CÁLCULO DIFERENCIAL E INTEGRAL

Javier Pérez González

Departamento de Análisis Matemático

Universidad de Granada

Asignatura: Cálculo
Curso: Primero
Titulación: Ingeniero de Telecomunicación

septiembre 2006

Índice general

Lección 1

Axiomas de los números reales. Desigualdades. Principio de inducción

Introducción

En esta lección quiero que entiendas la importancia de disponer de un “marco de referencia”. Trataré de explicarme. Para empezar, voy a proponerte unos ejercicios muy sencillos.

1. ¿Sabes probar que $0x = 0$? Inténtalo.
2. ¿Qué entiendes por $-x$? ¿Es cierto que $-x$ es negativo?
3. Escribe con palabras lo que afirma la igualdad $(-x)y = -xy$. ¿Sabes probarla?
4. Demuestra que si $x \neq 0$ entonces $x^2 > 0$ (en consecuencia $1 > 0$).
5. ¿Sabes por qué no se puede dividir por 0?
6. Seguro que sabes construir un segmento de longitud $\sqrt{2}$. ¿Y de longitud $\sqrt{3}$?
7. ¿Qué quiere decir que un número no es racional? Demuestra que $\sqrt{2}$ no es racional.

Supongo que hace ya tanto tiempo que conoces estas propiedades de los números que has olvidado cuándo las aprendiste. ¡Y ahora te piden que las *demuestres*! Puedo imaginar tu reacción ¿que demuestre que $0x = 0$?, ¡pero si eso es evidente! ¡siempre me han dicho que es así! ¿cómo se puede demostrar tal cosa?.

Pienso que muchas veces la dificultad de un ejercicio está en que no sabes qué es exactamente lo que se te pide que hagas; no te dan un marco claro de referencia. En estas situaciones lo más frecuente es *“quedarse colgado”* con la *mente en blanco* sin saber qué hacer. Para evitar ese peligro, en este curso vamos a dar un marco de referencia muy claro que va a consistir en unas propiedades de los números (axiomas, si quieres llamarlas así) que vamos a aceptar como punto de partida para nuestro estudio. Esas propiedades, junto con las reglas de inferencia lógica usuales y con definiciones apropiadas nos permitirán demostrar resultados (teoremas) que podremos usar para seguir avanzando. Simplificando un poco, puede decirse que en matemáticas no hay nada más que axiomas y teoremas (bueno, también hay conjeturas, proposiciones indecidibles...). Todo lo que se demuestra es un teorema; por ejemplo $0x = 0$ es un teorema. Ocurre que el nombre *teorema* se reserva para resultados que se consideran realmente importantes y que ha costado esfuerzo llegar a probarlos. Se usan también los términos: *corolario*, *lema*, *proposición* y otros. Pero la estructura de una *teoría matemática elaborada* se resume en un conjunto de axiomas y de teoremas que se deducen de ellos mediante reglas de inferencia lógica.

Es conveniente recordar las propiedades de los números reales porque son ellas las que nos permiten trabajar con desigualdades. Es muy fácil equivocarse al trabajar con desigualdades. Yo creo que en el bachillerato no se le da a este tema la importancia que merece. Fíjate que algunos de los conceptos más importantes del Cálculo se definen mediante desigualdades (por ejemplo, la definición de sucesión convergente o de límite de una función en un punto). Por ello, tan importante como saber realizar cálculos más o menos complicados, es aprender a manejar correctamente desigualdades, y la única manera de hacerlo es con la práctica mediante numerosos ejemplos concretos. Por supuesto, siempre *deben respetarse cuidadosamente las reglas generales que gobiernan las desigualdades entre números* y asegurarse de que se usan correctamente. Aparte de tales reglas no hay otros métodos generales que nos digan cómo tenemos que proceder en cada caso particular.

1.1. Números reales. Propiedades algebraicas y de orden

Como todos sabéis se distinguen distintas clases de números:

Los números naturales 1,2,3,... . El conjunto de todos ellos se representa por $\mathbb{N}$.

Los números enteros ...,-2,-1,0,1,2,... cuyo conjunto se representa por $\mathbb{Z}$.

Los números racionales que son cocientes de la forma p/q donde $p \in \mathbb{Z}, q \in \mathbb{N}$, cuyo conjunto representamos por $\mathbb{Q}$.

También conocéis otros números como $\sqrt{2}$, π, o el número e que no son números racionales y que se llaman, con una expresión no demasiado afortunada, "números irracionales". Pues bien, el conjunto formado por todos los números racionales e irracionales se llama conjunto de los números reales y se representa por $\mathbb{R}$.

Es claro que $\mathbb{N} \subset \mathbb{Z} \subset \mathbb{Q} \subset \mathbb{R}$.

Aunque los números que no son racionales pueden parecer un poco raros, no merece la pena, al menos por ahora, preocuparse por cómo son estos números; sino que lo realmente interesante es aprender a trabajar con ellos. Lo interesante del número $\sqrt{2}$ es que su cuadrado es igual a 2.

Pues bien, una de las cosas más llamativas de los números es que a partir de un pequeño grupo de propiedades pueden deducirse casi todas las demás. Vamos a destacar estas propiedades básicas que, naturalmente, hacen referencia a las dos operaciones fundamentales que se pueden hacer con los números: la suma y el producto. La suma de dos números reales x, y se escribe $x + y$, representándose el producto por xy. Las propiedades básicas a que nos referimos son las siguientes.

P1 [Propiedades asociativas] $(x + y) + z = x + (y + z)$; $(xy)z = x(yz)$ para todos x, y, z en $\mathbb{R}$.

P2 [Propiedades conmutativas] $x+y=y+x$; $xy=yx$ para todos x,y en $\mathbb{R}$.

P3 [Elementos neutros] El 0 y el 1 son tan importantes que enunciamos seguidamente sus propiedades:

$$0+x=x \quad ; \quad 1x=x \quad \text{para todo } x\in\mathbb{R}.$$

P4 [Elementos opuesto e inverso] Para cada número real x hay un número real llamado *opuesto de x*, que representamos por $-x$, tal que $x+(-x)=0$.

Para cada número real x distinto de 0, $x\neq 0$, hay un número real llamado *inverso de x*, que representamos por x^{-1}, tal que $xx^{-1}=1$.

P5 [Propiedad distributiva] $(x+y)z=xz+yz$ para todos x,y,z en $\mathbb{R}$.

Las propiedades anteriores son de tipo algebraico y, aunque son muy sencillas, a partir de ellas pueden *probarse* cosas tan familiares como que $0x=0$, o que $(-x)y=-(xy)$.

Pero los números tienen, además de las propiedades algebraicas, otras propiedades que suelen llamarse *propiedades de orden*. Como todos sabemos, los números suelen representarse como puntos de una recta en la que se fija un origen, el 0, de forma arbitraria. Los números que hay a la derecha de 0, se llaman *positivos* y el conjunto de todos ellos se representa por $\mathbb{R}^+$. Las propiedades básicas del orden son las siguientes.

P6 [Ley de tricotomía] Para cada número real x se verifica que o bien es $x=0$, o bien x es positivo, o bien su opuesto $-x$ es positivo.

P7 [Estabilidad de $\mathbb{R}^+$] La suma y el producto de números positivos es también un número positivo.

Suele escribirse $x-y$ en vez de $x+(-y)$. También, supuesto $y\neq 0$, se escribe x/y o $\frac{x}{y}$ en vez de xy^{-1}. Los opuestos de los números positivos, es decir los elementos del conjunto $\mathbb{R}=\{-x : x\in\mathbb{R}^+\}$, se llaman *números negativos*. Nótese que el 0 no es positivo ni negativo.

Para $x,y\in\mathbb{R}$ escribimos $x<y$ (léase x *es menor que* y) o $y>x$ (léase y *es mayor que* x) para indicar que $y-x\in\mathbb{R}^+$, y escribimos $x\leqslant y$ o $y\geqslant x$ para indicar que $y-x\in\mathbb{R}^+\cup\{0\}$. En adelante usaremos las notaciones: $\mathbb{R}^+_o=\mathbb{R}^+\cup\{0\}$, $\mathbb{R}^-_o=\mathbb{R}^-\cup\{0\}$ y $\mathbb{R}^*=\mathbb{R}\setminus\{0\}$. Nótese que si $x\in\mathbb{R}$ entonces $-x\in\mathbb{R}^+$.

1.1 Teorema (Reglas para trabajar con desigualdades). *Sean x,y,z números reales.*

1. $x\leqslant y$ *e* $y\leqslant z$ *implican que* $x\leqslant z$.
2. $x\leqslant y$ *e* $y\leqslant x$ *implican que* $x=y$.
3. *Se verifica exactamente una de las tres relaciones:* $x<y$, $x=y$, *o* $y<x$.
4. $x<y$ *implica que* $x+z<y+z$.
5. $x<y$, $z>0$ *implican que* $xz<yz$.
6. $x<y$, $z<0$ *implican que* $xz>yz$.

7. *$xy > 0$ si, y sólo si, x e y son los dos positivos o los dos negativos. En consecuencia si $x \neq 0$ es $x^2 > 0$ y, en particular, $1 > 0$.*

8. *$z > 0$ implica que $\dfrac{1}{z} > 0$.*

9. *Supuesto que x e y son los dos positivos o los dos negativos, se verifica que $x < y$ implica que $\dfrac{1}{y} < \dfrac{1}{x}$.*

Valor absoluto

El **valor absoluto** de un número $x \in \mathbb{R}$ se define como el número:

$$|x| = \begin{cases} x & \text{si } x \geqslant 0 \\ -x & \text{si } x \leqslant 0 \end{cases}$$

Para trabajar con valores absolutos es útil recordar que dado $x \in \mathbb{R}_o^+$, representamos por $\sqrt{x}$ al único número ***mayor o igual que cero*** cuyo cuadrado es igual a x. Puesto que, evidentemente, $|x|^2 = x^2$ y, *además*, $|x| \geqslant 0$, se tiene que $|x| = \sqrt{x^2}$.

La siguiente estrategia de procedimiento es de gran utilidad.

Dados $a, b \in \mathbb{R}_o^+$ para probar que $a = b$ es suficiente probar que $a^2 = b^2$ y para probar que $a < b$ es suficiente probar que $a^2 < b^2$.

Geométricamente, $|x|$ representa la distancia de x al origen, 0, en la recta real. De manera más general:

$$|x - y| = \text{distancia entre } x \text{ e } y$$

representa la longitud del segmento de extremos x e y.

1.2 Teorema (**Propiedades del valor absoluto**). *Para $x, y \in \mathbb{R}$ se verifica que:*

1. *$|xy| = |x||y|$;*

2. *$|x| \leqslant y$ es equivalente a $-y \leqslant x \leqslant y$;*

3. *$|x + y| \leqslant |x| + |y|$ y la igualdad se da si, y sólo si, $xy \geqslant 0$* (**desigualdad triangular**)*;*

4. *$\big||x| - |y|\big| \leqslant |x - y|$ y la igualdad se da si, y sólo si, $xy \geqslant 0$.*

1.2. Ejercicios

1. Sabiendo que $a + b > c + d$, $a > b$, $c > d$; ¿se verifica necesariamente alguna de las desigualdades: $a > c$, $a > d$, $b > c$ o $b > d$? Dar una prueba o un contraejemplo en cada caso.

2. Calcula para qué valores de x se verifica que:

i)	$\dfrac{2x-3}{x+2} < \dfrac{1}{3}$	ii)	$\dfrac{1}{x} + \dfrac{1}{1-x} > 0$	iii)	$x^2 - 5x + 9 > x$
iv)	$x^3(x-2)(x+3)^2 < 0$	v)	$x^2 \leqslant x$	vi)	$x^3 \leqslant x$
vii)	$x^2 - (a+b)x + ab < 0$			viii)	$3(x-a)a^2 < x^3 - a^3 < 3(x-a)x^2$

3. Prueba las siguientes desigualdades:

   i) $0 < x+y-xy < 1$ siempre que $0 < x < 1$, $0 < y < 1$.

   ii) $\dfrac{1}{x}+\dfrac{1}{a+b-x} < \dfrac{1}{a}+\dfrac{1}{b}$ siempre que $0 < a < x < b$.

4. Calcula para qué valores de x se verifica que:

i)	$\lvert x-5\rvert < \lvert x+1\rvert$	ii)	$\lvert x-1\rvert\,\lvert x+2\rvert = 3$	iii)	$\lvert x^2-x\rvert > 1$
iv)	$\lvert x-y+z\rvert = \lvert x\rvert - \lvert z-y\rvert$	v)	$\lvert x-1\rvert + \lvert x+1\rvert < 1$	vi)	$\lvert x+y+z\rvert = \lvert x+y\rvert + \lvert z\rvert$
vii)	$\lvert x\rvert - \lvert y\rvert = \lvert x-y\rvert$	viii)	$\lvert x+1\rvert < \lvert x+3\rvert$		

5. Dado que $\dfrac{s}{t} < \dfrac{u}{v} < \dfrac{x}{y}$ donde $t,v,y\in\mathbb{R}^+$, prueba que $\dfrac{s}{t} < \dfrac{s+u+x}{t+v+y} < \dfrac{x}{y}$. Generaliza este resultado.

6. Prueba cada una de las siguientes desigualdades y estudia, en cada caso, cuándo se da la igualdad.

   i) $2xy \leqslant x^2+y^2$.

   ii) $4xy \leqslant (x+y)^2$.

   iii) $x^2+xy+y^2 \geqslant 0$.

   iv) $(a^2+a+1)(b^2+b+1)(c^2+c+1) \geqslant 27abc$ donde $a>0, b>0, c>0$.

   Sugerencia: para probar i) considérese $(x-y)^2$. Las demás desigualdades pueden deducirse de i).

7. Demuestra los teoremas (1.1) y (1.2).

1.3. Principio de inducción matemática

El *Principio de inducción matemática* es un método que se usa para probar que ciertas propiedades matemáticas se verifican para todo número natural. Considera, por ejemplo, la siguiente igualdad en la que $n\in\mathbb{N}$:

$$1^2+2^2+3^2+\cdots+n^2 = \frac{1}{6}n(n+1)(2n+1)$$

Si le damos a n un valor, por ejemplo $n=2$, podemos comprobar fácilmente que la igualdad correspondiente es cierta. Si le damos a n el valor 1000 ya no es tan fácil comprobar esa igualdad y se le damos a n el valor 10^{1000} la cosa ya se pone realmente difícil. Pero nosotros queremos aún más, no nos conformamos con probar que esa igualdad es cierta para unos cuantos miles o millones de valores de n; no, queremos probar que es válida para *todo número natural* n. En estos casos es el *Principio de inducción matemática* el que viene en nuestra ayuda para salvarnos del apuro. Para nosotros el principio de inducción matemática es algo que aceptamos, es decir, puedes considerarlo como un axioma de la teoría que estamos desarrollando (aunque su formulación lo hace "casi evidente").

Principio de inducción matemática. Sea A un conjunto de números naturales, $A\subseteq\mathbb{N}$, y supongamos que:

i) $1 \in A$

ii) Siempre que un número n está en A se verifica que $n+1$ también está en A.

Entonces $A = \mathbb{N}$.

El Principio de Inducción Matemática es la herramienta básica para probar que una cierta propiedad $P(n)$ es verificada por todos los números naturales. Para ello se procede de la siguiente forma:

A) Comprobamos que el número 1 satisface la propiedad, esto es, que $P(1)$ es cierta.

B) Comprobamos que si un número n satisface la propiedad, *entonces* también el número $n+1$ la satisface. Es decir comprobamos que si $P(n)$ es cierta, *entonces* también lo es $P(n+1)$.

Nótese que en B) no se dice que se tenga que probar que $P(n)$ es cierta, sino que hay que *demostrar la implicación lógica* $P(n) \Longrightarrow P(n+1)$.

Si definimos el conjunto $A = \{n \in \mathbb{N} : P(n) \text{ es cierta}\}$, entonces el punto A) nos dice que $1 \in A$, y el punto B) nos dice que siempre que n está en A se verifica que $n+1$ también está en A. Concluimos que $A = \mathbb{N}$, o sea, que $P(n)$ es cierta para todo número natural n.

1.3 Ejemplo. Para cada número natural n, sea $P(n)$ la proposición *si el producto de* n *números positivos es igual a* 1*, entonces su suma es mayor o igual que* n.

Demostraremos por inducción que $P(n)$ es verdadera para todo $n \in \mathbb{N}$. Trivialmente $P(1)$ es verdadera. Supongamos que $P(n)$ es verdadera. Consideremos $n+1$ números positivos no todos iguales a 1 cuyo producto sea igual a 1. En tal caso alguno de dichos números, llamémosle x_1, tiene que ser menor que 1 y otro, al que llamaremos x_2, tiene que ser mayor que 1. Notando $x_3, \cdots, x_{n+1}$ los restantes números se tiene que:

$$(x_1x_2)x_3 \cdots x_{n+1} = 1$$

es decir, $x_1x_2, x_3, \cdots, x_{n+1}$ son n números positivos con producto igual a 1 por lo que:

$$x_1x_2 + x_3 + \cdots + x_{n+1} \geqslant n \qquad (1)$$

y como $0 < (1-x_1)(x_2-1)$, tenemos que:

$$x_1 + x_2 > 1 + x_1x_2 \qquad (2)$$

De (1) y (2) se sigue que:

$$x_1 + x_2 + x_3 + \cdots + x_{n+1} > n+1$$

Hemos probado así que $P(n+1)$ es verdadera. ♦

1.4 Teorema (Desigualdad de las medias). *Cualesquiera sean los números positivos* $a_1, a_2, \cdots, a_n$ *se verifica que:*

$$\sqrt[n]{a_1a_2\cdots a_n} \leqslant \frac{a_1+a_2+\cdots+a_n}{n}$$

y la igualdad se da si, y sólo si, $a_1 = a_2 = \cdots = a_n$.

Demostración. Basta poner $G=\sqrt[n]{a_1a_2\cdots a_n}$ y $x_i=\dfrac{a_i}{G}$, $1\leqslant i\leqslant n$, con lo cual $x_1x_2\cdots x_n=1$ por lo que $\sum_{i=1}^{n}x_i\geqslant n$ es decir $\sum_{i=1}^{n}a_i\geqslant nG$ y se da la igualdad solamente cuando $x_i=1$, para $i=1,2,\dots,n$; es decir, cuando $a_1=a_2=\cdots=a_n$. □

El principio de inducción matemática puede aplicarse en muchas situaciones en las que, a primera vista, no aparecen para nada los números naturales. Por ejemplo, una proposición referente a todos los polinomios podría probarse por inducción sobre el grado del polinomio. Un teorema sobre matrices cuadradas podría probarse por inducción sobre el orden de la matriz.

Probaremos a continuación una útil igualdad algebraica conocida como *fórmula del binomio de Newton.*

Para establecer esta igualdad necesitamos definir los llamados *coeficientes binómicos.* Dados dos números enteros $n\geqslant k\geqslant 0$ se define:

$$\binom{n}{k}=\frac{n!}{k!(n-k)!}\quad \text{donde}\quad n!=\prod_{p=1}^{n}p$$

Es decir, $n!$ es el producto de todos los números naturales menores o iguales que n. Se define también $0!=1$. La igualdad

$$\binom{n}{k-1}+\binom{n}{k}=\binom{n+1}{k}\qquad (1\leqslant k\leqslant n)\tag{1.1}$$

es de comprobación inmediata. A partir de ella se prueba fácilmente, por inducción sobre n, que $\binom{n}{k}$ es un número entero positivo.

1.5 Teorema (Fórmula del binomio de Newton). *Cualesquiera sean los números reales a,b y el número natural n se verifica que:*

$$(a+b)^n=\sum_{k=0}^{n}\binom{n}{k}a^{n-k}b^k.$$

Demostración. Para $n=1$ la igualdad del enunciado es trivialmente verdadera. Supongamos que dicha igualdad se verifica para $n\in\mathbb{N}$. Entonces:

$$\begin{aligned}
(a+b)^{n+1} &= (a+b)(a+b)^n=(a+b)\left[\sum_{k=0}^{n}\binom{n}{k}a^{n-k}b^k\right]\\
&= \sum_{k=0}^{n}\binom{n}{k}a^{n+1-k}b^k+\sum_{k=0}^{n}\binom{n}{k}a^{n-k}b^{k+1}=\\
&= \sum_{k=0}^{n}\binom{n}{k}a^{n+1-k}b^k+\sum_{k=1}^{n+1}\binom{n}{k-1}a^{n+1-k}b^k\\
&= a^{n+1}+b^{n+1}+\sum_{k=1}^{n}\left[\binom{n}{k}+\binom{n}{k-1}\right]a^{n+1-k}b^k=\\
&= \sum_{k=0}^{n+1}\binom{n+1}{k}a^{n+1-k}b^k
\end{aligned}$$

Lo que prueba la validez de la igualdad para $n+1$. En virtud del principio de inducción, concluimos que la igualdad del enunciado es cierta para todo $n \in \mathbb{N}$. $\square$

La inducción matemática es un proceso demostrativo

Considera la expresión $991n^2+1$. Si la evalúas para $n = 1, 2, 3, \ldots, 100000, \ldots$ no creo que consigas obtener valores de n que sean cuadrados perfectos. ¿Debemos concluir que *para todo número natural* n se verifica que $991n^2+1$ no es un cuadrado perfecto? Pues no. Entre los números de la forma $991n^2+1$ hay cuadrados perfectos... ¡el valor mínimo de n para el cual $991n^2+1$ es un cuadrado es $n = 12055735790331359447442538767$!

Con eso te indico que hay que ser precavido: no basta comprobar la veracidad de una expresión para unos cuantos valores de n para concluir que dicha expresión es cierta para todo n. La historia de las matemáticas está llena de este tipo de errores.

1.4. Ejercicios

1. Demuestra que $3^n - 1$ es divisible por 2 para todo $n \in \mathbb{N}$.
2. Demuestra que cualquier conjunto de número naturales, con un número finito de elementos, contiene un número natural máximo.
3. Demuestra que la fórmula
$$2+4+6+\cdots+2n = n^2+n+2$$
cumple con el segundo paso del principio de inducción matemática. Esto es, si la fórmula es verdadera para n, también lo es para $n+1$. Sin embargo, esta fórmula no es válida para $n = 1$. ¿Qué deduces de esto?
4. Teorema del mapa de dos colores: si se traza en una hoja de papel líneas rectas que empiezan y terminan en un borde de la hoja, este mapa puede ser coloreado con sólo dos colores sin que ninguna región adyacente tenga el mismo color.
5. ¿Dónde está el error en el siguiente razonamiento?

   A) En un conjunto formado por *una única* niña, todas los niñas de dicho conjunto tienen el mismo color de ojos.

   B) Supongamos que para todo conjunto formado por n niñas se verifica que todas las niñas del conjunto tienen el mismo color de ojos.

   Consideremos un conjunto formado por $n+1$ niñas. Quitamos una niña del conjunto y nos queda un conjunto formado por n niñas, las cuales, por la hipótesis de inducción, tienen el mismo color de ojos. Ahora devolvemos al conjunto la niña que habíamos sacado y sacamos otra. Volvemos a razonar como antes y deducimos que la niña que habíamos sacado también tiene el mismo color de ojos que las demás n niñas del conjunto. Por tanto las $n+1$ niñas tienen todas ellas igual color de ojos. Como hay una niña con ojos azules, deducimos que todas las niñas tiene ojos azules.

6. Prueba que para todo $n\in\mathbb{N}$ se verifica que:

   a) Todos los números de la forma n^3+5n son múltiplos de 6.

   b) Todos los números de la forma $3^{2n}-1$ son múltiplos de 8.

   c) Todos los números de la forma n^5-n son múltiplos de 5.

   d) 3 *no* divide a n^3-n+1,

   e) $1+\dfrac{1}{2}+\dfrac{1}{3}+\dfrac{1}{4}+\cdots+\dfrac{1}{2^n}\geqslant 1+\dfrac{n}{2}$

   f) $1+\dfrac{1}{1\cdot 3}+\dfrac{1}{3\cdot 5}+\dfrac{1}{5\cdot 7}+\cdots+\dfrac{1}{(2n-1)(2n+1)}=\dfrac{n}{2n+1}$

7. Dados n números positivos $a_1,a_2,\dots,a_n$ prueba que:

   i) $\dfrac{a_1}{a_2}+\dfrac{a_2}{a_3}+\cdots+\dfrac{a_{n-1}}{a_n}+\dfrac{a_n}{a_1}\geqslant n$;

   ii) $\dfrac{n}{1/a_1+1/a_2+\cdots+1/a_n}\leqslant\sqrt[n]{a_1a_2\cdots a_n}$;

   iii) $(a_1+a_2+\cdots+a_n)\left(\dfrac{1}{a_1}+\dfrac{1}{a_2}+\cdots+\dfrac{1}{a_n}\right)\geqslant n^2$.

   ¿Cuándo las desigualdades anteriores son igualdades?

   Sugerencia: Usar la desigualdad de las medias aritmética y geométrica.

8. Utiliza la desigualdad de las medias para probar que:

$$ab^n<\left(\frac{a+nb}{n+1}\right)^{n+1}\qquad\text{siendo } a>0,\ b>0,\ a\neq b,\ \text{ y } n\in\mathbb{N}.$$

   Deduce que para todo número natural n se verifica que:

$$\left(1+\frac{1}{n}\right)^n<\left(1+\frac{1}{n+1}\right)^{n+1},\ \text{ y }\ \left(1+\frac{1}{n+1}\right)^{n+2}<\left(1+\frac{1}{n}\right)^{n+1}$$

9. Sea $q\in\mathbb{N}$ y $a>0$. Prueba que el número $\dfrac{n^q}{(1+a)^n}$ es muy pequeño si n es muy grande.

10. Prueba que entre todos los rectángulos de perímetro dado el de mayor área es el cuadrado.

Lección 2

Funciones reales. Funciones elementales

Introducción

En esta lección vamos a estudiar con algún detalle un concepto teórico importante que es el de continuidad. En este curso se supone que ya tienes un conocimiento intuitivo de las funciones elementales (exponencial, logaritmo natural, trigonométricas), no obstante, si yo doy por sabido algo que tú desconoces harás muy bien en preguntar y yo haré lo posible por despejar tus dudas.

2.1. Funciones reales

Las funciones son las herramientas principales para la descripción matemática de una situación real. Todas las *fórmulas* de la Física no son más que funciones: expresan cómo ciertas magnitudes (por ejemplo el volumen de un gas) dependen de otras (la temperatura y la presión). El concepto de función es tan importante que muchas ramas de la matemática moderna se caracterizan por el tipo de funciones que estudian. No es de extrañar, por ello, que el concepto de función sea de una gran generalidad. Además, se trata de uno de esos conceptos cuyo contenido esencial es fácil de comprender pero difícil de formalizar.

La idea básica de función es la siguiente. Supongamos que tenemos dos conjuntos A y B; una función de A en B es una *regla* que *a cada elemento de* A *asocia un único elemento de* B.

En este curso estamos interesados principalmente en funciones entre conjuntos de números reales, es decir, A y B son subconjuntos de $\mathbb{R}$; con frecuencia $B = \mathbb{R}$. Estas funciones se llaman *funciones reales de una variable real*. En lo que sigue nos referiremos solamente a este tipo de funciones y, si no se especifica otra cosa, se entiende que $B = \mathbb{R}$. Por tanto, para darnos una función nos deben decir, en principio, el subconjunto A de $\mathbb{R}$ donde suponemos que la función está definida y la regla que asigna a cada número de A un único número real. El conjunto A recibe el nombre de *dominio* de la función.

Las funciones se representan por letras. En la práctica las letras más usadas son f, g y h, pero cualquiera otra es también buena. Si f es una función y x es un número que está en su dominio, se representa por $f(x)$ (léase "f de x") el número que f asigna a x, que se llama *imagen de* x *por* f *. Es muy importante en este curso distinguir entre* f *(una función) y* $f(x)$ *(un número real).*

Es importante advertir que las propiedades de una función depende de la regla que la define *y también de su dominio*, por ello *dos funciones que tienen distintos dominios se consideran distintas funciones aunque la regla que las defina sea la misma.*

Criterio de igualdad para funciones.

Dos funciones f y g son iguales cuando tienen igual dominio y $f(x) = g(x)$ para todo x en el dominio común.

Notemos también que aunque estamos acostumbrados a representar a las funciones mediante *fórmulas*, no siempre es posible hacerlo.

El símbolo $f : A \to \mathbb{R}$ se utiliza para indicar que f es una función *cuyo dominio es* A (se supone, como hemos dicho antes, que A es un subconjunto de $\mathbb{R}$)

Veamos unos ejemplos sencillos.

a) Sea $f : \mathbb{R} \to \mathbb{R}$ la función dada por $f(x) = x^2$.

b) Sea $g : \mathbb{R}^+ \to \mathbb{R}$ la función dada por $g(x) = x^2$.

c) Sea $h : \mathbb{R} \to \mathbb{R}$ la función dada por: $$h(x) = \begin{cases} 0, & \text{si } x \in \mathbb{Q} \\ 1, & \text{si } x \in \mathbb{R} \setminus \mathbb{Q} \end{cases}$$

d) Sea $f(x) = \dfrac{x^3 + 5x + 6}{x^2 - 1}$

Según lo antes dicho, las funciones en a) y b) son distintas. Nótese que la función definida en b) es creciente y la definida en a) no lo es.

La función definida en c) es llamada *función de Dirichlet*. Nótese que no es fácil calcular los valores de dicha función porque no siempre se sabe si un número real dado es racional o irracional. ¿Es $e + \pi$ racional? Pese a ello la función está correctamente definida.

En d) no nos dan explícitamente el dominio de f por lo que se entiende que f está definida siempre que $f(x)$ tenga sentido, es decir, siempre que, $x^2 - 1 \neq 0$, esto es, para $x \pm 1$.

El convenio del dominio

Cuando una función se define mediante una fórmula $\boxed{f(x) = \text{fórmula}}$ y el dominio no es explícito, se entiende que el dominio es el mayor conjunto de valores de x para los cuales la expresión $f(x)$ tiene sentido como número real. Éste es el llamado *dominio natural* de la función. Si queremos restringir el dominio natural de alguna manera, entonces debemos decirlo de forma explícita.

Usaremos la notación $dom(f)$ para representar el dominio de una función f (dicho dominio puede ser el natural o un subconjunto del mismo). El conjunto de todos los valores que toma una función, $\{f(x) : x \in dom(f)\}$, suele llamarse *rango o recorrido de* f, o simplemente, *la imagen* de f y lo representaremos por $imagen(f)$.

Ocurre que el dominio natural de muchas funciones es un *intervalo* o la unión de varios intervalos. Recordemos el concepto de intervalo y cuántos tipos diferentes hay.

2.1 Definición. Un conjunto $I \subseteq \mathbb{R}$ se llama un *intervalo* si siempre que dos números están en I todos los números comprendidos entre ellos dos también están en I. El conjunto vacío, Ø, se considera también como un intervalo.

Además de $\mathbb{R}$ y del Ø, hay los siguientes tipos de intervalos[1].

Intervalos que tienen dos puntos extremos a y b (donde $a \leqslant b$ son números reales):

$$
\begin{aligned}
[a,b] &= \{x \in \mathbb{R} : a \leqslant x \leqslant b\}\,; && \text{(intervalo cerrado)}\\
]a,b[&= \{x \in \mathbb{R} : a < x < b\}\,; && \text{(intervalo abierto)}\\
[a,b[&= \{x \in \mathbb{R} : a \leqslant x < b\}\,; && \text{(intervalo abierto a derecha y cerrado a izquierda)}\\
]a,b] &= \{x \in \mathbb{R} : a < x \leqslant b\}\,; && \text{(intervalo abierto a izquierda y cerrado a derecha)}
\end{aligned}
$$

Intervalos que tienen un único punto extremo $c \in \mathbb{R}$ llamado *origen* del intervalo:

$$
\begin{aligned}
]-\infty,c[&= \{x \in \mathbb{R} : x < c\}\;; && \text{(semirrecta abierta a la izquierda)}\\
]-\infty,c] &= \{x \in \mathbb{R} : x \leqslant c\}\;; && \text{(semirrecta cerrada a la izquierda)}\\
]c,+\infty[&= \{x \in \mathbb{R} : x > c\}\;; && \text{(semirrecta abierta a la derecha)}\\
[c,+\infty[&= \{x \in \mathbb{R} : x \geqslant c\}\;; && \text{(semirrecta cerrada a la derecha)}
\end{aligned}
$$

Como es la primera vez que aparecen, hay que decir que los símbolos $+\infty$ (léase: “más infinito”) y $-\infty$ (léase: “menos infinito"); son eso: símbolos. No son números. Cada vez que aparece uno de ellos en una situación determinada hay que recordar cómo se ha definido su significado para dicha situación. A veces, se escribe $\mathbb{R} =]-\infty,+\infty[$.

La mayoría de las funciones que vamos a usar en este curso pertenecen a la clase de las *funciones elementales*. Se llaman así porque pueden obtenerse a partir de ciertos tipos de funciones bien conocidas realizando las operaciones de suma, producto, cociente y composición de funciones.

Dadas dos funciones f y g se define su *función suma* (resp. *producto*) como la función que a cada número $x \in dom(f) \cap dom(g)$ asigna el número real $f(x)+g(x)$ (resp. $f(x)g(x)$). Dicha función se representa con el símbolo $f+g$ (resp. fg). Se define la función cociente de f por g como la función que a cada número $x \in dom(f) \cap dom(g)$ con $g(x) \neq 0$ asigna el número real $\dfrac{f(x)}{g(x)}$. Dicha función se representa con el símbolo $\dfrac{f}{g}$. También podemos multiplicar una función f por un

[1] Este resultado, en apariencia *evidente*, no podríamos *demostrarlo* con las herramientas de que disponemos hasta ahora.

número α para obtener la función αf que asigna a cada $x \in dom(f)$ el número $\alpha f(x)$. De todas formas, el producto de un número por una función puede considerarse como un caso particular del producto de funciones, pues se identifica el número α con la *función constante* que toma como único valor α.

Las propiedades de la suma y el producto de funciones son las que cabe esperar y su demostración es inmediata pues se reducen a las correspondientes propiedades de los números.

Cualesquiera sean las funciones f, g y h se verifica:

Propiedades asociativas. $(f+g)+h = f+(g+h)$; $\quad (fg)h = f(gh)$

Propiedades conmutativas. $f+g = g+f$; $\quad fg = gf$

Propiedad distributiva. $(f+g)h = fh+gh$

Composición de funciones

Supongamos que f y g son funciones verificando que $imagen(f) \subset dom(g)$. En tal caso, la función h dada por $h(x) = g(f(x))$ para todo $x \in dom(f)$ se llama *composición de* g *con* f y se representa por $g \circ f$. La composición de funciones es asociativa, esto es

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

Funciones inyectivas

Se dice que una función f es inyectiva en un conjunto $A \subseteq dom(f)$, si en puntos distintos de A toma valores distintos; es decir, $x,y \in A$ y $x \neq y$, entonces $f(x) \neq f(y)$. Se dice que f es inyectiva cuando es inyectiva en $dom(f)$.

La función inversa de una función inyectiva

Si f es una función inyectiva, puede definirse una nueva función $f^{-1} : imagen(f) \to \mathbb{R}$ que llamaremos *función inversa de* f, que a cada número $y \in imagen(f)$ asigna el único número $x \in dom(f)$ tal que $f(x) = y$. Equivalentemente $f^{-1}(f(x)) = x$ para todo $x \in dom(f)$, y también $f(f^{-1}(y)) = y$ para todo $y \in dom(f^{-1}) = imagen(f)$.

Funciones monótonas

Se dice que una función f es creciente (resp. decreciente) en un conjunto $A \subseteq dom(f)$, si f conserva (resp. invierte) el orden entre puntos de A, es decir, si $x,y \in A$ y $x \leqslant y$, entonces $f(x) \leqslant f(y)$ (resp. $f(x) \geqslant f(y)$). Se dice que f es creciente (resp. decreciente) cuando lo es en todo su dominio ($A = dom(f)$). Se dice que una función es *monótona* para indicar que es creciente o decreciente. Una función monótona e inyectiva se dice que es *estrictamente monótona*, pudiendo ser estrictamente creciente o estrictamente decreciente.

Gráfica de una función

La gráfica de una función f es el conjunto de pares de números $\{(x, f(x)) : x \in dom(f)\}$.

La gráfica de una función pone de manifiesto, a simple vista, muchas de sus propiedades. Para dibujar gráficas de funciones se precisan herramientas de cálculo que estudiaremos más adelante.

2.2. Estudio descriptivo de las funciones elementales[2]

Funciones polinómicas y funciones racionales

Las *funciones polinómicas o polinomios* son las funciones de la forma

$$P(x) = c_0 + c_1 x + c_2 x^2 + \cdots + c_n x^n$$

donde $c_0, c_1, \ldots, c_n$ son números reales llamados *coeficientes* del polinomio; $n \in \mathbb{N}$ es un número natural que, si $c_n \neq 0$, se llama grado del polinomio. Las funciones polinómicas tienen como dominio natural de definición la totalidad de $\mathbb{R}$ aunque con frecuencia nos interesará estudiar una función polinómica en un intervalo.

Mientras que la suma, el producto y la composición de funciones polinómicas es también una función polinómica, el cociente de funciones polinómica da lugar a las llamadas *funciones racionales*. Una función racional es una función de la forma:

$$R(x) = \frac{P(x)}{Q(x)}$$

donde P (el numerador) y Q (el denominador) son polinomios y Q no es el polinomio constante igual a 0. La función R tiene como dominio natural de definición el conjunto $\{x \in \mathbb{R} : Q(x) \neq 0\}$. Observa que las funciones polinómicas son también funciones racionales (con denominador constante 1).

Es inmediato que sumas, productos y cocientes de funciones racionales son también funciones racionales; y la composición de dos funciones racionales es también una función racional.

Raíces de un número

Dados un número real $x > 0$ y un número natural $k \geqslant 2$, hay un único número real *positivo*, $z > 0$, que verifica que $z^k = x$. Dicho número real z se llama la **raiz k-ésima o de orden** k de x y se representa por $\sqrt[k]{x}$ o por $x^{1/k}$.

Además, si $y > 0$, se verifica que:

i) $x < y$ si, y sólo si, $\sqrt[k]{x} < \sqrt[k]{y}$

ii) $\sqrt[k]{xy} = \sqrt[k]{x}\,\sqrt[k]{y}$

[2] El estudio de las funciones elementales que haremos aquí se complementa con el cuaderno de *Mathematica* que está en http://www.ugr.es/local/fjperez/funciones_elementales.nb.

Si $x<0$ y k es *impar* se define $\sqrt[k]{x}=-\sqrt[k]{|x|}$

Potencias racionales

Dados $x>0$, $p\in\mathbb{Z}$ y $q\in\mathbb{N}$, definimos $x^{p/q}=\sqrt[q]{x^p}$. Notemos que $(\sqrt[q]{x})^p=\sqrt[q]{x^p}$ pues

$$\left((\sqrt[q]{x})^p\right)^q=(\sqrt[q]{x})^{pq}=\left((\sqrt[q]{x})^q\right)^p=x^p$$

Naturalmente, si $p/q=m/n$ donde $m\in\mathbb{Z}$ y $n\in\mathbb{N}$, entonces se comprueba fácilmente que $x^{p/q}=x^{m/n}$. En consecuencia, si r es un número racional podemos definir, sin ambigüedad alguna, la potencia x^r por $x^r=x^{p/q}$, donde $p\in\mathbb{Z}$ y $q\in\mathbb{N}$ son tales que $r=p/q$.

Logaritmos

Vamos a hacer un estudio descriptivo de estas funciones. Nos limitaremos a recordar sus definiciones y propiedades básicas, dejando para más adelante un estudio riguroso de las mismas.

Dado un número $a>0$, $a\neq 1$, y un número $x>0$, se define el *logaritmo en base* a *de* x como el único número $y\in\mathbb{R}$ que verifica la igualdad $a^y=x$. El logaritmo en base a de x se representa por el símbolo $\log_a x$. Observa que, por definición, para todo $x>0$ es $a^{\log_a x}=x$.

El dominio de la función $\log_a$ es $\mathbb{R}^+$, y su imagen es $\mathbb{R}$. La función es estrictamente creciente si $a>1$ y estrictamente decreciente si $a<1$. La propiedad básica de los logaritmos es que convierten productos en sumas:

$$\log_a(xy)=\log_a x+\log_a y \qquad (x>0, y>0)$$

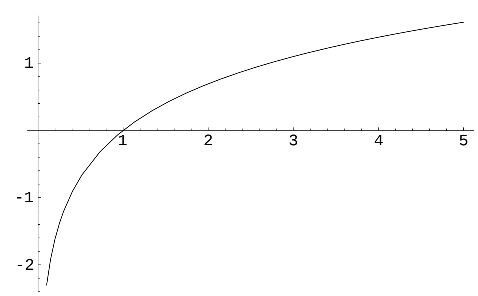


Figura 2.1: Función $\log_a(x)$, $(a>1)$

Los *logaritmos decimales* corresponden a tomar $a=10$ y los *logaritmos naturales*, también llamados *neperianos* (en honor de John Napier 1550-1617), corresponden a tomar como base el número e. El número e es un número irracional que puede aproximarse arbitrariamente por números de la forma $(1+1/n)^n$ para valores grandes de n. Un valor aproximado de e es $2,7182818284$.

En esta asignatura trabajaremos siempre, salvo que explícitamente se indique lo contrario, con la función logaritmo natural, que notaremos log (la notación, cada día más en desuso, “ln”, para dicha función no será usada en este curso).

Teniendo en cuenta que

$$\log_a x=\frac{\log x}{\log a} \qquad (x>0)$$

podemos deducir muy fácilmente las propiedades de la función logaritmo en base a a partir de las propiedades de la función logaritmo natural.

Exponenciales

La función inversa de la función $\log_a$ es la función exponencial de base a, que se representa por $\exp_a$. Por tanto, para cada $x\in\mathbb{R}$, $\exp_a(x)$ es, por definición, el único número positivo cuyo logaritmo en base a es igual a x: $\log_a(\exp_a(x)) = x$. Es fácil comprobar que si $r\in\mathbb{Q}$ entonces $\exp_a(r) = a^r$, por lo que se usa la notación $\exp_a(x) = a^x$.

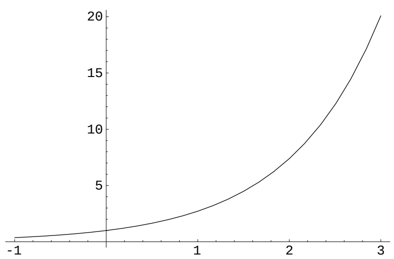


Figura 2.2: Función $\exp_a(x)$, $a > 0$

El dominio de la función $\exp_a$ es $\mathbb{R}$, y su imagen es $\mathbb{R}^+$. La función es estrictamente creciente si $a > 1$ y estrictamente decreciente si $a < 1$. La propiedad básica de $\exp_a$ es que convierten sumas en productos:

$$\exp_a(x+y) = \exp_a(x)\exp_a(y) \qquad (x,y\in\mathbb{R})$$

Dos funciones exponenciales cualesquiera, $\exp_a$ y $\exp_b$, están relacionadas por la igualdad:

$$\exp_b(x) = \exp_a(x\log_a b) \qquad (x\in\mathbb{R})$$

La función exponencial de base e, inversa de la función logaritmo natural, se notará simplemente por exp. Por tanto $\exp(x) = \mathrm{e}^x$. Con ello tenemos que:

$$\boxed{x^y = \mathrm{e}^{y\log x} \qquad (x > 0, y\in\mathbb{R})}$$

La letra e se eligió en honor del gran matemático Leonhard Euler (1707-1783). A primera vista puede parecer que no hay razones particulares para llamar *natural* al número e. Las razones matemáticas de esta elección se verán al estudiar la derivación. Sin embargo, hay muchos procesos de crecimiento que hacen del número e una base exponencial extremadamente útil e interesante. Veamos unos ejemplos.

Interés compuesto. Supongamos que invertimos un capital inicial, P, a una tasa de interés anual r (expresado en tanto por uno), ¿cuánto dinero tendremos cuando hayan pasado k años? Respuesta: depende de cómo se paguen los intereses. En el *interés simple* se paga el total de los intereses al terminar la inversión, por lo que el interés total producido es igual a Prk, y el capital final será igual a $P(1+rk)$.

Sin embargo, lo usual es que se paguen intereses en períodos más cortos de tiempo. Estos intereses se acumulan al capital inicial y producen, a su vez, nuevos intereses. Esto se conoce como *interés compuesto*. Por ejemplo, si el interés se paga n veces al año (trimestralmente $(n = 4)$, mensualmente $(n = 12)$, etcétera) al final del primer período tendremos $P(1+r/n)$, al final del segundo $P(1+r/n)^2$; al final del primer año $P(1+r/n)^n$, al final del k-ésimo año tendremos $P(1+r/n)^{nk}$.

Cuando n es muy grande, el número $(1+r/n)^n$ es aproximadamente igual a e^r. Precisamente, si los interese se acumulan instantáneamente al capital, lo que se conoce como *interés compuesto continuo*, entonces el capital al final del k-ésimo año viene dado por $P\,\mathrm{e}^{rk}$.

Crecimiento demográfico. Llamemos P_0 la población mundial actual, y sea λ la tasa anual de crecimiento expresada en tanto por uno, la cual suponemos que se mantiene constante. Notemos por $P(t)$ la población mundial pasados t años.

Pasado un año, la población será $P(1) \approxeq P_0 + \lambda P_0 = (1+\lambda)P_0$. Utilizamos el signo $\approxeq$ y no el $=$ porque hemos calculado el crecimiento de la población λP_0 como si esta fuese constantemente igual a P_0 en todo el año, lo que no es correcto.

Obtendríamos un resultado más exacto si consideramos el crecimiento de la población mensualmente. Como la tasa de crecimiento mensual es $\lambda/12$, pasado un mes la población será $(1+\frac{\lambda}{12})P_0$, y pasados doce meses $P(1) \approxeq \left(1+\dfrac{\lambda}{12}\right)^{12} P_0$. El cálculo sigue siendo aproximado, pues la población crece *continuamente*. Para obtener una mejor aproximación podríamos considerar días en vez de meses; en general si dividimos el año en n períodos, obtendríamos como aproximación:

$$P(1) \approxeq \left(1+\frac{\lambda}{n}\right)^{n} P_0$$

Cuanto mayor sea n menor será el error que cometemos. *Si hacemos que* n *crezca indefinidamente*, entonces el número $\left(1+\dfrac{\lambda}{n}\right)^{n}$ *se convierte en* e^{λ}, por lo que $P(1) = \mathrm{e}^{\lambda} P_0$. Si el período de tiempo es de t años, entonces $P(t) = P_0\,\mathrm{e}^{\lambda t}$.

Función potencia de exponente real a

Se llama así la función cuyo dominio es $\mathbb{R}^+$ que a cada $x>0$ asigna el número x^a. Puesto que $x^a = \exp(a\log x)$, las propiedades de esta función se deducen con facilidad de las propiedades de las funciones exponencial y logaritmo natural.

Funciones trigonométricas

Vamos a hacer un estudio descriptivo de estas funciones. Nos limitaremos a recordar sus definiciones y propiedades básicas, dejando para más adelante un estudio más riguroso de las mismas.

La palabra *tri-gono-metría* significa “medida de las figuras con tres esquinas”, es decir, de los triángulos. La trigonometría (plana) es el estudio de las relaciones entre las longitudes de los lados de un triángulo (plano) y las medidas de sus ángulos. Por ello, las funciones trigonométricas se definieron originalmente mediante triángulos rectángulos. No obstante, interesa definir dichas funciones usando la *circunferencia unidad*, es decir, la circunferencia centrada en 0 y de radio 1.

El concepto más específico de la trigonometría es el de *medida de un ángulo*. Para medir un ángulo llevamos su vértice al origen y medimos la longitud del arco de la circunferencia unidad que dicho ángulo intercepta, obtenemos así un número que llamamos la medida (absoluta, es decir no orientada) del ángulo en cuestión. Naturalmente, lo primero que hay que hacer para medir cualquier cosa es elegir una unidad de medida. Pues bien, para medir ángulos suelen usarse dos unidades de medida.

Hay una expresión que estamos acostumbrados a usar y cuyo significado conviene precisar. Me refiero a la expresión: “una circunferencia de radio r”. Cuando empleamos dicha expresión se sobreentiende que el radio r de la circunferencia es un número expresado en alguna unidad de medida de longitudes. *Es decir, la expresión "una circunferencia de radio r" presupone que hemos fijado una unidad de medida con la cual hemos medido r.*

Medida de ángulos en grados

Supongamos que tenemos una circunferencia de radio r. Para medir ángulos en grados sobre dicha circunferencia lo que hacemos es tomar como unidad de medida un arco cuya longitud sea igual a la longitud total de esa circunferencia $(2\pi r)$ dividida por 360. Un ángulo de un grado es el que intercepta en una circunferencia de radio r un arco cuya longitud es igual a $\frac{2\pi r}{360}$.

Medida de ángulos en radianes

Supongamos que tenemos una circunferencia de radio r. Para medir ángulos en radianes sobre dicha circunferencia lo que hacemos es tomar como unidad de medida un arco cuya longitud sea igual a la del radio. Un ángulo de un radián es el que intercepta en una circunferencia de radio r un arco cuya longitud es igual a r.

Las palabras "grado" y "radián" se usan tanto para referirse a los respectivos ángulos como a las medidas de sus arcos. Es así como debes interpretar la expresión "la longitud total de la circunferencia es 360 grados y también es igual a 2π radianes". Sería más exacto decir: "la longitud total de la circunferencia es 360 veces *la longitud de un arco de un grado* y también es igual a 2π veces *la longitud de un arco de un radián*". Evidentemente, la longitud de un arco de un radián es igual al radio de la circunferencia.

La relación entre grados y radianes viene dada por:

$$360 \text{ grados} = 2\pi \text{ radianes}$$

No hay que olvidar que *grados* y *radianes* no son otra cosa que *unidades de medida* de longitudes, al igual que lo son el metro y el centímetro. En la navegación y en la astronomía los ángulos se miden en grados, pero en Cálculo es preferible medirlos en radianes porque se simplifican las cuentas. Por ejemplo, la longitud de un arco de circunferencia se obtiene multiplicando la longitud del radio de dicha circunferencia por la medida *en radianes* del ángulo que corresponde a dicho arco.

Observa que la ventaja de medir arcos en radianes es que, en tal caso, la misma unidad con la que medimos el radio nos sirve para medir arcos. Por ejemplo, si el radio es 1 centímetro el radián también mide 1 centímetro; mientras que la medida de un grado en centímetros sería $2\pi/360 \simeq 0,0174533$.

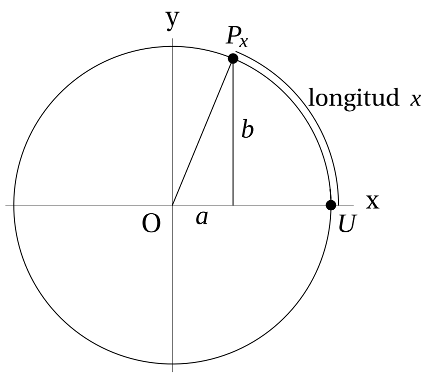
Convenio de los ángulos: usar radianes

De ahora en adelante, a menos que se establezca explícitamente otra unidad, supondremos que todos los ángulos están medidos en radianes.

Funciones seno y coseno

Hay dos funciones que suelen confundirse: el seno de un ángulo y el seno de un número. En geometría se habla del *seno de un ángulo* y en Cálculo usamos la expresión $\operatorname{sen}(\sqrt{2})$ para referirnos al *seno del número* $\sqrt{2}$. ¿Qué relación hay entre uno y otro? Antes que nada hay que decir que tanto el seno de un ángulo como el seno de un número *son números*, pero mientras que el seno de un ángulo tiene una sencilla definición geométrica, no es evidente, a priori, cómo se puede definir el seno de un número.

La idea consiste en asociar a cada número un (único) ángulo y definir el seno del número como el seno del ángulo que le corresponde. Es evidente que a cada número $x \geqslant 0$ le podemos asignar de manera única un ángulo "enrollando" el segmento $[0,x]$ sobre la circunferencia unidad, *en sentido contrario a las agujas del reloj*, de forma que el origen de dicho segmento coincida con el punto $U=(1,0)$ de la circunferencia. Obtenemos así un punto P_x de la circunferencia unidad. Pues bien, si las coordenadas de P_x son (a,b), se define:

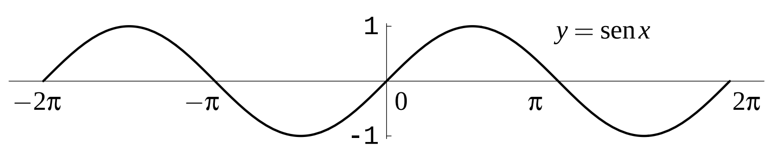


$$\operatorname{sen} x = \text{seno del ángulo}(\widehat{OUP_x}) = b$$
$$\cos x = \text{coseno del ángulo}(\widehat{OUP_x}) = a$$

Al ser igual a 2π la longitud de la circunferencia unidad, es claro que $P_{x+2\pi} = P_x$, por lo que $\operatorname{sen}(x) = \operatorname{sen}(x+2\pi)$ y $\cos(x) = \cos(x+2\pi)$. Observa también que si $0 \leqslant x < 2\pi$, entonces la *medida en radianes* del ángulo $\widehat{OUP_x}$ es igual a x, es decir:

$$\operatorname{sen}(x) = \text{ seno del ángulo de } x \text{ radianes } (0 \leqslant x < 2\pi)$$

Si $x < 0$ podemos proceder con el segmento $[x,0]$ de forma análoga a la anterior, con la diferencia de que ahora enrollamos dicho segmento sobre la circunferencia unidad *en el sentido de las agujas del reloj*, de forma que su extremo 0 coincida con el punto $U=(1,0)$ de la circunferencia. Obtenemos así un punto $P_x=(c,d)$ de la circunferencia unidad y se define, igual que antes $\operatorname{sen}(x)=d$, $\cos(x)=c$. Es fácil ver que si $P_x=(c,d)$, entonces $P_{-x}=(c,-d)$. Resulta así que $\operatorname{sen}(x) = -\operatorname{sen}(-x)$ y $\cos(x)=\cos(-x)$.



Observación

Podemos definir la función *seno en grados* sin más que interpretar que x es la medida en grados del ángulo que le corresponde. El hecho de que se use la misma notación para ambas funciones es la causa de muchos errores. Si notamos $\operatorname{sen}^{o}(x)$ el valor del seno del ángulo cuya media es x grados, y notamos $\operatorname{sen}^{r}(x)$ el valor del seno del ángulo cuya media es x radianes (es decir, la función que hemos definido antes); la relación entre ambas funciones viene dada por:

$$\operatorname{sen}^{o}(x)=\operatorname{sen}^{r}\frac{2\pi x}{360}=\operatorname{sen}^{r}\frac{\pi x}{180}$$

Es frecuente que $\operatorname{sen}^{o}(x)$ se escriba como $\operatorname{sen} x^{o}$. Por ejemplo $\operatorname{sen}(45^{o})$. A esta mala notación se deben las dudas que a veces surgen sobre el significado de $\operatorname{sen} x$ y que llevan a preguntar: "¿está x en grados o en radianes?", cuando lo que realmente debería preguntarse es "¿se trata de $\operatorname{sen}^{o}(x)$ o de $\operatorname{sen}^{r}(x)$?"; porque, en ambos casos, x es tan sólo un número al que no hay por qué ponerle ninguna etiqueta.

Insistimos, una última vez: en este curso de Cálculo el número $\operatorname{sen} x$ significará siempre $\operatorname{sen}^{r} x$. Por tanto $\operatorname{sen}(\pi/4)\neq\operatorname{sen}(45)$ (pero $\operatorname{sen}(\pi/4)=\operatorname{sen}^{o}(45)$).

Propiedades de las funciones seno y coseno

Las funciones seno y coseno son funciones reales cuyo dominio es todo $\mathbb{R}$. Las identidades básicas que dichas funciones verifican son:

$$\operatorname{sen}^2 x+\cos^2 x=1 \qquad (x\in\mathbb{R})$$

Como se ha dicho antes, las funciones seno y coseno son periódicas de período 2π:

$$\operatorname{sen}(x+2\pi)=\operatorname{sen} x,\quad \cos(x+2\pi)=\cos x\quad (x\in\mathbb{R})$$

La función seno es impar y la función coseno es par:

$$\operatorname{sen}(-x)=-\operatorname{sen} x,\quad \cos(-x)=\cos x\quad (x\in\mathbb{R})$$

Todas las propiedades anteriores se deducen fácilmente de las definiciones dadas. Las siguientes igualdades, conocidas como *fórmulas de adición*, se probarán más adelante:

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

La función seno se anula en los múltiplos enteros de π, es decir, en los puntos de la forma $k\pi$ donde k es un entero cualquiera. La función coseno se anula en los puntos de la forma $k\pi+\pi/2$ donde k es un entero cualquiera.

Las funciones **tangente** y **secante**, que se representan por tg y sec son las funciones definidas en el conjunto $\mathbb{R}\setminus\{k\pi+\pi/2 : k\in\mathbb{Z}\}=\{x\in\mathbb{R} : \cos x\neq 0\}$, por:

$$\operatorname{tg} x=\frac{\operatorname{sen} x}{\cos x},\quad \sec x=\frac{1}{\cos x}$$

Las funciones **cotangente** y **cosecante**, que se representan por cotg y csc son las funciones definidas en el conjunto $\mathbb{R}\setminus\{k\pi : k\in\mathbb{Z}\}=\{x\in\mathbb{R} : \operatorname{sen} x\neq 0\}$, por:

$$\operatorname{cotg} x=\frac{\cos x}{\operatorname{sen} x},\quad \csc x=\frac{1}{\operatorname{sen} x}$$

Las propiedades de estas funciones se deducen con facilidad de las propiedades del seno y del coseno. Por ejemplo, $\operatorname{tg}(x)=\operatorname{tg}(x+\pi)$; es decir, la función tangente es periódica de período π.

Las funciones arcoseno, arcocoseno y arcotangente

Lo primero que hay que decir es que ninguna de las funciones “seno”, “coseno”, “tangente”, es inyectiva pues todas ellas son periódicas y, por tanto, toman cada uno de sus valores en infinitos puntos; en consecuencia, ninguna de ellas tiene inversa. Por tanto, no debe decirse que las funciones *arcoseno*, *arcocoseno*, *arcotangente* sean las funciones inversas del seno, del coseno o de la tangente: eso no es cierto. Hecha esta observación imprescindible, pasemos a definir dichas funciones.

La función seno es estrictamente creciente en el intervalo $[-\pi/2,\pi/2]$ y en dicho intervalo toma todos los valores comprendidos entre -1 y 1, $\operatorname{sen}([-\pi/2,\pi/2])=[-1,1]$. En consecuencia, dado un número $x\in[-1,1]$ hay un único número $y\in[-\pi/2,\pi/2]$ tal que $\operatorname{sen} y=x$; dicho número y se representa por $\operatorname{arc\,sen} x$ y se llama el *arcoseno de* x. Es decir, el arcoseno es la función $\operatorname{arc\,sen}\colon[-1,1]\to\mathbb{R}$ definida por $\operatorname{sen}(\operatorname{arc\,sen} x)=x$ y $-\frac{\pi}{2}\leqslant\operatorname{arc\,sen} x\leqslant\frac{\pi}{2}$. Observa que la igualdad $\operatorname{arc\,sen}(\operatorname{sen} x)=x$, es cierta si, y sólo si, $-\pi/2\leqslant x\leqslant\pi/2$.

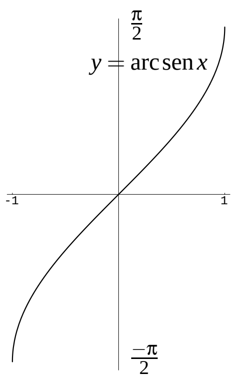


Figura 2.3: Función $\operatorname{arc\,sen} x$

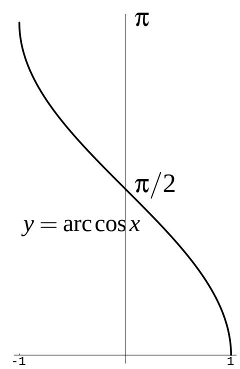


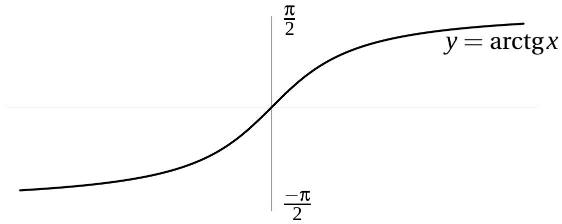
Figura 2.4: Función $\operatorname{arc\,cos} x$

La función coseno es estrictamente decreciente en el intervalo $[0,\pi]$ y en dicho intervalo toma todos los valores comprendidos entre -1 y 1. Por tanto, dado un número $x\in[-1,1]$, hay un único número $y\in[0,\pi]$ tal que $\cos y=x$; dicho número y se representa por $\operatorname{arc\,cos} x$ y se llama *arcocoseno de* x. Es decir, arcocoseno es la función $\operatorname{arc\,cos}\colon[-1,1]\to\mathbb{R}$ dada por $\cos(\operatorname{arc\,cos} x)=x$ y $0\leqslant\operatorname{arc\,cos} x\leqslant\pi$. Observa que la igualdad $\operatorname{arc\,cos}(\cos x)=x$, es cierta si, y sólo si, $0\leqslant x\leqslant\pi$.

La función tangente es estrictamente creciente en el intervalo $]-\pi/2,\pi/2[$ y en dicho intervalo toma todos los valores reales, $\operatorname{tg}(]-\pi/2,\pi/2[)=\mathbb{R}$. En consecuencia, dado un número $x\in\mathbb{R}$, hay un único número $y\in]-\pi/2,\pi/2[$ tal que $\operatorname{tg} y=x$; dicho número y se representa por $\operatorname{arc\,tg} x$ y se llama el *arcotangente de* x. Es decir, el arcotangente es la función:

$$\operatorname{arc\,tg}\colon\mathbb{R}\to\mathbb{R}\ \text{ definida por: }\ \operatorname{tg}(\operatorname{arc\,tg} x)=x,\quad -\frac{\pi}{2}<\operatorname{arc\,tg} x<\frac{\pi}{2}.$$

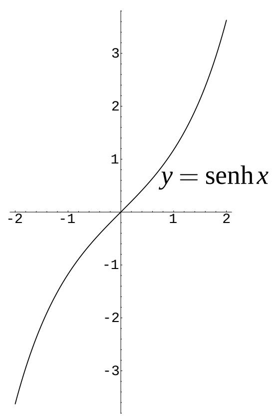
Observa que la igualdad $\operatorname{arc\,tg}(\operatorname{tg} x)=x$, es cierta si, y sólo si, $-\pi/2<x<\pi/2$.

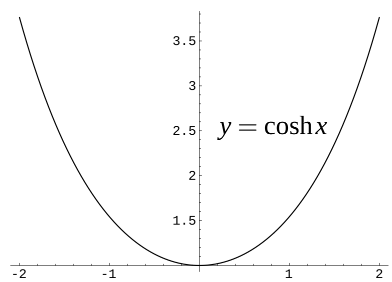


Las funciones hiperbólicas

Hay algunas combinaciones de las funciones $\exp(x)$ y $\exp(-x)$ que aparecen con tanta frecuencia que se les da nombre propio. Ellas son las funciones *seno hiperbólico*, representada por senh, y *coseno hiperbólico*, representada por cosh, y están definidas para todo $x\in\mathbb{R}$ por:

$$\cosh x = \frac{\mathrm{e}^x + \mathrm{e}^{-x}}{2}, \qquad \operatorname{senh} x = \frac{\mathrm{e}^x - \mathrm{e}^{-x}}{2}$$





Propiedades de las funciones seno hiperbólico y coseno hiperbólico

Las funciones seno hiperbólico y coseno hiperbólico son funciones reales cuyo dominio es todo $\mathbb{R}$. La identidad básica que dichas funciones verifican es:

$$\cosh^2 x - \operatorname{senh}^2 x = 1 \qquad (x\in\mathbb{R})$$

La función seno hiperbólico es impar y la función coseno hiperbólico es par:

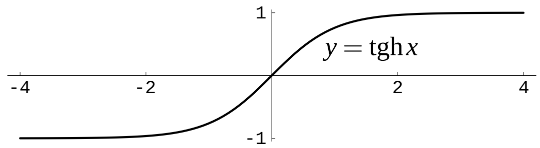
$$\operatorname{senh}(-x) = -\operatorname{senh} x, \quad \cosh(-x) = \cosh x \quad (x\in\mathbb{R})$$

La función seno hiperbólico es estrictamente creciente en $\mathbb{R}$. La función coseno hiperbólico es estrictamente creciente en $\mathbb{R}_o^+$.

Todas las propiedades anteriores se deducen fácilmente de las definiciones dadas.

La función **tangente hiperbólica** que se representa por tgh es la función definida para todo $x \in \mathbb{R}$ por:

$$\operatorname{tgh} x = \frac{\operatorname{senh} x}{\cosh x} = \frac{e^x - e^{-x}}{e^x + e^{-x}}$$

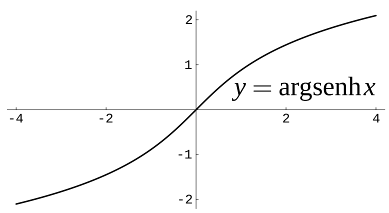


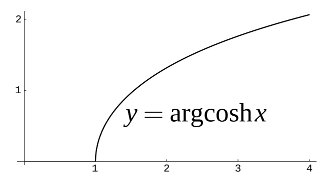
De forma análoga se definen las funciones cotangente, secante y cosecante hiperbólicas.

Las funciones hiperbólicas inversas

La función seno hiperbólico es una biyección de $\mathbb{R}$ sobre $\mathbb{R}$ cuya inversa, representada por, argsenh, (léase **argumento seno hiperbólico**) viene dada por:

$$\operatorname{argsenh} x = \log(x + \sqrt{x^2+1}) \quad (x \in \mathbb{R})$$



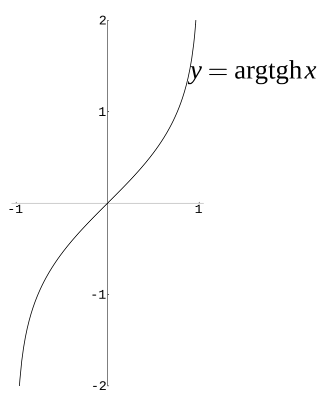


La función tangente hiperbólica es una biyección de $\mathbb{R}$ sobre el intervalo $]-1,1[$ cuya inversa, representada por, argtgh, (léase **argumento tangente hiperbólica**) es la función definida en el intervalo $]-1,1[$ por:

$$\operatorname{argtgh} x = \frac{1}{2}\log\left(\frac{1+x}{1-x}\right) \quad (-1 < x < 1)$$

La función coseno hiperbólico es inyectiva en $\mathbb{R}_o^+$ y su imagen es la semirrecta $[1,+\infty[$. La función, definida en $[1,+\infty[$, que a cada número $x \geqslant 1$ asigna el único número $y > 0$ tal que $\cosh y = x$, se llama **argumento coseno hiperbólico**, se representa por, argcosh, y viene dada por:

$$\operatorname{argcosh} x = \log(x + \sqrt{x^2-1}) \quad (x \geqslant 1)$$



La razón de por qué estas funciones se llaman hiperbólicas es que, al igual que los puntos de la circunferencia unidad pueden representarse en la forma $(\cos t, \operatorname{sen} t)$, los puntos en la rama derecha de la hipérbola unitaria $x^2 - y^2 = 1$ pueden representarse como $(\cosh t, \operatorname{senh} t)$.

Naturalmente, la importancia de las funciones trigonométricas procede de que multitud de fenómenos naturales son de naturaleza ondulatoria. Todos sabéis lo que es un electrocardiograma; pues bien, la gráfica que aparece en ese informe clínico no es más que superposiciones de gráficas de senos y cosenos.

Las funciones hiperbólicas, por su parte, también sirven para describir el movimiento de ondas en sólidos elásticos, o la forma que adoptan los cables eléctricos colgantes. Hay una hermosa curva llamada *catenaria* cuya ecuación es de la forma $y = a\cosh(x/a)$ (donde se entiende que a es una constante). La catenaria es la forma que adopta una cadena perfectamente flexible suspendida de sus extremos y bajo la acción de la gravedad.

2.3. Ejercicios

1. Compara $a^{\log b}$ con $b^{\log a}$.
2. Resuelve $\dfrac{1}{\log_x(a)} = \dfrac{1}{\log_b(a)} + \dfrac{1}{\log_c(a)} + \dfrac{1}{\log_d(a)}$
3. ¿Es correcto escribir $\log(x-1)(x-2) = \log(x-1) + \log(x-2)$?
4. Prueba que $\log(x + \sqrt{1+x^2}) + \log(\sqrt{1+x^2} - x) = 0$.
5. Resuelve $x^{\sqrt{x}} = (\sqrt{x})^x$.
6. Simplifica las expresiones $a^{\log(\log a)/\log a}$, $\log_a(\log_a(a^{a^x}))$.
7. Resuelve el sistema: $7(\log_y x + \log_x y) = 50$, $xy = 256$. Se supondrá que $x > y > 1$.
8. Indica cuál de los dos números $1{,}234{,}567^{6{,}334{,}568}$ y $1{,}234{,}568^{6{,}334{,}567}$ es el mayor.
9. Calcula los valores de x para los que se verifica la igualdad:
$$\log_x(10) + 2\log_{10x}(10) + \log_{190x}(70) = 0$$
10. Sea $f : \mathbb{R}^+ \to \mathbb{R}$ una función que verifica las propiedades:
    1. $f(xy) = f(x) + f(y)$ para todos x, y en $\mathbb{R}^+$;
    2. $f(x) > 0$ para todo $x > 1$;
    3. $f(e) = 1$.

    Demuestra que $f(x) = \log(x)$ para todo $x \in \mathbb{R}^+$.

    Sugerencias: a) Prueba primero que f es creciente y que $f(\mathrm{e}^r) = r$ para todo $r \in \mathbb{Q}$.

    b) Sea $\varphi(x) = f(\exp(x))$. Justifica que φ es estrictamente creciente. Supón que hay algún número a tal que $\varphi(a) \neq a$ y deduce una contradicción (utiliza que entre dos números reales cualesquiera siempre hay algún número racional).

11. Prueba las igualdades siguientes.

$$\cos(\operatorname{arc\,tg} x) = \frac{1}{\sqrt{1+x^2}} \qquad \operatorname{sen}(\operatorname{arc\,tg} x) = \frac{x}{\sqrt{1+x^2}}$$

$$\tan(\operatorname{arc\,sen} x) = \frac{x}{\sqrt{1-x^2}} \quad \forall x \in]-1,1[, \quad \operatorname{arc\,cos} x + \operatorname{arc\,sen} x = \frac{\pi}{2} \quad \forall x \in [-1,1]$$

12. Sean $a,b\in\mathbb{R}$ tales que $a^2+b^2=1$, $a \neq -1$. Definamos $\vartheta = 2\operatorname{arc\,tg}\dfrac{b}{a+1}$. Prueba que $\cos\vartheta = a$, $\operatorname{sen}\vartheta = b$.

13. Prueba por inducción la siguiente igualdad.

$$\operatorname{sen}\frac{x}{2}\,(\operatorname{sen} x + \operatorname{sen} 2x + \cdots + \operatorname{sen} nx) = \operatorname{sen}\frac{nx}{2}\operatorname{sen}\frac{n+1}{2}x$$

14. Prueba que $\operatorname{tg}(x+y) = \dfrac{\operatorname{tg} x + \operatorname{tg} y}{1 - \operatorname{tg} x \operatorname{tg} y}$. ¿Qué excepciones hay que hacer?.

15. Indica para qué valores de x e y se verifica la igualdad $\operatorname{arc\,tg} x + \operatorname{arc\,tg} y = \operatorname{arc\,tg}\dfrac{x+y}{1-xy}$.

Lección 3

Números complejos. Exponencial compleja

Introducción

Los números complejos son una herramienta básica de cálculo. Son especialmente útiles para trabajar con funciones sinusoidales, y por eso se hace uso constante de ellos siempre que representamos una señal por medio de dichas funciones, y no hay que olvidar que ése es el propósito básico de los *"métodos de Fourier"*. La *Transformada de Fourier Discreta*, una herramienta fundamental en el tratamiento digital de señales, toma valores complejos. Las *transformadas de Fourier y de Laplace* son funciones complejas. La *transformada* z, al igual que otras transformadas de uso frecuente, se define como una serie de números complejos. La función exponencial compleja desempeña un papel fundamental en el estudio de los sistemas LTI (sistemas lineales invariantes en el tiempo) y también en la teoría de las ecuaciones diferenciales lineales.

3.1. Operaciones básicas con números complejos

3.1 Definición. Consideremos en el conjunto $\mathbb{R}^2$ las operaciones de adición y producto definidas por

$$(a,b)+(c,d)=(a+c,b+d)$$
$$(a,b)(c,d)=(ac-bd,ad+bc)$$

Es muy fácil comprobar las propiedades asociativa, conmutativa y distributiva de las operaciones así definidas. El elemento neutro de la suma es $(0,0)$ y $(1,0)$ es la unidad del producto. Además, $(-a,-b)$ es el opuesto de (a,b), y todo $(a,b)\neq(0,0)$ tiene inverso

$$(a,b)\left(\frac{a}{a^2+b^2},\frac{-b}{a^2+b^2}\right)=(1,0)$$

Todas estas propiedades se resumen diciendo que $(\mathbb{R}^2,+,\cdot)$ (léase "el conjunto $\mathbb{R}^2$ con las operaciones de adición y producto") es un *cuerpo*. Dicho cuerpo se representa simbólicamente por $\mathbb{C}$ y sus elementos se llaman **números complejos**.

Comentarios a la definición

A los elementos de $\mathbb{R}^2$ se les llama unas veces *pares ordenados de números reales*, otras *vectores* o *puntos* y también *números complejos*. La razón de esto es que en $\mathbb{R}^2$ conviven varias estructuras cada una con su terminología propia. Por eso a los elementos de $\mathbb{R}^2$ se les llama *vectores* si se está considerando la estructura de espacio vectorial, *puntos* si fijamos la atención en la estructura topológica o afín, *pares ordenados* cuando estamos pensando en $\mathbb{R}^2$ como conjunto sin ninguna estructura particular y *números complejos* cuando se considera la estructura de cuerpo antes definida. Ocurre que estos términos se usan a veces en un mismo párrafo lo que puede resultar confuso. La regla que debes tener siempre presente es que todo **concepto matemático** tiene sentido propio dentro de una determinada **estructura matemática**. Por ello, a un elemento de $\mathbb{R}^2$ se le llama número complejo cuando se va a usar el producto antes definido que es lo que en realidad distingue a los números complejos de los vectores de $\mathbb{R}^2$.

Forma cartesiana de un número complejo

El símbolo usual (a,b) para representar pares ordenados no es conveniente para representar el número complejo (a,b). Para convencerte calcula $(1,-1)^4$. Representaremos los números complejos con un simbolismo más apropiado. Para ello hacemos la identificación $(a,0)=a$ y el número complejo $(0,1)$ lo representaremos por i. Con ello tenemos que

$$i^2=(0,1)(0,1)=(-1,0)=-1$$

Ahora podemos escribir

$$(a,b)=(a,0)+(0,b)=(a,0)+(b,0)(0,1)=a+bi$$

Se dice que a es la **parte real** y b es la **parte imaginaria** del número complejo $z=a+ib$ y escribimos $a=\operatorname{Re}(z)$, $b=\operatorname{Im}(z)$. El producto ahora es muy fácil de recordar pues

$$(a+ib)(c+id)=ac+i^2bd+i(ad+bc)=ac-bd+i(ad+bc)$$

Comentarios a la definición usual $i=\sqrt{-1}$

Acabamos de ver que $i^2=-1$ pero eso no nos permite escribir así, sin más ni más, que $i=\sqrt{-1}$. Fíjate lo que ocurre si ponemos $i=\sqrt{-1}$ y manejamos ese símbolo con las reglas a las que estamos acostumbrados:

$$i^2=-1=i\,i=\sqrt{-1}\sqrt{-1}=\sqrt{(-1)(-1)}=\sqrt{1}=1$$

Luego $1=-1$. Por tanto, las matemáticas son contradictorias y aquí hemos acabado.

Naturalmente, el error, procede de que estamos haciendo disparates. Fíjate que en la expresión $\sqrt{-1}$ *no puedes interpretar que* -1 *es el número real* -1 (porque, como sabes, los números reales negativos no tienen raíz cuadrada real), *sino que tienes que interpretar* -1 *como el número complejo* -1 (espero que ya tengas clara la diferencia). Resulta así que estamos usando raíces de números complejos *sin haberlas definido y dando por supuesto que dichas raíces verifican las mismas propiedades que las de los números reales positivos.*

Antes de escribir $\sqrt{-1}$ hay que definir qué significa $\sqrt{z}$ para $z\in\mathbb{C}$. Cuando lo hagamos veremos ¡sorpresa! que *la igualdad* $\sqrt{z}\,\sqrt{w}=\sqrt{zw}$, válida cuando $z,w\in\mathbb{R}^+$, *no es cierta en general cuando* $z,w\in\mathbb{C}$.

Todavía más disparatado es *definir* $i=\sqrt{-1}$ sin ni siquiera haber definido antes los números complejos. Sin embargo, y aunque parezca mentira, en muchos textos se define (porque sí, sin más explicaciones) $i=\sqrt{-1}$ y a continuación se dice que los números de la forma $a+ib$ son los números complejos. No es de extrañar que luego resulte que $1=-1$.

No hay un orden en $\mathbb{C}$ compatible con la estructura algebraica

Al ampliar $\mathbb{R}$ a $\mathbb{C}$ ganamos mucho pero también perdemos algo. Te recuerdo que $\mathbb{R}$ tiene dos estructuras: la algebraica y la de orden. Ambas estructuras están armoniosamente relacionadas. Pues bien, en $\mathbb{C}$ no hay nada parecido. Podemos definir relaciones de orden en $\mathbb{C}$, pero no hay ninguna de ellas que sea compatible con la estructura algebraica. Es decir, es imposible definir un concepto de número complejo positivo de forma que la suma y el producto de complejos positivos sea positivo. Por ello no se define en $\mathbb{C}$ ningún orden. Así que ya sabes: ¡nunca escribas desigualdades entre números complejos! Naturalmente, puedes escribir desigualdades entre las partes reales o imaginarias de números complejos, porque tanto la parte real como la parte imaginaria de un número complejo son números reales.

3.1.1. Representación gráfica. Complejo conjugado y módulo

Es usual interpretar el número complejo $x+iy$ como el vector del plano (x,y) y, en ese sentido, se habla del *plano complejo*. El eje horizontal recibe el nombre de *eje real*, y el eje vertical recibe el nombre de *eje imaginario*. Si $z=x+iy$ es un número complejo (con x e y reales), en-

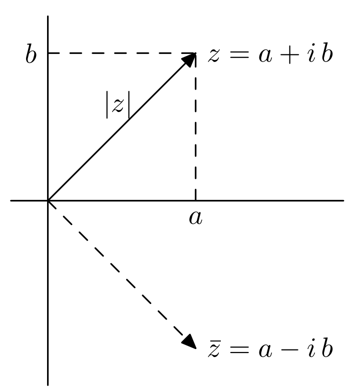


Figura 3.1: Representación de un número complejo

tonces el **conjugado** de z se define como:

$$\overline{z}=x-iy$$

y el **módulo** o **valor absoluto** de z, se define como:

$$|z|=\sqrt{x^2+y^2}$$

Geométricamente, $\bar{z}$ es la reflexión de z respecto al eje real, mientras que $|z|$ es la distancia euclídea del punto (x,y) a $(0,0)$ o, también, la longitud o **norma euclídea** del vector (x,y) (ver figura 3.1). La **distancia** entre dos números complejos z y w se define como $|z-w|$.

La representación gráfica de la suma es conocida. Dos números complejos $z=a+ib$ y $w=c+id$ determinan un paralelogramo cuya diagonal (ver figura 3.2) es $z+w$. Se comprueba fácil-

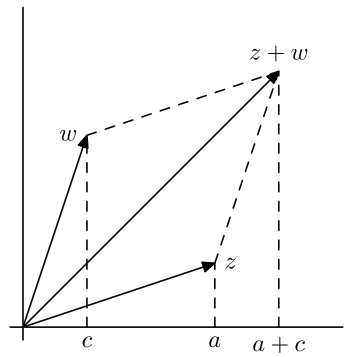


Figura 3.2: Suma de números complejos

mente que si z y w son números complejos se verifica que $\bar{\bar{z}}=z$, $\overline{z+w}=\bar{z}+\bar{w}$ y $\overline{zw}=\bar{z}\bar{w}$.

La igualdad $|z|^2=z\bar{z}$ que se deduce directamente de la definición de módulo de un número complejo, permite probar con facilidad que para todos $z,w\in\mathbb{C}$ es

$$\textbf{a)}\quad |zw|=|z|\,|w| \quad \text{y} \quad \textbf{b)}\quad |z+w|\leqslant|z|+|w|$$

También son de comprobación inmediata las desigualdades

$$\text{máx}\{|\operatorname{Re} z|,|\operatorname{Im} z|\}\leqslant|z|\leqslant|\operatorname{Re} z|+|\operatorname{Im} z| \tag{3.1}$$

3.1.2. Forma polar y argumentos de un número complejo

El uso de coordenadas polares en el plano facilita mucho los cálculos con productos de números complejos. Para cualquier número complejo $z=x+iy\neq 0$ podemos escribir

$$z=|z|\,\Big(\frac{x}{|z|}+i\frac{y}{|z|}\Big)$$

Como $\big(\frac{x}{|z|},\frac{y}{|z|}\big)$ es un punto de la circunferencia unidad, puede escribirse en la forma

$$\Big(\frac{x}{|z|},\frac{y}{|z|}\Big)=(\cos\vartheta,\operatorname{sen}\vartheta)$$

para algún número $\vartheta\in\mathbb{R}$. Resulta así que

$$z=|z|\,(\cos\vartheta+i\operatorname{sen}\vartheta)$$

Esta forma de expresar un número complejo recibe el nombre de **forma polar**, cuya interpretación gráfica vemos en la figura siguiente.

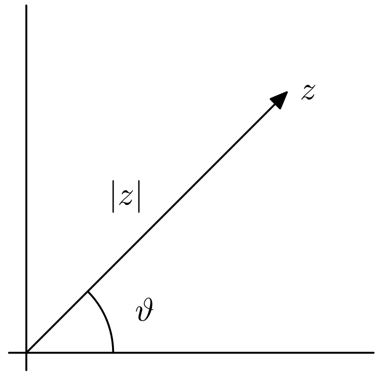


Figura 3.3: Forma polar de un número complejo

Dado $z \in \mathbb{C}$, $z \neq 0$, hay infinitos números $t \in \mathbb{R}$ que verifican la igualdad $z = |z|(\cos t, \operatorname{sen} t)$ cualquiera de ellos recibe el nombre de **argumento** de z. El conjunto de todos los argumentos de un número complejo no nulo se representa por $\operatorname{Arg}(z)$.

$$\operatorname{Arg}(z) = \{t \in \mathbb{R} : z = |z|(\cos t + i \operatorname{sen} t)\}$$

Observa que

$$s, t \in \operatorname{Arg}(z) \Longleftrightarrow \left\{ \begin{array}{c} \cos(t) = \cos(s) \\ \sin(t) = \sin(s) \end{array} \right\} \Longleftrightarrow s = t + 2k\pi \text{ para algún } k \in \mathbb{Z}$$

Por tanto, conocido un argumento $t_o \in \operatorname{Arg}(z)$ cualquier otro es de la forma $t_o + 2k\pi$ para algún $k \in \mathbb{Z}$, es decir, $\operatorname{Arg}(z) = t_o + 2\pi\mathbb{Z}$.

De entre todos los argumentos de un número complejo $z \neq 0$ hay uno único que se encuentra en el intervalo $]-\pi, \pi]$, se representa por $\arg(z)$ y se le llama **argumento principal** de z. No es difícil comprobar que el argumento principal de $z = x + iy \neq 0$ viene dado por:

$$\boxed{\arg(z) = \begin{cases} \operatorname{arc\,tg}(y/x) - \pi & \text{si } y \leqslant 0, x < 0 \\ -\pi/2 & \text{si } y \leqslant 0, x = 0 \\ \operatorname{arc\,tg}(y/x) & \text{si } x > 0 \\ \pi/2 & \text{si } y > 0, x = 0 \\ \operatorname{arc\,tg}(y/x) + \pi & \text{si } y \geqslant 0, x < 0 \end{cases}}$$

Observaciones a la definición de argumento principal

Puede parecer un poco extraña la forma de elegir el argumento principal de un número complejo. La elección que hemos hecho supone que medimos ángulos en el semiplano superior de 0 a π y en el semiplano inferior de 0 a $-\pi$.

Fíjate que si tomas un número complejo que esté situado en el tercer cuadrante $z = x + iy$ con $x < 0, y < 0$ y supones que y es próximo a 0, su argumento principal está próximo a $-\pi$, y si tomas un número complejo que esté situado en el segundo cuadrante, $w = x + iv$ con $x < 0, v > 0$, y supones que v es próximo a 0, su argumento principal está próximo a π. Además, la distancia

$|w-z|=|v-y|=v-y$ es tan pequeña como quieras. Esto nos dice que el argumento principal tiene una discontinuidad en el eje real negativo: salta de $-\pi$ a π cuando atravesamos dicho eje desde el tercer al segundo cuadrante.

Peor todavía dirás. Hasta cierto punto. Primero, la discontinuidad es inevitable. Si queremos elegir argumentos en un intervalo de longitud 2π, digamos $[\alpha,\alpha+2\pi[$, entonces dichos argumentos saltan de α a $\alpha+2\pi$ cuando atravesamos la semirrecta $(x,y)=\rho(\cos\alpha,\operatorname{sen}\alpha)$, $(\rho>0)$. En particular, si tomamos argumentos en el intervalo $[0,2\pi[$ (cosa que, a primera vista, parece lo razonable) nos encontramos con que entonces se produce una discontinuidad de dichos argumentos en *el eje real positivo*. Bien, sucede que *la extensión a* $\mathbb{C}$ de algunas funciones definidas en $\mathbb{R}^+$ (el logaritmo, las raíces) hace intervenir el argumento principal. Naturalmente, queremos que dichas extensiones sigan siendo continuas en $\mathbb{R}^+$ y ello justifica que tengamos que tomar argumentos principales de la forma en que lo hemos hecho: porque preferimos introducir una discontinuidad en $\mathbb{R}^-$ a perder la continuidad en $\mathbb{R}^+$.

Fórmula de De Moivre

Veamos cómo la forma polar permite hacer fácilmente productos de números complejos. Consideremos dos números complejos no nulos escritos en forma polar.

$$\begin{aligned} z &= |z|\,(\cos\vartheta+i\operatorname{sen}\vartheta) \\ w &= |w|\,(\cos\varphi+i\operatorname{sen}\varphi) \end{aligned}$$

Entonces

$$\begin{aligned} zw &= |z|\,|w|\,(\cos\vartheta+i\operatorname{sen}\vartheta)(\cos\varphi+i\operatorname{sen}\varphi)= \\ &= |zw|\,[(\cos\vartheta\cos\varphi-\operatorname{sen}\vartheta\operatorname{sen}\varphi)+i(\operatorname{sen}\vartheta\cos\varphi+\cos\vartheta\operatorname{sen}\varphi)]= \\ &= |zw|\,(cos(\vartheta+\varphi)+i\operatorname{sen}(\vartheta+\varphi)) \end{aligned}$$

Es decir: *para multiplicar dos números complejos se multiplican sus módulos y se suman sus argumentos.*

Así pues, el producto de dos números complejos es geométricamente un giro (pues se suman los argumentos de los números que estamos multiplicando) seguido de una homotecia (el producto de los módulos de ambos números).

Acabamos de ver que si $z,w\in\mathbb{C}^*$, $\vartheta\in\operatorname{Arg}(z)$ y $\varphi\in\operatorname{Arg}(w)$, entonces $\vartheta+\varphi\in\operatorname{Arg}(z+w)$. Es ahora fácil demostrar mediante inducción la siguiente fórmula, muy útil, conocida como fórmula de *De Moivre*.

3.2 Proposición (Fórmula de De Moivre). *Si z es un complejo no nulo, ϑ es un argumento de z y n es un número entero, se verifica que $n\vartheta\in\operatorname{Arg}(z^n)$, es decir:*

$$\boxed{z^n=\big(|z|\,(\cos\vartheta+i\operatorname{sen}\vartheta)\big)^n=|z|^n\,(\cos n\vartheta+i\operatorname{sen}n\vartheta),\qquad \vartheta\in\operatorname{Arg}(z),\ n\in\mathbb{Z}}$$

3.1.3. Raíces de un número complejo

Se trata ahora de resolver la ecuación $w^n = z$ donde n es un número natural, $n \geqslant 2$, y $z \neq 0$ es un número complejo conocido. Escribamos w en forma polar:

$$w = |w|\,(\cos\varphi + i\operatorname{sen}\varphi)$$

Ahora, usando la fórmula de De Moivre, podemos escribir la ecuación $w^n = z$ en la forma equivalente:

$$w^n = |w|^n\,(\cos n\varphi + i\operatorname{sen} n\varphi) = |z|\,(\cos\vartheta + i\operatorname{sen}\vartheta)$$

Donde $\vartheta = \arg z$. Esta igualdad se da cuando $|w|^n = |z|$ y $n\varphi = \vartheta + 2k\pi$ donde $k \in \mathbb{Z}$. Deducimos que $|w| = \sqrt[n]{|z|}$ (ojo: se trata de la raíz n–ésima de un número positivo, cosa ya conocida). Ahora bien, para cualquier número φ_k de la forma $\varphi_k = (\vartheta + 2k\pi)/n$ tenemos un número complejo

$$w_k = \sqrt[n]{|z|}(\cos\varphi_k + i\operatorname{sen}\varphi_k)$$

tal que $(w_k)^n = z$. Como una ecuación polinómica de grado n no puede tener más de n soluciones, se sigue que distintos valores de k deben dar lugar al mismo número w_k. Veamos:

$$w_k = w_q \Leftrightarrow \varphi_k - \varphi_q = 2m\pi \Leftrightarrow k - q = nm$$

Es decir, si k y q dan el mismo resto al dividirlos por n entonces $w_k = w_q$. Deducimos que para $k = 0, 1, 2, \dots, n-1$ obtenemos w_k distintos y cualquier otro w_q es igual a uno de ellos. Por tanto hay n raíces n–ésimas distintas de z.

Hemos obtenido que las n raíces n–ésimas de z vienen dadas por

$$\boxed{z_k = |z|^{1/n}\left(\cos\frac{\arg z + 2k\pi}{n} + i\operatorname{sen}\frac{\arg z + 2k\pi}{n}\right) \quad k = 0, 1, 2, \dots, n-1}$$

Observa que definiendo $u = \cos(2\pi/n) + i\operatorname{sen}(2\pi/n)$, los números $u_0 = 1$, u, $u^2, \dots, u^{n-1}$ son las raíces n–ésimas de la unidad. Podemos escribir las raíces n–ésimas de z en la forma $z_k = z_0\, u^k$. Como multiplicar por u es un giro de amplitud $2\pi/n$, deducimos que las n raíces de z se obtienen girando la raíz n–ésima principal, z_0, con giros sucesivos de amplitud $2\pi/n$. Es decir, si representamos todas las raíces n–ésimas de z obtenemos n puntos sobre una circunferencia de centro $(0,0)$ y radio $\sqrt[n]{|z|}$ que forman un polígono regular de n lados.

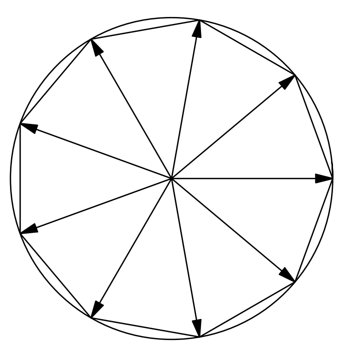


Figura 3.4: Raíces novenas de la unidad

De entre todas las raíces n–ésimas de z vamos a designar con el símbolo $\sqrt[n]{z}$ a la **raíz n-ésima principal**, que está definida por

$$\boxed{\sqrt[n]{z} = |z|^{1/n}\left(\cos\frac{\arg z}{n} + i\,\mathrm{sen}\,\frac{\arg z}{n}\right)}$$

Observa que en el caso particular de que z sea un número real positivo, entonces la raíz principal de z (considerado como número complejo) coincide con la raíz de z (considerado como número real positivo).

En general no es cierto que dados dos números complejos z y w entonces el producto de las raíces n-ésimas *principales* de z y de w sea igual a la raíz n-ésima *principal* de zw. Lo que sí es cierto es que el producto de dos raíces n-ésimas cualesquiera de z y de w es una raíz n-ésima de zw. Por tanto, $\sqrt[n]{z}\,\sqrt[n]{w}$, es **una** raíz n-ésima de zw pero no tiene por qué ser la principal.

Es fácil probar que

$$\sqrt[n]{z}\,\sqrt[n]{w} = \sqrt[n]{zw} \Longleftrightarrow -\pi < \arg(z) + \arg(w) \leqslant \pi \Longleftrightarrow \arg(zw) = \arg(z) + \arg(w)$$

Si $\operatorname{Re} z > 0$ $\operatorname{Re} w > 0$, entonces $-\pi < \arg(z) + \arg(w) < \pi$ por lo que, en este caso, $\sqrt[n]{z}\,\sqrt[n]{w} = \sqrt[n]{zw}$.

Para $n = 2, z = w = -1$, como $\arg(-1) = \pi$, tenemos que

$$\sqrt{-1} = \cos(\pi/2) + i\,\mathrm{sen}(\pi/2) = i$$

En este caso

$$\sqrt{-1}\,\sqrt{-1} = i\,i = -1 \neq \sqrt{(-1)(-1)} = \sqrt{1} = 1$$

es decir $\sqrt{-1}\,\sqrt{-1} = -1$ es una raíz cuadrada de 1 (porque $1 = (-1)(-1)$) pero no es la raíz cuadrada principal de 1.

Ahora ya sabes dónde está el error en lo que sigue:

$$-1 = i^2 = i\,i = \sqrt{-1}\,\sqrt{-1} = \sqrt{(-1)(-1)} = \sqrt{1} = 1$$

3.2. Ejercicios

1. Realiza las operaciones indicadas y expresa el resultado en la forma $a + ib$.

$$\text{i)}\ (7-2i)(5+3i) \quad \text{ii)}\ (i-1)^3 \quad \text{iii)}\ \overline{(1+i)(2+i)}(3+i) \quad \text{iv)}\ \frac{3+i}{2+i}$$
$$\text{v)}\ \frac{(4-i)(1-3i)}{-1+2i} \quad \text{vi)}\ (1+i)^{-2} \quad \text{vii)}\ \frac{1+2i}{2-i} \quad \text{viii)}\ i^2(1+i)^3$$

2. Calcula la parte real e imaginaria de las funciones:

$$\text{a)}\ f_1(z) = \overline{z}^2 \quad \text{b)}\ f_2(z) = z^3 \quad \text{c)}\ f_3(z) = \frac{1}{z} \quad \text{d)}\ f(z) = \frac{1}{1+z^2} \quad \text{e)}\ f_4(z) = \frac{z+i}{z-i}$$

3. Calcula las siguientes cantidades.

$$\text{a)}\ |(1+i)(2-i)| \quad \text{b)}\ \left|\frac{4-3i}{2-i\sqrt{5}}\right| \quad \text{c)}\ \left|(1+i)^{20}\right| \quad \text{d)}\ \left|\sqrt{2}+i(\sqrt{2}+1)\right|$$

4. Calcula los números complejos z tales que $\dfrac{1+z}{1-z}$ es:

a) Un número real; b) Un número imaginario puro.

5. Expresa en forma polar los siguientes números complejos.

$$\text{a)}\ -\sqrt{3}-i \quad \text{b)}\ -\sqrt{3}+i \quad \text{c)}\ \frac{3}{\sqrt{3}+i} \quad \text{d)}\ \frac{1+i\sqrt{3}}{(1+i)^2}$$

6. Expresa los siguientes números en la forma $a+ib$:

$$\text{a)}\ (-1+i\sqrt{3})^{11} \quad \text{b)}\ \left(\frac{1+i}{1-i}\right)^5 \quad \text{c)}\ \left(\frac{1+i\sqrt{3}}{1-i}\right)^6 \quad \text{d)}\ (-\sqrt{3}+i)^{13}$$

7. Supuesto que $|z|=1$, prueba que

$$\arg\left(\frac{z-1}{z+1}\right)=\begin{cases}\pi/2 & \text{si } \operatorname{Im} z>0\\ -\pi/2 & \text{si } \operatorname{Im} z<0\end{cases}$$

8. Resuelve la ecuación cuadrática $az^2+bz+c=0$ donde a,b,c, son números complejos conocidos y $a\neq 0$.

9. Calcula todas las soluciones de las siguientes ecuaciones:

$$\text{a)}\ z^3=1+i \quad \text{b)}\ z^4=i \quad \text{c)}\ z^3=-1+i\sqrt{3} \quad \text{d)}\ z^8=1 \quad \text{e)}\ z^2+\sqrt{32}\,iz-6i=0$$

10. Calcula las soluciones de las ecuaciones:

$$\text{a)}\ z^4+2z^3+7z^2-18z+26=0; \qquad \text{b)}\ z^4+(1+2i)z^2+2i=0$$

11. Demuestra la llamada “igualdad del paralelogramo”:

$$|z+w|^2+|z-w|^2=2(|z|^2+|w|^2) \quad (z,w\in\mathbb{C})$$

y explica su significado geométrico.

12. Prueba que $\left|\dfrac{z-a}{1-\overline{a}z}\right|<1$ si $|z|<1$ y $|a|<1$ y también si $|z|>1$ y $|a|>1$.

Sugerencia: Una estrategia básica para probar desigualdades entre *módulos* de números complejos consiste en elevar al cuadrado ambos miembros de la desigualdad.

13. Sea x un número real que no es múltiplo entero de 2π. Prueba las igualdades

$$\text{a)}\quad 1+\cos x+\cos 2x+\cdots+\cos nx = \cos\left(\frac{n}{2}x\right)\frac{\operatorname{sen}\left(\dfrac{n+1}{2}x\right)}{\operatorname{sen}\left(\dfrac{x}{2}\right)}$$

$$\text{b)}\quad \operatorname{sen} x+\operatorname{sen} 2x+\cdots+\operatorname{sen} nx = \operatorname{sen}\left(\frac{n}{2}x\right)\frac{\operatorname{sen}\left(\dfrac{n+1}{2}x\right)}{\operatorname{sen}\left(\dfrac{x}{2}\right)}$$

Sugerencia: Si llamamos A a la primera suma y B a la segunda, calcula $A+iB$ haciendo uso de la fórmula de De Moivre.

14. Haciendo uso de la fórmula de De Moivre prueba que:

a) $\operatorname{sen} 3\varphi = 3 \operatorname{sen}\varphi - 4\operatorname{sen}^3\varphi$;

b) $\cos 4\varphi = 8\cos^4\varphi - 8\cos^2\varphi + 1$.

15. Representar gráficamente los conjuntos de números complejos z que verifican:

$$
\begin{array}{llll}
|z-3| \leqslant 3; & 2 < |z-i| \leqslant 3; & |\arg z| < \pi/6; & |z-i| + |z+i| = 4 \\
|z-1| = |z-2i|; & \left|\dfrac{z-i}{z+2i}\right| = 2; & \operatorname{Im}(z^2) > 6; & |z-i| = \operatorname{Im} z + 1
\end{array}
$$

3.3. Funciones elementales complejas

Las funciones complejas no son más que las funciones definidas en subconjuntos de $\mathbb{R}^2$ con valores en $\mathbb{R}^2$ cuando en $\mathbb{R}^2$ consideramos su estructura compleja. Dado un conjunto $A \subset \mathbb{C}$, a toda función compleja $f : A \to \mathbb{C}$ se le asocian dos funciones reales: la función $u = \operatorname{Re} f$ "parte real de f" y la función $v = \operatorname{Im} f$ "parte imaginaria de f" definidas para todo $(x,y) = x+iy \in A$ por:

$$u(x,y) = \operatorname{Re} f(x+iy), \quad v(x,y) = \operatorname{Im} f(x+iy)$$

Naturalmente, $f(x+iy) = u(x,y) + i\,v(x,y)$.

3.3.1. La función exponencial

Definimos[1] la exponencial compleja de un número $z = x + i\,y$ como

$$\mathrm{e}^{x+iy} = \exp(x+iy) = \mathrm{e}^x\big(\cos y + i \operatorname{sen} y\big)$$

Observa que

$$|\,\mathrm{e}^z\,| = \mathrm{e}^{\operatorname{Re} z}, \quad \operatorname{Im} z \in \operatorname{Arg}(\mathrm{e}^z)$$

En particular, obtenemos la llamada *fórmula de Euler*:

$$\boxed{\mathrm{e}^{it} = \cos t + i \operatorname{sen} t \qquad (\text{para todo } t \in \mathbb{R})}$$

que establece una relación entre la exponencial compleja y las funciones trigonométricas. De la fórmula de Euler se deducen fácilmente las llamadas *ecuaciones de Euler*:

$$\boxed{\cos t = \frac{\mathrm{e}^{it} + \mathrm{e}^{-it}}{2}, \qquad \operatorname{sen} t = \frac{\mathrm{e}^{it} - \mathrm{e}^{-it}}{2i} \qquad (t \in \mathbb{R})}$$

Se prueba fácilmente que $\mathrm{e}^{z+w} = \mathrm{e}^z\,\mathrm{e}^w$ para todos $z, w \in \mathbb{C}$. Se deduce que para todo $z \in \mathbb{C}$ y todo $k \in \mathbb{Z}$ es

$$\mathrm{e}^z = \mathrm{e}^{z+2k\pi i}$$

Lo que nos dice que la exponencial compleja es una función **periódica** con período $2\pi i$. Naturalmente, esto supone una gran diferencia con la exponencial real que es una función inyectiva. Observa que la exponencial no se anula nunca pues $|\,\mathrm{e}^z\,| = \mathrm{e}^{\operatorname{Re} z} > 0$.

[1] Más adelante veremos la justificación de esta definición.

3.3.2. Logaritmos complejos

Dado un número complejo $z \neq 0$, hay infinitos números complejos w que satisfacen la ecuación $\mathrm{e}^w = z$. Cualquiera de ellos se llama **un logaritmo** de z. El conjunto de todos ellos lo representaremos por $\operatorname{Log} z$ y es el conjunto:

$$\boxed{\operatorname{Log} z = \{\log|z| + i(\arg(z) + 2k\pi), k \in \mathbb{Z}\}}$$

De entre todos ellos elegimos uno, llamado **logaritmo principal**, definido por

$$\boxed{\log z = \log|z| + i\arg(z) \qquad \text{para todo } z \in \mathbb{C}^*}$$

Observa que cualquier otro logaritmo de z es de la forma $\log(z) + i2k\pi$ para algún entero k. Es importante que observes que la igualdad

$$\log zw = \log z + \log w$$

que es válida para los logaritmos de los números reales positivos, no es siempre cierta para números complejos. Por ejemplo:

$$\log\left(\mathrm{e}^{i2\pi/3}\right) = i\frac{2\pi}{3},\ \log\left(\mathrm{e}^{i3\pi/4}\right) = i\frac{3\pi}{4},\ \log\left(\mathrm{e}^{i2\pi/3}\mathrm{e}^{i3\pi/4}\right) = \log\left(\mathrm{e}^{i17\pi/12}\right) = \log\left(\mathrm{e}^{-i7\pi/12}\right) = -i\frac{7\pi}{12}$$

Lo que está claro es que el número $\log z + \log w \in \operatorname{Log}(zw)$, es decir, $\log z + \log w$ es **un** logaritmo de zw pero no tiene por qué ser el logaritmo **principal** de zw.

3.3.3. Potencias complejas

Recuerda que dados dos números reales $a > 0$ y $b \in \mathbb{R}$, la potencia de base a y exponente b se define como $a^b = e^{b\log a}$. Ahora, dados $a, b \in \mathbb{C}$, con $a \neq 0$, sabemos que hay infinitos logaritmos de a, todos ellos son de la forma $\log a + i2k\pi$, con $k \in \mathbb{Z}$. Por ello, cualquier número complejo de la forma $\mathrm{e}^{b(\log a + i2k\pi)}$ donde $k \in \mathbb{Z}$, es **una** potencia de base a y exponente b. Representamos por $[a^b]$ el conjunto de todas ellas.

$$[a^b] = \left\{\mathrm{e}^{b(\log a + i2k\pi)} : k \in \mathbb{Z}\right\}$$

Se destaca una:

$$a^b = \mathrm{e}^{b\log a}$$

que se llama **valor principal** de la potencia de base a y exponente b. Observa que si $b = 1/n$ donde $n \in \mathbb{N}$, el número

$$a^{1/n} = \exp\left(\frac{1}{n}\log a\right) = \exp\left(\frac{\log a}{n} + i\frac{\arg a}{n}\right) = |z|^{1/n}\left(\cos\frac{\arg a}{n} + i\operatorname{sen}\frac{\arg a}{n}\right)$$

es el valor principal de la raíz n-ésima de a que antes hemos notado por $\sqrt[n]{a}$.

3.4. Ejerccios

1. Expresa los 8 números $\pm 1 \pm i$, $\pm\sqrt{3} \pm i$ en la forma $r\,\mathrm{e}^{i\varphi}$.

2. Calcula el módulo y los argumentos principales de los números

$$1+\mathrm{e}^{i\varphi},\ 1-\mathrm{e}^{i\varphi},\ -a\,\mathrm{e}^{i\varphi}$$

donde $|\varphi| \leqslant \pi$ y $a > 0$.

3. Calcula $\log z$ y $\operatorname{Log} z$ cuando z es uno de los números siguientes

$$i,\ -i,\ \mathrm{e}^{-3},\ \mathrm{e}^{5i},\ 4,\ -5\,\mathrm{e},\ 1+i$$

4. Calcula $\log(3i) + \log(-1+i\sqrt{3})$ y $\log\big(3i(-1+i\sqrt{3})\big)$.

5. Calcula $\log(-1-i) - \log i$ y $\log\left(\dfrac{-1-i}{i}\right)$.

6. Calcula

$$[(-4)^i],\ i^{-3i},\ [i^{2/\pi}],\ [i^i],\ 1^{2i},\ 3^{1-i},\ ((-i)^i)^i,\ (1+i)^{1+i}$$

7. Estudia, para $z\in\mathbb{C}^*$ y $n\in\mathbb{N}$, las igualdades:

$$\text{a)}\ \log(\exp(z)) = z;\ b)\ \exp(\log(z)) = z;\ c)\ \log(\sqrt[n]{z}) = \frac{\log(z)}{n};\ d)\ \log(z^n) = n\log(z).$$

8. Explica con detalle dónde está el error en las igualdades siguientes:

$$i = (-1)^{1/2} = [(-1)^3]^{1/2} = (-1)^{3/2} = i^3 = -i$$

Lección 4

Continuidad

Introducción

Para motivar la definición que vamos a dar de continuidad, consideremos una ley física de la forma $P = f(V)$, que relaciona los valores de una "variable independiente V" (podemos pensar que es el volumen de un gas) con otra "variable dependiente P" (podemos pensar que es la presión). Si queremos usar dicha ley, hemos de medir un valor V_o de la variable V, y es inevitable que al hacerlo cometamos algún error el cual, naturalmente, influye en el correspondiente valor de P, que ya no será exactamente igual a $P_o = f(V_o)$. Surge así la pregunta natural: ¿de qué forma el error en la medida de V afecta al valor resultante de P? Es claro que si para valores de V "muy próximos" a V_o obtengo valores de P muy diferentes entre sí, la ley "f" que relaciona V con P no tendrá ninguna utilidad práctica.

Puesto que los errores de medida son inevitables, no es razonable tratar de obtener "el verdadero valor P_o". Lo que sí puede hacerse es fijar una cota de error admisible para P (la cual dependerá de cada situación concreta); llamemos "ε" a dicha cota, $(\varepsilon > 0)$, y tratar de obtener otra cota de error "δ", $(\delta > 0)$, de tal forma que siempre que midamos V_o con un error menor que δ tengamos la seguridad de que el valor resultante para P se diferencia de P_o en menos que ε. Esto es, $|f(V) - f(V_o)| < \varepsilon$ siempre que $|V - V_o| < \delta$. Cuando esto efectivamente pueda hacerse para cualquier cota de error $\varepsilon > 0$ decimos que la ley "f" es continua en V_o. Observa que cabe esperar que la cota de error δ dependa del $\varepsilon > 0$ fijado en cada caso, y también de V_o.

Las ideas anteriores conducen, de forma natural, a la definición matemática de continuidad. En todo lo que sigue, la letra A representará un conjunto no vacío de números reales. En la práctica A será siempre un intervalo o una unión de intervalos. Recuerda que la notación $f : A \to \mathbb{R}$ quiere decir que f es una función real cuyo dominio es A. Es muy importante advertir que A no tiene por qué coincidir con el dominio natural de la función. Esto es así porque con frecuencia estamos interesados en estudiar propiedades de una función en una parte de su dominio natural. Además, la continuidad de f depende tanto de la "regla que la define" como del conjunto en donde estamos trabajando. Enseguida pondremos ejemplos para aclarar esto.

4.1 Definición (Continuidad en un punto). Una función $f : A \to \mathbb{R}$ se dice que es continua en un punto $a \in A$ si, para cada número $\varepsilon > 0$, se puede encontrar un número $\delta > 0$ (que, en general, dependerá de ε y de a) tal que para todo $x \in A$ con $|x - a| < \delta$ se verifica que $|f(x) - f(a)| < \varepsilon$.

La definición anterior suele escribirse, con abuso del formalismo lógico, de la siguiente forma:

$$\forall \varepsilon \in \mathbb{R}^+ \quad \exists \delta \in \mathbb{R}^+ : \quad \left.\begin{array}{c} |x-a| < \delta \\ x \in A \end{array}\right\} \Longrightarrow \quad |f(x) - f(a)| < \varepsilon$$

Observa cómo en esta definición el conjunto A tiene mucho protagonismo: sólo se consideran los valores de f en A, lo que le pueda pasara a f fuera de A no nos interesa.

Se dice que f es continua en un subconjunto $C \subseteq A$, si f es continua en todo punto de C.

No suele ser tarea fácil demostrar que una función dada es continua. Generalmente, lo que se hace es descomponer la función que queremos estudiar en otras más sencillas cuya continuidad ya es conocida previamente. Es por ello interesante saber qué tipo de operaciones realizadas con funciones continuas conducen a nuevas funciones continuas.

4.1.1. Propiedades básicas de las funciones continuas

4.2 Teorema. *Sean f, g funciones reales definidas en A. Se verifica que:*

1. *Las funciones $f + g$ y fg son continuas en todo punto de A en el que las dos funciones f y g sean continuas. En particular, las funciones suma y producto de funciones continuas son funciones continuas.*
2. *Si $g(x) \neq 0$ para todo $x \in A$, la función $\dfrac{1}{g}$ es continua en todo punto de A en el que g sea continua. En consecuencia, la función cociente de dos funciones continuas cuyo denominador no se anula nunca es una función continua.*

Las propiedades anteriores no son difíciles de demostrar y, sin embargo, son de gran utilidad.

4.3 Corolario. *Las funciones racionales son funciones continuas.*

De hecho, todas las funciones elementales que conoces son continuas en sus dominios naturales de definición.

Además de sumar y multiplicar funciones, también sabemos componerlas. Veamos cómo se comporta la continuidad respecto de la composición de funciones.

4.4 Teorema (Continuidad de una función compuesta). *Sean $f : A \to \mathbb{R}$ y $g : B \to \mathbb{R}$ funciones tales que $f(A) \subseteq B$. Supongamos que f es continua en un punto $a \in A$ y que g es continua en el punto $f(a)$. Entonces la función compuesta $g \circ f : A \to \mathbb{R}$ es continua en el punto a. En particular, si g es continua en $f(A)$, entonces $g \circ f$ es continua en todo punto de A en el que f sea continua. Más en particular, la composición de funciones continuas es una función continua.*

Demostración. Dado $\varepsilon > 0$, por la continuidad de g en $f(a)$, existe $\rho > 0$ tal que para todo $y \in B$ con $|y - f(a)| < \rho$ se tiene que $|g(y) - g(f(a))| < \varepsilon$. Ahora, por la continuidad de f en a, existe

$\delta > 0$ tal que para todo $x \in A$ con $|x - a| < \delta$ se tiene que $|f(x) - f(a)| < \rho$. Deducimos así que $|g(f(x)) - g(f(a))| < \varepsilon$ para todo $x \in A$ con $|x - a| < \delta$. Es decir, la función compuesta $g \circ f$ es continua en a.

La continuidad de una función en un punto permite obtener información sobre el comportamiento de la función en los puntos próximos al mismo. Estos resultados se llaman *locales*.

4.5 Teorema (**Conservación local del signo**). *Sea $f : A \to \mathbb{R}$ continua en un punto $a \in A$ con $f(a) \neq 0$. Entonces hay un número $r > 0$ tal que para todo $x \in A$ con $|x - a| < r$ se verifica que $f(x)f(a) > 0$. (Es decir, f es positiva (si $f(a) > 0$) o negativa (si $f(a) < 0$) en todos los puntos de un entorno de a)*

Demostración. Supondremos que $f(a) > 0$. Podemos entonces tomar $\varepsilon = f(a)/2$ para obtener, en virtud de la continuidad de f en a, un $r > 0$ tal que para todo $x \in A$ con $|x - a| < r$ se verifica que $|f(x) - f(a)| < f(a)/2$, lo que implica que $f(x) > f(a)/2 > 0$. El caso en que $f(a) < 0$ se reduce al anterior sin más que sustituir f por $-f$.

4.2. Teorema de Bolzano. Supremo e ínfimo

Si ahora mides 175cms. y hace 10 años medías 135cms., es seguro que en algún momento intermedio medías con exactitud 161cms. Si una entrada de cine cuesta 5 euros y hace 3 años costaba 4 euros, es seguro que en algún momento ir al cine costaba exactamente 4,99 euros. ¿Seguro? No, a ningún empresario de cine le parecería bien cobrar 4,99 euros por la entrada.

La diferencia está en que la talla de una persona es una función continua del tiempo y para pasar de 135cms. a 175cms. tiene que pasar por todos los valores intermedios, pero el precio de las entradas de cine no varía de forma continua con el tiempo y puede pasar "de golpe" de 4,5 euros a 5 euros.

La gráfica de una función continua en un intervalo, $f : [a, b] \to \mathbb{R}$, la imaginamos como una curva continua, por ello, si $f(a) < 0 < f(b)$, la gráfica de f tiene que atravesar el eje x para pasar de un punto situado por debajo de él a otro que se encuentra por encima y, por tanto, f *tiene que* anularse en algún punto entre a y b. Esto es precisamente lo que afirma el conocido *teorema* que sigue.

4.6 Teorema (**Teorema de los ceros de Bolzano**). *Toda función continua en un intervalo que toma valores positivos y negativos se anula en algún punto de dicho intervalo.*

Lo primero que llama la atención en este teorema es su *evidencia*. No está de más a este respecto recordar que, como decía Bertrand Russell, "en matemáticas la evidencia es enemiga de la corrección". Precisamente, el mérito de Bernard Bolzano (1781-1848) está en haber llamado la atención sobre la necesidad de *demostrar* muchas proposiciones, aparentemente evidentes, que se refieren a las funciones continuas. Podemos añadir, además, que suele ser particularmente difícil demostrar matemáticamente lo que nuestra intuición presenta como evidente; de hecho, *con las herramientas que tenemos hasta ahora no podemos demostrar el teorema.*

La función $f(x) = x^2 - 2$ es continua y $f(0) < 0 < f(2)$, el teorema de Bolzano asegura que existe un número positivo en el que f se anula. En otras palabras, el teorema prueba *la existencia* del número $\sqrt{2}$ y, como dicho número no es racional, deducimos que para probar el teorema se precisa usar alguna propiedad que NO tienen los números racionales. Pero todas las propiedades de los números reales que enunciamos en la primera lección las tienen también los números racionales. Concluimos que los números reales deberán tener otra propiedad que todavía no hemos considerado.

Comentamos el primer día que no debemos preocuparnos mucho *por lo que sea* el número $\sqrt{2}$, pero al menos deberíamos de tener alguna forma de *probar su existencia*; es decir, de las propiedades de los números reales se debería poder deducir que hay un número cuyo cuadrado es igual a 2. ¿Qué sabemos de $\sqrt{2}$? No es racional, pero podemos aproximarlo por racionales. Con una calculadora obtenemos sucesivas aproximaciones racionales de $\sqrt{2}$ por defecto:

$$1{,}41,\ 1{,}414,\ 1{,}4142,\ 1{,}41421,\ 1{,}414213, \dots$$

Es claro que $\sqrt{2}$ debe ser el *menor número mayor que todas ellas*. Pues bien, justamente necesitamos una propiedad que garantice la existencia de ese “menor número mayor que”. Nos vendrá bien introducir alguna terminología nueva.

4.7 Definición. Sea E un conjunto no vacío de números reales. Un número $z\in\mathbb{R}$ se dice que es un **mayorante o cota superior** (resp. **minorante o cota inferior**) de E si $x\leqslant z$ (resp. $z\leqslant x$) para todo $x\in E$.

Si hay algún elemento de E que también sea mayorante (resp. minorante) de E, dicho elemento es necesariamente único y se llama **máximo** (resp.
textbf mínimo) de E y lo representaremos por $\max(E)$ (resp. $\min(E)$).

Un conjunto que tiene algún mayorante (resp. minorante) se dice que está **mayorado o acotado superiormente** (resp. **minorado o acotado inferiormente**). Un conjunto que está mayorado y minorado se dice que está **acotado**.

Está claro que un conjunto puede no tener mínimo ni máximo. Los problemas de “optimización” consisten, justamente, en estudiar condiciones que garanticen la existencia de valores máximos y mínimos para funciones de diversas clases. La siguiente propiedad garantiza que *ciertos conjuntos* de números reales tienen mínimo.

P8 [Propiedad del supremo] Para todo conjunto de números reales no vacío y mayorado se verifica que el conjunto de sus mayorantes tiene mínimo.

4.8 Definición. Dado un conjunto $E\subseteq\mathbb{R}$, no vacío y mayorado, se llama **supremo o extremo superior** de E, al mínimo mayorante de E y lo notaremos por $\sup(E)$.

Con esta terminología lo que dice la propiedad del supremo es que todo conjunto de números reales no vacío y mayorado tiene supremo (pero nótese que el supremo no tiene por qué pertenecer al conjunto).

La propiedad del supremo es lo que distingue a los números reales de los racionales. Dicha propiedad se usa para probar la existencia de números reales que cumplen alguna determinada condición. La demostración del teorema de Bolzano es un ejemplo importante de ello.

Demostración del teorema de los ceros de Bolzano

Es suficiente probar que si $f : [a,b] \to \mathbb{R}$ es continua y $f(a) < 0 < f(b)$, entonces f se anula en algún punto del intervalo $]a,b[$. Una buena estrategia para demostrar un teorema es "darlo por demostrado" y *trabajar hacia atrás*. Tenemos que buscar un punto $c \in]a,b[$ tal que $f(c) = 0$. Por supuesto, puede haber muchos puntos donde f se anule (el teorema dice que *al menos hay uno*), pero de todos ellos el más fácil de caracterizar es el "primero", porque a la izquierda de él la función es siempre negativa. Esto lleva a considerar el conjunto E de los puntos $x \in [a,b]$ tales que f toma valores negativos en $[a,x]$:

$$E = \{x \in [a,b] : f(t) < 0 \ \text{ para todo } \ t \in [a,x]\}$$

Por su definición, tenemos que $E \subset [a,b]$ y $a \in E$. La propiedad del supremo nos dice que hay un número real, c, que es el supremo de E. Es evidente que $a \leqslant c \leqslant b$. La propiedad de conservación local del signo implica que existe algún $\delta > 0$ tal que $a + \delta < b - \delta$ y f es negativa en todos los puntos del intervalo $[a, a + \delta]$ y positiva en todos los puntos del intervalo $[b - \delta, b]$. Esto implica que $a < c < b$.

Veamos que $[a,c[\subset E$. Sea $a < x_o < c$. Como $x_o < c$ y c es el mínimo mayorante de E, tiene que existir algún punto $z_o \in E$ tal que $x_o < z_o \leqslant c$. Por tanto, si $t \in [a, x_o]$ también $t \in [a, z_o]$ y, como, $z_o \in E$, será $f(t) < 0$, luego $x_o \in E$. Nótese que hemos probado también que $f(x) < 0$ para todo $x \in [a,c[$. Finalmente, probaremos que $f(c) = 0$. Como a la izquierda de c la función f toma valores negativos y f es continua, deducimos que *no puede ser* $f(c) > 0$ y, por tanto, $f(c) \leqslant 0$. Pero tampoco puede ser $f(c) < 0$, pues entonces, por la conservación local del signo, habría un intervalo de la forma $[c - \rho, c + \rho] \subset [a,b]$ tal que $f(t) < 0$ para todo $t \in [c - \rho, c + \rho]$ lo que implica que en E hay puntos mayores que c lo que es contradictorio. Concluimos así que $f(c) = 0$.

Hay consecuencias de este teorema que están lejos de ser evidentes. Por ejemplo, puede probarse, con la ayuda del teorema de Bolzano, que si tenemos tres sólidos en el espacio (imagina que son tres bocadillos de muy distintos tamaños), es siempre posible encontrar un plano que los divida simultáneamente en partes iguales (puedes cortar a los tres bocatas exactamente por la mitad de un sólo tajo).

Un enunciado *equivalente* del teorema de Bolzano es el siguiente.

4.9 Teorema (Teorema del valor intermedio). *La imagen de un intervalo por una función continua es un intervalo.*

Hemos *demostrado* así la *evidencia* inicial: una función continua en un intervalo toma todos los valores comprendidos entre dos cualesquiera de sus valores.

Veamos algunas consecuencias sencillas del teorema de Bolzano.

4.10 Corolario (Existencia de raíces). *Dados $a > 0$ y $k \in \mathbb{N}$ hay un único número $c > 0$ tal que $c^k = a$.*

4.11 Corolario (Ceros de polinomios de grado impar). *Toda función polinómica de grado impar se anula en algún punto.*

A partir de la propiedad del supremo, se prueba con facilidad el siguiente resultado.

4.12 Proposición (Propiedad del ínfimo). *Para todo conjunto de números reales no vacío y minorado se verifica que el conjunto de sus minorantes tiene máximo.*

4.13 Definición. Dado un conjunto $E \subseteq \mathbb{R}$, no vacío y minorado, se llama **ínfimo o extremo inferior** de E, al máximo minorante de E y lo notaremos por $\operatorname{ínf}(E)$.

Con esta terminología lo que dice la propiedad del ínfimo es que todo conjunto de números reales no vacío y minorado tiene ínfimo (pero nótese que el ínfimo no tiene por qué pertenecer al conjunto).

4.3. Ejercicios

1. *a*) Da un ejemplo de una función continua cuya imagen no sea un intervalo.
   b) Da un ejemplo de una función definida en un intervalo cuya imagen sea un intervalo y que no sea continua.
   c) Da un ejemplo de una función continua en todo $\mathbb{R}$, no constante y cuya imagen sea un conjunto (obligatoriamente un intervalo) acotado.
   d) Da un ejemplo de una función continua en $[0,1[$ tal que $f([0,1[)$ no sea acotado.
   e) Da un ejemplo de una función continua definida en un intervalo abierto acotado y cuya imagen sea un intervalo cerrado y acotado.
2. Sea $f:[a,b] \to \mathbb{R}$ continua. Supongamos que $a \leqslant f(x) \leqslant b$ para todo x en $[a,b]$. Prueba que hay algún punto $c \in [a,b]$ tal que $f(c)=c$.
3. Sea $a>1$. Prueba que la ecuación $x+\mathrm{e}^{-x}=a$ tiene al menos una solución positiva y otra negativa.
4. Prueba que la ecuación $x+\mathrm{e}^{x}+\operatorname{arc\,tg} x=0$ tiene una sola raíz real. Da un intervalo de longitud uno en el que se encuentre dicha raíz.
5. Suponiendo que la temperatura varía de forma continua, prueba que siempre hay dos puntos antípodas en el ecuador terrestre que están a la misma temperatura.
6. Sea $f:[a,b] \to \mathbb{R}$ continua con $f(a)=f(b)$. Dado $n \in \mathbb{N}$, $n \geqslant 2$, prueba que hay algún punto $c \in [a, b-(b-a)/n]$ tal que $f(c)=f(c+(b-a)/n)$.
7. Un corredor recorre 6 kilómetros en 30 minutos. Demuestra que en algún momento de su carrera recorre 1 kilómetro en exactamente 5 minutos.
8. Un reloj averiado marca inicialmente un tiempo t_0. El reloj puede adelantar o atrasar, pero cuenta con exactitud períodos de 12 horas, es decir, pasadas 12 horas el reloj marca un tiempo t_0+12 horas. Demuestra que en algún momento dicho reloj mide con exactitud una hora.
9. Un automovilista sale de Granada hacia Madrid un sábado a las 8h de la mañana y el domingo inicia el regreso a la misma hora. Sabiendo que invirtió igual tiempo en ambos viajes, pruébese que en algún momento del domingo el automovilista se encuentra a igual distancia de Granada que a la que se encontraba el sábado en ese mismo momento.

10. Sean f,g funciones continuas que no se anulan en un intervalo I, tales que $(f(x))^2=(g(x))^2$ para todo $x\in I$. Prueba que o bien $f(x)=g(x)$ para todo $x\in I$, o bien $f(x)=-g(x)$ para todo $x\in I$. ¿Cuántas funciones hay $\varphi:\mathbb{R}\to\mathbb{R}$ continuas y verificando que $(\varphi(x))^2=x^2$ para todo $x\in\mathbb{R}$?.

11. Justifica que toda función polinómica de grado impar se anula en algún punto.

12. Sea $f:\mathbb{R}\to\mathbb{R}$ continua y decreciente. Prueba que hay un único $a\in\mathbb{R}$ tal que $f(a)=a$.

    Para probar desigualdades en las que intervienen supremos o ínfimos las siguientes observaciones, aunque evidentes, pueden ser útiles. Sea $C\subseteq\mathbb{R}$ un conjunto no vacío.

    (I) Si queremos probar que un número real x verifica que $\sup(C)\leqslant x$, lo que tenemos que hacer es probar que x es un mayorante de C.

    (II) Si queremos probar que un número real x verifica que $x\leqslant \operatorname{ínf}(C)$, lo que tenemos que hacer es probar que x es un minorante de C.

13. Sean A,B conjuntos no vacíos de números reales. Supongamos que $a\leqslant b$ para todo $a\in A$ y para todo $b\in B$. Prueba que $\sup A\leqslant \operatorname{ínf} B$.

14. Sean A, B, conjuntos no vacíos y acotados de números reales. Definamos

    $$A-B=\{a-b : a\in A,\ b\in B\};\ \ AB=\{ab : a\in A,\ b\in B\}$$

    Prueba que $\sup(A-B)=\sup A-\operatorname{ínf} B$ y, supuesto que $A\subset\mathbb{R}^+$ y $B\subset\mathbb{R}^+$, prueba que $\sup(AB)=\sup A\sup B$.

15. Sea A un conjunto no vacío de números reales. Para cada $x\in\mathbb{R}$ definamos la “distancia de x a A” por $\operatorname{dist}(x,A)=\operatorname{ínf}\{|x-a| : a\in A\}$. Prueba que para todos $x,y\in\mathbb{R}$ se verifica que:

    $$|\operatorname{dist}(x,A)-\operatorname{dist}(y,A)|\leqslant|x-y|$$

    Deduce que la aplicación $x\mapsto\operatorname{dist}(x,A)$ es continua.

16. Sea $f:\mathbb{R}\to\mathbb{R}$ continua, mayorada y tal que para todos $a,b\in\mathbb{R}$ con $a<b$, se verifica que $\sup f(]a,b[)=\sup f(\mathbb{R})$. Prueba que f es constante.

17. Sea $f:[a,b]\to\mathbb{R}$ una función continua tal que $f(a)<0$, $f(b)<0$ y $f(c)>0$ para algún $c\in]a,b[$. Prueba que hay dos números u, v tales que $a<u<v<b$, $f(u)=f(v)=0$ y $f(x)>0$ para todo $x\in]u,v[$.

18. Sea $f:[a,b]\to\mathbb{R}$ creciente. Supongamos que $a\leqslant f(x)\leqslant b$ para todo x en $[a,b]$. Prueba que hay algún punto $c\in[a,b]$ tal que $f(c)=c$.

Lección 5

Sucesiones

Introducción

Las sucesiones aparecen de manera natural en muchos cálculos que responden a un esquema iterativo. Por ejemplo, al dividir 2 entre 3 obtenemos $\frac{2}{3}=\frac{6}{10}+\frac{2}{3}\frac{1}{10}$, igualdad que podemos usar ahora para obtener

$$\frac{2}{3}=\frac{6}{10}+\left(\frac{6}{10}+\frac{2}{3}\frac{1}{10}\right)\frac{1}{10}=\frac{6}{10}+\frac{6}{10^2}+\frac{2}{3}\frac{1}{10^2}$$

y de nuevo

$$\frac{2}{3}=\frac{6}{10}+\frac{6}{10^2}+\left(\frac{6}{10}+\frac{2}{3}\frac{1}{10}\right)\frac{1}{10^2}=\frac{6}{10}+\frac{6}{10^2}+\frac{6}{10^3}+\frac{2}{3}\frac{1}{10^3}$$

y así podemos continuar tantas veces como queramos, obteniendo para cada $n\in\mathbb{N}$ la igualdad:

$$\frac{2}{3}=\sum_{k=1}^{n}\frac{6}{10^k}+\frac{2}{3}\frac{1}{10^n}.$$

Escribiendo $x_n=\sum_{k=1}^{n}\frac{6}{10^k}$ tenemos que $0<\frac{2}{3}-x_n=\frac{2}{3}\frac{1}{10^n}$. Nótese que, aunque los números x_n son *todos ellos distintos* de $2/3$, dada una cota de error arbitrariamente pequeña $\varepsilon>0$ y tomando $n_0\in\mathbb{N}$ de manera que $\frac{2}{3}\frac{1}{10^{n_0}}<\varepsilon$, deducimos que *para todo* número natural $n\geqslant n_0$ se verifica que $|x_n-2/3|<\varepsilon$, lo que se expresa escribiendo $2/3=\lim\limits_{n\to\infty}\{x_n\}$.

Este ejemplo está relacionado con la expresión decimal de $2/3$ que, como todos sabemos, es un decimal periódico con período igual a 6, lo que suele escribirse $2/3=0,\widehat{6}$ igualdad en la que, según se dice a veces, el símbolo $0,\widehat{6}$ debe interpretarse como que el 6 *se repite infinitas veces*. £Qué quiere decir esto? Lo que está claro es que, por mucho tiempo y paciencia que tengamos, nunca podremos escribir *infinitos* 6 uno detrás de otro... bueno, podríamos escribir algo como

$$\frac{2}{3}=0,\widehat{6}=0,6666666\ldots(\text{infinitos}\,6)$$

lo que tampoco sirve de mucho pues seguimos sin saber cómo se interpreta esta igualdad. Pues bien, para dar un significado matemático a lo que se quiere expresar con esa igualdad hay que recurrir al concepto de límite de una sucesión tal como hemos hecho antes.

Veamos otro ejemplo en esta misma línea. Vamos a intentar calcular aproximaciones racionales a $\sqrt{10}$. Si partimos inicialmente de un número $x > \sqrt{10}$, tendremos que $\frac{10}{x} < \sqrt{10} < x$. Pongamos $y = \frac{1}{2}\left(x + \frac{10}{x}\right)$. Entonces, en virtud de la desigualdad de las medias, $\sqrt{10} < y$, y como también $y < x$, deducimos que y está más cerca de $\sqrt{10}$ que x. Podemos ahora repetir este proceso sustituyendo x por y obteniendo una nueva aproximación mejor de $\sqrt{10}$. Nótese que si x es racional también lo será y. Esto sugiere que, partiendo de un valor inicial, por ejemplo $x_1 = 4$, calculemos $x_2 = \frac{1}{2}\left(x_1 + \frac{10}{x_1}\right)$, y después $x_3 = \frac{1}{2}\left(x_2 + \frac{10}{x_2}\right)$, y así podemos continuar tantas veces como queramos, obteniendo para cada $n \in \mathbb{N}$ un número x_n tal que

$$x_{n+1} = \frac{1}{2}\left(x_n + \frac{10}{x_n}\right)$$

con $x_1 = 4$. Con una calculadora manual obtenemos enseguida los valores $x_2 = 3{,}25$; $x_3 = 3{,}1634615$; $x_4 = 3{,}1622779$ con seis cifras decimales exactas:

$$0 < x_4 - \sqrt{10} = \frac{x_4^2 - 10}{x_4 + \sqrt{10}} < \frac{x_4^2 - 10}{6} < \frac{0{,}000005}{6} < \frac{1}{10^6}$$

es decir, x_4 coincide con $\sqrt{10}$ hasta la sexta cifra decimal. De hecho, como $x_n > \sqrt{10}$ tenemos que:

$$0 < x_{n+1} - \sqrt{10} = \frac{1}{2}\left(x_n + \frac{10}{x_n}\right) - \sqrt{10} < \frac{1}{2}x_n + \frac{1}{2}\sqrt{10} - \sqrt{10} = \frac{1}{2}(x_n - \sqrt{10})$$

de donde se sigue que $0 < x_{n+1} - \sqrt{10} < \frac{1}{2^n}(x_1 - \sqrt{10}) < \frac{1}{2^n}$, por tanto, dado cualquier $\varepsilon > 0$, y tomando $n_0 \in \mathbb{N}$ tal que $2^{-n_0} < \varepsilon$, deducimos que *para todo* número natural $n \geqslant n_0$ se verifica que $|x_n - \sqrt{10}| < \varepsilon$, lo que simbólicamente se expresa escribiendo $\sqrt{10} = \lim_{n\to\infty} \{x_n\}$.

En los ejemplos anteriores hemos dado por supuesto que ya tienes cierta familiaridad con los conceptos de "sucesión" y de "límite de una sucesión" de los cuales vamos a ocuparnos a continuación con detalle.

Sucesión de elementos de un conjunto

Sea A un conjunto no vacío. Una sucesión de elementos de A es una **aplicación** del conjunto $\mathbb{N}$ de los números naturales en A. En particular, una sucesión de números reales es una **aplicación** del conjunto $\mathbb{N}$ de los números naturales en el conjunto $\mathbb{R}$ de los números reales.

En todo lo que sigue solamente consideraremos sucesiones de números reales por lo que nos referiremos a ellas simplemente como "sucesiones".

Dada una sucesión $\varphi : \mathbb{N} \to \mathbb{R}$ suele emplearse una notación especial para representarla. Para $n \in \mathbb{N}$ suele notarse el número real $\varphi(n)$ en la forma $x_n = \varphi(n)$ (naturalmente la letra "x" nada tiene de especial y puede sustituirse por cualquier otra). La sucesión misma se representa por

$\varphi = \{x_n\}_{n\in\mathbb{N}}$, es decir, el símbolo $\{x_n\}_{n\in\mathbb{N}}$ debe interpretarse como la **aplicación** que a cada $n\in\mathbb{N}$ hace corresponder el número real x_n. Cuando no hay posibilidad de confusión escribimos simplemente $\{x_n\}$ en vez de $\{x_n\}_{n\in\mathbb{N}}$. Conviene insistir en que $\{x_n\}$ es, por definición, la **aplicación** de $\mathbb{N}$ en $\mathbb{R}$ dada por $n \mapsto x_n$. No hay que confundir la sucesión $\{x_n\}$, que es una aplicación, con su **conjunto imagen**, que es el subconjunto de $\mathbb{R}$ formado por todos los números x_n, el cual se representa por $\{x_n : n \in \mathbb{N}\}$. Por ejemplo, $\{(-1)^n\}$ y $\{(-1)^{n+1}\}$ son sucesiones distintas con el mismo conjunto imagen. El número x_n se llama *término n-ésimo* de la sucesión; para $n = 1, 2, 3$ se habla respectivamente de primero, segundo, tercer término de la sucesión.

5.1. Sucesiones de números reales

5.1 Definición. Una sucesión $\{x_n\}$ se dice que **converge** a un número real x si, dado cualquier número real $\varepsilon > 0$, existe un número natural m_ε tal que si n es cualquier número natural mayor o igual que m_ε se cumple que $|x_n - x| < \varepsilon$. Simbólicamente:

$$\forall\varepsilon > 0 \ \ \exists m_\varepsilon \in\mathbb{N} \ : n \geqslant m_\varepsilon \Rightarrow |x_n - x| < \varepsilon$$

Se dice también que el número x es **límite de la sucesión** $\{x_n\}$ y se escribe $\lim\limits_{n\to\infty}\{x_n\} = x$ o, simplemente, $\lim\{x_n\} = x$ e incluso, si no hay posibilidad de confusión, $\{x_n\} \to x$.

Se comprueba fácilmente que *una sucesión convergente tiene un único límite.*

En Matemáticas se dan definiciones para introducir nuevos conceptos y saber de qué estamos hablando, pero las definiciones no suelen ser útiles para el cálculo. Por eso no debes preocuparte si la definición anterior te parece difícil de aplicar en casos concretos. Debes hacer un esfuerzo por comprenderla pero no tendrás que usarla para hacer cálculos.

Estudiamos a continuación cómo se comportan las sucesiones convergentes respecto de las estructuras algebraica y de orden de $\mathbb{R}$.

5.2 Proposición. *Supongamos que* $\lim\{x_n\} = x$, $\lim\{y_n\} = y$ *y que existe* $m \in\mathbb{N}$ *tal que para todo* $n \geqslant m$ *se tiene que* $x_n \leqslant y_n$. *Entonces se verifica que* $x \leqslant y$.

Respecto al resultado anterior, de muy fácil demostración, conviene advertir que *aunque las desigualdades sean estrictas no puede asegurarse que* $\lim\{x_n\} = x$ *sea estrictamente menor que* $\lim\{y_n\} = y$. Por ejemplo, si $x_n = 0$ e $y_n = 1/n$, es claro que $x_n < y_n$ para todo $n \in \mathbb{N}$ pero $x = 0 = y$.

5.3 Proposición (Principio de las sucesiones encajadas). *Supongamos que* $\{x_n\}$, $\{y_n\}$, $\{z_n\}$ *son sucesiones tales que* $\lim\{x_n\} = \lim\{z_n\} = \alpha$ *y existe un número natural* m_0 *tal que para todo* $n \geqslant m_0$ *se verifica que* $x_n \leqslant y_n \leqslant z_n$, *entonces la sucesión* $\{y_n\}$ *es convergente y* $\lim\{y_n\} = \alpha$.

Demostración. Sea $\varepsilon > 0$. Por hipótesis existen m_1, m_2 tales que

$$\alpha - \varepsilon < x_p < \alpha + \varepsilon \quad \text{y} \quad \alpha - \varepsilon < z_q < \alpha + \varepsilon \tag{5.1}$$

para todo $p \geqslant m_1$ y todo $q \geqslant m_2$. Sea $m_3 = \max\{m_0, m_1, m_2\}$. Para todo $n \geqslant m_3$ las desigualdades (5.1) se cumplen para $p = q = n$, además como $n \geqslant m_0$ se tiene que $x_n \leqslant y_n \leqslant z_n$. Deducimos que, para todo $n \geqslant m_3$, se verifica que

$$\alpha - \varepsilon < x_n \leqslant y_n \leqslant z_n < \alpha + \varepsilon$$

y, por tanto, $\alpha-\varepsilon<y_n<\alpha+\varepsilon$, es decir, $\lim\{y_n\}=\alpha$. □

Una consecuencia inmediata de este resultado es que si cambiamos arbitrariamente un número finito de términos de una sucesión la nueva sucesión así obtenida es convergente si lo era la de partida y con su mismo límite.

El principio de las sucesiones encajadas es de gran utilidad y se usa con mucha frecuencia. Naturalmente, cuando apliquemos dicho principio a un caso concreto, la sucesión $\{y_n\}$ del enunciado será la que queremos estudiar y tendremos que ser capaces de “inventarnos” las sucesiones $\{x_n\}$ y $\{z_n\}$ de manera que se cumplan las condiciones del enunciado.

5.4 Definición. Una sucesión $\{x_n\}$ se dice que es:

Mayorada o acotada superiormente si su conjunto imagen está mayorado, es decir, si hay un número $\mu\in\mathbb{R}$ tal que $x_n\leqslant\mu$ para todo $n\in\mathbb{N}$.

Minorada o acotada inferiormente si su conjunto imagen está minorado, es decir, si hay un número $\lambda\in\mathbb{R}$ tal que $\lambda\leqslant x_n$ para todo $n\in\mathbb{N}$.

Acotada si su conjunto imagen está mayorado y minorado, equivalentemente, si hay un número $M\in\mathbb{R}^+$ tal que $|x_n|\leqslant M$ para todo $n\in\mathbb{N}$.

Creciente si $x_n\leqslant x_{n+1}$ para todo $n\in\mathbb{N}$.

Estrictamente creciente si $x_n<x_{n+1}$ para todo $n\in\mathbb{N}$.

Decreciente si $x_n\geqslant x_{n+1}$ para todo $n\in\mathbb{N}$.

Estrictamente decreciente si $x_n>x_{n+1}$ para todo $n\in\mathbb{N}$.

Monótona si es creciente o decreciente.

Estrictamente monótona si es estrictamente creciente o decreciente.

Observa que si una sucesión $\{x_n\}$ es creciente (resp. decreciente) entonces se verifica que $x_m\leqslant x_n$ (resp. $x_m\geqslant x_n$) siempre que $m\leqslant n$.

Conviene advertir que cuando se dice que una sucesión es monótona *no se excluye* la posibilidad de que, de hecho, sea estrictamente monótona. Es por ello que, en general, suele hablarse de sucesiones monótonas y tan sólo cuando tiene algún interés particular se precisa si son estrictamente monótonas.

5.5 Proposición. *Toda sucesión convergente está acotada.*

Demostración. Supongamos que $\lim\{x_n\}=x$. Todos los términos de $\{x_n\}$ a partir de uno en adelante estarán en el intervalo $]x-1,x+1[$, es decir, hay un número $m\in\mathbb{N}$ tal que para todo $n\geqslant m$ se verifica que $|x_n-x|<1$, lo que implica que

$$|x_n|\leqslant|x_n-x|+|x|<1+|x|\quad\text{para todo } n\geqslant m.$$

Tomando $M=\max\{1+|x|,|x_1|,\cdots,|x_m|\}$, tenemos que $|x_n|\leqslant M$ para todo $n\in\mathbb{N}$. □

La proposición anterior es útil a veces para probar que una sucesión *no* es convergente: para ello basta probar que no está acotada.

La proposición recíproca de la anterior no es cierta: la sucesión $\{(-1)^n\}$ es acotada y *no* es convergente. No obstante, hay un caso especial muy importante en que sí es cierta la recíproca.

5.6 Teorema. *Toda sucesión monótona y acotada es convergente. Más concretamente, si una sucesión* $\{x_n\}$ *es:*

i) Creciente y mayorada, entonces $\lim\{x_n\} = \beta$, *donde* $\beta = \sup\{x_n : n\in\mathbb{N}\}$.

ii) Decreciente y minorada, entonces $\lim\{x_n\} = \alpha$, *donde* $\alpha = \inf\{x_n : n\in\mathbb{N}\}$.

Demostración. Probaremos *i*) quedando la demostración de *ii*) como ejercicio. La hipótesis de que $\{x_n\}$ es mayorada garantiza, en virtud del principio del supremo, la existencia del número real $\beta = \sup\{x_n : n \in \mathbb{N}\}$. Dado $\varepsilon > 0$, tiene que existir un término x_m de la sucesión tal que $\beta - \varepsilon < x_m$. Puesto que la sucesión es creciente para todo $n \geqslant m$ se verificará que $x_m \leqslant x_n$, y por tanto $\beta - \varepsilon < x_n$. En consecuencia $\beta - \varepsilon < x_n < \beta + \varepsilon$ para todo $n \geqslant m$. Hemos probado así que $\lim\{x_n\} = \beta$. □

5.7 Ejemplo. La sucesión $\{x_n\}$ definida por $x_n = \displaystyle\sum_{k=n+1}^{2n} \frac{1}{k}$, es convergente.

En efecto, como

$$x_{n+1} - x_n = \frac{1}{2n+2} + \frac{1}{2n+1} - \frac{1}{n+1} > \frac{1}{2n+2} + \frac{1}{2n+2} - \frac{1}{n+1} = 0$$

se sigue que $x_{n+1} > x_n$ para todo $n \in \mathbb{N}$, es decir, es una sucesión creciente. Además

$$x_n \leqslant \frac{1}{n+1} + \overset{(n}{\cdots} + \frac{1}{n+1} = \frac{n}{n+1} < 1$$

por lo que también está mayorada. Concluimos, por el teorema anterior, que dicha sucesión es convergente. ♦

5.8 Ejemplo (El número e**).** En el ejercicio (8) hemos probado que la sucesión $x_n = \left(1 + \frac{1}{n}\right)^n$ es creciente y que la sucesión $y_n = \left(1 + \frac{1}{n}\right)^{n+1}$ es decreciente. Como $0 < y_n$, se sigue que $\{y_n\}$ es convergente. Puesto que

$$x_n = y_n \left(1 + \frac{1}{n}\right)^{-1} = y_n \frac{n}{n+1}$$

se sigue que $\{x_n\}$ también es convergente y $\lim\{x_n\} = \lim\{y_n\}$. El valor común de este límite es un número real que se representa con el símbolo e. Como consecuencia del teorema (5.6), se verifica que

$$\mathrm{e} = \sup\left\{ \left(1 + \frac{1}{n}\right)^n : n\in\mathbb{N} \right\} = \inf\left\{ \left(1 + \frac{1}{m}\right)^{m+1} : m\in\mathbb{N} \right\}$$

En particular, se verifica que

$$\left(1 + \frac{1}{n}\right)^n < \mathrm{e} < \left(1 + \frac{1}{m}\right)^{m+1} \qquad (n, m\in\mathbb{N})$$

♦

En los resultados anteriores han intervenido de manera esencial las propiedades de la estructura de orden de $\mathbb{R}$. Vamos a estudiar ahora el comportamiento de las sucesiones convergentes respecto de la adición y el producto de números reales. Los resultados que vamos a obtener, conocidos tradicionalmente con el nombre de *álgebra de límites*, son básicos para el estudio de la convergencia de sucesiones.

Dadas dos sucesiones $\{x_n\}$ e $\{y_n\}$, se define su **suma** como la sucesión $\{x_n + y_n\}$ y su **producto** como la sucesión $\{x_n y_n\}$.

5.9 Proposición. *El producto de una sucesión convergente a cero por una sucesión acotada es una sucesión convergente a cero.*

Demostración. Sea $\lim\{x_n\} = 0$, e $\{y_n\}$ acotada. Sea $c > 0$ tal que $|y_n| \leqslant c$ para todo $n \in \mathbb{N}$. Dado $\varepsilon > 0$, existe un número natural m tal que para todo $n \geqslant m$ se verifica que $|x_n| < \varepsilon/c$. Deducimos que, para todo $n \geqslant m$, se verifica que $|x_n y_n| = |x_n||y_n| < \dfrac{\varepsilon}{c} c = \varepsilon$, lo que prueba que $\lim\{x_n y_n\} = 0$. □

5.10 Proposición (Álgebra de límites). *Supongamos que* $\lim\{x_n\} = x$ *y* $\lim\{y_n\} = y$. *Entonces se verifica que:*

$$\lim\{x_n + y_n\} = x + y, \quad \lim\{x_n y_n\} = xy\,.$$

Si además suponemos que $y \neq 0$, *entonces* $\lim\{x_n/y_n\} = x/y$.

Demostración. Dado $\varepsilon > 0$, por hipótesis existen m_1, m_2 tales que

$$x - \varepsilon/2 < x_p < x + \varepsilon/2 \quad \text{y} \quad y - \varepsilon/2 < y_q < y + \varepsilon/2 \tag{5.2}$$

para todo $p \geqslant m_1$ y todo $q \geqslant m_2$. Sea $m_0 = \max\{m_1, m_2\}$. Para todo $n \geqslant m_0$ las desigualdades (5.2) se cumplen para $p = q = n$, por lo que, sumándolas término a término, deducimos que $x + y - \varepsilon < x_n + y_n < x + y + \varepsilon$ cualquiera sea $n \geqslant m_0$, lo que prueba que $\lim\{x_n + y_n\} = x + y$.

Teniendo en cuenta que $\lim\{(x_n - x)y_n\} = \lim\{x(y_n - y)\} = 0$, y la igualdad

$$x_n y_n - xy = (x_n - x)y_n + x(y_n - y)$$

deducimos que $\lim\{x_n y_n - xy\} = 0$, es decir, $\lim\{x_n y_n\} = xy$.

Finalmente, para probar que $\lim\{x_n/y_n\} = x/y$, probaremos que la sucesión

$$\left\{\frac{x_n}{y_n} - \frac{x}{y}\right\} = \left\{\frac{x_n y - y_n x}{y_n y}\right\}$$

converge a cero, para lo cual, teniendo en cuenta que $\lim\{x_n y - y_n x\} = xy - yx = 0$, bastará probar que la sucesión $\{1/y_n\}$ está acotada. Puesto que $\lim\{y_n\} = y$, se deduce de la desigualdad $||y_n| - |y|| \leqslant |y_n - y|$ que $\lim\{|y_n|\} = |y|$. Existirá, por tanto, un número $m_0 \in \mathbb{N}$ tal que para todo $n \geqslant m_0$ es $|y_n| > |y|/2$. Pongamos $K = \max\left\{\dfrac{1}{|y_1|}, \dfrac{1}{|y_2|}, \dots, \dfrac{1}{|y_{m_0}|}, \dfrac{2}{|y|}\right\}$. Se tiene entonces que $\dfrac{1}{|y_n|} \leqslant K$ para todo $n \in \mathbb{N}$. Hemos probado así que la sucesión $\{1/y_n\}$ está acotada, lo que concluye la demostración del teorema. □

Hay que leer con atención las hipótesis del teorema anterior para no hacer un uso incorrecto del mismo. En particular, no hay que olvidar que *la suma o el producto de dos sucesiones no convergentes puede ser una sucesión convergente.*

5.11 Definición. Sea $\{x_n\}$ una sucesión de números reales; dada una aplicación $\sigma:\mathbb{N}\to\mathbb{N}$ *estrictamente creciente*, la sucesión que a cada número natural n hace corresponder el número real $x_{\sigma(n)}$ se representa por $\{x_{\sigma(n)}\}$ y se dice que es una **sucesión parcial** de $\{x_n\}$. Nótese que $\{x_{\sigma(n)}\}$ no es otra cosa que la composición de las aplicaciones $\{x_n\}$ y σ, esto es, $\{x_{\sigma(n)}\}=\{x_n\}\circ\sigma$.

Se dice que un número real x es un **valor de adherencia** de la sucesión $\{x_n\}$ si hay alguna sucesión parcial de $\{x_n\}$ que converge a x.

5.12 Ejemplo. Representemos por $E(x)$ el mayor entero menor o igual que x. La sucesión $\{x_n\}$ dada por $x_n=n/5-E(n/5)$ para todo $n\in\mathbb{N}$, tiene a $0,1/5,2/5,3/5$ y $4/5$, como valores de adherencia.
En efecto, basta considerar que para cada $j\in\{0,1,2,3,4\}$, la sucesión parcial $\{x_{5n-j}\}_{n\in\mathbb{N}}$ viene dada por $x_{5n}=0$, para $j=0$, y $x_{5n-j}=1-j/5$ para $j=1,2,3,4$. ♦

Es fácil probar por inducción que si σ es una aplicación estrictamente creciente de $\mathbb{N}$ en $\mathbb{N}$ entonces se verifica que $\sigma(n)\geqslant n$ para todo $n\in\mathbb{N}$. Con ello se obtiene fácilmente el siguiente resultado.

5.13 Proposición. *Si* $\lim\{x_n\}=x$, *toda sucesión parcial de* $\{x_n\}$ *también converge a* x. *En particular, una sucesión convergente tiene como único valor de adherencia su límite.*

Observa que hay sucesiones, la de los números naturales por ejemplo, que no tienen *ningún* valor de adherencia. También puede ocurrir que una sucesión *tenga un único valor de adherencia y no sea convergente*. Por ejemplo, la sucesión dada para todo $n\in\mathbb{N}$ por $x_n=(1+(-1)^n)n+1/n$, no es convergente y tiene a 0 como único valor de adherencia. Vamos a ver a continuación que estos comportamientos no pueden darse con sucesiones acotadas.

5.14 Lema. *Toda sucesión tiene una sucesión parcial monótona.*

Demostración. Sea $\{x_n\}$ una sucesión y definamos

$$A=\{n\in\mathbb{N}: x_n\geqslant x_p \quad \text{para todo } p>n\}$$

Podemos visualizar el conjunto A como sigue. Consideremos en el plano los segmentos de extremos (n,x_n) y $(n+1,x_{n+1})$, $n=1,2,3,\dots$. Resulta así una línea poligonal infinita y podemos imaginar que dicha línea es el perfil de una cordillera cuyas cumbres y valles son los puntos (n,x_n). Imaginemos ahora que los rayos de luz del Sol, paralelos al eje de abscisas, iluminan dicha cordillera por el lado derecho (el Sol estaría, pues, situado en el infinito del eje de abscisas positivo). Pues bien, un número natural n pertenece al conjunto A si el punto (n,x_n) está iluminado y no pertenece a A si dicho punto está en sombra.

Supongamos que A es infinito. Entonces podemos definir una aplicación $\sigma:\mathbb{N}\to\mathbb{N}$ estrictamente creciente y tal que $\sigma(\mathbb{N})=A$ de la siguiente forma:

$$\begin{aligned}&\sigma(1)=\min(A)\\&\sigma(n+1)=\min\{p\in A:\sigma(n)<p\} \quad \text{para todo } n\in\mathbb{N}\end{aligned}$$

es decir la aplicación σ va eligiendo los elementos de A de menor a mayor empezando por el primero. Resulta ahora evidente que la sucesión parcial $\{x_{\sigma(n)}\}$ es decreciente (todos los puntos $(\sigma(n),x_{\sigma(n)})$ están iluminados y, por tanto, ninguno de ellos puede hacerle sombra a uno anterior).

Si A es finito podemos suponer que $A = \emptyset$. En tal caso, para todo $n \in \mathbb{N}$ hay algún $p > n$ tal que $x_n < x_p$ (pues todo punto (n, x_n) está en sombra). Podemos definir ahora una aplicación $\sigma : \mathbb{N} \to \mathbb{N}$ estrictamente creciente de la siguiente forma:

$$\begin{aligned} &\sigma(1) = 1 \\ &\sigma(n+1) = \min\{p \in \mathbb{N} : \sigma(n) < p \ \text{ y } \ x_{\sigma(n)} < x_p\} \ \text{ para todo } n \in \mathbb{N} \end{aligned}$$

Resulta ahora evidente que la sucesión parcial $\{x_{\sigma(n)}\}$ es creciente (porque cada punto $(\sigma(n), x_{\sigma(n)})$ deja en la sombra al anterior). □

5.15 Teorema (**Teorema de Bolzano - Weierstrass**). *Toda sucesión acotada de números reales tiene alguna sucesión parcial convergente.*

Demostración. Sea $\{x_n\}$ una sucesión acotada. En virtud el lema anterior, hay una sucesión parcial de $\{x_n\}$ que es monótona, dicha sucesión parcial está acotada por estarlo $\{x_n\}$ y, por tanto, es convergente. □

Si volvemos a leer la definición de sucesión convergente, parece que para estudiar la convergencia de una sucesión $\{x_n\}$ debemos ser capaces de "adivinar", de alguna manera, su posible límite. De hecho, una idea bastante extendida consiste en pensar que es lo mismo probar la convergencia de una sucesión que calcular su límite. Esto no es del todo correcto; son relativamente pocas las sucesiones convergentes cuyo límite puede efectivamente calcularse. Cuando se estudia la convergencia de una sucesión $\{x_n\}$, la mayoría de las veces, lo que conocemos es, justamente, la sucesión y, naturalmente, se desconoce su posible límite el cual pudiera, incluso, no existir. Por ello interesa tener *criterios de convergencia intrínsecos a la sucesión*, es decir, que no hagan intervenir a un objeto en principio *extraño* a ella como es su posible límite. Conocemos ya un criterio de convergencia intrínseco para sucesiones *monótonas*. Usando dicho criterio hemos probado la convergencia de la sucesión $x_n = \displaystyle\sum_{k=n+1}^{2n} \frac{1}{k}$ *sin necesidad de conocer su límite*.

A continuación vamos a establecer un criterio intrínseco de convergencia para sucesiones que es más general pues puede aplicarse a cualquier sucesión. Este criterio fué formulado por Bolzano en 1817 y también, independientemente, por Cauchy en 1821, y establece una condición necesaria y suficiente para la convergencia de una sucesión. Dicha condición se conoce con el nombre de *condición de Cauchy*.

5.16 Definición. Se dice que una sucesión $\{x_n\}$ satisface la **condición de Cauchy**, si para cada número positivo, $\varepsilon > 0$, existe un número natural m_ε, tal que para todos $p, q \in \mathbb{N}$ con $p \geqslant m_\varepsilon$ y $q \geqslant m_\varepsilon$ se verifica que $|x_p - x_q| < \varepsilon$.

La condición de Cauchy puede también expresarse de una manera equivalente, aunque formalmente distinta, como sigue:

Una sucesión $\{x_n\}$ satisface la condición de Cauchy, si para cada número positivo, $\varepsilon > 0$, existe un número natural m_ε, tal que para todo $p \geqslant m_\varepsilon$ y *para todo número natural* h, se verifica que $|x_{p+h} - x_p| < \varepsilon$.

5.17 Teorema (Teorema de complitud de $\mathbb{R}$). *Una sucesión de números reales es convergente si, y sólo si, verifica la condición de Cauchy.*

Demostración. Supongamos que $\{x_n\}$ verifica la condición de Cauchy. Probemos primero que $\{x_n\}$ está acotada. La condición de Cauchy implica que hay $m_0 \in \mathbb{N}$ tal que $|x_p - x_{m_0}| < 1$ para todo $p \geqslant m_0$, y como $|x_p| \leqslant |x_p - x_{m_0}| + |x_{m_0}|$, deducimos que $|x_p| < 1 + |x_{m_0}|$ para $p \geqslant m_0$. En consecuencia si definimos $M = \text{máx}\{|x_1|, |x_2|, \ldots, |x_{m_0}|, 1 + |x_{m_0}|\}$, obtenemos que $|x_n| \leqslant M$ para todo $n \in \mathbb{N}$.

El teorema de Bolzano-Weierstrass garantiza que hay un número real x y una sucesión parcial $\{x_{\sigma(n)}\}$ que converge a x. Probaremos que $\{x_n\}$ también converge a x. Dado $\varepsilon > 0$, existe $n_o \in \mathbb{N}$ tal que $|x_p - x_q| < \varepsilon/2$ siempre que $p, q \geqslant n_o$. También existe $n_1 \in \mathbb{N}$ tal que $|x_{\sigma(n)} - x| < \varepsilon/2$ siempre que $n \geqslant n_1$. Sea $m = \text{máx}\{n_o, n_1\}$. Para todo $n \geqslant m$ se tiene que $\sigma(n) \geqslant n \geqslant m$ por lo que

$$|x_n - x| \leqslant |x_n - x_{\sigma(n)}| + |x_{\sigma(n)} - x| < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon$$

lo que prueba que $\lim\limits_{n\to\infty}\{x_n\} = x$.

Recíprocamente, si $\{x_n\}$ es convergente y $\lim\{x_n\} = x$, dado $\varepsilon > 0$, hay un número $m_\varepsilon \in \mathbb{N}$ tal que para todo número natural $n \geqslant m_\varepsilon$ se tiene que $|x_n - x| < \varepsilon/2$. Deducimos que si p, q son números naturales mayores o iguales que m_ε entonces $|x_p - x_q| \leqslant |x_p - x| + |x - x_q| <$
$< \varepsilon/2 + \varepsilon/2 = \varepsilon$. Por tanto la sucesión $\{x_n\}$ verifica la condición de Cauchy. □

5.1.1. Sucesiones divergentes. Indeterminaciones en el cálculo de límites

5.18 Definición. Una sucesión $\{x_n\}$ se dice que es **positivamente divergente**, y escribimos $\{x_n\} \to +\infty$, si para todo número real $K > 0$ existe un número natural $m_K \in \mathbb{N}$, tal que para todo $n \in \mathbb{N}$ con $n \geqslant m_K$ se verifica que $x_n \geqslant K$.

Una sucesión $\{x_n\}$ se dice que es **negativamente divergente**, y escribimos $\{x_n\} \to -\infty$, si para todo número real $K < 0$ existe un número natural $m_K \in \mathbb{N}$, tal que para todo $n \in \mathbb{N}$ con $n \geqslant m_K$ se verifica que $x_n \leqslant K$.

Diremos que una sucesión es **divergente** para indicar que es positivamente o negativamente divergente.

En la siguiente proposición se exponen algunas propiedades elementales, pero importantes, de las sucesiones divergentes. Teniendo en cuenta que $\{x_n\} \to +\infty$ si, y sólo si, $\{-x_n\} \to -\infty$, es suficiente enunciar dichas propiedades para sucesiones positivamente divergentes.

5.19 Proposición. *i)* $\{|x_n|\} \to +\infty$ *si, y sólo si,* $\{1/x_n\} \to 0$.

ii) La suma de una sucesión positivamente divergente con una sucesión acotada es una sucesión positivamente divergente.

iii) La suma de una sucesión positivamente divergente con una sucesión minorada es otra sucesión positivamente divergente. En particular, la suma de dos sucesiones positivamente divergentes es otra sucesión positivamente divergente.

iv) El producto de dos sucesiones positivamente divergentes es otra sucesión positivamente divergente.

v) El producto de una sucesión positivamente divergente por una sucesión que converge a un número positivo es otra sucesión positivamente divergente.

Sucesiones de exponenciales y logaritmos

A continuación vamos a estudiar sucesiones de exponenciales y logaritmos. El resultado básico al respecto es el siguiente.

5.20 Proposición. *a) Para toda sucesión* $\{x_n\}$ *se verifica que:*

i) $\{x_n\} \to x\in\mathbb{R} \iff \{e^{x_n}\} \to e^x$.

ii) $\{x_n\} \to +\infty \iff \{e^{x_n}\} \to +\infty$.

iii) $\{x_n\} \to -\infty \iff \{e^{x_n}\} \to 0$.

b) Para toda sucesión de números positivos $\{x_n\}$ *se verifica que:*

iv) $\{x_n\} \to x > 0 \iff \{\log(x_n)\} \to \log x$.

v) $\{x_n\} \to +\infty \iff \{\log(x_n)\} \to +\infty$.

vi) $\{x_n\} \to 0 \iff \{\log(x_n)\} \to -\infty$.

Frecuentemente hay que estudiar la convergencia o divergencia de una suma o producto de dos sucesiones precisamente cuando las reglas que hemos visto en secciones anteriores no pueden aplicarse. Se trata de aquellos casos en que el comportamiento de las sucesiones $\{x_n + y_n\}$, $\{x_n y_n\}$ no está determinado por el de $\{x_n\}$ e $\{y_n\}$. Por ejemplo, si sabemos que $\{x_n\} \to +\infty$ y que $\{y_n\} \to -\infty$, ¿qué podemos decir del comportamiento de la sucesión $\{x_n + y_n\}$? Respuesta: absolutamente nada. Baste para convencerse de ello la consideración de los siguientes casos:

$$
\begin{array}{lll}
x_n = 2n, & y_n = -n; & \{x_n + y_n\} = \{n\} \to +\infty \\
x_n = n, & y_n = -2n; & \{x_n + y_n\} = \{-n\} \to -\infty \\
x_n = n+1, & y_n = -n; & \{x_n + y_n\} = \{1\} \to 1 \\
x_n = (-1)^n + n, & y_n = (-1)^n - n; & \{x_n + y_n\} = \{2(-1)^n\}
\end{array}
$$

En consecuencia, las sucesiones del tipo $\{x_n + y_n\}$ donde $\{x_n\} \to +\infty$, $\{y_n\} \to -\infty$, *requieren un estudio particular en cada caso*. Tales sucesiones suele decirse que son **una indeterminación del tipo** “$\infty - \infty$”.

Análogamente, si sabemos que $\{x_n\} \to 0$ y que $\{y_n\}$ es divergente, ello no proporciona ninguna información sobre el comportamiento de la sucesión $\{x_n y_n\}$; la cual se dice que es **una indeterminación del tipo** “$0\,\infty$”. Las indeterminaciones que aparecen al estudiar el cociente de dos sucesiones divergentes o de dos sucesiones que convergen a cero, las llamadas **indeterminaciones de los tipos** “∞/∞”, “$0/0$”, pueden reducirse a una indeterminación del tipo “$0\,\infty$”.

El siguiente resultado permite resolver en muchas ocasiones indeterminaciones de la forma “∞/∞”.

5.21 Teorema (Criterio de Stolz). *Sea* $\{y_n\}$ *una sucesión positivamente divergente y estrictamen-*

te creciente y sea $\{x_n\}$ *cualquier sucesión. Supongamos que*

$$\left\{\frac{x_{n+1}-x_n}{y_{n+1}-y_n}\right\}\to L$$

donde $L\in\mathbb{R}$*, o* $L=+\infty$*, o* $L=-\infty$*. Entonces se verifica también que* $\left\{\dfrac{x_n}{y_n}\right\}\to L$.

Del Criterio de Stolz se deducen dos útiles criterios para estudiar la convergencia de sucesiones de medias aritméticas o geométricas.

5.22 Proposición (Criterio de la media aritmética). *Supongamos que* $\{a_n\}\to L$ *donde* L *es un número real, o* $L=+\infty$*, o* $L=-\infty$*. Entonces se verifica que*

$$\left\{\frac{a_1+a_2+\cdots+a_n}{n}\right\}\to L.$$

5.23 Proposición (Criterio de la media geométrica). *Supongamos que* $\{a_n\}\to L$ *donde* $\{a_n\}$ *es una sucesión de números positivos y* L *es un número real o bien* $L=+\infty$*. Entonces se verifica que*

$$\left\{\sqrt[n]{a_1a_2\ldots a_n}\right\}\to L.$$

5.24 Corolario. *Supongamos que* $\left\{\dfrac{x_{n+1}}{x_n}\right\}\to L$ *donde* $\{x_n\}$ *es una sucesión de números positivos y* L *es un número real o bien* $L=+\infty$*. Entonces se verifica que* $\{\sqrt[n]{x_n}\}\to L$.

Hay otras indeterminaciones que surgen al considerar *sucesiones de potencias*, es decir, sucesiones de la forma $\{x_n^{y_n}\}$ donde $\{x_n\}$ es una sucesión de números positivos e $\{y_n\}$ es una sucesión cualquiera de números reales. Puesto que $x_n^{y_n}=\exp(y_n\log(x_n))$, teniendo en cuenta la proposición (5.20), la convergencia o divergencia de la sucesión $\{x_n^{y_n}\}$ vendrá determinada por la de $\{y_n\log(x_n)\}$; la cual, a su vez, está determinada en todos los casos por el comportamiento de las sucesiones $\{x_n\}$ e $\{y_n\}$, excepto cuando dicha sucesión $\{y_n\log(x_n)\}$ es una indeterminación del tipo "$0\,\infty$", lo que ocurre en los siguientes casos.

a) $\{x_n\}\to 1$, $\{|y_n|\}\to+\infty$ (indeterminación "1^∞")

b) $\{x_n\}\to+\infty$, $\{y_n\}\to 0$ (indeterminación "∞^0")

c) $\{x_n\}\to 0$, $\{y_n\}\to 0$ (indeterminación "0^0")

Cuando estudiemos las derivadas veremos algunas técnicas que permiten resolver en muchos casos estas indeterminaciones.

5.2. Sucesiones de números complejos

5.25 Definición. Una sucesión de números complejos $\{z_n\}$ se dice que **converge** a un número complejo z si, dado cualquier número real $\varepsilon>0$, existe un número natural m_ε tal que si n es cualquier número natural mayor o igual que m_ε se cumple que $|z_n-z|<\varepsilon$. Simbólicamente:

$$\forall\varepsilon>0\ \exists m_\varepsilon\in\mathbb{N}\ :\ n\geqslant m_\varepsilon\Rightarrow|z_n-z|<\varepsilon$$

Se dice que el número z es **límite de la sucesión** $\{z_n\}$ y se escribe $\lim_{n\to\infty}\{z_n\} = z$ o, simplemente, $\lim\{z_n\} = z$ e incluso, si no hay posibilidad de confusión, $\{z_n\} \to z$.

Observa que, en virtud de la definición dada, se verifica que

$$\{z_n\} \to z \iff |z_n - z| \to 0$$

Recordemos que $\max\{|\operatorname{Re} z|, |\operatorname{Im} z|\} \leqslant |z| \leqslant |\operatorname{Re} z| + |\operatorname{Im} z|$. Gracias a esta desigualdad tenemos que

$$\left.\begin{array}{l} |\operatorname{Re} z_n - \operatorname{Re} z| \\ |\operatorname{Im} z_n - \operatorname{Im} z| \end{array}\right\} \leqslant |z_n - z| \leqslant |\operatorname{Re} z_n - \operatorname{Re} z| + |\operatorname{Im} z_n - \operatorname{Im} z|$$

Deducimos que $|z_n - z| \to 0$ si, y sólo si, $|\operatorname{Re} z_n - \operatorname{Re} z| \to 0$ y $|\operatorname{Im} z_n - \operatorname{Im} z| \to 0$. Hemos probado así el siguiente resultado.

5.26 Proposición. *Una sucesión de números complejos* $\{z_n\}$ *es convergente si, y sólo si, las sucesiones de números reales* $\{\operatorname{Re} z_n\}$ *y* $\{\operatorname{Im} z_n\}$ *son convergentes. Además, en dicho caso*

$$\lim\{z_n\} = z \iff \operatorname{Re} z = \lim\{\operatorname{Re} z_n\} \quad y \quad \operatorname{Im} z = \lim\{\operatorname{Im} z_n\}$$

Gracias a este resultado el estudio de sucesiones de números complejos se reduce a estudiar la convergencia de dos sucesiones de números reales.

Los resultados obtenidos para sucesiones de números reales *en los que no interviene la estructura de orden* son también válidos para sucesiones de números complejos. Son válidos, en particular, los resultados relativos a álgebra de límites, el teorema de Bolzano–Weierstrass y el teorema de complitud.

5.3. Ejercicios

1. *a*) Sea $\{x_n\}$ una sucesión y supongamos que hay números $\rho \in]0,1[$, $p \in \mathbb{N}$, tales que para todo $n \geqslant p$ es $|x_{n+1}| \leqslant \rho |x_n|$. Prueba que $\lim\{x_n\} = 0$.

   b) Sea $\{x_n\}$ una sucesión de números no nulos verificando que $\lim \dfrac{|x_{n+1}|}{|x_n|} = \lambda$, donde $0 \leqslant \lambda < 1$. Prueba que $\lim\{x_n\} = 0$.

   Aplicación. Dados $a \in]-1,1[$, $k \in \mathbb{N}$, prueba que $\lim_{n\to\infty}\{n^k a^n\} = 0$.

2. Estudia la convergencia de las sucesiones siguientes.

$$x_n = \frac{2n + (-1)^n(n+2)}{7n+3} \qquad x_n = n\left(\frac{1+(-1)^n}{3}\right)^n \qquad x_n = n^2\left(\frac{1+n}{3n}\right)^n$$

$$x_n = \sqrt[n]{a^n + b^n} \ \ (a > 0, b > 0) \qquad x_n = \sum_{k=1}^{n} \frac{1}{\sqrt{k+n^2}} \qquad x_n = \sqrt[3]{n+1} - \sqrt[3]{n}$$

$$x_n = \sqrt{n^2+3n+2} - n \qquad x_n = \left(\sqrt{n^2+\sqrt{n}} - n\right)\left(\sqrt{n+1} + \sqrt{2n}\right) \qquad x_n = \frac{n!}{n^n}$$

   Sugerencia. En algunos casos puede usarse el principio de las sucesiones encajadas o el ejercicio anterior.

3. Utiliza la desigualdad de las medias para probar que $\sqrt[n]{n} \leqslant \dfrac{n-2+2\sqrt{n}}{n}$. Deduce que $\lim \sqrt[n]{n} = 1$.

4. Sea $x_n = \dfrac{1\cdot 3\cdot 5\cdots(2n-1)}{2\cdot 4\cdot 6\cdots 2n}$. Prueba que $x_n < \dfrac{1}{\sqrt{2n+1}}$. Deduce que $\lim\{x_n\} = 0$.

5. Como consecuencia inmediata de la formula del binomio de Newton, la desigualdad $(1+x)^n \geqslant 1+nx$ es válida para todo $x>0$. Usa dicha desigualdad para probar que la sucesión $n\left(\sqrt[n]{n+1} - \sqrt[n]{n}\right)$ converge a 0.

6. Supongamos que $\{a_n\} \to 0$. Prueba que $\lim \dfrac{\sqrt{1+a_n}-1}{a_n} = \dfrac{1}{2}$.

7. Sean $a_0, a_1, \dots, a_p$ números reales cuya suma es igual a cero. Justifica que
$$\lim_{n\to\infty}\left\{a_0\sqrt{n}+a_1\sqrt{n+1}+a_2\sqrt{n+2}+\cdots+a_p\sqrt{n+p}\right\} = 0$$
Sugerencia. Saca factor común $\sqrt{n}$, resta $a_0+a_1+\cdots+a_p$ y usa el ejercicio anterior.

8. Dados $0<b_1<a_1$, definamos para todo $n\in\mathbb{N}$:
$$a_{n+1} = \frac{a_n+b_n}{2}, \quad b_{n+1} = \sqrt{a_n b_n}.$$
Justifica que las sucesiones así definidas son monótonas y convergen al mismo número (que se llama *media aritmético-geométrica* de a_1 y b_1).

9. Estudia la convergencia de las siguientes sucesiones.

   a) $x_1 = 1$, $x_{n+1} = \dfrac{4+3x_n}{3+2x_n}$.

   b) Dado $a>0$, definimos $x_1 = \sqrt{a}$, $x_{n+1} = \sqrt{a+x_n}$

   c) Dado $a>0$, $a\neq 1$, definimos $x_1 = a$, $x_{n+1} = \dfrac{1}{2}\left(x_n + \dfrac{a}{x_n}\right)$.

   Sugerencia. Estudia en cada caso monotonía y acotación.

10. *a)* Para cada $n\in\mathbb{N}$ sea $x_n = 1+\frac{1}{2}+\cdots+\frac{1}{n}-\log(n)$, $y_n = x_n - \frac{1}{n}$. Prueba que $\{x_n\}$ es estrictamente decreciente e $\{y_n\}$ es estrictamente creciente. Deduce que ambas sucesiones convergen a un mismo número. Dicho número se llama la *constante de Euler*, se representa por la letra griega γ. No se sabe si dicho número es racional o irracional.

    b) Deduce que $\lim\limits_{n\to\infty} \dfrac{1+1/2+\cdots+1/n}{\log(n)} = 1$.

    c) Justifica que
$$\lim_{n\to\infty}\left\{\frac{1}{n+1}+\frac{1}{n+2}+\cdots+\frac{1}{2n}\right\} = \log 2.$$

11. ¿Puede existir alguna sucesión acotada, $\{x_n\}$, verificando que $|x_n - x_m| \geqslant 10^{-75}$ siempre que $n\neq m$? Razona tu respuesta.

12. Sea $\{x_n\}$ una sucesión de números reales y supongamos que existen $\rho\in]0,1[$, $p\in\mathbb{N}$, tales que $|x_{n+2}-x_{n+1}| \leqslant \rho|x_{n+1}-x_n|$ para todo $n\geqslant p$. Prueba que $\{x_n\}$ es convergente.

Sugerencia. Deduce primero que $|x_{n+2}-x_{n+1}| \leqslant \rho^n |x_2-x_1|$. Teniendo ahora en cuenta que para todos $n,h\in\mathbb{N}$ se verifica que:

$$\rho^{n+h}+\rho^{n+h-1}+\rho^{n+h-2}+\cdots+\rho^n < \frac{\rho^n}{1-\rho}$$

deduce que $\{x_n\}$ verifica la condición de Cauchy.

13. Sea I un intervalo cerrado (puede ser $I=\mathbb{R}$); $f: I\to\mathbb{R}$ una función, y supongamos que hay un número $\alpha\in]0,1[$ tal que

$$|f(x)-f(y)| \leqslant \alpha|x-y|, \quad \text{para todos } x,y \text{ en } I.$$

Supongamos además que $f(x)\in I$ para todo $x\in I$. Dado un punto $a\in I$, definamos $\{x_n\}$ por $x_1=a$, y $x_{n+1}=f(x_n)$ para todo $n\in\mathbb{N}$. Prueba que $\{x_n\}$ converge a un punto $x\in I$ que es el único punto fijo de f, es decir, $f(x)=x$.

14. Sea k un número natural. Calcula el límite de la sucesión $\dfrac{1^k+2^k+3^k+\cdots+n^k}{n^{k+1}}$.

15. Dadas dos funciones polinómicas P,Q, tales que el grado de Q es mayor o igual que el grado de P y $Q(n)\neq 0$ para todo $n\in\mathbb{N}$, justifica que la sucesión $\left\{\dfrac{P(n)}{Q(n)}\right\}$ es convergente y calcula su límite.

Lección 6

Continuidad en intervalos cerrados y acotados. Límite funcional

Introducción

Sabemos ya que la imagen, $f(I)$, de un intervalo I por una función continua f es un intervalo. Nos preguntamos ¿Es $f(I)$ un intervalo del *mismo tipo* que I? Enseguida nos damos cuenta de que no tiene por qué ser así.

1. $f(x) = x^2$; $f([-1,1[) = f(]-1,1]) = [0,1]$;
2. $f(x) = 1/x$; $f(]0,1]) = [1,+\infty[$; $f([1,+\infty[) =]0,1]$.
3. $f(x) = \operatorname{sen} x$; $f(]-\pi,\pi[= [-1,1]$.

Vemos así que la imagen por una función continua de un intervalo abierto, o semiabierto, o de una semirrecta, puede ser un intervalo de distinto tipo. Nos queda por considerar qué ocurre con los intervalos cerrados (y acotados), es decir, los de la forma $[a,b]$. Nótese que si $f : [a,b] \to \mathbb{R}$ es continua, para probar que $f([a,b])$ es un intervalo cerrado y acotado basta probar que el conjunto $f([a,b])$ tiene máximo y mínimo, es decir, que hay números $u,v \in [a,b]$ tales que para todo $x \in [a,b]$ es $f(u) \leqslant f(x) \leqslant f(v)$, pues entonces será $f([a,b]) = [f(u), f(v)]$.

6.1 Definición. Sea $f : B \to \mathbb{R}$. Se dice que f *está acotada en* E si el conjunto $f(B)$ está acotado. Se dice que f alcanza en B un **máximo** (resp. un **mínimo**) **absoluto** si el conjunto $f(B)$ tiene máximo (resp. mínimo), es decir, existe algún punto $c \in B$ (resp. $b \in B$) tal que $f(x) \leqslant f(c)$ (resp. $f(b) \leqslant f(x)$) para todo $x \in B$.

El siguiente resultada es importante. Su demostración se propone como ejercicio.

6.2 Proposición. *Sea $f : A \to \mathbb{R}$ una función continua en un punto $x \in A$. Entonces para toda sucesión $\{x_n\}$ de puntos de A que converge a x se verifica que la sucesión $\{f(x_n)\}$ converge a $f(x)$.*

$$\lim f(x_n) = f(\lim x_n) = f(x)$$

6.3 Teorema (Teorema de Weierstrass). *Toda función continua en un intervalo cerrado y acotado alcanza en dicho intervalo un máximo y un mínimo absolutos. En particular, toda función continua en un intervalo cerrado y acotado está acotada en dicho intervalo.*

Demostración. Sea $f:[a,b]\to\mathbb{R}$ una función continua. Queremos probar que hay algún punto $c\in[a,b]$ en el que f alcanza un máximo absoluto.

Empezaremos probando que f está acotada en $[a,b]$. Razonamos por contradicción: supondremos que f no está acotada y llegaremos a una contradicción. Si f no está acotada entonces para cada $n\in\mathbb{N}$ tiene que haber algún punto $x_n\in[a,b]$ tal que $f(x_n)\geqslant n$. La sucesión $\{x_n\}$ está acotada y, por el teorema de Bolzano-Weierstrass, tiene alguna parcial $y_n=x_{\sigma(n)}$ convergente. Sea $x=\lim\{x_n\}$. Como $a\leqslant y_n\leqslant b$ se deduce que $a\leqslant x\leqslant b$ y por tanto $x\in[a,b]$ (aquí es donde se usa el hecho de que $[a,b]$ es un intervalo cerrado). Usando la proposición anterior y la continuidad de f en x obtenemos que la sucesión $\{f(y_n)\}$ debe ser convergente. Pero esto es imposible porque $f(y_n)=f(x_{\sigma(n)})\geqslant\sigma(n)\geqslant n$ lo que nos dice que la sucesión $\{f(y_n)\}$ no está acotada y por tanto no puede ser convergente.

Como f está acotada, el conjunto $f[a,b]$ es un conjunto acotado de números reales y, por tanto, tiene supremo e ínfimo. Sea $\beta=\sup f[a,b]$. Probaremos que $\beta\in f[a,b]$. Para cada $n\in\mathbb{N}$ tiene que existir un $z_n\in[a,b]$ tal que $\beta-1/n\leqslant f(z_n)\leqslant\beta$. Repitiendo el razonamiento anterior obtenemos que la sucesión $\{z_n\}$ tiene alguna sucesión parcial convergente. Sin perder generalidad, podemos suponer que la propia $\{z_n\}$ es convergente. Sea $z=\lim\{z_n\}$. Entonces $z\in[a,b]$ (por ser un intervalo cerrado) y $f(z)=\lim f(z_n)=\beta$. En consecuencia f alcanza un máximo absoluto en el punto z.

Análogamente se prueba que $\alpha=\inf f[a,b]\in f[a,b]$.

Como aplicación del teorema de Weierstrass puede probarse el siguiente resultado.

6.4 Proposición. *Una función polinómica de grado par* $f(x)=\displaystyle\sum_{k=0}^{n}a_kx^k$ *alcanza un mínimo absoluto en* $\mathbb{R}$ *si el coeficiente líder es positivo,* $a_n>0$*, y alcanza un máximo absoluto en* $\mathbb{R}$ *si el coeficiente líder es negativo,* $a_n<0$.

6.1. Límite funcional

Sea I un intervalo, a un punto de I, y f una función definida en $I\setminus\{a\}$. Naturalmente, como f no está definida en a no tiene sentido hablar de la continuidad de f en a. Sin embargo, podemos preguntarnos ¿es posible encontrar un número $L\in\mathbb{R}$ tal que *definiendo* $f(a)=L$, la *nueva función* así obtenida sea continua en a? Para ello el número L tendría que cumplir la siguiente propiedad:

$$\forall\varepsilon\in\mathbb{R}^+\quad\exists\delta\in\mathbb{R}^+\ :\ \left.\begin{array}{c}0<|x-a|<\delta\\ x\in I\end{array}\right\}\Longrightarrow\quad|f(x)-L|<\varepsilon$$

donde la condición “$0<|x-a|$” es obligada porque la función f no está definida en a.

Podemos modificar un poco la situación anterior, suponiendo ahora que f está definida en todo el intervalo I pero no es continua en a. En este caso queremos cambiar el valor de f en a, es decir, encontrar, si es posible, un número $L \in \mathbb{R}$ tal que *definiendo* el valor de f en a igual a L, la *nueva función* así obtenida sea continua en a. La condición que tiene que cumplir dicho número L es exactamente la misma de antes.

Nótese que ahora la condición "$0 < |x-a|$" es obligada porque nuestra función f no está definida en a de "forma apropiada".

En los dos casos considerados la condición obtenida es la misma con independencia del hecho de que f esté o no definida en a y, en caso de estarlo, del posible valor que f pueda tener en a. Por ello, en lo que sigue consideraremos la siguiente situación.

Notación. En adelante, representaremos por I un intervalo; a será un punto de I, y f será una función que supondremos definida en $I \setminus \{a\}$ sin excluir la posibilidad de que dicha función pueda estar definida en todo el intervalo I lo cual, para nuestros propósitos actuales, carece de importancia.

6.5 Definición. Se dice que f tiene límite en el punto a si existe un número $L \in \mathbb{R}$ tal que se verifica lo siguiente:

$$\forall \varepsilon \in \mathbb{R}^+ \quad \exists \delta \in \mathbb{R}^+ : \quad \left.\begin{array}{c} 0 < |x-a| < \delta \\ x \in I \end{array}\right\} \Longrightarrow \quad |f(x) - L| < \varepsilon$$

Dicho número se llama **límite de** f **en** a y escribimos $\lim\limits_{x \to a} f(x) = L$.

Observa que la existencia del límite es independiente de que f esté o no definida en a y, en caso de estarlo, del valor que f pueda tener en a. También debe advertirse que en la definición de la *igualdad* $\lim\limits_{x \to a} f(x) = L$, sólo intervienen *desigualdades.*

Es fácil probar que el límite de una función en un punto, si existe, es único. Una consecuencia inmediata de la definición dada es el siguiente resultado.

6.6 Proposición. *Sea $f : I \to \mathbb{R}$ una función definida en un intervalo y sea $a \in I$. Equivalen las afirmaciones siguientes:*

i) f es continua en a.

ii) $\lim\limits_{x \to a} f(x) = f(a)$.

En la recta real es posible distinguir si nos acercamos "por la derecha" o "por la izquierda" a un punto. Ello conduce de forma natural a la consideración de los *límites laterales* que pasamos a definir.

Límites laterales de una función en un punto

Supongamos que:

A) El conjunto $\{x \in I : a < x\}$ no es vacío. En tal caso, se dice que f tiene *límite por la derecha* en a, si existe un número $\alpha \in \mathbb{R}$ tal que se verifica lo siguiente:

$$\forall \varepsilon \in \mathbb{R}^+ \quad \exists \delta \in \mathbb{R}^+ : \quad \left.\begin{array}{c} a < x < a + \delta \\ x \in I \end{array}\right\} \Longrightarrow \quad |f(x) - \alpha| < \varepsilon$$

Dicho número se llama **límite por la derecha de** f **en** a y, simbólicamente, escribimos $\displaystyle\lim_{\substack{x\to a\\ x>a}} f(x)=\alpha$.

B) El conjunto $\{x\in I : x<a\}$ no es vacío. En tal caso, se dice que f tiene *límite por la izquierda* en a, si existe un número $\beta\in\mathbb{R}$ tal que se verifica lo siguiente:

$$\forall\varepsilon\in\mathbb{R}^+\quad\exists\delta\in\mathbb{R}^+ :\quad \left.\begin{array}{c} a-\delta<x<a\\ x\in I\end{array}\right\}\Longrightarrow\quad |f(x)-\beta|<\varepsilon$$

Dicho número se llama **límite por la izquierda de** f **en** a y, simbólicamente, escribimos $\displaystyle\lim_{\substack{x\to a\\ x<a}} f(x)=\beta$.

Teniendo en cuenta las definiciones dadas, es inmediato que:

i) Si $a=\sup I$, entonces $\displaystyle\lim_{x\to a} f(x)=\lim_{\substack{x\to a\\ x<a}} f(x)$.

ii) Si $a=\inf I$, entonces $\displaystyle\lim_{x\to a} f(x)=\lim_{\substack{x\to a\\ x>a}} f(x)$.

iii) Si a no es un extremo de I, entonces equivalen las afirmaciones:

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

b) $\displaystyle\lim_{\substack{x\to a\\ x<a}} f(x)=\lim_{\substack{x\to a\\ x>a}} f(x)=L$.

6.2. Límites infinitos

Funciones divergentes en un punto

Se dice que f **es positivamente divergente** en a si se verifica lo siguiente:

$$\forall M\in\mathbb{R}^+\quad\exists\delta\in\mathbb{R}^+ :\quad \left.\begin{array}{c} 0<|x-a|<\delta\\ x\in I\end{array}\right\}\Longrightarrow\quad f(x)>M$$

Simbólicamente, escribimos $\displaystyle\lim_{x\to a} f(x)=+\infty$.

Se dice que f **es positivamente divergente por la izquierda** en a si se verifica lo siguiente:

$$\forall M\in\mathbb{R}^+\quad\exists\delta\in\mathbb{R}^+ :\quad \left.\begin{array}{c} a-\delta<x<a\\ x\in I\end{array}\right\}\Longrightarrow\quad f(x)>M$$

Simbólicamente, escribimos $\displaystyle\lim_{\substack{x\to a\\ x<a}} f(x)=+\infty$.

De forma análoga se definen los conceptos:

- "f **es positivamente divergente** por la derecha en a". Simbólicamente $\displaystyle\lim_{\substack{x\to a\\ x>a}} f(x)=+\infty$
- "f **es negativamente divergente** en a". Simbólicamente $\displaystyle\lim_{x\to a} f(x)=-\infty$.
- "f **es negativamente divergente por la izquierda o por la derecha** en a". Simbólicamente $\displaystyle\lim_{\substack{x\to a\\ x<a}} f(x)=-\infty\qquad\lim_{\substack{x\to a\\ x>a}} f(x)=-\infty$

Límites en infinito

Sea $f : I \to \mathbb{R}$ una función definida en un intervalo *no mayorado* I. Se dice que f tiene límite en $+\infty$ si existe un número $L \in \mathbb{R}$ tal que se verifica lo siguiente:

$$\forall \varepsilon \in \mathbb{R}^+ \quad \exists K \in \mathbb{R}^+ : \quad \left.\begin{array}{c} x > K \\ x \in I \end{array}\right\} \Longrightarrow \quad |f(x) - L| < \varepsilon$$

Dicho número se llama límite de f en $+\infty$ y escribimos $\lim\limits_{x \to +\infty} f(x) = L$.

Análogamente se define el límite en $-\infty$.

Funciones divergentes en infinito

Sea $f : I \to \mathbb{R}$ una función definida en un intervalo *no mayorado* I. Se dice que f es positivamente divergente en $+\infty$ si se verifica lo siguiente:

$$\forall M \in \mathbb{R}^+ \quad \exists K \in \mathbb{R}^+ : \quad \left.\begin{array}{c} x > K \\ x \in I \end{array}\right\} \Longrightarrow \quad f(x) > M$$

En cuyo caso escribimos $\lim\limits_{x \to +\infty} f(x) = +\infty$.

Llegados aquí, el lector no tendrá dificultad en precisar el significado de:

$$\lim_{x \to +\infty} f(x) = -\infty, \qquad \lim_{x \to -\infty} f(x) = +\infty, \qquad \lim_{x \to -\infty} f(x) = -\infty.$$

El siguiente resultado establece una importantísima relación entre el límite funcional y el límite de sucesiones.

6.7 Proposición. *Sea f una función y sean $a, L \in \mathbb{R} \cup \{+\infty, -\infty\}$. Equivalen las afirmaciones:*

i) $\lim\limits_{x \to a} f(x) = L$

ii) Para toda sucesión $\{x_n\}$ de puntos en el dominio de definición de f, tal que $\{x_n\} \to a$ con $x_n \neq a$, se verifica que $\{f(x_n)\} \to L$.

6.3. Discontinuidades. Álgebra de límites. Límites de funciones monótonas

Clasificación de las discontinuidades

Sea $f : I \to \mathbb{R}$ una función definida en un intervalo y sea $a \in I$.

- Si f tiene límite en a y $\lim\limits_{x \to a} f(x) \neq f(a)$, se dice que f tiene en el punto a una **discontinuidad evitable**.

- Si los dos límites laterales de f en a existen y son distintos:
$$\lim_{\substack{x\to a\\x<a}} f(x) \neq \lim_{\substack{x\to a\\x>a}} f(x)$$
se dice que f tiene en el punto a una **discontinuidad de salto**.

- Si alguno de los límites laterales no existe se dice que f tiene en el punto a una **discontinuidad esencial**.

Es evidente que el concepto de límite es, al igual que el de continuidad en un punto, un concepto local; la existencia del límite de una función en un punto a depende solamente del comportamiento de la función en los puntos próximos al punto a.

Es importante advertir que el concepto de límite lateral es un caso particular del concepto general de límite de una función en un punto. Por ello, cualquier resultado referente a límites de funciones en un punto puede ser convenientemente enunciado para límites laterales sin más que considerar la restricción de la función a la derecha o a la izquierda del punto en cuestión.

El siguiente resultado pone de manifiesto la compatibilidad de la "operación de paso al límite" con la estructura algebraica y de orden de $\mathbb{R}$.

6.8 Teorema (Álgebra de límites). *Supongamos que f y g tienen límite en a donde aceptamos que a puede ser un número real, o $+\infty$, o $-\infty$. Se verifica entonces que:*

i) Las funciones $f+g$ y fg tienen límite en a y
$$\lim_{x\to a}(f+g)(x) = \lim_{x\to a} f(x) + \lim_{x\to a} g(x), \quad \lim_{x\to a}(fg)(x) = \lim_{x\to a} f(x)\,\lim_{x\to a} g(x)$$

ii) Si $\lim\limits_{x\to a} f(x) \neq 0$, entonces $\lim\limits_{x\to a} \dfrac{1}{f(x)} = \dfrac{1}{\lim\limits_{x\to a} f(x)}$

iii) Si $f(x) \leqslant g(x)$ para todo $x \in I$, $x \neq a$, entonces $\lim\limits_{x\to a} f(x) \leqslant \lim\limits_{x\to a} g(x)$

iv) Supongamos que $f(x) \leqslant h(x) \leqslant g(x)$ para todo $x \in I$, $x \neq a$ y $\lim\limits_{x\to a} f(x) = \lim\limits_{x\to a} g(x) = L$. Entonces se verifica que h tiene límite en a y $\lim\limits_{x\to a} h(x) = L$.

En el siguiente resultado se establecen condiciones que garantizan la divergencia de una suma o de un producto.

6.9 Teorema. *Supongamos que f es positivamente divergente en a, $\lim\limits_{x\to a} f(x) = +\infty$, donde aceptamos que a puede ser un número real, o $+\infty$, o $-\infty$.*

i) Supongamos que hay un número $M \in \mathbb{R}$ tal que $g(x) \geqslant M$ para todo $x \in I$, $x \neq a$. Entonces $\lim\limits_{x\to a}(f+g)(x) = +\infty$.

ii) Supongamos que hay un número $M > 0$ tal que $g(x) \geqslant M$ para todo $x \in I$, $x \neq a$. Entonces $\lim\limits_{x\to a}(fg)(x) = +\infty$.

En el siguiente resultado se establecen condiciones que garantizan que un producto tenga límite igual a cero.

6.10 Teorema. *Supongamos que* $\lim_{x\to a} f(x) = 0$, *y que hay un número* $M > 0$ *tal que* $|g(x)| \leqslant M$ *para todo* $x \in I$, $x \neq a$. *Entonces* $\lim_{x\to a}(fg)(x) = 0$.

Un resultado establece que la continuidad permuta con el paso al límite.

6.11 Teorema. *Si* g *es continua en el punto* $L = \lim_{x\to a} f(x)$, *entonces* $\lim_{x\to a}(g\circ f)(x) = g(L)$. *Simbólicamente:*

$$\lim_{x\to a}(g\circ f)(x) = g(\lim_{x\to a} f(x))$$

6.4. Continuidad y monotonía

6.12 Teorema (Límites de una función monótona). *Sea* f *una función creciente definida en un intervalo* I.

i) Para todo punto $a \in I$ *que no sea un extremo de* I *se verifica que:*

$$\lim_{\substack{x\to a\\ x<a}} f(x) = \sup\{f(x) : x\in I, x < a\} \qquad \lim_{\substack{x\to a\\ x>a}} f(x) = \inf\{f(x) : x\in I, x > a\}$$

ii) Si $a \in \mathbb{R} \cup \{-\infty\}$ *es el extremo izquierdo de* I, *entonces:*

a) Si f *está minorada en* I *es* $\lim_{x\to a} f(x) = \inf\{f(x) : x \in I \setminus \{a\}\}$.

b) Si f *no está minorada en* I *es* $\lim_{x\to a} f(x) = -\infty$.

iii) Si $a \in \mathbb{R} \cup \{+\infty\}$ *es el extremo derecho de* I, *entonces:*

a) Si f *está mayorada en* I *es* $\lim_{x\to a} f(x) = \sup\{f(x) : x \in I \setminus \{a\}\}$.

b) Si f *no está mayorada en* I *es* $\lim_{x\to a} f(x) = +\infty$.

Demostración. Supongamos que $a\in I$ no es el extremo izquierdo de I, es decir que el conjunto $\{x\in I : x < a\}$ no es vacío. Entonces, el conjunto $B = \{f(x) : x\in I, x < a\}$ tampoco es vacío y, por ser f creciente, el número $f(a)$ es un mayorante de B. Sea $\alpha = \sup\{f(x) : x\in I, x < a\}$. Dado $\varepsilon > 0$, el número $\alpha - \varepsilon$ no puede ser mayorante de B, es decir, tiene que haber algún punto $x_o \in I$, $x_o < a$ tal que $\alpha - \varepsilon < f(x_o)$. Sea $\delta = a - x_o > 0$. Entonces para $a - \delta < x < a$, esto es, para $x_o < x < a$, se verifica que $\alpha - \varepsilon < f(x_o) \leqslant f(x) \leqslant \alpha$, lo que claramente implica que $\alpha - \varepsilon < f(x) < \alpha + \varepsilon$, es decir, $|f(x) - \alpha| < \varepsilon$. Hemos probado así que $\lim_{\substack{x\to a\\ x<a}} f(x) = \sup\{f(x) : x\in I, x < a\}$. Los demás casos se prueban de forma muy parecida y quedan como ejercicios. $\square$

6.13 Teorema (Discontinuidades de las funciones monótonas). *Sea* f *una función monótona en un intervalo. Entonces:*

i) En los puntos del intervalo que no son extremos del mismo, f *solamente puede tener discontinuidades de salto.*

ii) Si el intervalo tiene máximo o mínimo, f *puede tener en dichos puntos discontinuidades evitables.*

6.14 Teorema (Continuidad de una función monótona). *Una función monótona definida en un intervalo es continua si, y sólo si, su imagen es un intervalo.*

Demostración. En efecto, si $f : I \to \mathbb{R}$ es una función creciente en un intervalo I y suponemos que su imagen $f(I)$ es un intervalo entonces, si $a \in I$ no es un punto extremo de I, es decir, hay puntos $u, v \in I$ tales que $u < a < v$, tenemos que

$$\{f(x) : x \in I, x < a\} \supset [f(u), f(a)[, \quad \{f(x) : x \in I, x > a\} \supset]f(a), f(v)]$$

y deducimos que

$$\lim_{\substack{x\to a\\ x<a}} f(x) = f(a) = \lim_{\substack{x\to a\\ x>a}} f(x)$$

esto es, f es continua en a. Análogamente se prueba que si I contiene a alguno de sus extremos entonces f es continua también en esos puntos.

Teniendo en cuenta que la función inversa de una función estrictamente monótona es también estrictamente monótona (y del mismo tipo), se deduce de lo anterior el siguiente importante resultado.

6.15 Teorema. *La función inversa de una función estrictamente monótona y continua en un intervalo es también una función continua y estrictamente monótona.*

El siguiente resultado se demuestra haciendo uso del teorema de los ceros de Bolzano y será usado en el próximo capítulo para obtener una importante propiedad de las funciones con derivada distinta de cero. Su demostración no añade nada nuevo a lo que ya sabemos y por eso no la incluyo aquí; pero lo que se afirma en él es muy intuitivo: si una función es *continua e inyectiva en un intervalo* entonces es claro que *su gráfica no puede subir y bajar*, en consecuencia o siempre sube o siempre baja.

6.16 Teorema (Funciones continuas e inyectivas en intervalos). *Una función continua e inyectiva definida en un intervalo es estrictamente monótona.*

6.5. Indeterminaciones en el cálculo de límites

Frecuentemente hay que estudiar el límite de una suma o producto de dos funciones precisamente cuando las reglas que hemos visto anteriormente no pueden aplicarse. Se trata de aquellos casos en que el comportamiento de las funciones $f+g$, fg, no está determinado por el de f y g. Por ejemplo, si sabemos que $\lim_{x\to a} f(x) = +\infty$ y que $\lim_{x\to a} g(x) = -\infty$, ¿qué podemos decir en general del comportamiento en el punto a de la función $f+g$? Respuesta: absolutamente nada. En consecuencia, para calcular un límite del tipo $\lim_{x\to a}(f+g)(x)$ donde $\lim_{x\to a} f(x) = +\infty$ y $\lim_{x\to a} g(x) = -\infty$ *se requiere un estudio particular en cada caso.* Suele decirse que estos límites son **una indeterminación del tipo** "$\infty - \infty$".

Análogamente, si sabemos que $\lim_{x\to a} f(x) = 0$ y que la función g es divergente (positivamente o negativamente) en el punto a, ello no proporciona ninguna información sobre el comportamiento de la función fg en dicho punto. Cuando esto ocurre se dice que el límite $\lim_{x\to a}(fg)(x)$

es una indeterminación del tipo "0∞". Las indeterminaciones que aparecen al estudiar el cociente de dos funciones divergentes o de dos funciones con límite cero, es decir, las llamadas **indeterminaciones de los tipos** "∞/∞", "$0/0$", pueden reducirse a una indeterminación del tipo "0∞".

Todavía hemos de considerar nuevas indeterminaciones que van a surgir al considerar funciones de la forma $f(x)^{g(x)}$ donde f es una función que toma valores positivos y g es una función cualquiera. Puesto que:

$$f(x)^{g(x)} = \exp(g(x)\log f(x))$$

teniendo en cuenta los resultados anteriores, el límite $\lim\limits_{x\to a} f(x)^{g(x)}$ vendrá determinado por el límite $\lim\limits_{x\to a} g(x)\log f(x)$, el cual, a su vez, está determinado en todos los casos por el comportamiento en el punto a de las funciones f y g, excepto cuando dicho límite es una indeterminación del tipo "0∞", lo que ocurre en los siguientes casos:

a) $\lim\limits_{x\to a} f(x) = 1$, $\lim\limits_{x\to a} |g(x)| = +\infty$ (indeterminación "1^∞")

b) $\lim\limits_{x\to a} f(x) = +\infty$, $\lim\limits_{x\to a} g(x) = 0$ (indeterminación "∞^0")

c) $\lim\limits_{x\to a} f(x) = 0$, $\lim\limits_{x\to a} g(x) = 0$ (indeterminación "0^0")

Ni que decir tiene que no hay técnicas generales que permitan "resolver las indeterminaciones", ¡no serían tales si las hubiera! Es por ello que, los límites indeterminados, requieren un estudio particular en cada caso. Es un hecho que la mayoría de los límites que tienen algún interés matemático son límites indeterminados. Cuando estudiemos las derivadas obtendremos técnicas que permitirán calcular con comodidad dichos límites.

Límites de exponenciales y logaritmos

Los resultados que siguen son de gran utilidad para calcular límites. Su justificación se verá más adelante.

6.17 Proposición. *Sea a un número real o $+\infty$ o $-\infty$. En los apartados b1), b2) y b3) se supone que $f(x) > 0$.*

a1) $\lim\limits_{x\to a} f(x) = L \Longleftrightarrow \lim\limits_{x\to a} \mathrm{e}^{f(x)} = \mathrm{e}^L$.

a2) $\lim\limits_{x\to a} f(x) = +\infty \Longleftrightarrow \lim\limits_{x\to a} \mathrm{e}^{f(x)} = +\infty$.

a3) $\lim\limits_{x\to a} f(x) = -\infty \Longleftrightarrow \lim\limits_{x\to a} \mathrm{e}^{f(x)} = 0$.

b1) $\lim\limits_{x\to a} f(x) = L > 0 \Longleftrightarrow \lim\limits_{x\to a} \log f(x) = \log L$.

b2) $\lim\limits_{x\to a} f(x) = +\infty \Longleftrightarrow \lim\limits_{x\to a} \log f(x) = +\infty$.

b3) $\lim\limits_{x\to a} f(x) = 0 \Longleftrightarrow \lim\limits_{x\to a} \log f(x) = -\infty$.

El siguiente resultado es de gran importancia. En él se comparan los "*órdenes de crecimiento*" de las funciones logaritmo, potencias y exponenciales, resultando que el logaritmo crece más lentamente que cualquier potencia positiva y éstas, a su vez, crecen más lentamente que la exponencial.

6.18 Proposición. I) $\displaystyle\lim_{x\to+\infty}\frac{|\log x|^{\mu}}{x^{\alpha}}=0$ *para todos* $\alpha>0$ *y* $\mu\in\mathbb{R}$.

II) $\displaystyle\lim_{x\to 0}|x|^{\alpha}\big|\log|x|\big|^{\mu}=0$ *para todos* $\alpha>0$ *y* $\mu\in\mathbb{R}$.

III) $\displaystyle\lim_{x\to+\infty}\frac{x^{\alpha}}{e^{\mu x}}=0$ *para todos* $\alpha>0$ *y* $\mu>0$.

6.6. Ejercicios

1. Sea $f:[a,b]\to\mathbb{R}$ continua. Supongamos que para cada $x\in[a,b]$ hay algún $y\in[a,b]$ tal que $|f(y)|\leqslant\frac{9}{10}|f(x)|$. Prueba que f se anula en algún punto de $[a,b]$.

2. Sea $f:[a,b]\to\mathbb{R}$ continua. Prueba que la función $g:[a,b]\to\mathbb{R}$ dada por $g(x)=\max f([a,x])$, $(a\leqslant x\leqslant b)$, es continua.

3. Sea $f:[a,b]\to\mathbb{R}$ continua, pongamos $M=\max f([a,b])$, $m=\min f([a,b])$ y supongamos que $f(a)=f(b)$ y que $m<f(a)<M$. Prueba que f toma todo valor de $[f(a),M[\cup]m,f(a)]$ en al menos dos puntos de $[a,b]$.

4. La ecuación $ax^2+2x-1=0$ donde $a>-1$, $a\neq 0$ tiene dos soluciones que representaremos por $\lambda(a)$ y por $\mu(a)$. Calcula los límites de dichas funciones en $a=0$ y en $a=-1$.

5. Prueba que, dado $x\in\mathbb{R}$, la ecuación $\log t+t^5=x$ tiene una única solución, que representamos por $\varphi(x)$. Justifica que la función $x\mapsto\varphi(x)$, $(x\in\mathbb{R})$, así definida es continua.

6. Sea $f:[0,1]\to\mathbb{R}$ continua verificando que $|f(s)-f(t)|\geqslant|s-t|$ para todos $s,t\in[0,1]$, y $f(\{0,1\})=\{0,1\}$. Prueba que o bien es $f(x)=x$ para todo $x\in[0,1]$, o bien es $f(x)=1-x$ para todo $x\in[0,1]$.

7. Estudia los límites en $+\infty$ y en $-\infty$ de:

   a) Una función polinómica.

   b) Una función racional.

8. Justifica que una función polinómica de grado par o bien alcanza un máximo en $\mathbb{R}$ o bien alcanza un mínimo absoluto en $\mathbb{R}$.

Lección 7

Derivadas

Introducción

Los orígenes del Cálculo estuvieron motivados por el deseo de resolver diversos problemas vinculados al movimiento de los cuerpos, así como problemas de tipo geométrico de importancia en Óptica y problemas de cálculo de valores máximos y mínimos de una función dada. Simplificando podemos destacar dos problemas principales:

- Determinar la tangente a una curva en un punto (el problema de las tangentes).
- Determinar el área encerrada por una curva (el problema de las cuadraturas).

Son los conceptos de derivada e integral, respectivamente, los que permiten resolver satisfactoriamente dichos problemas. Mientras que el concepto de integral tiene sus raíces en la antigüedad clásica, la otra idea fundamental del Cálculo, la derivada, no se formuló hasta el siglo XVII. Fue el descubrimiento efectuado por Sir Isaac Newton (1642-1727) y Gottfried Wilhelm Leibnitz (1646-1716) de la relación entre estas dos ideas, tan dispares en apariencia, lo que inició el magnífico desarrollo del Cálculo. Si bien los trabajos de Newton y Leibnitz son decisivos por sus aportaciones e influencia, no hay que olvidar que ellos son el punto culminante de un largo proceso en el que han participado científicos de la talla de Johannes Kepler (1571-1630), René Descartes (1596-1650), Pierre de Fermat (1601-1665), John Wallis (1616-1703) e Isaac Barrow (1630-1677) entre otros.

7.1.1. Concepto de derivada. Interpretación física y geométrica

Para entender los resultados del Cálculo diferencial es necesario, antes que nada, comprender la idea básica del mismo: el concepto de derivada. La derivada de una función puede interpretarse geométricamente como la pendiente de una curva, y físicamente como una razón "instantánea" de cambio.

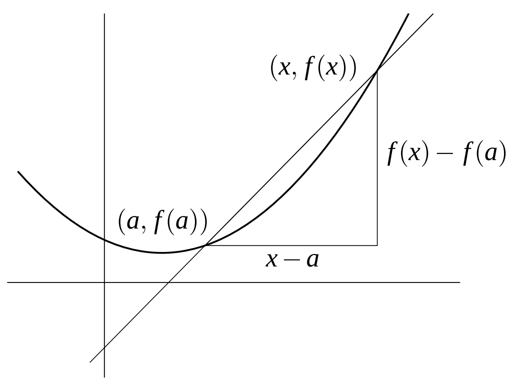
Tangente a una curva

A principios del siglo XVII no se sabía cómo calcular la tangente a una curva en un punto de la misma. Este problema se presentaba con frecuencia en mecánica, en óptica y en geometría.

Vamos a estudiar el concepto general de tangente a una curva en un punto dado. En general, no es un asunto sencillo hallar la pendiente de esta tangente. La razón es que, en principio, se necesita para ello otro punto, además del de tangencia. Supongamos que queremos hallar la tangente a la curva de ecuación cartesiana $y = f(x)$ en el punto $(a, f(a))$. La estrategia, usada primero por Pierre de Fermat y más tarde por Newton, consiste en aproximar la tangente por rectas secantes cuyas pendientes sí pueden calcularse directamente. En particular, considérese la recta que une el punto $(a, f(a))$ con un punto cercano, $(x, f(x))$, de la gráfica de f. Esta recta se llama una secante (recta que corta, pero no es tangente a la curva). La pendiente de esta secante es:

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

dicho número suele llamarse *cociente incremental de f en a*.



Nótese que una secante es una buena aproximación de la tangente, siempre que el punto $(x, f(x))$ esté muy próximo a $(a, f(a))$. Estas consideraciones llevan a *definir la tangente a la gráfica de f en el punto $(a, f(a))$ como la recta que pasa por dicho punto y cuya pendiente es igual al límite*:

$$\lim_{x\to a}\frac{f(x)-f(a)}{x-a}$$

supuesto, claro está, que dicho límite exista.

Razón de cambio

Muchas leyes de la Física, la Química, la Biología o la Economía, son funciones que relacionan una variable "dependiente" y con otra variable "independiente" x, lo que suele escribirse en la forma $y = f(x)$. Si la variable independiente cambia de un valor inicial a a otro x, la variable y lo hace de $f(a)$ a $f(x)$. La *razón de cambio promedio de $y = f(x)$ con respecto a x en el intervalo* $[a,x]$ es:

$$\text{Razón de cambio promedio } = \frac{f(x)-f(a)}{x-a}$$

Con frecuencia interesa considerar la razón de cambio en intervalos cada vez más pequeños. Esto lleva a definir lo que podemos llamar "*razón de cambio puntual de $y = f(x)$ con respecto a x en el punto a*" como:

$$\lim_{x\to a}\frac{f(x)-f(a)}{x-a}.$$

El ejemplo más conocido de esto que decimos es el de una partícula que se mueve a lo largo de una recta sobre la cual hemos elegido un origen. Sea $f(t)$ la distancia de la partícula al origen en el tiempo t. La razón de cambio promedio tiene en este caso una interpretación física natural. Es la *velocidad media* de la partícula durante el intervalo de tiempo considerado.

Parece intuitivo que, en cada instante, la partícula se mueve con una determinada *velocidad instantánea*. Pero la definición corriente de velocidad es en realidad una definición de velocidad media; la única definición razonable de velocidad instantánea es como la razón de cambio puntual. Es importante darse cuenta de que la velocidad instantánea es un concepto teórico, y una abstracción, que no corresponde exactamente a ninguna cantidad observable.

Notación. En lo que sigue las letras I, J representan intervalos no vacíos de números reales.

7.1 Definición. Se dice que **una función** $f : I \to \mathbb{R}$ **es derivable en un punto** $a \in I$, si existe el límite:

$$\lim_{x\to a} \frac{f(x)-f(a)}{x-a}.$$

Explícitamente, f es derivable en a si hay un número $L\in\mathbb{R}$ verificando que para cada número $\varepsilon > 0$ existe algún número $\delta > 0$ tal que para todo $x \in I$ con $x \neq a$ y $|x-a| < \delta$ se tiene que:

$$\left| \frac{f(x)-f(a)}{x-a} - L \right| \leqslant \varepsilon.$$

Dicho número L se llama **la derivada de** f **en** a y suele representarse por $f'(a)$ (notación debida a Lagrange) y también, a veces, por $\left.\frac{d f(x)}{dx}\right|_{x=a}$ (notación de Leibnitz).

Observaciones

i) El límite $\displaystyle\lim_{x\to a} \frac{f(x)-f(a)}{x-a}$ se escribe también en la forma $\displaystyle\lim_{h\to 0} \frac{f(a+h)-f(a)}{h}$.

ii) La derivabilidad de f en un punto $a \in I$ es una *propiedad local*, depende solamente del comportamiento de f en los puntos de I próximos al punto a. Concretamente, si J es cualquier *intervalo abierto* que contiene el punto a, se verifica que f es derivable en a si, y sólo si, la función restricción $f_{|I\cap J}$ es derivable en a y, por supuesto, en tal caso ambas funciones tienen la misma derivada en a.

7.1.2. Derivadas laterales

7.2 Definición. Se dice que f **es derivable por la izquierda en** a si existe el límite:

$$\lim_{\substack{x\to a\\ x<a}} \frac{f(x)-f(a)}{x-a}.$$

El valor de dicho límite se llama la **derivada por la izquierda de** f **en** a.

Análogamente se dice que f **es derivable por la derecha en** a, si existe el límite:

$$\lim_{\substack{x\to a\\ x>a}} \frac{f(x)-f(a)}{x-a}.$$

El valor de dicho límite se llama la **derivada por la derecha de** f **en** a.

Teniendo en cuenta la relación que hay entre el límite de una función en un punto y los límites laterales, es claro que:

i) Si $a = \max I$, entonces la derivabilidad de f en a es lo mismo que la derivabilidad por la izquierda de f en a.

ii) Si $a = \min I$, entonces la derivabilidad de f en a es lo mismo que la derivabilidad por la derecha de f en a.

iii) Si a no es un extremo de I, entonces equivalen las afirmaciones:

a) f es derivable en a.

b) Las derivadas por la izquierda y por la derecha de f en a existen y coinciden.

El siguiente resultado nos dice que la derivabilidad es una propiedad más fuerte que la continuidad.

7.3 Proposición. *Toda función derivable en un punto es continua en dicho punto.*

En efecto, si $f : I \to \mathbb{R}$ es derivable en a, de la igualdad:

$$f(x) = f(a) + (x-a)\frac{f(x)-f(a)}{x-a} \qquad (x \in I,\ x \neq a)$$

se sigue que $\lim_{x\to a} f(x) = f(a)$, es decir, f es continua en a.

7.4 Teorema (Reglas de derivación). *Sean $f\,g : I \to \mathbb{R}$ dos funciones. Se verifican las siguientes afirmaciones:*

i) La función suma $f+g$ y la función producto fg son derivables en todo punto $a \in I$ en el que f y g sean derivables; en tal caso las derivadas respectivas vienen dadas por:

$$(f+g)'(a) = f'(a) + g'(a); \qquad (fg)'(a) = f'(a)g(a) + f(a)g'(a)$$

ii) Si $g(x) \neq 0$ para todo $x \in I$, la función cociente f/g es derivable en todo punto $a \in I$ en el que f y g sean derivables en cuyo caso se verifica que:

$$\left(\frac{f}{g}\right)'(a) = \frac{f'(a)g(a) - f(a)g'(a)}{(g(a))^2}$$

7.5 Corolario. *Las funciones polinómicas son derivables en todo punto y las funciones racionales son derivables en todo punto de su conjunto natural de definición. Además la derivada de la función polinómica $f(x) = a_0 + a_1x + a_2x^2 + \cdots + a_nx^n$ en cada punto $x \in \mathbb{R}$ viene dada por:*

$$f'(x) = a_1 + 2a_2x + 3a_3x^2 + \cdots + na_nx^{n-1}$$

7.6 Teorema (Derivación de una función compuesta o regla de la cadena). *Sean $f : I \to \mathbb{R}$ y $g : J \to \mathbb{R}$ con $f(I) \subseteq J$, y sea $h = g \circ f : I \to \mathbb{R}$ la función compuesta. Supongamos que f es derivable en $a \in I$ y que g es derivable en $f(a)$. Entonces h es derivable en a y $h'(a) = g'(f(a))f'(a)$.*

En particular, si g es derivable en J, la función compuesta $h = g \circ f$ es derivable en todo punto de I donde f sea derivable.

Demostración. Pongamos $b = f(a)$. Tenemos que probar que $\displaystyle\lim_{x\to a}\frac{h(x)-h(a)}{x-a} = g'(b)f'(a)$. Por hipótesis se cumple que :

$$\lim_{y\to b}\frac{g(y)-g(b)}{y-b}\lim_{x\to a}\frac{f(x)-f(a)}{x-a} = g'(b)f'(a)$$

La idea de la demostración es hacer en esta igualdad la sustitución $y = f(x)$. Como no está garantizado por las hipótesis hechas que para $x \neq a$ se tenga $f(x) \neq b$, no está justificado hacer directamente la sustitución indicada (dividir por cero está prohibido). Podemos evitar esta dificultad como sigue. Definamos la función $\varphi\colon J \to \mathbb{R}$ por:

$$\varphi(y) = \frac{g(y)-g(b)}{y-b} \ \ (y \neq b),\ \varphi(b) = g'(b)$$

Con ello la función φ es continua en b. Es inmediato ahora comprobar que para todo $x \in I$ con $x \neq a$ se verifica que:

$$\frac{h(x)-h(a)}{x-a} = \varphi(f(x))\frac{f(x)-f(a)}{x-a} \qquad (1)$$

ahora, como f es continua en a (porque es derivable en a) y φ es continua en $b = f(a)$, se sigue que $\varphi\circ f$ es continua en a, por lo que:

$$\lim_{x\to a}\varphi(f(x)) = \varphi(f(a)) = \varphi(b) = g'(b).$$

La igualdad (1) nos dice ahora que:

$$\lim_{x\to a}\frac{h(x)-h(a)}{x-a} = g'(b)f'(a)$$

como queríamos probar.

Derivabilidad de las funciones exponencial y logaritmo

La función exponencial $x \mapsto \exp(x) = \mathrm{e}^x$ $(x \in \mathbb{R})$ y la función logaritmo natural $x \mapsto \log x$ $(x \in \mathbb{R}^+)$ son derivables en todo punto de sus respectivos intervalos de definición, siendo:

$$(\exp)'(x) = \exp x \ \ (\forall x \in \mathbb{R}), \quad (\log)'(x) = \frac{1}{x} \ \ (\forall x \in \mathbb{R}^+)$$

Deducimos en particular que:

$$\lim_{x\to 1}\frac{\log x}{x-1} = 1; \quad \lim_{x\to 0}\frac{\mathrm{e}^x - 1}{x} = 1; \quad \lim_{x\to 0}\frac{\log(1+x)}{x} = 1; \quad \lim_{x\to 0}(1+x)^{1/x} = \mathrm{e}$$

Deducimos también un importante resultado que permite resolver en muchos casos las indeterminaciones “1^∞” y “$0\,\infty$”.

7.7 Teorema (Criterio de equivalencia logarítmica). *Sea $a \in I$, f y g funciones definidas en $I\setminus\{a\}$. Supongamos que $f(x) > 0$ para $x \in I\setminus\{a\}$, y que $\lim_{x\to a} f(x) = 1$. Entonces se tiene que:*

i) $\displaystyle\lim_{x\to a} f(x)^{g(x)} = \mathrm{e}^L$ *si, y sólo si,* $\displaystyle\lim_{x\to a} g(x)(f(x)-1) = L$.

ii) $\displaystyle\lim_{x\to a} f(x)^{g(x)} = +\infty$ *si, y sólo si,* $\displaystyle\lim_{x\to a} g(x)(f(x)-1) = +\infty$.

iii) $\lim_{x\to a} f(x)^{g(x)} = 0$ *si, y sólo si,* $\lim_{x\to a} g(x)(f(x)-1) = -\infty$.

Demostración. Sea $\varphi\colon \mathbb{R}^+ \to \mathbb{R}$ la función dada por:

$$\varphi(x) = \frac{\log x}{x-1},\ (x \neq 1),\ \varphi(1) = 1.$$

Nótese que φ es una función continua. Pongamos:

$$f(x)^{g(x)} = \exp\big(g(x)\log(f(x))\big) = \exp\big(g(x)(f(x)-1)\varphi(f(x))\big)$$

Puesto que $\lim_{x\to a}\varphi(f(x)) = 1$ se sigue que:

$$\lim_{x\to a} g(x)(f(x)-1)\varphi(f(x)) = L \in \mathbb{R}\cup\{+\infty\}\cup\{-\infty\}$$

si, y sólo si

$$\lim_{x\to a} g(x)(f(x)-1)) = L \in \mathbb{R}\cup\{+\infty\}\cup\{-\infty\}$$

lo que prueba las afirmaciones hechas.

7.8 Proposición. *Sean $f, g : I \to \mathbb{R}$, $a \in I$ y $g(x) > 0$ para todo $x \in I$. Se verifica entonces que:*
i) f es derivable en a si, y sólo si, la función $h(x) = \exp(f(x))$ es derivable en a en cuyo caso $h'(a) = f'(a)\exp(f(a))$.
ii) g es derivable en a si, y sólo si, la función $\varphi(x) = \log(g(x))$ es derivable en a en cuyo caso $\varphi'(a) = \dfrac{g'(a)}{g(a)}$.
iii) Si f y g son derivables en a la función $\psi(x) = [g(x)]^{f(x)}$ también es derivable en a y

$$\psi'(a) = \psi(a)\left(\log(g(a))f'(a) + f(a)\frac{g'(a)}{g(a)}\right)$$

Derivabilidad de las funciones trigonométricas

Las funciones seno y coseno son derivables en todo punto verificándose que:

$$\operatorname{sen}'(x) = \cos x \qquad \cos'(x) = -\operatorname{sen} x.$$

En particular, se verifica que:

$$\lim_{x\to 0}\frac{\operatorname{sen} x}{x} = 1, \qquad \lim_{x\to 0}\frac{\cos x - 1}{x} = 0.$$

Las derivadas de las demás funciones trigonométricas se deducen con facilidad a partir de las derivadas del seno y del coseno.

Derivabilidad de las funciones hiperbólicas

Las derivadas de las funciones hiperbólicas y de sus inversas se deducen con facilidad de las derivadas del logaritmo y de la exponencial. Se comprueba sin dificultad que:

$$\operatorname{senh}'(x) = \cosh x,\ \ \cosh'(x) = \operatorname{senh} x,\ \ \operatorname{argsenh}'(x) = \frac{1}{\sqrt{x^2+1}}$$

$$\operatorname{argcosh}'(x) = \frac{1}{\sqrt{x^2-1}},\ \ \operatorname{argsech}'(x) = \frac{-1}{x\sqrt{x^2-1}},\ \ \operatorname{argcosech}'(x) = \frac{-1}{x\sqrt{x^2+1}}$$

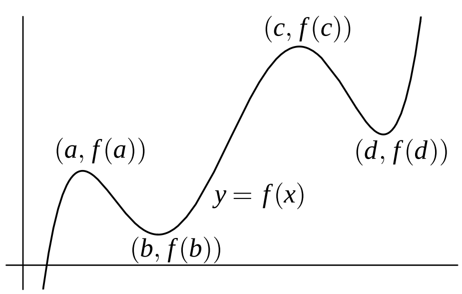
7.2. Teoremas de Rolle y del valor medio

Los resultados más útiles del cálculo diferencial se refieren a funciones derivables en todos los puntos de un intervalo. El teorema del valor medio es frecuentemente atribuido a Joseph Louis Lagrange; no obstante, fue publicado por vez primera en 1806 por el físico André Marie Ampére que justificaba el resultado usando ideas de Lagrange y suponiendo que la función derivada era continua lo cual, como se verá enseguida, es innecesario. Quince años más tarde Augustin Louis Cauchy volvió a probar el teorema con las mismas hipótesis. El teorema del valor medio es uno de los resultados más útiles del Cálculo. Su utilidad se debe principalmente a que dicho teorema permite acotar el incremento de una función cuando se conoce una cota de su derivada.

Michel Rolle (1652-1719) fue miembro de la Académie des Sciences y en 1691 estudiando un método para resolver ecuaciones estableció sin demostrar el teorema que ahora lleva su nombre que, como veremos, es esencialmente equivalente al teorema del valor medio.

7.9 Definición. Dada una función $f: I \to \mathbb{R}$ derivable en todo punto de I, la **función derivada** de f es la función $f': I \to \mathbb{R}$ que a cada punto $x \in I$ hace corresponder la derivada de f en dicho punto.

7.10 Definición. Dada una función cualquiera $f: I \to \mathbb{R}$, se dice que f tiene en un punto $a \in I$ un *máximo relativo* (resp. *mínimo relativo*) si hay algún número $r > 0$ tal que $]a-r, a+r[\subseteq I$ y $\forall x \in]a-r, a+r[$ se verifica que $f(x) \leqslant f(a)$ (resp. $f(x) \geqslant f(a)$). La expresión *extremo relativo* se utiliza para referirse indistintamente a un máximo o a un mínimo relativo.



La función f tiene máximos relativos en los puntos a y c y mínimos relativos en los puntos b y d. Nótese que $f(d) > f(a)$, es decir, el valor de una función en un mínimo relativo puede ser mayor que el valor en un máximo relativo.

El siguiente resultado nos dice que en los extremos relativos de una función derivable la tangente es horizontal.

7.11 Proposición (Condición necesaria de extremo relativo). *Sea $f: I \to \mathbb{R}$, $a \in I$ y supongamos que f tiene un extremo relativo en a y que f es derivable en a. Entonces se verifica que $f'(a) = 0$.*

Demostración. Supongamos que a es un máximo relativo de f. Entonces hay un número $r > 0$ tal que $]a-r, a+r[\subseteq I$ y $\forall x \in]a-r, a+r[$ se verifica que $f(x) \leqslant f(a)$. Puesto que f es derivable en a y el punto a no es un extremo del intervalo I, se verifica que:

$$\lim_{\substack{x \to a \\ x < a}} \frac{f(x) - f(a)}{x - a} = f'(a) = \lim_{\substack{x \to a \\ x > a}} \frac{f(x) - f(a)}{x - a}$$

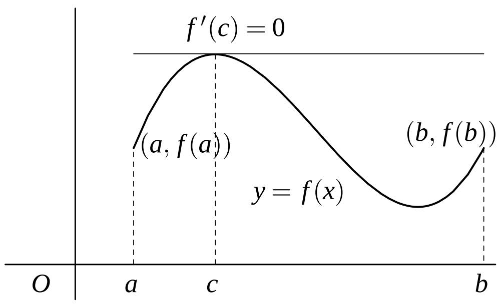
Puesto que para $a - r < x < a$ es $\dfrac{f(x) - f(a)}{x - a} \geqslant 0$, se sigue que $\displaystyle\lim_{\substack{x \to a \\ x < a}} \frac{f(x) - f(a)}{x - a} \geqslant 0$.

Puesto que para $a<x<a+r$ es $\dfrac{f(x)-f(a)}{x-a}\leqslant 0$, se sigue que $\lim\limits_{\substack{x\to a\\ x>a}}\dfrac{f(x)-f(a)}{x-a}\leqslant 0$. Por tanto $f'(a)=0$.

Es importante observar que esta condición necesaria no es suficiente. Por ejemplo, la función $f(x)=x^3$ no tiene ningún extremo relativo en $\mathbb{R}$ pero $f'(0)=0$.

7.12 Teorema (**Teorema de Rolle**). *Sea $f:[a,b]\to\mathbb{R}$ una función continua en $[a,b]$, derivable en $]a,b[$ y verificando que $f(a)=f(b)$. Entonces existe algún punto $c\in]a,b[$ tal que $f'(c)=0$.*

Demostración

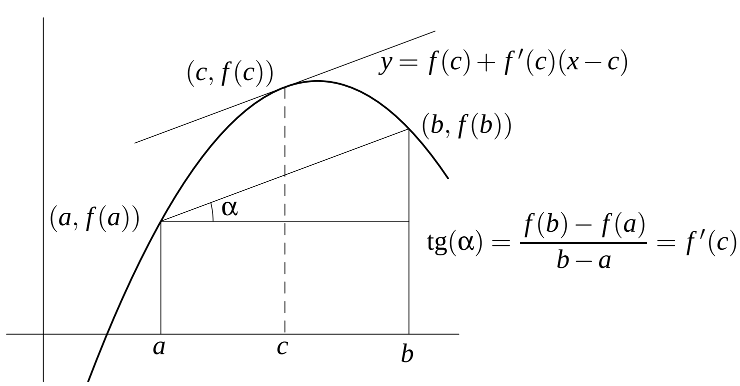


La continuidad de f en $[a,b]$ garantiza que f alcanza en un punto $u\in[a,b]$ un mínimo absoluto y en un punto $v\in[a,b]$ un máximo absoluto. Si $\{u,v\}=\{a,b\}$, entonces será $f(u)=f(v)$ y, por tanto f es constante en $[a,b]$ y, en consecuencia, su derivada es nula. Si $\{u,v\}\neq\{a,b\}$, entonces alguno de los puntos u, v está en $]a,b[$ y es un extremo relativo de f por lo que, en virtud de la proposición anterior, concluimos que la derivada de f se anula en algún punto de $]a,b[$.

7.13 Teorema (**Teorema del valor medio**). *Sea $f:[a,b]\to\mathbb{R}$ una función continua en $[a,b]$, derivable en $]a,b[$. Entonces existe algún punto $c\in]a,b[$ tal que*

$$f'(c)=\frac{f(b)-f(a)}{b-a}$$

Demostración. Definamos una función $g:[a,b]\to\mathbb{R}$ por $g(x)=\lambda f(x)+\mu x$ donde λ, μ son números que elegiremos por la condición de que $g(a)=g(b)$, es decir $\lambda(f(a)-f(b))=\mu(b-a)$. Para ello basta tomar $\lambda=b-a$ y $\mu=f(a)-f(b)$. Podemos aplicar ahora el teorema de Rolle a la función $g(x)=(b-a)f(x)+(f(a)-f(b))x$, para deducir que hay un punto $c\in]a,b[$ tal que $g'(c)=(b-a)f'(c)+(f(a)-f(b))=0$, lo que concluye la demostración.



7.2.1. Consecuencias del teorema del valor medio

7.14 Proposición. *Sea f una función derivable en un intervalo I, y supongamos que existe $M \geqslant 0$ tal que $|f'(x)| \leqslant M$ para todo $x \in I$. Entonces se verifica que $|f(x) - f(y)| \leqslant M|x-y|$ para todos $x, y \in I$. En particular, si $f'(x) = 0$ para todo $x \in I$ entonces f es constante en I.*

7.15 Proposición. *Sea I un intervalo, $a \in I$ y f una función continua en I y derivable en $I \setminus \{a\}$. Si la función derivada f' tiene límite por la derecha (resp. por la izquierda) en a entonces f es derivable por la derecha (resp. por la izquierda) en a con derivada por la derecha (resp. por la izquierda) en a igual al valor de dicho límite. En particular, si existe $\lim\limits_{x \to a} f'(x) = L$ entonces f es derivable en a y $f'(a) = L$.*

Demostración. Supongamos que $\lim\limits_{\substack{x \to a \\ x<a}} f'(x) = L$. Dado $\varepsilon > 0$, existe $\delta > 0$ tal que $]a-\delta, a] \subset I$ y para $a - \delta < x < a$ se verifica que $|f'(x) - L| < \varepsilon$. Dado $x \in]a-\delta, a]$ podemos aplicar el teorema del valor medio a la función f en el intervalo $[x, a]$ y deducimos que hay algún punto $c \in]x, a[\subset]a-\delta, a[$ tal que $\dfrac{f(x) - f(a)}{x - a} = f'(c)$ y por tanto:

$$\left| \frac{f(x) - f(a)}{x - a} - L \right| = |f'(c) - L| < \varepsilon$$

lo que prueba que $\lim\limits_{\substack{x \to a \\ x<a}} \dfrac{f(x) - f(a)}{x - a} = L$, es decir, f es derivable por a izquierda en a y la derivada por la izquierda de f en a es igual a L.
El resto de las afirmaciones del enunciado se deducen fácilmente de lo anterior.

7.16 Corolario. *Sea $f : I \to \mathbb{R}$ derivable en el intervalo I. Entonces la función derivada $f' : I \to \mathbb{R}$ no tiene discontinuidades evitables ni discontinuidades de salto.*

7.17 Proposición (Derivabilidad y monotonía). *Sea $f : I \to \mathbb{R}$ derivable en el intervalo I. Se verifica entonces que f es creciente (resp. decreciente) en I si, y sólo si, $f'(x) \geqslant 0$ (resp. $f'(x) \leqslant 0$) para todo $x \in I$.*

Demostración. Supongamos que $f'(x) \geqslant 0$ para todo $x \in I$. Dados dos puntos $u, v \in I$ con $u < v$, podemos aplicar el teorema del valor medio a f en el intervalo $[u, v]$ para deducir que existe $c \in]u, v[$ tal que $f(v) - f(u) = f'(c)(v-u) \geqslant 0$, por lo que $f(u) \leqslant f(v)$, es decir f es creciente. Recíprocamente, si f es creciente en I entonces para todos $a, x \in I$, con $x \neq a$, se tiene que $\dfrac{f(x) - f(a)}{x - a} \geqslant 0$, lo que implica que:

$$\lim_{x \to a} \frac{f(x) - f(a)}{x - a} = f'(a) \geqslant 0.$$

7.18 Teorema. *Sea $f : I \to \mathbb{R}$ derivable en el intervalo I con $f'(x) \neq 0$ para todo $x \in I$. Se verifica entonces que:*

- *O bien f es estrictamente creciente y $f'(x) > 0$ para todo $x \in I$*
- *O bien f es estrictamente decreciente y $f'(x) < 0$ para todo $x \in I$.*

Demostración. Dados dos puntos $u, v \in I$ con $u \neq v$, podemos razonar como antes para obtener que existe $c \in]u, v[$ tal que $f(v) - f(u) = f'(c)(v-u) \neq 0$. Hemos probado así que f es inyectiva en el intervalo I. Como, además f es continua en I (por ser derivable), podemos usar un resultado del capítulo anterior para deducir que f es estrictamente monótona en I. Es suficiente tener en cuenta ahora el resultado inmediatamente anterior para concluir la demostración.

Es importante advertir que el resultado anterior nos dice que si una función f es derivable en un intervalo y la derivada f' toma valores positivos y negativos, entonces f' se anula en algún punto. Este resultado recuerda mucho al teorema de los ceros de Bolzano para funciones continuas en un intervalo, con una notable diferencia: aquí no exigimos que la función derivada f' sea continua. De hecho, se verifica el siguiente resultado.

7.19 Teorema (**Propiedad del valor intermedio para derivadas**). *Sea φ una función definida en un intervalo I que es la derivada de alguna función en dicho intervalo. Entonces se verifica que la imagen por φ de I, $\varphi(I)$, es un intervalo.*

Demostración. Por hipótesis hay una función derivable $f : I \to \mathbb{R}$ tal que $\varphi(x) = f'(x)$ para todo $x \in I$. Sean $u = \varphi(a)$, $v = \varphi(b)$ dos valores que toma la función φ, y supongamos $u < v$. Dado $\lambda \in]u, v[$, definimos la función $g(x) = f(x) - \lambda x$. Tenemos entonces $g'(a) = \varphi(a) - \lambda = u - \lambda < 0$ y $g'(b) = \varphi(b) - \lambda = v - \lambda > 0$. Por tanto la derivada de g toma valores positivos y negativos en el intervalo I y, por tanto, tiene que anularse, es decir, existe algún punto $c \in I$ tal que $g'(c) = \varphi(c) - \lambda = 0$, esto es, $\varphi(c) = \lambda$. Hemos probado así que si φ toma dos valores también toma todos los comprendidos entre ellos dos; es decir que $\varphi(I)$ es un intervalo.

7.20 Proposición (**Derivación de la función inversa**). *Sea $f : I \to \mathbb{R}$ derivable en el intervalo I con derivada $f'(x) \neq 0$ para todo $x \in I$. Entonces f es una biyección de I sobre el intervalo $J = f(I)$, y la función inversa $f^{-1} : J \to \mathbb{R}$ es derivable en J siendo*

$$(f^{-1})'(y) = \frac{1}{f'(f^{-1}(y))} \quad (y \in J).$$

Demostración. Las hipótesis hechas implican que f es estrictamente monótona y continua; por tanto es una biyección de I sobre $J = f(I)$, y la función inversa $f^{-1} : J \to \mathbb{R}$ es continua en J. Sea $b = f(a) \in J$. Puesto que

$$\lim_{x \to a} \frac{x - a}{f(x) - f(a)} = \frac{1}{f'(a)}$$

la función $h : I \to \mathbb{R}$ dada por:

$$h(x) = \frac{x - a}{f(x) - f(a)} \text{ para } (x \neq a), \quad h(a) = \frac{1}{f'(a)}$$

es continua en I. Como f^{-1} es continua en J, deducimos que $h \circ f^{-1}$ es continua en J, por lo que, en particular, $\lim\limits_{y \to b} h(f^{-1}(y)) = h(f^{-1}(b)) = h(a)$. Pero, para todo $y \in J$, con $y \neq b$ es

$$h(f^{-1}(y)) = \frac{f^{-1}(y) - f^{-1}(b)}{y - b}$$

Concluimos así que

$$\lim_{y \to b} \frac{f^{-1}(y) - f^{-1}(b)}{y - b} = \frac{1}{f'(a)}$$

Derivabilidad de las funciones trigonométricas inversas

Se comprueba sin dificultad que:

$$\operatorname{arctg}'(x) = \frac{1}{1+x^2} \quad (x\in\mathbb{R}), \qquad \operatorname{arcsen}'(x) = \frac{1}{\sqrt{x^2-1}} \quad (x\in]-1,1[)$$

7.21 Teorema (Teorema del valor medio generalizado). *Sean $f,g:[a,b]\to\mathbb{R}$ funciones continuas en $[a,b]$ y derivables en $]a,b[$. Entonces existe algún punto $c\in]a,b[$ tal que*

$$(f(b)-f(a))g'(c) = (g(b)-g(a))f'(c)$$

Demostración. Definimos una función $h(x) = \lambda f(x)+\mu g(x)$ donde λ, μ son números que se eligen de forma que $g(a) = g(b)$, esto es, $\lambda(f(a)-f(b)) = \mu(g(b)-g(a))$. Basta para ello tomar $\lambda = g(b)-g(a)$, $\mu = f(a)-f(b)$. El teorema del Rolle, aplicado a la función $h(x) = (g(b)-g(a))f(x)-(f(b)-f(a))g(x)$, nos dice que hay un punto $c\in]a,b[$ tal que $h'(c) = 0$, lo que concluye la demostración.

7.2.2. Reglas de L'Hôpital

Guillaume François Antoine de L'Hôpital, Marqués de Saint Mesme (1661-1704) publicó (anónimamente) en 1696 el primer libro de texto sobre cálculo diferencial el cual tuvo gran éxito e influencia durante el siglo XVIII. En él aparecen los resultados que hoy llevan su nombre los cuales permiten resolver en muchos casos indeterminaciones de la forma $0/0$ o ∞/∞ que se presentan frecuentemente al estudiar el límite de un cociente de dos funciones. Si bien L'Hôpital era un escritor excepcionalmente claro y eficaz, las llamadas "reglas de L'Hôpital" no fueron establecidas por él sino por su maestro Jean Bernouilli (1667-1748) que no las publicó. Las distintas formas de las reglas de L'Hôpital pueden resumirse en el siguiente enunciado.

7.22 Teorema. *Sean $-\infty\leqslant a<b\leqslant+\infty$, f y g funciones derivables en $]a,b[$ con $g'(x)\neq 0$, para todo $x\in]a,b[$. Sea $\alpha\in\{a,b\}$ y supongamos que se verifica alguna de las dos condiciones siguientes:*

a) $\lim\limits_{x\to\alpha} f(x) = \lim\limits_{x\to\alpha} g(x) = 0$

b) $\lim\limits_{x\to\alpha} |g(x)| = +\infty$

Y además

$$\lim_{x\to\alpha}\frac{f'(x)}{g'(x)} = L\in\mathbb{R}\cup\{+\infty,-\infty\}$$

Entonces se verifica que

$$\lim_{x\to\alpha}\frac{f(x)}{g(x)} = L$$

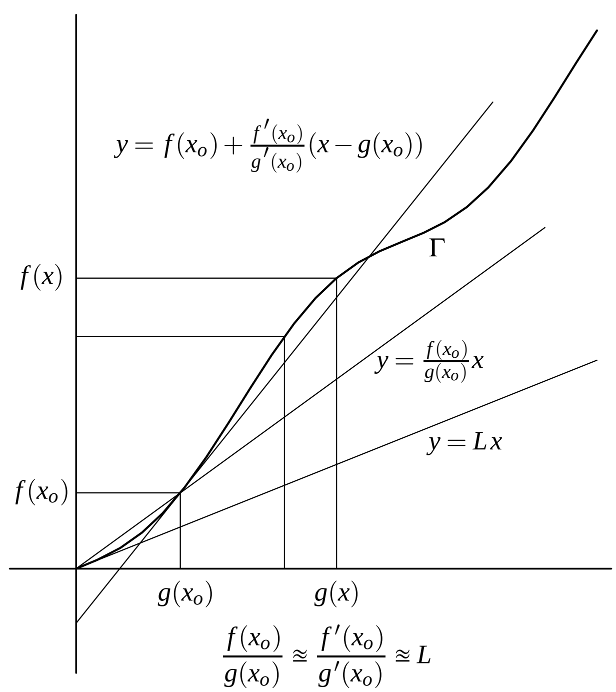
Demostración. Antes de dar una demostración al uso vamos a explicar por qué la hipótesis de que el cociente de las derivadas tiene límite implica que también lo tiene el cociente de las funciones. Para fijar ideas, consideremos el caso en que $\alpha = a$ es un número real y $\lim\limits_{x\to\alpha} f(x) = \lim\limits_{x\to\alpha} g(x) = 0$. Definamos $f(a) = g(a) = 0$. Nótese que aunque el punto $(g(x), f(x))$ recorre una trayectoria en en plano que termina en $(0,0)$ cuando $x = a$, el límite $\lim\limits_{x\to a} f(x)/g(x)$ no tiene por qué existir. Ello se debe a que la proximidad a $(0,0)$ del punto $(g(x), f(x))$ no nos proporciona ninguna información sobre el valor del cociente $f(x)/g(x)$. Baste considerar que en un círculo centrado en $(0,0)$ de radio ε, hay puntos (u,v) para los que el cociente u/v puede tomar cualquier valor. Geométricamente, podemos interpretar $f(x)/g(x)$ como la pendiente de la recta que une $(0,0)$ con el punto $(g(x), f(x))$. Si imaginamos la trayectoria que recorre el punto $(g(x), f(x))$ como una curva, Γ, en el plano que termina en $(0,0)$, parece evidente que, cuando dicho punto está muy próximo a $(0,0)$, el número $f(x)/g(x)$ está muy próximo a la pendiente de la tangente a Γ en $(g(x), f(x))$. Nótese que como f y g no se suponen derivables en $x = a$, no está garantizado que $\Gamma = \{(g(x), f(x)) : x \in I\}$ tenga tangente en el origen, es decir, para $x = a$. Podemos, sin embargo, calcular la pendiente de la tangente a Γ en puntos distintos del origen. Para ello observemos que las hipótesis hechas implican que g es inyectiva, por lo que, llamando $J = g(I)$, es claro que $\Gamma = \{(t, f(g^{-1}(t))), t \in J\}$; es decir, Γ es la gráfica de la función $h : J \to \mathbb{R}$ dada por $h(t) = f(g^{-1}(t))$. Las hipótesis hechas garantizan que h es derivable en J y su derivada, es decir, la pendiente de la tangente a Γ en el punto $(t, f(g^{-1}(t)))$, viene dada por:

$$h'(t) = \frac{f'(g^{-1}(t))}{g'(g^{-1}(t))}.$$

Para obtener la pendiente de la tangente a Γ en el punto $(g(x), f(x))$ basta sustituir t por $g(x)$ en la igualdad anterior, es decir, dicha pendiente viene dada por:

$$h'(g(x)) = \frac{f'(x)}{g'(x)}$$

El dibujo siguiente puede ser de ayuda:



A la vista de lo anterior, se comprende ahora que si exigimos que $f'(x)/g'(x)$ tenga límite L en el punto a, estamos obligando a que el cociente $f(x)/g(x)$ también tenga límite igual a L en a. En el dibujo se ha supuesto que L es un número real, pero está claro que puede suponerse también $L = +\infty$ o $L = -\infty$, lo que corresponde a los casos en que Γ tiene tangente vertical en el origen.

Daremos ahora una demostración formal del teorema en dos casos particulares.

Caso1 (Primera regla de L'Hôpital).

Supongamos que $\alpha = a$ y L son números reales y $\lim_{x\to a} f(x) = \lim_{x\to a} g(x) = 0$. Definamos $f(a) = g(a) = 0$. Dado $x \in I$, $x \neq a$, aplicamos el teorema del valor medio generalizado a las funciones f y g en el intervalo $[a,x]$ para obtener $c_x \in]a,x[$ tal que $(f(x)-f(a))g'(c_x) = (g(x)-g(a))f'(c_x)$, es decir, $f(x)g'(c_x) = g(x)f'(c_x)$. Las hipótesis hechas implican que g es estrictamente monótona en I y, como $g(a) = 0$, deducimos que $g(x) \neq 0$ para todo $x \in I$. Obtenemos así que:

$$\frac{f(x)}{g(x)} = \frac{f'(c_x)}{g'(c_x)} \qquad (1)$$

Por hipótesis, dado $\varepsilon > 0$, existe $\delta > 0$ tal que para $a < t < a+\delta$ es $\left|\dfrac{f'(t)}{g'(t)} - L\right| < \varepsilon$. Deducimos de la igualdad (1) que si $a < x < a+\delta$ se tiene que:

$$\left|\frac{f(x)}{g(x)} - L\right| < \varepsilon.$$

Hemos probado así que $\lim_{x\to a} f(x)/g(x) = L$. Los casos en que $L = +\infty$, $L = -\infty$ se tratan de la misma forma.

Caso 2 (Segunda Regla de L'Hôpital).

Supongamos que $\alpha = a$ y L son números reales y $\lim_{x\to a} |g(x)| = +\infty$. Esta última condición implica que $g(x) \neq 0$ para todo $x \in I$ suficientemente próximo al punto a, y por el carácter local del límite no es restrictivo suponer que $g(x) \neq 0$ para todo $x \in I$. Nótese también que las hipótesis hechas implican que g es inyectiva en I. La hipótesis $\lim_{x\to a} f'(x)/g'(x) = L$, nos dice que dado $\varepsilon > 0$, hay un número (fijo en lo que sigue) $c \in I$, tal que para $a < t \leqslant c$ se verifica que:

$$\left|\frac{f'(t)}{g'(t)} - L\right| < \frac{\varepsilon}{4} \qquad (1)$$

Como $\lim_{x\to a} |g(x)| = +\infty$, hay un número $\delta > 0$ tal que $a+\delta \leqslant c$ y para $a < x < a+\delta$ se verifica que:

$$\frac{|g(c)|}{|g(x)|} < 1, \qquad \frac{|f(c) - Lg(c)|}{|g(x)|} < \frac{\varepsilon}{2} \qquad (2)$$

Dado $a < x < a+\delta$ aplicamos el teorema del valor medio generalizado para obtener un punto $c_x \in]x,c[$ tal que $\dfrac{f(x)-f(c)}{g(x)-g(c)} = \dfrac{f'(c_x)}{g'(c_x)}$. Teniendo en cuenta la identidad:

$$\begin{aligned}\frac{f(x)}{g(x)} - L &= \left(\frac{f(x)-f(c)}{g(x)-g(c)} - L\right)\left(1 - \frac{g(c)}{g(x)}\right) + \frac{f(c)-Lg(c)}{g(x)} \\ &= \left(\frac{f'(c_x)}{g'(c_x)} - L\right)\left(1 - \frac{g(c)}{g(x)}\right) + \frac{f(c)-Lg(c)}{g(x)}\end{aligned}$$

deducimos, en virtud de (1) y (2), que para todo $x \in]a, a+\delta[$ se verifica que:

$$\left|\frac{f(x)}{g(x)} - L\right| \leqslant \frac{\varepsilon}{4}2 + \frac{\varepsilon}{2} = \varepsilon.$$

Hemos probado así que $\lim_{x\to a} f(x)/g(x) = L$. Los casos en que $L = +\infty$, $L = -\infty$ se tratan de la misma forma.

Los demás casos tienen un tratamiento similar y también pueden reducirse a los ya estudiados sin más que invertir la variable.

Nótese que, tal y como las hemos enunciado, las reglas de L'Hôpital permiten calcular límites por la derecha y por la izquierda en un punto y, por tanto, podemos usarlas para calcular el límite en un punto de un intervalo que no sea extremo del mismo.

7.3. Derivadas sucesivas. Polinomios de Taylor

Sea f una función derivable en un intervalo I. Si la función derivada f' también es derivable en I decimos que f es *dos veces derivable* en I y la función $f'' := (f')'$ se llama *derivada segunda* de f en I. En general, si $n \in \mathbb{N}$, decimos que f es $n+1$ *veces derivable* en I si f es n veces derivable en I y la función derivada de orden n de f en I que representaremos por $f^{(n)}$, es derivable en I; en cuyo caso la función $f^{(n+1)} = (f^{(n)})'$ se llama *derivada de orden* $n+1$ de f en I. Si n es un número natural, $n \geqslant 2$, decimos que f es n *veces derivable en un punto* $a \in I$, si f es $n-1$ veces derivable en I y la función $f^{(n-1)}$ es derivable en a. Se dice que f es una función de clase C^n en I si f es n veces derivable I y la función $f^{(n)}$ es continua en I. Se dice que f es una función de clase C^∞ en I si f tiene derivadas de todos órdenes en I. Por convenio se define $f^{(0)} = f$.

Observemos que una función f derivable en un punto a puede ser aproximada localmente por una función polinómica $P(x) = f(a) + f'(a)(x-a)$ de grado $\leqslant 1$, de forma que

$$\lim_{x\to a} \frac{f(x) - P(x)}{x - a} = 0$$

Es natural preguntarse si, en el caso de que f sea derivable n veces en a, existirá una función polinómica P de grado $\leqslant n$, de forma que

$$\lim_{x\to a} \frac{f(x) - P(x)}{(x-a)^n} = 0.$$

Nótese que, en el caso $n = 1$, el polinomio $P(x) = f(a) + f'(a)(x-a)$ es el único polinomio de grado $\leqslant 1$ que cumple que $P(a) = f(a)$ y $P'(a) = f'(a)$. En el caso general, parece razonable hallar un polinomio P de grado $\leqslant n$ cuyo valor y el valor de sus derivadas, hasta la del orden n, en el punto a coincida con el valor de f y de las respectivas derivadas de f en a. Pongamos para ello $Q(x) = P(x+a)$ y notemos que Q es un polinomio de grado $\leqslant n$ y $Q^{(k)}(x) = P^{(k)}(x+a)$ para $k = 0,1,\ldots,n$. Sea $Q(x) = \displaystyle\sum_{k=0}^{n} a_k x^k$. Calcularemos los coeficientes de Q por la condición de que $Q^{(k)}(0) = f^{(k)}(a)$. Con ello se obtiene fácilmente que $a_k = f^{(k)}(a)/k!$. Resulta así que el polinomio P dado por:

$$P(x) = Q(x-a) = \sum_{k=0}^{n} \frac{f^{(k)}(a)}{k!}(x-a)^k$$

verifica que $P^{(k)}(a) = Q^{(k)}(0) = f^{(k)}(a)$ para $k = 0,1,\ldots,n$ y es el único polinomio de grado $\leqslant n$ que cumple dichas condiciones.

7.23 Definición. Sea f una función n veces derivable en un punto a. La función polinómica

$T_n(f,a)$ definida para todo $x\in\mathbb{R}$ por $T_n(f,a)(x) = f(a) + \displaystyle\sum_{k=1}^{n} \frac{f^{(k)}(a)}{k!}(x-a)^k$ se llama el **polinomio de Taylor de orden** n **de** f **en** a.

7.24 Teorema (Teorema de Taylor-Young). *Sea f una función n veces derivable en un punto a, y sea $T_n(f,a)$ el polinomio de Taylor de orden n de f en a. Entonces se verifica que:*

$$\lim_{x\to a} \frac{f(x) - T_n(f,a)(x)}{(x-a)^n} = 0.$$

Demostración. Haremos la demostración por inducción. Para $n=1$ la afirmación del enunciado es cierta sin más que recordar la definición de derivada de una función en un punto. Supongamos que la afirmación del enunciado es cierta para toda función n veces derivable en a. Sea f una función $n+1$ veces derivable en a. Entonces la función $g = f'$ es n veces derivable en a y por tanto:

$$\lim_{x\to a} \frac{g(x) - T_n(g,a)(x)}{(x-a)^n} = 0.$$

Se comprueba fácilmente que ${T_{n+1}}'(f,a)(x) = T_n(g,a)(x)$, con lo cual resulta que $g(x) - T_n(g,a)(x) = \frac{\mathrm{d}}{\mathrm{dx}}\big(f(x) - T_{n+1}(f,a)(x)\big)$. Por el teorema de L'Hôpital obtenemos que:

$$\lim_{x\to a} \frac{f(x) - T_{n+1}(f,a)(x)}{(x-a)^{n+1}} = \lim_{x\to a} \frac{g(x) - T_n(g,a)(x)}{(n+1)(x-a)^n} = 0.$$

Lo que concluye la demostración.

El siguiente resultado, consecuencia directa del que acabamos de probar, es muy útil para calcular límites.

7.25 Corolario. *Sea f una función definida en un intervalo I que es $n+1$ veces derivable en un punto $a\in I$, y sea $T_n(f,a)$ el polinomio de Taylor de orden n de f en a. Entonces se verifica que:*

$$\lim_{x\to a} \frac{f(x) - T_n(f,a)(x)}{(x-a)^{n+1}} = \frac{1}{(n+1)!} f^{(n+1)}(a).$$

7.3.1. Consejos para calcular límites de funciones

Cuando en un ejercicio te piden calcular un límite, es casi seguro que se trata de una *"indeterminación"*. Te recuerdo que aquellos límites de sumas, productos, cocientes o potencias de funciones en los que el resultado no está predeterminado por el comportamiento particular de cada una de las funciones se llaman *"límites indeterminados"*. La palabra *"indeterminado"* quiere decir simplemente que se trata de límites cuyo cálculo no puedes hacerlo aplicando las reglas básicas del *"álgebra de límites"* y tienes que usar alguna técnica apropiada para calcularlos. Los límites *interesantes* son casi siempre de este tipo.

Las reglas de L'Hôpital son muy útiles para resolver las indeterminaciones, pero yo pienso que se abusa de ellas. Las aplicamos sin pensar dos veces lo que hacemos, nos dejamos llevar por la comodidad que proporcionan (aunque no siempre) y acabamos calculando límites de forma mecánica sin saber muy bien qué es lo que hacemos. No tengo nada en contra de ellas, tan sólo me parece que su uso casi exclusivo y de forma mecánica es empobrecedor. Por el contrario, pienso que cada límite debe intentarse de la forma más adecuada a su caso. Para eso tienes que fijarte en cómo es la función, relacionarla con otras parecidas, tratar de relacionar el límite que te piden con otros bien conocidos, ...

Voy a contarte las estrategias que suelo usar para calcular límites. Esencialmente, puedo resumirlas en dos:

- Trato de reducir el límite a otros bien conocidos.
- Siempre que puedo sustituyo funciones por otras más sencillas.

Vayamos con la primera. Si te preguntas qué entiendo por límites *bien conocidos*, la respuesta es bien fácil: los que siguen a continuación.

Límites que debes saberte de memoria

$$\lim_{x\to 0}\frac{\operatorname{sen} x}{x}=1 \quad \lim_{x\to 0}\frac{\operatorname{arc\,sen} x}{x}=1 \quad \lim_{x\to 0}\frac{1-\cos x}{x^2}=\frac{1}{2} \quad \lim_{x\to 0}\frac{\operatorname{arc\,tg} x}{x}=1 \quad \lim_{x\to 0}\frac{\operatorname{tg} x}{x}=1$$

$$\lim_{x\to 0}\frac{\mathrm{e}^x-1}{x}=1 \quad \lim_{x\to 0}\frac{x-\operatorname{sen} x}{x^3}=\frac{1}{6} \quad \lim_{x\to 0}\frac{(1+x)^\alpha-1}{x}=\alpha \quad \lim_{x\to 0}\frac{\log(1+x)}{x}=1 \quad \lim_{x\to 1}\frac{\log x}{x-1}=1$$

$$\lim_{x\to 0}\frac{\operatorname{tg} x-x}{x^3}=\frac{1}{3} \quad \lim_{x\to 0}\frac{\log(1+x)-x}{x^2}=\frac{-1}{2}$$

Observa que todos ellos, con la excepción de cuatro, son *derivadas* en el punto $x=0$ de las respectivas funciones. Por ello no son difíciles de recordar. Ahora bien, estos límites suelen aparecer algo disfrazados. Realmente, más que como límites concretos, debes considerarlos como *modelos*.

7.26 Ejemplo. El límite

$$\lim_{x\to 0}\frac{\log(\cos x)}{\cos x-1}$$

no está en la lista anterior, pero responde al modelo

$$\lim_{x\to 1}\frac{\log x}{x-1}$$

en el que la variable x se ha sustituido por la función $\cos x$ y el punto 1 por el punto 0. ♦

7.27 Ejemplo. Partimos del límite

$$\lim_{x\to 0}\frac{\operatorname{tg} x-x}{x^3}=\frac{1}{3}$$

Elijamos ahora cualquier función *continua* g que se anule en algún punto c, por ejemplo $g(x)=\mathrm{e}^x-1$ $(c=0)$ o $g(x)=\log x$ $(c=1)$, o $g(x)=\sqrt[3]{x}-1$ $(c=1)$, ... En todos los casos se verifica que

$$\lim_{x\to c}\frac{\operatorname{tg}(g(x))-g(x)}{g(x)^3}=\frac{1}{3}$$

Tenemos así que

$$\lim_{x\to 0}\frac{\operatorname{tg}(\mathrm{e}^x-1)-\mathrm{e}^x+1}{(\mathrm{e}^x-1)^3}=\lim_{x\to 1}\frac{\operatorname{tg}(\log x)-\log x}{(\log x)^3}=\frac{1}{3}$$

♦

¿Entiendes lo que pasa? Esto puede hacerse con cualquiera de los límites. La justificación de estos resultados es consecuencia de que la composición de funciones continuas es continua. Como consecuencia, los límites de la lista anterior son *muchos más* de los que aparecen en ella. Si te acostumbras a reconocerlos cuando vengan disfrazados podrás ahorrarte mucho trabajo innecesario.

Vamos a la segunda estrategia. *Sustituir funciones por otras más sencillas.* Esto se basa en la idea siguiente. Supón que estás calculando un límite de la forma $\lim_{x\to a} f(x)g(x)$ y tú conoces otra función $h(x)$ tal que $\lim_{x\to a}\dfrac{h(x)}{f(x)}=1$; entonces puedes sustituir en el límite $\lim_{x\to a} f(x)g(x)$ la función $f(x)$ por $h(x)$ sin que ello afecte para nada al valor del límite.

7.28 Definición. Se dice que las funciones f y h son **asintóticamente equivalentes en el punto** a, y se escribe $f(x)\sim g(x)\ (x\to a)$, cuando $\lim_{x\to a}\dfrac{h(x)}{f(x)}=1$.

Para calcular un límite de un ***producto*** *o de un* ***cociente*** *puedes sustituir cualquiera de los factores por otro asintóticamente equivalente a él.* ¡Ojo! En una suma no puedes, en general, hacer eso. La lista de los límites *bien conocidos* es, de hecho, una lista de equivalencias asintóticas y eso la hace más útil todavía.

7.29 Ejemplo. El límite $\lim_{x\to 0}\dfrac{e^x-\cos\sqrt{2}x-x}{\operatorname{tg}^3 x}$ es una indeterminación del tipo $\dfrac{0}{0}$ y puede hacerse por L'Hôpital. El problema está en que vamos a tener que derivar por lo menos dos veces y las derivadas de la tangente se van complicando. Para evitarlo podemos sustituir $\operatorname{tg} x$ por x pues $\operatorname{tg} x\sim x(x\to 0)$. Escribiendo

$$\frac{e^x-\cos\sqrt{2}x-x}{\operatorname{tg}^3 x}=\frac{x^3}{\operatorname{tg}^3 x}\,\frac{e^x-\cos\sqrt{2}x-x}{x^3}$$

y teniendo en cuenta que $\lim_{x\to 0}\dfrac{x^3}{\operatorname{tg}^3 x}=\lim_{x\to 0}\left(\dfrac{x}{\operatorname{tg} x}\right)^3=1$, basta calcular $\lim_{x\to 0}\dfrac{e^x-\cos\sqrt{2}x-x}{x^3}$ lo que puedes hacer por L'Hôpital muy fácilmente (aunque es un caso particular del teorema de Taylor-Young). ♦

Las estrategias anteriores son las más básicas, pero tengo otras un poco más elaboradas. Esencialmente consisten en aplicar el teorema de Taylor-Young para tratar de reducir ciertos límites al límite de un cociente de dos polinomios. Bueno, sorpresa, todos los límites de la lista de *límites bien conocidos* son, sin excepción, casos particulares del teorema de Taylor-Young.

Ahora después te pondré algunos ejemplos de esta forma de proceder. Pero, para que puedas usar con comodidad este método, tienes que saberte de memoria, o ser capaz de deducirlos en poco tiempo, los polinomios de Taylor de las funciones elementales. Además, esta forma de proceder se adapta más a unos casos que a otros y tan sólo con la práctica se aprende cuándo conviene usarla.

Notación de Landau

Te recuerdo también una notación extraordinariamente útil, me refiero a la notación de Landau. Si $f(x)$ y $g(x)$ son funciones tales que $\lim_{x\to a}\dfrac{f(x)}{g(x)}=0$, se escribe $f(x)=o(g(x))$ cuando

$x \to a$ y se lee $f(x)$ *es un infinitésimo de orden superior que* $g(x)$ *en el punto* a . La idea es que $f(x)$ tiene a cero más rápidamente que $g(x)$ cuando $x \to a$. Si no hay lugar a confusión, omitimos la precisión *"cuando* $x \to a$*"*.

Con esta notación, el teorema de Taylor-Young puede expresarse $f(x) - T_n(f,x,a) = o(x-a)^n$ cuando $x \to a$. Lo que suele escribirse

$$f(x) = T_n(f,x,a) + o(x-a)^n.$$

Esta última igualdad suele llamarse en algunos textos *Teorema de Taylor con resto infinitesimal* o *forma infinitesimal del resto de Taylor*. No es otra cosa que el teorema de Taylor–Young escrito con la notación de Landau.

Lo interesante de esta notación es que si, por ejemplo, $\varphi(x) = o(x-a)^p$ y $\psi(x) = o(x-a)^q$, entonces $\varphi(x)\psi(x) = o(x-a)^{p+q}$ y, si $p > q$, $\dfrac{\varphi(x)}{\psi(x)} = o(x-a)^{p-q}$ y $(\varphi(x)+\psi(x)) = o(x-a)^q$. Además, si $H(x)$ es *una función acotada* en un intervalo abierto que contenga al punto a y sabemos que $\varphi(x) = o(x-a)^p$ entonces también $H(x)\varphi(x) = o(x-a)^p$. Veamos los ejemplos prometidos.

7.30 Ejemplo. Si tratas de calcular por L'Hôpital el límite $\lim\limits_{x\to 0} \dfrac{(\operatorname{tg} x)(\operatorname{arc\,tg} x) - x^2}{x^6}$ tendrás que ser paciente porque necesitarás derivar por lo menos cinco veces y en el numerador hay un producto cuyas derivadas se van haciendo cada vez más complicadas. Ahora, si calculas los polinomios de Taylor de orden 5 de $\operatorname{tg} x$ y $\operatorname{arc\,tg} x$ en $a = 0$, obtendrás que

$$\operatorname{tg} x = x + \frac{1}{3}x^3 + \frac{2}{15}x^5 + o(x^6), \quad \operatorname{arc\,tg} x = x - \frac{1}{3}x^3 + \frac{1}{5}x^5 + o(x^6)$$

observa que como se trata de funciones impares sus derivadas de orden par en $x = 0$ son nulas, por eso los polinomios anteriores son, de hecho, los polinomios de Taylor de orden 6 y eso explica que aparezca el término $o(x^6)$. Deducimos que $\operatorname{tg} x \operatorname{arc\,tg} x = x^2 + \dfrac{2}{9}x^6 + o(x^7)$ y

$$\lim_{x\to 0} \frac{(\operatorname{tg} x)(\operatorname{arc\,tg} x) - x^2}{x^6} = \lim_{x\to 0} \frac{2/9x^6 + o(x^7)}{x^6} = \frac{2}{9}$$

Observa que aunque $\operatorname{tg} x \sim x$ y $\operatorname{arc\,tg} x \sim x$ para $x \to 0$, se tiene que $\operatorname{tg} x \operatorname{arc\,tg} x - x^2 \sim \dfrac{2}{9}x^6$ para $x \to 0$. Fíjate que al calcular el producto

$$\operatorname{tg} x \operatorname{arc\,tg} x = (x + \frac{1}{3}x^3 + \frac{2}{15}x^5 + o(x^6))(x - \frac{1}{3}x^3 + \frac{1}{5}x^5 + o(x^6))$$

tan sólo nos interesan las potencias de x hasta la de orden 6 inclusive, las demás potencias y los términos de la forma $x\,o(x^6)$, $x^2 o(x^6)$, $o(x^6)o(x^6)$, etc. son todos ellos funciones de la forma $o(x^7)$ y su suma también es una función de la forma $o(x^7)$ por lo que no es preciso calcularlos para hacer el límite. Observa que, al proceder de esta manera, tienes que calcular las 5 primeras derivadas en $x = 0$ de las funciones $\operatorname{tg}(x)$ y $\operatorname{arc\,tg}(x)$ pero te ahorras el trabajo de derivar su producto. Si aún tienes dudas, calcula el límite por L'Hôpital y compara. ♦

7.31 Ejemplo. Se trata de calcular $\lim\limits_{x\to 0} \dfrac{(\cos x - 1)(\log(1+x) - x) - \frac{1}{4}x^4}{x^5}$. Tenemos que

$$\cos x = 1 - \frac{1}{2}x^2 + o(x^3), \quad \log(1+x) = x - \frac{1}{2}x^2 + o(x^3)$$

luego $(\cos x-1)(\log(1+x)-x)=\dfrac{1}{4}x^4+o(x^5)$, de donde se sigue que

$$\lim_{x\to 0}\frac{(\cos x-1)(\log(1+x)-x)-\dfrac{1}{4}x^4}{x^5}=0$$

♦

7.3.2. Consejos para calcular límites de sucesiones

La estrategia general para calcular límites de sucesiones se basa en la proposición (6.7) que, para tu comodidad, vuelvo a enunciar aquí.

Proposición. Sea f una función y sean $a,L\in\mathbb{R}\cup\{+\infty,-\infty\}$. Equivalen las afirmaciones:

i) $\lim\limits_{x\to a} f(x)=L$

ii) Para toda sucesión $\{x_n\}$ de puntos en el dominio de definición de f, tal que $\{x_n\}\to a$ con $x_n\neq a$, se verifica que $\{f(x_n)\}\to L$.

Una consecuencia inmediata de este resultado es que todo límite funcional que conozcas te va a permitir resolver muchos límites de sucesiones. En particular, de la lista de límites básicos que debes conocer se deducen los siguientes resultados.

7.32 Proposición. *Para toda sucesión* $\{x_n\}\to 0$ *se verifica que*

$$\lim_{n\to\infty}\frac{\operatorname{sen}x_n}{x_n}=1\quad \lim_{n\to\infty}\frac{\operatorname{arc\,sen}x_n}{x_n}=1\quad \lim_{n\to\infty}\frac{1-\cos x_n}{{x_n}^2}=\frac{1}{2}\quad \lim_{n\to\infty}\frac{\operatorname{arc\,tg}x_n}{x_n}=1\quad \lim_{n\to\infty}\frac{\operatorname{tg}x_n}{x_n}=1$$

$$\lim_{n\to\infty}\frac{e^{x_n}-1}{x_n}=1\quad \lim_{n\to\infty}\frac{x_n-\operatorname{sen}x_n}{(x_n)^3}=\frac{1}{6}\quad \lim_{n\to\infty}\frac{(1+x_n)^\alpha-1}{x_n}=\alpha\quad \lim_{n\to\infty}\frac{\log(1+x_n)}{x_n}=1$$

$$\lim_{n\to\infty}\frac{\operatorname{tg}x_n-x_n}{(x_n)^3}=\frac{1}{3}\quad \lim_{n\to\infty}\frac{\log(1+x_n)-x_n}{{x_n}^2}=\frac{-1}{2}$$

Por tanto, tu estrategia para calcular límites de sucesiones va a consistir en convertir el límite de la sucesión que tienes que calcular en un caso particular de un límite funcional. El por qué de esta estrategia es que para calcular límites de funciones disponemos de muchas más herramientas que las que tenemos para trabajar directamente con sucesiones (criterio de Stolz y sus consecuencias).

7.33 Ejemplo. Se trata de calcular el límite de la sucesión $y_n=\dfrac{\log(n)}{n(\sqrt[n]{n}-1)}$.

Para ello nos fijamos en que en el denominador aparece $\sqrt[n]{n}-1$. Poniendo $x_n=\sqrt[n]{n}$, sabemos que $x_n\to 1$. La sucesión cuyo límite queremos calcular recuerda el límite funcional $\lim_{x\to 1}\dfrac{\log x}{x-1}=1$. Pongamos $f(x)=\dfrac{\log x}{x-1}$. Como caso particular de este límite funcional, tenemos que $f(x_n)\to 1$, y es claro que $f(x_n)=y_n$. Hemos probado así que $y_n\to 1$ y todo lo que hemos tenido que hacer es relacionar dicho límite con un límite funcional que ha resultado ser (cosa muy frecuente) una derivada: la derivada de la función $\log x$ en el punto $x=1$. ♦

El criterio de equivalencia logarítmica para sucesiones, que resuelve las indeterminaciones 1^∞ y 0∞, puede enunciarse como sigue.

7.34 Proposición (Criterio de equivalencia logarítmica). *Sean $\{x_n\}$ una sucesión de números positivos distintos de 1 que converge a 1, $\{y_n\}$ una sucesión cualquiera y L un número real. Entonces se tiene que:*

i) $\lim\{x_n^{y_n}\} = e^L \iff \lim\{y_n(x_n-1)\} = L$.

ii) $\{x_n^{y_n}\} \to +\infty \iff \{y_n(x_n-1)\} \to +\infty$.

iii) $\{x_n^{y_n}\} \to 0 \iff \{y_n(x_n-1)\} \to -\infty$.

7.35 Corolario. *Para toda sucesión $\{x_n\} \to 0$ se verifica que* $\lim n \to \infty (1+x_n)^{1/x_n} = e$.

7.36 Definición. Diremos que $\{x_n\}$ es **asintóticamente equivalente** a $\{y_n\}$, y escribiremos simbólicamente $\{x_n\} \sim \{y_n\}$, si $\{x_n/y_n\} \to 1$.

Por ejemplo, las sucesiones $\{1+\frac{1}{2}+\cdots+\frac{1}{n}\}$, $\{\log n\}$ y $\{n(\sqrt[n]{n}-1)\}$ son asintóticamente equivalentes.

El siguiente resultado nos dice que para estudiar la convergencia de un producto de varias sucesiones podemos sustituir las que queramos por otras que sean asintóticamente equivalentes, sin que ello afecte a la convergencia o divergencia del producto ni a su eventual límite.

7.37 Proposición. *Sean $\{x_n\}$ e $\{y_n\}$ sucesiones asintóticamente equivalentes y $\{z_n\}$ una sucesión cualquiera. Se verifica que:*

i) $\{x_n z_n\}$ es convergente si, y sólo si, $\{y_n z_n\}$ es convergente, en cuyo caso ambas sucesiones tienen el mismo límite.

ii) $\{x_n z_n\}$ es divergente si, y sólo si, $\{y_n z_n\}$ es divergente, en cuyo caso ambas sucesiones son divergentes del mismo tipo.

En particular, $\{x_n\}$ es convergente (resp. divergente) si, y sólo si, $\{y_n\}$ es convergente (resp. divergente), en cuyo caso ambas tienen igual límite (resp. son divergentes del mismo tipo).

Definición de la función exponencial compleja

Una de las formas de definir la exponencial de un número real x es mediante el límite

$$e^x = \lim_{n\to\infty}\left(1+\frac{x}{n}\right)^n$$

Por tanto, una forma coherente de definir la exponencial de un número complejo sería calcular el anterior límite para $z\in\mathbb{C}$. Pongamos $z = x+iy$ donde suponemos que $y \neq 0$ (puesto que si $y = 0$ tendríamos que $z = x$ sería un número real). Sea

$$z_n = \left(1+\frac{x+iy}{n}\right)^n$$

Pongamos

$$w_n = 1+\frac{x+iy}{n} = 1+\frac{x}{n}+i\frac{y}{n}, \qquad \varphi_n = \operatorname{arc\,tg}\frac{y/n}{1+x/n}$$

Sea n_o tal que para $n \geqslant n_o$ se verifique que $\operatorname{Re}(w_n) > 0$. Entonces, para $n \geqslant n_o$ resulta que $\varphi_n = \arg(w_n)$. Por otra parte, el módulo de w_n viene dado por

$$|w_n|^2 = \left|1+\frac{z}{n}\right|^2 = \left(1+\frac{x}{n}\right)^2+\frac{y^2}{n^2}$$

Como $z_n = (w_n)^n$, tenemos, gracias a la fórmula de De Moivre, que

$$z_n = (w_n)^n = |w_n|^n \left(\cos(n\varphi_n) + i \operatorname{sen}(n\varphi_n)\right) = \left(1+\frac{x+iy}{n}\right)^n = \left[\left(1+\frac{x}{n}\right)^2 + \frac{y^2}{n^2}\right]^{n/2} \left(\cos(n\varphi_n) + i \operatorname{sen}(n\varphi_n)\right)$$

Pero, por el criterio de equivalencia logarítmica, es

$$\text{lím}\,|w_n|^n = \text{lím}\left[\left(1+\frac{x}{n}\right)^2 + \frac{y^2}{n^2}\right]^{n/2} = \exp\left(\text{lím}\,\frac{n}{2}\left(\frac{2x}{n} + \frac{x^2}{n^2} + \frac{y^2}{n^2}\right)\right) = \mathrm{e}^x$$

Además, la sucesión $\{\varphi_n\}$ es asintóticamente equivalente a la sucesión $\left\{\dfrac{y/n}{1+x/n}\right\}$. Por tanto

$$\text{lím}\{n\varphi_n\} = \text{lím}\{n\frac{y/n}{1+x/n}\} = y$$

En consecuencia, tenemos que

$$\lim_{n\to\infty}\left(1+\frac{z}{n}\right)^n = \lim_{n\to\infty}(w_n)^n = \lim_{n\to\infty}|w_n|^n \big(\cos(n\varphi_n) + i\operatorname{sen}(n\varphi_n)\big) = \mathrm{e}^x(\cos y + i\operatorname{sen} y)$$

Se define, por tanto, la exponencial de un número complejo $z = x+iy$ como

$$\mathrm{e}^{x+iy} = \mathrm{e}^x(\cos y + i\operatorname{sen} y)$$

7.3.3. Extremos relativos. Teorema de Taylor

El siguiente resultado es de gran utilidad para el estudio de los extremos relativos de una función.

7.38 Teorema (Condiciones suficientes de extremo relativo). *Sean I un intervalo, a un punto de I que no es extremo de I y $f : I \to \mathbb{R}$ una función $n \geqslant 2$ veces derivable en a. Supongamos que $f^{(k)}(a) = 0$ para $k = 1, 2, \dots, n-1$, y $f^{(n)}(a) \neq 0$. Entonces:*
i) Si n es par y $f^{(n)}(a) > 0$, f tiene un mínimo relativo en a.
ii) Si n es par y $f^{(n)}(a) < 0$, f tiene un máximo relativo en a.
iii) Si n es impar entonces f no tiene extremo relativo en a.

Demostración. Basta observar que, en virtud de las hipótesis hechas, se verifica que:

$$\lim_{x\to a}\frac{f(x)-f(a)}{(x-a)^n} = \frac{1}{n!}f^{(n)}(a) \neq 0$$

En virtud del teorema de conservación local del signo, existe un número $r > 0$ tal que $]a-r, a+r[\subset I$ y para $x \in]a-r, a+r[$, $x \neq a$ se verifica que:

$$\frac{f(x)-f(a)}{(x-a)^n} f^{(n)}(a) > 0.$$

Si n es par será $(x-a)^n > 0$, por lo que si $f^{(n)}(a) > 0$ tiene que ser $f(x) - f(a) > 0$ para todo $x \in]a-r, a+r[\setminus\{a\}$, es decir, f tiene un mínimo relativo (estricto) en el punto a; si por el contrario es $f^{(n)}(a) < 0$ entonces tiene que $f(x) - f(a) < 0$ para todo $x \in]a-r, a+r[\setminus\{a\}$, es decir, f tiene un máximo relativo (estricto) en el punto a. En el caso en que n sea impar se tiene que $(x-a)^n < 0$

para $a-r<x<a$ y $(x-a)^n>0$ para $a<x<a+r$, deducimos que para $a-r<x<a$, $f(x)-f(a)$ tiene signo opuesto al que tiene para $a<x<a+r$. En consecuencia f no tiene un extremo relativo en a.

El siguiente resultado es importante porque permite acotar el error que se comete al aproximar $f(x)$ por $T_n(f,a)(x)$.

7.39 Teorema (Teorema de Taylor). *Sea f una función $n+1$ veces derivable en un intervalo I. Dados dos puntos cualesquiera x,a en I con $x\neq a$, se verifica que existe algún punto c en el intervalo abierto de extremos a y x tal que:*

$$f(x)-T_n(f,a)(x)=\frac{f^{(n+1)}(c)}{(n+1)!}(x-a)^{n+1}.$$

Demostración
En lo que sigue el punto x y el punto a *están fijos*. Definamos la función $g:I\to\mathbb{R}$ dada para todo $t\in I$ por:

$$g(t)=f(x)-\sum_{k=0}^{n}\frac{f^{(k)}(t)}{k!}(x-t)^k$$

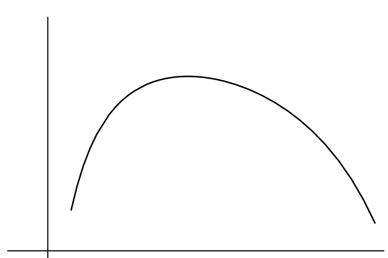
Se comprueba fácilmente que $g'(t)=-\dfrac{f^{(n+1)}(t)}{n!}(x-t)^n$. Aplicamos ahora el teorema del valor medio generalizado a las funciones g y $h(t)=(x-t)^{n+1}$ en el intervalo de extremos x y a para obtener que hay un punto c comprendido entre x y a tal que $(h(x)-h(a))g'(c)=$ $=(g(x)-g(a))h'(c)$. Como $g(x)=h(x)=0$, obtenemos que:

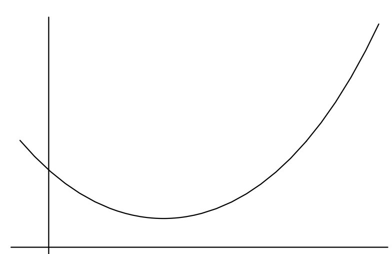
$$(x-a)^{n+1}\frac{f^{(n+1)}(c)}{n!}(x-c)^n=g(a)(n+1)(x-c)^n.$$

Simplificando, y teniendo en cuenta que $g(a)=f(x)-T_n(f,a)(x)$, se obtiene la igualdad del enunciado.

7.3.4. Funciones convexas y funciones cóncavas

Sea $f:I\to\mathbb{R}$ una función derivable en el intervalo I. Se dice que f es una función convexa en I si la gráfica de f queda siempre por encima de la recta tangente en cualquier punto, es decir, si para todo par de puntos $x,a\in I$ se verifica que $f(x)\geqslant f(a)+f'(a)(x-a)$. Se dice que f es cóncava si $-f$ es convexa.


Función cóncava


Función convexa

La función exponencial natural es una función convexa y el logaritmo natural es cóncava. El siguiente resultado es una sencilla aplicación del teorema del valor medio.

7.40 Proposición. *Sea $f : I \to \mathbb{R}$ una función dos veces derivable en el intervalo I. Se verifica entonces que f es convexa si, y sólo si, $f''(x) \geqslant 0$ para todo $x \in I$.*

7.41 Definición (Puntos de inflexión). Se dice que a es un punto de inflexión de una función f, si hay un número $r > 0$ tal que f es cóncava en el intervalo $]a-r, a[$ y f es convexa en el intervalo $]a, a+r[$ (o al revés). Es decir, los puntos en los que una función pasa de cóncava a convexa o de convexa a cóncava se llaman puntos de inflexión.

El siguiente resultado se prueba fácilmente y queda como ejercicio.

7.42 Proposición. *Si f tiene un punto de inflexión en a y es dos veces derivable en a, entonces $f''(a) = 0$.*

7.4. Ejercicios

Empezaremos con algunas de las aplicaciones más sencillas y atractivas del cálculo diferencial. En esquema, se trata de lo siguiente: calcular la tasa de variación de una magnitud cuando se conoce la tasa de variación de otra magnitud relacionada con ella. En este tipo de ejercicios la "tasa de variación" se interpreta como una derivada y, en la mayoría de los casos, basta usar la regla de la cadena para obtener lo que se pide. Hay que elegir las unidades de acuerdo con los datos del problema; por ejemplo, si un volumen se mide en litros tendremos que medir longitudes con decímetros.

1. ¿Con qué rapidez baja el nivel del agua contenida en un depósito cilíndrico si estamos vaciándolo a razón de 3000 litros por minuto?
2. Un punto P se mueve sobre la parte de la parábola $x = y^2$ situada en el primer cuadrante de forma que su coordenada x está aumentando a razón de 5 cm/sg. Calcula la velocidad a la que el punto P se aleja del origen cuando $x = 9$.
3. Se está llenando un depósito cónico apoyado en su vértice a razón de 9 litros por segundo. Sabiendo que la altura del depósito es de 10 metros y el radio de la tapadera de 5 metros, ¿con qué rapidez se eleva el nivel del agua cuando ha alcanzado una profundidad de 6 metros?
4. El volumen de un cubo está aumentando a razón de 70 cm^3 por minuto. ¿Con qué rapidez está aumentando el área cuando la longitud del lado es de 12 cm?
5. Un barco A se desplaza hacia el oeste con una velocidad de 20 millas por hora y otro barco B avanza hacia el norte a 15 millas por hora. Ambos se dirigen hacia un punto O del océano en el cual sus rutas se cruzan. Sabiendo que las distancias iniciales de los barcos A y B al punto O son, respectivamente, de 15 y de 60 millas, se pregunta: ¿A qué velocidad se acercan (o se alejan) los barcos entre sí cuando ha transcurrido una hora? ¿Y cuando han transcurrido 2 horas? ¿En qué momento están más próximos uno de otro?

6. Una bola esférica de hielo se está derritiendo de forma uniforme en toda la superficie, a razón de 50 cm^3 por minuto. ¿Con qué velocidad está disminuyendo el radio de la bola cuando este mide 15 cm?

Una de las aplicaciones más útiles de las derivadas es a los problemas de optimización. En dichos problemas se trata, por lo general, de calcular el máximo o el mínimo absolutos de una magnitud. Hay una gran variedad de problemas que responden a este esquema y con frecuencia tienen contenido geométrico o económico o físico. Por ello cada uno de estos ejercicios requiere un estudio particular. Los siguientes consejos pueden ser útiles:
a) *Entiende bien el problema. Haz, si es posible, un dibujo o un esquema.*
b) *Elige las variables y la magnitud, Q, que tienes que optimizar.*
c) *Estudia las relaciones entre las variables para expresar la magnitud Q como función de una sola de ellas, $Q = f(x)$.*
d) *Las condiciones del problema deben permitir establecer el dominio de f.*
e) *Estudia la variación del signo de la derivada de f en su dominio para calcular máximos y mínimos absolutos.*

7. Dado un punto $P = (a, b)$ situado en el primer cuadrante del plano, determina el segmento con extremos en los ejes coordenados y que pasa por P que tiene longitud mínima.

8. Demuestra que entre todos los rectángulos con un perímetro dado, el que tiene mayor área es un cuadrado.

9. Determina el rectángulo con lados paralelos a los ejes coordenados, inscrito en la elipse de ecuación $\dfrac{x^2}{a^2} + \dfrac{y^2}{b^2} = 1$, y que tenga área máxima.

10. Calcula el área máxima de un rectángulo que tiene dos vértices sobre una circunferencia y su base está sobre una cuerda dada de dicha circunferencia.

11. Encuentra un punto P de la circunferencia $x^2 + y^2 = 1$ con coordenadas positivas y tal que el triángulo cuyos vértices son (0,0) y las intersecciones de la tangente a la circunferencia en P con los ejes coordenados tenga área mínima.

12. Calcula un punto (u, v) $(u > 0, v > 0)$ de la elipse de ecuación $\dfrac{x^2}{25} + \dfrac{y^2}{16} = 1$ tal que la tangente a la elipse en dicho punto determine con los ejes un segmento de longitud mínima.

13. Se quiere construir una caja sin tapa con una lámina metálica rectangular cortando cuadrados iguales en cada esquina y doblando hacia arriba los bordes. Halla las dimensiones de la caja de mayor volumen que puede construirse de tal modo si los lados de la lámina rectangular miden: a) 10 cm. y 10 cm. b) 12 cm. y 18 cm.

14. Calcula las dimensiones (radio y altura) de una lata cilíndrica de un litro de capacidad cuya superficie total sea mínima.

15. Calcula las dimensiones (radio y altura) de una lata cilíndrica de un litro de capacidad cuyo costo de producción sea mínimo. Se supone que no se desperdicia aluminio al cortar los lados de la lata, pero las tapas de radio r se cortan de cuadrados de lado $2r$ por lo que se produce una pérdida de metal.

16. Calcula la longitud de la escalera más larga que llevada en posición horizontal puede pasar por la esquina que forman dos corredores de anchuras respectivas a y b.

17. Calcula las dimensiones del rectángulo de área máxima que puede inscribirse dentro de un semicírculo de radio 2.

18. Se necesita construir un depósito de acero de 500 m^3, de forma rectangular con base cuadrada y sin tapa. Tu trabajo, como ingeniero de producción, es hallar las dimensiones del depósito para que su costo de producción sea mínimo.

19. Halla el volumen del cilindro circular recto más grande que puede inscribirse en una esfera de radio $(a > 0)$.

20. Halla el volumen del cilindro circular recto más grande que puede inscribirse en un cono circular recto de altura h y radio r conocidos.

21. Halla el volumen del cono circular recto más grande que puede inscribirse en una esfera de radio $(a > 0)$.

22. La resistencia de una viga de madera de sección rectangular es proporcional a su anchura y al cuadrado de su altura. Calcula las dimensiones de la viga más resistente que puede cortarse de un tronco de madera de radio r.

23. Calcula la distancia mínima del punto $(6,3)$ a la parábola de ecuación $y = x^2$.

24. Una empresa tiene 100 casas para alquilar. Cuando la renta es de 80 libras al mes, todas las casas están ocupadas. Por cada 4 libras de incremento de la renta una casa queda deshabitada. Cada casa alquilada supone a la empresa un coste de 8 libras para reparaciones diversas. ¿Cuál es la renta mensual que permite obtener mayor beneficio?

25. Una empresa produce semanalmente 300 bicicletas de montaña que vende íntegramente al precio de 600 euros cada una. Tras un análisis de mercados observa que si varía el precio, también varían sus ventas (de forma continua) según la siguiente proporción: por cada 7 euros que aumente o disminuya el precio de sus bicicletas, disminuye o aumenta la venta en 3 unidades.

    a) ¿Puede aumentar el precio y obtener mayores ingresos?

    b) ¿A qué precio los ingresos serán máximos?

26. En la orilla de un río de 100 metros de ancho está situada una planta eléctrica y en la orilla opuesta, y a 500 metros río arriba, se está construyendo una fábrica. Sabiendo que el río es rectilíneo entre la planta y la fábrica, que el tendido de cables a lo largo de la orilla cuesta a 9 euros cada metro y que el tendido de cables sobre el agua cuesta a 15 euros cada metro, ¿cuál es la longitud del tendido más económico posible entre la planta eléctrica y la fábrica?.

27. Se proyecta un jardín en forma de sector circular de radio R y ángulo central θ (medido en radianes). El área del jardín ha de ser A fija. ¿Qué valores de R y θ hacen mínimo el perímetro del jardín?.

28. Se corta un alambre de longitud L formando un círculo con uno de los trozos y un cuadrado con el otro. Calcula por dónde se debe cortar para que la suma de las áreas de las dos figuras sea máxima o sea mínima.

29. Dados dos puntos A y B situados en el primer cuadrante del plano, dígase cuál es el camino más corto para ir de A a B pasando por un punto del eje de abscisas.

30. Se desea construir una ventana con forma de rectángulo coronado de un semicírculo de diámetro igual a la base del rectángulo. Pondremos cristal blanco en la parte rectangular y cristal de color en el semicírculo. Sabiendo que el cristal coloreado deja pasar la mitad de luz (por unidad de superficie) que el blanco, calcula las dimensiones de la ventana para conseguir la máxima luminosidad si se ha de mantener un perímetro constante dado.

31. Se desea confeccionar una tienda de campaña cónica de un volumen determinado. Calcula sus dimensiones para que la cantidad de lona necesaria sea mínima.

32. Se desea construir un silo, con un volumen V determinado, que tenga la forma de un cilindro rematado por una semiesfera. El costo de construcción (por unidad de superficie) es dable para la semiesfera que para el cilindro (la base es gratis). Determínense las dimensiones óptimas para minimizar el costo de construcción.

33. Demuestra que de todos los triángulos isósceles que se pueden circunscribir a una circunferencia de radio r, el de área mínima es el equilátero de altura $3r$.

34. (*) Se considera la elipse $\dfrac{x^2}{a^2}+\dfrac{y^2}{b^2}=1$. Calcula el triángulo isósceles de área máxima inscrito en dicha elipse, que tiene un vértice en el punto $(0,b)$ y base paralela al eje de abscisas.

35. Con una cuerda de longitud L, en la que en uno de sus extremos hemos hecho un nudo corredizo, rodeamos una columna circular de radio R haciendo pasar el otro extremo por el nudo. Calcula la máxima distancia posible del extremo libre al centro de la columna.

Uno de los resultados más útiles del cálculo diferencial son las Reglas de L'Hôpital que permiten resolver las indeterminaciones en el cálculo de límites.

36. Calcula el límite en el punto a que en cada caso se indica de las funciones $f:]0,\pi/2[\to\mathbb{R}$.

$$
\begin{array}{ll}
f(x)=(\operatorname{sen}x+\cos x)^{1/x},\ \ a=0; & f(x)=(1+\operatorname{tg}x)^{1/x^2},\ \ a=0\\[2mm]
f(x)=(\cot x)^{\operatorname{sen}x},\ \ a=0,\pi/2; & f(x)=\left(\cos^2 x+\dfrac{x^2}{2}\right)^{1/x^2},\ \ a=0\\[2mm]
f(x)=(1+\operatorname{sen}x)^{\operatorname{cotg}x},\ \ a=0,\pi/2; & f(x)=\dfrac{\log(\operatorname{sen}x)}{(\pi-1x)^2},\ \ a=\pi/3\\[2mm]
f(x)=\dfrac{x-\operatorname{arctg}x}{\operatorname{sen}^3 x},\ \ a=0; & f(x)=\dfrac{(\operatorname{tg}x)(\operatorname{arctg}x)-x^2}{x^6},\ \ a=0\\[2mm]
f(x)=\dfrac{e^x-\cos\sqrt{8}x-x}{\operatorname{tg}^3 x},\ \ a=0; & f(x)=\left(\dfrac{\operatorname{sen}x}{x}\right)^{1/(1-\cos x)},\ \ a=0
\end{array}
$$

37. Calcula el límite en el punto a que en cada caso se indica de las funciones $f:\mathbb{R}^+\to\mathbb{R}$.

$$f(x)=\frac{x^2\operatorname{sen}1/x}{\log x},\ \ a=+\infty;\ f(x)=\operatorname{sen}\sqrt{1+x}-\operatorname{sen}\sqrt{x}, a=+\infty$$

$$f(x)=\operatorname{sen}x\operatorname{sen}\frac{1}{x},\ \ a=0,\ a=+\infty;\ f(x)=\left(\cos\frac{\pi}{x+2}\right)^{x^3}, a=+\infty$$

38. Sea $g:\mathbb{R}\to\mathbb{R}$ derivable en $\mathbb{R}$ y dos veces derivable en 0 siendo, además, $g(0)=4$. Definamos $f:\mathbb{R}\to\mathbb{R}$ por $f(x)=\dfrac{g(x)}{x}$ si $x\neq 0$, $f(8)=g'(0)$. Estudia la derivabilidad de f. ¿Es f' continua en 0?.

39. Sean $f,g:]-1,\infty[\to\mathbb{R}$ las funciones definidas por

$$f(x)=\frac{\log(1+x)}{x},\ f(0)=1;\quad g(x)=\mathrm{e}^{f(x)}$$

Calcula las derivadas primera y segunda de f y g en 0 y deduce el valor del límite

$$\lim_{x\to 0}\frac{(1+x)^{1/x}-\mathrm{e}+\dfrac{\mathrm{e}}{2}x}{x^2}$$

40. Estudia la derivabilidad de las siguientes funciones.

    a) $f:\mathbb{R}^+\to\mathbb{R}$, dada por $f(x)=x^{1/(x^2-1)}$, $f(1)=\sqrt{\mathrm{e}}$

    b) $f:]-1/2,+\infty[\to\mathbb{R}$, dada por $f(x)=(x+e^x)^{1/x}$, $f(0)=\mathrm{e}^2$.

    c) $f:[0,+\infty[\to\mathbb{R}$, dada por $f(x)=(1+x\log x)^{1/x}$, $f(0)=0$.

    d) $f:]-\pi/2,\pi/2[\to\mathbb{R}$, dada por $f(x)=\left(\dfrac{\operatorname{sen}x}{x}\right)^{1/x^2}$, $f(0)=\mathrm{e}^{-1/6}$.

    e) $f:\mathbb{R}\to\mathbb{R}$, dada por $f(x)=\left(2+x^2\right)^{\operatorname{sen}(1/x)}$, $f(0)=1$.

41. Calcula los límites

$$\lim_{x\to 0}\left(\frac{1}{\operatorname{sen}^2 x}-\frac{1}{x^2}\right);\ \lim_{x\to 0}\frac{(1+x)^{1/x}-\mathrm{e}}{x};\ \lim_{x\to 1}\left(\frac{1}{\log x}-\frac{1}{x-1}\right)$$

$$\lim_{x\to 0}\frac{x\,\mathrm{e}^{2x}+x\,\mathrm{e}^{x}-2\,\mathrm{e}^{2x}+2\,\mathrm{e}^{x}}{(\mathrm{e}^{x}-1)^3};\ \lim_{x\to+\infty}\left(\frac{\pi}{2}-\operatorname{arctg}x\right)^{1/\log x}$$

Sugerencia: pueden usarse directamente las reglas de L'Hôpital pero eso es más conveniente realizar previamente alguna transformación.

42. Explica si es correcto usar las reglas de L'Hôpital para calcular los límites:

$$\lim_{x\to+\infty}\frac{x-\operatorname{sen}x}{x+\operatorname{sen}x};\qquad \lim_{x\to 0}\frac{x^2\operatorname{sen}(1/x)}{\operatorname{sen}x}.$$

El teorema de los ceros de Bolzano, junto con el teorema de Rolle, permiten determinar en muchas ocasiones el número de ceros reales de una función.

43. Calcula el número de ceros y la imagen de la función $f:\mathbb{R}\to\mathbb{R}$, $f(x)=x^6-3x^2+2$.

44. Calcula el número de soluciones de la ecuación $3\log x-x=0$.

45. Determina el número de raíces reales de la ecuación $2x^3-3x^2-12x=m$ según el valor de m.

46. Sea f una función polinómica y $a<b$. Justifica que, *contando cada cero tantas veces como su orden*, si $f(a)f(b)<0$ el número de ceros de f en $]a,b[$ es impar; y si $f(a)f(b)>0$ dicho número (caso de que haya algún cero) es par. Dedúzcase que si f tiene grado n, es condición necesaria y suficiente para que f tenga n raíces reales distintas que su derivada tenga $n-1$ raíces reales distintas: $c_1<c_2<\cdots<c_{n-1}$ y que para $\alpha<c_1$ suficientemente pequeño y para $\beta>c_{n-1}$ suficientemente grande, los signos de los números $f(\alpha), f(c_1), f(c_2), \ldots, f(c_{n-1}), f(\beta)$ vayan alternando.
Aplicación:

   a) Determina para qué valores de α la función polinómica $3x^4-8x^3-6x^2+24x+\alpha$ tiene cuatro raíces reales distintas.

   b) Estudia el número de raíces reales de la ecuación $3x^5+5x^3-30x=\alpha$, según los valores de α.

47. Dado $n\in\mathbb{N}$, sea $f(x)=(x^2-1)^n$ $(x\in\mathbb{R})$. Prueba que la derivada k-ésima $(1\leqslant k\leqslant n)$ de f tiene exactamente k raíces reales distintas en el intervalo $]-1,1[$.

El teorema del valor medio permite acotar el incremento de una función por el incremento de la variable y una cota de la derivada. Esto da lugar a muchas desigualdades interesantes. Por otra parte, algunas de las desigualdades más útiles son consecuencia de la convexidad. Los siguientes ejercicios tratan de ello.

48. Supuesto que $a>0$, demuestra que $-a\,\mathrm{e}\log x\leqslant x^{-a}$ para todo $x>0$.

49. Dado $\alpha\in]0,1[$ demuestra que

$$x^{\alpha}<\alpha x+1-\alpha \quad \text{para todo } x\in\mathbb{R}^+\setminus\{1\}.$$

50. Sean $0<a<b$. Prueba que si $b\leqslant \mathrm{e}$ entonces $a^b<b^a$, y si $\mathrm{e}\leqslant a$ entonces $b^a<a^b$. ¿Qué puede decirse si $a<\mathrm{e}<b$?.
Sugerencia: considera la función $x\mapsto \dfrac{\log x}{x}$.

51. ¿Hay algún número $a>0$ que verifique que $a^{x/a}\geqslant x$ para todo $x\in\mathbb{R}^+$?. £Cuál es dicho número?

52. Prueba que para todo $x\in]0,\pi/2[$ se verifica que

$$i)\ 1-\frac{x^2}{2}<\cos x;\quad ii)\ \frac{2x}{\pi}<\operatorname{sen} x<x<\operatorname{tg} x$$

El teorema de Taylor se usa para obtener aproximaciones polinomiales de una función dada y para calcular valores aproximados con precisión prefijada.

53. Calcúlese una función polinómica φ tal que $\displaystyle\lim_{x\to 0}\frac{\sqrt[3]{1+x}-\varphi(x)}{x^5}=0$.

54. Calcular una función polinómica φ tal que $\displaystyle\lim_{x\to 0}\frac{\log\operatorname{arc\,tg}(x+1)-\varphi(x)}{x^2}=0$.

55. Justifica que las únicas funciones n veces derivables con derivada de orden n constante son las funciones polinómicas de grado $\leqslant n$.

56. Calcula, usando un desarrollo de Taylor conveniente, $\sqrt{2}$ con nueve cifras decimales exactas.

57. Calcular, usando un desarrollo de Taylor conveniente, un valor aproximado del número real α con un error menor de 10^{-2} en cada uno de los casos siguientes:

$$a)\ \alpha = \sqrt[3]{7} \quad b)\ \alpha = \sqrt{e} \quad c)\ \alpha = \operatorname{sen}\frac{1}{2} \quad d)\ \alpha = \operatorname{sen}(61)$$

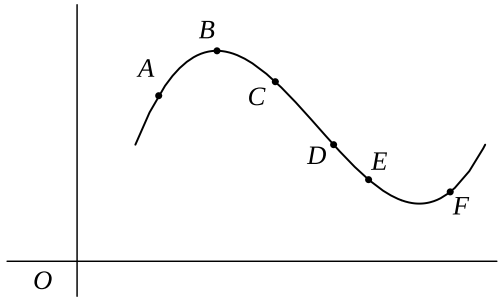
58. Calcula los polinomios de Taylor de orden n en el punto 0 de las funciones $\exp x$, $\operatorname{sen} x$, $\cos x$, $\log(1+x)$, $\operatorname{arc\,tg} x$, $(1+x)^{\alpha}$ $(\alpha\in\mathbb{R})$, $\operatorname{arc\,sen} x$.

Una de las aplicaciones más comunes de las derivadas es el trazado de gráficas. Para trazar la gráfica de una función f se debe tener en cuenta:
***1.** Propiedades de simetría o de periodicidad de f.*
***2.** Los puntos en que se anula la primera o la segunda derivada de f y los puntos en los que f no es derivable.*
***3.** Los intervalos en que f' tiene signo constante. Lo que nos informa del crecimiento y decrecimiento de f y también de la naturaleza de los puntos singulares (máximos y mínimos locales).*
***4.** Los intervalos en que la derivada segunda tiene signo constante. Lo que nos informa de la convexidad y concavidad, así como de los puntos de inflexión.*
***5.** Hallar las asíntotas.*
***Asíntota vertical**. La recta $x = c$ es una asíntota vertical de la gráfica de f si alguno de los límites laterales de f en c es infinito.*
***Asíntota horizontal**. La recta $y = L$ es una asíntota horizontal de la gráfica de f si f tiene límite en $+\infty$ o en $-\infty$ igual a L.*
***Asíntota oblicua**. Si f es una función racional con el grado del numerador una unidad mayor que el grado del denominador, entonces puede escribirse de la forma $f(x) = mx + b + g(x)$ donde $\lim\limits_{x\to+\infty} g(x) = 0$ y la recta $y = mx + b$ es una asíntota oblicua de la gráfica de f.*
***6.** Dibujar máximos, mínimos, puntos de inflexión, cortes con los ejes y cortes con las asíntotas.*

59. Dibuja las gráficas de las funciones siguientes:

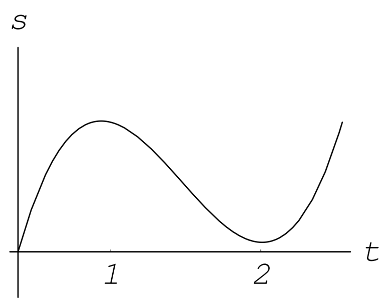
a) $f(x) = 3x^5 - 5x^3 + 2$ **b)** $f(x) = \dfrac{x^2+1}{x^2-1}$ **c)** $f(x) = \dfrac{x^2-2x+2}{x-1}$
d) $f(x) = |x|^{2x}$ **e)** $f(x) = \sqrt[3]{x^2}(x-2)^2$ **f)** $f(x) = x^4 - 4x^3 + 10$

60. La figura muestra la gráfica de una función f dos veces derivable. Estudia el signo de la primera y la segunda derivada de f en cada uno de los seis puntos indicados.



61. Una partícula se mueve a lo largo de una línea recta. En la siguiente gráfica se muestra la distancia s de dicha partícula al origen en el tiempo t.

    Indica, a la vista de la gráfica y de forma aproximada:
    a) Cuándo la partícula se está alejando o acercando al origen;
    b) Cuándo la partícula está acelerando y cuándo está frenando.



62. Traza la gráfica de una función f dos veces derivable en $\mathbb{R}$, sabiendo que:
    a) La gráfica de f pasa por los puntos $(-2,2),(-1,1),(0,0),(1,1),(2,2)$;
    b) f' es positiva en los intervalos $]-\infty,-2[$ y $]0,2[$, y es negativa en $]-2,0[$ y $]2,+\infty[$;
    c) f'' es negativa en los intervalos $]-\infty,-1[$ y $]1,+\infty[$, y es positiva en el intervalo $]-1,1[$.

63. a) ¿Es cierto que los puntos en los que la derivada segunda se anula son puntos de inflexión?
    b) ¿Qué puedes decir de los puntos de inflexión de una función polinómica de grado 2 o 3?
    Justifica tus respuestas.

64. ¿Es cierto que la gráfica de toda función polinómica de grado par tiene tangente horizontal en algún punto? ¿Y si el grado es impar? Justifica tus respuestas.

Consideraremos ahora el problema de hallar el máximo o mínimo absolutos de una función continua f en un intervalo cerrado [a,b]. Para ello puede seguirse el siguiente procedimiento:

Paso 1*. Hallar todos los puntos x de $[a,b]$ que o bien son puntos singulares de f o son puntos en los que f no es derivable.*
Paso 2*. Calcular el valor de f en cada uno de los puntos obtenidos en el Paso 1 y también en a y en b.*

***Paso 3**. Comparar los valores obtenidos en el Paso 2. El mayor de todos ello será el máximo absoluto de f en $[a,b]$ y el menor será el mínimo absoluto de f en $[a,b]$.*

65. Calcula los valores máximo y mínimo de las siguientes funciones en los intervalos que se indican:

    a) $f(x)=x^3-x^2-8x+1$ en el intervalo $[-2,2]$.

    b) $\dfrac{x+1}{x^2+1}$ en el intervalo $[-1,2]$.

    c) $f(x)=\dfrac{1}{2}(\operatorname{sen}^2 x+\cos x)+2\operatorname{sen} x-x$ en el intervalo $[0,\pi/2]$.

    d) $f(x)=\sqrt[3]{x^2}\,(5-2x)$ en el intervalo $[-1,2]$.

    e) $f(x)=-x^3+12x+5$ en el intervalo $[-3,3]$.

Cuando una función no está definida en un intervalo cerrado hay que estudiar el signo de la derivada si queremos calcular máximos o mínimos absolutos cuya existencia habrá que justificar.

66. Calcula el mínimo valor de $\sum_{k=1}^{n}(x-a_k)^2$ donde $a_1,a_2,\cdots a_n$ son números reales dados.

67. Calcula la imagen de $f:\mathbb{R}^+\to\mathbb{R}$ dada por $f(x)=x^{1/x}$.

68. Sea $f:\mathbb{R}\to\mathbb{R}$ la función definida por $f(x)=\mathrm{e}^{-1/x^2}$ para $x\neq 0$, y $f(0)=0$. Estudia la continuidad y derivabilidad de f y calcula su imagen.

Los ejercicios que siguen son de cálculo de límites de sucesiones. Deberás usar los criterios de Stolz y de las medias aritmética y geométrica y el criterio de equivalencia logarítmica. En general, debes seguir la estrategia básica de relacionar un límite de una sucesión con un límite funcional apropiado.

69. Supongamos que $\{x_n\}\to 0$, siendo $-1<x_n\neq 0$, y sea $\alpha\in\mathbb{R}^*$. Prueba que $\{(1+x_n)^\alpha-1\}$ es asintóticamente equivalente a $\{\alpha x_n\}$.

70. Prueba que la sucesión $\{\log n!\}$ es asintóticamente equivalente a $\{n\log n\}$.

71. Prueba que la sucesión $\left\{\sqrt[n]{1+1/n^\alpha}-1\right\}$ es asintóticamente equivalente a $\left\{1/n^{\alpha+1}\right\}$, donde $\alpha>0$.

72. Calcula los límites de las sucesiones $\{x_n\}$ definidas por:

    a) $x_n=\dfrac{1^\alpha+2^\alpha+3^\alpha+\cdots+n^\alpha}{n^{\alpha+1}}$, donde $\alpha>-1$.

    b) $x_n=\sqrt[k]{(n+a_1)(n+a_2)\cdots(n+a_k)}\,-\,n$, donde $k\in\mathbb{N}$, $a_j\in\mathbb{R}$, $1\leqslant j\leqslant k$.

    c) $x_n=\left(\dfrac{\alpha\sqrt[n]{a}+\beta\sqrt[n]{b}}{\alpha+\beta}\right)^n$ donde $a>0$, $b>0$ y $\alpha,\beta\in\mathbb{R}$, $\alpha+\beta\neq 0$.

    d) $x_n=\left(\dfrac{1+2^{p/n}+3^{p/n}+\cdots+p^{p/n}}{p}\right)^n$, donde $p\in\mathbb{N}$.

e) $x_n = n\left(\dfrac{1+2^k+3^k+\cdots+n^k}{n^{k+1}} - \dfrac{1}{k+1}\right)$, donde $k\in\mathbb{N}$.

f) $x_n = \left(\dfrac{3}{4}\dfrac{1+3^2+5^2+\cdots+(2n-1)^2}{n^3}\right)^{n^2}$

g) $x_n = n\left[\left(1+\dfrac{1}{n^3\log(1+1/n)}\right)^n - 1\right]$

h) $x_n = \dfrac{1}{n}\left(n+\dfrac{n-1}{2}+\dfrac{n-2}{3}+\cdots+\dfrac{2}{n-1}+\dfrac{1}{n}-\log(n!)\right)$

73. Calcula los límites de las sucesiones $\{x_n\}$ definidas por:

a) $x_n = \dfrac{\log\left(1+\frac{1}{2}+\cdots+\frac{1}{n}\right)}{\log(\log n)}$; **b)** $x_n = \dfrac{e\sqrt{e}\sqrt[3]{e}\cdots\sqrt[n]{e}}{n}$

c) $x_n = \left(1+\dfrac{\log n}{n^\alpha}\right)^n$ $(\alpha>0)$; **d)** $x_n = \left(\dfrac{\log(n+2)}{\log(n+1)}\right)^{n\log n}$

e) $x_n = \dfrac{1}{n}\displaystyle\sum_{k=1}^{n}\dfrac{1}{k}\log\prod_{j=1}^{k}\left(1+\dfrac{1}{j}\right)^j$; **f)** $x_n = \dfrac{(2\sqrt[n]{n}-1)^n}{n^2}$

g) $x_n = \log n\left[\left(\dfrac{\log(n+1)}{\log n}\right)^n - 1\right]$; **h)** $x_n = \sqrt[n]{\dfrac{(pn)!}{(qn)^{pn}}}$ $(p,q\in\mathbb{N})$

74. Sabiendo que $\{a_n\}\to a$, calcula el límite de las sucesiones:

a) $x_n = n(\sqrt[n]{a_n}-1)$

b) $x_n = \dfrac{\exp(a_1)+\exp(a_2/2)+\cdots+\exp(a_n/n)-n}{\log n}$

c) $x_n = \dfrac{a_1+a_2/2+\cdots+a_n/n}{\log n}$

75. Sea $\{x_n\}$ una sucesión de números positivos tal que $\left\{\dfrac{x_{n+1}}{x_n}\right\}\to L\in\mathbb{R}^+$. Calcula el límite de la sucesión

$$\sqrt[n]{\frac{x_n}{\sqrt[n]{x_1x_2\cdots x_n}}}.$$

76. Sea $\{x_n\}$ una sucesión de números positivos, α un número real, y supongamos que $\{n^\alpha x_n\}\to L\in\mathbb{R}_o^+$. Calcula el límite de la sucesión $n^\alpha\sqrt[n]{x_1x_2\cdots x_n}$.

77. Sean a, b números positivos; definamos $x_k = a+(k-1)b$ para cada $k\in\mathbb{N}$ y sea G_n la media geométrica de $x_1, x_2,\dots,x_n$ y A_n su media aritmética. Calcula el límite de la sucesión $\dfrac{G_n}{A_n}$.

78. Sea $\{x_n\}\to x$, $\{y_n\}\to y$, $x\neq y$. Definamos $z_{2n-1}=x_n$, y $z_{2n}=y_n$. Justifica que la sucesión

$$\left\{\frac{z_1+z_2+\cdots+z_n}{n}\right\}$$

es convergente.

79. Sean $\{x_n\}$, $\{y_n\}$ sucesiones de números positivos tales que $\{(x_n)^n\}\to x>0$ $\{(y_n)^n\}\to y>0$. Dados $\alpha,\beta\in\mathbb{R}^+$, con $\alpha+\beta=1$, calcula $\text{lím}(\alpha x_n+\beta y_n)^n$.

80. Sea $\{a_n\}$ una sucesión de números positivos tal que $\{a_1+a_2+\cdots+a_n\}$ es divergente, y sea $\{b_n\}\to L$, donde L puede ser un número real o $\pm\infty$. Justifica que

$$\left\{\frac{a_1b_1+a_2b_2+\cdots+a_nb_n}{a_1+a_2+\cdots+a_n}\right\}\to L.$$

Aplicación. Supuesto que $\{x_n\}\to x$, calcula $\displaystyle \lim \frac{1}{2^n}\sum_{k=1}^{n}\binom{n}{k}x_k$.

Acabamos esta larga relación con algunos ejercicios que me ha parecido que no encajaban propiamente en ninguno de los apartados anteriores.

81. Supongamos que f es derivable en a, g es continua en a y $f(a)=0$. Prueba que fg es derivable en a.

82. Sea $f:[a,b]\to\mathbb{R}$ derivable y f' creciente. Prueba que la función $g:]a,b]\to\mathbb{R}$ dada para todo $x\in]a,b]$ por $g(x)=\dfrac{f(x)-f(a)}{x-a}$, es creciente.

83. Sea $f:[a,b]\to\mathbb{R}$ continua en $[a,b]$ y derivable dos veces en $]a,b[$. Supongamos que el segmento de extremos $(a,f(a))$, $(b,f(b))$ corta a la gráfica de f en un punto $(c,f(c))$ con $a<c<b$. Demuestra que existe algún punto $d\in]a,b[$ tal que $f''(d)=0$.
Sugerencia: interpreta gráficamente el enunciado.

84. Justifica que existe una función $g:\mathbb{R}\to\mathbb{R}$ derivable y que verifica que $g(x)+\mathrm{e}^{g(x)}=x$ para todo $x\in\mathbb{R}$. Calcula $g'(1)$ y $g'(1+\mathrm{e})$.

Lección 8

Integral de Riemann

Introducción

El cálculo integral tiene sus orígenes en problemas de *cuadraturas* en los que se trataba de calcular áreas de regiones planas limitadas por una o varias curvas. Se atribuye a Eudoxo (ca. 370 A.C.) la invención del método de *exhaución*, una técnica para calcular el área de una región aproximándola por una sucesión de polígonos de forma que en cada paso se mejora la aproximación anterior. Arquímides (287-212 A.C.) perfeccionó este método y, entre otros resultados, calculó el área de un segmento de parábola y el volumen de un segmento de paraboloide, así como el área y el volumen de una esfera.

Sorprende que, siendo tan antiguos sus orígenes, la primera definición matemática de integral no fuera dada hasta el siglo XIX por Augustin Louis Cauchy (1789-1857). Una posible explicación es que, durante los siglos XVII y XVIII, la integración fue considerada como la operación inversa de la derivación; el cálculo integral consistía esencialmente en el cálculo de primitivas. Naturalmente, se conocía la utilidad de las integrales para calcular áreas y volúmenes, pero los matemáticos de la época consideraban estas nociones como dadas de forma intuitiva y no vieron la necesidad de precisar su significación matemática. Los trabajos de Joseph Fourier (1768-1830) sobre representación de funciones por series trigonométricas hicieron que el concepto de función evolucionara, desde la idea restrictiva de función como fórmula, hasta la definición moderna de función dada por Dirichlet en 1837. Para entender el significado de la integral de estas nuevas funciones más generales se vio la necesidad de precisar matemáticamente los conceptos de área y de volumen.

La originalidad de Cauchy es que unió dos ideas, la de límite y la de área, para dar una definición matemática de integral. Poco después Georg F.B. Riemann (1826-1866) generalizó la definición de integral dada por Cauchy. La teoría de la integral de Riemann fue un avance importante pero, desde un punto de vista matemático, insuficiente. Hubo que esperar hasta el siglo XX para que Henri Lebesgue (1875-1941) estableciera en su libro *Leçons sur l'intégration et la recherché des fonctions primitives* (1904) los fundamentos de una teoría matemáticamente satisfactoria de la integración.

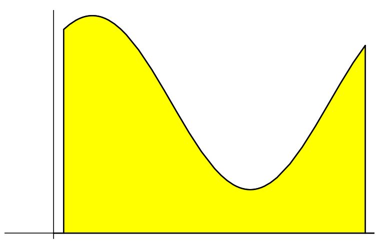
La integración es una de las herramientas más versátiles del Cálculo, sus aplicaciones no se limitan a calcular áreas de regiones planas o volúmenes de sólidos, también se utiliza para calcular longitudes de curvas, centros de masas, momentos de inercia, áreas de superficies, para representar magnitudes físicas como el trabajo, la fuerza ejercida por una presión, o la energía potencial en un campo de fuerzas.

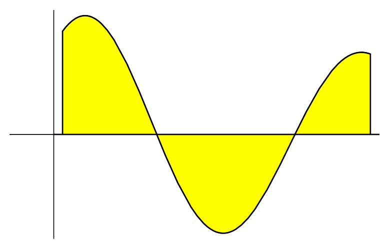
En este curso vamos a estudiar la integración desde un punto de vista esencialmente práctico. Nos interesa la integral como herramienta de cálculo y para ese propósito es suficiente la integral de Riemann.

Todo lo que sigue puedes verlo también como página Web y en formato de cuaderno de *Mathematica* en el sitio http://wwww.ugr.es/local/fjperez.

8.1.1. Sumas de Riemann

Sea $f:[a,b]\to\mathbb{R}$ una función acotada. Representaremos por $G(f,a,b)$ la región del plano comprendida entre la gráfica $y=f(x)$, el eje de abscisas y las rectas $x=a$ y $x=b$. Aquí puedes ver dos de estas regiones coloreadas en amarillo.





Nos proponemos calcular el área de regiones de este tipo. Puesto que, en general, $G(f,a,b)$ no puede descomponerse en triángulos o rectángulos, no hay una *fórmula* que nos permita calcular directamente su área.

En situaciones como esta, una estrategia básica consiste en obtener *soluciones aproximadas que permitan definir el valor exacto del área como límite de las mismas*. Fíjate que, al proceder así, *estamos definiendo dicho valor exacto*, es decir, *estamos dando una definición matemática del concepto intuitivo de área*[1]. Naturalmente, queremos que dicha definición sea lo más general posible, *lo que depende del tipo de soluciones aproximadas que elijamos.* Las aproximaciones consideradas en la teoría de la integral de Lebesgue conducen a un concepto de área muy general. En lo que sigue vamos a considerar las aproximaciones que conducen a la integral de Riemann.

Parte positiva y parte negativa de una función

Como los conceptos que vamos a introducir se interpretan con más facilidad cuando la función f es positiva, es conveniente tener bien presente en lo que sigue el siguiente artificio

[1] Ello trae como consecuencia inevitable que haya regiones extrañas en el plano que, según la definición dada, no tengan área.

que permite representar cualquier función como diferencia de dos funciones positivas.

Cualquier función f puede escribirse como diferencia de dos funciones positivas:

$$f^+(x) = \frac{|f(x)| + f(x)}{2} \qquad f^-(x) = \frac{|f(x)| - f(x)}{2}$$

Es claro que $f(x) = f^+(x) - f^-(x)$ y que $f^+(x) \geqslant 0$, $f^-(x) \geqslant 0$. La función f^+ se llama **parte positiva** de f, y la función f^- se llama **parte negativa** de f. Si $f(x) \geqslant 0$ se tiene que $f(x) = f^+(x)$ y $f^-(x) = 0$; mientras que si $f(x) \leqslant 0$ se tiene que $f(x) = -f^-(x)$ y $f^+(x) = 0$. Fíjate que, a pesar de su nombre y de la forma en que se simboliza, la función f^- es una función positiva. También es consecuencia de las definiciones dadas que $|f(x)| = f^+(x) + f^-(x)$.

En lo que sigue, representaremos el valor exacto (que aún no hemos definido) del área de la región $G(f,a,b)$ por $\lambda(G(f,a,b))$ (la letra "λ" alude a la inicial de "Lebesgu"). En la integral de Riemann, el área buscada se aproxima por rectángulos de la siguiente forma. Primero, se divide el intervalo $[a,b]$ en un número finito de subintervalos $[x_{k-1},x_k]$, $1 \leqslant k \leqslant n$, cuyas longitudes pueden ser distintas y con la única condición de que no se solapen:

$$a = x_0 < x_1 < x_2 < \cdots < x_{n-1} < x_n = b$$

Se dice que estos puntos constituyen una **partición** de $[a,b]$. A continuación se elige en cada subintervalo un punto $t_k \in [x_{k-1},x_k]$, y se forma el rectángulo cuya base es el intervalo $[x_{k-1},x_k]$ y altura igual a $f(t_k)$. Dicho rectángulo está en el semiplano superior si $f(t_k) > 0$ y en el semiplano inferior si $f(t_k) < 0$. Finalmente se forma la suma $\displaystyle\sum_{k=1}^{n} f(t_k)(x_k - x_{k-1})$.

8.1 Definición. Dada una partición $P = \{a = x_0, x_1, x_2, \ldots, x_{n-1}, x_n = b\}$ del intervalo $[a,b]$, y un punto $t_k \in [x_{k-1},x_k]$ en cada uno de los intervalos de la misma, el número

$$\sigma(f,P) = \sum_{k=1}^{n} f(t_k)(x_k - x_{k-1})$$

se llama una **suma de Riemann** de f para la partición P.

Observaciones

• Fíjate que, como hay libertad para elegir los puntos $t_k \in [x_{k-1},x_k]$, para cada partición P hay infinitas sumas de Riemann.

• Cuando la función f es positiva, $\sigma(f,P)$ es una aproximación del área de la región $G(f,a,b)$. Simbólicamente $\sigma(f,P) \approx \lambda(G(f,a,b))$.

• Cuando la función f toma valores positivos y negativos podemos escribir

$$\begin{aligned}\sigma(f,P) &= \sum_{k=1}^{n} f(t_k)(x_k - x_{k-1}) = \sum_{k=1}^{n} (f^+(t_k) - f^-(t_k))(x_k - x_{k-1}) = \\ &= \sum_{k=1}^{n} f^+(t_k)(x_k - x_{k-1}) - \sum_{k=1}^{n} f^-(t_k)(x_k - x_{k-1}) = \sigma(f^+,P) - \sigma(f^-,P)\end{aligned}$$

En este caso $\sigma(f,P)$ es una aproximación del área de $G(f^+,a,b)$ menos el área de $G(f^-,a,b)$. Simbólicamente $\sigma(f,P) \approx \lambda(G(f^+,a,b)) - \lambda(G(f^-,a,b))$.

En la siguiente figura pueden apreciarse estas aproximaciones.

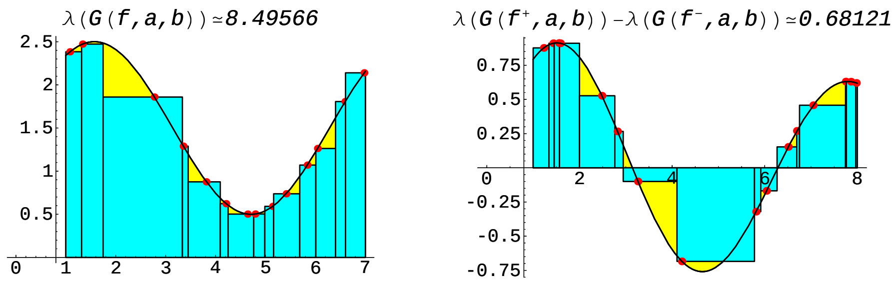


Figura 8.1: Aproximación del área por sumas de Riemann

8.2 Definición. Dada una partición $P = \{a = x_0, x_1, x_2, \ldots, x_{n-1}, x_n = b\}$ del intervalo $[a,b]$, definamos $M_k = \sup f[x_{k-1}, x_k]$, $m_k = \inf f[x_{k-1}, x_k]$. Los números

$$S(f,P) = \sum_{k=1}^{n} M_k(x_k - x_{k-1}), \quad I(f,P) = \sum_{k=1}^{n} m_k(x_k - x_{k-1})$$

se llaman, respectivamente, **suma superior y suma inferior** de f para la partición P. Es para definir estas sumas para lo que se precisa que f esté acotada en $[a,b]$.

Observaciones

• Puesto que para todo $t_k \in [x_{k-1}, x_k]$ es $m_k \leqslant f(t_k) \leqslant M_k$, deducimos que para toda suma de Riemann, $\sigma(f,P)$ de f para la partición P se cumple que $I(f,P) \leqslant \sigma(f,P) \leqslant S(f,P)$.

• Para cada partición hay una única suma superior y otra inferior.

• Cuando f es positiva $S(f,P)$ es un *valor aproximado por exceso* de $\lambda(G(f,a,b))$, y $I(f,P)$ es un *valor aproximado por defecto* de $\lambda(G(f,a,b))$.

• Cuando la función f toma valores positivos y negativos $S(f,P)$ es un *valor aproximado por exceso* de $\lambda(G(f^+,a,b)) - \lambda(G(f^-,a,b))$, y $I(f,P)$ es un *valor aproximado por defecto* de $\lambda(G(f^+,a,b)) - \lambda(G(f^-,a,b))$.

En las siguientes figuras pueden apreciarse estas aproximaciones.

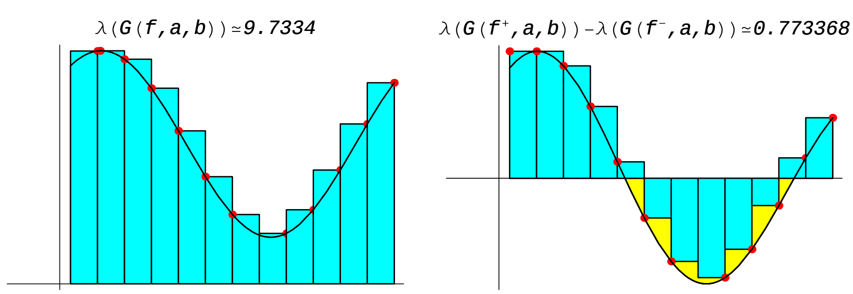


Figura 8.2: Aproximación del área por sumas superiores

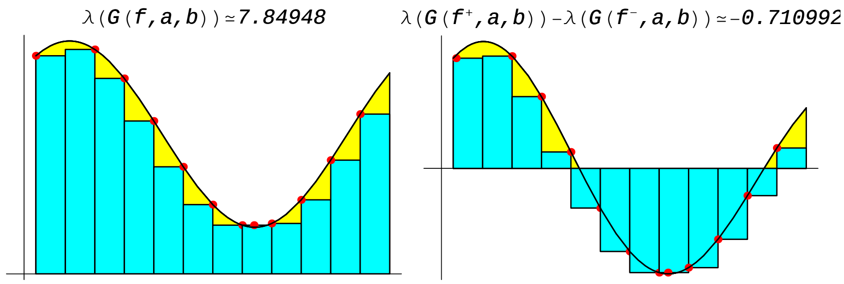


Figura 8.3: Aproximación del área por sumas inferiores

8.1.2. Definición y propiedades básicas de la integral

Supongamos que la función f es positiva en $[a,b]$. Es claro que, en tal caso, el valor exacto del área de la región $G(f,a,b)$ debe verificar que $I(f,P) \leqslant \lambda(G(f,a,b)) \leqslant S(f,P)$ para toda partición P de $[a,b]$. Tenemos, en consecuencia, ***dos*** candidatos para $\lambda(G(f,a,b))$, a saber:

$$\lambda(G(f,a,b)) = \text{ínf}\{S(f,P) : P \in \mathfrak{P}[a,b]\}, \quad \lambda(G(f,a,b)) = \sup\{I(f,P) : P \in \mathfrak{P}[a,b]\}$$

Donde hemos representado por $\mathfrak{P}[a,b]$ el conjunto de *todas* las particiones de $[a,b]$. Llegados aquí, podemos ya dar la definición principal de la teoría de la integral de Riemann.

8.3 Definición. Sea f una función acotada y positiva en $[a,b]$. Se dice que el conjunto $G(f,a,b)$ **tiene área** cuando

$$\text{ínf}\{S(f,P) : P \in \mathfrak{P}[a,b]\} = \sup\{I(f,P) : P \in \mathfrak{P}[a,b]\}$$

Dicho valor común es, por definición, el valor del área y lo representaremos por $\lambda(G(f,a,b))$. Cuando esto ocurre, se dice también que la función f **es integrable Riemann** en $[a,b]$ y, por definición, la integral de f en $[a,b]$ es igual a $\lambda(G(f,a,b))$. Simbólicamente escribimos

$$\int_a^b f(x)\,\mathrm{d}x = \lambda(G(f,a,b))$$

En el caso general en que la función f toma valores positivos y negativos, se dice que f es integrable Riemann en $[a,b]$ cuando lo son las funciones f^+ y f^-, en cuyo caso se define la integral de f en $[a,b]$ como el número:

$$\int_a^b f(x)\,\mathrm{d}x = \lambda(G(f^+,a,b)) - \lambda(G(f^-,a,b))$$

Observaciones

- No te confundas con la notación. El símbolo $\int_a^b f(x)\,\mathrm{d}x$ representa un número. La variable x que figura en él se suele decir que es una *variable muda*. Naturalmente, la letra x no tiene ningún significado especial y puede sustituirse por la que tú quieras o no poner ninguna; por ejemplo

$$\int_a^b f(t)\,\mathrm{d}t, \quad \int_a^b f(s)\,\mathrm{d}s, \quad \int_a^b f$$

son tres formas de escribir lo mismo. Volveremos sobre esta notación cuando estudiemos técnicas de integración.

• La definición anterior debes entenderla como una primera aproximación matemática al concepto intuitivo de área. Aunque pueda parecerte extraño, el concepto de área (y de integral) que acabamos de definir es bastante restrictivo.

• En el caso en que la función f toma valores positivos y negativos, observa que la gráfica de f^- se obtiene por simetría respecto al eje de abscisas de las partes de la gráfica de f en las que $f(x) < 0$. Como regiones simétricas respecto de una recta tienen la misma área, se sigue que:

$$\lambda(G(f,a,b)) = \lambda(G(f^+,a,b)) + \lambda(G(f^-,a,b)) = \lambda(G(f^+ + f^-,a,b)) = \lambda(G(|f|,a,b)) = \int_a^b |f(x)|\,\mathrm{d}x$$

Seamos prácticos. ¿Cómo podemos, a partir de la definición dada, calcular $\int_a^b f(x)\,\mathrm{d}x$? Una primera idea en este sentido consiste en observar que cuanto mayor sea el número de intervalos de la partición y más pequeña la anchura de cada uno de ellos cabe esperar que la aproximación obtenida sea mejor. Para precisar esta idea, definimos **el paso de una partición** P, y lo representamos por $\delta(P)$, como *la mayor de las longitudes de los subintervalos de dicha partición.*

8.4 Teorema (Convergencia de las sumas integrales). *Sea $f : [a,b] \to \mathbb{R}$ una función integrable, $\{P_n\}$ una sucesión de particiones de $[a,b]$ tal que $\{\delta(P_n)\} \to 0$ y $\sigma(f,P_n)$ una suma de Riemann de f para la partición P_n. Se verifica entonces que*

$$\lim_{n\to\infty} S(f,P_n) = \lim_{n\to\infty} \sigma(f,P_n) = \lim_{n\to\infty} I(f,P_n) = \int_a^b f(x)\,\mathrm{d}x$$

Este resultado permite en algunos casos particulares y con bastante esfuerzo e ingenio calcular ciertas integrales. Como más adelante aprenderemos a calcular integrales con facilidad, es más interesante usar dicho resultado *sensu contrario* para calcular los límites de ciertas sucesiones.

Ha llegado el momento de preguntarse por condiciones que garanticen que una función es integrable Riemann. Nos vamos a contentar con una respuesta parcial a esta pregunta, que es suficiente para nuestros propósitos.

8.5 Teorema (Condiciones suficientes de integrabilidad Riemann). *Sea $f : [a,b] \to \mathbb{R}$. Cada una de las siguientes condiciones garantizan que f es integrable Riemann en $[a,b]$.*

i) f está acotada en $[a,b]$ y tiene un número finito de discontinuidades en $[a,b]$. En particular, toda función continua en un intervalo cerrado y acotado es integrable en dicho intervalo.

ii) f es monótona en $[a,b]$.

Teniendo en cuenta que $\sigma(\alpha f + \beta g, P) = \alpha\sigma(f,P) + \beta\sigma(g,P)$, cualesquiera sean las funciones f, g y los números α, β, se deduce, haciendo uso del resultado sobre convergencia de sumas integrales, que la integral es lineal. Esta propiedad, junto con otras propiedades básicas de las integrales se recogen en el siguiente resultado.

8.6 Proposición (Propiedades básicas de la integral). *i)* ***Linealidad****. Si f,g son integrables en $[a,b]$ y α,β son números reales, se verifica que la función $\alpha f+\beta g$ también es integrable en $[a,b]$ y*

$$\int_a^b (\alpha f(x)+\beta g(x))\,\mathrm{d}x = \alpha\int_a^b f(x)\,\mathrm{d}x+\beta\int_a^b g(x)\,\mathrm{d}x$$

ii) ***Conservación del orden****. Si f,g son integrables en $[a,b]$ y $f(x)\leqslant g(x)$ para todo $x\in[a,b]$, entonces se verifica que*

$$\int_a^b f(x)\,\mathrm{d}x \leqslant \int_a^b g(x)\,\mathrm{d}x$$

En particular, si f es integrable en $[a,b]$ y $m\leqslant f(x)\leqslant M$ para todo $x\in[a,b]$, entonces se verifica la siguiente ***acotación fundamental***

$$m(b-a)\leqslant \int_a^b f(x)\,\mathrm{d}x \leqslant M(b-a)$$

iii) Si f es integrable en $[a,b]$ también $|f|$ (función valor absoluto de f) es integrable en $[a,b]$ y se verifica la desigualdad

$$\left|\int_a^b f(x)\,\mathrm{d}x\right| \leqslant \int_a^b |f(x)|\,\mathrm{d}x$$

iv) El producto de funciones integrables Riemann también es una función integrable Riemann.

v) ***Aditividad respecto del intervalo****. Sea $a<c<b$. Una función f es integrable en $[a,b]$ si, y sólo si, es integrable en $[a,c]$ y en $[c,b]$, en cuyo caso se verifica la igualdad*

$$\int_a^b f(x)\,\mathrm{d}x = \int_a^c f(x)\,\mathrm{d}x + \int_c^b f(x)\,\mathrm{d}x$$

8.1.3. El Teorema Fundamental del Cálculo

Dada una función integrable $f:[a,b]\to\mathbb{R}$, podemos definir una nueva función $F:[a,b]\to\mathbb{R}$ por

$$F(x)=\int_a^x f(t)\,\mathrm{d}t \qquad \text{para todo } x\in[a,b]$$

Observa que aquí la variable es x – el límite superior de la integral. Por eso, es obligado no usar la misma letra x como variable de la función f en el integrando. $F(x)$ es la integral de la función f en el intervalo $[a,x]$.

Sabemos que $F(x)=\lambda(G(f^+,a,x))-\lambda(G(f^-,a,x))$. Por supuesto, si f es una positiva entonces $F(x)=\lambda(G(f,a,x))$ es el área de la región del plano limitada por la gráfica de f, el eje de abscisas y las rectas verticales $X=a$, $X=x$. No debes olvidar en lo que sigue que $F(x)=\int_a^x f(t)\,\mathrm{d}t$ se ha definido en términos de áreas. A la función F la llamaremos *la función área de f*.

A veces hay que considerar funciones de la forma $H(x)=\int_c^x f(t)\,\mathrm{d}t$ en donde $a<c<b$ y $x\in[a,b]$; por lo que es necesario precisar lo que se entiende por $\int_c^x f(t)\,\mathrm{d}t$ cuando $x<c$. El convenio que se hace es que

$$\int_u^v f(t)\,\mathrm{d}t = -\int_v^u f(t)\,\mathrm{d}t$$

cualquiera sean los números u y v. La justificación de este convenio es que, con él, la igualdad

$$\int_x^y f(t)\,\mathrm{d}t + \int_y^z f(t)\,\mathrm{d}t + \int_z^x f(t)\,\mathrm{d}t = 0$$

se cumple cualesquiera sean los puntos x,y,z del intervalo $[a,b]$. Compruébalo.

Nuestro próximo objetivo va a consistir en invertir el proceso que nos ha llevado de f a $F(x)=\int_a^x f(t)\,\mathrm{d}t$. Nuestro problema es: ¿Cómo podemos recuperar la función f a partir del conocimiento de la función área de f? El resultado que sigue, uno de los más útiles del Cálculo, establece una relación entre dos conceptos aparentemente lejanos entre sí: el concepto de área y el de tangente a una curva.

8.7 Teorema (Teorema Fundamental del Cálculo). *Sea $f:[a,b]\to\mathbb{R}$ una función integrable y definamos $F:[a,b]\to\mathbb{R}$ por*

$$F(x)=\int_a^x f(t)\,\mathrm{d}t \qquad \textit{para todo } x\in[a,b]$$

Entonces:

i) F es continua en $[a,b]$.

ii) En todo punto c de $[a,b]$ en el que f sea continua se verifica que F es derivable en dicho punto siendo $F'(c)=f(c)$. En particular, si f es continua en $[a,b]$, entonces F es derivable en $[a,b]$ y $F'(x)=f(x)$ para todo $x\in[a,b]$.

Demostración.

i) Como f es integrable debe estar acotada. Sea $M>0$ tal que $|f(x)|\leqslant M$ para todo $x\in[a,b]$. Entonces, si $x<y$ son puntos de $[a,b]$ tenemos que

$$|F(y)-F(x)| = \left|\int_x^y f(t)\,\mathrm{d}t\right| \leqslant \int_x^y |f(t)|\,\mathrm{d}t \leqslant M(y-x)$$

Por la misma razón, si suponemos que $y<x$, tendremos que $|F(y)-F(x)|\leqslant M(y-x)$. Estas dos desigualdades nos dicen que $|F(y)-F(x)|\leqslant M|y-x|$ para todo par de puntos $x,y\in[a,b]$. De esta desigualdad se sigue inmediatamente la continuidad de F en $[a,b]$.

ii) Pongamos

$$\frac{F(x)-F(c)}{x-c}-f(c)=\frac{F(x)-F(c)-(x-c)f(c)}{x-c}=\frac{\int_c^x f(t)\,\mathrm{d}t-\int_c^x f(c)\,\mathrm{d}t}{x-c}=\frac{\int_c^x (f(t)-f(c))\,\mathrm{d}t}{x-c}$$

Dado, $\varepsilon > 0$, la continuidad de f en c nos dice que hay un $\delta > 0$ tal que para todo $t \in [a,b]$ con $|t-c| < \delta$ se tiene que $|f(t) - f(c)| < \varepsilon$. Tomemos ahora un punto cualquiera $x \in [a,b]$ tal que $|x-c| < \delta$. Entonces es claro que para todo t comprendido entre x y c se tendrá que $|t-c| < \delta$ y, por tanto, $|f(t) - f(c)| < \varepsilon$ por lo que

$$\left| \int_c^x (f(t) - f(c))\,\mathrm{d}t \right| \leqslant \varepsilon\, |x-c|$$

Deducimos que para todo $x \in [a,b]$ tal que $|x-c| < \delta$, y $x \neq c$, se verifica que

$$\left| \frac{F(x) - F(c)}{x-c} - f(c) \right| = \left| \frac{\int_c^x (f(t) - f(c))\,\mathrm{d}t}{x-c} \right| \leqslant \frac{\varepsilon\,|x-c|}{|x-c|} = \varepsilon$$

Hemos probado así que $\lim\limits_{x \to c} \dfrac{F(x) - F(c)}{x-c} = f(c)$, esto es, F es derivable en c y $F'(c) = f(c)$. □

8.8 Definición. Dada un función $h\colon [a,b] \to \mathbb{R}$, cualquier función $H\colon [a,b] \to \mathbb{R}$ que sea continua en $[a,b]$, derivable en $]a,b[$ y verifique que $H'(x) = h(x)$ para todo $x \in]a,b[$, se llama una **primitiva** de f en el intervalo $[a,b]$.

Es importante advertir que no todas las funciones tienen primitivas. Por ejemplo, una *condición necesaria* que debe cumplir una función para tener primitivas es que dicha función tenga la propiedad del valor intermedio pues, como recordarás, las funciones derivadas tienen esa propiedad. También, como consecuencia del teorema del valor medio, es inmediato que **dos primitivas de una función en un mismo intervalo se diferencian en una constante**. Por ello, si conocemos una primitiva de una función en un intervalo las conocemos todas.

Una consecuencia muy importante del Teorema Fundamental del Cálculo es que ***toda función continua en un intervalo tiene primitivas en dicho intervalo***. Además, el teorema nos dice que la función área, esto es, la función $F(x) = \int_a^x f(t)\,\mathrm{d}t$, es la primitiva de la función continua $f\colon [a,b] \to \mathbb{R}$ que se anula en a, $F(a) = 0$. **Es importante que aprecies que este es un teorema de existencia**; es la definición que hemos dado de área - y por consiguiente de integral - lo que nos ha permitido **construir** la función primitiva de f. No lo olvides: **la integración es una potente herramienta para construir nuevas funciones.**

El Teorema Fundamental del Cálculo proporciona una técnica para calcular la integral de una *función continua* en un intervalo $[a,b]$. Para ello lo que hacemos es calcular una primitiva de f en $[a,b]$. Si h es una tal primitiva, entonces las funciones $F(x) = \int_a^x f(t)\,\mathrm{d}t$, y $h(x) - h(a)$ son dos primitivas de f en $[a,b]$ que coinciden en un punto, pues ambas se anulan en a. Deducimos que $F(x) = h(x) - h(a)$ para todo $x \in [a,b]$ y, por tanto, $F(b) = \int_a^b f(t)\,\mathrm{d}t = h(b) - h(a)$. Podemos generalizar este resultado como sigue.

8.9 Teorema (Regla de Barrow). *Sea $f\colon [a,b] \to \mathbb{R}$ integrable y supongamos que h es una primitiva de f en $[a,b]$. Entonces*

$$\int_a^b f(t)\,\mathrm{d}t = h(b) - h(a)$$

Demostración. Sea $P=\{a=x_0,x_1,x_2,\dots,x_{n-1},x_n=b\}$ una partición de $[a,b]$. Aplicando el teorema de valor medio, tenemos que

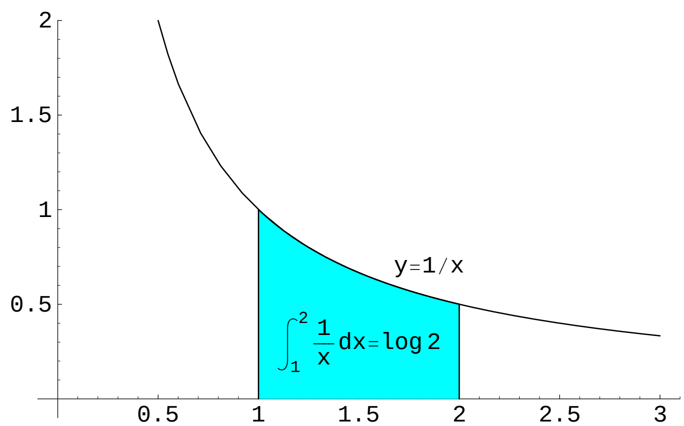
$$h(b)-h(a)=\sum_{k=1}^{n}(h(x_k)-h(x_{k-1}))=\sum_{k=1}^{n}f(t_k)(x_k-x_{k-1})=\sigma(f,P)$$

La igualdad anterior nos dice que para toda partición P de $[a,b]$ hay *alguna* suma de Riemann de f asociada a dicha partición, $\sigma(f,P)$, que es igual a $h(b)-h(a)$. Si ahora tomamos una sucesión $\{P_n\}$ de particiones del intervalo $[a,b]$ tales que $\delta(P_n)\to 0$, tenemos que $h(b)-h(a)=\sigma(f,P_n)$ para alguna suma de Riemann, $\sigma(f,P_n)$ de f asociada a la partición P_n. Pero sabemos que $\sigma(f,P_n)\to\int_a^b f$, por lo que obtenemos que $h(b)-h(a)=\int_a^b f$. □

Fíjate que en la regla de Barrow no se supone que f sea continua sino tan sólo que es integrable y que, además, tiene una primitiva.

8.1.4. Las funciones logaritmo y exponencial

Quiero convencerte de que muchas veces el cálculo integral proporciona la interpretación más intuitiva de una función. Considera, por ejemplo, la función logaritmo natural. Quizás sepas expresar $\log 2$ como límite de una sucesión o algo parecido; pero, ¿puedes representar de alguna forma intuitiva el número $\log 2$? ¿Sabrías representar gráficamente el número $\log 2$? En la siguiente gráfica puedes *ver* el número $\log 2$.



Espero que estés de acuerdo conmigo: la forma más fácil e intuitiva de imaginar el número $\log t$ es como el área de la región plana limitada por la curva $y=1/x$, las rectas $y=1$, $y=t$, y el eje de abscisas. Dicha área se considera positiva si $t>1$ y negativa si $t<1$. Dicho de otra forma

$$\log t=\int_1^t \frac{1}{x}\,dx$$

Es frecuente interpretar esta igualdad de la siguiente forma: la función $\log x$ es derivable y $\log' x=1/x$; por tanto $\displaystyle\int_1^t \frac{1}{x}\,dx=\log t-\log 1=\log t$. ¡Parece que hemos probado algo! Y no es así porque en este razonamiento estamos usando que la función logaritmo es derivable y eso es algo que no hemos probado. Todavía peor: ni siquiera hemos dado una definición de la función logaritmo que permita probar las propiedades de dicha función. Usualmente se define $\log x$ como el número y que verifica que $e^y = x$. La existencia de ese número y está lejos de ser evidente. El propio número e tiene que ser definido de alguna forma apropiada.

Hago estas reflexiones para que te des cuenta de que lo que sabes de las funciones logaritmo, exponencial, trigonométricas ..., es un conocimiento sin una base matemática correcta. De estas funciones conoces, porque te lo han dicho, su comportamiento; pero no creo que nunca hayas demostrado sus propiedades, ni siquiera que conozcas una definición matemáticamente correcta de las mismas. Bueno, no quiero que pienses que tus profesores te ocultan información, lo que ocurre es que una definición correcta de estas funciones requiere herramientas matemáticas que no tienen cabida en las enseñanzas medias. Precisamente, el Teorema Fundamental del Cálculo permite definir estas funciones de forma fácil, elegante y correcta.

Olvida ahora todo lo que sepas de la función logaritmo natural. ¿Lo has olvidado ya? Sigamos.

8.10 Definición. La función *logaritmo natural* es la función $\log : \mathbb{R}^+ \to \mathbb{R}$ dada por

$$\log t = \int_1^t \frac{1}{x}\,\mathrm{d}x \qquad \text{para todo } t > 0$$

Propiedades de la función logaritmo

El Teorema Fundamental del Cálculo nos dice que la función *logaritmo natural* es *derivable* (y por tanto continua) y que $\log' t = 1/t$. Como la derivada es positiva, deducimos que dicha función es *estrictamente creciente*.

Dado $a > 0$, sea $h(x) = \log(ax)$. Entonces $h'(x) = a/(ax) = 1/x$. Luego la función $h(x) - \log(x)$ tiene derivada nula en $\mathbb{R}^+$, por lo que es constante y, como para $x = 1$ es igual a $\log a$, se sigue que $h(x) - \log(x) = \log a$. Hemos probado así que $\log(ax) = \log a + \log x$ para todo $a > 0$ y para todo $x > 0$.

Observa que en poco más de tres líneas hemos obtenido ya las propiedades principales del logaritmo. Sigamos nuestro estudio.

De lo ya visto se sigue que $\log(2^n) = n \log 2$ para todo número entero n. De aquí se deduce que la función *logaritmo natural* no está mayorada ni minorada y, como es estrictamente creciente, concluimos que $\lim\limits_{x\to 0} \log x = -\infty$ y $\lim\limits_{x\to+\infty} \log x = +\infty$. Por tanto, podemos afirmar que dicha función es una *biyección estrictamente creciente de* $\mathbb{R}^+$ *sobre* $\mathbb{R}$.

Representemos provisionalmente por $\varphi : \mathbb{R} \to \mathbb{R}$ la función inversa del logaritmo. Dicha función se llama función *exponencial natural*. El teorema de derivación de la inversa nos dice que φ es derivable y para todo $x \in \mathbb{R}$ es

$$\varphi'(x) = \frac{1}{\log'(\varphi(x))} = \varphi(x)$$

Ahora, dados,$x, y \in \mathbb{R}$, sean $a, b, \in \mathbb{R}^+$ tales que $x = \log a$, $y = \log b$. Entonces

$$\varphi(x + y) = \varphi(\log a + \log b) = \varphi(\log(ab)) = ab = \varphi(x)\varphi(y)$$

Hemos probado así que $\varphi(x + y) = \varphi(x)\varphi(y)$ para todos $x, y \in \mathbb{R}$. De esta igualdad se deduce fácilmente que apara todo número racional r se verifica que $\varphi(r) = \varphi(1)^r$. El número $\varphi(1)$ se representa con la letra e, es decir, es el número definido por la igualdad $\log \mathrm{e} = \int\limits_1^{\mathrm{e}} \frac{1}{x}\,\mathrm{d}x = 1$. Con ello apara todo número racional r se tiene que $\varphi(r) = \mathrm{e}^r$, por lo que se usa la notación $\varphi(x) = \mathrm{e}^x$ para representar a la función exponencial.

Fíjate con qué facilidad y elegancia hemos obtenido las propiedades principales de las funciones logaritmo y exponencial naturales.

Así mismo, podemos ***definir*** la función *arcotangente* de la forma

$$\operatorname{arc\,tg} x = \int\limits_0^x \frac{1}{1+t^2}\,\mathrm{d}t$$

Lo que constituye un punto de partida para definir las demás funciones trigonométricas. Este proceso está desarrollado con detalle en el libro de *Michael Spivak* Calculo etcétera. Veremos más adelante otro procedimiento más directo para definir las funciones trigonométricas.

8.2. Integrales impropias de Riemann

Una de las limitaciones de la teoría de la integral de Riemann es que en ella se consideran funciones acotadas en intervalos acotados. Queremos evitar estas limitaciones y considerar funciones no acotadas o intervalos no acotados. Los siguientes ejemplos indican el camino a seguir.

8.11 Ejemplo. La función $f(x) = \frac{1}{\sqrt{x}}$ no está acotada en el intervalo $]0,1]$. Como $h(x) = 2\sqrt{x}$ es una primitiva de f en $[0,1]$, para todo $t \in]0,1]$ se tiene que

$$\int\limits_t^1 \frac{1}{\sqrt{x}}\,\mathrm{d}x = h(1) - h(t) = 2 - 2\sqrt{t}$$

Por tanto

$$\lim_{t\to 0} \int\limits_t^1 \frac{1}{\sqrt{x}}\,\mathrm{d}x = 2$$

Es natural definir

$$\int\limits_0^1 \frac{1}{\sqrt{x}}\,\mathrm{d}x = 2$$

♦

8.12 Ejemplo. Para todo $t > 0$ se tiene que

$$\int\limits_0^t \mathrm{e}^{-x}\,\mathrm{d}x = 1 - \mathrm{e}^{-t} \Longrightarrow \lim_{t\to+\infty} \int\limits_0^t \mathrm{e}^{-x}\,\mathrm{d}x = 1$$

Por ello es natural definir

$$\int\limits_0^{+\infty} \mathrm{e}^{-x}\,\mathrm{d}x = 1$$

♦

En el primer ejemplo hemos considerado una función no acotada y en el segundo un intervalo no acotado.

8.13 Definición. Sea $f : [c,b[\to \mathbb{R}$ una función continua en el intervalo $[c,b[$, donde suponemos que $c\in\mathbb{R}$ y que b un número real mayor que c o bien $b = +\infty$. Se define la integral impropia de Riemann de f en $[c,b[$ como el límite

$$\int_c^b f(x)\,\mathrm{d}x = \lim_{t\to b}\int_c^t f(x)\,\mathrm{d}x \tag{8.1}$$

Supuesto, claro está, que dicho límite exista y sea un número real, en cuyo caso se dice también que la integral de f es convergente en $[c,b[$.

Sea $f :]a,c] \to \mathbb{R}$ una función continua en el intervalo $]a,c]$, donde suponemos que $c\in\mathbb{R}$ y que a un número real menor que c o bien $a = -\infty$. Se define la integral impropia de Riemann de f en $]a,c]$ como el límite

$$\int_a^c f(x)\,\mathrm{d}x = \lim_{t\to a}\int_t^c f(x)\,\mathrm{d}x \tag{8.2}$$

Supuesto, claro está, que dicho límite exista y sea un número real, en cuyo caso se dice también que la integral de f es convergente en $]a,c]$.

Cuando el límite (8.1) o (8.2) existe y es igual a $+\infty$ (resp. $-\infty$) se dice que la respectiva integral es positivamente o negativamente divergente.

Sea $f :]a,b[\to \mathbb{R}$ una función continua en el intervalo $]a,b[$, donde $-\infty \leqslant a < b \leqslant +\infty$. Sea $c\in\mathbb{R}$ con $a < c < b$. Se dice que la integral de f es convergente en $]a,b[$ cuando las integrales de f en $]a,c]$ y en $[c,b[$ son convergentes, en cuyo caso se define

$$\int_a^b f(x)\,\mathrm{d}x = \int_a^c f(x)\,\mathrm{d}x + \int_c^b f(x)\,\mathrm{d}x \tag{8.3}$$

8.14 Ejemplo. Sea $a \neq 1$. Se tiene que

$$\int_1^t \frac{1}{x^a}\,\mathrm{d}x = \frac{t^{1-a}}{1-a} - \frac{1}{1-a}$$

Deducimos que

$$\int_1^{+\infty} \frac{1}{x^a}\,\mathrm{d}x = \begin{cases} \dfrac{1}{a-1} & \text{si } a > 1 \\ +\infty & \text{si } a < 1 \end{cases}$$

Análogamente

$$\int_0^1 \frac{1}{x^a}\,\mathrm{d}x = \begin{cases} \dfrac{1}{1-a} & \text{si } a < 1 \\ +\infty & \text{si } a > 1 \end{cases}$$

♦

8.15 Ejemplo. Sea $a \neq 1$. Usando la técnica de integración por partes, que estudiaremos más adelante, es fácil calcular una primitiva de la función $f(x) = \dfrac{\log x}{x^a}$. Comprueba que

$$F(x) = \frac{x^{1-a}(-1+(1-a)\log x)}{(1-a)^2}$$

es una primitiva de f en $\mathbb{R}^+$. Por tanto $\int\limits_1^t f(x)\,\mathrm{d}x = F(t) - F(1)$. En consecuencia

$$\int\limits_1^{+\infty} \frac{\log x}{x^a}\,\mathrm{d}x = \begin{cases} \dfrac{1}{(1-a)^2} & \text{si } a > 1 \\ +\infty & \text{si } a < 1 \end{cases}$$

Análogamente

$$\int\limits_0^{1} \frac{\log x}{x^a}\,\mathrm{d}x = \begin{cases} -\dfrac{1}{(1-a)^2} & \text{si } a < 1 \\ -\infty & \text{si } a > 1 \end{cases}$$

♦

8.2.1. Criterios de convergencia para integrales

Naturalmente, no siempre vamos a disponer de una primitiva expresable por medio de funciones elementales, bien porque no exista o porque su cálculo efectivo sea muy complicado. Por ello, interesa conocer condiciones que aseguren la convergencia de una integral sin necesidad de conocer una primitiva elemental. Lógicamente, estas condiciones no nos permitirán calcular el valor numérico de la integral; tan sólo nos dirán si es o no convergente. El caso en que la función integrando es positiva es particularmente sencillo de estudiar.

8.16 Proposición (Criterio básico de convergencia). *Sea f continua y positiva en $[c,b[$. Entonces, la integral de f en $[c,b[$ es convergente si, y sólo si, la función $F(x) = \int\limits_c^x f(t)\,\mathrm{d}t$ está mayorada en $[c,b[$, en cuyo caso*

$$\int\limits_c^b f(t)\,\mathrm{d}t = \sup\left\{ \int\limits_c^x f(t)\,\mathrm{d}t \, : x \in [c,b[\right\}$$

En otro caso la integral de f en $[c,b[$ es positivamente divergente.

Las afirmaciones hechas son consecuencia de que, por ser f positiva en $[c,b[$, la función $F(x) = \int\limits_c^x f(t)\,\mathrm{d}t$ es creciente en $[c,b[$.

El siguiente criterio es consecuencia inmediata del anterior.

8.17 Proposición (Criterio de comparación). *Sean f y g continuas y positivas en $[c,b[$. Supongamos que la integral de g en $[c,b[$ es convergente y que $f(x) \leqslant g(x)$ para todo $x \in [c,b[$. Entonces la integral de f en $[c,b[$ también es convergente.*

De este criterio se deduce fácilmente el siguiente.

8.18 Proposición (Criterio límite de comparación). *Sean f y g continuas y positivas en $[c,b[$. Supongamos que*

$$\lim_{x \to b} \frac{f(x)}{g(x)} = \rho \in \mathbb{R}^+$$

Entonces las integrales de f y g en $[c,b[$ ambas convergen o ambas divergen positivamente.

8.19 Definición. Se dice que la integral de f es **absolutamente convergente** en un cierto intervalo cuando la integral de la función $|f|$ es convergente en dicho intervalo.

Naturalmente, los criterios de convergencia antes vistos para integrales de funciones positivas, pueden usarse para estudiar la convergencia absoluta de la integral de cualquier función. Por ello, el siguiente resultado, que no demostraremos, es de gran utilidad.

8.20 Teorema. *Si la integral de f es absolutamente convergente, entonces la integral de f también es convergente.*

Ejercicios

1. Estudia la convergencia de la integral

$$I = \int_0^{+\infty} x^{\alpha} \frac{x+\operatorname{sen} x}{x-\operatorname{sen} x}\,\mathrm{d}x$$

   Según los valores de $\alpha \in \mathbb{R}$.

8.3. Técnicas de cálculo de Primitivas

Introducción

Para calcular $\int_a^b f(x)dx$ donde f es una función continua, hay que calcular una primitiva de f, evaluarla en a y en b y hacer la diferencia. Pero, ¿para qué calcular una primitiva? ¿no sabemos ya que una primitiva de f es la función $F(x) = \int_a^x f(t)dt$? Y, naturalmente, cualquier otra será de la forma $F(x)+C$ donde C es una constante. ¿Qué interés tiene entonces el cálculo de primitivas de funciones continuas? Respuesta: desde un punto de vista teórico ninguno. Ahora, si lo que queremos es aplicar la regla de Barrow para calcular el número $\int_a^b f(x)dx$, entonces la primitiva $F(x) = \int_a^x f(t)dt$ no nos sirve para nada porque si la evaluamos en a y en b y hacemos la diferencia obtenemos una identidad perfectamente inútil para nuestros propósitos. Lo que necesitamos es conocer una primitiva de f que sea realmente evaluable, es decir que al evaluarla en a y en b proporcione valores numéricos.

En otros términos, **el problema del cálculo de primitivas consiste en tratar de expresar la *"primitiva trivial"* $F(x) = \int_a^x f(t)dt$ por medio de funciones elementales**[2] **que permitan una evaluación efectiva de la integral**. Para eso sirven las técnicas de cálculo de primitivas.

Pero no hay que olvidar que, si bien la derivada de una función elemental también es una función elemental, es frecuente que una función elemental no tenga primitivas que puedan expresarse por medio de funciones elementales. Esto ocurre, por ejemplo, con las funciones e^{-x^2}, $\dfrac{\operatorname{sen} x}{x}$, $\operatorname{sen}(x^2)$, $\sqrt{x^3+1}$, y muchas más. En tales casos la forma más sencilla de representar una primitiva de f es justamente mediante la función $F(x) = \int_a^x f(t)dt$ y, para obtener valores concretos de dicha función hay que recurrir a métodos numéricos de cálculo de integrales.

[2] Las funciones que se obtienen por medio de sumas, productos, cocientes y composiciones a partir de las funciones racionales, exponenciales, logarítmicas, trigonométricas y sus inversas, se llaman funciones elementales.

En lo que sigue vamos a considerar algunos tipos de funciones elementales cuyas primitivas también pueden expresarse por medio de funciones elementales y pueden calcularse con procedimientos más o menos sistemáticos.

Para leer lo que sigue necesitas tener papel y un bolígrafo a mano para ir haciendo los ejercicios que se proponen. A calcular primitivas se aprende practicando; la imprescindible agilidad en los cálculos la lograrás haciendo decenas de ejercicios. Fíjate que, en la mayoría de los casos, se trata de ejercicios en los que tan sólo tienes que aplicar una técnica general a un caso particular. Esto es tan *"fácil"* que lo saben hacer los programas de cálculo simbólico, como *Mathematica*, *Derive*, *Mapple* y otros. Cuando se logre fabricar una calculadora de bolsillo que pueda ejecutar estos programas quizás ya no sea imprescindible aprender a calcular primitivas, pero hasta que llegue ese momento sigue siendo necesario que aprendas a calcular primitivas con agilidad. Sería lamentable que, por no saber calcular una primitiva, no puedas resolver una sencilla ecuación diferencial, ni calcular una probabilidad, ni el área de una superficie,... Las aplicaciones del cálculo integral son tan variadas, que el tiempo que dediques a la práctica del cálculo de primitivas será más rentable de lo que ahora puedas imaginar.

Observaciones sobre la notación y terminología usuales

Para representar una primitiva de una función f, suele usarse la notación $\int f(x)dx$. Así, por ejemplo, se escribe $\displaystyle\int \frac{1}{x-a}\,dx = \log|x-a|$. Esta notación es algo imprecisa porque no especifica el intervalo en que se considera definida f. En el ejemplo anterior hay que interpretar que la función $\dfrac{1}{x-a}$ está definida en uno de los intervalos $]-\infty, a[$ o $]a,+\infty[$ y elegir la primitiva correspondiente. Estos pequeños inconvenientes están compensados por la comodidad en los cálculos que proporciona esta notación. Es frecuente también, aunque no lo haremos en lo que sigue (pero mira el ejercicio 3), añadir una constante arbitraria, C, y escribir $\displaystyle\int \frac{1}{x-a}\,dx = \log|x-a|+C$.

La integral de una función en un intervalo, $\int_a^b f(x)dx$, se llama a veces *"integral definida"* de f (y es un número), y al símbolo $\int f(x)dx$ se le llama *"integral indefinida"* o, simplemente, *"integral"* de f (y representa una primitiva cualquiera de f). Aunque esto puede ser confuso, no olvides que, *cuando hablamos de calcular la integral* $\int f(x)dx$ *lo que realmente queremos decir es que queremos calcular una primitiva de* f.

Como ya sabes, en los símbolos $\int f(x)dx$ o $\int_a^b f(x)dx$ la letra "x" puede sustituirse por cualquier otra y el símbolo "dx" (que se lee *"diferencial x"*) sirve para indicar la variable de integración. Esto es muy útil si la función f contiene parámetros. Por ejemplo, son muy diferentes las integrales $\int x^y dx$ y $\int x^y dy$.

Te recuerdo también que, si $y = y(x)$ es una función de x, suele usarse la notación $dy = y'dx$ que es útil para mecanizar algunos cálculos pero que no tiene ningún significado especial: es una forma de indicar que y' es la derivada de y respecto a x.

Finalmente, si φ es una función, se usa la notación $\varphi(x)\Big|_{x=c}^{x=d}$ o sencillamente, $\varphi(x)\Big|_c^d$ para indicar el número $\varphi(d)-\varphi(c)$, y usaremos la notación $\varphi(x)\Big|_{x\to a}^{x\to b}$ para indicar $\lim_{x\to b}\varphi(x) - \lim_{x\to a}\varphi(x)$. Esta notación es cómoda cuando estudiamos convergencia de integrales.

8.3.1. Integración por partes

Si u y v son funciones con derivada primera continua en un intervalo, por la regla de derivación para un producto sabemos que: $(u(x)\,v(x))' = u'(x)v(x) + u(x)v'(x)$. Deducimos que la función producto $u\,v$ es una primitiva de la función $u'v + v'u$, es decir, $\int (u'(x)v(x) + u(x)v'(x))dx = u(x)\,v(x)$. Lo que suele escribirse en la forma:

$$\int u\,dv = u\,v - \int v\,du$$

Por supuesto, esta igualdad podemos usarla para calcular integrales definidas:

$$\int_c^d u(x)v'(x)dx = u(x)v(x)\Big|_{x=c}^{x=d} - \int_c^d v(x)u'(x)dx \tag{8.4}$$

Finalmente, si u y v están definidas en un intervalo abierto de extremos $-\infty \leqslant a < b \leqslant +\infty$ *y existen los límites* $\lim\limits_{x\to a} u(x)v(x)$ y $\lim\limits_{x\to b} u(x)v(x)$, entonces la igualdad (8.4) nos dice que las integrales $\int_a^b v(x)u'(x)dx$ y $\int_a^b u(x)v'(x)dx$ ambas convergen o ninguna converge y, cuando son convergentes se verifica que:

$$\int_a^b u(x)v'(x)dx = u(x)v(x)\Big|_{x\to a}^{x\to b} - \int_a^b v(x)u'(x)dx \tag{8.5}$$

Naturalmente, si queremos usar este método para calcular una integral $\int f(x)dx$ lo primero que hay que hacer es expresar $f(x) = u(x)w(x)$ de forma que el cálculo de $v(x)$ por la condición, $v'(x) = w(x)$, es decir la integral $v(x) = \int w(x)dx$, sea inmediata. Tenemos entonces

$$\int f(x)dx = \int u(x)w(x)dx = \int u(x)v'(x)dx = u(x)v(x) - \int v(x)u'(x)dx \tag{8.6}$$

Veamos algunas situaciones en las que este método puede aplicarse con éxito.

- Cuando la integral $\int v(x)u'(x)dx$ es inmediata. Por ejemplo, para calcular una integral $\int f(x)dx$ en la que la derivada de $f(x)$ es más sencilla que la propia función, como es el caso de $\log x$, $\operatorname{arc\,sen} x$, $\operatorname{arc\,tg} x$. Entonces conviene tomar $u(x) = f(x)$ y $v'(x) = w(x) = 1$ en (8.6).

8.21 Ejemplo.

$$\int \operatorname{arc\,tg} x\,dx = \begin{bmatrix} u = \operatorname{arc\,tg} x \to d\,u = \dfrac{1}{1+x^2}\,dx \\ d\,v = dx \to v = x \end{bmatrix} = x\operatorname{arc\,tg} x - \int \frac{x}{1+x^2}\,dx = x\operatorname{arc\,tg} x + \frac{1}{2}\log(1+x^2)$$

♦

- Cuando la integral $\int v(x)u'(x)dx$ es del mismo tipo que la integral de partida, pero más sencilla, de manera que reiterando el proceso se llega a una integral inmediata. Este es el caso cuando $f(x)$ es de la forma $P(x)\,\mathrm{e}^{ax}$, $P(x)\operatorname{sen}(ax)$, $P(x)\cos(ax)$, donde $P(x)$ es una función polinómica. En todos los casos se elige $u(x) = P(x)$, y $v'(x) = \mathrm{e}^{ax}$, $v'(x) = \operatorname{sen}(ax)$, $v'(x) = \cos(ax)$.

8.22 Ejemplo.

$$\int P(x)\,\mathrm{e}^{ax}\,dx = \begin{bmatrix} u = P(x) \to du = P'(x)dx \\ d\,v = \mathrm{e}^{ax}\,dx \to v = \dfrac{\mathrm{e}^{ax}}{a} \end{bmatrix} = P(x)\frac{\mathrm{e}^{ax}}{a} - \frac{1}{a}\int P'(x)e^{ax}dx$$

La última integral es *del mismo tipo que la primera pero con el grado del polinomio rebajado en una unidad*. El proceso se repite tantas veces como sea necesario. ♦

- Cuando la integral $\int v(x)u'(x)dx$ es parecida a la de partida, de forma que al volver a aplicar el proceso la integral de partida se repite y es posible despejarla de la igualdad obtenida.

8.23 Ejemplo.

$$\int \cos(\log x)dx = \left[\begin{array}{l} u = \cos(\log x) \to du = -\dfrac{1}{x}\operatorname{sen}(\log x)dx \\ dv = dx \to v = x \end{array}\right] = x\cos(\log x) + \int \operatorname{sen}(\log x)dx =$$

$$= \left[\begin{array}{l} u = \operatorname{sen}(\log x) \to du = \dfrac{1}{x}\cos(\log x)dx \\ dv = dx \to v = x \end{array}\right] = x\cos(\log x) + x\operatorname{sen}(\log x) - \int \cos(\log x)dx$$

deducimos que $\displaystyle\int \cos(\log x)dx = \frac{x}{2}\big(\cos(\log x) + \operatorname{sen}(\log x)\big)$. ♦

Ejercicios

1. Calcular las integrales:

$$\int_1^2 \log x\,dx\,,\quad \int s^2\,\mathrm{e}^{2s}\,ds\,,\quad \int \operatorname{arc\,sen} x\,dx\,,\,,\quad \int_1^4 \sqrt{t}\,\log t\,dt\,,\quad \int_1^e (\log x)^2 dx$$

$$\int x^3\,\mathrm{e}^{x^2}\,dx\,,\quad \int \log(x^2+1)dx\,,\quad \int_0^{\pi/4} \frac{\vartheta}{\cos^2\vartheta}\,d\vartheta\,,\quad \int x^2\operatorname{sen} x\,dx\,,\quad \int_1^e \cos^2(\log x)dx$$

2. Calcular las integrales $\displaystyle\int \mathrm{e}^{ax}\cos(bx)\,dx$, y $\displaystyle\int \mathrm{e}^{ax}\operatorname{sen}(bx)\,dx$. Y deducir, para $a > 0$ el valor de $\displaystyle\int_0^{+\infty} \mathrm{e}^{-ax}\cos(bx)\,dx$ y $\displaystyle\int_0^{+\infty} \mathrm{e}^{-ax}\operatorname{sen}(bx)\,dx$.

3. Explica la aparente contradicción

$$\int \frac{1}{\operatorname{sen} x\cos x}\,dx = \int \frac{\operatorname{cotg} x}{\cos^2 x}\,dx = \int \operatorname{cotg} x\,\operatorname{tg}' x\,dx = \operatorname{cotg} x\,\operatorname{tg} x - \int \operatorname{tg} x\,\operatorname{cotg}' x\,dx = 1 + \int \frac{\operatorname{tg} x}{\operatorname{sen}^2 x}\,dx$$

$$= 1 + \int \frac{1}{\operatorname{sen} x\cos x}\,dx. \quad \text{Luego } 1 = 0.$$

Ahora que estás empezando a hacer ejercicios de cálculo de primitivas es una buena práctica que compruebes los resultados. Además es muy sencillo: basta derivar la primitiva que has obtenido.

Integración por recurrencia

La técnica de integración por partes permite en algunas ocasiones relacionar una integral de la forma $I_n = \int f(x,n)dx$ en la que interviene un parámetro n (con frecuencia un número natural) con otra del mismo tipo en la que el parámetro ha disminuido en una o en dos unidades. Las expresiones así obtenidas se llaman fórmulas de reducción o de recurrencia y permiten el cálculo efectivo de la integral cuando se particularizan valores del parámetro. Los siguientes ejemplos son ilustrativos de esta forma de proceder.

8.24 Ejemplo.

$$\int (\log x)^n dx = \left[\begin{array}{l} u = (\log x)^n \to du = n\dfrac{(\log x)^{n-1}}{x}\,dx \\ dv = dx \to v = x \end{array}\right] = x(\log x)^n - n\int (\log x)^{n-1}dx$$

♦

8.25 Ejemplo.

$$I_n = \int x^n \,\mathrm{e}^{ax}\, dx = \left[\begin{array}{c} u = x^n \to du = n x^{n-1} \\ dv = \mathrm{e}^{ax}\, dx \to v = \dfrac{\mathrm{e}^{ax}}{a}\, dx \end{array}\right] = \frac{1}{a}\left(x^n \,\mathrm{e}^{ax} - n I_{n-1}\right)$$

♦

8.26 Ejemplo.

$$I_n = \int \operatorname{sen}^n x\, dx = \left[\begin{array}{l} u = \operatorname{sen}^{n-1} x \to du = (n-1)\operatorname{sen}^{n-2} x \cos x\, dx \\ dv = \operatorname{sen} x\, dx \to v = -\cos x \end{array}\right] =$$

$$= -\cos x \operatorname{sen}^{n-1} x + (n-1)\int \operatorname{sen}^{n-2} x \cos^2 x\, dx = -\cos x \operatorname{sen}^{n-1} x + (n-1)\int \operatorname{sen}^{n-2} x\, dx - (n-1) I_n$$

Y deducimos fácilmente que $\displaystyle\int \operatorname{sen}^n x\, dx = -\frac{1}{n}\cos x \operatorname{sen}^{n-1} x + \frac{n-1}{n}\int \operatorname{sen}^{n-2} x\, dx$. En particular, $\displaystyle\int_0^{\pi/2} \operatorname{sen}^n x\, dx = \frac{n-1}{n}\int_0^{\pi/2} \operatorname{sen}^{n-2} x\, dx$. De aquí se obtienen las igualdades

$$\int_0^{\pi/2} \operatorname{sen}^{2n+1} x\, dx = \frac{2\cdot 4\cdot 6\cdots 2n}{3\cdot 5\cdot 7\cdots (2n+1)}, \qquad \int_0^{\pi/2} \operatorname{sen}^{2n} x\, dx = \frac{1\cdot 3\cdot 5\cdots (2n-1)}{2\cdot 4\cdot 6\cdots 2n}\,\frac{\pi}{2}$$

♦

Ejercicios

4. Calcula, haciendo uso de los resultados anteriores, las integrales

$$\int (\log x)^3 dx, \quad \int x^4 \,\mathrm{e}^x\, dx, \quad \int_0^{\pi/2} \operatorname{sen}^4 x\, dx, \quad \int \operatorname{sen}^5 x\, dx$$

5. Prueba las siguientes relaciones de recurrencia

(a) $\displaystyle I_n = \int \cos^n x\, dx = \frac{1}{n}\big(\cos^{n-1} x \operatorname{sen} x + (n-1) I_{n-2}\big)$ (b) $\displaystyle I_n = \int \operatorname{tg}^n x\, dx = \frac{1}{n-1}\operatorname{tg}^{n-1} x - I_{n-2}$

6. Prueba la igualdad $\displaystyle I_n = \int \frac{1}{(1+x^2)^n}\, dx = \frac{x}{(2n-2)(1+x^2)^{n-1}} + \frac{2n-3}{2n-2} I_{n-1}$

Sugerencias: $\displaystyle I_n = \int \frac{(1+x^2)-x^2}{(1+x^2)^n}\, dx = I_{n-1} - \int \frac{x^2}{(1+x^2)^n}\, dx$. Ahora:

$$\int \frac{x^2}{(1+x^2)^n}\, dx = \left[\begin{array}{lcl} u = x & \to & du = dx \\ dv = \dfrac{x}{(1+x^2)^n}\, dx & \to & v = \dfrac{1}{2(n-1)}\,\dfrac{1}{(1+x^2)^{n-1}} \end{array}\right] = \cdots$$

7. (*) Estudia la convergencia de la integral

$$I_n = \int_0^{+\infty} \frac{x^{2n-1}}{(1+x^2)n+3}\, \mathrm{d}x \qquad (n \geqslant 1)$$

Prueba que para $n \geqslant 2$ es $\displaystyle I_n = \frac{n-1}{n+2} I_{n-1}$. Calcula I_1, I_2 e I_3.

8.3.2. Integración por sustitución o cambio de variable

Sean $g\colon J\to\mathbb{R}$ una función con derivada primera continua en un intervalo J que toma valores en un intervalo I, y f una función continua en I. Sea F una primitiva de f en I, y pongamos $H=F\circ g$. Tenemos, por la regla de la cadena, que $H'(t)=F'(g(t))g'(t)=f(g(t))g'(t)$, es decir, la función H es una primitiva en J de la función $h(t)=f(g(t))g'(t)$. Si c, d son puntos de J, deducimos que

$$\int_c^d f(g(t))g'(t)\,dt=H(d)-H(c)=F(g(d))-F(g(c))=\int_{g(c)}^{g(d)} f(x)\,dx$$

Esta igualdad se conoce con el nombre de *"fórmula de integración por sustitución o cambio de variable"*. En ella se supone que queremos calcular, por ejemplo, la integral $\int_a^b f(x)\,dx$ y lo que hacemos es la sustitución $x=g(t)$, con lo que $dx=g'(t)dt$ y se eligen c y d por la condición de que $g(c)=a$, $g(d)=b$. Naturalmente, esto tiene interés cuando la función $f(g(t))g'(t)$ es más fácil de integrar que la función f. Simbólicamente este proceso suele representarse en la forma

$$\int_a^b f(x)\,dx=\begin{bmatrix} x=g(t),\ dx=g'(t)dt\\ a=g(c),\ b=g(d)\end{bmatrix}=\int_c^d f(g(t))g'(t)\,dt$$

Para el caso de integrales indefinidas este proceso de sustitución de representa de forma menos precisa y se escribe simplemente

$$\int f(x)\,dx=\begin{bmatrix} x=g(t)\\ dx=g'(t)dt\end{bmatrix}=\int f(g(t))g'(t)\,dt$$

En este contexto, es frecuente calcular $\int f(g(t))g'(t)\,dt=H(t)$, y escribir $\int f(x)\,dx=H(t)$, igualdad que no tiene mucho sentido si no se especifica también la relación entre las variables t y x, escribiendo "$\int f(x)\,dx=H(t)$ *donde* $x=g(t)$". Desde luego, el conocimiento de $H(t)$ y de la relación $x=g(t)$ es suficiente para calcular integrales definidas de f, pero también podemos *"deshacer el cambio"* para obtener una primitiva de f. Para eso la función g debe ser una biyección de J sobre I con derivada no nula. En tal caso, la función $F(x)=H(g^{-1}(x))$ es una primitiva de f en I. En efecto:

$$F'(x)=H'(g^{-1}(x))(g^{-1})'(x)=f(g(g^{-1}(x)))g'(g^{-1}(x))(g^{-1})'(x)=f(x)g'(g^{-1}(x))\frac{1}{g'(g^{-1}(x))}=f(x)$$

No olvides que la fórmula del cambio de variables puede usarse en un sentido (de izquierda a derecha) o en otro (de derecha a izquierda) según convenga.

Puede ocurrir que al hacer un cambio de variable en una *"integral corriente"* obtengamos una *"integral impropia"*. No hay que preocuparse porque **para estudiar la convergencia de una integral pueden hacerse cambios de variable *biyectivos*: ello no altera la eventual convergencia de la integral ni su valor**.

8.27 Ejemplo. Con frecuencia se hacen cambios de variable para quitar radicales.

$$\int_{2/\sqrt{3}}^{2}\frac{1}{x^2\sqrt{x^2+4}}\,dx=\begin{bmatrix} x=2\operatorname{tg}t,\ dx=\dfrac{2}{\cos^2 t}\\ 2/\sqrt{3}=2\operatorname{tg}(\pi/6),\ 2=2\operatorname{tg}(\pi/4)\end{bmatrix}=\frac14\int_{\pi/6}^{\pi/4}\frac{\cos t}{\operatorname{sen}^2 t}\,dt=\frac14\left[\frac{-1}{\operatorname{sen}t}\right]_{\pi/6}^{\pi/4}=\frac{2-\sqrt{2}}{4}$$

♦

8.28 Ejemplo. Un cambio de variable en una integral impropia. Consideremos la integral:

$$\int_a^b \frac{1}{\sqrt{(x-a)(b-x)}}\,dx$$

Suponemos que $a<b$. El cambio que hacemos consiste en llevar el intervalo $]-1,1[$ al $]a,b[$ por una biyección del tipo $g(t)=\alpha t+\beta$. Las condiciones $g(-1)=a$, $g(1)=b$ nos dan que $\alpha=(b-a)/2$, $\beta=(b+a)/2$. Con ello:

$$\int_a^b \frac{1}{\sqrt{(x-a)(b-x)}}\,dx = \begin{bmatrix} x=g(t),\ dx=\dfrac{b-a}{2} \\ a=g(-1),\ b=g(1) \end{bmatrix} = \int_{-1}^{1}\frac{dt}{\sqrt{1-t^2}}=\pi$$

♦

Ejercicios

7. Calcular las siguientes integrales utilizando el cambio de variable indicado

$$\int_0^{\pi/4}\frac{\operatorname{sen}^3 x}{\cos^4 x}\,dx \quad x=\operatorname{arc}\cos t\,;\quad \int_{-\pi/4}^{\pi/4}\frac{\operatorname{sen}^2 x}{\cos^4 x}\,dx \quad x=\operatorname{arc}\operatorname{tg} t\,;\quad \int_1^{+\infty}\frac{dx}{\mathrm{e}^x+1}\quad x=\log t$$

8. Calcular las integrales

$$\int\sqrt{4-x^2}\,dx\,,\quad \int\frac{dx}{x\sqrt{x^2-1}}\,,\quad \int_{\mathrm{e}}^{\mathrm{e}^4}\frac{dx}{x\sqrt{\log x}}\,,\quad \int_1^4\frac{1}{x^2}\sqrt{1+\frac{1}{x}}\,dx\,,\quad \int\frac{\mathrm{e}^x+3\,\mathrm{e}^{2x}}{2+\mathrm{e}^x}\,dx$$

9. Sea $a>0$. Prueba que si f es impar, es decir, $f(-x)=-f(x)$, entonces $\int_{-a}^{a}f(t)dt=0$. Y si f es una función par, es decir, $f(-x)=f(x)$, entonces $\int_{-a}^{a}f(t)dt=2\int_0^a f(t)dt$.

8.3.3. Integración de funciones racionales

Dadas dos funciones polinómicas $P(x)$ y $Q(x)$, queremos calcular $\int\frac{P(x)}{Q(x)}dx$. Si el grado de P es mayor o igual que el de Q, podemos dividir los dos polinomios obteniendo

$$\frac{P(x)}{Q(x)}=H(x)+\frac{G(x)}{Q(x)},$$

donde $H(x)$ y $G(x)$ son polinomios y el grado de G es menor que el grado de Q. Por tanto, *supondremos siempre que el grado de* P *es menor que el grado de* Q. *Supondremos también que el coeficiente líder del polinomio* Q *es* 1. La técnica para calcular la integral consiste en descomponer la fracción $\frac{P(x)}{Q(x)}$ en otras más sencillas llamadas *"fracciones simples"*. Estudiaremos dos formas de hacerlo: el método de los coeficientes indeterminados y una variante del mismo conocida como Método de Hermite.

Paso 1. Descomposición del denominador en factores irreducibles

Descomponemos el denominador, $Q(x)$, como producto de factores de grado 1 y factores de grado 2 irreducibles:

$$Q(x) = (x-a_1)^{\alpha_1} \cdots (x-a_n)^{\alpha_n} (x^2+b_1x+c_1)^{\beta_1} \cdots (x^2+b_mx+c_m)^{\beta_m} \tag{8.7}$$

Observaciones

• Esto se dice muy pronto, pero puede ser muy difícil de hacer si no imposible. Afortunadamente, en los casos prácticos esta descomposición o se conoce o es muy fácil de realizar.

• En la descomposición (8.7) cada a_j es una raíz real de orden α_j del polinomio Q, y los factores cuadráticos del tipo $(x^2+b_jx+c_j)^{\beta_j}$ corresponden a raíces complejas conjugadas de orden β_j. Tales factores cuadráticos son irreducibles, es decir, su discriminante es negativo o, lo que es igual, $x^2+b_jx+c_j > 0$ para todo $x \in \mathbb{R}$.

Paso 2

Método de los coeficientes indeterminados

Escribimos el cociente $\dfrac{P(x)}{Q(x)}$ como suma de fracciones de la siguiente forma:

• Por cada raíz real a_j de orden α_j escribimos α_j fracciones cuyos numeradores son constantes A_{k_j} que hay que determinar, y los denominadores son de la forma $(x-a_j)^{k_j}$ donde k_j toma valores de 1 hasta α_j.

• Por cada factor cuadrático irreducible $(x^2+b_jx+c_j)^{\beta_j}$ escribimos β_j fracciones cuyos numeradores son de la forma $B_{k_j}x+C_{k_j}$ siendo B_{k_j} y C_{k_j} constantes que hay que determinar, y los denominadores son de la forma $(x^2+b_jx+c_j)^{k_j}$ donde k_j toma valores de 1 hasta β_j.

• La descomposición es de la forma:

$$\frac{P(x)}{Q(x)} = \sum_{j=1}^{n} \left[\sum_{k_j=1}^{\alpha_j} \frac{A_{k_j}}{(x-a_j)^{k_j}} \right] + \sum_{j=1}^{m} \left[\sum_{k_j=1}^{\beta_j} \frac{B_{k_j}x + C_{k_j}}{(x^2+b_jx+c_j)^{k_j}} \right] \tag{8.8}$$

Método de Hermite

Escribimos el cociente $\dfrac{P(x)}{Q(x)}$ de la siguiente forma:

$$\begin{aligned}\frac{P(x)}{Q(x)} &= \frac{A_1}{x-a_1} + \cdots + \frac{A_n}{x-a_n} + \frac{B_1x+C_1}{x^2+b_1x+c_1} + \cdots + \frac{B_mx+C_m}{x^2+b_mx+c_m} + \\ &+ \frac{d}{dx}\left(\frac{F(x)}{(x-a_1)^{\alpha_1-1}\cdots(x-a_n)^{\alpha_n-1}(x^2+b_1x+c_1)^{\beta_1-1}\cdots(x^2+b_mx+c_m)^{\beta_m-1}} \right)\end{aligned}$$

donde $A_1,\dots,A_n,B_1,\dots,B_m,C_1,\dots,C_m$ son coeficientes que tenemos que determinar y, en la fracción que aparece con una derivada, $F(x)$ es un polinomio genérico de grado uno menos que el denominador. En resumen, se trata de escribir $\dfrac{P(x)}{Q(x)}$ como suma de fracciones simples, una por cada factor, más la derivada de un cociente que tiene por denominador lo que queda de $Q(x)$. Observa que en ambos métodos hay que calcular tantos coeficientes como el grado de Q.

Paso 3. Determinación de los coeficientes

Tanto en un caso como en otro, se reducen todas las fracciones a común denominador (que será $Q(x)$), y se iguala a $P(x)$ al numerador resultante. Esto nos producirá un sistema de ecuaciones cuya resolución nos dará el valor de todos los coeficientes. Naturalmente, en el método de Hermite hay que efectuar la derivada antes de reducir a común denominador.

Observaciones

• En ambos métodos tenemos que calcular el mismo número de coeficientes pero en el método de Hermite la obtención del sistema de ecuaciones es un poco más trabajosa debido a la presencia de la derivada.

• El método de Hermite es interesante de aplicar cuando hay factores cuadráticos de orden elevado (raíces imaginarias múltiples).

Paso 4. Integración de las fracciones simples

En el método de Hermite, una vez escrita la función racional $\dfrac{P(x)}{Q(x)}$ de la forma anterior, es fácil calcular su integral:

$$\int \frac{P(x)}{Q(x)}\,dx = \int \frac{A_1}{x-a_1}\,dx + \cdots + \int \frac{B_1x+C_1}{x^2+b_1x+c_1}\,dx + \cdots +$$
$$+ \frac{F(x)}{(x-a_1)^{\alpha_1-1}\cdots(x-a_n)^{\alpha_n-1}(x^2+b_1x+c_1)^{\beta_1-1}\cdots(x^2+b_mx+c_m)^{\beta_m-1}}$$

Sólo nos queda saber calcular las integrales que hemos dejado pendientes:

• $\displaystyle\int \frac{A}{x-a}\,dx = A\log|x-a|.$

• $\displaystyle\int \frac{Bx+C}{x^2+bx+c}\,dx$. Siempre se puede escribir $x^2+bx+c=(x-d)^2+k^2$, con lo que descomponemos nuestra integral en dos:

$$\int \frac{Bx+C}{x^2+bx+c}\,dx = \int \frac{Bx+C}{(x-d)^2+k^2}\,dx = \int \frac{B(x-d)+C+Bd}{(x-d)^2+k^2}\,dx =$$
$$= \int \frac{B(x-d)}{(x-d)^2+k^2}\,dx + \int \frac{C+Bd}{(x-d)^2+k^2}\,dx =$$
$$= \frac{B}{2}\log\left((x-d)^2+k^2\right) + (C+Bd)\int \frac{dx}{(x-d)^2+k^2}$$

y la última integral es inmediata (del tipo arcotangente) si hacemos el cambio de variable $t=\dfrac{x-d}{k}$.

En el método de los coeficientes indeterminados aparecen también, cuando hay raíces múltiples, otros dos tipos de fracciones elementales:

• Fracciones del tipo $\dfrac{A}{(x-a)^k}$ donde $k\in\mathbb{N}$ y $k\geqslant 2$, correspondientes a raíces reales múltiples, las cuales no ofrecen dificultad pues

• $\displaystyle\int \frac{A}{(x-a)^k}\,dx = -\frac{A}{k-1}\frac{1}{(x-a)^{k-1}}.$

• Fracciones del tipo $\dfrac{Bx+C}{(x^2+bx+c)^k}$ donde $k\in\mathbb{N}$ y $k\geqslant 2$, correspondientes a raíces imaginarias múltiples, la integración de las cuales ofrece bastante dificultad a partir de $k\geqslant 3$. Suelen hacerse usando la fórmula de reducción del ejercicio número 6.

8.29 Ejemplo. Se trata de calcular $\displaystyle\int \frac{x^2-2}{x^3(x^2+1)^2}\,dx$. Como hay raíces imaginarias múltiples aplicaremos el método de Hermite.

$$\frac{x^2-2}{x^3(x^2+1)^2} = \frac{A}{x} + \frac{Bx+C}{x^2+1} + \frac{d}{dx}\left(\frac{ax^3+bx^2+cx+d}{x^2(x^2+1)}\right)$$

Realizando la derivada y reduciendo a común denominador, obtenemos un sistema de ecuaciones cuya solución es

$$a = 0, \qquad b = 5/2, \qquad c = 0, \qquad d = 1, \qquad A = 5, \qquad B = -5, \qquad C = 0;$$

por lo tanto

$$\int \frac{x^2-2}{x^3(x^2+1)^2}\,dx = \frac{(5/2)x^2+1}{x^2(x^2+1)} + 5\log x - \frac{5}{2}\log(x^2+1).$$

♦

8.30 Ejemplo. Calcular la integral $\displaystyle\int_2^{+\infty} \frac{x+1}{x(x+1)(x^2+1)}\,dx$. Aplicaremos el método de los coeficientes indeterminados.

$$\frac{x+1}{x(x+1)(x^2+1)} = \frac{A}{x} + \frac{B}{x+1} + \frac{Cx+D}{x^2+1}$$

Reduciendo a común denominador obtenemos:

$$\frac{x+1}{x(x+1)(x^2+1)} = \frac{-A+(A+B-D)x+(-A-C+D)x^2+(A+B+C)x^3}{x(x+1)(x^2+1)}$$

Identificando coeficientes resulta el sistema de ecuaciones lineales:

$$\left.\begin{aligned} A+B+C &= 0 \\ -A-C+D &= 0 \\ A+B-D &= 1 \\ -A &= 1 \end{aligned}\right\} \Rightarrow \left\{\begin{array}{ll} A=-1 & B=1 \\ C=0 & D=-1 \end{array}\right.$$

Deducimos que:

$$\int_2^t \frac{x+1}{x(x+1)(x^2+1)}\,dx = \int_2^t \frac{dx}{x-1} - \int_2^t \frac{dx}{x} - \int_2^t \frac{dx}{x^2+1} = \log\left(2\frac{t-1}{t}\right) - \operatorname{arctg} t + \operatorname{arctg} 2$$

Por tanto:

$$\int_2^{+\infty} \frac{x+1}{x(x+1)(x^2+1)}\,dx = \log 2 - \frac{\pi}{2} + \operatorname{arctg} 2$$

♦

Ejercicios

10. Calcular las siguientes integrales

$$\int \frac{2-x^2}{x^3-3x^2}\,dx, \quad \int_{-1/2}^{1/2} \frac{dx}{x^4-1}\,dx, \quad \int \frac{x^4+6x^3-7x^2-4x-3}{x^3-2x^2+x-2}\,dx$$

$$\int_1^{+\infty} \frac{x-1}{x^3-3x^2+x+5}\,dx, \quad \int_{-\infty}^{+\infty} \frac{dx}{(x^2-2x+2)^2}\,dx, \quad \int_0^1 \frac{dx}{1+x^4}\,dx$$

$$\int \frac{x^2}{(x^4-1)^2}\,dx, \quad \int \frac{dx}{x(1+x^4)}, \quad \int \frac{3x^2+30}{x^4+2x^2-8}\,dx, \quad \int \frac{x^2}{(x^2+1)^2}\,dx$$

8.3.4. Integración por racionalización

Acabamos de ver que la primitiva de una función racional siempre puede expresarse mediante funciones elementales. Nos vamos a ocupar ahora de algunos tipos de funciones no racionales cuyas integrales se pueden transformar, por medio de un cambio de variable, en integrales de funciones racionales. Se dice entonces que la integral de partida se ha *racionalizado* y esta técnica se conoce como *"integración por racionalización"*. Conviene advertir que los cambios de variable que siguen son los que la práctica ha confirmado como más útiles en general, pero que en muchas ocasiones la forma concreta de la función que queremos integrar sugiere un cambio de variable específico que puede ser más eficaz.

En lo que sigue, representaremos por $R=R(x,y)$ una función racional de dos variables, es decir, un cociente de funciones polinómicas de dos variables. Te recuerdo que una función polinómica de dos variables es una función de la forma $P(x,y)=\displaystyle\sum_{i=0}^{n}\sum_{j=0}^{m}c_{ij}x^iy^j$.

Integración de funciones del tipo $R(\operatorname{sen}x,\cos x)$

Las integrales del tipo $\int R(\operatorname{sen}x,\cos x)dx$ donde $R=R(x,y)$ una función racional de dos variables, se racionalizan con el cambio de variable $t=\operatorname{tg}(x/2)$. Con lo que

$$\operatorname{sen}x=\frac{2t}{1+t^2},\qquad \cos x=\frac{1-t^2}{1+t^2},\qquad dx=\frac{2\,dt}{1+t^2}$$

Con ello resulta:

$$\int R(\operatorname{sen}x,\cos x)dx=\Big[t=\operatorname{tg}(x/2)\Big]=\int R\left(\frac{2t}{1+t^2},\frac{1-t^2}{1+t^2}\right)\frac{2\,dt}{1+t^2}$$

8.31 Ejemplo.

$$\int\frac{dx}{\operatorname{sen}x-\operatorname{tg}x}=\int\frac{\cos x\,dx}{\operatorname{sen}x\cos x-\operatorname{sen}x}=\Big[\operatorname{tg}x/2=t\Big]=\cdots=\int\frac{t^2-1}{2t^3}\,dt$$
$$=\frac{1}{4t^2}+\frac{\log t}{2}=\frac{1}{4\operatorname{tg}^2(x/2)}+\frac{1}{2}\log|\operatorname{tg}(x/2)|.$$

♦

Casos particulares

• Cuando $R(-\operatorname{sen}x,-\cos x)=R(\operatorname{sen}x,\cos x)$ se dice que *"R es par en seno y coseno"*. En este caso es preferible el cambio $\operatorname{tg}x=t$. Con lo que

$$\operatorname{sen}x=\frac{t}{\sqrt{1+t^2}},\qquad \cos x=\frac{1}{\sqrt{1+t^2}},\qquad dx=\frac{dt}{1+t^2}$$

En el caso particular de tratarse de una integral del tipo

$$\int \operatorname{sen}^n x\cos^m x\,dx$$

con n y m números enteros *pares*, es preferible simplificar la integral usando las identidades

$$\cos^2 x=\frac{1+\cos 2x}{2}\qquad \operatorname{sen}^2 x=\frac{1-\cos 2x}{2}.$$

• Cuando $R(-\operatorname{sen} x, \cos x) = -R(\operatorname{sen} x, \cos x)$ se dice que *"R es impar en seno"* y el cambio $\cos x = t$ suele ser eficaz.

• Cuando $R(\operatorname{sen} x, -\cos x) = -R(\operatorname{sen} x, \cos x)$ se dice que *"R es impar en coseno"* y el cambio $\operatorname{sen} x = t$ suele ser eficaz.

8.32 Ejemplo. Calcular $\mathbf{I} = \int \operatorname{sen}^2 x \cos^2 x \, dx$. Tenemos:

$$\begin{aligned}
\mathbf{I} &= \int (1-\cos^2 x)\cos^2 x \, dx = \int \cos^2 x \, dx - \int \cos^4 x \, dx = \int \frac{1+\cos 2x}{2} dx - \int \left(\frac{1+\cos 2x}{2}\right)^2 dx \\
&= \frac{x}{2} + \frac{\operatorname{sen} 2x}{4} - \frac{1}{4}\int (1 + 2\cos 2x + \cos^2 2x) dx = \frac{x+\operatorname{sen} 2x}{4} - \frac{x}{4} - \frac{1}{2}\int \cos 2x \, dx - \frac{1}{4}\frac{1+\cos 4x}{2} dx \\
&= \frac{x+\operatorname{sen} 2x}{4} - \frac{\operatorname{sen} 2x}{4} - \frac{x}{8} - \frac{\operatorname{sen} 4x}{32} = \frac{1}{8}\left(x - \frac{\operatorname{sen} 4x}{4}\right)
\end{aligned}$$

♦

8.33 Ejemplo.

$$\begin{aligned}
\int \frac{\cos^3 x}{\operatorname{sen}^2 x} dx &= \int \frac{(1-\operatorname{sen}^2 x)\cos x \, dx}{\operatorname{sen}^2 x} = \left[\begin{array}{c} t = \operatorname{sen} x \\ dt = \cos x \, dx \end{array}\right] = \int \frac{1-t^2}{t^2} dt \\
&= \frac{-1}{t} - t = \frac{-1}{\operatorname{sen} t} - \operatorname{sen} t.
\end{aligned}$$

♦

8.34 Ejemplo. Sea $\mathbf{I} = \int \frac{\operatorname{sen}^2 x \cos x}{\operatorname{sen} x + \cos x} dx$. Se trata de una función par en seno y en coseno. Haciendo $t = \operatorname{tg} x$, obtenemos:

$$\mathbf{I} = \int \frac{t^2}{(t+1)(t^2+1)^2} dt$$

Aplicando el método de Hermite escribimos:

$$\frac{t^2}{(t+1)(t^2+1)^2} = \frac{A}{t+1} + \frac{Bt+C}{t^2+1} + \frac{d}{dx}\left(\frac{\alpha t + \beta}{t^2+1}\right)$$

Haciendo la derivada y reduciendo a común denominador obtenemos:

$$\frac{t^2}{(t+1)(t^2+1)^2} = \frac{A+C+\beta+(B+C-2\alpha+\beta)t+(2A+B+C-2\alpha-\beta)t^2+(B+C-\beta)t^3+(A+B)t^4}{(t+1)(t^2+1)^2}$$

Identificando coeficientes resulta el sistema de ecuaciones lineales:

$$\left.\begin{array}{rl}
A+C+\beta & = 0 \\
B+C-2\alpha+\beta & = 0 \\
2A+B+C-2\alpha-\beta & = 1 \\
B+C-\beta & = 0 \\
A+B & = 0
\end{array}\right\} \Rightarrow \left\{\begin{array}{ll}
A = 1/4 & B = -1/4 \\
C = 0 & D = -1 \\
\alpha = -1/4 & \beta = -1/4
\end{array}\right.$$

Deducimos que:

$$\mathbf{I} = \frac{1}{4}\log|t+1| - \frac{1}{8}\log(t^2+1) - \frac{1}{4}\frac{1+t}{1+t^2} = \frac{1}{4}\log|\operatorname{sen} x + \cos x| - \frac{1}{4}\cos x(\operatorname{sen} x + \cos x)$$

♦

• Cuando la función $R(\operatorname{sen} x, \cos x)$ sea de la forma

$$\operatorname{sen}(ax+b)\operatorname{sen}(cx+d), \quad \operatorname{sen}(ax+b)\cos(cx+d), \quad \cos(ax+b)\cos(cx+d)$$

puede resolverse la integral usando las fórmulas:

$$\operatorname{sen}\alpha\cos\beta = \frac{\operatorname{sen}(\alpha+\beta)+\operatorname{sen}(\alpha-\beta)}{2}, \quad \operatorname{sen}\alpha\operatorname{sen}\beta = \frac{\cos(\alpha-\beta)-\cos(\alpha+\beta)}{2}$$

$$\cos\alpha\cos\beta = \frac{\cos(\alpha-\beta)+\operatorname{sen}(\alpha+\beta)}{2}$$

8.35 Ejemplo.

$$\int \operatorname{sen}(3x)\cos(2x)\,dx = \frac{1}{2}\int \operatorname{sen}(5x)\,dx + \frac{1}{2}\int \operatorname{sen} x\,dx = -\frac{1}{10}\cos(5x) - \frac{1}{2}\cos x$$

♦

• Integrales de la forma $\int \operatorname{tg}^n x\,dx$, $\int \operatorname{cotg}^n x\,dx$. Se reducen a una con grado inferior separando $\operatorname{tg}^2 x$ o $\operatorname{cotg}^2 x$ y sustituyéndola por $\sec^2 x - 1$ o $\operatorname{cosec}^2 x - 1$.

8.36 Ejemplo.

$$\begin{aligned}\int \operatorname{tg}^5 x\,dx &= \int \operatorname{tg}^3 x\operatorname{tg}^2 x\,dx = \int \operatorname{tg}^3 x(\sec^2 x-1)\,dx = \int \operatorname{tg}^3 x\sec^2 x\,dx - \int \operatorname{tg}^3 x\,dx\\ &= \frac{\operatorname{tg}^4 x}{4} - \int \operatorname{tg}^3 x\,dx = \frac{\operatorname{tg}^4 x}{4} - \int \operatorname{tg} x\operatorname{tg}^2 x\,dx = \frac{\operatorname{tg}^4 x}{4} - \int \operatorname{tg} x(\sec^2 x-1)\,dx\\ &= \frac{\operatorname{tg}^4 x}{4} - \int \operatorname{tg} x\sec^2 x\,dx + \int \operatorname{tg} x\,dx = \frac{\operatorname{tg}^4 x}{4} - \frac{1}{2}\operatorname{tg}^2 x + \log|\cos x|\end{aligned}$$

♦

Ejercicios

11. Calcular las integrales

$$\int \frac{1}{a+b\cos x}\,dx, \quad \int_0^{\pi} \frac{1}{\cos x + 2\operatorname{sen} x + 3}\,dx, \quad \int \frac{1-2\cos x}{5-4\cos x}\,dx, \quad \int \frac{dx}{\cos x}, \quad \int \frac{1}{\operatorname{sen} x\cos x}\,dx$$

12. Calcular las integrales

$$\int \frac{dx}{\operatorname{sen}^2 x\cos^2 x}, \quad \int_0^{\pi/4} \frac{\cos(3x+4)}{\sqrt{1+\operatorname{tg}^2(x+2)}}\,dx, \quad \int \frac{dx}{(1+\operatorname{sen} x)\cos x}, \quad \int \operatorname{sen}^2 x\cos^3 x\,dx, \quad \int \frac{dx}{\cos^3 x}$$

13. Calcular $\int_{-\pi}^{\pi} \operatorname{sen} px\cos qx\,dx$, $\int_{-\pi}^{\pi} \operatorname{sen} px\operatorname{sen} qx\,dx$, $\int_{-\pi}^{\pi} \cos px\cos qx\,dx$ donde p y q son enteros.

14. Para $x\in\mathbb{R}$, y $n\in\mathbb{N}$, definamos $F(x) = \dfrac{a_o}{2} + \sum_{\substack{k=-n\\ k\neq 0}}^{n} (a_k\cos kx + b_k\operatorname{sen} kx)$. Para $-n \leqslant p \leqslant n$ prueba que:

$$a_p = \frac{1}{\pi}\int_{-\pi}^{\pi} F(x)\cos px\,dx \quad \text{y} \quad b_p = \frac{1}{\pi}\int_{-\pi}^{\pi} F(x)\operatorname{sen} px\,dx$$

Integrales del tipo $\int R\big(x, [L(x)]^r, [L(x)]^s, \dots\big)\, dx$

Donde $L(x) = \dfrac{ax+b}{cx+d}$, $a, b, c, d \in \mathbb{R}$ con $ad - bc \neq 0$ y $r, s, \dots$ son números racionales.

Se racionalizan con el cambio $t^q = L(x)$ donde q es el mínimo común denominador de las fracciones $r, s, \dots$. Pues entonces tenemos que

$$x = \frac{d\,t^q - b}{a - c\,t^q} = r(t)$$

y la integral se transforma en

$$\int R(r(t), t^{rq}, t^{sq}, \dots) r'(t)\, dt$$

en la que el integrando es una función racional de t.

8.37 Ejemplo. Sea $\mathbf{I} = \displaystyle\int \left(\frac{x+1}{x-1}\right)^{1/3} \frac{1}{1+x}\, dx$. El cambio de variable $\dfrac{x+1}{x-1} = t^3$ racionaliza la integral pues se tiene que $x = \dfrac{t^3+1}{t^3-1}$, con lo que:

$$\mathbf{I} = -3\int \frac{1}{t^3-1}\, dt = \int \left(\frac{t+2}{t^2+t+1} - \frac{1}{t-1}\right) dt = \frac{1}{2}\log\left(\frac{t^2+t+1}{(t-1)^2}\right) + \sqrt{3}\operatorname{arc\,tg}\frac{2t+1}{\sqrt{3}}$$

donde $t = \sqrt[3]{\dfrac{x+1}{x-1}}$. ♦

Integrales binomias

Se llaman así las de la forma

$$\int x^{\alpha}(a + bx^{\beta})^{\gamma} dx$$

donde α, β, γ son números racionales y a, b números reales todos ellos distintos de cero. Haciendo la sustitución

$$x^{\beta} = t, \quad x = t^{\frac{1}{\beta}}, \quad dx = \frac{1}{\beta}\, t^{\frac{1}{\beta}-1}$$

la integral se transforma en

$$\frac{1}{\beta}\int t^{\frac{\alpha+1}{\beta}-1}(a+bt)^{\gamma} dt$$

que es de la forma $\int t^r (a+bt)^{\gamma} dt$ donde $r = \dfrac{\alpha+1}{\beta} - 1$. Esta integral es del tipo de las consideradas en el apartado anterior cuando el número:

- γ es entero, pues es de la forma $\int R(t, t^r) dt$
- r es entero, pues es de la forma $\int R\big(t, (a+bt)^{\gamma}\big) dt$
- $\gamma + r$ es entero, pues es de la forma $\displaystyle\int \left(\frac{a+bt}{t}\right)^{\gamma} t^{\gamma+r} dt$

P.L. Chebyshev probó que si no se da ninguna de estas circunstancias la integral no puede expresarse por medio de funciones elementales.

8.38 Ejemplo. Sea $\mathbf{I}=\int x\sqrt{x^{2/3}+2}\,dx$. En este caso es $\alpha=1$, $\beta=2/3$, $\gamma=1/2$ y $\dfrac{\alpha+1}{\beta}=3$. Deducimos que la primitiva buscada puede expresarse por funciones elementales. Haciendo $x^{2/3}=t$ obtenemos $\mathbf{I}=\dfrac{3}{2}\int t^2\sqrt{t+2}\,dt$, la cual se racionaliza haciendo $t+2=s^2$ $(s>0)$, con lo que $\mathbf{I}=3\int(s^2-2)^2 s\,ds$ que es inmediata. ♦

Integrales del tipo $\int R(\mathrm{e}^x)dx$

Se racionalizan con el cambio $x=\log t$. Un caso particular de este es el de las integrales de la forma $\int R(\cosh x,\operatorname{senh} x)dx$ que también admiten un tratamiento parecido al de las trigonométricas.

8.39 Ejemplo. Sea $\mathbf{I}=\displaystyle\int\frac{2}{\operatorname{senh} x+\operatorname{tgh} x}dx$. Desarrolla los cálculos para comprobar que

$$\mathbf{I}=[x=\log t]=\int\frac{2(1+t^2)}{(t-1)(1+t)^3}dt=\log\left(\operatorname{tgh}\left(\frac{x}{2}\right)\right)-\frac{1}{1+\cosh x}$$

Por otra parte, como la función $\dfrac{2}{\operatorname{senh} x+\operatorname{tgh} x}$ es impar en $\operatorname{senh} x$, también podemos proceder como sigue

$$\mathbf{I}=[t=\cosh x]=\int\frac{2t}{(-1+t)(1+t)^2}dt=-\frac{1}{1+\cosh x}+\frac{1}{2}\log(-1+\cosh x)-\frac{1}{2}\log(1+\cosh x)$$

Por supuesto, puedes comprobar que las dos primitivas encontradas son de hecho iguales. ♦

Integración de funciones del tipo $R(x,\sqrt{ax^2+bx+c})$

Una integral de la forma $\int R(x,\sqrt{ax^2+bx+c})dx$ puede racionalizarse por medio de las sustituciones siguientes.

- Si el trinomio ax^2+bx+c tiene dos raíces reales α y β, distintas entonces

$$\sqrt{ax^2+bx+c}=[a(x-\alpha)(x-\beta)]^{1/2}=(x-\alpha)\left[\frac{a(x-\beta)}{x-\alpha}\right]^{1/2}$$

Donde, por comodidad, hemos supuesto que $x-\alpha>0$. Deducimos que la sustitución

$$\frac{a(x-\beta)}{x-\alpha}=t^2\ \ (t>0),\qquad x=\frac{\alpha t^2-\beta\alpha}{t^2-a}=r(t)$$

transforma la integral en $\int R\big(r(t),(r(t)-\alpha)t\big)r'(t)dt$ donde el integrando es una función racional de t.

- Si el trinomio ax^2+bx+c no tiene raíces reales, entonces debe ser $ax^2+bx+c>0$ para todo $x\in\mathbb{R}$, en particular $c>0$. La sustitución:

$$\sqrt{ax^2+bx+c}=tx+\sqrt{c},\qquad x=\frac{b-2t\sqrt{c}}{t^2-a}=g(t)$$

transforma la integral en $\int R\big(g(t),tg(t)+\sqrt{c}\big)g'(t)dt$ donde el integrando es una función racional de t.

Las sustituciones anteriores se conocen como *sustituciones de Euler*.

8.40 Ejemplo. Cálcular $\int \frac{x}{(7x-10-x^2)^{3/2}}\,dx$. Observa que, si $R(x,y)=\frac{x}{y^3}$, la integral que nos piden es $\int R(x,\sqrt{7x-10-x^2})dx$ del tipo que acabamos de considerar.

Como $7x-10-x^2=(x-2)(5-x)$, tenemos que

$$\int \frac{x}{(7x-10-x^2)^{3/2}}\,dx=\left[x=\frac{5+2t^2}{1+t^2}\right]=-\frac{6}{27}\int \frac{5+2t^2}{t^2}\,dt=-\frac{2}{9}\left(-\frac{5}{t}+2t\right)$$

donde $t=\dfrac{(7x-10-x^2)^{1/2}}{x-2}$. ♦

8.41 Ejemplo. $\int \frac{1}{(1+x)\sqrt{1+x+x^2}}\,dx$. Haciendo la sustitución $\sqrt{1+x+x^2}=x+t$, es decir $x=\dfrac{t^2-1}{1-2t}$ tenemos

$$\int \frac{1}{(1+x)\sqrt{1+x+x^2}}\,dx=\left[x=\frac{t^2-1}{1-2t}\right]=\int \frac{2}{t^2-2t}\,dt=\int\left(\frac{-1}{t}+\frac{1}{t-2}\right)dt=-\log t+\log|t-2|$$

donde $t=\sqrt{1+x+x^2}-x$. ♦

También es posible transformar una integral del tipo $\int R(x,\sqrt{ax^2+bx+c})dx$ en otra de la forma $\int F(\operatorname{sen}x,\cos x)dx$ donde F es una función racional de dos variables las cuales ya hemos estudiado. Para ello se sigue el siguiente procedimiento.

• Con un primer cambio de variable, de la forma $x=\alpha t+\beta$ que después explicaremos, se transforma la integral $\int R(x,\sqrt{ax^2+bx+c})dx$ en otra de alguna de las formas

$$\textbf{a)}\ \int G(t,\sqrt{t^2-1})dt,\quad \textbf{b)}\ \int G(t,\sqrt{1-t^2})dt,\quad \textbf{c)}\ \int G(t,\sqrt{1+t^2})dt$$

donde G es una función racional de dos variables. Los cambios de variable respectivos

$$\textbf{a)}\ x=\sec u,\quad \textbf{b)}\ x=\operatorname{sen}u,\quad \textbf{c)}\ x=\operatorname{tg}u$$

convierten las integrales anteriores en otras de la forma $\int F(\operatorname{sen}x,\cos x)dx$ donde F es una función racional de dos variables.

Alternativamente, en el caso **a)** puede hacerse también $x=\cosh u$, y en el caso **c)** $x=\operatorname{senh}u$, lo que transforma dichas integrales en otras del tipo $\int T(e^x)dx$ donde T es una función racional de una variable, que ya han sido estudiadas.

Nos queda por explicar cómo se hace el primer cambio de variable.

• Si el trinomio $h(x)=ax^2+bx+c$ tiene dos raíces reales $\alpha<\beta$, lo que se hace es transformar dicho trinomio en otro que tenga como raíces -1 y 1. Para ello llevamos -1 a α y 1 a β mediante una función de la forma $\varphi(t)=\lambda t+\mu$. Las condiciones $\varphi(-1)=\alpha$, $\varphi(1)=\beta$, determinan que $\lambda=\dfrac{\beta-\alpha}{2}$, $\mu=\dfrac{\beta+\alpha}{2}$. Con el cambio

$$x=\varphi(t)=\frac{\beta-\alpha}{2}t+\frac{\beta+\alpha}{2}$$

tenemos que $h(\varphi(t))=a\dfrac{(\beta-\alpha)^2}{4}(t^2-1)$. Ahora, si $a>0$, deducimos que

$$\int R(x,\sqrt{ax^2+bx+c})dx=[x=\varphi(t)]=\int R\left(\varphi(t),\sqrt{a}\,\frac{(\beta-\alpha)}{2}\sqrt{t^2-1}\right)\frac{\beta-\alpha}{2}\,dt$$

que es del tipo **a)** anterior. Si $a<0$, entonces

$$\int R(x,\ \sqrt{ax^2+bx+c})dx = [x=\varphi(t)] = \int R\left(\varphi(t),\ \sqrt{-a}\,\frac{(\beta-\alpha)}{2}\ \sqrt{1-t^2}\right)\frac{\beta-\alpha}{2}\,dt$$

que es del tipo **b)** anterior.

• Si el trinomio ax^2+bx+c no tiene raíces reales, entonces debe ser $d=4ac-b^2>0$ y también $a>0$. Poniendo $\gamma=\dfrac{\sqrt{d}}{2\sqrt{a}}$, podemos escribir:

$$ax^2+bx+c=\left(\sqrt{a}x+\frac{b}{2\sqrt{a}}\right)^2+c-\frac{b^2}{4a}=\left(\sqrt{a}x+\frac{b}{2\sqrt{a}}\right)^2+\gamma^2=\gamma^2\left[\left(\frac{\sqrt{a}}{\gamma}x+\frac{b}{2\sqrt{a}\gamma}\right)^2+1\right]=$$
$$=\gamma^2\left[\left(\frac{2a}{\sqrt{d}}x+\frac{b}{\sqrt{d}}\right)^2+1\right]$$

El cambio

$$\frac{2a}{\sqrt{d}}x+\frac{b}{\sqrt{d}}=t\,,\ \text{esto es}\ ,\ x=\frac{\sqrt{d}\,t-b}{2a}=\phi(t)$$

transforma la integral en

$$\int R(x,\ \sqrt{ax^2+bx+c})dx=[x=\phi(t)]=\int R\big(\phi(t),\gamma\sqrt{t^2+1}\big)\frac{\sqrt{d}}{2a}\,dt$$

que es del tipo **c)** anterior.

Casos particulares

• Las integrales de la forma $\displaystyle\int\frac{P(x)}{\sqrt{ax^2+bx+c}}\,dx$ donde $P(x)$ es una función polinómica pueden resolverse con facilidad por el *método de reducción*. Se procede de la siguiente forma.

Escribimos

$$\frac{P(x)}{\sqrt{ax^2+bx+c}}=\frac{d}{dx}\left(Q(x)\,\sqrt{ax^2+bx+c}\right)+\frac{C}{\sqrt{ax^2+bx+c}}$$

donde $Q(x)$ es un polinomio, cuyos coeficientes hay que calcular, de grado una unidad menos que el polinomio $P(x)$ y C es una constante que también hay que calcular. Observa que la igualdad anterior puede escribirse

$$P(x)=Q'(x)(ax^2+bx+c)+\frac{1}{2}\,Q(x)(2ax+b)+C$$

y a la derecha queda un polinomio de igual grado que $P(x)$ lo que permite identificar coeficientes. Una vez calculados el polinomio Q y la constante C tenemos que

$$\int\frac{P(x)}{\sqrt{ax^2+bx+c}}\,dx=Q(x)\,\sqrt{ax^2+bx+c}+C\int\frac{1}{\sqrt{ax^2+bx+c}}\,dx$$

con lo que todo se reduce a calcular una integral de la forma $\displaystyle\int\frac{1}{\sqrt{ax^2+bx+c}}\,dx$. Haciendo uso de los cambios antes visto, esta integral, salvo constantes, puede escribirse de alguna de las formas

$$\int\frac{1}{\sqrt{1-t^2}}\,dt=\operatorname{arc\,sen}(t)\,,\quad\int\frac{1}{\sqrt{1+t^2}}\,dt=\operatorname{argsenh}(t)\,,\quad\int\frac{1}{\sqrt{t^2-1}}\,dt=\operatorname{argcosh}(t)$$

Recuerda que $\operatorname{argsenh}(t) = \log\left(t + \sqrt{t^2+1}\right)$ y $\operatorname{argcosh}(t) = \log\left(t + \sqrt{t^2-1}\right)$.

• Finalmente, las integrales de la forma

$$\int \frac{1}{(x-\alpha)^k \sqrt{ax^2+bx+c}}\,dx$$

se reducen a las del tipo anterior con el cambio $x - \alpha = \dfrac{1}{t}$.

Ejercicios

15. Calcular la integrales

$$\int \frac{x+3}{\sqrt{x^2+2x+2}}\,dx, \qquad \int \frac{x^2}{\sqrt{2x-x^2}}\,dx, \qquad \int \frac{1}{x^2\sqrt{x^2-x+1}}\,dx, \qquad \int \sqrt{2ax-x^2}\,dx$$

$$\int \frac{1}{(1-x^2)\sqrt{1+x^2}}\,dx, \qquad \int \frac{1}{x^2\sqrt[3]{(4+x^3)^5}}\,dx, \qquad \int x^{7/2}(1-x^3)^{-2}dx, \qquad \int \frac{\sqrt{x^2+9x}}{x^2}\,dx$$

$$\int \frac{1}{2\operatorname{senh}x - \cosh x}\,dx, \qquad \int \sqrt[3]{x}(1+\sqrt{x})^{-2}dx, \qquad \int \frac{\sqrt{5-8x-4x^2}}{x+5/2}\,dx, \qquad \int x^{-4}(1+x^2)^{-1/2}dx$$

8.4. Aplicaciones de la integral

Con una integral puedes calcular magnitudes tan diversas como áreas, volúmenes, longitudes de curvas, el trabajo realizado por una fuerza, la masa de un sólido, momentos de inercia, el campo eléctrico, el flujo de un fluido a través de una superficie y muchas más. Es notable, sin embargo, que la forma de proceder sea casi siempre la misma, y consiste en expresar el valor exacto de la magnitud que se quiere calcular como un límite de sumas de Riemann, para deducir, a partir de ellas, la integral cuyo cálculo proporciona la solución del problema. Podrás comprobar en lo que sigue que esta técnica es bastante sencilla e intuitiva. Con un poco de práctica tú mismo podrás aplicarla con éxito en situaciones distintas de las que aquí se consideran.

Todo lo que sigue está también en formato de cuaderno de *Mathematica* en el sitio http://wwww.ugr.es/local/fjperez.

8.4.1. Cálculo de áreas planas

Te recuerdo que si $f : [a,b] \to \mathbb{R}$ es una función continua, representamos por $G(f,a,b)$ la región del plano comprendida entre la curva $y = f(x)$, el eje de abscisas y las rectas $x = a$, $x = b$. Como sabes, el área de dicha región viene dada por

$$\lambda(G(f,a,b)) = \int_a^b |f(x)|\,\mathrm{d}x$$

Es interesante interpretar la integral que proporciona el área de la siguiente forma. Observa que $|f(x)|$ es la *longitud* del segmento intersección de $G(f,a,b)$ con la recta vertical que pasa por $(x,0)$, es decir, $|f(x)|$ es la longitud de la *sección vertical* de $G(f,a,b)$ por el punto $(x,0)$, y *el área de la región* $G(f,a,b)$ *es igual a la integral de las longitudes de sus secciones.* Intuitivamente: integrando longitudes obtenemos áreas. Como el área es invariante por rotaciones, este resultado es también válido si consideramos secciones por rectas paralelas a una recta cualquiera dada. Deducimos así el siguiente resultado.

Principio de Cavalieri. *El área de una región plana es igual a la integral de las longitudes de sus secciones por rectas paralelas a una recta dada.*

Veamos cómo se aplica este principio en algunos casos concretos.

Regiones de tipo I

Supongamos que f, g son funciones continuas y llamemos Ω a la región del plano comprendida entre las curvas $y = f(x)$ e $y = g(x)$ para $a \leqslant x \leqslant b$. Se dice que Ω es una región de tipo I. Es evidente que las longitudes de las secciones verticales de Ω son iguales a $|f(x) - g(x)|$ por lo que su área viene dada por

$$\int_a^b |f(x) - g(x)|\,\mathrm{d}x \tag{8.9}$$

Observa que esta integral expresa el área de Ω como límite de las sumas de Riemann

$$\sum_{k=1}^{n} |f(t_k) - g(t_k)|\,(x_k - x_{k-1})$$

lo que tiene una sencilla interpretación que puedes ver en la siguiente figura.

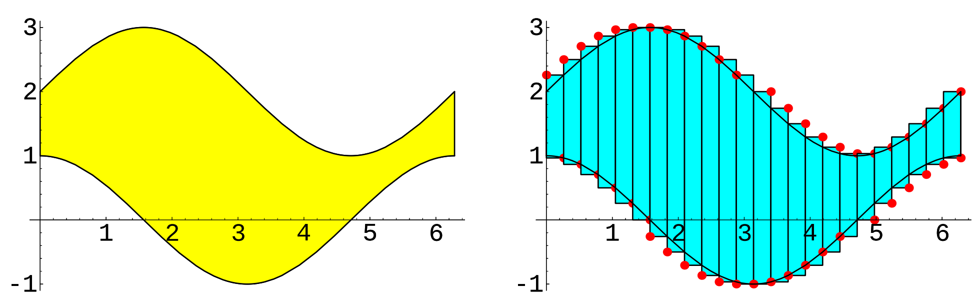


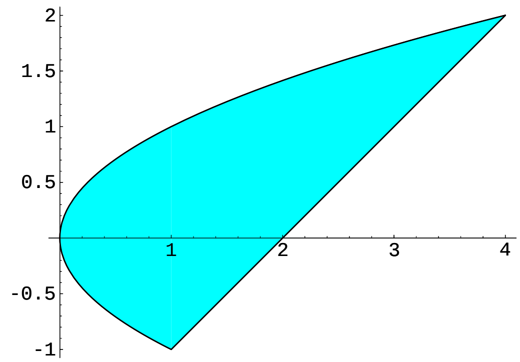
Figura 8.4: Región de tipo I

Cuando la función $f - g$ no tiene signo constante en el intervalo $[a,b]$, para calcular la integral (8.9) se descompone dicho intervalo en intervalos en los que la función $f - g$ es siempre positiva o siempre negativa, lo que permite quitar el valor absoluto en el integrando.

A veces interesa expresar una región de tipo I como unión de dos o más regiones de tipo I disjuntas y más sencillas, entonces su área es la suma de las áreas de cada una de dichas regiones.

8.42 Ejemplo. Vamos a calcular el área de la región Ω comprendida entre la parábola $y^2 = x$ y la recta $y = x - 2$.

Calculamos los puntos de corte de la recta y la parábola resolviendo la ecuación $x = (x-2)^2$, cuyas soluciones son $a = 1$, $b = 4$. Puedes ver representada la región Ω en azul en la siguiente figura.



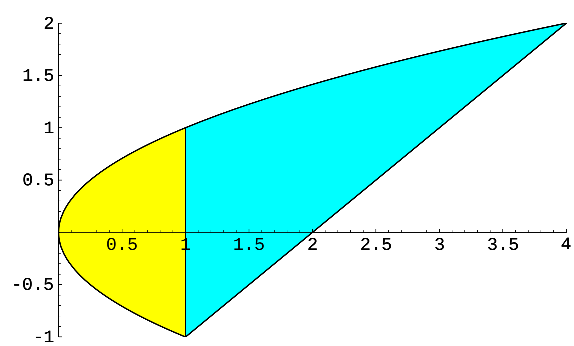
Podemos considerar Ω como una región de tipo I. La función cuya gráfica limita a Ω por arriba es $g(x) = \sqrt{x}$. La función cuya gráfica limita a Ω por abajo viene dada por

$$f(x) = \begin{cases} -\sqrt{x} & 0 \leqslant x \leqslant 1 \\ x-2 & 1 \leqslant x \leqslant 4 \end{cases}$$

En consecuencia

$$\lambda(\Omega) = \int_0^4 |g(x) - f(x)|\,\mathrm{d}x = \int_0^1 (\sqrt{x} - (\sqrt{x}))\,\mathrm{d}x + \int_1^4 (\sqrt{x} - (x-2))\,\mathrm{d}x = \frac{9}{2}$$

Observa que podemos ver Ω como unión de dos regiones de tipo I como se indica en la siguiente figura.



Y lo que hemos hecho antes ha sido calcular el área de cada una de estas dos regiones. ♦

Regiones de tipo II

Supongamos que f, g son funciones continuas y llamemos Ω a la región del plano comprendida entre las curvas $x = f(y)$ y $x = g(y)$ para $a \leqslant y \leqslant b$. Se dice que Ω es una región de tipo II. Es evidente que las longitudes de las secciones horizontales de Ω son iguales a $|f(y) - g(y)|$ por lo que su área viene dada por

$$\int_a^b |f(y) - g(y)|\,\mathrm{d}y \tag{8.10}$$

lo que tiene una sencilla interpretación que puedes ver en la siguiente figura.

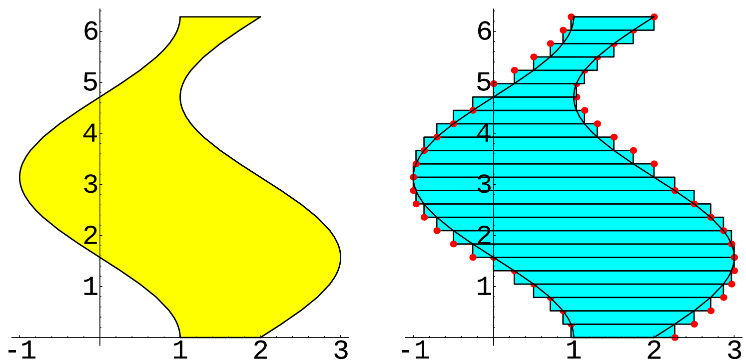


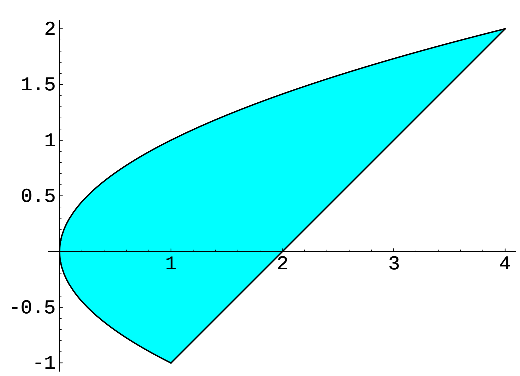
Figura 8.5: Región de tipo II

Es importante advertir que la distinción entre regiones de tipo I y de tipo II es tan sólo una cuestión de conveniencia. No son conjuntos de distinta naturaleza sino formas distintas de describir un conjunto. En la práctica te vas a encontrar siempre con regiones que puedes considerar tanto de tipo I como de tipo II y deberás elegir la descripción que más facilite el cálculo de la correspondiente integral.

De todas formas, no debes olvidar que basta cambiar la variable x por la variable y para convertir una región de tipo II en otra de tipo I. Geométricamente, lo que hacemos es una simetría respecto a la recta $y = x$, lo que deja invariante el área. Por tanto, si en un ejercicio resulta conveniente considerar la región cuya área quieres calcular como una región de tipo II y te encuentras más cómodo trabajando con regiones de tipo I, basta con que cambies los nombres de las variables.

Observa que las figuras (8.4) y (8.5) son simétricas respecto de la recta $y = x$.

8.43 Ejemplo. La región del ejemplo (8.42) puedes considerarla como una región de tipo II.



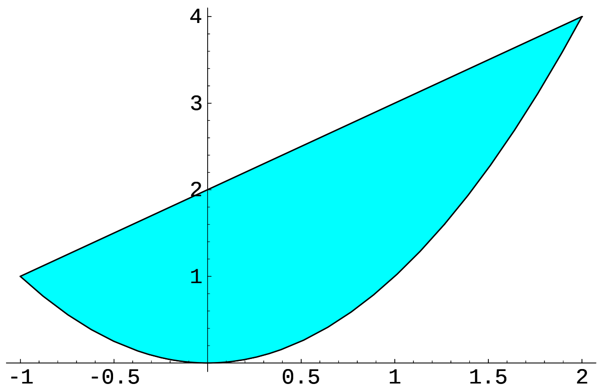
La curva que limita esta región por la derecha es la gráfica de la recta $x = y+2$ y la curva que limita esta región por la derecha es la gráfica de la parábola $x = y^2$. La variable y está comprendida entre -1 y 2.

$$\Omega=\left\{(x,y): y^2\leqslant x\leqslant y+2,\ -1\leqslant y\leqslant 2\right\}$$

Tenemos que

$$\lambda(\Omega)=\int_{-1}^{2}(y+2-y^2)\,\mathrm{d}y=\frac{9}{2}$$

También puedes transformar directamente Ω en una región de tipo I más sencilla mediante una simetría. Aquí la tienes.



Aunque la región así obtenida no es la misma Ω tiene, sin embargo, igual área que Ω pues ambas regiones se transforman una en otra por medio de una simetría respecto de la recta $y=x$.

♦

8.4.2. Ejercicios

1. Calcula el área de las regiones del plano limitadas por las siguientes curvas.

   a) $x=12y^2-12y^3, \quad x=2y^2-2y.$

   b) $y=-x^2-2x, \quad y=x^2-4, \quad -3\leqslant x\leqslant 1.$

   c) $y=x^2, \quad x+y=2, \quad x\geqslant 0,\ y\geqslant 0.$

   d) $x+y^2=3, \quad 4x+y^2=4.$

   e) $y=\sec^2 x, \quad y=\operatorname{tg}^2 x, \quad -\pi/4\leqslant x\leqslant \pi/4.$

   f) $\dfrac{x^2}{4}+\dfrac{y^2}{9}=1.$

   g) $(y-x)^2=x-3, \quad x=7.$

   h) $y=(\log x)^2, \quad 0<x\leqslant \mathrm{e}.$

   i) $y^2=\dfrac{1-x}{1+x}, \quad x=-1.$

   j) $y=x\,\mathrm{e}^{-x}, \quad y=x^2\,\mathrm{e}^{-x}, \quad x\geqslant 0.$

   k) $y^2=x, \quad x^2+y^2=8.$

Curvas en el plano

Seguramente te imaginas una curva en el plano como una línea continua que puede dibujarse de un trazo, sin levantar el lápiz del papel. Esa idea es esencialmente correcta. Las circunferencias, las elipses, las cardioides son todas ellas curvas. Faltaría más. Ninguna de ellas puedes representarla por una igualdad de la forma $y = f(x)$. Las curvas que pueden representarse por una ecuación cartesiana del tipo $y = f(x)$ son curvas muy particulares pues son gráficas de funciones. No olvides que cuando dices "sea la curva dada por la ecuación $y = f(x)$" te estás refiriendo a la curva cuya imagen es el conjunto de puntos del plano $\{(x,y) : x \in [a,b], y = f(x)\}$ es decir, a la gráfica de f.

Si lo piensas un momento verás que muy pocas curvas son gráficas. Para que una curva sea una gráfica es necesario que cualquier recta vertical la corte a lo más en un solo punto; ninguna curva cerrada cumple esta condición. Precisamente entre las curvas cerradas se encuentran algunas de las curvas más interesantes, a ellas pertenecen los distintos tipos de óvalos y lemniscatas, las astroides, las cardioides y muchas más.

Vamos a ver ahora una forma de representar curvas planas mucho más general que las ecuaciones cartesianas del tipo $y = f(x)$ que sólo sirven para representar curvas que también son gráficas. Para empezar, consideremos una curva que viene dada por una ecuación cartesiana de la forma $y = f(x)$ donde $x \in [a,b]$. Nuestra curva es, por tanto, la imagen de la aplicación $\gamma : [a,b] \to \mathbb{R}^2$ definida por $\gamma(x) = (x, f(x))$ para todo $x \in [a,b]$. Intuitivamente, cuando x recorre el intervalo $[a,b]$, el punto $(x, f(x))$ recorre la curva. Es fácil generalizar esta situación sin perder la idea intuitiva de curva. Lo esencial es que podamos describir las coordenadas de los puntos de la curva como funciones continuas de un parámetro. En la situación que estamos considerando se tiene que $y = f(x)$, es decir, la segunda coordenada es función continua de la primera. La generalización consiste en que ambas coordenadas sean funciones continuas de un parámetro. Llegamos así a la definición siguiente.

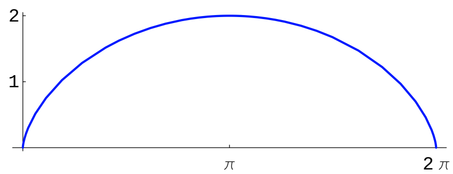
8.44 Definición. Una curva en el plano es una aplicación continua $\gamma : [a,b] \to \mathbb{R}^2$.

Si $\gamma(t) = (x(t), y(t))$, decimos que $x = x(t)$, $y = y(t)$ son las **ecuaciones paramétricas** de la curva. El punto $\gamma(a)$ es el origen y $\gamma(b)$ el extremo de la curva. Si $\gamma(a) = gamma(b)$ se dice que la curva es **cerrada**. Se dice que una curva γ es **simple** si no se corta a sí misma, es decir, si para $s, t \in [a,b]$ con $s \neq t$ se verifica que $\gamma(s) \neq \gamma(t)$. Una curva cerrada se llama simple si la función γ es inyectiva en $]a,b[$.

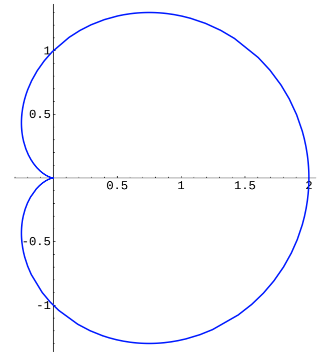
8.45 Ejemplo. • La curva de ecuaciones paramétricas $x(t) = a + r\cos t$, $y(t) = b + R\operatorname{sen} t$ donde $0 \leqslant t \leqslant 2\pi$ es una elipse cuyo centro es el punto (a,b) y semiejes de longitudes r y R. Cuando $r = R$ se trata de una circunferencia.

• La curva de ecuaciones paramétricas $x(t) = r(t - \operatorname{sen} t)$, $y(t) = r(1 - \cos t)$ para $0 \leqslant t \leqslant 2\pi$ es la **cicloide**. Es la curva que describiría una chincheta clavada en una rueda de radio r que avanza girando sin deslizar.

• La curva de ecuaciones paramétricas $x(t) = \cos t(1 + \cos t)$, $y(t) = \operatorname{sen} t(1 + \cos t)$ para $0 \leqslant t \leqslant 2\pi$ se llama **cardioide**. Es la curva que describe un punto fijo del borde de un círculo que rueda sin deslizar sobre otro del mismo radio.



Cicloide



Cardioide

♦

Área encerrada por una curva

Sea Ω la región rodeada por una curva cerrada simple $\gamma(t) = (x(t), y(t))$, $a \leqslant t \leqslant b$, y supongamos que las funciones $x(t), y(t)$ tienen primera derivada continua. En estas condiciones se verifica que el área de Ω viene dada por

$$\lambda(\Omega) = \left| \int_a^b x(t)y'(t)\,\mathrm{d}t \right| = \left| \int_a^b x'(t)y(t)\,\mathrm{d}t \right| \tag{8.11}$$

(la igualdad entre las dos integrales se deduce fácilmente integrando por partes y teniendo en cuenta que por ser γ una curva cerrada es $x(b)y(b) - x(a)y(a) = 0$).

Ejercicios

1. Calcula el área encerrada por la elipse $x(t) = a + r\cos t$, $y(t) = b + R\operatorname{sen} t$ donde $0 \leqslant t \leqslant 2\pi$.
2. Calcula el área encerrada por la cardioide $x(t) = \cos t(1 + \cos t)$, $y(t) = \operatorname{sen} t(1 + \cos t)$ para $0 \leqslant t \leqslant 2\pi$.

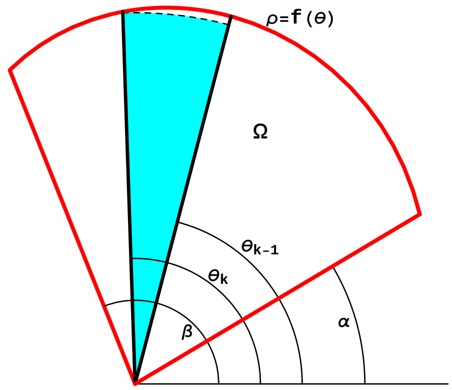
Áreas planas en coordenadas polares

Dado un punto $(x, y) \neq (0, 0)$, hay un único par de números (ρ, ϑ) tales que $\rho > 0$, $-\pi < \vartheta \leqslant \pi$, que verifican las igualdades $x = \rho\cos\vartheta$, $y = \rho\operatorname{sen}\vartheta$. Dichos números se llaman **coordenadas polares** del punto (x, y). Si consideras el número complejo $x + iy$, entonces ρ es su módulo y ϑ es su argumento principal.

Una curva puede venir dada en coordenadas polares por medio de una ecuación de la forma $\rho = f(\vartheta)$ donde $f : [\alpha, \beta] \to \mathbb{R}$ es una función continua. Esta forma de representar una curva no es más que la parametrización dada por

$$\begin{cases} x(\vartheta) = f(\vartheta)\cos\vartheta \\ y(\vartheta) = f(\vartheta)\operatorname{sen}\vartheta \end{cases} \qquad (\alpha \leqslant \vartheta \leqslant \beta) \tag{8.12}$$

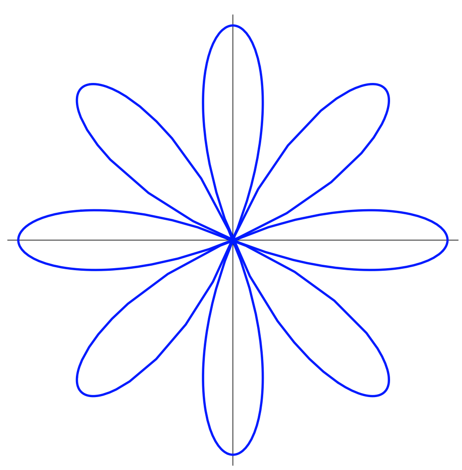
Queremos calcular el área de la región del plano $\Omega = \{(\rho\cos\vartheta, \rho\cos\vartheta) : 0 < \rho \leqslant f(\vartheta),\ \alpha \leqslant \vartheta \leqslant \beta\}$.



Para ello lo que hacemos es aproximar Ω por medio de sectores circulares. Recuerda que el área de un sector circular de radio ρ y amplitud φ (medida en radianes) es igual a $\dfrac{1}{2}\rho^2\varphi$. Consideramos para ello una partición $\{\alpha = \vartheta_0, \vartheta_1, \vartheta_2, \dots, \vartheta_{n-1}, \vartheta_n = \beta\}$ de $[\alpha, \beta]$ y formamos la suma $\displaystyle\sum_{k=1}^{n} \frac{1}{2} f(\vartheta_k)^2(\vartheta_k - \vartheta_{k-1})$. Como el número $\dfrac{1}{2} f(\vartheta_k)^2(\vartheta_k - \vartheta_{k-1})$ es el área del sector circular, representado en azul en la figura, de radio $f(\vartheta_k)$ y amplitud igual a $\vartheta_k - \vartheta_{k-1}$, es claro que la suma anterior representa una aproximación del área de Ω. Como $\displaystyle\sum_{k=1}^{n} \frac{1}{2} f(\vartheta_k)^2(\vartheta_k - \vartheta_{k-1})$ es una suma de Riemann de la función $\vartheta \mapsto \dfrac{1}{2} f(\vartheta)^2$, se sigue que el área de Ω viene dada por la integral

$$\frac{1}{2}\int_{\alpha}^{\beta} f(\vartheta)^2\, d\vartheta$$

Con frecuencia, las ecuaciones en coordenadas polares se usan para representar distintos tipos de curvas simétricas llamadas "rosa". Por ejemplo, aquí tienes una rosa de 8 hojas o lazos, cuya ecuación en coordenadas polares es $\rho = \cos(4\vartheta)$, $0 \leqslant \vartheta \leqslant 2\pi$.

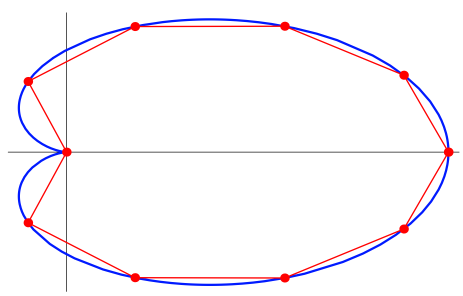


Ejercicios

1. Calcula el área de la región del plano rodeada por un lazo de la lemniscata de ecuación polar $\rho^2 = \cos(2\vartheta)$, $(-\pi/4 \leqslant \vartheta \leqslant \pi/4)$.
2. Calcula el área limitada por el arco de la espiral de Arquímides $\rho = a\,\vartheta$, $a > 0$, comprendido entre $\vartheta = 0$ y $\vartheta = \pi$.
3. Calcula el área encerrada por el lazo interior de la curva $\rho = \dfrac{1}{2} + \cos\vartheta$.
4. Hallar el área encerrada por una de las hojas de la rosa $\rho = 2\cos(2\vartheta)$.
5. Calcular el área del lóbulo del folium de Descartes de ecuación $x^3 + y^3 - 3axy = 0$, $a > 0$. Sugerencia. Expresa la ecuación en coordenadas polares.
6. Calcula el área de la región común a las dos elipses
$$\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1, \quad \frac{x^2}{b^2} + \frac{y^2}{a^2} = 1$$
Sugerencia. Representa gráficamente las elipses. Usa la simetría polar para simplificar los cálculos y pasar a coordenadas polares.

8.4.3. Longitud de un arco de curva

Se trata de calcular la longitud de la curva plana γ dada por las ecuaciones paramétricas
$$\begin{cases} x = x(t) \\ y = y(t) \end{cases} \qquad (a \leqslant t \leqslant b)$$
donde suponemos que $x(t)$, $y(t)$ tienen derivada primera continua. Para ello aproximamos la curva por poligonales inscritas en ella. Cada partición $\{a = t_0, t_1, t_2, \dots, t_{n-1}, t_n = b\}$ induce una poligonal cuyos vértices son los puntos $\gamma(t_k) = (x(t_k), y(t_k))$, $(0 \leqslant k \leqslant n)$.



La longitud de dicha poligonal viene dada por
$$\sum_{k=1}^{n} \sqrt{(x(t_k) - x(t_{k-1}))^2 + (y(t_k) - y(t_{k-1}))^2} \approx \sum_{k=1}^{n} \sqrt{x'(s_k)^2 + y'(s_k)^2}\,(t_k - t_{k-1})$$
Donde hemos usado el teorema del valor medio y la continuidad de las derivadas. Pero esta es una suma de Riemann de la función $t \mapsto \sqrt{x'(t)^2 + y'(t)^2}$. Deducimos que la longitud de la curva γ viene dada por
$$\ell(\gamma) = \int_a^b \sqrt{x'(t)^2 + y'(t)^2}\,\mathrm{d}t \tag{8.13}$$

Para el caso particular de que la curva sea la gráfica de una función $y = f(x)$, esto es $\gamma(x) = (x, f(x))$, entonces su longitud viene dada por

$$\ell(\gamma) = \int_a^b \sqrt{1 + f'(x)^2}\,dx$$

Para el caso particular de que la curva venga dada por una parametrización polar de la forma (8.12), su longitud viene dada por

$$\ell(\gamma) = \int_\alpha^\beta \sqrt{f(\vartheta)^2 + f'(\vartheta)^2}\,d\vartheta$$

Si interpretamos que la curva $\gamma(t) = (x(t), y(t))$ es la *función de trayectoria* seguida por un móvil, entonces la *velocidad* de dicho móvil en cada instante t viene dada por el vector derivada $\gamma'(t) = (x'(t), y'(t))$, y la *rapidez* es la norma euclídea de dicho vector, es decir $\sqrt{x'(t)^2 + y'(t)^2}$. La igualdad (8.13) tiene ahora una interpretación clara: la distancia recorrida por un móvil se obtiene integrando la rapidez. Volveremos sobre esto más adelante.

Ejercicios

1. Calcula la longitud del arco de catenaria $y = \cosh x$ entre $x = 0$ y $x = 1$.
2. Calcula la longitud de un arco de la cicloide $x(t) = t - \operatorname{sen} t$, $y(t) = 1 - \cos t$, $(0 \leqslant t \leqslant 2\pi)$.
3. Calcular la longitud del arco de curva $y = x^2 + 4$, entre $x = 0$ y $x = 3$.
4. Calcula la longitud de la astroide $\left(\dfrac{x}{a}\right)^{2/3} + \left(\dfrac{y}{a}\right)^{2/3} = 1$, $a > 0$.

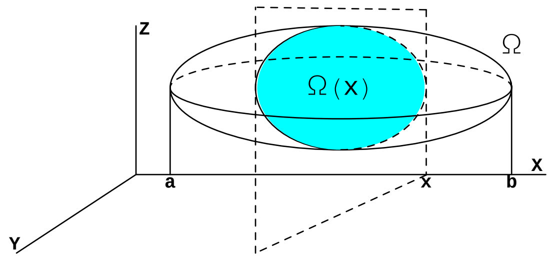
   Sugerencia. Obtener las ecuaciones paramétricas de la astroide y tener en cuenta la simetría.
5. Calcula la longitud de la cardioide $\rho = 3(1 + \cos\vartheta)$, $(0 \leqslant \vartheta \leqslant 2\pi)$.
6. Calcula la longitud de la curva $y = \dfrac{x^4 + 48}{24x}$ donde $2 \leqslant x \leqslant 4$.
7. Calcula la longitud de la curva $y = \log(1 - x^2)$, donde $1/3 \leqslant x \leqslant 2/3$.

8.4.4. Volúmenes de sólidos

Al igual que podemos calcular áreas de regiones planas integrando las longitudes de sus secciones por rectas paralelas a una dada, podemos también calcular volúmenes de regiones en $\mathbb{R}^3$ integrando las áreas de sus secciones por planos paralelos a uno dado. Este resultado es un caso particular del teorema de Fubini que veremos al estudiar integrales múltiples.

Cálculo de volúmenes por secciones planas. *El volumen de una región en aes igual a la integral del área de sus secciones por planos paralelos a uno dado.*

Para justificar esta afirmación, sea Ω una región en $\mathbb{R}^3$ como la de la figura.



Representemos por $\Omega(x)$ la sección de Ω por el plano perpendicular al eje OX en el punto $(x,0,0)$. Sea $V(x)$ el volumen de la parte de Ω que queda a la izquierda de dicho plano y sea $\lambda(\Omega(x))$ el área de la sección $\Omega(x)$. Observa que la situación es totalmente análoga a la considerada en el Teorema Fundamental del Cálculo: allí teníamos la función área cuya derivada era la longitud de la sección. No debe sorprenderte por ello que ahora resulte que la derivada de la función volumen, $V(x)$, sea el área de la sección. En efecto, sea $h>0$. Suponiendo, naturalmente, que la función $x\mapsto\lambda(\Omega(x))$ es continua, tenemos que

$$\min\{\lambda(\Omega(t)): x\leqslant t\leqslant x+h\}\,h\leqslant V(x+h)-V(x)\leqslant \max\{\lambda(\Omega(t)): x\leqslant t\leqslant x+h\}\,h$$

de donde se deduce que $\lim\limits_{h\to 0}\dfrac{V(x+h)-V(x)}{h}=\lambda(\Omega(x))$. Hemos obtenido así que $V'(x)=\lambda(\Omega(x))$. Deducimos que el volumen de Ω viene dado por la integral

$$Vol(\Omega)=\int_a^b\lambda(\Omega(x))\,\mathrm{d}x \tag{8.14}$$

El razonamiento anterior se ha hecho para secciones por planos verticales al eje OX, es decir planos paralelos al plano YZ; pero el resultado obtenido también es válido para secciones por planos paralelos a un plano dado.

Podemos llegar también a este resultado considerando sumas de Riemann. Para ello aproximamos la región Ω por cilindros de la siguiente forma. Consideremos una partición

$$\{a=x_0,x_1,x_2,\dots,x_{n-1},x_n=b\}$$

de $[a,b]$. La parte de Ω comprendida entre los planos perpendiculares al eje OX por los puntos $(x_{k-1},0,0)$ y $(x_k,0,0)$ puede aproximarse por un cilindro de altura x_k-x_{k-1} y base $\Omega(x_k)$ cuyo volumen es igual $\lambda(\Omega(x_k))(x_k-x_{k-1})$. La suma de los volúmenes de todos estos cilindros, $\displaystyle\sum_{k=1}^n\lambda(\Omega(x_k))(x_k-x_{k-1})$ es por tanto una aproximación del volumen de Ω, pero dicha suma es una suma de Riemann de la función $x\mapsto\lambda(\Omega(x))$, por lo que el volumen de Ω viene dado por $\displaystyle\int_a^b\lambda(\Omega(x))\,\mathrm{d}x$.

Vamos a estudiar algunos casos en los que es fácil calcular el área de las secciones de Ω.

Volumen de un cuerpo de revolución

Los cuerpos de revolución o sólidos de revolución son regiones de $\mathbb{R}^3$ que se obtienen girando una región plana alrededor de una recta llamada eje de giro.

• Método de los discos

Sea $f:[a,b]\to\mathbb{R}$ una función continua. Girando la región del plano comprendida entre la curva $y=f(x)$, el eje de abscisas y las rectas $x=a$ y $x=b$, alrededor del eje OX obtenemos un sólido de revolución Ω. Es evidente que la sección, $\Omega(x)$, de Ω por el plano perpendicular al eje OX en el punto $(x,0,0)$, es un disco contenido en dicho plano de centro $(x,0,0)$ y radio $|f(x)|$. Por tanto el área de $\Omega(x)$ es $\lambda(\Omega(x))=\pi f(x)^2$; en consecuencia el volumen de Ω es igual a

$$Vol(\Omega)=\pi\int_a^b f(x)^2\,dx$$

El volumen del sólido de revolución, Ω, obtenido girando alrededor del eje OX una región de tipo I definida por dos funciones continuas $f,g:[a,b]\to\mathbb{R}$ tales que $0\leqslant f(x)\leqslant g(x)$ para todo $x\in[a,b]$, viene dado por

$$Vol(\Omega)=\pi\int_a^b (g(x)^2-f(x)^2)\,dx$$

Una expresión similar se obtiene para el volumen de un sólido de revolución obtenido girando alrededor del eje OY una región de tipo II.

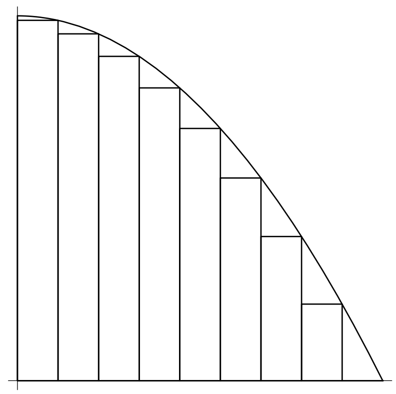
Ejercicios

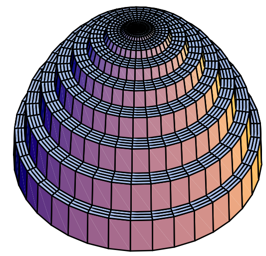
1. Calcula el volumen de la esfera obtenida girando la circunferencia $x^2+y^2=R^2$ alrededor del eje OX.
2. Calcula el volumen del cono circular recto de altura h y radio de la base R obtenido girando la recta $y=Rx/h$ entre $x=0$ y $x=h$.
3. Calcula el volumen del sólido engendrado al girar alrededor del eje OX la parte de la curva $y=\operatorname{sen}^2 x$ comprendida entre 0 y π.
4. Calcula el volumen del sólido engendrado al girar alrededor del eje OX la gráfica de la función $f:[0,+\infty[\to\mathbb{R}$ dada por $f(x)=\dfrac{18x}{x^2+9}$.
5. Calcula el volumen del sólido engendrado al girar la región limitada por la parábola $y^2=4x$ y la recta $x=4$ alrededor de dicha recta.
6. Calcula el volumen del sólido engendrado al girar la región limitada por las parábolas $y^2=x, x^2=y$ alrededor del eje OX.
7. Calcula el volumen del elipsoide $\dfrac{x^2}{a^2}+\dfrac{y^2}{b^2}+\dfrac{z^2}{c^2}=1$.
8. Calcula el volumen limitado por el paraboloide $\dfrac{x^2}{9}+\dfrac{y^2}{16}=z$ y el plano $z=7$.

• Método de las capas o de los tubos

Consideremos una función positiva $f:[a,b]\to\mathbb{R}$ y la región $G(f,a,b)$ limitada por la gráfica de dicha función y las rectas verticales $x=a$, $x=b$. Girando dicha región alrededor del eje OY obtenemos un sólido de revolución, Ω, cuyo volumen podemos aproximar considerando rectángulos verticales inscritos en la gráfica de f y girándolos alrededor del eje OY.

En la siguiente figura puedes ver rectángulos inscritos en la gráfica de la parábola $y = 1 - x^2$ en el intervalo $[0,1]$ y el cuerpo de revolución que engendran al girarlos alrededor del eje OY.





Naturalmente, la aproximación va mejorando a medida que aumentamos el número de puntos de división del intervalo.

Consideremos una partición $\{a = x_0, x_1, x_2, \dots, x_{n-1}, x_n = b\}$ de $[a,b]$. Al girar alrededor del eje OY un rectángulo vertical cuya base es el intervalo $[x_{k-1}, x_k]$ y altura $f(x_k)$, obtenemos una lámina de un cilindro circular recto, esto es, un *tubo* cuya base tiene área $\pi(x_k^2 - x_{k-1}^2)$ y altura $f(x_k)$, cuyo volumen es, por tanto, igual a

$$\pi(x_k^2 - x_{k-1}^2) f(x_k) = \pi(x_k - x_{k-1})(x_k + x_{k-1}) f(x_k) = x_k f(x_k)(x_k - x_{k-1}) + x_{k-1} f(x_k)(x_k - x_{k-1})$$

La suma de todos ellos es igual a

$$\sum_{k=1}^{n} \pi x_k f(x_k)(x_k - x_{k-1}) + \sum_{k=1}^{n} \pi x_{k-1} f(x_k)(x_k - x_{k-1})$$

Pero estas dos sumas son sumas de Riemann de la función $x \mapsto \pi x f(x)$, deducimos que el volumen de Ω viene dado por

$$Vol(\Omega) = 2\pi \int_a^b x f(x)\,\mathrm{d}x$$

Esto es lo que se conoce como *método de las capas o de los tubos*. Puedes adaptar fácilmente esta expresión para el caso de que el eje de giro sea la recta vertical $x = c$. En general, si notamos por $R(x)$ el “radio de giro” de la lámina, entonces

$$Vol(\Omega) = 2\pi \int_a^b R(x) f(x)\,\mathrm{d}x$$

Ejercicios

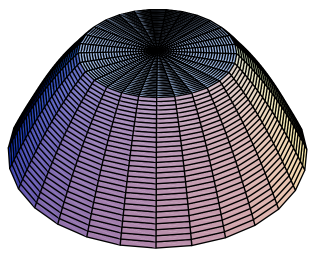
1. Calcula el volumen del toro engendrado al girar el círculo de centro $(a,0)$ y radio $R < a$ alrededor del eje OY.

2. Calcular el volumen del sólido Ω engendrado al girar la región limitada por las parábolas $y = x^2$, $x = y^2$ alrededor del eje OY.

3. Calcular el volumen del toro engendrado al girar el círculo de centro 0 y radio 3 alrededor de la recta $x = 6$.

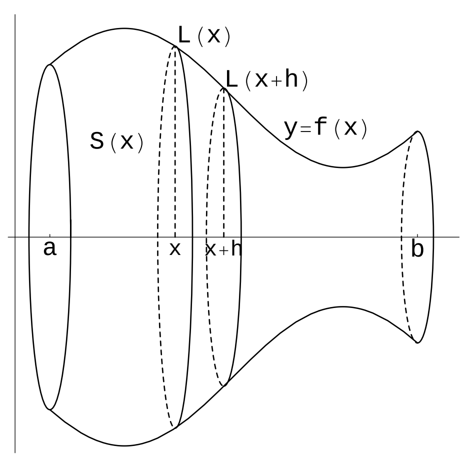
4. Calcular el volumen del sólido Ω engendrado al girar la región limitada por las parábolas $y = x^2$, $x = y^2$ alrededor la recta $x = 4$.

5. Un flan tiene forma de tronco de paraboloide de revolución, siendo r y $2r$ los radios de sus bases y h su altura. Calcular su volumen volumen y el volumen de la porción obtenida al cortarlo verticalmente desde un punto del borde superior.



8.4.5. Área de una superficie de revolución

Una superficie de revolución se obtiene girando una curva dada alrededor de una recta. Sea $f : [a,b] \to \mathbb{R}$ una función con derivada primera continua. Girando la gráfica de dicha función alrededor del eje OX obtenemos una superficie de revolución, Γ. Fíjate en la siguiente representación.



Sea $S(x)$ el área de la parte de la superficie comprendida entre los planos $X = a$, y $X = x$. Representemos por $L(x)$ la longitud de la gráfica de f entre a y x. Recuerda que $L(x) = \int\limits_a^x \sqrt{1 + f'(t)^2}\,dt$.

Sea $h > 0$. Teniendo en cuenta que el área lateral de un cilindro circular recto es igual a la longitud de la base por la altura, se deduce que

$$2\pi \min\{f(t) : x \leqslant t \leqslant x+h\}(L(x+h) - L(x)) \leqslant S(x+h) - S(x) \leqslant 2\pi \max\{f(t) : x \leqslant t \leqslant x+h\}(L(x+h) - L(x))$$

Por tanto

$$2\pi\,\text{mín}\{f(t): x \leqslant t \leqslant x+h\}\frac{L(x+h)-L(x)}{h} \leqslant \frac{S(x+h)-S(x)}{h} \leqslant 2\pi\,\text{máx}\{f(t): x \leqslant t \leqslant x+h\}\frac{L(x+h)-L(x)}{h}$$

Y tomando límite para $h \to 0$ se sigue que

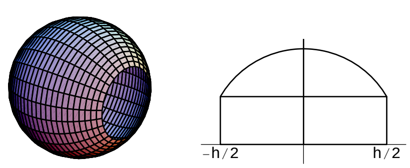
$$S'(x) = 2\pi f(x)L'(x) = 2\pi f(x)\sqrt{1+f'(x)^2}$$

Luego el área de la superficie Γ viene dada por

$$\lambda(\Gamma) = 2\pi\int_a^b f(x)\sqrt{1+f'(x)^2}\,\mathrm{d}x$$

Ejercicios

1. Calcula el área de una superficie esférica de radio R.
2. Calcula el área de la superficie de revolución engendrada al girar la curva $y = x^3$, $0 \leqslant x \leqslant 1$, alrededor del eje OX.
3. Calcula el área de la superficie de revolución engendrada al girar la curva $x^{2/3}+y^{2/3} = a^{2/3}$, $a > 0$, alrededor del eje OX.
4. Calcular el área de la superficie de revolución engendrada al girar la elipse $\dfrac{x^2}{a^2}+\dfrac{y^2}{b^2} = 1$ alrededor del eje OY.
5. Calcular el área de la superficie de revolución engendrada al girar la catenaria $y = \cosh x$, $0 \leqslant x \leqslant 1$, alrededor del eje OX.
6. Al girar alrededor del eje OX el segmento de parábola $y = \sqrt{x}$, $0 \leqslant x \leqslant a$, engendra un tronco de paraboloide de revolución cuya superficie tiene área igual a la de una esfera de radio $\sqrt{13/12}$. Se pide calcular el valor de a.
7. Se perfora una esfera de radio r con un agujero cilíndrico (ver figura) de modo que el anillo esférico resultante tiene altura h. Calcula el volumen del anillo y el área de la superficie total del anillo.



8. Comprueba que el área de la superficie de revolución (llamada horno de Gabriel) engendrada al girar la curva $y = 1/x$, $1 \leqslant x \leqslant +\infty$, alrededor del eje OX es infinita (por tanto sería necesaria una cantidad infinita de pintura si quisiéramos pintarla) pero el volumen del sólido de revolución engendrado es finito(por tanto podemos llenarlo con una cantidad finita de pintura). Comenta a tu gusto esta aparente paradoja.

9. Calcula el área de un espejo parabólico de 3 metros de diámetro y 1 metro de fondo.

10. Calcula el volumen de una esfera de radio 3 en la que, siguiendo un diámetro, se ha perforado un agujero cilíndrico de radio $r < 3$. Calcula el área de la superficie total del solido obtenido. Calcula los valores de r para los que dicha área alcanza sus valores extremos.

Lección 9

Series

Introducción

Las series son una de las herramientas más útiles del Análisis Matemático. Las *series de Fourier* permiten representar señales complejas como superposición de señales sinusoidales, las *series de Taylor* permiten representar funciones analíticas como límites de funciones polinómicas, la transformada z es una *serie de Laurent*. Al igual que la integral, las series son una poderosa herramienta para definir funciones; muchas funciones se definen por medio de series.

En esta lección vamos a estudiar los conceptos y resultados básicos de la teoría de series. Lo único que necesitas para entender lo que sigue es comprender bien el concepto de sucesión convergente al que ya hemos dedicado suficiente atención en la lección correspondiente. Puede ser conveniente que vuelvas a repasarlo antes de seguir.

9.1. Conceptos básicos

En lo que sigue vamos a considerar sucesiones numéricas que pueden ser de números reales o complejos. Naturalmente, todo resultado que se enuncie para números complejos también será válido, en particular, para números reales. En una primera lectura puedes suponer, si te es más cómodo, que se trata de sucesiones de números reales.

9.1 Definición. Dada una sucesión (de números reales o complejos), $\{z_n\}$, podemos formar a partir de ella otra sucesión, $\{S_n\}$, cuyos términos se obtienen *sumando consecutivamente* los términos de $\{z_n\}$, es decir:

$$S_1 = z_1,\ S_2 = z_1 + z_2,\ S_3 = z_1 + z_2 + z_3, \dots,\ S_n = z_1 + z_2 + \cdots + z_n, \dots$$

La sucesión $\{S_n\}$ así obtenida se llama *serie de término general* z_n y es costumbre representarla por $\sum\limits_{n\geqslant 1} z_n$ o, más sencillamente, $\sum z_n$. El número S_n se llama *suma parcial de orden* n de la serie $\sum z_n$.

Debe quedar claro desde ahora que ***una serie es una sucesión cuyos términos se obtienen sumando consecutivamente los términos de otra sucesión***.

Ni que decir tiene que, siendo las series sucesiones, *los conceptos y resultados vistos para sucesiones conservan su misma significación cuando se aplican a series*. En particular, es innecesario volver a definir qué se entiende cuando se dice que una serie es "convergente". Si una serie $\sum_{n\geqslant 1} z_n$ es convergente se usa el símbolo $\sum_{n=1}^{\infty} z_n$ para representar el *límite de la serie* que suele llamarse *suma de la serie*. Naturalmente $\sum_{n=1}^{\infty} z_n$ es el número definido por

$$\sum_{n=1}^{\infty} z_n = \text{lím}\{S_n\} = \lim_{n\to\infty} \sum_{k=1}^{n} z_k$$

Observa que si la serie $\sum z_n$ converge entonces la sucesión $z_n = \sum_{j=1}^{n} z_j - \sum_{j=1}^{n-1} z_j$ es diferencia de dos sucesiones que convergen al mismo límite y por tanto converge a cero.

9.2 Proposición (**Condición necesaria para la convergencia de una serie**)**.** *Para que la serie $\sum z_n$ sea convergente es necesario que* $\text{lím}\{z_n\} = 0$.

9.3 Ejemplo (**Serie geométrica**)**.** Dado $z\in\mathbb{C}$, la sucesión $\{1+z+z^2+\cdots+z^n\}$ se llama serie geométrica de razón z. Observa que dicha serie se obtiene sumando consecutivamente los términos de la sucesión $\{1, z, z^2, z^3, \ldots, z^n, \ldots\}$. Es costumbre representar la serie geométrica de razón z con el símbolo $\sum_{n\geqslant 0} z^n$. Dicha serie converge si, y sólo si, $|z| < 1$, en cuyo caso se verifica que

$$\sum_{n=0}^{\infty} z^n = \frac{1}{1-z}.$$

Todas las afirmaciones hechas se deducen de que si $z \neq 1$, se tiene:

$$\sum_{k=0}^{n} z^k = 1+z+z^2+\cdots+z^n = \frac{1}{1-z} - \frac{z^{n+1}}{1-z} \tag{9.1}$$

si $|z| < 1$ entonces $\lim_{n\to\infty} \dfrac{z^{n+1}}{1-z} = 0$ y obtenemos que

$$\sum_{n=0}^{\infty} z^n = \lim_{n\to\infty} \sum_{k=0}^{n} z^k = \frac{1}{1-z} \qquad (|z| < 1)$$

Si $|z| \geqslant 1$ entonces la sucesión $\{z^n\}$ no converge a 0, por lo que, en virtud de la proposición anterior, deducimos que la serie $\sum_{n\geqslant 0} z^n$ no converge. ♦

9.4 Ejemplo (**Serie armónica**)**.** Se llama así la serie de término general $1/n$; es decir, la serie $\{1+1/2+\cdots+1/n\}$. Se verifica que la serie armónica diverge positivamente

$$\sum_{n=1}^{\infty} \frac{1}{n} = \lim_{n\to\infty} \{1+1/2+\cdots+1/n\} = +\infty$$

En efecto, para todo $n \in \mathbb{N}$ tenemos que

$$\log n = \int_1^n \frac{1}{x}\,dx = \sum_{j=1}^{n-1} \int_j^{j+1} \frac{1}{x}\,dx \leqslant \sum_{j=1}^{n-1} \int_j^{j+1} \frac{1}{j}\,dx = \sum_{j=1}^{n-1} \frac{1}{j} < 1 + \frac{1}{2} + \cdots + \frac{1}{n-1} + \frac{1}{n}$$

y por tanto

$$\lim_{n\to\infty} \{1 + 1/2 + \cdots + 1/n\} \geqslant \lim_{n\to\infty} \log n = +\infty \Longrightarrow \sum_{n=1}^{\infty} \frac{1}{n} = +\infty$$

Este resultado es también consecuencia directa de que, según vimos al estudiar las sucesiones, sabemos que

$$\lim_{n\to\infty} \frac{1 + 1/2 + 1/3 + \cdots + 1/n}{\log n} = 1$$

♦

9.5 Ejemplo (Serie armónica alternada). Se llama así la serie de término general $\dfrac{(-1)^{n-1}}{n}$; es decir, la serie $\displaystyle\sum_{n\geqslant 1} \frac{(-1)^{n-1}}{n}$. Se verifica que la serie armónica alternada es convergente y su suma es igual a $\log 2$.

$$\sum_{n=1}^{\infty} \frac{(-1)^{n-1}}{n} = \log 2$$

En efecto, sustituyendo z por $-x$ en la igualdad (9.1), obtenemos la siguiente igualdad válida para todo $n \in \mathbb{N}$ y todo $x \neq -1$:

$$\frac{1}{1+x} = 1 - x + x^2 - x^3 + \cdots + (-1)^n x^n + (-1)^{n+1} \frac{x^{n+1}}{1+x} \tag{9.2}$$

Integrando esta igualdad entre 0 y 1 tenemos que:

$$\log 2 = 1 - \frac{1}{2} + \frac{1}{3} - \frac{1}{4} + \cdots + (-1)^n \frac{1}{n+1} + (-1)^{n+1} \int_0^1 \frac{x^{n+1}}{1+x}\,dx = \sum_{k=1}^{n} \frac{(-1)^{k-1}}{k} + (-1)^{n+1} \int_0^1 \frac{x^{n+1}}{1+x}\,dx$$

De donde

$$\left| \log 2 - \sum_{k=1}^{n} \frac{(-1)^{k-1}}{k} \right| = \int_0^1 \frac{x^{n+1}}{1+x}\,dx \leqslant \int_0^1 x^{n+1} = \frac{1}{n+2}$$

Y deducimos que

$$\lim_{n\to\infty} \left| \log 2 - \sum_{k=1}^{n} \frac{(-1)^{k-1}}{k} \right| = 0 \Longrightarrow \log 2 = \sum_{n=1}^{\infty} \frac{(-1)^{n-1}}{n}$$

♦

El siguiente ejemplo te ayudará a entender el concepto de serie convergente. Vamos a ver que modificando el orden de los términos en una serie convergente podemos obtener otra serie convergente con distinta suma.

Reordenando términos en la serie armónica alternada podemos obtener otra serie con distinta suma.

Como hemos vista, la serie armónica alternada es la sucesión que se obtiene sumando *consecutivamente* los términos de la sucesión

$$\left\{\frac{(-1)^{n-1}}{n}\right\}=\left\{1,-\frac{1}{2},\frac{1}{3},-\frac{1}{4},\frac{1}{5},-\frac{1}{6},\frac{1}{7},-\frac{1}{8},\frac{1}{9},-\frac{1}{10},\frac{1}{11},-\frac{1}{12},\dots\dots\right\} \tag{9.3}$$

Vamos a cambiar el orden de los términos en esta sucesión poniendo uno positivo seguido de dos negativos manteniendo sus posiciones relativas. Obtenemos así la sucesión

$$\left\{1,-\frac{1}{2},-\frac{1}{4},\frac{1}{3},-\frac{1}{6},-\frac{1}{8},\frac{1}{5},-\frac{1}{10},-\frac{1}{12},\frac{1}{7},-\frac{1}{14},-\frac{1}{16},\dots\dots\right\} \tag{9.4}$$

Cuya serie asociada, obtenida sumando *consecutivamente* sus términos, es la sucesión $\{S_n\}$ dada por:

$$\begin{aligned}
S_1 &= 1\\
S_2 &= 1-\frac{1}{2}\\
S_3 &= 1-\frac{1}{2}-\frac{1}{4}\\
S_4 &= 1-\frac{1}{2}-\frac{1}{4}+\frac{1}{3}\\
S_5 &= 1-\frac{1}{2}-\frac{1}{4}+\frac{1}{3}-\frac{1}{6}\\
S_6 &= 1-\frac{1}{2}-\frac{1}{4}+\frac{1}{3}-\frac{1}{6}-\frac{1}{8}\\
\dots\dots &= \dots\dots\\
S_9 &= 1-\frac{1}{2}-\frac{1}{4}+\frac{1}{3}-\frac{1}{6}-\frac{1}{8}+\frac{1}{5}-\frac{1}{10}-\frac{1}{12}\\
\dots\dots &= \dots\dots\\
S_{3n} &= \sum_{j=1}^{n}\left(\frac{1}{2j-1}-\frac{1}{4j-2}-\frac{1}{4j}\right)
\end{aligned}$$

Tenemos que

$$\begin{aligned}
S_{3n} &= 1-\frac{1}{2}-\frac{1}{4}+\frac{1}{3}-\frac{1}{6}-\frac{1}{8}+\frac{1}{5}-\frac{1}{10}-\frac{1}{12}+\dots\dots+\frac{1}{2n-1}-\frac{1}{4n-2}-\frac{1}{4n}\\
&= \left(1-\frac{1}{2}\right)-\frac{1}{4}+\left(\frac{1}{3}-\frac{1}{6}\right)-\frac{1}{8}+\left(\frac{1}{5}-\frac{1}{10}\right)-\frac{1}{12}+\dots\dots+\left(\frac{1}{2n-1}-\frac{1}{4n-2}\right)-\frac{1}{4n}\\
&= \frac{1}{2}-\frac{1}{4}+\frac{1}{6}-\frac{1}{8}+\frac{1}{10}-\frac{1}{12}+\dots\dots+\frac{1}{2(2n-1)}-\frac{1}{4n}\\
&= \frac{1}{2}\left(1-\frac{1}{2}+\frac{1}{3}-\frac{1}{4}+\frac{1}{5}-\frac{1}{6}+\dots\dots+\frac{1}{2n-1}-\frac{1}{2n}\right)\\
&= \frac{1}{2}\sum_{j=1}^{n}\frac{(-1)^{j-1}}{j}
\end{aligned}$$

Deducimos que

$$\lim_{n\to\infty} S_{3n}=\frac{1}{2}\lim_{n\to\infty}\sum_{j=1}^{n}\frac{(-1)^{j-1}}{j}=\frac{1}{2}\log 2$$

Es claro que $\lim\{S_{3n}-S_{3n-1}\}=\lim\{S_{3n}-S_{3n-2}\}=0$ de donde se sigue que

$$\lim\{S_n\}=\frac{1}{2}\log 2$$

Es decir, hemos probado que la serie obtenida reordenando los términos de la serie armónica alternada por el criterio de sumar uno positivo seguido de dos negativos, es convergente y su suma es $\frac{1}{2}\log 2$.

La *suma* de una serie convergente no es una suma

El ejemplo anterior pone claramente de manifiesto que la *suma* de una serie convergente no es una suma en el sentido usual de la palabra, es decir, no es una suma algebraica de números. Observa que los conjuntos de números (9.3) y (9.4) son los mismos pero las series correspondientes tienen *distinta* suma; la primera tiene *suma* $\log 2$ y la segunda $\frac{1}{2}\log 2$. Si la suma de una serie consistiera en sumar los infinitos términos de una sucesión, entonces el orden en que los sumáramos sería indiferente porque la suma de números tiene la propiedad conmutativa. Debes tener claro, por tanto, que cuando calculas la suma de una serie no estás haciendo una suma infinita sino que estás calculando un *límite de una sucesión* cuyos términos se obtienen sumando consecutivamente los términos de otra sucesión dada. Insisto: calcular la suma de una serie no es una operación algebraica, no consiste en sumar infinitos términos, es un proceso analítico que supone un límite.

La particularidad del estudio de las series

Ahora viene la pregunta del millón: si las series no son nada más que sucesiones, ¿por qué dedicarles una atención especial? La respuesta a esta pregunta es que en el estudio de las series hay una *hipótesis implícita* que los libros silencian. A saber: se supone que las series son sucesiones demasiado difíciles de estudiar *directamente*.

La característica que distingue el estudio de las series es la siguiente: se trata de deducir propiedades de la serie $\{S_n\}=\{z_1+z_2+\cdots+z_n\}$, a partir del comportamiento de $\{z_n\}$. Es decir, los resultados de la teoría de series dan información sobre la sucesión $\{S_n\}$ haciendo hipótesis sobre la sucesión $\{z_n\}$. ¿Por qué esto es así?, ¿no sería más lógico, puesto que lo que queremos es estudiar la serie $\{S_n\}$, hacer hipótesis directamente sobre ella? La razón de esta forma de proceder es que, por lo general, no se conoce una expresión de $S_n=z_1+z_2+\cdots+z_n$ que permita hacer su estudio de forma directa; es decir, la suma $z_1+z_2+\cdots+z_n$ no es posible “realizarla” en la práctica. Por ello, en el estudio de las series se supone implícitamente que *la sucesión* $\{z_n\}$ *es el dato que podemos utilizar*. Naturalmente, esto hace que el estudio de las series se preste a muchas confusiones porque, aunque su objetivo es obtener propiedades de la serie $\{S_n\}$, las hipótesis hacen siempre referencia a la sucesión $\{z_n\}$.

Si lo piensas un poco, esta forma de proceder no es del todo nueva. Ya estás acostumbrado a usar la derivada de una función para estudiar propiedades de la función; pues bien, la situación aquí es parecida: para estudiar la serie $\{z_1+z_2+\cdots+z_n\}$ (la función) estudiamos la sucesión $\{z_n\}$ (la derivada). Un buen ejemplo de esto que digo son los criterios de convergencia que veremos seguidamente.

Otra dificultad adicional en el estudio de las series es la notación tan desafortunada que se emplea. En la mayoría de los textos se representa con el mismo símbolo, $\sum_{n=1}^{\infty} z_n$, la serie (que es una sucesión) y su suma (que es un límite que no siempre existe). Esto es un disparate: se está confundiendo una sucesión con un número. ¿Es lo mismo la sucesión $\{1/n\}$ que el número 0 que es su límite? En ninguna parte verás escrita la igualdad disparatada $\{1/n\} = 0$ ¿Por qué entonces, al tratar con series, se confunde el número $\sum_{n=1}^{\infty} \frac{1}{2^n} = 1$ con $\sum_{k\geqslant 1} \frac{1}{2^k}$ que es la sucesión $\left\{\sum_{k=1}^{n} \frac{1}{2^k}\right\}$?

Quizás esto se debe a que, parece increíble pero es cierto, no hay acuerdo general para representar de forma apropiada la serie de término general z_n. La notación que estamos usando aquí, $\sum_{n\geqslant 1} z_n$, tiene la ventaja de que es clara y evita las confusiones que estoy comentando pero no la verás en los libros. Advertido quedas.

Todavía queda una última sorpresa. Estamos de acuerdo en que las series son sucesiones. ¿Muy especiales? En absoluto. Toda sucesión podemos verla, si así nos interesa, como una serie. Pues toda sucesión $\{z_n\}$ es la serie definida por la sucesión de sus diferencias, esto es, por la sucesión $\{w_n\}$ dada por

$$w_1 = z_1,\, w_2 = z_2 - z_1,\, w_3 = z_3 - z_2, \dots, w_{n+1} = z_{n+1} - z_n, \dots$$

Es claro que $z_n = \sum_{j=1}^{n} w_j$. Por tanto, toda sucesión podemos considerarla como una serie.

Creo que con lo dicho ya puedes hacerte una idea correcta de lo que son las series. Insisto en esto porque en los libros encontrarás disparates para todos los gustos. Algunos textos definen una serie como... ¡un par de sucesiones!, otros dicen que una serie es... ¡una suma infinita! En fin.

9.2. Criterios de convergencia para series de términos positivos

Una serie $\sum a_n$ tal que $a_n \geqslant 0$ para todo $n \in \mathbb{N}$, se dice que es una *serie de términos positivos*. Observa que una serie de términos positivos es una sucesión creciente por lo que o bien es convergente (cuando está mayorada) o es positivamente divergente. Observa el parecido con los criterios de convergencia para integrales de funciones positivas.

9.6 Proposición (Criterio básico de comparación). *Sean $\sum_{n\geqslant 1} a_n$ y $\sum_{n\geqslant 1} b_n$ dos series de términos positivos. Supongamos que hay un número $k \in \mathbb{N}$ tal que $a_n \leqslant b_n$ para todo $n > k$. Entonces se verifica que si la serie $\sum_{n\geqslant 1} b_n$ es convergente, también $\sum_{n\geqslant 1} a_n$ es convergente.*

Demostración. Pongamos $A_n = a_1 + a_2 + \cdots + a_n$, $B_n = b_1 + b_2 + \cdots + b_n$. Las hipótesis hechas implican que para todo $n > k$ es $A_n \leqslant B_n + A_k$. Deducimos que si $\{B_n\}$ está mayorada también lo está $\{A_n\}$. □

9.7 Proposición (Criterio límite de comparación). *Sean* $\sum_{n\geqslant 1} a_n$ *y* $\sum_{n\geqslant 1} b_n$ *dos series de términos positivos, y supongamos que*

$$\lim \frac{a_n}{b_n} = L \in \mathbb{R}_o^+ \cup \{+\infty\}.$$

a) *Si* $L = +\infty$ *y* $\sum_{n\geqslant 1} b_n$ *es divergente también* $\sum_{n\geqslant 1} a_n$ *es divergente.*

b) *Si* $L = 0$ *y* $\sum_{n\geqslant 1} b_n$ *es convergente también* $\sum_{n\geqslant 1} a_n$ *es convergente.*

c) *Si* $L \in \mathbb{R}^+$ *las series* $\sum_{n\geqslant 1} a_n$ *y* $\sum_{n\geqslant 1} b_n$ *son ambas convergentes o ambas divergentes.*

Demostración. Supongamos que $L \in \mathbb{R}^+$. Sea $0 < \alpha < L < \beta$. Todos los términos de la sucesión $\{a_n/b_n\}$, a partir de uno en adelante, están en el intervalo $]\alpha, \beta[$, es decir, existe $k \in \mathbb{N}$ tal que para todo $n \geqslant k$ es $\alpha < a_n/b_n < \beta$, y, por tanto, $\alpha b_n < a_n < \beta b_n$. Concluimos, por el criterio de comparación, que la convergencia de una de las series implica la convergencia de la otra. Queda, así, probado el punto c) del enunciado. Los puntos a) y b) se prueban de manera parecida. □

9.8 Proposición (Criterio integral). *Sea* $f : [1, +\infty[\to \mathbb{R}$ *una función positiva y decreciente. Entonces se verifica que*

$$\sum_{k=2}^{n+1} f(k) \leqslant \int_1^{n+1} f(x)\,\mathrm{d}x \leqslant \sum_{k=1}^{n} f(k)$$

En consecuencia, la serie $\sum_{n\geqslant 1} f(n)$ *y la integral* $\int_1^{+\infty} f(x)\,\mathrm{d}x$ *ambas convergen o ambas divergen.*

Demostración. Por ser f decreciente se tiene que $f(k+1) \leqslant f(x) \leqslant f(k)$ para todo $x \in [k, k+1]$. Integrando, deducimos que

$$f(k+1) \leqslant \int_k^{k+1} f(x)\,\mathrm{d}x \leqslant f(k)$$

Sumando estas desigualdades desde $k = 1$ hasta $k = n$, obtenemos la desigualdad del enunciado. □

Unas series de términos positivos muy útiles para comparar con otras series son las siguientes.

9.9 Proposición (Series de Riemann). *Dado un número real* α*, la serie* $\sum_{n\geqslant 1} \frac{1}{n^\alpha}$ *se llama serie de Riemann de exponente* α*. Dicha serie es convergente si, y sólo si,* $\alpha > 1$.

Demostración. Para que se cumpla la condición necesaria de convergencia es preciso que sea $\alpha > 0$. Supuesto que esto es así, podemos aplicar el criterio integral a la función $f(x) = 1/x^\alpha$ y tener en cuenta los resultados vistos en el ejemplo (8.14). □

Si en el criterio de comparación por paso al límite hacemos $b_n = 1/n^\alpha$, obtenemos el siguiente criterio de convergencia.

9.10 Proposición (**Criterio de Prinsheim**). *Sea* $\sum_{n\geqslant 1} a_n$ *una serie de términos positivos,* α *un número real y supongamos que* $\{n^\alpha a_n\} \to L \in \mathbb{R}_o^+ \cup \{+\infty\}$. *Entonces:*

i) *Si* $L = +\infty$ *y* $\alpha \leqslant 1$, $\sum_{n\geqslant 1} a_n$ *es divergente.*

ii) *Si* $L = 0$ *y* $\alpha > 1$, $\sum_{n\geqslant 1} a_n$ *es convergente.*

iii) *Si* $L \in \mathbb{R}^+$, $\sum_{n\geqslant 1} a_n$ *converge si* $\alpha > 1$ *y diverge si* $\alpha \leqslant 1$.

Vamos a estudiar a continuación dos criterios de convergencia que se aplican a series que pueden compararse con una serie geométrica. Puesto que la serie geométrica de término general $a_n = x^n$, donde $x > 0$, converge si $\dfrac{a_{n+1}}{a_n} = x < 1$, esto nos lleva, en el caso general de una serie términos positivos $\sum_{n\geqslant 1} a_n$, a considerar el comportamiento de la sucesión $\{a_{n+1}/a_n\}$.

9.11 Proposición (**Criterio del cociente o de D'Alembert (1768)**). *Supongamos que* $a_n > 0$ *para todo* $n \in \mathbb{N}$ *y que*

$$\lim \frac{a_{n+1}}{a_n} = L \in \mathbb{R}_o^+ \cup \{+\infty\}.$$

a) *Si* $L < 1$ *la serie* $\sum_{n\geqslant 1} a_n$ *es convergente;*

b) *Si* $L > 1$ *o si* $L = +\infty$, *entonces* $\sum_{n\geqslant 1} a_n$ *es divergente.*

Demostración. a) Sea λ un número tal que $L < \lambda < 1$. La definición de límite implica que existe $n_o \in \mathbb{N}$ tal que para todo $n \geqslant n_o$ se verifica que

$$a_n = \frac{a_n}{a_{n-1}} \frac{a_{n-1}}{a_{n-2}} \cdots \frac{a_{n_o+1}}{a_{n_o}} a_{n_o} \leqslant \lambda^{n-n_o} a_{n_o} = \frac{a_{n_o}}{\lambda^{n_o}} \lambda^n$$

y como, por ser $0 < \lambda < 1$, la serie $\sum_{n\geqslant 1} \lambda^n$ es convergente, deducimos en virtud del criterio de comparación, que $\sum_{n\geqslant 1} a_n$ es convergente.

b) Si $L > 1$ entonces, tomando λ tal que $1 < \lambda < L$ y razonando como antes, obtenemos que para todo $n \geqslant n_o$ es $a_n \geqslant \frac{a_{n_o}}{\lambda^{n_o}} \lambda^n$ y, como, $\lambda > 1$, se sigue que la sucesión $\{a_n\}$ diverge positivamente y, con mayor razón, la serie $\sum_{n\geqslant 1} a_n$ diverge positivamente. □

Análogamente, puesto que la serie geométrica de término general $a_n = x^n$, donde $x > 0$, converge si $\sqrt[n]{a_n} = x < 1$, esto nos lleva, en el caso general de una serie de términos positivos $\sum_{n\geqslant 1} a_n$, a considerar el comportamiento de la sucesión $\{\sqrt[n]{a_n}\}$.

9.12 Proposición (**Criterio de la raíz o de Cauchy (1821)**). *Supongamos que para todo* $n \in \mathbb{N}$ *es* $a_n \geqslant 0$, *y que*

$$\lim \sqrt[n]{a_n} = L \in \mathbb{R}_o^+ \cup \{+\infty\}.$$

a) *Si* $L < 1$ *la serie* $\sum\limits_{n\geqslant 1} a_n$ *es convergente;*

b) *Si* $L > 1$ *o si* $L = +\infty$, *entonces* $\sum\limits_{n\geqslant 1} a_n$ *es divergente.*

Demostración. a) Sea λ un número tal que $L < \lambda < 1$. La definición de límite implica que existe $n_o \in \mathbb{N}$ tal que para todo $n \geqslant n_o$ es $\sqrt[n]{a_n} \leqslant \lambda$, es decir, $a_n \leqslant \lambda^n$. Puesto que $0 < \lambda < 1$, la serie $\sum\limits_{n\geqslant 1} \lambda^n$ es convergente y, en virtud del criterio de comparación, se sigue que $\sum\limits_{n\geqslant 1} a_n$ es convergente.

b) Si $L > 1$ entonces, tomando λ tal que $1 < \lambda < L$ y razonando como antes, obtenemos que para todo $n \geqslant n_o$ es $a_n \geqslant \lambda^n$ y, como, $\lambda > 1$, se sigue que la sucesión $\{a_n\}$ diverge positivamente y, con mayor razón, la serie $\sum\limits_{n\geqslant 1} a_n$ diverge positivamente. □

Cuando $\lim \dfrac{a_{n+1}}{a_n} = 1$, también es $\lim \sqrt[n]{a_n} = 1$. En esta situación los criterios del cociente y de la raíz no proporcionan información suficiente sobre el comportamiento de la serie $\sum\limits_{n\geqslant 1} a_n$. Por ejemplo, para las series de Riemann, $a_n = 1/n^\alpha$, se tiene que $\lim \dfrac{a_{n+1}}{a_n} = 1$ cualquiera sea α. Observa que estos criterios solamente pueden proporcionar información sobre la convergencia de series que pueden compararse con una serie geométrica. El siguiente criterio suele aplicarse cuando fallan los anteriores.

9.13 Proposición (**Criterio de Raabe (1832)**). *Supongamos que* $a_n > 0$ *para todo* $n \in \mathbb{N}$, *y pongamos* $R_n = n\left(1 - \dfrac{a_{n+1}}{a_n}\right)$.

i) *Si* $\{R_n\} \to L$, *donde* $L > 1$ *o* $L = +\infty$, *la serie* $\sum\limits_{n\geqslant 1} a_n$ *es convergente.*

ii) *Si* $\{R_n\} \to L$, *donde* $L < 1$ *o* $L = -\infty$, *o bien si existe algún* $k \in \mathbb{N}$ *tal que* $R_n \leqslant 1$ *para todo* $n \geqslant k$, *entonces la serie* $\sum\limits_{n\geqslant 1} a_n$ *es divergente.*

Demostración. i) Las hipótesis hechas implican que existen $\alpha > 1$ y $n_o \in \mathbb{N}$ tales que para todo $n \geqslant n_o$ es $R_n \geqslant \alpha$. Sea $\delta = \alpha - 1 > 0$. Tenemos que

$$R_n - 1 = (n-1) - n\frac{a_{n+1}}{a_n} \geqslant \delta \qquad (n \geqslant n_o)$$

por lo que

$$a_n \leqslant \frac{1}{\delta}\big((n-1)a_n - na_{n+1}\big) \qquad (n \geqslant n_o).$$

Como $a_n > 0$, deducimos que $na_{n+1} < (n-1)a_n$ para todo $n \geqslant n_o$. Así la sucesión $\{na_{n+1}\}$ es decreciente para $n \geqslant n_o$ y, como es de números positivos, deducimos que es convergente. Sea $\gamma = \lim\{na_{n+1}\} = \inf\{na_{n+1} : n \geqslant n_o\}$. Tenemos que

$$\sum_{j=n_o}^{n} a_n \leqslant \frac{1}{\delta}\big((n_o - 1)a_{n_o} - na_{n+1}\big) \leqslant \frac{1}{\delta}\big((n_o - 1)a_{n_o} - \gamma\big)$$

y, por el criterio general de convergencia para series de términos positivos, deducimos que $\sum\limits_{n\geqslant 1} a_n$ es convergente (lo que, dicho sea de paso, implica que $\gamma = 0$).

ii) Si $R_n \leqslant 1$ para todo $n \geqslant k$, entonces $(n-1)a_n - na_{n+1} \leqslant 0$ y resulta que la sucesión $\{na_{n+1}\}$ es creciente para $n \geqslant k$, luego $na_{n+1} \geqslant ka_{k+1}$, es decir, para todo $n \geqslant k$, es $a_{n+1} \geqslant ka_{k+1}\dfrac{1}{n}$ y, por el criterio de comparación, deducimos que $\sum\limits_{n\geqslant 1} a_n$ es divergente. □

Los criterios de convergencia que acabamos de estudiar hacen siempre hipótesis sobre la sucesión $\{a_n\}$ para obtener información sobre el comportamiento de la serie $\sum\limits_{n\geqslant 1} a_n$. Ya dijimos antes que esto es típico del estudio de las series. Pero no lo olvides: no estamos estudiando la sucesión $\{a_n\}$ sino la sucesión $\sum\limits_{n\geqslant 1} a_n = \{a_1 + a_2 + \cdots + a_n\}$.

Parece que estos criterios son muy restrictivos pues solamente pueden aplicarse a series de términos positivos, pero no es así porque estos criterios pueden aplicarse para estudiar la **convergencia absoluta** de una serie. Precisemos este concepto.

9.14 Definición. Se dice que una serie (de números reales o complejos) $\sum z_n$ es **absolutamente convergente** si la serie de términos positivos

$$\sum_{n\geqslant 1} |z_n| = \{|z_1| + |z_2| + \cdots + |z_n|\}$$

es convergente.

El siguiente resultado, que no demostraremos, es muy importante en el estudio de las series.

9.15 Teorema. *Toda serie absolutamente convergente es convergente.*

Merece la pena explicar esto con detalle. Que la serie $\sum\limits_{n\geqslant 1} z_n$ converge absolutamente quiere decir que es convergente la sucesión

$$\sum_{n\geqslant 1} |z_n| = \{|z_1| + |z_2| + \cdots + |z_n|\}$$

Y el teorema anterior afirma que esto implica la convergencia de la sucesión

$$\sum_{n\geqslant 1} z_n = \{z_1 + z_2 + \cdots + z_n\}$$

¡Son sucesiones muy diferentes!

Naturalmente, si una serie $\{z_1 + z_2 + \cdots + z_n\}$ converge, también converge la sucesión que se obtiene tomando valores absolutos $\{|z_1 + z_2 + \cdots + z_n|\}$ pero esta sucesión ***no es igual*** a $\{|z_1| + |z_2| + \cdots + |z_n|\}$. Por eso *puede ocurrir que una serie sea convergente pero no sea absolutamente convergente*. La serie armónica alternada es un ejemplo de serie convergente que no es absolutamente convergente.

Criterios de convergencia no absoluta

Cuando una serie no es absolutamente convergente se utilizan los siguientes criterios para estudiar su convergencia.

9.16 Teorema. *Sea* $\{a_n\}$ *una sucesión de números reales y* $\{z_n\}$ *una sucesión de números reales o complejos.*

Criterio de Dirichlet. *Si* $\{a_n\}$ *es monótona y converge a cero y la serie* $\sum z_n$ *tiene sumas parciales acotadas, entonces* $\sum a_n z_n$ *converge.*

Criterio de Abel. *Si* $\{a_n\}$ *es monótona y acotada y la serie* $\sum z_n$ *converge, entonces* $\sum a_n z_n$ *es convergente.*

Podemos particularizar el criterio de Dirichlet para *series alternas*, es decir, series del tipo $\sum\limits_{n\geqslant 1}(-1)^{n+1}a_n$ donde $a_n \geqslant 0$ para todo $n\in\mathbb{N}$.

9.17 Proposición (Criterio de Leibnitz). *Si la sucesión* $\{a_n\}$ *es monótona y convergente a cero, entonces la serie* $\sum\limits_{n\geqslant 1}(-1)^{n+1}a_n$ *es convergente.*

Estrategia para estudiar la convergencia de una serie

Para estudiar la convergencia de una serie $\sum z_n$ numérica lo primero que debes hacer es estudiar la convergencia absoluta, es decir la convergencia de la serie de términos positivos $\sum |z_n|$, para lo que se aplican los criterios de convergencia para series de términos positivos. Si la serie $\sum |z_n|$ converge hemos acabado. Cuando la serie $\sum |z_n|$ no converge se aplican los criterios de Dirichlet o de Abel para estudiar directamente la convergencia de la serie $\sum z_n$.

9.2.1. Ejercicios

1. Estudia la convergencia de las siguientes series.

$$\sum_{n\geqslant 1}\log\left(\frac{(n+1)^2}{n(n+2)}\right) \qquad \sum_{n\geqslant 1}\frac{(n+1)^n}{n^{n+2}} \qquad \sum_{n\geqslant 1}n^{-1-1/n}$$

$$\sum_{n\geqslant 1}\frac{(n+1)^n}{3^n n!} \qquad \sum_{n\geqslant 1}\frac{1}{n!}\left(\frac{n}{e}\right)^n \qquad \sum_{n\geqslant 1}\left(\frac{1}{n}-\log(1+1/n)\right)$$

$$\sum_{n\geqslant 1}a^{\log n}\ \ (a>0) \qquad \sum_{n\geqslant 2}\frac{n^{\log n}}{(\log n)^n} \qquad \sum_{n\geqslant 1}\left(e-\left(1+1/n^2\right)^{n^2}\right)$$

$$\sum_{n\geqslant 1}(\sqrt[n]{2}-1)^n \qquad \sum_{n\geqslant 1}\left(1-\frac{1}{\sqrt{n}}\right)^n \qquad \sum_{n\geqslant 1}\left(\frac{n^2+1}{n^2+n+1}\right)^{n^2}$$

$$\sum_{n\geqslant 1}\alpha^{\sum_{j=1}^{n}1/j},\ (\alpha>0) \qquad \sum_{n\geqslant 1}n^{\alpha}\left(\sqrt[n]{n+1/n}-\sqrt[n]{n}\right),\ (\alpha\in\mathbb{R}) \qquad \sum_{n\geqslant 1}\left(1-\frac{\alpha\log n}{n}\right)^n,\ (\alpha\in\mathbb{R})$$

2. Estudia la convergencia de las siguientes series.

$$\sum_{n\geqslant 1}(n^{1/n^2}-1); \qquad \sum_{n\geqslant 1}(\sqrt[3]{n+1}-\sqrt[3]{n})\log\left(\frac{n+1}{n}\right)$$

$$\sum_{n\geqslant 1}\left(\sqrt[4]{n+1}-\sqrt[4]{n}\right)a^{\log n},\ (a>0); \qquad \sum_{n\geqslant 1}\left(\frac{2\cdot 4\cdot 6\cdots(2n)}{5\cdot 7\cdots(2n+3)}\right)^{\alpha},\ (\alpha\in\mathbb{R})$$

$$\sum_{n\geqslant 1}\frac{1}{n}\left(e-(1+1/n)^n\right); \qquad \sum_{n\geqslant 1}\left(n^{n^{\alpha}}-1\right),\ (\alpha\in\mathbb{R})$$

$$\sum_{n\geqslant 1}n^{\alpha}\left(1+\frac{1}{2}+\cdots+\frac{1}{n}\right),\ (\alpha\in\mathbb{R}); \qquad \sum_{n\geqslant 1}n^{\alpha}\exp\Big(-\beta\sum_{k=1}^{n}\frac{1}{k}\Big),\ (\alpha,\beta\in\mathbb{R})$$

3. Estudia la convergencia de las series.

a) $\displaystyle\sum_{n\geqslant 1}\frac{3^n n!}{\sqrt[3]{n}\,5\cdot 8\cdot 11\cdots(5+3n)}$

b) $\displaystyle\sum_{n\geqslant 1}\left(\frac{2\cdot 3\cdots(n+2)}{5\cdot 6\cdots(n+5)}\right)^{1/2}$

c) $\displaystyle\sum_{n\geqslant 1}(2-\sqrt{2})(2-\sqrt[3]{2})\cdots(2-\sqrt[n]{2})$

d) $\displaystyle\sum_{n\geqslant 1}\frac{n!}{a(a+1)(a+2)\cdots(a+n)n^{\alpha}},\ (a>0,\alpha\in\mathbb{R})$

e) $\displaystyle\sum_{n\geqslant 1}\log(1+1/n)a^{\log n},\ (a>0)$

f) $\displaystyle\sum_{n\geqslant 1}(\sqrt{n+1}-\sqrt{n})^{\alpha}(\log(1+1/n))^{\beta},\ (\alpha,\beta\in\mathbb{R})$

g) $\displaystyle\sum_{n\geqslant 1}\left(\frac{(1+\alpha)(3+\alpha)(5+\alpha)\cdots(2n-1+\alpha)}{(2+\beta)(4+\beta)(6+\beta)\cdots(2n+\beta)}\right)^{\rho},\ (\alpha,\beta,\rho\in\mathbb{R}^+)$

4. Estudia la convergencia de las series:

i) $\displaystyle\sum_{n\geqslant 0}\frac{1}{(1+i)^n}$ ii) $\displaystyle\sum_{n\geqslant 1}\frac{\cos n+i\,\operatorname{sen} n}{n}$

iii) $\displaystyle\sum_{n\geqslant 1}\frac{\cos n+i\,\operatorname{sen} n}{n^2}$ iv) $\displaystyle\sum_{n\geqslant 1}\frac{\cos\frac{\pi}{n}+i\,\operatorname{sen}\frac{\pi}{n}}{n}$

v) $\displaystyle\sum_{n\geqslant 1}\frac{(2+i)^n}{(1+2i)^n}\frac{1}{n}$ vi) $\displaystyle\sum_{n\geqslant 1}\frac{1}{\sqrt{n}}\left(\frac{1+i\sqrt{3}}{2}\right)^n$

vii) $\displaystyle\sum_{n\geqslant 1}\left(\cos\frac{\pi}{n^2}+i\,\operatorname{sen}\frac{\pi}{n^2}\right)$ viii) $\displaystyle\sum_{n\geqslant 0}\frac{(3+4i)^n}{2i(4+3i)^n+7}$

5. Sea $\rho\in\mathbb{R}$ con $|\rho|<1$ y $\vartheta\in\mathbb{R}$. Calcula los límites $\displaystyle\sum_{n=0}^{\infty}\rho^n\cos(n\vartheta)$ y $\displaystyle\sum_{n=0}^{\infty}\rho^n\operatorname{sen}(n\vartheta)$.

6. Estudia la convergencia absoluta y la convergencia no absoluta de las series:

$$\sum_{n\geqslant 1}(-1)^{n+1}\frac{\log(n+2)}{n+2} \qquad \sum_{n\geqslant 1}(-1)^{n}\frac{1}{n^{\alpha}+(-1)^{n}},\ (\alpha\in\mathbb{R})$$
$$\sum_{n\geqslant 1}(-1)^{n+1}\frac{2+(-1)^{n}}{n} \qquad \sum_{n\geqslant 1}(-1)^{n+1}\frac{\sqrt{n}}{n+100}$$
$$\sum_{n\geqslant 1}\log\left(1+\frac{(-1)^{n+1}}{n}\right) \qquad \sum_{n\geqslant 1}(-1)^{n+1}\left(\frac{1\cdot 3\cdot 5\cdots(2n-1)}{2\cdot 4\cdot 6\cdots 2n}\right)^{\alpha},\ (\alpha\in\mathbb{R})$$

9.3. Series de potencias

9.18 Definición. Dadas una sucesión de números complejos $\{c_n\}_{n\in\mathbb{N}_o}$ y un número $a\in\mathbb{C}$, se llama **serie de potencias** centrada en a a la sucesión

$$\{c_o+c_1(z-a)+c_2(z-a)^2+c_3(z-a)^3+\cdots+c_n(z-a)^n\}$$

en la cual z representa un número complejo arbitrario. Dicha serie se representa simbólicamente por $\sum_{n\geqslant 0}c_n(z-a)^n$. La sucesión $\{c_n\}$ recibe el nombre de sucesión de coeficientes de la serie.

Dados $a\in\mathbb{C}$ y $r>0$, definimos $D(a,r)=\{z\in\mathbb{C}:|z-a|<r\}$. El conjunto $D(a,r)$ se llama disco abierto de centro a y radio r.

El primer problema que plantea una serie de potencias, $\sum_{n\geqslant 0}c_n(z-a)^n$, es estudiar para qué valores de z dicha serie es convergente. El siguiente resultado es clave a este respecto.

9.19 Lema (Lema de Abel). *Sea $\rho>0$ un número positivo tal que la sucesión $\{c_n\rho^n\}$ está acotada. Entonces para todo $z\in D(a,\rho)$ se verifica que la serie $\sum_{n\geqslant 0}c_n(z-a)^n$ converge absolutamente.*

Demostración. Sea $z\in D(a,\rho)$. Pongamos $r=|z-a|<\rho$. Puesto que $\{c_n\rho^n\}$ está acotada existe una constante $M>0$ tal que $|c_n\rho^n|\leqslant M$ para todo número natural n. Tenemos que

$$|c_n(z-a)^n|=\left|c_n\rho^n\frac{(z-a)^n}{\rho^n}\right|=|c_n\rho^n|\,\frac{|z-a|^n}{\rho^n}\leqslant M\left(\frac{|z-a|}{\rho}\right)^n=M\left(\frac{r}{\rho}\right)^n$$

Pongamos $\{\alpha_n\}=\{M(r/\rho)^n\}$. Como la serie $\sum_{n\geqslant 0}\alpha_n$ converge por ser una serie geométrica de razón r/ρ menor que 1, el criterio de comparación nos dice que la serie $\sum_{n\geqslant 0}|c_n(z-a)^n|$ converge, es decir, la serie $\sum_{n\geqslant 0}c_n(z-a)^n$ es absolutamente convergente. □

El lema anterior conduce de manera natural a considerar el más grande $\rho>0$ tal que la sucesión $\{a_n\rho^n\}$ está acotada. Introducimos para ello el conjunto

$$A=\{\rho\geqslant 0:\{c_n\rho^n\}\text{ está acotada}\}$$

y definimos $R=\sup(A)\in\mathbb{R}_0^+\cup\{+\infty\}$. Pueden presentarse los casos:

- $R=0$. Entonces $A=\{0\}$, luego si $z\neq a$ la sucesión $\{c_n(z-a)^n\}$ no está acotada y, en particular, no converge a cero. Por tanto la serie de potencias $\sum_{n\geqslant 0} c_n(z-a)^n$ no converge. En este caso se dice que *la serie de potencias es trivial* pues solamente converge para $z=a$.
- $0<R<+\infty$. En este caso la serie $\sum_{n\geqslant 0} c_n(z-a)^n$ es absolutamente convergente para todo $z\in D(a,R)$. Para $|z-a|>R$ se tiene que la sucesión $\{c_n(z-a)^n\}$ no está acotada y, por tanto, la serie no converge.

  Nada puede afirmarse en general del comportamiento de la serie en la frontera del disco $D(a,R)$.
- Si $R=+\infty$ la serie converge absolutamente para todo $z\in\mathbb{C}$.

Al número R se le llama *radio de convergencia* de la serie. Nótese que R sólo depende de la sucesión $\{c_n\}$ de coeficientes de la serie y no del centro de la serie. Dada una serie de potencias no trivial, $\sum c_n(z-a)^n$, llamaremos *dominio de convergencia de la serie* al conjunto $D(a,R)$ donde R es el radio de convergencia de la serie (naturalmente, si $R=+\infty$ entonces $D(a,+\infty)=\mathbb{C}$).

El siguiente resultado, consecuencia directa de aplicar los criterios del cociente o de la raíz a la serie $\sum_{n\geqslant 0} |c_n(z-a)^n|$, permite en muchos casos calcular el radio de convergencia.

9.20 Proposición. *Supongamos que se cumple alguna de las igualdades siguientes:*

- $\lim \dfrac{|c_{n+1}|}{|c_n|} = L \in \mathbb{R}_0^+ \cup \{+\infty\}$
- $\lim \{\sqrt[n]{|c_n|}\} = L \in \mathbb{R}_0^+ \cup \{+\infty\}$

Entonces $R=1/L$ *con los convenios:* $R=0$ *si* $L=+\infty$ *y* $R=+\infty$ *si* $L=0$.

El siguiente resultado es muy útil para calcular la suma de una serie de potencias en puntos de la frontera de su disco de convergencia.

9.21 Teorema (Continuidad radial). *Sea* $\sum c_n z^n$ *una serie de potencias con radio de convergencia* $0<R<+\infty$. *Sea* $f:D(0,R)\to\mathbb{C}$ *la función suma de la serie y supongamos que la serie es convergente en un punto* z_0 *de la frontera del disco* $D(0,R)$. *Entonces se verifica que*

$$\lim_{\substack{r\to 1\\ 0<r<1}} f(rz_0) = \sum_{n=0}^{\infty} c_n z_0^n$$

En lo que sigue vamos a estudiar la derivabilidad de funciones definidas por series de potencias, por ello supondremos que se trata de series de potencias reales, es decir, los coeficientes de la serie son números reales y la variable, que representaremos por la letra x, solamente puede tomar valores reales.

Consideremos, pues, una serie de potencias reales $\sum_{n\geqslant 0} c_n(x-a)^n$ con radio de convergencia $R>0$. El dominio de convergencia de esta serie es el intervalo $\Omega=\{x\in\mathbb{R} : |x-a|<R\} =]a-R,a+R[$, con el convenio de que dicho intervalo es todo $\mathbb{R}$ cuando el radio de convergencia es $R=+\infty$. Como consecuencia del estudio realizado, sabemos que la serie es convergente absolutamente para todo $x\in]a-R,a+R[$ y, por tanto, es convergente para todo $x\in]a-R,a+$

$R[$. Podemos definir la *función suma de la serie*, $f:]a-R,a+R[\to\mathbb{R}$ que viene dada para cada $x\in]a-R,a+R[$ por

$$f(x)=\sum_{n=0}^{\infty}c_n(x-a)^n$$

Observa que $f(x)$ viene dada como un límite de una sucesión de funciones polinómicas. Cabe esperar que las propiedades de las funciones polinómicas se "hereden" de alguna forma por f.

9.22 Teorema (Teorema de derivación de una serie de potencias). *Sea $a\in\mathbb{R}$, $\sum_{n\geqslant 0}c_n(x-a)^n$ una serie de potencias no trivial y Ω su dominio de convergencia. Sea $f:\Omega\to\mathbb{R}$ la función suma de la serie*

$$f(x)=\sum_{n=0}^{\infty}c_n(x-a)^n \qquad x\in\Omega$$

Entonces f es indefinidamente derivable en Ω y, para cada $k\in\mathbb{N}$ su derivada k-ésima se obtiene derivando k veces la serie término a término, esto es:

$$f^{(k)}(x)=\sum_{n=k}^{\infty}n(n-1)\cdots(n-k+1)c_n(x-a)^{n-k} \qquad x\in\Omega$$

En particular $f^{(k)}(a)=k!\,c_k$ o, lo que es lo mismo

$$c_k=\frac{f^{(k)}(a)}{k!} \qquad \textit{para todo } k\in\mathbb{N}\cup\{0\}$$

9.23 Definición. Sea f una función indefinidamente derivable en un intervalo $\Omega\subset\mathbb{R}$ y sea $a\in\Omega$. La serie de potencias

$$\sum_{n\geqslant 0}\frac{f^{(n)}(a)}{n!}(x-a)^n$$

se llama *serie de Taylor* de f en el punto a.

Acabamos de ver en el teorema de derivación que si $\sum_{n\geqslant 0}c_n(x-a)^n$ es una serie de potencias no trivial, y f es su función suma, entonces se verifica que $c_k=\dfrac{f^{(k)}(a)}{k!}$ para todo $k=0,1,2,\dots$. Esto nos dice que la serie de partida *es* la serie de Taylor en el punto a de su función suma.

9.24 Corolario. *Las únicas series de potencias no triviales son series de Taylor (de su función suma).*

Convergencia de las series de Taylor

Las siguientes preguntas surgen de manera inmediata: ¿Son no triviales las series de Taylor de una función indefinidamente derivable? ¿La función suma de una serie de Taylor de una función indefinidamente derivable coincide con dicha función? La respuesta a ambas preguntas es que no. Si partimos de una *función real* indefinidamente derivable y formamos su serie de Taylor en un punto, no hay garantía de que dicha serie tenga radio de convergencia no nulo o de que, en caso de que sea convergente en un intervalo, su suma sea igual a la función de partida.

La herramienta básica que permite estudiar la convergencia de una serie de Taylor es el teorema de Taylor. Supongamos que $f : I \to \mathbb{R}$ es una función indefinidamente derivable en un intervalo abierto I y sea $a \in I$. La serie de Taylor de f en a viene dada por $\displaystyle\sum_{n\geqslant 0} \frac{f^{(n)}(a)}{n!}(x-a)^n$, es decir, se trata de la sucesión $\displaystyle\left\{\sum_{k=0}^{n} \frac{f^{(k)}(a)}{k!}(x-a)^k\right\}$, pero $T_n(f,a)(x) = \displaystyle\sum_{k=0}^{n} \frac{f^{(k)}(a)}{k!}(x-a)^k$ es el polinomio de Taylor de orden n de f en a. Es decir, *la serie de Taylor de f en a es la sucesión de los polinomios de Taylor de f en a*. El teorema de Taylor afirma que dado $x \in I$ y $n \in \mathbb{N}$ hay un punto c (que dependerá de x y de n) comprendido entre x y a de manera que se verifica la igualdad

$$f(x) - T_n(f,a)(x) = \frac{f^{(n+1)}(c)}{(n+1)!}(x-a)^{n+1} \tag{9.5}$$

En esta igualdad el lado de la derecha parece mucho más manejable que el de la izquierda. Si somos capaces de probar que

$$\lim_{n\to\infty} \frac{f^{(n+1)}(c)}{(n+1)!}(x-a)^{n+1} = 0 \tag{9.6}$$

habremos probado que $\displaystyle\lim_{n\to\infty}(f(x) - T_n(f,a)(x)) = 0$, esto es $f(x) = \displaystyle\lim_{n\to\infty} T_n(f,a)(x)$, es decir, habremos probado que para ese valor de x la serie de Taylor de f en a converge a $f(x)$. La dificultad está en que en el límite (9.6) el punto c no está fijo: depende de n (además de x, pero el punto x está fijo en este razonamiento). El que se verifique la igualdad en (9.6) dependerá del comportamiento de las derivadas sucesivas de f. El caso más sencillo de todos es cuando dichas derivadas están acotadas en el intervalo I, es decir, hay un número $M > 0$ (independiente de n) tal que se verifica

$$\left|f^{(n+1)}(t)\right| \leqslant M \qquad \forall t \in I,\ \forall n \in \mathbb{N} \tag{9.7}$$

Cuando esto es así, como $\lim_{n\to\infty} \dfrac{(x-a)^{n+1}}{(n+1)!} = 0$ cualquiera sea $x \in \mathbb{R}$, se deduce que la serie de Taylor de f centrada en a es convergente en todo punto $x \in I$ y su suma es igual a $f(x)$.

Desarrollos en serie de las funciones elementales

Función exponencial.

Las derivadas de la función exponencial natural, $\exp(x)$, en $x = 0$ valen todas 1. Por tanto, el polinomio de Taylor de orden n en el punto $a = 0$ de la función exponencial es

$$T_n(\exp,0)(x) = 1 + \sum_{k=1}^{n} \frac{1}{k!}x^k$$

Sea ahora $x \in \mathbb{R}$, fijo en lo que sigue. Tomemos $\rho > 0$ de forma que $x \in]-\rho,\rho[$. El teorema de Taylor afirma que hay un número c comprendido entre x y 0 y, por tanto $c \in]-\rho,\rho[$, tal que

$$\exp(x) = 1 + \sum_{k=1}^{n} \frac{1}{k!}x^k + \frac{\exp(c)}{(n+1)!}x^{n+1}$$

Como $\exp(c) < \exp(\rho)$ deducimos que

$$\left|\exp(x) - 1 - \sum_{k=1}^{n} \frac{1}{k!}x^k\right| < \exp(\rho)\frac{x^{n+1}}{(n+1)!}$$

De esta forma hemos eliminado el punto c y como $\lim\limits_{n\to\infty} \dfrac{x^{n+1}}{(n+1)!} = 0$, obtenemos que

$$\exp(x) = \sum_{k=0}^{\infty} \frac{1}{k!} x^k \tag{9.8}$$

En particular, para $x = 1$ tenemos que $\mathrm{e} = \sum\limits_{n=0}^{\infty} \dfrac{1}{n!}$.

Funciones seno y coseno.

Como $\operatorname{sen}'(x) = \cos(x) = \operatorname{sen}(\frac{\pi}{2} + x)$, se sigue que $\operatorname{sen}^{(n)}(x) = \operatorname{sen}(\frac{n\pi}{2} + x)$. En particular, $\operatorname{sen}^{(n)}(0) = \operatorname{sen}(\frac{n\pi}{2})$. Por tanto el polinomio de Taylor de orden n de la función seno en el punto $a = 0$ es

$$T_n(\operatorname{sen}, 0)(x) = \sum_{k=1}^{n} \frac{\operatorname{sen}(\frac{k\pi}{2})}{k!} x^k$$

Como para k par es $\operatorname{sen}(\frac{k\pi}{2}) = 0$ y para k impar $k = 2q - 1$ es $\operatorname{sen}(\frac{(2q-1)\pi}{2}) = (-1)^{q+1}$, resulta que

$$T_{2n-1}(\operatorname{sen}, 0)(x) = T_{2n}(\operatorname{sen}, 0)(x) = \sum_{k=1}^{n} \frac{(-1)^{k+1}}{(2k-1)!} x^{2k-1}$$

Análogamente para la función coseno

$$T_{2n}(\cos, 0)(x) = T_{2n+1}(\cos, 0)(x) = \sum_{k=0}^{n} \frac{(-1)^k}{(2k)!} x^{2k}$$

Como las funciones seno y coseno tienen derivadas acotadas en todo $\mathbb{R}$, se sigue por lo antes visto que

$$\operatorname{sen} x = \sum_{n=0}^{\infty} \frac{(-1)^n}{(2n+1)!} x^{2n+1} \qquad (x \in \mathbb{R}) \tag{9.9}$$

$$\cos x = \sum_{n=0}^{\infty} \frac{(-1)^n}{(2n)!} x^{2n} \qquad (x \in \mathbb{R}) \tag{9.10}$$

Serie binomial de Newton.

Pongamos $f(x) = (1+x)^{\alpha}$. Tenemos que $f^{(n)}(x) = \alpha(\alpha-1)(\alpha-2)\cdots(\alpha-n+1)(1+x)^{\alpha-n}$. Por lo que el polinomio de Taylor de orden n de f en $a = 0$ es

$$T_n(f, 0)(x) = 1 + \sum_{k=1}^{n} \frac{\alpha(\alpha-1)(\alpha-2)\cdots(\alpha-k+1)}{k!} x^k$$

Cualquiera sea el *número real* α y el número natural k se define

$$\binom{\alpha}{k} = \frac{\alpha(\alpha-1)(\alpha-2)\cdots(\alpha-k+1)}{k!}$$

Por convenio $\binom{\alpha}{0} = 1$. Con ello podemos escribir

$$T_n(f, 0)(x) = \sum_{k=0}^{n} \binom{\alpha}{k} x^k$$

Queremos probar la igualdad

$$(1+x)^{\alpha} = \sum_{n=0}^{\infty} \binom{\alpha}{n} x^n \qquad x \in]-1,1[\tag{9.11}$$

En este caso el teorema de Taylor no es el camino más sencillo y es preferible razonar como sigue.

La serie de potencias $\sum_{n\geqslant 0} \binom{\alpha}{n} x^n$ tiene radio de convergencia 1 porque, poniendo $a_n = \binom{\alpha}{n}$, tenemos que $\dfrac{|a_{n+1}|}{|a_n|} = \dfrac{|\alpha - n|}{n+1} \to 1$. Definamos $\varphi :]-1,1[\to \mathbb{R}$ por

$$\varphi(x) = \sum_{n=0}^{\infty} \binom{\alpha}{n} x^n, \qquad x \in]-1,1[$$

El teorema de derivación nos dice que φ es derivable en $]-1,1[$. Ahora es fácil comprobar que la función $h(x) = \varphi(x)(1+x)^{-\alpha}$ tiene derivada nula en $]-1,1[$, por lo que es constante en dicho intervalo y, como $h(0)=1$, concluimos que $h(x)=1$ para todo $x \in]-1,1[$, es decir, $(1+x)^{\alpha} = \varphi(x)$ para todo $x \in]-1,1[$, como queríamos probar.

identidades que se deducen de la serie geométrica

Partiendo de la serie geométrica

$$\sum_{n=0}^{\infty} x^n = \frac{1}{1-x} \qquad |x| < 1 \tag{9.12}$$

podemos derivarla para obtener

$$\sum_{n=1}^{\infty} n x^{n-1} = \frac{1}{(1-x)^2} \qquad |x| < 1$$

Igualdad que, a su vez, podemos volver a derivar y obtenemos

$$\sum_{n=2}^{\infty} n(n-1) x^{n-2} = \frac{2}{(1-x)^3} \qquad |x| < 1$$

Así podemos seguir. Pero también podemos invertir el proceso: en vez de derivar la serie (9.12), podemos construir una primitiva suya.

El cálculo de una primitiva de una función definida por una serie de potencias es inmediato. Si la función viene dada por

$$f(x) = \sum_{n=0}^{\infty} c_n (x-a)^n \qquad |x-a| < R$$

Entonces, en virtud del teorema de derivación, la primitiva de f que coincide con f en el punto $x = a$ es

$$F(x) = f(a) + \sum_{n=0}^{\infty} \frac{c_n}{n+1} (x-a)^{n+1} \qquad |x-a| < R$$

De la igualdad (9.12), se deduce ahora que

$$-\log(1-x)=\sum_{n=0}^{\infty}\frac{1}{n+1}x^{n+1}\qquad |x|<1$$

Y, cambiando x por $-x$ en esta igualdad, obtenemos

$$\log(1+x)=\sum_{n=0}^{\infty}\frac{(-1)^n}{n+1}x^{n+1}\qquad |x|<1$$

Definición de las funciones trigonométricas

Las series de potencias proporcionan la herramienta adecuada para definir con comodidad la función seno. Olvida ahora todo lo que sabes de la función seno. ¿Lo has olvidado ya? Sigamos.

Como la serie de potencias $\sum_{n\geqslant 0}\frac{(-1)^n}{(2n+1)!}x^{2n+1}$ tiene radio de convergencia $R=+\infty$ la suma de dicha serie está definida en todo $\mathbb{R}$.

9.25 Definición. La función seno es la función $\operatorname{sen}:\mathbb{R}\to\mathbb{R}$ definida para todo $x\in\mathbb{R}$ por

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

A partir de esta definición se obtienen las propiedades usuales de la función seno. La función coseno se define como la derivada de la función seno. Las demás funciones trigonométricas se definen, como puedes imaginar, a partir del seno y el coseno. El número π se define como el mínimo número positivo en el que se anula el seno (hay que probar su existencia con ayuda del teorema de Bolzano).

No es mi propósito proseguir este estudio sino solamente dejarte indicado el camino.

Ejercicios

1. Estudia la convergencia las siguientes series de potencias.

$$\textbf{a)}\ \sum_{n\geqslant 1}\frac{z^n}{n!}\qquad \textbf{b)}\ \sum_{n\geqslant 1}\frac{(n+1)^n}{n^{n+1}}z^n\qquad \textbf{c)}\ \sum_{n\geqslant 1}n^{\alpha}z^n$$

$$\textbf{d)}\ \sum_{n\geqslant 1}\frac{n^n}{n!}z^n\qquad \textbf{e)}\ \sum_{n\geqslant 1}\frac{3\cdot 5\cdots(3n+1)}{5\cdot 10\cdots 5n}z^n\qquad \textbf{f)}\ \sum_{n\geqslant 1}\frac{z^n}{1+1/2+\cdots+1/n}$$

   Estudia en los casos c)y f), el comportamiento de la serie en los puntos de la circunferencia unidad.

2. Dada la serie de potencias

$$\sum_{n\geqslant 1}\left(n^2+\frac{1}{n}\right)(x-1)^n$$

   Calcula su radio de convergencia. Estudia la convergencia en los extremos del intervalo de convergencia. Calcula su suma.

3. Calcula el desarrollo de Taylor en $a=0$ de ña función
$$f(x)=\frac{x^2-3x+1}{x^2-5x+6}$$
e indica su dominio de convergencia.

   Sugerencia. Usa la descomposición en fracciones simples.

4. Sea la serie de potencias $\sum_{n\geqslant 0}(2^{n+1}-n)x^n$. Calcula su dominio de convergencia y su suma.

5. Sea la serie de potencias $\sum_{n\geqslant 0}\left(\frac{(-1)^n}{(n+1)2^n}-\frac{n}{3^n}\right)x^n$. Calcula su dominio de convergencia y su suma.

6. Calcula, por medio de series de potencias primitivas de las funciones $\operatorname{sen}(x^2)$, $\exp(x^2)$, $\sqrt{1+x^3}$.

Lección 10

Cálculo diferencial en $\mathbb{R}^n$

10.1. Estructura euclídea y topología de $\mathbb{R}^n$

Como sabes, $\mathbb{R}^n$ es un espacio vectorial en el que suele destacarse la llamada base canónica formada por los vectores $\{\mathbf{e_1}, \mathbf{e_2}, \dots, \mathbf{e_n}\}$ donde $\mathbf{e_k}$ es el vector cuyas componentes son todas nulas excepto la que ocupa el lugar k que es igual a 1. Dados dos vectores $\mathbf{x} = (x_1, x_2, \dots, x_n)$ $\mathbf{y} = (y_1, y_2, \dots, y_n)$ se define su producto escalar a por:

$$\langle \mathbf{x} \,|\, \mathbf{y} \rangle = \sum_{k=1}^{n} x_k y_k = x_1 y_1 + x_2 y_2 + \cdots + x_n y_n$$

Este producto escalar se llama *producto escalar euclídeo*. Observa que el producto escalar de dos vectores no es un vector sino un número real. La notación $\mathbf{x \dot y}$ es frecuentemente usada en los libros de Física para representar el producto escalar de los vectores $\mathbf{x}$ e $\mathbf{y}$. Las dos notaciones son útiles y las usaremos en lo que sigue.

Las siguientes propiedades del producto escalar se deducen fácilmente de la definición:

- $\langle \mathbf{x} \,|\, \mathbf{y} \rangle = \langle \mathbf{y} \,|\, \mathbf{x} \rangle$ para todos $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$ (simetría).
- $\langle \alpha\mathbf{x} + \beta\mathbf{y} \,|\, \mathbf{z} \rangle = \alpha \langle \mathbf{x} \,|\, \mathbf{z} \rangle + \beta \langle \mathbf{y} \,|\, \mathbf{z} \rangle$ para todos $\alpha, \beta \in \mathbb{R}$ y para todos $\mathbf{x}, \mathbf{y}, \mathbf{z} \in \mathbb{R}^n$ (linealidad).

La norma euclídea de un vector $\mathbf{x}$ se define por

$$\|\mathbf{x}\| = \sqrt{\langle \mathbf{x} \,|\, \mathbf{x} \rangle} = \sqrt{\sum_{k=1}^{n} x_k^2}$$

Desigualdad de Cauchy-Schwarz.

Para todos $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$ se verifica que $|\langle \mathbf{x} \,|\, \mathbf{y} \rangle| \leqslant \|\mathbf{x}\| \, \|\mathbf{y}\|$. Además, supuesto que $\mathbf{x}$ e $\mathbf{y}$ no son nulos, la igualdad $|\langle \mathbf{x} \,|\, \mathbf{y} \rangle| = \|\mathbf{x}\| \, \|\mathbf{y}\|$ equivale a que hay un número $\lambda \in \mathbb{R}$ tal que $\mathbf{x} = \lambda \mathbf{y}$ (es decir, los vectores $\mathbf{x}$ e $\mathbf{y}$ están en una misma recta que pasa por el origen).

Desigualdad triangular.

Para todos $\mathbf{x},\mathbf{y}\in\mathbb{R}^n$ *se verifica que* $\|\mathbf{x}+\mathbf{y}\| \leqslant \|\mathbf{x}\|+\|\mathbf{y}\|$. *Además, supuesto que* $\mathbf{x}$ *e* $\mathbf{y}$ *no son nulos, la igualdad* $\|\mathbf{x}+\mathbf{y}\| = \|\mathbf{x}\|+\|\mathbf{y}\|$ *equivale a que hay un número* $\lambda>0$ *tal que* $\mathbf{x}=\lambda\mathbf{y}$ *(es decir, los vectores* $\mathbf{x}$ *e* $\mathbf{y}$ *están en una misma semirrecta que pasa por el origen).*

10.1 Definición. Se dice que los vectores $\mathbf{x}$ e $\mathbf{y}$ son **ortogonales**, y escribimos $\mathbf{x}\perp\mathbf{y}$, cuando su producto escalar es cero. Se dice que un vector $\mathbf{x}$ es ortogonal a un conjunto de vectores $E\subset\mathbb{R}^n$ cuando $\mathbf{x}$ es ortogonal a todo vector en E. Un conjunto de vectores no nulos que son mutuamente ortogonales se dice que es un **conjunto ortogonal** de vectores; si, además, los vectores tienen todos norma 1 se dice que es un **conjunto ortonormal** de vectores. Una base vectorial que también es un conjunto ortogonal (ortonormal) se llama una **base ortogonal** (ortonormal).

Si $\mathbf{x}$ e $\mathbf{y}$ son vectores no nulos, el vector

$$\prod_{\mathbf{y}}(\mathbf{x}) = \frac{\langle \mathbf{x}\,|\,\mathbf{y}\rangle}{\langle \mathbf{y}\,|\,vy\rangle}\mathbf{y}$$

se llama **proyección ortogonal de x sobre y**.

Puedes comprobar que el vector $\mathbf{x}-\prod_{\mathbf{y}}(\mathbf{x})$ es ortogonal a $\mathbf{y}$. En particular, si $\mathbf{y}$ es un **vector unitario** (de norma 1) entonces el vector $\mathbf{x}-\langle \mathbf{x}\,|\,\mathbf{y}\rangle\mathbf{y}$ es ortogonal a $\mathbf{y}$.

Ejercicios

1. Prueba la desigualdad de Cauchy-Schwarz.

   Sugerencia. Comprueba que la ecuación $\langle \mathbf{x}-\lambda\mathbf{y}\,|\,\mathbf{x}-\lambda\mathbf{y}\rangle=0$, en la que λ es un número real arbitrario y $\mathbf{x}$ e $\mathbf{y}$ son vectores que se suponen fijos, es un trinomio de segundo grado en la variable λ. Ten en cuenta que dicho trinomio toma siempre valores mayores o iguales que cero (£por qué?) lo que proporciona información sobre su discriminante.

2. Prueba la desigualdad triangular.

   Sugerencia. Una estrategia para probar desigualdades entre normas euclídeas es elevar al cuadrado. La desigualdad $\|\mathbf{x}+\mathbf{y}\|^2 \leqslant \big(\|\mathbf{x}\|+\|\mathbf{y}\|\big)^2$ es equivalente a la desigualdad triangular pero es muy fácil de probar desarrollando el término $\|\mathbf{x}+\mathbf{y}\|^2=\langle \mathbf{x}+\mathbf{y}\,|\,\mathbf{x}+\mathbf{y}\rangle$ y usando la desigualdad de Cauchy-Schwarz.

3. **Teorema de Pitágoras**. Prueba que los vectores $\mathbf{x}$ e $\mathbf{y}$ son ortogonales si, y solo si, $\|\mathbf{x}+\mathbf{y}\|^2=\|\mathbf{x}\|^2+\|\mathbf{y}\|^2$.

4. Prueba que el vector $\mathbf{x}-\prod_{\mathbf{y}}(\mathbf{x})$ es ortogonal a $\mathbf{y}$.

10.2 Definición. Dados dos vectores $\mathbf{x}$ e $\mathbf{y}$, el número $\|\mathbf{x}-\mathbf{y}\|$ se llama la **distancia** (euclídea) entre $\mathbf{x}$ e $\mathbf{y}$.

- Dados $\mathbf{x}\in\mathbb{R}^n$ y $r>0$, definimos $B(\mathbf{x},r)=\{\mathbf{y}\in\mathbb{R}^n : \|x-y\|<r\}$.

- Un conjunto $E\subset\mathbb{R}^n$ se dice que es un **conjunto abierto** si para todo punto $\mathbf{x}\in E$ se verifica que hay un número $r_{\mathbf{x}}>0$ tal que $B(\mathbf{x},r_{\mathbf{x}})\subset E$. Por convenio, el conjunto vacío, $\emptyset$, se considera abierto.

- Es fácil comprobar que los conjuntos de la forma $B(\mathbf{x}, r)$ son conjuntos abiertos. El conjunto $B(\mathbf{x}, r)$ se llama **bola abierta** de centro $\mathbf{x}$ y radio r.

- Un conjunto $F \subset \mathbb{R}^n$ se dice que es un **conjunto cerrado** si su complemento $\mathbb{R}^n \setminus F$ es un conjunto abierto.

- Dados $\mathbf{x} \in \mathbb{R}^n$ y $r \geqslant 0$, definimos $\overline{B}(\mathbf{x}, r) = \{\mathbf{y} \in \mathbb{R}^n : \|x - y\| \leqslant r\}$. Es fácil comprobar que $\overline{B}(\mathbf{x}, r)$ es un conjunto cerrado. Se llama **bola cerrada** de centro $\mathbf{x}$ y radio r.

- Se dice que un conjunto $E \subset \mathbb{R}^n$ es **acotado** cuando hay un número $M > 0$ tal que $\|\mathbf{x}\| \leqslant M$ para todo $\mathbf{x} \in E$.

- Se dice que un conjunto $K \subset \mathbb{R}^n$ es **compacto** cuando es cerrado y acotado.

- Sea $E \subset \mathbb{R}^n$. Decimos que un punto $\mathbf{x} \in \mathbb{R}^n$ es **adherente** al conjunto E si toda bola abierta centrada en $\mathbf{x}$ tiene puntos de E. El conjunto de todos los puntos adherentes a E se llama la **adherencia** de E y se representa por $\overline{E}$.

- Sea $E \subset \mathbb{R}^n$. Decimos que un punto $\mathbf{x} \in \mathbb{R}^n$ es un punto de **acumulación** del conjunto E si toda bola abierta centrada en $\mathbf{x}$ tiene puntos de E *distintos* de $\mathbf{x}$. El conjunto de todos los puntos de acumulación de E se llama la **acumulación** de E y se representa por E'.

- Sea $E \subset \mathbb{R}^n$. El conjunto de todos los puntos adherentes a E y a $\mathbb{R}^n \setminus E$ se llama la **frontera** de E y se representa por $\operatorname{Fr}(()E)$.

- Sea $E \subset \mathbb{R}^n$. Decimos que un punto $\mathbf{x} \in E$ es un punto **interior** al conjunto E si hay alguna bola abierta centrada en $\mathbf{x}$ contenida en E.

- Dados $\mathbf{x} \in \mathbb{R}^n$ y $r > 0$, el conjunto $S(\mathbf{x}, r) = \{\mathbf{y} \in \mathbb{R}^n : \|x - y\| = r\}$ se llama **esfera** de centro $\mathbf{x}$ y radio r.

- Representaremos por Π_j la aplicación $\Pi_j : \mathbb{R}^n \to \mathbb{R}$ que a cada vector $\mathbf{x} = (x_1, x_2, \ldots, x_n) \in \mathbb{R}^n$ hace corresponder su coordenada j-ésima en la base canónica.

$$\Pi_j(\mathbf{x}) = \Pi_j((x_1, x_2, \ldots, x_n)) = x_j$$

Las aplicaciones Π_j, $1 \leqslant j \leqslant n$, así definidas se llaman las **proyecciones canónicas**.

Ejercicios

1. Prueba que $B(\mathbf{x}, r)$ es un conjunto abierto.
2. Prueba que todo conjunto abierto es unión de bolas abiertas.
3. Prueba que la intersección de dos conjuntos abiertos es un conjunto abierto.
4. Prueba que la unión de conjuntos abiertos es un conjunto abierto.
5. Prueba que $\overline{B}(\mathbf{x}, r)$ es un conjunto cerrado.
6. Prueba que la intersección de conjuntos cerrados es un conjunto cerrado.
7. Da ejemplos de conjuntos que no sean abiertos ni cerrados.
8. Prueba que $\overline{E} = E \cup \operatorname{Fr}(()E)$.

10.1.1. Sucesiones en $\mathbb{R}^n$

10.3 Definición. Una sucesión $\{\mathbf{x}_m\}$ de puntos de $\mathbb{R}^n$ se dice que es **convergente** si hay un vector $\mathbf{x}\in\mathbb{R}^n$ tal que $\|\mathbf{x}_m-\mathbf{x}\|\to 0$. En tal caso escribimos $\lim_{m\to\infty}\{\mathbf{x}_m\}=\mathbf{x}$ o, simplemente, $\{\mathbf{x}_m\}\to\mathbf{x}$ y decimos que $\mathbf{x}$ es el **límite** de la sucesión $\{\mathbf{x}_m\}$.

Una sucesión $\{\mathbf{x}_m\}$ de puntos de $\mathbb{R}^n$ se dice que es **acotada** si ha un número $M>0$ tal que $\|\mathbf{x}_m\|\leqslant M$ para todo $m\in\mathbb{N}$.

Teniendo en cuenta la desigualdad

$$\text{máx}\{|x_k-y_k| : 1\leqslant k\leqslant n\}\leqslant\|\mathbf{x}-\mathbf{y}\|\leqslant\sum_{k=1}^{n}|x_k-y_k| \tag{10.1}$$

Se deduce fácilmente que $\{\mathbf{x}_m\}\to\mathbf{x}$ si, y sólo si, $\{\Pi_j(\mathbf{x}_m)\}\to\Pi_j(\mathbf{x})$ para $(1\leqslant j\leqslant n)$, esto es, **la convergencia en $\mathbb{R}^n$ equivale a la convergencia por coordenadas**.

10.4 Teorema (Teorema de Bolzano – Weierstrass). *Toda sucesión acotada de puntos de $\mathbb{R}^n$ tiene alguna sucesión parcial convergente.*

10.5 Teorema (Caracterización de los conjuntos compactos). *Un conjunto $E\subset\mathbb{R}^n$ es compacto si, y sólo si, toda sucesión de puntos de E tiene alguna sucesión parcial que converge a un punto de E.*

10.1.2. Campos escalares. Continuidad y límite funcional

Reciben el nombre de *campos escalares* las funciones definidas en subconjuntos de $\mathbb{R}^n$ que toman valores en $\mathbb{R}$. Un campo escalar es, por tanto, una función real que depende de n variables.

Un campo escalar de una variables es, simplemente, una función real de variable real; un campo escalar de dos variables es una función definida en un subconjunto del plano que toma valores reales; un campo escalar de tres variables es una función definida en un subconjunto del espacio que toma valores reales.

Los campos escalares de una o dos variables se pueden visualizar por medio de sus representaciones gráficas que son, respectivamente, curvas en el plano o superficies en el espacio. No es posible visualizar campos escalares de tres o más variables porque sus gráficas están en espacios de dimensión mayor o igual que 4.

Naturalmente, los campos escalares se pueden sumar y multiplicar al igual que lo hacemos con las funciones reales.

10.6 Definición. Sea f un campo escalar definido en un conjunto $E\subset\mathbb{R}^n$ y sea $\mathbf{a}\in E$. Se dice que f es **continuo** en $\mathbf{a}$ si para todo $\varepsilon>0$ existe un $\delta>0$ tal que se verifica $\|f(\mathbf{x})-f(\mathbf{a})\|<\varepsilon$ siempre que $\mathbf{x}\in E$ y $\|\mathbf{x}-\mathbf{a}\|<\varepsilon$.

Se dice que f es continuo en un conjunto $A\subset E$ si f es continuo en todo punto $\mathbf{a}\in A$.

Un ejemplo de campo escalar continuo lo proporcionan las proyecciones canónicas Π_j pues se tiene que

$$\left|\Pi_j(\mathbf{x})-\Pi_j(\mathbf{y})\right|=\left|x_j-y_j\right|\leqslant\|\mathbf{x}-\mathbf{y}\|$$

de donde se deduce enseguida la continuidad de Π_j.

10.7 Proposición. *a) Si f y g son campos escalares definidos en un conjunto $E \subset \mathbb{R}^n$, se verifica que los campos escalares $f+g$ y fg son continuos en todo punto de E donde f y g sean continuos. Y si f no se anula en E, el campo escalar $1/f$ es continua en todo punto de E donde f sea continuo.*

b) Sea f un campo escalar definido en un conjunto $E \subset \mathbb{R}^n$ y sea h una función real de variable real continua definida en un intervalo I que contiene la imagen de f, $I \supset f(E)$. Entonces la función compuesta $h \circ g$ es continua en todo punto de E donde f sea continua.

Los campos escalares más sencillos son las funciones polinómicas de varias variables. Dichas funciones se obtienen como sumas de productos de las proyecciones canónicas y son, por tanto, continuas.

Para $n=3$ las proyecciones canónicas son

$$\Pi_1((x,y,z)) = x, \quad \Pi_2((x,y,z)) = y, \quad \Pi_3((x,y,z)) = z$$

Un producto de estas funciones es una función de la forma $f(x,y,z) = x^m y^p z^q$ donde m, p, q son números naturales o nulos. Las funciones polinómicas en tres variables son sumas de este tipo de funciones.

Las funciones racionales de n variables son las funciones de la forma

$$R(x_1, x_2, \dots, x_n) = \frac{P(x_1, x_2, \dots, x_n)}{Q(x_1, x_2, \dots, x_n)}$$

Donde $P(x_1, x_2, \dots, x_n)$ y $Q(x_1, x_2, \dots, x_n)$ son funciones polinómicas de n variables. El dominio natural de definición de una función racional es el conjunto de puntos donde no se anula el denominador $\Omega = \{\mathbf{x} \in \mathbb{R}^n : Q(\mathbf{x}) \neq 0\}$. Las funciones racionales son continuas en su conjunto natural de definición.

Componiendo funciones continuas reales de una variable con funciones polinómicas y racionales en varias variables obtenemos muchísimos ejemplos de campos escalares continuos. Aquí tienes unos pocos.

$$f(x,y) = \operatorname{sen}(xy), \quad f(x,y) = \log(1+x^2+y^2), \quad f(x,y,z) = \frac{1+xy^2+xz^2}{1+\operatorname{arc\,tg}(xyz)}, \quad f(x,y,z) = \cos(\sqrt{y^2+z^2})$$

El siguiente resultado establece la relación entre la continuidad y el límite secuencial.

10.8 Proposición. *Sea Sea f un campo escalar definido en un conjunto $E \subset \mathbb{R}^n$ y sea $\mathbf{a} \in E$. Equivalen las siguientes afirmaciones:*

a) f es continua en $\mathbf{a}$.

b) Para toda sucesión $\{\mathbf{x}_n\}$ de puntos de E tal que $\{\mathbf{x}_n\} \to \mathbf{a}$ se verifica que $\{f(\mathbf{x}_n)\} \to f(\mathbf{a})$.

El siguiente resultado se demuestra de la misma forma que su análogo para funciones reales.

10.9 Teorema (Teorema de Weierstrass). *Todo campo escalar continuo en un conjunto compacto alcanza en dicho conjunto un valor máximo absoluto y un valor mínimo absoluto.*

Dicho de otra forma, si $K \subset \mathbb{R}^n$ es un conjunto compacto y f es un campo escalar continuo en K, entonces hay puntos $\mathbf{a} \in K$, $\mathbf{b} \in K$ tales que $f(\mathbf{c}a) \leqslant f(\mathbf{x}) \leqslant f(\mathbf{b})$ para todo $\mathbf{x} \in K$.

10.10 Definición. Sea f un campo escalar definido en un conjunto $E \subset \mathbb{R}^n$ y sea $\mathbf{a} \in E'$. Se dice que f **tiene límite en a** si hay un número $L \in \mathbb{R}$ con la propiedad de que para todo $\varepsilon > 0$ existe un $\delta > 0$ tal que se verifica $\|f(\mathbf{x}) - L\| < \varepsilon$ siempre que $\mathbf{x} \in E$ y $0 < \|\mathbf{x} - \mathbf{a}\| < \varepsilon$. Simbólicamente escribimos $\lim\limits_{\mathbf{x} \to \mathbf{a}} f(\mathbf{x}) = L$. El número L se llama *límite* de f en $\mathbf{a}$.

El siguiente resultado establece la relación entre el límite funcional y el límite secuencial.

10.11 Proposición. *Sea Sea f un campo escalar definido en un conjunto $E \subset \mathbb{R}^n$ y sea $\mathbf{a} \in E'$. Equivalen las siguientes afirmaciones:*

a) $\lim\limits_{\mathbf{x} \to \mathbf{a}} f(\mathbf{x}) = L$.

b) Para toda sucesión $\{\mathbf{x}_n\}$ de puntos de E distintos *de* $\mathbf{a}$, *tal que $\{\mathbf{x}_n\} \to \mathbf{a}$ se verifica que $\{f(\mathbf{x}_n)\} \to L$.*

10.1.3. Curvas en $\mathbb{R}^n$

Una curva en $\mathbb{R}^n$ es una aplicación continua $\gamma : [a,b] \to \mathbb{R}^n$. El punto $\gamma(a)$ se llama *origen* y el punto $\gamma(b)$ *extremo* de la curva. Naturalmente, como $\gamma(t) \in \mathbb{R}^n$ podremos expresarlo por medio de sus componentes en la base canónica que serán funciones de t.

$$\gamma(t) = (\gamma_1(t), \gamma_2(t), \ldots, \gamma_n(t))$$

Las funciones $\gamma_k(t)$ se llaman funciones componentes de γ. Se dice que γ es derivable en un punto t cuando todas sus funciones componentes son derivables en dicho punto, en cuyo la derivada de γ en t es, por definición, el vector

$$\gamma'(t) = (\gamma_1'(t), \gamma_2'(t), \ldots, \gamma_n'(t))$$

Dado un punto $\mathbf{a} = \gamma(t_0)$ tal que $\gamma'(t_0) \neq \mathbf{0}$, se define la **recta tangente** a γ en el punto $\mathbf{a}$ (aunque es más apropiado decir *en el punto* t_0) como la recta de ecuación paramétrica $\mathbf{a} + t\gamma'(t_0)$, es decir, la recta que pasa por $\mathbf{a}$ con vector de dirección $\gamma'(t_0)$.

Cuando se interpreta $\gamma(t)$ como la función de trayectoria de un móvil, entonces su **velocidad** en un instante t es el vector $\gamma'(t)$ y su **rapidez** es $\|\gamma'(t)\|$. La distancia que recorre dicho móvil entre dos instantes $t = a$ y $t = b$ viene dada por

$$\int_a^b \|\gamma'(t)\| \, \mathrm{d}t \, .$$

10.12 Definición. Un conjunto abierto $\Omega \subset \mathbb{R}^n$ con la propiedad de que cualesquiera dos de sus puntos pueden unirse por una curva que queda dentro de Ω se llama un **dominio**.

Intuitivamente, un dominio es un conjunto abierto *de un solo trozo*. Los dominios desempeñan en $\mathbb{R}^n$ un papel similar al de los intervalos en $\mathbb{R}$.

10.1.4. Derivadas parciales. Vector gradiente

Acabamos de ver que los conceptos de continuidad y límite para funciones reales de una variable se generalizan fácilmente para campos escalares de varias variables. No ocurre lo mismo con el concepto de derivabilidad el cual no puede generalizarse de forma inmediata. La razón es que el concepto de derivabilidad hace intervenir la división de números reales, pues una derivada es un límite de cocientes incrementales, y en $\mathbb{R}^n$ no podemos dividir por vectores, es decir, la estructura algebraica de $\mathbb{R}^n$ no permite generalizar algo parecido a un "cociente incremental". Si f es un campo escalar, la expresión

$$\frac{f(\mathbf{x})-f(\mathbf{a})}{\mathbf{x}-\mathbf{a}}$$

no tiene ningún sentido.

Otra diferencia importante es que en la recta real, $\mathbb{R}$, solamente podemos acercarnos a un punto de ella a través de la propia recta, mientras que en $\mathbb{R}^n$ para $n \geqslant 2$ hay muchísimas más posibilidades de acercarse a un punto dado; por ejemplo, podemos acercarnos a través de cualquier curva que pase por dicho punto. Surge así una primera idea que consiste en acercarse a un punto dado a través de una recta dada. Parece que esta situación es más parecida a lo que conocemos para funciones reales de una variable.

10.13 Definición. Una **dirección** en $\mathbb{R}^n$ es un vector de norma 1.

- Dados un punto $\mathbf{a}\in\mathbb{R}^n$ y una dirección $\mathbf{u}$, la recta que pasa por $\mathbf{a}$ con dirección $\mathbf{u}$ es la imagen de la aplicación $\gamma\colon\mathbb{R}^n\to\mathbb{R}$ dada por $\gamma(t)=\mathbf{a}+t\mathbf{u}$, es decir, es el conjunto de puntos $\{\mathbf{a}+t\mathbf{u} : t\in\mathbb{R}\}$.

- Sea f un campo escalar definido en un conjunto abierto $E\subset\mathbb{R}^n$, sea $\mathbf{a}\in E$ y $\mathbf{u}$ una dirección. Se define la **derivada de f en a en la dirección u** como el límite

$$D_{\mathbf{u}} f(a) = \lim_{t\to 0}\frac{f(\mathbf{a}+t\,\mathbf{u})-f(\mathbf{a})}{t} \tag{10.2}$$

supuesto, claro está, que dicho límite exista.

- Las derivada direccional de un campo escalar f en un punto $\mathbf{a}$ en la dirección del vector $\mathbf{e_k}$ de la base canónica, se llama derivada parcial de f en $\mathbf{a}$ respecto a la variable k-ésima. Está definida por

$$\begin{aligned} D_{\mathbf{e_k}} f(a) &= \lim_{t\to 0}\frac{f(\mathbf{a}+t\,\mathbf{e_k})-f(\mathbf{a})}{t} = \lim_{t\to 0}\frac{f(a_1,\dots,a_k+t,\dots,a_n)-f(a_1,\dots,a_k,\dots,a_n)}{t} \\ &= \lim_{x_k\to a_k}\frac{f(a_1,\dots,x_k,\dots,a_n)-f(a_1,\dots,a_k,\dots,a_n)}{x_k-a_k} \end{aligned} \tag{10.3}$$

y se representa con los símbolos $D_k f(\mathbf{a})$ y $\dfrac{\partial f}{\partial x_k}(\mathbf{a})$.

Observa que las derivadas que acabamos de definir son derivadas de funciones reales de una variable real pues, para calcular la derivada de un campo escalar f en un punto $\mathbf{a}$ en la dirección $\mathbf{u}$ lo que se hace es derivar en $t=0$ la función $t\mapsto f(\mathbf{a}+t\,\mathbf{u})$ que es una función real de una variable real.

Observa que la segunda igualdad de (10.3) nos dice que, *para calcular la derivada parcial $D_k f(\mathbf{a})$, lo que se hace es derivar f respecto a la variable k-ésima considerando fijas las demás variables*. Por eso se llaman derivadas *parciales*.

Interpretación geométrica de las derivadas parciales

Es importante que entiendas el significado de las derivadas parciales de una función en un punto. Para poder visualizarlo vamos a considerar un campo escalar f de dos variables definido en $E \subset \mathbb{R}^2$. Fijemos un punto (a,b). Las derivadas parciales de f en (a,b) son, por definición

$$\begin{aligned} D_1 f(a,b) &= \lim_{t\to 0}\frac{f(a+t,b)-f(a,b)}{t} = \lim_{x\to a}\frac{f(x,b)-f(a,b)}{x-a} \\ D_2 f(a,b) &= \lim_{t\to 0}\frac{f(a,b+t)-f(a,b)}{t} = \lim_{y\to b}\frac{f(a,y)-f(a,b)}{y-b} \end{aligned}$$

Es decir, lo que hacemos es derivar las funciones parciales $x \mapsto f(x,b)$ y $y \mapsto f(a,y)$ en los puntos $x=a$ e $y=b$ respectivamente.

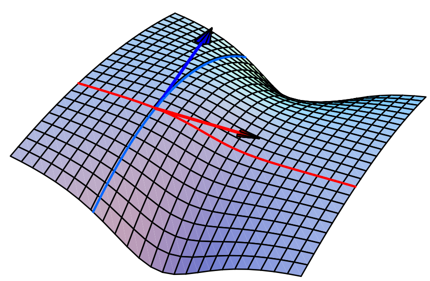
La gráfica de f, es decir, el conjunto $S = \{(x,y,f(x,y)) : (x,y)\in E\}$ es una superficie en $\mathbb{R}^3$. Las funciones

$$\gamma_1(x) = (x,b,f(x,b)), \quad \gamma_2(y) = (a,y,f(a,y))$$

son curvas contenidas en dicha superficie que pasan por el punto (a,b). Dichas curvas se obtienen cortando la superficie S por los planos $y=b$ y $x=a$ respectivamente. Los vectores tangentes a dichas curvas en los puntos $\gamma_1(a)$ y $\gamma_2(b)$ son, respectivamente

$$\gamma_1{}'(a) = (1,0,D_1 f(a,b)), \quad \gamma_2{}'(b) = (0,1,D_2 f(a,b))$$

En la siguiente figura se ha representado la gráfica de f y las curvas obtenidas cortándola por los planos $x=a$ e $y=b$ junto a sus vectores tangentes en el punto (a,b)



Cuando un campo escalar f tiene derivadas parciales en todos los puntos de un conjunto $E \subset \mathbb{R}^n$, podemos definir las *funciones derivadas parciales* de f, $D_k f : E \to \mathbb{R}$ que a cada punto $\mathbf{x}\in E$ hace corresponder el número $D_k f(\mathbf{x})$. Dichas funciones son también campos escalares.

10.14 Definición. Sea f un campo escalar. Se define el **vector gradiente** de f en un punto $\mathbf{a}$ como el vector

$$\nabla f(\mathbf{a}) = \big(D_1 f(\mathbf{a}), D_2 f(\mathbf{a}), \dots, D_n f(\mathbf{a})\big)$$

supuesto, claro está, que dichas derivadas parciales existan.

Supongamos que f es una función real de una variable real. La derivabilidad de f en un punto $a\in\mathbb{R}$ se expresa por

$$\lim_{x\to a}\frac{f(x)-f(a)}{x-a} = f'(a) \Longleftrightarrow \lim_{x\to a}\frac{f(x)-f(a)-f'(a)(x-a)}{x-a} = 0$$

Recuerda que la recta de ecuación cartesiana $y = f(a) + f'(a)(x-a)$ es la recta tangente a la gráfica de f en el punto $(a, f(a))$.

Si ahora f es un campo escalar definido en un conjunto $E \subset \mathbb{R}^n$, cuyo vector gradiente $\nabla f(\mathbf{a})$ está definido en un punto $a \in E$, podemos considerar el hiperplano en $\mathbb{R}^{n+1}$ de ecuación cartesiana $x_{n+1} = f(\mathbf{a}) + \langle \nabla f(\mathbf{a}) \,|\, \mathbf{x} - \mathbf{a} \rangle$. Este hiperplano pasa por el punto $(\mathbf{a}, f(\mathbf{a})) \in \mathbb{R}^{n+1}$ y es la generalización natural de la recta tangente a la gráfica de una función. Observa el parecido formal entre las expresiones

$$y = f(a) + f'(a)(x-a), \qquad x_{n+1} = f(\mathbf{a}) + \langle \nabla f(\mathbf{a}) \,|\, \mathbf{x} - \mathbf{a} \rangle$$

Ambas representan hiperplanos (un hiperplano en $\mathbb{R}^2$ es una recta) y la segunda se deduce de la primera sustituyendo la derivada por el vector gradiente y el producto usual de números reales por el producto escalar de vectores. Esto nos lleva a la siguiente definición.

10.15 Definición. Sea f un campo escalar definido en un conjunto $E \subset \mathbb{R}^n$ y sea $\mathbf{a}$ un punto interior de E. Supongamos que está definido el vector gradiente $\nabla f(\mathbf{a})$. Se dice que f es **diferenciable** en $\mathbf{a}$ si se verifica que

$$\lim_{\mathbf{x} \to \mathbf{a}} \frac{f(\mathbf{x}) - f(\mathbf{a}) - \langle \nabla f(\mathbf{a}) \,|\, \mathbf{x} - \mathbf{a} \rangle}{\|\mathbf{x} - \mathbf{a}\|} = 0 \tag{10.4}$$

Definamos

$$R(\mathbf{x}, \mathbf{a}) = \frac{f(\mathbf{x}) - f(\mathbf{a}) - \langle \nabla f(\mathbf{a}) \,|\, \mathbf{x} - \mathbf{a} \rangle}{\|\mathbf{x} - \mathbf{a}\|}$$

La igualdad (10.4) dice que $\lim\limits_{\mathbf{x} \to \mathbf{a}} R(\mathbf{x}, \mathbf{a}) = 0$. Con lo que, otra forma equivalente de escribir la igualdad (10.4) es la siguiente.

$$f(\mathbf{x}) = f(\mathbf{a}) + \langle \nabla f(\mathbf{a}) \,|\, \mathbf{x} - \mathbf{a} \rangle + R(\mathbf{x}, \mathbf{a}) \,\|\mathbf{x} - \mathbf{a}\| \quad \text{donde} \quad \lim_{\mathbf{x} \to \mathbf{a}} R(\mathbf{x}, \mathbf{a}) = 0 \tag{10.5}$$

10.16 Definición. Sea f un campo escalar diferenciable en un punto $\mathbf{a}$. El hiperplano en $\mathbb{R}^{n+1}$ de ecuación cartesiana

$$x_{n+1} = f(\mathbf{a}) + \langle \nabla f(\mathbf{a}) \,|\, \mathbf{x} - \mathbf{a} \rangle$$

se llama hiperplano tangente a f en $\mathbf{a}$ o **hiperplano tangente** a la gráfica de f en el punto $(\mathbf{a}, f(\mathbf{a}))$.

10.17 Proposición. *Sea f un campo escalar diferenciable en un punto $\mathbf{a}$ y sea $\mathbf{u}$ una dirección en $\mathbb{R}^n$. Entonces se verifica que*

$$D_{\mathbf{u}} f(\mathbf{a}) = \langle \nabla f(\mathbf{a}) \,|\, \mathbf{u} \rangle$$

Demostración. En la igualdad (10.5) pongamos $\mathbf{x} = \mathbf{a} + t\,\mathbf{u}$ con lo que obtenemos

$$f(\mathbf{a} + t\,\mathbf{u}) \;=\; f(\mathbf{a}) + \langle \nabla f(\mathbf{a}) \,|\, t\,\mathbf{u} \rangle + R(\mathbf{a} + t\,\mathbf{u}, \mathbf{a}) \,\|t\,\mathbf{u}\| = f(\mathbf{a}) + t \,\langle \nabla f(\mathbf{a}) \,|\, \mathbf{u} \rangle + R(\mathbf{a} + t\,\mathbf{u}, \mathbf{a}) \,|t| \Longrightarrow$$

$$\lim_{t \to 0} \frac{f(\mathbf{a} + t\,\mathbf{u}) - f(\mathbf{a})}{t} = \lim_{t \to 0} R(\mathbf{a} + t\,\mathbf{u}, \mathbf{a}) \frac{|t|}{t} = 0 \qquad \square$$

10.18 Corolario. *Sea f un campo escalar diferenciable en un punto $\mathbf{a}$ con vector gradiente no nulo en $\mathbf{a}$.*

a) La dirección en la que la derivada direccional de f en a es máxima es la dirección dada por el gradiente, es decir, la dirección $\mathbf{u} = \dfrac{\nabla f(\mathbf{a})}{\|\nabla f(\mathbf{a})\|}$.

b) La dirección en la que la derivada direccional de f en a es mínima es la dirección opuesta a la dada por el gradiente, es decir, la dirección $\mathbf{v} = -\dfrac{\nabla f(\mathbf{a})}{\|\nabla f(\mathbf{a})\|}$.

Demostración. Las afirmaciones hechas son consecuencia de la proposición anterior y de la desigualdad de Cauchy–Schwarz, pues para toda dirección $\mathbf{w}$ se tiene que

$$\left|\left\langle \nabla f(\mathbf{a}) \,\middle|\, \mathbf{w} \right\rangle\right| \leqslant \|\nabla f(\mathbf{a})\| \, \|\mathbf{w}\| = \|\nabla f(\mathbf{a})\|$$

Y la igualdad se da si, y solo si, hay un número $\lambda \in \mathbb{R}$ tal que $\mathbf{w} = \lambda \nabla f(\mathbf{a})$. Tomando normas en esta igualdad se deduce que $|\lambda| = 1/\|\nabla f(\mathbf{a})\|$, es decir las direcciones $\mathbf{w}$ que hacen máximo $|D_{\mathbf{w}} f(\mathbf{a})| = \left|\left\langle \nabla f(\mathbf{a}) \,\middle|\, \mathbf{w} \right\rangle\right|$ son $\mathbf{u} = \dfrac{\nabla f(\mathbf{a})}{\|\nabla f(\mathbf{a})\|}$ y $\mathbf{v} = -\dfrac{\nabla f(\mathbf{a})}{\|\nabla f(\mathbf{a})\|}$.

Para la primera se tiene que

$$D_{\mathbf{u}} f(\mathbf{a}) = \left\langle \nabla f(\mathbf{a}) \,\middle|\, \frac{\nabla f(\mathbf{a})}{\|\nabla f(\mathbf{a})\|} \right\rangle = \frac{1}{\|\nabla f(\mathbf{a})\|} \left\langle \nabla f(\mathbf{a}) \,\middle|\, \nabla f(\mathbf{a}) \right\rangle = \|\nabla f(\mathbf{a})\|$$

que es el valor máximo que puede tener una derivada direccional.

Análogamente, para la segunda se tiene que

$$D_{\mathbf{v}} f(\mathbf{a}) = -\|\nabla f(\mathbf{a})\|$$

que es el valor mínimo que puede tener una derivada direccional. □

El resultado anterior nos dice que el vector gradiente en un punto señala la dirección en la que el campo tiene máximo crecimiento en dicho punto. Mientras que en la dirección opuesta a la del vector gradiente en un punto el campo tiene máximo decrecimiento.

10.19 Proposición. *Sean f un campo escalar definido en un conjunto $E \subset \mathbb{R}^n$ y γ una curva en $\mathbb{R}^n$ que toma valores en el conjunto E. Supongamos que γ es derivable en un punto t_0 y que f es diferenciable en el punto $\mathbf{a} = \gamma(t_0) \in E$. Entonces se verifica que la función $h(t) = f(\gamma(t))$ es derivable en t_0 y su derivada viene dada por*

$$h'(t_0) = \left\langle \nabla f(\mathbf{a}) \,\middle|\, \gamma'(t_0) \right\rangle = \sum_{k=1}^{n} D_k f(\mathbf{a}) \gamma_k{}'(t_0) \tag{10.6}$$

Demostración. Se tiene que

$$h(t) - h(t_0) = f(\gamma(t)) - f(\gamma(t_0)) = \left\langle \nabla f(\mathbf{a}) \,\middle|\, \gamma(t) - \gamma(t_0) \right\rangle + R(\gamma(t), \gamma(t_0)) \, \|\gamma(t) - \gamma(t_0)\|$$

Dividiendo por $t - t_0$ tenemos

$$\frac{h(t) - h(t_0}{t - t_0} = \frac{f(\gamma(t)) - f(\gamma(t_0))}{t - t_0} = \left\langle \nabla f(\mathbf{a}) \,\middle|\, \frac{\gamma(t) - \gamma(t_0)}{t - t_0} \right\rangle + R(\gamma(t), \gamma(t_0)) \frac{\|\gamma(t) - \gamma(t_0)\|}{t - t_0}$$

Teniendo en cuenta que $\displaystyle\lim_{t \to t_0} \frac{\gamma(t) - \gamma(t_0)}{t - t_0} = \gamma'(t_0)$ se deduce que

$$\lim_{t \to t_0} \frac{h(t) - h(t_0)}{t - t_0} = \left\langle \nabla f(\mathbf{a}) \,\middle|\, \gamma'(t_0) \right\rangle$$

como queríamos demostrar. □

Que un campo escalar tenga derivadas parciales en un punto es una propiedad muy débil. Por ejemplo, el campo escalar $f(x,y)=\dfrac{xy}{x^2+y^2}$, $f(0,0)=0$ tiene derivadas parciales nulas en $(0,0)$ pero no es continuo en dicho punto. La propiedad de ser diferenciable es mucho más fuerte que tener derivadas parciales. Por ejemplo, es fácil probar que **un campo escalar diferenciable en un punto es continuo en dicho punto**. El siguiente resultado proporciona una condición suficiente de diferenciabilidad muy útil.

10.20 Teorema (Condición suficiente de diferenciabilidad). *Un campo escalar que tiene derivadas parciales continuas en un conjunto abierto es diferenciable en todo punto de dicho conjunto.*

En la práctica suele suponerse que los campos escalares tienen derivadas parciales continuas. Esta hipótesis garantiza que son diferenciables y es suficiente para justificar la mayoría de los resultados que siguen.

Es sabido que una función derivable en un intervalo con derivada nula es constante. Para campos escalares hay un resultado análogo. Observa la hipótesis de que el campo esté definido en un *dominio*.

10.21 Proposición. *Un campo escalar definido en un dominio con derivadas parciales nulas en todo punto del mismo es constante.*

En la siguiente sección te digo cómo calcular rectas y planos tangentes a curvas y superficies considerando las distintas formas en que éstas pueden venir dadas. Mi propósito es esencialmente práctico, a saber, que entiendas la forma de proceder en cada caso; por lo que no me preocupo de justificar con detalle todo lo que digo.

10.1.5. Rectas tangentes y planos tangentes

Curvas en el plano

Una curva Γ en el plano puede venir dada de tres formas:

a) Como la *gráfica de una función* $y=f(x)$ donde $x\in I$ siendo I un intervalo de $\mathbb{R}$.

$$\Gamma=\{(x,f(x)):x\in I\}$$

b) Por medio de *ecuaciones paramétricas* $\gamma(t)=(x(t),y(t))$.

$$\Gamma=\gamma(I)=\{(x(t),y(t)):t\in I\}$$

c) De *forma implícita* como el conjunto de puntos $g(x,y)=0$ donde se anula una función diferenciable de dos variables.

$$\Gamma=\left\{(x,y)\in\mathbb{R}^2:g(x,y)=0\right\}$$

Suele usarse la siguiente terminología. Si $h(x,y)$ es un campo escalar diferenciable, las curvas de ecuación implícita $h(x,y)=c$ o, lo que es igual $h(x,y)-c=0$, donde c es una constante, se llaman *curvas de nivel*. Dichas curvas se obtienen cortando la gráfica de h con planos de la forma $z=c$. Estas curvas son las que ves representadas en los mapas topográficos.

Observa que **a)** es un caso particular de **c)** (basta considerar $g(x,y) = f(x) - y$) y también es un caso particular de **b)** (basta considerar $\gamma(x) = (x, f(x))$).

La tangente en un punto de Γ viene dada en cada caso como sigue.

a′) La tangente en un punto $(a,b) = (a, f(a)) \in \Gamma$ es la recta de ecuación cartesiana $y - b = f'(a)(x - a)$. El vector $(1, f'(a))$ es tangente a Γ en el punto (a,b) y el vector $(f'(a), -1)$ es ortogonal a Γ en el punto (a,b).

b′) La tangente en un punto $\gamma(t_0) = (a,b) \in \Gamma$ es la recta de ecuaciones paramétricas

$$(x,y) = \gamma(t_0) + t\gamma'(t_0) = (a,b) + t(x'(t_0), y'(t_0))$$

El vector $\gamma'(t_0) = (x'(t_0), y'(t_0))$ es tangente a Γ en (a,b).

c′) La tangente en un punto $(a,b) \in \Gamma$ es la recta de ecuación implícita

$$\left\langle \nabla g(a,b) \,\middle|\, (x-a, y-b) \right\rangle = 0$$

Se supone que $\nabla g(a,b) \neq 0$ pues en otro caso, la tangente en (a,b) no está definida. El vector gradiente $\nabla g(a,b)$ es ortogonal a Γ en el punto (a,b).

Estas últimas afirmaciones requieren alguna justificación. Para ello, supongamos que conocemos una *representación paramétrica local* de Γ en torno al punto (a,b). Es decir, hay una curva de la forma $\alpha(t) = (\alpha_1(t), \alpha_2(t)) \in \Gamma$ que pasa por el punto (a,b) y que es derivable[1]. Pongamos $\alpha(t_0) = (a,b)$. Por lo visto en **b′)**, sabemos que la tangente a Γ en (a,b) es la recta que pasa por el punto (a,b) con vector de dirección $\alpha'(t_0)$. Pongamos $h(t) = g(\alpha(t))$. En virtud de la igualdad (10.6), tenemos que $h'(t) = \left\langle \nabla g(\alpha(t)) \,\middle|\, \alpha'(t) \right\rangle$. Pero $h(t) = 0$, por lo que $h'(t) = \left\langle \nabla g(\alpha(t)) \,\middle|\, \alpha'(t) \right\rangle = 0$. Resulta así que el vector $\nabla g(\alpha(t))$ es ortogonal al vector tangente $\alpha'(t)$. En particular, el vector $\nabla g(a,b)$ es ortogonal al vector $\alpha'(t_0)$ tangente a Γ en (a,b). Concluimos que la recta que pasa por (a,b) y tiene como vector ortogonal $\nabla g(a,b)$ es la recta tangente a Γ en (a,b), pero dicha recta es justamente la recta de ecuación cartesiana $\left\langle \nabla g(a,b) \,\middle|\, (x-a, y-b) \right\rangle = 0$.

De lo antes visto, merece la pena destacar la siguiente propiedad.

El vector gradiente $\nabla g(x,y)$ de un campo escalar es ortogonal en todo punto (x,y) (en el que $\nabla g(x,y) \neq \mathbf{0}$) a la curva de nivel que pasa por dicho punto.

Superficies en $\mathbb{R}^3$

Una superficie S en el espacio $\mathbb{R}^3$ puede venir dada de tres formas:

a) Como la gráfica de una función $y = f(x,y)$ donde $(x,y) \in A$ siendo A un conjunto de $\mathbb{R}^2$.

$$S = \{(x, y, f(x,y)) : (x,y) \in A\}$$

b) Por medio de ecuaciones paramétricas $\gamma(s,t) = (x(s,t), y(s,t), z(s,t))$ donde $(s,t) \in A \subset \mathbb{R}^2$.

$$S = \gamma(A) = \{(x(s,t), y(s,t), z(s,t)) : (s,t) \in A\}$$

[1] El teorema de la función implícita, que se verá más adelante, garantiza la existencia de dicha curva siempre que el vector gradiente $\nabla g(a,b) \neq \mathbf{0}$.

c) De forma implícita como el conjunto de puntos $g(x,y,z)=0$ donde se anula una función diferenciable de tres variables.

$$S=\left\{(x,y,z)\in\mathbb{R}^3 : g(x,y,z)=0\right\}$$

Observa que **a)** es un caso particular de **c)** (basta considerar $g(x,y,z)=f(x,y)-z$) y también es un caso particular de **b)** (basta considerar $\gamma(s,t)=(s,t,f(s,t))$). El plano tangente en un punto de S viene dada en cada caso como sigue.

a′) El plano tangente en un punto $(a,b,c)=(a,b,f(a,b))\in S$ es el plano de ecuación cartesiana

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

Los vectores $\left(1,0,\dfrac{\partial f}{\partial x}(a,b)\right)$ y $\left(0,1,\dfrac{\partial f}{\partial y}(a,b)\right)$ son tangentes a S en (a,b,c) y el vector

$$\left(\frac{\partial f}{\partial x}(a,b),\frac{\partial f}{\partial y}(a,b),-1\right)$$

es ortogonal a S en el punto (a,b,c).

b′) El plano tangente en un punto $\gamma(s_0,t_0)=(a,b,c)\in S$ es el plano de ecuaciones paramétricas

$$(x,y,z)=\gamma(s_0,t_0)+s\frac{\partial\gamma}{\partial s}(s_0,t_0)+t\frac{\partial\gamma}{\partial t}(s_0,t_0)$$

Donde

$$\frac{\partial\gamma}{\partial s}(s_0,t_0)=\left(\frac{\partial x}{\partial s}(s_0,t_0),\frac{\partial y}{\partial s}(s_0,t_0),\frac{\partial z}{\partial s}(s_0,t_0)\right)$$

y

$$\frac{\partial\gamma}{\partial t}(s_0,t_0)=\left(\frac{\partial x}{\partial t}(s_0,t_0),\frac{\partial y}{\partial t}(s_0,t_0),\frac{\partial z}{\partial t}(s_0,t_0)\right)$$

Dichos vectores son tangentes a S en (a,b,c).

c′) El plano tangente en un punto $(a,b,c)\in S$ es el plano de ecuación implícita

$$\left\langle \nabla g(a,b,c)\,\middle|\,(x-a,y-b,z-c)\right\rangle=0$$

Se supone que $\nabla g(a,b,c)\neq 0$ pues en otro caso, el plano tangente a S en (a,b,c) no está definido. El vector gradiente $\nabla g(a,b,c)$ es ortogonal a S en el punto (a,b,c).

Si $g(x,y,z)$ es un campo escalar, las superficies de ecuación implícita $g(x,y,z)=c$ o, lo que es igual $g(x,y,z)-c=0$, donde c es una constante, se llaman *superficies de nivel* (cuando el campo se interpreta como un potencial se llaman *superficies equipotenciales*). De lo dicho en **c′)**, se sigue que *el vector gradiente* $\nabla g(x,y,z)$ *es ortogonal en todo punto* (x,y,z) *(en el que* $\nabla g(x,y,z)\neq\mathbf{0}$*) a la superficie de nivel que pasa por dicho punto.*

Curvas en $\mathbb{R}^3$

Una curva Γ en el espacio puede venir dada de dos formas.

a) Como intersección de dos superficies S_1 y S_2.

b) Por medio de ecuaciones paramétricas $\gamma(t) = (x(t), y(t), z(t))$ donde $t \in I \subset \mathbb{R}$ e I es un intervalo.

$$\Gamma = \gamma(I) = \{(x(t), y(t), z(t)) : t \in I\}$$

La tangente en un punto de Γ viene dada en cada caso como sigue.

a′) La tangente en un punto $(a,b,c) \in \Gamma$ es la recta intersección de los planos tangentes a S_1 y a S_2 en (a,b,c). Por ejemplo, si las superficies vienen dadas por sus ecuaciones implícitas.

$$\begin{cases} S_1 = \left\{(x,y,z) \in \mathbb{R}^3 : f(x,y,z) = 0\right\} \\ S_2 = \left\{(x,y,z) \in \mathbb{R}^3 : g(x,y,z) = 0\right\} \end{cases} \qquad \Gamma = \left\{(x,y,z) \in \mathbb{R}^3 : g(x,y,z) = f(x,y,z) = 0\right\}$$

Entonces, las ecuaciones implícitas de la recta tangente son

$$\begin{cases} \big\langle \nabla f(a,b,c) \,\big|\, (x-a, y-b, z-c) \big\rangle = 0 \\ \big\langle \nabla g(a,b,c) \,\big|\, (x-a, y-b, z-c) \big\rangle = 0 \end{cases}$$

Donde se supone que los vectores gradiente $\nabla f(a,b,c)$, $\nabla g(a,b,c)$ son linealmente independientes pues, en otro caso, la recta tangente a la curva Γ en (a,b,c) no está definida.

b′) La tangente en un punto $\gamma(t_0) = (a,b,c) \in \Gamma$ es la recta de ecuaciones paramétricas

$$(x,y,z) = \gamma(t_0) + t\gamma'(t_0) = (a,b,c) + t(x'(t_0), y'(t_0), z'(t_0))$$

El vector $\gamma'(t_0) = (x'(t_0), y'(t_0), z'(t_0))$ es tangente a Γ en (a,b,c).

Derivadas parciales de orden superior

Supongamos un campo escalar f que tiene derivadas parciales $D_k f$ en un conjunto $E \subset \mathbb{R}^n$. Las funciones $D_k f$ son también campos escalares que podemos, cuando se dejen, volver a derivar parcialmente en puntos de E. Obtenemos de esta forma las *derivadas parciales de segundo orden* de f, es decir las funciones $D_j(D_k f)$, que se representan simbólicamente de las formas

$$D_{jk} f(\mathbf{x}), \quad \frac{\partial^2 f}{\partial x_j \partial x_k}(\mathbf{x}), \quad \frac{\partial^2 f}{\partial x_k^2}(\mathbf{x})$$

De forma análoga se definen las derivadas parciales de tercer orden de f como las derivadas parciales de las derivadas parciales de segundo orden de f y se representan por

$$D_{jkm} f(\mathbf{x}), \quad \frac{\partial^3 f}{\partial x_j \partial x_k \partial x_m}(\mathbf{x}); \quad \frac{\partial^3 f}{\partial x_k^3}(\mathbf{x}); \quad \frac{\partial^3 f}{\partial x_k^2 \partial x_j}(\mathbf{x})$$

Es natural preguntarse si el orden en que se realizan las derivadas debe ser o no tenido en cuenta. Afortunadamente, en la mayoría de los casos podemos olvidarlo porque se verifica el siguiente resultado.

10.22 Definición. Se dice que un campo escalar f es de clase C^k en un abierto $E \subset \mathbb{R}^n$ si f tiene derivadas parciales de orden k continuas en E.

10.23 Teorema. *Las derivadas parciales de orden menor o igual que k de un campo escalar de clase C^k solamente dependen del número de veces que se deriva parcialmente respecto de cada variable, pero el orden en que se realicen dichas derivaciones no afecta para nada al resultado final.*

10.1.6. Ejercicios

Como para calcular derivadas parciales de una función de varias variables se consideran fijas todas las variables menos aquella respecto a la que se deriva, calcular derivadas parciales es lo mismo que derivar funciones de una variable. Solamente debes tener cuidado para darte cuenta qué tipo de función es la que tienes que derivar porque ello puede depender de la variable respecto de la que derivas. Por ejemplo, la función $f(x,y)=x^y$ cuando fijas y (para derivar respecto a x) es una función potencia (la variable está en la base y el exponente está fijo) y cuando fijas x (para derivar respecto a y) es una función exponencial (la variable está en el exponente y la base está fija).

Te recuerdo que es muy frecuente, sobre todo en libros de Física e ingenierías diversas, representar las funciones por letras. Así, lo que los matemáticos solemos escribir $f(x,y)=\cos(xy)+xy^2$, para indicar que f es una función de dos variables x e y cuyo valor en el punto (x,y) viene dado por $\cos(xy)+xy^2$, suele expresarse de forma menos precisa en la forma $z=\cos(xy)+xy^2$, cuyo significado es exactamente el mismo que el anterior cambiando f por z. Naturalmente, en vez de z puede usarse cualquier otro símbolo que sea distinto de x e y. Tienes que acostumbrarte a esta notación y entender cuándo una letra representa una variable y cuándo representa una función.

1. Calcula las derivadas parciales de primer orden de los campos escalares:

   (a) $f(x,y)=x^2y+z^2x+y\,\mathrm{sen}(xz)$ (b) $z=(x^2+y^3)\,\mathrm{e}^{-xy}$ (c) $w=x\,\mathrm{e}^z+z\,\mathrm{e}^y+xyz$.

2. Calcula las derivadas parciales de primer y segundo orden del campo $f(x,y,z)=\dfrac{xy}{1+y^2+z^2}$.

3. Calcula las derivadas parciales de primer y segundo orden de los campos escalares:

   (a) $z=\mathrm{sen}\left(\cos\left(\mathrm{e}^{xy}\right)\right)$ (b) $w=\log\left(4+\mathrm{arc\,tg}(x/y)\right)$ (c) $u=\mathrm{tg}\left((xy)^z\right)$ (d) $v=\mathrm{arc\,tg}\left(z^{xy}\right)$

   Te recuerdo que una dirección viene dada por un vector de norma euclídea 1. Si $\mathbf{a}$ y $\mathbf{b}$ son puntos de $\mathbb{R}^n$ la dirección del punto $\mathbf{a}$ hacia el punto $\mathbf{b}$ viene dada por el vector $\dfrac{\mathbf{b}-\mathbf{a}}{\|\mathbf{b}-\mathbf{a}\|}$.

4. Calcula la derivada direccional de $f(x,y)=\log(1+\sqrt{x^2+y^2})$ en el punto $(1,2)$ en la dirección hacia el origen.

5. Calcula la derivada direccional de $z(x,y)=\mathrm{arc\,tg}\left(\dfrac{xy}{x^2+y^2}\right)$ en el punto $(1,1)$ en la dirección hacia el punto $(2,1)$.

6. Calcula valores de a, b y c para que la derivada direccional de la función

$$f(x,y,z)=axy^2+byz+cz^2x^3$$

   en el punto $(1,2,-1)$ tenga un valor máximo igual a 64 en la dirección del eje OZ.

7. Calcula la ecuación de la recta tangente y de la recta normal a la elipse de ecuación

$$\frac{x^2}{a^2}+\frac{y^2}{b^2}=1$$

   en un punto (u,v) de la misma.

8. Considera la curva dada por las ecuaciones paramétricas $x(t)=\mathrm{e}^t+\cos t$, $y(t)=\mathrm{e}^{-t}+\operatorname{sen} t$. Calcula la ecuación de la recta tangente en el punto $(x(0),y(0))$.

9. Calcula, para los siguientes campos escalares, el vector normal en P_0 a la curva de nivel que pasa por dicho punto.

   a) $f(x,y)=\operatorname{arctg}\left(\dfrac{y}{\sqrt{1+x^2+y^2}}\right) \qquad P_0=(1,1)$.

   b) $f(x,y)=\dfrac{\operatorname{sen}(x+y)}{2+\cos(x-y)} \qquad P_0=(\pi/2,\pi/4)$.

10. Calcula la derivada de $h(x,y)=\dfrac{x-y}{1+\log(1+x^2y^2)}$ en el punto $(-1,-1)$ en la dirección dada por el vector ortogonal (de norma 1) en el punto $(1,1)$ a la curva de nivel del campo $f(x,y)=xy^3+x^3y$ que pasa por dicho punto.

11. Calcula las ecuaciones del plano tangente y de la recta normal a cada una de las siguientes superficies en el punto P_o indicado.

$$z^2-2x^2-2y^2-12=0, \quad P_o(1,-1,4); \qquad z-\log(x^2+y^2)=0, \quad P_o(1,0,0)$$
$$x^2+y^2+z^3-2x+4y+3z+1=0, \quad P_o(3,4,-3); \qquad 4-x^2-4z^2=y, \quad P_o(0,0,1)$$
$$z(xy-1)-(x+y)=0, \quad P_o(1,2,3); \qquad z+\mathrm{e}^z+2x+2y-x^2-y^2-3=0, \quad P_o(1,1+\sqrt{e},1)$$

12. Halla la ecuación de la tangente a la curva dada como intersección del elipsoide $x^2+4y^2+2z^2=27$ y el hiperboloide $x^2+y^2-2z^2=11$ en el punto $(3,-2,1)$.

13. Calcula la ecuación de la recta tangente a la curva definida por la intersección de las superficies $z=xy$, $x^2+y^2-2z=4$ en el punto $(3,1,3)$. Comprueba el resultado expresando la curva por sus ecuaciones paramétricas.

14. Calcula la ecuación de la recta tangente a la curva definida por la intersección de las superficies $4xz=(x+z)y$, $3z^2+y=5x$ en el punto $(1,2,1)$.

10.1.7. Extremos relativos

10.24 Definición. Sea f un campo escalar definido en un conjunto $E\subset\mathbb{R}^n$. Se dice que f tiene un **máximo relativo** (resp. **mínimo relativo**) en un punto $\mathbf{a}\in E$, si $\mathbf{a}$ es un punto interior de E y existe un número $r>0$ tal que $B(\mathbf{a},r)\subset E$ y $f(\mathbf{x})\leqslant f(\mathbf{a})$ (resp. $f(\mathbf{a})\leqslant f(\mathbf{x})$) para todo $\mathbf{x}\in B(\mathbf{a},r)$. Cuando estas desigualdades se verifican de forma estricta se dice que el máximo o el mínimo relativo es estricto.

Los puntos en los que f tiene un máximo o un mínimo relativos se llaman **extremos relativos** de f.

10.25 Proposición (Condición necesaria de extremo relativo). *Sea f un campo escalar definido en un conjunto $E\subset\mathbb{R}^n$ y supongamos que f tiene un extremo relativo en un punto $\mathbf{a}\in E$ y además que el vector gradiente de f en $\mathbf{a}$ está definido. Entonces se verifica que $\nabla f(\mathbf{a})=\mathbf{0}$. Es decir, las derivadas parciales de primer orden de f en $\mathbf{a}$ son todas nulas.*

Demostración. Supongamos que f tiene un máximo relativo en $\mathbf{a}$ y sea $r>0$ tal que $B(\mathbf{a},r)\subset E$ y $f(\mathbf{x})\leqslant f(\mathbf{a})$ para todo $\mathbf{x}\in B(\mathbf{a},r)$. Definamos $\varphi:]-r,r[\to\mathbb{R}$ por $\varphi(t)=f(\mathbf{a}+t\mathbf{e_k})$. La función φ está definida en el intervalo $]-r,r[$ pues para todo $t\in]-r,r[$ se tiene que $\|\mathbf{a}+t\mathbf{e_k}-\mathbf{a}\|=|t|<r$ por lo que $\mathbf{a}+t\mathbf{e_k}\in B(\mathbf{a},r)\subset E$. Además, para todo $t\in]-r,r[$ se tiene que $\varphi(t)=f(\mathbf{a}+t\mathbf{e_k})\leqslant f(\mathbf{a})=\varphi(0)$. Luego φ tiene en $t=0$ un máximo relativo. Además como, por hipótesis, existe $D_kf(\mathbf{a})$, tenemos que φ es derivable en $t=0$. Luego $\varphi'(0)=0$, pero $\varphi'(0)=D_kf(\mathbf{a})$. □

10.26 Definición. Los puntos donde se anula el gradiente de un campo escalar f se llaman **puntos críticos** de f. Los puntos críticos de un campo escalar que no son extremos relativos se llaman **puntos de silla**.

Si f es un campo escalar diferenciable, en los puntos críticos el hiperplano tangente es “horizontal”.

La condición necesaria de extremo relativo no es suficiente. Por ejemplo, el campo escalar $f(x,y)=x^2-y^2$ tiene un punto crítico en $(0,0)$, pero no tiene extremo relativo en dicho punto pues en toda bola centrada en $(0,0)$ toma valores positivos y negativos.

Al igual que para funciones de una variable, la derivada segunda proporciona una condición suficiente de extremo relativo, para campos escalares de varias variables las derivadas parciales de segundo orden nos van a permitir dar una condición suficiente de extremo relativo. Necesitaremos para ello el siguiente resultado.

10.27 Proposición. *Sea f un campo escalar definido en un conjunto $E\subset\mathbb{R}^n$ y supongamos que f tiene derivadas parciales de segundo orden continuas en un punto $\mathbf{a}$ interior de E. Sea $r>0$ tal que $B(\mathbf{a},r)\subset E$. Entonces para todo $\mathbf{x}$ con $\|\mathbf{x}\|<r$ se tiene que*

$$f(\mathbf{a}+\mathbf{x})=f(\mathbf{a})+\sum_{k=1}^{n}D_kf(\mathbf{a})x_k+\frac{1}{2}\sum_{j=1}^{n}\sum_{k=1}^{n}D_{j,k}f(\mathbf{a})x_kx_j+\|\mathbf{x}\|^2\varphi(\mathbf{x}) \quad \textit{con } \lim_{\mathbf{x}\to\mathbf{0}}\varphi(\mathbf{x})=0 \tag{10.7}$$

Demostración. Fijemos el vector $\mathbf{x}$ en las condiciones del enunciado y definamos la función $h_{\mathbf{x}}(t)=f(\mathbf{a}+t\mathbf{x})$. Dicha función está definida en un intervalo abierto $I\supset[-1,1]$ y es dos veces derivable en $t=0$. El teorema de Taylor–Young dice que

$$h_{\mathbf{x}}(t)=h_{\mathbf{x}}(0)+h_{\mathbf{x}}'(0)t+\frac{1}{2}h_{\mathbf{x}}''(0)t^2+t^2r(t,\mathbf{x}) \tag{10.8}$$

con $\lim_{t\to0}r(t,\mathbf{x})=0$. Pongamos $\gamma(t)=\mathbf{a}+t\mathbf{x}$, con lo cual $h_{\mathbf{x}}(t)=f(\gamma(t))$. Por (10.6) tenemos que

$$h_{\mathbf{x}}'(t)=\sum_{j=1}^{n}D_jf(\gamma(t))\gamma'_j(t)=\sum_{j=1}^{n}D_jf(\gamma(t))x_j \tag{10.9}$$

Donde hemos tenido en cuenta que las componentes de γ son $\gamma_j(t)=a_j+tx_j$. En particular

$$h_{\mathbf{x}}'(0)=\sum_{j=1}^{n}D_jf(\mathbf{a})x_j \tag{10.10}$$

Volviendo a derivar la igualdad (10.9) en $t=0$, aplicando otra vez la misma regla de derivación a los campos escalares $D_jf(\gamma(t))$, obtenemos

$$h_{\mathbf{x}}''(0)=\sum_{j=1}^{n}\left(\sum_{k=1}^{n}D_{j,k}f(\mathbf{a})x_k\right)x_j=\sum_{j=1}^{n}\sum_{k=1}^{n}D_{j,k}f(\mathbf{a})x_kx_j \tag{10.11}$$

Sustituyendo las igualdades (10.10) y (10.11) en (10.8) y haciendo $t=1$ obtenemos

$$f(\mathbf{a}+\mathbf{x})=f(\mathbf{a})+\sum_{k=1}^{n}D_k f(\mathbf{a})x_k+\frac{1}{2}\sum_{j=1}^{n}\sum_{k=1}^{n}D_{j,k}f(\mathbf{a})x_k x_j+r(1,\mathbf{x})$$

Solo queda probar que $r(1,\mathbf{x})$ puede escribirse en la forma $r(1,\mathbf{x})=\|\mathbf{x}\|^2\varphi(\mathbf{x})$ con $\lim_{\mathbf{x}\to\mathbf{0}}\varphi(\mathbf{x})=0$ pero esto vamos a dejarlo para otra ocasión. □

10.28 Definición. Sea f un campo escalar de n variables que tiene derivadas parciales de segundo orden continuas en un punto $\mathbf{a}$. La matriz $n\times n$

$$H(f,\mathbf{a})=\big(D_{ij}f(\mathbf{a})\big)_{1\leqslant i,j\leqslant n}$$

se llama **matriz hessiana** de f en $\mathbf{a}$.

Observa que la matriz hessiana es simétrica porque $D_{ij}f(\mathbf{a})=D_{ji}f(\mathbf{a})$. En consecuencia, dicha matriz define una forma cuadrática, que representaremos por $Q(f,\mathbf{a})$, que viene dada para todo $\mathbf{x}=(x_1,x_2,\dots,x_n)\in\mathbb{R}^n$ por

$$Q(f,\mathbf{a})(\mathbf{x})=\mathbf{x}\cdot H(f,\mathbf{a})\cdot\mathbf{x}^t=\sum_{j=1}^{n}\sum_{k=1}^{n}D_{j,k}f(\mathbf{a})x_k x_j$$

donde el punto "·" indica producto matricial y $\mathbf{x}^t$ es el vector columna $\mathbf{x}$. Con esta notación podemos escribir la igualdad (10.7) en la forma

$$f(\mathbf{a}+\mathbf{x})=f(\mathbf{a})+\big\langle\nabla f(\mathbf{a})\,\big|\,\mathbf{x}\big\rangle+\frac{1}{2}Q(f,\mathbf{a})(\mathbf{x})+\|\mathbf{x}\|^2\varphi(\mathbf{x})\quad\text{donde}\quad\lim_{\mathbf{x}\to\mathbf{0}}\varphi(\mathbf{x})=0\tag{10.12}$$

Si suponemos que $\mathbf{a}$ es un punto crítico de f podemos escribir

$$f(\mathbf{a}+\mathbf{x})=f(\mathbf{a})+\frac{1}{2}Q(f,\mathbf{a})(\mathbf{x})+\|\mathbf{x}\|^2\varphi(\mathbf{x})\quad\text{donde}\quad\lim_{\mathbf{x}\to\mathbf{0}}\varphi(\mathbf{x})=0\tag{10.13}$$

De donde se sigue que

$$\frac{f(\mathbf{a}+\mathbf{x})-f(\mathbf{a})}{\|\mathbf{x}\|^2}=\frac{1}{2\,\|\mathbf{x}\|^2}Q(f,\mathbf{a})(\mathbf{x})+\varphi(\mathbf{x})\quad\text{donde}\quad\lim_{\mathbf{x}\to\mathbf{0}}\varphi(\mathbf{x})=0$$

Teniendo en cuenta que las formas cuadráticas son polinomios homogéneos de grado 2, es decir, $Q(f,\mathbf{a})(\lambda\mathbf{x})=\lambda^2 Q(f,\mathbf{a})(\mathbf{x})$, se tiene que $\dfrac{1}{2\,\|\mathbf{x}\|^2}Q(f,\mathbf{a})(\mathbf{x})=\dfrac{1}{2}Q(f,\mathbf{a})(\mathbf{x}/\|\mathbf{x}\|)$. Resulta así la igualdad

$$\frac{f(\mathbf{a}+\mathbf{x})-f(\mathbf{a})}{\|\mathbf{x}\|^2}=\frac{1}{2}Q(f,\mathbf{a})(\mathbf{x}/\|\mathbf{x}\|)+\varphi(\mathbf{x})\quad\text{donde}\quad\lim_{\mathbf{x}\to\mathbf{0}}\varphi(\mathbf{x})=0\tag{10.14}$$

10.29 Definición. Una forma cuadrática $Q(\mathbf{x})=\sum_{i,j=1}^{n}\alpha_{ij}x_i x_j$ se llama:

- **Positiva definida** si $Q(\mathbf{x})>0$ para todo $\mathbf{x}\in\mathbb{R}^n$ con $\mathbf{x}\neq\mathbf{0}$.
- **Semidefinida positiva** si $Q(\mathbf{x})\geqslant 0$ para todo $\mathbf{x}\in\mathbb{R}^n$.
- **Positiva negativa** si $Q(\mathbf{x})<0$ para todo $\mathbf{x}\in\mathbb{R}^n$ con $\mathbf{x}\neq\mathbf{0}$.
- **Semidefinida negativa** si $Q(\mathbf{x})\leqslant 0$ para todo $\mathbf{x}\in\mathbb{R}^n$.
- **No definida** si hay vectores $\mathbf{x}$ para los que $Q(\mathbf{x})>0$ y hay vectores $\mathbf{x}$ para los que $Q(\mathbf{x})<0$.

10.30 Teorema. *Sea f un campo escalar definido en un conjunto $E \subset \mathbb{R}^n$ y supongamos que f tiene derivadas parciales de segundo orden continuas en un punto $\mathbf{a}$ interior de E que además es un punto crítico de f. Sea $Q(f,\mathbf{a})$ la forma cuadrática asociada a la matriz hessiana de f en $\mathbf{a}$.*

$$Q(f,\mathbf{a})(\mathbf{x}) = \mathbf{x}\cdot H(f,\mathbf{a})\cdot \mathbf{x}^t = \sum_{j=1}^{n}\sum_{k=1}^{n} D_{j,k}f(\mathbf{a})x_k x_j$$

a) Si la forma cuadrática $Q(f,\mathbf{a})$ es definida positiva entonces f tiene en $\mathbf{a}$ un mínimo relativo estricto.

b) Si la forma cuadrática $Q(f,\mathbf{a})$ es definida negativa entonces f tiene en $\mathbf{a}$ un máximo relativo estricto.

c) Si la forma cuadrática $Q(f,\mathbf{a})$ es no definida entonces f tiene un punto de silla en $\mathbf{a}$.

d) Si f tiene un máximo relativo en $\mathbf{a}$ entonces la forma cuadrática $Q(f,\mathbf{a})$ es semidefinida negativa.

e) Si f tiene un mínimo relativo en $\mathbf{a}$ entonces la forma cuadrática $Q(f,\mathbf{a})$ es semidefinida positiva.

Demostración. Como $Q(f,\mathbf{a})$ es una función polinómica y, por tanto, continua, y la esfera unidad de $\mathbb{R}^n$, $S(\mathbf{0},1)=\{\mathbf{u}\in\mathbb{R}^n : \|\mathbf{u}\|=1\}$, es un conjunto compacto, en virtud del teorema de Weierstrass, dicha función alcanza un mínimo valor y un máximo valor en $S(\mathbf{0},1)$. Sea

$$m = \operatorname{mín}\{Q(f,\mathbf{a})(\mathbf{u}) : \|\mathbf{u}\|=1\}, \quad M = \operatorname{máx}\{Q(f,\mathbf{a})(\mathbf{u}) : \|\mathbf{u}\|=1\}$$

a) Supongamos que $Q(f,\mathbf{a})$ es definida positiva. Entonces se tiene que $m>0$. y, por la igualdad (10.14), tenemos que

$$\frac{f(\mathbf{a}+\mathbf{x})-f(\mathbf{a})}{\|\mathbf{x}\|^2} = \frac{1}{2}Q(f,\mathbf{a})(\mathbf{x}/\|\mathbf{x}\|)+\varphi(\mathbf{x}) \geqslant \frac{m}{2}+\varphi(\mathbf{x}) \quad \text{donde} \quad \lim_{\mathbf{x}\to\mathbf{0}}\varphi(\mathbf{x})=0$$

La condición $\lim\limits_{\mathbf{x}\to\mathbf{0}}\varphi(\mathbf{x})=0$ garantiza la existencia de un número $s>0$ tal que $|\varphi(\mathbf{x})|<m/4$ siempre que $0<\|\mathbf{x}\|<s$. En consecuencia, si en la desigualdad anterior suponemos que $0<\|\mathbf{x}\|<s$, se tiene

$$\frac{f(\mathbf{a}+\mathbf{x})-f(\mathbf{a})}{\|\mathbf{x}\|^2} \geqslant \frac{m}{2}+\varphi(\mathbf{x}) > \frac{m}{2}-\frac{m}{4}=\frac{m}{4}>0$$

Deducimos que $f(\mathbf{a}+\mathbf{x})-f(\mathbf{a})>0$ para todo $\mathbf{x}$ con $0<\|\mathbf{x}\|<s$. O, lo que es igual, $f(\mathbf{z})-f(\mathbf{a})>0$ para todo $\mathbf{z}$ tal que $0<\|\mathbf{z}-\mathbf{a}\|<s$. Lo que prueba que f tiene en $\mathbf{a}$ un mínimo relativo estricto.

Los demás puntos se prueban de forma parecida. □

Para poder usar el resultado anterior hay que saber clasificar una forma cuadrática. Hay varios procedimientos sencillos para ello. Los dos que siguen a continuación son los que me parecen más cómodos.

Clasificación de formas cuadráticas

Sean $\mathcal{A}=\left(a_{ij}\right)_{1\leqslant i,j\leqslant n}$ una matriz simétrica de números reales y

$$Q_{\mathcal{A}}(\mathbf{x}) = \mathbf{x}\cdot\mathcal{A}\cdot\mathbf{x}^t = \sum_{i,j=1}^{n} a_{ij}xx^j \tag{10.15}$$

la forma cuadrática definida por $\mathcal{A}$. Los *valores propios* de $\mathcal{A}$ son las raíces del polinomio característico $p(\lambda)$, que se define como el determinante de la matriz $\mathcal{A}-\lambda I$:

$$p(\lambda)=\left|\mathcal{A}-\lambda I\right|$$

Es sabido que, en la situación que estamos considerando, las raíces de dicho polinomio son todas reales.

Sean λ_j $(1\leqslant j\leqslant n)$ los valores propios de $\mathcal{A}$. Se demuestra que hay una base $\mathbf{B}=\{\mathbf{u}_1,\mathbf{u}_2,\dots,\mathbf{u}_n\}$ en $\mathbb{R}^n$ tal que para todo vector $\mathbf{x}\in\mathbb{R}^n$ se tiene que

$$Q_{\mathcal{A}}(\mathbf{x})=\sum_{j=1}^{n}\lambda_j x_j^2$$

donde $(x_1,x_2,\dots,x_n)$ son los coordenadas del vector $\mathbf{x}$ en la base $\mathbf{B}$. De aquí se siguen los siguientes criterios.

- La forma cuadrática $Q_{\mathcal{A}}$ es definida positiva si, y sólo si, todos los valores propios de $\mathcal{A}$ son positivos.
- La forma cuadrática $Q_{\mathcal{A}}$ es definida negativa si, y sólo si, todos los valores propios de $\mathcal{A}$ son negativos.
- La cuadrática $Q_{\mathcal{A}}$ es no definida si, y sólo si, $\mathcal{A}$ tiene valores propios positivos y negativos.
- La forma cuadrática $Q_{\mathcal{A}}$ es semidefinida positiva si, y sólo si, todos los valores propios de $\mathcal{A}$ son mayores o iguales que 0.
- La forma cuadrática $Q_{\mathcal{A}}$ es semidefinida negativa si, y sólo si, todos los valores propios de $\mathcal{A}$ son menores o iguales que 0.

Para aplicar estos criterios no es preciso calcular los valores propios de $\mathcal{A}$ sino solamente saber cuántos de ellos son positivos, negativos o nulos. Afortunadamente, hay un criterio que nos proporciona esta información sin más que observar los coeficientes del polinomio característico.

10.31 Proposición (Regla de los signos de Descartes). *Sea $f(x)=a_nx^n+a_{n-1}x^{n-1}+\cdots+a_1x+a_0$ un polinomio con coeficientes reales y cuyas raíces son todas números reales. Se verifica entonces que:*

a) El número de raíces positivas de f (contando multiplicidades) es igual al número de cambios de signo en la sucesión $(a_n,a_{n-1},\dots,a_1,a_0)$ de los coeficientes de f.

b) El número de raíces negativas de f (contando multiplicidades) es igual al número de cambios de signo en la sucesión $((-1)^na_n,(-1)^{n-1}a_{n-1},\dots,-a_1,a_0)$ de los coeficientes de $f(-x)$.

Para contar los cambios de signo en la sucesión de coeficientes se saltan los coeficientes nulos. Por ejemplo, si $f(x)=2x^6+x^5-x^3+x^2-5$, la sucesión de coeficientes de f es $(2,1,0,-1,1,0,-1)$ cuyo número de cambios de signo es 3.

10.32 Corolario. *Sea $p(\lambda)$ el polinomio característico de la matriz hessiana de f en $\mathbf{a}$. Entonces.*

- *Si $p(\lambda)$ tiene grado n, todos sus coeficientes son distintos de cero y tienen igual sigo, se verifica que f tiene un máximo relativo estricto en $\mathbf{a}$.*

- *Si $p(\lambda)$ tiene grado n, todos sus coeficientes son distintos de cero y van alternando su sigo, se verifica que f tiene un mínimo relativo estricto en* $\mathbf{a}$.

- *Si $p(\lambda)$ tiene grado n, sus coeficientes nulos van seguidos y llegan hasta el término independiente y los coeficientes no nulos tienen igual signo o van alternando su sigo, entonces no puede afirmarse nada.*

- *En todos los demás casos, f tiene un punto de silla en* $\mathbf{a}$.

Otro criterio para estudiar el carácter de la forma cuadrática (10.15) se basa en la sucesión de signos de los *menores principales* de la matriz $\mathcal{A}$. El menor principal de orden k es el determinante $\Delta_k = \left|a_{i,j}\right|_{1\leqslant i,j\leqslant k}$. Se verifica que:

- *Si todos los determinantes principales son positivos la forma cuadrática es definida positiva.*

- *Si los determinantes principales son todos distintos de cero y van alternando signo siendo el primero de ellos negativo, la forma cuadrática es definida negativa.*

- *Si los determinantes principales son nulos a partir de uno de ellos en adelante y los no nulos son positivos o van alternando signo siendo el primero de ellos negativo, no puede afirmarse nada.*

- *En los demás casos la forma cuadrática es no definida.*

Observa que cuando la dimensión n es par, si el determinante de la matriz $\mathcal{A}$ es negativo entonces la forma es no definida.

Podemos particularizar este criterio para el caso de dos dimensiones.

Sea $A\subset\mathbb{R}^2$ un conjunto abierto y sea f un campo escalar definido en A que tiene derivadas parciales de segundo orden continuas. Supongamos que $(a,b)\in A$ es un punto crítico de f y sea

$$H(f,(a,b)) = \begin{pmatrix} \dfrac{\partial^2 f}{\partial x^2}(a,b) & \dfrac{\partial^2 f}{\partial x\partial y}(a,b) \\ \dfrac{\partial^2 f}{\partial x\partial y}(a,b) & \dfrac{\partial^2 f}{\partial y^2}(a,b) \end{pmatrix}$$

la matriz hessiana de f en (a,b) y notemos $\det H(f,(a,b))$ su determinante.

- Si $\det H(f,(a,b))>0$ y $\dfrac{\partial^2 f}{\partial x^2}(a,b)>0$ entonces f tiene en (a,b) un mínimo relativo estricto.
- Si $\det H(f,(a,b))>0$ y $\dfrac{\partial^2 f}{\partial x^2}(a,b)<0$ entonces f tiene en (a,b) un máximo relativo estricto.
- Si $\det H(f,(a,b))<0$ entonces f no tiene extremo relativo en (a,b). Se dice que (a,b) es un punto de silla de f.
- Cuando $\det H(f,(a,b))=0$ el conocimiento de la matriz hessiana no permite decidir si hay o no hay extremo relativo en (a,b). Cuando esto sucede puede ser interesante estudiar el comportamiento de las curvas $f(a,t+b)$ y $f(a+t,b)$. Si alguna de dichas curvas no tiene extremo relativo o tienen extremos relativos de distinta naturaleza en $t=0$, podemos concluir que en (a,b) no hay extremo relativo de f.

Ejercicios

1. Determinar los extremos relativos de las funciones:

$f(x,y)=2x^3+6xy^2-3x^2+3y^2$; $\quad f(x,y)=x^2-2xy^2+y^4-y^5$;

$f(x,y)=\dfrac{x^2y^2-8x+y}{xy}$; $\quad f(x,y)=2x^2+y^2+8x-6y+20$;

$f(x,y)=-x^3+4xy-2y^2+1$; $\quad f(x,y)=\cos(x)\cos(y)$

$f(x,y)=2x+y+x^2+xy+y^3$; $\quad f(x,y)=x^2y^2-x^2-y^2$;

$f(x,y)=x\log y-x$ $\quad f(x,y)=2x^4+y^4-4x^2-2y^2$;

$f(x,y)=xy(1-x-y)$; $\quad f(x,y)=-4x^3+6x^2y+3y^4-4y^3$

$f(x,y,z)=x^2+y^2+3z^2+yz+2xz-xy$; $\quad f(x,y,z)=(x^2+z^2)\,\mathrm{e}^{x(y^2+z^2+1)}$;

$f(x,y,z)=xy+xz+yz$; $\quad f(x,y,z)=(x+z^2)\,\mathrm{e}^{-x(y^2+z^2+1)}$

2. Trazar un plano que pase por el punto $(1,2,3)$ y que forme con los ejes coordenados un tetraedro de volumen mínimo (el volumen del tetraedro es un tercio del área de la base por la altura).

3. **Recta de mínimos cuadrados**. Dados n puntos $(x_i,y_i)\in\mathbb{R}^2$, determinar los números α y β para que la cantidad $\displaystyle\sum_{i=1}^{n}\Big(y_i-\alpha x_i-\beta\Big)^2$ sea mínima.

4. Dados m puntos $\mathbf{a}_i\in\mathbb{R}^n$, calcular el valor mínimo de la función $f(\mathbf{x})=\sum_{i=1}^{n}\|\mathbf{x}-\mathbf{a}_i\|^2$.

10.1.8. Funciones vectoriales. Matriz jacobiana

Una función vectorial es cualquier función que toma valores en un espacio vectorial de dimensión mayor mayor que 1. Las curvas en el plano o en el espacio son funciones vectoriales de una variable. Ahora nos interesa considerar funciones vectoriales de varias variables.

10.33 Definición. Sean $f_1,f_2,\dots,f_m$ campos escalares definidos en un subconjunto $E\subset\mathbb{R}^n$. La función $\mathbf{F}:E\to\mathbb{R}^m$ definida para todo $\mathbf{x}=(x_1,x_2,\dots,x_n)\in E$ por

$$\mathbf{F}(\mathbf{x})=\big(f_1(\mathbf{x}),f_2(\mathbf{x}),\dots,f_m(\mathbf{x})\big)$$

es una *función vectorial* de n variables y m *componentes*. Suele escribirse $\mathbf{F}=(f_1,f_2,\dots,f_m)$. El nombre de *campo vectorial* se aplica a aquellas funciones vectoriales que tienen igual número de variables que de componentes, esto es, para funciones definidas en un subconjunto de un espacio vectorial y que toman valores en dicho espacio vectorial.

10.34 Definición. Sea $\mathbf{F}=(f_1,f_2,\dots,f_m):E\to\mathbb{R}^m$, donde $E\subset\mathbb{R}^n$, una función vectorial de n variables y m componentes. Sea $\mathbf{a}$ un punto interior de E. Se dice que $\mathbf{F}$ es **diferenciable** en $\mathbf{a}$ si los campos escalares $f_1,f_2,\dots,f_m$ componentes de $\mathbf{F}$ son diferenciables en $\mathbf{a}$. En tal caso, la matriz cuyas filas son los vectores gradiente $\nabla f_i(\mathbf{a})$, esto es la matriz de m filas y n columnas $\big(D_j f_i(\mathbf{a})\big)_{\substack{1\leqslant i\leqslant m\\1\leqslant j\leqslant n}}$, se llama **matriz jacobiana** de f en $\mathbf{a}$ y se representará por $J(f,\mathbf{a})$.

La aplicación lineal $D\mathbf{F}(\mathbf{a}):\mathbb{R}^n\to\mathbb{R}^m$ definida para todo $\mathbf{x}\in\mathbb{R}^n$ por

$$D\mathbf{F}(\mathbf{a})(\mathbf{x})=J(f,\mathbf{a})\cdot\mathbf{x}^t$$

donde "$\cdot$" indica producto matricial y $\mathbf{x}^t$ es el vector columna $\mathbf{x}$, se llama **diferencial** de $\mathbf{F}$ en $\mathbf{a}$.

En términos del producto escalar, podemos escribir para todo $\mathbf{x}\in\mathbb{R}^n$:

$$D\mathbf{F}(\mathbf{a})(\mathbf{x}) = \big(\langle\nabla f_1(\mathbf{a})\,|\,\mathbf{x}\rangle, \langle\nabla f_2(\mathbf{a})\,|\,\mathbf{x}\rangle, \dots, \langle\nabla f_m(\mathbf{a})\,|\,\mathbf{x}\rangle\big) \in\mathbb{R}^m$$

Es fácil deducir a partir de esta igualdad y de la definición de campo escalar diferenciable que se verifica

$$\lim_{\mathbf{x}\to\mathbf{0}} \frac{\mathbf{F}(\mathbf{x}) - \mathbf{F}(\mathbf{a}) - D\mathbf{F}(\mathbf{a})(\mathbf{x}-\mathbf{a})}{\|\mathbf{x}-\mathbf{a}\|} = \mathbf{0}$$

10.35 Teorema (Regla de la cadena). *Sean $\mathbf{F}: E \to \mathbb{R}^m$, $E \subset \mathbb{R}^n$, y $\mathbf{G}: A \to \mathbb{R}^n$, $A \subset \mathbb{R}^q$, funciones vectoriales tales que $\mathbf{G}(A) \subset E$ de manera que la composición $\mathbf{H} = \mathbf{F}\circ\mathbf{G} : A \to \mathbb{R}^m$ está definida. Supongamos que $\mathbf{G}$ es diferenciable en un punto $\mathbf{a}\in A$ y que $\mathbf{F}$ es diferenciable en el punto $\mathbf{G}(\mathbf{a})\in E$. Entonces se verifica que la función compuesta $\mathbf{H}$ es diferenciable en $\mathbf{a}$, y su diferencial viene dada como la composición de las respectivas diferenciales:*

$$D\mathbf{H}(\mathbf{a}) = D\mathbf{F}(\mathbf{G}(\mathbf{a}))\circ D\mathbf{G}(\mathbf{a}) \tag{10.16}$$

Observa que la composición tiene sentido pues $D\mathbf{G}(\mathbf{a}): \mathbb{R}^q \to \mathbb{R}^n$ y $D\mathbf{F}((\mathbf{G}(\mathbf{a})): \mathbb{R}^n \to \mathbb{R}^m$, por lo que la composición es una aplicación lineal de $\mathbb{R}^q$ a $\mathbb{R}^m$, como debe ser pues $\mathbf{H}$ es una función vectorial de q variables y m componentes.

La expresión de la igualdad (10.16) por medio de matrices jacobianas es

$$J(\mathbf{H},\mathbf{a}) = J(\mathbf{F},\mathbf{G}(\mathbf{a}))\cdot J(\mathbf{G},\mathbf{a}) \tag{10.17}$$

Poniendo $\mathbf{H} = (h_1, h_2, \dots, h_m)$, $\mathbf{F} = (f_1, f_2, \dots, f_m)$, $\mathbf{G} = (g_1, g_2, \dots, g_q)$; notando las variables por $\mathbf{x} = (x_1, x_2, \dots, x_n)\in\mathbb{R}^n$, $\mathbf{y} = (y_1, y_2, \dots, y_m)\in\mathbb{R}^q$, y escribiendo $\mathbf{b} = \mathbf{G}(\mathbf{a})$, tenemos que

$$\left(\frac{\partial h_i}{\partial y_j}(\mathbf{a})\right)_{\substack{1\leqslant i\leqslant m\\ 1\leqslant j\leqslant q}} = \left(\frac{\partial f_i}{\partial x_k}(\mathbf{b})\right)_{\substack{1\leqslant i\leqslant m\\ 1\leqslant k\leqslant n}} \cdot \left(\frac{\partial g_k}{\partial y_j}(\mathbf{a})\right)_{\substack{1\leqslant k\leqslant n\\ 1\leqslant j\leqslant q}} \qquad \mathbf{b} = \mathbf{G}(\mathbf{a})$$

De donde se sigue

$$\frac{\partial h_i}{\partial y_j}(\mathbf{a}) = \sum_{k=1}^{n} \frac{\partial f_i}{\partial x_k}(\mathbf{b})\frac{\partial g_k}{\partial y_j}(\mathbf{a}) \qquad \mathbf{b} = \mathbf{G}(\mathbf{a}) \qquad (1\leqslant i\leqslant m,\ 1\leqslant j\leqslant q) \tag{10.18}$$

Esta igualdad constituye la **regla de la cadena para derivadas parciales** y es importante que aprendas a aplicarla y que entiendas lo que dice. Voy a intentar facilitarte las cosas.

Primero, lo más frecuente es que $\mathbf{F}$ sea un campo escalar. Supongamos, pues, que en lo anterior, $\mathbf{F} = f$ es un campo escalar, en cuyo caso $h = f\circ\mathbf{G}$ también es un campo escalar. La igualdad (10.18) queda ahora

$$\frac{\partial h}{\partial y_j}(\mathbf{a}) = \sum_{k=1}^{n} \frac{\partial f}{\partial x_k}(\mathbf{b})\frac{\partial g_k}{\partial y_j}(\mathbf{a}) \qquad \mathbf{b} = \mathbf{G}(\mathbf{a}) \qquad (1\leqslant j\leqslant q) \tag{10.19}$$

En esta igualdad se interpreta que la función $\mathbf{G}: A \to E \subset \mathbb{R}^n$ lo que hace es un *"cambio de variables"*. Hablando familiarmente, podemos decir, que las "variables antiguas" de la función f, esto es las $\mathbf{x} = (x_1, x_2, \dots, x_n)\in E$ se han sustituido por "variable nuevas" $\mathbf{y} = (y_1, y_2, \dots, y_q)\in A$ y la función f se ha "expresado en estas nuevas variables" dando lugar a la función h. La relación entre unas variables y otras viene dada por

$$x_k = g_k(y_1, y_2, \dots, y_q), \qquad 1 \leqslant k \leqslant n \tag{10.20}$$

De esta manera podemos interpretar la igualdad (10.19) en la forma siguiente:

> Para derivar la función nueva h, respecto a una nueva variable y_j, se deriva la función antigua f respecto a cada una de sus variables x_k y se multiplica por la derivada de cada una de ellas $x_k = g_k(y_1, y_2, \dots, y_q)$ respecto a la variable y_j.

Ya se ve que la situación está pidiendo que hagamos algunas simplificaciones que, además, son las que se hacen siempre en la práctica porque, aunque son algo confusas facilitan mucho los cálculos.

Lo primero que se hace es identificar las funciones g_k que introducen las nuevas coordenadas con las coordenadas antiguas x_k, es decir, vemos las coordenadas antiguas como funciones de las nuevas y esto lo escribimos en la forma siguiente.

$$x_k = x_k(y_1, y_2, \dots, y_q), \qquad 1 \leqslant k \leqslant n \tag{10.21}$$

Con esta notación, la igualdad (10.19) queda como sigue.

$$\frac{\partial h}{\partial y_j}(\mathbf{a}) = \sum_{k=1}^{n} \frac{\partial f}{\partial x_k}(\mathbf{b}) \frac{\partial x_k}{\partial y_j}(\mathbf{a}) \qquad \mathbf{b} = \mathbf{G}(\mathbf{a}) \qquad (1 \leqslant j \leqslant q) \tag{10.22}$$

Observa el doble papel que desempeña a la derecha de esta igualdad la letra x_k; cuando se deriva respecto de ella representa una variable y cuando ella se deriva respecto de una variable nueva representa una función.

La igualdad (10.22) ya es bastante fácil de recordar pero todavía se siguen haciendo en la práctica, sobre en todo en los textos de Física que suelen usar notaciones muy desafortunadas, algunas simplificaciones adicionales (y peligrosas). A saber: no se distingue entre la función f y la función h porque, como suele decirse en esos textos aludidos, son *"la misma función expresada en distintas variables"*. Haciendo la identificación de f con h nos queda lo siguiente.

$$\frac{\partial f}{\partial y_j}(\mathbf{a}) = \sum_{k=1}^{n} \frac{\partial f}{\partial x_k}(\mathbf{b}) \frac{\partial x_k}{\partial y_j}(\mathbf{a}) \qquad \mathbf{b} = \mathbf{G}(\mathbf{a}) \qquad (1 \leqslant j \leqslant q) \tag{10.23}$$

Aquí la letra f desempeña un doble papel: a la izquierda es la función compuesta y a la derecha es la función dada en sus variable iniciales.

Todavía suele darse un pasito más que consiste en representar la función f con una letra que suele usarse para representar variables; a saber, la letra z. Esto es frecuente también en textos de Física. Vamos a hacerlo así.

$$\frac{\partial z}{\partial y_j}(\mathbf{a}) = \sum_{k=1}^{n} \frac{\partial z}{\partial x_k}(\mathbf{b}) \frac{\partial x_k}{\partial y_j}(\mathbf{a}) \qquad \mathbf{b} = \mathbf{G}(\mathbf{a}) \qquad (1 \leqslant j \leqslant q) \tag{10.24}$$

Todavía hay algo que podemos simplificar. Habrás observado que siempre indico la relación que hay entre los puntos $\mathbf{b}$ y $\mathbf{a}$. Eso es muy importante para entender lo que se hace. Hay que saber dónde se evalúan las derivadas parciales de cada función. Pues bien, eso no se indica *jamás* en textos de Física. Nunca se indica en dónde se evalúan las derivadas parciales. Así que vamos a suprimirlo.

$$\frac{\partial z}{\partial y_j} = \sum_{k=1}^{n} \frac{\partial z}{\partial x_k}\frac{\partial x_k}{\partial y_j} \qquad (1 \leqslant j \leqslant q) \tag{10.25}$$

Debes de familiarizarte con esta igualdad y saber reconocer en ella la igualdad de partida. Y no olvides la forma en que se evalúa esta igualdad. Lo vuelvo a poner.

$$\frac{\partial z}{\partial y_j}(\mathbf{y}) = \sum_{k=1}^{n} \frac{\partial z}{\partial x_k}(\mathbf{G}(\mathbf{y}))\frac{\partial x_k}{\partial y_j}(\mathbf{y}) \qquad (1 \leqslant j \leqslant q) \tag{10.26}$$

Si tuviéramos que volver a derivar en esta igualdad respecto a una variable y_k se derivaría como de costumbre: la derivada de una suma es la suma de las derivadas y para derivar el producto se aplica la regla usual. Pero hay un detalle muy importante y es que la función $\dfrac{\partial z}{\partial x_k}(\mathbf{G}(\mathbf{y}))$ *vuelve a ser la función compuesta del campo escalar* $\dfrac{\partial z}{\partial x_k}$ *con la función* $\mathbf{G}$. *Por tanto para derivarla hay que aplicarle la misma regla que hemos aplicado para derivar* z como función compuesta y que nos ha llevado a la igualdad anterior. Es por eso que el cálculo de derivadas parciales de segundo orden en funciones compuestas suele ser bastante engorroso y es fácil equivocarse si no se sabe lo que se hace.

10.36 Ejemplo. Vamos a calcular $\dfrac{\partial z}{\partial x}$ siendo $z = u^2 + v^5 + 3uv$ donde $u = x^2 + y^2$, $v = \operatorname{sen}(xy)$.

Así es como suelen enunciarse estos ejercicios y debes entender bien el enunciado. Nos están dando una función de las variables (u, v) a la que llaman z. Esto es la letra z representa una función, a saber, $z = u^2 + v^5 + 3uv$. Nos están dando un *cambio de variables* por medio de las igualdades $u = x^2 + y^2$, $v = \operatorname{sen}(xy)$. Y nos piden calcular $\dfrac{\partial z}{\partial x}$. Esto último ya nos dice claramente que debemos ver z como función de x e y, es decir, la letra z en $\dfrac{\partial z}{\partial x}$ es la función que nos dan *después de sustituir en ella las nuevas variables*, o sea, la función compuesta de $z = u^2 + v^5 + 3uv$ con $\mathbf{G}(x, y) = (x^2 + y^2, \operatorname{sen}(xy))$.

Sabemos que

$$\frac{\partial z}{\partial x} = \frac{\partial z}{\partial u}\frac{\partial u}{\partial x} + \frac{\partial z}{\partial v}\frac{\partial v}{\partial x} = (2u + 3v)2x + (5v^4 + 3u)y\cos(xy)$$

Si lo dejamos así escrito parece que $\dfrac{\partial z}{\partial x}$ depende de 4 variables. Pero no es así porque en la igualdad anterior las variables son x e y (las nuevas variables) mientras que u y v (las antiguas variables) vienen dadas por $u = x^2 + y^2$, $v = \operatorname{sen}(xy)$. Por tanto, es mejor hacer la sustitución, con lo que resulta

$$\frac{\partial z}{\partial x} = (2(x^2 + y^2) + 3\operatorname{sen}(xy))2x + (5\operatorname{sen}^4(xy) + 3x^2 + y^2)y\cos(xy)$$

que nos da el valor de la derivada parcial de la función compuesta en un punto (x, y). En este caso es muy sencillo calcular la función compuesta. Hazlo y comprueba el resultado obtenido. ♦

Ejercicios

Consideremos una función de dos variables x e y, $z=z(x,y)$, y supongamos que expresamos x e y en función de nuevas variables u y v, lo que indicamos en la forma $x=x(u,v)$, $y=y(u,v)$. De esta forma la función z es función (función compuesta) de las "variables libres" u y v, a través de las "variables dependientes" x e y. Se trata de calcular las derivadas parciales de z respecto de las nuevas variables u y v. La regla para hacerlo es la siguiente: para derivar una función

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

respecto de una nueva variable, se deriva z respecto de cada una de las antiguas variables y se multiplica por la derivada de cada antigua variable respecto de la nueva variable. Se entiende mejor si lo escribimos simbólicamente

$$\frac{\partial z}{\partial u}=\frac{\partial z}{\partial x}\frac{\partial x}{\partial u}+\frac{\partial z}{\partial y}\frac{\partial y}{\partial u}$$

En esta igualdad debes darte cuenta de que a la izquierda, como estamos derivando respecto a u, la letra z representa a la función compuesta $z=z(x(u,v),y(u,v))$ y la derivada está calculada en un punto (u,v). En la parte derecha de la igualdad la letra z representa la función dada $z=z(x,y)$ y las letras x e y representan variables (cuando se deriva respecto de ellas) y funciones (cuando se derivan respecto de u). Debe entenderse que cuando se sustituye un valor de (u,v) en la igualdad los valores de x e y deben sustituirse por $x=x(u,v)$, $y=y(u,v)$.

1. Sea $z=\cos(xy)+\mathrm{e}^{y-1}\cos x$ donde $x=u^2+v, y=u-v^2$. Calcular $\dfrac{\partial z}{\partial u}$ en el punto $(u,v)=(1,1)$.
2. Sea $u=(x+y)^4+y^2(z+x)^3$ donde $x=rs\,\mathrm{e}^{-t},\ y=rs\log(1+t^2),\ z=r^2s\cos t$. Calcula $\dfrac{\partial u}{\partial s}$ cuando $r=2, s=1, t=0$.
3. Sea $z=f(x,y)$, y pongamos $x=u^2+v^2$, $y=u/v$. Calcular las derivadas parciales de de z respecto de las nuevas variables u y v en función de las derivadas parciales de z respecto de x e y.
4. Sea $u=x^4y+y^2z^3+\varphi(x/y)$, donde
$$\begin{cases} x=1+rs\,\mathrm{e}^{t}\\ y=rs^2\,\mathrm{e}^{-t}\\ z=r^2s\operatorname{sen}t\end{cases}$$
Calcular $\dfrac{\partial u}{\partial s}$ cuando $r=2, s=1, t=0$, sabiendo que $\varphi'(3/2)=-1$.
5. Sea $z=f(x,y)$ donde $x=s^4+r^4, y=2rs^2$. Calcula $\dfrac{\partial z}{\partial x}(2,2)$ y $\dfrac{\partial z}{\partial y}(2,2)$. Siendo $\dfrac{\partial z}{\partial r}(1,1)=-2$ y $\dfrac{\partial z}{\partial s}(1,1)=3$.
6. Prueba que la función $F(x,y)=f(\frac{y}{x^2-y^2})$, donde f es una función real derivable, verifica la igualdad
$$(x^2+y^2)\frac{\partial F}{\partial x}+2xy\frac{\partial F}{\partial y}=0$$

7. Prueba que la función $F(u,v) = f(uv,(u^2-v^2)/2)$, donde $f:\mathbb{R}^2\to\mathbb{R}$ es una función diferenciable, verifica la igualdad

$$(u^2+v^2)\left[\left(\frac{\partial f}{\partial x}\right)^2+\left(\frac{\partial f}{\partial y}\right)^2\right]=\left(\frac{\partial F}{\partial u}\right)^2+\left(\frac{\partial F}{\partial v}\right)^2$$

8. Sea $z = f(x,y)$, donde $x = \rho\cos\vartheta$, $y = \rho\operatorname{sen}\vartheta$. Calcula $\partial z/\partial\rho$ y $\partial z/\partial\vartheta$ y prueba que

$$\left(\frac{\partial z}{\partial x}\right)^2+\left(\frac{\partial z}{\partial y}\right)^2=\left(\frac{\partial z}{\partial \rho}\right)^2+\frac{1}{\rho^2}\left(\frac{\partial z}{\partial \vartheta}\right)^2$$

9. Sea $g(s,t) = f(s^2-t^2,t^2-s^2)$. Prueba la igualdad $t\dfrac{\partial g}{\partial s}+s\dfrac{\partial g}{\partial t}=0$.

10. Sea $u = f(x,y)$ donde $x = \mathrm{e}^s\cos t$, $y = \mathrm{e}^s\operatorname{sen}t$. Justifica que

$$\frac{\partial^2 u}{\partial x^2}+\frac{\partial^2 u}{\partial y^2}=\mathrm{e}^{-2s}\left(\frac{\partial^2 u}{\partial s^2}+\frac{\partial^2 u}{\partial t^2}\right)$$

11. Sea $z = f(x,y)$, donde $x = \rho\cos\vartheta$, $y = \rho\operatorname{sen}\vartheta$. Prueba que

$$\frac{\partial^2 z}{\partial x^2}+\frac{\partial^2 z}{\partial y^2}=\frac{\partial^2 z}{\partial \rho^2}+\frac{1}{\rho^2}\frac{\partial^2 z}{\partial \vartheta^2}+\frac{1}{\rho}\frac{\partial z}{\partial \rho}$$

12. Sea $z = f(x,y)$ donde $x = x(u,v)$, $y = y(u,v)$. Prueba que

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

E indica la forma e que se evalúan estas funciones.

13. Una función se llama homogénea de grado $n\in\mathbb{N}$ si $f(tx,ty) = t^n f(x,y)$. Prueba que en tal caso se verifica la igualdad

$$x\frac{\partial f}{\partial x}+y\frac{\partial f}{\partial y}=nf(x,y)$$

14. Sean las funciones $f(x,y,z) = (\mathrm{e}^x+y^2,\lambda\mathrm{e}^z+y)$, $g(u,v) = v^2+\log u$ para $(u,v)\in\mathbb{R}\times\mathbb{R}^+$. ¿Qué valor debe tener λ para que la derivada direccional máxima de $g\circ f$ en $(0,0,0)$ sea igual a 1?

10.1.9. Extremos condicionados

En la teoría de extremos relativos se supone que las variables pueden tomar valores en cualquier punto de un conjunto abierto, es decir, pueden *"moverse libremente"* en dicho conjunto. En muchos, por no decir que en la mayoría, de los problemas reales las variables no tienen tanta libertad y están obligadas a satisfacer ciertas condiciones que en Física suelen llamarse *''ligaduras"*. Por ejemplo, supongamos que un móvil se mueve en una curva Γ dada por la intersección de dos superficies; para cada punto $(x,y,z)\in\Gamma$ la energía cinética del móvil viene dada por una función conocida $f(x,y,z)$ y queremos calcular los puntos de la trayectoria donde dicha energía es máxima o mínima. En esta situación las variables x, y, z no son libres sino que deben satisfacer la condición $(x,y,z)\in\Gamma$. Otro ejemplo; supongamos que la temperatura en un punto (x,y,z) de la superficie terrestre viene dada por una función $T(x,y,z)$ y queremos calcular los puntos de mayor y menor temperatura. Aquí las variables tampoco son libres pues deben verificar una condición de la forma $x^2+y^2+z^2=R^2$ donde R es el radio de la Tierra. Igualmente, en problemas de optimización de costes o beneficios las variables están siempre sometidas a restricciones que dependen de las condiciones de producción o del mercado.

Es importante que comprendas la diferencia entre un problema de extremos relativos "libres" y un problema de extremos condicionados. Considera el siguiente ejemplo.

10.37 Ejemplo. La función $f(x,y)=xy\mathrm{e}^{x^2+y^2}$ tiene un único punto crítico, el origen, que es un punto de silla. Por tanto dicha función no tiene extremos relativos en $\mathbb{R}^2$. Supongamos que imponemos a las variables la condición $x^2+y^2=1$ y queremos calcular el máximo valor de $f(x,y)$ cuando se verifique que $x^2+y^2=1$. Fíjate en que el problema es completamente distinto. Ahora solamente nos interesan los valores que toma la función $f(x,y)$ en el conjunto

$$K=\left\{(x,y)\in\mathbb{R}^2 : x^2+y^2=1\right\}$$

Sabemos que dicho conjunto es un conjunto compacto (es cerrado – porque coincide con su frontera – y acotado); además la función f es continua, por tanto podemos asegurar, de entrada, que tiene que haber algún punto $(a,b)\in K$ en el cual la función f alcanza su mayor valor en K (y tiene que haber otro donde alcance su menor valor en K). Calcular dicho punto es, en este caso, muy sencillo pues para $(x,y)\in K$ se tiene que $f(x,y)=\mathrm{e}\,xy$. Como para $(x,y)\in K$ se tiene que $y=\pm\sqrt{1-x^2}$ y los valores negativos de f no nos interesan porque queremos calcular el mayor valor que toma en K, se sigue que

$$\text{máx}\{f(x,y):(x,y)\in K\}=\text{máx}\left\{\mathrm{e}\,x\sqrt{1-x^2} : -1\leqslant x\leqslant 1\right\}$$

Nuestro problema se ha convertido en calcular el máximo absoluto de la función $h(x)=\mathrm{e}\,x\sqrt{1-x^2}$ para $-1\leqslant x\leqslant 1$. ♦

De hecho, tú has resuelto ejercicios de extremos condicionados aunque no seas consciente de ello. Por ejemplo, seguro que alguna vez has resuelto el siguiente ejercicio.

10.38 Ejemplo. Entre todos los rectángulos cuyo perímetro es igual a 16 calcular el que tiene área máxima.

Este ejercicio puedes plantearlo como sigue. Sea $f(x,y)=xy$ la función que da el área de un rectángulo cuyos lados tienen longitudes x e y. Se trata de calcular el máximo de $f(x,y)$ cuando las variables verifican la condición $2x+2y=16$. Por tanto, es un problema de extremos condicionados. Seguro que ahora recuerdas algunos otros ejercicios parecidos a este que has hecho sin saber que estabas haciendo problemas de extremos condicionados. La razón es clara: la condición que nos dan es tan sencilla que permite despejar una variable en función de la otra, $y=8-x$, con lo que nuestra función se convierte en $xy=x(8-x)$ y el problema queda reducido a calcular el mayor valor de $x(8-x)$ cuando $-8\leqslant x\leqslant 8$. ♦

Los ejemplos anteriores ponen de manifiesto que *los problemas de extremos condicionados en los que puede utilizarse la condición que nos dan para despejar una variable en función de* *otra, se reducen fácilmente a problemas de extremos de funciones de una variable.* Pero supongamos ahora que cambiamos la condición del ejemplo 1 por la siguiente:

$$x - e^x + y + e^y + \sin(1 + xy) = 2$$

La cosa se complica porque ahora es imposible usar la condición impuesta para despejar una variable en función de la otra. Ahora sí tenemos un auténtico problema de extremos condicionados.

Lo antes dicho para funciones de dos variables puedes generalizarlo para funciones de tres variables. Por ejemplo el problema de calcular las dimensiones de un ortoedro de volumen igual a 8 para que su superficie lateral sea mínima, puedes plantearlo como sigue: calcular el máximo de

$$f(x,y,z) = 2xy + 2xz + 2yz$$

(la función que da la superficie lateral de un ortoedro cuyos lados tiene longitudes x, y, z) con la condición $xyz = 8$. Se trata de un problema de extremos condicionados, pero la condición dada permite despejar una variable en función de las otras dos $z = 8/(xy)$ con lo que nuestra función queda $2xy + 2xz + 2yz = xy + 16/y + 16/x$ función de la que hay que calcular su mínimo absoluto cuando $0 < x$, $0 < y$. Hemos convertido así el problema en uno de extremos relativos de una función de dos variables. Pero si cambiamos la condición anterior por la siguiente

$$x^2yz^3 + \operatorname{sen}(1 + xz) + y - e^{yx} = 1$$

o bien, si imponemos dos condiciones como las siguientes:

$$\log(1 + x^2 + y^2) + \sin(1 + xz) - 1 = 0, \qquad e^{1+y+x+z} + \cos(xyz) + x^2z^2 - 3 = 0$$

entonces no podemos usar esa condición (o condiciones) para despejar una variable (o dos variables) en función de las otras (de la otra).

La teoría de extremos condicionados te dice cómo proceder en este tipo de problemas independientemente de que la condición (o condiciones) que nos den sea más o menos fácil y permita o no despejar variables. El resultado básico de esa teoría, que proporciona una ***condición necesaria*** de extremo condicionado, es el teorema de Lagrange. Para facilitar su comprensión, en vez de dar un enunciado general, lo enuncio en los tres casos que se presentan con mayor frecuencia. Antes de enunciarlo conviene dar la definición de extremo local condicionado.

10.39 Definición. Sea f un campo escalar de n variables y S un subconjunto de $\mathbb{R}^n$. Se dice que f tiene un máximo (resp. mínimo) local condicionado (por la condición $x \in S$) en un punto $a \in S$, si hay un número $r > 0$ tal que para todo $x \in B(x,r) \cap S$ se verifica que $f(a) \geqslant f(x)$ (resp. $f(a) \leqslant g(x)$). Cuando f tiene en a un máximo o un mínimo local condicionado (por la condición $x \in S$) se dice que f tiene un extremo condicionado en a.

Teorema de Lagrange

En lo que sigue supondremos que las funciones que intervienen tienen derivadas parciales de primer orden continuas.

a) Consideremos el problema de calcular los extremos locales una función de dos variables $f(x,y)$ cuando las variables están obligadas a moverse en una curva Γ dada por $g(x,y)=0$:

$$\Gamma=\left\{(x,y)\in\mathbb{R}^2 : g(x,y)=0\right\}$$

Es decir, se trata de un problema de extremos condicionados por la condición $(x,y)\in\Gamma$ o, equivalentemente, $g(x,y)=0$.

Además de las condiciones de derivabilidad que se han supuesto al principio, hay que suponer que el vector gradiente de g no se anula en los puntos de Γ. En estas hipótesis, para que un punto $(a,b)\in\Gamma$ sea un extremo local condicionado de f, es necesario que los vectores gradiente de f y de g en el punto (a,b) sean linealmente dependientes; es decir, que exista un número real λ_0 tal que

$$\nabla f(a,b)+\lambda_0\nabla g(a,b)=0 \Longleftrightarrow \begin{cases} \dfrac{\partial f}{\partial x}(a,b)+\lambda_0\dfrac{\partial g}{\partial x}(a,b)=0 \\[2mm] \dfrac{\partial f}{\partial y}(a,b)+\lambda_0\dfrac{\partial g}{\partial y}(a,b)=0 \end{cases}$$

Como debe cumplirse también que $g(a,b)=0$, para recordar estas tres condiciones que debe cumplir el punto (a,b) se suele definir una nueva función de tres variables, llamada función de Lagrange, por

$$F(x,y,\lambda)=f(x,y)+\lambda g(x,y)$$

y las condiciones anteriores nos dicen que el punto (a,b,λ_0) es un punto crítico de la función de Lagrange, es decir, es solución del sistema de ecuaciones (llamado sistema de Lagrange):

$$\begin{cases} \dfrac{\partial F}{\partial x}(x,y,\lambda)=\dfrac{\partial f}{\partial x}(x,y)+\lambda\dfrac{\partial g}{\partial x}(x,y)=0 \\[2mm] \dfrac{\partial F}{\partial y}(x,y,\lambda)=\dfrac{\partial f}{\partial y}(x,y)+\lambda\dfrac{\partial g}{\partial y}(x,y)=0 \\[2mm] \dfrac{\partial F}{\partial \lambda}(x,y,\lambda)=g(x,y)=0 \end{cases}$$

b) Consideremos el problema de calcular los extremos locales una función de tres variables $f(x,y,z)$ cuando las variables están obligadas a moverse en una superficie S dada por $g(x,y,z)=0$:

$$S=\left\{(x,y,z)\in\mathbb{R}^3 : g(x,y,z)=0\right\}$$

Es decir, se trata de un problema de extremos condicionados por la condición $(x,y,z)\in S$ o, equivalentemente, $g(x,y,z)=0$.

Además de las condiciones de derivabilidad que se han supuesto al principio, hay que suponer que el vector gradiente de g no se anula en los puntos de S. En estas hipótesis, para que un punto $(a,b,c)\in S$ sea un extremo local condicionado de f, es necesario que los vectores gradiente de f y de g en el punto (a,b,c) sean linealmente dependientes; es decir, que exista un número real λ_0 tal que

$$\nabla f(a,b,c)+\lambda_0\nabla g(a,b,c)=0 \Longleftrightarrow \begin{cases} \dfrac{\partial f}{\partial x}(a,b,c)+\lambda_0\dfrac{\partial g}{\partial x}(a,b,c)=0 \\[2mm] \dfrac{\partial f}{\partial y}(a,b,c)+\lambda_0\dfrac{\partial g}{\partial y}(a,b,c)=0 \\[2mm] \dfrac{\partial f}{\partial z}(a,b,c)+\lambda_0\dfrac{\partial g}{\partial z}(a,b,c)=0 \end{cases}$$

Como debe cumplirse también que $g(a,b,c)=0$, para recordar estas cuatro condiciones que debe cumplir el punto (a,b,c) se suele definir una nueva función de cuatro variables, llamada función de Lagrange, por

$$F(x,y,z,\lambda)=f(x,y,z)+\lambda g(x,y,z)$$

y las condiciones anteriores nos dicen que el punto (a,b,c,λ_0) es un punto crítico de la función de Lagrange, es decir, es solución del sistema de ecuaciones (llamado sistema de Lagrange):

$$\left\{\begin{array}{l}\dfrac{\partial F}{\partial x}(x,y,z,\lambda)=\dfrac{\partial f}{\partial x}(x,y,z)+\lambda\dfrac{\partial g}{\partial x}(x,y,z)=0\\[2mm] \dfrac{\partial F}{\partial y}(x,y,z,\lambda)=\dfrac{\partial f}{\partial y}(x,y,z)+\lambda\dfrac{\partial g}{\partial y}(x,y,z)=0\\[2mm] \dfrac{\partial F}{\partial z}(x,y,z,\lambda)=\dfrac{\partial f}{\partial z}(x,y,z)+\lambda\dfrac{\partial g}{\partial z}(x,y,z)=0\\[2mm] \dfrac{\partial F}{\partial \lambda}(x,y,z,\lambda)=g(x,y,z)=0\end{array}\right.$$

c) Consideremos el problema de calcular los extremos locales una función de tres variables $f(x,y,z)$ cuando las variables están obligadas a moverse en una curva Γ dada por $g(x,y,z)=h(x,y,z)=0$:

$$\Gamma=\left\{(x,y,z)\in\mathbb{R}^3 : g(x,y,z)=h(x,y,z)=0\right\}$$

Es decir, se trata de un problema de extremos condicionados por la condición $(x,y,z)\in\Gamma$ o, equivalentemente, $g(x,y,z)=h(x,y,z)=0$.

Además de las condiciones de derivabilidad que se han supuesto al principio, hay que suponer que los vectores gradiente de g y de h son linealmente independientes en todo punto de Γ. En estas hipótesis, para que un punto $(a,b,c)\in\Gamma$ sea un extremo local condicionado de f, es necesario que los vectores gradiente de f, g y h en el punto (a,b,c) sean linealmente dependientes; es decir, que existan números reales λ_0,μ_0 tales que

$$\nabla f(a,b,c)+\lambda_0\nabla g(a,b,c)+\mu_0\nabla h(a,b,c)=0\Longleftrightarrow\left\{\begin{array}{l}\dfrac{\partial f}{\partial x}(a,b,c)+\lambda_0\dfrac{\partial g}{\partial x}(a,b,c)+\mu_0\dfrac{\partial h}{\partial x}(a,b,c)=0\\[2mm] \dfrac{\partial f}{\partial y}(a,b,c)+\lambda_0\dfrac{\partial g}{\partial y}(a,b,c)+\mu_0\dfrac{\partial h}{\partial y}(a,b,c)=0\\[2mm] \dfrac{\partial f}{\partial z}(a,b,c)+\lambda_0\dfrac{\partial g}{\partial z}(a,b,c)+\mu_0\dfrac{\partial h}{\partial z}(a,b,c)=0\end{array}\right.$$

Como debe cumplirse también que $g(a,b,c)=h(a,b,c)=0$, para recordar estas cinco condiciones que debe cumplir el punto (a,b,c) se suele definir una nueva función de cinco variables, llamada función de Lagrange, por

$$F(x,y,z,\lambda,\mu)=f(x,y,z)+\lambda g(x,y,z)+\mu\, h(x,y,z)$$

y las condiciones anteriores nos dicen que el punto (a,b,c,λ_0,μ_0) es un punto crítico de la función de Lagrange, es decir, es solución del sistema de ecuaciones (llamado sistema de Lagrange):

$$\begin{cases} \dfrac{\partial F}{\partial x}(x,y,z,\lambda,\mu) = \dfrac{\partial f}{\partial x}(x,y,z) + \lambda\dfrac{\partial g}{\partial x}(x,y,z) + \mu\dfrac{\partial h}{\partial x}(x,y,z) = 0 \\ \dfrac{\partial F}{\partial y}(x,y,z,\lambda,\mu) = \dfrac{\partial f}{\partial y}(x,y,z) + \lambda\dfrac{\partial g}{\partial y}(x,y,z) + \mu\dfrac{\partial h}{\partial y}(x,y,z) = 0 \\ \dfrac{\partial F}{\partial z}(x,y,z,\lambda,\mu) = \dfrac{\partial f}{\partial z}(x,y,z) + \lambda\dfrac{\partial g}{\partial z}(x,y,z) + \mu\dfrac{\partial h}{\partial z}(x,y,z) = 0 \\ \dfrac{\partial F}{\partial \lambda}(x,y,z,\lambda,\mu) = g(x,y,z) = 0 \\ \dfrac{\partial F}{\partial \mu}(x,y,z,\lambda,\mu) = h(x,y,z) = 0 \end{cases}$$

Esta es la teoría que debes saber referente a extremos condicionados. El método que hemos descrito se conoce como ***método de los multiplicadores de Lagrange*** porque las variables λ, μ que se introducen se llaman multiplicadores de Lagrange.

La situación que consideraremos en los ejercicios será la siguiente: deberás calcular el máximo o el mínimo absolutos de los valores de una función cuando las variables están sometidas a una condición como las que hemos considerado anteriormente (las variables deben estar en una curva Γ en el plano, o en una superficie S en el espacio, o en una curva Γ dada como intersección de dos superficies) donde, **además la curva Γ o la superficie S, según sea el caso, son conjuntos compactos** (lo que deberás justificar en cada caso). En esta situación, el teorema de Weierstrass asegura que hay puntos de Γ o S en los que la función alcanza un máximo y un mínimo absolutos, es decir, son puntos en los que la función toma el mayor valor o el menor valor de todos los valores que toma en Γ o S. Para calcular dichos puntos lo único que debes hacer es calcular los puntos críticos de la función de Lagrange y calcular el valor de la función en cada uno de ellos, aquél punto (o puntos, puede haber más de uno) donde la función tome el mayor valor será el punto donde se alcanza el máximo absoluto; aquél punto (o puntos, puede haber más de uno) donde la función tome el menor valor será donde se alcanza el mínimo absoluto.

Finalmente, incluyo, por complitud, un resultado que establece condiciones suficientes de extremo condicionado. No creo que tengas que usarlo.

Condiciones suficientes de extremo condicionado

Supongamos que f es un campo escalar de n variables con derivadas parciales continuas de segundo orden. Sean g_j, $1 \leqslant j \leqslant m$, campos escalares de n variables con derivadas parciales de segundo orden continuas y definamos $M = \{\mathbf{x} : g_j(\mathbf{x}) = 0, 1 \leqslant j \leqslant m\}$. Se supone que en todo punto $\mathbf{x} \in M$ los vectores gradiente $\nabla g_j(\mathbf{x})$ son linealmente independientes. Pongamos $\mathbf{G} = (g_1, g_2, \dots, g_m)$ y $\lambda = (\lambda_1, \lambda_2, \dots, \lambda_m)$. Sea

$$F(\mathbf{x},\lambda) = f(\mathbf{x}) + \langle \mathbf{G}(\mathbf{x}) \,|\, \lambda \rangle$$

la función de Lagrange y sea $(\mathbf{a},\mu)$ un punto crítico de la misma. Consideremos el siguiente polinomio

$$p(z)=\begin{vmatrix} \mathbf{0}_{m\times m} & J(\mathbf{G},\mathbf{a}) \\ J(\mathbf{G},\mathbf{a})^t & \left(\dfrac{\partial^2 F}{\partial x_i \partial x_j}(\mathbf{a},\mu)\right)_{\substack{1\leqslant i\leqslant n\\ 1\leqslant j\leqslant n}} -z\mathbf{I} \end{vmatrix}$$

- Si $p(z)$ es de grado $n-m$ y todos sus coeficientes son positivos o negativos, entonces **a** es un máximo local condicionado de f.
- Si $p(z)$ es de grado $n-m$ y todos sus coeficientes son distintos de cero y van alternando su signo, entonces **a** es un mínimo local condicionado de f.
- Si $p(z)$ es de grado $n-m$ sus coeficientes nulos están seguidos y llegan hasta el término independiente y los no nulos o bien tienen todos igual signo o van alternando su signo, no se puede decir nada.
- En otro caso **a** no es extremo condicionado de f.

Ejercicios

1. Calcular el valor mayor y el valor menor que toma la función $f(x,y,z)=xyz$ en los puntos del elipsoide $x^2+4y^2+9z^2=3$.
2. Calcular el valor mayor y el valor menor que toma la función $f(x,y,z)=y^2+4z^2-4yz-2xz-2xy$ en los puntos del elipsoide $2x^2+3y^2+6z^2=1$.
3. Determinar los puntos sobre la curva $x^2y=2$ más próximos al origen.
4. Hallar el punto de la recta intersección de los planos $x-y=2$ y $x-2z=4$, que está más próximo al origen.
5. Calcular el punto $P(x,y,z)$ en el plano de ecuación $2x+y-z=5$ que está más cerca del origen.
6. El plano $x+y+z=24$ corta al paraboloide $z=x^2+y^2$ en una elipse. Calcula los puntos más altos y más bajos de dicha elipse.
7. Utiliza el método de los multiplicadores de Lagrange para calcular un punto de la elipse de ecuación
$$\frac{x^2}{a^2}+\frac{y^2}{b^2}=1$$
tal que el segmento determinado por la intersección de la tangente a la elipse en dicho punto con los ejes coordenados tenga longitud mínima.
8. Dado el elipsoide
$$\frac{x^2}{a^2}+\frac{y^2}{b^2}+\frac{z^2}{c^2}=1$$
calcular un punto de coordenadas positivas tal que el plano tangente al elipsoide en dicho punto determine con los ejes coordenados un tetraedro de volumen mínimo.

9. Hallar los puntos de la curva
$$\begin{cases} x^2 - xy + y^2 - z^2 = 1 \\ x^2 + y^2 = 1 \end{cases}$$
que están más próximos al origen de coordenadas.

10. Calcular la mínima distancia del origen a la superficie de ecuación $xy^2z^3 = 2$.

11. Calcular los valores máximo y mínimo de la función $f(x,y,z) = xyz$ cuando el punto (x,y,z) pertenece a la curva definida por la intersección del plano $x+y+z=0$ y la esfera $x^2+y^2+z^2-1=0$.

12. Calcular la mínima distancia entre la recta $x+y=4$ y la circunferencia $x^2+y^2=1$.

13. Calcular la mínima distancia entre la recta $x-y=2$ y la parábola $y=x^2$.

14. Calcula la distancia mínima entre la elipse $x^2+2y^2=6$ y la recta $x+y=5$.

15. El área de una caja rectangular sin tapa es de 108cm^2. Calcular sus dimensiones para que el volumen sea máximo.

Cálculo de extremos en conjuntos compactos

En este tipo de ejercicios se trata de calcular el máximo o el mínimo absolutos de una función f con derivadas parciales continuas en un conjunto compacto K formado por la unión de un conjunto abierto acotado y de su frontera, $K = U \cup Fr(U)$. En este tipo de ejercicios la existencia de dichos extremos está asegurada de antemano en virtud del teorema de Weierstrass. Se trata realmente de dos problemas, pues lo que hay que hacer es estudiar los extremos relativos de f en el abierto U (un problema de extremos relativos) y estudiar los extremos locales condicionados de f en $Fr(U)$. Si la frontera de U está definida de forma apropiada (es una curva o una superficie) éste último es un problema de extremos condicionados. Cuando la frontera de U está dada por condiciones sencillas que permiten despejar variables puede hacerse un estudio directo sin necesidad de recurrir a la teoría de extremos condicionados.

1. Calcular los extremos absolutos de $f(x,y) = (x^2+2y^2)\,\mathrm{e}^{-x^2-y^2}$ en el disco $x^2+y^2 \leqslant 4$.

2. Calcular los valores máximos y mínimos absolutos de $f(x,y,z) = xy^2z^3$ en la bola $x^2+y^2+z^2 \leqslant 1$.

3. Hallar los extremos absolutos de $f(x,y) = x^2+3y^2$ en el círculo $x^2-2x+y^2-3 \leqslant 0$.

4. Hallar los extremos absolutos de la función $f(x,y) = x^2y^3(1-x-y)$ en el conjunto
$$K = \{(x,y) : |x|+|y| \leqslant 1\}$$

5. (*) Hallar los extremos absolutos de $f(x,y) = x^2+y^2-xy-x-y$ en el conjunto
$$K = \left\{(x,y) \in \mathbb{R}^2 : x \geqslant 0, y \geqslant 0, x+y \leqslant 3\right\}$$

6. Calcula los extremos absolutos del campo escalar $f(x,y,z) = x+y+z$ en el conjunto
$$A = \left\{(x,y,z) \in \mathbb{R}^3 : x^2+y^2 \leqslant z \leqslant 1\right\}.$$

10.1.10. Derivación de funciones implícitamente definidas

Sea $f(x,y)$ una función de dos variables con derivadas parciales de primer orden continuas y consideremos la ecuación $f(x,y)=0$. Las soluciones de dicha ecuación representan una curva en el plano. Bueno, hablando con propiedad pueden representar algo más general que una curva. Para que te convenzas de ello basta que consideres la ecuación

$$f(x,y)=(x^2+y^2-1)(2(x-1)^2+3(y-2)^2-1)(y-x^2)=0$$

la función f se anula en los puntos de la circunferencia $x^2+y^2=1$, de la parábola $y=x^2$ y de la elipse $2(x-1)^2+3(y-2)^2=1$. Por tanto la ecuación $f(x,y)=0$ representa la unión de todas esas curvas.

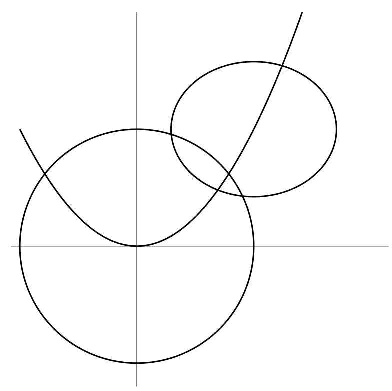


Figura 10.1: Conjunto dado por $f(x,y)=0$

Ese conjunto (ver figura (10.1)) no es exactamente una curva pero ***localmente*** se parece a una curva. La palabra "localmente" quiere decir que si fijamos un punto (a,b) tal que $f(a,b)=0$ entonces hay una bola abierta centrada en (a,b) de radio positivo, $B((a,b),r)$ tal que el corte de dicha bola con el conjunto de puntos $V=\{(x,y):f(x,y)=0\}$ es una curva, donde la palabra "curva" tiene el significado que le hemos dado en el apartado dedicado al cálculo de rectas tangentes. De hecho, no es cierto que la condición anterior se verifique para todos los puntos (a,b) tales que $f(a,b)=0$. Dicha condición falla en los puntos donde se cortan dos de las curvas cuya unión forma V, pues es claro que en dichos puntos el conjunto V no parece localmente una curva. Pues bien, dichos puntos son justamente los puntos donde se anula el vector gradiente de f. En dichos puntos la recta tangente no está definida. Este ejemplo te ayudará a entender lo que sigue.

Volvamos al caso general de una función de dos variables $f(x,y)$ con derivadas parciales continuas de primer orden. Consideremos ahora la ecuación $f(x,y)=0$ desde otro punto de vista. Intuitivamente, *una* ecuación es *una* condición que debe ligar a *una* de las variables, es decir, que si en la igualdad $f(x,y)=0$ se fija un valor de x entonces el valor de y queda determinado de manera única por dicho valor de x. A veces esto es verdad como en el siguiente ejemplo. Consideremos

$$f(x,y)=y^3+y\,\mathrm{e}^x+\operatorname{sen}x$$

Fijado un valor de x la ecuación $f(x,y)=0$ es un polinomio de tercer grado en y que tiene una única solución real pues su derivada respecto de y es $3y^2+\mathrm{e}^x$ que no se anula. Es decir, en este caso es cierto que la igualdad

$$y^3 + y\,e^x + \operatorname{sen} x = 0 \tag{10.27}$$

define de manera única a y como función de x, en el sentido de que fijado un valor de x, hay un único $y = \varphi(x)$ que verifica dicha igualdad, esto es, la función $\varphi(x)$ está definida por la condición:

$$\varphi(x)^3 + \varphi(x)\,e^x + \operatorname{sen} x = 0 \tag{10.28}$$

Se dice que la función φ **está implícitamente definida** por la igualdad (10.27). Puedes calcular con *Mathematica* el valor de dicha función y comprobarás que es bastante complicada. El hecho es que la mejor forma de trabajar con la función φ es la igualdad (10.28) que la define. Por ejemplo, si queremos calcular la derivada de φ en un punto basta con que derivemos dicha igualdad para obtener

$$3\varphi'(x)\varphi(x)^2 + \varphi'(x)\,e^x + \varphi(x)\,e^x + \cos x = 0$$

lo que permite calcular $\varphi'(x)$ en función de $\varphi(x)$.

En general, no es cierto que una igualdad de la forma $f(x,y) = 0$ permita despejar una variable en función de la otra. Para convencerte, considera el primer ejemplo que pusimos. Ni tan siquiera una igualdad tan sencilla como $x^2 + y^2 - 1 = 0$ permite despejar una variable como función de la otra pues es claro que para cada valor que fijemos de una variable (comprendido entre -1 y 1) hay *dos* posibles valores de la otra que verifican dicha igualdad.

Relacionemos ahora los dos puntos de vista que hemos considerado. Pongamos

$$\Gamma = \left\{(x,y) \in \mathbb{R}^2 : f(x,y) = 0\right\}$$

Si la igualdad $f(x,y) = 0$ permitiera despejar y en función de x, es decir, definiera una función $y = \varphi(x)$ por la condición

$$f(x,y) = 0 \Longleftrightarrow y = \varphi(x)$$

entonces se tendría que (llamando I al intervalo donde está definida φ)

$$\Gamma = \left\{(x,y) \in \mathbb{R}^2 : f(x,y) = 0\right\} = \{(x, \varphi(x)) : x \in I\}$$

es decir, el conjunto Γ sería la gráfica de φ, que, como sabemos, es un tipo muy particular de curva. Pero ya hemos visto que el conjunto Γ puede ser una "curva" mucho más general que la gráfica de una función. Pero incluso en este caso, dicha "curva" es *localmente*, excepto en los puntos donde se anula el gradiente, una gráfica de una función.

Las consideraciones anteriores se pueden llevar al caso de una función de tres variables $f(x,y,z)$ considerando ahora la "superficie" definida por la ecuación $f(x,y,z) = 0$. La pregunta ahora es si fijados un valor de x y otro de y queda determinado de manera única un valor de $z = \varphi(x,y)$ que verifica dicha ecuación. En caso afirmativo tendríamos que la superficie de ecuación $f(x,y,z) = 0$ coincidiría con la gráfica de φ. Ya puedes suponer que esto no es cierto en general pues la mayoría de las "superficies" no son gráficas de funciones.

El siguiente resultado, conocido como teorema de la función implícita, nos dice lo que podemos afirmar en general en una situación como la que estamos considerando.

Teorema de la función implícita

Suponemos que las funciones que consideramos en lo que sigue tienen derivadas parciales de primer orden continuas.

a) Consideremos primero el caso de una función $f(x,y)$ de dos variables. Sea

$$\Gamma = \left\{(x,y)\in\mathbb{R}^2 : f(x,y)=0\right\}$$

Supongamos que $(a,b)\in\Gamma$ y se verifica que

$$\frac{\partial f}{\partial y}(a,b)\neq 0$$

Entonces existe una función $\varphi : I\to\mathbb{R}$, definida en un intervalo I tal que $a\in I$ y $\varphi(a)=b$, que verifica que $f(x,\varphi(x))=0$ para todo $x\in I$. La función φ se dice que está implícitamente definida por la ecuación $f(x,y)=0$. Dicha función es derivable en I y su derivada se calcula derivando la igualdad $f(x,\varphi(x))=0$ respecto a x con lo que se obtiene

$$\frac{\partial f}{\partial x}(x,\varphi(x))+\frac{\partial f}{\partial y}(x,\varphi(x))\varphi'(x)=0\Longrightarrow \varphi'(x)=\frac{-\dfrac{\partial f}{\partial x}(x,\varphi(x))}{\dfrac{\partial f}{\partial y}(x,\varphi(x))}$$

Además tenemos que

$$\Gamma\cap(I\times\varphi(I))=\left\{(x,y)\in\mathbb{R}^2 : f(x,y)=0\right\}\cap(I\times\varphi(I))=\{(x,\varphi(x)) : x\in I\}$$

es decir, Γ es ***localmente*** en el punto (a,b) una curva que viene dada por la gráfica de φ.

b) Consideremos ahora el caso de una función $f(x,y,z)$ de tres variables. Sea

$$S=\left\{(x,y,z)\in\mathbb{R}^3 : f(x,y,z)=0\right\}$$

Supongamos que $(a,b,c)\in S$ y se verifica que

$$\frac{\partial f}{\partial z}(a,b,c)\neq 0$$

Entonces existe una función $\varphi : U\to\mathbb{R}$, definida en un abierto $U\subset\mathbb{R}^2$ con $(a,b)\in U$ y $\varphi(a,b)=c$, que verifica que $f(x,y,\varphi(x,y))=0$ para todo $(x,y)\in U$. La función φ se dice que está implícitamente definida por la ecuación $f(x,y,z)=0$. Dicha función tiene derivadas parciales continuas en U y sus derivadas parciales se calculan derivando la igualdad $f(x,y,\varphi(x,y))=0$ parcialmente respecto a x e y con lo que se obtiene

$$\frac{\partial f}{\partial x}(x,y,\varphi(x,y))+\frac{\partial f}{\partial z}(x,y,\varphi(x,y))\frac{\partial\varphi}{\partial x}(x,y)=0\Longrightarrow\frac{\partial\varphi}{\partial x}(x,y)=\frac{-\dfrac{\partial f}{\partial x}(x,y,\varphi(x,y))}{\dfrac{\partial f}{\partial z}(x,y,\varphi(x,y))}$$

$$\frac{\partial f}{\partial y}(x,y,\varphi(x,y))+\frac{\partial f}{\partial z}(x,y,\varphi(x,y))\frac{\partial\varphi}{\partial y}(x,y)=0\Longrightarrow\frac{\partial\varphi}{\partial y}(x,y)=\frac{-\dfrac{\partial f}{\partial y}(x,y,\varphi(x,y))}{\dfrac{\partial f}{\partial z}(x,y,\varphi(x,y))}$$

Además tenemos que

$$S \cap (U \times \varphi(U)) = \left\{(x,y,z) \in \mathbb{R}^3 : f(x,y,z) = 0\right\} \cap (U \times \varphi(U)) = \{(x,y,\varphi(x,y)) : (x,y) \in U\}$$

es decir, S es ***localmente*** en el punto (a,b,c) una superficie que viene dada por la gráfica de φ.

El teorema de la función implícita es mucho más general pero nos limitaremos a los casos considerados. En las hipótesis hechas pueden admitirse variaciones. La hipótesis que hay que hacer siempre es que el vector gradiente de f no sea cero en el punto considerado. En el caso **a)** puede suponerse igualmente que

$$\frac{\partial f}{\partial x}(a,b) \neq 0$$

y la conclusión es que x puede expresarse localmente como función de y, es decir, que hay una función $\psi : J \to \mathbb{R}$ definida en un intervalo J tal que $b \in J$ y $\psi(b) = a$ que verifica que $f(\psi(y), y) = 0$ para todo $y \in J$. Lo que sigue ya lo puedes suponer.

Análogamente, en el caso **b)** puede suponerse, por ejemplo que

$$\frac{\partial f}{\partial x}(a,b,c) \neq 0$$

entonces es la variable x la que queda definida localmente de forma implícita como función de y, z. Tú mismo puedes completar el enunciado en este caso. Todo esto nos da más libertad para elegir la variable que queremos expresar como función de las otras, basta con que la derivada parcial respecto de dicha variable sea distinta de cero.

En la práctica el teorema de la función implícita se aplica en la forma que te explico en los siguientes ejemplos.

10.40 Ejemplo. Comprobar que la ecuación

$$xyz + \operatorname{sen}(z-6) - 2(x+y+x^2y^2) = 0$$

define a z como función implícita de (x,y) en un entorno de $(1,1)$, con $z(1,1) = 6$. Comprobar que $(1,1)$ es un punto crítico de la función $z = z(x,y)$.

Solución. Pongamos $f(x,y,z) = xyz + \operatorname{sen}(z-6) - 2(x+y+x^2y^2)$ que tiene derivadas parciales continuas de todo orden. Tenemos que $\dfrac{\partial f}{\partial z} = xy + \cos(z-6)$, por lo que $\dfrac{\partial f}{\partial z}(1,1,6) = 2 \neq 0$. Como, además, $f(1,1,6) = 0$, el teorema de la función implícita garantiza que hay una función con derivadas parciales continuas, $(x,y) \mapsto z(x,y)$, definida en un entorno, U, de $(1,1)$ tal que $z(1,1) = 6$, y

$$f(x,y,z(x,y)) = 0 \ \text{ para todo } \ (x,y) \in U.$$

Derivando esta identidad tenemos que:

$$\frac{\partial f}{\partial x} + \frac{\partial f}{\partial z}\frac{\partial z}{\partial x} = yz - 2(1+2xy^2) + (xy + \cos(z-6))\frac{\partial z}{\partial x} = 0 \qquad (1)$$

$$\frac{\partial f}{\partial y} + \frac{\partial f}{\partial z}\frac{\partial z}{\partial y} = xz - 2(1+2x^2y) + (xy + \cos(z-6))\frac{\partial z}{\partial y} = 0 \qquad (2)$$

Donde las derivadas parciales de la función implícita $z = z(x,y)$ están calculadas en un punto $(x,y) \in U$ y las de f están calculadas en el punto $(x,y,z(x,y))$. Haciendo $x = y = 1$, $z = z(1,1) = 6$, en las igualdades anteriores, se obtiene que $\dfrac{\partial z}{\partial x}(1,1) = \dfrac{\partial z}{\partial y}(1,1) = 0$, esto es, $(1,1)$ es un punto crítico de $z = z(x,y)$. ♦

El ejemplo anterior es todavía demasiado explícito, nos dice muy claramente lo que hay que hacer. Lo más frecuente es que nos encontremos con ejercicios como el siguiente.

10.41 Ejemplo. Sabiendo que

$$y\cos(xz) + x^3\,\mathrm{e}^{zy} - z + 1 = 0 \tag{10.29}$$

Calcular $\dfrac{\partial z}{\partial x}(x,y)$ y particularizar para el punto $(x,y) = (0,0)$.

Solución. En un ejercicio como este lo más fácil es que en la igualdad (10.29) sustituyas mentalmente $z = z(x,y)$ y la veas como

$$y\cos\big(xz(x,y)\big) + x^3\,\mathrm{e}^{z(x,y)y} - z(x,y) + 1 = 0 \tag{10.30}$$

es decir, supones que has calculado para valores de x e y dados la solución respecto a z de la igualdad (10.29). Esta solución (que de hecho no es posible expresar de forma explícita, esto es, que no puede calcularse) la representamos por $z = z(x,y)$ y es la función implícita definida por la igualdad (10.29) (el teorema de la función implícita *que es un teorema de existencia* garantiza que dicha función existe). Ahora derivamos en la igualdad (10.30) respecto a x para obtener

$$-y\,\mathrm{sen}\big(xz(x,y)\big)\left(z(x,y) + x\frac{\partial z}{\partial x}(x,y)\right) + 3x^2\,\mathrm{e}^{z(x,y)y} + x^3 y\frac{\partial z}{\partial x}(x,y)\,\mathrm{e}^{z(x,y)y} - \frac{\partial z}{\partial x}(x,y) = 0$$

de donde

$$\frac{\partial z}{\partial x}(x,y) = \frac{yz(x,y)\,\mathrm{sen}\big(xz(x,y)\big) - 3x^2\,\mathrm{e}^{z(x,y)y}}{x^3y\,\mathrm{e}^{z(x,y)y} - xy\,\mathrm{sen}(xz(x,y)) - 1}$$

Naturalmente, esta igualdad tiene sentido siempre que el denominador de la fracción sea distinto de cero. Puedes comprobar que si llamas $f(x,y,z) = y\cos(xz) + x^3\,\mathrm{e}^{zy} - z + 1$ entonces la igualdad anterior es precisamente

$$\frac{-\dfrac{\partial f}{\partial x}(x,y,z)}{\dfrac{\partial f}{\partial z}(x,y,z)}$$

calculada en el punto $(x,y,z(x,y))$. Para $(x,y) = (0,0)$ se tiene que $z(0,0)$ viene dado por la ecuación que se obtiene haciendo $x = 0$ e $y = 0$ en la igualdad (10.29) de donde se sigue $z(0,0) = 1$. Además

$$\frac{\partial f}{\partial z}(0,0,z(0,0)) = \frac{\partial f}{\partial z}(0,0,1) = -1 \neq 0$$

Por lo que

$$\frac{\partial z}{\partial x}(0,0) = \frac{0}{-1} = 0$$

♦

Ejercicios

1. Calcular las derivadas parciales de primer orden de la función $z = z(x,y)$ definida implícitamente por $yz^4 + x^2z^3 - \mathrm{e}^{xyz} = 0$. Particularizar para el punto $(x,y) = (1,0)$.

2. Calcular las derivadas parciales de primer orden de la función $z = z(x,y)$ definida implícitamente por $z^3 + z\mathrm{e}^x + \cos y = 0$.

3. Calcular las derivadas parciales de primer orden de la función $z = z(x,y)$ dada implícitamente por $3x^2y^2 + 2z^2xy - 2zx^3 + 4zy^3 - 4 = 0$, en el punto $(2,1)$ siendo $z(2,1) = 2$.

4. Supongamos que la igualdad

$$\int_{xy}^{y+z} g(t)dt + \int_{3x+y}^{z^2} h(t)dt = 0$$

donde g y h son funciones reales derivables, define a z como función implícita de x,y. Calcular las derivadas parciales de primer orden de $z = z(x,y)$.

5. Supongamos que la igualdad $F(x,y,z) = 0$ determina implícitamente funciones diferenciables $x = x(y,z)$, $y = y(x,z)$, $z = z(x,y)$. Probar que $\dfrac{\partial x}{\partial y}\dfrac{\partial y}{\partial z}\dfrac{\partial z}{\partial x} = -1$.

6. Calcular la derivada de la función $y = y(x)$ definida implícitamente por

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

Particularizar para $x = 1$ sabiendo que $y(1) = 1$.

7. Calcular la derivada de la función $y = y(x)$ definida implícitamente por

$$y\log(x^2 2 + y^2) - 2xy = 0$$

Particularizar para $x = 0$ sabiendo que $y(0) = 1$.