Fundamentos del Cálculo

Rubén Flores Espinoza
Marco Antonio Valencia Arvizu
Guillermo Dávila Rascón
Martín Gildardo García Alvarado

Proyecto FOMIX
CONACYT, Gobierno del Estado
Clave: SON-2004-C02-008

Publicado por Editorial GARABATOS
Febrero, 2008
ISBN: 970-9920-18-5
Tiraje: 1000 ejemplares

Contenido

Presentación

La invención del Cálculo en el último cuarto del siglo XVII representa un hito en la historia de las matemáticas; puede decirse con toda certeza que ahí inician las matemáticas modernas, pues este acontecimiento dio origen al desarrollo de múltiples ramas de las matemáticas, mantuvo prácticamente la exclusividad del trabajo de los matemáticos durante un siglo, y aún los ocupa en sus múltiples ramificaciones y aplicaciones. Antes del Cálculo, las matemáticas sólo servían para describir lo fijo y estático, con él se pudo describir el movimiento y lo dinámico; estableciendo una comparación, podría decirse que antes del Cálculo las matemáticas sólo proporcionaban fotografías de la realidad, y después de él, películas. Además de describir el movimiento, el Cálculo llegó para resolver y unificar los problemas de cálculo de áreas y volúmenes, el trazo de tangentes a curvas y la obtención de valores máximos y mínimos, proporcionando una metodología general para la solución de todos estos problemas; también permitió definir el concepto de continuidad y manejar procesos infinitos. El resultado fue que el Cálculo y sus derivaciones pronto encontraron múltiples aplicaciones y sirvieron para modelar procesos en todos los ámbitos científicos, empezando por la física y las ciencias naturales, hasta llegar a las ciencias sociales. Por todas estas razones, el conocimiento y manejo del Cálculo marca una diferencia cualitativa muy importante en la formación de una persona y en su capacidad para utilizar las matemáticas en otras ciencias y la ingeniería. Podemos afirmar, sin lugar a dudas, que un buen curso de Cálculo cambia la percepción del estudiante universitario.

A escala mundial, la enseñanza y el aprendizaje del Cálculo Diferencial e Integral presenta una severa problemática debido a los altos índices de reprobación y deserción de estudiantes en los cursos básicos de esa materia a nivel de licenciatura. En términos generales, tanto en los países industrializados como en los países en desarrollo se reportan índices de reprobación y deserción superiores al 50%, lo que representa un costo muy elevado en recursos y en oportunidades desaprovechadas.

Siendo el Cálculo una disciplina fundamental en la formación de ingenieros, técnicos y científicos, el problema educativo que presenta nos impulsa a la búsqueda de estrategias y metodologías, tanto disciplinarias como de carácter pedagógico, que permitan asegurar estándares apropiados para poblaciones crecientes de estudiantes.

Los malos resultados que se presentan en el aprovechamiento y desempeño escolar en los cursos de Cálculo se pueden considerar como producto de las dificultades y características de los conceptos y métodos propios de esta rama de las matemáticas y de la insuficiencia de profesores y recursos pedagógicos de apoyo a su enseñanza y aprendizaje. Al masificarse la educación universitaria, la homogenización de los niveles de formación en Cálculo Diferencial e Integral a nivel universitario se presenta como uno de los grandes retos nacionales ante el imperativo de estandarizar la calidad del sistema educativo y facilitar la integración exitosa de los egresados a los mercados de profesionistas que soportan el desarrollo económico y social.

Ante esta situación, un grupo de profesores del Departamento de Matemáticas de la Universidad de Sonora, encabezados por el Doctor Rubén Flores Espinoza, hemos propuesto un conjunto de estrategias para la homogenización y certificación de los cursos de matemáticas a nivel estatal, en el marco de un proyecto apoyado por el Fondo Mixto CONACYT-Gobierno del Estado de Sonora.

Como primera estrategia para la homogenización de los programas de Cálculo en las instituciones de educación superior en Sonora, se aborda el problema del uso del libro obligatorio en los cursos de esta materia. Este problema constituye, en general, una de las más notables deficiencias en la organización y atención de los cursos básicos en el sistema universitario en México. Al no establecerse textos básicos obligatorios que incluyan y desarrollen los contenidos completos de los programas, se deja al estudiante sin una guía para su trabajo personal, a la vez que se propicia la discrecionalidad en el cumplimiento de los programas, se dificulta el establecimiento y evaluación de los estándares de calidad y se vuelve al estudiante más dependiente del profesor. Para contribuir a resolver la problemática anterior, el texto que aquí se presenta desarrolla en forma completa los distintos conceptos, métodos y aplicaciones del Cálculo que son necesarios y suficientes para una formación de calidad en ciencias e ingeniería. Este texto permitirá a todos los estudiantes y profesores de la materia, contar con un referente completo sobre los contenidos y tópicos del cálculo, así como con un amplio conjunto de ejemplos, ejercicios y problemas para el estudio y entrenamiento personal, los cuales se ampliarán en un problemario aparte.

El segundo elemento estratégico para la homogenización de los cursos de Cálculo a nivel superior contemplado en el proyecto antes citado, consiste en la constitución de un Sistema de Entrenamiento y Evaluación en Línea que tiene por propósito el poner a disposición de estudiantes y profesores un sistema electrónico basado en el software MAPLE TA 30 de apoyo a la elaboración, aplicación y evaluación automática de exámenes y pruebas, diseñados de un amplio banco de reactivos y problemas sobre los distintos tópicos de la materia. Este sistema permite la aplicación de exámenes simultáneos a grandes conjuntos de estudiantes de distintas instituciones, lo cual permitirá establecer y conocer los niveles de calidad de la formación en esta materia.

En este texto, intitulado *Fundamentos del Cálculo*, se incluyen todos los tópicos de un programa básico en Cálculo Diferencial e Integral de funciones reales de una variable real. El texto presenta una estructura acorde al desarrollo histórico del Cálculo y orienta sus aplicaciones a la descripción y estudio de las leyes dinámicas que constituyen su verdadero poder y que lo han significado como la invención matemática de mayor impacto en el desarrollo de la ciencia y la tecnología en toda la historia.

Varias particularidades importantes distinguen este libro de la gran cantidad de textos sobre esta materia. En primer lugar, ha sido escrito en un lenguaje llano y familiar, con un buen número de observaciones y notas que buscan motivar y explicar el sentido de los conceptos y resultados y llamar la atención sobre puntos y detalles importantes. También se ha procurado mostrar las características del razonamiento y el discurso matemático presentando los conceptos con todo rigor pero sin caer en sofisticaciones formales que a veces dificultan el aprendizaje, e incluyendo demostraciones completas de todos los resultados. En este sentido, se puede considerar el texto como una iniciación al análisis matemático.

Por otro lado, el texto incluye un buen número de las aplicaciones del Cálculo, principalmente las orientadas a la descripción y estudio de los fenómenos gobernados por leyes dinámicas o de movimiento. Con ese propósito se incluye el estudio de problemas cuyo planteamiento remite a ecuaciones dadas en términos de los conceptos y operaciones del Cálculo y cuya solución requiere el uso y manejo de las reglas de derivación y el conocimiento de los distintos tipos de funciones. En particular, se incluye el tratamiento completo de las ecuaciones diferenciales de segundo orden con coeficientes constantes, por ser éstas las de mayor aplicabilidad en problemas básicos de mecánica y otras disciplinas.

Por la precisión con que se presentan los conceptos, el cuidado puesto en las demostraciones y el énfasis que se hace en los fundamentos del Cálculo, este texto cumple con todo lo necesario para la formación de los estudiantes en el área de ciencias. Al mismo tiempo, por los temas abordados, las técnicas desarrolladas y las aplicaciones presentadas, resulta idóneo para las carreras de ingeniería, pues no solamente incluye las técnicas para la localización de máximos y mínimos, el cálculo de longitudes, áreas y volúmenes, la determinación de presiones y la ubicación de centros de gravedad, sino que también proporciona elementos para comprender mejor las relaciones estáticas y dinámicas entre variables y construir modelos matemáticos que describan cuantitativa y cualitativamente los patrones de comportamiento surgidos de la observación.

El capítulo primero incluye una historia breve del Cálculo a partir de su invención en el siglo XVII y se describen las etapas sucesivas de su desarrollo, hasta llegar a la época actual. Este referente histórico del texto se complementa mediante notas de pie de página con datos alusivos a personajes cuyas aportaciones aparecen en los demás capítulos.

El capítulo segundo está dedicado a una presentación del sistema de los números reales y sus propiedades a partir de su representación como expansiones decimales. Este enfoque permite, desde un principio, poner al estudiante en contacto con nuevos entes matemáticos expresados como conjuntos infinitos de símbolos sobre los cuales se opera y argumenta en preparación a la posterior formalización de los conceptos fundamentales de límite y convergencia de sucesiones. En este capítulo se presenta la propiedad de completez o continuidad, que hace de los números reales el sistema algebraico adecuado para la descripción de las magnitudes que toman valores continuos. Aunque esta presentación es en parte intuitiva, la formalización del uso de esas representaciones que involucran un número infinito de dígitos puede lograrse con los resultados del último capítulo, referente a series.

El capítulo tercero está dedicado al concepto de función, el cual se introduce como una relación entre variables o atributos, para después abstraer su esencia como regla de correspondencia entre conjuntos de números reales. Este enfoque facilita el descubrimiento y construcción de funciones en contextos tanto de la vida real como de origen matemático, en campos como la geometría o el álgebra.

En el capítulo cuarto se introducen los *Fundamentos del Cálculo* a partir de los conceptos de sucesión y convergencia; se incluyen demostraciones completas de los principales resultados básicos del análisis matemático, procurando evitar complicaciones o sofisticaciones formales en la medida de lo posible. El capítulo incluye varios comentarios sobre aspectos finos en la definición y sentido del concepto de continuidad de funciones y su relación con las propiedades de los números.

El capítulo quinto aborda el concepto de derivada de una función en un punto como la razón de cambio puntual o instantánea; se comenta el significado geométrico y dinámico de la derivada y se presentan las reglas de derivación para las diferentes operaciones entre funciones, así como su generalización a derivadas de orden superior.

El capítulo sexto muestra, a través del teorema del valor medio y sus consecuencias, el poder de la derivada en la descripción cualitativa del comportamiento de las funciones, y concluye con la aproximación polinomial que proporciona el teorema de Taylor.

En el capítulo séptimo se caracteriza la función exponencial a partir de las propiedades de su función derivada. Este enfoque muestra cómo aparecen nuevas familias de funciones a partir del estudio de leyes dinámicas y facilita la introducción de la familia de funciones de tipo exponencial y logarítmico, a la vez que nos prepara para el capítulo octavo, donde se aborda el problema del cálculo de antiderivadas o integrales indefinidas.

Por otra parte, en el capítulo noveno se estudia el concepto de integral de Riemann y sus propiedades cuando se aplica a funciones continuas, concepto surgido al aplicar el método exhaustivo o de agotamiento al cálculo del área bajo la gráfica de una función. También se muestra, con el teorema fundamental del Cálculo, cómo el proceso de integración permite "integrar o sumar" las variaciones infinitesimales de una función a lo largo de un intervalo para obtener la variación neta de la función en ese intervalo. En el caso particular del movimiento de una partícula, hace posible calcular el desplazamiento total de la partícula en un intervalo de tiempo, a partir de las velocidades instantáneas mostradas durante ese intervalo.

En el capítulo décimo se incluyen algunas de las aplicaciones más comunes de la integral al cálculo de áreas y volúmenes, lo mismo que al cálculo de presiones de fluidos sobre superficies.

El undécimo capítulo constituye a la vez una introducción a las ecuaciones diferenciales y un ejemplo más elaborado de la aplicación del Cálculo; en él abordamos la solución de ecuaciones diferenciales de segundo orden con coeficientes constantes, cuyas aplicaciones en las ciencias naturales son de primera importancia.

En el duodécimo y último capítulo, se presentan el concepto de serie y los criterios más relevantes para decidir sobre su convergencia, para concluir con la presentación de la familia de las funciones analíticas, o sea las funciones expresables como series de potencias, y la demostración de que constituyen una familia cerrada bajo la operación de derivación, lo que resulta de gran trascendencia en varias áreas de las matemáticas y sus aplicaciones.

Como se señaló antes, este texto se elaboró en el marco del proyecto Homogenización y certificación de los programas de matemáticas de las instituciones de educación superior en Sonora, con registro SON-2004-C02-008, apoyado con los recursos del Fondo Mixto CONACYT-Gobierno del Estado de Sonora. Los autores expresan aquí su agradecimiento al CESUES y a la Universidad de la Sierra por su apoyo institucional a la realización del proyecto, así como a distintas personas que contribuyeron de maneras diversas a la realización de este trabajo, especialmente al Delegado de CONACYT en Sonora, Ing. Francisco Javier Ceballos y a su colaboradora, Lic. Laura Petra Reyes Medina. Agradecemos también a los CC.PP. Ricardo Efrén Espinoza, Angélica Pereida Hoyos y Blanca Irene López Fimbres, por su apoyo en la gestión administrativa al interior de la Universidad de Sonora durante el desarrollo de este proyecto. A Eduardo Tellechea Armenta, Jacobo Núñez Urías, José Luis Díaz Gómez y José Ramón Jiménez Rodríguez, profesores del Departamento de Matemáticas de la Universidad de Sonora, nuestro reconocimiento por sus comentarios y observaciones, y a Manuel Francisco Ocejo Montaño, por su participación en la captura del texto.

Los autores
Hermosillo, Sonora, México
Diciembre del 2007

Capítulo 1

Una historia breve del cálculo

1.1 El siglo XVII: Newton y Leibniz

El Cálculo Diferencial e Integral ha sido reconocido como el instrumento más efectivo para la investigación científica que jamás hayan producido las matemáticas. Concebido para el estudio del cambio, el movimiento y la medición de áreas y volúmenes, el cálculo es la invención que caracteriza la revolución científica del siglo XVII. Su creación se debe al trabajo independiente de dos matemáticos, el inglés Isaac Newton (1642-1727) y el alemán Gottfried Wilhelm Leibniz (1646-1716), quienes publicaron sus investigaciones entre los años de 1680 y 1690. Leibniz en 1684, en la revista *Acta Eruditorum*, y Newton en 1687, en su gran obra *Principia Mathematica Philosophiae Naturalis*.

Sir Isaac Newton
(1642–1727)

Gotfried Whilhelm Leibniz
(1646–1716)

El cálculo se desarrolló a partir de las técnicas infinitesimales utilizadas para resolver dos tipos de problemas: el cálculo de áreas y volúmenes y el cálculo de tangentes a curvas. Arquímedes de Siracusa (287 a.C.-212 a.C), desde tiempos antiguos, había realizado los avances más significativos sobre esos problemas, aplicando el método exhaustivo o de agotamiento para la determinación de áreas y volúmenes y obteniendo importantes resultados sobre el cálculo de tangentes para ciertas curvas particulares. En la primera mitad del siglo XVII, se renovó el interés por esos problemas clásicos y varios matemáticos como Bonaventura Cavalieri (1598-1647), John Wallis (1616-1703), Pierre de Fermat (1601-1665), Gilles de Roberval (1602-1675) e Isaac Barrow (1630-1677), lograron avances que prepararon el camino para la obra de Leibniz y Newton.

A partir de la utilización del método cartesiano[1] para sintetizar los resultados y técnicas desarrollados previamente para el cálculo de áreas y tangentes de curvas, Newton y Leibniz inventaron los métodos y algoritmos que hacen del cálculo una herramienta aplicable a clases generales de problemas. Sus contribuciones en la creación del cálculo difieren en origen, desarrollo e influencia y merecen ser tratadas separadamente.

Newton, hijo de granjeros, nació en Lincolnshire, Inglaterra, en el día de Navidad de 1642 y llegó en 1669 a ocupar, en la Universidad de Cambridge, la Cátedra Lucasiana como profesor de matemáticas. En sus primeras investigaciones introdujo las series infinitas de potencias en una variable x para reformular resultados previos de John Wallis y bajo la influencia de su profesor Isaac Barrow utilizó infinitesimales para mostrar la relación inversa entre el cálculo de áreas y el cálculo de tangentes. Las operaciones de derivación e integración de funciones y su relación recíproca, emergen como un proceso analítico que puede ser aplicado al estudio general de las curvas.

En la presentación de sus ideas, Newton recurre a argumentos basados en el movimiento y la dinámica de los cuerpos. Así, las variables son vistas como algo que cambia o fluye con el tiempo (*fluente*) y a su derivada o razón de cambio con respecto al tiempo la llama su *fluxión*. El problema básico del cálculo es, para Newton, el estudio de las relaciones entre fluentes y sus fluxiones. En 1671, Newton concluye su tratado sobre el método de fluxiones que no es publicado sino hasta 1736, casi diez años después de su muerte, ocurrida en 1727.

En su libro *Principios Matemáticos de la Filosofía Natural*, escrito en 1687, Newton estudia la dinámica de las partículas y establece las bases matemáticas para el cálculo de razones de cambio mediante una teoría geométrica de los límites. Utilizando estos conceptos, desarrolla su teoría de gravitación y reformula las leyes de Kepler para el movimiento de los cuerpos celestes. En su libro, Newton expresa magnitudes y razones de cambio en términos de cantidades geométricas, tanto de tipo finito como infinitesimal, tratando deliberadamente de evitar el uso del lenguaje algebraico. Esta reticencia de Newton a usar los métodos algebraicos, limitó su influencia en el campo de las matemáticas e hizo necesario reformular sus contribuciones en términos del cálculo de Leibniz.

G. W. Leibniz fue el hijo de un profesor de filosofía y nació en la ciudad de Leipzig, Alemania, en 1646. Ingresó a la universidad a la edad de quince años y

[1] Por René Descartes (1596-1650), quien inventó la geometría analítica, independientemente de Pierre de Fermat, y la dió a conocer en 1637 en su obra *La Géométrie*.

obtuvo el doctorado en filosofía a la edad de 21 años. El interés de Leibniz por las matemáticas nació en 1672 durante una visita a París, donde el matemático holandés Christiaan Huygens (1629-1695) lo introdujo al estudio de la teoría de curvas. Después de varios años de estudio bajo la dirección de Huygens, Leibniz investigó las relaciones entre la suma y la diferencia de sucesiones infinitas de números y dedujo varias fórmulas famosas.

Leibniz se interesó en las cuestiones de lógica y de notación para la investigación formal, y su cálculo infinitesimal es el ejemplo supremo, en todas las ciencias y las matemáticas, de un sistema de notación y terminología perfectamente adaptado a su objeto de estudio. En el sentido anterior, Leibniz formalizó, con su notación, las propiedades y reglas fundamentales de los procesos de derivación e integración, haciendo de su aplicación a los más variados problemas, un ejercicio de rutina que un estudiante puede aprender desde sus primeros años. Su primera publicación sobre el cálculo diferencial apareció en 1684, en el *Acta Eruditorum*, bajo el título *Un nuevo método para máximos y mínimos así como para el cálculo de tangentes que incluyen cantidades tanto fraccionales como irracionales y un notable tipo de cálculo para todo esto*. En este artículo, Leibniz introduce la diferencial dx y las reglas básicas del cálculo diferencial $\mathrm{d}(x+y) = \mathrm{d}x + \mathrm{d}y$ y $\mathrm{d}(xy) = x\mathrm{d}y + y\mathrm{d}x$. Dos años después, publica su segundo artículo *Sobre una geometría oculta*, donde introduce y explica el significado del símbolo $\int$ de integración y aplica el poder del cálculo para estudiar curvas trascendentes y deriva una fórmula analítica para la cicloide.

El vigoroso empuje de Leibniz al estudio y desarrollo del nuevo cálculo, el espíritu didáctico de sus escritos y su habilidad para relacionarse con otros investigadores contribuyeron a fortalecer su gran influencia en las matemáticas. Mantuvo una estrecha colaboración con otros estudiosos de su época, incluyendo los hermanos Juan (1667-1748) y Jacobo Bernoulli (1654-1705), quienes se convirtieron en los principales usuarios, investigadores y promotores del nuevo método, Pierre Varignon y Guillaume François Antoine de L'Hospital (1661-1704), este último, autor del primer libro de texto de cálculo diferencial publicado, en 1696. En 1700, Leibniz convence a Federico I de Prusia para crear la Academia de Ciencias de Brandenburgo (después Real Academia de Berlín) de la cual será su presidente vitalicio. En contraste, el aislamiento y la lentitud mostrada por Newton para difundir sus ideas y descubrimientos redujo su presencia en las matemáticas europeas de ese tiempo y aunque un buen número de matemáticos ingleses continuó desarrollando el cálculo, su programa resultó inferior al desarrollado por Leibniz.

1.2 El siglo XVIII: Euler y Lagrange

El siglo XVIII es denominado "El siglo del Análisis Matemático". De 1700 a 1800 se dió la consolidación del cálculo y sus aplicaciones a las ciencias naturales, particularmente a la Mecánica. Con ese desarrollo, vino la especialización y el nacimiento de nuevas ramas de las matemáticas, tales como: la Teoría de Ecuaciones Diferenciales, ordinarias y parciales, el Cálculo de Variaciones, la Teoría de Series y la Geometría Diferencial. Las aplicaciones del análisis incluyen ahora la Teoría de Vibraciones, la Dinámica de Partículas, la Teoría de Cuerpos Rígidos, la Mecánica de Cuerpos Elásticos y Deformables y la Mecánica de Fluidos. A partir de entonces, se distinguen las matemáticas puras de las matemáticas aplicadas.

El desarrollo del análisis matemático en el siglo XVIII está documentado en los trabajos presentados en las Academias de París, Berlín, San Petersburgo y otras, así como en los tratados expositorios publicados en forma independiente. Las figuras dominantes de este periodo son el matemático suizo Leonhard Euler (1707-1783) y el matemático italo-francés Joseph-Louis Lagrange (1736-1813).

Leonhard Euler
(1707–1783)

Joseph Louis Lagrange
(1736-1813)

Euler nació en Basilea, Suiza, y completó se educación universitaria a la edad de quince años. Es considerado el matemático más prolífico de todos los tiempos, sus obras abarcan casi setenta y cinco volúmenes y contienen contribuciones fundamentales a casi todas las ramas de las matemáticas y sus aplicaciones. La carrera profesional de Euler se desarrolló en la Real Academia de San Petersburgo, Rusia (1727-1741 y 1766-1783) y en la Academia de Berlín (1741-1766).

La obra de Euler en dos volúmenes intitulada *Introducción al análisis infinitesimal*, publicada en 1748, da lugar al nacimiento del llamado Análisis Matemático como rama de esta disciplina, análoga al Álgebra y la Geometría. El Análisis Matemático es construido a partir del concepto fundamental de función y de los procesos infinitos desarrollados para la representación y estudio de las funciones. En esa gran obra, por primera vez se presenta el estudio sistemático de las funciones exponenciales y de las funciones trigonométricas como funciones numéricas, así como el estudio de las funciones transcendentes elementales mediante sus desarrollos en series infinitas. A esa primera obra de Euler, siguieron dos obras más, en 1755 y 1768, sobre el cálculo diferencial e integral, respectivamente, que constituyen la fuente original de los actuales libros y textos sobre el cálculo y las ecuaciones diferenciales.

El enfoque analítico de Euler recibió un gran impulso de la otra gran figura del siglo XVIII, el matemático Joseph Louis Lagrange, quien a la muerte de Euler, en

1783, lo reemplazó como el matemático líder de Europa. Aplicando métodos puramente analíticos, Lagrange extendió y perfeccionó el Cálculo de Variaciones y a partir de sus aplicaciones a la mecánica, sentó los fundamentos de la llamada Mecánica Analítica. En 1788 se publicó su famoso tratado *Mecánica Analítica* en donde, aplicando las ideas del cálculo de variaciones, presenta los fundamentos analíticos de la mecánica. En el prefacio de su tratado, Lagrange declara que en su exposición sólo recurre a argumentos analíticos, sin dibujos, figuras o razonamientos mecánicos. Es decir, Lagrange hace de la mecánica una rama del análisis matemático.

Para fines del siglo XVIII había preocupación en Europa por los fundamentos del cálculo y del análisis. Los argumentos basados en la teoría de fluxiones de Newton y en la idea de infinitamente pequeño mostraban serias inconsistencias que fueron puntualmente señaladas por el obispo anglicano irlandés George Berkeley (1685-1753) en 1734. Afrontando la situación anterior, Lagrange publicó en 1797 su obra *Teoría de funciones analíticas* en la cual pretende presentar un desarrollo completo del cálculo de funciones sin recurrir a los conceptos de límite o de cantidad infinitesimal. El enfoque de Lagrange se basa en considerar que las funciones son representables como series de potencias, cuyos coeficientes definen las derivadas de los distintos órdenes. En este tratado, Lagrange sienta las bases para la aproximación de funciones por polinomios y da la forma del residuo denominada Residuo de Lagrange.

1.3 El siglo XIX: Cauchy, Riemann y Weierstrass

Al finalizar el siglo XVIII, los matemáticos habían ya detectado distintas limitaciones e incongruencias en las bases sobre las que se había desarrollado hasta entonces el cálculo diferencial e integral. Los trabajos de Jean D'Alembert (1717-1783) sobre la cuerda vibrante y de Joseph Fourier (1768-1830) sobre la Teoría Analítica del Calor, de 1807, remitían a la necesidad de considerar clases más amplias de funciones que las meramente representables como series de potencias a la manera de Lagrange. En ese momento, emerge la necesidad de aclarar las propiedades de continuidad y de integrabilidad de las funciones, así como las condiciones de convergencia para series de funciones.

El concepto de continuidad de una función aparece explícitamente definido, por primera vez, en el trabajo del matemático checo Bernhard Bolzano (1781-1848), pero es el matemático francés Augustin Louis Cauchy (1789-1857) quien desarrolla en su generalidad la teoría de funciones continuas y formula los conceptos y procesos fundamentales del cálculo para ese tipo de funciones en los términos en que actualmente se presentan. En sus tres grandes obras *Curso de Análisis* (1821), *Resumen de Lecciones sobre el Cálculo Infinitesimal* (1822) y *Lecciones sobre el Cálculo Diferencial* (1829), Cauchy hace una exposición rigurosa del cálculo basándose en el concepto fundamental de límite de una función. En particular, define la derivada de una función como el límite de cocientes de los incrementos de las variables y demuestra sus distintas propiedades; presenta el *teorema del valor medio* y sus aplicaciones a la aproximación de funciones por polinomios; establece rigurosamente los criterios para la existencia de máximos y mínimos de funciones; define la integral definida de una función continua en un intervalo mediante el límite de sumas asociadas a particiones de ese intervalo; y formula, con todo rigor, el llamado *teorema fundamental del cálculo,* estableciendo la relación inversa que existe entre los procesos de derivación e integración de funciones.

El siguiente avance en la evolución histórica del cálculo, se debe a Bernhard F. Riemann (1826-1866), quien introdujo las funciones esencialmente discontinuas en el desarrollo del cálculo, extendiendo el proceso de integración a este tipo de funciones, con importantes consecuencias sobre los conceptos primarios de longitud, área y volumen de conjuntos. A pesar de los grandes esfuerzos por dotar al análisis matemático

Augustin Louis Cauchy (1789–1857)

Bernhard Riemann (1826–1866)

Karl Weierstrass (1815-1897)

de bases sólidas, a mediados del siglo XIX varias suposiciones sobre la estructura de los números reales utilizadas en la prueba de las propiedades importantes de las funciones continuas, y otras suposiciones, como por ejemplo la existencia de derivada en casi todos los puntos para toda función continua, son señaladas críticamente y desmentidas por contundentes contraejemplos dados por matemáticos como el mismo Bolzano y el alemán Karl Weierstrass (1815-1897) quienes, por ejemplo, logran exhibir funciones continuas que no poseen derivada en punto alguno. Ese tipo de situaciones, obliga a los matemáticos al estudio y construcción del sistema de los números reales a partir del sistema de los números naturales. El año de 1872 registra la publicación, casi simultánea, de construcciones de los números reales debidas a Georg Cantor (1845-1918), Richard Dedekind (1831-1916) y Edward Heine (1821-1881), basadas en los conceptos de límite y sucesiones, previamente desarrollados.

La construcción de los números reales es el paso decisivo hacia la aritmetización del análisis matemático, que permite al mismo Karl Weierstrass dar la definición de límite en términos de las meras estructuras algebraicas y de orden de los números reales, y con ello los conceptos y procesos propios del cálculo quedan debidamente justificados y adquieren la presentación definitiva con que hoy son expuestos en los libros de texto y demás trabajos matemáticos.

1.4 El siglo XX: Lebesgue y Robinson

Finalmente, es necesario decir que el siglo XX registra dos nuevos avances en el desarrollo del análisis: la *integral de Lebesgue*, debida al francés Henri Lebesgue (1875-1941), y el *Análisis no-Estándar*, debido básicamente a Abraham Robinson (1918-1974).

El concepto de integral desarrollado por Cauchy se aplica a funciones continuas, pero aunque éste fue generalizado después, por Riemann, a funciones con cierto tipo de discontinuidades, el espacio de las funciones integrables no es cerrado bajo los procesos de convergencia y de límite de sucesiones de funciones, lo que restringe su aplicablidad a otras ramas de la matemática.

Basado en trabajos del italiano Giuseppe Peano (1858-1932) y del francés Camille Jordan (1838-1922), Henri Lebesgue logró dar, en 1920, una definición de conjunto medible y de medida que generalizan, en la recta, las nociones de intervalo y de longitud de un intervalo, respectivamente. Con base en estos nuevos conceptos, Lebesgue introdujo una nueva clase de funciones llamadas funciones medibles, para las cuales adquiere sentido una nueva definición de integral, definida como el límite de integrales de funciones que toman valores constantes en conjuntos medibles. En este sentido, la integral de Lebesgue es una generalización de la integral de Riemann, que se obtiene como el límite de integrales de funciones que toman valores constantes sobre intervalos.

Henri Lebesgue
(1875–1941)

Abraham Robinson
(1918–1974)

La clase de las funciones integrables en el sentido de Lebesgue tiene propiedades inmejorables para los propósitos del análisis matemático en tanto que límites de sucesiones y series convergentes de funciones de este tipo resultan ser también funciones integrables. La nueva teoría de la medida e integración sienta las bases para el desarrollo de la Teoría Matemática de la Probabilidad y la Estadística, que tanta importancia tienen en la ciencia actual.

El otro desarrollo importante del análisis del siglo XX fué presentado en 1960 por Abraham Robinson, seguido de su libro *Análisis no Estándar*, en el que se retoma el problema de la aritmetización del análisis a partir del concepto de número y de magnitud infinitamente pequeña. A partir de construcciones basadas en la teoría de conjuntos, Robinson introdujo el concepto de número hiperreal con lo que logra dar un significado preciso a los "infinitamente pequeños" que Euler usaba en sus argumentos y demostraciones. Con ello, los procesos de límite y de convergencia del análisis son sustituidos por operaciones y procedimientos meramente algebraicos en la clase de los números hiperreales.

Aunque la nueva formulación de Robinson da lugar a un cálculo más simple, la construcción de los números hiperreales es muy elaborada y los libros en los que se expone el cálculo no estándar no han logrado tener éxito en los niveles matemáticos medio y básico.

Capítulo 2

Los números reales

El sistema de los números reales es la estructura algebraica adecuada al propósito del cálculo diferencial e integral. Son precisamente los atributos y las relaciones expresables en términos de este tipo de números, los objetos de estudio de esa rama de las matemáticas. Las propiedades especiales del sistema de los números reales permiten definir los conceptos fundamentales para la descripción y estudio del cambio y el movimiento.

La presentación que aquí se hace del sistema de los números reales, se basa en el concepto de expansión decimal, utilizado en la vida diaria para representar y operar con números y magnitudes. Así, cada número real se identifica con una sucesión infinita de dígitos separados por un punto decimal y el conjunto de tales objetos resulta ser una extensión del conjunto de los números racionales, los cuales quedan identificados con las llamadas expansiones periódicas. Las operaciones de suma y multiplicación, y la relación de orden entre los números racionales se extienden de manera natural, preservando sus propiedades algebraicas y de orden, al conjunto de los números reales.

La propiedad que distingue al sistema de los números reales del sistema de los números racionales es la propiedad de continuidad o completez. Esta propiedad, de carácter geométrico o topológico, es la que permite dar un sentido preciso a los conceptos fundamentales de límite y continuidad, sobre los cuales se desarrolla el cálculo diferencial e integral.

2.1 Expansiones decimales

Desde la escuela primaria, hemos aprendido a representar y a manejar las medidas y las cantidades mediante números expresados en el sistema decimal, es decir, mediante la utilización de sucesiones de los dígitos $0, 1, 2, 3, 4, 5, 6, 7, 8, 9$ que forman lo que llamamos la expansión decimal del número de que se trate.

Las expansiones decimales a cuyo uso nos acostumbramos en los primeros niveles de educación, sólo constan de un número finito de dígitos separados por un punto decimal. Por ejemplo, la expansión

$$A = 123.7584$$

representa al número

$$A = 1 \cdot 10^2 + 2 \cdot 10^1 + 3 \cdot 10^0 + 7 \cdot 10^{-1} + 5 \cdot 10^{-2} + 8 \cdot 10^{-3} + 4 \cdot 10^{-4}.$$

Para ese tipo de expansiones, se desarrollan algoritmos para realizar las operaciones básicas de la aritmética y posteriormente, ya en la escuela secundaria, se incluyen expansiones negativas, sobre las cuales se extienden las operaciones aritméticas valiéndose de la regla de los signos

$$(-A) \cdot (-B) = +(A \cdot B)$$
$$(-A) \cdot (+B) = -(A \cdot B),$$

para cada par de expansiones decimales A, B.

Otro tipo de expansiones que también nos son familiares, son las que aparecen al construir la representación decimal de los números racionales m/n, donde m y n son enteros, con $n \neq 0$, y que resultan ser *expansiones infinitas y periódicas,* pues tienen la propiedad de presentar un bloque de dígitos que se repite indefinidamente a la derecha a partir de un cierto lugar de la expansión. Por ejemplo,

$$\frac{1}{3} = 0.3333 \cdots 33 \cdots$$

o

$$\frac{29}{7} = 4.142857142857 \cdots 142857 \cdots$$

Ejemplo 2.1 Para ilustrar cómo se genera la expansión decimal periódica de un número racional, construyamos paso a paso, como ejemplo, la expansión decimal del número racional

$$D = \frac{4}{7}$$

aplicando el algoritmo de la división que aprendimos en la escuela primaria. Al realizar esa operación, vamos obteniendo en cada etapa un dígito del cociente y un residuo r mayor o igual a cero y menor que el divisor 7, de tal manera que al efectuar a lo más 7 veces el procedimiento de división, forzosamente tendrá que repetirse, por primera vez, alguno de los residuos obtenidos en las etapas anteriores, con la consiguiente repetición de los dígitos correspondientes en la expansión decimal del cociente que se está construyendo. Así, en el caso de 4/7, al aplicar el algoritmo de la división, tal como se muestra en la figura, se obtienen, en el primer paso, 0 unidades en el cociente y residuo 4; en el segundo paso se obtienen 5 décimos en el cociente y residuo 5; en el tercer paso se obtienen 7 centésimos en el cociente y residuo 1, y así, sucesivamente, hasta llegar al séptimo paso, en el que se obtienen 8 millonésimos en el cociente y residuo 4, tal como lo tuvimos en el primer paso.

Luego, a partir del octavo paso, se repite la sucesión de residuos, dando lugar a una repetición del bloque de dígitos 571428, obteniéndose así la expansión decimal que representa al número 4/7:

$$\frac{4}{7} = 0.571428571428\cdots571428\cdots$$

$$
\begin{array}{rl}
 & 0.571428\widehat{571428}\cdots \longleftarrow \text{Bloque que se repite}\\
 & 7\,\overline{)\,4}\\
\text{Primer residuo} \longrightarrow & 40\\
 & \quad 50\\
 & \qquad 10\\
 & \qquad\quad 30\\
 & \qquad\qquad 20\\
 & \qquad\qquad\quad 60\\
\text{Se repite} \longrightarrow & \qquad\qquad\qquad 40\\
 & \qquad\qquad\qquad\quad \ddots
\end{array}
$$
◁

En este punto, lo notable no sólo es que la expansión decimal de todo número racional sea una expansión periódica, sino que más aún, cada expansión decimal periódica es la expansión decimal de algún número racional, estableciéndose así una equivalencia entre ambos conjuntos de objetos. Enseguida mostramos, con un ejemplo, cómo se encuentra el número racional que corresponde a una expansión periódica dada.

Ejemplo 2.2 Si queremos encontrar el número racional que corresponde a la expansión decimal periódica

$$D = -2.83434\cdots3434\cdots,$$

procedemos a multiplicarla por 10 y luego por 1000 y obtenemos las siguientes expresiones, que tienen los mismos dígitos a la derecha del punto decimal

$$\begin{aligned} 10 \cdot D &= \quad -28.3434\cdots34\cdots \\ 1000 \cdot D &= -2834.3434\cdots34\cdots \end{aligned}$$

Al restar la primera expansión de la segunda, obtenemos

$$990 \cdot D = -2806,$$

por lo que

$$D = -\frac{2806}{990}.$$
◁

Notación. Escribiremos las expansiones decimales periódicas en forma simplificada omitiendo los dígitos después de la primera aparición del bloque de dígitos que se repite y marcando con una línea superior dicho bloque. Por ejemplo, la expresión

$$3.23\overline{45}$$

representa la expansión decimal periódica

$$3.234545\cdots 45\cdots \qquad \triangleleft$$

A los números que no se pueden expresar como un cociente de números enteros se les llama *números irracionales* y por lo que mostramos anteriormente, sus expansiones decimales no pueden ser periódicas. El conjunto de los números irracionales se denota por $\mathbb{I}$. Un ejemplo de número irracional es la raíz cuadrada de 2. Esta afirmación se justifica en el ejemplo siguiente.

Ejemplo 2.3 Para probar que $\sqrt{2}$ no puede expresarse como cociente de dos números naturales, argumentaremos por contradicción, es decir, supondremos que es cierto lo contrario, que existen números primos relativos a, b (es decir, sin divisores comunes) tales que

$$\sqrt{2} = \frac{a}{b}.$$

Elevando al cuadrado, tenemos que

$$2 = \frac{a^2}{b^2},$$

o, equivalentemente,

$$2b^2 = a^2. \tag{2.1}$$

Pero (2.1) implica que el número a^2 es un número par, por lo que a debe ser un número par (ya que el cuadrado de un número par es un número par y el cuadrado de un número impar es impar). Por lo tanto, a se puede escribir en la forma

$$a = 2c, \tag{2.2}$$

para algún número entero c. Sustituyendo ahora (2.2) en (2.1), tenemos

$$2b^2 = 4c^2,$$

y, consecuentemente,

$$b^2 = 2c^2,$$

es decir, b^2 es un número par y por lo tanto b tiene que ser a su vez un número par y, por consiguiente, tanto a como b son números pares, lo cual es falso pues supusimos desde el principio que a y b no tenían divisores en común. Luego, la suposición es falsa y por lo tanto $\sqrt{2}$ no es un número racional. $\triangleleft$

Es relativamente sencillo generar números irracionales, como se muestra en el ejemplo siguiente.

Ejemplo 2.4 La expansiones decimales

$$A = 23.010010001\cdots 1\overbrace{00\cdots 0}^{i-\text{veces}}1\;\overbrace{00\cdots 0}^{(i+1)-\text{veces}}\;1\cdots$$

$$B = -2.454554555\cdots 4\overbrace{55\cdots 5}^{i-\text{veces}}4\;\overbrace{55\cdots 5}^{(i+1)-\text{veces}}\;4\cdots$$

corresponden a números irracionales. ◁

Tomando en cuenta la discusión anterior, tenemos la definición siguiente.

Definición 2.1 *Una **expansión decimal** A, es una expresión de la forma*

$$A = \pm a_k a_{k-1} \cdots a_1 a_0 . b_1 b_2 \cdots b_{r-1} b_r \cdots$$

*donde $a_k, a_{k-1}, \ldots, a_0$ y $b_1, b_2, \ldots, b_{r-1}, b_r, \cdots$ son algunos de los dígitos $\{0, 1, 2, \ldots, 8, 9\}$. Al punto después del dígito a_0 se le llama **punto decimal** de la expansión. Si la expansión decimal va precedida del signo + se dice que la expansión decimal es **positiva** y si va precedida del signo - se le llama **expansión decimal negativa**.*

Nota Importante:
Cada expansión decimal se extiende a la derecha del punto decimal, mientras que a la izquierda del punto decimal sólo consta de un número finito de dígitos.

2.2 El Sistema de los Números Reales

Se define *el conjunto* $\mathbb{R}$ *de los números reales* como el *conjunto de las expansiones decimales*, sobre el cual se establece el siguiente criterio de igualdad: Dos expansiones decimales A y B son iguales (representan el mismo número real) si se presenta alguna de las dos situaciones siguientes:

1. A y B constan de los mismos dígitos y estos ocupan el mismo orden, o

2. A y B constan de los mismos dígitos hasta un cierto lugar r y enseguida la expansión de uno de ellos continúa en la forma

$$\pm a_k a_{k-1} \cdots a_0 . b_0 b_1 \cdots b_r b_{r+1} \overline{9}$$

con $b_{r+1} \neq 9$, mientras que la expansión del otro es de la forma

$$\pm a_k a_{k-1} \cdots a_0 . b_0 b_1 \cdots b_r (b_{r+1} + 1) \overline{0}$$

Ejemplo 2.5 Las expansiones $1.34\overline{9}$ y $1.35\overline{0}$ son, por definición, iguales y representan el mismo número real. ◁

Nota Importante:
En general, en la definición de las operaciones y propiedades de los números reales siempre evitaremos escribir expansiones decimales con bloques repetidos de nueves.

2.2.1 Operaciones con los números reales

Las operaciones con los números reales, son las usuales de suma y multiplicación que empezamos a manejar desde la escuela primaria. De hecho, en la escuela secundaria aprendemos los métodos o algoritmos para sumar y multiplicar expansiones decimales finitas tanto positivas como negativas y sabemos cómo construir la expansión decimal correspondiente a la suma o al producto, a partir de la suma y producto de los dígitos y la posición que éstos ocupan en las expansiones decimales que se pretende operar.

Antes de introducir las operaciones entre expansiones decimales infinitas, para cada expansión $A = \pm a_k a_{k-1} \cdots a_0.b_1 \cdots b_r b_{r+1} \cdots$ definimos su *expansión truncada de orden* r, con $r \geqslant 0$, como la expansión decimal periódica

$$A_r = \pm a_k a_{k-1} \cdots a_0.b_1 \cdots b_r \overline{0}$$

que consta de los mismos dígitos que la expansión de A hasta el lugar r después del punto decimal, y todos los dígitos siguientes a la derecha son cero. La expansión truncada de orden r se puede escribir también en términos de sumas de potencias del número 10 en la forma usual

$$\begin{aligned} A_r &= \pm a_k a_{k-1} \cdots a_0.b_1 \cdots b_r \overline{0} \\ &= \pm\left(a_k 10^k + a_{k-1} 10^{k-1} + \cdots + a_1 10 + a_0 + \frac{b_1}{10} + \frac{b_2}{10^2} + \cdots + \frac{b_r}{10^r}\right). \end{aligned}$$

Nota Importante:
Un número real está totalmente determinado si se conocen sus expansiones truncadas de cualquier orden y viceversa. Observe que la expansión decimal truncada de orden cero es el número entero a la izquierda del punto decimal de la expansión decimal inicial.

Para sumar dos expansiones decimales $A = \pm a_k a_{k-1} \cdots a_0.b_1 \cdots b_r b_{r+1} \cdots$ y $B = \pm c_j c_{j-1} \cdots c_0.d_1 \cdots d_r d_{r+1} \cdots$ y formar la expansión decimal correspondiente a la suma $A+B$, se procede como sigue: Para cada orden $r = 0, 1, 2, \cdots$ la expansión truncada de orden r de la suma $A + B$ se define como la expansión truncada de orden r de la suma de las expansiones truncadas de orden $r + 1$ de A y B.

Por ejemplo, si queremos sumar las expansiones decimales $A = 2.\overline{95}$ y $B = 1.2020020002 \cdots 200 \cdots 02 \cdots$, la expansión suma $A + B$ es aquélla que tiene por expansiones decimales truncadas de los distintos órdenes, las siguientes:

$$\begin{aligned}
(A+B)_0 &= 4.\overline{0}\\
(A+B)_1 &= 4.1\overline{0}\\
(A+B)_2 &= 4.16\overline{0}\\
(A+B)_3 &= 4.165\overline{0}\\
(A+B)_4 &= 4.1659\overline{0}\\
&\vdots
\end{aligned}$$

que se forman sumando, de acuerdo a la definición, las expansiones truncadas correspondientes de los números iniciales.

Análogamente, para multiplicar las dos expansiones decimales A y B y formar la expansión decimal correspondiente al producto $A \cdot B$, se procede como sigue:

1. Se determina cuántos dígitos a la izquierda del punto decimal tiene cada uno de los factores. Digamos que A tiene m dígitos y B tiene n dígitos a la izquierda del punto decimal.

2. Se multiplica la expansión truncada de orden $n+1$ de A con la expansión truncada de orden $m+1$ de B y la expansión truncada de orden cero del producto de estas será la expansión truncada de orden cero de la expansión decimal de $A \cdot B$,

3. Para determinar la expansión truncada de orden $r > 0$ de $A \cdot B$, se multiplica las expansión truncada de orden $n+r+1$ de A por la expansión truncada de orden $m+r+1$ de B y la expansión truncada de orden r de ese producto de expansiones truncadas se toma como la expansión truncada de orden r del producto $A \cdot B$.

Ejemplo 2.6 Para multiplicar las expansiones decimales

$$\begin{aligned}
A &= 12.\overline{34},\\
B &= -253.2020020002\cdots,
\end{aligned}$$

las expansiones truncadas de $A \cdot B$ se determinan de acuerdo a la definición anterior, en la forma siguiente:

$$\begin{aligned}
(A\cdot B)_0 &= -3125.\overline{0},\\
(A\cdot B)_1 &= -3125.3\overline{0},\\
(A\cdot B)_2 &= -3125.38\overline{0},\\
&\vdots
\end{aligned}$$

etcétera. ◁

A partir de las definiciones de suma y multiplicación de los números reales, dadas en términos de sus expansiones, enlistamos sus propiedades principales.

Sean A, B, C números reales y $+ : \mathbb{R} \times \mathbb{R} \to \mathbb{R}$, $\cdot : \mathbb{R} \times \mathbb{R} \to \mathbb{R}$ las operaciones de suma y multiplicación. Entonces se cumple:

(S1) $A + B = B + A$ (Conmutatividad de la suma)

(S2) $A + (B + C) = (A + B) + C$ (Asociatividad de la suma)

(S3) $0 + A = A$ (Existencia de neutro bajo la suma)

(S4) $A + (-A) = 0$ (Existencia de inversos aditivos bajo la suma)

(M1) $A \cdot B = B \cdot A$ (Conmutatividad de la multiplicación)

(M2) $A \cdot (B \cdot C) = (A \cdot B) \cdot C$ (Asociatividad de la multiplicación)

(M3) $1 \cdot A = A$ (Existencia de neutro bajo la multiplicación)

(M4) Si $A \neq 0$ existe A^{-1} tal que $A \cdot A^{-1} = 1$ (Existencia de inversos multiplicativos)

(M5) $A \cdot (B + C) = A \cdot B + A \cdot C$ (Distributividad de la multiplicación respecto a la suma)

Cuando un conjunto S posee dos operaciones (suma y multiplicación) que tienen las propiedades (S1)–(S4) y (M1)–(M5), arriba mencionadas, se dice que tiene *estructura algebraica de campo* y así, se habla del *campo de los números reales*.

2.2.2 El orden de los números reales

El conjunto de los números reales se descompone en tres subconjuntos mutuamente ajenos:

(*i*) los reales positivos, $\mathbb{R}^+$, formados de las expansiones decimales positivas,

(*ii*) los reales negativos, $\mathbb{R}^-$, formado por las expansiones decimales negativas, y

(*iii*) el conjunto $\{0\}$ formado por la expansión cero.

La descomposición anterior da lugar a la llamada *ley de tricotomía para el orden*, que estipula que cada número real A tiene una y sólo una de las siguientes propiedades: ó A es positivo, ó A es negativo, ó A es el número cero.

El conjunto $\mathbb{R}^+$ de los reales positivos tiene la propiedad de que tanto la suma como la multiplicación de cualesquiera dos de sus elementos, es nuevamente un real positivo.

Definición 2.2 *Se dice que un número real A* **es mayor que** *el número real B (o equivalentemente, que B es* **menor que** *A) si $A - B$ es un real positivo. Para denotar que A es mayor que B escribiremos $A > B$.*

Nótese que, según la definición 2.2, todo número real negativo es menor que cero.

En términos de sus expansiones decimales respectivas, una expansión positiva A es mayor que otra expansión positiva B si al recorrer sus dígitos de izquierda a derecha existe un lugar k, tal que ambas constan de los mismos dígitos hasta ese lugar y el dígito $k + 1$ de A es mayor que el dígito $k + 1$ de B.

De las propiedades de los números reales positivos, se deduce la validez de la siguiente proposición:

Proposición 2.1

a) Si $A > B$ y $C > D$, entonces $A + C > B + D$.

b) Si $A > B$ y $C > 0$ entonces $A \cdot C > B \cdot C$.

c) Si $A > B$ y $C < 0$ entonces $A \cdot C < B \cdot C$.

DEMOSTRACIÓN. Para probar a), tenemos que si $A > B$ y $C > D$, por definición, esto significa que $A - B \in \mathbb{R}^+$ y $C - D \in \mathbb{R}^+$. Luego, de las propiedades de los números positivos, concluimos que $(A - B) + (C - D) \in \mathbb{R}^+$. Por lo tanto $(A + C) - (B + D) \in \mathbb{R}^+$, lo cual significa que

$$(A + C) > (B + D).$$

Para probar b), basta notar que si $A > B$ y $C > 0$ se tiene que $A - B \in \mathbb{R}^+$ y de las propiedades de los números positivos se sigue que $(A - B) \cdot C > 0$, es decir, $A \cdot C - B \cdot C > 0$ o equivalentemente, $A \cdot C > B \cdot C$.

Finalmente, para demostrar la validez de c), observamos que si $A - B \in \mathbb{R}^+$ y $C \in \mathbb{R}^-$ se tiene que $-C \in \mathbb{R}^+$ y $(A - B) \cdot (-C) \in \mathbb{R}^+$, por lo que $A \cdot C < B \cdot C$. ■

Otra propiedad importante que posee la relación de orden entre los números reales, es la propiedad de *arquimedianidad*[1] que se enuncia en los términos siguientes:

Propiedad de Arquimedianidad. Si A y B son números reales tales que $0 < A < B$, entonces existe un número natural n tal que $n \cdot A > B$.

En el conjunto $\mathbb{R}$ de los números reales, tanto el subconjunto de los números racionales $\mathbb{Q}$ como el subconjunto de los números irracionales $\mathbb{I}$, se distribuyen de

[1]Así llamada en honor de Arquímedes de Siracusa.

manera *densa* en $\mathbb{R}$. Esto quiere decir, que dados dos números reales distintos $A < B$, siempre existen números racionales y números irracionales que son mayores que A y menores que B. Esto significa que tanto los racionales como los irracionales se encuentran tan cerca como se quiera de cualquier número real.

Ejemplo 2.7 Si $A = 2.34526789\cdots$ y $B = 2.34612387\cdots$, se tiene que $A < B$ y la expansión $C = 2.346\overline{0}$ es un número racional mayor que A y menor que B, mientras que el número $D = 2.346001000100001\cdots$ que no muestra ningún bloque de dígitos que se repita, es un número irracional mayor que A y menor que B. ◁

Haciendo uso de los conceptos de orden entre los números reales, se introduce la definición de *intervalo abierto con extremo izquierdo* A *y extremo derecho* B como el subconjunto de números reales dado por

$$(A, B) = \{x \in \mathbb{R} \text{ tales que } A < x < B\}$$

y la definición de *intervalo cerrado con extremo izquierdo* A *y extremo derecho* B como el subconjunto de números reales dado por

$$[A, B] = \{x \in \mathbb{R} \text{ tales que } A \leqslant x \leqslant B\}\,.$$

Análogamente, se definen los intervalos

$$\begin{aligned}
(A, \infty) &= \{x \in \mathbb{R} \text{ tales que } x > A\}\,,\\
(-\infty, A) &= \{x \in \mathbb{R} \text{ tales que } x < A\}\,,\\
[A, \infty) &= \{x \in \mathbb{R} \text{ tales que } x \geqslant A\}\,,\\
(-\infty, A] &= \{x \in \mathbb{R} \text{ tales que } x \leqslant A\}\,.
\end{aligned}$$

2.2.3 Valor absoluto de un número real

Introduciremos ahora el concepto de *métrica* o de *distancia* entre los números reales. Para ello, presentamos el concepto de *valor absoluto* de un número real mediante la siguiente definición.

> ***Definición 2.3*** *Si A es un número real, el* **valor absoluto de** *A se define como el número $|A|$ tal que:*
>
> $$|A| = \begin{cases} A & \textit{si } A \geqslant 0,\\ -A & \textit{si } A < 0. \end{cases}$$
>
> *Note que el valor absoluto de un número es siempre un número no-negativo.*

Ejemplo 2.8 $|2.\overline{31}| = 2.\overline{31}$, $\quad |-12.54230\cdots| = 12.54230\cdots$ ◁

El valor absoluto de un número real tiene las propiedades que se enuncian en la proposición siguiente.

Proposición 2.2 *Los enunciados siguientes son verdaderos:*

1. $|A \cdot B| = |A||B|$ *para cualesquiera* A, $B \in \mathbb{R}$.
2. *Si* $|B| < A$, *entonces* $-A < B$ *y* $B < A$; *recíprocamente, si* $-A < B$ *y* $B < A$ *y* $A > 0$, *entonces* $|B| < A$;
3. *Si* $|B| > A$, *entonces* $A < B$ *ó* $B < -A$; *recíprocamente, si* $A < B$ *o* $B < -A$, *entonces* $|B| > A$.
4. $|A + B| \leqslant |A| + |B|$, *para todos* $A, B \in \mathbb{R}$. (Desigualdad del triángulo)

DEMOSTRACIÓNLa demostración de los puntos 1, 2 y 3 se sigue directamente de la definición de valor absoluto. La demostración del punto 4 (la desigualdad del triángulo) es como sigue. Primero, tenemos que para cualesquiera $A, B \in \mathbb{R}$

$$-|A| \leqslant A \leqslant |A|, \qquad -|B| \leqslant B \leqslant |B|.$$

Enseguida, sumando término a término, tenemos

$$-(|A| + |B|) \leqslant A + B \leqslant |A| + |B|,$$

lo cual significa, en virtud de 2, que

$$|A + B| \leqslant |A| + |B|,$$

como se afirma en 4. ∎

La *distancia*, $d(A, B)$, entre dos números reales A y B se define como el valor absoluto de la diferencia de los dos números, esto es

$$d(A, B) = |A - B|.$$

De las propiedades del valor absoluto se deducen las siguientes propiedades para la distancia entre los números reales:
Para cualesquiera números reales A, B y C se cumple:

(a) $d(A, B) \geqslant 0$ y $d(A, B) = 0$ si y sólo si $A = B$ (Positividad definida)

(b) $d(A, B) = d(B, A)$ (Simetría)

(c) $d(A, B) \leqslant d(A, C) + d(C, B)$ (Desigualdad del triángulo)

Enseguida presentamos dos ejemplos sobre la aplicación de las propiedades del valor absoluto de un número.

Ejemplo 2.9 Represente, como unión de intervalos, el conjunto $\mathcal{A}$ de los números reales x tales que
$$|3x+2| \leqslant 8 \quad \text{y} \quad |-x+3| > 1.$$

Solución. Consideremos primero el conjunto de los números reales x tales que
$$|3x+2| \leqslant 8.$$

De acuerdo al punto 2 de la proposición 2.2, esto es equivalente a
$$3x+2 \leqslant 8 \quad \text{y} \quad 3x+2 \geqslant -8,$$

lo cual implica que
$$3x \leqslant 6 \quad \text{y} \quad 3x \geqslant -10,$$

por lo que
$$x \leqslant 2 \quad \text{y} \quad x \geqslant -\frac{10}{3}.$$

Resumiendo, hemos probado que si x es tal que $|3x+2| \leqslant 8$, entonces
$$x \in \left[-\frac{10}{3}, 2\right].$$

Por otro lado, si un número real x satisface la desigualdad
$$|-x+3| > 1,$$

entonces debe satisfacer las desigualdades
$$-x+3 > 1 \qquad \text{o} \qquad -x+3 < -1,$$

lo cual implica que
$$-x > -2 \qquad \text{o} \qquad -x < -4,$$

es decir,
$$x \in (-\infty, 2) \cup (4, \infty).$$

Finalmente, los números reales x que satisfacen ambas desigualdades serán aquéllos en la intersección del intervalo $[-10/3, 2]$ con el conjunto $(-\infty, 2) \cup (4, \infty)$. Es decir,
$$\mathcal{A} = \left[-\frac{10}{3}, 2\right] \cap ((-\infty, 2) \cup (4, \infty)) = \left[-\frac{10}{3}, 2\right). \qquad \triangleleft$$

Ejemplo 2.10 Escriba como unión de intervalos el conjunto
$$\mathcal{A} = \{x \in \mathbb{R} \text{ tales que } 1 < |-2x+3| \leqslant 2\}.$$

Solución. Si $x \in \mathcal{A}$, entonces

$$|-2x+3| \leqslant 2 \ \text{ y } \ |-2x+3| > 1.$$

Si $|-2x+3| \leqslant 2$, entonces

$$-2x+3 \leqslant 2 \ \text{ y } \ -2x+3 \geqslant -2.$$

Sumando -3 a ambos lados en cada desigualdad y luego dividiendo por -2, se obtiene

$$x \geqslant \frac{1}{2} \ \text{ y } \ x \leqslant \frac{5}{2},$$

o, lo que es lo mismo,

$$x \in \left[\frac{1}{2}, \frac{5}{2}\right].$$

Por otro lado, si $|-2x+3| > 1$, se tiene que

$$-2x+3 > 1 \ \text{ o } \ -2x+3 < -1,$$

lo cual implica que

$$x < 1 \qquad \text{o} \qquad x > 2,$$

y por lo tanto,

$$x \in (-\infty, 1) \cup (2, \infty).$$

De lo anterior deducimos que

$$\mathcal{A} = \left[\frac{1}{2}, \frac{5}{2}\right] \bigcap \big((-\infty, 1) \cup (2, \infty)\big) = \left[\frac{1}{2}, 1\right) \bigcup \left(2, \frac{5}{2}\right]. \qquad \triangleleft$$

2.3 Completez de los números reales

La propiedad que hace del sistema de los números reales un sistema numérico apropiado para representar variables que toman un continuo de valores, es la llamada *propiedad de completez* (o de continuidad). Intuitivamente, esto quiere decir que el conjunto $\mathbb{R}$ es un conjunto sin cortes, cuyos elementos están dispuestos según un orden "continuo". En esta sección explicamos estos conceptos.

Definición 2.4 *Un conjunto $\mathcal{A}$ no vacío de números reales se dice* ***acotado superiormente*** *por $M \in \mathbb{R}$ si*

$$x \leqslant M \ \textit{ para todo } x \in \mathcal{A}.$$

Al número M se le llama una ***cota superior*** *para $\mathcal{A}$.*
Análogamente un conjunto $\mathcal{A}$ de reales se dice ***acotado inferiormente*** *por $N \in \mathbb{R}$ si*

$$N \leqslant x \ \ \textit{ para todo } x \in A.$$

Al número N se le llama una ***cota inferior*** *para el conjunto A.*

Propiedad de completez (o continuidad) de los números reales: *Para cada subconjunto $\mathcal{A}$ no vacío de números reales acotado superiormente, existe un número real S tal que:*

1. *S es cota superior de $\mathcal{A}$.*
2. *S es menor o igual que cualquier otra cota superior de $\mathcal{A}$.*

*Al número S se le denomina **mínima cota superior** o **supremum de** $\mathcal{A}$ y se denota*

$$S = \sup \mathcal{A}.$$

Análogamente, para cada subconjunto $\mathcal{A}$ no vacío de números reales acotado inferiormente, existe un número real I tal que:

1. *I es cota inferior de $\mathcal{A}$.*
2. *I es mayor o igual que cualquier otra cota inferior de $\mathcal{A}$.*

*Al número real I se le denomina la **máxima cota inferior** o **infimum de** $\mathcal{A}$ y se denota*

$$I = \inf \mathcal{A}.$$

Demostración. Daremos la construcción, paso a paso, del supremum para cualquier subconjunto de reales positivos acotados superiormente. Esto es suficiente para probar la propiedad de completez, pues para conjuntos arbitrarios acotados superiormente, la construcción se puede hacer de la misma forma con modificaciones mínimas.

Sea $\mathcal{A}$ un conjunto no vacío de expansiones decimales positivas acotadas superiormente. Mostraremos la existencia del supremum de $\mathcal{A}$ mediante la siguiente construcción.

Primer paso: Siendo $\mathcal{A}$ acotado superiormente, la parte entera de los elementos de $\mathcal{A}$ es un conjunto de números enteros acotados superiormente y de los cuales podemos sin ambigüedad, determinar el entero mayor. Ese número será la parte entera del supremum S de $\mathcal{A}$.

Segundo paso: Enseguida determinamos el subconjunto $\mathcal{A}_1$ de $\mathcal{A}$ formado de los elementos de $\mathcal{A}$ cuya parte entera es igual al máximo valor que toma la parte entera de los elementos de $\mathcal{A}$. (Note que todos los elementos de $\mathcal{A}_1$ tienen la misma parte entera). Ahora nos fijamos en el primer dígito después del punto decimal de los elementos de $\mathcal{A}_1$ y tomamos el dígito mayor (esto último se consigue sin problema pues sólo se tiene que escoger entre los diez dígitos posibles). Ese dígito mayor será el primer dígito del supremum S después del punto decimal.

Tercer paso: Enseguida determinamos el subconjunto $\mathcal{A}_2$ de $\mathcal{A}_1$ formado de los elementos de $\mathcal{A}_1$ cuyo primer dígito después del punto decimal es igual al máximo valor que toma ese primer dígito en los elementos de $\mathcal{A}_1$. (Note que los elementos de $\mathcal{A}_2$ todos tienen la misma parte entera y el mismo primer dígito después del punto decimal). Ahora nos fijamos en el segundo dígito a la derecha del punto decimal de los elementos de $\mathcal{A}_2$ y tomamos el dígito mayor, ésto último se consigue sin problema pues sólo se tiene que escoger entre los diez dígitos posibles. Ese dígito mayor será el segundo dígito de S después del punto decimal.

Repitiendo el procedimiento anterior, se va construyendo de manera sucesiva la expansión decimal correspondiente al supremum de $\mathcal{A}$, lo que prueba que el conjunto de los números reales posee la propiedad de continuidad. ∎

La propiedad de completez se puede establecer también sin hacer referencia explícita al supremum o al infimum, como lo hacemos a continuación.

Propiedad de completez de los números reales (segunda versión): *Si $I_n = [a_n, b_n]$, para $n = 1, 2, ...,$ es una sucesión de intervalos cerrados de $\mathbb{R}$ tales que*

$$I_{n+1} \subset I_n \text{ para cada } n = 1, 2...,$$

entonces la intersección de todos ellos es un conjunto no vacío; es decir,

$$\bigcap_{n=1}^{\infty} I_n \neq \phi.$$

En particular, si para cada $n = 1, 2, \ldots$ se tiene que

$$b_n - a_n = \left(\frac{1}{10}\right)^n,$$

entonces existe un único número real P que pertenece a cada uno de los intervalos, es decir, tal que $P = \cap_{n=1}^{\infty} I_n$.

NOTA IMPORTANTE:
Si $\mathcal{A}$ es un conjunto de expansiones decimales periódicas acotado superiormente, su supremum no necesariamente será una expansión periódica, como se ilustra en el ejemplo siguiente.

Ejemplo 2.11 El conjunto de expansiones decimales periódicas de la forma

$$\begin{aligned}
&0.1\overline{0}\\
&0.1011\overline{0}\\
&0.101101110\overline{0}\\
&0.1011011101111\overline{0}\\
&\ldots
\end{aligned}$$

constituye un conjunto acotado superiormente por el número 1 y su supremum es la expansión decimal no periódica

$$0.10110111011110\cdots 10\overbrace{111\cdots 1}^{i-\text{veces}}0\ \overbrace{111\cdots 1}^{(i+1)-\text{veces}}\ 0\cdots$$

Este ejemplo nos muestra que el conjunto de los números reales, a diferencia del campo de los racionales, es un conjunto continuo o "sin cortes". $\triangleleft$

2.4 La Recta Real

El sistema de los números reales que hemos presentado, tiene como modelo geométrico el conjunto de puntos de una recta ideal L, sobre la cual se han identificado un punto arbitrario A con el número real cero y otro punto arbitrario B, a la derecha del anterior, con el número real 1.

Antes de establecer la correspondencia entre las expansiones decimales y los puntos de la recta L, recordaremos el procedimiento para dividir con regla y compás un segmento de recta $\overline{AB}$ en diez subsegmentos iguales. El procedimiento es como sigue: Se traza en el plano una recta M distinta de L y que la corte en el punto A. Con el compás se toma una distancia arbitraria d y con esa misma abertura y punto inicial A se trazan consecutivamente sobre M los diez puntos $C_1, C_2, ..., C_{10}$, como se muestra en la figura 2.1.

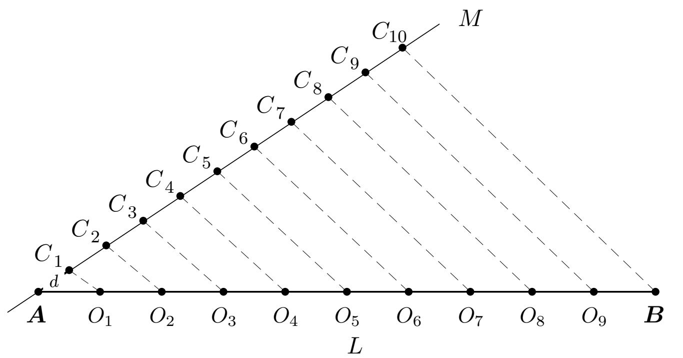


Figura 2.1 División de un segmento en 10 partes iguales

Enseguida se traza la recta que pasa por B y C_{10} y las rectas paralelas a esa recta que pasan por $C_1, C_2, ...C_9$ y se determinan sus puntos de corte $O_1, O_2, ...O_9$ con el segmento AB. Como los triángulos AO_iC_i con $i = 1, 2, ...9$ son triángulos semejantes (debido a que tienen los mismos ángulos), entonces los puntos $O_1, O_2, ...O_9$ dividen a AB en diez segmentos de igual longitud.

Procedamos ahora a identificar los puntos de la recta ideal L con el conjunto de las expansiones decimales. Mostraremos primero cómo a cada uno de los puntos de L, a la derecha del punto cero, se le asocia una expansión decimal positiva y, de manera recíproca, cómo a cada expansión decimal positiva le corresponde un único punto de la recta. La identificación de la parte de la recta a la izquierda del punto cero con las expansiones decimales negativas se hace de manera análoga.

Tomando el compás con abertura igual a la distancia entre los puntos asociados al cero y al uno, y marcando con esa abertura sucesivamente hacia la derecha, se fijan los puntos correspondientes a los números naturales y con ellos, los intervalos de la forma $[a, a+1]$ con a un número natural. Enseguida, realizando el procedimiento con regla y compás dado anteriormente, se subdivide cada intervalo $[a, a+1]$, en diez subintervalos de longitud $1/10$ de la forma $[a + b_1/10, a + (b_1 + 1)/10]$ con $b_1 = 0, 1, 2, \ldots, 9$. Luego, se toma a ese intervalo para obtener los subintervalos de longitud $1/10^2$, de la misma forma: $[a+b_1/10+b_2/10^2, a+b_1/10+(b_2+1)/10^2]$ con $b_1, b_2 = 0, 1, 2, \ldots, 9$. Repitiendo este procedimiento k veces, se determinan todos los intervalos con extremos racionales de la forma

$$[a.b_1 b_2 \cdots b_k \ , \ a.b_1 b_2 \cdots (b_k + 1)]$$

con a entero y $b_1, b_2, \cdots, b_k = 0, 1, 2, \ldots, 9$. Note que al ir tomando k los valores $0, 1, 2, 3, \ldots$, se obtendrán todos los intervalos que tienen por extremos los puntos asociados a las expansiones truncadas positivas.

Habiéndose construido los intervalos anteriores sobre la recta ideal, se procede a asociar a cada expansión decimal $a.b_1 b_2 \cdots b_k \cdots$ el punto P de la recta ideal que se encuentra en la intersección de los intervalos cerrados anidados

$$[a.b_1, a.(b_1+1)], [a.b_1 b_2, a.b_1(b_2+1)], [a.b_1 b_2 b_3, a.b_1 b_2(b_3+1)],$$

etcétera, formados con los dígitos correspondientes de la expansión decimal inicial. Aquí suponemos que la recta ideal tiene la propiedad de que la intersección de intervalos anidados de longitud cada vez más pequeña e igual a una potencia de $1/10$, tienen por intersección un punto. Esta última, es una manera de suponer que la recta ideal no tiene agujeros, es decir, que se forma de un continuo de puntos. A dicha recta también se le llama la *recta real,* o *recta numérica.*

Recíprocamente, a cada punto P de la recta ideal se le asocia la expansión construida en la forma siguiente: Primero se determina a qué intervalo de longitud 1 y de la forma $[a, a+1]$, con a entero, pertenece el punto P. Esto determina la parte entera de la expansión decimal correspondiente a P. Enseguida, se divide el intervalo anterior en diez subintervalos de longitud $1/10$ de la forma $[a+b_1/10, a+(b_1+1)10]$ para $b_1 = 1, 2, \ldots, 9$ y se determina el valor del dígito b_1 tal que $P \in [a + b_1/10, a + (b_1+1)/10]$. Este dígito b_1 será el primer dígito a la izquierda de la expansión decimal asociada a P. A continuación se repite el proceso anterior, subdividiendo en cada paso al intervalo anterior en diez subintervalos y agregando a la derecha de la expansión en formación, el dígito correspondiente al extremo izquierdo del subintervalo al cual pertenece P. Este proceso nos permite conocer cada uno de los dígitos de la expansión decimal asociada a P y, por lo tanto, el número real asociado a P.

Ejercicios y problemas del capítulo

1. a) Sean las expansiones $A = 2.3458$, $B = -3.2568$ y $C = -1.35802$. Calcule las expansiones de $A + B$, $A \cdot B$, $B \cdot C$ y $B - C$.

   b) Escriba la expansión decimal correspondiente a los números racionales siguientes y determine, en cada caso, el bloque de dígitos que se repite:

$$i)\ \frac{23}{7}, \qquad ii)\ -\frac{57}{4}, \qquad iii)\ \frac{2491}{990}.$$

2. Dé los números racionales cuya expansión decimal es cada una de las siguientes: $A = 2.34\overline{210}$, $B = 37.2856\overline{0}$, $C = -13.3\overline{45}$.

3. Demuestre que $\sqrt{5}$ y $\sqrt{3}$ no son números racionales.

4. Encuentre la expansión truncada de orden 5 de la suma $A + B$ de los reales

$$A = 1.28288288828888\cdots$$
$$B = 12.\overline{253}$$

5. Encuentre un número irracional entre los números $0.0001\overline{0}$ y $0.001\overline{0}$.

6. Demuestre que si A es un número real distinto de cero, entonces $A^2 > 0$.

7. Sea $A = 2.13113\overline{0}$, encuentre un racional y un irracional cuya distancia a A sea menor que $1/10^4$.

8. Pruebe que el conjunto de números de la forma $a + b\sqrt{2}$ con $a, b \in \mathbb{Q}$ forman un campo que contiene a $\sqrt{2}$.

9. Demuestre que entre cada par de números racionales distintos siempre hay un número irracional y que entre cada par de irracionales distintos existe siempre un número racional.

10. Escriba como unión de intervalos ajenos los conjuntos

    (a) $\mathcal{A} = \{x \in \mathbb{R}$ tales que $|2x - 4| \leqslant 6$ y $|x - 3| > 1\}$
    (b) $\mathcal{B} = \{x \in \mathbb{R}$ tales que $|x - 3| \leqslant 2|x|\}$
    (c) $\mathcal{C} = \{x \in \mathbb{R}$ tales que $|2x - 3| > 2$ y $|x - 5| < 1\}$
    (d) $\mathcal{D} = \{x \in \mathbb{R}$ tales que $|x - 2| < 1$ y $|2x - 1| > 2\}$
    (e) $\mathcal{E} = \{x \in \mathbb{R}$ tales que $1 < |3x + 2| \leqslant 5\}$
    (f) $\mathcal{F} = \{x \in \mathbb{R}$ tales que $x > 0$ y $|x - 3| > 2|x|\}$.

11. Demuestre que $|A| - |B| \leqslant |A - B|$ para todo $A, B \in \mathbb{R}$.

12. Demuestre las siguientes afirmaciones para conjuntos acotados superiormente:

    (a) Si $\mathcal{A} \subset \mathcal{B} \subset \mathbb{R}$ entonces $\sup \mathcal{A} \leqslant \sup \mathcal{B}$,

    (b) $\sup(\mathcal{A} + \mathcal{B}) = \sup \mathcal{A} + \sup \mathcal{B}$,

    (c) $\sup \mathcal{A} = -\inf(-(A))$.

13. Conteste "falso" o "verdadero".

    ____ Si A es un número real distinto de cero, existe otro número real tal que $A \cdot B = 2$.

    ____ Existe un número real A tal que $A^2 + A + 1 = 0$.

    ____ Si A y B son números reales con $A < B$, entonces $A < \dfrac{A+B}{2} < B$.

Capítulo 3

Variables y funciones

El cambio y el movimiento se manifiestan siempre en el marco de una relación entre dos o más objetos o variables que toman entre sí distintas configuraciones o valores relativos. En particular, el tipo de cambio o movimiento que describe y estudia el cálculo diferencial es el que se presenta en relaciones entre variables que toman un continuo de valores numéricos reales. A tales variables se les denomina variables reales y a las relaciones entre ellas, funciones reales.

En este capítulo se introducen los conceptos fundamentales de variable real y función real de variable real y se definen las operaciones básicas entre éstas. Se introduce, además, el concepto de gráfica de una función, el cual permite identificar funciones con curvas en el plano cartesiano, estableciéndose el vínculo para la interpretación geométrica de los conceptos y procesos del cálculo.

Finalmente, se presentan algunas de las familias de funciones elementales más importantes en las aplicaciones.

3.1 El concepto de variable y el de función

Una *variable* es una propiedad o un atributo que puede tomar uno o varios valores dados por los elementos de un conjunto. Por ejemplo, la variable "estado civil de una persona", toma valores en el conjunto cuyos elementos son: soltero, casado, viudo, divorciado, etc. Cuando el conjunto de los valores posibles de una variable es un subconjunto o un intervalo de números reales, se le denomina *variable real.* Como ejemplos de variables reales tenemos, "la magnitud o medida del lado de los cuadrados ", "la medida de la estatura de las personas", " la medida de la velocidad de un automóvil", etc. Al conjunto de los valores posibles que puede tomar una variable x se le llama el *dominio de la variable* x, por ejemplo, "la medida de la diagonal del cuadrado" es una variable real cuyo dominio es el conjunto de los números reales positivos, mientras que el dominio de la variable " la temperatura de un cuerpo", es el conjunto de los números reales, la cual toma valores negativos para temperaturas bajo cero.

A las variables reales, las denotaremos con letras, como $x, y, z, t, \ldots$

Por ejemplo:

1. t : "valor de la medida del tiempo",
2. x : "valor de la magnitud del lado del cuadrado",
3. y : "valor del área del cuadrado",
4. z : "número real mayor que cero y menor que uno".

NOTA IMPORTANTE:
Así como el conjunto de valores "soltero, casado, viudo, divorciado, etc." define la variable "estado civil de la persona", también en general podemos identificar a cada variable real con el conjunto de sus valores.

En los distintos campos del conocimiento y de la experiencia, encontramos pares o conjuntos de variables relacionadas entre sí, en el sentido de que el valor de una o de varias de ellas depende del valor o los valores de las otras.

Ejemplo 3.1 Consideremos las variables

y : "valor del área del cuadrado,"
x : "valor de la longitud del lado del cuadrado".

Estas dos variables numéricas están relacionadas en el sentido de que a cada valor de una de ellas corresponde un valor para la otra. Así, si la variable x toma el valor numérico a, el valor correspondiente de la variable y es a^2. En este caso, la regla f que establece la relación o correspondencia entre las variables y y x se simboliza escribiendo en forma compacta

$$y = f(x) = x^2.$$

En esta notación, el símbolo $f(x)$ representa el valor de y cuando el valor de la otra variable es x. Así, si el valor de la variable x es 30, el valor correspondiente de la variable y será 900. Esto se escribe poniendo

$$y = f(30) = 900$$

y se interpreta "cuando el valor de la variable x es 30, el valor de la variable y es igual a 900". $\triangleleft$

Cuando la relación entre dos variables x y y es tal que a cada valor de la primera, corresponde un único valor de la segunda, se dice que esta última está en función de la primera. A la variable x se le llama *variable independiente* y a la variable y se le llama *variable dependiente*. Es este tipo de relaciones entre variables reales las que son el objeto de estudio del cálculo y a las que dedicaremos este capítulo.

Definición 3.1 *Una* ***función real de variable real*** [1] *es una regla de correspondencia f entre dos variables reales x y y, tal que a cada valor de la primera corresponde, según esa regla, un* ***único valor*** *de la segunda. Al dominio de la variable independiente x se le llama el* ***dominio de la función*** *y al dominio de la variable dependiente y se le llama el* ***contradominio de la función.***

Ejemplo 3.2 Sea x la variable con dominio los números reales distintos de cero y y la variable "número real". La regla que asocia a cada valor x el número $y = f(x) = 1/x$, define una función cuyo dominio es el conjunto de los números reales distintos de cero y cuyo contradominio es el conjunto de los números reales. $\triangleleft$

NOTA IMPORTANTE:
En la definición de función, el punto clave es que la regla de correspondencia f que define la función, sea tal que a cada valor de la variable independiente le corresponda uno y sólo un valor de la variable dependiente. Por ejemplo, si x es la variable con dominio los números reales, la regla $y = f(x) = \sqrt{x}$, en general, no define de manera unívoca un valor para la variable y, ya que si el valor de x es un número negativo no es posible extraer raíz cuadrada y por lo tanto no está definido el valor que le corresponde a la variable dependiente, lo mismo sucede si el valor de la variable independiente x es un número positivo, ya que existen dos valores posibles para $\sqrt{x}$, uno el negativo del otro, y la regla no especifica cuál de los dos valores tomar. Como se observa, en ambos casos el valor de la variable independiente x no determina de manera unívoca el valor de la variable y. Una manera de definir correctamente la función "raíz cuadrada de x" es considerando como variable independiente la variable x = "número real mayor o igual a cero" y como regla de correspondencia la expresión

$$y = f(x) = +\sqrt{x}$$

que nos dice que para cada valor de la variable x se tome la raíz cuadrada positiva de ese valor.

Para denotar una función, es necesario señalar claramente sus distintos elementos, como son:

(*i*) la variable independiente x y su dominio,

(*ii*) la variable dependiente y y su dominio y

(*iii*) la regla de correspondencia f que permite calcular o conocer el valor de la variable dependiente conociendo el valor de la variable independiente.

[1] La definición moderna de *función* se debe al matemático alemán de ascendencia belga J. P. G. Lejeune Dirichlet (1805-1859), quien la formuló en 1837.

Al escribir la regla de correspondencia, se denota por $y = f(x)$ el valor de la variable dependiente y correspondiente al valor x de la variable independiente y se hace explícita la regla f para calcular y a partir del valor x. De aquí en adelante, adoptando la notación establecida por el matemático J. L. Lagrange, denotaremos cada función real escribiendo

$$f : A \to B$$
$$y = f(x),$$

donde A es el dominio de la variable independiente o dominio de la función, B el contradominio de la función o dominio de la variable dependiente y f denota la regla de correspondencia entre los valores de las variables. Cuando la situación anterior se presente, también diremos que "la variable y está en función de la variable x."

Una función f se puede también visualizar de manera intuitiva como un mecanismo que recibe como entrada un valor de x, el cual es procesado por el mecanismo, dando lugar a un valor de salida $f(x)$ que es el valor de la variable dependiente que corresponde a x, como se ilustra en la figura 3.1.

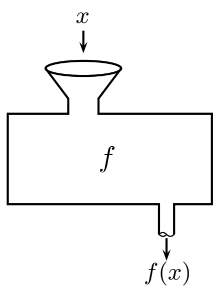


Figura 3.1 Función como proceso

Ejemplo 3.3 La función f con variable dependiente y :"valor del área del cuadrado" y con variable independiente x :"longitud del lado del cuadrado", se escribe simbólicamente

$$f : [0, \infty) \to [0, \infty)$$
$$y = f(x) = x^2.$$

El símbolo $f : [0, \infty) \to [0, \infty)$ significa que la función asocia a cada valor del dominio $[0, \infty)$ de la variable x, un valor para y en el intervalo $[0, \infty)$, y en el segundo renglón se hace explícita la regla de correspondencia entre los valores de las variables, señalando que el valor de y correspondiente al valor x de la variable independiente, es igual a x^2. $\triangleleft$

Dada una función $f : A \to B$ y $x \in A$, al número $f(x)$ se le llama *imagen de x bajo f* y al conjunto

$$f(A) = \{f(a) \text{ con } \ a \in A\} \subset B$$

se le llama la *imagen de la función* f y se denota $\operatorname{Im} f$.

Algunas veces, la regla de correspondencia entre los valores de la variable independiente x y los de la variable dependiente se expresa con distintas fórmulas, dependiendo del valor de la variable independiente.

Ejemplo 3.4 La función

$$f : [1,4] \to \mathbb{R},$$
$$f(x) = \begin{cases} 2x+1 & \text{si } x \in [1,2], \\ x^2 & \text{si } x \in (2,3), \\ 2 & \text{si } x \in [3,4], \end{cases}$$

está definida con tres expresiones distintas según dónde toma su valor la variable x. Note que, en el ejemplo 3.4, la imagen de f es el conjunto $\{2\} \cup [3,9)$. $\triangleleft$

Una variable real puede depender de dos o más variables reales.

Ejemplo 3.5 Dado un rectángulo, consideremos las variables
p : "valor del área del rectángulo",
u : "valor de lo largo del rectángulo", y
v : "valor de lo ancho del rectángulo".

Entonces podemos escribir a la variable p en función de las variables u y v de tal manera que se tiene una función de dos variables que se expresa en la forma

$$p : [0,\infty) \times [0,\infty) \to [0,\infty),$$
$$p(u,v) = u \cdot v.$$ $\triangleleft$

Cuando el valor de una variable depende de los valores de dos o más variables, se dice que se tiene una *función de varias variables reales.*

NOTA IMPORTANTE:
Si bien al tener una función $y = f(x)$ a cada valor de la variable independiente x le corresponde un único valor $y = f(x)$ de la variable dependiente y, es posible que a valores distintos de x, la regla de correspondencia les asocie el mismo valor de la variable y. Por ejemplo, la función real

$$f : \mathbb{R} \to \mathbb{R}$$
$$y = f(x) = x^2$$

le hace corresponder a cada par de valores a y $-a$ de la variable x, el mismo valor a^2 de la variable y.

Las funciones que a valores distintos de la variable independiente x, hacen corresponder valores distintos de la variable dependiente y, juegan en la teoría un papel importante, en tanto permiten definir la regla de correspondencia inversa entre la variable y y la variable x. Fijaremos este último concepto mediante la siguiente definición.

Definición 3.2 *Una función real de variable real $f : A \to B$ con $y = f(x)$, se dice* ***inyectiva*** *(o* ***uno a uno****) si a valores distintos de la variable independiente les hace corresponder valores distintos para la variable dependiente. En lenguaje simbólico, esta definición se escribe: $f : A \to B$, es inyectiva o uno a uno si*

$$a, b \in A \text{ con } a \neq b \text{ implica } f(a) \neq f(b).$$

Note que si $y = f(x)$ es inyectiva, entonces para cada valor c de la variable y perteneciente a la imagen de la función $y = f(x)$, existe un único valor a de la variable independiente x tal que $f(a) = c$. Al único valor a de la variable x que corresponde al valor c de la variable y se le denota $a = f^{-1}(c)$. Esto determina una nueva función f^{-1} cuya variable independiente es la variable y y su variable dependiente, x. A esta nueva función se le llama *la función inversa de la función* $y = f(x)$ y se denota

$$f^{-1} : \operatorname{Im} f \subset B \to A$$
$$x = f^{-1}(y)$$

Ejemplo 3.6 La función $f : \mathbb{R} \to \mathbb{R}$ definida mediante la regla

$$y = f(x) = 3x + 5$$

es una función inyectiva, ya que si $a \neq b$, se tiene

$$f(a) - f(b) = 3a + 5 - 3b - 5 = 3(a - b) \neq 0.$$

En este caso, la función inversa $x = f^{-1}(y)$ toma la forma

$$x = f^{-1}(y) = \frac{y - 5}{3}$$

y tiene por dominio $\mathbb{R}$. ◁

Ejemplo 3.7 La función

$$f : (0, 1) \to \mathbb{R},$$
$$f(x) = x^2 + 2$$

es una función uno a uno y tiene por función inversa

$$f^{-1} : (2, 3) \to (0, 1)$$
$$f^{-1}(y) = +\sqrt{y - 2}.$$
◁

Ejemplo 3.8 La función

$$g : \mathbb{R} \to \mathbb{R}$$

$$g(x) = x^3 + 1$$

es una función uno a uno, como se verifica a continuación. Si x_1 y x_2 son elementos distintos del dominio de la variable x, es decir $x_1 \neq x_2$, entonces los valores de la función $y = g(x)$ en esos puntos serán respectivamente

$$g(x_1) = x_1^3 + 1 \text{ y } g(x_2) = x_2^3 + 1.$$

Si ahora calculamos su diferencia

$$g(x_1) - g(x_2) = x_1^3 - x_2^3 = (x_1 - x_2)(x_1^2 + x_1x_2 + x_2^2)$$

observamos que por elección $(x_1 - x_2) \neq 0$ y por otro lado

$$(x_1^2 + x_1x_2 + x_2^2) \neq 0 \text{ si } x_1 \neq x_2.$$

Luego, tendremos que

$$g(x_1) - g(x_2) \neq 0,$$

lo que significa que $g(x_1) \neq g(x_2)$, y por lo tanto la función $y = g(x)$ es uno a uno. En este caso la función inversa

$$x = g^{-1}(y)$$

queda definida mediante la regla

$$x = g^{-1}(y) = \sqrt[3]{y - 1}$$

y su dominio de definición es el conjunto

$$\text{Dominio de } g^{-1} = \text{Imagen de } g = \mathbb{R}. \qquad \triangleleft$$

Entre las principales funciones inyectivas o uno a uno, se encuentran las *funciones monótonas*, mismas que se clasifican en:

(*i*) *funciones crecientes*, definidas como aquellas funciones reales $f : A \to \mathbb{R}$ que preservan el orden de los valores de la variable independiente, es decir, para $x_1, x_2 \in A$ con $x_1 > x_2$ se tiene

$$f(x_1) > f(x_2), \quad \text{y}$$

(*ii*) *funciones decrecientes* si sus imágenes invierten el orden de los valores de la variable independiente, es decir,

$$f(x_1) < f(x_2) \text{ si } x_1 > x_2.$$

3.1.1 Gráfica de una función

Una función real de variable real $f : A \to B$, se representa gráficamente por un conjunto de puntos en el plano mediante el llamado *método de coordenadas de Descartes.*[2] Este método consiste en fijar primeramente un sistema de rectas reales perpendiculares debidamente numeradas y orientadas, de modo que se corten ambas en su punto correspondiente al número cero. Al plano, junto con las dos rectas así dispuestas, se le llama *plano cartesiano.* A una de las rectas, escogida y orientada arbitrariamente se le llama *eje de las abscisas* y a la otra se le llama *eje de las ordenadas* y se toma como su orientación positiva, la que le da el eje de las abscisas al rotarse noventa grados en sentido contrario al de las manecillas del reloj y caer sobre ese segundo eje. Cada punto P del plano cartesiano se identifica con la pareja ordenada de números (a, b) donde a, llamado la *abscisa de* P, es el número real correspondiente al punto sobre el eje de las abscisas donde corta la perpendicular bajada de P a ese eje, y b llamado la *ordenada de* P, es el número real correspondiente al punto sobre el eje de las ordenadas donde corta la perpendicular bajada de P a ese eje. A la pareja ordenada (a, b) que identifica al punto P, se le llaman las *coordenadas de* P relativas al plano cartesiano inicial.

La gráfica de $f : A \to B$ se construye de la manera siguiente: En la recta del eje de las abscisas se marca el dominio A de la variable independiente x. Para cada número real $a \in A$, se determina el punto en el plano cartesiano de abscisa a y ordenada $y = f(a)$. Este proceso identifica al conjunto del plano denominado *gráfica de la función* f, definido así:

$$\text{Gráfica de } f = \big\{(a, f(a)) \text{ con } a \in A\big\}.$$

En la figura 3.2 se ilustra este conjunto.

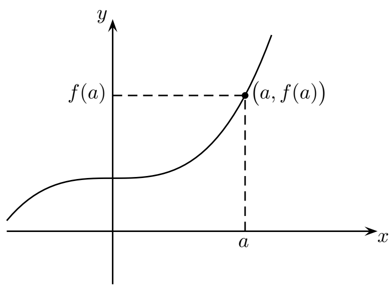


Figura 3.2 Gráfica de f

Ejemplo 3.9 En la figura 3.3 se muestra la gráfica de la función real

$$f(x) = \begin{cases} 2x - 3, & \text{si } x \in [2, 3) \\ -x + 6, & \text{si } x \in [3, 4], \end{cases}$$

cuyo dominio es el intervalo $[2, 4]$. ◁

[2]Por René Descartes.

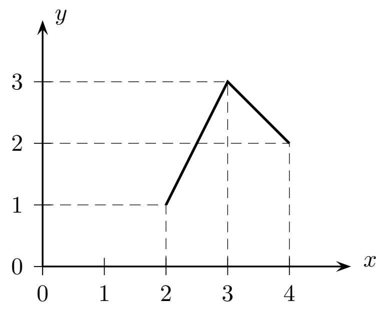


Figura 3.3 Gráfica de $f(x)$ para el ejemplo 3.9

Nota Importante:
La gráfica de una función sólo puede tener un punto sobre cada recta perpendicular al eje de las abscisas. (¿Por qué?)

3.2 Operaciones con funciones

1. Suma y multiplicación de funciones
Sea $A \subset \mathbb{R}$ y consideremos dos funciones reales $f : A \to \mathbb{R}$ y $g : A \to \mathbb{R}$.

a) A la función $f + g : A \to \mathbb{R}$ definida por

$$(f + g)(x) = f(x) + g(x), \quad \text{para cada } x \in A,$$

se le llama *función suma* de las funciones f y g.

b) A la función $f \cdot g : A \to \mathbb{R}$ definida por

$$(f \cdot g)(x) = f(x) \cdot g(x) \text{ para cada } x \in A$$

se le denomina *función multiplicación* o *función producto* de f y g.

Ejemplo 3.10 La función suma $f+g$, de las funciones $f : (0, 1) \to \mathbb{R}$ y $g : (0, 1) \to \mathbb{R}$ definidas por $f(x) = x^2 + x + 1$ y $g(x) = x^3 + 2x$, es la función $f + g : (0, 1) \to \mathbb{R}$ definida por

$$(f + g)(x) = x^3 + x^2 + 3x + 1.$$

El valor de la función suma $f + g$ en $x = 2$ es

$$(f + g)(2) = f(2) + g(2) = 19. \qquad \triangleleft$$

Ejemplo 3.11 Para las funciones reales $f : (0,1) \to \mathbb{R}$ y $g : (0,1) \to \mathbb{R}$ definidas por $f(x) = x^2 + x + 1$ y $g(x) = x^3$, la función producto es la función $f \cdot g : (0,1) \to \mathbb{R}$

$$(f \cdot g)(x) = x^3(x^2 + x + 1). \qquad \triangleleft$$

Análogamente a la definición de suma y producto de dos funciones reales $f : A \to \mathbb{R}$ y $g : A \to \mathbb{R}$, se definen *la diferencia*, $f - g : A \to \mathbb{R}$ y *el cociente*, $\dfrac{f}{g} : \{x \in A \text{ con } g(x) \neq 0\} \to \mathbb{R}$ de dos funciones como dos nuevas funciones cuyas reglas de correspondencia toman la forma:

$$(f - g)(x) = f(x) - g(x) \quad \text{para cada } x \in A,$$

y

$$\frac{f}{g}(x) \overset{\text{def}}{=} \frac{f(x)}{g(x)}.$$

Ejemplo 3.12 El cociente de las funciones

$$f(x) = 2x + 1 \quad \text{y} \quad g(x) = x^2 - 1$$

sólo está definido para valores de la variable independiente distintos de 1 y -1. Es decir, la función cociente

$$\frac{f}{g}(x) = \frac{2x + 1}{x^2 - 1}$$

tiene por dominio el conjunto $\mathbb{R} - \{1, -1\}$. $\triangleleft$

2. La composición de funciones

Otra operación propia entre funciones, es la llamada *composición* de funciones y que presentamos a continuación.

Si y denota una variable que depende de una variable x bajo cierta función $f : A \to \mathbb{R}$ y z es una variable que depende de la variable y bajo otra función $g : B \to \mathbb{R}$, con $\mathrm{Im} f \subset B$, entonces para cada valor de la variable x en A, la aplicación de la regla de correspondencia $y = f(x)$ seguida de la regla de correspondencia $z = g(y)$, establece una regla de correspondencia o función entre la variable x y la variable z que se denota

$$g \circ f : A \to \mathbb{R},$$

$$z = g(f(x)).$$

A la función $g \circ f : A \to \mathbb{R}$, se le llama *función composición* de la función $y = f(x)$ con la función $z = g(y)$. Nótese que para que sea posible construir la función composición $g \circ f$, los valores $f(x)$ que toma la variable y deben pertenecer al dominio de la segunda función $z = g(y)$. Podemos esquematizar la composición de funciones mediante el diagrama siguiente:

$$x \xrightarrow{f} y = f(x) \xrightarrow{g} z = g(f(x)).$$

Ejemplo 3.13 Sean las variables reales x, y, z y las funciones dadas por las reglas de correspondencia

$$y = f(x) = \sqrt{x+2} \quad \text{y} \qquad z = g(y) = y^2 + y + 1.$$

La función composición $z = (g \circ f)(x) = g(f(x))$ toma la forma

$$\begin{aligned} z &= (g \circ f)(x) = g(f(x)) \\ &= g(\sqrt{x+2}) \\ &= \sqrt{x+2} + x + 3. \end{aligned}$$

◁

Ejemplo 3.14 Sean las variables

r : "valor del radio del círculo",
l : "valor de la longitud de la circunferencia",
z : "valor del área del círculo",

y consideremos las funciones

$$l = f(r) = 2\pi r \quad \text{y} \quad z = g(l) = \frac{l^2}{4\pi}$$

que expresan la longitud de la circunferencia como función del radio y el área del círculo como función de la longitud de la circunferencia, respectivamente; entonces la función composición

$$z = (g \circ f)(r) = g(f(r)) = g(2\pi r) = \pi r^2$$

expresa el área del círculo como función del radio. ◁

Ejemplo 3.15 Considere las funciones reales de variable real $f : (0, 1) \to (5, 8)$ y $g : (0, \infty) \to \mathbb{R}$ definidas por $= f(x) = x^2+x+5$ y $z = g(y) = y+2$, respectivamente.

La función composición $g \circ f$ como función de la variable independiente x es la función $g \circ f : (0, 1) \to \mathbb{R}$ definida por

$$\begin{aligned} z &= (g \circ f)(x) = g(f(x)) \\ &= g(x^2 + x + 5) \\ &= x^2 + x + 7. \end{aligned}$$

Así, por ejemplo, $(g \circ f)(2) = 13$. ◁

Ejemplo 3.16 Si $f : \mathbb{R} \to \mathbb{R}$ y $g : \mathbb{R} \to \mathbb{R}$ son las funciones reales

$$f(x) = x^2 \quad \text{y} \quad g(y) = y + 1,$$

entonces $g \circ f : \mathbb{R} \to \mathbb{R}$ está dada por la expresión

$$(g \circ f)(x) = x^2 + 1,$$

mientras que la función $f \circ g : \mathbb{R} \to \mathbb{R}$ se escribe como

$$(f \circ g)(y) = (y + 1)^2.$$

◁

La operación de composición tiene las propiedades siguientes, suponiendo que las funciones cumplan con los requisitos para llevar a cabo esas operaciones:

a) $f \circ (g + h) = f \circ g + f \circ h$

b) $(f \circ g)^{-1} = g^{-1} \circ f^{-1}$.

3.3 Funciones racionales y trigonométricas

A las funciones reales de variable real definidas en toda la recta y de la forma

$$f(x) = x^n + a_1 x^{n-1} + \cdots + a_{n-1}x + a_n$$

donde $a_1, \cdots, a_n$ son números reales, se les llama *funciones polinomiales.*

En los casos particulares:

$$f(x) = ax + b \text{ con } a, b \in \mathbb{R},$$

la función se llama *función lineal* y

$$f(x) = ax^2 + bx + c \quad \text{con } a, b, c \in \mathbb{R},$$

la función se llama *función cuadrática.*

La suma, el producto y la composición de funciones polinomiales es una función polinomial.

Al cociente de dos funciones polinomiales se le conoce como *función racional*:

$$r(x) = \frac{p(x)}{q(x)} = \frac{x^n + a_1 x^{n-1} + \cdots + a_{n-1}x + a_n}{x^m + b_1 x^{m-1} + \cdots + b_{m-1}x + b_m},$$

y su dominio de definición es el conjunto de los números reales excepto las raíces del denominador $q(x)$, es decir, se excluyen los números $a \in \mathbb{R}$ tales que $q(a) = 0$.

Nótese que si el grado del polinomio $q(x)$ en el denominador es menor que el grado del polinomio $p(x)$ en el numerador, se puede reducir la expresión para $r(x)$ a la forma

$$r(x) = a(x) + \frac{x^k + a_1 x^{k-1} + \cdots + a_{k-1}x + a_k}{x^s + b_1 x^{s-1} + \cdots + b_{s-1}x + b_s}$$

con $s \geq k$ y donde $a(x)$ es una función polinomial.

3.3.1 Medición de ángulos: radianes

En los cursos elementales de trigonometría, la unidad de medida para los ángulos es el grado, entendido como la medida del ángulo central que resulta de dividir el círculo en 360 partes iguales. Para efectos del cálculo, la medida de los ángulos se hace con números reales. Para lo anterior, se asigna como medida de un ángulo,

Positivo *Negativo*

Figura 3.4 Signo de ángulos

el número real correspondiente a la longitud con signo, del arco que subtiende el ángulo sobre la circunferencia de radio unitario, como se muestra en la figura 3.4. Se considera sentido positivo para el arco si se recorre en sentido contrario al de las manecillas del reloj.

A partir de la convención anterior, el ángulo de longitud de arco 2π es el ángulo de 360 grados. A la medida del ángulo que subtiende un arco de longitud 1 lo llamaremos *radián* y su medida en grados es de $\dfrac{360}{2\pi} \approx 57.29$ grados. En general, si el ángulo θ mide x grados, entonces su longitud de arco o medida en radianes será de $x\left(\dfrac{2\pi}{360}\right)$ radianes y, recíprocamente, si la longitud en radianes del ángulo θ es de y radianes, su medida en grados será de $y\left(\dfrac{360}{2\pi}\right)$ grados.

Al utilizar longitudes de arco para medir ángulos, tiene sentido hablar de ángulos de cualquier longitud de arco, tanto positiva como negativa.

Ejemplo 3.17 El ángulo de medida -100 radianes corresponde al que se obtiene al recorrer $\dfrac{100}{2\pi}$ veces la circunferencia unitaria en el sentido de las manecillas del reloj. $\triangleleft$

Ejemplo 3.18 Los ángulos de 30, 45 y 60 grados miden respectivamente $\dfrac{\pi}{6}, \dfrac{\pi}{4}$ y $\dfrac{\pi}{3}$ radianes. $\triangleleft$

3.3.2 Las funciones trigonométricas

Consideremos en un plano cartesiano la circunferencia de centro el origen y radio unitario. Para cada número real x, consideremos el arco de circunferencia de longitud x, con extremo el punto inicial $A = (1, 0)$. Sea P_x el extremo final del arco $\overset{\frown}{AP_x}$ de longitud x. Se definen las funciones seno y coseno de la forma siguiente:

$$\text{sen} : \mathbb{R} \to [-1, 1], \qquad \text{sen}\, x = \text{ordenada de } P_x$$

y

$$\cos : \mathbb{R} \to [-1, 1], \qquad \cos x = \text{abscisa de } P_x.$$

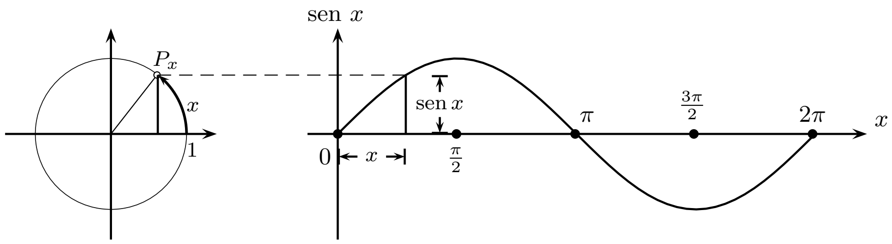


Figura 3.5 Gráfica de $\operatorname{sen} x$

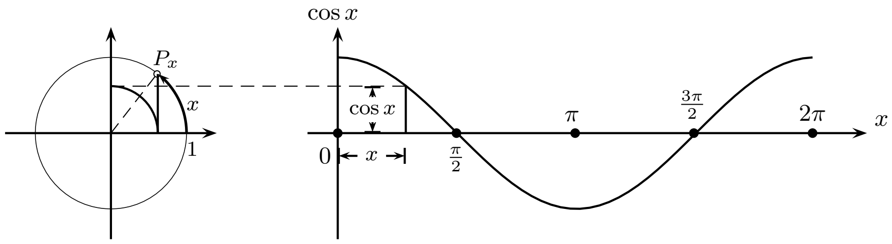


Figura 3.6 Gráfica de $\cos x$

Ver figuras 3.5 y 3.6, respectivamente.

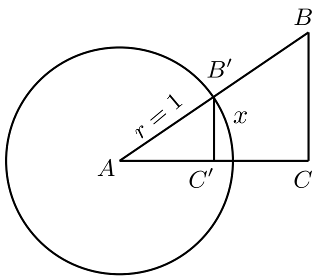
NOTA IMPORTANTE:
Si AB es la hipotenusa de un triángulo rectángulo de vértices A, B y C, y si la medida del ángulo BAC es x radianes, el valor de $\operatorname{sen} x$ es la razón entre el cateto BC, opuesto al ángulo x, y la hipotenusa AB, y el valor de $\cos x$ es la razón entre el cateto AC, adyacente al ángulo x, y la hipotenusa AB, como se muestra en la figura 3.7.

Las funciones $\operatorname{sen} x$ y $\cos x$ tienen las propiedades siguientes:

1. $\operatorname{sen}(-x) = -\operatorname{sen} x, \quad \cos(-x) = \cos x$
2. $\operatorname{sen}(x + 2\pi) = \operatorname{sen} x, \quad \cos(x + 2\pi) = \cos x$
3. $\operatorname{sen}(0) = 0, \quad \operatorname{sen}\left(\frac{\pi}{2}\right) = 1, \quad \operatorname{sen} \pi = 0$
4. $\cos(0) = 1, \quad \cos\left(\frac{\pi}{2}\right) = 0, \quad \cos(\pi) = -1$
5. $\operatorname{sen}(x + y) = \operatorname{sen} x \cos y + \operatorname{sen} y \cos x,$

   $\cos(x + y) = \cos x \cos y - \operatorname{sen} x \operatorname{sen} y$
6. $\cos\left(x - \frac{\pi}{2}\right) = \operatorname{sen} x, \quad \operatorname{sen}\left(x + \frac{\pi}{2}\right) = \cos x$

7. $\cos^2 x + \operatorname{sen}^2 x = 1$.



$$\operatorname{sen} x = B'C' = \frac{BC}{AB}$$

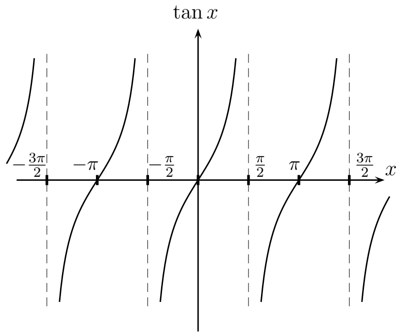
$$\cos x = AC' = \frac{AC}{AB}$$

Figura 3.7 El seno y el coseno de x

A las funciones generadas mediante las operaciones entre las funciones $\operatorname{sen} x$ y $\cos x$ se les denomina *funciones trigonométricas* y tienen como propiedad fundamental su periodicidad (propiedad 2 de la lista anterior). A partir de las funciones seno y coseno se definen las siguientes cuatro funciones trigonométricas adicionales:

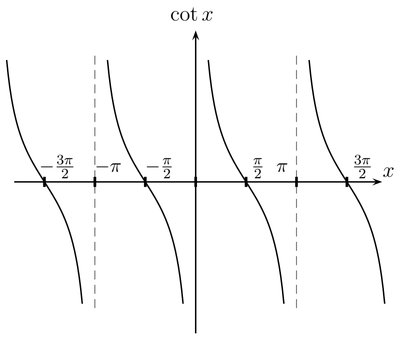
1. La función tangente

$$\tan : \mathbb{R} \setminus \{n\pi \pm \frac{\pi}{2}, n \in \mathbb{N}\} \to \mathbb{R}$$
$$\tan x = \frac{\operatorname{sen} x}{\cos x}$$



2. La función cotangente

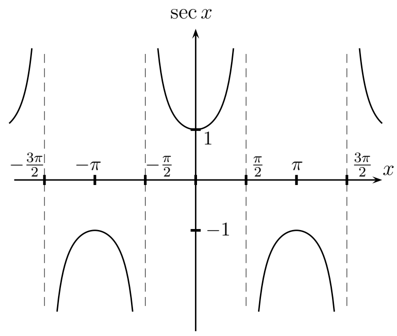
$$\cot : \mathbb{R} \setminus \{n\pi, n \in \mathbb{N}\} \to \mathbb{R}$$
$$\cot x = \frac{\cos x}{\operatorname{sen} x}$$



3. La función secante

$$\sec : \mathbb{R} \setminus \{n\pi \pm \frac{\pi}{2}, n \in \mathbb{N}\} \to \mathbb{R}$$

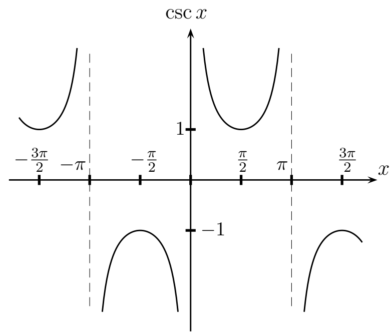
$$\sec x = \frac{1}{\cos x}$$



4. La función cosecante

$$\csc : \mathbb{R} \setminus \{n\pi, n \in \mathbb{N}\} \to \mathbb{R}$$

$$\csc x = \frac{1}{\operatorname{sen} x}$$



3.3.3 Las funciones trigonométricas inversas

La función $f(x) = \operatorname{sen} x$ es uno a uno en los intervalos abiertos

$$I_n = \left(n\pi - \frac{\pi}{2}, n\pi + \frac{\pi}{2}\right) \quad \text{con } n = 0, \pm 1, \pm 2, \ldots$$

y por lo tanto, en cada intervalo I_n existe su función inversa, llamada *función arco seno*,

$$\operatorname{arcsen} : (-1, 1) \to \left(n\pi - \frac{\pi}{2}, n\pi + \frac{\pi}{2}\right)$$

$$\operatorname{arcsen} x = y, \text{ donde } \operatorname{sen} y = x.$$

Análogamente, la función $f(x) = \cos x$ es uno a uno en cada uno de los intervalos

$$J_n = (n\pi, n\pi + \pi) \text{ con } n = 0, \pm 1, \pm 2, \cdots$$

y en cada intervalo J_n está definida la función inversa de $\cos x$ que se llama *función arco coseno,*

$$\arccos : (-1, 1) \to (n\pi, n\pi + \pi),$$

$$\arccos x = y, \quad \text{donde} \quad \cos y = x.$$

Las gráficas correspondientes a los intervalo I_0 y J_0 de estas funciones aparecen en las figuras 3.8(a) y 3.8(b), respectivamente.

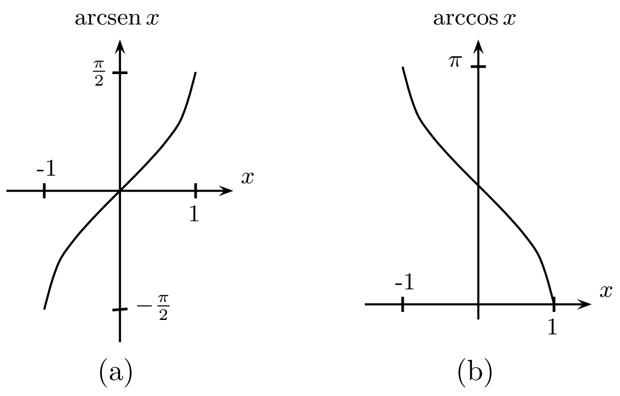


Figura 3.8 Las funciones arcsen$x$ y arccos$x$

La función $f(x) = \tan x$ es también uno a uno en los intervalos abiertos I_n y por lo tanto, en cada intervalo I_n existe su función inversa, llamada *función arco tangente*,

$$\arctan : \mathbb{R} \to \left(n\pi - \frac{\pi}{2}, n\pi + \frac{\pi}{2} \right)$$
$$\arctan\, x = y, \text{ donde } \tan y = x.$$

La gráfica de la función $\arctan x$ correspondiente al intervalo I_0 se muestra en la figura 3.9

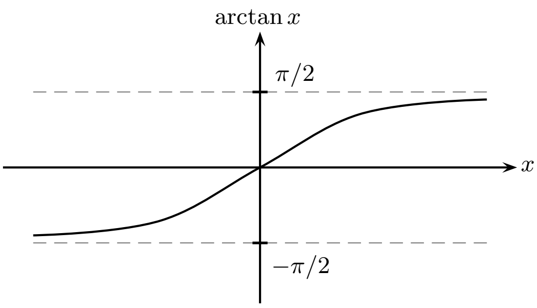


Figura 3.9 La función $f(x) = \arctan x$

De manera similar se definen las funciones inversas de las funciones trigonométricas $\cot x$, $\sec x$ y $\csc x$, en aquellos dominios en las que estas funciones sean inyectivas.

Ejercicios y problemas del capítulo

1. Considere la regla de correspondencia $f(x) = \dfrac{1}{x^2 - 1}$. ¿Cuál debe ser el dominio de la variable independiente x para que la regla defina correctamente una función real de variable real? ¿Cuál es la imagen de la función?

2. Considere las variables reales siguientes.

   y: "valor del área del triángulo equilátero,"
   x: "valor de la longitud del lado del triángulo equilátero,"
   z: "valor del perímetro del triángulo equilátero",
   w: "valor de la altura del triángulo equilátero".

   Escriba:

   (a) la variable y en función de la variable x,
   (b) la variable y en función de la variable z,
   (c) la variable w en función de la variable y,
   (d) la variable x en función de la variable z,
   (e) la variable x en función de la variable y.

3. Pruebe que las funciones siguientes son uno a uno (es decir, inyectivas) y determine su función inversa.

   (a) $f : \mathbb{R} \to [1, \infty), \quad y = f(x) = x^2 + x + 1$
   (b) $l : [0, 2] \to \mathbb{R}, \; y = l(x) = x^2 + x + 1$
   (c) $h : \mathbb{R} \to \mathbb{R}, \; y = h(x) = 2x + 3$
   (d) Pruebe que toda función monótona (creciente o decreciente) es uno a uno.

4. En la figura 3.10 aparece la gráfica de la función $y = f(x)$.

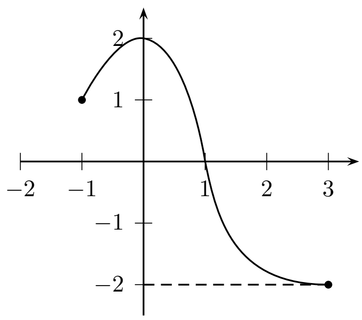


Figura 3.10 Gráfica de la función del ejercicio 4

Conteste las preguntas siguientes:

(a) ¿Cuál es el dominio de f?
(b) ¿Cuál es la imagen de f?
(c) ¿En qué intervalo es f creciente?
(d) ¿En qué intervalo es decreciente?
(e) ¿Tiene f función inversa en el conjunto $(-1, 3)$? Si ése es el caso, ¿cuál es el dominio y cuál el contradominio de la inversa?

5. Dibuje la gráfica de cada una de las siguientes funciones con dominio y contradominio el conjunto de los números reales.

(a) $f(x) =\mid 2x + 3 \mid$
(b) $f(x) = [x]$ =máximo entero menor o igual a x.
(c) $f(x) = f(x) = x - [x]$
(d) $f(x) = x+ \mid x^2 + 3x + 1 \mid$.

6. Una función $f : \mathbb{R} \to \mathbb{R}$ toma los valores $f(0) = 1$ y $f(2) = 1$ y su gráfica está formada por segmentos de recta con pendiente -1 si $x < 0$, pendiente 0 en [0,2] y pendiente 1 si $x > 2$. Dibuje la gráfica de la función g en cada uno de los casos siguientes:

(a) $g(x) = f(x)$
(b) $g(x) = -f(-x)$
(c) $g(x) = f(2x)$
(d) $g(x) = f(x + 2)$
(e) $g(x) = f(3x - 2)$.

7. Exprese la regla de correspondencia o función $y = f(x)$ entre la variable x y la variable y en los casos siguientes.

(a) x: "valor del perímetro de un triángulo equilátero",
y: "valor del área de un triángulo equilátero".
(b) x: "valor del ángulo formado por los lados iguales de un triángulo isósceles de perímetro 12",
y: "valor del área de un triángulo isósceles de perímetro 12".
(c) x: "valor del área de una esfera",
y: "valor del radio de la esfera en cuestión."

8. Una función $f(x)$ definida en todo $\mathbb{R}$ se dice que es una *función par* si

$$f(x) = f(-x)$$

para cada $x \in \mathbb{R}$. Análogamente se dice que es una *función impar* si

$$f(x) = -f(-x).$$

(a) Diga para qué valores del número natural n, la función $f(x) = x^n$ es una función impar.

(b) Pruebe que para cada función $f(x)$, la función

$$p(x) = \frac{1}{2}(f(x) + f(-x))$$

es una función par.

(c) Pruebe que toda función definida en $\mathbb{R}$ se puede escribir como suma de una función par y una función impar.

9. Exprese en grados la medida de los ángulos que subtienden arcos en la circunferencia unitaria de longitud: (a) 20 radianes, (b) -12 radianes, (c) $\pi/6$ radianes, (d) 7π radianes.

10. Sea $h : \mathbb{R} \to \mathbb{R}$ la función racional

$$h(x) = \frac{x^2 + 2x + 3}{x^4 + 3}.$$

(a) Encuentre los valores de x tales que $h(x) = 0$.

(b) Encuentre la imagen de h.

11. Sean las funciones reales de variable real

$$f(x) = ax + b \quad \text{y} \quad g(x) = cx + d.$$

(a) ¿Cuándo se tiene $f \circ g = g \circ f$?

(b) ¿Cuándo $f \circ g = f$?

(c) ¿Cuándo $f \circ g = g$?

12. A partir de las propiedades de las funciones $\operatorname{sen} x$ y $\cos x$, deduzca las fórmulas trigonométricas siguientes.

(a) $\cos(x + y) = \cos x \cos y - \operatorname{sen} x \operatorname{sen} y$

(b) $\tan^2(x) = \sec^2(x) - 1$

(c) $\tan(x + y) = \dfrac{\tan x + \tan y}{1 - \tan x \tan y}$

(d) $\cos(\operatorname{arcsen}(y)) = \pm\sqrt{1 - y^2}$

(e) $\operatorname{sen}(\arccos(y)) = \pm\sqrt{1 - y^2}$

(f) $\tan(\operatorname{arcsen} x) = \dfrac{x}{\sqrt{1 - x^2}}$

(g) $\operatorname{arccsc}(\tan x) = \dfrac{\sqrt{1 + x^2}}{x}$

13. Determine los intervalos en los cuales las funciones $\tan x$, $\sec x$, $\csc x$ y $\cot x$ tienen función inversa.

14. Demuestre que $|a \operatorname{sen} x + b \cos x| \leqslant \sqrt{a^2 + b^2}$ para toda $x \in \mathbb{R}$.

Capítulo 4

Fundamentos del Cálculo

El límite de una función real de variable real $y = f(x)$ *en un punto* x_0 *de su dominio es el concepto fundamental del cálculo. Es ésta una noción asociada al comportamiento de los valores de la función en los puntos vecinos del punto* x_0*, que permite definir la idea de continuidad y los conceptos fundamentales de derivada e integral de una función.*

La definición de límite es reconocida como una de las máximas expresiones del discurso matemático moderno y su manejo es imprescindible para una clara comprensión del cálculo y sus aplicaciones. En este capítulo se presenta y se estudia la noción de límite a partir del concepto de convergencia de sucesiones, definiéndose luego las funciones continuas como aquéllas que preservan precisamente la convergencia de sucesiones. La diferencia entre los conceptos y métodos del cálculo y los que usualmente se manejan en el álgebra y la geometría, radica en que los primeros se definen en términos de propiedades o procesos con conjuntos infinitos.

Este capítulo incluye las demostraciones de los resultados básicos del análisis matemático y se presentan así para ir introduciendo al estudiante en el manejo de las técnicas de argumentación y prueba propias de esta área de las matemáticas.

4.1 Sucesiones reales

> Una *sucesión real* es un conjunto de números reales ordenado mediante el conjunto de los números naturales.

En otras palabras, una sucesión es un conjunto de números reales etiquetados con números naturales, de tal manera que la etiqueta especifica el lugar o el orden que ocupa cada elemento en la sucesión. La etiqueta, al ser un número natural, nos señala cuál es el primer elemento de la sucesión, cuál el segundo, cuál el tercero, etc. Para definir una sucesión real, se necesita especificar los números que la integran y el lugar que ocupan según el orden de sus etiquetas.

Ejemplo 4.1 La sucesión con primer elemento el número 1, con segundo elemento el número $1/2$, con tercer elemento el número $1/3$, y así, en general con i-ésimo elemento el número $1/i$, para los valores $i = 1, 2, \ldots$, se puede representar escribiendo

$$1_{\text{primero}}, \left(\frac{1}{2}\right)_{\text{segundo}}, \left(\frac{1}{3}\right)_{\text{tercero}}, \ldots, \left(\frac{1}{i}\right)_{i-\text{ésimo}}, \cdots$$

Los puntos sucesivos al final significan que la sucesión se extiende de acuerdo al orden creciente de las etiquetas. ◁

Una manera usual de describir una sucesión real consiste en dar la regla, o fórmula, que nos permita conocer, para cada valor $i = 1, 2, \ldots$, de la etiqueta, el número que lleva dicha etiqueta. Esto se hace denotando por s_i el número que ocupa el i-ésimo lugar, para cada uno de los lugares $i = 1, 2, \ldots$ y mostrando cómo se calcula el valor de s_i en términos del valor i de su etiqueta. Por ejemplo, la sucesión

$$s_i = \frac{1}{i} \quad \text{para} \quad i = 1, 2, \ldots$$

representa, en forma compacta, la sucesión del ejemplo 4.1. De esa manera quedan determinados todos los elementos de la sucesión y podemos saber directamente cuál es el número que ocupa cada lugar. Por ejemplo, el elemento que ocupa el lugar 130 es el número $1/130$. Note que siendo el número de etiquetas infinito, cada sucesión consta de un número infinito de números que pueden en principio repetirse o ser el mismo número para varios lugares.

En general, para denotar una sucesión escribiremos entre llaves el número que ocupa el i-ésimo lugar y fuera de las llaves, como subíndice, escribiremos $i = 1$ y como supraíndice el símbolo ∞ para significar que la etiqueta toma valores sobre todos los números naturales.

Ejemplo 4.2

1. La sucesión

$$s_i = (-1)^i \quad \text{para} \quad i = 1, 2, \ldots$$

es una sucesión cuyos elementos sólo toman dos valores.

2. El símbolo

$$\left\{\frac{i}{i^2 - 8}\right\}_{i=1}^{\infty}$$

representa la sucesión real

$$\left\{\frac{-1}{7}, \frac{-1}{2}, 3, \frac{1}{2}, \cdots, \frac{i}{i^2 - 8}, \cdots\right\}.$$

3. En la sucesión real $\left\{(-1)^i\sqrt{i^2+1}\right\}_{i=1}^{\infty}$ el número $\sqrt{101}$ ocupa el décimo lugar.

◁

Asociado al concepto de sucesión, se tiene de manera natural el concepto de *subsucesión de una sucesión,* entendida como una nueva sucesión cuyos elementos forman un subconjunto de la primera y su orden como elementos de la subsucesión preserva el orden que esos mismos elementos tenían en la sucesión inicial. Es decir, si en la subsucesión un elemento es posterior a otro, como elementos de la sucesión inicial también el primer elemento era posterior al segundo. Esto lo escribiremos de manera precisa con la definición siguiente.

Definición 4.1 *1. La sucesión de etiquetas $\{m_i\}_{i=1}^{\infty}$ es* ***creciente*** *si siempre que $j > k$ se tiene $m_j > m_k$ para j, k números naturales.*

2. Una sucesión $\{s_i\}_{i=1}^{\infty}$ se dice ***subsucesión*** *de una sucesión $\{a_i\}_{i=1}^{\infty}$, si existe una sucesión creciente de etiquetas $\{m_i\}_{i=1}^{\infty}$ tal que $s_i = a_{m_i}$ para $i = 1, 2, \ldots$*

Ejemplo 4.3 1. La sucesión $\left\{\sqrt{2i+3}\right\}_{i=1}^{\infty}$ cuyos elementos son $\sqrt{5}$, $\sqrt{7}$, $\sqrt{9}$, $\sqrt{11}$, $\cdots, \sqrt{2i+3}, \ldots$ es una subsucesión de la sucesión $\left\{\sqrt{i}\right\}_{i=1}^{\infty}$ cuyos elementos son $\sqrt{1}, \sqrt{2}, \sqrt{3}, \ldots$ En este caso, la sucesión creciente de etiquetas que dan lugar a la subsucesión es $\{2i+3\}_{i=1}^{\infty}$, de tal manera que el elemento de la subsucesión con etiqueta i es el elemento de la sucesión que tiene etiqueta $2i+3$ para $i = 1, 2, \ldots$

2. La sucesión $\{c_i\}_{i=1}^{\infty} = \{6i+1\}_{i=1}^{\infty}$ es subsucesión de $\{a_i\}_{i=1}^{\infty} = \{2i+1\}_{i=1}^{\infty}$, donde la relación entre las etiquetas es $c_i = a_{3i}$ para cada $i = 1, 2, \ldots$ $\triangleleft$

Dadas dos sucesiones de números reales $\{a_i\}_{i=1}^{\infty}$ y $\{b_i\}_{i=1}^{\infty}$, podemos sumar o multiplicar término a término estas sucesiones para formar nuevas sucesiones reales. Así, a la sucesión $\{s_i\}_{i=1}^{\infty}$ cuyo i-ésimo término se forma sumando el i-ésimo término de la sucesión $\{a_i\}_{i=1}^{\infty}$ con el i-ésimo término de la sucesión $\{b_i\}_{i=1}^{\infty}$,

$$s_i = a_i + b_i \text{ para } i = 1, 2, \ldots$$

se le llama *sucesión suma* de las dos sucesiones iniciales y se denota

$$\{s_i\}_{i=1}^{\infty} = \{a_i\}_{i=1}^{\infty} + \{b_i\}_{i=1}^{\infty} = \{a_i + b_i\}_{i=1}^{\infty}.$$

Análogamente, a la sucesión $\{p_i\}_{i=1}^{\infty}$ cuyo i-ésimo término se forma multiplicando el i-ésimo término de la sucesión $\{a_i\}_{i=1}^{\infty}$ con el i-ésimo término de la sucesión $\{b_i\}_{i=1}^{\infty}$,

$$p_i = a_i b_i \text{ para } i = 1, 2, \cdots$$

se le llama *sucesión producto* de las sucesiones iniciales y se denota

$$\{p_i\}_{i=1}^{\infty} = \{a_i\}_{i=1}^{\infty} \cdot \{b_i\}_{i=1}^{\infty} = \{a_i b_i\}_{i=1}^{\infty}.$$

Las operaciones de suma y producto de sucesiones, heredan las propiedades de campo de las operaciones de los números reales, como son la conmutatividad, asociatividad, distributividad, existencia de neutro aditivo y multiplicativo, existencia de inverso aditivo y cuando la sucesión está formada de números distintos de cero, la existencia de inverso multiplicativo.

4.2 Convergencia de sucesiones

Se denota con $(L - r, L + r)$ y se llama *intervalo abierto con centro en el número real L y radio $r > 0$*, al conjunto

$$(L - r, L + r) = \{x \in \mathbb{R} \text{ tales que } |x - L| < r\}.$$

Se dice que una sucesión $\{s_i\}_{i=1}^{\infty}$ de números reales es *convergente a un número real L* si los elementos de la sucesión se aproximan al número L "tanto como se quiera" a medida que crece la magnitud de las etiquetas de esos elementos. En términos precisos, lo anterior se enuncia de la siguiente manera:

Definición 4.2 *La sucesión de números reales $\{s_i\}_{i=1}^{\infty}$* ***converge al número*** *L si para cada intervalo I con centro L existe una etiqueta N_I tal que todos los elementos de la sucesión cuya etiqueta es posterior a N_I pertenecen a dicho intervalo. Cuando el enunciado anterior es verdadero, escribimos simbólicamente*

$$\{s_i\}_{i=1}^{\infty} \;\to\; L$$

o

$$\lim_{i\to\infty} s_i = L.$$

Al número L se le llama ***límite de la sucesión*** $\{s_i\}_{i=1}^{\infty}$.

Dado que un intervalo con centro L queda determinado por su radio r, la definición 4.2 se puede parafrasear como sigue:

"$\{s_i\}_{i=1}^{\infty}$ converge a L si para cada $r > 0$, existe un número natural N_r tal que si $i > N_r$ entonces $|s_i - L| < r$."

Nota Importante:
Para fijar mejor la definición anterior, considere las observaciones siguientes.

1. *La definición de sucesión convergente no dice cómo encontrar el límite de una sucesión, sino sólo qué propiedad define al límite de una sucesión. En ese sentido, la definición sólo dice qué debemos hacer para comprobar que un cierto número es efectivamente el límite de la sucesión.*

2. *La aproximación y acumulación de los elementos de la sucesión al límite L se expresa mediante la pertenencia a intervalos abiertos centrados en L y cada intervalo queda definido cuando se da el valor de su radio.*

3. *Para comprobar que un número L es el límite de una sucesión, la definición nos exige que para cada intervalo I centrado en L encontremos una etiqueta N_I a partir de la cual todos los elementos de la sucesión con etiqueta mayor pertenezcan a ese intervalo. Esa etiqueta, cuya existencia hay que mostrar, no es la única con esa propiedad ya que cualquier otra mayor que ella también tendrá esa propiedad.*

4. *El valor de la etiqueta a partir de la cual los elementos de la sucesión pertenecen a un intervalo dado, depende del tamaño o radio de ese intervalo y mientras más pequeño sea ese radio, en general más grande tendrá que ser la etiqueta.*

5. *La convergencia de una sucesión a un número L no se altera si se modifica el valor o el orden de cualquier número finito de elementos de la sucesión. Esto es así porque la convergencia de una sucesión es una propiedad que sólo tiene que ver con el comportamiento de sus elementos a la larga, es decir cuando su orden crece indefinidamente y no depende de lo que suceda con sus primeros elementos.*

La siguiente observación, que enunciamos en forma de lema, nos será de gran utilidad en lo que sigue.

Lema 4.1 *Si una sucesión $\{s_i\}_{i=1}^{\infty}$ converge a L, entonces $\{s_i - L\}_{i=1}^{\infty}$ es una sucesión convergente a cero y, recíprocamente, si la sucesión $\{s_i - L\}_{i=1}^{\infty}$ converge a cero, entonces la sucesión $\{s_i\}_{i=1}^{\infty}$ converge a L.*

NOTA IMPORTANTE:

1. *Toda sucesión constante es convergente: $\{s_i\}_{i=1}^{\infty} = \{a\}_{i=1}^{\infty}$ converge a a.*
2. *Toda subsucesión de una sucesión convergente a L es convergente a L.*

Ejemplo 4.4 La sucesión $\left\{\dfrac{3i+1}{2i+8}\right\}_{i=1}^{\infty}$ es una sucesión convergente a $3/2$. Para probar lo anterior, aplicando la definición de convergencia, estimamos primero la distancia del i-ésimo término de la sucesión al número $3/2$:

$$\begin{aligned}\left|s_i - \frac{3}{2}\right| &= \left|\frac{3i+1}{2i+8} - \frac{3}{2}\right| = \left|\frac{2(3i+1) - 3(2i+8)}{2(2i+8)}\right| \\ &= \left|\frac{-22}{2(2i+8)}\right| = \left|\frac{11}{(2i+8)}\right|.\end{aligned}$$

Esta distancia se hace cada vez más pequeña a medida de que la etiqueta i crece. Si consideramos ahora el intervalo $I = \left(\dfrac{3}{2} - r, \dfrac{3}{2} + r\right)$ con $r > 0$, según la estimación anterior, el elemento s_i de la sucesión pertenecerá al intervalo I, si su etiqueta i es tal que

$$\left|\frac{11}{(2i+8)}\right| < r,$$

o, lo que es lo mismo, siempre que su etiqueta i sea mayor que la etiqueta N_r dada por

$$N_r = \text{primer natural mayor que } \frac{1}{2}\left(\frac{11}{r} - 8\right).$$

Es decir, se tendrá que

$$\left|s_i - \frac{3}{2}\right| < r$$

siempre que

$$i > \text{ primer natural mayor que } \frac{1}{2}\left(\frac{11}{r} - 8\right),$$

lo cual, según la definición 4.2, significa que la sucesión converge al número $3/2$. ◁

Por ejemplo, si escogemos inicialmente $r = 1/10^6$, los elementos de la sucesión distarán de $3/2$ en menos de una millonésima de unidad si su etiqueta es posterior a la etiqueta

$$N_{\frac{1}{10^6}} = \frac{1}{2}((11)(10^6) - 8) = 5499996.$$

Ejemplo 4.5 La sucesión $\{-1, 1, -1, \cdots\}$ es un ejemplo de una sucesión no convergente. Esto es así, pues si un número L fuera su límite, tendríamos una contradicción que explicamos a continuación. En primer lugar tendríamos que L no podría ser distinto de 1 y -1, pues si ese no fuera el caso, al escoger un intervalo con centro en L con radio r menor que la mitad de la distancia más pequeña de L a 1 y -1, no podríamos encontrar una etiqueta a partir de la cual los elementos de la sucesión pertenecieran a dicho intervalo. Por otro lado si supusiéramos que L es el número 1, tomando $r = 1/2$ tampoco podríamos encontrar una etiqueta a partir de la cual todos los elementos de la sucesión pertenecieran al intervalo de centro $L = 1$ y radio $1/2$, ya que los elementos de la forma $(-1)^i$ con i un número impar están alejados en más de $1/2$ de $L = 1$. Análogamente, podríamos repetir el argumento si supusiéramos que L es igual a -1. Luego, no existe valor real de L que satisfaga la definición de límite de la sucesión; por lo tanto, ésta no es convergente. ◁

4.2.1 Propiedades de las sucesiones convergentes

Enseguida enunciaremos y demostraremos algunas de las propiedades más importantes de las sucesiones convergentes.

Una sucesión de números reales $\{a_i\}_{i=1}^{\infty}$ se dice *acotada* si todos sus elementos se encuentran dentro de un intervalo cerrado de la forma $[-M, M]$ con M algún número mayor que cero, es decir, si

$$-M \leqslant a_i \leqslant M \quad \text{para toda } i = 1, 2, \ldots$$

Proposición 4.2 *Toda sucesión* $\{a_i\}_{i=1}^{\infty}$ *de números reales convergente es acotada.*

Demostración. La proposición se prueba a partir del argumento siguiente: Primero denotemos por L el límite de esa sucesión y por I el intervalo con centro L y radio $r = 1$. Por ser convergente la sucesión, tendremos que existe una etiqueta N a partir de la cual todos los elementos de la sucesión a_i con $i > N$ pertenecen al intervalo I, o lo que es lo mismo, satisfacen

$$|a_i - L| < 1 \quad \text{para toda } i > N. \tag{4.1}$$

Ahora, aplicando la desigualdad del triángulo en (4.1), se tiene

$$|a_i| - |L| \leqslant |a_i - L| < 1$$

y entonces

$$|a_i| < 1 + |L| \text{ para toda } i > N.$$

Tomando ahora $C = \text{máx } \{|a_1|, |a_2|, \ldots, |a_N|\}$ y $M = \text{máx } \{1 + |L|, C\}$ tendremos que

$$|a_i| < M \text{ para todo } \ i = 1, 2, \ldots,$$

lo cual muestra que todos los términos de la sucesión pertenecen al intervalo $[-M, M]$. ∎

Proposición 4.3 *El producto de una sucesión* $\{a_i\}_{i=1}^{\infty}$ *acotada por una sucesión* $\{b_i\}_{i=1}^{\infty}$ *convergente a cero es una sucesión convergente a cero.*

Demostración. Sea $M > 0$ una cota para $\{a_i\}_{i=1}^{\infty}$, es decir,

$$|a_i| < M \quad \text{para toda } i = 1, 2, \ldots$$

La distancia al cero del i-ésimo término de la sucesión producto $\{a_i b_i\}_{i=1}^{\infty}$ para $i = 1, 2, \ldots$ satisface

$$|a_i b_i| \leqslant |a_i||b_i| \leqslant M|b_i|, \tag{4.2}$$

es decir, es siempre menor que M veces la distancia del término b_i a cero. Luego, si tomamos el intervalo con centro 0 y radio $r > 0$, sabemos por la convergencia a cero de la sucesión $\{b_i\}_{i=1}^{\infty}$, que existe un índice N_r tal que

$$|b_i| < \frac{r}{M} \quad \text{para toda} \quad i > N_r. \tag{4.3}$$

Combinando (4.2) y (4.3) tendremos

$$|a_i b_i| < r \quad \text{para toda} \quad i > N_r,$$

lo que muestra que los términos de la sucesión producto distan de cero en menos que r a partir de la etiqueta N_r. Siendo $r > 0$ arbitrario, esto significa que la sucesión producto $\{a_i b_i\}_{i=1}^{\infty}$ converge a cero. ∎

Proposición 4.4 *Si las sucesiones* $\{a_i\}_{i=1}^{\infty}$ *y* $\{b_i\}_{i=1}^{\infty}$ *convergen a cero, entonces la sucesiones* $\{a_i + b_i\}_{i=1}^{\infty}$ *y* $\{a_i - b_i\}_{i=1}^{\infty}$ *también convergen a cero.*

Demostración. Por definición, para probar que $\{a_i + b_i\}_{i=1}^{\infty}$ converge a cero, tenemos que probar que si escogemos arbitrariamente un intervalo de centro 0 y radio r mayor que cero, entonces podemos encontrar una etiqueta N tal que

$$|a_i + b_i| < r$$

siempre que $i > N$. Para encontrar tal N, razonamos de la forma siguiente: Como las sucesiones $\{a_i\}_{i=1}^{\infty}$ y $\{b_i\}_{i=1}^{\infty}$ convergen a cero, para el intervalo $(-\frac{r}{2}, \frac{r}{2})$ y para cada una de ellas, podremos encontrar etiquetas N_1 y N_2 tales que

$$|a_i| < \frac{r}{2} \text{ si } i > N_1$$

y

$$|b_i| < \frac{r}{2} \text{ si } i > N_2.$$

Ahora, si tomamos $N = \text{máx}\{N_1, N_2\}$, o sea N igual a la mayor de las dos etiquetas N_1 y N_2, simultáneamente se cumplirá que

$$|a_i| < \frac{r}{2} \text{ y } |b_i| < \frac{r}{2} \text{ si } i > N$$

y aplicando la desigualdad del triángulo, tendremos que los términos de la sucesión $\{a_i + b_i\}_{i=1}^{\infty}$ satisfarán

$$|a_i + b_i| \leqslant |a_i| + |b_i| < r, \quad \text{si } i > N,$$

lo cual comprueba que podemos encontrar una etiqueta a partir de la cual todos los términos de la sucesión suma $\{a_i + b_i\}_{i=1}^{\infty}$ disten de cero en menos que r. De lo anterior se sigue que $\{a_i + b_i\}_{i=1}^{\infty} \to 0$.

Para probar la convergencia a cero de la sucesión $\{a_i - b_i\}_{i=1}^{\infty}$, el razonamiento es el mismo, pues también se tiene $|a_i - b_i| < |a_i| + |b_i| < r$, si $i > N$. ■

Como consecuencia directa de la proposición 4.4, tenemos la proposición siguiente.

Proposición 4.5 *Una sucesión convergente* $\{s_i\}_{i=1}^{\infty}$ *tiene un sólo límite.*

Demostración. Para demostrar la validez de esta proposición, recurriremos al método de demostración conocido como “de reducción al absurdo”: Supongamos que el enunciado a demostrar es falso y que existe una sucesión $\{s_i\}_{i=1}^{\infty}$ que converge al mismo tiempo a dos números distintos L_1 y L_2 con $L_1 \neq L_2$. Como consecuencia de la suposición anterior, se tendría que las sucesiones $\{s_i - L_1\}_{i=1}^{\infty}$y $\{s_i - L_2\}_{i=1}^{\infty}$ convergen a cero, y como consecuencia del lema anterior tendríamos que su diferencia

$$\{s_i - L_1 - s_i + L_2\}_{i=1}^{\infty} = \{L_2 - L_1\}_{i=1}^{\infty} \to 0$$

converge a cero, lo cual no es posible pues la sucesión $\{L_2 - L_1\}_{i=1}^{\infty}$ es constante y converge a $L_2 - L_1 \neq 0$. Luego, no se puede suponer que la proposición sea falsa porque se llegaría a un absurdo, por lo tanto la proposición tiene que ser verdadera. ■

Enseguida mostraremos que, bajo las operaciones de suma y producto de sucesiones, se preserva la propiedad de convergencia. Esto significa que al sumar y multiplicar sucesiones convergentes obtenemos de nuevo sucesiones convergentes.

Teorema 4.6 *Si* $\{a_i\}_{i=1}^{\infty}, \{b_i\}_{i=1}^{\infty}$ *son sucesiones convergentes y* $\lim\limits_{i\to\infty} a_i = L$ *y* $\lim\limits_{i\to\infty} b_i = M$ *entonces:*

a) La sucesión suma $\{a_i + b_i\}_{i=1}^{\infty} = \{a_i\}_{i=1}^{\infty} + \{b_i\}_{i=1}^{\infty}$ *es convergente y su límite es igual a la suma de los límites de* $\{a_i\}_{i=1}^{\infty}$ *y* $\{b_i\}_{i=1}^{\infty}$

$$\lim_{i\to\infty} (a_i + b_i)_i = L + M.$$

b) La sucesión producto $\{a_i b_i\}_{i=1}^{\infty} = \{a_i\}_{i=1}^{\infty} \cdot \{b_i\}_{i=1}^{\infty}$ *es convergente y su límite es igual al producto de los límites de* $\{a_i\}_{i=1}^{\infty}$ *y* $\{b_i\}_{i=1}^{\infty}$

$$\lim_{i\to\infty} a_i b_i = L \cdot M.$$

c) Si $\{a_i\}_{i=1}^{\infty}$ *es tal que* $a_i \neq 0$ *para* $i = 1, 2\ldots$ *y converge a* $L \neq 0$, *la sucesión* $\left\{\dfrac{1}{a_i}\right\}_{i=1}^{\infty}$ *converge a* $\dfrac{1}{L}$.

Demostración Para probar el inciso a), observamos que si $\lim\limits_{i\to\infty} a_i = L$ y $\lim\limits_{i\to\infty} b_i = M$, entonces las sucesiones

$$\{a_i - L\}_{i=1}^{\infty} \quad \text{y} \quad \{b_i - M\}_{i=1}^{\infty}$$

convergen ambas a cero. La proposición 4.4 implica entonces que su suma

$$\{a_i + b_i - (L + M)\}_{i=1}^{\infty}$$

converge también a cero. Pero esto quiere decir que $\{a_i + b_i\}_{i=1}^{\infty}$ converge a $L + M$, con lo cual se prueba a).

Para probar b) observemos que

$$\{a_i b_i - LM\}_{i=1}^{\infty} = \{a_i b_i - a_i M + a_i M - LM\}_{i=1}^{\infty} \tag{4.4}$$

donde hemos sumado y restado el término $a_i M$ a $a_i b_i - LM$ para cada $i = 1, 2, \ldots$ Ahora, (4.4) se puede escribir como sigue:

$$\{a_i b_i - a_i M + a_i M - LM\}_{i=1}^{\infty} = \{a_i\}_{i=1}^{\infty} \{b_i - M\}_{i=1}^{\infty} + \{M\}_{i=1}^{\infty} \{a_i - L\}_{i=1}^{\infty}$$

y en el lado derecho, el producto de las sucesiones $\{a_i\}_{i=1}^{\infty}\{b_i - M\}_{i=1}^{\infty}$ converge a cero pues $\{a_i\}_{i=1}^{\infty}$ es acotada por ser convergente y $\{b_i - M\}_{i=1}^{\infty}$ converge a cero. El otro término, $\{M\}_{i=1}^{\infty}\{a_i - L\}_{i=1}^{\infty}$, también converge a cero, por las mismas razones. Luego

$$\{a_ib_i - LM\}_{i=1}^{\infty}$$

converge a cero, o lo que es lo mismo,

$$\lim_{i\to\infty} a_ib_i = LM.$$

La prueba de c) se tiene a partir de escribir

$$\left\{\frac{1}{a_i} - \frac{1}{L}\right\}_{i=1}^{\infty} = \left\{\frac{L-a_i}{a_iL}\right\}_{i=1}^{\infty} = \left\{\frac{1}{a_iL}\right\}_{i=1}^{\infty}\{L-a_i\}_{i=1}^{\infty}$$

y notar que la sucesión $\left\{\frac{1}{a_iL}\right\}_{i=1}^{\infty}$ es acotada pues como $\lim_{i\to\infty} a_i = L \neq 0$, tomando el intervalo con centro L y radio $r = |L|/3$ existirá una etiqueta N tal que

$$|a_i - L| < \frac{|L|}{3} \quad \text{si } i > N$$

y aplicando la desigualdad $|a_i - L| \geqslant |L| - |a_i|$ tendremos

$$|a_i| \geqslant \frac{2|L|}{3} \quad \text{si } i > N$$

y, por lo tanto,

$$\frac{1}{|a_i||L|} \leqslant \frac{3}{2L^2} \quad \text{si } i > N.$$

Es decir, la sucesión $\{1/a_iL\}_{i=1}^{\infty}$ es acotada y entonces la sucesión de la derecha en la expresión

$$\left\{\frac{1}{a_i} - \frac{1}{L}\right\}_{i=1}^{\infty} = \left\{\frac{1}{a_iL}\right\}_{i=1}^{\infty}\{L-a_i\}_{i=1}^{\infty}$$

converge a cero y, por lo tanto,

$$\left\{\frac{1}{a_i}\right\}_{i=1}^{\infty} \text{ converge a } \frac{1}{L}. \qquad \blacksquare$$

Nota Importante:

En el inciso c) del teorema 4.6, la condición $L \neq 0$ es necesaria para que el límite de la sucesión $\{1/a_i\}_{i=1}^{\infty}$ exista y sea $1/L$. Si se remueve esa condición, el inciso c) deja de ser válido, como lo muestra la sucesión $\{1/i\}_{i=1}^{\infty}$ que converge a cero pero su recíproca $\left\{\frac{1}{1/i}\right\}_{i=1}^{\infty} = \{i\}_{i=1}^{\infty}$ no converge, pues crece sin límite.

Finalizamos esta subsección con el resultado siguiente, que es de gran utilidad y que es consecuencia de la propiedad de continuidad de los números reales.

Proposición 4.7 *Toda sucesión acotada posee una subsucesión convergente.*[1]

Demostración. Sea $\{c_i\}_{i=1}^{\infty}$ una sucesión cuyos elementos pertenecen al intervalo $[a,b]$. Dividamos ahora el intervalo $[a,b]$ en los intervalos de igual longitud, $\left[a, \dfrac{a+b}{2}\right]$ y $\left[\dfrac{a+b}{2}, b\right]$. De aquí podemos concluir que en al menos uno de esos intervalos habrá una infinidad de puntos de la sucesión $\{c_n\}$. Supongamos que eso ocurre en el primer subintervalo $\left[a, \dfrac{a+b}{2}\right]$. Denotemos este subintervalo por $[a_1, b_1]$, dividámoslo a su vez en los subintervalos de igual longitud $\left[a_1, \dfrac{a_1+b_1}{2}\right]$ y $\left[\dfrac{a_1+b_1}{2}, b_1\right]$, y escojamos alguno de ellos que contenga una infinidad de elementos de la sucesión $\{c_i\}$ y denotémoslo por $[a_2, b_2]$. Repitamos este proceso, dando lugar a una sucesión $I_i = [a_i, b_i]$ de intervalos cerrados anidados de longitud $\dfrac{b-a}{2^i}$ para $i = 1, 2, \ldots$ Los extremos izquierdos $\{a_i\}_{i=1}^{\infty}$ de los intervalos I_k forman una sucesión creciente y acotada de reales y los extremos derechos $\{b_i\}_{i=1}^{\infty}$ forman una sucesión decreciente y acotada. En virtud de la propiedad de continuidad de los números reales, la sucesión de los extremos izquierdos deberá converger a su supremum y la sucesión de extremos derechos convergerá a su infimum. Como la diferencia entre los extremos correspondientes $\{b_i - a_i\}_{i=1}^{\infty}$ tiende a cero por construcción, tendremos que ambas sucesiones convergen a un mismo punto real $c \in [a,b]$. Ahora en cada intervalo I_i escojamos un elemento $c_{m(i)} \in I_i$ de la sucesión $\{c_i\}$, de tal manera que $m(1) < m(2) < \cdots < m(k) < m(k+1) < \cdots$; se puede hacer ésto pues en cada intervalo I_i existe una infinidad de elementos de $\{c_i\}_{i=1}^{\infty}$. Luego, la subsucesión $\left\{c_{m(k)}\right\}$ está bien definida y es convergente al punto c. ■

4.3 Sucesiones monótonas

Una sucesión de números reales $\{s_i\}_{i=1}^{\infty}$ se dice *creciente* si para cada par de etiquetas $k, j = 1, 2 \cdots$ con $k > j$ se tiene que $s_k > s_j$. Análogamente, la sucesión se dice *decreciente* si $k > j$ implica $s_k < s_j$. A las sucesiones crecientes o decrecientes se les llama genéricamente *sucesiones monótonas*.

Ejemplo 4.6 La sucesión $\left\{\dfrac{i-1}{i}\right\}_{i=1}^{\infty}$ es una sucesión creciente, pues si $k > j$ se tiene que

$$s_k - s_j = \frac{k-1}{k} - \frac{j-1}{j} = \frac{k-j}{kj} > 0,$$

es decir,

$$s_k > s_j. \qquad \triangleleft$$

[1]Esta proposición es conocida como *teorema de Bolzano-Weierstrass.*

La propiedad más importante de las sucesiones reales monótonas es que si son acotadas entonces son convergentes. Este resultado es, de hecho, equivalente también a la propiedad de continuidad de los números reales y, por lo tanto, juega un papel relevante en la fundamentación del cálculo.

Teorema 4.8 *Cada sucesión monótona y acotada de números reales es convergente.*

DEMOSTRACIÓN. Supongamos que $\{s_i\}_{i=1}^{\infty}$ es una sucesión real creciente y acotada. En virtud de la propiedad de continuidad de los reales que discutimos en el capítulo 2, existe un número real L que es el supremum del conjunto formado con los elementos de la sucesión, es decir,

$$s_i \leqslant L \quad \text{para } i = 1, 2, \ldots$$

y si $s_i \leqslant M$ para $i = 1, 2, \ldots$ para algún otro número real M, entonces $L \leqslant M$. Probaremos entonces que

$$\lim_{i\to\infty} s_i = L.$$

Para ello, sea $I = (L - r, L + r)$ el intervalo con centro L y radio $r > 0$. Note que no existen elementos de la sucesión en $[L, L + r)$ pues L es cota superior de la sucesión, pero sí deberá existir un elemento s_N de la sucesión en $(L - r, L]$ ya que si así no fuera, entonces $L - r$ sería una cota superior de la sucesión menor que L, lo cual no es posible pues hemos supuesto que L es la mínima cota superior o supremum de la sucesión. Luego, deberá existir un elemento $s_N \in (L - r, L]$ y, al ser la sucesión creciente, todos sus elementos s_i con etiqueta i mayor que N, se encontrarán a la derecha de s_N y consecuentemente pertenecerán al intervalo I. Como I es un intervalo de radio arbitrario, hemos probado que $L = \lim_{i\to\infty} s_i$. ■

El resultado siguiente es de gran utilidad en la prueba de varias proposiciones claves en la materia. Este resultado es también consecuencia de la propiedad de continuidad de $\mathbb{R}$; más aún, es equivalente a esa propiedad.

Proposición 4.9 *Si A es un conjunto acotado de puntos de un intervalo cerrado $[a, b]$ y S es su mínima cota superior, entonces existe una sucesión no-decreciente de puntos de A que converge a S.*

DEMOSTRACIÓN. Si $S \in A$, tomando la sucesión constante $\{a_i\}_{i=1}^{\infty} = \{S\}_{i=1}^{\infty}$ tendremos una sucesión no-decreciente de puntos de A convergente a S. Si $S \notin A$ debe existir un elemento a_1 de A con $a_1 < S$. Como $S = \sup A$, existe $a_2 \in A$ con $a_1 < a_2 < S$ y $S - a_2 < \frac{1}{2}$ ya que si así no fuera, entonces $S - 1/2$ sería el supremum de A, lo cual no puede ser verdadero. Análogamente, y por la misma razón que antes, deberá existir un elemento a_3 de A tal que $a_2 < a_3 < S$ y $S - a_3 < \frac{1}{3}$. Lo anterior nos permite construir una sucesión creciente $a_1 < a_2 < \cdots < a_n < \cdots < S$ de elementos de A tales que $S - a_n < 1/n$ para cada natural n. Luego, $\lim_{n\to\infty} a_n = S$. ■

4.3.1 Criterio de convergencia de Cauchy

El criterio de convergencia de Cauchy[2] proporciona una manera de constatar la convergencia de una sucesión sin que necesariamente se conozca su límite. Este es un resultado de gran utilidad en el análisis matemático y aquí lo presentamos en los términos siguientes.

Definición* 4.3** *Se dice que una sucesión* $\{a_i\}_{i=1}^{\infty}$ *satisface la* ***propiedad de Cauchy *si para cada* $r > 0$ *existe una etiqueta* N_r *tal que para cualquier par de etiquetas* $m, n > N_r$ *se tiene que* $|a_m - a_n| < r$.

Teorema 4.10 *Una sucesión de números reales* $\{a_i\}_{i=1}^{\infty}$ *es convergente si y sólo si satisface la propiedad de Cauchy.*

DEMOSTRACIÓN. *Suficiencia:* Supongamos que la sucesión $\{a_i\}_{i=1}^{\infty}$ satisface la propiedad de Cauchy. En particular, si tomamos $r = 1$ existe una etiqueta N_1 tal que si $m, n \geqslant N_1$ se cumple

$$|a_m - a_n| < 1,$$

es decir, en particular

$$|a_m - a_{N_1}| < 1 \text{ si } m > N_1,$$

lo que implica que

$$|a_m| < 1 + |a_{N_1}| \text{ si } m > N_1,$$

y entonces el número

$$M = \text{máx}\,\{|a_1|, |a_2|, \ldots, |a_{N_1-1}|, 1 + |a_{N_1}|\}$$

será tal que

$$-M \leqslant a_i \leqslant M \quad \text{para todo } i = 1, 2, \ldots,$$

lo cual significa que la sucesión es acotada y todos sus elementos pertenecen al intervalo cerrado $[-M, M]$. Luego, en virtud de la proposición 4.7, posee una subsucesión $\{a_{m_i}\}_{i=1}^{\infty}$ convergente, digamos a un número L. Probaremos ahora que la sucesión $\{a_i\}_{i=1}^{\infty}$ converge a L. Para ello consideremos el intervalo arbitrario $(L-r, L+r)$. En primer lugar, dado que la subsucesión $\{a_{m_i}\}_{i=1}^{\infty}$ converge a L, existirá una etiqueta N_r tal que si $m_i > N_r$, se tiene que

$$|a_{m_i} - L| < \frac{r}{2}.$$

[2]Augustin Louis Cauchy, mencionado en el capítulo primero, quien fue pionero en el estudio de la convergencia y la divergencia de series infinitas; hizo aportaciones en ecuaciones diferenciales y física matemática, entre otras áreas.

Por otro lado, como la sucesión $\{a_i\}_{i=1}^{\infty}$ es de Cauchy, existe una etiqueta W tal que si j y k son etiquetas con $j > W$ y $k > W$ entonces

$$|a_j - a_k| < \frac{r}{2}.$$

Tomando entonces la etiqueta Q como la mayor de las etiquetas m_W y N_r, tenemos que la distancia de todos los elementos de la sucesión $\{a_i\}_{i=1}^{\infty}$ con etiqueta $s > Q$ al número L es tal que

$$\begin{aligned} |a_s - L| &\leqslant |a_s - a_{m_W} + a_{m_W} - L| \\ &\leqslant |a_s - a_{m_W}| + |a_{m_W} - L| < \frac{r}{2} + \frac{r}{2} = r \end{aligned}$$

ya que

$$|a_s - a_{m_W}| < \frac{r}{2},$$

pues s y m_W son mayores que Q, la cual es una etiqueta mayor que N_r y

$$|a_{m_W} - L| < \frac{r}{2},$$

pues $m_W > W$. Luego, los elementos de la sucesión de Cauchy $\{a_i\}_{i=1}^{\infty}$ caen en $(L - r, L + r)$ a partir de la etiqueta Q, y habiéndose tomado $r > 0$ arbitrario, se tiene entonces que

$$\lim_{i \to \infty} a_i = L.$$

Necesidad: Esto quiere decir que si la sucesión $\{a_i\}_{i=1}^{\infty}$ es convergente, necesariamente es de Cauchy. Esto es más fácil de probar ya que, si $r > 0$, por ser $\{a_i\}_{i=1}^{\infty}$ convergente, digamos a un número L, existe una etiqueta N_r tal que

$$|a_i - L| < \frac{r}{2} \text{ si } i > N_r.$$

Luego si j y k son dos etiquetas mayores que N_r se tendrá

$$|a_j - a_k| \leqslant |a_j - L| + |a_k - L| < \frac{r}{2} + \frac{r}{2} = r,$$

lo cual nos dice que la sucesión $\{a_i\}_{i=1}^{\infty}$ tiene la propiedad de Cauchy. ■

4.4 Límite de una función en un punto

Definición 4.4 *Sea $f(x)$ una función real definida en un intervalo abierto (a,b) y x_0 un punto del intervalo cerrado $[a,b]$. Diremos que el* ***límite*** [3] ***de la función*** *$f(x)$* ***cuando la variable*** *x* ***tiende a*** *x_0* ***tiene el valor*** *L si para cada sucesión $\{x_i\}_{i=1}^{\infty}$ de elementos de (a,b), distintos de x_0, y que converge a x_0, se tiene que la sucesión $\{f(x_i)\}_{i=1}^{\infty}$ converge a L. Simbólicamente, denotaremos lo anterior escribiendo*

$$\lim_{x\to x_0} f(x) = L.$$

En el caso de funciones $f(x)$ definidas en intervalos de la forma (a,∞) se dirá que L es *el límite de $f(x)$ cuando la variable x tiende a ∞*, si para cada sucesión $\{x_i\}_{i=1}^{\infty}$ de elementos de (a,∞) creciente y no acotada se tiene que la sucesión $\{f(x_i)\}_{i=1}^{\infty}$ converge a L. Tal caso se denota

$$\lim_{x\to\infty} f(x) = L.$$

En lenguaje común, podemos decir que L es el límite de una función $f(x)$ cuando la variable independiente tiende al valor x_0, si al tomar valores de la variable independiente x cada vez más cercanos a x_0, los valores correspondientes $y = f(x)$ de la variable dependiente bajo la función son "cada vez más cercanos a L". A la expresión "cada vez más cercanos" le hemos dado significado con el concepto de sucesión convergente y con la condición de que la aproximación al punto x_0 se haga mediante puntos distintos de él.

NOTA IMPORTANTE:

1. *De acuerdo a la definición 4.4, el punto x_0 donde se define el límite de la función, no tiene que estar necesariamente en el dominio de la función; lo importante es que x_0 sea límite de puntos del dominio de la función. Por ejemplo, se puede hablar del límite de la función en los puntos extremos a y b, aunque la función sólo esté definida en (a,b).*

2. *Cuando se dice que la variable independiente x toma valores cada vez más cercanos al valor x_0, se están considerando las dos posibilidades: que x se aproxime a x_0 por la izquierda y que se aproxime por la derecha. Si $\{x_n\}_{n=0}^{\infty}$ es una sucesión que se aproxima a x_0 por la izquierda (es decir, $x_n - x_0 < 0$*

[3] Una definición equivalente fué dada alrededor de 1850 por Karl Weierstrass, llamada *citerio ε-δ para la existencia del límite*, y que enunciamos a continuación: $\lim_{x\to x_0} f(x) = L$ si para cada $\varepsilon > 0$ existe $\delta(\varepsilon)$ tal que si $0 < |x - x_0| < \delta$ entonces $|f(x) - L| < \varepsilon$.

para $n = 1, 2, \ldots$*) y la sucesión* $\{f(x_n)\}_{n=1}^{\infty}$ *converge al número* L^-*, entonces escribimos*

$$\lim_{x \to x_0^-} f(x) = L^-.$$

Si $\{z_n\}_{n=0}^{\infty}$ *es una sucesión que se aproxima a* x_0 *por la derecha (es decir,* $z_n - x_0 > 0$ *para* $n = 1, 2, \ldots$*) y la sucesión* $\{f(z_n)\}_{n=1}^{\infty}$ *converge al número* L^+*, entonces escribimos*

$$\lim_{x \to x_0^+} f(x) = L^+.$$

Los números L^- *y* L^+ *se llaman límite por la izquierda y límite por la derecha, respectivamente, de* $f(x)$ *en* x_0*.*

3. *Para que el límite de una función no exista en un punto* x_0*, es suficiente que* $L^+ \neq L^-$*, o que uno de estos límites laterales no exista, o que ninguno exista, como se ilustra en el ejemplo 4.7.*

Ejemplo 4.7 Consideremos las funciones $f_i : \mathbb{R} \to \mathbb{R}$, $i = 1, 2, 3$, definidas mediante

$$
\begin{aligned}
f_1(x) &= \begin{cases} 1 & \text{si } x \geqslant 0 \\ -1 & \text{si } x < 0 \end{cases} \\
f_2(x) &= \begin{cases} 1/x & \text{si } x > 0 \\ -1 & \text{si } x \leqslant 0 \end{cases} \\
f_3(x) &= \begin{cases} 1/x & \text{si } x \neq 0 \\ 0 & \text{si } x = 0 \end{cases}
\end{aligned}
$$

cuya gráficas aparecen en la figura 4.1.

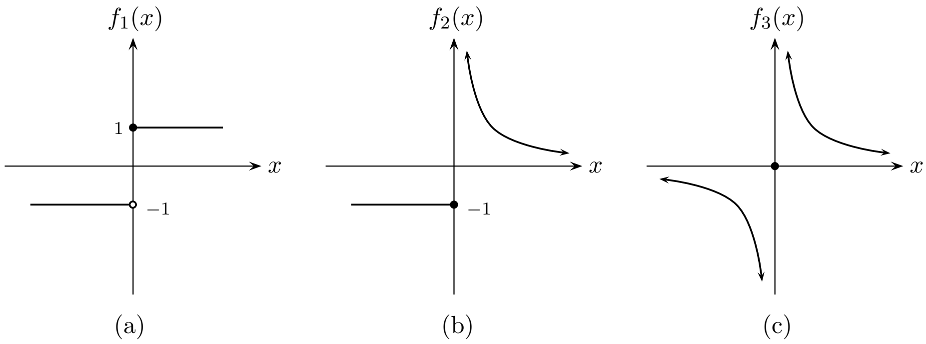


Figura 4.1 Funciones del ejemplo 4.7

Tomemos las sucesiones

$$\{x_i\}_{i=1}^{\infty} = \left\{\frac{1}{i}\right\}_{i=1}^{\infty}, \qquad \{z_i\}_{i=1}^{\infty} = \left\{\frac{-1}{i}\right\}_{i=1}^{\infty}.$$

Nótese que $x_i \to 0^+$ y $z_i \to 0^-$ si $i \to \infty$.

Calculamos las sucesiones correspondientes definidas con los valores de cada una de las funciones y obtenemos lo siguiente:

1. En el primer caso, $\{f_1(x_i)\}_{i=1}^{\infty} = \{1\}_{i=1}^{\infty} \to 1 = L^+$ si $i \to \infty$;
   $\{f_1(z_i)\}_{i=1}^{\infty} = \{-1\}_{i=1}^{\infty} \to -1 = L^-$ si $i \to \infty$.
   Como $L^+ \neq L^-$, se concluye que $\lim\limits_{x\to 0} f_1(x)$ no existe.

2. En el segundo caso, $\{f_2(x_i)\}_{i=1}^{\infty} = \left\{\dfrac{1}{x_i}\right\}_{i=1}^{\infty}$ no converge (es decir, L^+ no existe), mientras que $\{f_1(z_i)\}_{i=1}^{\infty} = \{-1\}_{i=1}^{\infty}$ converge a -1 ($= L^-$) cuando $i \to \infty$. Entonces $\lim\limits_{x\to 0} f_2(x)$ no existe.

3. En el tercer caso, las sucesiones $\{f_3(x_i)\}_{i=1}^{\infty} = \left\{\dfrac{1}{x_i}\right\}_{i=1}^{\infty}$ y $\{f_3(z_i)\}_{i=1}^{\infty} = \left\{\dfrac{1}{z_i}\right\}_{i=1}^{\infty}$ no convergen. Entonces $\lim\limits_{x\to 0} f_3(x)$ no existe. $\triangleleft$

Ejemplo 4.8 La función $f(x) = x^2 + 2x + 3$ definida en $\mathbb{R}$ tiene por límite en el punto $x_0 = 2$ al valor 11 ya que si $\{x_i\}_{i=1}^{\infty}$ converge a 2, la sucesión $\{f(x_i)\}_{i=1}^{\infty}$ es la sucesión

$$\{f(x_i)\}_{i=1}^{\infty} = \left\{x_i^2 + 2x_i + 3)\right\}_{i=1}^{\infty} = \left\{x_i^2\right\}_{i=1}^{\infty} + \{2x_i\}_{i=1}^{\infty} + \{3\}_{i=1}^{\infty},$$

la cual converge al valor 11 en virtud del teorema 4.6. $\triangleleft$

Ejemplo 4.9 Límites trigonométricos básicos:

$$(a) \quad \lim_{x\to 0} \operatorname{sen} x = 0 \qquad\qquad (b) \quad \lim_{x\to 0} \cos x = 1$$

$$(c) \quad \lim_{x\to 0} \frac{\operatorname{sen} x}{x} = 1 \qquad\qquad (d) \quad \lim_{x\to 0} \frac{1 - \cos x}{x} = 1.$$

Para probar estas igualdades, considérese la figura 4.2.

(a) Por definición, las coordenadas del punto A (sobre el círculo unitario) son $(\cos x, \operatorname{sen} x)$, donde x es la longitud del arco AD. Se observa que el segmento $AB = \operatorname{sen} x$ no es mayor que x, la longitud del arco AD, para valores pequeños de x. Por tanto

$$-x \leqslant \operatorname{sen} x \leqslant x,$$

y

$$\lim_{x\to 0^+} \operatorname{sen} x = 0$$

(ver ejercicio 5). Para valores pequeños y negativos de x tenemos:

$$x \leqslant \operatorname{sen} x \leqslant -x,$$

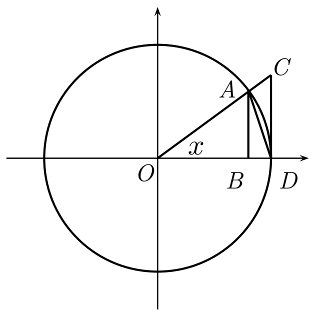


Figura 4.2 Límites trigonométricos del ejemplo 4.9

y

$$\lim_{x\to 0^-} \operatorname{sen} x = 0.$$

Por tanto,

$$\lim_{x\to 0} \operatorname{sen} x = 0.$$

(b) Nótese que el punto B se aproxima al punto D cuando x se aproxima a 0 (por la izquierda o por la derecha). Luego,

$$\lim_{x\to 0} \cos x = 1.$$

(c) La siguiente desigualdad es obvia para valores pequeños de x :

área del triángulo $OAB \leqslant$ área del sector $OAD \leqslant$ área del triángulo OCD,

o

$$\frac{1}{2}\cos x \operatorname{sen} x \leqslant \frac{x}{2} \leqslant \frac{1}{2}\frac{\operatorname{sen} x}{\cos x}. \tag{4.5}$$

Supongamos que x es positivo. Multiplicando cada parte de la desigualdad (4.5) por $2/\operatorname{sen} x$ obtenemos

$$\cos x \leqslant \frac{x}{\operatorname{sen} x} \leqslant \frac{1}{\cos x}.$$

Entonces, en el límite tenemos

$$\lim_{x\to 0^+} \frac{x}{\operatorname{sen} x} = 1,$$

y, al tomar los recíprocos,

$$\lim_{x\to 0^+} \frac{\operatorname{sen} x}{x} = 1.$$

Como

$$\frac{\operatorname{sen}(-x)}{-x} = \frac{\operatorname{sen} x}{x},$$

se deduce que
$$\lim_{x\to 0^-} \frac{\operatorname{sen} x}{x} = 1.$$
Por tanto
$$\lim_{x\to 0} \frac{\operatorname{sen} x}{x} = 1.$$

(d)

$$\begin{aligned}\lim_{x\to 0} \frac{1-\cos x}{x} &= \lim_{x\to 0} \frac{(1-\cos x)(1+\cos x)}{x(1+\cos x)} \\ &= \lim_{x\to 0} \frac{1-\cos^2 x}{x}\frac{1}{1+cosx} \\ &= \lim_{x\to 0} \frac{\operatorname{sen} x}{x}\frac{\operatorname{sen} x}{1+\cos x} \\ &= 1\cdot\frac{0}{2} = 0.\end{aligned}$$ ◁

De la definición de límite de una función y las propiedades de las sucesiones convergentes respecto de las operaciones de suma y producto, se sigue el teorema siguiente, que justifica el penúltimo paso de la demostración anterior.

Teorema 4.11 *Sean $f:(a,b)\to\mathbb{R}$ y $g:(a,b)\to\mathbb{R}$ funciones reales y $x_0\in[a,b]$.*

1. *Si*
$$\lim_{x\to x_0} f(x) = L \quad y \quad \lim_{x\to x_0} g(x) = M,$$
entonces,
$$\lim_{x\to x_0} (f+g)(x) = L+M \quad y \quad \lim_{x\to x_0} (f\cdot g)(x) = LM.$$

2. *Si $\lim_{x\to x_0} f(x) = L$ y $L\neq 0$,*
$$\lim_{x\to x_0} \frac{1}{f(x)} = \frac{1}{L}.$$

4.5 Continuidad de funciones

El concepto de límite que hemos presentado nos permite definir en términos precisos la noción de continuidad en un punto para una función real de variable real. La continuidad de una función es una propiedad de carácter local, en tanto se refiere al valor en un punto y al comportamiento de la función alrededor de ese punto. En lenguaje llano, una función real de variable real $y = f(x)$ se dice *continua* en un punto x_0 de su dominio, si en puntos x cercanos a x_0 los valores de la función $y = f(x)$ son cercanos a $f(x_0)$. En términos del concepto de límite, la definición de continuidad se escribe así:

Definición 4.5 *Una función real $y = f(x)$ definida en un intervalo I se dice* ***continua en un punto*** *$x_0 \in I$ si para cada sucesión $\{x_i\}_{i=1}^{\infty}$ de elementos de I convergente a x_0, la sucesión $\{f(x_i)\}_{i=1}^{\infty}$ es una sucesión convergente a $f(x_0)$. Es decir,*

$$\lim_{x \to x_0} f(x) = f(x_0).$$

Si la función $f(x)$ es continua en cada uno de los puntos de su dominio I, se dice que la función es ***continua en el intervalo*** *I.*

Nota Importante:
La continuidad en un punto x_0 del dominio de una función puede no existir, ya sea porque el límite de la función en ese punto no exista (ver ejemplo 4.10), o sí exista, pero sea distinto del valor $f(x_0)$ de la función en el punto (ver ejemplo 4.11).

Ejemplo 4.10 La función $f : [-1, 1] \to \mathbb{R}$ definida por

$$f(x) = \begin{cases} \dfrac{1}{x} & \text{si} \quad x \neq 0 \\ 0 & \text{si} \quad x = 0 \end{cases}$$

es continua en $x_0 \neq 0$ y no es continua en $x_0 = 0$ pues el $\lim_{x \to 0} f(x)$ no existe. (La gráfica de $f(x)$ aparece en la figura 4.1(c).) $\triangleleft$

Ejemplo 4.11 La función $f : (-1, 1) \to \mathbb{R}$ definida mediante

$$f(x) = \begin{cases} |x| & \text{si} \quad x \neq 0 \\ -1 & \text{si} \quad x = 0, \end{cases}$$

es continua en todo su dominio salvo en $x_0 = 0$ ya que $\lim_{x \to 0} f(x) = 0 \neq f(0) = -1$. $\triangleleft$

Ejemplo 4.12 La función $f : \mathbb{R} \to \mathbb{R}$ definida mediante

$$f(x) = \begin{cases} 1 & \text{si} \quad x \text{ es racional} \\ 0 & \text{si} \quad x \text{ es irracional} \end{cases}$$

no es continua en punto alguno de $\mathbb{R}$. $\triangleleft$

Enseguida probaremos que la propiedad de continuidad en un punto se preserva bajo las operaciones de suma, producto y composición de funciones.

Teorema 4.12 *Para un intervalo I, sean $f : I \to \mathbb{R}$ y $g : I \to \mathbb{R}$ funciones reales y $x_0 \in I$.*

1. *Si f y g son continuas en x_0, entonces las funciones suma y producto $f + g$ y $f \cdot g$ son continuas en x_0.*
2. *Si $h : J \to \mathbb{R}$, $f(I) \subset J$, f es continua en x_0 y h es continua en $f(x_0)$, entonces la función composición*

$$h \circ f : (a, b) \to \mathbb{R}$$

es continua en x_0.

Demostración. La prueba de a) se sigue directamente de las propiedades de los límites de funciones respecto de la suma y producto de funciones. Para probar b), tomemos una sucesión $\{x_i\}_{i=1}^{\infty}$ de elementos de I convergente a x_0. Como f es continua en x_0, se tiene que $\{f(x_i)\}_{i=1}^{\infty}$ converge a $f(x_0)$ y en vista de que h es una función continua en $f(x_0)$, tendremos que $\{h(f(x_i))\}_{i=1}^{\infty} = \{(h \circ f)(x_i)\}_{i=1}^{\infty}$ convergerá a $h(f(x_0)) = (h \circ f)(x_0)$. Esto prueba que siempre que una sucesión converja a x_0, la sucesión formada por sus imágenes bajo la función composición $h \circ f$ convergerá a $(h \circ f)(x_0)$, es decir $h \circ f$ es continua en x_0. ■

Una familia muy importante de funciones continuas son las llamadas funciones *lipschitzianas,*[4] que presentamos a continuación.

Definición* 4.6** *Una función real $f : A \subset \mathbb{R} \to \mathbb{R}$ se dice* ***lipschitziana *si existe $L > 0$ tal que*

$$|f(x) - f(y)| \leqslant L|x - y|,$$

para cualesquiera $x, y \in A$.

4.6 Continuidad en intervalos compactos

Concluimos este capítulo, enlistando y probando las propiedades principales de las funciones continuas definidas en un *intervalo compacto*, es decir, en un intervalo cerrado y acotado. La prueba de estas propiedades se basa generalmente en el principio de continuidad de la recta real y en la propiedad que tiene todo intervalo cerrado y acotado de contener el límite de cada sucesión convergente formada de

[4] Llamadas así en honor de Rudolf O. S. Lipschitz (1832-1903), matemático alemán, quien descubrió que esta propiedad garantiza la unicidad de la solución de la ecuación diferencial $y' = f(x, y)$.

puntos de ese intervalo.

Teorema 4.13 *Sea $f(x) : [a,b] \to \mathbb{R}$ una función continua en $[a,b]$. Entonces:*

1. *La función $f(x)$ es una función acotada en $[a,b]$, es decir existe $M > 0$ tal que*
$$|f(x)| \leqslant M \text{ para todo } x \in [a,b].$$
2. *La función $f(x)$ alcanza en $[a,b]$ su supremum y su infimum, es decir, existen x_1 y $x_2 \in [a,b]$ tales que $f(x_1) \geqslant f(x)$ y $f(x_2) \leqslant f(x)$ para toda $x \in [a,b]$.*
3. *(Propiedad del valor intermedio) Si $f(a) > 0$ y $f(b) < 0$ entonces existe $x_c \in (a,b)$ tal que $f(x_c) = 0$.*
4. *La función $f(x)$ es **uniformemente continua** en $[a,b]$, es decir para cada $r > 0$ existe $\delta > 0$ tal que siempre que $x, z \in [a,b]$ con $|x - z| < \delta$, se cumple que $|f(x) - f(z)| < r$.*

Demostración. Probaremos el punto 1 por contradicción. Supongamos que existiese una función $f(x)$ continua en un intervalo cerrado $[a,b]$ y ésta no fuera acotada; esto significaría que para cada número natural n, existiría un punto $c_n \in [a,b]$ tal que $|f(c_n)| > n$. Siendo $\{c_n\}_{n=1}^{\infty}$ una sucesión acotada, en virtud de la Proposición 4.7, podemos extraer de $\{c_n\}_{n=1}^{\infty}$ una subsucesión $\left\{c_{m(n)}\right\}_{n=1}^{\infty}$ convergente a un punto $c \in [a,b]$. Por otro lado, por construcción, la sucesión $\left\{f(c_{m(n)})\right\}_{n=1}^{\infty}$ no es convergente ya que $|f(c_{m(n)})| > m(n)$, lo cual da lugar a una contradicción. Luego, la suposición que hicimos al principio es falsa y entonces toda función continua en $[a,b]$ es acotada.

Para demostrar la validez del punto 2, denotemos por
$$S = \sup\{f(x) \text{ para } x \in [a,b]\}.$$
Por la definición de supremum, existe una sucesión creciente $\{f(x_i)\}_{i=1}^{\infty}$ convergente a S. Consideremos ahora la sucesión $\{x_i\}_{i=1}^{\infty}$. Esta sucesión es un conjunto infinito de puntos de $[a,b]$ y por el mismo argumento que usamos en la prueba del punto 1, podremos extraer de ella una subsucesión $\left\{x_{m(i)}\right\}_{i=1}^{\infty}$ convergente a un punto $x = c$. Siendo la función continua en c, tendremos que $\lim_{i\to\infty} f(x_{m(i)}) = f(c)$. Pero $\lim_{i\to\infty} f(x_{m(i)}) = \lim_{i\to\infty} f(x_i) = S$ por ser $\left\{x_{m(i)}\right\}_{i=1}^{\infty}$ subsucesión de $\{x_i\}_{i=1}^{\infty}$. Luego, $f(c) = S$, es decir el supremum S se alcanza en el punto $x = c$.

Para probar el punto 3, realicemos el proceso siguiente: en un primer paso, determinemos el punto medio $m_1 = (a+b)/2$ de $[a,b]$. Si $f(m_1) = 0$, habremos probado el enunciado del teorema. Si $f(m_1) > 0$, consideramos el intervalo $[m_1, b] \subset [a,b]$, y si $f(m_1) < 0$, entonces consideraremos el intervalo $[a, m_1] \subset [a,b]$. Supongamos que $f(m_1) > 0$, en tal caso, determinemos enseguida el punto medio $m_2 = (b + m_1)/2$

de $[m_1, b]$ y evaluemos $f(m_2)$. Si $f(m_2) = 0$ habremos probado el enunciado 3. Si por el contrario, $f(m_2) > 0$, consideramos el intervalo $[m_2, b] \subset [m_1, b] \subset [a, b]$, y si $f(m_2) < 0$ entonces consideramos el intervalo $[a, m_2] \subset [a, m_1] \subset [a, b]$ y volvemos a tomar el punto medio m_3 y repetimos el proceso. Finalmente, habremos construido una sucesión I_n de intervalos cerrados anidados $I_n \subset I_{n-1} \subset I_{n-2} \subset \cdots \subset [a, b]$ de longitud $(b-a)/2^n$ y tales que el valor de $f(x)$ en el extremo izquierdo de cada intervalo es positivo y el valor de $f(x)$ en cada extremo derecho es negativo. Teniendo en cuenta que la sucesión $\{a_n\}$ formada con los extremos izquierdos de I_n y la sucesión $\{b_n\}$ formada con los extremos derechos de los intervalos I_n son sucesiones monótonas con $b_n - a_n < (b-a)/2^n$, ambas sucesiones deberán converger a un mismo punto c,

$$c = \lim_{n\to\infty} a_n = \lim_{n\to\infty} b_n$$

y por la continuidad de f en c deberemos tener

$$f(c) = \lim_{n\to\infty} f(a_n) = \lim_{n\to\infty} f(b_n).$$

Como $f(a_n) > 0$ y $f(b_n) < 0$ para $n = 1, 2, \ldots$, concluimos que simultáneamente se tendrá que $f(c) \geqslant 0$ y $f(c) \leqslant 0$ y por lo tanto $f(c) = 0$, con lo que se prueba el punto 3.

La prueba de la continuidad uniforme la daremos también por contradicción. Supongamos así que lo afirmado en el punto 4 no es cierto. Esto quiere decir que existirá un número $\varepsilon_0 > 0$ y dos sucesiones $\{a_i\}_{i=1}^{\infty}$y $\{b_i\}_{i=1}^{\infty}$ de elementos de $[a, b]$ tales que

$$b_i > a_i,\ b_i - a_i < \frac{1}{i},\ \text{y}\ \ |f(b_i) - f(a_i)| > \varepsilon_0 \text{ para } i = 1, 2, \ldots$$

Repitiendo el argumento que utilizamos al probar el punto 1, podemos afirmar la existencia de una sucesión formada con elementos de la forma $a_{m(i)}$ con $m(1) < m(2) < \cdots m(k) < \cdots$, la cual es convergente a algún número $c \in [a, b]$. Fijándonos ahora en la sucesión correspondiente $b_{m(i)}$, podemos asegurar que también convergerá al mismo número c ya que $b_{m(i)} - a_{m(i)} < 1/m(i)$. Siendo $f(x)$ continua en c deberemos tener que $\lim_{i\to\infty} f(a_{m(i)}) = \lim_{i\to\infty} f(b_{m(i)}) = f(c)$, lo cual contradice el hecho de que $|f(b_{m(i)}) - f(a_{m(i)})| > \varepsilon_0$ para toda $i = 1, 2, \ldots$ que habíamos supuesto verdadero para las sucesiones $\{a_i\}_{i=1}^{\infty}$y $\{b_i\}_{i=1}^{\infty}$. Por lo tanto la suposición hecha es falsa y lo afirmado por el Teorema en el punto 4, es verdadero. ■

El resultado siguiente es consecuencia de la propiedad del valor intermedio para funciones continuas.

Corolario 4.14 *Si $f : [a, b] \to \mathbb{R}$ es continua e inyectiva, entonces f es monótona.*

Demostración. Supongamos que f no es monótona en $[a, b]$. Entonces existen $a \leqslant x_1 < x_2 < x_3 \leqslant b$ tales que $f(x_1) < f(x_2)$ y $f(x_2) > f(x_3)$, o $f(x_1) > f(x_2)$

y $f(x_2) < f(x_3)$. Supongamos que se da el primer caso. Como no es posible que $f(x_3) = f(x_1)$, se tendrá que $f(x_3) > f(x_1)$, o $f(x_3) < f(x_1)$. En el primer caso podemos asegurar la existencia de un número c tal que $f(x_1) < f(x_3) < c < f(x_2)$. Consideremos la función continua $h(x) = f(x) - c$. Nótese que $h(x_1) < 0$, $h(x_2) > 0$ y $h(x_3) < 0$. Por la propiedad del valor intermedio, existirán números $y_c \in (x_1, x_2)$ y $z_c \in (x_2, x_3)$ tales que $h(y_c) = h(z_c)$. Pero esto implicaría que $f(y_c) = f(z_c)$, contrario a la suposición de que f es inyectiva. ■

Corolario 4.15 *Si una función continua f:$[a, b] \to \mathbb{R}$ tiene inversa, ésta es continua.*

Nota Importante:

1. *En el teorema 4.13, la condición de que el dominio de continuidad de la función sea un intervalo cerrado es imprescindible. Basta considerar como contraejemplo la función $f : (0, 1) \to \mathbb{R}$ definida mediante*

$$f(x) = \frac{1}{x},$$

que obviamente ni es acotada ni es uniformemente continua en $(0, 1)$.

2. *La propiedad del valor intermedio da lugar al llamado "método de bisección" para determinar soluciones a ecuaciones de la forma $f(x) = 0$ con $f : \mathbb{R} \to \mathbb{R}$ una función continua. El método proporciona soluciones con el grado de aproximación que se desee y consiste de los siguientes pasos:*

   1ro. *Encuentre $a, b \in \mathbb{R}$ con $a < b$ y tales que $f(a) > 0$ y $f(b) < 0$. En virtud de la propiedad del valor intermedio, existirá una solución a $f(x) = 0$ en $[a, b]$.*

   2do. *Tome el punto medio $(a + b)/2$ del intervalo $[a, b]$ y considere el valor $f((a + b)/2)$. Si $f((a + b)/2) = 0$ habremos encontrado una solución a la ecuación. Si $f((a + b)/2) > 0$, se considera el intervalo $[(a + b)/2, b]$ y por el mismo argumento anterior, existirá en $[(a + b)/2, b]$ una solución para la ecuación. En caso contrario, si $f((a + b)/2) < 0$, existirá solución en el intervalo $[a, (a + b)/2]$. Supongamos que se tiene $f((a + b)/2) < 0$.*

   3ro. *Tome el punto medio $(3a + b)/4$ del intervalo $[a, (a + b)/2]$ y evalúe la función en ese punto, $f((3a + b)/4)$ y, según sea su signo, tome el intervalo $[a, (3a + b)/4]$ si $f((3a + b)/4) < 0$ o el intervalo $[(3a + b)/4, (a + b)/2]$ si $f((3a + b)/4) > 0$ y en el nuevo intervalo, que es de longitud $(b - a)/8$, deberá existir una solución. Prosiguiendo con este proceso, en el n-ésimo paso se tendrá situada una solución en un intervalo de longitud $(b-a)/2^n$, lo que nos permite conocer una solución de la ecuación con un error menor que $(b - a)/2^n$.*

Ejemplo 4.13 Encuentre una solución a la ecuación $x^3 - 15x + 1 = 0$ con aproximación de $1/100$.

Solución:

1ro. Tomando $a = 0$ y $b = 1$ tenemos $f(0) = 1 > 0$ y $f(1) = -13 < 0$ y siendo $f(x) = x^3 - 15x + 1$ continua, existirá una solución a la ecuación en el intervalo $I_1 = [0, 1]$. Este primer paso, nos determina el número de etapas que serán necesarias para obtener una solución a la ecuación con un error pedido. En nuestro caso, deberemos de llevar al cabo al menos 8 bisecciones para tener un error menor que $\frac{1}{2^8} < \frac{1}{100}$.

2do. El punto medio de I_1 es el punto $x = \frac{1}{2}$ y $f(\frac{1}{2}) < 0$; luego, habrá una solución en el intervalo $I_2 = [0, \frac{1}{2}]$.

3ro. El punto medio de I_2 es el punto $x = \frac{1}{4}$ y $f(\frac{1}{4}) < 0$; así, la solución se halla en $I_3 = [0, \frac{1}{4}]$

4to. El punto medio de I_3 es $x = \frac{1}{8}$ y $f(\frac{1}{8}) < 0$; luego, la solución se encuentra en $I_4 = [0, \frac{1}{8}]$.

5to. El punto medio de I_4 es $x = \frac{1}{16}$ y $f(\frac{1}{16}) > 0$, y entonces la solución se localiza en el intervalo $I_5 = [\frac{1}{16}, \frac{1}{8}]$.

6to. El punto medio de I_5 es $x = \frac{3}{32}$ y $f(\frac{3}{32}) < 0$, y la solución se encuentra en $I_6 = [\frac{1}{16}, \frac{3}{32}]$.

7mo. El punto medio de I_6 es $x = \frac{5}{64}$ y $f(\frac{5}{64}) < 0$; luego, la solución se ubica en $I_7 = [\frac{1}{16}, \frac{5}{64}]$

8vo. El punto medio de I_7 es $x = \frac{9}{128}$ con $f(\frac{9}{128}) < 0$ y tendremos una solución en $[\frac{8}{128}, \frac{9}{128}]$.

De lo anterior concluimos que el valor $x = \frac{17}{256}$ es un número que satisface la ecuación con un error menor que $1/2^8$, es decir, menor que $1/100$. $\triangleleft$

Ejercicios y problemas del capítulo

1. Para cada una de las sucesiones siguientes, escriba sus elementos en una lista, ordenados hacia la derecha.

   (a) $\left\{\dfrac{i+1}{i^2+2}\right\}_{i=1}^{\infty}$

   (b) $\left\{(-1)^i\right\}_{i=1}^{\infty}$

   (c) $\{\operatorname{sen}(\pi i)\}_{i=1}^{\infty}$

   (d) $\{a_i\}_{i=1}^{\infty}$ donde $a_i = 0$ si i es impar y $a_i = (\frac{i}{2})^2$ si i es par.

2. Explique por qué la sucesión $\left\{\operatorname{sen}(i^2+2)\right\}_{i=1}^{\infty}$ es subsucesión de $\{\operatorname{sen} i\}_{i=1}^{\infty}$ y diga qué lugar ocupa en la sucesión el octavo elemento de la subsucesión.

3. Escriba los primeros tres elementos de la subsucesión de la sucesión $\left\{\operatorname{sen}(i^2+i)\right\}_{i=1}^{\infty}$ generada por la elección creciente de etiquetas dada por la sucesión $\left\{i^3\right\}_{i=1}^{\infty}$. ¿Cuál es el octavo término de la subsucesión y qué lugar ocupaba en la sucesión? ¿Cuál es el j-ésimo elemento de la subsucesión?

4. Escriba la definición de sucesión no-convergente.

5. (a) Si $\{a_n\}_{n=1}^{\infty}$ es una sucesión convergente de términos no negativos (es decir, $a_n \geqslant 0$), demuestre que su límite es no negativo.

   (b) Demuestre que si $\{a_n\}_{n=1}^{\infty}, \{b_n\}_{n=1}^{\infty}$ y $\{c_n\}_{n=1}^{\infty}$ son sucesiones convergentes, y si $a_n \leqslant b_n \leqslant c_n$ para toda $n = 1, 2, \ldots,$ entonces $\lim\limits_{n\to\infty} a_n \leqslant \lim\limits_{n\to\infty} b_n \leqslant \lim\limits_{n\to\infty} c_n$. Deduzca que si $\lim\limits_{n\to\infty} a_n = \lim\limits_{n\to\infty} c_n = L$ entonces $\lim\limits_{n\to\infty} b_n = L$.

6. Considere la sucesión $\{a_i\}_{i=1}^{\infty} = \left\{\dfrac{2i+3}{i+5}\right\}_{i=1}^{\infty}$. Encuentre una etiqueta a partir de la cual los elementos de esa sucesión disten de $L = 2$ en menos que $1/10^3$. Pruebe que $\lim\limits_{i\to\infty} a_i = 2$.

7. Demuestre que toda subsucesión de una sucesión convergente es convergente y tienen el mismo límite.

8. Dé un ejemplo de una sucesión no convergente que tenga subsucesiones convergentes y otro que no tenga subsucesiones convergentes.

9. Muestre que la sucesión $\left\{i^3\right\}_{i=1}^{\infty}$ es creciente.

10. (a) Demuestre que la suma y composición de funciones lipschitzianas es lipschitziana.

    (b) Demuestre que si f y g son funciones acotadas y lipschitzianas en $A \subset \mathbb{R}$, entonces su producto es una función lipschitziana.

(c) Demuestre que cada función lipschitziana es continua en todos los puntos de su dominio.

11. Sea $\{a_i\}_{i=1}^{\infty}$ una sucesión real con $\lim_{i\to\infty} a_i = L$. Pruebe que la sucesión $\{|a_i|\}_{i=1}^{\infty}$ converge a $|L|$.

12. Pruebe que si la función $f(x)$ es continua en $x = x_0$, también será continua en ese punto la función $|f(x)|$.

13. Calcule los límites siguientes, si existen, y si no es el caso, diga por qué.

    (a) $\lim_{x\to 0} x \operatorname{sen} \dfrac{1}{x}$

    (b) $\lim_{x\to 0} \operatorname{sen} \dfrac{1}{x}$

    (c) $\lim_{x\to 2} \dfrac{x^2-4}{x-2}$

14. Conteste "falso" o "verdadero".

    Si $\lim_{x\to a} f(x) = L$, entonces:

    (a) f está definida en $x = a$,

    (b) $f(a) = L$,

    (c) f es continua en $x = a$.

    Si f es continua en $\mathbb{R}$, entonces

    (a) $f^2(x)$ es continua en $\mathbb{R}$,

    (b) $f(x+2)$ es continua en $\mathbb{R}$.

    Si $f : (a, b) \to \mathbb{R}$ es una función continua en (a, b), entonces

    (a) f es no acotada,

    (b) f es acotada pero no alcanza en (a,b) su valor máximo.

15. Aplicando la propiedad del valor intermedio para funciones continuas, demuestre que la ecuación $x^3 + x^2 - x - 4 = 0$ tiene una solución en el intervalo $[1, 2]$.

16. Pruebe que si una función continua $f:[a, b] \to \mathbb{R}$ tiene inversa, ésta es continua.

17. Sea $f(x)$ una función continua en $[0, 1]$ y tal que $0 \leqslant f(x) \leqslant 1$ para toda $x \in [0, 1]$. Demuestre que existe $c \in [0, 1]$ tal que $f(c) = c$. (Aplique la propiedad del valor intermedio a la función $g(x) = f(x) - x$.)

Capítulo 5

Medida de la razón de cambio: la derivada

La derivada de una función real de variable real $y = f(x)$ es el concepto que da sentido matemático a la razón de cambio puntual o movimiento instantáneo. Tomando en cuenta que, en una función, a cada variación de la variable independiente con respecto a un valor x_0, corresponde una variación de la variable dependiente con respecto al valor $f(x_0)$, la derivada define la razón de cambio puntual (o instantáneo) en x_0 como el límite de los cocientes de las variaciones de esas variables cuando la variación de la variable independiente tiende a cero.

En el caso de la función de posición de un cuerpo físico con respecto al tiempo, la derivada corresponde a la noción de velocidad instantánea, que así resulta definida como el límite de las velocidades promedio tomadas en intervalos de tiempo cuya duración tiende a cero. Las características de la derivada hacen de ésta el concepto adecuado para la formulación de las leyes dinámicas en las ciencias naturales.

Por otro lado, para la curva en el plano cartesiano que define la gráfica de una función, la derivada es el valor de la pendiente de la recta tangente a la curva en el punto correspondiente, obteniéndose así una interpretación geométrica para la derivada que sienta las bases para el estudio analítico de curvas y superficies.

En este capítulo, a partir de su definición, se deducen las propiedades principales de la derivada y las reglas para su cálculo cuando intervienen las distintas operaciones entre funciones. Se introduce también el concepto de derivada de orden superior y se calcula la función derivada de las principales funciones elementales.

5.1 Definición de derivada

Consideremos una función real de variable real $y = f(x)$ definida en un intervalo abierto (a, b). Sea x_0 un elemento de (a, b). La manera natural de comparar la variación que muestra el valor de la variable dependiente $y = f(x)$ cuando el valor de la variable independiente x experimenta en x_0 una variación $h \neq 0$, es considerar el cociente

$$\frac{f(x_0 + h) - f(x_0)}{h}. \tag{5.1}$$

Al cociente (5.1) lo llamaremos *cociente diferencial de la función en el punto* x_0 *correspondiente a la variación de magnitud* h *de la variable independiente* y lo denotaremos con $\dfrac{\Delta f}{\Delta x}(x_0)(h)$.

Ejemplo 5.1 Para la función

$$y = f(x) = x^2,$$

el cociente diferencial en el punto $x_0 = 2$ correspondiente a la variación de magnitud $h = \frac{1}{10}$, es el número

$$\frac{\Delta f}{\Delta x}(2)\left(\frac{1}{10}\right) = \frac{f\big(2+\frac{1}{10}\big) - f(2)}{\frac{1}{10}} = \frac{(2+\frac{1}{10})^2 - 4}{\frac{1}{10}} = \frac{41}{10}.$$

Análogamente, el cociente diferencial de la función en el punto $x_0 = 2$ correspondiente a una variación de magnitud $h = -2/5$ es el número

$$\frac{\Delta f}{\Delta x}(2)\left(-\frac{2}{5}\right) = \frac{f\big(2-\frac{2}{5}\big) - f(2)}{-\frac{2}{5}} = \frac{(2-\frac{2}{5})^2 - 4}{-\frac{2}{5}} = \frac{36}{10}.$$

En general, para una variación de magnitud h en el punto $x_0 = 2$, el cociente diferencial de la función $y = f(x) = x^2$ toma el valor

$$\frac{\Delta f}{\Delta x}(2)(h) = \frac{(2+h)^2 - (2)^2}{h} = 4 + h. \qquad \triangleleft$$

Ejemplo 5.2 Supongamos que una partícula se mueve a lo largo de una línea recta y que la variable $d = d(t)$ mide su posición en el instante t con respecto a un punto de referencia dado. El cociente diferencial en un tiempo t_0 y para una variación de magnitud h en el tiempo, es la distancia recorrida en el intervalo de tiempo $[t_0, t_0 + h]$ dividida por el tiempo transcurrido (es decir, h):

$$\frac{\Delta d}{\Delta t}(t_0)(h) = \frac{d(t_0+h) - d(t_0)}{h}, \qquad h \neq 0.$$

A este valor se le conoce como *velocidad media de la partícula en el intervalo de tiempo* $[t_0, t_0 + h]$ (o el intervalo $[t_0 + h, t_0]$ si $h < 0$):

$$\text{Velocidad media en } [t_0, t_0 + h] = \frac{\Delta d}{\Delta t}(t_0)(h). \qquad \triangleleft$$

Nota Importante:

La magnitud h de la variación de la variable independiente siempre se toma distinta de cero y el valor del cociente diferencial de una función en un punto x_0 depende del valor de h. En el sentido anterior, el cociente diferencial de una función en el punto x_0 es una función de la magnitud h de la variación de la variable independiente.

A partir del concepto de cociente diferencial en un punto o valor x_0, enunciamos a continuación la definición de *derivada.*

Definición 5.1 *Sea $y = f(x)$ una función real definida en el intervalo (a, b) y x_0 un valor de la variable independiente en (a, b). Se define* ***la derivada de la función*** *$f(x)$* ***en el punto*** *x_0 como el límite (cuando éste exista) del cociente diferencial de la función en x_0 cuando la magnitud h de la variación de la variable independiente tiende a cero. Cuando existe la derivada de la función $y = f(x)$ en un punto x_0 ésta se denota[1] con $\dfrac{\mathrm{d}f}{\mathrm{d}x}(x_0)$; es decir,*

$$\frac{\mathrm{d}f}{\mathrm{d}x}(x_0) = \lim_{h\to 0} \frac{\Delta f}{\Delta x}(x_0)(h) = \lim_{h\to 0} \frac{f(x_0 + h) - f(x_0)}{h},$$

y se dice que la función es ***derivable en*** *x_0.*

Ejemplo 5.3 Para la función $f(x) = x^2$, la derivada en el punto $x_0 = 2$ tiene el valor 4, pues aplicando la definición se obtiene

$$\frac{\mathrm{d}(x^2)}{\mathrm{d}x}(2) = \lim_{h\to 0} \frac{(2 + h)^2 - (2)^2}{h} = \lim_{h\to 0}(4 + h) = 4. \qquad \triangleleft$$

Ejemplo 5.4 La función $f : \mathbb{R} \to \mathbb{R}$, definida por

$$f(x) = \begin{cases} 3x + 5 & \text{si} \quad x < 1 \\ x^2 + x & \text{si} \quad x \geqslant 1, \end{cases}$$

tiene derivada en $x_0 = 1$, pues si $\{h_n\}_{n=1}^{\infty}$ es una sucesión de variaciones cuyas magnitud tiende a cero, la sucesión de cocientes diferenciales toma la forma

$$\begin{aligned} \frac{\Delta f}{\Delta x}(1)(h_n) &= \frac{f(1 + h_n) - f(1)}{h_n} \\ &= \begin{cases} \dfrac{(1 + h_n)^2 + (1 + h_n) - 2}{h_n} = h_n + 3 & \text{si} \quad h_n > 0 \\ \dfrac{3(1 + h_n) + 5 - 8}{h_n} = 3 & \text{si} \quad h_n < 0. \end{cases} \end{aligned}$$

Al tomar el límite cuando $h_n \to 0$, obtenemos que la sucesión de cocientes diferenciales converge a 3. Luego, $\dfrac{\mathrm{d}f}{\mathrm{d}x}(1) = 3$. $\triangleleft$

[1] Esta notación se debe a Leibniz. Otra notación es $\dot{y}(x_0)$, utilizada por Newton y que se reserva actualmente para los casos en los que variable independiente es el tiempo. Una tercera notación es $y'(x_0)$ o $f'(x_0)$, debida a Lagrange, que es muy utilizada en ecuaciones diferenciales.

Nota Importante:
Es posible que una función no posea derivada en algunos de los puntos de su dominio. Por ejemplo, la función $f : \mathbb{R} \to \mathbb{R}$

$$f(x) = |x|$$

no posee derivada en $x = 0$, ya que el cociente diferencial toma la forma

$$\frac{\Delta |x|}{\Delta x}(0)(h) = \begin{cases} \dfrac{|h|}{h} = 1 & si \quad h < 0 \\ \dfrac{|h|}{h} = -1 & si \quad h > 0. \end{cases}$$

Luego, si tomamos una sucesión convergente a cero de variaciones positivas $\{h_n\}_{n=1}^{\infty}$ ($h_n > 0$), la sucesión de los cocientes diferenciales convergerá a 1, mientras que si tomamos una sucesión convergente a cero de variaciones negativas $\{h_n\}_{n=1}^{\infty}$ ($h_n < 0$) la sucesión de los cocientes diferenciales convergerá a -1. Al no existir un límite único cuando las variaciones h tienden a cero, la función no tiene derivada en $x = 0$. En términos de la gráfica de la función alrededor del punto $x = 0$, se observa que esa curva no tiene una recta tangente en el punto (0,0) pues mientras las rectas secantes correspondientes a incrementos positivos de la variable independiente tienen todas pendiente 1, las rectas secantes correspondientes a incrementos negativos de la variable x tienen todas pendiente -1 y no se define una recta tangente única para ambos lados del punto $x = 0$.

Si la función $y = f(x)$ tiene por dominio un intervalo cerrado y acotado $[a, b]$, se dice que es derivable en el punto extremo a si para cada sucesión $\{h_n\}_{n=1}^{\infty}$de variaciones positivas $h_n > 0$, la sucesión correspondiente de cocientes diferenciales converge. A ese valor se le llama *derivada por la derecha en $x = a$*. Análogamente se define la *derivada por la izquierda* en el extremo b del intervalo $[a, b]$. En general, diremos que la función f es derivable en todos los puntos de un intervalo I, si lo es en los puntos interiores y en los extremos posee derivada por la derecha o por la izquierda, según sea el caso.

Una propiedad natural que tiene toda función que es derivable en un punto, es que es continua en ese punto. Probaremos esto último con la proposición siguiente.

Proposición 5.1 *Si $f : (a, b) \to \mathbb{R}$, $y = f(x)$ es una función derivable en x_0, entonces f es una función continua en x_0.*

Demostración. Debemos demostrar que $\lim_{x \to x_0} f(x) = f(x_0)$ o, equivalentemente, que $\lim_{x \to x_0}(f(x) - f(x_0)) = 0$. Tomemos una sucesión $\{x_i\}_{i=1}^{\infty}$ convergente a x_0 con $x_i \neq x_0$ para $i = 1, 2, \ldots,$ y consideremos la sucesión de variaciones $\{h_i\}_{i=1}^{\infty} = \{x_i - x_0\}_{i=1}^{\infty}$, la cual es una sucesión convergente a cero. Entonces, para probar que la función $y = f(x)$ es continua en x_0, debemos mostrar que la sucesión

$\{f(x_0+h_i)-f(x_0)\}_{i=1}^{\infty}$ es convergente a cero. Al escribir

$$\{f(x_0+h_i)-f(x_0)\}_{i=1}^{\infty} = \left\{\frac{f(x_0+h_i)-f(x_0)}{h_i}\right\}_{i=1}^{\infty} \{h_i\}_{i=1}^{\infty}$$

vemos que la sucesión $\left\{\dfrac{f(x_0+h_i)-f(x_0)}{h_i}\right\}_{i=1}^{\infty}$ es convergente a $\dfrac{\mathrm{d}f}{\mathrm{d}x}(x_0)$ ya que, por hipótesis, la función $y=f(x)$ tiene derivada en x_0. Por lo tanto, en el lado derecho tenemos un producto de sucesiones convergentes, de las cuales una de ellas es convergente a cero. Luego, por el teorema 4.6, el producto de las sucesiones será convergente a cero, con lo cual se demuestra que

$$\{f(x_0+h_i)-f(x_0)\}_{i=1}^{\infty} \to 0,$$

y por lo tanto, la función $y=f(x)$ es continua en x_0. ■

Nota Importante:
La proposición anterior afirma que la continuidad en el punto es necesaria para la existencia de la derivada, sin embargo, la continuidad en el punto no es suficiente para asegurar la existencia de la derivada, como lo muestra la función $f(x)=|x|$ en el punto $x=0$, la cual a pesar de ser continua en $x=0$, no es derivable en ese punto.

5.1.1 Interpretación geométrica de la derivada

En términos de la gráfica de la función $y=f(x)$, para cada variación de magnitud h de la variable independiente con respecto al valor inicial x_0, el cociente diferencial $\dfrac{\Delta f}{\Delta x}(x_0)(h)=\dfrac{f(x_0+h)-f(x_0)}{h}$ es la pendiente de la recta secante a la gráfica de la función por los puntos $(x_0,f(x_0))$ y $(x_0+h,f(x_0+h))$ como se observa en la figura 5.1.

Así, la derivada en el punto x_0 es el límite de las pendientes de las rectas secantes cuando el segundo punto $(x_0+h,f(x_0+h))$ sobre la gráfica se toma cada vez más cercano al punto inicial $(x_0,f(x_0))$. En los términos geométricos anteriores, la derivada de $y=f(x)$ en el punto x_0 coincide con la pendiente de la recta tangente a la gráfica de la función en el punto $(x_0,f(x_0))$. Por lo contrario, el que la función $y=f(x)$ no posea derivada en el punto x_0 significa que la curva que define la gráfica de la función no tiene recta tangente en el punto $(x_0,f(x_0))$. Este último es el caso de la gráfica de la función $f(x)=|x|$ en el punto $(0,0)$.

En el caso en el que $d=d(t)$ sea una función de posición, la derivada en un instante t_0 es el número al cual tienden las velocidades medias en intervalos de la forma $[t_0,t_0+h]$, cuando la duración h del intervalo tiende a cero, y se interpreta como la *velocidad instantánea en t_0*. Dicho de otra manera, la velocidad instantánea en el instante t_0 es el límite de las velocidades promedio tomadas sobre intervalos de tiempo alrededor de t_0 con duración cada vez más y más pequeña.

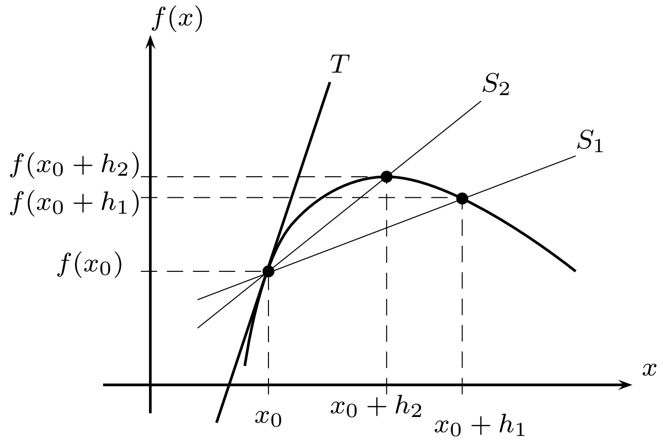


Figura 5.1 Las rectas secantes $S_1, S_2, \ldots,$ tienden a T, la recta tangente a la gráfica en el punto $(x_0, f(x_0))$, cuando $h \to 0$.

5.1.2 Derivada de algunas funciones elementales

A continuación calcularemos la derivada de algunas funciones elementales utilizando directamente la definición 5.1, de derivada en un punto. Después, luego de haber presentado las principales reglas de derivación, aumentaremos la lista de funciones y sus derivadas mediante la aplicación de esas reglas.

En lo que sigue, x_0 es un valor de la variable independiente en el que la función en cuestión es diferenciable y $\{h_i\}_{i=1}^{\infty}$ es una sucesión de variaciones tal que $h_i \neq 0$ para todo $i = 1, 2, \ldots,$ y $\{h_i\}_{i=1}^{\infty} \to 0$.

1. *La derivada en x_0 de la función constante $f(x) = c$ es igual a cero:*

$$\frac{\mathrm{d}c}{\mathrm{d}x}(x_0) = 0. \tag{5.2}$$

   Para probarlo, consideramos la sucesión correspondiente de cocientes diferenciales

$$\frac{\Delta c}{\Delta x}(x_0)(h_i) = \frac{c-c}{h_i} = 0.$$

   Esta sucesión es constante; su valor es cero para $i = 1, 2, \ldots,$ y, por lo tanto, converge a cero. Luego,

$$\frac{\mathrm{d}c}{\mathrm{d}x}(x_0) = 0.$$

2. *La derivada en x_0 de la función identidad $f(x) = x$ es igual a uno:*

$$\frac{\mathrm{d}x}{\mathrm{d}x}(x_0) = 1.$$

La sucesión correspondiente de cocientes diferenciales,

$$\frac{\Delta x}{\Delta x}(x_0)(h_i) = \frac{x_0 + h_i - x_0}{h_i} = \frac{h_i}{h_i} = 1$$

es constante e igual a uno y, por lo tanto, converge a uno. Luego,

$$\frac{\mathrm{d}x}{\mathrm{d}x}(x_0) = 1.$$

3. *La derivada en x_0 de la función $f(x) = x^2$ es igual a $2x_0$:*

$$\frac{\mathrm{d}x^2}{\mathrm{d}x}(x_0) = 2x_0.$$

La sucesión correspondiente de cocientes diferenciales toma la forma

$$\frac{\Delta x^2}{\Delta x}(x_0)(h_i) = \frac{(x_0 + h_i)^2 - x_0^2}{h_i} = 2x_0 + h_i,$$

y, tomando el límite, tenemos que

$$\lim_{h_i \to 0} \frac{\Delta x^2}{\Delta x}(x_0)(h_i) = \lim_{h_i \to 0}(2x_0 + h_i) = 2x_0.$$

Así,

$$\frac{\mathrm{d}x^2}{\mathrm{d}x}(x_0) = 2x_0.$$

4. *La derivada en x_0 de la función $f(x) = x^k$, donde k un número natural, es igual a kx_0^{k-1} :*

$$\frac{\mathrm{d}x^k}{\mathrm{d}x}(x_0) = kx_0^{k-1}. \tag{5.3}$$

Consideramos la sucesión correspondiente de cocientes diferenciales

$$\frac{\Delta x^k}{\Delta x}(x_0)(h_i) = \frac{(x_0 + h_i)^k - x_0^k}{h_i}.$$

Por el teorema del binomio tenemos

$$(x_0 + h_i)^k = \sum_{j=0}^{j=k} \frac{k!}{j!(k-j)!} x_0^{k-j} h_i^j,$$

y, entonces, el cociente diferencial es

$$\begin{aligned}
\frac{\Delta x^k}{\Delta x}(x_0)(h_i) &= \frac{1}{h_i}\left[\sum_{j=0}^{j=k} \frac{k!}{j!(k-j)!} x_0^{k-j} h_i^j - x_0^k\right] \\
&= \frac{1}{h_i}\sum_{j=1}^{j=k} \frac{k!}{j!(k-j)!} x_0^{k-j} h_i^j = \\
&= kx_0^{k-1} + \frac{k(k-1)}{2} x_0^{k-2} h_i \\
&\quad + \frac{k(k-1)(k-3)}{3!} x_0^{k-3} h_i^2 + \cdots + kx_0 h_i^{k-1} + h_i^k.
\end{aligned}$$

Calculando el límite cuando $h_i \to 0$ obtendremos

$$\begin{aligned}\lim_{h_i \to 0} \frac{\Delta x^k}{\Delta x}(x_0)(h_i) &= \lim_{h_i \to 0} \Big[k x_0^{k-1} + \frac{k(k-1)}{2} x_0^{k-2} h_i \\ &\quad + \frac{k(k-1)(k-3)}{3!} x_0^{k-3} h_i^2 + \cdots + k x_0 h_i^{k-1} + h_i^k \Big] \\ &= k x_0^{k-1},\end{aligned}$$

ya que el límite de cada uno de los términos que contiene alguna potencia de h_i es cero.

5. *La derivada de la función* $f(x) = \operatorname{sen} x$ *en el punto* x_0 *es igual a* $\cos x_0$:

$$\frac{\mathrm{d} \operatorname{sen} x}{\mathrm{d} x}(x_0) = \cos x_0. \tag{5.4}$$

Consideramos la sucesión correspondiente de cocientes diferenciales

$$\begin{aligned}\frac{\Delta \operatorname{sen} x}{\Delta x}(x_0)(h_i) &= \frac{\operatorname{sen}(x_0 + h_i) - \operatorname{sen} x_0}{h_i} = \\ &= \frac{\operatorname{sen} x_0 \cos h_i + \cos x_0 \operatorname{sen} h_i - \operatorname{sen} x_0}{h_i} = \\ &= \cos x_0 \frac{\operatorname{sen} h_i}{h_i} + \operatorname{sen} x_0 \left(\frac{\cos h_i - 1}{h_i} \right).\end{aligned}$$

Tomando en cuenta que

$$\lim_{h_i \to 0} \frac{\operatorname{sen} h_i}{h_i} = 1 \quad \text{y} \quad \lim_{h_i \to 0} \frac{\cos h_i - 1}{h_i} = 0,$$

tenemos

$$\frac{\mathrm{d} \operatorname{sen} x}{\mathrm{d} x}(x_0) = \lim_{h_i \to 0} \frac{\Delta \operatorname{sen} x}{\Delta x}(x_0)(h_i) = \cos x_0.$$

6. *La derivada de la función* $f(x) = \csc x = \dfrac{1}{\operatorname{sen} x}$ *en el punto* $x_0 \neq k\pi$ *para* $k = 0, \pm 1, \pm 2, \cdots$ *es igual a* $-\cot x_0 \csc x_0$:

$$\frac{\mathrm{d} \csc(x)}{\mathrm{d} x}(x_0) = -\cot x_0 \csc x_0.$$

Consideramos la sucesión correspondiente de cocientes diferenciales

$$\begin{aligned}\frac{\Delta \csc x}{\Delta x}(x_0)(h_i) &= \frac{\dfrac{1}{\operatorname{sen}(x_0 + h_i)} - \dfrac{1}{\operatorname{sen} x_0}}{h_i} = \\ &= \frac{\operatorname{sen} x_0 - \operatorname{sen}(x_0 + h_i)}{h_i \operatorname{sen}(x_0 + h_i) \operatorname{sen} x_0} = \\ &= \frac{\operatorname{sen} x_0 - \operatorname{sen}(x_0 + h_i)}{h_i} \frac{1}{\operatorname{sen}(x_0 + h_i)} \frac{1}{\operatorname{sen} x_0}.\end{aligned}$$

Tomando en cuenta que

$$\lim_{h_i \to 0} \frac{\operatorname{sen}(x_0 + h_i) - \operatorname{sen} x_0}{h_i} = \cos x_0$$

y que $\operatorname{sen} x_0 \neq 0$, se tiene

$$\lim_{h_i \to 0} \frac{1}{\operatorname{sen}(x_0 + h_i)} = \frac{1}{\operatorname{sen} x_0}.$$

Ahora, tomando límite al cociente diferencial cuando $h_i \to 0$ y aplicando las propiedades de las sucesiones convergentes, se obtiene

$$\frac{\mathrm{d} \csc x}{\mathrm{d}x}(x_0) = -\frac{\cos x_0}{\operatorname{sen}^2 x_0} = -\csc x_0 \cdot \cot x_0.$$

5.1.3 Reglas básicas de la derivación de funciones

Cinco son las reglas básicas de derivación para funciones construidas utilizando las operaciones básicas entre funciones derivables. Aplicando esas reglas, podremos calcular las derivadas de todas aquellas funciones que se forman sumando, multiplicando o componiendo funciones derivables.

Teorema 5.2 *Sean $f : (a, b) \to \mathbb{R}$ y $g : (a, b) \to \mathbb{R}$ funciones derivables en un punto $x_0 \in (a, b)$ y sea $q : (c, d) \to \mathbb{R}$ una función derivable en el punto $f(x_0) \in (c, d)$, donde $f(a, b) \subset (c, d)$. Entonces son válidas las reglas siguientes:*

1. Regla de derivación de la suma de funciones.

   La función suma $f + g : (a, b) \to \mathbb{R}$ es derivable en x_0 y

   $$\frac{\mathrm{d}(f + g)}{\mathrm{d}x}(x_0) = \frac{\mathrm{d}f}{\mathrm{d}x}(x_0) + \frac{\mathrm{d}g}{\mathrm{d}x}(x_0). \tag{5.5}$$

2. Regla de Leibniz o de derivación de la multiplicación de funciones.

   La función $f \cdot g : (a, b) \to \mathbb{R}$ es derivable en x_0 y

   $$\frac{\mathrm{d}(f \cdot g)}{\mathrm{d}x}(x_0) = f(x_0)\frac{\mathrm{d}g}{\mathrm{d}x}(x_0) + g(x_0)\frac{\mathrm{d}f}{\mathrm{d}x}(x_0). \tag{5.6}$$

   Consecuencia inmediata de (5.6) y (5.2) es la siguiente.

   Si $c \in \mathbb{R}$, entonces

   $$\frac{\mathrm{d}(c\,f)}{\mathrm{d}x}(x_0) = c\frac{\mathrm{d}f}{\mathrm{d}x}(x_0). \tag{5.7}$$

3. Regla de derivación de un cociente de funciones.

   Si f es derivable en x_0 y $f(x_0) \neq 0$, entonces la función $\dfrac{1}{f}$ está definida en un intervalo alrededor de x_0, es derivable en x_0 y

$$\frac{\mathrm{d}\left(\dfrac{1}{f}\right)}{\mathrm{d}x}(x_0) = -\frac{\dfrac{\mathrm{d}f}{\mathrm{d}x}(x_0)}{f(x_0)^2}. \tag{5.8}$$

4. Regla de la cadena (o de derivación de funciones compuestas).

   La función composición $q \circ f : (a,b) \to \mathbb{R}$ es derivable en el punto x_0 y

$$\frac{\mathrm{d}(q \circ f)}{\mathrm{d}x}(x_0) = \frac{\mathrm{d}q}{\mathrm{d}y}(f(x_0)) \cdot \frac{\mathrm{d}f}{\mathrm{d}x}(x_0). \tag{5.9}$$

5. Regla de derivación de la función inversa.

   Sea $f : [a,b] \to \mathbb{R}$ una función continua y derivable en el punto $x_0 \in (a,b)$ con $\dfrac{\mathrm{d}f}{\mathrm{d}x}(x_0) \neq 0$. Si f posee función inversa $f^{-1} : [c,d] \to [a,b]$, entonces la inversa f^{-1} es derivable en $y_0 = f(x_0)$ y

$$\frac{\mathrm{d}f^{-1}}{\mathrm{d}y}(f(x_0)) = \frac{1}{\dfrac{\mathrm{d}f}{\mathrm{d}x}(x_0)}, \tag{5.10}$$

   o

$$\frac{\mathrm{d}f^{-1}}{\mathrm{d}y}(y_0) = \frac{1}{\dfrac{\mathrm{d}f}{\mathrm{d}x}(f^{-1}(y_0))}. \tag{5.11}$$

Demostración. Sea $h_i \neq 0, i = 1, 2, \ldots,$ y $\{h_i\}_{i=1}^{\infty} \to 0$.

1. La sucesión de cocientes diferenciales en x_0 de la función suma $f + g$ correspondientes a esas variaciones, toma la forma

$$\begin{aligned}\frac{\Delta(f+g)}{\Delta x}(x_0)(h_i) &= \frac{(f+g)(x_0+h_i) - (f+g)(x_0)}{h_i} = \\ &= \frac{f(x_0+h_i) - f(x_0) + g(x_0+h_i) - g(x_0)}{h_i} = \\ &= \frac{\Delta f}{\Delta x}(x_0)(h_i) + \frac{\Delta g}{\Delta x}(x_0)(h_i).\end{aligned}$$

Tomando en cuenta que para las funciones $y = f(x)$ y $z = g(x)$ la sucesión de cocientes diferenciales tiende a sus respectivas derivadas en el punto x_0 cuando el incremento h_i tiende a cero, es decir, que

$$\lim_{h_i \to 0} \frac{\Delta f}{\Delta x}(x_0)(h_i) = \frac{\mathrm{d}f}{\mathrm{d}x}(x_0) \quad \text{y} \quad \lim_{h_i \to 0} \frac{\Delta g}{\Delta x}(x_0)(h_i) = \frac{\mathrm{d}g}{\mathrm{d}x}(x_0),$$

y observando que la sucesión de cocientes diferenciales para la función suma $f+g$ es la suma de las sucesiones de cocientes diferenciales para f y g, respectivamente, al aplicar el teorema 4.6 tenemos que

$$\lim_{h_i \to 0} \frac{\Delta(f+g)}{\Delta x}(x_0)(h_i) = \frac{\mathrm{d}f}{\mathrm{d}x}(x_0) + \frac{\mathrm{d}g}{\mathrm{d}x}(x_0) = \frac{\mathrm{d}(f+g)}{\mathrm{d}x}(x_0).$$

2. Para probar la regla de Leibniz (fórmula (5.6)), escribamos la sucesión de cocientes diferenciales correspondiente al producto de funciones $(f \cdot g)(x)$

$$\begin{aligned}\frac{\Delta(f \cdot g)}{\Delta x}(x_0)(h_i) &= \frac{(f \cdot g)(x_0+h_i) - (f \cdot g)(x_0)}{h_i} = \\ &= \frac{f(x_0+h_i) \cdot g(x_0+h_i) - f(x_0) \cdot g(x_0)}{h_i}.\end{aligned}$$

Al sumar y restar el término $f(x_0+h_i) \cdot g(x_0)$ en el numerador, obtenemos

$$\begin{aligned}\frac{\Delta(f \cdot g)}{\Delta x}(x_0)(h_i) &= \frac{1}{h_i}\Big[f(x_0+h_i)g(x_0+h_i) - f(x_0+h_i)g(x_0) \\ &\quad + f(x_0+h_i)g(x_0) - f(x_0)g(x_0)\Big] \\ &= f(x_0+h_i)\frac{g(x_0+h_i)-g(x_0)}{h_i} + g(x_0)\frac{f(x_0+h_i)-f(x_0)}{h_i} \\ &= f(x_0+h_i)\frac{\Delta g}{\Delta x}(x_0)(h_i) + g(x_0)\frac{\Delta f}{\Delta x}(x_0)(h_i),\end{aligned}$$

de donde se obtiene que

$$\begin{aligned}\left\{\frac{\Delta(f \cdot g)}{\Delta x}(x_0)(h_i)\right\}_{i=1}^{\infty} &= \{f(x_0+h_i)\}_{i=1}^{\infty}\left\{\frac{\Delta g}{\Delta x}(x_0)(h_i)\right\}_{i=1}^{\infty} + \\ &\quad + \{g(x_0)\}_{i=1}^{\infty}\left\{\frac{\Delta f}{\Delta x}(x_0)(h_i)\right\}_{i=1}^{\infty}.\end{aligned}$$

Las sucesiones de la derecha son convergentes por las razones siguientes:

$\{f(x_0+h_i)\}_{i=1}^{\infty}$ converge a $f(x_0)$ pues $y = f(x)$ es continua en x_0;

$\left\{\frac{\Delta g}{\Delta x}(x_0)(h_i)\right\}_{i=1}^{\infty}$ converge a $\frac{\mathrm{d}g}{\mathrm{d}x}(x_0)$ pues $z = g(x)$ es, por hipótesis, derivable en x_0;

$\{g(x_0)\}_{i=1}^{\infty}$ converge a $g(x_0)$ por ser una sucesión constante;

$\left\{\frac{\Delta f}{\Delta x}(x_0)(h_i)\right\}_{i=1}^{\infty}$ converge a $\frac{\mathrm{d}f}{\mathrm{d}x}(x_0)$ pues $y = f(x)$ es, por hipótesis, derivable en x_0.

Entonces, por el teorema 4.6, la sucesión de cocientes diferenciales de la función producto, $\left\{\frac{\Delta(f\cdot g)}{\Delta x}(x_0)(h_i)\right\}_{i=1}^{\infty}$, convergerá, y así tendremos

$$\begin{aligned}\frac{\mathrm{d}(f\cdot g)}{\mathrm{d}x}(x_0) &= \lim_{h_i\to 0}\left\{\frac{\Delta(f\cdot g)(x)}{\Delta x}(x_0)(h_i)\right\}_{i=1}^{\infty} = \\ &= g(x_0)\frac{\mathrm{d}f}{\mathrm{d}x}(x_0) + f(x_0)\frac{\mathrm{d}g}{\mathrm{d}x}(x_0).\end{aligned}$$

3. Para probar (5.8), escribamos la sucesión de cocientes diferenciales correspondiente a la función $\frac{1}{f}$:

$$\begin{aligned}\frac{\Delta\left(\frac{1}{f}\right)}{\Delta x}(x_0)(h_i) &= \frac{1}{h_i}\left(\frac{1}{f(x_0+h_i)} - \frac{1}{f(x_0)}\right) = \\ &= -\frac{f(x_0+h)-f(x_0)}{h_i}\frac{1}{f(x_0+h_i)}\frac{1}{f(x_0)} = \\ &= -\frac{\Delta f}{\Delta x}(x_0)(h_i)\frac{1}{f(x_0+h_i)}\frac{1}{f(x_0)}.\end{aligned}$$

Tomando en cuenta que $y = f(x)$ es derivable en x_0 y que $f(x_0) \neq 0$, tenemos

$$\lim_{h_i\to 0}\frac{\Delta f}{\Delta x}(x_0)(h_i) = \frac{\mathrm{d}f}{\mathrm{d}x}(x_0),$$

y, por la continuidad de f en x_0,

$$\lim_{h_i\to 0}\frac{1}{f(x_0+h_i)} = \frac{1}{f(x_0)}.$$

Tomando el límite cuando $i \to \infty$ en la sucesión de cocientes diferenciales, y aplicando el teorema 4.6, tenemos:

$$\frac{\mathrm{d}\left(\frac{1}{f}\right)}{\mathrm{d}x}(x_0) = \lim_{h_i\to 0}\frac{\Delta\left(\frac{1}{f}\right)}{\Delta x}(x_0)(h_i) = -\frac{1}{f(x_0)^2}\frac{\mathrm{d}f}{\mathrm{d}x}(x_0).$$

4. La sucesión de cocientes diferenciales correspondiente a la composición de funciones $w = (q \circ f)(x)$ es

$$\frac{\Delta(q\circ f)}{\Delta x}(x_0)(h_i) = \frac{q(f(x_0+h_i)) - q(f(x_0))}{h_i}.$$

Sea ahora $\{k_i\}_{i=1}^{\infty}$ la sucesión

$$\{k_i\}_{i=1}^{\infty} = \{f(x_0 + h_i) - f(x_0)\}_{i=1}^{\infty}.$$

Note que $\{k_i\}_{i=1}^{\infty} \to 0$ ya que f es continua en x_0. Aquí distinguiremos dos casos:

(i) Existe una etiqueta N tal que $k_i \neq 0$ para $i > N$. En este caso, podremos dividir entre k_i y hacer para $i > N$ la siguiente estimación:

$$\begin{aligned}
\frac{\Delta(q \circ f)}{\Delta x}(x_0)(h_i) &= \frac{q(f(x_0 + h_i)) - q(f(x_0))}{h_i} = \\
&= \frac{q(f(x_0) + k_i)) - q(f(x_0))}{h_i} \cdot \frac{k_i}{k_i} = \\
&= \frac{q(f(x_0) + k_i)) - q(f(x_0))}{k_i} \cdot \frac{k_i}{h_i} = \\
&= \frac{q(f(x_0) + k_i)) - q(f(x_0))}{k_i} \cdot \frac{f(x_0 + h_i) - f(x_0)}{h_i} = \\
&= \frac{\Delta q}{\Delta y}(f(x_0))(k_i) \cdot \frac{\Delta f}{\Delta x}(x_0)(h_i),
\end{aligned}$$

donde $y = f(x)$.

Tomando el límite de los cocientes diferenciales $\dfrac{\Delta(q \circ f)}{\Delta x}(x_0)(h_i)$ cuando $h_i \to 0$, se tendrá

$$\lim_{h_i \to 0} \frac{\Delta(q \circ f)}{\Delta x}(x_0)(h_i) = \frac{\mathrm{d}(q \circ f)}{\mathrm{d}x}(x_0) = \frac{\mathrm{d}q}{\mathrm{d}y}(f(x_0)) \cdot \frac{\mathrm{d}f}{\mathrm{d}x}(x_0)$$

puesto que, por hipótesis,

$\dfrac{\Delta q}{\Delta y}(f(x_0))(k_i)$ converge a $\dfrac{\mathrm{d}q}{\mathrm{d}y}(f(x_0))$, y

$\dfrac{\Delta f}{\Delta x}(x_0)(h_i)$ converge a $\dfrac{\mathrm{d}f}{\mathrm{d}x}(x_0)$.

(ii) Si existiera una sucesión de etiquetas $\{m_i\}_{i=1}^{\infty} \to \infty$ tales que

$$k_{m_i} = f(x_0 + h_{m_i}) - f(x_0) = 0,$$

se tendría entonces que $\dfrac{\mathrm{d}f}{\mathrm{d}x}(x_0) = 0$ y la sucesión de cocientes de la composición sería una sucesión de ceros,

$$\frac{\Delta(q \circ f)}{\Delta x}(x_0)(h_{m_i}) = \frac{q(f(x_0 + h_{m_i})) - q(f(x_0))}{h_{m_i}} = 0,$$

por lo que

$$\lim_{h_i \to 0} \frac{\Delta(q \circ f)}{\Delta x}(x_0)(h_{m_i}) = \frac{\mathrm{d}q}{\mathrm{d}y}(f(x_0)) \cdot \frac{\mathrm{d}f}{\mathrm{d}x}(x_0).$$

Es decir, en ambos casos se obtiene la misma fórmula para el límite de los cocientes diferenciales, probándose así la validez de la regla de la cadena (5.9).

5. Como f es continua y posee inversa, en virtud del corolario 4.14, f es una función monótona que supondremos creciente. Para probar que la función inversa $x = f^{-1}(y)$ es derivable en $y_0 = f(x_0)$, tomemos una sucesión de variaciones $\{f(x_0) + k_i\}_{i=1}^{\infty} \to 0$ de la variable y con $k_i \neq 0$. Sea ahora la sucesión $\{h_i\}_{i=1}^{\infty}$ definida para cada $i = 1, 2, ..$ como aquel valor $h_i \neq 0$ tal que

$$f(x_0 + h_i) = f(x_0) + k_i.$$

La existencia de las $h_i \neq 0$ que satisfagan la relación anterior está garantizada, pues f es uno a uno, y cada número $f(x_0) + k_i$ pertenece a la imagen de f. Escribamos ahora el cociente diferencial para $x = f^{-1}(y)$ en $f(x_0)$ en la forma

$$\begin{aligned}\frac{\Delta f^{-1}}{\Delta y}(f(x_0))(k_i) &= \frac{f^{-1}(f(x_0)+k_i) - f^{-1}(f(x_0))}{k_i} = \\ &= \frac{h_i}{f(x_0+h_i) - f(x_0)} = \frac{1}{\dfrac{f(x_0+h_i)-f(x_0)}{h_i}},\end{aligned}$$

donde hemos tomando en cuenta que $f^{-1}(f(x)) = x$ para cada $x \in [a, b]$ y hemos sustituido las relaciones

$$x_0 + h_i = f^{-1}(f(x_0 + h_i)) = f^{-1}(f(x_0) + k_i), \quad \text{para } i = 1, 2, \ldots$$

Ahora, considerando que la función inversa es continua en $f(x_0)$, tenemos

$$\lim_{i\to\infty}(x_0 + h_i) = \lim_{i\to\infty}\left\{f^{-1}(f(x_0 + h_i))\right\}_{i=1}^{\infty} = x_0,$$

ya que $\lim_{i\to\infty} f(x_0 + h_i) = f(x_0)$ y entonces $\{h_i\}_{i=1}^{\infty} \to 0$. Luego, aplicando el criterio de convergencia para cocientes de sucesiones (teorema 4.6), podemos escribir

$$\begin{aligned}\frac{\mathrm{d}f^{-1}}{\mathrm{d}y}(f(x_0)) &= \lim_{h_i\to 0}\frac{1}{\dfrac{f(x_0+h_i)-f(x_0)}{h_i}} = \\ &= \frac{1}{\lim\limits_{h_i\to 0}\dfrac{f(x_0+h_i)-f(x_0)}{h_i}} = \frac{1}{\dfrac{\mathrm{d}f}{\mathrm{d}x}(x_0)}.\end{aligned}$$ ■

Ejemplo 5.5 La derivada de la función $\sqrt{x}$ en $x_0 \neq 0$ es el número

$$\frac{\mathrm{d}\sqrt{x}}{\mathrm{d}x}(x_0) = \frac{1}{2\sqrt{x_0}}$$

ya que, al ser $y = f(x) = \sqrt{x}$ la función inversa de la función $x = f^{-1}(y) = y^2$, en $(0, \infty)$ podemos aplicar la fórmula de derivación para la función inversa y obtener

$$\frac{\mathrm{d}\sqrt{x}}{\mathrm{d}x}(x_0) = \frac{1}{\dfrac{\mathrm{d}y^2}{\mathrm{d}y}(\sqrt{x_0})} = \frac{1}{2\sqrt{x_0}}.$$ ◁

5.1.4 Derivadas de funciones racionales, trigonométricas y trigonométricas inversas

Aplicando directamente las reglas de derivación, a continuación enlistamos las fórmulas de derivación de algunas funciones racionales, trigonométricas y sus inversas.

1. *Para m entero y x_0 real con $x_0^{\frac{1}{m}}$ bien definido, se tiene*

$$\frac{\mathrm{d}(x^{\frac{1}{m}})}{\mathrm{d}x}(x_0) = \frac{1}{m}x_0^{\frac{1}{m}-1}.$$

Tomando en cuenta que la función $y = f(x) = x^{\frac{1}{m}}$ es la función inversa de $x = f^{-1}(y) = y^m$, aplicamos la regla de derivación para funciones inversas (5.11) y la fórmula para la derivada de la función $f^{-1}(y) = y^m$ (5.3) para obtener

$$\frac{\mathrm{d}x^{\frac{1}{m}}}{\mathrm{d}x}(x_0) = \frac{1}{\frac{\mathrm{d}y^m}{\mathrm{d}y}(x_0^{\frac{1}{m}})} = \frac{1}{m(x_0^{\frac{1}{m}})^{m-1}} = \frac{1}{m}x_0^{\frac{1}{m}-1}.$$

2. *Para cada $x_0 \in \mathbb{R}$, la función $p(x) = a_0x^n + a_1x^{n-1} + \cdots + a_{n-1}x + a_n$ con $a_i \in \mathbb{R}$, para $i = 0, 1, \ldots, n$ es derivable y*

$$\frac{\mathrm{d}p}{\mathrm{d}x}(x_0) = na_0x_0^{n-1} + (n-1)a_{n-1}x_0^{n-2} + \cdots + 2a_{n-2}x_0 + a_{n-1}.$$

3. *Si $y = f(x)$ y $z = g(x)$ son dos funciones derivables en x_0 y $g(x_0) \neq 0$, la función cociente $(f/g)(x) = \dfrac{f(x)}{g(x)}$ es derivable en $x = x_0$ y*

$$\frac{\mathrm{d}(f/g)}{\mathrm{d}x}(x_0) = \frac{g(x_0)\frac{\mathrm{d}f}{\mathrm{d}x}(x_0) - f(x_0)\frac{\mathrm{d}g(x)}{\mathrm{d}x}(x_0)}{g(x_0)^2}.$$

4. *Para toda $x_0 \in \mathbb{R}$ la función $\cos x$ es derivable y*

$$\frac{\mathrm{d}\cos x}{\mathrm{d}x}(x_0) = -\operatorname{sen}(x_0).$$

De las propiedades de las funciones trigonométricas (en concreto, las propiedades 5 y 6 de la subsección 3.3.2) tenemos que

$$\cos x = \operatorname{sen}\left(x + \frac{\pi}{2}\right).$$

Definiendo las funciones $y = f(x) = x + \dfrac{\pi}{2}$ y $q(y) = \operatorname{sen} y$, podemos escribir $\cos x = (q \circ f)(x)$. Por (5.4),

$$\frac{\mathrm{d}q}{\mathrm{d}y}(f(x_0)) = \cos(f(x_0)) = \cos\left(x_0 + \frac{\pi}{2}\right) = -\operatorname{sen}(x_0),$$

y, por (5.3), $\dfrac{\mathrm{d}f}{\mathrm{d}x}(x_0) = 1$. Al calcular la derivada, usando la regla (5.9), obtendremos

$$\frac{\mathrm{d}\cos x}{\mathrm{d}x}(x_0) = \frac{\mathrm{d}q}{\mathrm{d}y}(f(x_0))\frac{\mathrm{d}f}{\mathrm{d}x}(x_0) = -\operatorname{sen}(x_0).$$

5. *Para* $x_0 \neq \pm\dfrac{\pi}{2}, \pm\dfrac{3}{2}\pi, \pm\dfrac{5}{2}\pi, \ldots,$ *la función* $\tan x$ *es derivable y*

$$\frac{\mathrm{d}\tan x}{\mathrm{d}x}(x_0) = \sec^2 x_0.$$

Aplicando la regla de Leibniz (5.6) y la regla de derivación de un cociente (5.8), obtenemos

$$\begin{aligned}
\frac{\mathrm{d}\tan x}{\mathrm{d}x}(x_0) &= \frac{\mathrm{d}}{\mathrm{d}x}\left(\frac{\operatorname{sen} x}{\cos x}\right)(x_0) \\
&= \frac{1}{\cos x_0}\frac{\mathrm{d}\operatorname{sen} x}{\mathrm{d}x}(x_0) + \operatorname{sen} x_0 \frac{\mathrm{d}}{\mathrm{d}x}\left(\frac{1}{\cos x}\right)(x_0) \\
&= 1 + \frac{\operatorname{sen}^2 x_0}{\cos^2 x_0} = \sec^2 x_0.
\end{aligned}$$

6. *Para* $y_0 \neq \pm 1$, *la función* $\operatorname{arcsen} x$ *es derivable y*

$$\frac{\mathrm{d}\operatorname{arcsen} y}{\mathrm{d}y}(y_0) = \frac{1}{\sqrt{1-y_0^2}}.$$

Aplicando (5.11), se tiene

$$\begin{aligned}
\frac{\mathrm{d}\operatorname{arcsen} y}{\mathrm{d}y}(y_0) &= \frac{1}{\dfrac{\mathrm{d}\operatorname{sen} x}{\mathrm{d}x}(\operatorname{arcsen} y_0)} \\
&= \frac{1}{\cos(\operatorname{arcsen} y_0)} = \frac{1}{\sqrt{1-y_0^2}}.
\end{aligned}$$

7. *Para* $y_0 \in \mathbb{R}$ *la función* $\arctan y$ *es derivable y*

$$\frac{\mathrm{d}\arctan y}{\mathrm{d}y}(y_0) = 1 - y_0^2.$$

Aplicando (5.11) se tiene

$$\begin{aligned}
\frac{\mathrm{d}\arctan y}{\mathrm{d}y}(y_0) &= \frac{1}{\dfrac{\mathrm{d}\tan x}{\mathrm{d}x}(\arctan y_0)} = \\
&= \frac{1}{\sec^2(\arctan y_0)} = 1 - y_0^2.
\end{aligned}$$

8. *Para* $y_0 \in \mathbb{R}$ *con* $y_0 \neq \pm 1$, *la función* $\arccos x$ *es derivable y*

$$\frac{\mathrm{d} \arccos y}{\mathrm{d} y}(y_0) = \frac{-1}{\sqrt{1-y_0^2}}.$$

Aplicando la regla de derivación para la función inversa, se tiene

$$\frac{\mathrm{d} \arccos y}{\mathrm{d} y}(y_0) = \frac{1}{\dfrac{\mathrm{d} \cos x}{\mathrm{d} x}(\arccos y_0)} = \\ = \frac{-1}{\operatorname{sen}(\arccos y_0)} = \frac{-1}{\sqrt{1-y_0^2}}.$$

5.2 Derivadas de orden superior

Sea $y = f(x) : (a,b) \to \mathbb{R}$ una función que tiene derivada en cada uno de los puntos de su dominio de definición. La función que hace corresponder a cada valor de la variable dependiente x el número $\dfrac{\mathrm{d} f}{\mathrm{d} x}(x)$ se llama *función derivada de* $y = f(x)$ y se denota por

$$\frac{\mathrm{d} f}{\mathrm{d} x} : (a,b) \to \mathbb{R}.$$

Ejemplo 5.6 La función derivada de la función $f(x) = x^2 + x + 3$ es la función $\dfrac{\mathrm{d} f}{\mathrm{d} x}(x) = 2x + 1$. ◁

Nota Importante:
Los conceptos “derivada de $y = f(x)$ *en el punto* x_0*” y “función derivada de la función* f *en* (a,b)*” son diferentes. En el primer caso, el concepto se refiere a un número, mientras que en el último se refiere a una función. La función derivada* $\dfrac{\mathrm{d} f}{\mathrm{d} x}$ *de* $y = f(x)$ *es la función que a cada valor de la variable independiente* x *le asocia el valor de la derivada* $\dfrac{\mathrm{d} f}{\mathrm{d} x}(x)$ *en ese valor. En este sentido, la función derivada nos proporciona la ley de cambio que gobierna la relación entre las variables* x *y* y.

Análogamente al concepto de derivada en un punto x_0, se define la *segunda derivada de* $y = f(x)$ *en el punto* x_0 como la derivada en x_0 de su función derivada $\dfrac{\mathrm{d} f}{\mathrm{d} x}(x)$. A la segunda derivada en x_0 se le denota[2] con el símbolo $\dfrac{\mathrm{d}^2 f}{\mathrm{d} x^2}(x_0)$; es decir,

$$\frac{\mathrm{d}^2 f}{\mathrm{d} x^2}(x_0) = \lim_{h \to 0} \frac{\dfrac{\mathrm{d} f}{\mathrm{d} x}(x_0 + h) - \dfrac{\mathrm{d} f}{\mathrm{d} x}(x_0)}{h} = \frac{\mathrm{d}\left(\dfrac{\mathrm{d} f}{\mathrm{d} x}\right)}{\mathrm{d} x}(x_0).$$

[2]Esta es la notación propuesta por Leibniz; otras notaciones para la segunda derivada en x_0 son $y''(x_0)$ o $f''(x_0)$, y $\ddot{y}(x_0)$, de Lagrange y Newton, respectivamente.

En el caso de la función de posición $y = d(t)$ de un móvil con respecto al tiempo, al valor de la segunda derivada en un tiempo t_0 se le llama *aceleración en* t_0.

Ejemplo 5.7 La segunda derivada de la función $f(x) = x^3 + x^2 + x$ en el punto $x_0 = 2$, es la derivada de la función $\dfrac{\mathrm{d}f}{\mathrm{d}x}(x) = 3x^2 + 2x + 1$ en el punto $x_0 = 2$, es decir $\dfrac{\mathrm{d}^2 f}{\mathrm{d}x^2}(x) = 6x + 2$. En particular, $\dfrac{\mathrm{d}^2 f}{\mathrm{d}x^2}(2) = 14$. $\triangleleft$

De manera recursiva, se define la *k-ésima derivada de $y(x)$ en el punto x_0* como la derivada en x_0 de la $(k-1)$-ésima función derivada, y se denota[3] $\dfrac{\mathrm{d}^k f}{\mathrm{d}x^k}(x)$.

Por otra parte, la *función segunda derivada* es la función derivada de la función derivada: $\dfrac{\mathrm{d}}{\mathrm{d}x}\left(\dfrac{\mathrm{d}f}{\mathrm{d}x}\right)$, y la denotamos $\dfrac{\mathrm{d}f}{\mathrm{d}x} : (a,b) \to \mathbb{R}$. De manera recursiva definimos la *función k-ésima derivada* y la denotamos $\dfrac{\mathrm{d}^k f}{\mathrm{d}x^k} : (a,b) \to \mathbb{R}$.

5.3 Diferencial de una función

Si una función $y = f(x)$ tiene derivada en el punto x_0, a la función lineal

$$\begin{gathered} \mathrm{d}f(x_0) : \mathbb{R} \to \mathbb{R} \\ \mathrm{d}f(x_0)(x) = \frac{\mathrm{d}f}{\mathrm{d}x}(x_0)x \end{gathered}$$

se le llama *diferencial de $y = f(x)$ en el punto x_0*. A la diferencial también se le llama *aproximación lineal de la función $f(x)$ alrededor del punto x_0*, ya que si denotamos con $e(x)$ al error entre la función $f(x)$ y la función afín

$$\ell_{x_0}(x) = f(x_0) + \mathrm{d}f(x_0)(x - x_0),$$

es decir,

$$e(x) \stackrel{\text{def}}{=} f(x) - \ell_{x_0}(x) = f(x) - f(x_0) - \mathrm{d}f(x_0)(x - x_0),$$

entonces

$$\lim_{x \to x_0} \frac{e(x)}{|x - x_0|} = 0.$$

Es decir, cuando $x - x_0$ tiende a cero, el error $e(x)$ tiende a cero más rápidamente que $x - x_0$. En términos de la gráfica de $y = f(x)$, la función $y = \ell_{x_0}(x)$ es tal que su gráfica es la recta tangente a la gráfica de $y = f(x)$ en el punto $(x_0, f(x_0))$ (ver figura 5.2).

[3] Otra notación para la k-ésima derivada es $y^{(k)}(x)$ o $f^{(k)}(x)$, con paréntesis, para distinguirla de la potencia $y^k(x)$ o $f^k(x)$.

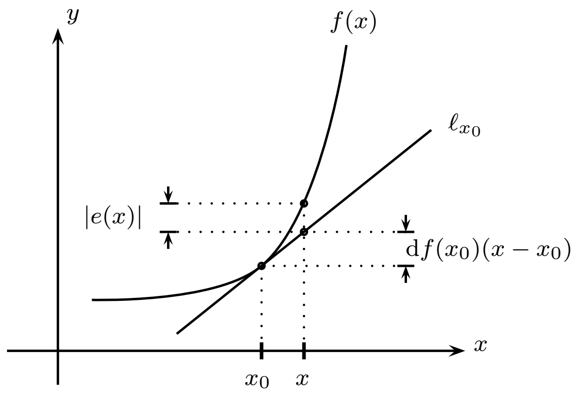


Figura 5.2 Interpretación geométrica de la diferencial de la función $y = f(x)$ en el punto $(x_0, f(x_0))$.

5.4 Cálculo de razones de cambio

Enseguida presentamos algunos ejemplos típicos donde la derivada de una función se expresa en términos de las derivadas de otras a las cuales está ligada funcionalmente.

Ejemplo 5.8 En un tiempo t_0, el ancho de un rectángulo crece a una velocidad de 3 cm/seg y su diagonal crece a razón de 2 cm/seg. ¿Con qué velocidades crecen el perímetro y el área del rectángulo si en t_0 su ancho es de 5 cm y su largo es de 10 cm?

Solución:

Si denotamos por $a(t)$ y $l(t)$ las funciones que a cada tiempo t le asocian el valor del ancho y de la diagonal de un rectángulo, respectivamente, el perímetro $p(t)$ y el área $A(t)$ como funciones del tiempo se escriben, respectivamente, en términos de $a(t)$ y $l(t)$ en la forma

$$p(t) = 2(a(t) + \sqrt{l(t)^2 - a(t)^2}) \quad \text{y} \quad A(t) = a(t)\sqrt{l(t)^2 - a(t)^2},$$

y la información dada en el enunciado del problema es

$$a(t_0) = 5, \qquad l(t_0) = \sqrt{100 + 25}, \qquad \frac{\mathrm{d}a(t)}{\mathrm{d}t}(t_0) = 3, \quad \text{y} \quad \frac{\mathrm{d}l(t)}{\mathrm{d}t}(t_0) = 2. \tag{5.12}$$

Aplicando las reglas de derivación en $t = t_0$, se tiene

$$\frac{dp(t)}{dt}(t_0) = 2\frac{da(t)}{dt}(t_0) + \frac{2l(t_0)\frac{\mathrm{d}l(t)}{\mathrm{d}t}(t_0) - 2a(t_0)\frac{\mathrm{d}a(t)}{\mathrm{d}t}(t_0)}{\sqrt{l(t_0)^2 - a(t_0)^2}}$$

y, sustituyendo (5.12) se obtiene que el perímetro crece en $t = t_0$ a razón de

$$\frac{\mathrm{d}p(t)}{\mathrm{d}t}(t_0) = 3 + 2\sqrt{5}.$$

La derivada del área es

$$\frac{\mathrm{d}A(t)}{\mathrm{d}t}(t_0) = a(t_0)\frac{l(t_0)\frac{\mathrm{d}l(t)}{\mathrm{d}t}(t_0) - a(t_0)\frac{\mathrm{d}a(t)}{\mathrm{d}t}(t_0)}{\sqrt{l(t_0)^2 - a(t_0)^2}} + \sqrt{l(t_0)^2 - a(t_0)^2}\frac{\mathrm{d}a(t)}{\mathrm{d}t}(t_0).$$

Usando (5.12) se obtiene que el área crece en $t = t_0$ a razón de

$$\frac{\mathrm{d}A(t)}{\mathrm{d}t}(t_0) = 5\left(\sqrt{5} + \frac{9}{2}\right). \qquad \triangleleft$$

Ejemplo 5.9 Calcule la razón de cambio del área A de un triángulo rectángulo isósceles con respecto a su perímetro p.

Solución. Obtendremos la respuesta de dos maneras.

a) Una primera forma de resolver el problema es determinar el área de un triángulo rectángulo isósceles como función del perímetro. Si l denota el cateto del triángulo y p el perímetro, se tiene

$$p(l) = 2l + \sqrt{2}l,$$

cuya función inversa es

$$l(p) = \frac{p}{2+\sqrt{2}}.$$

Por otra parte, el área A como función del cateto toma la forma

$$A(l) = \frac{l^2}{2},$$

y componiendo con la función $l(p)$, se tiene para el área la expresión en términos de la variable p

$$A(p) = (A \circ l)(p) = A(l(p));$$

aplicando la regla de la cadena,

$$\frac{\mathrm{d}A}{\mathrm{d}p}(p) = \frac{\mathrm{d}A}{\mathrm{d}l}(l(p))\frac{\mathrm{d}l}{\mathrm{d}p}(p) = \frac{p}{2+\sqrt{2}}\frac{1}{2+\sqrt{2}} = \frac{p}{(2+\sqrt{2})^2}.$$

b) Otra forma de resolver este problema es expresar el área como función del perímetro y luego tomar la derivada. Así,

$$A(p) = \frac{1}{2}\left(\frac{p}{2+\sqrt{2}}\right)^2$$

y, derivando, obtenemos

$$\frac{\mathrm{d}A}{\mathrm{d}p}(p) = \frac{p}{(2+\sqrt{2})^2}. \qquad \triangleleft$$

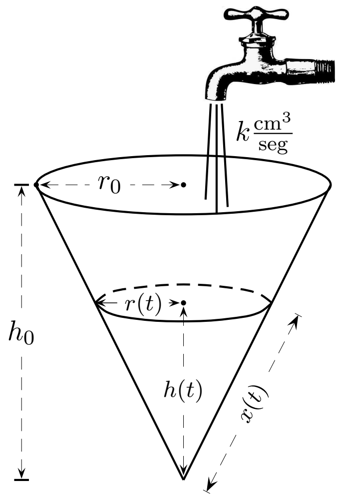


Figura 5.3 $x(t)$ = nivel del agua medido sobre la pared del cono en el tiempo t.

Ejemplo 5.10 Un cono recto de radio r_0 y altura h_0 se llena con un chorro de agua que arroja $k\,\text{cm}^3/\text{seg}$ (ver figura 5.3). Diga con qué velocidad crece el nivel del agua medido sobre la pared del cono cuando se ha llenado la mitad del volumen del cono.

Solución: Para cada t, denotemos por $h(t)$ la altura del agua medida sobre el eje del cono y $r(t)$ el radio del círculo que forma el nivel superior del agua. La relación entre $h(t)$ y $r(t)$, debido a la geometría del cono, es

$$\frac{r(t)}{h(t)} = \frac{r_0}{h_0}.$$

El volumen de un cono $V(t)$ en el tiempo t es

$$V(t) = \frac{\pi}{3} r(t)^2 h(t)$$

y en el caso particular del cono del problema, toma la forma

$$V(t) = \frac{\pi r_0^2}{3h_0^2} h^3(t).$$

Calculando la derivada de $V(t)$, se tiene

$$\frac{\mathrm{d}V}{\mathrm{d}t}(t) = \frac{\pi r_0^2}{h_0^2} h^2(t) \frac{\mathrm{d}h}{\mathrm{d}t}(t).$$

Denotando por t_m el tiempo transcurrido para llenar la mitad del cono, tenemos que la altura del nivel del agua $h(t_m)$ en ese momento es

$$h(t_m) = \frac{1}{\sqrt[3]{2}} h_0,$$

y tomando en cuenta que

$$k = \frac{\mathrm{d}V}{\mathrm{d}t}(t_m) = \frac{\pi r_0^2}{\sqrt[3]{4}} \frac{\mathrm{d}h}{\mathrm{d}t}(t_m),$$

tenemos que la variación de la altura del agua en t_m es

$$\frac{\mathrm{d}h}{\mathrm{d}t}(t_m) = \frac{\sqrt[3]{4}k}{\pi r_0^2}.$$

Por otro lado, sobre la pared del cono el nivel del agua $x(t)$ está relacionado con la altura del agua sobre el eje del cono en la forma

$$x(t) = \frac{h(t)}{h_0}\sqrt{r_0^2 + h_0^2}$$

y entonces la velocidad $\frac{\mathrm{d}x}{\mathrm{d}t}(t_m)$ con que crece el nivel del agua medido sobre la pared del cono, cuando se ha llenado la mitad del cono, es

$$\frac{\mathrm{d}x}{\mathrm{d}t}(t_m) = \frac{\sqrt{r_0^2 + h_0^2}}{h_0} \frac{\mathrm{d}h}{\mathrm{d}t}(t_m) = \frac{\sqrt[3]{4}k}{\pi r_0^2 h_0}\sqrt{r_0^2 + h_0^2}. \qquad \triangleleft$$

Ejercicios y problemas del capítulo

1. (a) Sea la función $y(x) = x|x|$. Calcule y grafique su función derivada en el intervalo $[-1, 1]$.
   (b) Diga en qué puntos tiene derivada la función $f(x) = |x^2 - 1|$.
   (c) Calcule la derivada de la función $f(x) = \operatorname{sen}(g(x) + 2)$ en el punto $x = 3$ si $g(3) = \dfrac{\pi - 12}{6}$ y $\dfrac{\mathrm{d}g}{\mathrm{d}x}(3) = -4$.

2. Sean $y = f(x)$ y $z = g(x)$ funciones derivables en cada punto de $\mathbb{R}$ tales que $\dfrac{\mathrm{d}f}{\mathrm{d}x}(2) = 3$, $\dfrac{\mathrm{d}g}{\mathrm{d}x}(2) = -3$, $f(2) = 1$ y $g(2) = 2$. Calcule:
   (a) $\dfrac{\mathrm{d}(f + g)}{\mathrm{d}x}(2)$
   (b) $\dfrac{\mathrm{d}(f \cdot g)}{\mathrm{d}x}(2)$
   (c) $\dfrac{\mathrm{d}\left(\dfrac{f}{g}\right)}{\mathrm{d}x}(2)$

3. (a) Calcule la ecuación de la recta tangente a la gráfica de la función $y = f(x) = x^2$ en el punto $(1, 1)$.
   (b) Sobre la elipse $\dfrac{x^2}{9} + \dfrac{y^2}{4} = 1$, encuentre los puntos donde la recta tangente tiene una inclinación de $\pi/4$ radianes.

4. ¿Con qué ángulo se corta la elipse $\dfrac{x^2}{a^2} + \dfrac{y^2}{b^2} = 1$ con la parábola $y = x^2$? (Es decir, ¿qué ángulo forman las rectas tangentes en el punto de intersección de las curvas?)

5. Calcule la derivada de la función $f(x) = \cos(\operatorname{sen}(\cos x))$ en el punto $x = \operatorname{arcsen}\dfrac{1}{2}$.

6. Diga en cuántos cm^3 por segundo crece el volumen de un cilindro si su área es de 100 cm^2 y crece 1 cm^2 por segundo y su altura es de 15 cm y decrece 3 cm por segundo.

7. Calcule la variación de la pendiente de la recta tangente a la parábola $y = x^2$ con respecto a la variación de la abscisa del punto de tangencia en el punto $(2, 4)$.

8. Sea $h(x) = (f \circ g)(x))$ donde f y g son funciones derivables. Encuentre la fórmula para $\dfrac{\mathrm{d}^2h}{\mathrm{d}x^2}(x_0)$.

9. Sea $f(x)$ una función derivable para cada $x \in \mathbb{R}$. Demuestre los siguientes enunciados:

   (a) Si $f(x)$ es una función par, entonces $\dfrac{\mathrm{d}f}{\mathrm{d}x}(x)$ es una función impar.

   (b) Si $f(x)$ es una función impar, entonces $\dfrac{\mathrm{d}f}{\mathrm{d}x}(x)$ es una función par.

10. Sea $g(x) = x^3 + x$. Calcule la derivada de la función inversa de $g(x)$ en $x = 2$.

11. Si el ángulo que forman los lados iguales de un triángulo isósceles de altura constante e igual a 10 crece a razón de $\pi/64$ radianes por segundo, ¿cómo varía el área del triángulo en el momento en que su área es 1?

12. Sean las curvas $y = x^2 + 4$ y $y = x^3$. Encuentre los puntos sobre la segunda curva donde la recta tangente a esa curva es perpendicular a la recta tangente a la primera curva en el punto $(-1, 5)$.

13. Considere un cono recto invertido de radio 4 m y altura 10 m en el que se vierte agua a razón de 2 cm^3/seg. Calcule la velocidad con que aumenta el nivel del agua cuando se ha llenado la mitad del recipiente.

14. Considere un cilindro de radio r y altura h. Si en un tiempo t_0 el área lateral del cilindro varía a razón de 2 cm/seg, diga cómo debe variar el radio con respecto al tiempo para que el volumen del cilindro permanezca constante.

15. Considere la curva en el plano $y = x^2 + x + 2$ para $x > 0$ y denote por P un punto de ella. Calcule la razón de cambio del ángulo que hace la tangente a la curva en el punto P con el eje de las abscisas con respecto al valor de la abscisa del punto de tangencia.

16. Considere la familia de parábolas con coeficientes dependientes del tiempo

$$y(x;t) = a(t)x^2 + c(t).$$

   Si el coeficiente $a(t)$ crece a razón de 2 cm/seg y el coeficiente $c(t)$ decrece a razón de -1 cm/seg, ¿con qué velocidad se desplaza el punto en donde las parábolas cortan al eje de las abscisas cuando $a = 1$, $c = -2$?

Capítulo

6 Teorema del valor medio y sus aplicaciones

El teorema del valor medio es uno de los resultados más importantes del análisis matemático. Para las funciones derivables, este teorema compara en términos de su derivada, la variación total de la variable dependiente con respecto a la variación de la variable independiente. Sus distintos corolarios y generalizaciones proporcionan criterios útiles para la descripción cualitativa del comportamiento de la función, tanto globalmente como en la vecindad de cada punto de su dominio.

Iniciamos este capítulo explicando, en lenguaje común, el contenido y significado del teorema del valor medio. Después de su prueba rigurosa, se discuten varias de sus aplicaciones a la descripción de la gráfica de la función y a la reconstrucción de ésta a partir de la información sobre sus derivadas. Finalmente, se presenta su generalización a funciones con derivadas de orden superior, para dar paso al teorema de Taylor y sus estimaciones para la aproximación de funciones derivables mediante polinomios. Con las herramientas anteriores, se dan criterios para la clasificación de los puntos del dominio de una función derivable y se abordan los problemas de cálculo de valores máximos y mínimos, que son de gran importancia en las aplicaciones.

Este capítulo es clave para el buen manejo y aplicación de los conceptos del cálculo, tanto a los problemas propios del análisis matemático como de aplicación en las demás áreas de la ciencia y la tecnología.

6.1 Motivaciones

Consideremos el movimiento de un auto (incluyendo la posibilidad de marchar en reversa o con velocidad negativa) a lo largo de la carretera Hermosillo-Guaymas y supongamos que tiene 100 kilómetros de longitud[1]. Si ese automóvil inicia su viaje en Hermosillo y recorre la distancia al puerto de Guaymas en una hora, entonces podemos afirmar que al menos una vez durante el recorrido, el velocímetro del auto

[1]La distancia real es de 130 kilómetros, aproximadamente.

habrá marcado 100 Km/h (100 kilómetros por hora). Este hecho evidente constituye el contenido del teorema del valor medio.

Una manera equivalente de describir la situación anterior es afirmar que la velocidad promedio que desarrolla un auto al recorrer la distancia entre Hermosillo y Guaymas se leerá al menos una vez en el velocímetro del auto. Es decir, si hacemos T horas en recorrer los 100 kilómetros que separan esas ciudades, entonces el velocímetro del auto al menos una vez marcará el valor $(100/T)$ Km/h.

Como consecuencias directas del teorema del valor medio, tenemos las siguientes conclusiones:

a) Si el velocímetro siempre marca velocidades mayores que cero o positivas, entonces la distancia a la que se encuentra el auto de la ciudad de Hermosillo aumenta con el paso del tiempo. Si, por el contrario, el velocímetro siempre marca velocidad negativa, el auto se alejará cada vez más en sentido contrario a la dirección a Guaymas.

b) Si el velocímetro, durante un recorrido, da siempre una lectura de la velocidad entre a Km/h y b Km/h, con $0 \leqslant a \leqslant b$, entonces, en cada intervalo de tiempo de duración T horas, el auto recorre una distancia mayor o igual que aT y menor o igual que bT.

c) Si un auto que inicia su recorrido en Hermosillo regresa a su punto de partida después de T horas, entonces al menos una vez durante ese lapso el auto se detuvo, es decir su velocímetro debió marcar 0 Km/h.

d) Si el velocímetro de un auto marca 0 Km/h durante un cierto intervalo de tiempo, entonces el auto permanece en la misma posición durante ese intervalo de tiempo.

Como el lector notará, los ejemplos y observaciones anteriores son evidentes y hasta parecen afirmaciones triviales sobre el movimiento de los automóviles. Sin embargo, cuando se generalizan a funciones entre variables reales que tengan derivadas en todos los puntos de un intervalo cerrado $[a, b]$, adquieren un contenido matemático más profundo que permite enunciar verdades generales sobre las funciones derivables.

6.2 El teorema del valor medio

Para funciones reales de una variable real, el *teorema del valor medio* se enuncia en los términos siguientes.

Teorema 6.1 (del valor medio) *Si $f(x) : [a, b] \to \mathbb{R}$ es una función continua que tiene derivada en cada uno de los puntos de (a, b), entonces existe $c \in (a, b)$ tal que*

$$f(b) - f(a) = \frac{\mathrm{d}f}{\mathrm{d}t}(c)(b - a).$$

Demostración. Probaremos primero el teorema suponiendo que la función f toma el mismo valor en los extremos del intervalo $[a, b]$, es decir

$$f(a) = f(b). \tag{6.1}$$

Bajo la hipótesis (6.1) el teorema se denomina *teorema de Rolle*.[2] Demostraremos que existe un punto c tal que $a < c < b$ y $\frac{\mathrm{d}f}{\mathrm{d}x}(c) = 0$. Consideremos los siguientes dos casos posibles para f :

a) f es constante en (a, b), en cuyo caso en cualquier punto $c \in (a, b)$ se tendría $\frac{\mathrm{d}f}{\mathrm{d}x}(c) = 0$, ya que la derivada de una función constante es en todo punto igual a cero.

b) f no es constante en $[a, b]$ y entonces, por la continuidad de la función, tendrá que existir al menos un punto $c \in (a, b)$ donde la función f alcance su valor máximo o su valor mínimo en el intervalo $[a, b]$. Esta última afirmación es cierta, pues si f alcanzara tanto su valor máximo como su valor mínimo en los extremos del intervalo $[a, b]$, siendo $f(a) = f(b)$, se tendría que f es constante en $[a, b]$ y ese no es ahora el caso, luego o el valor máximo o el valor mínimo de f en $[a, b]$ se alcanza en el interior de $[a, b]$. Supongamos que es el valor máximo de f el que se alcanza en un punto c del intervalo abierto (a, b), es decir,

$$c \in (a, b) \quad \text{y} \quad f(c) \geqslant f(x) \quad \text{para toda } x \in [a, b].$$

Evaluando la derivada de f en c tendremos

$$\frac{\mathrm{d}f}{\mathrm{d}x}(c) = \lim_{h \to 0} \frac{f(c + h) - f(c)}{h}.$$

Ahora veamos el signo de $\frac{\mathrm{d}f}{\mathrm{d}x}(c)$. Si la magnitud h de la variación de la variable independiente en $x = c$ se toma mayor que cero, el cociente diferencial tendrá signo negativo

$$\frac{f(c + h) - f(c)}{h} \leqslant 0 \quad \text{si} \quad h > 0,$$

ya que $f(c+h) - f(c) \leqslant 0$ para todo $c+h \in [a, b]$, puesto que $f(c)$ es el valor máximo de la función en el intervalo. En vista de lo anterior, al tomar h valores positivos que

[2] Por Michel Rolle (1652-1719), matemático francés, quien lo publicó en 1691.

tienden a cero, los cocientes diferenciales correspondientes serán siempre menores o iguales a cero y deberán tener como límite un número menor o igual que cero, es decir

$$\lim_{h\to 0^+} \frac{f(c+h)-f(c)}{h} = \frac{\mathrm{d}f}{\mathrm{d}x}(c) \leqslant 0.$$

Análogamente, si la magnitud de la variación h es negativa, el valor del cociente será siempre mayor que cero,

$$\frac{f(c+h)-f(c)}{h} \geqslant 0, \quad \text{si} \quad h < 0,$$

y, por lo tanto, su límite será también mayor o igual a cero:

$$\lim_{h\to 0^-} \frac{f(c+h)-f(c)}{h} = \frac{\mathrm{d}f}{\mathrm{d}x}(c) \geqslant 0.$$

De las dos estimaciones anteriores, concluimos que, si existe la derivada en $x = c$, se debe tener

$$\frac{\mathrm{d}f}{\mathrm{d}x}(c) = 0,$$

con lo que hemos probado la validez del teorema del valor medio en el caso particular de que $f(a) = f(b)$.

En el caso general, si $f(a) \neq f(b)$, consideremos la función auxiliar $g : [a, b] \to \mathbb{R}$ definida por

$$g(x) = f(x) - f(a) - \frac{f(b)-f(a)}{b-a}(x-a),$$

la cual es una función derivable en (a, b) y tal que

$$g(b) = g(a) = 0.$$

Luego, en virtud del caso anterior, existirá $c \in (a, b)$ tal que $\frac{\mathrm{d}g}{\mathrm{d}x}(c) = 0$, y escribiendo esto último en términos de la función original f, se tiene

$$0 = \frac{\mathrm{d}g}{\mathrm{d}x}(c) = \frac{\mathrm{d}f}{\mathrm{d}x}(c) - \frac{f(b)-f(a)}{b-a},$$

es decir,

$$f(b) - f(a) = \frac{\mathrm{d}f}{\mathrm{d}x}(c)(b-a),$$

donde $c \in (a, b)$, y así hemos probado el teorema en general. ■

Una generalización muy importante del teorema del valor medio es la siguiente.

Teorema 6.2 (del valor medio generalizado [3]) *Sean $f(x)$ y $g(x)$ dos funciones continuas en el intervalo $[a, b]$ y derivables en (a, b); supongamos, además, que $\frac{\mathrm{d}g}{\mathrm{d}x}(x) \neq 0$ para toda $x \in (a, b)$. Entonces existe un punto $c \in (a, b)$ tal que*

$$\frac{f(b) - f(a)}{g(b) - g(a)} = \frac{\frac{\mathrm{d}f}{\mathrm{d}x}(c)}{\frac{\mathrm{d}g}{\mathrm{d}x}(c)}.$$

Note que si $\frac{\mathrm{d}g}{\mathrm{d}x}(x) \neq 0$ para toda $x \in (a, b)$, entonces $g(b) - g(a) \neq 0$.

Demostración. Consideremos la función

$$h(x) = (f(b) - f(a))(g(x) - g(a)) - (g(b) - g(a))(f(x) - f(a)),$$

la cual es continua en $[a, b]$ y derivable en (a, b). Por lo tanto, en virtud del teorema del valor medio, existe $c \in (a, b)$ tal que

$$h(b) - h(a) = \frac{\mathrm{d}h}{\mathrm{d}x}(c)(b - a)$$

y escribiendo esta última expresión en términos de las funciones $f(x)$ y $g(x)$, se tiene

$$0 = \left[(f(b) - f(a))\frac{\mathrm{d}g}{\mathrm{d}x}(c) - (g(b) - g(a))\frac{\mathrm{d}f}{\mathrm{d}x}(c)\right](b - a),$$

o lo que es lo mismo,

$$\frac{f(b) - f(a)}{g(b) - g(a)} = \frac{\frac{\mathrm{d}f}{\mathrm{d}x}(c)}{\frac{\mathrm{d}g}{\mathrm{d}x}(c)},$$

como se quería demostrar. ■

6.3 Aplicaciones del teorema del valor medio

Como consecuencia directa del teorema del valor medio, se tienen los siguientes corolarios que recogen y generalizan las afirmaciones que hicimos en la sección 6.1 para el caso del movimiento de un automóvil.

[3] Este resultado se debe a Cauchy y, por eso, a veces se le llama *teorema de Cauchy*.

6.3.1 Significado del signo de la derivada

En primer lugar, presentamos el siguiente corolario que consigna la información que el signo de la función derivada proporciona sobre la función y su comportamiento. La demostración es directa a partir del teorema del valor medio y se deja al lector.

Corolario 6.3 *Sea $f : (a,b) \to \mathbb{R}$ una función derivable.*

a) Si $\dfrac{\mathrm{d}f}{\mathrm{d}x}(x) > 0$ para todo $x \in (a,b)$, entonces $f(x_1) > f(x_2)$ siempre que $x_1 > x_2$ y $x_1, x_2 \in (a,b)$. Es decir, f es una función creciente en (a,b).

b) Si $\dfrac{\mathrm{d}f}{\mathrm{d}x}(x) < 0$ para toda $x \in (a,b)$, entonces $f(x_1) < f(x_2)$ siempre que $x_1 > x_2$ y $x_1, x_2 \in (a,b)$. Es decir, f es una función decreciente en (a,b).

c) Si $\dfrac{\mathrm{d}f}{\mathrm{d}x}(x) = 0$ para toda $x \in (a,b)$, entonces $f(x) =$ constante en (a,b).

d) Si $f(x)$ y $g(x)$ son funciones derivables definidas en un mismo intervalo (a,b) y en cada punto $x \in (a,b)$ se tiene que

$$\frac{\mathrm{d}f}{\mathrm{d}x}(x) = \frac{\mathrm{d}g}{\mathrm{d}x}(x),$$

entonces

$$f(x) = g(x) + c$$

para toda $x \in (a,b)$ y algún $c \in \mathbb{R}$. Es decir, las funciones f y g difieren por una constante.

Ejemplo 6.1 La función $f(x) = \frac{1}{16}x^2(3x^2 + 4x - 12)$, definida en todo $\mathbb{R}$, tiene por función derivada

$$\frac{\mathrm{d}f}{\mathrm{d}x}(x) = \frac{3}{4}x^3 + \frac{3}{4}x^2 - \frac{3}{2}x = \frac{3}{4}x(x+2)(x-1).$$

Luego, el signo de la función derivada es

a) positivo, si $x \in (-2,0) \cup (1,\infty)$,

b) cero si $x = -2, 0, 1$,

c) negativo si $x \in (-\infty,-2) \cup (0,1)$.

Aplicando el corolario 6.3, tenemos que la función f es creciente en $(-2,0) \cup (1,\infty)$ y decreciente en $(-\infty,-2) \cup (0,1)$. Además, f tiene un mínimo local en $x = -2$ con valor $f(-2) = -2$, un máximo local en $x = 0$ con valor $f(0) = 0$ y un mínimo local en $x = 1$ con valor $f(1) = -5/16$. El mínimo absoluto se alcanza en $x = -2$, mientras que no posee máximo absoluto pues crece sin límite cuando $x \to \pm\infty$. (Ver figura 6.1.) $\triangleleft$

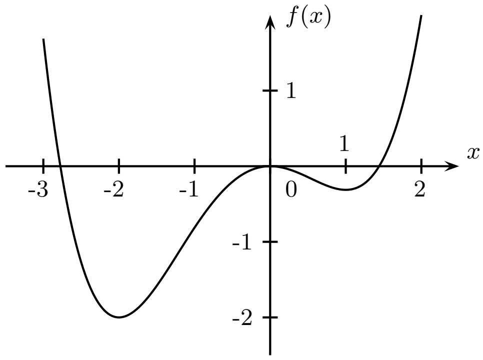


Figura 6.1 Gráfica de la función $f(x) = \dfrac{1}{16}x^2(3x^2 + 4x - 12)$

6.3.2 La función segunda derivada

Tomando en cuenta que la función segunda derivada es la función derivada de la función derivada de f, el conjunto de enunciados, teoremas y corolarios anteriores son aplicables a la segunda derivada, dando lugar a una mayor información sobre la función inicial. En el caso de que f sea la función de posición de un auto, la segunda derivada se llama *función de aceleración* o *de variación de la velocidad con respecto al tiempo.*

Teorema 6.4 *Sea $f : (a, b) \to \mathbb{R}$ una función que tiene segunda derivada en (a, b) y $c \in (a, b)$. Entonces, para cada punto $x_0 \in (a, b)$, la diferencia o error $R(x_0) = f(x_0) - P(x_0)$ entre $f(x_0)$ y el valor en x_0 del polinomio de primer grado $P(x) = f(c) + \dfrac{\mathrm{d}f}{\mathrm{d}x}(c)(x - c)$ en x_0, es tal que*

$$R(x_0) = \frac{\mathrm{d}^2 f}{\mathrm{d}x^2}(x_c)(x_0 - x_c)(x_0 - c),$$

donde $x_c \in (c, x_0)$.

Demostración. Definamos la función

$$g(x) = f(x) + \frac{\mathrm{d}f}{\mathrm{d}x}(x)(x_0 - x).$$

Nótese que

$$\frac{\mathrm{d}g}{\mathrm{d}x}(x) = \frac{\mathrm{d}^2 f}{\mathrm{d}x^2}(x)(x_0 - x).$$

Si $x_0 \in [c, b)$, al aplicar el teorema del valor medio en el intervalo $[c, x_0]$ a la función g obtenemos

$$g(x_0) - g(c) = \frac{\mathrm{d}g}{\mathrm{d}x}(x_c)(x_0 - c), \tag{6.2}$$

donde $x_c \in (c, x_0)$. Al escribir la expresión (6.2) en términos de la función f, se tiene que

$$f(x_0) - f(c) - \frac{\mathrm{d}f}{\mathrm{d}x}(c)(x_0 - c) = \frac{\mathrm{d}^2 f}{\mathrm{d}x^2}(x_c)(x_0 - x_c)(x_0 - c),$$

lo que prueba el teorema. ■

Corolario 6.5 *Bajo las hipótesis del teorema 6.4 y suponiendo además que la función segunda derivada es acotada por un número M, es decir, que*

$$\left|\frac{\mathrm{d}^2 f}{\mathrm{d}x^2}(x)\right| \leqslant M \quad \textit{para todo} \quad x \in (a, b),$$

se tiene la siguiente estimación para el error $R(x_0)$ en cada punto $x_0 \in (a, b)$

$$|R(x_0)| = \left|f(x_0) - f(c) - \frac{\mathrm{d}f}{\mathrm{d}x}(c)(x_0 - c)\right| \leqslant M|x_0 - c|^2.$$

Corolario 6.6 *Sea $f : (a, b) \to \mathbb{R}$ una función con segunda derivada en (a, b). Entonces los siguientes enunciados son válidos:*

1. *Si $\dfrac{\mathrm{d}^2 f}{\mathrm{d}x^2}(x) = 0$ para cada $x \in (a, b)$, entonces $f(x) = mx + n$, con m, n constantes. Es decir, f es una función afín.*

2. *Si $\dfrac{\mathrm{d}^2 f}{\mathrm{d}x^2}(x) > 0$ (o $\dfrac{\mathrm{d}^2 f}{\mathrm{d}x^2}(x) < 0$) para toda $x \in (a, b)$, la gráfica de la función f se encuentra de un mismo lado de la recta tangente*

$$y = f(c) + \frac{\mathrm{d}f}{\mathrm{d}x}(c)(x - c)$$

en el punto $(c, f(c))$, ya que

$$f(x) - f(c) - \frac{\mathrm{d}f}{\mathrm{d}x}(c)(x - c) = \frac{\mathrm{d}^2 f}{\mathrm{d}x^2}(x_c)(x - x_c)(x - c) > 0 \ (< 0).$$

Si, además, $\dfrac{\mathrm{d}f}{\mathrm{d}x}(x) > 0$, entonces la función es creciente de manera cóncava hacia arriba en (a, b). En el caso de que $\dfrac{\mathrm{d}^2 f}{\mathrm{d}x^2}(x) < 0$ y $\dfrac{\mathrm{d}f}{\mathrm{d}x}(x) > 0$, la función f es creciente de manera cóncava hacia abajo en (a, b).

6.3.3 Curvatura de curvas en el plano

Consideremos en el plano cartesiano la curva de forma $y = f(x)$ para $x \in (a, b)$, donde f es una función con segunda derivada continua. En cada punto $P = (x_0, f(x_0))$ de la curva, a la recta N_0 cuya ecuación es

$$y = f(x_0) - \frac{1}{\dfrac{\mathrm{d}f}{\mathrm{d}x}(x_0)}(x - x_0) \tag{6.3}$$

se le llama la *recta normal a f por P*. Esta recta es perpendicular a la recta tangente por P a $f(x)$. Si $\dfrac{\mathrm{d}f}{\mathrm{d}x}(x_0) = 0$, la ecuación de N_0 es $x = x_0$.

Para puntos $Q = (x_0 + h, f(x_0 + h))$ con $h \neq 0$ cercanos a P, la recta normal N_h a f por Q toma la forma

$$y = f(x_0 + h) - \frac{1}{\dfrac{\mathrm{d}f}{\mathrm{d}x}(x_0 + h)}(x - x_0 - h). \tag{6.4}$$

Resolviendo las ecuaciones (6.3) y (6.4) para (x, y), se obtienen las coordenadas del punto de intersección $C_h = (x_h, y_h)$ de N_0 y N_h en la forma

$$\begin{aligned}
x_h &= x_0 - \frac{\dfrac{\mathrm{d}f}{\mathrm{d}x}(x_0 + h)\dfrac{\mathrm{d}f}{\mathrm{d}x}(x_0)}{\dfrac{\mathrm{d}f}{\mathrm{d}x}(x_0 + h) - \dfrac{\mathrm{d}f}{\mathrm{d}x}(x_0)}\left[f(x_0 + h) - f(x_0) + \frac{h}{\dfrac{\mathrm{d}f}{\mathrm{d}x}(x_0 + h)}\right],\\
y_h &= f(x_0) + \frac{\dfrac{\mathrm{d}f}{\mathrm{d}x}(x_0 + h)}{\dfrac{\mathrm{d}f}{\mathrm{d}x}(x_0 + h) - \dfrac{\mathrm{d}f}{\mathrm{d}x}(x_0)}\left[f(x_0 + h) - f(x_0) + \frac{h}{\dfrac{\mathrm{d}f}{\mathrm{d}x}(x_0 + h)}\right].
\end{aligned}$$

Aplicando el teorema del valor medio en $[x_0, x_0 + h]$, a las funciones $f(x)$ y $\dfrac{\mathrm{d}f}{\mathrm{d}x}(x)$, escribimos

$$\begin{aligned}
f(x_0 + h) - f(x_0) &= \frac{\mathrm{d}f}{\mathrm{d}x}(x_1)h,\\
\frac{\mathrm{d}f}{\mathrm{d}x}(x_0 + h) - \frac{\mathrm{d}f}{\mathrm{d}x}(x_0) &= \frac{\mathrm{d}^2 f}{\mathrm{d}x^2}(x_2)h,
\end{aligned}$$

donde $x_1, x_2 \in (x_0, x_0+h)$. Así, las coordenadas del punto C_h, se reescriben

$$\begin{aligned} x_h &= x_0 - \frac{\dfrac{\mathrm{d}f}{\mathrm{d}x}(x_0+h)\dfrac{\mathrm{d}f}{\mathrm{d}x}(x_0)}{\dfrac{\mathrm{d}^2 f}{\mathrm{d}x^2}(x_2)h}\left[\frac{\mathrm{d}f}{\mathrm{d}x}(x_1)h + \frac{h}{\dfrac{\mathrm{d}f}{\mathrm{d}x}(x_0+h)}\right] \\ &= x_0 - \frac{\dfrac{\mathrm{d}f}{\mathrm{d}x}(x_0+h)\dfrac{\mathrm{d}f}{\mathrm{d}x}(x_0)}{\dfrac{\mathrm{d}^2 f}{\mathrm{d}x^2}(x_2)}\left[\frac{\mathrm{d}f}{\mathrm{d}x}(x_1) + \frac{1}{\dfrac{\mathrm{d}f}{\mathrm{d}x}(x_0+h)}\right], \\ y_h &= f(x_0) + \frac{\dfrac{\mathrm{d}f}{\mathrm{d}x}(x_0+h)}{\dfrac{\mathrm{d}^2 f}{\mathrm{d}x^2}(x_2)h}\left[\frac{\mathrm{d}f}{\mathrm{d}x}(x_1)h + \frac{h}{\dfrac{\mathrm{d}f}{\mathrm{d}x}(x_0+h)}\right] \\ &= f(x_0) + \frac{\dfrac{\mathrm{d}f}{\mathrm{d}x}(x_0+h)}{\dfrac{\mathrm{d}^2 f}{\mathrm{d}x^2}(x_2)}\left[\frac{\mathrm{d}f}{\mathrm{d}x}(x_1) + \frac{1}{\dfrac{\mathrm{d}f}{\mathrm{d}x}(x_0+h)}\right]. \end{aligned}$$

Haciendo ahora tender h a cero, se tiene $x_1 \to x_0$ y $x_2 \to x_0$; luego el punto

$$C_0 = \lim_{h\to 0} C_h = (\lim_{h\to 0} x_h, \lim_{h\to 0} y_h)$$

toma la forma

$$C_0 = (x_0, f(x_0)) + \frac{1}{\dfrac{\mathrm{d}^2 f}{\mathrm{d}x^2}(x_0)}\left[1 + \left(\frac{\mathrm{d}f}{\mathrm{d}x}(x_0)\right)^2\right]\Big(-\frac{\mathrm{d}f}{\mathrm{d}x}(x_0), 1\Big).$$

Al punto C_0 se le llama el *centro de curvatura de f en el punto P* y a su distancia ρ al punto P, dada por

$$\rho = \frac{1}{\left|\dfrac{\mathrm{d}^2 f}{\mathrm{d}x^2}(x_0)\right|}\left[1 + \left(\frac{\mathrm{d}f}{\mathrm{d}x}(x_0)\right)^2\right]^{\frac{3}{2}},$$

se le llama el *radio de curvatura de f en P*. Finalmente, se define la *curvatura* $\kappa(x_0)$ *en el punto* $(x_0, f(x_0))$ de la curva como

$$\kappa(x_0) = \frac{1}{\rho}.$$

En particular, una circunferencia de radio r tiene curvatura constante $1/r$. Por esa razón, al círculo con centro en C_0 y radio ρ se le llama *círculo de curvatura* (o *círculo osculador*) de la curva en P.

Ejemplo 6.2 Usando los cálculos anteriores para la función $y = f(x) = x^3 - x^2 + 1$ en $x_0 = 1$, encontramos lo siguiente:

la recta normal N_0 en el punto $P(x_0, f(x_0)) = (1,1)$ es $y = -x + 2$;

la recta normal N_h en el punto $Q(x_0 + h, f(x_0 + h))$ es

$$y = -\frac{x}{1 + 4h + 3h^2} + 1 + h + 2h^2 + h^3 + \frac{3}{h+1};$$

el centro de curvatura es $C_0 = (h,k) = \left(\frac{1}{2}, \frac{3}{2}\right)$;

el radio de curvatura es $\rho = \frac{\sqrt{2}}{2}$. (Ver figura 6.2.) ◁

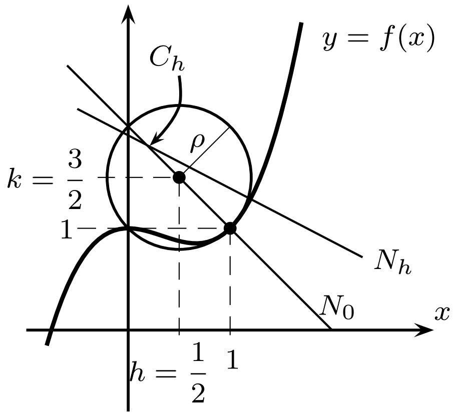


Figura 6.2 El círculo de curvatura para la función $f(x) = x^3 - x^2 + 1$ en $x_0 = 1$

6.4 El teorema de Taylor

Los valores de una función y de sus derivadas de orden superior en un punto x_0 de su dominio proporcionan información importante sobre su comportamiento alrededor de ese punto. Por ejemplo, si todas las derivadas de orden k de la función f se anulan en un intervalo, entonces la función es un polinomio de grado k, como lo mostramos en la proposición siguiente.

Proposición 6.7 *Si $f : (a,b) \to \mathbb{R}$ es una función que posee derivada de orden k en cada punto $x \in (a,b)$ con*

$$\frac{\mathrm{d}^k f}{\mathrm{d}x^k}(x) = 0 \quad \textit{para} \quad x \in (a,b),$$

y x_0 es un punto arbitrario de (a,b), entonces $f(x)$ es un polinomio de grado $k-1$ de la forma

$$f(x) = a_0 + a_1(x - x_0) + \cdots + a_{k-1}(x - x_0)^{k-1},$$

donde $a_0 = f(x_0)$, $a_j = \dfrac{1}{j!}\dfrac{\mathrm{d}^j f}{\mathrm{d}x^j}(x_0)$, *para* $j = 1, 2, \ldots, k-1$.

Demostración. Consideremos la función

$$\begin{aligned} g(x) &= f(x) + \frac{\mathrm{d}f}{\mathrm{d}x}(x)(x_0 - x) + \frac{1}{2!}\frac{\mathrm{d}^2 f}{\mathrm{d}x^2}(x)(x_0 - x)^2 \\ &\quad + \cdots + \frac{1}{(k-1)!}\frac{\mathrm{d}^{k-1} f}{\mathrm{d}x^{k-1}}(x)(x_0 - x)^{k-1}. \end{aligned}$$

Derivando, obtenemos

$$\frac{\mathrm{d}g}{\mathrm{d}x}(x) = 0.$$

Por lo tanto, g es constante y

$$g(x) = f(x_0),$$

y podemos escribir

$$\begin{aligned} f(x_0) &= f(x) + \frac{\mathrm{d}f}{\mathrm{d}x}(x)(x_0 - x) + \frac{1}{2!}\frac{\mathrm{d}^2 f}{\mathrm{d}x^2}(x)(x_0 - x)^2 \\ &\quad + \cdots + \frac{1}{(k-1)!}\frac{\mathrm{d}^{k-1} f}{\mathrm{d}x^{k-1}}(x)(x_0 - x)^{k-1}. \end{aligned}$$

Como x y x_0 son puntos arbitrarios de (a, b), los podemos intercambiar y obtener

$$\begin{aligned} f(x) &= f(x_0) + \frac{\mathrm{d}f}{\mathrm{d}x}(x_0)(x - x_0) + \frac{1}{2!}\frac{\mathrm{d}^2 f}{\mathrm{d}x^2}(x_0)(x - x_0)^2 \\ &\quad + \cdots + \frac{1}{(k-1)!}\frac{\mathrm{d}^{k-1} f}{\mathrm{d}x^{k-1}}(x_0)(x - x_0)^{k-1}. \end{aligned}$$ ■

En el caso de funciones con derivadas de orden superior, el teorema del valor medio permite dar una estimación de la diferencia entre la función y el polinomio construido como en la proposición 6.7. Este nuevo resultado se conoce como *teorema de Taylor,*[4] el cual presentamos a continuación.

Consideremos una función real $f : [a, b] \to \mathbb{R}$ que tenga derivadas de orden k en cada punto de (a, b). Tomemos un punto $c \in (a, b)$ y definamos el polinomio $P_c^k(x)$ de grado k mediante

$$P_c^k(x) = f(c) + \frac{\mathrm{d}f}{\mathrm{d}x}(c)(x - c) + \frac{1}{2!}\frac{\mathrm{d}^2 f}{\mathrm{d}x^2}(c)(x - c)^2 + \cdots + \frac{1}{k!}\frac{\mathrm{d}^k f}{\mathrm{d}x^k}(c)(x - c)^k.$$

Directamente podemos verificar que los valores de $P_c^k(x)$ y f, así como los de sus derivadas hasta orden k, coinciden en el punto c. Es decir,

$$P_c^k(c) = f(c), \quad \frac{\mathrm{d}P_c^k}{\mathrm{d}x}(c) = \frac{\mathrm{d}f}{\mathrm{d}x}(c), \quad \cdots, \quad \frac{\mathrm{d}^k P_c^k}{\mathrm{d}x^k}(c) = \frac{\mathrm{d}^k f}{\mathrm{d}x^k}(c).$$

[4]Brook Taylor (1685-1731), matemático inglés. Enunció, en 1712, el teorema que lleva su nombre.

Al polinomio P_c^k se le llama el *polinomio de Taylor de grado k en el punto c de la función f.*

Teorema 6.8 (de Taylor) *Sea $f : [a,b] \to \mathbb{R}$ con derivadas hasta de orden $k+1$ en (a,b) y consideremos $P_c^k(x)$ su polinomio de Taylor de grado k alrededor del punto $c \in (a,b)$. Entonces, para cada punto $x_0 \in (a,b)$ la diferencia*

$$R(x_0) = f(x_0) - P_c^k(x_0)$$

entre el valor de f y el valor del polinomio de Taylor $P_c^k(x_0)$ en el punto $x = x_0$, es tal que

$$R(x_0) = \frac{1}{k!}\frac{\mathrm{d}^{k+1}f}{\mathrm{d}x^{k+1}}(x_c)(x_0 - x_c)^k(x_0 - c),$$

donde x_c es un punto entre c y x_0. A la función $R(x)$ se le llama también la ***función residuo de orden*** *k.*

Demostración. Si $x_0 \in [c,b)$, aplicando el teorema del valor medio en el intervalo $[c,x_0]$ a la función

$$\begin{aligned} g(x) &= f(x) + \frac{\mathrm{d}f}{\mathrm{d}x}(x)(x_0 - x) + \cdots + \frac{1}{k!}\frac{\mathrm{d}^k f}{\mathrm{d}x^k}(x)(x_0 - x)^k = \\ &= f(x) + \sum_{i=1}^{k}\frac{1}{i!}\frac{\mathrm{d}^i f}{\mathrm{d}x^i}(x)(x_0 - x)^i, \end{aligned}$$

se tiene

$$g(x_0) - g(c) = \frac{\mathrm{d}g}{\mathrm{d}x}(x_c)(x_0 - c),$$

donde $x_c \in (c, x_0)$. Por otro lado,

$$g(x_0) - g(c) = R(x_0).$$

Calculando directamente el término $\frac{\mathrm{d}g}{\mathrm{d}x}(x_c)$, se tiene

$$\begin{aligned} \frac{\mathrm{d}g}{\mathrm{d}x}(x_c) \quad &= \quad \frac{\mathrm{d}f}{\mathrm{d}x}(x_c) + \sum_{i=1}^{k}\frac{1}{i!}\frac{\mathrm{d}^{i+1}f}{\mathrm{d}x^{i+1}}(x_c)(x_0 - x_c)^i \\ &\quad -\sum_{i=1}^{k}\frac{1}{(i-1)!}\frac{\mathrm{d}^i f}{\mathrm{d}x^i}(x_c)(x_0 - x_c)^{i-1}. \end{aligned} \tag{6.5}$$

Reacomodando los índices de la última suma en la forma

$$\sum_{i=1}^{k}\frac{1}{(i-1)!}\frac{\mathrm{d}^i f}{\mathrm{d}x^i}(x_c)(x_0 - x_c)^{i-1} = \sum_{i=0}^{k-1}\frac{1}{i!}\frac{\mathrm{d}^{i+1}f}{\mathrm{d}x^{i+1}}(x_c)(x_0 - x_c)^i,$$

y sustituyendo en (6.5), obtenemos que el error $R(x_0)$ es

$$R(x_0) = \frac{\mathrm{d}g}{\mathrm{d}x}(x_c)(x_0 - c) = \frac{1}{k!}\frac{\mathrm{d}^{k+1}f}{\mathrm{d}x^{k+1}}(x_c)(x_0 - x_c)^k(x_0 - c),$$

donde x_c es un punto en el intervalo (c, x_0). ■

Nota Importante:

1. *El teorema 6.8 nos da explícitamente el valor de la diferencia entre el valor de la función f en el punto x_0 y el valor del polinomio de Taylor en ese mismo punto. Esa diferencia está dada en términos de la $(k+1)$-ésima derivada en un punto intermedio x_c entre el punto x_0 y el punto c donde se calculan los coeficientes del polinomio de Taylor. Ese punto x_c depende del punto x_0 y su existencia la asegura el teorema del valor medio.*

2. *Si la función derivada de orden $(k+1)$ de la función $f(x)$ es acotada en (a, b), es decir, si existe $M > 0$ tal que*

$$\left|\frac{\mathrm{d}^{k+1}f}{\mathrm{d}x^{k+1}}(x)\right| < M \quad \text{para toda } x \in (a, b),$$

entonces se tiene la estimación siguiente para la función error:

$$|R(x_0)| < \frac{1}{k!}M|x_0 - c|^{k+1}$$

para cada $x_0 \in (a, b)$.

3. *El valor de la función residuo $R(x_0)$ depende de la distancia del punto x_0 al punto c, donde se determina el polinomio, y de la cota M de la $(k+1)$-ésima derivada en el intervalo $[c, x_0]$.*

4. *También se acostumbra escribir el residuo en la forma*

$$x_0 - x_c = x_0 - c - \theta(x_0 - c) = (1 - \theta)(x_0 - c)$$

donde $0 < \theta < 1$, y

$$R(x_0) = \frac{1}{k!}\frac{\mathrm{d}^{k+1}f}{\mathrm{d}x^{k+1}}(x_c)(1 - \theta)^k(x_0 - c)^{k+1}.$$

Esta fórmula para el residuo se le atribuye a Cauchy.

La aproximación de una función por su polinomio de Taylor permite calcular el valor de la función en puntos cercanos a un punto donde se pueda conocer explícitamente el valor de la función y sus derivadas. Enseguida se muestra este tipo de aplicaciones.

Ejemplo 6.3 Aplicando la fórmula para el residuo, calcule el error que se comete al evaluar $\sqrt{10}$ usando el polinomio de Taylor de orden 1 en el punto $x = 9$.

Solución: Sea $f(x) = \sqrt{x}$ y consideremos su polinomio de Taylor de orden 1 alrededor de $x = 9$,

$$P_9^1(x) = 3 + \frac{1}{6}(x - 9).$$

Por el teorema de Taylor, se tiene que

$$\sqrt{x} = 3 + \frac{1}{6}(x - 9) + R(x)$$

con

$$R(x) = \frac{\mathrm{d}^2 f}{\mathrm{d}x^2}(x_c)(x - x_c)(x - 9), \text{ donde } x_c \in [9, x].$$

Luego,

$$\sqrt{10} \approx 3 + \frac{1}{6},$$

con un error $R(10)$ acotado por

$$|R(10)| \leqslant \left| \frac{\mathrm{d}^2 \sqrt{x}}{\mathrm{d}x^2}(x_c) \right| \leqslant \frac{1}{4} \frac{1}{\sqrt{x_c^3}} \leqslant \frac{1}{4(27)},$$

ya que $x_c \in [9, 10]$. En conclusión,

$$\sqrt{10} \approx 3.16 \qquad \triangleleft$$

con un error menor que 0.01.

Ejemplo 6.4 Para $x \in (-1, 1)$, encuentre el polinomio de Taylor que aproxima a la función $\operatorname{sen} x$ con un error menor que 10^{-3}. Calcule con dos cifras decimales el valor de $\operatorname{sen}(1/2)$.

Solución: Como los puntos de $(-1, 1)$ son cercanos a $x = 0$, donde se conoce exactamente el valor de las derivadas de la función $\operatorname{sen} x$, entonces los polinomios de Taylor en cero son los apropiados para aproximar a esa función en puntos de $(-1, 1)$. Notemos primero que todas las derivadas de la función $\operatorname{sen} x$ son funciones acotadas:

$$\left| \frac{\mathrm{d}^n \operatorname{sen} x}{\mathrm{d}x^n}(x) \right| \leqslant 1, \quad \text{para toda } x \in \mathbb{R} \text{ y } n \text{ natural.}$$

Para k natural, el polinomio de Taylor de orden $2k + 1$ en el punto $x = 0$, toma la forma

$$P_0^{2k+1}(x) = x - \frac{1}{3!}x^3 + \frac{1}{5!}x^5 - \cdots + \frac{(-1)^k}{(2k+1)!}x^{2k+1}$$

y la diferencia $R(x)$ entre $\operatorname{sen} x$ y $P_0^{2k+1}(x)$ es

$$\operatorname{sen} x - P_0^{2k+1}(x) = R(x)$$

con
$$|R(x)| \leqslant \frac{1}{(2k+1)!}|x|^{2k+2}.$$
Luego, para $x \in (-1,1)$, se tiene $|R(x)| \leqslant \dfrac{1}{(2k+1)!}$ y, por lo tanto, $|R(x)| \leqslant 10^{-3}$ si $k=3$. Así, el polinomio de Taylor de grado menor que aproxima a $\operatorname{sen} x$ para todo $x \in (-1,1)$ con un error menor que 10^{-3} es
$$P_0^7(x) = x - \frac{1}{3!}x^3 + \frac{1}{5!}x^5 - \frac{1}{7!}x^7.$$
Para el valor de sen(1/2), el error $R(1/2) = \operatorname{sen}(1/2) - P_0^{2k+1}(1/2)$ es tal que
$$|R(1/2)| \leqslant \frac{1}{2^{2k+2}(2k+1)!}.$$
Entonces, para estimar sen(1/2) con 2 cifras decimales, basta tomar $k=2$:
$$\operatorname{sen}(1/2) \approx \frac{1}{2} - \frac{1}{48} = \frac{23}{48} = 0.47\ldots$$ ◁

6.4.1 Puntos regulares, críticos y de inflexión

Los puntos del dominio de una función real $f(x)$ que posee una o más derivadas continuas, se clasifican en:

puntos regulares, que son aquéllos donde la derivada es distinta de cero, y

puntos críticos o *puntos singulares,* donde la derivada es igual a cero.

La propiedad característica de un punto regular es que existe un intervalo con centro en ese punto, donde la función es monótona y, por lo tanto, posee ahí función inversa. Esto se debe a que, al ser la derivada de la función en ese punto distinta de cero, por la continuidad de la función derivada, debe ser distinta de cero en un intervalo alrededor de ese punto y, por el teorema del valor medio, en ese intervalo la función será monótona y por lo tanto tendrá función inversa en el intervalo en cuestión.

Si la función $f : [a,b] \to \mathbb{R}$ es una función continua, a un punto $c \in (a,b)$ se le dice *punto máximo local* (o *punto mínimo local*), si existe un intervalo I con centro en c tal que
$$f(c) \geqslant f(x) \quad \text{para todo } x \in I$$
(o $f(c) \leqslant f(x)$ para todo $x \in I$, respectivamente). Note que los puntos máximos y mínimos locales tienen esa propiedad solamente en un intervalo alrededor de ese punto. Es posible que el máximo global de la función $f(x)$ se alcance en alguno de los puntos extremos del intervalo inicial $[a,b]$ y que en ellos la derivada tenga un valor distinto de cero, como se ve en la figura 6.3. Lo importante es que si un punto interior $c \in (a,b)$ es máximo o mínimo local de $f(x)$, necesariamente la derivada de f se debe anular en c y ese punto será un punto crítico. Para demostrar lo anterior,

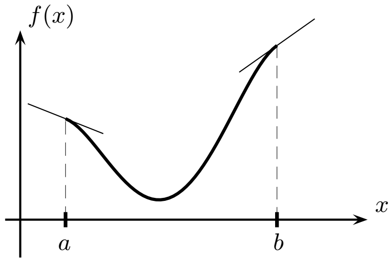


Figura 6.3 $f(b)$ es máximo y $\dfrac{\mathrm{d}f}{\mathrm{d}x}(b) > 0$

supongamos que $c \in (a, b)$ es un máximo local en el intervalo $I = (c - \varepsilon, c + \varepsilon)$. Si calculamos la derivada de f en $x = c$, tendremos

$$\frac{\mathrm{d}f}{\mathrm{d}x}(c) = \lim_{h \to 0} \frac{f(c+h) - f(c)}{h}.$$

Si la sucesión $\{h_n\}_{n=1}^{\infty}$ es de números positivos, la sucesión de los cocientes

$$\left\{ \frac{f(c+h_n) - f(c)}{h_n} \right\}_{n=1}^{\infty}$$

es una sucesión de números negativos pues el numerador es negativo ya que $f(c) \geqslant f(c+h_n)$ y el denominador h_n es positivo. Luego, $\displaystyle\lim_{n\to\infty} \frac{f(c+h_n) - f(c)}{h_n}$ será menor o igual a cero. Análogamente, si tomamos la sucesión $\{h_n\}_{n=1}^{\infty}$ de números negativos que convergen a cero, la sucesión de cocientes $\dfrac{f(c+h_n) - f(c)}{h_n}$ será una sucesión de números positivos y $\displaystyle\lim_{n\to\infty} \frac{f(c+h_n) - f(c)}{h_n}$ deberá ser entonces mayor o igual a cero. Como ese límite es precisamente la derivada $\dfrac{\mathrm{d}f}{\mathrm{d}x}(c)$, tenemos entonces que la única manera en que ambas sucesiones converjan a un mismo número es que ese número sea igual a cero. Con esta argumentación concluimos que si $x = c$ es un punto máximo o mínimo local de f, entonces se cumple que

$$\frac{\mathrm{d}f}{\mathrm{d}x}(c) = 0.$$

Nota Importante:
La condición de que la derivada de una función f se anule en los puntos máximos y mínimos locales es una condición necesaria que cumplen ese tipo de puntos. Sin embargo, por sí sola no es una condición suficiente para asegurar que esos puntos sean máximos o mínimos locales. Por ejemplo, la derivada de la función $f(x) = x^3$ se anula en cero; sin embargo, ese punto no es máximo ni mínimo local pues la función crece a su derecha y decrece a su izquierda.

La observación anterior nos lleva de manera natural a plantear el problema de encontrar condiciones adicionales al anulamiento de la derivada en un punto para poder asegurar que ese punto sea un máximo o mínimo local. Una de esas condiciones, se tiene en términos del signo de la segunda derivada, como sigue.

Proposición 6.9 *Sea $f : [a, b] \to \mathbb{R}$ una función que tiene función segunda derivada continua y sea $c \in (a, b)$ un punto crítico $\left(\frac{\mathrm{d}f}{\mathrm{d}x}(c) = 0\right)$. Entonces, si $\frac{\mathrm{d}^2 f}{\mathrm{d}x^2}(c) < 0$, el punto c es un máximo local y si $\frac{\mathrm{d}^2 f}{\mathrm{d}x^2}(c) > 0$, el punto c es un mínimo local.*

Demostración. Probaremos únicamente el primer caso; el segundo se demuestra de manera análoga y se deja al lector. Como $\frac{\mathrm{d}^2 f}{\mathrm{d}x^2}(x)$ es una función continua con $\frac{\mathrm{d}^2 f}{\mathrm{d}x^2}(c) < 0$, existe un intervalo $I = (c - \varepsilon, c + \varepsilon)$ con centro en c y radio $\varepsilon > 0$ tal que $\frac{\mathrm{d}^2 f}{\mathrm{d}x^2}(x) < 0$ para toda $x \in I$. Más aún, para cada $x \in (c - \varepsilon, c + \varepsilon)$ se tiene, por el teorema de Taylor de orden 2, la estimación

$$f(x) - f(c) - \frac{\mathrm{d}f}{\mathrm{d}x}(c)(x - c) = \frac{\mathrm{d}^2 f}{\mathrm{d}x^2}(x_c)(x - c)(x - x_c),$$

donde x_c es un punto entre x y c. Tomando en cuenta que $\frac{\mathrm{d}f}{\mathrm{d}x}(c) = 0$ y que $\frac{\mathrm{d}^2 f}{\mathrm{d}x^2}(x_c) < 0$, resulta que

$$f(x) - f(c) < 0, \quad \text{para toda } x \in I,$$

lo cual muestra que $f(c)$ es el valor máximo de f en el intervalo $(c - \varepsilon, c + \varepsilon)$ y c es un máximo local de f. ∎

Finalmente, en esta sección presentamos el concepto de *punto de inflexión.*

Un punto $x = c$ se llama *punto de inflexión* de una función $y = f(x)$ derivable en c, si existe un intervalo abierto I que contiene a c tal que la función $R(x) = f(x) - f(c) - \frac{\mathrm{d}f}{\mathrm{d}x}(c)(x - c)$ es negativa si $x < c$ y positiva si $x > c$, o recíprocamente: es negativa si $x > c$ y positiva si $x < c$.

Para funciones que poseen segunda derivada continua, si $x = c$ es un punto de inflexión existe un intervalo abierto I que contiene a c donde la función cambia de concavidad; es decir,

$$\frac{\mathrm{d}^2 f}{\mathrm{d}x^2}(x) < 0 \ \text{ si } \ x < c \ \text{ y } \ \frac{\mathrm{d}^2 f}{\mathrm{d}x^2}(x) > 0 \ \text{ si } \ x > c, \ \text{ o}$$

$$\frac{\mathrm{d}^2 f}{\mathrm{d}x^2}(x) > 0 \ \text{ si } \ x < c \ \text{ y } \ \frac{\mathrm{d}^2 f}{\mathrm{d}x^2}(x) < 0 \ \text{ si } \ x > c.$$

En este caso, el teorema del valor intermedio aplicado a la función segunda derivada implica que si $x = c$ es un punto de inflexión de f, la segunda derivada de f se anula en ese punto, es decir $\dfrac{\mathrm{d}^2 f}{\mathrm{d}x^2}(c) = 0$.

Nota Importante:
En un punto de inflexión de una función f no necesariamente se anula su derivada. Por ejemplo, para la función $f(x) = x^3 - x$, se tiene $\dfrac{\mathrm{d}f}{\mathrm{d}x}(0) = -1$ y el punto $x = 0$ es punto de inflexión, ya que la función segunda derivada, $\dfrac{\mathrm{d}^2 f}{\mathrm{d}x^2}(x) = 6x$, es negativa si $x < 0$ y positiva si $x > 0$.

Finalizamos esta sección con el siguiente teorema, que nos proporciona un criterio general para caracterizar los puntos críticos y los puntos de inflexión de una función.

Teorema 6.10 *Sea $f : [a, b] \to \mathbb{R}$ una función que tiene derivadas continuas de orden r en (a, b) y $c \in (a, b)$. Supongamos que*

$$\frac{\mathrm{d}^2 f}{\mathrm{d}x^2}(c) = \frac{\mathrm{d}^3 f}{\mathrm{d}x^3}(c) = \cdots = \frac{\mathrm{d}^{r-1} f}{\mathrm{d}x^{r-1}}(c) = 0, \tag{6.6}$$

pero $\dfrac{\mathrm{d}^r f}{\mathrm{d}x^r}(c) \neq 0$. Entonces:

a) Si $\dfrac{\mathrm{d}f}{\mathrm{d}x}(c) = 0$ y r es par con $\dfrac{\mathrm{d}^r f}{\mathrm{d}x^r}(c) > 0$, el punto c es un punto mínimo local;

b) Si $\dfrac{\mathrm{d}f}{\mathrm{d}x}(c) = 0$ y r es par con $\dfrac{\mathrm{d}^r f}{\mathrm{d}x^r}(c) < 0$, el punto c es un máximo local;

c) Si r es impar, c es un punto de inflexión.

Demostración. Aplicando el teorema de Taylor (teorema 6.8), se tiene para cada $x \in (a, b)$ la estimación siguiente:

$$\begin{aligned} f(x) - f(c) - \frac{\mathrm{d}f}{\mathrm{d}x}(c)(x-c) - \cdots - \frac{1}{(r-1)!}\frac{\mathrm{d}^{r-1} f}{\mathrm{d}x^{r-1}}(c)(x-c)^{r-1} \\ = \frac{1}{(r-1)!}\frac{\mathrm{d}^r f}{\mathrm{d}x^r}(x_c)(x-x_c)^{r-1}(x-c). \end{aligned} \tag{6.7}$$

Debido a (6.6), la expresión (6.7) toma la forma

$$f(x) - f(c) - \frac{\mathrm{d}f}{\mathrm{d}x}(c)(x-c) = \frac{1}{(r-1)!}\frac{\mathrm{d}^r f}{\mathrm{d}x^r}(x_c)(x-x_c)^{r-1}(x-c),$$

donde $x - x_c > 0$ si $x > c$ y $x - x_c < 0$ si $x < c$. En el caso de que $\dfrac{\mathrm{d}f}{\mathrm{d}x}(c) = 0$ y $\dfrac{\mathrm{d}^r f}{\mathrm{d}x^r}(c) > 0$, de la continuidad de $\dfrac{\mathrm{d}^r f}{\mathrm{d}x^r}(x)$ en $x = c$ podemos asegurar la existencia de un intervalo de la forma $(c-\varepsilon, c+\varepsilon)$ donde $\dfrac{\mathrm{d}^r f}{\mathrm{d}x^r}(x) > 0$ para toda $x \in (c-\varepsilon, c+\varepsilon)$ y, en particular, $\dfrac{\mathrm{d}^r f}{\mathrm{d}x^r}(x_c) > 0$ pues $|x_c - c| < \varepsilon$. Por lo tanto, siendo r un número par, el lado derecho de (6.7) será

$$f(x) - f(c) = \frac{1}{(r-1)!}\frac{\mathrm{d}^r f}{\mathrm{d}x^r}(x_c)(x - x_c)^{r-1}(x - c) > 0$$

para cada $x \in (c-\varepsilon, c+\varepsilon)$, lo cual significa que $x = c$ es un punto mínimo local pues

$$f(x) > f(c) \text{ para toda } x \in (c-\varepsilon, c+\varepsilon).$$

En el caso de que $\dfrac{\mathrm{d}^r f}{\mathrm{d}x^r}(x_c) < 0$ un argumento análogo al anterior prueba que $x = c$ tiene que ser un punto máximo local.

Finalmente, para probar c) supongamos que $\dfrac{\mathrm{d}^r f}{\mathrm{d}x^r}(x) > 0$. Por la continuidad de $\dfrac{\mathrm{d}^r f}{\mathrm{d}x^r}(x)$ en $x = c$, existe un intervalo de la forma $(c-\varepsilon, c+\varepsilon)$ donde $\dfrac{\mathrm{d}^r f}{\mathrm{d}x^r}(x) > 0$ para toda $x \in (c-\varepsilon, c+\varepsilon)$ y al ser r un número impar, el término de la derecha en (6.7) tiene la forma

$$\frac{1}{(r-1)!}\frac{\mathrm{d}^r f}{\mathrm{d}x^r}(x_c)(x - x_c)^{r-1}(x - c),$$

que toma signos contrarios al evaluarlo en puntos x a la derecha y a la izquierda de c, lo que significa que c es un punto de inflexión. ■

Ejemplo 6.5 Los puntos del dominio de la función $f : \mathbb{R} \to \mathbb{R}$ con $f(x) = 8x^9 - 9x^8$ se clasifican de la forma siguiente:

De la función primera derivada

$$\frac{\mathrm{d}f}{\mathrm{d}x}(x) = 72x^7(x-1)$$

concluimos que el conjunto A de puntos regulares de f es

$$A = \mathbb{R} - \{0, 1\}.$$

La función segunda derivada es

$$\frac{\mathrm{d}^2 f}{\mathrm{d}x^2}(x) = 504x^6(x-1) + 72x^7 = 72x^6(8x-7).$$

Concluimos que $x = 1$ es mínimo local, pues $\dfrac{\mathrm{d}f}{\mathrm{d}x}(1) = 0$ y $\dfrac{\mathrm{d}^2 f}{\mathrm{d}^2 x}(1) > 0$. Finalmente, para el punto $x = 0$, tenemos

$$\frac{\mathrm{d}^2 f}{\mathrm{d}x^2}(0) = 0,\ \frac{\mathrm{d}^3 f}{\mathrm{d}x^3}(0) = 0, \cdots,\ \frac{\mathrm{d}^7 f}{\mathrm{d}x^7}(0) = 0,\ \frac{\mathrm{d}^8 f}{\mathrm{d}x^8}(0) < 0$$

y por lo tanto $x = 0$ es un máximo local. La gráfica de f aparece en la figura 6.4.◁

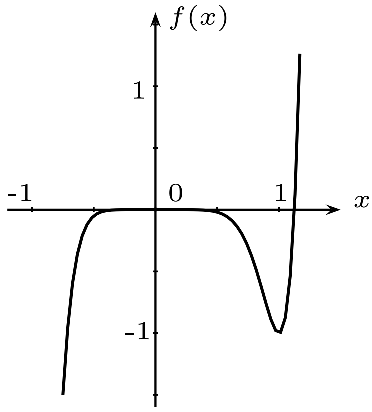


Figura 6.4 Gráfica de la función $f(x) = 8x^9 - 9x^8$

Ejemplo 6.6 Supongamos que la gráfica de la derivada de la función $y = f(x)$ en el intervalo $(0, 10)$ es la que se muestra en la figura 6.5.

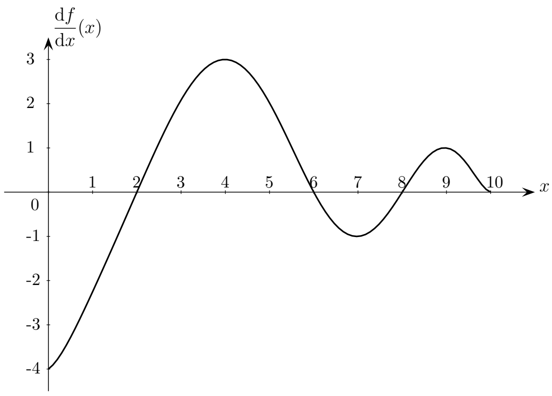


Figura 6.5 Gráfica de $\dfrac{\mathrm{d}f}{\mathrm{d}x}(x)$

Podemos observar las características siguientes de la función f:

a) Es creciente en los intervalos $[2, 6]$ y $[8, 10)$ ya que en esos intervalos $\dfrac{\mathrm{d}f}{\mathrm{d}x}(x)$ es positiva.

b) Es decreciente en los intervalos $(0,2]$ y $[6,8]$, puesto que en ellos $\frac{\mathrm{d}f}{\mathrm{d}x}(x)$ es negativa.

c) Es cóncava hacia arriba en los intervalos $(0,4]$ y $[7,9]$, pues ahí $\frac{\mathrm{d}^2f}{\mathrm{d}x^2}(x) > 0$.

d) Es cóncava hacia abajo en los intervalos $[4,7]$ y $[9,10)$, ya que $\frac{\mathrm{d}^2f}{\mathrm{d}x^2}(x) < 0$ en esos intervalos.

e) Los puntos $x = 2, 8$ son mínimos locales pues $\frac{\mathrm{d}f}{\mathrm{d}x}(2) = \frac{\mathrm{d}f}{\mathrm{d}x}(8) = 0$, $\frac{\mathrm{d}^2f}{\mathrm{d}x^2}(2) > 0$ y $\frac{\mathrm{d}^2f}{\mathrm{d}x^2}(8) > 0$.

f) El punto $x = 6$ es un máximo local pues $\frac{\mathrm{d}f}{\mathrm{d}x}(6) = 0$ y $\frac{\mathrm{d}^2f}{\mathrm{d}x^2}(6) < 0$.

h) Los puntos $x = 4, 7, 9$ son puntos de inflexión ya que $\frac{\mathrm{d}^2f}{\mathrm{d}x^2}(4) = \frac{\mathrm{d}^2f}{\mathrm{d}x^2}(7) = \frac{\mathrm{d}^2f}{\mathrm{d}x^2}(9) = 0$ y $\frac{\mathrm{d}^3f}{\mathrm{d}x^3}(4) \neq 0$, $\frac{\mathrm{d}^3f}{\mathrm{d}x^3}(7) \neq 0$, $\frac{\mathrm{d}^3f}{\mathrm{d}x^3}(9) \neq 0$.

i) La gráfica de f tal que $f(2) = 0$ aparece en la figura 6.6. ◁

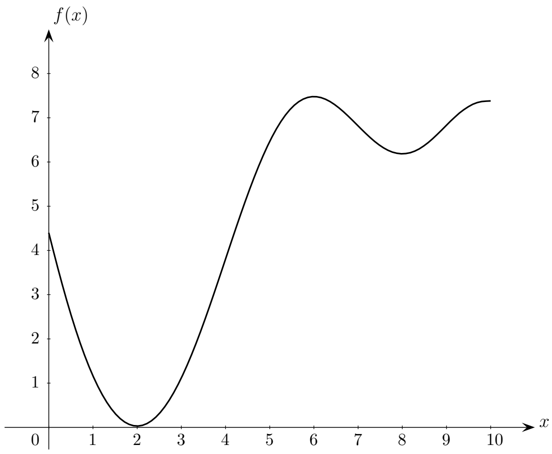


Figura 6.6 Gráfica de f

Ejemplo 6.7 Determine las dimensiones del cilindro de área mínima (incluyendo sus tapaderas) de volumen 1000 cm^3.

Solución. Cada cilindro de volumen 1000 cm^3 está totalmente determinado por el valor de su radio r o de su altura h ya que se tiene la relación

$$\pi r^2 h = 10^3$$

Si tomamos el valor del radio r como la variable que define a los cilindros de volumen 10^3 cm^3, entonces su altura $h(r)$ está dada por

$$h(r) = \frac{10^3}{\pi r^2}, \quad \text{para} \quad r > 0.$$

El área A de sus lados y tapaderas es la función $A : (0, \infty) \to \mathbb{R}$ de la variable independiente r y está dada por

$$A(r) = 2\pi r h(r) + 2\pi r^2 = \frac{2 \cdot 10^3}{r} + 2\pi r^2.$$

Entonces, el problema planteado consiste en determinar el valor de r para el cual $A(r)$ es mínima. Lo anterior nos lleva a determinar los mínimos locales de la función y después a tomar aquél de ellos en el que $A(r)$ tome el valor menor. Los mínimos locales se encuentran entre los puntos r tales que

$$\frac{\mathrm{d}A}{\mathrm{d}r}(r) = -\frac{2 \cdot 10^3}{r^2} + 4\pi r = 0,$$

es decir,

$$r = 10\sqrt[3]{\frac{1}{2\pi}}.$$

Tomando en cuenta que el valor de la segunda derivada en el único punto crítico es

$$\frac{\mathrm{d}^2A}{\mathrm{d}r^2}\left(10\sqrt[3]{\frac{1}{2\pi}}\right) = 120\pi > 0,$$

se tiene que dicho punto es mínimo local y entonces el cilindro de área mínima de volumen 1000 cm^3 es aquél cuyas dimensiones son

$$r = 10\sqrt[3]{\frac{1}{2\pi}}, \quad h = \frac{10}{\sqrt[3]{\frac{\pi}{4}}},$$

y el valor del área mínima es

$$A_{\min} = 300\sqrt[3]{2\pi}.$$

Observe que, tomando el radio muy pequeño o muy grande, el área del cilindro de volumen 1000 cm^3 se puede hacer tan grande como se quiera. $\triangleleft$

Ejemplo 6.8 Encuentre las dimensiones del rectángulo de área máxima que puede inscribirse en la elipse

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1,$$

y tiene uno de sus lados sobre el eje mayor. Calcule el área máxima.

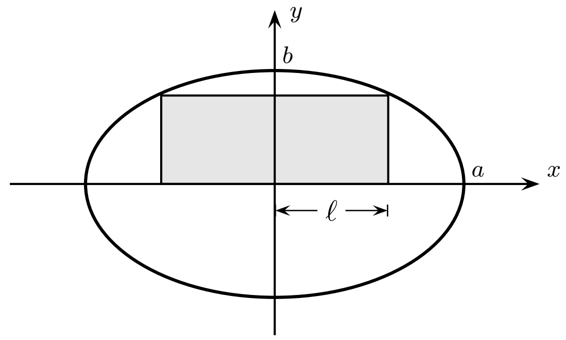


Figura 6.7 El rectángulo del ejemplo 6.8

Solución. De la figura 6.7 observamos que cada rectángulo en cuestión está determinado por el valor del segmento ℓ señalado en la figura y su área es la función $A : (0, a) \to \mathbb{R}$, de la variable ℓ, definida mediante la regla

$$A(\ell) = 2\ell b\sqrt{1 - \frac{\ell^2}{a^2}}, \quad 0 < \ell < a.$$

Luego, el valor de $\ell_{\max}$ que da lugar al rectángulo inscrito de mayor área deberá anular la derivada

$$\frac{\mathrm{d}A}{\mathrm{d}\ell}(\ell_{\max}) = \frac{2b}{a}\frac{a^2 - 2\ell_{\max}^2}{\sqrt{a^2 - \ell_{\max}^2}} = 0.$$

Despejando $\ell_{\max}$ de la fórmula anterior, se tiene

$$\ell_{\max} = \frac{1}{2}\sqrt{2}a.$$

Note que el valor de la segunda derivada en $\ell_{\max}$ es negativa y, por lo tanto, corresponde a un punto máximo local. Las dimensiones del rectángulo inscrito de área máxima son

$$\text{base} = \sqrt{2}a, \quad \text{altura} = \frac{\sqrt{2}}{2}b,$$

y el área máxima,

$$A_{\max} = ab.$$

Observe que cuando ℓ tiende a cero, el área del rectángulo inscrito correspondiente tiende a cero, es decir, el valor mínimo del área de los rectángulos inscritos en la elipse es cero. ◁

Ejemplo 6.9 Un deportista se encuentra en un punto A al borde de un lago circular de radio r Km y desea llegar al punto C, diametralmente opuesto a A. (Ver figura 6.8).

Si puede correr a razón de v_c Km por hora y remar en un bote a razón de v_r Km por hora, ¿en qué ángulo θ con relación al diámetro debe remar para luego correr sobre el borde del lago para alcanzar el punto opuesto en el menor tiempo posible?

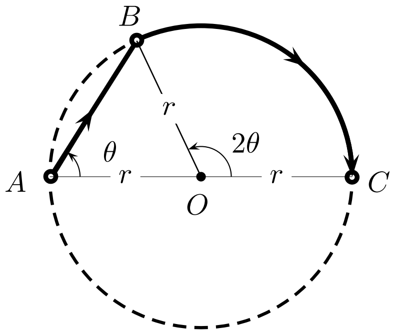


Figura 6.8 Diagrama para el ejemplo 6.9

Solución. Cada trayectoria posible está determinada por el ángulo θ y, correspondientemente, tiene asociado un tiempo $t(\theta)$ de recorrido. La distancia recorrida en el agua es el lado AB del triángulo isósceles AOB, la cual es igual a $2r\cos\theta$, mientras que la distancia recorrida en tierra es igual a la longitud del arco subtendido por el ángulo central 2θ, la cual es igual a $2r\theta$. Entonces, el tiempo de recorrido es la función $t : \left[0, \frac{\pi}{2}\right] \to \mathbb{R}$ dada por

$$t(\theta) = \frac{2r\cos\theta}{v_r} + \frac{2r\theta}{v_c}.$$

Luego, el ángulo $\theta_{\min}$ correspondiente al tiempo mínimo de recorrido debe anular la derivada de $t(\theta)$, es decir

$$\frac{\mathrm{d}t}{\mathrm{d}\theta}(\theta_{\min}) = \frac{-2r\operatorname{sen}\theta_{\min}}{v_r} + \frac{2r}{v_c} = 0.$$

Note que si $0 < \dfrac{v_r}{v_c} < 1$, se tiene un único punto crítico en $\left(0, \dfrac{\pi}{2}\right)$, dado por

$$\theta_{\min} = \operatorname{arcsen}\left(\frac{v_r}{v_c}\right)$$

y corresponde a un mínimo local, ya que el valor de la segunda derivada en ese punto toma un valor negativo,

$$\frac{\mathrm{d}^2t}{\mathrm{d}\theta^2}(\theta_{\min}) = \frac{-2r\cos\theta_{\min}}{v_r} < 0.$$

El valor del tiempo correspondiente a $\theta_{\min}$ es

$$\begin{aligned} t(\theta_{\min}) &= \frac{2r\cos\left(\operatorname{arcsen}\left(\frac{v_r}{v_c}\right)\right)}{v_r} + \frac{2r\operatorname{arcsen}\left(\frac{v_r}{v_c}\right)}{v_c} \\ &= 2r\left[\frac{1}{v_r}\sqrt{1-\left(\frac{v_r}{v_c}\right)^2} + \frac{1}{v_c}\operatorname{arcsen}\left(\frac{v_r}{v_c}\right)\right], \end{aligned}$$

y define el mínimo absoluto de la función en el intervalo $[0, \frac{\pi}{2}]$, ya que en los extremos $\theta = 0, \dfrac{\pi}{2}$ se tiene $t(0) = \dfrac{2r}{v_r}$ y $t\left(\dfrac{\pi}{2}\right) = \dfrac{r\pi}{v_c}$, los valores de la función $t(\theta)$ deben ser superiores a $t(\theta_{\min})$ por haber sólo un punto crítico en $(0, \frac{\pi}{2})$.

Note que si $\dfrac{v_r}{v_c} \geqslant 1$, entonces el valor mínimo de $t(\theta)$ se alcanza en el extremo izquierdo de $[0, \frac{\pi}{2}]$ con $\theta_{\min} = 0$ y con valor $t(\theta_{\min}) = \dfrac{2r}{v_r}$. ◁

6.4.2 Reglas de L'Hospital

A veces es necesario calcular el límite de cocientes de funciones en puntos en los cuales el límite de cada una de las funciones es cero o tienden a $\pm\infty$ simultáneamente. Por ejemplo, límites del tipo

$$\lim_{x\to 0} \frac{\operatorname{sen} x}{x}, \quad \lim_{x\to 0} \frac{\sqrt{\dfrac{1}{x^3} + x}}{\dfrac{1}{x} + 2}, \quad \lim_{x\to a} \frac{\tan(x-a)}{\csc(x-a)}.$$

En estos casos, el teorema del valor medio generalizado nos permite examinar el comportamiento de las funciones a partir del comportamiento de sus derivadas, proporcionando dos criterios muy útiles en tales casos, que se denominan *reglas de L'Hospital.*[5]

Proposición 6.11 (Primera regla de L'Hospital) *Sean $f(x)$ y $g(x)$ dos funciones derivables en (a,b) y $\dfrac{\mathrm{d}g}{\mathrm{d}x}(x) \neq 0$ para toda $x \in (a,b)$. Si se tiene que*

$$\lim_{x\to a} f(x) = \lim_{x\to a} g(x) = 0 \quad y \quad \lim_{x\to a} \frac{\dfrac{\mathrm{d}f}{\mathrm{d}x}(x)}{\dfrac{\mathrm{d}g}{\mathrm{d}x}(x)} = L,$$

donde $L \in \mathbb{R}$, o $L = \pm\infty$, entonces

$$\lim_{x\to a} \frac{f(x)}{g(x)} = L.$$

Demostración. Obsérvese que si extendemos el dominio de definición de $f(x)$ y $g(x)$ al punto $x = a$ definiéndolas ahí como $f(a) = g(a) = 0$, se tiene una extensión como funciones continuas a $[a,b)$. Aplicando el teorema 6.2 del valor medio

[5] En honor de Guillaume de L'Hospital, mencionado en el capítulo primero.

generalizado en $[a, x]$, podemos escribir

$$\frac{f(x)-f(a)}{g(x)-g(a)} = \frac{f(x)}{g(x)} = \frac{\dfrac{\mathrm{d}f}{\mathrm{d}x}(c_x)}{\dfrac{\mathrm{d}g}{\mathrm{d}x}(c_x)}.$$

donde $c_x \in (a, x)$. Tomando límite cuando $x \to a$, se tiene que $c_x \to a$ y de la hipótesis, obtenemos

$$\lim_{x\to a} \frac{f(x)}{g(x)} = \lim_{x\to a} \frac{\dfrac{\mathrm{d}f}{\mathrm{d}x}(c_x)}{\dfrac{\mathrm{d}g}{\mathrm{d}x}(c_x)} = L,$$

con lo que se prueba la proposición. ■

Nota Importante:
Cuando escribimos $\lim_{x\to a} h(x)$ *y la función* $h(x)$ *está definida en* (a, b), *nos referimos a que la variable* x *tiende a* a *con valores en el intervalo* (a, b). *Esto se escribe también en la forma* $\lim_{x\to a^+} h(x)$ *para denotar que la variable se acerca de derecha a izquierda hacia el punto* a.

Proposición 6.12 (Segunda regla de L'Hospital) *Sean* $f(x)$ *y* $g(x)$ *dos funciones derivables en* (a, b) *y* $\dfrac{\mathrm{d}g}{\mathrm{d}x}(x) \neq 0$ *para toda* $x \in (a, b)$. *Si se tiene que* $\lim\limits_{x\to a} f(x) = \pm\infty$, $\lim\limits_{x\to a} g(x) = \pm\infty$ *y*

$$\lim_{x\to a} \frac{\dfrac{\mathrm{d}f}{\mathrm{d}x}(x)}{\dfrac{\mathrm{d}g}{\mathrm{d}x}(x)} = L \quad \text{con } L \in \mathbb{R} \text{ o } L = \pm\infty,$$

entonces

$$\lim_{x\to a} \frac{f(x)}{g(x)} = L.$$

Demostración. Consideremos $a < x < t$. Luego, en $[a, t]$ tendremos la estimación

$$\frac{f(x)-f(t)}{g(x)-g(t)} = \frac{\dfrac{\mathrm{d}f}{\mathrm{d}x}(c_{x,t})}{\dfrac{\mathrm{d}g}{\mathrm{d}x}(c_{x,t})}$$

donde $c_{t,x} \in (x, t)$. Haciendo uso de esta expresión, podemos escribir

$$\frac{f(x)}{g(x)} - L = \frac{\dfrac{\mathrm{d}f}{\mathrm{d}x}(c_{x,t})}{\dfrac{\mathrm{d}g}{\mathrm{d}x}(c_{x,t})} - L + \frac{1}{g(x)}\left[f(t) - g(t)\frac{\dfrac{\mathrm{d}f}{\mathrm{d}x}(c_{x,t})}{\dfrac{\mathrm{d}g}{\mathrm{d}x}(c_{x,t})} \right],$$

y tomando el valor absoluto, se tiene

$$\left|\frac{f(x)}{g(x)} - L\right| \leqslant \left|\frac{\frac{\mathrm{d}f}{\mathrm{d}x}(c_{x,t})}{\frac{\mathrm{d}g}{\mathrm{d}x}(c_{x,t})} - L\right| + \left|\frac{1}{g(x)}\right| \left|f(t) - g(t)\frac{\frac{\mathrm{d}f}{\mathrm{d}x}(c_{x,t})}{\frac{\mathrm{d}g}{\mathrm{d}x}(c_{x,t})}\right|.$$

Cuando t tiende al punto a, el punto intermedio $c_{x,t}$ tiende a a y se cumple

$$\lim_{x\to a}\left|\frac{\frac{\mathrm{d}f}{\mathrm{d}x}(c_{x,t})}{\frac{\mathrm{d}g}{\mathrm{d}x}(c_{x,t})} - L\right| = 0 \quad \text{por hipótesis},$$

$$\lim_{x\to a}\left|f(t) - g(t)\frac{\frac{\mathrm{d}f}{\mathrm{d}x}(c_{x,t})}{\frac{\mathrm{d}g}{\mathrm{d}x}(c_{x,t})}\right| = |f(t) - g(t)L|,$$

$$\lim_{x\to a}\left|\frac{1}{g(x)}\right| = 0,$$

lo cual muestra que

$$\lim_{x\to a}\frac{f(x)}{g(x)} - L = 0,$$

y la proposición queda probada. ∎

NOTA IMPORTANTE:

La proposición 6.12 es válida aún en el caso $a = \pm\infty$. Basta definir $f_1(x) = f\left(\frac{1}{x}\right)$ y $g_1(x) = g\left(\frac{1}{x}\right)$ y aplicar la Regla de L'Hospital para esas nuevas funciones cuando $x \to 0^+$ si $a = +\infty$ o $x \to 0^-$ si $a = -\infty$.

Ejemplo 6.10 Aplicando la regla de L'Hospital, los límites siguientes se calculan directamente:

a) $\lim_{x\to\infty} x \operatorname{sen}\frac{1}{x} = 1.$

Reescribiendo $x \operatorname{sen}\frac{1}{x} = \dfrac{\operatorname{sen}\frac{1}{x}}{\frac{1}{x}}$ y tomando en cuenta que

$$\lim_{x\to\infty}\frac{1}{x} = \lim_{x\to\infty}\operatorname{sen}\frac{1}{x} = 0,$$

tenemos que

$$\lim_{x\to\infty}\frac{\frac{\mathrm{d}\operatorname{sen}\frac{1}{x}}{\mathrm{d}x}}{\frac{\mathrm{d}\frac{1}{x}}{\mathrm{d}x}} = \lim_{x\to\infty}\frac{-\frac{1}{x^2}\cos\frac{1}{x}}{-\frac{1}{x^2}} = \lim_{x\to\infty}\cos\frac{1}{x} = 1.$$

Luego, aplicando la regla de l'Hospital, se tiene que $\lim\limits_{x\to\infty} x \operatorname{sen}\dfrac{1}{x} = 1$.

b) $\lim\limits_{x\to\infty} x(2\arctan x - \pi) = -2$.

Reescribiendo

$$x(2\arctan x - \pi) = \frac{2\arctan x - \pi}{\frac{1}{x}}$$

y tomando en cuenta que

$$\lim_{x\to\infty} \frac{1}{x} = \lim_{x\to\infty} (2\arctan x - \pi) = 0,$$

tenemos que

$$\lim_{x\to\infty} \frac{\dfrac{\mathrm{d}(2\arctan x - \pi)}{\mathrm{d}x}}{\dfrac{\mathrm{d}\frac{1}{x}}{\mathrm{d}x}} = \lim_{x\to\infty} \frac{\dfrac{2}{1+x^2}}{-\dfrac{1}{x^2}} = \lim_{x\to\infty} \frac{-2x^2}{1+x^2} = -2.$$

Luego, aplicando la regla de l'Hospital, se tiene que

$$\lim_{x\to\infty} x(2\arctan x - \pi) = -2. \qquad \triangleleft$$

Ejercicios y problemas del capítulo

1. Aplicando el teorema del valor medio, establezca la estimación

$$|\operatorname{sen} x - \operatorname{sen} y| \leqslant |x-y| \quad \text{para cualesquiera} \quad x, y \in \mathbb{R}.$$

2. Pruebe, utilizando el teorema del valor medio, que toda función con derivada continua en un intervalo cerrado $[a,b]$ es lipschitziana. (Sugerencia: recuerde que toda función continua en un intervalo cerrado y acotado es acotada.)

3. Si la gráfica de la función velocidad $v = v(t)$ de un automóvil, se ve cualitativamente como en la figura 6.9, dibuje la gráfica de la posición del auto con respecto al tiempo si la posición en el tiempo $t = 1$ era de 100 metros medidos a partir del inicio de la carretera. Describa el movimiento del auto.

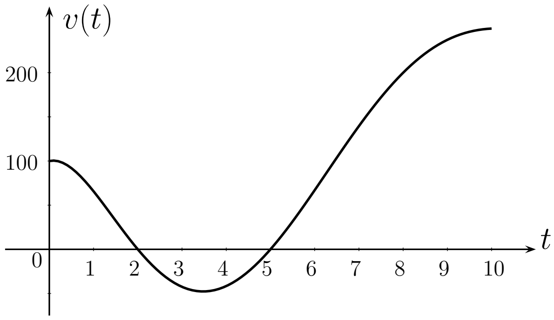


Figura 6.9 Gráfica de la función velocidad del ejercicio 3

4. Determine los intervalos de monotonía y los puntos máximos y mínimos locales de la función $f(x) = 7x^9 - 18x^7$. Con la información anterior, dibuje la gráfica de f.

5. Sean $P(x)$ y $Q(x)$ dos polinomios de grado 2 tales que

$$P(1) = Q(1), \quad \frac{\mathrm{d}P}{\mathrm{d}x}(1) = \frac{\mathrm{d}Q}{\mathrm{d}x}(1), \quad \frac{\mathrm{d}^2P}{\mathrm{d}x^2}(1) = \frac{\mathrm{d}^2Q}{\mathrm{d}x^2}(1).$$

   Pruebe que los dos polinomios son iguales.

6. Determine los puntos regulares, clasifique los puntos críticos y encuentre los puntos de inflexión de la función

$$f(x) = \frac{1}{5}x^5 + x^3 + x + 2.$$

7. Demuestre que si $f(x) : \mathbb{R} \to \mathbb{R}$ es una función cuya tercera derivada se anula en todos los puntos, entonces $f(x)$ es un polinomio de grado dos.

8. Haciendo uso del teorema de Taylor, calcule $\sqrt[3]{1.5}$ con tres cifras decimales.

9. Sea $f(x)$ una función continua en el intervalo $[-3,3]$ y tal que su primera y segunda funciones derivadas tienen las características señaladas.

$$\frac{\mathrm{d}f}{\mathrm{d}x}(x) : \begin{cases} > 0, & \text{si } x \in (-3,-1) \\ = 0, & \text{si } x = -1 \\ < 0, & \text{si } x \in (-1,1) \cup (1,3) \end{cases}$$

$$\frac{\mathrm{d}^2 f}{\mathrm{d}x^2}(x) : \begin{cases} < 0, & \text{si } x \in (-3,0) \cup (1,3) \\ = 0, & \text{si } x = 0 \\ > 0, & \text{si } x \in (0,1) \end{cases}$$

   (a) ¿En qué puntos de $[-3,3]$ tiene $f(x)$ máximos o mínimos locales?

   (b) ¿Qué puntos de $[-3,3]$ son puntos de inflexión?

   (c) Si $f(0) = 0$, dibuje la gráfica de $f(x)$ señalando sus intervalos de monotonía, sus concavidades, etc.

10. ¿En qué puntos de la curva $y = -x^3 + 3x^2 + 1$ tiene la pendiente de la tangente por ese punto el valor máximo?

11. De las ventanas con perímetro 10 m y cuya forma es la unión de un rectángulo y un semicírculo cuyo diámetro coincide con el lado superior del rectángulo, encuentre aquélla que deja pasar la mayor cantidad de luz.

12. Encuentre las dimensiones del cono circular recto de volumen mínimo que pueda contener a la esfera de radio 4.

13. Para cada uno de los casos siguientes, diga si existe una función dos veces derivable que satisfaga las propiedades enunciadas. (Justifique sus respuestas.)

   (a) $\dfrac{\mathrm{d}f}{\mathrm{d}x}(x) > 0$, $\dfrac{\mathrm{d}^2 f}{\mathrm{d}x^2}(x) > 0$ y $f(x) < 0$ para todo $x \in \mathbb{R}$.

   (b) $\dfrac{\mathrm{d}^2 f}{\mathrm{d}x^2}(x) > 0$ y $\dfrac{\mathrm{d}f}{\mathrm{d}x}(x) < 0$ para todo $x \in \mathbb{R}$.

   (c) $\dfrac{\mathrm{d}^2 f}{\mathrm{d}x^2}(x) > 0$ y $f(x) < 0$ para todo $x \in \mathbb{R}$.

14. Aplicando las reglas de L'Hospital, evalúe los límites siguientes:

   (a) $\displaystyle\lim_{x \to \pi} \frac{\operatorname{sen}^2 x}{x - \pi}$

   (b) $\displaystyle\lim_{x \to 0+} \frac{\operatorname{sen}^2 x}{x - \tan x}$

   (c) $\displaystyle\lim_{x \to 0} \frac{\operatorname{arcsen} x}{\arctan x}$

   (d) $\displaystyle\lim_{x \to 0} \frac{3 \operatorname{sen} x - \operatorname{sen} 3x}{3 \tan x - \tan 3x}$

15. Si $f(x)$ es una función con derivadas continuas de orden cuatro, diga cuáles de los siguientes enunciados son verdaderos y cuáles son falsos, dando en este último caso un ejemplo para probarlo.

    (a) Si f es no decreciente en (a, b), entonces $\dfrac{\mathrm{d}f}{\mathrm{d}x}(x) \geqslant 0$ para toda $x \in (a, b)$.

    (b) Si $\dfrac{\mathrm{d}f}{\mathrm{d}x}(x) \neq 0$ para toda $x \in (a, b)$, entonces f no tiene máximos ni mínimos en (a, b).

    (c) Si f tiene dos máximos locales en (a, b) entonces tiene un mínimo local en (a, b).

    (d) Si $x_0 \in (a, b)$ es un punto de inflexión de f, entonces $\dfrac{\mathrm{d}f}{\mathrm{d}x}(x_0) = 0$.

    (e) Si $\dfrac{\mathrm{d}f}{\mathrm{d}x}$ es creciente en (a, b), entonces f no tiene puntos de inflexión en (a, b).

    (f) Si $x_0 \in [a, b]$ y es mínimo absoluto de f en $[a, b]$, entonces $\dfrac{\mathrm{d}f}{\mathrm{d}x}(x_0) = 0$.

    (g) Si $\dfrac{\mathrm{d}^2 f}{\mathrm{d}x^2} > 0$ en (a, b) entonces f es una función monótona en (a, b).

    (h) Si $\dfrac{\mathrm{d}^4 f}{\mathrm{d}x^4} = 0$ en $\mathbb{R}$, entonces f no es acotada en $\mathbb{R}$.

Capítulo 7

La función exponencial y sus aplicaciones

Así como introducimos algunas familias de funciones elementales recurriendo a relaciones de carácter algebraico, geométrico o trigonométrico, hay algunas funciones cuya presentación es más natural a partir de la ley de cambio o de movimiento que satisfacen. Un ejemplo relevante es el caso de la llamada función exponencial, que se define como aquella función $f(x)$ cuya derivada en cada punto es igual al valor de la función en ese mismo punto,

$$\frac{\mathrm{d}f}{\mathrm{d}x}(x) = f(x),$$

y además su valor en $x = 0$ es $f(0) = 1$.

En las aplicaciones, distintos fenómenos y modelos dinámicos involucran funciones cuya razón de cambio en cada punto depende proporcionalmente del valor de la función en ese punto. Como ejemplos más destacados, se tienen los llamados fenómenos de difusión de calor o de epidemias y los de desintegración y decaimiento. En este capítulo, basados en lo que hemos estudiado sobre las funciones derivables, presentamos estos modelos y las propiedades de las funciones exponenciales y sus inversas.

7.1 La función exponencial

Analicemos las propiedades principales de las funciones que satisfacen la ecuación

$$\frac{\mathrm{d}f}{\mathrm{d}x}(x) = f(x). \tag{7.1}$$

La cuestión que es necesario establecer primero, es la existencia de soluciones de la ecuación diferencial (7.1) distintas de la solución $f(x) = 0$, denominada *solución trivial*. Esto constituye el enunciado del llamado *teorema de existencia de ecuaciones diferenciales* cuya demostración está más allá del alcance de este texto. En vista de lo anterior, supondremos de antemano la existencia de soluciones de la ecuación diferencial (7.1), definidas en todo $\mathbb{R}$ y distintas de la trivial. Bajo la suposición anterior, deduciremos las principales propiedades de esas funciones y a partir de ellas daremos su forma explícita, que nos permitirá calcular su valor en cada punto de $\mathbb{R}$.

Proposición 7.1 *El conjunto de todas las funciones definidas en $\mathbb{R}$ que satisfacen la ecuación diferencial (7.1) tiene las propiedades siguientes:*

1. Existe una solución de (7.1) distinta de la solución trivial $f(x) = 0$ para $x \in \mathbb{R}$.

2. Si $f_1(x)$ y $f_2(x)$ satisfacen la ecuación (7.1), entonces también la satisface toda función de la forma $h(x) = af_1(x) + f_2(x)$ con $a \in \mathbb{R}$.

3. Si $f(x)$ satisface (7.1) y $f(x_0) = 0$ para algún punto $x_0 \in \mathbb{R}$, entonces $f(x) = 0$ para toda $x \in \mathbb{R}$.

4. (Unicidad de soluciones de (7.1).) Si $f_1(x)$ y $f_2(x)$ satisfacen (7.1) y $f_1(x_0) = f_2(x_0)$ para algún punto x_0, entonces $f_1(x) = f_2(x)$ para toda $x \in \mathbb{R}$.

Demostración. La validez del punto 1 es consecuencia del teorema de existencia para ecuaciones diferenciales. Para probar el punto 2, nótese que

$$\begin{aligned}\frac{\mathrm{d}h}{\mathrm{d}x}(x) &= \frac{\mathrm{d}}{\mathrm{d}x}(af_1(x) + f_2(x)) = a\frac{df_1}{\mathrm{d}x}(x) + \frac{\mathrm{d}f_2}{\mathrm{d}x}(x)\\ &= af_1(x) + f_2(x) = h(x).\end{aligned}$$

Lo anterior nos permite afirmar que si $f(x)$ es la solución a (7.1) con $f(0) = 1$, entonces la función $h(x) = af(x)$ será solución de (7.1) con $h(0) = a$.

Para demostrar el punto 3, consideremos primero los puntos $x \in \mathbb{R}$ con $|x - x_0| < 1$. Aplicando el teorema del valor medio a la función solución f en el intervalo $[x_0, x]$, se tiene

$$f(x) = f(x) - f(x_0) = \frac{\mathrm{d}f}{\mathrm{d}x}(x_1)(x - x_0) = f(x_1)(x - x_0), \tag{7.2}$$

con $x_0 < x_1 < x$.

Aplicando de nuevo el teorema del valor medio a la misma función $f(x)$ en el intervalo $[x_0, x_1]$, obtenemos

$$f(x_1) = f(x_1) - f(x_0) = \frac{\mathrm{d}f}{\mathrm{d}x}(x_2)(x_1 - x_0) = f(x_2)(x_1 - x_0), \tag{7.3}$$

con $x_0 < x_2 < x_1$.

Sustituyendo (7.3) en (7.2) se obtiene que

$$f(x) = f(x_2)(x_1 - x_0)(x - x_0).$$

Repitiendo este argumento k veces, encontramos puntos $x_1, x_2, \ldots, x_k$, tales que $x_0 < x_i < x$ para $i = 1, \ldots, k$ y

$$f(x) = f(x_k)(x_{k-1} - x_0)(x_{k-2} - x_0) \cdots (x_1 - x_0).$$

Tomando valor absoluto, se tiene

$$|f(x)| \leqslant M|x - x_0|^{k-1},$$

donde M es una cota para $f(x)$ en el intervalo $[x_0, x]$, la cual existe pues hemos supuesto que $f(x)$ tiene derivada en ese intervalo y por lo tanto es continua y está acotada en $[x_0, x]$. Como $|x - x_0| < 1$ se tendrá que $|x - x_0|^k \to 0$ cuando $k \to \infty$, por lo que $f(x) = 0$ para cada x con $|x - x_0| < 1$, y por la continuidad de $f(x)$ en $\mathbb{R}$, se tendrá que $f(x_0 + 1) = 0$. Repitiendo los argumentos anteriores y cambiando el punto x_0 por el punto $x_0 + 1$, probamos igualmente que $f(x) = 0$ para cada x en $[x_0, x_0 + 2]$. Prosiguiendo de esta forma, concluimos que $f(x) = 0$ para toda $x \in \mathbb{R}$.

Para la prueba del punto 4, observemos que la función $h(x) = (f_1 - f_2)(x)$ es también, en virtud del punto 2, solución de (7.1) y se anula en el punto x_0, luego, como consecuencia del punto 3, se tiene $f_1(x) = f_2(x)$ para toda $x \in \mathbb{R}$. ■

Nota Importante:

1. *Conjuntamente, los puntos 2, 3 y 4 de la proposición 7.1 aseguran que cada solución no trivial de (7.1) genera todas las soluciones al multiplicarse por un número real arbitrario.*
2. *Si* $f(x)$ *es solución de (7.1), su función derivada* $\dfrac{\mathrm{d}f}{\mathrm{d}x}(x)$ *también lo es.*

A partir de la proposición 7.1, damos la definición siguiente:

Definición 7.1 *A la única solución de (7.1) tal que* $f(0) = 1$ *se le llama* ***función exponencial*** *y su valor en el punto* $x \in \mathbb{R}$ *lo denotaremos por* $\exp x$. *El valor* $\exp 1$ *se denota con* e, *notación introducida por Leonhard Euler.*

Proposición 7.2 (Propiedades de la función exponencial)

1. *La función exponencial es siempre mayor que cero, monótona creciente y cóncava hacia arriba.*
2. *La función exponencial tiene las propiedades siguientes:*
   a) $\exp(x + x_0) = \exp x_0 \exp x$ *para cualesquiera* $x, x_0 \in \mathbb{R}$.
   b) *Para cada* $x \in \mathbb{R}$ *se tiene*
   $$\exp x = \lim_{k\to\infty} \left(1 + x + \frac{1}{2!}x^2 + \frac{1}{3!}x^3 + \cdots + \frac{1}{k!}x^k\right).$$
   c) $\displaystyle\lim_{x\to\infty} \frac{x^k}{\exp x} = 0$, *para cada* k *natural.*

Demostración. El enunciado en 1 es verdadero ya que por la proposición 7.1, ninguna solución de (7.1) distinta de cero en un punto puede anularse en punto alguno y, por lo tanto, si $f(0) = 1$ se tiene $f(x) > 0$ para toda x. Por otro lado, dado que la función exponencial es igual a su derivada, ésta también será siempre positiva y por consiguiente f es una función creciente. Finalmente, la segunda derivada también será positiva y entonces la gráfica de f sera cóncava hacia arriba, como se muestra en la figura 7.1.

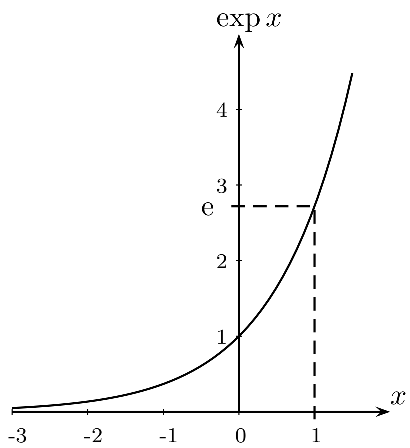


Figura 7.1 La función exponencial $f(x) = \exp x$

Para probar 2.a), tenemos que para cada x_0, la función $g(x) = \exp(x + x_0)$ es también solución de (7.1), ya que aplicando la regla de la cadena se tiene

$$\frac{\mathrm{d}g}{\mathrm{d}x}(x) = \frac{\mathrm{d}}{\mathrm{d}x}\exp(x + x_0) = \exp(x + x_0) = g(x)$$

y, además, $g(0) = \exp x_0$. Como la función $\exp x_0 \exp x$, en virtud de lo demostrado en el punto 2 de la proposición 7.1, es también solución de (7.1) y coincide con $g(x)$ en el punto x_0, por la unicidad de soluciones con un mismo valor en un punto dado, tendremos

$$g(x) = \exp(x + x_0) = \exp x_0 \exp x \quad \text{para toda } x \in \mathbb{R}.$$

La prueba de 2.b) se sigue del hecho de que cada solución de (7.1) tiene derivadas de todos los órdenes, y entonces, escribiendo su desarrollo de Taylor de orden n alrededor de $x = 0$, se tiene

$$\exp x = 1 + x + \frac{1}{2!}x^2 + \cdots + \frac{1}{n!}x^n + R_n(x)$$

$$\text{con} \quad |R_n(x)| \leqslant \frac{1}{n!}\exp x_c |x|^{n+1},$$

donde $x_c \in [-x, x]$. Por otro lado, al ser $\exp x$ una función continua en $[-x, x]$, existe $M > 0$ tal que

$$\exp y < M \quad \text{para toda} \quad y \in [-x, x].$$

Entonces, para cada etiqueta n,

$$|R_n(x)| \leqslant \frac{M}{n!}|x|^{n+1}$$

y

$$\left| \exp x - \left(1 + x + \frac{1}{2!}x^2 + \cdots + \frac{1}{n!}x^n\right) \right| \leqslant \frac{M}{n!}|x|^{n+1}.$$

Por otro lado, la sucesión $\{a_n\}_{n=1}^{\infty} = \left\{ \frac{|x|^{n+1}}{n!} \right\}_{n=1}^{\infty}$ converge a cero, ya que $a_{n+1} = \frac{|x|}{n+1} a_n$ para toda n, y entonces para toda etiqueta n tal que $n+1 > 2|x|$, se tendrá

$$a_{n+1} \leqslant \frac{1}{2} a_n.$$

Por tanto, la sucesión $\{a_n\}_{n=1}^{\infty}$ converge a zero. Tomando en cuenta lo anterior, se tiene

$$\exp x = \lim_{n\to\infty} \sum_{k=0}^{n} \frac{1}{k!} x^k,$$

y con ello se prueba 2.b).

Para demostrar 2.c), usamos la segunda regla de L'Hospital como sigue.

Para $k = 1$, se tiene

$$\lim_{x\to\infty} x = \lim_{x\to\infty} \exp x = \infty$$

y

$$\lim_{x\to\infty} \frac{\frac{\mathrm{d}x}{\mathrm{d}x}}{\frac{\mathrm{d}\exp x}{\mathrm{d}x}} = \lim_{x\to\infty} \frac{1}{\exp x} = 0,$$

luego, por la segunda regla de L'Hospital, se tiene que $\lim_{x\to\infty} \frac{x}{\exp x} = 0$.

Análogamente, para $k = 2$, tenemos

$$\lim_{x\to\infty} \frac{\mathrm{d}x^2}{\mathrm{d}x}(x) = \lim_{x\to\infty} \exp x = \infty$$

y

$$\lim_{x\to\infty} \frac{2x}{\exp x} = 2 \lim_{x\to\infty} \frac{x}{\exp x} = 0,$$

por el caso anterior. Luego, aplicando nuevamente la regla de L'Hospital, se obtiene $\lim_{x\to\infty} \frac{x^2}{\exp x} = 0$. Repitiendo este argumento k veces, podemos afirmar que para todo $k \in \mathbb{N}$ se tiene $\lim_{x\to\infty} \frac{x^k}{\exp x} = 0$. ■

Nota Importante:

1. *La interpretación geométrica de la propiedad 2.c) de la proposición 7.2 es que la función exponencial crece más rápidamente que cualquier potencia de x cuando $x \to \infty$.*

2. *De 2.a) de la proposición 7.2 se sigue que*

$$\begin{aligned}\exp n &= (\exp 1)^n \quad \textit{para cada } n \in \mathbb{N},\\ \exp\left(\frac{n}{m}\right) &= \exp(1)^{\frac{n}{m}} = \sqrt[m]{(\exp 1)^n} \quad \textit{para cualesquiera } n,\ m\ \in \mathbb{N}.\end{aligned}$$

3. *De 2.b) se tiene que*

$$\mathrm{e} = \exp 1 = \lim_{n\to\infty} \sum_{k=0}^{n} \frac{1}{k!}.$$

4. *Tomando en cuenta que la función exponencial comparte varias propiedades con la operación exponenciación, se acostumbra denotarla también con el símbolo* e^x*; es decir,*

$$\mathrm{e}^x = \exp x, \quad x \in \mathbb{R}.$$

7.2 La función logaritmo natural

La función exponencial tiene función inversa definida en el conjunto de los reales positivos. A su función inversa se le llama *función logaritmo natural* y se denota

$$\begin{aligned}\ln &: (0,\infty) \to \mathbb{R}\\ \ln x &= \exp^{-1} x.\end{aligned}$$

La función $\ln x$ tiene las propiedades siguientes:

1. $\ln(xy) = \ln x + \ln y$
2. $\dfrac{\mathrm{d}\ln}{\mathrm{d}x}(x) = \dfrac{1}{x}$
3. Es creciente y cóncava hacia abajo en toda la semirrecta $(0,\infty)$
4. $\ln(1+x) = x - \frac{1}{2}x^2 + \frac{1}{3}x^3 - \cdots (-1)^{k-1}\frac{1}{k}x^k + R_k(x)$.

   con $R_k(x) = (-1)^k(1+x_c)^{-k-1}(x-x_c)^k x$, donde $x_c \in (0,x)$.

La gráfica de la función $f(x) = \ln x$ se muestra en la figura 7.2.

Nota Importante:

En la propiedad 4, el residuo $R_k(x)$ tiende a cero cuando $k \to \infty$ para $x \in (-1,0)$

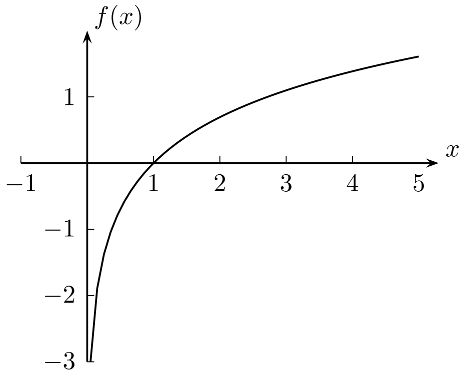


Figura 7.2 La función logaritmo natural $f(x) = \ln x$

y la sucesión de sumas converge al valor de $\ln(1 + x)$. *Por otro lado, tomando en cuenta que* $\ln x = -\ln \dfrac{1}{x}$ *se concluye que la fórmula representa el logaritmo natural para todo punto de su dominio.*

7.3 Funciones de tipo exponencial

En general, se definen las *funciones de tipo exponencial* como aquéllas de la forma

$$f(x) = \exp a(x) = e^{a(x)},$$

donde $a(x)$ es una función derivable.

Entre las funciones de tipo exponencial se distinguen las siguientes:

1. Para cada $r \in \mathbb{R}$, $x^r : (0, \infty) \to \mathbb{R}$, $\quad x^r \stackrel{\text{def}}{=} \mathrm{e}^{r \ln x}$.
2. Para cada $a > 0$, $a^x : (0, \infty) \to \mathbb{R}$, $\quad a^x \stackrel{\text{def}}{=} \mathrm{e}^{x \ln a}$.
3. $x^x : (0, \infty) \to \mathbb{R}$, $\quad x^x \stackrel{\text{def}}{=} \mathrm{e}^{x \ln x}$.
4. En general, si $f : \mathbb{R} \to \mathbb{R}$ y $g : \mathbb{R} \to \mathbb{R}$ son funciones derivables y $f(x) > 0$ para toda $x \in \mathbb{R}$, se define $f^g : \mathbb{R} \to \mathbb{R}$ mediante $f^g(x) \stackrel{\text{def}}{=} \mathrm{e}^{g(x) \ln f(x)}$.

7.4 Aplicaciones de la función exponencial

En este apartado presentamos tres ejemplos relevantes de leyes dinámicas que dan lugar a comportamientos de tipo exponencial.

Ejemplo 7.1 (Dinámica de Poblaciones) Denotemos por $p(t)$ la función que representa el tamaño de una población en el tiempo t, y de forma simplificada, supongamos una tasa β de crecimiento constante, es decir, cada 100 unidades de población dan lugar por reproducción a β nuevas unidades de población por unidad de tiempo. Supongamos además una política de remoción que consista en retirar (por muertes o emigraciones) de la población en el tiempo t un total de $g(t)$ unidades de población por unidad de tiempo. En este caso la ley de cambio de la población toma la forma

$$\frac{\mathrm{d}}{\mathrm{d}t}p(t) = \alpha p(t) - g(t) \quad \text{con} \quad \alpha = \frac{\beta}{100}.$$

Si la razón de retiro es constante, $g(t) = \rho$, describa la evolución de la población a partir de la población inicial. Calcule el tiempo que llevará para que la población inicial se duplique.

Solución:
Para buscar la solución $p(t)$ de la ecuación $\frac{\mathrm{d}}{\mathrm{d}t}p(t) = \alpha p(t) - \rho$, podemos multiplicar ambos lados de la ecuación por el *factor integrante* $\mathrm{e}^{-\alpha t}$,

$$\mathrm{e}^{-\alpha t}\frac{\mathrm{d}}{\mathrm{d}t}p(t) - \mathrm{e}^{-\alpha t}p(t) = -\mathrm{e}^{-\alpha t}\rho,$$

lo cual nos permite reescribir la ecuación en la forma

$$\frac{\mathrm{d}}{\mathrm{d}t}(\mathrm{e}^{-\alpha t}p(t)) = \frac{\mathrm{d}}{\mathrm{d}t}\frac{\rho}{\alpha}\mathrm{e}^{-\alpha t},$$

de donde se obtiene

$$p(t) = c\mathrm{e}^{\alpha t} + \frac{\rho}{\alpha},$$

con

$$c = p(0) - \frac{\rho}{\alpha}.$$

Entonces

$$p(t) = \left(p(0) - \frac{\rho}{\alpha}\right)\mathrm{e}^{\alpha t} + \frac{\rho}{\alpha}. \tag{7.4}$$

Existen tres escenarios con respecto al comportamiento de la población.

1. Si $p(0) - \frac{\rho}{\alpha} > 0$ entonces la población crecerá exponencialmente y se duplicará en un tiempo t_0 dado por $t_0 = \frac{1}{\alpha}\ln\frac{2p(0) - \rho/\alpha}{p(0) - \rho/\alpha}$. Por ejemplo, supongamos que inicialmente hay 10 individuos ($p_0 = 10$) y que la tasa de reproducción es $\beta = 20$. Si la razón de retiro es $\rho = 1$, entonces la población en el tiempo t es $p(t) = 5\mathrm{e}^{0.2t} + 5$ y el instante en el que la población se duplica es $t_0 = 5\ln 3$. Ver figura 7.3(a).

2. Si $p(0) - \rho/\alpha = 0$, la población es constante, $p(t) = \rho/\alpha$. Por ejemplo, si la población inicial y la tasa de reproducción se mantienen en $p_0 = 10$ y $\beta = 20$ pero la razón de retiro aumenta a $\rho = 2$, entonces la población se mantendrá constante, $p(t) = p(0) = 10$. Ver figura 7.3(b)

3. Si $p(0) - \dfrac{\rho}{\alpha} < 0$ la población se extinguirá en un tiempo finito $t_0 = \dfrac{1}{\alpha} \ln \dfrac{\frac{\rho}{\alpha}}{\frac{\rho}{\alpha} - p(0)}$. Por ejemplo, si la tasa de retiro sobrepasa el valor de 2, digamos $\rho = 3$, entonces $p(0) - \dfrac{\rho}{\alpha} = -5$. En este caso, la población en cada instante t está dada por $p(t) = -5(\mathrm{e}^{0.2t} - 3)$ y se extinguirá cuando $t_0 = 5 \ln 3$. Ver figura 7.3(c) ◁

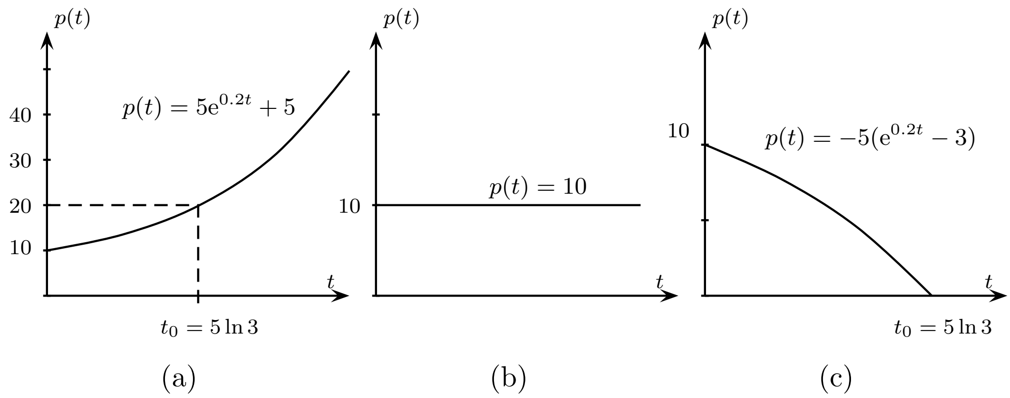


Figura 7.3 Comportamientos posibles de la función exponencial (7.4) para distintos valores de ρ, con $p(0)$ y α fijos

Ejemplo 7.2 (Ley de enfriamiento de Newton) Es un hecho conocido que la temperatura de un cuerpo de material homogéneo situado en un medio a temperatura constante, experimenta un proceso de enfriamiento o calentamiento que a la larga lo lleva a adquirir la misma temperatura del medio ambiente. La rapidez con la cual tiene lugar el enfriamiento o calentamiento en cada tiempo es proporcional a la diferencia entre la temperatura del cuerpo en ese instante y la temperatura del medio ambiente. A este hecho físico se le conoce como *ley del enfriamiento de Newton*. Nuestro problema consiste en deducir, a partir de esa ley, la temperatura del cuerpo como función del tiempo conocidas las temperaturas iniciales del cuerpo y del medio, así como la constante de difusión térmica del material que forma el cuerpo en cuestión. Si la temperatura inicial del cuerpo es tres veces mayor que la temperatura ambiente, ¿en qué tiempo será el doble de la del medio ambiente?

Solución:

Si denotamos por $T(t)$ la función de temperatura del cuerpo en el tiempo t, por T_m la temperatura del medio ambiente y si $\rho > 0$ es la constante de difusión térmica del material, la ley de enfriamiento estipula que

$$\frac{\mathrm{d}T}{\mathrm{d}t}(t) = \rho(T_m - T(t)).$$

Introduzcamos la función $f(t) = T(t) - T_m$; entonces f satisface la ecuación

$$\frac{\mathrm{d}f}{\mathrm{d}t}(t) = -\rho f(t),$$

cuyas soluciones son de la forma

$$f(t) = f(0)\exp(-\rho t).$$

Entonces, en términos de la función $T(t)$, tendremos

$$T(t) = T_m + (T(0) - T_m)\exp(-\rho t). \tag{7.5}$$

De la fórmula (7.5) se deduce que si la temperatura inicial del cuerpo, $T(0)$, es mayor que la temperatura del medio ambiente, aquélla decrecerá exponencialmente con el paso del tiempo y tenderá al valor T_m de la temperatura ambiente. Por ejemplo, si $T(0) = 3T_m$, de (7.5) tendremos que $T(t) = 2T_m$ si

$$2T_m = T_m + 2T_m\exp(-\rho t).$$

Luego, el tiempo t_0 en el cual la temperatura del cuerpo será el doble de la temperatura del medio ambiente es

$$t_0 = -\frac{1}{\rho}\ln\frac{1}{2}.$$

◁

Ejemplo 7.3 (La ecuación logística) En la naturaleza, la dinámica de una población está determinada básicamente por su tasa de crecimiento, que en general depende tanto de la especie misma como de las limitantes del medio representadas por la disponibilidad de alimentos y la existencia de depredadores. Si suponemos que no existen depredadores y el medio sólo puede sostener a un número fijo L de unidades de población, se considera entonces que la tasa α de crecimiento es una función del tiempo $\alpha(t)$ directamente proporcional a la diferencia entre la máxima población sustentable L y la población existente en el tiempo t, es decir

$$\alpha(t) = k(L - p(t)),$$

donde k es una constante. En este caso, la dinámica de crecimiento de la población $p(t)$ es gobernada por la ecuación diferencial

$$\frac{\mathrm{d}p}{\mathrm{d}t}(t) = \beta\left(1 - \frac{p(t)}{L}\right)p(t), \tag{7.6}$$

donde $\beta = \dfrac{Lk}{100}$. A la ecuación anterior se le llama *ecuación logística.*

Se puede verificar (ver ejercicio 9) que la función

$$p(t) = \frac{Lp(0)}{p(0) + (L - p(0))e^{-kt}}$$

es solución de la ecuación (7.6). Note que

$$\lim_{t \to \infty} p(t) = L.$$

Por esta razón, en el modelo (7.6), el número L se interpreta como la máxima población sustentable por el medio. $\triangleleft$

Ejercicios y problemas del capítulo

1. Encuentre, en términos de la función exponencial, la función $f(x)$ tal que $f(0) = 2$ y
$$\frac{\mathrm{d}f}{\mathrm{d}x}(x) = af(x), \text{ con } a \in \mathbb{R}.$$

2. Usando las propiedades de las funciones exponenciales, derive las siguientes fórmulas y propiedades.

   (a) $\dfrac{\mathrm{d}x^r}{\mathrm{d}x}(x) = rx^{r-1}$

   (b) $\dfrac{\mathrm{d}a^x}{\mathrm{d}x}(x) = (\ln a)a^x$

   (c) $\dfrac{\mathrm{d}x^x}{\mathrm{d}x}(x) = (1 + \ln x)x^x$

   (d) $\dfrac{\mathrm{d}}{\mathrm{d}x}(f(x)^{g(x)})(x) = \left(\dfrac{\mathrm{d}g}{\mathrm{d}x}(x) \ln f(x) + \dfrac{g(x)}{f(x)}\right) f(x)^{g(x)}$

3. Observe que las funciones x^r, a^x, x^x y $f(x)^{g(x)}$ son todas no-negativas. ¿Cuáles de ellas son uno a uno?

4. (*Funciones hiperbólicas*) La función exponencial e^x se descompone como suma de las funciones $\cosh x \overset{\text{def}}{=} \frac{1}{2}(e^x + e^{-x})$ y $\text{senh} x \overset{\text{def}}{=} \frac{1}{2}(e^x - e^{-x})$, llamadas *coseno hiperbólico* y *seno hiperbólico de x*, respectivamente, cuyas gráficas aparecen en la figura 7.4.

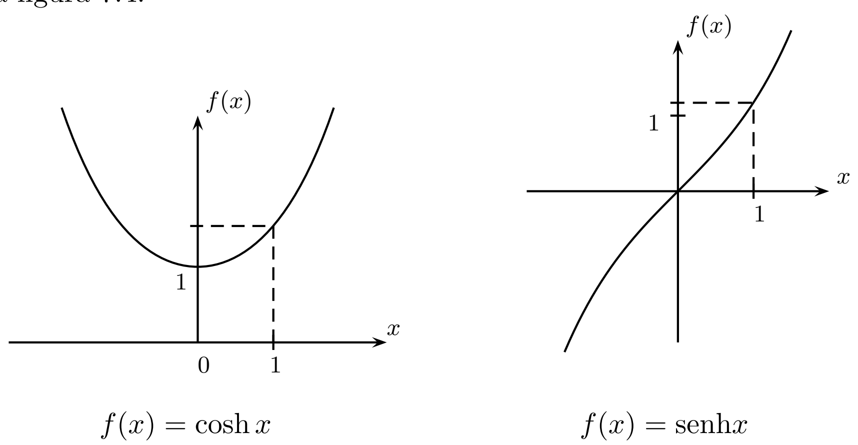


Figura 7.4 Las funciones hiperbólicas
($\cosh(1) \approx 1.543$, $\text{senh}(1) \approx 1.175$)

Pruebe las afirmaciones siguientes:

   (a) $\cosh x$ es una función par y $\text{senh} x$ es una función impar.

(b) $\cosh^2 x - \mathrm{senh}^2 x = 1$

(c) $\dfrac{\mathrm{d}\cosh x}{\mathrm{d}x} = \mathrm{senh} x, \quad \dfrac{\mathrm{dsenh} x}{\mathrm{d}x} = \cosh x$

(d) $\cosh(x+y) = \cosh x \cosh y + \mathrm{senh} x \,\mathrm{senh}\, y$

$\mathrm{senh}(x+y) = \mathrm{senh} x \cosh y + \mathrm{senh} y \cosh x$

(e) Diga en qué intervalo la función $\cosh x$ es uno a uno y demuestre que su función inversa, denotada $\cosh^{-1} y$, tiene la expresión siguiente:

$$\cosh^{-1} y = \ln(y + \sqrt{y^2 - 1}) \quad \text{para } y \geqslant 1.$$

5. Muestre que las funciones $\cosh kx$ y $\mathrm{senh} kx$ satisfacen ambas la ecuación diferencial $\dfrac{\mathrm{d}^2 y}{\mathrm{d}x^2}(x) - k^2 y(x) = 0$.

6. A partir de la función exponencial, encuentre una función de la forma $h(x) = \exp(f(x))$ tal que
$$\frac{\mathrm{d}h}{\mathrm{d}x}(x) = xh(x).$$

7. Una curva pasa por el punto $(2,3)$ y la pendiente de su recta tangente en cada punto es igual al doble de la ordenada del punto. Encuentre la ecuación de la curva.

8. Sea $x \geqslant 0$ y n un número natural. Aplicando el teorema del valor medio, muestre que
$$\ln\left(1 + \frac{x}{n}\right)^n = \frac{nx}{n + x^*},$$
donde $x^* \in [0, x]$. Haciendo uso de esta estimación, demuestre que
$$\lim_{n\to\infty} \ln\left(1 + \frac{x}{n}\right)^n = x$$
y deduzca la fórmula
$$\mathrm{e}^x = \lim_{n\to\infty} (1 + \frac{x}{n})^n.$$
Esta última expresión es, a veces, usada como definición de la función e^x.

9. (a) Sustituyendo directamente en (7.6), muestre que la función
$$p(t) = \frac{Lp(0)}{p(0) + (L - p(0))\mathrm{e}^{-kt}}$$
satisface esa ecuación. Note que cuando el tiempo crece, la población tiende a la población máxima que el medio puede sustentar.

(b) Haciendo uso de la solución de la ecuación logística, diga en qué tiempo se duplica la población $p(t)$ si la población inicial $p(0)$ es igual a $L/3$.

(c) Encuentre el tiempo en que la población alcanza su valor máximo.

10. (Fechamiento por carbono radioactivo.) Las plantas y los animales adquieren de la atmósfera o del alimento que ingieren un isótopo radioactivo del carbono denominado *carbono catorce*. Este isótopo es tal que por desintegración radioactiva pierde la mitad de su masa en un tiempo de 5730 años. Las plantas y los animales vivos restituyen continuamente el carbono catorce emitido, dejando de hacerlo en cuanto mueren, de tal manera que a partir de entonces el carbono empieza a perder su masa con rapidez directamente proporcional a la masa presente en cada instante. Calcule la constante de proporcionalidad y diga qué edad tiene un pergamino que ha perdido por radioactividad el 74% de la masa que actualmente tiene el carbono catorce en el tipo de plantas de las cuales se hizo ese pergamino.

Capítulo 8

La integral indefinida

En la presencia de fenómenos de cambio o movimiento, es a veces más viable conocer o deducir la ley de cambio a la que obedece la variación relativa de las variables involucradas, que la función misma entre esas variables. Es decir, a veces se conoce la derivada de la función o relaciones que satisfacen las derivadas, pero no se conoce la función misma. Por ejemplo, en el caso del movimiento de un automóvil, es a menudo más fácil estimar la velocidad o la aceleración durante un cierto intervalo de tiempo, que la función de posición del vehículo en cada instante. Una idea de la velocidad se puede tener, por ejemplo, observando el velocímetro desde dentro del mismo vehículo. Algo similar se tiene en el caso del movimiento que muestran los cuerpos ante la presencia de una fuerza externa y que se manifiesta, según las leyes del movimiento de Newton, en términos de variaciones de la velocidad del cuerpo con respecto al tiempo en forma proporcional a la magnitud y dirección de la fuerza actuante. En este caso, el problema consiste en deducir la posición del cuerpo con respecto al tiempo a partir del comportamiento de su segunda derivada.

Al problema de determinar la forma y los valores de una función a partir del conocimiento de su derivada o de una ecuación que involucra a sus derivadas se le llama problema de integración y es el problema fundamental de la teoría de las ecuaciones diferenciales. En este sentido, el problema de integración es el problema inverso al de derivación o de cálculo de derivadas.

En este capítulo se inicia el estudio de los problemas de integración a partir del concepto de integral indefinida y se muestra cómo las distintas reglas de derivación dan lugar a métodos de integración que permiten resolver problemas como los arriba citados.

8.1 Antiderivadas e integrales indefinidas

Sea una función $g(x)$ definida y continua en un intervalo abierto (o cerrado) I. Nos preguntamos sobre las funciones $f(x)$ definidas en I que tienen como función derivada a la función $g(x)$. Lo anterior se plantea en términos de la ecuación diferencial

$$\frac{\mathrm{d}f}{\mathrm{d}x}(x) = g(x), \;\; x \in I, \tag{8.1}$$

donde la incógnita o indeterminada es la función $f(x)$. En caso de existir, cada función $f(x)$ que satisface la ecuación (8.1) se llama *antiderivada* o *primitiva de* $g(x)$ en el intervalo I.

Nota Importante:
Si $f_1(x)$ es una antiderivada de la función $g(x)$ en un intervalo I, también lo es la función

$$f_2(x) = f_1(x) + a,$$

donde a es cualquier número real. El recíproco del resultado anterior es también cierto, ya que si $f_1(x)$ y $f_2(x)$ son antiderivadas de la función $g(x)$ en el intervalo I, se tiene

$$\frac{\mathrm{d}(f_2 - f_1)}{\mathrm{d}x}(x) = \frac{\mathrm{d}f_2}{\mathrm{d}x}(x) - \frac{\mathrm{d}f_1}{\mathrm{d}x}(x) = g(x) - g(x) = 0,$$

lo que implica, en virtud del teorema del valor medio, que $f_2(x)-f_1(x) = a =$constante para toda $x \in I$.

Si $g(x)$ es una función definida en un intervalo abierto o cerrado I, a la familia de antiderivadas o primitivas en I se le denomina la *integral indefinida de $g(x)$ en I*, y se denota por el símbolo

$$\int g(x)\mathrm{d}x.$$

En otros términos, la integral indefinida $\int g(x)\mathrm{d}x$ de una función $g(x)$ es la familia de funciones

$$\int g(x)\mathrm{d}x = \left\{ f : I \to \mathbb{R} \quad \text{tales que} \quad \frac{\mathrm{d}f}{\mathrm{d}x}(x) = g(x) \right\}. \tag{8.2}$$

De la definición de antiderivada se desprenden las afirmaciones siguientes.

1. Si f_1 y f_2 están en la familia $\int g(x)\mathrm{d}x$, entonces $f_1(x) - f_2(x)$ es una función constante. Esto implica que si $f_0(x)$ es una antiderivada de $g(x)$, entonces la familia (8.2) es
$$\int g(x)\mathrm{d}x = \left\{ f_0(x) + c, \quad c \in \mathbb{R} \right\}.$$
2. La antiderivada de la función $f(x) = 0$ es la familia
$$\int 0\,\mathrm{d}x = \left\{ c, \quad \text{con} \quad c \in \mathbb{R} \right\}.$$
3. Si $g_1(x)$ y $g_2(x)$ son antiderivadas de una misma función, entonces $g_1(x) = g_2(x) + c$, con $c \in \mathbb{R}$.

4. La integral indefinida de la función derivada $\frac{\mathrm{d}f}{\mathrm{d}x}(x)$ es la familia de las funciones $f(x)+c$, con $c \in \mathbb{R}$, es decir

$$\left(\int \frac{\mathrm{d}f}{\mathrm{d}x}(x)\mathrm{d}x\right)(y) = \{f(y)+c, \quad c \in \mathbb{R}\}\,. \tag{8.3}$$

Debido a (8.3) se dice que el cálculo de la integral indefinida de una función es la operación inversa (o recíproca) de la operación de derivación.

Notación: Denotaremos con $\int g(x)\mathrm{d}x$ no a la familia (8.2) de antiderivadas de g sino a una antiderivada (es decir, a algún elemento de (8.2)) y para no sobrecargar la notación, en lugar de escribir $\left(\int g(x)\mathrm{d}x\right)(x) = h(x)$ escribiremos simplemente $\int g(x)\mathrm{d}x = h(x)$. Usando este convenio, escribiremos (8.3) en la forma

$$\int \frac{\mathrm{d}f}{\mathrm{d}x}(x)\mathrm{d}x = f(x)+c.$$

Nota Importante:

Para determinar unívocamente una de las funciones primitivas o antiderivadas a que da lugar la integral indefinida de una función, basta fijar el valor de esa primitiva en un punto. En este caso la constante queda determinada y, por lo tanto, la función primitiva que se buscaba.

Ejemplo 8.1 Determinemos la antiderivada $f(x)$ de la función $g(x) = 3x^2 + x$ tal que $f(1) = 3$.

En este caso, la antiderivada $f(x)$ es un elemento de la integral indefinida

$$f(x) \in \int (3x^2+x)\mathrm{d}x = \left\{x^3 + \frac{1}{2}x^2 + c\right\},$$

y para fijar el valor de c que corresponde al elemento $f(x)$ de esa familia tal que $f(1)=3$, la constante c deberá ser tal que

$$f(1) = \frac{3}{2} + c = 3,$$

es decir,

$$c = \frac{3}{2}$$

y

$$f(x) = x^3 + \frac{1}{2}x^2 + \frac{3}{2}. \qquad \triangleleft$$

A continuación presentamos la integral indefinida de algunas funciones elementales $g(x)$.

Función $g(x)$	Integral indefinida $\int g(x)\mathrm{d}x$
a	$ax + c$
$a_0 + a_1 x + \cdots + a_n x^n$	$a_0 x + \frac{1}{2}a_1 x^2 + \cdots + \frac{1}{n+1}a_n x^{n+1} + c$
$(x+b)^{\frac{m}{n}}$ $\quad b \in \mathbb{R}, \quad m, n$ enteros	$\frac{n}{m+n}(x+b)^{\frac{m+n}{n}} + c$
$\operatorname{sen}(bx) \quad b \in \mathbb{R}$	$-\frac{1}{b}\cos(bx) + c$
$\cos(bx) \quad b \in \mathbb{R}$	$\frac{1}{b}\operatorname{sen}(bx) + c$

Ejemplo 8.2 La función $g(x) : \mathbb{R} \setminus \{0\} \to \mathbb{R}$ con $g(x) = \dfrac{1}{x}$ tiene por integral indefinida en $(-\infty, 0)$ (o en $(0, \infty)$) a las funciones

$$\int \frac{1}{x}\mathrm{d}x = \ln|x| + c. \qquad \triangleleft$$

Ejemplo 8.3 La integral indefinida de $g(x) = |3x - 5|$, es la familia de funciones

$$\int |3x - 5|\mathrm{d}x = f(x) + c,$$

donde

$$f(x) = \begin{cases} 5x - \dfrac{3}{2}x^2 & \text{si } x < \dfrac{5}{3} \\ \dfrac{3}{2}x^2 - 5x + \dfrac{25}{3} & \text{si } x \geq \dfrac{5}{3}. \end{cases}$$

La antiderivada o función primitiva $h(x)$ de $g(x) = |3x - 5|$ en $\mathbb{R}$ tal que $h(1) = 2$, es

$$h(x) = \begin{cases} 5x - \dfrac{3}{2}x^2 - \dfrac{3}{2} & \text{si } x < \dfrac{5}{3} \\ \dfrac{3}{2}x^2 - 5x \ + \dfrac{41}{6} & \text{si } x \geq \dfrac{5}{3}. \end{cases} \qquad \triangleleft$$

8.2 Métodos de integración

Dado que los procesos de derivación y de cálculo de la integral indefinida son operaciones inversas, cada regla o fórmula de derivación da lugar a una regla o método para el cálculo de la integral indefinida de funciones continuas. A estos métodos se les conoce como *métodos de integración.*

Antes de desarrollar los métodos de integración inducidos por las reglas de derivación para productos y composición de funciones que merecen una atención detallada, enunciamos primeramente un par de propiedades que son consecuencia directa de las reglas de derivación (5.7) y (5.5), respectivamente.

1. Si $c \in \mathbb{R}$ y f es una función que tiene antiderivada, entonces la función cf tiene antiderivada y
$$\int cf(x)\mathrm{d}x = c\int f(x)\mathrm{d}x.$$
2. Si f y g son dos funciones que tienen antiderivada entonces la función $f+g$ tiene antiderivada y
$$\int (f(x)+g(x))\mathrm{d}x = \int f(x)\mathrm{d}x + \int g(x)\mathrm{d}x.$$

8.2.1 Integración por partes

La fórmula de Leibniz para la derivación de un producto de funciones (fórmula (5.6) da lugar al llamado *método de integración por partes,*[1] que presentamos a continuación.

Proposición 8.1 (Integración por partes) *Sean $p(x)$ y $q(x)$ funciones derivables. Entonces*
$$\int p(x)\frac{\mathrm{d}q}{\mathrm{d}x}(x)\mathrm{d}x = p(x)q(x) - \int \frac{\mathrm{d}p}{\mathrm{d}x}(x)\, q(x)\mathrm{d}x + c.$$

Demostración. Derivando el producto $(pq)(x) = p(x)q(x)$, tenemos
$$\frac{\mathrm{d}(pq)}{\mathrm{d}x}(x) = p(x)\frac{\mathrm{d}q}{\mathrm{d}x}(x) + q(x)\frac{\mathrm{d}p}{\mathrm{d}x}(x).$$
Entonces,
$$p(x)q(x) = \int p(x)\frac{\mathrm{d}q}{\mathrm{d}x}(x) + \int q(x)\frac{\mathrm{d}p}{\mathrm{d}x}(x) + c,$$
de donde se sigue el resultado. ■

Nota Importante:
El método de integración por partes se aplica al cálculo de la integral indefinida de productos de dos funciones cuando la integral indefinida de una de ellas es conocida. En tal caso, el método reduce el problema al cálculo de la integral indefinida de otro producto de funciones que en muchos casos es más fácil de resolver.

[1]Este método fue desarrollado por Brook Taylor.

Ejemplo 8.4 Calculemos la integral indefinida de las funciones

1. $g(x) = x \operatorname{sen} x$,
2. $g(x) = \cos^2 x$,
3. $g(x) = e^{ax} \cos bx$ con $a \neq 0$ y $b \neq 0$.

1. Haciendo $p(x) = x$ y $q(x) = -\cos x$, tenemos $x \operatorname{sen} x = p(x)\dfrac{\mathrm{d}q}{\mathrm{d}x}(x)$ y aplicando la fórmula de integración por partes, se obtiene

$$\begin{aligned}\int (x \operatorname{sen} x)\mathrm{d}x &= \int (p(x)\frac{\mathrm{d}}{\mathrm{d}x}q(x)\mathrm{d}x \\ &= -x\cos x + \int \cos x \mathrm{d}x \\ &= -x\cos x + \operatorname{sen} x + c.\end{aligned}$$

2. Haciendo $p(x) = \cos x$ y $q(x) = \operatorname{sen} x$, se tiene $\cos^2 x = p(x)\dfrac{\mathrm{d}q}{\mathrm{d}x}(x)$ y aplicando la fórmula de integración por partes obtenemos

$$\begin{aligned}\int \cos^2 x \mathrm{d}x &= \int \cos x \frac{\mathrm{d} \operatorname{sen} x}{\mathrm{d}x}\mathrm{d}x \\ &= \cos x \operatorname{sen} x + \int \operatorname{sen}^2 x \mathrm{d}x = \\ &= \cos x \operatorname{sen} x + \int (1-\cos^2 x)\mathrm{d}x,\end{aligned}$$

de donde

$$\int \cos^2 x \mathrm{d}x = \frac{1}{2}(\cos x \operatorname{sen} x + x) + c.$$

3. Haciendo $p(x) = \mathrm{e}^{ax}$ y $q(x) = \dfrac{1}{b}\operatorname{sen} bx$, y aplicando la fórmula de integración por partes, se tiene

$$\int \mathrm{e}^{ax}\cos bx \mathrm{d}x = \int p(x)\left(\frac{\mathrm{d}q}{\mathrm{d}x}(x)\right)\mathrm{d}x = \frac{1}{b}\mathrm{e}^{ax}\operatorname{sen} bx - \frac{a}{b}\int \mathrm{e}^{ax}\operatorname{sen} bx \mathrm{d}x. \quad (8.4)$$

Calculando ahora $\int \mathrm{e}^{ax}\operatorname{sen} bx \mathrm{d}x$, aplicando nuevamente el método de integración por partes y haciendo

$$u(x) = \mathrm{e}^{ax} \quad \text{y} \quad v(x) = -\frac{1}{b}\cos bx,$$

tenemos

$$\int \mathrm{e}^{ax}\operatorname{sen} bx \mathrm{d}x = \int u(x)\frac{dv}{\mathrm{d}x}(x)\mathrm{d}x = -\frac{1}{b}\mathrm{e}^{ax}\cos bx + \frac{a}{b}\int \mathrm{e}^{ax}\cos bx \mathrm{d}x. \quad (8.5)$$

Sustituyendo (8.5) en (8.4),

$$\int \mathrm{e}^{ax}\cos bx \mathrm{d}x = \frac{1}{b}\mathrm{e}^{ax}\operatorname{sen} bx - \frac{a}{b}\left[-\frac{1}{b}\mathrm{e}^{ax}\cos bx + \frac{a}{b}\int \mathrm{e}^{ax}\cos bx \mathrm{d}x\right],$$

y agrupando términos, se tiene

$$\left(1+\frac{a^2}{b^2}\right)\int \mathrm{e}^{ax}\cos bx \mathrm{d}x = \frac{1}{b}\mathrm{e}^{ax}\operatorname{sen} bx + \frac{a}{b^2}(\mathrm{e}^{ax}\cos bx).$$

Finalmente,

$$\int \mathrm{e}^{ax}\cos bx = \frac{\mathrm{e}^{ax}}{a^2+b^2}\Big(b\operatorname{sen} bx + a\cos bx\Big) + c. \qquad \triangleleft$$

8.2.2 Integración por sustitución

La regla de la cadena o de derivación de una composición de funciones da lugar al *método de integración por sustitución,* que a continuación presentamos.

Proposición 8.2 (Integración por sustitución) *Si $g(x)$ y $h(y)$ son derivables y $f(y)$ es una antiderivada de $h(y)$, entonces*

$$\int h(g(x))\left(\frac{\mathrm{d}g}{\mathrm{d}x}(x)\right)\mathrm{d}x = \left(\int h(y)\mathrm{d}y\right)\circ g(x) = f(g(x)) + c.$$

Demostración. Efectivamente, si derivamos la función $f(g(x))+c$ usando la regla de la cadena y consideramos que

$$\int h(y)\mathrm{d}y = f(y) + c \text{ para } y\in I,$$

obtendremos

$$\frac{\mathrm{d}}{\mathrm{d}x}(f(g(x))+c) = \frac{\mathrm{d}f}{\mathrm{d}y}(g(x))\cdot\frac{\mathrm{d}g}{\mathrm{d}x}(x) = h(g(x))\frac{\mathrm{d}g}{\mathrm{d}x}(x),$$

lo que prueba la fórmula de integración por sustitución. ■

Nota Importante:
El método de integración por sustitución también se puede escribir en la forma

$$\int h(y)\mathrm{d}y = \left(\int h(g(x))\big(\frac{\mathrm{d}g}{\mathrm{d}x}(x)\big)\mathrm{d}x\right)\circ(g^{-1}(y))$$

donde se supone que $g(x)$ es una función monótona. En este caso, cuando se presenta el problema de calcular la integral indefinida de una función de la forma $h(g(x))\dfrac{\mathrm{d}g}{\mathrm{d}x}(x)$, se hace la sustitución formal, o simbólica,

$$y = g(x), \qquad \mathrm{d}y = \frac{\mathrm{d}g}{\mathrm{d}x}(x)\mathrm{d}x,$$

para escribir directamente

$$\int h(g(x))\left(\frac{\mathrm{d}g}{\mathrm{d}x}\right)\mathrm{d}x = \left(\int h(y)\mathrm{d}y\right) \circ g(x),$$

y reducir así el cálculo a la integral $\displaystyle\int h(y)\mathrm{d}y$.

Ejemplo 8.5 Para calcular la integral indefinida de la función

$$h(x) = (x+1)\operatorname{sen}(x^2+2x),$$

observamos la presencia de la composición de la función $\operatorname{sen} y$ con la función $g(x) = (x^2+2x)$. Luego, haciendo la sustitución

$$y = x^2 + 2x, \qquad \mathrm{d}y = (2x+2)\mathrm{d}x,$$

podemos escribir

$$\begin{aligned}\int h(x)\mathrm{d}x &= \frac{1}{2}\int \operatorname{sen}(g(x))\frac{\mathrm{d}g}{\mathrm{d}x}\mathrm{d}x = \frac{1}{2}\left(\int \operatorname{sen} y\mathrm{d}y\right) \circ (g(x)) = \\ &= -\frac{1}{2}\cos(g(x)) + c = -\frac{1}{2}\cos(x^2+2x) + c,\end{aligned}$$

ya que la función $-\cos y$ es una antiderivada de $\operatorname{sen} y$. ◁

Ejemplo 8.6 La integral indefinida $\int \operatorname{sen} x \cos^m x\mathrm{d}x$ se obtiene haciendo la sustitución

$$y = \cos x, \qquad \mathrm{d}y = -\operatorname{sen} x\mathrm{d}x,$$

y escribiendo

$$\begin{aligned}\int \operatorname{sen} x\cos^m x\mathrm{d}x &= -\left(\int y^m\mathrm{d}y\right) \circ (g(x)) \\ &= \left(-\frac{y^{m+1}}{m+1}\right) \circ (\cos x) \\ &= -\frac{1}{m+1}\cos^{m+1} x + c.\end{aligned}$$

◁

Ejemplo 8.7 La función $f(x)$ tal que

$$\frac{\mathrm{d}f}{\mathrm{d}x}(x) = \operatorname{arcsen}(x) \quad \text{y} \quad f(0) = 2$$

se determina aplicando primero el método de integración por partes en la forma

$$\int (\operatorname{arcsen} x)\mathrm{d}x = \int (\operatorname{arcsen} x)\frac{\mathrm{d}x}{\mathrm{d}x}\mathrm{d}x = x \operatorname{arcsen} x - \int \left(x \frac{\mathrm{d} \operatorname{arcsen} x}{\mathrm{d}x} \right)\mathrm{d}x =$$
$$= x \operatorname{arcsen} x - \int \frac{x}{\sqrt{1-x^2}}\mathrm{d}x$$

y calculando la integral indefinida del último término mediante la sustitución

$$y = 1 - x^2, \qquad \mathrm{d}y = -2x\mathrm{d}x,$$

para obtener

$$\int \frac{x}{\sqrt{1-x^2}}\mathrm{d}x = \left(-\frac{1}{2}\int \frac{1}{\sqrt{y}}\mathrm{d}y \right) \circ (1-x^2) =$$
$$= -\sqrt{1-x^2} + c.$$

Luego, la integral indefinida de $\operatorname{arcsen}(x)$ es

$$\int (\operatorname{arcsen} x)\mathrm{d}x = x \operatorname{arcsen} x + \sqrt{1-x^2} + c.$$

Finalmente, tomando la constante $c = 1$ tendremos la solución de la ecuación diferencial

$$f(x) = x\operatorname{arcsen} x + \sqrt{1-x^2} + 1.$$

Nota Importante:
Si $f(x)$ es una primitiva de $g(x)$ y $z(x)$ es una primitiva de $f(x)$ entonces

$$\frac{d^2 z}{\mathrm{d}x^2}(x) = g(x).$$

Lo anterior muestra cómo determinar, mediante el cálculo reiterado de integrales indefinidas, las funciones que tienen como segunda derivada una función previamente dada.

Ejemplo 8.8 El cálculo de la función $z(x)$ tal que

$$\frac{\mathrm{d}^2 z}{\mathrm{d}x^2}(x) = x + \operatorname{sen} x \quad \text{con} \quad z(0) = 1 \text{ y } \frac{\mathrm{d}z}{\mathrm{d}x}(0) = 1,$$

se obtiene encontrando primeramente la integral indefinida de $g(x) = x + \operatorname{sen} x$,

$$\int g(x)\mathrm{d}x = \frac{1}{2}x^2 - \cos x + c$$

donde c es una constante y, enseguida, la integral indefinida de la antiderivada anterior, para obtener

$$z(x) = \int \left(\int g(x)\mathrm{d}x \right) \mathrm{d}x = \frac{1}{6}x^3 - \operatorname{sen} x + cx + b,$$

con b constante. Al requerir que $z(0) = 1$ y $\dfrac{dz}{\mathrm{d}x}(0) = 1$, se obtiene

$$z(0) = b = 1$$

y

$$\frac{dz}{\mathrm{d}x}(0) = -\cos 0 + c = 1.$$

Entonces la función buscada es

$$z(x) = \frac{1}{6}x^3 - \operatorname{sen} x + 2x + 1. \qquad \triangleleft$$

8.2.3 Integración por sustitución trigonométrica

Para la familia de funciones de la forma $h(\sqrt{1-x^2})$, $h(\sqrt{1+x^2})$ o $h(\sqrt{x^2-1})$ donde $h(y)$ es una función continua, el método de sustitución nos permite remitir el cálculo de su integral indefinida a la integral de una función trigonométrica. Presentamos a continuación esos casos.

Primer caso: Sea $h(y)$ una función continua. Para calcular la integral indefinida de la forma $\int h(\sqrt{1-x^2})\mathrm{d}x$, se hace la sustitución

$$\theta(x) = \operatorname{arcsen} x, \quad x \in (-1, 1)$$

y se reescribe la integral indefinida en la forma

$$\int h(\sqrt{1-x^2})\mathrm{d}x = \int h(\cos\theta(x))\mathrm{d}x. \tag{8.6}$$

Tomando en cuenta que

$$\frac{d\theta}{\mathrm{d}x} = \frac{1}{\sqrt{1-x^2}} = \frac{1}{\cos\theta(x)},$$

escribimos la integral indefinida en el lado derecho de (8.6) en la forma

$$\int h(\cos\theta(x))\mathrm{d}x = \int h(\cos\theta(x))\cos\theta(x)\frac{d\theta}{\mathrm{d}x}\mathrm{d}x.$$

Aplicando la fórmula de integración por sustitución, obtenemos

$$\boxed{\int h(\sqrt{1-x^2})\mathrm{d}x = \left(\int h(\cos\theta)\cos\theta \mathrm{d}\theta \right) \circ (\operatorname{arcsen} x). \qquad (8.7)}$$

Ejemplo 8.9 La integral indefinida

$$\int \frac{\sqrt{9-x^2}}{x^2}\mathrm{d}x$$

se obtiene mediante sustitución trigonométrica. Primero hacemos la sustitución

$$y = \frac{1}{3}x, \qquad \mathrm{d}y = \frac{1}{3}\mathrm{d}x,$$

para obtener

$$\int \frac{\sqrt{9-x^2}}{x^2}\mathrm{d}x = \frac{1}{3}\int \frac{\sqrt{1-\left(\frac{1}{3}x\right)^2}}{\left(\frac{1}{3}x\right)^2}\mathrm{d}x = \left(\int \frac{\sqrt{1-x^2}}{x^2}\mathrm{d}x\right)\circ\left(\frac{x}{3}\right),$$

y la integral a calcular toma la forma

$$\int \frac{\sqrt{1-x^2}}{x^2}\mathrm{d}x = \int h(\sqrt{1-x^2})\mathrm{d}x \quad \text{ con } \quad h(y) = \frac{y}{1-y^2}.$$

Aplicando ahora la fórmula (8.7) se tiene

$$\int \frac{\sqrt{1-x^2}}{x^2}\mathrm{d}x = \left(\int \frac{\cos^2\theta}{\operatorname{sen}^2\theta}d\theta\right)\circ(\operatorname{arcsen} x) = -\cot(\operatorname{arcsen} x) - \arcsin x + c$$

y, finalmente,

$$\int \frac{\sqrt{9-x^2}}{x^2}\mathrm{d}x = -\frac{\sqrt{9-x^2}}{x} - \operatorname{arcsen}\frac{x}{3} + c. \qquad \triangleleft$$

Segundo caso: Sea $h(y)$ una función continua. Para calcular una integral indefinida de la forma $\int h(\sqrt{1+x^2})\mathrm{d}x$, se hace la sustitución

$$\theta(x) = \arctan x, \quad x \in (-\infty, \infty)$$

y se reescribe la integral indefinida en la forma

$$\int h(\sqrt{1+x^2})\mathrm{d}x = \int h\left(\sqrt{1+\tan^2\theta(x)}\right)\mathrm{d}x = \int h(\sec\theta(x))\mathrm{d}x.$$

Tomando en cuenta que

$$\frac{d\theta}{\mathrm{d}x} = \frac{1}{1+x^2} = \cos^2\theta(x),$$

escribimos la última integral indefinida en la forma

$$\int h(\sec\theta(x))\mathrm{d}x = \int h(\sec\theta(x))\sec^2\theta(x)\frac{d\theta}{\mathrm{d}x}\mathrm{d}x,$$

y aplicando la fórmula de integración por sustitución se tiene

$$\int h(\sqrt{1+x^2})\mathrm{d}x = \left(\int h(\sec\theta)\sec^2\theta d\theta\right)\circ(\arctan x). \tag{8.8}$$

Ejemplo 8.10 Para calcular la integral indefinida

$$\int \frac{1}{x^2\sqrt{x^2+4}}\mathrm{d}x,$$

hacemos la sustitución

$$y = \frac{x}{2}, \quad \mathrm{d}y = \frac{1}{2}\mathrm{d}x$$

y reescribimos la integral en la forma

$$\int \frac{1}{x^2\sqrt{x^2+4}}\mathrm{d}x = \int \frac{1}{4\left(\frac{x}{2}\right)^2\sqrt{\left(\frac{x}{2}\right)^2+1}}\mathrm{d}\left(\frac{x}{2}\right) = \frac{1}{4}\left(\int \frac{1}{x^2\sqrt{x^2+1}}\mathrm{d}x\right)\circ\left(\frac{x}{2}\right).$$

Entonces, la antiderivada que debemos calcular es

$$\int \frac{1}{x^2\sqrt{x^2+1}}\mathrm{d}x = \int h(\sqrt{x^2+1})\mathrm{d}x, \quad \text{con} \quad h(y) = \frac{1}{y(y^2-1)}.$$

Aplicando la fórmula (8.8) obtenemos

$$\begin{aligned}
\int \frac{1}{x^2\sqrt{x^2+1}}\mathrm{d}x &= \int h(\sqrt{x^2+1})\mathrm{d}x = \left(\int \frac{\sec^2\theta}{\tan^2\theta\sec\theta}\mathrm{d}\theta\right)\circ(\arctan x)\\
&= \left(\int \frac{\cos\theta}{\operatorname{sen}^2\theta}\mathrm{d}\theta\right)\circ(\arctan x) = -\frac{1}{\operatorname{sen}\theta}\circ(\arctan x)\\
&= -\frac{\sqrt{1+x^2}}{x} + c,
\end{aligned}$$

por lo que, finalmente,

$$\int \frac{1}{x^2\sqrt{x^2+4}}\mathrm{d}x = \frac{1}{4}\left(\int \frac{1}{x^2\sqrt{x^2+1}}\mathrm{d}x\right)\circ\left(\frac{x}{2}\right) = -\frac{\sqrt{4+x^2}}{4x} + c. \qquad \triangleleft$$

Tercer caso: Sea $h(y)$ una función continua. Para calcular la integral indefinida de la forma $\int h(\sqrt{x^2-1})\mathrm{d}x$, se hace la sustitución

$$\theta(x) = \operatorname{arcsec} x, \quad x \in (-1, 1)$$

y se reescribe la integral indefinida en la forma

$$\int h(\sqrt{x^2-1})\mathrm{d}x = \int h(\tan\theta(x))\mathrm{d}x.$$

Tomando en cuenta que

$$\frac{d\theta}{\mathrm{d}x} = \frac{d\,\mathrm{arcsec}\,x}{\mathrm{d}x} = \frac{1}{x\sqrt{x^2-1}} = \frac{1}{\sec\theta(x)\tan\theta(x)},$$

escribimos la última integral indefinida en la forma

$$\int h(\tan\theta(x))\mathrm{d}x = \int h(\tan\theta(x))\sec\theta(x)\tan\theta(x)\frac{d\theta}{\mathrm{d}x}\mathrm{d}x$$

y, aplicando la fórmula de integración por sustitución, se tiene

$$\boxed{\int h(\sqrt{x^2-1})\mathrm{d}x = \left(\int h(\tan\theta)\sec\theta\tan\theta d\theta\right)\circ(\mathrm{arcsec}\,x). \qquad (8.9)}$$

Ejemplo 8.11 Para calcular la integral indefinida $\displaystyle\int \frac{1}{\sqrt{x^2-a^2}}\mathrm{d}x$ con $a>0$, hacemos la sustitución

$$y = \frac{1}{a}x, \qquad \mathrm{d}y = \frac{1}{a}\mathrm{d}x,$$

y encontramos que

$$\int \frac{1}{\sqrt{x^2-a^2}}\mathrm{d}x = \int \frac{1}{a\sqrt{\left(\frac{x}{a}\right)^2-1}}\mathrm{d}x = \left(\int \frac{1}{\sqrt{x^2-1}}\mathrm{d}x\right)\circ\left(\frac{x}{a}\right).$$

Enseguida nos remitimos al cálculo de la integral indefinida $\displaystyle\int \frac{1}{\sqrt{x^2-1}}\mathrm{d}x$, que es de la forma

$$\int \frac{1}{\sqrt{x^2-1}}\mathrm{d}x = \int h(\sqrt{x^2-1})\mathrm{d}x \quad \text{con} \quad h(y) = \frac{1}{y}.$$

Aplicando la fórmula (8.9), se tiene

$$\begin{aligned}\int \frac{1}{\sqrt{x^2-1}}\mathrm{d}x &= \left(\int \sec\theta d\theta\right)\circ(\mathrm{arcsec}\,x) = \\ &= \ln|\sec\theta+\tan\theta|\circ(\mathrm{arcsec}\,x) = \\ &= \ln|x+\sqrt{x^2-1}|\end{aligned}$$

y, finalmente,

$$\int \frac{1}{\sqrt{x^2-a^2}}\mathrm{d}x = \ln|x+\sqrt{x^2-a^2}| - \ln a + c. \qquad \triangleleft$$

8.2.4 Integración de funciones racionales

En esta subsección presentamos el llamado *método de fracciones parciales* para el cálculo de la integral indefinida de funciones racionales. Una función racional es una función de la forma
$$r(x) = \frac{p(x)}{q(x)},$$
donde $p(x)$ y $q(x)$ son polinomios. La función racional $r(x)$, puede reescribirse, después de aplicar el algoritmo de la división para polinomios, en la forma

$$r(x) = w(x) + \frac{s(x)}{t(x)}$$

donde $w(x), s(x)$ y $t(x)$ son polinomios y el grado de $s(x)$ es menor que el grado de $t(x)$. Lo anterior nos permite reducir el cálculo de la integral indefinida de funciones racionales a aquéllas para las que el grado del polinomio del numerador es estrictamente menor que el grado del polinomio del denominador.

Los teoremas siguientes sobre polinomios y funciones racionales son dos resultados importantes de álgebra, que permiten resolver el problema del cálculo de antiderivadas de funciones racionales.

Teorema 8.3 (Fundamental del álgebra) *Cada polinomio $t(x)$ con coeficientes en los números reales se factoriza en la forma*

$$\begin{aligned} t(x) = c(x-a_1)^{m_1}(x-a_2)^{m_2}\cdots(x-a_s)^{m_s}\cdot \\ \cdot[(x-b_1)^2+k_1^2]^{r_1}[(x-b_2)^2+k_2^2]^{r_2}\cdots[(x-b_q)^2+k_q^2]^{r_q}, \end{aligned}$$

donde
$$m_1 + m_2 \cdots + m_s + 2r_1 + \cdots + 2r_q = \mathit{grado}\ t(x).$$
Note que los factores de la forma $[(x-b)^2+k^2]^{r_i}$ son polinomios de segundo grado con raíces complejas.

Teorema 8.4 *Sean $s(x)$ y $t(x)$ polinomios tales que el grado de $s(x)$ es menor que el grado de $t(x)$. Entonces la función racional $r(x) = \dfrac{s(x)}{t(x)}$ puede expresarse como suma de funciones racionales de la forma*

$$\frac{e}{(x-a)^s} \quad o \quad \frac{cx+d}{[(x-b)^2+k^2]^r}, \tag{8.10}$$

donde $a, b, c, d, k, e \in \mathbb{R}$ y $r, s \in \mathbb{N}$. A las funciones (8.10) se les denomina ***fracciones parciales****. Más aún, por cada factor del polinomio $t(x)$ de la forma $(x-a)^m$ se tienen m sumandos de la forma*

$$\frac{e_1}{(x-a)} + \frac{e_2}{(x-a)^2} + \cdots + \frac{e_m}{(x-a)^m},$$

y cada uno de los términos de la forma $[(x-b)^2+k^2]^r$ *da lugar a* r *sumandos de la forma*

$$\frac{c_1x+d_1}{[(x-b)^2+k^2]}+\frac{c_2x+d_2}{[(x-b)^2+k^2]^2}+\cdots+\frac{c_rx+d_r}{[(x-b)^2+k^2]^r}.$$

Basado en la descomposición anterior, el teorema siguiente caracteriza la integral indefinida de toda función racional.

Teorema 8.5 *La integral indefinida de una función racional es una suma de funciones de la forma*

$$\int\frac{s(x)}{t(x)}\mathrm{d}x=U(x)+a\ln V(x)+b\arctan W(x)+c$$

donde $U(x),V(x)$ *y* $W(x)$ *son funciones racionales y* $a,b,c\in\mathbb{R}$.

Demostración Tomando en cuenta que toda función racional se puede expresar como suma de fracciones parciales de la forma $\dfrac{a}{(x-b)^s}$ y $\dfrac{cx+d}{[(x-e)^2+k^2]^r}$ donde s,r son números naturales y a,b,c,d,e,k son constantes reales, la integral indefinida de esas funciones se reduce al cálculo de la integral indefinida de las fracciones parciales. Para la evaluación de $\displaystyle\int\frac{a}{(x-b)^s}\mathrm{d}x$ se tiene, directamente por sustitución,

$$\int\frac{a}{(x-b)^s}\mathrm{d}x=\begin{cases} a\ln|x-b| \text{ si } s=1,\\ \dfrac{a}{1-s}\dfrac{1}{(x-b)^{s-1}} \text{ si } s>1.\end{cases}$$

El cálculo de la integral indefinida para los términos de la forma

$$\frac{cx+d}{[(x-e)^2+k^2]^r}$$

se reduce a calcular primitivas para funciones racionales de la forma

$$\frac{1}{[(x-e)^2+k^2]^r}\quad \text{o}\quad \frac{x-e}{[(x-e)^2+k^2]^r}.$$

Para el cálculo de la integral indefinida de las funciones de forma

$$\frac{1}{[(x-e)^2+k^2]^r},$$

se procede definiendo las integrales

$$I_j=\int\frac{1}{[(x-e)^2+k^2]^j}\mathrm{d}x,\quad j=1,2,\ldots,$$

las cuales se calculan recursivamente mediante las fórmulas

$$I_1 = \frac{1}{k^2}\arctan\left(\frac{x-e}{k}\right)$$

$$I_j = \frac{1}{2k^2(j-1)}\left[(2j-3)I_{j-1} + \frac{x-e}{[(x-e)^2+k^2]^{j-1}}\right], \quad j = 2, 3\ldots$$

Finalmente, la integral indefinida de las fracciones parciales de la forma

$$\frac{x-e}{[(x-e)^2+k^2]^r}$$

se calcula directamente sustituyendo $y = (x-e)^2+k^2$, $\mathrm{d}y = 2(x-e)\mathrm{d}x$, para obtener

$$\int \frac{x-e}{[(x-e)^2+k^2]^r}\mathrm{d}x = \frac{1}{2(1-r)}\frac{1}{[(x-e)^2+k^2]^{r-1}}. \qquad \blacksquare$$

Ejemplo 8.12 La integral indefinida de la función racional $r(x) = \dfrac{1}{x^3+1}$ se calcula con la descomposición en fracciones parciales

$$\frac{1}{x^3+1} = \frac{1}{(x+1)(x^2-x+1)} = \frac{e}{x+1} + \frac{cx+d}{x^2-x+1},$$

donde

$$c = -\frac{1}{3}, \quad d = \frac{2}{3}, \quad e = \frac{1}{3}.$$

Por lo tanto,

$$\int \frac{1}{x^3+1}\mathrm{d}x = \frac{1}{3}\int \frac{1}{x+1}\mathrm{d}x + \frac{-1}{3}\int \frac{x-2}{x^2-x+1}\mathrm{d}x =$$

$$= \frac{1}{3}\ln|x+1| - \frac{1}{6}\ln(x^2-x+1) + \frac{1}{\sqrt{3}}\arctan\left(\frac{2x-1}{\sqrt{3}}\right) + c. \qquad \triangleleft$$

Ejercicios y problemas del capítulo

1. Encuentre una antiderivada de la función $f(x) = xe^x$ ensayando con funciones de la forma $y(x) = Axe^x + Be^x$.

2. Aplicando el método de integración por partes, calcule las integrales indefinidas siguientes.

   (a) $\int \operatorname{arcsen} x \mathrm{d}x$

   (b) $\int \ln x \mathrm{d}x$

   (c) $\int x^2 \operatorname{sen} x \mathrm{d}x$

   (d) $\int \cos^4 x \mathrm{d}x$

3. Calcule las integrales indefinidas siguientes.

   (a) $\int |x| \mathrm{d}x$

   (b) $\int (|x-1| + |2x+1|) \mathrm{d}x$

   (c) $\displaystyle\int \left(\sum_{i=1}^{k} a_i x^i \right) \mathrm{d}x$

4. Calcule las integrales indefinidas siguientes.

   (a) $\displaystyle\int \frac{x^2+1}{x^3+8} \mathrm{d}x$

   (b) $\displaystyle\int \frac{x}{(x^2-x+1)^2} \mathrm{d}x$

   (c) $\displaystyle\int \frac{1}{x^4-a^4} \mathrm{d}x$

   (d) $\displaystyle\int \frac{1}{\mathrm{e}^{2x} - 4\mathrm{e}^x + 4} \mathrm{d}x$

5. Deduzca la fórmula recursiva siguiente:

$$\int \cos^n x \mathrm{d}x = \frac{1}{n} \cos^{n-1} x \operatorname{sen} x + \frac{n-1}{n} \int \cos^{n-2} x \mathrm{d}x, \quad n \text{ natural.}$$

6. Encuentre $y(x)$ tal que $y(1) = 0$, $\dfrac{\mathrm{d}y}{\mathrm{d}x}(1) = 1$ y $\dfrac{d^2y}{\mathrm{d}x}(x) = |x|$.

7. Dibuje la gráfica de las funciones que forman la integral indefinida de la función $f(x)$ cuya gráfica aparece en la figura 8.1.

8. Haciendo sustituciones trigonométricas, calcule las integrales indefinidas siguientes.

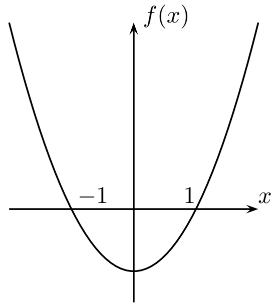


Figura 8.1 Gráfica de $f(x)$ para el ejercicio 7

(a) $\displaystyle\int \frac{x^2}{(x^2+a^2)^{\frac{3}{2}}}\mathrm{d}x$

(b) $\displaystyle\int x^3\sqrt{9-x^2}\mathrm{d}x$

(c) $\displaystyle\int x\sqrt{25+x^2}\mathrm{d}x$

9. Verifique la fórmulas para la integral indefinida de las funciones racionales siguientes:

(a) $\displaystyle\int \frac{1}{x^2+1}\mathrm{d}x = \arctan x + c$

(b) $\displaystyle\int \frac{x}{x^2+1}\mathrm{d}x = \int \frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}x}\ln(x^2+1)\mathrm{d}x = \ln\sqrt{x^2+1}+c$

(c) $\displaystyle\int \frac{1}{(x^2+1)^2}\mathrm{d}x = \int \frac{x^2+1-x^2}{(x^2+1)^2}\mathrm{d}x = \frac{1}{2}\left(\arctan x + \frac{x}{x^2+1}\right)+c$

(d) $\displaystyle\int \frac{x}{(x^2+1)^2}\mathrm{d}x = \int \frac{-1}{2}\frac{\mathrm{d}}{\mathrm{d}x}\frac{1}{(x^2+1)}\mathrm{d}x = \frac{-1}{2}\frac{1}{x^2+1}+c$

(e) $\displaystyle\int \frac{1}{a\,\mathrm{sen}\,x+b\cos x}\mathrm{d}x = \frac{1}{\sqrt{a^2+b^2}}\ln\left|\tan\frac{x+\arctan(b/a)}{2}\right|+c$

10. Haciendo la sustitución $y^2=\dfrac{x-q}{x-p}$, calcule la integral indefinida

$$\int\sqrt{(x-p)(x-q)}\mathrm{d}x.$$

11. Calcule las integrales indefinidas siguientes.

(a) $\displaystyle\int \frac{\mathrm{sen}(\ln x)}{x^2}\mathrm{d}x$

(b) $\displaystyle\int \frac{x^3}{\sqrt{1-x^2}}\mathrm{d}x$

(c) $\displaystyle\int \tan^4(\pi x)\mathrm{d}x$

(d) $\displaystyle\int \frac{1}{x^{\frac{1}{3}}-1}\mathrm{d}x$

12. Dibuje la gráfica de la función $f(x)$ si $f(1)=0$ y $\dfrac{\mathrm{d}f}{\mathrm{d}x}(1)=1$ y la gráfica de $\dfrac{\mathrm{d}^2f}{\mathrm{d}x^2}(x)$ es la que aparece en la figura 8.2

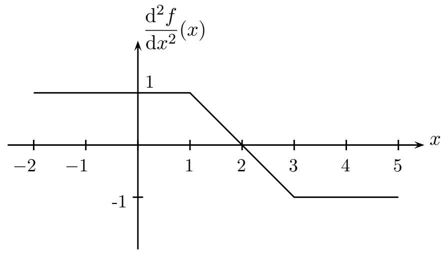


Figura 8.2 Gráfica de $\dfrac{\mathrm{d}^2f}{\mathrm{d}x^2}(x)$ para el ejercicio 12

13. Mediante el cálculo de la integral indefinida de la función $g(x)$, encuentre la función $f(x)$ tal que

$$\frac{\mathrm{d}f}{\mathrm{d}x}(x)=g(x)$$

y que además satisfaga la condición señalada a la derecha, para los siguientes casos:

(a) $g(x):(-3,\infty)\to\mathbb{R}, \quad g(x)=\dfrac{1}{x+3}$ y $f(0)=1$

(b) $g(x):(1,\infty)\to\mathbb{R}, \quad g(x)=\dfrac{x^2+5x+6}{x-1}$ y $f(2)=0$

(c) $g(x):\left(-\dfrac{\pi}{2},\dfrac{\pi}{2}\right)\to\mathbb{R}, \quad g(x)=\tan x$ y $f(0)=0.$

Capítulo 9

La integral definida

El otro concepto central del cálculo de funciones reales es el concepto de integral de una función sobre un intervalo. El proceso de integración permite "integrar o sumar" las variaciones infinitesimales de una función a lo largo de un intervalo para obtener la variación neta de la función en ese intervalo. En el caso particular del movimiento de una partícula, hace posible calcular el desplazamiento neto de la partícula en un intervalo de tiempo, a partir de las velocidades instantáneas mostradas durante ese intervalo.

Desde un enfoque geométrico, el valor de la integral de una función en un intervalo es igual al área de la región delimitada por su gráfica y el eje de las abscisas, considerando con signo negativo el área de la región que queda por debajo del eje. La relación entre los dos enfoques anteriores la proporciona el llamado "teorema fundamental del cálculo", al establecer que las operaciones de derivación e integración de funciones son procesos inversos.

Para introducir aquí la noción de integral de una función, se aplica el método de agotamiento para el cálculo del área bajo la gráfica de la función sobre un intervalo. Dicho método aproxima el área de un conjunto irregular mediante sumas de áreas de rectángulos, de tal manera que en el "límite" se alcanza el área exacta del conjunto en cuestión.

Primeramente introduciremos el concepto de integral para las funciones continuas no negativas definidas en un intervalo cerrado y acotado, y luego haremos algunas generalizaciones a funciones continuas por segmentos y a intervalos no acotados. En la presentación y justificación de los resultados y conceptos, utilizaremos fuertemente las propiedades básicas de las funciones continuas definidas en intervalos cerrados que quedaron establecidas al final del capítulo cuarto.

9.1 La definición de integral definida

Una *partición* $\mathcal{P}$ de un intervalo cerrado $[a, b]$ es un conjunto finito de puntos de $[a, b]$ que contiene a los extremos del intervalo $[a, b]$.

Cada partición admite dos orientaciones: una *orientación positiva,* dada por el orden que a la partición impone el orden de los números reales y cuyo elemento inicial es el extremo izquierdo del intervalo $[a, b]$, y que denotamos

$$a = x_0 < x_1 < \cdots < x_n = b,$$

y otra *orientación negativa,* que es contraria a la impuesta por el orden en los números reales y cuyo elemento inicial es el extremo derecho del intervalo $[a, b]$, y que denotamos

$$b = y_0 > y_1 > \cdots > y_n = a,$$

donde $y_0 = x_n, y_1 = x_{n-1}, \ldots, y_n = a.$

Una partición que consiste de $n + 1$ puntos divide al intervalo $[a, b]$ en n subintervalos $[x_{i-1}, x_i], i = 1, 2, \ldots, n$, de longitudes

$$\Delta x_i = x_i - x_{i-1}.$$

A la longitud máxima de los subintervalos $[x_{i-1}, x_i]$ se le llama *norma de la partición* $\mathcal{P}$, y se denota por $|\mathcal{P}|$.

Al conjunto de las particiones de $[a, b]$ lo denotaremos $\Omega([a, b])$.

Una partición $\mathcal{Q}$ de $[a, b]$ se dice un *refinamiento* de la partición $\mathcal{P}$ si $\mathcal{P} \subset \mathcal{Q}$.

Sea $f : [a, b] \to \mathbb{R}$ una función real, continua y no negativa.

Definición 9.1 *Sea $\mathcal{P} \in \Omega([a, b])$. Para cada $i = 1, 2, \ldots, n$, sea x_i^* un punto arbitrario en $[x_{i-1}, x_i]$. Al número*

$$S(f, \mathcal{P}, \{x_i^*\}) = \sum_{i=1}^{n} f(x_i^*)(x_i - x_{i-1})$$

se le llama la ***suma de Riemann correspondiente a la partición $\mathcal{P}$ y a la elección de puntos intermedios*** $\{x_i^*\}_{i=1}^{k}$.

Como casos particulares y muy importantes de sumas de Riemann se tienen aquéllas en las que los puntos intermedios x_i^* se toman de tal manera que $x_i^* = x_i^{\max}$, donde

$$f(x_i^{\max}) = \max \{f(x),\ x \in [x_{i-1}, x_i]\},$$

o $x_i^* = x_i^{\min}$, donde

$$f(x_i^{\min}) = \min \{f(x),\ x \in [x_{i-1}, x_i]\},$$

es decir, $x_i^{\max}$ $(x_i^{\min})$ es aquel punto en $[x_{i-1}, x_i]$ en el que la función f alcanza su máximo (mínimo). A las sumas de Riemann correspondientes a esas elecciones de los puntos en cada subintervalo, se les llama *suma superior de Darboux*[1]*-Riemann* y *suma inferior de Darboux-Riemann correspondientes a la partición* $\mathcal{P}$ y se denotan, respectivamente, por

$$\bar{S}(f,\mathcal{P}) = \sum_{i=1}^{k} f(x_i^{\max})(x_i - x_{i-1})$$

y

$$\underline{S}(f,\mathcal{P}) = \sum_{i=1}^{k} f(x_i^{\min})(x_i - x_{i-1}).$$

Ejemplo 9.1 Supongamos que la gráfica de la función $y = f(x)$, el intervalo $[a, b]$ y la partición $\mathcal{P} = \{x_0 = a, x_1, x_2, x_3, x_4 = b\}$ son los que se muestran en la figura 9.1

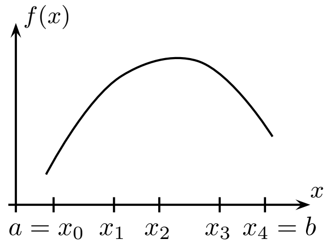
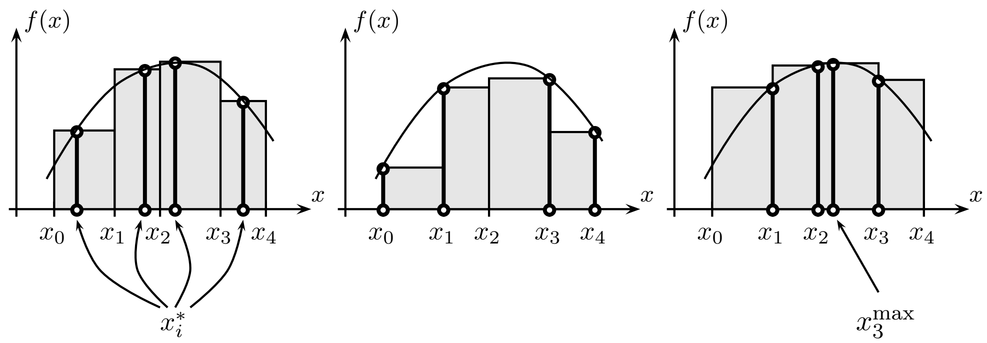


Figura 9.1 Una partición de $[a, b]$

Supongamos que tomamos los puntos intermedios $x_i^* \in [x_{i-1}, x_i]$, con $i = 1, 2, 3, 4$, como se muestra en la figura 9.2(a). Entonces las sumas de Riemann $S(f, \mathcal{P}, \{x_i^*\})$, $\underline{S}(f,\mathcal{P})$ y $\overline{S}(f,\mathcal{P})$, corresponden a las regiones sombreadas en las figuras 9.2(a), (b) y (c), respectivamente.



(a) $S(f, \mathcal{P}, \{x_i^*\})$ (b) $\underline{S}(f,\mathcal{P})$ (c) $\overline{S}(f,\mathcal{P})$

Figura 9.2 Sumas de Darboux-Riemann

[1]Jean Gaston Darboux (1842-1917) hizo importantes contribuciones a la geometría diferencial y al análisis.

Nótese que, en este caso particular, se tiene $x_1^{\min} = x_0, x_2^{\min} = x_1$, $x_3^{\min} = x_3$ y $x_4^{\min} = x_4$, mientras que $x_1^{\max} = x_1$, $x_2^{\max} = x_2$ y $x_4^{\max} = x_3$. Por otra parte, $x_3^{\max}$ está señalado en la figura 9.2(c). ◁

Puesto que para cualquier elección de $x_i^* \in [x_{i-1}, x_i]$ se cumple que

$$f(x_i^{\min}) \leqslant f(x_i^*) \leqslant f(x_i^{\max}),$$

entonces, para cada $\mathcal{P} \in \Omega([a,b])$, se tiene

$$\underline{S}(f,\mathcal{P}) \leqslant S(f,\mathcal{P},\{x_i^*\}) \leqslant \overline{S}(f,\mathcal{P}).$$

Nota Importante:

1. *Si $f(x)$ es una función no-negativa, es decir $f(x) \geqslant 0$ para toda $x \in [a,b]$, entonces la suma superior de Darboux-Riemann correspondiente a la partición $\mathcal{P}$, es la suma de las áreas de los rectángulos cuyas bases son los segmentos $[x_{i-1}, x_i]$ y cuya altura es la máxima ordenada de los puntos de la gráfica de f sobre el segmento $[x_{i-1}, x_i]$ para $i = 1, 2, \ldots n$. En ese caso, la región bajo la curva está contenida en la unión de rectángulos y su área será menor o igual a la suma de las áreas de éstos. Una situación análoga se tiene con la suma inferior de Darboux-Riemann asociada a la partición $\mathcal{P}$, donde el conjunto bajo la gráfica contiene a la unión de rectángulos cuyas áreas dan la suma inferior. También es de observarse que para cualquier elección de los puntos $x_i^* \in [x_{i-1}, x_i]$, la suma de Riemann $S(f,\mathcal{P},\{x_i^*\})$ es un número entre las sumas inferior y superior de Darboux-Riemann.*

2. *Si la función f toma en $[a,b]$ tanto valores positivos como negativos, la contribución a cada suma de los subintervalos sobre los que f toma valores negativos, será también un número negativo y entonces, en el cálculo de la suma de Riemann, la suma del área de los rectángulos que quedan por abajo del eje de las abscisas se resta del área de los rectángulos que quedan por arriba de esos ejes. Vea la figura 9.3.*

Para mostrar cómo tiene lugar el proceso de aproximación a medida que se toman particiones de $[a,b]$ cada vez más finas, probaremos primero el lema siguiente.

Lema 9.1 *Sean $\mathcal{P}, \mathcal{Q} \in \Omega([a,b])$ y $f : [a,b] \to \mathbb{R}$ una función continua. Entonces*

1. $\overline{S}(f,\mathcal{P}) \geqslant \underline{S}(f,\mathcal{P})$
2. $\overline{S}(f,\mathcal{P} \cup \mathcal{Q}) \leqslant \overline{S}(f,\mathcal{P})$
3. $\underline{S}(f,\mathcal{P} \cup \mathcal{Q}) \geqslant \underline{S}(f,\mathcal{P})$
4. $\sup\limits_{x\in[a,b]} f(x)(b-a) \geqslant \overline{S}(f,\mathcal{P}) \geqslant \underline{S}(f,\mathcal{P}) \geqslant \inf\limits_{x\in[a,b]} f(x)(b-a)$

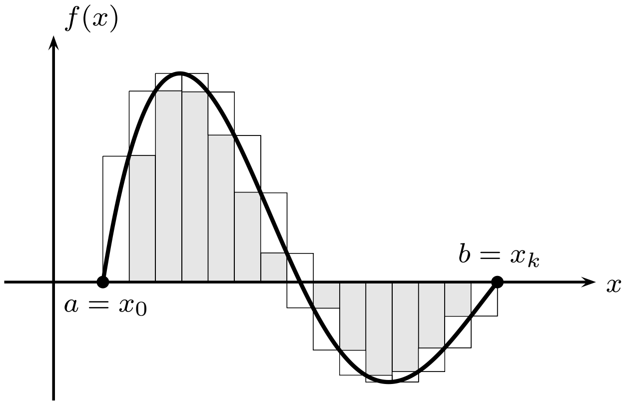


Figura 9.3 Sumas superiores e inferiores de Darboux-Riemann

5. $\overline{S}(f, \mathcal{P}) \geqslant \underline{S}(f, \mathcal{Q})$
6. $\inf\limits_{\mathcal{T} \in \Omega([a,b])} \overline{S}(f, \mathcal{T}) \geqslant \sup\limits_{\mathcal{T} \in \Omega([a,b])} \underline{S}(f, \mathcal{T})$

Demostración.

Sean $\mathcal{P} = \{a = x_0 < x_2 < \cdots x_k = b\}$ y $\mathcal{Q} = \{a = z_0 < z_1 < \cdots < z_r = b\}$. Como $f(x_i^{\max}) \geqslant f(x_i^{\min})$ para cada $i = 1, \ldots, k$, se tiene la validez del punto 1.

Para probar el punto 2, basta observar que al considerar la partición $\mathcal{P} \cup \mathcal{Q}$, el subintervalo $[x_{i-1}, x_i]$ de $\mathcal{P}$ contiene a varios elementos z_k de $\mathcal{Q}$ en la forma $x_{i-1} \leqslant z_j < z_{j+1} < \cdots z_s \leqslant x_i$ y da lugar en $\overline{S}(f, \mathcal{P} \cup \mathcal{Q})$ a varios sumandos que corresponden al área de los rectángulos asociados a $\mathcal{P} \cup \mathcal{Q}$ cuyas bases están contenidas en $[x_{i-1}, x_i]$ y cuyas alturas $\sup_{x \in [z_{s-1}, z_s]} f(x)$ son menores o iguales que $\sup_{x \in [x_{i-1}, x_i]} f(x)$. Puesto que, en general, $\sup_M f(x) \leqslant \sup_N f(x)$ si $M \subset N$, entonces $\overline{S}(f, \mathcal{P} \cup \mathcal{Q}) \leqslant \overline{S}(f, \mathcal{P})$, lo cual prueba el punto 2.

La prueba del punto 3 es de forma análoga a la prueba del punto 2 si observamos que

$$\inf_M f(x) \geqslant \inf_N f(x) \quad \text{si} \quad M \subset N.$$

La validez del punto 4 es evidente, y el punto 5 se sigue de notar que

$$\overline{S}(f, \mathcal{P}) \geqslant \overline{S}(f, \mathcal{P} \cup \mathcal{Q}) \geqslant \underline{S}(f, \mathcal{P} \cup \mathcal{Q}) \geqslant \underline{S}(f, \mathcal{Q}).$$

La validez del punto 6 es consecuencia directa de la validez del punto 5. ■

Nota Importante:

Los puntos 1 y 5 en el lema anterior implican que cada suma superior de Darboux-Riemann es cota superior para el conjunto de todas las sumas inferiores de Darboux-Riemann posibles y análogamente, cada suma inferior de Darboux-Riemann es cota *inferior de todas las sumas superiores de Darboux-Riemann posibles. Por otro lado, los puntos 2 y 3 significan que al ir refinando cada vez más una partición mediante la incorporación de nuevos puntos, se genera una sucesión creciente de sumas inferiores y una sucesión decreciente de sumas superiores. El punto clave aquí es que a medida que se toman particiones con norma cada vez más pequeña, las sumas de Riemann convergen todas a un mismo número, lo cual motiva la definición siguiente.*

Definición 9.2 *Para cada función $f : [a,b] \to \mathbb{R}$ continua, se define la* ***integral de f en el intervalo*** *$[a,b]$ como el número real $\int_a^b f(x)\mathrm{d}x$ dado por*

$$\int_a^b f(x)\mathrm{d}x = \lim_{n\to\infty} S(f, \mathcal{P}_n, \{x_i^{n*}\})$$

donde

$$\mathcal{P}_n = \left\{a = x_0^n < x_2^n < \cdots < x_{k(n)}^n = b\right\} \text{ para } n = 1, 2, \ldots$$

es una sucesión de particiones de $[a,b]$ cuya norma $|\mathcal{P}_n|$ tiende a cero y $\{x_i^{n*}\}$ es una elección arbitraria de puntos $x_i^{n*} \in [x_{i-1}^n, x_i^n]$ para cada $i = 1, \ldots, k(n)$ y $n = 1, 2, \ldots$

Nota Importante:

1. *El símbolo $\int$ es una deformación del símbolo $\sum$. La expresión $f(x)\mathrm{d}x$ denota el área de un rectángulo de base un incremento infinitesimal $\mathrm{d}x$ de la variable x y altura $f(x)$. Los extremos inferior a y superior b en el signo $\int_a^b$ denotan el sentido en que se recorre el intervalo $[a,b]$.*

2. *Para que la definición de integral definida sea válida, debemos probar que si f es una función continua, entonces todas las sucesiones de sumas de Riemann correspondientes a particiones cuya norma tiende a cero, tienen un mismo límite. Para probar ese hecho, demostraremos en la proposición 9.2, que si la función f es continua, entonces el infimum de las sumas superiores de Darboux-Riemann coincide con el supremum de las sumas inferiores y, por lo tanto, cualquier sucesión de sumas de Riemann correspondientes a particiones con norma que tiende a cero convergen a un mismo número real. Ese número es la integral definida de f en $[a,b]$ y corresponde, si f es no-negativa, al valor del área del conjunto bajo la gráfica de la función sobre $[a,b]$.*

Proposición 9.2 *Si $f : [a,b] \to \mathbb{R}$ es una función continua, entonces*

$$\inf_{\mathcal{P}\in\Omega([a,b])} \overline{S}(f,\mathcal{P}) = \sup_{\mathcal{P}\in\Omega([a,b])} \underline{S}(f,\mathcal{P}).$$

Más aún, si $\mathcal{P}_n = \{a = x_0^n < x_2^n < \ldots < x_k^n = b\}$ para $n = 1, 2, \ldots,$ es una sucesión de particiones de $[a,b]$ cuya norma $|\mathcal{P}_n|$ tiende a cero, entonces se tiene que

$$\lim_{n\to\infty} S(f, \mathcal{P}_n, \{x_i^{n*}\}) = \inf_{\mathcal{P}\in\Omega([a,b])} \overline{S}(f,\mathcal{P}) = \sup_{\mathcal{P}\in\Omega([a,b])} \underline{S}(f,\mathcal{P}),$$

donde $\{x_i^{n*}\}$ *es una elección arbitraria de puntos* $x_i^{n*} \in [x_{i-1}^n, x_i^n]$ *y* $S(f, \mathcal{P}_n, \{x_i^*\})$ *es la suma de Riemann asociada a* $\mathcal{P}_n$ *y a la elección* $\{x_i^{n*}\}$.

Demostración. Al ser f continua en $[a, b]$, en virtud del teorema 4.13, es uniformemente continua, es decir, para cada $\varepsilon > 0$ existe $\delta > 0$ tal que si x y $y \in [a, b]$ y $|x - y| \leqslant \delta$ entonces

$$|f(x) - f(y)| \leqslant \frac{\varepsilon}{b-a}.$$

Por lo tanto, si $\mathcal{P}$ es una partición tal que los subintervalos en los cuales divide a $[a, b]$ son de longitud menor o igual a δ, se tiene que en cada subintervalo $[x_{i-1}, x_i]$ se cumple

$$f(x_i^{\max}) - f(x_i^{\min}) \leqslant \frac{\varepsilon}{b-a}.$$

Consecuentemente, las sumas superiores e inferiores respectivas satisfacen

$$\overline{S}(f, \mathcal{P}) - \underline{S}(f, \mathcal{P}) \leqslant \varepsilon.$$

Esto muestra que para cada número $\varepsilon > 0$, toda partición $\mathcal{P}$ de $[a, b]$ tal que $|\mathcal{P}| < \delta$, es tal que las suma superior e inferior de Darboux-Riemann respectivas difieren entre sí en menos que ε. Teniendo en cuenta la estimación anterior, podemos escribir

$$\inf_{\mathcal{T}\in\Omega([a,b])} \overline{S}(f, \mathcal{T}) \leqslant \overline{S}(f, \mathcal{P}) \leqslant \underline{S}(f, \mathcal{P}) + \varepsilon \leqslant \sup_{\mathcal{T}\in\Omega([a,b])} \underline{S}(f, \mathcal{T}) + \varepsilon,$$

y por lo tanto,

$$\inf_{\mathcal{T}\in\Omega([a,b])} \overline{S}(f, \mathcal{T}) \leqslant \sup_{\mathcal{T}\in\Omega([a,b])} \underline{S}(f, \mathcal{T}) + \varepsilon.$$

Como ε es un número positivo arbitrario, concluimos que

$$\inf_{\mathcal{T}\in\Omega([a,b])} \overline{S}(f, \mathcal{T}) \leqslant \sup_{\mathcal{T}\in\Omega([a,b])} \underline{S}(f, \mathcal{T}),$$

y esto, junto con el punto 6 del lema 9.1, implica que

$$\inf_{\mathcal{T}\in\Omega([a,b])} \overline{S}(f, \mathcal{T}) = \sup_{\mathcal{T}\in\Omega([a,b])} \underline{S}(f, \mathcal{T}).$$

Sea ahora una sucesión $\mathcal{P}_n$ de particiones con $|\mathcal{P}_n| \to 0$. Dado $\varepsilon > 0$ existe N tal que $|\mathcal{P}_n| < \delta$ si $n > N$ y entonces podemos escribir

$$\inf_{\mathcal{T}\in\Omega([a,b])} \overline{S}(f, \mathcal{T}) \leqslant \overline{S}(f, \mathcal{P}_n) \leqslant \underline{S}(f, \mathcal{P}_n) + \varepsilon, \quad \text{para} \quad n > N$$

y, por lo tanto,

$$\inf_{\mathcal{T}\in\Omega([a,b])} \overline{S}(f, \mathcal{T}) \leqslant \overline{S}(f, \mathcal{P}_n) \leqslant \sup_{\mathcal{T}\in\Omega([a,b])} \underline{S}(f, \mathcal{T}) + \varepsilon, \quad \text{para} \quad n > N.$$

Luego, hemos probado que

$$\lim_{n\to\infty} \overline{S}(f,\mathcal{P}_n) = \sup_{\mathcal{T}\in\Omega([a,b])} \underline{S}(f,\mathcal{T}).$$

Análogamente, se tiene

$$\lim_{n\to\infty} \underline{S}(f,\mathcal{P}_n) = \inf_{\mathcal{T}\in\Omega([a,b])} \overline{S}(f,\mathcal{T}).$$

Finalmente, puesto que para cualquier elección de puntos $\{x_i^{n*}\}$ se tiene

$$\overline{S}(f,\mathcal{P}_n) \geqslant S(f,\mathcal{P}_n,\{x_i^{n*}\}) \geqslant \underline{S}(f,\mathcal{P}_n),$$

haciendo $n \to \infty$ se obtiene,

$$\lim_{n\to\infty} S(f,\mathcal{P}_n,\{x_i^{n*}\}) = \sup_{\mathcal{T}\in\Omega([a,b])} \underline{S}(f,\mathcal{T}) = \inf_{\mathcal{T}\in\Omega([a,b])} \overline{S}(f,\mathcal{T}),$$

con lo cual hemos demostrado que todas las sucesiones de sumas de Riemann correspondientes a particiones con norma que tienden a cero, tienen un mismo límite, que se denomina la *integral definida de* f *en* $[a,b]$. ■

Nota Importante:
Al cambiar la orientación de una partición $\mathcal{P}$, *la suma de Riemann cambia de signo al invertirse los extremos inicial y final de cada subintervalo. Para señalar la orientación que se ha dado a las particiones que se han utilizado para calcular la integral, escribiremos siempre en la parte inferior del signo* $\int f(x)\mathrm{d}x$ *el extremo inicial del intervalo de integración y en la parte superior el extremo final, definidos éstos de acuerdo a la orientación dada a las particiones. Tomando en cuenta lo anterior, tenemos la relación*

$$\int_b^a f(x)\mathrm{d}x = -\int_a^b f(x)\mathrm{d}x.$$

Ejemplo 9.2 A partir de la definición, calculemos la integral de la función $f(x) = cx$ en el intervalo $[a,b]$.

Tomemos la sucesión particular de particiones de $[a,b]$ con norma que tiende a cero dada mediante $\mathcal{P}_n = a < a+h < a+2h < \cdots < a+nh = b$, donde $h = \dfrac{b-a}{n}$. En este caso, los intervalos en que se divide $[a,b]$, tienen por extremo derecho a $x_i = a+ih$, con $i = 1,\ldots,n$. Elijamos ahora los puntos intermedios $x_i^* = a+ih-\dfrac{h}{2} \in [x_{i-1},x_i]$ y calculemos las sumas de Riemann correspondientes, obteniendo

$$\begin{aligned} S(f,\mathcal{P}_n,\{x_i^*\}) &= \sum_{i=1}^n c\left(a+ih-\frac{h}{2}\right)h = ca(b-a) + \frac{c(b-a)^2}{n^2}\sum_{i=1}^n\left(i-\frac{1}{2}\right) \\ &= ca(b-a) + \frac{c(b-a)^2}{n^2}\left(\frac{n(n+1)}{2}-\frac{n}{2}\right) = \frac{1}{2}c(b^2-a^2), \end{aligned}$$

y tomando límite cuando $n \to \infty$ tendremos

$$\int_a^b cx\mathrm{d}x = \frac{1}{2}c(b^2 - a^2). \qquad \triangleleft$$

9.1.1 Propiedades de la integral definida

En la siguiente proposición enlistamos las propiedades principales de la integral definida.

Proposición 9.3 *Sean $f, g : [a, b] \to \mathbb{R}$ funciones continuas en $[a, b]$. Entonces*

1. *Para cada $\lambda \in \mathbb{R}$,*
$$\int_a^b (\lambda f + g)(x)\mathrm{d}x = \lambda \int_a^b f(x)\mathrm{d}x + \int_a^b g(x)\mathrm{d}x \qquad \text{(Linealidad)}$$
2. *Para cada $c \in (a, b)$,*
$$\int_a^b f(x)\mathrm{d}x = \int_a^c f(x)\mathrm{d}x + \int_c^b f(x)\mathrm{d}x \qquad \text{(Aditividad del intervalo)}$$
3. $\displaystyle \inf_{x\in[a,b]} f(x)(b-a) \leqslant \int_a^b f(x)\mathrm{d}x \leqslant \sup_{x\in[a,b]} f(x)(b-a)$
4. *Si $f(x) \geqslant g(x)$ en $[a, b]$, entonces $\int_a^b f(x)\mathrm{d}x \geqslant \int_a^b g(x)\mathrm{d}x$.*
5. $\displaystyle \left|\int_a^b f(x)\mathrm{d}x\right| \leqslant \int_a^b |f(x)|\mathrm{d}x$

Demostración. Para probar 1, basta observar que para cada sucesión de particiones $\mathcal{P}_n$ de $[a, b]$ con norma que tiende a cero y cada elección de puntos intermedios $\{x_i^{n*}\}$, se cumple que

$$S(\lambda f + g, \mathcal{P}_n, \{x_i^{n*}\}) = \lambda S(f, \mathcal{P}_n, \{x_i^{n*}\}) + S(g, \mathcal{P}_n, \{x_i^{n*}\}),$$

y tomando en cuenta la convergencia de las sucesiones de la derecha, se tendrá

$$\int_a^b (\lambda f + g)(x)\mathrm{d}x = \lambda \int_a^b f(x)\mathrm{d}x + \int_a^b g(x)\mathrm{d}x.$$

Para la prueba de 2, obsérvese que si $\mathcal{P}_n$ y $\mathcal{R}_n$ son sucesiones de particiones con norma que tiende a cero de $[a, c]$ y $[c, b]$, entonces su unión $\mathcal{P}_n \cup \mathcal{R}_n$ define una sucesión de particiones de $[a, b]$ tales que para cualquier elección de puntos intermedios se tiene

$$S(f, \mathcal{P}_n \cup \mathcal{R}_n, \{x_i^{n*}\}) = S(f, \mathcal{P}_n, \{x_i^{n*}\}) + S(f, \mathcal{R}_n, \{x_i^{n*}\}).$$

Al tomar límite cuando $n \to \infty$, de las sucesiones anteriores se obtiene

$$\int_a^b f(x)\mathrm{d}x = \int_a^c f(x)\mathrm{d}x + \int_c^b f(x)\mathrm{d}x,$$

con lo que se prueba el punto 2.

La demostración del punto 3 es consecuencia de la siguiente estimación para cada partición $\mathcal{P}$ de $[a, b]$:

$$\sup_{x\in[a,b]} f(x)(b-a) \geqslant S(f, \mathcal{P}_n, \{x_i^{n*}\}) \geqslant \inf_{x\in[a,b]} f(x)(b-a).$$

La demostración del punto 4 se tiene al observar que si $k : [a, b] \to R$ es una función continua con $k(x) \geqslant 0$ para $x \in [a, b]$, entonces

$$\inf_{x\in[a,b]} k(x)(b-a) \geqslant 0,$$

y, de lo probado en el punto 3, se sigue que $\int_a^b k(x)\mathrm{d}x \geqslant 0$. Tomando en cuenta lo anterior, si $f(x) \geqslant g(x)$ en $[a, b]$, se tiene $(f-g)(x) \geqslant 0$ y entonces

$$\int_a^b f(x)\mathrm{d}x - \int_a^b g(x)\mathrm{d}x \geqslant 0.$$

Finalmente, la prueba del punto 5 se sigue directamente del punto 3 y de considerar que si f es continua, también lo es $|f(x)|$ y

$$-|f(x)| \leqslant f(x) \leqslant |f(x)|;$$

luego,

$$-\int_a^b |f(x)|\mathrm{d}x \leqslant \int_a^b f(x)\mathrm{d}x \leq \int_a^b |f(x)|\mathrm{d}x,$$

lo cual implica que

$$\left|\int_a^b f(x)\mathrm{d}x\right| \leqslant \int_a^b |f(x)|\mathrm{d}x.$$

■

Corolario 9.4 (Teorema del valor medio para integrales) *Sea $f : [a, b] \to \mathbb{R}$ continua. Entonces existe $c \in [a, b]$ tal que*

$$\int_a^b f(x)\mathrm{d}x = f(c)(b-a).$$

Demostración. Del punto 3 de la proposición anterior, tenemos que

$$\inf_{x\in[a,b]} f(x) \leqslant \frac{1}{(b-a)} \int_a^b f(x)\mathrm{d}x \leqslant \sup_{x\in[a,b]} f(x).$$

Tomando ahora en cuenta que toda función continua en un intervalo cerrado $[a, b]$, alcanza cada valor intermedio entre su valor máximo y su valor mínimo, y como $\dfrac{1}{(b-a)}\displaystyle\int_a^b f(x)\mathrm{d}x$ es un valor entre esos valores extremos de la función, entonces existe $c \in [a, b]$ tal que

$$f(c) = \frac{1}{(b-a)}\int_a^b f(x)\mathrm{d}x,$$

con lo cual se prueba el corolario. ■

Corolario 9.5 (Segundo teorema del valor medio para integrales) *Sean $f(x)$ y $g(x)$ funciones continuas en $[a, b]$ y $f(x) \geqslant 0$. Entonces, existe $c \in (a, b)$ tal que*

$$\int_a^b f(x)g(x)\mathrm{d}x = g(c)\int_a^b f(x)\mathrm{d}x.$$

Demostración A partir de la estimación

$$f(x)\min_{x\in[a,b]} g(x) \leqslant f(x)g(x) \leqslant f(x)\max_{x\in[a,b]} g(x),$$

podemos escribir

$$\min_{x\in[a,b]} g(x)\int_a^b f(x)\mathrm{d}x \leqslant \int_a^b f(x)g(x)\mathrm{d}x \leqslant \max_{x\in[a,b]} g(x)\int_a^b f(x)\mathrm{d}x.$$

Enseguida, aplicando el teorema del valor intermedio para funciones continuas, existirá $c \in (a, b)$ tal que

$$\int_a^b f(x)g(x)\mathrm{d}x = g(c)\int_a^b f(x)\mathrm{d}x. \quad ■$$

9.2 El teorema fundamental del cálculo

Aparentemente, el cálculo de una integral definida es un proceso difícil de implementar, pues se requiere tomar en cuenta el comportamiento de la función a lo largo de todo el intervalo de integración. Sin embargo, cuando se conoce una primitiva o antiderivada de la función, el cálculo de la integral en el intervalo se reduce a una mera valuación de esa primitiva en sus extremos, tal como lo mostramos en la proposición siguiente.

Proposición 9.6 *Sea $f : [a, b] \to \mathbb{R}$ continua y $g : [a, b] \to \mathbb{R}$ una primitiva de f, es decir $\frac{\mathrm{d}g}{\mathrm{d}x}(x) = f(x)$. Entonces*

$$\int_a^b f(x)\mathrm{d}x = \int_a^b \frac{dg}{\mathrm{d}x}(x)\mathrm{d}x = g(b) - g(a).$$

Demostración. Consideremos una partición $\mathcal{P} = \{a = x_0 < x_2 < \cdots < x_k = b\}$ y tomemos en cada subintervalo $[x_{i-1}, x_i]$ un punto intermedio x_i^* tal que

$$g(x_i) - g(x_{i-1}) = \frac{dg}{\mathrm{d}x}(x_i^*)(x_i - x_{i-1}).$$

La existencia de tal punto lo asegura el teorema de valor medio aplicado a la función $g(x)$ en el intervalo $[x_{i-1}, x_i]$. Luego, evaluando la suma de Riemann correspondiente, se tiene

$$\begin{aligned} S(f, \mathcal{P}, \{x_i^*\}\,) &= \sum_{i=1}^{k} \frac{dg}{\mathrm{d}x}(x_i^*)(x_i - x_{i-1}) = \\ &= \sum_{i=1}^{k} (g(x_i) - g(x_{i-1})) = g(b) - g(a), \end{aligned}$$

lo cual muestra que $S(f, \mathcal{P}, \{x_i^*\}\,)$ es un número que no depende de la partición $\mathcal{P}$. Luego, el límite de la sumas de Riemann correspondientes a una sucesión de particiones $\mathcal{P}_n$ con norma que tiende a cero y donde los puntos intermedios x_i^* se eligen como se hizo arriba, será el número $g(b) - g(a)$ y por lo tanto se tendrá

$$\int_a^b f(x)\mathrm{d}x = g(b) - g(a). \qquad \blacksquare$$

Nota Importante:
Hemos probado que si una función continua tiene una primitiva, entonces la integral de la primera sobre cada intervalo cerrado es igual a la diferencia de valores de la primitiva en sus extremos. Tomando en cuenta que dos primitivas en un intervalo difieren por una constante, el valor de la diferencia de valores $g(b) - g(a)$ es independiente de la primitiva que se utilice.

El resultado anterior es la primera parte del llamado *teorema fundamental del cálculo*.[2] La parte restante afirma que cada función continua tiene una primitiva.

[2] La idea original de este teorema se debe a Isaac Barrow, matemático mencionado en el capítulo primero.

Proposición 9.7 *Si $f : [a,b] \to \mathbb{R}$ es una función continua, la función*

$$g : [a,b] \to \mathbb{R},$$

definida mediante la expresión

$$g(x) = \int_a^x f(s)\mathrm{d}s,$$

es una primitiva de f.

Demostración. Mostremos primero que $g(x)$ es derivable.

$$\begin{aligned}
\lim_{h\to 0} \frac{g(x+h)-g(x)}{h} &= \lim_{h\to 0} \frac{1}{h}\left(\int_a^{x+h} f(s)\mathrm{d}s - \int_a^x f(s)\mathrm{d}s \right) \\
&= \lim_{h\to 0} \frac{1}{h}\left(\int_a^{x} f(s)\mathrm{d}s + \int_x^{x+h} f(s)\mathrm{d}s - \int_a^x f(s)\mathrm{d}s \right) \\
&= \lim_{h\to 0} \frac{1}{h} \int_x^{x+h} f(s)\mathrm{d}s.
\end{aligned}$$

Aplicando el teorema de valor medio para integrales, tenemos que

$$\frac{1}{h}\int_x^{x+h} f(s)\mathrm{d}s = f(x_h), \text{ con } x_h \in [x, x+h]$$

Tomando límite cuando $h \to 0$ y considerando que f es continua en x, obtenemos

$$\lim_{h\to 0} \frac{g(x+h)-g(x)}{h} = \lim_{h\to 0} f(x_h) = f(x).$$

Esto prueba que la función $g(x)$ es derivable y $\dfrac{\mathrm{d}g}{\mathrm{d}x}(x) = f(x)$. Luego, g es una primitiva de f en $[a,b]$. ■

Nota Importante:

1. *El teorema fundamental del cálculo en cierta manera establece que las operaciones de derivación y de integración son procesos recíprocos, en el sentido siguiente:*

$$\int_a^x \frac{\mathrm{d}f}{\mathrm{d}x}(s)\mathrm{d}s = f(x) - f(a),$$

$$\frac{\mathrm{d}\left(\int_a^x f(s)\mathrm{d}s \right)}{\mathrm{d}x}(x) = f(x).$$

2. *El teorema fundamental del cálculo proporciona un método para el cálculo de integrales definidas de funciones continuas, remitiendo al problema del cálculo de primitivas o antiderivadas.*

Ejemplo 9.3 Utilizando el teorema fundamental del cálculo, podemos evaluar la integral definida $\int_0^{\frac{\pi}{2}} (\operatorname{sen} x + x)\mathrm{d}x$, observando que la función $g(x) = -\cos x + \frac{1}{2}x^2$ es una antiderivada de la función $\operatorname{sen} x + x$. Luego,

$$\int_0^{\frac{\pi}{2}} (\operatorname{sen} x + x)\mathrm{d}x = g\left(\frac{\pi}{2}\right) - g(0) = \frac{1}{8}\pi^2 + 1. \qquad \triangleleft$$

Ejemplo 9.4 Si en un tanque vacío de 5000 metros cúbicos de capacidad se vierte agua a una razón de $\frac{t}{100}\frac{m^3}{seg}$, ¿en cuánto tiempo se llena el tanque?

Solución: Si denotamos por $V(t)$ la función que en el tiempo t es igual al volumen de agua en el tanque, se tiene como dato que

$$\frac{\mathrm{d}V}{\mathrm{d}t}(t) = \frac{t}{100},$$

y por lo tanto, aplicando el teorema fundamental del cálculo, tenemos

$$V(t) = \int_0^t \frac{s}{100}\mathrm{d}s = \frac{1}{200}t^2.$$

El tanque se llenará en el tiempo T tal que

$$V(T) = 5000 = \frac{1}{200}T^2,$$

es decir,

$$T = 1000 \text{ segundos}. \qquad \triangleleft$$

9.3 Integrales impropias

En este apartado presentamos los conceptos de integral impropia y de integral para funciones seccionalmente continuas. En el primer caso, se trata de extender el concepto de integral a intervalos abiertos o no acotados, mientras que en el último, se trata de definir la integral para funciones que contienen un número finito de puntos de discontinuidad.

Definición 9.3 *Una función continua $f(x)$ definida en un intervalo de la forma $[a,\infty)$ se dice integrable si para cada sucesión $\{b_i\}_{i=1}^{\infty}$ que tiende a ∞ se tiene que la sucesión de integrales definidas $\left\{\int_a^{b_i} f(x)\mathrm{d}x\right\}_{i=1}^{\infty}$ converge a un límite común $L \in \mathbb{R}$. Al número L se le llama la* ***integral impropia de f en*** *$[a,\infty)$ y se denota con $\int_a^{\infty} f(x)\mathrm{d}x$; es decir,*

$$\int_a^{\infty} f(x)\mathrm{d}x = \lim_{x\to\infty} \int_a^{x} f(s)\mathrm{d}s.$$

NOTA IMPORTANTE:
Cuando decimos que la sucesión $\{b_i\}_{i=1}^{\infty}$ tiende a ∞, entendemos que para cada natural M existe una etiqueta N tal que $b_i \geqslant M$ para todo $i \geqslant N$.

Ejemplo 9.5 La función $f(x) = \operatorname{sen} x$ no tiene integral impropia en $[0,\infty)$ ya que si tomamos la sucesión $\{2\pi k\}_{k=1}^{\infty}$ se tiene que

$$\lim_{k\to\infty} \int_0^{2\pi k} \operatorname{sen} x\mathrm{d}x = 0,$$

pero si tomamos la sucesión $\left\{2\pi k + \frac{\pi}{2}\right\}_{k=1}^{\infty}$ se tiene

$$\lim_{k\to\infty} \int_0^{2\pi k+\frac{\pi}{2}} \operatorname{sen} x\mathrm{d}x = 1. \qquad \triangleleft$$

Ejemplo 9.6 La integral impropia $\int_1^{\infty} x^{-p}\mathrm{d}x$, donde p es un número entero, es convergente para $p > 1$ y divergente para $p \leqslant 1$. La afirmación para $p > 1$ se deduce directamente del cálculo

$$\int_1^{\infty} x^{-p}\mathrm{d}x = \lim_{t\to\infty} \int_1^{t} x^{-p}\mathrm{d}x = \frac{1}{1-p} \lim_{t\to\infty} (t^{-p+1} - 1) = \frac{1}{1-p},$$

mientras que si $p = 1$, se tiene

$$\int_1^{\infty} x\mathrm{d}x = \lim_{t\to\infty} \int_1^{t} x\mathrm{d}x = \frac{1}{2} \lim_{t\to\infty} (t^2 - 1) \ \to \infty,$$

y si $p < 1$,

$$\int_1^{\infty} x^{-p}\mathrm{d}x = \lim_{t\to\infty} \int_1^{t} x^{-p}\mathrm{d}x = \frac{1}{1-p} \lim_{t\to\infty} (t^{-p+1} - 1) \to \infty.$$

◁

Análogamente, se define la *integral impropia de una función continua $f(x)$ en $(-\infty, b]$* como el límite común, si existe, de las sucesiones de la forma $\left\{\int_{a_i}^{b} f(x)\mathrm{d}x\right\}_{i=1}^{\infty}$ para toda sucesión $\{a_i\}_{i=1}^{\infty}$ que tiende a $-\infty$. La integral impropia de f en $(-\infty, b]$ se denota con $\int_{-\infty}^{b} f(x)\mathrm{d}x$; es decir,

$$\int_{-\infty}^{b} f(x)\mathrm{d}x = \lim_{x\to-\infty}\int_{x}^{b} f(s)\mathrm{d}s.$$

Finalmente, diremos que una función $f(x)$, continua en $(-\infty, \infty)$, tiene *integral impropia en* $(-\infty, \infty)$, si para algún número real a, la función f posee integrales impropias en $(-\infty, a]$ y en $[a, \infty)$. A la suma de tales integrales se le llama la *integral impropia de f en* $(-\infty, \infty)$ y se denota con $\int_{-\infty}^{\infty} f(x)\mathrm{d}x$; es decir,

$$\int_{-\infty}^{\infty} f(x)\mathrm{d}x = \int_{-\infty}^{a} f(x)\mathrm{d}x + \int_{a}^{\infty} f(x)\mathrm{d}x.$$

Nota Importante:

1. *La definición de integral impropia en toda la recta $(-\infty, \infty)$, no depende del valor a.*

2. *A las integrales impropias de funciones en intervalos no acotados se les denomina "integrales impropias de primera clase". Cuando una integral impropia existe, también se dice que la integral impropia converge y en caso contrario se dice que la integral impropia diverge.*

De manera similar al caso de intervalos no acotados, si $f(x)$ es una función continua definida en un intervalo de la forma $[a, b)$, se dice que tiene *integral impropia en* $[a, b)$ si para toda sucesión de reales $\{b_i\}_{i=1}^{\infty}$ con términos $a < b_i < b$ y tal que $\lim_{i\to\infty} b_i = b$, se tiene que la sucesión de integrales definidas $\left\{\int_{a}^{b_i} f(x)\mathrm{d}x\right\}_{i=1}^{\infty}$ converge a un valor L, independiente de la sucesión $\{b_i\}_{i=1}^{\infty}$ que se tome. Al número L se le llama *integral impropia de f en* $[a, b)$ y se denota con $\int_{a}^{b} f(x)\mathrm{d}x$; es decir,

$$\int_{a}^{b} f(x)\mathrm{d}x = \lim_{b_i\to b}\int_{a}^{b_i} f(x)\mathrm{d}x.$$

Para el caso de funciones continuas sobre intervalos de la forma $(a, b]$, la definición de integral impropia es similar. Finalmente si $f(x)$ es una función continua en un intervalo abierto de la forma (a, b), se dice que tiene *integral impropia en* (a, b) si para algún $c \in (a, b)$ las integrales impropias $\int_a^c f(x)\mathrm{d}x$ y $\int_c^b f(x)\mathrm{d}x$ existen y en tal caso se define la *integral impropia de* f *en* (a, b) como la suma

$$\int_a^b f(x)\mathrm{d}x = \int_a^c f(x)\mathrm{d}x + \int_c^b f(x)\mathrm{d}x.$$

Por su utilidad en probar la existencia de integrales impropias, damos aquí el siguiente criterio de comparación.

Proposición 9.8 (Criterio de comparación) *Sean* $f(x)$ *y* $g(x)$ *dos funciones no-negativas definidas en* $(-\infty, \infty)$ *y tales que* $0 \leqslant f(x) \leqslant g(x)$ *para toda* $x \in (-\infty, \infty)$. *Entonces, si* $g(x)$ *tiene integral impropia en* $(-\infty, \infty)$, *también la tiene la función* $f(x)$ *y se cumple*

$$\int_{-\infty}^{\infty} f(x)\mathrm{d}x \leqslant \int_{-\infty}^{\infty} g(x)\mathrm{d}x.$$

Demostración. Basta mostrar que bajo las hipótesis anteriores la función $f(x)$ tiene integral impropia en $[0, \infty)$. Tomemos una sucesión $\{b_i\}_{i=1}^{\infty}$ que tienda a ∞. Por la no-negatividad de las funciones y de la condición $0 \leqslant f(x) \leqslant g(x)$ se sigue que

$$\int_0^{b_i} f(x)\mathrm{d}x \leqslant \int_0^{b_i} g(x)\mathrm{d}x.$$

Como las dos sucesiones $\left\{\int_0^{b_i} f(x)\mathrm{d}x\right\}_{i=1}^{\infty}$ y $\left\{\int_0^{b_i} g(x)\mathrm{d}x\right\}_{i=1}^{\infty}$ son positivas y la segunda convergente a, digamos, un número L, se tiene que la sucesión $\left\{\int_0^{b_i} f(x)\mathrm{d}x\right\}_{i=1}^{\infty}$ será convergente (ver ejercicio 5, capítulo 4), y

$$\int_0^{\infty} f(x)\mathrm{d}x \leqslant \int_0^{\infty} g(x)\mathrm{d}x.$$

Análogamente se muestra que $\displaystyle\int_{-\infty}^{0} f(x)\mathrm{d}x \leqslant \int_{-\infty}^{0} g(x)\mathrm{d}x$ y, consiguientemente, la proposición queda probada. ∎

9.4 Integración de funciones continuas por secciones

La integral definida que hemos presentado es válida solamente para funciones continuas en un intervalo cerrado y acotado. Concluimos este capítulo presentando la generalización de ese concepto a la familia de *funciones seccionalmente continuas,* que son aquellas funciones que tienen un número finito de puntos de discontinuidad. En este caso, el dominio de definición de esas funciones se puede escribir como unión de un número finito de intervalos abiertos ajenos, cuyos extremos los ocupan los puntos de discontinuidad.

Ejemplo 9.7 La función $f : [0,3] \to \mathbb{R}$ con

$$f(x) = \begin{cases} \sqrt{1-x^2} & \text{si } x \in [0,1] \\ 2 & \text{si } x \in (1,2] \\ x-2 & \text{si } x \in (2,3] \end{cases}$$

es una función seccionalmente continua cuyos puntos de discontinuidad son $x = 1$ y $x = 2$. ◁

> ***Definición 9.4*** *Una función seccionalmente continua $f(x)$ se dice integrable en un dominio si posee integral impropia en cada uno de los intervalos abiertos en que sus discontinuidades dividen a ese dominio. A la suma de tales integrales se le llama la **integral de la función seccionalmente continua.***

Ejemplo 9.8 La función $f : (0,2] \to \mathbb{R}$ con

$$f(x) = \begin{cases} \dfrac{1}{\sqrt{x}} & \text{si } 0 < x \leqslant 1 \\ x-1 & \text{si } 1 < x \leqslant 2 \end{cases}$$

es seccionalmente continua y su integral se calcula como sigue:

$$\begin{aligned} \int_0^2 f(x)\mathrm{d}x &= \int_0^1 \frac{1}{\sqrt{x}}\mathrm{d}x + \int_1^2 (x-1)\mathrm{d}x \\ &= \lim_{c \to 0^+} \int_c^1 \frac{1}{\sqrt{x}}\mathrm{d}x + \left(\frac{x^2}{2} - x\right)\Bigg|_1^2 \\ &= \lim_{c \to 0^+} 2\sqrt{c}\Big|_0^1 + \frac{1}{2} \\ &= \frac{5}{2}. \end{aligned}$$

◁

Ejercicios y problemas del capítulo

1. Haciendo uso de las propiedades de la integral definida, resuelva los problemas siguientes:

   (a) Sea f función continua tal que $\displaystyle\int_1^3 f(x)\mathrm{d}x = 3$, $\displaystyle\int_2^5 f(x)\mathrm{d}x = -2$. Calcule $\displaystyle\int_3^5 f(x)\mathrm{d}x$ y $\displaystyle\int_2^1 f(x)\mathrm{d}x$.

   (b) Demuestre que $\displaystyle\sqrt{3} \leqslant \frac{1}{2}\int_1^3 \sqrt{x^2+x+1}\mathrm{d}x \leqslant \sqrt{11}$.

   (c) Demuestre que $\displaystyle\left|\int_0^\pi \operatorname{sen}\sqrt{x}\,\mathrm{d}x\right| \leqslant \pi$.

2. Calcule las integrales

   (a) $\displaystyle\int_0^2 |t^2-t|dt$

   (b) $\displaystyle\int_{-1}^1 x\operatorname{sgn}x\,\mathrm{d}x$, donde $\operatorname{sgn}x = \dfrac{x}{|x|}$.

3. Un automóvil, durante un recorrido de 30 minutos, registra en su velocímetro que la velocidad instantánea $v(t)$ en el tiempo t es igual a

$$v(t) = t^2 + 1.$$

   ¿Cuál es el desplazamiento neto del automóvil durante ese intervalo de tiempo?

4. Sea $f(x)$ una función continua, $g(x)$ y $h(x)$ funciones derivables. Deduzca las fórmulas de derivación siguientes:

   (a)
$$\frac{\mathrm{d}}{\mathrm{d}x}\int_a^{g(x)} f(s)ds = f(g(x))\frac{\mathrm{d}g}{\mathrm{d}x}(x)$$

   (b)
$$\frac{\mathrm{d}}{\mathrm{d}x}\int_{h(x)}^{g(x)} f(s)ds = f(g(x))\frac{\mathrm{d}g}{\mathrm{d}x}(x) - f(h(x))\frac{\mathrm{d}h}{\mathrm{d}x}(x)$$

   (c)
$$\frac{\mathrm{d}}{\mathrm{d}x}\left(\int_0^b f(sx)\mathrm{d}s\right)(1) = -\int_0^b f(s)\mathrm{d}s + bf(b).$$

5. Sea $f(x) : (p,q) \to \mathbb{R}$ una función con k derivadas en el intervalo (p,q). Sea $a \in (p,q)$. Por el teorema fundamental del cálculo, para cada $c \in (p,q)$ podemos escribir

$$f(a) - f(c) = \int_c^a \frac{\mathrm{d}f}{\mathrm{d}x}(s)\mathrm{d}s.$$

(a) Utilizando dos veces la fórmula de integración por partes, escriba la integral de la derecha en la forma

$$\begin{aligned} f(a) - f(c) &= (s-a)\frac{\mathrm{d}f}{\mathrm{d}x}(s)\bigg|_c^a - \int_c^a (s-a)\frac{\mathrm{d}^2 f}{\mathrm{d}x^2}(s)ds \\ &= (a-c)\frac{\mathrm{d}f}{\mathrm{d}x}(c) + \frac{1}{2}(a-c)^2\frac{\mathrm{d}^2 f}{\mathrm{d}x^2}(c) + \frac{1}{2}\int_c^a (s-a)^2\frac{\mathrm{d}^3 f}{\mathrm{d}x^3}(s)ds. \end{aligned}$$

Como esto vale para dos puntos a, c de (p, q) arbitrarios, podemos escribir en lugar de a el valor x y en lugar de c el valor a para obtener

$$f(x) = f(a) + (x-a)\frac{\mathrm{d}f}{\mathrm{d}x}(a) + \frac{1}{2}(x-a)^2\frac{\mathrm{d}^2 f}{\mathrm{d}x^2}(a)$$

$$+\frac{1}{2}\int_a^x (s-x)^2\frac{\mathrm{d}^3 f}{\mathrm{d}x^3}(s)\mathrm{d}s.$$

Al término

$$R_a(x) = \frac{1}{2}\int_a^x (s-x)^2\frac{\mathrm{d}^3 f}{\mathrm{d}x^3}(s)ds,$$

se le llama *residuo de Taylor*[3] *de orden tres en forma integral.*

(b) Deduzca la forma integral del residuo de Taylor de orden n.

6. Sea $a > 0$, diga para qué valores de $p \in \mathbb{R}$ la función de tipo exponencial $f(x) = x^{-p}$ tiene integral impropia en $[a, \infty)$ y para qué valores de p tiene integral impropia en $(0, a)$.

7. Encuentre la integral de la función $f(x) : (0, 1) \to \mathbb{R}$, dada por

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

8. Utilizando el criterio de comparación (proposición 9.8):

   (a) Pruebe que la función e^{-x^2} tiene integral impropia en todo el intervalo $(-\infty, \infty)$. Sugerencia: Note que $0 \leqslant \mathrm{e}^{-x^2} \leqslant e^{-x}$ para toda $x \in [1, \infty)$ y que $0 \leqslant e^{-x^2} \leqslant 1$ para toda $x \in [0, 1]$.

   (b) Pruebe que la función $f(x) = \dfrac{1}{\sqrt{x + x^3}}$ tiene integral impropia en el intervalo $(0, \infty)$. Sugerencia: compare la función en cuestión con las funciones $\dfrac{1}{\sqrt{x}}$ y $\dfrac{1}{\sqrt{x^3}}$ en los subintervalos apropiados.

9. Calcule las integrales impropias siguientes:

   (a) $\displaystyle\int_0^1 x \ln x \mathrm{d}x$ (b) $\displaystyle\int_0^\infty x\mathrm{e}^{-x}\mathrm{d}x$ (c) $\displaystyle\int_0^{\frac{\pi}{2}} \tan x \mathrm{d}x$

[3] Por Brook Taylor, citado en los capítulos 1, 6 y 8.

10. Calcule el área de la región bajo la curva $y = \dfrac{4}{2x+1} - \dfrac{2}{x+2}$, arriba del eje de las abscisas y a la derecha de $x = 1$.

Capítulo

10

Aplicaciones de la integral definida

En este capítulo presentamos algunas de las aplicaciones elementales del cálculo integral a la geometría, la física y la ingeniería. En estos ejemplos se ilustra cómo los métodos del cálculo integral dan sentido preciso a las técnicas de agotamiento para el cálculo de sumas o resultantes de efectos infinitesimales.

10.1 Cálculo de áreas, volúmenes y longitudes

10.1.1 Áreas de regiones delimitadas por curvas suaves

Para $f_1 : [a, b] \to \mathbb{R}$ y $f_2 : [a, b] \to \mathbb{R}$, dos funciones continuas tales que

$$f_1(x) \leq f_2(x) \text{ para } x \in [a, b],$$

consideremos el problema del cálculo del área de la región $\mathcal{A}$ del plano delimitada por las gráficas de f_1 y f_2 y definida mediante el conjunto

$$\mathcal{A} = \{(x, y) \text{ tales que } a \leq x \leq b,\ f_1(x) \leq y \leq f_2(x)\},$$

el cual se muestra en la figura 10.1.

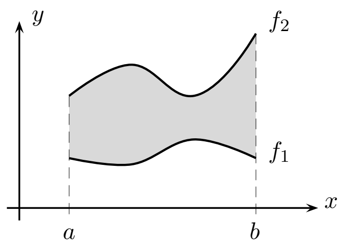


Figura 10.1 Área delimitada por las funciones f_1 y f_2 y las rectas $x = a$ y $x = b$

Este problema es equivalente al cálculo del área de la región delimitada por la gráfica de la función no-negativa g dada por

$$g(x) = f_2(x) - f_1(x), \quad x \in [a, b],$$

y el eje de las abscisas. Con base en las propiedades de la integral definida, vemos que el área de la región $\mathcal{A}$ se calcula con la fórmula

$$\text{Área de } \mathcal{A} = \int_a^b g(x)\mathrm{d}x = \int_a^b (f_2(x) - f_1(x))\mathrm{d}x.$$

Ejemplo 10.1 Calculemos el área de la región $\mathcal{A}$ del plano delimitada por la curva $y = 3 - x^2$ y la recta $y = -x + 1$.

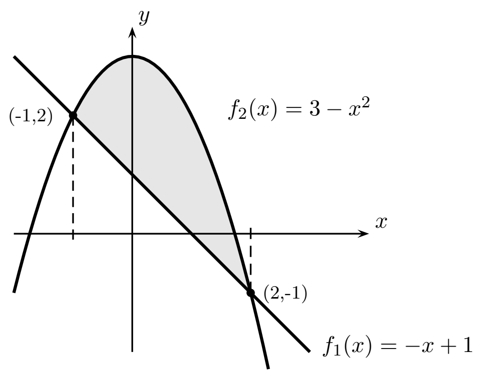


Figura 10.2 La región $\mathcal{A}$ del ejemplo 10.1

La región $\mathcal{A}$, como se muestra en la figura 10.2, está delimitada por las funciones $f_1(x) = -x + 1$ y $f_2(x) = 3 - x^2$ sobre el intervalo $[-1, 2]$, y por lo tanto, su área es

$$\text{Área de } \mathcal{A} = \int_{-1}^{2} (-x^2 + x + 2)\mathrm{d}x = \frac{9}{2}. \qquad \triangleleft$$

Nota Importante:
Para calcular el área de regiones más generales se puede utilizar el procedimiento anterior descomponiendo la región original en partes delimitadas por gráficas de dos curvas y rectas paralelas al eje de las ordenadas, calculándolas como en el caso anterior, y sumando después las áreas de esas regiones.

A veces conviene expresar las fronteras de la región cuya área se desea calcular, como gráficas de funciones de la variable y e integrar sobre intervalos en el eje de las ordenadas.

Ejemplo 10.2 El área de la región $\mathcal{B}$ delimitada por el eje de las ordenadas y las curvas $y = \operatorname{sen} x$ y $y = \cos x$,

$$\mathcal{B} = \left\{ (x, y) \text{ con } x \in \left[0, \frac{\pi}{4}\right], \quad \operatorname{sen} x \leqslant y \leqslant \cos x \right\}$$

(ver figura 10.3), se puede expresar como la suma del área de la región bajo la gráfica de la función $h_1(y) = \operatorname{arcsen} y$ sobre el intervalo $\left[0, \frac{\sqrt{2}}{2}\right]$ del eje de las ordenadas y

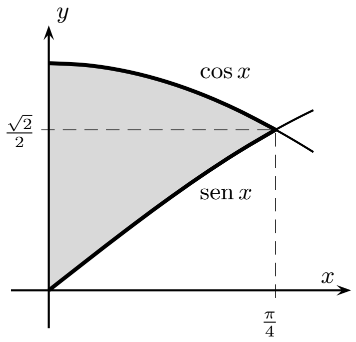


Figura 10.3 La región $\mathcal{B}$ del ejemplo 10.2

el área de la región bajo la gráfica de la función $h_1(y) = \arccos y$ sobre el intervalo $\left[\frac{\sqrt{2}}{2}, 1\right]$ del eje de las ordenadas.

Utilizando esta descomposición, se obtiene

$$\text{Área de } \mathcal{B} = \int_0^{\frac{\sqrt{2}}{2}} \operatorname{arcsen} y \mathrm{d}y + \int_{\frac{\sqrt{2}}{2}}^1 \arccos y \mathrm{d}y = \sqrt{2} - 1.$$

Por otro lado, esa misma región $\mathcal{B}$ se puede ver como delimitada por las gráficas de las funciones $f_1(x) = \cos x$ y $f_2(x) = \operatorname{sen} x$ sobre el intervalo $\left[0, \frac{\pi}{4}\right]$ y para su área se tiene

$$\text{Área de } \mathcal{B} = \int_0^{\frac{\pi}{4}} (\cos x - \operatorname{sen} x)\mathrm{d}x = (\operatorname{sen} x + \cos x)\,|_0^{\frac{\pi}{4}} = \sqrt{2} - 1. \qquad \triangleleft$$

10.1.2 Volúmenes de sólidos de revolución

En este apartado mostramos cómo se aplica la integral definida al cálculo del volumen de sólidos de revolución en dos casos importantes en las aplicaciones.

Primer caso: Sea $f : [a, b] \to \mathbb{R}$ una función continua y consideremos el sólido de revolución S que se genera al girar alrededor del eje de las abscisas la región delimitada por la gráfica de f sobre el intervalo $[a, b]$. Ver la figura 10.4.

Específicamente, el sólido de revolución S es el conjunto

$$S = \left\{(x, y, z) \text{ tales que } y^2 + z^2 \leq (f(x))^2 \text{ con } a \leqslant x \leqslant b\right\}.$$

Para calcular el volumen del sólido S observamos que cada uno de los subintervalos definidos por una partición $\mathcal{P} = \{a = x_0 < x_1 < \cdots < x_k = b\}$ de $[a, b]$ genera, al rotar alrededor del eje de las abscisas, un cilindro de altura Δx_i y radio $|f(x_i^*)|$. Ver figura 10.5.

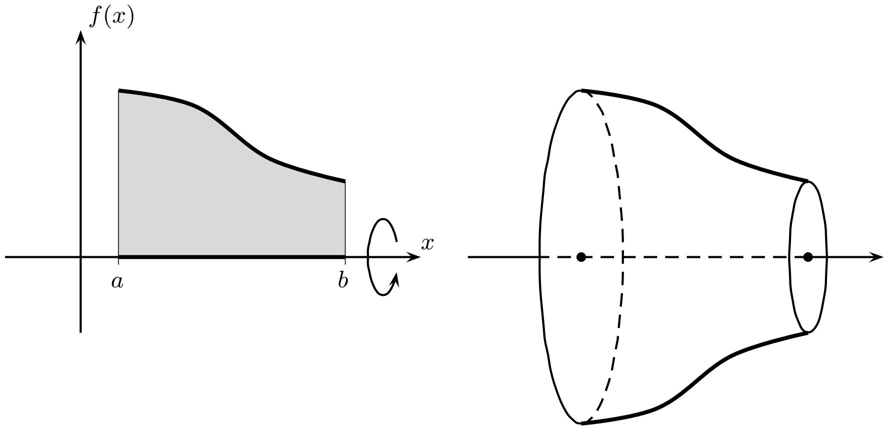


Figura 10.4 Sólido de revolución

El volumen de la unión de tales cilindros,

$$V(f, \mathcal{P}, \{x_i^*\}) = \pi \sum_{i=1}^{k} f^2(x_i^*)\Delta x_i, \tag{10.1}$$

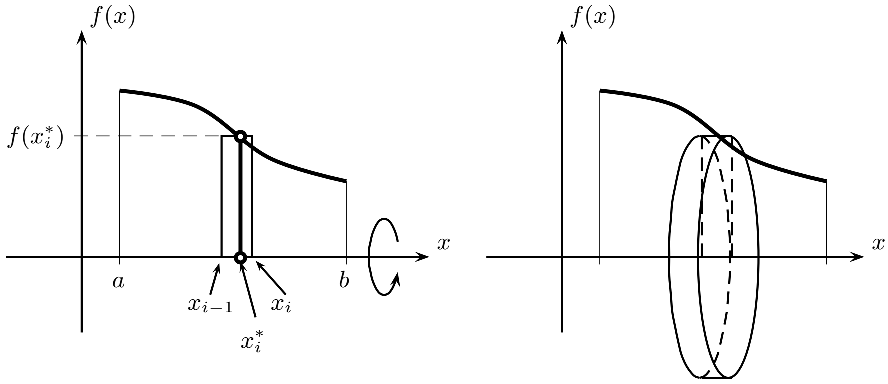


Figura 10.5 Cálculo del volumen de un sólido de revolución

aproxima al volumen de S y corresponde a la suma de Riemann de la función $\pi f^2(x)$ asociada a la partición $\mathcal{P}$. En el límite, cuando $k \to \infty$, las sumas (10.1) convergen al número $\displaystyle\int_a^b \pi f^2(x)\mathrm{d}x$, que es el volumen del sólido de revolución S. Es decir,

$$V = \text{Volumen de } S = \pi \int_a^b f^2(x)\mathrm{d}x. \tag{10.2}$$

Ejemplo 10.3 El volumen del elipsoide de revolución

$$E = \left\{(x, y, z) \text{ tales que } \frac{x^2}{a^2} + \frac{y^2}{b^2} + \frac{z^2}{b^2} \le 1\right\}$$

es generado al rotar la gráfica de la función

$$f : [-a, a] \to \mathbb{R}$$
$$f(x) = b\sqrt{1 - \frac{x^2}{a^2}}$$

y el cálculo del volumen nos arroja la expresión

$$V = \pi b^2 \int_{-a}^{a} \left(1 - \frac{x^2}{a^2}\right) \mathrm{d}x = \frac{4}{3}\pi b^2 a.^1 \qquad \triangleleft$$

Segundo caso: Para sólidos de revolución generados por la rotación de la gráfica de una función continua $f : [a, b] \to \mathbb{R}$, $a \geqslant 0$, alrededor del eje de las ordenadas, como se muestra en la figura 10.6, el cálculo de su volumen puede realizarse de manera análoga.

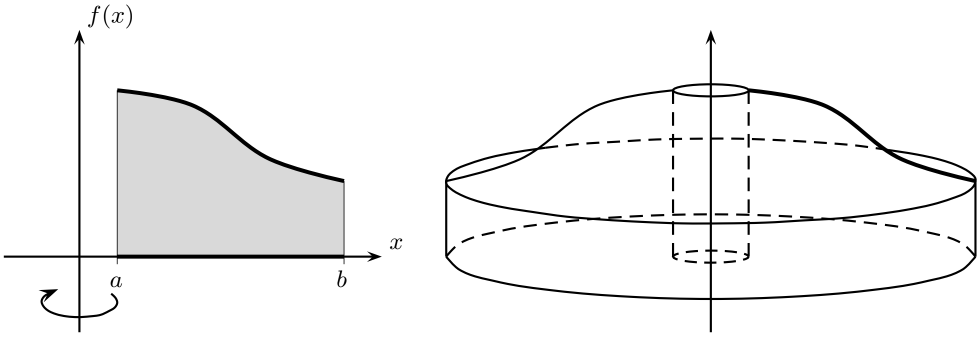


Figura 10.6 Sólido de revolución

De manera específica, cada subintervalo $[x_{i-1}, x_i]$ de una partición $\mathcal{P} = \{a = x_0 < x_1 < \cdots < x_k = b\}$ de $[a, b]$ genera, al girar alrededor del eje de las ordenadas, un anillo cilíndrico de espesor Δx_i y altura $|f(x_i^*)|$, donde $x_i^* \in [x_{i-1}, x_i]$, como se muestra en la figura 10.7.

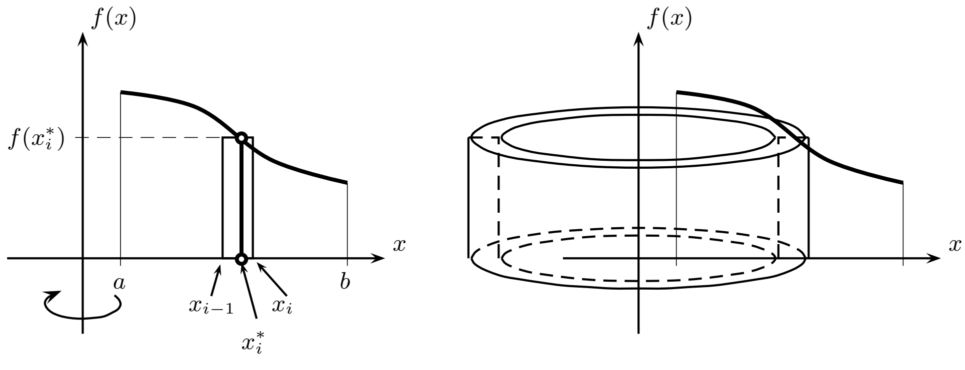


Figura 10.7 Cálculo del volumen de un sólido de revolución

[1] Nótese que si $a = b$, E es una esfera y V su volumen.

La suma de los volúmenes de estos anillos cilíndricos es

$$V(f,\mathcal{P},\{x_i^*\}) = \pi\sum_{i=1}^{k}|f(x_i^*)|(x_{i+1}^2 - x_i^2) = \pi\sum_{i=1}^{k}(x_{i+1}+x_i)|f(x_i^*)|\Delta x_i. \qquad (10.3)$$

En el límite, cuando $k \to \infty$, la suma (10.3) converge al número

$$V = \text{Volumen de } S = 2\pi\int_a^b x|f(x)|\mathrm{d}x, \qquad (10.4)$$

que es el volumen exacto del sólido de revolución S.

Ejemplo 10.4 Consideremos el sólido que se genera al rotar alrededor del eje de las ordenadas la gráfica de la curva $f(x) = 2x^2 - x^3$ con $x \in [0,2]$. Aplicando la fórmula (10.4) obtenemos

$$V = 2\pi\int_0^2 x|f(x)|\mathrm{d}x = 2\pi\int_0^2 (2x^3 - x^4)\mathrm{d}x = \frac{16}{5}\pi. \qquad \triangleleft$$

10.1.3 Longitudes de curvas

El cálculo de longitudes de curvas en el plano cartesiano es otra aplicación importante de la integral definida. Si consideramos la curva Γ dada por la ecuación

$$y = f(x) \text{ para } x \in [a,b],$$

donde $f(x)$ es una función con derivada continua, podemos calcular la longitud de Γ aproximándola por curvas poligonales, es decir, formadas por segmentos de recta, cuya longitud es directamente calculable mediante la fórmula de la distancia entre dos puntos del plano cartesiano. Esto puede realizarse tomando una partición $\mathcal{P}$ $= \{a = x_0 < x_1 < \cdots < x_k = b\}$ de $[a,b]$ y construyendo la curva poligonal $s(x)$ que pasa por los puntos $(x_i, f(x_i))$ para $i = 1, \cdots k$, de la gráfica de Γ, como se muestra en la figura 10.8.

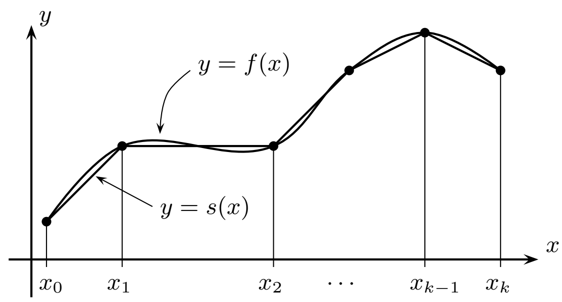


Figura 10.8 Aproximación de la longitud de arco

La longitud del segmento s_i determinado por los puntos $(x_{i-1}, f(x_{i-1}))$ y

$(x_i, f(x_i))$, es igual a

$$\text{longitud de } s_i = \sqrt{\Delta x_i^2 + (f(x_i) - f(x_{i-1}))^2}. \tag{10.5}$$

Aplicando el teorema del valor medio a la función $f(x)$ en cada subintervalo $[x_{i-1}, x_i]$, podemos escribir

$$f(x_i) - f(x_{i-1}) = \frac{\mathrm{d}f}{\mathrm{d}x}(x_i^*)\Delta x_i, \tag{10.6}$$

donde $x_i^* \in [x_{i-1}, x_i]$. Sustituyendo (10.6) en (10.5) y sumando se obtiene

$$\text{longitud de } s = \sum_{i=1}^{k} s_i = \sum_{i=1}^{k} \sqrt{1 + \left(\frac{\mathrm{d}f}{\mathrm{d}x}(x_i^*)\right)^2}\Delta x_i,$$

que es la suma de Riemann de la función $\sqrt{1 + \left(\frac{\mathrm{d}f}{\mathrm{d}x}(x)\right)^2}$ correspondiente a la partición $\mathcal{P}$ y a la elección de puntos intermedios $\{x_i^*\}$. Tomando el límite de esas sumas cuando $|\mathcal{P}| \to 0$, obtenemos

$$\text{longitud de } \Gamma = \int_a^b \sqrt{1 + \left(\frac{\mathrm{d}f}{\mathrm{d}x}(x)\right)^2}\,\mathrm{d}x. \tag{10.7}$$

Ejemplo 10.5 La longitud de la curva Γ en el plano cuya ecuación es $y = x^3$, $x \in [a, b]$, es

$$\text{longitud de } \Gamma = \int_a^b \sqrt{1 + 9x^4}\mathrm{d}x. \qquad \triangleleft$$

Nota Importante:

Frecuentemente, y aún en aplicaciones del cálculo a problemas sencillos, la integración de funciones elementales es un problema difícil. Más aún, se ha demostrado que muchas funciones elementales no poseen antiderivadas que se puedan expresar en términos de una cantidad finita de funciones elementales. Por ejemplo, al calcular la longitud de la elipse

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1,$$

se obtiene, aplicando la fórmula (10.7) a $f(x) = \frac{b}{a}\sqrt{a^2 - x^2}$, *que*

$$\begin{aligned} \textit{longitud de la elipse} &= 4\int_0^a \sqrt{1 + \frac{b^2}{a^2}\frac{x^2}{a^2 - x^2}}\mathrm{d}x \\ &= 4a\int_0^{\frac{\pi}{2}} \sqrt{1 + \left(\frac{b^2}{a^2} - 1\right)\operatorname{sen}^2\theta}\, d\theta. \end{aligned} \tag{10.8}$$

Sin embargo, la antiderivada de la función $g(\theta) = \sqrt{1 + \left(\dfrac{b^2}{a^2} - 1\right) \operatorname{sen}^2 \theta}$ *no es expresable en términos de una cantidad finita de funciones elementales. A las integrales de la forma (10.8) se les llama integrales elípticas. Para calcularlas se deben usar otros métodos (numéricos, por ejemplo).*

10.2 Área de superficies de revolución

Antes de abordar el problema del cálculo del área de una superficie de revolución, recordemos que el área lateral de un cono circular recto de radio r y altura h se puede calcular aproximándola por una suma de áreas de triángulos de base $r\Delta\theta$ y altura $\sqrt{h^2 + r^2}$, como en la figura 10.9, de tal manera que

$$\text{Área del cono} = \int_0^{2\pi} \frac{1}{2}\sqrt{r^2 + h^2}\, r \, \mathrm{d}\theta = \pi r \sqrt{h^2 + r^2}. \tag{10.9}$$

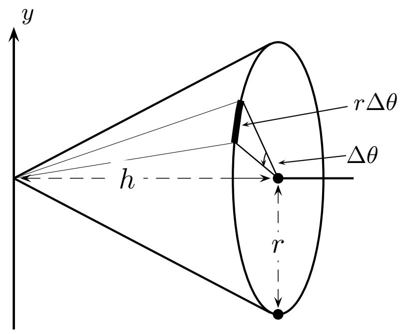


Figura 10.9 Un cono circular recto

A partir de la fórmula (10.9), deducimos que el área lateral de un cono recto truncado de base mayor de radio R y base menor de radio r y altura h, como el que se muestra en la figura 10.10, es

$$\text{Área del cono truncado recto} = \pi(R + r)\sqrt{h^2 + (R - r)^2}. \tag{10.10}$$

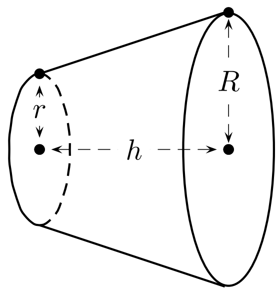


Figura 10.10 Un cono truncado recto

Con la fórmula (10.10) podemos calcular el área externa de un sólido aproximando su área lateral por una suma de áreas laterales de conos truncados cuyas paredes aproximan la superficie exterior del sólido. Más explícitamente, consideremos la superficie de revolución generada al rotar alrededor del eje de las abscisas la gráfica de la función no-negativa f sobre el intervalo $[a,b]$. Véase la figura 10.11.

Para cada partición $\mathcal{P} = \{a = x_0 < x_1 < \cdots < x_k = b\}$ de $[a,b]$ consideremos la curva poligonal $s(x)$ que pasa por los puntos $(x_i, f(x_i))$ para $i = 1, 2, \ldots, k$. Al rotar la curva $s(x)$, cada subintervalo $[x_{i-1}, x_i]$, genera un cono circular truncado cuyas bases tienen por radios $R = f(x_i)$ y $r = f(x_{i-1})$ y altura $h = x_i - x_{i-1}$, como se ve en la figura 10.11.

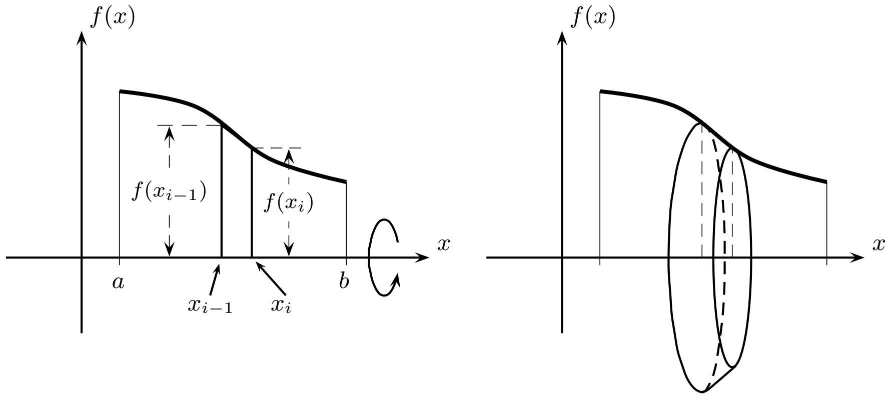


Figura 10.11 Cálculo del área de una superficie de revolución

El área de los conos truncados generados por la partición $\mathcal{P}$ toma el valor:

$$\begin{array}{c}\text{Suma}\\ \text{de áreas laterales}\\ \text{de conos}\end{array} = \pi \sum_{i=1}^{k} (f(x_i) + f(x_{i-1}))\sqrt{\Delta x_i^2 + (f(x_i) - f(x_{i-1}))^2}. \quad (10.11)$$

Aplicando el teorema del valor medio a la función f en cada subintervalo $[x_{i-1}, x_i]$ podemos escribir

$$f(x_i) - f(x_{i-1}) = \frac{\mathrm{d}f}{\mathrm{d}x}(x_i^*)\Delta x_i \quad (10.12)$$

para algún $x_i^* \in (x_{i-1}, x_i)$. Sustituyendo (10.12) en (10.11), tendremos,

$$\begin{array}{c}\text{suma}\\ \text{de áreas laterales}\\ \text{de conos}\end{array} = \pi \sum_{i=1}^{k} (f(x_i) + f(x_{i-1}))\sqrt{1 + \left(\frac{\mathrm{d}f}{\mathrm{d}x}(x_i^*)\right)^2}\,\Delta x_i,$$

la cual es una suma de Riemann de la función

$$\ell(x) = 2\pi f(x)\sqrt{1 + \left(\frac{\mathrm{d}f}{\mathrm{d}x}(x_i^*)\right)^2}$$

correspondiente a la partición $\mathcal{P}$ y a la elección de puntos intermedios x_i^*, $i = 1, 2, \ldots, k$. Al tomar particiones de $[a,b]$ con norma tendiente a cero, la suma de áreas laterales de los conos truncados converge a la integral de la función $\ell(x)$; es decir,

$$\begin{array}{c}\text{área de la}\\ \text{superficie de revolución}\end{array} = 2\pi \int_a^b f(x)\sqrt{1+\left(\frac{\mathrm{d}f}{\mathrm{d}x}(x)\right)^2}\,\mathrm{d}x. \tag{10.13}$$

Ejemplo 10.6 El área de la esfera corresponde al área de la superficie de revolución generada al rotar la gráfica de la función

$$f(x) = \sqrt{r^2 - x^2}, \quad x \in [-r, r]$$

alrededor del eje de las abscisas. Aplicando la fórmula (10.13) se tiene:

$$\begin{aligned}\begin{array}{c}\text{área de la esfera}\\ \text{de radio } r\end{array} &= 2\pi \int_{-r}^{r} \sqrt{r^2-x^2}\sqrt{1+\frac{x^2}{r^2-x^2}}\mathrm{d}x\\ &= 2\pi \int_{-r}^{r} r\mathrm{d}x = 4\pi r^2.\end{aligned}$$

◁

10.3 Centros de masa y presión de fluidos

En este apartado presentamos algunas aplicaciones de la integral a problemas del cálculo de centros de masa (o centroides) de varillas y regiones planas.

10.3.1 Centroides de varillas y regiones planas

Consideremos una varilla de longitud L de densidad variable, de tal manera que la masa del material que forma la varilla por unidad de longitud es una función $\rho(x)$, donde $x \in [0, L]$. Si colocamos la varilla sobre un pivote colocado en el punto x_M, como se muestra en la figura 10.12, y consideramos la acción de la fuerza de gravedad sobre cada uno de sus puntos, la varilla tenderá a rotar en la dirección de las manecillas del reloj por efecto de la fuerza de palanca generada por el peso de los puntos a la derecha del pivote y en sentido contrario por la fuerza de palanca de los puntos a la izquierda de x_M.

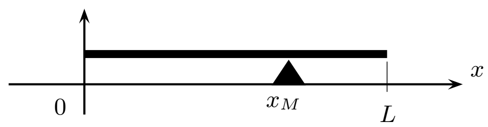


Figura 10.12 Centroide de una varilla

La fuerza de palanca o “torca” ejercida por cada masa puntual de la varilla es igual al producto de la distancia de ese punto al pivote, multiplicado por el peso del material concentrado en ese punto. La resultante de la suma de las torcas ejercidas por los puntos a la derecha e izquierda del pivote, tiene un efecto final sobre la varilla, haciéndola girar en el sentido correspondiente al signo de la torca resultante. Se define el *centro de masa de la varilla* (o *centroide*) como aquel punto-pivote x_M con respecto al cual la torca resultante ejercida por todos los puntos de la varilla es cero.

Para el cálculo de la torca resultante respecto a un punto-pivote x_M, se recurre al método de agotamiento, considerando a la varilla como formada por un conjunto finito de segmentos dispuestos a lo largo de la varilla y donde cada uno de ellos tiene densidad constante e igual al valor de la función $\rho(x)$ en el punto x del segmento, elegido arbitrariamente. Bajo esa aproximación, la suma de las torcas ejercidas por esos segmentos tenderá, cuando la longitud de los segmentos es cada vez más pequeña, a la torca de la varilla alrededor del punto-pivote x_M. Para realizar lo anterior, tomamos una partición $\mathcal{P} = \{0 = x_0 < x_1 < \cdots < x_k = L\}$ de $[0, L]$ y consideramos la suma

$$T(\rho, \mathcal{P}, \{x_i^*\}) = \sum_{i=1}^{k} \rho(x_i^*)(x_i^* - x_M)\Delta x_i, \tag{10.14}$$

donde x_i^* es un punto arbitrario de $[x_{i-1}, x_i]$ para cada $i = 1, 2, \ldots, k$. La expresión (10.14) es la suma de Riemann para la función $\rho(x)(x - x_M)$ correspondiente a la partición $\mathcal{P}$ y la elección de puntos $\{x_i^*\}$, y por lo tanto, para cada elección del punto pivote x_M, el valor de la integral $\int_0^L \rho(x)(x - x_M)\mathrm{d}x$ corresponde a la torca ejercida por la varilla alrededor de ese punto. Luego, el centro de masa buscado corresponderá al punto x_M tal que

$$\int_0^L \rho(x)(x - x_M)\mathrm{d}x = 0,$$

o explícitamente,

$$x_M = \frac{\int_0^L \rho(x)x\mathrm{d}x}{\int_0^L \rho(x)\mathrm{d}x}. \tag{10.15}$$

Ejemplo 10.7 La determinación del centro de masa de una varilla de longitud L y función de densidad lineal $\rho(x) = ax + b$ con $a, b > 0$ y $x \in [0, L]$, se obtiene directamente de la fórmula (10.15) y se sitúa a una distancia x_M del extremo izquierdo dada por

$$x_M = \frac{\int_0^L (ax + b)x\mathrm{d}x}{\int_0^L (ax + b)\mathrm{d}x} = \frac{\frac{1}{3}aL^2 + \frac{1}{2}bL}{\frac{1}{2}aL + b}.$$

Se deduce inmediatamente que el centro de masa de una varilla de densidad constante (es decir, haciendo $a = 0$) se localiza en el punto medio de la varilla. $\triangleleft$

Enseguida, veamos cómo obtener el centro de masa, o *centroide*, de una superficie S delimitada por curvas suaves y hecha de un material de densidad constante. Aquí también el centroide se caracteriza por ser aquel punto en el plano donde la resultante de las torcas medidas con respecto a ese punto como pivote tienen resultante cero, de tal manera de que si se sostiene la superficie con un pivote en ese punto, ésta permanecerá en equilibrio.

Supongamos que S está dada por la región bajo la gráfica de una función $f(x)$ sobre un intervalo $[a,b]$, como se muestra en la figura 10.13(a).

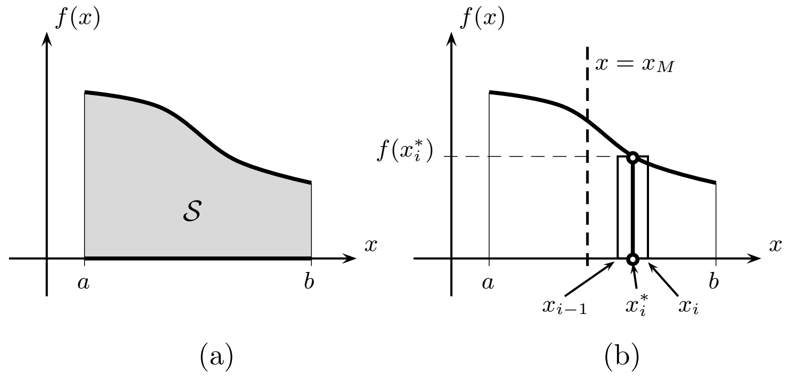


Figura 10.13 Cálculo del centroide de una región delimitada por una curva

Para determinar las coordenadas del centroide, primero buscaremos la recta de la forma $x = x_M$, con respecto a la cual la torca ejercida por los puntos de la superficie, tanto a la derecha como a la izquierda de dicha recta, es cero. Esto quiere decir que la superficie se mantendrá en equilibrio y no rotará alrededor del eje definido por la recta perpendicular al eje de las abscisas con ecuación $x = x_M$. Para calcular la torca total alrededor de la recta pivote, tomamos particiones $\mathcal{P} = \{a = x_0 < x_1 < \cdots < x_k = b\}$ de $[a,b]$ y consideramos la torca ejercida por cada uno de los rectángulos de base $[x_{i-1}, x_i]$ y altura $|f(x_i^*)|$, con x_i^* elegido arbitrariamente en el subintervalo $[x_{i-1}, x_i]$, obteniendo así las sumas de la forma

$$T(f, \mathcal{P}, \{x_i^*\}) = \sum_{i=1}^{k} \rho |f(x_i^*)|(x_i^* - x_M)\Delta x_i,$$

que corresponden a las sumas de Riemann para la función

$$h(x) = \rho |f(x)|(x - x_M).$$

Tomando el límite de tales sumas cuando las particiones tienen norma tendiente a cero, obtenemos que la posición de la recta $x = x_M$ con respecto a la cual la torca total es cero, debe satisfacer la relación

$$\int_a^b \rho |f(x)|(x - x_M)\mathrm{d}x = 0;$$

es decir,

$$x_M = \frac{\int_a^b \rho|f(x)|x\mathrm{d}x}{\int_a^b \rho|f(x)|\mathrm{d}x} = \frac{\int_a^b |f(x)|x\mathrm{d}x}{\int_a^b |f(x)|\mathrm{d}x}.$$

Para el cálculo de la ordenada y_M del centroide consideremos, para la partición $\mathcal{P} = \{a = x_0 < x_1 < \cdots < x_k = b\}$, las franjas verticales bajo la gráfica de f sobre cada subintervalo $[x_{i-1}, x_i]$. La longitud de cada franja es, aproximadamente, $f(x_i^*)$, con $x_i^* \in [x_{i-1}, x_i]$ y tienen su centroide en el punto $(x_i^*, f(x_i^*)/2)$. Consideremos ahora el conjunto S de puntos $(x_i^*, f(x_i^*)/2)$ para $i = 1, 2, \ldots, k$ con una masa concentrada en cada uno de ellos igual a $\rho|f(x_i^*)|\Delta x_i$. Calculemos ahora la torca que hace el conjunto S con respecto a una recta horizontal $y = y_M$, la cual toma la forma

$$\sum_{i=1}^{k} \rho\left(\frac{1}{2}f(x_i^*) - y_M\right)|f(x_i^*)|\Delta x_i.$$

Tomando particiones cada vez más finas, tenemos que la torca total del conjunto bajo la gráfica de f es

$$\int_a^b \rho\left(\frac{1}{2}f(x) - y_M\right)|f(x)|\mathrm{d}x,$$

misma que se anulará si

$$y_M = \frac{\int_a^b \rho f(x)|f(x)|\mathrm{d}x}{2\int_a^b \rho|f(x)|\mathrm{d}x} = \frac{\int_a^b f(x)|f(x)|\mathrm{d}x}{2\int_a^b |f(x)|\mathrm{d}x}. \tag{10.16}$$

Finalmente, tomando en cuenta que la masa total M de la superficie es $M = \int_a^b \rho|f(x)|\mathrm{d}x$, concluimos que las coordenadas de su centroide son

$$x_M = \frac{1}{M}\int_a^b \rho x|f(x)|\mathrm{d}x, \quad y_M = \frac{1}{2M}\int_a^b \rho f(x)|f(x)|\mathrm{d}x. \tag{10.17}$$

Ejemplo 10.8 Calcule las coordenadas del centro de masa de la región delimitada por las gráficas de dos funciones $f : [a, b] \to \mathbb{R}$ y $g : [a, b] \to \mathbb{R}$.

Solución. Un argumento análogo al anterior para el cálculo de la abcisa x_M del centro de masa nos remite a la fórmula

$$x_M = \frac{\int_a^b x|f(x) - g(x)|\mathrm{d}x}{\int_a^b |f(x) - g(x)|\mathrm{d}x},$$

mientras que para la ordenada y_M se tiene

$$y_M = \frac{\int_a^b x|f(x) + g(x)||f(x) - g(x)|\mathrm{d}x}{2\int_a^b |f(x) - g(x)|\mathrm{d}x}. \qquad \triangleleft$$

Ejemplo 10.9 (Teorema de Pappus[2]) En el caso de un sólido de revolución generado al girar la gráfica de una función no-negativa $f : [a, b] \to \mathbb{R}$ alrededor del eje de las abscisas, sabemos que su volumen está dado por la fórmula

$$V = \pi \int_a^b f^2(x)\mathrm{d}x. \tag{10.18}$$

Por otro lado, sustituyendo (10.18) en la fórmula (10.16) para la ordenada y_M del centroide de la región delimitada por la gráfica de f y el eje de las abscisas se obtiene la relación

$$2\pi y_M \int_a^b f(x)\mathrm{d}x = \pi \int_a^b f^2(x)\mathrm{d}x,$$

la cual verifica el llamado *teorema de Pappus*, que establece que el volumen de un sólido de revolución generado al rotar una región plana R alrededor de una recta a la cual no intersecta, es igual al área de la región R que lo genera, multiplicado por la distancia que recorre el centroide de R al efectuar la rotación. ◁

10.3.2 Presión de líquidos sobre superficies

Como consecuencia de su peso y de su naturaleza deformable, los líquidos se "recargan" sobre las superficies con las que están en contacto y ejercen una determinada presión. En cada punto de una superficie y sobre cada sección de ella suficientemente pequeña que contiene a ese punto, la presión (o fuerza ejercida por unidad de área) que ejerce el líquido es la misma en cualquier dirección y su magnitud es igual al peso de la columna de líquido sobre esa unidad de área. A una profundidad de h unidades, la presión ejercida por el líquido es

$$p(h) = \rho g h,$$

donde g es la aceleración producida por la fuerza de gravedad y ρ es la densidad del líquido. A la ley anterior se le conoce como *principio de Pascal*[3] .

Veamos ahora cómo se aplica la integral definida para el cálculo de la presión ejercida por un líquido sobre superficies de distinta geometría.

Ejemplo 10.10 Consideremos un cilindro de radio r que descansa en el fondo de un estanque de h metros de profundidad con $2r < h$. Véase la figura 10.14(a). Nos interesa calcular la presión que ejerce el agua sobre cada una de sus tapaderas.

Para calcular la presión, situaremos un par de ejes de coordenadas de tal manera que la tapa en cuestión esté dada por el conjunto

$$\{(x, y) \quad \text{con} \quad x^2 + y^2 \leqslant r^2\},$$

como se indica en la figura 10.14(b). Tomemos una partición $\mathcal{P} = \{-r = y_0 < y_1 < \cdots < y_k = r\}$ del intervalo $[-r, r]$ sobre el eje de las ordenadas. Cada subintervalo

[2]Pappus (290-350, aprox.), quien vivió en Alejandría.

[3]Por Blas Pascal (1623-1662), matemático francés, su descubridor.

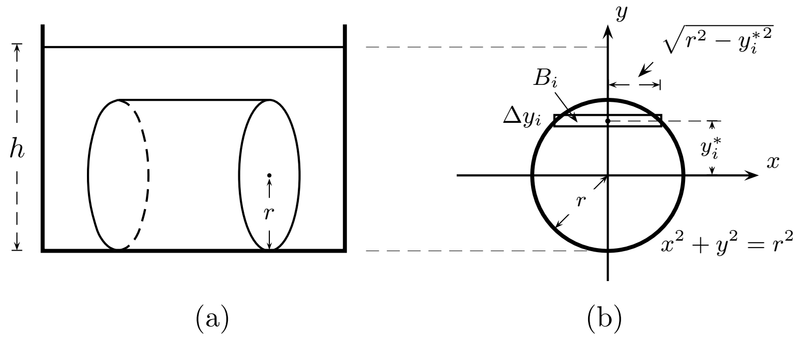


Figura 10.14 Presión sobre la tapa de un tanque

de la partición $[y_{i-1}, y_i]$ da lugar a una banda horizontal B_i cuya área es, aproximadamente, $2\sqrt{r^2 - y_i^{*2}}\Delta y_i$, donde $y_i^* \in [y_{i-1}, y_i]$. Por el principio de Pascal, sobre la banda B_i el agua, cuya densidad es $\rho = 1$, ejerce una presión constante con valor approximado

$$2g\sqrt{r^2 - y_i^{*2}}(h - r - y_i^*)\Delta y_i.$$

Sumando las presiones ejercidas sobre cada una de las bandas B_i para $i = 1, 2, \ldots, k$, se tiene

$$2g\sum_{i=1}^{k}\sqrt{r^2 - y_i^{*2}}(h - r - y_i^*)\Delta y_i. \tag{10.19}$$

Cuando $|\mathcal{P}| \to 0$ las sumas de Riemann (10.19) tienden a

$$P = 2g\int_{-r}^{r}\sqrt{r^2 - y^2}(h - r - y)\mathrm{d}y, \tag{10.20}$$

que es el valor exacto de la presión del agua sobre cada tapa del tanque. ◁

Ejemplo 10.11 Consideremos el problema de calcular la presión que un líquido ejerce sobre las paredes del recipiente que lo contiene. Supongamos que el recipiente está lleno de agua ($\rho = 1$) y que tiene la forma de una superficie de revolución generada al rotar la gráfica de la función no negativa $f(y)$, con $y \in [c, d]$, alrededor del eje de las ordenadas, como se muestra en la figura 10.15.

Aproximemos la pared del recipiente con elementos de superficie, como en el caso del cálculo del área de una superficie de revolución. Aplicando el principio de Pascal se obtiene que la presión total es

$$P = 2\pi g\int_{c}^{d}(d - y)f(y)\sqrt{1 + \left(\frac{\mathrm{d}f}{\mathrm{d}y}(y)\right)^2}\mathrm{d}y. \tag{10.21}$$

En caso de que el agua sólo alcance una altura $k < d$, la integral en (10.21) se calcula únicamente en el intervalo $[c, k]$.

Como aplicaciones directas de (10.21), presentamos los casos siguientes.

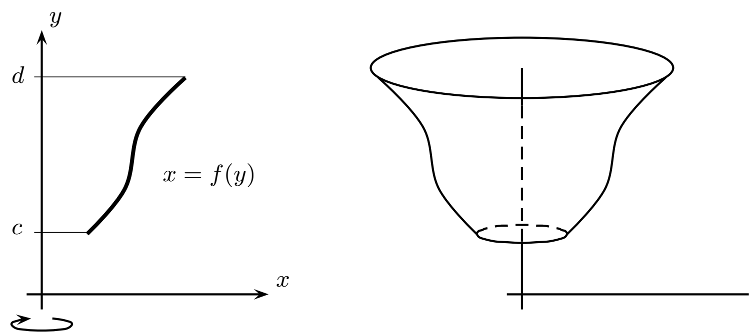


Figura 10.15 Presión sobre las paredes de un recipiente

1. Si $f(y) = r$ entonces el recipiente es cilíndrico y la presión sobre las paredes del cilindro es
$$P = 2\pi gr \int_c^d (d - y)\mathrm{d}y = \pi grh^2.$$

2. Si el recipiente es un cono recto de radio r y altura h con vértice en el suelo, entonces la función que lo genera, por rotación, es $f(y) = \dfrac{r}{h}y$ y la presión total sobre las paredes es
$$P = 2\pi g \int_0^h (h - y)\frac{r}{h^2}y\sqrt{r^2 + h^2}\mathrm{d}y = \frac{1}{3}\pi grh\sqrt{r^2 + h^2}.$$

3. Si el recipiente es esférico de radio r, entonces $f(y) = \sqrt{r^2 - y^2}$ y la presión sobre las paredes es
$$P = 2\pi g \int_{-r}^r (r - y)r\mathrm{d}y = 4\pi gr^3. \qquad \triangleleft$$

Ejercicios y problemas del capítulo

1. Calcule el área de la región delimitada por la curva cerrada $y^2 = x^2 - x^4$.
2. Calcule el área de la región delimitada por la elipse $\dfrac{x^2}{a^2} + \dfrac{y^2}{b^2} = 1$.
3. Calcule el área de la región comprendida entre las parábolas $y = -x^2 + 2$ y $y = 2x^2 - 3$.
4. Calcule el volumen de un cono recto de altura h y radio r.
5. Demuestre que la longitud de la circunferencia de radio r es igual a $2\pi r$.
6. Calcule el área del elipsoide de revolución de eje mayor a y eje menor b.
7. Calcule el área de la superficie de revolución generada al rotar la curva $y = \dfrac{1}{x}$ sobre el intervalo $(0, \infty)$.
8. Calcule el área de la superficie de revolución generada al girar el círculo $x^2 + y^2 = r^2$ alrededor de la recta $y = r$.
9. La superficie de una cortina de una presa está inclinada y forma un ángulo de $30°$ con la vertical, tiene la forma de un trapezoide isósceles de 50 metros de coronamiento y 25 metros de ancho en el fondo con una altura inclinada de 35 metros medidos sobre la pared. Calcule la presión del agua sobre la cortina cuando la presa está llena.
10. Sea S la región delimitada por las curvas $y = x^m$ y $y = x^n$ para $x \in [0, 1]$, donde m y n son enteros y $0 \leqslant n < m$.
    (a) Dibuje la región S.
    (b) Calcule las coordenadas del centroide de S.
    (c) Determine para qué valores de n y m el centroide de S no está contenido en S.
11. Un tanque de agua que descansa horizontalmente tiene extremos con la forma de la región que se extiende entre la parábola $y = \frac{x^2}{2}$ y la recta $y = 12$. Calcule la presión que ejerce el agua sobre cada extremo si su nivel está a 2 metros de altura.
12. Calcule el centroide de cada una de las superficies siguientes.
    (a) Un cuadrante de un círculo de radio r.
    (b) Una pieza formada de un rectángulo de altura h coronado por medio de un semicírculo de radio r.
    (c) Una pieza de forma de triángulo rectángulo con catetos de longitudes a y b.

Capítulo

11 Ecuaciones diferenciales elementales y aplicaciones

El gran impacto que ha tenido el cálculo diferencial e integral en las ciencias y tecnología modernas se debe al desarrollo de la teoría de las ecuaciones diferenciales. El concepto de ecuación diferencial es el apropiado para la formulación de las leyes dinámicas que gobiernan los fenómenos naturales y el control de los procesos de la industria y la tecnología.

De manera análoga al concepto de ecuación algebraica, que se introduce en el ámbito de los números y las operaciones aritméticas, una ecuación diferencial es una expresión en términos de funciones y las operaciones propias de ellas, que incluyen la toma de derivadas de cualquier orden. En el caso de las ecuaciones diferenciales, las incógnitas son funciones y el problema de encontrarlas o "despejarlas" es el objetivo de los métodos de solución o integración de esas ecuaciones.

Entre las ecuaciones diferenciales más importantes en las aplicaciones, están las ecuaciones de movimiento de Newton, la ecuación de Navier-Stokes para la dinámica de fluidos, las ecuaciones del electromagnetismo de Maxwell y muchas otras en las diferentes áreas de la ciencia.

En este capítulo, y como una introducción a la teoría de las ecuaciones diferenciales, se estudia la familia de las ecuaciones diferenciales lineales con coeficientes constantes de primero y segundo orden y se muestra cómo se aplican los métodos del cálculo desarrollados previamente para calcular sus soluciones. Se presentan varias aplicaciones al movimiento de los cuerpos en la vecindad de la superficie de la Tierra.

11.1 El concepto de ecuación diferencial

Las ecuaciones algebraicas, como las ecuaciones de segundo grado o los sistemas de ecuaciones lineales, que conocemos desde la escuela secundaria, son expresiones que incluyen operaciones algebraicas entre números e incógnitas igualadas al número cero. Por ejemplo, la ecuación de segundo grado

$$x^2 + ax + b = 0,$$

o el sistema de dos ecuaciones lineales no homogéneo en dos incógnitas y con coeficientes reales

$$\begin{cases} ax + by - c = 0 \\ dx + ey - f = 0. \end{cases}$$

Resolver estas ecuaciones, significa encontrar valores para las variables o incógnitas x, y de tal manera que al sustituirlos en las expresiones se obtenga una identidad.

Análogamente al caso algebraico, una ecuación diferencial es una expresión entre funciones e incógnitas relacionadas mediante operaciones propias de funciones, que incluye las de derivación de los distintos órdenes.

Por ejemplo, la expresión

$$\frac{\mathrm{d}^2 y}{\mathrm{d}x^2} + x\frac{\mathrm{d}y}{\mathrm{d}x} - (x^2+2)\cos x = 0, \quad x \in (a,b), \tag{11.1}$$

define una ecuación diferencial en (a, b) donde la incógnita es la función $y(x)$, la cual aparece en primera y segunda derivada sumadas y multiplicadas con funciones conocidas. Al orden máximo de derivación al cual se somete la función incógnita se le llama el *orden de la ecuación;* así, (11.1) es una ecuación diferencial de segundo orden.

Resolver la ecuación diferencial significa encontrar las funciones $y(x)$ definidas en (a, b) tales que al sustituirlas en la ecuación se obtenga una identidad entre los términos a la derecha e izquierda del signo de igualdad.

Ejemplo 11.1 La función

$$y(x) = x \operatorname{sen} x$$

es una solución de la ecuación diferencial

$$\frac{\mathrm{d}^2 y}{\mathrm{d}x^2} + x\frac{\mathrm{d}y}{\mathrm{d}x} - (x^2+2)\cos x = 0$$

en $(0, 1)$, ya que para cada $x \in (0, 1)$

$$\frac{\mathrm{d}y}{\mathrm{d}x}(x) = \operatorname{sen} x + x\cos x$$

y

$$\frac{\mathrm{d}^2 y}{\mathrm{d}x^2}(x) = 2\cos x - x \operatorname{sen} x,$$

y al sustituir en la ecuación diferencial obtenemos

$$(2\cos x - x\operatorname{sen} x) + x(\operatorname{sen} x + x\cos x) - (x^2+2)\cos x = 0.$$

11.2 La ecuación $y'(x) + a(x)y(x) = f(x)$

Sean $a(x)$ y $f(x)$, funciones continuas en un intervalo (a, b). A la ecuación diferencial

$$\frac{\mathrm{d}y}{\mathrm{d}x} + a(x)y = f(x) \tag{11.2}$$

se le llama *ecuación diferencial lineal no-homogénea de primer orden.*

Para encontrar todas las soluciones de la ecuación (11.2), tomemos una antiderivada $g(x)$ de $a(x)$:

$$\frac{\mathrm{d}g}{\mathrm{d}x}(x) = a(x).$$

Multiplicando ambos lados de la ecuación por la función $\mathrm{e}^{g(x)}$, se tiene

$$\mathrm{e}^{g(x)}\left[\frac{\mathrm{d}y}{\mathrm{d}x} + a(x)y\right] = \mathrm{e}^{g(x)}f(x). \tag{11.3}$$

Observando que el lado izquierdo de (11.3) es la derivada del producto entre $\mathrm{e}^{g(x)}$ y $y(x)$, escribimos (11.3) en la forma

$$\frac{\mathrm{d}}{\mathrm{d}x}(\mathrm{e}^{g(x)}y(x)) = \mathrm{e}^{g(x)}f(x).$$

Luego, tomando $x_0 \in (a, b)$ e integrando,

$$\mathrm{e}^{g(x)}y(x) = \mathrm{e}^{g(x_0)}y(x_0) + \int_{x_0}^{x} \mathrm{e}^{g(t)}f(t)\mathrm{d}t,$$

de donde se obtiene la solución general de la ecuación (11.2), en la forma

$$y(x) = \mathrm{e}^{g(x_0)-g(x)}y(x_0) + \int_{x_0}^{x} \mathrm{e}^{g(t)-g(x)}f(t)\mathrm{d}t.$$

Ejemplo 11.2 Considere la ecuación diferencial en $\mathbb{R}$

$$\frac{\mathrm{d}y}{\mathrm{d}x} + (\cos x)y = \operatorname{sen} x \cos x.$$

Una antiderivada de $a(x) = \cos x$ es la función $g(x) = \operatorname{sen} x$. Luego, tomando $x_0 = 0$, las soluciones $y(x)$ de la ecuación son

$$y(x) = \mathrm{e}^{-\operatorname{sen} x}y(0) + \int_0^x \mathrm{e}^{\operatorname{sen} t - \operatorname{sen} x} \operatorname{sen} t \cos t\mathrm{d}t$$

e integrando por partes, se obtiene

$$y(x) = y(0)\mathrm{e}^{-\operatorname{sen} x} + \operatorname{sen} x - 1.$$

Note que el valor de la solución en $x = 0$, determina totalmente la solución. $\triangleleft$

11.3 La ecuación $y''(x) + by'(x) + ay(x) = f(x)$

Veamos ahora cómo los métodos desarrollados hasta aquí se aplican a la resolución de la *ecuación lineal no-homogénea de segundo orden con coeficientes constantes*

$$\frac{\mathrm{d}^2 y}{\mathrm{d}x^2} + b\frac{\mathrm{d}y}{\mathrm{d}x} + ay = f(x), \tag{11.4}$$

donde a, b son constantes reales y $f(x)$ es una función continua arbitraria.

El estudio de (11.4) lo iniciaremos con el caso homogéneo:

$$\frac{\mathrm{d}^2 y}{\mathrm{d}x^2} + b\frac{\mathrm{d}y}{\mathrm{d}x} + ay = 0. \tag{11.5}$$

Las soluciones de la ecuación (11.5) tienen las propiedades siguientes.

1. La función $y(x) = 0$ es solución de (11.5).

2. Si $y(x)$ es solución de (11.5), entonces la función $z(x) = y(x - x_0)$ también lo es.

3. Si $y_1(x)$ y $y_2(x)$ son dos soluciones de (11.5), entonces la combinación lineal

$$y(x) = \alpha y_1(x) + \beta y_2(x)$$

con $\alpha, \beta \in \mathbb{R}$ es también solución de (11.5).

4. Si $y(x)$ es solución de la ecuación (11.5), también lo es la función $\dfrac{\mathrm{d}y}{\mathrm{d}x}(x)$.

11.3.1 La ecuación $y''(x) - cy(x) = 0$

En este apartado, encontraremos las soluciones de la ecuación diferencial homogénea de segundo orden

$$\frac{\mathrm{d}^2 y}{\mathrm{d}x^2}(x) - cy(x) = 0, \tag{11.6}$$

donde c es una constante real.

Distinguiremos los tres casos siguientes:

1. $c = 0$.

   Las soluciones de la ecuación $\dfrac{\mathrm{d}^2 y}{dx^2}(x) = 0$, se obtienen directamente, y son de la forma

$$y(x) = \alpha x + \beta$$

   para cualquier par $\alpha, \beta \in \mathbb{R}$.

2. $c > 0$.

   Las soluciones de la ecuación $\dfrac{\mathrm{d}^2 y}{\mathrm{d}x^2}(x) = cy(x)$ se pueden encontrar ensayando con funciones de la forma $y(x) = \mathrm{e}^{rx}$ y observando que, al sustituirlas en la ecuación, se tiene la relación

$$r^2 \mathrm{e}^{rx} = c\mathrm{e}^{rx},$$

   la cual se satisfará si el número r tiene alguno de los valores siguientes

$$r_1 = \sqrt{c}, \qquad r_2 = -\sqrt{c}.$$

   Por lo tanto, las funciones $y_1(x) = \mathrm{e}^{\sqrt{c}x}$ y $y_2(x) = \mathrm{e}^{-\sqrt{c}x}$ son dos soluciones particulares y entonces cualquier combinación lineal de esas dos soluciones,

$$y(x) = \alpha \mathrm{e}^{\sqrt{c}x} + \beta \mathrm{e}^{-\sqrt{c}x}$$

   con α y β constantes arbitrarias, también es solución.

3. $c < 0$.

   Ahora la ecuación se escribe en la forma

$$\frac{\mathrm{d}^2 y}{\mathrm{d}x^2}(x) = -|c|y(x)$$

   y observamos que las funciones $y_1(x) = \operatorname{sen}\sqrt{|c|}x$ y $y_2(x) = \cos\sqrt{|c|}x$ son dos soluciones particulares y las combinaciones lineales

$$y(x) = \alpha \operatorname{sen}(\sqrt{|c|}x) + \beta \cos(\sqrt{|c|}x),$$

   con α y β constantes arbitrarias, forman una familia de soluciones.

Enseguida mostraremos que las familias de soluciones que hemos encontrado para los distintos casos de la ecuación (11.6) nos dan todas sus soluciones posibles. Para probar esta afirmación, demostraremos primero que si una solución $y(x)$ de la ecuación (11.6) es tal que en el punto x_0 se anula su valor y el de su derivada, es decir $y(x_0) = 0$ y $\dfrac{\mathrm{d}y}{\mathrm{d}x}(x_0) = 0$, entonces $y(x) = 0$ para todo $x \in \mathbb{R}$.

Lema 11.1 *Sea $y(x)$ una función definida en $\mathbb{R}$ tal que $\dfrac{\mathrm{d}^2 y}{\mathrm{d}x^2}(x) = cy(x)$ con $c \in \mathbb{R}$ y además $y(x_0) = 0$ y $\dfrac{\mathrm{d}y}{\mathrm{d}x}(x_0) = 0$. Entonces $y(x) = 0$ para todo $x \in \mathbb{R}$.*

Demostración. Sea $x \in [x_0, x_0 + a)$ con $0 < a < 1$ y $0 < a|c| < 1$. Por el teorema del valor medio, podemos escribir

$$y(x) - y(x_0) = \frac{\mathrm{d}y}{\mathrm{d}x}(x_1)(x - x_0) \tag{11.7}$$

donde $x_1 \in (x_0, x)$. Análogamente, aplicando el teorema del valor medio a la función $\frac{\mathrm{d}y}{\mathrm{d}x}(x)$ en el intervalo $[x_0, x_1]$, podemos escribir

$$\frac{\mathrm{d}y}{\mathrm{d}x}(x_1) - \frac{\mathrm{d}y}{\mathrm{d}x}(x_0) = \frac{\mathrm{d}^2y}{\mathrm{d}x^2}(x_2)(x_1 - x_0) \tag{11.8}$$

donde $x_2 \in (x_0, x_1)$. Combinando (11.7) y (11.8) se tiene

$$y(x) = \frac{\mathrm{d}^2y}{\mathrm{d}x^2}(x_2)(x_1 - x_0)(x - x_0), \tag{11.9}$$

y considerando que $\frac{\mathrm{d}^2y}{\mathrm{d}x^2}(x) = cy(x)$, se tiene que

$$y(x) = cy(x_2)(x_1 - x_0)(x - x_0). \tag{11.10}$$

Repitiendo ahora el argumento para el punto x_2, obtendremos

$$y(x_2) = cy(x_4)(x_3 - x_0)(x_2 - x_0) \tag{11.11}$$

con $x_3, x_4 \in [x_0, x_2)$ y, sustituyendo (11.11) en (11.10) tenemos

$$y(x) = c^2y(x_4)(x_3 - x_0)(x_2 - x_0)(x_1 - x_0)(x - x_0).$$

Repitiendo el argumento anterior k veces y tomando valor absoluto, arribaremos a la estimación

$$|y(x)| \leq Mc^k a^{2k}$$

donde M es cota de $y(x)$ en $[x_0, x_0 + a]$. Tomando límite cuando $k \to \infty$ se tiene que $y(x) = 0$ para $x \in [x_0, x_0 + a)$. Se sigue, por continuidad, que $y(x_0 + a) = 0$ y $\frac{\mathrm{d}y}{\mathrm{d}x}(x_0+a) = 0$. Podemos repetir el argumento anterior en el intervalo $[x_0+a, x_0+2a)$ y concluir que el intervalo donde se anula la función $y(x)$ se extiende indefinidamente a la derecha de x_0. Por un argumento similar se extiende también indefinidamente a la izquierda de x_0, con lo cual se prueba que $y(x) = 0$ para toda $x \in \mathbb{R}$. ■

A partir del lema 11.1, se sigue que si dos soluciones de la ecuación (11.6) y sus derivadas toman el mismo valor en un punto, entonces coinciden en todos los puntos de $\mathbb{R}$, ya que su diferencia, al ser también solución y anularse junto con su derivada en un punto, se anulará en todo $\mathbb{R}$, lo cual significa que ambas soluciones coincidirán y serán entonces la misma. A este resultado se le conoce como el *teorema de unicidad de la solución* para el problema de condición inicial.

Tomando en cuenta el lema 11.1 y sus consecuencias sobre la unicidad de la solución, observamos que si se establece de antemano en un punto inicial x_0 el valor de una solución a la ecuación (11.6) y el de su derivada y se encuentra alguna solución que tome los mismos valores en x_0, entonces esa será la única solución con esas propiedades. Como las familias de soluciones que hemos encontrado para los distintos casos de la ecuación (11.6) dependen de dos coeficientes o parámetros, entonces, dados dos valores prescritos para la solución y su derivada en un punto, siempre podremos encontrar dentro de esas familias, una función con esos valores. Comprobemos esto con algunos ejemplos.

Ejemplo 11.3 La solución de la ecuación $\dfrac{\mathrm{d}^2y}{\mathrm{d}x^2}(x) = y(x)$ tal que $y(1) = 2$ y $\dfrac{\mathrm{d}y}{\mathrm{d}x}(1) = -1$, se encuentra buscando en la familia de soluciones $y(x) = \alpha\mathrm{e}^x + \beta\mathrm{e}^{-x}$ aquélla que satisfaga tales condiciones. Para ello, α y β deben satisfacer

$$\begin{aligned} \alpha\mathrm{e} + \beta\mathrm{e}^{-1} &= 2 \\ \alpha\mathrm{e} - \beta\mathrm{e}^{-1} &= -1, \end{aligned}$$

es decir, $\alpha = \frac{1}{2}\mathrm{e}^{-1}$ y $\beta = \frac{3}{2}\mathrm{e}$. Luego, la solución buscada es

$$y(x) = \frac{1}{2}\mathrm{e}^{x-1} + \frac{3}{2}\mathrm{e}^{1-x}. \qquad \triangleleft$$

Ejemplo 11.4 La solución de la ecuación $\dfrac{\mathrm{d}^2y}{\mathrm{d}x^2}(x) = -4y(x)$ tal que $y(0) = 1$ y $\dfrac{\mathrm{d}y}{\mathrm{d}x}(0) = -1$, se encuentra buscando en la familia de soluciones $y(x) = \alpha\cos 2x + \beta\operatorname{sen} 2x$, aquélla que satisfaga esas condiciones; es decir, α y β deben satisfacer

$$\alpha = 1, \qquad \beta = -\frac{1}{2}.$$

Entonces, la solución es la función $y(x) = \cos 2x - \frac{1}{2}\operatorname{sen} 2x$. $\triangleleft$

Ahora veamos que cualquier ecuación homogénea de segundo orden con coeficientes constantes es equivalente a alguna de las anteriores, en el sentido de que un cambio de variable la reduce a uno de estos casos. En concreto, si $y(t)$ es solución de la ecuación diferencial (11.5) y sustituimos la función

$$z(x) = y(x)\mathrm{e}^{\frac{b}{2}x}$$

en la ecuación (11.5), tendremos

$$\frac{\mathrm{d}^2}{\mathrm{d}x^2}\big(\mathrm{e}^{-\frac{b}{2}x}z(x)\big) + b\frac{\mathrm{d}}{\mathrm{d}x}\big(\mathrm{e}^{-\frac{b}{2}x}z(x)\big) + a\mathrm{e}^{-\frac{b}{2}x}z(x) = 0.$$

De aquí se obtiene que la ecuación diferencial para $z(x)$ es

$$\frac{\mathrm{d}^2z}{\mathrm{d}x^2} = \left(\frac{b^2}{4} - a\right)z,$$

que es una ecuación de la forma (11.6) con

$$c = \frac{b^2}{4} - a.$$

Ejemplo 11.5 Para la ecuación $\dfrac{\mathrm{d}^2y}{\mathrm{d}x}+2\dfrac{\mathrm{d}y}{\mathrm{d}x}+3y=0$ tal que $y(2)=0$ y $\dfrac{\mathrm{d}y}{\mathrm{d}x}(2)=1$, se tiene que $c=\dfrac{b^2}{4}-a=-2$. Para encontrar la solución resolvemos primero la ecuación

$$\frac{\mathrm{d}^2z}{\mathrm{d}x^2}(x)=-2z(x),$$

cuyas soluciones son de la forma

$$z(x)=\alpha\cos\sqrt{2}x+\beta\operatorname{sen}\sqrt{2}x$$

y dan lugar a la familia de soluciones de la ecuación inicial, que es de la forma

$$y(x)=\alpha\mathrm{e}^{-x}\cos\sqrt{2}x+\beta\mathrm{e}^{-x}\operatorname{sen}\sqrt{2}x.$$

Para calcular los valores de α y β, usamos las condiciones iniciales:

$$\begin{aligned} 0 &= y(2)=\alpha\mathrm{e}^{-2}\cos 2\sqrt{2}+\beta\mathrm{e}^{-2}\operatorname{sen}2\sqrt{2} \\ 1 &= \frac{\mathrm{d}y}{\mathrm{d}x}(2)=-\mathrm{e}^{-2}(\alpha-\beta\sqrt{2})\cos 2\sqrt{2}-\mathrm{e}^{-2}(\beta+\alpha\sqrt{2})\operatorname{sen}2\sqrt{2}), \end{aligned}$$

de donde

$$\alpha=-\frac{\mathrm{e}^2}{\sqrt{2}}\operatorname{sen}2\sqrt{2},\qquad \beta=\frac{\mathrm{e}^2}{\sqrt{2}}\cos 2\sqrt{2}.$$

Así que la solución es

$$y(x)=\frac{\mathrm{e}^{2-x}}{\sqrt{2}}\operatorname{sen}(\sqrt{2}(x-2)).$$ ◁

Resumimos la discusión anterior en el teorema siguiente.

Teorema 11.2 *La soluciones $y(x)$ de la ecuación diferencial lineal de segundo orden homogénea con coeficientes constantes $\frac{\mathrm{d}^2y}{\mathrm{d}x^2}+b\frac{\mathrm{d}y}{\mathrm{d}x}+ay=0$ están definidas para $x\in\mathbb{R}$ y son:*

$$y(x)=\begin{cases} \mathrm{e}^{-\frac{b}{2}x}\left[\alpha\mathrm{e}^{\frac{\sqrt{b^2-4a}}{2}x}+\beta\mathrm{e}^{\frac{-\sqrt{b^2-4a}}{2}x}\right] & \text{si } b^2-4a>0,\\ \mathrm{e}^{-\frac{b}{2}x}\big(\alpha+\beta x\big) & \text{si } b^2-4a=0,\\ \mathrm{e}^{-\frac{b}{2}x}\left[\alpha\cos\left(\sqrt{a-\frac{b^2}{4}}x\right)+\beta\operatorname{sen}\left(\sqrt{a-\frac{b^2}{4}}x\right)\right] & \text{si } b^2-4a<0, \end{cases}$$

con $\alpha,\beta\in\mathbb{R}$.

NOTA IMPORTANTE:

1. *Si $y_1(x)$ y $y_2(x)$ son soluciones de la ecuación homogénea 11.5, a la función*

$$W_{y_1,y_2}=y_1(x)\frac{\mathrm{d}y_2}{\mathrm{d}x}(x)-y_2(x)\frac{\mathrm{d}y_1}{\mathrm{d}x}(x)$$

se le llama wronskiano[1] *de* y_1 *y* y_2, *y tiene la propiedad de que*

$$\frac{\mathrm{d}W}{\mathrm{d}x}(0) = -bW,$$

es decir, $W(x) = W(0)\mathrm{e}^{-bx}$, *de donde se deduce que si* $W(0) \neq 0$ *entonces* $W(x) \neq 0$ *para toda* $x \in \mathbb{R}$.

2. *Si* $y_3(x)$ *y* $y_4(x)$ *es otro par de soluciones de 11.5, entonces* $W_{y_3,y_4} = cW_{y_1,y_2}$ *para alguna* $c \in \mathbb{R}$.

11.3.2 Método de variación de constantes

Para encontrar todas las soluciones de la ecuación de segundo orden no-homogénea

$$\frac{\mathrm{d}^2y}{\mathrm{d}x^2} + b\frac{\mathrm{d}y}{\mathrm{d}x} + ay = f(x), \tag{11.12}$$

hagamos primero la observación siguiente: Si $y_1(x)$ y $y_2(x)$ son dos soluciones de (11.12) entonces su diferencia $y_1(x) - y_2(x)$ es solución de la ecuación homogénea (11.5). Tomando en cuenta esto, bastará conocer una solución particular $y_p(x)$ de la ecuación no-homogénea para conocer todas sus soluciones, las cuales serán de la forma

$$y(x) = y_p(x) + \text{solución general de la ecuación homogénea.}$$

Esta observación reduce el problema de encontrar todas las soluciones de la ecuación (11.12), al cálculo de sólo una solución particular de dicha ecuación. Para encontrar una solución particular de (11.12) consideraremos el *método de variación de constantes*[2] que consiste en construir una solución particular para la ecuación no-homogénea (11.12) a partir de las soluciones de la ecuación homogénea (11.5). Si denotamos por $y_1(x)$ y $y_2(x)$ un par de soluciones de la ecuación lineal homogénea cuyas combinaciones lineales generan todas las soluciones de esa ecuación y ahora buscamos una solución particular de la ecuación no-homogénea (11.12) de la forma

$$y_p(x) = z_1(x)y_1(x) + z_2(x)y_2(x) = \sum_{i=1}^{2} z_i(x)y_i(x) \tag{11.13}$$

donde $z_1(x)$ y $z_2(x)$ son dos funciones a determinar, al sustituir (11.13) en (11.12), tendremos que

$$f(x) = \frac{\mathrm{d}^2y_p}{\mathrm{d}x^2}(x) + b\frac{\mathrm{d}y_p}{\mathrm{d}x}(x) + ay_p,$$

lo cual da lugar, para las funciones $z_1(x)$ y $z_2(x)$, a la expresión

$$f(x) = \sum_{i=1}^{2}\left(\frac{\mathrm{d}^2}{\mathrm{d}x^2}(z_i(x)y_i(x)) + b\frac{\mathrm{d}}{\mathrm{d}x}z_i(x)y_i(x) + az_i(x)y_i(x)\right). \tag{11.14}$$

[1] Por J. M. Höené de Wronski (1778-1853), matemático polaco de ascendencia checa.

[2] A este método también se le llama *método de variación de parámetros* o *método de Lagrange*, por Joseph-Louis Lagrange.

Al tomar en cuenta que

$$\frac{\mathrm{d}^2 y_i}{\mathrm{d}x^2}(x) + b\frac{\mathrm{d}y_i}{\mathrm{d}x}(x) + ay_i(x) = 0$$

para cada $i = 1, 2$, la ecuación (11.14) se simplifica en la forma

$$\sum_{i=1}^{2}\left(\frac{\mathrm{d}^2 z_i}{\mathrm{d}x^2}(x)y_i(x) + b\frac{\mathrm{d}z_i}{\mathrm{d}x}(x)y_i(x) + 2\frac{\mathrm{d}z_i}{\mathrm{d}x}(x)\frac{\mathrm{d}y_i}{\mathrm{d}x}(x)\right) = f(x). \tag{11.15}$$

Como $z_1(x)$ y $z_2(x)$ son dos funciones a determinar, podemos imponer la condición de que

$$\frac{\mathrm{d}z_1}{\mathrm{d}x}(x)y_1(x) + \frac{\mathrm{d}z_2}{\mathrm{d}x}(x)y_2(x) = 0, \tag{11.16}$$

en cuyo caso se tendrá también que

$$\sum_{i=1}^{2}\left(\frac{\mathrm{d}^2 z_i}{\mathrm{d}x^2}(x)y_i(x) + \frac{\mathrm{d}z_i}{\mathrm{d}x}(x)\frac{\mathrm{d}y_i}{\mathrm{d}x}(x)\right) = 0$$

y la expresión (11.15) se reduce a

$$\frac{\mathrm{d}z_1}{\mathrm{d}x}(x)\frac{\mathrm{d}y_1}{\mathrm{d}x}(x) + \frac{\mathrm{d}z_2}{\mathrm{d}x}(x)\frac{\mathrm{d}y_2}{\mathrm{d}x}(x) = f(x),$$

que, junto con la ecuación (11.16), constituye, para cada valor $x \in \mathbb{R}$, un sistema lineal no-homogéneo de dos ecuaciones algebraicas para las incógnitas $\frac{dz_1}{\mathrm{d}x}(x)$ y $\frac{\mathrm{d}z_2}{\mathrm{d}x}(x)$ y cuyas soluciones son

$$\frac{\mathrm{d}z_1}{\mathrm{d}x}(x) = \frac{-f(x)y_2(x)}{y_1(x)\frac{\mathrm{d}y_2}{\mathrm{d}x}(x) - y_2(x)\frac{\mathrm{d}y_1}{\mathrm{d}x}(x)}$$

y

$$\frac{\mathrm{d}z_2}{\mathrm{d}x}(x) = \frac{f(x)y_1(x)}{y_1(x)\frac{\mathrm{d}y_2}{\mathrm{d}x}(x) - y_2(x)\frac{\mathrm{d}y_1}{\mathrm{d}x}(x)}.$$

El denominador $y_1(x)\frac{\mathrm{d}y_2}{\mathrm{d}x}(x) - y_2(x)\frac{\mathrm{d}y_1}{dx}(x)$ es el wronskiano W_{y_1,y_2} y es distinto de cero para toda $x \in \mathbb{R}$. Integrando las expresiones para $\frac{\mathrm{d}z_1}{\mathrm{d}x}(x)$ y $\frac{\mathrm{d}z_2}{\mathrm{d}x}(x)$, se tiene

$$z_1(x) = \int_{x_0}^{x} \frac{-f(t)y_2(t)}{y_1(t)\frac{\mathrm{d}y_2}{\mathrm{d}x}(t) - y_2(t)\frac{\mathrm{d}y_1}{\mathrm{d}x}(t)}\mathrm{d}t = \int_{x_0}^{x} \frac{-f(t)y_2(t)}{W_{y_1,y_2}}\mathrm{d}t,$$

$$z_2(x) = \int_{x_0}^{x} \frac{f(t)y_1(t)}{y_1(t)\frac{\mathrm{d}y_2}{\mathrm{d}x}(t) - y_2(t)\frac{\mathrm{d}y_1}{\mathrm{d}x}(t)}\mathrm{d}t = \int_{x_0}^{x} \frac{f(t)y_1(t)}{W_{y_1,y_2}}\mathrm{d}t.$$

Para resumir estos cálculos, enunciamos el teorema siguiente.

Teorema 11.3 *Si $y_1(y), y_2(x)$ son soluciones de la ecuación diferencial (11.12) y $W_{y_1,y_2}(x_0) \neq 0$ entonces la función*

$$y_p = \left[\int_{x_0}^{x} \frac{-f(t)y_2(t)}{W_{y_1,y_2}} \mathrm{d}t \right] y_1(x) + \left[\int_{x_0}^{x} \frac{f(t)y_1(t)}{W_{y_1,y_2}} \mathrm{d}t \right] y_2(x)$$

es una solución particular de la ecuación no-homogénea (11.12).

Ejemplo 11.6 Encontremos todas las soluciones de la ecuación

$$\frac{\mathrm{d}^2 y}{\mathrm{d}x^2} + b\frac{\mathrm{d}y}{\mathrm{d}x} + ay = f(x), \tag{11.17}$$

cuando $b^2 - 4a > 0$. En este caso, las soluciones de la ecuación homogénea son de la forma $\alpha \mathrm{e}^{r_1 x} + \beta \mathrm{e}^{r_2 x}$ con

$$\begin{aligned} r_1 &= \frac{-b + \sqrt{b^2 - 4a}}{2}, \\ r_2 &= \frac{-b - \sqrt{b^2 - 4a}}{2}, \end{aligned}$$

y una solución particular $y_p(x)$ de la ecuación no-homogénea es

$$y_p(x) = z_1(x)\mathrm{e}^{r_1 x} + z_2(x)\mathrm{e}^{r_2 x}$$

con

$$\begin{aligned} z_1(x) &= \frac{1}{\sqrt{b^2 - 4a}} \int_{x_0}^{x} f(t)\mathrm{e}^{-r_1 t} dt, \\ z_2(x) &= \frac{-1}{\sqrt{b^2 - 4a}} \int_{x_0}^{x} f(t)\mathrm{e}^{-r_2 t} dt. \end{aligned}$$

Con base en lo anterior enunciamos el teorema siguiente.

Teorema 11.4 *Las soluciones $y(x)$ de la ecuación diferencial de segundo orden con coeficientes constantes y no-homogénea (11.17) toman, en los distintos casos, la forma siguiente:*

1. Si $b^2 - 4a > 0$,
$$y(x) = \alpha \mathrm{e}^{r_1 x} + \beta \mathrm{e}^{r_2 x} + \frac{1}{\sqrt{b^2 - 4a}} \int_{x_0}^{x} f(t)[\mathrm{e}^{r_1(x-t)} - \mathrm{e}^{r_2(x-t)}]\mathrm{d}t,$$
donde $\alpha, \beta \in \mathbb{R}$ y
$$r_1 = \frac{-b + \sqrt{b^2 - 4a}}{2}, \qquad r_2 = \frac{-b - \sqrt{b^2 - 4a}}{2}.$$

2. Si $b^2 - 4a = 0$,
$$y(x) = \mathrm{e}^{\frac{-b}{2}x}\big[\alpha + \beta x\big] - \int_{x_0}^{x} f(t)(t - x)\mathrm{e}^{\frac{b}{2}(t-x)}\mathrm{d}t,$$
donde $\alpha, \beta \in \mathbb{R}$.

3. Si $b^2 - 4a < 0$,
$$y(x) = \mathrm{e}^{\frac{-b}{2}x}\big[\alpha \cos qx + \beta \operatorname{sen} qx\big] + \frac{1}{q} \int_{x_0}^{x} f(t)\mathrm{e}^{\frac{b}{2}(t-x)} \operatorname{sen}(q(x-t))\mathrm{d}t,$$
donde $q = \dfrac{1}{2}\sqrt{|b^2 - 4a|}$.

Ejemplo 11.7 Resuelva la ecuación diferencial
$$\frac{\mathrm{d}^2y}{\mathrm{d}x^2} - \frac{\mathrm{d}y}{\mathrm{d}x} - 2y = \mathrm{e}^{-x}.$$
La ecuación homogénea tiene la forma $\dfrac{\mathrm{d}^2y}{\mathrm{d}x^2} - \dfrac{\mathrm{d}y}{\mathrm{d}x} - 2y = 0$. Dado que $(b^2 - 4a) = 9 > 0$, las soluciones $w(x)$ de la ecuación homogénea son $w(x) = \alpha\mathrm{e}^{2x} + \beta\mathrm{e}^{-x}$. Calculemos ahora la solución particular de la forma
$$y_p(x) = z_1(x)\mathrm{e}^{2x} + z_2(x)\mathrm{e}^{-x}.$$
Las derivadas de las funciones $z_1(x)$ y $z_2(x)$ satisfacen el sistema
$$\begin{aligned} \frac{\mathrm{d}z_1}{\mathrm{d}x}(x)\mathrm{e}^{2x} + \frac{\mathrm{d}z_2}{\mathrm{d}x}(x)\mathrm{e}^{-x} &= 0 \\ 2\frac{\mathrm{d}z_1}{\mathrm{d}x}(x)\mathrm{e}^{2x} - \frac{\mathrm{d}z_2}{\mathrm{d}x}(x)\mathrm{e}^{-x} &= \mathrm{e}^{-x}, \end{aligned}$$
de donde se obtiene
$$\frac{\mathrm{d}z_1}{\mathrm{d}x}(x) = \frac{1}{3}\mathrm{e}^{-3x}, \qquad \frac{\mathrm{d}z_2}{\mathrm{d}x}(x) = -\frac{1}{3}.$$

Entonces

$$z_1(x) = -\frac{1}{9}\mathrm{e}^{-3x} \qquad z_2(x) = -\frac{1}{3}x.$$

La solución particular $y_p(x)$ es

$$y_p(x) = -\frac{1}{9}\mathrm{e}^{-x} - \frac{1}{3}x\mathrm{e}^{-x}.$$

Finalmente, las soluciones $y(x)$ de la ecuación son

$$y(x) = \alpha\mathrm{e}^{2x} + \beta\mathrm{e}^{-x} - \frac{1}{3}x\mathrm{e}^{-x}. \qquad \triangleleft$$

11.4 Leyes de movimiento de Newton

Entre las más importantes aplicaciones del cálculo diferencial está la descripción matemática del movimiento de los cuerpos materiales sometidos a la acción de fuerzas externas. El mismo concepto de fuerza pudo ser definido identificando su acción sobre los cuerpos físicos en cada instante como proporcional a la variación que experimenta su velocidad con respecto al tiempo (aceleración) como consecuencia de la presencia de dichas fuerzas.

En términos generales, se supone que el movimiento de las partículas tiene lugar en el marco de un sistema de referencia fijo, con respecto al cual se realizan las mediciones de las posiciones de los cuerpos o partículas, relativas al fluir del tiempo, que se considera independiente de todo observador.

Isaac Newton formuló las leyes que gobiernan el movimiento físico de los cuerpos en los términos siguientes:

Primera (*ley de inercia*): Los cuerpos mantienen su estado de reposo o de velocidad constante en ausencia de fuerzas externas.

Segunda (*ley de movimiento*): Si $y(t)$ representa la función de posición de un cuerpo durante un intervalo de tiempo (a, b), entonces la fuerza $F(t, y(t))$ que experimenta en el tiempo t hallándose en la posición $y(t)$ es proporcional a la segunda derivada $\frac{\mathrm{d}^2y}{\mathrm{d}t^2}(t)$ de la función posición en el tiempo t, es decir,

$$F(t, y(t)) = m\frac{\mathrm{d}^2y}{\mathrm{d}t^2}(t).$$

La constante de proporcionalidad m se conoce como la *masa* del cuerpo y depende de las propiedades de la materia que forma ese cuerpo.

1. **Caída bajo la acción de la gravedad**. Un cuerpo situado en una vecindad de la superficie de la Tierra, experimenta una fuerza de magnitud constante W igual a su peso y dirigida perpendicularmente hacia el suelo. A tal fuerza se le denomina *fuerza de gravedad.*

Si $y(t)$ denota la posición de un cuerpo en el tiempo t medida sobre la vertical, la segunda ley de movimiento de Newton se expresa mediante la ecuación diferencial

$$m\frac{\mathrm{d}^2y}{\mathrm{d}t^2}(t) = -W,$$

donde m es la masa del cuerpo y $\dfrac{W}{m} = g = 9.8$ m/seg^2. Esto último significa que la aceleración de los cuerpos en caída libre es constante e igual a $-g$. A la constante g se le llama *aceleración debida a la fuerza de gravedad.*

Resolviendo la ecuación diferencial, se tiene que la función de posición $y(t)$ es

$$y(t) = -\frac{1}{2}gt^2 + \frac{\mathrm{d}y}{\mathrm{d}t}(0)t + y(0),$$

donde $\dfrac{\mathrm{d}y}{\mathrm{d}t}(0)$ y $y(0)$ son la velocidad y la posición del cuerpo en el tiempo $t = 0$ sobre la vertical.

Ejemplo 11.8 Si un cuerpo se arroja desde una altura de 100 metros con una velocidad inicial hacia arriba de 20 m/seg, calcule el tiempo y la velocidad con que golpeará el suelo.

Solución:

De acuerdo a la segunda ley de movimiento de Newton, si denotamos por $y(t)$ la posición del cuerpo en el tiempo t, tendremos

$$y(t) = -4.9t^2 + 20t + 100$$

y entonces el cuerpo golpeará el suelo en el tiempo t_0 tal que $y(t_0) = 0$, es decir, cuando

$$-4.9t_0^2 + 20t_0 + 100 = 0,$$

o

$$t_0 = \frac{20}{9.8}(1 + \sqrt{5.9}).$$

La velocidad que llevará el cuerpo al chocar será de

$$\frac{\mathrm{d}y}{\mathrm{d}t}(t_0) = -\sqrt{5.9}.$$

El signo negativo significa que la velocidad es en dirección contraria al sentido positivo en que se miden las alturas sobre la Tierra.

2. **Caída bajo la acción de la gravedad con fricción del aire.** Si se toma en cuenta el efecto de la presencia del aire sobre la caída de un cuerpo, experimentalmente se ha observado que el aire opone una fuerza de resistencia proporcional a la velocidad que lleva el cuerpo y en sentido contrario a esa

velocidad. Esta nueva fuerza, llamada fuerza de fricción, modifica la ley de caída libre, de tal manera que la ecuación de movimiento toma ahora la forma

$$\frac{\mathrm{d}^2 y}{\mathrm{d}t^2}(t) + \frac{k}{m}\frac{\mathrm{d}y}{\mathrm{d}t} = -g, \tag{11.18}$$

donde $k > 0$ es una constante.

Aplicando el teorema 11.4, observamos que las soluciones de la ecuación diferencial (11.18) son de la forma

$$y(t) = \alpha \mathrm{e}^{-\frac{k}{m}t} - \frac{gm}{k}t + \beta, \tag{11.19}$$

donde las constantes α, β están determinadas por la posición y la velocidad del cuerpo en el tiempo $t = 0$,

$$\begin{aligned} \alpha &= -\left(\frac{\mathrm{d}y}{\mathrm{d}t}(0) + \frac{gm}{k}\right)\frac{m}{k}, \\ \beta &= y(0) + \left(\frac{\mathrm{d}y}{\mathrm{d}t}(0) + \frac{gm}{k}\right)\frac{m}{k}. \end{aligned}$$

Note que la acción de la fuerza de fricción hace que, a la larga, la velocidad de caída sea constante e igual $-\dfrac{gm}{k}$.

3. **Movimiento bajo la fuerza de un resorte**. Consideramos, sobre una mesa lisa, sin fricción, un resorte que tiene fijo uno de sus extremos. Sujetemos un cuerpo de masa m al extremo libre del resorte como se muestra en la figura 11.1 Experimentalmente se ha determinado que la fuerza que ejerce el resorte

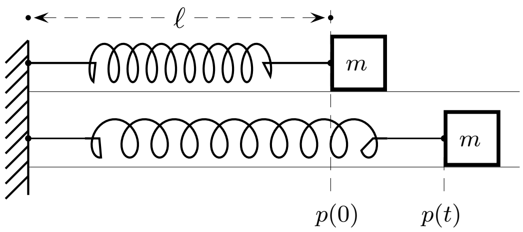


Figura 11.1

sobre el cuerpo en un tiempo t, es proporcional y en sentido contrario a la deformación (estiramiento o contracción) que muestra el resorte con respecto a su longitud normal. La constante de proporcionalidad, llamada *constante de restitución,* es un número positivo k que sólo depende de las características materiales del resorte y no cambia con respecto al tiempo.

El problema del movimiento bajo la acción del resorte consiste en determinar para un cuerpo, que supondremos de masa m, su posición como función del tiempo, conocidas la posición inicial $p(0)$ y la velocidad inicial $\dfrac{\mathrm{d}p}{\mathrm{d}t}(0)$ de ese cuerpo.

En virtud de la segunda ley de movimiento de Newton, si la posición $p(t)$ del cuerpo es medida desde el punto de cero deformación del resorte, tal como se muestra en la figura 11.1, la función $p(t)$ debe satisfacer, en cada tiempo, la relación siguiente:

$$\frac{\mathrm{d}^2 p}{\mathrm{d}t^2}(t) = -\omega^2 p(t), \tag{11.20}$$

donde

$$\omega^2 = \frac{k}{m}.$$

Las soluciones de la ecuación (11.20), de acuerdo al teorema 11.4, tienen la forma

$$p(t) = a \operatorname{sen} \omega t + b \cos \omega t,$$

donde

$$a = \frac{1}{\omega}\frac{\mathrm{d}p}{\mathrm{d}t}(0), \qquad b = p(0).$$

La fórmula para la posición se puede escribir en la forma

$$p(t) = \frac{1}{\omega}\sqrt{\left(\frac{\mathrm{d}p}{\mathrm{d}t}(0)\right)^2 + \omega^2 p^2(0)}\,\operatorname{sen}(\omega t + \phi),$$

con

$$\phi = \operatorname{arcsen}\frac{p(0)}{\sqrt{\left(\frac{\mathrm{d}p}{\mathrm{d}t}(0)\right)^2 + \omega^2 p^2(0)}}.$$

A la constante

$$A \stackrel{\text{def}}{=} \frac{1}{\omega}\sqrt{\left(\frac{\mathrm{d}p}{\mathrm{d}t}(0)\right)^2 + \omega^2 p^2(0)}$$

se le llama *amplitud del movimiento* y a ϕ, *fase.* Observe que bajo la fuerza del resorte, el cuerpo describe un movimiento periódico con periodo $T = \dfrac{2\pi}{\omega}$ y amplitud A. En la figura 11.2 se muestra la gráfica de la función $p(t)$.

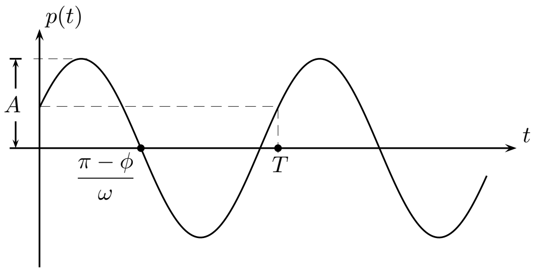


Figura 11.2 La función posición $p(t)$

4. **Movimiento con fricción bajo la fuerza de un resorte y ante la presencia de una fuerza externa.** Consideremos ahora el movimiento de un cuerpo sometido a la fuerza de un resorte, en un medio que ofrece una fuerza de fricción $F_f(t)$, proporcional a la velocidad del cuerpo

$$F_f(t) = -\rho \frac{\mathrm{d}p}{\mathrm{d}t}(t),$$

y ante la presencia de una fuerza externa $f(t)$ dependiente del tiempo. Supondremos, además, que $\rho^2 < 4\omega^2$, ya que en otro caso el movimiento del cuerpo no es oscilatorio.

Con las condiciones anteriores, la ecuación de movimiento toma la forma

$$\frac{\mathrm{d}^2 p}{\mathrm{d}t^2}(t) = -\omega^2 p(t) - \rho \frac{\mathrm{d}p}{\mathrm{d}t}(t) + f(t),$$

donde $p(t)$ representa la posición del cuerpo medida desde la posición de cero deformación del resorte. Aplicando el teorema 11.4, tenemos

$$p(t) = \alpha \mathrm{e}^{\frac{-\rho}{2}t} \cos qt + \beta \mathrm{e}^{\frac{-\rho}{2}t} \operatorname{sen} qt + \frac{1}{q}\int_0^t f(s)\mathrm{e}^{\frac{-\rho}{2}(s-t)} \operatorname{sen}(q(t-s))ds,$$

donde

$$q = \frac{1}{2}\sqrt{\rho^2 - 4\omega^2}.$$

Como casos particulares importantes, se presentan los siguientes.

(a) $\rho = 0$ y $f\left(t + \frac{2\pi}{\omega}\right) = f(t)$, es decir, el movimiento es sin fricción y la fuerza externa es una fuerza periódica con la misma frecuencia que la de las soluciones del resorte libre. En esta situación la función de posición es de la forma

$$p(t) = \alpha \cos \omega t + \beta \operatorname{sen} \omega t + \frac{1}{\omega}\int_0^t f(s) \operatorname{sen}(\omega(t-s))ds,$$

y si consideramos que la fuerza externa es periódica,

$$f(s) = A \cos \omega s,$$

entonces todas las soluciones se escriben

$$p(t) = \alpha \cos \omega t + \beta \operatorname{sen} \omega t + \frac{A}{\omega}\left[\frac{t}{2} \operatorname{sen} \omega t \;+ \frac{1}{2\omega} \cos \omega t\right].$$

En particular, la solución con $p(0) = 0$ y $\frac{\mathrm{d}p}{\mathrm{d}t}(0) = 1$, es

$$p(t) = \frac{1}{\omega}\left(\frac{At}{2} + 1\right) \operatorname{sen} \omega t. \tag{11.21}$$

Note que cuando el tiempo crece, el desplazamiento del cuerpo crece sin límite y el resorte terminará por romperse. En la figura 11.3 se muestra este comportamiento.

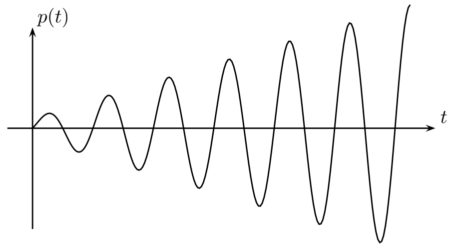


Figura 11.3 Gráfica de la función (11.21)

(b) $f(t) = 0$, es decir no existe fuerza externa y a la fuerza del resorte se suma la fuerza de fricción. En este caso, la ecuación de movimiento es

$$\frac{\mathrm{d}^2 p}{\mathrm{d}t^2}(t) = -\omega^2 p(t) - \rho \frac{\mathrm{d}p}{\mathrm{d}t}(t),$$

cuyas soluciones son de la forma

$$p(t) = \mathrm{e}^{\frac{-\rho}{2}t}\left[\alpha \cos qt + \beta \operatorname{sen} qt\right], \tag{11.22}$$

donde

$$q = \frac{1}{2}\sqrt{\rho^2 - 4\omega^2}.$$

Note que en este último caso, el cuerpo oscila con una amplitud que decrece exponencialmente y, así, cuando el tiempo crece, el cuerpo tiende a la posición $p = 0$. En la figura 11.4 se muestra la gráfica de la función (11.22).

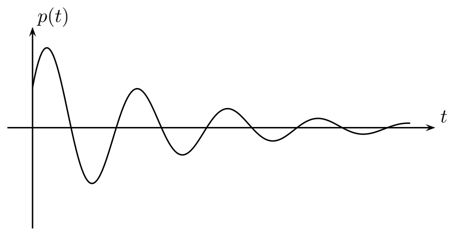


Figura 11.4 Gráfica de la función (11.22)

Ejercicios y problemas del capítulo

1. Pruebe que si $y(x)$ es una solución de la ecuación

$$\frac{\mathrm{d}y}{\mathrm{d}x}(x) = xy(x) + 1, \qquad x \in (0,1),$$

entonces la función $z(x) = y^2(x)$ es solución de la ecuación

$$\frac{\mathrm{d}z}{\mathrm{d}x}(x) = 2(xz(x) + \sqrt{z}), \qquad x \in (0,1).$$

2. Dada una ecuación diferencial de la forma

$$a(x)\frac{\mathrm{d}^2y}{\mathrm{d}x^2}(x) + b(x)\frac{\mathrm{d}y}{\mathrm{d}x}(x) + y(x) = 0,$$

encuentre funciones $a(x)$ y $b(x)$ tales que $y_1(x) = x$ y $y_2(x) = x^2$ sean soluciones de la ecuación.

3. (a) Resuelva la ecuación

$$y(x)\frac{\mathrm{d}y}{\mathrm{d}x}(x) = x.$$

   (b) Encuentre las soluciones $y(x)$ de la ecuación anterior tales que
      i. $y(2) = 1$;
      ii. $y(2) = -1$;
      iii. $y(-2) = -1$.

   (c) Dibuje la gráfica de las soluciones del punto *b*).

4. Encuentre todas las soluciones de las ecuaciones de primer orden siguientes.

   (a) $\dfrac{\mathrm{d}y}{\mathrm{d}x} + \mathrm{e}^x y = 3\mathrm{e}^x$

   (b) $\dfrac{\mathrm{d}y}{\mathrm{d}x} + 2xy = x\mathrm{e}^{-x^2}$

   (c) $\dfrac{\mathrm{d}y}{\mathrm{d}x} - (\tan x)y = \mathrm{e}^{\operatorname{sen} x}$ para $x \in (0, \frac{\pi}{2})$

   (d) $\dfrac{\mathrm{d}y}{\mathrm{d}x} + 2y = f(x)$, para $x \in \mathbb{R}$, donde $f(x) = 1 - |x|$ para $|x| \leq 1$ y $f(x) = 0$ si $|x| > 1$.

5. Considere la *ecuación de Bernoulli*[3]

$$\frac{\mathrm{d}y}{\mathrm{d}x}(x) + a(x)y(x) = f(x)y^k(x), \quad k \text{ constante.}$$

[3] Se refiere a Jacobo Bernoulli, quien la propuso en 1695 y fue resuelta por su hermano Juan Bernoulli. La sustitución $z = y^{1-k}$, que la linealiza, se debe a Leibniz.

Si $y(x)$ es solución de la ecuación anterior, pruebe que la función

$$z(x) = y^{1-k}(x)$$

es solución de

$$\frac{\mathrm{d}z}{\mathrm{d}x}(x) + (1-k)a(x)z(x) = (1-k)f(x).$$

Con la información anterior, encuentre todas las soluciones de la ecuación de Bernoulli

$$\frac{\mathrm{d}y}{\mathrm{d}x}(x) - 2xy(x) = xy^2(x).$$

6. Si $y(x)$ es solución de la ecuación

$$\frac{\mathrm{d}^2y}{\mathrm{d}x^2} = -y,$$

diga de qué ecuación diferencial es solución la función

$$z(x) = \frac{1}{2}y^2(x) + \frac{1}{2}\left(\frac{\mathrm{d}y}{\mathrm{d}x}\right)^2(x).$$

7. Muestre que si $y(x)$ es solución de la ecuación $\frac{\mathrm{d}^2y}{\mathrm{d}x^2} + b\frac{\mathrm{d}y}{\mathrm{d}x} + ay = 0$, entonces la función $z(x) = \mathrm{e}^{\frac{b}{2}x}y(x)$ es solución de la ecuación

$$\frac{\mathrm{d}^2z}{\mathrm{d}x^2} = \left(\frac{b^2}{4} - a\right)z.$$

8. Encuentre la solución $y(x)$ de la ecuación

$$\frac{\mathrm{d}^2y}{\mathrm{d}x^2}(x) + 2\frac{\mathrm{d}y}{\mathrm{d}x}(x) + 4y(x) = 1$$

tal que $y(0) = 2$ y $\frac{\mathrm{d}y}{\mathrm{d}x}(0) = -1$.

9. Encuentre todas las soluciones de las ecuaciones dadas a continuación.

   (a) $\frac{\mathrm{d}^2y}{\mathrm{d}x^2}(x) + 4y(x) = \cos x$

   (b) $6\frac{\mathrm{d}^2y}{\mathrm{d}x^2}(x) + 5\frac{\mathrm{d}y}{\mathrm{d}x}(x) - 6y(x) = x$

   (c) $\frac{\mathrm{d}^2y}{\mathrm{d}x^2}(x) - \frac{\mathrm{d}y}{\mathrm{d}x}(x) + 5y(x) = 3\mathrm{e}^{-x} + 2x^2$

10. Encuentre la curva $y = f(x)$ que pasa por el punto $(1, 1)$ y la pendiente de su normal en cada punto $(x, f(x))$ es igual a $\frac{x}{f(x)}$.

Capítulo

12 Series

La posibilidad de dar sentido a la suma de un número infinito ordenado de números reales, es una de las consecuencias importantes del concepto fundamental de límite de una sucesión. Esta nueva operación, propia del cálculo, da lugar a la noción de serie numérica, que de hecho se ha manejado desde el primer capítulo de este texto, con la representación decimal de los números reales. Ahora, con el apoyo de la teoría de los límites, se retoma ese concepto y se demuestran los resultados básicos sobre convergencia de series numéricas, incluyendo los criterios más importantes para comprobar la existencia de tal propiedad. Finalmente, se introduce la familia de las funciones analíticas, definidas como aquéllas que son representables como series de potencias. Esta clase de funciones resulta ser de gran importancia en varias áreas del análisis matemático y sus aplicaciones.

12.1 Definición de serie y su suma

Cada sucesión de números reales $\{a_i\}_{i=1}^{\infty}$ da lugar a la sucesión de *sumas parciales* $\{s_j\}_{j=1}^{\infty}$ definida por

$$s_j = a_1 + a_2 + \cdots + a_j = \sum_{i=1}^{j} a_i, \quad \text{para} \quad j = 1, 2, \ldots$$

A esta sucesión de sumas se le llama la *serie inducida por la sucesión* $\{a_i\}_{i=1}^{\infty}$ y se denota con el símbolo

$$\sum_{i=1}^{\infty} a_i.$$

A la sucesión inicial $\{a_i\}_{i=1}^{\infty}$ se le dice *sucesión de sumandos de la serie* $\sum_{i=1}^{\infty} a_i$.

Ejemplo 12.1 1. La sucesión real $\left\{\dfrac{1}{i}\right\}_{i=1}^{\infty}$ induce la *serie armónica*

$$\{b_j\}_{j=1}^{\infty} = \left\{\sum_{i=1}^{j} \frac{1}{i}\right\}_{j=1}^{\infty} = \sum_{i=1}^{\infty} \frac{1}{i}.$$

2. La sucesión $\left\{r^i\right\}_{i=1}^{\infty}$, donde r es un número real, induce la *serie geométrica*

$$\{s_j\}_{j=1}^{\infty} = \left\{\sum_{i=1}^{j} r^i\right\}_{j=1}^{\infty} = \sum_{i=1}^{\infty} r^i. \qquad \triangleleft$$

Para la serie geométrica con $r \neq 1$, el j-ésimo término de la sucesión s_j de sumas parciales está dado por

$$s_j = r + r^2 + \cdots + r^j = \frac{r(1-r^j)}{1-r}, \quad j = 1, 2, \ldots,$$

como se deduce directamente de la relación

$$rs_j - s_j = r^{j+1} - r.$$

Para $r = 1$, la serie geométrica coincide con la sucesión de números naturales $\{j\}_{j=1}^{\infty}$.

Tomando en cuenta que una serie es una sucesión de números reales, se dice que la serie $\sum_{i=1}^{\infty} a_i$ es *convergente a* S (o *que tiene suma* S) si la sucesión correspondiente de sumas parciales $\left\{\sum_{i=1}^{j} a_i\right\}_{j=1}^{\infty}$ tiene por límite al número S. En tal caso, se escribe

$$\sum_{i=1}^{\infty} a_i = \lim_{j\to\infty} \sum_{i=1}^{j} a_i = S.$$

Si una serie no es convergente, se le llama serie *divergente*.

Nota Importante:
Si la serie $\sum_{i=1}^{\infty} a_i$ es convergente, denotamos a su suma S con el mismo símbolo que la serie, es decir $S = \sum_{i=1}^{\infty} a_i$.

Ejemplo 12.2 La serie geométrica $\displaystyle\sum_{i=1}^{\infty} a_i = \left\{\sum_{i=1}^{j} r^i\right\}_{j=1}^{\infty} = \left\{\frac{r(1-r^j)}{1-r}\right\}_{j=1}^{\infty}$ es convergente si $|r| < 1$ y tiene por límite o suma el número

$$S = \frac{r}{1-r}.$$

Si $|r| \geqslant 1$, la serie geométrica es divergente. $\triangleleft$

Ejemplo 12.3 La serie $\sum_{i=1}^{\infty}(-1)^i$ no es convergente, ya que la sucesión de sumas parciales correspondiente toma los valores

$$s_j = \begin{cases} 0 \text{ si } j \text{ es par} \\ -1 \text{ si } j \text{ es impar} \end{cases}$$

y, por lo tanto, no converge. $\triangleleft$

Nota Importante:
Si se modifica un número finito de sumandos de una serie convergente, la serie resultante seguirá siendo convergente.

12.2 Propiedades de las series convergentes

Las propiedades de las sucesiones convergentes que se presentaron en el capítulo 3, se trasladan de manera automática al caso de series.

Proposición 12.1 *Las propiedades principales de las series convergentes son:*

1. *La sucesión de sumandos de toda serie convergente es una sucesión convergente a cero.*

2. *La suma y la multiplicación por un número real de series convergentes es una serie convergente. Más aún, si* $\sum_{i=1}^{\infty} a_i = S$ *y* $\sum_{i=1}^{\infty} b_i = M$, *entonces* $\sum_{i=1}^{\infty}(\lambda a_i + b_i) = \lambda S + M$.

Demostración. Si $\sum_{i=1}^{\infty} a_i = S$, las sucesiones de sumas parciales$\{s_j\}_{j=1}^{\infty} = \left\{\sum_{i=1}^{j} a_i\right\}$ y $\{p_j\}_{j=1}^{\infty} = \left\{\sum_{i=1}^{j+1} a_i\right\}$ convergen ambas a S y por lo tanto, su diferencia, que es la sucesión de términos $\{a_{j+1}\}_{j=1}^{\infty}$ deberá converger a cero, es decir,

$$\lim_{j\to\infty} a_{j+1} = \lim_{j\to\infty} p_j - \lim_{j\to\infty} s_j = 0.$$

Por lo tanto, la sucesión de los sumandos de toda serie convergente es necesariamente convergente a cero.

Para probar la validez del punto 2, observemos que las series $\sum_{i=1}^{\infty}(a_i + b_i)$ y $\sum_{i=1}^{\infty} \lambda a_i$, son la suma y el producto por un escalar de sucesiones convergentes y, por lo tanto, serán convergentes a la suma y al producto por el escalar de los límites correspondientes a esas sucesiones. ■

Nota Importante:
La convergencia a cero de los términos de una serie no es condición suficiente para asegurar la convergencia de la serie misma. Basta considerar como contraejemplo

a la serie armónica $\sum_{i=1}^{\infty} \frac{1}{i}$*, la cual, a pesar de que la sucesión de sumandos tiende a cero, tiene una sucesión de sumas parciales que crece sin límite, como se muestra a partir de las estimaciones siguientes:*

$$\frac{1}{3} + \frac{1}{4} > \frac{2}{4} = \frac{1}{2}$$
$$\frac{1}{5} + \frac{1}{6} + \frac{1}{7} + \frac{1}{8} > 4\left(\frac{1}{8}\right) > \frac{1}{2}$$
$$\vdots$$
$$\frac{1}{2^n + 1} + \frac{1}{2^n + 2} + \cdots + \frac{1}{2^{n+1}} > 2^n(\frac{1}{2^{n+1}}) = \frac{1}{2}$$

es decir, la sucesión $s_j = \sum_{i=1}^{j} \frac{1}{i}$, $j = 1, 2, \ldots$ *es tal que*

$$s_{2^{n+1}} - s_{2^n} > \frac{1}{2}$$

y por lo tanto

$$s_{2^n} > \frac{n}{2},$$

lo cual muestra que la sucesión de sumas parciales no es acotada y, consecuentemente, la serie armónica no es convergente.

Antes de discutir los diversos criterios de convergencia, introduciremos algunas de las familias de series de mayor importancia:

a) Una serie de números reales $\sum_{i=1}^{\infty} a_i$ se dice que es una *serie positiva* si $a_i > 0$ para $i = 1, 2, \ldots$

b) A una serie de la forma

$$\sum_{i=1}^{\infty} (-1)^{i+1} a_i, \quad \text{con} \quad a_i > 0 \quad \text{para} \quad i = 1, 2, \ldots,$$

se le llama *serie alternante.*

c) Una serie $\sum_{i=1}^{\infty} a_i$ se dice *serie telescópica* si existe una sucesión real $\{b_i\}_{i=1}^{\infty}$ tal que

$$a_i = b_{i+1} - b_i \quad \text{para} \quad i = 1, 2, \ldots$$

En el caso de las series telescópicas, cada una de sus sumas parciales s_k para $k = 1, 2, \ldots$ es de la forma

$$s_k = \sum_{i=1}^{k} a_i = b_{k+1} - b_1$$

y, por lo tanto, la serie telescópica será convergente si la sucesión correspondiente $\{b_k\}_{k=1}^{\infty}$ que la genera es convergente. En ese caso,

$$\sum_{i=1}^{\infty} a_i = \lim_{k\to\infty} b_{k+1} - b_1.$$

Ejemplo 12.4 La serie $\displaystyle\sum_{n=1}^{\infty} \frac{1}{n(n+1)}$ es telescópica ya que

$$\frac{1}{n(n+1)} = -\frac{1}{n+1} + \frac{1}{n}$$

y, por lo tanto, su suma es

$$\sum_{n=1}^{\infty} \frac{1}{n(n+1)} = 1. \qquad \triangleleft$$

12.3 Series positivas

Consideremos ahora una serie de términos positivos $\sum_{i=1}^{\infty} a_i$. La característica más notable de las series positivas es que su sucesión de sumas parciales es una sucesión creciente y positiva de reales y por lo tanto, de acuerdo al teorema 4.8, será convergente si y sólo si es acotada, es decir, si y sólo si existe $M > 0$ tal que

$$\left|\sum_{i=1}^{j} a_i\right| \leqslant M, \quad j = 1, 2, \ldots$$

Desafortunadamente, este criterio de convergencia es a menudo difícil de aplicar y resulta más conveniente recurrir al llamado *criterio de comparación*, que consiste en comparar una serie positiva con otra serie cuya convergencia o divergencia se conoce, pudiéndose deducir de esto la convergencia o divergencia de la serie inicial.

Criterio de Comparación. Si $\sum_{i=1}^{\infty} a_i$ y $\sum_{i=1}^{\infty} b_i$ son series de términos positivos tales que:

a) $\sum_{i=1}^{\infty} a_i = S$ y

b) $b_i \leqslant a_i$, para $i \geqslant N$ con $n \in \mathbb{N}$,

entonces, la serie $\sum_{i=1}^{\infty} b_i$ es convergente y $\displaystyle\sum_{i=1}^{\infty} b_i \leqslant S$. Análogamente, si la serie $\sum_{i=1}^{\infty} a_i$ es divergente y $a_i \leqslant b_i$, para $i \geqslant N$, entonces la serie $\sum_{i=1}^{\infty} b_i$ es divergente.

Demostración. Para mostrar la validez del criterio de comparación, observemos que bajo las hipótesis se tiene, para cada $j > N$, que

$$\sum_{i=N}^{j} a_i \leqslant \sum_{i=N}^{\infty} a_i = S - \sum_{i=1}^{N-1} a_i$$

y entonces obtenemos la estimación

$$\sum_{i=1}^{j} b_i = \sum_{i=1}^{N-1} b_i + \sum_{i=N}^{j} b_i \leqslant \sum_{i=1}^{N-1} b_i + \sum_{i=N}^{j} a_i \leqslant S - \sum_{i=1}^{N-1} a_i + \sum_{i=1}^{N-1} b_i,$$

lo cual muestra que la sucesión de sumas parciales $\sum_{i=1}^{j} b_i$, $j = 1, 2, \ldots$ es acotada por el número $S - \sum_{i=1}^{N-1} a_i + \sum_{i=1}^{N-1} b_i$ y, por consiguiente, la serie $\sum_{i=1}^{\infty} b_i$ es convergente.

Por otro lado, si la serie $\sum_{i=1}^{\infty} a_i$ es divergente y $a_i \leqslant b_i$ para $i \geqslant N$, entonces

$$\sum_{i=1}^{j} a_i = \sum_{i=1}^{N-1} a_i + \sum_{i=N}^{j} a_i \leqslant \sum_{i=1}^{N-1} a_i + \sum_{i=N}^{j} b_i$$

y por lo tanto, la sucesión $\sum_{i=N}^{j} b_i$ crecerá sin límite cuando $j \to \infty$, ya que es mayor que la sucesión de números positivos $\sum_{i=1}^{j} a_i$, que diverge. ■

De la aplicación del criterio de comparación con respecto a la serie geométrica, se deduce el llamado *criterio del cociente* que se debe al matemático francés Jean Le Rond D'Alembert y que se enuncia en los términos siguientes:

Criterio del Cociente (de D'Alembert). Si la sucesión de números positivos $\{a_i\}_{i=1}^{\infty}$ es tal que

$$\lim_{i\to\infty} \frac{a_{i+1}}{a_i} = L,$$

entonces la serie $\sum_{i=1}^{\infty} a_i$ es convergente si $L < 1$ y es divergente si $L > 1$.

Demostración. Supongamos primero que $L < 1$. De la definición de límite, dado $r = \dfrac{1-L}{2} < 1$, existe un número natural N tal que

$$\left|\frac{a_{i+1}}{a_i} - L\right| \leqslant \frac{1-L}{2}, \text{ para } i \geqslant N. \tag{12.1}$$

Luego,

$$a_{i+1} \leqslant \left(\frac{1+L}{2}\right) a_i, \text{ para } i \geqslant N,$$

y la aplicación reiterada de esta última estimación, nos lleva a

$$a_{i+1} \leqslant \left(\frac{1+L}{2}\right)^{i-N+1} a_N, \text{ para } i \geqslant N.$$

Esto significa que cada término a_{i+1} con $i \geqslant N$ de la serie inicial, es menor o igual que el término $b_{i+1} = cr^{i+1}$ con $r = \left(\dfrac{1+L}{2}\right)$ y $c = \left(\dfrac{1+L}{2}\right)^{-N} a_N$. Al ser $r < 1$, la serie geométrica $\sum_{i=1}^{\infty} r^i$ es convergente y mayor, término a término, que la serie inicial $\sum_{i=1}^{\infty} a_i$ a partir de $i = N$. Entonces, por el criterio de comparación, se deduce la convergencia de la serie $\sum_{i=1}^{\infty} a_i$. Si, por lo contrario, se tuviera $L > 1$, tomando $r = \dfrac{L-1}{2}$ y aplicando la estimación (12.1), se tendrá

$$\frac{a_{i+1}}{a_i} - L \geqslant -\frac{L-1}{2} \text{ para } i \geqslant N,$$

o equivalentemente,

$$a_{i+1} \geqslant \left(\frac{L+1}{2}\right) a_i \text{ para } i \geqslant N.$$

Aplicando repetidamente la estimación anterior, se obtiene que

$$a_{N+k} \geqslant \left(\frac{L+1}{2}\right)^k a_N \quad \text{para } k = 1, 2, \ldots$$

y, tomando en cuenta que $\dfrac{L+1}{2} > 1$, se tendrá

$$\sum_{i=N}^{\infty} a_i \geqslant a_N \sum_{i=N}^{\infty} \left(\frac{L+1}{2}\right)^k$$

y, por el criterio de comparación, la serie $\sum_{i=1}^{\infty} a_i$ será divergente. ■

NOTA IMPORTANTE:
Si para una serie positiva $\Sigma_{i=1}^{\infty} a_i$ se tiene $\lim\limits_{i\to\infty} \dfrac{a_{i+1}}{a_i} = 1$, entonces no se puede asegurar la convergencia ni la divergencia de tal serie. Para mostrar lo anterior, basta hacer notar que tanto la serie convergente de términos positivos $\displaystyle\sum_{n=1}^{\infty} \frac{1}{n(n+1)}$ como la serie divergente de términos positivos $\displaystyle\sum_{n=1}^{\infty} \frac{1}{n}$, son tales que su sucesión de cocientes $\left\{\dfrac{a_{i+1}}{a_i}\right\}_{i=1}^{\infty}$ es convergente a 1.

Ejemplo 12.5 La serie $\sum_{n=1}^{\infty} n^3 e^{-n}$ es convergente ya que al aplicar el criterio del cociente se tiene

$$\lim_{n\to\infty} \frac{(n+1)^3 \mathrm{e}^{-n-1}}{n^3 \mathrm{e}^{-n}} = \frac{1}{\mathrm{e}} < 1. \qquad \triangleleft$$

Criterio de Cauchy. Si la sucesión de números positivos $\{a_i\}_{i=1}^{\infty}$ es tal que

$$\lim_{n\to\infty} \sqrt[n]{a_n} = L,$$

entonces la serie $\sum_{n=1}^{\infty} a_n$ es convergente si $L < 1$ y es divergente si $L > 1$.

Demostración. Si $\lim_{n\to\infty} \sqrt[n]{a_n} = L < 1$, existe $r < 1$ y N natural tal que $\sqrt[n]{a_n} < r < 1$ para $n > N$. Luego,

$$a_n < r^n \quad \text{si} \quad n > N$$

y, por el criterio de comparación, se sigue que $\sum_{i=1}^{\infty} a_i$ es convergente. Por el contrario, si $\lim_{n\to\infty} \sqrt[n]{a_n} = L > 1$, existirán $r > 1$ y N natural tales que $\sqrt[n]{a_n} > r > 1$ para $n > N$; luego,

$$a_n > r^n \quad \text{si} \quad n > N$$

y, aplicando el criterio de comparación, se tendrá que $\sum_{i=1}^{\infty} a_i$ diverge. ∎

Criterio de comparación al límite. Sean $\{a_i\}_{i=1}^{\infty}$ y $\{b_i\}_{i=1}^{\infty}$ sucesiones de números positivos y

$$\lim_{i\to\infty} \frac{a_i}{b_i} = L,$$

donde L es un número no-negativo o ∞:

a) Si $L < \infty$ y $\sum_{i=1}^{\infty} b_i$ converge, entonces $\sum_{i=1}^{\infty} a_i$ también converge.

b) Si $L > 0$ y $\sum_{i=1}^{\infty} b_i$ diverge, entonces $\sum_{i=1}^{\infty} a_i$ también diverge.

Demostración. Si $L < \infty$, entonces existe una etiqueta N tal que

$$\frac{a_i}{b_i} \leqslant L + 1 \quad \text{para} \quad i > N\ ;$$

luego,

$$a_i \leqslant (L+1)b_i \quad \text{para} \quad i > N,$$

y aplicando el criterio de comparación, se concluye que la serie $\sum_{i=1}^{\infty} a_i$ es convergente. Para probar el punto 2, observamos que si $L > 0$, existirá una etiqueta N tal que

$$\frac{a_i}{b_i} > \frac{L}{2}, \quad \text{para } i > N;$$

por lo tanto,

$$b_i < \frac{2}{L} a_i, \quad \text{para } i > N,$$

y entonces, si $\sum_{i=1}^{\infty} b_i$ diverge, aplicando el criterio de comparación concluimos que $\sum_{i=1}^{\infty} a_i$ deberá también diverger. Note que si $\lim_{i\to\infty} \frac{a_i}{b_i} = \infty$, se tendría que $\frac{a_i}{b_i} > 1$ para toda etiqueta mayor que un cierto natural K. Luego, $b_i < a_i$ para $i > K$, y por el criterio de comparación se concluye que, si $\sum_{i=1}^{\infty} b_i$ diverge, entonces $\sum_{i=1}^{\infty} a_i$ diverge. ■

Ejemplo 12.6 La serie $\sum_{n=1}^{\infty}(\sqrt[2]{n^2+1} - n)$ es divergente ya que para cada $n \in \mathbb{N}$ se tiene la comparación

$$\sqrt{n^2+1} - n = \frac{1}{\sqrt{n^2+1}+n} > \frac{1}{2\sqrt{n^2+1}} > \frac{1}{2n+2}.$$

Dado que la serie $\sum_{n=1}^{\infty} \frac{1}{2n+2}$ es divergente, se tiene que la primera, por el criterio de comparación, será también divergente. ◁

Ejemplo 12.7 La serie $\displaystyle\sum_{i=1}^{\infty} \operatorname{sen}\frac{1}{i}$ es divergente ya que al comparar con la serie inducida por la sucesión $\displaystyle\left\{\frac{1}{i}\right\}_{i=1}^{\infty}$ calculando el límite,

$$\lim_{i\to\infty} \frac{\operatorname{sen}\frac{1}{i}}{\frac{1}{i}} = 1,$$

luego, al ser $\displaystyle\sum_{i=1}^{\infty} \frac{1}{i}$ divergente, la serie $\displaystyle\sum_{i=1}^{\infty} \operatorname{sen}\frac{1}{i}$ será divergente. ◁

Criterio de la Integral. Si la sucesión $\{a_i\}_{i=1}^{\infty}$ es tal que a partir de una etiqueta N se tiene $a_i = f(i)$ para $i = N, N+1, \ldots$, donde $f : [N, \infty) \to \mathbb{R}$ es una función continua, positiva y $f(x) \leqslant f(y)$ si $x \geqslant y$, entonces

$$\sum_{i=N}^{\infty} a_i \qquad \text{y} \qquad \int_N^{\infty} f(x)\mathrm{d}x,$$

convergen ambas o divergen ambas.

Demostración. Comparando, para cada k natural, el término $\sum_{i=N}^{N+k} a_i$ con la integral $\int_N^{N+k} f(x)\mathrm{d}x$, se tiene en virtud del carácter no-creciente de f, la siguiente estimación:

$$\sum_{i=N}^{N+k} f(i+1) \leqslant \int_N^{N+k} f(x)\mathrm{d}x \leqslant \sum_{i=N}^{N+k} f(i).$$

Note que los términos $\sum_{i=N}^{N+k} f(i+1)$, $\sum_{i=N}^{N+k} f(i)$ son las sumas inferior y superior de Darboux-Riemann, respectivamente, asociadas a la partición
$\mathcal{P} = \{N, N+1, N+2, \ldots, N+k\}$ del intervalo $[N, N+k]$. Véase la figura 12.1.

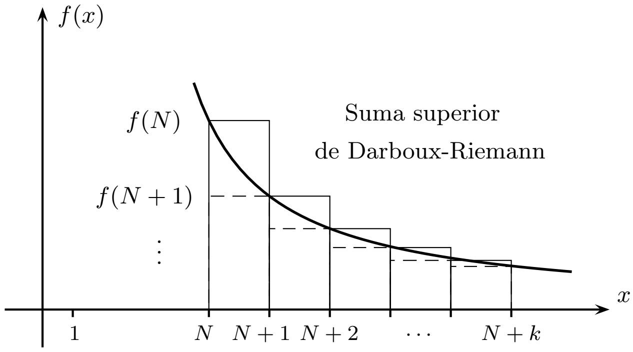


Figura 12.1 Sumas inferior y superior de Darboux-Riemann

Tomando límites cuando $k \to \infty$, se tiene que

$$\sum_{i=N}^{\infty} a_{i+1} \leqslant \int_{N}^{\infty} f(x)\mathrm{d}x \leqslant \sum_{i=N}^{\infty} a_i,$$

lo cual demuestra que la serie $\sum_{i=N}^{\infty} a_i$ y la integral impropia $\int_N^{\infty} f(x)\mathrm{d}x$, o ambas convergen o ambas divergen. ■

Ejemplo 12.8 Una aplicación importante del criterio de la integral es a la serie

$$\sum_{i=1}^{\infty} \frac{1}{i^p}, \tag{12.2}$$

donde p es un número real fijo. Comparando con la función $f(x) = x^{-p}$, se tiene que $i^{-p} = f(i)$, luego, tomando en cuenta que

$$\int_1^{\infty} x^{-p}\mathrm{d}x = \frac{1}{1-p} x^{-p+1}\Big|_1^{\infty},$$

se concluye que la serie (12.2) converge si $p > 1$ y diverge si $p \leqslant 1$. ◁

12.4 Series absolutamente convergentes

Una serie $\sum_{i=1}^{\infty} a_i$ se dice *absolutamente convergente* si la serie de números no negativos $\sum_{i=1}^{\infty} |a_i|$ es convergente. La importancia de este tipo de series radica en el resultado siguiente.

Proposición 12.2 *Toda serie absolutamente convergente es convergente.*

Demostración. Sea $\sum_{i=1}^{\infty} a_i$ una serie absolutamente convergente y consideremos las sucesiones de sumas parciales $\{s_i\}_{i=1}^{\infty} = \left\{\sum_{j=1}^{i} a_j\right\}_{i=1}^{\infty}$ y $\{p_i\}_{i=1}^{\infty} = \left\{\sum_{j=1}^{i} |a_j|\right\}_{i=1}^{\infty}$. Por hipótesis, la sucesión de términos no negativos $\{p_i\}_{i=1}^{\infty}$ es convergente y, por lo tanto, está acotada por el número $M > 0$, digamos. Aplicando la desigualdad del triángulo se tiene, para cada $i = 1, 2, \ldots,$ que

$$\left|\sum_{j=1}^{i} a_j\right| \leqslant \sum_{j=1}^{i} |a_j| \leqslant M,$$

lo que muestra que también la sucesión de sumas parciales $\{s_i\}_{i=1}^{\infty}$ es acotada por M. Ahora formemos la serie con términos no-negativos

$$\sum_{i=1}^{\infty} a_i^{+}$$

donde $a_i^{+} = a_i$ si $a_i \geqslant 0$ y $a_i^{+} = 0$ si $a_i < 0$ y, análogamente, la serie de términos no-negativos

$$\sum_{i=1}^{\infty} a_i^{-},$$

donde $a_i^{-} = -a_i$ si $a_i < 0$ y $a_i^{-} = 0$ si $a_i \geqslant 0$. Estas dos últimas series son tales que sus respectivas sucesiones de sumas parciales son acotadas y, por lo tanto, al ser no-negativas, las dos series son convergentes. Por otro lado, la serie inicial se puede escribir en la forma

$$\sum_{i=1}^{\infty} a_i = \sum_{i=1}^{\infty} a_i^{+} - \sum_{i=1}^{\infty} a_i^{-},$$

es decir, es la diferencia de dos series convergentes. Luego, la serie $\sum_{i=1}^{\infty} a_i$ es también convergente. ■

Ejemplo 12.9 La serie $\displaystyle\sum_{n=1}^{\infty} \frac{(-1)^n}{n^2}$ es una serie convergente ya que es absolutamente convergente pues

$$\sum_{n=1}^{\infty} \left|\frac{(-1)^n}{n^2}\right| = \sum_{n=1}^{\infty} \frac{1}{n^2}$$

y ésta última es convergente. ◁

Nota Importante:

1. *Si una serie $\sum_{i=1}^{\infty} a_i$ es absolutamente convergente, se tiene la siguiente estimación para su suma:*

$$-\sum_{i=1}^{\infty}|a_i| \leqslant \sum_{i=1}^{\infty} a_i \leqslant \sum_{i=1}^{\infty}|a_i|.$$

2. *Es oportuno hacer notar, que si bien es cierto que toda sucesión absolutamente convergente es convergente, no necesariamente una serie que no es absolutamente convergente es divergente, por ejemplo, la serie*

$$\sum_{i=1}^{\infty} \frac{(-1)^n}{n}$$

no es absolutamente convergente pero sí es convergente, como lo probaremos después de la sección siguiente.

12.5 Los criterios de Abel y Dirichlet

Los criterios que hasta ahora hemos presentado se refieren a la convergencia o divergencia de series de términos positivos. En este apartado incluiremos dos criterios que se aplican a series que no convergen absolutamente y se basan en la llamada *fórmula de Abel*[1] para las sumas parciales. Incluimos su aplicación a las series alternantes

Teorema 12.3 (Fórmula de Abel) *Si $\{a_n\}_{n=1}^{\infty}$ y $\{b_n\}_{n=1}^{\infty}$ son dos sucesiones y $\{s_n\}_{n=1}^{\infty}$ es la sucesión de sumas parciales*

$$s_n = a_1 + a_2 + \cdots + a_n$$

entonces, para cada $n \in \mathbb{N}$, se tiene la siguiente identidad:

$$\begin{aligned}\sum_{k=1}^{n} a_k b_k &= \sum_{k=1}^{n}(s_k - s_{k-1})b_k \\ &= \sum_{k=1}^{n} s_k b_k - \sum_{k=1}^{n} s_k b_{k+1} + s_n b_{n+1} \\ &= \sum_{k=1}^{n} s_k(b_k - b_{k+1}) + s_n b_{n+1}, \qquad (12.3)\end{aligned}$$

donde hemos tomado $s_0 = 0$. Por lo tanto, $\sum_{k=1}^{\infty} a_k b_k$ converge si $\sum_{k=1}^{\infty} s_k(b_k - b_{k+1})$ y $\{s_n b_{n+1}\}_{n=1}^{\infty}$ convergen.

[1] Por Niels Henrik Abel (1802-1829), matemático noruego.

Nota Importante:
La fórmula de Abel es análoga a la fórmula de integración por partes.

Corolario 12.4 (Criterio de Dirichlet [2]) *Si la sucesión de las sumas parciales de la serie $\sum_{k=1}^{\infty} a_k$ está acotada y la sucesión $\{b_n\}_{n=1}^{\infty}$ es decreciente y converge a cero, entonces la serie $\sum_{k=1}^{\infty} a_k b_k$ es convergente.*

Demostración. Sea $s_n = a_1 + a_2 + \cdots + a_n$ y M una cota para esa sucesión; es decir,

$$| s_n | \leq M \quad \text{para todo } n \in \mathbb{N}.$$

Siendo $\{b_n\}_{n=1}^{\infty}$ decreciente, se tiene

$$| s_k(b_k - b_{k+1}) | \leq M(b_k - b_{k+1})$$

y, entonces, la serie $\sum_{k=1}^{\infty} s_k(b_k - b_{k+1})$ es convergente, pues es absolutamente convergente, ya que

$$\sum_{k=1}^{\infty} | s_k(b_k - b_{k+1}) | \leq M \sum_{k=1}^{\infty}(b_k - b_{k+1}) = Mb_1.$$

Por otro lado,

$$\lim_{n\to\infty} | s_n b_{n+1} | \leq M \lim_{n\to\infty} | b_{n+1} | = 0$$

y, por lo tanto, tomando en cuenta la fórmula de Abel (12.3), concluimos que la serie $\sum_{k=1}^{\infty} a_k b_k$ converge. ■

Corolario 12.5 (Criterio de Abel) *Si la serie $\sum_{k=1}^{\infty} a_k$ converge y la sucesión $\{b_n\}_{n=1}^{\infty}$ es decreciente y convergente, entonces la serie $\sum_{k=1}^{\infty} a_k b_k$ es convergente.*

Demostración. Repitiendo la argumentación dada en el corolario anterior y tomando en cuenta la convergencia de $\{s_n b_n\}_{n=1}^{\infty}$ se concluye la validez de este corolario. ■

[2] Por Lejeune Dirichlet.

Corolario 12.6 (Criterio de convergencia para series alternantes) *Toda serie alternante* $\sum_{i=1}^{\infty}(-1)^{i+1}a_i$, *donde* $a_i > 0$ *para* $i = 1, 2, \ldots,$ *que satisface las condiciones*

$$a_1 \geqslant a_2 \geqslant \cdots \geqslant a_i \geqslant \cdots \quad y \quad \lim_{i\to\infty} a_i = 0$$

es convergente. Más aún, la sucesión de sumas parciales $s_n = \sum_{i=1}^{n}(-1)^{i+1}a_i$ *para* $n = 1, 2, \cdots$ *satisface la desigualdad*

$$s_{2n} \leq \sum_{i=1}^{\infty}(-1)^{i+1}a_i \leq s_{2n+1}.$$

Demostración. La convergencia de la serie $\sum_{i=1}^{\infty}(-1)^{i+1}a_i$ se sigue directamente de la aplicación del criterio de Dirichlet. La estimación de la desigualdad es consecuencia del comportamiento de las sumas parciales pares e impares. Concretamente, la sucesión de sumas parciales pares s_{2n} es creciente ya que

$$s_{2n+2} = s_{2n} + a_{2n+1} - a_{2n+2} \geqslant s_{2n}$$

Análogamente, la subsucesión de sumas parciales s_{2n+1} para las etiquetas impares es decreciente ya que

$$s_{2n+3} = s_{2n+1} - a_{2n+2} + a_{2n+3} \leqslant s_{2n+1}.$$

Siendo ambas sucesiones convergentes a la suma $\sum_{i=1}^{\infty}(-1)^{i+1}a_i$, se sigue la validez de la desigualdad dada en el enunciado del corolario. ■

Ejemplo 12.10 Las series alternantes $\displaystyle\sum_{i=1}^{\infty}\frac{(-1)^n}{\sqrt{n}}$ y $\displaystyle\sum_{i=1}^{\infty}\frac{(-1)^n}{n}\sqrt{\frac{2n+1}{5n-1}}$ son convergentes. La primera satisface las condiciones del criterio de Dirichlet, mientras que la segunda, las del criterio de Abel. ◁

12.6 Series de potencias

En esta sección consideraremos series numéricas que dependen de una variable real x, en la forma siguiente:

$$\sum_{i=1}^{\infty} a_i x^i. \tag{12.4}$$

A este tipo de series se les llama *series de potencias* y como veremos, definen en cierto intervalo, una función con derivadas de todos los órdenes. Esta familia de funciones es de gran importancia para el cálculo y sus aplicaciones.

El primer concepto asociado a las series de potencias es el concepto de *intervalo de convergencia* que se define como el máximo intervalo abierto I con centro en cero tal que para cada valor $x \in I$, la serie numérica $\sum_{i=1}^{\infty} a_i x^i$ es absolutamente convergente.

Para justificar el concepto anterior, supongamos que al tomar $x = x_0$, la serie $\sum_{i=1}^{\infty} a_i x_0^i$ es convergente. Si tomamos ahora otro valor y con $|y| < |x_0|$, tendremos que la serie $\sum_{i=1}^{\infty} a_i y^i$ resulta ser absolutamente convergente, ya que

$$\sum_{i=1}^{\infty} |a_i y^i| \leqslant \sum_{i=1}^{\infty} |a_i||\frac{y}{x_0}|^i |x_0|^i \leqslant K \sum_{i=1}^{\infty} r^i$$

donde $r = |\frac{y}{x_0}| < 1$ y K es cota de la sucesión de términos de $\left\{a_i x_0^i\right\}_{i=1}^{\infty}$, misma que existe en virtud de la convergencia de la serie de potencias en $x = x_0$, es decir,

$$|a_i x_0^i| \leqslant K \quad \text{para } i = 1, 2, \ldots$$

Tomando en cuenta lo anterior, concluimos que

$$\sum_{i=1}^{\infty} |a_i y^i| \leqslant K \sum_{i=1}^{\infty} r^i \text{ con } r < 1$$

y, por lo tanto, aplicando el criterio de comparación entre series positivas, concluimos que la serie es absolutamente convergente.

Nota Importante:
Al valuar una serie de potencias en $x = 0$, se tiene una serie convergente a cero. Más aún, es posible que sólo en ese punto la serie sea convergente; por ejemplo, tenemos el caso de la serie

$$\sum_{n=1}^{\infty} n! x^n$$

que sólo es absolutamente convergente en $x = 0$, ya que para cualquier otro valor de x se tiene

$$\lim_{n\to\infty} \frac{(n+1)!|x|^{n+1}}{n!|x|^n} = \lim_{n\to\infty} (n+1)|x| > 1,$$

y la serie sólo converge para $x = 0$. El otro caso extremo se tiene cuando la serie es absolutamente convergente para todo $x \in \mathbb{R}$, como es el caso de la serie de potencias

$$\sum_{n=1}^{\infty} \frac{1}{n!} x^n,$$

cuyo intervalo de convergencia es $(-\infty, \infty)$.

Ejemplo 12.11 La serie de potencias $\sum_{n=1}^{\infty} \frac{1}{n} x^n$ tiene por intervalo de convergencia a $(-1,1)$ ya que si $|x|<1$ se tiene que

$$\lim_{n\to\infty} \frac{n}{(n+1)} |x| < 1$$

y, por el criterio del cociente para series positivas, la serie es convergente. Note que si $x=1$ la serie de potencias no es convergente. ◁

Como se ha mostrado, una serie de potencias define una función $f(x)$ cuyo dominio es el intervalo de convergencia, que denotaremos

$$f(x) = \sum_{n=1}^{\infty} a_n x^n. \tag{12.5}$$

Enseguida mostraremos que $f(x)$ es una función derivable y calcularemos tanto su derivada como sus primitivas.

Lema 12.7 *Si la serie de potencias* $\sum_{n=1}^{\infty} a_n x^n$ *es absolutamente convergente en* $(-c,c)$, *entonces la serie*

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

es también absolutamente convergente en $(-c,c)$.

Demostración. Para cada $x_0 \in (-c,c)$, probaremos que la serie $\sum_{k=1}^{\infty} n a_n x_0^{n-1}$ es absolutamente convergente en $x=x_0$. Sea $r>0$ con $|x_0|+r<c$. Tomando en cuenta que $|x_0|+r \in (-c,c)$, se tiene que la serie

$$\sum_{n=1}^{\infty} |a_n|(|x_0|+r)^n$$

es absolutamente convergente. Por otro lado, por el desarrollo de una potencia de un binomio, se tiene que

$$(|x_0|+r)^n = |x_0|^n + n|x_0|^{n-1}r + \sum_{k=2}^{n} \binom{n}{k} |x_0|^{n-k} r^k$$

y, por lo tanto,

$$n|x_0|^{n-1} \leqslant \frac{1}{r}(|x_0|+r)^n$$

y

$$\sum_{n=1}^{\infty} n|a_n||x_0|^{n-1} \leqslant \frac{1}{r} \sum_{n=1}^{\infty} |a_n|(|x_0|+r)^n.$$

Tomando en cuenta que la serie de la derecha es convergente pues $|x_0|+r \in (-c,c)$, por el criterio de comparación, tenemos que la serie positiva de la izquierda deberá ser también convergente, con lo que se prueba el lema. ∎

Nota Importante:
El lema anterior nos permite concluir que cada una de las series de potencias obtenidas tomando derivadas de orden superior, término a término, de una serie de potencias inicial, serán también absolutamente convergentes en el mismo intervalo $(-c,c)$. Es decir, las series

$$\sum_{n=2}^{\infty} n(n-1)a_n x^{n-2}, \quad \sum_{n=3}^{\infty} n(n-1)(n-2)a_n x^{n-3}, \quad \cdots$$

son absolutamente convergentes en $(-c,c)$.

Apoyados en el lema 12.7 probaremos la proposición siguiente.

Proposición 12.8 *La función $f(x)=\sum_{n=1}^{\infty} a_n x^n$ es una función derivable en el interior $(-c,c)$ de su intervalo de convergencia I y su función derivada tiene la forma*

$$\frac{df}{dx}(x) = \sum_{n=1}^{\infty} n a_n x^{n-1}, \quad x \in I.$$

Demostración. Para probar la proposición anterior, sean $x_0 \in (-c,c)$, h y r tales que $r>0$, $x_0+r \in (-c,c)$ y $|h|<r$. Demostraremos que

$$\lim_{h\to 0}\left[\frac{f(x_0+h)-f(x_0)}{h} - \sum_{n=1}^{\infty} n a_n x_0^{n-1}\right] = 0. \tag{12.6}$$

Aplicando el teorema del valor medio, podemos escribir, para cada $n=1,2,\ldots,$

$$\frac{(x_0+h)^n - x_0^n}{h} = n x_{n,h}^{n-1}$$

donde $x_{n,h} \in (x_0, x_0+h)$. Sustituyendo esas estimaciones en (12.6), se tiene

$$\frac{f(x_0+h)-f(x_0)}{h} - \sum_{n=1}^{\infty} n a_n x_0^{n-1} = \sum_{k=1}^{\infty} n a_n (x_{n,h}^{n-1} - x_0^{n-1}). \tag{12.7}$$

Aplicando nuevamente el teorema del valor medio, podemos escribir

$$(x_{n,h}^{n-1} - x_0^{n-1}) = (n-1) z_{n,h}^{n-2}(x_{n,h} - x_0),$$

donde $z_{n,h} \in (x_0, x_{n,h})$. Sustituyendo esta nueva estimación en (12.7), tenemos

$$\frac{f(x_0+h)-f(x_0)}{h} - \sum_{n=1}^{\infty} n a_n x_0^{n-1} = \sum_{k=2}^{\infty} n(n-1) a_n z_{n,h}^{n-2} (x_{n,h} - x_0) \leqslant$$
$$\leqslant |h| \sum_{k=2}^{\infty} n(n-1) a_n z_{n,h}^{n-2}.$$

Tomando en cuenta que la serie

$$\sum_{k=2}^{\infty} n(n-1)|a_n||z_{n,h}|^{n-2} \leqslant \sum_{k=2}^{\infty} n(n-1)|a_n|(|x_0|+r)^{n-2}$$

es absolutamente convergente para $x = |x_0| + r \in (-c, c)$, como consecuencia de la aplicación reiterada del lema 12.7, se tiene

$$\lim_{h\to 0} \left| \frac{f(x_0+h)-f(x_0)}{h} - \sum_{n=1}^{\infty} n a_n x_0^{n-1} \right| = 0,$$

lo que prueba que

$$\frac{df}{dx}(x_0) = \sum_{n=1}^{\infty} n a_n x_0^{n-1}. \qquad \blacksquare$$

Nota Importante:
La proposición 12.8 nos permite afirmar que cada función dada como serie de potencias tiene derivadas de todos los órdenes y esas funciones derivadas se obtienen derivando término a término la serie inicial.

Proposición 12.9 *La función $f(x) = \sum_{n=1}^{\infty} a_n x^n$ tiene por antiderivada en el interior $(-c, c)$ de su intervalo de convergencia I a la función $g(x)$ dada por la serie de potencias*

$$g(x) = \sum_{n=1}^{\infty} \frac{1}{n+1} a_n x^{n+1}, \quad x \in (-c, c).$$

Demostración. Si $\sum_{n=1}^{\infty} |a_n||x|^n$converge para $x \in (-c, c)$, dado que para todo $n = 1, 2, \ldots$ se tiene que

$$\left| \frac{1}{n+1} a_n x^n \right| \leqslant |a_n x^n|,$$

podemos afirmar, en virtud del criterio de comparación, que la serie

$$\sum_{n=1}^{\infty} \left| \frac{1}{n+1} a_n x^n \right|$$

es también convergente para $x \in (-c, c)$. De lo anterior, se deduce que la serie

$$\sum_{n=1}^{\infty} \left| \frac{1}{n+1} a_n x^{n+1} \right| = |x| \sum_{n=1}^{\infty} \left| \frac{1}{n+1} a_n x^n \right|$$

también converge para cada $x \in (-c, c)$. Como la serie inicial $\sum_{n=1}^{\infty} a_n x^n$ se forma derivando término a término la serie $\displaystyle\sum_{n=1}^{\infty} \frac{1}{n+1} a_n x^{n+1}$, aplicando la proposición 12.8, se tiene

$$\frac{\mathrm{d}g}{\mathrm{d}x}(x) = f(x),$$

lo que demuestra que

$$\int \sum_{n=1}^{\infty} a_n x^n \mathrm{d}x = \sum_{n=1}^{\infty} \int a_n x^n \mathrm{d}x. \qquad \blacksquare$$

Nota Importante:

1. *Las series de potencias también aparecen en la forma*

$$\sum_{n=1}^{\infty} a_n (x - x_0)^n$$

   y lo que hemos dicho y mostrado se extiende de manera natural trasladando el origen $x = 0$ al punto $x = x_0$.

2. *A las funciones representables como series de potencias se les llama* funciones analíticas. *Estas funciones tienen derivadas de todos los órdenes y cada k-suma parcial de la serie de potencias corresponde a su polinomio de Taylor de orden k. Una observación importante aquí es que no toda función que tenga derivadas de todos los órdenes es representable como serie de potencias. Enseguida daremos el ejemplo típico al respecto.*

Ejemplo 12.12 La función

$$f(x) = \begin{cases} 0 & \text{si } x \leqslant 0 \\ \mathrm{e}^{-\frac{1}{x^2}} & \text{si } x > 0 \end{cases}$$

tiene derivadas de todos los órdenes en $x = 0$, con valor cero, ya que $\dfrac{\mathrm{d}}{\mathrm{d}x}\mathrm{e}^{-\frac{1}{x^2}} = \dfrac{2}{x^3}\mathrm{e}^{-\frac{1}{x^2}}$ si $x > 0$ y dado que $\displaystyle\lim_{x \to 0} \frac{2}{x^3}\mathrm{e}^{-\frac{1}{x^2}} = 0$, se tiene

$$\frac{\mathrm{d}f}{\mathrm{d}x}(x) = \begin{cases} 0 & \text{si } x \leqslant 0 \\ \dfrac{2}{x^3}\mathrm{e}^{-\frac{1}{x^2}} & \text{si } x > 0. \end{cases}$$

Aplicando repetidamente un argumento análogo al anterior, se prueba que f tiene derivadas de todos los órdenes y éstas se anulan en cero. Lo anterior implica que el polinomio cero es el polinomio de Taylor de cualquier orden de esa función y si esa función fuera representable en serie de potencias, ésta debería ser la serie idénticamente cero, lo cual no es posible pues $f(x) \neq 0$ para toda $x > 0$. La situación anterior se presenta así debido a que los residuos de Taylor no convergen a cero cuando el orden del polinomio crece. $\triangleleft$

Ejemplo 12.13

1. $\exp x = 1 + \sum_{n=1}^{\infty} \frac{1}{n!} x^n$ para $x \in (-\infty, \infty)$

2. $\operatorname{sen} x = \sum_{n=1}^{\infty} \frac{1}{(2n-1)!} x^{2n-1}$ para $x \in (-\infty, \infty)$ $\triangleleft$

Ejercicios y problemas del capítulo

1. Diga cuáles de las series siguientes convergen:

   (a) $\displaystyle\sum_{i=1}^{\infty}(\sqrt{2})^i$ (b) $\displaystyle\sum_{n=1}^{\infty}\frac{1}{n}$ (c) $\displaystyle\sum_{n=1}^{\infty}\frac{(-1)^n}{n^2}$

   (d) $\displaystyle\sum_{n=1}^{\infty}a_n\frac{n!}{(-100)^n}$ (e) $\displaystyle\sum_{n=1}^{\infty}\frac{1}{\sqrt{n}}$

2. Conteste "falso" o "verdadero" a cada uno de los enunciados siguientes.

   (a) Si $a_n > c > 0$ para cada n, entonces la serie $\sum_{n=1}^{\infty} a_n$ diverge.

   (b) Si $a_n < 0$ y $\sum_{n=1}^{\infty} a_n$ converge, entonces $\sum_{n=1}^{\infty} a_n^2$ converge.

   (c) Si $\sum_{n=1}^{\infty} a_n$ converge, entonces $\displaystyle\sum_{n=1}^{\infty}\frac{1}{a_n}$ diverge.

   (d) Si $\lim_{n\to\infty} a_n = 0$, entonces $\sum_{n=1}^{\infty} a_n^2$ converge.

   (e) Si $\sum_{n=1}^{\infty} a_n$ converge, entonces $\sum_{n=1}^{\infty}(-1)^n a_n$ converge.

   (f) Si $\sum_{n=1}^{\infty} a_n$ es absolutamente convergente, entonces $\sum_{n=1}^{\infty}(-1)^n a_n$ es también absolutamente convergente.

   (g) Si $\sum_{n=1}^{\infty} a_n$ converge y $\sum_{n=1}^{\infty}(-1)^n a_n$ converge absolutamente, entonces $\sum_{n=1}^{\infty} a_n$ converge absolutamente.

3. Diga para qué valores de x las series siguientes convergen.

   (a) $\displaystyle\sum_{n=1}^{\infty}\frac{x^n}{\sqrt{n+1}}$ (b) $\displaystyle\sum_{n=1}^{\infty}\frac{1}{n}\left(1+\frac{1}{x}\right)^n$

   (c) $\displaystyle\sum_{n=1}^{\infty}\frac{\operatorname{sen}^n x}{n}$ (d) $\displaystyle\sum_{n=1}^{\infty}\frac{(2x+3)^n}{n^{\frac{1}{3}}4^n}$

4. Determine el intervalo de convergencia y su suma en los puntos de ese intervalo, para cada una de las series de potencias siguientes.

   (a) $\displaystyle 1 - 4x + 16x^2 - 64x^3 + \cdots = \sum_{n=0}^{\infty}(-1)^n(4x)^n$

   (b) $\displaystyle 1\times 3 - 2\times 4x + 3\times 5x^2 - 4\times 6x^3 + \cdots = \sum_{n=1}^{\infty}(-1)^{n-1}n(n+2)x^{n-1}$

   (c) $\displaystyle 2 + 4x^2 + 6x^4 + 8x^6 + 10x^8 + \cdots = \sum_{n=1}^{\infty}2nx^{2(n-1)}$

5. Diga en qué intervalo son válidas las siguiente representaciones para las funciones dadas a la izquierda.

   (a) $\dfrac{1}{1-x} = 1 + x^2 + x^3 + x^4 + \cdots$

   (b) $\ln(1+x) = \displaystyle\sum_{n=1}^{\infty} \frac{(-1)^{n-1}}{n} x^n$

6. Dé un ejemplo de una serie alternante $\sum_{n=1}^{\infty}(-1)^n a_{n+1}$, $a_n > 0$, que sea divergente con $\lim_{n\to\infty} a_n = 0$. (Sugerencia: Considere la serie $1 - \frac{1}{4} + \frac{1}{3} - \frac{1}{16} + \frac{1}{5} - \cdots$.)

7. Demuestre que si $\sum_{n=1}^{\infty} a_n$ es una serie positiva convergente con $a_n < a_{n-1}$ para cada n natural, entonces $\lim_{n\to\infty} na_n = 0$.

8. Considere la serie de potencias $\sum_{n=1}^{\infty} a_n x^n$. Aplicando el criterio del cociente de D'Alembert, pruebe que si $\displaystyle\lim_{n\to\infty}\left|\frac{a_{n+1}}{a_n}\right| = r$, entonces el radio de convergencia de la serie de potencias es igual a $\dfrac{1}{r}$. Análogamente, aplicando el criterio de Cauchy, pruebe que si $\lim_{n\to\infty} \sqrt[n]{|\, a_n \,|} = r$, entonces el radio de convergencia de la serie de potencias es igual a $\dfrac{1}{r}$. Haciendo uso de los criterios anteriores, calcule el radio de convergencia de las series

   (a) $\displaystyle\sum_{n=1}^{\infty} \frac{x^{2n}}{\sqrt{n+2}}$ (b) $\displaystyle\sum_{n=1}^{\infty} \frac{(4x+1)^n}{n^n}$ (c) $\displaystyle\sum_{n=1}^{\infty} \frac{x^n}{2^n \ln n}$

9. Diga por qué la serie $\displaystyle\sum_{n=1}^{\infty} \frac{2}{10^{n^2}}$ converge a un número irracional.

10. A partir de la representación de la función $f(x) = \dfrac{1}{1-x}$ en el intervalo $(-1, 1)$,

$$\frac{1}{1-x} = \sum_{n=0}^{\infty} x^n,$$

encuentre la representación de las funciones siguientes como series de potencias en x y señale el dominio de validez de esa representación.

   (a) $g(x) = \dfrac{1}{2-x}$ (b) $g(x) = \dfrac{1}{(2-x)^2}$

   (c) $g(x) = \dfrac{1-x}{1+x}$ (d) $g(x) = \ln(2-x)$

Bibliografía

[1] Adams, R. A. *Calculus, A complete course,* Addison-Wesley Publishers Limited, 1991.

[2] Apostol, T. *Calculus,* segunda edición, Ed. Reverté, Barcelona, 1968.

[3] Apostol, T. (Editor), *Selected papers on Calculus,* Mathematical Association of America, 1968.

[4] Fraga, R. (Editor), *Resources for Calculus Collection, Vol. 2: Calculus Problems for a new century,* Mathematical Association of America, 1993.

[5] Kuratowski, K. *Introducción al Cálculo,* Editorial Limusa-Wiley, S.A. México, 1970.

[6] Leithold, L. *El cálculo,* séptima edición, Oxford University Press, México, 1998.

[7] Spivak, M. *Cálculo infinitesimal,* vol. 1 y vol. 2, Editorial Reverté S.A., Barcelona, 1998.

[8] Stewart, J. *Cálculo de una variable, Trascendentes tempranas,* cuarta edición, Thomson editores, S.A. de C.V., 2001.

[9] Tellechea, E. *Cálculo diferencial e integral II,* Colección Textos Académicos, No. 23. Editorial UNISON, 2001.

[10] Zill, D. *Cálculo con geometría analítica,* Grupo Editorial Iberoamericano, México, 1987.

Índice