la función original integrando su derivada. Más precisamente:

$$f(x) - f(a) = \int_a^x f'(t)dt \quad (a \leq x \leq b). \tag{4.4}$$

En la versión elemental de este teorema uno supone que f es diferenciable en todo punto de $[a,b]$ y que f' es una función continua. De este manera, la prueba de (4.4) es sencilla.

Si extendemos (4.4) al ámbito de la integral de Lebesgue, ¿es suficiente suponer que $f' \in \mathcal{L}^1[a,b]$ en vez de que f' sea continua?

Aunque supongamos que f es continua en $[a,b]$, diferenciable en casi todo punto de $[a,b]$, y $f' \in \mathcal{L}^1[a,b]$, no implica que se verifique (4.4). Como prueba de ello, veamos el ejemplo de la función de Cantor φ:

Sea $C_0 = [0,1]$ y para $n \geq 1$ designamos por C_n a la unión de los 2^n intervalos compactos obtenidos al suprimir de los subintervalos C_{n-1} los tercios intermedios. El *conjunto de Cantor* es la intersección

$$C = \bigcap_{n=0}^{\infty} C_n.$$

Es bien conocido que el conjunto de Cantor es compacto, no numerable y de medida nula.

Para cada $n \in \mathbb{N} \cup \{0\}$, la función

$$g_n = \frac{3^n}{2^n} \mathcal{X}_{C_n}$$

es no negativa y simple, por tanto $g_n \in \mathcal{L}^1(\mu)$ y su primitiva

$$f_n(x) = \int_0^x g_n(t)dt$$

es una función monótona y absolutamente continua (más adelante presentamos este concepto), cuya imagen es el intervalo $[0, 1]$, y cuya derivada es nula en todos los puntos de intervalo abierto $\mathbb{R} \setminus C_n$ (ver figura 4.1).

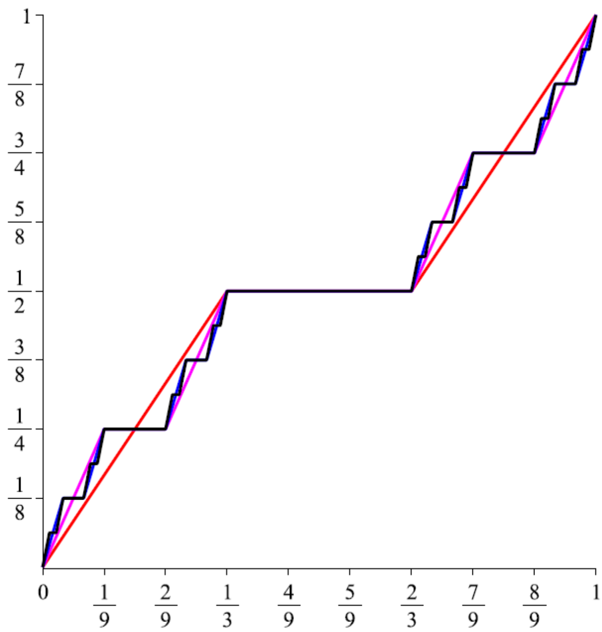

Figura 4.1: Primeras funciones f_n

Si I es uno de los 2^n intervalos cuya unión es C_n, se tiene que

$$\int_I g_n(t)dt = \int_I g_{n+1}(t)dt = \frac{1}{2^n}$$

entonces $f_{n+1}(x) = f_n(x)$ para los $x \notin C_n$, y

$$|f_n(x) - f_{n+1}(x)| \leq \int_I |g_n - g_{n+1}| < \frac{1}{2^{n-1}} \quad x \in C_n.$$

Por lo tanto, $\{f_n\}$ converge uniformemente a la función de Cantor φ, que es continua, monótona, tal que $\varphi(0) = 0$, $\varphi(1) = 1$ y $\varphi'(x) = 0$ para todo $x \notin C$. Como $m(C) = 0$, se tiene que $\varphi' = 0$ en casi todo punto.

Ahora vamos a ver una serie de resultados que son necesarios para que se verifique (4.4).

Definición 4.2.1. Dada una función $f\colon I \to \mathbb{K}$ ($\mathbb{K} = \mathbb{R}$ o $\mathbb{C}$), llamamos variación de f sobre I a

$$V_a^b f = \sup \left\{ \sum_{i=1}^{N} |f(t_i) - f(t_{i-1})| : \quad a = t_0 < t_1 < \ldots < t_N = b \right\}.$$

Si el número $V_a^b f$ es finito, decimos que es de *variación acotada* en I (abreviadamente, VA en I). Si no es de variación acotada, escribimos $V_a^b f = +\infty$.

Observación 4.2.2. Si f es de VA en $[a, b]$ lo es en cualquier subintervalo compacto, por tanto, tiene sentido definir

$$T_f(x) = V_a^x(f), \quad x \in [a, b].$$

Teorema 4.2.3. Supongamos que $f\colon I \to \mathbb{R}$ es VA con $I = [a, b]$, entonces la funciones T_f, $T_f + f$, $T_f - f$ son crecientes y VA en I.

Observación 4.2.4. Si f es de VA, existen funciones g y h acotadas y crecientes, por ejemplo $g = (T_f + f)/2$ y $h = (T_f - f)/2$, tales que $f = g - h$. En consecuencia, toda función real de variación acotada en un intervalo compacto es derivable en casi todo punto del intervalo.

Lo mismo para una función compleja, pues verifica la propiedad VA cuando lo hacen la parte real y la parte imaginaria.

Definición 4.2.5. Se dice que una función compleja definida en un intervalo $I = [a, b]$, es *absolutamente continua* en I (abreviadamente, f es AC en I) si a cada $\varepsilon > 0$ le corresponde un $\delta > 0$ tal que

$$\sum_{i=1}^{n} |f(\beta_i) - f(\alpha_i)| < \varepsilon$$

para todo n y para cualquier colección disjunta de segmentos, $(\alpha_1, \beta_1), \ldots,$ (α_n, β_n) en I cuyas longitudes satisfagan

$$\sum_{i=1}^{n} (\beta_i - \alpha_i) < \delta.$$

Observación 4.2.6.

i) Si una función f es AC, entonces f es continua.

ii) Si una función f es AC, entonces f es de VA.

El reciproco en ambos casos no es cierto.

Con todo lo expuesto, llegamos a nuestro principal objeto:

Teorema 4.2.7. Sea $I = [a, b]$, y sea $f : I \to \mathbb{R}$ continua y monótona. Las siguientes afirmaciones son equivalentes:

a) f es AC en I.

b) f lleva conjuntos de medida nula en conjuntos de medida nula.

c) f es diferenciable es casi todo punto de I, $f' \in \mathcal{L}^1([a, b])$, y

$$f(x) - f(a) = \int_a^x f'(t)dt \quad (a \leq x \leq b).$$

En el siguiente teorema se obtiene (4.4) a partir de un conjunto diferente de hipótesis, pues se supone la diferenciabilidad en todo punto de $[a, b]$.

Teorema 4.2.8. Si $f \colon [a, b] \to \mathbb{R}$ es diferenciable en todo punto de $[a, b]$ y $f' \in \mathcal{L}^1([a, b])$, entonces

$$f(x) - f(a) = \int_a^x f'(t)dt \quad (a \leq x \leq b).$$

Nota. Para ver las demostraciones de estos teoremas, consultar ([28], pag. 165-169) o ([15], pag. 100-107).

Bibliografía

[1] Apostol, T.M. - *Análisis Matemático*, 2ª ed., Editorial Reverté, S.A. 1991.

[2] Austin, D. - *A geometric proof of the Lebesgue differentiation theorem*, Proceedings of the American Mathematical Society, Vol.16 nº2 (1965) pag.220-221.

[3] Billingsley, P. - *Van Der Waerden's Continuous Nowhere Differentiable Function*, The American Mathematical Monthly, Vol.89 nº9 (1982) pag.691.

[4] Blanton, J.D. - *Fondations of Differential Calculus*, (Tradución del original de Euler), Springer 2000.

[5] Botsko, M.W. - *An Elementary Proof of Legesgue's Differentiation Theorem*, The American Mathematical Monthly, Vol.110 nº 9 (2003) pag.834-839.

[6] Botsko, M.W. - *Exactly Which Bounded Darboux Functions are Derivatives?*, The American Mathematical Monthly, Vol.114 nº3 (2007) pag.242-246.

[7] Boyer, C.B. - *Historia de la matemática*, Alianza Universidad Textos 94, 1986.

[8] Bressoud, D.M. - *A radical approach to Real Analysis*, M.A.A. 2007

[9] Cater, F.S. - *On Van Der Waerden's Nowhere Differentiable Function*, The American Mathematical Monthly, Vol.91 nº5 (1984) pag.307-308.

[10] Collette, J.P. - *Historia de las matemáticas II*, Siglo XXI 1985.

[11] Diethelm, K. - *The Analysis of Fractional Differential Equations*, LNM, Springer 2010.

[12] Dunham, W. - *Touring the Calculus Gallery*, The American Mathematical Monthly, Vol.112 nº1 (2005) pag.1-19.

[13] Duoandikoetxea, J. - *200 años de convergencia de las series de Fourier*, La gaceta de la RSME vol 10.3 (2007).

[14] Euler, L. - *Institutiones calculi differentialis.*

[15] Folland, G.B. - *Real Analysis: modern techniques and their applications*, 2nd ed., John Wiley and Sons 1999.

[16] Gray, J. - *The Real and the Complex: A History of Analysis in the 19th century*, Springer 2015.

[17] Hailpern, B. - *Continuous non-diferentiable functions*, Pi Mu Epsilon Journal, Vol.6 nº5 (1976) pag.249-260.

[18] Hewit, E. and Stromberg, K. - *Real and Abstsract Analysis*, Springe 1965.

[19] Hyksova, M. - *Karel Rychlik and Bernard Bolzano* (documento digital).

[20] Kline, M. - *El pensamiento matemático de la Antigüiedad a nuestros días* (3 volumnes), Alianza Universidad 1992.

[21] Kopel, D. - *A New Extension of the Derivative*, The American Mathematical Monthly, Vol.97 nº3 (1990) pag.230-233.

[22] Lagarias, J.C. and Maddock, Z. - *Level Sets of the Takagi Function: Generic Level Sets*, Indiana University Mathematics Journal, Vol.60 nº6 (2011) pag.1857-1884.

[23] Laugwitz, D. - *On the Historical Development of Ifinitesimal Mathematics*, The American Mathematical Monthly, Vol.104 nº5 (1997) pag.447-456.

[24] Laugwitz, D. - *On the Historical Development of Ifinitesimal Mathematics*, The American Mathematical Monthly, Vol.104 nº7 (1997) pag.654-663.

[25] Lebesgue, H. - *Leçons sur l'intégration et la recherche des fonctions primitives*, Gauthier-Villars 1904.

[26] Mitrea, D. - *Distributions, Partial Differential Equations, and Harmonic Analysis*, Springer 2013.

[27] Royden, H.L. and Fitzpatrick, P.M. - *Real Analysis*, Pearson 2010.

[28] Rudin, W. - *Análisis Real y Complejo*, 3ª ed., McGraw-Hill 1988.

[29] Rudin, W. - *Functional Analysis*, 2nd ed., McGraw-Hill 1973.

[30] Sagan, H. - *An Elementary Proof that Schoenberg's Space-Filling Curve Is Nowhere Differentiable*, Mathematics Magazine, Vol.65 nº2 (1992) pag.125-128.

[31] Sánchez Fernández, C. y Valdés Castro, C. - *Las funciones, un paseo por su historia*, 19 ciencia abierta 2007.

[32] Stahl, S. - *Real Analysis a radical approach*, John Wiley 1999.

[33] Wussing, H. - *Lecciones de Historia de la Matemáticas*, Siglo XXI 1998.

Índice alfabético

Cronograma

En la siguiente tabla se relatan cronológicamente algunos de los hechos recogidos en esta memoria.

1669	Newton	*De Analysi per Aequationes Numero Terminorum Infinitas* (no publicado hasta 1711).
1672	Newton	*Methodus Fuxionum et Serierum Infinitarum* (no publicado hasta 1736).
1673	Leibniz	Usó la palabra *función* por primera vez.
1676	Newton	*Tractatus de Quadratura Curvarum* (no publicado hasta 1704).
1684	Leibniz	*Nova methodus pro maximis et minimis, itemque tangetibus, qua nec irrationales quantitates moratur.*
1686	Leibniz	*De geometria recondita et analysi indivisibilum atque infinitarum.*
1695	Leibniz	Regla de la diferenciación multiple del producto.
1696	L'Hôpital	*Analyse des infinitement petits pour l'intelligence des lignes courbes.*
1718	J. Bernoulli	Definción de función. Notación φx para las funciones
1734	Euler	Notación actual de la función.
1748	Euler	*Introductio in analysin infinitorum.* Primera definición de función.
1755	Euler	*Institutiones calculi differentialis.* Segunda definición de función.
1759	Lagrange	*Investigaciones sobre la naturaleza y la propagación del sonido.*
1797	Lagrange	*Théorie des fonctions analytiques.* Definición de función y principios del cálculo diferencial.
1806	Ampère	Intento de probar la existencia general de derivadas.
1821	Cauchy	*Course d'Analyse.* Definición de función.
1822	Fourier	*Théorie des fonctions analytiques.* Defínción de función.

1823	Cauchy	*Résumé des leçons sur le calcul infinitésimal.*
1829	Cauchy	*Leçons sur le calcul différential.*
1829	Dirichelt	Condiciones suficientes de convergencia de series.
1834	Bolzano	*Funktionenlehre.* Ejemplo de función continua y no diferenciable en ningún punto (descubierta sobre 1830, y publicada en 1930 por K. Rychlik)
1837	Dirichlet	Definición de función.
1860	Cellérier	Ejemplo de función continua y no diferenciable en ningún punto (no publicada hasta 1890).
1861	Riemann	Ejemplo de función continua y no diferenciable en ningún intervalo abierto.
1867	Riemann	Definición de función.
1872	Weierstrass	Ejemplo de función continua y no diferenciable en ningún punto (publicada en 1875 por Paul du Bois-Reymond).
1878	Dini	*Fundamenti per la teorica dell funzioni di variabili reali.* Definición de derivadas.
1890	Peano	Ejemplo de curva que rellena todo el espacio.
1891	Hilbert	Ejemplo de curva que rellena todo el espacio.
1903	Takagi	Ejemplo de función continua y no diferenciable en ningún punto.
1904	Lebesgue	*Leçons sur l'intégration et la recherche des fonctions primitives.* Teorema de diferenciación.
1930	Var der Waerden	Ejemplo de función continua y no diferenciable en ningún punto.
1930	Dirac	*Principios de mecánica cuántica.* Defınción y propiedades de la función delta.
1938	Schoenberg	Ejemplo de curva que rellena todo el espacio.